(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
03 June 2021 (03.06.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/108765 A1

(51) International Patent Classification:
C12N 9/02 (2006.01) *G01N 21/76* (2006.01)

(21) International Application Number:
PCT/US2020/062499

(22) International Filing Date:
27 November 2020 (27.11.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/941,255 27 November 2019 (27.11.2019) US

(71) Applicant: **PROMEGA CORPORATION** [US/US]; 2800 Woods Hollow Road, Madison, Wisconsin 53713 (US).

(72) Inventors: **KINCAID, Virginia**; c/o Promega Corporation, 2800 Woods Hollow Road, Madison, Wisconsin 53713 (US). **ENCELL, Lance P.**; c/o Promega Corporation, 2800 Woods Hollow Road, Madison, Wisconsin 53713 (US). **HALL, Mary**; c/o Promega Corporation, 2800 Woods Hollow Road, Madison, Wisconsin 53713 (US). **KILLORAN, Michael**; c/o Promega Corporation, 2800 Woods Hollow Road, Madison, Wisconsin 53713 (US). **WOOD, Keith**; c/o Promega Corporation, 2800 Woods Hollow Road, Madison, Wisconsin 53713 (US). **DART, Melanie**; c/o Promega Corporation, 2800 Woods Hollow Road, Madison, Wisconsin 53713 (US). **HOANG, Trish**; c/o Promega Corporation, 2800 Woods Hollow Road, Madison, Wisconsin 53713 (US).

(74) Agent: **STAPLE, David W.**; Casimir Jones, S.C., 2275 Deming Way, Suite 310, Middleton, Wisconsin 53562 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title:** MULTIPARTITE LUCIFERASE PEPTIDES AND POLYPEPTIDES

(57) **Abstract:** Provided herein are bioluminescent polypeptides and compositions and methods for the assembly of a tripartite or multipartite bioluminescent complex. In particular embodiments, a bioluminescent complex is formed upon the interaction of three or more peptide and/or polypeptide components.

WO 2021/108765 A1

MULTIPARTITE LUCIFERASE PEPTIDES AND POLYPEPTIDES

CROSS-REFERENCE TO RELATED APPLICATION

The present application claims priority to U.S. Provisional Patent Application Serial No. 62/941,255 filed November 27, 2019, which is hereby incorporated by reference in its entirety.

FIELD

Provided herein are bioluminescent polypeptides and compositions and methods for the assembly of tripartite or multipartite bioluminescent complexes. In particular embodiments, a bioluminescent complex is formed upon the interaction of three or more peptide and/or polypeptide components.

BACKGROUND

Biological processes and analyte detection rely on the co-localization and interactions between molecules, macromolecules, and molecular complexes. In order to understand such processes, and to develop techniques and compounds to manipulate them for research, clinical, and other practical applications, it is necessary to have tools available to detect and monitor these co-localizations/interactions. The study of these interactions, particularly under physiological conditions (e.g., at normal expression levels for monitoring protein interactions) or in complex sample matrices (e.g. blood samples, environmental samples), requires high sensitivity.

SUMMARY

Provided herein are bioluminescent polypeptides and compositions and methods for the assembly of a tripartite or multipartite bioluminescent complex. In particular embodiments, a bioluminescent complex is formed upon the interaction of three or more peptide and/or polypeptide components.

Experiments conducted during development of embodiments herein demonstrate the assembly of a bioluminescent complex, capable of generating luminescence in the presence of an appropriate substrate (e.g., a coelenterazine or a coelenterazine analog substrate), from complementary polypeptide(s) and peptide(s) that collectively span the the length (or >75% of the length, >80% of the length, >85% of the length, >90% of the length, >95% of the

length, or more) of a luciferase base sequence (or collectively comprise at least 40% sequence identity to a luciferase base sequence (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75% >80%, >85%, >90%, >95%, or more). In some embodiments, "complementary" polyptpide(s) and peptide(s) are separate molecules that each correspond to a portion of a luciferase base sequence. Through structural complementarity, they assemble to form a bioluminescent complex.

Additional experiments were conducted during development of embodiments herein to develop monomeric bioluminescent polypeptides with enhanced characteristics (e.g., stability, luminescence, etc.).

In some embodiments, the complementary polypeptide(s) and peptide(s) are fragments of a luciferase base sequence that assmble to form a bioluminescent complex. In some embodiments, the fragments collectively comprise the full length of the luciferase base sequence. In some embodiments, the fragments collectively comprise at least 75% of the full length of the luciferase base sequence (e.g., >75% of the length, >80% of the length, >85% of the length, >90% of the length, >95% of the length, or more).

In some embodiments, the complementary polypeptide(s) and peptide(s) are variants of portions of a luciferase base sequence individually comprising at least 40% sequence identity to the corresponding portion of the luciferase base sequence (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75% >80%, >85%, >90%, >95%, or more) that assmble to form a bioluminescent complex. In some embodiments, the complementary polypeptide(s) and peptide(s) are variants of portions of a luciferase base sequence collectively comprising at least 40% sequence identity to the entire luciferase base sequence (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75% >80%, >85%, >90%, >95%, or more) that assemble to form a bioluminescent complex. In some embodiments, the fragments collectively comprise the full length of the luciferase base sequence. In some embodiments, the complementary polypeptide(s) and peptide(s) collectively comprise at least 75% of the full length of the luciferase base sequence (e.g., >75% of the length, >80% of the length, >85% of the length, >90% of the length, >95% of the length, or more).

Examples of luciferase base sequences include SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 788, and SEQ ID NO: 789. Some embodiments herein provide a polypeptide component that is a fragment of the luciferase base sequence (e.g., SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 788, and SEQ ID NO: 789) or a variant thereof (e.g., >40%, >45%,

>50%, >55%, >60%, >65%, >70%, >75% >80%, >85%, >90%, >95% sequence identity), and one or more complementary peptide(s) and/or polypeptide(s) that collectvily span the remainder of the luciferase base sequence. For example, if a luciferase base sequence is 170 amino acid residues in length, an exemplary polypeptide component may be, for example 102, 124, 133, or 148 amino acids in length, and 1, 2, 3, 4, 5, or more complementary peptides correspond to the remaining 68, 46, 37, or 22 amino acids. In some embodiments, each polypeptide component individually comprises at least 40% sequence identity (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75% >80%, >85%, >90%, >95%, or more) to the corresponding portion of the luciferase base sequence.

In some embodiments, provided herein are systems or kits comprising comprising: (a) a polypeptide component comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to a polypeptide fragment of SEQ ID NO: 788 or SEQ ID NO: 789; and (b) one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to the complementary portion of SEQ ID NO: 788 or SEQ ID NO: 789; wherein a bioluminescent signal produced by a bioluminescent complex assembled from the polypeptide component and one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when compared to a bioluminescent signal produced by the polypeptide component or one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides and the coelenterazine substrate alone. In some embodiments, the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 790, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 794. In some embodiments, the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 791, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 795. In some embodiments, the

polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 792, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 796. In some embodiments, the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 793, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 797. In some embodiments, the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 790, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 798. In some embodiments, the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 791, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 799. In some embodiments, the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 792, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 800. In some embodiments, the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 793, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID 801. In some embodiments, the bioluminescent signal is substantially increased when the polypeptide component associates with the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides.

In some embodiments, polypeptide component and/or one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides comprise amino acid sequences that are not a naturally occurring sequences or fragments thereof. In some embodiments, polypeptide component and/or one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides comprise a non-natural amino acid, an amino acid analog, and/or peptoid amino acids. In some embodiments, the polypeptide component and/or one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides are present as fusions with one or more additional amino acid sequences. In some embodiments, the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety. In some embodiments, the additional amino acid sequence is a binding moiety selected from the group consisting of antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide. In some embodiments, the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide. In some embodiments, the additional amino acid sequence is a protein of interest and is a candidate drug target. In some embodiments, provided herein are bioluminescent complexes comprising the polypeptide component and one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides of the systems or kits described herein.

In some embodiments, provided herein are systems or kits comprising two or more peptide, dipeptide, tripeptide and/or polypeptide components collectively comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 788 or SEQ ID NO: 789; wherein a bioluminescent signal produced by the bioluminescent complex in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when compared to a bioluminescent

signal produced by the polypeptide or one or more complementary peptides and the coelenterazine substrate alone. In some embodiments, a system of kit comprises a polypeptide component having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 790, and the one or more complementary peptides, dipeptides, and or tripeptides collectively having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 794. In some embodiments, the polypeptide comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 791, and the one or more complementary peptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 795. In some embodiments, the polypeptide comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 792, and the one or more complementary peptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 796. In some embodiments, the polypeptide comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 793, and the one or more complementary peptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 797. In some embodiments, the polypeptide comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 790, and the one or more complementary peptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 798. In some embodiments, the polypeptide comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 791, and the one or more complementary peptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 799. In some embodiments, the polypeptide comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 792, and the one or more complementary peptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%,

80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 800. In some embodiments, the polypeptide comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 793, and the one or more complementary peptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 801. In some embodiments, the bioluminescent signal is substantially increased when the polypeptide associates with the one or more complementary peptides. In some embodiments, the polypeptide and/or one or more complementary peptides comprise amino acid sequences that are not a naturally occurring sequences or fragments thereof. In some embodiments, polypeptide and/or one or more complementary peptides comprise a non-natural amino acid, an amino acid analog, and/or peptoid amino acids. In some embodiments, the polypeptide and/or one or more complementary peptides are present as fusions with one or more additional amino acid sequences. In some embodiments, the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety. In some embodiments, the additional amino acid sequence is a binding moiety selected from the group consisting of antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide. In some embodiments, the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism with a second co-localization polypeptide. In some embodiments, the additional amino acid sequence is a protein of interest and is a candidate drug target. In some embodiments, provided herein are bioluminescent complexes comprising the two or more peptide, dipeptide, tripeptide, and/or polypeptide components of the systems or kits described herein.

In some embodiments, provided herein are methods comprising: (a) combining: (i)a polypeptide component comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to a polypeptide fragment of SEQ ID NO: 788 or SEQ ID NO: 789; (ii) one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to the complementary portion of SEQ ID NO: 788 or SEQ ID NO: 789; and (iii) a coelenterazine or a coelenterazine analog substrate; and (b) detecting luminescence, wherein a greater level of luminescence compared to a level of luminescence produced by the polypeptide component and a coelenterazine or a coelenterazine analog alone indicates formation of a bioluminescent complex of the polypeptide component and the one or more complementary peptides. In some embodiments, one or more of the polypeptide component and the first and second peptides are expressed in a cell, added to a cell exogenously, and/or added to a sample. In some embodiments, (i) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 790 and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 794; (ii) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 791, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 795; (iii) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 792, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 796; (iv) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 793, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to

SEQ ID NO: 797; (v) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 790, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 798; (vi) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 791, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 799; (vii) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 792, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 800; or (viii) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 793, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID 801.

In some embodiments, provided herein are methods comprising: (a) combining: (i) two or more peptide, dipeptide, tripeptide, and/or polypeptide components collectively comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to the full length of SEQ ID NO: 788 or SEQ ID NO: 789; and (ii) a coelenterazine or a coelenterazine analog substrate; and (b) detecting luminescence, wherein a greater level of luminescence compared to a level of luminescence produced by the peptide, dipeptide, tripeptide, and/or polypeptide components and the coelenterazine or coelenterazine analog indicates formation of a bioluminescent complex of the peptide and polypeptide components. In some embodiments, one or more of the polypeptide components and the first and second peptides may be expressed in a cell, added to a cell exogenously, and/or added to a sample. In some embodiments, (i) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%,

80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 790, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 794; (ii) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 791, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 795; (iii) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 792, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 796; (iv) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 793, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 797; (v) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 790, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 798; (vi) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 791, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 799; (vii) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 792, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 800; or

(viii) the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID NO: 793, and the one or more complementary peptides, dipeptides, tripeptides, and/or polypeptides collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to SEQ ID 801.

In some embodiments, provided herein are methods of detecting an interaction between a first molecular entity and a second molecular entity, the method comprising: (a) tagging the first molecular entity with a first peptide, dipeptide, or tripeptide tag; (b) tagging the second molecular entity with a second peptide, dipeptide, or tripeptide tag; (c) combining the tagged first molecular entity and the tagged second molecular entity and/or allowing the tagged first molecular entity and the tagged second molecular entity to come into contact with one another; (d) adding one or more peptide, dipeptide, tripeptide, and/or polypeptide components, wherein the first peptide, dipeptide, or tripeptide tag, the second peptide, dipeptide, or tripeptide tag, and the peptide, dipeptide, tripeptide, and/or polypeptide components collectively comprise an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with the entirety of SEQ ID NO: 788 or 789, and capable of assembling to form a bioluminescent complex; (e) adding a coelenterazine or a coelenterazine analog substrate; and (f) detecting a luminescent signal produced by the bioluminescent complex, wherein the magnitude of the luminescent signal correlates to the strength of the interaction between the first molecular entity and the second molecular entity. In some embodiments, the first molecular entity and/or the second molecular entity is a protein of interest or a peptide of interest and tagging comprises generating a fusion of the first molecular entity and/or the second molecular entity with the first tag and/or second tag. In some embodiments, the first molecular entity and/or the second molecular entity is a small molecule and tagging comprises directly or indirectly linking the first molecular entity and/or the second molecular entity with the first tag and/or second tag. In some embodiments, one of the first molecular entity and the second molecular entity is a drug or drug candidate and the other is a drug target or candidate drug target, and the bioluminescent signal indicates binding of the drug or drug candidate to the other is a drug target or candidate drug target. In some embodiments, combining the tagged first molecular entity and the tagged second molecular entity comprises expressing one or both within a cell and/or adding one or both to a cell.

In some embodiments, provided herein are methods of detecting an interaction between a first protein or peptide entity and a second protein or peptide entity with a cell comprising, the method comprising: (a) expressing within the cell a fusion comprising the first protein or peptide entity and a first peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a first portion of SEQ ID NO: 788 or 789; (b) expressing within the cell a fusion comprising the second protein or peptide entity and a second peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a second portion of SEQ ID NO: 788 or 789; (c) expressing within the cell one or more peptide, dipeptide, tripeptide, and/or polypeptide components comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a third portion of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with the entirety of SEQ ID NO: 788 or 789 and are configured to produce a bioluminescent complex upon interaction of the first protein or peptide entity and the second protein or peptide entity; (d) adding a coelenterazine or a coelenterazine analog substrate to the cell; and (e) detecting a luminescent signal produced by the bioluminescent complex, wherein the magnitude of the luminescent signal correlates to the strength of the interaction between the first protein or peptide entity and the second protein or peptide entity.

In some embodiments, provided herein are methods of detecting co-localization of a first molecular entity and a second molecular entity, the method comprising: (a) tagging the first molecular entity with a first peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a first portion of SEQ ID NO: 788 or 789; (b) tagging the second molecular entity with a second peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a second portion of SEQ ID NO: 788 or 789; (c) combining the tagged first molecular entity and the tagged second molecular entity in the same system; (d) adding one or more peptide,

dipeptide, tripeptide, and/or polypeptide components to the system, the components having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a third portion of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with the entirety of SEQ ID NO: 788 or 789, wherein the first peptide tag, the second peptide tag, and components are configured to produce a bioluminescent complex upon co-localization of the first molecular entity and the second molecular entity; (e) adding a coelenterazine or a coelenterazine analog substrate to the system; and (f) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates co-localization of the first molecular entity and the second molecular entity within the system, and/or wherein the magnitude of the luminescent signal correlates to the amount of co-localization within the system of the first molecular entity and the second molecular entity. In some embodiments, the system comprises a cell, tissue, organ, whole organism, and/or a biochemical, non-cellular sample. In some embodiments, the first molecular entity and/or the second molecular entity is a protein of interest or a peptide of interest, and tagging comprises generating a fusion of the first molecular entity and/or the second molecular entity with the first tag and/or peptide tag. In some embodiments, the first molecular entity and/or the second molecular entity is a small molecule and tagging comprises directly or indirectly linking the first molecular entity and/or the second molecular entity with the first tag and/or second tag. In some embodiments, combining the tagged first molecular entity and the tagged second molecular entity comprises expressing one or both within the system and/or adding one or both to the system.

In some embodiments, provided herein are methods of detecting co-localization of a first protein or peptide entity and a second protein or peptide entity with a cell comprising, the method comprising: (a) expressing within the cell a fusion comprising the first protein or peptide entity and a first peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a first portion of SEQ ID NO: 788 or 789; (b) expressing within the cell a fusion comprising the second protein or peptide entity and a second peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or

more) sequence identity with a second portion of SEQ ID NO: 788 or 789; (c) expressing with the cell one or more peptide, dipeptide, tripeptide, or polypeptide components having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a third portion of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with the entirety of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components are configured to produce a bioluminescent complex upon co-localization of the first protein or peptide entity and the second protein or peptide entity; (d) adding a coelenterazine or a coelenterazine analog substrate to the cell; and (e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates co-localization of the first protein or peptide entity and the second protein or peptide entity within the cell, and/or wherein the magnitude of the luminescent signal correlates to the amount of co-localization within the system of the first protein or peptide entity and the second protein or peptide entity.

In some embodiments, provided herein are methods of detecting a target molecule, wherein the target molecule displays a first antigen, epitope, or sequence and a distinct second antigen, epitope, or sequence, the method comprising: (a) contacting a sample containing the target molecule with (i) a first primary binding moiety that recognizes the first antigen, epitope, or sequence and (ii) a second primary binding moiety that recognizes the second antigen, epitope, or sequence and allowing the first and second primary binding moieties to bind to the first and second antigens, epitopes, or sequences; (b) contacting the sample with (i) a first secondary binding moiety conjugated to a first tag and (ii) a second secondary binding moiety conjugated to second tag, wherein the first secondary binding moiety recognizes the first primary binding moiety and the second secondary binding moiety recognizes the second primary binding moiety, wherein the first or second tags comprises amino acid sequences having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with first and second portions of SEQ ID NO: 788 or 789; (c) allowing the first and second secondary binding moieties to bind to the first and second primary binding moieties; (d) contacting the sample with comprising one or more peptide, dipeptide, tripeptide, and/or polypeptide components having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more)

sequence identity with a third portion of SEQ ID NO: 788 or 789; wherein the first tag, the second tag, and the components collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with the entirety of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components are configured to produce a bioluminescent complex upon interaction; (d) contacting the sample with a coelenterazine or a coelenterazine analog substrate; and (e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates the presence of the target molecule, and/or wherein the magnitude of the luminescent signal correlates to the amount of target molecule within the sample. In some embodiments, the binding moieties are independently selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the target molecule is a protein, nucleic acid, or small molecule. In some embodiments, the sample is *in vitro* or *in vivo*.

In some embodiments, provided herein are methods of detecting a target molecule, wherein the target molecule displays a first antigen, epitope, or sequence and a distinct second antigen, epitope, or sequence, the method comprising: (a) contacting the sample with (i) a first binding moiety conjugated to a first tag and (ii) a second binding moiety conjugated to second tag, wherein the first secondary binding moiety recognizes the first antigen, epitope, or sequence, and the second binding moiety recognizes the second antigen, epitope, or sequence, wherein the first tag comprises an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a first portion of SEQ ID NO: 788 or 789, and wherein the second tag comprises an amino acid sequence with a first portion of SEQ ID NO: 788 or 789; (b) allowing the first and second binding moieties to bind to the first and second antigens, epitope, or sequences; (c) contacting the sample with a peptide, dipeptide, tripeptide, or polypeptide component having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with a third portion of SEQ ID NO:

788 or 789, wherein the first tag, the second tag, and the components collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with the entirety of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components are configured to produce a bioluminescent complex upon interaction; (d) contacting the sample with a coelenterazine or a coelenterazine analog substrate; and (e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates the presence of the target molecule, and/or wherein the magnitude of the luminescent signal correlates to the amount of target molecule within the sample. In some embodiments, the binding moieties are independently selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the target molecule is a protein, nucleic acid, or small molecule. In some embodiments, the sample is *in vitro*, *in vivo,* or a biochemical sample.

In some embodiments, provided herein are peptides, dipeptides, tripeptides, and/or polypeptides listed in Table 1, Table 9, or Table 10. In some embodiments, a single peptide, dipeptide, tripeptide, or polypeptide listed in Table 1, Table 9, or Table 10 is provided (e.g., as a reagent, as a tag, etc.). In some embodiments, a pair (2) or set (e.g., 2, 3, 4, 5, or more) of peptides, dipeptides, tripeptides, and/or polypeptides listed in Table 1, Table 9, or Table 10 are provided. In particular, pairs or sets of the peptides, dipeptides, tripeptides, and/or polypeptides are provided that are complementary and are capable of forming a bioluminescent complex upon interaction (e.g., facilitated, unfacilitated) with one another.

In some embodiments, the polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to one of SEQ ID NOS: 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 933, 935, 937, 939, 941, 943, 945, 947, 949, 951, 953, 955, 957, 959, 961, 963, 965, 967, 969, 117, 119, 121, 123, 125, 127, 129, 131, 596, 598, 600, 602, 604, 606, 608, 610, 612, 614, 616, 618, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648,

650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 766, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 802, 804, 806, 808, 813, 815, or 829. In some embodiments, a polypeptide comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to one of SEQ ID NOS: 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 933, 935, 937, 939, 941, 943, 945, 947, 949, 951, 953, 955, 957, 959, 961, 963, 965, 967, 969, 117, 119, 121, 123, 125, 127, 129, 131, 596, 598, 600, 602, 604, 606, 608, 610, 612, 614, 616, 618, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 766, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 802, 804, 806, 808, 813, 815, or 829 is provided with one or more peptides or dipeptides capable of forming a bioluminescent complex. In some embodiments, suitable fragments of SEQ ID NOS: 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 933, 935, 937, 939, 941, 943, 945, 947, 949, 951, 953, 955, 957, 959, 961, 963, 965, 967, 969, 117, 119, 121, 123, 125, 127, 129, 131, 596, 598, 600, 602, 604, 606, 608, 610, 612, 614, 616, 618, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746, 748, 750, 752, 754, 756, 758, 760, 762, 764, 766, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 802, 804, 806, 808, 813, 815, or 829 of polypeptides having % or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to one of SEQ ID NOS: 909, 911, 913, 915, 917, 919, 921, 923, 925, 927, 929, 931, 933, 935, 937, 939, 941, 943, 945, 947, 949, 951, 953, 955, 957, 959, 961, 963, 965, 967, 969, 117, 119, 121, 123, 125, 127, 129, 131, 596, 598, 600, 602, 604, 606, 608, 610, 612, 614, 616, 618, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, 668, 670, 672, 674, 676, 678, 680, 682, 684, 686, 688, 690, 692, 694, 696, 698, 700, 702, 704, 706, 708, 710, 712, 714, 716, 718, 720, 722, 724, 726, 728, 730, 732, 734, 736, 738, 740, 742, 744, 746,

748, 750, 752, 754, 756, 758, 760, 762, 764, 766, 768, 770, 772, 774, 776, 778, 780, 782, 784, 786, 802, 804, 806, 808, 813, 815, or 829 are provided. In some embodiments, such fragments are capable of forming bioluminescent complexes with a suitable set of peptides, dipeptides, tripeptides, polypeptides, etc. provided herein.

In some embodiments, the peptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to one of SEQ ID NOS: 900-907. In some embodiments, a polypeptide comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity to one of SEQ ID NOS: 900-907 is provided with one or more peptides, dipeptides, tripeptides, polypeptides, etc., capable of forming a bioluminescent complex.

In some embodiments, provided herein are peptides, dipeptides, tripeptides, and/or polypeptides having at least 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) seqeunce identity with one or more of the peptides, dipeptides, tripeptides, and/or polypeptides listed in Table 1, Table 9, or Table 10. In some embodiments, a single peptide, dipeptide, tripeptide, or polypeptide having at least 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) seqeunce identity with one or more of the peptides, dipeptides, tripeptides, and/or polypeptides listed in Table 1, Table 9, or Table 10 is provided (e.g., as a reagent, as a tag, etc.). In some embodiments, a pair (2) or set (e.g., 2, 3, 4, 5, or more) of peptides, dipeptides, tripeptides, and/or polypeptides having at least 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) seqeunce identity with one or more of the peptides, dipeptides, tripeptides, and/or polypeptides listed in Table 1, Table 9, or Table 10 is provided are provided. In particular, pairs or sets of the peptides, dipeptides, tripeptides, and/or polypeptides are provided that are complementary and are capable of forming a bioluminescent complex upon interaction (e.g., facilitated, unfacilitated) with one another.

In some embodiments, provided herein are polypeptides comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with one of SEQ ID NO: 790, 791, 792, or 793. In some embodiments, the polypeptide is provided alone or as a pair/set with complementary peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, fusions of polypeptides herein with proteins of interest, intearction elements, colocalization

elements, etc., are provided. In some embodiments, nucleic acids and vectors encoding the polypeptides and fusions thereof or provided.

In some embodiments, provided herein are peptides comprising SEQ ID NO: 817, 818, 819, 13, 15, 23, or 25. In some embodiments, provided herein are peptides comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with one of SEQ ID NO: 817, 818, 819, 13, 15, 23, or 25. In some embodiments, the peptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, fusions of peptides herein with proteins of interest, intearction elements, colocalization elements, etc., are provided. In some embodiments, nucleic acids and vectors encoding the peptides and fusions thereof or provided. In some embodiments, molecules of interest and/or proteins of interest are tagged with a peptide herein.

In some embodiments, provided herein is a β6-7-like dipeptide comprising SEQ ID NOS: 817 and 818. In some embodiments, provided herein is a β6-7-like dipeptide having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NOS: 817 and 818. In some embodiments, the dipeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, nucleic acids and vectors encoding the dipeptides and fusions thereof or provided. In some embodiments, molecules of interest and/or proteins of interest are tagged with a dipeptide herein.

In some embodiments, provided herein is a β7-8-like dipeptide comprising SEQ ID NOS: 818 and 819. In some embodiments, provided herein is a β7-8-like dipeptide having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NOS: 818 and 819. In some embodiments, the dipeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, the dipeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, nucleic acids and vectors encoding the

dipeptides and fusions thereof or provided. In some embodiments, molecules of interest and/or proteins of interest are tagged with a dipeptide herein.

In some embodiments, provided herein is a β8-9-like dipeptide comprising SEQ ID NOS: 819/23 or 819/25. In some embodiments, provided herein is a β8-9-like dipeptide having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with NOS: 819/23 or 819/25. In some embodiments, the dipeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, the dipeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, nucleic acids and vectors encoding the dipeptides and fusions thereof or provided. In some embodiments, molecules of interest and/or proteins of interest are tagged with a dipeptide herein.

In some embodiments, provided herein is a β9-10-like dipeptide comprising SEQ ID NOS: 23/13, 23/15, 25/13 or 25/15. In some embodiments, provided herein is a β8-9-like dipeptide having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with NOS: SEQ ID NOS: 23/13, 23/15, 25/13 or 25/15. In some embodiments, the dipeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, the dipeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, nucleic acids and vectors encoding the dipeptides and fusions thereof or provided. In some embodiments, molecules of interest and/or proteins of interest are tagged with a dipeptide herein.

In some embodiments, provided herein is a β6-8-like tripeptide comprising SEQ ID NOS: 817-819. In some embodiments, provided herein is a β6-8-like tripeptide having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with NOS: SEQ ID NOS: 817-819. In some embodiments, the tripeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, the tripeptide is provided alone or as a pair/set with complementary

polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, nucleic acids and vectors encoding the tripeptides and fusions thereof or provided. In some embodiments, molecules of interest and/or proteins of interest are tagged with a tripeptide herein.

In some embodiments, provided herein is a β7-9-like tripeptide comprising SEQ ID NOS: 818/819/23 or 818/819/25. In some embodiments, provided herein is a β7-9-like tripeptide having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with NOS: SEQ ID NOS: 818/819/23 or 818/819/25. In some embodiments, the tripeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, the tripeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, nucleic acids and vectors encoding the tripeptides and fusions thereof or provided. In some embodiments, molecules of interest and/or proteins of interest are tagged with a tripeptide herein.

In some embodiments, provided herein is a β8-10-like tripeptide comprising SEQ ID NOS: 819/23/13, 819/23/15, 819/25/13, or 819/25/15. In some embodiments, provided herein is a β7-9-like tripeptide having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NOS: 819/23/13, 819/23/15, 819/25/13, or 819/25/15. In some embodiments, the tripeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, the tripeptide is provided alone or as a pair/set with complementary polypeptide and/or other peptide(s), dipeptide(s), and/or tripeptide for the formation of a bioluminescent complex. In some embodiments, nucleic acids and vectors encoding the tripeptides and fusions thereof or provided. In some embodiments, molecules of interest and/or proteins of interest are tagged with a tripeptide herein.

In some embodiments, provided herein are peptides comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%,

85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the peptide contacts a second peptide consisting of SEQ ID NO: 25 and a polypeptide complement consisting of SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 when compared to a bioluminescent signal produced by the peptide and the coelenterazine or coelenterazine analog substrate alone. In some embodiments, the bioluminescent signal is substantially increased when the peptide associates with the second peptide and the polypeptide complement. In some embodiments, the peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 6 and/or SEQ ID NO: 9, wherein the traits are selected from: affinity for the second peptide and the polypeptide complement or enhanced expression, solubility, stability, and/or bioluminescent activity when combined with the second peptide and the polypeptide complement. In some embodiments, the amino acid sequence is not a naturally occurring protein (e.g., not SEQ ID NO: 1), not a mutant version thereof (e.g., not SEQ ID NO: 3), not a fragment of a naturally occurring protein (e.g., not SEQ ID NOS: 5-7), and not a fragment of a mutant version thereof (e.g., not one of SEQ ID NOS: 8-10). In some embodiments, the amino acid sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids. In some embodiments, a peptide is chemically conjugated to a linker, reactive moiety, detection element (e.g., fluorophore), interaction/binding element, etc.

In some embodiments, provided herein are fusion polypeptides (e.g., genetic fusions (or alternatively, chemical conjugations or synthetically produced)) comprising a peptide described in the preceding paragraph and an additional amino acid sequence or compound (e.g. small molecule drug). In some embodiments, the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and/or a binding moiety. In some embodiments, the additional amino acid sequence is a binding moiety selected from the group consisting of an antibody (e.g., polyclonal, monoclonal, and/or recombinant), antibody fragment,protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, a Ig binding domain of protein L, protein M, an Ig binding domain of protein M, peptide nucleic acid, DARPin, affimer, a purified protein (e.g., an analyte or a protein that binds to an analyte), and analyte binding domain(s) of proteins.

In some embodiments, the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide. In some embodiments, the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism with a second co-localization polypeptide. In some embodiments, the additional amino acid sequence is a protein of interest and is a candidate drug target.

In some embodiments, provided herein are peptides comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the peptide contacts a second peptide consisting of SEQ ID NO: 23 and a polypeptide complement consisting of SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 when compared to a bioluminescent signal produced by the peptide and the coelenterazine or coelenterazine analog substrate alone. In some embodiments, the bioluminescent signal is substantially increased when the peptide associates with the second peptide and the polypeptide complement. In some embodiments, the peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 7 and/or SEQ ID NO: 10, wherein the traits are selected from: affinity for the second peptide and the polypeptide complement or enhanced expression, solubility, stability, and/or bioluminescent activity when combined with the second peptide and the polypeptide complement. In some embodiments, the amino acid sequence is not a naturally occurring protein (e.g., not SEQ ID NO: 1), not a mutant version thereof (e.g., not SEQ ID NO: 3), not a fragment of a naturally occurring protein (e.g., not SEQ ID NOS: 5-7), and not a fragment of a mutant version thereof (e.g., not one of SEQ ID NOS: 8-10). In some embodiments, the amino acid sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids. In some embodiments, a peptide is chemically conjugated to a linker, reactive moiety, detection element (e.g., fluorophore), interaction/binding element, etc.

In some embodiments, provided herein are fusion polypeptides (e.g., genetic fusions, synthetically-produced fusions, chemical conjugates, enzymatic conjugates, etc.) comprising a peptide described in the preceding paragraph and an additional amino acid sequence. In

some embodiments, the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety. In some embodiments, the additional amino acid sequence is a binding moiety selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, a Ig binding domain of protein L, protein M, an Ig binding domain of protein M, peptide nucleic acid, DARPin, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide. In some embodiments, the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide. In some embodiments, the additional amino acid sequence is a protein of interest and is a candidate drug target.

In some embodiments, provided herein are compositions comprising: (a) a first peptide comprising an amino acid sequence having greater than 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more), but less than 100% sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10; and (b) a second peptide comprising an amino acid sequence having greater than 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more), but less than 100% sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6, SEQ ID NO: 9, and SEQ ID NO: 29; wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the first peptide contacts the second peptide and a polypeptide complement consisting of SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 when compared to a bioluminescent signal produced by the first peptide and/or the second peptide and the coelenterazine substrate alone. In some embodiments, the bioluminescent signal is substantially increased when the first peptide associates with the second peptide and the polypeptide complement. In some embodiments, the first peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 7 and/or SEQ ID NO: 10, and the

second peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 6, SEQ ID NO: 9, and SEQ ID NO: 29, wherein the traits are selected from: affinity for the second peptide and the polypeptide complement or enhanced expression, solubility, stability, and/or bioluminescent activity when combined with the second peptide and the polypeptide complement. In some embodiments, the amino acid sequence of the first and/or second peptide is not a naturally occurring protein or a fragment thereof. In some embodiments, the amino acid sequence of the first and/or second peptide contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

In some embodiments, provided herein are compositions comprising fusion polypeptides comprising the first and second peptides of described in the preceding paragraph and an additional amino acid sequence. In some embodiments, the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety. In some embodiments, the additional amino acid sequence is a binding moiety selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, a Ig binding domain of protein L, protein M, an Ig binding domain of protein M, peptide nucleic acid, DARPin, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide. In some embodiments, the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide. In some embodiments, the additional amino acid sequence is a protein of interest and is a candidate drug target.

In some embodiments, provided herein are polypeptides comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the polypeptide contacts a

first peptide consisting of SEQ ID NO: 23 and a second peptide consisting of SEQ ID NO: 25 when compared to a bioluminescent signal produced by the peptide and the coelenterazine or a coelenterazine analog substrate alone. In some embodiments, the bioluminescent signal is substantially increased when the polypeptide associates with the first and second peptides. In some embodiments, the polypeptide exhibits enhancement of one or more traits compared to a polypeptide of SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the traits are selected from: affinity for the first and/or second peptides or enhanced expression, solubility, stability, and/or bioluminescent activity when combined with the first and second peptides. In some embodiments, the amino acid sequence is not a naturally occurring protein (e.g., not SEQ ID NO: 1), not a mutant version thereof (e.g., not SEQ ID NO: 3), not a fragment of a naturally occurring protein (e.g., not SEQ ID NOS: 5-7), and not a fragment of a mutant version thereof (e.g., not one of SEQ ID NOS: 8-10). In some embodiments, the amino acid sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

In some embodiments, provided herein are fusion polypeptides (e.g., genetic fusions, synthetically-produced fusions, chemical conjugates, enzymatic conjugates, etc.) comprising a polypeptide described in the preceding paragraph and an additional amino acid sequence, nucleic acid sequence, or other fused or appended molecule. In some embodiments, the additional sequence or other molecule is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety. In some embodiments, the additional sequence or other molecule is a binding moiety selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, a Ig binding domain of protein L, protein M, an Ig binding domain of protein M, peptide nucleic acid, DARPin, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the additional sequence or other fused or appended molecule is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide. In some embodiments, the additional sequence or other fused or appended molecule is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism with a

second co-localization polypeptide. In some embodiments, the additional sequence or other fused or appended molecule is a protein of interest and is a candidate drug target.

In some embodiments, provided herein are polypeptides comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, and/or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the polypeptide contacts a first peptide consisting of SEQ ID NO: 23 and a second peptide consisting of SEQ ID NO: 25 when compared to a bioluminescent signal produced by the peptide and the coelenterazine or coelenterazine analog substrate alone. In some embodiments, the bioluminescent signal is substantially increased when the polypeptide associates with the first and second peptides. In some embodiments, the polypeptide exhibits enhancement of one or more traits compared to a polypeptide of SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the traits are selected from: affinity for the first and/or second peptides or enhanced expression, solubility, stability, and/or bioluminescent activity when combined with the first and second peptides. In some embodiments, the amino acid sequence is not a naturally occurring protein (e.g., not SEQ ID NO: 1), not a mutant version thereof (e.g., not SEQ ID NO: 3), not a fragment of a naturally occurring protein (e.g., not SEQ ID NOS: 5-7), and not a fragment of a mutant version thereof (e.g., not one of SEQ ID NOS: 8-10). In some embodiments, the amino acid sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

In some embodiments, provided herein are fusion polypeptides (e.g., genetic fusions, synthetically-produced fusions, chemical conjugates, enzymatic conjugates, etc.) comprising a peptide described in the preceding paragraph and an additional amino acid sequence. In some embodiments, the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety. In some embodiments, the additional amino acid sequence is a binding moiety selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, a Ig binding domain of protein L, protein M, an Ig binding domain of protein M, peptide

nucleic acid, DARPin, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide. In some embodiments, the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism with a second co-localization polypeptide. In some embodiments, the additional amino acid sequence is a protein of interest and is a candidate drug target.

In some embodiments, provided herein are β9/β10-like dipeptides comprising an amino acid sequence having greater than 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more), but less than 100% sequence identity with SEQ ID NO: 35 and less than 100% sequence identity with SEQ ID NO: 205 and SEQ ID NO: 206, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the peptide contacts a polypeptide complement consisting of SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 when compared to a bioluminescent signal produced by the peptide and the coelenterazine or a coelenterazine analog substrate alone. In some embodiments, a dipeptide (e.g., β_9/β_{10}-like dipeptide) associates (e.g., forms a bioluminescent complex) with a polypeptide component described herein (e.g., β_{1-8}-like polypeptide) without facilitation (e.g., from interaction elements). In other embodiments, a dipeptide (e.g., β_9/β_{10}-like dipeptide) and polypeptide component described herein (e.g., β_{1-8}-like polypeptide) will not form a bioluminescent complex without facilitation (e.g., from interaction elements), but will associate (e.g., form a bioluminescent complex) with facilitation from appropriate interaction elements. In some embodiments, the bioluminescent signal is substantially increased when the peptide associates with the polypeptide complement. In some embodiments, the peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 205 and/or SEQ ID NO: 206, wherein the traits are selected from: affinity for the polypeptide complement or enhanced expression, solubility, stability, and/or bioluminescent activity when combined with the polypeptide complement. In some embodiments, the amino acid sequence is not a naturally occurring protein or a fragment thereof. In some embodiments, the amino acid

sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

In some embodiments, provided herein are fusion polypeptides (e.g., genetic fusions, synthetically-produced fusions, chemical conjugates, enzymatic conjugates, etc.) comprising the β9/β10-like dipeptides described herein and an additional amino acid sequence. In some embodiments, the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety. In some embodiments, the additional amino acid sequence or other fused or appended molecule is a binding moiety selected from the group consisting of antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, a Ig binding domain of protein L, protein M, an Ig binding domain of protein M, peptide nucleic acid, DARPin, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the additional amino acid sequence or other fused or appended molecule is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide. In some embodiments, the additional amino acid sequence or other fused or appended molecule is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism with a second co-localization polypeptide. In some embodiments, the additional amino acid sequence or other fused or appended molecule is a protein of interest and is a candidate drug target.

In some embodiments, provided herein are nucleic acids and/or vectors coding for the peptides, polypeptides, and/or fusion polypeptides described herein. In some embodiments, provided herein are cells expressing nucleic acids and/or vectors coding for the peptides, polypeptides, and/or fusion polypeptides described herein. In some embodiments, synthetic production of the peptides, polypeptides, and/or fusion polypeptides described herein is provided. In some embodiments, the peptides, polypeptides, and/or fusion polypeptides described herein are chemically conjugated to additional moieties (e.g., interaction elements, co-localization elements, proteins of interest, molecules of interest, etc.).

In some embodiments, provided herein are bioluminescent complexes comprising: (a) a polypeptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%,

50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8; (b) a first peptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9; and (c) a second peptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10; wherein the bioluminescent complex produces substantially increased bioluminescence in the presence of a coelenterazine or a coelenterazine analog substrate when compared to a coelenterazine or a coelenterazine analog substrate in the presence of: the polypeptide alone, the first peptide alone, the second peptide alone, and any two of the polypeptide, the first peptide, and the second peptide. In some embodiments, the first peptide is a first peptide tag, wherein the second peptide is a second peptide tag, and wherein the first and second peptide tags are each linked to moieties that are independently selected from the group consisting of a molecule of interest, a peptide of interest, a protein of interest, an interaction element, a co-localization element, or a binding moiety. In some embodiments, the first peptide tag or the second peptide tag is linked to a drug or drug candidate, and the other peptide tag is linked to a drug target or candidate drug target, and wherein the intensity of the bioluminescence from the bioluminescent complex correlates to the affinity of the drug or drug candidate for the drug target or candidate drug target. In some embodiments, the first peptide tag is linked to a first interaction element, and the second peptide tag is linked to a second interaction element, and wherein the intensity of the bioluminescence from the bioluminescent complex correlates to the affinity of the first interaction element for the second interaction element under the conditions assayed (e.g., in some embodiments, the combination of the first peptide, second peptide, polypeptide component, and substrate do not form the bioluminescent complex (and produce significant light output (e.g., above background)) in the absence of an interaction between interaction elements). In some embodiments, the first peptide tag is linked to a first co-localization element, and the second peptide tag is linked to a second co-localization element, and wherein substantially increased bioluminescence indicates co-localization, but

not necessarily interaction, of the first co-localization element and the second co-localization element, under the conditions assayed.

In some embodiments, the peptides and polypeptide provided herein are not fragments of larger (e.g., pre-existing) proteins. In other embodiments, one or more peptides and/or polypeptides provided herein are fragments of larger (e.g., pre-existing) proteins.

In some embodiments, provided herein are methods comprising: (a) combining: (i) a first peptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9, (ii) a second peptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10, (iii) a polypeptide component comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon interaction of the first molecular entity and the second molecular entity, and (iv) a coelenterazine or a coelenterazine analog substrate; and (b) detecting luminescence, wherein a greater level of luminescence compared to a level of luminescence produced by the polypeptide component and a coelenterazine or a coelenterazine analog alone indicates formation of a bioluminescent complex of the polypeptide component and the first and second peptides. In some embodiments, one or more of the polypeptide component and the first and second peptides are expressed in a cell, added to a cell exogenously, and/or added to a sample.

In some embodiments, provided herein are methods of detecting an interaction between a first molecular entity and a second molecular entity, the method comprising: (a) tagging the first molecular entity with a first peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9; (b) tagging the second molecular entity with a second peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%,

45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10; (c) combining the tagged first molecular entity and the tagged second molecular entity; (d) adding a polypeptide component comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon interaction of the first molecular entity and the second molecular entity; (e) adding a coelenterazine or a coelenterazine analog substrate; and (f) detecting a luminescent signal produced by the bioluminescent complex, wherein the magnitude of the luminescent signal correlates with (e.g., is proportional to, is directly proportional to, etc.) the number of, strength of, favorability of, and/or stability of the interaction(s)) between the first molecular entity and the second molecular entity. In some embodiments, catalytic efficiency, substrate turnover, and/or specific activity of the resulting bioluminescent complex correlates with (e.g., is proportional to, is directly proportional to, etc.) the number of, strength of, favorability of, and/or stability of the interaction(s)) between the first molecular entity and the second molecular entity. In some embodiments, the first molecular entity and/or the second molecular entity is a protein of interest or a peptide of interest, and tagging comprises generating a fusion (or synthetic conjugation) of the first molecular entity and/or the second molecular entity with the first peptide tag and/or second peptide tag. In some embodiments, the first molecular entity and/or the second molecular entity is a small molecule, and tagging comprises directly or indirectly linking the first molecular entity and/or the second molecular entity with the first peptide tag and/or second peptide tag. In some embodiments, one of the first molecular entity and the second molecular entity is a drug or drug candidate, and the other is a drug target or candidate drug target, and the bioluminescent signal indicates binding of the drug or drug candidate to the other is a drug target or candidate drug target. In some embodiments, combining the tagged first molecular entity and the tagged second molecular entity comprises expressing one or both within a cell and/or adding one or both to a cell. In some embodiments, combining the tagged first molecular entity and the tagged second molecular entity is performed *in vitro*, in a non-cellular sample, etc. In some embodiments, the affinity of a drug or candidate drug for

a drug target or candidate drug target is determined using the systems and methods herein by titrating unlabeled drug target or candidate drug target into the system. In some embodiments, two or more of steps (a)-(f) are performed concurrently. In some embodiments, two or more of steps (a)-(f) are performed separately.

In some embodiments, provided herein are method of performing a competition assay to detect an interaction between a first molecular entity and a second molecular entity, the method comprising: (a) combining: (i) a tracer comprising the first molecular entity tagged with a first peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9, (ii) the second molecular entity tagged with a second peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10, (iii) a coelenterazine or a coelenterazine analog substrate, (iv) a polypeptide component comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, and (v) a sample suspected of containing untagged first molecular entity; wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex and produce a bioluminescent signal in the presence of the coelenterazine or a coelenterazine analog substrate; (b) detecting the bioluminescent signal produced by the bioluminescent complex; and (c) comparing the bioluminescent signal produced in the presence of the sample with a control bioluminescent signal produced in the absence of the sample, wherein a decrease in the bioluminescent signal indicates the presence or amount of untagged first molecular entity in the sample. In some embodiments, the first molecular entity is a small molecule or peptide (e.g., drug or candidate drug). In some embodiments, the second molecular entity is a drug target or candidate drug target (e.g., a protein).

In some embodiments, provided herein are methods of detecting an interaction between a first protein, peptide, or molecular entity and a second protein, peptide, or molecular entity within a cell comprising, the method comprising: (a) expressing within the

cell (or adding to a cell or other system (e.g., non-cellular sample)), a fusion (e.g., genetic fusion, synthetic fusion, chemical conjugation, etc.) comprising the first protein, peptide, or molecular entity and a first peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9; (b) expressing within the cell (or adding to a cell or other system (e.g., non-cellular sample)), a fusion (e.g., genetic fusion, synthetic fusion, chemical conjugation, etc.) comprising the second protein, peptide, or molecular entity and a second peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10; (c) expressing with the cell (or adding to a cell or other system (e.g., non-cellular sample)), a polypeptide component comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon interaction of the first protein, peptide, or molecular entity and the second protein, peptide, or molecular entity; (d) adding a coelenterazine or a coelenterazine analog substrate to the cell; and (e) detecting a luminescent signal produced by the bioluminescent complex, wherein the magnitude of the luminescent signal correlates to the strength of the interaction between the first protein, peptide, or molecular entity and the second protein, peptide, or molecular entity. In some embodiments, two or more of steps (a)-(e) are performed concurrently. In some embodiments, two or more of steps (a)-(e) are performed separately.

In some embodiments, provided herein are methods of detecting co-localization of a first molecular entity and a second molecular entity, the method comprising: (a) tagging the first molecular entity with a first peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9; (b) tagging the second molecular entity with a second peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%,

55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10; (c) combining the tagged first molecular entity and the tagged second molecular entity in the same system; (d) adding a polypeptide component to the system, the polypeptide components comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon co-localization of the first molecular entity and the second molecular entity; (e) adding a coelenterazine or a coelenterazine analog substrate to the system; and (f) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates co-localization of the first molecular entity and the second molecular entity within the system, and/or wherein the magnitude of the luminescent signal correlates to the amount of co-localization within the system of the first molecular entity and the second molecular entity. In some embodiments, the system comprises a cell, tissue, organ, or whole organism. In some embodiments, the first molecular entity and/or the second molecular entity is a protein of interest or a peptide of interest, and tagging comprises generating a fusion (e.g., genetic fusion, synthetic fusion, chemical conjugation, enzymatic conjugation, etc.) of the first molecular entity and/or the second molecular entity with the first peptide tag and/or second peptide tag. In some embodiments, the first molecular entity and/or the second molecular entity is a small molecule and tagging comprises directly or indirectly linking the first molecular entity and/or the second molecular entity with the first peptide tag and/or second peptide tag. In some embodiments, combining the tagged first molecular entity and the tagged second molecular entity is performed *in vitro*, in a non-cellular sample, etc. In some embodiments, combining the tagged first molecular entity and the tagged second molecular entity comprises expressing one or both within the system and/or adding one or both to the system. In some embodiments, two or more of steps (a)-(f) are performed concurrently. In some embodiments, two or more of steps (a)-(f) are performed separately.

In some embodiments, provided herein are methods of detecting co-localization of a first protein, peptide, or molecular entity and a second protein, peptide, or molecular entity

within a cell the method comprising: (a) expressing within the cell a fusion comprising the first protein or peptide entity and a first peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9; (b) expressing within the cell a fusion comprising the second protein or peptide entity and a second peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10; (c) expressing with the cell a polypeptide component comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, and/or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon co-localization of the first protein or peptide entity and the second protein or peptide entity; (d) adding a coelenterazine or a coelenterazine analog substrate to the cell; and (e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates co-localization of the first protein or peptide entity and the second protein or peptide entity within the cell, and/or wherein the magnitude of the luminescent signal correlates to the amount of co-localization within the system of the first protein or peptide entity and the second protein or peptide entity. In some embodiments, two or more of steps (a)-(e) are performed concurrently. In some embodiments, two or more of steps (a)-(e) are performed separately.

In some embodiments, provided herein are kits comprising: (a) a first binding moiety conjugated to a first peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9; and (b) a second binding moiety conjugated to second peptide tag comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10. In some

embodiments, the first and second binding moieties are independently selected from the group consisting of an antibody (e.g., polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the first and second binding moieties are primary binding moieties configured to bind to antigens, epitopes, or sequences on the same target entity. In some embodiments, the first and second binding moieties are secondary binding moieties configured to bind to antigens, epitopes, or sequences on primary binding moieties. In some embodiments, kits further comprise a polypeptide reagent comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, and/or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8. In some embodiments, kits further comprise a coelenterazine or a coelenterazine analog.

In some embodiments, provided herein are methods of detecting a target molecule, wherein the target molecule displays a first antigen, epitope, or sequence and a distinct second antigen, epitope, or sequence the method comprising: (a) contacting a sample containing the target molecule with (i) a first primary binding moiety that recognizes the first antigen, epitope, or sequence and (ii) a second primary binding moiety that recognizes the second antigen, epitope, or sequence and allowing the first and second primary binding moieties to bind to the first and second antigens, epitopes, or sequences; (b) contacting the sample with (i) a first secondary binding moiety conjugated or fused to a first peptide tag and (ii) a second secondary binding moiety conjugated or fused to second peptide tag, wherein the first secondary binding moiety recognizes the first primary binding moiety and the second secondary binding moiety recognizes the second primary binding moiety, wherein the first or second peptide tag comprises an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 (and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9), and wherein the other of the first or second peptide tag comprises an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%,

85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 (and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10) and allowing the first and second secondary binding moieties to bind to the first and second primary binding moieties; (c) contacting the sample with comprising an polypeptide component having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, and/or SEQ ID NO: 302 (and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8), wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon interaction; (d) contacting the sample with a coelenterazine or a coelenterazine analog substrate; and (e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates the presence of the target molecule, and/or wherein the magnitude of the luminescent signal correlates to the amount of target molecule within the sample. In some embodiments, the binding moieties are independently selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the target molecule is a protein, peptide, nucleic acid, chemical, or drug. In some embodiments, the sample is *in vitro* or *in vivo*.

In some embodiments, provided herein are methods of detecting a target molecule, wherein the target molecule displays a first antigen, epitope, or sequence and a distinct second antigen, epitope, or sequence, the method comprising: (a) contacting the sample with (i) a first binding moiety conjugated or fused to a first peptide tag and (ii) a second binding moiety conjugated or fused to second peptide tag, wherein the first binding moiety recognizes the first antigen, epitope, or sequence and the second binding moiety recognizes the second antigen, epitope, or sequence, wherein the first or second peptide tag comprises an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 (and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9), and wherein the other of the first

or second peptide tag comprises an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 (and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10) and allowing the first and second binding moieties to bind to the first and second antigens, epitopes, or sequences; (c) contacting the sample with a polypeptide component having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, and/or SEQ ID NO: 302 (and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8), wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon interaction; (d) contacting the sample with a coelenterazine or a coelenterazine analog substrate; and (e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates the presence of the target molecule, and/or wherein the magnitude of the luminescent signal correlates to the amount of target molecule within the sample. In some embodiments, the binding moieties are independently selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins. In some embodiments, the target molecule is a protein, peptide, nucleic acid, chemical, or drug. In some embodiments, the sample is *in vitro* or *in vivo*.

In some embodiments, provided herein are methods comprising: (a) combining: (i) a peptide component comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 35 and less than 100% sequence identity with SEQ ID NO: 205 and SEQ ID NO: 206, (ii) a polypeptide component comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, SEQ ID NO: 21, and/or SEQ ID NO: 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, and (iii) a coelenterazine or a coelenterazine analog substrate, wherein the peptide component and the

polypeptide component are configured to produce a bioluminescent complex upon interaction; and (b) detecting luminescence, wherein a greater level of luminescence compared to a level of luminescence produced by the polypeptide component and a coelenterazine or a coelenterazine analog alone indicates formation of a bioluminescent complex of the polypeptide component with the peptide. In some embodiments, the peptide is a fusion (e.g., genetic, synthetic, chemical conjugate, enzymatic conjugate, etc.) with a first interaction element, and the polypeptide component is a fusion (e.g., genetic, synthetic, chemical conjugate, enzymatic conjugate, etc.) with a second interaction element, wherein the peptide and the polypeptide component form a bioluminescent complex upon interaction of the interaction elements, but do not form a bioluminescent complex in the absence of an interaction between the interaction elements. In some embodiments, the peptide and the polypeptide component form a bioluminescent complex in the absence of facilitation (e.g., by interaction elements. In some embodiments, the peptide is a fusion or conjugate (e.g., genetic, synthetic, chemical conjugate, enzymatic conjugate, etc.) with a protein, peptide, or molecule of interest (e.g., not an interaction element) and/or the polypeptide component is a fusion or conjugate (e.g., genetic, synthetic, chemical conjugate, enzymatic conjugate, etc.) with a protein, peptide, or molecule of interest (e.g., not an interaction element). In some embodiments, the peptide component and the polypeptide component form a bioluminescent complex upon co-localization (e.g., in a sample, in a cell, in a tissue, in a subject, etc.) without facilitation by interaction elements. In some embodiments, the peptide component and the polypeptide component form a bioluminescent complex upon facilitation by interaction elements but not without facilitation.

In some embodiments, provided herein are compositions comrpising a monomeric polypeptide comprising an amino acid sequence with greater than 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, or ranges therebetween), but less than 100% sequence identity with SEQ ID NO: 788 or 789 and less than 100% sequence identity with SEQ ID NO: 1 or 3; wherein the polypeptide exhibits luminescence in the presence of a coelenterazine or a coelenterazine analog substrate. In some embodiments, the peptide comprising an amino acid sequence with greater than 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, or ranges therebetween) sequence identity with one or more of SEQ ID NOS: 780, 782, 784, 786, 802, 804, 806, 808, 813, 815, or 829. In some embodiments, the polypeptide further comrpises an additional amino acid

sequence. In some embodiments, fusion protiens of the polypeptides described herein are provided.

In some embodiments, provided herein are nucleic acisd comprising a sequence coding for a comprising an amino acid sequence with greater than 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, or ranges therebetween), but less than 100% sequence identity with SEQ ID NO: 788 or 789 and less than 100% sequence identity with SEQ ID NO: 1 or 3; wherein the polypeptide exhibits luminescence in the presence of a coelenterazine or a coelenterazine analog substrate. In some embodiments, a nucleic acid codes for an amino acid sequence with greater than 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, or ranges therebetween) sequence identity with one or more of SEQ ID NOS: 780, 782, 784, 786, 802, 804, 806, 808, 813, 815, or 829. In some embodiments, nucleic acids coding for fusion protiens of the polypeptides described herein are provided.

In some embodiments, provided herein are methods of detecting bioluminescence comprising contacting a monomeric bioluminescent polypeptide described herein with a coelenterazine or a coelenterazine analog substrate and detecting luminescence.

In some embodiments, provided herein are polypeptides comrpising an amino acid sequence with greater than 40% sequence identity with SEQ ID NO: 15 linked to the N-terminal end of an amino acid sequence with greater than 40% sequence identity with one of SEQ ID NOS: 17, 21, or 302; wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the polypeptide contacts a peptide or polytpeptide comprising an amino acid sequce consisting of SEQ ID NO: 23 when compared to a bioluminescent signal produced by the polypeptide and the coelenterazine substrate alone. In some embodiments, nucleic acids comprising a sequence coding for the polypeptides described herein are provided. In some embodiments, fusions of the polypeptides described herein and additional amino acid sequences are provided.

In some embodiments, provided herein are methods comprising (a) contacting a polypeptide comrpising an amino acid sequence with greater than 40% sequence identity with SEQ ID NO: 15 linked to the N-terminal end of an amino acid sequence with greater than 40% sequence identity with one of SEQ ID NOS: 17, 21, or 302 with a coelenterazine or a coelenterazine analog substrate and a peptide or polypeptide comprising an amino acid

sequence with 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, or ranges therebetween) sequecne identity with SEQ ID NO: 23; and (b) detecting luminescence.

In some embodiments, provided herein are systems comprising: (a) a sensor polypeptide comprising a first amino acid sequence having at least 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, or ranges therebetween) sequence identity with SEQ ID NO: 11 linked to an amino acid sequence that localizes in a specific cellular location; and (b) a protein of interest linked to a peptide tag comprising an amino acid sequence having at least 40% (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, or ranges therebetween) sequence identity to SEQ ID NO: 23; wherein a bioluminescent complex is formed between the polypeptide and the peptide tag upon colocalization within the specific cellular location. In some embodiments, the specific cellular location is selected from plasma membrane, nucleus, mitochondria, and the endoplasmic reticulum. In some embodiments, provided herein are methods comprising (a) expressing a system described herein in a cell; (b) contacting the cell with a coelenterazine or a coelenterazine analog substrate; and (c) detecting luminescnce, wherein an increase in luminescene indicates formation of the bioluminescent complex and co-localization of the polypeptide and peptide tag. In some embodiments, the method further comprises a step of inducing translocation protein of interest linked to a peptide tag to the specific cellular location.

BRIEF DESCRIPTION OF THE DRAWINGS

The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawings will be provided by the Office upon request and payment of the necessary fee.

Figure 1. Graph demonstrating that a polypeptide lacking β9 and β10 portions (LgTrip 2098; SEQ ID NO: 17) exhibits reduced background luminescence compared to LgBiT (SEQ ID NO: 11) and is activated by complementation with a peptide corresponding to β9 and β10.

Figure 2. Graph demonstrating activation of LgTrip 2098 (SEQ ID NO: 17) by separate peptide corresponding to β9 and β10, respectively.

Figure 3. Graph depicting the relative stability of exemplary LgTrip 2098 mutants.

Figure 4. Graph depicting the relative luminescent activity of amino acid site saturation at position 42 of LgTrip 2098 mutants.

Figure 5. Graph depicting the relative stability of amino acid changes at position 42 of LgTrip 2098 mutants.

Figure 6A-C. Graph depicting the relative luminescent activity of amino acid changes at (A) position 4, (B) position 30, and (C) position 106 of LgTrip 2098 mutants.

Figure 7. Graph depicting the relative luminescent activity of amino acid changes at (A) position 101, (B) position 117, (C) position 127, (D) position 120, and (E) position 126 of LgTrip 3092 mutants.

Figure 8. Graph depicting the relative stability of LgTrip 2098 (WT)(SEQ ID NO: 31), LgTrip 3092 (SEQ ID NO: 19), and LgBiT (SEQ ID NO: 11) at 37ºC.

Figure 9. Graph depicting the relative stability of LgTrip variants at (A) 42ºC and (B) 60ºC.

Figure 10. Graphs depicting (A) titration of various LgTrip variants with SmTrip9 pep286 (SEQ ID NO: 37), (B) titration of various LgTrip variants with SmTrip10 pep86 (SEQ ID NO: 25), and (C) the affinity of various LgTrip variants for SmTrip9 pep286 and SmTrip10 pep86.

Figure 11A-B. Graphs depicting the (A) stability (half-life) and (B) relative stability of various LgTrip variants at 60ºC.

Figure 12A-B. Graphs depicting the kinetic profiles of LgTrip variants in the presence of SmTrip9 pep286 (SEQ ID NO: 37) and SmTrip10 pep86 (HiBiT; SEQ ID NO: 25) compared to NanoLuc (SEQ ID NO: 3) and LgBiT (SEQ ID NO: 11) and SmTrip10 pep86 (HiBiT; SEQ ID NO: 25) (A) assayed in TBS+0.01% BSA and (B) assayed with NanoGlo® assay buffer.

Figure 13A-B. Graphs depicting facilitated complementation of various tripartite systems via rapamycin-induced formation of a FRB/FKBP complex: (A) SmTrip10 pep86 (SEQ ID NO:25), SmTrip9 pep245 (SEQ ID NO: 23), and LgTrip 2098 (SEQ ID NO: 31); and (B) SmBiT (SEQ ID NO: 13), SmTrip9 pep245 (SEQ ID NO: 23), and LgTrip 2098 (SEQ ID NO: 31).

Figure 14A-B. Graphs comparing the stability at 37°C of LgBiT (SEQ ID NO: 11) and LgTrip 2098 (WT) (SEQ ID NO: 31 in (A) TBS+0.01% BSA and (B) Passive Lysis Buffer (PLB).

Figure 15A-B. Graphs comparing stability of LgBiT (SEQ ID NO: 11), LgTrip 3546 (SEQ ID NO: 51), and NanoLuc (SEQ ID NO: 31) at 60ºC; (A) time course and (B) half-life.

Figure 16A-B. Graphs comparing LgBiT (SEQ ID NO: 11), NanoLuc (SEQ ID NO: 3), and LgTrip 3546 (SEQ ID NO: 51), and LgTrip 2098 (WT) (SEQ ID NO: 31) (A) in the presence NaCl and (B) after 26 hour exposure to NaCl.

Figure 17A-B. Graphs comparing LgBiT (SEQ ID NO: 11), NanoLuc (SEQ ID NO: 3), and LgTrip 3546 (SEQ ID NO: 51) and LgTrip 2098 (WT) (SEQ ID NO: 31) variants (A) in the presence urea and (B) after 26 hour exposure to urea.

Figure 18A-B. Graphs comparing LgBiT (SEQ ID NO: 11), NanoLuc (SEQ ID NO: 3), and LgTrip 3546 (SEQ ID NO: 51) and LgTrip 2098 (WT) (SEQ ID NO: 31) variants (A) at varying pH and (B) after 26 hour exposure to varying pH.

Figure 19. Graph comparing the autoluminescence of LgBiT (SEQ ID NO: 11) and LgTrip 3546 (SEQ ID NO: 51).

Figure 20A-B. Graph comparing the luminescence of LgBiT (SEQ ID NO: 11) + SmTrip 10 pep86 (HiBiT; SEQ ID NO: 25), LgBiT (SEQ ID NO: 11) + pep263 (SEQ ID NO: 35) (β9/β10 dipeptide), and LgTrip 3546 (SEQ ID NO: 51) + pep263 (β9/β10 dipeptide) (SEQ ID NO: 35): (A) RLU and (B) signal/background (S/B).

Figure 21A-C. Facilitated complementation of LgTrip 2098 (SEQ ID NO: 31) and LgTrip 3546 (SEQ ID NO: 51), respectively with SmTrip10 pep86 (SEQ ID NO: 25) and SmTrip9 pep245 (SEQ ID NO: 23): (A) schematic of assay system, (B) RLU, and (C) signal/background (S/B).

Figure 22. Graph and table comparing the affinities of various SmTrip10 sequences for LgTrip 3546 (SEQ ID NO: 51) and SmTrip9 pep286 (SEQ ID NO: 37).

Figure 23. Graph and table comparing the activation of LgTrip 2098 (SEQ ID NO: 31) and LgTrip 3546 (SEQ ID NO: 51) by standard-orientation (pep263) (SEQ ID NO: 35) and inverse-orientation (pep326) (SEQ ID NO: 179) dipeptides.

Figure 24. Graph and table depicting activation of complement polypeptides by dipeptides comprising the HiBiT or SmBiT sequence. Dipeptide with HiBiT sequence pep263 (SEQ ID NO: 35) or Dipeptide with SmBiT sequence pep274 (SEQ ID NO: 147)

Figure 25A-B. (A) Graph depicting luminescence resulting from complementation of various combinations of polypeptide components (with additions or deletions relative to LgTrip 3546) with SmTrip9 pep286 (SEQ ID NO: 37) and various β10-like peptides (SmTrip10 peptides); (B) Graph depicting luminescence resulting from complementation of

LgTrip 3546 (SEQ ID NO: 51) and SmTrip9 pep286 (SEQ ID NO: 37) with various β10-like peptides (SmTrip10 peptides).

Figure 26A-C. (A-C) Graphs depicting luminescence produced by polypeptide/peptide combinations having overlap (relative to a base luciferase sequence) between the polypeptide component and a peptide corresponding to the β9-strand or between the β9 and β10-like peptides.

Figure 27A-B. Figures and tables depicting luminescence resulting from (A) the titration of various β9-like peptides (SmTrip9 peptides) in the present of constant LgTrip 3546 (SEQ ID NO: 51) and SmTrip10 pep86 (SEQ ID NO: 25) concentrations, and (B) the titration of SmTrip10 pep86 (SEQ ID NO: 25) in the presence of constant concentrations of LgTrip 3546 (SEQ ID NO: 51) and various β9-like peptides (SmTrip9 peptides).

Figure 28. Figure and table depicting luminescence resulting from the titration of various β10-like peptides (SmTrip 10 peptides) in the present of constant concentrations of LgTrip 3546 (SEQ ID NO: 51) and SmTrip9 pep286 (SEQ ID NO: 37).

Figure 29A-B. Figures and tables depicting titration of β9-like peptides (A) SmTrip9 pep286 (SEQ. ID 37) and (B) SmTrip9 pep287 (SEQ ID NO: 148) in the presence of constant concentration of various β10-like peptides (SmTrip10 peptides) and LgTrip 3546 (SEQ ID NO: 51) polypeptide component.

Figure 30. Graph depicting the effect of construct orientation (β9-FKBP, FKBP-β9, β10-FKBP, FKBP-β10, β9-FRB, FRB-β9, β10-FRB, or FRB-β10) on facilitated complementation in HEK293 cells.

Figure 31. Graph depicting the effect of construct orientation (β9-FKBP, FKBP-β9, β10-FKBP, FKBP-β10, β9-FRB, FRB-β9, β10-FRB, or FRB-β10) on facilitated complementation in E. coli cells.

Figure 32. Graph depicting calculated Kd values for various β10-like peptides with LgTrip 3546 (SEQ ID NO: 51) and SmTrip9 pep286 (SEQ ID NO: 37).

Figure 33. Graph and table depicting luminescence from combinations of components having varied split sites between the polypeptide component LgTrip 3546 (SEQ ID NO: 51) and the β9-like peptide.

Figure 34. Graph depicting luminescence from combinations of components with sequence gaps and/or overlaps between various LgTrip polypeptide components and SmTrip9 pep286 (SEQ ID NO: 37).

Figure 35. Graph depicting luminescence from NanoTrip component combinations with gaps and/or overlaps in sequence between the β9-like peptides (SmTrip9 peptides) and polypeptide component LgTrip 3546 (SEQ ID NO: 51) in the presence of SmTrip10 pep86 (HiBiT; SEQ ID NO: 25).

Figure 36. Table depicting a biochemical analysis of β9-like peptide (SmTrip9 peptides) length influence on β9-like peptide affinity and maximum light output with LgTrip 3546 (SEQ ID NO: 51) and SmTrip10 pep86.

Figure 37. Table depicting a biochemical analysis of β9-like peptide (SmTrip9 peptides) length influence on HiBiT affinity and maximum light output with LgTrip 3546 (SEQ ID NO: 51) and SmTrip10 pep86 (SEQ ID NO: 25).

Figure 38. Table depicting Kd and Bmax of β9-like SmTrip9 pep286 (SEQ ID NO: 37) point mutants with LgTrip 3546 (SEQ ID NO: 51) and SmTrip10 pep86 (SEQ ID NO: 25).

Figure 39. Table depicting the effect of various solubility tags on β9-like peptide affinity with LgTrip 3546 (SEQ ID NO: 51) and SmTrip9 pep86 (SEQ ID NO: 25).

Figure 40. Table depicting the effect of various C-terminal extension sequences on β9-like or β10-like peptide affinity and maximum light output. β9-like peptide titrations (pep286 (SEQ ID NO: 37), pep292 (SEQ ID NO: 153), pep297 (SEQ ID NO: 157), pep302 (SEQ ID NO: 161)) and β10-like peptide SmTrip10 pep86 (HiBiT; SEQ ID NO: 25) are depicted.

Figure 41. Graph depicting the effect of FRB-β10 construct linker length (15, 10, or 5 Gly/Ser residues), linker composition (with or without Ala-Ile), hexahistidine tag inclusion, and β10 composition (SmTrip10 pep86 (SEQ ID NO: 25) or SmTrip10 pep289 (SEQ ID NO: 150)) on facilitated complementation in E. coli lysates with LgTrip 3546 (SEQ ID NO: 51).

Figure 42. Graph depicting the effect of FRB-β10 construct linker length (15, 10, or 5 Gly/Ser residues), linker composition (with or without Ala-Ile), hexahistidine tag inclusion, and β10 composition SmTrip10 pep86 (SEQ ID NO: 25) or SmTrip10 pep289 (SEQ ID NO: 150) on facilitated complementation in HEK lysates with LgTrip 3546 (SEQ ID NO: 51).

Figure 43. Graph depicting the effect of β9 sequence truncations and extensions and construct orientation (β9-FKBP or FKBP-β9) on facilitated complementation with FRB-SmTrip10 pep86 (β10) (SEQ ID NO: 25) in E. coli lysates with LgTrip 3546 (SEQ ID NO: 51).

Figure 44. Graph depicting the effect of β9 sequence truncations and extensions and construct orientation (β9-FKBP or FKBP-β9) on facilitated complementation with FRB-SmTrip10 pep289 (β10) (SEQ ID NO: 150) in E. coli lysates with LgTrip 3546 (SEQ ID NO: 51).

Figure 45. Graph depicting the effect of β9 sequence truncations, extensions, and construct orientation (β9-FKBP or FKBP-β9) on facilitated complementation with FRB-SmTrip10 pep86 (β10) (SEQ ID NO: 25) in E. coli lysates with LgTrip 3546 (SEQ ID NO: 51).

Figure 46. Graph depicting the effect of β9 sequence truncations, extensions, and construct orientation (β9-FKBP or FKBP-β9) on facilitated complementation with FRB-SmTrip10 pep289 (β10) (SEQ ID NO 150) in E. coli lysates with LgTrip 3546 (SEQ ID NO: 51).

Figure 47. Graph depicting the effect of β9 sequence truncations, extensions, and construct orientation (β9-FKBP or FKBP-β9) on fold induction (facilitated complementation/spontaneous complementation) with FRB-β10 (SmTrip10 pep86 (SEQ ID NO: 25) or SmTrip10 pep289 (SEQ ID NO: 150)) in E. coli lysates (Summary of Figures 45 and 46) with LgTrip 3546 (SEQ ID NO: 51).

Figure 48. Graph depicting the effect of β9 sequence truncations and extensions and construct orientation (β9-FKBP or FKBP-β9) on facilitated complementation with FRB-SmTrip10 pep86 (SEQ ID NO: 25) in HEK293 lysates with LgTrip 3546 (SEQ ID NO: 51).

Figure 49. Graph depicting the effect of β9 sequence truncations and extensions and construct orientation (β9-FKBP or FKBP-β9) on facilitated complementation with FRB-SmTrip10 pep289 (SEQ ID NO: 150) in HEK293 lysates. (LgTrip 3546 (SEQ ID NO: 51).

Figure 50. Graph depicting the effect of β9 sequence truncations, extensions, and construct orientation (β9-FKBP or FKBP-β9) on fold induction (facilitated complementation/spontaneous complementation) with FRB-β10 (SmTrip10 pep86 (SEQ ID NO: 25) or SmTrip10pep289 (SEQ ID NO: 150)) in HEK293 lysates (Summary of Figures 48 and 49). (LgTrip 3546 (SEQ ID NO: 51)).

Figure 51A-D. Schematic illustrations depicting exemplary protein-protein interaction assays or analyte detection assays using binding moieties tagged with peptides.

Figure 51E-H. Schematic illustrations depicting exemplary immunoassays using components and reagents described herein: (A) direct immunoassay, (B) indirect

immunoassay, (C) competition direct immunoassay, and (D) competition indirect immunoassay.

Figure 52. Schematic illustration of an exemplary multiplexed tripartite lateral flow assay. Such an assay finds use, for example, in the detection of pathogens.

Figure 53. Schematic illustration of an exemplary multiplexed tripartite lateral flow assay. Such an assay finds use, for example, in the detection of antiviral antibodies.

Figure 54. Schematic illustration of an exemplary antibody detection assay.

Figure 55. Schematic illustration of an exemplary bead-based assay.

Figure 56. Schematic illustration of an exemplary nucleic acid detection assay.

Figure 57. Graph depicting FRB-FKBP facilitated complementation in E. coli lysates with AI (Ala-Ile) dipeptide absent from linker in constructs denoted by **. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 58. Graph depicting FRB-FKBP facilitated complementation in HEK293 lysates with AI sequence dipeptide absent from linker in constructs denoted by **. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 59. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FRB-SmTrip10 pep86 and C-terminally extended FKBP-SmTrip9 peptides in E. coli lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 60. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FRB-SmTrip10 pep289 and C-terminally extended FKBP-SmTrip9 peptides in E. coli lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 61. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FRB-SmTrip10 pep86 and C-terminally extended FKBP-SmTrip9 peptides in HEK293 lysate. (LgTrip 3546 (SEQ ID NO: 51))

Figure 62. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FRB-SmTrip10 pep86 and SmTrip9 peptide sequence trunctions and extensions in FKBP fusions in E. coli lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 63. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FRB-SmTrip10 pep289 and SmTrip9 peptide sequence trunctions and extensions in FKBP fusions in E. coli lysate. (LgTrip 3546 (SEQ ID NO: 51)) .

Figure 64. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FRB-SmTrip10 pep86 and SmTrip9 peptide sequence trunctions and extensions in FKBP fusions in HEK293 lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 65. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FRB-SmTrip10 pep86 and SmTrip9 peptide sequence trunctions and extensions in FKBP fusions in HEK293 lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 66. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FKBP-SmTrip9 solubility variants and FRB-SmTrip10 pep86 in E. coli lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 67. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FKBP-SmTrip9 solubility variants and FRB-SmTrip10 pep289 in E. coli lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 68. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FKBP-SmTrip9 solubility variants and FRB-SmTrip10 pep86 in HEK293 lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 69. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FKBP-SmTrip9 solubility variants and FRB-SmTrip10 pep289 in HEK293 lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 70. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FKBP-SmTrip9 solubility variants and FRB-SmTrip pep86 in E. coli lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 71. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FKBP-SmTrip9 solubility variants and FRB-SmTrip pep289 in E. coli lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 72. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FKBP-SmTrip9 solubility variants and FRB-SmTrip pep86 in HEK293 lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 73. Graph depicting FRB-FKBP facilitation of luminescent complex formation with FKBP-SmTrip9 solubility variants and FRB-SmTrip pep289 in HEK293 lysate. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 74. Table listing affinity and Bmax of synthetic SmTrip9 solubility variants with C-terminal extensions. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 75. Table listing affinity and Bmax of synthetic SmTrip9 solubility variants with C-terminal extensions. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 76. Table listing Kd and Bmax of synthetic SmTrip9 variants with differentially blocked termini. (LgTrip 3546 (SEQ ID NO: 51)).

Figure 77. Table listing the solubility of synthetic SmTrip9 peptides.

Figure 78A-B. (A) Graph depicting the affinity of SmTrip9 pep286 (SEQ ID NO: 37) for SmTrip10 pep86 (HiBiT)/LgTrip fusions (SEQ ID NO: 210 and 212). (B) Graph depicting the affinity of SmTrip9 pep759 (SEQ ID NO: 496) for various SmTrip10 pep86 (HiBiT)/LgTrip fusions.

Figure 79. Graphs depicting bioluminescence following an 18 hour exposure to increasing detergent concentrations. NanoLuc (SEQ ID NO: 3), LgBiT (SEQ ID NO: 11), (LgTrip 3546 (SEQ ID NO: 51)).

Figure 80. Graphs depicting enzyme activity in the presence of increasing detergent concentrations. NanoLuc (SEQ ID NO: 3), LgBiT (SEQ ID NO: 11), LgTrip 3546 (SEQ ID NO: 51).

Figure 81. Graph demonstrating the reversibility of FRB-FKBP facilitated bioluminescent complex formation with LgBiT (SEQ ID NO: 11) and LgTrip 3546 (SEQ ID NO: 51).

Figure 82. Table listing results of titration of various SmTrip10 peptides in the presence of constant SmTrip9 pep286 (SEQ ID NO: 37) and LgTrip 3546 (SEQ ID NO: 51) .

Figure 83. Table listing results of titration of various SmTrip10 peptides in the presence of constant SmTrip9 pep286 (SEQ ID NO: 37) and LgTrip 3546 (SEQ ID NO: 51) titration

Figure 84. Graph depicting bioluminescence from Antares-type fusions (LgTrip 3546) with SmTrip9 pep263 (SEQ ID NO: 35) and SmTrip10 pep86 (SEQ ID NO: 25) or SmTrip10 pep86 + SmTrip9 pep286 (SEQ ID NO: 37).

Figures 85A-B. Graphs depicting emission spectra from Antares-type fusions (LgTrip 3546) (SEQ ID NO: 51) with SmTrip9 pep263 (SEQ ID NO: 35) and SmTrip pep86 (HiBiT; SEQ ID NO: 25) or SmTrip10 pep86 (HiBiT; SEQ ID NO: 25) + SmTrip9 pep286 (SEQ ID NO: 37).

Figure 86. Graphs depicting titration of LgBiT (SEQ ID NO: 11) and LgTrip 3546 (SEQ ID NO: 51) with "dark" dipeptide 272 (SEQ ID NO: 146) in the presence of dipeptide pep263 (SEQ ID NO: 35).

Figure 87. Graphs comparing the inhibition of dark dipeptides 272 (SEQ ID NO: 146) and 273 (SEQ ID NO: 298) with LgTrip 3546 (SEQ ID NO: 51) and LgBiT (SEQ ID NO: 11).

Figure 88. Graph depicting inhibition of LgBiT (SEQ ID NO: 11) and LgTrip 3546 (SEQ ID NO: 51) with dark BiT167 (SEQ ID NO: 300).

Figure 89. Graph depicting FRB-FKBP facilitation of luminescent complex formation in E. coli lysate with FKBP-SmTrip9 pep434 (SEQ ID NO: 230) variants' complementation with LgTrip 3546 (SEQ ID NO: 51) and FRB-HiBiT (SEQ ID NO: 25).

Figure 90. Graph depicting FRB-FKBP facilitation of luminescent complex formation in E. coli lysate with SmTrip9 pep434 (SEQ ID NO: 230) variants' complementation with LgTrip 3546 (SEQ ID NO: 51) and FRB-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO: 150).

Figure 91. (A) Graph depicting FRB-FKBP facilitation of luminescent complex. formation in HEK293 lysate with SmTrip9 pep435 (SEQ ID NO: 231) and SmTrip9 pep434 (SEQ ID NO: 230) variants' complementation with LgTrip 3546 (SEQ ID NO: 51) and FRB-SmTrip10 pep86 (HiBiT; SEQ ID NO: 25). (B) Graph depicting FRB-FKBP facilitation of luminescent complex formation in HEK293 lysate with SmTrip9 pep435 (SEQ ID NO: 231) and SmTrip9 pep434 (SEQ ID NO: 230) variants' complementation with LgTrip 3546 (SEQ ID NO: 51) and FRB-SmTrip10 pep289 (SEQ ID NO: 150).

Figure 92. Table depicting the results of a FRB-FKBP assay screen with SmTrip9s 823 and 840.

Figure 93. Table listing Kd and Bmax of synthetic SmTrip9 pep435 (SEQ ID NO: 231) and SmTrip9 pep434 (SEQ ID NO: 230) variants with LgTrip 3546 (SEQ ID NO: 51).

Figure 94. Graph demonstration of wt LgTrip 2098 (SEQ ID NO: 31) and LgTrip 3546 (SEQ ID NO: 51) with pep263 (SEQ ID NO: 35) or pep331 (SEQ ID NO: 301) as bioluminescence reagents for detecting endogenously tagged (e.g., by CRISPR/Cas9) GAPDH.

Figure 95. Exemplary SmTrip10 chemical conjugates. (A) Example of SmTrip10 with N-terminal cysteine modification for disulfide bond formation on solvent exposed or protected cysteine targets on proteins/peptides/DNA and RNA oligonucleotides/small

molecules or proteins/peptides/DNA and RNA oligonucleotides/small molecules that have been prepared with maleimide for reaction with a thiol such as cysteine or N-hydroxysuccinimide esters (NHS-ester) for reaction with an amine such as lysine. (B) Examplary SmTrip10 with N-terminal azido-lysine modification for copper catalyzed or copper free 1,3-dipolar cycloaddition reactions ("Click") with unstrained or strained alkyne targets separately installed on proteins/peptides/ DNA and RNA oligonucleotides/small molecules. (C) Examplary SmTrip10 with N-terminal N-hydroxylsuccinimide ester (NHS-ester) for general conjugation to nucleophiles (e.g., lysines, other primary amines) on proteins/peptides/ DNA and RNA oligonucleotides/small molecules. Nucleophiles can be present on unmodified proteins/oligos/small molecules or may be chemically added for the purposes of this conjugation. (D) Examplary SmTrip10 with an N-terminal propargyl glycine modification for copper catalyzed or copper free 1,3-dipolar cycloaddition reactions ("Click") with azide, diazo, or tetrazine targets separately introduced chemically or biologically on proteins/peptides/DNA and RNA oligonucleotides/small molecules. (E) Examplary SmTrip10 with a N-terminal propargyl glycine modification for copper catalyzed or copper free 1,3-dipolar cycloaddition reactions ("Click") and a C-terminal fluorophore (e.g., BODIPY dye).

Figure 96. Exemplary SmTrip9 pep286 chemical conjugates. (A) Example of SmTrip9-286 with C-terminal azido-lysine modification for copper catalyzed or copper free 1,3-dipolar cycloaddition reactions ("Click") with unstrained or strained alkyne targets separately introduced chemically or biologically on proteins/peptides/DNA and RNA oligonucleotides/small molecules. (B) Example of SmTrip9 pep286 with C-terminal propargyl glycine modification for copper catalyzed or copper free 1,3-dipolar cycloaddition reactions ("Click") with azide, diazo, tetrazine targets separately introduced chemically or biologically on proteins/peptides/DNA and RNA oligonucleotides/small molecules. (C) Example of SmTrip9 pep286 with C-terminal cysteine modification for disulfide bond formation on solvent exposed or protected cysteine targets on proteins/peptides/DNA and RNA oligonucleotides/small molecules or proteins/peptides/DNA and RNA oligonucleotides/small molecules that have been prepared with maleimide handles or N-hydroxysuccinimide esters. (D) Example of SmTrip9 pep286 with C-terminal cysteine modification and a N-terminal BODIPY dye. The dye is not limited to BODIPY and could be any fluorophore, BRET partner, or FRET dye/quencher partner. Dyes can be incorporated

with any other combination of conjugation handles prepared on the C-terminus. (E) Example of SmTrip9 pep286 with C-terminal N-hydroxysuccinimide esters (NHS-ester) for general conjugation to nucleophilic targets (e.g., lysines) on proteins/peptides/DNA and RNA oligonucleotides/small molecules or proteins/peptides/DNA and RNA oligonucleotides/small molecules. (F) Example of SmTrip9 pep286 with C-terminal N-hydroxysuccinimide ester (NHS-ester) for general conjugation to nucleophilic targets (i.e. lysines) on proteins/peptides/DNA and RNA oligonucleotides/small molecules or proteins/peptides/DNA and RNA oligonucleotides/small molecules.

Figure 97. Exemplary SmTrip9 pep521 chemical conjugates. (A) Example of SmTrip9 pep521 with C-terminal azido-lysine modification and a N-terminal BODIPY dye. The dye is not limited to BODIPY and could be any fluorophore, BRET partner, or FRET dye/quencher partner. Dyes can be incorporated with any other combination of conjugation handles prepared on the C-terminus. (B) Example of SmTrip9 pep521with C-terminal azido-lysine modification for copper catalyzed or copper free 1,3-dipolar cycloaddition reactions ("Click") with unstrained or strained alkyne targets separately introducedchemically or biologically on proteins/peptides/ DNA and RNA oligonucleotides/small molecules. (C) Example of SmTrip9 pep521 with C-terminal propargyl glycine modification for copper catalyzed or copper free 1,3-dipolar cycloaddition reactions ("Click") with azide, diazo, tetrazine targets separately introduced chemically or biologicallyon proteins/peptides/ DNA and RNA oligonucleotides/small molecules. (D) Example of SmTrip9 pep521 with C-terminal cysteine modification for disulfide bond formation on solvent exposed or protected cysteine targets on proteins/peptides/DNA and RNA oligonucleotides/small molecules or proteins/peptides/DNA and RNA oligonucleotides/small molecules that have been prepared with maleimide handles or an NHS-ester. (E) Example of SmTrip9 pep521 with C-terminal N-hydroxysuccinimide ester (NHS-ester) for general conjugation to nucleophilic targets (i.e. lysines) on proteins/peptides/DNA and RNA oligonucleotides/small molecules. (F) Example of SmTrip9 pep521 with C-terminal N-hydroxysuccinimide ester (NHS-ester) for general conjugation to nucleophilic targets (i.e. lysines) on proteins/peptides/ DNA and RNA oligonucleotides/small molecules.

Figure 98. Exemplary SmTrip9 pep524 chemical conjugates. (A) Example of SmTrip9 pep524 with C-terminal azido-lysine modification and a N-terminal BODIPY dye. The dye is not limited to BODIPY and could be any fluorophore, BRET partner, or FRET

dye/quencher partner. Dyes can be incorporated with any other combination of conjugation handles on the C-terminus. (B) Example of SmTrip9 pep524 with C-terminal azido-lysine modification for copper catalyzed or copper free 1,3-dipolar cycloaddition reactions ("Click") with unstrained or strained alkyne targets separately introduced chemically or biologically on proteins/peptides/DNA and RNA oligonucleotides/small molecules. (C) Example of SmTrip9 pep524 with C-terminal propargyl glycine modification for copper catalyzed or copper free 1,3-dipolar cycloaddition reactions ("Click") with azide, diazo, tetrazine targets separately introduced chemically or biologically on proteins/peptides/DNA and RNA oligonucleotides/small molecules. (D) Example of SmTrip9 pep524 with C-terminal cysteine modification for disulfide bond formation on solvent exposed or protected cysteine targets on proteins/peptides/ DNA and RNA oligonucleotides/small molecules or proteins/peptides/DNA and RNA oligonucleotides/small molecules that have been prepared with maleimide handles or an NHS-ester. (E) Example of SmTrip9 pep524 with C-terminal N-hydroxysuccinimide ester (NHS-ester) for general conjugation to nucleophilic targets (i.e. lysines) on proteins/peptides/ DNA and RNA oligonucleotides/small molecules. (F) Example of SmTrip9 pep524 with C-terminal N-hydroxysuccinimide ester (NHS-ester) for general conjugation to nucleophilic targets (i.e. lysines) on proteins/peptides/DNA and RNA oligonucleotides/small molecules.

Figure 99A-B. Exemplary peptide-oligomer probes. Peptides displaying reactive azido groups are conjugated to oligonucleotides displaying reactive alkyne groups to form exemplary peptide-oligomer probes.(A) Peptide oligomer conjugate of SmTrip9 pep286 (w/azido group) conjugated to a DNA oligomer containing 5'-terminal alkyne functionality via a copper "click" 1,3-cycloaddition. (B) Peptide oligomer conjugate of SmTrip10 pep86 (HiBiT) (w/azido group) conjugated to a DNA oligomer containing 3'-terminal alkyne functionality via a copper "click" 1,3-cycloaddition.

Figure 100. Graph depicting a screen of SmTrip9 G147 site-saturation variants.

Figure 101. Graph depicting a screen of SmTrip9 K148 site-saturation variants.

Figure 102. Graph depicting a screen of SmTrip9 M149 site-saturation variants.

Figure 103. Graph depicting a screen of SmTrip9 L150 site-saturation variants.

Figure 104. Graph depicting a screen of SmTrip9 F151 site-saturation variants.

Figure 105. Graph depicting a screen of SmTrip9 R152 site-saturation variants.

Figure 106. Graph depicting a screen of SmTrip9 V153 site-saturation variants.

Figure 107. Graph depicting a screen of SmTrip9 T154 site-saturation variants.

Figure 108. Graph depicting a screen of SmTrip9 I155 site-saturation variants.

Figure 109. Graph depicting a screen of SmTrip9 N156 site-saturation variants.

Figure 110. Graph depicting a screen of SmTrip9 S157 site-saturation variants.

Figure 111. Graph depicting a screen of SmTrip9 W158 site-saturation variants.

Figure 112. Graph depicting a screen of SmTrip9 K149 site-saturation variants.

Figure 113. Table of the results of FRB-FKBP facilitated complementation in E. coli lysates with SmTrip9 pep435/434.

Figure 114. Table of the results of FRB-FKBP facilitated complementation in E. coli lysates with SmTrip9 pep435/434.

Figure 115. Table of the results of FRB-FKBP facilitated complementation in E. coli lysates with SmTrip9 pep435/434.

Figure 116. Table of the results FRB-FKBP facilitated complementation assay screen with combinational SmTrip9 variants.

Figure 117. Table of the results FRB-FKBP facilitated complementation assay screen with combinational SmTrip9 variants.

Figure 118. Table of the results FRB-FKBP facilitated complementation assay screen with combinational SmTrip9 variants.

Figure 119. Table of the results FRB-FKBP facilitated complementation assay screen with combinational SmTrip9 variants.

Figure 120. Table of the results FRB-FKBP facilitated complementation assay screen with combinational SmTrip9 variants.

Figure 121. Table of the results FRB-FKBP facilitated complementation assay screen with combinational SmTrip9 variants.

Figure 122. Table of the results FRB-FKBP facilitated complementation assay screen with combinational SmTrip9 variants.

Figure 123. Table of Kd and Bmax of SmTrip9 synthetic peptides.

Figure 124. Table of Kd and Bmax of SmTrip9 synthetic peptides.

Figure 125. Table of Kd and Bmax of SmTrip9 synthetic peptides.

Figure 126. Table of Kd and Bmax of SmTrip9 synthetic peptides.

Figure 127. Table of Kd and Bmax of SmTrip9 synthetic peptides.

Figure 128. Table of Kd and Bmax of SmTrip9 synthetic peptides.

Figure 129. Table of Kd and Bmax of SmTrip9 synthetic peptides.

Figure 130. Table of Kd and Bmax of SmTrip9 synthetic peptides.

Figure 131. Table of Solubility of synthetic SmTrip9 peptides.

Figure 132. Graph of biochemical co-titration of SmTrip9 synthetic peptides and pep289.

Figure 133. Garph of biochemical co-titration of SmTrip9 synthetic peptides and pep289.

Figure 134. Graph of biochemical co-titration of SmTrip9 and SmTrip 10 synthetic peptides.

Figure 135. Graph of biochemical co-titration of pep521 and alternative SmTrip 10 synthetic peptides.

Figure 136. SDS PAGE gel of strand removal (purification) from LgTrip 3546 template.

Figure 137. Graphs of strand removal proteins with various combinations of peptides.

Figure 138. graphs of strands 6, 7, 8, 9, or 10 removal (purification) from LgTrip 3546 template.

Figure 139. Graphs of Kd and Bmax values of the dipeptide titrations.

Figure 140. Schematic depicting the approach taken to develop a solution-based homogeneous, quantitative assay for anti-TNFa biologic agents Remicade, Humira, and Enbrel using tripartite protein G and TNFa fusion proteins.

Figure 141. Graphs depicting quantitative analysis of TNFa inhibitor dose responses via facilitated complementation with SmTrip9 pep521-protein G (SEQ ID NO: 268) and TNFa-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO:150) fusion proteins with purified LgTrip 3546 (SEQ ID NO: 51) in a solution-based homogeneous assay.

Figure 142. Graphs depicting quantitative analysis of 10nM infliximab via facilitated complementation with SmTrip9 pep521-protein G (SEQ ID NO: 268) and TNFa-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO:150) fusion proteins with purified LgTrip 3546 (SEQ ID NO: 51) in the presence of complex sample matrices including human serum and urine using a solution-based homogenous assay.

Figure 143. Graph depicting the binding kinetics of signal generation measuring 100pM of infliximab via facilitated complementation with SmTrip9 pep521-protein G (SEQ ID NO: 268) and TNFa-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO:150) fusion proteins with purified LgTrip 3546 (SEQ ID NO: 51) in a solution-based homogenous assay.

Figure 144. Graph depicting signal generation measuring 10nM of infliximab via facilitated complementation of different SmTrip9 pep(X)-protein G variants and TNFa-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO:150) fusion proteins with purified LgTrip 3546 (SEQ ID NO: 51) in a solution-based homogenous assay.

Figure 145. Schematic depicting the approach taken to develop a homogenous cell-based, quantitative assay for anti-EGFR biologic agents panitumumab and cetuximab using SmTrip9-protein G fusion proteins and HEK293 cells expressing SmTrip10 pep289-EGFR (SEQ ID NO:150).

Figure 146. Graph depicting quantitation of Panitumumab via facilitated complementation with SmTrip9 pep521-protein G (SEQ ID NO: 268) fusion protein and SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) expressing cells with purified LgTrip 3546 (SEQ ID NO: 51) in a cell-based homogeneous assay.

Figure 147. Graph depicting the real time binding kinetics of signal generation measuring Cetuximab via facilitated complementation with SmTrip9 pep521-protein G (SEQ ID NO: 268) fusion protein and SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) expressing cells with purified LgTrip 3546 (SEQ ID NO: 51) in a cell-based homogeneous assay.

Figure 148. Graph depicting signal generation measuring 1nM of panitumumab via facilitated complementation of different SmTrip9 pep(X)-protein G variants and SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) expressing cells paired with purified LgTrip 3546 (SEQ ID NO: 51) in a cell-based homogenous assay.

Figure 149. Graph depicting quantitation of panitumumab dose response via facilitated complementation of different SmTrip9 pep(X)-protein G variants and SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) expressing cells paired with purified LgTrip 3546 (SEQ ID NO: 51) in a cell-based homogenous assay.

Figure 150. Graphs depicting quantitation of human IL-1beta using Halotag-SmTrip9 pep521 (SEQ ID NO: 268) and HaloTag-SmTrip10 pep289 (SEQ ID NO: 150) chemically-

labeled paired antibodies in a solution-based homogeneous assay. Real-time binding kinetics for human Troponin using NanoTrip chemically labeled paired antibodies.

Figure 151. Graphs depicting real time binding kinetics for quantitation of human Troponin using Halotag-SmTrip9 pep521 (SEQ ID NO: 268) and HaloTag-SmTrip10 pep289 (SEQ ID NO: 150) chemically labeled paired antibodies in a solution-based homogeneous assay.

Figure 152. Specialized peptides responsible to direct proteins to specific cellular compartments were fused to LgBiT-HaloTag. (A) LgBiT-membrane sensor: LgBiT is in green and nucleus is in blue. (B) LgBiT-nuclear sensor: LgBiT is in green and nucleus is in blue. (C) LgBiT-mitochondria sensor: LgBiT is in green, MitoTracker is in red, and nucleus is in blue. (D) LgBiT-ER sensor: LgBiT is in green, ER marker is in red, and nucleus is in blue.

Figure 153. Translocation assay. POI is endogenously tagged with HiBiT. Upon stimulation, the POI translocates to a different cellular compartment, for example the nucleus. A LgBiT-nuclear sensor could be used to detect this translocation event as the HiBiT meets LgBiT resulting in luminescence signal.

Figure 154. Membrane translocation assay with wild-type LgBiT sensor. PKCα-HiBiT cell line was transfected with wild-type LgBiT-membrane sensor. Due to the strong interaction between LgBiT and HiBiT, the spontaneous complementation occurs, leading to no response to PMA stimuli.

Figure 155. Membrane translocation assay with LgBiT* sensor (SEQ ID NO: 979). PKCα-HiBiT cell line was transfected with different amount of DNA encoding LgBiT*-membrane sensor. Upon PMA treatment, PKCα-HiBiT migrates to the plasma membrane, where the LgBiT*-membrane sensor anchors. The assembly between HiBiT and LgBiT* produces luminescence signal, and the signal is proportional to the amount of PKCα-HiBiT on the membrane. The assay is sensitive and robust. Titration of PMA yielded similar EC_{50} (EC_{50} = 2.0 nM) regardless of the amount of sensor input. Fold response is varied between 12- to 19-fold depending on the amount of sensor input.

Figure 156. Nuclear translocation assay with LgBiT* sensor. p65-HiBiT cell line was transfected with DNA encoding LgBiT*-nuclear sensor. Addition of TNFα recruits p65 to the nucleus where LgBiT*-nuclear sensor localizes. Complementation occurs between HiBiT and LgBiT* to produce light. The signal intensity reflects the concentration of p65 in the nucleus.

Titration of TNFα yielded EC_{50} of 0.7 ng/mL with fold-response of 4. Real-time measurement showed that it requires 30 min to reach the maximum accumulation of p65 in the nucleus upon stimulation.

Figure 157A-B. (A) Graph and (B) table depicting affinity and Bmax of LgBiT mutants with HiBiT.

Figure 158. Graph depicting affinity of LgBiT mutant lysates for HiBiT.

Figure 159. Graph and table depicting the relative bioluminescence, produced in the presence of substrate, of various combinations of truncated luciferase polypeptide and peptide components.

Figure 160. LgTrip was diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol and pep289 was added to 25 uM. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (20x Kd) and titration of VS-HiBiT.

Figure 161. LgTrip was diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol and pep289 was added to 25 uM. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (20x Kd) and titration of VS-HiBiT.

Figure 162. Dipeptides were diluted to 5 uM and diluted serially 5-fold using TBS + 0.01% BSA + 0.01% Tergitol with 0.2 nM of LgTrip as the diluent. Samples were incubated 10 minutes at room temperature and added to assay plates in triplicate. One-to-one vol:vol of TBS + 0.01% BSA + 0.01% Tergitol with 20x diluted live cell substrate was added to samples, and plates were read on a GloMax® luminometer after 10 minutes.

Figure 163. LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 4 hours at 37°C with shaking. Ten microliters of each induction was diluted into 250 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5). Eighty

microliters of lysates were diluted further in 2 ml lysis buffer. A 10-fold dilution series of pep788 (SEQ ID 414) was performed stating at 10 µM peptide using Nano-Glo with 50 uM furimazine as the diluent. Peptide dilutions and lysates were mixed 1:1 vol:vol, incubated 10 min at room temperature, and luminescence was read.

Figure 164. LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 4 hours at 37°C with shaking. Ten microliters of each induction was diluted into 250 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5). Eighty microliters of lysates were diluted further in 2 ml lysis buffer. A 5-fold dilution series of pep759 (SEQ ID 496) was performed stating at 50 µM peptide using Nano-Glo with 50 uM furimazine and 50 µM pep289 (SEQ ID 826) as the diluent. Peptide dilutions and lysates were mixed 1:1 vol:vol, incubated 10 min at room temperature, and luminescence was read.

Figure 165. LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 20 hours at 25°C with shaking. Twenty microliters of each induction was diluted into 40 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5) and lysed for 15 min at room temperature. Lysates were diluted 1,000-fold into 1X TBS + 0.01% BSA. Fifty microliters of each sample was transferred into a PCR plate and incubated at 80°C for 1.5 hrs in a thermocycler. Controls were incubated on ice. Samples were equilibrated to room temperature and diluted 1:100 into 1X TBS + 0.01% BSA. Twenty-five microliters of each sample were transferred into assay plates and mixed with 25 ul of 400nM pep788 (SEQ ID 414) in TBS + 0.01% BSA + 20x diluted live cell substrate. Samples were incubated 10 minutes at room temperature, and luminescence was read.

Figure 166. LgTrip variants were purified using the Promega MagneHis™ Protein Purification system according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT.

Figure 167. LgTrip variants were purified using the Promega MagneHis™ Protein Purification system according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT.

Figure 168. LgTrip variants were purified using the Promega MagneHis™ Protein Purification system according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (12.5 uM) and titration of VS-HiBiT.

Figure 169. MagneHis purified LgTrip variants were diluted to 20 nM in TBS+0.01% BSA. In triplicate, 100µl aliquots of each sample were loaded into 200µl thin walled PCR tubes. Samples were incubated at 70°C in thermal cycler. Samples were removed at various time-points and equilibrated to room temperature. Samples were diluted to 0.2 nM (5 µl in 495 µl) in TBS+0.01% BSA. 25 µl of each diluted sample was combined with 25 µl of TBS + 0.01% BSA +20x diluted live cell substrate furimazine + 400 nM pep788 (SEQ ID NO: 414). Samples were incubated for 10 minutes and then read on GMM+. Half-life was calculated by non-linear regression.

Figure 170. Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10s were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. PLB assay reagent was prepared with 444 nM of a MagneHis purified LgTrip variants, 90x diluted FRB-SmTrip10 culture, +/- 35 nM Rapamycin. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB, and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of Nano-Glo containing 50 uM furimazine was

added to assay plates wells, and luminescence was read on GloMax® luminometer after 5 minutes.

Figure 171. Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10s were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. PLB assay reagent was prepared with 444 nM of a MagneHis purified LgTrip variant, 90x diluted FRB-SmTrip10 culture, +/- 35 nM Rapamycin. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB, and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of Nano-Glo containing 50 uM furimazine was added to assay plates wells, and luminescence was read on GloMax® luminometer after 5 minutes.

Figure 172. LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 20 hours at 25°C with shaking. Twenty microliters of each induction was diluted into 40 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5) and lysed for 15 min at room temperature. Lysates were diluted 1,000-fold into 1X TBS + 0.01% BSA. Fifty microliters of each sample was transferred into a PCR plate and incubated at 70°C for 1.5 hrs in a thermocycler. Controls were incubated on ice. Samples were equilibrated to room temperature and diluted 1:100 into 1X TBS + 0.01% BSA. Twenty-five microliters of each sample were transferred into assay plates and mixed with 25 ul of 400nM pep788 (SEQ ID 414) in TBS + 0.01% BSA + 20x diluted live cell furimazine substrate. Samples were incubated 10 minutes at room temperature, and luminescence was read.

Figure 173. LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 20 hours at 25°C with shaking. Twenty microliters of each induction was diluted into 40 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5) and lysed for 15 min at room temperature. Lysates were diluted 1,000-fold into 1X TBS + 0.01% BSA. Fifty microliters of each sample was transferred into a PCR plate and incubated at two temperature gradients, either 75-100°C for 10 min or 50-75°C for 1.5 hr, in a Veritas thermocycler. Controls were incubated on ice. Samples were equilibrated to room temperature and diluted 1:100 into 1X TBS + 0.01% BSA. Twenty-five microliters of each

sample were transferred into assay plates and mixed with 25 ul of 400nM pep788 (SEQ ID 414) in TBS + 0.01% BSA + 20x diluted live cell furimazine substrate. Samples were incubated 10 minutes at room temperature, and luminescence was read.

Figure 174. MagneHis purified LgTrip variants were diluted to 20 nM in TBS+0.01% BSA. In triplicate, 100µl aliquots of each sample were loaded into 200µl thin walled PCR tubes. Samples were incubated at 70°C in a thermal cycler. Samples were removed at various time-points and equilibrated to room temperature. Samples were diluted to 0.2 nM (5 in 495µl) in TBS+0.01% BSA. 25 µl of each diluted sample was combined with 25 µl of TBS + 0.01% BSA + 20x diluted live cell substrate furimazine + 400 nM pep788 (SEQ ID NO: 414). Samples were incubated for 10 minutes and then read on GMM+.

Figure 175. LgTrip variants were purified using the Promega MagneHis™ Protein Purification system according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (12.5 uM) and titration of VS-HiBiT.

Figure 176. LgTrip variants were purified using the Promega MagneHis™ Protein Purification system according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (12.5 uM) and titration of VS-HiBiT.

Fiogure 177. LgTrip variants were purified using the Promega MagneHis™ Protein Purification system according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated

10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT.

Figure 178. LgTrip variants were purified according to using the Promega MagnaHisMagneHis™ Protein Purification system according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT.

Figure 179. LgTrip variantss were purified according to using the Promega MagnaHisMagneHis™ Protein Purification system according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (20x Kd) and titration of VS-HiBiT.

Figure 180. LgTrip variants were purified according to using the Promega MagnaHisMagneHis™ Protein Purification system according to the manufacturer's protocoland diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (20x Kd) and titration of VS-HiBiT.

Figure 181. LgTrip variant cultures were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin and induced ~20 hr at 25°C with shaking. Cells were diluted 1000x in

PLB assay reagent (0.3X PLB + 25mM HEPES pH 7.5) and lysed for 20 min. Five-fold serial dilutions of SmTrip9 pep840 were performed in Nano-Glo + 50 uM furimazine + 25 uM pep289 and mixed 1:1 vol:vol with LgTrip lysates. Samples were incubated 10 min at room temperature and read on GloMax® luminometer.

Figure 182. LgTrip variants were purified according to using the Promega MagnaHisMagneHis™ Protein Purification system according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT.

Figure 183. MagneHis purified LgTrip variants were diluted to 20 nM in TBS + 0.01% BSA. In triplicate, 100µl aliquots of each sample were loaded into 200µl thin walled PCR tubes. Samples were incubated at 70°C in a thermal cycler. Samples were removed at various time-points and equilibrated to room temperature. Samples were diluted to 0.2 nM (5 in 495µl) in TBS + 0.01% BSA. 25 µl of each diluted sample was combined with 25 µl of TBS + 0.01% BSA + 20x diluted live cell substrate furimazine + 200 nM pep900 (SEQ ID NO: 907). Samples were incubated for 10 minutes and then read on GMM+. Half-life was calculated by non-linear regression.

Figure 184. Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10s were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. PLB assay reagent was prepared with 444 nM of a MagneHis purified LgTrip variant, 90x diluted FRB-SmTrip10 culture, +/- 35 nM Rapamycin. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB, and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of Nano-Glo containing 50 uM furimazine was added to assay plates wells, and luminescence was read on GloMax after 5 minutes.

Figure 185. LgTrip variant cultures were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and

100ug/ml ampicillin and induced ~20 hr at 25°C with shaking. Cells were diluted 5000x in 0.3x PLB assay reagent and lysed for 20 min. Five-fold serial dilutions of dipeptides pep788, pep900, or SmTrip9 pep840 with saturating pep289 were performed in Nano-Glo + 50 uM furimazine and mixed 1:1 vol:vol with LgTrip variant lysates. Samples were incubated 10 min at room temperature and read on GloMax® luminometer. Bmax was calculated by non-linear regression.

Figure 186. LgTrip variants were diluted to 20 nM in 2 ml TBS+0.01% BSA, and 100ul of each sample aliquoted into duplicate rows of 96-well PCR plates (two plates). Plates were incubated for 3 hrs at high (75-100°C) or low (50-75°C) temperature gradient in a Veritas thermocycler. Samples were placed at 70°C and then aliquots moved to RT at various timepoints. Samples were equilibrated to RT. Samples at each timepoint were mixed with pipette and then diluted 1:100 into TBS + 0.01% BSA(5ul into 495ul). 25ul of each sample was aliquoted into a white assay plate, and 25ul of 200nM pep788 or pep900 in TBS + 0.01% BSA + 20x dil LCS added. The plate was incubated for 10 minutes and then luminescence read on GMM+.

Figure 187. Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10s were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. PLB assay reagent was prepared with 444 nM of a MagneHis purified LgTrip variant, 90x diluted FRB-SmTrip10 culture, and +/- 35 nM Rapamycin. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB, and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of Nano-Glo containing 50 uM furimazine was added to assay plates wells, and luminescence was read on GloMax® luminometer after 5 minutes.

Figure 188. LgTrip 3546 was diluted to 1nM in OptiMEM+10% FBS. 12μM solutions of strand 9 peptides 521 and 693 were prepared in OptiMem+10% FBS. Each strand 9 dilution was used to prepare 3-fold dilution series of each strand 10 peptide starting at 20μM. (pep86=HiBiT, pep289=VS HiBiT, pep691=HiBiT RR, pep692=VSHiBiT RR). 90μl of each dilution series was transferred to a white assay plate and then 10μl of the 1μM stock of LgTrip 3546 added. The plate was placed on an orbital shaker set to 600RPM for 30 minutes. Detection reagent of OptiMEM + 10% FBS consisting of 10mM DTT and 50uM

Furimazine was prepared, and 11µl added to the samples. The plate was placed on orbital shaker and mixed for 5 minutes at room temperature. The plate was read on a GloMax® Multi+ luminometer. Kd and Bmax were calculated using GraphPad Prism one site specific binding.

Figure 189. Purification of Proteins. Cultures for each sample were started from an isolated colony in LB + 30ug/ml KAN and grown for 20 hours at 37°C. Cultures were diluted 1:100 (500ul in 50ml) of LB supplemented with 30ug/ml KAN, 0.1% Rhamnose, and 0.15% glucose. Cultures were grown for 20 hours at 25°C. Cells were spun and resuspended in 9ml 100mM HEPES pH 7.5 + 1ml FastBreak Lysis Buffer + 200ul RQ DNAse I and placed on orbital mixer for 30 minutes at 4°C (An aliquot of the total lysate was saved). Each sample was spun (7000RPM 15 minutes) to clear lysate, and MagneHis™ purification system (Promega V8550) used to purify proteins. 1ml of MagneHis™ magnetic resin was added to each cleared lysate and placed on orbital mixer for 10 minutes. Each sample was washed three times with wash buffer and then 500ul of elution buffer was used 2 times to recover samples. The samples were dialyzed into 1X TBS using 10,000 MWCO dialysis units (Thermo 88404) for 2 hours.

Figure 190. Luminescnece comparison of NanoLuc (ATG-462) with monomeric LgBiT-SmBiT proteins. Purified protein was diluted to 0.2nM in TBS + 0.01% BSA. 50ul of each sample was combined with 50ul of 50uM Fz in Nano-Glo® buffer (N113, N112), and luminescence measured on GMM+ luminometer three minutes after substrate addition. NanoLuc was ~2X brighter compared to monomeric constructs

Figure 191. Km determination. Purified protein was diluted to 0.2nM in TBS + 0.01% BSA. A 2-fold dilution series of Fz in Nano-Glo® buffer starting at 50uM (40ul in 2ml) was prepared, then 1ml to 1ml dilutions. 50ul of each sample in duplicate was combined with 50ul of the titration series, incubated for three minutes at RT, and then luminescence read on GMM+

Figure 192. Stability of monomeric LgBiT-SmBiT protein compared to NanoLuc (ATG-462). Each sample was diluted to 2nM in TBS + 0.01% BSA, aliquoted into multiple wells of a 96-well PCR plate, and placed in a thermocycler set for 30 minutes with temperature gradient. After a 30-minute incubation, 5ul of each sample was combined with 45ul of TBS + 0.01% BSA, 50ul of Fz diluted into Nano-Glo® buffer added, incubated for 3 minutes, and luminescence measured on a GMM+ luminometer . Temperature gradient A:

54°, 57°, 60°, 63°, 66°, and 70°C. Temperature gradient B: 55°, 60°, 65°, 70°, 75°, and 80°C. Temperature gradient C: 65°, 70°, 75°, 80°, 85°, and 90°C.

Figure 193. Thermostability comparison. Each sample was diluted to 2nM in TBS + 0.01% BSA, aliquoted into multiple wells of two-96-well PCR plateand placed in thermocyclers set for 30 minutes with temperature gradient. (75°, 80°, 85°, 90°, 95°, and 100°C). After a 30-minute incubation, 5ul of each sample was combined with 45ul of TBS + 0.01% BSA, 50ul of Fz diluted into Nano-Glo® buffer added, incubated for 3 minutes, and luminescence read a GMM+ luminomether. Position 159G provides enhanced stability compared to 159S.

Figure 194. Stbility comparison of LgBiT-SmBiT variants with NanoLuc. Samples were diluted to 2nM, aliquoted into wells of a 96-well PCR plate, and placed in a thermocycler set to 60°C. At various timepoints, an aliquot was removed and placed on ice. After all samples were incubated, samples were allowed to equilibrate to RT, and 5ul of each sample combined with 45ul of TBS + 0.01% BSA. 50ul of Nano-Glo® buffer + 50uM Fz was added, the plated incubated 3 minutes, and then luminescence measure on GMM+ luminometer. The 60°C stability curve is consistent with temperature gradients curves with LgBiT-SmBiT slightly more stable compared to NanoLuc, and LgBiT-HiBiT significantly more stable compared to NanoLuc.

Figure 195. Stability of monomeric LgBiT-SmBiT at elevated temperature. Samples were diluted to 200nM in TBS + 0.01% BSA and then further diluted to 0.2nM (4ul in 4ml). 50ul of ATG-462 or ATG-3564 was combined with 50ul of Nano-Glo® buffer + 50uM Fz or 50ul of TBS + 0.01%BSA + 20uM Fz in wells of a thin wall PCR tray, and the tray placed in a Veritas thermocycler set with gradient 55°, 60°, 65°, 70°, 75°, and 80°C. At various timepoints, samples were removed, and luminescence measure on GMM+ luminometer.

Figure 196. Stability of monomeric LgBiT-SmBiT at elevated temperature for 30 minutes. Samples were diluted to 200nM in TBS + 0.01% BSA and then further diluted to 0.2nM (4ul in 4ml). 50ul of ATG-462 or ATG-3564 was combined with 50ul of Nano-Glo® buffer + 50uM Fz or 50ul of TBS + 0.01%BSA + 20uM Fz into wells of a thin wall PCR tray, and the tray placed in a Veritas thermocycler set with gradient 55°, 60°, 65°, 70°, 75°, and 80°C. Luminescence was measure at 30 minutes.

Figure 197. Measurement of Km and Vmax for purified NanoLuc® mutants. Purified protein was diluted to 0.2nM in TBS+0.01% BSA. A 2-fold dilution series of Fz in Nano-

Glo® buffer starting at 50uM (40ul in 2ml), then 1ml to 1ml dilutions. 50ul of each sample in duplicate was combined with 50ul of the titration series, incubated for three minutes at RT, and luminescence read on GMM+ luminometer.

Figure 198. Thermal challenge.

Figure 199. Thermal challenge with stable variants. Each sample was diluted to 2nM in TBS + 0.01% BSA, aliquoted into multiple wells of two-96-well PCR plates, and placed in a thermocyclers set for 30 minutes with temperature gradient. (70°, 75°, 80°, 85°, 90°, and 95°C). After 30-minute incubation, 5ul of each sample was combined with 45ul of TBS + 0.01% BSA, 50ul of Fz diluted into Nano-Glo® buffer added, incubated for 3 minutes, and luminescence measured on a GMM+ luminometer.

Figure 200. Comparison of pep521 and pep 40 with detector proteins. Each protein was diluted to 20nM in TBS + 0.01% BSA. A 3X dilution series of pep521 and pep840 was prepared in TBS + 0.01% BSA starting at 20uM. 50ul of each enzyme dilution in duplicate was combined with 50ul of each peptide titration and incubated for 10 minutes on shaker for pre-equilibration. An assay buffer was prepared by diluting Nano-Glo Live Cell Substrate (LCS; Promega N205) 30-fold into TBS + 0.01% BSA, 100ul added to each well, incubated 5 minutes, and luminescece measured on GMM+ luminometer. The background reading was obtained from samples that contained no peptide.

Figure 201. Comparison of strand 9 detector proteins with pep840. Each protein was diluted to 200nM in TBS + 0.01% BSA. A 3X dilution series of pep840 was prepared in TBS + 0.01% BSA starting at 0.5uM. 50ul of each enzyme dilution in quadruplicate was combined with 50ul of each peptide titration and incubated for 10 minutes on shaker for pre-equilibration. An assay buffer was prepared by diluting Nano-Glo LCS (Promega N205) 30-fold into TBS + 0.01% BSA, 100ul added to each well, incubated 5 minutes, and luminescence measured on GMM+ luminometer. The background reading was obtained from samples that contained no peptide.

Figure 202. Test linker series. Overnight cultures of each sample were prepared in LB + 100ug/ml amp. The following day, cultures were diluted 1:20 (150ul to 3ml) in LB + 0.1% Rhamnose + 100ug/ml amp, grown for 4 hours at 37°C, and then lysed with PLB lysis buffer (0.3X Passive Lysis Buffere (Promega) + 25mM HEPES pH 7.5) (500ul of lysate to 4.5ml of PLB lysis buffer). To assay, lysates were diluted 1:100 into TBS + 0.01% BSA. A 3X dilution series of strand 9 (pep840) starting at 10nM was prepared,50ul combined with 50ul

of the diluted lysate for each sample, and incubated for 20 minutes at RT on orbital shaker set to 600rpm. 100ul of NanoGlo® buffer +50uM Furimazine (N113) was added, and luminescence measured on GMM+ luminometer. Each clone, with the exception of ATG-5485, 5AA linker,produced similar RLU values to ATG-4992.

Figure 203. Linker comparison: 8GS (ATG-4992) vs. 11GS (ATG-5490). ATG-4992 and ATG-5490 were purified using the MagneHis™ purification system and diluted to 100nM in CO_2 independent media + 10% FBS. A 3-fold serial dilution of pept840 was prepared starting at 10nM in Nano-Glo® buffer + 50uM furimazine. 50ul of each enzyme dilution was combined in quadruplicate with 50ul the peptide titration. Luminescence was measured over time in a GMM+ luminometer. The data plotted is from the 50-minute kinetic read. ATG-5490 had a ~2-fold higher signal compared to ATG-4992

Figure 204. Kd of mutants of ATG-4992 and ATG-5490. Overnight cultures of each sample in LB+100ug/ml amp were prepared. The following day, cultures were diluted 1:20 (150ul to 3ml) in LB + 0.1% Rhamnose + 100ug/ml amp, grown for 4 hours at 37°C, and lysed with PLB lysis buffer (500ul of lysate to 4.5ml of PLB lysis buffer). To assay, lysates were diluted 1:100 into TBS + 0.01% BSA. A 3X dilution series of strand 9 (pep840) was prepared starting at 40uM, 50ul combined with 50ul of the diluted lysate for each sample, and incubated for 20 minutes at RT on orbital shaker set to 600rpm. 100ul of Nano-Glo® buffer + 50uM Furimazine (N113), and luminescence measured on GMM+ luminometer.

Figure 205. Comparison of mutants of ATG-4992 and ATG-5490. Proteins were purified (MagneHis™ Purification System, Promega) and diluted to 100nM in CO_2 independent media + 10% FBS. A 3-fold dilution series was prepared with pep840 starting at 2nM in Nano-Glo® buffer + 50uM furimazine. 50ul of each enzyme dilution was combined in triplicate with 50ul of each peptide titration in the Furimazine/NanoGlo solution. Luminescence was measured every 5 minutes. Data is plotted from the 60-minute timepoint. Clones ATG-5515 and ATG-5517 had lower calculated Kd values and showed higher RLU values when paired with pep840 and slightly higher signal to background

Figure 206. Biochemical analysis. Samples were purified using MagneHis™ Protein Purification System (Promega V8500) and diluted to 0.2nM in TBS + 0.01% BSA. A 3X dilution series of VS-HiBiT peptide (pep289) was prepared in TBS + 0.01% Tergitol. One series was prepared starting at 400nM and the other at 20uM. 50ul of each diluted enzyme (0.2nM) was combined with 50ul of each peptide titration and incubated on shaker for 10

minutes. 100ul of Fz, diluted 250- fold in TBS + 0.01% BSA was added, the plate placed on shaker for 5 minutes, and luminescence measured on a GMM+ luminometer. Kd and Bmax was calculated using GraphPad prism one site specific binding regression.

Figure 207. Activity at various pH. Each sample of purified enzyme was diluted to 0.2nM in TBS + 0.01% BSA, 20uM VS HiBiT (pep289) was added to each solution, and incubated at RT for 20 minutes. 990ul of each pH buffer was aliquoted into wells of a deep well plate, 10ul of furimazine (Promega, N113) added, and 50ul of each enzyme/peptide dilution with 50ul of the pH buffer/furimazine solution added. Samples were incubated for 12 minutes at RT, and luminescence measured on GMM+ luminometer. Data was normalized to the pH 8.49 sample.

Figure 208. Biochemical comparison of ATG-5823, ATG-5824, ATG-5825, and ATG-5146 (pep 263). Proteins (ATG-5823, ATG-5824, and ATG-5825) were purified using MageHis™ purification system (Promega). ATG-5146 was purified using an AKTA with Nickel Sepharose column. Purified proteins were diluted first to 200nM in TBS + 0.01% BSA and then further diluted to 0.2nM in TBS + 0.01% BSA. A three-fold dilution series of pep263 was prepared starting at 100nM in TBS + 0.01% BSA + 0.02% Tergitol. 50ul of each enzyme was combined with 50ul of the peptide dilution series. Samples were incubated on orbital shaker (600RPM) for 10 minutes. After incubation, 100ul of LCS (N205 Promega) was diluted 1:30 into TBS + 0.01% BSA and added to each sample. Samples were incubated for 3 minutes at RT, and then luminescence was measured luminescence was measured on a GloMaxMulti+.

Figure 209. Biochemical comparison of ATG-5826 and ATG-5827 with pep86. Proteins were purified using MagneHis™ purification system (Promega V8500). Purified proteins were diluted first to 200nM in TBS + 0.01% BSA and then further diluted to 0.2nM in TBS + 0.01% BSA. Two, 2-fold titration series of pep86 were prepared starting at 1uM and 100nM in TBS + 0.01% BSA + 0.02% Tergitol. 50ul of ATG-5826 and ATG-5827 were combined with 50ul of the peptide dilution series that started at 1uM. 50ul of LgBiT protein (Promega N401C) was combined with the peptide titration series that started at 100nM. Samples were incubated on orbital shaker (600RPM) for 10 minutes. After incubation, 100ul of LCS (N205 Promega) was diluted 1:30 into TBS + 0.01% BSA and added to each sample. Samples were incubated for 3 minutes at RT, and then luminescence was measured luminescence was measured on a GloMaxMulti+.

Figure 210. Biochemical comparison of ATG-5826 and ATG-5827 with pep114. Proteins were purified using MagneHis™ purification system (Promega V8500). Purified proteins were diluted first to 200nM in TBS + 0.01% BSA and then further diluted to 0.2nM in TBS + 0.01% BSA. Two, 2-fold titration series of pep114 was prepared starting at 1mM in TBS + 0.01% BSA + 0.02% Tergitol. 50ul of ATG-5826, ATG-5827, and LgBiT protein (Promega N401C) were combined with the peptide titration series. Samples were incubated on orbital shaker (600RPM) for 30 minutes. After incubation, 100ul of LCS (N205 Promega) was diluted 1:30 into TBS + 0.01% BSA and added to each sample. Samples were incubated for 3 minutes at RT, and then luminescence was measured luminescence was measured on a GloMaxMulti+.

Figure 211. SDS PAGE analysis. Each protein was diluted to 0.1ug/ml in TBS+1X SDS loading dye. Samples were heated to 70°C for 5 minutes and then 3ul was loaded to an SDS PAGE gel.

Figure 212. Exemplary assay formats for the detection of SARS-COV-2.

Figure 213. Exemplary reagents used to screen reactive peptides. SulfoSE reacts with lysines on target proteins, PEG6 linker and pyridinium sulfonic acid (PSA) provide solubility, and PSA has UV signature that can be used to measure labeling density.

Figure 214. Spike-in antibody titration of alternative labeling strategies for detection of SARS-COV-2.

Figure 215. Purified SARS-CoV-2 Nucleocapsid antigen detection.

Figure 216. Samples tested with the lyophilized antigen assay in handheld assay. Patient nasal swab was placed into reagent tube breaking the foil seal at the bottom of the tube. Buffer containing the reagent cap was placed onto the reagent tube containing sample, which locks. The buffer capsule was cracked to release and then shook. The reagent tube was inserted into a handheld luminometer, and the sample analyzed.

Figure 217. Monomeric NanoBiT with Fluoro-Fz. Each enzyme to be tested was diluted into TBS + 0.01% BSA. A titration series with either FZ (N205) or JRW-1677 was prepared in either TBS (starting at 20uM) or NanoGlo (starting at 25uM). Each substrate was serially diluted with either TBS + 0.01% BSA or Nano-Glo® buffer (Promega N112). 50ul of each enzyme dilution was combined with 50ul of each substrate titration, incubated for 3 minutes, and then measured luminescence on GMM+ luminometer. Kinetic parameters for Fz and JRW-1667 were calculated. NanoLuc (ATG-462) produced higher RLU values with

both Fz and JRW-1667 (~10-20 fold) compared to the Monomeric NanoBiT constructs when TBS was used to dilute substrates. NanoLuc and the monomeric constructs showed similar RLU values for Fz/NanoGlo buffer, but only NanoLuc showed improved luminescence with JRW-1667. Although the monomeric NanoBiT constructs showed lower RLU values with both Fz and JRW-1667, RLU values were similar for the two buffers and two substrates. Km values are lower in TBS + 0.01% BSA compared to NanoGlo, but generally the Km values were similar for each condition except ATG-3562 (Monomeric LgBiT-SmBiT), which showed a lower calculated Km for all conditions tested.

Figure 218. HaloTag-LgBiT* - mammalian cell expression and fluorescence imaging. Immunofluoresence images of HaloTag-LgBiT* in HeLa cells showed its ubiquitous distribution in the cells. Left: red channel only, Middle: blue channel only; Right: overlay.

Figure 219. Nuclear sensor - mammalian cell expression and fluorescence imaging. Immunofluoresence images of nuclear sensor in HeLa cells showed its nuclear localization. Left: red channel only, Middle: blue channel only; Right: overlay.

Figure 220. Mitochondrial sensor - mammalian cell expression and fluorescence imaging. Immunofluoresence images of mitochondria sensor in HeLa cells showed its localization in mitochondria matrix. DNA stain in blue. Left: red channel only, Middle: green channel only; Right: overlay.

Figure 221. ER sensor - mammalian cell expression and fluorescence imaging. Immunofluorescence images of endoplasmic reticulumn sensor (ER) in Hela cells showed its ER localization. DNA stain in blue. Left: green channdel only; Middle: red channel only; Right: overlay.

Figure 222. Golgi sensor - mammalian cell expression and fluorescence imaging. Immunofluorescence images of Golgi sensor in Hela cells showed its Golgi localization. Golgi-GFP marker only stained for cis-Golgi. A portion of our Golgi sensor did not co-localize with Golgi-GFP marker, suggesting our Golgi sensor might reside in both cis- and trans-Golgi. DNA stain in blue. Left: red channdel only; Middle: green channel only; Right: overlay.

Figure 223. Lysosome sensor - mammalian cell expression and fluorescence imaging. Immunofluoresence images of lysosome sensor in HeLa cells showed its localization to the lysosome. DNA stain in blue. Left: red channel only, Middle: green channel only; Right: overlay.

Figure 224. Membrane sensor - mammalian cell expression and fluorescence imaging. Immunofluoresence images of membrane sensor in HeLa cells showed its localization to the plasma membrane. DNA stain in blue. Left: red channel only, Middle: green channel only; Right: overlay.

Figure 225. Membrane sensor - mammalian cell expression and luinescence assays. Translocation of PKCα under PMA stimulation. PKCα-HiBiT clone (HeLa) was transiently transfected with membrane sensors. Transfected cells were treated with phorbol-12-myristate-13 acetate (PMA) and measured kinetically. All 3 membrane sensors resulted in similar EC_{50}, suggesting the affinity of LgBiT variants to HiBiT do not effect translocation events of PKCα.

Figure 226. Membrane/ HaloTag-LgBiT sensor - mammalian cell expression and luinescence assays. Reversibility of HiBiT-LgBiT* complex. Increased PMA concentration recruits more PKCα to the plasma membrane, resulting in higher luminescence signal, as more PKCα-HiBiT complementing with LgBiT* membrane sensor. Conversely, less PKCα accumulates in the cytosol; thus, decreasing luminescence signal as pairing with HaloTag-LgBiT (cytosolic sensor).

Figure 227. Translocation event of PKCα. Kinetic measurement of PKCα translocation event. Accumulation of PKCα to the plasma membrane can be measured kinetically, when paired with membrane sensor. Under PMA stimulation, PKCα reaches its maximum accumulation to the plasma membrane at t =16 min.

DEFINITIONS

Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of embodiments described herein, some preferred methods, compositions, devices, and materials are described herein. However, before the present materials and methods are described, it is to be understood that this invention is not limited to the particular molecules, compositions, methodologies or protocols herein described, as these may vary in accordance with routine experimentation and optimization. It is also to be understood that the terminology used in the description is for the purpose of describing the particular versions or embodiments only, and is not intended to limit the scope of the embodiments described herein.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention

belongs. However, in case of conflict, the present specification, including definitions, will control. Accordingly, in the context of the embodiments described herein, the following definitions apply.

As used herein and in the appended claims, the singular forms "a", "an" and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a peptide" is a reference to one or more peptides and equivalents thereof known to those skilled in the art, and so forth.

As used herein, the term "and/or" includes any and all combinations of listed items, including any of the listed items individually. For example, "A, B, and/or C" encompasses A, B, C, AB, AC, BC, and ABC, each of which is to be considered separately described by the statement "A, B, and/or C."

As used herein, the term "comprise" and linguistic variations thereof denote the presence of recited feature(s), element(s), method step(s), etc. without the exclusion of the presence of additional feature(s), element(s), method step(s), etc. Conversely, the term "consisting of" and linguistic variations thereof, denotes the presence of recited feature(s), element(s), method step(s), etc. and excludes any unrecited feature(s), element(s), method step(s), etc., except for ordinarily-associated impurities. The phrase "consisting essentially of" denotes the recited feature(s), element(s), method step(s), etc. and any additional feature(s), element(s), method step(s), etc. that do not materially affect the basic nature of the composition, system, or method. Many embodiments herein are described using open "comprising" language. Such embodiments encompass multiple closed "consisting of" and/or "consisting essentially of" embodiments, which may alternatively be claimed or described using such language.

As used herein, the term "substantially" means that the recited characteristic, parameter, and/or value need not be achieved exactly, but that deviations or variations, including for example, tolerances, measurement error, measurement accuracy limitations and other factors known to skill in the art, may occur in amounts that do not preclude the effect the characteristic was intended to provide. A characteristic or feature that is substantially absent (e.g., substantially non-luminescent) may be one that is within the noise, beneath background, below the detection capabilities of the assay being used, or a small fraction (e.g., <1%, <0.1%, <0.01%, <0.001%, <0.00001%, <0.000001%, <0.0000001%) of the significant

characteristic (e.g., luminescent intensity of a bioluminescent protein or bioluminescent complex).

As used herein, the term "bioluminescence" refers to production and emission of light by a chemical reaction catalyzed by, or enabled by, an enzyme, protein, protein complex, or other biomolecule (e.g., bioluminescent complex). In typical embodiments, a substrate for a bioluminescent entity (e.g., bioluminescent protein or bioluminescent complex) is converted into an unstable form by the bioluminescent entity; the substrate subsequently emits light.

As used herein the term "complementary" refers to the characteristic of two or more structural elements (e.g., peptide, polypeptide, nucleic acid, small molecule, etc.) of being able to hybridize, dimerize, or otherwise form a complex with each other. For example, a "complementary peptide and polypeptide" are capable of coming together to form a complex. Complementary elements may require assistance (facilitation) to form a complex (e.g., from interaction elements), for example, to place the elements in the proper conformation for complementarity, to place the elements in the proper proximity for complementarity, to co-localize complementary elements, to lower interaction energy for complementary, to overcome insufficient affinity for one another, etc.

As used herein, the term "complex" refers to an assemblage or aggregate of molecules (e.g., peptides, polypeptides, etc.) in direct and/or indirect contact with one another. In one aspect, "contact," or more particularly, "direct contact" means two or more molecules are close enough so that attractive noncovalent interactions, such as Van der Waal forces, hydrogen bonding, ionic and hydrophobic interactions, and the like, dominate the interaction of the molecules. In such an aspect, a complex of molecules (e.g., peptides and polypeptide) is formed under assay conditions such that the complex is thermodynamically favored (e.g., compared to a non-aggregated, or non-complexed, state of its component molecules). As used herein the term "complex," unless described as otherwise, refers to the assemblage of two or more molecules (e.g., peptides, polypeptides or a combination thereof).

As used herein, the term "non-luminescent" refers to an entity (e.g., peptide, polypeptide, complex, protein, etc.) that exhibits the characteristic of not emitting a detectable amount of light in the visible spectrum (e.g., in the presence of a substrate). For example, an entity may be referred to as non-luminescent if it does not exhibit detectable luminescence in a given assay. As used herein, the term "non-luminescent" is synonymous with the term "substantially non-luminescent. In some embodiments, an entity is considered

"non-luminescent" if any light emission is sufficiently minimal so as not to create interfering background for a particular assay.

As used herein, the terms "non-luminescent peptide" and "non-luminescent polypeptide" refer to peptides and polypeptides that exhibit substantially no luminescence (e.g., in the presence of a substrate), or an amount that is beneath the noise (e.g., 100-fold, 200-fold, 500-fold, $1x10^3$-fold, $1x10^4$-fold, $1x10^5$-fold, $1x10^6$-fold, $1x10^7$-fold, etc.) when compared to a significant signal (e.g., a bioluminescent complex) under standard conditions (e.g., physiological conditions, assay conditions, etc.) and with typical instrumentation (e.g., luminometer, etc.). In some embodiments, such non-luminescent peptides and polypeptides assemble, according to the criteria described herein, to form a bioluminescent complex.

As used herein, the term "interaction element" refers to a moiety that assists or facilitates the bringing together of non-luminescent elements to form a bioluminescent complex. In some embodiments, a pair of interaction elements (a.k.a. "interaction pair") is attached to a pair of non-luminescent elements (e.g., non-luminescent peptides), and the attractive interaction between the two interaction elements facilitates formation of the bioluminescent complex; although the present invention is not limited to such a mechanism, and an understanding of the mechanism is not required to practice the invention. Interaction elements may facilitate formation of the bioluminescent complex by any suitable mechanism (e.g., bringing non-luminescent elements into close proximity, placing a non-luminescent element in proper conformation for stable interaction, reducing activation energy for complex formation, combinations thereof, etc.). An interaction element may be a protein, polypeptide, peptide, small molecule, cofactor, nucleic acid, lipid, carbohydrate, antibody, etc. An interaction pair may be made of two of the same interaction elements (i.e., homopair) or two different interaction elements (i.e., heteropair). In the case of a heteropair, the interaction elements may be the same type of moiety (e.g., polypeptides) or may be two different types of moieties (e.g., polypeptide and small molecule). In some embodiments, in which complex formation by the interaction pair is studied, an interaction pair may be referred to as a "target pair" or a "pair of interest," and the individual interaction elements are referred to as "target elements" (e.g., "target peptide," "target polypeptide," etc.) or "elements of interest" (e.g., "peptide of interest," "polypeptide or interest," etc.).

As used herein, the term "low affinity" describes an intermolecular interaction between two or more (e.g., three) entities that is too weak to result in significant complex

formation between the entities, except at concentrations substantially higher (e.g., 2-fold, 5-fold, 10-fold, 100-fold, 1000-fold, or more) than physiologic or assay conditions, or with facilitation from the formation of a second complex of attached elements (e.g., interaction elements).

As used herein, the term "high affinity" describes an intermolecular interaction between two or more (e.g., three) entities that is of sufficient strength to produce detectable complex formation under physiologic or assay conditions, without facilitation from the formation of a second complex of attached elements (e.g., interaction elements).

As used herein, the term "co-localization element" refers to a moiety that facilitates co-localization of non-luminescent elements. In some embodiments, a set of non-luminescent elements has sufficient affinity to form a complex when the non-luminescent elements are co-localized at sufficient concentration. In such embodiments, a set (e.g., pair) of co-localization elements (a.k.a. "co-localization pair") is attached to a pair of non-luminescent elements (e.g., non-luminescent peptides), and the co-localization (e.g., within a cellular compartment, within a tissue, within a solution, on a solid matrix support, etc.) of the two co-localization elements facilitates co-localization of the non-luminescent elements, thereby facilitating formation of the bioluminescent complex; although the present invention is not limited to such a mechanism, and an understanding of the mechanism is not required to practice the invention. In some embodiments, due to the capacity of the non-luminescent elements to self-assemble into a luminescent complex, the co-localization elements need not directly interact to facilitate complex formation. A co-localization element may be a protein, polypeptide, peptide, small molecule, cofactor, nucleic acid, lipid, carbohydrate, antibody, etc. A co-localization pair may be made of two of the same co-localization elements (i.e., homopair) or two different co-localization elements (i.e., heteropair). In the case of a heteropair, the co-localization elements may be the same type of moiety (e.g., polypeptides) or may be two different types of moieties (e.g., polypeptide and small molecule). In some embodiments, in which the localization of the co-localization pair is studied, a co-localization pair may be referred to as a "target pair" or a "pair of interest," and the individual co-localization elements are referred to as "target elements" (e.g., "target peptide," "target polypeptide," etc.) or "elements of interest" (e.g., "peptide of interest," "polypeptide or interest," etc.).

As used herein, the term "coelenterazine" refers to naturally-occurring ("native") coelenterazine. As used herein, the term "coelenterazine analog" or "coelenterazine derivative" refers to synthetic (e.g., derivative or variant) and natural analogs thereof, including furimazine, coelenterazine-n, coelenterazine-f, coelenterazine-h, coelenterazine-hcp, coelenterazine-cp, coelenterazine-c, coelenterazine-e, coelenterazine-fcp, bis-deoxycoelenterazine ("coelenterazine-hh"), coelenterazine-i, coelenterazine-icp, coelenterazine-v, and 2-methyl coelenterazine, in addition to those disclosed in WO 2003/040100; U.S. Application Ser. No. 12/056,073 (paragraph [0086]); U.S. Pat. No. 8,669,103; WO 2012/061529, U.S. Pat. Pub. 2017/0233789 and U.S. Pat. Pub. 2018/0030059; the disclosures of which are incorporated by reference herein in their entireties. In some embodiments, coelenterazine analogs include pro-substrates such as, for example, those described in U.S. Application Ser. No. 12/056,073; U.S. Pub. No. 2012/0707849; U.S. Pub. No. 2014/0099654; herein incorporated by reference in their entireties.

As used herein, the term "preexisting protein" refers to an amino acid sequence that was in physical existence prior to a certain event or date. A "peptide that is not a fragment of a preexisting protein" is a short amino acid chain that is not a fragment or sub-sequence of a protein (e.g., synthetic or naturally-occurring) that was in physical existence prior to the design and/or synthesis of the peptide.

As used herein, the term "fragment" refers to a peptide or polypeptide that results from dissection or "fragmentation" of a larger whole entity (e.g., protein, polypeptide, enzyme, etc.), or a peptide or polypeptide prepared to have the same sequence as such. Therefore, a fragment is a subsequence of the whole entity (e.g., protein, polypeptide, enzyme, etc.) from which it is made and/or designed. A peptide or polypeptide that is not a subsequence of a preexisting whole protein is not a fragment (e.g., not a fragment of a preexisting protein). A peptide or polypeptide that is "not a fragment of a preexisting bioluminescent protein" is an amino acid chain that is not a subsequence of a protein (e.g., natural or synthetic) that: (1) was in physical existence prior to design and/or synthesis of the peptide or polypeptide, and (2) exhibits substantial bioluminescent activity.

As used herein, the term "subsequence" refers to peptide or polypeptide that has 100% sequence identify with a portion of another, larger peptide or polypeptide. The subsequence is a perfect sequence match for a portion of the larger amino acid chain.

The term "amino acid" refers to natural amino acids, unnatural amino acids, and amino acid analogs, all in their D and L stereoisomers, unless otherwise indicated, if their structures allow such stereoisomeric forms.

Natural amino acids include alanine (Ala or A), arginine (Arg or R), asparagine (Asn or N), aspartic acid (Asp or D), cysteine (Cys or C), glutamine (Gln or Q), glutamic acid (Glu or E), glycine (Gly or G), histidine (His or H), isoleucine (Ile or I), leucine (Leu or L), Lysine (Lys or K), methionine (Met or M), phenylalanine (Phe or F), proline (Pro or P), serine (Ser or S), threonine (Thr or T), tryptophan (Trp or W), tyrosine (Tyr or Y) and valine (Val or V).

Unnatural amino acids include, but are not limited to, pentafluorophenylalanine ("Z"), azetidinecarboxylic acid, 2-aminoadipic acid, 3-aminoadipic acid, beta-alanine, naphthylalanine ("naph"), aminopropionic acid, 2-aminobutyric acid, 4-aminobutyric acid, 6-aminocaproic acid, 2-aminoheptanoic acid, 2-aminoisobutyric acid, 3-aminoisbutyric acid, 2-aminopimelic acid, tertiary-butylglycine ("tBuG"), 2,4-diaminoisobutyric acid, desmosine, 2,2'-diaminopimelic acid, 2,3-diaminopropionic acid, N-ethylglycine, N-ethylasparagine, homoproline ("hPro" or "homoP"), hydroxylysine, allo-hydroxylysine, 3-hydroxyproline ("3Hyp"), 4-hydroxyproline ("4Hyp"), isodesmosine, allo-isoleucine, N-methylalanine ("MeAla" or "Nime"), N-alkylglycine ("NAG") including N-methylglycine, N-methylisoleucine, N-alkylpentylglycine ("NAPG") including N-methylpentylglycine. N-methylvaline, naphthylalanine, norvaline ("Norval"), norleucine ("Norleu"), octylglycine ("OctG"), ornithine ("Orn"), pentylglycine ("pG" or "PGly"), pipecolic acid, thioproline ("ThioP" or "tPro"), homoLysine ("hLys"), and homoArginine ("hArg"). Unnatural reactive amino acids are described in, for example, Boutureira, O. and G. J. Bernardes (2015) "Advances in chemical protein modification." Chem Rev 115(5): 2174-2195; herein incorporated by reference in its entirety.

The term "amino acid analog" refers to a natural or unnatural amino acid where one or more of the C-terminal carboxy group, the N-terminal amino group and side-chain bioactive group has been chemically blocked, reversibly or irreversibly, or otherwise modified to another bioactive group. For example, aspartic acid-(beta-methyl ester) is an amino acid analog of aspartic acid; N-ethylglycine is an amino acid analog of glycine; or alanine carboxamide is an amino acid analog of alanine. Other amino acid analogs include methionine sulfoxide, methionine sulfone, S-(carboxymethyl)-cysteine, S-(carboxymethyl)-cysteine sulfoxide and S-(carboxymethyl)-cysteine sulfone. Amino acid analogs may

comprise amino acids with various protecting groups (Isidro-Llobet, A., et al. (2009). "Amino Acid-Protecting Groups." Chemical Reviews 109(6): 2455-2504; herein incorporated by reference in its entirety).

As used herein, unless otherwise specified, the terms "peptide" and "polypeptide" refer to polymer compounds of two or more amino acids joined through the main chain by peptide amide bonds (--C(O)NH--). The term "peptide" typically refers to short amino acid polymers (e.g., chains having fewer than 30 amino acids), whereas the term "polypeptide" typically refers to longer amino acid polymers (e.g., chains having more than 30 amino acids).

As used herein, unless otherwise specified, the term "dipeptide" refers to a peptide or small polypeptide (e.g., <70 amino acids, <60 amino acids, <50 amino acids, etc.) comprising two peptide segments (e.g., corresponding to two beta strands of a luciferase (e.g., a "β9/β10 dipeptide," corresponding to the β9 and β10 strands of an OgLuc luciferase polypeptide), fused/attached diretly or indirectly (e.g., via a linker (e.g., peptide linker (e.g., 1-10 amino acids (e.g., a single glycine)))).

As used herein, unless otherwise specified, the term "tripeptide" refers to a peptide or small polypeptide (e.g., <100 amino acids, <90 amino acids, <80 amino acids, etc.) comprising three peptide segments (e.g., corresponding to three beta strands of a luciferase (e.g., a "β8-10 tripeptide," corresponding to the β8-10 strands of an OgLuc luciferase polypeptide), fused/attached diretly or indirectly (e.g., via a linker (e.g., peptide linker (e.g., 1-10 amino acids (e.g., a single glycine)))).

As used herein, terms "peptidomimetic" and "peptide mimetic" refer to peptide-like or polypeptide-like molecules that emulate a sequence derived from a protein or peptide. A peptidomimetic may contain amino acids analogs, peptoid amino acids, and/or non-amino acid components either exclusively or in combination with amino acids (e.g., natural or non-natural amino acids). Examples of peptidomimitecs include chemically modified peptides/polypeptides, peptoids (side chains are appended to the nitrogen atom of the peptide backbone rather than to the α-carbons), β-peptides (amino group bonded to the β carbon rather than the α carbon), etc.

As used herein, the term "peptoid" refers to a class of peptidomimetics where the side chains are functionalized on the nitrogen atom of the peptide backbone rather than to the α-carbon.

As used herein, the term "artificial" refers to compositions and systems that are designed or prepared by man and are not naturally occurring. For example, an artificial peptide, peptoid, or nucleic acid is one comprising a non-natural sequence (e.g., a peptide without 100% identity with a naturally-occurring protein or a fragment thereof).

As used herein, a "conservative" amino acid substitution refers to the substitution of an amino acid in a peptide or polypeptide with another amino acid having similar chemical properties such as size or charge. For purposes of the present disclosure, each of the following eight groups contains amino acids that are conservative substitutions for one another:

1) Alanine (A) and Glycine (G);

2) Aspartic acid (D) and Glutamic acid (E);

3) Asparagine (N) and Glutamine (Q);

4) Arginine (R) and Lysine (K);

5) Isoleucine (I), Leucine (L), Methionine (M), and Valine (V);

6) Phenylalanine (F), Tyrosine (Y), and Tryptophan (W);

7) Serine (S) and Threonine (T); and

8) Cysteine (C) and Methionine (M).

Naturally occurring residues may be divided into classes based on common side chain properties, for example: polar positive (or basic) (histidine (H), lysine (K), and arginine (R)); polar negative (or acidic) (aspartic acid (D), glutamic acid (E)); polar neutral (serine (S), threonine (T), asparagine (N), glutamine (Q)); non-polar aliphatic (alanine (A), valine (V), leucine (L), isoleucine (I), methionine (M)); non-polar aromatic (phenylalanine (F), tyrosine (Y), tryptophan (W)); proline and glycine; and cysteine. As used herein, a "semi-conservative" amino acid substitution refers to the substitution of an amino acid in a peptide or polypeptide with another amino acid within the same class.

In some embodiments, unless otherwise specified, a conservative or semi-conservative amino acid substitution may also encompass non-naturally occurring amino acid residues that have similar chemical properties to the natural residue. These non-natural residues are typically incorporated by chemical peptide synthesis rather than by synthesis in biological systems. These include, but are not limited to, peptidomimetics and other reversed or inverted forms of amino acid moieties. Embodiments herein may, in some embodiments, be limited to natural amino acids, non-natural amino acids, and/or amino acid analogs.

Non-conservative substitutions may involve the exchange of a member of one class for a member from another class.

As used herein, the term "sequence identity" refers to the degree two polymer sequences (e.g., peptide, polypeptide, nucleic acid, etc.) have the same sequential composition of monomer subunits. The term “sequence similarity” refers to the degree with which two polymer sequences (e.g., peptide, polypeptide, nucleic acid, etc.) have similar polymer sequences. For example, similar amino acids are those that share the same biophysical characteristics and can be grouped into the families, e.g., acidic (e.g., aspartate, glutamate), basic (e.g., lysine, arginine, histidine), non-polar (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan) and uncharged polar (e.g., glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine). The “percent sequence identity” (or “percent sequence similarity”) is calculated by: (1) comparing two optimally aligned sequences over a window of comparison (e.g., the length of the longer sequence, the length of the shorter sequence, a specified window), (2) determining the number of positions containing identical (or similar) monomers (e.g., same amino acids occurs in both sequences, similar amino acid occurs in both sequences) to yield the number of matched positions, (3) dividing the number of matched positions by the total number of positions in the comparison window (e.g., the length of the longer sequence, the length of the shorter sequence, a specified window), and (4) multiplying the result by 100 to yield the percent sequence identity or percent sequence similarity. For example, if peptides A and B are both 20 amino acids in length and have identical amino acids at all but 1 position, then peptide A and peptide B have 95% sequence identity. If the amino acids at the non-identical position shared the same biophysical characteristics (e.g., both were acidic), then peptide A and peptide B would have 100% sequence similarity. As another example, if peptide C is 20 amino acids in length and peptide D is 15 amino acids in length, and 14 out of 15 amino acids in peptide D are identical to those of a portion of peptide C, then peptides C and D have 70% sequence identity, but peptide D has 93.3% sequence identity to an optimal comparison window of peptide C. For the purpose of calculating “percent sequence identity” (or “percent sequence similarity”) herein, any gaps in aligned sequences are treated as mismatches at that position.

Any peptide/polypeptides described herein as having a particular percent sequence identity or similarity (e.g., at least 70%) with a reference sequence ID number, may also be expressed as having a maximum number of substitutions (or terminal deletions) with respect

to that reference sequence. For example, a sequence having at least Y% sequence identity (e.g., 90%) with SEQ ID NO:Z (e.g., 100 amino acids) may have up to X substitutions (e.g., 10) relative to SEQ ID NO:Z, and may therefore also be expressed as "having X (e.g., 10) or fewer substitutions relative to SEQ ID NO:Z."

As used herein, the term "wild-type," refers to a gene or gene product (e.g., protein, polypeptide, peptide, etc.) that has the characteristics (e.g., sequence) of that gene or gene product isolated from a naturally occurring source, and is most frequently observed in a population. In contrast, the term "mutant" or "variant" refers to a gene or gene product that displays modifications in sequence when compared to the wild-type gene or gene product. It is noted that "naturally-occurring variants" are genes or gene products that occur in nature, but have altered sequences when compared to the wild-type gene or gene product; they are not the most commonly occurring sequence. "Artificial variants" are genes or gene products that have altered sequences when compared to the wild-type gene or gene product and do not occur in nature. Variant genes or gene products may be naturally occurring sequences that are present in nature, but not the most common variant of the gene or gene product, or "synthetic," produced by human or experimental intervention.

As used herein, the term "physiological conditions" encompasses any conditions compatible with living cells, e.g., predominantly aqueous conditions of a temperature, pH, salinity, chemical makeup, etc. that are compatible with living cells.

As used herein, the term "sample" is used in its broadest sense. In one sense, it is meant to include a specimen or culture obtained from any source, as well as biological and environmental samples. Biological samples may be obtained from animals (including humans) and encompass fluids, solids, tissues, and gases. Biological samples include blood products, such as plasma, serum, and the like. Sample may also refer to cell lysates or purified forms of the enzymes, peptides, and/or polypeptides described herein. Cell lysates may include cells that have been lysed with a lysing agent or lysates such as rabbit reticulocyte or wheat germ lysates. Sample may also include cell-free expression systems. Environmental samples include environmental material such as surface matter, soil, water, crystals, and industrial samples. Such examples are not however to be construed as limiting the sample types applicable to the present invention.

As used herein, the terms "fusion," "fusion polypeptide," and "fusion protein" refer to a chimeric protein containing a first protein or polypeptide of interest (e.g., substantially non-

luminescent peptide) joined to a second different peptide, polypeptide, or protein (e.g., interaction element).

As used herein, the terms "conjugated" and "conjugation" refer to the covalent attachment of two molecular entities (e.g., post-synthesis and/or during synthetic production). The attachment of a peptide or small molecule tag to a protein or small molecule, chemically (e.g., "chemically" conjugated) or enzymatically, is an example of conjugation.

The term "binding moiety" refers to a domain that specifically binds an antigen or epitope independently of a different epitope or antigen binding domain. A binding moiety may be an antibody, antibody fragment, a receptor domain that binds a target ligand, proteins that bind to immunoglobulins (e.g., protein A, protein G, protein A/G, protein L, protein M), a binding domain of a proteins that bind to immunoglobulins (e.g., protein A, protein G, protein A/G, protein L, protein M), oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins etc. Table A provides a lists of exemplary binding moieties that could be used singly or in various combinations in methods, systems, and assays (e.g., immunoassays) herein.

Table A. Exemplary binding moieties

Binding Moiety A	Binding Moiety B
Protein A	Protein A
Ig Binding domain of protein A	Ig binding domain of protein A
Protein G	Protein G
Ig Binding domain of protein G	Ig binding domain of protein G
Protein L	Protein L
Ig Binding domain of protein L	Ig binding domain of protein L
Protein M	Protein M
Ig Binding domain of protein M	Ig binding domain of protein M
polyclonal antibody against analyte X	polyclonal antibody: same antibody or second polyclonal antibody recognizing same target analyte X
monoclonal antibody	monoclonal antibody recognizing different epitope on same target analyte X
recombinant antibody	recombinant antibody recognizing different epitope on same target analyte X
scFv	scFv recognizing different epitope on same target analyte X
variable light chain (V_L) of antibody (monoclonal, recombinant, polyclonal) recognizing target analyte X	variable heavy chain (V_H) of same antibody (monoclonal, recombinant, polyclonal) recognizing target analyte X
protein (e.g. receptor) binding domain 1 that binds to analyte X	protein (e.g. receptor) binding domain 2 that binds to analyte X

(Fab) fragment	(Fab) fragment from different antibody recognizing different epitope to same target analyte X
Fab' fragment	Fab' from different antibody recognizing different epitope to same target analyte X
Fv fragment	Fv from different antibody recognizing different epitope to same target analyte X
F(ab')2 fragment	F(ab')2 from different antibody recognizing different epitope to same target analyte X
oligonucleotide probe	oligonucleotide probe to same DNA or RNA target but recognizing non-overlapping sequence
DARPin	DARPin recognizing non-overlapping domain of same target
peptide nucleic acid	peptide nucleic acid recognizing same DNA or RNA target but non-overlapping sequence
aptamer	aptamer binding to same target analyte X but recognizing non-overlapping sequence or epitope
affimer	aptamer binding to same target analyte X but recognizing different epitope

As used herein, the term "antibody" refers to a whole antibody molecule or a fragment thereof (e.g., fragments such as Fab, Fab', and $F(ab')_2$, variable light chain, variable heavy chain, Fv, it may be a polyclonal or monoclonal or recombinant antibody, a chimeric antibody, a humanized antibody, a human antibody, etc. As used herein, when an antibody or other entity "specifically recognizes" or "specifically binds" an antigen or epitope, it preferentially recognizes the antigen in a complex mixture of proteins and/or macromolecules, and binds the antigen or epitope with affinity which is substantially higher than to other entities not displaying the antigen or epitope. In this regard, "affinity which is substantially higher" means affinity that is high enough to enable detection of an antigen or epitope which is distinguished from entities using a desired assay or measurement apparatus. Typically, it means binding affinity having a binding constant (K_a) of at least 10^7 M^{-1} (e.g., $>10^7$ M^{-1}, $>10^8$ M^{-1}, $>10^9$ M^{-1}, $>10^{10}$ M^{-1}, $>10^{11}$ M^{-1}, $>10^{12}$ M^{-1}, $>10^{13}$ M^{-1}, etc.). In certain such embodiments, an antibody is capable of binding different antigens so long as the different antigens comprise that particular epitope. In certain instances, for example, homologous proteins from different species may comprise the same epitope.

As used herein, the term "antibody fragment" refers to a portion of a full-length antibody, including at least a portion of the antigen binding region or a variable region. Antibody fragments include, but are not limited to, Fab, Fab', $F(ab')_2$, Fv, scFv, Fd, variable light chain, variable heavy chain, diabodies, and other antibody fragments that retain at least a portion of the variable region of an intact antibody. See, e.g., Hudson et al. (2003) Nat.

Med. 9:129-134; herein incorporated by reference in its entirety. In certain embodiments, antibody fragments are produced by enzymatic or chemical cleavage of intact antibodies (e.g., papain digestion and pepsin digestion of antibody) produced by recombinant DNA techniques, or chemical polypeptide synthesis. For example, a "Fab" fragment comprises one light chain and the C_{H1} and variable region of one heavy chain. The heavy chain of a Fab molecule cannot form a disulfide bond with another heavy chain molecule. A "Fab'" fragment comprises one light chain and one heavy chain that comprises additional constant region, extending between the C_{H1} and C_{H2} domains. An interchain disulfide bond can be formed between two heavy chains of a Fab' fragment to form a "$F(ab')_2$" molecule. An "Fv" fragment comprises the variable regions from both the heavy and light chains, but lacks the constant regions. A single-chain Fv (scFv) fragment comprises heavy and light chain variable regions connected by a flexible linker to form a single polypeptide chain with an antigen-binding region. Exemplary single chain antibodies are discussed in detail in WO 88/01649 and U.S. Pat. Nos. 4,946,778 and 5,260,203; herein incorporated by reference in their entireties. In certain instances, a single variable region (e.g., a heavy chain variable region or a light chain variable region) may have the ability to recognize and bind antigen. Other antibody fragments will be understood by skilled artisans.

As used herein, the term "peptide tag" refers to a peptide that may be attached (e.g., post-synthesis or during synthetic production) or fused to another entity (e.g., protein of interest, molecule of interest, interaction element, co-localization element, etc.). The peptide tag may or may not be attached to another entity. Typically, as used herein, a peptide tag is capable of forming a bioluminescent complex with another peptide tag and a polypeptide under appropriate conditions. In embodiments in which a peptide tag is attached to another entity, a peptide tag is chemically conjugated to another molecule (e.g., peptide, polypeptide, nucleic acid, other small molecules or macromolecules), chemically synthesized to be a part of another molecule, or genetically fused to another peptide or polypeptide molecule, etc.

As used herein, the term "polypeptide component" is used synonymously with the term "polypeptide component of a bioluminescent complex." Typically, as used herein, a polypeptide component is capable of forming a bioluminescent complex with a pair of peptide tags, under appropriate conditions.

As used herein, the term "an Oplophorus luciferase" ("an OgLuc") refers to a luminescent polypeptide having significant sequence identity, structural conservation, and/or

the functional activity of the luciferase produce by and derived from the deep-sea shrimp *Oplophorus gracilirostris*. In particular, an OgLuc polypeptide refers to a luminescent polypeptide having significant sequence identity, structural conservation, and/or the functional activity of the mature 19 kDa subunit of the Oplophorus luciferase protein complex (e.g., without a signal sequence) such as SEQ ID NOs: 1 (WT OgLuc) and 3 (NanoLuc), which comprises 10 β strands (β1, β2, β3, β4, β5, β6, β7, β8, β9, β10) and utilize substrates such as coelenterazine or coelenterazine derivatives to produce luminescence.

As used herein, the term "β9-like peptide" refers to a peptide (or peptide tag) comprising significant sequence identity, structural conservation, and/or the functional activity of the β (beta) 9 strand of an OgLuc polypeptide. In particular, a β9-like peptide is a peptide capable of structurally complementing an OgLuc polypeptide lacking a β9 strand resulting in enhanced luminescence of the complex compared to the OgLuc polypeptide in the absence of the β9-like peptide. Other "βX-like peptides" may be similarly named (e.g., β1-like, β2-like, β3-like, β4-like, β5-like, β6-like, β7-like, β8-like, β9-like).

As used herein, the term "β10-like peptide" refers to a peptide (or peptide tag) comprising significant sequence identity, structural conservation, and/or the functional activity of the β (beta) 10 strand of an OgLuc polypeptide. In particular, a β10-like peptide is a peptide capable of structurally complementing an OgLuc polypeptide lacking a β10 strand resulting in enhanced luminescence of the complex compared to the OgLuc polypeptide in the absence of the β10-like peptide. Other "βX-like peptides" may be similarly named (e.g., β1-like, β2-like, β3-like, β4-like, β5-like, β6-like, β7-like, β8-like, β9-like).

As used herein, the term "β_{1-8}-like polypeptide" refers to a polypeptide bearing sequence and structural similarity to β (beta) strands 1-8 of an OgLuc polypeptide, but lacking β (beta) strands 9 and 10. Other "β_{Y-Z}-like polypeptides" may be similarly named (e.g., β_{1-4}-like polypeptide, β_{2-8}-like polypeptide, β_{5-10}-like polypeptide, etc.).

As used herein, the term "NANOLUC" refers to an artificial luciferase or bioluminescent polypeptide produced commercially by the Promega Corporation and corresponding to SEQ ID NO: 3.

As used herein, the term "LgBiT" refers to a polypeptide corresponding to β_{1-9}-like polypeptide that finds use in, for example, binary complementation to form a bioluminescent complex and corresponds to SEQ ID NO: 11.

As used herein, the term "SmBiT" refers to a peptide corresponding to β_{10}-like peptide that finds use in, for example, binary complementation to form a bioluminescent complex, but has low affinity for LgBiT (e.g., requires facilitation for complex formation) and corresponds to SEQ ID NO: 13.

As used herein, the term "HiBiT" refers to a peptide corresponding to β_{10}-like peptide that finds use in, for example, binary complementation to form a bioluminescent complex, but has low affinity for LgBiT (e.g., requires facilitation for complex formation) and corresponds to SEQ ID NO: 15. HiBiT is has the same sequence as "SmHiTrip10" (SEQ ID NO: 25) and "pep86," terms which may be used interchangeably (also SmTrip10 pep86, etc.).

As used herein, the term "LgTrip" refers to a polypeptide corresponding to β_{1-8}-like polypeptide that corresponds to SEQ ID NO: 17 and finds use in, for example, tripartite complementation with β_9-like and β_{10}-like peptides to form a bioluminescent complex, or bianry complementation, with a β_{9-10}-like dipeptide to form a bioluminescent complex. LgTrip variants include: LgTrip 2098 (w/ His tag: SEQ ID NO: 31; w/o His tag: SEQ ID NO: 304) and LgTrip 3546 (w/ His tag: SEQ ID NO: 51; w/o His tag: SEQ ID NO: 302).

As used herein, the term "SmTrip10" refers to a peptide corresponding to β_{10}-like peptide that finds use in, for example, tripartite complementation to form a bioluminescent complex.

As used herein, the term "SmTrip9" refers to a peptide corresponding to β_9-like peptide that finds use in, for example, tripartite complementation to form a bioluminescent complex.

DETAILED DESCRIPTION

Provided herein are bioluminescent polypeptides and compositions and methods for the assembly of a tripartite or multipartite bioluminescent complex. In particular embodiments, a bioluminescent complex is formed upon the interaction of three or more peptide and/or polypeptide components.

Experiments conducted during development of embodiments herein demonstrate that a tripartite luciferase comprising two small peptide elements (e.g., a β10-like peptide and β9-like peptide) and one polypeptide element (e.g., β_{1-8}-like polypeptide) assemble to form a luminescent complex. Experiments conducted during development of embodiments herein further demonstrate the formation of a bioluminescent complex from up to five fragments of a luciferase (or variants of such fragments), such as a polypeptide fragment (or variants

thereof) and one or more peptide, dipeptide, or tripeptide fragments (or variants of such fragments).

The commercially-available NANOLUC luciferase (Promega Corporation) comprises 10 β (beta) strands (β1, β2, β3, β4, β5, β6, β7, β8, β9, β10). U.S. Pat. No. 9,797,889 (herein incorporated by reference in its entirety) describes development and use of a complementation system comprising a β_{1-9}-like polypeptide and a β_{10}-like peptide (the actual polypeptide and peptide sequences in U.S. Pat. No. 9,797,889 differ from the corresponding sequences in NANOLUC and wild-type native OgLuc).

In experiments conducted during development of embodiments herein, a β_{1-9}-like polypeptide was further split by removal of the β9 strand. The remaining portion (a β_{1-8}-like polypeptide) is referred to herein as LgTrip 2098 (SEQ ID NO: 17; or SEQ ID NO: 31 (with His tag)). Experiments attempted to reconstitute a luminescent complex from LgTrip and two peptides corresponding to the β9 (SmTrip9 pep245; SEQ ID NO: 23) and β10 (SmTrip10 pep86; HiBit, a β10 sequence optimized for use in a high affinity bipartite system; SEQ ID NO: 15) strands. Experiments demonstrated that LgTrip 2098 (SEQ ID NO: 17; or SEQ ID NO: 31 (with His tag)) expressed poorly in *E. coli*, was unstable, and was susceptible to surface inactivation. Experiments were conducted during development of embodiments herein to develop artificial variants that exhibit one or more (e.g., all) of enhanced stability, enhanced expression, enhanced activity, enhanced molecular interactions, etc., and is capable of being used in a system to reconstitute a bioluminescent complex with peptides corresponding to the β9 (e.g., β9-like peptides (e.g., SmTrip9 pep245; SEQ ID NO: 23)) and β10 (e.g., β10-like peptides (e.g., SmTrip10 pep86; HiBiT; SEQ ID NO: 25)) strands. Experiments conducted during development of embodiments herein demonstrate, for example, that LgTrip 3092 (SEQ ID NO: 19) or LgTrip 3546 (SEQ ID NO: 51) are capable of forming a luminescent complex with suitable β9-like (e.g., SmTrip9 pep245; SEQ ID NO: 23) and β10-like (e.g., SmTrip10 pep86; HiBiT; SEQ ID NO: 25) peptides. Experiments were conducted during development of embodiments herein to develop artificial polypeptide components (e.g., SEQ ID NOs: 19, 21, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, and additional variants thereof) and peptide tags (e.g., the peptides listed in Table 1 and additional variants thereof) with enhanced characteristics for luminescent complex reconstitution.

Further experiments conducted during development of embodiments herein demonstrate that NANOLUC-based bioluminescent complexes can be formed using constructs comprising other polypeptide components (e.g., β_{1-5}-like, β_{1-6}-like, β_{1-7}-like, etc.) and corresponding combinations of complimentary peptides (e.g., β_6-like, β_7-like, β_8-like, β_9-like, β_{10}-like), dipeptides (e.g., β_{6-7}-like, β_{7-8}-like, β_{8-9}-like, β_{9-10}-like), tripeptides (e.g., β_{6-8}-like, β_{7-9}-like, β_{8-10}-like), polypeptides (e.g., β_{6-10}-like, β_{6-9}-like, β_{7-10}-like, etc.) derived from the NANOLUC-based, NanoBiT-based, and NanoTrip-based systems, polypeptides, and peptide described herein. The experiments conducted during development of embodiments herein demonstrate the formation of a bioluminescent complex from two or more (e.g., 2, 3, 4, 5, etc.) peptide and polypeptide components that collectively comprise the full lenth of a luciferase contruct (e.g., a full lngth luciferase polypeptide comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, or ranges therebetween) sequence identity with SEQ ID NO: 788 or 789).

In some embodiments, provided herein are compositions and methods for the assembly of a bioluminescent complex from two peptide tags (e.g., β9-like (e.g., SmTrip9) and β10-like (e.g., SmTrip10) peptides) and a polypeptide component (e.g., β_{1-8}-like (e.g., LgTrip) polypeptide).

In some embodiments, provided herein are compositions and methods for the assembly of a bioluminescent complex from a polypeptide component (e.g., a β_{1-5}-like, β_{1-6}-like , β_{1-7}-like , or β_{1-8}-like polypeptide), and complementary peptide(s) (e.g., β_6-like, β_7-like, β_8-like, β_9-like, β_{10}-like), dipeptide(s) (e.g., β_{6-7}-like, β_{7-8}-like, β_{8-9}-like, β_{9-10}-like), tripeptide (e.g., β_{6-8}-like, β_{7-9}-like, β_{8-10}-like), and/or polypeptides (e.g., β_{6-10}-like, β_{6-9}-like, β_{7-10}-like, etc.).

In some embodiments, one or more (e.g., two, three, four, five, etc.) of the peptide tags and the polypeptide component are not fragments of a preexisting protein (e.g., are not structurally-complementary subsequences of a known polypeptide sequence). However, in other embodiments, one or more of the peptide tags and the polypeptide component may be fragments of a known or existing protein, polypeptide, or peptide. In certain embodiments, the bioluminescent activity of the polypeptide component (of the bioluminescent complex) is enhanced (e.g., 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 10^2-fold, 10^3-fold, 10^4-fold, 10^5-fold, 10^6-fold, or more) via structural complementation with the two peptide tags. In some embodiments, provided herein are peptide (peptide tags)/polypeptide elements that are

capable of assembling into a bioluminescent complex for the purpose of, for example, detecting and monitoring molecular interactions (e.g., protein-protein, protein-DNA, protein-RNA interactions, RNA-DNA, protein-small molecule, RNA-small-molecule, DNA-DNA, RNA-RNA, PNA-DNA, PNA-RNA, etc.). In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptide thereof) are fused, or otherwise linked to interaction elements. In particular embodiments, the peptide/dipeptide/tripeptide tags and polypeptide components, when for the purpose of detecting/monitoring molecular interactions, do not form a complete bioluminescent complex without facilitation by the interaction between interaction elements. However, upon interaction (e.g., binding) of the interaction elements to each other (or to a target molecule or complex), formation of the bioluminescent complex is facilitated. In some embodiments, the bioluminescent signal from the bioluminescent complex (or the capacity to produce such a signal in the presence of substrate) serves as a reporter for the formation of a complex by the interaction elements. If an interaction complex is formed, then a bioluminescent complex is formed, and a bioluminescent signal is detected/measured/monitored (e.g., in the presence of substrate). If an interaction complex fails to form (e.g., due to unfavorable conditions, due to unstable interaction between the interaction elements, due to incompatible interaction elements), then a bioluminescent complex does not form, and a bioluminescent signal is not produced (e.g., in the presence of substrate). In some embodiments, the bioluminescent signal from the bioluminescent complex (or the capacity to produce such a signal in the presence of substrate) serves as a reporter for the binding of the interaction elements to a target. If target-binding occurs, then a bioluminescent complex is formed and a bioluminescent signal is detected/measured/monitored (e.g., in the presence of substrate). If target-binding fails to occur (e.g., due to unfavorable conditions, due to unstable interaction between an interaction element and target, due to the absence of target, etc.), then a bioluminescent complex does not form and a bioluminescent signal is not produced.

In certain embodiments, interaction elements are two molecules of interest (e.g., protein(s) of interest, small molecule(s) of interest, etc.). For example, assays can be performed to detect the interaction of two molecules of interest by tethering each one to separate peptide/dipeptide/tripeptide tag (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides

thereof). If the molecules of interest interact (e.g., transiently interact, stably interact, etc.), the peptide/dipeptide/tripeptide tags are brought into close proximity in a suitable conformation, and a bioluminescent complex is formed between the peptide/dipeptide/tripeptide tags and the polypeptide component of the bioluminescent complex (and bioluminescent signal is produced/detected (e.g., in the presence of substrate)). In the absence of an interaction between the molecules of interest, the peptide/dipeptide/tripeptide tags are not brought into close proximity and/or arranged in an orientation to facilitate complex formation with the polypeptide component of the bioluminescent complex, the bioluminescent complex is not formed, and a bioluminescent signal is not produced (in the presence of substrate). Such embodiments can be used to study the effect of inhibitors on complex formation, the effect of mutations on complex formation, the effect of conditions (e.g., temperature, pH, etc.) on complex formation, the interaction of a small molecule (e.g., potential therapeutic) with a target molecule, etc.

In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptide thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) are provided that are capable of assembling into a bioluminescent complex without facilitation by interaction elements. In such embodiments, a bioluminescent complex will form when the peptide/dipeptide/tripeptide tags and polypeptide component are together within the same sample, subcellular compartment, cell, tissue, etc. (e.g., co-localized). In some embodiments, provided herein peptide/dipeptide/tripeptide (tags)/polypeptide elements that are capable of assembling into a bioluminescent complex for use in detecting and monitoring co-localization (e.g., without molecular interaction) of molecular elements (e.g., protein(s), nucleic acid(s), small molecule(s), lipid, carbohydrate, cellular structure, etc.). In some embodiments, a bioluminescent complex is formed from peptide/dipeptide/tripeptide tags and a polypeptide component that collectively span a full β1-like to β10-like sequence. In some embodiments, the peptide/dipeptide/tripeptide tags are fused or otherwise linked to co-localization elements. In particular embodiments, particularly for the purpose of detecting/monitoring co-localization (e.g., without molecular interaction), the peptide/dipeptide/tripeptide tags and polypeptide components are capable of forming a bioluminescent complex without facilitation (e.g., without interaction elements). Upon co-localization (e.g., within the same cell, on the same surface, with the same cellular

compartment, within the same tissue, etc.) of the co-localization elements (e.g., fused to the peptide/dipeptide/tripeptide tags), formation of the bioluminescent complex (from the peptide/dipeptide/tripeptide tags and the polypeptide component), with or without interaction of the co-localization elements, is facilitated. In some embodiments, the bioluminescent signal from the bioluminescent complex (or the capacity to produce such a signal in the presence of substrate) serves as a reporter for co-localization of the co-localization elements. If the co-localization elements co-localize, then a bioluminescent complex of the polypeptide component and the peptide/dipeptide/tripeptide tags fused to the co-localization elements is formed, and a bioluminescent signal is detected/measured/monitored (e.g., in the presence of substrate). If the co-localization elements do not co-localize, then a bioluminescent complex does not form, and a bioluminescent signal is not produced (e.g., in the presence of substrate).

In certain embodiments, the co-localization pair comprises two molecules of interest (e.g., protein(s) of interest, small molecule(s) of interest, etc.). For example, assays can be performed to detect the co-localization (e.g., within a cell, within a cellular compartment, within a tissue, etc.) of two molecules of interest by tethering each one to a separate dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof). If the molecules of interest co-localize, the peptide tags are brought into close proximity in a suitable conformation, and a bioluminescent complex is formed with the polypeptide component polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide), and bioluminescent signal is produced/detected (e.g., in the presence of substrate). In the absence of co-localization of the molecules of interest, the polypeptide components and peptide/dipeptide/tripeptide tags tags do not interact to form a complex, and a bioluminescent signal is not produced (e.g., in the presence of substrate). Such embodiments can be used to study co-localization of molecules of interest under various conditions.

In some embodiments, systems, assays, and devices comprising dipeptide/tripeptide tags tags e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) are provided for the detection of an analyte (e.g., small molecule, peptide, protein, antibody, nucleic acid, etc.) in a sample. In some embodiments, peptide/dipeptide/tripeptide tags are tethered or fused

with detection/binding agents (e.g., binding moiety, binding sequence, etc.) that recognize the target analyte, target analytes, secondary analytes that are bound by the target analyte, secondary binding agents that bind to primary binding agents, etc. In some embodiments, various combinations of peptide/dipeptide/tripeptide tags tethered/fused to the aforementioned detection/binding agents are used in assays and devices for the detection/quantification/identification of analytes in a sample. Exemplary systems that find use in assays and devices are depicted in, for example, Figures 51-56 and described herein.

In some embodiments, provided herein are compositions and methods for the assembly of a bioluminescent complex from a dipeptide (e.g., a β9/β10-like dipeptide) and a polypeptide component (e.g., β_{1-8}-like (e.g., LgTrip) polypeptide). In some embodiments, the dipeptide and the polypeptide component are not fragments of a preexisting protein (e.g., are not structurally-complementary subsequences of a known polypeptide sequence). However, in other embodiments, the dipeptide and/or the polypeptide component may be fragments of a known or existing protein, polypeptide, or peptide. In certain embodiments, the bioluminescent activity of the polypeptide component (of the bioluminescent complex) is enhanced (e.g., 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 10^2-fold, 10^3-fold, 10^4-fold, 10^5-fold, 10^6-fold, or more) via structural complementation with the dipeptide. In some embodiments, a β_{1-8}-like polypeptide exhibits lower background luminescence than a β_{1-9}-like polypeptide. In some embodiments, a β_{1-8}-like polypeptide exhibits increased thermal and chemical stability compared to a β_{1-9}-like polypeptide.

In some embodiments, provided herein are bioluminescent complexes, including but not limited to those comprising any of the following combinations of peptide, dipeptide, tripeptide, and polypeptide components:

- β1-5-like polypeptide + β6-like peptide + β7-like peptide + β8-like peptide + β9-like peptide + β10-like peptide;
- β1-5-like polypeptide + β6-like peptide + β7-like peptide + β8-like peptide + β9/10-like dipeptide;
- β1-5-like polypeptide + β6-like peptide + β7/8-like dipeptide + β9/10-like dipeptide;
- β1-5-like polypeptide + β6/7/8-like tripeptide + β9/10-like dipeptide;
- β1-5-like polypeptide + β6-like peptide + β7/8/9-like tripeptide + β10-like peptide;
- β1-6-like polypeptide + β7-like peptide + β8-like peptide + β9-like peptide + β10-like peptide;

- β1-6-like polypeptide + β7-like peptide + β8-like peptide + β9/10-like dipeptide;
- β1-6-like polypeptide + β7/8-like dipeptide + β9/10-like dipeptide;
- β1-6-like polypeptide + β6/7/8-like dipeptide + β9-like peptide + β10-like peptide;
- β1-6-like polypeptide + β7/8/9-like tripeptide + β10-like peptide;
- β1-7-like polypeptide + β8-like peptide + β9-like peptide + β10-like peptide;
- β1-7-like polypeptide + β8-like peptide + β9/10-like dipeptide;
- β1-7-like polypeptide + 8/9-like dipeptide + β10-like peptide;
- β1-7-like polypeptide + β8/9/10-like tripeptide;
- β1-8-like polypeptide + β9-like peptide + β10-like peptide;
- β1-8-like polypeptide + β9/10-like dipeptide;
- β1-5-like polypeptide + β6-10-like polypeptide;
- β1-5-like polypeptide + β6-9-like polypeptide + β10-like peptide; and
- β1-5-like polypeptide + β7-10-like polypeptide + β6-like peptide.

The above combinations are not limiting and other combinations of peptide, dipeptide, tripeptide, and polypeptide components are within the scope herein.

In some embodiments, a β1-5-like polypeptide comprises positions 1-102 of SEQ ID NO: 788. In some embodiments, a β1-6-like polypeptide comprises positions 1-124 of SEQ ID NO: 788. In some embodiments, a β1-7-like polypeptide comprises positions 1-133 of SEQ ID NO: 788. In some embodiments, a β1-8-like polypeptide comprises positions 1-148 of SEQ ID NO: 788.

In some embodiments, a set of β5-10-like peptides/dipeptide/tripeptides/polypeptide collectively comprise positions 103-170 of SEQ ID NO: 788 or 789. In some embodiments, a set of β6-10-like peptides/dipeptide/tripeptides/polypeptide collectively comprise positions 125-170 of SEQ ID NO: 788 or 789. In some embodiments, a set of β7-10-like peptides/dipeptide/tripeptides/polypeptide collectively comprise positions 134-170 of SEQ ID NO: 788 or 789. In some embodiments, a set of β8-10-like peptides/dipeptide/tripeptides/polypeptide collectively comprise positions 149-170 of SEQ ID NO: 788 or 789.

In some embodiments, one or more components of a bioluminescent complex span partial beta strands of the base luciferases (e.g., OgLuc, NANOLUC, SEQ ID NO: 788, SEQ ID NO: 789, etc.) described herein. The separations between peptide, dipeptide, tripeptide, and polypeptide components may reside at the split points between the beta strands or may appear at a position -1, -2, -3, -4, -5, +1, +2, +3, +4, +5, or more from the split points

identified by the sequences herein. In some embodiments, peptide, dipeptide, tripeptide, and polypeptide components that span the full sequence of a base luciferases (e.g., OgLuc, NANOLUC, SEQ ID NO: 788, SEQ ID NO: 789, etc.) described herein are capable of forming a bioluminescent complex, even if the split points for the components are not between the beta strands.

For example, a split site between β5 and β6 may occur between positions 102 and 103 of SEQ ID NO: 788, or in some embodiments such a split site may occur at a position up to 5 residues before or after that position (e.g., after position 96, 97, 98, 99, 100, 101, 103, 104, 105, 106, 107). In some embodiments, a split site between β6 and β7 may occur between positions 124 and 125 of SEQ ID NO: 788, or in some embodiments such a split site may occur at a position up to 5 residues before or after that position (e.g., after position 118, 119, 120, 121, 122, 123, 125, 126, 127, 128, 129). In some embodiments, a split site between β7 and β8 may occurr between positions 133 and 134 of SEQ ID NO: 788, or in some embodiments such a split site may occur at a position up to 5 residues before or after that position (e.g., after position 127, 128, 129, 130, 131, 132, 134, 135, 136, 137, 138). In some embodiments, a split site between β8 and β9 may occur between positions 148 and 149 of SEQ ID NO: 788, or in some embodiments such a split site may occur at a position up to 5 residues before or after that position (e.g., after position 142, 143, 144, 145, 146, 147, 149, 150, 151, 152, 153).

In some embodiments, two peptide, dipeptide, tripeptide, and polypeptide components that are sequentially adjacent within the base luciferase (e.g., OgLuc, NANOLUC, SEQ ID NO: 788, SEQ ID NO: 789, etc.) sequence comprise all of the amino acids of that corresponding portion of the base sequence. In some embodiments, one or more (e.g., 1, 2, 3, 4, 5, or more) amino acids adjaecnt to the split point in the base sequence are absent from the corresponding peptide, dipeptide, tripeptide, and/or polypeptide components.

In some embodiments, provided herein are peptides comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%) sequence identity with one of the following:

- β6-like – GVTPNKLNYFGRPYEGIAVFDG (SEQ ID NO: 802);
- β7-like – KKITTTGTL (SEQ ID NO: 803
- β8-like – WNGNKIIDERLITPD (SEQ ID NO: 804

- β9-like – GSMLFRVTINS (SEQ ID NO: 805
- β10-like (Hi affinity) – VSGWRLFKKIS (SEQ ID NO: 806 and
- β10-like (Lo affinity) – VTGYRLFEEIL (SEQ ID NO: 807

In some embodiments, provided herein are dipeptides comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%) sequence identity with one of the following:

- β6/7-like – GVTPNKLNYFGRPYEGIAVFDGKKITTTGTL (SEQ ID NO: 808
- β7/8-like – KKITTTGTLWNGNKIIDERLITPD (SEQ ID NO: 809);
- β8/9-like – WNGNKIIDERLITPDGSMLFRVTINS (SEQ ID NO: 810);
- β9/10-like (Hi affinity) – GSMLFRVTINSVSGWRLFKKIS (SEQ ID NO: 811); and
- β9/10-like (Lo affinity) – GSMLFRVTINSVTGYRLFEEIL (SEQ ID NO: 812).

In some embodiments, provided herein are tridipeptides comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%) sequence identity with one of the following:

- β6/7/8-like – GVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPD (SEQ ID NO: 813);
- β7/8/9-like – KKITTTGTLWNGNKIIDERLITPDGSMLFRVTINS (SEQ ID NO: 814);
- B8/9/10-like (Hi affinity) – WNGNKIIDERLITPDGSMLFRVTINSVSGWRLFKKIS (SEQ ID NO: 815); and
- B8/9/10-like (Lo affinity) – WNGNKIIDERLITPDGSMLFRVTINSVTGYRLFEEIL (SEQ ID NO: 816).

In some embodiments, provided herein are polypeptide comprising 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%) sequence identity with one of the following:

- β1-5-like – MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKI DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVID (SEQ ID NO: 790);

- β6-10-like (Hi affinity) – GVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDGSMLFRVTINSVSGWRLFKKIS (SEQ ID NO: 794);
- β6-10-like (Lo affinity) – GVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDGSMLFRVTINSVTGYRLFEEIL (SEQ ID NO: 798);
- β6-9-like – GVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDGSMLFRVTINS (SEQ ID NO: 829);
- β7-10-like (Hi affinity) – KKITTTGTLWNGNKIIDERLITPDGSMLFRVTINSVSGWRLFKKIS (SEQ ID NO: 795); and
- β7-10-like (Lo affinity) – KKITTTTGTLWNGNKIIDERLITPDGSMLFRVTINSVTGYRLFEEIL (SEQ ID NO: 799).

In some embodiments, a polypeptide component (e.g., of a set of peptides/polypeptides, or a bioluminescent complex, etc.) comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%) sequence identity with one of SEQ ID NOS: 788, 789, 790, 791, 792, and 793.

In some embodiments, peptide/dipeptide/tripeptide components (e.g., tags) (e.g., of a set of peptides/polypeptides, or a bioluminescent complex, etc.) collectively comprise 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%) sequence identity with one of SEQ ID NOS: 794, 795, 796, 797, 798, 799, 800, and 801.

In some embodiments, provided herein are sets of components and complexes of the peptides, dipeptides, tripeptides, and polypeptides listed above. In particular embodiments, sets of components are selected that span all ten of the beta strands of a base luciferase sequence.

In some embodiments, the interaction, co-localization, detection, and other methods, assays, and technologies described for use with the two-peptide tag systems herein (e.g., β9-like (e.g., SmTrip9) peptide, β10-like (e.g., SmTrip10) peptides and polypeptide component ((e.g., β1-8-like (e.g., LgTrip) polypeptide)), also find use with the dipeptide systems described herein (e.g., β9/10-like dipeptide and polypeptide component). In some

embodiments, a dipeptide has high affinity for a polypeptide component; in such embodiments, a bioluminescent complex forms when the dipeptide and polypeptide component are brought into contact (e.g., co-localize, are added to the sample sample, etc.) without facilitation. In some embodiments, a dipeptide has low affinity for a polypeptide component; in such embodiments, a bioluminescent complex will not form when the dipeptide and polypeptide component are brought into contact (e.g., co-localize, are added to the sample sample, etc.) without facilitation. Like the two-peptide tag systems herein (e.g., β9-like (e.g., SmTrip9) peptide, β10-like (e.g., SmTrip10) peptides and polypeptide component (e.g., β1-8-like (e.g., LgTrip) polypeptide)), dipeptide/polypeptide pairs of varying affinities may be selected for different applications. In some embodiments, systems, methods, and assays for two-component complementation systems are described in U.S. Pat. No. 9,797,890 (herein incorporated by reference in its entirety), and all such systems, methods, and assays find use with the dipeptide/polypeptide systems described herein.

In some embodiments, the interaction, co-localization, detection, and other methods, assays, and technologies described for use with the two-peptide tag systems herein (e.g., β9-like (e.g., SmTrip9) peptide, β10-like (e.g., SmTrip10) peptides and polypeptide component ((e.g., β1-8-like (e.g., LgTrip) polypeptide)), also find use with systems comprising any suitbale combination fo peptides, dipeptides, trippeptides, and polypeptides, as described herein. In some embodiments, the components have high affinity for one another; in such embodiments, a bioluminescent complex forms when the components are brought into contact (e.g., co-localize, are added to the sample sample, etc.) without facilitation. In some embodiments, one or more of the components have low affinity for one or more of the other components; in such embododiments, a bioluminescent complex will not form when the components are brought into contact (e.g., co-localize, are added to the sample sample, etc.) without facilitation. Like the two-peptide tag systems herein (e.g., β9-like (e.g., SmTrip9) peptide, β10-like (e.g., SmTrip10) peptides and polypeptide component (e.g., β1-8-like (e.g., LgTrip) polypeptide)), the other systems described herein may be provided with varying affinities for different applications. In some embodiments, systems, methods, and assays for two-component complementation systems are described in U.S. Pat. No. 9,797,890 (herein incorporated by reference in its entirety), and all such systems, methods, and assays find use with the various peptide, dipeptide, tripeptide, and polypeptide systems described herein.

In some embodiments, provided herein are complementary panels of interchangeable peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) that have variable affinities and luminescence upon formation of bioluminescent complexes therefrom (e.g., a high-affinity/high-luminescence, a moderate-affinity/high-luminescence, a low-affinity/moderate-luminescence, etc.). Utilizing different combinations of peptide/dipeptide/tripeptide tags and polypeptide components provides an adaptable system comprising various sets ranging from lower to higher affinities, luminescence, expression level, stability, solubility, and other variable characteristics. This adaptability allows the detection/monitoring/identification/quantification of analytes, molecular interactions, co-localization, and/or other characteristics to be fine-tuned to the specific molecule(s) of interest and/or conditions to be studied and expands the range of molecular interactions and/or co-localizations that can be detected/monitored/identified/quantified to include interactions with very high or low affinities. Further provided herein are methods by which non-luminescent elements and panels of non-luminescent elements are developed and tested.

In some embodiments, due to the small size of the tags (e.g., peptide tags) herein (e.g., compared to larger polypeptides and proteins), they are resistant to denaturation (they have no tertiary structure required for function).

In some embodiments, peptide/dipeptide/tripeptide tags and a polypeptide components may be selected based on the molecules or proteins of interest to be studied. In some embodiments, different peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) may require different strength, duration, and/or stability of an interaction complex (e.g., complex of interaction elements) to result in bioluminescent complex formation. In some embodiments, a highly stable interaction complex is required to produce a detectable bioluminescent signal (e.g., in the presence of a substrate). In other embodiments, even a weak or transient interaction complex results in bioluminescent complex formation. In still other embodiments, a bioluminescent complex will form in the absence of an interaction complex as long as the peptide/dipeptide/tripeptide tags and

polypeptide component are co-localized. In some embodiments, the strength or extent of an interaction complex is directly proportional to the strength of the resulting bioluminescent signal. Some peptide/dipeptide/tripeptide tags/polypeptide component sets produce a detectable signal when combined with an interaction complex with a high millimolar dissociation constant (e.g., K_d>100 mM). Other peptide/dipeptide/tripeptide tags/polypeptide component sets require an interaction pair with a low millimolar (e.g., K_d<100 mM), micromolar (e.g., K_d<1 mM), nanomolar (e.g., K_d<1 μM), or even picomolar (e.g., K_d<1 nM) dissociation constant in order to produce a bioluminescent complex with a detectable signal.

In some embodiments, the peptide/dipeptide/tripeptide tags and/or polypeptide components herein are not fragments of a pre-existing protein (e.g., a pre-existing bioluminescent protein). In some embodiments, none of the peptide/dipeptide/tripeptide tags and polypeptide component used to form a complex are fragments of a pre-existing protein (e.g., the same pre-existing protein, a pre-existing bioluminescent protein, etc.). In some embodiments, neither the peptide tags (e.g., β9-like (e.g., SmTrip9) and β10-like (e.g., SmTrip10) peptides; β9/β10-like dipeptides; etc.) nor polypeptide component (e.g., β_{1-8}-like (e.g., LgTrip) polypeptide)) that assemble together to form a bioluminescent complex are fragments of a pre-existing protein (e.g., the same pre-existing protein, a pre-existing bioluminescent protein, etc.). In some embodiments, the peptide/dipeptide/tripeptide tags or polypeptide component of a bioluminescent complex for use in embodiments of the present invention is not a subsequence of a preexisting protein. In some embodiments, non-luminescent elements for use in embodiments described herein do not comprise structurally-complementary subsequences of a preexisting protein.

In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) herein are non-luminescent or substantially non-luminescent in isolation (e.g., in the presence or absence of substrate). In some embodiments, the peptide/dipeptide/tripeptide tags herein are non-luminescent or substantially non-luminescent when associated together, in the absence of the polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) (e.g., in the presence or absence of substrate). In some embodiments, a polypeptide component is non-luminescent or substantially non-luminescent in isolation (e.g., in the presence or absence of substrate). In some embodiments, a single peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like,

β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and the polypeptide component are non-luminescent or substantially non-luminescent in the absence of the second or third or fourth peptide/dipeptide/tripeptide tag (e.g., in the presence or absence of substrate). In certain embodiments, when placed in suitable conditions (e.g., physiological conditions), multiple peptide/dipeptide/tripeptide tags and a polypeptide component interact to form a bioluminescent complex and produce a bioluminescent signal in the presence of substrate.

In certain embodiments, an interaction element and/or co-localization element and a peptide/dipeptide/tripeptide tag are attached, fused, linked, connected, etc. In typical embodiments, a first peptide/dipeptide/tripeptide tag and a first interaction element (or first co-localization element) are attached to each other, and a second peptide/dipeptide/tripeptide tag and a second interaction element (or second co-localization element) are attached to each other. Attachment of peptide/dipeptide/tripeptide tags to interaction elements (or co-localization elements) may be achieved by any suitable mechanism, chemistry, linker, etc. The interaction elements (or co-localization elements) and peptide/dipeptide/tripeptide tags are typically attached through covalent connection, but non-covalent linking of the two elements is also provided. In some embodiments, the peptide/dipeptide/tripeptide tags and interaction elements (or co-localization elements) are directly connected and, in other embodiments, they are connected by a linker. In some embodiments, the peptide/dipeptide/tripeptide tags and interaction elements (or co-localization elements) are provided as genetic/recombinant fusions. In some embodiments, endogenous tagging with the peptide/dipeptide/tripeptide tags herein (e.g., under endogenous regulatory control), allows for monitoring of normal cellular functions with the tools described herein. For example, a protien of interest may be endogenously tagged (e.g., using CRISPR/Cas9) with a high affinity β9/β10-like dipeptide, and then spontaneous complementation with LgTrip (or a variant thereof) is monitored in a cell, animal, lysate, etc. In other embodiments, the peptide tags and interaction elements (or co-localization elements) are connected by chemical modification/conjugation, such as by Native chemical ligation, Staudinger ligation, "traceless" Staudinger ligation, amide coupling, methods that employ activated esters, methods to target lysine, tyrosine and cysteine residues, imine bond formation (with and without *ortho*-boronic acid), boronic acid/diol interactions, disulfide bond formation, copper/copper free azide, diazo, and tetrazine "click" chemistry, UV promoted thiolene

conjugation, diazirine photolabeling, Diels-Alder cycloaddition, metathesis reaction, Suzuki cross-coupling, thiazolidine (Step-4) coupling, streptavidin/biotin complementation, HaloTag/chloroalkane substrate complementation, etc. In some embodiments, peptide/dipeptide/tripeptide tags and interaction elements (or co-localization elements) are produced synthetically (e.g., solid-state synthesis, solution-phase synthesis, etc.). In some embodiments, interaction elements (or co-localization elements) are produced (e.g., synthetically or recombinantly) or obtained (e.g., from crude lysate, extracted proteins, purified proteins, etc.) by any suitable means.

In some embodiments, in which the interaction element (or co-localization element) is a peptide or polypeptide, a peptide/dipeptide/tripeptide tag (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and an interaction element (or co-localization element) are contained within a single amino acid chain. In some embodiments, a single amino acid chain comprises, consists of, or consists essentially of a peptide/dipeptide/tripeptide tag and an interaction element (or co-localization element). In some embodiments, a single amino acid chain comprises, consists of, or consists essentially of a peptide/dipeptide/tripeptide tag, an interaction element (or co-localization element), optionally one or more an N-terminal sequence, a C-terminal sequence, regulatory elements (e.g., promoter, translational start site, etc.), and a linker sequence. In some embodiments, the peptide/dipeptide/tripeptide tag and interaction element (or co-localization element) are contained within a fusion polypeptide. In some embodiments, the first fusion of peptide/dipeptide/tripeptide tag and interaction element (or co-localization element) and the second fusion of peptide/dipeptide/tripeptide tag and interaction element (or co-localization element) are expressed separately; however, in other embodiments, a fusion protein is expressed that comprises or consist of both of the interaction (or co-localization) and peptide/dipeptide/tripeptide tags.

In some embodiments, a first fusion protein comprising a first peptide/dipeptide/tripeptide tag (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and first interaction element as well as a second fusion protein comprising a second peptide/dipeptide/tripeptide tag (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and second interaction element are expressed within the same cells. In some embodiments, a first fusion

protein comprising a first peptide/dipeptide/tripeptide tag and first co-localization element as well as a second fusion protein comprising a second peptide/dipeptide/tripeptide tag and second co-localization element are expressed within the same cells. In some embodiments, the first and second fusion proteins are purified and/or isolated from the cells. In some embodiments, the interaction and/or co-localization of the fusion proteins is assayed within the cells. In some embodiments, the interaction and/or co-localization of the fusion proteins is assayed within a lysate of the cells. In other embodiments, first and second fusion proteins are expressed in separate cells and combined (e.g., following purification and/or isolation, following fusion of the cells or portions of the cells, by transfer of a fusion protein from one cell to another, or by secretion of one or more fusion proteins into the extracellular medium) for signal detection. In some embodiments, one or more fusion proteins are expressed in cell lysate (e.g., rabbit reticulocyte lysate) or in a cell-free system. In some embodiments, one or more fusion proteins are expressed from the genome of a virus or other cellular pathogen. In some embodiments, the polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) and any other peptide/dipeptide/tripeptide components (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) for complex formation (with the first and second fusion proteins) is expressed in the same cell or cell lysate as one or both of the tag-containing fusion proteins. In some embodiments, the peptide/dipeptide/tripeptide/polypeptide components for complex formation with the peptide/dipeptide/tripeptide tags (within the first and second fusion proteins) are expressed in a different cell or cell lysate as one or both of the peptide-tag-containing fusion proteins. In some embodiments, the peptide/dipeptide/tripeptide/polypeptide components for complex formation with the peptide/dipeptide/tripeptide tags (within the first and second fusion proteins) is added to a cell, cell lysate, or other sample comprising the peptide-tag-containing fusion proteins.

In some embodiments, the systems (e.g., peptide/dipeptide/tripeptide tags, peptide/dipeptide/tripeptide/polypeptide components, substrates, vectors, etc.) and methods herein find use in the analysis of a sample (e.g., detection/quantification/identification/monitoring of co-localization, a molecular interaction, a target, etc.). In some embodiments, one or more of the components of a system herein are added to and/or provided or expressed within a sample. Suitable samples that may find use in

embodiments herein include, but are not limited to: blood, plasma, sera, urine, saliva, cells, cell lysates, tissues, tissue homogenates, purified nucleic acids, stool, vaginal secretions, cerebrospinal fluid, allantoic fluid, water, biofilm, soil, dust, food, beverage, agriculture products, plants, etc.

In certain embodiments, nucleic acids, DNA, RNA, vectors, etc. are provided that encode the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide), fusion polypeptides, fusion proteins, etc. described herein. Such nucleic acids and vectors may be used for expression, transformation, transfection, injection, etc.

In some embodiments, a peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and interaction, co-localization element, or binding agent are connected by a linker. In some embodiments, a linker connects the signal and interaction or co-localization elements while providing a desired amount of space/distance between the elements. In some embodiments, a linker allows both the signal and interaction elements to form their respective complexes (e.g., luminescent complex and interaction complex) simultaneously. In some embodiments, a linker assists the interaction element in facilitating the formation of a luminescent complex. In some embodiments, when an interaction complex is formed, the linkers that connect each peptide/dipeptide/tripeptide tag to their respective interaction elements position the peptide tags at the proper distance and conformation to form a bioluminescent complex. In some embodiments, an interaction or co-localization element and peptide/dipeptide/tripeptide tag are held in close proximity (e.g., <4 monomer units) by a linker. In some embodiments, a linker provides a desired amount of distance (e.g., 1, 2, 3, 4, 5, 6… 10… 20, or more monomer units) between peptide tags and interaction elements (e.g., to prevent undesirable interactions between peptide/dipeptide/tripeptide tags and interaction or co-localization elements, for steric considerations, to allow proper orientation of non-luminescent element upon formation of interaction complex, to allow propagation of a complex-formation from interaction complex to luminescent complex, etc.). In certain embodiments, a linker provides appropriate attachment chemistry between the peptide/dipeptide/tripeptide tags and interaction elements.

A linker may also improve the synthetic process of making the peptide/dipeptide/tripeptide tag and interaction or co-localization element (e.g., allowing them to be synthesized as a single unit, allowing post synthesis connection of the two elements, etc.).

In some embodiments, a linker is any suitable chemical moiety capable of linking, connecting, or tethering a peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) or polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) to an interaction element or co-localization element. In some embodiments, a linker is a polymer of one or more repeating or non-repeating monomer units (e.g., nucleic acid, amino acid, carbon-containing polymer, carbon chain, etc.). When a peptide/dipeptide/tripeptide tag and an interaction, co-localization element, or binding agent are part of a fusion protein, a linker (when present) is typically an amino acid chain. When a peptide/dipeptide/tripeptide tag and interaction element, co-localization element, or binding agent are tethered together after the expression of the individual elements, a linker may comprise any chemical moiety with functional (or reactive) groups at either end that are reactive with functional groups on the peptide tag and interaction or co-localization elements, respectively. Any suitable moiety capable of tethering the signal and interaction elements, co-localization element, and/or binding agent may find use as a linker.

A wide variety of linkers may be used. In some embodiments, the linker is a single covalent bond. In some embodiments, the linker comprises a linear or branched, cyclic or heterocyclic, saturated or unsaturated, structure having 1-20 nonhydrogen atoms (e.g., C, N, P, O and S) and is composed of any combination of alkyl, ether, thioether, imine, carboxylic, amine, ester, carboxamide, sulfonamide, hydrazide bonds and aromatic or heteroaromatic bonds. In some embodiments, linkers are longer than 20 non-hydrogen atoms (e.g. 21 non-hydrogen atoms, 25 non-hydrogen atoms, 30 non-hydrogen atoms, 40 non-hydrogen atoms, 50 non-hydrogen atoms, 100 non-hydrogen atoms, etc.) In some embodiments, the linker comprises 1-50 non-hydrogen atoms (in addition to hydrogen atoms) selected from the group of C, N, P, O, and S (e.g. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40,41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 non-hydrogen atoms).

The scope of embodiments herein is not limited by the types of linkers available. The peptide/dipeptide/tripeptide tags, polypeptide components, and interaction elements, co-

localization elements, or binding agents are linked either directly (e.g. linker consists of a single covalent bond) or linked via a suitable linker. Embodiments are not limited to any particular linker group. A variety of linker groups are contemplated, and suitable linkers could comprise, but are not limited to, alkyl groups, methylene carbon chains, ether, polyether, alkyl amide linker, a peptide linker, a modified peptide linker, a Poly(ethylene glycol) (PEG) linker, a streptavidin-biotin or avidin-biotin linker, polyaminoacids (e.g. polylysine), functionalized PEG, polysaccharides, glycosaminoglycans, dendritic polymers (WO93/06868 and by Tomalia et al. in Angew. Chem. Int. Ed. Engl. 29:138-175 (1990), herein incorporated by reference in their entireties), PEG-chelant polymers (W94/08629, WO94/09056 and WO96/26754, herein incorporated by reference in their entireties), oligonucleotide linker, phospholipid derivatives, alkenyl chains, alkynyl chains, disulfide, or a combination thereof. In some embodiments, the linker is cleavable (e.g., enzymatically (e.g., TEV protease site), chemically, photoinduced, etc.

In some embodiments, a peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide), recognition element, interaction element, co-localization element, binding agent, analyte, substrate, etc. is attached (e.g., via any suitable chemistry) to, or contained within, a solid surface or matrix. In some embodiments, one or more system components are attached (e.g., via any suitable chemistry) to, or contained within, a solid surface or matrix and other components are added (e.g., in solution (e.g., in a sample)) to the solid surface or matrix. Suitable solid surfaces include, but are not limited to: beads (e.g., magnetic beads), chips, tubes, plates, particles, membranes, paper, etc. In some embodiments, solid surfaces/matrix is made of any suitable materials, such as: Ahlstrom CytoSep, Cellulose nitrate, Cellulose acetate, Cellulose (e.g., Whatman FTA-DMPK-A, B, and C cards; Whatman ET 3/Chr; Whatman protein saver 903 cards; Whatman Grade 1 filter paper; Whatman FTA Elute; Ahlstrom 226 specimen collection paper; etc.), Noviplex Plasma Prep Cards, Polypropylene membrane, PVDF, Nitrocellulose membrane (Millipore Nitrocellular Hi Flow Plus) Polytetrafluoroethylene film, Mixed cellulose esters, Glass fiber media (e.g., Whatman uniflter plates glass fiber filter membrane, Agilent dried matrix spotting cards, Ahlstrom grade 8950, etc.), Plastic (e.g., Polyester, Polypropylene, Polythersulfene, poly (methacrylate), Acrylic polymers, polytetrafluoreten, etc.), natural and

synthetic polymers (e.g., mixture of polymers, co-block polymers, etc.), sugars (e.g., pullulan, trehalose, maltose, sucrose, cellulose, etc.), polyamides (e.g., natural (e.g., wool, silk, etc.), synthetic (e.g., aramids, nylon, etc.), etc.), metals (e.g., aluminum, cadmium, chromium, cobalt, copper, iron, manganese, nickel, platinum, palladium, rhodium, silver, gold, tin, titanium, tungsten, vanadium, zinc, etc.), alloys (e.g., alloys of aluminium (e.g., Al-Li, alumel, duralumin, magnox, zamak, etc.), alloys of iron (e.g., steel, stainless steel, surgical stainless steel, silicon steel, tool steel, cast iron, Spiegeleisen, etc.), alloys of cobalt (e.g., stellite, talonite, etc.), alloys of nickel (e.g., German silver, chromel, mu-metal, monel metal, nichrome, nicrosil, nisil, nitinol, etc.), alloys of copper (e.g., beryllium copper, billon, brass, bronze, phosphor bronze, constantan, cupronickel, bell metal, Devarda's alloy, gilding metal, nickel silver, nordic gold, prince's metal, tumbaga, etc.), alloys of silver (e.g., sterling silver, etc.), alloys of tin (e.g., Britannium, pewter, solder, etc.), alloys of gold (electrum, white gold, etc.), amalgam, etc.), ELISPot plates, Immunoassay plates, Tissue culture plates, etc.

In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) of a luminescent complex are provided with less than 100% sequence identity and/or similarity to any portion of an existing luciferase (e.g., a firefly luciferase, a Renilla luciferase, an Oplophorus luciferase, enhanced Oplophorus luciferases as described in U.S. Pat. App. 2010/0281552 and U.S. Pat. App. 2012/0174242, herein incorporated by reference in their entireties). Certain embodiments involve the formation of bioluminescent complexes of peptide/dipeptide/tripeptide tags and a polypeptide component with less than 100% sequence identity with all or a portion (e.g., 8 or more amino acids, less than about 25 amino acids for peptides) of SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence). Certain embodiments involve the formation of bioluminescent complexes from peptide/dipeptide/tripeptide tags and a polypeptide component with less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity with all or a portion (e.g., 8 or more amino acids, less than about 25 amino acids for peptides) of SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence). In some embodiments, peptide/dipeptide/tripeptide tags and a polypeptide

component are provided with less than 100% sequence similarity with a portion (e.g., 8 or more amino acids, less than about 25 amino acids for peptides) of SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence). In some embodiments, peptide/dipeptide/tripeptide tags and a polypeptide component are provided with less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence similarity with a portion (e.g., 8 or more amino acids, less than about 25 amino acids for peptides) of SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence). In some embodiments, peptide/dipeptide/tripeptide tags are provided that have less than 100% sequence identity and/or similarity with about a 25 amino acid or less portion of SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence), wherein two of such peptides form a bioluminescent complex when combined under appropriate conditions (e.g., stabilized by an interaction pair, brought into proximity by co-localization elements, etc.) with a polypeptide component having less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity and/or similarity with another portion SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence). In some embodiments, peptide/dipeptide/tripeptide tags are provided that have less than 100% sequence identity and/or similarity with about a 25 amino acid or less portion of SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence), wherein a pair of such peptide tags form a bioluminescent complex when combined under appropriate conditions (e.g., stabilized by an interaction pair, brought into proximity by co-localization elements, etc.) with a polypeptide component having less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity and/or similarity with another portion SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence). In some embodiments, peptide/dipeptide/tripeptide tags are provided that have less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity and/or similarity with about a 25 amino acid

or less portion of SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence), wherein a pair of such peptides form a bioluminescent complex when combined under appropriate conditions (e.g., stabilized by an interaction pair, brought into proximity by co-localization elements, etc.) with a polypeptide having less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity and/or similarity with another portion of SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence). Similarly, polypeptide components are provided that have less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with a portion of SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence), wherein such polypeptide components form a bioluminescent complex when combined under appropriate conditions with a pair of peptide tags having less than 100%, but optionally more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity and/or similarity with another portion SEQ ID NO: 1 (e.g., complete wild type Oplophorus luciferase sequence) and/or SEQ ID NO: 3 (e.g., complete NANOLUC sequence). In some embodiments, peptide tags with less than 100% sequence identity or similarity with SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, and/or SEQ ID NO: 10 are provided. In some embodiments, peptide tags with less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO:9, and/or SEQ ID NO: 10 are provided. In some embodiments, peptide tags with less than 100% sequence identity or similarity with SEQ ID NO: 23, SEQ ID NO: 25, and/or SEQ ID NO: 29 are provided. In some embodiments, peptide tags with less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 23, SEQ ID NO: 25, and/or SEQ ID NO: 29 are provided. In some embodiments, polypeptide components with less than 100% sequence identity or similarity with SEQ ID NO: 5 and/or SEQ ID NO: 8 are provided. In some embodiments, polypeptide components with less than 100% sequence identity or similarity with SEQ ID NO: 17 and/or SEQ ID NO: 27 are provided. In some embodiments, polypeptide components

with less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 5, SEQ ID NO: 8, and/or SEQ ID NO: 27 are provided. In some embodiments, polypeptide components with less than 100%, but more than 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 17 are provided.

In some embodiments, one or more (e.g., all) peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) in a set, kit, or system herein comprise 100% sequence identity with a portion of a luciferase (e.g., SEQ ID NO: 1, SEQ ID NO: 3, etc.).

In some embodiments, peptide tags (e.g., β9-like (e.g., SmTrip9) and β10-like (e.g., SmTrip10) peptides; β9/β10-like dipeptides; etc.) that find use in embodiments of the present invention include peptides with one or more amino acid substitutions, deletions, or additions from SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO:9, SEQ ID NO: 10, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 29. In some embodiments, a peptide tag comprises at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 13. In some embodiments, a peptide tag comprises 6 or fewer (e.g., 6, 5, 4, 3, 2, 1, or ranges there between) substitutions (e.g., conservative substitutions, semi-conservative substitutions, non-conservative substations, etc.) relative to SEQ ID NO: 13. In some embodiments, a peptide tag comprises at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 25. In some embodiments, a peptide tag comprises 6 or fewer (e.g., 6, 5, 4, 3, 2, 1, or ranges there between) substitutions (e.g., conservative substitutions, semi-conservative substitutions, non-conservative substations, etc.) relative to SEQ ID NO: 25. In some embodiments, a peptide tag comprises at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 23. In some embodiments, a peptide tag comprises 6 or fewer (e.g., 6, 5, 4, 3, 2, 1, or ranges there between) substitutions (e.g., conservative substitutions, semi-conservative substitutions, non-conservative substations, etc.) relative to SEQ ID NO: 23. In some embodiments, a

peptide tag comprises at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 25. In some embodiments, a peptide tag comprises 6 or fewer (e.g., 6, 5, 4, 3, 2, 1, or ranges there between) substitutions (e.g., conservative substitutions, semi-conservative substitutions, non-conservative substations, etc.) relative to SEQ ID NO: 25.

In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) that find use in embodiments of the present invention include the peptides, dipeptides, tripeptides, and polypeptides disclosed herein and in the tables provided herein. In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) that find use in embodiments of the present invention comprise one or more amino acid substitutions, deletions, or additions relative to the peptides, dipeptides, tripeptides, and polypeptides disclosed herein and in the tables provided herein. In some embodiments, a peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) that find use in embodiments of the present invention comprise at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with the peptides, dipeptides, tripeptides, and polypeptides disclosed herein and in the tables provided herein.

In some embodiments, dipeptides and tripeptides that find use in embodiments herein comprise any suitable combinations of the peptides described herein and/or listed in the tables herein.

In some embodiments, a peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) or a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) is linked (e.g., chemically) or fused to one or more additional elements (e.g., recognition element, interaction element, co-localization element, detectable

element (e.g., a fluorophore (e.g., to facilitate BRET)), protein of interest, HALOTAG, etc.). In some embodiments, a peptide/dipeptide/tripeptide tag or polypeptide component is linked or fused to a cyOFP (e.g., in an Antares construct such as those described in U.S. Pat. No. 9,908,918; herein incorporated by reference in its entirety) or other fluorescent protein (e.g., to facilitate BRET). In some embodiments, a peptide/dipeptide/tripeptide tag or polypeptide component comprises one or more chemical modifications and/or unnatural amino acids or amino acid analogs to facilitate chemical conjugation of the polypeptide component with additional elements. In some embodiments, provided herein is a single peptide/dipeptide/tripeptide tag or polypeptide component fused to an acceptor fluorescent protein. In some embodiments, two or more peptide/dipeptide/tripeptide and/or polypeptide components are fused to an acceptor fluorescent protein (e.g., sandwhich fusion). In some embodiments, a peptide/dipeptide/tripeptide tag or polypeptide component is fused to two or more acceptor fluorescent protein (e.g., sandwhich fusion). In some embodiments, a LgTrip polypeptide (e.g., a β_{1-8}-like polypeptide desccribed herein) is fused to a single fluorescent protein (e.g., cyOFP) or placed between two fluorescent proteins (e.g., two copies of a cyOFP) in a sandwhcih fusion.

In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) that find use in the present invention incorporate reactive groups suitable for chemical conjugation to an additional element (e.g., recognition element, interaction element, etc.). These reactive groups may be present on the N-terminus, C-terminus, or within the sequence. These reactive groups may optionally be attached to the peptide with a linker. In some cases, these peptide/dipeptide/tripeptide bearing reactive groups may be synthesized using standard synthesis and incorporated on an unnatural amino acid bearing the desired group. In some cases, the reactive group may be present on a natural amino acid (e.g. the sulfhydryl of cysteine). The additional element intended to react with a peptide tag bearing a reactive group may be a protein, an antibody, a nucleic acid, a small molecule such as a drug or a fluorophore or a surface. The peptide/dipeptide/tripeptide tag may incorporate a reactive group that is designed to react specifically with a reactive partner that has been chemically or biologically introduced on the additional element using bioorthogonal, or click, chemistry. An exemplary click reaction is copper catalyzed click where the peptide tag bears an alkyne or an azide, and the additional element bears the complementary group. Mixing

these two species together in the presence of an appropriate copper catalyst causes the peptide to be covalently conjugated to the additional element through a triazole. Many other bioorthogonal reactions have been reported (for example Patterson, D. M., et al. (2014). "Finding the Right (Bioorthogonal) Chemistry." ACS Chemical Biology 9(3): 592-605.), and peptide tags and additional elements incorporating complementary reactive species are embodiments of the present invention.

Another embodiment of the present invention are peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and/or a polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) bearing reactive groups that react with naturally occurring amino acids. Exemplary reactive groups include maleimides for reaction with cysteine and succinimidyl esters for reaction with lysine. A more comprehensive list of reactive groups can be found in Koniev, O. and A. Wagner (2015). "Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation." Chem Soc Rev 44: 5495-5551. These reactive groups may be chemically or biologically introduced on a peptide/dipeptide/tripeptide/polypeptide through peptide synthesis or through other chemical modification of a peptide tag. In some embodiments, the peptide tag exists in a protected form (Isidro-Llobet, A., et al. (2009). "Amino Acid-Protecting Groups." Chemical Reviews 109(6): 2455-2504; herein incorporated by reference in its entirety), preventing the peptide/dipeptide/tripeptide/polypeptide itself from reacting with the reactive group. These reactive groups may react with a protein in a selective fashion or in a random fashion, yielding either one conjugate or a mixture of conjugates. In some embodiments, either a defined single conjugate or a mixture can be used successfully in this invention.

Examples of peptides (e.g., β9-like (e.g., SmTrip9) and β10-like (e.g., SmTrip10) peptides; β9/β10-like dipeptides; etc.) described herein bearing reactive groups suitable for chemical conjugation to an additional element (e.g., recognition element, interaction element, etc.) are displayed in Figures 95-98. Other combinations of reactive groups and peptides/dipeptides/tripeptides/polypeptides are within the scope herein.

In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-

like (e.g., LgTrip) polypeptide) described herein is fused or conjugated to a detectable element such as a fluorophore or fluorescent protein. In such embodiments, complementation to form the bioluminescent complex, and the resultant bioluminescence, results in BRET and excitement of/emission from the attached detectable element (e.g., fluorophore or fluorescent protein). In such embodiments, the bioluminescent complex is a BRET energy donor, and the detectable element (e.g., fluorophore or fluorescent protein) attached to a component of the complex (e.g., peptide tag or polypeptide component) is the BRET energy acceptor.

Suitable fluorophores for use in a BRET system with the tripartite/multipartite complementation systems described herein include, but are not limited to: xanthene derivatives (e.g., fluorescein, rhodamine, Oregon green, eosin, Texas red, etc.), cyanine derivatives (e.g., cyanine, indocarbocyanine, oxacarbocyanine, thiacarbocyanine, merocyanine, etc.), naphthalene derivatives (e.g., dansyl and prodan derivatives), oxadiazole derivatives (e.g., pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, etc.), pyrene derivatives (e.g., cascade blue), oxazine derivatives (e.g., Nile red, Nile blue, cresyl violet, oxazine 170, etc.), acridine derivatives (e.g., proflavin, acridine orange, acridine yellow, etc.), arylmethine derivatives (e.g., auramine, crystal violet, malachite green, etc.), tetrapyrrole derivatives (e.g., porphin, phtalocyanine, bilirubin, etc.), CF dye (Biotium), BODIPY (Invitrogen), ALEXA FLOUR (Invitrogen), DYLIGHT FLUOR (Thermo Scientific, Pierce), ATTO and TRACY (Sigma Aldrich), FluoProbes (Interchim), DY and MEGASTOKES (Dyomics), SULFO CY dyes (CYANDYE, LLC), SETAU AND SQUARE DYES (SETA BioMedicals), QUASAR and CAL FLUOR dyes (Biosearch Technologies), SURELIGHT DYES (APC, RPE, PerCP, Phycobilisomes)(Columbia Biosciences), APC, APCXL, RPE, BPE (Phyco-Biotech), autofluorescent proteins (e.g., YFP, RFP, mCherry, mKate), quantum dot nanocrystals, etc. In some embodiments, a fluorophore is a rhodamine analog (e.g., carboxy rhodamine analog), such as those described in U.S. Pat. App. Ser. No. 13/682,589, herein incorporated by reference in its entirety.

In other embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) described herein is fused or conjugated to a detectable element such as a fluorophore or fluorescent protein (e.g., green fluorescent protein (GFP), enhanced GFP (EGFP), cyan fluorescent protein (CFP), yellow fluorescent protein (YFP),

and variants thereof. In other embodiments, a peptide tag or polypeptide component described herein is fused or conjugated to a cyan-excitable orange-red fluorescent protein (CyOFP), such as those described in U.S. Pat. No. 9,908,918; herein incorporated by reference in its entirety. In some embodiments, the CyOFP and BRET systems described in U.S. Pat. No. 9,908,918 find use with the peptide tags and/or polypeptide components described herein (e.g., CyOFP-(β_9-like peptide), CyOFP-(β_{10}-like peptide), CyOFP-(β_{1-8}-like polypeptide), CyOFP-(β_{9-10}-like peptide), CyOFP-(β_9-like peptide)-CyOFP, CyOFP-(β_{10}-like peptide)-CyOFP, CyOFP-(β_{1-8}-like polypeptide)-CyOFP, CyOFP-(β_{9-10}-like peptide) - CyOFP, etc.). In some embodiments, such systems comprising CyOFP linked to peptide/dipeptide/tripeptide tags and/or polypeptide components herein may be referred to herein as "Antares constructs" or "Antares systems." Such BRET systems are particularly useful in certain imaging applications (Schaub, F. X., et al. (2015) "Fluorophore-NanoLuc BRET Reporters Enable Sensitive In Vivo Optical Imaging and Flow Cytometry for Monitoring Tumorigenesis." Cancer Research 75(23): 5023-5033; herein incorporated by reference in its entirety).

In other embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) described herein are linked to fluorophores (e.g., directly or via a linker) for use in a constitutive BRET system (e.g., an Antares-like system). In constitutive BRET systems, the emission spectrum is shifted from the bioluminescence spectrum toward that of the fluorophore (e.g., for better sensitivity, lower scattering, desired emission wavelength, etc.). In other embodiments, peptide/dipeptide/tripeptide tags and/or polypeptide components described herein find use as functional sensors (e.g., for monitoring cellular/intracellular/intercellular processes (e.g., for detecting calcium flux or voltage (Suzuki, K., et al. (2016). "Five colour variants of bright luminescent protein for real-time multicolour bioimaging." Nature Communications 7: 13718.; Inagaki, S., et al. (2017). "Genetically encoded bioluminescent voltage indicator for multi-purpose use in wide range of bioimaging." Sci Rep 7: 42398; herein incorporated by reference in their entireties)), for imaging, for optogenetics, etc.).

In some embodiments, two or more of the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides,

and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) are attached to a single interaction element that can access multiple conformations. In one conformation, the peptide/dipeptide/tripeptide tags and the polypeptide are unable to form a luminescent complex. Upon changing conformation, i.e., in response to a stimulus, the peptide/dipeptide/tripeptide tags are brought into a conformation where they can form a bioluminescent complex. As an example, a SmTrip9 peptide and a SmTrip10 peptide can be conjugated to calmodulin such that they do not form a luminescent complex even in the presence of LgTrip and furimazine. Upon exposure to calcium, the conformational change of calmodulin bring the SmTrip9 peptide and SmTrip10 peptide into a position whereupon addition of LgTrip makes a complex that is bioluminescent in the presence of furimazine. Many other biosensors for calcium and other stimuli (pH, voltage, etc.) are known in the literature.

In some embodiments, systems herein find use in multiplexable analyte detection. In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) that find use in the present invention are conjugated to both an interaction element and a reporter element. In some embodiments, the interaction element is an antibody or the like, and the reporter element is a small molecule fluorophore. In some embodiments, antibodies to different pathogens (e.g. Zika virus, Dengue virus, etc.) are conjugated to a peptide/dipeptide/tripeptide tag (e.g., a SmTrip9 peptide) and a fluorophore with a different and distinguishable wavelength. In this embodiment, the luminescent complex that is formed upon the antibody binding to its antigen emits light at the emission wavelength of the bound fluorophore due to bioluminescence resonance energy transfer. This allows the antibodies to all be present in the same well, device, etc., and the identity of the antigen detected to be determined by the color of the light emitted by the luminescent complex formed.

In some embodiments, polypeptide components (e.g., β1-8-like (e.g., LgTrip) polypeptide) that find use in embodiments of the present invention include polypeptides with one or more amino acid substitutions, deletions, or additions from SEQ ID NO: 5 and/or SEQ ID NO: 8. In some embodiments, a polypeptide component comprises at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 11. In some embodiments,

polypeptide component comprises 100 or fewer (e.g., 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or ranges there between) substitutions (e.g., conservative substitutions, semi-conservative substitutions, non-conservative substations, etc.) relative to SEQ ID NO: 11. In some embodiments, a polypeptide component comprises at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 17. In some embodiments, a polypeptide component comprises 100 or fewer (e.g., 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or ranges there between) substitutions (e.g., conservative substitutions, semi-conservative substitutions, non-conservative substations, etc.) relative to SEQ ID NO: 17. In some embodiments, a polypeptide component comprises at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 17, SEQ ID NO: 21, and/or SEQ ID NO: 302. In some embodiments, a polypeptide component comprises 100 or fewer (e.g., 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or ranges there between) substitutions (e.g., conservative substitutions, semi-conservative substitutions, non-conservative substations, etc.) relative to SEQ ID NO: 17, SEQ ID NO: 21, and/or SEQ ID NO: 302. In some embodiments, a polypeptide component comprises at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 788. In some embodiments, polypeptide component comprises 100 or fewer (e.g., 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or ranges there between) substitutions (e.g., conservative substitutions, semi-conservative substitutions, non-conservative substations, etc.) relative to SEQ ID NO: 788. In some embodiments, a polypeptide component comprises at least 40% (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >98%, >99%) sequence identity or similarity with SEQ ID NO: 789. In some embodiments, polypeptide component comprises 100 or fewer (e.g., 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or ranges there between) substitutions (e.g., conservative substitutions, semi-conservative substitutions, non-conservative substations, etc.) relative to SEQ ID NO: 789.

In some embodiments, a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) is linked (e.g., chemically) or fused to one or more additional elements (e.g., recognition element, interaction element, co-localization element,

detectable element (e.g., a fluorophore (e.g., to facilitate BRET)), protein of interest, HALOTAG, etc.). In some embodiments, a polypeptide component is linked or fused to a cyOFP (e.g., in an Antares construct such as those described in U.S. Pat. No. 9,908,918; herein incorporated by reference in its entirety) or other fluorescent protein (e.g., to facilitate BRET). In some embodiments, a polypeptide component comprises one or more chemical modifications and/or unnatural amino acids or amino acid analogs to facilitate chemical conjugation of the polypeptide component with additional elements.

In some embodiments, a peptide tag (e.g., β9-like (e.g., SmTrip9) and β10-like (e.g., SmTrip10) peptides; β9/β10-like dipeptides; etc.) and/or peptide component is not identical to and/or is not exact subsequences of one or more (e.g., all) of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 17, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 51, SEQ ID NO: 302 (or any combinations thereof). In other embodiments, a peptide tag and/or peptide component is identical to and/or is an exact subsequences one or more (e.g., all) of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 17, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, and/or SEQ ID NO: 29, SEQ ID NO: 51, SEQ ID NO: 302 (or any combinations thereof).

In some embodiments, a polypeptide component (e.g., β1-8-like (e.g., LgTrip) polypeptide) corresponds to and comprises substantial sequence identity (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >99%, 100%) with a portion of SEQ ID NO: 3. For example, in some embodiments, a polypeptide component corresponds to, and comprises substantial sequence identity (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >99%, 100%) with positions 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 through 142, 143, 144, 145, 146, 147, 148, 149, 150, 152, 153 of SEQ ID NO: 3 (e.g., positions 1-148).

In some embodiments, a peptide tag (β10-like (e.g, SmTrip10) peptide) corresponds to and comprises substantial sequence identity (e.g., >40%, >50%, >60%, >70%, >80%, >90%, 100%) with a portion of SEQ ID NO: 3. For example, in some embodiments, a peptide tag corresponds to, and comprises substantial sequence identity (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >99%, 100%) with

positions 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 163, or 164 through 166, 167, 168, 169, 170, or 171 of SEQ ID NO: 3 (e.g., positions 160-171).

In some embodiments, a peptide tag (β9-like (e.g., SmTrip9) peptide) corresponds to and comprises substantial sequence identity (e.g., >40%, >50%, >60%, >70%, >80%, >90%, 100%) with a portion of SEQ ID NO: 3. For example, in some embodiments, a peptide tag corresponds to, and comprises substantial sequence identity (e.g., >40%, >45%, >50%, >55%, >60%, >65%, >70%, >75%, >80%, >85%, >90%, >95%, >99%, 100%) with positions 142, 143, 144, 145, 146, 147, 148, 149, 150, 152, 153 through 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 163, or 164 of SEQ ID NO: 3.

In some embodiments, a polypeptide component e.g., β1-8-like (e.g., LgTrip) polypeptide), a first peptide tag (β9-like (e.g., SmTrip9) peptide) and a second peptide tag (β10-like (e.g., SmTrip10) peptide) together correspond to and comprise substantial sequence identity (e.g., >40%, >50%, >60%, >70%, >80%, >90%, 100%) with at least 90% of the length of SEQ ID NO: 3.

In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) comprise one or more substitutions relative to SEQ ID NO: 1 and/or SEQ ID NO: 3. For example, in some embodiments, a polypeptide component comprises 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99%) sequence identity with SEQ ID NO: 3 or a portion thereof (e.g., SEQ ID NO: 11, SEQ ID NO: 17, etc.), but comprise a substitution at one or more of positions 4, 30, 42, and/or 106 relative to SEQ ID NO: 17. In some embodiments, a polypeptide component comprises an E4D substitution relative to SEQ ID NO: 17. In some embodiments, a polypeptide component comprises an A, D, E, G, K, L, M, N, Q, S, T, V, or Y at position 30 relative to SEQ ID NO: 17. In some embodiments, a polypeptide component comprises an A, C, F, G, I, L, M, S, T, or V at position 42 relative to SEQ ID NO: 17. In some embodiments, a polypeptide component comprises a D, K, or Q at position 106 relative to SEQ ID NO: 17.

In some embodiments, a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) is an artificial sequence that comprises 70% or greater (e.g., 75%, 80%, 85%, 90%, 95%, 100%, or ranges there between) sequence identity and/or sequence similarity with one or more of SEQ ID NOS: 19, 21, 39, 41, 43, 45, 47, 49, 51, 53,

55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, and 131 (or the r β1-5-like, β1-6-like, or β1-7-like portion thereof). In some embodiments, a polypeptide component is an artificial sequence that comprises all or a portion (e.g., 50 amino acids, 60 amino acids, 70 amino acids, 80 amino acids, 90 amino acids, 100 amino acids, 110 amino acids, 120 amino acids, 130 amino acids, 140, or more, or ranges there between) of one of SEQ ID NOs: 19, 21, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, and 131 (or the r β1-5-like, β1-6-like, or β1-7-like portion thereof). In some embodiments, a polypeptide component is a sequence consisting of one of SEQ ID NOs: 19, 21, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, and 131 (or the r β1-5-like, β1-6-like, or β1-7-like portion thereof).

In some embodiments, a peptide/dipeptide/tripeptide tag ((e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) is an artificial sequence that comprises 70% or greater (e.g., 75%, 80%, 85%, 90%, 95%, 100%, or ranges there between) sequence identity and/or sequence similarity with one or more of the peptide sequences listed in Table 1, Table 9, Table 10, or dipeptide/tripeptide combinations thereof. In some embodiments, a peptide/dipeptide/tripeptide tag component is an artificial sequence that comprises all or a portion (e.g., 5 amino acids, 6 amino acids, 7 amino acids, 8 amino acids, 9 amino acids, 10 amino acids, 11 amino acids, 12 amino acids, 13 amino acids, 14, or more, or ranges there between) of one of the peptide sequences listed in Table 1, Table 9, Table 10, or dipeptide/tripeptide combinations thereof. In some embodiments, a peptide/dipeptide/tripeptide tag component is a sequence consisting of one of the peptide sequences listed in Table 1, Table 9, Table 10, or dipeptide/tripeptide combinations thereof.

Although referred to herein as peptide/dipeptide/tripeptide e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof), in some embodiments, one or more of the peptide/dipeptide/tripeptide components of a bioluminescent complex within the scope herein are not attached to an interaction element, co-localization element, binding agent, protein of interest, molecule of

interest, or any other moiety. In some embodiments, one or both of the peptide/dipeptide/tripeptide components interact with the polypeptide and other peptide/dipeptide/tripeptide components to form a luminescent complex without being fused or otherwise tethered to another element.

In some embodiments, a peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) of a luminescent complex, a co-localization element, and/or an interaction element comprises a synthetic peptide/polypeptide, a peptide/polypeptide containing one or more non-natural amino acids, a peptide/polypeptide containing one or more amino acid analogs, a peptide/polypeptide mimetic, a conjugated synthetic peptide (e.g., conjugated to a functional group (e.g., fluorophore, luminescent substrate, etc.)), etc.

Provided herein are compositions and methods that are useful in a variety of fields including basic research, medical research, molecular diagnostics, etc. Although the reagents and assays described herein are not limited to any particular applications, and any useful application should be viewed as being within the scope of the present invention, the following are exemplary assays, kits, fields, experimental set-ups, etc. that make use of the presently claimed invention.

Typical applications that make use of embodiments herein involve the monitoring/detection of protein dimerization (e.g., heterodimers, homodimers), protein-protein interactions, protein-RNA interactions, protein-DNA interactions, antibody (or other recognition element) binding to a target, nucleic acid hybridization, protein-small molecule interactions, analyte quantitation or detection, or any other combinations of molecular entities. In an exemplary embodiment, a first entity of interest is attached to a first peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a second entity of interest is attached to the second peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof). If a detectable signal is produced under the particular assay conditions (e.g., in the presence of a polypeptide component of the luminescent complex (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) and a coelenterazine or a coelenterazine analog substrate), then interaction

and/or co-localization of the first and second entities is inferred. Such assays are useful for monitoring molecular interactions and/or localization under any suitable conditions (e.g., in vitro, in vivo, in situ, whole animal, etc.), and find use in, for example, drug discovery, elucidating molecular pathways, studying equilibrium or kinetic aspects of complex assembly, high throughput screening, proximity sensor, etc.

In some embodiments, the systems and methods provided herein are useful for the detection, quantification, analysis, characterization, etc. of: an analyte, analytes, co-localization of analytes, and/or molecular interaction of analytes. In some embodiments, a peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) is tethered/fused to an analyte. In some embodiments, a peptide/dipeptide/tripeptide tag is tethered/fused to a recognition element agent that binds to a target analyte.

Suitable analytes that find use (e.g., are analyzed) in embodiments herein include, but are not limited to: nucleic acids (e.g., DNA, RNA, miRNA, etc.), proteins (ex: bacterial antigens, viral antigens, biomarkers, antibodies, etc.), small molecules, toxins, biomarkers, environmental or food contaminants, surfactants, pathogens (e.g., viral antigens and proteins, bacterial antigens and proteins, etc.), drugs (e.g., therapeutic drugs, drugs of abuse, etc.), vitamins, cytokines, antibodies (e.g., autoantibodies, infectious disease exposure, therapeutic drug monitoring, anti-HLA transplantation rejection, etc.), cells, cell receptor proteins, biomarker based diagnostics, cell free nucleic acids and non-cell free nucleic acids (e.g., DNA, RNA, mRNA, miRNA, etc.), nucleic acid SNPs, extracted nucleic acids, non-amplified nucleic acid samples, genomic DNA, ssDNA, bacterial resistance genes, immunocomplexes (e.g., antigen:antibody complex; antigen:complement complex, etc.), blood sugars, hormones, metabolites, microbes, parasites, enzymes, transcription factors, metal ions/heavy metals, etc.

Suitable recognition elements or binding moieties that find use (e.g., fused/tethered to a peptide tag, binding to an analyte, etc.) in embodiments herein, include, but are not limited to: antibodies (e.g., monoclonal, polyclonal, recombinant, animal derived, autoantibody, biotherapeutic, etc.), antibody variable heavy chain, antibody variable light chain, antibody binding fragment (Fab) [F(ab)'2], camelid, single chain variable fragment (scFv), monomeric proteins, receptor domains, affibodies, monobodies, natural and derivatized nucleic acid aptamers (e.g., RNA aptamer, DNA aptamer, chemical modified aptamer, etc.), peptide

nucleic acids (PNA), locked nucleic acids (LNA), hexitol nucleic acids (HNA), protein A, G, L, M and/or domains thereof, sequence specific oligonucleotide probes (e.g., DNA probe, RNA probe, etc.), small molecule drug, antibody-oligonucleotide conjugates, darpins, nanobodies, affimers, adhirons, anticalins, phage, magnetic particles (e.g., labeled directly or labeled with a tagged recognition element), nanoparticles (e.g., polystyrene nanospheres, etc.) labeled directed or labeled with a tagged recognition element, streptavidin, antigens, etc. In some embodiments, a peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) is linked to an oligonucleotide recognition element or binding moiety. Such constructs may find use in nucleic acid (e.g., DNA, RNA, etc.) complementation and/or detection. Exemplary peptide/oligomer probes are depicted in Figure 99. In such exempalry constructs, a peptide/dipeptide/tripeptide comprising a reactive group (e.g., azido group (e.g., N-terminal, C-terminal, internal, etc.) or other reactive group herein) is conjugated to an oligonucleotide comprising a complementary reactive group (e.g., alkyne group (e.g., 5'-terminal, 3'-terminal, interanl, etc.) or other reactive group herein). In an exemplary embodiment, peptide oligonuceleotide probes are prepard by combing components and reagents (e.g., oligonucleotide (1 mg,161 nmol, in water); triethylammonium acetate buffer (40 uL, 1M in water); aminoguanidine hydrochloride (8 uL, 50 mM in water); peptide (2.8 mg, 1.93 umol, in DMSO); copper(II) TBTA solution (10 mM in 1:1 water/DMSO); ascorbic acid solution (50 mM in water); final volume is 300 uL, 1:1 Water:DMSO); vortexing and heat for 30 min at 60°C; filtering using Illustra NAP-5 column; exchanging buffer into TE buffer that is RNase and DNase free; and storing at -20°C. In embodiments, in which a molecular interaction is being monitored/detected, peptide/dipeptide/tripeptide tags and a polypeptide component are selected that have affinities for each other such that a significant increase in signal is detectable/measurable upon interaction (e.g., binding) of the associated first and second entities. In some embodiments, one or both (or more) peptide/dipeptide/tripeptide tags have sufficiently low affinity for the other peptide tag and/or the polypeptide component that only background luminescence is detected in the absence of the interaction (e.g., binding) between the associated first and second entities. In other embodiments, the peptide/dipeptide/tripeptide tags and polypeptide component will form a complex and produce a signal in the absence of interaction between

the associated first and second entities, but the signal is increased (e.g., 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 10^2-fold, 10^3-fold, 10^4-fold, 10^5-fold, 10^6-fold, or more or ranges there between) upon interaction (e.g., binding) of the associated first and second entities.

In embodiments in which a co-localization is being monitored/detected, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) are selected that have affinities for each other, such that a signal from the luminescent complex is detectable/measurable even in the absence of an interaction (e.g., binding) of the associated first and second entities. In such embodiments, if the associated first and second entities co-localize (e.g., in the same tissue, in the same cell, in the same subcellular compartment, etc.), the peptide/dipeptide/tripeptide tags and polypeptide component will form a complex and emit a signal (in the presence of coelenterazine or a coelenterazine analog), whether or not the first and second entities interact with each other. In some embodiments, two or more (e.g., both, all) of the peptide/dipeptide/tripeptide tags have sufficiently high affinity for the other components that luminescence is detected in the absence of the interaction (e.g., binding) between the associated first and second entities. In some embodiments, no significant increase in signal is detected upon interaction of the first and second entities. In other embodiments, the peptide/dipeptide/tripeptide tags and polypeptide component will form a complex and produce a signal in the absence of interaction between the associated first and second entities, but the signal is increased (e.g., 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 10^2-fold, 10^3-fold, 10^4-fold, 10^5-fold, 10^6-fold, or more or ranges there between) upon interaction (e.g., binding) of the associated first and second entities.

In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) having known characteristics (e.g., spectral characteristics, mutual affinity, etc.) are used to elucidate the affinity of, or understand the interaction of, an interaction pair of interest. In other embodiments, a well-characterized interaction pair is used to determine the characteristics (e.g., spectral characteristics, mutual affinity, etc.) of

one or more elements of a set of peptide/dipeptide/tripeptide tags and a polypeptide component. In some embodiments, peptide/dipeptide/tripeptide tags and a polypeptide component having known characteristics (e.g., spectral characteristics, mutual affinity, etc.) are used to characterize/monitor the co-localization of a co-localization par of interest (e.g., under desired conditions).

Embodiments described herein may find use in drug screening and/or drug development. For example, the interaction of a small molecule drug or an entire library of small molecules with a target protein of interest (e.g., therapeutic target) is monitored under one or more relevant conditions (e.g., physiological conditions, disease conditions, etc.). Such an assay may comprise a first peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) attached to a drug candidate (or a library of candidates) and a second peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) attached to a therapeutic target; luminescence in the present of the polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) and substrate indicates interaction and/or co-localization of the candidate and target.

Some embodiments herein find use in the diagnostic or criminal setting for monitoring for drugs (e.g., drugs of abuse in human) as well as for therapeutic drug monitoring of patients in biological samples. For example, two peptide/dipeptide/tripeptide tagged binding moieties (e.g., binding moieties separately tagged with peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) that recognize a drug analyte facilitate such embodiments. In some embodiments, a competitive dispalcement assay utilizing a peptide/dipeptide/tripeptide-tagged target in a system described herein to identify untagged target in a sample finds use in embodiments herein. Some embodiments find use in detecting environmental contamination, for example, soil samples, water supply, etc. being contaminated by a specific drug or other specific contaminant (e.g., small molecule contaminant).

In other embodiments, the ability of a drug (e.g., small molecule drug) or an entire library of drugs (e.g., small molecules) to enhance or inhibit the interactions between two entities (e.g., receptor and ligand, protein-protein, etc.) is assayed (e.g., by gain or loss of the

bioluminescent signal). In some embodiments, drug screening applications are carried out in a high through-put format to allow for the detection of the binding of thousands, or tens of thousands, of different molecules to a target, or to test the effect of those molecules on the binding of other entities.

In some embodiments, provided herein is the detection of molecular interactions in living organisms (e.g., bacteria, yeast, eukaryotes, mammals, primates, human, etc.) and/or cells. In some embodiments, pep peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) fused to interaction (target) polypeptides are co-expressed in a cell or whole organism, and a signal is detected in the presence of a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) and substrate (e.g., coelenterazine or coelenterazine analog), wherein the signal is correlated to the formation of the interaction complex. In some embodiments, cells are transiently and/or stably transformed or transfected with vector(s) coding for fusions comprising peptide tags and interaction elements. In some embodiments, CRISPR is utilized to generate cells that express fusions comprising peptide/dipeptide/tripeptide tags and interaction elements. In some embodiments, fusions (e.g., of a cellular target and a peptide/dipeptide/tripeptide or polypeptide described herein) generated by CRISPR replace endogenous protein (e.g., non-fused cellular target) and are regulated in a similar manner to endogenous protein. In some embodiments, such endogenous taging is used to monitor the level of the endogenously tagged protein, especially in complex systems such as live cells, whole organisms, etc. In some embodiments, transgenic organisms are generated that code for the necessary fusions (e.g., fusions comprising peptide tags and interaction elements) for carrying out the assays described herein. In other embodiments, vectors are injected into whole organisms.

In some embodiments, provided herein is the detection of molecular co-localization in living organisms (e.g., bacteria, yeast, eukaryotes, mammals, primates, human, etc.) and/or cells. In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) fused to co-localization (target) polypeptides are co-expressed in a cell or whole organism, and a signal is detected in the presence of a polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) and substrate (e.g., coelenterazine or coelenterazine analog), wherein the signal is correlated to the co-

localization of the co-localization elements. In some embodiments, cells are transiently and/or stably transformed or transfected with vector(s) coding for fusions comprising peptide tags and co-localization elements. In some embodiments, CRISPR is utilized to generate cells that express fusions comprising peptide tags and co-localization elements. In some embodiments, transgenic organisms are generated that code for the necessary fusions (e.g., fusions comprising peptide tags and co-localization elements) for carrying out the assays described herein. In other embodiments, vectors are injected into whole organisms.

In certain embodiments, cells are engineered to express one or more peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof), polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide), or fusions thereof (e.g., with cellular targets) using gene transfer technology or other engineering techniques. For example, the cells may be genetically engineered to express one or more peptide/dipeptide/tripeptide tags, polypeptide components, or fusions thereof (e.g., with cellular targets) using gene editing methodologies such as CRISPR (clustered regularly interspaced short palindromic repeat). The terms "CRISPR" or "CRISPR-Cas9," as used herein, refer to the various CRISPR-Cas9 and -CPF1 (and other) systems that can be programmed to target specific stretches of a genome and to edit DNA at precise locations. CRISPR-Cas9 gene editing systems are based on the RNA-guided Cas9 nuclease from the type II prokaryotic clustered regularly interspaced short palindromic repeats (CRISPR) adaptive immune system (see, e.g., Jinek et al., Science, 337: 816 (2012); Gasiunas et al, Proc. Natl. Acad. Set U.S.A., 109, E2579 (2012); Garneau et al., Nature, 468: 67 (2010); Deveau et al., Annu. Rev. Microbiol, 64: 475 (2010); Horvath and Barrangou, Science, 327: 167 (2010); Makarova et al., Nat. Rev. Microbiol., 9, 467 (2011); Bhaya et al., Annu. Rev. Genet., 45, 273 (201 1); and Cong et al., Science, 339: 819-823 (2013); herein incorporated by reference in their entireties). CRISPR gene editing systems have been developed to enable targeted modifications to a specific gene of interest in eukaryotic cells (see, e.g., Cong et al., supra; Xiao-Jie et al., J. Med. Genet., 52(5): 289-96 (2015); U.S. Patent 8,697,359; Xie et al., Genome Res., 24(9): 1526-1533 (2014); Huang et al., Stem Cells, 33(5): 1470-1479 (2015); Smith et al., Molecular Therapy, 23(3): 570- 577 (2015); and U.S. Patent Application Publication 2014/0068797; herein incorporated by reference in their entireties). Methods for utilizing CRISPR technology for gene editing are described in, for

example, Barrangou et al., Science 315, 1709-1712 (2007); Bolotin et al., Microbiology, 151, 2551-2561 (2005); Brouns et al., Science 321, 960-964 (2008); Cong et al., supra; Deltcheva et al., Nature 471, 602-607 (2011); Gasiunas et al., supra; Hale et al., Cell 139, 945-956 (2009); Jinek et al., Science 337, 816-821 (2012); Makarova et al., Biology Direct 2006, 1:7 (2006); Mali et al., Science 339, 823-826 (2013); Marraffini et al., Science 322, 1843-1845 (2008); Mojica et al., J Mol Evol 60, 174-182 (2005); Pourcel et al., Microbiology 151, 653-663 (2005); and Sapranauskas et al., Nucl. Acids Res. 39, gkr606–gkr9282 (2011); herein incorporated by reference in their entireties.

In some embodiments, one or more peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) are employed as a protein tag (e.g., within cells, within a whole animal). In such embodiments, the complement components to the peptide/dipeptide/tripeptide tag(s) (e.g., polypeptide components, the other peptide/dipeptide/tripeptide tag, substrate) are applied to the system (e.g., cells, animal, etc.) (e.g., as part of a reagent) to detect/quantify the presence of tagged proteins.

In some embodiments, the small size of the peptide tags herein (e.g., β9-like (e.g., SmTrip9) and β10-like (e.g., SmTrip10) peptides) is useful for protein tagging.

In some embodiments, the components of the bioluminescent complexes herein (e.g., peptide/dipeptide/tripeptide tags herein (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof), polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) are stable enough to exist in a suitable buffer for extended periods of time (e.g., in the presence of coelenterazine or a coelenterazine analog (e.g., furimazine) substrate). In certain embodiments, components of the bioluminescent complexes herein (e.g., peptide/dipeptide/tripeptide tags, polypeptide components, etc.) exhibit minimal detectable luminescence in the absence of the complementing components (e.g., even in the presence of coelenterazine or coelenterazine analog (e.g., furimazine) substrate). In some embodiments, optimized buffer conditions are utilized to meet criteria necessary for protein tagging.

The compositions and methods provided herein, as well as any techniques or technologies based thereon find use in a variety of applications and fields.

Provided herein are methods for the design and/or optimization of peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide), and the bioluminescent complexes that form therefrom. Any suitable method for the design of non-luminescent pairs/groups that are consistent with embodiments described herein, and/or panels thereof, is within the scope herein. In some embodiments, characteristics of peptide/dipeptide/tripeptide tags and polypeptide components, and combinations thereof are optimized by substitutions (e.g., substitution of natural amino acids, non-natural amino acids, amino acid analogs, etc.); such characteristics include, but are not limited to structural stability (e.g., of the peptide/dipeptide/tripeptide tag or polypeptide component, of a complex, etc.), expression, stickiness (e.g., to tubes, wells, etc.), brightness (or complexes formed therefrom), affinity for other components of the bioluminescent complex, solubility, thermal and chemical stability, low autoluminescence, etc.

In certain embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and a polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) are designed de novo to lack luminescence individually and exhibit luminescence upon association. In such embodiments, the strength of the interaction between the non-luminescent elements is insufficient to produce a bioluminescent signal in the absence of interaction elements to facilitate formation of the bioluminescent complex. In other embodiments, peptide/dipeptide/tripeptide tags and polypeptide components and/or bioluminescent complexes thereof are rationally designed, for example, using a bioluminescent protein as a starting point. For example, such methods may comprise: (a) aligning the sequences of three or more related proteins; (b) determining a consensus sequence for the related proteins; (c) providing fragments (e.g., one or more peptides/dipeptides/tripeptides and a polypeptide) of a bioluminescent protein that is related to the ones from which the consensus sequence was determined, wherein the fragments are individually substantially non-luminescent but exhibit luminescence upon interaction of the fragments; and (d) testing the fragments for the absence of luminescence when unassociated and luminescence upon association of the non-luminescent pair. In some embodiments, the

fragments are mutated at one or more positions (e.g., in vitro, in silico, etc.), wherein said mutations alter the sequences of the fragments and result in optimization of characteristics.

In some embodiments, a peptide/dipeptide/tripeptide tag is a 'dark peptide,' or one that forms a complex with the other peptide tag and polypeptide components (e.g., with low or high affinity), but produces minimal or no luminescence. In some embodiments, a high affinity dark peptide/dipeptide/tripeptide finds use in inverse complementation or gain of signal assays for biosensors or for measuring inhibitors. In some embodiments, a low affinity dark peptide/dipeptide/tripeptide is used to bring down background luminescence of a complex for the detection of binding of a high affinity bright peptide/dipeptide/tripeptide tag to the complex.

In some embodiments, a peptide/dipeptide/tripeptide tag is a 'quencher peptide,' or one that contains a quencher moiety (e.g., DAB), and the quencher absorbs the light/energy produced by either or both of a polypeptide component (e.g., the signal produced independent of a complementing peptide/dipeptide/tripeptide tags) and/or bioluminescent complex.

In some embodiments, the luminescent complexes herein find use in systems, methods, assays, devices, etc. that utilize BRET between the complex and a fluorophore (e.g., small molecule fluorophore, fluorescent protein (e.g., cyOFP)). In some embodiments, a fluorophore (e.g., small molecule fluorophore, fluorescent protein (e.g., cyOFP)) is linked or fused to an analyte, cellular target, etc. In some embodiments, a fluorophore (e.g., small molecule fluorophore, fluorescent protein (e.g., cyOFP)) is linked or fused to a peptide/dipeptide/tripeptide tag and/or polypeptide component. In some embodiments, energy is transferred from a bioluminescent complex to an energy acceptor. In certain embodiments, an energy acceptor is a fluorophore or other detectable chromophore. Suitable fluorophores include, but are not limited to: xanthene derivatives (e.g., fluorescein, rhodamine, Oregon green, eosin, Texas red, etc.), cyanine derivatives (e.g., cyanine, indocarbocyanine, oxacarbocyanine, thiacarbocyanine, merocyanine, etc.), naphthalene derivatives (e.g., dansyl and prodan derivatives), oxadiazole derivatives (e.g., pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, etc.), pyrene derivatives (e.g., cascade blue), oxazine derivatives (e.g., Nile red, Nile blue, cresyl violet, oxazine 170, etc.), acridine derivatives (e.g., proflavin, acridine orange, acridine yellow, etc.), arylmethine derivatives (e.g., auramine, crystal violet, malachite green, etc.), tetrapyrrole derivatives (e.g., porphin, phtalocyanine, bilirubin, etc.), CF dye (Biotium), BODIPY (Invitrogen), ALEXA FLOUR (Invitrogen), DYLIGHT FLUOR

(Thermo Scientific, Pierce), ATTO and TRACY (Sigma Aldrich), FluoProbes (Interchim), DY and MEGASTOKES (Dyomics), SULFO CY dyes (CYANDYE, LLC), SETAU AND SQUARE DYES (SETA BioMedicals), QUASAR and CAL FLUOR dyes (Biosearch Technologies), SURELIGHT DYES (APC, RPE, PerCP, Phycobilisomes)(Columbia Biosciences), APC, APCXL, RPE, BPE (Phyco-Biotech), autofluorescent proteins (e.g., YFP, RFP, mCherry, mKate), quantum dot nanocrystals, etc. In some embodiments, a fluorophore is a rhodamine analog (e.g., carboxy rhodamine analog), such as those described in U.S. patent application Ser. No. 13/682,589, herein incorporated by reference in its entirety. In some embodiments, a fluorophore is a small molecule fluorophore; embodiments herein reciting a fluorophore may be read as or limited to a small molecule fluorophore. In some embodiments, a fluorophore is a fluorescent protein (e.g., cyOFP, GFP, CFP, etc.; embodiments herein reciting a fluorophore may be read as or limited to a fluorescent protein (e.g., cyOFP, GFP, CFP, etc.).

In various embodiments, the bioluminescent complexes described herein, and components thereof, find use in a variety of different immunoassay concepts. For example, in some embodiments, a peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) is tethered/fused to a primary or secondary antibody to provide a method of detection for a particular analyte. As another example, a peptide tag is tethered/fused to an antibody-binding protein (e.g., protein A or protein G) and used to detect a specific antibody bound to a particular analyte (e.g., wherein the analyte is bound to the complementary peptide tag). As another example, a peptide/dipeptide/tripeptide tag is tethered/fused to streptavidin and used to detect a specific biotinylated antibody bound to a particular analyte (e.g., wherein the analyte is bound to the complementary peptide tag). As yet another example, peptide/dipeptide/tripeptide tags are tethered/fused to primary and secondary antibodies, where the primary antibody recognizes a particular analyte, and the secondary antibody recognizes the primary antibody. As still another example, a peptide/dipeptide/tripeptide tag is tethered/fused to an analyte and used in a competitive sandwich ELISA format. A peptide/dipeptide/tripeptide tag is tethered/fused conjugated to an analyte may also be used to detect antibodies capable of binding the analyte.

Various embodiments herein find use in small molecule detection via immunoassay. Exemplary embodiments comprise the use of a small molecule directly (e.g., identical or

similar to the target small molecule) labeled with a first peptide/dipeptide/tripeptide described herein and a binding moiety for the target small molecule fused or linked to a peptide/dipeptide/tripeptide described herein. In the presence of polypeptide component and substrate (e.g., coelenterazine or coelenterazine analog), a bioluminescent signal is produced by the system. When the system is exposed to a sample (e.g., biological sample, environmental sample, etc.), the bioluminescent signal will be reduced if the small molecule target is present in the sample (the labeled small molecule will be competed out of the complex allowing, in some cases, quantitation of the small molecule target). Alternative configurations for such assays are also within the scope herein. In some embodiments, the target small molecule is a toxin (e.g., mycotoxin, etc.), metabolite (e.g., amino acid, glucose molecule, fatty acid, nucleotide, cholesterol, steroid, etc.), vitamin (e.g., vitamin A, vitamin B1, vitamin B2, Vitamin B3, vitamin B5, vitamin B7, vitamin B9, vitamin B12, vitamin C, vitamin D, vitamin E, vitamin H or vitamin K, etc.), coenzyme or cofactor (e.g., coenzyme A, coenzyme B, coenzyme M, coenzyme Q, cytidine triphosphate, acetyl coenzyme A, reduced nicotinamide adenine dinucleodtide (NADH), nicotinamide adenine (NAD+), nucleotide adenosine monophosphoate, nucleotide adenosine triphosphate, glutathione, heme, lipoamide, molybdopterin, 3'-phosphoadenosine-5'-phsphosulfate, pyrroloquinoline quinone, tetrahydrobiopterin, etc.), biomarker or antigen (e.g., erythropoietin (EPO), ferritin, folic acid, hemoglobin, alkaline phosphatase, transferrin, apolipoprotein E, CK, CKMB, parathyroid hormone, insulin, cholesteryl ester transfer protein (CETP), cytokines, cytochrome c, apolipoprotein AI, apolipoprotein AII, apolipoprotein BI, apolipoprotein B-100, apolipoprotein B48, apolipoprotein CII, apolipoprotein CIII, apolipoprotein E, triglycerides, HD cholesterol, LDL cholesterol, lecithin cholesterol acyltransferase, paraxonase, alanine aminotransferase (ALT), asparate transferase (AST), CEA, HER-2, bladder tumor antigen, thyroglobulin, alpha-fetoprotein, PSA, CA 125, CA 19.9, CA 15.3, leptin, prolactin, osteoponitin, CD 98, fascin, troponin I, CD20, HER2, CD33, EGFR, VEGFA, etc.), drug (cannabinoid (e.g., tetrahydrocannabinol (THC), cannabidiol (CBD) and cannabinol (CBN), etc.), opioid (e.g., heroin, opium, fentanyl, etc.), stimulant (e.g., cocaine, amphetamine, methamphetamine, etc.), club drug (e.g., MDMA, flunitrazepam, gama-hydroxybutyrate, etc.), dissociative drug (e.g., ketamine, phencyclidine, salvia, dextromethorphan, etc.), hallucinogens (e.g., LSD, mescaline, psilocybin, etc.), etc.), explosive (e.g., 2,4,6-trinitrotoluene (TNT) and hexahydro-1,3,5-trinitro-1,3,5-triazine

(RDX), pentaerythritol tetranitrate (PETN), etc.), toxic chemical (e.g., tabun (GA), sarin (GB), soman (GD), cyclosarin (GF), 2-(dimethylamino)ethyl N, N-dimethylphosphoramidofluroidate (GV), VE, VG, VM, VP, VR, VS, or VX nerve agent), etc.

The systems and methods described herein find use in a wide variety of applications and formats. The following are non-exhaustive exemplary examples of methods and formats utilizing the systems described herein.

- In some embodiments, provided herein are intracellular two protein systems for dynamic protein-protein interaction analysis with SmTrip peptide-labeled proteins expressed as fusions via traditional transfection or endogenously tagged proteins via CRISPR; LgTrip can be used as a detection reagent either by co-transfection, of LgTrip, providing a stable cell line expressing LgTrip, or providing LgTrip in the detection reagent and adding it to lysed cells expressing SmTrip-labeled proteins.(Figure 51A).
- In some embodiments, provided herein are intracellular three protein systems for dynamic protein-protein interaction analysis with SmTrip- and LgTrip-labeled proteins expressed as fusions via traditional transfection or as endogenously-tagged proteins generated via CRISPR (Figure 51B).
- In some embodiments, provided herein are target specific assays to measure analyte X with binding moiety A and binding moiety B (See Table A; purified genetic fusions or chemically conjugated SmTrip9 or SmTrip10 peptide) for a gain of signal assay (e.g. diagnostic test, non-cellular, etc.) (Figure 51C).
- In some embodiments, provided herein are target specific competition assays for analyte measurement through loss of signal (e.g. diagnostic test, noncellular, etc.) (Figure 51D). Such a system use a purified binding moiety A (e.g., purified genetic fusion or chemically conjugated comprising synthetic SmTrip9 or SmTrip10 peptide) that binds the tagged target analyte to generate light in the presence of LgTrip and a coelenterazine substrate or coelenterazine analog, which may be provided as part of a detection reagent. In the presence of sample analyte X, SmTrip9 or SmTrip10 peptide will compete with the sample analyte X to cause a loss of signal specific to the presence of the sample analyte in the sample and proportional to the concentration of the analyte.

- In some embodiments, two or three of the peptide tags peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) are linked (or fused) to recognition elements for proximal, but non-overlapping (mutually exclusive or distinct), epitopes on the same target analyte. A signal generated from the luminescent complex (e.g., in the presence of a substrate) indicates the presence of the target analyte.
- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in immunoassays or various formats. Immunoassays employing the peptide/dipeptide/tripeptide tags and polypeptide components herein are not limited to full length antibodies and may also employ antibody fragments or non-antibody binding moieties (e.g., DARPins, aptamers, affimers, etc.). In an exemplary direct immunoassay (See, e.g., Figure 51E), two monoclonal or recombinant antibodies (mAbs or rAbs) against an analyte are labeled with β9-like (e.g., SmTrip9) and β10-like (e.g., SmTrip10) peptide tags; a polypeptide component (e.g., β1-8-like (e.g., LgTrip) polypeptide) of the luminescent complex is included as part of detection reagent (e.g., with substrate). For an exemplary indirect immunoassay (See, e.g., Figure 51F), generic reagents labeled with β9-like (e.g., SmTrip9) and β10-like (SmTrip10) peptide tags are used in combination with any paired antibody system specific to an analyte (e.g., mAb or rAb +Biotin-pAb, Biotin-mAb, or Biotin-rAb etc.); a polypeptide component of the luminescent complex is included as part of detection reagent (e.g., with substrate). An exemplary competition direct immunoassay (See, e.g., Figure 51G) is provided by labeling one antibody with a first peptide tag (β9- (e.g, SmTrip9) or β10-like (e.g., SmTrip10) peptide) and labeling a analyte with a second peptide tag (β10- (e.g., SmTrip10) or β9-like (e.g., SmTrip9) peptide); a polypeptide component (e.g., β1-8-like (e.g., LgTrip) polypeptide) of the luminescent complex is included as part of detection reagent (e.g., with substrate); loss of signal indicates the presence of unlabeled target analyte. To provide a competition indirect immunoassay (See, e.g.,

Figure 51H), one antibody is labeled with a first peptide tag (β9-(e.g., SmTrip9) or β10-like (e.g., SmTrip10) peptide), a generic binding reagent (e.g., streptavidin) is labeled with a second peptide tag (β10-(e.g., SmTrip10) or β9-like (e.g., SmTrip9) peptide), and analyte is labeled with a binding moiety for the generic binding reagent (e.g., biotin); a polypeptide component of the luminescent complex is included as part of detection reagent (e.g., with substrate); loss of signal indicates the presence of unlabeled test analyte. Alternative immunoassays utlizing other peptide/dipeptide/tripeptide/polypeptide combinations described herein are within the scpe of the present invention.

- In some embodiments, provided herein are homogeneous assays using peptide/dipeptide/tripeptide tag-labelled (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) recognition elements with the polypeptide component (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) as component of a detection reagent (e.g., along with a substrate for the luminescent complex).
- In some embodiments, provided herein are homogeneous assays utilizing peptide/dipeptide/tripeptide-tag-labelled (e.g., SmTrip9, SmTrip10, etc.) and/or polypeptide-component-labelled (e.g., LgTrip variants) recognition elements. In some embodiments, homogeneous assays are provided for detection/quantification of a single analyte or multiple analytes.
- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in sandwich hybridization assays (e.g. non-target amplified, amplified, etc.).
- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in the detection of analyte(s) in liquid/solution phase or solid phase.
- In some embodiments, the pepeptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or

dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in surface-based assays (e.g., plate-based (e.g., microtiter plate), paper-based (e.g., Whatman protein saver 903 cards), plastic-based, swab-based, cuvette-based, membrane-based (e.g., PVDF, nitrocellulose, etc.), etc.

- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in lateral flow and other capillary driven based methods. In some embodiments, such lateral flow assays allow multiplexed detection/identification/characterization of analytes (e.g., pathogens). In some embodiments, lateral flow assays find use in performing immunoassays described herein.
    - An exemplary multiplexed tripartite lateral flow assay for the detection and identification of pathogens using tripartite antibody fusions in a direct immunoassay is depicted in Figure 52. In this example, a set of monoclonal or recombinant antibodies (mAbs or rAbs), each fused to a peptide tag (e.g., β10-like (e.g., SmTrip10) peptides) are added to a liquid sample, the sample is passed over a detection window comprising a second set of mAbs or rAbs, each fused to a peptide tag (e.g., β9-like (e.g., SmTrip9) peptides), immobilized in lanes within the detection window, and each recognizing a distinct epitope on the same target as one of the mAbs or rAbs in the liquid sample. When the liquid sample is passed through the detection window in the presence of a polypeptide component and substrate (e.g., preloaded in the detection window, added with the sample, added separately to the device, etc.), luminescence in a particular lane indicates the binding of mAbs or rAbs to separate epitopes on a target, and thereby provide for detection and identification of the target. The above described assay, and alternatives thereof utilizing the systems and methods herein, may find use in providing various detection panels (e.g., Respiratory Panel: Streptococcus, Pseudomonas, Mycobacterium, Staphylococcus; Urinary Tract Panel: E. Coli, Klebsiella, Enterobacter, Streptococcus; Food Borne Panel: Shigella,

Campylobacter, Salmonella, E. Coli, Listeria; Waste Water Management: Coliform panel; Panel for strain identification within one type of bacteria; etc.), as well as for other applications (e.g., toxin detection).

- An exemplary multiplexed tripartite lateral flow assay for the detection and identification of anti-viral antibodies, e.g., for disease diagnosis, using tripartite antibody fusions in a direct immunoassay is depicted in Figure 53. In this example, a sample is added to the lateral flow device and allowed to flow into a conjugation zone (e.g., pad). The conjugation zone comprises a generic antibody-binding agent (e.g., Protein L), tethered or fused to a first peptide tag (e.g., β10-like (e.g., SmTrip10) peptide). If antibodies are present in the sample, they will be bound by the labeled antibody-binding agent. A detection window of the device comprises separate lanes, each comprising distinct immobilized viral antigens tethered or fused to a second peptide tag (e.g., β9-like (e.g., SmTrip9) peptide). As the labeled antibody flows from the conjugation zone into the detection window, the antibodies will bind to appropriate antigens, binging the peptide tags into proximity and producing a luminescent signal in the presence of the polypeptide component and substrate (e.g., preloaded in the detection window, added with the sample, added separately to the device, etc.). Such a device and assay would allow detection and discrimination of multiple viruses and viral infections using a single device/assay. For example, Zika, Dengue, and Chicungkunga could all be independently detected using a single test.

- In some embodiments, the details of the above lateral flow assays are carried out in a plate-based format for solution phase assay (e.g., with the binding moiety combinations in wells provided with a map). In some embodiments, such an assay is performed in a multiplexed dot blot/spot array assay format. In some embodiments, any multiplexed assays described herein in a particular format (e.g., lateral flow) may also be performed in alterantive formats described herein or understood in the field (e.g., dot blot, spot array, etc.).
- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-

like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in aerosol-based detection (e.g., (1) protease to lyse cells, (2) spray detection reagents, (3) visualize to detect/quantify).

- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use with isothermal amplification of nucleic acids.
- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use with rapid cycling PCR detection of nucleic acids.
- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in the detection of protein-protein interaction (e.g., between 2 proteins, between 3 proteins, etc.).
- In some embodiments, analysis of the assays and methods described herein is performed using stationary or portable devices and readers, a luminometer plate reader or handheld reader, smart phone camera or CCD camera, etc.
- In some embodiments, analyte is detected/quantified via gain of signal through recognition elements via luminometer or imaging based techniques.
- In some embodiments, analyte is detected/quantified measured via loss of signal through competitive displacement via luminometer or imaging based techniques.
- In some embodiments, systems and methods herein allow for multiple tags on each recognition element, either genetically or through chemical conjugation, thereby providing signal amplification by adding increasing stoichiometry of peptide tags per recognition element.
- In some embodiments, systems and methods herein find use in detection of native proteins in heterogeneous solutions.

- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in nucleic acid detection, for example, peptide-tagged complimentary recognition elements hybridize to a nucleic acid target sequence in tandem.
- In some embodiments, assays are provided herein for the detection of an antibody (e.g., antibody as analyte). One such assay is depicted in Figure 54. A first peptide tag (e.g., β9-like (e.g., SmTrip9) peptide) is fused or tethered to an antibody-binding protein (e.g., Protein L) and a second peptide tag (e.g., β10-like (e.g., SmTrip10) peptide) is fused or tethered to the analyte that the antibody is specific to; a polypeptide component (e.g., β1-8-like (e.g., LgTrip) polypeptide) of the luminescent complex is included as part of detection reagent (e.g., with substrate); presence of the analyte specific antibody in a sample results in complex formation and luminescence.
- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in FISH-like applications utilizing bioluminescence or BRET for detection/quantification.
- In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in detection of nucleic acids (e.g., single stranded and/or double stranded DNA and/or RNA) via, for example amplification-free detection of nucleic acids. For example, as depicted in Figure 56, a pair of peptide tag-labelled nucleic acid probes, when hybridized to nearby locations on a nucleic acid target, will allow formation of a luminescent complex, facilitated by complementation with the nucleic acid target. Such and assay could be performed on a solid surface, in solution, for a single nucleic acid target, or multiplexed (e.g., using an array).
- In some embodiments, peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides

and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use in lab-on-chip and/or microfluidics applications.

- In some embodiments, systems and methods herein find use in heterogeneous assays, such as, immunoassays (e.g., PCR amplification combined with homogeneous immunoassay analysis).

In some embodiments, a peptide/dipeptide/tripeptide-based sensor is provided is which a chemical (e.g., removal of a blocking moiety) or enzymatic (e.g., proteolytic cleavage) event is required to render a peptide tag capable of bioluminescent complex formation. For example, a protease is required to cleave a blocked dipeptide (e.g., incapable of bioluminescent complex formation) into two non-blocked peptides capable of complementation. In some embodiments, the peptide/dipeptide/tripeptide tags (e.g., β6-like, β7-like, β8-like, β9-like (e.g., SmTrip9), and/or β10-like (e.g., SmTrip10) peptides, and/or dipeptides and tripeptides thereof) and polypeptide components (e.g., β1-5-like, β1-6-like, β1-7-like, β1-8-like (e.g., LgTrip) polypeptide) herein find use with bead-based assays, utilizing magnetic enrichment for increased assay sensitivity. One such assay is depicted in Figure 55. In such an assay, a magnetic particle is conjugated to a first peptide tag (e.g., β9-like (e.g., SmTrip9) peptide) and to a first binding agent directed to a first epitope on an analyte; a non-magnetic particle (e.g., polystyrene particle) is conjugated to a second peptide tag (e.g., β10-like (e.g., SmTrip10) peptide) and to a second binding agent directed to the first epitope or a second epitope on an analyte. The beads are combined with a sample, along with a polypeptide component (e.g., β1-8-like (e.g., LgTrip) polypeptide) of the luminescent complex. Magnetic separation is used to capture the magnetic beads and any components of the sample or other reagents bound thereto. Luminescence of the magnetically-captured elements is then detected in the presence of substrate for the luminescent complex. If the analyte is present in the sample, both the magnetic and non-magnetic beads will be captured, resulting in the capture of the luminescent complex. In the absence of analyte, the non-magnetic beads will not be captured, and the luminescent complex will not be formed. The above applications and formats are exemplary and non-limiting. Other embodiments consistent with the description herein are within the scope of the present invention. Systems comprising and method utilizing peptides, dipeptides, and polypeptides bearing structural (although not necessarily sequence identity) and functional correlation to portions of

NanoLuc® commercial luciferase and/or natural luciferase from *Oplophorus gracilirostris*, and bioluminescent complexes formed by complementation thereof, are described herein. In particualr, detailed description is provided of complementaion between β_{1-8}-like (e.g., LgTrip) polypeptides and either β_9-like (e.g., SmTrip9) and β_{10}-like (e.g., SmTrip10) peptides or $\beta_{9/10}$-like dipeptides. However, embodiments herein are not limited to complementaion between β_{1-8}-like polypeptides (e.g., LgTrip) and β_9-like (e.g., SmTrip9) and β_{10}-like (e.g., SmTrip10) peptides or $\beta_{9/10}$-like dipeptides. In some embodiments, eptides, dipeptides, and polypeptides bearing structural (although not necessarily sequence identity) and functional correlation to portions of NanoLuc® commercial luciferase and/or natural luciferase from *Oplophorus gracilirostris* are provided. For example, also provided herein are systems and methods for complementation between a β_{1-5}-like polypeptide and β_{6-10}-like polypeptide; between a β_{1-2}-like dipeptide and β_{3-10}-like polypeptide; between a β_1-like peptide, β_2-like peptide and β_{3-10}-like polypeptide; between a β_{7-8}-like dipeptide and $\beta_{9-10/1-6}$-like polypeptide fusion; between a β_{1-7}-like polypeptide and β_8-like, β_9-like, and β_{10}-like peptides; and/or between a β_{1-6}-like polypeptide and β_7-like, β_8-like, β_9-like, and β_{10}-like peptides.

In some embodiments, the peptides, dipeptides, triprptides, and/or polypeptides herein find use in translocation assays. In some embodiments, a translocation assay is composed of two components: a complementary polypeptide sensor (e.g., LgBiT-based, LgTrip-based, etc.) and a peptide/dipeptide/tripeptide-tagged protein of interest (POI). A variety of LgBiT sensors were genrated that localize at specific cellular compartments such as plasma membrane, nucleus, mitochondria and endoplasmic reticulum (ER) (Figure 152). These LgBiT sensors can be introduced to cells via transfection or establishment of stable cell lines. The POI is endogenously tagged with peptide/dipeptide/tripeptide complementary to the polypeptide (e.g., LgBiT) . Under stimuli, the POI translocates to a different cellular compartment where the polypeptide (e.g., LgBiT) sensor resides, complementation occurs leading to the assembly of pepdtide/polypeptide compl;ex (e.g., HiBiT•LgBiT) to yield luminescence signal (Figure 153). Thus, the translocation activity of POI is quantitatively measured via luminescence output. Experiments conducted during development of embodiments of the translocation assay are described in Example 89.

As described further herein, translocation assays can be designed as part of a modular system that includes, for example, the following components: 1) Cell lines expressing HiBiT-

tagged proteins from endogenous loci; and 2) a LgBiT localization sensor. This principal design allows for the correlation of a luminescent signal with the presence of the target protein at a specific cellular localization. The sensors can include localization-specific sequences fused to a LgBiT affinity variant, as well as to HaloTag for validation and imaging of the target compartment. Localization-specific sequences can be short peptides that are derived from known protein cellular markers. In some embodiments, when short peptides are not specific enough to exclude the LgBiT sensor to the cellular compartment of choice, full length protein marker can be used to fuse to the sensor. DNA sequences encoding LgBiT localization sensors can be cloned into a circular double-stranded DNA plasmid, which can be delivered into HiBiT cell line. In some embodiments, the delivery method can be via lipid based transfection using Fugene HD transfection reagent, or via viral transduction using Lenti viral particles or BacMam. In some embodiments, when transfected/transduced, cells are treated by an agonist, and HiBiT-tagged protein translocates to the cellular compartment where LgBiT sensor resides; complementation occurs and produces luminescence. In some embodiment, the agonists are small molecules that stimulate translocation of the target protein.

EXPERIMENTAL

Example 1

Further truncated version of (LgBiT) is activated by peptide

The luminescence of LgBiT (background and in the presence of complementary SmTrip10 pep86)(SEQ ID NO: 15, 25) was compared with a further truncated polypeptide (LgTrip 2098, SEQ ID NO: 17) lacking both the β10 and β9 strands of the full-length luciferase (background and in the presence of complementary pep263 (SEQ. ID 35)) (Figure 1).

E. coli KRX harboring LgBiT (SEQ ID NO: 11) or LgTrip 2098 (SEQ ID NO: 17) were grown for 20 h from a single colony in LB+amp (50 ug/mL) at 30ºC (275 rpm) in a volume of 50 mL. From these cultures, 100x dilutions were made into the same media and the cultures grown at 37C (275 rpm) for 3 h and then cooled to 25ºC before adding rhamnose (inducing agent for protein overexpression) to a final concentration of 0.2%. Cultures were then grown (induced) for 22 h at 25ºC (275 rpm) at which time cultures were harvested, and the resulting pellets stored at -20ºC until processing. To lyse cells, pellets were removed from

-20°C, resuspended in 50 mL of PBS pH 7.2, and taken through 3 sequential freeze thaw cycles (-70°C to 22°C), centrifuged to produce soluble fractions, and then kept cold (on ice) until assaying. Lysates and peptide(s) (25 nm final concentration) were incubated together for 10 minutes at 25°C prior to the addition of NanoGlo® reagent . After addition of reagent, plates were incubated for another 5 min at 35°C and read over time to measure luminescence (RLU) using a Tecan Infinite F500 plate reader.

Experiments conducted during development of embodiments herein demonstrate that both LgBiT (SEQ ID NO: 11) and LgTrip 2098 (SEQ ID NO: 17) produce some background luminescence, but the level is much higher for LgBiT. Data shows that both LgBiT and LgTrip 2098 produce more luminescence in the presence of their respective complementary peptide. The magnitude of the gain in signal in the presence of peptide is greater for LgTrip 2098. These data demonstrate that the further truncated LgBiT (and with the A51G substitution) is activated by a single complementary peptide corresponding to the β10 and β9 beta strands that are absent from LgTrip 2098.

Example 2

LgTrip 2098 is activated by pair of separate β9 and β10-like peptides

The luminescence of LgTrip 2098 (SEQ ID NO: 31) was monitored over time in the presence of separate peptides corresponding to the β10 and β9 portions of the full-length luciferase (SmTrip10 pep86 (SEQ ID NO: 25) and SmTrip9 pep245) (SEQ ID NO: 23) (Figure 2). Similar experimental protocols were used as in Example 1; however, a 10x concentrated lysate was used, Peptides SmTrip10 pep86 and SmTrip9 pep245 were used at 500 nM, and 0.001% Prionex added to reactions. Experiments conducted during development of embodiments herein demonstrate that LgTrip 2098 (SEQ ID NO: 31) is activated by the addition of SmTrip10 pep86 and SmTrip9 pep245. Controls with no peptides added or only one of the peptides added produced near the background of the plate reader.

Example 3

LgTrip mutagenesis – round 1 (luminescence)

Overnight cultures used for sequencing were used to inoculate cultures (30ul of cells in 3ml of media+0.1% Rhamnose+0.15% glucose). Cells were grown overnight at 25°C for 20 hours. Cells were diluted 10ul into 250ul of Passive Lysis Buffer (PLB) and allowed to lyse for 5 minutes. The lysate was mixed and then diluted 1:100 into PLB lysis buffer (0.3X PLB, 25mM HEPES pH 7.5, 0.001U/ml RQ DNase 1 (10ul in 990ul). 50ul of the diluted

lysate was combined with 50ul of NanoGlo® buffer + 2uM pep263 (SmTrip9-10 dipeptide) (SEQ ID NO: 35) at a 1uM final concentration (saturating dipeptide concentration). Samples were incubated for 5 minutes, read on GloMax® Multi+ (GMM+) luminometer, and normalized to LgTrip 2098 (SEQ ID NO: 31) (Table 2).

Table 2. Relative luminescence of LgTrip variants compared to LgTrip 2098.

Clone #	Cell plate	Sequence	Secondary screen(normalize to 2098)
#7		F1L	1.1
#10		Q42L	1.8
#14		I44V, E63D, L142Q	2.5
#16		L3OS	8.3
#19		N17D	3.0
#22		Y16C, I56T	2.6
#35		L142Q	1.8
#38		T2S, M106K	1.8
#39		E4D, V27A	3.7
#42		E4D	2.5

Example 4

LgTrip mutagenesis – round 1 (stability)

Experiments were conducted during development of embodiments herein to determine the stability of HisLink purified LgTrip mutants. LgTrip 2098 (SEQ ID NO: 31) , LgTrip 2098 (RH) (SEQ ID NO: 31) (column purified LgTrip 2098), #10, #14, #16, #19, #22, #35, #38, #39, and #42 polypeptides (Table 3) were diluted 1:1000 into PLB lysis buffer (2ul into 2ml). 100ul of each sample was transferred into one column of wells in a 96-well PCR tray. Samples were incubated at 37°C, and aliquots were remove at various time-points. Samples were placed on ice when thermal treatment was complete. When all samples were processed, the PCR tray was equilibrated to room temperature. Samples were mixed and then diluted 1:100 in PLB lysis buffer (5ul into 495ul buffer). 50ul of each sample was combined with 50ul of NanoGlo® buffer reagent + 2uM pep263. The plate was incubated for 5 minutes and then read on GMM+. Results are depicted in Figure 3. Stability studies

identified position 42 of LgTrip 2098 (SEQ ID NO: 31) as a position of interest for further analysis.

Table 3. Experimental nomenclature for LgTrip mutants (mutations relative to LgTrip 2098).

Clone #	Sequence
#10	Q42L
#14	I44V, E63D, L142Q
#16	L30S
#19	N17D
#22	Y16C, I56T
#35	L142Q
#38	T2S, M106K
#39	E4D, V27A
#42	E4D

Example 5

Position 42 site saturation (luminescence)

Experiments were conducted during development of embodiments herein to optimize the identity of the amino acid at position 42 of LgTrip 2098 (SEQ ID NO: 31) (Figure 4). *E. Coli* cultures (3ml) were prepared for each sample and grown overnight at 37°C in LB media+100ug/ml ampicillin. Cultures were then diluted in quadruplicate at a 20X concentration (10µl in 200µl) into induction media (LB+ampicllin+0.1% Rhamnose). Samples were grown at 37°C for 6 hours. Samples were then lysed with 0.3X PLB, 25mM HEPES pH 7.5, and 0.001U/ml RQ1 DNase (10µl of cells to 250µl of Lysis buffer). 50µl of the lysate was then combined with 50µl of NanoGlo® buffer+50µM furimazine+20nM of dipeptide 263 (SEQ ID NO: 35). Samples were measured on a BMG Clariostar luminometer. RLU values were normalized to LgTrip 2098 (SEQ ID NO: 31).) (Figure 4)

Example 6

37°C stability of purified LgTrip position 42 mutants

Experiments were conducted during development of embodiments herein to determine the stability of position 42 in LgTrip 2098 mutants (Figure 5). Polypeptides were diluted to 20nM in TBS+0.01% BSA. In triplicate, 100µl aliquots of each sample were

loaded into 200µl thin walled PCR tubes. Samples were incubated at 37°C in thermal cycler. Samples were removed at various time-points, placed on ice, and then allowed to equilibrate to room temperature. Samples were diluted to 0.2nM (5 in 495µl) in PLB lysis buffer (0.3X PLB+25mM HEPES pH 7.5). 50µl of each diluted sample was combined with 50µl of 50µM Furimazine+6µM pep263 (SEQ ID NO: 35) in NanoGlo® buffer. Samples were incubated for 10 minutes and then read on GMM+. Half-life was calculated by non-linear regression (Figure 5).

Example 7

Site saturation of LgTrip

Experiments were conducted during development of embodiments herein to optimize the identity of the amino acid at various positions of LgTrip 2098 (SEQ ID NO: 31) (Figure 6). *E. coli* cultures (3ml) were prepared for each sample and grown overnight at 37°C in LB media+100µg/ml ampicillin. Cultures were then diluted in quadruplicate to a 20X concentration (10µl in 200µl) into induction media (LB+ampicllin+0.1% Rhamnose). Samples were grown at 37°C for 6 hours. Samples were then lysed with 0.3X PLB+25mM HEPES pH 7.5+0.001U/ml DNase (10µl of cells to 250µl of Lysis buffer). 50µl of the lysate was then combined with 50µl of NanoGlo® buffer+50µM furimazine+20nM of dipeptide 263 (SEQ ID NO: 35). Samples were measured on a BMG Clariostar luminometer. RLU values were normalized to LgTrip 2098 (SEQ ID NO: 31)(Figure 6).

Example 8

Mutations on LgTrip 3092 template

Experiments were conducted during development of embodiments herein to determine the effect of various amino acid substitutions relative to the LgTrip 3092 (SEQ ID NO: 19) variant (Table 4). *E. coli* cultures (3ml) were prepared for each sample and grown overnight at 37°C in LB media+100ug/ml ampicillin. Cultures were then diluted in quadruplicate to a 20X concentration (10µl in 200µl) into induction media (LB+ampicllin+0.1% Rhamnose). Samples were grown at 37°C for 6 hours. Samples were then lysed with 0.3X PLB+25mM HEPES pH 7.5+0.001U/ml DNase. (10µl of cells to 250µl of Lysis buffer). 50µl of the lysate was then combined with 50µl of NanoGlo® buffer+50µM furimazine+2nM of dipeptide 263 (SEQ ID NO: 35). The mutant lysates were further diluted 1:100 in PLB (5µl in 495µl). 50µl of the diluted lysate was added to 50µl of NanoGlo® buffer+50µM furimazine+6µM pep263 or 50µl of TBS+20µM LCS (furimazine)+6µM

pep263 (SEQ ID NO: 35). Samples were measured on a GMM+ after a 10 minute incubation. RLU values were normalized to LgTrip 3092 (SEQ ID NO: 19)

Table 4. Relative luminescence of LgTrip variants compared to LgTrip 3092.

Sample	LCS 6uM 263	Nglo 6uM 263	Nglo 2nM 263
ATG 3092	1.0	1.0	1.0
V127A	4.2	3.5	3.7
Y16S	1.3	1.3	3.1
V119A	1.4	1.2	2.0
V127A+T128A	4.9	3.7	3.7
I117N	2.9	2.1	2.5
F120S	2.0	1.6	2.1
G122D	1.2	1.3	2.3
N105S	1.2	1.4	1.6
T126S	2.3	1.4	2.1
G101E	2.3	1.4	3.9
V36E+V102D+E115D	2.4	1.4	1.7

Example 9

Site saturation of LgTrip 3092 template

Experiments were conducted during development of embodiments herein to optimize the identity of the amino acid at various positions of LgTrip 3092 (SEQ ID NO: 19). *E. coli* cultures (3ml) were prepared for each sample and grown overnight at 37°C in LB media+100ug/ml ampicillin. Cultures were then diluted in quadruplicate to a 20X concentration (10µl in 200µl) into induction media (LB+ampicllin+0.1% Rhamnose). Samples were grown at 37°C for 6 hours. Samples were then lysed with 0.3X PLB+25mM HEPES pH 7.5+0.001U/ml DNase (10µl of cells to 250µl of Lysis buffer). 50µl of the lysate was then combined with 50µl of NanoGlo® buffer+50µM furimazine+2nM of dipeptide 263 (1nM final). Samples were measured on a BMG Clariostar luminometer. RLU values were normalized to LgTrip 3092 (Figure 7).

Example 10

Stability of LgTrip 2098 and LgTrip 3092 compared to LgBiT

Experiments were conducted during development of embodiments herein to compare the stability of positions in LgTrip 2098 (SEQ ID NO: 31) and LgTrip 3092 (SEQ ID NO:

33) to LgBiT (SEQ ID NO: 11) (Figure 8). Purified LgTrip 2098, LgTrip 3092, and LgBiT samples were diluted to 20nM in TBS+0.01% BSA. 100µl of each sample was aliquoted into 200µl thin walled PCR tubes. Samples were incubated at 37°C in thermal cycler. Samples were removed at various time-points and placed on ice. Samples were equilibrated to RT and then diluted to 0.2nM (5µl in 495µl) in PLB lysis buffer (0.3X PLB+25mM HEPES pH 7.5). 50µl of each sample was diluted with 50µl of 50µM Furimazine +6µM pep263 (SEQ ID NO: 35) in NanoGlo® buffer. Samples were incubated for 10 minutes and then read on GMM+ (Figure 8).

Example 11

Stability of LgTrip variants at 42°C and 60°C

Experiments were conducted during development of embodiments herein to compare the stability of LgTrip variants at 42°C and 60°C (Figure 9). The LgTrip variant samples were diluted to 20nM in TBS+0.01% BSA. 100µl aliquots were added into 200µl thin walled PCR tubes. Samples were incubated at 42°C or 60°C in thermal cycler. Samples were removed at various time-points and placed on ice. Samples were equilibrated to RT and then each diluted to 0.2nM (5µl in 495µl) in PLB lysis buffer (0.3X PLB+25mM HEPES pH 7.5). 50µl of each diluted sample was combined with 50µl of 50uM Furimazine+6µM pep263 (SEQ ID NO: 35) in NanoGlo® buffer. Samples were incubated for 10 minutes and then read on GMM+ (Figure 9).

Example 12

Affinity of LgTrip variants with SmTrip9 and SmTrip10

Experiments were conducted during development of embodiments herein to determine the affinity of various LgTrip variants for the SmTrip9- and SmTrip10-like peptides (Figure 10).

For the SmTrip9 pep286 (SEQ ID NO: 37) titration, purified LgTrip samples were diluted to 2nM in TBS+0.01% BSA+0.005% IGEPAL. Assay reagent containing TBS+0.01% BSA+0.005% IGEPAL+20µM furimazine+200µM SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) was prepared. 4uM of SmTrip9 pep286 (SEQ ID NO: 37) was added to the assay reagent and then serially diluted 500µl to 500µl in assay reagent containing Furimazine+200µM SmTrip10 pep86 (HiBiT; SEQ ID NO: 15). 25ul of each peptide titration was added to 25ul of diluted LgTrip solution. Luminescence was read on a plate reader at 10 and 15 minutes.

For the SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) titration, purified LgTrip samples were diluted to 2nM in TBS+0.01% BSA+0.005% IGEPAL. Assay reagent containing TBS+0.01% BSA+0.005% IGEPAL+20µM furimazine+4µM SmTrip9 pep286 (SEQ ID NO: 37) was prepared. 200uM of SmTrip10 pep86 (SmHiTrip; SEQ ID NO:25) was added to then assay reagent and then serially diluted 500µl to 500µl in assay reagent containing Furimazine+4µM SmTrip9 pep286 (SEQ ID NO: 37). 25ul of each peptide titration was added to 25ul of diluted LgTrip solutions. Luminescence was read on plate reader at 10 and 15 minutes.

Example 13

Stability of LgTrip variants (60°C)

Experiments were conducted during development of embodiments herein to compare the stability of LgTrip variants at 60°C (Figure 11). Purified LgTrip mutants were diluted to 20nM in TBS+0.01% BSA. 100µl of each sample was aliquoted into 200µl thin walled PCR tubes. Samples were incubated at 60°C in thermal cycler and then were removed at various time-points, placed on ice, equilibrated to room temperature, and then diluted to 0.2nM (5µl in 495µl) in PLB lysis buffer (0.3X PLB+25mM HEPES pH 7.5). 50µl of each diluted sample was combined with 50µl of 50µM Furimazine+6µM pep263 (SEQ ID NO: 35) in NanoGlo® buffer. Samples were incubated for 10 minutes and then read on GMM+. Half-life was calculated using GraphPad Prism non-linear regression (One-Phase Decay plateau set to zero).

Example 14

Comparison of kinetic profiles of LgBiT and LgTrip variants

Experiments were conducted during development of embodiments herein to compare the kinetic profiles of various LgTrip variants with NanoLuc® luciferase (SEQ ID NO: 3) and a LgBiT (SEQ ID NO: 11)/HiBiT (SmTrip10 pep86) (SEQ ID NO: 25) two component system (Figure 12). NanoLuc, LgBiT, LgTrip 2098 (SEQ ID NO: 31), LgTrip 3092 (SEQ ID NO: 33), and LgTrip 3546 (SEQ ID NO: 51) were diluted to 20µM in TBS+0.01% BSA+0.005% IGEPAL. Samples were diluted 1:100 (2µl in 198µl) and then 1:1000 (10µl in 10ml) or $10E^{-5}$ dilution total to a 0.2nM final concentration. To the LgBiT polypeptide, 200nM of SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) was added. To the LgTrip variants, 200µM SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) and 20µM SmTrip9 pep286 (SEQ ID NO: 37) were added. Samples were incubated for 15 minutes, 50ul of each enzyme/peptide

dilution combined with either TBS+0.01% BSA+20μM Live Cell Substrate (LCS; Promega Cat. No. N205) or NanoGlo® buffer+50μM Furimazine, and immediately read on a GMM+ luminometer.

Example 15

Detecting protein-protein interactions with a tripartite system

Experiments were conducted during development of embodiments herein to demonstrate the use of a tripartite complementation system in detecting protein-protein interactions (Figure 13). Lysates containing FRB and FKBP fused to one each of SmTrip9 pep245 and SmTrip10 pep86 (Figure 13A) or SmTrip9 pep245 and SmBiT (Figure 13B) were added to purified LgTrip 2098 (SEQ ID NO: 31). Formation of the FRB/FKBP complex was induced with rapamycin and facilitated complementation of the tripartite system was monitored by luminescence.

Cell and lysate preparation. Cultures of each FRB-FKBP construct were grown overnight in LB+100ug/ml ampicillin. Cultures were induced (30μl in 3ml of culture) in LB+0.1% rhamnose+0.15% glucose+100ug/ml ampicillin and grow for 24 hours at 25°C. 200μl of 10X Fastbreak Cell Lysis Reagent (Promega) was added to 2ml of culture 0.001U/ml RQ1 DNase. Cultures were incubated for 30 min at 4°C on a rotating mixer and then spun at 3500 rpm for 30 min. at 4°C. Cleared lysate was removed and placed into new tubes, frozen, and stored at -70°C.

Assay. Lysates were thawed, diluted 1:10 into TBS+0.1% BSA, and appropriate lysates combined. The lysates were divided, and 30nM rapamycin was added to one of the aliquots. 25μl of each lysate was combined with 25μl of LgTrip 2098 (SEQ ID NO: 31), diluted to 200nM in TBS+0.01% BSA, and incubated for one hour. 50μl of NanoGlo® Buffer+50μM furimazine was added, and luminescence was read on GMM+.

Example 16

Random library preparation and screening

A random library of variant LgTrip polypeptides (using template LgTrip 2098) (SEQ ID NO: 31) was generated and screened for complementation with the β9/β10 dipeptide (SEQ ID NO: 35)(pep263).

Template DNA from LgTrip 2098 (SEQ ID NO: 32) was diluted to 10ug/ml in water. Diversify™ PCR Random Mutagenesis Kit (63070-ClonTech) was used to prepare a random library of mutants. Library amplification products were band isolated and purified using

WIZARD SV Gel and PCR Clean-Up System (A9281; Promega), cloned into pF4Ag, transformed into KRX competent cells (Promega), and plated onto LB agarose plates. Colonies were picked and place into wells of a 96-well plates with LB+ampicillin, and samples were grown overnight at 37°C with shaking. Overnight cultures were diluted 1:20 into induction media (LB+0.1% Rhamnose+0.15% glucose+100ug/ml ampicillin), and cultures were grown for 2-6 hours at 37°C. 10ul of cells were added to 250ul of PLB lysis buffer (0.3X PLB, 25mM HEPES pH 7.0, 0.001U/ml DNase). 50ul of cell lysate was combined with 50ul of assay buffer (NanoGlo® buffer+50uM Furimazine+0.2nM of pep263). Plates were incubated for 5 minutes after reagent addition and then samples were read on ClarioStar luminometer. Clones that had improved luminescence compared to the template clone were selected for additional screening.

Approximately 6,000 LgTrip 2098-based variant clones were further screened, and favorable mutation sites were evaluated with site saturation mutagenesis. Favorable mutations following saturation mutagenesis were combined to derive the LgTrip clone LgTrip 3092 (SEQ ID NO: 19). Screening was repeated using LgTrip 3092 (SEQ ID NO: 19) as a template, and the resulting clone was LgTrip 3546 (SEQ ID NO: 51).

Example 17

Purification of LgTrip clones

50ml cultures of LgTrip mutants were induced in LB+0.1% Rhamnose+0.15% Glucose+amp. Cultures were spun and re-suspended in 4.5ml of Hepes pH 7.5+0.001U/ml DNase. 500ul of FastBreak™ Cell Lysis Reagent (Promega; V8571) was added, and samples were incubated on a rotating mixer for 1 hour at 4°C. Samples were spun to clear lysate, and supernatant was transferred to a new tube. Using the HisLink™ Spin Protein Purification System, 500ul of HisLink™ Protein Purification Resin (Promega; V8821) was added to the samples, incubated for 2 hours at 4°C on a rotating mixer, washed with HisLink wash/binding buffer, and eluted with elution buffer. Slide-A-Lyzer dialysis columns were used to exchange buffer to TBS.

Example 18

Stability comparison

Experiments were conducted during development of embodiments herein to compare the stability of activity of LgBiT (SEQ ID NO: 11) and LgTrip 2098 (SEQ ID NO: 31) over time (Figure 14). Diluted purified LgTrip 2098 and LgBiT to 20nM in TBS+0.01% BSA or

in 0.3X PLB+25mM HEPES pH 7.5. Aliquoted 100µl of each sample into 200µl thin walled PCR tubes. Incubated samples at 37°C in thermal cycler, removed at various time-points, and placed on ice. Samples were equilibrated to room temperature and then diluted each sample to 0.2nM (5µl in 495µl) in PLB lysis buffer (0.3X PLB+25mM HEPES pH 7.5). 50µl of each diluted sample was combined with 50µl of 50µM Furimazine+6µM pep263 (SEQ ID NO: 35) in NanoGlo® buffer. Samples were incubated for 10 minutes and then read on GMM+. Calculated half-life using GraphPad Prism non-linear regression (One-Phase Decay plateau set to zero).

Example 19

Stability of LgTrip in TBS+minimal BSA carrier

Experiments were conducted during development of embodiments herein to determine the stability of the activity of NanoLuc® (SEQ ID NO: 3), LgBiT (SEQ ID NO: 11), and LgTrip 3546 (SEQ ID NO: 51) in TBS and a minimal BSA carrier over time (Figure 15). NanoLuc, LgBiT, and LgTrip 3546 were diluted to 20nM in TBS+0.01% BSA. 100µl of each were aliquoted into 200µl thin walled PCR tubes. Samples were incubated at 60°C in thermal cycler, removed at various time-points, and placed on ice. Samples were equilibrated to room temperature and diluted to 0.2nM (5µl in 495µl) in PLB lysis buffer (0.3X PLB+25mM HEPES pH 7.5). 50µl of each diluted sample was combined with 50µl of 50µM Furimazine+6µM pep263 (SEQ ID NO: 35) in NanoGlo® buffer. Samples were incubated for 10 minutes and then read on GMM+. Half-life was calculated using GraphPad Prism non-linear regression (One-Phase Decay plateau set to zero).

Example 20

Effect of salt on activity

Experiments were conducted during development of embodiments herein to determine the effect of salt concentration on the activity of NanoLuc® (SEQ ID NO: 3), LgBiT (SEQ ID NO: 11) LgTrip 2098 (SEQ ID NO: 31) and LgTrip 3546 (SEQ ID NO: 51) (Figure 16).

To test activity in the presence of salt, each enzyme was diluted to 1uM in TBS+0.01% BSA+0.01% Tergitol, and further diluted to 2nM in TBS+0.01% BSA+0.01% Tergitol . 4uM of pep 263 (SEQ ID NO: 35) was added to LgBiT, LgTrip 3546 (SEQ ID NO: 51), and LgTrip 2098 (SEQ ID NO: 31) and incubated for 30 minutes. A two-fold titration series was prepared for each, starting with 5M NaCl in 25mM Tris pH 7.5+0.01% Tergitol.

10uM furimazine was added to each sample of the titration series, and 5ul of each enzyme or enzyme+pep263 (SEQ ID NO: 35) were combined with 45ul of each substrate additive mixture. Plates were incubated for 3 minutes and then read on GMM+.

To test the effect of prolonged exposure to salt, each enzyme was diluted to 1uM, and a two-fold titration series was prepared starting with 5M NaCl in 25mM Tris pH 7.5+0.01% Tergitol. 2ul of each enzyme was added to 198ul of the NaCl titration (10nM final of each enzyme). The "no" additive control was TBS+0.01% BSA+0.01% Tergitol. Samples were incubated for 26 hours. After incubation, samples were diluted 1:10,000 into TBS+0.01% BSA+0.01% Tergitol (10ul in 990 two times). 4uM pep263 (SEQ ID NO: 35) was added to LgTrip 2098 (SEQ ID NO: 31), LgTrip 3546 (SEQ ID NO: 51), and LgBiT (SEQ ID NO: 11) in the second dilution. 50ul of each sample was combined with 50ul of NanoGlo® buffer+50uM Furimazine. Plates were incubated for 3 minutes and then read on GMM+.

Example 21

Effect of urea on activity

Experiments were conducted during development of embodiments herein to determine the effect of urea concentration on the activity of NanoLuc (SEQ ID NO: 3), LgBiT (SEQ ID NO: 11), LgTrip 2098 (SEQ ID NO: 31), and LgTrip 3546 (SEQ ID NO: 51) (Figure 17).

To test activity in the presence of urea, each enzyme was diluted to 1uM in TBS+0.01% BSA+0.01% Tergitol and further diluted to 2nM in TBS+0.01% BSA+0.01% Tergitol. 4uM of pep263 (SEQ ID NO: 35) was added to LgBiT (SEQ ID NO: 11), LgTrip 3546 (SEQ ID NO: 51), and LgTrip 2098 (SEQ ID NO: 31) and incubated for 30 minutes. A two-fold titration series was prepared for each, starting with 10M urea in 25mM Tris pH 7.5+0.01% Tergitol. 10uM furimazine was added to each sample of the titration series, and 5ul of each enzyme or enzyme+pep263 (SEQ ID NO: 35) were combined with 45ul of each substrate additive mixture. Plates were incubated for 3 minutes and then read on GMM+.

To test the effect of prolonged exposure to urea, each enzyme was diluted to 1uM, and a two-fold titration series was prepared starting with 10M urea in 25mM Tris pH 7.5+0.01% Tergitol. 2ul of each enzyme was added to 198ul of the urea titration (10nM final of each enzyme). The "no" additive control was TBS+0.01% BSA+0.01% Tergitol. Samples were incubated for 26 hours. After incubation, samples were diluted 1:10,000 into TBS+0.01% BSA+0.01% Tergitol (10ul in 990 ul two times). 4uM pep263 (SEQ ID NO: 35)

was added to LgTrip 2098 (SEQ ID NO: 31), LgTrip 3546 (SEQ ID NO: 51), and LgBiT (SEQ ID NO: 11) in the second dilution. 50ul of each sample was combined with 50ul of NanoGlo® buffer+50uM Furimazine. Plates were incubated for 3 minutes and then read on GMM+.

The results demonstrate that NanoLuc and NanoBiT are more susceptible to inactivation by urea compared to LgTrip 3546, while LgTrip 2098 is the least effected by urea. The exposure results demonstrate that LgTrip 3546, LgTrip 2098, and LgBiT regain activity upon prolonged treatment with urea indicating that activity of these polypeptide may be negatively affected by contaminating proteins, and that denaturation of these contaminants enhances activity.

Example 22

Effect of pH on activity

Experiments were conducted during development of embodiments herein to determine the effect of pH on the activity of NanoLuc® (SEQ ID NO: 3), LgBiT (SEQ ID NO: 11), LgTrip 2098 (WT) (SEQ ID NO: 31), and LgTrip 3546 (SEQ ID NO: 51) (Figure 18).

A universal buffer was prepared containing 25mM of each: NaCitrate, MES, PIPES, HEPES, TAPS, and Thiourea in 0.5% Tergitol. The buffer was divided into 8 aliquots of 20 ml, and NaOH or HCl was added to create a pH titration series.

To test effect of pH on activity, each enzyme was diluted to 1uM and then diluted to 0.4nM in 3ml of TBS+0.01% BSA+0.01% Tergitol. 4uM pep263 (SEQ ID NO: 35) was added to LgBiT, LgTrip 2098, and LgTrip 3546. Assay reagent for each pH buffer tested (Table 5) (20ul of furimazine in 980 of buffer). 10ul of each enzyme/peptide dilution was combined with 50ul of assay reagent. Plates were incubated for 3 minutes and read on GMM+.

Table 5. Buffers.

Component	MW(g/mol)	g
Na Citrate	294.1	1.47
MES	195.24	0.98
PIPES	302.37	1.51
Hepes	238.3	1.19
TAPS	243.3	1.22
Thiourea	76.12	0.53

To test the effect of prolonged exposure varying pH, each enzyme was diluted to 1uM in TBS+0.01% BSA+0.01% Tergitol, which was then diluted to 20 nM in each buffer. T=0 samples were mixed and then diluted 1:10 into 200mM HEPES pH 7.5+0.01% BSA+0.01% Tergitol and stored at 4ºC. T=8 samples were mixed and then diluted 1:10 into 200mM HEPES pH 7.5+0.01% BSA+0.01% Tergitol store at 4ºC. T=24 samples were mixed and then diluted 1:10 into 200mM HEPES pH 7.5+0.01% BSA+0.01% Tergitol store at 4ºC. To perform the assay, LgTrip and LgBiT were diluted 1:10 in TBS+0.01% BSA+0.01% Tergitol+4uM SmTrip10 pep286 (SEQ ID NO: 35), and NanoLuc was diluted into TBS+0.01% BSA+0.01% Tergitol. 40ul of each sample was diluted with 40ul of NanoGlo® buffer+40uM furimazine, incubated for 3 minutes, and then read on GMM+.

The results indicate that LgTrip is resistant to a wide pH range.

Example 23

Autoluminescence

Experiments were conducted during development of embodiments herein to compare the autoluminescence of LgBiT (SEQ ID NO: 11) and LgTrip 3546 (SEQ ID NO: 51). Each was diluted to 3uM in DPBS+0.01% BSA. Three-fold serial dilutions were prepared of each in DPBS+0.01% BSA (300 ul to 700ul) and placed in wells of a 96-well plate. The last row of the plate contained the furimazine controls (n=12). 50ul of each titration (or controls) were combined with 50ul of NanoGlo® buffer+50uM furimazine, incubated for 3 minutes, and then read on GMM+. LgTrip (i.e., LgTrip 3546) exhibited significantly reduced autoluminescence compared to LgBiT (Figure 19).

Example 24

Complementation of LgTrip with β9/β10 dipeptide

Experiments were conducted during development of embodiments herein to determine the capacity of a β9/β10 dipeptide (e.g., pep263) (SEQ ID NO: 35) to form a

bioluminescent complex with either LgTrip 3546 (SEQ ID NO: 51) or LgBiT (SEQ ID NO: 11). The luminescence of such a complex was compared with the luminescence of complexes of LgTrip 3546 with β9/β10 dipeptide and LgBiT with either SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) or the dipeptide, pep263 (SEQ ID NO: 35) (Figure 20). SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) and pep263 (SEQ ID NO: 35) were diluted to 2nM in H_2O. A 3X dilution series was prepared of each peptide in TBS+0.01% BSA starting at 300nM. Solutions of 200nM LgTrip 3546 and LgBiT were prepared in NanoGlo® buffer+50uM furimazine (NanoGlo® reagent). 50ul of each titration were combined with 50ul of each NanoGlo® reagent, and luminescence was read after a five minute incubation. Signal/background was calculated by dividing the amount of peptide dependent RLU by the background reading. Results demonstrate that the dipeptide has a K_d ~2-3X higher than HiBiT which accounts for lower RLU at low peptide concentration. The background of LgBiT decreases signal to background. RLU values for LgBiT/dipeptide and LgTrip/dipeptide were equal.

Example 25

Comparison of LgTrip 2098 & LgTrip 3546 complementation with SmTrip10 and SmTrip9

Experiments were conducted during development of embodiments herein to demonstrate complementation of LgTrip 2098 (SEQ ID NO: 31) & LgTrip 3546 (SEQ ID NO: 51), respectively, with SmTrip10 and SmTrip9 peptides, facilitated by the rapamycin-induced binding of SmTrip9 pep245-bound FKBP to SmTrip10 pep86-bound Frb (Figure 21). FKBP-SmTrip9 pep245, SmTrip10 pep86-Frb, and LgTrip 3546 or LgTrip 2098 were transiently transfected into HEK293 cells (20,000 cells per well/96-well plate). Samples were exposed to serial dilutions of rapamycin (to induce FKBP/Frb complex formation) and 10μM furimazine, and luminescence was measured. Results demonstrate that the affinity of SmTrip10 pep86 is ~10X lower for LgTrip 3546 compared to LgTrip 2098.

Example 26

Affinity of various SmTrip10 sequences

Experiments were conducted during development of embodiments herein forming luminescent complexes between various SmTrip10 pep286 (HiBiT; SEQ ID NO: 25) sequences and LgTrip 3546 (SEQ ID NO: 51)/SmTrip9 pep286 (SEQ ID NO: 37) (Figure 22). Enzymes were diluted to 200nM in TBS+0.01% BSA+0.01% Tergitol, serial dilutions

(100ul into 900ul) were prepared in TBS+0.01% BSA+0.01% Tergitol (2X to make 2nM), and 2nM sample was diluted 1:10 into TBS+0.01% BSA+0.01% Tergitol (500ul into 4.5ml). A 2X dilution series was prepared of each SmTrip10-like peptide in TBS+0.01% BSA+0.01% Tergitol+20uM of SmTrip9 pep286 (SEQ ID NO: 37) starting at 100uM. 50ul of diluted LgTrip 3546 (SEQ ID NO: 51) was combined with 50ul of the peptide titration and incubated for 10 minutes at room temperature. 100ul of TBS+0.01% BSA+0.01% Tergitol+20uM Furimazine was added to each sample, samples incubated for 10 minutes, and then read on GMM+.

Example 27

Inverse dipeptide

Experiments were conducted during development of embodiments herein to compare the capacity of dipeptides having opposite beta strand order (e.g., β9-β10 vs. β10-β9) to activate complement polypeptides (Figure 23). LgTrip 3546 (SEQ ID NO: 51) and LgTrip 2098 (SEQ ID NO: 31) were diluted to 200nM in TBS+0.01% BSA+0.01% Tergitol, and serial dilutions of each were prepared (100ul into 900ul) in TBS+0.01% BSA+0.01% Tergitol. The 2nM sample was diluted 1:10 into TBS+0.01% BSA+0.01% Tergitol (500ul into 4.5ml). 20uM stocks of each dipeptide (pep326 (SEQ ID NO: 179) and pep263 (SEQ ID NO: 35)) were prepared in TBS+0.01% BSA+0.01% Tergitol. 2X serial dilutions of each peptide were prepared TBS+0.01% BSA+0.01% Tergitol (250ul in 250ul). 50ul diluted LgTrip 2098 and LgTrip 3546 was combined with 50ul of the each peptide series and incubated at room temperature for 20 minutes. 100ul of LCS (Live cell substrate; Promega Catalog No. N205) in TBS (20uM) was added, and samples were incubated for 3 minutes and then read on GMM+.

Example 28

Comparison of dipeptides comprised of SmTrip9 (SEQ ID NO: 23) and either SmHiTrip (SEQ ID NO: 25) or SmBiT (SEQ ID NO: 13) for the SmTrip10 component

Experiments were conducted during development of embodiments herein to compare the capacity of dipeptides comprising the SmHiTrip (SEQ ID NO:25) or SmBiT (SEQ ID NO:13) sequence to activate complement polypeptides (Figure 24). LgBiT, LgTrip 2098, and LgTrip 2899 (SEQ ID NO: 364) were diluted to 200nM into TBS+0.01% BSA. Polypeptides were further diluted 1:100 into NanoGlo® buffer+50uM Furimazine (30ul in 3 ml). Pep263 (SEQ ID NO: 35) and pep274 (SEQ ID NO: 147) were diluted into TBS+0.01%

BSA to 5uM. 50ul of each LgBiT/LgTrip dilution were combined with 50ul of peptide dilution, incubated 5 minutes, and then read on GMM+.

Example 29

Additions/deletions of C-terminus of LGTrip 3546

Experiments were conducted during development of embodiments herein to determine the effect of C-terminal additions/deletions and/or corresponding additions/deletions to peptide tags on complementation and luminescence (Figure 25). Exemplary tested peptides and polypeptides are listed in Table 6.

Table 6. Peptide tags and polypeptide components comprising additions/deletions

SmTrip9 SmTrip10
LITPDGSMLFRVTINSVSGWRLFKKIS Note: SmTrip9 peptides contain additional SSWKR sequence at their N-termini.

LgTrip	C-term	SmTrip9	ID	SmTrip10	ID
ATG-3575 (aka LgTrip+GS)	--LITPDGS	MLFRVTINS	292	VSGWRLFKKIS	86
ATG-3572 (aka LgTrip+G)	--LITPDG	SMLFRVTINS	291	VSGWRLFKKIS	86
ATG-3573 (aka LgTrip-D)	--LITP	DGSMLFRVTINS	293	VSGWRLFKKIS	86
ATG-3574 (aka LgTrip-PD	--LIT	PDGSMLFRVTINS	294	VSGWRLFKKIS	86
ATG-3575 (aka LgTrip+GS)	--LITPDGS	MLFRVTINSV	297	SGWRLFKKIS	219
ATG-3572 (aka LgTrip+G)	--LITPDG	SMLFRVTINSV	296	SGWRLFKKIS	219
ATG-3573 (aka LgTrip-D)	--LITP	DGSMLFRVTINSV	298	SGWRLFKKIS	219
ATG-3574 (aka LgTrip-PD	--LIT	PDGSMLFRVTINSV	299	SGWRLFKKIS	219
ATG-3575 (aka LgTrip+GS)	--LITPDGS	MLFRVTINSVS	302	GWRLFKKIS	206
ATG-3572 (aka LgTrip+G)	--LITPDG	SMLFRVTINSVS	301	GWRLFKKIS	206
ATG-3573 (aka LgTrip-D)	--LITP	DGSMLFRVTINSVS	303	GWRLFKKIS	206
ATG-3574 (aka LgTrip-PD	--LIT	PDGSMLFRVTINSVS	304	GWRLFKKIS	206
ATG-3575 (aka LgTrip+GS)	--LITPDGS	MLFRVTIN	308	SVSGWRLFKKIS	157
ATG-3572 (aka LgTrip+G)	--LITPDG	SMLFRVTIN	307	SVSGWRLFKKIS	157
ATG-3573 (aka LgTrip-D)	--LITP	DGSMLFRVTIN	309	SVSGWRLFKKIS	157
ATG-3574 (aka LgTrip-PD)	--LIT	PDGSMLFRVTIN	310	SVSGWRLFKKIS	157
ATG-3575 (aka LgTrip+GS)	--LITPDGS	MLFRVTI	312	NSVSGWRLFKKIS	158
ATG-3572 (aka LgTrip+G)	--LITPDG	SMLFRVTI	311	NSVSGWRLFKKIS	158
ATG-3573 (aka LgTrip-D)	--LITP	DGSMLFRVTI	313	NSVSGWRLFKKIS	158
ATG-3574 (aka LgTrip-PD)	--LIT	PDGSMLFRVTI	314	NSVSGWRLFKKIS	158
ATG-3546 (aka LgTrip)	--LITPD	GSMLFRVTINSV	295	SGWRLFKKIS	219
ATG-3546 (aka LgTrip)	--LITPD	GSMLFRVTINSVS	300	GWRLFKKIS	206
ATG-3546 (aka LgTrip)	--LITPD	GSMLFRVTIN	305	SVSGWRLFKKIS	157
ATG-3546 (aka LgTrip)	--LITPD	GSMLFRVTI	306	NSVSGWRLFKKIS	158

Addition/deletion polypeptides were grown in 50ml cultures, pelleted, and resuspended in 10ml of 100mm HEPES pH 7.5+0.001U/ml DNase. 1ml of Fastbreak Cell Lysis Reagent (Promega Corporation) and 1ml of HisLink Resin (Promega Corporation) were added and incubated on a rotating shaker for 3 hours at 4°C. Resin was allowed to settle, and samples were washed 4X with 100mM HEPES pH 7.5+10mM Imidazole. Polypeptides were eluted twice into 500ul HisLink Elution buffer. Thermo dialysis tubes were used to equilibrate to 1X TBS.

Enzymes were diluted to 200nM in TBS+0.01% BSA+0.01% Tergitol, and serial dilutions were prepared (100ul into 900ul) in TBS+0.01% BSA+0.01% Tergitol. 2nM samples were diluted 1:10 into TBS+0.01% BSA+0.01% Tergitol (500ul into 4.5ml). SmTrip9- and SmTrip10-like peptides were combined with a polypeptide complement according to Table 7 and incubated for 10 minutes at room temperature. 100ul of TBS+0.01% BSA+0.01% Tergitol+20uM furimazine was added, incubated for10 minutes, and then read on GMM+.

Table 7. Polypeptide/peptide combinations tested.

Group 1	SmTrip 9	SmTrip10
3546	286	86
3575	292	86
3572	291	86
3573	293	86
3574	294	86

Group 2	SmTrip 9	SmTrip10
3546	286	86
3575	297	219
3572	296	219
3573	298	219
3574	299	219

Group 3	SmTrip 9	SmTrip10
3546	286	86
3575	302	206
3572	302	206
3573	303	206
3574	304	206

Group 4	SmTrip 9	SmTrip10
3546	286	86
3575	312	158
3572	312	158
3573	312	158
3574	312	158

Group 5	SmTrip 9	SmTrip10
3546	286	86
3546	295	219
3546	300	206
3546	305	157
3546	306	158

Example 30

Polypeptide/peptide and/or peptide/peptide overlap

Experiments were conducted during development of embodiments herein to determine the sequence overlap between the polypeptide component and a peptide corresponding to the β-strand or between the two peptides (Figure 26). In such experiments, a polypeptide component and peptide, or the two peptides, each comprise amino acids corresponding to the same amino acids in a base luciferase.

Polypeptides were diluted to 200nM in TBS+0.01% BSA+0.01% Tergitol. Serial dilutions (100ul into 900ul) were prepared in TBS+0.01% BSA+0.01% Tergitol. 2nM of each polypeptide sample were diluted 1:10 into TBS+0.01% BSA+0.01% Tergitol (500 into 4.5ml). Polypeptides and peptides were combined, i.e., 50ul of each LgTrip mutant with 50ul of the peptide, according to Table 8. Reactions were incubated for 10 minutes at room temperature. Next, 100ul of TBS+0.01% BSA+0.01% Tergitol+20uM furimazine was added, and reactions were incubated for 10 more minutes prior to reading on a GMM+ luminometer.

Table 8. Polypeptide/peptide combinations tested.

Group1	SmTrip9	SmTrip10
3546	286	86
3546	286	219
3546	286	206
3546	286	157
3546	286	158

Group2	SmTrip9	SmTrip10
3575	286	86
3575	286	219
3575	286	206
3575	286	157
3575	286	158

Group3	SmTrip9	SmTrip10
3546	293	86
3546	293	219
3546	293	206
3546	293	157
3546	293	158

Group4	SmTrip9	SmTrip10
3575	293	86
3575	293	86
3575	293	86
3575	293	86
3575	293	86

Group5	SmTrip9	SmTrip10
3546	294	86
3546	294	219
3546	294	206
3546	294	157
3546	294	158

Group6	SmTrip9	SmTrip10
3575	294	86
3575	294	219
3575	294	206
3575	294	157
3575	294	158

Example 31

The identity of the β9-peptide alters the affinity of the β10-peptide

Experiments were conducted during development of embodiments herein to determine the sequence overlap between the polypeptide component and a peptide corresponding to the β-strand or between the two peptides (Figure 27). These results show that the sequence of the β9 strand peptide (SmTrip9) can impact the affinity of the β10 strand peptide (SmTrip10).

SmTrip9 titration

LgTrip 3546 (SEQ ID NO: 51) was diluted to 200nM in TBS+0.01% BSA+0.01% Tergitol, and serial dilutions were prepared (100ul into 900ul) in TBS+0.01% BSA+0.01% Tergitol. 2nM polypeptide samples were diluted 1:10 into TBS+0.01% BSA+0.01% Tergitol (500into 4.5ml).

A 2X dilution series was prepared of each SmTrip9 peptide in TBS+0.01% BSA+0.01% Tergitol+100uM of SmTrip10 pep86 (SEQ ID NO: 25) starting at 20uM. 50ul of diluted LgTrip 3546 (SEQ ID NO: 51) was combined with 50ul of each peptide titration. Reactions were incubated for 10 minutes at room temperature. 100ul of TBS+0.01% BSA+0.01% Tergitol+20uM furimazine was added, the reaction was incubated for 10 more minutes, and read on a GMM+.

SmTrip10 titration

LgTrip 3546 (SEQ ID NO: 51) was diluted to 200nM in TBS+0.01% BSA+0.01% Tergitol, and serial dilutions were prepared (100ul into 900ul) in TBS+0.01% BSA+0.01% Tergitol. 2nM polypeptide samples were diluted 1:10 into TBS+0.01% BSA+0.01% Tergitol (500into 4.5ml).

A 2X dilution series was prepared of SmTrip10 pep86 (SEQ ID NO: 25) in TBS+0.01% BSA+0.01% Tergitol+20uM of SmTrip9-like peptides starting at 100uM. 50ul of diluted LgTrip 3546 was combined with 50ul of each peptide titration. Reactions were incubated for 10 minutes at room temperature. 100ul of TBS+0.01% BSA+0.01% Tergitol+20uM furimazine was added, reaction was incubated for 10 more minutes, and then read on GMM+.

Results (Figure 27)

SmTrip 9 peptide variants were titrated in the presence of constant SmTrip10 ep86 (SEQ ID 15) (Figure 27a), and then SmTrip10 pep86 was titrated in the presence of saturating amounts of each SmTrip9 variant peptide. (Figure 27b) This shows that the affinity of the SmTrip10 sequence can be altered depending on the SmTrip9 sequence. The experiments demonstrate that identity of the β9-like peptide (e.g., SmTrip9) can influence the affinity of the β10-like peptide (e.g., SmTrip10) for the polypeptide component. SmTrip9 pep293 (SEQ ID NO: 154) and SmTrip9 pep294 (SEQ ID NO: 155) sequences have overlap with the C-terminus of LgTrip 3546 (SEQ ID NO: 51) and show a decrease in affinity compared to SmTrip9 pep286 (SEQ ID NO: 37) (no overlap), but also decrease affinity of SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) for LgTrip 3546. SmTrip9 pep298 (SEQ ID

NO: 158) and SmTrip9 pep299 (SEQ ID NO: 159) sequences overlap with the C-terminus of LgTrip 3546 and the N-terminus of SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) and decrease the affinity of SmTrip10 pep86 (HiBiT) for or LgTrip 3546.

Example 32

Effect of β10-peptide identity on the affinity of the β10 peptide component to the polypeptide and β9-peptide.

Experiments were conducted during development of embodiments herein to determine the how sequence overlaps or sequence gaps between the polypeptide component and a peptide corresponding to the β-strands or between the two peptides influence the affinity of the β10-like (e.g., SmTrip10) peptides (Figure 28). LgTrip 3546 (SEQ ID NO: 51) was diluted to 200nM in TBS+0.01% BSA+0.01% Tergitol, and serial dilutions were prepared (100ul into 900ul) in TBS+0.01% BSA+0.01% Tergitol. 2nM polypeptide samples were diluted 1:10 into TBS+0.01% BSA+0.01% Tergitol (500 into 4.5ml). A 2X dilution series was prepared of each SmTrip10-like peptide in TBS+0.01% BSA+0.01% Tergitol+20uM of SmTrip9 pep286 (SEQ ID NO: 37) starting at 100uM. 50ul of diluted LgTrip 3546 was combined with 50ul of each peptide titration. Reactions were incubated for 10 minutes at room temperature. 100ul of TBS+0.01% BSA+0.01% Tergitol+20uM furimazine was added, reaction was incubated for 10 more minutes, and then read on GMM+.

Example 33

Measure of affinity of β9-like peptides in the presence of various saturating β10-like peptides

Experiments were conducted during development of embodiments herein to determine the how the affinity of β9-like (e.g., SmTrip9) peptides are impacted in the presence of constant concentrations of various β10-like (e.g., SmTrip10) peptides with LgTrip 3546 (SEQ ID NO: 51) (Figure 29). Polypeptide component LgTrip 3546 was diluted to 200nM in TBS+0.01% BSA+0.01% Tergitol, serial dilutions were prepared (100ul into 900ul) in TBS+0.01% BSA+0.01% Tergitol (2X to make 2nM), and 2nM sample was diluted 1:10 into TBS+0.01% BSA+0.01% Tergitol (500µl into 4.5ml). A 2X dilution series was prepared of each β9-like peptide (SmTrip9 pep286 (SEQ ID NO: 37) and SmTrip9 pep287 (SEQ ID NO: 148)) in TBS+0.01% BSA+0.01% Tergitol+100uM of each SmTrip10-like peptide starting at 20uM. 50ul of diluted LgTrip 3546 (SEQ ID NO: 51) was combined with 50ul of each peptide titration. The reactions were incubated for 10 minutes at room

temperature, 100ul of TBS+0.01% BSA+0.01% Tergitol+20uM furimazine was added, and the reactions were incubated for another 10 minutes and then read on GMM+.

Example 34

Effect of construct orientation on facilitated complementation in HEK293 cells

Experiments were conducted during development of embodiments herein to determine the effect the orientation of interaction elements (FRB and FKBP) relative to the peptide tags on complementation with LgTrip 3546 (SEQ ID NO: 51) (Figure 30). Un-induced signal greater than 100,000 RLU is indicative of background contamination, which decreases the apparent fold-response.

HEK293 cells were grown overnight at 37°C with 5% CO_2. Cells were transfected with 3 ug DNA (SmTrip9 pep245-FKBP, FKBP-SmTrip9 pep245, SmTrip10 pep86-FKBP, FKBP-SmTrip10 pep86, SmTrip9 pep245-FRB, FRB-SmTrip9 pep245, SmTrip10 pep86-FRB, or FRB-SmTrip10 pep86 construct) per well using FuGENE protocol. Cells were washed in DPBS. 1 ml DPBS was added, cells were frozen at -80C for ~10 min, and thawed at room temperature. Lysates were cleared by centrifugation for 10 minutes, diluted 1:10 in TBS+200nM LgTrip (+/- 30nM RAP), and incubated for 30 min at room temperature. 50ul of each sample was combined with 50ul of TBS + 20uM Furimazine, and luminescence was read at 5 minutes.

Example 35

Effect of construct orientation on facilitated complementation in E. coli

Experiments were conducted during development of embodiments herein to determine the effect the orientation of interaction elements (FRB and FKBP) relative to the peptide tags on complementation with LgTrip 3546 (SEQ ID NO: 51) (Figure 31).

Overnight cultures of each construct were prepared in LB+100ug/ml ampicillin. Cultures were diluted 1:100 into induction media (LB+amp+0.1% rhamnose+0.15% glucose, cells were grown for 20 hours at 25°C, and lysed with PLB lysis buffer (0.3X PLB, 25mM Hepes pH 7.5, 0.001U/ul Rq1 DNase; 250ul of cells, 750ul PLB) for 15 minutes. Cells were diluted 5X into CO_2 independent media+10% FBS that contains 200nM LgTrip 3546 and +/- 30nM RAP. Reactions were incubated for 30 minutes at room temperature, combined with equal volumes of NanoGlo+50uM furimazine (50ul to 50ul), incubated for 5 minutes, and then read on GMM+

Example 36

Kd measurement for various β10-like peptides

Experiments were conducted during development of embodiments herein to measure Kd values for various β10-like peptides with LgTrip 3546 (SEQ ID NO: 51) and SmTrip9 pep286 (SEQ ID NO: 37) (Figure 32). A solution was prepared of 20uM SmTrip9 pep286 (SEQ ID NO: 37) in TBS+0.005%+0.01% BSA. 3X serial dilutions were prepared of SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25), SmTrip10 pep288 (SEQ ID NO: 149), SmTrip10 pep289 (SEQ ID NO: 150), and SmTrip10 pep290 (SEQ ID NO: 151) (150ul in 350ul TBS+0.01% BSA+286 solution starting at 100uM). 20nM LgTrip 3546 solutions were prepared in TBS+0.01% BSA, and then diluted 1:10 in TBS+0.01% BSA. 25ul of each peptide solution was combined with 2.5ul of the LgTrip 3546 solutions. Reactions were incubated for 10 minutes, 28ul of TBS+20uM LCS (Promega Catalog No. N205) was added, incubated for 10 minutes, and then read on GMM+. This experiment shows that the addition of either "V" or "VS" to the N-terminus of SEQ ID NO: 25 increases the affinity of the SmTrip10-like peptide compared to SmTrip10 pep86 (HiBiT).

Example 37

Effect of Polypeptide/β9 split site on luciferase light output

Experiments were conducted during development of embodiments herein to analyze the effect of moving the split site between the polypeptide component and the SmTrip9-like peptide (Figure 33). Polypeptide components with varied C-terminal extensions or deletions were diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol, and 50 uM SmTrip10 pep86 (SEQ ID NO: 25) was added to each. SmTrip9 pep286 (SEQ ID NO: 37) was added to 10 uM in the SmTrip10 pep86+LgTrip solutions, and samples were incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM furimazine detection reagent was added at 1:1, and luminescence was read. All synthetic SmTrip9 peptides contained the N-terminal solubility tag SSWKR.

Example 38

Effect of sequence gaps and overlaps between LgTrip C-terminus and SmTrip9 pep286 on luciferase light output

Experiments were conducted during development of embodiments herein to analyze the effect of gaps and/or overlaps between the polypeptide component and the SmTrip9-like peptide (Figure 34). Polypeptide components with varied C-terminal extensions or deletions

were diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol, and 50 uM SmTrip10 pep86 (SEQ ID NO: 25) was added to each. 10 uM of a SmTrip9 pep286 was added to SmTrip10 pep86+LgTrip solutions, and the reactions were incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM furimazine detection reagent was added at 1:1, and luminescence was read.

Example 39

Effect of SmTrip9 sequence gaps and overlaps with LgTrip 3546 and SmTrip10 pep 86 (HiBiT) on luciferase light output

Experiments were conducted during development of embodiments herein to analyze the effect of gaps and/or overlaps between the SmTrip9-like peptide and the polypeptide component (e.g., LgTrip) and/or SmTrip10-like peptide (Figure 35). LgTrip 3546 (SEQ ID NO: 51) was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol, and 50 uM SmTrip10 pep86 (SEQ ID NO: 25) was added to each. SmTrip9 pep286 (SEQ ID NO: 37) was added to 10 uM in SmTrip10 pep86+LgTrip solutions, and samples were incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM furimazine detection reagent was added at 1:1, and luminescence was read. All synthetic SmTrip9 peptides contained the N-terminal solubility tag, SSWKR.

Example 40

Biochemical analysis (Kd and Bmax) of SmTrip9 peptide length variants

Experiments were conducted during development of embodiments herein to analyze complementation of SmTrip9 peptides of different lengths with LgTrip 3546 (SEQ ID NO: 51) and SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 15) (Figures 36-37).

SmTrip9 titration (Figure 36)

LgTrip 3546 (SEQ ID NO: 51) polypeptide was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol. 100 uM SmTrip10 pep86 was prepared in TBS+0.01% BSA+0.01% Tergitol. 20 uM solutions of each SmTrip9-like peptide were added to the SmTrip10 pep86 solution. 2X serial dilutions were prepared of each SmTrip9 peptide solution using the SmTrip10 pep86 solution as a diluent. Peptide dilutions and LgTrip 3546 solution were combined 1:1 and incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM furimazine detection reagent was added (1:1), and luminescence was detected.

HiBiT (SmTrip10) titration (Figure 37)

LgTrip 3546 (SEQ ID NO: 51) polypeptide was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol. 20 uM SmTrip9-like peptide solutions were prepared in TBS+0.01% BSA+0.01% Tergitol for each SmTrip9-like peptide to be tested. 100 uM solutions of SmTrip10 pep86 was added to each SmTrip9-like peptide solution. 2X serial dilutions were prepared of SmTrip10 pep86 using each SmTrip9-like peptide solution as a diluent. Peptide dilutions and LgTrip 3546 solution were combined, 1:1, and incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM furimazine detection reagent was added (1:1), and luminescence was detected.

Example 41

Biochemical affinity and Bmax of SmTrip9 pep286 point mutants

Experiments were conducted during development of embodiments herein to analyze the affinity of SmTrip9 pep286 (SEQ ID NO: 37) point mutants for LgTrip 3546 (SEQ ID NO: 51) and SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25). LgTrip 3546 polypeptide was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol. 100 uM SmTrip10 pep86 was prepared in TBS+0.01% BSA+0.01% Tergitol. 20 uM solutions of each SmTrip9-like peptide were added to the SmTrip10 pep86 solution. 2X serial dilutions were prepared of each SmTrip9-like peptide solution using the SmTrip10 pep86 solution as a diluent. Equal volumes of peptide dilutions and LgTrip solution were combined and incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM furimazine detection reagent was added (1:1, vol:vol), and luminescence was detected (Figure 38) to determine Kd and Bmax of each SmTrip9-like peptide.

Example 42

Effect of SmTrip9 solubility tags on biochemical affinity and Bmax

Experiments were conducted during development of embodiments herein to analyze the affinity of SmTrip9-like peptides with alternative solubility tags (Figure 39). LgTrip 3546 (SEQ ID NO: 51) was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol. 100 uM SmTrip10/pep86 solution was prepared in TBS+0.01% BSA+0.01% Tergitol. 20 uM solutions were prepared of each SmTrip9-like peptide in the SmTrip10 pep86 solution. 2X serial dilutions were prepared of each SmTrip9-like peptide using the SmTrip10 pep86 solution as a diluent. Equal volumes of peptide dilutions were combined with the LgTrip 3546 solution, and reactions were incubated for 10 minutes. TBS+0.01% BSA+0.01%

Tergitol+20uM furimazine detection reagent was added 1:1 vol:vol to the reactions, and luminescence was read after 10 minutes of incubation.

Example 43

C-terminal extension sequences

Experiments were conducted during development of embodiments herein to analyze the affinity of SmTrip9-like peptides with C-terminal sequence extensions (Figure 40).

SmTrip9 peptide titration

LgTrip 3546 (SEQ ID NO: 51) polypeptide was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol. 100 uM SmTrip10 pep86 was prepared in TBS+0.01% BSA+0.01% Tergitol. 20 uM solutions of each SmTrip9-like peptide were added to the SmTrip10 pep86 solution. 2X serial dilutions were prepared of each SmTrip9-like peptide solution using the SmTrip10 pep86 solution as a diluent. Peptide dilutions and LgTrip 3546 solution were combined 1:1 and incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM furimazine detection reagent was added (1:1), and luminescence was detected.

SmTrip10 pep 86 (HiBiT) titration

LgTrip 3546 (SEQ ID NO: 51) polypeptide was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol. 20 uM SmTrip9-like peptide solutions were prepared in TBS+0.01% BSA+0.01% Tergitol for SmTrip9-like peptide to be tested. 100 uM solutions of SmTrip10 pep86 (SEQ ID NO: 25) was added each SmTrip9-like solution. 2X serial dilutions were prepared of SmTrip10 pep86 using each SmTrip9-like peptide solution as a diluent. Peptide dilutions and LgTrip 3546 solution were combined 1:1 and incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM furimazine detection reagent was added (1:1), and luminescence was detected.

Example 44

Measurement of FRB-FKBP facilitated complementation using FRB-SmTrip10 variants and FKBP fused SmTrip9 pep245 in KRX E. coli lysates

Overnight cultures of FRB-SmTrip10 variants, FKBP-SmTrip9 pep245, and SmTrip9 pep245-FKBP were grown in LB+100ug/ml ampicillin from glycerol stocks. Cells were diluted 1:100 in LB + 0.15% glucose + 0.1% rhamnose + Amp, and shook for 20 hours at 25°C. Cultures were diluted 1:4 in PLB and incubated 15 min at room temperature to lyse cells. SmTrip9/SmTrip10 peptide combinations were mixed 1:1 (vol:vol). Mixtures were diluted 1:5 into PLB + 200 nM LgTrip 3546 (SEQ ID NO: 51) with or without 30 nM

rapamycin, and reactions were incubated for 30 minutes at room temperature. Each reaction was combined with 50ul of NanoGlo® buffer+50uM Furimazine, and luminescence was measured at 5 minutes. Results for fold induction (+rap signal/-rap signal) are depicted in Figure 41. FRB-SmTrip10 variant peptide constructs possessed varied linker lengths, linker content (with or without alanine-isoleucine), and either contained or lacked a hexahistidine tag.

Example 45

Measurement of FRB-FKBP facilitated complementation using FRB-SmTrip10 variants and FKBP fused SmTrip9 pep245 in HEK lysates

Overnight cultures of FRB-SmTrip10 variants, FKBP-SmTrip9 pep245, and SmTrip9 pep245-FKBP were grown at 37°C with 5% CO_2. Cells were transfected with 1 ug DNA (FKBP or FRB construct) per well using FuGENE protocol. Cells were washed in 1 ml DPBS. 1 ml DPBS was added and cultures were frozen at -80°C for 10 min. Cultures were thawed at room temperature to lyse cells. Lysates were cleared by centrifugation for 10 min and diluted 2-fold in PLB. SmTrip9/SmTrip10 peptide combinations were mixed 1:1 (vol:vol). Mixtures were diluted 1:5 into PLB + 200 nM LgTrip 3546 (SEQ ID NO: 51) with or without 30 nM rapamycin, and reactions were incubated for 30 minutes at room temperature. Each reaction was combined with 50ul of NanoGlo®buffer+50uM Furimazine, and luminescence was measured at 5 minutes. Results for fold induction (+rap signal/-rap signal) are depicted in Figure 42. FRB-SmTrip10 variant peptide constructs possessed varied linker lengths, linker content (with or without alanine-isoleucine), and either contained or lacked a hexahistidine tag.

Example 46

Measurement of FRB-FKBP facilitated complementation using FRB-SmTrip10 pep86 (HiBiT)/SmTrip10 pep289 (VS-HiBiT) and SmTrip9 sequences fused to FKBP in both orientations in KRX E. coli lysates

Overnight cultures for FRB-SmTrip10 pep86 (HiBiT; SEQ ID NO: 25) or FRB-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO: 150) and SmTrip9-like peptide sequences fused to FKBP were grown in LB+100ug/ml ampicillin from glycerol stocks. Cells were diluted 1:100 in LB+0.15% glucose+0.1% rhamnose+Amp and shaken for 20 hours at 25°C. Cultures were diluted 1:4 in PLB and incubated 15 min at room temperature to lyse. SmTrip9/SmTrip10 peptide combinations were mixed 1:1 (vol:vol). Mixtures were diluted

1:5 into PLB + 200 nM LgTrip 3546 (SEQ ID NO: 51) with or without 30 nM rapamycin, and reactions were incubated for 30 minutes at room temperature. Each reaction was combined with 50ul of NanoGlo® buffer+50uM Furimazine, and luminescence was measured at 5 minutes. Results are depicted in Figures 43-47.

Example 47

Measurement of FRB-FKBP facilitated complementation using FRB-SmTrip10 pep86 (HiBiT)/SmTrip10 pep289 (VS-HiBiT) and SmTrip9 sequences fused to FKBP in both orientations in HEK lysates

Overnight cultures for FRB-SmTrip10 pep86 (HiBiT; SEQ ID NO: 25) or FRB-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO: 150) and SmTrip9-like peptide sequences fused to FKBP were grown at 37°C with 5% CO_2. Cells were transfected with 1 ug DNA (FKBP or FRB construct) per well using FuGENE protocol. Cells were washed in 1 ml DPBS. 1 ml DPBS was added and cultures were frozen at -80°C for 10 min. Cultures were thawed at room temperature to lyse cells. Lysates were cleared by centrifugation for 10 min, and diluted 2-fold in PLB. SmTrip9/SmTrip10 peptide combinations were mixed 1:1 (vol:vol). Mixtures were diluted 1:5 into PLB + 200 nM LgTrip 3546 (SEQ ID NO: 51) with or without 30 nM rapamycin, and reactions were incubated for 30 minutes at room temperature. Each reaction was combined with 50ul of NanoGlo® buffer+50uM Furimazine, and luminescence was measured at 5 minutes. Results are depicted in Figures 48-50.

Example 48

Measurement of FRB-FKBP facilitated complementation using FRB-SmTrip10 pep86 (HiBiT)/SmTrip10 pep289 (VS-HiBiT) and SmTrip9 sequences fused to FKBP in E. coli lysates

Overnight cultures for FRB-SmTrip10 pep86 (HiBiT; SEQ ID NO: 25) or FRB-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO: 150) and SmTrip9-like peptide sequences fused to FKBP were grown in LB+100ug/ml ampicillin from glycerol stocks. Cells were diluted 1:100 in LB + 0.15% glucose + 0.1% rhamnose + Amp and shook for 20 hours at 25°C. Cultures were diluted 1:4 in PLB and incubated 15 min at room temperature to lyse. SmTrip9/SmTrip10 peptide combinations were mixed 1:1 (vol:vol). Mixtures were diluted 1:5 into PLB + 200 nM LgTrip 3546 (SEQ ID NO: 51) with or without 30 nM rapamycin, and reactions were incubated for 30 minutes at room temperature. Each reaction was combined with 50ul of NanoGlo® buffer+50uM Furimazine, and luminescence was

measured at 5 minutes. Results are depicted in Figures 57, 59, 60, 62-63, 66-67, and 70-71. In Figure 57, ** indicates that alanine-isoluecine (AI) in the linker directly upstream of SmTrip9 peptides or SmTrip10 peptides has been removed. Alanine-isoluecine is absent from C-terminal FKBP or FRB fusions with SmTrip9 peptide or SmTrip10 peptides, respectively, in all subsequent figures.

Example 49

Measurement of FRB-FKBP facilitated complementation using FRB-SmTrip10 pep86 (HiBiT)/SmTrip10 pep289 (VS-HiBiT) and SmTrip9 sequences fused to FKBP in HEK293 lysates

Overnight cultures for FRB-SmTrip10 pep86 (HiBiT; SEQ ID NO: 25) or FRB SmTrip10 pep289(VS-HiBiT; SEQ ID NO: 150) and SmTrip9-like peptide sequences fused to FKBP were grown at 37°C with 5% CO_2. Cells were transfected with 3 ug DNA (FKBP or FRB construct) per well using FuGENE protocol. Cells were washed in 1 ml DPBS. 1 ml DPBS was added and cultures were frozen at -80°C for 10 min. Cultures were thawed at room temperature to lyse cells. Lysates were cleared by centrifugation for 10 min, and diluted 2-fold in PLB. SmTrip9/SmTrip10 peptide combinations were mixed 1:1 (vol:vol). Mixtures were diluted 1:5 into PLB + 200 nM LgTrip 3546 (SEQ ID NO: 51) with or without 30 nM rapamycin, and reactions were incubated for 30 minutes at room temperature. Each reaction was combined with 50ul of NanoGlo® buffer+50uM Furimazine, and luminescence was measured at 5 minutes. Results are depicted in Figures 58, 61, 64-65, 68-69, and 72-73. In Figure 58, ** indicates that alanine-isoluecine (AI) in the linker directly upstream of SmTrip9 peptides or SmTrip10 peptides has been removed. Alanine-isoluecine is absent from C-terminal FKBP or FRB fusions with SmTrip9 peptides or SmTrip10 peptides, respectively, in all subsequent figures.

Example 50

Biochemical analysis (Kd and Bmax) of varied SmTrip9 sequences

Results are depicted in Figures 74-76.

SmTrip9-like Peptide Titrations

LgTrip 3546 (SEQ ID NO: 51) was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol. 100 uM SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) was prepared in TBS+0.01% BSA+0.01% Tergitol. 20 uM solutions were prepared of each SmTrip9-like peptide in the SmTrip10 pep86 solution. 2X serial dilutions of each SmTrip9-like peptide

were prepared using the SmTrip10 pep86 solution as a diluent. Peptide dilutions were combined with LgTrip 3546 (SEQ ID NO: 51) solution, 1:1, and incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM Furimazine (Fz) detection reagent was added, 1:1. Luminescence was read at 10 min.

SmTrip10 pep86 (HiBiT) Titrations

LgTrip 3546 (SEQ ID NO: 51) was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol. 20 uM SmTrip9-like peptide solutions were prepared in TBS+0.01% BSA+0.01% Tergitol. SmTrip10 pep86 was added to 100 uM in each SmTrip9-like peptide solution. 2X serial dilutions of each SmTrip10 pep86 (SEQ ID NO: 25) were prepared using the SmTrip9-like peptide solutions as a diluent. Peptide dilutions were combined with LgTrip 3546 (SEQ ID NO: 51) solution 1:1 and incubated for 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM Furimazine (Fz) detection reagent was added, 1:1. Luminescence was read at 10 min.

Example 51

Solubility of Synthetic SmTrip9 peptides

Synthetic peptides were synthesized by Peptide2.0 with termini blocked (N-terminal acetylation and C-terminal amidation) unless otherwise noted. Peptides were dissolved in nuclease-free water to ~1 mM and mixed on rotater at 4°C for 30 min. Following centrifugation for 10 min at top speed, peptides were diluted 1:50 in water and quantified on NanoDrop. Peptides were stored at -20°C until use. Peptides were deemed soluble if they remained in solution after 3 freeze/thaw cycles in which peptides were thawed in a 22°C water bath, kept at 4°C, and frozen at -20°C. Solubility of synthetic peptides is depicted in Figure 77.

Example 52

Circulalry permuted LgBiT

SmTrip9/pep286 affinity and Bmax for SmTrip10 pep 86 (HiBiT)-LgTrip 3546 fusions

A fusion polypeptide comprising a SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) sequence fused to the front of LgTrip 3546 (SEQ ID NO: 51) was generated and experiments were conducted to monitor complex formation and luminescence of the SmTrip10 pep86 HiBiT-LgTrip fusions with SmTrip9 pep286 (SEQ ID NO: 37) (Figure 78).

Overnight cultures were grown in LB+100ug/ml ampicillin from glycerol stocks. Cells were diluted 1:100 in LB+0.15% glucose+0.1% rhamnose+Amp and shaken for 20 hr at

25°C. 800 µl culture was lysed in FastBreak and each SmTrip10 pep86-LgTrip fusion was purified using the HisLink protocol. 2-fold SmTrip9 pep286 (SEQ ID NO: 37) serial dilutions starting at 10 uM were made in TBS + 0.01% Tergitol + 0.01% BSA containing 0.2 nM SmTrip10 pep86-LgTrip fusion (ATG 3745 (SEQ ID NO: 211) or ATG 3746 (SEQ ID NO: 213)). Reactions were pre-incubated for 10 minutes at room temperature. TBS+0.01% Tergitol+0.01% BSA with 20 uM Furimazine (Fz) was added to samples 1:1 (vol:vol). Luminescence was read on GloMax® luminometer at 10 min.

SmTrip9/pep759 affinity for vaious SmTrip10 pep 86 (HiBiT)-LgTrip 3546 fusions

From pellets of induced cell culture, pellets were resuspended in 1/10 of the original culture volume (e.g. a 50 mL culture would be resuspended in 5 mL) using 1X TBS + 0.01% BSA. A lysis buffer was prepared using 100 parts Fast Break® Buffer, 10 Parts RQ1 RNAse free DNAse, and 1 part 1 M DTT (e.g. 650 µL Fast Break® Buffer + 65 µL RQ1 RNAse free DNAse, and 6.5 µL 1 M DTT or equivalent scaling). 1 part Lysis buffer was added to 9 parts cell suspension (e.g. 33.3 µL Lysis buffer + 300 uL suspension) in a 15 mL tube. Incubated at 4°C for 30 minutes while mixing (using a rotary shaker). A 4 µM solution of pep 759 was prepared in 1X TBS + 0.01 % BSA. 50 uL of 4 µM pep759 was added to 50 µL of each lysate in a 96 well plate in triplicate. 50 µL of each lysate was separately mixed with 50 µL of 1X TBS + 0.01 % BSA buffer in triplicate. NanoGlo® Reagent aws prepared by mixing 100 parts NanoGlo® Buffer with 1 part Furimazine (e.g. 10 mL buffer + 100 uL furimazine). 100 uL of NanoGlo® reagent was added to each well. Luminescence was measured using Glomax® Multi instrument kinetic cycles. Luminescence measurements were compared after about 29 minutes. Luminescence readings for samples with pep759 were divided by the corresponding measurement of the same lysate without pep759. Results are depicted in figure 78B. Two batches of cultures were used to generate data: one was from inductions of 50 mL cultures (the right side, ATG-4808 through and including ATG-4632) and the other was from inductions of 3 mL cultures (left side, starting with ATG-4815 through and including ATG-3746). Some constructs were present in both tests (ATG-2623, ATG-3745, ATG-3746, ATG-4632).

Example 53

Detergent titration

Experiments were conducted during development of embodiments herein to determine the impact of various detergents on NanoLuc (SEQ ID NO: 3), LgBiT (SEQ ID NO: 11), and LgTrip 3546 (SEQ ID NO: 51) complexes with the dipeptide, pep263 (SEQ ID NO: 35).

Exposure experiments

500ul of 20mM SDS or 2mM CDTA or 5% Tergitol was added to a deep well plate. 3X serial dilutions were prepared of each detergent in TBS+0.01% BSA (150ul in 350ul). 100ul of each dilution was combined with 100ul of either 2nM NanoLuc, LgBiT, or LgTrip, and samples were incubated for 18 hours. Samples were diluted 1:100 in TBS+0.01% BSA (5ul in 495 ul). 50ul of each sample was combined in triplicate with 50ul of TBS+0.01% BSA+20uM Furimazine (Fz) for NanoLuc or TBS+0.01% BSA+20uM Furimazine (Fz)+2uM pep263 for LgBiT and LgTrip. Luminescence of samples was read on GMM+ 3 minutes after reagent addition. Results of prolonged exposure to detergent on LgBiT, LgTrip 3546, and NanoLuc are depicted in Figure 79.

Activity experiments

20ml of 20uM Fz was prepared in TBS+0.01% BSA. 2ml of 20mM SDS and 2mM of CDTA and 5% Tergitol were added to a deep well plate. 20uM Fz was added to each sample (8ul). 2X serial dilutions were prepared of each detergent in 20uM Fz solution (1ml to 1ml). A solution of 400pM NanoLuc in TBS+0.01% BSA was prepared. A solution of 400pM LgBiT+1uM pep263 (SEQ ID NO: 35) in TBS+0.01% BSA was prepared. A solution of 400pM LgTrip 3546 (SEQ ID NO: 51)+1uM pep263 (SEQ ID NO: 35) in TBS+0.01% BSA was prepared. 50ul of each enzyme solution was combined with 50ul of the detergent titrations, placed in luminometer, and read after a 3 minute incubation at RT. Results of LgBiT, LgTrip and NanoLuc activity in the presence of detergent are depicted in Figure 80.

Example 54

Reversibility of FRB-FKBP facilitated complex formation

Experiments were conducted during development of embodiments herein to demonstrate the reversibility of bioluminescent complex formation. Media was aspirated from a T75 growth flask of HEK293 cells. Cells were washed with 10ml of DPBS and trypsinized by adding 3ml of Tryple Express Trypsin. After a 3 minute incubation at 37°C,

10ml of growth media (DMDM+10% FBS) was added to the flask, mixing cells with pipette. Cells were pelleted at 200 rpm for 5 minutes. Media was aspirated and replaced by fresh media. Cells were counted on a T20 cell counter and diluted to 200,000 cells/ml. 3ml of the cell suspension was added to each well of a six well plate. Cells were grown overnight at 37°C with 5% CO_2. To transfect the cells, DNA was diluted for each construct to a concentration of 100ng/ul and 3.3ug of DNA was added in a final volume of 155ul of OptiMEM for each construct (FKBP-SmBiT, FRB-LgBiT, FRB-SmTrip10 pep86 (HiBiT), FKBP-SmTrip9 pep245). 9.9ul of FugeneHd was added to the diluted DNA and incubated for 15 minutes. 150ul of each DNA complex was then added to cells plated in a 6 well plate. Cells were grown overnight at 37°C with 5% CO_2. After aspirating media, cells were washed once with DPBS (Life Technologies Cat. No. 14190) and then frozen in a fresh 1ml of DPBS at -80°C. The samples were then thawed to lyse cells. FRB and FKBP constructs for NanoBiT (FKBP-SmBiT+FRB-LgBiT) and NanoTrip (FRB-SmTrip10 pep86 (HiBiT)+FKBP-SmTrip9 pep245+200nM purified LgTrip 3546 (SEQ ID NO: 51)) were combined and incubated with 30nM Rapamyacin for 30 minutes. A dilution series of FK506 was prepared in DMSO starting at 10mM. 3-fold serial dilutions were performed in DMSO (30ul into 70ul). 200ul of each FRB-FKBP combination was aliquoted into 8wells of a 96 well PCR tray. Upon addition of 2ul of the FK506 dilution series, each sample was incubated at 37°C for 6 hours. 50ul of each sample was combined with 50ul of TBS+0.01% BSA+20uM Furimazine (Fz), incubated for 3 minutes, and read on GMM+. Results are depicted in Figure 81.

Example 55

LgTrip/SmTrip9 titration with SmTrip10 peptides

Experiments were conducted during development of embodiments herein to analyze titrations of LgTrip 3546 (SEQ ID NO: 51) and SmTrip9 pep286 (SEQ ID NO: 37) with various SmTrip10 peptides. Data was normalized to SmTrip10 pep86 (HiBiT) values. SmTrip10 pep86 (HiBiT) is SmHiTrip10 (SEQ ID NO: 25).

Peptide stocks were diluted to 250uM in water. A SmTrip9 pep286 (SEQ ID NO: 37) solution (10uM in final reaction) was prepared in OptiMEM+10% FBS. A 2-fold serial dilution of each SmTrip10 peptide was performed in the OptiMEM solution containing SmTrip9 pep286. The highest concentration of the SmTrip10 peptide was 15uM (final in reaction). A 10X stock (1nM) of LgTrip 3546 (SEQ ID NO: 51) was prepared in

OptiMEM+10% FBS, and 10ul was added to 90ul of each of the SmTrip10 peptide titrations. Samples were incubated for 30 minutes on an orbital shaker set to 500rpm. 2ml of detection reagent (OptiMEM+10% FBS+20ul of 1M DTT+80ul of 5mM Furimazine) was prepared. 10ul of detection reagent was added to each LgTrip 3546:peptide solution, and plates were placed on an orbital shaker. Plates were read at 5 minutes and 10 minutes. SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) and SmTrip10 pep289 (SEQ ID NO: 150) were used as controls on each of the 4 plates. Results are depicted in the tables of Figures 82-83.

Example 56

Antares constructs

Experiments were conducted during development of embodiments herein to demonstrate the complementation systems described herein in the context of the Antares BRET system comprising one or more CyOFP fluorescent proteins linked to a component of the systems described herein.

Samples were purified using HisLink Resin: 10ml of 100mM HEPES pH 7.5, 1ml of FastBreak Cell Lysis Reagent and 50u of DNase were added, and samples were placed on rotating mixer for 45 minutes and then spun at 7,000rpm for 20 minutes. Next, 1ml of HisLink resin was added to each sample, and samples were washed 3X with 5ml of binding wash buffer, eluted with 300ul of elution buffer, and dialyzed against TBS (2 hours, TBS replace, 2 more hours). Samples were diluted to 100nM in TBS+0.01% BSA and then further diluted to 0.4nM by adding 4ul to 996ul of TBS+0.01% BSA. 3X serial dilutions were prepared by transferring 300ul to 700ul . 10ml of 2uM dipeptide pep263 (SEQ ID NO: 35) was prepared in TBS+0.01% BSA. 10ml of 400pM SmTrip10 pep86 (SEQ ID NO: 25) was prepared in TBS+0.01% BSA. 10ml of 1uM SmTrip9 pep286 (SEQ ID NO: 37) and 10uM SmTrip10 pep86 were prepared. 50ul of each enzyme was combined with either TBS or dipeptide solution (all samples in triplicate on two plates). Antares fusions with LgBiT and LgTrip 3546 samples were combined with SmTrip9 pep286+SmTrip10 pep86. Samples were incubated for 1 hour at RT. 100ul of 20um furimazine was added in TBS+0.01% BSA+2mM ATT. Plates were incubated for 3 minutes and then read on GMM+. Results are depicted in graphs of Figures 84-85.

Example 57

"Dark" dipeptide 272

Experiments were conducted during development of embodiments herein to compare titration series with "Dark" dipeptide 272 (SEQ ID NO: 146) with LgBiT (SEQ ID NO: 11) and LgTrip 3546 (SEQ ID NO: 51) in the presence of 0.1nM pep 263. LgBiT and LgTrip 3456 were diluted to 200nM in TBS+0.01% BSA and +/- 0.4nM of dipeptide pep263 (SEQ ID NO: 35) and incubated for 10 minutes. A 3X dilution series of dipeptide pep272 was prepared starting at 40nM (at this concentration, LgBiT showed inhibition at high concentrations, so Kd value could not be calculated; a new titration series was created starting at 4nM pep272 for LgBiT to obtain a Kd value). 50ul of the peptide dilution series was added to an assay plate followed by addition of 50ul of the LgBiT and LgTrip 3546 dilutions. Samples were incubated for 1 hour at room temperature. After addition of 100ul of NanoGlo+50uM Furimazine (Fz), plates were incubated for 5 minutes and luminescence was read on GMM+. Results are depicted in Figure 86.

Example 58

Comparison of dark dipeptides pep272 and pep273

LgBiT (SEQ ID NO: 11) and LgTrip 3546 (SEQ ID NO: 51) were diluted to 200nM in TBS + 0.01% BSA with +/- 0.4nM of dipeptide pep263 (SEQ ID NO: 35) or +/- 0.4nM didpeptide pep264 (SEQ ID NO: 299) and incubated for 10 minutes. 3X dilution series of didpeptide pep272 (SEQ ID NO: 146) and dipeptide pep273 (SEQ ID NO: 298) were prepared starting at 40nM using the dipeptide pep263 dilution as a diluent for pep272 and the dipeptide pep264 dilution as a diluent for pep273. 50ul of the LgBiT and LgTrip 3546 dilutions was combined with 50ul of the pep272/273 titration series and incubated at room temperature for 2 hours. After addition of 100ul of NanoGlo® buffer+50uM Fz, plates were incubated at room temperature for 5 minutes, and luminescence was read on GMM+. Results are depicted in Figure 87.

Example 59

DarkBiT pep167

Solutions with 200nM LgBiT (SEQ ID NO: 11) and LgTrip 3546 (SEQ ID NO: 51) were prepared. 0.2nM SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25) was added to LgBiT solutions, and 1uM of SmTrip9 pep286 (SEQ ID NO: 37) with 200nM of SmTrip10 pep86 was added to LgTrip 3546 solutions. A dark bit (pep167) (SEQ ID NO: 300) titration was prepared starting at 12uM in TBS+0.01% BSA. 50ul of the dark bit titration was combined

with 50ul of the LgBiT or LgTrip 3546/pep167 dilutions and incubated for 1 hour. After addition of 100ul of NanoGlo® buffer+50uM Furimazine (Fz), plates were incubated 10 minutes, and luminescence was read on GMM+. Results are depicted in Figure 88.

Example 60

FRB-FKBP facilitated complementation in E. coli lysates with SmTrip9 pep435/434 variants

Cultures were grown overnight in LB+100ug/ml ampicillin from glycerol stocks, and cells were diluted 1:100 in LB + 0.15% glucose + 0.1% rhamnose + Amp. After 20 hr shaking at 25ºC, cells were diluted 1:4 in PLB and incubated 15 min at room temperature to lyse. Lysates of SmTrip9/SmTrip10 peptide combinations of interest were mixed 1:1 vol:vol and diluted 1:5 in PLB+ 200 nM LgTrip 3546 (SEQ ID NO: 51) with or without 30 nM rapamycin. Samples were incubated for 30 minutes at room temperature and combined 1:1 (vol:vol) with NanoGlo® buffer containing 50uM Furimazine. Luminescence was read at 5 minutes. Results are depicted in Figures 89-90.

Example 61

FRB-FKBP facilitated complementation

FRB-FKBP facilitated complementation in HEK lysates with SmTrip9 pep435 and pep434 variants

600,000 cells were added to each well of 6-well plates in DMEM+1% FBS. Cells were grown overnight at 37ºC with 5% CO_2 and transfected with 3 µg DNA (FKBP or FRB construct) per well using FuGENE protocol. Following overnight incubation at 37ºC with 5% CO_2, cells were washed with DPBS. After aspiration, 1 ml of fresh DPBS was added to each well and plates were frozen at -80ºC for ~10 min. Plates were thawed at room temperature to lyse cells and lysates were cleared by centrifuging 10 min and removing supernatant. Lysates were diluted 2-fold in PLB and SmTrip9/SmTrip10 peptide combinations of interest were mixed 1:1 (vol:vol). Mixtures were then diluted 1:5 in PLB + 200 nM LgTrip 3546 (SEQ ID NO: 51) with or without 30 nM rapamycin. Samples were incubated for 30 minutes at room temperature and combined 1:1 vol:vol with NanoGlo® buffer containing 50uM Furimazine. Luminescence was read at 5 minutes. Results are depicted in Figure 91.

FRB-FKBP assay screen with SmTrip9s 823 and 840

Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10s were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1%

rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. PLB assay reagent was prepared with 444 nM LgTrip 3546, 90x diluted FRB-SmTrip10 culture, +/- 35 nM Rapamycin. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of NanoGlo containing 50 uM furimazine was added to assay plates wells and luminescence was read on GloMax after 5 minutes. Results are depicted in Figure 92.

Example 62

Determination of Kd of pep434 and pep435 variants

LgTrip 3546 (SEQ ID NO: 51) was diluted to 0.2 nM in TBS+0.01% BSA+0.01% Tergitol. 20 uM solutions of each SmTrip9-like peptides were prepared with 100 uM SmTrip10 pep86 (SEQ ID NO: 25) in TBS+0.01% BSA+0.01% Tergitol. 2-fold serial dilutions of each SmTrip9-like peptide were performed using the 100 uM SmTrip10 pep86 solution as a diluent. Peptide dilutions were combined with LgTrip 3546 solution 1:1 (vol:vol) and incubated 10 minutes. TBS+0.01% BSA+0.01% Tergitol+20uM Furimazine (Fz) detection reagent was added to LgTrip/peptide solutions 1:1 vol:vol and luminescence was read at 10 min. Results are depicted in Figure 93.

Example 63

Detection of CRISPR-tagged dipeptide-GAPDH using LgTrip 3546

Experiments were conducted during development of embodiments herein to demonstrate that both LgTrip 2098 (SEQ ID NO: 31) and LgTrip 3546 (SEQ ID NO: 51) find use as bioluminescence reagents for detecting endogenously tagged GAPDH (Tagged with SmTrip10 pep86 (SmHiTrip; SEQ ID NO: 25).

HeLa cells were edited using CRISPR/Cas9 to express endogenous GAPDH C-terminal fusions to the indicated peptide. The edited HeLa cells were plated at a density of approximately 20,000 cells per well of a solid white assay plate in 100 µl of DMEM/10% FBS. Cells were then incubated in the presence of 100 µl of NanoGlo® HiBiT Lytic Buffer (Promega) containing NanoGlo® HiBiT Lytic Buffer and 200 nM of LgTrip for 10 min. Luminescence was recorded using a GloMax® Discover with 0.5s integration time. Relative cell numbers were determined using the CellTiter® Glo Luminescent Cell Viability Assay (Promega) according to manufacturer's protocol. Data are represented as average relative light units normalized to cell number, with variability expressed as standard deviation.

Results are depicted in Figure 94.

Example 64

Site-Saturation Screen of SmTrip9

Experiments were conducted during development of embodiments herein to identify beneficial amino acid substitutions in SmTrip9.

Genetic site-saturation libraries were generated using primers with randomized codons at the indicated positions in SmTrip9. KRX *E. coli* was transformed with pooled genetic variants, plated onto LB + ampicillin agar, and grown overnight at 37°C. Individual colonies were picked and placed into 96-well culture plates containing LB + 100 ug/ml ampicillin. Cultures were grown overnight at 37°C with shaking. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. Assay reagent was prepared by adding 444 nM LgTrip (SEQ ID NO: 51), 90x diluted FRB-VS-HiBiT culture, and +/- 35 nM rapamycin to 25 mM HEPES with 0.3x Passive Lysis Buffer (PLB) and DNase. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB, and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of NanoGlo® buffer containing 50 uM furimazine was added to assay plates wells, and luminescence was read on GloMax® luminometer after 5 minutes.

Results are depicted in Figures 100-112.

Example 65

FRB-FKBP facilitated complementation in E. coli lysates with SmTrip9 pep435/434 variants

Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10s were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. Cultures were diluted 1:4 in PLB and incubated 15 min at room temperature to lyse cells. SmTrip9 and SmTrip10 dilutions were mixed 1:1 (vol:vol) for combinations of interest. Mixtures were diluted 1:5 into PLB + 200 nM LgTrip , with or without 30 nM rapamycin. Samples were incubated 30 min at room temperature. Fifty microliters of NanoGlo® buffer containing 50 uM furimazine was added to assay plates wells, and luminescence was read on GloMax® luminometer after 5 minutes. Results are depicted in Figures 113-115.

Example 66

FRB-FKBP Facilitated Complementation Assay Screen with Combinational SmTrip9 variants

Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10s were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. PLB assay reagent was prepared with 444 nM LgTrip, 90x diluted FRB-SmTrip10 culture, +/- 35 nM Rapamycin. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB, and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of NanoGlo® buffer containing 50 uM furimazine was added to assay plates wells, and luminescence was read on GloMax® luminometer after 5 minutes. Results are depicted in Figures 116-122.

Example 67

Determination of Kd and Bmax of SmTrip9 synthetic peptides

LgTrip was diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol and pep289 was added to 25 uM. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT. Results are depicted in Figures 123-130.

Example 68

Determination of Solubility of Synthetic SmTrip9 peptides

Synthetic peptides were ordered from Peptide2.0 with termini blocked (N-terminal acetylation and C-terminal amidation) unless otherwise noted. Peptides were dissolved in nuclease-free water and stored at -20°C. Stocks were thawed in 22°C water bath, centrifuged, and kept at 4°C until use. Results are depicted in Figures 131.

Example 69

Biochemical co-titration of SmTrip9 synthetic peptides and pep289

LgTrip was diluted to 200 nM in 25 mM HEPES with 0.3x Passive Lysis Buffer (PLB) and DNase. SmTrip9 peptides and pep289 were diluted to 100 uM and co-titrated serially 6-fold in PLB. Samples were incubated 10 minutes at room temperature. Most concentrated samples were diluted 50-100-fold in PLB. Samples were aliquoted in triplicate into assay plates and mixed 1:1 vol:vol with NanoGlo® buffer + 50uM furimazine. Luminescence was read after 10 minutes on ClarioStar or GloMax® instruments. Results are depicted in Figures 132-133.

Example 70

Biochemical co-titration of SmTrip9 and SmTrip 10 synthetic peptides

LgTrip was diluted to 200 nM in 25 mM HEPES with 0.3x Passive Lysis Buffer (PLB) and DNase. SmTrip9 and SmTrip10 peptides were diluted to 100 uM and co-titrated serially 6-fold in PLB. Samples were incubated 10 minutes at room temperature. Most concentrated samples were diluted 50-100-fold in PLB. Samples were aliquoted in triplicate into assay plates and mixed 1:1 vol:vol with NanoGlo® buffer + 50uM furimazine. Luminescence was read after 10 minutes on ClarioStar or GloMax® instruments. Results are depicted in Figure 134.

Example 71

Biochemical co-titration of pep521 and alternative SmTrip 10 synthetic peptides

LgTrip was diluted to 200 nM in 25 mM HEPES with 0.3x Passive Lysis Buffer (PLB) and DNase. SmTrip10 peptides and pep521 were diluted to 100 uM and co-titrated serially 6-fold in PLB. Samples were incubated 10 minutes at room temperature. Most concentrated samples were diluted 50-100-fold in PLB. Samples were aliquoted in triplicate into assay plates and mixed 1:1 vol:vol with NanoGlo® buffer + 50uM furimazine. Luminescence was read after 10 minutes on ClarioStar or GloMax® instruments. Results are depicted in Figure 135.

Example 72

Strand removal (purification) from LgTrip 3546 template

A single colony from each clone was grown for 18 hours at 37°C in LB+100ug/ml ampicillin. The overnight culture was diluted 1:100 into 50ml of Terrific Broth+0.1% Rhamnose+100ug/ml ampicillin. After 48 hours of growth at 15°C, cells were pelleted and resuspended in 10ml of 100mM HEPES pH 7.5+0.001U/ml DNase. 1ml of FastBreak® Lysis Buffer was added to each sample, and then samples incubated on a rotating mixer at 4°C for 1 hour. A cleared lysate was prepared by centrifugation of 7,000 RPM for 10 minutes.

Purification of the strands using the MagneHis purification system: 300ul µl of MagneHis resin (Promega) was added to each sample, and then samples mixed 20 times and placed on a magnetic stand. The supernatant was removed, and the resin was washed two times with column wash buffer. Samples were eluted in 600ul of elution buffer. Samples were then placed in a dialysis apparatus to exchange with TBS. Identification of the strand removal proteins was observed via SDS PAGE as depicted in Figure 136.

Example 73

Strand Removal Proteins with Various Combinations of Peptides

200µl of OptiMEM+10% FBS was added to multiple wells of a multi-well plate. Peptide combinations were added to a final concentration of 10µM with each to be assayed separately with each strand removal protein. Each strand removal protein was diluted to 20nM (2nM for LgTrip 3546) in OptiMEM+10% FBS. 20µl of each strand removal peptide was added to the designated peptide combination, samples e mixed, and 45µl aliquoted in triplicate into wells of a white assay plate (Costar 3600). After 15-minute incubation at RT, 5µl of detection reagent (100uM Fz (Promega LCS N205)) was added to each sample. Samples were placed on an orbital shaker for 30 seconds, and then luminescence was measured every 2 minutes for 1 hour. Luminescence is reported as peak height of the kinetic read. Background is OptiMEM+10% FBS+detection reagent.

As demonstrated in Figure 137, there was no signal over background for strand removal proteins 7, 8, 9, 10 when added as separate peptides. Two of the three peptide combinations gave ~2X signal over background ((8+9) dipeptide+7+10) or ((7+8)dipeptide+9+10). One of the 3 peptide combinations gave ~10X signal over background (((9+10) dipeptide+7+8) The two dipeptide combination of (10+9)+(7+8) gave signal of ~4.5 logs over background. It is likely that the peptide combinations that gave the

greatest signal have the highest affinity. Lower affinity combinations could produce light in a facilitated complementation assay. Figure 137D demonstrates that peptides with alternative split sites (e.g., mid beta strand) are capable of forming a biolumiunescent complex.

Example 74

Strands 6, 7, 8, 9, or 10 Removal (purification) from LgTrip 3546 template

400µl of OptiMEM+10% FBS was added to multiple wells of a deep well 96-well plate. Peptide combinations were added to a final concentration of 10µM each peptide to be assayed separately with either ATG-3929 or LgTrip. The peptide solutions were then divided. To one of the peptide aliquots, 20ul of either 20nM ATG-3929 or 2nM LgTrip was added to the designated peptide combination, samples mixed, and 45µl of the +/- peptide samples aliquoted in triplicate into wells of a white assay plate (Costar 3600). After a 15-minute incubation at RT, 5µl of detection reagent (100uM Fz in OptiMEM+10% FBS (Promega LCS N205)) was added to each sample. Samples were placed on an orbital shaker for 5 minutes. Background for each sample is OptiMEM+10% FBS+peptide dilutions+detection reagent.

As demonstrated in Figure 138, sample ATG-3929 with strands (9+10)+(7+8)+6 shows ~2X signal over background. On the other hand, the sample with two peptides (6+7+8)+(9+10) showed ~300X over background.

Note that spontaneous complementation is not visible for samples with more than 3 peptides. It is possible that the affinity is not high enough affinity of the peptides is not high enough to produce light. It is possible that if the peptides are brought together through facilitated complementation with a fusion partner that it would be possible to obtain signal.

Example 75

Dipeptide Titrations

Dipeptides were diluted to 5 uM and diluted serially 5-fold using TBS + 0.01% BSA + 0.01% Tergitol with 0.2 nM of LgTrip as the diluent. Samples were incubated 10 minutes at room temperature and added to wells of assay plates in triplicate. One-to-one vol:vol of TBS + 0.01% BSA + 0.01% Tergitol with 20x diluted live cell substrate was added to samples and plates were read on a GloMax luminometer after 10 minutes. Figure 139A-B demonstate the Kd and Bmax values from the dipeptide titrations.

Fold response of Binary NanoTrip in mammalian cells

Growth media was removed from confluent flasks of cells. (HEK293 and Hela). Cells were washed with 10ml of DPBS and then 3ml of TrypLE Express trypsin was added to cells. Cells were incubated for 3 minutes at 37°C. 10ml of growth media was added and then cells were spun at 200RCF for 5 minutes. Media was replaced and cells were resuspend in 10ml of growth media. Cells were counted and diluted to 200,000/ml. 100ul of cells were plated into each well of a white assay plates and grow overnight at 37°C with CO_2. The next day 100ng/ul DNA from FRB and FKBP fusions of LgTrip (3546) and various dipeptide in each orientation were combined. 263 samples started at 1:10 dilution in carrier DNA or 10ng/ul. DNA samples were then diluted serially into carrier DNA (10ul to 90µl in 100ng/ul carrier DNA) Next 20ul of each DNA dilution was added to 83ul of OptiMEM. Samples were mixed and then 6.6 ul of Viafect transfection reagent was added to each sample. Samples were incubated for 20 minutes at RT and then 5ul of transfection complex was added to 6 wells of cells for each FRB-FKBP orientation. Plates were then grown overnight at at 37°C with CO_2. The next day Rapamycin (RAP) was added to 3 of the wells for each sample to a final concentration of 100nM. Samples were placed on orbital shaker for 1 minute and then Incubated at $_{37C}$ for 30 minutes. After incubation, 100ul of NanoGlo+50uM Fz was added to each sample (+RAP and –RAP) and then samples were placed on orbital shaker for 5 minutes. Luminescent measurements were acquired using a Glomax Discover luminometer. Fold response was calculated by dividing RLU values from the +RAP sample by the RLU values from the –RAP samples. Results are depicted I Figure 139C-E. Dipeptide fusions that have lower affinity to LgTrip produce a greater fold response compared to samples with higher affinity.

Example 76

Development of a Tripartite Quantitative Assay for anti-TNFa Biologic Agents using Tripartite Fusion Proteins

Infliximab (Remicade), Adalimumab (Humira), and Etanercept (Enbrel) are TNFa inhibitors that all bind human TNFa and also all contain a human IgG1 Fc. A quantitative assay was deveeloped for all 3 TNFa inhibitors by expressing and purifying SmTrip9- or SmTrip10-protein G and TNFa fusion proteins which serve as the binding components to the TNFa inhibitor (Figure 140). The Protein G fusion protein will bind to the conserved IgG1 Fc region of the TNFa inhibitor. The Inhibitor will bind to the TNFa fusion protein bringing the SmTrip9 and SmTrip10 into close proximity. In the presence of LgTrip, the

bioluminescent complex will form creating the signal that is proportional to the amount of TNFa inhibitor present. All reporter tag configurations were tested with SmTrip9 or SmTrip10 expressed on the N- or C-terminal of Protein G and TNFa with either a 4gly-ser or 15gly-ser linker. The optimal pairing resulting from screening all orientations was SmTrip9-15gly/ser-protein G with TNFa-15gly/ser-SmTrip10.

Methods for making the fusion proteins

A fusion protein comprising of SmTrip9 pep521 (SEQ ID NO: 268) sequence followed by a linker of 15 glycine-serine repeat was fused to the N-terminus of Protein G was expressed and purified. A second fusion protein comprising of SmTrip10 pep289 (SEQ ID NO: 150) sequence was fused to the C terminus of human TNFa separated by a linker of 15 glycine-serine repeat was also expressed and purified. Streak plates from glycerol stocks of KRX transformed E. coli cells were created on LB plates with Ampicillin (100ug/ml) and allowed to incubate overnight at 37°C. A single colony was inoculated into 3mls of SOC media + AMP and incubated shaking (275 rpm) overnight at 37°C. The cells were lysed and the plasmid DNA was collected. Shuffle competent E. coli cells were transformed with 100ng of plasmid DNA, spread onto pre-warmed selection plates, and allowed to incubate overnight at 30°C. A colony was selected and inoculated into a 50ml volume of LB containing ampicillin. The cultures were incubated overnight at 37°C shaking before being diluted 1:100 into 500mL of LB medium containing ampicillin. These flasks were allowed to incubate at 37°C while shaking until the OD600 reached 0.6-0.8. Cells were induced by addition of IPTG at a final concentration of 1mM and allowed to incubate overnight at 25°C while shaking. Cells were harvested, centrifuged, and resuspended in 50mL extraction and lysis buffer at 4°C with mixing. Three cycles of freeze/thaw were performed followed by addition of RQI DNase. The total lysate was transferred to a think 50mL centrifuge tube and spun at 10,000xg for 30 minutes at 4°C. 20mM Imidazole/350mM NaCl was added prior to loading onto a nickel column. Fusion proteins were washed and eluted off the columns in a 5 step elution process with increasing imidazole. Samples were dialyzed against TBS and final stock proteins were stored in 50% glycerol in TBS at -20°C.

Example 77

Homogeneous Quantitative Analysis of TNFa Inhibitors Infliximab, Adalimumab, and Etanercept using SmTrip9 pep521-protein G and TNFa-SmTrip10 pep289 Fusion Proteins

Experiments were conducted during development of embodiments herein to determine the ability of NanoTrip fusion proteins to quantitate TNFa inhibitors in a homogeneous assay. The results show that protein G and TNFa NanoTrip fusion proteins together with LgTrip display great sensitivity and range for quantitating infliximab, adalimumab, and etanercept.

A 2X stock of the TNFa inhibitors was generated in assay buffer, serially diluted 1:2 to create a dose response, and 50ul/well was added to a non-binding surface treated, 96 well solid-white plate (Costar 3600). A 2X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9 pep521-protein G (SEQ ID NO: 268) (final 10nM) + TNFa-SmTrip10 pep289 (SEQ ID NO: 150) (final 10nM) was created in assay buffer, and 50ul/well added. Plates were allowed to incubate at room temperature for 90 minutes. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, 25ul/well added to the plate for a final concentration of 10uM, allowed to incubate for ~ 5 minutes, and luminescence measured using a GloMax® Discover. Assay buffer consisted of Blocker BSA (10%) (Thermo) diluted in PBS (pH 7.0) to a final of 0.01% BSA in PBS. Samples were tested in triplicate/plate, and n=3 independent experiments run. Data as demonstrated in Figure 141 was analyzed for limit of detection (LOD), limit of quantitation (LOQ), and upper limit of quantitation (ULOQ).

Example 78

Homogenous Quantitative Analysis of Infliximab in Complex Sample Matrices such as Human Serum and Urine

Experiments were conducted during development of embodiments herein to determine the ability of NanoTrip fusion proteins to quantitate infliximab in the presence of the complex sample matrices of normal human IgG depleted serum, normal pooled human AB serum, and pooled normal human urine in a homogenous assay. Results indicate that the NanoTrip system was largely unaffected by the presence of urine nor the presence of serum proteins with the exception of endogenous IgG as expected.

A 2X stock containing 20nM Infliximab in presence of the human sample matrix to be tested was created by diluting with assay buffer, and 50ul/well added to a non-binding

surface treated, 96 well solid-white plate (Costar 3600). A 2X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9 pep521-protein G (SEQ ID NO: 268) (final 10nM) + TNFa-SmTrip10 pep289 (SEQ ID NO: 150) (final 10nM) was created in assay buffer, and 50ul/well added. Plates were allowed to incubate at room temperature for 90 minutes. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, 25ul/well added to the plate for a final concentration of 10uM, allowed to incubate for ~ 5 minutes, and luminescence measured using a GloMax® Discover. Assay buffer consisted of Blocker BSA (10%) (Thermo) diluted in PBS (pH 7.0) to a final of 0.01% BSA in PBS. Samples were tested in triplicate. Data as demonstrated in Figure 142 is displayed as signal/background.

Example 79

Kinetic Analysis of Signal Generation via Facilitated Complementation of SmTrip9 pep521-protein G (SEQ ID NO: 268) and TNFa-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO:150) Fusion Proteins with Purified LgTrip 3546 (SEQ ID NO: 51) in the Presence of 100pM of Infliximab in a Solution Phase, Homogenous Assay.

Experiments were conducted during development of embodiments herein to determine the binding kinetics of the Protein G/TNFa NanoTrip system to quantitate 100pM of Infliximab in a solution phase, homogenous assay. Results show that signal generation is immediate and sustained indicating rapid binding kinetics of the fusion proteins to infliximab as well as LgTrip to the SmTrip9 and SmTrip10 fusion proteins.

A 2X stock of Infliximab (100pM final) was generated in assay buffer, and 50ul/well added to a non-binding surface treated, 96 well solid-white plate (Costar 3600). A 2X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9 pep521-protein G (SEQ ID NO: 268) (final 10nM) + TNFa-SmTrip10 pep289 (SEQ ID NO: 150) (final 10nM) was created in assay buffer, and 50ul/well added. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, and 25ul/well added to the plate for a final concentration of 10uM. All reagents were added, and the plate immediately placed on a GloMax® Discover to read luminescence over time. Assay buffer consisted of Blocker BSA (10%) (Thermo) diluted in PBS (pH 7.0) to a final of 0.01% BSA in PBS. Samples were tested in triplicate.

Results are depicted in Figure 143.

Example 80

Testing SmTrip9-Protein G Variants for Their Ability to Measure Infliximab via Facilitated Complementation with TNFa-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO:150) Fusion Proteins Purified LgTrip 3546 (SEQ ID NO: 51) in a Solution Phase, Homogenous Assay

Experiments were conducted during development of embodiments herein to determine the ability of other SmTrip9 variants expressed as a fusion proteins to protein G to measure Infliximab via facilitated complementation with TNFa-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO:150) fusion proteins purified LgTrip 3546 (SEQ ID NO: 51) in a solution phase, homogenous assay. Results show that all of the SmTrip9 pep(x)-Protein G variants tested were able to generate signal.

A 2X stock of Infliximab (10nM final) was generated in assay buffer, and 50ul/well added to a non-binding surface treated, 96 well solid-white plate (Costar 3600). A 2X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9 pep(x)-Protein G (final 10nM) + TNFa-SmTrip10 pep289 (SEQ ID NO: 150) (final 10nM) was created in assay buffer, and 50ul/well added. Plates were allowed to incubate at room temperature for 90 minutes. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, 25ul/well added to the plate for a final concentration of 10uM, allowed to incubate for ~ 5 minutes, and luminescence was measured using a GloMax® Discover. Assay buffer consisted of Blocker BSA (10%) (Thermo) diluted in PBS (pH 7.0) to a final of 0.01% BSA in PBS. Samples were tested in triplicate. Results are depicted in Figure 144.

Example 81

Homogeneous Quantitative Infliximab Testing SmTrip9 pep(X)-Protein G Variants and TNFa-SmTrip10 pep289 Fusion Proteins

Experiments were conducted during development of embodiments herein to demonstrate the ability of different SmTrip9 pep(X)-Protein G variants to quantitate Infliximab via facilitated complementation with TNFa-SmTrip10 pep289 (VS-HiBiT; SEQ ID NO:150) fusion proteins with purified LgTrip 3546 (SEQ ID NO: 51) in a solution phase, homogeneous assay. Results show that all SmTrip9 variants were able to quantitate infliximab.

A 2X stock of Infliximab (10nM final) was generated in assay buffer, and 50ul/well added to a non-binding surface treated, 96 well solid-white plate (Costar 3600). A 2X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9

pep(x)-protein G (final 10nM) + TNFa-SmTrip10 pep289 (SEQ ID NO: 150) (final 10nM) was created in assay buffer, and 50ul/well added. Plates were allowed to incubate at room temperature for 90 minutes. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, 25ul/well added to the plate for a final concentration of 10uM, allowed to incubate for ~ 5 minutes, and luminescence was measured using a GloMax® Discover. Assay buffer consisted of Blocker BSA (10%) (Thermo) diluted in PBS (pH 7.0) to a final of 0.01% BSA in PBS. Samples were tested in triplicate.

Example 82

Development of a Tripartite Quantitative Assay for anti-EGFR Biologic Agents Using Tripartite Fusion Proteins in a Cell-Based Assay.

We developed a quantitative, cell-based assay for panitumumab and cetuximab representing a phase separation or surface chemistry like assay. Using purified SmTrip9-Protein G fusion proteins that will bind to the conserved human IgG Fc region of the EGFR inhibitor, the Inhibitor will bind to the SmTrip10-EGFR fusion protein that is expressed on the cell surface bringing the SmTrip9 and SmTrip10 into close proximity. In the presence of LgTrip, the bioluminescent complex will form creating the signal that is proportional to the amount of EGFR inhibitor present. All reporter tag configurations were tested with SmTrip9 or SmTrip10 expressed on the N- or C-terminal of protein G or on the N terminal of EGFR with either a 4gly-ser or 15gly-ser linker. The optimal pairing resulting from screening all orientations was SmTrip9-4gly/ser-protein G with EGFR-15gly/ser-SmTrip10.

Results are depicted in Figure 145.

Example 83

Quantitation of Panitumumab via Facilitated Complementation with SmTrip9 pep521-protein G (SEQ ID NO: 268) Fusion Protein and SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) Expressing Cells with Purified LgTrip 3546 (SEQ ID NO: 51) in a Cell-Based Homogeneous Assay.

Experiments were conducted during development of embodiments herein to determine the ability of NanoTrip fusion proteins to quantitate the EGFR inhibitor panitumumab in a cell-based homogeneous assay. The results show that SmTrip9 pep521-protein G (SEQ ID NO: 268) purified protein, SmTrip10 pep289-EGFR (SEQ ID NO:150) expressing cells, and LgTrip 3546 (SEQ ID NO: 51) display great sensitivity and range for quantitating panitumumab.

HEK293 cells were maintained in growth medium (DMEM) supplemented with 10% fetal bovine serum (FBS, Hyclone) at 37°C/5% CO_2 in a humidified tissue culture incubator. Transient reverse transfection were performed by first diluting the expression construct for the SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) into Opti-MEM containing carrier DNA (PGEM-3ZF(-)) at a mass ratio of 1:10. The transfection reagent:DNA complex was prepared by adding FuGENE HD transfection reagent at a ratio of 1:3 (mg DNA per mL FuGENE HD) followed by 15 minutes incubation at room temperature. The resulting transfection:DNA complex was then mixed with a HEK293 cell suspension (2 x 10^5 cells/ml) in growth medium at a ratio of 1:20 (vol/vol), followed by incubation for 18-20 hours at 37°C/5% CO_2 in humidified tissue culture incubator.

HEK293 cells expressing the SmTrip10 pep289-EGFR (SEQ ID NO: 150) fusion protein were harvested using Trypsin-EDTA, washed in growth medium, and resuspended in Opti-MEM at a concentration of 4.5 x 10^5 cells/ml. 50ul of cells/well (20,000 cells/well) are added to a non-binding surface, solid white 96 well plate (Costar 3600). A 4X stock of Panitumumab was generated in Opti-MEM, serially diluted in Opti-MEM to create dose response, and 25ul/well added. A 4X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9 pep521-protein G (SEQ ID NO: 268) (final 5nM) was created in Opti-MEM, and 25ul/well added. Plates were allowed to incubate for 1 hour at 37°C. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, 25ul/well added to the plate for a final concentration of 10uM, and luminescence was measured on a GloMax® Discover. Samples were tested in triplicate. N=3 independent experiments.

Results are depicted in figure 146.

Example 84

Real-Time Binding Kinetic Analysis of Signal Generation via Facilitated Complementation of SmTrip9 pep521-Protein G (SEQ ID NO: 268) Purified Fusion Protein and SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) Expressing HEK293 Cells Paired with Purified LgTrip 3546 (SEQ ID NO: 51) in the Presence of Increasing Doses of Cetuximab in a Cell-Based Homogeneous Assay.

Experiments were conducted during development of embodiments herein to determine the binding kinetics of the Protein G/EGFR NanoTrip system to quantitate Cetuximab in a cell-based homogenous assay. Results show that the luminescent signal

increases with time in accordance with the formation of the luciferase complex. Signal generation is also dose dependent.

HEK293 cells were maintained in growth medium (DMEM) supplemented with 10% fetal bovine serum (FBS, Hyclone) at 37°C/5% CO_2 in a humidified tissue culture incubator. Transient reverse transfection were performed by first diluting the expression construct for the SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) into Opti-MEM containing carrier DNA (PGEM-3ZF(-)) at a mass ratio of 1:10. The transfection reagent:DNA complex was prepared by adding FuGENE HD transfection reagent at a ratio of 1:3 (mg DNA per mL FuGENE HD) followed by 15 minutes incubation at room temperature. The resulting transfection:DNA complex was then mixed with a HEK293 cell suspension (2 x 10^5 cells/ml) in growth medium at a ratio of 1:20 (vol/vol), followed by incubation for 18-20 hours at 37°C/5% CO_2 in humidified tissue culture incubator.

HEK293 cells expressing the SmTrip10 pep289-EGFR fusion protein were harvested using Trypsin-EDTA, washed in growth medium, and resuspended in Opti-MEM at a concentration of 4.5 x 10^5 cells/ml. 50ul of cells/well (20,000 cells/well) were added to a non-binding surface, solid white 96 well plate (Costar 3600). A 4X stock of cetuximab was generated in Opti-MEM, and 25ul/well added. A 4X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 10uM) + SmTrip9 pep521-protein G (SEQ ID NO: 268) (final 780pM) was created in Opti-MEM, and 25ul/well added. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, and 25ul/well added to the plate for a final concentration of 10uM. All reagents were added, and the plate was immediately placed on a GloMax® Discover to read luminescence over time. Samples were tested in triplicate.

Results are depicted in Figure 147.

Example 85

Testing SmTrip9-Protein G Variants for Their Ability to Measure Panitumumab via Facilitated Complementation with SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) Expressing Cell Paired with Purified LgTrip 3546 (SEQ ID NO: 51) in a Cell-Based Homogenous Assay.

Experiments were conducted during development of embodiments herein to determine the ability of other SmTrip9 variants expressed as a fusion proteins to protein G to measure Panitumumab via facilitated complementation with SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) expressing cells paired with purified LgTrip 3546 (SEQ ID NO: 51)

in a cell-based homogenous assay. Results show that all of the SmTrip9 pep(x)-protein G variants tested were able to generate signal.

HEK293 cells were maintained in growth medium (DMEM) supplemented with 10% fetal bovine serum (FBS, Hyclone) at 37°C/5% CO_2 in a humidified tissue culture incubator. Transient reverse transfection were performed by first diluting the expression construct for the SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) into Opti-MEM containing carrier DNA (PGEM-3ZF(-)) at a mass ratio of 1:10. The transfection reagent:DNA complex was prepared by adding FuGENE HD transfection reagent at a ratio of 1:3 (mg DNA per mL FuGENE HD) followed by 15 minutes incubation at room temperature. The resulting transfection:DNA complex was then mixed with a HEK293 cell suspension (2 x 10^5 cells/ml) in growth medium at a ratio of 1:20 (vol/vol), followed by incubation for 18-20 hours at 37°C/5% CO_2 in humidified tissue culture incubator.

HEK293 cells expressing the SmTrip10 pep289-EGFR (SEQ ID NO: 150) fusion protein were harvested using Trypsin-EDTA, washed in growth medium, and resuspended in Opti-MEM at a concentration of 4.5 x 10^5 cells/ml. 50ul of cells/well (20,000 cells/well) were added to a non-binding surface, solid white 96 well plate (Costar 3600). A 4X stock of Panitumumab (final 1nM) was generated in Opti-MEM, and 25ul/well added. A 4X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9 pep(X)-protein G (final 10nM) was created in Opti-MEM, and 25ul/well added. Plates were allowed to incubate for 1 hour at 37C. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, 25ul/well added to the plate for a final concentration of 10uM, and luminescence was measured on a GloMax® Discover. Samples were tested in triplicate. N=3 independent experiments. Results are depicted in Figure 148.

Example 86

Testing SmTrip9-Protein G Variants for Their Ability to Measure Panitumumab via Facilitated Complementation with SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) Expressing Cell Paired with Purified LgTrip 3546 (SEQ ID NO: 51) in a cell-Based Homogenous Assay.

Experiments were conducted during development of embodiments herein to determine the ability of other SmTrip9 variants expressed as a fusion proteins to protein G to measure panitumumab via facilitated complementation with SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) expressing cells paired with purified LgTrip 3546 (SEQ ID NO: 51)

in a cell-based homogenous assay. Results show that all of the SmTrip9 pep(x)-protein G variants tested were able to quantitate panitumumab in a dose response analysis.

HEK293 cells were maintained in growth medium (DMEM) supplemented with 10% fetal bovine serum (FBS, Hyclone) at 37°C/5% CO_2 in a humidified tissue culture incubator. Transient reverse transfection were performed by first diluting the expression construct for the SmTrip10 pep289-EGFR (VS-HiBiT; SEQ ID NO:150) into Opti-MEM containing carrier DNA (PGEM-3ZF(-)) at a mass ratio of 1:10. The transfection reagent:DNA complex was prepared by adding FuGENE HD transfection reagent at a ratio of 1:3 (mg DNA per mL FuGENE HD) followed by 15 minutes incubation at room temperature. The resulting transfection:DNA complex was then mixed with a HEK293 cell suspension (2 x 10^5 cells/ml) in growth medium at a ratio of 1:20 (vol/vol), followed by incubation for 18-20 hours at 37°C/5% CO_2 in humidified tissue culture incubator.

HEK293 cells expressing the SmTrip10 pep289-EGFR (SEQ ID NO: 150) fusion protein were harvested using Trypsin-EDTA, washed in growth medium, and resuspended in Opti-MEM at a concentration of 4.5 x 10^5 cells/ml. 50ul of cells/well (20,000 cells/well) were added to a non-binding surface, solid white 96 well plate (Costar 3600). A 4X stock of Panitumumab (final 1nM) was generated in Opti-MEM, and 25ul/well added. A 4X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9 pep(X)-protein G (final 10nM) was created in Opti-MEM, and 25ul/well added. Plates were allowed to incubate for 1 hour at 37°C. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, 25ul/well added to the plate for a final concentration of 10uM, and luminescence was measured on a GloMax® Discover. Samples were tested in triplicate. Results are depicted in Figure 149.

Example 87

Quantitation of Human IL-1beta using NanoTrip Chemically Labeled-Paired Antibodies

Experiments were conducted during development of embodiments herein to demonstrate the use of paired monocloncal antibodies that have been chemically conjugated with NanoTrip peptides to quantitation human IL-1beta. This model system consists of two monoclonal mouse antibodies that recognize IL-1beta at different epitopes. HaloTag-SmTrip9 pep521 (SEQ ID NO: 268) was chemically conjugated to one of the antibodies, and HaloTag-SmTrip10 pep289 (SEQ ID NO: 150) was chemically conjugated to the other

antibody. In the presence of IL-1beta, the two antibodies bind to the IL-1beta thus bringing the two tags in close proximity. Addition of LgTrip 3546 (SEQ ID NO: 51) completes the complementation, and a luminescent signal is generated.

HaloTag-SmTrip9 and HaloTag-SmTrip10 fusion proteins are expressed and purified. Anti-IL-1beta mouse monoclonal antibody clone 508A 4A2 (Thermo) is labeled with the HaloTag-SmTrip9 pep521 (SEQ ID NO: 268) and anti-IL-1beta mouse monoclonal antibody clone 508A 7G8 (Thermo) is labeled with the HaloTag-SmTrip10 pep289 (SEQ ID NO: 150). The unlabeled antibodies are prepped by first doing a buffer exchange into 10mM $NaHCO_3$ (pH 8.5) using a Zeba column. Antibodies are then primed with a 20-fold excess of HaloTag® Succinimidyl Ester (04) Ligand (Promega) and allowed to incubate at room temperature for 2 hours. A buffer exchange is done 2x using Zeba columns to remove free linker. The primed antibodies are incubated with a 4-fold excess of HaloTag-SmTrip9 or HaloTag-SmTrip10 overnight at 4C while mixing. HaloLink® Resin is used to remove any free HaloTag® fusion proteins.

A 2X stock of recombinant human IL-1beta was generated in assay buffer, serially diluted 1:2 to create a dose response, and 50ul/well added to a non-binding surface treated, 96 well solid-white plate (Costar 3600). A 2X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9 pep521 labeled 4A2 clone (SEQ ID NO: 268) (final 100ng/ml) + SmTrip10 pep289 labeled 7G8 clone (SEQ ID NO: 150) (final 100 ng/ml) was created in assay buffer, and 50ul/well added. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, 25ul/well added to the plate for a final concentration of 10uM, and luminescence measured in real-time using a GloMax® Discover. Assay buffer consisted of Blocker BSA (10%) (Thermo) diluted in PBS (pH 7.0) to a final of 0.01% BSA in PBS. Samples were tested in triplicate. Data displayed is the signal that was read at the 20 minute time point.

Results are depicted in Figure 150.

Example 88

Real-time Binding Kinetics for Human Troponin using NanoTrip Chemically-labeled Paired Antibodies

Experiments were conducted during development of embodiments herein to demonstrate the use of paired monocloncal antibodies that have been chemically conjugated with NanoTrip peptides to quantitation human Troponin. This model system consists of two

monoclonal mouse antibodies that recognize Troponin at different epitopes. HaloTag-SmTrip9 pep521 (SEQ ID NO: 268) was chemically conjugated to one of the antibodies, and HaloTag-SmTrip10 pep289 (SEQ ID NO: 150) was chemically conjugated to the other antibody. In the presence of Troponin, the two antibodies bind to the Troponin thus bringing the two tags in close proximity. Addition of LgTrip 3546 (SEQ ID NO: 51) completes the complementation and a luminescent signal is generated.

HaloTag-SmTrip9 and HaloTag-SmTrip10 fusion proteins are expressed and purified. Anti-troponin mouse monoclonal antibody 10-T79C (Fitzgerald) is labeled with the HaloTag-SmTrip10 pep289 (SEQ ID NO: 150), and anti-troponin mouse monoclonal antibody 10-T79F (Fitzgerald) is labeled with the HaloTag-SmTrip9 pep521 (SEQ ID NO: 268). The unlabeled antibodies are prepped by first doing a buffer exchange into 10mM NaHCO3 (pH 8.5) using a Zeba column. Antibodies are then primed with a 20-fold excess of HaloTag® Succinimidyl Ester (04) Ligand (Promega) and allowed to incubate at room temperature for 2 hours. A buffer exchange is done 2x using Zeba columns to remove free linker. The primed antibodies are incubated with a 4-fold excess of HaloTag-SmTrip9 or HaloTag-SmTrip10 overnight at 4°C while mixing. HaloLink® Resin is used to remove any free HaloTag® fusion proteins.

A 2X stock of recombinant human Troponin (final 1ug/ml) was generated in assay buffer, and 50ul/well added to a non-binding surface treated, 96 well solid-white plate (Costar 3600). A 2X master mix of the purified LgTrip 3546 (SEQ ID NO: 51) (final concentration 1uM) + SmTrip9 pep521 labeled 10-T79F clone (SEQ ID NO: 268) (final 1ug/ml) + SmTrip10 pep289 labeled 10-T79C clone (SEQ ID NO: 150) (final 1ug/ml) was created in assay buffer, and 50ul/well added. A 5X stock of Nano-Glo® Live Cell Substrate in assay buffer, 25ul/well was added to the plate for a final concentration of 10uM, and luminescence was measured in real-time using a GloMax® Discover. Assay buffer consisted of Blocker BSA (10%) (Thermo) diluted in PBS (pH 7.0) to a final of 0.01% BSA in PBS. Samples were tested in triplicate.

Results are depicted in Figure 151.

Example 89

Translocation assay

HiBiT exhibits a very high affinity for the LgBiT polypeptide (K_D = 1 nM) and other similar complementary polypeptides. The strong interaction between the two fragments

would drive complementation without any stimuli (Figure 154), which would be unsuitable for a translocation assay. A study was conducted to determine the optimal affinity between two components (e.g., peptide and polypeptide) of a translocation assays. The optimal affinity was found to be in the range of 280 nM to 1300 nM. A quadruple mutations in LgBiT (E11K/I44M/N135V/L150S), referred to as LgBiT*, reduces its interaction with HiBiT by ~1000 fold (K_D = 1296 nM), renderring the HiBiT/LgBiT* pair well-suited for a translocation assay. Two different translocation assays were designed and tested.

A membrane translocation assay was developed to measure PKCα translocation from cytosol to the plasma membrane under PMA stimulus. PKCα was endogenously tagged with HiBiT at the C-terminus in HeLa cells. The clones of edited cells were isolated, and the best clone with the highest luminescence signal was chosen to perform the assay. LgBiT*-membrane sensor was introduced to the clone using transfection method. Addition of PMA recruits PKCα-HiBiT to the plasma membrane, where HiBiT meets LgBiT* to produce light. Titration of PMA yielded 12- to 19-fold increase in response depending on the amount of LgBiT* transfected (Figure 155).

A nuclear translocation assay was developed using measuring p65 movement from cytosol to the nucleus under TNFα stimulus. The nuclear translocation assay was set up similar to the membrane translocation assay. Specifically, p65 was endogenously tagged at the C-terminus in HeLa cells, and LgBiT*-nuclear sensor was introduced to p65-HiBiT cell line via transfection method. Treatment of TNFα promotes translocation of p65-HiBiT to the nucleus, where complementation occurs between HiBiT and LgBiT* to yield luminescence signal. Titration of TNFα resulted in 4-fold increase in response (Figure 156A). The assay allows measurement of protein translocation in real time. As shown in Figure 156B, it takes approximately 30 minutes for p65 to migrate to the nucleus upon stimulation of TNFα, which is consistent with findings in the literature.

Example 90

Comparison Kd and Bmax values of LgBiT mutants with HiBiT

A solution of HiBiT peptide was prepared starting at 1.22uM in OptiMEM+10% FBS. Serially diluted the peptide dilution 3-fold into OptiMEM+10% FBS. (300ul in 700ul). Diluted purified LgBiT or LgBiT mutant into OptiMEM+10% FBS to a concentration of 2nM. 90ul of the peptide solution was combined with 10ul of the LgBiT dilution (0.2nM LgBiT final). Samples were incubated on an orbital shaker for 30 minutes, and then 11ul of

100uM furimazine in OptiMEM+10% FBS added. Samples were placed on an orbital shaker for 5 minutes, and then luminescence read using a GloMax Multi+ luminometer. Bmax and Kd was calculated with GraphPad Prism using one site specific binding non-linear regression (Figure 157A-B).

Example 91

Affinity of LgBiT mutant lysates for HiBiT

Grew 37°C overnight cultures of LgBiT and each LgBiT mutant. Diluted each culture 1:100 into LB+0.1% Rhamnose and 0.15% glucose. Grew for 20 hours at 25°C. Lysates of each culture were prepared by diluting equal volumes of induced cultures with PLB lysis buffer. (PLB lysis buffer is 0.3X PLB+25mM HEPES pH 7.5). Each lysate was then diluted10,000X into PLB lysis buffer. A dilution series of synthetic HiBiT peptide starting at 300nM was prepared into NanoGlo® Assay buffer+50uM furimazine. 50ul of each diluted lysate was combined with 50ul of the peptide/NanoGlo® titration. Samples were incubated for 3 minutes, and then luminescence read samples on a GloMax® multi+ luminometer (Figure 158).

Example 92

Bioluminescence from complexes of truncated LgTrip3546 with complementary polypeptide

495ul of OptiMEM+10% FBS was aliquoted into deep well plate. Peptides 846 and 847 were diluted to 20uM in 500ul of OptiMEM+10% FBS. Two-fold serial dilutions were prepared for each peptide. Then 200ul of each dilution series was transferred to a new row and then ATG-3929 was added (2nM final) to the 846 titration series and ATG-4794 was added to the 847 titration. As a control 250nM pep 263 was added to 0.2nM of LgTrip 3546. 50ul of each sample was aliquoted in triplicate into a white assay plate and then plate was incubated with shaking for 10 minutes. After incubation 6ul of 10X FZ LCS (167ul of LCS in 833ul of OptiMEM+10% FBS. Samples were mixed for 30 sec on orbital shaker then read immediately, and several times after to get peak luminescence for each condition. The mixtures of 846+ATG-3929 and 847+ATG-4794 peaked at a concentration of 1.25uM peptide. Results are depicted in Figure 159.

Example 93

B9 Titration

LgTrip was diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol, and pep289 was added to 25 uM. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (20x Kd) and titration of VS-HiBiT. Results are depicted in Figure 160.

Example 94

Pep289 Titration

LgTrip was diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol, and pep289 was added to 25 uM. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (20x Kd) and titration of VS-HiBiT. Results are depicted in Figure 161.

Example 95

Dipeptide Affinity Determination

Dipeptides pep263, pep788, and pep900 were diluted to 5 uM and diluted serially 5-fold using TBS + 0.01% BSA + 0.01% Tergitol with 0.2 nM of LgTrip as the diluent. Samples were incubated 10 minutes at room temperature and added to assay plates in triplicate. One-to-one vol:vol of TBS + 0.01% BSA + 0.01% Tergitol with 20x diluted live cell furimazine substrate was added to samples, and plates were read on a GloMax® luminometer after 10 minutes. Results are depicted in Figure 162.

Example 96

Bmax Determination with LgTrip Variants and Pep788

LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 4 hours at 37°C with shaking. Ten microliters of each induction sample was diluted into 250 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5). Eighty microliters of lysates were diluted further in 2 ml lysis buffer. A 10-fold dilution series of pep788 (SEQ ID 414) was performed stating at 10 μM peptide using Nano-Glo with 50 uM furimazine as the

diluent. Peptide dilutions and lysates were mixed 1:1 vol:vol, incubated 10 min at room temperature, and luminescence was read. Results are depicted in Figure 163.

Example 97

Bmax Determination with LgTrip Variants and Pep759

LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 4 hours at 37°C with shaking. Ten microliters of each induction sample was diluted into 250 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5). Eighty microliters of lysates were diluted further in 2 ml lysis buffer. A 5-fold dilution series of pep759 (SEQ ID 496) was performed stating at 50 µM peptide using Nano-Glo with 50 uM furimazine and 50 µM pep289 (SEQ ID 826) as the diluent. Peptide dilutions and lysates were mixed 1:1 vol:vol, incubated 10 min at room temperature, and luminescence was read. Results are depicted in Figure 164.

Example 98

Thermal Stability of LgTrip Variants

LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 20 hours at 25°C with shaking. Twenty microliters of each induction was diluted into 40 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5) and lysed for 15 min at room temperature. Lysates were diluted 1,000-fold into 1X TBS + 0.01% BSA. Fifty microliters of each sample was transferred into a PCR plate and incubated at 80°C for 1.5 hrs in a thermocycler. Controls were incubated on ice. Samples were equilibrated to room temperature and diluted 1:100 into 1X TBS + 0.01% BSA. Twenty-five microliters of each sample were transferred into assay plates and mixed with 25 ul of 400nM pep788 (SEQ ID 414) in TBS + 0.01% BSA + 20x diluted live cell furimazine substrate. Samples were incubated 10 minutes at room temperature, and luminescence was read. Results are depicted in Figure 165.

Example 99

Kd and Bmax Determinations of LgTrip Variants with Pep788

LgTrip variants were purified using the Promega MagneHis™ Protein Purification System according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep788. This solution was used as the diluent for 5-fold serial

dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT. Results are depicted in Figure 166.

Example 100

Kd and Bmax Determinations of LgTrip Variants with Pep840

LgTrips variants were purified using the Promega MagneHis™ Protein Purification System according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep840. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT. Results are depicted in Figure 167.

Example 101

Kd and Bmax Determinations of LgTrip Variants with Pep289 and Saturating Pep840

LgTrip variants were purified using the Promega MagneHis™ Protein Purification System according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289 + 12.5 uM Pep840. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT. Results are depicted in Figure 168.

Example 102

Half-Life Determination of LgTrip Variants

MagneHis purified LgTrip variants were diluted to 20 nM in TBS+0.01% BSA. In triplicate, 100µl aliquots of each sample were loaded into 200µl thin walled PCR tubes. Samples were incubated at 70°C in thermal cycler. Samples were removed at various time-points and equilibrated to room temperature. Samples were diluted to 0.2 nM (5 in 495µl) in TBS+0.01% BSA. 25 µl of each diluted sample was combined with 25 µl of TBS + 0.01% BSA +20x diluted live cell substrate furimazine + 400 nM pep788 (SEQ ID NO: 414). Samples were incubated for 10 minutes and then read on GMM+. Half-life was calculated by non-linear regression. Results are depicted in Figure 169.

Example 103

Rapamycin Assay with LgTrip Variants

Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10 were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. PLB assay reagent was prepared with 444 nM of a MagneHis purified LgTrip variant, 90x diluted FRB-SmTrip10 culture, +/- 35 nM Rapamycin. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB, and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of Nano-Glo containing 50 uM furimazine was added to assay plates wells, and luminescence was read on GloMax® luminometer after 5 minutes. Results are depicted in Figures 170-171.

Example 104

Thermal Stability of the Lysates of LgTrip Variants

LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 20 hours at 25°C with shaking. Twenty microliters of each induction was diluted into 40 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5) and lysed for 15 min at room temperature. Lysates were diluted 1,000-fold into 1X TBS + 0.01% BSA. Fifty microliters of each sample was transferred into a PCR plate and incubated at 70°C for 1.5 hrs in a thermocycler. Controls were incubated on ice. Samples were equilibrated to room temperature and diluted 1:100 into 1X TBS + 0.01% BSA. Twenty-five microliters of each sample were transferred into assay plates and mixed with 25 ul of 400nM pep788 (SEQ ID

414) in TBS + 0.01% BSA + 20x diluted live cell substrate. Samples were incubated 10 minutes at room temperature, and luminescence was read. Results are depicted in Figure 172.

Example 105

Thermal Stability of the Lysates of LgTrip Variants at Different Temperature Gradients

LgTrip variants were grown overnight at 37°C in LB with 100 ug/ml ampicillin. Cells were diluted 20-fold into induction media (LB with 100 ug/ml ampicillin and 0.1% rhamnose w/v) and induced 20 hours at 25°C with shaking. Twenty microliters of each induction was diluted into 40 ul lysis buffer (0.3X PLB + 25mM HEPES pH 7.5) and lysed for 15 min at room temperature. Lysates were diluted 1,000-fold into 1X TBS + 0.01% BSA. Fifty microliters of each sample was transferred into a PCR plate and incubated at two temperature gradients, either 75-100°C for 10 min or 50-75°C for 1.5 hr, in a Veritas thermocycler. Controls were incubated on ice. Samples were equilibrated to room temperature and diluted 1:100 into 1X TBS + 0.01% BSA. Twenty-five microliters of each sample were transferred into assay plates and mixed with 25 ul of 400nM pep788 (SEQ ID 414) in TBS + 0.01% BSA + 20x diluted live cell furimazine substrate. Samples were incubated 10 minutes at room temperature, and luminescence was read. Results are depicted in Figure 173.

Example 106

Thermal Stability of Purified LgTrip Variants

MagneHis purified LgTrip variants were diluted to 20 nM in TBS+0.01% BSA. In triplicate, 100µl aliquots of each sample were loaded into 200µl thin walled PCR tubes. Samples were incubated at 70°C in thermal cycler. Samples were removed at various time-points and equilibrated to room temperature. Samples were diluted to 0.2 nM (5 in 495µl) in TBS+0.01% BSA. 25 µl of each diluted sample was combined with 25 µl of TBS + 0.01% BSA +20x diluted live cell substrate furimazine + 400 nM pep788 (SEQ ID NO: 414). Samples were incubated for 10 minutes and then read on GMM+. Results are depicted in Figure 174.

Example 107

Kd and Bmax Determinations of LgTrip Variants with Pep521 and Saturating VS-HiBiT

LgTrip variants were purified using the Promega MagneHis™ Protein Purification System according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA

+ 0.01% Tergitol + 25 uM pep521. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. Results are depicted in Figure 175.

Example 108

Kd and Bmax Determinations of LgTrip Variants with Pep840 and Saturating VS-HiBiT

LgTrip variants were purified using the Promega MagneHis™ Protein Purification System according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep840. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. Results are depicted in Figure 176.

Example 109

Kd and Bmax Determinations of SmTrip9 Variants with LgTrip Variant ATG-3546 or ATG-5146 and Saturating Pep289

LgTrip ATG-3546 and ATG-5146 were purified using the Promega MagneHis™ Protein Purification System according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min and luminescence was read. Results are depicted in Figures 177 and 178.

Example 109

Kd and Bmax Determinations of Pep289 with LgTrip Variant ATG-3546 or ATG-5146 and Saturating SmTrip9

LgTrip ATG-3546 and ATG-5146 were purified using the Promega MagneHis™ Protein Purification System according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at

room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min and luminescence was read. To determine VS-HiBiT Kd, the same protocol was followed, but with saturating SmTrip9 (25 uM) and titration of VS-HiBiT. Results are depicted in Figures 179 and 180.

Example 110

Kd and Bmax Determinations of Lysates of LgTrip Variants with Pep840 Titration

LgTrip cultures were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin and induced ~20 hr at 25°C with shaking. Cells were diluted 1000x in PLB assay reagent (0.3X PLB + 25mM HEPES pH 7.5) and lysed for 20 min. Five-fold serial dilutions of SmTrip9 pep840 were performed in NanoGlo + 50 uM furimazine + 25 uM pep289 and mixed 1:1 vol:vol with LgTrip lysates. Samples were incubated 10 min at room temperature and read on GloMax® luminometer. Results are depicted in Figures 181.

Example 111

Kd and Bmax Determinations of Purified LgTrip Variants with Pep840 Titration

LgTrip variants were purified using the Promega MagneHis™ Protein Purification System according to the manufacturer's protocol and diluted to 0.2 nM in TBS + 0.01% BSA + 0.01% Tergitol + 25 uM pep289. This solution was used as the diluent for 5-fold serial dilution series of SmTrip9 peptides. Samples were equilibrated 10 min at room temperature and aliquoted into assay plates in triplicate. TBS + 0.01% BSA + 0.01% Tergitol containing 20 uM furimazine was added to samples in 1:1 vol:vol ratio. Plates were incubated 10 min, and luminescence was read. Results are depicted in Figures 182.

Example 112

Half-Life Determination of Purified LgTrip Variants

MagneHis purified LgTrip variants were diluted to 20 nM in TBS+0.01% BSA. In triplicate, 100µl aliquots of each sample were loaded into 200µl thin walled PCR tubes. Samples were incubated at 70°C in thermal cycler. Samples were removed at various time-points and equilibrated to room temperature. Samples were diluted to 0.2 nM (5 in 495µl) in TBS+0.01% BSA. 25 µl of each diluted sample was combined with 25 µl of TBS + 0.01%

BSA +20x diluted live cell substrate furimazine + 200 nM pep900 (SEQ ID NO: 907). Samples were incubated for 10 minutes and then read on GMM+. Half-life was calculated by non-linear regression. Results are depicted in Figures 183.

Example 113

Rapamycin Assay with LgTrip Variants ATG-3546 and ATG-5146

Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10 were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. PLB assay reagent was prepared with 444 nM of a MagneHis purified LgTrip variant, 90x diluted FRB-SmTrip10 culture, +/- 35 nM Rapamycin. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB, and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of Nano-Glo containing 50 uM furimazine was added to assay plates wells, and luminescence was read on GloMax® luminometer after 5 minutes. Results are depicted in Figures 184.

Example 114

Bmax Determinations of Lystates of LegTrip Variants with Pep788, Pep900, or Pep840

LgTrip variant cultures were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin and induced ~20 hr at 25°C with shaking. Cells were diluted 5000x in 0.3x PLB assay reagent and lysed for 20 min. Five-fold serial dilutions of dipeptides pep788, pep900, or SmTrip9 pep840 with saturating pep289 were performed in Nano-Glo + 50 uM furimazine and mixed 1:1 vol:vol with LgTrip lysates. Samples were incubated 10 min at room temperature and read on GloMax® luminometer. Bmax was calculated by non-linear regression. Results are depicted in Figures 185.

Example 115

Thermal Stability of LgTrip Variants at Different Temperature Gradients

LgTrip variants were diluted to 20 nM in 2 ml TBS+0.01% BSA. 100ul of each sample was aliquoted into duplicate rows of 96-well PCR plates (make two plates). Plates were incubated for 3 hrs at high (75-100°C) or low (50-75°C) temperature gradients in a Veritas thermocycler. Samples were placed at 70°C and then aliquots moved to rt at various

timepoints. Samples were mixed at each timepoint with a pipette and then diluted 1:100 into TBS + 0.01% BSA. (5ul into 495ul). 25ul of each sample was then aliquoted into a white assay plate. 25ul of 200nM pep788 or pep900 in TBS + 0.01% BSA + 20x diluted live cell furimaizine substrate was added. The plate was incubated for 10 minutes and then read on GMM+. Results are depicted in Figures 186.

Example 116

Rapamycin Assay with LgTrip Variants

Cultures of FKBP_SmTrip9 variants and FRB-SmTrip10 were grown overnight in LB + 100 ug/ml ampicillin at 37°C. Cells were diluted 1:20 in LB with 0.15% glucose, 0.1% rhamnose, and 100ug/ml ampicillin. Cultures were induced ~20 hr at 25°C with shaking. PLB assay reagent was prepared with 444 nM of a MagneHis purified LgTrip variant, 90x diluted FRB-SmTrip10 culture, +/- 35 nM Rapamycin. Ninety microliters of assay reagent was added to each well of 96-well assay plates. FKBP_SmTrip9 cultures were diluted 1:10 in PLB, and 10 ul was added to assay plates. Samples were incubated 30 min at room temperature. One hundred microliters of Nano-Glo containing 50 uM furimazine was added to assay plates wells, and luminescence was read on GloMax® luminometer after 5 minutes. Results are depicted in Figures 187.

Example 117

Kd and Bmax Determinations of Pep691 and Pep692

LgTrip 3546 was diluted to 1nM in OptiMEM+10% FBS. 12µM solutions of strand 9 peptides 521 and 693 were prepared in OptiMem+10% FBS. Each strand 9 dilution was used to prepare 3-fold dilution series of each strand 10 peptide starting at 20µM. (pep86=HiBiT, pep289=VS HiBiT, pep691=HiBiT RR, pep692=VSHiBiT RR). 90µl of each dilution series was transferred to a white assay plate and then 10µl of the 1µM stock of LgTrip 3546 added. The plate was placed on an orbital shaker set to 600RPM for 30 minutes. Detection reagent of OptiMEM+10% FBS consisting of 10mM DTT and 50uM Furimazine was prepared, and 11µl added to the samples. The plate was placed on orbital shaker and mixed for 5 minutes at room temperature. The plate was read on a GloMax® Multi+ luminometer. Kd and Bmax were calculated using GraphPad Prism one site specific binding. Results are depicted in Figures 188.

Example 118

Purification of Monomeric LgBiT-SmBiT Clones

50ml cultures of an isolated colony of each monomeric LgBiT-SmBiT clone were grown in LB+30ug/ml Kanamycin for 20 hours at 37°. Cultures were then diluted 1:100 (500ul in 50ml) in LB supplmented with 30ug/ml Kanamycin, 0.1% Rhamnose, and 0.15% Glucose and grown at 25°C for 20 hours. Cultures were spun and re-suspended in 9ml of 100mMHepes pH 7.5+1ml FastBreak™ Cell Lysis Reagent (Promega; V8571)+200ul RQ DNase 1 (Promega). Samples were incubated on a orbital mixer for 30 min. at 4°C. A aliquot was saved for "Total Lysate" sample. Samples were spun to clear lysate (7000RPM for 15 minutes), and supernatant was transferred to a new tube. Using the HisLink™ Spin Protein Purification System (Promega; V8550), 1ml of HisLink™ Protein Purification Resin was added to each cleared lysated, incubated for 10 minutes at 4°C on an orbital mixer, washed 3x with HisLink wash/binding buffer, and eluted with 500ul elution buffer two times to recover samples.

Figure 189 demonstrates the amount of soluble and purified protein from each monomeric LgBiT-SmBiT clone, and Table 12 lists the constructs used.

Table 12. Constructs

ATG-462	NanoLuc	G
ATG-3561	Monomeric LgBiT-SmBiT	S
ATG-3562	Monomeric LgBiT-SmBiT	G
ATG-3563	Monomeric LgBiT-HiBiT	S
ATG-3564	Monomeric LgBiT-HiBiT	G

Example 119

Luminescence Determination of Monomeric LgBiT-SmBiT Clones

Purified monomeric LgBiT-SmBiT proteins from Example 118 were diluted to 0.2nM in TBS+0.01% BSA. Diluted protein was then combined with 50ul of Furimazine (Promega; N113) in NanoGlo® buffer (Promega; N112). Luminescence was read on a GMM+ 3 minutes after substrate addition.

Figure 190 demonstrates that NanoLuc (ATG-462) is 2x brighter than the monomeric LgBiT-SmBiT proteins.

Example 120

Substrate Utilization by Monomeric LgBiT-SmBiT Clones

Purified monomeric LgBiT-SmBiT proteins from Example 118 were diluted to 0.2nM in TBS+0.01% BSA. A 2-fold dilution series, starting at 50uM (40ul in 2ml) then 1ml to 1ml, of Furimazine (Promega; N113) in NanoGlo® buffer (Promega; N112) was prepared. 50ul of each purified monomeric LgBiT-SmBiT was combined, in duplicate, with 50ul of the tirated substrate series. Samples were incubated for 3 minutes at RT, and luminescence read on a GMM.

Figure 191 demonstrates that each monomeric LgBiT-SmBiT protein utilizes Furimazine similarly.

Example 121

Temperature Gradient

Experiments were conducted during development of embodiments herein to determine the impact of temperature on monomeric LgBiT-SmBiT variants.

Purified monomeric LgBiT-SmBiT proteins from Example 118 were diluted to 2nM in TBS+0.01% BSA. Each diluted sample was then aliquoted into multiple wells of a 96-well PCR plate. The plate was placed in a thermocycler set for 30 minutes with a temperature gradient:

Temperature gradient A: 54, 57, 60, 63, 66, 70°C,

Temperature gradient B: 55, 60, 65, 70, 75, 80°C, or

Temperature gradient C: 65, 70, 75, 80, 85, 90°C.

After the 30 minute incubation, 5ul of each sample was combined wth 45ul of TBS+0.01% BSA, 50ul of Furimazine in NanoGlo® buffer added, incubated for 3 minutes at RT, and luminescence detected on a GMM+.

Figure 192 demonstrates that the monomeric LgBiT-SmBiT protein are significantly more stable compared to NanoLuc (ATG-462).

Example 122

Temperature Challenge

Experiments were conducted during development of embodiments herein to determine the impact of high temperature on monomeric LgBiT-SmBiT variants.

Purified monomeric LgBiT-SmBiT proteins from Example 118 were diluted to 2nM in TBS+0.01% BSA. Each diluted sample was then aliquoted into multiple wells of a 96-well PCR plate. The plate was placed in a thermocycler set for 30 minutes with a temperature gradient of 75, 80, 85, 90, 95, 100°C.

After the 30 minute incubation, 5ul of each sample was combined wth 45ul of TBS+0.01% BSA, 50ul of Furimazine in NanoGlo® buffer added, incubated for 3 minutes at RT, and luminescence detected on a GMM+.

Figure 193 demonstrates that the 159G amino acid change (ATG-3563) provides enhanced thermostability over the 159S amino acid change (ATG-3564) in the monomeric LgBiT-SmBiT protein are significantly more stable compared to NanoLuc (ATG-462).

Example 123

Stability of Monomeric LgBiT-SmBiT at 60°C

Experiments were conducted during development of embodiments herein to determine the stability of monomeric LgBiT-SmBiT variants at 60°C.

Purified monomeric LgBiT-SmBiT proteins from Example 118 were diluted to 2nM in TBS+0.01% BSA. Each diluted sample was then aliquoted into multiple wells of a 96-well PCR plate, and the placed placed in a thermocycler set at 60°C. At various timepoints, aliquots were removed and kept on ice. After all sample timepoints had been collected, samples were equilibrated to RT.

Once equilibrated, 5ul of each sample was combined wth 45ul of TBS+0.01% BSA, 50ul of Furimazine in NanoGlo® buffer added, incubated for 3 minutes at RT, and luminescence detected on a GMM+.

Figure 194 demonstrates consistent results with the temperature gradient assay. The monomeric LgBiT-SmBiT variants are more stable than NanoLuc.

Example 124

Stability of Monomeric LgBiT-SmBiT with Reagent at Elevated Temperature

Experiments were conducted during development of embodiments herein to determine the stability of monomeric LgBiT-SmBiT variants with a reagent at elevated temperature.

Purified monomeric LgBiT-SmBiT proteins from Example 118 were diluted to 200nM in TBS+0.01% BSA and then further diluted to 0.2nM (4ul in 4ml). 50ul of ATG-462 or ATG-3564 was mixed with either 50ul of 50uM Furimazine in NanoGlo™ buffer or 50ul of 20uM Furimazine in TBS+0.01% BSA and placed into wells of a thin-walled 96-well PCR plate. The tray was placed in a Veritas thermocycler set at a temperature gradient of 55, 60, 65, 70, 75, 80°C. At various timepoints (FIG. 195) or after 30 (FIG. 196) minutes, aliquots were removed, and luminescence detected on a GMM+.

Example 125

Kd and Vmax Determinations of NanoLuc Variants

Purified NanoLuc variants were diluted to 0.2nM in TBS+0.01% BSA. A 2-fold dilution series of Furimazine in NanoLuc® buffer starting at 50uM (40ul in 2ml), and then 1ml to 1ml dilutions. 50ul of sample, in duplicate, was mixed with the 50ul of the titration series. Samples were inclubated for 3 minutes at RT, and luminescence detected on a GMM+ (FIG. 197).

Example 126

Temperature Challenge of NanoLuc Variants

Experiments were conducted during development of embodiments herein to determine the impact of high temperature on NanoLuc variants.

Purified NanoLuc variant proteins were diluted to 2nM in TBS+0.01% BSA. Each diluted sample was then aliquoted into multiple wells of two 96-well PCR plates. The plates were placed in a thermocycler set for 30 minutes with a temperature gradient of 60, 65, 70, 75, 80, 85°C.

After the 30 minute incubation, 5ul of each sample was combined wth 45ul of TBS+0.01% BSA, 50ul of Furimazine in NanoGlo® buffer added, incubated for 3 minutes at RT, and luminescence detected on a GMM+ (FIG. 198).

The temperature challenge was repeated at a higher temperature gradient 70, 75, 80, 85, 90, 95°C with the most stable clones identified in the above temperature challenges (FIG. 199).

Example 127

Variants on ATG-5333 and ATG-5344 Screen in the Context of a Fusion for Cell-Based Translocation Assay

a) Luminescence Detection

Cultures of each variant were started from a single colony picked into 200ul LB+ampicillin media into wells of 96 well plates and grown at 37°C for 20 hours with shaking. The following day induction cultures were prepared by diluting 10ul of culture into 200ul of LB+100ug/ml ampicillin+0.1% Rhamnose and grown for 3 hours at 37°C. A lysate was prepared from the induced cultures as follows: 10ul of cells was transferred to 190ul of Passive Lysis Buffer (PLB)(0.3X PLB+25mM HEPES pH 7.5) and incubated for 5 minutes. 50ul of lysate was transferred into two assay plates and then 50ul of PLB assay buffer+20uM Furimazine with either 6uM or 0.2uM (3uM or 0.1uM final) of Pep289 was added. Assay plates were incubated for 5 minutes, and then luminescence was measured. A ratio was calculated by dividing RLU from the 3uM samples by the RLU values of the 0.1uM samples.

Table 13. Variants on ATG-5333

ATG#	3uM	0.1uM	ratio
5338	1.1	0.30	3.7
5340	1.3	0.50	2.6
5407	1.9	2.70	0.7
5408	0.3	0.06	5.0
5411	1.4	1.40	1.0
5413	3.6	2.98	1.2
5414	1.09	0.54	2.0
5416	1.34	0.28	4.8
5417	1.62	0.80	2.0
5418	1.61	0.77	2.1
5419	1.47	0.36	4.1
5420	0.63	0.10	6.3
5421	1.97	1.40	1.4

Table 14. Variants on ATG-5344

ATG#	3uM	0.1uM	ratio
5338	1.1	0.30	3.7
5340	1.3	0.50	2.6
5407	1.9	2.70	0.7
5408	0.3	0.06	5.0
5411	1.4	1.40	1.0
5413	3.6	2.98	1.2
5414	1.09	0.54	2.0
5416	1.34	0.28	4.8
5417	1.62	0.80	2.0
5418	1.61	0.77	2.1
5419	1.47	0.36	4.1
5420	0.63	0.10	6.3
5421	1.97	1.40	1.4

b) Kd and Bmax Calculation with Pep289

2ml of induction media (LB+100ug/ml AMP+0.1% Rhamnose) was innoculated with 100ul of an overnight culture of each variant. Cells were grown for 3 hours at 37°C. 250ul of the cells were diluted in 5ml of PLB and incubated for ~10 minutes at RT. 20ml of 20uM Fz in PLB (80ul) was prepared, and three 3X titration series of Pep289 prepared in the Fz reagent (50uM in 2ml) (10ul of 5mM sample in 990ul, then 300ul in 700ul, and then the 3-dilution series combined. 50ul of cell lysate was combined with 50ul of Pep289 titration, incubated for 5 minutes, and then luminescence read on GMM+.

Table 15. Calculated Bmax and Kd for LgBiT mutants

ATG#	Sequence	Bmax	Kd
5333	LgBiT'	1	1
5339	K11N	0.58	13.0
5340	R152Q	0.9	2.3
5341	V135A	0.3	2.3
5408	N156D	0.4	20.2
5409	H57Q	0.88	2.2
5432	L3H	0.2	44.0
5433	T13S	0.42	3.8
5434	P93H	0.75	1.0
5435	F120L	0.30	2.5
5437	S157R	1.0	0.6
5438	H86L	0.82	1.1
5439	M149V	0.07	51.3
5456	K11L	0.3	0.6
5457	K11R	0.67	1.0
5458	K11Y	0.4	1.2
5459	K11N+R152Q	0.4	16.9
5460	K11N+N156D	0.3	25.9
5491	K11Q	0.9	0.6
5492	K11M	0.3	0.6
5493	K11H	0.7	2.8
5494	K11F	0.3	2.2
5495	K11W	0.2	2.7
5505	V135A+R152Q	2.0	1.2
5506	V135A+R152Q+N156D	1.5	4.1
5507	P93H+V135A	1.9	0.5
5508	P93H+R152Q	0.8	3.0
5509	P93H+R152Q+N156D	0.7	6.8
5510	P93H+N135A+R152Q	1.7	1.5
5511	P93H+N135G+R152Q	0.9	3.9

Example 128

Site saturation of ATG-5534 template

Experiments were conducted during development of embodiments herein to optimize the identity of the amino acid at various positions of ATG-5534 (SEQ ID NO: 978).

a) E. coli cultures were started from a single colony picked into 200ul LB+ampicillin media into wells of 96-well plates and grown at 37°C for 20 hours with shaking. The following day, induction cultures were prepared by diluting 10ul of culture into 200ul of LB+100ug/ml ampicillin+0.1% Rhamnose and grown for 3 hours at 37°C. A lysate was prepared from the induced cultures as follows: 10ul of cells was transferred to 190ul of PLB lysis buffer (0.3X PLB+25mM HEPES pH 7.5) and incubated for 5 minutes. 50ul of lysate was transferred into two assay plates and then 50ul of PLB assay buffer+20uM Furimazine with either 6uM or 0.2uM (3uM or 0.1uM final) of pep289 was added. Assay plates were incubated for 5 minutes, and then luminescence was measured. Bmax and Kd values were determined for each mutant.

Table 16. Calculated Bmax and Kd for site saturation at position 44 of LgBiT mutant ATG-5810

ATG#	Mutant	Bmax	Kd
5654	M44V	**1.8**	**0.4**
5655	M44I	**4.3**	**0.13**
5676	M44K	0.03	19.4
5677	M44E	0.02	15.7
5678	M44A	0.4	2.1
5679	M44C	0.4	1.1
5680	M44W	0.4	0.1
5681	M44G	0.1	11.9
5682	M44H	0.4	2.5
5683	M44S	0.2	2.9
5684	M44Q	0.4	0.8
5685	M44R	0.001	0.0
5686	M44T	0.4	1.3
5687	M44Y	0.2	1.1
5688	M44L	0.9	0.8
5689	M44P	0.0	3.4
5690	M44F	0.8	1.0

b) Comparison in E.coli and mammalian cells

i) E. Coli (Kd and Bmax)

2ml of induction media (LB+100ug/ml AMP+0.1% Rhamnose) was inoculated with 100ul of an overnight culture for each mutant. Cells were grown for 3 hours at 37°C in a tube. 250ul of cells were diluted in 5ml of PLB lysis buffer and incubated for ~10 minutes at RT. 20ml of 20uM Furimazine in PLB lysis buffer (80ul) was prepared, and three 3X titration series of pep289 in the Furimazine reagent (50uM in 2ml) (10ul of 5mM sample in 990ul, then 300ul in 700ul with the 3-dilution series combined). 50ul of cell lysate was mixed with 50ul of pep289 titration, incubated for 5 minutes at RT, and then luminescence read on GMM+.

ii) Mammalian Cell Expression

a) Transfection protocol: Media from HeLa cells (PKCα-HiBiT clone) that were grown to confluency in a T-150 flask was aspriated, and cells washed with 10ml DPBS. (Life Technologies 14190). The DPBS was aspirated, and 4ml of TryPLE Express Trypsin (Life Technologies 12604) added. Cells were incubated for 2-3 minutes at 37°C, then resuspend in 16ml of growth media (DMEM Life Technologies 11995)+10% FBS (VWR 89510-194).

Cells were spun at 200RPM for 5 minutes, supernatant aspirated, and 20ml of DMEM+10% FBS added. Cells were counted and then diluted to 1,000,000 cells per ml. 1mL of cells with 3mL of DMEM+10%FBS were plated into 6 cm dishes and incubated for 24h.

b) Transfection complex: For each DNA sample to be tested, 10ug of DNA (1ug of DNA encoding membrane sensor, and 9 ug of carrier DNA) (Promega E4882) combined and mixed with 400ul of OptiMem (Life Technologies 11058), 30µl of FugeneHD transfection reagent (Promega E2311) added, and incubated for 10 min at ambient temperature. Transfection complex was added to the plated cells and incubated for 24 h.

c) Cell-based luminescence assay: Media was aspirated from the transfected cells HeLa cells, and then the cells were washed with 5ml DPBS. (Life Technologies 14190). The DPBS was aspirated, and 0.75ml of TryPLE Express Trypsin (Life Technologies 12604) added. Cells were incubated for 2-3 minutes at 37°C, and then resuspended in 4ml of growth media (DMEM Life Technologies 11995)+10% FBS (VWR 89510-194). 100µl of each sample was added to each well of a white 96 well assay plate (Corning 3917). For each sample, 30 wells in total were used, 10 wells per row with 3 rows. Cells were incubated for another 16-24h.

The following morning, growth media was aspirated and replaced with 90µl of CO_2 independent media (Life Technologies 18045)+10% FBS containing 1.1x NanoGlo® Live Cell Substrate (Promega N2012), incubated for 10 min, and 10µl of titrated PMA compound added to each well. (See, e.g., FIGS. 225-227). Luminescence was detected on a GlomaxMulti+ luminometer set to 37°C with a kinetic run over the course of 1 hour. Unless otherwise stated, the reported luminescence is the peak height of the kinetic run.

Table 17. RLU, S/B, Bmax, and Kd Values Normalized to ATG-5534

		Membrane sensor		*E. Coli*	
Sample	Mutation	RLU	S/B	Bmax	Kd
ATG-5534	WT	1	4.5	1	1
ATG-5652	Y16F+Q20P+Q152H	0.3	5.2	1.7	2.3
ATG-5653	M106R+Y114F	0.6	3.8	2	1
ATG-5654	M44V	4.4	2.4	1.8	0.4
ATG-5655	M44I	5.2	2.1	4.3	0.13
ATG-5656	M106R	1	3.2	3.5	0.76
ATG-5657	M106T	0.7	3.6	2.2	0.77
ATG-5658	M106R	0.8	4.1	3.2	0.61
ATG-5659	M106K+K136E	0.7	4.4	3.1	1.58
ATG-5660	V36G	0.8	3.3	3.7	1.1
ATG-5661	M106L+I138K	0.6	4.2	3.7	0.68
ATG-5688	M44L	1.72	3.1	0.89	0.76
ATG-5690	M44F	0.24	3.5	0.79	0.98

Example 129

Strand 9 detector polypeptides

A polypetide construct was prepared by placing a strand 10 sequence at the N-terminus of a LgTrip polypeptide. This configuration produces increased luminescence in the presence of strand 9 sequences. Experiments were conducted during development of embodiments herein to determine the limits of detection for the strand 9 peptides being tested (pep521 and pep840).

a) Strand 9 Detector Proteins with pep521 and pep840

Each strand 9 detector protein was diluted to 20nM in TBS+0.01% BSA. A 3X dilution series of pep521 and pep840 was diluted in TBS+0.01% BSA starting at 20uM. 50ul of each enzyme dilution in duplicate was combined with 50ul of each peptide titration and incubated for 10 minutes on a shaker for pre-equilibration. An assay buffer was prepared by diluting Nano-Glo Live Cell Substrate (Furimazine; Promega; N205) 30-fold into TBS+0.01% BSA, 100ul added to each well, incubated for 5 minutes at RT, and luminescence then read on GMM+. The background reading was obtained from samples that contained no peptide.

FIG. 200 shows that all three strand 9 detector proteins bound more tightly to pep840 than to pep521.

b) Strand 9 Detector Proteins with pep840

Each strand 9 detector protein was diluted to 200nM in TBS+0.01% BSA. A 3X dilution series of pep840 was diluted in TBS+0.01% BSA starting at 0.5uM. 50ul of each enzyme dilution in quadruplicate was combined with 50ul of each peptide titration and incubated for 10 minutes on a shaker for pre-equilibration. An assay buffer was prepared by diluting Nano-Glo Live Cell Substrate (Furimazine; Promega; N205) 30-fold into TBS+0.01% BSA, 100ul added to each well, incubated for 5 minutes at RT, and luminescence then read on GMM+. The background reading was obtained from samples that contained no peptide.

Regarding FIG. 20, the panel on the left shows a titration of Strand 9 peptide 840 in the presence of 50nM of three strand 9 detector constructs. The panel on the right shows the background (no peptide 840) for each strand 9 detector and the signal to background ratio at 0.1nM peptide 840 for each strand 9 detector.

Example 130

Linker Test Series

a) Overnight cultures of circularly-permuted, strand 9 detector variants were diluted to 1:20 (150ul to 3ml) in LB+0.1% rhamnose+100ug/ml ampicilin, grown for 4 hours at 37°C, and then lysed in PLB lysis buffer (0.3X PLB+25mM HEPES pH 7.5) (500ul of lysate to 4.5ml of PLB lysis buffer). To assay, lysates were diluted 1:100 into TBS+0.01% BSA. A 3x dilution series of pep840 (strand 9) was prepared starting at 10nM. 50ul of each dilute lysate was combined with 50ul of peptide titration and incubated for 20 minutes at RT on a shaker set at 600rpm. 100ul Nano-Glo® buffer + 50uM Furimazine was added to each well, and luminescence then read on GMM+. The linkers used for each construct tested were: ATG-4992, 8GS; ATG-5485, 5GS; ATG-5486, 6GS; ATG-5487, 7GS; ATG-5488, 9GS; ATG-5489, 10GS; and ATG-5490, 11GS.

FIG. 202 shows that each clone, other than ATG-5485, which has a 5AA linker, produced similar luminescence to ATG-4992. The experiments demonstrate that linker length between strand 10, and LgTrip did not play a significant role in the detection of strand 9 sequences.

b) 8GS Linker (ATG-4992) vs. 11GS (ATG-5490) Linker

ATG-4992 and ATG-5490 proteins were purified using the MagneHis purification system. (Promega). Purified protein for 4992 and 5490 were diluted to 100nM in CO_2 independent media+10% FBS. A 3-fold serial dilution of pep840 was prepared starting at 10nM in Nano-Glo® buffer + 50uM furimazine(Promega N113). 50ul of each enzyme dilution was combined in quadruplicate with 50ul the peptide titration. Luminescence was measured over time on a GMM+ luminometer.

FIG. 203 show the data plotted is from the 50-minute kinetic read and that 5490, which has the longer linker, provided 2-fold more luminescence over 4992.

Example 131

Determination of Kd of ATG-4992 and ATG-5490 variants

Overnight cultures of each variant were prepared in LB+100ug/ml ampicillin. The following day, cultures were diluted 1:20 (150ul to 3ml) in LB+0.1% Rhamnose+100ug/ml ampicillin. Cultures were grown for 4 hours at 37°C and then lysed with PLB lysis buffer (0.3X PLB (Promega)+25mM HEPES pH 7.5)(500ul of lysate to 4.5ml of PLB lysis buffer). To assay, lysates were diluted 1:100 into TBS+0.01% BSA. A 3X dilution series of strand 9 (pep840) was prepared starting at 40uM. 50ul of the titration series was combined with 50ul

of the each diluted lysate, incubated for 20 minutes at RT on orbital shaker set to 600rpm, 100ul of Nano-Glo® buffer + 50uM Furimazine (Promega; N113) added, and luminescence read on GMM+. Results are depicted in FIG. 204.

Example 132

Comparison of ATG-4992 and ATG-5490 variants

ATG-4992 and ATG-5490 variant proteins were purified using the MagneHis purification system. (Promega). Purified proteins were then diluted to 100nM in CO_2 independent media+10% FBS. A 3-fold serial dilution of pep840 was prepared starting at 2nM in Nano-Glo® buffer + 50uM furimazine(Promega N113). 50ul of each enzyme dilution was combined in triplicate with 50ul the peptide titration. Luminescence was measured every 5 minutes on a GMM+ luminometer. Results are depicted in FIG. 205 as data plotted from the 60-minute timepoint and show that variants 5515 and 5517 have lower calculated Kd values (Example 131) and show higher RLU signal to background values when paired with pep840.

Example 133

Screen of ATG-4166 Variants

Table 18. LgBiT sequences in different backbones. (C-terminal His Tag for protein purification or sensor fusion (pH domain-GSSG-HaloTag-GSSG-LgBiT mutant).

pF1A (His tag for purification)	Equivalent sequence (sensor backbone)	Sequence compared to LgBiT
ATG-2623	LgBiT	LgBiT
ATG-4166	ATG-5333	LgBiT+E11K+I44M+N135V+L150S
ATG-5810	ATG-5534	LgBiT+E11K+I44M+L150S+P93H+V135A+R152Q
ATG-5819	ATG-5655	LgBiT+E11K+L150S+P93H+N135A+R152Q
ATG-5820	ATG-5654	LgBiT+E11K+I44V+L150S+P93H+N135A+R152Q
ATG-5821	ATG-5690	LgBiT+E11K+I44F+L150S+P93H+N135A+R152Q
ATG-5822	ATG-5688	LgBiT+E11K+I44L+L150S+P93H+N135A+R152Q

i) Bmax and Kd Determination

ATG-4166 variant proteins were purified using the MagneHis purification system. (Promega). Purified proteins were then diluted to 0.2nM in TBS+0.01% BSA. A 3-fold serial dilution of pep289 (VS-HiBiT) was prepared in TBS+0.01% Tergitol with one series starting at 400nM and the other at 20uM. 50ul of each enzyme dilution (0.2nM) was combined with 50ul the peptide titration and incubated on a shaker for 10 minutes. 100ul of Furimazine (N113) diluted 250-fold in TBS+0.0.1% Tergitol was added, and samples again placed on a shaker for 5 miuntes. Luminescence was measured on a GMM+ luminometer. Bmax and Kd were calculated, and results depicted in FIG. 206.

ii) Activity at Various pH

ATG-4166 variant proteins were diluted to 0.2nM in TBS+0.01% BSA. 20uM pep289 (VS-HiBiT) to each variant protein sample and incubate at RT for 20 minutes. 990ul of each pH buffer solution was added to a well a deep well plate. 10ul of Furimazine (N113) was added to each well, and 50ul of each variant protein/peptide sample added to each well. The plate was incubated at RT for 12 minutes, and luminescence was measured on a GMM+ luminometer. Activity was calculated, and results depicted in FIG. 207.

pH Buffer Series Preparation: Components listed in the table below were mixed in 400ml of water. 30ml of the buffer was added to twelve 50ml tubes, and either NaOH or HCl used to create the necessary pH.

Table 19. Formulation for universal pH buffer

Component	Stock	Amount	Unit	Concentration (400ml)
Na Citrate	Powder	5.88	g	50mM
MES Hydrate	Powder	3.9	g	50mM
PIPES	Powder	6.05	g	50mM
HEPES	1M	20	ml	50mM
TAPS	Powder	4.87	g	50mM
Tergitol	100%	1	ml	0.25
Mazu	100%	0.1	ml	0.025
ATT	powder	57	mg	1mM

Example 134

Screen of ATG-5823, ATG-5824, and ATG-5825 Variants

Variant proteins (ATG-5823,ATG-5824, and ATG-5825) were purified using Magne His purification system (Promega). ATG-5146 was purified using an AKTA with Nickel Sepharose column. Purified proteins were diluted first to 200nM in TBS+0.01% BSA and

then further diluted to 0.2nM in TBS+0.01% BSA. A three-fold dilution series of pep263 was prepared starting at 100nM in TBS+0.01% BSA+0.02% tergitol. 50ul of each enzyme was combined with 50ul of the peptide dilution series. Samples were incubated on orbital shaker (600RPM) for 10 minutes. After incubation, 100ul of LCS (N205; Promega) was diluted 1:30 into TBS+0.01% BSA and added to each sample. Samples were incubated for 3 minutes at RT, and then luminescence measured on a GloMaxMulti+. Bmax and Kd were calculated, and results depicted in FIG. 208 show similar Bmax and Kd for these variants.

Example 135

Screen of ATG-5826 and ATG-5827 Variants

i. pep86

Variant proteins (ATG-5826 and ATG-5827) were purified using MagneHis purification system (Promega V8500). Purified proteins were diluted first to 200nM in TBS+0.01% BSA and then further diluted to 0.2nM in TBS+0.01% BSA. Two 2-fold titration series of pep86 were prepared starting at 1uM and 100nM in TBS+0.01% BSA+0.02% Tergitol. 50ul of ATG-5826 and ATG-5827 were combined with 50ul of the peptide dilution series that started at 1uM. 50ul of LgBiT protein (Promega; N401C) was combined with the peptide titration series that started at 100nM. Samples were incubated on orbital shaker (600RPM) for 10 minutes. After incubation 100ul of LCS (N205; Promega) was diluted 1:30 into TBS+0.01% BSA and added to each sample. Samples were incubated for 3 minutes at RT, and then luminescence measured on a GloMaxMulti+. Bmax and Kd were calculated, and results depicted in FIG. 209 show similar Bmax and Kd for these variants. Clones with 760 sequence for strand 9 showed a significantly higher Kd. This shows that variants such as these two containing the 760 sequence have higher Kd valued.

ii. pep114

Variant proteins (ATG-5826 and ATG-5827) were purified using MageHis purification system (Promega V8500). Purified proteins were diluted first to 200nM in TBS+0.01% BSA and then further diluted to 0.2nM in TBS+0.01% BSA. Two 2-fold titration series of pep114 was prepared starting at 1mM in TBS+0.01% BSA+0.02% tergitol. 50ul of ATG-5826, ATG-5827, and LgBiT protein (Promega N401C) were combined with the peptide titration series. Samples were incubated on orbital shaker (600RPM) for 30 minutes. After incubation, 100ul of LCS (N205 Promega) was diluted 1:30 into TBS+0.01% BSA and added to each sample. Samples were incubated for 3 minutes at RT, and then luminescence measured on a

GloMaxMulti+. Bmax and Kd were calculated, and results depicted in FIG. 210 show similar trend for Bmax values compared to pep86 (ATG-5826>LgBiT>ATG-5827). Both ATG-5826 and ATG-5827 have slightly lower calculated Kd values compared to LgBiT/pep114, but significantly higher Kd values compared to LgBiT/HiBiT(pep86).

Example 136

SDS-PAGE of ATG-5823, ATG-5824, ATG-5825, ATG-5826, and ATG-5827 Variants

Each variant protein was diluted to 0.1ug/ml in TBS+1X SDS loading dye. Samples were heated to 70°C for 5 minutes and then 3ul (0.3ug) was loaded to an SDS PAGE gel (BioRad Criterion). Results depicted in FIG. 211.

Example 137

Spike-in Antibody Titration Using 3 Labeling Methods:

NanoTrip, Sulfhydryl-Trip Labeling, and NHS-CA-HaloTag-Trip labeling

Figure 212 provides a demonstration of each labeling method.

i) NanoTrip-Genetic fusion

A two-fold antibody titration starting at 4ug/mL was prepared in PBS+0.01% BSA or PBS+20% human serum leaving the 24th well as the "no antibody" control. (Pool equal amounts of Sino antibodies D1-D5). 50ul of each antibody titration was added to wells of a white assay plate in triplicate. A master mix for each combination with each strand 9 and 10 protein and 2μM of LgTrip was prepared. 125ng/ml was used for ATG-5547 and ATG-5546 and 500ng/ml for ATG-5541. 50ul of the master mix for each combination (ATG-5546+ATG-5541 or ATG5547+ATG-5541) was added to the antibody titrations and incubated for 60 minutes at RT. A detection reagent was made by diluting NanoGlo® Live Cell Substate (N205) 1:30 in PBS+0.01% BSA, and 100ul added to each sample. Plates were incubated for three minutes, and then luminescence measured on a GloMax Discover. RLU readings were divided by the "no antibody" control to obtain Signal/Background readings.

ii) Sulfhydryl-Trip and NHS-CA-HaloTag-Trip labeling

FIG. 214 demonstrates that three different NanoTrip detection methods are capable of detecting SARS-COV/COV2 antibody.

Sino Biologicals (SARS-CoV/SARS-CoV2 spike)Antibodies:

D1	40150-D001	HA14MA0604
D2	40150-D002	HA14FE2802
D3	40150-D003	HA14FE2803
D4	40150-D004	MA14AP0203
D5	40150-D005	HA14FE2502

About 50ul 2X RBD reagents + 2X LgTrip prepared in 0.01% PBSA were added to wells of a 96-well assay plated. For the NHS-CA-HaloTag-Trip labeling, RBD reagent was prepared as a final concentration/well of 8ng/ml RBD-HaloTag-HiBiT + 16 ng/ml RBD-HaloTag-SmTrip9 (pep840) + 1uM LgTrip ATG-5146. For Suldhydryl-Trip labeling, RBD reagent was prepared as a final concentration/well of 15ng/ml RBD-HiBiT + 15 ng/ml SulfoSE-PEG6-RBD-SmTrip9 (pep840)-PSA (FIG. 213) + 1uM LgTrip ATG-5146. 50ul 2X pooled antibodies prepared in either 0.01% PBSA or 20% serum diluted in 0.01% PBSA was added to each sample and incubated for 45 minutes. 20uM NanoGlo® Live Cell Substrate was prepared in 0.01% PBSA, 100ul added to each well, and total luminescence read on a luminometer.

Example 138

SARS-CoV-2 Nucleocapsid Titration

Anti-nucleocapsid Ab clone 9547 (Meridian Biosciences) and anti-nucleocapsid Ab clone 9548 (Meridian Biosciences) were labeled with HaloTag-SmTrip9(pep840) and HaloTag-VSHiBiT, respectfully. 25ul/well of a 4x cocktail of Abs + LgTrip ATG-5146 was added to wells of a non-binding surface, solid white 96 well microtiter plate (Costar 3600) for a final concentration/well of 30ng/ml Ab-SmTrip9 + 60 ng/ml Ab-HiBiT + 1uM LgTrip ATG-5146. 25ul/well of a 4x solution of recombinant nucleocapsid protein (Meridian Biosciences Cat # 9560) was added to each well followed by the addition of 50ul/well of a 2x solution of NanoGlo® Live Cell Substrate for final concentration/well of 10uM substrate. Plates were incubate for 15 minutes, and total luminescence measured on GlowMax luminometer.

FIG. 215 demonstrates ternary NanoLuc-labeled antibodies detect SARS-CoV-2 Nucleocapsid protein.

Example 139

Point of Care Swabs for SARS-CoV-2 Nucleocapsid Protein in Nasopharyngeal Swab Samples

A stock solution containing 120 ng/ml Anti-nucleocapsid antibody HaloTag-SmTrip9, 240 ng/ml Anti-nucleocapsid antibody HaloTag-HiBiT, 4uM LgTrip ATG-5146, 40uM

Furimazine in ethanol, 1.2mM azo-thiothymine (ATT), 1.2 mM ascorbic acid, 0.6% pullulan w/v, 4.8 mM HEPES buffer pH 8.0, 21.6 mM Glycine, 4.8mM histidine, 6 mg/ml sucrose, and 0.0024% Polysorbate 80 was created. 100ul of the stock solution was dispensed into plastic swab jackets and loaded onto the lyophilizer (Virtis Genesis 12EL) with shelves pre-set to 4.0°C. Upon evacuation of the system, the lyophilization process was performed between shelf temperatures of -25°C and +25°C. The ice sublimation phase lasted 8hr, and the bound water desorption phase lasted 16hr. At the end of the lyophilization process, the swabs were back filled with N° and sealed by inserting a second plastic swab jacket manually.

Swabs containing the lyophilized assay reagents were then rehydrated with 100ul of nasopharyngeal swab samples + 300ul PBS containing 0.01% BSA. Total luminescence was measured on a handheld luminometer (Prom4ega) at times 15, 30, 45, and 60 minutes and plotted.

FIG. 216 show the results from 3 PCR confirmed negative (NS46, NS47, and NS52) and 3 PCR confirmed positive samples (PS46, PS49, and PS56).

Example 140

Monomeric NanoBiT with Fluoro-Fz

Each enzyme to be tested was diluted into TBS+0.01% BSA. A titration series with either Furimazine (N205) or JRW-1677 was made either TBS (start at 20uM) or NanoGlo® buffer (start at 25uM; Promega N112). Each substrate was serially diluted with either TBS+0.01% BSA or NanoGlo® buffer. 50ul of each enzyme dilution as combined with 50ul of each substrate titration. Plates were incubated for 3 minutes, and then luminescence measured on GMM+ luminometer. GraphPad Prism was used to generate a non-linear regression using Michaelis-Menten Least squares fit.

FIG. 217 provide the kinetic parameters (Vmax and Km) for Furimazine and JRW-1667. NanoLuc (ATG-462) produced higher RLU values with both Furimazine and JRW-1667 (~10-20 fold) compared to the Monomeric NanoBiT constructs when TBS is used to dilute substrates. NanoLuc and the monomeric constructs showed similar RLU values in Furimazin/NanoGlo® buffer, but only NanoLuc showed improved luminescence with JRW-1667. Although the monomeric NanoBiT constructs showed lower RLU values with both Furimazine and JRW-1667, RLU values were similar for the two buffers and two substrates.

Km values were lower in TBS+0.01% BSA compared to NanoGlo, but generally the Km values were similar for each condition except ATG-3562 (Monomeric LgBiT-SmBiT), which shows a lower calculated Km for all conditions tested.

Example 141

Mammalian cell expression and fluorescence imaging

Transfection protocol: Preparing cells: Aspirated media from HeLa cells that were grown to confluency in a T-75 flask and washed cells with 10ml DPBS. (Life Technologies 14190). Aspirated DPBS and added 2ml of TryPLE Express Trypsin (Life Technologies 12604). Incubated cells for 2-3 minutes at 37°C then resuspended cells in 8ml of growth media (DMEM Life Technologies 11995)+10% FBS (VWR 89510-194). Spun cells at 200RPM for 5 minutes. Aspirated supernatant and added 10ml of DMEM+10% FBS. Counted cells and then diluted to 100,000 cells per ml. Plated 3 mL of cells to each well of 6 well plate. Incubated cells for 24 h.

Transfection complex: For each DNA sample to be tested, combined and mixed 2.5µg of DNA (0.25µg of DNA encoding sensor, and 2.25 µg of carrier DNA) (Promega E4882) with 100ul of OptiMem (Life Technologies 11058). Next added 7.5µl of Fugene HD transfection reagent (Promega E2311) and incubated transfection complex for 10 min at ambient temperature. Added transfection complex to each well containing plated cells. Incubated cells for 24 h.

Fluorescence Imaging: Replated transfected cells to 8-well chamber slides (MatTek glass bottom). Aspirated media from transfected cells HeLa cells and washed cells with 3ml DPBS. (Life Technologies 14190). Aspirated DPBS and add 0.5ml of TryPLE Express Trypsin (Life Technologies 12604). Incubatde cells for 2-3 minutes at 37°C then resuspended cells in 3ml of growth media (DMEM Life Technologies 11995)+10% FBS (VWR 89510-194). Counted cells and then diluted to 50,000 cells per mL. Plated 500µl of each well. For cellular markers that are delivered by BacMam transduction, after 3h of plating, add 10% (v/v) of CellLight Plasma Membrane-GFP (Thermo Fisher C10607) or CellLight Golgi-GFP (Thermo Fisher C10592) or CellLight Lysosome-GFP (Thermo Fisher C10507). Incubated cells for another 16-24h.

The following morning, aspirated growth media and replaced with 400µl of FluoroBrite DMEM+10%FBS (Life Technologies A1896701)+10% FBS. Prepared 5X dilution of Janelia Fluor 549 HaloTag (Promega GA1110) in FluoroBrite DMEM+10%FBS,

and add 100 µl to each well. Incubated for 30 min. Washed cells with 500µl FluoroBrite DMEM+10%FBS for 30 min. Cells were stained with appropriate cellular markers. Staining procedure was followed by manufacture recommendation. Stained mitochondria with MitoTracker Green FM (Thermo Fisher M7514). Stain endoplasmic reticulum with ER-Tracker Red (Thermo Fisher E34250). Counterstain cells with the nuclear probe NucBlue Live ReadyProbes reagent (Thermo Fisher R37605). Image cells with C2 laser scanning confocal microscope (Nikon).

Results are depicted in images of Figures 218-224.

Example 142

Mammalian cell expression and luminescence assays

Transfection protocol: Preparing cells: Aspirated media from HeLa cells (PKCα-HiBiT clone) that were grown to confluency in a T-150 flask and wash cells with 10ml DPBS. (Life Technologies 14190). Aspirated DPBS and added 4ml of TryPLE Express Trypsin (Life Technologies 12604). Incubatde cells for 2-3 minutes at 37°C then resuspended cells in 16ml of growth media (DMEM Life Technologies 11995)+10% FBS (VWR 89510-194). Spun cells at 200RPM for 5 minutes. Aspirated supernatant and added 20ml of DMEM+10% FBS. Counted cells and then diluted to 1,000,000 cells per ml. Plated 1 mL of cells and 3 mL of DMEM+10%FBS in 6 cm dish. Incubated cells for 24 h.

Transfection complex: For each DNA sample to be tested, combined and mixed 10ug of DNA (1ug of DNA encoding membrane sensor, and 9 ug of carrier DNA) (Promega E4882) with 400ul of OptiMem (Life Technologies 11058). Next, added 30µl of FugeneHD transfection reagent (Promega E2311) and incubated transfection complex for 10 min at ambient temperature. Added transfection complex to 6 cm dish containing plated cells. Incubated cells for 24 h.

Cell-based luminescence assay: Replated transfected cells to white 96 well assay plate. Aspirated media from transfected cells HeLa cells in 6 cm dish and washed cells with 5ml DPBS. (Life Technologies 14190). Aspirated DPBS and add 0.75ml of TryPLE Express Trypsin (Life Technologies 12604). Incubated cells for 2-3 minutes at 37°C then resuspended cells in 4ml of growth media (DMEM Life Technologies 11995)+10% FBS (VWR 89510-194). Plated 100µl of each sample in 30 wells of a white 96 well assay plate. Incubated cells for another 16-24h.

The following morning, aspirated growth media and replaced with 90µl of CO2 independent media (Life Technologies 18045)+10% FBS containing 1.1X NanoGlo Live Cell Substrate (Promega N2012). Incubated for 10 min. Added 10µl of titrated PMA compound to each well. (Final PMA concentration is presented in the graph). Placed plate in a GlomaxMulti+ luminometer set to 37°C and ran a kinetic over the course of 1 hours. Unless otherwise stated, the reported luminescence is the peak height of the kinetic run.

Results depicted in Figures 225-227.

Example 143

LgTrip 3546 Optimization

Experiments were conducted during development of embodiments herein to optimize the identity of the amino acid at various positions of LgTrip 3546 (SEQ ID NO: 51), as shown below in Table 20. *E. coli* cultures (200µl) were prepared for each sample and grown overnight at 37°C in LB media+100µg/ml ampicillin. Cultures were then diluted in quadruplicate to a 20X concentration (10µl in 200µl) into induction media (LB+ampicllin+0.1% Rhamnose). Samples were grown at 37°C for 3 hours. Samples were then lysed with 0.3X PLB+25mM HEPES pH 7.5+0.001U/ml DNase (10µl of cells to 250µl of Lysis buffer). 50µl of the lysate was then combined with 50µl of NanoGlo® buffer+50µM furimazine+50nM of dipeptide 788 (SEQ ID NO: 51). Samples were measured on a BMG Clariostar luminometer. RLU values were normalized to LgTrip 3546 (SEQ ID NO: 51).

Table 20. Amino Acid Substitutions for LgTrip 3546 Optimization

```
01        10         20         30         40         50         60         70
||        |          |          |          |          |          |          |
VFTLDDFVGDW EQTAAYNLDQ VLEQGGVSSL LQNLAVSVTP IMRIVRSGEN ALKIDIHVII PYEGLSADQM

71      80         90        100        110        120        130        140
|       |          |         |          |          |          |          |
AQIEEVFKVV YPVDDHHFKV ILPYGTLVID GVTPNKLNYF GRPYEGIAVF DGKKITTTGT LWNGNKIIDE

141  146
|    |
RLITPD
```

Amino acid substitution	Fold-improvement over 3546
V0D	1.90
T13E	2.62
S28P	2.12
S29G	1.83
S29P	2.02
L30F	1.66

L30T	1.65
L30Y	1.95
T39K	1.81
K53S	1.58
Y62N	1.46
G64E	1.88
Y81E	1.97
H86D	2.52
H87N	1.83
K89E	1.89
Y94I	2.18
L97V	1.85
I99A	1.60
I99L	1.49
I99M	1.57
V102E	1.85
R112H	2.37
E115D	2.38
I117Q	2.06
I117T	1.76
V119A	1.66
V119D	1.72
V119E	1.62
V119G	1.81
V119N	1.89
K123E	1.95
K123G	1.57
W132P	1.96
L142C	1.93
T144D	1.81
P145D	1.48

SEQUENCES

The following polypeptide sequences each comprise an N-terminal methionine residue or corresponding ATG codon; polypeptide sequences lacking the N-terminal methionine residue or corresponding ATG codon are also within the scope herein and are incorporated herein by reference.

The following peptide sequences (and the peptide sequences of Table 1) each lack an N-terminal methionine residue; peptide sequences comprising an N-terminal methionine residue are also within the scope herein and are incorporated herein by reference.

Some embodiments described herein make reference to a His-tagged (or non-His-tagged) sequence within Table 1; alterantive embodiments utilizing a non-His-tagged (or His-

tagged) version of the sequence, either appearing in Table 1 or not listed, are within the scope herein.

Table 1. Exemplary peptide, dipeptide, and polypeptide sequences.

SEQ ID NO	Name	Sequence
1	WT OgLuc	MFTLADFVGDWQQTAGYNQDQVLEQGGLSSLFQALGV SVTPIQKVVLSGENGLKADIHVIIPYEGLSGFQMGLIEMIF KVVYPVDDHHFKIILHYGTLVIDGVTPNMIDYFGRPYPGI AVFDGKQITVTGTLWNGNKIYDERLINPDGSLLFRVTIN GVTGWRLCENILA
2	WT OgLuc	atggtgtttaccttggcagatttcgttggagactggcaacagacagctggatacaaccaag atcaagtgttagaacaaggaggattgtctagtctgttccaagccctgggagtgtcagtcac cccaatccagaaagttgtgctgtctggggagaatgggttaaaagctgatattcatgtcatca tcccttacgagggactcagtggttttcaaatgggtctgattgaaatgatcttcaaagttgttta cccagtggatgatcatcatttcaagattattctccattatggtacactcgttattgacggtgtg acaccaaacatgattgactactttggacgcccttaccctggaattgctgtgtttgacggcaa gcagatcacagttactggaactctgtggaacggcaacaagatctatgatgagcgcctgat caacccagatggttcactcctcttccgcgttactatcaatggagtcaccggatggcgccttt gcgagaacattcttgcc
3	NanoLuc	MKHHHHHHAIAMVFTLEDFVGDWRQTAGYNLDQVLEQ GGVSSLFQNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLS GDQMGQIEKIFKVVYPVDDHHFKVILHYGTLVIDGVTPN MIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINP DGSLLFRVTINGVTGWRLCERILAV
4	NanoLuc	atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttgg ggactggcgacagacagccggctacaacctggaccaagtccttgaacagggaggtgtg tccagtttgtttcagaatctcggggtgtccgtaactccgatccaaaggattgtcctgagcgg tgaaaatgggctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcggcga ccaaatgggccagatcgaaaaaatttttaaggtggtgtaccctgtggatgatcatcacttta aggtgatcctgcactatggcacactggtaatcgacggggttacgccgaacatgatcgact atttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacag gaccctgtggaacggcaacaaaattatcgacgagcgcctgatcaaccccgacggctcc ctgctgttccgagtaaccatcaacggagtgaccggctggcggctgtgcgaacgcattctg gcggtt
5	WT OgLuc Lg	MFTLADFVGDWQQTAGYNQDQVLEQGGLSSLFQALGV SVTPIQKVVLSGENGLKADIHVIIPYEGLSGFQMGLIEMIF KVVYPVDDHHFKIILHYGTLVIDGVTPNMIDYFGRPYPGI AVFDGKQITVTGTLWNGNKIYDERLINPD
6	WT OgLuc β9	GSLLFRVTIN
7	WT OgLuc β10	GVTGWRLCENILA

8	WT NanoLuc Lg	MVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNLG VSVTPIQRIVLSGENGLKIDIHVIIPYEGLSGDQMGQIEKIF KVVYPVDDHHFKVILHYGTLVIDGVTPNMIDYFGRPYEG IAVFDGKKITVTGTLWNGNKIIDERLINPD
9	WT NanoLuc β9	GSLLFRVTINV
10	WT NanoLuc β10	GVTGWRLCERILA
11	LgBit	MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYE GIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTI NSHHHHHH
12	LgBit	atggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcat caccatcaccac
13	SmBit	VTGYRLFEEIL
14	SmBit	gtgaccggctaccggctgttcgaggagattctg
15	HiBit	VSGWRLFKKIS
16	HiBit	gtgagcggctggcggctgttcaagaagattagc
17	LgTrip 2098	MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYE GIAVFDGKKITVTGTLWNGNKIIDERLITPD
18	LgTrip 2098	atggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggattgtccggagcggtgaaaatgCctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgac
19	LgTrip 3092 His	MKHHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPD

20	LgTrip 3092 His	atgaaacatcaccatcaccatcatgtcttcacactcgaagatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatccaaaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacatgctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgac
21	LgTrip 3092	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITVTGTLWNGNKIIDERLITPD
22	LgTrip 3092	atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgac
23	SmTrip9	GSMLFRVTINS
24	SmTrip9	ggctccatgctgttccgagtaaccatcaacagc
25	SmHiTrip10	VSGWRLFKKIS
26	SmHiTrip10	gtgagcggctggcggctgttcaagaagattagc
27	5P-B9	MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLFQNLA VSVTPIQRIVLSGENALKIDIHVIIPYEGLSADQMAQIEKIF KVVYPVDDHHFKVILHYGTLVIDGVTPNMINYFGRPYEG IAVFDGKKITVTGTLWNGNKIIDERLITPD
28	5P-B9	atggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgtttcagaatctcgccgtgtccgtaac tccgatccaaaggattgtcctgagcggtgaaaatgccctgaagatcgacatccatgtcatc atcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaaaaaatttttaaggtg gtgtaccctgtggatgatcatcactttaaggtgatcctgcactatggcacactggtaatcga cggggttacgccgaacatgatcaactatttcggacggccgtatgaaggcatcgccgtgtt cgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgacg agcgcctgatcacccccgac
29	5P(147-157)	GSMLFRVTINV
30	5P(147-157)	ggctccatgctgttccgagtaaccatcaac
31	LgTrip 2098 His	MKHHHHHHVFTLEDFVGDWEQTAAYNLDQVLEQGGVS SLLQNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQ MAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLN YFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPD

32	LgTrip 2098 His	atgaaacatcaccatcaccatcatgtcttcacactcgaagatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatccaaaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacatgctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgac
35	SmTrip9/10 Dipeptide (pep263)	GSMLFRVTINSVSGWRLFKKIS
36	SmTrip9/10 Dipeptide (pep263)	ggctccatgctgttccgagtaaccatcaacagcgtgagcggctggcggctgttcaagaag attagc
37	SmTrip9+ (pep286)	SSWKRGSMLFRVTINS
38	SmTrip9+ (pep286)	Agcagctggaagcgcggctccatgctgttccgagtaaccatcaacagc
39	LgTrip 3440	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGDTPNKL NYFGRPYDGIAVFDGKKITVTGTLWNGNKIIDERLITPD
40	LgTrip 3440	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggatacgccgaacaagctgaactatttcggacgg ccgtatgatggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgac
41	LgTrip 3121	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPSKL NYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPD
42	LgTrip 3121	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgagcaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgac
43	LgTrip 3482	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGFAVFDGKKITVTGTLWNGNKIIDERLITPD

44	LgTrip 3482	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcttcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtgga acggcaacaaaattatcgacgagcgcctgatcacccccgac
45	LgTrip 3497	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVCDGKKITVTGTLWNGNKIIDERLITPD
46	LgTrip 3497	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgtgcgacggcaaaaagatcactgtaacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgac
47	LgTrip 3125	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKISVTGTLWNGNKIIDERLITPD
48	LgTrip 3125	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcgggcggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatctctgtaacagggaccctgtgga acggcaacaaaattatcgacgagcgcctgatcacccccgac
49	LgTrip 3118	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITATGTLWNGNKIIDERLITPD
50	LgTrip 3118	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgcaacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgac
51	LgTrip 3546	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPD

52	LgTrip 3546	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgac
53	LgTrip 3546+G (ATG 3572)	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITTTGTLWNGNKIIDERLITPDG
54	LgTrip 3546+G (ATG 3572)	atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacggc
55	LgTrip 3546-D (ATG 3573)	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITTTGTLWNGNKIIDERLITP
56	LgTrip 3546-D (ATG 3573)	atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcaccccc
57	LgTrip 3546-PD (ATG 3574)	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITTTGTLWNGNKIIDERLIT
58	LgTrip 3546-PD (ATG 3574)	atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacc
59	LgTrip 3546+GS (ATG 3575)	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITTTGTLWNGNKIIDERLITPDGS

60	LgTrip 3546+GS (ATG 3575)	atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacggcagc
61	-V_LgBiT (ATG3618)	MFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAV SVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVF KVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYE GIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTI NSHHHHHH
62	-V_LgBiT (ATG3618)	atgttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctggac caagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactc cgatccaaaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatca tcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtg gtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcga cggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgtt cgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgacg agcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcatc accatcaccactaa
63	-VF_LgBiT (ATG3619)	MTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS VTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFK VVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYEGI AVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINS HHHHHH
64	-VF_LgBiT (ATG3619)	atgacactcgaagatttcgttggggactgggaacagacagccgcctacaacctggacca agtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccg atccaaaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcc cgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgt accctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacgg gttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgttcga cggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgacgagc gcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcatcacc atcaccactaa
65	-VFT_LgBiT (ATG3620)	MLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSV TPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKV VYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYEGIA VFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINSH HHHHH
66	-VFT_LgBiT (ATG3620)	atgctcgaagatttcgttggggactgggaacagacagccgcctacaacctggaccaagt ccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatc caaaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccg tatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtac cctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggt tacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacgg caaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgacgagcgcc

		tgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcatcaccatca ccactaa
67	-VFTL_LgBiT (ATG3621)	MEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVT PIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVV YPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYEGIAV FDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINSHH HHHH
68	-VFTL_LgBiT (ATG3621)	atggaagatttcgttggggactgggaacagacagccgcctacaacctggaccaagtcctt gaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaa ggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatga aggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgt ggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacg ccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaa aagatcactgtaacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgat caccccgacggctccatgctgttccgagtaaccatcaacagccatcatcaccatcacca ctaa
69	(M)FKKIS-GSSG-LgBiT (ATG3632)	MFKKISGSSGVFTLEDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNML NYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDG SMLFRVTINSHHHHHH
70	(M)FKKIS-GSSG-LgBiT (ATG3632)	atgttcaagaagattagcggctcgagcggtgtcttcacactcgaagatttcgttggggact gggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagt ttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccggagcggtgaaa atgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaat ggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtg atcctgccctatggcacactggtaatcgacggggttacgccgaacatgctgaactatttcg gacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggacc ctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgacggctccatgct gttccgagtaaccatcaacagccatcatcaccatcaccactaa
71	(M)KKIS-GSSG-LgBiT (ATG3633)	MKKISGSSGVFTLEDFVGDWEQTAAYNLDQVLEQGGVS SLLQNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQ MAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLN YFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGS MLFRVTINSHHHHHH
72	(M)KKIS-GSSG-LgBiT (ATG3633)	atgaagaagattagcggctcgagcggtgtcttcacactcgaagatttcgttggggactgg gaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagttt gctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccggagcggtgaaaa tgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatg gcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgat cctgccctatggcacactggtaatcgacggggttacgccgaacatgctgaactatttcgga cggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccct

		gtggaacggcaacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgtt ccgagtaaccatcaacagccatcatcaccatcaccactaa
73	(M)KIS-GSSG-LgBiT (ATG3634)	MKISGSSGVFTLEDFVGDWEQTAAYNLDQVLEQGGVSS LLQNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQM AQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNY FGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSM LFRVTINSHHHHHH
74	(M)KIS-GSSG-LgBiT (ATG3634)	atgaagattagcggctcgagcggtgtcttcacactcgaagatttcgttggggactgggaac agacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgc agaatctcgccgtgtccgtaactccgatccaaaggattgtccggagcggtgaaaatgccc tgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccc agatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctg ccctatggcacactggtaatcgacggggttacgccgaacatgctgaactatttcggacgg ccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtg gaacggcaacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttcc gagtaaccatcaacagccatcatcaccatcaccactaa
75	(M)IS-GSSG-LgBiT (ATG3635)	MISGSSGVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSL LQNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMA QIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYF GRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSML FRVTINSHHHHHH
76	(M)IS-GSSG-LgBiT (ATG3635)	atgattagcggctcgagcggtgtcttcacactcgaagatttcgttggggactgggaacaga cagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcaga atctcgccgtgtccgtaactccgatccaaaggattgtccggagcggtgaaaatgccctga agatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccaga tcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccct atggcacactggtaatcgacggggttacgccgaacatgctgaactatttcggacggccgt atgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaac ggcaacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagta accatcaacagccatcatcaccatcaccactaa
77	(M)S-GSSG-LgBiT (ATG3636)	MSGSSGVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLL QNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMA QIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYF GRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSML FRVTINSHHHHHH
78	(M)S-GSSG-LgBiT (ATG3636)	atgagcggctcgagcggtgtcttcacactcgaagatttcgttggggactgggaacagaca gccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaat ctcgccgtgtccgtaactccgatccaaaggattgtccggagcggtgaaaatgccctgaag atcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatc gaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctat ggcacactggtaatcgacggggttacgccgaacatgctgaactatttcggacggccgtat gaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacg

		gcaacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcatcaccatcaccactaa
79	LgTrip + GSM (ATG3722)	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDGSM
80	LgTrip + GSM (ATG3722)	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgacggcagcatgtaa
81	LgTrip + GSML (ATG3723)	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDGSML
82	LgTrip + GSML (ATG3723)	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgacggcagcatgctgtaa
83	LgTrip + GSMLF (ATG3724)	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDGSMLF
84	LgTrip + GSMLF (ATG3724)	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgacggcagcatgctgttctaa

85	LgTrip – TPD (ATG3725)	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLI
86	LgTrip – TPD (ATG3725)	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatctaa
87	LgTrip – ITPD (ATG3726)	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERL
88	LgTrip – ITPD (ATG3726)	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgtaa
89	LgTrip – LITPD (ATG3727)	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDER
90	LgTrip – LITPD (ATG3727)	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtgg aacggcaacaaaattatcgacgagcgctaa
91	FRB-15GS-AI-86 (ATG1634)	MVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPL HAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMK SGNVKDLTQAWDLYYHVFRRISGGSGGGGSGGSSSGGA IVSGWRLFKKIS
92	FRB-15GS-AI-86 (ATG1634)	atggtggccatcctctggcatgagatgtggcatgaaggcctggaagaggcatctcgtttgt actttggggaaaggaacgtgaaaggcatgtttgaggtgctggagcccttgcatgctatgat ggaacggggccccccagactctgaaggaaacatcctttaatcaggcctatggtcgagattt aatggaggcccaagagtggtgcaggaagtacatgaaatcagggaatgtcaaggacctc acccaagcctgggacctctattatcatgtgttccgacgaatcagtggtggttcaggtggtg gcgggagcggtggctcgagcagcggtggagcgatcgtgagcggctggcggctgttca agaagattagctaa

93	FRB-15GS-AI-289 (ATG3586)	MVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPL HAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMK SGNVKDLTQAWDLYYHVFRRISGGSGGGGSGGSSSGGA IVSVSGWRLFKKIS
94	FRB-15GS-AI-289 (ATG3586)	atggtggccatcctctggcatgagatgtggcatgaaggcctggaagaggcatctcgtttgt actttggggaaaggaacgtgaaaggcatgtttgaggtgctggagcccttgcatgctatgat ggaacggggcccccagactctgaaggaaacatcctttaatcaggcctatggtcgagattt aatggaggcccaagagtggtgcaggaagtacatgaaatcagggaatgtcaaggacctc acccaagcctgggacctctattatcatgtgttccgacgaatcagtggtggttcaggtggtg gcgggagcggtggctcgagcagcggtggagcgatcgttagcgttagcggctggcgcct gttcaagaagatcagctaa
95	FRB-15GS-AI-86-His6 (ATG3743)	MVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPL HAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMK SGNVKDLTQAWDLYYHVFRRISGGSGGGGSGGSSSGGA IVSGWRLFKKISHHHHHH
96	FRB-15GS-AI-86-His6 (ATG3743)	atggtggccatcctctggcatgagatgtggcatgaaggcctggaagaggcatctcgtttgt actttggggaaaggaacgtgaaaggcatgtttgaggtgctggagcccttgcatgctatgat ggaacggggcccccagactctgaaggaaacatcctttaatcaggcctatggtcgagattt aatggaggcccaagagtggtgcaggaagtacatgaaatcagggaatgtcaaggacctc acccaagcctgggacctctattatcatgtgttccgacgaatcagtggtggttcaggtggtg gcgggagcggtggctcgagcagcggtggagcgatcgtgagcggctggcggctgttca agaagattagccatcatcaccatcaccactaa
97	FRB-15GS-AI-289-His6 (ATG3744)	MVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPL HAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMK SGNVKDLTQAWDLYYHVFRRISGGSGGGGSGGSSSGGA IVSVSGWRLFKKISHHHHHH
98	FRB-15GS-AI-289-His6 (ATG3744)	atggtggccatcctctggcatgagatgtggcatgaaggcctggaagaggcatctcgtttgt actttggggaaaggaacgtgaaaggcatgtttgaggtgctggagcccttgcatgctatgat ggaacggggcccccagactctgaaggaaacatcctttaatcaggcctatggtcgagattt aatggaggcccaagagtggtgcaggaagtacatgaaatcagggaatgtcaaggacctc acccaagcctgggacctctattatcatgtgttccgacgaatcagtggtggttcaggtggtg gcgggagcggtggctcgagcagcggtggagcgatcgttagcgtgagcggctggcggc tgttcaagaagattagccatcatcaccatcaccactaa
99	His6-FRB-5GS-86 (ATG3760)	MKHHHHHHVAILWHEMWHEGLEEASRLYFGERNVKG MFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQE WCRKYMKSGNVKDLTQAWDLYYHVFRRISGGSGGVSG WRLFKKIS
100	His6-FRB-5GS-86 (ATG3760)	atgaaacatcaccatcaccatcatgtggccatcctctggcatgagatgtggcatgaaggcc tggaagaggcatctcgtttgtactttggggaaaggaacgtgaaaggcatgtttgaggtgct ggagcccttgcatgctatgatggaacggggcccccagactctgaaggaaacatcctttaa tcaggcctatggtcgagatttaatggaggcccaagagtggtgcaggaagtacatgaaatc agggaatgtcaaggacctcacccaagcctgggacctctattatcatgtgttccgacgaatc agtggtggttcaggtggtgtgagcggctggcggctgttcaagaagattagctaa
101	His6-FRB-10GS-86 (ATG3761)	MKHHHHHHVAILWHEMWHEGLEEASRLYFGERNVKG MFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQE WCRKYMKSGNVKDLTQAWDLYYHVFRRISGGSGGGGS GGVSGWRLFKKIS

102	His6-FRB-10GS-86 (ATG3761)	atgaaacatcaccatcaccatcatgtggccatcctctggcatgagatgtggcatgaaggcc tggaagaggcatctcgtttgtactttggggaaaggaacgtgaaaggcatgtttgaggtgct ggagcccttgcatgctatgatggaacggggcccccagactctgaaggaaacatcctttaa tcaggcctatggtcgagatttaatggaggcccaagagtggtgcaggaagtacatgaaatc agggaatgtcaaggacctcacccaagcctgggacctctattatcatgtgttccgacgaatc agtggtggttcaggtggtggcgggagcggtggcgtgagcggctggcggctgttcaaga agattagctaa
103	His6-FRB-15GS-86 (ATG3762)	MKHHHHHHVAILWHEMWHEGLEEASRLYFGERNVKG MFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQE WCRKYMKSGNVKDLTQAWDLYYHVFRRISGGSGGGGS GGSSSGGVSGWRLFKKIS
104	His6-FRB-15GS-86 (ATG3762)	atgaaacatcaccatcaccatcatgtggccatcctctggcatgagatgtggcatgaaggcc tggaagaggcatctcgtttgtactttggggaaaggaacgtgaaaggcatgtttgaggtgct ggagcccttgcatgctatgatggaacggggcccccagactctgaaggaaacatcctttaa tcaggcctatggtcgagatttaatggaggcccaagagtggtgcaggaagtacatgaaatc agggaatgtcaaggacctcacccaagcctgggacctctattatcatgtgttccgacgaatc agtggtggttcaggtggtggcgggagcggtggctcgagcagcggtggagtgagcggct ggcggctgttcaagaagattagctaa
105	His6-FRB-5GS-289 (ATG3763)	MKHHHHHHVAILWHEMWHEGLEEASRLYFGERNVKG MFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQE WCRKYMKSGNVKDLTQAWDLYYHVFRRISGGSGGVSV SGWRLFKKIS
106	His6-FRB-5GS-289 (ATG3763)	atgaaacatcaccatcaccatcatgtggccatcctctggcatgagatgtggcatgaaggcc tggaagaggcatctcgtttgtactttggggaaaggaacgtgaaaggcatgtttgaggtgct ggagcccttgcatgctatgatggaacggggcccccagactctgaaggaaacatcctttaa tcaggcctatggtcgagatttaatggaggcccaagagtggtgcaggaagtacatgaaatc agggaatgtcaaggacctcacccaagcctgggacctctattatcatgtgttccgacgaatc agtggtggttcaggtggtgttagcgttagcggctggcgcctgttcaagaagatcagctaa
107	His6-FRB-10GS-289 (ATG3764)	MKHHHHHHVAILWHEMWHEGLEEASRLYFGERNVKG MFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQE WCRKYMKSGNVKDLTQAWDLYYHVFRRISGGSGGGGS GGVSVSGWRLFKKIS
108	His6-FRB-10GS-289 (ATG3764)	atgaaacatcaccatcaccatcatgtggccatcctctggcatgagatgtggcatgaaggcc tggaagaggcatctcgtttgtactttggggaaaggaacgtgaaaggcatgtttgaggtgct ggagcccttgcatgctatgatggaacggggcccccagactctgaaggaaacatcctttaa tcaggcctatggtcgagatttaatggaggcccaagagtggtgcaggaagtacatgaaatc agggaatgtcaaggacctcacccaagcctgggacctctattatcatgtgttccgacgaatc agtggtggttcaggtggtggcgggagcggtggcgttagcgttagcggctggcgcctgtt caagaagatcagctaa
109	His6-FRB-15GS-289 (ATG3765)	MKHHHHHHVAILWHEMWHEGLEEASRLYFGERNVKG MFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQE WCRKYMKSGNVKDLTQAWDLYYHVFRRISGGSGGGGS GGSSSGGVSVSGWRLFKKIS

110	His6-FRB-15GS-289 (ATG3765)	atgaaacatcaccatcaccatcatgtggccatcctctggcatgagatgtggcatgaaggcc tggaagaggcatctcgtttgtactttggggaaaggaacgtgaaaggcatgtttgaggtgct ggagcccttgcatgctatgatggaacggggccccccagactctgaaggaaacatcctttaa tcaggcctatggtcgagatttaatggaggcccaagagtggtgcaggaagtacatgaaatc agggaatgtcaaggacctcacccaagcctgggacctctattatcatgtgttccgacgaatc agtggtggttcaggtggtggcgggagcggtggctcgagcagcggtggagttagcgttag cggctggcgcctgttcaagaagatcagctaa
111	SmTrip9-FKBP fusion template (ATG780)	M— GSMLFRVTINS — SSSGGGGSGGGSSGGGVQVETISPGDGRTFPKRGQTCVV HYTG MLEDGKKFDSSRDRNKPFKFMLGKQEVIRGWEEGVAQ MSVGQRAKLTISPDYAYGATGHPGIIPPHATLVFDVELL KLE
112	SmTrip9-FKBP fusion template (ATG780)	atgggctccatgctgttccgagtaaccatcaacagctcgagttcaggtggtggcgggagc ggtggagggagcagcggtggaggagtgcaggtggaaaccatctccccaggagacgg gcgcaccttccccaagcgcggccagacctgcgtggtgcactacaccgggatgcttgaag atggaaagaaatttgattcctcccgggacagaaacaagccctttaagtttatgctaggcaa gcaggaggtgatccgaggctgggaagaaggggttgcccagatgagtgtgggtcagag agccaaactgactatatctccagattatgcctatggtgccactgggcacccaggcatcatc ccaccacatgccactctcgtcttcgatgtggagcttctaaaactggaataa
113	FKBP-SmTrip9 fusion template (ATG777)	MGVQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKFD SSRDRNKPFKFMLGKQEVIRGWEEGVAQMSVGQRAKLT ISPDYAYGATGHPGIIPPHATLVFDVELLKLEGGSGGGGS GGSSSGGAI— GSMLFRVTINS
114	FKBP-SmTrip9 fusion template (ATG777)	Atgggagtgcaggtggaaaccatctccccaggagacgggcgcaccttccccaagcgc ggccagacctgcgtggtgcactacaccgggatgcttgaagatggaaagaaatttgattcc tcccgggacagaaacaagccctttaagtttatgctaggcaagcaggaggtgatccgagg ctgggaagaaggggttgcccagatgagtgtgggtcagagagccaaactgactatatctc cagattatgcctatggtgccactgggcacccaggcatcatcccaccacatgccactctcgt cttcgatgtggagcttctaaaactggaaggtggttcaggtggtggcgggagcggtggctc gagcagcggtggagcgatcggctccatgctgttccgagtaaccatcaacagc
115	LgBiT (ATG2623)	MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYE GIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTI NSHHHHHH
116	LgBiT (ATG2623)	atggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcat caccatcaccactaa
133	pep78	NVSGWRLFKKISN
134	pep79	NVTGYRLFKKISN

135	pep80	VSGWRLFKKISN
136	pep81	SGWRLFKKISN
137	pep82	GWRLFKKISN
138	pep99	VTGYRLFEKIS
139	pep219	SGWRLFKKIS
140	pep225	VSGWRL
141	pep226	VSGWRLF
142	pep227	VSGWRLFK
143	pep228	VSGWRLFKK
144	pep229	VSGWRLFKKI
145	pep243	VSGWRLYKKIS
146	pep272	GSMLFRVTINSVSGWALFKKIS
147	pep274	GSMLFRVTINSVTGYRLFEEIL
148	pep287 (WT SmTrip9)+Cterm solubility tag	GSMLFRVTINSSSWKR
149	pep288	VSGVSGWRLFKKIS
150	pep289	VSVSGWRLFKKIS
151	pep290	VVSGWRLFKKIS
152	pep291	SSWKRSMLFRVTINS
153	pep292	SSWKRMLFRVTINS
154	pep293	SSWKRDGSMLFRVTINS
155	pep294	SSWKRPDGSMLFRVTINS
156	pep296	SSWKRSMLFRVTINSV
157	pep297	SSWKRMLFRVTINSV
158	pep298	SSWKRDGSMLFRVTINSV
159	pep299	SSWKRPDGSMLFRVTINSV
160	pep301	SSWKRSMLFRVTINSVS
161	pep302	SSWKRMLFRVTINSVS
162	pep303	SSWKRDGSMLFRVTINSVS
163	pep304	SSWKRPDGSMLFRVTINSVS
164	pep305	SSWKRGSMLFRVTIN
165	pep306	SSWKRGSMLFRVTI
166	pep307	SSWKRSMLFRVTIN
167	pep308	SSWKRMLFRVTIN
168	pep309	SSWKRDGSMLFRVTIN
169	pep310	SSWKRPDGSMLFRVTIN
170	pep311	SSWKRSMLFRVTI
171	pep312	SSWKRMLFRVTI
172	pep313	SSWKRDGSMLFRVTI

173	pep314	SSWKRPDGSMLFRVTI
174	pep316	VSGWRLFKKISVFTL
175	pep317	VSGWRLFKKISVFT
176	pep318	VSGWRLFKKISVF
177	pep319	VSGWRLFKKISV
178	pep320	VSGWRLCKKIS
179	pep326	VSGWRLFKKISGSMLFRVTINS
180	pep380	SSWKRLFRVTINS
181	pep383	SSWKRFRVTINS
182	pep386	SSWKRRVTINS
183	pep389	SSWKRTPDGSMLFRVTINS
184	pep392	SSWKRITPDGSMLFRVTINS
185	pep395	SSWKRLITPDGSMLFRVTINS
186	pep396	SSRGSMLFRVTINSWK
187	pep397	SKRGSMLFRVTINSWS
188	pep398	SWRGSMLFRVTINS
189	pep400	SSRGSMLFRVTIWK
190	pep401	SSWKRGSMLYRVTINS
191	pep402	SSWKRGSMLWRVTINS
192	pep403	SSWKRGSMLHRVTINS
193	pep404	SSWKRGSLLFRVTINS
194	pep405	SSWKRGSKLFRVTINS
195	pep406	SSWKRGSRLFRVTINS
196	pep407	SSWKRGSFLFRVTINS
197	pep408	SSWKRGSWLFRVTINS
198	pep409	SSWKRGSMLFRVSINS
199	pep410	SSWKRGSMLFRVQINS
200	pep411	SSWKRGSMLFRVNINS
201	SmTrip9-286 with cysteine	SSWKRGSMLFRVTINSC
202	HiBit with cysteine	CVSGWRLFKKIS
203	SmTrip9-286 with azide	SSWKRGSMLFRVTINSK(Aza)
204	HiBit with azide	(aza)KVSGWRLFKKIS
205	WT OgLuc dipeptide	GSLLFRVTINGVTGWRLCENILA
206	WT NanoLuc dipeptide	GSLLFRVTINVGVTGWRLCERILA
207	pep157	SVSGWRLFKKIS
208	pep158	NSVSGWRLFKKIS

209	pep206	GWRLFKKIS
210	HiBiT-His-LgTrip3546 (ATG 3745)	Atggtgagcggctggcggctgttcaagaagattagccaccatcaccatcaccatcatcac ttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctggaccaa gtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgat catgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatccc gtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgta ccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggg gttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgac ggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgacgagc gcctgatcacccccgactaa
211	HiBiT-His-LgTrip3546 (ATG 3745)	MVSGWRLFKKISHHHHHHHHFTLDDFVGDWEQTAAYN LDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIH VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTL VIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNK IIDERLITPD
212	His-HiBiT-GSSG-LgTrip3546 (ATG 3746)	Atgaaacatcaccatcaccatcatgtgagcggctggcggctgttcaagaagattagcgg cagctccggtttcacactcgacgatttcgttggggactgggaacagacagccgcctacaa cctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtcc gtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccat gtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgttt aaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggt aatcgacggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcg ccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaatt atcgacgagcgcctgatcacccccgactaa
213	His-HiBiT-GSSG-LgTrip3546 (ATG 3746)	MKHHHHHHVSGWRLFKKISGSSGFTLDDFVGDWEQTA AYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALK IDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILP YGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLW NGNKIIDERLITPD
214	FRB-15GS-86, no AI in linker (ATG3768)	Atggtggccatcctctggcatgagatgtggcatgaaggcctggaagaggcatctcgtttg tactttggggaaaggaacgtgaaaggcatgtttgaggtgctggagcccttgcatgctatga tggaacggggcccccagactctgaaggaaacatcctttaatcaggcctatggtcgagattt aatggaggcccaagagtggtgcaggaagtacatgaaatcagggaatgtcaaggacctc acccaagcctgggacctctattatcatgtgttccgacgaatcagtggtggttcaggtggtg gcgggagcggtggctcgagcagcggtggagtgagcggctggcggctgttcaagaaga ttagctaa
215	FRB-15GS-86, no AI in linker (ATG3768)	MVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPL HAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMK SGNVKDLTQAWDLYYHVFRRISGGSGGGGSGGSSSGGV SGWRLFKKIS
216	FRB-15GS-289 (ATG3769)	Atggtggccatcctctggcatgagatgtggcatgaaggcctggaagaggcatctcgtttg tactttggggaaaggaacgtgaaaggcatgtttgaggtgctggagcccttgcatgctatga tggaacggggcccccagactctgaaggaaacatcctttaatcaggcctatggtcgagattt aatggaggcccaagagtggtgcaggaagtacatgaaatcagggaatgtcaaggacctc acccaagcctgggacctctattatcatgtgttccgacgaatcagtggtggttcaggtggtg gcgggagcggtggctcgagcagcggtggagttagcgttagcggctggcgcctgttcaa gaagatcagctaa

217	FRB-15GS-289 (ATG3769)	MVAILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPL HAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMK SGNVKDLTQAWDLYYHVFRRISGGSGGGGSGGSSSGGV SVSGWRLFKKIS
218	FKBP-SmTrip9 fusion template, no AI in linker (ATG3770)	atgggagtgcaggtggaaaccatctccccaggagacgggcgcaccttccccaagcgcg gccagacctgcgtggtgcactacaccgggatgcttgaagatggaaagaaatttgattcctc ccgggacagaaacaagccctttaagtttatgctaggcaagcaggaggtgatccgaggct gggaagaaggggttgcccagatgagtgtgggtcagagagccaaactgactatatctcca gattatgcctatggtgccactgggcacccaggcatcatcccaccacatgccactctcgtctt cgatgtggagcttctaaaactggaaggtggttcaggtggtggcgggagcggtggctcga gcagcggtgga
219	FKBP-SmTrip9 fusion template, no AI in linker (ATG3770)	MGVQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKFD SSRDRNKPFKFMLGKQEVIRGWEEGVAQMSVGQRAKLT ISPDYAYGATGHPGIIPPHATLVFDVELLKLEGGSGGGGS GGSSSGG
220	295	GSMLFRVTINSV
221	300	GSMLFRVTINSVS
222	412	MLFRVTINSVSG
223	413	MLFRVTINSVSGW
224	415	MLFRVTINSVSGWK
225	416	MLFRVTINSVSGWR
226	418	GSMLFRVTINSVSG
227	419	GSMLFRVTINSVSGW
228	422	GSMLFRVTINSVSGWR
229	423	GSMLFRVTINSVSGWK
230	434	GSMLFRVTIWK
231	435	GSMLFRVTINSWK
232	477	MLFRVTINSWK
233	478	MLFRVTINSWS
234	479	MLFRVTIWS
235	480	MLFRVTIWK
236	481	MLFRVKINS
237	482	GSMLFRVTINSWS
238	483	GSMLFRVKINS
239	484	GSMLFRVTIWS
240	485	MLFRVNINS
241	486	MLFRVWINS
242	487	LLFRVKINS
243	488	FLFRVTINS
244	295	SSWKRGSMLFRVTINSV
245	300	SSWKRGSMLFRVTINSVS

246	412	SSWKRMLFRVTINSVSG
247	413	SSWKRMLFRVTINSVSGW
248	414	SSWKRMLFRVTINSVSGWR
249	415	SSWKRMLFRVTINSVSGWK
250	417	MLFRVTINSVSGWK
251	418	SSWKRGSMLFRVTINSVSG
252	419	SSWKRGSMLFRVTINSVSGW
253	420	SSWKRGSMLFRVTINSVSGWR
254	421	SSWKRGSMLFRVTINSVSGWK
255	424	SSWKRGSYLFRVTINS
256	425	SSWKRGSMLFRVKINS
257	426	SSWKRGSMLFRVRINS
258	427	SSWKRGSMLFRVWINS
259	428	SSKRGSMLFRVTIWSV
260	429	SSKRGSMLFRVTIWSVS
261	430	SSWRGSMLFRVTIKS
262	431	KRSSGSMLFRVTIWS
263	432	SSKRMLFRVTIWS
264	433	KRSSMLFRVTIWS
265	445	GSMKFRVTINSWK
266	450	GSMLFRKTINSWK
267	455	GSMLFRVTKNSWK
268	521	GKMLFRVTINSWK
269	522	GKMLFRVTIWK
270	523	GSMKFRVTINSWK
271	524	GSMKFRVTIWK
272	525	GRMLFRVTINSWK
273	526	GRMLFRVTIWK
274	527	GSMRFRVTINSWK
275	528	GSMRFRVTIWK
276	529	GDMLFRVTINSWK
277	530	GDMLFRVTIWK
278	531	GSMDFRVTINSWK
279	532	GSMDFRVTIWK
280	533	GEMLFRVTINSWK
281	535	GSMEFRVTINSWK
282	536	GSMEFRVTIWK
283	538	GSMLFRVTIWKVK
284	539	GSMLFRVTIWSVK
285	540	GSMLFRVTIWSK
286	541	GSMLFRVTIWKWK

287	542	GSMLFRVTIWKK
288	245	GSMLFRVTINS
289	292.x	MLFRVTINS
290	297.x	MLFVTINSV
291	302.x	MLFRVTINSVS
292	305.x	GSMLFRVTIN
293	306.x	GSMLFRVTI
294	307.x	SMLFRVTIN
295	308.x	MLFRVTIN
296	312.x	MLFRVTI
297	399	SSKRGSMLFRVTIWS
298	273	GSMLFRVTINSGVSGWALFKKIS
299	264	GSMLFRVTINSGVSGWRLFKKIS
300	167	VSGWALFKKIS
301	331	GSMLFRVTINSVSGVSGWRLFKKIS
302	LgTrip 3546 (no His6)	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITTTGTLWNGNKIIDERLITPD
303	LgTrip 3546 (no His6)	atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgac
304	LgTrip 2098 (no His6)	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITVTGTLWNGNKIIDERLITPD
305	LgTrip 2098 (no His6)	atggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgac
306	157	SVSGWRLFKKIS
307	158	NSVSGWRLFKKIS
308	206	GWRLFKKIS
309	264	GSMLFRVTINSGVSGWRLFKKIS
310	489	GSMLFRVTINSWK (N-term unblocked)

311	490	GSMLFRVTINSWK (C-term unblocked)
312	491	GSMLFRVTINSWK (Both unblocked)
313	492	GSMLFRVTINKWK
314	493	GSMLFRVTIKSWK
315	494	GSMLFRVTINRWK
316	495	GSMLFRVTIRSWK
317	496	GSMLFRVTINDWK
318	497	GSMLFRVTIDSWK
319	498	GSMLFRVTINEWK
320	499	GSMLFRVTIESWK
321	465	GSMRFRVTINSWK (Both termini unblocked)
322	466	GSMDFRVTINSWK (Both termini unblocked)
323	467	GSMEFRVTINSWK (Both termini unblocked)
324	468	GSMLFRRTINSWK (Both termini unblocked)
325	469	GSMLFRDTINSWK (Both termini unblocked)
326	470	GSMLFRETINSWK (Both termini unblocked)
327	472	GSMLFRVTDNSWK (Both termini unblocked)
328	473	GSMLFRVTENSWK (Both termini unblocked)
329	474	GSMKFRVTINSWK (Both termini unblocked)
330	475	GSMLFRKTINSWK (Both termini unblocked)
331	476	GSMLFRVTKNSWK (Both termini unblocked)
332	436	GSMLFRVTINS (N-term unblocked)
333	437	GSMLFRVSINS (N-term unblocked)
334	438	GSMLFRVNINS (N-term unblocked)
335	439	GSKLFRVTINS (N-term unblocked)
336	440	GSRLFRVTINS (N-term unblocked)
337	441	GSMWFRVTINS (N-term unblocked)
338	442	GSMSFRVTINS (N-term unblocked)
339	443	GSMNFRVTINS (N-term unblocked)
340	444	GSMKFRVTINS (N-term unblocked)
341	446	GSMLFRWTINS (N-term unblocked)
342	447	GSMLFRSTINS (N-term unblocked)
343	448	GSMLFRNTINS (N-term unblocked)
344	449	GSMLFRKTINS (N-term unblocked)
345	451	GSMLFRVTWNS (N-term unblocked)
346	452	GSMLFRVTSNS (N-term unblocked)
347	453	GSMLFRVTNNS (N-term unblocked)
348	454	GSMLFRVTKNS (N-term unblocked)
349	456	GSMLFRVTIKS (N-term unblocked)

350	Antares ATG 3802	MKHHHHHHVSKGEELIKENMRSKLYLEGSVNGHQFKCT HEGEGKPYEGKQTNRIKVVEGGPLPFAFDILATHFMYGS KVFIKYPADLPDYFKQSFPEGFTWERVMVFEDGGVLTAT QDTSLQDGELIYNVKVRGVNFPANGPVMQKKTLGWEPS TETMYPADGGLEGRCDKALKLVGGGHLHVNFKTTYKS KKPVKMPGVHYVDRRLERIKEADNETYVEQYEHAVAR YSNLGGGFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLF QNLGVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSGDQMG QIEKIFKVVYPVDDHHFKVILHYGTLVIDGVTPNMIDYFG RPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLF RVTINGVTGWRLCERILARHELIKENMRSKLYLEGSVNG HQFKCTHEGEGKPYEGKQTNRIKVVEGGPLPFAFDILAT HFMYGSKVFIKYPADLPDYFKQSFPEGFTWERVMVFED GGVLTATQDTSLQDGELIYNVKVRGVNFPANGPVMQKK TLGWEPSTETMYPADGGLEGRCDKALKLVGGGHLHVN FKTTYKSKKPVKMPGVHYVDRRLERIKEADNETYVEQY EHAVARYSNLGGGMDELYK
351	Antares ATG 3802	atgaaacatcaccatcaccatcatgtgagcaagggagaagaacttataaaagaaaacatg cggtctaaactgtacctcgagggctccgtcaatgggcaccagtttaagtgtacccacgag ggtgagggaaagccctatgaggggaagcagacaaaccgcatcaaggtcgtcgaaggg ggacccctcccgtttgcctttgatatcttggctactcactttatgtacggaagcaaagttttcat aaagtatcctgccgaccttcctgattattttaaacagtcatttcccgagggtttcacatggga aagggtcatggtgtttgaggatggaggcgtgctcactgcaactcaggacacctcactgca ggacggcgagctgatctacaatgtgaaggtccggggtgtaaacttccctgccaacgggc ctgtaatgcagaagaagaccctgggatgggagccgtccaccgaaaccatgtaccctgct gatggtgggctggagggccgatgtgacaaggctctgaagctcgttggaggtggtcatttg cacgtaaatttcaagacaacttacaagagcaaaaaacccgtaaaaatgcccggggttcatt acgttgacagaaggcttgaacgcataaaggaagctgataacgagacatacgtggagcag tacgagcacgccgttgcccggtactcaaacctgggggtggctttacactggaggattttg tgggagattggagacagacagccggctacaatctggatcaggtgctggaacaaggagg agtgtcttctctgtttcagaatctgggagtgagcgtgacacctatccagaggatcgtgctgt ctggcgagaatggactgaagatcgatattcacgtgatcatcccctacgaaggcctgtctgg agaccagatgggccagattgagaagatcttcaaagtggtgtatcctgtggacgatcacca cttcaaggtgatcctgcactacggcaccctggtgattgatggagtgacacctaacatgatc gactacttcggaagaccttacgagggaatcgccgtgttcgacggaaagaagatcaccgt gacaggaacactgtggaatggaaacaagatcatcgacgagcggctgatcaaccctgatg gatctctgctgttcagagtgaccatcaacggagtgacaggatggagactgtgcgagaga attctggctagacatgagctaatcaaggaaaatatgagaagtaagctatacttagaggggt ccgtcaacggtcaccagtttaaatgcactcatgaaggtgaggggaaaccttatgaaggta agcagactaatcgaataaaagtggtcgagggcggtcctctgccattcgctttcgatattctg gccactcactttatgtatgggtctaaggtctttattaaataccccgctgatttgccagactactt taaacagtccttccctgaaggattcacatgggagcgggtgatggtgttcgaggatggagg cgttcttactgcaactcaggatacttccttgcaagacggggaactgatctacaacgttaagg tccgcggcgtcaatttcccagccaatggtccagtgatgcagaagaaaaccttggggtggg agccctcaacggagacaatgtaccctgcggacggcgggcttgagggtagatgtgataag gcattgaaactcgtcgggggcggccaccttcatgtgaatttcaagactacatataaaagta aaaaaccagtcaagatgcctggagtgcactacgtggatcgtaggttggagaggataaaa

		gaagccgacaacgaaacttatgtagagcaatatgagcacgccgtggctcgttattccaact tgggcggaggaatggatgaactgtacaag
352	Antares (LgBiT) ATG 3803	MKHHHHHHVSKGEELIKENMRSKLYLEGSVNGHQFKCT HEGEGKPYEGKQTNRIKVVEGGPLPFAFDILATHFMYGS KVFIKYPADLPDYFKQSFPEGFTWERVMVFEDGGVLTAT QDTSLQDGELIYNVKVRGVNFPANGPVMQKKTLGWEPS TETMYPADGGLEGRCDKALKLVGGGHLHVNFKTTYKS KKPVKMPGVHYVDRRLERIKEADNETYVEQYEHAVAR YSNLGGGFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLL QNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMA QIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYF GRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSML FRVTINSRHELIKENMRSKLYLEGSVNGHQFKCTHEGEG KPYEGKQTNRIKVVEGGPLPFAFDILATHFMYGSKVFIK YPADLPDYFKQSFPEGFTWERVMVFEDGGVLTATQDTS LQDGELIYNVKVRGVNFPANGPVMQKKTLGWEPSTETM YPADGGLEGRCDKALKLVGGGHLHVNFKTTYKSKKPV KMPGVHYVDRRLERIKEADNETYVEQYEHAVARYSNL GGGMDELYK

353	Antares (LgBiT) ATG 3803	atgaaacatcaccatcaccatcatgtgagcaagggagaagaacttataaaagaaaacatg cggtctaaactgtacctcgagggctccgtcaatgggcaccagtttaagtgtacccacgag ggtgagggaaagccctatgaggggaagcagacaaaccgcatcaaggtcgtcgaaggg ggacccctcccgtttgcctttgatatcttggctactcactttatgtacggaagcaaagttttcat aaagtatcctgccgaccttcctgattattttaaacagtcatttcccgagggtttcacatggga aagggtcatggtgtttgaggatggaggcgtgctcactgcaactcaggacacctcactgca ggacggcgagctgatctacaatgtgaaggtccggggtgtaaacttccctgccaacgggc ctgtaatgcagaagaagaccctgggatgggagccgtccaccgaaaccatgtaccctgct gatggtgggctggagggccgatgtgacaaggctctgaagctcgttggaggtggtcatttg cacgtaaatttcaagacaacttacaagagcaaaaaacccgtaaaaatgcccggggttcatt acgttgacagaaggcttgaacgcataaaggaagctgataacgagacatacgtggagcag tacgagcacgccgttgcccggtactcaaacctgggggggtggcttcacactcgaagatttc gttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggag gtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagc gccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcat cactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacatgc tgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgt aacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgac ggctccatgctgttccgagtaaccatcaacagcagacatgagctaatcaaggaaaatatg agaagtaagctatacttagaggggtccgtcaacggtcaccagtttaaatgcactcatgaag gtgaggggaaaccttatgaaggtaagcagactaatcgaataaaagtggtcgagggcggt cctctgccattcgctttcgatattctggccactcactttatgtatgggtctaaggtctttattaaa taccccgctgatttgccagactactttaaacagtccttccctgaaggattcacatgggagcg ggtgatggtgttcgaggatggaggcgttcttactgcaactcaggatacttccttgcaagac ggggaactgatctacaacgttaaggtccgcggcgtcaatttcccagccaatggtccagtg atgcagaagaaaaccttggggtgggagccctcaacggagacaatgtaccctgcggacg gcgggcttgagggtagatgtgataaggcattgaaactcgtcgggggcggccaccttcat gtgaatttcaagactacatataaaagtaaaaaaccagtcaagatgcctggagtgcactacg tggatcgtaggttggagaggataaaagaagccgacaacgaaacttatgtagagcaatatg agcacgccgtggctcgttattccaacttgggcggaggaatggatgaactgtacaag
354	Antares (LgTrip 3546) ATG 3804	MKHHHHHHVSKGEELIKENMRSKLYLEGSVNGHQFKCT HEGEGKPYEGKQTNRIKVVEGGPLPFAFDILATHFMYGS KVFIKYPADLPDYFKQSFPEGFTWERVMVFEDGGVLTAT QDTSLQDGELIYNVKVRGVNFPANGPVMQKKTLGWEPS TETMYPADGGLEGRCDKALKLVGGGHLHVNFKTTYKS KKPVKMPGVHYVDRRLERIKEADNETYVEQYEHAVAR YSNLGGGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLL QNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMA QIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYF GRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDRHELI KENMRSKLYLEGSVNGHQFKCTHEGEGKPYEGKQTNRI KVVEGGPLPFAFDILATHFMYGSKVFIKYPADLPDYFKQ SFPEGFTWERVMVFEDGGVLTATQDTSLQDGELIYNVK VRGVNFPANGPVMQKKTLGWEPSTETMYPADGGLEGR CDKALKLVGGGHLHVNFKTTYKSKKPVKMPGVHYVDR RLERIKEADNETYVEQYEHAVARYSNLGGGMDELYK

355	Antares (LgTrip 3546) ATG 3804	atgaaacatcaccatcaccatcatgtgagcaagggagaagaacttataaaagaaaacatg cggtctaaactgtacctcgagggctccgtcaatgggcaccagtttaagtgtacccacgag ggtgagggaaagccctatgaggggaagcagacaaaccgcatcaaggtcgtcgaaggg ggacccctcccgtttgcctttgatatcttggctactcactttatgtacggaagcaaagttttcat aaagtatcctgccgaccttcctgattattttaaacagtcatttcccgagggtttcacatggga aagggtcatggtgtttgaggatggaggcgtgctcactgcaactcaggacacctcactgca ggacggcgagctgatctacaatgtgaaggtccggggtgtaaacttccctgccaacgggc ctgtaatgcagaagaagaccctgggatgggagccgtccaccgaaaccatgtaccctgct gatggtgggctggagggccgatgtgacaaggctctgaagctcgttggaggtggtcatttg cacgtaaatttcaagacaacttacaagagcaaaaaacccgtaaaaatgcccggggttcatt acgttgacagaaggcttgaacgcataaaggaagctgataacgagacatacgtggagcag tacgagcacgccgttgcccggtactcaaacctgggggggtggcttcacactcgacgatttc gttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggag gtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccgg agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagc gccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcat cactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaag ctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcact accacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccg acagacatgagctaatcaaggaaaatatgagaagtaagctatacttagaggggtccgtca acggtcaccagtttaaatgcactcatgaaggtgaggggaaaccttatgaaggtaagcaga ctaatcgaataaaagtggtcgagggcggtcctctgccattcgctttcgatattctggccact cactttatgtatgggtctaaggtctttattaaatacccccgctgatttgccagactactttaaaca gtccttccctgaaggattcacatgggagcgggtgatggtgttcgaggatggaggcgttctt actgcaactcaggatacttccttgcaagacggggaactgatctacaacgttaaggtccgc ggcgtcaatttcccagccaatggtccagtgatgcagaagaaaaccttggggtgggagcc ctcaacggagacaatgtaccctgcggacggcgggcttgagggtagatgtgataaggcat tgaaactcgtcgggggcggccaccttcatgtgaatttcaagactacatataaaagtaaaaa accagtcaagatgcctggagtgcactacgtggatcgtaggttggagaggataaaagaag ccgacaacgaaacttatgtagagcaatatgagcacgccgtggctcgttattccaacttggg cggaggaatggatgaactgtacaag
356	ATG 3815	MKHHHHHHFTLEDFVGDWEQTAAYNLDQVLEQGGVSS LLQNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQM AQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNY FGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSM LFRVTINSGGSGGSSGELIKENMRSKLYLEGSVNGHQFK CTHEGEGKPYEGKQTNRIKVVEGGPLPFAFDILATHFMY GSKVFIKYPADLPDYFKQSFPEGFTWERVMVFEDGGVLT ATQDTSLQDGELIYNVKVRGVNFPANGPVMQKKTLGW EPSTETMYPADGGLEGRCDKALKLVGGGHLHVNFKTTY KSKKPVKMPGVHYVDRRLERIKEADNETYVEQYEHAV ARYSNLGGGMDELYK

357	ATG 3815	atgaaacatcaccatcaccatcatttcacactcgaagatttcgttggggactgggaacaga cagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcaga atctcgccgtgtccgtaactccgatccaaaggattgtccggagcggtgaaaatgccctga agatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccaga tcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccct atggcacactggtaatcgacggggttacgccgaacatgctgaactatttcggacggccgt atgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaac ggcaacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagta accatcaacagcggaggctcaggtggatcctcaggtgagctaatcaaggaaaatatgag aagtaagctatacttagaggggtccgtcaacggtcaccagtttaaatgcactcatgaaggt gaggggaaaccttatgaaggtaagcagactaatcgaataaaagtggtcgagggcggtcc tctgccattcgctttcgatattctggccactcactttatgtatgggtctaaggtctttattaaata ccccgctgatttgccagactactttaaacagtccttccctgaaggattcacatgggagcgg gtgatggtgttcgaggatggaggcgttcttactgcaactcaggatacttccttgcaagacg gggaactgatctacaacgttaaggtccgcggcgtcaatttcccagccaatggtccagtgat gcagaagaaaaccttggggtgggagccctcaacggagacaatgtaccctgcggacggc gggcttgagggtagatgtgataaggcattgaaactcgtcgggggcggccaccttcatgtg aatttcaagactacatataaaagtaaaaaaccagtcaagatgcctggagtgcactacgtgg atcgtaggttggagaggataaaagaagccgacaacgaaacttatgtagagcaatatgag cacgccgtggctcgttattccaacttgggcggaggaatggatgaactgtacaag
358	ATG 3816	MKHHHHHHFTLEDFVGDWEQTAAYNLDQVLEQGGVSS LLQNLAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQM AQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNY FGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSM LFRVTINSRHELIKENMRSKLYLEGSVNGHQFKCTHEGE GKPYEGKQTNRIKVVEGGPLPFAFDILATHFMYGSKVFI KYPADLPDYFKQSFPEGFTWERVMVFEDGGVLTATQDT SLQDGELIYNVKVRGVNFPANGPVMQKKTLGWEPSTET MYPADGGLEGRCDKALKLVGGGHLHVNFKTTYKSKKP VKMPGVHYVDRRLERIKEADNETYVEQYEHAVARYSN LGGGMDELYK
359	ATG 3816	Atgaaacatcaccatcaccatcatttcacactcgaagatttcgttggggactgggaacaga cagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcaga atctcgccgtgtccgtaactccgatccaaaggattgtccggagcggtgaaaatgccctga agatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccaga tcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccct atggcacactggtaatcgacggggttacgccgaacatgctgaactatttcggacggccgt atgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaac ggcaacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagta accatcaacagcagacatgagctaatcaaggaaaatatgagaagtaagctatacttagag gggtccgtcaacggtcaccagtttaaatgcactcatgaaggtgaggggaaaccttatgaa ggtaagcagactaatcgaataaaagtggtcgagggcggtcctctgccattcgctttcgata ttctggccactcactttatgtatgggtctaaggtctttattaaataccccgctgatttgccagac tactttaaacagtccttccctgaaggattcacatgggagcgggtgatggtgttcgaggatg gaggcgttcttactgcaactcaggatacttccttgcaagacggggaactgatctacaacgtt aaggtccgcggcgtcaatttcccagccaatggtccagtgatgcagaagaaaaccttggg gtgggagccctcaacggagacaatgtaccctgcggacggcgggcttgagggtagatgt gataaggcattgaaactcgtcgggggcggccaccttcatgtgaatttcaagactacatata

		aaagtaaaaaaccagtcaagatgcctggagtgcactacgtggatcgtaggttggagagg ataaaagaagccgacaacgaaacttatgtagagcaatatgagcacgccgtggctcgttatt ccaacttgggcggaggaatggatgaactgtacaag
360	ATG 3817	MKHHHHHHFTLDDFVGDWEQTAAYNLDQVLEQGGVSS LLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQ MAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKLN YFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDGG SGGSSGELIKENMRSKLYLEGSVNGHQFKCTHEGEGKPY EGKQTNRIKVVEGGPLPFAFDILATHFMYGSKVFIKYPA DLPDYFKQSFPEGFTWERVMVFEDGGVLTATQDTSLQD GELIYNVKVRGVNFPANGPVMQKKTLGWEPSTETMYPA DGGLEGRCDKALKLVGGGHLHVNFKTTYKSKKPVKMP GVHYVDRRLERIKEADNETYVEQYEHAVARYSNLGGG MDELYK
361	ATG 3817	Atgaaacatcaccatcaccatcatttcacactcgacgatttcgttggggactgggaacaga cagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcaga atctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaa gatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagat cgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccct atggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggccgt atgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaac ggcaacaaaattatcgacgagcgcctgatcacccccgacggaggctcaggtggatcctc aggtgagctaatcaaggaaaatatgagaagtaagctatacttagaggggtccgtcaacgg tcaccagtttaaatgcactcatgaaggtgaggggaaaccttatgaaggtaagcagactaat cgaataaaagtggtcgagggcggtcctctgccattcgctttcgatattctggccactcacttt atgtatgggtctaaggtctttattaaatacccccgctgatttgccagactactttaaacagtcctt ccctgaaggattcacatgggagcgggtgatggtgttcgaggatggaggcgttcttactgc aactcaggatacttccttgcaagacggggaactgatctacaacgttaaggtccgcggcgt caatttcccagccaatggtccagtgatgcagaagaaaaccttggggtgggagccctcaac ggagacaatgtaccctgcggacggcgggcttgagggtagatgtgataaggcattgaaac tcgtcgggggcggccaccttcatgtgaatttcaagactacatataaaagtaaaaaaccagt caagatgcctggagtgcactacgtggatcgtaggttggagaggataaaagaagccgaca acgaaacttatgtagagcaatatgagcacgccgtggctcgttattccaacttgggcggag gaatggatgaactgtacaag
362	ATG 3818	MKHHHHHHFTLDDFVGDWEQTAAYNLDQVLEQGGVSS LLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQ MAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKLN YFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDRH ELIKENMRSKLYLEGSVNGHQFKCTHEGEGKPYEGKQT NRIKVVEGGPLPFAFDILATHFMYGSKVFIKYPADLPDYF KQSFPEGFTWERVMVFEDGGVLTATQDTSLQDGELIYN VKVRGVNFPANGPVMQKKTLGWEPSTETMYPADGGLE GRCDKALKLVGGGHLHVNFKTTYKSKKPVKMPGVHYV DRRLERIKEADNETYVEQYEHAVARYSNLGGGMDELYK

363	ATG 3818	Atgaaacatcaccatcaccatcatttcacactcgacgatttcgttggggactgggaacaga cagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcaga atctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaa gatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagat cgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccct atggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggccgt atgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaac ggcaacaaaattatcgacgagcgcctgatcacccccgacagacatgagctaatcaagga aaatatgagaagtaagctatacttagaggggtccgtcaacggtcaccagtttaaatgcact catgaaggtgaggggaaaccttatgaaggtaagcagactaatcgaataaaagtggtcga gggcggtcctctgccattcgctttcgatattctggccactcactttatgtatgggtctaaggtc tttattaaataccccgctgatttgccagactactttaaacagtccttccctgaaggattcacat gggagcgggtgatggtgttcgaggatggaggcgttcttactgcaactcaggatacttcctt gcaagacggggaactgatctacaacgttaaggtccgcggcgtcaatttcccagccaatg gtccagtgatgcagaagaaaaccttggggtgggagccctcaacggagacaatgtaccct gcggacggcgggcttgagggtagatgtgataaggcattgaaactcgtcgggggcggcc accttcatgtgaatttcaagactacatataaaagtaaaaaaccagtcaagatgcctggagtg cactacgtggatcgtaggttggagaggataaaagaagccgacaacgaaacttatgtaga gcaatatgagcacgccgtggctcgttattccaacttgggcggaggaatggatgaactgta caag
364	LgTrip 2899 (LgTrip 2098+Q42L)	MKHHHHHHVFTLEDFVGDWEQTAAYNLDQVLEQGGVS SLLQNLAVSVTPILRIVRSGENALKIDIHVIIPYEGLSADQ MAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLN YFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPD
365	LgTrip 2899 (LgTrip 2098+Q42L)	atgaaacatcaccatcaccatcatgtcttcacactcgaagatttcgttggggactgggaaca gaccgccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcctaaggattgtccggagcggtgaaaatgccctg aagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccag atcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgcc ctatggcacactggtaatcgacggggttacgccgaacatgctgaactatttcggacggcc gtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtgga acggcaacaaaattatcgacgagcgcctgatcacccccgac
820	ATG-3930	atgAAACATCACCATCACCATCATgtcTTCACACTCGACG ATTTCGTTGGGGACTGGGAACAGACAGCCGCCTACAA CCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGT TTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATCAT GAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGATC GACATCCATGTCATCATCCCGTATGAAGGTCTGAGCG CCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGGT GGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCC TGCCCTATGGCACACTGGTAATCGACGGGGTTACGCC GAACAAGCTGAACTATTTCGGACGGCCGTATGAAGGC ATCGCCGTGTTCGACGGCTAA

821	ATG-3930	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDG
822	SmTrip9-15GS-ProteinG (ATG 4002)	gggagctccGGTGGTGGCGGGAGCGGAGGTGGAGGctcgAG CGGTATGACGTATAAGTTAATCCTTAATGGTAAAACA TTGAAAGGCGAGACAACTACTGAAGCTGTTGATGCTG CTACTGCAGAAAAAGTCTTCAAACAATACGCTAACGA CAACGGTGTTGACGGTGAATGGACTTACGACGATGCG ACGAAAACCTTTACGGTCACCGAAAAACCAGAAGTGA TCGATGCGTCTGAATTAACACCAGCCGTGACAACTTA CAAACTTGTTATTAATGGTAAAACATTGAAAGGCGAA ACAACTACTGAGGCTGTTGATGCTGCTACTGCAGAGA AGGTGTTCAAACAATATGCGAATGACAACGGTGTTGA CGGTGAGTGGACTTACGACGATGCGACTAAGACCTTT ACAGTTACTGAAAAACCAGAAGTGATCGATGCGTCTG AGTTAACACCAGCCGTGACAACTTACAAACTTGTTATT AATGGTAAAACATTGAAAGGCGAAACAACTACTAAA GCAGTAGACGCAGAAACTGCGGAGAAGGCCTTCAAA CAATACGCTAACGACAACGGTGTTGATGGTGTTTGGA CTTATGATGATGCCACAAAAACCTTTACGGTAACTGA GCATCATCACCATCACCACTAA
823	SmTrip9-15GS-ProteinG (ATG 4002)	GSSGGGGSGGGGSSGMTYKLILNGKTLKGETTTEAVDA ATAEKVFKQYANDNGVDGEWTYDDATKTFTVTEKPEVI DASELTPAVTTYKLVINGKTLKGETTTEAVDAATAEKVF KQYANDNGVDGEWTYDDATKTFTVTEKPEVIDASELTP AVTTYKLVINGKTLKGETTTKAVDAETAEKAFKQYAND NGVDGVWTYDDATKTFTVTEHHHHHH
830	ATG-3929	atgAAACATCACCATCACCATCATgtcTTCACACTCGACG ATTTCGTTGGGGACTGGGAACAGACAGCCGCCTACAA CCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGT TTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATCAT GAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGATC GACATCCATGTCATCATCCCGTATGAAGGTCTGAGCG CCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGGT GGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCC TGCCCTATGGCACACTGGTAATCGACGGGGTTACGCC GAACAAGCTGAACTATTTCGGATAA

831	ATG-3929	Mkhhhhhhvftlddfvgdweqtaaynldqvleqggvssllqnlavsvtpimrivrsg enalkidihviipyeglsadqmaqieevfkvvypvddhhfkvilpygtlvidgvtpnk lnyfg
832	ATG-3930	atgAAACATCACCATCACCATCATgtcTTCACACTCGACG ATTTCGTTGGGGACTGGGAACAGACAGCCGCCTACAA CCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGT TTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATCAT GAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGATC GACATCCATGTCATCATCCCGTATGAAGGTCTGAGCG CCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGGT GGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCC TGCCCTATGGCACACTGGTAATCGACGGGGTTACGCC GAACAAGCTGAACTATTTCGGACGGCCGTATGAAGGC ATCGCCGTGTTCGACGGCTAA
833	ATG-3930	Mkhhhhhhvftlddfvgdweqtaaynldqvleqggvssllqnlavsvtpimrivrsg enalkidihviipyeglsadqmaqieevfkvvypvddhhfkvilpygtlvidgvtpnk lnyfgrpyegiavfdg
834	ATG-3931	atgAAACATCACCATCACCATCATgtcTTCACACTCGACG ATTTCGTTGGGGACTGGGAACAGACAGCCGCCTACAA CCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGT TTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATCAT GAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGATC GACATCCATGTCATCATCCCGTATGAAGGTCTGAGCG CCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGGT GGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCC TGCCCTATGGCACACTGGTAATCGACGGGGTTACGCC GAACAAGCTGAACTATTTCGGACGGCCGTATGAAGGC ATCGCCGTGTTCGACGGCAAAAAGATCACTACCACAG GGACCCTGTAA
835	ATG-3931	Mkhhhhhhvftlddfvgdweqtaaynldqvleqggvssllqnlavsvtpimrivrsg enalkidihviipyeglsadqmaqieevfkvvypvddhhfkvilpygtlvidgvtpnk lnyfgrpyegiavfdgkkitttgtl

836	ATG-3932	atgAAACATCACCATCACCATCATgtcTTCACACTCGACG ATTTCGTTGGGGACTGGGAACAGACAGCCGCCTACAA CCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGT TTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATCAT GAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGATC GACATCCATGTCATCATCCCGTATGAAGGTCTGAGCG CCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGGT GGTGTACCCTGTGGATGATCATCACTTTAAGGTGATCC TGCCCTATGGCACACTGGTAATCGACGGGGTTACGCC GAACAAGCTGAACTATTTCGGACGGCCGTATGAAGGC ATCGCCGTGTTCGACGGCAAAAAGATCACTACCACAG GGACCCTGTGGAACGGCTAA
837	ATG-3932	Mkhhhhhhvftlddfvgdweqtaaynldqvleqggvssllqnlavsvtpimrivrsg enalkidihviipyeglsadqmaqieevfkvvypvddhhfkvilpygtlvidgvtpnk lnyfgrpyegiavfdgkkitttgtlwng
838	ATG-4808	Atggtttccgtgagcggctggcggctgttcaagaagattagcttcacactcgacgatttcg ttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggagg tgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccgga gcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcg ccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatc actttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagct gaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactac cacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgact aa
839	ATG-4808	Mvsvsgwrlfkkisftlddfvgdweqtaaynldqvleqggvssllqnlavsvtpimri vrsgenalkidihviipyeglsadqmaqieevfkvvypvddhhfkvilpygtlvidgv tpnklnyfgrpyegiavfdgkkitttgtlwngnkiiderlitpd
840	ATG-4809	Atggtttccgtgagcggctggcggctgttcaagaagattagcggcagctccggtttcaca ctcgacgatttcgttggggactgggaacagacagccgcctacaacctggaccaagtcctt gaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatga ggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatga aggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgt ggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacg ccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaa aaagatcactaccacagggaccctgtggaacggcaacaaaattatcgacgagcgcctga tcacccccgactaa

841	ATG-4809	MVSVSGWRLFKKISGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPD
842	ATG-4810	Atggtttccgtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctcgagcggtttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgactaa
843	ATG-4810	MVSVSGWRLFKKISGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPD
844	ATG-4811	Atggtttccgtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctcgagcggtggctcgagcggtttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgactaa
845	ATG-4811	MVSVSGWRLFKKISGSSGGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPD
846	ATG-4812	Atggtttccgtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctcgagcggtggctcgagcggtggctcgagcggtttcacactcgacgatttcgttggggactggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcg

		gacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggac cctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgactaa
847	ATG-4812	MVSVSGWRLFKKISGSSGGSSGGSSGGSSGFTLDDFVGD WEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRS GENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDH HFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKIT TTGTLWNGNKIIDERLITPD
848	ATG-4813	Atggtttccgtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctc gagcggtggctcgagcggtggctcgagcggtggctcgagcggtttcacactcgacgatt tcgttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacaggga ggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccg gagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgag cgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatca tcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaag ctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcact accacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccg actaa
849	ATG-4813	MVSVSGWRLFKKISGSSGGSSGGSSGGSSGGSSGFTLDD FVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMR IVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPV DDHHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDG KKITTTGTLWNGNKIIDERLITPD
850	ATG-4814	Atggtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctcgagcg gtggctcgagcggtggctcgagcggtggctcgagcggtttcacactcgacgatttcgttg gggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgt gtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccggagc ggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgcc gaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcact ttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagctga actatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactacca cagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgactaa

851	ATG-4814	MVSGWRLFKKISGSSGGSSGGSSGGSSGGSSGFTLDDFV GDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIV RSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDD HHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKI TTTGTLWNGNKIIDERLITPD
852	ATG-4815	Atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacgtttccgtgagcggctggcggctgttcaagaagattagc taa
853	ATG-4815	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITTTGTLWNGNKIIDERLITPDVSVSGWRLF KKIS
854	ATG-4816	Atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacggctcgagcggtgtttccgtgagcggctggcggctgtt caagaagattagctaa
855	ATG-4816	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITTTGTLWNGNKIIDERLITPDGSSGVSVSG WRLFKKIS
856	ATG-4817	Atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgac

		gagcgcctgatcacccccgacggctcgagcggtggctcgagcggtgtttccgtgagcg gctggcggctgttcaagaagattagctaa
857	ATG-4817	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITTTGTLWNGNKIIDERLITPDGSSGGSSGV SVSGWRLFKKIS
858	ATG-4818	Atggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcat catcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggt ggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacggctcgagcggtggctcgagcggtgtgagcggctgg cggctgttcaagaagattagctaa
859	ATG-4818	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNKLNYFGRPYE GIAVFDGKKITTTGTLWNGNKIIDERLITPDGSSGGSSGV SGWRLFKKIS
860	ATG-4819	Atggtttccgtgagcggctggcggctgttcaagaagattagcttcacactcgacgatttcg ttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggagg tgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccgga gcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcg ccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatc actttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagct gaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactac cacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgac catcaccatcaccatcattaa
861	ATG-4819	MVSVSGWRLFKKISFTLDDFVGDWEQTAAYNLDQVLEQ GGVSSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGL SADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTP NKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLIT PDHHHHHH

862	ATG-4820	Atggtttccgtgagcggctggcggctgttcaagaagattagcggcagctccggtttcaca ctcgacgatttcgttggggactgggaacagacagccgcctacaacctggaccaagtcctt gaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatga ggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatga aggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgt ggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacg ccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaa aaagatcactaccacagggaccctgtggaacggcaacaaaattatcgacgagcgcctga tcacccccgaccatcaccatcaccatcattaa
863	ATG-4820	MVSVSGWRLFKKISGSSGFTLDDFVGDWEQTAAYNLDQ VLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIP YEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVID GVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIID ERLITPDHHHHHH
864	ATG-4821	Atggtttccgtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctc gagcggtttcacactcgacgatttcgttggggactgggaacagacagccgcctacaacct ggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgta actccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtc atcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaag gtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatc gacggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgt gttcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcg acgagcgcctgatcacccccgaccatcaccatcaccatcattaa
865	ATG-4821	MVSVSGWRLFKKISGSSGGSSGFTLDDFVGDWEQTAAY NLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDI HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYG TLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNG NKIIDERLITPDHHHHHH
866	ATG-4822	Atggtttccgtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctc gagcggtggctcgagcggtttcacactcgacgatttcgttggggactgggaacagacag ccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatct cgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccctgaagat cgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcga agaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatgg cacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggccgtatga

		aggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaacggc aacaaaattatcgacgagcgcctgatcacccccgaccatcaccatcaccatcattaa
867	ATG-4822	MVSVSGWRLFKKISGSSGGSSGGSSGFTLDDFVGDWEQ TAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENA LKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVI LPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTL WNGNKIIDERLITPDHHHHHH
868	ATG-4823	Atggtttccgtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctc gagcggtggctcgagcggtggctcgagcggtttcacactcgacgatttcgttggggactg ggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtt tgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaa atgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaat ggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtg atcctgccctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcg gacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggac cctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgaccatcaccatc accatcattaa
869	ATG-4823	MVSVSGWRLFKKISGSSGGSSGGSSGGSSGFTLDDFVGD WEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRS GENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDH HFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKIT TTGTLWNGNKIIDERLITPDHHHHHH
870	ATG-4824	Atggtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctcgagcg gtggctcgagcggtggctcgagcggtggctcgagcggtttcacactcgacgatttcgttg gggactgggaacagacagccgcctacaacctggaccaagtccttgaacagggaggtgt gtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccggagc ggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgcc gaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcact ttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaagctga actatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactacca cagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccgaccat caccatcaccatcattaa

871	ATG-4824	MVSGWRLFKKISGSSGGSSGGSSGGSSGGSSGFTLDDFV GDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRIV RSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDD HHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKI TTTGTLWNGNKIIDERLITPDHHHHHH
872	ATG-4825	Atggtttccgtgagcggctggcggctgttcaagaagattagcggctcgagcggtggctc gagcggtggctcgagcggtggctcgagcggtggctcgagcggtttcacactcgacgatt tcgttggggactgggaacagacagccgcctacaacctggaccaagtccttgaacaggga ggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccg gagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgag cgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggatgatca tcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacaag ctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcact accacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcacccccg accatcaccatcaccatcattaa
873	ATG-4825	MVSVSGWRLFKKISGSSGGSSGGSSGGSSGGSSGFTLDD FVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMR IVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPV DDHHFKVILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDG KKITTTGTLWNGNKIIDERLITPDHHHHHH
874	ATG-4826	Atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaac agacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgc agaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgacgtttccgtgagcggctg gcggctgttcaagaagattagctaa
875	ATG-4826	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDV SVSGWRLFKKIS

876	ATG-4827	Atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaac agacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgc agaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgacggctcgagcggtgtttc cgtgagcggctggcggctgttcaagaagattagctaa
877	ATG-4827	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDG SSGVSVSGWRLFKKIS
878	ATG-4828	Atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaac agacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgc agaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgacggctcgagcggtggct cgagcggtgtgagcggctggcggctgttcaagaagattagctaa
879	ATG-4828	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDG SSGGSSGVSGWRLFKKIS
880	ATG-4829	Atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaac agacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgc agaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgacggggttacgccgaacaagctgaactatttcggacggc cgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtgg aacggcaacaaaattatcgacgagcgcctgatcacccccgacggctcgagcggtggct cgagcggtgtttccgtgagcggctggcggctgttcaagaagattagctaa

881	ATG-4829	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDG SSGGSSGVSVSGWRLFKKIS
882	ATG-2623	atggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcat caccatcaccactaa
883	ATG-2623	MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYE GIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTI NSHHHHHH
884	ATG-3745	atggtgagcggctggcggctgttcaagaagattagccaccatcaccatcaccatcatcact tcacactcgacgatttcgttggggactgggaacagacagccgcctacaacctggaccaa gtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgat catgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatccc gtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgta ccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggg gttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgac ggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattatcgacgagc gcctgatcacccccgactaa
885	ATG-3745	MVSGWRLFKKISHHHHHHHHFTLDDFVGDWEQTAAYN LDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIH VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTL VIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNK IIDERLITPD
886	ATG-3746	atgaaacatcaccatcaccatcatgtgagcggctggcggctgttcaagaagattagcggc agctccggtttcacactcgacgatttcgttggggactgggaacagacagccgcctacaac ctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccg taactccgatcatgaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgt catcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaa ggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaa tcgacggggttacgccgaacaagctgaactatttcggacggccgtatgaaggcatcgcc

		gtgttcgacggcaaaaagatcactaccacagggaccctgtggaacggcaacaaaattat cgacgagcgcctgatcacccccgactaa
887	ATG-3746	MKHHHHHHVSGWRLFKKISGSSGFTLDDFVGDWEQTA AYNLDQVLEQGGVSSLLQNLAVSVTPIMRIVRSGENALK IDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILP YGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLW NGNKIIDERLITPD
888	ATG-4632	atggtgagcggctggcggctgttcaagaagattagcggcagctccggtttcacactcgac gatttcgttggggactgggaacagacagccgcctacaacctggaccaagtccttgaaca gggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggatt gtccggagcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggt ctgagcgccgaccaaatggcccagatcgaagaggtgtttaaggtggtgtaccctgtggat gatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccga acaagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaag atcactaccacagggaccctgtggaacggcaacaaaattatcgacgagcgcctgatcac ccccgaccatcaccatcaccatcattaa
889	ATG-4632	MVSGWRLFKKISGSSGFTLDDFVGDWEQTAAYNLDQVL EQGGVSSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYE GLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGV TPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERL ITPDHHHHHH
890	ATG-2757	atggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacgagaacaaaattatcgac gagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcat caccatcaccactaa
891	ATG-2757	MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYE GIAVFDGKKITVTGTLWNENKIIDERLITPDGSMLFRVTI NSHHHHHH

892	ATG-2760	atggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggattgtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcgttaaaattatcgacg agcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcatc accatcaccactaa
893	ATG-2760	MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEV FKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYE GIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMLFRVTI NSHHHHHH
894	ATG-3882	atggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggatggtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcat caccatcaccactaa
895	ATG-3882	MVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQIEE VFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP YEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRV TINSHHHHHH
896	ATG-3901	atggtcttcacactcgaagatttcgttggggactggaagcagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggatggtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgac gagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcat caccatcaccactaa

897	ATG-3901	MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQIEE VFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP YEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRV TINSHHHHHH
898	ATG-3945	atggtcttcacactcgaagatttcgttggggactggaagcagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggatggtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacgacgtcaaaattatcgac gagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacagccatcat caccatcaccactaa
899	ATG-3945	MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQIEE VFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP YEGIAVFDGKKITVTGTLWNDVKIIDERLITPDGSMLFRV TINSHHHHHH
890	ATG-3984	atggtcttcacactcgaagatttcgttggggactggaagcagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggatggtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacgacgtcaaaattatcgac gagcgcctgatcacccccgacggctccatgtccttccgagtaaccatcaacagccatcat caccatcaccactaa
891	ATG-3984	MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQIEE VFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP YEGIAVFDGKKITVTGTLWNDVKIIDERLITPDGSMSFRV TINSHHHHHH
892	ATG-4147	atggtcttcacactcgaagatttcgttggggactggaagcagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggatggtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaacaaaattatcgac

		gagcgcctgatcacccccgacggctccatgtccttccgagtaaccatcaacagccatcat caccatcaccactaa
893	ATG-4147	MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQIEE VFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP YEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMSFRV TINSHHHHHH
894	ATG-4166	atggtcttcacactcgaagatttcgttggggactggaagcagacagccgcctacaacctg gaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaa ctccgatccaaaggatggtccggagcggtgaaaatgccctgaagatcgacatccatgtca tcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagaggtgtttaagg tggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcg acggggttacgccgaacatgctgaactatttcggacggccgtatgaaggcatcgccgtgt tcgacggcaaaaagatcactgtaacagggaccctgtggaacggcgtcaaaattatcgac gagcgcctgatcacccccgacggctccatgtccttccgagtaaccatcaacagccatcat caccatcaccactaa
895	ATG-4166	MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLA VSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQIEE VFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRV TINSHHHHHH
908	ATG-5037	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCTATGGCACACTGGTAATCGACGGGGTTACGC CGAACAAGCTGAACTATTTCGGACACCCGTATGAAGG CATCGCCGTGTTCGACGGCAAAAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCACCCCCGACTAA

909	ATG-5037	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGHPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPD
910	ATG-5038	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCTATGGCACACTGGTAATCGACGGGGTTACGC CGAACAAGCTGAACTATTTCGGACGGCCGTATGAAGG CATCGCCGTGTTCGACGGCGAGAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCACCCCCGACTAA
911	ATG-5038	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGEKITTTGTLWNGNKIIDERLITPD
912	ATG-5039	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCTATGGCACACTGGTAATCGACGGGGTTACGC CGAACAAGCTGAACTATTTCGGACGGCCGTATGAAGG CATCGCCGTGTTCGACGGCAAAAAGATCACTACCACA GGGACCCTGCCTAACGGCAACAAAATTATCGACGAGC GCCTGATCACCCCCGACTAA

913	ATG-5039	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLPNGNKIIDERLITPD
914	ATG-5040	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCTATGGCACACTGGTAATCGACGGGGTTACGC CGAACAAGCTGAACTATTTCGGACGGCCGTATGAAGG CATCGCCGTGTTCGACGGCAAAAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCGATCCCGACTAA
915	ATG-5040	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLIDPD
916	ATG-5041	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCTATGGCACACTGGTAATCGACGGGGTTACGC CGAACAAGCTGAACTATTTCGGACGGCCGTATGAAGG CATCGCCGTGTTCGACGGCAAAAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCACCGATGACTAA

917	ATG-5041	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITDD
918	ATG-5135	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCTATGGCACACTGGTAATCGACGGGGTTACGC CGAACAAGCTGAACTATTTCGGACACCCGTATGAAGG CATCGCCGTGTTCGACGGCGAGAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCACCCCCGACTAA
919	ATG-5135	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLITPD
920	ATG-5146	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCTATGGCACACTGGTAATCGACGGGGTTACGC CGAACAAGCTGAACTATTTCGGACACCCGTATGAAGG CATCGCCGTGTTCGACGGCGAGAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCGATCCCGACTAA

921	ATG-5146	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLIDPD
922	ATG-5158	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCTATGGCACACTGGTAATCGACGGGGTTACGC CGAACAAGCTGAACTATTTCGGACACCCGTATGAAGG CATCGCCGTGTTCGACGGCGAGAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCGATGATGACTAA
923	ATG-5158	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNKL NYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLIDDD
924	ATG-5260	ATGAAACATCACCATCACCATCATGATTTCACACTCG ACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTA CAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCC AGTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGAT CATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAG ATCGACATCCATGTCATCATCCCGTATGAAGGTCTGA GCGCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAA GGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG ATCCTGCCCATCGGCACACTGGTAATCGACGGGGTTA CGCCGAACAAGCTGAACTATTTCGGACACCCGTATGA AGGCATCGCCGTGTTCGACGGCGAGAAGATCACTACC ACAGGGACCCTGTGGAACGGCAACAAAATTATCGACG AGCGCCTGATCGATCCCGACTAA

925	ATG-5260	MKHHHHHHDFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPIGTLVIDGVTPNKL NYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLIDPD
926	ATG-5266	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCATCGGCACACTGGTAATCGACGGGGAGACGC CGAACAAGCTGAACTATTTCGGACACCCGTATGAAGG CATCGCCGTGTTCGACGGCGAGAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCGATCCCGACTAA
927	ATG-5266	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPIGTLVIDGETPNKL NYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLIDPD
928	ATG-5267	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCATCGGCACACTGGTAATCGACGGGGTTACGC CGAACAAGCTGAACTATTTCGGACACCCGTATGAAGG CATCGCCGATTTCGACGGCGAGAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCGATCCCGACTAA

929	ATG-5267	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPIGTLVIDGVTPNKL NYFGHPYEGIADFDGEKITTTGTLWNGNKIIDERLIDPD
930	ATG-5278	ATGAAACATCACCATCACCATCATGTCTTCACACTCGA CGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCA GTTTGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATC ATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGG TGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC CTGCCCATCGGCACACTGGTAATCGACGGGGAGACGC CGAACAAGCTGAACTATTTCGGACACCCGTATGAAGG CATCGCCGATTTCGACGGCGAGAAGATCACTACCACA GGGACCCTGTGGAACGGCAACAAAATTATCGACGAGC GCCTGATCGATCCCGACTAA
931	ATG-5278	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPIGTLVIDGETPNKL NYFGHPYEGIADFDGEKITTTGTLWNGNKIIDERLIDPD
932	ATG-4794	atgaaacatcaccatcaccatcatgtcttcacactcgacgatttcgttggggactgggaaca gacagccgcctacaacctggaccaagtccttgaacagggaggtgtgtccagtttgctgca gaatctcgccgtgtccgtaactccgatcatgaggattgtccggagcggtgaaaatgccct gaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggccca gatcgaagaggtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgc cctatggcacactggtaatcgac
933	ATG-4794	MKHHHHHHVFTLDDFVGDWEQTAAYNLDQVLEQGGV SSLLQNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSAD QMAQIEEVFKVVYPVDDHHFKVILPYGTLVID

Table 9. Exemplary peptide sequences.

Pep ID	SEQ ID NO.	Sequence
86	824	VSGWRLFKKIS
229	825	VSGWRLFKKI
289	826	VSVSGWRLFKKIS
521	827	GKMLFRVTINSWK
543	366	WNGNKIIDERLITPD
544	367	KKITTTGTLWNGR
545	368	RPYEGIAVFDGK
591	369	GKMLFRVTIWKVSVSGWRLFKKIS
592	370	GKMLFRVTIWKVSGWRLFKKIS
593	371	GSMKFRVTINSWKVSVSGWRLFKKIS
594	372	GSMKFRVTINSWKVSGWRLFKKIS
595	373	GSMKFRVTINSWKNVTGYRLFKKISN
596	374	GSMKFRVTINSWKVTGYRLFEKIS
597	375	GSMKFRVTIWKVSVSGWRLFKKIS
598	376	GSMKFRVTIWKVSGWRLFKKIS
599	377	GRMLFRVTINSWKVSVSGWRLFKKIS
600	378	GRMLFRVTINSWKVSGWRLFKKIS
601	379	GRMLFRVTIWKVSVSGWRLFKKIS
602	380	GRMLFRVTIWKVSGWRLFKKIS
603	381	GSMLFRVTINSVSVSGWRLFKKIS
604	382	GSMLFKVTINSVSGWRLFKKIS
605	383	GSMLFQVTINSVSGWRLFKKIS

606	384	GSMLFEVTINSVSGWRLFKKIS
607	385	GSMLFNVTINSVSGWRLFKKIS
608	386	GRPYEGIAVFDGKKITTTGTL
609	387	GSMKFRVTINSWKVTGYRLFEKES
610	388	GSMKFRVTINSWKVEGYRLFEKIS
611	389	KKITTTGTLWNGNKIIDERLITPD
612	390	WNGNKIIDERLITPDGSMLFRVTINS
671	391	GKMLFRVTIQKWK
668	392	GKMLFRVTIGKWK
727	393	GKMLFRVTIGRWK
669	394	GKMLFRVTIGNWK
674	395	GKMLFRVTIQNWK
702	396	GKMLFRVTIDKWK
703	397	GKMLFRVTIEKWK
705	810	EKMLFRVTIESWK
724	811	EKLLFRVTIESWK
725	812	EKLLFRVTIESYK
730	398	GKMLFRVTIERWK
731	399	GKMLFRVTIDRWK
738	400	DKMLFRVTIQKWK
739	401	DKMLFRVTIGKWK
848	402	DKMLFRVTIGRWK
740	403	DKMLFRVTIGNWK
741	404	DKMLFRVTIQNWK

732	405	DKMLFRVTIDKWK
742	406	DKMLFRVTIEKWK
735	407	DKMLFRVTIERWK
733	408	DKMLFRVTIDRWK
759	816	DKLLFTVTIEKYK
798	409	RPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPD
849	410	EKMLFRVTIQKWK
708	411	EKMLFRVTIGKWK
709	412	EKMLFRVTIGRWK
775	413	DKMLFTVTIQKVSGWRLFKKIS
788	414	DKLLFTVTIEKVSGWRLFKKIS
789	415	DKLLFTVTIEKWKVSGWRLFKKIS
790	416	DKLLFTVTIEKYKVSGWRLFKKIS
792	417	DKLLFTVTIEKYKVSVSGWRLFKKIS
795	418	KKMLFRVTIQKVSGWRLFKKIS
797	419	KKMLFRVTIQKWKVSVSGWRLFKKIS
796	420	KKMLFRVTIQKWKVSGWRLFKKIS
804	421	DKLLFTVTIGKVSGWRLFKKIS
805	422	DKLLFTVTIGKYKVSGWRLFKKIS
806	423	DKLLFTVTIGKYKVSVSGWRLFKKIS
807	424	DKLLFTVTIGKWKVSVSGWRLFKKIS
808	425	DKLLFTVTIQKVSGWRLFKKIS
813	426	KKMLFTVTIQKVSGWRLFKKIS
816	427	KKLLFRVTIQKVSGWRLFKKIS

825	428	DKLLFTVTIEKVSGWRLFKKI
826	429	DKLLFTVTIEKYKVSVSGWRLFKKI
827	430	DRLLFTVTIERVSGWRLFKKIS
831	431	EKLLFTVTIEKVSGWRLFKKIS
832	432	KKLLFTVTIGKVSGWRLFKKIS
833	433	GSMRFRVTINSWRVTGYRLFERES
834	434	GSMKFRVTINSVTGYRLFEKES
844	435	KKITTTGTLWNGNKIID
845	436	ERLITPDGSMLFRVTINSVSGWRLFKKIS
846	437	GRPYEGIAVDFGKKITTTGTLWNGNKIIDERLITPDGSML FRVTINSVSGWRLFKKIS
847	438	GVTPNKLNYFGRPYEGIAVDFGKKITTTGTLWNGNKIID ERLITPDGSMLFRVTINSVSGWRLFKKIS
850	439	EKMLFRVTIGNWK
851	440	EKMLFRVTIQNWK
706	441	EKMLFRVTIDKWK
707	442	EKMLFRVTIEKWK
737	443	EKMLFRVTIERWK
736	444	EKMLFRVTIDRWK
760	445	KKMLFRVTIQKWK
852	446	KKMLFRVTIGKWK
853	447	KKMLFRVTIGRWK
854	448	KKMLFRVTIGNWK
855	449	KKMLFRVTIQNWK
856	450	KKMLFRVTIDKWK

857	451	KKMLFRVTIEKWK
858	452	KKMLFRVTIERWK
859	453	KKMLFRVTIDRWK
860	454	RKMLFRVTIQKWK
861	455	RKMLFRVTIGKWK
862	456	RKMLFRVTIGRWK
863	457	RKMLFRVTIGNWK
864	458	RKMLFRVTIQNWK
865	459	RKMLFRVTIDKWK
866	460	RKMLFRVTIEKWK
867	461	RKMLFRVTIERWK
868	462	RKMLFRVTIDRWK
656	463	EQMLFRVTINSWK
869	464	SRMLFRVTINSWK
533	465	GEMLFRVTINSWK
690	466	GKMKFRVTINSWK
678	467	GKMLFRVKINSWK
679	468	GKMLFRVRINSWK
681	469	GKMLFRVDINSWK
663	470	GKMLFRVTIDSWK
743	471	GKMLFRVTINKWK
714	472	EKMLFKVTIQKWK
870	473	EKMLFTVTIQKWK
871	474	EKMLFKVTIDKWK

872	475	EKMLFTVTIDKWK
873	476	EKMLFKVTIGRWK
744	477	DKMLFKVTIQKWK
745	478	DKMLFTVTIQKWK
874	479	DKMLFKVTIDKWK
875	480	DKMLFTVTIDKWK
876	481	GKMLFKVTIEKWK
877	482	GKMLFTVTIEKWK
748	483	DKMLFKVTIGKWK
749	484	DKMLFTVTIGKWK
878	485	DKMLFKVTIGNWK
879	486	DKMLFKVTIQNWK
781	487	GKMLFKVTINKWK
782	488	GKMLFTVTINKWK
752	489	DKMLFKVTIEKWK
753	490	DKMLFTVTIEKWK
750	491	DKLLFKVTIGKWK
786	492	DKMLFTVTINKWK
756	493	DKLLFTVTIQKWK
757	494	DKLLFTVTIQKYK
758	495	DKLLFTVTIEKWK
759	496	DKLLFTVTIEKYK
793	497	DKLLFTVTIGKWK
794	498	DKLLFTVTIGKYK

799	499	DKLLFTVTINKWK
800	500	DKLLFTVTINKYK
780	501	GKMLFRVTINS
765	502	DKMLFTVTIQK
779	503	DKMLFKVTIQK
820	504	DKLLFTVTIGK
819	505	DKMLFTVTIGK
822	506	DKMLFTVTIEK
821	507	DKLLFTVTIEK
627	508	*DKMLFRVTINSWK
628	509	*EKMLFRVTINSWK
629	510	*RKMLFRVTINSWK
630	511	*KKMLFRVTINSWK
631	512	*HKMLFRVTINSWK
632	513	*GLMLFRVTINSWK
633	514	*GQMLFRVTINSWK
634	515	*GTMLFRVTINSWK
635	516	*GKLLFRVTINSWK
636	517	*GKMLFKVTINSWK
637	518	*GKMLFRVTIQSWK
638	519	*GKMLFRVTIDSWK
639	520	*GKMLFRVTIGSWK
640	521	*GKMLFRVTINTWK
641	522	*GKMLFRVTINNWK

642	523	*GKMLFRVTINQWK
643	524	*GKMLFRVTINPWK
644	525	*GKMLFRVTINKWK
645	526	*GKMLFRVTINSWQ
646	527	*GKMLFRVTINSWN
647	528	*GKMLFRVTINSWT
648	529	*GKMLFRVTINSWH
649	530	*GKMLFRVTINSWP
650	531	*GKMLFRVTINSWR
677	532	GKMKFRVTIDSWK
680	533	GKMLFRVEINSWK
682	534	GKMLFRVQINSWK
683	535	GKMKFRVKINSWK
684	536	GKMKFRVRINSWK
685	537	GKMKFRVEINSWK
686	538	GKMKFRVDINSWK
687	539	GKMKFRVQINSWK
688	540	GKMKFRVNINSWK
689	541	GKMKFRVSINSWK
613	542	GKMLFRVNINSWK
614	543	GKMLFRVSINSWK
615	544	GKMLFRVWINSWK
616	545	GKMSFRVTINSWK
617	546	GKMWFRVTINSWK

618	547	GKMNFRVTINSWK
619	548	GSMLFRVTINSYK
620	549	GKMLFRVTINSYK
621	550	GKMLFRVTIKSWK
622	551	GKMLFRVTIESWK
716	552	GKMKFRVTIQSWK
717	553	GKMKFRVTIESWK
718	554	GKMKFRVTIKSWK
719	555	GKMKFRVTIRSWK
651	556	RLMLFRVTINSWK
652	557	RQMLFRVTINSWK
653	558	KLMLFRVTINSWK
654	559	KQMLFRVTINSWK
655	560	ELMLFRVTINSWK
657	561	DLMLFRVTINSWK
658	562	DQMLFRVTINSWK
659	563	DKMLFRVTINSWK
660	564	EKMLFRVTINSWK
661	565	RKMLFRVTINSWK
662	566	KKMLFRVTINSWK
665	567	GKMLFRVTIGSWK
667	568	GKMLFRVTINKWK
670	569	GKMLFRVTISKWK
671	570	GKMLFRVTIQKWK

672	571	GKMLFRVTITKWK
673	572	GKMLFRVTIKKWK
675	573	GKMLFKVTINSWK
676	574	RLMLFRVTIGKWK
701	575	GKMLFRVTINRWK
710	576	EKMLFTVTIGKWK
711	577	EKLLFTVTIGKWK
712	578	EKMLFTVTIGRWK
720	579	EKMLFTVTIEKWK
722	580	DKMLFRVTIESWK
726	581	EKLLFRVTIGKYK
746	582	DKLLFKVTIQKWK
747	583	DKLLFKVTIQKYK
751	584	DKLLFKVTIGKYK
754	585	DKLLFKVTIEKWK
755	586	DKLLFKVTIEKYK
761	587	KKLLFRVTIQKWK
762	588	DRMLFRVTIQRWR
766	589	ERMLFRVTIGRWR
768	590	GRMLFRVTINRWR
770	591	DRMLFRVTIERWR
783	592	DKMLFKVTIQKYK
784	593	DKMLFRVTINKWK
785	594	DKMLFKVTIEKYK

787	595	DKMLFKVTINKWK
693	900	GRMLFRVTINSWR
691	901	VSGWRLFRRIS
895	902	GRLLFVVVIERYR
937	903	VSGWRLFRRISC
938	904	GRMLFRVTINSWRC
939	905	GRLLFTVTIERYRC
840	906	GKLLFVVVIEKYK
900	907	GKLLFVTIEKVSGWRLFKKIS
*Terminus unblocked		

Table 10. Exemplary luciferase base sequences

Pep ID	**SEQ ID NO.**	**Sequence**
LgTrip 3546 – WT strand 9 – HiBiT	788	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS VTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKV VYPVDDHHFKVILPYGTLVID**GVTPNKLNYFGRPYEGIAV FDG**KKITTTGTL**WNGNKIIDERLITPD**GSMLFRVTINS**VSG WRLFKKIS**
LgTrip 3546 – WT strand 9 – SmBiT	789	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS VTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKV VYPVDDHHFKVILPYGTLVID**GVTPNKLNYFGRPYEGIAV FDG**KKITTTGTL**WNGNKIIDERLITPD**GSMLFRVTINS**VTG YRLFEEIL**
LgTrip 3546 (1-5)	790	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS VTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKV VYPVDDHHFKVILPYGTLVID

LgTrip 3546 (1-6)	791	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS VTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKV VYPVDDHHFKVILPYGTLVID**GVTPNKLNYFGRPYEGIAV FDG**
LgTrip 3546 (1-7)	792	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS VTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKV VYPVDDHHFKVILPYGTLVID**GVTPNKLNYFGRPYEGIAV FDG**KKITTTGTL
LgTrip 3546 (1-8)	793	MVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS VTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKV VYPVDDHHFKVILPYGTLVID**GVTPNKLNYFGRPYEGIAV FDG**KKITTTGTL**WNGNKIIDERLITPD**
LgTrip 3546 (strands 6-8) – WT strand 9 – HiBiT	794	**GVTPNKLNYFGRPYEGIAVFDG**KKITTTGTL**WNGNKIID ERLITPD**GSMLFRVTINS**VSGWRLFKKIS**
LgTrip 3546 (strands 7-8) – WT strand 9 – HiBiT	795	KKITTTGTL**WNGNKIIDERLITPD**GSMLFRVTINS**VSGWR LFKKIS**
LgTrip 3546 (strand 8) – WT strand 9 – HiBiT	796	**WNGNKIIDERLITPD**GSMLFRVTINS**VSGWRLFKKIS**
WT strand 9 – HiBiT	797	GSMLFRVTINS**VSGWRLFKKIS**

LgTrip 3546 (strands 6-8) – WT strand 9 – SmBiT	798	**GVTPNKLNYFGRPYEGIAVFDG**KKITTTGTL**WNGNKIID ERLITPD**GSMLFRVTINS**VTGYRLFEEIL**
LgTrip 3546 (strands 7-8) – WT strand 9 – SmBiT	799	KKITTTGTL**WNGNKIIDERLITPD**GSMLFRVTINS**VTGYRL FEEIL**
LgTrip 3546 (strand 8) – WT strand 9 – SmBiT	800	**WNGNKIIDERLITPD**GSMLFRVTINS**VTGYRLFEEIL**
WT strand 9 – SmBiT	801	GSMLFRVTINS**VTGYRLFEEIL**
β6-like	817	**GVTPNKLNYFGRPYEGIAVFDG**
β7-like	818	KKITTTGTL
β8-like	819	**WNGNKIIDERLITPD**

Table 11. Exemplary polypeptides

Name	Polypeptide construct description
ATG-2623	LgBiT-6His
ATG-3745	HiBiT-8His-LgTrip
ATG-3746	6His-HiBiT-4GS-LgTrip
ATG-4632	HiBiT-4GS-LgTrip-6His
ATG-4808	VS-HiBiT-0GS-LgTrip

ATG-4809 VS-HiBiT-4GS-LgTrip

ATG-4810 VS-HiBiT-8GS-LgTrip

ATG-4811 VS-HiBiT-12GS-LgTrip

ATG-4812 VS-HiBiT-16GS-LgTrip

ATG-4813 VS-HiBiT-20GS-LgTrip

ATG-4814 HiBiT-20GS-LgTrip

ATG-4815 LgTrip-0GS-VS-HiBiT

ATG-4816 LgTrip-4GS-VS-HiBiT

ATG-4817 LgTrip-8GS-VS-HiBiT

ATG-4818 LgTrip-8GS-HiBiT

ATG-4819 VS-HiBIT-0GS-LgTrip-6His

ATG-4820 VS-HiBiT-4GS-LgTrip-6His

ATG-4821 VS-HiBiT-8GS-LgTrip-6His

ATG-4822 VS-HiBiT-12GS-LgTrip-6His

ATG-4823 VS-HiBiT-16GS-LgTrip-6His

ATG-4824 HiBiT-20GS-LgTrip-6His

ATG-4825 VS-HiBiT-20GS-LgTrip-6His

ATG-4826 6His-LgTrip-0GS-VS-HiBiT

ATG-4827 6His-LgTrip-4GS-VS-HiBiT

ATG-4828 6His-LgTrip-8GS-HiBiT

ATG-4829 6His-LgTrip-8GS-VS-HiBiT

Mutations on 5333

ATG-5538 (I41T) Nucleotide (SEQ ID NO: 934)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGACCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA

TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5538 (I41T) Amino Acid (SEQ ID NO: 935)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPTQRMVRSGENALKI
DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGR
PYEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5339 (K11N) Nucleotide (SEQ ID NO: 936)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAACCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5339 (K11N) Amino Acid (SEQ ID NO: 937)

VFTLEDFVGDWNQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5340 (R152Q) Nucleotide (SEQ ID NO: 938)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCAAGTAACCATCAACAGC

ATG-5340 Amino Acid (SEQ ID NO: 939)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFQVTINS

ATG-5407 V135A Nucleotide (SEQ ID NO: 940)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5407 V135A Amino Acid (SEQ ID NO: 941)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFRVTINS

ATG-5408 D156N Nucleotide (SEQ ID NO: 942)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAACCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCGACAGC

ATG-5408 D156N Amino Acid (SEQ ID NO: 943)
VFTLEDFVGDWNQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTIDS

ATG-5409 H57Q Nucleotide (SEQ ID NO: 944)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG

TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCAAGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCC
AGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGT
GATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAA
CTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCAC
TGTAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCAC
CCCCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5409 H57Q Amino Acid (SEQ ID NO: 945)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IQVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5411 N33K+I155V Nucleotide (SEQ ID NO: 946)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAAACTCGCC
GTGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAG
ATCGACATCCACGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCC
AGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGT
GATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAA
CTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCAC
TGTAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCAC
CCCCGACGGCTCCATGTCCTTCCGAGTAACCGTCAACAGC

ATG-5411 N33K+I155V Amino Acid (SEQ ID NO: 947)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQKLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTVNS

ATG-5412 I54V+L127A Nucleotide (SEQ ID NO: 948)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGG
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
CAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5412 I54V+L127A Amino Acid (SEQ ID NO: 949)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKV
DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGR
PYEGIAVFDGKKITATGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5413 M44L+V135A Nucleotide (SEQ ID NO: 950)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGTTGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5413 M44L+V135A Amino Acid (SEQ ID NO: 951)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRLVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFRVTINS

ATG-5414 V119A Nucleotide (SEQ ID NO: 952)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGCGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5414 V119A (Amino Acid) (SEQ ID NO: 953)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAAFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5416 D9A+H57Q Nucleotide (SEQ ID NO: 954)

GTCTTCACACTCGAAGATTTCGTTGGGGCCTGGAAACAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT

CGACATCCAAGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5416 D9A+H57Q (Amino Acid) (SEQ ID NO: 955)

VFTLEDFVGAWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IQVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5417 N33D+I41T Nucleotide (SEQ ID NO: 956)

GTCTTCACACTCGAAGATTTCGTAGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGGATCTCGCCG
TGTCCGTAACTCCGACCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5417 N33D+I41T Amino Acid (SEQ ID NO: 957)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQDLAVSVTPTQRMVRSGENALKI
DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGR
PYEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5418 Q32R+I155T Nucleotide (SEQ ID NO: 958)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAGACAGACTGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCGGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCACCAACAGC

ATG-5418 Q32R+I155T mino Acid (SEQ ID NO: 959)

VFTLEDFVGDWRQTAAYNLDQVLEQGGVSSLLRNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTTNS

ATG-5419 D19V+M106T+V120L Nucleotide (SEQ ID NO: 960)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGTCCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACACGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGCTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5419 D19V+M106T+V120L Amino Acid (SEQ ID NO: 961)

VFTLEDFVGDWKQTAAYNLVQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNTLNYFGRP
YEGIAVLDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5420 I41N+E63G Nucleotide (SEQ ID NO: 962)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGAACCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGGAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TGACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5420 I41N+E63G Amino Acid (SEQ ID NO: 963)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPNQRMVRSGENALKI
DIHVIIPYGGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFG
RPYEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5421 N50S Nucleotide (SEQ ID NO: 964)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAACCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAGTGCCCTGAAGA

TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5421 N50S Amino Acid (SEQ ID NO: 965)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGESALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5432 L3H Nucleotide (SEQ ID NO: 966)

GTCTTCACACACGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5432 Amino Acid (SEQ ID NO: 967)

VFTHEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5433 T13S Nucleotide (SEQ ID NO: 968)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGTCAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5433 T13S Amino Acid (SEQ ID NO: 969)

VFTLEDFVGDWKQSAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5434 P93H Nucleotide (SEQ ID NO: 970)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5434 P93H Amino Acid (SEQ ID NO: 117)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5435 F120L Nucleotide (SEQ ID NO: 118)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGCTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATAGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5435 F120L Amino Acid (SEQ ID NO: 119)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVLDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5437 S157R Nucleotide (SEQ ID NO: 120)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA

TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACCGC

ATG-5437 S157R Amino Acid (SEQ ID NO: 121)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINR

ATG-5438 H86L Nucleotide (SEQ ID NO: 122)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCTTCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5438 H86L Amino Acid (SEQ ID NO: 123)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDLHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5439 M149V Nucleotide (SEQ ID NO: 124)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCGTGTCCTTCCGAGTAACCATCAACAGC

ATG-5439 M149V Amino Acid (SEQ ID NO: 125)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSVSFRVTINS

ATG-5440 I59V Nucleotide (SEQ ID NO: 126)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTTGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCGTCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5440 I59V Amino Acid (SEQ ID NO: 127)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVVIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGR
PYEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5441 D19A Nucleotide (SEQ ID NO: 128)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGCCCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5441 D19A Amino Acid (SEQ ID NO: 129)

VFTLEDFVGDWKQTAAYNLAQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5442 Q69L+T144S Nucleotide (SEQ ID NO: 130)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA

TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCTAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTATTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCTCCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5442 Q69L+T144S Amino Acid (SEQ ID NO: 131)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADLMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLISPDGSMSFRVTINS

ATG-5456 K11Y Nucleotide (SEQ ID NO: 132)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGTACCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5456 K11Y Nucleotide Amino Acid (SEQ ID NO: 596)

VFTLEDFVGDWYQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5457 K11R Nucleotide (SEQ ID NO: 597)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGGCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5457 K11R Amino Acid (SEQ ID NO: 598)

VFTLEDFVGDWRQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5458 K11L Nucleotide (SEQ ID NO: 599)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGCTGCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5458 K11L Amino Acid (SEQ ID NO: 600)

VFTLEDFVGDWLQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5459 R152Q Nucleotide (SEQ ID NO: 601)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAACCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5459 R152Q Amino Acid (SEQ ID NO: 602)

VFTLEDFVGDWNQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFQVTINS

ATG-5460 N156D Nucleotide (SEQ ID NO: 603)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAACCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT

CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCGACAGC

ATG-5460 N156D Amino Acid (SEQ ID NO: 604)

VFTLEDFVGDWNQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTIDS

ATG-5491 K11Q Nucleotide (SEQ ID NO: 605)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGCAGCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5491 K11Q Amino Acid (SEQ ID NO: 606)

VFTLEDFVGDWQQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5492 K11M Nucleotide (SEQ ID NO: 607)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGATGCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5492 K11M Amino Acid (SEQ ID NO: 608)

VFTLEDFVGDWMQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5493 K11H Nucleotide (SEQ ID NO: 609)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGCACCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5493 K11H Amino Acid (SEQ ID NO: 610)

VFTLEDFVGDWHQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5494 K11F Nucleotide (SEQ ID NO: 611)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGTTCCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5494 K11F Amino Acid (SEQ ID NO: 612)

VFTLEDFVGDWFQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5495 K11W Nucleotide (SEQ ID NO: 613)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGTGGCAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGAT

CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5495 K11W Amino Acid (SEQ ID NO: 614)

VFTLEDFVGDWWQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFRVTINS

ATG-5505 V135A+R152Q Nucleotide (SEQ ID NO: 615)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATAGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTACGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5505 V135A+R152Q Amino Acid (SEQ ID NO: 616)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5506 V135A+R152Q+N156D Nucleotide (SEQ ID NO: 617)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATAGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTACGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCAGGTAACCATCGACAGC

ATG-5506 V135A+R152Q+N156D Amino Acid (SEQ ID NO: 618)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTIDS

ATG-5507 P93H+V135A Nucleotide (SEQ ID NO: 619)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATAGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTACGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCGAGTAACCATCAACAGC

ATG-5507 P93H+V135A Amino Acid (SEQ ID NO: 620)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFRVTINS

ATG-5508 P93H+R152Q Nucleotide (SEQ ID NO: 621)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5508 P93H+R152Q Amino Acid (SEQ ID NO: 622)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFQVTINS

ATG-5509 P93H+R152Q+N156D Nucleotide (SEQ ID NO: 623)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA

TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGTCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCAGGTAACCATCGACAGC

ATG-5509 P93H+R152Q+N156D Amino Acid (SEQ ID NO: 624)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGVKIIDERLITPDGSMSFQVTIDS

ATG-5510 P93H+V135A+R152Q Nucleotide (SEQ ID NO: 625)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATAGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTACGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5510 P93H+V135A+R152Q Amino Acid (SEQ ID NO: 626)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5511 P93H+V135G+R152Q Nucleotide (SEQ ID NO: 627)

GTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATAGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGAACT
ATTTCGGACGGCCGTACGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTG
TAACAGGGACCCTGTGGAACGGCGGCAAAATTATCGACGAGCGCCTGATCACCC
CCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5511 P93H+V135G+R152Q Amino Acid (SEQ ID NO: 628)

VFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKID
IHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGGKIIDERLITPDGSMSFQVTINS

Mutations on 5344

ATG-5534 P93H+V135A+R152Q (codon optimized 5510) Nucleotide (SEQ ID NO: 629)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGAtaaactcg

ATG-5534 P93H+V135A+R152Q (codon optimized 5510) Amino Acid (SEQ ID NO: 630)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5535 P93H+V135A+R152Q+N156D Nucleotide (SEQ ID NO: 631)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAGACTCG

ATG-5535 P93H+V135A+R152Q+N156D Amino Acid (SEQ ID NO: 632)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTID

ATG-5536 E4D+V135A+R152Q Nucleotide (SEQ ID NO: 633)

TTCACTTTGGACGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACCATATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5536 E4D+V135A+R152Q Amino Acid (SEQ ID NO: 634)

FTLDDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5537 E4D+Q42M+P93H+V135A+R152Q Nucleotide (SEQ ID NO: 635)

TTCACTTTGGACGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATAATGCGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCGT

ATG-5537 E4D+Q42M+P93H+V135A+R152Q Amino Acid (SEQ ID NO: 636)

FTLDDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIMRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

Mutations on 5534

ATG-5652 5534+Y16F Q20P Q152H Nucleotide (SEQ ID NO: 637)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATTCAACTTAG
ACCCAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT

AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGCTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCACGTTACGATAAACTCG

ATG-5652 5534+Y16F Q20P Q152H Amino Acid (SEQ ID NO: 638)

FTLEDFVGDWKQTAAFNLDPVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFHVTINS

ATG-5653 5534+M106R,Y114F,E4E Nucleotide (SEQ ID NO: 639)

TTCACTTTGGAGGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACAGGTTAAACTATT
TCGGTAGACCTTTTGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTA
CTGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAG
ACGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5653 5534+M106R,Y114F,E4E Amino Acid (SEQ ID NO: 640)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNRLNYFGRPF
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5654 5534+M44V Nucleotide (SEQ ID NO: 641)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTGTGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
AGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5654 5534+M44V (amino acid) (SEQ ID NO: 642)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRVVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5655 5534+M44I Nucleotide (SEQ ID NO: 643)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTATAGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5655 5534+M44I Amino Acid (SEQ ID NO: 644)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRIVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5656 5534+A35A,P40P,M106R Nucleotide (SEQ ID NO: 645)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCTGTTAG
TGTTACGCCAATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACAGGTTAAACTATT
TCGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTA
CTGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAG
ACGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5656 5534+A35A,P40P,M106R amino acid (SEQ ID NO: 646)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNRLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5657 5534+M106T Nucleotide (SEQ ID NO: 647)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACACGTTAAACTATT
TCGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTA
CTGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAG
ACGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5657 5534+M106T amino acid (SEQ ID NO: 648)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNTLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5659 5534+M106K+L30L+K136E Nucleotide (SEQ ID NO: 649)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTGTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACAAGTTAAACTATT
TCGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTA
CTGGTACGTTATGGAACGGAGCCGAGATAATAGACGAGAGATTAATAACGCCAG
ACGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5659 5534+M106K+L30L+K136E amino acid (SEQ ID NO: 650)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNKLNYFGRP
YEGIAVFDGKKITVTGTLWNGAEIIDERLITPDGSMSFQVTINS

ATG-5660 5534+V36G,N50N,A14A Nucleotide (SEQ ID NO: 651)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCGGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGGTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAACGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5660 5534+V36G,N50N,A14A amino acid (SEQ ID NO: 652)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAGSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5661 5534+A15A,A67A,M106L,L107L,I138K Nucleotide (SEQ ID NO: 653)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCGTACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCTGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACTTGTTGAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAAAAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5661 5534+A15A,A67A,M106L,L107L,I138K amino acid (SEQ ID NO: 654)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNLLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIKDERLITPDGSMSFQVTINS

ATG-5662 5534+D5G, A35T, L92F, D100V, R141G, L1142S Nucleotide (SEQ ID NO: 655)

TTCACTTTGGAAGGTTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGACAGTTAG
TGTTACGCCTATACAACGTATGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTTCACTATGGAACTTTAGTTATAGTCGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGGGATCAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5662 5534+D5G, A35T, L92F, D100V, R141G, L1142S amino acid (SEQ ID NO: 656)

FTLEGFVGDWKQTAAYNLDQVLEQGGVSSLLQNLTVSVTPIQRMVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVIFHYGTLVIVGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDEGSITPDGSMSFQVTINS

ATG-5676 5534+M44K Nucleotide (SEQ ID NO: 657)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTAAGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGAtaaactcg

ATG-5676 5534+M44K amino acid (SEQ ID NO: 658)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRKVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5677 5534+M44E Nucleotide (SEQ ID NO: 659)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTGAGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5677 5534+M44E amino acid (SEQ ID NO: 660)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQREVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5678 5534+M44A Nucleotide (SEQ ID NO: 661)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTGCGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5678 5534+M44A amino acid (SEQ ID NO: 662)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRAVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5679 M44C Nucleotide (SEQ ID NO: 663)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTTGTGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5679 M44C amino acid (SEQ ID NO: 664)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRCVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5680 M44W Nucleotide (SEQ ID NO: 665)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTTGGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCGT

ATG-5680 M44W amino acid (SEQ ID NO: 666)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRWVRSGENALKIDI
HVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRP
YEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5681 M44G Nucleotide (SEQ ID NO: 667)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG

TGTTACGCCTATACAACGTGGGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5681 M44G amino acid (SEQ ID NO: 668)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRGVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5682 M44H Nucleotide (SEQ ID NO: 669)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTCATGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5682 M44H amino acid (SEQ ID NO: 670)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRHVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5683 M44S Nucleotide (SEQ ID NO: 671)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTAGTGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5683 M44S amino acid (SEQ ID NO: 672)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRSVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5684 M44Q Nucleotide (SEQ ID NO: 673)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTCAGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5684 M44Q Amino acid (SEQ ID NO: 674)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRQVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5685 5534+M44R Nucleotide (SEQ ID NO: 675)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTCGGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5685 5534+M44R amino acid (SEQ ID NO: 676)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRRVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5686 5534+M44T Nucleotide (SEQ ID NO: 677)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTACGGTGAGATCGGGAGAAAATGCATTAAAAATAGA

CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5686 5534+M44T amino acid (SEQ ID NO: 678)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRTVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5687 5534+M44Y Nucleotide (SEQ ID NO: 679)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTTATGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5687 5534+M44Y Amino acid (SEQ ID NO: 680)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRYVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5688 5534+M44L Nucleotide (SEQ ID NO: 681)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTCTGGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5688 5534+M44L amino acid (SEQ ID NO: 682)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRLVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5689 5534+M44P Nucleotide (SEQ ID NO: 683)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTCCTGTGAGATCGGGAGAAAATGCATTAAAAATAGA
CATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGAT
AGAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATA
TTACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTT
CGGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTAC
TGGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGA
CGGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5689 5534+M44P amino acid (SEQ ID NO: 684)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRPVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5590 5534+M44F Nucleotide (SEQ ID NO: 685)

TTCACTTTGGAAGATTTCGTGGGTGACTGGAAACAGACTGCAGCATACAACTTAG
ACCAAGTGTTGGAACAGGGTGGAGTTAGTTCGCTTTTACAGAATTTGGCAGTTAG
TGTTACGCCTATACAACGTTTTGTGAGATCGGGAGAAAATGCATTAAAAATAGAC
ATACATGTGATAATACCTTATGAAGGATTATCGGCAGACCAAATGGCACAGATA
GAAGAGGTTTTTAAAGTTGTTTACCCAGTTGATGATCATCACTTTAAAGTTATATT
ACACTATGGAACTTTAGTTATAGACGGTGTGACTCCTAACATGTTAAACTATTTC
GGTAGACCTTATGAAGGAATAGCAGTTTTCGACGGAAAGAAAATAACGGTTACT
GGTACGTTATGGAACGGAGCCAAGATAATAGACGAGAGATTAATAACGCCAGAC
GGAAGTATGAGTTTCCAGGTTACGATAAACTCG

ATG-5590 5534+M44F Amino acid (SEQ ID NO: 686)

FTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRFVRSGENALKIDIH
VIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPY
EGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5788 Nucleotide (SEQ ID NO: 687)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag

gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5788 Amino Acid (SEQ ID NO: 688)

MKHHHHHHAIAMVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5789 Nucleotide (SEQ ID NO: 689)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcaccccgacaagtccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5789 Amino Acid (SEQ ID NO: 690)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDKSMLFRVTINGVS
GWRLFKKIS

ATG-5790 Nucleotide (SEQ ID NO: 691)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcaccccgacggctccatgctgttccgagtaaccatcaacaaggttagcggctggcgcctgt
tcaagaagatcagc

ATG-5790 Amino Acid (SEQ ID NO: 692)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINKVS
GWRLFKKIS

ATG-5791 Nucleotide (SEQ ID NO: 693)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcgagaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5791 Amino Acid (SEQ ID NO: 694)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGEKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5792 Nucleotide (SEQ ID NO: 695)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
aagctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5792 Amino Acid (SEQ ID NO: 696)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNKLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5793 Nucleotide (SEQ ID NO: 697)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc

aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatccagggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5793 Amino Acid (SEQ ID NO: 698)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTIQGVS
GWRLFKKIS

ATG-5794 Nucleotide (SEQ ID NO: 699)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacaagaagatgctgttccgagtaaccatccagaaggttagcggctggcgcctg
ttcaagaagatcagc

ATG-5794 Amino Acid (SEQ ID NO: 700)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDKKMLFRVTIQKVS
GWRLFKKIS

ATG-5795 Nucleotide (SEQ ID NO: 701)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5795 Amino Acid (SEQ ID NO: 702)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIM
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5796 Nucleotide (SEQ ID NO: 703)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacacccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggca
acaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgtt
caagaagatcagc

ATG-5796 Amino Acid (SEQ ID NO: 704)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGHPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5797 Nucleotide (SEQ ID NO: 705)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggcaagatgctgttccgagtaaccatcaacggcgttagcggctggcgcctg
ttcaagaagatcagc

ATG-5797 Amino Acid (SEQ ID NO: 706)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGKMLFRVTINGVS
GWRLFKKIS

ATG-5798 Nucleotide (SEQ ID NO: 707)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcgatcccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5798 Amino Acid (SEQ ID NO: 708)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLIDPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5799 Nucleotide (SEQ ID NO: 709)

Atgaaacatcaccatcaccatcatgcgatcgccatggatttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5799 Amino Acid (SEQ ID NO: 710)

MKHHHHHHAIAMDFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5800 Nucleotide (SEQ ID NO: 711)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactaccacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5800 Amino Acid (SEQ ID NO: 712)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5801 Nucleotide (SEQ ID NO: 713)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag

gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgcccatcggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5801 Amino Acid (SEQ ID NO: 714)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPIGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5802 Nucleotide (SEQ ID NO: 715)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgtgaccggctggcggctg
tgcgaacgcattctg

ATG-5802 Amino Acid (SEQ ID NO: 716)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVT
GWRLCERIL

ATG-5803 Nucleotide (SEQ ID NO: 717)

Atgaaacatcaccatcaccatcatgcgatcgccatggatttcacactcgacgatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgcccatcggcacactggtaatcgacggggttacgccgaac
aagctgaactatttcggacacccgtatgaaggcatcgccgtgttcgacggcgagaagatcactaccacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcgatcccgacaagaagatgctgttccgagtaaccatccagaaggttagcggctggcgcctgt
tcaagaagatcagc

ATG-5803 Amino Acid (SEQ ID NO: 718)

MKHHHHHHAIAMDFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIM
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPIGTLVIDG

VTPNKLNYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLIDPDKKMLFRVTIQKVS
GWRLFKKIS

ATG-5804 Nucleotide (SEQ ID NO: 719)

Atgaaacatcaccatcaccatcatgcgatcgccatggatttcacactcgacgatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatcatgaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgcccatcggcacactggtaatcgacggggttacgccgaac
aagctgaactatttcggacacccgtatgaaggcatcgccgtgttcgacggcgagaagatcactaccacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcgatcccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5804 Amino Acid (SEQ ID NO: 720)

MKHHHHHHAIAMDFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIM
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPIGTLVIDG
VTPNKLNYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLIDPDKKMLFRVTIQKVS
GWRLFKKIS

Example 52-Circularly permuted LgBiT

ATG-4992 PEP78-8GS-LGTRIP3546-6HIS Nucleotide (SEQ ID NO: 721)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCC
GCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAG
AATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATG
CCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCA
AATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCA
CTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAAC
AAGCTGAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAA
AAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGC
CTGATCACCCCCGACCATCACCATCACCATCAT

ATG-4992 PEP78-8GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 722)

MNVSGWRLFKKISNGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNL
AVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILP
YGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHH
H

ATG-4993 PEP79-8GS-LGTRIP3546-6HIS nucleotide (SEQ ID NO: 723)

ATGAACGTGACCGGCTACCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCC
GCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAG
AATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATG
CCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCA
AATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCA
CTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAAC
AAGCTGAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAA
AAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGC
CTGATCACCCCCGACCATCACCATCACCATCAT

ATG-4993 PEP79-8GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 724)

MNVTGYRLFKKISNGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNL
AVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILP
YGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHH
H

ATG-4994 PEP99-8GS-LGTRIP3546-6HIS nucleotide (SEQ ID NO: 725)

ATGGTGACCGGCTACCGGCTGTTCGAGAAGATTAGCGGCTCGAGCGGTGGCTCG
AGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGAT
CACCCCCGACCATCACCATCACCATCAT

ATG-4994 PEP99-8GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 726)

MVTGYRLFEKISGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAV
SVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYG
TLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-4995 PEP128-8GS-LGTRIP3546-6HIS Nucleotide (SEQ ID NO: 727)

ATGGTGACCGGCTACCGGCTGTTCGAGAAGATTCTGGGCTCGAGCGGTGGCTCG
AGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGA

AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGAT
CACCCCCGACCATCACCATCACCATCAT

ATG-4995 PEP128-8GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 728)

MVTGYRLFEKILGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAV
SVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYG
TLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-4996 HIBIT-8GS-LGTRIP3546-6HIS nucleotide (SEQ ID NO: 729)

ATGGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCGGCTCGAGCGGTGGCTCG
AGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGAT
CACCCCCGACCATCACCATCACCATCAT

ATG-4996 HIBIT-8GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 730)

MVSGWRLFKKISGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAV
SVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYG
TLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-4997 HIBIT-STOG-8GS-LGTRIP-3546-6HIS nucleotide (SEQ ID NO: 731)

ATGGTGAGCGGCTGGCGGCTGTTCAAGAAGATTGGCGGCTCGAGCGGTGGCTCG
AGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGAT
CACCCCCGACCATCACCATCACCATCAT

ATG-4997 HIBIT-STOG-8GS-LGTRIP-3546-6HIS amino acid (SEQ ID NO: 732)

MVSGWRLFKKIGGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA
VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPY
GTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-4998 HIBIT-ISTOSG-8GS-LGTRIP-3546-6HIS nucleotide (SEQ ID NO: 733)

ATGGTGAGCGGCTGGCGGCTGTTCAAGAAGTCGGGCGGCTCGAGCGGTGGCTCG
AGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGAT
CACCCCCGACCATCACCATCACCATCAT

ATG-4998 HIBIT-ISTOSG-8GS-LGTRIP-3546-6HIS amino acid (SEQ ID NO: 734)

MVSGWRLFKKSGGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA
VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPY
GTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-4999 HIBIT-DELETE IS-8GS-LGTRIP-3546-6HIS nucleotide (SEQ ID NO: 735)

ATGGTGAGCGGCTGGCGGCTGTTCAAGAAGGGCTCGAGCGGTGGCTCGAGCGGT
TTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTACAACCTG
GACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGTGT
CCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGATCG
ACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCAGAT
CGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGATC
CTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCTGAACTATT
TCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTACCA
CAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCACCCCCG
ACCATCACCATCACCATCAT

ATG-4999 HIBIT-DELETE IS-8GS-LGTRIP-3546-6HIS nucleotide amino acid (SEQ ID NO: 736)

MVSGWRLFKKGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS
VTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGT
LVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-5000 HIBIT-DELETE S-8GS-LGTRIP-3546-6HIS nucleotide (SEQ ID NO: 737)

ATGGTGAGCGGCTGGCGGCTGTTCAAGAAGATTGGCTCGAGCGGTGGCTCGAGC
GGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTACAAC
CTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCG
TGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGA
TCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCTGAAC
TATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACT
ACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCACC
CCCGACCATCACCATCACCATCAT

ATG-5000 HIBIT-DELETE S-8GS-LGTRIP-3546-6HIS amino acid (SEQ ID NO: 738)

MVSGWRLFKKIGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS
VTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGT
LVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-5001 VSHIBIT-ISTOSG-8GS-LGTRIP-3546-6HIS nucleotide (SEQ ID NO: 739)

ATGGTTTCCGTGAGCGGCTGGCGGCTGTTCAAGAAGTCGGGCGGCTCGAGCGGT
GGCTCGAGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCC
GCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAG
AATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATG
CCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCA
AATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCA
CTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAAC
AAGCTGAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAA
AAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGC
CTGATCACCCCCGACCATCACCATCACCATCAT

ATG-5001 VSHIBIT-ISTOSG-8GS-LGTRIP-3546-6HIS amino acid (SEQ ID NO: 740)

MVSVSGWRLFKKSGGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQN
LAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVIL
PYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHH
HH

ATG-5002 VSHIBIT-STOG-8GS-LGTRIP-3546-6HIS nucleotide (SEQ ID NO: 741)

ATGGTTTCCGTGAGCGGCTGGCGGCTGTTCAAGAAGATTGGCGGCTCGAGCGGT
GGCTCGAGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCC
GCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAG
AATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATG
CCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCA
AATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCA
CTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAAC
AAGCTGAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAA
AAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGC
CTGATCACCCCCGACCATCACCATCACCATCAT

ATG-5002 VSHIBIT-STOG-8GS-LGTRIP-3546-6HIS amino acid (SEQ ID NO: 742)

MVSVSGWRLFKKIGGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNL
AVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILP
YGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHH
H

ATG-5003 VSHIBIT-DELETE IS-8GS-LGTRIP-3546-6HIS nucleotide (SEQ ID NO: 743)

ATGGTTTCCGTGAGCGGCTGGCGGCTGTTCAAGAAGGGCTCGAGCGGTGGCTCG
AGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGAT
CACCCCCGACCATCACCATCACCATCAT

ATG-5003 VSHIBIT-DELETE IS-8GS-LGTRIP-3546-6HIS amino acid (SEQ ID NO: 744)

MVSVSGWRLFKKGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA
VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPY
GTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-5004 VSHIBIT-DELETE S-8GS-LGTRIP-3546-6HIS nucleotide (SEQ ID NO: 745)

ATGGTTTCCGTGAGCGGCTGGCGGCTGTTCAAGAAGATTGGCTCGAGCGGTGGCT
CGAGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCT
ACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCT
CGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTG
AAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATG
GCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTA

AGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGC
TGAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGA
TCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGA
TCACCCCCGACCATCACCATCACCATCAT

ATG-5004 VSHIBIT-DELETE S-8GS-LGTRIP-3546-6HIS amino acid (SEQ ID NO: 746)

MVSVSGWRLFKKIGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA
VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPY
GTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-5485 PEP78-5GS-LGTRIP3546-6HIS nucleotide (SEQ ID NO: 747)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
TCGTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTACAACC
TGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCGCCGT
GTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGAAGAT
CGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCCCA
GATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCTGAAC
TATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACT
ACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCACC
CCCGACCATCACCATCACCATCAT

ATG-5485 PEP78-5GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 748)

MNVSGWRLFKKISNGSSGSFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVS
VTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGT
LVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-5486 PEP78-6GS-LGTRIP3546-6HIS nucleotide (SEQ ID NO: 749)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGAT
CACCCCCGACCATCACCATCACCATCAT

ATG-5486 PEP78-6GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 750)

MNVSGWRLFKKISNGSSGGSFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLA
VSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPY
GTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHHH

ATG-5487 PEP78-7GS-LGTRIP3546-6HIS nucleotide (SEQ ID NO: 751)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCCGCCT
ACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCT
CGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATGCCCTG
AAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATG
GCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTA
AGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAACAAGC
TGAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGA
TCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGCCTGA
TCACCCCCGACCATCACCATCACCATCAT

ATG-5487 PEP78-7GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 752)

MNVSGWRLFKKISNGSSGGSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNL
AVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILP
YGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHHHH
H

ATG-5488 PEP78-9GS-LGTRIP3546-6HIS nucleotide (SEQ ID NO: 753)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACA
GCCGCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGC
AGAATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAA
ATGCCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGA
CCAAATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCA
TCACTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCG
AACAAGCTGAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGC
AAAAAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAG
CGCCTGATCACCCCCGACCATCACCATCACCATCAT

ATG-5488 PEP78-9GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 754)

MNVSGWRLFKKISNGSSGGSSGGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQ
NLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVI
LPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHH
HHH

ATG-5489 PEP78-10GS-LGTRIP3546-6HIS nucleotide (SEQ ID NO: 755)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTGGTAGCTTCACACTCGACGATTTCGTTGGGGACTGGGAACAG
ACAGCCGCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTG
CTGCAGAATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTG
AAAATGCCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGC
CGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGA
TCATCACTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACG
CCGAACAAGCTGAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGAC
GGCAAAAAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGAC
GAGCGCCTGATCACCCCCGACCATCACCATCACCATCAT

ATG-5489 PEP78-10GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 756)

MNVSGWRLFKKISNGSSGGSSGGSFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQ
NLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVI
LPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHHH
HHH

ATG-5490 PEP78-11GS-LGTRIP3546-6HIS nucleotide (SEQ ID NO: 757)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTGGTAGCTCGTTCACACTCGACGATTTCGTTGGGGACTGGGAAC
AGACAGCCGCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTT
TGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGG
TGAAAATGCCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC
GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGAT
GATCATCACTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTA
CGCCGAACAAGCTGAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCG
ACGGCAAAAAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCG
ACGAGCGCCTGATCACCCCCGACCATCACCATCACCATCAT

ATG-5490 PEP78-11GS-LGTRIP3546-6HIS amino acid (SEQ ID NO: 758)

MNVSGWRLFKKISNGSSGGSSGGSSFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLL
QNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFK
VILPYGTLVIDGVTPNKLNYFGRPYEGIAVFDGKKITTTGTLWNGNKIIDERLITPDHH
HHHH

ATG-5513 4992+R112H+K123E nucleotide (SEQ ID NO: 759)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCC
GCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAG
AATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATG
CCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCA

AATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCA
CTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAAC
AAGCTGAACTATTTCGGACACCCGTATGAAGGCATCGCCGTGTTCGACGGCGAG
AAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGC
CTGATCACCCCCGACCATCACCATCACCATCAT

ATG-5513 4992+R112H+K123E amino acid (SEQ ID NO: 760)

MNVSGWRLFKKISNGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNL
AVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILP
YGTLVIDGVTPNKLNYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLITPDHHHHH
H

ATG-5514 4992+R112H+T144D nucleotide (SEQ ID NO: 761)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCC
GCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAG
AATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATG
CCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCA
AATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCA
CTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAAC
AAGCTGAACTATTTCGGACACCCGTATGAAGGCATCGCCGTGTTCGACGGCAAA
AAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGC
CTGATCGATCCCGACCATCACCATCACCATCAT

ATG-5514 4992+R112H+T144D amino acid (SEQ ID NO: 762)

MNVSGWRLFKKISNGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNL
AVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILP
YGTLVIDGVTPNKLNYFGHPYEGIAVFDGKKITTTGTLWNGNKIIDERLIDPDHHHHH
H

ATG-5515 4992+R112H+K123E+T144D nucleotide (SEQ ID NO: 763)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTTTCACACTCGACGATTTCGTTGGGGACTGGGAACAGACAGCC
GCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAG
AATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGGTGAAAATG
CCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCA
AATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCA
CTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTACGCCGAAC
AAGCTGAACTATTTCGGACACCCGTATGAAGGCATCGCCGTGTTCGACGGCGAG
AAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCGACGAGCGC
CTGATCGATCCCGACCATCACCATCACCATCAT

ATG-5515 4992+R112H+K123E+T144D amino acid (SEQ ID NO: 764)

MNVSGWRLFKKISNGSSGGSSGFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNL
AVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILP
YGTLVIDGVTPNKLNYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLIDPDHHHHH
H

ATG-5516 5490+R112H+K123E nucleotide (SEQ ID NO: 765)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTGGTAGCTCGTTCACACTCGACGATTTCGTTGGGGACTGGGAAC
AGACAGCCGCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTT
TGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGG
TGAAAATGCCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC
GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGAT
GATCATCACTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTA
CGCCGAACAAGCTGAACTATTTCGGACACCCGTATGAAGGCATCGCCGTGTTCG
ACGGCGAGAAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCG
ACGAGCGCCTGATCACCCCCGACCATCACCATCACCATCAT

ATG-5516 5490+R112H+K123E amino acid (SEQ ID NO: 766)

MNVSGWRLFKKISNGSSGGSSGGSSFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLL
QNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFK
VILPYGTLVIDGVTPNKLNYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLITPDHH
HHHH

ATG-5517 5490+R112H+K123E+T144D nucleotide (SEQ ID NO: 767)

ATGAACGTGAGCGGCTGGCGGCTGTTCAAGAAGATTAGCAACGGCTCGAGCGGT
GGCTCGAGCGGTGGTAGCTCGTTCACACTCGACGATTTCGTTGGGGACTGGGAAC
AGACAGCCGCCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTT
TGCTGCAGAATCTCGCCGTGTCCGTAACTCCGATCATGAGGATTGTCCGGAGCGG
TGAAAATGCCCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGC
GCCGACCAAATGGCCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGAT
GATCATCACTTTAAGGTGATCCTGCCCTATGGCACACTGGTAATCGACGGGGTTA
CGCCGAACAAGCTGAACTATTTCGGACACCCGTATGAAGGCATCGCCGTGTTCG
ACGGCGAGAAGATCACTACCACAGGGACCCTGTGGAACGGCAACAAAATTATCG
ACGAGCGCCTGATCGATCCCGACCATCACCATCACCATCAT

ATG-5517 5490+R112H+K123E+T144D amino acid (SEQ ID NO: 768)
MNVSGWRLFKKISNGSSGGSSGGSSFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLL
QNLAVSVTPIMRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFK
VILPYGTLVIDGVTPNKLNYFGHPYEGIAVFDGEKITTTGTLWNGNKIIDERLIDPDHH
HHHH

ATG-5810 Nucleotide (SEQ ID NO: 769)

ATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAGCAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCCAAAGGATGGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTGTAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGAT
CACCCCCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5810 Amino Acid (SEQ ID NO: 770)

MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKI
DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFG
RPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5819 Nucleotide (SEQ ID NO: 771)

ATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAGCAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCCAAAGGATAGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTGTAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGAT
CACCCCCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5819 Amino Acid (SEQ ID NO: 772)

MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRIVRSGENALKI
DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFG
RPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5820 Nucleotide (SEQ ID NO: 773)

ATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAGCAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCCAAAGGGTGGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA

GGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTGTAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGAT
CACCCCCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5820 Amino Acid (SEQ ID NO: 774)

MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRVVRSGENALKI
DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFG
RPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5821 Nucleotide (SEQ ID NO: 775)

ATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAGCAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCCAAAGGTTTGTCCGGAGCGGTGAAAATGCCCTGAA
GATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGGCC
CAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGG
TGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCTGA
ACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCA
CTGTAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGATCA
CCCCCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5821 Amino Acid (SEQ ID NO: 776)

MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRFVRSGENALKI
DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFG
RPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

ATG-5822 Nucleotide (SEQ ID NO: 777)

ATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGAAGCAGACAGCCGCCTAC
AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGCTGCAGAATCTCG
CCGTGTCCGTAACTCCGATCCAAAGGCTGGTCCGGAGCGGTGAAAATGCCCTGA
AGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGCCGACCAAATGG
CCCAGATCGAAGAGGTGTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAA
GGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGCT
GAACTATTTCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGAT
CACTGTAACAGGGACCCTGTGGAACGGCGCCAAAATTATCGACGAGCGCCTGAT
CACCCCCGACGGCTCCATGTCCTTCCAGGTAACCATCAACAGC

ATG-5822 Amino Acid (SEQ ID NO: 778)

MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRLVRSGENALKI
DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFG
RPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

Monomeric LgBiT sequences:

ATG-5872 Nucleotide (SEQ ID NO: 779)

Atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgacgatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
aagctgaactatttcggacacccgtatgaaggcatcgccgtgttcgacggcgagaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5872 Amino Acid (SEQ ID NO: 780)

MKHHHHHHAIAMVFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNKLNYFGHPYEGIAVFDGEKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5873 Nucleotide (SEQ ID NO: 781)

atgaaacatcaccatcaccatcatgcgatcgccatggatttcacactcgacgatttcgttggggactgggaacagacagccgcctacaa
cctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgga
gcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagagg
tgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaaca
agctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcgagaagatcactgtaacagggaccctgtggaacggca
acaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgtt
caagaagatcagc

ATG-5873 Amino Acid (SEQ ID NO: 782)

MKHHHHHHAIAMDFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNKLNYFGRPYEGIAVFDGEKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5874 Nucleotide (SEQ ID NO: 783)

atgaaacatcaccatcaccatcatgcgatcgccatggatttcacactcgacgatttcgttggggactgggaacagacagccgcctacaa
cctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgga

gcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagagg
tgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaaca
agctgaactatttcggacacccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggca
acaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgtt
caagaagatcagc

ATG-5874 Amino Acid (SEQ ID NO: 784)

MKHHHHHHAIAMDFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNKLNYFGHPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5875 Nucleotide (SEQ ID NO: 785)

Atgaaacatcaccatcaccatcatgcgatcgccatggatttcacactcgacgatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
atgctgaactatttcggacacccgtatgaaggcatcgccgtgttcgacggcgagaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5875 Amino Acid (SEQ ID NO: 786)

MKHHHHHHAIAMDFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGHPYEGIAVFDGEKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5876 Nucleotide (SEQ ID NO: 787)

Atgaaacatcaccatcaccatcatgcgatcgccatggatttcacactcgaagatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
aagctgaactatttcggacacccgtatgaaggcatcgccgtgttcgacggcgagaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5876 Amino Acid (SEQ ID NO: 802)

MKHHHHHHAIAMDFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG

VTPNKLNYFGHPYEGIAVFDGEKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5877 Nucleotide (SEQ ID NO: 803)

Atgaaacatcaccatcaccatcatgcgatcgccatggatttcacactcgacgatttcgttggggactgggaacagacagccgcctaca
acctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgg
agcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagag
gtgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaac
aagctgaactatttcggacacccgtatgaaggcatcgccgtgttcgacggcgagaagatcactgtaacagggaccctgtggaacggc
aacaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgt
tcaagaagatcagc

ATG-5877 Amino Acid (SEQ ID NO: 804)

MKHHHHHHAIAMDFTLDDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNKLNYFGHPYEGIAVFDGEKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5878 Nucleotide (SEQ ID NO: 805)

Atgaaacatcatcaccatcaccacgcagaaatcggtactggctttccattcgacccccattatgtggaagtcctgggcgagcgcatgc
actacgtcgatgttggtccgcgcgatggcacccctgtgctgttcctgcacggtaacccgacctcctcctacgtgtggcgcaacatcatc
ccgcatgttgcaccgacccatcgctgcattgctccagacctgatcggtatgggcaaatccgacaaaccagacctgggttatttcttcgac
gaccacgtccgcttcatggatgccttcatcgaagccctgggtctggaagaggtcgtcctggtcattcacgactggggctccgctctggg
tttccactgggccaagcgcaatccagagcgcgtcaaaggtattgcatttatggagttcatccgccctatcccgacctgggacgaatggc
cagaatttgcccgcgagaccttccaggccttccgcaccaccgacgtcggccgcaagctgatcatcgatcagaacgtttttatcgagggt
acgctgccgatgggtgtcgtccgcccgctgactgaagtcgagatggaccattaccgcgagccgttcctgaatcctgttgaccgcgagc
cactgtggcgcttcccaaacgagctgccaatcgccggtgagccagcgaacatcgtcgcgctggtcgaagaatacatggactggctgc
accagtcccctgtcccgaagctgctgttctggggcaccccaggcgttctgatcccaccggccgaagccgctcgcctggccaaaagcc
tgcctaactgcaaggctgtggacatcggcccgggtctgaatctgctgcaagaagacaacccggacctgatcggcagcgagatcgcg
cgctggctgtcgacgctcgagatttccggcgagccaaccactagcgataacggctccatgctgttccgagtaaccatcaacagcgtta
gcggctggcgcctgttcaagaagatcagc

ATG-5878 Amino Acid (SEQ ID NO: 806)

MKHHHHHHAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVW
RNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHD
WGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLII
DQNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVAL
VEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQED
NPDLIGSEIARWLSTLEISGEPTTSDNGSMLFRVTINSVSGWRLFKKIS

ATG-5879 Nucleotide (SEQ ID NO: 807)

atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactggaagcagacagccgcctacaa
cctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgga
gcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagagg
tgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacat
gctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaa
caaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgttc
aagaagatcagc

ATG-5879 Amino Acid (SEQ ID NO: 808)

MKHHHHHHAIAMVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5880 Nucleotide (SEQ ID NO: 809)

atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactggcgccagacagccgcctacaa
cctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgga
gcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagagg
tgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacat
gctgaactatttcggacggccgtatgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaa
caaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgttc
aagaagatcagc

ATG-5880 Amino Acid (SEQ ID NO: 813)

MKHHHHHHAIAMVFTLEDFVGDWRQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5881 Nucleotide (SEQ ID NO: 814)

atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaa
cctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgga
gcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagagg
tgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacat
gctgaactatttcggacggccgttcgaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggcaa
caaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgttc
aagaagatcagc

ATG-5881 Amino Acid (SEQ ID NO: 815)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPFEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5882 Nucleotide (SEQ ID NO: 828)

atgaaacatcaccatcaccatcatgcgatcgccatggtcttcacactcgaagatttcgttggggactgggaacagacagccgcctacaa
cctggaccaagtccttgaacagggaggtgtgtccagtttgctgcagaatctcgccgtgtccgtaactccgatccaaaggattgtccgga
gcggtgaaaatgccctgaagatcgacatccatgtcatcatcccgtatgaaggtctgagcgccgaccaaatggcccagatcgaagagg
tgtttaaggtggtgtaccctgtggatgatcatcactttaaggtgatcctgccctatggcacactggtaatcgacggggttacgccgaacat
gctgaactatttcggacggccgtgggaaggcatcgccgtgttcgacggcaaaaagatcactgtaacagggaccctgtggaacggca
acaaaattatcgacgagcgcctgatcacccccgacggctccatgctgttccgagtaaccatcaacggcgttagcggctggcgcctgtt
caagaagatcagc

ATG-5882 Amino Acid (SEQ ID NO: 829)

MKHHHHHHAIAMVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQ
RIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTLVIDG
VTPNMLNYFGRPWEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTINGVS
GWRLFKKIS

ATG-5546 IL6-3xFlag-optRBD-15GS-SmTrip9 (840) Nucleotide (SEQ ID NO: 971)

atgaactccttctccacaagcgccttcggtccagttgccttctccctgggcctgctcctggtgttgcctgctgccttccctgccccagacta
caaagaccatgacggtgattataaagatcatgacatcgattacaaggatgacgatgacaagagagtccaacctactgaatctattgttag
atttcctaatattactaatctttgtccttttggcgaggtgttcaatgccaccagattcgcctctgtgtacgcctggaaccggaagcggatca
gcaattgcgtggccgactactccgtgctgtacaactccgccagcttcagcaccttcaagtgctacggcgtgtcccctaccaagctgaac
gacctgtgcttcacaaacgtgtacgccgacagcttcgtgatccggggagatgaagtgcggcagattgcccctggacagacaggcaa
gatcgccgactacaactacaagctgcccgacgacttcaccggctgtgtgattgcctggaacagcaacaacctggactccaaagtcgg
cggcaactacaattacctgtaccggctgttccggaagtccaatctgaagcccttcgagcgggacatctccaccgagatctatcaggcc
ggcagcacccctgtaacggcgtggaaggcttcaactgctacttcccactgcagtcctacggctttcagcccacaaatggcgtgggct
atcagccctacagagtggtggtgctgagcttcgaactgctgcatgcccctgccacagtgtgcggccctaagaaaagcaccaatctcgt
gaagaacaaatgcgtgaacttcgggagctccggtggtggcgggagcggaggtggaggctcgagcggtggcaagctcctgttcacg
gtaaccatcgagaagtataag

ATG-5546 Amino Acid (SEQ ID NO: 972)

MNSFSTSAFGPVAFSLGLLLVLPAAFPAPDYKDHDGDYKDHDIDYKDDDDKRVQPT
ESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYG
VSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNS
NNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSY
GFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFGSSGGGGSGG
GGSSGGKLLFTVTIEKYK

ATG-5541 nucleotide IL6-VSHiBiT-optRBD-3xFlag (SEQ ID NO: 973)

atgaactccttctccacaagcgccttcggtccagttgccttctccctgggcctgctcctggtgttgcctgctgccttccctgccccagtctc
cgtgagcggctggcggctgttcaagaagattagcgggagctccggtggtggcgggagcggaggtggaggctcgagcggtagagt
ccaacctactgaatctattgttagatttcctaatattactaatctttgtccttttggcgaggtgttcaatgccaccagattcgcctctgtgtacg
cctggaaccggaagcggatcagcaattgcgtggccgactactccgtgctgtacaactccgccagcttcagcaccttcaagtgctacgg
cgtgtcccctaccaagctgaacgacctgtgcttcacaaacgtgtacgccgacagcttcgtgatccggggagatgaagtgcggcagatt
gcccctggacagacaggcaagatcgccgactacaactacaagctgcccgacgacttcaccggctgtgtgattgcctggaacagcaa
caacctggactccaaagtcggcggcaactacaattacctgtaccggctgttccggaagtccaatctgaagcccttcgagcgggacatc
tccaccgagatctatcaggccggcagcacccccttgtaacggcgtggaaggcttcaactgctacttcccactgcagtcctacggctttca
gcccacaaatggcgtgggctatcagccctacagagtggtggtgctgagcttcgaactgctgcatgcccctgccacagtgtgcggccct
aagaaaagcaccaatctcgtgaagaacaaatgcgtgaacttcgactacaaagaccatgacggtgattataaagatcatgacatcgatta
caaggatgacgatgacaag

ATG-5541 Amino Acid (SEQ ID NO: 974)

MNSFSTSAFGPVAFSLGLLLVLPAAFPAPVSVSGWRLFKKISGSSGGGGSGGGGSSGR
VQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTF
KCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVI
AWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFP
LQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFDYKDH
DGDYKDHDIDYKDDDDK

ATG-5547 Nucleotide IL6-SmTrip9 (840)-optRBD-3xFlag (SEQ ID NO: 975)

atgaactccttctccacaagcgccttcggtccagttgccttctccctgggcctgctcctggtgttgcctgctgccttccctgccccaggca
agctcctgttcacggtaaccatcgagaagtataaggggagctccggtggtggcgggagcggaggtggaggctcgagcggtagagt
ccaacctactgaatctattgttagatttcctaatattactaatctttgtccttttggcgaggtgttcaatgccaccagattcgcctctgtgtacg
cctggaaccggaagcggatcagcaattgcgtggccgactactccgtgctgtacaactccgccagcttcagcaccttcaagtgctacgg
cgtgtcccctaccaagctgaacgacctgtgcttcacaaacgtgtacgccgacagcttcgtgatccggggagatgaagtgcggcagatt
gcccctggacagacaggcaagatcgccgactacaactacaagctgcccgacgacttcaccggctgtgtgattgcctggaacagcaa
caacctggactccaaagtcggcggcaactacaattacctgtaccggctgttccggaagtccaatctgaagcccttcgagcgggacatc
tccaccgagatctatcaggccggcagcacccccttgtaacggcgtggaaggcttcaactgctacttcccactgcagtcctacggctttca
gcccacaaatggcgtgggctatcagccctacagagtggtggtgctgagcttcgaactgctgcatgcccctgccacagtgtgcggccct
aagaaaagcaccaatctcgtgaagaacaaatgcgtgaacttcgactacaaagaccatgacggtgattataaagatcatgacatcgatta
caaggatgacgatgacaag

ATG-5547 Amino Acid (SEQ ID NO: 976)

MNSFSTSAFGPVAFSLGLLLVLPAAFPAPGKLLFTVTIEKYKGSSGGGGSGGGGSSGR
VQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTF
KCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVI
AWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFP
LQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFDYKDH
DGDYKDHDIDYKDDD

ATG-5534 Nucleotide (SEQ ID NO: 977)

atggactcgggccgggacttcctgaccctgcacggcctacaggatgatgaggatctacaggcgctgctgaagggcagccagctcct
gaaggtgaagtccagctcatggaggagagagcggttctacaagttgcaggaggactgcaagaccatctggcaggagtcccgcaag
gtcatgcggaccccggagtcccagctgttctccatcgaggacattcaggaggtgcgaatggggcaccgcacggagggtctggagaa
gttcgcccgtgatgtgcccgaggaccgctgcttctccattgtcttcaaggaccagcgcaatacactagacctcatcgccccatcgccag
ctgatgcccagcactgggtgctggggctgcacaagatcatccaccactcaggctccatggaccagcgtcagaagctacagcactgg
attcactcctgcttgcgaaaagctgacaaaaacaaggacaacaagatgagcttcaaggagctgcagaacttcctgaaggagctcaac
atccagctcggctcttcaggcgcagaaatcggtactggctttccattcgacccccattatgtggaagtcctgggcgagcgcatgcacta
cgtcgatgttggtccgcgcgatggcacccctgtgctgttcctgcacggtaacccgacctcctcctacgtgtggcgcaacatcatcccgc
atgttgcaccgacccatcgctgcattgctccagacctgatcggtatgggcaaatccgacaaaccagacctgggttatttcttcgacgacc
acgtccgcttcatggatgccttcatcgaagccctgggtctggaagaggtcgtcctggtcattcacgactggggctccgctctgggtttcc
actgggccaagcgcaatccagagcgcgtcaaaggtattgcatttatggagttcatccgccctatcccgacctgggacgaatggccaga
atttgcccgcgagaccttccaggccttccgcaccaccgacgtcggccgcaagctgatcatcgatcagaacgtttttatcgagggtacgc
tgccgatgggtgtcgtccgcccgctgactgaagtcgagatggaccattaccgcgagccgttcctgaatcctgttgaccgcgagccact
gtggcgcttcccaaacgagctgccaatcgccggtgagccagcgaacatcgtcgcgctggtcgaagaatacatggactggctgcacc
agtcccctgtcccgaagctgctgttctggggcaccccaggcgttctgatcccaccggccgaagccgctcgcctggccaaaagcctgc
ctaactgcaaggctgtggacatcggcccgggtctgaatctgctgcaagaagacaacccggacctgatcggcagcgagatcgcgcgc
tggctgtcgacgctggagatttccggcggctcgagcggtgtcTTCACTTTGGAAGATTTCGTGGGTGACTG
GAAACAGACTGCAGCATACAACTTAGACCAAGTGTTGGAACAGGGTGGAGTTAG
TTCGCTTTTACAGAATTTGGCAGTTAGTGTTACGCCTATACAACGTATGGTGAGA
TCGGGAGAAAATGCATTAAAAATAGACATACATGTGATAATACCTTATGAAGGA
TTATCGGCAGACCAAATGGCACAGATAGAAGAGGTTTTTAAAGTTGTTTACCCAG
TTGATGATCATCACTTTAAAGTTATATTACACTATGGAACTTTAGTTATAGACGGT
GTGACTCCTAACATGTTAAACTATTTCGGTAGACCTTATGAAGGAATAGCAGTTT
TCGACGGAAAGAAAATAACGGTTACTGGTACGTTATGGAACGGAGCCAAGATAA
TAGACGAGAGATTAATAACGCCAGACGGAAGTATGAGTTTCCAGGTTACGAtaaact
cg

ATG-5534 Amino Acid (SEQ ID NO: 978)

MDSGRDFLTLHGLQDDEDLQALLKGSQLLKVKSSSWRRERFYKLQEDCKTIWQESR
KVMRTPESQLFSIEDIQEVRMGHRTEGLEKFARDVPEDRCFSIVFKDQRNTLDLIAPSP
ADAQHWVLGLHKIIHHSGSMDQRQKLQHWIHSCLRKADKNKDNKMSFKELQNFLK
ELNIQLGSSGAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVW
RNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHD
WGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLII
DQNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVAL
VEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQED
NPDLIGSEIARWLSTLEISGGSSGVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQ
NLAVSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKV
ILHYGTLVIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGS
MSFQVTINS

LgBiT* Amino Acid (SEQ ID NO: 979)

MVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKI
DIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFG
RPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

HaloTag-LgBiT* Amino Acid (SEQ ID NO: 980)

MAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPHVAP
THRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWGSALGFH
WAKRNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQNVFIEGT
LPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDW
LHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPDLIGSEI
ARWLSTLEISGGSSGVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPI
QRMVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVI
DGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS

LgBiT*-Membrane Sensor Amino Acid (SEQ ID NO: 981)

MDSGRDFLTLHGLQDDEDLQALLKGSQLLKVKSSSWRRERFYKLQEDCKTIWQESR
KVMRTPESQLFSIEDIQEVRMGHRTEGLEKFARDVPEDRCFSIVFKDQRNTLDLIAPSP
ADAQHWVLGLHKIIHHSGSMDQRQKLQHWIHSCLRKADKNKDNKMSFKELQNFLK
ELNIQLGSSGAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVW
RNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHD
WGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLII
DQNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVAL
VEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQED
NPDLIGSEIARWLSTLEISGGSSGVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQ
NLAVSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKV
ILHYGTLVIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGS
MSFQVTINS

LgBiT*-Nuclear Sensor Amino Acid (SEQ ID NO: 982)

MPKKKRKVAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWR
NIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDW
GSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQ
NVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVE
EYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNP
DLIGSEIARWLSTLEISGGSSGVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNL
AVSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVIL
HYGTLVIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMS
FQVTINSPKKKRKVPKKKRKVPPPVSKRESKSRSRSKSPPKSPEEEGAVSS

LgBiT*-Mitochondria Sensor Amino Acid (SEQ ID NO: 983)

MSVLTPLLLRGLTGSARRLPVPRAKIHSLGDPMSVLTPLLLRGLTGSARRLPVPRAKI
HSLGKLMSVLTPLLLRGLTGSARRLPVPRAKIHSLGDPMSVLTPLLLRGLTGSARRLP
VPRAKIHSLGKLRILQSTVPRARDPPVATMAEIGTGFPFDPHYVEVLGERMHYVDVG
PRDGTPVLFLHGNPTSSYVWRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVR
FMDAFIEALGLEEVVLVIHDWGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPE
FARETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREP
LWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKS
LPNCKAVDIGPGLNLLQEDNPDLIGSEIARWLSTLEISGGSSGVFTLEDFVGDWKQTA

AYNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMA
QIEEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPYEGIAVFDGKKITVT
GTLWNGAKIIDERLITPDGSMSFQVTINS

LgBiT*-Endoplasmic Reticulum Sensor Amino Acid (SEQ ID NO: 984)

MLLSVPLLLGLLGLAVAVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQNLAVSV
TPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYGTL
VIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTI
NSGSSGMAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWRN
IIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWG
SALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQN
VFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEE
YMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPD
LIGSEIARWLSTLEISGKDEL

LgBiT*-Golgi Sensor Amino Acid (SEQ ID NO: 985)

MAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPHVAP
THRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWGSALGFH
WAKRNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQNVFIEGT
LPMGVVRPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDW
LHQSPVPKLLFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPDLIGSEI
ARWLSTLEISGGSSGVFTLEDFVGDWKQTAAYNLDQVLEQGGVSSLLQNLAVSVTPI
QRMVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILHYGTLVI
DGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGAKIIDERLITPDGSMSFQVTINS
SGLRSRGEPQQSFSEAQQQLCNTRQEVNELRKLLEEERDQRVAAENALSVAEEQIRR
LEHSEWDSSRTPIIGSCGTQEQALLIDLTSNSCRRTRSGVGWKRVLRSLCHSRTRVPL
LAAIYFLMIHVLLILCFTGHL

LgBiT*-Lysosome Sensor Amino Acid (SEQ ID NO: 986)

MAAPGSARRPLLLLLLLLLLGLMHCASAAMFMVKNGNGTACIMANFSAAFSVNYD
TKSGPKNMTFDLPSDATVVLNRSSCGKENTSDPSLVIAFGRGHTLTLNFTRNATRYS
VQLMSFVYNLSDTHLFPNASSKEIKTVESITDIRADIDKKYRCVSGTQVHMNNVTVT
LHDATIQAYLSNSSFSRGETRCEQDRPSPTTAPPAPPSPSPSPVPKSPSVDKYNVSGTN
GTCLLASMGLQLNLTYERKDNTTVTRLLNINPNKTSASGSCGAHLVTLELHSEGTTV
LLFQFGMNASSSRFFLQGIQLNTILPDARDPAFKAANGSLRALQATVGNSYKCNAEE
HVRVTKAFSVNIFKVWVQAFKVEGGQFGSVEECLLDENSMLIPIAVGGALAGLVLIV
LIAYLVGRKRSHAGYQTIKDPPVATMAEIGTGFPFDPHYVEVLGERMHYVDVGPRD
GTPVLFLHGNPTSSYVWRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFM
DAFIEALGLEEVVLVIHDWGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPEFA
RETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPLTEVEMDHYREPFLNPVDREPL
WRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWGTPGVLIPPAEAARLAKSL
PNCKAVDIGPGLNLLQEDNPDLIGSEIARWLSTLEISGGSSGVFTLEDFVGDWKQTAA
YNLDQVLEQGGVSSLLQNLAVSVTPIQRMVRSGENALKIDIHVIIPYEGLSADQMAQI
EEVFKVVYPVDDHHFKVILHYGTLVIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGT
LWNGAKIIDERLITPDGSMSFQVTINS

CLAIMS

1. A peptide comprising an amino acid sequence having greater than 40% but less than 100% sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6, SEQ ID NO: 9, and/or SEQ ID NO: 29, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the peptide contacts a second peptide consisting of SEQ ID NO: 25 and a polypeptide complement consisting of SEQ ID NO: 17 when compared to a bioluminescent signal produced by the peptide and the coelenterazine substrate alone.

2. The peptide of claim 1, wherein the bioluminescent signal is substantially increased when the peptide associates with the second peptide and the polypeptide complement.

3. The peptide of claim 1, wherein the peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 6 and/or SEQ ID NO: 9, wherein the traits are selected from: affinity for the second peptide and the polypeptide complement, expression, solubility, stability, and bioluminescent activity when combined with the second peptide and the polypeptide complement.

4. The peptide of claim 1, wherein the amino acid sequence is not a naturally occurring protein or a fragment thereof.

5. The peptide of claim 4, wherein the amino acid sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

6. A nucleic acid comprising a sequence coding for a peptide of claim 1.

7. A fusion polypeptide comprising the peptide of claim 1 and an additional amino acid sequence.

8. The fusion polypeptide of claim 7, wherein the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety.

9. The fusion polypeptide of claim 8, wherein the additional amino acid sequence is a binding moiety selected from the group consisting of antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

10. The fusion polypeptide of claim 8, wherein the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide.

11. The fusion polypeptide of claim 8, wherein the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide.

12. The fusion polypeptide of claim 8, wherein the additional amino acid sequence is a protein of interest and is a candidate drug target.

13. A nucleic acid comprising a sequence coding for the fusion polypeptide of claim 7.

14. A peptide comprising an amino acid sequence having greater than 40% but less than 100% sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the peptide contacts a second peptide consisting of SEQ ID NO: 23 and a polypeptide complement consisting of SEQ ID NO: 17 when compared to a bioluminescent signal produced by the peptide and the coelenterazine substrate alone.

15. The peptide of claim 14, wherein the bioluminescent signal is substantially increased when the peptide associates with the second peptide and the polypeptide complement.

16. The peptide of claim 14, wherein the peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 7 and/or SEQ ID NO: 10, wherein the traits are selected from: affinity for the second peptide and the polypeptide complement, expression, solubility, stability, and bioluminescent activity when combined with the second peptide and the polypeptide complement.

17. The peptide of claim 14, wherein the amino acid sequence is not a naturally occurring protein or a fragment thereof.

18. The peptide of claim 17, wherein the amino acid sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

19. A nucleic acid comprising a sequence coding for a peptide of claim 14.

20. A fusion polypeptide comprising the peptide of claim 14 and an additional amino acid sequence.

21. The fusion polypeptide of claim 20, wherein the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety.

22. The fusion polypeptide of claim 21, wherein the additional amino acid sequence is a binding moiety selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified

protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

23. The fusion polypeptide of claim 21, wherein the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide.

24. The fusion polypeptide of claim 21, wherein the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide.

25. The fusion polypeptide of claim 21, wherein the additional amino acid sequence is a protein of interest and is a candidate drug target.

26. A nucleic acid comprising a sequence coding for the fusion polypeptide of claim 20.

27. A composition comprising:

(a) a first peptide comprising an amino acid sequence having greater than 40% but less than 100% sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10; and

(b) a second peptide comprising an amino acid sequence having greater than 40% but less than 100% sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6, SEQ ID NO: 9, and SEQ ID NO: 29;

wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the first peptide contacts the second peptide and a polypeptide complement consisting of SEQ ID NO: 17 when compared to a bioluminescent signal produced by the first peptide and/or the second peptide and the coelenterazine substrate alone.

28. The composition of claim 27, wherein the bioluminescent signal is substantially increased when the first peptide associates with the second peptide and the polypeptide complement.

29. The composition of claim 27, wherein the first peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 7 and/or SEQ ID NO: 10 and the second peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 6, SEQ ID NO: 9, and SEQ ID NO: 29, wherein the traits are selected from: affinity for the second peptide and the polypeptide complement, expression, solubility, stability, and bioluminescent activity when combined with the second peptide and the polypeptide complement.

30. The composition of claim 27, wherein the amino acid sequence of the first and/or second peptide is not a naturally occurring protein or a fragment thereof.

31. The composition of claim 30, wherein the amino acid sequence of the first and/or second peptide contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

32. A composition comprising nucleic acid comprising sequence coding for the first and second peptides of claim 27.

33. A composition comprising fusion polypeptides comprising the first and second peptides of claim 27 and an additional amino acid sequence.

34. The composition of claim 33, wherein the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety.

35. The composition of claim 34, wherein the additional amino acid sequence is a binding moiety selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L,

a Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

36. The composition of claim 34, wherein the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide.

37. The composition of claim 34, wherein the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide.

38. The composition of claim 34, wherein the additional amino acid sequence is a protein of interest and is a candidate drug target.

39. A composition comprising nucleic acid comprising sequences coding for the composition of claim 33.

40. A polypeptide comprising an amino acid sequence having greater than 40% but less than 100% sequence identity with SEQ ID NO: 17 and less than 100% sequence identity with SEQ ID NO: 5, SEQ ID NO: 8, and/or SEQ ID NO: 27, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the polypeptide contacts a first peptide consisting of SEQ ID NO: 23 and a second peptide consisting of SEQ ID NO: 25 when compared to a bioluminescent signal produced by the polypeptide and the coelenterazine substrate alone.

41. The polypeptide of claim 40, wherein the bioluminescent signal is substantially increased when the polypeptide associates with the first and second peptides.

42. The polypeptide of claim 40, wherein the polypeptide exhibits enhancement of one or more traits compared to a polypeptide of SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the traits are selected from: affinity for the first and/or second peptides, expression, solubility, stability, and bioluminescent activity when combined with the first and second peptides.

43. The polypeptide of claim 40, wherein the amino acid sequence is not a naturally occurring protein or a fragment thereof.

44. The polypeptide of claim 43, wherein the amino acid sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

45. A nucleic acid comprising a sequence coding for a polypeptide of claim 40.

46. A fusion polypeptide comprising the polypeptide of claim 40 and an additional amino acid sequence.

47. The fusion polypeptide of claim 46, wherein the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety.

48. The fusion polypeptide of claim 47, wherein the additional amino acid sequence is a binding moiety selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, a Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

49. The fusion polypeptide of claim 47, wherein the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction

polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide.

50. The fusion polypeptide of claim 47, wherein the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide.

51. The fusion polypeptide of claim 47, wherein the additional amino acid sequence is a protein of interest and is a candidate drug target.

52. A nucleic acid comprising a sequence coding for the fusion polypeptide of claim 46.

53. A polypeptide comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 51 and/or 302 and less than 100% sequence identity with SEQ ID NO: 5, SEQ ID NO: 8, and SEQ ID NO: 27, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the polypeptide contacts a first peptide consisting of SEQ ID NO: 23 and a second peptide consisting of SEQ ID NO: 25 when compared to a bioluminescent signal produced by the peptide and the coelenterazine substrate alone.

54. The polypeptide of claim 53, wherein the bioluminescent signal is substantially increased when the polypeptide associates with the first and second peptides.

55. The polypeptide of claim 53, wherein the polypeptide exhibits enhancement of one or more traits compared to a polypeptide of SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the traits are selected from: affinity for the first and/or second peptides, expression, solubility, stability, and bioluminescent activity when combined with the first and second peptides.

56. The polypeptide of claim 53, wherein the amino acid sequence is not a naturally occurring protein or a fragment thereof.

57. The polypeptide of claim 56, wherein the amino acid sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

58. A nucleic acid comprising a sequence coding for a polypeptide of claim 53.

59. A fusion polypeptide comprising the polypeptide of claim 53 and an additional amino acid sequence.

60. The fusion polypeptide of claim 59, wherein the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety.

61. The fusion polypeptide of claim 60, wherein the additional amino acid sequence is a binding moiety selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, a Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

62. The fusion polypeptide of claim 60, wherein the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide.

63. The fusion polypeptide of claim 60, wherein the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide.

64. The fusion polypeptide of claim 60, wherein the additional amino acid sequence is a protein of interest and is a candidate drug target.

65. A nucleic acid comprising a sequence coding for the fusion polypeptide of claim 59.

66. A bioluminescent complex comprising:

(a) a polypeptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, 21, or 302 and less than 100% sequence identity with SEQ ID NO: 5 and/or SEQ ID NO: 8;

(b) a first peptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and/or SEQ ID NO: 9; and

(c) a second peptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10;

wherein the bioluminescent complex produces substantially increased bioluminescence in the presence of a coelenterazine or a coelenterazine analog substrate when compared to a coelenterazine or a coelenterazine analog substrate in the presence of: the polypeptide alone, the first peptide alone, the second peptide alone, and any two of the polypeptide, the first peptide and the second peptide.

67. The bioluminescent complex of claim 66, wherein the first peptide is a first peptide tag, wherein the second peptide is a second peptide tag, and wherein the first and second peptide tags are each linked to moieties that are independently selected from the group consisting of a molecule of interest, a peptide of interest, a protein of interest, an interaction element, a co-localization element, or a binding moiety.

68. The bioluminescent complex of claim 67, wherein the first peptide tag or the second peptide tag is linked to a drug or drug candidate, and the other peptide tag is linked to a drug

target or candidate drug target, and wherein the intensity of the bioluminescence from the bioluminescent complex correlates to the affinity of the drug or drug candidate for the drug target or candidate drug target.

69. The bioluminescent complex of claim 67, wherein the first peptide tag is linked to a first interaction element and the second peptide tag is linked to a second interaction element, and wherein the intensity of the bioluminescence from the bioluminescent complex correlates to the affinity of the first interaction element for the second interaction element under the conditions assayed.

70. The bioluminescent complex of claim 67, wherein the first peptide tag is linked to a first co-localization element and the second peptide tag is linked to a second co-localization element, and wherein substantially increased bioluminescence indicates co-localization, but not necessarily interaction, of the first co-localization element and the second co-localization element, under the conditions assayed.

71. A method comprising:

(a) combining:

(i) a first peptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9,

(ii) a second peptide comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10,

(iii) a polypeptide component comprising an amino acid sequence having 40% or greater (e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more) sequence identity with SEQ ID NO: 17, 21, and/or 302 and less than 100% sequence identity with SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to

produce a bioluminescent complex upon interaction of the first molecular entity and the second molecular entity, and

(iv) a coelenterazine or a coelenterazine analog substrate; and

(b) detecting luminescence, wherein a greater level of luminescence compared to a level of luminescence produced by the polypeptide component and a coelenterazine or a coelenterazine analog alone indicates formation of a bioluminescent complex of the polypeptide component and the first and second peptides.

72. The method of claim 71, wherein one or more of the polypeptide component and the first and second peptides are expressed in a cell, added to a cell exogenously, and/or added to a sample.

73. A method of detecting an interaction between a first molecular entity and a second molecular entity, the method comprising:

(a) tagging the first molecular entity with a first peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and/or SEQ ID NO: 9;

(b) tagging the second molecular entity with a second peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10;

(c) combining the tagged first molecular entity and the tagged second molecular entity;

(d) adding a polypeptide component comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 17, 21, and/or 302 and less than 100% sequence identity with SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon interaction of the first molecular entity and the second molecular entity;

(e) adding a coelenterazine or a coelenterazine analog substrate; and

(f) detecting a luminescent signal produced by the bioluminescent complex, wherein the magnitude of the luminescent signal correlates to the strength of the interaction between the first molecular entity and the second molecular entity.

74. The method of claim 73, wherein the first molecular entity and/or the second molecular entity is a protein of interest or a peptide of interest, and tagging comprises generating a fusion of the first molecular entity and/or the second molecular entity with the first peptide tag and/or second peptide tag.

75. The method of claim 73, wherein the first molecular entity and/or the second molecular entity is a small molecule and tagging comprises directly or indirectly linking the first molecular entity and/or the second molecular entity with the first peptide tag and/or second peptide tag.

76. The method of claim 73, wherein one of the first molecular entity and the second molecular entity is a drug or drug candidate and the other is a drug target or candidate drug target, and the bioluminescent signal indicates binding of the drug or drug candidate to the other is a drug target or candidate drug target.

77. The method of claim 73, wherein combining the tagged first molecular entity and the tagged second molecular entity comprises expressing one or both within a cell and/or adding one or both to a cell.

78. A method of detecting an interaction between a first protein or peptide entity and a second protein or peptide entity with a cell comprising, the method comprising:

(a) expressing within the cell a fusion comprising the first protein or peptide entity and a first peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9;

(b) expressing within the cell a fusion comprising the second protein or peptide entity and a second peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10;

(c) expressing with the cell a polypeptide component comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 17, 21, and/or 302 and

less than 100% sequence identity with SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon interaction of the first protein or peptide entity and the second protein or peptide entity;

(d) adding a coelenterazine or a coelenterazine analog substrate to the cell; and

(e) detecting a luminescent signal produced by the bioluminescent complex, wherein the magnitude of the luminescent signal correlates to the strength of the interaction between the first protein or peptide entity and the second protein or peptide entity.

79. A method of detecting co-localization of a first molecular entity and a second molecular entity, the method comprising:

(a) tagging the first molecular entity with a first peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and/or SEQ ID NO: 9;

(b) tagging the second molecular entity with a second peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10;

(c) combining the tagged first molecular entity and the tagged second molecular entity in the same system;

(d) adding a polypeptide component to the system, the polypeptide components comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 17, 21, and/or 302 and less than 100% sequence identity with SEQ ID NO: 5 and SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon co-localization of the first molecular entity and the second molecular entity;

(e) adding a coelenterazine or a coelenterazine analog substrate to the system; and

(f) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates co-localization of the first molecular entity and the second molecular entity within the system, and/or wherein the magnitude of the luminescent signal correlates to the amount of co-localization within the system of the first molecular entity and the second molecular entity.

80. The method of claim 79, wherein the system comprises a cell, tissue, organ, whole organism, a biochemical, non-cellular sample.

81. The method of claim 79, wherein the first molecular entity and/or the second molecular entity is a protein of interest or a peptide of interest, and tagging comprises generating a fusion of the first molecular entity and/or the second molecular entity with the first peptide tag and/or second peptide tag.

82. The method of claim 79, wherein the first molecular entity and/or the second molecular entity is a small molecule and tagging comprises directly or indirectly linking the first molecular entity and/or the second molecular entity with the first peptide tag and/or second peptide tag.

83. The method of claim 79, wherein combining the tagged first molecular entity and the tagged second molecular entity comprises expressing one or both within the system and/or adding one or both to the system.

84. A method of detecting co-localization of a first protein or peptide entity and a second protein or peptide entity with a cell comprising, the method comprising:

(a) expressing within the cell a fusion comprising the first protein or peptide entity and a first peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and/or SEQ ID NO: 9;

(b) expressing within the cell a fusion comprising the second protein or peptide entity and a second peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10;

(c) expressing with the cell a polypeptide component comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 17, 21, and/or 302 and less than 100% sequence identity with SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce

a bioluminescent complex upon co-localization of the first protein or peptide entity and the second protein or peptide entity;

(d) adding a coelenterazine or a coelenterazine analog substrate to the cell; and

(e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates co-localization of the first protein or peptide entity and the second protein or peptide entity within the cell, and/or wherein the magnitude of the luminescent signal correlates to the amount of co-localization within the system of the first protein or peptide entity and the second protein or peptide entity.

85. A kit comprising:

(a) a first binding moiety conjugated to a first peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and/or SEQ ID NO: 9; and

(b) a second binding moiety conjugated to second peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10.

86. The kit of claim 85, wherein the first binding moieties are independently selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

87. The kit of claim 86, wherein the first and second binding moieties are primary binding moieties configured to bind to antigens, epitopes, or sequences on the same target entity.

88. The kit of claim 86, wherein the first and second binding moieties are secondary binding moieties configured to bind to antigens, epitopes, or sequences on primary binding moieties.

89. The kit of claim 86, further comprising a polypeptide reagent comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 17, 21, and/or 302 and less than 100% sequence identity with SEQ ID NO: 5 and/or SEQ ID NO: 8.

90. The kit of claim 89, further comprising a coelenterazine or a coelenterazine analog.

91. A method of detecting a target molecule, wherein the target molecule displays a first antigen, epitope, or sequence and a distinct second antigen, epitope, or sequence, the method comprising:

(a) contacting a sample containing the target molecule with (i) a first primary binding moiety that recognizes the first antigen, epitope, or sequence and (ii) a second primary binding moiety that recognizes the second antigen, epitope, or sequence, and allowing the first and second primary binding moieties to bind to the first and second antigens, epitopes, or sequences;

(b) contacting the sample with (i) a first secondary binding moiety conjugated to a first peptide tag and (ii) a second secondary binding moiety conjugated to second peptide tag, wherein the first secondary binding moiety recognizes the first primary binding moiety and the second secondary binding moiety recognizes the second primary binding moiety, wherein the first or second peptide tag comprises an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and/or SEQ ID NO: 9, and wherein the other of the first or second peptide tag comprises an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10, and allowing the first and second secondary binding moieties to bind to the first and second primary binding moieties;

(c) contacting the sample with comprising an polypeptide component having 40% or greater sequence identity with SEQ ID NO: 17, 21, and/or 302 and less than 100% sequence identity with SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon interaction;

(d) contacting the sample with a coelenterazine or a coelenterazine analog substrate; and

(e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates the presence of the target molecule, and/or wherein the magnitude of the luminescent signal correlates to the amount of target molecule within the sample.

92. The method of claim 91, wherein the binding moieties are independently selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

93. The method of claim 91, wherein the target molecule is a protein, nucleic acid, or small molecule.

94. The method of claim 91, wherein the sample is *in vitro* or *in vivo*.

95. A method of detecting a target molecule, wherein the target molecule displays a first antigen, epitope, or sequence and a distinct second antigen, epitope, or sequence, the method comprising:

(a) contacting the sample with (i) a first binding moiety conjugated to a first peptide tag and (ii) a second binding moiety conjugated to second peptide tag, wherein the first secondary binding moiety recognizes the first antigen, epitope, or sequence and the second binding moiety recognizes the second antigen, epitope, or sequence, wherein the first or second peptide tag comprises an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and/or SEQ ID NO: 9, and wherein the other of the first or second peptide tag comprises an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and/or SEQ ID NO: 10, and allowing the

first and second binding moieties to bind to the first and second antigens, epitope, or sequences;

(c) contacting the sample with comprising an polypeptide component having 40% or greater sequence identity with SEQ ID NO: 17, 21, and/or 302 and less than 100% sequence identity with SEQ ID NO: 5 and/or SEQ ID NO: 8, wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex upon interaction;

(d) contacting the sample with a coelenterazine or a coelenterazine analog substrate; and

(e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates the presence of the target molecule, and/or wherein the magnitude of the luminescent signal correlates to the amount of target molecule within the sample.

96. The method of claim 95, wherein the binding moieties are independently selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

97. The method of claim 95, wherein the target molecule is a protein, nucleic acid, or small molecule.

98. The method of claim 95, wherein the sample is *in vitro*, *in vivo,* or a biochemical sample.

99. A β9/β10-like dipeptide comprising an amino acid sequence having greater than 40% but less than 100% sequence identity with SEQ ID NO: 35 and less than 100% sequence identity with SEQ ID NO: 205 and SEQ ID NO: 206, wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is

substantially increased when the peptide contacts a polypeptide complement consisting of SEQ ID NO: 17, 21, and/or 302 when compared to a bioluminescent signal produced by the peptide and the coelenterazine substrate alone.

100. The β9/β10-like dipeptide of claim 99, wherein the bioluminescent signal is substantially increased when the peptide associates with the polypeptide complement.

101. The β9/β10-like dipeptide of claim 99, wherein the peptide exhibits enhancement of one or more traits compared to a peptide of SEQ ID NO: 205 and/or SEQ ID NO: 206, wherein the traits are selected from: affinity for the polypeptide complement, expression, solubility, stability, and bioluminescent activity when combined with the the polypeptide complement.

102. The β9/β10-like dipeptide of claim 99, wherein the amino acid sequence is not a naturally occurring protein or a fragment thereof.

103. The β9/β10-like dipeptide of claim 102, wherein the amino acid sequence contains a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

104. A nucleic acid comprising a sequence coding for a β9/β10-like dipeptide of claim 99.

105. A fusion polypeptide comprising the β9/β10-like dipeptide of claim 99 and an additional amino acid sequence.

106. The fusion polypeptide of claim 105, wherein the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety.

107. The fusion polypeptide of claim 106, wherein the additional amino acid sequence is a binding moiety selected from the group consisting of antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, a Ig binding domain of protein L, protein M, an Ig binding domain of protein

M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

108. The fusion polypeptide of claim 106, wherein the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide.

109. The fusion polypeptide of claim 106, wherein the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide.

110. The fusion polypeptide of claim 106, wherein the additional amino acid sequence is a protein of interest and is a candidate drug target.

111. A nucleic acid comprising a sequence coding for the fusion polypeptide of claim 105.

112. A method of performing a competition assay to detect an interaction between a first molecular entity and a second molecular entity, the method comprising:

(a) combining:

(i) a tracer comprising the first molecular entity tagged with a first peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 23 and less than 100% sequence identity with SEQ ID NO: 6 and SEQ ID NO: 9,

(ii) the second molecular entity tagged with a second peptide tag comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 25 and less than 100% sequence identity with SEQ ID NO: 7 and SEQ ID NO: 10,

(iii) a coelenterazine or a coelenterazine analog substrate,

(iv) a polypeptide component comprising an amino acid sequence having 40% or greater sequence identity with SEQ ID NO: 17, 21, and/or 302 and less than 100% sequence identity with SEQ ID NO: 5 and/or SEQ ID NO: 8, and

(v) a sample suspected of containing untagged first molecular entity;

wherein the first peptide tag, the second peptide tag, and the polypeptide component are configured to produce a bioluminescent complex and produce a bioluminescent signal in the presence of the coelenterazine or a coelenterazine analog substrate

(b) detecting the bioluminescent signal produced by the bioluminescent complex;

(c) comparing the bioluminescent signal produced in the presence of the sample with a control bioluminescent signal produced in the absence of the sample, wherein a decrease in the biluminescent signal indicates the presence or amount of untagged first molecular entity int the sample.

113. The method of claim 112, wherein the first molecular entity is a small molecule or peptide.

114. The method of claim 112, wherein the second molecular entity is a drug target or candidate drug target.

115. A system or kit comprising comprising:

(a) a polypeptide component comprising 40% or greater sequence identity to a polypeptide fragment of SEQ ID NO: 788 or SEQ ID NO: 789; and

(b) one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprising 40% or greater sequence identity to the complementary portion of SEQ ID NO: 788 or SEQ ID NO: 789;

wherein a bioluminescent signal produced by a bioluminescent complex assembled from the polypeptide component and one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when compared to a bioluminescent signal produced by the polypeptide component or one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides and the coelenterazine substrate alone.

116. The system or kit comprising of claim 115, wherein the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 790 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 794.

117. The system or kit comprising of claim 115, wherein the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 791 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 795.

118. The system or kit comprising of claim 115, wherein the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 792 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 796.

119. The system or kit comprising of claim 115, wherein the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 793 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 797.

120. The system or kit comprising of claim 115, wherein the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 790 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 798.

121. The system or kit comprising of claim 115, wherein the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 791 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 799.

122. The system or kit comprising of claim 115, wherein the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 792 and the one or more

complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 800.

123. The system or kit comprising of claim 115, wherein the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 793 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID 801.

124. The system or kit comprising of claim 115, wherein the bioluminescent signal is substantially increased when the polypeptide component associates with the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides.

125. The system or kit comprising of claim 115, wherein polypeptide component and/or one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides comprise amino acid sequences that are not a naturally occurring sequences or fragments thereof.

126. The system or kit comprising of claim 125, wherein polypeptide component and/or one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides comprise a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

127. The system or kit comprising of claim 115, wherein the polypeptide component and/or one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides are present as fusions with one or more additional amino acid sequences.

128. The system or kit comprising of claim 127, wherein the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety.

129. The system or kit comprising of claim 127, wherein the additional amino acid sequence is a binding moiety selected from the group consisting of antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain

of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

130. The system or kit comprising of claim 127, wherein the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide.

131. The system or kit comprising of claim 127, wherein the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide.

132. The system or kit comprising of claim 127, wherein the additional amino acid sequence is a protein of interest and is a candidate drug target.

133. A bioluminescent complex comprising the polypeptide component and one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides of the system or kit of one of claims 115-132.

134. A system or kit comprising two or more peptide, dipeptide, tripeptide and/or polypeptide components collectively comprising 40% or greater sequence identity to SEQ ID NO: 788 or SEQ ID NO: 789; wherein a bioluminescent signal produced by the bioluminescent complex in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when compared to a bioluminescent signal produced by the polypeptide or one or more complementary peptides and the coelenterazine substrate alone.

135. The system or kit of claim 134, comprising a polypeptide component having 40% or greater sequence identity to SEQ ID NO: 790 and one or more complementary peptides,

dipeptides, and or tripeptides collectively having 40% or greater sequence identity to SEQ ID NO: 794.

136. The system or kit of claim 134, wherein the polypeptide comprises 40% or greater sequence identity to SEQ ID NO: 791 and the one or more complementary peptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 795.

137. The system or kit of claim 134, wherein the polypeptide comprises 40% or greater sequence identity to SEQ ID NO: 792 and the one or more complementary peptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 796.

138. The system or kit of claim 134, wherein the polypeptide comprises 40% or greater sequence identity to SEQ ID NO: 793 and the one or more complementary peptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 797.

139. The system or kit of claim 134, wherein the polypeptide comprises 40% or greater sequence identity to SEQ ID NO: 790 and the one or more complementary peptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 798.

140. The system or kit of claim 134, wherein the polypeptide comprises 40% or greater sequence identity to SEQ ID NO: 791 and the one or more complementary peptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 799.

141. The system or kit of claim 134, wherein the polypeptide comprises 40% or greater sequence identity to SEQ ID NO: 792 and the one or more complementary peptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 800.

142. The system or kit of claim 134, wherein the polypeptide comprises 40% or greater sequence identity to SEQ ID NO: 793 and the one or more complementary peptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 801.

143. The system or kit of claim 134, wherein the bioluminescent signal is substantially increased when the polypeptide associates with the one or more complementary peptides.

144. The system or kit of claim 134, wherein polypeptide and/or one or more complementary peptides comprise amino acid sequences that are not a naturally occurring sequences or fragments thereof.

145. The system or kit of claim 144, wherein polypeptide and/or one or more complementary peptides comprise a non-natural amino acid, an amino acid analog, and/or peptoid amino acids.

146. The system or kit of claim 134, wherein the polypeptide and/or one or more complementary peptides are present as fusions with one or more additional amino acid sequences.

147. The system or kit of claim 146, wherein the additional amino acid sequence is selected from the group consisting of a protein of interest, an interaction element, a co-localization element, and a binding moiety.

148. The system or kit of claim 146, wherein the additional amino acid sequence is a binding moiety selected from the group consisting of antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

149. The system or kit of claim 146, wherein the additional amino acid sequence is a first interaction polypeptide that is configured to form a complex with a second interaction polypeptide upon contact of the first interaction polypeptide and the second interaction polypeptide.

150. The system or kit of claim 146, wherein the additional amino acid sequence is a first co-localization polypeptide that is configured to co-localize within a cellular compartment, a cell, a tissue, or an organism within a with a second co-localization polypeptide.

151. The system or kit of claim 146, wherein the additional amino acid sequence is a protein of interest and is a candidate drug target.

152. A method comprising:

(a) combining:

(i) a polypeptide component comprising 40% or greater sequence identity to a polypeptide fragment of SEQ ID NO: 788 or SEQ ID NO: 789;

(ii) one or more complementary peptides, dipeptide, tripeptides, and/or polypeptides collectively comprising 40% or greater sequence identity to the complementary portion of SEQ ID NO: 788 or SEQ ID NO: 789; and

(iii) a coelenterazine or a coelenterazine analog substrate; and

(b) detecting luminescence, wherein a greater level of luminescence compared to a level of luminescence produced by the polypeptide component and a coelenterazine or a coelenterazine analog alone indicates formation of a bioluminescent complex of the polypeptide component and the one or more complementary peptides.

153. The method of claim 152, wherein one or more of the polypeptide component and the first and second peptides are expressed in a cell, added to a cell exogenously, and/or added to a sample.

154. The method of claim 152, wherein:

(i) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 790 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 794;

(ii) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 791 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 795;

(iii) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 792 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 796;

(iv) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 793 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 797;

(v) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 790 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 798;

(vi) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 791 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 799;

(vii) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 792 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 800; or

(viii) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 793 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID 801.

155. A method comprising:

(a) combining:

(i) two or more peptide, dipeptide, tripeptide, and/or polypeptide components collectively comprising 40% or greater sequence identity to the full length of SEQ ID NO: 788 or SEQ ID NO: 789; and

(ii) a coelenterazine or a coelenterazine analog substrate; and

(b) detecting luminescence, wherein a greater level of luminescence compared to a level of luminescence produced by the peptide, dipeptide, tripeptide, and/or polypeptide components and a coelenterazine or a coelenterazine analog indicates formation of a bioluminescent complex of the peptide and polypeptide components.

156. The method of claim 155, wherein one or more of the polypeptide component and the first and second peptides are expressed in a cell, added to a cell exogenously, and/or added to a sample.

157. The method of claim 155, wherein:

(i) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 790 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 794;

(ii) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 791 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 795;

(iii) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 792 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 796;

(iv) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 793 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 797;

(v) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 790 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 798;

(vi) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 791 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 799;

(vii) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 792 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID NO: 800; or

(viii) the polypeptide component comprises 40% or greater sequence identity to SEQ ID NO: 793 and the one or more complementary peptides, dipeptides, tripeptide, and/or polypeptides collectively comprise 40% or greater sequence identity to SEQ ID 801.

158. A method of detecting an interaction between a first molecular entity and a second molecular entity, the method comprising:

(a) tagging the first molecular entity with a first peptide, dipeptide, or tripeptide tag;

(b) tagging the second molecular entity with a second peptide, dipeptide, or tripeptide tag;

(c) combining the tagged first molecular entity and the tagged second molecular entity and/or allowing the tagged first molecular entity and the tagged second molecular entity to come into contact with one another;

(d) adding peptide, dipeptide, tripeptide, and/or polypeptide components, wherein the first peptide, dipeptide, or tripeptide tag, the second peptide, dipeptide, or tripeptide tag, and the peptide, dipeptide, tripeptide, and/or polypeptide components collectively comprise an amino acid sequence having 40% or greater sequence identity with the entirety of SEQ ID NO: 788 or 789, and capable for assembling to form a bioluminescent complex;

(e) adding a coelenterazine or a coelenterazine analog substrate; and

(f) detecting a luminescent signal produced by the bioluminescent complex, wherein the magnitude of the luminescent signal correlates to the strength of the interaction between the first molecular entity and the second molecular entity.

159. The method of claim 158, wherein the first molecular entity and/or the second molecular entity is a protein of interest or a peptide of interest, and tagging comprises generating a fusion of the first molecular entity and/or the second molecular entity with the first tag and/or second tag.

160. The method of claim 158, wherein the first molecular entity and/or the second molecular entity is a small molecule and tagging comprises directly or indirectly linking the first molecular entity and/or the second molecular entity with the first tag and/or second tag.

161. The method of claim 158, wherein one of the first molecular entity and the second molecular entity is a drug or drug candidate and the other is a drug target or candidate drug target, and the bioluminescent signal indicates binding of the drug or drug candidate to the other is a drug target or candidate drug target.

162. The method of claim 158, wherein combining the tagged first molecular entity and the tagged second molecular entity comprises expressing one or both within a cell and/or adding one or both to a cell.

163. A method of detecting an interaction between a first protein or peptide entity and a second protein or peptide entity with a cell comprising, the method comprising:

(a) expressing within the cell a fusion comprising the first protein or peptide entity and a first peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater sequence identity with a first portion of SEQ ID NO: 788 or 789;

(b) expressing within the cell a fusion comprising the second protein or peptide entity and a second peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater sequence identity with a second portion of SEQ ID NO: 788 or 789;

(c) expressing with the cell peptide, dipeptide, tripeptide, and/or polypeptide components comprising an amino acid sequence having 40% or greater sequence identity with a third portion of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components collectively comprise 40% or greater sequence identity with the entirety of SEQ ID NO: 788 or 789, and are configured to produce a bioluminescent complex upon interaction of the first protein or peptide entity and the second protein or peptide entity;

(d) adding a coelenterazine or a coelenterazine analog substrate to the cell; and

(e) detecting a luminescent signal produced by the bioluminescent complex, wherein the magnitude of the luminescent signal correlates to the strength of the interaction between the first protein or peptide entity and the second protein or peptide entity.

164. A method of detecting co-localization of a first molecular entity and a second molecular entity, the method comprising:

(a) tagging the first molecular entity with a first peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater sequence identity with a first portion of SEQ ID NO: 788 or 789;

(b) tagging the second molecular entity with a second peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater sequence identity with a second portion of SEQ ID NO: 788 or 789;

(c) combining the tagged first molecular entity and the tagged second molecular entity in the same system;

(d) adding peptide, dipeptide, tripeptide, and/or polypeptide components to the system, the components having 40% or greater sequence identity with a third portion of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components collectively comprise 40% or greater sequence identity with the entirety of SEQ ID NO: 788 or 789, wherein the first peptide tag, the second peptide tag, and components are configured to produce a bioluminescent complex upon co-localization of the first molecular entity and the second molecular entity;

(e) adding a coelenterazine or a coelenterazine analog substrate to the system; and

(f) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates co-localization of the first molecular entity and the second molecular entity within the system, and/or wherein the magnitude of the luminescent signal correlates to the amount of co-localization within the system of the first molecular entity and the second molecular entity.

165. The method of claim 164, wherein the system comprises a cell, tissue, organ, whole organism, a biochemical, non-cellular sample.

166. The method of claim 164, wherein the first molecular entity and/or the second molecular entity is a protein of interest or a peptide of interest, and tagging comprises generating a fusion of the first molecular entity and/or the second molecular entity with the first tag and/or peptide tag.

167. The method of claim 164, wherein the first molecular entity and/or the second molecular entity is a small molecule and tagging comprises directly or indirectly linking the first molecular entity and/or the second molecular entity with the first tag and/or second tag.

168. The method of claim 164, wherein combining the tagged first molecular entity and the tagged second molecular entity comprises expressing one or both within the system and/or adding one or both to the system.

169. A method of detecting co-localization of a first protein or peptide entity and a second protein or peptide entity with a cell comprising, the method comprising:

(a) expressing within the cell a fusion comprising the first protein or peptide entity and a first peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater sequence identity with a first portion of SEQ ID NO: 788 or 789;

(b) expressing within the cell a fusion comprising the second protein or peptide entity and a second peptide, dipeptide, or tripeptide tag comprising an amino acid sequence having 40% or greater sequence identity with a second portion of SEQ ID NO: 788 or 789;

(c) expressing with the cell one or more peptide, dipeptide, tripeptide, or polypeptide components having 40% or greater sequence identity with a third portion of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components collectively comprise 40% or greater sequence identity with the entirety of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components are configured to produce a bioluminescent complex upon co-localization of the first protein or peptide entity and the second protein or peptide entity;

(d) adding a coelenterazine or a coelenterazine analog substrate to the cell; and

(e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates co-localization of the first protein or peptide entity and the second protein or peptide entity within the cell, and/or wherein the magnitude of the luminescent signal correlates to the amount of co-localization within the system of the first protein or peptide entity and the second protein or peptide entity.

170. A method of detecting a target molecule, wherein the target molecule displays a first antigen, epitope, or sequence and a distinct second antigen, epitope, or sequence, the method comprising:

(a) contacting a sample containing the target molecule with (i) a first primary binding moiety that recognizes the first antigen, epitope, or sequence and (ii) a second primary binding moiety that recognizes the second antigen, epitope, or sequence, and allowing the first and second primary binding moieties to bind to the first and second antigens, epitopes, or sequences;

(b) contacting the sample with (i) a first secondary binding moiety conjugated to a first tag and (ii) a second secondary binding moiety conjugated to second tag, wherein the first secondary binding moiety recognizes the first primary binding moiety and the second secondary binding moiety recognizes the second primary binding moiety, wherein the first or second tags comprises amino acid sequences having 40% or greater sequence identity with first and second portions of SEQ ID NO: 788 or 789;

(c) allowing the first and second secondary binding moieties to bind to the first and second primary binding moieties;

(d) contacting the sample with comprising one or more peptide, dipeptide, tripeptide, and/or polypeptide components having 40% or greater sequence identity with a third portion of SEQ ID NO: 788 or 789; wherein the first tag, the second tag, and the components collectively comprise 40% or greater sequence identity with the entirety of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components are configured to produce a bioluminescent complex upon interaction;

(d) contacting the sample with a coelenterazine or a coelenterazine analog substrate; and

(e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates the presence of the target molecule, and/or wherein the magnitude of the luminescent signal correlates to the amount of target molecule within the sample.

171. The method of claim 170, wherein the binding moieties are independently selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

172. The method of claim 170, wherein the target molecule is a protein, nucleic acid, or small molecule.

173. The method of claim 170, wherein the sample is *in vitro* or *in vivo*.

174. A method of detecting a target molecule, wherein the target molecule displays a first antigen, epitope, or sequence and a distinct second antigen, epitope, or sequence, the method comprising:

(a) contacting the sample with (i) a first binding moiety conjugated to a first tag and (ii) a second binding moiety conjugated to second tag, wherein the first secondary binding moiety recognizes the first antigen, epitope, or sequence and the second binding moiety recognizes the second antigen, epitope, or sequence, wherein the first tag comprises an amino acid sequence having 40% or greater sequence identity with a first portion of SEQ ID NO: 788 or 789, and wherein the second tag comprises an amino acid sequence with a first portion of SEQ ID NO: 788 or 789;

(b) allowing the first and second binding moieties to bind to the first and second antigens, epitope, or sequences;

(c) contacting the sample with a peptide, dipeptide, tripeptide, or polypeptide component having 40% or greater sequence identity with a third portion of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components collectively comprise 40% or greater sequence identity with the entirety of SEQ ID NO: 788 or 789, wherein the first tag, the second tag, and the components are configured to produce a bioluminescent complex upon interaction;

(d) contacting the sample with a coelenterazine or a coelenterazine analog substrate; and

(e) detecting a luminescent signal produced by the bioluminescent complex, wherein the presence of luminescent signal above background indicates the presence of the target molecule, and/or wherein the magnitude of the luminescent signal correlates to the amount of target molecule within the sample.

175. The method of claim 174, wherein the binding moieties are independently selected from the group consisting of an antibody (polyclonal, monoclonal, and/or recombinant), antibody fragment, protein A, an Ig binding domain of protein A, protein G, an Ig binding domain of protein G, protein A/G, an Ig binding domain of protein A/G, protein L, an Ig binding domain of protein L, protein M, an Ig binding domain of protein M, oligonucleotide

probe, peptide nucleic acid, DARPin, aptamer, affimer, a purified protein (either the analyte itself or a protein that binds to the analyte), and analyte binding domain(s) of proteins.

176. The method of claim 174, wherein the target molecule is a protein, nucleic acid, or small molecule.

177. The method of claim 174, wherein the sample is *in vitro*, *in vivo,* or a biochemical sample.

178. A composition comrpising a polypeptide comprising an amino acid sequence with greater than 40% but less than 100% sequence identity with SEQ ID NO: 788 or 789 and less than 100% sequence identity with SEQ ID NO: 1 or 3; wherein the polypeptide exhibits luminescence in the presence of a coelenterazine or a coelenterazine analog substrate.

179. The composition of claim 178, wherein the peptide comprising an amino acid sequence with greater than 40% sequence identity with one or more of SEQ ID NOS: 780, 782, 784, 786, 802, 804, 806, 808, 813, 815, or 829.

180. A nucleic acid comprising a sequence coding for a poly peptide of claim 178.

181 A fusion polypeptide comprising the polypeptide of claim 178 and an additional amino acid sequence.

182 A nucleic acid comprising a sequence coding for the fusion polypeptide of claim 181.

183. A method comprising (a) contacting the composiution of claim 178 with a coelenterazine or a coelenterazine analog substrate; and (b) detecting luminescence.

184. A polypeptide comprisingcomrpising an amino acid sequence with greater than 40% sequence identity with SEQ ID NO: 15 linked to the N-terminal end of an amino acid sequence with greater than 40% sequence identity with one of SEQ ID NOS: 17, 21, or 302;

wherein a bioluminescent signal produced in the presence of a coelenterazine or a coelenterazine analog substrate is substantially increased when the polypeptide contacts a peptide or polytpeptide comprising an amino acid sequce consisting of SEQ ID NO: 23 when compared to a bioluminescent signal produced by the polypeptide and the coelenterazine substrate alone.

185. A nucleic acid comprising a sequence coding for a poly peptide of claim 184.

186. A fusion polypeptide comprising the polypeptide of claim 184 and an additional amino acid sequence.

187. A nucleic acid comprising a sequence coding for the fusion polypeptide of claim 181.

188. A method comprising (a) contacting the polypeptide of claim 184 with a coelenterazine or a coelenterazine analog substrate and a peptide or polypeptide comprising an amino acid sequence with 40% or greater sequecne identity with SEQ ID NO: 23; and (b) detecting luminescence.

189. A system comprising:

(a) a sensor polypeptide comprising a first amino acid sequence having at least 40% sequence identity with SEQ ID NO: 11 linked to an amino acid sequence that localizes in a specific cellular location; and

(b) a protein of interest linked to a peptide tag comprising an amino acid sequence having at least 40% sequence identity to SEQ ID NO: 23;

wherein a bioluminescent complex is formed between the polypeptide and the peptide tag upon colocalization within the specific cellular location.

190. The system of claim 189, wherein the specific cellular location is slected from plasma membrane, nucleus, mitochondria, and the endoplasmic reticulum.

191. A method comprising (a) expressing the system of claim 189 in a cell; (b) contacting the cell with a enterazine or a coelenterazine analog substrate; and (c) detecting luminescnce, wherein an increase in luminescene indicates formation of the bioluminescent complex and co-localitions of the polypeptide and peptide tag.

192. The method of claim 191, further comprising a step of inducing translocation protein of interest linked to a peptide tag to the specific cellular location.

FIG. 1

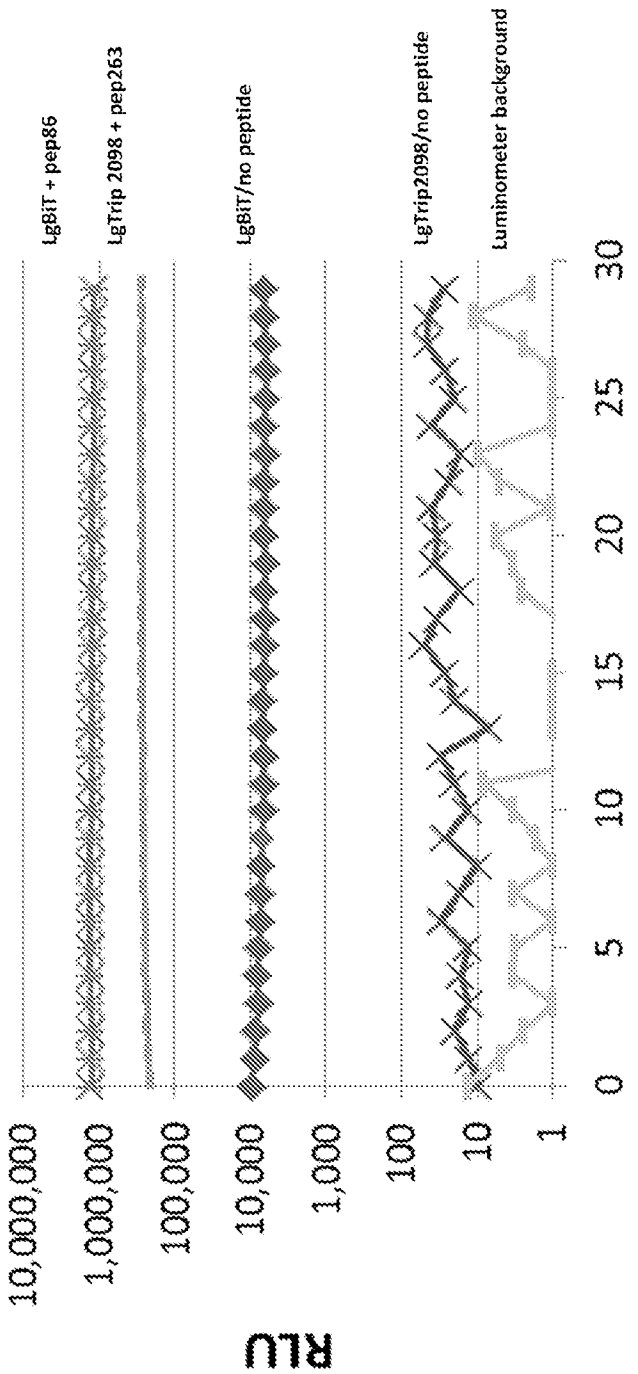

LgBiT + pep86
LgTrip 2098 + pep263
LgBiT/no peptide
LgTrip2098/no peptide
Luminometer background
RLU
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
0
5
10
15
20
25
30
Time (min)

FIG. 2

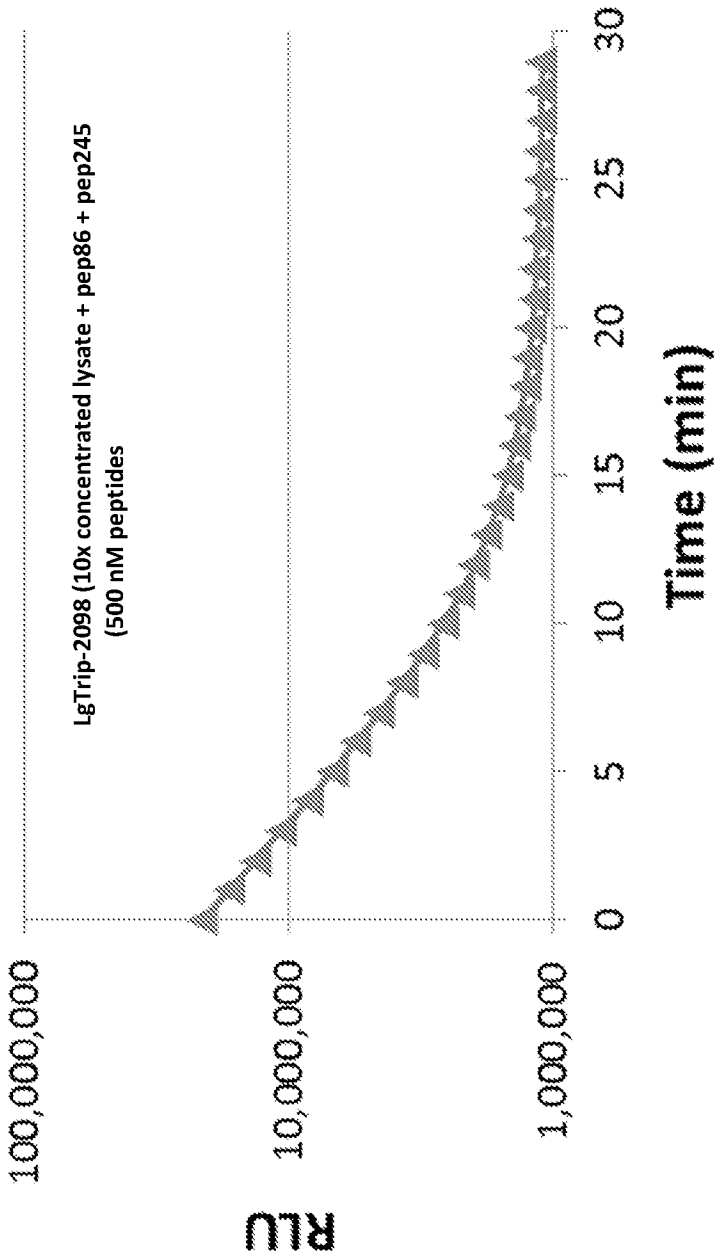

LgTrip-2098 (10x concentrated lysate + pep86 + pep245
(500 nM peptides
RLU
100,000,000
10,000,000
1,000,000
0
5
10
15
20
25
30
Time (min)

FIG. 3

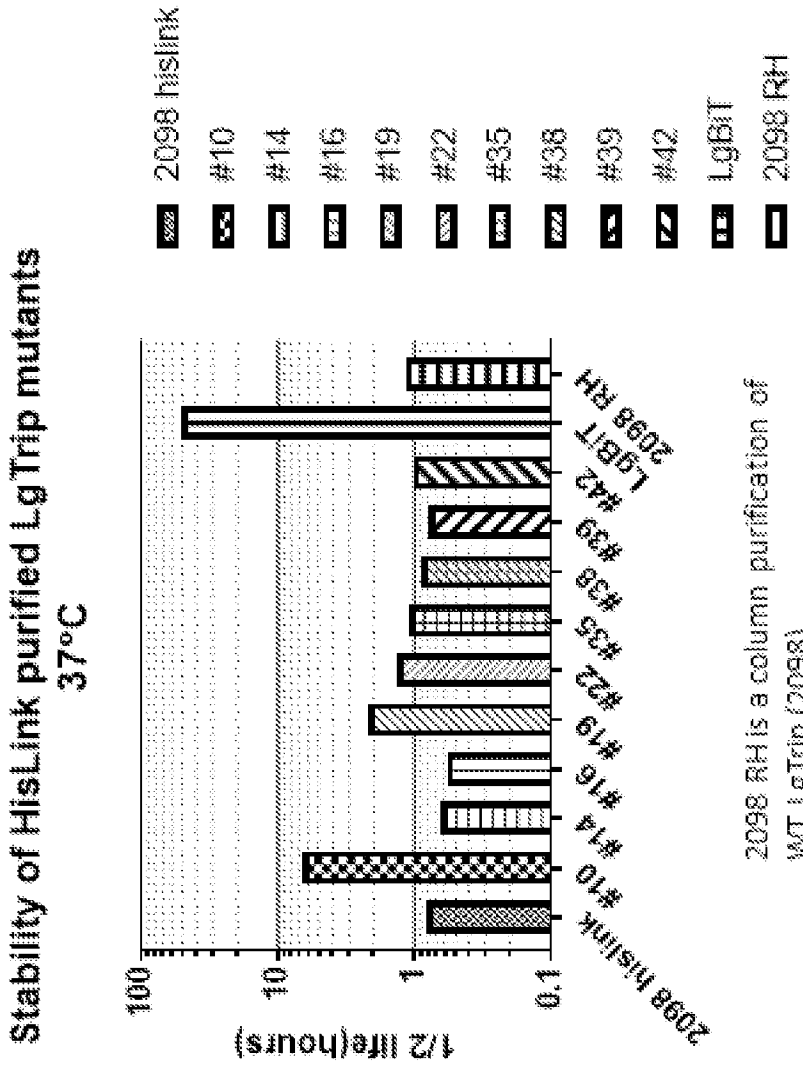

Stability of HisLink purified LgTrip mutants
37°C
1/2 life(hours)
100
10
1
0.1
2098 hislink
#10
#14
#16
#19
#22
#35
#38
#39
#42
LgBiT
2098 RH
2098 hislink
#10
#14
#16
#19
#22
#35
#38
#39
#42
LgBiT
2098 RH
2098 RH is a column purification of
WT LgTrip (2098).

FIG. 4

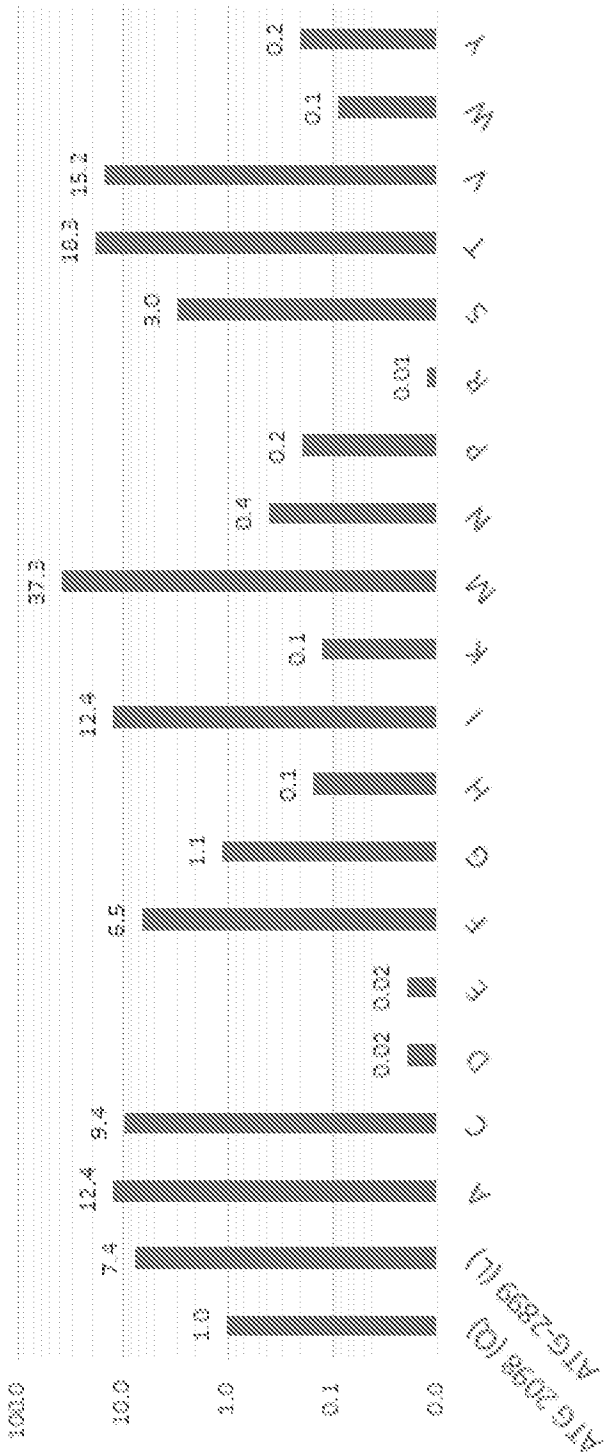

Position 42 site saturation LgTrip
6 hour induction (37C)
ATG-7098 (Q)
ATG-2889 (L)
100.0
10.0
1.0
0.1
0.0
A V C D E F G H I K L M N P Q R S T W Y
1.0 7.4 12.4 9.4 0.02 0.02 6.5 1.1 0.1 12.4 0.1 37.3 0.4 0.2 0.01 3.0 10.2 15.2 0.1 0.2

FIG. 5

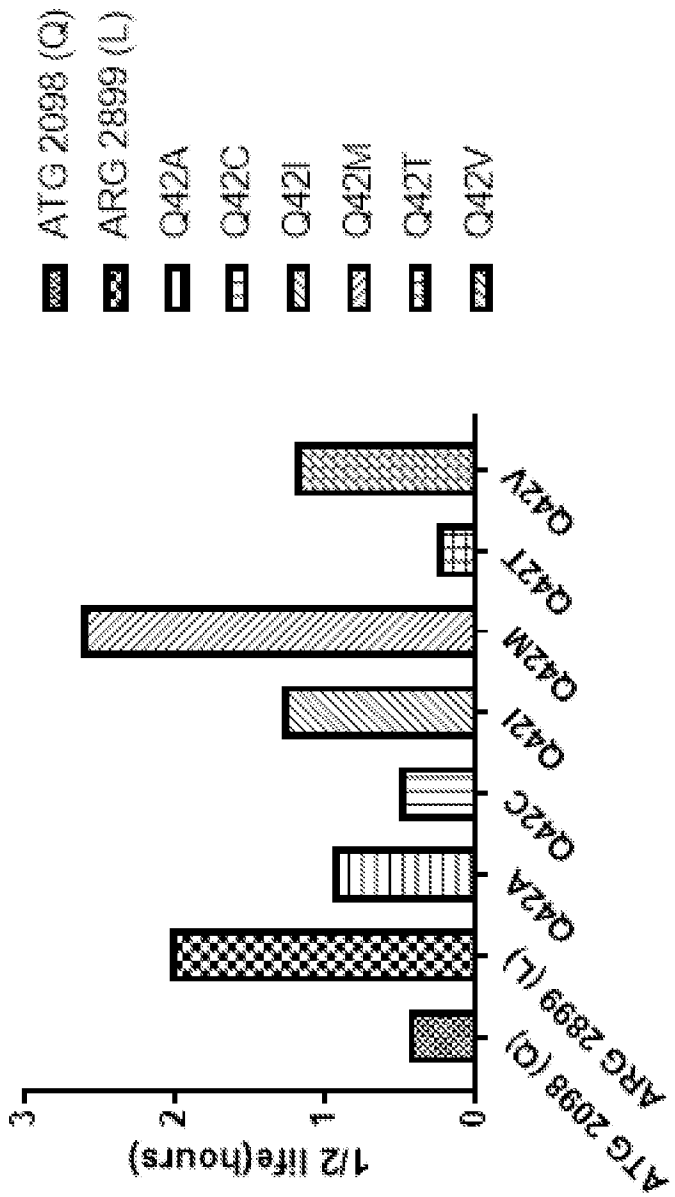

ATG 2098 (Q)
ARG 2899 (L)
Q42A
Q42C
Q42I
Q42M
Q42T
Q42V
37°C stability
1/2 life(hours)
0
1
2
3
ATG 2098 (Q)
ARG 2899 (L)
Q42A
Q42C
Q42I
Q42M
Q42T
Q42V

FIG. 6A

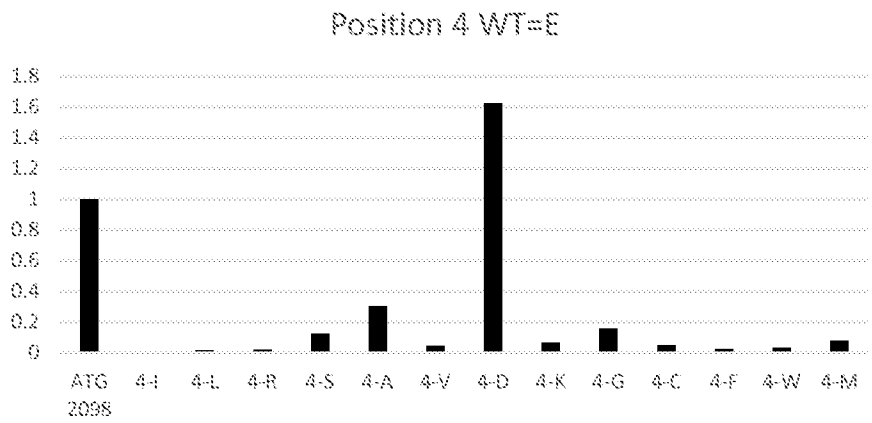


FIG. 6B

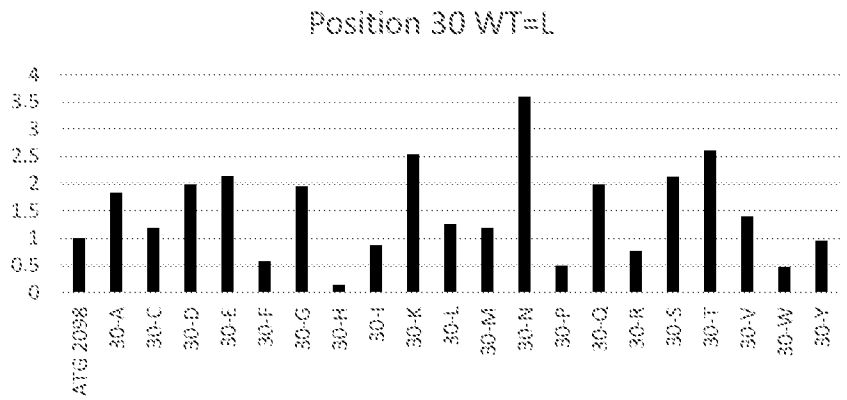


FIG. 6C

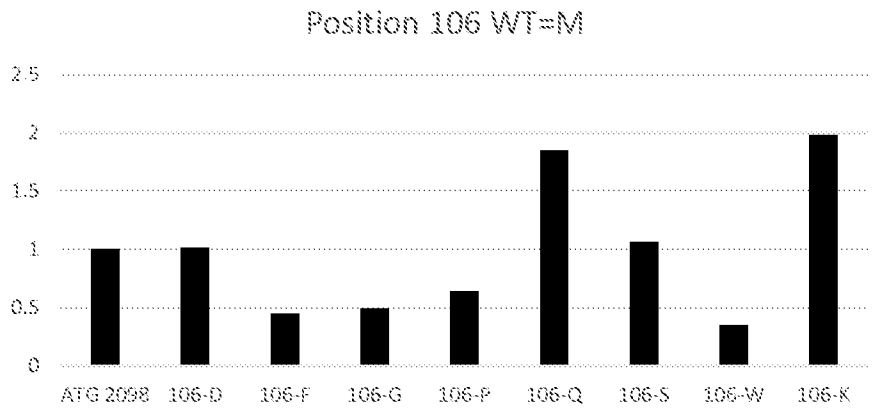


FIG. 7A

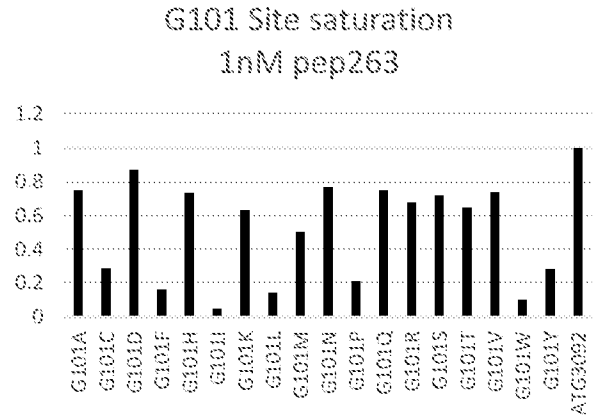


FIG. 7B

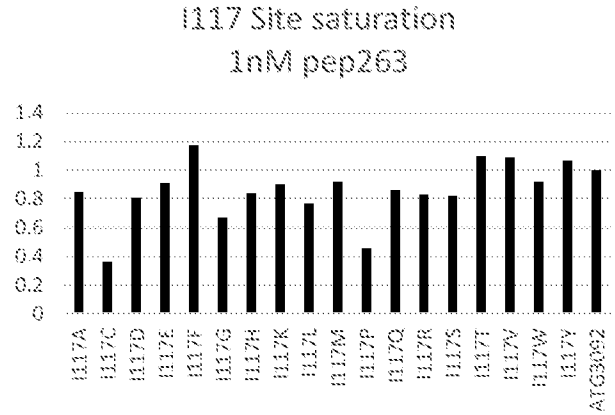


FIG. 7C

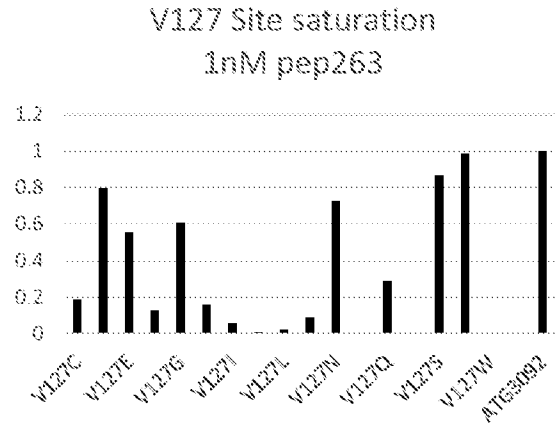


FIG. 7D

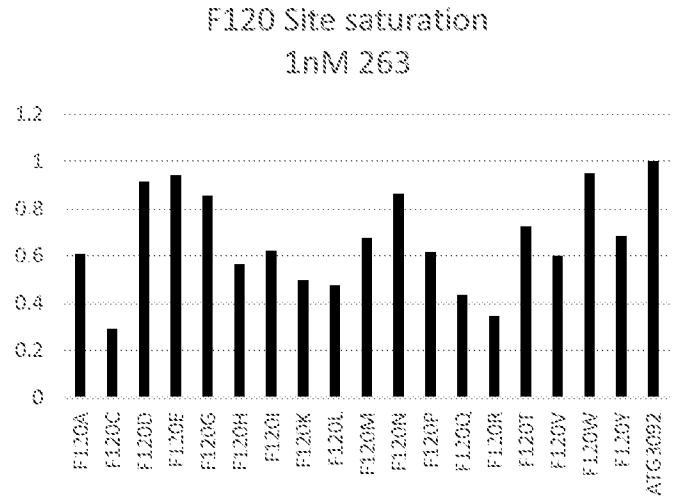

F120 Site saturation
1nM 263
1.2
1
0.8
0.6
0.4
0.2
0
F120A
F120C
F120D
F120E
F120G
F120H
F120I
F120K
F120L
F120M
F120N
F120P
F120Q
F120R
F120T
F120V
F120W
F120Y
ATG3092

FIG. 7E

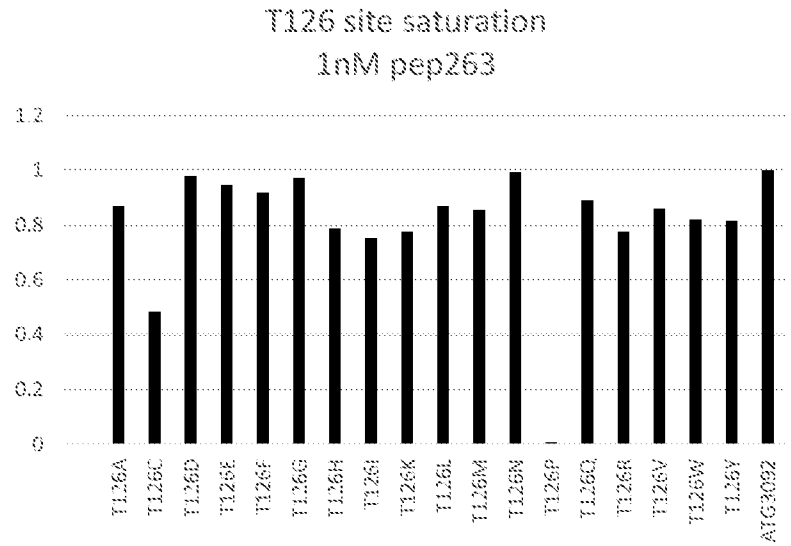

T126 site saturation
1nM pep263
1.2
1
0.8
0.6
0.4
0.2
0
T126A
T126C
T126D
T126E
T126F
T126G
T126H
T126I
T126K
T126L
T126M
T126N
T126P
T126Q
T126R
T126V
T126W
T126Y
ATG3092

FIG. 8

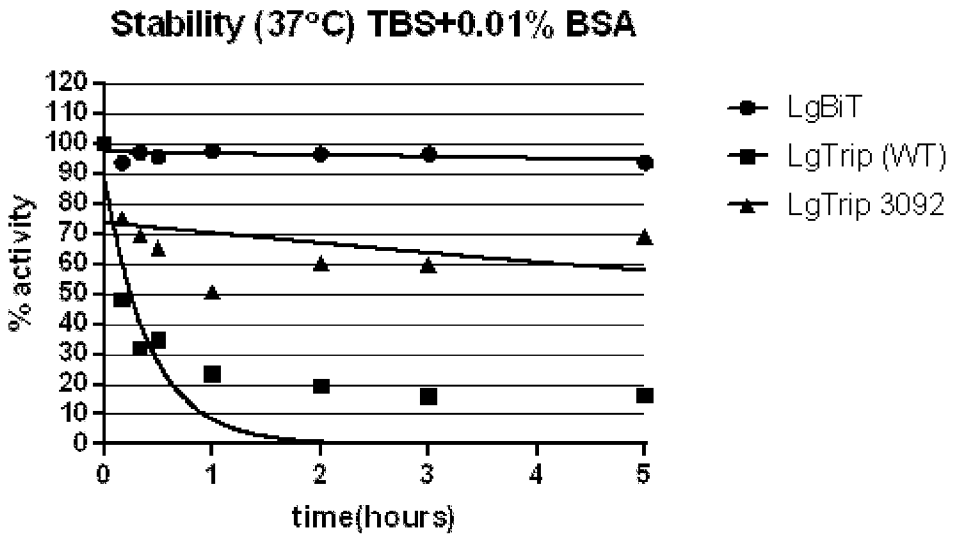
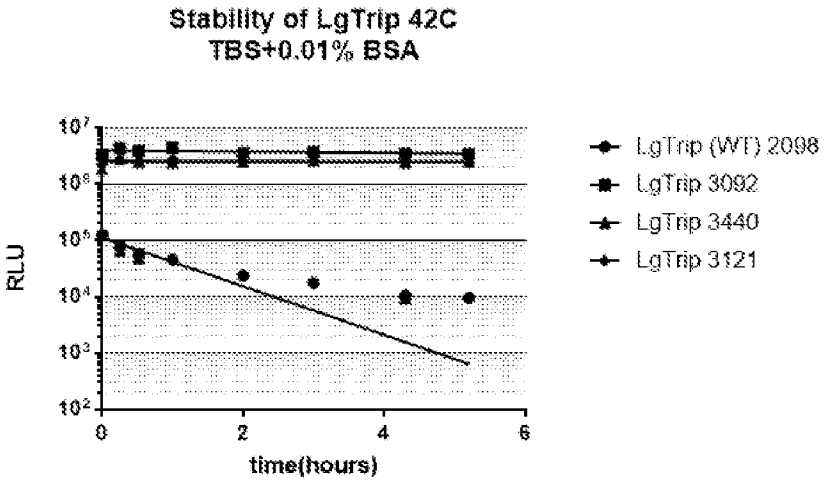

Stability (37°C) TBS+0.01% BSA
% activity
time(hours)
LgBiT
LgTrip (WT)
LgTrip 3092

FIG. 9A



ATG 3121=3092+N105S
ATG 3440=3092+V102D+E115D
ATG 3092=Q42M+E4E+M106K)
ATG 2098=W.T. LgTrip

FIG. 9B

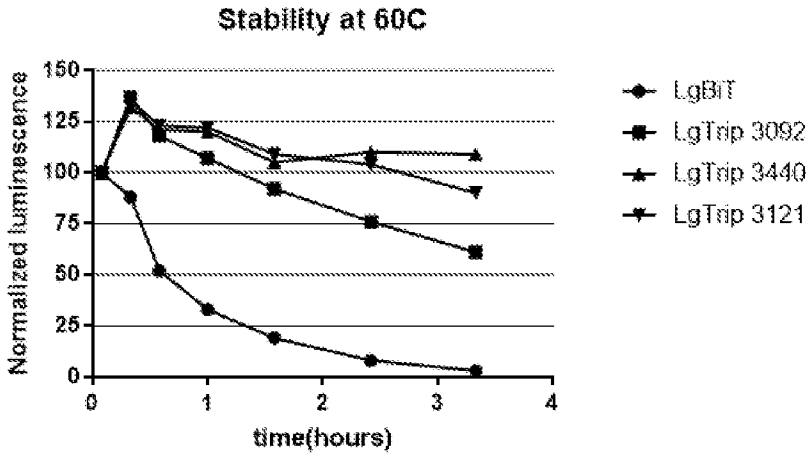
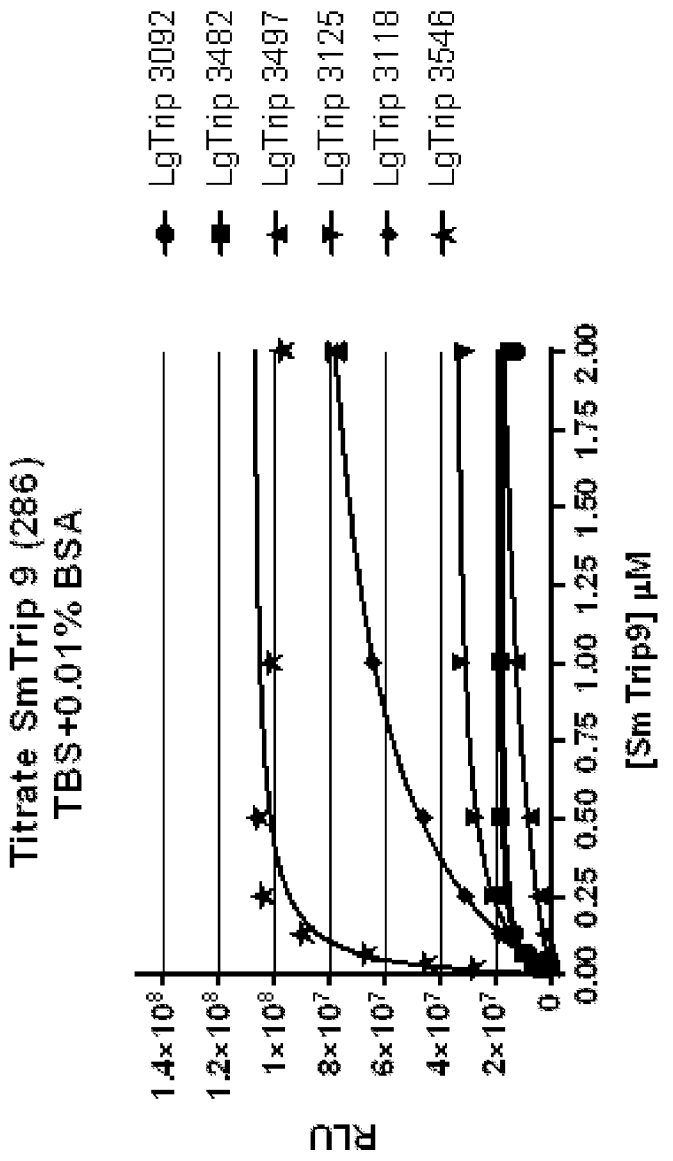
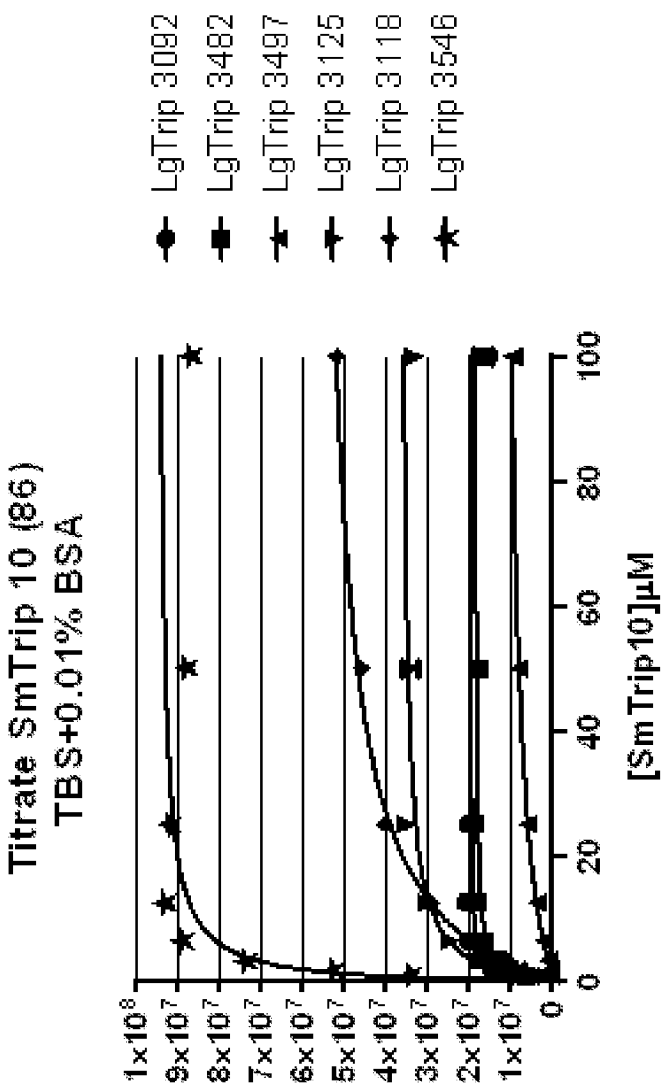


FIG. 10A



ATG	Mutation compared to 3092
ATG 3092	
ATG 3482	I3 I7F
ATG 3497	I3 I28D
ATG 3125	I3 I26S
ATG 3118	V3 I27A
ATG 3546	V3 I27T

FIG. 10B



ATG	Mutation compared to 3092
ATG 3092	
ATG 3482	I121F
ATG 3497	P128Q
ATG 3125	T135S
ATG 3118	V127A
ATG 3546	V127T

FIG. 10C

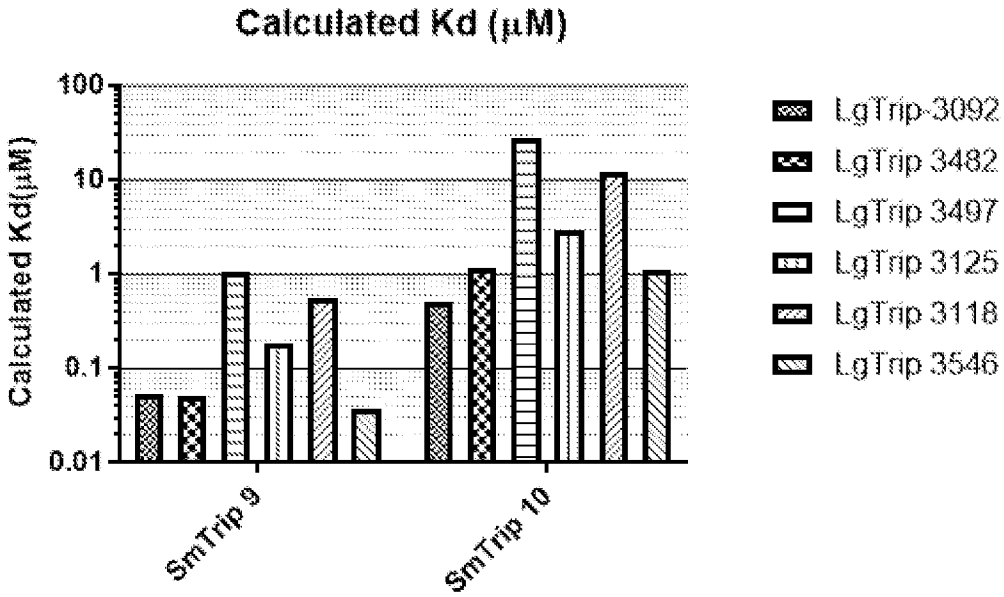
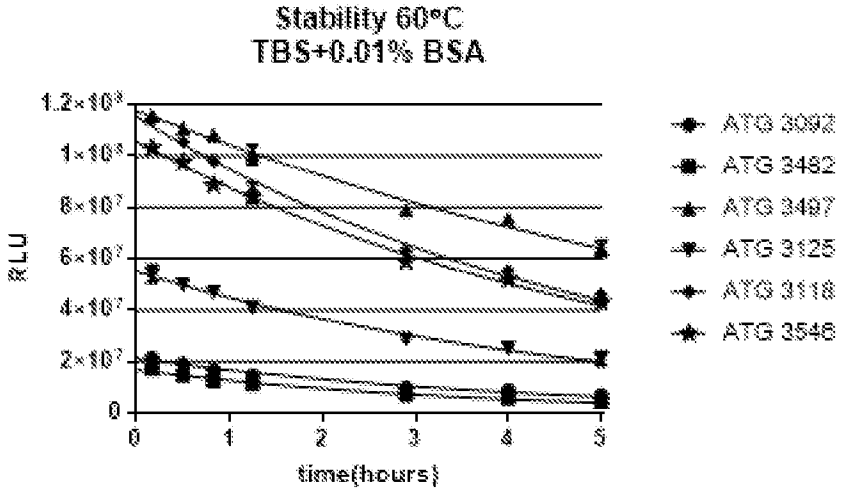

Calculated Kd (μM)
Calculated Kd(μM)
100
10
1
0.1
0.01
SmTrip 9
SmTrip 10
LgTrip-3092
LgTrip 3482
LgTrip 3497
LgTrip 3125
LgTrip 3118
LgTrip 3546

FIG. 11A



	ATG 3092	ATG 3482	ATG 3497	ATG 3125	ATG 3118	ATG 3548
Half Life	2.81	2.389	5.686	3.391	3.587	3.748

FIG. 11B

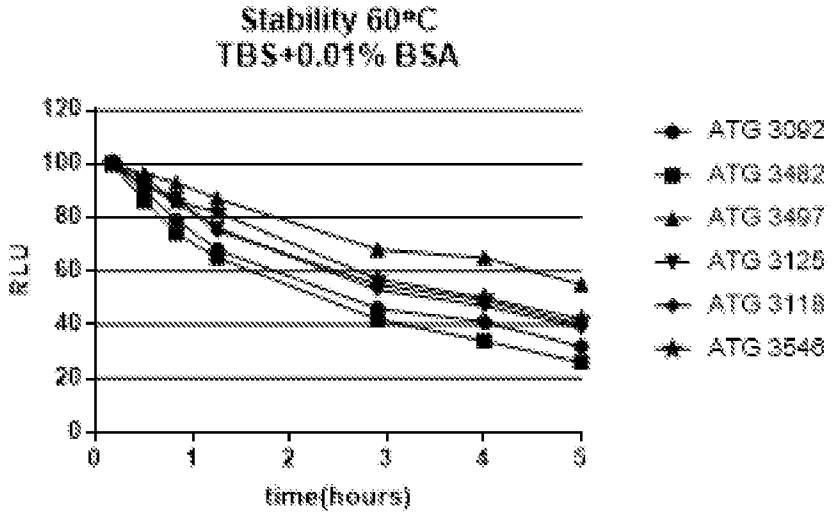


FIG. 12A

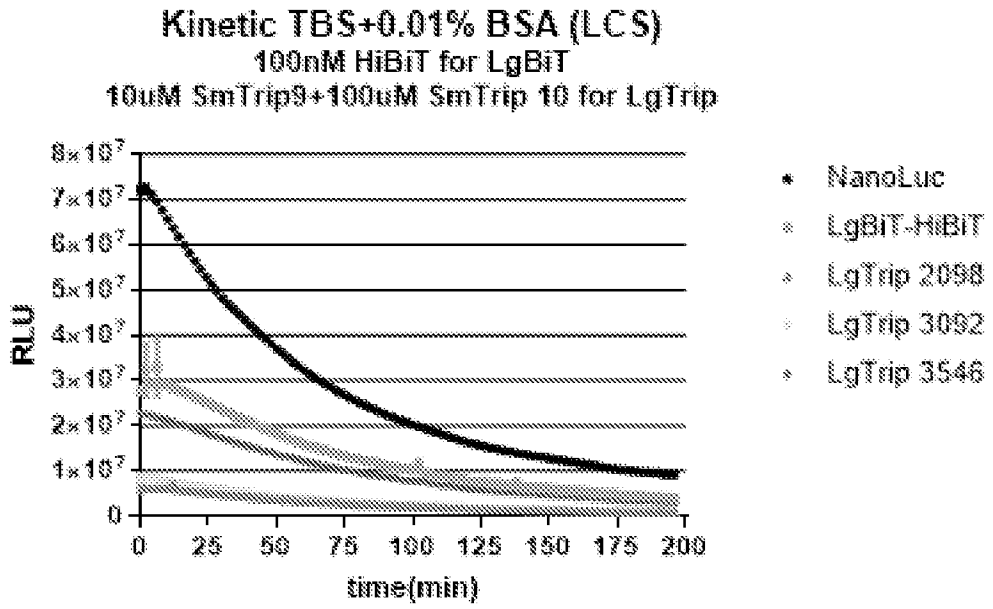


FIG. 12B

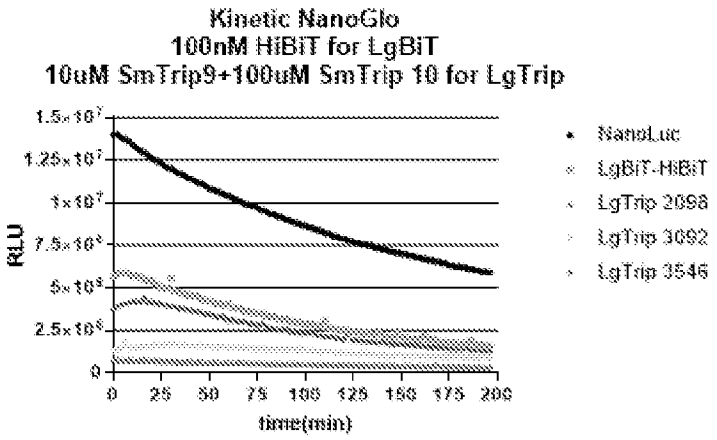


FIG. 13A

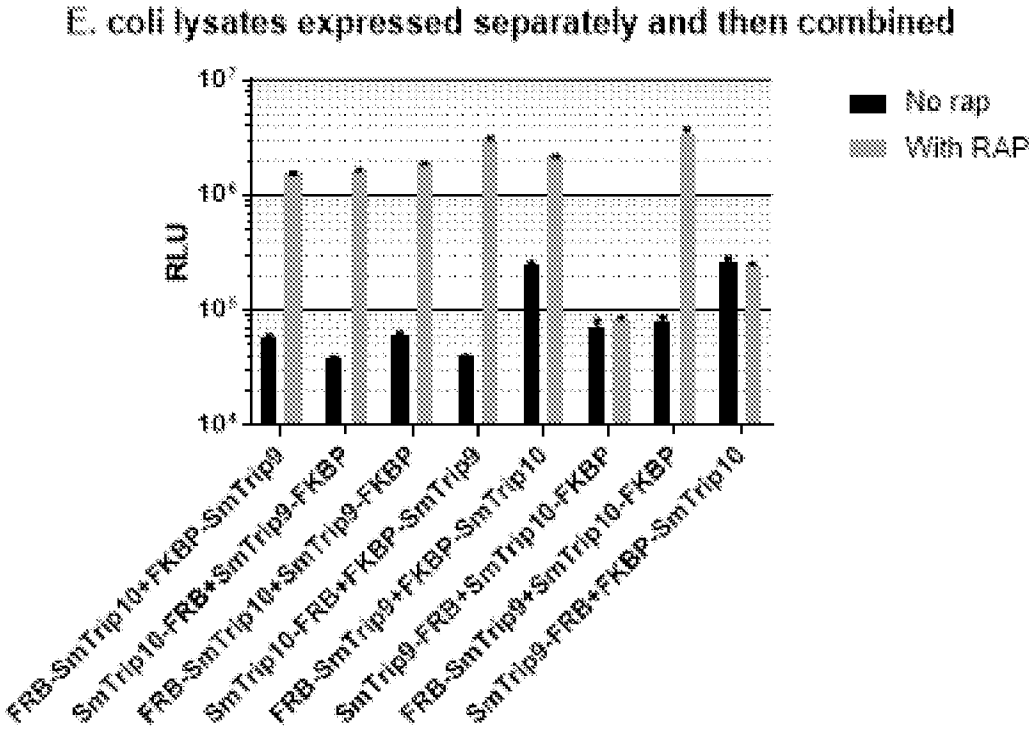


FIG. 13B

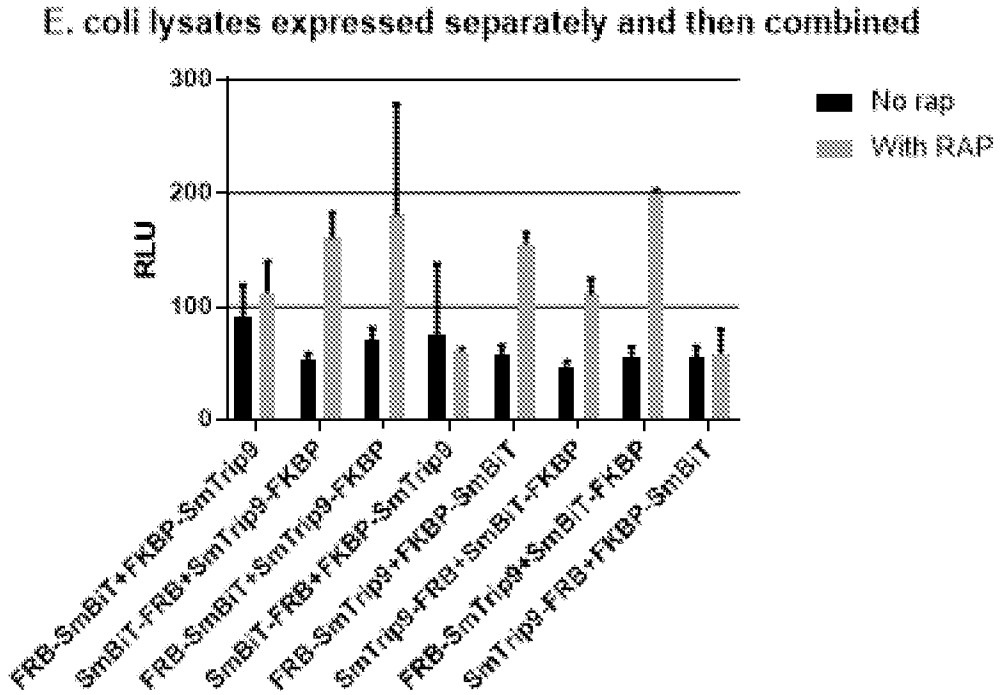


FIG. 14A

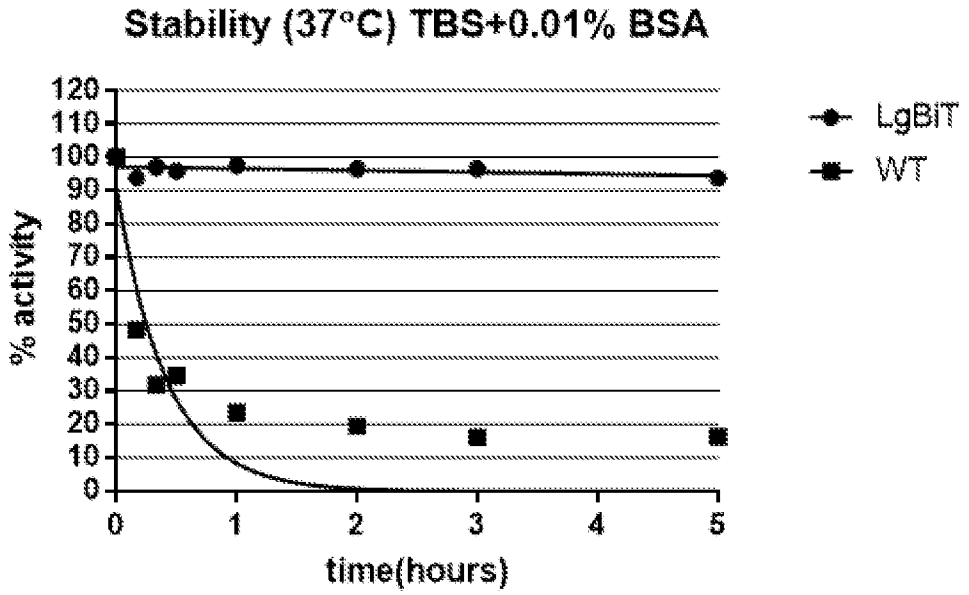


FIG. 14B

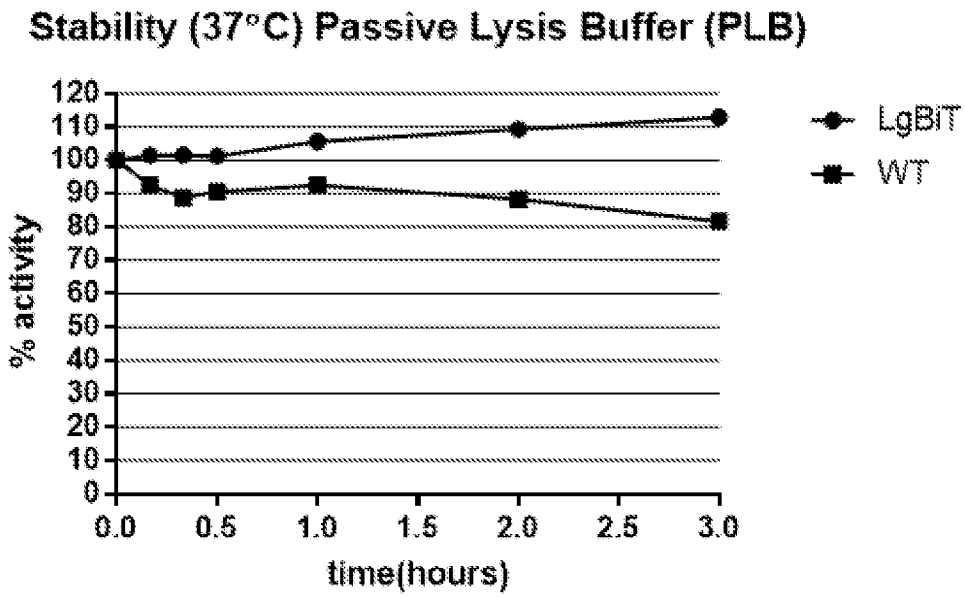


FIG. 15A

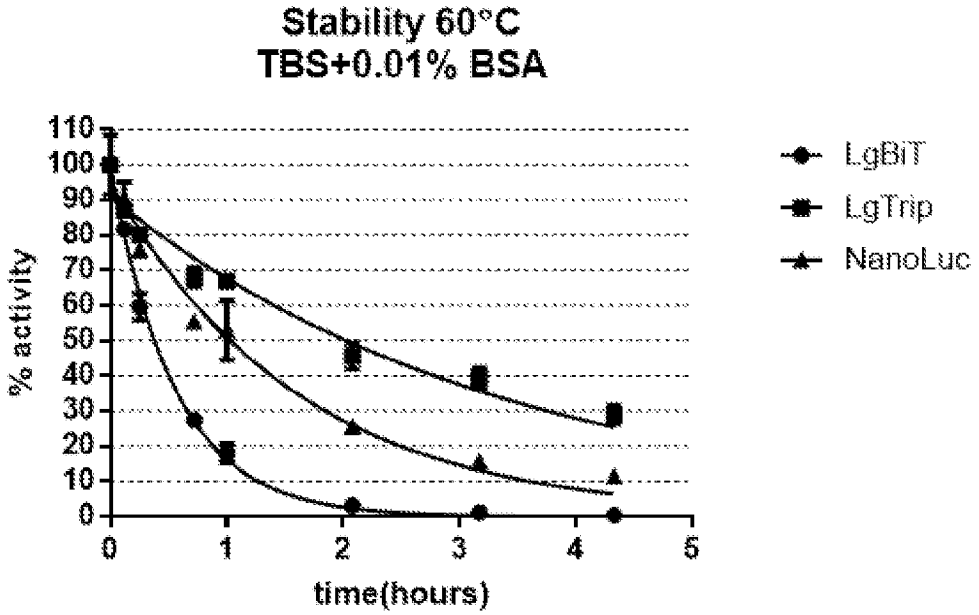


FIG. 15B

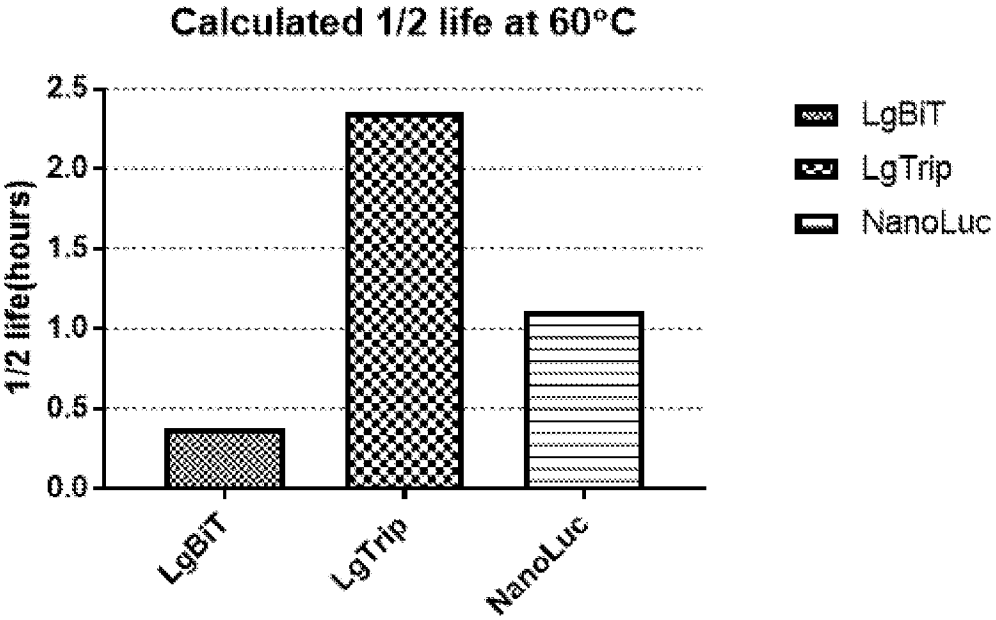
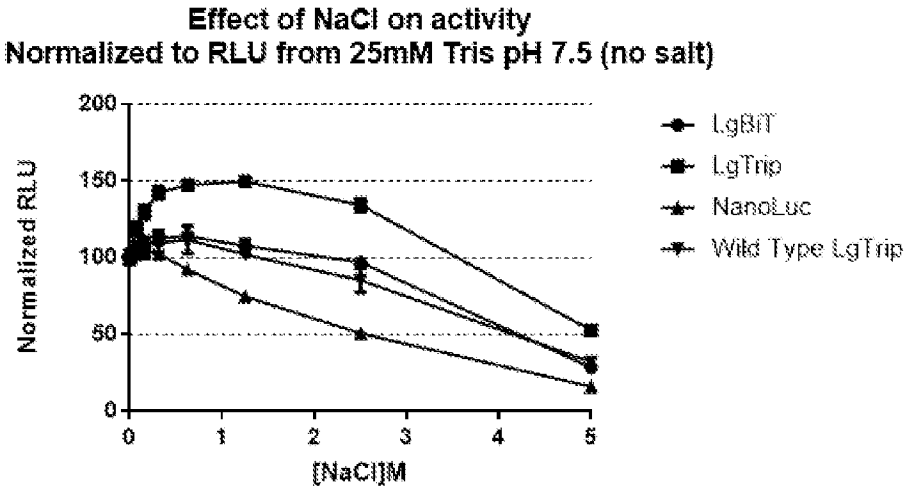
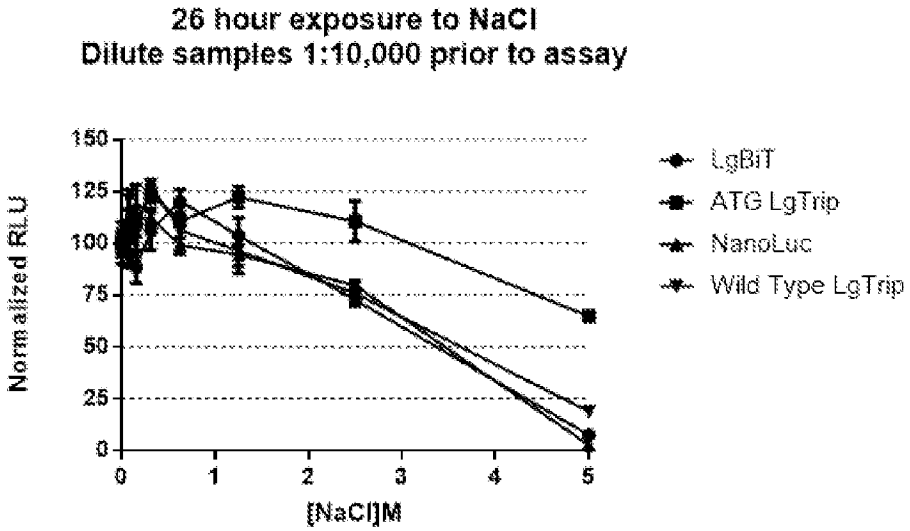


FIG. 16A



Activity=Salt in the furimazine assay buffer

FIG. 16B



Exposure=Incubate in presence of salt titration, then dilute and assay under standard assay conditions

FIG. 17A

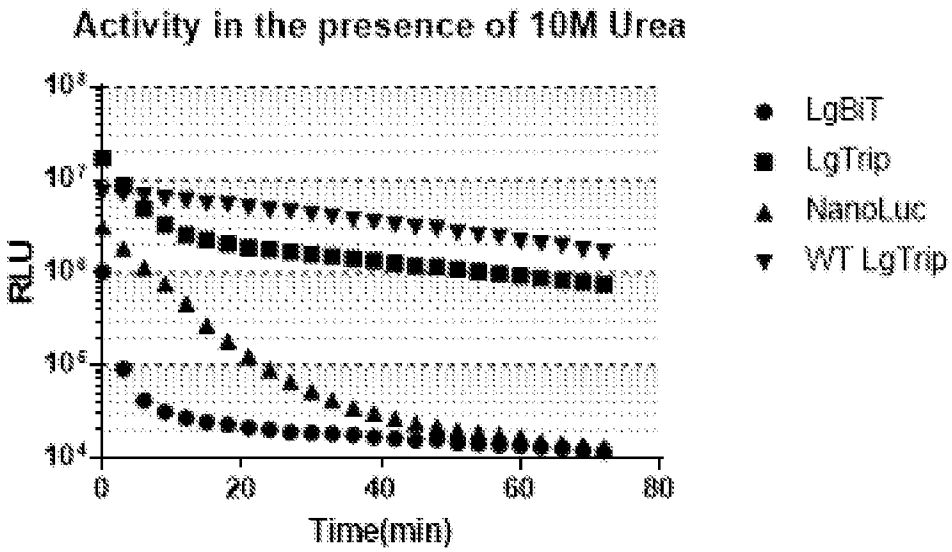


FIG. 17B

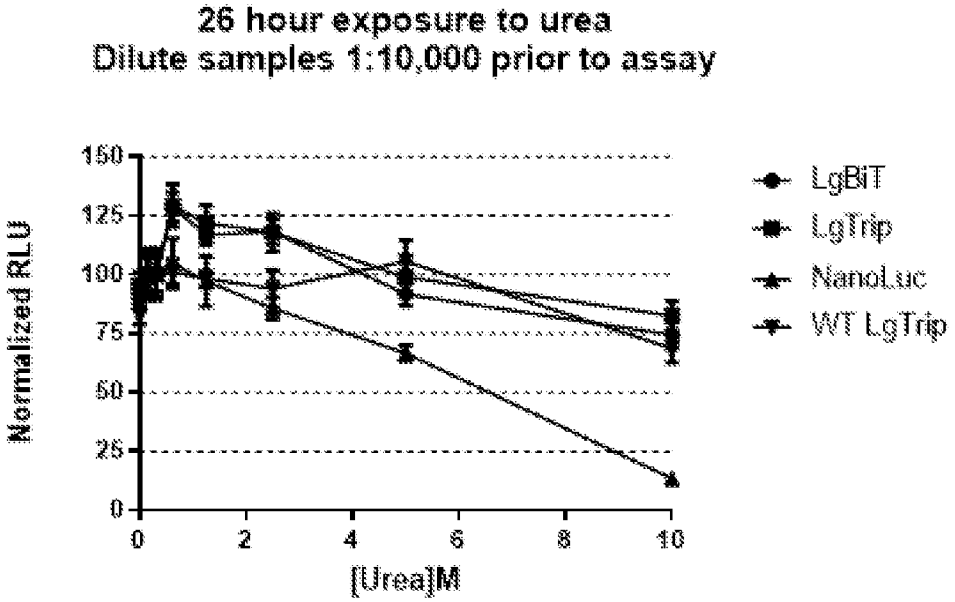


FIG. 18A

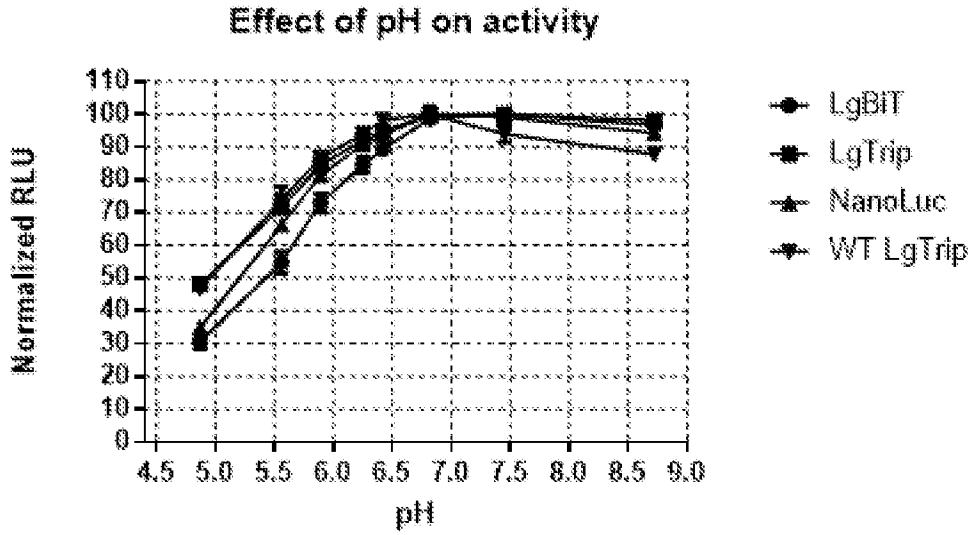


FIG. 18B

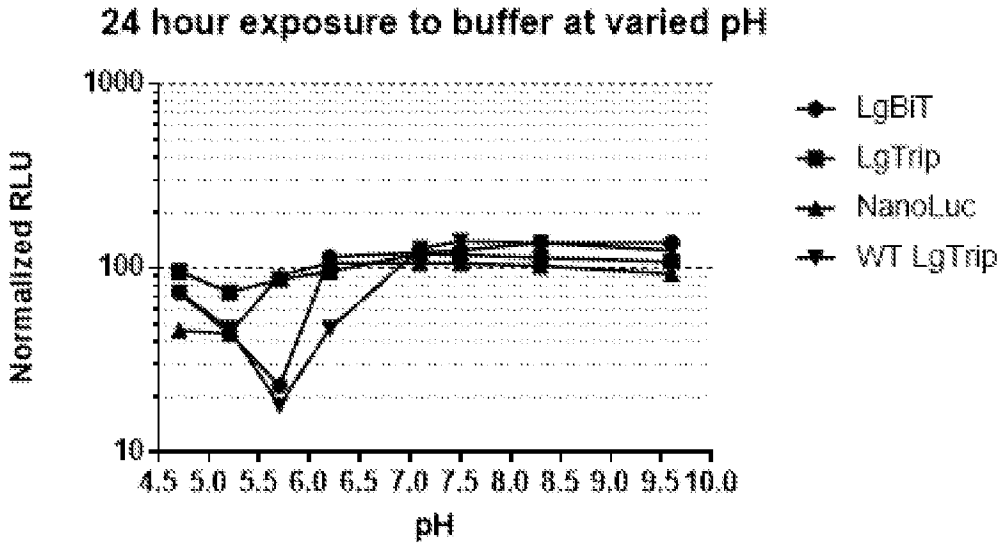
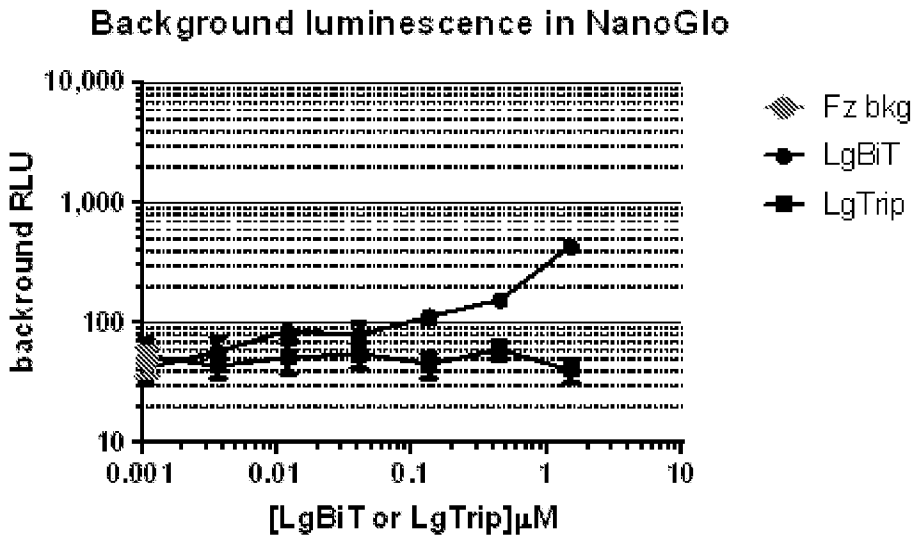


FIG. 19


Background luminescence in NanoGlo
10,000
1,000
100
10
backround RLU
0.001
0.01
0.1
1
10
[LgBiT or LgTrip]µM
Fz bkg
LgBiT
LgTrip

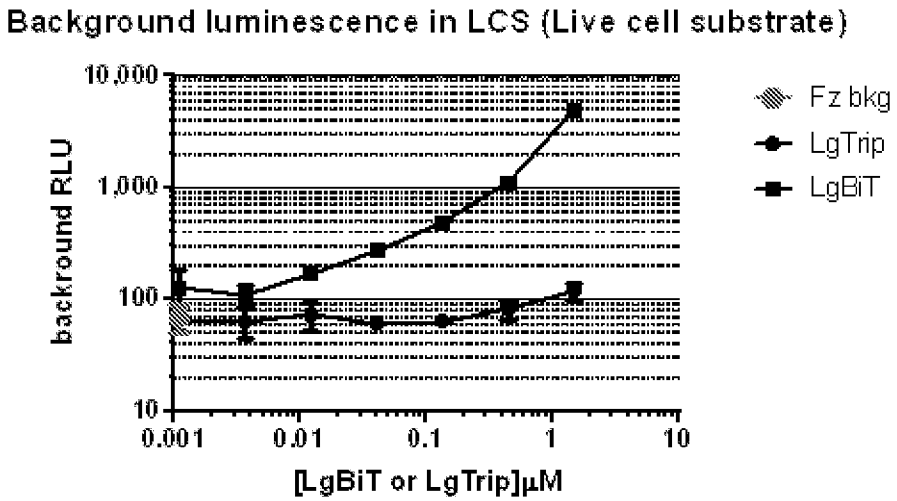

Background luminescence in LCS (Live cell substrate)
10,000
1,000
100
10
backround RLU
0.001
0.01
0.1
1
10
[LgBiT or LgTrip]µM
Fz bkg
LgTrip
LgBiT

FIG. 20A

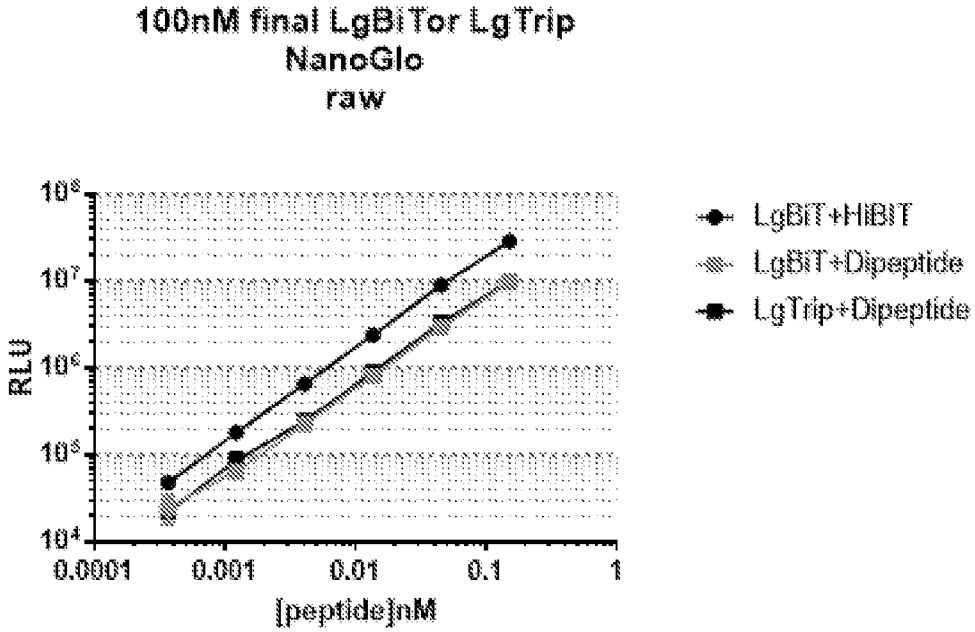


FIG. 20B

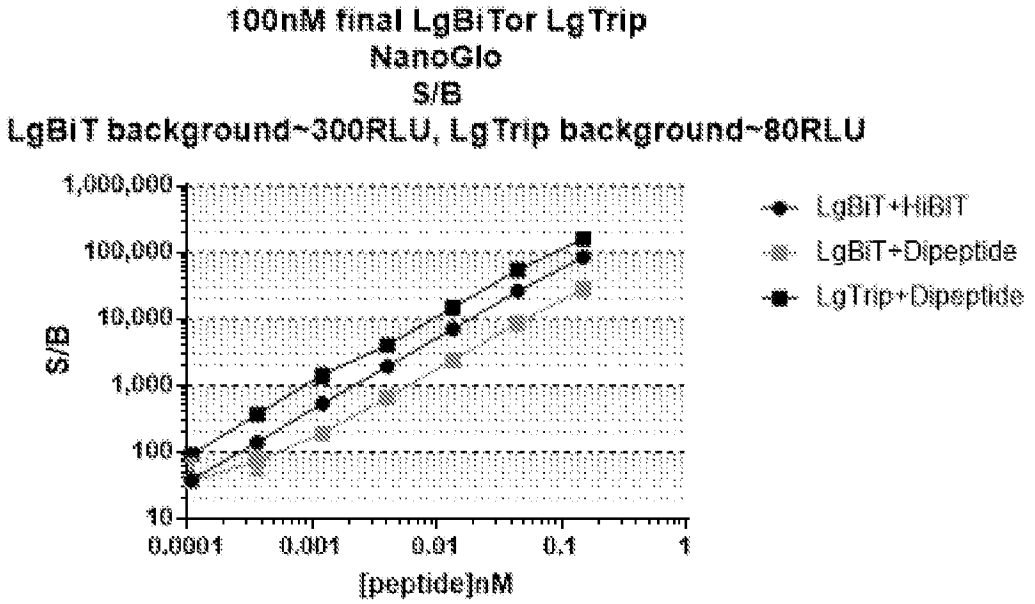


FIG. 21A

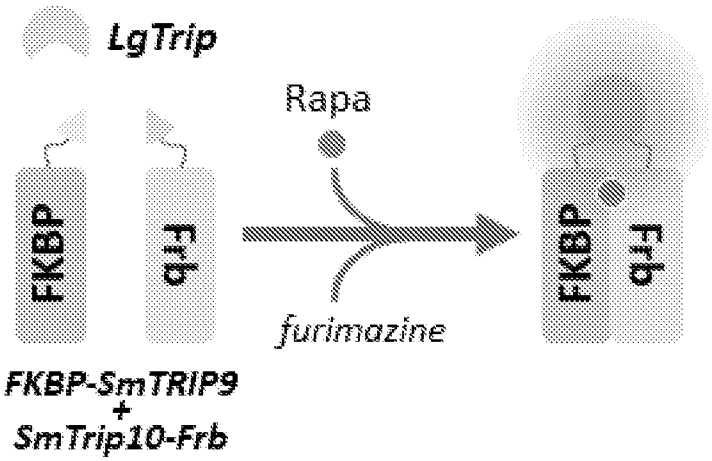
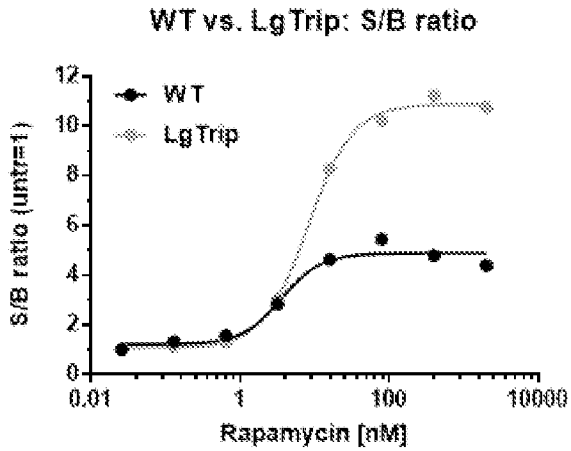
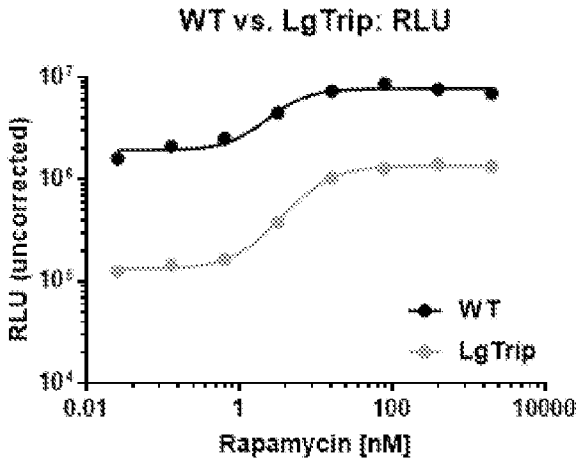


FIG. 21B



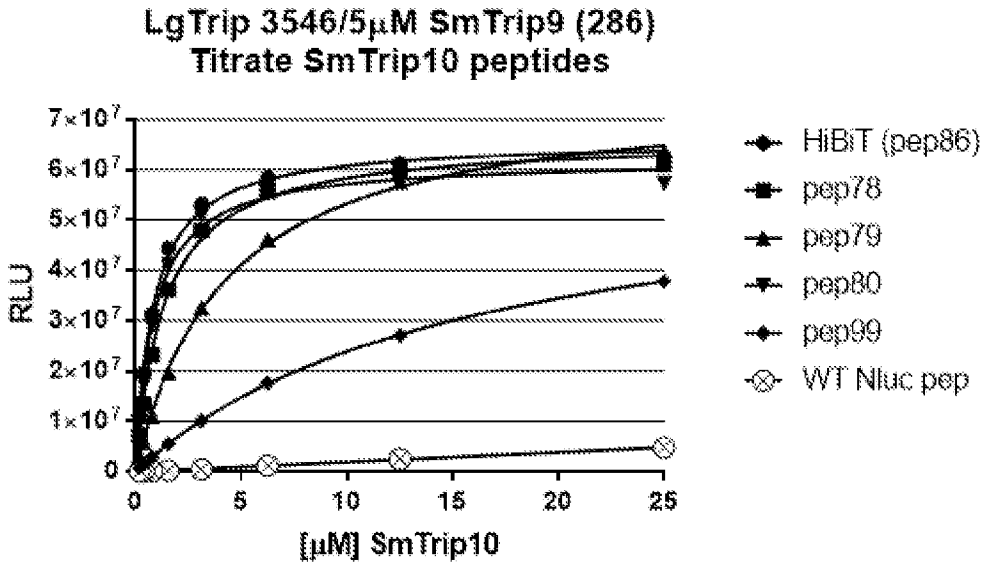
	WT	LgTrip
EC50	3.552	8.131

FIG. 21C



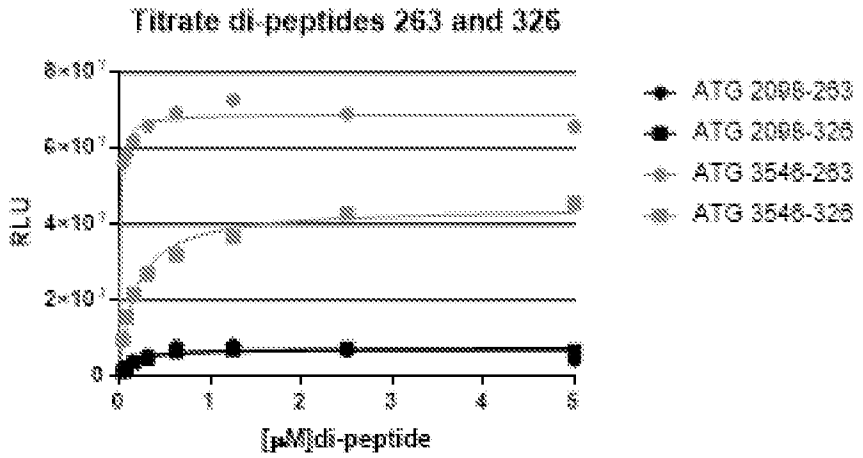
	WT	LgTrip
EC50	3.551	8.13

FIG. 22



	HiBiT (pep86)	pep78	pep79	pep80	pep99
Kd	0.8432	1.354	4.226	0.8704	15.94

FIG. 23



	ATG 2098-263	ATG 2098-326	ATG 3548-263	ATG 3548-326
Kd	0.1042	0.168	0.01081	0.1782

FIG. 24

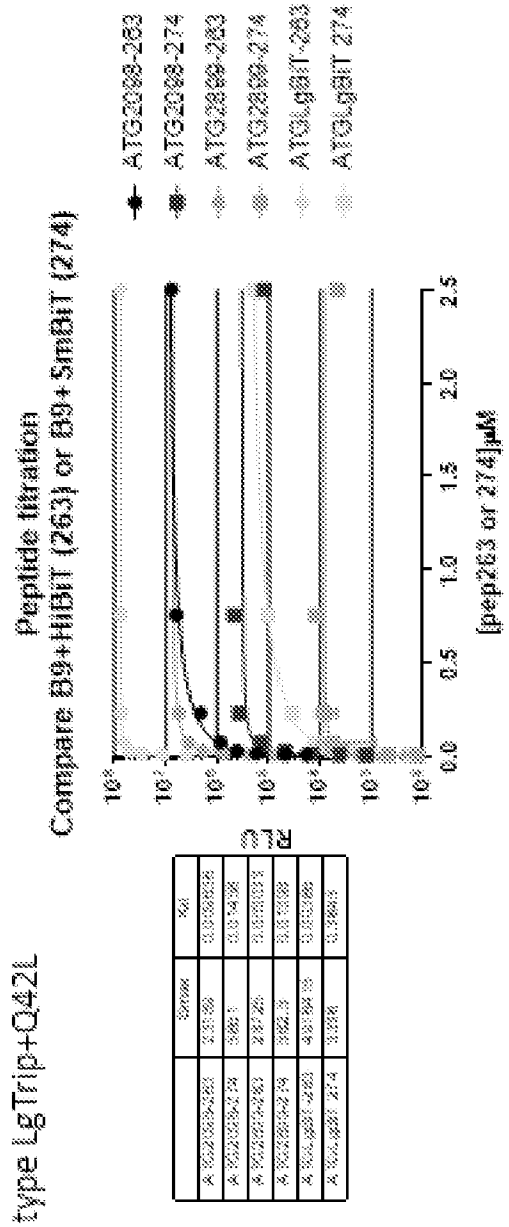

ATG 2899=Wild type LgTrip+Q42L
Peptide titration
Compare B9+HiBiT (263) or B9+SmBiT (274)
ATG2268-263
ATG2268-274
ATG2899-263
ATG2899-274
ATGLgBT-263
ATGLgBT 274
RLU
[pep263 or 274]µM
0.0
0.5
1.0
1.5
2.0
2.5

FIG. 25A

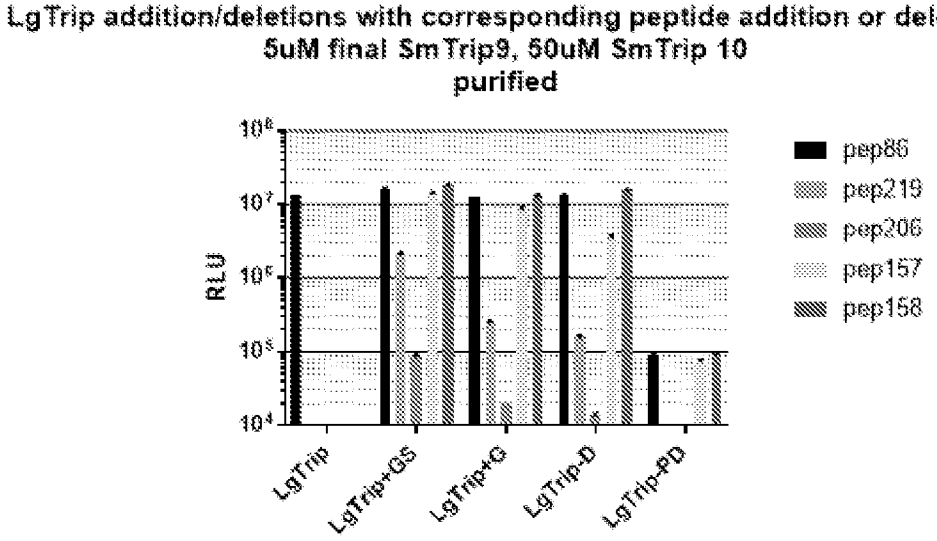


FIG. 25B

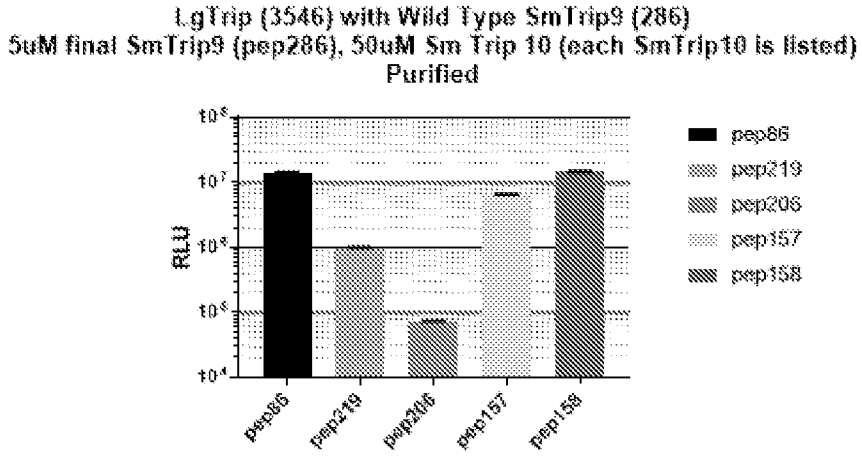


FIG. 26A

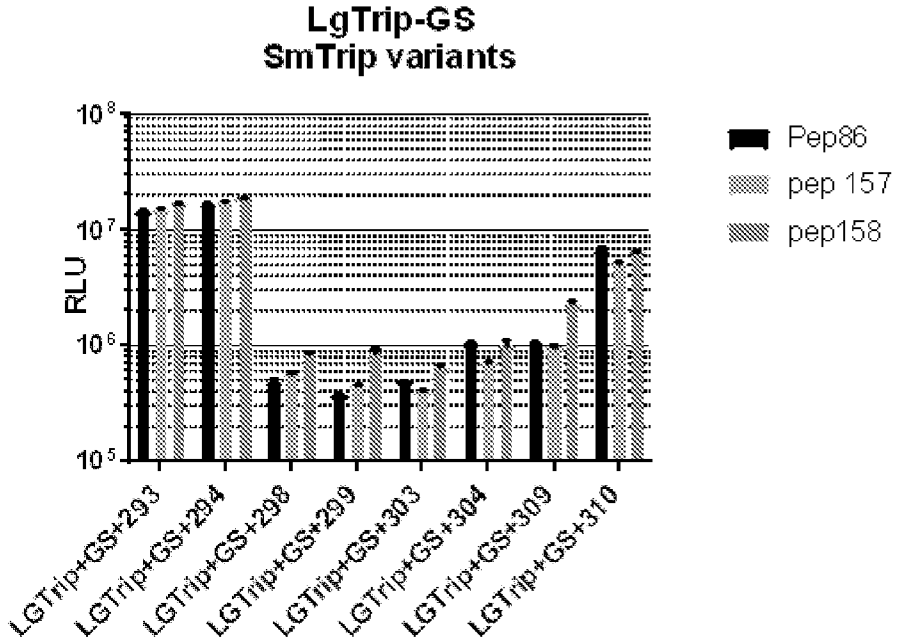


FIG. 26B

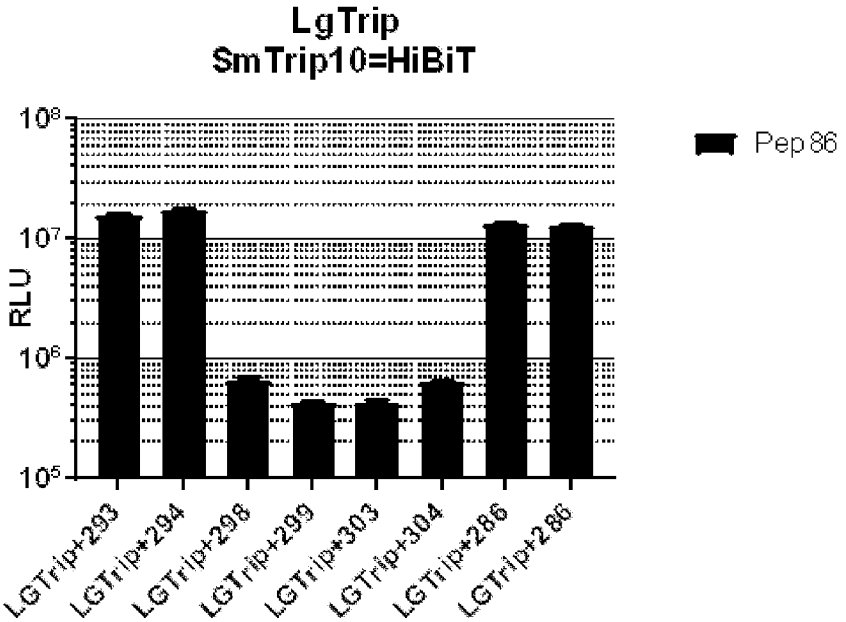


FIG. 26C

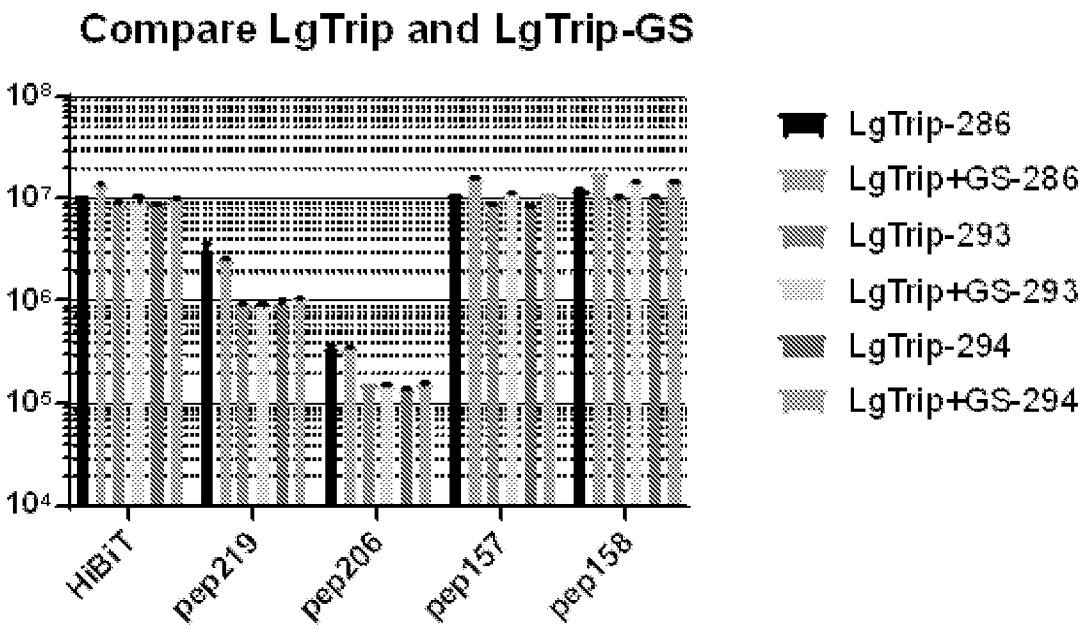
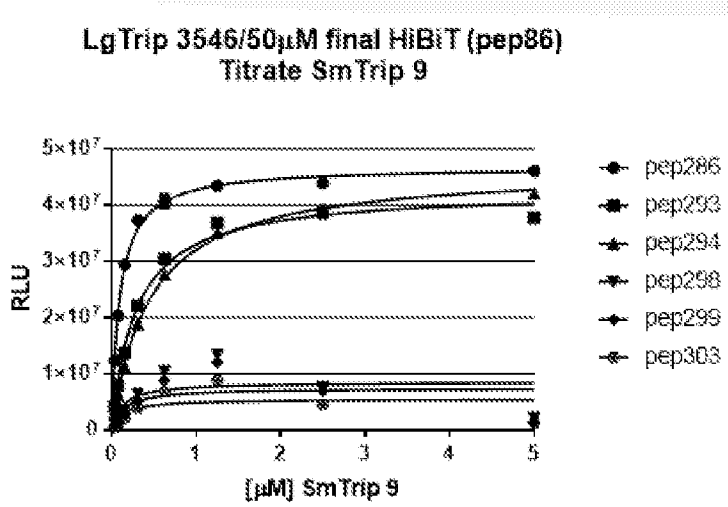

Compare LgTrip and LgTrip-GS
10^8
10^7
10^6
10^5
10^4
HiBiT
pep219
pep206
pep157
pep158
LgTrip-286
LgTrip+GS-286
LgTrip-293
LgTrip+GS-293
LgTrip-294
LgTrip+GS-294

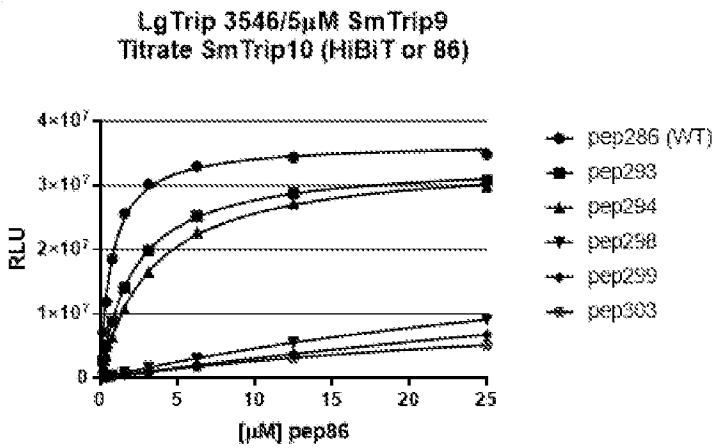
FIG. 27A



	pep286	pep293	pep294	pep298	pep299	pep303
Kd	0.09691	0.2926	0.4541	0.1292	0.1362	0.1321

Titration of β9 peptide in the presence of constant β10 and polypeptide

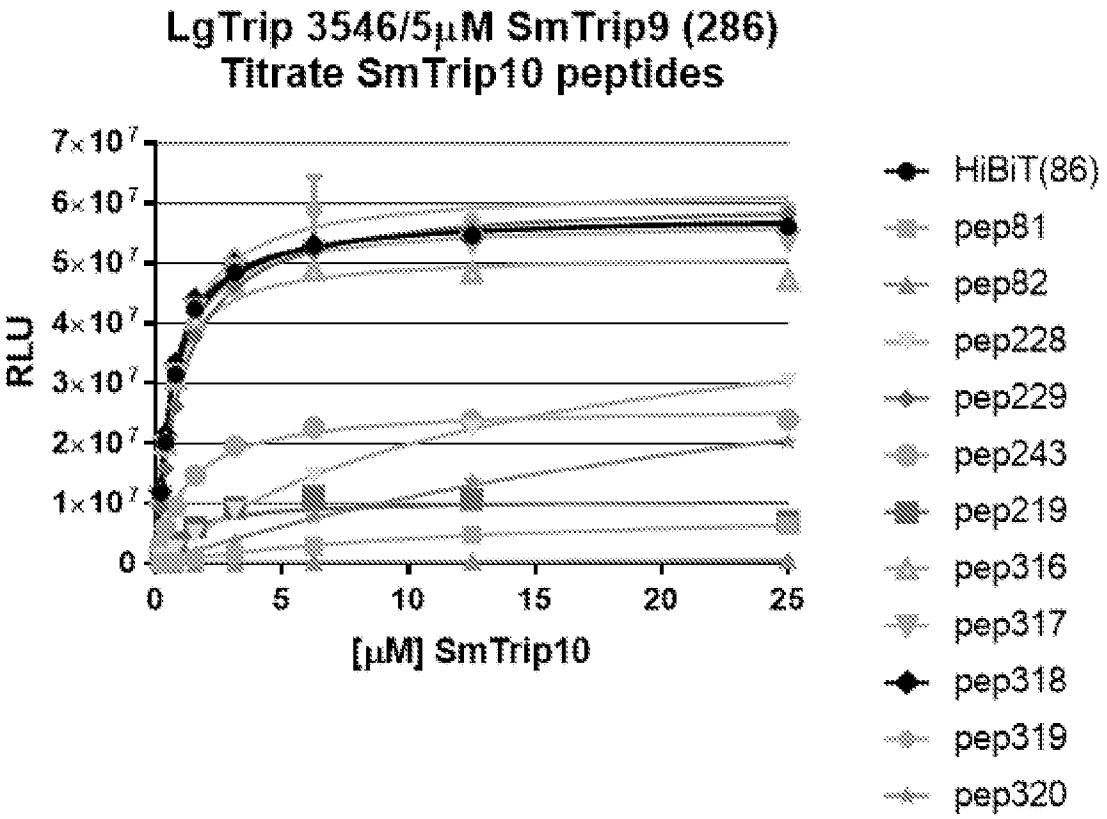
FIG. 27B



	pep286 (WT)	pep293	pep294	pep298	pep299	pep303
Kd	0.7467	2.174	3.311	45.2	97.08	42.38

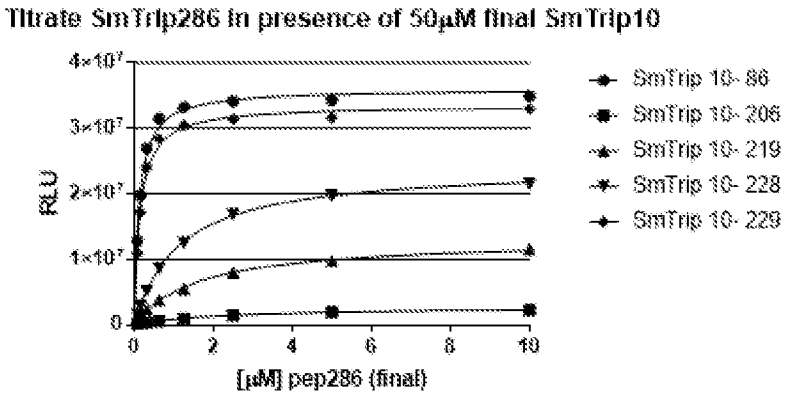
Titration of β10 peptide in the presence of constant β9 and polypeptide

FIG. 28



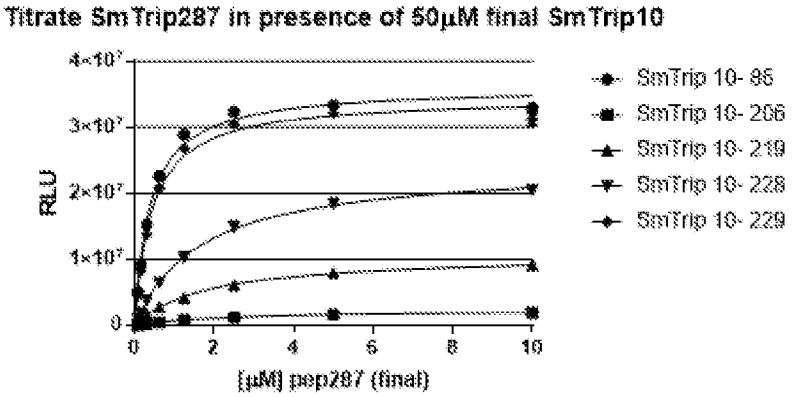
	Bmax	Kd
HiBiT(86)	58330355	0.6791
pep81	10033932	14.28
pep82	1093350	21.84
pep228	47058110	13.7
pep229	60603642	1.012
pep243	26193641	1.236
pep219	10913886	1.105
pep316	51521729	0.5413
pep317	57096749	0.654
pep318	57910752	0.6036
pep319	62774228	0.766
pep320	48479473	33.67

FIG. 29A



	SmTrip 10- 86	SmTrip 10- 206	SmTrip 10- 219	SmTrip 10- 228	SmTrip 10- 229
Kd	0.1225	2.183	1.614	1.13	0.1409

FIG. 29B



	SmTrip 10- 86	SmTrip 10- 206	SmTrip 10- 219	SmTrip 10- 228	SmTrip 10- 229
Kd	0.3971	2.54	1.833	1.656	0.4337

FIG. 30

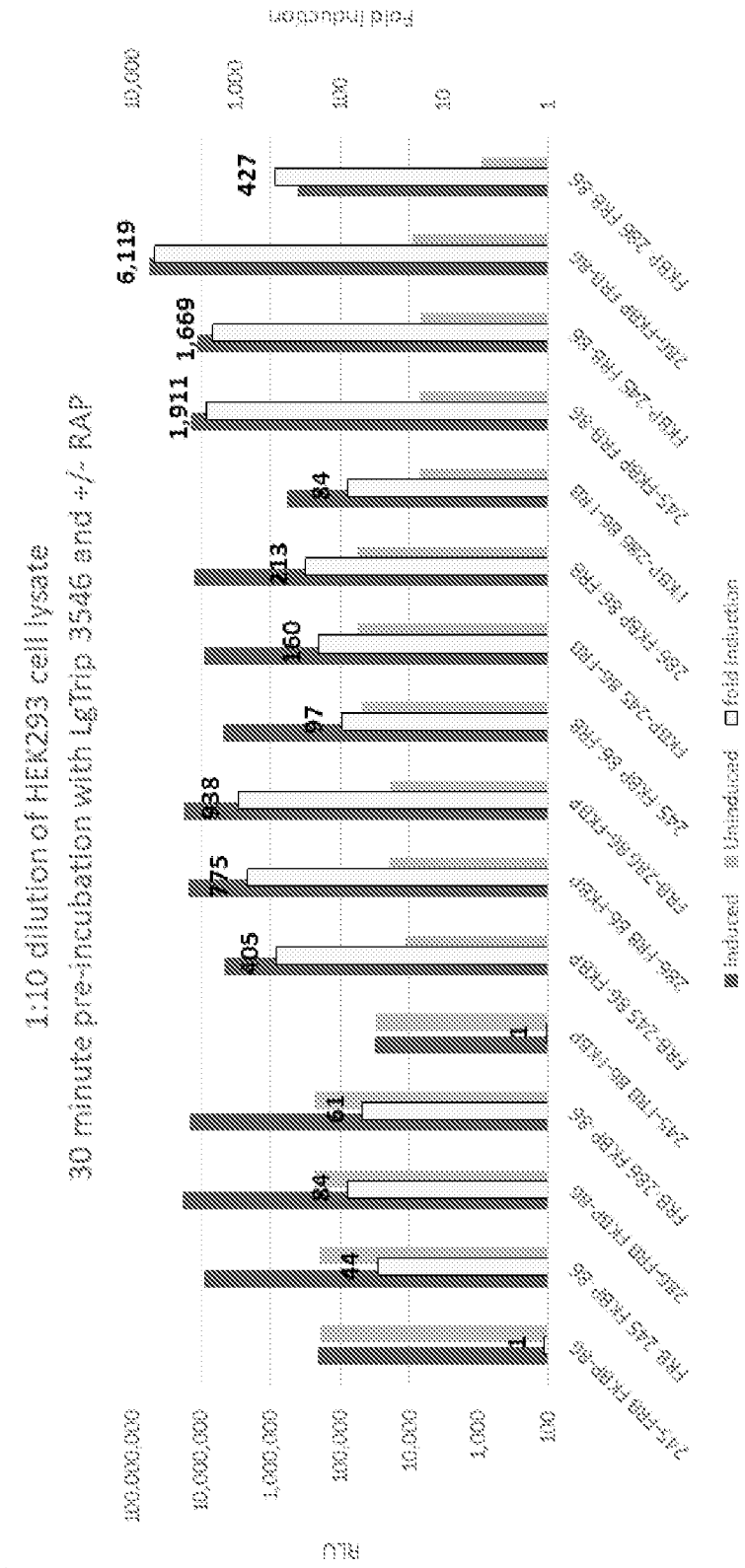

1:10 dilution of HEK293 cell lysate
30 minute pre-incubation with LgTrip 3546 and +/- RAP
RLU
Fold Induction
Induced
Uninduced
fold induction
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
10,000
1,000
100
10
1
6,119
1,911
1,669
427
213
160
97
84
61
44
34
1
1
1
2
5
5
8
5
7
5

FIG. 31

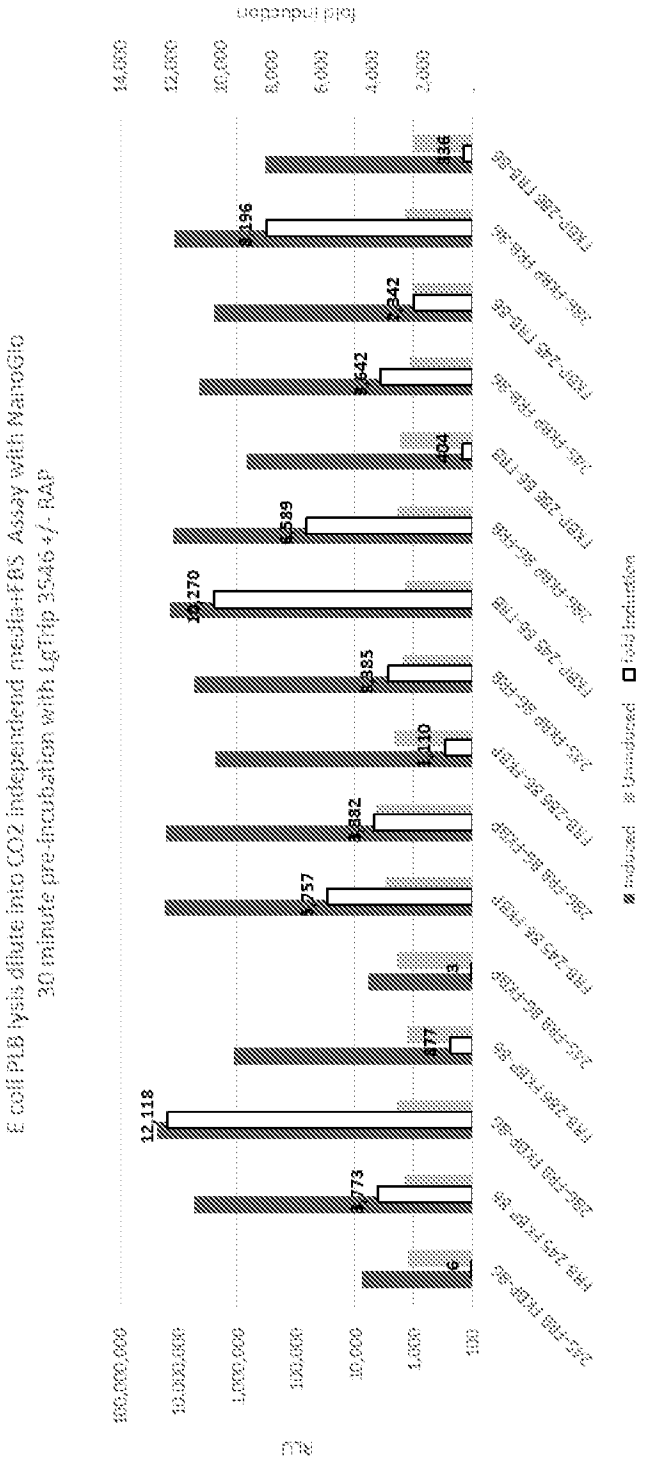

E coli PLB lysis dilute into CO2 independent media+FBS Assay with ManoGlo
30 minute pre-incubation with LgTrip 35486 +/- RAP
RLU
fold induction
Induced
Uninduced
fold induction
100
1,000
10,000
100,000
1,000,000
10,000,000
100,000,000
2,000
4,000
6,000
8,000
10,000
12,000
14,000
6
6,773
12,118
77
3
1,757
5,382
1,310
5,385
10,270
5,589
434
5,642
1,242
8,196
136

FIG. 32

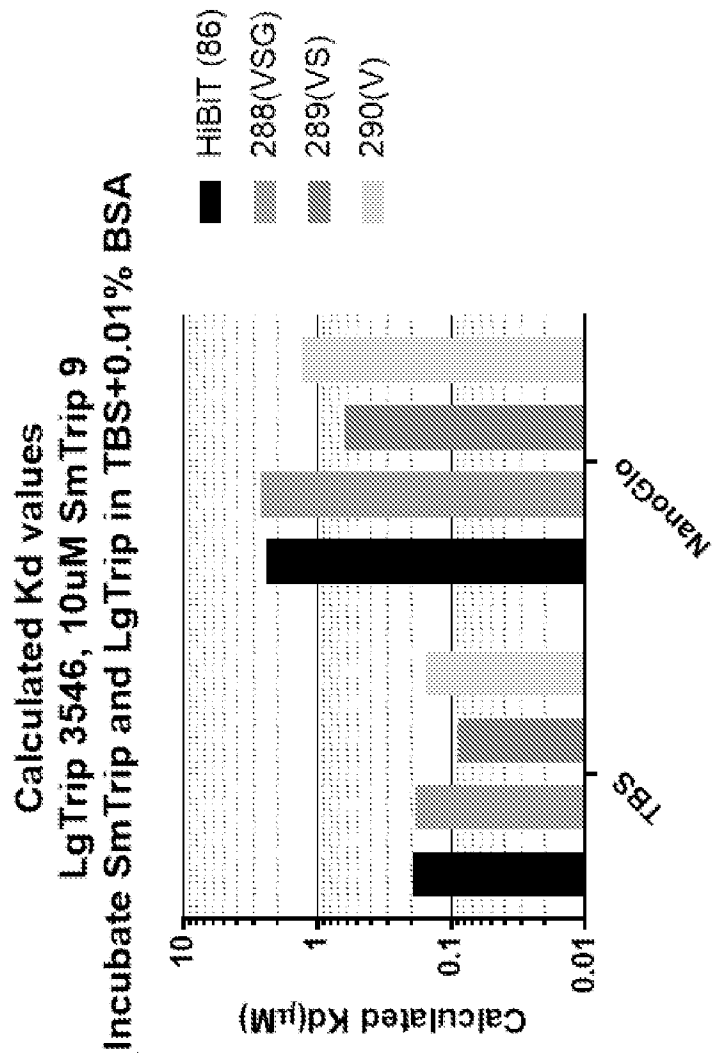
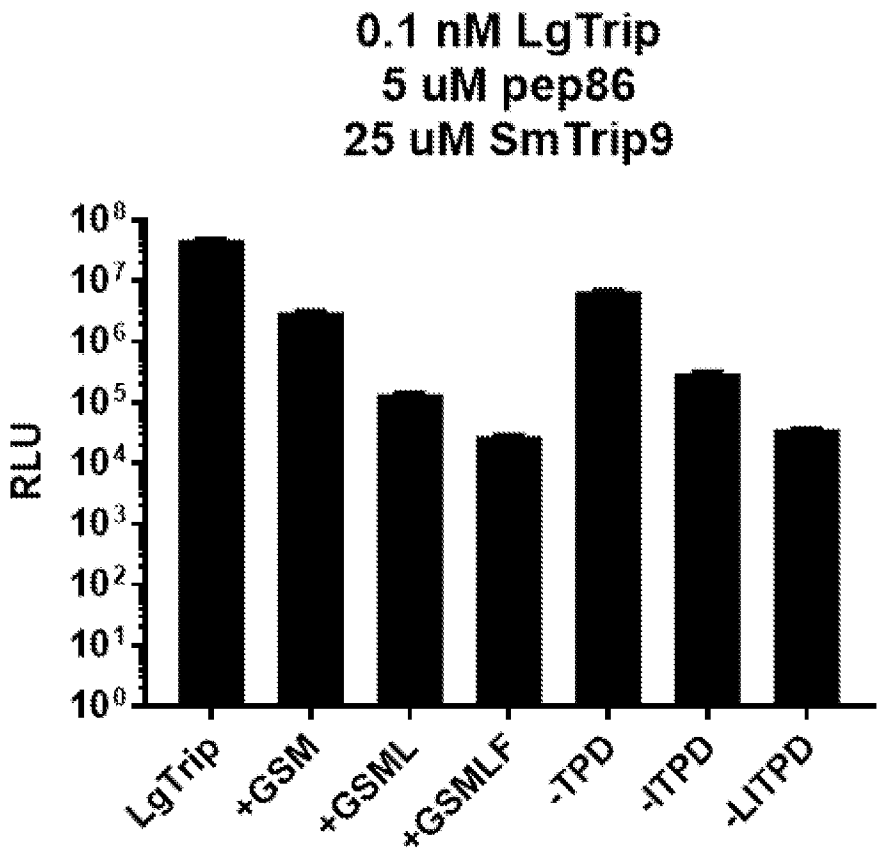
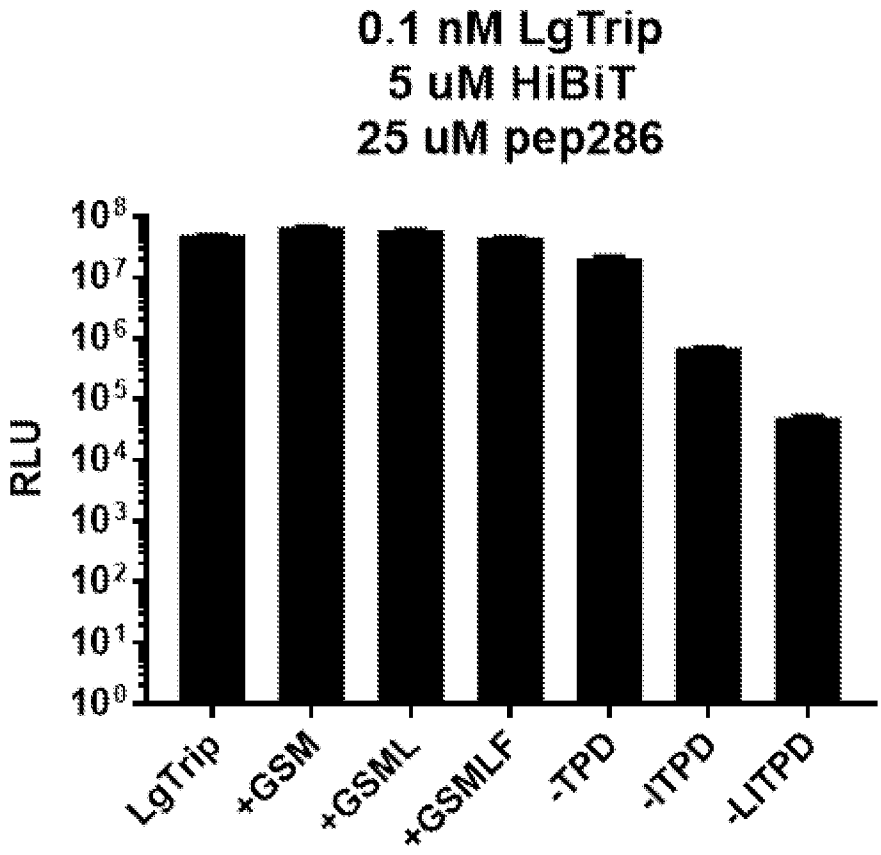

Calculated Kd values
LgTrip 3546, 10uM SmTrip 9
(Incubate SmTrip and LgTrip in TBS+0.01% BSA
HiBiT (86)
288(VSG)
289(VS)
290(V)
Calculated Kd(μM)
10
1
0.1
0.01
TBS
NanoGlo

FIG. 33



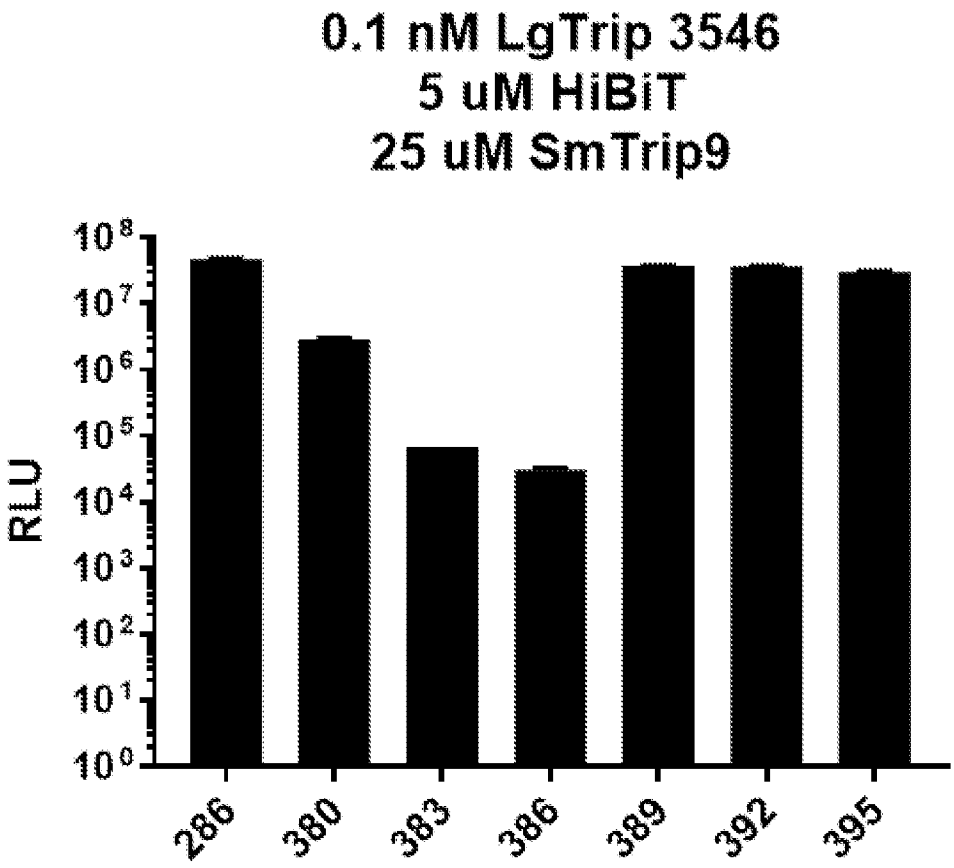
ATG#	LgTrip	SmTrip 9	Pep Sequence	SEQ ID NO.
3546	LgTrip	286	SSWKRGSMLFRVTINS	37
3722	LgTrip + GSM	380	SSWKRLFRVTINS	180
3723	LgTrip + GSML	383	SSWKRFRVTINS	181
3724	LgTrip + GSMLF	386	SSWKRRVTINS	182
3725	LgTrip – TPD	389	SSWKRTPDGSMLFRVTINS	183
3726	LgTrip – ITPD	392	SSWKRITPDGSMLFRVTINS	184
3727	LgTrip - LITPD	395	SSWKRLITPDGSMLFRVTINS	185

FIG. 34



ATG#	LgTrip
3546	LgTrip
3722	LgTrip + GSM
3723	LgTrip + GSML
3724	LgTrip + GSMLF
3725	LgTrip – TPD
3726	LgTrip – ITPD
3727	LgTrip - LITPD

FIG. 35



SmTrip 9	Pep Sequence	SEQ ID NO.
286	SSWKRGSMLFRVTINS	37
380	SSWKRLFRVTINS	180
383	SSWKRFRVTINS	181
386	SSWKRRVTINS	182
389	SSWKRTPDGSMLFRVTINS	183
392	SSWKRITPDGSMLFRVTINS	184
395	SSWKRLITPDGSMLFRVTINS	185

FIG. 36

SmTrip9 Titrations with 0.05 nM LgTrip and 25 uM HiBiT				
Pep ID	**Pep Sequence**	**SEQ ID NO.**	**SmTrip9 Kd (uM)**	**Bmax (RLU)**
286	SSWKRGSMLFRVTINS	37	0.09	1.65E7
292	SSWKRMLFRVTINS	153	0.10	1.65E7
305	SSWKRGSMLFRVTIN	164	0.60	1.16E7
306	SSWKRGSMLFRVTI	165	0.17	1.77E7
307	SSWKRSMLFRVTIN	166	0.10	1.47E7
308	SSWKRMLFRVTIN	167	0.11	1.55E7
312	SSWKRMLFRVTI	171	0.05	1.77E7

FIG. 37

HiBiT Titrations with 0.05 nM LgTrip and 5 uM SmTrip9 pep286				
Pep ID	**Pep Sequence**	**SEQ ID NO.**	**SmTrip9 Kd (uM)**	**Bmax (RLU)**
286	SSWKRGSMLFRVTINS	37	0.47	1.75E7
292	SSWKRMLFRVTINS	153	0.47	1.69E7
305	SSWKRGSMLFRVTIN	164	3.81	1.55E7
306	SSWKRGSMLFRVTI	165	0.66	1.86E7
307	SSWKRSMLFRVTIN	166	1.22	1.42E7
308	SSWKRMLFRVTIN	167	0.60	1.59E7
312	SSWKRMLFRVTI	171	0.25	1.79E7

FIG. 38

SmTrip9 Titrations with 0.1 nM LgTrip and 25 uM HiBiT				
SEQ ID NO:	Pep ID	Sequence	SmTrip9 Kd (uM)	Bmax (RLU)
37	286	SSWKRGSMLFRVTINS	0.19	3.45E7
190	401	SSWKRGSMLYRVTINS	0.16	1.69E6
191	402	SSWKRGSMLWRVTINS	0.02	2.29E5
192	403	SSWKRGSMLHRVTINS	0.07	1.01E6
193	404	SSWKRGSLLFRVTINS	0.18	3.41E7
194	405	SSWKRGSKLFRVTINS	33.07	7.12E6
195	406	SSWKRGSRLFRVTINS	9.23	2.65E6
196	407	SSWKRGSFLFRVTINS	0.65	1.51E7
197	408	SSWKRGSWLFRVTINS	3.93	4.64E6
198	409	SSWKRGSMLFRVSINS	0.47	3.48E7
199	410	SSWKRGSMLFRVQINS	0.46	2.58E7
200	411	SSWKRGSMLFRVNINS	1.69	2.99E7

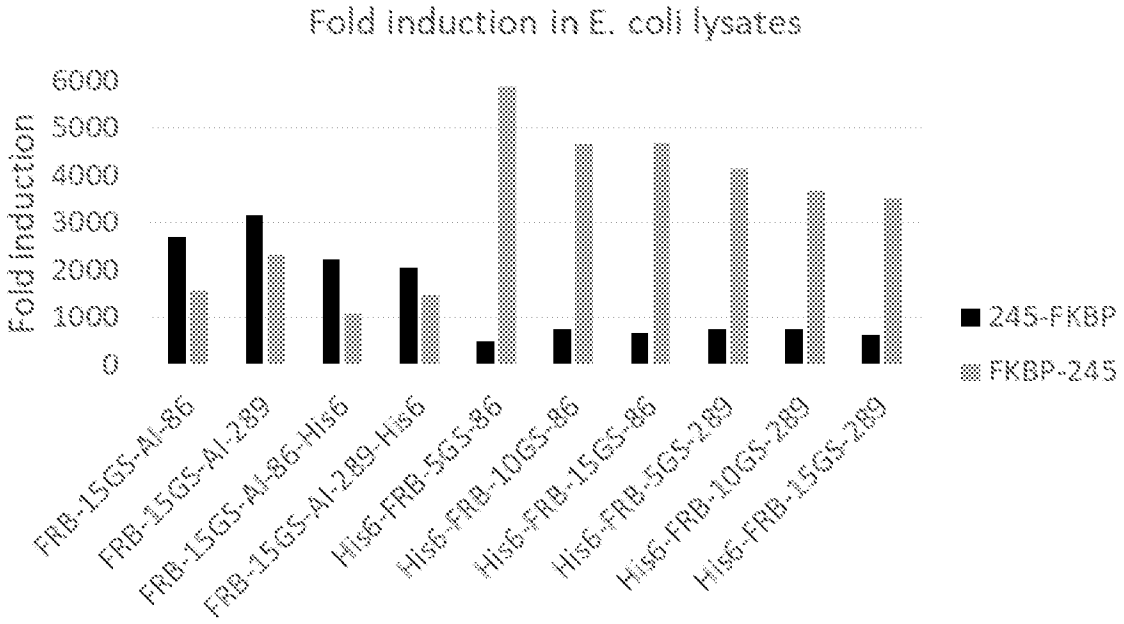
FIG. 39

SmTrip9 Titrations with 0.1 nM LgTrip and 25 uM HiBiT				
SEQ ID NO:	Pep ID	Sequence	SmTrip9 Kd (uM)	Bmax (RLU)
37	286	SSWKRGSMLFRVTINS	0.19	3.45E7
153	292	SSWKRMLFRVTINS	0.21	2.93E7
157	297	SSWKRMLFRVTINSV	0.2	2.06E7
160	302	SSWKRMLFRVTINSVS	0.52	1.27E7
186	396	SSRGSMLFRVTINSWK	0.66	2.89E7
187	397	SKRGSMLFRVTINSWS	0.37	2.80E7
188	398	SWRGSMLFRVTINS	1.28	2.33E7
297	399	SSKRGSMLFRVTIWS	0.12	4.02E7
189	400	SSRGSMLFRVTIWK	0.07	4.06E7

FIG. 40

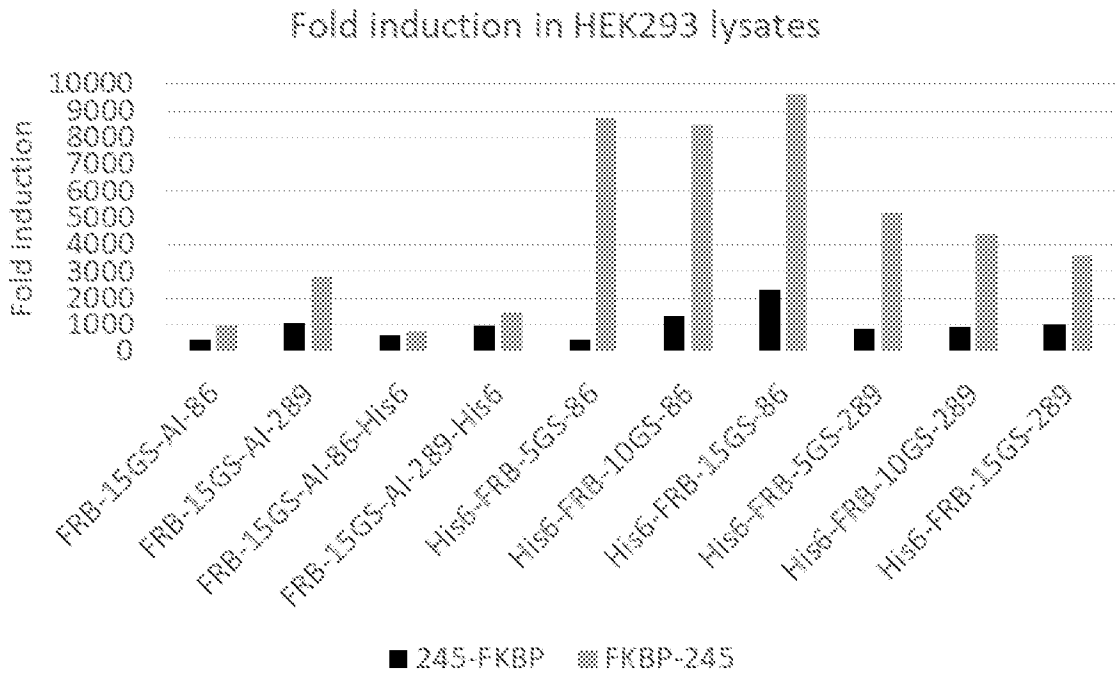
	SmTrip9 or HiBiT Titrations with 0.1 nM LgTrip			
SEQ ID NO:	Pep ID	Sequence	SmTrip9 Kd in uM (Bmax in RLU)	HiBiT Kd in uM (Bmax in RLU)
37	286	SSWKRGSMLFRVTINS	0.19 (3.45E7)	0.73 (3.66E7)
153	292	SSWKRMLFRVTINS	0.21 (2.93E7)	1.55 (3.05E7)
157	297	SSWKRMLFRVTINSV	0.2 (2.06E7)	5.09 (1.36E7)
160	302	SSWKRMLFRVTINSVS	0.52 (1.27E7)	8.15 (1.75E7)

FIG. 41



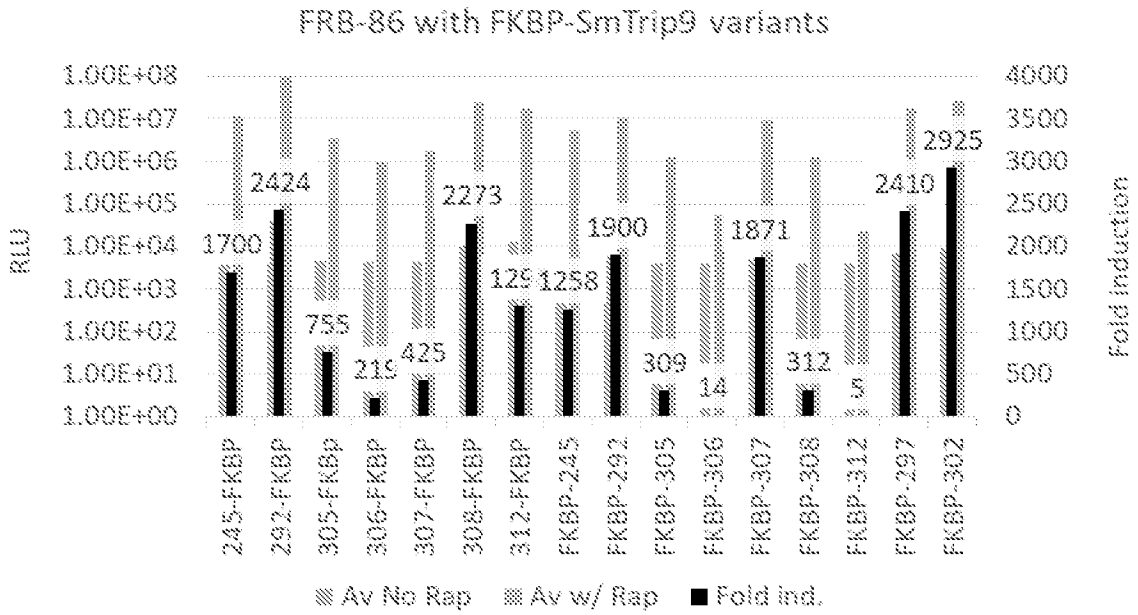
ATG#	Detail
1634	FRB-15GS-AI-86
3586	FRB-15GS-AI-289
3743	FRB-15GS-AI-86-His6
3744	FRB-15GS-AI-289-His6
3760	His6-FRB-5GS-86
3761	His6-FRB-10GS-86
3762	His6-FRB-15GS-86
3763	His6-FRB-5GS-289
3764	His6-FRB-10GS-289
3765	His6-FRB-15GS-289

FIG. 42



ATG#	Detail
1634	FRB-15GS-AI-86
3586	FRB-15GS-AI-289
3743	FRB-15GS-AI-86-His6
3744	FRB-15GS-AI-289-His6
3760	His6-FRB-5GS-86
3761	His6-FRB-10GS-86
3762	His6-FRB-15GS-86
3763	His6-FRB-5GS-289
3764	His6-FRB-10GS-289
3765	His6-FRB-15GS-289

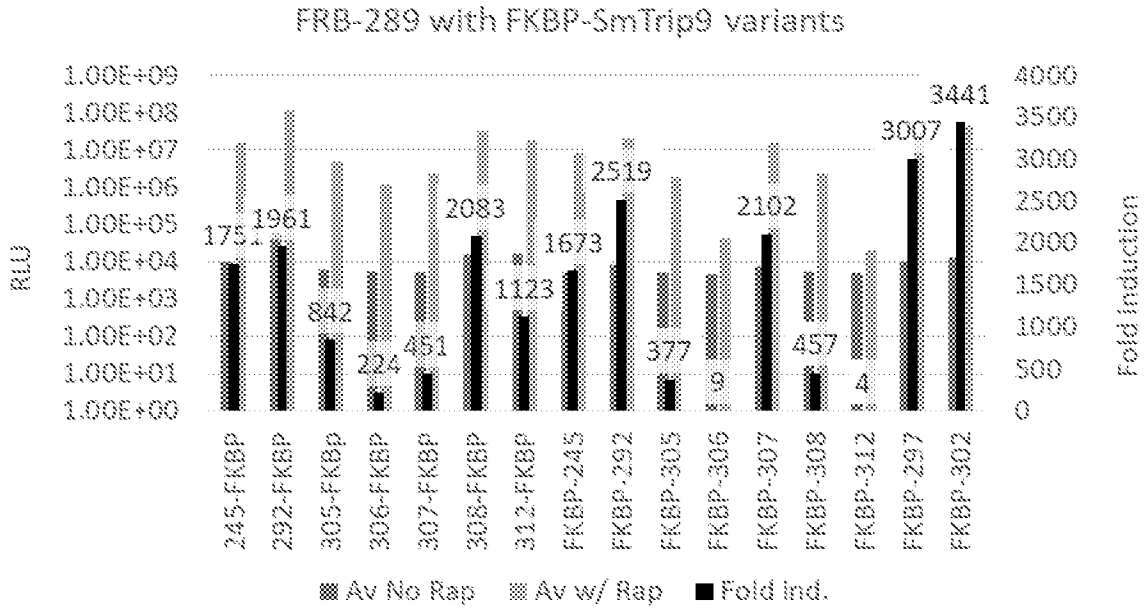
FIG. 43



Pep	Sequence
245	GSMLFRVTINS
292.x	MLFRVTINS
297.x	MLFRVTINSV
302.x	MLFRVTINSVS
305.x	GSMLFRVTIN
306.x	GSMLFRVTI
307.x	SMLFRVTIN
308.x	MLFRVTIN
312.x	MLFRVTI

Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS
305.x	292	GSMLFRVTIN
306.x	293	GSMLFRVTI
307.x	294	SMLFRVTIN
308.x	295	MLFRVTIN
312.x	296	MLFRVTI

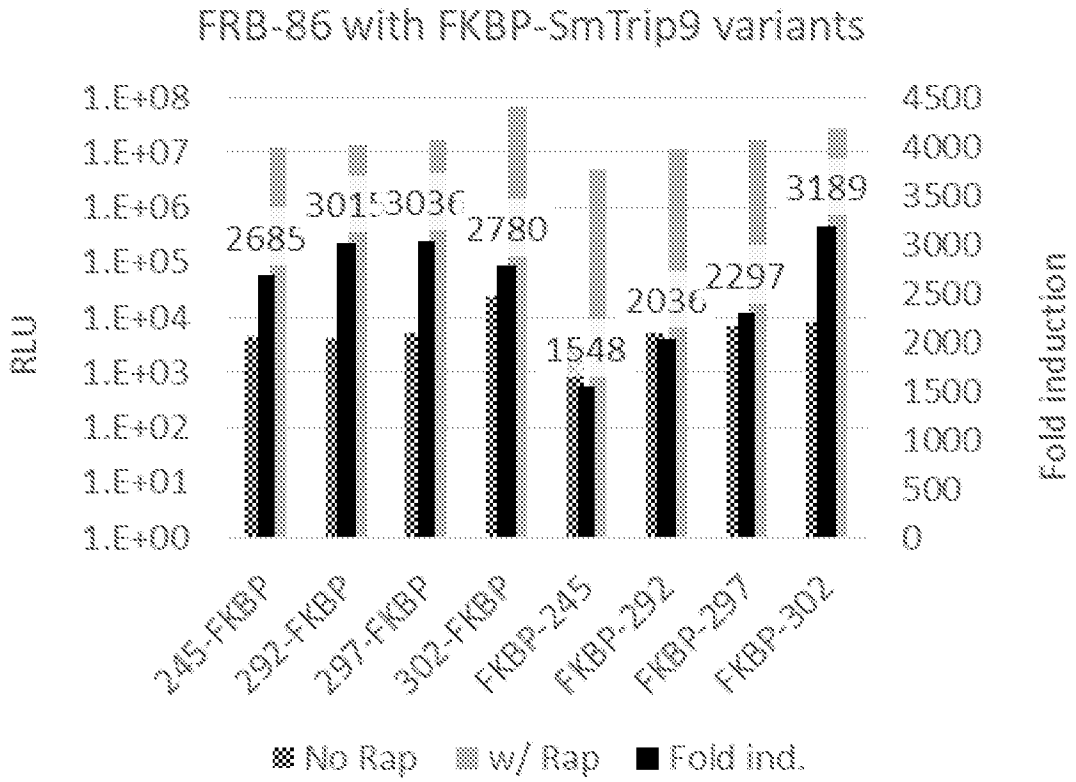
FIG. 44



Pep	Sequence
245	GSMLFRVTINS
292.x	MLFRVTINS
297.x	MLFRVTINSV
302.x	MLFRVTINSVS
305.x	GSMLFRVTIN
306.x	GSMLFRVTI
307.x	SMLFRVTIN
308.x	MLFRVTIN
312.x	MLFRVTI

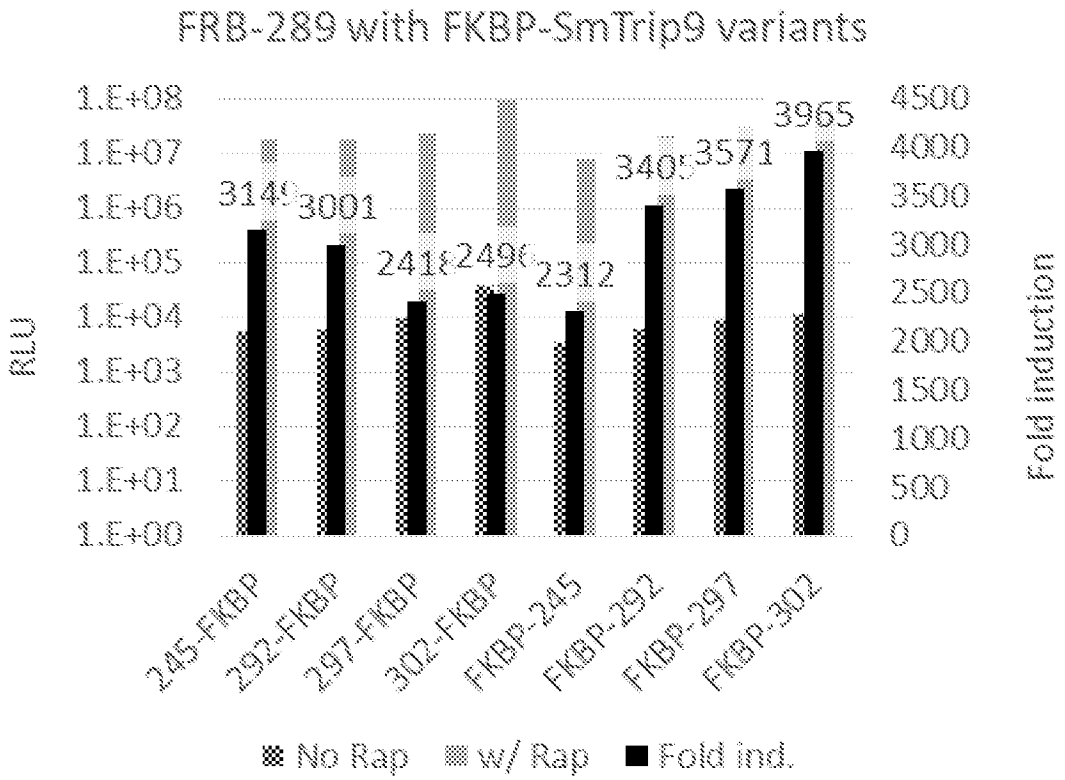
Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS
305.x	292	GSMLFRVTIN
306.x	293	GSMLFRVTI
307.x	294	SMLFRVTIN
308.x	295	MLFRVTIN
312.x	296	MLFRVTI

FIG. 45



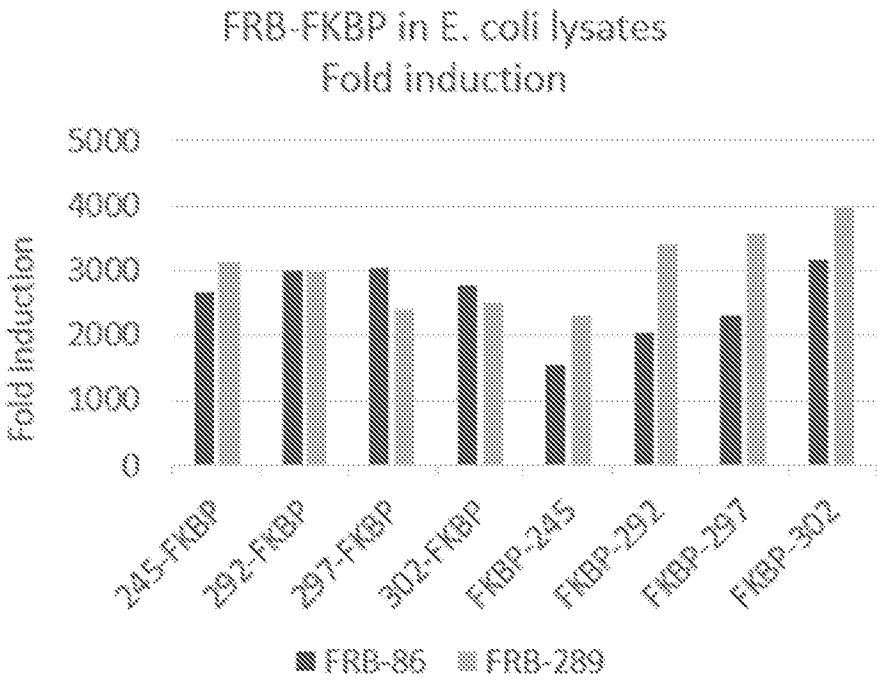
Pep ID	SEQ ID NO.	Sequence
89	824	VSGWRLFKKIS
289	826	VSVSGWRLFKKIS
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS

FIG. 46



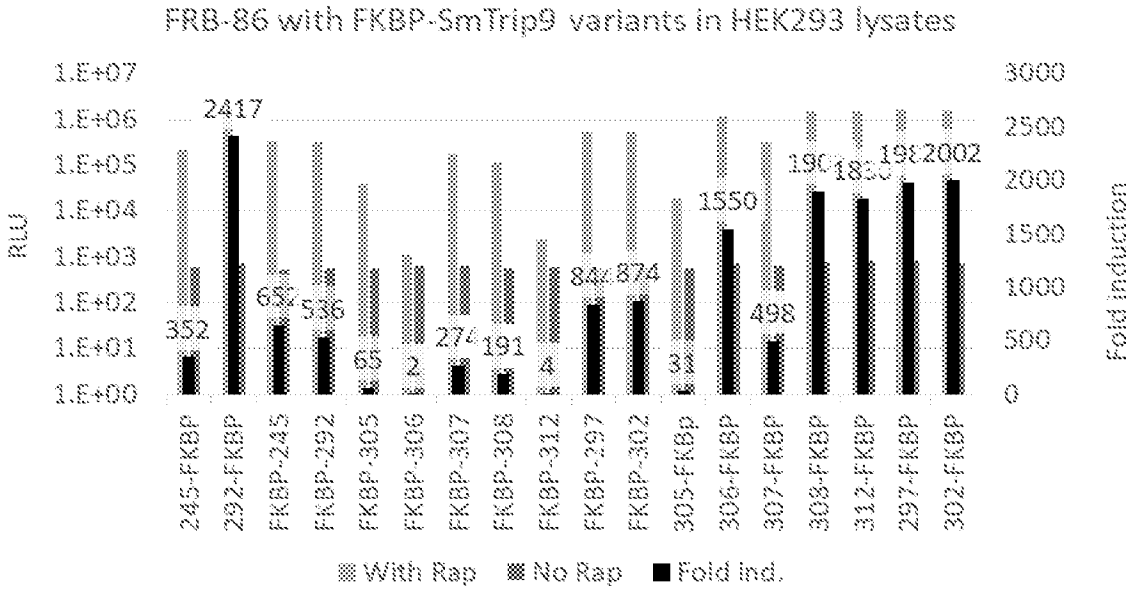
Pep ID	SEQ ID NO.	Sequence
89	824	VSGWRLFKKIS
289	826	VSVSGWRLFKKIS
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS

FIG. 47



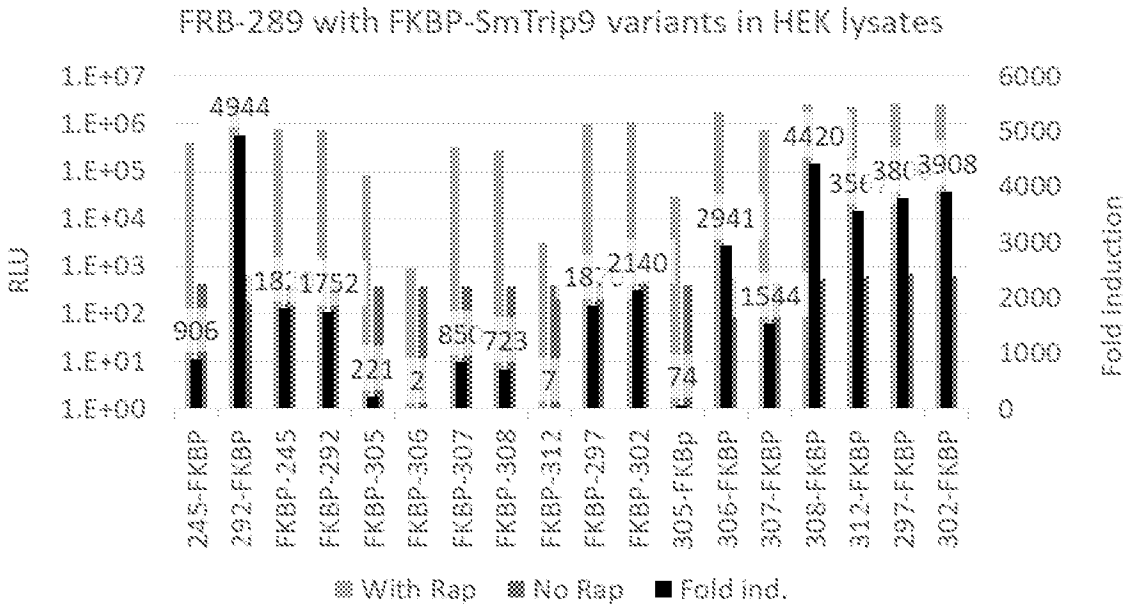
Pep ID	SEQ ID NO.	Sequence
89	824	VSGWRLFKKIS
289	826	VSVSGWRLFKKIS
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS

FIG. 48



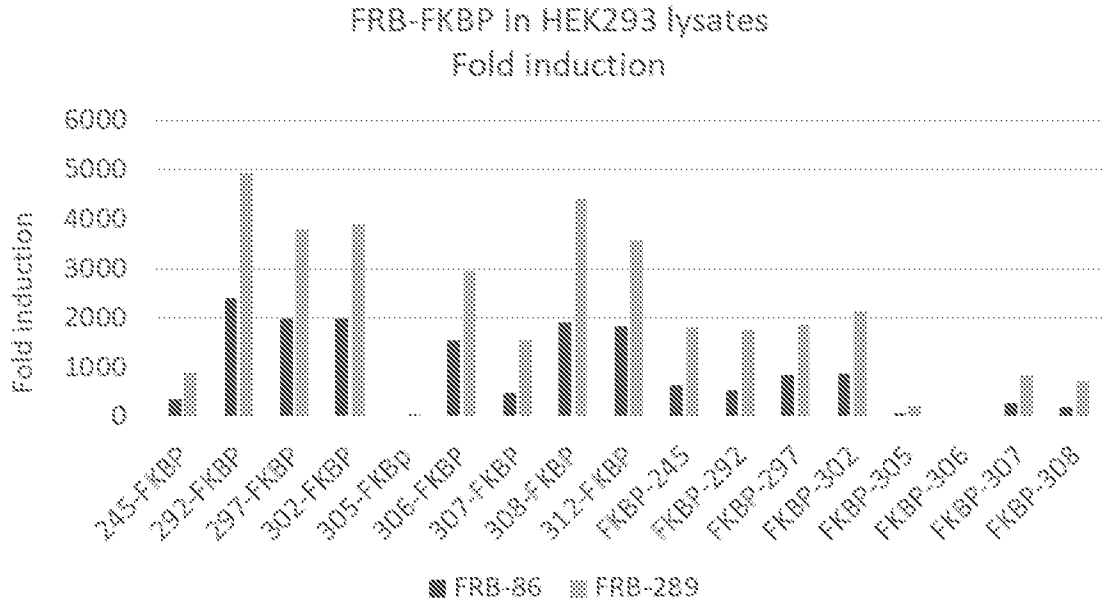
Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS
305.x	292	GSMLFRVTIN
306.x	293	GSMLFRVTI
307.x	294	SMLFRVTIN
308.x	295	MLFRVTIN
312.x	296	MLFRVTI

FIG. 49



Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS
305.x	292	GSMLFRVTIN
306.x	293	GSMLFRVTI
307.x	294	SMLFRVTIN
308.x	295	MLFRVTIN
312.x	296	MLFRVTI

FIG. 50



Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS
305.x	292	GSMLFRVTIN
306.x	293	GSMLFRVTI
307.x	294	SMLFRVTIN
308.x	295	MLFRVTIN
312.x	296	MLFRVTI

FIG. 51A

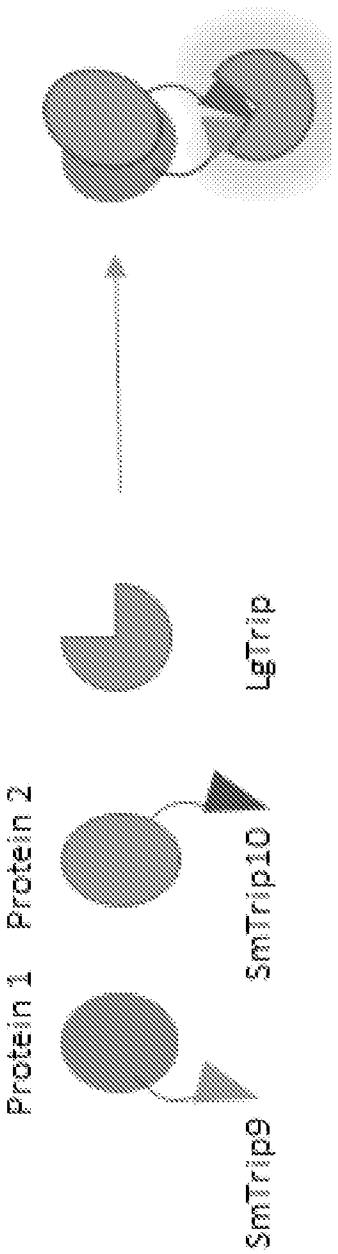


FIG. 51B

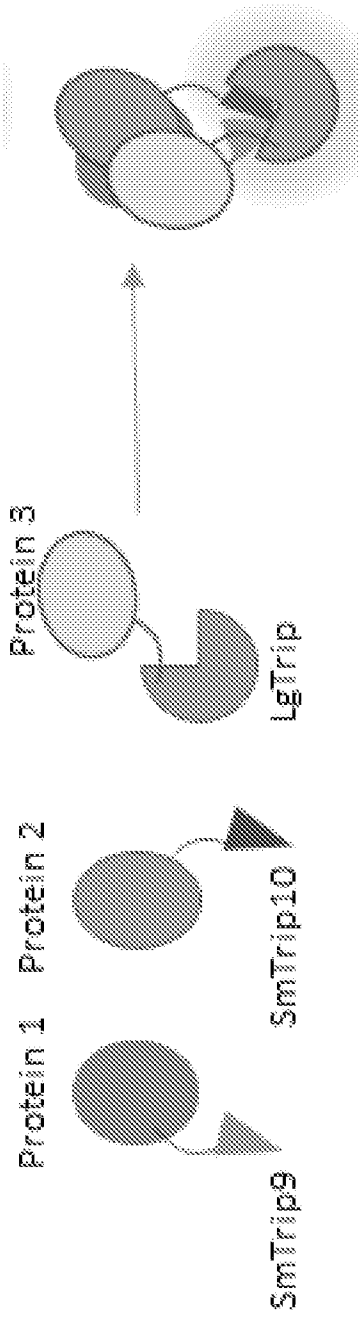


FIG. 51C

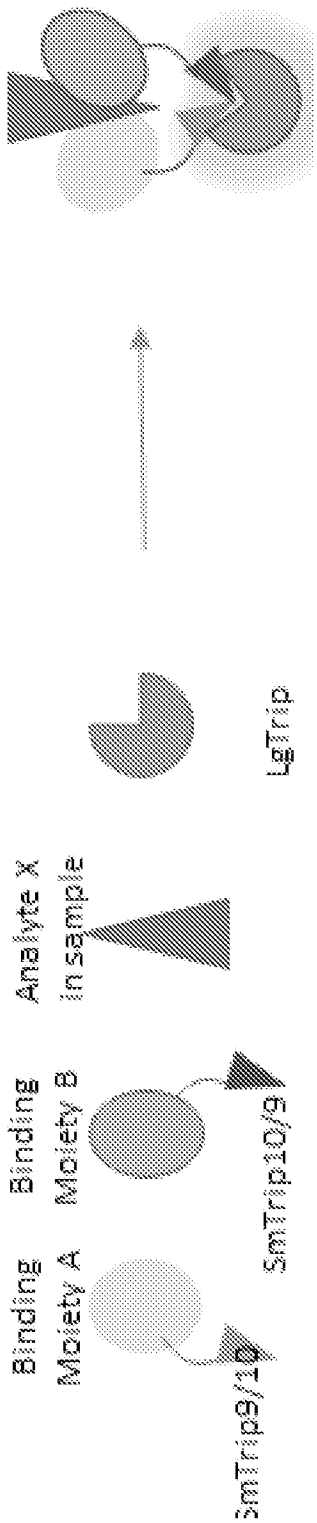
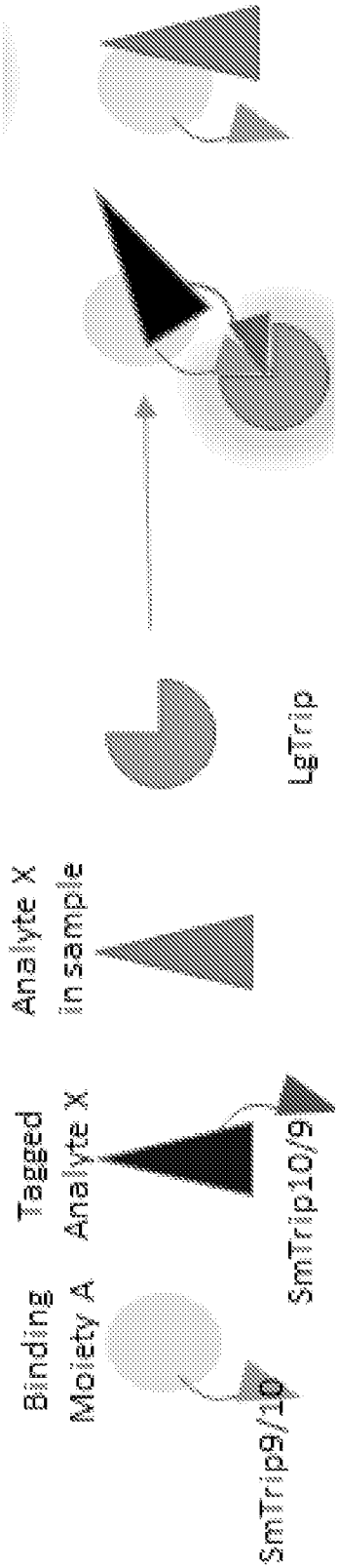


FIG. 51D



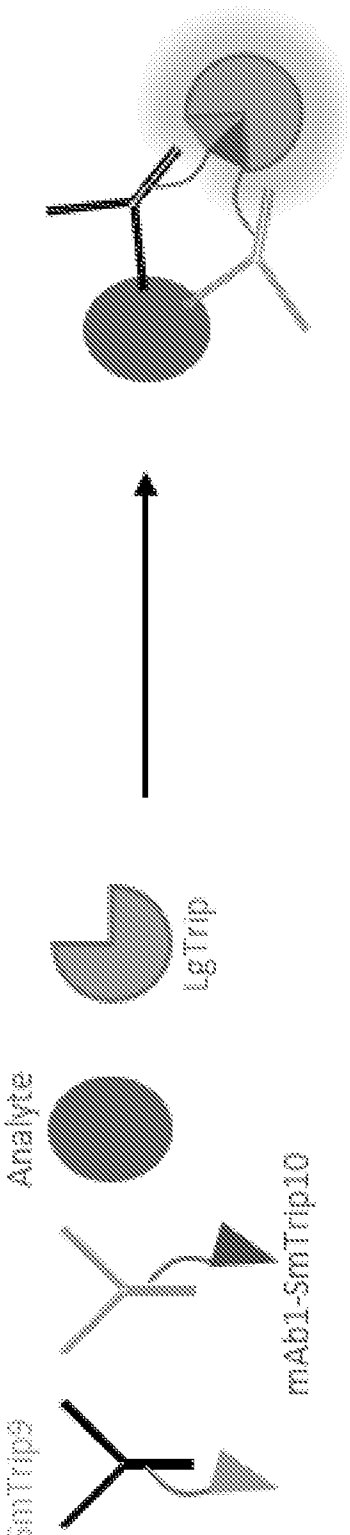


FIG. 51E

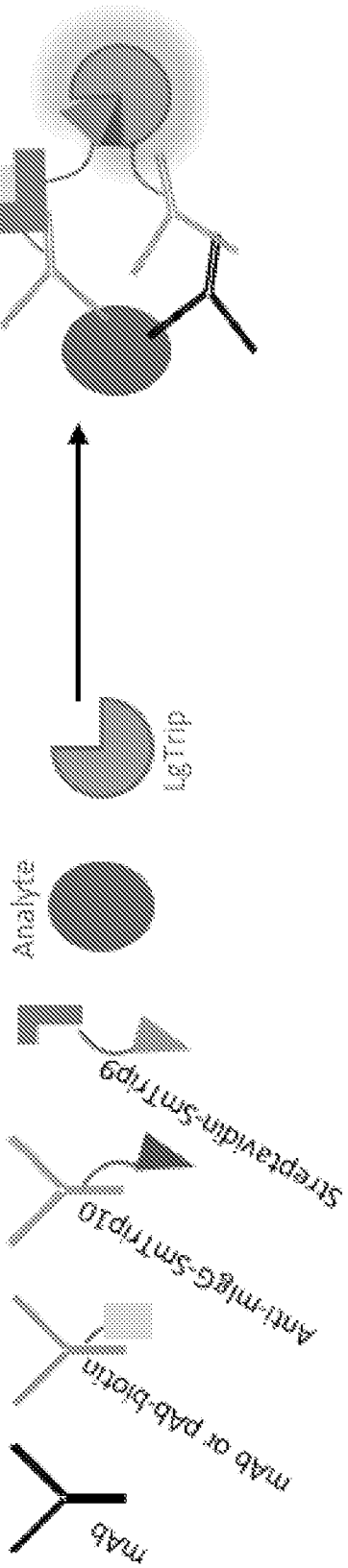


FIG. 51F

FIG. 51G

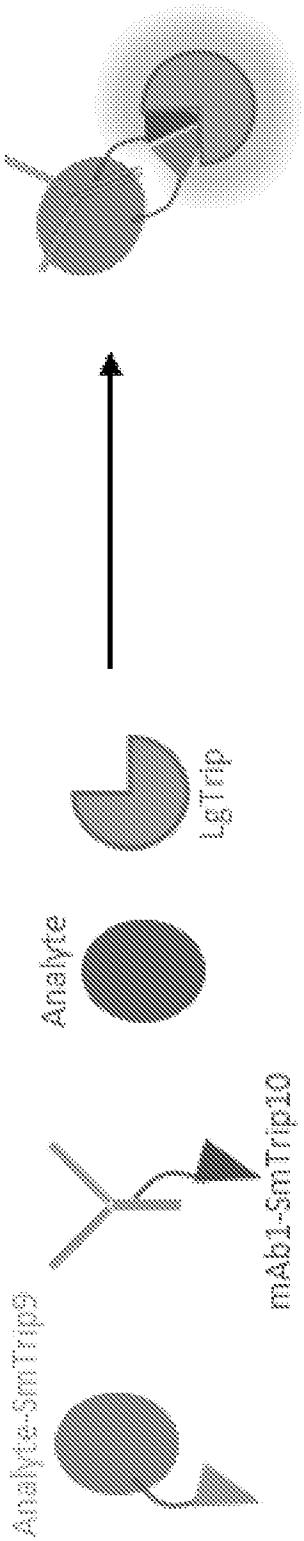


FIG. 51H

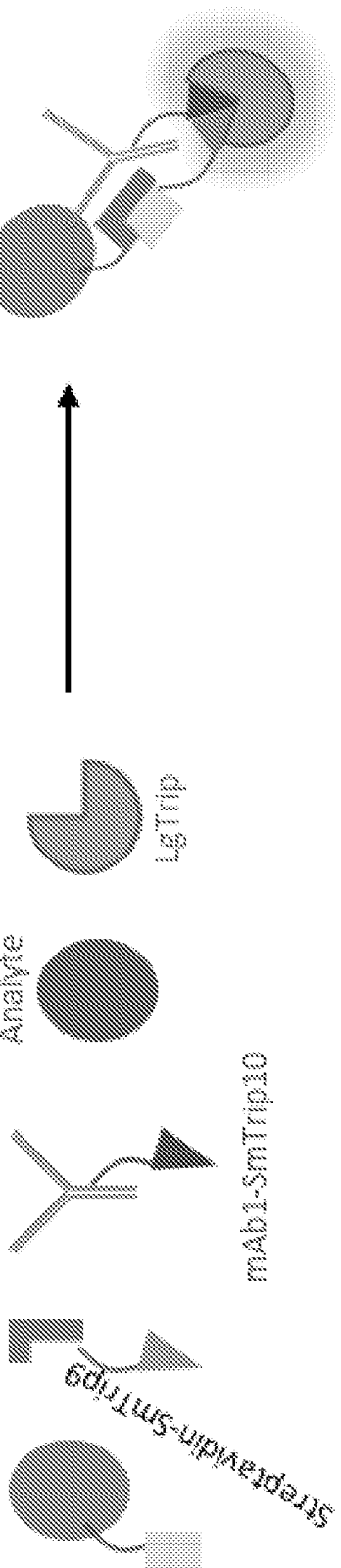


FIG. 52

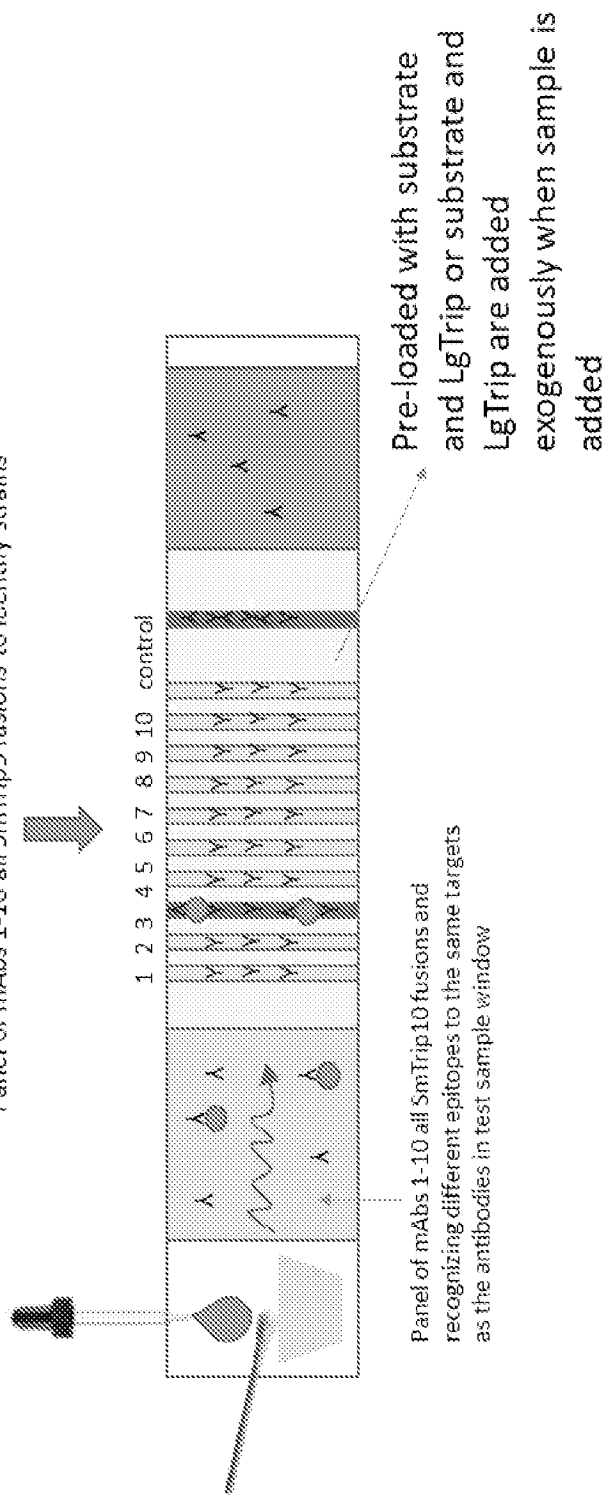

Panel of mAbs 1-10 all SmTrip9 fusions to identify strains
1 2 3 4 5 6 7 8 9 10 control
Pre-loaded with substrate and LgTrip or substrate and LgTrip are added exogenously when sample is added
Panel of mAbs 1-10 all SmTrip10 fusions and recognizing different epitopes to the same targets as the antibodies in test sample window

FIG. 53

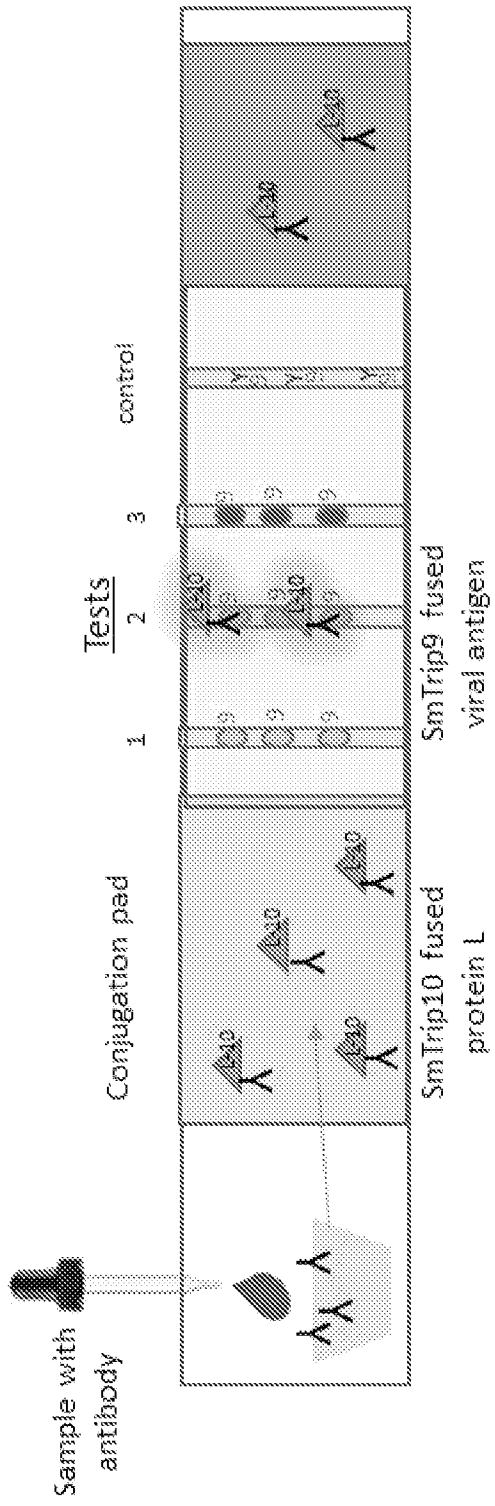

Sample with antibody
Conjugation pad
Tests
1
2
3
control
SmTrip10 fused protein L
SmTrip9 fused viral antigen

FIG. 54

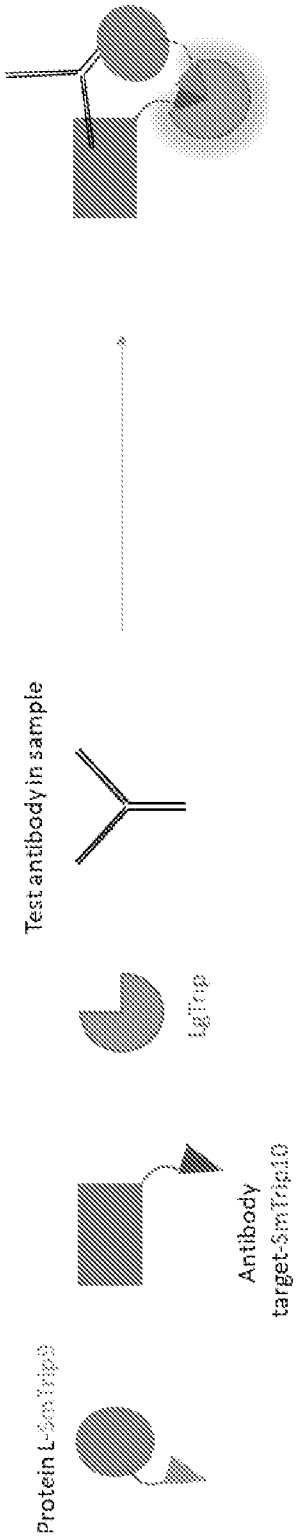

Test antibody in sample
IgTrap
Antibody target-SmTrap10
Protein L-SmTrap3

FIG. 55

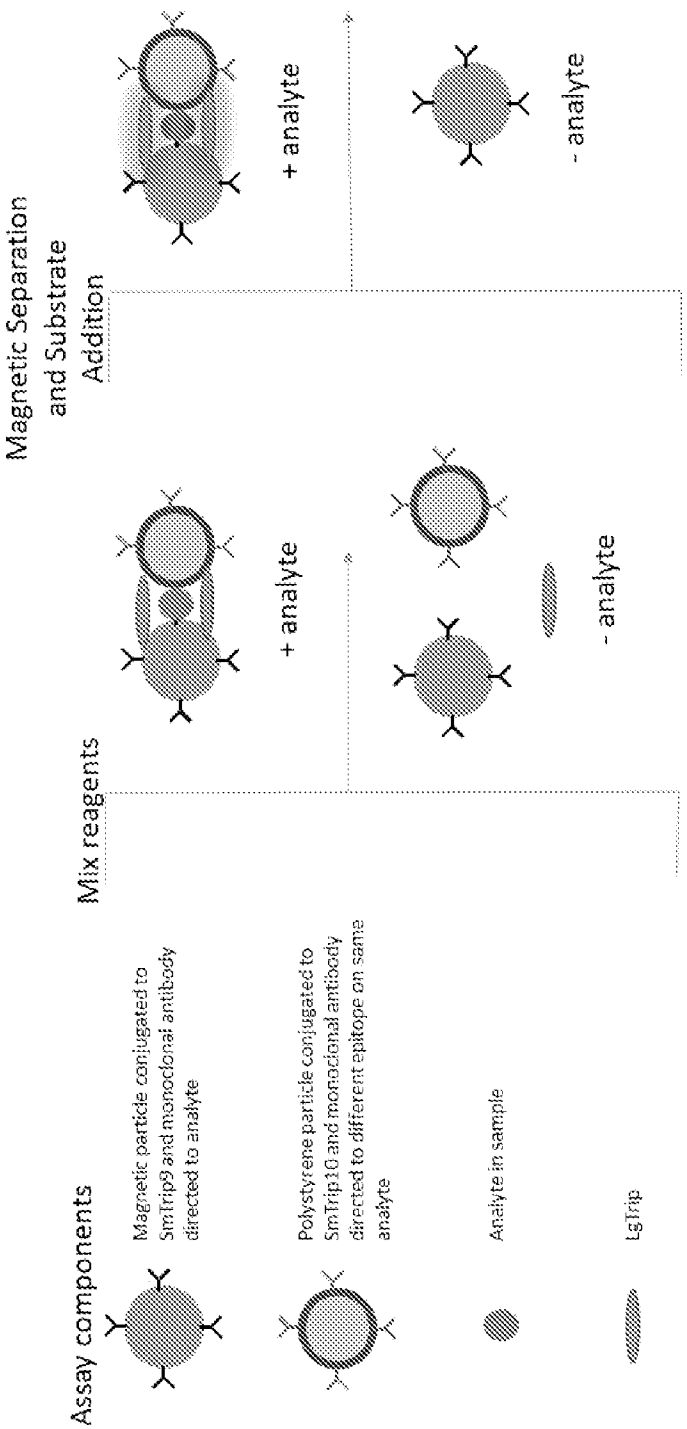

Assay components
Magnetic particle conjugated to SmTrip9 and monoclonal antibody directed to analyte
Polystyrene particle conjugated to SmTrip10 and monaclonal antibody directed to different epitope on same analyte
Analyte in sample
LgTrip
Mix reagents
+ analyte
- analyte
Magnetic Separation and Substrate Addition
+ analyte
- analyte

FIG. 56

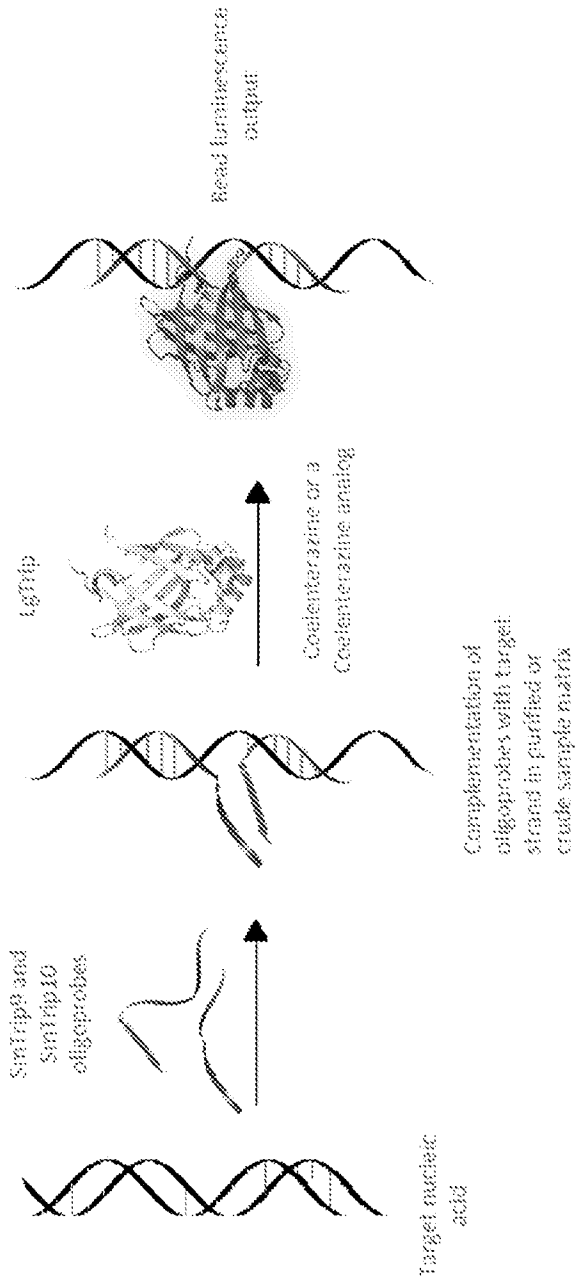
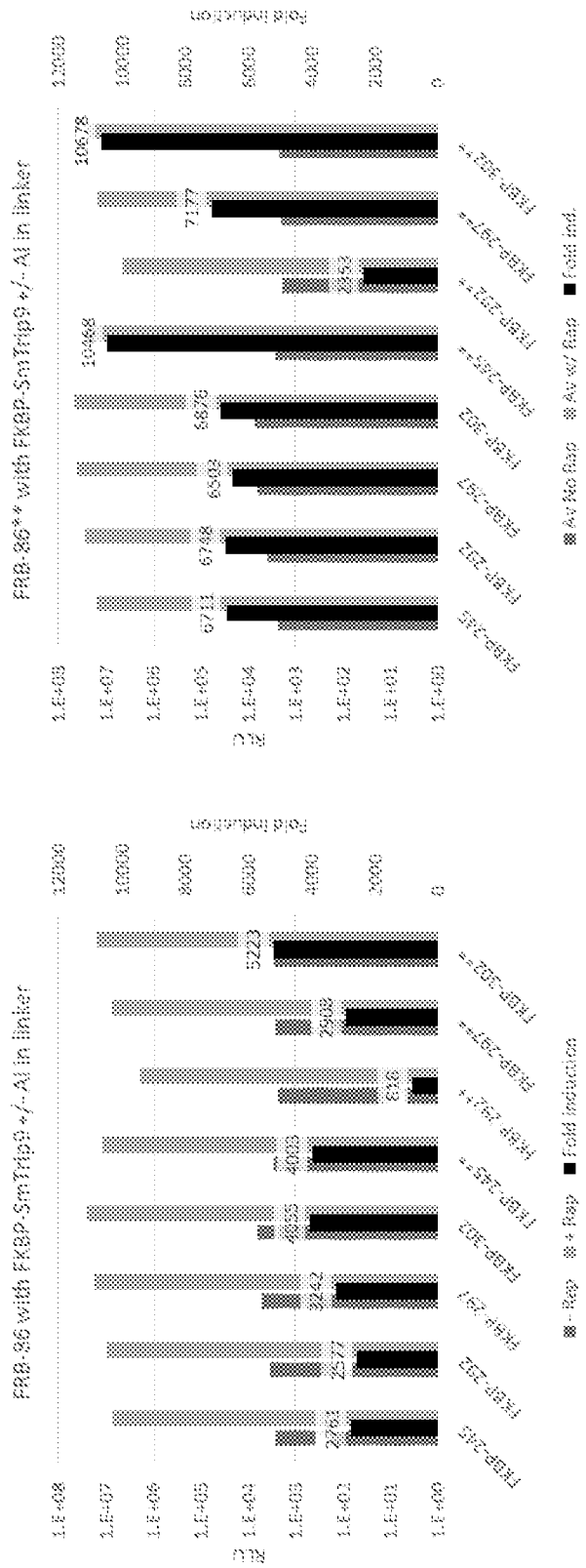
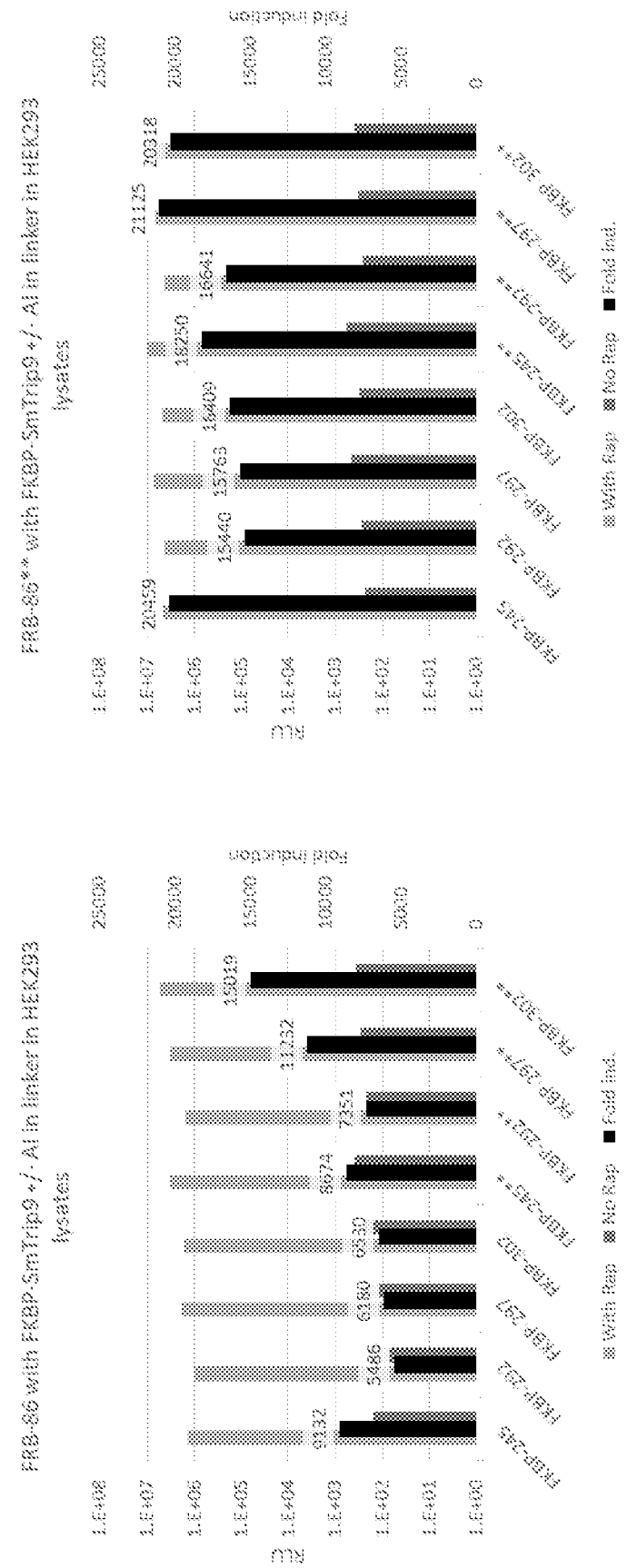

Read luminescence output
LgBiT
Coelenterazine or a Coelenterazine analog
Complementation of oligoprobes with target strand in purified or crude sample matrix
SmTrip and SmTrip10 oligoprobes
Target nucleic acid

FIG. 57



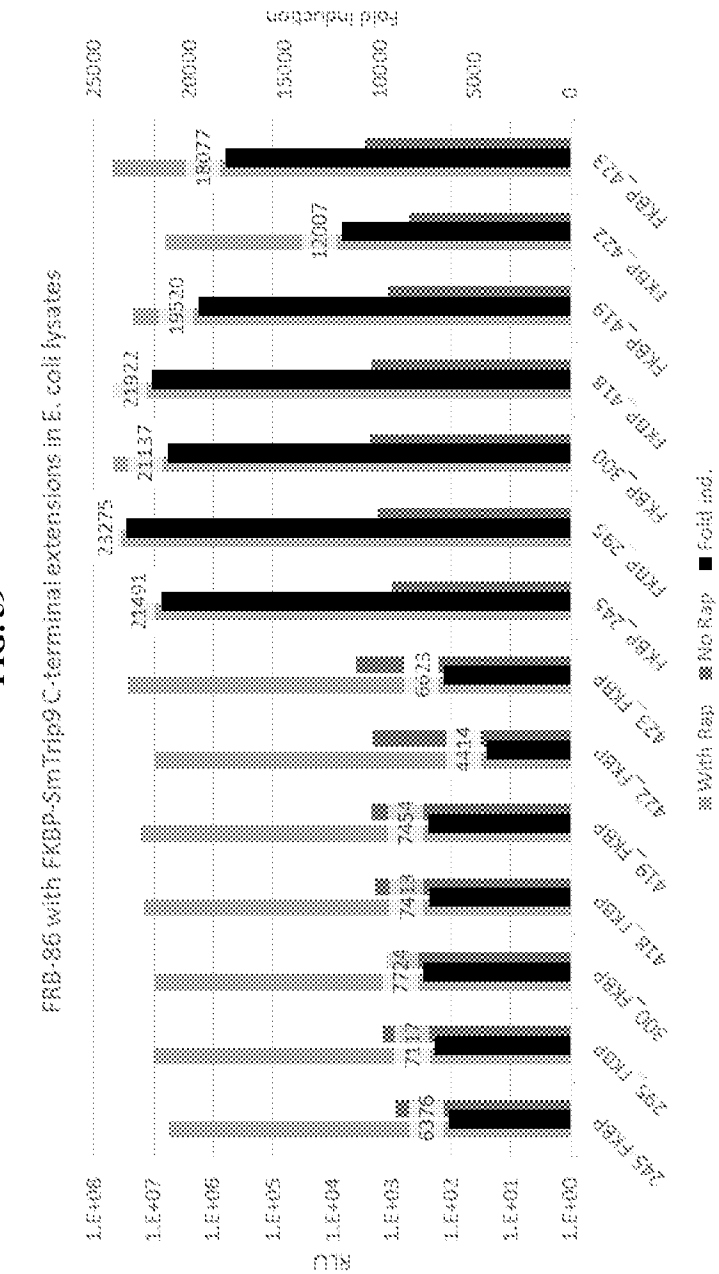
Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS

FIG. 58



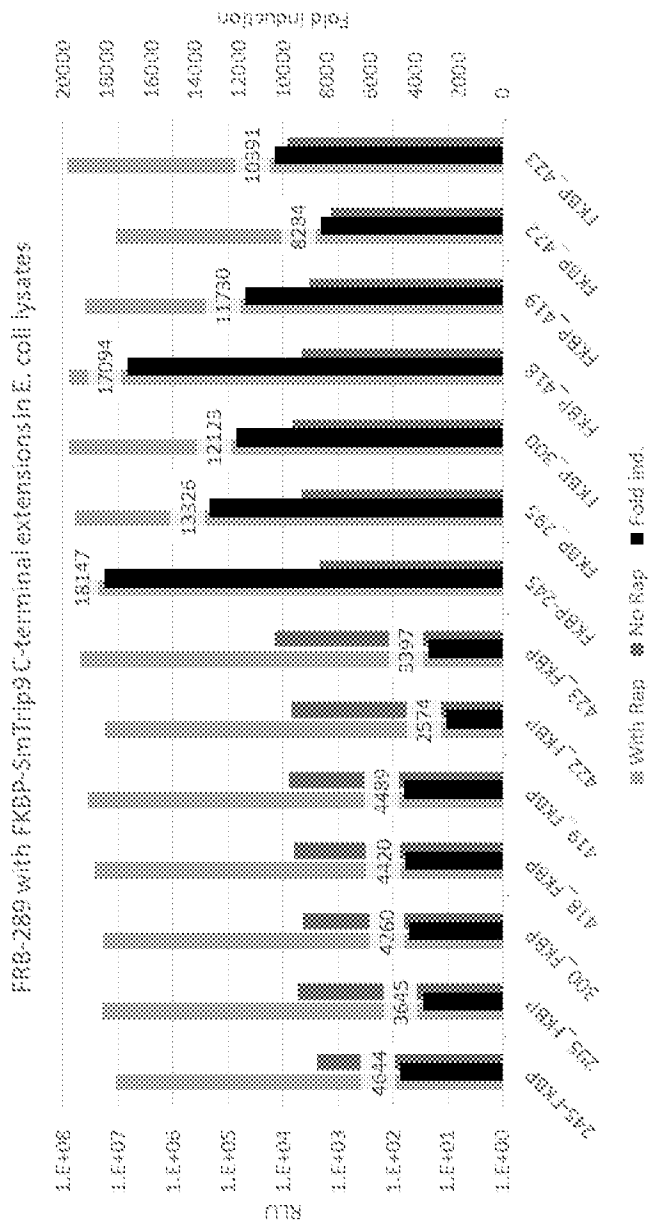
Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
292.x	289	MLFRVTINS
297.x	290	MLFRVTINSV
302.x	291	MLFRVTINSVS

FIG. 59



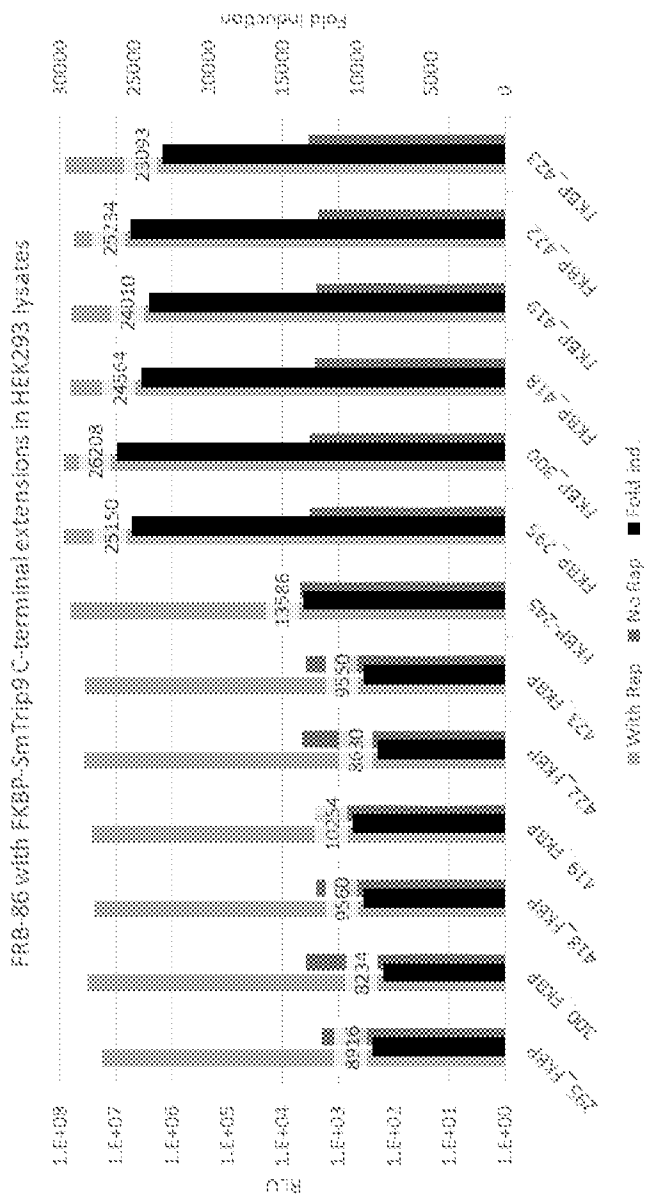
Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
295.x	220	GSMLFRVTINSV
300.x	221	GSMLFRVTINSVS
418.x	226	GSMLFRVTINSVSG
419.x	227	GSMLFRVTINSVSGW
422	228	GSMLFRVTINSVSGWR
423	229	GSMLFRVTINSVSGWK

FIG. 60



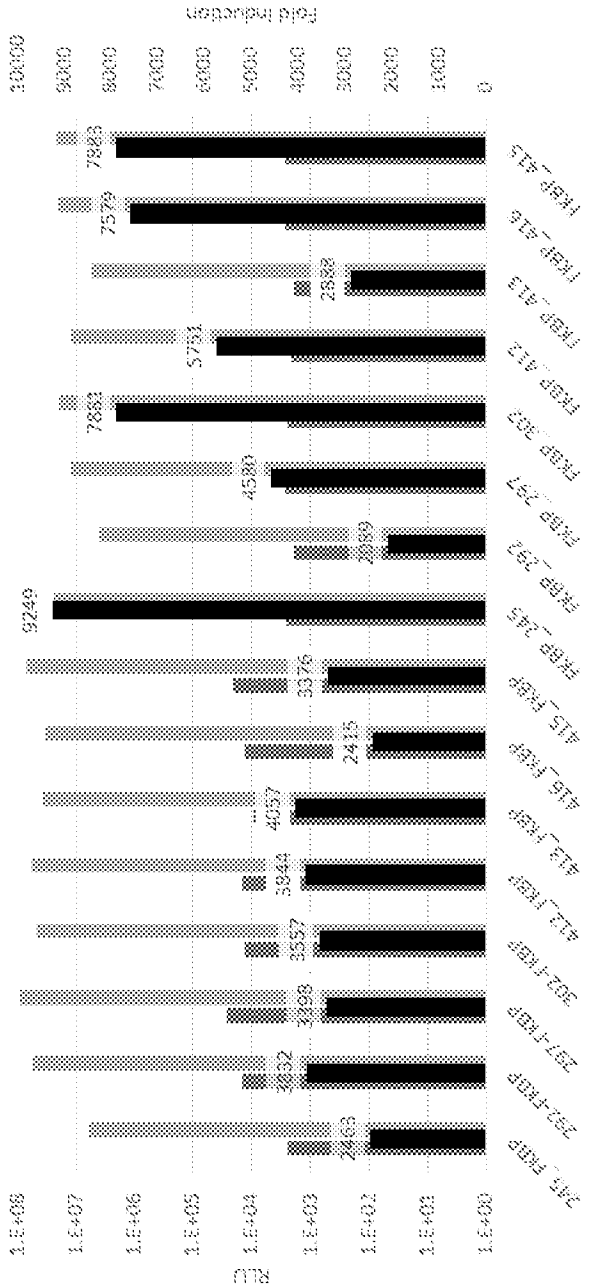
Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
295.x	220	GSMLFRVTINSV
300.x	221	GSMLFRVTINSVS
418.x	226	GSMLFRVTINSVSG
419.x	227	GSMLFRVTINSVSGW
422	228	GSMLFRVTINSVSGWR
423	229	GSMLFRVTINSVSGWK

FIG. 61



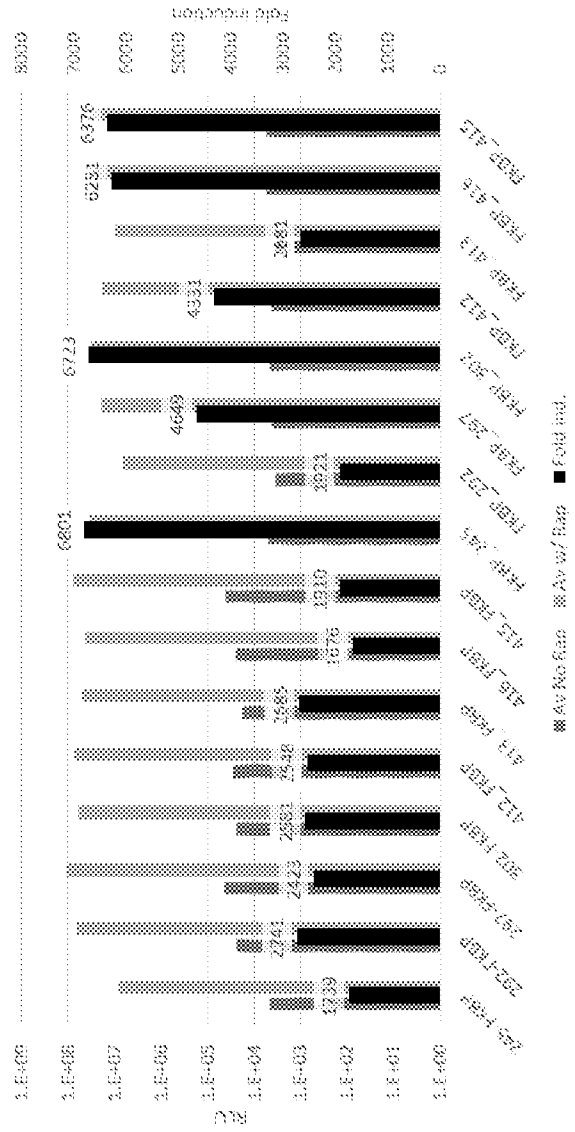
Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
295.x	220	GSMLFRVTINSV
300.x	221	GSMLFRVTINSVS
418.x	226	GSMLFRVTINSVSG
419.x	227	GSMLFRVTINSVSGW
422	228	GSMLFRVTINSVSGWR
423	229	GSMLFRVTINSVSGWK

FIG. 62



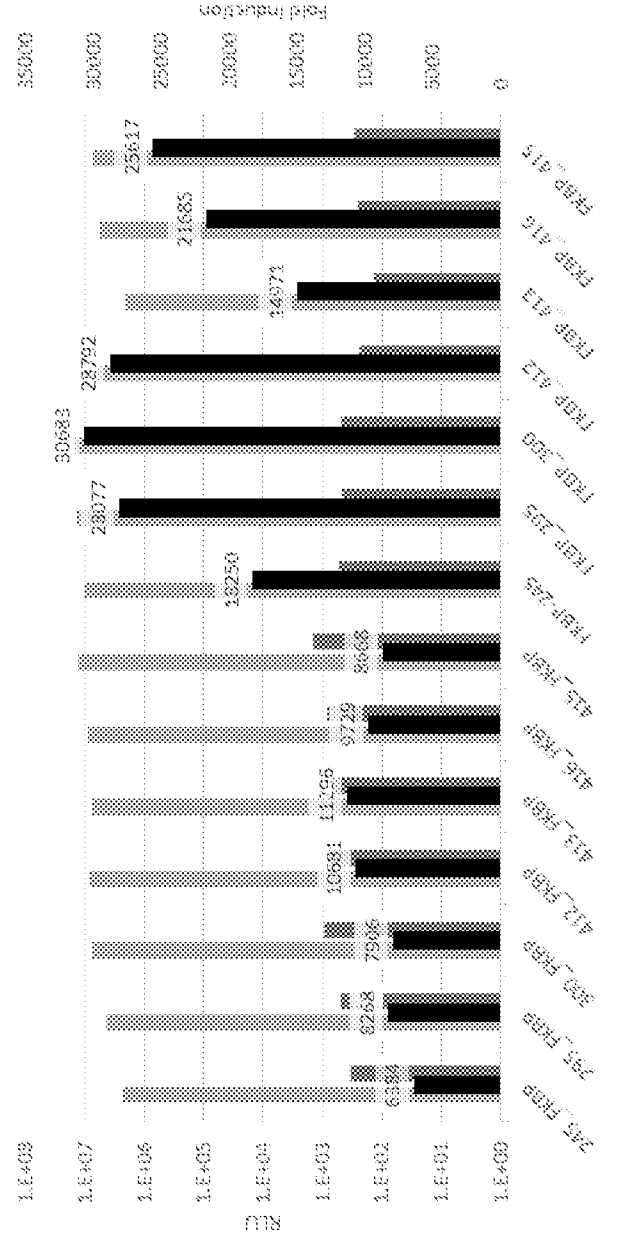
Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
295.x	220	GSMLFRVTINSV
300.x	221	GSMLFRVTINSVS
416	225	MLFRVTINSVSGWR
413.x	222	MLFRVTINSVSGW
412.x	222	MLFRVTINSVSG
415.x	224	MLFRVTINSVSGWK

FIG. 63



Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
295.x	220	GSMLFRVTINSV
300.x	221	GSMLFRVTINSVS
416	225	MLFRVTINSVSGWR
413.x	222	MLFRVTINSVSGW
412.x	222	MLFRVTINSVSG
415.x	224	MLFRVTINSVSGWK

FIG. 64



Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
295.x	220	GSMLFRVTINSV
300.x	221	GSMLFRVTINSVS
416	225	MLFRVTINSVSGWR
413.x	222	MLFRVTINSVSGW
412.x	222	MLFRVTINSVSG
415.x	224	MLFRVTINSVSGWK

FIG. 65

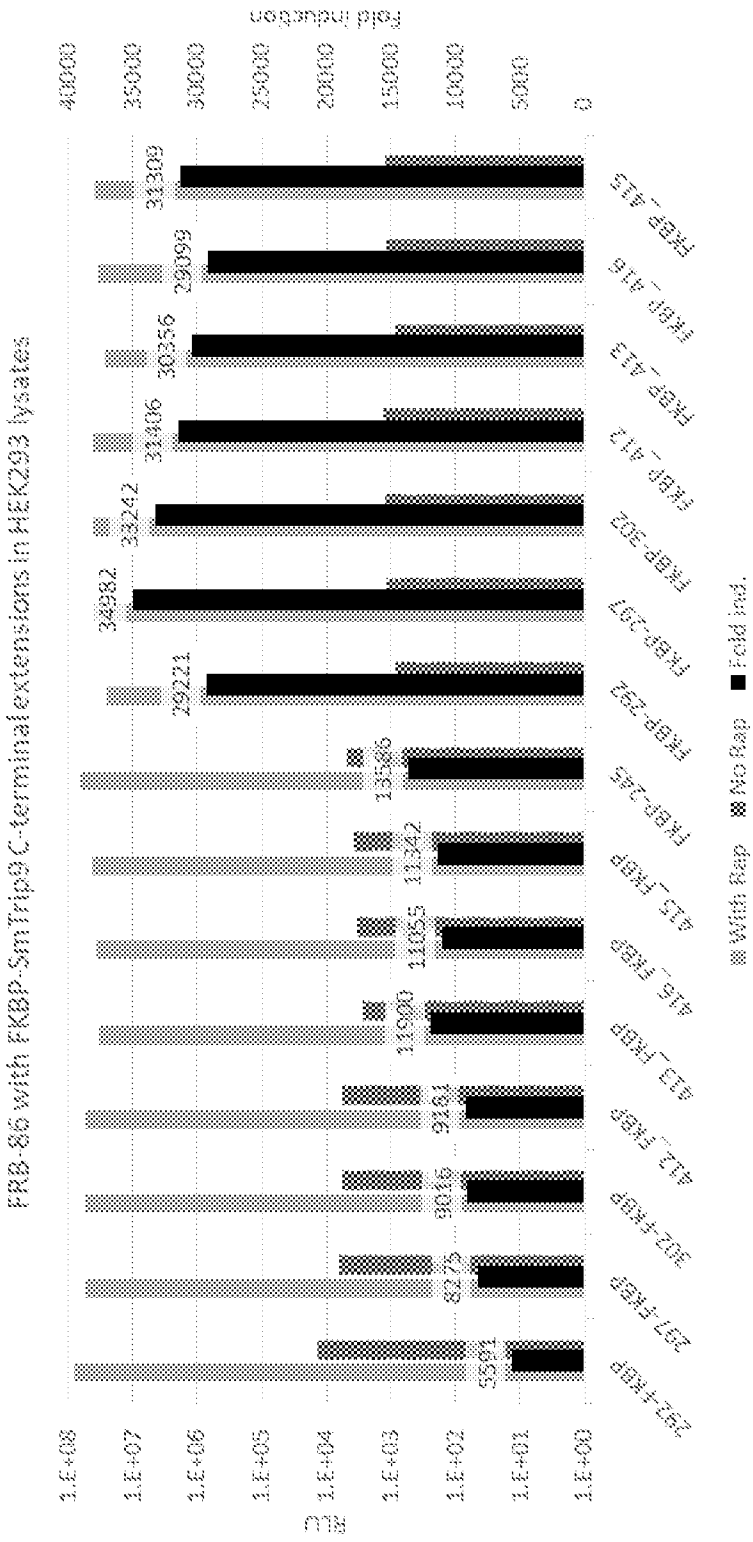
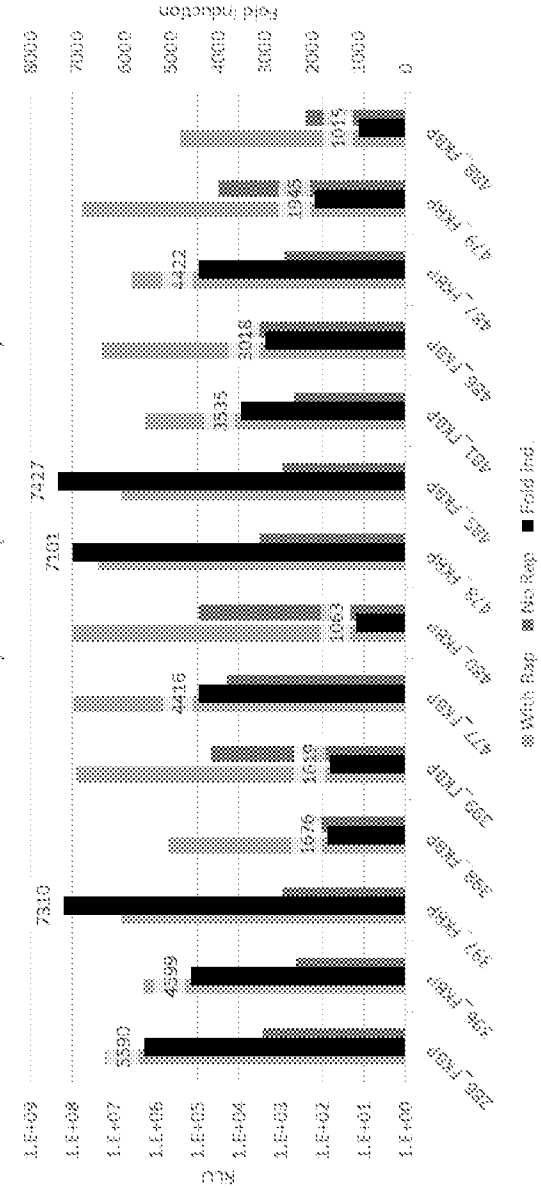
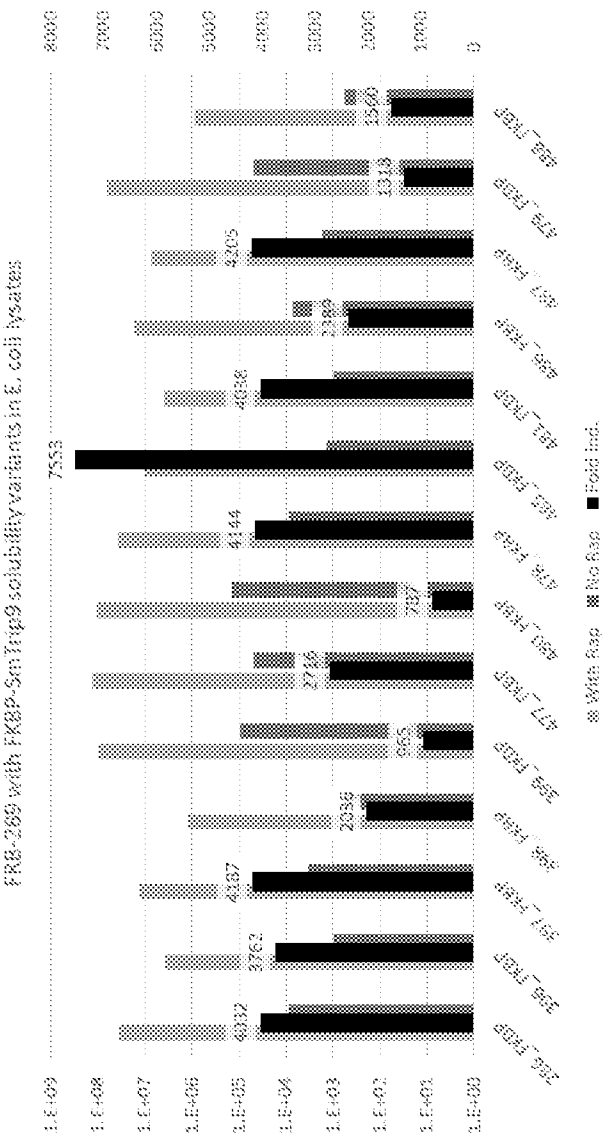

ERB-86 with FKBP-SmTrip9 C-terminal extensions in HEK293 lysates
RLU
Fold Induction
With Rap
No Rap
Fold Ind.
1.E+08
1.E+07
1.E+06
1.E+05
1.E+04
1.E+03
1.E+02
1.E+01
1.E+00
40000
35000
30000
25000
20000
15000
10000
5000
0
5561
8475
8046
9161
11000
11055
11242
13586
29221
34982
33242
31406
30356
29099
31309
292-FKBP
297-FKBP
302-FKBP
412_FKBP
413_FKBP
416_FKBP
415_FKBP
FKBP-292
FKBP-297
FKBP-302
FKBP_413
FKBP_412
FKBP_416
FKBP_415

FIG. 66



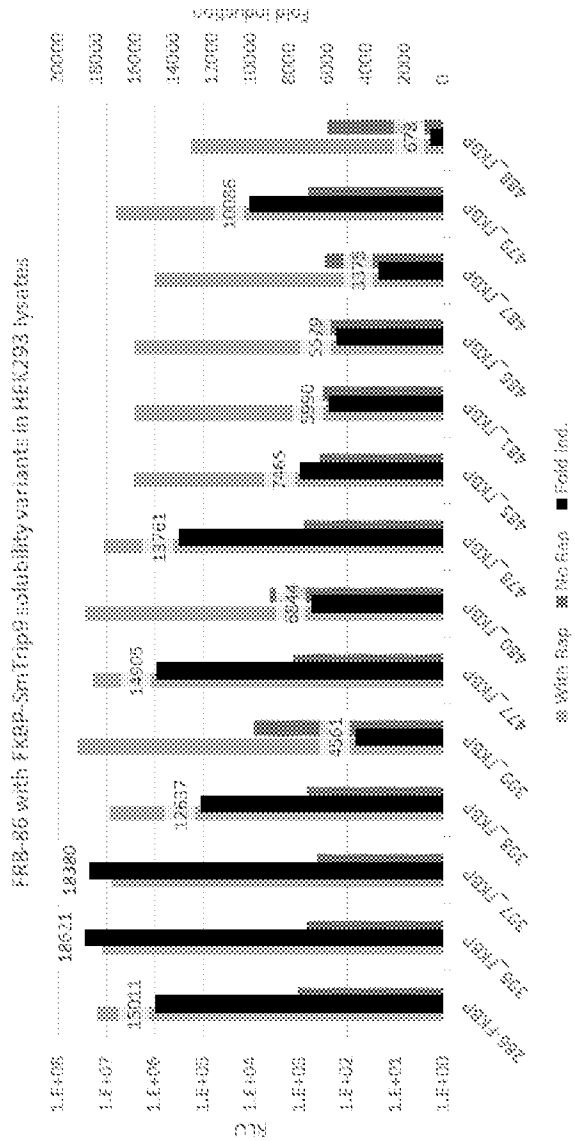
Pep ID	SEQ ID NO.	Sequence
286	37	SSWKRGSMLFRVTINS
396	186	SSRGSMLFRVTINSWK
397	187	SKRGSMLFRVTINSWS
398	188	SWRGSMLFRVTINS
399	297	SSKRGSMLFRVTIWS
477	232	MLFRVTINSWK
478	233	MLFRVTINSWS
479	234	MLFRVTIWS
480	235	MLFRVTIWK
481	236	MLFRVKINS
485	240	MLFRVNINS
486	241	MLFRVWINS
487	242	LLFRVKINS
488	243	FLFRVTINS

FIG. 67



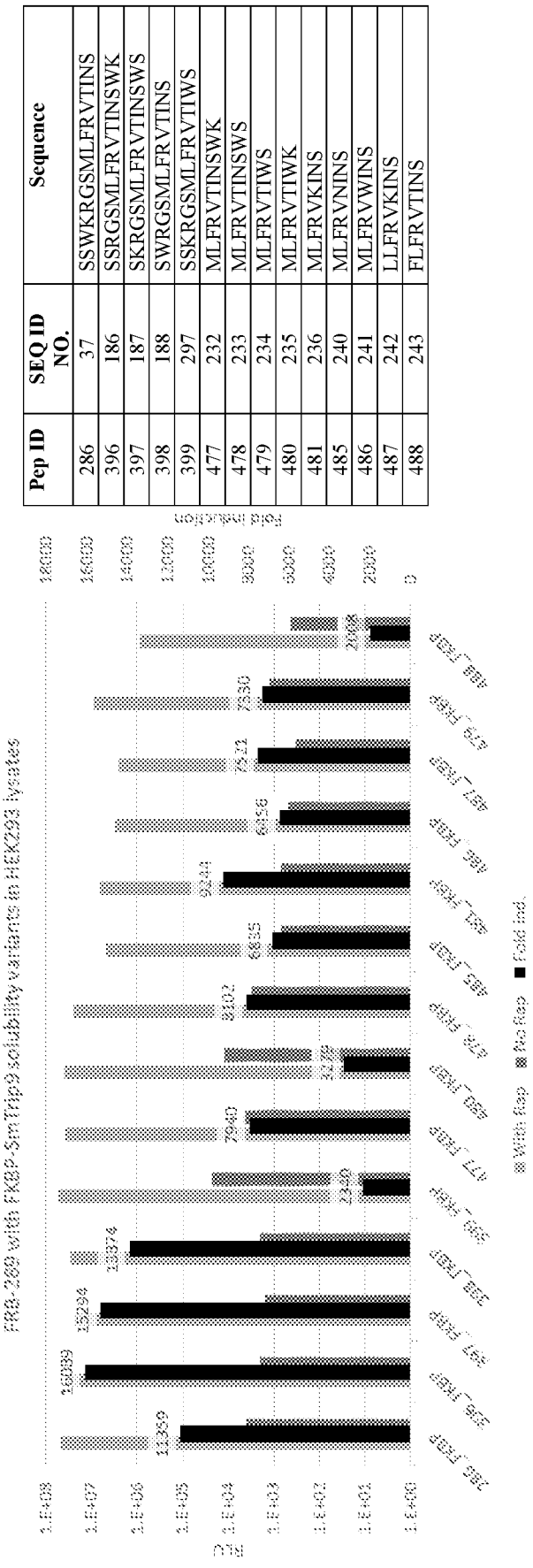
Pep ID	SEQ ID NO.	Sequence
286	37	SSWKRGSMLFRVTINS
396	186	SSRGSMLFRVTINSWK
397	187	SKRGSMLFRVTINSWS
398	188	SWRGSMLFRVTINS
399	297	SSKRGSMLFRVTIWS
477	232	MLFRVTINSWK
478	233	MLFRVTINSWS
479	234	MLFRVTIWS
480	235	MLFRVTIWK
481	236	MLFRVKINS
485	240	MLFRVNINS
486	241	MLFRVWINS
487	242	LLFRVKINS
488	243	FLFRVTINS

FIG. 68



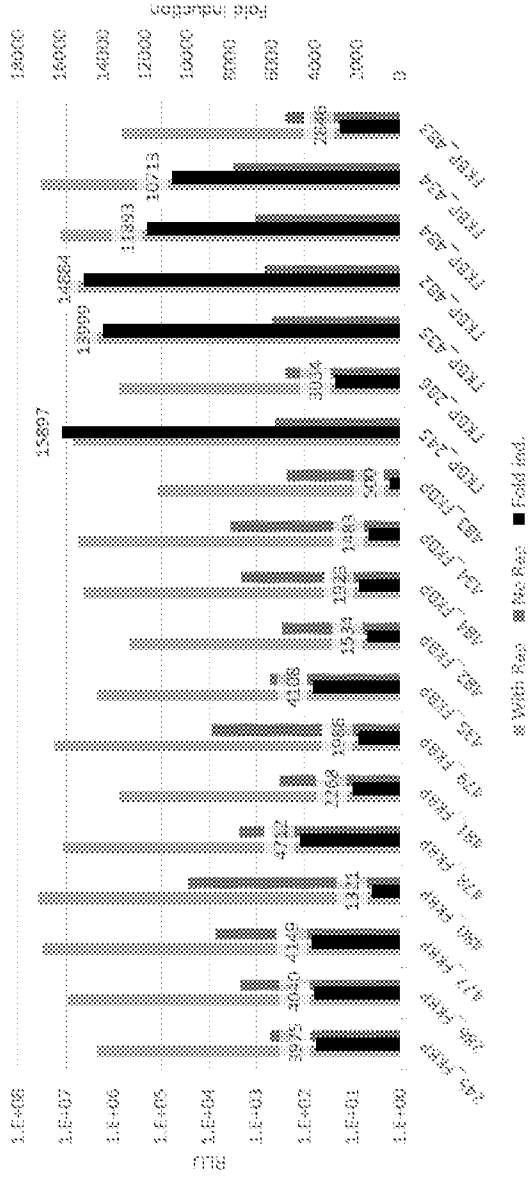
Pep ID	SEQ ID NO.	Sequence
286	37	SSWKRGSMLFRVTINS
396	186	SSRGSMLFRVTINSWK
397	187	SKRGSMLFRVTINSWS
398	188	SWRGSMLFRVTINS
399	297	SSKRGSMLFRVTIWS
477	232	MLFRVTINSWK
478	233	MLFRVTINSWS
479	234	MLFRVTIWS
480	235	MLFRVTIWK
481	236	MLFRVKINS
485	240	MLFRVNINS
486	241	MLFRVWINS
487	242	LLFRVKINS
488	243	FLFRVTINS

FIG. 69



Pep ID	SEQ ID NO.	Sequence
286	37	SSWKRGSMLFRVTINS
396	186	SSRGSMLFRVTINSWK
397	187	SKRGSMLFRVTINSWS
398	188	SWRGSMLFRVTINS
399	297	SSKRGSMLFRVTIWS
477	232	MLFRVTINSWK
478	233	MLFRVTINSWS
479	234	MLFRVTIWS
480	235	MLFRVTIWK
481	236	MLFRVKINS
485	240	MLFRVNINS
486	241	MLFRVWINS
487	242	LLFRVKINS
488	243	FLFRVTINS

FIG. 70



Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
286	37	SSWKRGSMLFRVTINS
434	187	SKRGSMLFRVTINSWS
435	188	SWRGSMLFRVTINS
477	232	MLFRVTINSWK
478	233	MLFRVTINSWS
479	234	MLFRVTIWS
480	235	MLFRVTIWK
481	236	MLFRVKINS
482	237	GSMLFRVTINSWS
483	238	GSMLFRVKINS
484	239	GSMLFRTIWS

Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
286	37	SSWKRGSMLFRVTINS
434	187	SKRGSMLFRVTINSWS
435	188	SWRGSMLFRVTINS
477	232	MLFRVTINSWK
478	233	MLFRVTINSWS
479	234	MLFRVTIWS
480	235	MLFRVTIWK
481	236	MLFRVKINS
482	237	GSMLFRVTINSWS
483	238	GSMLFRVKINS
484	239	GSMLFRTIWS

FIG. 71

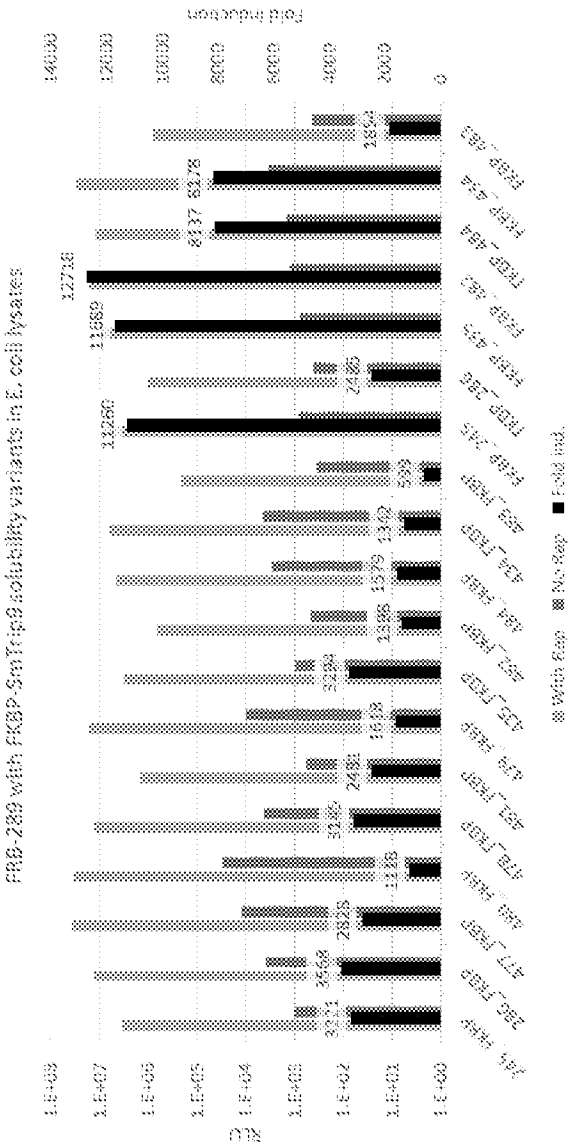
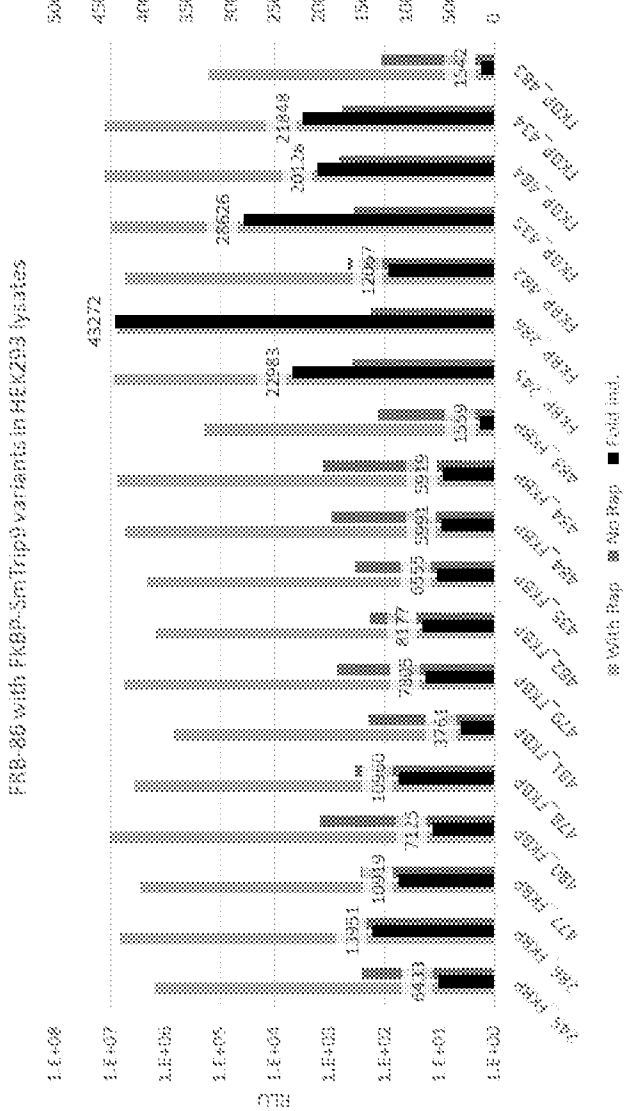


FIG. 72



Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
286	37	SSWKRGSMLFRVTINS
434	187	SKRGSMLFRVTINSWS
435	188	SWRGSMLFRVTINS
477	232	MLFRVTINSWK
478	233	MLFRVTINSWS
479	234	MLFRVTIWS
480	235	MLFRVTIWK
481	236	MLFRVKINS
482	237	GSMLFRVTINSWS
483	238	GSMLFRVKINS
484	239	GSMLFRTIWS

Pep ID	SEQ ID NO.	Sequence
245	23	GSMLFRVTINS
286	37	SSWKRGSMLFRVTINS
434	187	SKRGSMLFRVTINSWS
435	188	SWRGSMLFRVTINS
477	232	MLFRVTINSWK
478	233	MLFRVTINSWS
479	234	MLFRVTIWS
480	235	MLFRVTIWK
481	236	MLFRVKINS
482	237	GSMLFRVTINSWS
483	238	GSMLFRVKINS
484	239	GSMLFRTIWS

FIG. 73

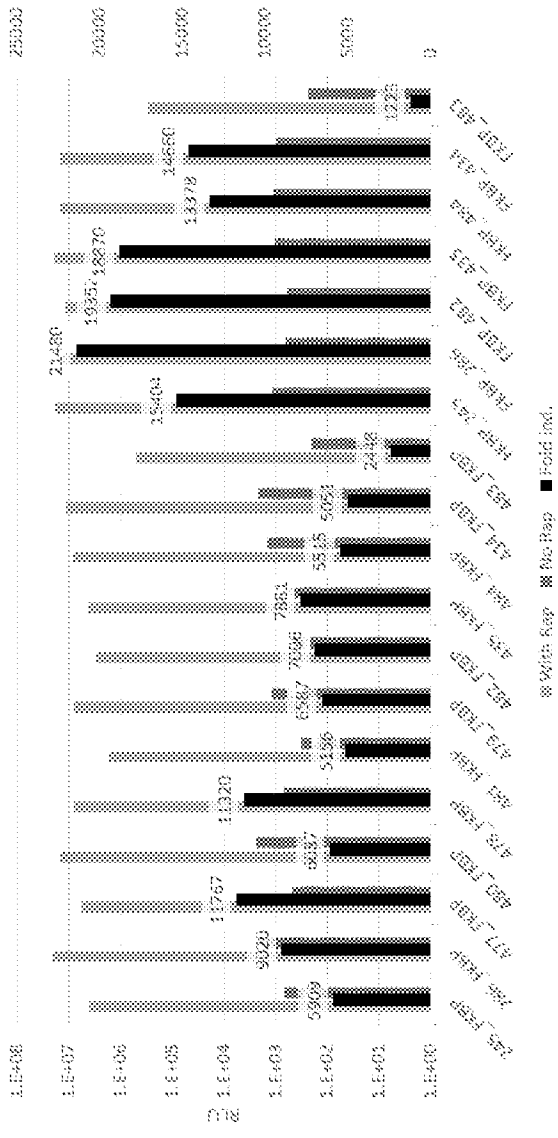


FIG. 74

Biochemical analysis of synthetic SmTrip9 variants				
Pep ID	**SEQ ID NO.:**	**Sequence**	**SmTrip9 Kd in uM (Bmax in RLU)**	**HiBiT Kd in uM (Bmax in RLU)**
286	37	SSWKRGSMLFRVTINS	0.10 (2.47E7)	4.05 (2.76E7)
416	225	MLFRVTINSVSGWR	3.54 (9.38E6)	29.88 (1.13E7)
417	224	MLFRVTINSVSGWK	1.15 (1.30E7)	19.56 (1.68E7)
422	228	GSMLFRVTINSVSGWR	0.24 (2.05E7)	4.91 (2.48E7)
423	229	GSMLFRVTINSVSGWK	0.18 (2.09E7)	3.48 (2.52E7)
434	230	GSMLFRVTIWK	0.03 (2.35E7)	0.52 (2.38E7)
435	231	GSMLFRVTINSWK	0.28 (2.27E7)	6.11 (2.34E7)

FIG. 75

Biochemical analysis of synthetic SmTrip9 Variants				
Pep ID	**SEQ ID NO.:**	**Sequence**	**SmTrip9 (uM)**	**Bmax (RLU)**
286	37	SSWKRGSMLFRVTINS	0.10	2.47E7
412	246	SSWKRMLFRVTINSVSG	0.11	2.24E7
413	247	SSWKRMLFRVTINSVSGW	0.15	2.20E7
414	248	SSWKRMLFRVTINSVSGWR	0.16	2.27E7
415	249	SSWKRMLFRVTINSVSGWK	0.18	2.34E7
418	226	GSMLFRVTINSVSG	0.31	2.20E7
419	227	GSMLFRVTINSVSGW	0.33	2.03E7
420	251	SSWKRGSMLFRVTINSVSG	0.29	2.27E7
421	252	SSWKRGSMLFRVTINSVSGW	0.26	2.38E7
424	253	SSWKRGSMLFRVTINSVSGWR	3.33	4.03E6
425	254	SSWKRGSMLFRVTINSVSGWK	0.85	2.19E7
426	255	SSWKRGSYLFRVTINS	0.29	1.76E7
427	256	SSWKRGSMLFRVKINS	0.23	2.46E7
428	257	SSWKRGSMLFRVRINS	0.04	2.38E7
430	261	SSWRGSMLFRVTIKS	0.28	2.35E7
431	262	KRSSGSMLFRVTIWS	0.03	2.38E7
432	263	SSKRMLFRVTIWS	0.01	2.39E7
433	264	KRSSMLFRVTIWS	0.02	2.56E7
422	228	GSMLFRVTINSVSGWR	0.24	2.05E7
423	229	GSMLFRVTINSVSGWK	0.18	2.09E7
434	230	GSMLFRVTIWK	0.03	2.35E7
435	231	GSMLFRVTINSWK	0.28	2.27E7
492	313	GSMLFRVTINKWK	0.71	1.89E07
493	314	GSMLFRVTIKSWK	0.47	2.19E07
494	315	GSMLFRVTINRWK	0.86	1.93E07
495	316	GSMLFRVTIRSWK	0.48	1.98E07
496	317	GSMLFRVTINDWK	2.17	9.85E06
497	318	GSMLFRVTIDSWK	5.25	1.76E07
498	319	GSMLFRVTINEWK	1.02	1.22E07
499	320	GSMLFRVTIESWK	4.25	1.71E07

FIG. 76

Pep ID	SEQ ID NO.:	Sequence	SmTrip9 (uM)	Bmax (RLU)
286	37	*SSWKRGSMLFRVTINS	0.10	2.75E7
445	265	*GSMKFRVTINSWK	1.02	1.84E7
450	266	*GSMLFRKTINSWK	1.14	4.92E4
455	267	*GSMLFRVTKNSWK	0.55	3.57E4
332	436	*GSMLFRVTINS	0.24	2.48E7
333	437	*GSMLFRVSINS	0.33	2.32E7
334	438	*GSMLFRVNINS	5.67	2.46E7
335	439	*GSKLFRVTINS	9.86	3.92E5
336	440	*GSRLFRVTINS	0.01	1.39E4
337	441	*GSMWFRVTINS	6.15	9.89E6
338	442	*GSMSFRVTINS	7.03	1.79E7
339	443	*GSMNFRVTINS	8.97	5.11E6
340	444	*GSMKFRVTINS	0.73	1.69E7
341	446	*GSMLFRWTINS	0.31	7.35E4
342	447	*GSMLFRSTINS	4.84	6.92E5
343	448	*GSMLFRNTINS	3.96	1.47E5
344	449	*GSMLFRKTINS	0.12	2.20E4
345	451	*GSMLFRVTWNS	2.24	5.51E5
346	452	*GSMLFRVTSNS	9.99	1.32E6
347	453	*GSMLFRVTNNS	14.58	2.51E6
348	454	*GSMLFRVTKNS	0.21	2.60E4
349	456	*GSMLFRVTIKS	0.07	2.56E7
310	489	*GSMLFRVTINSWK	0.51	1.38E07
311	490	GSMLFRVTINSWK*	12.49	8.55E06
312	491	*GSMLFRVTINSWK*	0.26	1.33E07
321	465	*GSMRFRVTINSWK*	10.57	1.47E7
322	466	*GSMDFRVTINSWK*	0.66	4.43E4
323	467	*GSMEFRVTINSWK*	NA	NA
324	468	*GSMLFRRTINSWK*	NA	1.27E4
325	469	*GSMLFRDTINSWK*	NA	1.29E4
326	470	*GSMLFRETINSWK*	NA	1.26E4
327	472	*GSMLFRVTDNSWK*	0.09	2.60E4
328	473	*GSMLFRVTENSWK*	0.02	1.70E4
329	474	*GSMKFRVTINSWK*	0.24	1.35E7
330	475	*GSMLFRKTINSWK*	1.97	1.64E5
331	476	*GSMLFRVTKNSWK*	0.97	5.63E4
*Terminus not blocked, NA= not applicable				

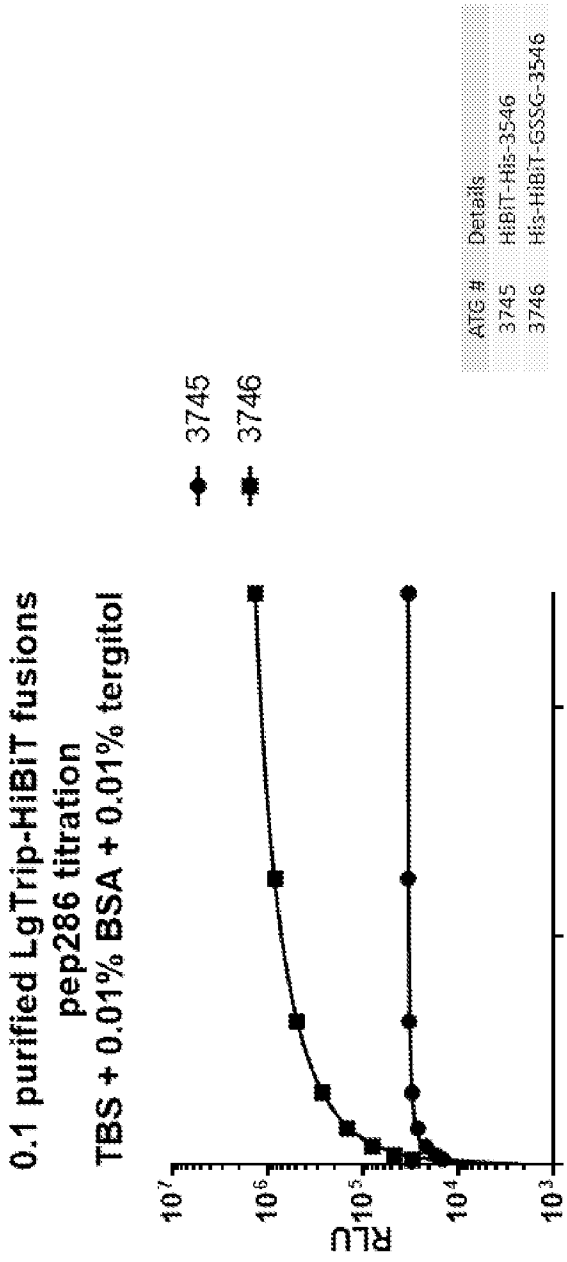
FIG. 77A

Solubility of Synthetic SmTrip9 peptides			
Pep ID	**SEQ ID NO.:**	**Sequence**	**Solubility#**
286	37	SSWKRGSMLFRVTINS	Y
292	153	SSWKRMLFRVTINS	N
297	157	SSWKRMLFRVTINSV	Y
302	161	SSWKRMLFRVTINSVS	Y
305	164	SSWKRGSMLFRVTIN	Y
306	165	SSWKRGSMLFRVTI	Y
307	166	SSWKRSMLFRVTIN	Y
308	167	SSWKRMLFRVTIN	N
312	171	SSWKRMLFRVTI	Y
396	186	SSRGSMLFRVTINSWK	Y
397	187	SKRGSMLFRVTINSWS	Y
398	188	SWRGSMLFRVTINS	Y
400	189	SSRGSMLFRVTIWK	Y
401	190	SSWKRGSMLYRVTINS	Y
402	191	SSWKRGSMLWRVTINS	Y
403	192	SSWKRGSMLHRVTINS	Y
404	193	SSWKRGSLLFRVTINS	Y
405	194	SSWKRGSKLFRVTINS	Y
406	195	SSWKRGSRLFRVTINS	Y
407	196	SSWKRGSFLFRVTINS	Y
408	197	SSWKRGSWLFRVTINS	Y
409	198	SSWKRGSMLFRVSINS	Y
410	199	SSWKRGSMLFRVQINS	Y
411	200	SSWKRGSMLFRVNINS	Y
412	246	SSWKRMLFRVTINSVSG	Y
413	247	SSWKRMLFRVTINSVSGW	N
414	248	SSWKRMLFRVTINSVSGWR	Y
415	249	SSWKRMLFRVTINSVSGWK	Y
416	225	MLFRVTINSVSGWR	N
417	250	MLFRVTINSVSGWK	Y
418	251	SSWKRGSMLFRVTINSVSG	N
419	252	SSWKRGSMLFRVTINSVSGW	Y
420	253	SSWKRGSMLFRVTINSVSGWR	Y
421	254	SSWKRGSMLFRVTINSVSGWK	Y
422	228	GSMLFRVTINSVSGWR	Y
423	229	GSMLFRVTINSVSGWK	Y
424	255	SSWKRGSYLFRVTINS	Y
425	256	SSWKRGSMLFRVKINS	Y
426	257	SSWKRGSMLFRVRINS	Y
427	258	SSWKRGSMLFRVWINS	Y
428	259	SSKRGSMLFRVTIWSV	Y
430	261	SSWRGSMLFRVTIKS	Y
431	262	KRSSGSMLFRVTIWS	Y
432	263	SSKRMLFRVTIWS	Y
433	264	KRSSMLFRVTIWS	Y
434	230	GSMLFRVTIWK	Y
435	231	GSMLFRVTINSWK	N
445	265	GSMKFRVTINSWK	Y
450	266	GSMLFRKTINSWK	Y
455	267	GSMLFRVTKNSWK	Y
#in water at ~1mM after multiple freeze/thaws *terminus unblocked			

FIG. 77B

Solubility of Synthetic SmTrip9 peptides			
Pep ID	**SEQ ID NO.:**	**Sequence**	**Solubility#**
521	268	GKMLFRVTINSWK	Y
522	269	GKMLFRVTIWK	Y
523	270	GSMKFRVTINSWK	Y
524	271	GSMKFRVTIWK	Y
525	272	GRMLFRVTINSWK	N
526	273	GRMLFRVTIWK	Y
527	274	GSMRFRVTINSWK	N
528	275	GSMRFRVTIWK	Y
529	276	GDMLFRVTINSWK	Y
530	277	GDMLFRVTIWK	Y
531	278	GSMDFRVTINSWK	Y
532	279	GSMDFRVTIWK	Y
533	280	GEMLFRVTINSWK	Y
535	281	GSMEFRVTINSWK	Y
536	282	GSMEFRVTIWK	Y
538	283	GSMLFRVTIWKVK	Y
539	284	GSMLFRVTIWSVK	Y
540	285	GSMLFRVTIWSK	Y
541	286	GSMLFRVTIWKWK	N
542	287	GSMLFRVTIWKK	Y
436	332	*GSMLFRVTINS	Y
437	333	*GSMLFRVSINS	Y
438	334	*GSMLFRVNINS	Y
439	335	*GSKLFRVTINS	Y
440	336	*GSRLFRVTINS	Y
441	337	*GSMWFRVTINS	N
442	338	*GSMSFRVTINS	Y
443	339	*GSMNFRVTINS	Y
444	340	*GSMKFRVTINS	Y
446	341	*GSMLFRWTINS	Y
447	342	*GSMLFRSTINS	Y
448	343	*GSMLFRNTINS	Y
449	344	*GSMLFRKTINS	Y
451	345	*GSMLFRVTWNS	Y
452	346	*GSMLFRVTSNS	Y
453	347	*GSMLFRVTNNS	Y
454	348	*GSMLFRVTKNS	Y
456	349	*GSMLFRVTIKS	Y
489	310	*GSMLFRVTINSWK	Y
490	311	GSMLFRVTINSWK*	N
491	312	*GSMLFRVTINSWK*	Y
465	465	*GSMRFRVTINSWK*	Y
466	466	*GSMDFRVTINSWK*	Y
467	467	*GSMEFRVTINSWK*	Y
468	468	*GSMLFRRTINSWK*	Y
469	469	*GSMLFRDTINSWK*	N
470	470	*GSMLFRETINSWK*	N
472	472	*GSMLFRVTDNSWK*	Y
473	473	*GSMLFRVTENSWK*	Y
474	474	*GSMKFRVTINSWK*	Y
475	475	*GSMLFRKTINSWK*	Y
476	476	*GSMLFRVTKNSWK*	Y
#in water at ~1mM after multiple freeze/thaws *terminus unblocked			

FIG. 78A



	3745	3746
Bmax	33862	3282750
Kd	0.07149	7.127

FIG. 78B

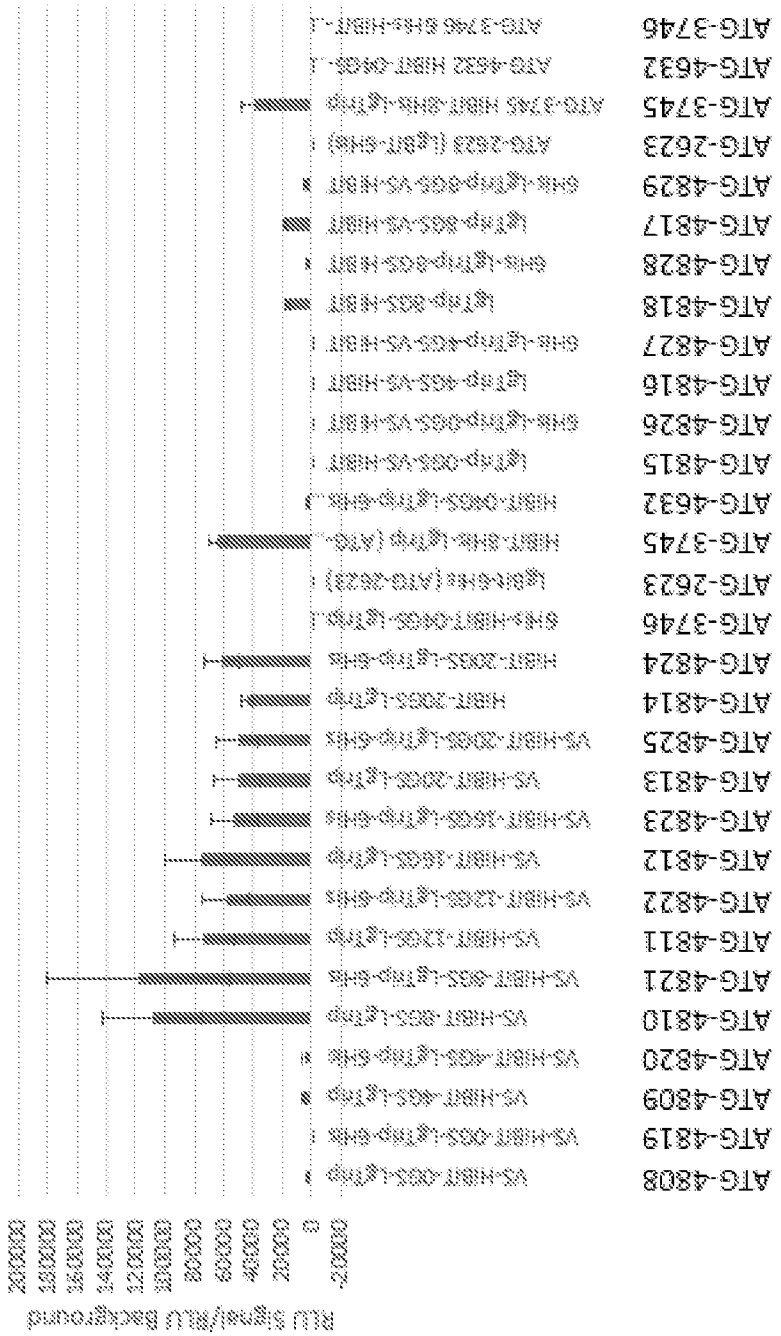

Signal Over Background
RLU Signal/RLU Background
0
20000
40000
60000
80000
100000
120000
140000
160000
180000
200000
ATG-4808 VS-HBST-ΔDS-LgTnp
ATG-4819 VS-HBST-ΔDS-LgTnp-6His
ATG-4809 VS-HBST-ΔGS-LgTnp
ATG-4820 VS-HBST-ΔGS-LgTnp-6His
ATG-4810 VS-HBST-ΔSS-LgTnp
ATG-4821 VS-HBST-ΔSS-LgTnp-6His
ATG-4811 VS-HBST-12GS-LgTnp
ATG-4822 VS-HBST-12GS-LgTnp-6His
ATG-4812 VS-HBST-16GS-LgTnp
ATG-4823 VS-HBST-16GS-LgTnp-6His
ATG-4813 VS-HBST-20GS-LgTnp
ATG-4825 VS-HBST-20GS-LgTnp-6His
ATG-4814 HBST-20GS-LgTnp
ATG-4824 HBST-20GS-LgTnp-6His
ATG-3746 6His-HBST-OMGS-LgTnp
ATG-2623 LgBirA-6His (ATG-2623)
ATG-3745 HBST-SHis-LgTnp (ATG-...
ATG-4632 HBST-OMGS-LgTnp-6His
ATG-4815 LgTnp-QDS-VS-HBST
ATG-4826 6His-LgTnp-QDS-VS-HBST
ATG-4816 LgTnp-4GS-VS-HBST
ATG-4827 6His-LgTnp-4GS-VS-HBST
ATG-4818 LgTnp-8GS-HBST
ATG-4828 6His-LgTnp-8GS-HBST
ATG-4817 LgTnp-8GS-VS-HBST
ATG-4829 6His-LgTnp-8GS-VS-HBST
ATG-2623 ATG-2623 (LgBT-6His)
ATG-3745 ATG-3745 HBST-SHis-LgTnp
ATG-4632 ATG-4632 HBST-OMGS-...
ATG-3746 ATG-3746 6His-HBST-...

FIG. 79

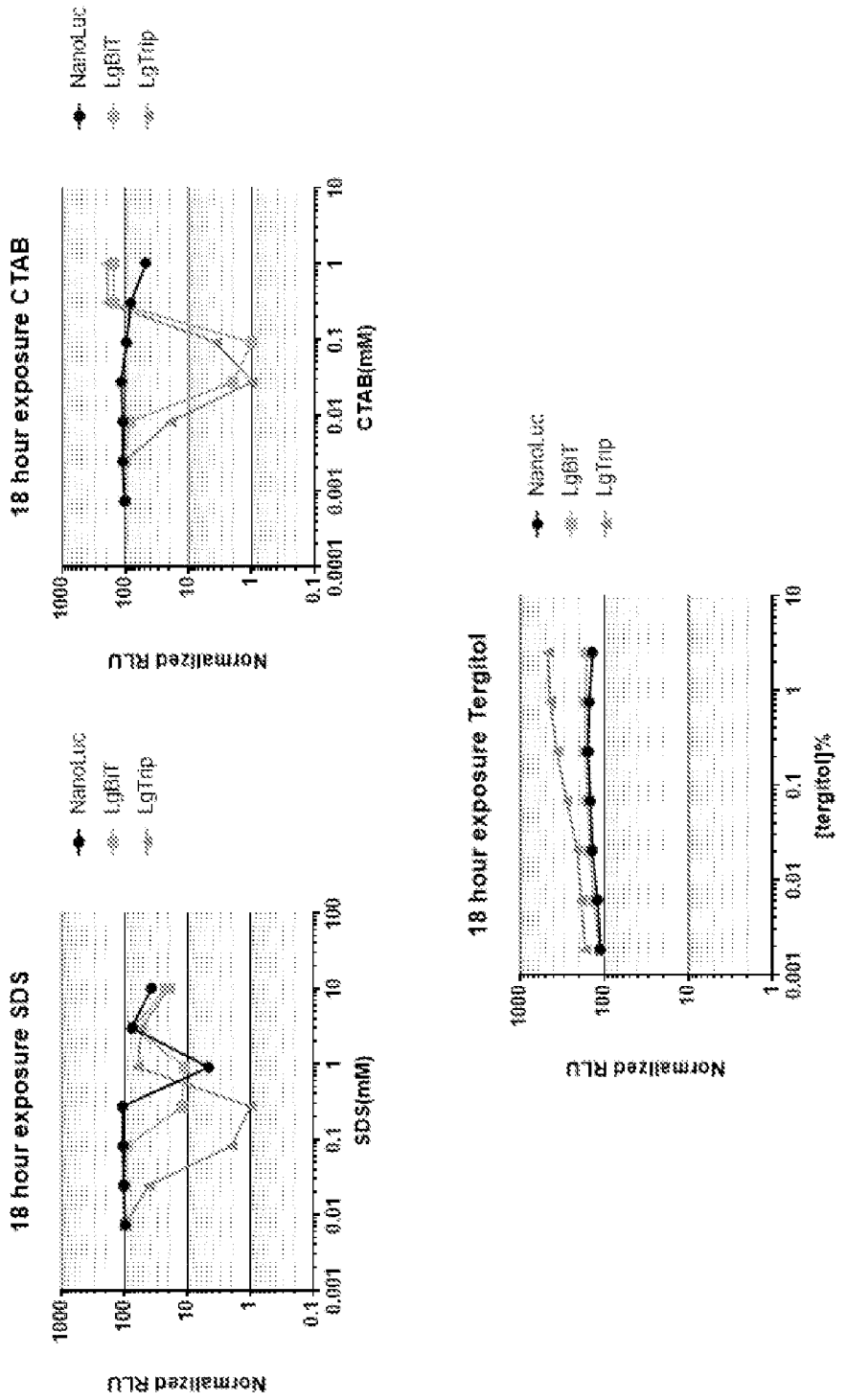

18 hour exposure SDS
NanoLuc
LgBiT
LgTrip
Normalized RLU
SDS(mM)
18 hour exposure CTAB
CTAB(mM)
18 hour exposure Tergitol
[Tergitol]%

FIG. 80

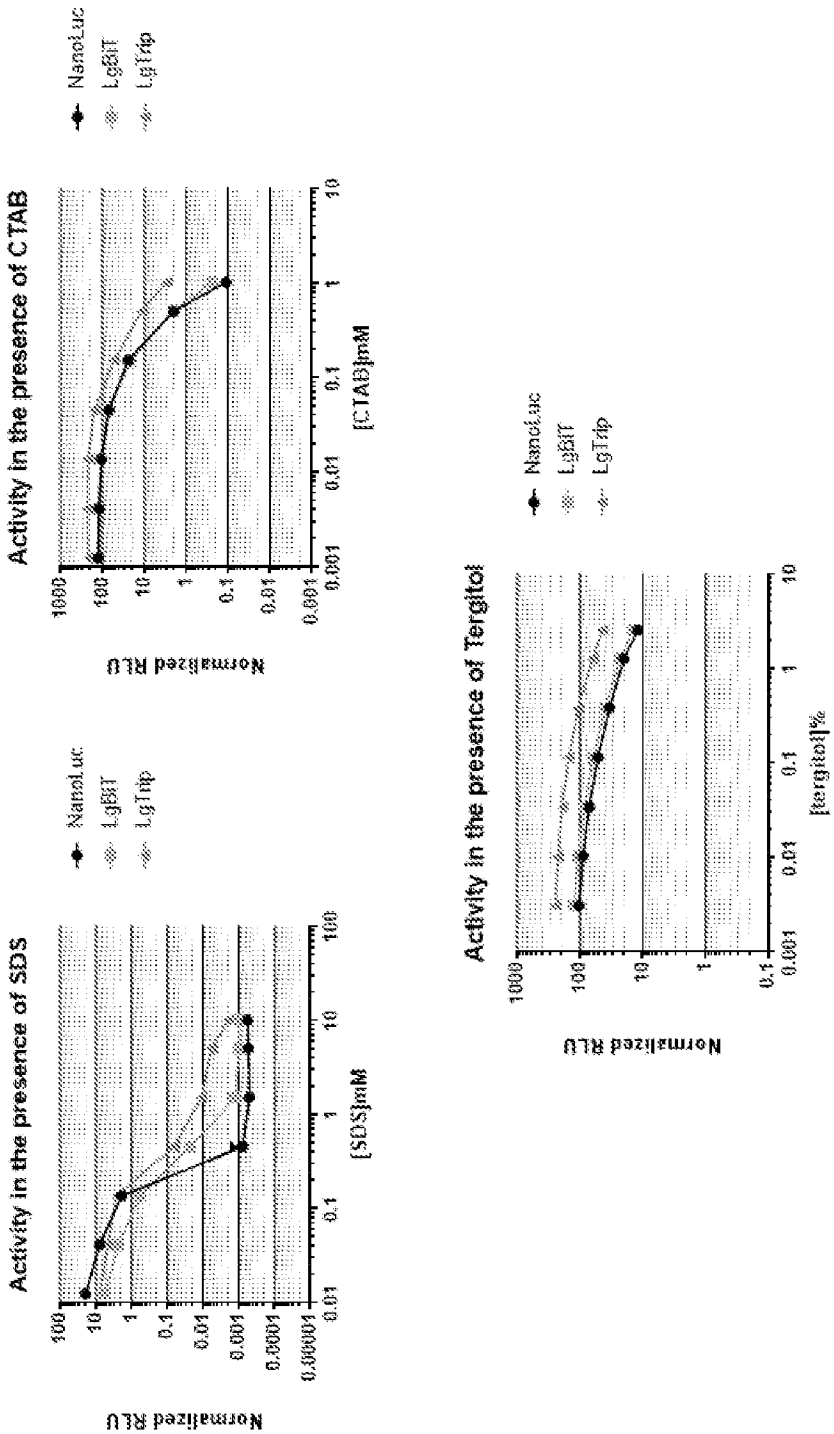

Activity in the presence of SDS
NanoLuc
LgBiT
LgTrip
Normalized RLU
[SDS]mM
Activity in the presence of CTAB
NanoLuc
LgBiT
LgTrip
Normalized RLU
[CTAB]mM
Activity in the presence of Tergitol
NanoLuc
LgBiT
LgTrip
Normalized RLU
[Tergitol]%

FIG. 81

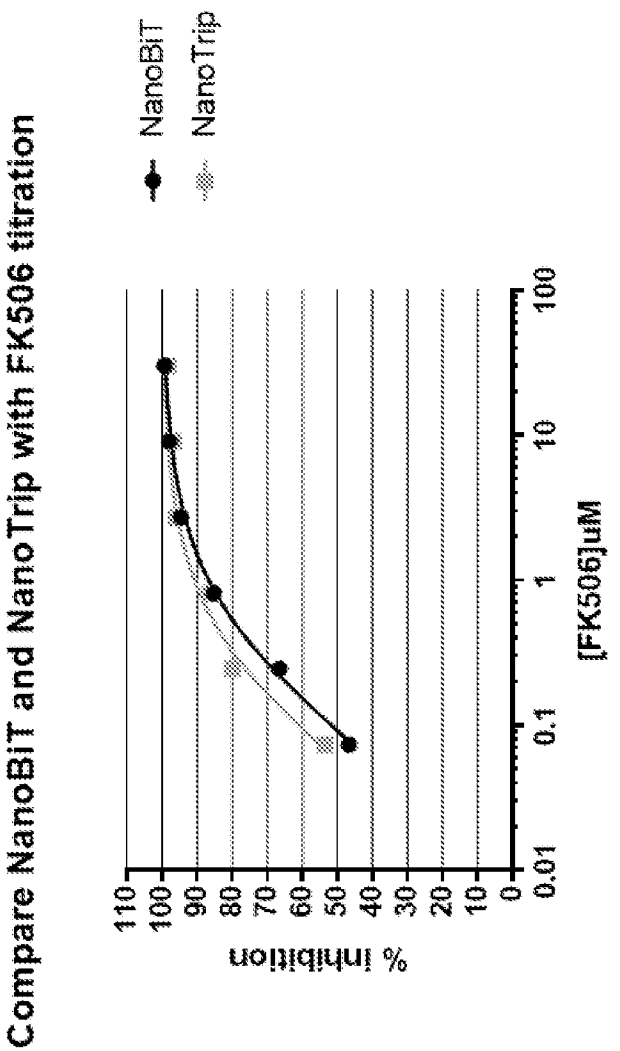

Compare NanoBiT and NanoTrip with FK506 titration
NanoBiT
NanoTrip
% inhibition
110
100
90
80
70
60
50
40
30
20
10
0
0.01
0.1
1
10
100
[FK506]uM

FIG. 82

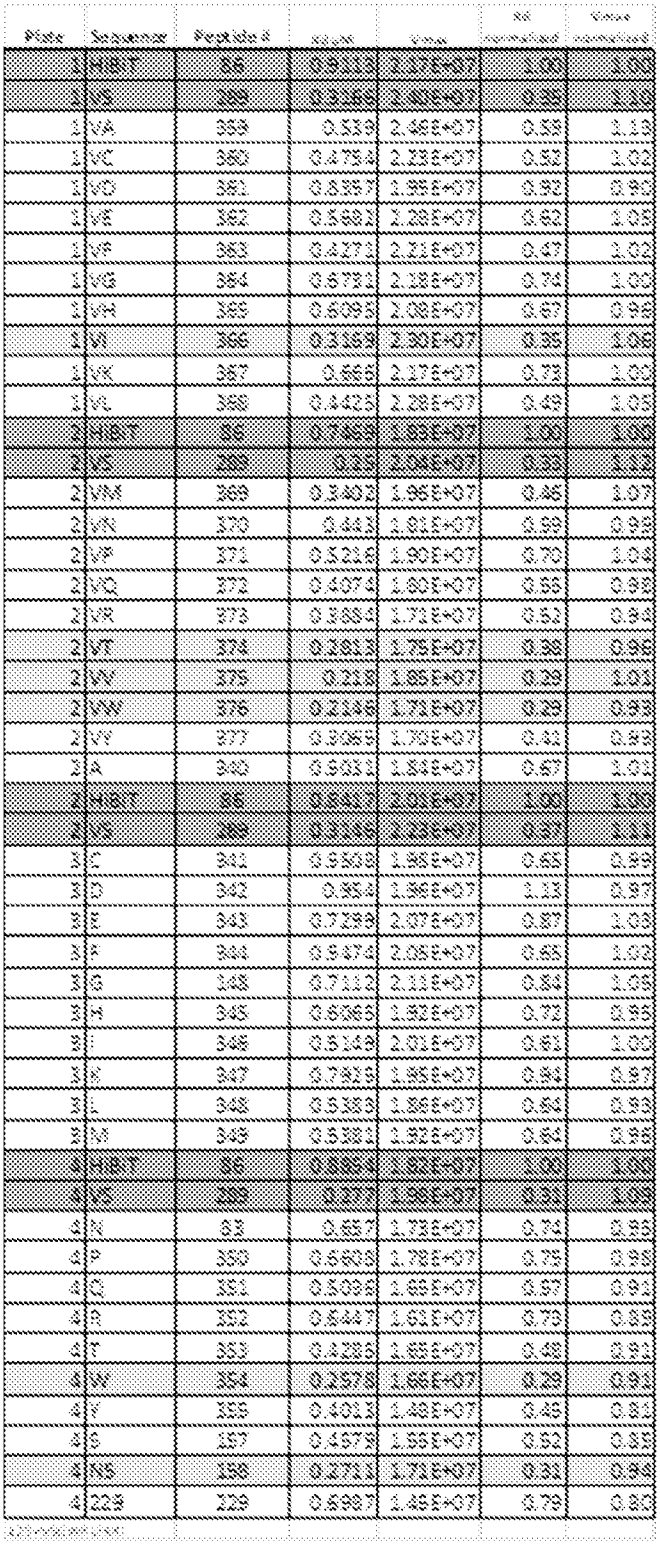
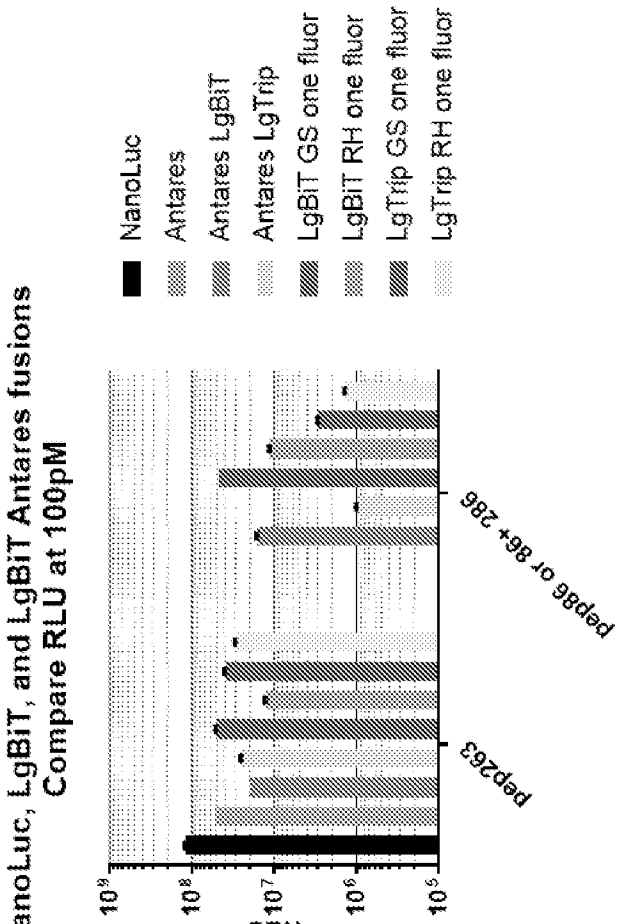


Plate	Sequence	Peptide #	[illegible]	[illegible]	[illegible]	[illegible]
1	HIGH	86	0.9113	2.17E+07	1.00	1.00
1	VS	369	0.3158	2.30E+07	0.48	1.06
1	VA	359	0.5339	2.46E+07	0.53	1.13
1	VC	360	0.4754	2.23E+07	0.52	1.03
1	VD	361	0.8397	1.95E+07	0.92	0.90
1	VE	362	0.5682	2.28E+07	0.62	1.05
1	VF	363	0.4271	2.21E+07	0.47	1.02
1	VG	364	0.6731	2.28E+07	0.74	1.05
1	VH	365	0.6095	2.08E+07	0.67	0.96
1	VI	366	0.3169	2.30E+07	0.35	1.06
1	VK	367	0.6668	2.17E+07	0.73	1.00
1	VL	368	0.4425	2.28E+07	0.49	1.05
2	HIGH	86	0.7663	1.93E+07	1.00	1.00
2	VS	369	0.23	2.04E+07	0.30	1.06
2	VM	369	0.3402	1.96E+07	0.46	1.01
2	VN	370	0.443	1.91E+07	0.59	0.99
2	VP	371	0.5216	1.90E+07	0.70	1.04
2	VQ	372	0.4074	1.80E+07	0.55	0.93
2	VR	373	0.3884	1.71E+07	0.52	0.94
2	VT	374	0.2613	1.75E+07	0.36	0.96
2	VV	375	0.218	1.85E+07	0.29	1.01
2	VW	376	0.2146	1.71E+07	0.29	0.93
2	VY	377	0.3068	1.70E+07	0.41	0.88
3	A	340	0.5031	1.84E+07	0.67	1.01
3	HIGH	86	0.6427	2.01E+07	1.00	1.00
3	VS	369	0.3148	2.23E+07	0.37	1.11
3	C	341	0.5503	1.96E+07	0.65	0.99
3	D	342	0.854	1.96E+07	1.13	0.97
3	E	343	0.7299	2.07E+07	0.87	1.03
3	F	344	0.5474	2.05E+07	0.65	1.02
3	G	345	0.7112	2.11E+07	0.84	1.05
3	H	345	0.6086	1.93E+07	0.72	0.95
3	I	346	0.5149	2.01E+07	0.61	1.00
3	K	347	0.7419	1.95E+07	0.84	0.97
3	L	348	0.5383	1.86E+07	0.64	0.92
3	M	349	0.5381	1.92E+07	0.64	0.96
4	HIGH	86	0.8854	1.82E+07	1.00	1.00
4	VS	369	0.277	1.96E+07	0.31	1.08
4	N	83	0.657	1.73E+07	0.74	0.95
4	P	350	0.6609	1.78E+07	0.75	0.98
4	Q	351	0.5036	1.65E+07	0.57	0.91
4	R	352	0.5447	1.61E+07	0.79	0.88
4	T	353	0.4286	1.65E+07	0.48	0.91
4	W	354	0.2578	1.68E+07	0.29	0.92
4	Y	355	0.4013	1.48E+07	0.45	0.81
4	9	357	0.4579	1.55E+07	0.52	0.85
4	NS	358	0.2711	1.71E+07	0.31	0.94
4	229	229	0.6987	1.49E+07	0.79	0.83

FIG. 83

Extension	peptide ID	Kd (uM)	Bmax	Kd normalized	Bmax normalized
HiBiT	86	0.68	2.50E+07	1.00	1.00
VS	289	0.23	2.92E+07	0.34	1.17
AW	457	0.19	2.75E+07	0.29	1.10
GW	458	0.19	2.88E+07	0.27	1.15
SW	459	0.19	2.90E+07	0.28	1.16
LW	460	0.22	2.84E+07	0.32	1.14
IW	461	0.20	2.70E+07	0.30	1.08
CW	462	0.38	2.72E+07	0.57	1.09
WW	463	0.46	2.62E+07	0.68	1.05
YW	464	0.22	2.77E+07	0.33	1.11

FIG. 84



2465	NanoLuc
3802	Antares NanoLuc
3803	Antares LgBiT
3804	Antares LgTrip
3815	LgBiT-8GS-CyOFP1
3816	LgBiT-RH-CyOFP1
3817	LgTrip-8GS-CyOFP1
3818	LgTrip-RH-CyOFP1

FIG. 85A

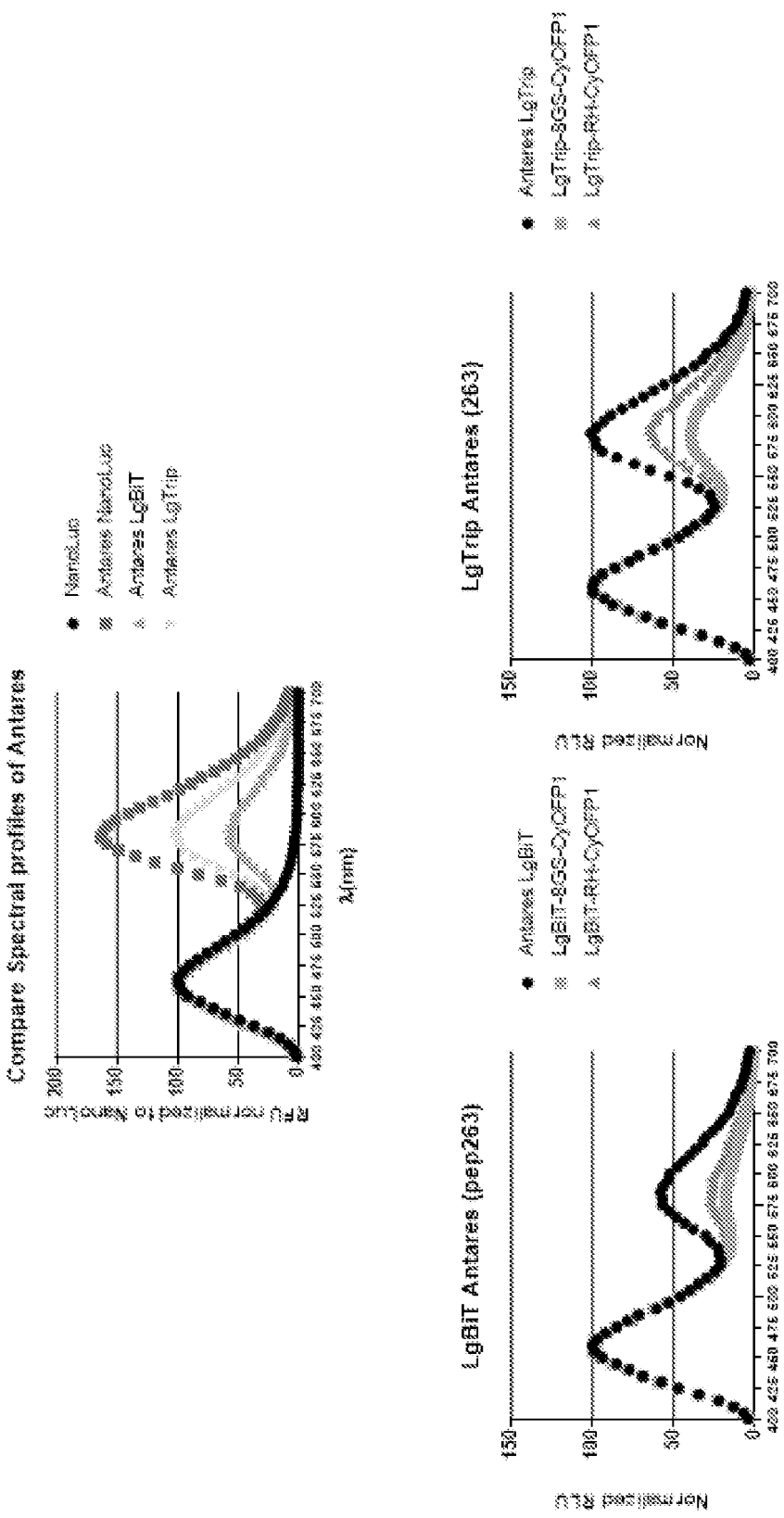
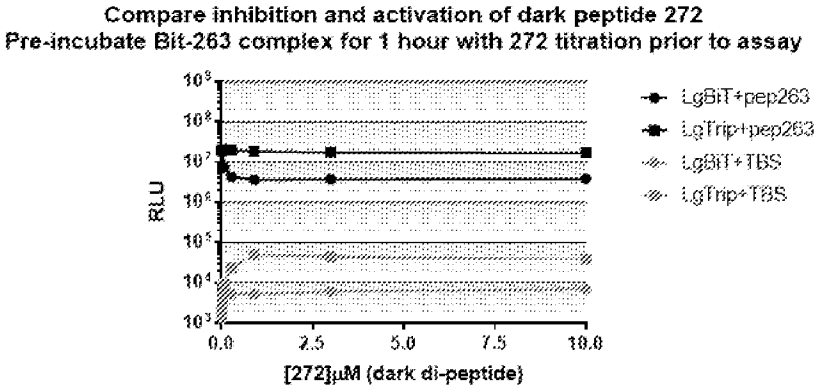
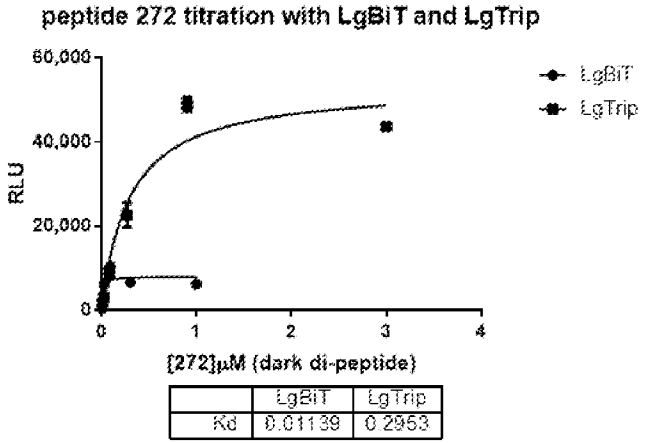

Compare Spectral profiles of Antares
NanoLuc
Antares NanoLuc
Antares LgBiT
Antares LgTrip
Normalized to NanoLuc
λ(nm)
LgBiT Antares (pep263)
Antares LgBiT
LgBiT-8GS-CyOFP1
LgBiT-RH-CyOFP1
Normalized RLU
λ(nm)
LgTrip Antares (263)
Antares LgTrip
LgTrip-8GS-CyOFP1
LgTrip-RH-CyOFP1
Normalized RLU
λ(nm)

FIG. 85B


LgBiT Antares (HiBiT)
LgTrip Antares (HiBiT + SmTrip3)
Antares LgBiT
LgBiT-8GS-CyOFP1
LgBiT-RH-CyOFP1
Antares LgTrip
LgTrip-8GS-CyOFP1
LgTrip-RH-CyOFP1
Normalized RLU
λ (nm)

FIG. 86

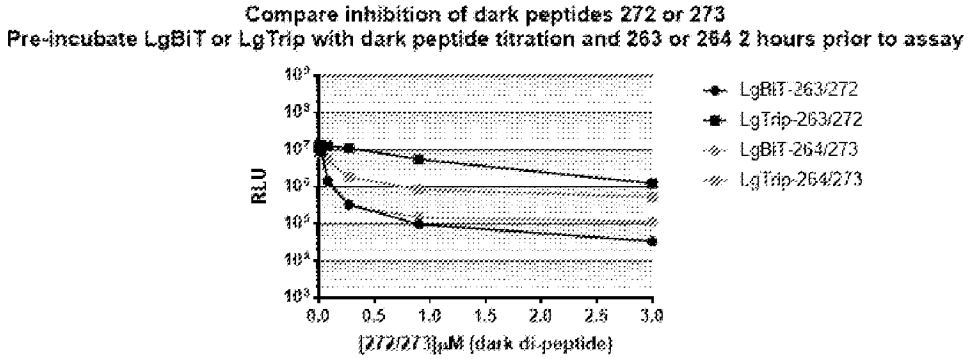


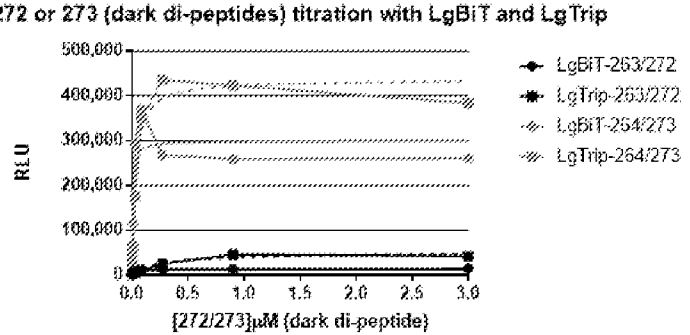


	LgBiT	LgTrip
Kd	0.01139	0.2953

Pep 272 GSMLFRVTINSVSGWALFKKIS
(SEQ ID NO.: 146)

FIG. 87

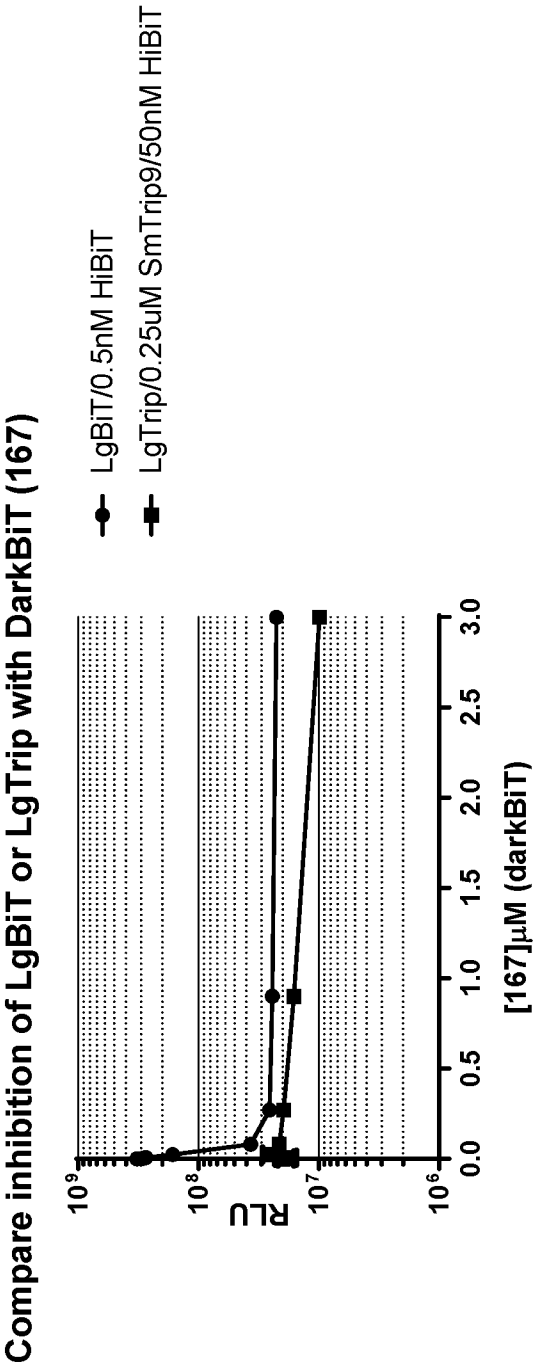




	LgBiT-263/272	LgTrip-263/272	LgBiT-264/273	LgTrip-264/273
Bmax	12729	50050	300271	436084
Kd	0.01524	0.2554	0.006736	0.02945

Peptide 273 GSMLFRVTINSGVSGWALFKKIS
Peptide 264 GSMLFRVTINSGVSGWRLFKKIS

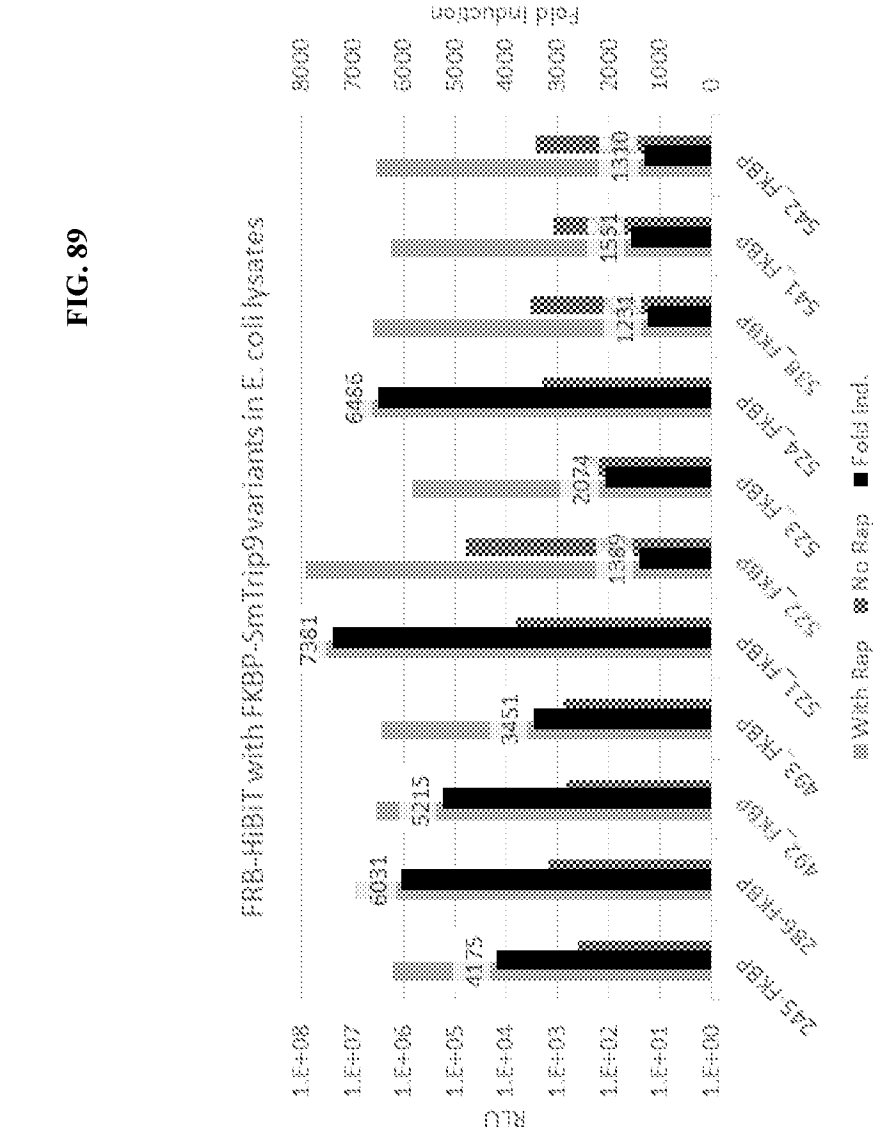
FIG. 88



SEQ ID NO.: 300
Peptide 167 VSGWALFKKIS

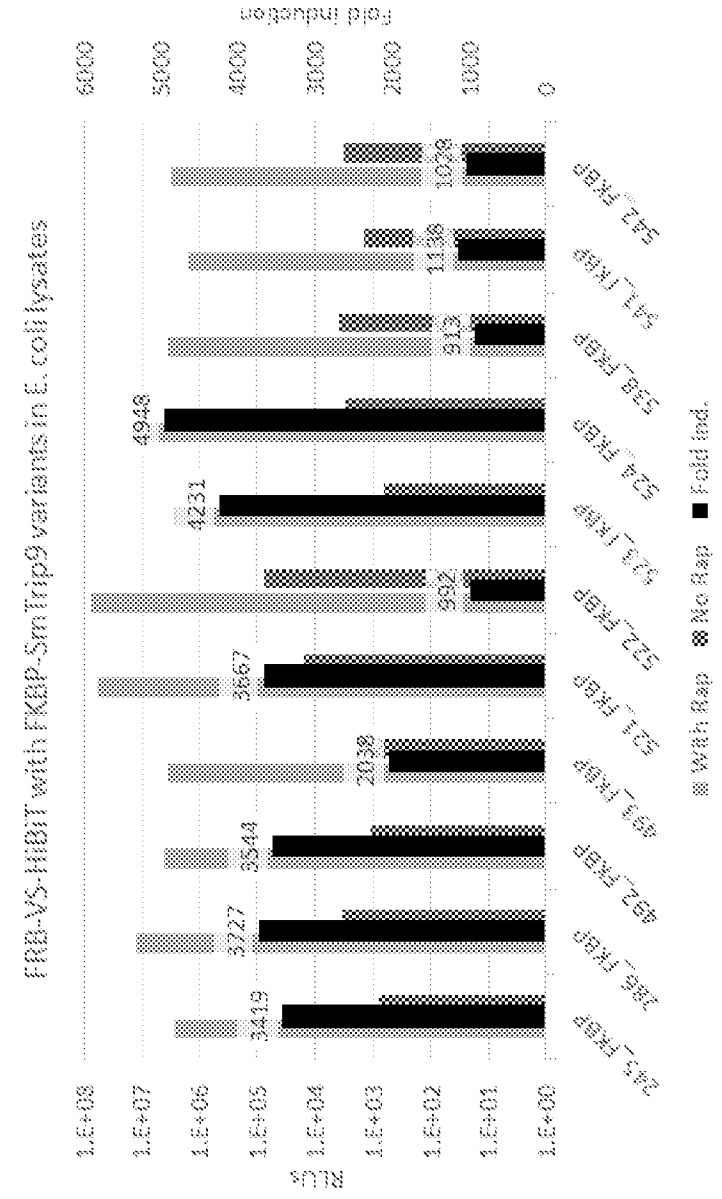
Pep ID	Sequence	SEQ ID NO.
245	GSMLFRVTINS	23
286	SSWKRGSMLFRVTINS	37
492	GSMLFRVTINKWK	313
493	GSMLFRVTIKSWK	314
521	GKMLFRVTINSWK	827
522	GKMLFRVTIWK	269
523	GSMKFRVTINSWK	270
524	GSMKFRVTIWK	271
538	GSMLFRVTIWKVK	283
541	GSMLFRVTIWKWK	286
542	GSMLFRVTIWKK	287

FIG. 89



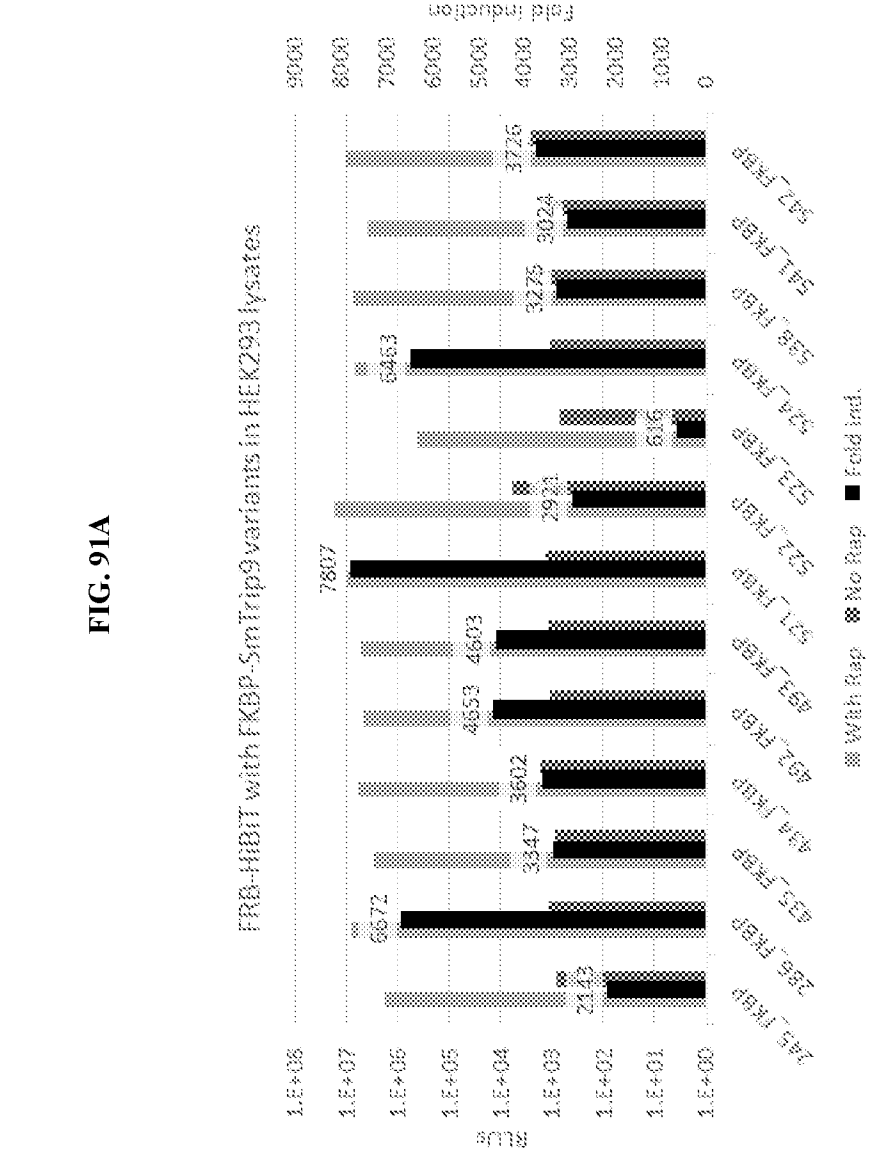
Pep ID	Sequence	SEQ ID NO.
245	GSMLFRVTINS	23
286	SSWKRGSMLFRVTINS	37
435	GSMLFRVTINSWK	310
434	GSMLFRVTIWK	230
492	GSMLFRVTINKWK	313
493	GSMLFRVTIKSWK	314
521	GKMLFRVTINSWK	827
522	GKMLFRVTIWK	269
523	GSMKFRVTINSWK	270
524	GSMKFRVTIWK	271
538	GSMLFRVTIWKVK	283
541	GSMLFRVTIWKWK	286
542	GSMLFRVTIWKK	287

FIG. 90



Pep ID	Sequence	SEQ ID NO.
245	GSMLFRVTINS	23
286	SSWKRGSMLFRVTINS	37
435	GSMLFRVTINSWK	310
434	GSMLFRVTIWK	230
492	GSMLFRVTINKWK	313
493	GSMLFRVTIKSWK	314
521	GKMLFRVTINSWK	827
522	GKMLFRVTIWK	269
523	GSMKFRVTINSWK	270
524	GSMKFRVTIWK	271
538	GSMLFRVTIWKVK	283
541	GSMLFRVTIWKWK	286
542	GSMLFRVTIWKK	287

FIG. 91A



Pep ID	Sequence	SEQ ID NO.
245	GSMLFRVTINS	23
286	SSWKRGSMLFRVTINS	37
435	GSMLFRVTINSWK	310
434	GSMLFRVTIWK	230
492	GSMLFRVTINKWK	313
493	GSMLFRVTIKSWK	314
521	GKMLFRVTINSWK	827
522	GKMLFRVTIWK	269
523	GSMKFRVTINSWK	270
524	GSMKFRVTIWK	271
538	GSMLFRVTIWKVK	283
541	GSMLFRVTIWKWK	286
542	GSMLFRVTIWKK	287

FIG. 91B

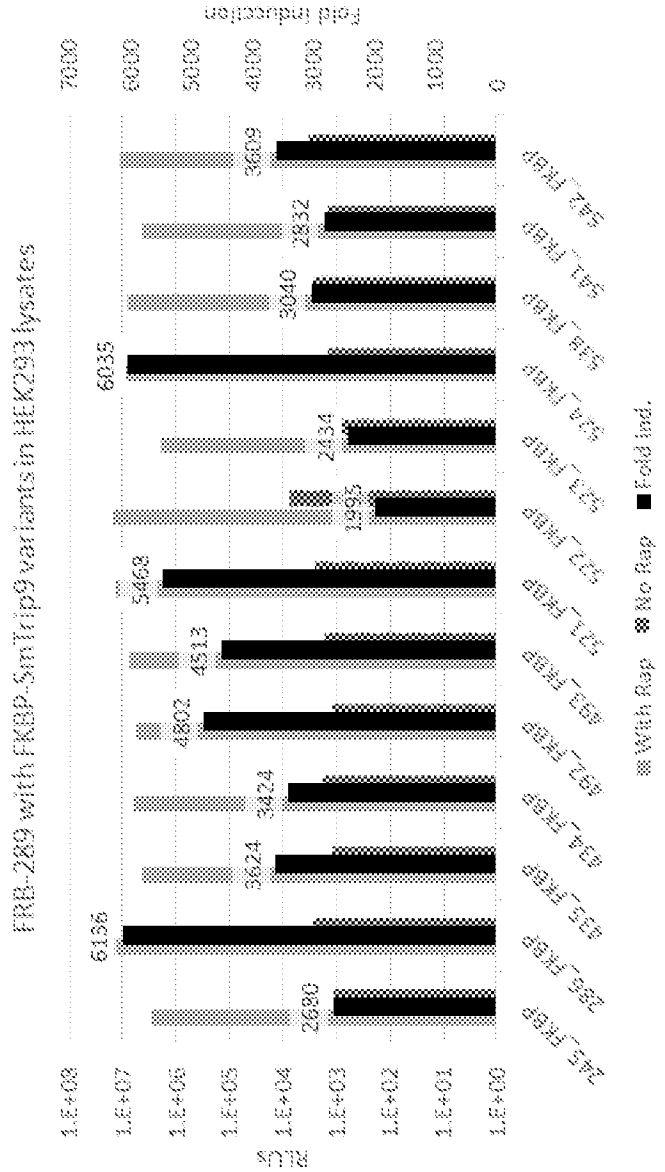


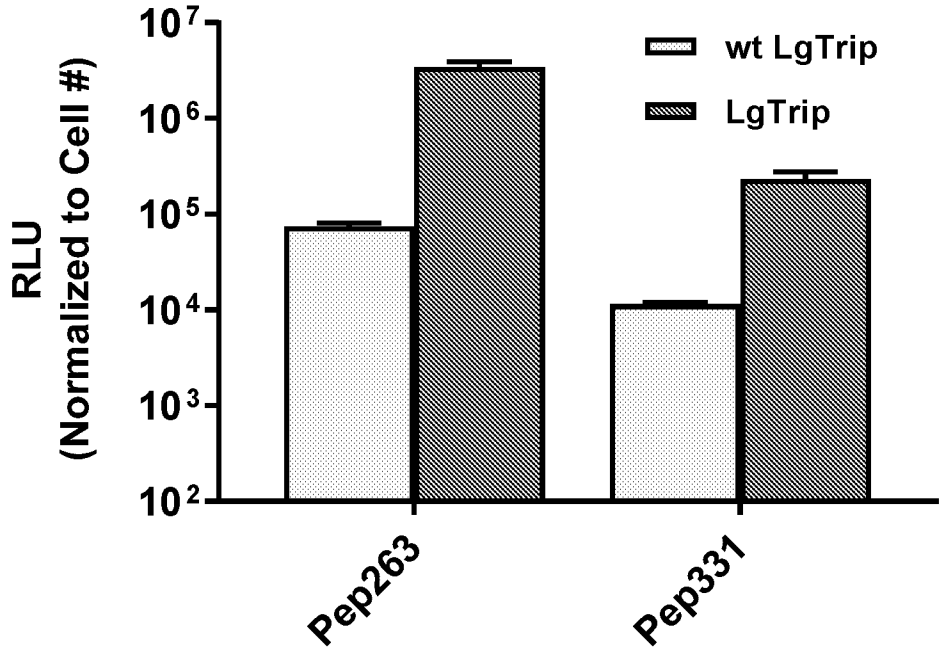
FIG. 92

FKBP construct				FRB_289 + FKBP_SmTrip9			FRB_86 + FKBP_SmTrip9			FRB_229 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response	With Rap	No Rap	Response	With Rap	No Rap	Response
3770	245	GSMLFRVTINS	23	0.53 +/- 0.04	0.24 +/- 0.02	2.26 +/- 0.11	0.41 +/- 0.03	0.16 +/- 0.01	2.53 +/- 0.07	0.21 +/- 0.01	0.10 +/- 0.01	2.21 +/- 0.34
3937	521	GKMLFRVTINSWK	827	1.00 +/- 0.04	1.00 +/- 0.07	1.00 +/- 0.04	0.85 +/- 0.02	0.70 +/- 0.01	1.21 +/- 0.05	0.49 +/- 0.03	0.48 +/- 0.03	1.02 +/- 0.13
4543	759	DKLLFTVTIEKYK	496	0.35 +/- 0.01	0.10 +/- 0.01	3.70 +/- 0.58	0.37 +/- 0.04	0.09 +/- 0.02	4.33 +/- 1.07	0.27 +/- 0.01	0.05 +/- 0.01	5.02 +/- 0.71
4792	823	EKLLFTVTIEKYK	596	0.21 +/- 0.09	0.08 +/- 0.00	2.47 +/- 0.99	0.24 +/- 0.09	0.06 +/- 0.02	4.13 +/- 0.37	0.17 +/- 0.06	0.05 +/- 0.00	3.30 +/- 0.93
4793	840	GKLLFTVTIEKYK	597	0.59 +/- 0.04	0.26 +/- 0.03	2.25 +/- 0.13	0.68 +/- 0.04	0.24 +/- 0.01	2.82 +/- 0.09	0.46 +/- 0.04	0.15 +/- 0.01	3.17 +/- 0.21

FIG. 93

Biochemical analysis of synthetic SmTrip9 variants				
Pep ID	Sequence	SEQ ID NO.	SmTrip9 Kd (uM)	Bmax (RLU)
286	SSWKRGSMLFRVTINS	37	0.19	2.57E+07
434	GSMLFRVTIWK	230	0.13	2.49E+07
521	GKMLFRVTINSWK	268	0.55	2.50E+07
522	GKMLFRVTIWK	269	0.02	2.47E+07
523	GSMKFRVTINSWK	270	1.34	1.51E+07
524	GSMKFRVTIWK	271	0.14	2.51E+07
525	GRMLFRVTINSWK	272	0.37	2.37E+07
526	GRMLFRVTIWK	273	0.01	2.32E+07
527	GSMRFRVTINSWK	274	0.72	1.06E+07
528	GSMRFRVTIWK	275	0.11	2.19E+07
529	GDMLFRVTINSWK	276	2.60	4.05E+06
530	GDMLFRVTIWK	277	0.54	2.23E+07
531	GSMDFRVTINSWK	278	1.74	1.32E+07
532	GSMDFRVTIWK	279	0.43	4.43E+06
533	GEMLFRVTINSWK	280	0.34	5.26E+06
535	GSMEFRVTINSWK	281	0.95	8.56E+06
536	GSMEFRVTIWK	282	0.18	1.63E+07
538	GSMLFRVTIWKVK	283	0.12	2.28E+07
539	GSMLFRVTIWSVK	284	0.03	1.82E+07
540	GSMLFRVTIWSK	285	0.15	2.11E+07
541	GSMLFRVTIWKWK	286	0.13	1.76E+07
542	GSMLFRVTIWKK	287	0.07	2.35E+07

FIG. 94


RLU
(Normalized to Cell #)
10^7
10^6
10^5
10^4
10^3
10^2
wt LgTrip
LgTrip
Pep263
Pep331

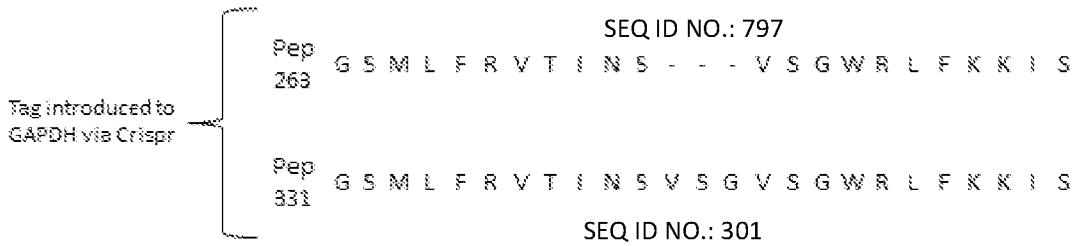
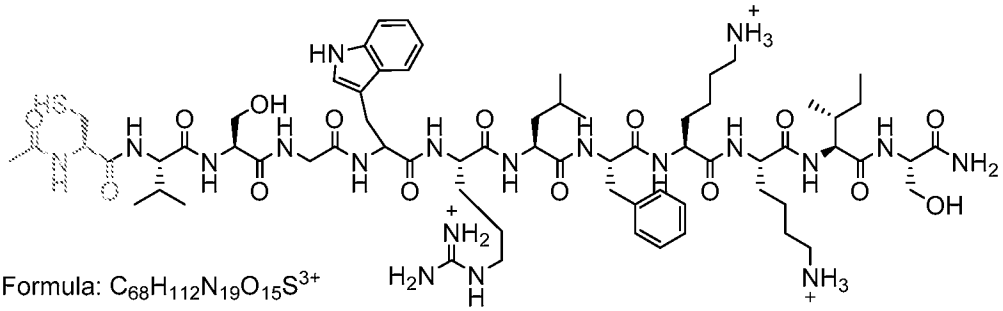
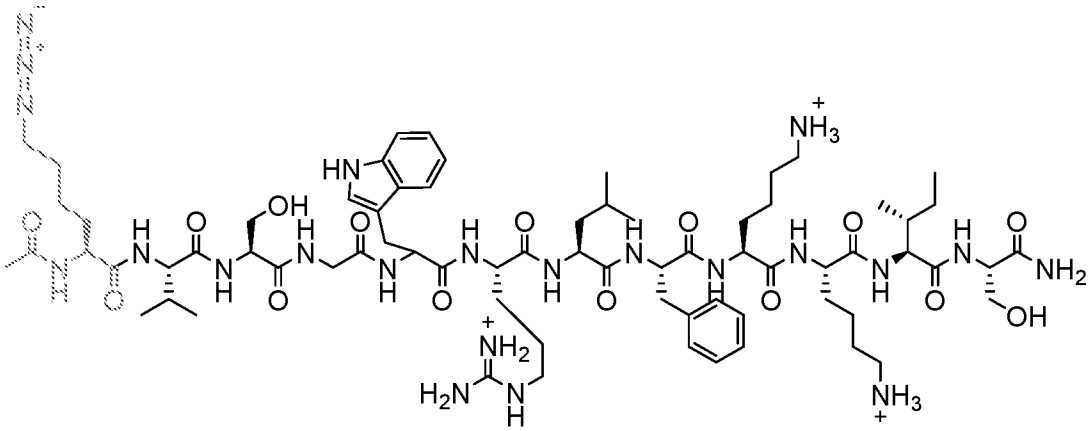

SEQ ID NO.: 797
Pep 263
G S M L F R V T I N S - - - V S G W R L F K K I S
Tag introduced to GAPDH via Crispr
Pep 331
G S M L F R V T I N S V S G V S G W R L F K K I S
SEQ ID NO.: 301

FIG. 95A



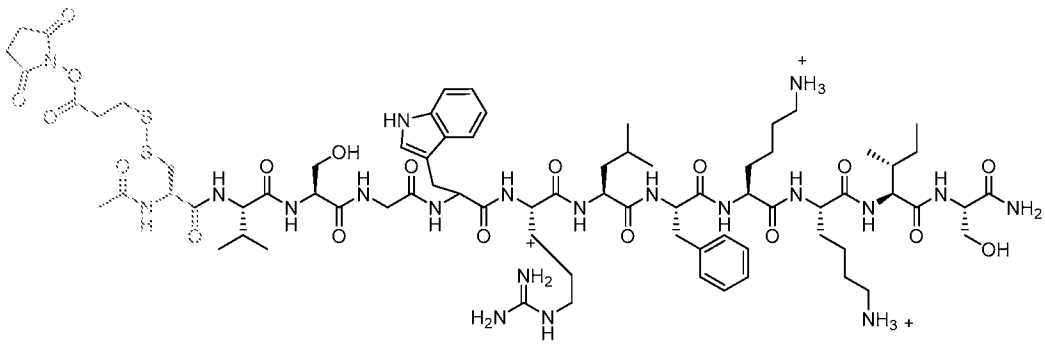
Chemical Formula: $C_{68}H_{112}N_{19}O_{15}S^{3+}$
Exact Mass: 1466.83

FIG. 95B



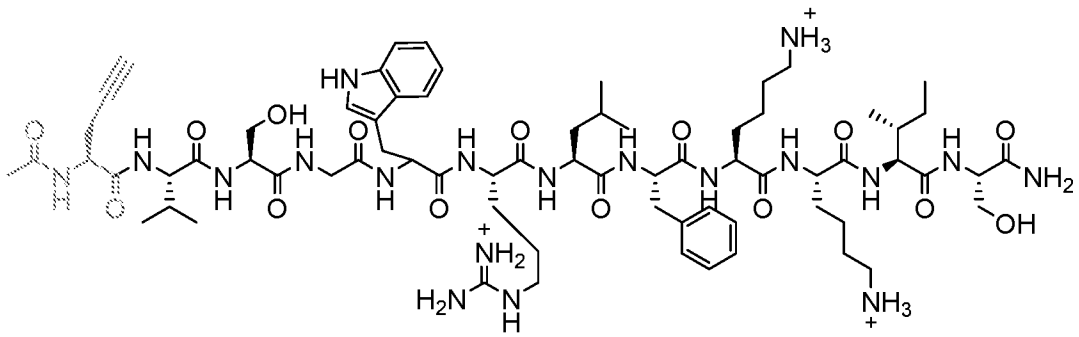
Chemical Formula: $C_{71}H_{117}N_{22}O_{15}^{3+}$
Exact Mass: 1517.91

FIG. 95C



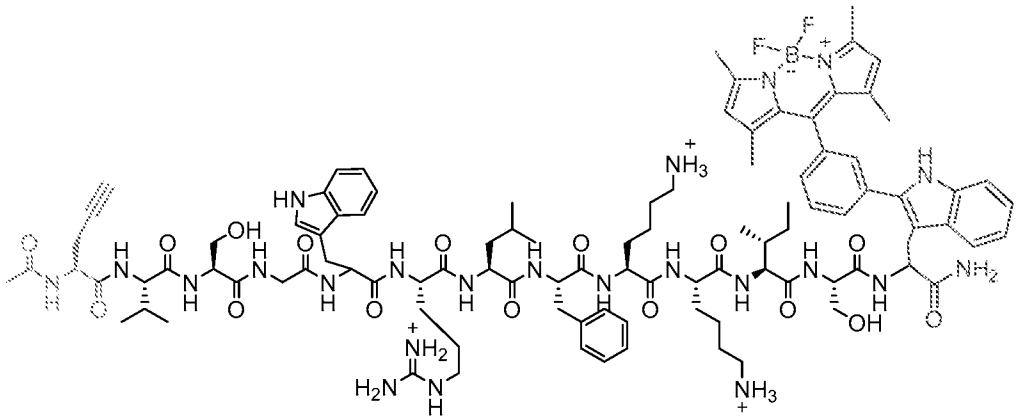
Chemical Formula: $C_{75}H_{118}N_{20}O_{19}S_2^{4+}$
Exact Mass: 1666.83

FIG. 95D



Chemical Formula: $C_{70}H_{112}N_{19}O_{15}^{3+}$
Exact Mass: 1458.86

FIG. 95E



Chemical Formula: $C_{100}H_{139}BF_2N_{23}O_{16}^{3+}$
Exact Mass: 1967.08

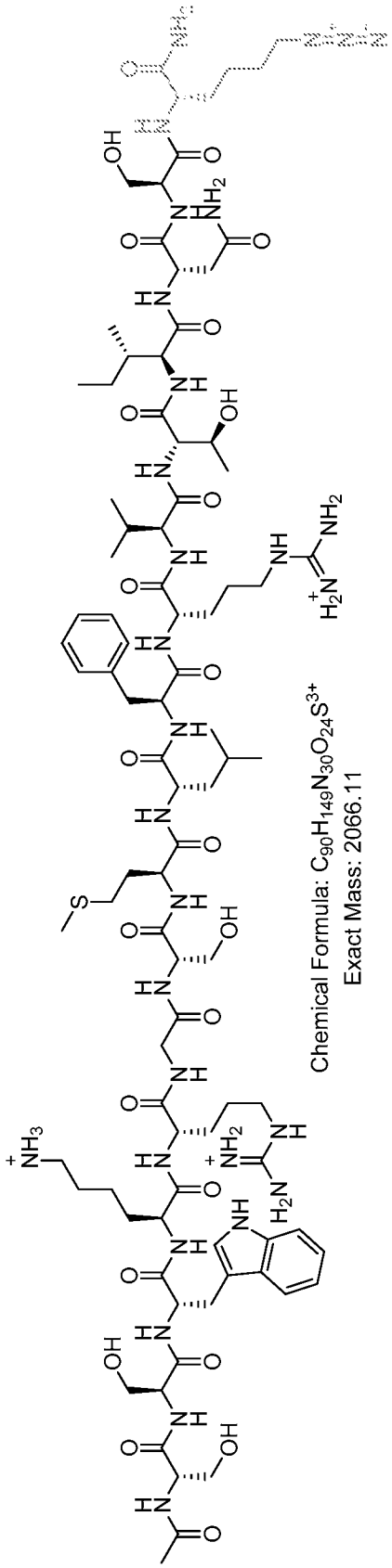


FIG. 96A

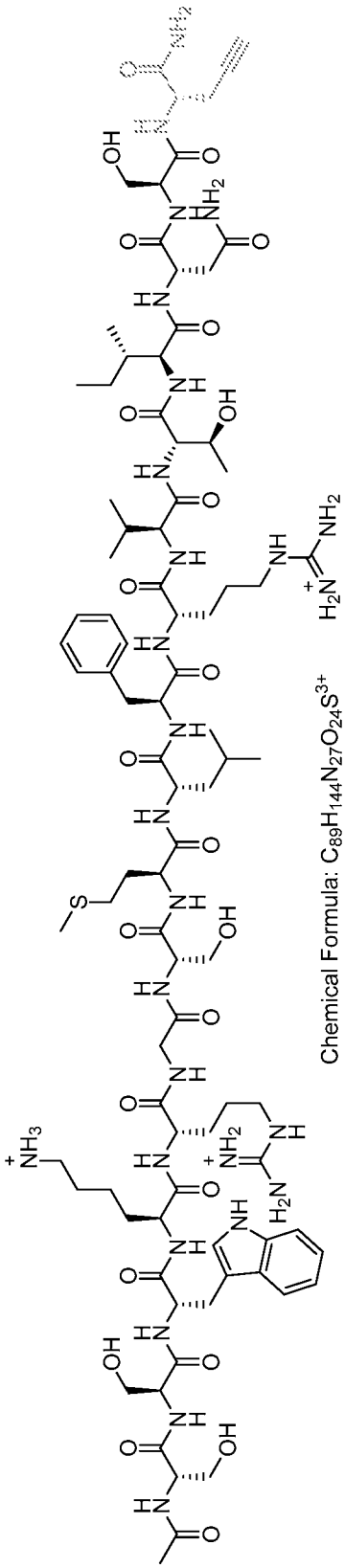
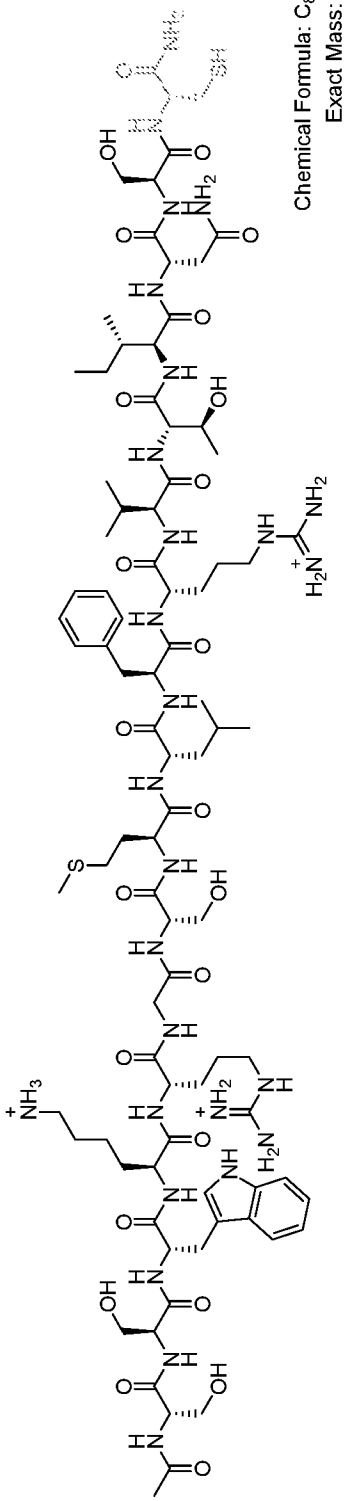
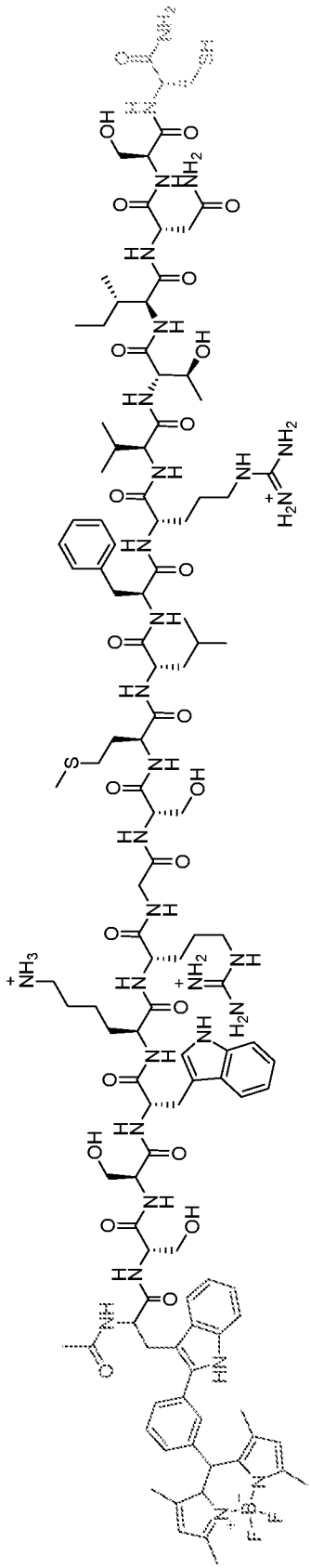


FIG. 96B



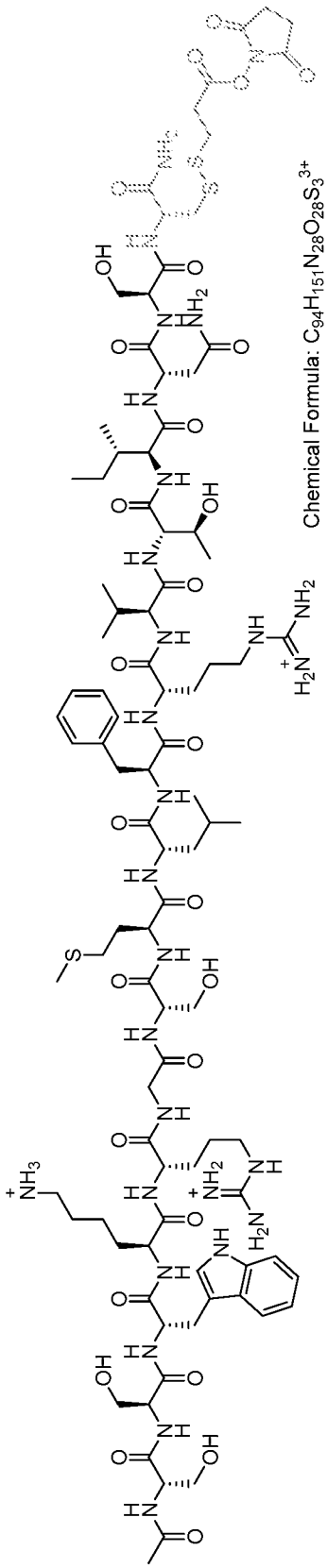
Chemical Formula: $C_{87}H_{144}N_{27}O_{24}S_2^{3+}$
Exact Mass: 2015.03

FIG. 96C



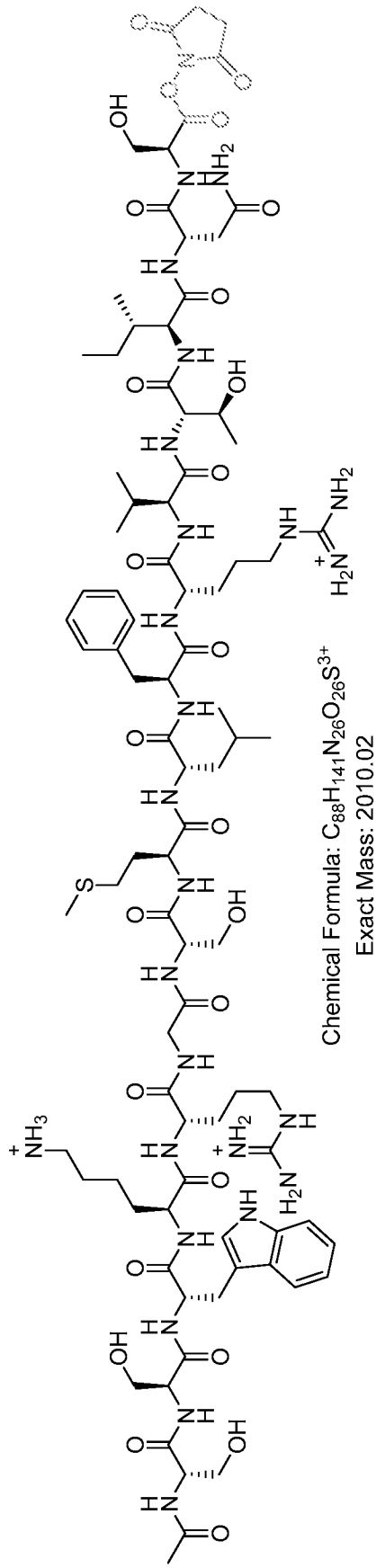
Chemical Formula: $C_{117}H_{173}BF_2N_{31}O_{25}S_2^{3+}$
Exact Mass: 2525.27

FIG. 96D



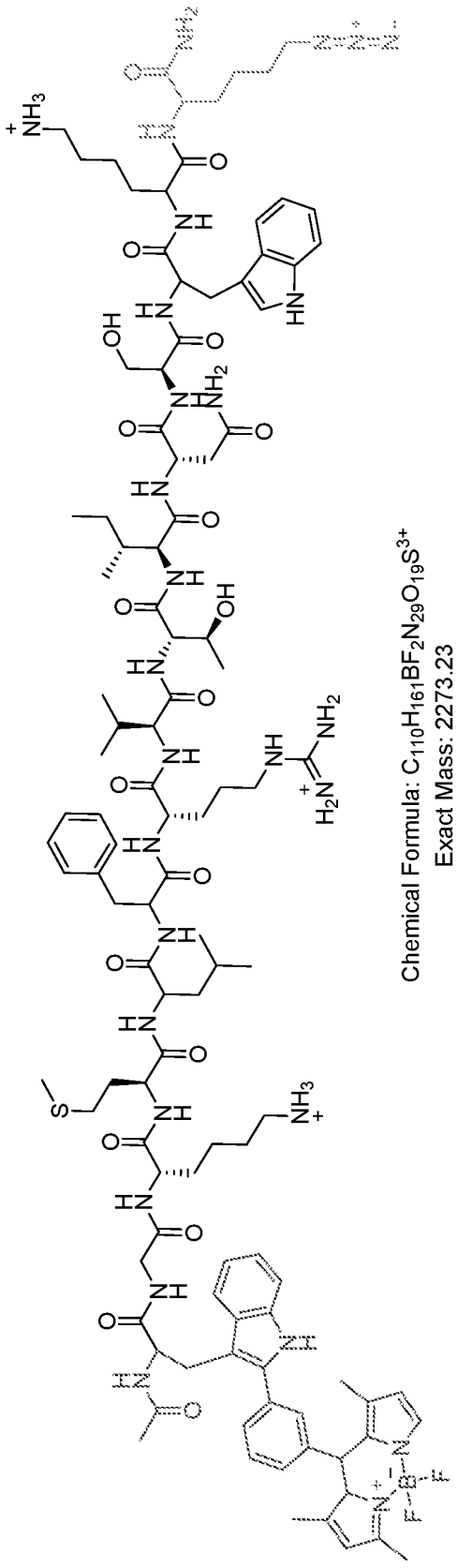
Chemical Formula: $C_{94}H_{151}N_{28}O_{28}S_3^{3+}$
Exact Mass: 2216.04

FIG. 96E



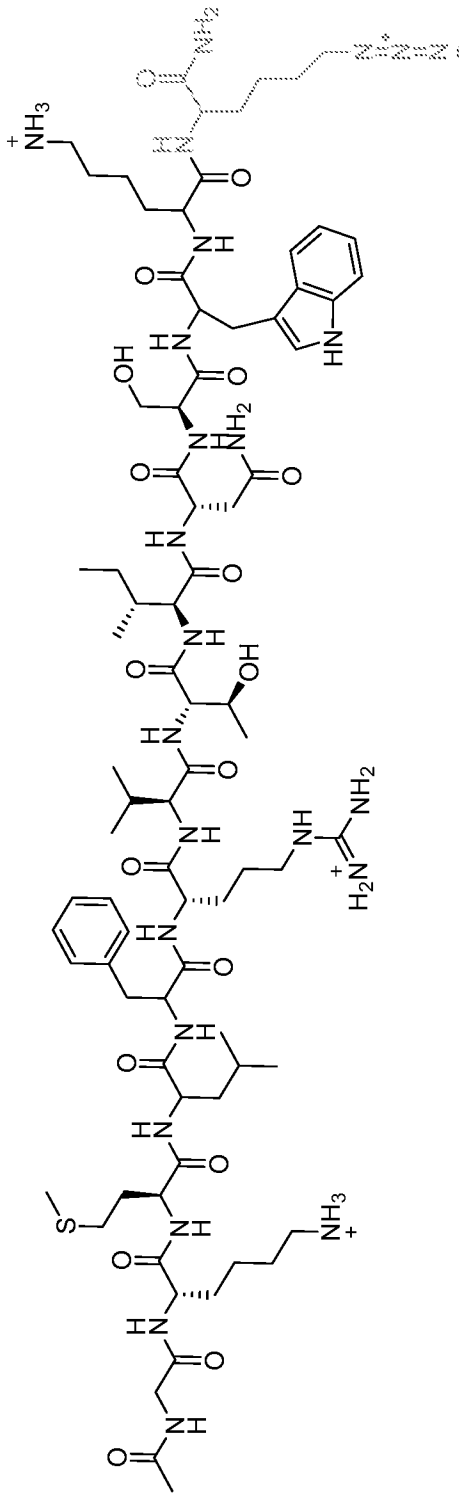
Chemical Formula: $C_{88}H_{141}N_{26}O_{26}S^{3+}$
Exact Mass: 2010.02

FIG. 96F



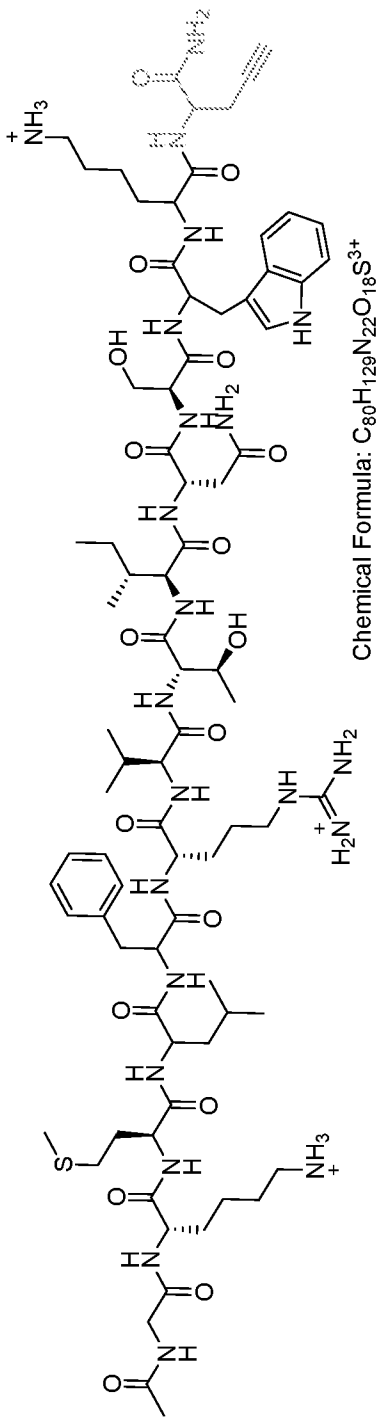
Chemical Formula: $C_{110}H_{161}BF_2N_{29}O_{19}S^{3+}$
Exact Mass: 2273.23

FIG. 97A



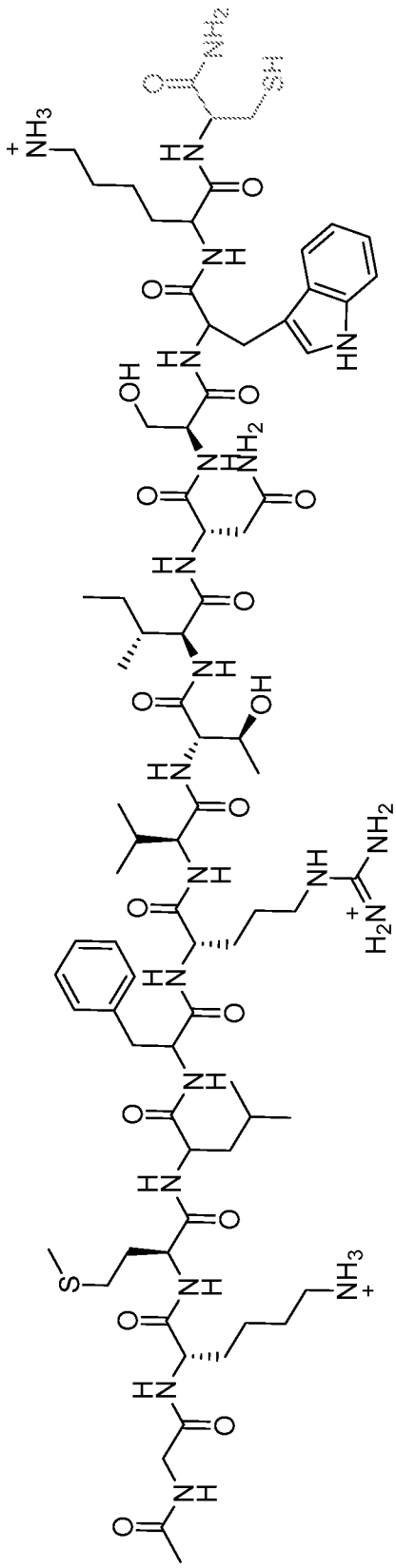
Chemical Formula: $C_{81}H_{134}N_{25}O_{18}S^{3+}$
Exact Mass: 1777.00

FIG. 97B



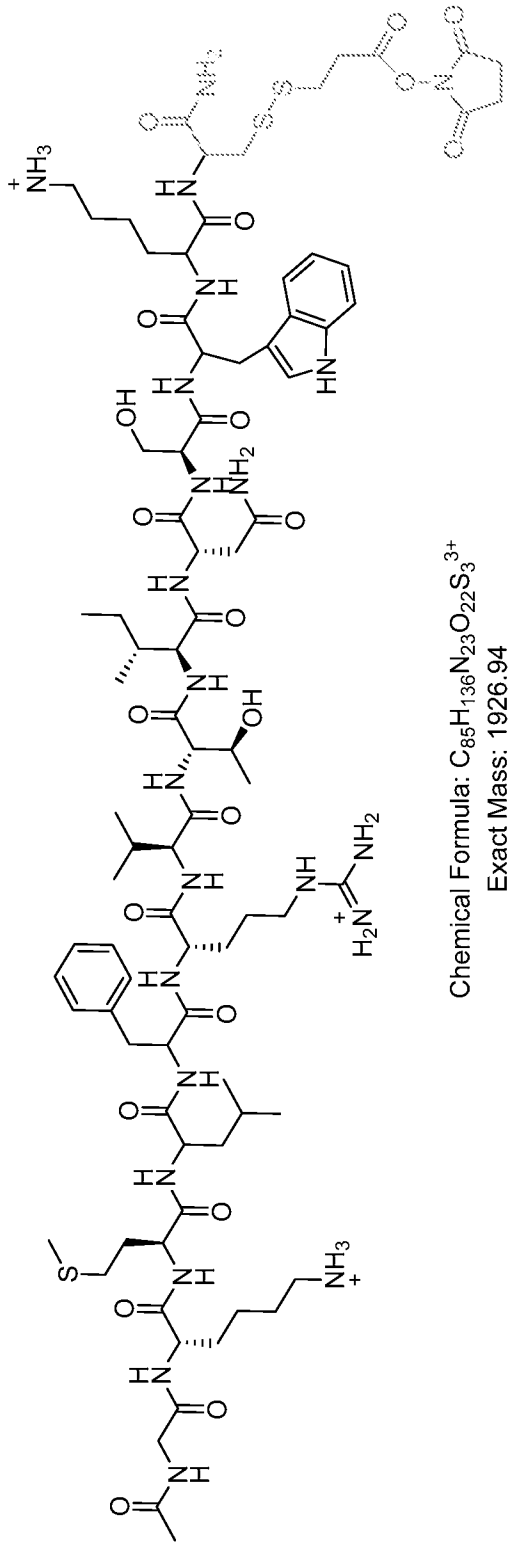
Chemical Formula: $C_{80}H_{129}N_{22}O_{18}S^{3+}$
Exact Mass: 1717.96

FIG. 97C



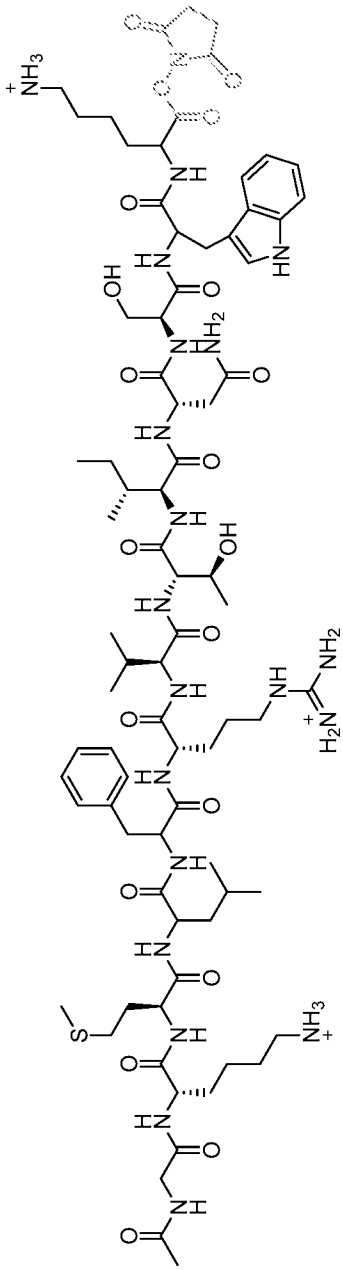
Chemical Formula: $C_{78}H_{129}N_{22}O_{18}S_2^{3+}$
Exact Mass: 1725.93

FIG. 97D



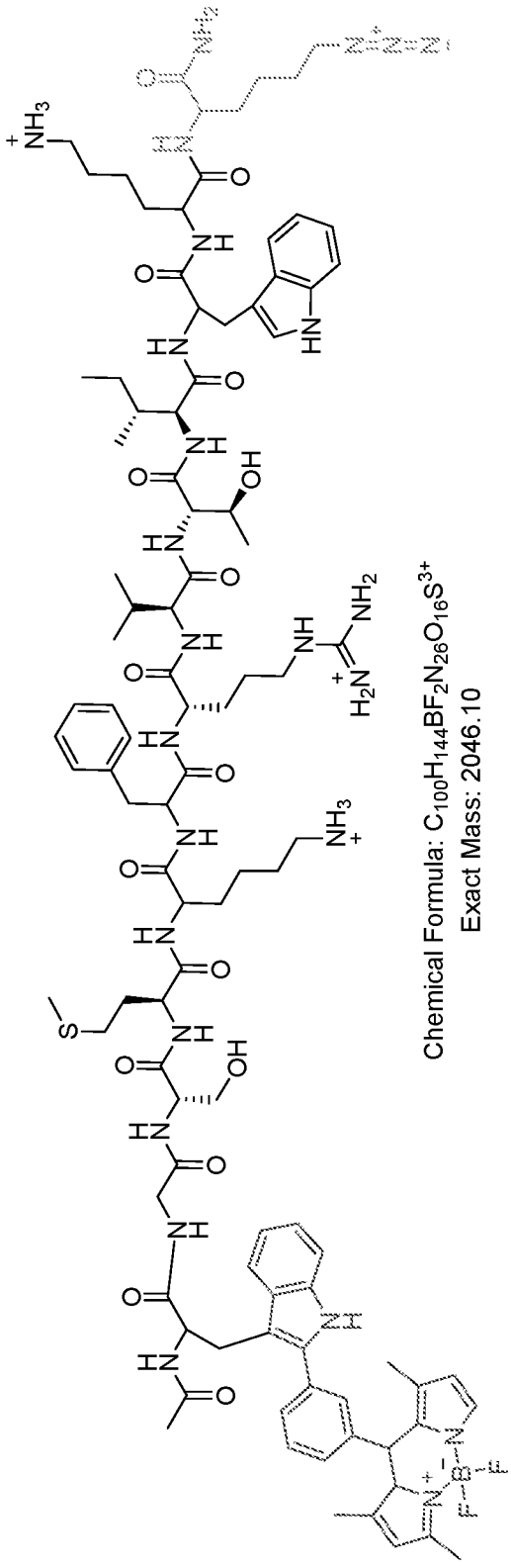
Chemical Formula: $C_{85}H_{136}N_{23}O_{22}S_3^{3+}$
Exact Mass: 1926.94

FIG. 97E



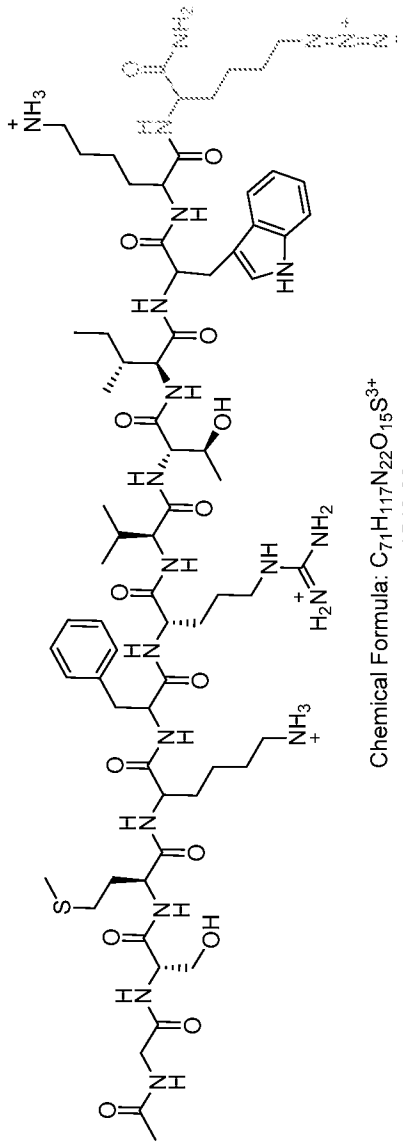
Chemical Formula: $C_{79}H_{126}N_{21}O_{20}S^{3+}$
Exact Mass: 1720.92

FIG. 97F



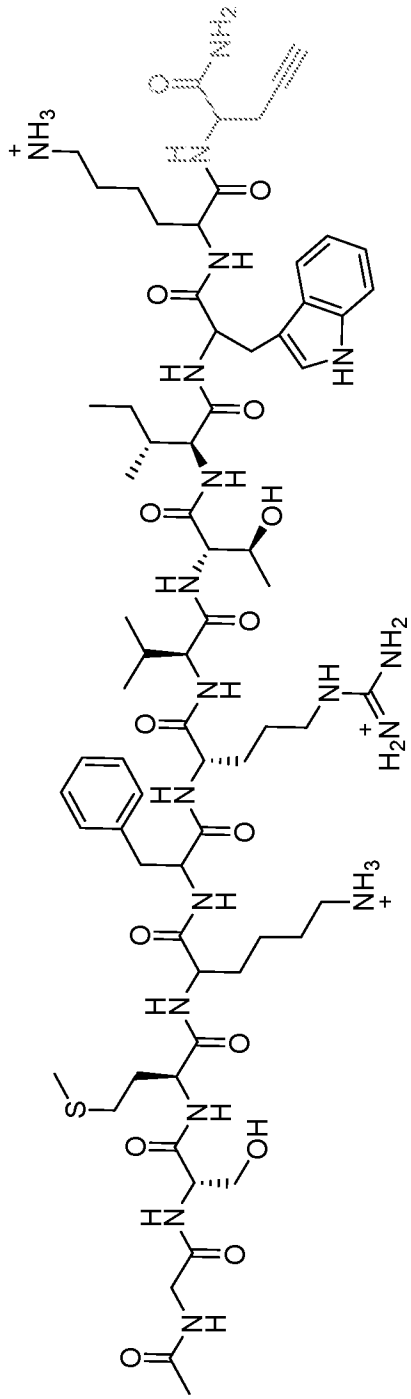
Chemical Formula: $C_{100}H_{144}BF_2N_{26}O_{16}S^{3+}$
Exact Mass: 2046.10

FIG. 98A



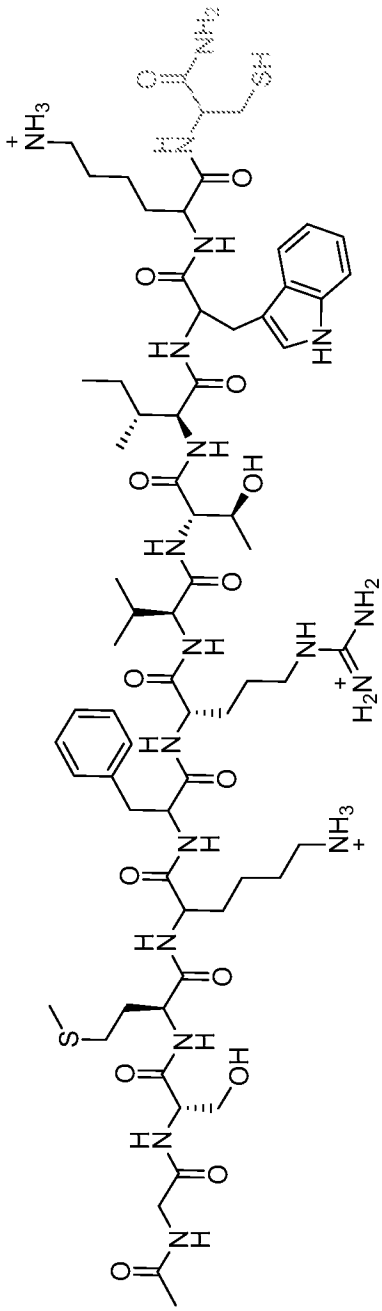
Chemical Formula: $C_{71}H_{117}N_{22}O_{15}S^{3+}$
Exact Mass: 1549.88

FIG. 98B



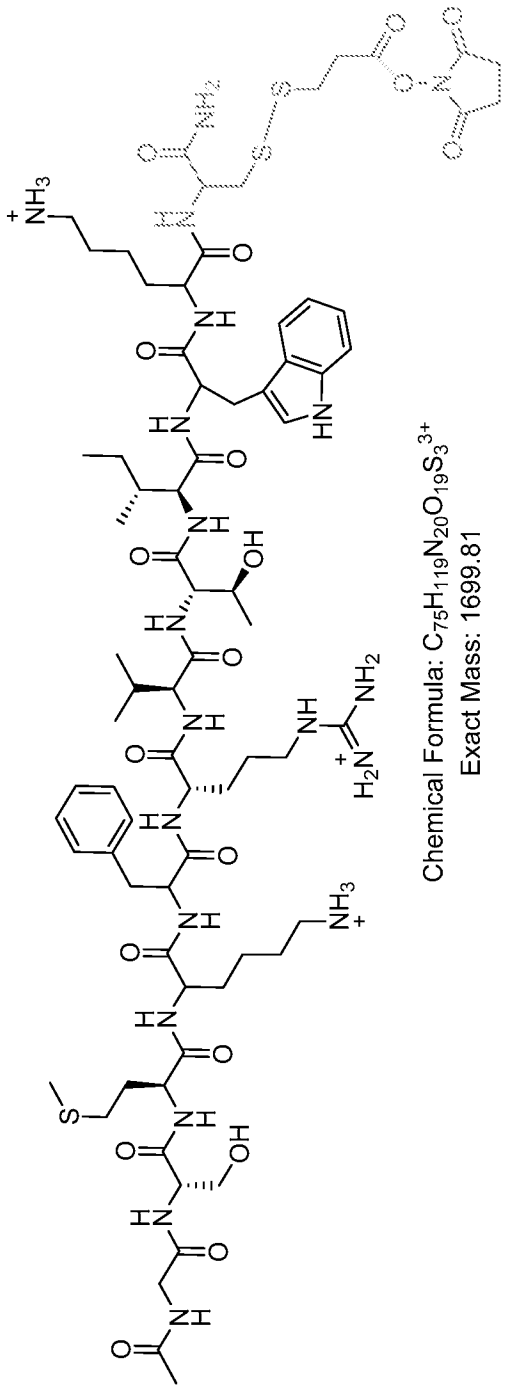
Chemical Formula: $C_{70}H_{112}N_{19}O_{15}S^{3+}$
Exact Mass: 1490.83

FIG. 98C



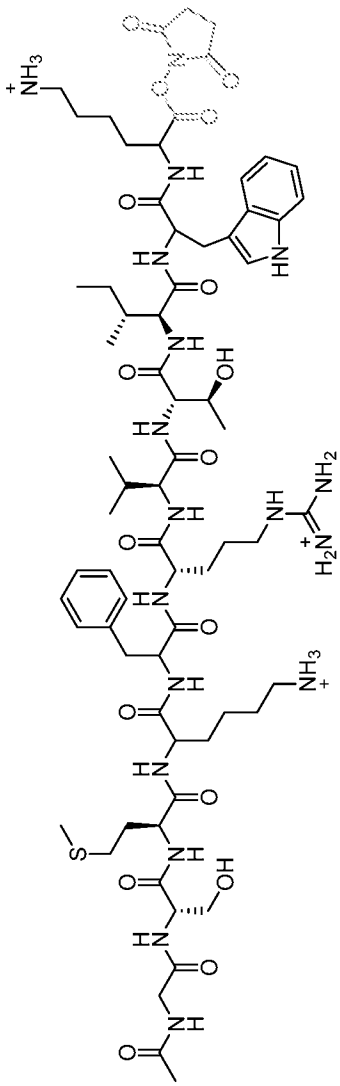
Chemical Formula: $C_{68}H_{112}N_{19}O_{15}S_2^{3+}$
Exact Mass: 1498.80

FIG. 98D



Chemical Formula: $C_{75}H_{119}N_{20}O_{19}S_3^{3+}$
Exact Mass: 1699.81

FIG. 98E



Chemical Formula: $C_{69}H_{109}N_{18}O_{17}S^{3+}$
Exact Mass: 1493.79

FIG. 98F

FIG. 99A

SmTrip9-286 with C-term azide

Trip9-286 peptide

SEQ ID NO.: 33

aatcgagaaccatgacac

Expected mass: 8229.27

Measured mass: 8229.4

FIG. 99B

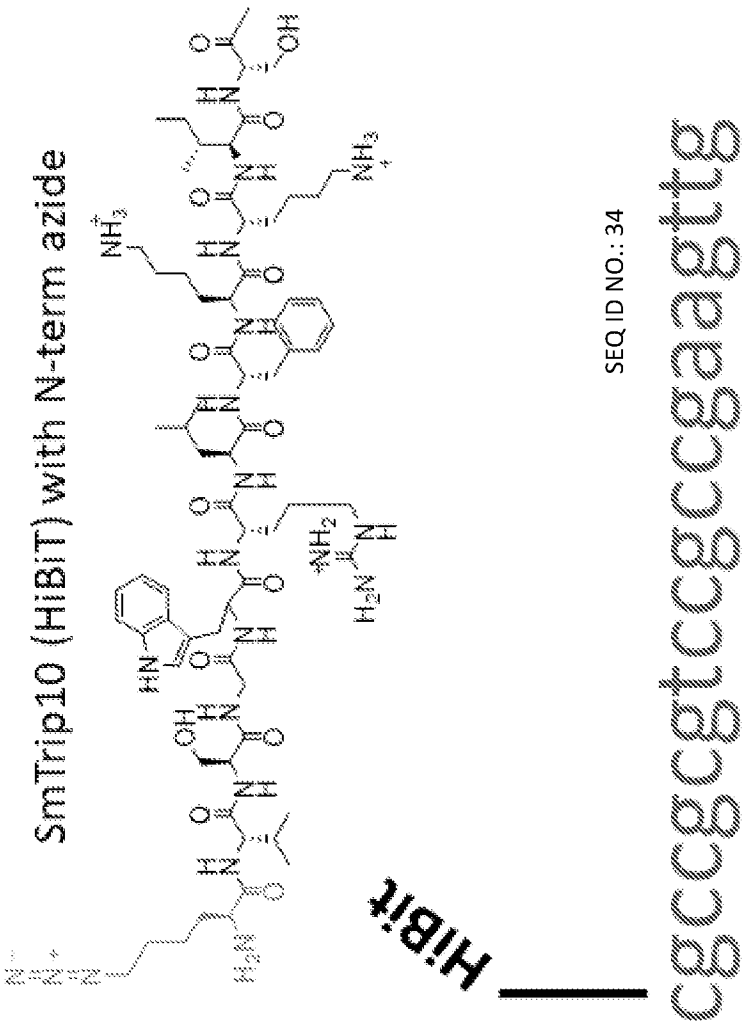

SmTrip10 (HiBiT) with N-term azide
HiBiT
SEQ ID NO.: 34
ccccgtccccgaatg
Expected mass: 7673
Measured mass: 7673.5

FIG. 100

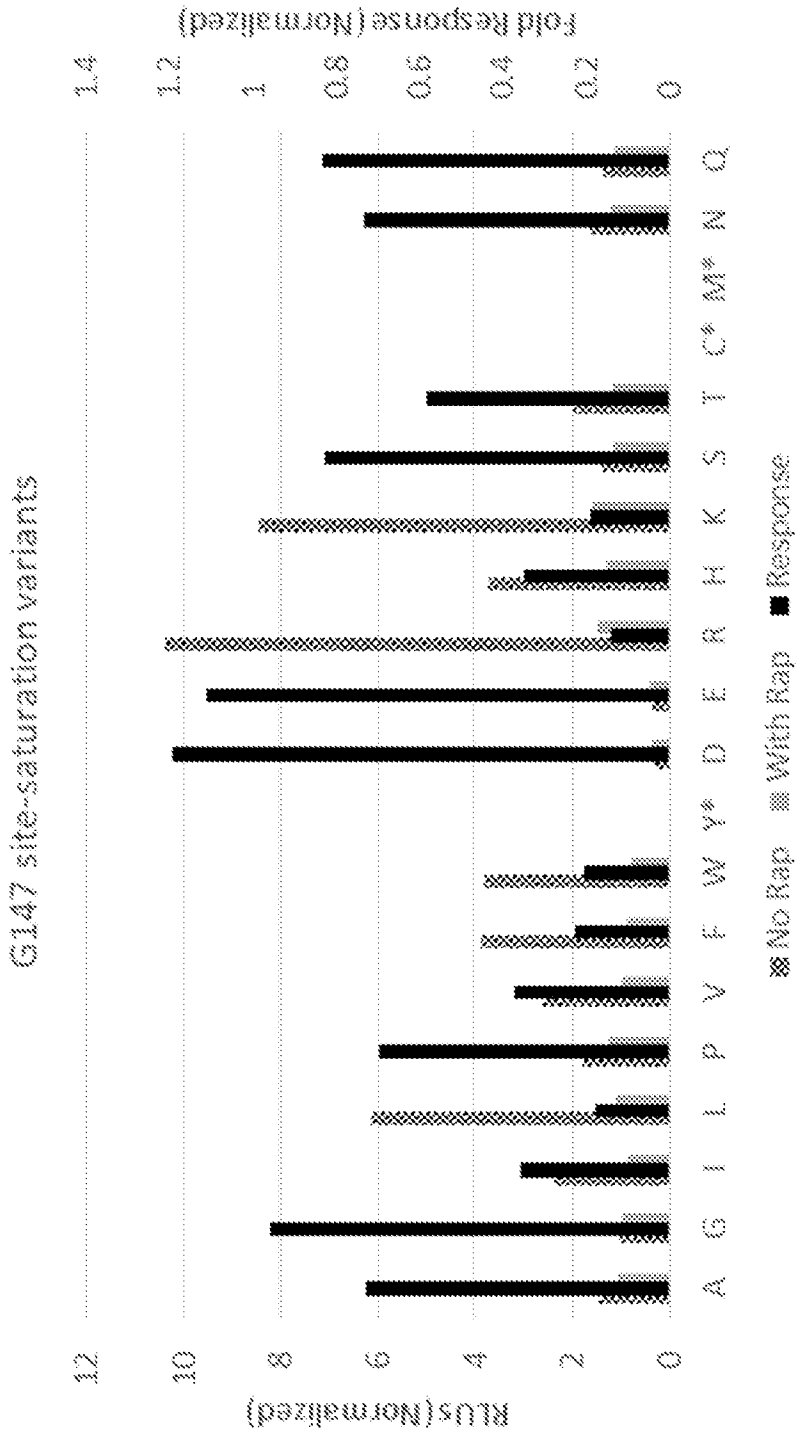

G147 site-saturation variants
RLUs (Normalized)
Fold Response (Normalized)
No Rap
With Rap
Response
A G I L P V F W Y* D E R H K S T C* M* N Q
0 2 4 6 8 10 12
0 0.2 0.4 0.6 0.8 1 1.2 1.4

FIG. 101

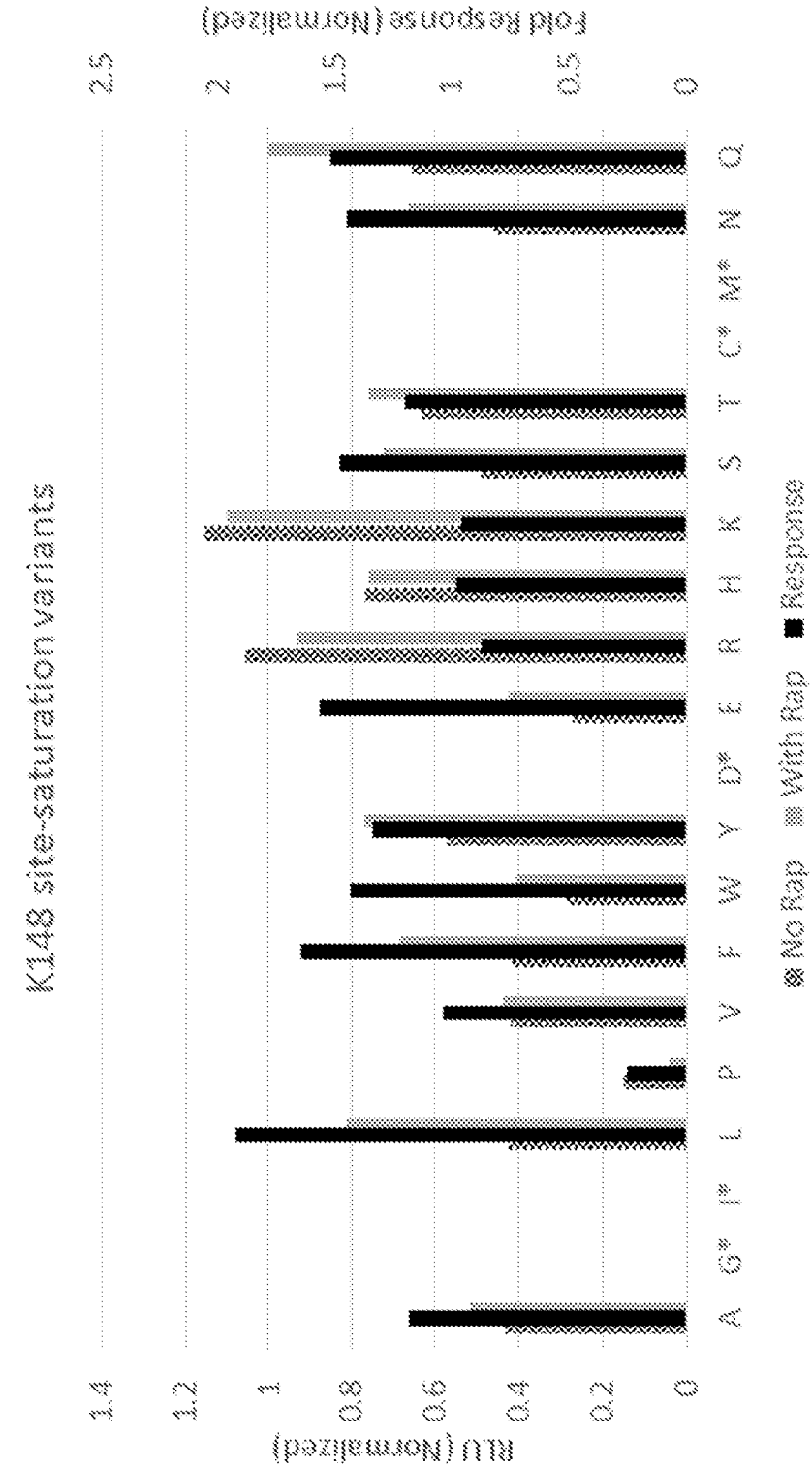

K148 site-saturation variants
RLU (Normalized)
Fold Response (Normalized)
No Rap
With Rap
Response
A G* I* L P V F W Y D* E R H K S T C* M* N Q

FIG. 102

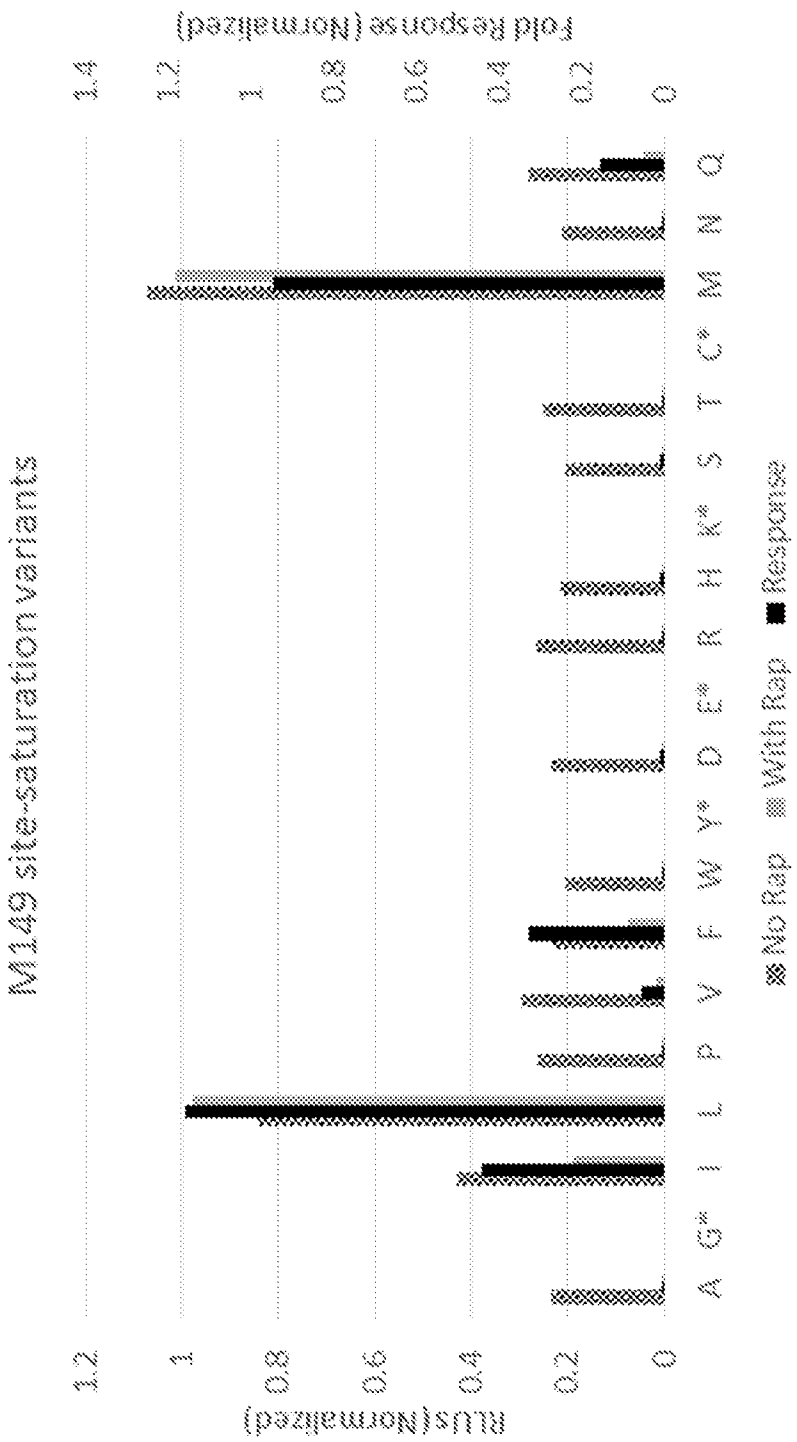

M149 site-saturation variants
RLUs (Normalized)
Fold Response (Normalized)
No Rap
With Rap
Response
A G* I L P V F W Y* D E* R H K* S T C* M N Q
0 0.2 0.4 0.6 0.8 1 1.2
0 0.2 0.4 0.6 0.8 1 1.2 1.4

FIG. 103

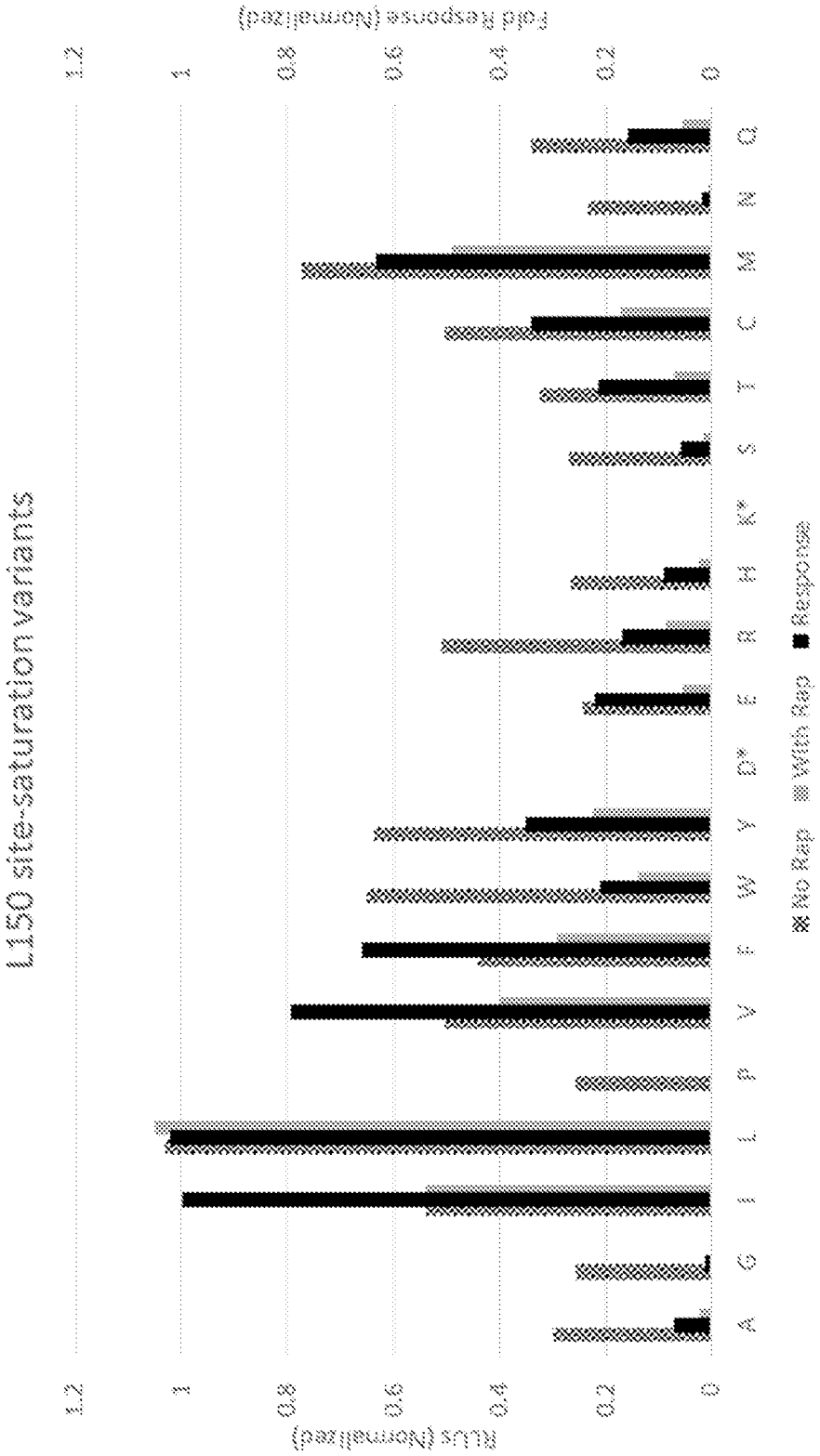

L150 site-saturation variants
Fold Response (Normalized)
RLUs (Normalized)
1.2
1
0.8
0.6
0.4
0.2
0
A G I L P V F W Y D* E R H K* S T C M N Q
No Rep
With Rep
Response

FIG. 104

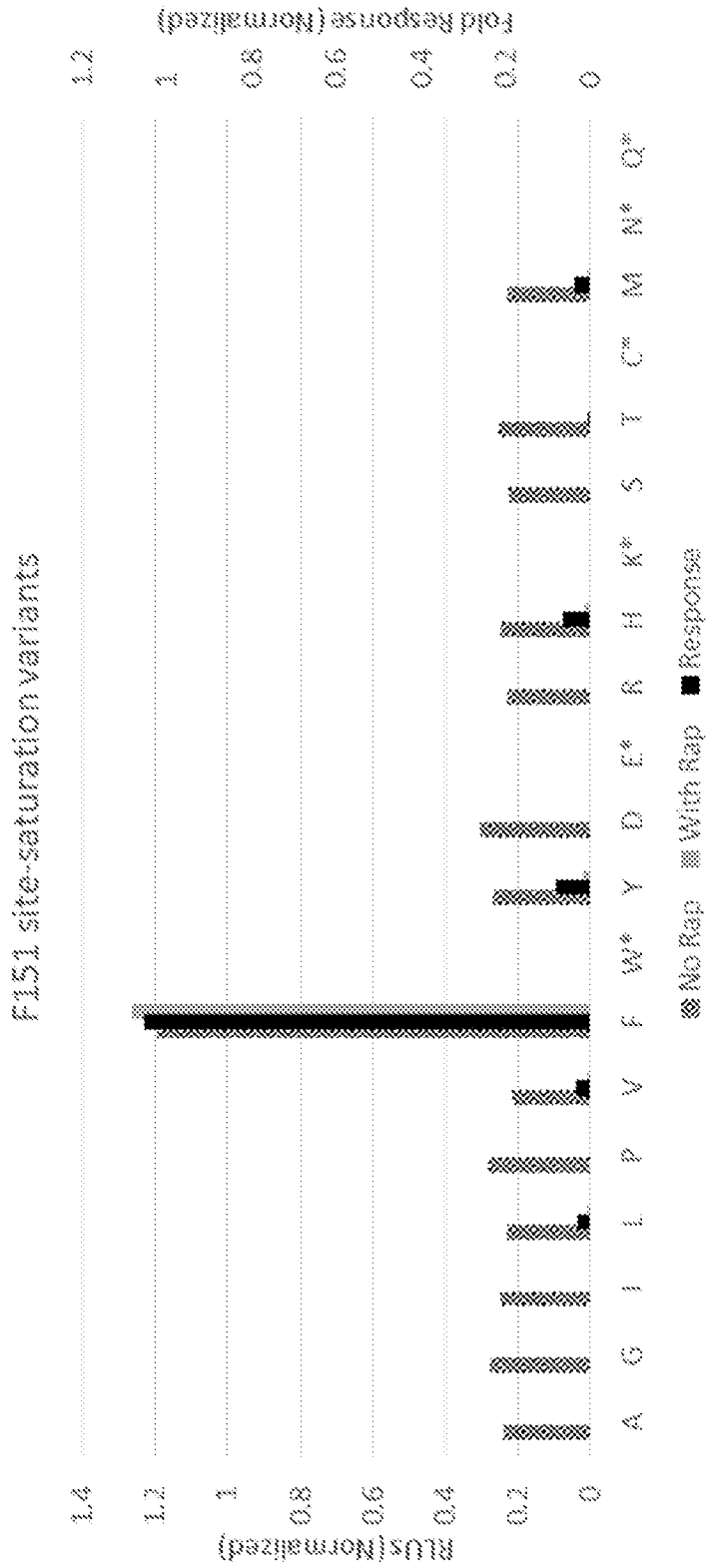

F151 site-saturation variants
RLUs (Normalized)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
A G I L P V F W* Y D E* R H K* S T C* M N* Q*
No Rap
With Rap
Response
Fold Response (Normalized)
1.2
1
0.8
0.6
0.4
0.2
0

FIG. 105

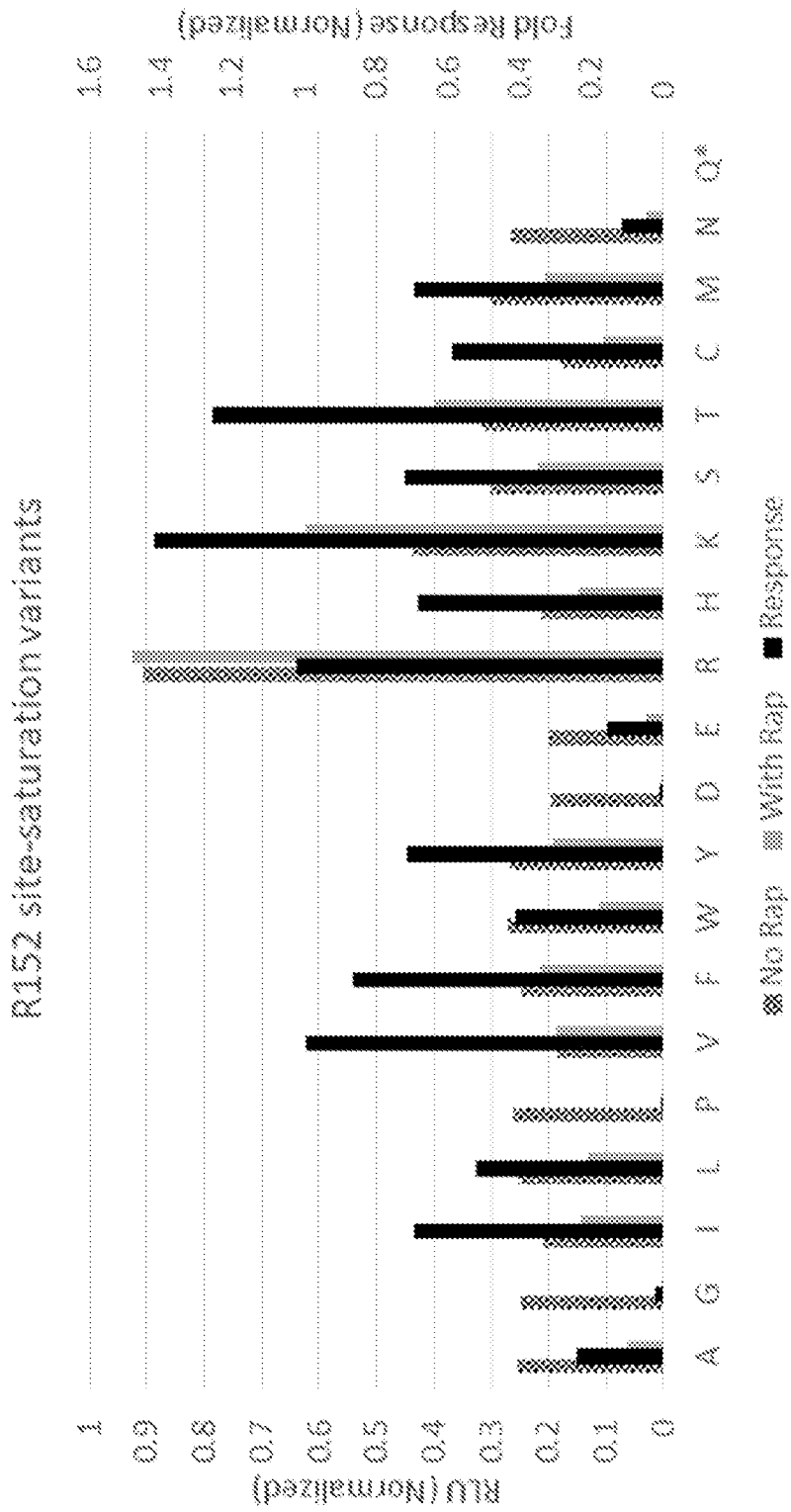

R152 site-saturation variants
RLU (Normalized)
Fold Response (Normalized)
No Rap
With Rap
Response
A G I L P V F W Y D E R H K S T C M N Q*

FIG. 106

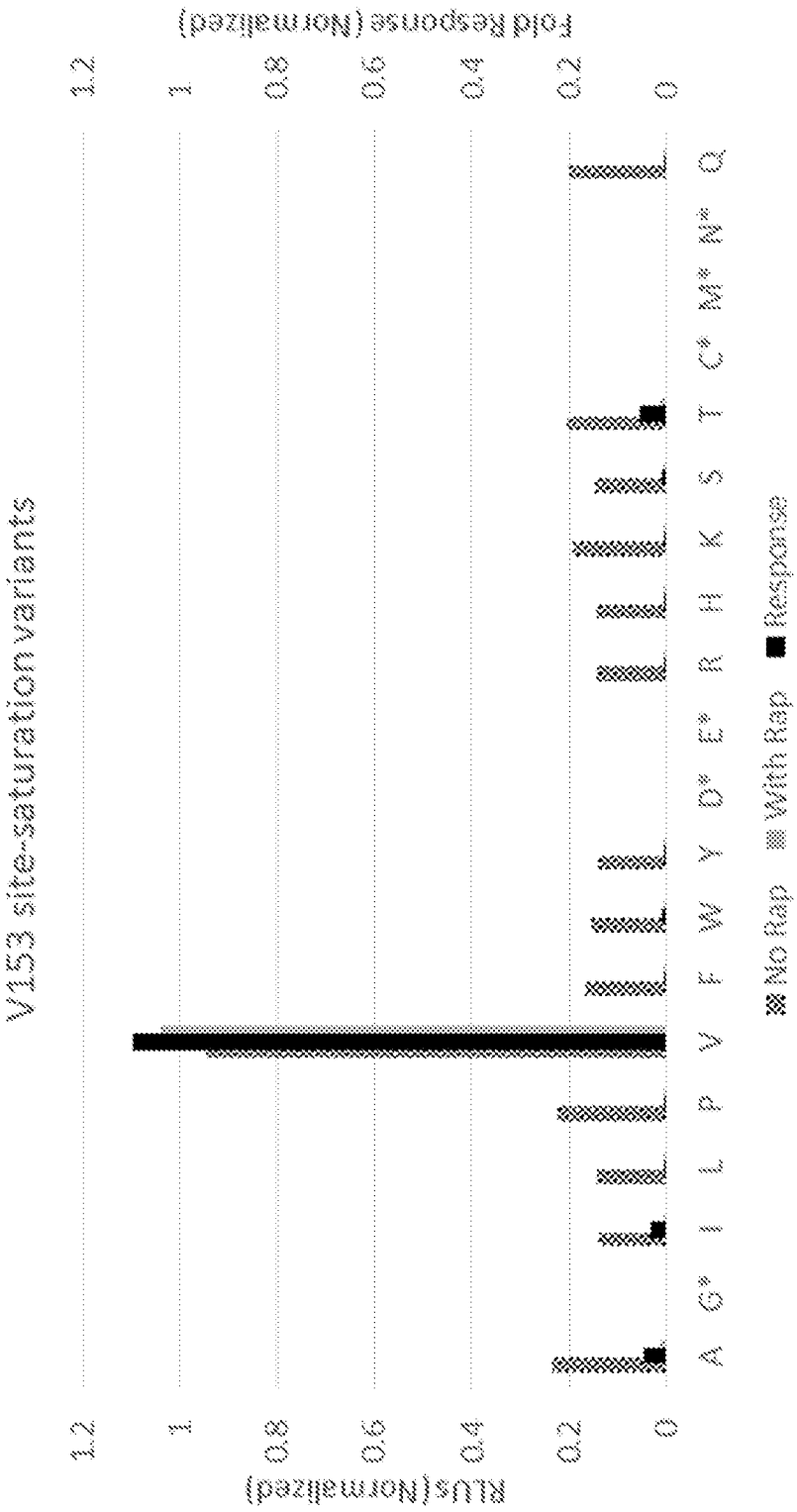

V153 site-saturation variants
RLUs (Normalized)
Fold Response (Normalized)
0
0.2
0.4
0.6
0.8
1
1.2
A G* I L P V F W Y D* E* R H K S T C* M* N* Q
No Rap
With Rap
Response

FIG. 107

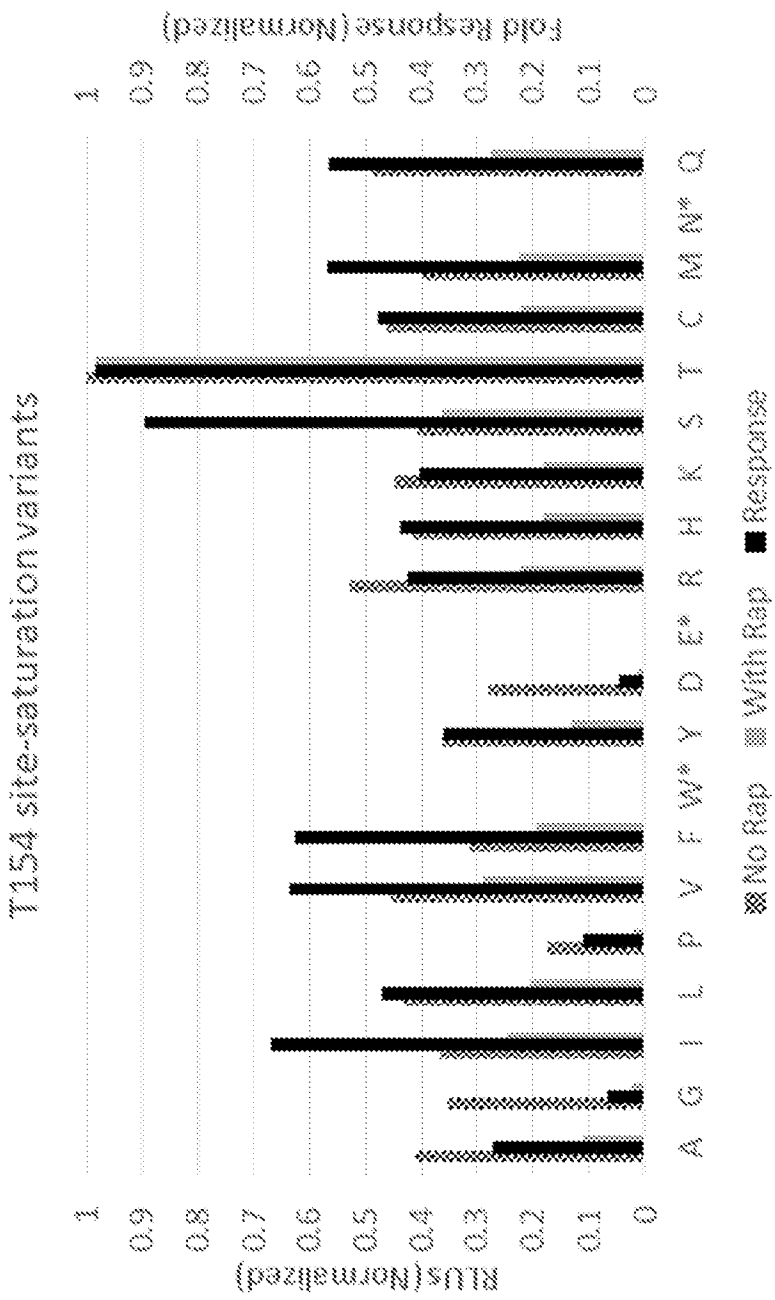

T154 site-saturation variants
RLUs (Normalized)
Fold Response (Normalized)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
A G I L P V F W* Y D E* R H K S T C M N* Q
No Rap
With Rap
Response

FIG. 108

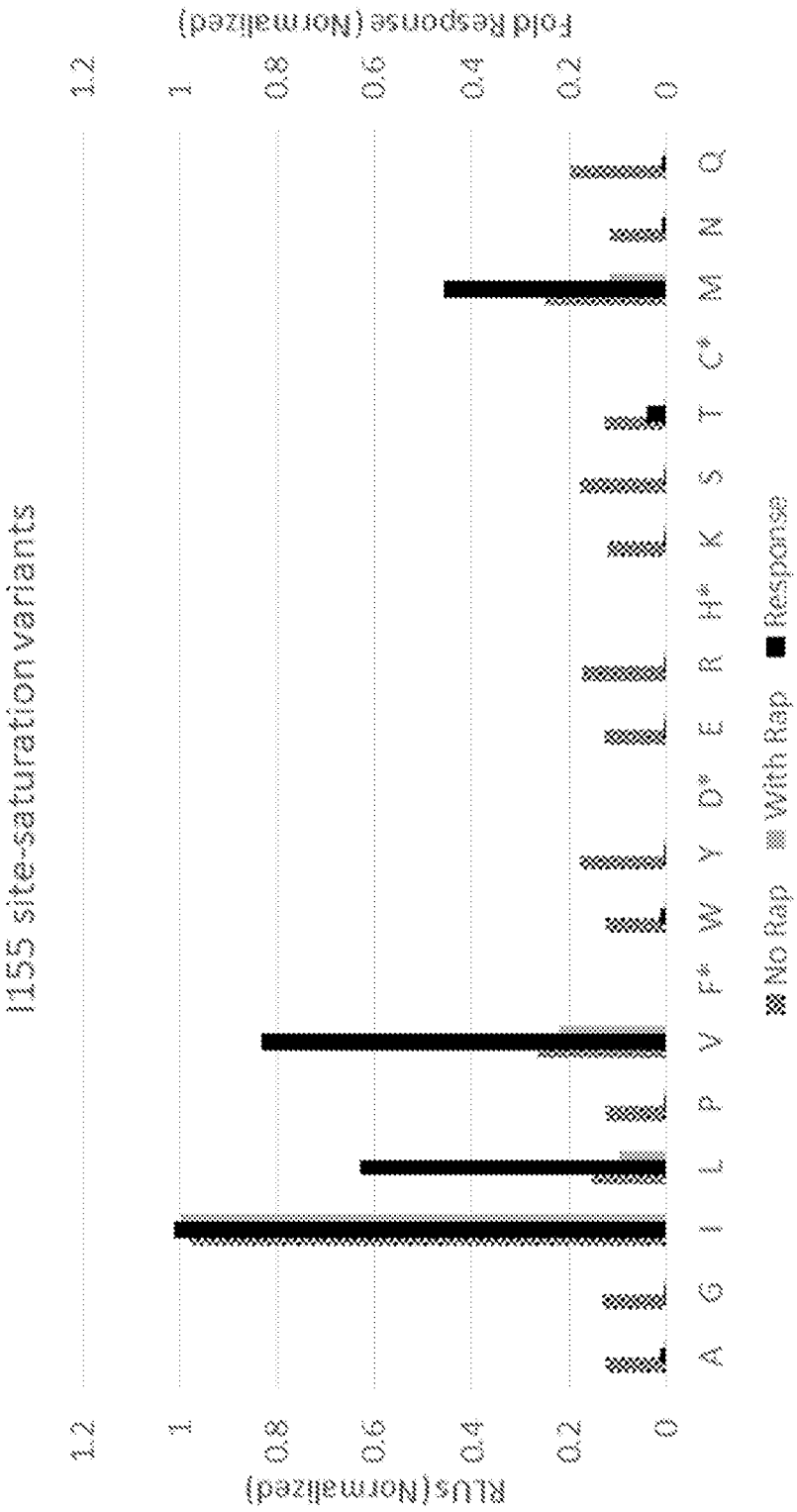

I155 site-saturation variants
RLUs (Normalized)
Fold Response (Normalized)
1.2
1
0.8
0.6
0.4
0.2
0
A G I L P V F* W Y D* E R H* K S T C* M N Q
No Rap
With Rap
Response

FIG. 109

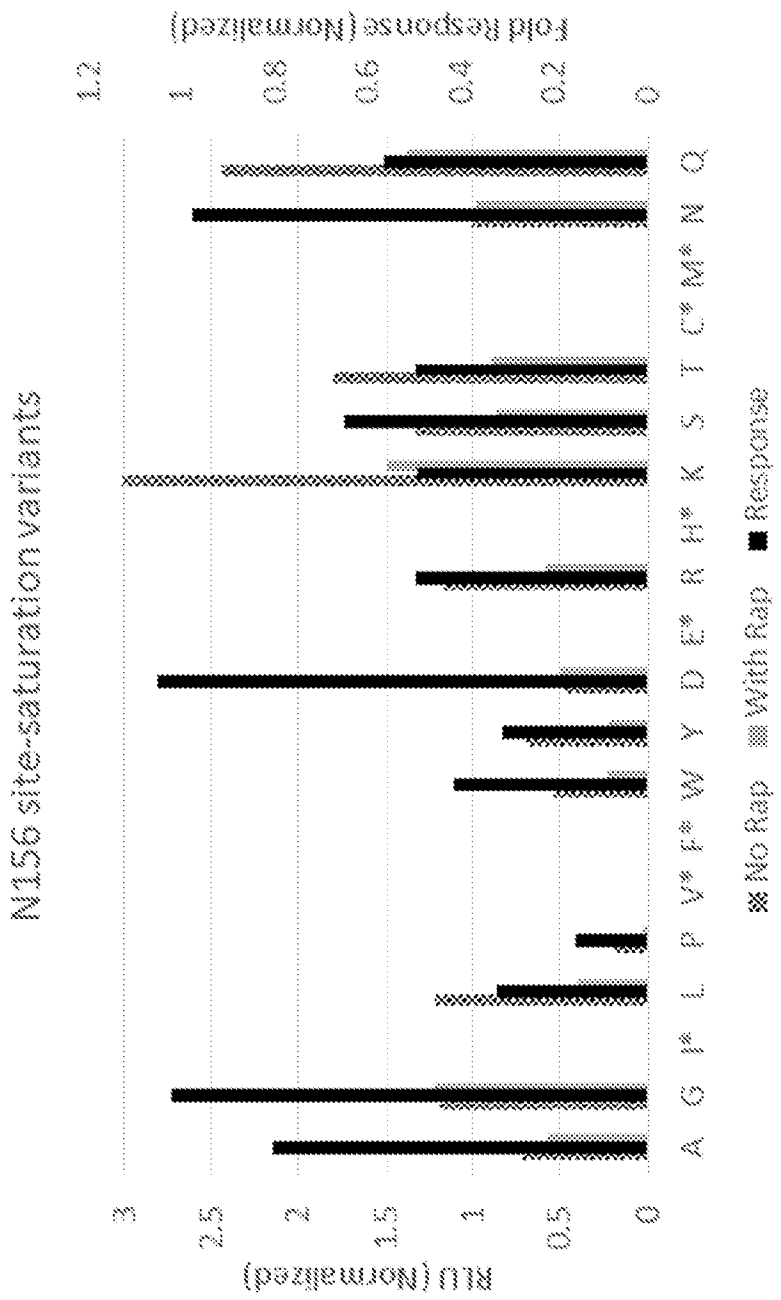

N156 site-saturation variants
RLU (Normalized)
Fold Response (Normalized)
No Rap
With Rap
Response
A G I* L P V* F* W Y D E* R H* K S T C* M* N Q

FIG. 110

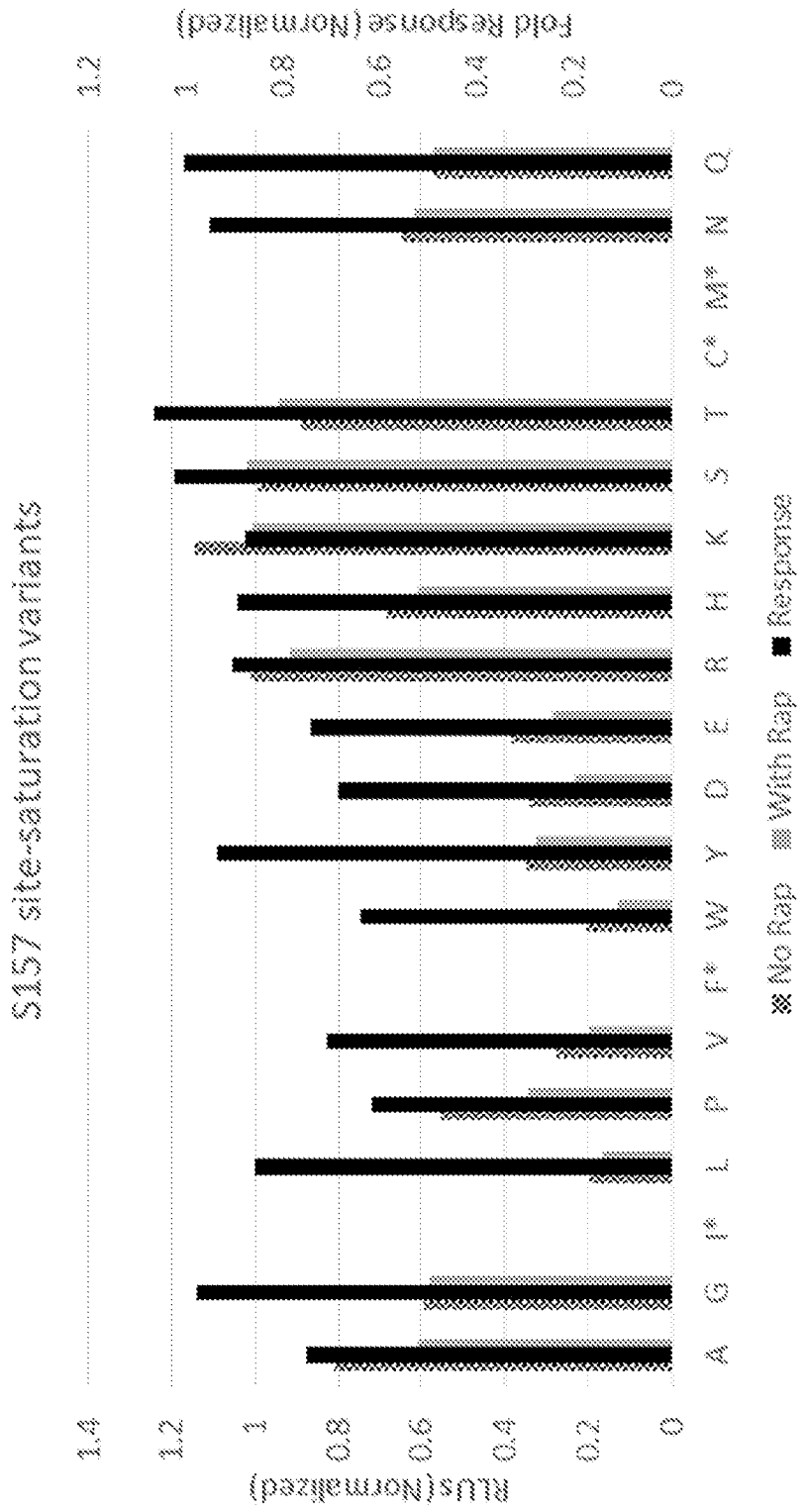

S157 site-saturation variants
Fold Response (Normalized)
RLUs (Normalized)
No Rap
With Rap
Response
A G I* L P V F* W Y D E R H K S T C* M* N Q

FIG. 111

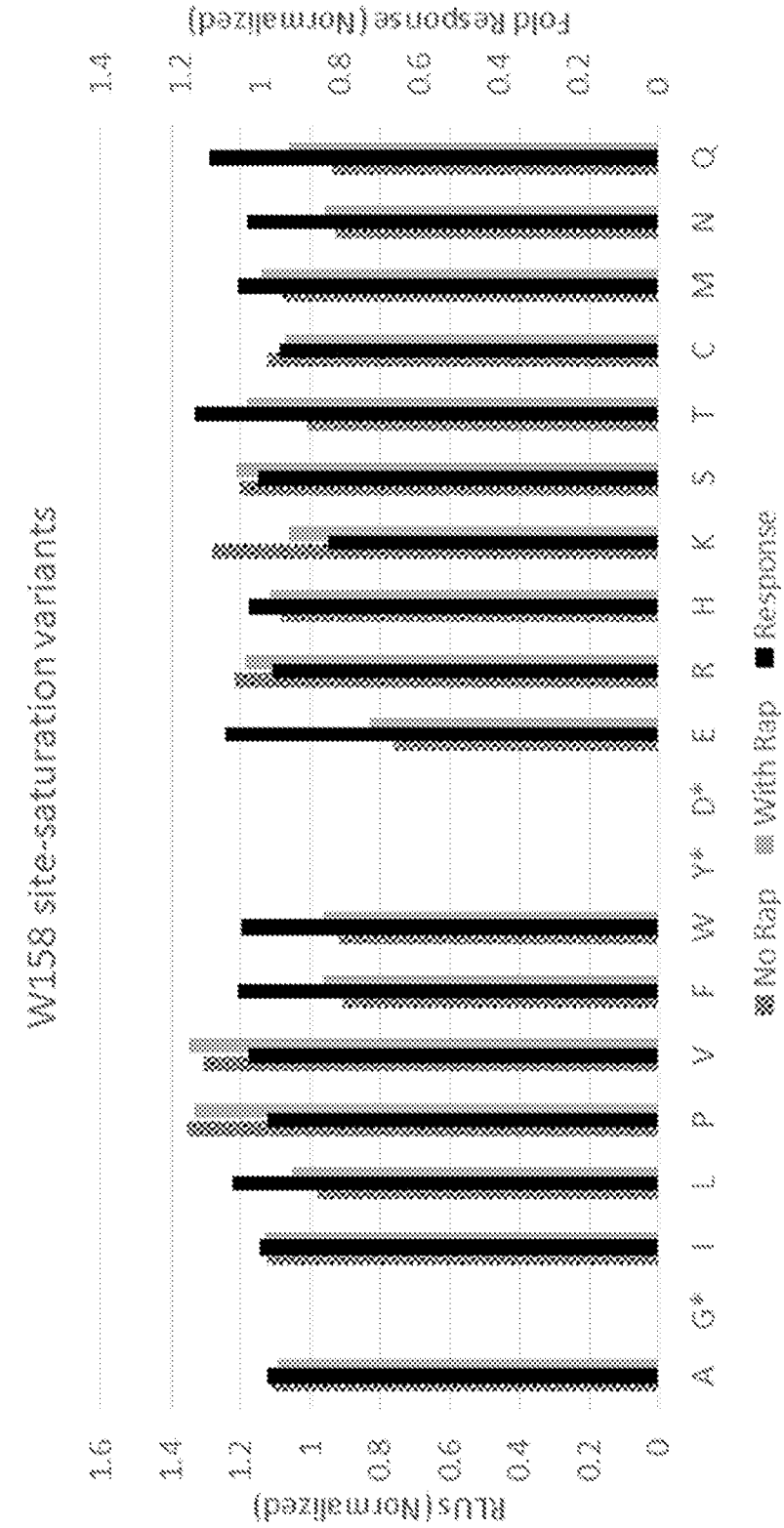

W158 site-saturation variants
RLUs (Normalized)
Fold Response (Normalized)
No Rap
With Rap
Response
A G* I L P V F W Y* D* E R H K S T C M N Q

FIG. 112

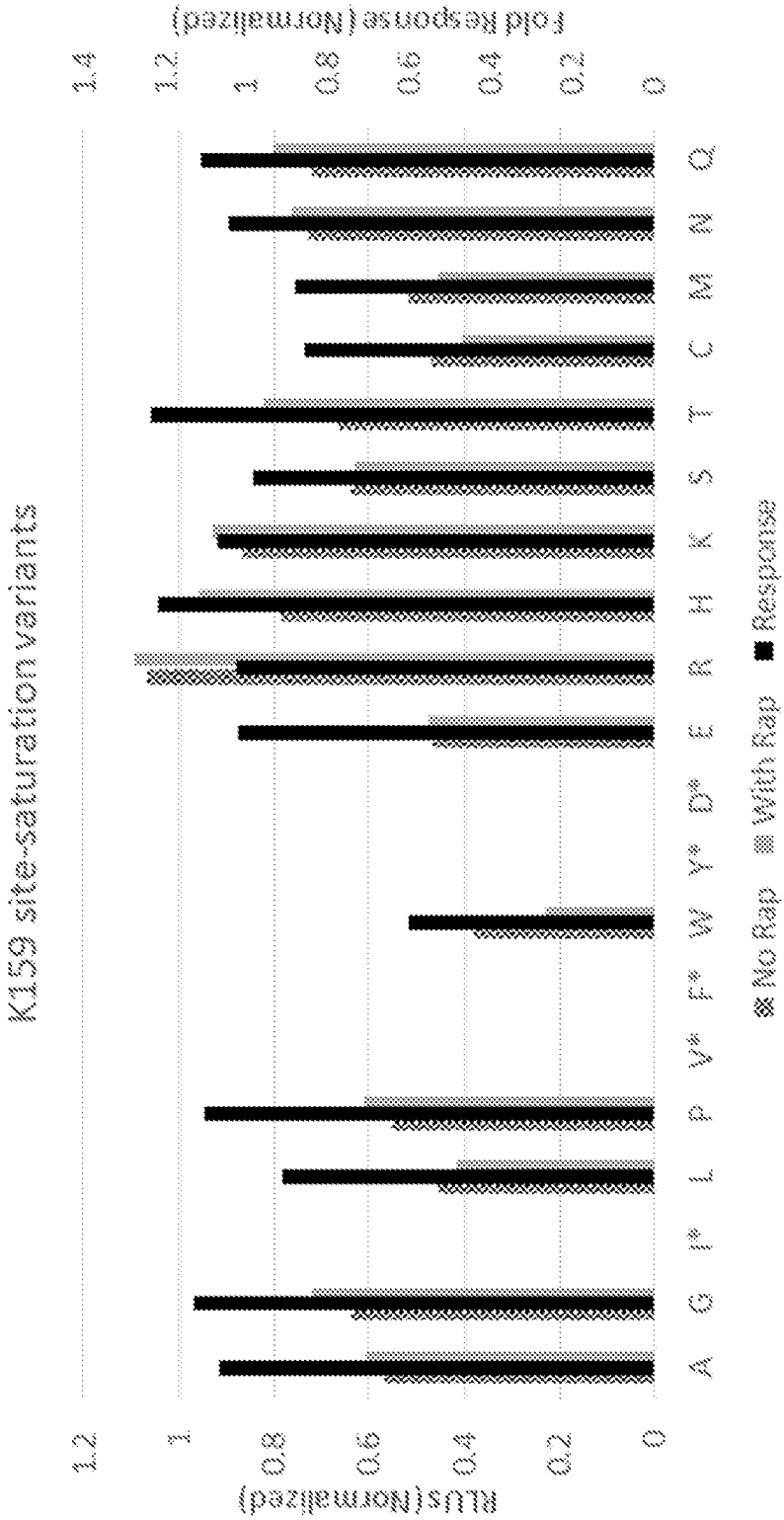

K159 site-saturation variants
Fold Response (Normalized)
RLUs (Normalized)
No Rap
With Rap
Response
A G I* L P V* F* W Y* D* E R H K S T C M N Q

FIG. 113

SmTrip9 on FKBP N-term

FKBP construct				FRB_289 + SmTrip9_FKBP			FRB_86 + SmTrip9_FKBP		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response	With Rap	No Rap	Response
2825	245	GSMLFRVTINS	23	0.05	0.08	0.63	0.04	0.06	0.63
3637	286	SSWKRGSMLFRVTINS	37	0.27	0.17	1.53	0.17	0.21	0.82
3823	435	GSMLFRVTINSWK	231	0.05	0.08	0.66	0.05	0.06	0.82
3826	434	GSMLFRVTIWK	230	0.11	0.25	0.41	0.12	0.48	0.25
3864	492	GSMLFRVTINKWK	313	0.09	0.09	1.02	0.10	0.10	0.95
3865	493	GSMLFRVTIKSWK	314	0.08	0.13	0.61	0.07	0.12	0.57
3866	521	GKMLFRVTINSWK	827	1.00	1.00	1.00	1.00	1.00	1.00
3867	522	GKMLFRVTIWK	269	1.30	4.79	0.27	1.68	10.80	0.16
3868	523	GSMKFRVTINSWK	265	0.04	0.05	0.65	0.01	0.04	0.26
3869	524	GSMKFRVTIWK	271	0.32	0.18	1.80	0.29	0.25	1.17
3870	538	GSMLFRVTIWKVK	283	0.10	0.35	0.28	0.10	0.66	0.15
3871	541	GSMLFRVTIWKWK	286	0.04	0.13	0.33	0.05	0.23	0.21
3872	542	GSMLFRVTIWKK	287	0.08	0.28	0.28	0.09	0.50	0.18

FIG. 114

SmTrip9 on FKBP C-term
Data normalized to 521

FKBP construct				FRB_289 + FKBP_SmTrip9			FRB_86 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response	With Rap	No Rap	Response
3770	245	GSMLFRVTINS	23	0.38	0.48	0.81	0.40	0.53	0.75
3639	286	SSWKRGSMLFRVTINS	37	0.05	0.32	0.14	0.05	0.42	0.13
3828	434	GSMLFRVTIWK	230	1.38	1.55	0.89	2.47	2.42	1.02
3829	435	GSMLFRVTINSWK	231	0.86	0.72	1.20	0.86	0.76	1.13
3936	493	GSMLFRVTIKSWK	314	1.12	1.10	1.02	1.34	1.14	1.18
3937	521	GKMLFRVTINSWK	827	1.00	1.00	1.00	1.00	1.00	1.00
3938	522	GKMLFRVTIWK	269	1.16	2.30	0.50	2.09	3.52	0.59
3939	523	GSMKFRVTINSWK	265	0.02	0.33	0.06	0.01	0.42	0.03
3940	524	GSMKFRVTIWK	271	0.33	0.57	0.57	0.38	0.73	0.52
3942	499	GSMLFRVTIESWK	320	1.45	1.24	1.17	1.33	1.02	1.31

FIG. 115

FKBP construct				FRB_289 + SmTrip9_FKBP			FRB_86 + SmTrip9_FKBP		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response	With Rap	No Rap	Response
3770	245	GSMLFRVTINS	23	0.05	0.10	0.55	0.04	0.15	0.26
3639	286	SSWKRGSMLFRVTINS	37	0.15	0.16	0.94	0.10	0.19	0.51
3828	434	GSMLFRVTIWK	230	0.12	0.35	0.35	0.16	0.64	0.26
3829	435	GSMLFRVTINSWK	231	0.06	0.11	0.58	0.06	0.11	0.49
3936	493	GSMLFRVTIKSWK	314	0.06	0.13	0.44	0.06	0.15	0.36
3937	521	GKMLFRVTINSWK	827	1.00	1.00	1.00	1.00	1.00	1.00
3938	522	GKMLFRVTIWK	269	1.25	5.87	0.21	1.87	9.63	0.19
3939	523	GSMKFRVTINSWK	265	0.03	0.08	0.37	0.01	0.11	0.07
3940	524	GSMKFRVTIWK	271	0.24	0.22	1.06	0.24	0.29	0.81
3942	499	GSMLFRVTIESWK	320	0.14	0.18	0.76	0.11	0.16	0.70

FIG. 116A

FKBP construct				FRB_289 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response
3937	521	GKMLFRVTINSWK	827	1.00	1.00	1.00
4262	671	GKMLFRVTIQKWK	391	1.53	3.60	0.43
4259	668	GKMLFRVTIGKWK	392	1.32	2.30	0.57
4260	727	GKMLFRVTIGRWK	393	1.24	2.47	0.50
4257	669	GKMLFRVTIGNWK	394	1.02	1.35	0.75
4263	674	GKMLFRVTIQNWK	395	0.90	1.50	0.60
4339	702	GKMLFRVTIDKWK	396	0.58	0.67	0.87
4340	703	GKMLFRVTIEKWK	397	0.78	1.65	0.47
4341	730	GKMLFRVTIERWK	398	0.00	0.25	0.01
4342	731	GKMLFRVTIDRWK	399	0.67	0.73	0.92
4343	738	DKMLFRVTIQKWK	400	0.82	0.57	1.44
4344	739	DKMLFRVTIGKWK	401	0.51	0.45	1.13
4345	848	DKMLFRVTIGRWK	402	0.63	0.47	1.33
4346	740	DKMLFRVTIGNWK	403	0.55	0.41	1.35
4347	741	DKMLFRVTIQNWK	404	0.58	0.43	1.33
4348	732	DKMLFRVTIDKWK	405	0.23	0.33	0.70
4349	742	DKMLFRVTIEKWK	406	0.58	0.48	1.21
4350	735	DKMLFRVTIERWK	407	0.29	0.37	0.79
4351	733	DKMLFRVTIDRWK	408	0.32	0.34	0.93

4352	849	EKMLFRVTIQKWK	410	0.80	0.65	1.23
4353	708	EKMLFRVTIGKWK	411	0.78	0.59	1.33
4354	709	EKMLFRVTIGRWK	412	0.69	0.54	1.28
4355	850	EKMLFRVTIGNWK	439	0.62	0.47	1.32
4356	851	EKMLFRVTIQNWK	440	0.60	0.51	1.17
4357	706	EKMLFRVTIDKWK	441	0.32	0.37	0.87
4358	707	EKMLFRVTIEKWK	442	0.59	0.52	1.13
4359	737	EKMLFRVTIERWK	443	0.26	0.40	0.65
4360	736	EKMLFRVTIDRWK	444	0.32	0.37	0.86
4361	760	KKMLFRVTIQKWK	445	1.35	27.31	0.05
4362	852	KKMLFRVTIGKWK	446	1.34	16.45	0.08
4363	853	KKMLFRVTIGRWK	447	1.06	15.19	0.07
4364	854	KKMLFRVTIGNWK	448	1.48	12.45	0.12
4365	855	KKMLFRVTIQNWK	449	1.78	18.75	0.09
4366	856	KKMLFRVTIDKWK	450	1.23	4.82	0.26
4367	857	KKMLFRVTIEKWK	451	1.24	14.92	0.08
4368	858	KKMLFRVTIERWK	452	0.45	6.05	0.07
4369	859	KKMLFRVTIDRWK	453	0.98	4.05	0.24
4370	860	RKMLFRVTIQKWK	454	1.30	27.15	0.05
4371	861	RKMLFRVTIGKWK	455	1.50	20.53	0.07
4372	862	RKMLFRVTIGRWK	456	1.20	19.17	0.06
4373	863	RKMLFRVTIGNWK	457	1.35	13.87	0.10
4374	864	RKMLFRVTIQNWK	458	1.27	18.12	0.07
4375	865	RKMLFRVTIDKWK	459	0.85	4.61	0.18
4376	866	RKMLFRVTIEKWK	460	0.80	14.05	0.06

FIG. 116B

4377	867	RKMLFRVTIERWK	461	0.36	6.68	0.05
4378	868	RKMLFRVTIDRWK	462	0.82	4.84	0.17
4291	656	EQMLFRVTINSWK	463	0.49	0.41	1.18
4292	869	SRMLFRVTINSWK	464	0.94	1.50	0.63
4293	533	GEMLFRVTINSWK	465	0.57	0.39	1.44
4331	690	GKMKFRVTINSWK	466	0.12	0.43	0.27
4332	678	GKMLFRVKINSWK	467	0.15	0.50	0.30
4333	679	GKMLFRVRINSWK	468	0.26	0.65	0.40
4334	681	GKMLFRVDINSWK	469	0.00	0.39	0.00

FIG. 116C

FIG. 117

FKBP construct				FRB_289 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response
3937	521	GKMLFRVTINSWK	827	1.00	1.00	1.00
4339	702	GKMLFRVTIDKWK	396	0.90	0.85	1.05
4340	703	GKMLFRVTIEKWK	397	0.96	1.71	0.56
4341	730	GKMLFRVTIERWK	398	0.00	0.32	0.01
4342	731	GKMLFRVTIDRWK	399	0.74	0.75	0.99
4343	738	DKMLFRVTIQKWK	400	1.11	0.64	1.72
4344	739	DKMLFRVTIGKWK	401	1.02	0.54	1.89
4345	848	DKMLFRVTIGRWK	402	1.00	0.61	1.64
4346	740	DKMLFRVTIGNWK	403	0.75	0.43	1.74
4347	741	DKMLFRVTIQNWK	404	0.74	0.50	1.48
4348	732	DKMLFRVTIDKWK	405	0.38	0.37	1.02
4349	742	DKMLFRVTIEKWK	406	0.63	0.44	1.42
4350	735	DKMLFRVTIERWK	407	0.27	0.38	0.72
4351	733	DKMLFRVTIDRWK	408	0.34	0.35	0.98
4352	849	EKMLFRVTIQKWK	410	1.11	0.73	1.51
4353	708	EKMLFRVTIGKWK	411	1.07	0.64	1.67
4354	709	EKMLFRVTIGRWK	412	0.78	0.57	1.36
4355	850	EKMLFRVTIGNWK	439	0.76	0.47	1.62
4356	851	EKMLFRVTIQNWK	440	0.74	0.48	1.56
4357	706	EKMLFRVTIDKWK	441	0.41	0.36	1.14
4358	707	EKMLFRVTIEKWK	442	0.59	0.47	1.25
4359	737	EKMLFRVTIERWK	443	0.24	0.35	0.69
4360	736	EKMLFRVTIDRWK	444	0.41	0.41	1.02
4185	663	GKMLFRVTIDSWK	470	0.53	0.56	0.95
4196	743	GKMLFRVTINKWK	471	1.15	1.24	0.92

FIG. 118

FKBP construct				FRB_289 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response
3937	521	GKMLFRVTINSWK	827	1.00	1.00	1.00
4339	702	GKMLFRVTIDKWK	396	0.62	0.71	0.88
4340	703	GKMLFRVTIEKWK	397	0.96	1.47	0.65
4341	730	GKMLFRVTIERWK	398	0.00	0.42	0.00
4342	731	GKMLFRVTIDRWK	399	0.62	0.74	0.83
4343	738	DKMLFRVTIQKWK	400	0.84	0.63	1.34
4344	739	DKMLFRVTIGKWK	401	0.66	0.61	1.09
4345	848	DKMLFRVTIGRWK	402	0.71	0.65	1.09
4346	740	DKMLFRVTIGNWK	403	0.54	0.52	1.03
4347	741	DKMLFRVTIQNWK	404	0.56	0.52	1.07
4348	732	DKMLFRVTIDKWK	405	0.27	0.45	0.61
4349	742	DKMLFRVTIEKWK	406	0.57	0.52	1.10
4350	735	DKMLFRVTIERWK	407	0.25	0.45	0.56
4351	733	DKMLFRVTIDRWK	408	0.24	0.39	0.62
4352	849	EKMLFRVTIQKWK	410	0.81	0.66	1.24
4353	708	EKMLFRVTIGKWK	411	0.70	0.62	1.13
4354	709	EKMLFRVTIGRWK	412	0.67	0.59	1.14
4355	850	EKMLFRVTIGNWK	439	0.59	0.50	1.18
4356	851	EKMLFRVTIQNWK	440	0.62	0.57	1.08
4357	706	EKMLFRVTIDKWK	441	0.30	0.40	0.74
4358	707	EKMLFRVTIEKWK	442	0.50	0.55	0.91
4359	737	EKMLFRVTIERWK	443	0.23	0.41	0.56
4360	736	EKMLFRVTIDRWK	444	0.30	0.44	0.68
4185	663	GKMLFRVTIDSWK	470	0.46	0.58	0.79
4196	743	GKMLFRVTINKWK	471	0.92	1.03	0.89

FIG. 119

FKBP construct				FRB_289 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response
3937	521	GKMLFRVTINSWK	827	1.00	1.00	1.00
4343	738	DKMLFRVTIQKWK	400	0.98	0.56	1.76
4349	742	DKMLFRVTIEKWK	406	0.58	0.45	1.31
4196	743	GKMLFRVTINKWK	471	1.21	1.23	0.98
4344	739	DKMLFRVTIGKWK	401	0.63	0.38	1.66
4409	714	EKMLFKVTIQKWK	472	0.68	0.35	1.93
4410	870	EKMLFTVTIQKWK	473	0.42	0.25	1.69
4411	871	EKMLFKVTIDKWK	474	0.14	0.24	0.60
4412	872	EKMLFTVTIDKWK	475	0.11	0.20	0.55
4413	873	EKMLFKVTIGRWK	476	0.60	0.31	1.93
4415	744	DKMLFKVTIQKWK	477	0.57	0.28	2.03
4416	745	DKMLFTVTIQKWK	478	0.46	0.25	1.82
4417	874	DKMLFKVTIDKWK	479	0.16	0.17	0.96
4418	875	DKMLFTVTIDKWK	480	0.12	0.16	0.76
4419	876	GKMLFKVTIEKWK	481	0.72	0.67	1.07
4420	877	GKMLFTVTIEKWK	482	0.48	0.35	1.38
4421	748	DKMLFKVTIGKWK	483	0.57	0.26	2.17
4422	749	DKMLFTVTIGKWK	484	0.33	0.20	1.69
4423	878	DKMLFKVTIGNWK	485	0.39	0.23	1.70
4424	879	DKMLFKVTIQNWK	486	0.32	0.25	1.28
4425	781	GKMLFKVTINKWK	487	0.77	0.47	1.61
4426	782	GKMLFTVTINKWK	488	0.59	0.23	2.52

FIG. 120

FKBP construct				FRB_289 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response
3937	521	GKMLFRVTINSWK	827	1.00	1.00	1.00
3770	245	GSMLFRVTINS	288	0.53	0.32	1.67
4343	738	DKMLFRVTIQKWK	400	0.98	0.53	1.85
4349	742	DKMLFRVTIEKWK	406	0.77	0.37	2.08
4196	743	GKMLFRVTINKWK	471	1.14	1.37	0.83
4344	739	DKMLFRVTIGKWK	401	0.85	0.37	2.33
4409	714	EKMLFKVTIQKWK	472	0.69	0.28	2.49
4410	870	EKMLFTVTIQKWK	473	0.60	0.19	3.18
4411	871	EKMLFKVTIDKWK	474	0.19	0.13	1.46
4412	872	EKMLFTVTIDKWK	475	0.14	0.11	1.27
4413	873	EKMLFKVTIGRWK	476	0.64	0.24	2.72
4415	744	DKMLFKVTIQKWK	477	0.67	0.21	3.20
4416	745	DKMLFTVTIQKWK	478	0.53	0.13	4.06
4417	874	DKMLFKVTIDKWK	479	0.20	0.10	1.98
4418	875	DKMLFTVTIDKWK	480	0.14	0.09	1.53
4419	876	GKMLFKVTIEKWK	481	0.78	0.70	1.11
4420	877	GKMLFTVTIEKWK	482	0.63	0.27	2.35
4421	748	DKMLFKVTIGKWK	483	0.59	0.19	3.02
4422	749	DKMLFTVTIGKWK	484	0.41	0.11	3.77
4423	878	DKMLFKVTIGNWK	485	0.45	0.14	3.29
4424	879	DKMLFKVTIQNWK	486	0.42	0.17	2.52
4425	781	GKMLFKVTINKWK	487	0.67	0.40	1.69
4426	782	GKMLFTVTINKWK	488	0.59	0.17	3.51

FIG. 121

FKBP construct				FRB_289 + FKBP_SmTrip9			FRB_86 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response	With Rap	No Rap	Response
3937	521	GKMLFRVTINSWK	827	1.00 +/- 0.12	1.00 +/- 0.16	1.00 +/- 0.10	1.09 +/- 0.09	0.64 +/- 0.12	1.69 +/- 0.07
3770	245	GSMLFRVTINS	288	0.61 +/- 0.15	0.26 +/- 0.10	2.45 +/- 0.33	0.51 +/- 0.05	0.11 +/- 0.03	4.53 +/- 0.13
4416	745	DKMLFTVTIQKWK	478	0.56 +/- 0.10	0.06 +/- 0.01	8.72 +/- 0.65	0.78 +/- 0.11	0.08 +/- 0.02	9.99 +/- 0.90
4422	749	DKMLFTVTIGKWK	484	0.40 +/- 0.03	0.04 +/- 0.00	9.59 +/- 0.39	0.63 +/- 0.03	0.06 +/- 0.01	10.84 +/- 0.54
4460	752	DKMLFKVTIEKWK	489	0.43 +/- 0.03	0.08 +/- 0.01	5.52 +/- 0.31	0.51 +/- 0.01	0.08 +/- 0.01	6.72 +/- 0.37
4461	753	DKMLFTVTIEKWK	490	0.49 +/- 0.06	0.06 +/- 0.01	8.80 +/- 1.20	0.66 +/- 0.08	0.05 +/- 0.01	12.05 +/- 0.44
4466	750	DKLLFKVTIGKWK	491	0.47 +/- 0.01	0.08 +/- 0.01	5.71 +/- 0.54	0.71 +/- 0.02	0.09 +/- 0.01	7.42 +/- 0.21
4467	786	DKMLFTVTINKWK	492	0.26 +/- 0.02	0.03 +/- 0.00	8.97 +/- 1.17	0.35 +/- 0.01	0.03 +/- 0.00	10.54 +/- 0.38
4468	756	DKLLFTVTIQKWK	493	0.50 +/- 0.10	0.07 +/- 0.01	7.23 +/- 0.70	0.71 +/- 0.10	0.07 +/- 0.01	10.71 +/- 0.36
4472	757	DKLLFTVTIQKYK	494	0.43 +/- 0.04	0.05 +/- 0.01	8.28 +/- 0.66	0.60 +/- 0.06	0.05 +/- 0.01	11.68 +/- 0.77
4542	758	DKLLFTVTIEKWK	495	0.37 +/- 0.03	0.05 +/- 0.01	6.99 +/- 1.24	0.50 +/- 0.02	0.05 +/- 0.01	10.21 +/- 0.56
4543	759	DKLLFTVTIEKYK	496	0.38 +/- 0.04	0.04 +/- 0.00	9.36 +/- 0.47	0.48 +/- 0.04	0.05 +/- 0.00	9.63 +/- 0.24
4544	793	DKLLFTVTIGKWK	497	0.50 +/- 0.09	0.05 +/- 0.01	9.63 +/- 0.94	0.78 +/- 0.10	0.06 +/- 0.01	12.93 +/- 0.53
4545	794	DKLLFTVTIGKYK	498	0.26 +/- 0.03	0.02 +/- 0.00	11.19 +/- 1.09	0.38 +/- 0.04	0.03 +/- 0.01	12.97 +/- 0.40
4546	799	DKLLFTVTINKWK	499	0.28 +/- 0.05	0.03 +/- 0.00	9.95 +/- 2.42	0.36 +/- 0.04	0.03 +/- 0.00	11.60 +/- 1.03
4547	800	DKLLFTVTINKYK	500	0.17 +/- 0.01	0.02 +/- 0.01	7.95 +/- 2.83	0.23 +/- 0.01	0.02 +/- 0.00	10.95 +/- 0.86

FIG. 122A

FKBP construct				FRB_289 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response
3770	245	GSMLFRVTINS	288	0.52 +/- 0.06	0.23 +/- 0.07	2.30 +/- 0.47
3937	521	GKMLFRVTINSWK	827	1.00 +/- 0.00	1.00 +/- 0.00	1.00 +/- 0.00
4455	780	GKMLFRVTINS	501	0.58 +/- 0.04	0.37 +/- 0.10	1.65 +/- 0.52
4416	745	DKMLFTVTIQKWK	478	0.63 +/- 0.19	0.28 +/- 0.32	6.34 +/- 6.31
4456	765	DKMLFTVTIQK	502	0.30 +/- 0.06	0.04 +/- 0.00	7.33 +/- 1.56
4415	744	DKMLFKVTIQKWK	477	0.60 +/- 0.05	0.16 +/- 0.03	3.93 +/- 1.03
4436	779	DKMLFKVTIQK	503	0.40 +/- 0.06	0.11 +/- 0.07	5.19 +/- 3.57
4544	793	DKLLFTVTIGKWK	497	0.42 +/- 0.00	0.18 +/- 0.20	6.36 +/- 6.96
4618	820	DKLLFTVTIGK	504	0.19 +/- 0.03	0.03 +/- 0.00	7.44 +/- 0.03
4422	749	DKMLFTVTIGKWK	484	0.49 +/- 0.04	0.07 +/- 0.01	7.21 +/- 1.89
4619	819	DKMLFTVTIGK	505	0.22 +/- 0.02	0.05 +/- 0.04	7.15 +/- 4.92
4461	753	DKMLFTVTIEKWK	490	0.52 +/- 0.02	0.20 +/- 0.22	7.29 +/- 7.94
4620	822	DKMLFTVTIEK	506	0.21 +/- 0.03	0.03 +/- 0.00	7.54 +/- 0.96
4542	758	DKLLFTVTIEKWK	495	0.39 +/- 0.03	0.07 +/- 0.00	5.34 +/- 0.80
4543	759	DKLLFTVTIEKYK	496	0.45 +/- 0.04	0.09 +/- 0.06	7.01 +/- 4.69
4621	821	DKLLFTVTIEK	507	0.21 +/- 0.00	0.10 +/- 0.11	5.94 +/- 6.30

FIG. 122B

FKBP construct				FRB_86 + FKBP_SmTrip9			FRB_229 + FKBP_SmTrip9		
ATG#	Pep ID	Pep Sequence	SEQ ID NO.	With Rap	No Rap	Response	With Rap	No Rap	Response
3770	245	GSMLFRVTINS	288	0.37 +/- 0.01	0.09 +/- 0.01	3.90 +/- 0.45	0.22 +/- 0.01	0.04 +/- 0.01	5.39 +/- 0.64
3937	521	GKMLFRVTINSWK	827	0.77 +/- 0.02	0.52 +/- 0.08	1.49 +/- 0.19	0.64 +/- 0.06	0.29 +/- 0.02	2.22 +/- 0.08
4455	780	GKMLFRVTINS	501	0.45 +/- 0.02	0.20 +/- 0.02	2.27 +/- 0.16	0.31 +/- 0.01	0.08 +/- 0.00	3.75 +/- 0.02
4416	745	DKMLFTVTIQKWK	478	0.56 +/- 0.01	0.09 +/- 0.04	6.94 +/- 3.13	0.53 +/- 0.11	0.06 +/- 0.04	10.92 +/- 4.89
4456	765	DKMLFTVTIQK	502	0.36 +/- 0.06	0.03 +/- 0.00	12.46 +/- 3.06	0.25 +/- 0.05	0.02 +/- 0.00	15.12 +/- 0.57
4415	744	DKMLFKVTIQKWK	477	0.56 +/- 0.03	0.12 +/- 0.05	5.19 +/- 2.29	0.41 +/- 0.02	0.06 +/- 0.01	6.88 +/- 0.65
4436	779	DKMLFKVTIQK	503	0.43 +/- 0.01	0.06 +/- 0.01	6.95 +/- 0.73	0.24 +/- 0.00	0.03 +/- 0.00	7.49 +/- 0.81
4544	793	DKLLFTVTIGKWK	497	0.48 +/- 0.01	0.07 +/- 0.03	7.10 +/- 3.14	0.42 +/- 0.03	0.04 +/- 0.02	12.75 +/- 6.40
4618	820	DKLLFTVTIGK	504	0.22 +/- 0.01	0.02 +/- 0.00	11.31 +/- 1.60	0.17 +/- 0.04	0.01 +/- 0.00	14.91 +/- 1.83
4422	749	DKMLFTVTIGKWK	484	0.58 +/- 0.01	0.07 +/- 0.04	9.94 +/- 4.88	0.47 +/- 0.01	0.03 +/- 0.00	16.98 +/- 0.78
4619	819	DKMLFTVTIGK	505	0.27 +/- 0.03	0.03 +/- 0.00	8.91 +/- 0.28	0.18 +/- 0.01	0.02 +/- 0.00	12.57 +/- 2.66
4461	753	DKMLFTVTIEKWK	490	0.43 +/- 0.03	0.07 +/- 0.03	6.47 +/- 1.99	0.43 +/- 0.01	0.05 +/- 0.03	11.88 +/- 6.85
4620	822	DKMLFTVTIEK	506	0.18 +/- 0.02	0.02 +/- 0.01	10.41 +/- 1.84	0.14 +/- 0.01	0.01 +/- 0.00	14.54 +/- 0.17
4542	758	DKLLFTVTIEKWK	495	0.39 +/- 0.09	0.06 +/- 0.02	7.91 +/- 4.43	0.36 +/- 0.03	0.03 +/- 0.00	13.82 +/- 0.02
4543	759	DKLLFTVTIEKYK	496	0.44 +/- 0.00	0.06 +/- 0.02	7.60 +/- 2.22	0.35 +/- 0.03	0.03 +/- 0.00	13.69 +/- 0.63
4621	821	DKLLFTVTIEK	507	0.19 +/- 0.00	0.03 +/- 0.01	7.30 +/- 2.11	0.17 +/- 0.01	0.02 +/- 0.01	11.52 +/- 5.27

FIG. 123

Biochemical analysis of synthetic SmTrip9 variants					
Pep ID	Sequence	SEQ ID NO.	SmTrip9 Kd (uM)	VS-HiBiT Kd (uM)	Bmax (RLU)
521	*GKMLFRVTINSWK	827	0.06	1.21	3.25E+07
627	*DKMLFRVTINSWK	508	0.29	11.76	1.29E+07
628	*EKMLFRVTINSWK	509	0.10	0.76	2.80E+07
629	*RKMLFRVTINSWK	510	0.02	1.23	2.99E+07
630	*KKMLFRVTINSWK	511	0.02	1.64	2.98E+07
631	*HKMLFRVTINSWK	512	0.03	ND	2.65E+07
632	*GLMLFRVTINSWK	513	0.08	ND	1.20E+07
633	*GQMLFRVTINSWK	514	0.09	ND	1.88E+07
634	*GTMLFRVTINSWK	515	0.08	ND	1.27E+07
635	*GKLLFRVTINSWK	516	0.06	1.54	2.81E+07
636	*GKMLFKVTINSWK	517	0.17	2.10	3.10E+07
637	*GKMLFRVTIQSWK	518	0.02	1.78	3.13E+07
638	*GKMLFRVTIDSWK	519	0.26	0.96	2.95E+07
639	*GKMLFRVTIGSWK	520	0.07	1.71	2.88E+07
640	*GKMLFRVTINTWK	521	0.13	ND	2.91E+07
641	*GKMLFRVTINNWK	522	0.24	ND	2.87E+07
642	*GKMLFRVTINQWK	523	0.14	ND	2.62E+07
643	*GKMLFRVTINPWK	524	0.08	ND	3.00E+07
644	*GKMLFRVTINKWK	525	0.09	ND	3.05E+07
645	*GKMLFRVTINSWQ	526	0.26	ND	4.96E+06
646	*GKMLFRVTINSWN	527	1.46	ND	8.61E+06
647	*GKMLFRVTINSWT	528	0.18	ND	2.47E+07
648	*GKMLFRVTINSWH	529	0.19	ND	2.64E+07
649	*GKMLFRVTINSWP	530	0.08	ND	2.97E+07
650	*GKMLFRVTINSWR	531	0.09	ND	2.53E+07
*Terminus unblocked ND = Not Determined					

FIG. 124

Biochemical analysis of synthetic SmTrip9 variants					
Pep ID	Sequence	SEQ ID NO.	SmTrip9 Kd (uM)	VS-HiBiT Kd (uM)	Bmax (RLU)
521	GKMLFRVTINSWK	827	0.18	2.11	2.25E+07
677	GKMKFRVTIDSWK	532	1.47	1.41	3.16E+07
678	GKMLFRVKINSWK	467	0.74	1.14	3.01E+07
679	GKMLFRVRINSWK	468	0.36	1.13	2.86E+07
680	GKMLFRVEINSWK	533	30.23	ND	3.72E+07
681	GKMLFRVDINSWK	469	3.21	ND	4.74E+06
682	GKMLFRVQINSWK	534	0.64	ND	1.91E+07
683	GKMKFRVKINSWK	535	6.72	ND	2.38E+07
684	GKMKFRVRINSWK	536	2.23	ND	2.09E+07
685	GKMKFRVEINSWK	537	8.63	ND	3.62E+07
686	GKMKFRVDINSWK	538	26.76	ND	2.62E+07
687	GKMKFRVQINSWK	539	2.52	ND	2.81E+07
688	GKMKFRVNINSWK	540	9.04	ND	2.65E+07
689	GKMKFRVSINSWK	541	2.97	ND	3.11E+07
690	GKMKFRVTINSWK	466	0.68	1.59	3.21E+07
ND = Not Determined					

FIG. 125

Biochemical analysis of synthetic SmTrip9 variants					
Pep ID	**Sequence**	**SEQ ID NO.**	**SmTrip9 Kd (uM)**	**VS-HiBiT Kd (uM)**	**Bmax (RLU)**
521	GKMLFRVTINSWK	827	0.51	ND	2.58E+07
613	GKMLFRVNINSWK	542	5.25	ND	3.23E+07
614	GKMLFRVSINSWK	543	2.58	ND	3.66E+07
615	GKMLFRVWINSWK	544	0.35	ND	1.75E+07
616	GKMSFRVTINSWK	545	1.23	ND	2.31E+07
617	GKMWFRVTINSWK	546	0.32	ND	1.20E+07
618	GKMNFRVTINSWK	547	2.69	ND	1.10E+07
619	GSMLFRVTINSYK	548	1.51	ND	2.57E+07
620	GKMLFRVTINSYK	549	0.19	ND	2.66E+07
621	GKMLFRVTIKSWK	550	0.04	1.12	3.11E+07
622	GKMLFRVTIESWK	551	7.24	33.48	3.41E+07
716	GKMKFRVTIQSWK	552	0.29	ND	3.36E+07
717	GKMKFRVTIESWK	553	1.14	ND	3.26E+07
718	GKMKFRVTIKSWK	554	0.38	ND	3.36E+07
719	GKMKFRVTIRSWK	555	0.37	ND	3.27E+07
ND = Not Determined					

FIG. 126

Biochemical analysis of synthetic SmTrip9 variants					
Pep ID	Sequence	SEQ ID NO.	SmTrip9 Kd (uM)	VS-HiBiT Kd (uM)	Bmax (RLU)
521	GKMLFRVTINSWK	827	0.17	4.20	2.51E+07
651	RLMLFRVTINSWK	556	0.04	ND	2.27E+07
652	RQMLFRVTINSWK	557	0.03	ND	2.86E+07
653	KLMLFRVTINSWK	558	0.16	ND	2.78E+07
654	KQMLFRVTINSWK	559	0.06	ND	2.86E+07
655	ELMLFRVTINSWK	560	15.86	ND	7.15E+06
656	EQMLFRVTINSWK	463	6.26	ND	1.52E+07
657	DLMLFRVTINSWK	561	0.07	ND	6.26E+05
658	DQMLFRVTINSWK	562	4.38	ND	2.38E+06
659	DKMLFRVTINSWK	563	0.58	31.91	3.65E+06
660	EKMLFRVTINSWK	564	0.19	9.75	1.83E+07
661	RKMLFRVTINSWK	565	0.01	4.34	3.19E+07
662	KKMLFRVTINSWK	566	0.02	2.72	3.09E+07
663	GKMLFRVTIDSWK	470	0.25	2.40	3.21E+07
665	GKMLFRVTIGSWK	567	0.14	6.16	2.81E+07
667	GKMLFRVTINKWK	568	0.07	2.98	3.63E+07
668	GKMLFRVTIGKWK	392	0.06	2.78	3.06E+07
669	GKMLFRVTIGNWK	394	0.38	3.06	2.82E+07
670	GKMLFRVTISKWK	569	0.03	3.96	3.11E+07
671	GKMLFRVTIQKWK	570	0.02	3.15	3.67E+07
672	GKMLFRVTITKWK	571	0.01	3.19	3.18E+07
673	GKMLFRVTIKKWK	572	0.04	2.85	3.53E+07
675	GKMLFKVTINSWK	573	0.17	8.06	2.61E+07
676	RLMLFRVTIGKWK	574	0.02	4.09	2.92E+07
ND = Not Determined					

FIG. 127

Biochemical analysis of synthetic SmTrip9 variants					
Pep ID	Sequence	SEQ ID NO.	SmTrip9 Kd (uM)	VS-HiBiT Kd (uM)	Bmax (RLU)
521	GKMLFRVTINSWK	827	0.21	2.18	2.82E+07
659	DKMLFRVTINSWK	563	0.76	24.39	5.24E+06
660	EKMLFRVTINSWK	564	0.30	4.38	2.35E+07
667	GKMLFRVTINKWK	568	0.16	0.79	4.07E+07
668	GKMLFRVTIGKWK	392	0.13	0.75	3.55E+07
671	GKMLFRVTIQKWK	570	0.06	0.65	4.21E+07
675	GKMLFKVTINSWK	573	0.47	9.71	1.65E+07
620	GKMLFRVTINSYK	549	0.12	2.27	2.82E+07
622	GKMLFRVTIESWK	551	2.88	1.73	3.16E+07
701	GKMLFRVTINRWK	575	0.12	1.01	3.69E+07
703	GKMLFRVTIEKWK	397	0.11	1.18	3.61E+07
707	EKMLFRVTIEKWK	442	1.12	0.90	3.23E+07
708	EKMLFRVTIGKWK	411	0.28	1.18	3.27E+07
709	EKMLFRVTIGRWK	412	0.18	1.61	2.99E+07
710	EKMLFTVTIGKWK	576	1.29	5.24	1.94E+07
711	EKLLFTVTIGKWK	577	0.50	5.52	1.40E+07
712	EKMLFTVTIGRWK	578	0.37	6.09	1.66E+07
714	EKMLFKVTIQKWK	472	0.22	1.18	3.34E+07
720	EKMLFTVTIEKWK	579	0.36	4.29	2.19E+07
722	DKMLFRVTIESWK	580	17.79	10.61	2.23E+07
726	EKLLFRVTIGKYK	581	0.64	1.49	2.63E+07
ND = Not Determined					

FIG. 128

Biochemical analysis of synthetic SmTrip9 variants					
Pep ID	Sequence	SEQ ID NO.	SmTrip9 Kd (uM)	VS-HiBiT Kd (uM)	Bmax (RLU)
521	GKMLFRVTINSWK	827	0.20	3.56	3.03E+07
738	DKMLFRVTIQKWK	400	0.23	3.06	3.64E+07
739	DKMLFRVTIGKWK	401	0.20	2.28	3.41E+07
740	DKMLFRVTIGNWK	403	3.65	59.23	1.93E+07
741	DKMLFRVTIQNWK	404	12.58	33.67	1.01E+07
742	DKMLFRVTIEKWK	406	0.75	1.44	3.55E+07
743	GKMLFRVTINKWK	471	0.10	1.33	3.98E+07
744	DKMLFKVTIQKWK	477	1.13	4.65	2.74E+07
745	DKMLFTVTIQKWK	478	0.16	7.52	2.30E+07
746	DKLLFKVTIQKWK	582	1.10	6.23	1.53E+07
747	DKLLFKVTIQKYK	583	0.10	4.25	2.76E+07
748	DKMLFKVTIGKWK	483	0.44	2.24	3.43E+07
749	DKMLFTVTIGKWK	484	1.70	8.26	1.92E+07
750	DKLLFKVTIGKWK	491	0.35	3.34	2.32E+07
751	DKLLFKVTIGKYK	584	0.61	3.28	2.55E+07
752	DKMLFKVTIEKWK	489	0.32	1.58	3.26E+07
753	DKMLFTVTIEKWK	490	0.12	7.59	2.01E+07
754	DKLLFKVTIEKWK	585	0.75	1.52	2.35E+07
755	DKLLFKVTIEKYK	586	1.45	1.20	2.70E+07
ND = Not Determined					

FIG. 129

Biochemical analysis of synthetic SmTrip9 variants					
Pep ID	Sequence	SEQ ID NO.	SmTrip9 Kd (uM)	VS-HiBiT Kd (uM)	Bmax (RLU)
521	GKMLFRVTINSWK	827	0.32	2.34	2.74E+07
756	DKLLFTVTIQKWK	493	0.48	4.26	1.73E+07
757	DKLLFTVTIQKYK	494	0.36	3.67	1.91E+07
758	DKLLFTVTIEKWK	495	0.19	6.34	1.84E+07
759	DKLLFTVTIEKYK	496	0.33	4.93	2.25E+07
760	KKMLFRVTIQKWK	445	0.00	1.61	4.48E+07
761	KKLLFRVTIQKWK	587	0.02	1.49	4.00E+07
762	DRMLFRVTIQRWR	588	0.12	2.83	3.45E+07
766	ERMLFRVTIGRWR	589	0.07	6.49	2.86E+07
768	GRMLFRVTINRWR	590	0.10	1.96	3.97E+07
770	DRMLFRVTIERWR	591	0.35	3.74	2.53E+07
738	DKMLFRVTIQKWK	400	0.11	2.37	3.97E+07
742	DKMLFRVTIEKWK	406	1.00	0.68	3.89E+07
743	GKMLFRVTINKWK	471	0.10	1.36	4.36E+07
745	DKMLFTVTIQKWK	478	0.20	5.84	2.46E+07
753	DKMLFTVTIEKWK	490	0.15	4.73	2.16E+07
ND = Not Determined					

FIG. 130

Biochemical analysis of synthetic SmTrip9 variants					
Pep ID	Sequence	SEQ ID NO.	SmTrip9 Kd (uM)	VS-HiBiT Kd (uM)	Bmax (RLU)
521	GKMLFRVTINSWK		0.48	3.95	2.43E+07
756	DKLLFTVTIQKWK		0.95	5.52	1.43E+07
757	DKLLFTVTIQKYK		0.19	7.02	1.87E+07
758	DKLLFTVTIEKWK		0.52	6.82	1.45E+07
759	DKLLFTVTIEKYK		0.56	4.94	1.87E+07
760	KKMLFRVTIQKWK		0.02	2.45	4.25E+07
781	GKMLFKVTINKWK		0.38	1.20	4.22E+07
782	GKMLFTVTINKWK		0.36	2.78	3.18E+07
783	DKMLFKVTIQKYK		0.24	2.20	3.24E+07
784	DKMLFRVTINKWK		0.41	11.60	1.67E+07
785	DKMLFKVTIEKYK		0.38	1.28	3.64E+07
786	DKMLFTVTINKWK		0.50	35.65	2.54E+06
787	DKMLFKVTINKWK		0.73	24.17	1.23E+07
793	DKLLFTVTIGKWK		0.54	5.65	1.27E+07
794	DKLLFTVTIGKYK		0.23	5.96	1.01E+07
799	DKLLFTVTINKWK		0.21	42.69	1.51E+06
800	DKLLFTVTINKYK		9.75	88.68	6.49E+05
743	GKMLFRVTINKWK		0.11	1.19	4.05E+07
745	DKMLFTVTIQKWK		0.30	25.63	2.66E+06
749	DKMLFTVTIGKWK		1.06	6.57	1.34E+07
ND = Not Determined					

FIG. 131A

Solubility of synthetic SmTrip9 peptides			
Pep ID	Sequence	SEQ ID NO.	Solubility#
521	*GKMLFRVTINSWK		Y
627	*DKMLFRVTINSWK		N
628	*EKMLFRVTINSWK		Y
629	*RKMLFRVTINSWK		Y
630	*KKMLFRVTINSWK		Y
631	*HKMLFRVTINSWK		N
632	*GLMLFRVTINSWK		N
633	*GQMLFRVTINSWK		N
634	*GTMLFRVTINSWK		N
635	*GKLLFRVTINSWK		Y
636	*GKMLFKVTINSWK		Y
637	*GKMLFRVTIQSWK		Y
638	*GKMLFRVTIDSWK		Y
639	*GKMLFRVTIGSWK		Y
640	*GKMLFRVTINTWK		Y
641	*GKMLFRVTINNWK		Y
642	*GKMLFRVTINQWK		Y
643	*GKMLFRVTINPWK		Y
644	*GKMLFRVTINKWK		Y
645	*GKMLFRVTINSWQ		Y
646	*GKMLFRVTINSWN		Y
647	*GKMLFRVTINSWT		Y
648	*GKMLFRVTINSWH		Y
649	*GKMLFRVTINSWP		Y
650	*GKMLFRVTINSWR		N
677	GKMKFRVTIDSWK		Y
678	GKMLFRVKINSWK		Y
679	GKMLFRVRINSWK		Y
680	GKMLFRVEINSWK		Y
681	GKMLFRVDINSWK		N
682	GKMLFRVQINSWK		Y
683	GKMKFRVKINSWK		Y
684	GKMKFRVRINSWK		Y
685	GKMKFRVEINSWK		Y
686	GKMKFRVDINSWK		Y
687	GKMKFRVQINSWK		Y
688	GKMKFRVNINSWK		Y
689	GKMKFRVSINSWK		Y
690	GKMKFRVTINSWK		Y
#in water at ~1 mM after multiple freeze/thaws *Terminus unblocked			

FIG. 131B

Solubility of synthetic SmTrip9 peptides			
Pep ID	Sequence	SEQ ID NO.	Solubility#
521	GKMLFRVTINSWK	827	Y
613	GKMLFRVNINSWK	542	Y
614	GKMLFRVSINSWK	543	Y
615	GKMLFRVWINSWK	544	Y
616	GKMSFRVTINSWK	545	Y
617	GKMWFRVTINSWK	546	Y
618	GKMNFRVTINSWK	547	Y
619	GSMLFRVTINSYK	548	Y
620	GKMLFRVTINSYK	549	Y
621	GKMLFRVTIKSWK	550	Y
622	GKMLFRVTIESWK	551	Y
716	GKMKFRVTIQSWK	552	Y
717	GKMKFRVTIESWK	553	Y
718	GKMKFRVTIKSWK	554	Y
719	GKMKFRVTIRSWK	555	Y
651	RLMLFRVTINSWK	827	N
652	RQMLFRVTINSWK	556	N
653	KLMLFRVTINSWK	557	N
654	KQMLFRVTINSWK	558	N
655	ELMLFRVTINSWK	559	N
656	EQMLFRVTINSWK	560	N
657	DLMLFRVTINSWK	463	N
658	DQMLFRVTINSWK	561	N
659	DKMLFRVTINSWK	562	N
660	EKMLFRVTINSWK	563	Y
661	RKMLFRVTINSWK	564	Y
662	KKMLFRVTINSWK	565	Y
663	GKMLFRVTIDSWK	566	Y
665	GKMLFRVTIGSWK	470	Y
667	GKMLFRVTINKWK	567	Y
668	GKMLFRVTIGKWK	568	Y
669	GKMLFRVTIGNWK	392	Y
670	GKMLFRVTISKWK	394	Y
671	GKMLFRVTIQKWK	569	Y
672	GKMLFRVTITKWK	570	Y
673	GKMLFRVTIKKWK	571	Y
675	GKMLFKVTINSWK	572	Y
676	RLMLFRVTIGKWK	573	Y
701	GKMLFRVTINRWK	575	Y
703	GKMLFRVTIEKWK	397	Y
707	EKMLFRVTIEKWK	442	Y
708	EKMLFRVTIGKWK	411	Y
709	EKMLFRVTIGRWK	412	Y
710	EKMLFTVTIGKWK	576	Y
711	EKLLFTVTIGKWK	577	Y
712	EKMLFTVTIGRWK	578	Y
714	EKMLFKVTIQKWK	472	Y
720	EKMLFTVTIEKWK	579	Y
722	DKMLFRVTIESWK	580	Y
726	EKLLFRVTIGKYK	581	Y
#in water at ~1 mM after multiple freeze/thaws *Terminus unblocked			

FIG. 131C

Solubility of synthetic SmTrip9 peptides			
Pep ID	Sequence	SEQ ID NO.	Solubility#
521	GKMLFRVTINSWK	827	Y
613	GKMLFRVNINSWK	542	Y
614	GKMLFRVSINSWK	543	Y
615	GKMLFRVWINSWK	544	Y
616	GKMSFRVTINSWK	545	Y
617	GKMWFRVTINSWK	546	Y
618	GKMNFRVTINSWK	547	Y
619	GSMLFRVTINSYK	548	Y
620	GKMLFRVTINSYK	549	Y
621	GKMLFRVTIKSWK	550	Y
622	GKMLFRVTIESWK	551	Y
716	GKMKFRVTIQSWK	552	Y
717	GKMKFRVTIESWK	553	Y
718	GKMKFRVTIKSWK	554	Y
719	GKMKFRVTIRSWK	555	Y
651	RLMLFRVTINSWK	827	N
652	RQMLFRVTINSWK	556	N
653	KLMLFRVTINSWK	557	N
654	KQMLFRVTINSWK	558	N
655	ELMLFRVTINSWK	559	N
656	EQMLFRVTINSWK	560	N
657	DLMLFRVTINSWK	463	N
658	DQMLFRVTINSWK	561	N
659	DKMLFRVTINSWK	562	N
660	EKMLFRVTINSWK	563	Y
661	RKMLFRVTINSWK	564	Y
662	KKMLFRVTINSWK	565	Y
663	GKMLFRVTIDSWK	566	Y
665	GKMLFRVTIGSWK	470	Y
667	GKMLFRVTINKWK	567	Y
668	GKMLFRVTIGKWK	568	Y
669	GKMLFRVTIGNWK	392	Y
670	GKMLFRVTISKWK	394	Y
671	GKMLFRVTIQKWK	569	Y
672	GKMLFRVTITKWK	570	Y
673	GKMLFRVTIKKWK	571	Y
675	GKMLFKVTINSWK	572	Y
676	RLMLFRVTIGKWK	573	Y
701	GKMLFRVTINRWK	575	Y
703	GKMLFRVTIEKWK	397	Y
707	EKMLFRVTIEKWK	442	Y
708	EKMLFRVTIGKWK	411	Y
709	EKMLFRVTIGRWK	412	Y
710	EKMLFTVTIGKWK	576	Y
711	EKLLFTVTIGKWK	577	Y
712	EKMLFTVTIGRWK	578	Y
714	EKMLFKVTIQKWK	472	Y
720	EKMLFTVTIEKWK	579	Y
722	DKMLFRVTIESWK	580	Y
726	EKLLFRVTIGKYK	581	Y
#in water at ~1 mM after multiple freeze/thaws *Terminus unblocked			

FIG. 132

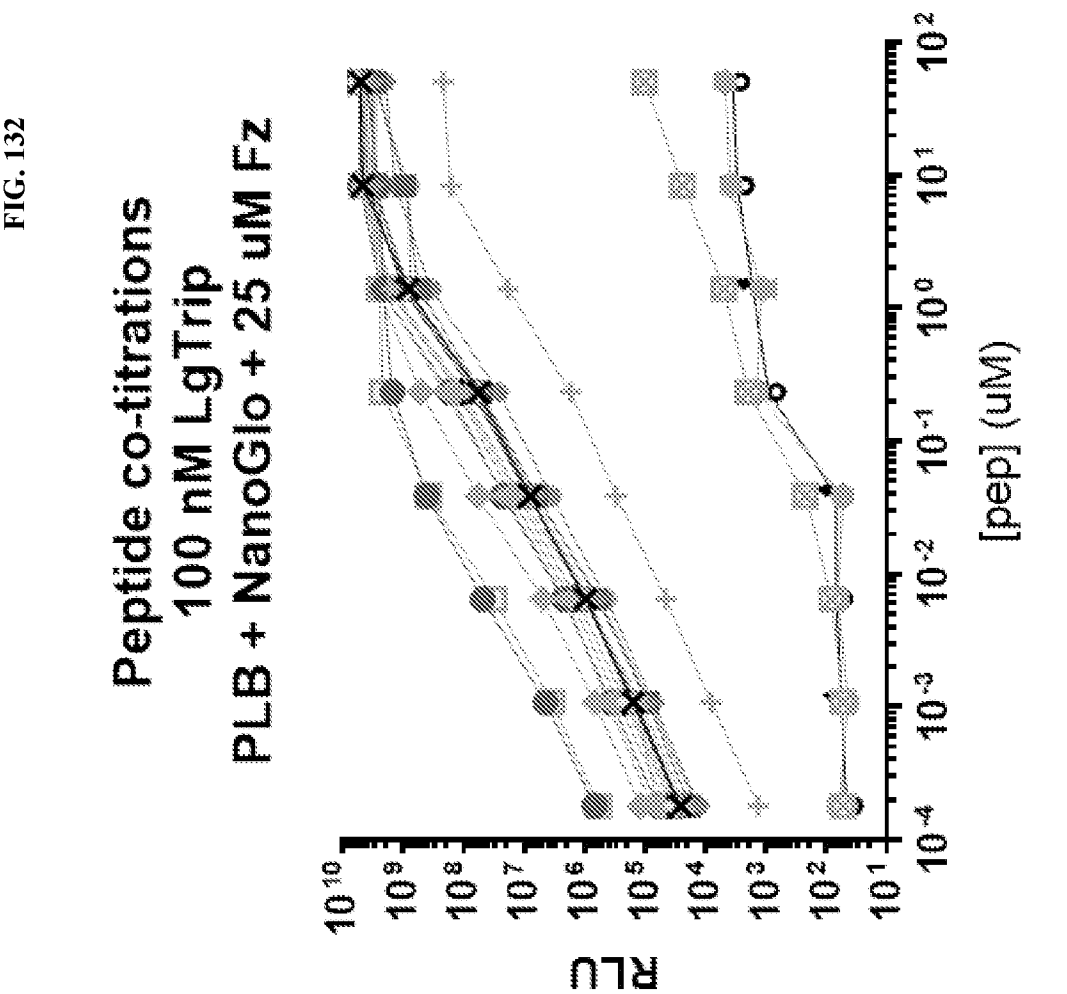

Peptide co-titrations
100 nM LgTrip
PLB + NanoGlo + 25 uM Fz
RLU
[pep] (uM)
743
745
749
747
752
744
LgTrip + 289
LgTrip + 521
LgTrip only
757
755
753
758
759
756
521
760
761
738
742

FIG. 133

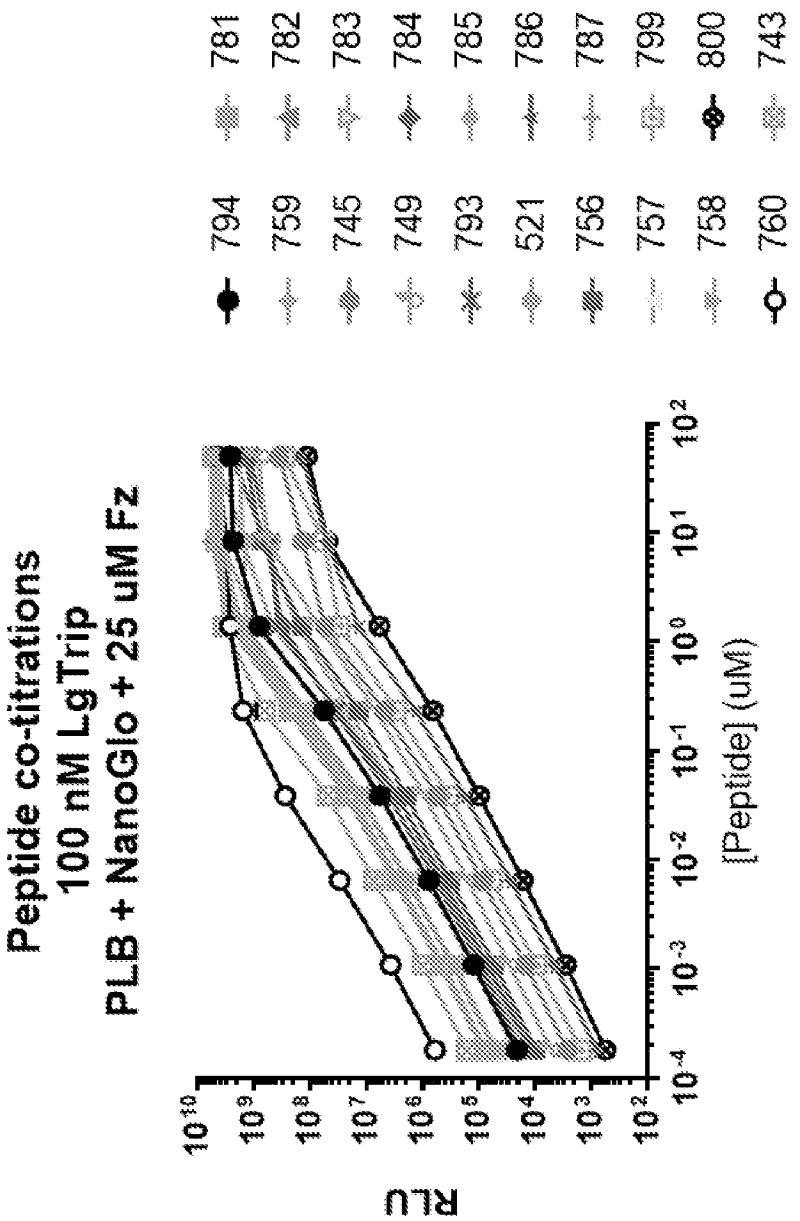

Peptide co-titrations
100 nM LgTrip
PLB + NanoGlo + 25 uM Fz
RLU
[Peptide] (uM)
794
759
745
749
793
521
756
757
758
760
781
782
783
784
785
786
787
799
800
743

FIG. 134

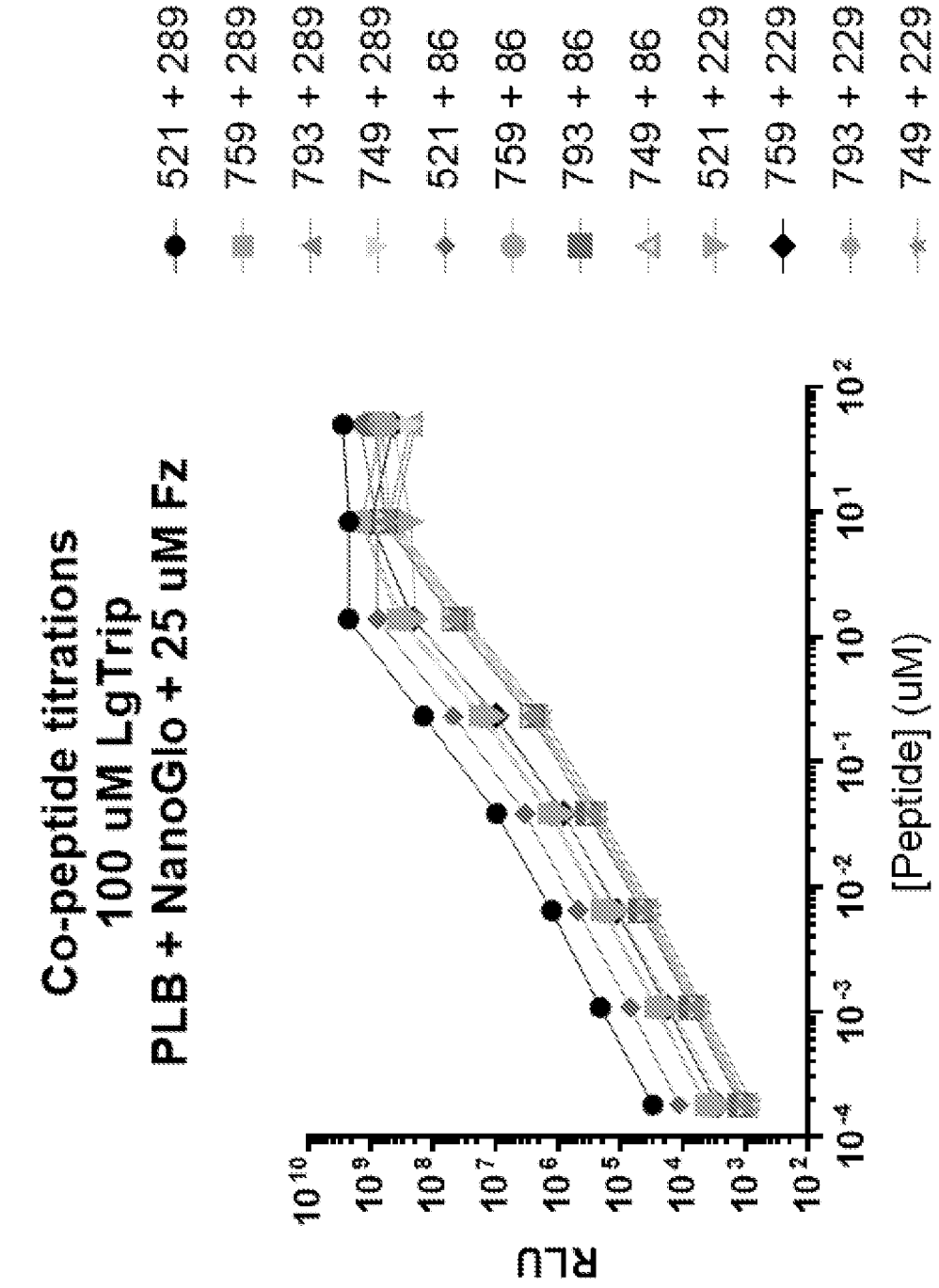

Co-peptide titrations
100 uM LgTrip
PLB + NanoGlo + 25 uM Fz
521 + 289
759 + 289
793 + 289
749 + 289
521 + 86
759 + 86
793 + 86
749 + 86
521 + 229
759 + 229
793 + 229
749 + 229
RLU
[Peptide] (uM)

FIG. 135

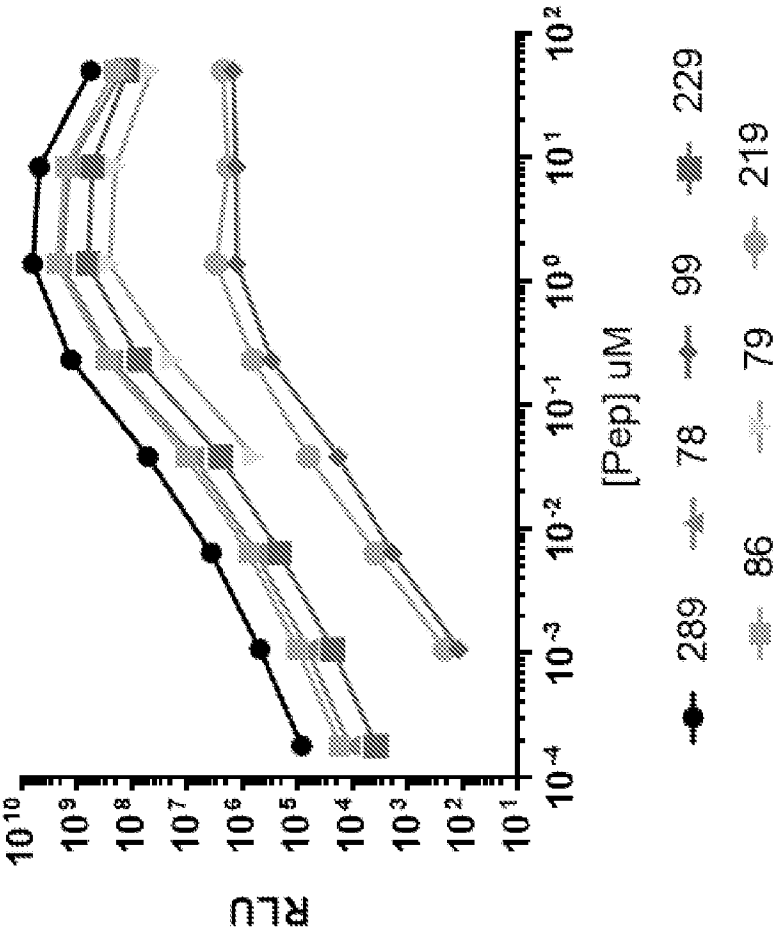

Biochemical peptide co-titration
100 nM LgTrip in PLB + NanoGlo
RLU
[Pep] uM
289
78
99
229
86
79
219

FIG. 136

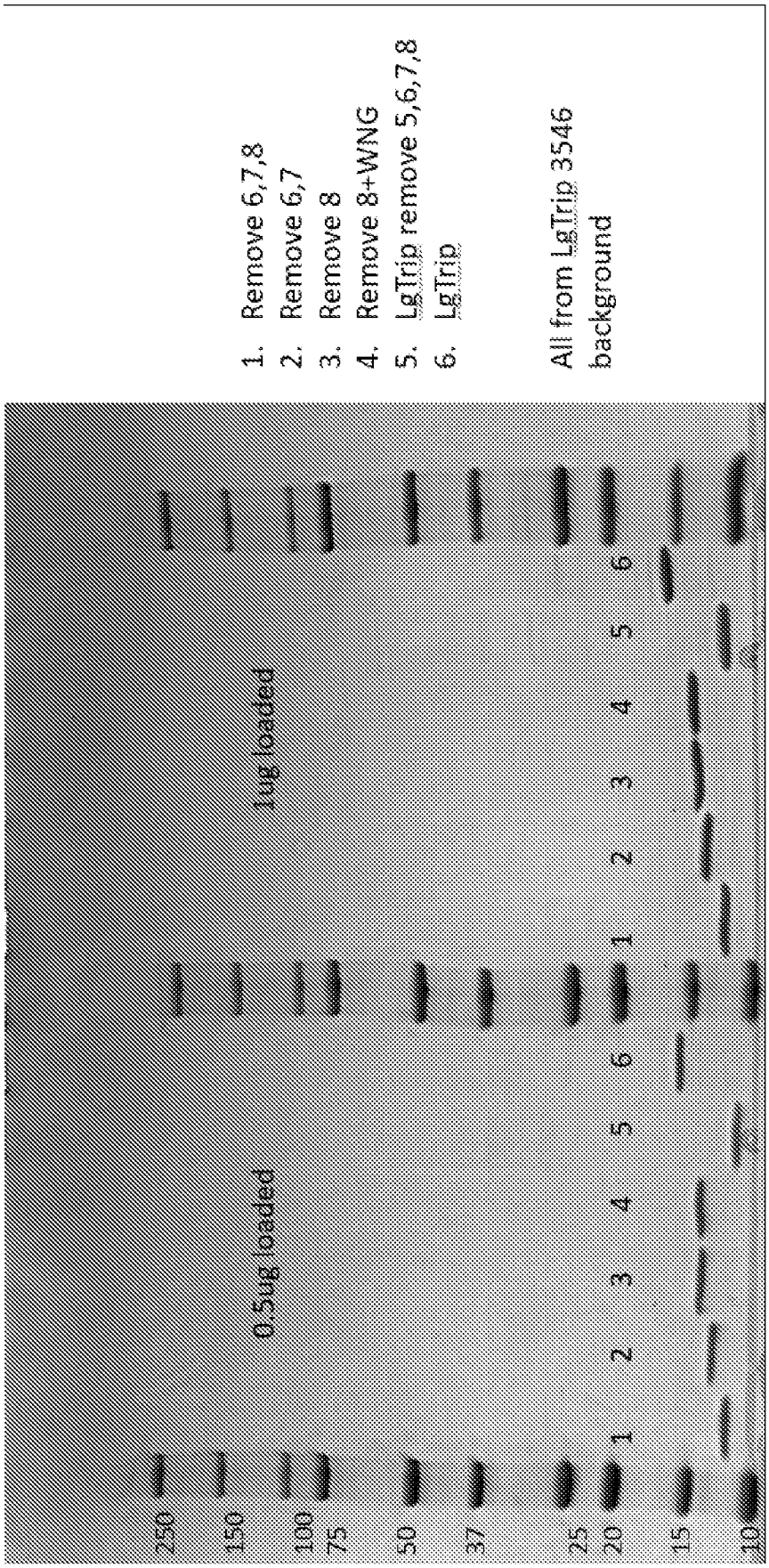

1. Remove 6,7,8
2. Remove 6,7
3. Remove 8
4. Remove 8+WNG
5. LgTrip remove 5,6,7,8
6. LgTrip
All from LgTrip 3546 background
0.5ug loaded
1ug loaded
250
150
100
75
50
37
25
20
15
10
1
2
3
4
5
6

FIG. 137A

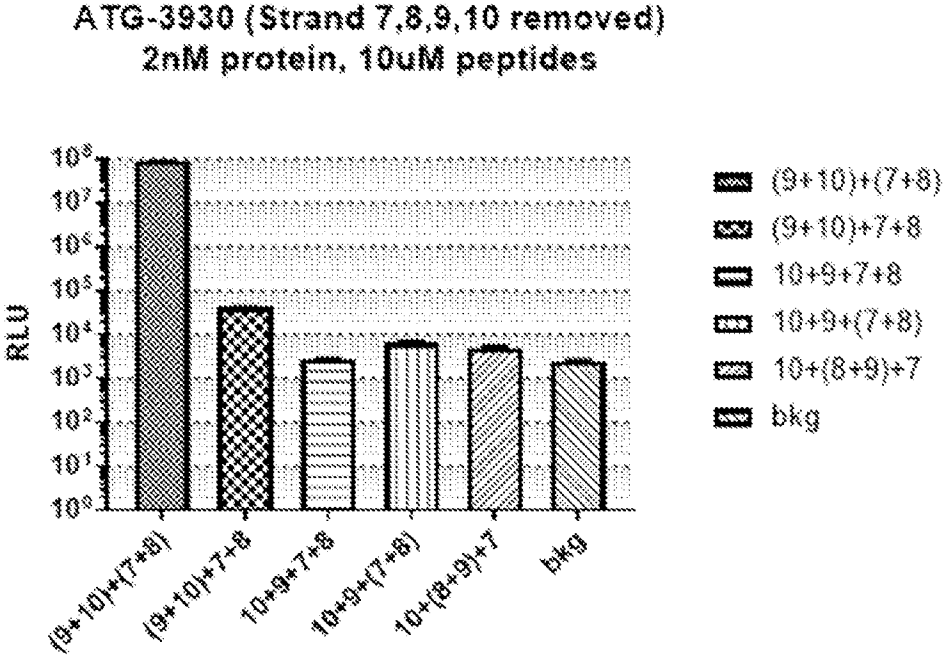


FIG. 137B

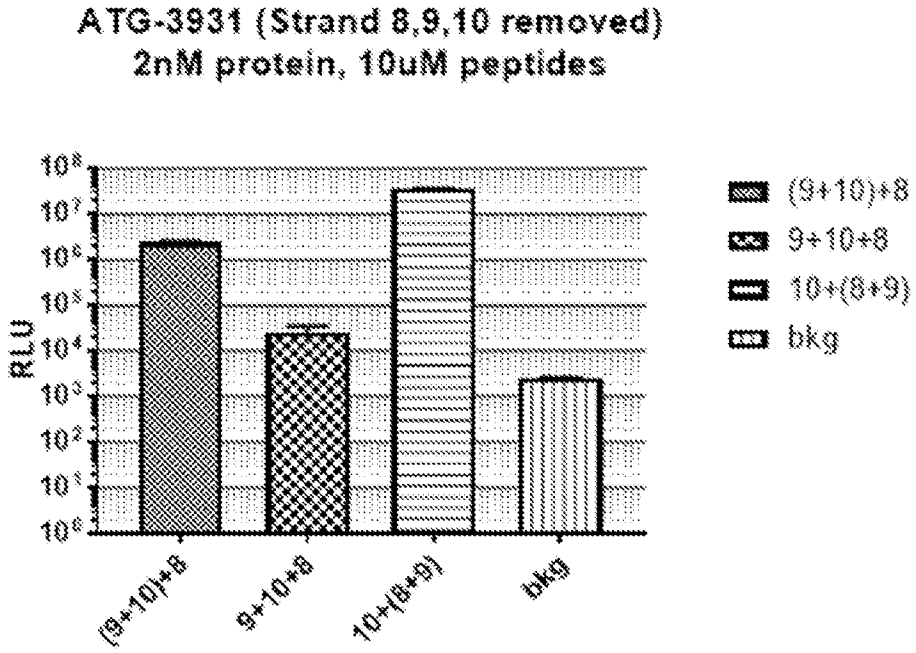


FIG. 137C

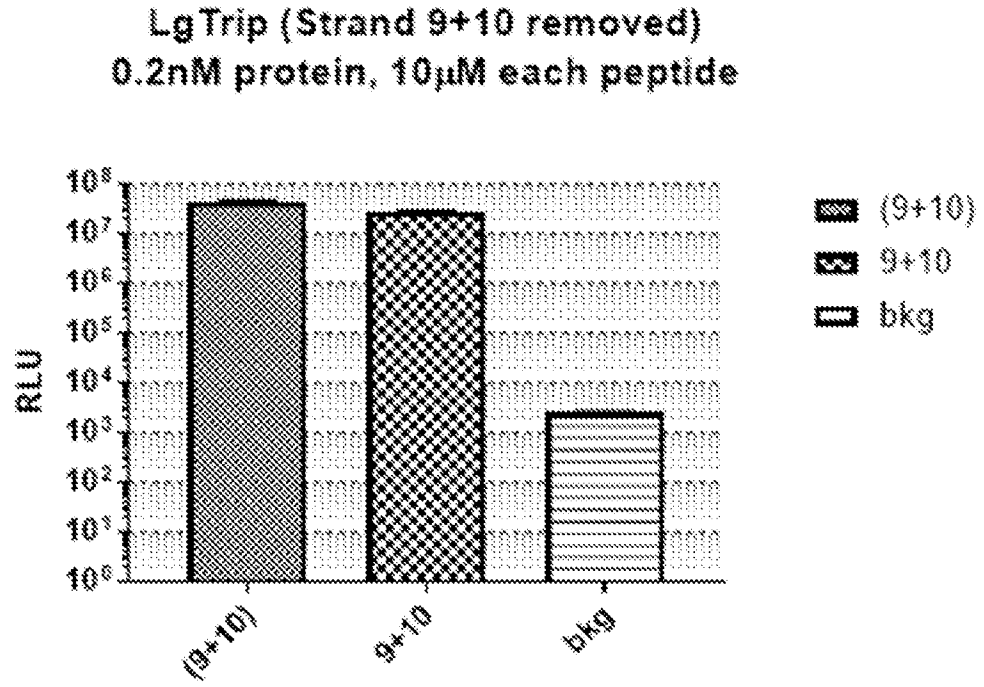


FIG. 137D

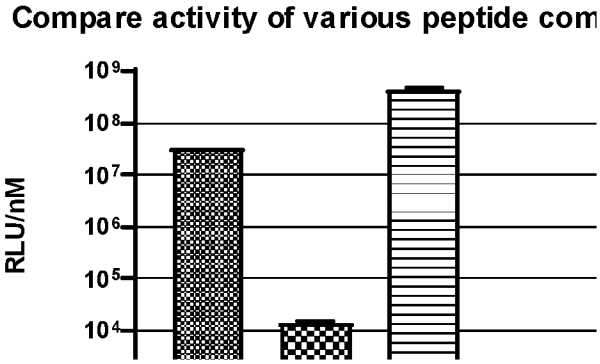


FIG. 138A

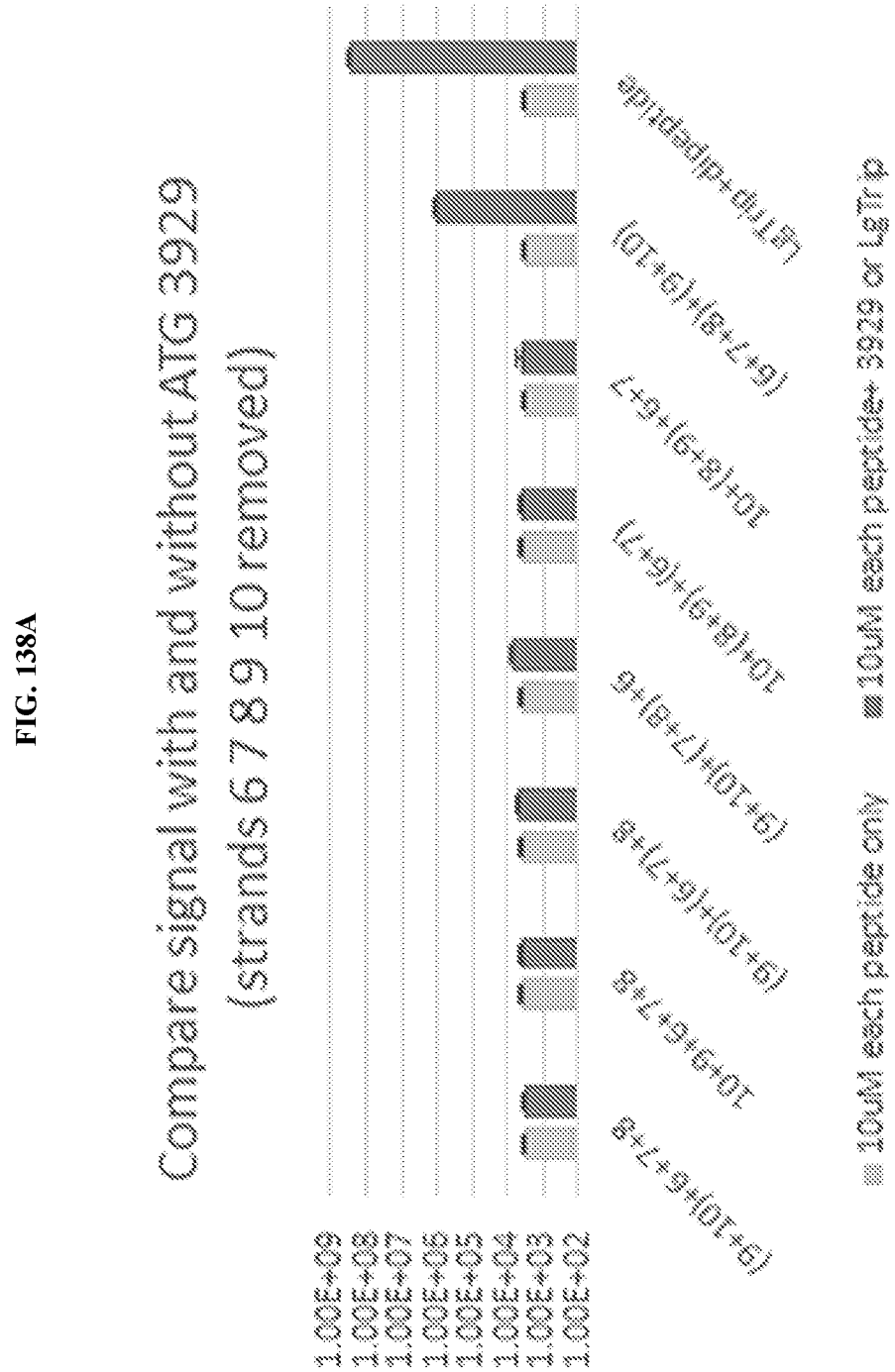

Compare signal with and without ATG 3929
(strands 6 7 8 9 10 removed)
1.00E+09
1.00E+08
1.00E+07
1.00E+06
1.00E+05
1.00E+04
1.00E+03
1.00E+02
6+7+9+(10T+6)
8+7+9+6+10T
8+(7+9)+(10T+6)
9+(8+7)+(10T+6)
(7+9)+(6+8)+10T
7+9+(6+8)+10T
(8T+6)+(8+7+9)
spiradadip+6p+(8T+6)
10uM each peptide only
10uM each peptide+ 3929 or LgTrip

FIG. 138B

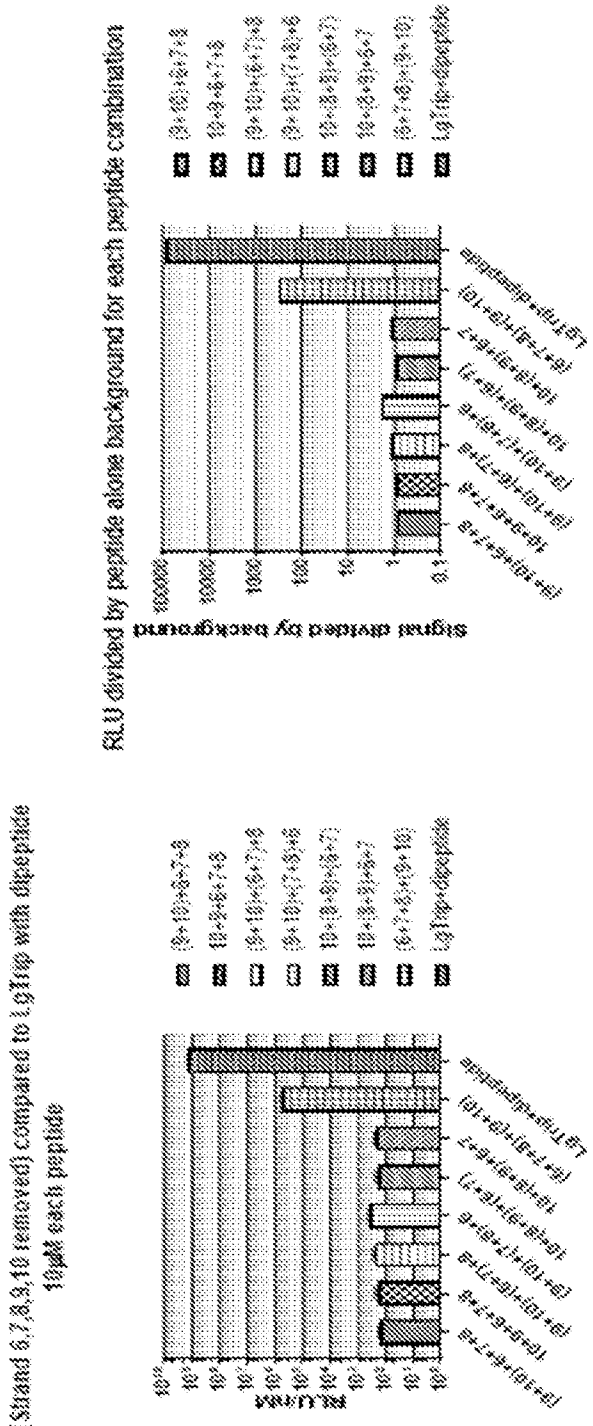

RLU divided by peptide alone background for each peptide combination
Signal divided by background
Activity of 3328 (Strand 6,7,8,9,10 removed) compared to LgTrip with dipeptide
10pM each peptide
RLU/rxn
LgTrip+dipeptide

FIG. 139A

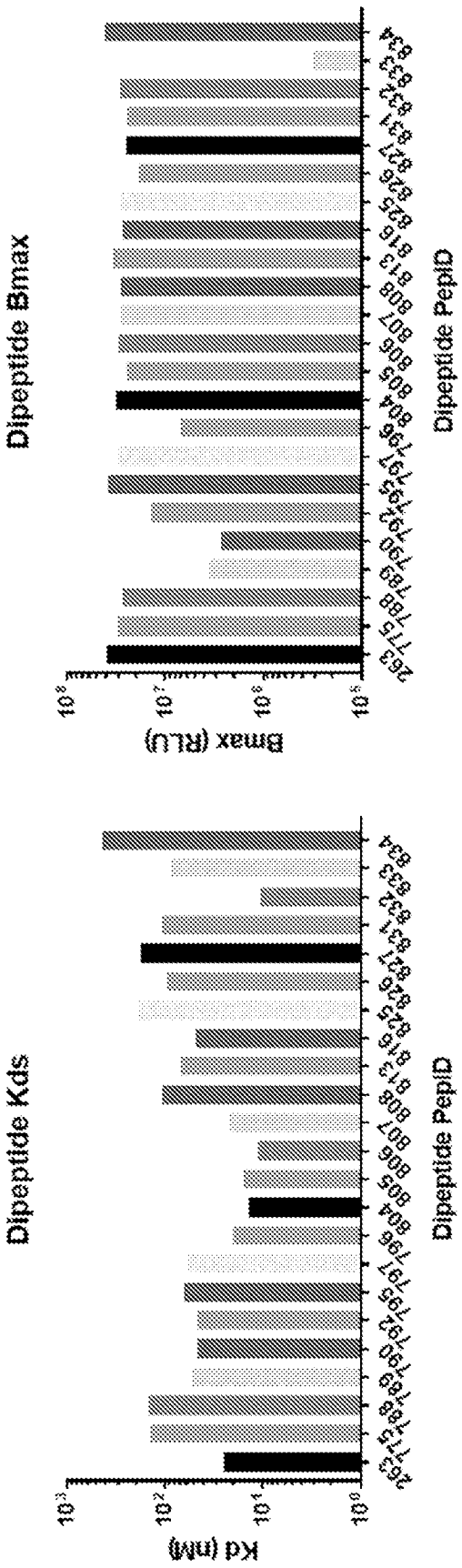

Dipeptide Bmax
Bmax (RLU)
Dipeptide PepID
Dipeptide Kds
Kd (nM)
Dipeptide PepID

FIG. 139B

peptide ID	Bmax	Kd
263	3.01E+07	11.59
592	3.00E+07	119.1
595	2.06E+07	56.02
603	1.76E+06	31.17
604	2.78E+07	10.05
605	2.76E+07	19.61
606	2.83E+07	46.39
607	3.42E+07	20.18
593	2.34E+07	95.48
594	1.19E+07	115
598	3.24E+07	178.1
599	2.05E+07	122.1
600	1.56E+07	83.39
602	2.67E+07	103.2
591	3.44E+06	44.5
596	8.79E+06	648.3
597	1.59E+06	86.75
601	3.12E+06	47.6
609	1.26E+07	4308
610	3.60E+06	535.8

peptide ID	Bmax	Kd
263	1.00	1.0
592	1.00	10.3
595	0.68	4.8
603	0.06	2.7
604	0.93	0.9
605	0.92	1.7
606	0.94	4.0
607	1.14	1.7
593	0.78	8.2
594	0.40	9.9
598	1.08	15.4
599	0.68	10.5
600	0.52	7.2
602	0.89	8.9
591	0.11	3.8
596	0.29	55.9
597	0.05	7.5
601	0.10	4.1
609	0.42	371.7
610	0.12	46.2

FIG. 139C

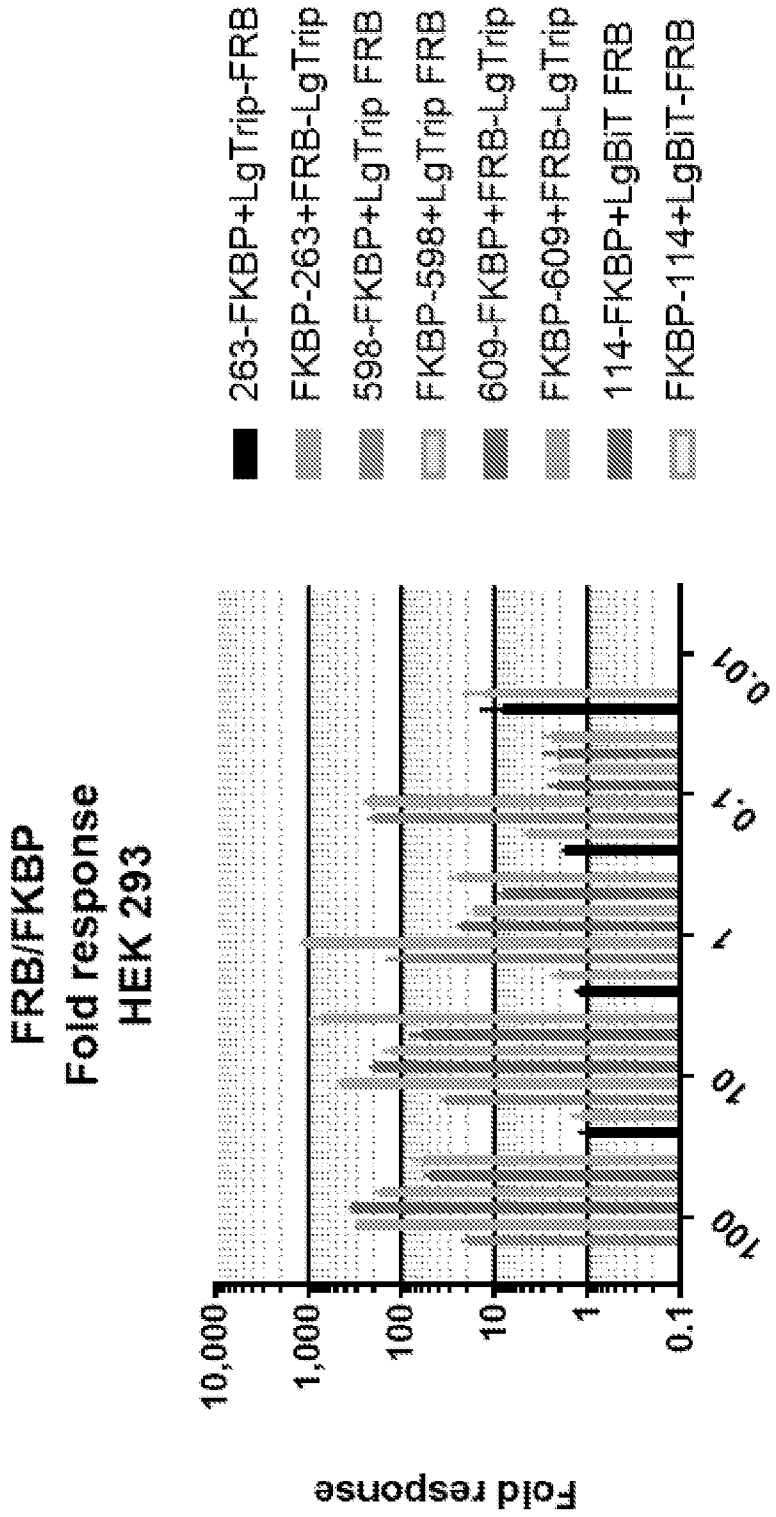

FRB/FKBP
Fold response
HEK 293
263-FKBP+LgTrip-FRB
FKBP-263+FRB-LgTrip
598-FKBP+LgTrip FRB
FKBP-598+LgTrip FRB
609-FKBP+FRB-LgTrip
FKBP-609+FRB-LgTrip
114-FKBP+LgBiT FRB
FKBP-114+LgBiT-FRB
Fold response
10,000
1,000
100
10
1
0.1
100
10
1
0.1
0.01

FIG. 139D

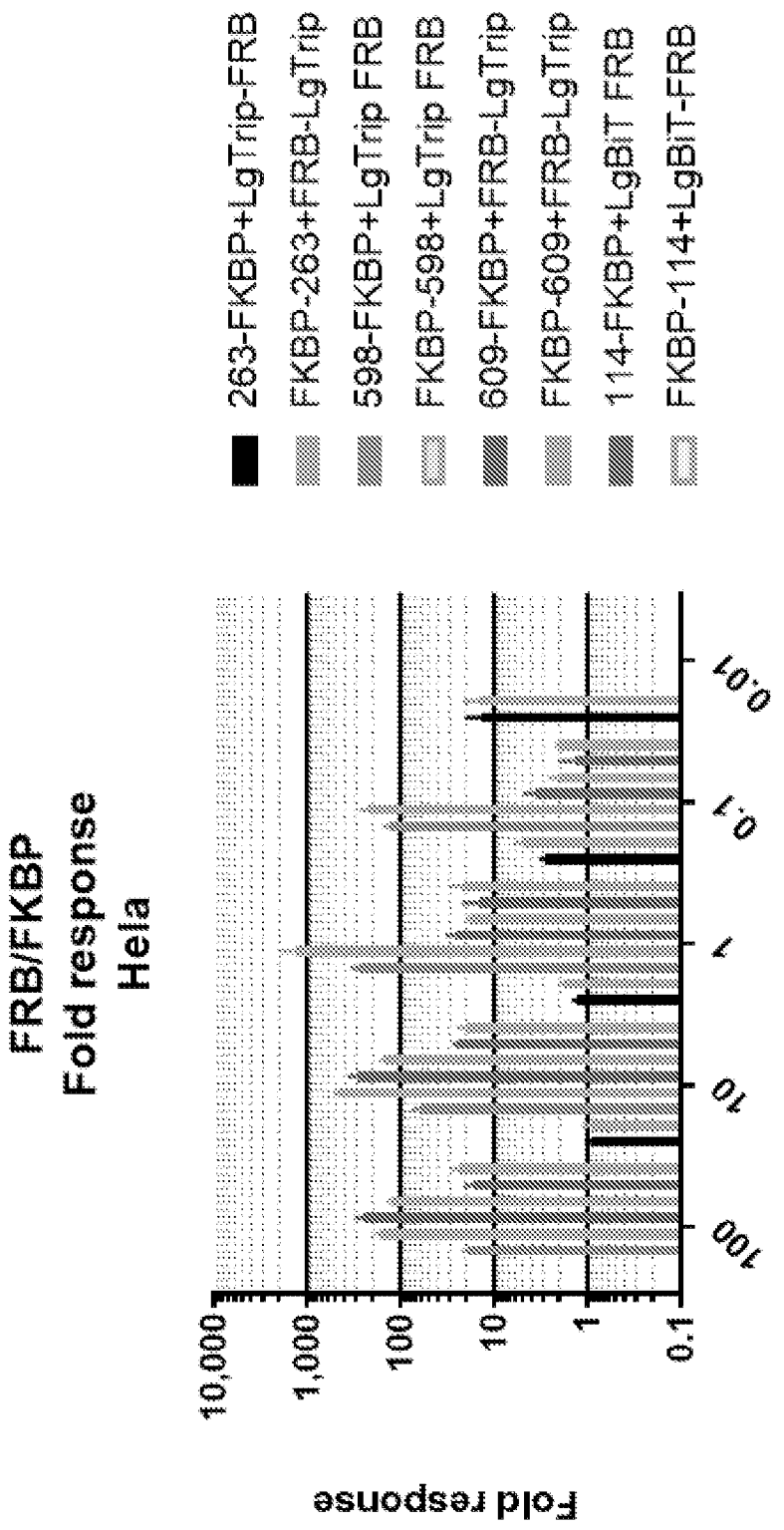

FRB/FKBP
Fold response
Hela
263-FKBP+LgTrip-FRB
FKBP-263+FRB-LgTrip
598-FKBP+LgTrip FRB
FKBP-598+LgTrip FRB
609-FKBP+FRB-LgTrip
FKBP-609+FRB-LgTrip
114-FKBP+LgBiT FRB
FKBP-114+LgBiT-FRB
Fold response
10,000
1,000
100
10
1
0.1
100
10
1
0.1
0.01

FIG. 139E

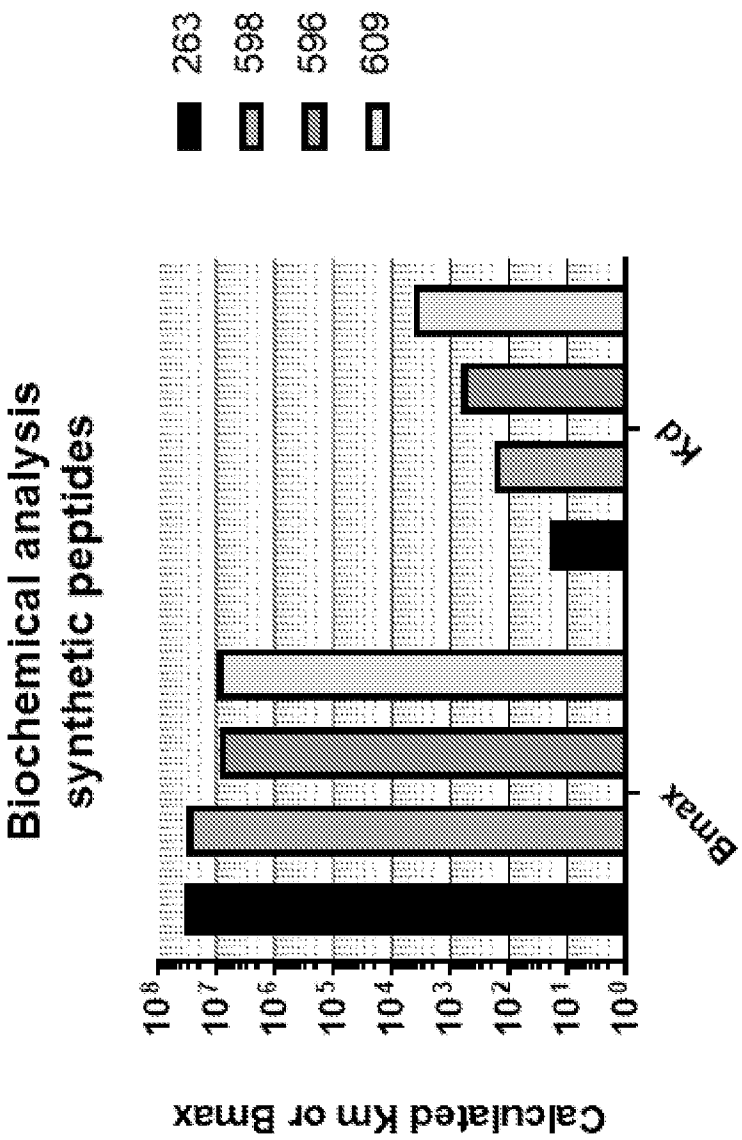

263
598
596
609
Biochemical analysis
synthetic peptides
Calculated Km or Bmax
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
10^0
Bmax
Kd

FIG. 140

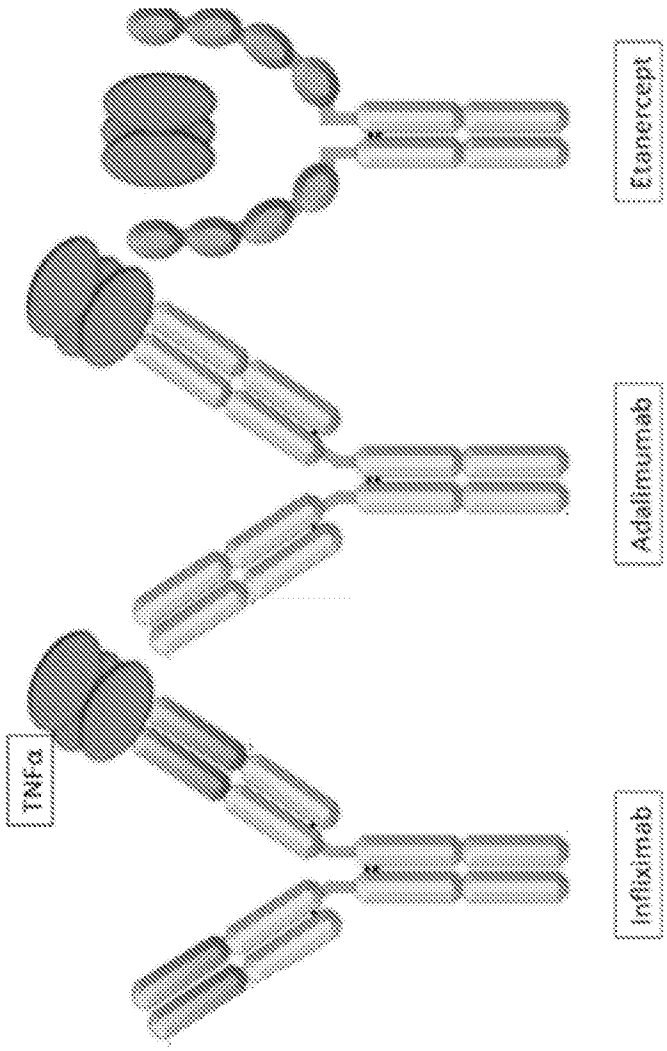
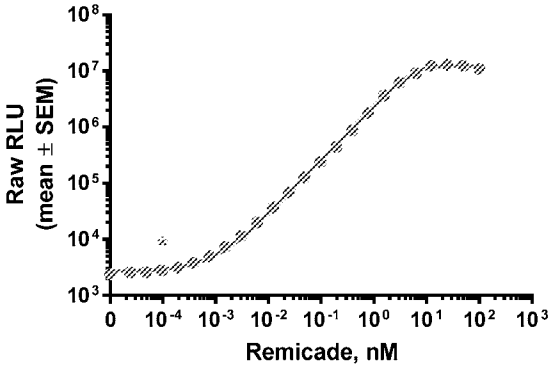


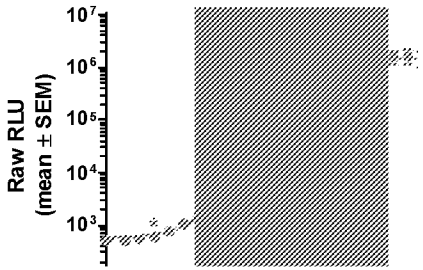
FIG. 141

Infliximab



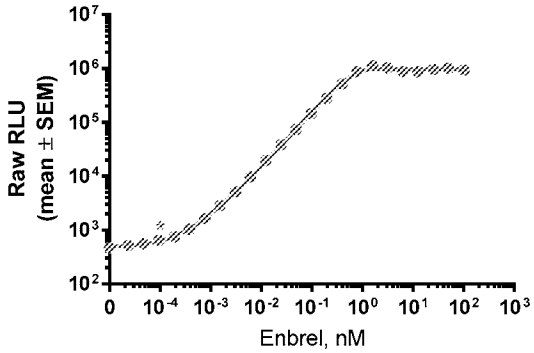
	Trip (521)
LOD*	0.10 pM
LOQ	1.0 pM
ULOQ	17,500 pM

Adalimumab



	Trip (521)
LOD*	0.10 pM
LOQ	0.80 pM
ULOQ	6876 pM

Etanercept



	Trip (521)
LOD*	0.10 pM
LOQ	0.3 pM
ULOQ	1319 pM

$LOD = mean_{blank} + 2*SD_{blank}$

$LOQ = mean_{blank} + 10*SD_{blank}$

$ULOQ = mean_{[high\ in\ linear\ dynamic\ range]} + 10*SD_{[high\ in\ linear\ dynamic\ range]}$

These RLU values are then interpolated from the dose response curve to obtain concentration values

FIG. 142

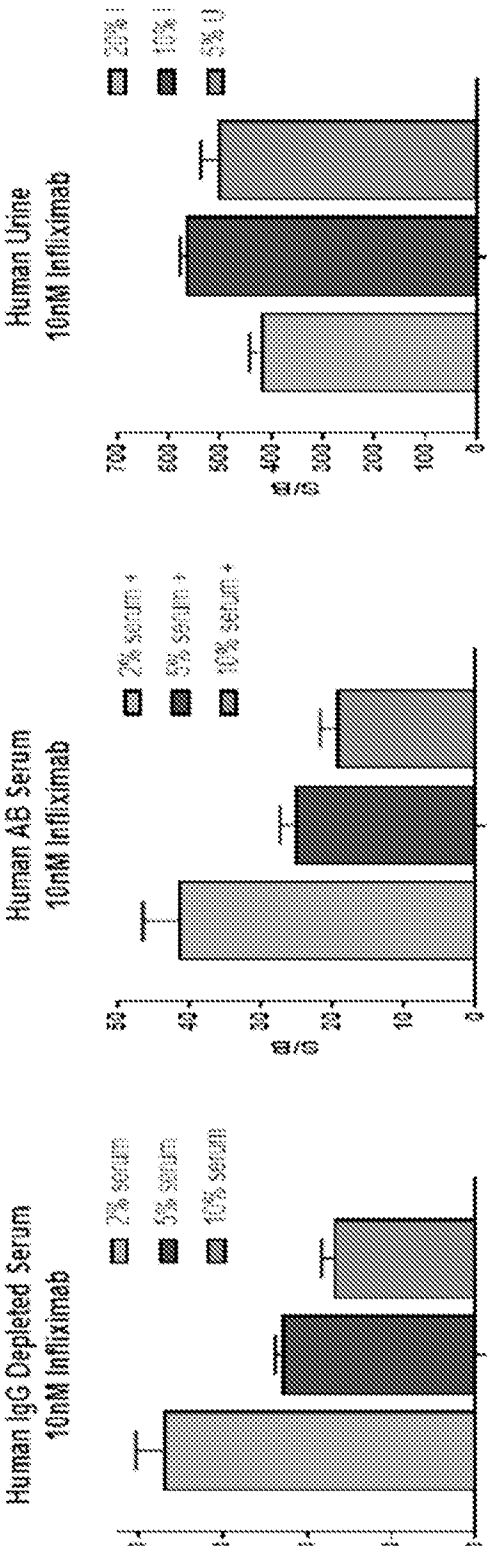

Human IgG Depleted Serum
10nM Infliximab
2% serum
5% serum
10% serum
Human AB Serum
10nM Infliximab
Human Urine
10nM Infliximab

FIG. 143

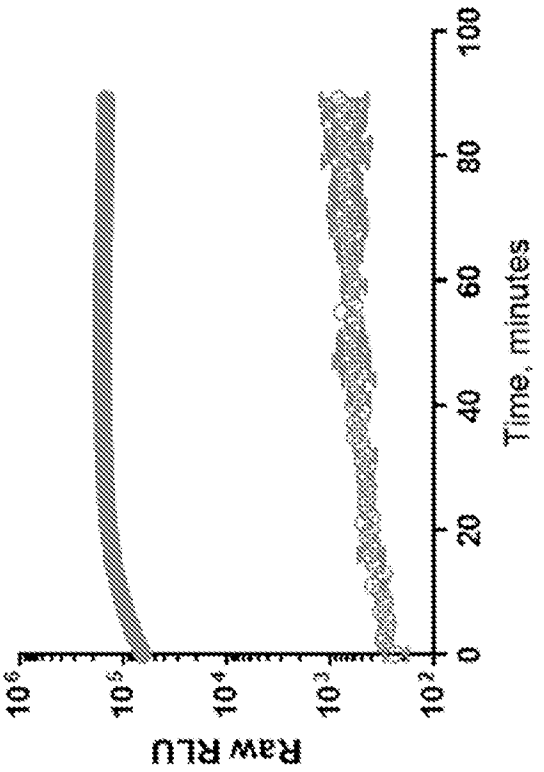

Real Time Binding Kinetics
Infliximab
NanoTrip + 100pM infliximab
NanoTrip neg control
Raw RLU
10^6
10^5
10^4
10^3
10^2
0
20
40
60
80
100
Time, minutes

FIG. 144

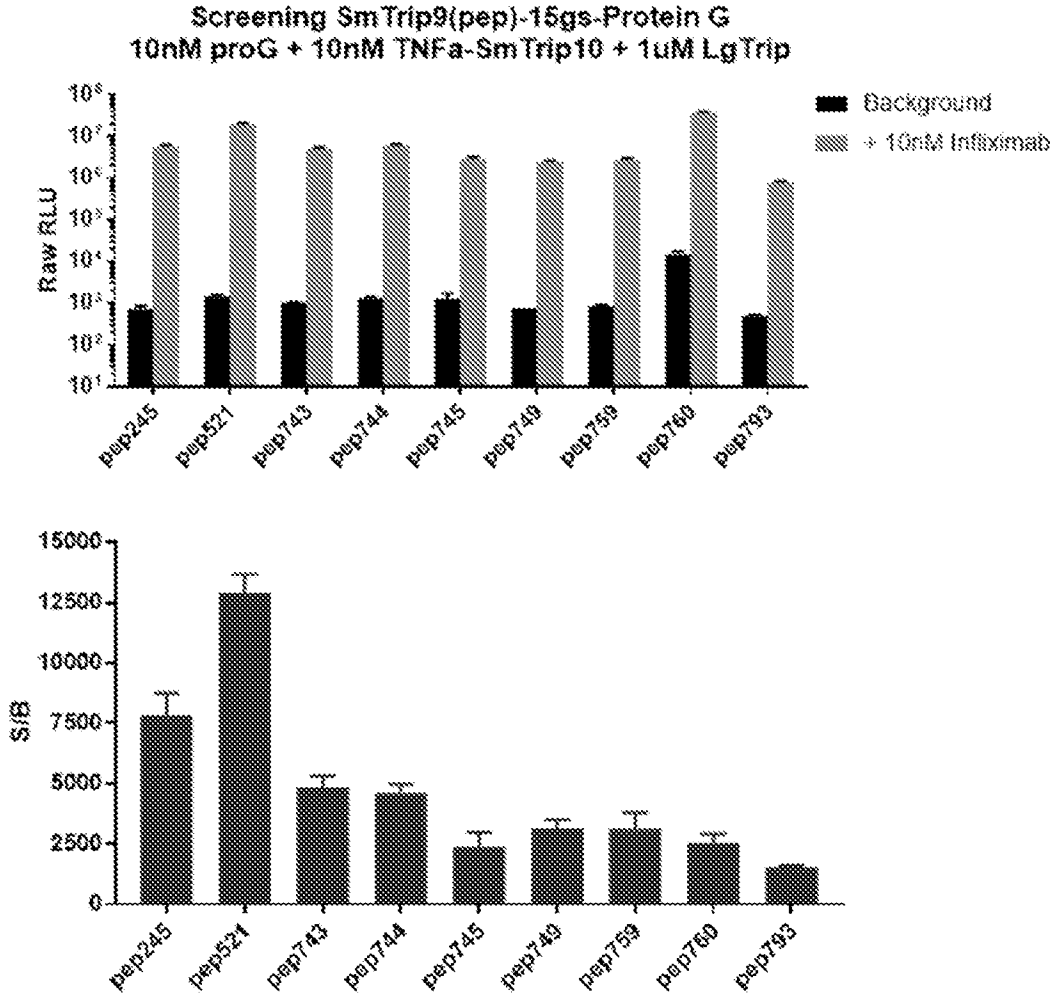

Screening SmTrip9(pep)-15gs-Protein G
10nM proG + 10nM TNFa-SmTrip10 + 1uM LgTrip
Background
+ 10nM Infliximab
Raw RLU
pep245
pep521
pep743
pep744
pep745
pep749
pep759
pep760
pep793
S/B
15000
12500
10000
7500
5000
2500
0

FIG. 145

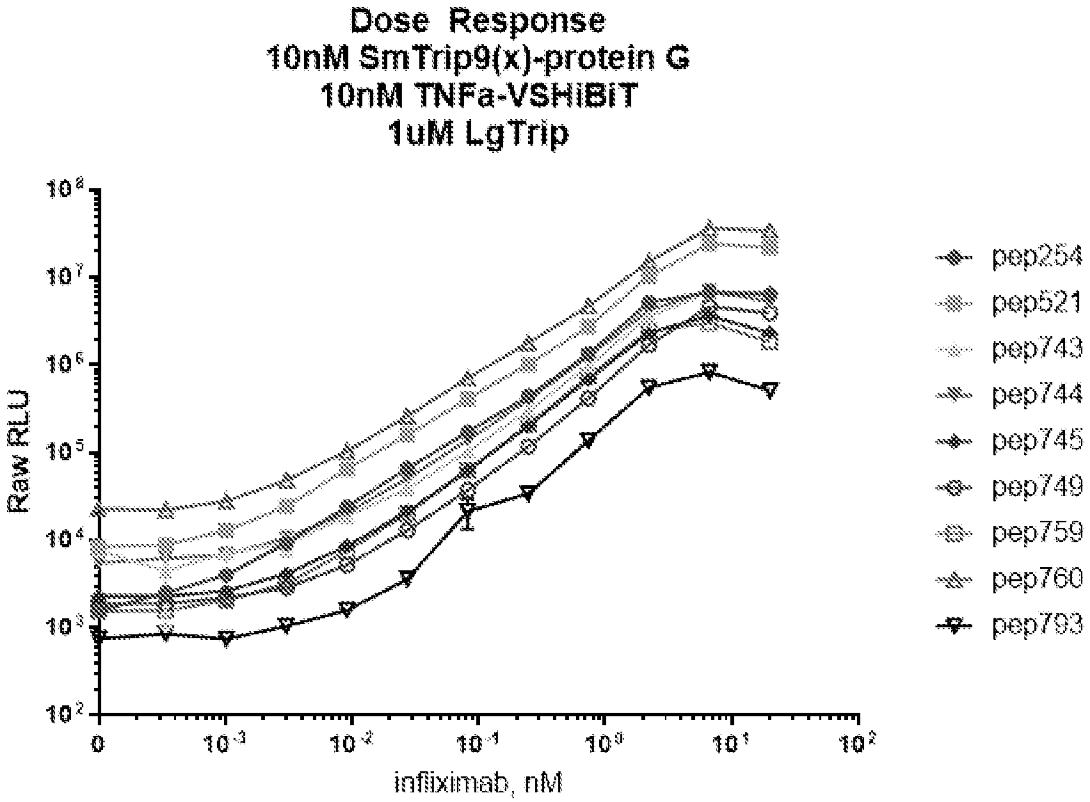

Dose Response
10nM SmTrip9(x)-protein G
10nM TNFa-VSHiBiT
1uM LgTrip
Raw RLU
infliximab, nM
pep254
pep521
pep743
pep744
pep745
pep749
pep759
pep760
pep793

FIG. 146

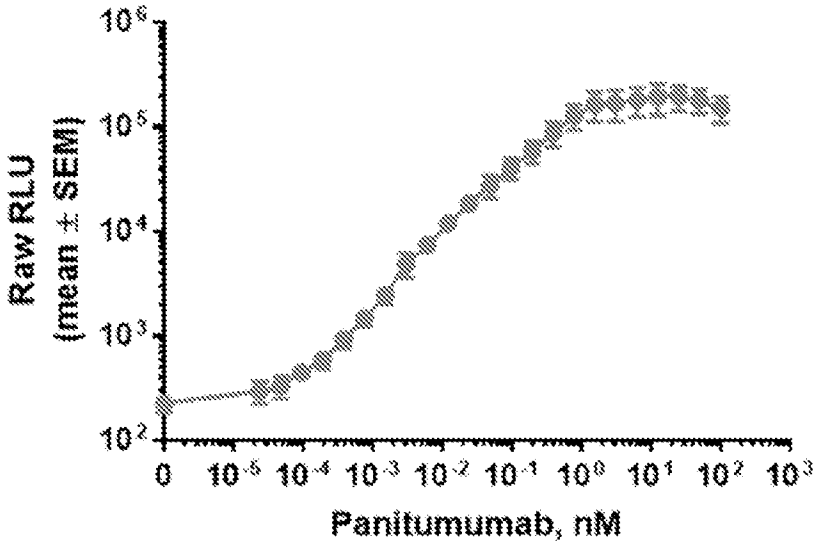

Raw RLU
(mean ± SEM)
10^6
10^5
10^4
10^3
10^2
0
10^-5
10^-4
10^-3
10^-2
10^-1
10^0
10^1
10^2
10^3
Panitumumab, nM

FIG. 147

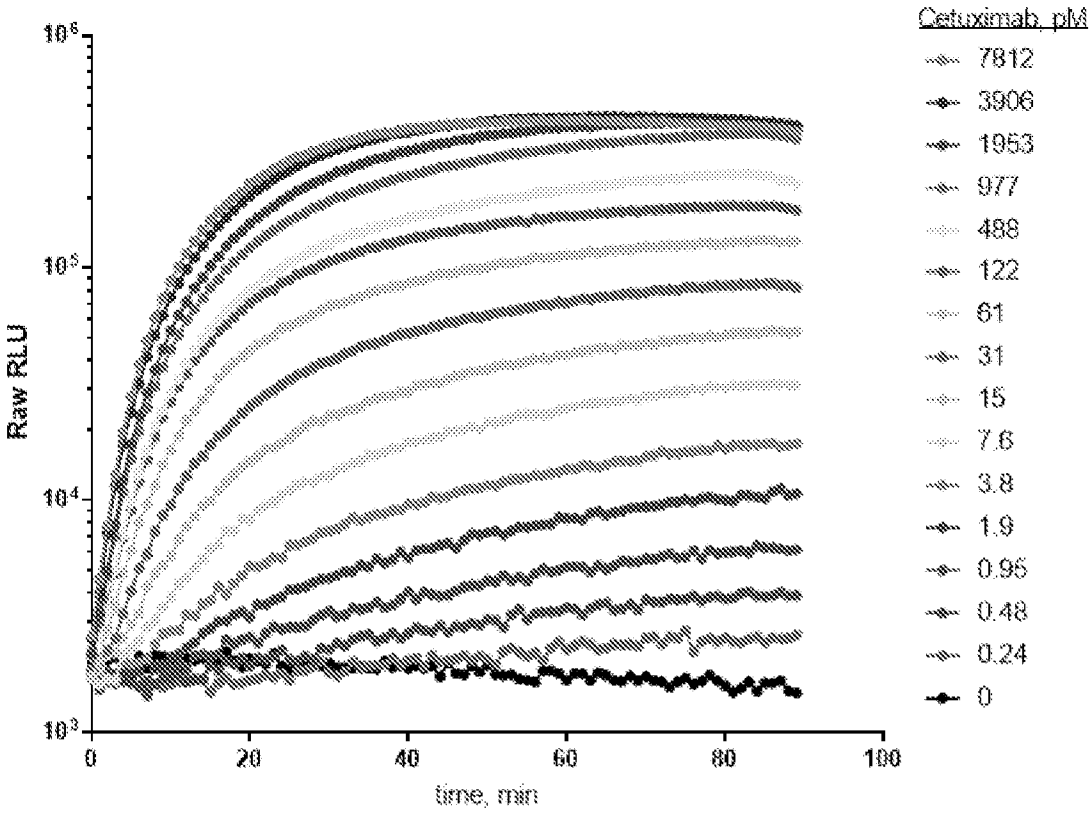

Cetuximab, pM
7812
3906
1953
977
488
122
61
31
15
7.6
3.8
1.9
0.95
0.48
0.24
0
Raw RLU
10^6
10^5
10^4
10^3
0
20
40
60
80
100
time, min

FIG. 148

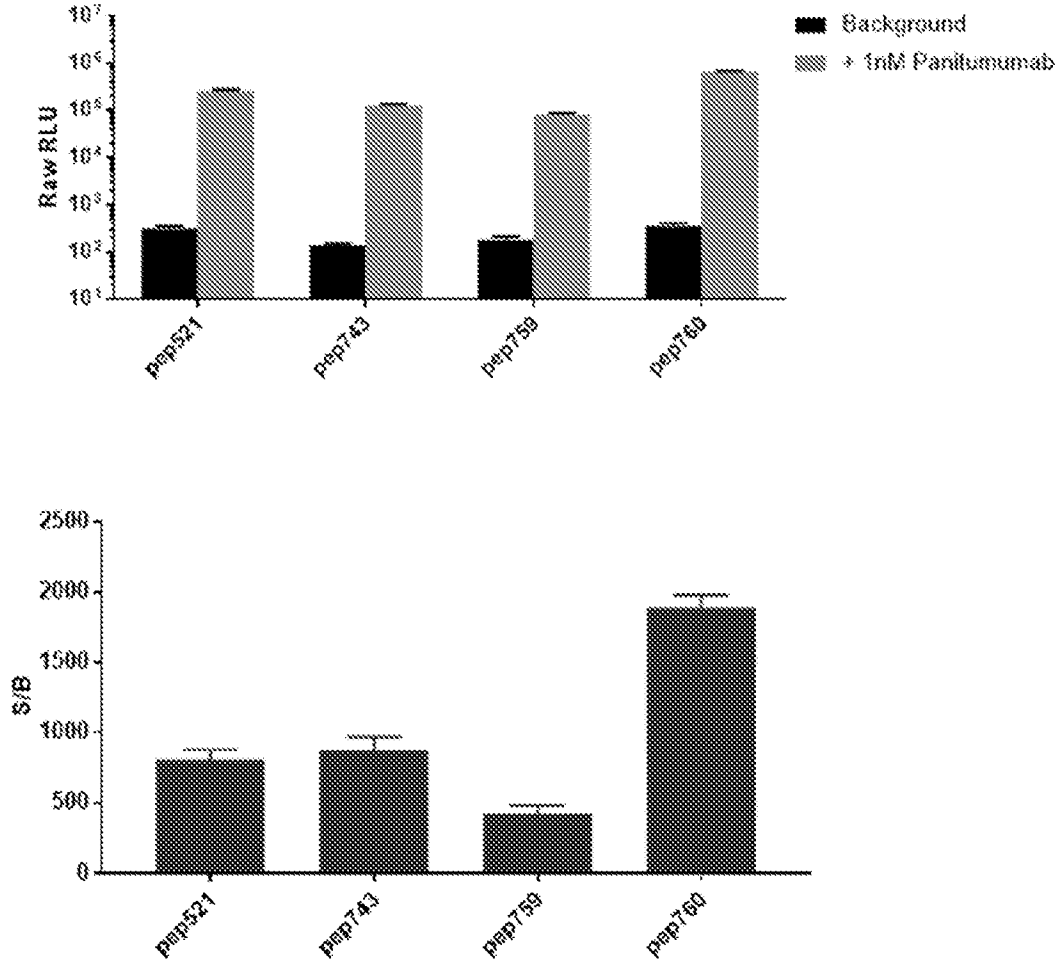

Background
+ 1nM Panitumumab
Raw RLU
pep521
pep743
pep759
pep760
S/B
2500
2000
1500
1000
500
0

FIG. 149

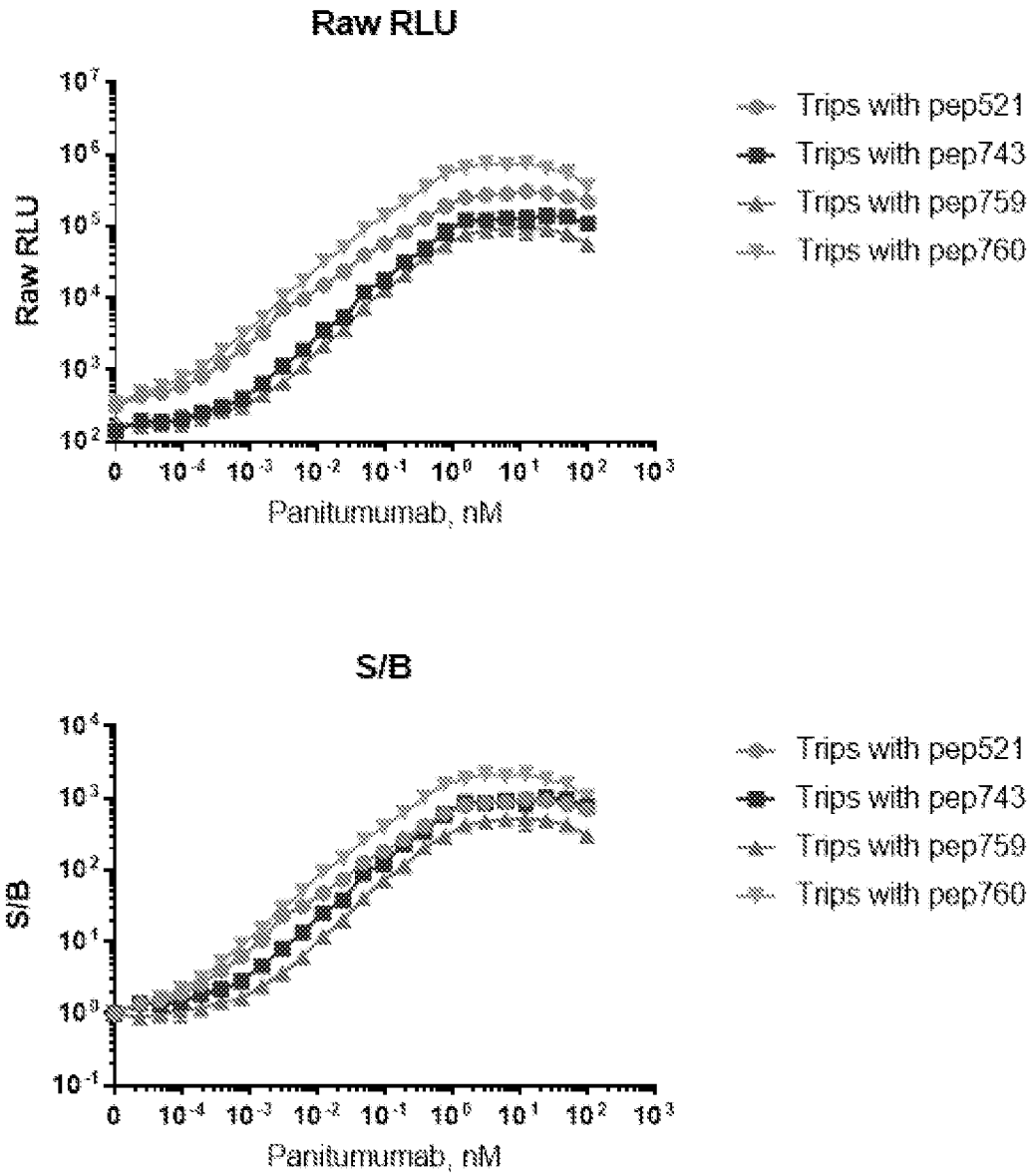

Raw RLU
Raw RLU
Panitumumab, nM
Trips with pep521
Trips with pep743
Trips with pep759
Trips with pep760
S/B
S/B
Panitumumab, nM
Trips with pep521
Trips with pep743
Trips with pep759
Trips with pep760

FIG. 150

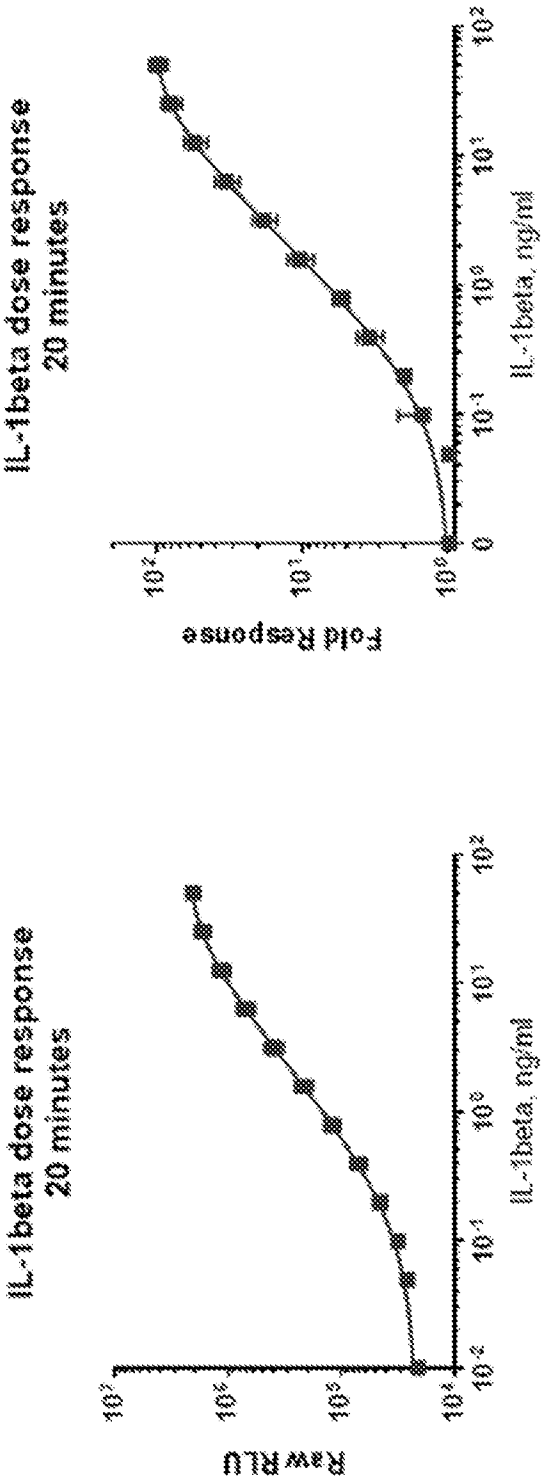

IL-1beta dose response
20 minutes
Raw RLU
IL-1beta, ng/ml
IL-1beta dose response
20 minutes
Fold Response
IL-1beta, ng/ml

FIG. 151

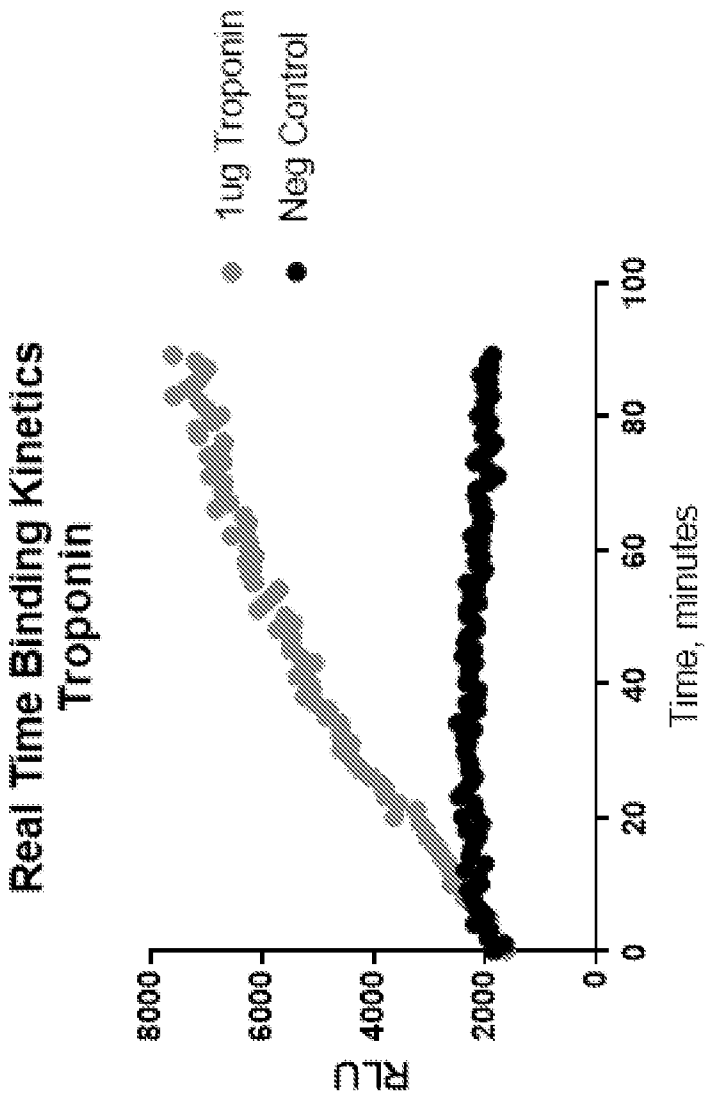

Real Time Binding Kinetics
Troponin
1ug Troponin
Neg Control
RLU
8000
6000
4000
2000
0
0
20
40
60
80
100
Time, minutes

FIG. 152

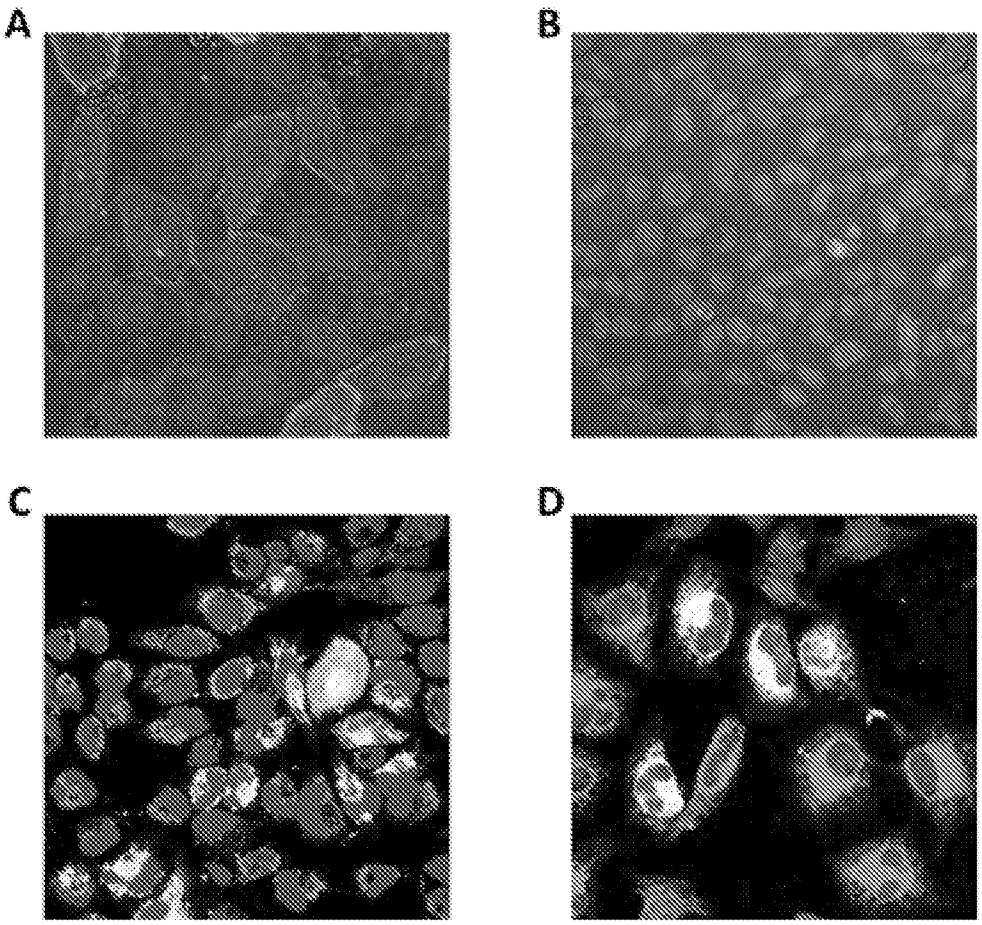


FIG. 153

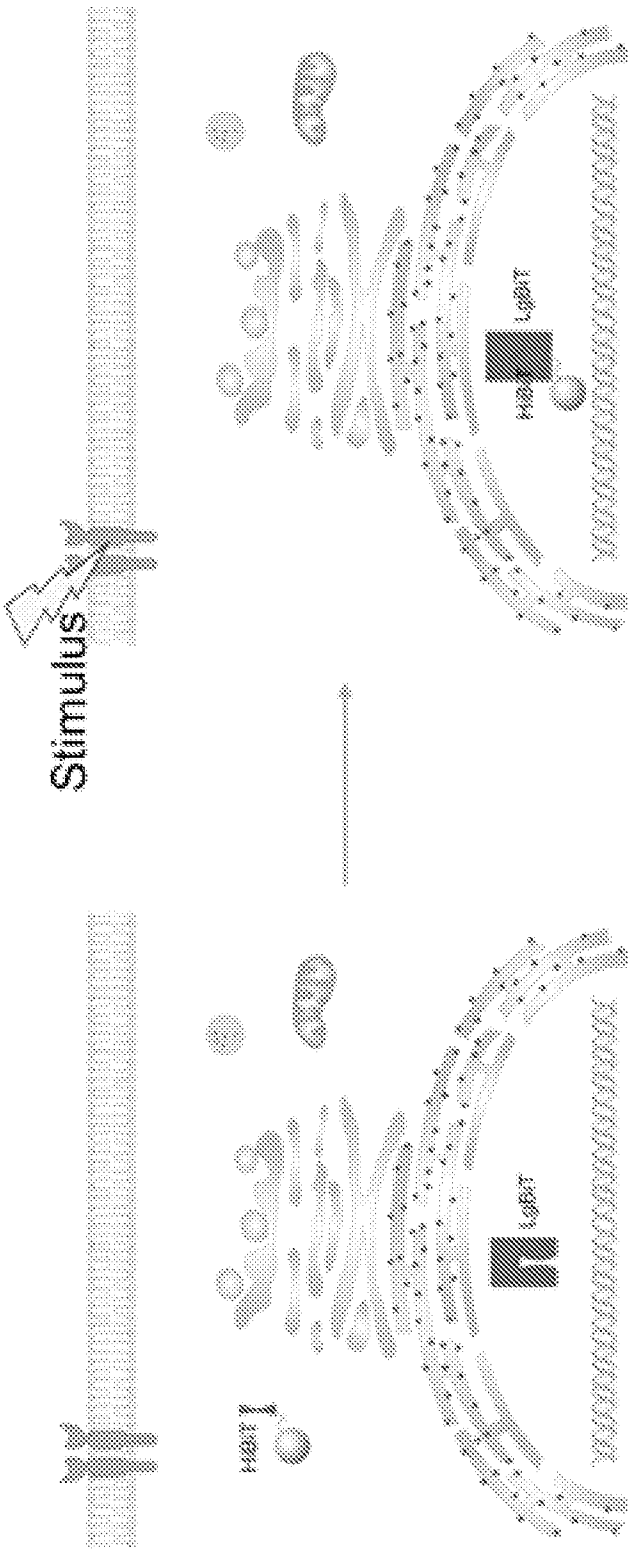


FIG. 154

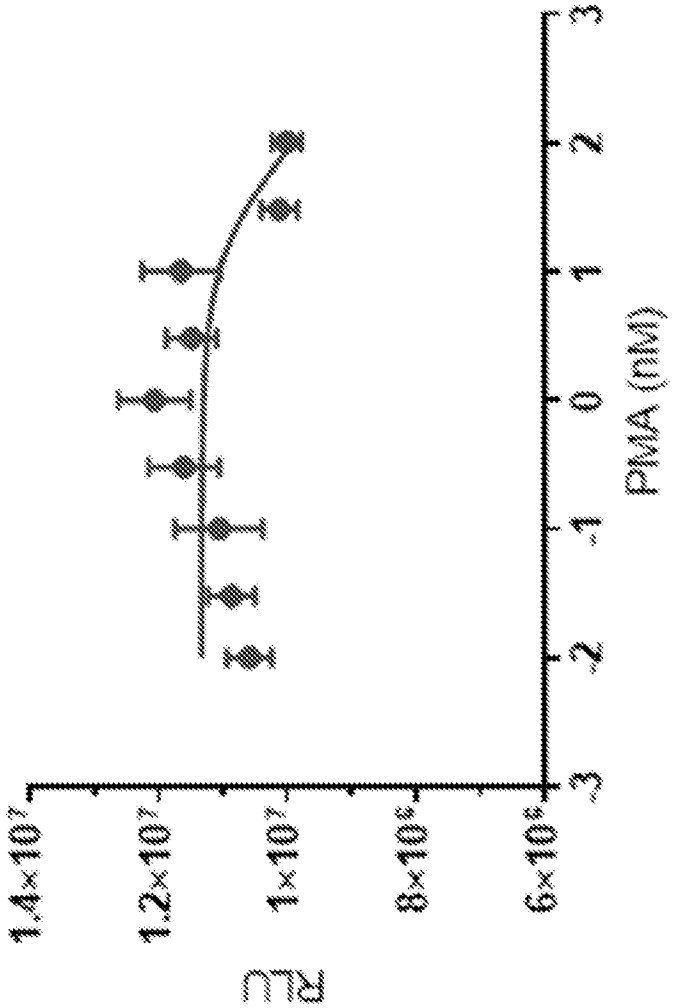


FIG. 155

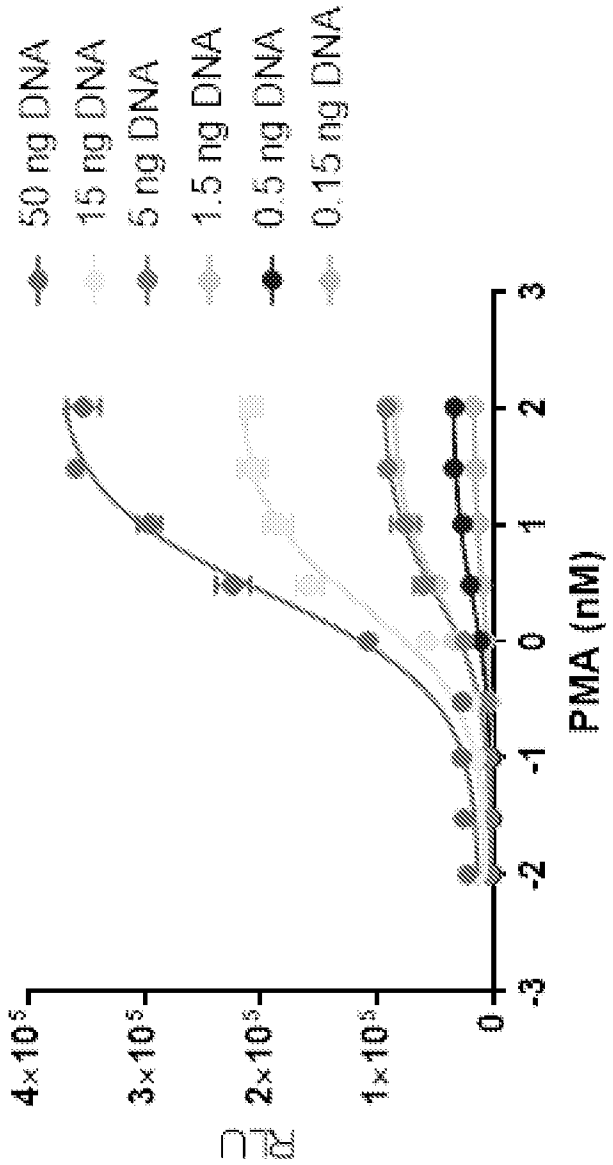

50 ng DNA
15 ng DNA
5 ng DNA
1.5 ng DNA
0.5 ng DNA
0.15 ng DNA
RLU
4×10^5
3×10^5
2×10^5
1×10^5
0
-3
-2
-1
0
1
2
3
PMA (nM)

FIG. 156

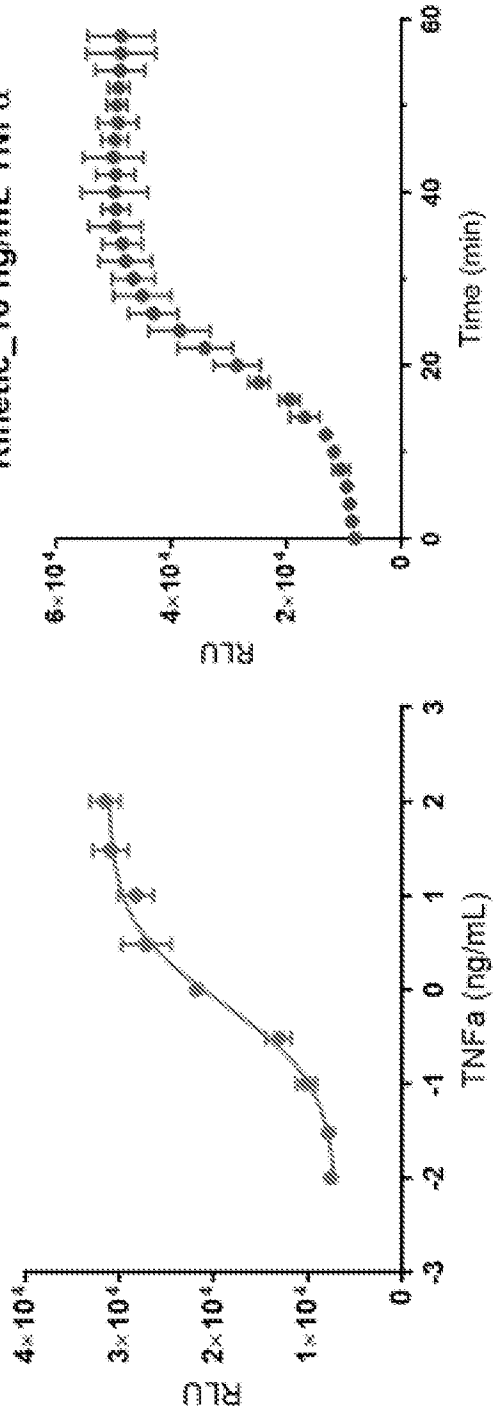


FIG. 157A

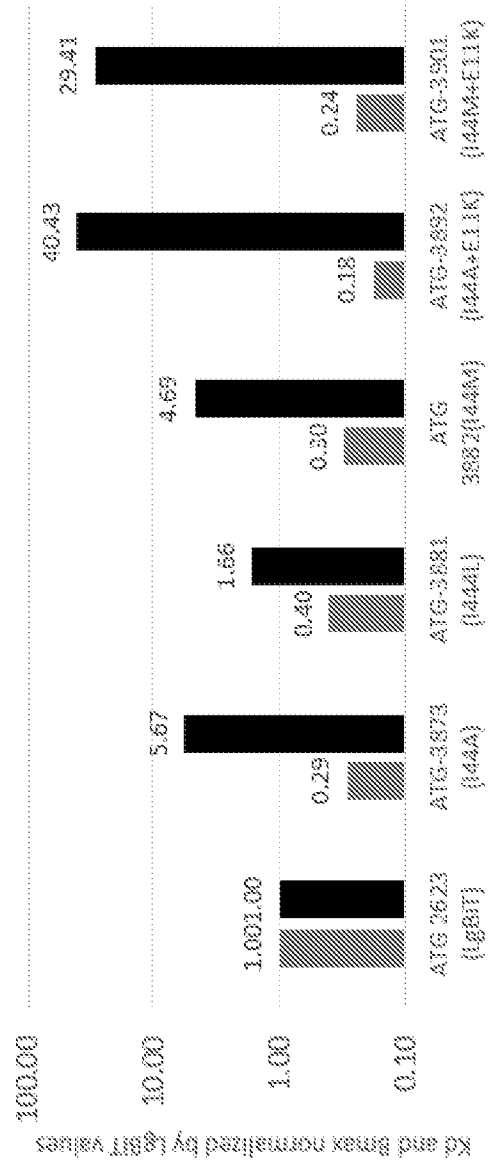
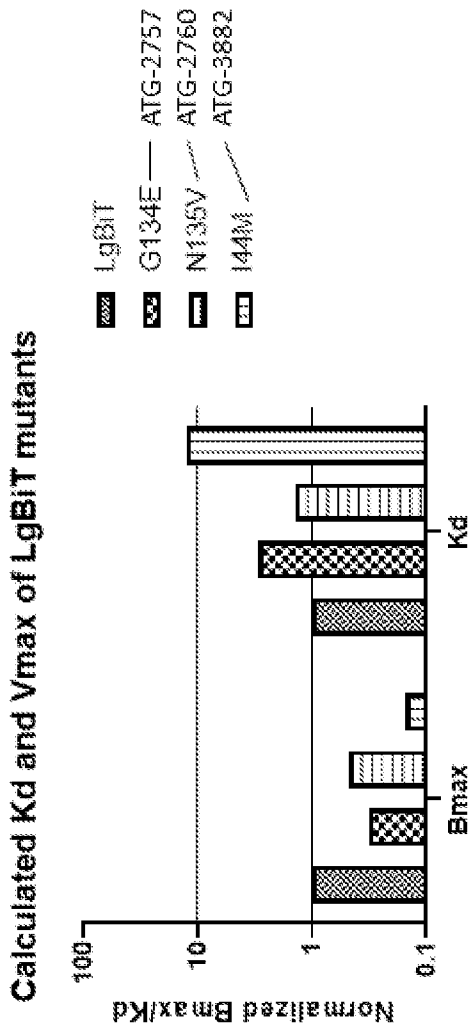

100.00
10.00
1.00
0.10
Kd and Bmax normalized by IgG1T values
1.00
1.00
0.29
5.67
0.40
1.66
0.30
4.69
0.18
40.43
0.24
29.41
ATG 2623 (IgG1T)
ATG-3873 (144A)
ATG-3881 (144M)
ATG 3887(144M)
ATG-3892 (144A+E11K)
ATG-3901 (144M+E11K)
Bmax
Kd

FIG. 157B

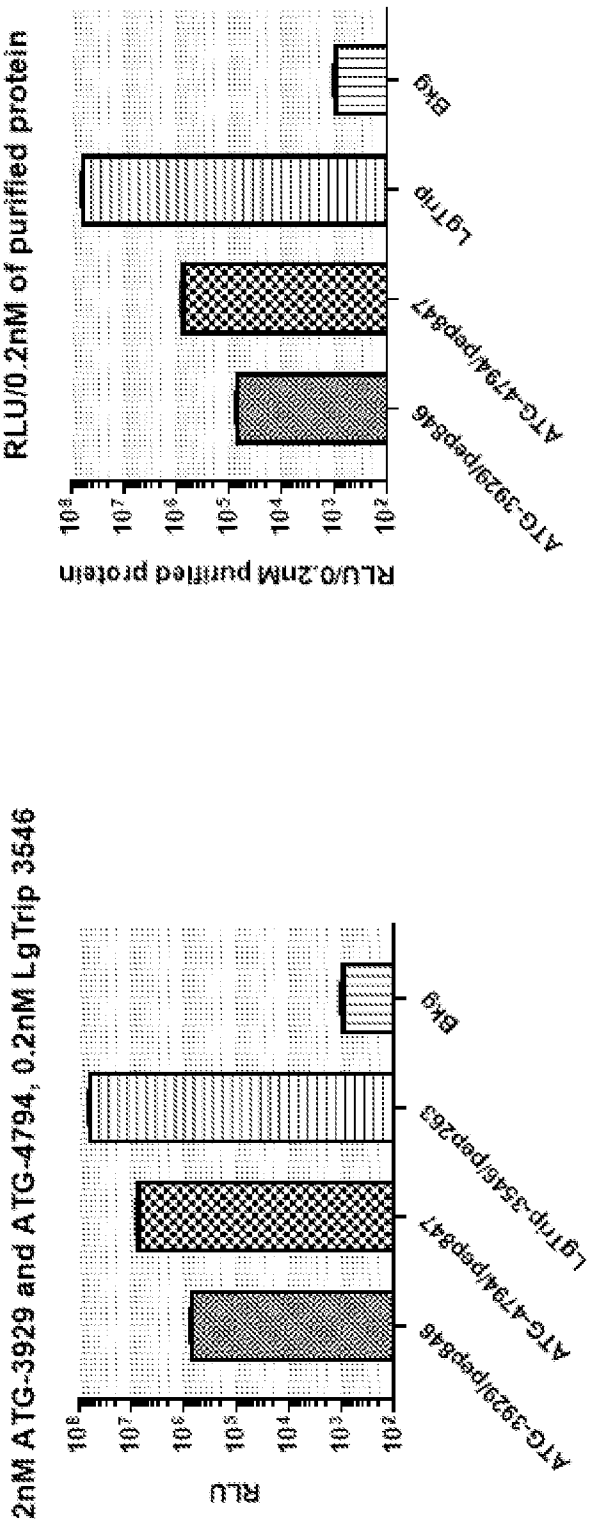
Clone	Mutations	Kd (normalized)	Bmax (normalized
LgBiT		1	1
ATG-2724	I44M	3	0.57
ATG-3901	**I44M+E11K**	**16**	**0.53**
ATG-3945	I44M+E11K+G134D+N135V	44	0.23
ATG-3984	**I44M+E11K+G134D+N135V+L150S**	**1958**	**0.12**
ATG-4147	**I44M+E11K+L150S**	**613**	**0.26**
ATG-4166	**I44M+E11K+N135V+L150S**	**1296**	**0.12**

FIG. 158



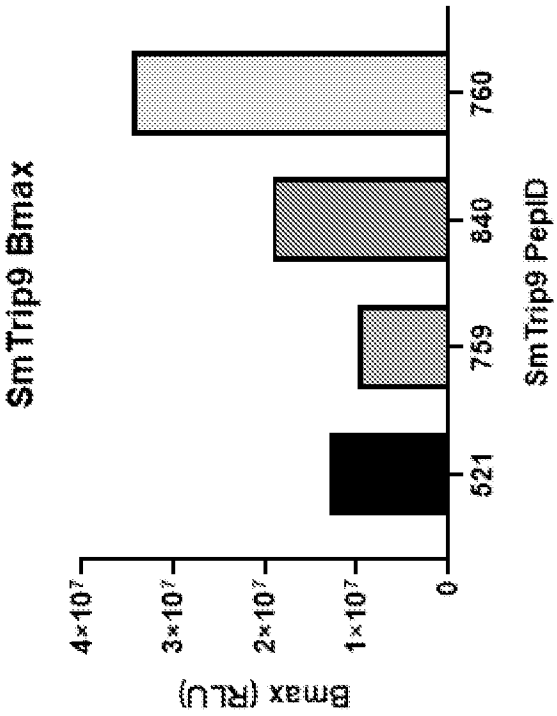
	LgBiT	G134E	N135V	I44M
Bmax	1.0	0.3	0.5	0.2
Kd	1.0	2.9	1.4	12.6

FIG. 159



Sample ID	Sample ID	peptide purity
ATG-3929	LgTrip3546 remove strands 6,7,8	
ATG-4794	LgTrip3546 remove strands 5,6,7,8	
ATG-3546	LgTrip	
pep846	Synthetic strands 6-10	71%
pep847	Synthetic strands 5-10	>95%
pep263	Synthetic peptide strands 9+10	>95%

FIG. 160


SmTrip9 Bmax
Bmax (RLU)
4×10^7
3×10^7
2×10^7
1×10^7
0
521
759
840
760
SmTrip9 PeptID

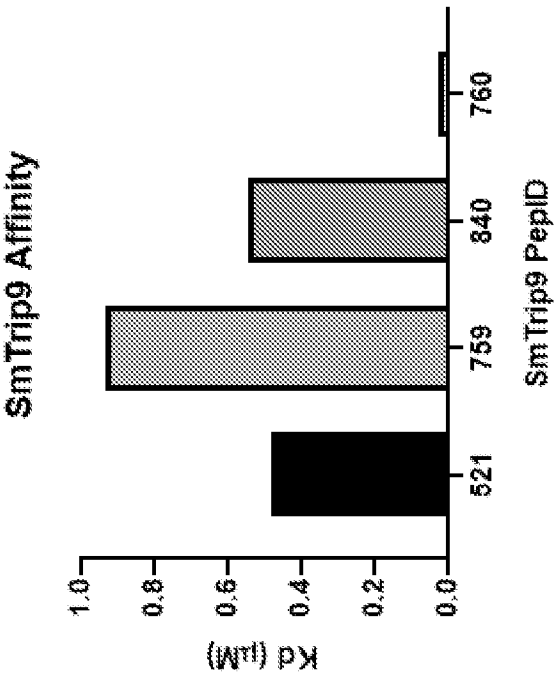
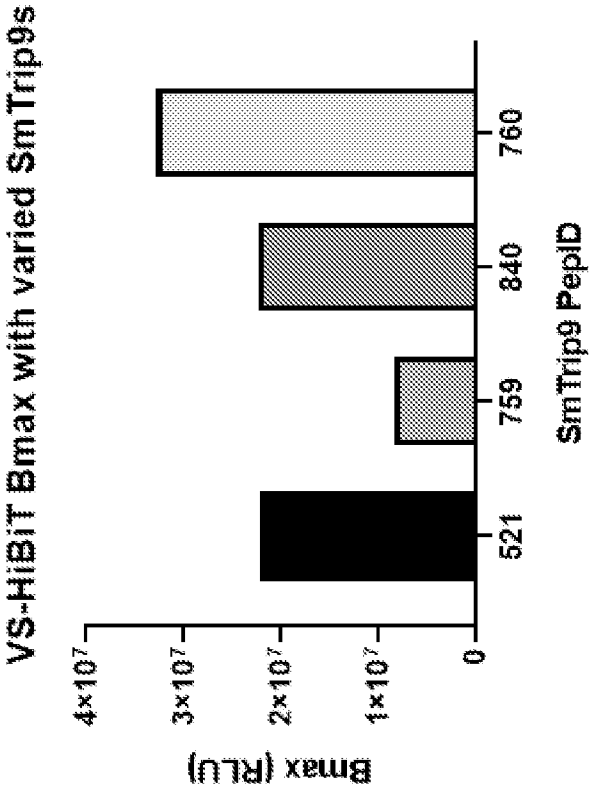

SmTrip9 Affinity
Kd (μM)
1.0
0.8
0.6
0.4
0.2
0.0
521
759
840
760
SmTrip9 PeptID

FIG. 161


VS-HiBiT Bmax with varied SmTrip9s
Bmax (RLU)
4×10^7
3×10^7
2×10^7
1×10^7
0
521
759
840
760
SmTrip9 PepID

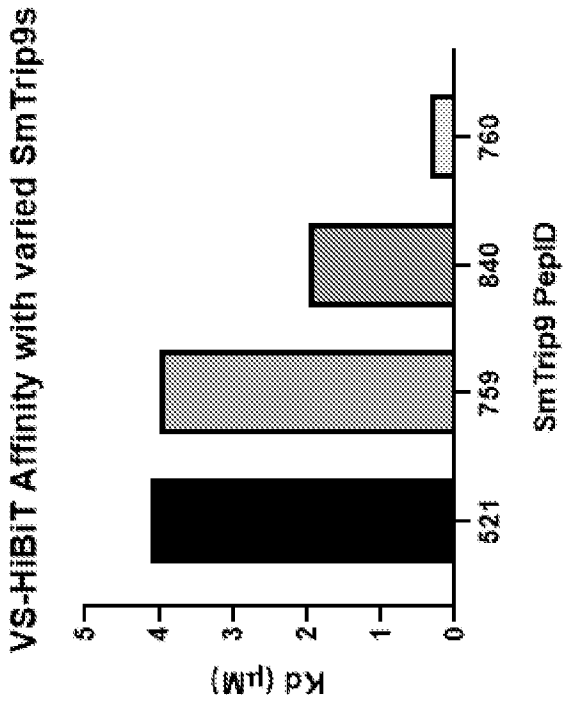

VS-HiBiT Affinity with varied SmTrip9s
Kd (μM)
5
4
3
2
1
0
521
759
840
760
SmTrip9 PepID

FIG. 162

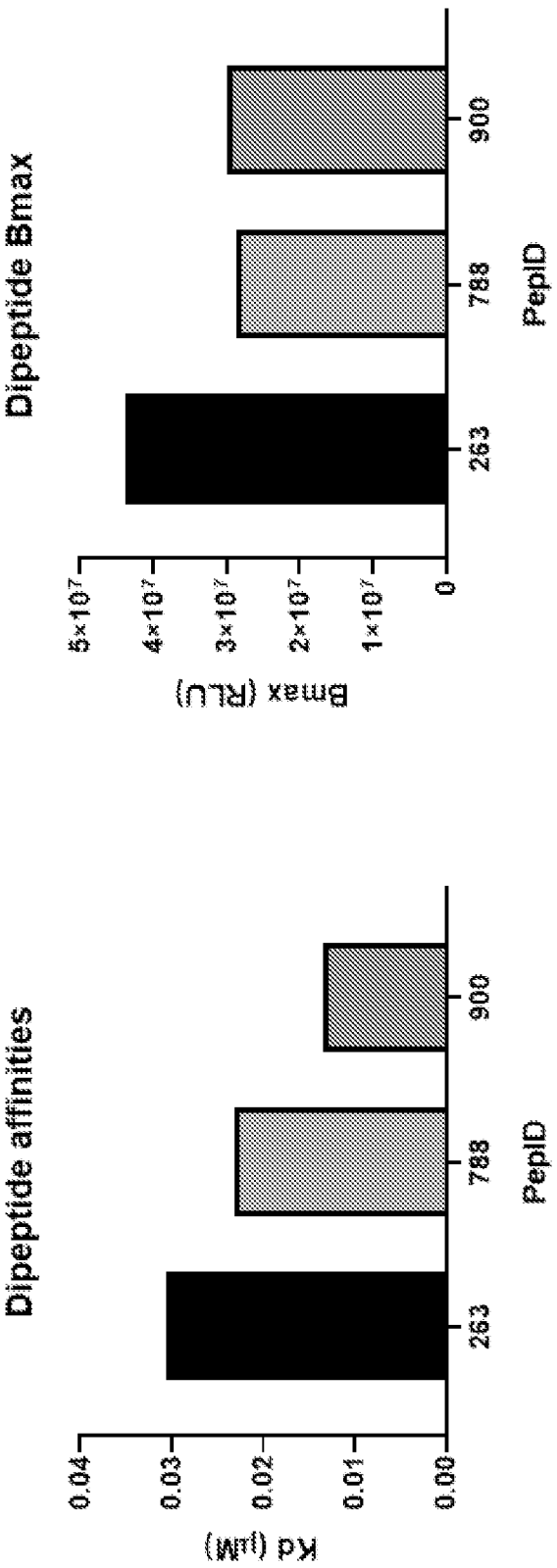

Dipeptide Bmax
Bmax (RLU)
5×10^7
4×10^7
3×10^7
2×10^7
1×10^7
0
263
788
900
PepID
Dipeptide affinities
Kd (μM)
0.04
0.03
0.02
0.01
0.00
263
788
900
PepID

FIG. 163

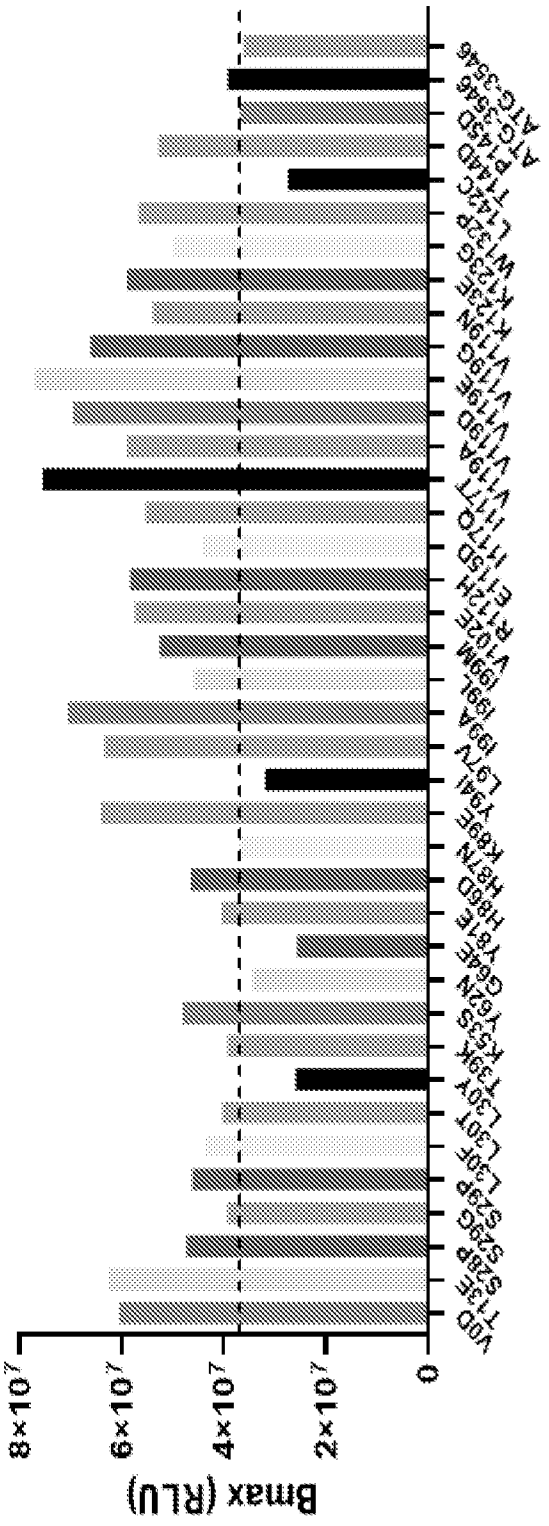

Pep788 Bmax with LgTrip lysates
Bmax (RLU)
8×10^7
6×10^7
4×10^7
2×10^7
0
VDD
T13E
S28P
S29G
S29P
L30F
L30T
T39Y
T39K
K53S
Y62N
G64E
Y84E
H86D
H87N
K89E
Y94I
L97V
I99A
I99L
I99M
V102E
R113H
E115D
I117D
I117T
V119A
V119D
V119E
V119G
K119N
K123E
K123G
W132P
L142C
T144C
P145D
ATG-3546
ATG-3546

FIG. 164

FIG. 165

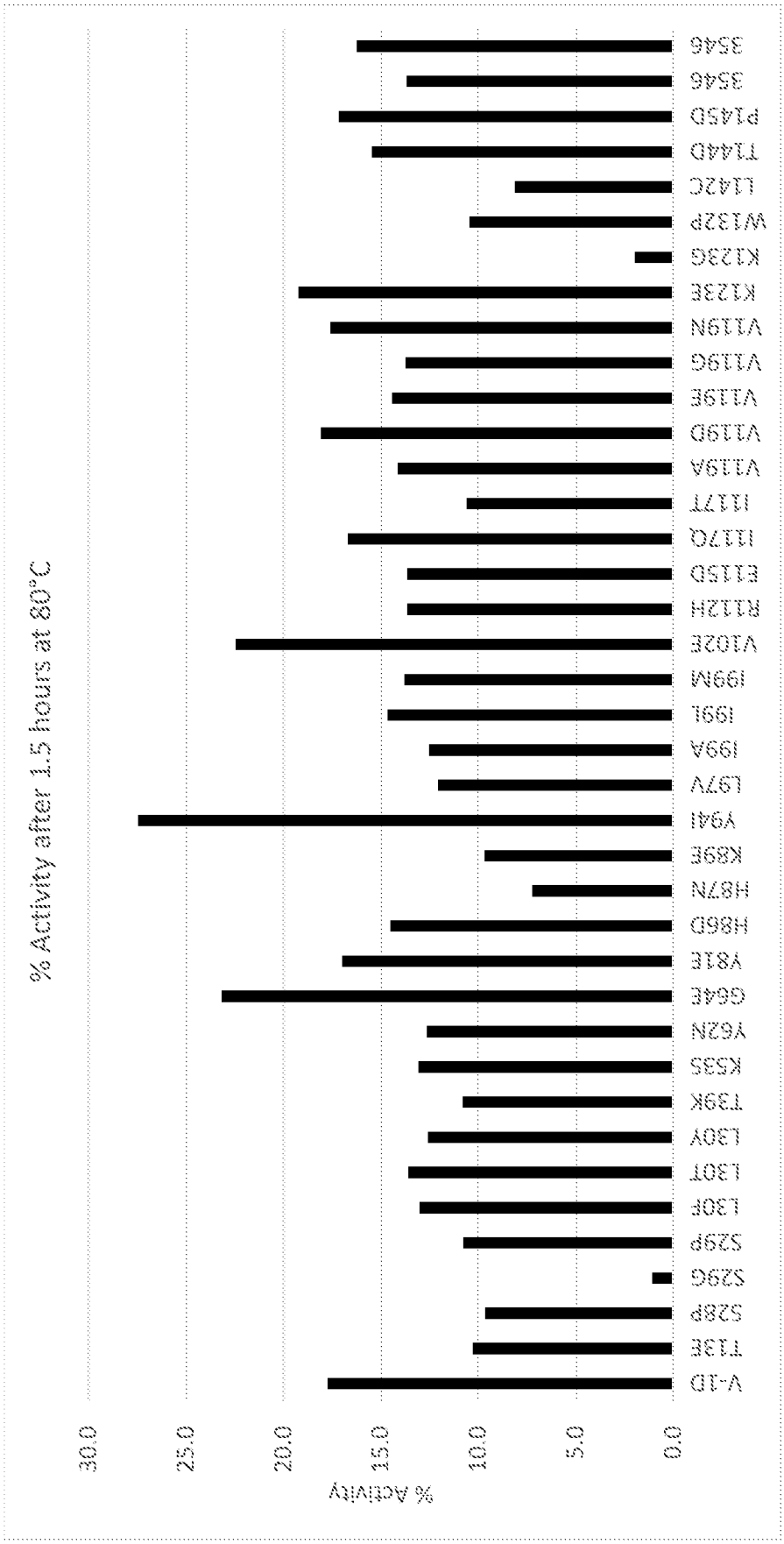

% Activity after 1.5 hours at 80°C
% Activity
30.0
25.0
20.0
15.0
10.0
5.0
0.0
V-1D
T13E
S28P
S29G
S29P
L30P
L30T
L30V
T39K
K53S
Y6ZN
G64E
Y81E
H86D
H87N
K89E
Y94I
L97V
I99A
I99L
I99M
V102E
R112H
E115D
I117Q
I117T
V119A
V119D
V119E
V119G
V119N
K123E
K123G
W132P
L142C
T144D
P145D
354G
354G

FIG. 166

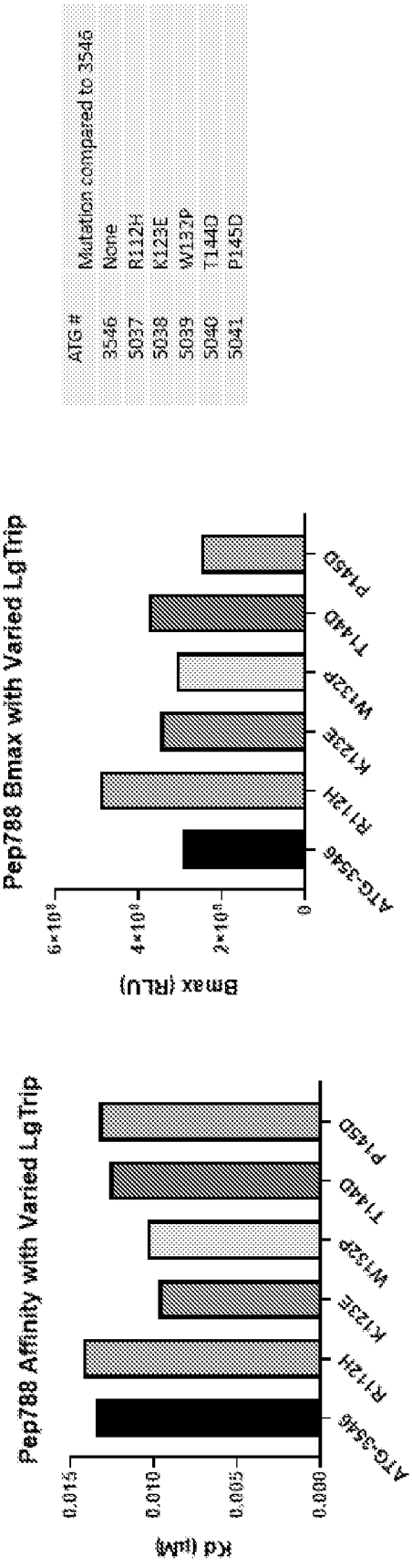
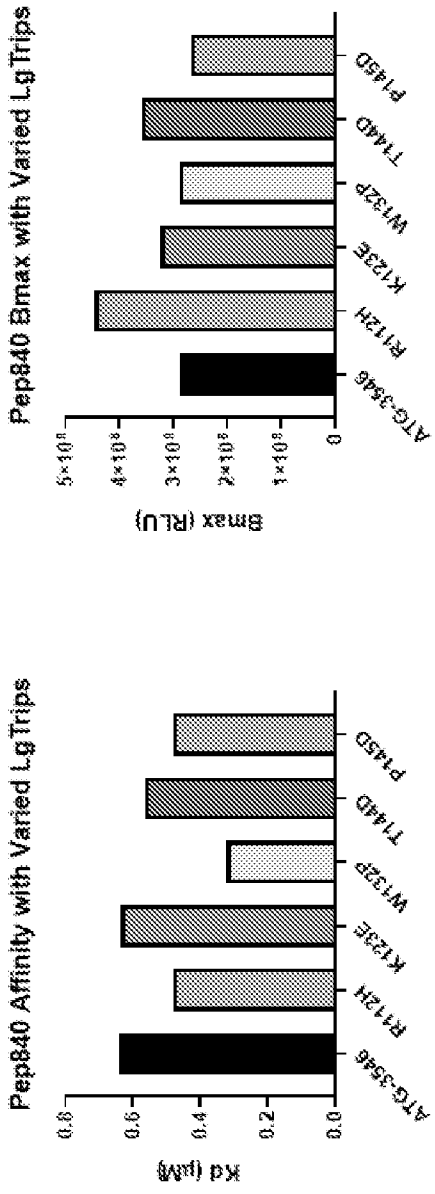
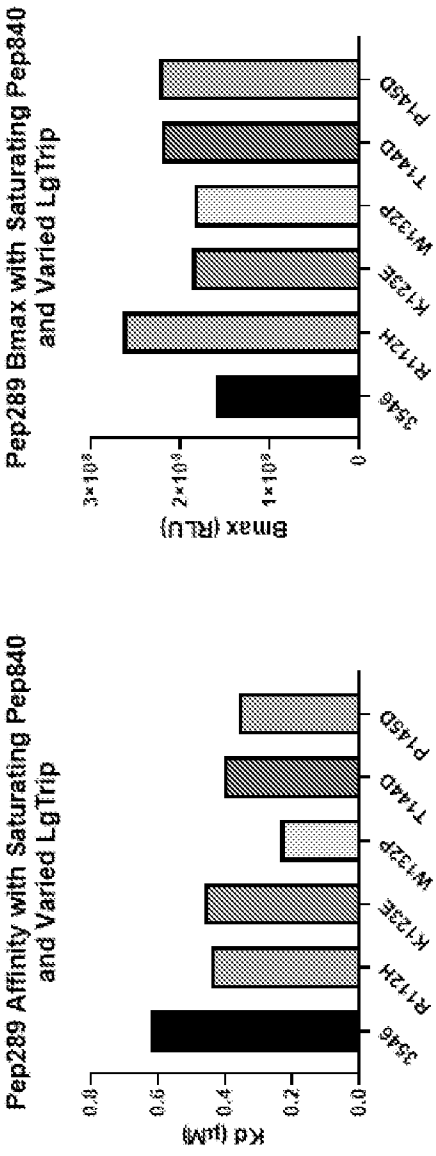

ATG #
Mutation compared to 3546
3546
None
5037
R112H
5038
K123E
5039
W132P
5040
T144D
5041
P145D
Pep788 Bmax with Varied Lg Trip
Bmax (RLU)
6×10⁸
4×10⁸
2×10⁸
0
ATG-3546
R112H
K123E
W132P
T144D
P145D
Pep788 Affinity with Varied Lg Trip
Kd (uM)
0.015
0.010
0.005
0.000
ATG-3546
R112H
K123E
W132P
T144D
P145D

FIG. 167



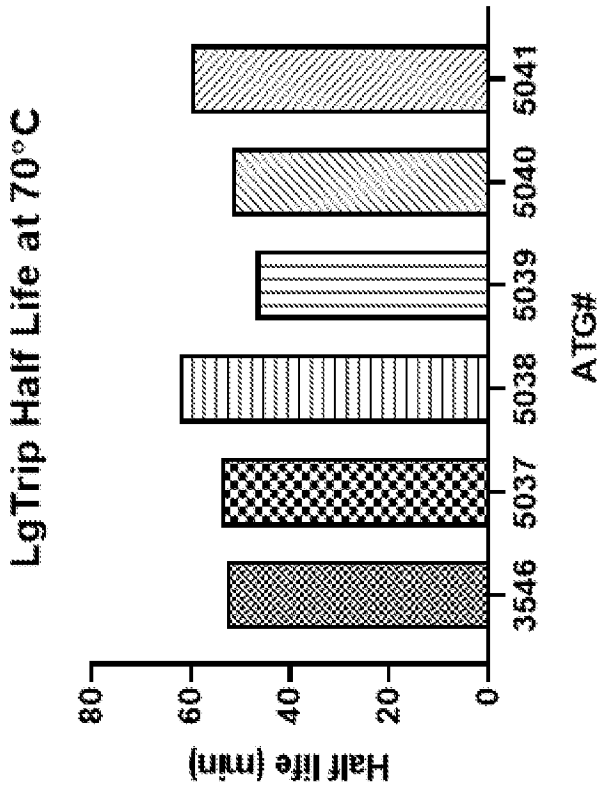
ATG #	Mutation compared to 3546
3546	None
5037	R112H
5038	K123E
5039	W132P
5040	T144D
5041	P145D

FIG. 168



ATG #	Mutation compared to 3546
3546	None
5037	R112H
5038	K123E
5039	W132P
5040	T144D
5041	P145D

FIG. 169



ATG #	Mutation compared to 3546
3546	None
5037	R112H
	K123E
5038	
5039	W132P
5040	T144D
5041	P145D

FIG. 170

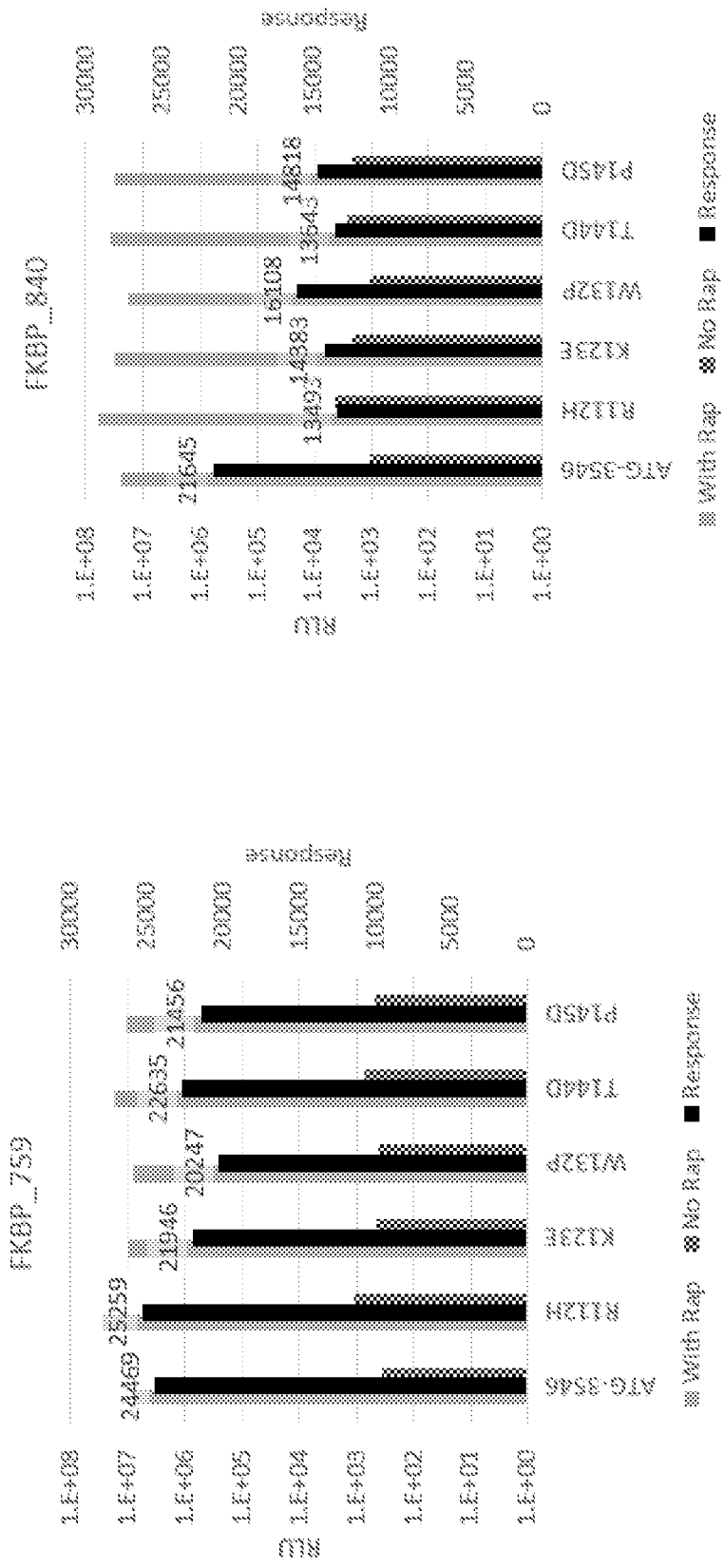

FKBP_840
RLU
Response
1.E+08
1.E+07
1.E+06
1.E+05
1.E+04
1.E+03
1.E+02
1.E+01
1.E+00
30000
25000
20000
15000
10000
5000
0
21645
13497
14383
16108
13643
14618
ATG-3546
R112H
K123E
W132P
T144D
P145D
With Rap
No Rap
Response
FKBP_759
RLU
Response
24469
26259
21946
20247
22635
21456
ATG-3546
R112H
K123E
W132P
T144D
P145D
With Rap
No Rap
Response

FIG. 171

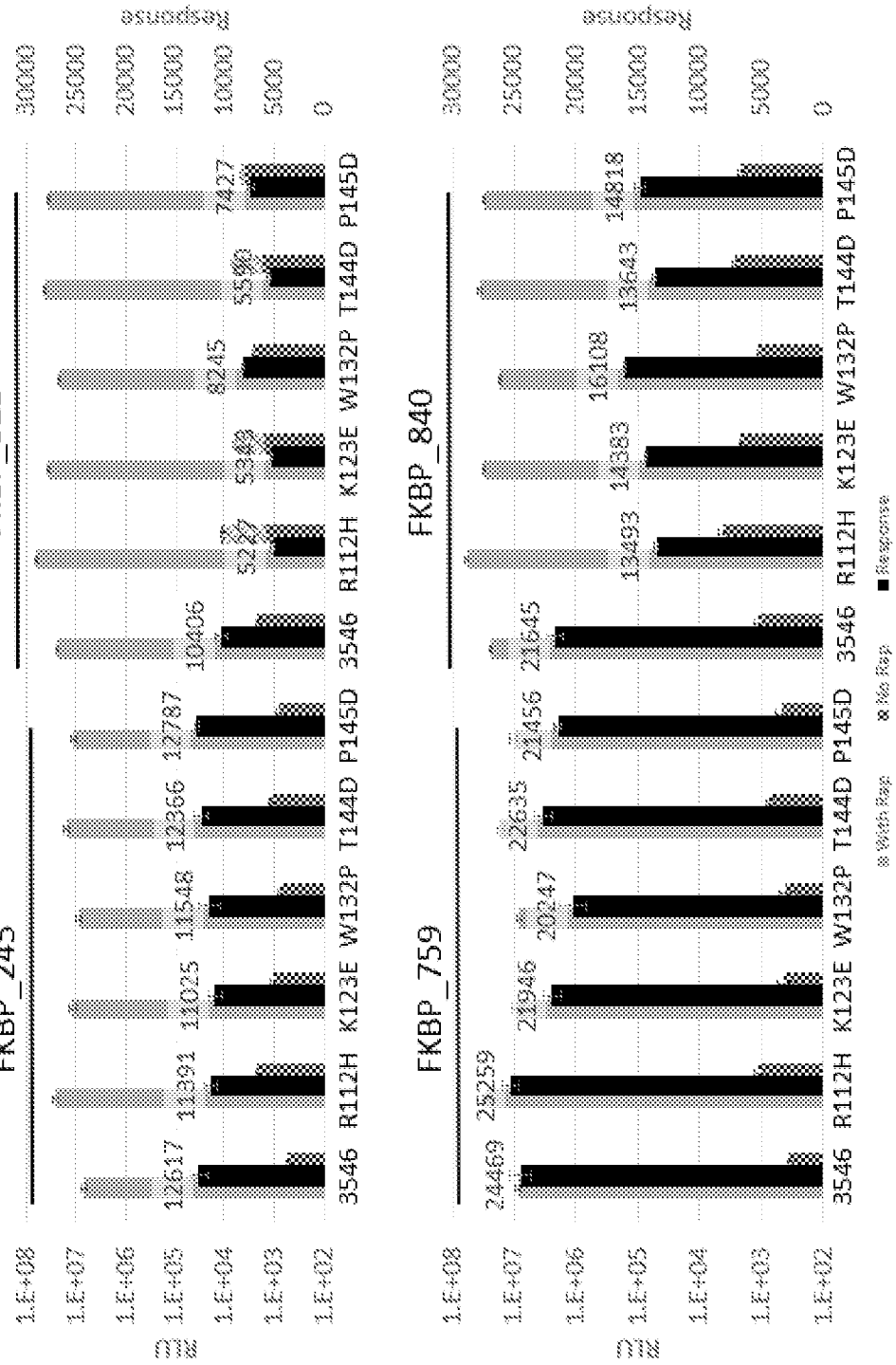

FKBP_245
FKBP_521
FKBP_759
FKBP_840
RLU
Response
1.E+08
1.E+07
1.E+06
1.E+05
1.E+04
1.E+03
1.E+02
30000
25000
20000
15000
10000
5000
0
12617
11391
11025
11548
12366
12787
10406
6245
7427
24469
25259
21946
20247
22635
21456
21645
13493
14383
16108
13643
14818
3546
R112H
K123E
W132P
T144D
P145D
With flap
No flap
Response

FIG. 172

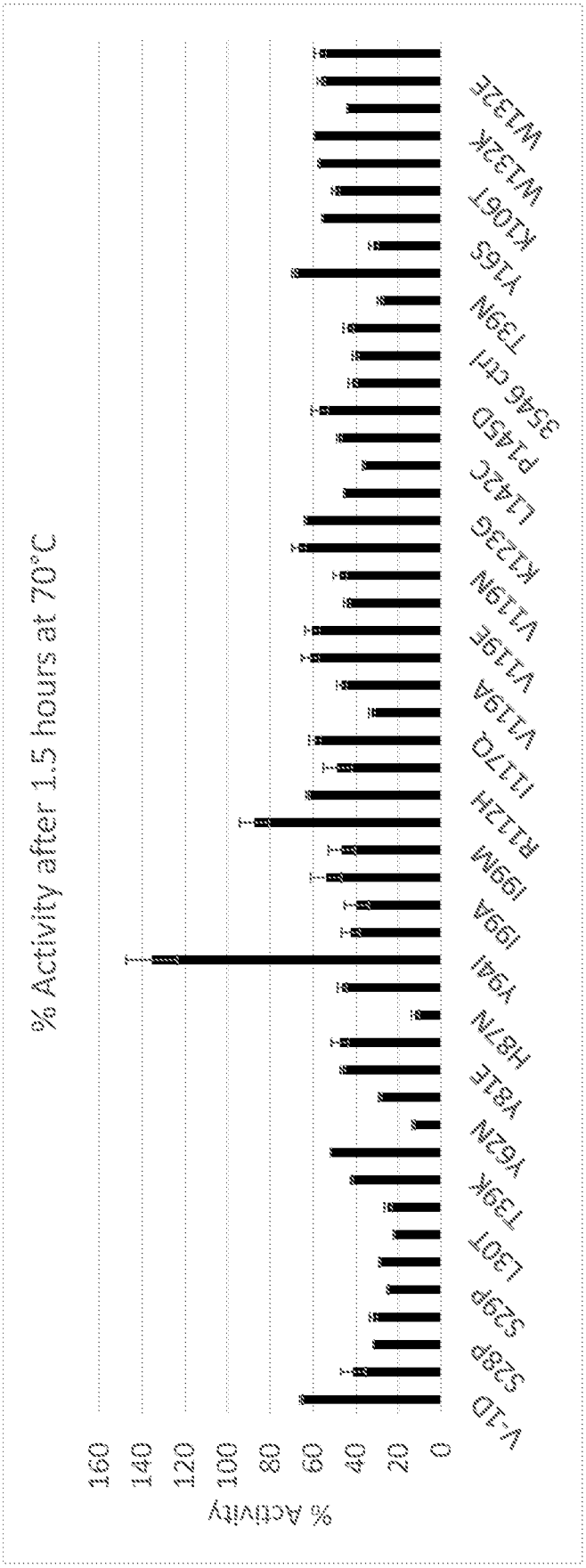

% Activity after 1.5 hours at 70°C
% Activity
0
20
40
60
80
100
120
140
160
V-1D
S28P
S29P
L30T
T39K
Y62N
V83E
H87N
Y94I
V96I
I99M
R112H
I117Q
V119A
V136E
V139N
K235G
L242C
P246D
3546 cm
T39N
V165
K206T
W132K
W132E

FIG. 173

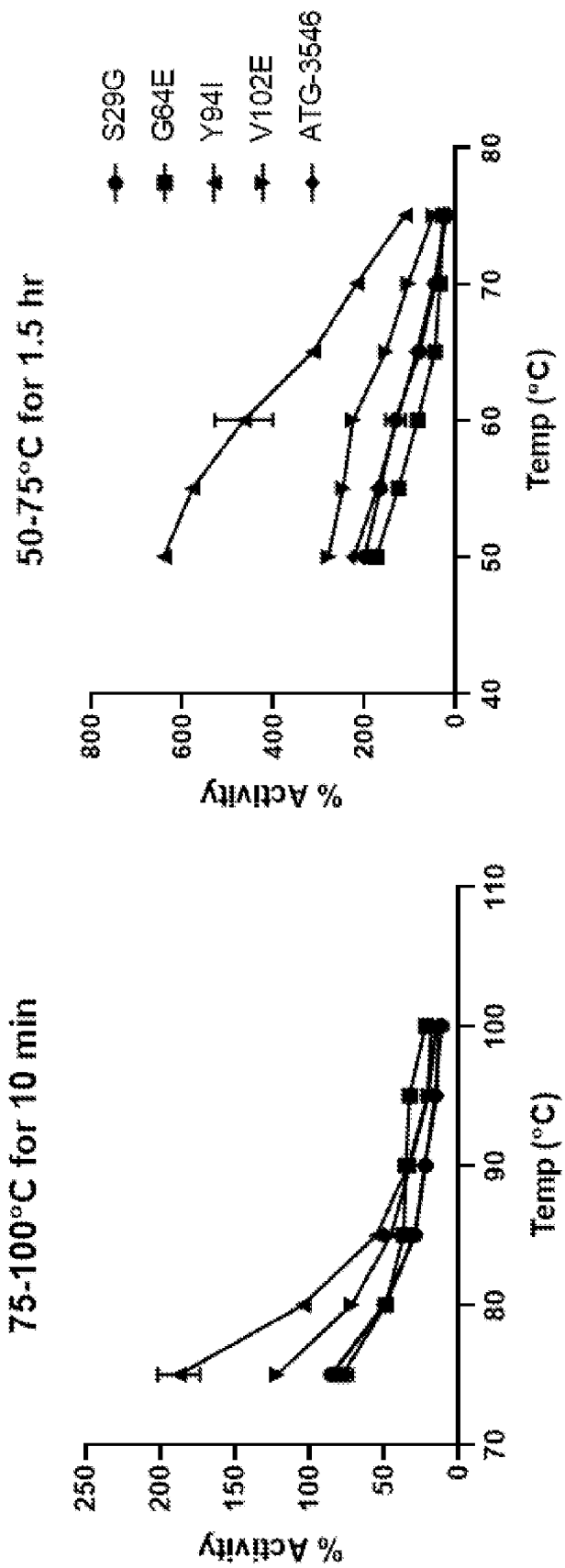

50-75°C for 1.5 hr
S29G
G64E
Y94I
V102E
ATG-3546
% Activity
Temp (°C)
75-100°C for 10 min
% Activity
Temp (°C)

FIG. 174

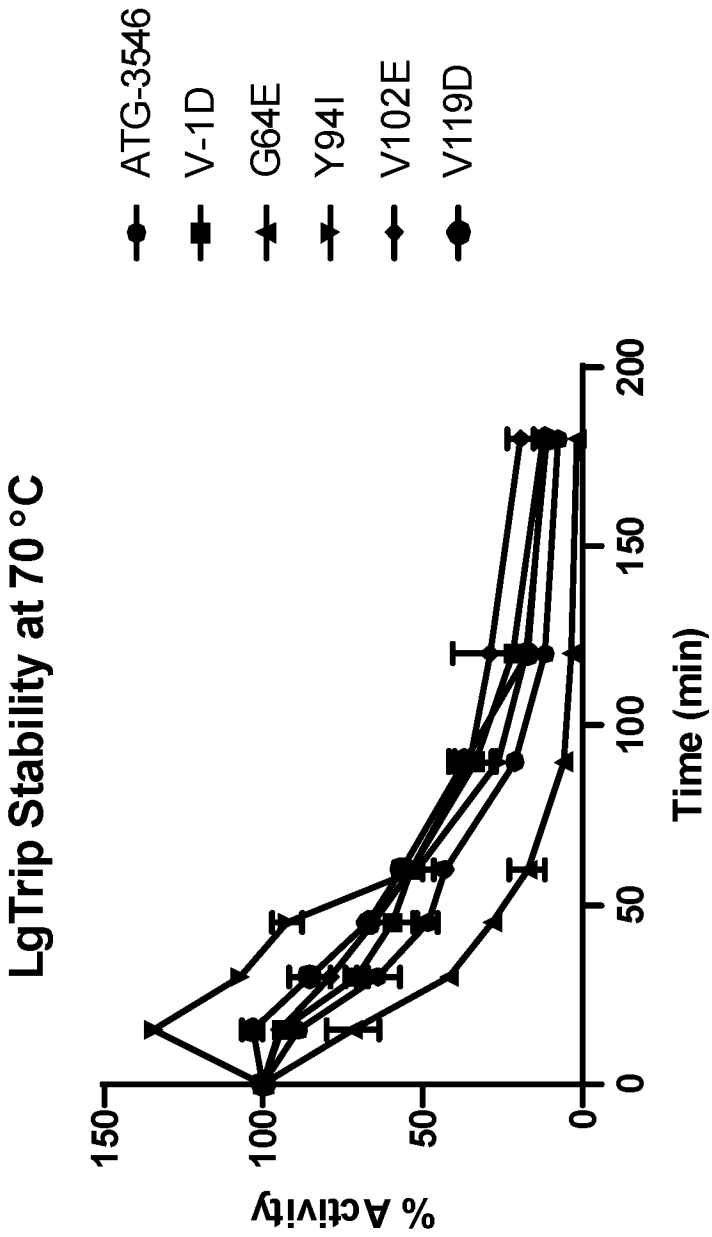
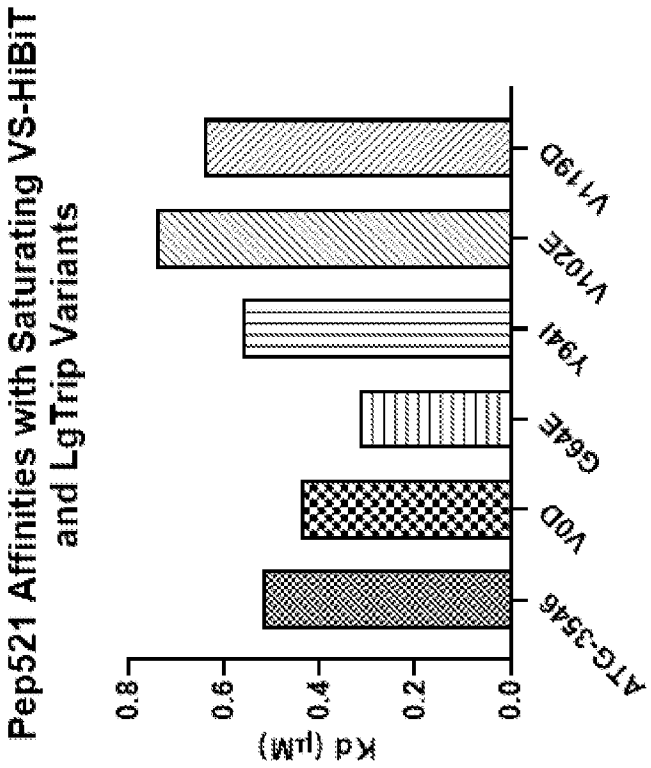

LgTrip Stability at 70 °C
ATG-3546
V-1D
G64E
Y94I
V102E
V119D
% Activity
150
100
50
0
0
50
100
150
200
Time (min)

FIG. 175


Pep521 Affinities with Saturating VS-HiBiT and LgTrip Variants
Kd (μM)
0.0
0.2
0.4
0.6
0.8
ATG-3546
V0D
G64E
Y94I
V102E
V119D

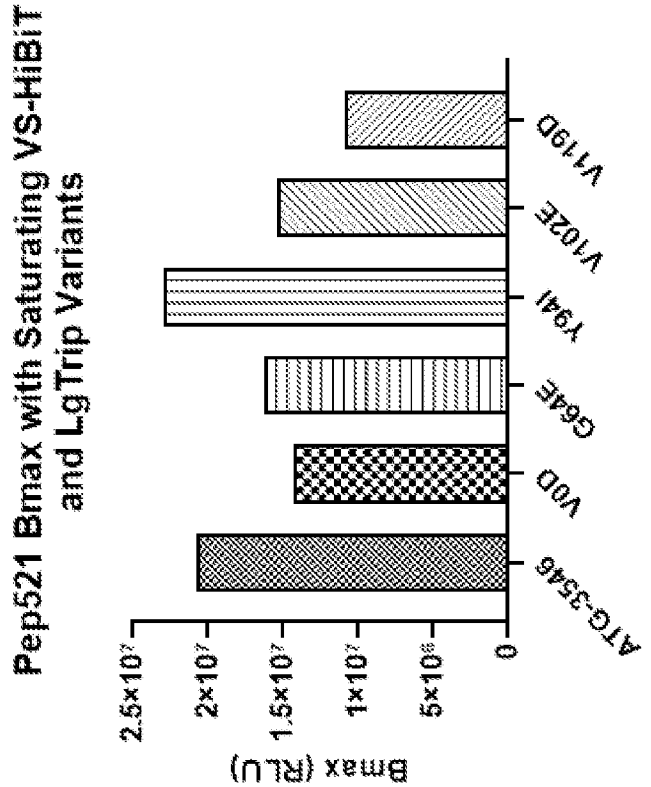
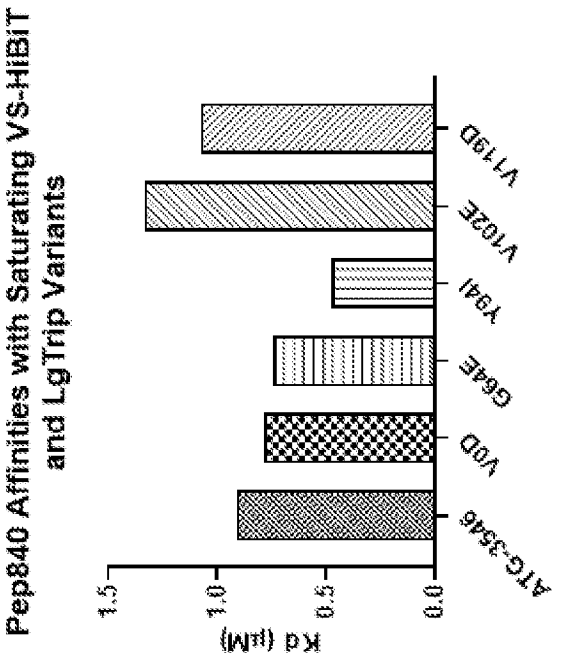

Pep521 Bmax with Saturating VS-HiBiT and LgTrip Variants
Bmax (RLU)
0
5×10⁶
1×10⁷
1.5×10⁷
2×10⁷
2.5×10⁷
ATG-3546
V0D
G64E
Y94I
V102E
V119D

FIG. 176


Pep840 Affinities with Saturating VS-HiBiT and LgTrip Variants
Kd (μM)
1.5
1.0
0.5
0.0
ATG-3546
V0D
G64E
Y94I
V102E
V119D

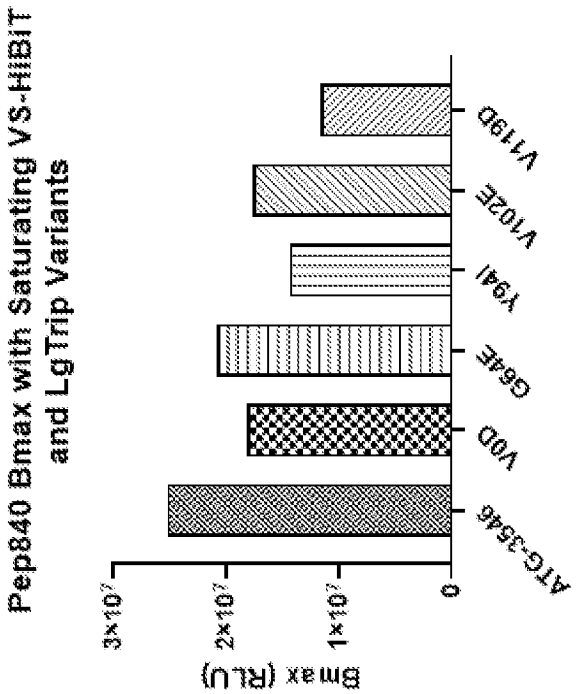

Pep840 Bmax with Saturating VS-HiBiT and LgTrip Variants
Bmax (RLU)
3×10^7
2×10^7
1×10^7
0
ATG-3546
V0D
G64E
Y94I
V102E
V119D

FIG. 177

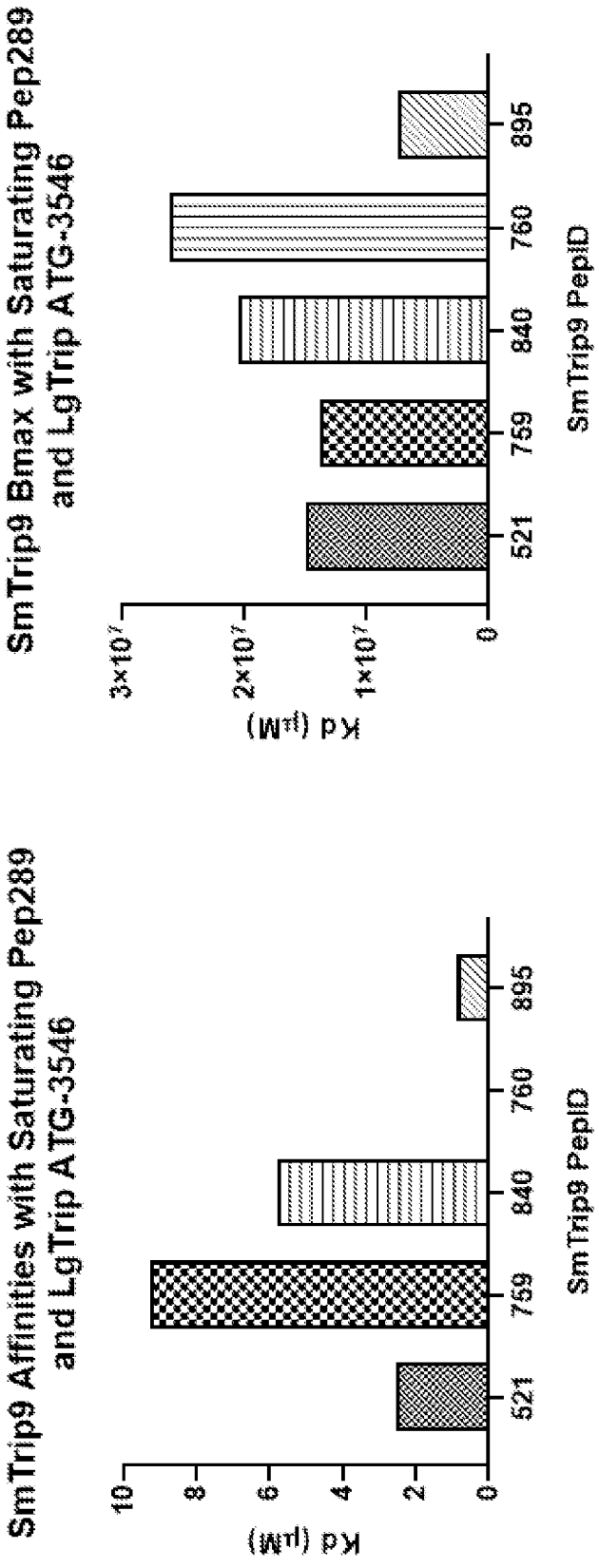
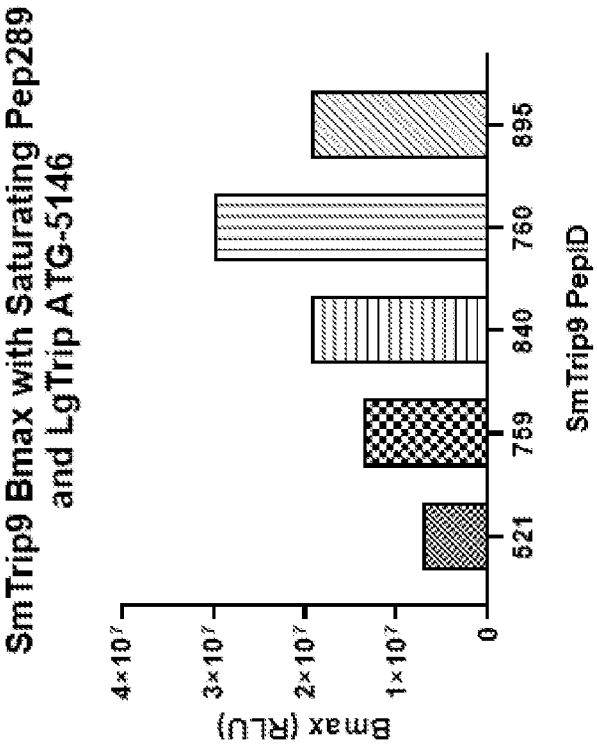

SmTrip9 Bmax with Saturating Pep289 and LgTrip ATG-3546
Kd (μM)
3×10⁷
2×10⁷
1×10⁷
0
521
759
840
760
895
SmTrip9 PepID
SmTrip9 Affinities with Saturating Pep289 and LgTrip ATG-3546
Kd (μM)
10
8
6
4
2
0
521
759
840
760
895
SmTrip9 PepID

FIG. 178


SmTrip9 Bmax with Saturating Pep289 and LgTrip ATG-5146
Bmax (RLU)
4×10⁷
3×10⁷
2×10⁷
1×10⁷
0
521
759
840
760
895
SmTrip9 PepID

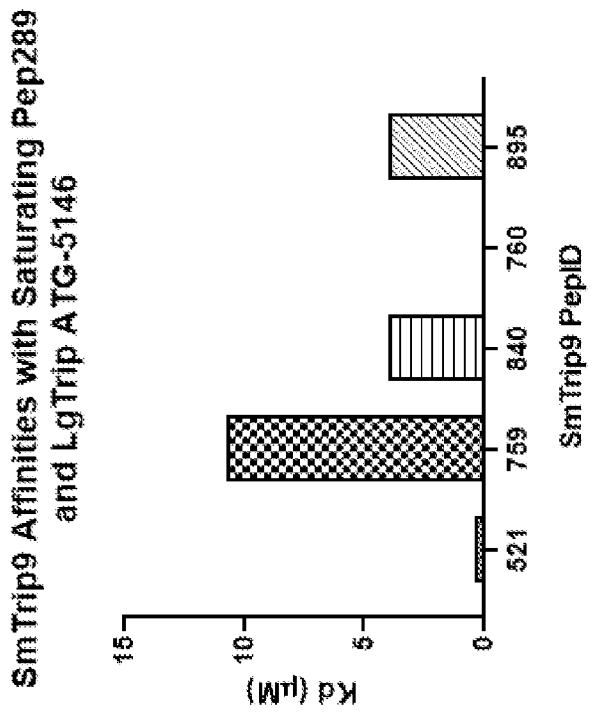
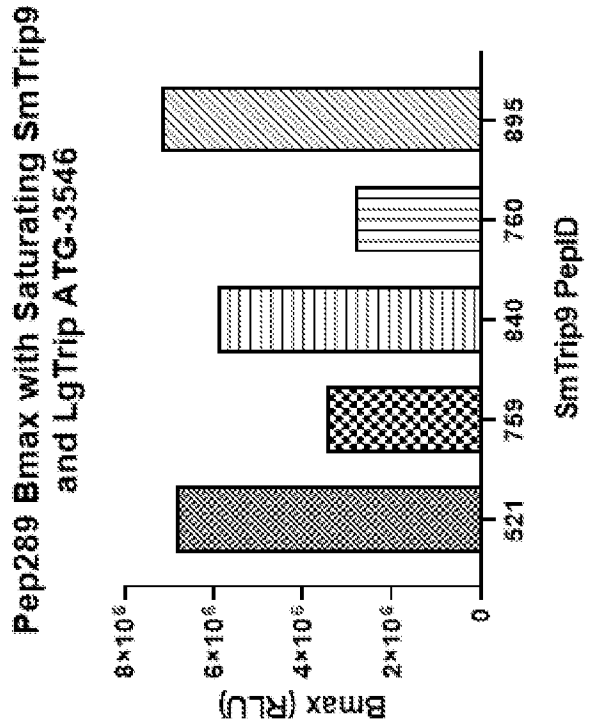

SmTrip9 Affinities with Saturating Pep289 and LgTrip ATG-5146
Kd (μM)
15
10
5
0
521
759
840
760
895
SmTrip9 PepID

FIG. 179


Pep289 Bmax with Saturating SmTrip9 and LgTrip ATG-3546
Bmax (RLU)
8×10^6
6×10^6
4×10^6
2×10^6
0
521
759
840
760
895
SmTrip9 PeptD

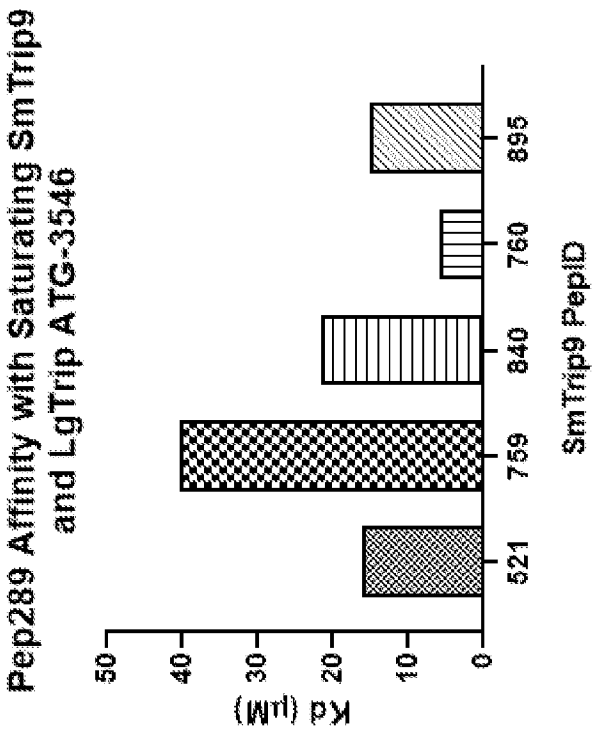
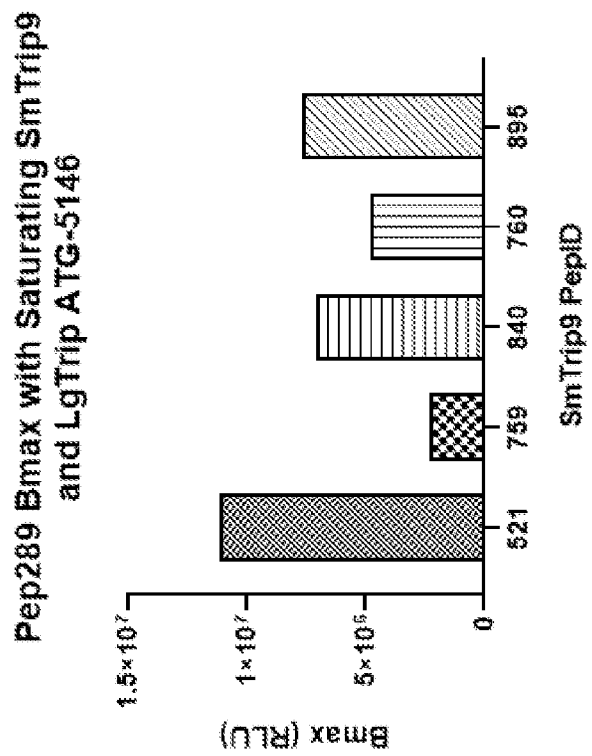

Pep289 Affinity with Saturating SmTrip9 and LgTrip ATG-3546
Kd (μM)
50
40
30
20
10
0
521
759
840
760
895
SmTrip9 PeptD

FIG. 180


Pep289 Bmax with Saturating SmTrip9 and LgTrip ATG-5146
Bmax (RLU)
1.5×10⁷
1×10⁷
5×10⁶
0
521
759
840
760
895
SmTrip9 PepID

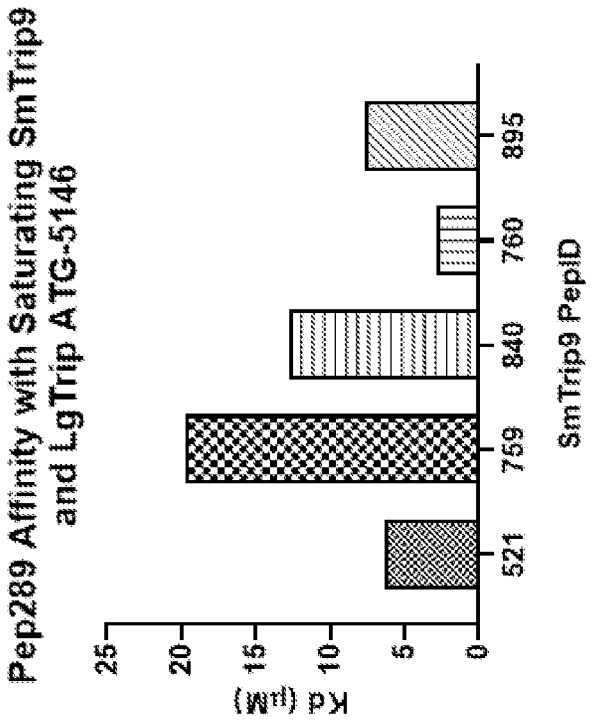

Pep289 Affinity with Saturating SmTrip9 and LgTrip ATG-5146
Kd (μM)
25
20
15
10
5
0
521
759
840
760
895
SmTrip9 PepID

FIG. 181

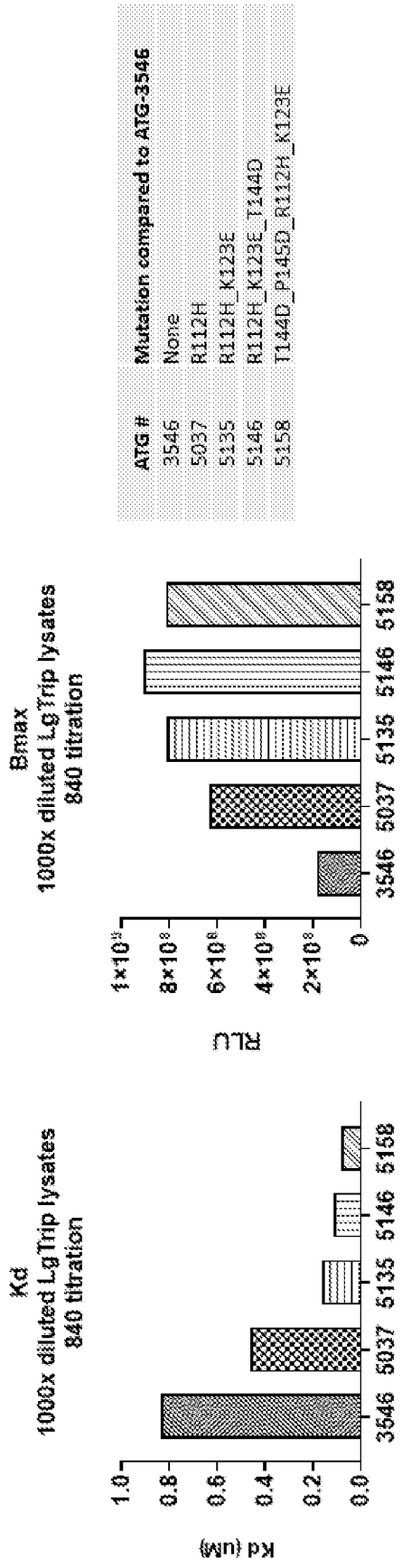
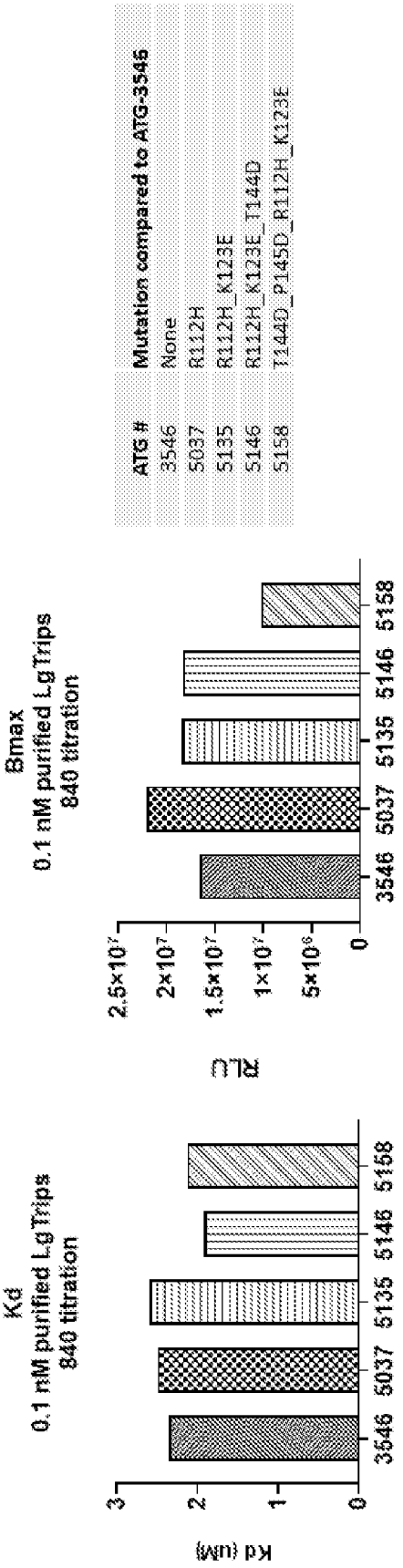

Kd
1000x diluted LgTrip lysates
840 titration
Kd (uM)
1.0
0.8
0.6
0.4
0.2
0.0
3546
5037
5135
5146
5158
Bmax
1000x diluted LgTrip lysates
840 titration
RLU
1x10^9
8x10^8
6x10^8
4x10^8
2x10^8
0
3546
5037
5135
5146
5158
ATG # Mutation compared to ATG-3546
3546 None
5037 R112H
5135 R112H_K123E
5146 R112H_K123E_T144D
5158 T144D_P145D_R112H_K123E

FIG. 182



ATG #	Mutation compared to ATG-3546
3546	None
5037	R112H
5135	R112H_K123E
5146	R112H_K123E_T144D
5158	T144D_P145D_R112H_K123E

FIG. 183

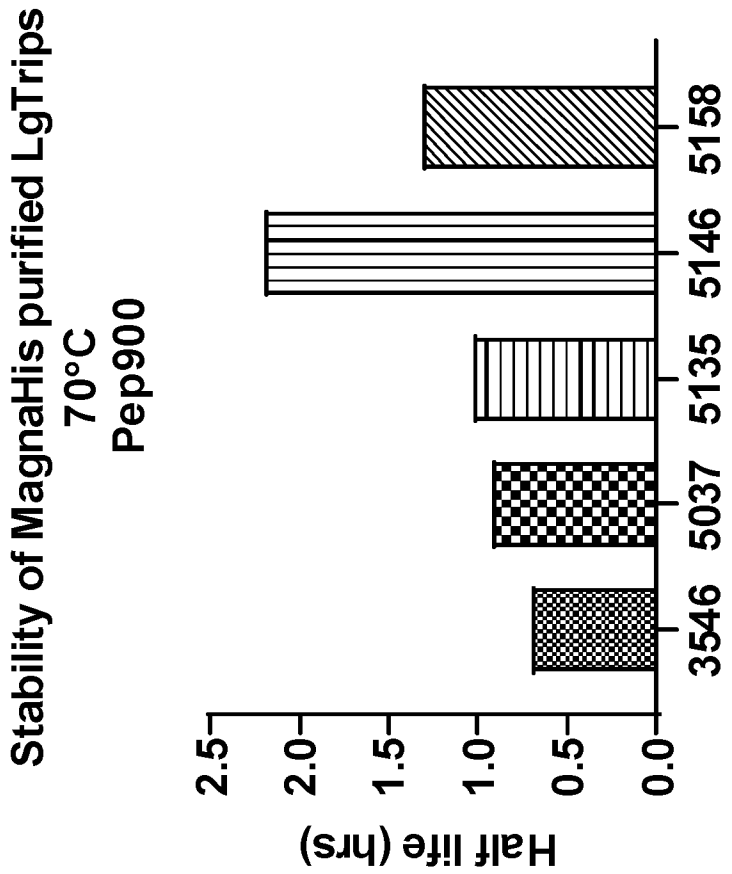

Stability of MagnaHis purified LgTrips
70°C
Pep900
Half life (hrs)
2.5
2.0
1.5
1.0
0.5
0.0
3546
5037
5135
5146
5158

FIG. 184

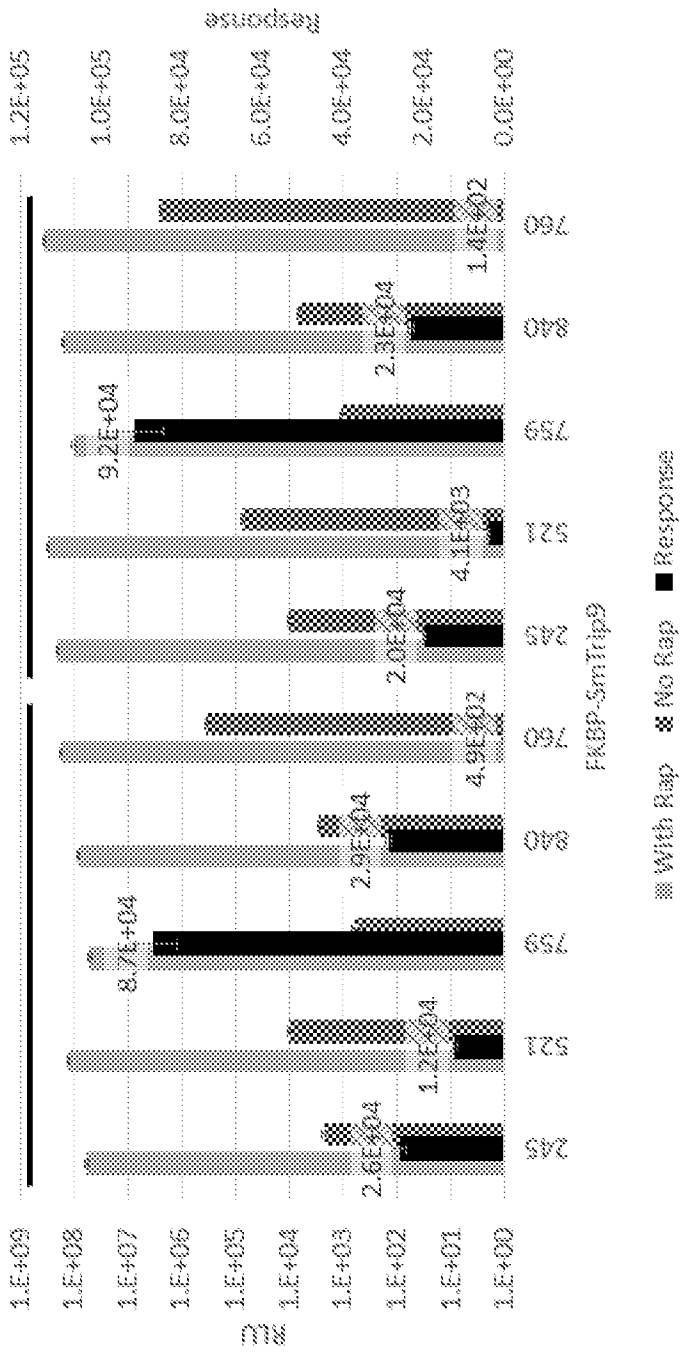

ATG-3546
ATG-5146
RLU
Response
1.E+09
1.E+08
1.E+07
1.E+06
1.E+05
1.E+04
1.E+03
1.E+02
1.E+01
1.E+00
1.2E+05
1.0E+05
8.0E+04
6.0E+04
4.0E+04
2.0E+04
0.0E+00
245
521
759
840
760
245
521
759
840
760
FABP-SmTrip9
With Rap
No Rap
Response

FIG. 185

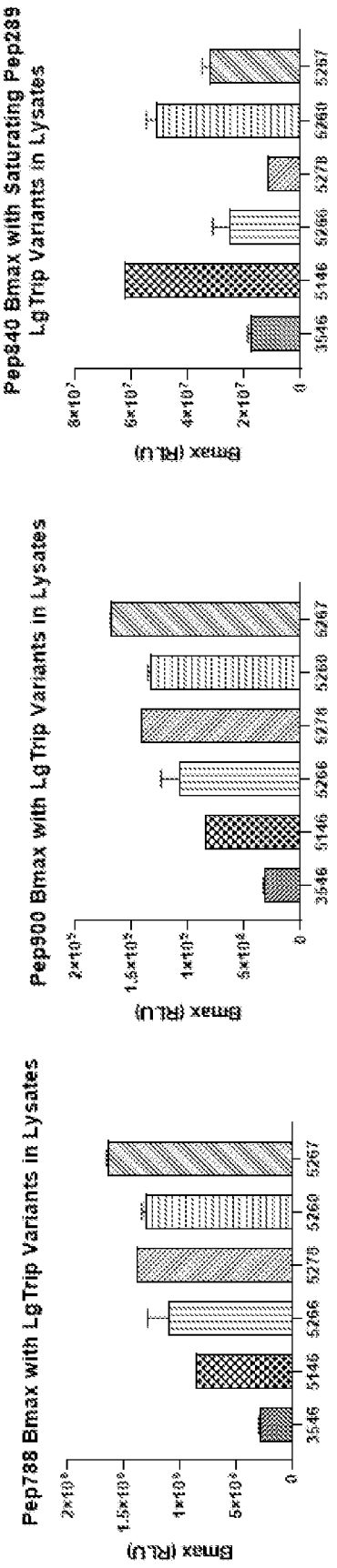

Pep840 Bmax with Saturating Pep289 Lg Trip Variants in Lysates
Bmax (RLU)
8x10^7
6x10^7
4x10^7
2x10^7
0
3546
5146
5266
5278
5260
5267
Pep900 Bmax with Lg Trip Variants in Lysates
Bmax (RLU)
2x10^5
1.5x10^5
1x10^5
5x10^4
0
3546
5146
5266
5278
5268
5267
Pep788 Bmax with Lg Trip Variants in Lysates
Bmax (RLU)
2x10^5
1.5x10^5
1x10^5
5x10^4
0
3546
5146
5266
5278
5260
5267

FIG. 186

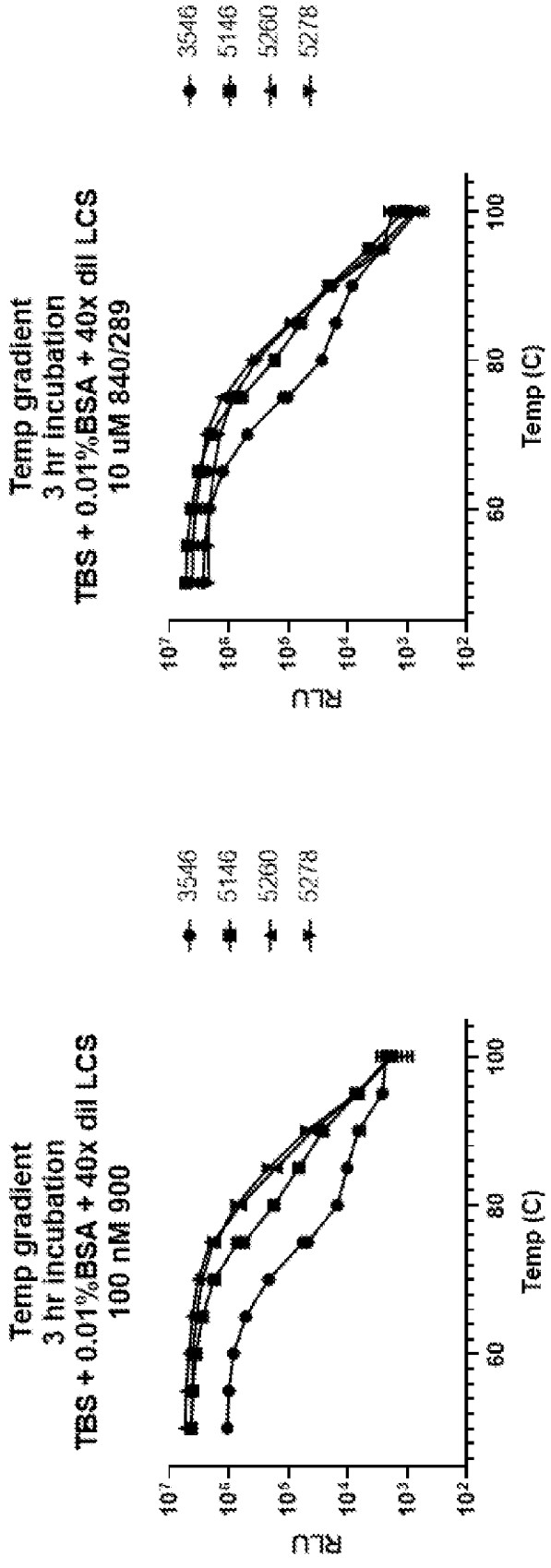

Temp gradient
3 hr incubation
TBS + 0.01%BSA + 40x dil LCS
100 nM 900
Temp gradient
3 hr incubation
TBS + 0.01%BSA + 40x dil LCS
10 uM 840/289
RLU
Temp (C)
3546
5146
5260
5278

FIG. 187

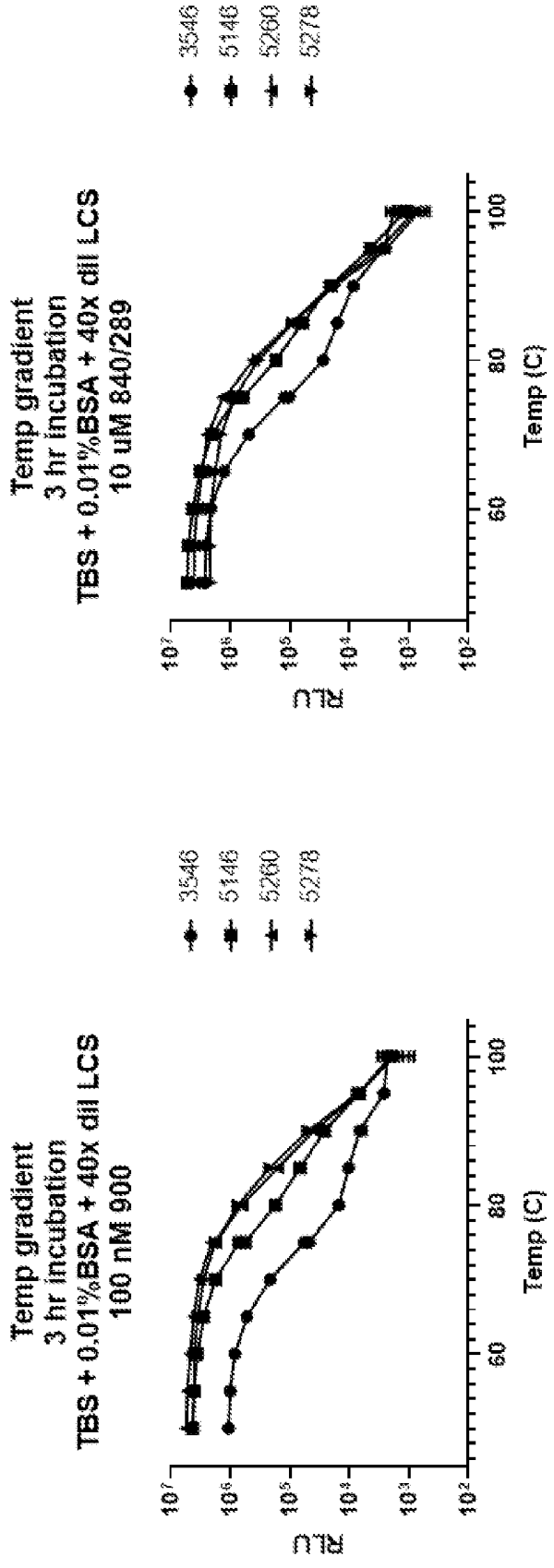
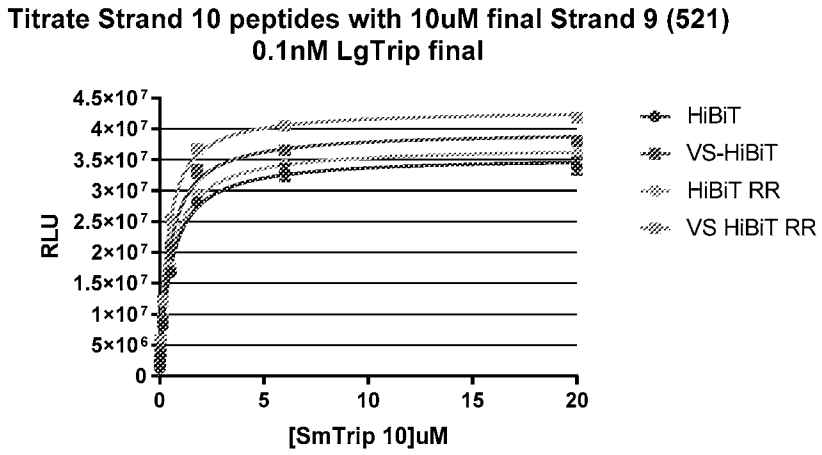
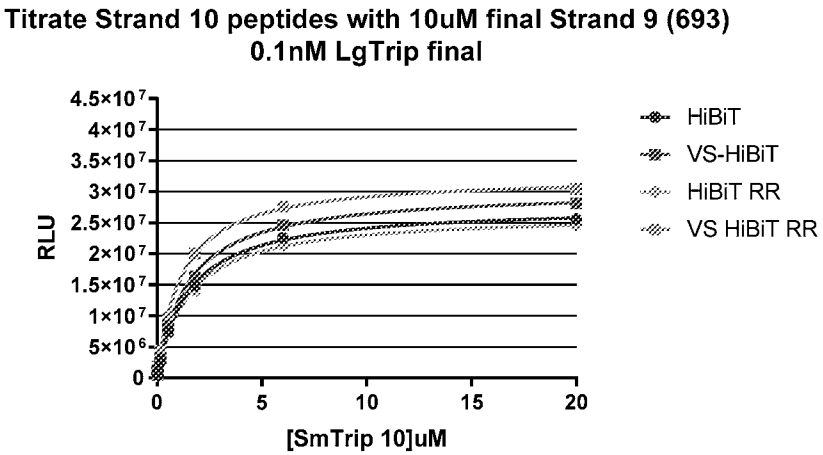

Temp gradient
3 hr incubation
TBS + 0.01%BSA + 40x dil LCS
100 nM 900
RLU
Temp (C)
3546
5146
5260
5278
Temp gradient
3 hr incubation
TBS + 0.01%BSA + 40x dil LCS
10 uM 840/289
RLU
Temp (C)
3546
5146
5260
5278

FIG. 188



	HiBiT	VS-HiBiT	HiBiT RR	VS HiBiT RR
Bmax	35518888	39630782	37182707	43098486
Kd	0.5453	0.4670	0.5318	0.3713



	HiBiT	VS-HiBiT	HiBiT RR	VS HiBiT RR
Bmax	27688586	30279910	26458890	32495666
Kd	1.484	1.463	1.497	1.149

FIG. 189

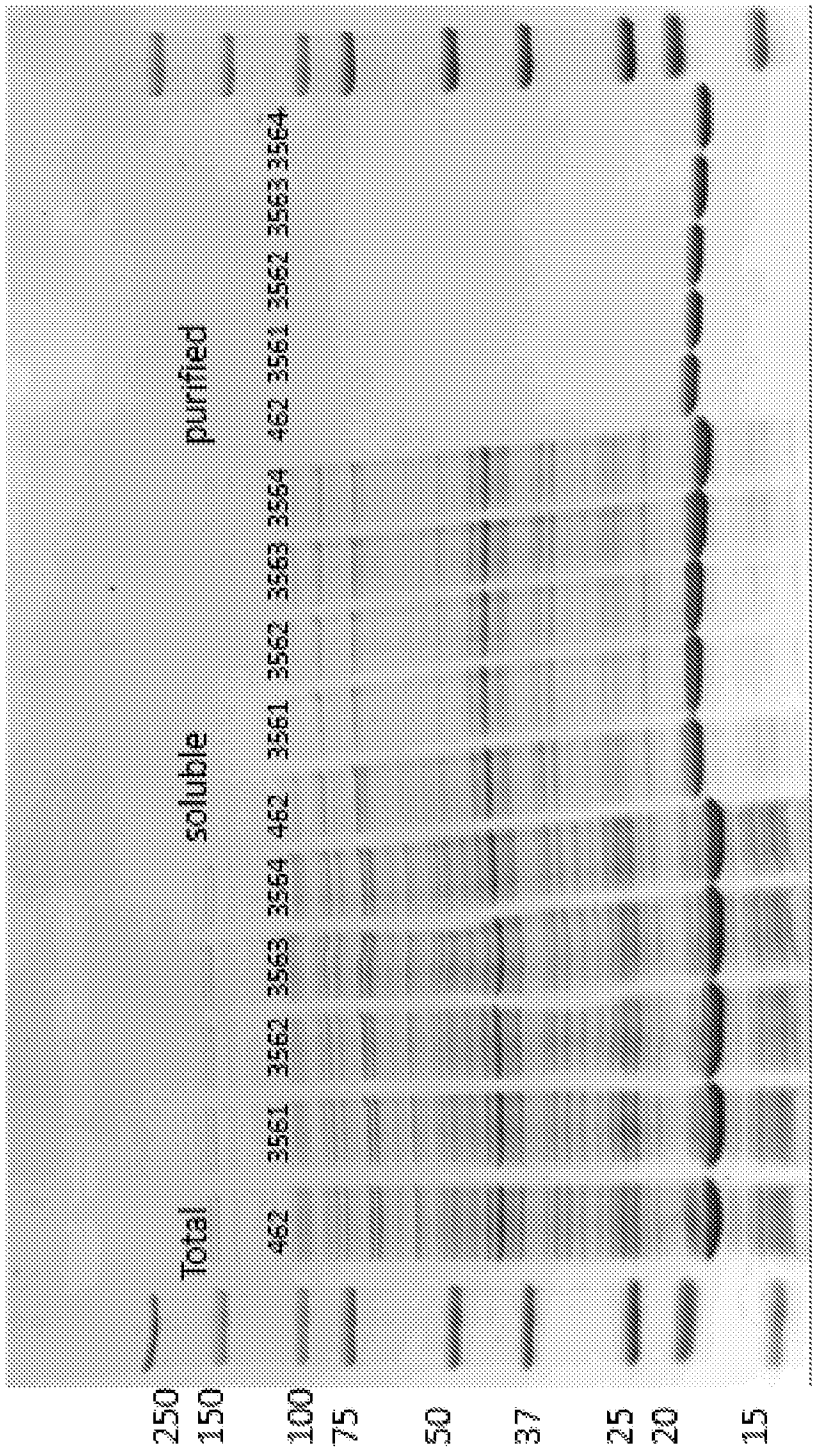

Total
462
3561
3562
3563
3564
soluble
462
3561
3562
3563
3564
purified
462
3561
3562
3563
3564
250
150
100
75
50
37
25
20
15

FIG. 190

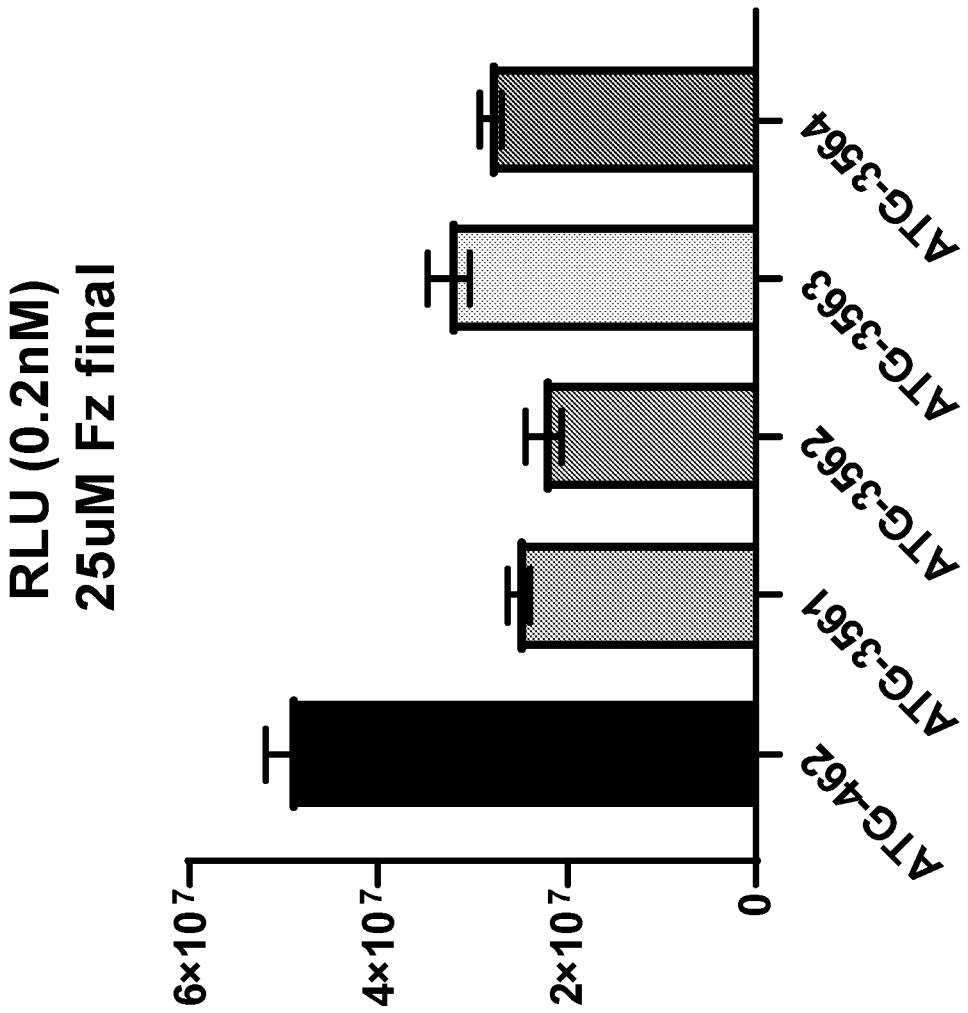
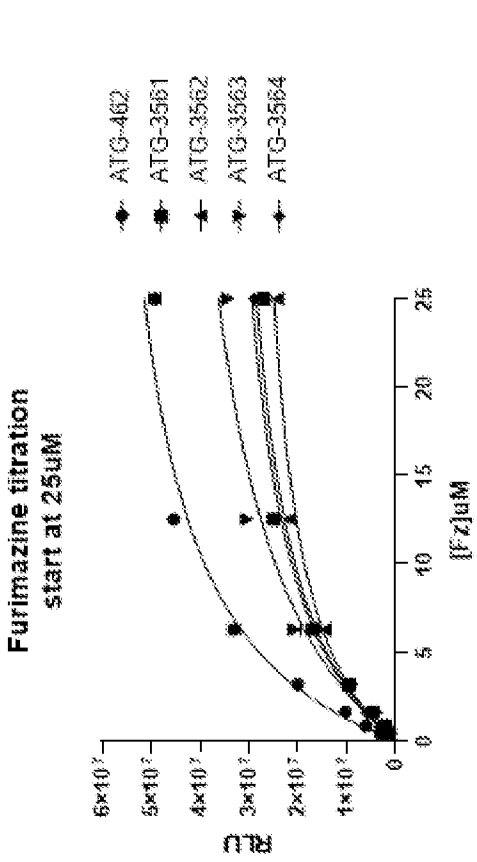

RLU (0.2nM)
25uM Fz final
6×10^7
4×10^7
2×10^7
0
ATG-462
ATG-3561
ATG-3562
ATG-3563
ATG-3564

FIG. 191



	ATG-462	ATG-3561	ATG-3562	ATG-3563	ATG-3564
Vmax	6.50E+07	3.80E+07	3.15E+07	5.28E+07	3.92E+07
Km	6.6	8.4	7.1	11.4	8.4

FIG. 192A-C

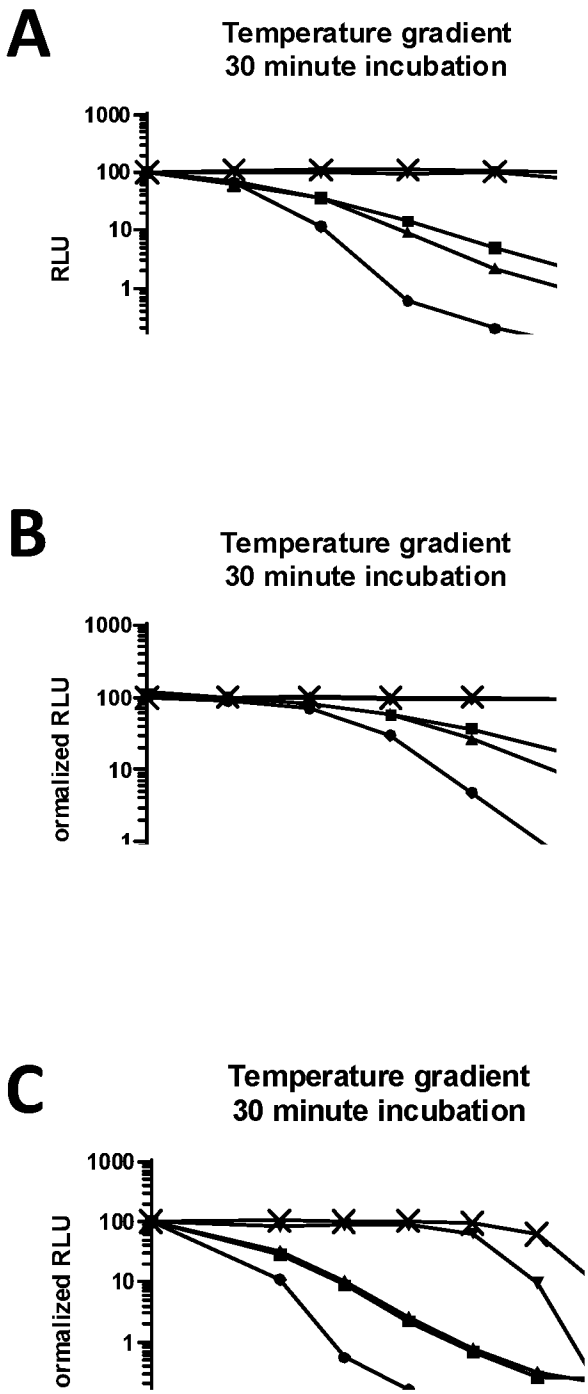

A
Temperature gradient
30 minute incubation
RLU
1000
100
10
1
B
Temperature gradient
30 minute incubation
ormalized RLU
1000
100
10
1
C
Temperature gradient
30 minute incubation
ormalized RLU
1000
100
10
1

FIG. 193

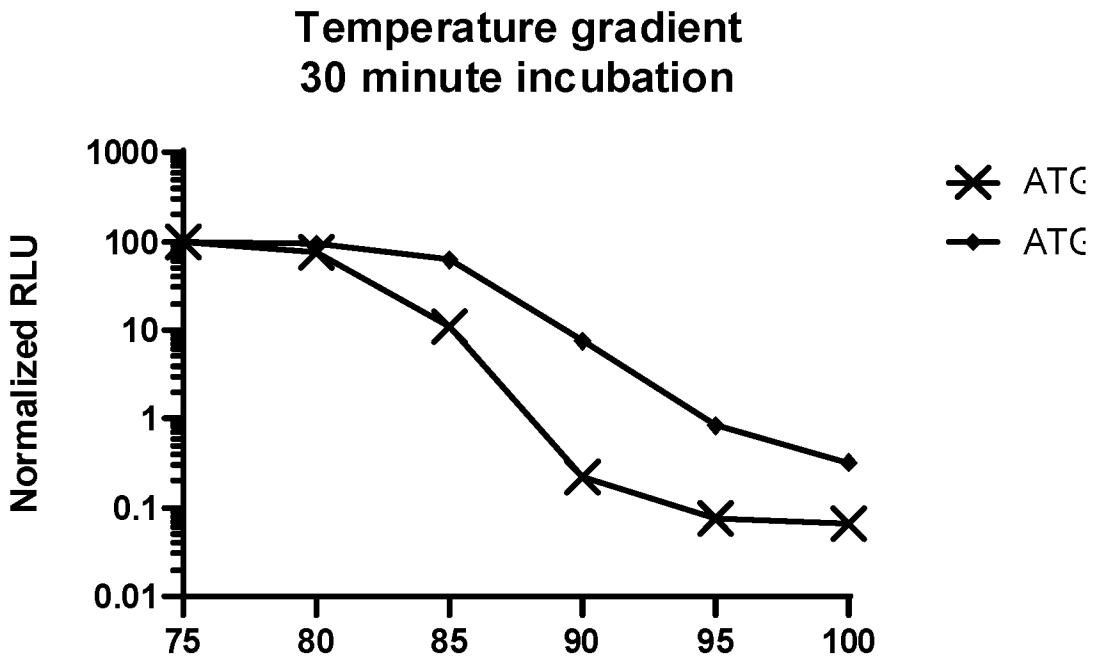
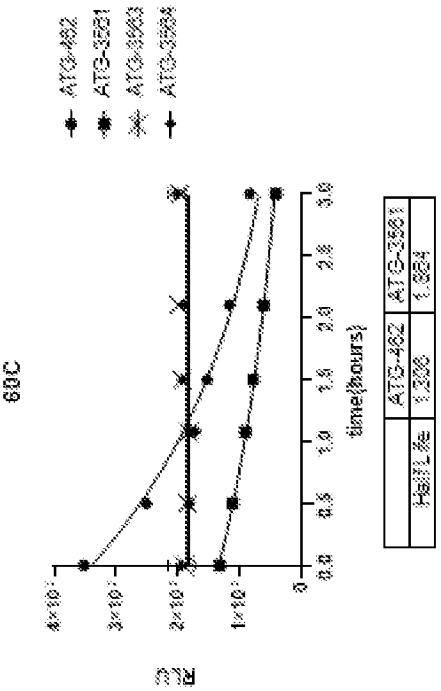

Temperature gradient
30 minute incubation
Normalized RLU
1000
100
10
1
0.1
0.01
75
80
85
90
95
100
ATG
ATG

FIG. 194



	ATG-462	ATG-3581
Half-life	1.206	1.854

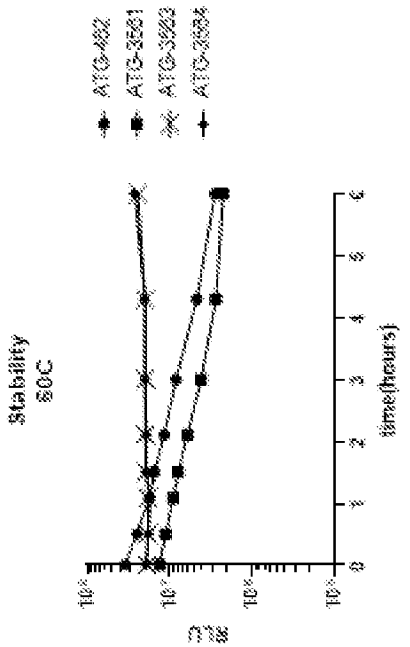


FIG. 195

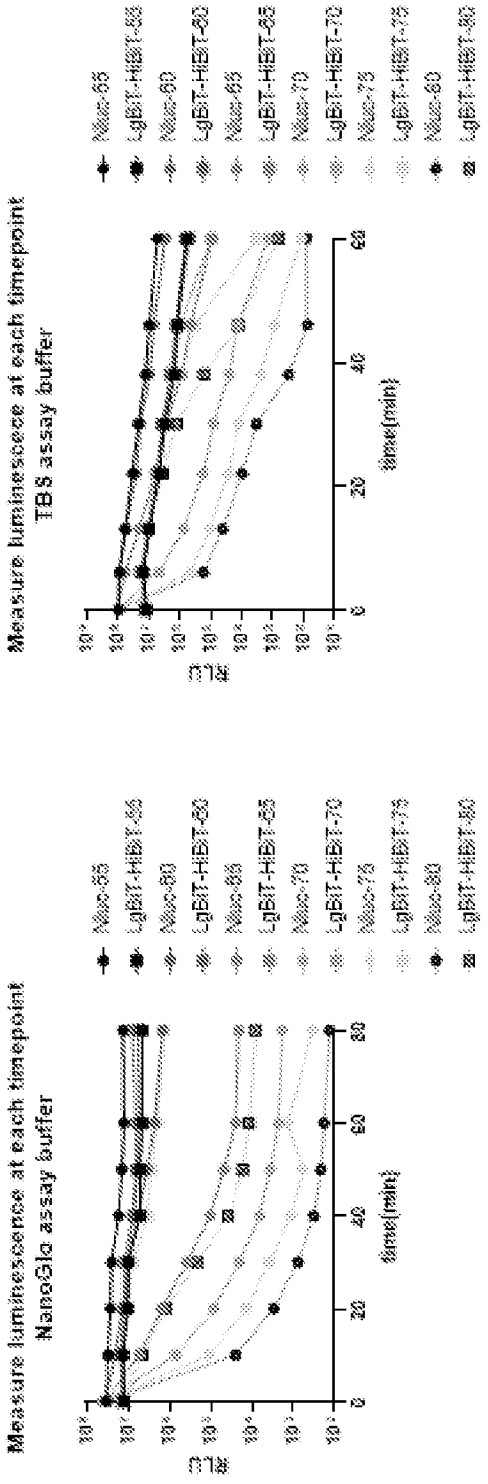
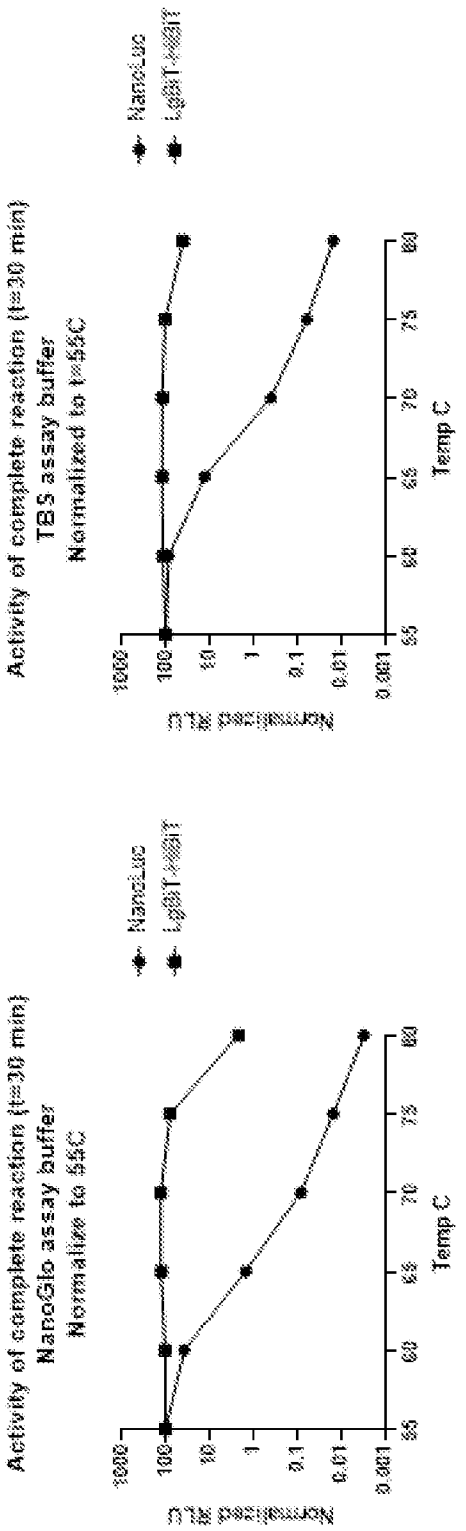

Measure luminescence at each timepoint
NanoGlo assay buffer
RLU
Time(min)
Nluc-65
LgBiT-HiBiT-65
Nluc-60
LgBiT-HiBiT-60
Nluc-65
LgBiT-HiBiT-65
Nluc-70
LgBiT-HiBiT-70
Nluc-75
LgBiT-HiBiT-75
Nluc-80
LgBiT-HiBiT-80
Measure luminescence at each timepoint
TBS assay buffer
RLU
Time(min)
Nluc-65
LgBiT-HiBiT-65
Nluc-60
LgBiT-HiBiT-60
Nluc-65
LgBiT-HiBiT-65
Nluc-70
LgBiT-HiBiT-70
Nluc-75
LgBiT-HiBiT-75
Nluc-80
LgBiT-HiBiT-80

FIG. 196


Activity of complete reaction (t=30 min)
NanoGlo assay buffer
Normalize to 55C
NanoLuc
LgBiT+SmBiT
Normalized RLU
1000
100
10
1
0.1
0.01
0.001
55
60
65
70
75
80
Temp C
Activity of complete reaction (t=30 min)
TBS assay buffer
Normalized to t=55C
NanoLuc
LgBiT+SmBiT
Normalized RLU
1000
100
10
1
0.1
0.01
0.001
55
60
65
70
75
80
Temp C

Sample	Vmax	Km
ATG-462	6.3E+07	5.4
ATG-3564	3.1E+07	3.0
ATG-5788	3.2E+07	4.4
ATG-5789	1.9E+07	3.1
ATG-5790	2.2E+07	4.1
ATG-5791	2.5E+07	3.3
ATG-5792	3.0E+07	5.5
ATG-5793	2.7E+07	4.8
ATG-5794	2.0E+07	5.1
ATG-5795	3.5E+07	2.6
ATG-5796	2.3E+07	2.1
ATG-5797	2.5E+07	3.9
ATG-5798	2.7E+07	4.3
ATG-5799	2.2E+07	3.4
ATG-5800	2.8E+07	4.6
ATG-5801	2.0E+07	2.1
ATG-5802	5.2E+07	4.9
ATG-5803	2.2E+07	1.9
ATG-5804	3.8E+07	2.3

FIG. 197

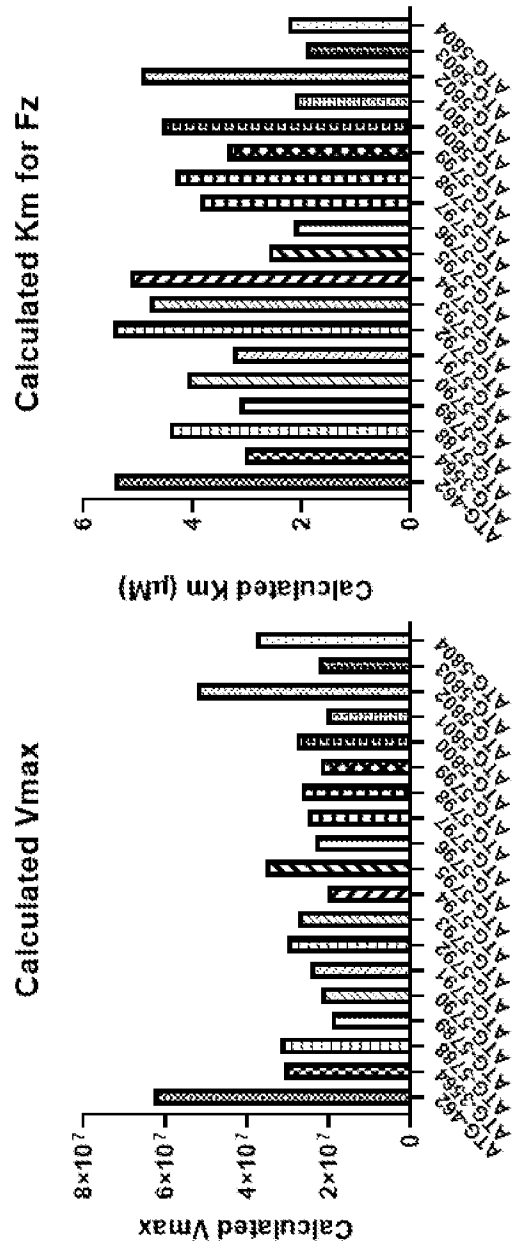
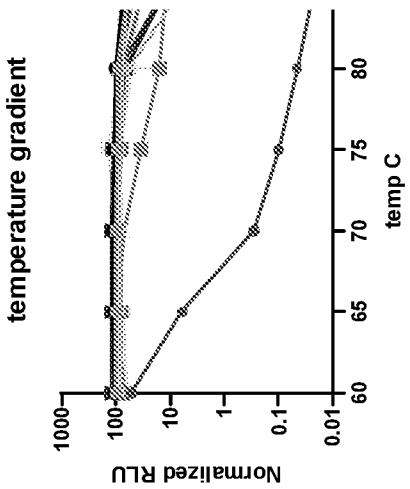


FIG. 198


temperature gradient
Normalized RLU
1000
100
10
1
0.1
0.01
60
65
70
75
80
temp C

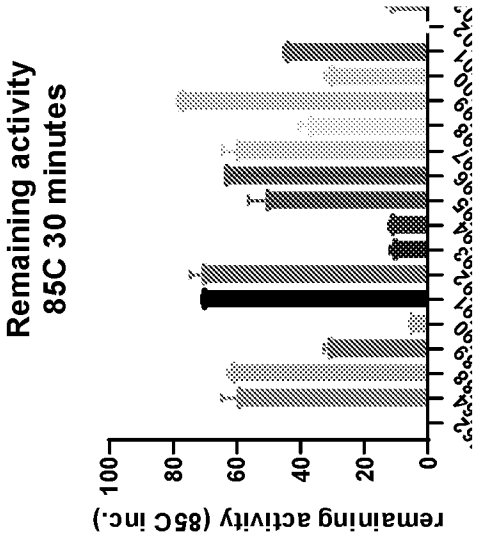

Remaining activity
85C 30 minutes
remaining activity (85C inc.)
100
80
60
40
20
0

FIG. 199

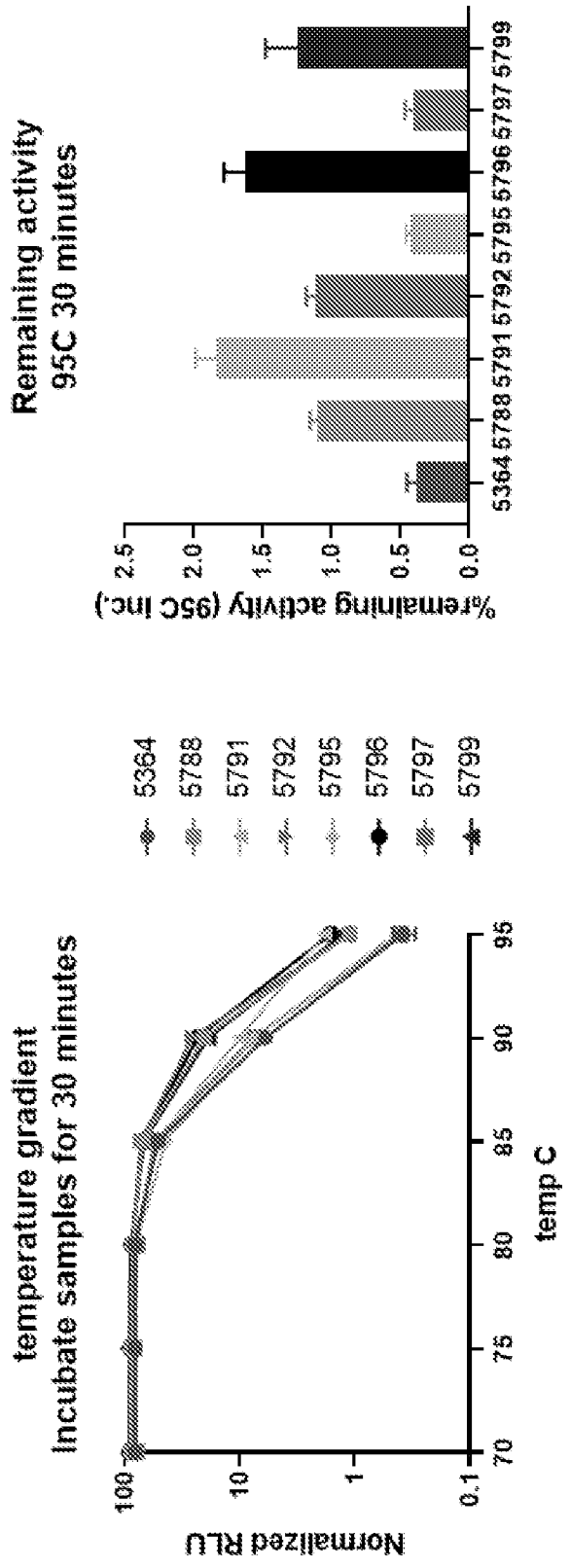

temperature gradient
Incubate samples for 30 minutes
Normalized RLU
100
10
1
0.1
70
75
80
85
90
95
temp C
5364
5788
5791
5792
5795
5796
5797
5799
Remaining activity
95C 30 minutes
%remaining activity (95C inc.)
2.5
2.0
1.5
1.0
0.5
0.0
5364 5788 5791 5792 5795 5796 5797 5799

FIG. 200

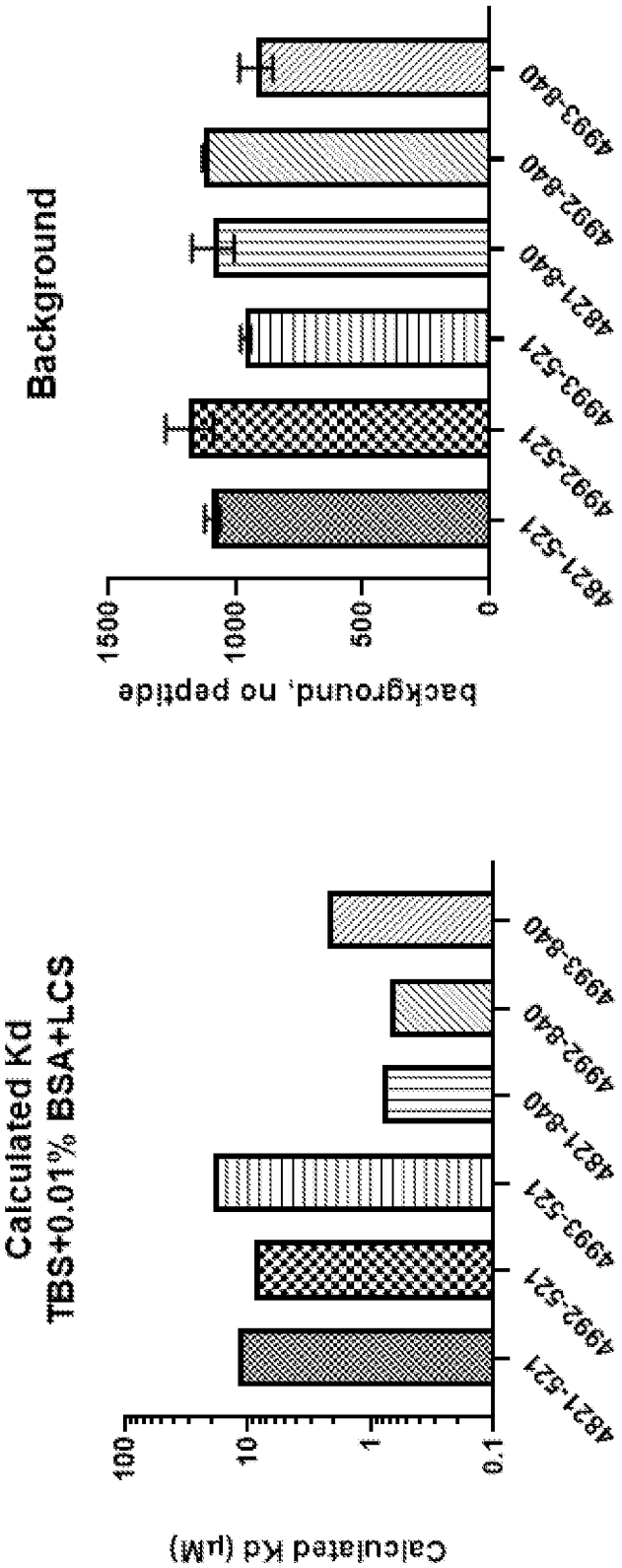

Background
background, no peptide
1500
1000
500
0
4821-521
4992-521
4993-521
4821-840
4992-840
4993-840
Calculated Kd
TBS+0.01% BSA+LCS
Calculated Kd (μM)
100
10
1
0.1
4821-521
4992-521
4993-521
4821-840
4992-840
4993-840

FIG. 201

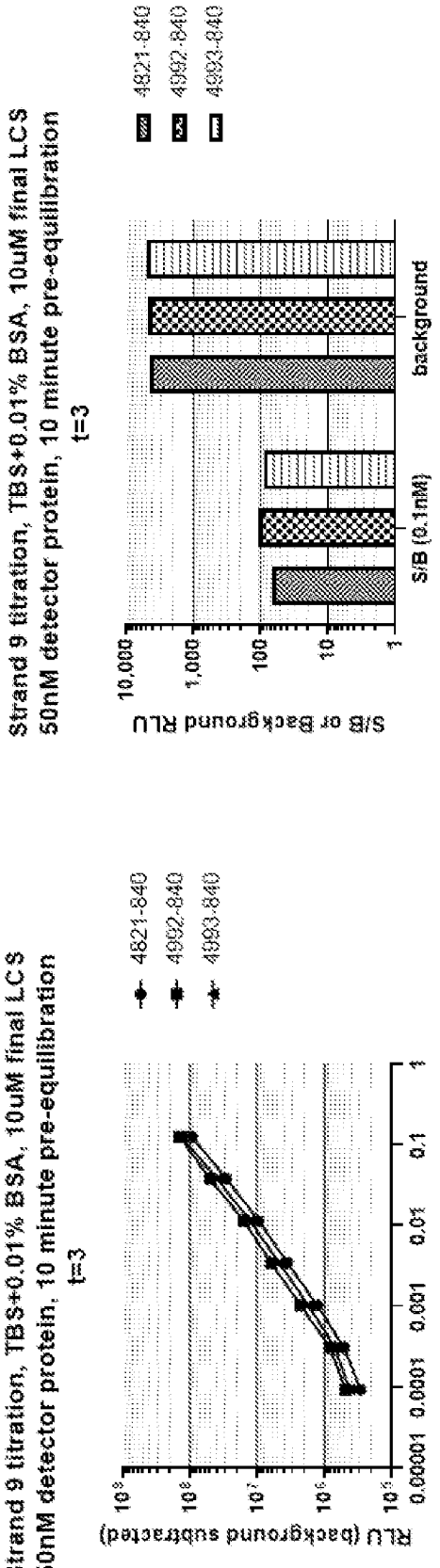

Strand 9 titration, TBS+0.01% BSA, 10uM final LCS
50nM detector protein, 10 minute pre-equilibration
t=3
4821-840
4992-840
4993-840
RLU (background subtracted)
[840]uM
Strand 9 titration, TBS+0.01% BSA, 10uM final LCS
50nM detector protein, 10 minute pre-equilibration
t=3
4821-840
4992-840
4993-840
S/B or Background RLU
S/B (0.1nM)
background

FIG. 202

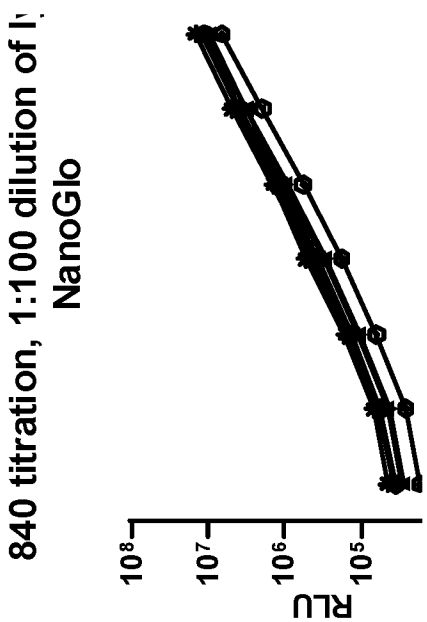


FIG. 203

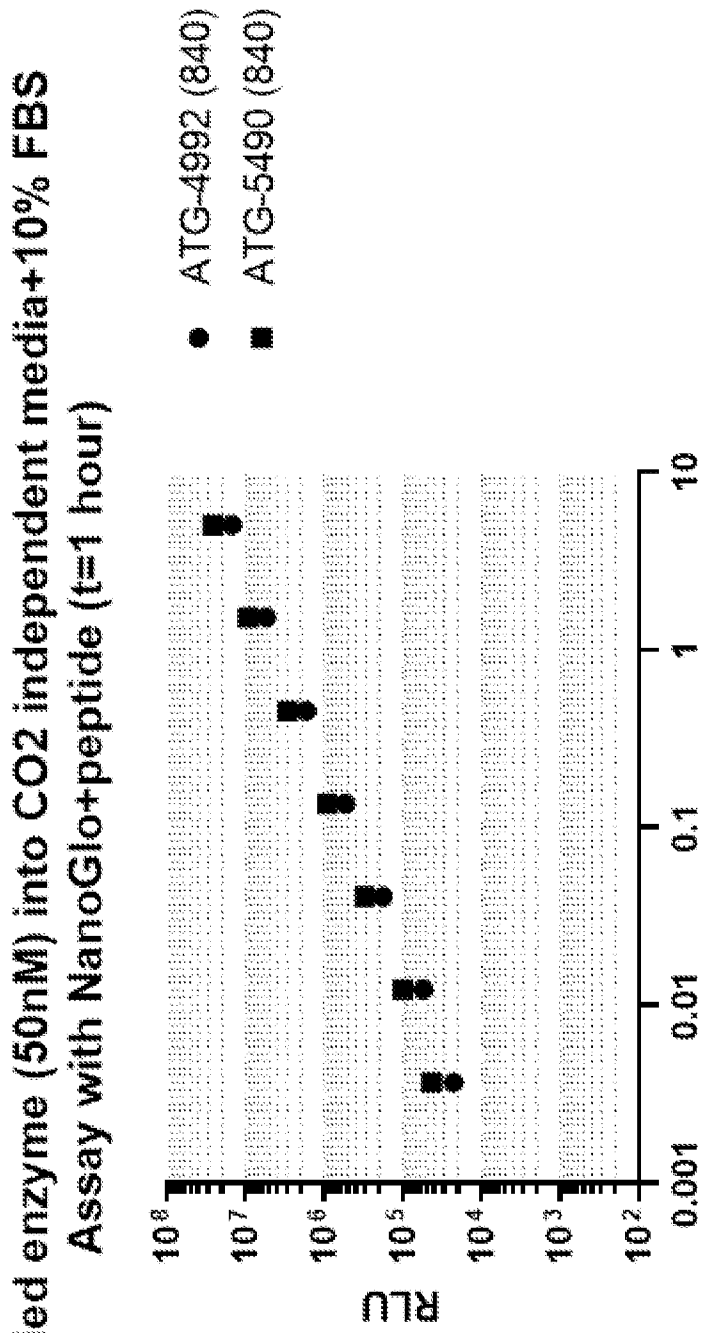
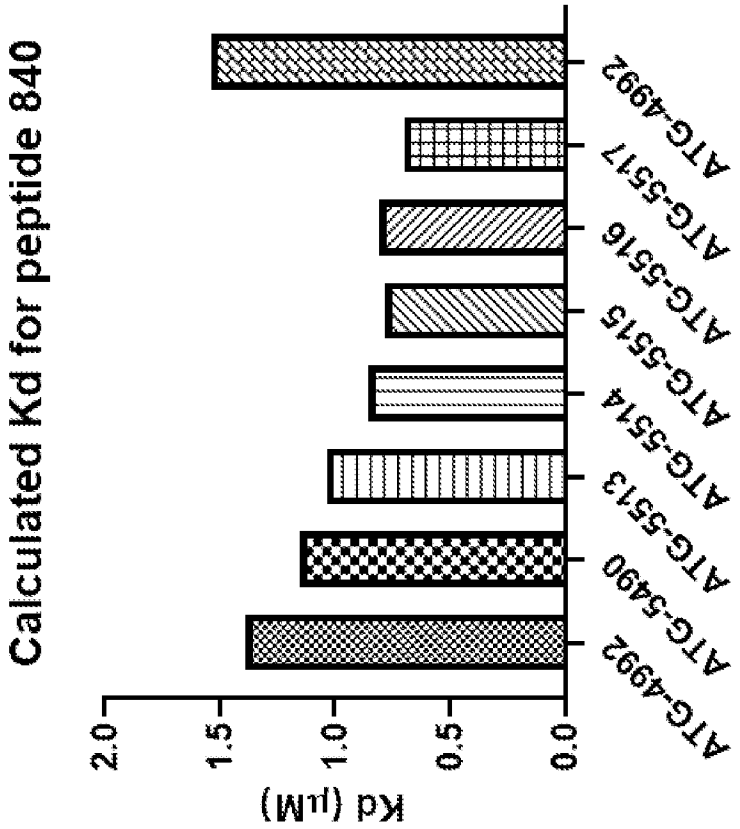

Dilute purified enzyme (50nM) into CO2 independent media+10% FBS
Assay with NanoGlo+peptide (t=1 hour)
ATG-4992 (840)
ATG-5490 (840)
RLU
[peptide]nM
0.001
0.01
0.1
1
10

FIG. 204



Sample	Linker length
ATG-4992	8
ATG-5490	11
ATG-5513	8
ATG-5514	8
ATG-5515	11
ATG-5516	11
ATG-5517	11

FIG. 205

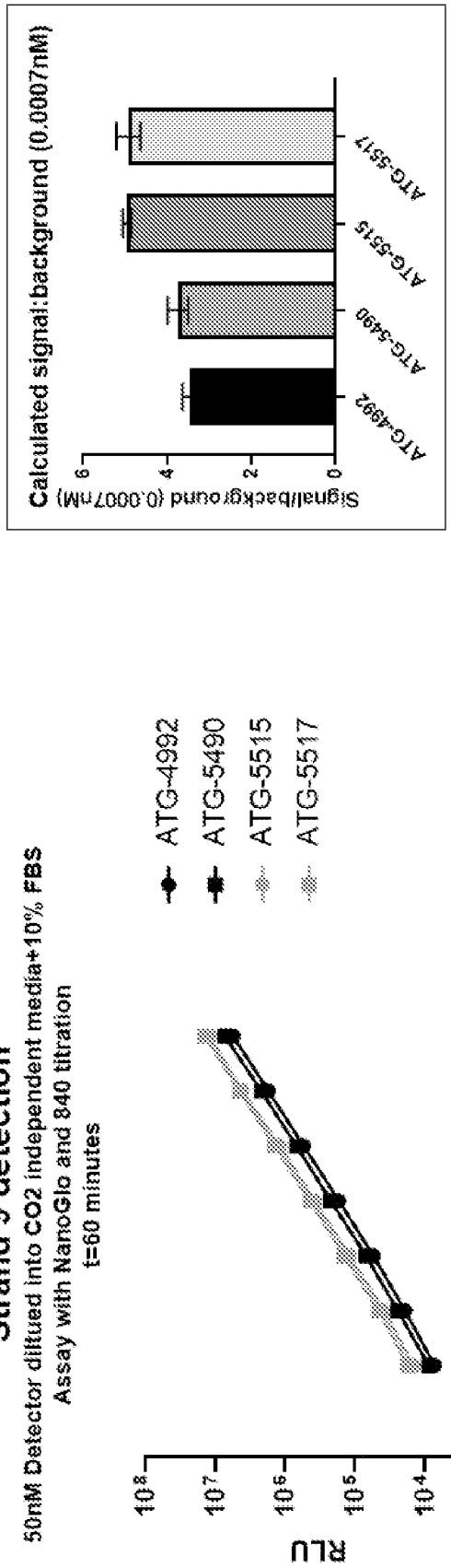
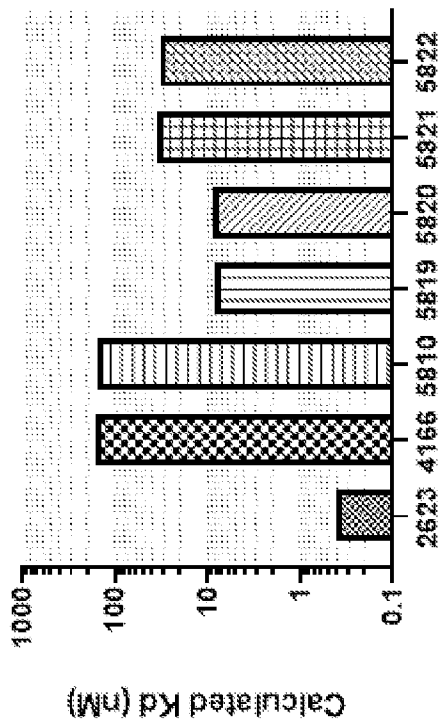

Strand 9 detection
50nM Detector diluted into CO2 independent media+10% FBS
Assay with NanoGlo and 840 titration
t=60 minutes
ATG-4992
ATG-5490
ATG-5515
ATG-5517
RLU
[pep840]µM
Calculated signal:background (0.0007nM)
Signal/background (0.0007nM)

FIG. 206


Calculated Kd
Calculated Kd (nM)
1000
100
10
1
0.1
2623
4166
5810
5819
5820
5821
5822

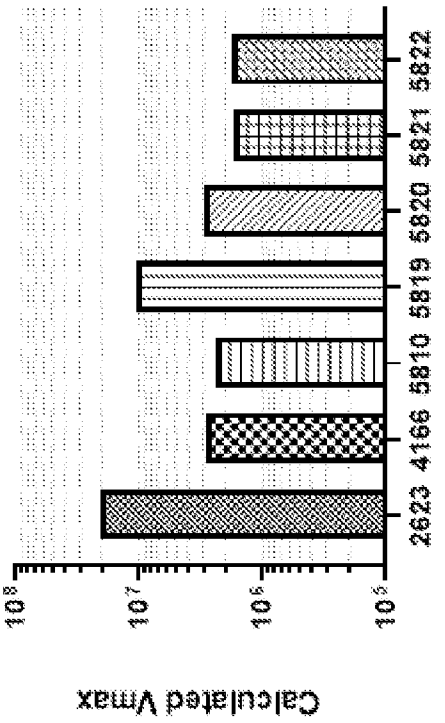

Calculated Bmax
Calculated Vmax
10^8
10^7
10^6
10^5
2623
4166
5810
5819
5820
5821
5822

FIG. 207

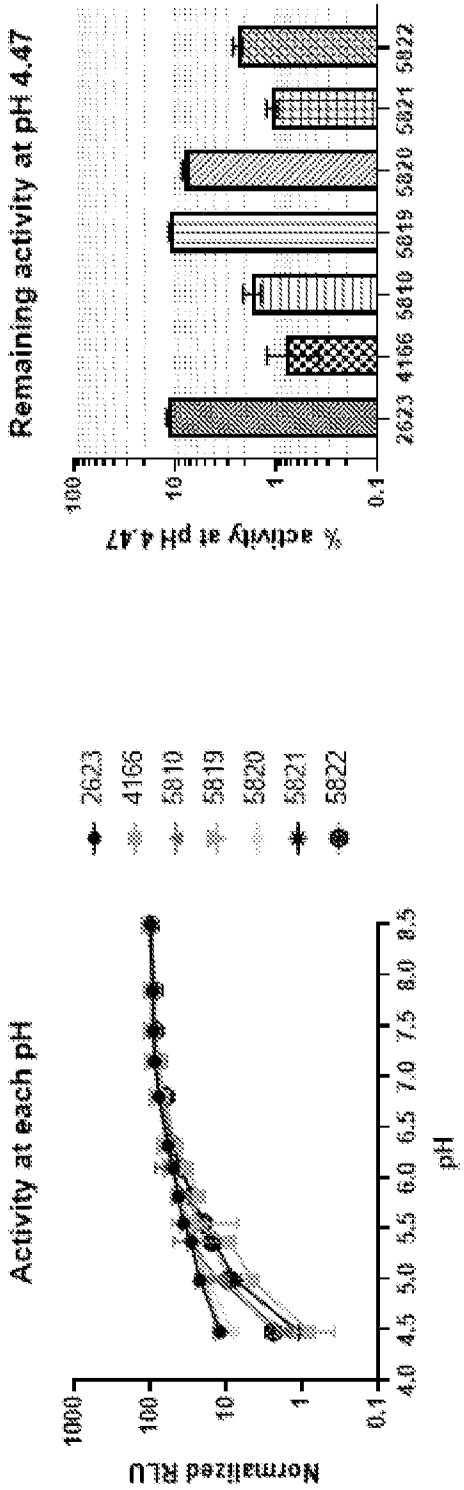

Remaining activity at pH 4.47
% activity at pH 4.47
100
10
1
0.1
2623
4166
5810
5819
5820
5821
5822
Activity at each pH
Normalized RLU
1000
100
10
1
0.1
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
pH
2623
4166
5810
5819
5820
5821
5822

FIG. 208

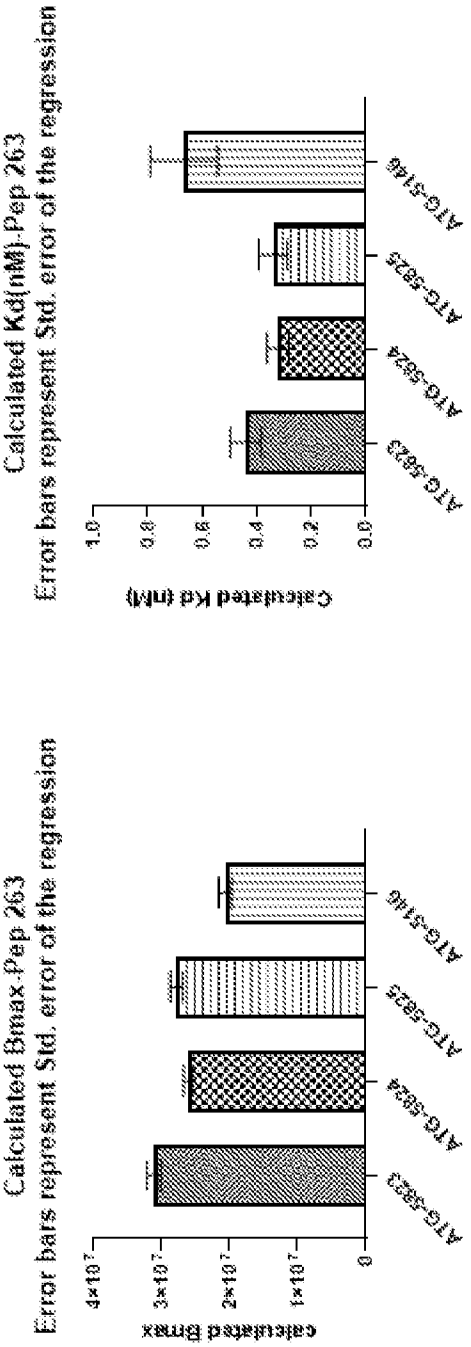

Calculated Kd(nM)-Pep 263
Error bars represent Std. error of the regression
Calculated Kd (nM)
1.0
0.8
0.6
0.4
0.2
0.0
ATG-5823
ATG-5824
ATG-5825
ATG-5146
Calculated Bmax-Pep 263
Error bars represent Std. error of the regression
calculated Bmax
4×10^7
3×10^7
2×10^7
1×10^7
0
ATG-5823
ATG-5824
ATG-5825
ATG-5146

FIG. 209

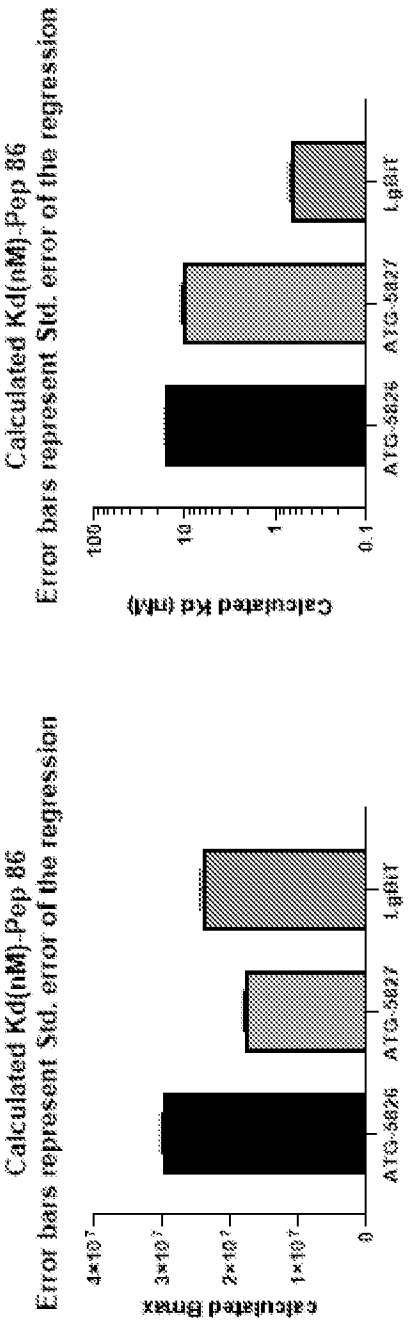
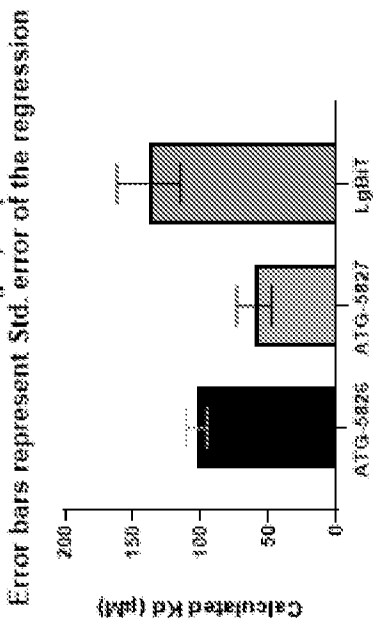

Calculated Kd(nM)-Pep 86
Error bars represent Std. error of the regression
Calculated Kd (nM)
100
10
1
0.1
ATG-5826
ATG-5827
IgBiT
Calculated Kd(nM)-Pep 86
Error bars represent Std. error of the regression
calculated Bmax
4×10⁷
3×10⁷
2×10⁷
1×10⁷
0
ATG-5826
ATG-5827
IgBiT

FIG. 210


Calculated Kd(μM).Pep 114
Error bars represent Std. error of the regression
Calculated Kd (μM)
200
150
100
50
0
ATG-5826
ATG-5827
LgBiT

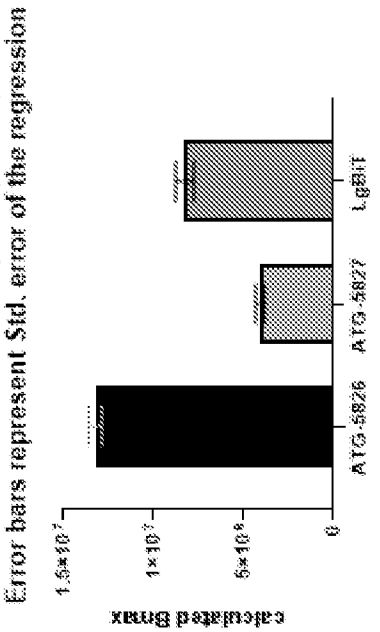

Calculated Bmax.Pep 114 (SmBiT)
Error bars represent Std. error of the regression
calculated Bmax
1.5×10⁷
1×10⁷
5×10⁶
0
ATG-5826
ATG-5827
LgBiT

FIG. 211

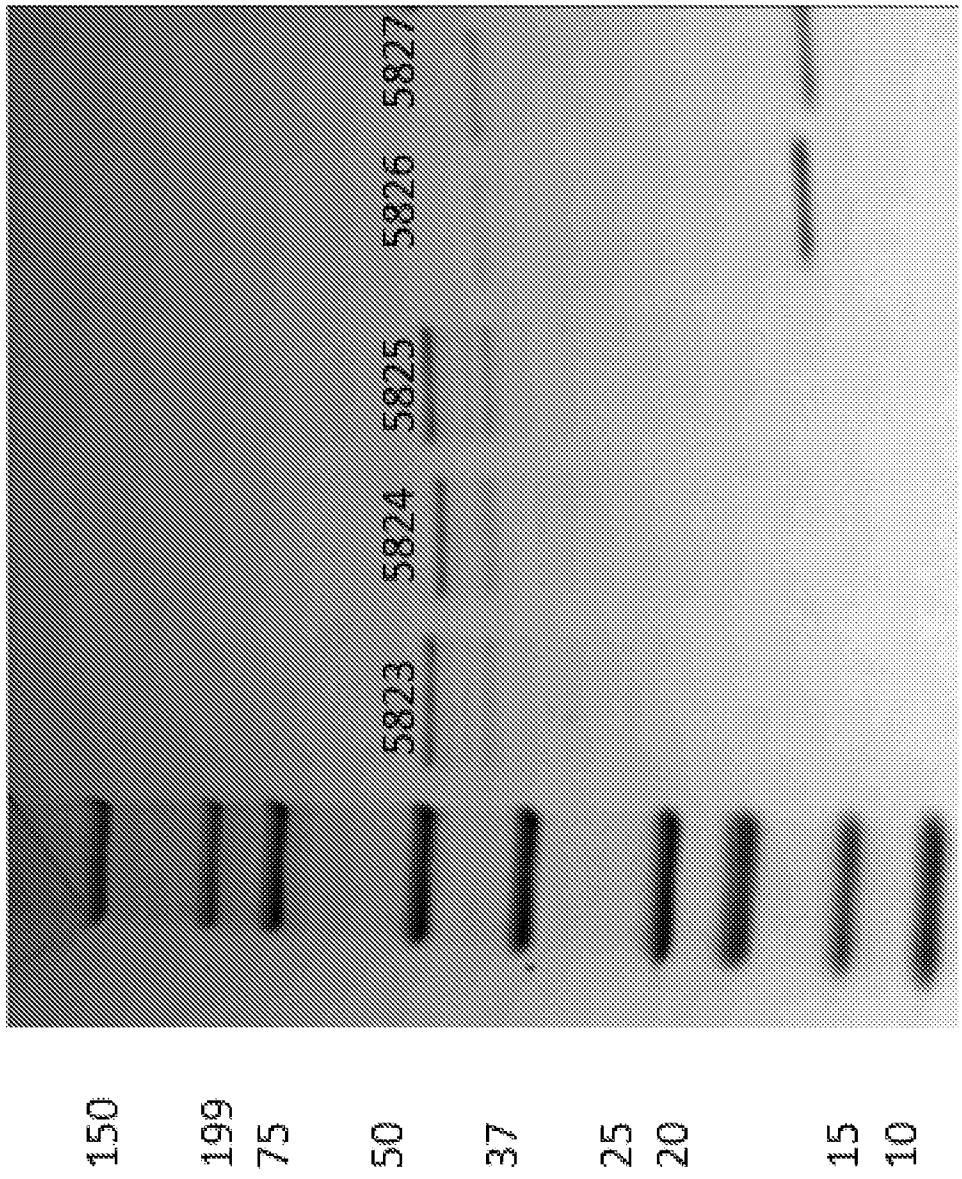


FIG. 212

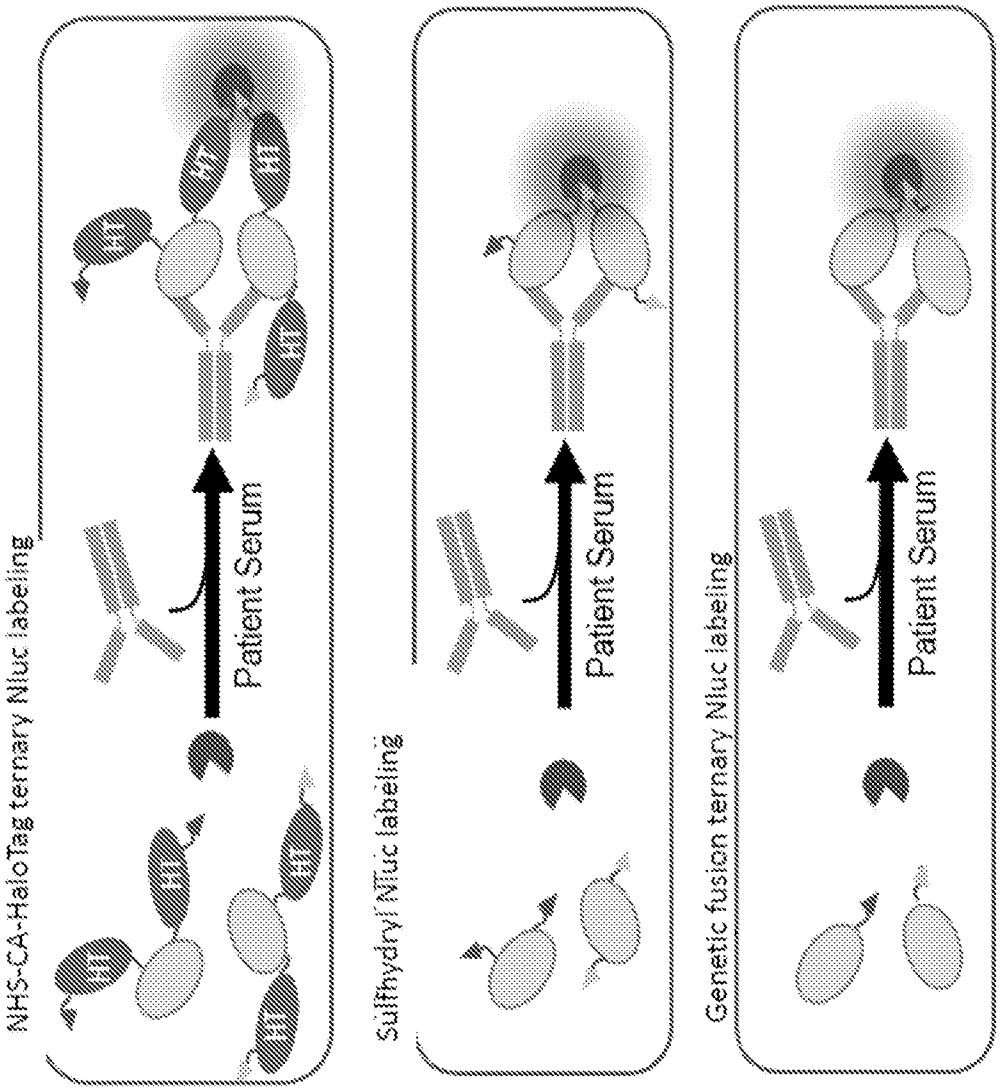
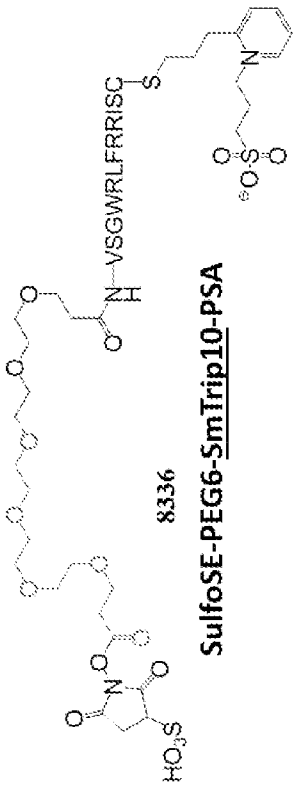

NHS-CA-HaloTag ternary Nluc labeling
Patient Serum
Sulfhydryl Nluc labeling
Patient Serum
Genetic fusion ternary Nluc labeling
Patient Serum

FIG. 213


VSGWRLFRRISC
8336
SulfoSE-PEG6-SmTrip10-PSA

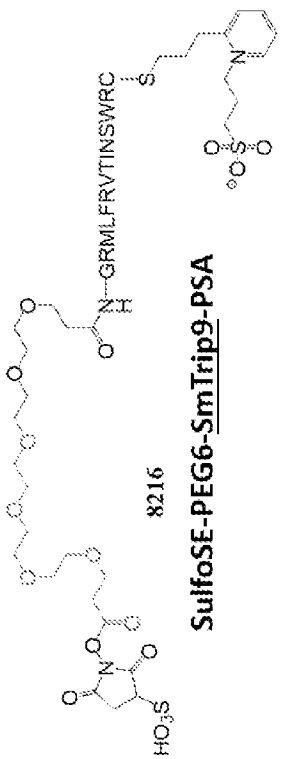

GRMLFRVTINSWRC
8216
SulfoSE-PEG6-SmTrip9-PSA

FIG. 214

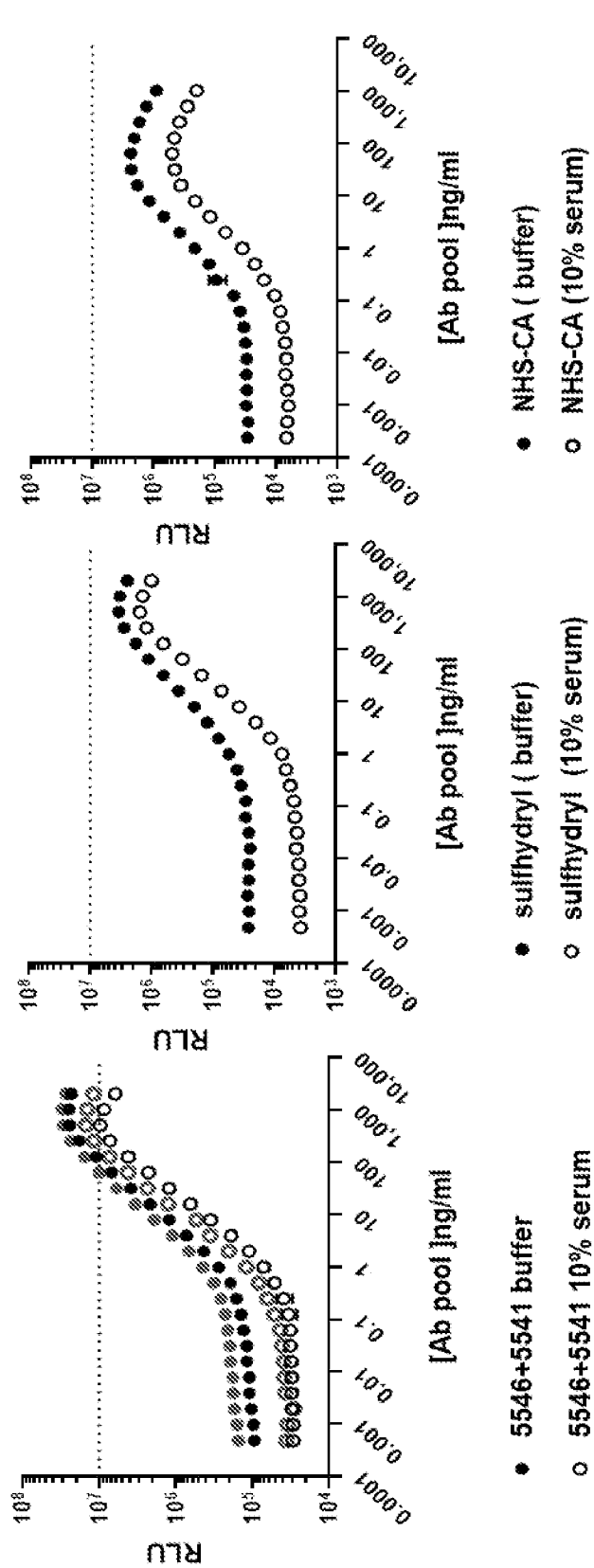
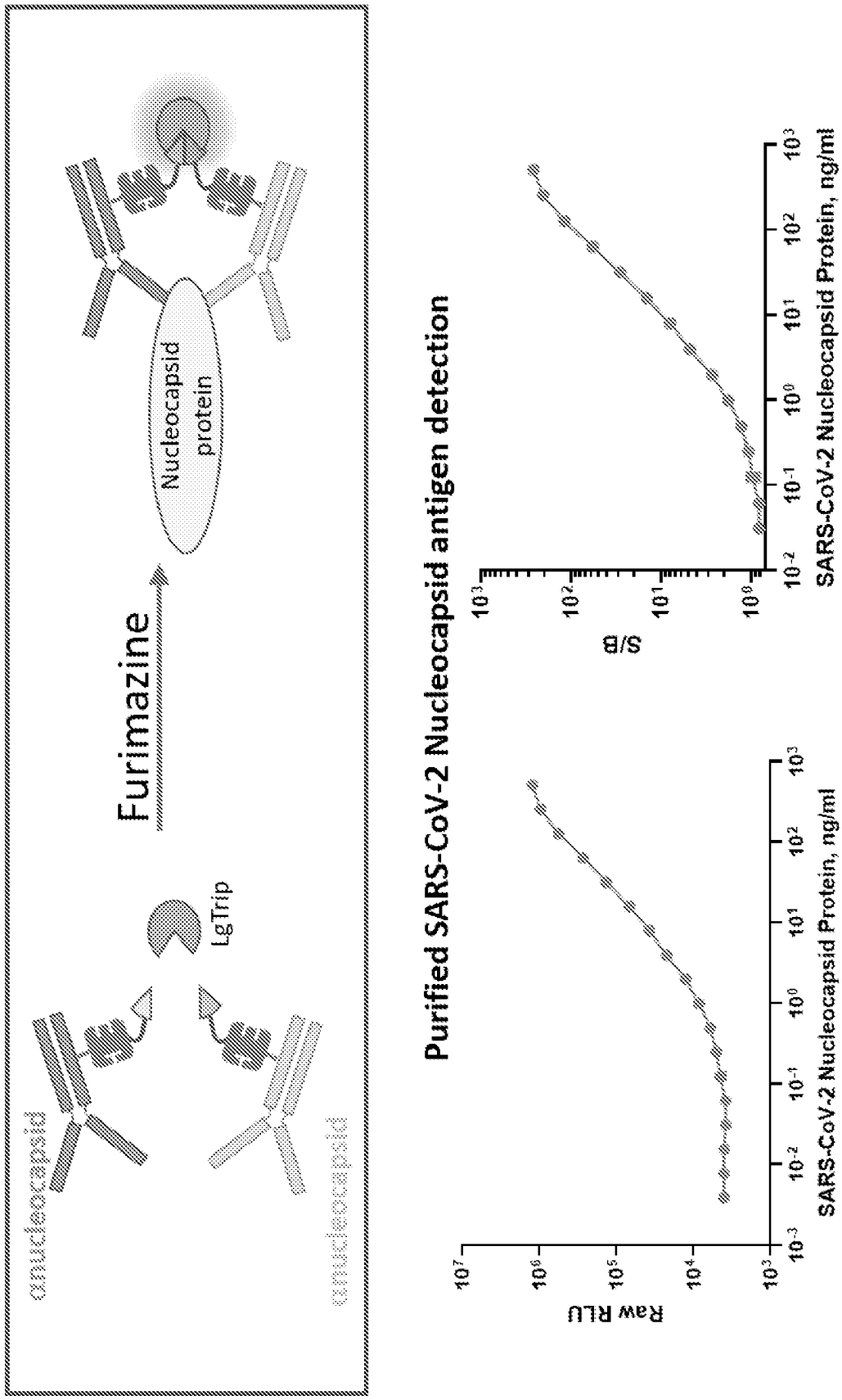

Genetic fusions
sulfhydryl-Trips labeling
NHS-CA-HaloTag-Trip labeling
RLU
[Ab pool]ng/ml
5546+5541 buffer
5546+5541 10% serum
5547+5541 (buffer)
5547+5541 10% serum
sulfhydryl (buffer)
sulfhydryl (10% serum)
NHS-CA (buffer)
NHS-CA (10% serum)

FIG. 215


anucleocapsid
LgTrip
Furimazine
Nucleocapsid protein
anucleocapsid
Purified SARS-CoV-2 Nucleocapsid antigen detection
Raw RLU
SARS-CoV-2 Nucleocapsid Protein, ng/ml
S/B
SARS-CoV-2 Nucleocapsid Protein, ng/ml

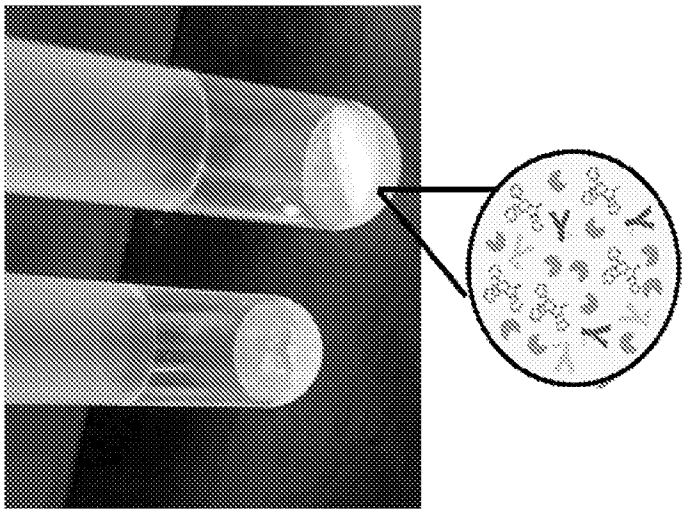


FIG. 216

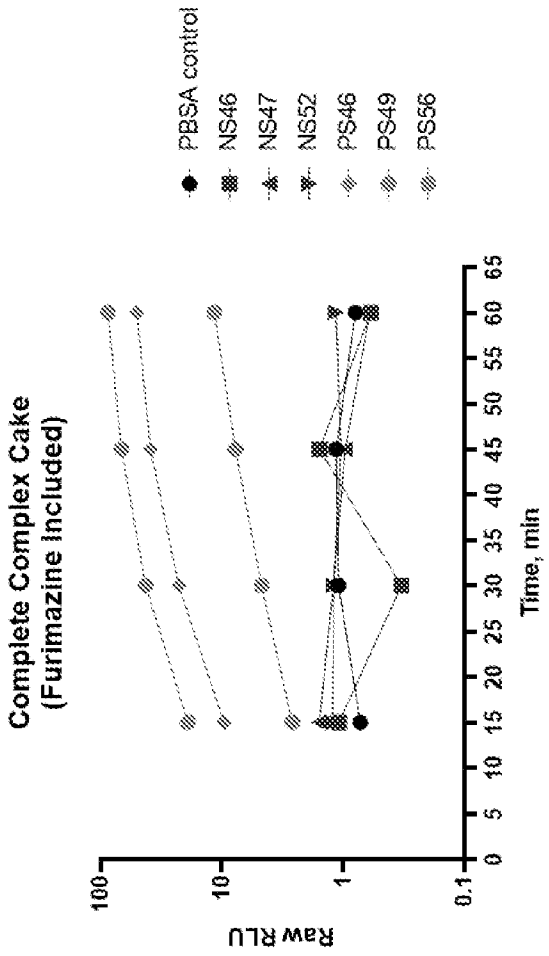

Complete Complex Cake
(Furimazine Included)
PBSA control
NS46
NS47
NS52
PS46
PS49
PS56
Raw RLU
100
10
1
0.1
0
5
10
15
20
25
30
35
40
45
50
55
60
65
Time, min

FIG. 217

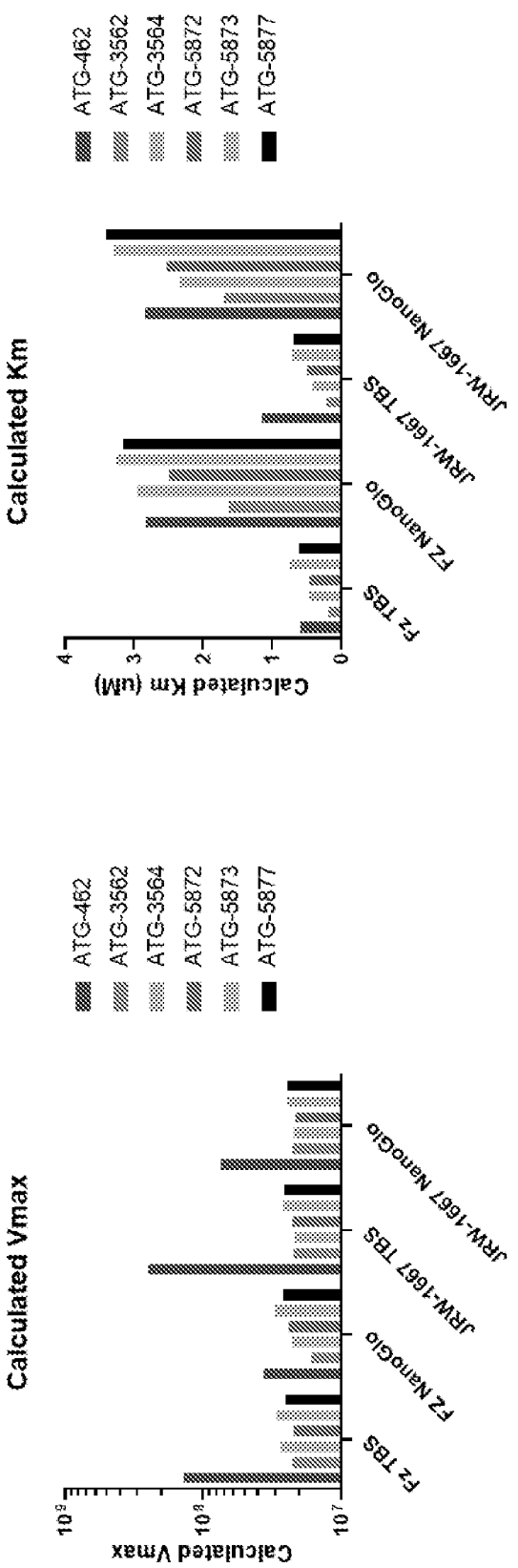

Calculated Km
Calculated Km (uM)
0
1
2
3
4
FZ TBS
FZ NanoGlo
JRW-1667 TBS
JRW-1667 NanoGlo
ATG-462
ATG-3562
ATG-3564
ATG-5872
ATG-5873
ATG-5877
Calculated Vmax
10^8
10^8
10^7
FZ TBS
FZ NanoGlo
JRW-1667 TBS
JRW-1667 NanoGlo
ATG-462
ATG-3562
ATG-3564
ATG-5872
ATG-5873
ATG-5877

FIG. 218

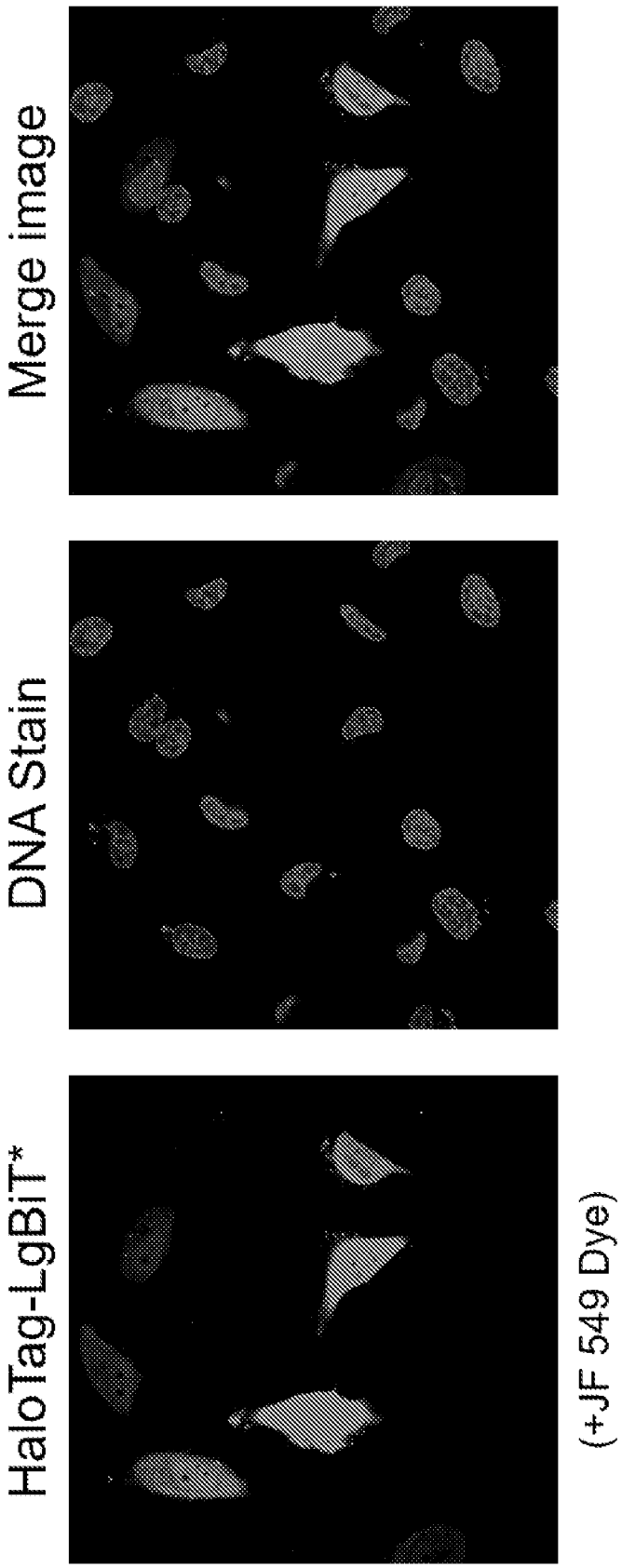


FIG. 219

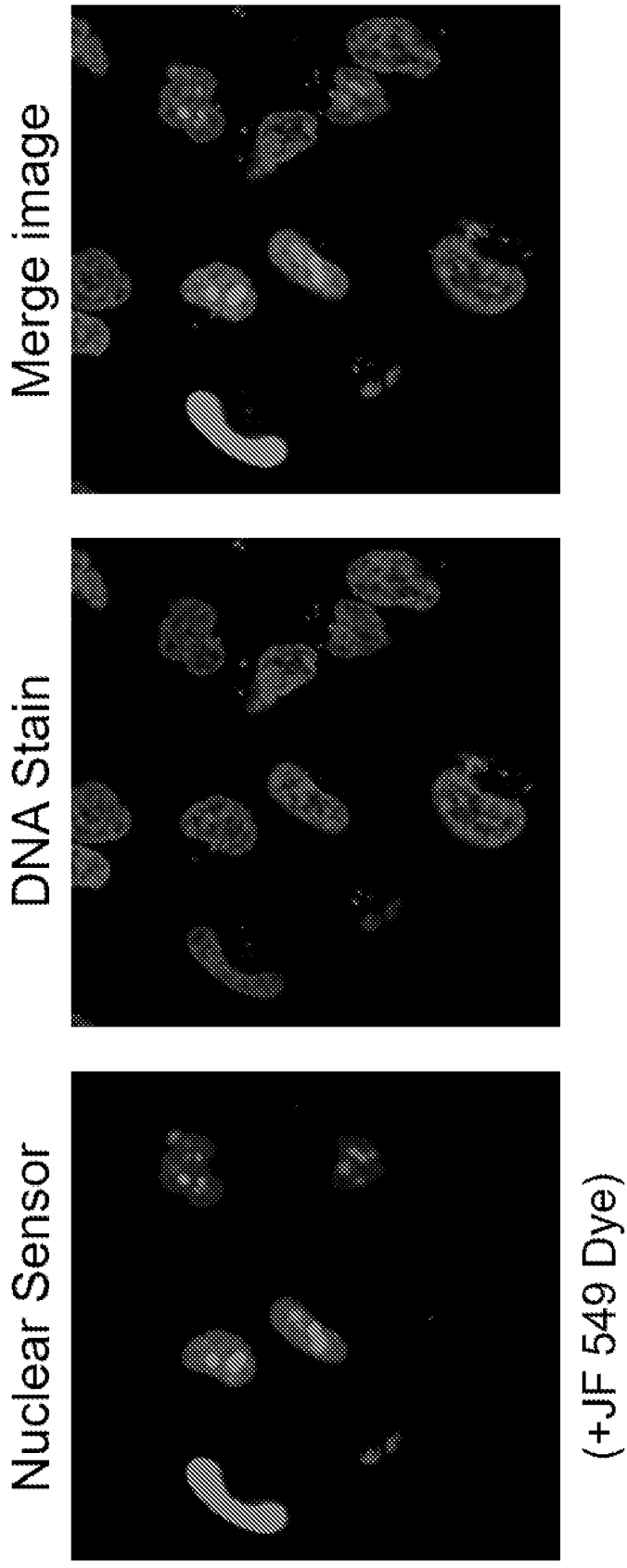


FIG. 220

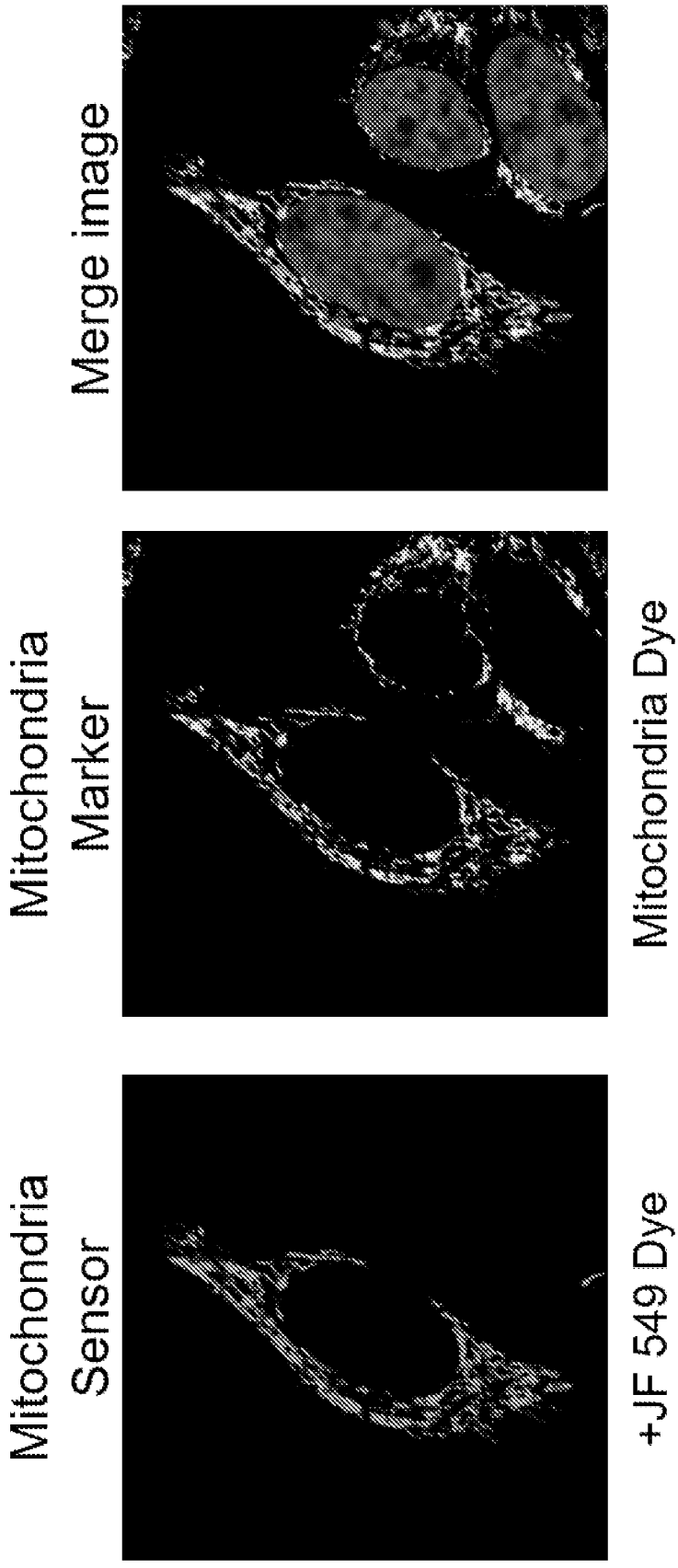

Merge image
Mitochondria Marker
Mitochondria Dye
Mitochondria Sensor
+JF 549 Dye

FIG. 221

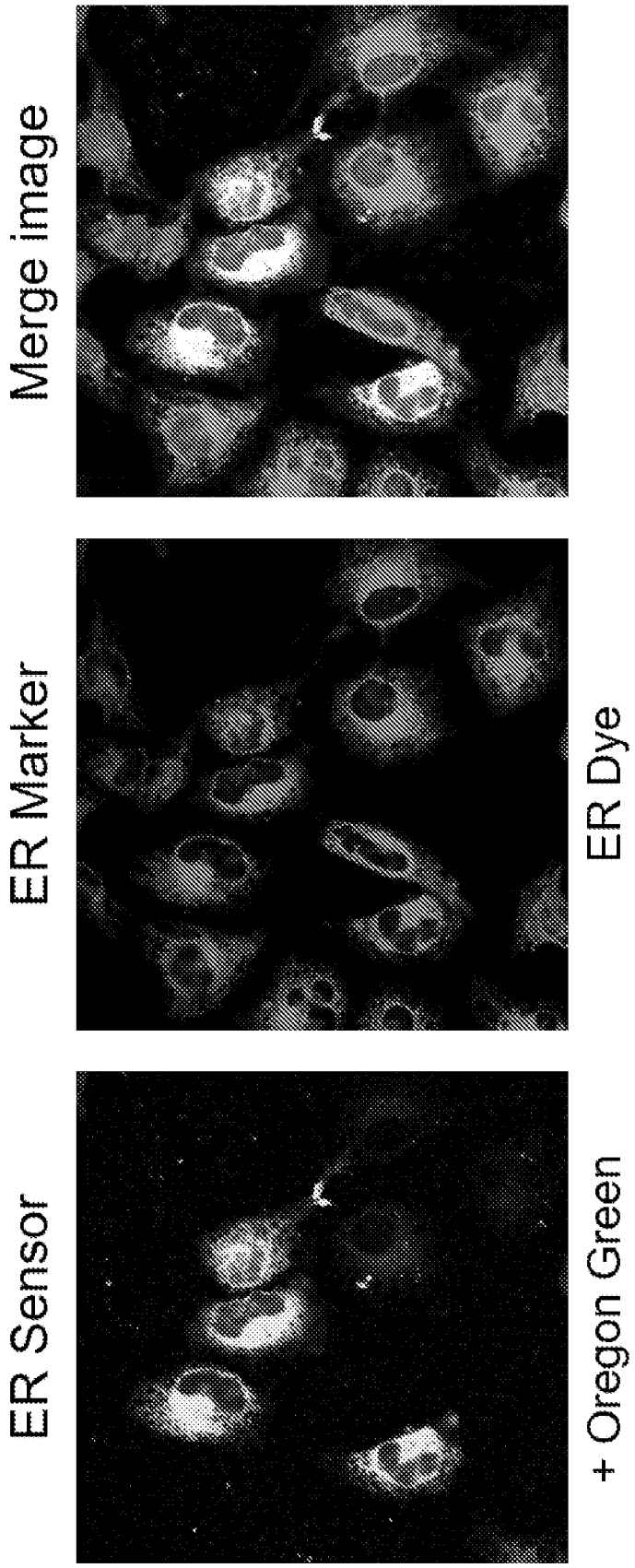


FIG. 222

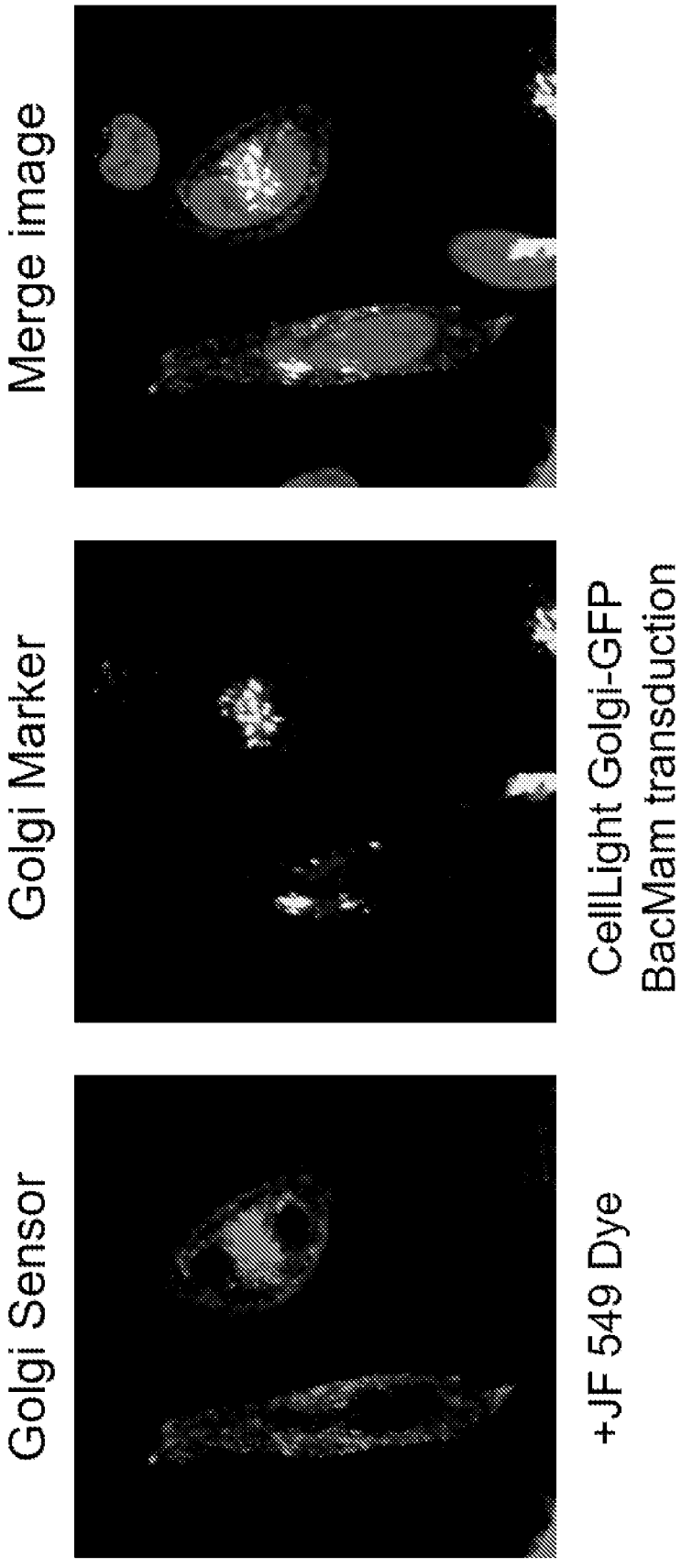

Golgi Sensor
+JF 549 Dye
Golgi Marker
CellLight Golgi-GFP
BacMam transduction
Merge image

FIG. 223

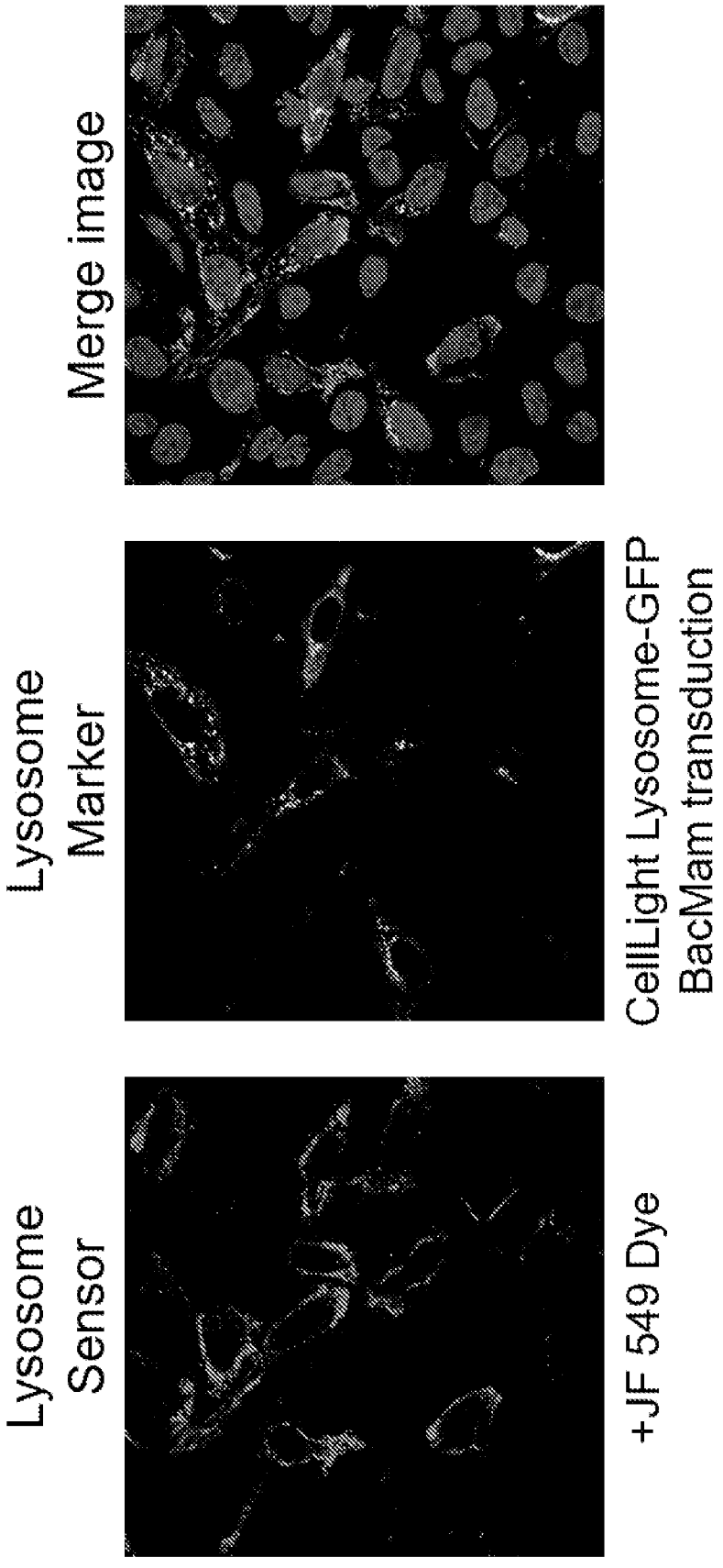

Lysosome Sensor
+JF 549 Dye
Lysosome Marker
CellLight Lysosome-GFP BacMam transduction
Merge image

FIG. 224

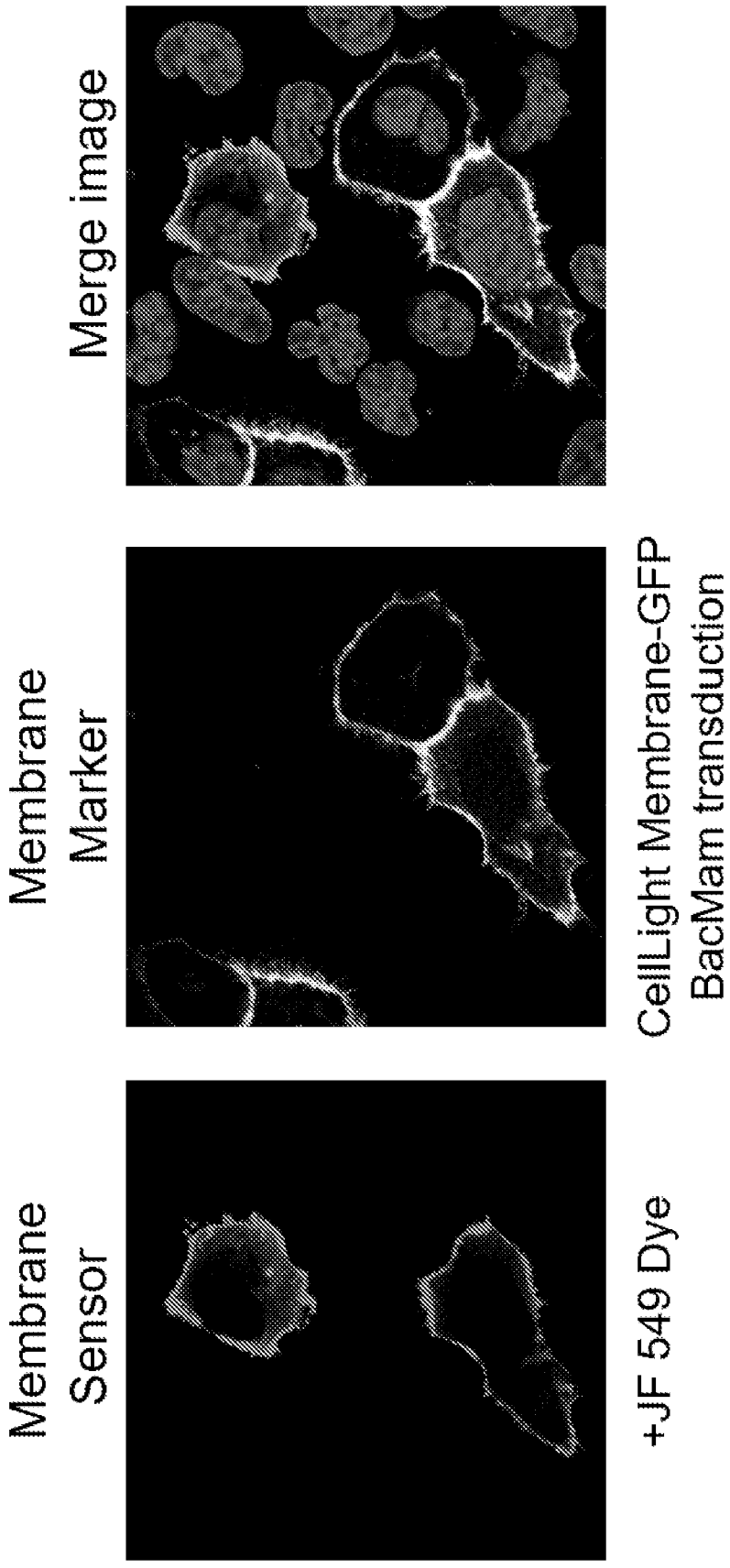
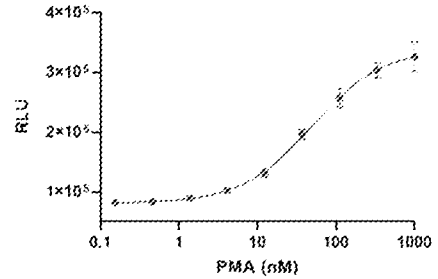
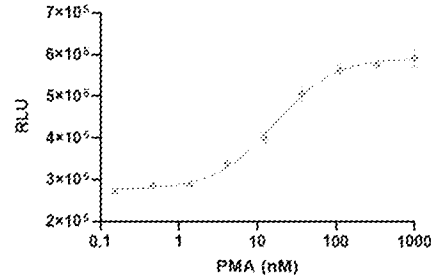

Membrane Sensor
+JF 549 Dye
Membrane Marker
CellLight Membrane-GFP BacMam transduction
Merge image

FIG. 225

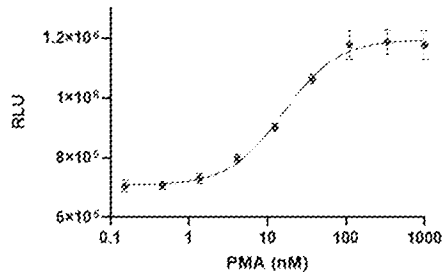
Transient transfected with Membrane Sensor (ATG 5534)



Transient transfected with Membrane Sensor (ATG 5688)



Transient transfected with Membrane Sensor (ATG 5655)



	EC_{50} (nM)
ATG 5534	48
ATG 5688	17
ATG 5655	16

Under PMA stimulus, PKCα translocates from cytosol to plasma member in a dose dependent manner

FIG. 226

	EC_{50} (nM)
ATG 5534	48
HaloTag-LgBiT	42

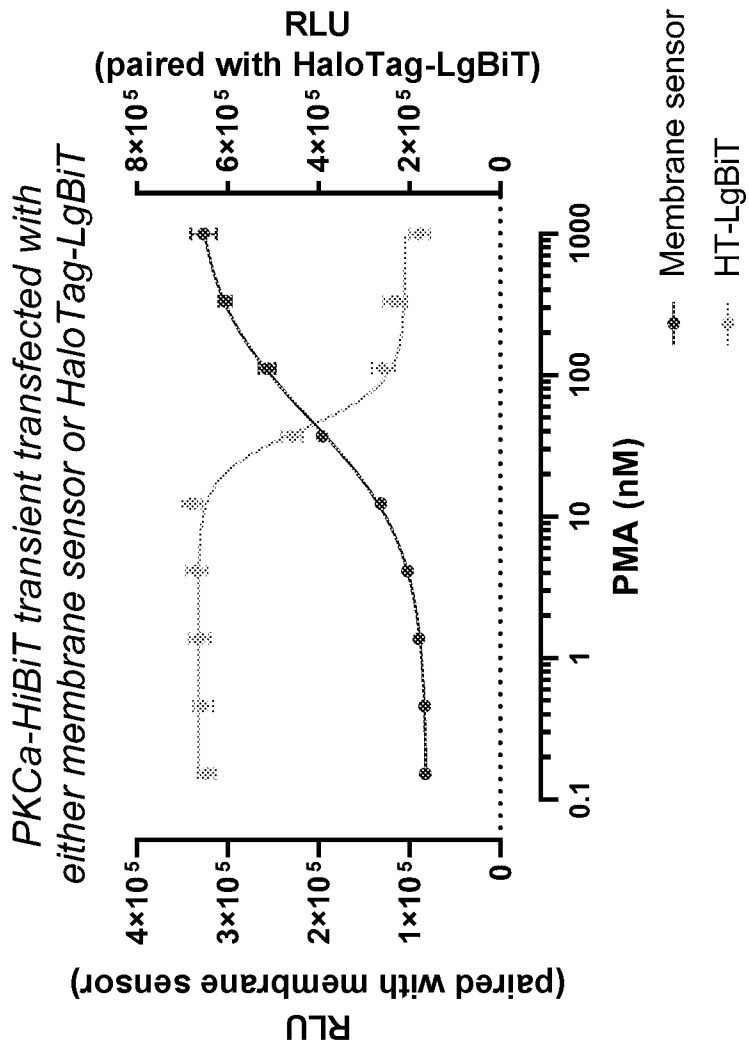
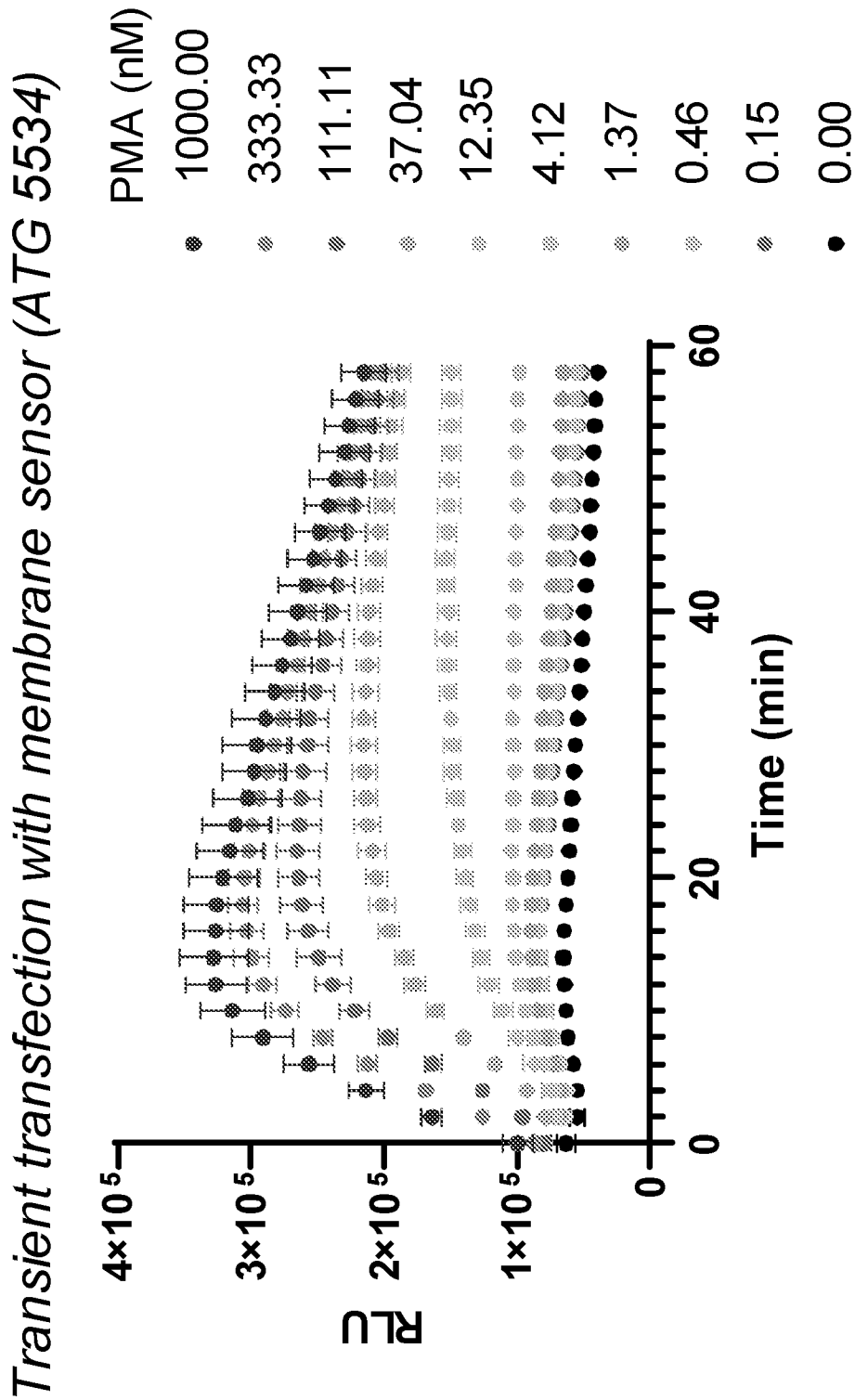


FIG. 227


PKCa-HiBiT
Transient transfection with membrane sensor (ATG 5534)
PMA (nM)
1000.00
333.33
111.11
37.04
12.35
4.12
1.37
0.46
0.15
0.00
RLU
4×10^5
3×10^5
2×10^5
1×10^5
0
0
20
40
60
Time (min)

SEQUENCE LISTING

<110> PROMEGA CORPORATION

<120> MULTIPARTITE LUCIFERASE PEPTIDES AND POLYPEPTIDES

<130> PRMG-38144.601

<150> US 62/941,255
<151> 2019-11-27

<160> 986

<170> PatentIn version 3.5

<210> 1
<211> 170
<212> PRT
<213> Oplophorus gracilirostris

<400> 1

```
Met Phe Thr Leu Ala Asp Phe Val Gly Asp Trp Gln Gln Thr Ala Gly
1               5                  10                  15


Tyr Asn Gln Asp Gln Val Leu Glu Gln Gly Gly Leu Ser Ser Leu Phe
            20                  25                  30


Gln Ala Leu Gly Val Ser Val Thr Pro Ile Gln Lys Val Val Leu Ser
        35                  40                  45


Gly Glu Asn Gly Leu Lys Ala Asp Ile His Val Ile Ile Pro Tyr Glu
    50                  55                  60


Gly Leu Ser Gly Phe Gln Met Gly Leu Ile Glu Met Ile Phe Lys Val
65                  70                  75                  80


Val Tyr Pro Val Asp Asp His His Phe Lys Ile Ile Leu His Tyr Gly
                85                  90                  95


Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Ile Asp Tyr Phe Gly
            100                 105                 110


Arg Pro Tyr Pro Gly Ile Ala Val Phe Asp Gly Lys Gln Ile Thr Val
        115                 120                 125
```

Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Tyr Asp Glu Arg Leu Ile
130 135 140

Asn Pro Asp Gly Ser Leu Leu Phe Arg Val Thr Ile Asn Gly Val Thr
145 150 155 160

Gly Trp Arg Leu Cys Glu Asn Ile Leu Ala
165 170

<210> 2
<211> 513
<212> DNA
<213> Oplophorus gracilirostris

<400> 2
atggtgttta ccttggcaga tttcgttgga gactggcaac agacagctgg atacaaccaa 60

gatcaagtgt tagaacaagg aggattgtct agtctgttcc aagccctggg agtgtcagtc 120

accccaatcc agaaagttgt gctgtctggg gagaatgggt taaaagctga tattcatgtc 180

atcatccctt acgagggact cagtggtttt caaatgggtc tgattgaaat gatcttcaaa 240

gttgtttacc cagtggatga tcatcatttc aagattattc tccattatgg tacactcgtt 300

attgacggtg tgacaccaaa catgattgac tactttggac gcccttaccc tggaattgct 360

gtgtttgacg gcaagcagat cacagttact ggaactctgt ggaacggcaa caagatctat 420

gatgagcgcc tgatcaaccc agatggttca ctcctcttcc gcgttactat caatggagtc 480

accggatggc gcctttgcga gaacattctt gcc 513

<210> 3
<211> 183
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 3

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Arg Gln Thr Ala Gly Tyr Asn Leu Asp
20 25 30

```
Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Phe Gln Asn Leu Gly
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Leu Ser Gly Glu Asn Gly
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Gly
65                  70                  75                  80

Asp Gln Met Gly Gln Ile Glu Lys Ile Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Ile Asp Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Asn Pro Asp Gly
145                 150                 155                 160

Ser Leu Leu Phe Arg Val Thr Ile Asn Gly Val Thr Gly Trp Arg Leu
                165                 170                 175

Cys Glu Arg Ile Leu Ala Val
            180
```

<210> 4
<211> 549
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 4
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggc gacagacagc cggctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgt ttcagaatct cggggtgtcc gtaactccga tccaaaggat tgtcctgagc 180

ggtgaaaatg ggctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcggc 240

gaccaaatgg gccagatcga aaaaattttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcacta tggcacactg gtaatcgacg gggttacgcc gaacatgatc 360

gactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcaa ccccgacggc 480

tccctgctgt tccgagtaac catcaacgga gtgaccggct ggcggctgtg cgaacgcatt 540

ctggcggtt 549

<210> 5
<211> 147
<212> PRT
<213> Oplophorus gracilirostris

<400> 5

Met Phe Thr Leu Ala Asp Phe Val Gly Asp Trp Gln Gln Thr Ala Gly
1 5 10 15

Tyr Asn Gln Asp Gln Val Leu Glu Gln Gly Gly Leu Ser Ser Leu Phe
20 25 30

Gln Ala Leu Gly Val Ser Val Thr Pro Ile Gln Lys Val Val Leu Ser
35 40 45

Gly Glu Asn Gly Leu Lys Ala Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Gly Phe Gln Met Gly Leu Ile Glu Met Ile Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Ile Ile Leu His Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Ile Asp Tyr Phe Gly
100 105 110

Arg Pro Tyr Pro Gly Ile Ala Val Phe Asp Gly Lys Gln Ile Thr Val
115 120 125

```
Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Tyr Asp Glu Arg Leu Ile
    130                 135                 140


Asn Pro Asp
145


<210>  6
<211>  10
<212>  PRT
<213>  Oplophorus gracilirostris

<400>  6

Gly Ser Leu Leu Phe Arg Val Thr Ile Asn
1               5                   10


<210>  7
<211>  13
<212>  PRT
<213>  Oplophorus gracilirostris

<400>  7

Gly Val Thr Gly Trp Arg Leu Cys Glu Asn Ile Leu Ala
1               5                   10


<210>  8
<211>  148
<212>  PRT
<213>  Oplophorus gracilirostris

<400>  8

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Arg Gln Thr Ala
1               5                   10                  15


Gly Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
            20                  25                  30


Phe Gln Asn Leu Gly Val Ser Val Thr Pro Ile Gln Arg Ile Val Leu
        35                  40                  45


Ser Gly Glu Asn Gly Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
    50                  55                  60
```

```
Glu Gly Leu Ser Gly Asp Gln Met Gly Gln Ile Glu Lys Ile Phe Lys
65                  70                  75                  80


Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr
                85                  90                  95


Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Ile Asp Tyr Phe
            100                 105                 110


Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125


Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140


Ile Asn Pro Asp
145
```

<210> 9
<211> 11
<212> PRT
<213> Oplophorus gracilirostris

<400> 9

```
Gly Ser Leu Leu Phe Arg Val Thr Ile Asn Val
1               5                   10
```

<210> 10
<211> 13
<212> PRT
<213> Oplophorus gracilirostris

<400> 10

```
Gly Val Thr Gly Trp Arg Leu Cys Glu Arg Ile Leu Ala
1               5                   10
```

<210> 11
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 11

```
Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1               5                  10                  15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
            20                  25                  30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg
        35                  40                  45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
    50                  55                  60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65                  70                  75                  80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His
145                 150                 155                 160

His His His His His
                165
```

<210> 12
<211> 495
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 12
atggtcttca cactcgaaga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgctgttcc gagtaaccat caacagccat 480

catcaccatc accac 495

<210> 13
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 13

Val Thr Gly Tyr Arg Leu Phe Glu Glu Ile Leu
1 5 10

<210> 14
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 14
gtgaccggct accggctgtt cgaggagatt ctg 33

<210> 15
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 15

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 16
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 16
gtgagcggct ggcggctgtt caagaagatt agc 33

<210> 17
<211> 148
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 17

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe

100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp
145

<210> 18
<211> 444
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 18
atggtcttca cactcgaaga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgac 444

<210> 19
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 19

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 20
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 20
atgaaacatc accatcacca tcatgtcttc acactcgaag atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc caaaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acatgctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactgtaac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac 465

<210> 21
<211> 148
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 21

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp
145

<210> 22
<211> 444
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 22
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgac 444

<210> 23
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 23

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 24
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 24

ggctccatgc tgttccgagt aaccatcaac agc 33

<210> 25
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 25

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 26
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 26
gtgagcggct ggcggctgtt caagaagatt agc 33

<210> 27
<211> 148
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 27

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Phe Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Leu
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

```
Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Lys Ile Phe Lys
65                  70                  75                  80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Ile Asn Tyr Phe
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140

Ile Thr Pro Asp
145
```

<210> 28
<211> 444
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 28
atggtcttca cactcgaaga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgtttc agaatctcgc cgtgtccgta 120

actccgatcc aaaggattgt cctgagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaaa aatttttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgcactatgg cacactggta 300

atcgacgggg ttacgccgaa catgatcaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgac 444

<210> 29
<211> 11
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 29

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Val
1 5 10

<210> 30
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 30
ggctccatgc tgttccgagt aaccatcaac 30

<210> 31
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 31

Met Lys His His His His His His Val Phe Thr Leu Glu Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His

85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 32
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 32
atgaaacatc accatcacca tcatgtcttc acactcgaag atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc caaaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acatgctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactgtaac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac 465

<210> 33
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 33

aagtcgagaa ccatgaccac 20

<210> 34
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 34
cgccgcgtcc gccgaagttg 20

<210> 35
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 35

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 36
<211> 66
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 36
ggctccatgc tgttccgagt aaccatcaac agcgtgagcg gctggcggct gttcaagaag 60

attagc 66

<210> 37
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 37

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 38
<211> 48
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 38
agcagctgga agcgcggctc catgctgttc cgagtaacca tcaacagc 48

<210> 39
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 39

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Asp Thr

```
           100                 105                 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Asp Gly Ile Ala Val
        115                 120                 125

Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn
    130                 135                 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145                 150                 155
```

<210> 40
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 40

```
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa       60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg      120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc      180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc      240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc      300

ctgccctatg gcacactggt aatcgacggg gatacgccga acaagctgaa ctatttcgga      360

cggccgtatg atggcatcgc cgtgttcgac ggcaaaaaga tcactgtaac agggaccctg      420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac                      465
```

<210> 41
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 41

```
Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1               5                  10                  15
```

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Ser Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 42
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 42
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga gcaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactgtaac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac 465

<210> 43
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 43

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Phe Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 44
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 44
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcttcgc cgtgttcgac ggcaaaaaga tcactgtaac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac 465

<210> 45
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 45

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp

50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Cys Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 46
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 46
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgtgcgac ggcaaaaaga tcactgtaac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac 465

<210> 47

<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 47

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Ser Val Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 48
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 48
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcggg 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tctctgtaac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac 465

<210> 49
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 49

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
            20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
        35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
    50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
                85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Ala Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 50
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 50
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactgcaac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac 465

<210> 51
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 51

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val

1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 52
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 52
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac 465

<210> 53
<211> 149
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 53

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu

130 135 140

Ile Thr Pro Asp Gly
145

<210> 54
<211> 447
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 54
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggc 447

<210> 55
<211> 147
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 55

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

```
Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
    50                  55                  60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65                  70                  75                  80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140

Ile Thr Pro
145
```

<210> 56
<211> 441
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 56

```
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg       60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta      120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc      180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag      240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta      300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc      360

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc      420

gacgagcgcc tgatcacccc c                                                441
```

<210> 57
<211> 146
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 57

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr
145

<210> 58
<211> 438
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic

<400> 58
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacc 438

<210> 59
<211> 150
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 59

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr

85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser
145 150

<210> 60
<211> 450
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 60
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggcagc 450

<210> 61
<211> 164
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 61

```
Met Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala
1               5                  10                  15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20                  25                  30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser
        35                  40                  45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    50                  55                  60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65                  70                  75                  80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85                  90                  95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115                 120                 125

Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140

Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His His
145                 150                 155                 160

His His His His
```

<210> 62
<211> 495
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 62

atgttcacac tcgaagattt cgttggggac tgggaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggattgtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcaacaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg ctgttccgag taaccatcaa cagccatcat 480

caccatcacc actaa 495

<210> 63
<211> 163
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 63

Met Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg

100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His His His
145 150 155 160

His His His

<210> 64
<211> 492
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 64
atgacactcg aagatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa 60

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg 120

atccaaagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc 180

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg 240

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac 300

ggggttacgc cgaacatgct gaactatttc ggacggccgt atgaaggcat cgccgtgttc 360

gacggcaaaa agatcactgt aacagggacc ctgtggaacg gcaacaaaat tatcgacgag 420

cgcctgatca cccccgacgg ctccatgctg ttccgagtaa ccatcaacag ccatcatcac 480

catcaccact aa 492

<210> 65
<211> 162
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 65

Met Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn
1 5 10 15

Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn
20 25 30

Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu
35 40 45

Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu
50 55 60

Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr
65 70 75 80

Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu
85 90 95

Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro
100 105 110

Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly
115 120 125

Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro
130 135 140

Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His His His His
145 150 155 160

His His

<210> 66
<211> 489
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 66
atgctcgaag atttcgttgg ggactgggaa cagacagccg cctacaacct ggaccaagtc 60

cttgaacagg gaggtgtgtc cagtttgctg cagaatctcg ccgtgtccgt aactccgatc 120

caaaggattg tccggagcgg tgaaaatgcc ctgaagatcg acatccatgt catcatcccg 180

tatgaaggtc tgagcgccga ccaaatggcc cagatcgaag aggtgtttaa ggtggtgtac 240

cctgtggatg atcatcactt taaggtgatc ctgccctatg gcacactggt aatcgacggg 300

gttacgccga acatgctgaa ctatttcgga cggccgtatg aaggcatcgc cgtgttcgac 360

ggcaaaaaga tcactgtaac agggaccctg tggaacggca acaaaattat cgacgagcgc 420

ctgatcaccc ccgacggctc catgctgttc cgagtaacca tcaacagcca tcatcaccat 480

caccactaa 489

<210> 67
<211> 161
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 67

Met Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu
1 5 10 15

Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu
20 25 30

Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn
35 40 45

Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser
50 55 60

Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro
65 70 75 80

Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val
85 90 95

Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr
100 105 110

Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr
115 120 125

Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
130 135 140

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His His His His His
145 150 155 160

His

<210> 68
<211> 486
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 68
atggaagatt tcgttgggga ctgggaacag acagccgcct acaacctgga ccaagtcctt 60

gaacagggag gtgtgtccag tttgctgcag aatctcgccg tgtccgtaac tccgatccaa 120

aggattgtcc ggagcggtga aaatgccctg aagatcgaca tccatgtcat catcccgtat 180

gaaggtctga gcgccgacca aatggcccag atcgaagagg tgtttaaggt ggtgtaccct 240

gtggatgatc atcactttaa ggtgatcctg ccctatggca cactggtaat cgacggggtt 300

acgccgaaca tgctgaacta tttcggacgg ccgtatgaag gcatcgccgt gttcgacggc 360

aaaaagatca ctgtaacagg gaccctgtgg aacggcaaca aaattatcga cgagcgcctg 420

atcacccccg acggctccat gctgttccga gtaaccatca acagccatca tcaccatcac 480

cactaa 486

<210> 69
<211> 174
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 69

Met Phe Lys Lys Ile Ser Gly Ser Ser Gly Val Phe Thr Leu Glu Asp
1   5   10   15

Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val
            20   25   30

Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser
        35   40   45

Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys
    50   55   60

Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln
65   70   75   80

Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp
                85   90   95

His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly
            100   105   110

Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile
        115   120   125

Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn
    130   135   140

Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met
145   150   155   160

Leu Phe Arg Val Thr Ile Asn Ser His His His His His His
                165   170

<210> 70
<211> 525
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 70
atgttcaaga agattagcgg ctcgagcggt gtcttcacac tcgaagattt cgttggggac 60

tgggaacaga cagccgccta caacctggac caagtccttg aacagggagg tgtgtccagt 120

ttgctgcaga atctcgccgt gtccgtaact ccgatccaaa ggattgtccg gagcggtgaa 180

aatgccctga agatcgacat ccatgtcatc atcccgtatg aaggtctgag cgccgaccaa 240

atggcccaga tcgaagaggt gtttaaggtg gtgtaccctg tggatgatca tcactttaag 300

gtgatcctgc cctatggcac actggtaatc gacggggtta cgccgaacat gctgaactat 360

ttcggacggc cgtatgaagg catcgccgtg ttcgacggca aaaagatcac tgtaacaggg 420

accctgtgga acggcaacaa aattatcgac gagcgcctga tcacccccga cggctccatg 480

ctgttccgag taaccatcaa cagccatcat caccatcacc actaa 525

<210> 71
<211> 173
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 71

Met Lys Lys Ile Ser Gly Ser Ser Gly Val Phe Thr Leu Glu Asp Phe
1   5   10   15

Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu
            20   25   30

Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val
        35   40   45

Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile
    50   55   60

Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met
65   70   75   80

Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His

85 90 95

His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val
100 105 110

Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala
115 120 125

Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly
130 135 140

Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu
145 150 155 160

Phe Arg Val Thr Ile Asn Ser His His His His His His
165 170

<210> 72
<211> 522
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 72
atgaagaaga ttagcggctc gagcggtgtc ttcacactcg aagatttcgt tggggactgg 60

gaacagacag ccgcctacaa cctggaccaa gtccttgaac agggaggtgt gtccagtttg 120

ctgcagaatc tcgccgtgtc cgtaactccg atccaaagga ttgtccggag cggtgaaaat 180

gccctgaaga tcgacatcca tgtcatcatc ccgtatgaag gtctgagcgc cgaccaaatg 240

gcccagatcg aagaggtgtt taaggtggtg taccctgtgg atgatcatca ctttaaggtg 300

atcctgccct atggcacact ggtaatcgac ggggttacgc cgaacatgct gaactatttc 360

ggacggccgt atgaaggcat cgccgtgttc gacggcaaaa agatcactgt aacagggacc 420

ctgtggaacg gcaacaaaat tatcgacgag cgcctgatca cccccgacgg ctccatgctg 480

ttccgagtaa ccatcaacag ccatcatcac catcaccact aa 522

<210> 73
<211> 172
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 73

Met Lys Ile Ser Gly Ser Ser Gly Val Phe Thr Leu Glu Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe
145 150 155 160

Arg Val Thr Ile Asn Ser His His His His His His
165 170

<210> 74
<211> 519
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic

<400> 74
atgaagatta gcggctcgag cggtgtcttc acactcgaag atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc caaaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acatgctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactgtaac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgacggctc catgctgttc 480

cgagtaacca tcaacagcca tcatcaccat caccactaa 519

<210> 75
<211> 171
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 75

Met Ile Ser Gly Ser Ser Gly Val Phe Thr Leu Glu Asp Phe Val Gly
1 5 10 15

Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln
20 25 30

Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro
35 40 45

Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile
50 55 60

His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln
65 70 75 80

```
Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe
                85                  90                  95

Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro
            100                 105                 110

Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe
        115                 120                 125

Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn Lys
    130                 135                 140

Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe Arg
145                 150                 155                 160

Val Thr Ile Asn Ser His His His His His His
                165                 170
```

<210> 76
<211> 516
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 76

```
atgattagcg gctcgagcgg tgtcttcaca ctcgaagatt tcgttgggga ctgggaacag       60

acagccgcct acaacctgga ccaagtcctt gaacagggag gtgtgtccag tttgctgcag      120

aatctcgccg tgtccgtaac tccgatccaa aggattgtcc ggagcggtga aaatgccctg      180

aagatcgaca tccatgtcat catcccgtat gaaggtctga gcgccgacca aatggcccag      240

atcgaagagg tgtttaaggt ggtgtaccct gtggatgatc atcactttaa ggtgatcctg      300

ccctatggca cactggtaat cgacggggtt acgccgaaca tgctgaacta tttcggacgg      360

ccgtatgaag gcatcgccgt gttcgacggc aaaaagatca ctgtaacagg gaccctgtgg      420

aacggcaaca aaattatcga cgagcgcctg atcacccccg acggctccat gctgttccga      480

gtaaccatca acagccatca tcaccatcac cactaa                                 516
```

<210> 77

<211> 170
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 77

Met Ser Gly Ser Ser Gly Val Phe Thr Leu Glu Asp Phe Val Gly Asp
1 5 10 15

Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly
20 25 30

Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile
35 40 45

Gln Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His
50 55 60

Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile
65 70 75 80

Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys
85 90 95

Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn
100 105 110

Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp
115 120 125

Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile
130 135 140

Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val
145 150 155 160

Thr Ile Asn Ser His His His His His His
165 170

<210> 78

<211> 513
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 78
atgagcggct cgagcggtgt cttcacactc gaagatttcg ttggggactg ggaacagaca 60

gccgcctaca acctggacca agtccttgaa cagggaggtg tgtccagttt gctgcagaat 120

ctcgccgtgt ccgtaactcc gatccaaagg attgtccgga gcggtgaaaa tgccctgaag 180

atcgacatcc atgtcatcat cccgtatgaa ggtctgagcg ccgaccaaat ggcccagatc 240

gaagaggtgt ttaaggtggt gtaccctgtg gatgatcatc actttaaggt gatcctgccc 300

tatggcacac tggtaatcga cggggttacg ccgaacatgc tgaactattt cggacggccg 360

tatgaaggca tcgccgtgtt cgacggcaaa aagatcactg taacagggac cctgtggaac 420

ggcaacaaaa ttatcgacga gcgcctgatc acccccgacg gctccatgct gttccgagta 480

accatcaaca gccatcatca ccatcaccac taa 513

<210> 79
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 79

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala

65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met
145 150 155

<210> 80
<211> 477
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 80
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgacggcag catgtaa 477

<210> 81
<211> 159
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 81

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu
145 150 155

<210> 82
<211> 480
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 82

atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgacggcag catgctgtaa 480

<210> 83
<211> 160
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 83

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

```
Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
        115                 120                 125


Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
    130                 135                 140


Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe
145                 150                 155                 160


<210>  84
<211>  483
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  84
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa       60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg      120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc      180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc      240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc      300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga      360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg      420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgacggcag catgctgttc      480

taa                                                                    483


<210>  85
<211>  152
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  85

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1               5                   10                  15
```

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile
145 150

<210> 86
<211> 459
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 86
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatctaa 459

<210> 87
<211> 151
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 87

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu
145 150

<210> 88
<211> 456
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 88
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgtaa 456

<210> 89
<211> 150
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 89

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp

```
    50                  55                  60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65                  70                  75                  80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
                85                  90                  95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
            100                 105                 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
        115                 120                 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
    130                 135                 140

Lys Ile Ile Asp Glu Arg
145                 150


<210>  90
<211>  453
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  90
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa       60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg      120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc      180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc      240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc      300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga      360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg      420

tggaacggca acaaaattat cgacgagcgc taa                                   453


<210>  91
```

<211> 123
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 91

Met Val Ala Ile Leu Trp His Glu Met Trp His Glu Gly Leu Glu Glu
1 5 10 15

Ala Ser Arg Leu Tyr Phe Gly Glu Arg Asn Val Lys Gly Met Phe Glu
20 25 30

Val Leu Glu Pro Leu His Ala Met Met Glu Arg Gly Pro Gln Thr Leu
35 40 45

Lys Glu Thr Ser Phe Asn Gln Ala Tyr Gly Arg Asp Leu Met Glu Ala
50 55 60

Gln Glu Trp Cys Arg Lys Tyr Met Lys Ser Gly Asn Val Lys Asp Leu
65 70 75 80

Thr Gln Ala Trp Asp Leu Tyr Tyr His Val Phe Arg Arg Ile Ser Gly
85 90 95

Gly Ser Gly Gly Gly Gly Ser Gly Gly Ser Ser Ser Gly Gly Ala Ile
100 105 110

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
115 120

<210> 92
<211> 372
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 92
atggtggcca tcctctggca tgagatgtgg catgaaggcc tggaagaggc atctcgtttg 60

tactttgggg aaaggaacgt gaaaggcatg tttgaggtgc tggagccctt gcatgctatg 120

atggaacggg gcccccagac tctgaaggaa acatccttta atcaggccta tggtcgagat 180

ttaatggagg cccaagagtg gtgcaggaag tacatgaaat cagggaatgt caaggacctc 240

acccaagcct gggacctcta ttatcatgtg ttccgacgaa tcagtggtgg ttcaggtggt 300

ggcgggagcg gtggctcgag cagcggtgga gcgatcgtga gcggctggcg gctgttcaag 360

aagattagct aa 372

<210> 93
<211> 125
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 93

Met Val Ala Ile Leu Trp His Glu Met Trp His Glu Gly Leu Glu Glu
1 5 10 15

Ala Ser Arg Leu Tyr Phe Gly Glu Arg Asn Val Lys Gly Met Phe Glu
20 25 30

Val Leu Glu Pro Leu His Ala Met Met Glu Arg Gly Pro Gln Thr Leu
35 40 45

Lys Glu Thr Ser Phe Asn Gln Ala Tyr Gly Arg Asp Leu Met Glu Ala
50 55 60

Gln Glu Trp Cys Arg Lys Tyr Met Lys Ser Gly Asn Val Lys Asp Leu
65 70 75 80

Thr Gln Ala Trp Asp Leu Tyr Tyr His Val Phe Arg Arg Ile Ser Gly
85 90 95

Gly Ser Gly Gly Gly Gly Ser Gly Gly Ser Ser Ser Gly Gly Ala Ile
100 105 110

Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
115 120 125

<210> 94

<211> 378
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 94
atggtggcca tcctctggca tgagatgtgg catgaaggcc tggaagaggc atctcgtttg 60

tactttgggg aaaggaacgt gaaaggcatg tttgaggtgc tggagccctt gcatgctatg 120

atggaacggg gcccccagac tctgaaggaa acatccttta atcaggccta tggtcgagat 180

ttaatggagg cccaagagtg gtgcaggaag tacatgaaat cagggaatgt caaggacctc 240

acccaagcct gggacctcta ttatcatgtg ttccgacgaa tcagtggtgg ttcaggtggt 300

ggcgggagcg gtggctcgag cagcggtgga gcgatcgtta gcgttagcgg ctggcgcctg 360

ttcaagaaga tcagctaa 378

<210> 95
<211> 129
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 95

Met Val Ala Ile Leu Trp His Glu Met Trp His Glu Gly Leu Glu Glu
1 5 10 15

Ala Ser Arg Leu Tyr Phe Gly Glu Arg Asn Val Lys Gly Met Phe Glu
20 25 30

Val Leu Glu Pro Leu His Ala Met Met Glu Arg Gly Pro Gln Thr Leu
35 40 45

Lys Glu Thr Ser Phe Asn Gln Ala Tyr Gly Arg Asp Leu Met Glu Ala
50 55 60

Gln Glu Trp Cys Arg Lys Tyr Met Lys Ser Gly Asn Val Lys Asp Leu
65 70 75 80

Thr Gln Ala Trp Asp Leu Tyr Tyr His Val Phe Arg Arg Ile Ser Gly

85 90 95

Gly Ser Gly Gly Gly Gly Ser Gly Gly Ser Ser Ser Gly Gly Ala Ile
100 105 110

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser His His His His His
115 120 125

His

<210> 96
<211> 390
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 96
atggtggcca tcctctggca tgagatgtgg catgaaggcc tggaagaggc atctcgtttg 60

tactttgggg aaaggaacgt gaaaggcatg tttgaggtgc tggagccctt gcatgctatg 120

atggaacggg gcccccagac tctgaaggaa acatccttta atcaggccta tggtcgagat 180

ttaatggagg cccaagagtg gtgcaggaag tacatgaaat cagggaatgt caaggacctc 240

acccaagcct gggacctcta ttatcatgtg ttccgacgaa tcagtggtgg ttcaggtggt 300

ggcgggagcg gtggctcgag cagcggtgga gcgatcgtga gcggctggcg gctgttcaag 360

aagattagcc atcatcacca tcaccactaa 390

<210> 97
<211> 131
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 97

Met Val Ala Ile Leu Trp His Glu Met Trp His Glu Gly Leu Glu Glu
1 5 10 15

Ala Ser Arg Leu Tyr Phe Gly Glu Arg Asn Val Lys Gly Met Phe Glu

```
            20                  25                  30

Val Leu Glu Pro Leu His Ala Met Met Glu Arg Gly Pro Gln Thr Leu
        35                  40                  45

Lys Glu Thr Ser Phe Asn Gln Ala Tyr Gly Arg Asp Leu Met Glu Ala
    50                  55                  60

Gln Glu Trp Cys Arg Lys Tyr Met Lys Ser Gly Asn Val Lys Asp Leu
65                  70                  75                  80

Thr Gln Ala Trp Asp Leu Tyr Tyr His Val Phe Arg Arg Ile Ser Gly
                85                  90                  95

Gly Ser Gly Gly Gly Gly Ser Gly Gly Ser Ser Ser Gly Gly Ala Ile
            100                 105                 110

Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser His His His
        115                 120                 125

His His His
    130
```

<210> 98
<211> 396
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 98

```
atggtggcca tcctctggca tgagatgtgg catgaaggcc tggaagaggc atctcgtttg       60

tactttgggg aaaggaacgt gaaaggcatg tttgaggtgc tggagccctt gcatgctatg      120

atggaacggg gcccccagac tctgaaggaa acatccttta atcaggccta tggtcgagat      180

ttaatggagg cccaagagtg gtgcaggaag tacatgaaat cagggaatgt caaggacctc      240

acccaagcct gggacctcta ttatcatgtg ttccgacgaa tcagtggtgg ttcaggtggt      300

ggcgggagcg gtggctcgag cagcggtgga gcgatcgtta gcgtgagcgg ctggcggctg      360

ttcaagaaga ttagccatca tcaccatcac cactaa                                396
```

<210> 99
<211> 118
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 99

Met Lys His His His His His His Val Ala Ile Leu Trp His Glu Met
1 5 10 15

Trp His Glu Gly Leu Glu Glu Ala Ser Arg Leu Tyr Phe Gly Glu Arg
20 25 30

Asn Val Lys Gly Met Phe Glu Val Leu Glu Pro Leu His Ala Met Met
35 40 45

Glu Arg Gly Pro Gln Thr Leu Lys Glu Thr Ser Phe Asn Gln Ala Tyr
50 55 60

Gly Arg Asp Leu Met Glu Ala Gln Glu Trp Cys Arg Lys Tyr Met Lys
65 70 75 80

Ser Gly Asn Val Lys Asp Leu Thr Gln Ala Trp Asp Leu Tyr Tyr His
85 90 95

Val Phe Arg Arg Ile Ser Gly Gly Ser Gly Gly Val Ser Gly Trp Arg
100 105 110

Leu Phe Lys Lys Ile Ser
115

<210> 100
<211> 357
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 100
atgaaacatc accatcacca tcatgtggcc atcctctggc atgagatgtg gcatgaaggc 60

ctggaagagg catctcgttt gtactttggg gaaaggaacg tgaaaggcat gtttgaggtg 120

ctggagccct tgcatgctat gatggaacgg ggcccccaga ctctgaagga aacatccttt 180

aatcaggcct atggtcgaga tttaatggag gcccaagagt ggtgcaggaa gtacatgaaa 240

tcagggaatg tcaaggacct cacccaagcc tgggacctct attatcatgt gttccgacga 300

atcagtggtg gttcaggtgg tgtgagcggc tggcggctgt tcaagaagat tagctaa 357

<210> 101
<211> 123
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 101

Met Lys His His His His His His Val Ala Ile Leu Trp His Glu Met
1 5 10 15

Trp His Glu Gly Leu Glu Glu Ala Ser Arg Leu Tyr Phe Gly Glu Arg
20 25 30

Asn Val Lys Gly Met Phe Glu Val Leu Glu Pro Leu His Ala Met Met
35 40 45

Glu Arg Gly Pro Gln Thr Leu Lys Glu Thr Ser Phe Asn Gln Ala Tyr
50 55 60

Gly Arg Asp Leu Met Glu Ala Gln Glu Trp Cys Arg Lys Tyr Met Lys
65 70 75 80

Ser Gly Asn Val Lys Asp Leu Thr Gln Ala Trp Asp Leu Tyr Tyr His
85 90 95

Val Phe Arg Arg Ile Ser Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
100 105 110

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
115 120

<210> 102

<211> 372
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 102
atgaaacatc accatcacca tcatgtggcc atcctctggc atgagatgtg gcatgaaggc 60

ctggaagagg catctcgttt gtactttggg gaaaggaacg tgaaaggcat gtttgaggtg 120

ctggagccct tgcatgctat gatggaacgg ggcccccaga ctctgaagga aacatccttt 180

aatcaggcct atggtcgaga tttaatggag gcccaagagt ggtgcaggaa gtacatgaaa 240

tcagggaatg tcaaggacct cacccaagcc tgggacctct attatcatgt gttccgacga 300

atcagtggtg gttcaggtgg tggcgggagc ggtggcgtga gcggctggcg gctgttcaag 360

aagattagct aa 372

<210> 103
<211> 128
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 103

Met Lys His His His His His His Val Ala Ile Leu Trp His Glu Met
1 5 10 15

Trp His Glu Gly Leu Glu Glu Ala Ser Arg Leu Tyr Phe Gly Glu Arg
20 25 30

Asn Val Lys Gly Met Phe Glu Val Leu Glu Pro Leu His Ala Met Met
35 40 45

Glu Arg Gly Pro Gln Thr Leu Lys Glu Thr Ser Phe Asn Gln Ala Tyr
50 55 60

Gly Arg Asp Leu Met Glu Ala Gln Glu Trp Cys Arg Lys Tyr Met Lys
65 70 75 80

Ser Gly Asn Val Lys Asp Leu Thr Gln Ala Trp Asp Leu Tyr Tyr His

85 90 95

Val Phe Arg Arg Ile Ser Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
100 105 110

Ser Ser Ser Gly Gly Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
115 120 125

<210> 104
<211> 387
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 104
atgaaacatc accatcacca tcatgtggcc atcctctggc atgagatgtg gcatgaaggc 60

ctggaagagg catctcgttt gtactttggg gaaaggaacg tgaaaggcat gtttgaggtg 120

ctggagccct tgcatgctat gatggaacgg ggcccccaga ctctgaagga aacatccttt 180

aatcaggcct atggtcgaga tttaatggag gcccaagagt ggtgcaggaa gtacatgaaa 240

tcagggaatg tcaaggacct cacccaagcc tgggacctct attatcatgt gttccgacga 300

atcagtggtg gttcaggtgg tggcgggagc ggtggctcga gcagcggtgg agtgagcggc 360

tggcggctgt tcaagaagat tagctaa 387

<210> 105
<211> 120
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 105

Met Lys His His His His His His Val Ala Ile Leu Trp His Glu Met
1 5 10 15

Trp His Glu Gly Leu Glu Glu Ala Ser Arg Leu Tyr Phe Gly Glu Arg
20 25 30

Asn Val Lys Gly Met Phe Glu Val Leu Glu Pro Leu His Ala Met Met

35 40 45

Glu Arg Gly Pro Gln Thr Leu Lys Glu Thr Ser Phe Asn Gln Ala Tyr
50 55 60

Gly Arg Asp Leu Met Glu Ala Gln Glu Trp Cys Arg Lys Tyr Met Lys
65 70 75 80

Ser Gly Asn Val Lys Asp Leu Thr Gln Ala Trp Asp Leu Tyr Tyr His
85 90 95

Val Phe Arg Arg Ile Ser Gly Gly Ser Gly Gly Val Ser Val Ser Gly
100 105 110

Trp Arg Leu Phe Lys Lys Ile Ser
115 120

<210> 106
<211> 363
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 106
atgaaacatc accatcacca tcatgtggcc atcctctggc atgagatgtg gcatgaaggc 60

ctggaagagg catctcgttt gtactttggg gaaaggaacg tgaaaggcat gtttgaggtg 120

ctggagccct tgcatgctat gatggaacgg ggcccccaga ctctgaagga aacatccttt 180

aatcaggcct atggtcgaga tttaatggag gcccaagagt ggtgcaggaa gtacatgaaa 240

tcagggaatg tcaaggacct cacccaagcc tgggacctct attatcatgt gttccgacga 300

atcagtggtg gttcaggtgg tgttagcgtt agcggctggc gcctgttcaa gaagatcagc 360

taa 363

<210> 107
<211> 125
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 107

Met Lys His His His His His His Val Ala Ile Leu Trp His Glu Met
1 5 10 15

Trp His Glu Gly Leu Glu Glu Ala Ser Arg Leu Tyr Phe Gly Glu Arg
20 25 30

Asn Val Lys Gly Met Phe Glu Val Leu Glu Pro Leu His Ala Met Met
35 40 45

Glu Arg Gly Pro Gln Thr Leu Lys Glu Thr Ser Phe Asn Gln Ala Tyr
50 55 60

Gly Arg Asp Leu Met Glu Ala Gln Glu Trp Cys Arg Lys Tyr Met Lys
65 70 75 80

Ser Gly Asn Val Lys Asp Leu Thr Gln Ala Trp Asp Leu Tyr Tyr His
85 90 95

Val Phe Arg Arg Ile Ser Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
100 105 110

Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
115 120 125

<210> 108
<211> 378
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 108
atgaaacatc accatcacca tcatgtggcc atcctctggc atgagatgtg gcatgaaggc 60

ctggaagagg catctcgttt gtactttggg gaaaggaacg tgaaaggcat gtttgaggtg 120

ctggagccct tgcatgctat gatggaacgg ggcccccaga ctctgaagga aacatccttt 180

aatcaggcct atggtcgaga tttaatggag gcccaagagt ggtgcaggaa gtacatgaaa 240

tcagggaatg tcaaggacct cacccaagcc tgggacctct attatcatgt gttccgacga 300

atcagtggtg gttcaggtgg tggcgggagc ggtggcgtta gcgttagcgg ctggcgcctg 360

ttcaagaaga tcagctaa 378

<210> 109
<211> 130
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 109

Met Lys His His His His His His Val Ala Ile Leu Trp His Glu Met
1 5 10 15

Trp His Glu Gly Leu Glu Glu Ala Ser Arg Leu Tyr Phe Gly Glu Arg
20 25 30

Asn Val Lys Gly Met Phe Glu Val Leu Glu Pro Leu His Ala Met Met
35 40 45

Glu Arg Gly Pro Gln Thr Leu Lys Glu Thr Ser Phe Asn Gln Ala Tyr
50 55 60

Gly Arg Asp Leu Met Glu Ala Gln Glu Trp Cys Arg Lys Tyr Met Lys
65 70 75 80

Ser Gly Asn Val Lys Asp Leu Thr Gln Ala Trp Asp Leu Tyr Tyr His
85 90 95

Val Phe Arg Arg Ile Ser Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
100 105 110

Ser Ser Ser Gly Gly Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys
115 120 125

Ile Ser
130

<210> 110
<211> 393
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic

<400> 110
atgaaacatc accatcacca tcatgtggcc atcctctggc atgagatgtg gcatgaaggc 60

ctggaagagg catctcgttt gtactttggg gaaaggaacg tgaaaggcat gtttgaggtg 120

ctggagccct tgcatgctat gatggaacgg ggcccccaga ctctgaagga aacatccttt 180

aatcaggcct atggtcgaga tttaatggag gcccaagagt ggtgcaggaa gtacatgaaa 240

tcagggaatg tcaaggacct cacccaagcc tgggacctct attatcatgt gttccgacga 300

atcagtggtg gttcaggtgg tggcgggagc ggtggctcga gcagcggtgg agttagcgtt 360

agcggctggc gcctgttcaa gaagatcagc taa 393

<210> 111
<211> 134
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 111

Met Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Ser Ser Ser Gly
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Ser Ser Gly Gly Gly Val Gln Val Glu
20 25 30

Thr Ile Ser Pro Gly Asp Gly Arg Thr Phe Pro Lys Arg Gly Gln Thr
35 40 45

Cys Val Val His Tyr Thr Gly Met Leu Glu Asp Gly Lys Lys Phe Asp
50 55 60

Ser Ser Arg Asp Arg Asn Lys Pro Phe Lys Phe Met Leu Gly Lys Gln
65 70 75 80

Glu Val Ile Arg Gly Trp Glu Glu Gly Val Ala Gln Met Ser Val Gly
85 90 95

```
Gln Arg Ala Lys Leu Thr Ile Ser Pro Asp Tyr Ala Tyr Gly Ala Thr
            100                 105                 110


Gly His Pro Gly Ile Ile Pro Pro His Ala Thr Leu Val Phe Asp Val
        115                 120                 125


Glu Leu Leu Lys Leu Glu
    130


<210>  112
<211>  405
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  112
atgggctcca tgctgttccg agtaaccatc aacagctcga gttcaggtgg tggcgggagc       60

ggtggaggga gcagcggtgg aggagtgcag gtggaaacca tctccccagg agacgggcgc      120

accttcccca agcgcggcca gacctgcgtg gtgcactaca ccgggatgct tgaagatgga      180

aagaaatttg attcctcccg ggacagaaac aagcccttta agtttatgct aggcaagcag      240

gaggtgatcc gaggctggga agaaggggtt gcccagatga gtgtgggtca gagagccaaa      300

ctgactatat ctccagatta tgcctatggt gccactgggc acccaggcat catcccacca      360

catgccactc tcgtcttcga tgtggagctt ctaaaactgg aataa                      405


<210>  113
<211>  136
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  113

Met Gly Val Gln Val Glu Thr Ile Ser Pro Gly Asp Gly Arg Thr Phe
1               5                   10                  15


Pro Lys Arg Gly Gln Thr Cys Val Val His Tyr Thr Gly Met Leu Glu
            20                  25                  30
```

```
Asp Gly Lys Lys Phe Asp Ser Ser Arg Asp Arg Asn Lys Pro Phe Lys
        35                  40                  45

Phe Met Leu Gly Lys Gln Glu Val Ile Arg Gly Trp Glu Glu Gly Val
    50                  55                  60

Ala Gln Met Ser Val Gly Gln Arg Ala Lys Leu Thr Ile Ser Pro Asp
65                  70                  75                  80

Tyr Ala Tyr Gly Ala Thr Gly His Pro Gly Ile Ile Pro Pro His Ala
                85                  90                  95

Thr Leu Val Phe Asp Val Glu Leu Leu Lys Leu Glu Gly Gly Ser Gly
            100                 105                 110

Gly Gly Gly Ser Gly Gly Ser Ser Ser Gly Gly Ala Ile Gly Ser Met
        115                 120                 125

Leu Phe Arg Val Thr Ile Asn Ser
    130                 135
```

<210> 114
<211> 408
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 114
atgggagtgc aggtggaaac catctcccca ggagacgggc gcaccttccc caagcgcggc 60

cagacctgcg tggtgcacta caccgggatg cttgaagatg gaaagaaatt tgattcctcc 120

cgggacagaa acaagccctt taagtttatg ctaggcaagc aggaggtgat ccgaggctgg 180

gaagaagggg ttgcccagat gagtgtgggt cagagagcca aactgactat atctccagat 240

tatgcctatg gtgccactgg gcacccaggc atcatcccac cacatgccac tctcgtcttc 300

gatgtggagc ttctaaaact ggaaggtggt tcaggtggtg gcgggagcgg tggctcgagc 360

agcggtggag cgatcggctc catgctgttc cgagtaacca tcaacagc 408

<210> 115

<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 115

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His
145 150 155 160

His His His His His
165

<210> 116

<211> 498
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 116
atggtcttca cactcgaaga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgctgttcc gagtaaccat caacagccat 480

catcaccatc accactaa 498

<210> 117
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 117

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val

65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 118
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 118
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ctcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatagac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 119
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 119

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1   5   10   15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20   25   30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
        35   40   45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    50   55   60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65   70   75   80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85   90   95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100   105   110

Arg Pro Tyr Glu Gly Ile Ala Val Leu Asp Gly Lys Lys Ile Thr Val
        115   120   125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130   135   140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145   150   155

<210> 120
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 120

gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa ccgc 474

<210> 121
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 121

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Arg
145 150 155

<210> 122
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 122
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatct tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 123
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 123

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu

20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp Leu His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 124
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 124
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccgtg tccttccgag taaccatcaa cagc 474

<210> 125
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 125

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Val Ser Phe Arg Val Thr Ile Asn Ser

145 150 155

<210> 126
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 126
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atcttgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcgtc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 127
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 127

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Val Ile Pro Tyr Glu
50 55 60

```
Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65                  70                  75                  80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85                  90                  95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115                 120                 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145                 150                 155
```

<210> 128
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 128
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggcc 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 129
<211> 158
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 129

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Ala Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 130
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 130
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccta atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgta 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tctcccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 131
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 131

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Leu Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly

100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Ser Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 132
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 132
gtcttcacac tcgaagattt cgttggggac tggtaccaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 133
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 133

Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn
1 5 10

<210> 134
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 134

Asn Val Thr Gly Tyr Arg Leu Phe Lys Lys Ile Ser Asn
1 5 10

<210> 135
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 135

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn
1 5 10

<210> 136
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 136

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn
1 5 10

<210> 137
<211> 10
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 137

Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn

1 5 10

<210> 138
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 138

Val Thr Gly Tyr Arg Leu Phe Glu Lys Ile Ser
1 5 10

<210> 139
<211> 10
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 139

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 140
<211> 6
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 140

Val Ser Gly Trp Arg Leu
1 5

<210> 141
<211> 7
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 141

Val Ser Gly Trp Arg Leu Phe
1 5

<210> 142
<211> 8
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 142

Val Ser Gly Trp Arg Leu Phe Lys
1 5

<210> 143
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 143

Val Ser Gly Trp Arg Leu Phe Lys Lys
1 5

<210> 144
<211> 10
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 144

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile
1 5 10

<210> 145
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 145

Val Ser Gly Trp Arg Leu Tyr Lys Lys Ile Ser
1 5 10

<210> 146
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 146

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Ala
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 147
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 147

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Thr Gly Tyr Arg
1 5 10 15

Leu Phe Glu Glu Ile Leu
20

<210> 148
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 148

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Ser Ser Trp Lys Arg

1 5 10 15

<210> 149
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 149

Val Ser Gly Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 150
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 150

Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 151
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 151

Val Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 152
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 152

Ser Ser Trp Lys Arg Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 153
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 153

Ser Ser Trp Lys Arg Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 154
<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 154

Ser Ser Trp Lys Arg Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn
1 5 10 15

Ser

<210> 155
<211> 18
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 155

Ser Ser Trp Lys Arg Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile
1 5 10 15

Asn Ser

<210> 156
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 156

Ser Ser Trp Lys Arg Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val
1 5 10 15

<210> 157
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 157

Ser Ser Trp Lys Arg Met Leu Phe Arg Val Thr Ile Asn Ser Val
1 5 10 15

<210> 158
<211> 18
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 158

Ser Ser Trp Lys Arg Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn
1 5 10 15

Ser Val

<210> 159
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 159

Ser Ser Trp Lys Arg Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile
1 5 10 15

Asn Ser Val

<210> 160
<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 160

Ser Ser Trp Lys Arg Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val
1 5 10 15

Ser

<210> 161
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 161

Ser Ser Trp Lys Arg Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser
1 5 10 15

<210> 162
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 162

Ser Ser Trp Lys Arg Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn

1 5 10 15

Ser Val Ser

<210> 163
<211> 20
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 163

Ser Ser Trp Lys Arg Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile
1 5 10 15

Asn Ser Val Ser
20

<210> 164
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 164

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn
1 5 10 15

<210> 165
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 165

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile
1 5 10

<210> 166

<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 166

Ser Ser Trp Lys Arg Ser Met Leu Phe Arg Val Thr Ile Asn
1 5 10

<210> 167
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 167

Ser Ser Trp Lys Arg Met Leu Phe Arg Val Thr Ile Asn
1 5 10

<210> 168
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 168

Ser Ser Trp Lys Arg Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn
1 5 10 15

<210> 169
<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 169

Ser Ser Trp Lys Arg Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile
1 5 10 15

Asn

<210> 170
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 170

Ser Ser Trp Lys Arg Ser Met Leu Phe Arg Val Thr Ile
1 5 10

<210> 171
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 171

Ser Ser Trp Lys Arg Met Leu Phe Arg Val Thr Ile
1 5 10

<210> 172
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 172

Ser Ser Trp Lys Arg Asp Gly Ser Met Leu Phe Arg Val Thr Ile
1 5 10 15

<210> 173
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 173

Ser Ser Trp Lys Arg Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile
1 5 10 15

<210> 174
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 174

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Val Phe Thr Leu
1 5 10 15

<210> 175
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 175

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Val Phe Thr
1 5 10

<210> 176
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 176

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Val Phe
1 5 10

<210> 177
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 177

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Val
1   5   10

<210> 178
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 178

Val Ser Gly Trp Arg Leu Cys Lys Lys Ile Ser
1   5   10

<210> 179
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 179

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser Met Leu Phe
1   5   10   15

Arg Val Thr Ile Asn Ser
            20

<210> 180
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 180

Ser Ser Trp Lys Arg Leu Phe Arg Val Thr Ile Asn Ser
1   5   10

<210> 181
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 181

Ser Ser Trp Lys Arg Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 182
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 182

Ser Ser Trp Lys Arg Arg Val Thr Ile Asn Ser
1 5 10

<210> 183
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 183

Ser Ser Trp Lys Arg Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr
1 5 10 15

Ile Asn Ser

<210> 184
<211> 20
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 184

Ser Ser Trp Lys Arg Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val
1 5 10 15

Thr Ile Asn Ser
20

<210> 185
<211> 21
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 185

Ser Ser Trp Lys Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe Arg
1 5 10 15

Val Thr Ile Asn Ser
20

<210> 186
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 186

Ser Ser Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10 15

<210> 187
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 187

Ser Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Trp Ser

1 5 10 15

<210> 188
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 188

Ser Trp Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 189
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 189

Ser Ser Arg Gly Ser Met Leu Phe Arg Val Thr Ile Trp Lys
1 5 10

<210> 190
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 190

Ser Ser Trp Lys Arg Gly Ser Met Leu Tyr Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 191
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 191

Ser Ser Trp Lys Arg Gly Ser Met Leu Trp Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 192
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 192

Ser Ser Trp Lys Arg Gly Ser Met Leu His Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 193
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 193

Ser Ser Trp Lys Arg Gly Ser Leu Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 194
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 194

Ser Ser Trp Lys Arg Gly Ser Lys Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 195
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 195

Ser Ser Trp Lys Arg Gly Ser Arg Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 196
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 196

Ser Ser Trp Lys Arg Gly Ser Phe Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 197
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 197

Ser Ser Trp Lys Arg Gly Ser Trp Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

<210> 198
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 198

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Ser Ile Asn Ser
1 5 10 15

<210> 199
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 199

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Gln Ile Asn Ser
1   5   10   15

<210> 200
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 200

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Asn Ile Asn Ser
1   5   10   15

<210> 201
<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 201

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1   5   10   15

Cys

<210> 202
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 202

Cys Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1   5   10

<210> 203
<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 203

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

Lys

<210> 204
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 204

Lys Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 205
<211> 23
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 205

Gly Ser Leu Leu Phe Arg Val Thr Ile Asn Gly Val Thr Gly Trp Arg
1 5 10 15

Leu Cys Glu Asn Ile Leu Ala
20

<210> 206
<211> 24
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 206

Gly Ser Leu Leu Phe Arg Val Thr Ile Asn Val Gly Val Thr Gly Trp
1 5 10 15

Arg Leu Cys Glu Arg Ile Leu Ala
20

<210> 207
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 207

Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 208
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 208

Asn Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 209
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 209

Gly Trp Arg Leu Phe Lys Lys Ile Ser

1 5

<210> 210
<211> 501
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 210
atggtgagcg gctggcggct gttcaagaag attagccacc atcaccatca ccatcatcac 60

ttcacactcg acgatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa 120

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg 180

atcatgagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc 240

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg 300

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac 360

ggggttacgc cgaacaagct gaactatttc ggacggccgt atgaaggcat cgccgtgttc 420

gacggcaaaa agatcactac cacagggacc ctgtggaacg gcaacaaaat tatcgacgag 480

cgcctgatca cccccgacta a 501

<210> 211
<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 211

Met Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser His His His His
1 5 10 15

His His His His Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln
20 25 30

Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser
35 40 45

Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile

50 55 60

Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile
65 70 75 80

Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val
85 90 95

Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu
100 105 110

Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn
115 120 125

Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys
130 135 140

Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu
145 150 155 160

Arg Leu Ile Thr Pro Asp
165

<210> 212
<211> 510
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 212
atgaaacatc accatcacca tcatgtgagc ggctggcggc tgttcaagaa gattagcggc 60

agctccggtt tcacactcga cgatttcgtt ggggactggg aacagacagc cgcctacaac 120

ctggaccaag tccttgaaca gggaggtgtg tccagtttgc tgcagaatct cgccgtgtcc 180

gtaactccga tcatgaggat tgtccggagc ggtgaaaatg ccctgaagat cgacatccat 240

gtcatcatcc cgtatgaagg tctgagcgcc gaccaaatgg cccagatcga agaggtgttt 300

aaggtggtgt accctgtgga tgatcatcac tttaaggtga tcctgcccta tggcacactg 360

gtaatcgacg gggttacgcc gaacaagctg aactatttcg gacggccgta tgaaggcatc 420

gccgtgttcg acggcaaaaa gatcactacc acagggaccc tgtggaacgg caacaaaatt 480

atcgacgagc gcctgatcac ccccgactaa 510

<210> 213
<211> 169
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 213

Met Lys His His His His His His Val Ser Gly Trp Arg Leu Phe Lys
1 5 10 15

Lys Ile Ser Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp
20 25 30

Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly
35 40 45

Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile
50 55 60

Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His
65 70 75 80

Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile
85 90 95

Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys
100 105 110

Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn
115 120 125

Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp
130 135 140

Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile
145 150 155 160

Ile Asp Glu Arg Leu Ile Thr Pro Asp
165

<210> 214
<211> 366
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 214
atggtggcca tcctctggca tgagatgtgg catgaaggcc tggaagaggc atctcgtttg 60

tactttgggg aaaggaacgt gaaaggcatg tttgaggtgc tggagccctt gcatgctatg 120

atggaacggg gcccccagac tctgaaggaa acatccttta atcaggccta tggtcgagat 180

ttaatggagg cccaagagtg gtgcaggaag tacatgaaat cagggaatgt caaggacctc 240

acccaagcct gggacctcta ttatcatgtg ttccgacgaa tcagtggtgg ttcaggtggt 300

ggcgggagcg gtggctcgag cagcggtgga gtgagcggct ggcggctgtt caagaagatt 360

agctaa 366

<210> 215
<211> 121
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 215

Met Val Ala Ile Leu Trp His Glu Met Trp His Glu Gly Leu Glu Glu
1 5 10 15

Ala Ser Arg Leu Tyr Phe Gly Glu Arg Asn Val Lys Gly Met Phe Glu
20 25 30

Val Leu Glu Pro Leu His Ala Met Met Glu Arg Gly Pro Gln Thr Leu
35 40 45

Lys Glu Thr Ser Phe Asn Gln Ala Tyr Gly Arg Asp Leu Met Glu Ala
50 55 60

```
Gln Glu Trp Cys Arg Lys Tyr Met Lys Ser Gly Asn Val Lys Asp Leu
65                  70                  75                  80


Thr Gln Ala Trp Asp Leu Tyr Tyr His Val Phe Arg Arg Ile Ser Gly
                85                  90                  95


Gly Ser Gly Gly Gly Gly Ser Gly Gly Ser Ser Ser Gly Gly Val Ser
            100                 105                 110


Gly Trp Arg Leu Phe Lys Lys Ile Ser
        115                 120


<210>  216
<211>  372
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  216
atggtggcca tcctctggca tgagatgtgg catgaaggcc tggaagaggc atctcgtttg       60

tactttgggg aaaggaacgt gaaaggcatg tttgaggtgc tggagccctt gcatgctatg      120

atggaacggg gcccccagac tctgaaggaa acatccttta atcaggccta tggtcgagat      180

ttaatggagg cccaagagtg gtgcaggaag tacatgaaat cagggaatgt caaggacctc      240

acccaagcct gggacctcta ttatcatgtg ttccgacgaa tcagtggtgg ttcaggtggt      300

ggcgggagcg gtggctcgag cagcggtgga gttagcgtta gcggctggcg cctgttcaag      360

aagatcagct aa                                                          372


<210>  217
<211>  123
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  217

Met Val Ala Ile Leu Trp His Glu Met Trp His Glu Gly Leu Glu Glu
1               5                   10                  15
```

Ala Ser Arg Leu Tyr Phe Gly Glu Arg Asn Val Lys Gly Met Phe Glu
20 25 30

Val Leu Glu Pro Leu His Ala Met Met Glu Arg Gly Pro Gln Thr Leu
35 40 45

Lys Glu Thr Ser Phe Asn Gln Ala Tyr Gly Arg Asp Leu Met Glu Ala
50 55 60

Gln Glu Trp Cys Arg Lys Tyr Met Lys Ser Gly Asn Val Lys Asp Leu
65 70 75 80

Thr Gln Ala Trp Asp Leu Tyr Tyr His Val Phe Arg Arg Ile Ser Gly
85 90 95

Gly Ser Gly Gly Gly Gly Ser Gly Gly Ser Ser Ser Gly Gly Val Ser
100 105 110

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
115 120

<210> 218
<211> 369
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 218
atgggagtgc aggtggaaac catctcccca ggagacgggc gcaccttccc caagcgcggc 60

cagacctgcg tggtgcacta caccgggatg cttgaagatg gaaagaaatt tgattcctcc 120

cgggacagaa acaagccctt taagtttatg ctaggcaagc aggaggtgat ccgaggctgg 180

gaagaagggg ttgcccagat gagtgtgggt cagagagcca aactgactat atctccagat 240

tatgcctatg gtgccactgg gcacccaggc atcatcccac cacatgccac tctcgtcttc 300

gatgtggagc ttctaaaact ggaaggtggt tcaggtggtg gcgggagcgg tggctcgagc 360

agcggtgga 369

<210> 219

<211> 123
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 219

Met Gly Val Gln Val Glu Thr Ile Ser Pro Gly Asp Gly Arg Thr Phe
1 5 10 15

Pro Lys Arg Gly Gln Thr Cys Val Val His Tyr Thr Gly Met Leu Glu
20 25 30

Asp Gly Lys Lys Phe Asp Ser Ser Arg Asp Arg Asn Lys Pro Phe Lys
35 40 45

Phe Met Leu Gly Lys Gln Glu Val Ile Arg Gly Trp Glu Glu Gly Val
50 55 60

Ala Gln Met Ser Val Gly Gln Arg Ala Lys Leu Thr Ile Ser Pro Asp
65 70 75 80

Tyr Ala Tyr Gly Ala Thr Gly His Pro Gly Ile Ile Pro Pro His Ala
85 90 95

Thr Leu Val Phe Asp Val Glu Leu Leu Lys Leu Glu Gly Gly Ser Gly
100 105 110

Gly Gly Gly Ser Gly Gly Ser Ser Ser Gly Gly
115 120

<210> 220
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 220

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val
1 5 10

<210> 221
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 221

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser
1   5   10

<210> 222
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 222

Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly
1   5   10

<210> 223
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 223

Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp
1   5   10

<210> 224
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 224

Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Lys

1 5 10

<210> 225
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 225

Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Arg
1 5 10

<210> 226
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 226

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly
1 5 10

<210> 227
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 227

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp
1 5 10 15

<210> 228
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 228

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Arg
1 5 10 15

<210> 229
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 229

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Lys
1 5 10 15

<210> 230
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 230

Gly Ser Met Leu Phe Arg Val Thr Ile Trp Lys
1 5 10

<210> 231
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 231

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 232
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 232

Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 233
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 233

Met Leu Phe Arg Val Thr Ile Asn Ser Trp Ser
1 5 10

<210> 234
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 234

Met Leu Phe Arg Val Thr Ile Trp Ser
1 5

<210> 235
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 235

Met Leu Phe Arg Val Thr Ile Trp Lys
1 5

<210> 236
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 236

Met Leu Phe Arg Val Lys Ile Asn Ser
1   5

<210> 237
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 237

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Trp Ser
1   5   10

<210> 238
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 238

Gly Ser Met Leu Phe Arg Val Lys Ile Asn Ser
1   5   10

<210> 239
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 239

Gly Ser Met Leu Phe Arg Val Thr Ile Trp Ser
1   5   10

<210> 240
<211> 9
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 240

Met Leu Phe Arg Val Asn Ile Asn Ser
1 5

<210> 241
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 241

Met Leu Phe Arg Val Trp Ile Asn Ser
1 5

<210> 242
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 242

Leu Leu Phe Arg Val Lys Ile Asn Ser
1 5

<210> 243
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 243

Phe Leu Phe Arg Val Thr Ile Asn Ser
1 5

<210> 244

<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 244

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

Val

<210> 245
<211> 18
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 245

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

Val Ser

<210> 246
<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 246

Ser Ser Trp Lys Arg Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser
1 5 10 15

Gly

<210> 247

<211> 18
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 247

Ser Ser Trp Lys Arg Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser
1 5 10 15

Gly Trp

<210> 248
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 248

Ser Ser Trp Lys Arg Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser
1 5 10 15

Gly Trp Arg

<210> 249
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 249

Ser Ser Trp Lys Arg Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser
1 5 10 15

Gly Trp Lys

<210> 250

<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 250

Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Lys
1 5 10

<210> 251
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 251

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

Val Ser Gly

<210> 252
<211> 20
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 252

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10 15

Val Ser Gly Trp
20

<210> 253
<211> 21
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 253

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1   5   10   15

Val Ser Gly Trp Arg
            20

<210> 254
<211> 21
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 254

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1   5   10   15

Val Ser Gly Trp Lys
            20

<210> 255
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 255

Ser Ser Trp Lys Arg Gly Ser Tyr Leu Phe Arg Val Thr Ile Asn Ser
1   5   10   15

<210> 256
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 256

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Lys Ile Asn Ser
1 5 10 15

<210> 257
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 257

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Arg Ile Asn Ser
1 5 10 15

<210> 258
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 258

Ser Ser Trp Lys Arg Gly Ser Met Leu Phe Arg Val Trp Ile Asn Ser
1 5 10 15

<210> 259
<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 259

Ser Ser Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Trp Ser Val
1 5 10 15

<210> 260
<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 260

Ser Ser Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Trp Ser Val
1 5 10 15

Ser

<210> 261
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 261

Ser Ser Trp Arg Gly Ser Met Leu Phe Arg Val Thr Ile Lys Ser
1 5 10 15

<210> 262
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 262

Lys Arg Ser Ser Gly Ser Met Leu Phe Arg Val Thr Ile Trp Ser
1 5 10 15

<210> 263
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 263

Ser Ser Lys Arg Met Leu Phe Arg Val Thr Ile Trp Ser
1 5 10

<210> 264

<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 264

Lys Arg Ser Ser Met Leu Phe Arg Val Thr Ile Trp Ser
1 5 10

<210> 265
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 265

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 266
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 266

Gly Ser Met Leu Phe Arg Lys Thr Ile Asn Ser Trp Lys
1 5 10

<210> 267
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 267

Gly Ser Met Leu Phe Arg Val Thr Lys Asn Ser Trp Lys
1 5 10

<210> 268
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 268

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 269
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 269

Gly Lys Met Leu Phe Arg Val Thr Ile Trp Lys
1 5 10

<210> 270
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 270

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 271
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 271

Gly Ser Met Lys Phe Arg Val Thr Ile Trp Lys

1 5 10

<210> 272
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 272

Gly Arg Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 273
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 273

Gly Arg Met Leu Phe Arg Val Thr Ile Trp Lys
1 5 10

<210> 274
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 274

Gly Ser Met Arg Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 275
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 275

Gly Ser Met Arg Phe Arg Val Thr Ile Trp Lys
1 5 10

<210> 276
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 276

Gly Asp Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 277
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 277

Gly Asp Met Leu Phe Arg Val Thr Ile Trp Lys
1 5 10

<210> 278
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 278

Gly Ser Met Asp Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 279
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 279

Gly Ser Met Asp Phe Arg Val Thr Ile Trp Lys
1 5 10

<210> 280
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 280

Gly Glu Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 281
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 281

Gly Ser Met Glu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 282
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 282

Gly Ser Met Glu Phe Arg Val Thr Ile Trp Lys
1 5 10

<210> 283
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 283

Gly Ser Met Leu Phe Arg Val Thr Ile Trp Lys Val Lys
1   5   10

<210> 284
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 284

Gly Ser Met Leu Phe Arg Val Thr Ile Trp Ser Val Lys
1   5   10

<210> 285
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 285

Gly Ser Met Leu Phe Arg Val Thr Ile Trp Ser Lys
1   5   10

<210> 286
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 286

Gly Ser Met Leu Phe Arg Val Thr Ile Trp Lys Trp Lys
1   5   10

<210> 287
<211> 12
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 287

Gly Ser Met Leu Phe Arg Val Thr Ile Trp Lys Lys
1 5 10

<210> 288
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 288

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 289
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 289

Met Leu Phe Arg Val Thr Ile Asn Ser
1 5

<210> 290
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 290

Met Leu Phe Val Thr Ile Asn Ser Val
1 5

<210> 291

<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 291

Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser
1   5   10

<210> 292
<211> 10
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 292

Gly Ser Met Leu Phe Arg Val Thr Ile Asn
1   5   10

<210> 293
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 293

Gly Ser Met Leu Phe Arg Val Thr Ile
1   5

<210> 294
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 294

Ser Met Leu Phe Arg Val Thr Ile Asn
1   5

<210> 295
<211> 8
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 295

Met Leu Phe Arg Val Thr Ile Asn
1   5

<210> 296
<211> 7
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 296

Met Leu Phe Arg Val Thr Ile
1   5

<210> 297
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 297

Ser Ser Lys Arg Gly Ser Met Leu Phe Arg Val Thr Ile Trp Ser
1   5   10   15

<210> 298
<211> 23
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 298

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Gly Val Ser Gly Trp

1 5 10 15

Ala Leu Phe Lys Lys Ile Ser
20

<210> 299
<211> 23
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 299

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Gly Val Ser Gly Trp
1 5 10 15

Arg Leu Phe Lys Lys Ile Ser
20

<210> 300
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 300

Val Ser Gly Trp Ala Leu Phe Lys Lys Ile Ser
1 5 10

<210> 301
<211> 25
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 301

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Val Ser
1 5 10 15

Gly Trp Arg Leu Phe Lys Lys Ile Ser

20 25

<210> 302
<211> 148
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 302

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp
145

<210> 303

<211> 444
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 303
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgac 444

<210> 304
<211> 148
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 304

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp
145

<210> 305
<211> 444
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 305
atggtcttca cactcgaaga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgac 444

<210> 306
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 306

Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 307
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 307

Asn Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 308
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 308

Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5

<210> 309
<211> 23
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 309

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Gly Val Ser Gly Trp
1 5 10 15

Arg Leu Phe Lys Lys Ile Ser
20

<210> 310

<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 310

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 311
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 311

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 312
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 312

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 313
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 313

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 314
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 314

Gly Ser Met Leu Phe Arg Val Thr Ile Lys Ser Trp Lys
1 5 10

<210> 315
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 315

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Arg Trp Lys
1 5 10

<210> 316
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 316

Gly Ser Met Leu Phe Arg Val Thr Ile Arg Ser Trp Lys
1 5 10

<210> 317
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 317

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Asp Trp Lys

1 5 10

<210> 318
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 318

Gly Ser Met Leu Phe Arg Val Thr Ile Asp Ser Trp Lys
1 5 10

<210> 319
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 319

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Glu Trp Lys
1 5 10

<210> 320
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 320

Gly Ser Met Leu Phe Arg Val Thr Ile Glu Ser Trp Lys
1 5 10

<210> 321
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 321

Gly Ser Met Arg Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 322
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 322

Gly Ser Met Asp Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 323
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 323

Gly Ser Met Glu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 324
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 324

Gly Ser Met Leu Phe Arg Arg Thr Ile Asn Ser Trp Lys
1 5 10

<210> 325
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 325

Gly Ser Met Leu Phe Arg Asp Thr Ile Asn Ser Trp Lys
1 5 10

<210> 326
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 326

Gly Ser Met Leu Phe Arg Glu Thr Ile Asn Ser Trp Lys
1 5 10

<210> 327
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 327

Gly Ser Met Leu Phe Arg Val Thr Asp Asn Ser Trp Lys
1 5 10

<210> 328
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 328

Gly Ser Met Leu Phe Arg Val Thr Glu Asn Ser Trp Lys
1 5 10

<210> 329
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 329

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys
1   5   10

<210> 330
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 330

Gly Ser Met Leu Phe Arg Lys Thr Ile Asn Ser Trp Lys
1   5   10

<210> 331
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 331

Gly Ser Met Leu Phe Arg Val Thr Lys Asn Ser Trp Lys
1   5   10

<210> 332
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 332

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser
1   5   10

<210> 333
<211> 11
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 333

Gly Ser Met Leu Phe Arg Val Ser Ile Asn Ser
1 5 10

<210> 334
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 334

Gly Ser Met Leu Phe Arg Val Asn Ile Asn Ser
1 5 10

<210> 335
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 335

Gly Ser Lys Leu Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 336
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 336

Gly Ser Arg Leu Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 337

<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 337

Gly Ser Met Trp Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 338
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 338

Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 339
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 339

Gly Ser Met Asn Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 340
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 340

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser
1 5 10

<210> 341
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 341

Gly Ser Met Leu Phe Arg Trp Thr Ile Asn Ser
1   5   10

<210> 342
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 342

Gly Ser Met Leu Phe Arg Ser Thr Ile Asn Ser
1   5   10

<210> 343
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 343

Gly Ser Met Leu Phe Arg Asn Thr Ile Asn Ser
1   5   10

<210> 344
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 344

Gly Ser Met Leu Phe Arg Lys Thr Ile Asn Ser

1 5 10

<210> 345
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 345

Gly Ser Met Leu Phe Arg Val Thr Trp Asn Ser
1 5 10

<210> 346
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 346

Gly Ser Met Leu Phe Arg Val Thr Ser Asn Ser
1 5 10

<210> 347
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 347

Gly Ser Met Leu Phe Arg Val Thr Asn Asn Ser
1 5 10

<210> 348
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 348

Gly Ser Met Leu Phe Arg Val Thr Lys Asn Ser
1 5 10

<210> 349
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 349

Gly Ser Met Leu Phe Arg Val Thr Ile Lys Ser
1 5 10

<210> 350
<211> 634
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 350

Met Lys His His His His His His Val Ser Lys Gly Glu Glu Leu Ile
1 5 10 15

Lys Glu Asn Met Arg Ser Lys Leu Tyr Leu Glu Gly Ser Val Asn Gly
20 25 30

His Gln Phe Lys Cys Thr His Glu Gly Glu Gly Lys Pro Tyr Glu Gly
35 40 45

Lys Gln Thr Asn Arg Ile Lys Val Val Glu Gly Gly Pro Leu Pro Phe
50 55 60

Ala Phe Asp Ile Leu Ala Thr His Phe Met Tyr Gly Ser Lys Val Phe
65 70 75 80

Ile Lys Tyr Pro Ala Asp Leu Pro Asp Tyr Phe Lys Gln Ser Phe Pro
85 90 95

Glu Gly Phe Thr Trp Glu Arg Val Met Val Phe Glu Asp Gly Gly Val

```
            100                 105                 110

Leu Thr Ala Thr Gln Asp Thr Ser Leu Gln Asp Gly Glu Leu Ile Tyr
        115                 120                 125

Asn Val Lys Val Arg Gly Val Asn Phe Pro Ala Asn Gly Pro Val Met
    130                 135                 140

Gln Lys Lys Thr Leu Gly Trp Glu Pro Ser Thr Glu Thr Met Tyr Pro
145                 150                 155                 160

Ala Asp Gly Gly Leu Glu Gly Arg Cys Asp Lys Ala Leu Lys Leu Val
                165                 170                 175

Gly Gly Gly His Leu His Val Asn Phe Lys Thr Thr Tyr Lys Ser Lys
            180                 185                 190

Lys Pro Val Lys Met Pro Gly Val His Tyr Val Asp Arg Arg Leu Glu
        195                 200                 205

Arg Ile Lys Glu Ala Asp Asn Glu Thr Tyr Val Glu Gln Tyr Glu His
    210                 215                 220

Ala Val Ala Arg Tyr Ser Asn Leu Gly Gly Gly Phe Thr Leu Glu Asp
225                 230                 235                 240

Phe Val Gly Asp Trp Arg Gln Thr Ala Gly Tyr Asn Leu Asp Gln Val
                245                 250                 255

Leu Glu Gln Gly Gly Val Ser Ser Leu Phe Gln Asn Leu Gly Val Ser
            260                 265                 270

Val Thr Pro Ile Gln Arg Ile Val Leu Ser Gly Glu Asn Gly Leu Lys
        275                 280                 285

Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Gly Asp Gln
    290                 295                 300

Met Gly Gln Ile Glu Lys Ile Phe Lys Val Val Tyr Pro Val Asp Asp
305                 310                 315                 320
```

```
His His Phe Lys Val Ile Leu His Tyr Gly Thr Leu Val Ile Asp Gly
               325                 330                 335

Val Thr Pro Asn Met Ile Asp Tyr Phe Gly Arg Pro Tyr Glu Gly Ile
            340                 345                 350

Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn
        355                 360                 365

Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Asn Pro Asp Gly Ser Leu
    370                 375                 380

Leu Phe Arg Val Thr Ile Asn Gly Val Thr Gly Trp Arg Leu Cys Glu
385                 390                 395                 400

Arg Ile Leu Ala Arg His Glu Leu Ile Lys Glu Asn Met Arg Ser Lys
               405                 410                 415

Leu Tyr Leu Glu Gly Ser Val Asn Gly His Gln Phe Lys Cys Thr His
            420                 425                 430

Glu Gly Glu Gly Lys Pro Tyr Glu Gly Lys Gln Thr Asn Arg Ile Lys
        435                 440                 445

Val Val Glu Gly Gly Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Thr
    450                 455                 460

His Phe Met Tyr Gly Ser Lys Val Phe Ile Lys Tyr Pro Ala Asp Leu
465                 470                 475                 480

Pro Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg
               485                 490                 495

Val Met Val Phe Glu Asp Gly Gly Val Leu Thr Ala Thr Gln Asp Thr
            500                 505                 510

Ser Leu Gln Asp Gly Glu Leu Ile Tyr Asn Val Lys Val Arg Gly Val
        515                 520                 525

Asn Phe Pro Ala Asn Gly Pro Val Met Gln Lys Lys Thr Leu Gly Trp
```

```
    530                 535                 540

Glu Pro Ser Thr Glu Thr Met Tyr Pro Ala Asp Gly Gly Leu Glu Gly
545                 550                 555                 560

Arg Cys Asp Lys Ala Leu Lys Leu Val Gly Gly Gly His Leu His Val
                565                 570                 575

Asn Phe Lys Thr Thr Tyr Lys Ser Lys Lys Pro Val Lys Met Pro Gly
            580                 585                 590

Val His Tyr Val Asp Arg Arg Leu Glu Arg Ile Lys Glu Ala Asp Asn
        595                 600                 605

Glu Thr Tyr Val Glu Gln Tyr Glu His Ala Val Ala Arg Tyr Ser Asn
    610                 615                 620

Leu Gly Gly Gly Met Asp Glu Leu Tyr Lys
625                 630
```

<210> 351
<211> 1902
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 351

```
atgaaacatc accatcacca tcatgtgagc aagggagaag aacttataaa agaaaacatg       60

cggtctaaac tgtacctcga gggctccgtc aatgggcacc agtttaagtg tacccacgag      120

ggtgagggaa agccctatga ggggaagcag acaaaccgca tcaaggtcgt cgaaggggga      180

cccctcccgt ttgcctttga tatcttggct actcacttta tgtacggaag caaagttttc      240

ataaagtatc ctgccgacct tcctgattat tttaaacagt catttcccga gggtttcaca      300

tgggaaaggg tcatggtgtt tgaggatgga ggcgtgctca ctgcaactca ggacacctca      360

ctgcaggacg gcgagctgat ctacaatgtg aaggtccggg gtgtaaactt ccctgccaac      420

gggcctgtaa tgcagaagaa gaccctggga tgggagccgt ccaccgaaac catgtaccct      480

gctgatggtg ggctggaggg ccgatgtgac aaggctctga agctcgttgg aggtggtcat      540
```

ttgcacgtaa atttcaagac aacttacaag agcaaaaaac ccgtaaaaat gcccggggtt 600

cattacgttg acagaaggct tgaacgcata aaggaagctg ataacgagac atacgtggag 660

cagtacgagc acgccgttgc ccggtactca aacctggggg gtggctttac actggaggat 720

tttgtgggag attggagaca gacagccggc tacaatctgg atcaggtgct ggaacaagga 780

ggagtgtctt ctctgtttca gaatctggga gtgagcgtga cacctatcca gaggatcgtg 840

ctgtctggcg agaatggact gaagatcgat attcacgtga tcatccccta cgaaggcctg 900

tctggagacc agatgggcca gattgagaag atcttcaaag tggtgtatcc tgtggacgat 960

caccacttca aggtgatcct gcactacggc accctggtga ttgatggagt gacacctaac 1020

atgatcgact acttcggaag accttacgag ggaatcgccg tgttcgacgg aaagaagatc 1080

accgtgacag gaacactgtg gaatggaaac aagatcatcg acgagcggct gatcaaccct 1140

gatggatctc tgctgttcag agtgaccatc aacggagtga caggatggag actgtgcgag 1200

agaattctgg ctagacatga gctaatcaag gaaaatatga gaagtaagct atacttagag 1260

gggtccgtca acggtcacca gtttaaatgc actcatgaag gtgaggggaa accttatgaa 1320

ggtaagcaga ctaatcgaat aaaagtggtc gagggcggtc ctctgccatt cgctttcgat 1380

attctggcca ctcactttat gtatgggtct aaggtcttta ttaaataccc cgctgatttg 1440

ccagactact ttaaacagtc cttccctgaa ggattcacat gggagcgggt gatggtgttc 1500

gaggatggag gcgttcttac tgcaactcag gatacttcct tgcaagacgg ggaactgatc 1560

tacaacgtta aggtccgcgg cgtcaatttc ccagccaatg gtccagtgat gcagaagaaa 1620

accttggggt gggagccctc aacggagaca atgtaccctg cggacggcgg gcttgagggt 1680

agatgtgata aggcattgaa actcgtcggg ggcggccacc ttcatgtgaa tttcaagact 1740

acatataaaa gtaaaaaacc agtcaagatg cctggagtgc actacgtgga tcgtaggttg 1800

gagaggataa aagaagccga caacgaaact tatgtagagc aatatgagca cgccgtggct 1860

cgttattcca acttgggcgg aggaatggat gaactgtaca ag 1902

<210> 352
<211> 622
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 352

```
Met Lys His His His His His His Val Ser Lys Gly Glu Glu Leu Ile
1               5                  10                  15

Lys Glu Asn Met Arg Ser Lys Leu Tyr Leu Glu Gly Ser Val Asn Gly
            20                  25                  30

His Gln Phe Lys Cys Thr His Glu Gly Glu Gly Lys Pro Tyr Glu Gly
        35                  40                  45

Lys Gln Thr Asn Arg Ile Lys Val Val Glu Gly Gly Pro Leu Pro Phe
    50                  55                  60

Ala Phe Asp Ile Leu Ala Thr His Phe Met Tyr Gly Ser Lys Val Phe
65                  70                  75                  80

Ile Lys Tyr Pro Ala Asp Leu Pro Asp Tyr Phe Lys Gln Ser Phe Pro
                85                  90                  95

Glu Gly Phe Thr Trp Glu Arg Val Met Val Phe Glu Asp Gly Gly Val
            100                 105                 110

Leu Thr Ala Thr Gln Asp Thr Ser Leu Gln Asp Gly Glu Leu Ile Tyr
        115                 120                 125

Asn Val Lys Val Arg Gly Val Asn Phe Pro Ala Asn Gly Pro Val Met
    130                 135                 140

Gln Lys Lys Thr Leu Gly Trp Glu Pro Ser Thr Glu Thr Met Tyr Pro
145                 150                 155                 160

Ala Asp Gly Gly Leu Glu Gly Arg Cys Asp Lys Ala Leu Lys Leu Val
                165                 170                 175

Gly Gly Gly His Leu His Val Asn Phe Lys Thr Thr Tyr Lys Ser Lys
            180                 185                 190

Lys Pro Val Lys Met Pro Gly Val His Tyr Val Asp Arg Arg Leu Glu
        195                 200                 205
```

```
Arg Ile Lys Glu Ala Asp Asn Glu Thr Tyr Val Glu Gln Tyr Glu His
    210                 215                 220

Ala Val Ala Arg Tyr Ser Asn Leu Gly Gly Gly Phe Thr Leu Glu Asp
225                 230                 235                 240

Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val
                245                 250                 255

Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser
            260                 265                 270

Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys
        275                 280                 285

Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln
    290                 295                 300

Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp
305                 310                 315                 320

His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly
                325                 330                 335

Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile
            340                 345                 350

Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn
        355                 360                 365

Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met
    370                 375                 380

Leu Phe Arg Val Thr Ile Asn Ser Arg His Glu Leu Ile Lys Glu Asn
385                 390                 395                 400

Met Arg Ser Lys Leu Tyr Leu Glu Gly Ser Val Asn Gly His Gln Phe
                405                 410                 415

Lys Cys Thr His Glu Gly Glu Gly Lys Pro Tyr Glu Gly Lys Gln Thr
```

```
            420                 425                 430

Asn Arg Ile Lys Val Val Glu Gly Gly Pro Leu Pro Phe Ala Phe Asp
        435                 440                 445

Ile Leu Ala Thr His Phe Met Tyr Gly Ser Lys Val Phe Ile Lys Tyr
    450                 455                 460

Pro Ala Asp Leu Pro Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly Phe
465                 470                 475                 480

Thr Trp Glu Arg Val Met Val Phe Glu Asp Gly Gly Val Leu Thr Ala
                485                 490                 495

Thr Gln Asp Thr Ser Leu Gln Asp Gly Glu Leu Ile Tyr Asn Val Lys
            500                 505                 510

Val Arg Gly Val Asn Phe Pro Ala Asn Gly Pro Val Met Gln Lys Lys
        515                 520                 525

Thr Leu Gly Trp Glu Pro Ser Thr Glu Thr Met Tyr Pro Ala Asp Gly
    530                 535                 540

Gly Leu Glu Gly Arg Cys Asp Lys Ala Leu Lys Leu Val Gly Gly Gly
545                 550                 555                 560

His Leu His Val Asn Phe Lys Thr Thr Tyr Lys Ser Lys Lys Pro Val
                565                 570                 575

Lys Met Pro Gly Val His Tyr Val Asp Arg Arg Leu Glu Arg Ile Lys
            580                 585                 590

Glu Ala Asp Asn Glu Thr Tyr Val Glu Gln Tyr Glu His Ala Val Ala
        595                 600                 605

Arg Tyr Ser Asn Leu Gly Gly Gly Met Asp Glu Leu Tyr Lys
    610                 615                 620
```

<210> 353
<211> 1866
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic

<400> 353
atgaaacatc accatcacca tcatgtgagc aagggagaag aacttataaa agaaaacatg 60

cggtctaaac tgtacctcga gggctccgtc aatgggcacc agtttaagtg tacccacgag 120

ggtgagggaa agccctatga ggggaagcag acaaaccgca tcaaggtcgt cgaaggggga 180

cccctcccgt ttgcctttga tatcttggct actcacttta tgtacggaag caaagttttc 240

ataaagtatc ctgccgacct tcctgattat tttaaacagt catttcccga gggtttcaca 300

tgggaaaggg tcatggtgtt tgaggatgga ggcgtgctca ctgcaactca ggacacctca 360

ctgcaggacg gcgagctgat ctacaatgtg aaggtccggg gtgtaaactt ccctgccaac 420

gggcctgtaa tgcagaagaa gaccctggga tgggagccgt ccaccgaaac catgtaccct 480

gctgatggtg ggctggaggg ccgatgtgac aaggctctga agctcgttgg aggtggtcat 540

ttgcacgtaa atttcaagac aacttacaag agcaaaaaac ccgtaaaaat gcccggggtt 600

cattacgttg acagaaggct tgaacgcata aaggaagctg ataacgagac atacgtggag 660

cagtacgagc acgccgttgc ccggtactca aacctggggg gtggcttcac actcgaagat 720

ttcgttgggg actgggaaca gacagccgcc tacaacctgg accaagtcct tgaacaggga 780

ggtgtgtcca gtttgctgca gaatctcgcc gtgtccgtaa ctccgatcca aaggattgtc 840

cggagcggtg aaaatgccct gaagatcgac atccatgtca tcatcccgta tgaaggtctg 900

agcgccgacc aaatggccca gatcgaagag gtgtttaagg tggtgtaccc tgtggatgat 960

catcacttta aggtgatcct gccctatggc acactggtaa tcgacggggt tacgccgaac 1020

atgctgaact atttcggacg gccgtatgaa ggcatcgccg tgttcgacgg caaaaagatc 1080

actgtaacag ggaccctgtg gaacggcaac aaaattatcg acgagcgcct gatcaccccc 1140

gacggctcca tgctgttccg agtaaccatc aacagcagac atgagctaat caaggaaaat 1200

atgagaagta agctatactt agaggggtcc gtcaacggtc accagtttaa atgcactcat 1260

gaaggtgagg ggaaacctta tgaaggtaag cagactaatc gaataaaagt ggtcgagggc 1320

ggtcctctgc cattcgcttt cgatattctg gccactcact ttatgtatgg gtctaaggtc 1380

tttattaaat accccgctga tttgccagac tactttaaac agtccttccc tgaaggattc 1440

acatgggagc gggtgatggt gttcgaggat ggaggcgttc ttactgcaac tcaggatact 1500

tccttgcaag acggggaact gatctacaac gttaaggtcc gcggcgtcaa tttcccagcc 1560

aatggtccag tgatgcagaa gaaaaccttg gggtgggagc cctcaacgga gacaatgtac 1620

cctgcggacg gcgggcttga gggtagatgt gataaggcat tgaaactcgt cgggggcggc 1680

caccttcatg tgaatttcaa gactacatat aaaagtaaaa aaccagtcaa gatgcctgga 1740

gtgcactacg tggatcgtag gttggagagg ataaaagaag ccgacaacga aacttatgta 1800

gagcaatatg agcacgccgt ggctcgttat tccaacttgg gcggaggaat ggatgaactg 1860

tacaag 1866

<210> 354
<211> 611
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 354

Met Lys His His His His His His Val Ser Lys Gly Glu Glu Leu Ile
1 5 10 15

Lys Glu Asn Met Arg Ser Lys Leu Tyr Leu Glu Gly Ser Val Asn Gly
20 25 30

His Gln Phe Lys Cys Thr His Glu Gly Glu Gly Lys Pro Tyr Glu Gly
35 40 45

Lys Gln Thr Asn Arg Ile Lys Val Val Glu Gly Gly Pro Leu Pro Phe
50 55 60

Ala Phe Asp Ile Leu Ala Thr His Phe Met Tyr Gly Ser Lys Val Phe
65 70 75 80

Ile Lys Tyr Pro Ala Asp Leu Pro Asp Tyr Phe Lys Gln Ser Phe Pro
85 90 95

Glu Gly Phe Thr Trp Glu Arg Val Met Val Phe Glu Asp Gly Gly Val
100 105 110

```
Leu Thr Ala Thr Gln Asp Thr Ser Leu Gln Asp Gly Glu Leu Ile Tyr
        115                 120                 125

Asn Val Lys Val Arg Gly Val Asn Phe Pro Ala Asn Gly Pro Val Met
    130                 135                 140

Gln Lys Lys Thr Leu Gly Trp Glu Pro Ser Thr Glu Thr Met Tyr Pro
145                 150                 155                 160

Ala Asp Gly Gly Leu Glu Gly Arg Cys Asp Lys Ala Leu Lys Leu Val
                165                 170                 175

Gly Gly Gly His Leu His Val Asn Phe Lys Thr Thr Tyr Lys Ser Lys
            180                 185                 190

Lys Pro Val Lys Met Pro Gly Val His Tyr Val Asp Arg Arg Leu Glu
        195                 200                 205

Arg Ile Lys Glu Ala Asp Asn Glu Thr Tyr Val Glu Gln Tyr Glu His
    210                 215                 220

Ala Val Ala Arg Tyr Ser Asn Leu Gly Gly Gly Phe Thr Leu Asp Asp
225                 230                 235                 240

Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val
                245                 250                 255

Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser
            260                 265                 270

Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys
        275                 280                 285

Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln
    290                 295                 300

Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp
305                 310                 315                 320

His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly
```

```
                325                 330                 335

Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile
            340                 345                 350

Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn
        355                 360                 365

Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Arg His Glu
    370                 375                 380

Leu Ile Lys Glu Asn Met Arg Ser Lys Leu Tyr Leu Glu Gly Ser Val
385                 390                 395                 400

Asn Gly His Gln Phe Lys Cys Thr His Glu Gly Glu Gly Lys Pro Tyr
                405                 410                 415

Glu Gly Lys Gln Thr Asn Arg Ile Lys Val Val Glu Gly Gly Pro Leu
            420                 425                 430

Pro Phe Ala Phe Asp Ile Leu Ala Thr His Phe Met Tyr Gly Ser Lys
        435                 440                 445

Val Phe Ile Lys Tyr Pro Ala Asp Leu Pro Asp Tyr Phe Lys Gln Ser
    450                 455                 460

Phe Pro Glu Gly Phe Thr Trp Glu Arg Val Met Val Phe Glu Asp Gly
465                 470                 475                 480

Gly Val Leu Thr Ala Thr Gln Asp Thr Ser Leu Gln Asp Gly Glu Leu
                485                 490                 495

Ile Tyr Asn Val Lys Val Arg Gly Val Asn Phe Pro Ala Asn Gly Pro
            500                 505                 510

Val Met Gln Lys Lys Thr Leu Gly Trp Glu Pro Ser Thr Glu Thr Met
        515                 520                 525

Tyr Pro Ala Asp Gly Gly Leu Glu Gly Arg Cys Asp Lys Ala Leu Lys
    530                 535                 540
```

```
Leu Val Gly Gly Gly His Leu His Val Asn Phe Lys Thr Thr Tyr Lys
545                 550                 555                 560


Ser Lys Lys Pro Val Lys Met Pro Gly Val His Tyr Val Asp Arg Arg
                565                 570                 575


Leu Glu Arg Ile Lys Glu Ala Asp Asn Glu Thr Tyr Val Glu Gln Tyr
            580                 585                 590


Glu His Ala Val Ala Arg Tyr Ser Asn Leu Gly Gly Gly Met Asp Glu
        595                 600                 605


Leu Tyr Lys
    610
```

<210> 355
<211> 1833
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 355

```
atgaaacatc accatcacca tcatgtgagc aagggagaag aacttataaa agaaaacatg       60

cggtctaaac tgtacctcga gggctccgtc aatgggcacc agtttaagtg tacccacgag      120

ggtgagggaa agccctatga ggggaagcag acaaaccgca tcaaggtcgt cgaaggggga      180

cccctcccgt ttgcctttga tatcttggct actcacttta tgtacggaag caaagttttc      240

ataaagtatc ctgccgacct tcctgattat tttaaacagt catttcccga gggtttcaca      300

tgggaaaggg tcatggtgtt tgaggatgga ggcgtgctca ctgcaactca ggacacctca      360

ctgcaggacg gcgagctgat ctacaatgtg aaggtccggg gtgtaaactt ccctgccaac      420

gggcctgtaa tgcagaagaa gaccctggga tgggagccgt ccaccgaaac catgtaccct      480

gctgatggtg ggctggaggg ccgatgtgac aaggctctga agctcgttgg aggtggtcat      540

ttgcacgtaa atttcaagac aacttacaag agcaaaaaac ccgtaaaaat gcccggggtt      600

cattacgttg acagaaggct tgaacgcata aaggaagctg ataacgagac atacgtggag      660

cagtacgagc acgccgttgc ccggtactca aacctggggg gtggcttcac actcgacgat      720
```

ttcgttgggg actgggaaca gacagccgcc tacaacctgg accaagtcct tgaacaggga 780

ggtgtgtcca gtttgctgca gaatctcgcc gtgtccgtaa ctccgatcat gaggattgtc 840

cggagcggtg aaaatgccct gaagatcgac atccatgtca tcatcccgta tgaaggtctg 900

agcgccgacc aaatggccca gatcgaagag gtgtttaagg tggtgtaccc tgtggatgat 960

catcacttta aggtgatcct gccctatggc acactggtaa tcgacggggt tacgccgaac 1020

aagctgaact atttcggacg gccgtatgaa ggcatcgccg tgttcgacgg caaaaagatc 1080

actaccacag ggaccctgtg gaacggcaac aaaattatcg acgagcgcct gatcaccccc 1140

gacagacatg agctaatcaa ggaaaatatg agaagtaagc tatacttaga ggggtccgtc 1200

aacggtcacc agtttaaatg cactcatgaa ggtgagggga aaccttatga aggtaagcag 1260

actaatcgaa taaaagtggt cgagggcggt cctctgccat tcgctttcga tattctggcc 1320

actcacttta tgtatgggtc taaggtcttt attaaatacc ccgctgattt gccagactac 1380

tttaaacagt ccttccctga aggattcaca tgggagcggg tgatggtgtt cgaggatgga 1440

ggcgttctta ctgcaactca ggatacttcc ttgcaagacg gggaactgat ctacaacgtt 1500

aaggtccgcg gcgtcaattt cccagccaat ggtccagtga tgcagaagaa aaccttgggg 1560

tgggagccct caacggagac aatgtaccct gcggacggcg ggcttgaggg tagatgtgat 1620

aaggcattga aactcgtcgg gggcggccac cttcatgtga atttcaagac tacatataaa 1680

agtaaaaaac cagtcaagat gcctggagtg cactacgtgg atcgtaggtt ggagaggata 1740

aaagaagccg acaacgaaac ttatgtagag caatatgagc acgccgtggc tcgttattcc 1800

aacttgggcg gaggaatgga tgaactgtac aag 1833

<210> 356
<211> 401
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 356

Met Lys His His His His His His Phe Thr Leu Glu Asp Phe Val Gly
1 5 10 15

Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln

```
            20                  25                  30

Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro
        35                  40                  45

Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile
    50                  55                  60

His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln
65                  70                  75                  80

Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe
                85                  90                  95

Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro
            100                 105                 110

Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe
        115                 120                 125

Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn Lys
    130                 135                 140

Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe Arg
145                 150                 155                 160

Val Thr Ile Asn Ser Gly Gly Ser Gly Gly Ser Ser Gly Glu Leu Ile
                165                 170                 175

Lys Glu Asn Met Arg Ser Lys Leu Tyr Leu Glu Gly Ser Val Asn Gly
            180                 185                 190

His Gln Phe Lys Cys Thr His Glu Gly Glu Gly Lys Pro Tyr Glu Gly
        195                 200                 205

Lys Gln Thr Asn Arg Ile Lys Val Val Glu Gly Gly Pro Leu Pro Phe
    210                 215                 220

Ala Phe Asp Ile Leu Ala Thr His Phe Met Tyr Gly Ser Lys Val Phe
225                 230                 235                 240
```

```
Ile Lys Tyr Pro Ala Asp Leu Pro Asp Tyr Phe Lys Gln Ser Phe Pro
                245                 250                 255

Glu Gly Phe Thr Trp Glu Arg Val Met Val Phe Glu Asp Gly Gly Val
            260                 265                 270

Leu Thr Ala Thr Gln Asp Thr Ser Leu Gln Asp Gly Glu Leu Ile Tyr
        275                 280                 285

Asn Val Lys Val Arg Gly Val Asn Phe Pro Ala Asn Gly Pro Val Met
    290                 295                 300

Gln Lys Lys Thr Leu Gly Trp Glu Pro Ser Thr Glu Thr Met Tyr Pro
305                 310                 315                 320

Ala Asp Gly Gly Leu Glu Gly Arg Cys Asp Lys Ala Leu Lys Leu Val
                325                 330                 335

Gly Gly Gly His Leu His Val Asn Phe Lys Thr Thr Tyr Lys Ser Lys
            340                 345                 350

Lys Pro Val Lys Met Pro Gly Val His Tyr Val Asp Arg Arg Leu Glu
        355                 360                 365

Arg Ile Lys Glu Ala Asp Asn Glu Thr Tyr Val Glu Gln Tyr Glu His
    370                 375                 380

Ala Val Ala Arg Tyr Ser Asn Leu Gly Gly Gly Met Asp Glu Leu Tyr
385                 390                 395                 400

Lys
```

<210> 357
<211> 1203
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 357

atgaaacatc accatcacca tcatttcaca ctcgaagatt tcgttgggga ctgggaacag 60

acagccgcct acaacctgga ccaagtcctt gaacagggag gtgtgtccag tttgctgcag 120

aatctcgccg tgtccgtaac tccgatccaa aggattgtcc ggagcggtga aaatgccctg 180

aagatcgaca tccatgtcat catcccgtat gaaggtctga gcgccgacca aatggcccag 240

atcgaagagg tgtttaaggt ggtgtaccct gtggatgatc atcactttaa ggtgatcctg 300

ccctatggca cactggtaat cgacggggtt acgccgaaca tgctgaacta tttcggacgg 360

ccgtatgaag gcatcgccgt gttcgacggc aaaaagatca ctgtaacagg gaccctgtgg 420

aacggcaaca aaattatcga cgagcgcctg atcacccccg acggctccat gctgttccga 480

gtaaccatca acagcggagg ctcaggtgga tcctcaggtg agctaatcaa ggaaaatatg 540

agaagtaagc tatacttaga ggggtccgtc aacggtcacc agtttaaatg cactcatgaa 600

ggtgagggga aaccttatga aggtaagcag actaatcgaa taaaagtggt cgagggcggt 660

cctctgccat tcgctttcga tattctggcc actcacttta tgtatgggtc taaggtcttt 720

attaaatacc ccgctgattt gccagactac tttaaacagt ccttccctga aggattcaca 780

tgggagcggg tgatggtgtt cgaggatgga ggcgttctta ctgcaactca ggatacttcc 840

ttgcaagacg gggaactgat ctacaacgtt aaggtccgcg gcgtcaattt cccagccaat 900

ggtccagtga tgcagaagaa aaccttgggg tgggagccct caacggagac aatgtaccct 960

gcggacggcg ggcttgaggg tagatgtgat aaggcattga aactcgtcgg gggcggccac 1020

cttcatgtga atttcaagac tacatataaa agtaaaaaac cagtcaagat gcctggagtg 1080

cactacgtgg atcgtaggtt ggagaggata aaagaagccg acaacgaaac ttatgtagag 1140

caatatgagc acgccgtggc tcgttattcc aacttgggcg gaggaatgga tgaactgtac 1200

aag 1203

<210> 358
<211> 395
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 358

Met Lys His His His His His His Phe Thr Leu Glu Asp Phe Val Gly

```
1               5                   10                  15

Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln
            20                  25                  30

Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro
        35                  40                  45

Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile
    50                  55                  60

His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln
65                  70                  75                  80

Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe
                85                  90                  95

Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro
            100                 105                 110

Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe
        115                 120                 125

Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn Lys
    130                 135                 140

Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe Arg
145                 150                 155                 160

Val Thr Ile Asn Ser Arg His Glu Leu Ile Lys Glu Asn Met Arg Ser
                165                 170                 175

Lys Leu Tyr Leu Glu Gly Ser Val Asn Gly His Gln Phe Lys Cys Thr
            180                 185                 190

His Glu Gly Glu Gly Lys Pro Tyr Glu Gly Lys Gln Thr Asn Arg Ile
        195                 200                 205

Lys Val Val Glu Gly Gly Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala
    210                 215                 220
```

```
Thr His Phe Met Tyr Gly Ser Lys Val Phe Ile Lys Tyr Pro Ala Asp
225                 230                 235                 240


Leu Pro Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu
                245                 250                 255


Arg Val Met Val Phe Glu Asp Gly Gly Val Leu Thr Ala Thr Gln Asp
            260                 265                 270


Thr Ser Leu Gln Asp Gly Glu Leu Ile Tyr Asn Val Lys Val Arg Gly
        275                 280                 285


Val Asn Phe Pro Ala Asn Gly Pro Val Met Gln Lys Lys Thr Leu Gly
    290                 295                 300


Trp Glu Pro Ser Thr Glu Thr Met Tyr Pro Ala Asp Gly Gly Leu Glu
305                 310                 315                 320


Gly Arg Cys Asp Lys Ala Leu Lys Leu Val Gly Gly Gly His Leu His
                325                 330                 335


Val Asn Phe Lys Thr Thr Tyr Lys Ser Lys Lys Pro Val Lys Met Pro
            340                 345                 350


Gly Val His Tyr Val Asp Arg Arg Leu Glu Arg Ile Lys Glu Ala Asp
        355                 360                 365


Asn Glu Thr Tyr Val Glu Gln Tyr Glu His Ala Val Ala Arg Tyr Ser
    370                 375                 380


Asn Leu Gly Gly Gly Met Asp Glu Leu Tyr Lys
385                 390                 395


<210>  359
<211>  1185
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  359
```

atgaaacatc accatcacca tcatttcaca ctcgaagatt tcgttgggga ctgggaacag 60

acagccgcct acaacctgga ccaagtcctt gaacagggag gtgtgtccag tttgctgcag 120

aatctcgccg tgtccgtaac tccgatccaa aggattgtcc ggagcggtga aaatgccctg 180

aagatcgaca tccatgtcat catcccgtat gaaggtctga gcgccgacca aatggcccag 240

atcgaagagg tgtttaaggt ggtgtaccct gtggatgatc atcactttaa ggtgatcctg 300

ccctatggca cactggtaat cgacggggtt acgccgaaca tgctgaacta tttcggacgg 360

ccgtatgaag gcatcgccgt gttcgacggc aaaaagatca ctgtaacagg gaccctgtgg 420

aacggcaaca aaattatcga cgagcgcctg atcacccccg acggctccat gctgttccga 480

gtaaccatca acagcagaca tgagctaatc aaggaaaata tgagaagtaa gctatactta 540

gaggggtccg tcaacggtca ccagtttaaa tgcactcatg aaggtgaggg gaaaccttat 600

gaaggtaagc agactaatcg aataaaagtg gtcgagggcg gtcctctgcc attcgctttc 660

gatattctgg ccactcactt tatgtatggg tctaaggtct ttattaaata ccccgctgat 720

ttgccagact actttaaaca gtccttccct gaaggattca catgggagcg ggtgatggtg 780

ttcgaggatg gaggcgttct tactgcaact caggatactt ccttgcaaga cggggaactg 840

atctacaacg ttaaggtccg cggcgtcaat ttcccagcca atggtccagt gatgcagaag 900

aaaaccttgg ggtgggagcc ctcaacggag acaatgtacc ctgcggacgg cgggcttgag 960

ggtagatgtg ataaggcatt gaaactcgtc gggggcggcc accttcatgt gaatttcaag 1020

actacatata aaagtaaaaa accagtcaag atgcctggag tgcactacgt ggatcgtagg 1080

ttggagagga taaaagaagc cgacaacgaa acttatgtag agcaatatga gcacgccgtg 1140

gctcgttatt ccaacttggg cggaggaatg gatgaactgt acaag 1185

<210> 360
<211> 390
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 360

Met Lys His His His His His His Phe Thr Leu Asp Asp Phe Val Gly
1 5 10 15

```
Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln
            20                  25                  30

Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro
        35                  40                  45

Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile
    50                  55                  60

His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln
65                  70                  75                  80

Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe
                85                  90                  95

Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro
            100                 105                 110

Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe
        115                 120                 125

Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys
    130                 135                 140

Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Gly Ser Gly Gly Ser
145                 150                 155                 160

Ser Gly Glu Leu Ile Lys Glu Asn Met Arg Ser Lys Leu Tyr Leu Glu
                165                 170                 175

Gly Ser Val Asn Gly His Gln Phe Lys Cys Thr His Glu Gly Glu Gly
            180                 185                 190

Lys Pro Tyr Glu Gly Lys Gln Thr Asn Arg Ile Lys Val Val Glu Gly
        195                 200                 205

Gly Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Thr His Phe Met Tyr
    210                 215                 220

Gly Ser Lys Val Phe Ile Lys Tyr Pro Ala Asp Leu Pro Asp Tyr Phe
```

225 230 235 240

Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Val Met Val Phe
245 250 255

Glu Asp Gly Gly Val Leu Thr Ala Thr Gln Asp Thr Ser Leu Gln Asp
260 265 270

Gly Glu Leu Ile Tyr Asn Val Lys Val Arg Gly Val Asn Phe Pro Ala
275 280 285

Asn Gly Pro Val Met Gln Lys Lys Thr Leu Gly Trp Glu Pro Ser Thr
290 295 300

Glu Thr Met Tyr Pro Ala Asp Gly Gly Leu Glu Gly Arg Cys Asp Lys
305 310 315 320

Ala Leu Lys Leu Val Gly Gly Gly His Leu His Val Asn Phe Lys Thr
325 330 335

Thr Tyr Lys Ser Lys Lys Pro Val Lys Met Pro Gly Val His Tyr Val
340 345 350

Asp Arg Arg Leu Glu Arg Ile Lys Glu Ala Asp Asn Glu Thr Tyr Val
355 360 365

Glu Gln Tyr Glu His Ala Val Ala Arg Tyr Ser Asn Leu Gly Gly Gly
370 375 380

Met Asp Glu Leu Tyr Lys
385 390

<210> 361
<211> 1170
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 361
atgaaacatc accatcacca tcatttcaca ctcgacgatt tcgttgggga ctgggaacag 60

acagccgcct acaacctgga ccaagtcctt gaacagggag gtgtgtccag tttgctgcag 120

aatctcgccg tgtccgtaac tccgatcatg aggattgtcc ggagcggtga aaatgccctg 180

aagatcgaca tccatgtcat catcccgtat gaaggtctga gcgccgacca aatggcccag 240

atcgaagagg tgtttaaggt ggtgtaccct gtggatgatc atcactttaa ggtgatcctg 300

ccctatggca cactggtaat cgacggggtt acgccgaaca agctgaacta tttcggacgg 360

ccgtatgaag gcatcgccgt gttcgacggc aaaaagatca ctaccacagg gaccctgtgg 420

aacggcaaca aaattatcga cgagcgcctg atcacccccg acggaggctc aggtggatcc 480

tcaggtgagc taatcaagga aaatatgaga agtaagctat acttagaggg gtccgtcaac 540

ggtcaccagt ttaaatgcac tcatgaaggt gaggggaaac cttatgaagg taagcagact 600

aatcgaataa aagtggtcga gggcggtcct ctgccattcg ctttcgatat tctggccact 660

cactttatgt atgggtctaa ggtctttatt aaataccccg ctgatttgcc agactacttt 720

aaacagtcct tccctgaagg attcacatgg gagcgggtga tggtgttcga ggatggaggc 780

gttcttactg caactcagga tacttccttg caagacgggg aactgatcta caacgttaag 840

gtccgcggcg tcaatttccc agccaatggt ccagtgatgc agaagaaaac cttggggtgg 900

gagccctcaa cggagacaat gtaccctgcg gacggcgggc ttgagggtag atgtgataag 960

gcattgaaac tcgtcggggg cggccacctt catgtgaatt tcaagactac atataaaagt 1020

aaaaaaccag tcaagatgcc tggagtgcac tacgtggatc gtaggttgga gaggataaaa 1080

gaagccgaca acgaaactta tgtagagcaa tatgagcacg ccgtggctcg ttattccaac 1140

ttgggcggag gaatggatga actgtacaag 1170

<210> 362
<211> 384
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 362

Met Lys His His His His His His Phe Thr Leu Asp Asp Phe Val Gly
1 5 10 15

Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln

```
            20                  25                  30

Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro
        35                  40                  45

Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile
    50                  55                  60

His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln
65                  70                  75                  80

Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe
                85                  90                  95

Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro
            100                 105                 110

Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe
        115                 120                 125

Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys
    130                 135                 140

Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Arg His Glu Leu Ile Lys
145                 150                 155                 160

Glu Asn Met Arg Ser Lys Leu Tyr Leu Glu Gly Ser Val Asn Gly His
                165                 170                 175

Gln Phe Lys Cys Thr His Glu Gly Glu Gly Lys Pro Tyr Glu Gly Lys
            180                 185                 190

Gln Thr Asn Arg Ile Lys Val Val Glu Gly Gly Pro Leu Pro Phe Ala
        195                 200                 205

Phe Asp Ile Leu Ala Thr His Phe Met Tyr Gly Ser Lys Val Phe Ile
    210                 215                 220

Lys Tyr Pro Ala Asp Leu Pro Asp Tyr Phe Lys Gln Ser Phe Pro Glu
225                 230                 235                 240
```

Gly Phe Thr Trp Glu Arg Val Met Val Phe Glu Asp Gly Gly Val Leu
245 250 255

Thr Ala Thr Gln Asp Thr Ser Leu Gln Asp Gly Glu Leu Ile Tyr Asn
260 265 270

Val Lys Val Arg Gly Val Asn Phe Pro Ala Asn Gly Pro Val Met Gln
275 280 285

Lys Lys Thr Leu Gly Trp Glu Pro Ser Thr Glu Thr Met Tyr Pro Ala
290 295 300

Asp Gly Gly Leu Glu Gly Arg Cys Asp Lys Ala Leu Lys Leu Val Gly
305 310 315 320

Gly Gly His Leu His Val Asn Phe Lys Thr Thr Tyr Lys Ser Lys Lys
325 330 335

Pro Val Lys Met Pro Gly Val His Tyr Val Asp Arg Arg Leu Glu Arg
340 345 350

Ile Lys Glu Ala Asp Asn Glu Thr Tyr Val Glu Gln Tyr Glu His Ala
355 360 365

Val Ala Arg Tyr Ser Asn Leu Gly Gly Gly Met Asp Glu Leu Tyr Lys
370 375 380

<210> 363
<211> 1152
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 363
atgaaacatc accatcacca tcatttcaca ctcgacgatt tcgttgggga ctgggaacag 60

acagccgcct acaacctgga ccaagtcctt gaacagggag gtgtgtccag tttgctgcag 120

aatctcgccg tgtccgtaac tccgatcatg aggattgtcc ggagcggtga aaatgccctg 180

aagatcgaca tccatgtcat catcccgtat gaaggtctga gcgccgacca aatggcccag 240

atcgaagagg tgtttaaggt ggtgtaccct gtggatgatc atcactttaa ggtgatcctg 300

ccctatggca cactggtaat cgacggggtt acgccgaaca agctgaacta tttcggacgg 360

ccgtatgaag gcatcgccgt gttcgacggc aaaaagatca ctaccacagg gaccctgtgg 420

aacggcaaca aaattatcga cgagcgcctg atcacccccg acagacatga gctaatcaag 480

gaaaatatga gaagtaagct atacttagag gggtccgtca acggtcacca gtttaaatgc 540

actcatgaag gtgaggggaa accttatgaa ggtaagcaga ctaatcgaat aaaagtggtc 600

gagggcggtc ctctgccatt cgctttcgat attctggcca ctcactttat gtatgggtct 660

aaggtcttta ttaaataccc cgctgatttg ccagactact ttaaacagtc cttccctgaa 720

ggattcacat gggagcgggt gatggtgttc gaggatggag gcgttcttac tgcaactcag 780

gatacttcct tgcaagacgg ggaactgatc tacaacgtta aggtccgcgg cgtcaatttc 840

ccagccaatg gtccagtgat gcagaagaaa accttggggt gggagccctc aacggagaca 900

atgtaccctg cggacggcgg gcttgagggt agatgtgata aggcattgaa actcgtcggg 960

ggcggccacc ttcatgtgaa tttcaagact acatataaaa gtaaaaaacc agtcaagatg 1020

cctggagtgc actacgtgga tcgtaggttg gagaggataa aagaagccga caacgaaact 1080

tatgtagagc aatatgagca cgccgtggct cgttattcca acttgggcgg aggaatggat 1140

gaactgtaca ag 1152

<210> 364
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 364

Met Lys His His His His His His Val Phe Thr Leu Glu Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Leu Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 365
<211> 465
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 365
atgaaacatc accatcacca tcatgtcttc acactcgaag atttcgttgg ggactgggaa 60

cagaccgccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc ctaaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acatgctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactgtaac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgac 465

<210> 366
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 366

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
1 5 10 15

<210> 367
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 367

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Arg
1 5 10

<210> 368
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 368

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys
1 5 10

<210> 369
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 369

Gly Lys Met Leu Phe Arg Val Thr Ile Trp Lys Val Ser Val Ser Gly

1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 370
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 370

Gly Lys Met Leu Phe Arg Val Thr Ile Trp Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 371
<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 371

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys Val Ser Val
1 5 10 15

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
20 25

<210> 372
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 372

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys Val Ser Gly

1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 373
<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 373

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys Asn Val Thr
1 5 10 15

Gly Tyr Arg Leu Phe Lys Lys Ile Ser Asn
20 25

<210> 374
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 374

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys Val Thr Gly
1 5 10 15

Tyr Arg Leu Phe Glu Lys Ile Ser
20

<210> 375
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 375

Gly Ser Met Lys Phe Arg Val Thr Ile Trp Lys Val Ser Val Ser Gly

1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 376
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 376

Gly Ser Met Lys Phe Arg Val Thr Ile Trp Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 377
<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 377

Gly Arg Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys Val Ser Val
1 5 10 15

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
20 25

<210> 378
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 378

Gly Arg Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys Val Ser Gly

1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 379
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 379

Gly Arg Met Leu Phe Arg Val Thr Ile Trp Lys Val Ser Val Ser Gly
1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 380
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 380

Gly Arg Met Leu Phe Arg Val Thr Ile Trp Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 381
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 381

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Val Ser Gly

1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 382
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 382

Gly Ser Met Leu Phe Lys Val Thr Ile Asn Ser Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 383
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 383

Gly Ser Met Leu Phe Gln Val Thr Ile Asn Ser Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 384
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 384

Gly Ser Met Leu Phe Glu Val Thr Ile Asn Ser Val Ser Gly Trp Arg

1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 385
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 385

Gly Ser Met Leu Phe Asn Val Thr Ile Asn Ser Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 386
<211> 21
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 386

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
1 5 10 15

Thr Thr Gly Thr Leu
20

<210> 387
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 387

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys Val Thr Gly

1 5 10 15

Tyr Arg Leu Phe Glu Lys Glu Ser
20

<210> 388
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 388

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys Val Glu Gly
1 5 10 15

Tyr Arg Leu Phe Glu Lys Ile Ser
20

<210> 389
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 389

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
1 5 10 15

Asp Glu Arg Leu Ile Thr Pro Asp
20

<210> 390
<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 390

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly

1 5 10 15

Ser Met Leu Phe Arg Val Thr Ile Asn Ser
20 25

<210> 391
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 391

Gly Lys Met Leu Phe Arg Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 392
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 392

Gly Lys Met Leu Phe Arg Val Thr Ile Gly Lys Trp Lys
1 5 10

<210> 393
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 393

Gly Lys Met Leu Phe Arg Val Thr Ile Gly Arg Trp Lys
1 5 10

<210> 394
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 394

Gly Lys Met Leu Phe Arg Val Thr Ile Gly Asn Trp Lys
1   5   10

<210> 395
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 395

Gly Lys Met Leu Phe Arg Val Thr Ile Gln Asn Trp Lys
1   5   10

<210> 396
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 396

Gly Lys Met Leu Phe Arg Val Thr Ile Asp Lys Trp Lys
1   5   10

<210> 397
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 397

Gly Lys Met Leu Phe Arg Val Thr Ile Glu Lys Trp Lys
1   5   10

<210> 398
<211> 13
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 398

Gly Lys Met Leu Phe Arg Val Thr Ile Glu Arg Trp Lys
1 5 10

<210> 399
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 399

Gly Lys Met Leu Phe Arg Val Thr Ile Asp Arg Trp Lys
1 5 10

<210> 400
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 400

Asp Lys Met Leu Phe Arg Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 401
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 401

Asp Lys Met Leu Phe Arg Val Thr Ile Gly Lys Trp Lys
1 5 10

<210> 402

<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 402

Asp Lys Met Leu Phe Arg Val Thr Ile Gly Arg Trp Lys
1 5 10

<210> 403
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 403

Asp Lys Met Leu Phe Arg Val Thr Ile Gly Asn Trp Lys
1 5 10

<210> 404
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 404

Asp Lys Met Leu Phe Arg Val Thr Ile Gln Asn Trp Lys
1 5 10

<210> 405
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 405

Asp Lys Met Leu Phe Arg Val Thr Ile Asp Lys Trp Lys
1 5 10

<210> 406
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 406

Asp Lys Met Leu Phe Arg Val Thr Ile Glu Lys Trp Lys
1 5 10

<210> 407
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 407

Asp Lys Met Leu Phe Arg Val Thr Ile Glu Arg Trp Lys
1 5 10

<210> 408
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 408

Asp Lys Met Leu Phe Arg Val Thr Ile Asp Arg Trp Lys
1 5 10

<210> 409
<211> 35
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 409

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val

1 5 10 15

Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile
20 25 30

Thr Pro Asp
35

<210> 410
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 410

Glu Lys Met Leu Phe Arg Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 411
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 411

Glu Lys Met Leu Phe Arg Val Thr Ile Gly Lys Trp Lys
1 5 10

<210> 412
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 412

Glu Lys Met Leu Phe Arg Val Thr Ile Gly Arg Trp Lys
1 5 10

<210> 413

<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 413

Asp Lys Met Leu Phe Thr Val Thr Ile Gln Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 414
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 414

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 415
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 415

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Trp Lys Val Ser Gly
1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 416

<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 416

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Tyr Lys Val Ser Gly
1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 417
<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 417

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Tyr Lys Val Ser Val
1 5 10 15

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
20 25

<210> 418
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 418

Lys Lys Met Leu Phe Arg Val Thr Ile Gln Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 419

<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 419

Lys Lys Met Leu Phe Arg Val Thr Ile Gln Lys Trp Lys Val Ser Val
1 5 10 15

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
20 25

<210> 420
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 420

Lys Lys Met Leu Phe Arg Val Thr Ile Gln Lys Trp Lys Val Ser Gly
1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 421
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 421

Asp Lys Leu Leu Phe Thr Val Thr Ile Gly Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 422

<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 422

Asp Lys Leu Leu Phe Thr Val Thr Ile Gly Lys Tyr Lys Val Ser Gly
1 5 10 15

Trp Arg Leu Phe Lys Lys Ile Ser
20

<210> 423
<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 423

Asp Lys Leu Leu Phe Thr Val Thr Ile Gly Lys Tyr Lys Val Ser Val
1 5 10 15

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
20 25

<210> 424
<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 424

Asp Lys Leu Leu Phe Thr Val Thr Ile Gly Lys Trp Lys Val Ser Val
1 5 10 15

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
20 25

<210> 425

<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 425

Asp Lys Leu Leu Phe Thr Val Thr Ile Gln Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 426
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 426

Lys Lys Met Leu Phe Thr Val Thr Ile Gln Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 427
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 427

Lys Lys Leu Leu Phe Arg Val Thr Ile Gln Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 428

<211> 21
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 428

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Val Ser Gly Trp Arg
1   5   10   15

Leu Phe Lys Lys Ile
            20

<210> 429
<211> 25
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 429

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Tyr Lys Val Ser Val
1   5   10   15

Ser Gly Trp Arg Leu Phe Lys Lys Ile
            20   25

<210> 430
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 430

Asp Arg Leu Leu Phe Thr Val Thr Ile Glu Arg Val Ser Gly Trp Arg
1   5   10   15

Leu Phe Lys Lys Ile Ser
            20

<210> 431

<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 431

Glu Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 432
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 432

Lys Lys Leu Leu Phe Thr Val Thr Ile Gly Lys Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 433
<211> 24
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 433

Gly Ser Met Arg Phe Arg Val Thr Ile Asn Ser Trp Arg Val Thr Gly
1 5 10 15

Tyr Arg Leu Phe Glu Arg Glu Ser
20

<210> 434

<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 434

Gly Ser Met Lys Phe Arg Val Thr Ile Asn Ser Val Thr Gly Tyr Arg
1 5 10 15

Leu Phe Glu Lys Glu Ser
20

<210> 435
<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 435

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
1 5 10 15

Asp

<210> 436
<211> 29
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 436

Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile
1 5 10 15

Asn Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
20 25

<210> 437

<211> 58
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 437

Gly Arg Pro Tyr Glu Gly Ile Ala Val Asp Phe Gly Lys Lys Ile Thr
1 5 10 15

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
20 25 30

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val
35 40 45

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
50 55

<210> 438
<211> 68
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 438

Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly
1 5 10 15

Ile Ala Val Asp Phe Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp
20 25 30

Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser
35 40 45

Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Arg Leu Phe
50 55 60

Lys Lys Ile Ser
65

<210> 439
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 439

Glu Lys Met Leu Phe Arg Val Thr Ile Gly Asn Trp Lys
1 5 10

<210> 440
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 440

Glu Lys Met Leu Phe Arg Val Thr Ile Gln Asn Trp Lys
1 5 10

<210> 441
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 441

Glu Lys Met Leu Phe Arg Val Thr Ile Asp Lys Trp Lys
1 5 10

<210> 442
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 442

Glu Lys Met Leu Phe Arg Val Thr Ile Glu Lys Trp Lys

1 5 10

<210> 443
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 443

Glu Lys Met Leu Phe Arg Val Thr Ile Glu Arg Trp Lys
1 5 10

<210> 444
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 444

Glu Lys Met Leu Phe Arg Val Thr Ile Asp Arg Trp Lys
1 5 10

<210> 445
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 445

Lys Lys Met Leu Phe Arg Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 446
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 446

Lys Lys Met Leu Phe Arg Val Thr Ile Gly Lys Trp Lys
1 5 10

<210> 447
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 447

Lys Lys Met Leu Phe Arg Val Thr Ile Gly Arg Trp Lys
1 5 10

<210> 448
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 448

Lys Lys Met Leu Phe Arg Val Thr Ile Gly Asn Trp Lys
1 5 10

<210> 449
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 449

Lys Lys Met Leu Phe Arg Val Thr Ile Gln Asn Trp Lys
1 5 10

<210> 450
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 450

Lys Lys Met Leu Phe Arg Val Thr Ile Asp Lys Trp Lys
1 5 10

<210> 451
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 451

Lys Lys Met Leu Phe Arg Val Thr Ile Glu Lys Trp Lys
1 5 10

<210> 452
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 452

Lys Lys Met Leu Phe Arg Val Thr Ile Glu Arg Trp Lys
1 5 10

<210> 453
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 453

Lys Lys Met Leu Phe Arg Val Thr Ile Asp Arg Trp Lys
1 5 10

<210> 454
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 454

Arg Lys Met Leu Phe Arg Val Thr Ile Gln Lys Trp Lys
1   5   10

<210> 455
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 455

Arg Lys Met Leu Phe Arg Val Thr Ile Gly Lys Trp Lys
1   5   10

<210> 456
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 456

Arg Lys Met Leu Phe Arg Val Thr Ile Gly Arg Trp Lys
1   5   10

<210> 457
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 457

Arg Lys Met Leu Phe Arg Val Thr Ile Gly Asn Trp Lys
1   5   10

<210> 458
<211> 13
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 458

Arg Lys Met Leu Phe Arg Val Thr Ile Gln Asn Trp Lys
1 5 10

<210> 459
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 459

Arg Lys Met Leu Phe Arg Val Thr Ile Asp Lys Trp Lys
1 5 10

<210> 460
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 460

Arg Lys Met Leu Phe Arg Val Thr Ile Glu Lys Trp Lys
1 5 10

<210> 461
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 461

Arg Lys Met Leu Phe Arg Val Thr Ile Glu Arg Trp Lys
1 5 10

<210> 462

<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 462

Arg Lys Met Leu Phe Arg Val Thr Ile Asp Arg Trp Lys
1 5 10

<210> 463
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 463

Glu Gln Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 464
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 464

Ser Arg Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 465
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 465

Gly Glu Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 466
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 466

Gly Lys Met Lys Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 467
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 467

Gly Lys Met Leu Phe Arg Val Lys Ile Asn Ser Trp Lys
1 5 10

<210> 468
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 468

Gly Lys Met Leu Phe Arg Val Arg Ile Asn Ser Trp Lys
1 5 10

<210> 469
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 469

Gly Lys Met Leu Phe Arg Val Asp Ile Asn Ser Trp Lys

1 5 10

<210> 470
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 470

Gly Lys Met Leu Phe Arg Val Thr Ile Asp Ser Trp Lys
1 5 10

<210> 471
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 471

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 472
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 472

Glu Lys Met Leu Phe Lys Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 473
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 473

Glu Lys Met Leu Phe Thr Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 474
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 474

Glu Lys Met Leu Phe Lys Val Thr Ile Asp Lys Trp Lys
1 5 10

<210> 475
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 475

Glu Lys Met Leu Phe Thr Val Thr Ile Asp Lys Trp Lys
1 5 10

<210> 476
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 476

Glu Lys Met Leu Phe Lys Val Thr Ile Gly Arg Trp Lys
1 5 10

<210> 477
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 477

Asp Lys Met Leu Phe Lys Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 478
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 478

Asp Lys Met Leu Phe Thr Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 479
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 479

Asp Lys Met Leu Phe Lys Val Thr Ile Asp Lys Trp Lys
1 5 10

<210> 480
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 480

Asp Lys Met Leu Phe Thr Val Thr Ile Asp Lys Trp Lys
1 5 10

<210> 481
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 481

Gly Lys Met Leu Phe Lys Val Thr Ile Glu Lys Trp Lys
1   5   10

<210> 482
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 482

Gly Lys Met Leu Phe Thr Val Thr Ile Glu Lys Trp Lys
1   5   10

<210> 483
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 483

Asp Lys Met Leu Phe Lys Val Thr Ile Gly Lys Trp Lys
1   5   10

<210> 484
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 484

Asp Lys Met Leu Phe Thr Val Thr Ile Gly Lys Trp Lys
1   5   10

<210> 485
<211> 13
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 485

Asp Lys Met Leu Phe Lys Val Thr Ile Gly Asn Trp Lys
1 5 10

<210> 486
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 486

Asp Lys Met Leu Phe Lys Val Thr Ile Gln Asn Trp Lys
1 5 10

<210> 487
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 487

Gly Lys Met Leu Phe Lys Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 488
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 488

Gly Lys Met Leu Phe Thr Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 489

<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 489

Asp Lys Met Leu Phe Lys Val Thr Ile Glu Lys Trp Lys
1 5 10

<210> 490
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 490

Asp Lys Met Leu Phe Thr Val Thr Ile Glu Lys Trp Lys
1 5 10

<210> 491
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 491

Asp Lys Leu Leu Phe Lys Val Thr Ile Gly Lys Trp Lys
1 5 10

<210> 492
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 492

Asp Lys Met Leu Phe Thr Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 493
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 493

Asp Lys Leu Leu Phe Thr Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 494
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 494

Asp Lys Leu Leu Phe Thr Val Thr Ile Gln Lys Tyr Lys
1 5 10

<210> 495
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 495

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Trp Lys
1 5 10

<210> 496
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 496

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Tyr Lys

1 5 10

<210> 497
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 497

Asp Lys Leu Leu Phe Thr Val Thr Ile Gly Lys Trp Lys
1 5 10

<210> 498
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 498

Asp Lys Leu Leu Phe Thr Val Thr Ile Gly Lys Tyr Lys
1 5 10

<210> 499
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 499

Asp Lys Leu Leu Phe Thr Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 500
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 500

Asp Lys Leu Leu Phe Thr Val Thr Ile Asn Lys Tyr Lys
1   5   10

<210> 501
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 501

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser
1   5   10

<210> 502
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 502

Asp Lys Met Leu Phe Thr Val Thr Ile Gln Lys
1   5   10

<210> 503
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 503

Asp Lys Met Leu Phe Lys Val Thr Ile Gln Lys
1   5   10

<210> 504
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 504

Asp Lys Leu Leu Phe Thr Val Thr Ile Gly Lys
1 5 10

<210> 505
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 505

Asp Lys Met Leu Phe Thr Val Thr Ile Gly Lys
1 5 10

<210> 506
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 506

Asp Lys Met Leu Phe Thr Val Thr Ile Glu Lys
1 5 10

<210> 507
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 507

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys
1 5 10

<210> 508
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 508

Asp Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1   5   10

<210> 509
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 509

Glu Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1   5   10

<210> 510
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 510

Arg Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1   5   10

<210> 511
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 511

Lys Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1   5   10

<210> 512
<211> 13
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 512

His Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 513
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 513

Gly Leu Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 514
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 514

Gly Gln Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 515
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 515

Gly Thr Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 516

<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 516

Gly Lys Leu Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 517
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 517

Gly Lys Met Leu Phe Lys Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 518
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 518

Gly Lys Met Leu Phe Arg Val Thr Ile Gln Ser Trp Lys
1 5 10

<210> 519
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 519

Gly Lys Met Leu Phe Arg Val Thr Ile Asp Ser Trp Lys
1 5 10

<210> 520
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 520

Gly Lys Met Leu Phe Arg Val Thr Ile Gly Ser Trp Lys
1 5 10

<210> 521
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 521

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Thr Trp Lys
1 5 10

<210> 522
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 522

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Asn Trp Lys
1 5 10

<210> 523
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 523

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Gln Trp Lys

1 5 10

<210> 524
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 524

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Pro Trp Lys
1 5 10

<210> 525
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 525

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 526
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 526

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Gln
1 5 10

<210> 527
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 527

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Asn
1 5 10

<210> 528
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 528

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Thr
1 5 10

<210> 529
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 529

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp His
1 5 10

<210> 530
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 530

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Pro
1 5 10

<210> 531
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 531

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Arg
1 5 10

<210> 532
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 532

Gly Lys Met Lys Phe Arg Val Thr Ile Asp Ser Trp Lys
1 5 10

<210> 533
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 533

Gly Lys Met Leu Phe Arg Val Glu Ile Asn Ser Trp Lys
1 5 10

<210> 534
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 534

Gly Lys Met Leu Phe Arg Val Gln Ile Asn Ser Trp Lys
1 5 10

<210> 535
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 535

Gly Lys Met Lys Phe Arg Val Lys Ile Asn Ser Trp Lys
1   5   10

<210> 536
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 536

Gly Lys Met Lys Phe Arg Val Arg Ile Asn Ser Trp Lys
1   5   10

<210> 537
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 537

Gly Lys Met Lys Phe Arg Val Glu Ile Asn Ser Trp Lys
1   5   10

<210> 538
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 538

Gly Lys Met Lys Phe Arg Val Asp Ile Asn Ser Trp Lys
1   5   10

<210> 539
<211> 13
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 539

Gly Lys Met Lys Phe Arg Val Gln Ile Asn Ser Trp Lys
1 5 10

<210> 540
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 540

Gly Lys Met Lys Phe Arg Val Asn Ile Asn Ser Trp Lys
1 5 10

<210> 541
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 541

Gly Lys Met Lys Phe Arg Val Ser Ile Asn Ser Trp Lys
1 5 10

<210> 542
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 542

Gly Lys Met Leu Phe Arg Val Asn Ile Asn Ser Trp Lys
1 5 10

<210> 543

<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 543

Gly Lys Met Leu Phe Arg Val Ser Ile Asn Ser Trp Lys
1 5 10

<210> 544
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 544

Gly Lys Met Leu Phe Arg Val Trp Ile Asn Ser Trp Lys
1 5 10

<210> 545
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 545

Gly Lys Met Ser Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 546
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 546

Gly Lys Met Trp Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 547
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 547

Gly Lys Met Asn Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 548
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 548

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Tyr Lys
1 5 10

<210> 549
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 549

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser Tyr Lys
1 5 10

<210> 550
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 550

Gly Lys Met Leu Phe Arg Val Thr Ile Lys Ser Trp Lys

1 5 10

<210> 551
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 551

Gly Lys Met Leu Phe Arg Val Thr Ile Glu Ser Trp Lys
1 5 10

<210> 552
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 552

Gly Lys Met Lys Phe Arg Val Thr Ile Gln Ser Trp Lys
1 5 10

<210> 553
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 553

Gly Lys Met Lys Phe Arg Val Thr Ile Glu Ser Trp Lys
1 5 10

<210> 554
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 554

Gly Lys Met Lys Phe Arg Val Thr Ile Lys Ser Trp Lys
1 5 10

<210> 555
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 555

Gly Lys Met Lys Phe Arg Val Thr Ile Arg Ser Trp Lys
1 5 10

<210> 556
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 556

Arg Leu Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 557
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 557

Arg Gln Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 558
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 558

Lys Leu Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 559
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 559

Lys Gln Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 560
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 560

Glu Leu Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 561
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 561

Asp Leu Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 562
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 562

Asp Gln Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1   5   10

<210> 563
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 563

Asp Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1   5   10

<210> 564
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 564

Glu Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1   5   10

<210> 565
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 565

Arg Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1   5   10

<210> 566
<211> 13
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 566

Lys Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 567
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 567

Gly Lys Met Leu Phe Arg Val Thr Ile Gly Ser Trp Lys
1 5 10

<210> 568
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 568

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 569
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 569

Gly Lys Met Leu Phe Arg Val Thr Ile Ser Lys Trp Lys
1 5 10

<210> 570

<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 570

Gly Lys Met Leu Phe Arg Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 571
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 571

Gly Lys Met Leu Phe Arg Val Thr Ile Thr Lys Trp Lys
1 5 10

<210> 572
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 572

Gly Lys Met Leu Phe Arg Val Thr Ile Lys Lys Trp Lys
1 5 10

<210> 573
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 573

Gly Lys Met Leu Phe Lys Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 574
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 574

Arg Leu Met Leu Phe Arg Val Thr Ile Gly Lys Trp Lys
1 5 10

<210> 575
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 575

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Arg Trp Lys
1 5 10

<210> 576
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 576

Glu Lys Met Leu Phe Thr Val Thr Ile Gly Lys Trp Lys
1 5 10

<210> 577
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 577

Glu Lys Leu Leu Phe Thr Val Thr Ile Gly Lys Trp Lys

1 5 10

<210> 578
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 578

Glu Lys Met Leu Phe Thr Val Thr Ile Gly Arg Trp Lys
1 5 10

<210> 579
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 579

Glu Lys Met Leu Phe Thr Val Thr Ile Glu Lys Trp Lys
1 5 10

<210> 580
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 580

Asp Lys Met Leu Phe Arg Val Thr Ile Glu Ser Trp Lys
1 5 10

<210> 581
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 581

Glu Lys Leu Leu Phe Arg Val Thr Ile Gly Lys Tyr Lys
1   5   10

<210> 582
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 582

Asp Lys Leu Leu Phe Lys Val Thr Ile Gln Lys Trp Lys
1   5   10

<210> 583
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 583

Asp Lys Leu Leu Phe Lys Val Thr Ile Gln Lys Tyr Lys
1   5   10

<210> 584
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 584

Asp Lys Leu Leu Phe Lys Val Thr Ile Gly Lys Tyr Lys
1   5   10

<210> 585
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 585

Asp Lys Leu Leu Phe Lys Val Thr Ile Glu Lys Trp Lys
1 5 10

<210> 586
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 586

Asp Lys Leu Leu Phe Lys Val Thr Ile Glu Lys Tyr Lys
1 5 10

<210> 587
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 587

Lys Lys Leu Leu Phe Arg Val Thr Ile Gln Lys Trp Lys
1 5 10

<210> 588
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 588

Asp Arg Met Leu Phe Arg Val Thr Ile Gln Arg Trp Arg
1 5 10

<210> 589
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 589

Glu Arg Met Leu Phe Arg Val Thr Ile Gly Arg Trp Arg
1   5   10

<210> 590
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 590

Gly Arg Met Leu Phe Arg Val Thr Ile Asn Arg Trp Arg
1   5   10

<210> 591
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 591

Asp Arg Met Leu Phe Arg Val Thr Ile Glu Arg Trp Arg
1   5   10

<210> 592
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 592

Asp Lys Met Leu Phe Lys Val Thr Ile Gln Lys Tyr Lys
1   5   10

<210> 593
<211> 13
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 593

Asp Lys Met Leu Phe Arg Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 594
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 594

Asp Lys Met Leu Phe Lys Val Thr Ile Glu Lys Tyr Lys
1 5 10

<210> 595
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 595

Asp Lys Met Leu Phe Lys Val Thr Ile Asn Lys Trp Lys
1 5 10

<210> 596
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 596

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Tyr Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu

20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 597
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 597
gtcttcacac tcgaagattt cgttggggac tggcggcaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 598
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 598

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Arg Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser

145 150 155

<210> 599
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 599
gtcttcacac tcgaagattt cgttggggac tggctgcaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 600
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 600

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Leu Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

```
Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65                  70                  75                  80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85                  90                  95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115                 120                 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145                 150                 155
```

<210> 601
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 601
gtcttcacac tcgaagattt cgttggggac tggaaccaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccagg taaccatcaa cagc 474

<210> 602
<211> 158
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 602

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Asn Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 603
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 603
gtcttcacac tcgaagattt cgttggggac tggaaccaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcga cagc 474

<210> 604
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 604

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Asn Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly

100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asp Ser
145 150 155

<210> 605
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 605
gtcttcacac tcgaagattt cgttggggac tggcagcaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 606
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 606

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Gln Gln Thr Ala Ala
1 5 10 15

```
Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20                  25                  30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
        35                  40                  45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    50                  55                  60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65                  70                  75                  80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85                  90                  95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115                 120                 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145                 150                 155
```

<210> 607
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 607
gtcttcacac tcgaagattt cgttggggac tggatgcaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 608
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 608

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Met Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 609
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 609
gtcttcacac tcgaagattt cgttggggac tggcaccaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 610
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 610

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp His Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu

50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn
145 150 155

<210> 611
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 611
gtcttcacac tcgaagattt cgttggggac tggttccaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 612

<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 612

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Phe Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 613
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 613
gtcttcacac tcgaagattt cgttggggac tggtggcaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 614
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 614

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Trp Gln Thr Ala Ala
1   5   10   15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20   25   30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
        35   40   45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    50   55   60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65   70   75   80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85   90   95

```
Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110


Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115                 120                 125


Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140


Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145                 150                 155


<210>  615
<211>  474
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  615
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac       60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact      120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc      180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tagaagaggt gtttaaggtg      240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc      300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtacgaagg catcgccgtg      360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgccaa aattatcgac      420

gagcgcctga tcacccccga cggctccatg tccttccagg taaccatcaa cagc            474


<210>  616
<211>  158
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  616

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
```

1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 617
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 617
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tagaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtacgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgccaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccagg taaccatcga cagc 474

<210> 618
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 618

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile

130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asp Ser
145 150 155

<210> 619
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 619
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tagaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc actatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtacgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgccaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 620
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 620

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

```
Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    50                  55                  60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65                  70                  75                  80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly
                85                  90                  95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115                 120                 125

Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145                 150                 155
```

<210> 621
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 621

```
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac       60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact      120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc      180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg      240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc actatggcac actggtaatc      300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg      360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac      420

gagcgcctga tcacccccga cggctccatg tccttccagg taaccatcaa cagc            474
```

<210> 622
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 622

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 623
<211> 474
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic

<400> 623
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc actatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccagg taaccatcga cagc 474

<210> 624
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 624

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly

85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asp Ser
145 150 155

<210> 625
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 625
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tagaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc actatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtacgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgccaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccagg taaccatcaa cagc 474

<210> 626
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 626

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 627
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 627
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tagaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc actatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtacgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcggcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccagg taaccatcaa cagc 474

<210> 628
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 628

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Gly Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 629
<211> 470
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 629
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc 470

<210> 630
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 630

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly

```
        35                  40                  45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
    50                  55                  60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65                  70                  75                  80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
                85                  90                  95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
            100                 105                 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
        115                 120                 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
    130                 135                 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145                 150                 155


<210>  631
<211>  471
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  631
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa       60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct      120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata      180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt      240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac      300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc      360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag      420
```

agattaataa cgccagacgg aagtatgagt ttccaggtta cgatagactc g 471

<210> 632
<211> 156
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 632

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asp
145 150 155

<210> 633

<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 633
ttcactttgg acgatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattaccat atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 634
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 634

Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

```
Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr
                85                  90                  95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
            100                 105                 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
        115                 120                 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
    130                 135                 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145                 150                 155
```

<210> 635
<211> 472
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 635
ttcactttgg acgatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

ataatgcgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc gt 472

<210> 636
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 636

Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 637
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 637
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcattcaa cttagaccca 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ctatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccacgtta cgataaactc g 471

<210> 638
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 638

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Phe
1 5 10 15

Asn Leu Asp Pro Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr

115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe His Val Thr Ile Asn Ser
145 150 155

<210> 639
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 639
ttcactttgg aggatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacaggtt aaactatttc ggtagacctt ttgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 640
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 640

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Arg Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Phe Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 641
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 641
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtg tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tacaggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 642
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 642

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Val Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 643
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 643
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta tagtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 644
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 644

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val

65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 645
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 645
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggctgttag tgttacgcca 120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacaggtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 646
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 646

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1   5   10   15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
            20   25   30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
        35   40   45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
    50   55   60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65   70   75   80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
                85   90   95

Leu Val Ile Asp Gly Val Thr Pro Asn Arg Leu Asn Tyr Phe Gly Arg
            100   105   110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
        115   120   125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
    130   135   140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145   150   155

<210> 647
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 647

ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacacgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 648
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 648

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Thr Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 649
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 649
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctg ttacagaatt tggcagttag tgttacgcct 120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacaagtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccgagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 650
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 650

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln

20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Glu Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 651
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 651
ttcactttgg aagatttcgt gggtgactgg aaacagactg cggcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcaggtag tgttacgcct 120

atacaacgta tggtgagatc gggagaaaac gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 652
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 652

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Gly Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser

145 150 155

<210> 653
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 653
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcgtacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc tgaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacttgtt gaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aaaagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 654
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 654

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

```
Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65                  70                  75                  80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
                85                  90                  95

Leu Val Ile Asp Gly Val Thr Pro Asn Leu Leu Asn Tyr Phe Gly Arg
            100                 105                 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
        115                 120                 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Lys Asp Glu Arg Leu Ile Thr
    130                 135                 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145                 150                 155
```

<210> 655
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 655

```
ttcactttgg aaggtttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa       60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tgacagttag tgttacgcct      120

atacaacgta tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata      180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt      240

tacccagttg atgatcatca ctttaaagtt atatttcact atggaacttt agttatagtc      300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc      360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag      420

ggatcaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g               471
```

<210> 656
<211> 157
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 656

Phe Thr Leu Glu Gly Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Thr Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Phe His Tyr Gly Thr
85 90 95

Leu Val Ile Val Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Gly Ser Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 657
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 657
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta aggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 658
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 658

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Lys Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg

100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 659
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 659
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtg aggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 660
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 660

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
            20   25   30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Glu Val Arg Ser Gly
        35   40   45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
    50   55   60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65   70   75   80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
                85   90   95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
            100   105   110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
        115   120   125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
    130   135   140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145   150   155

<210> 661
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 661
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa   60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct   120

atacaacgtg cggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata   180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt   240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 662
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 662

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ala Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 663
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 663
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtt gtgtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 664
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 664

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Cys Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly

50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 665
<211> 472
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 665
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtt gggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc gt 472

<210> 666

<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 666

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Trp Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 667
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 667
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtg gggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 668
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 668

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1   5   10   15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
            20   25   30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Gly Val Arg Ser Gly
        35   40   45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
    50   55   60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65   70   75   80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
                85   90   95

```
Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
            100                 105                 110


Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
        115                 120                 125


Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
    130                 135                 140


Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145                 150                 155


<210>  669
<211>  471
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  669
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa       60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct      120

atacaacgtc atgtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata      180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt      240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac      300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc      360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag      420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g               471


<210>  670
<211>  157
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  670

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
```

1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg His Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 671
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 671
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta gtgtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 672
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 672

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ser Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr

130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 673
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 673
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtc aggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 674
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 674

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Gln Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
    50   55   60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65   70   75   80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
                85   90   95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
            100   105   110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
        115   120   125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
    130   135   140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145   150   155

<210> 675
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 675
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtc gggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 676
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 676

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Arg Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 677
<211> 471
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic

<400> 677
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgta cggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 678
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 678

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Thr Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr

85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 679
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 679
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtt atgtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 680
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 680

```
Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1               5                   10                  15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
            20                  25                  30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Tyr Val Arg Ser Gly
        35                  40                  45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
    50                  55                  60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65                  70                  75                  80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
                85                  90                  95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
            100                 105                 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
        115                 120                 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
    130                 135                 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145                 150                 155
```

<210> 681
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 681
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtc tggtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 682
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 682

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Leu Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 683
<211> 471
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 683
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa 60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct 120

atacaacgtc ctgtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata 180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt 240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac 300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc 360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag 420

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 684
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 684

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Pro Val Arg Ser Gly

```
        35                  40                  45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
    50                  55                  60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65                  70                  75                  80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
                85                  90                  95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
            100                 105                 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
        115                 120                 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
    130                 135                 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145                 150                 155


<210>  685
<211>  471
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  685
ttcactttgg aagatttcgt gggtgactgg aaacagactg cagcatacaa cttagaccaa       60

gtgttggaac agggtggagt tagttcgctt ttacagaatt tggcagttag tgttacgcct      120

atacaacgtt ttgtgagatc gggagaaaat gcattaaaaa tagacataca tgtgataata      180

ccttatgaag gattatcggc agaccaaatg gcacagatag aagaggtttt taaagttgtt      240

tacccagttg atgatcatca ctttaaagtt atattacact atggaacttt agttatagac      300

ggtgtgactc ctaacatgtt aaactatttc ggtagacctt atgaaggaat agcagttttc      360

gacggaaaga aaataacggt tactggtacg ttatggaacg gagccaagat aatagacgag      420
```

agattaataa cgccagacgg aagtatgagt ttccaggtta cgataaactc g 471

<210> 686
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 686

Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr
1 5 10 15

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
20 25 30

Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Phe Val Arg Ser Gly
35 40 45

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
50 55 60

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
65 70 75 80

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr
85 90 95

Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg
100 105 110

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr
115 120 125

Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr
130 135 140

Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 687

<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 687
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga cgatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 688
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 688

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 689
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 689
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacaag 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 690
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 690

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Lys
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 691
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 691
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacaag gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 692
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 692

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu

```
1               5                   10                  15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
            20                  25                  30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Lys Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 693
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 693
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcgagaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 694
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 694

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val

85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Glu Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 695
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 695
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacaagctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

<210> 696
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 696

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu

165 170 175

Phe Lys Lys Ile Ser
180

<210> 697
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 697
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catccagggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 698
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 698

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

```
Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Gln Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 699
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 699
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacaag 480

aagatgctgt tccgagtaac catccagaag gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 700
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 700

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Lys
145 150 155 160

Lys Met Leu Phe Arg Val Thr Ile Gln Lys Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 701
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 701
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tcatgaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 702
<211> 181
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 702

```
Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1               5                  10                  15


Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
            20                  25                  30


Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35                  40                  45


Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60


Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80


Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95


Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110


Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125


Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140


Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160


Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175


Phe Lys Lys Ile Ser
            180
```

<210> 703
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 703
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacacccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 704
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 704

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala

50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly His Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 705
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 705
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

aagatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 706
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 706

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu

130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Lys Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 707
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 707
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcga tcccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 708
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 708

```
Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1               5                  10                  15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
            20                  25                  30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 709
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 709
atgaaacatc accatcacca tcatgcgatc gccatggatt tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 710
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 710

Met Lys His His His His His His Ala Ile Ala Met Asp Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
            20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

```
Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 711
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 711

```
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt       60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg      120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc      180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc      240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac      300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg      360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactacc      420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc      480
```

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 712
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 712

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 713
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 713
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccat cggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 714
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 714

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp

```
            20                  25                  30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Ile Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 715
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 715
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gtgaccggct ggcggctgtg cgaacgcatt 540

ctg 543

<210> 716
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 716

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile

```
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Thr Gly Trp Arg Leu
                165                 170                 175

Cys Glu Arg Ile Leu
            180
```

<210> 717
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 717

```
atgaaacatc accatcacca tcatgcgatc gccatggatt tcacactcga cgatttcgtt       60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg      120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tcatgaggat tgtccggagc      180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc      240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac      300

tttaaggtga tcctgcccat cggcacactg gtaatcgacg gggttacgcc gaacaagctg      360

aactatttcg gacacccgta tgaaggcatc gccgtgttcg acggcgagaa gatcactacc      420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcga tcccgacaag      480

aagatgctgt tccgagtaac catccagaag gttagcggct ggcgcctgtt caagaagatc      540

agc                                                                    543
```

<210> 718

<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 718

Met Lys His His His His His His Ala Ile Ala Met Asp Phe Thr Leu
1 5 10 15

Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Ile Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp Lys
145 150 155 160

Lys Met Leu Phe Arg Val Thr Ile Gln Lys Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser

180

<210> 719
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 719
atgaaacatc accatcacca tcatgcgatc gccatggatt tcacactcga cgatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tcatgaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccat cggcacactg gtaatcgacg gggttacgcc gaacaagctg 360

aactatttcg gacacccgta tgaaggcatc gccgtgttcg acggcgagaa gatcactacc 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcga tcccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 720
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 720

Met Lys His His His His His His Ala Ile Ala Met Asp Phe Thr Leu
1 5 10 15

Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Ile Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp Lys
145 150 155 160

Lys Met Leu Phe Arg Val Thr Ile Gln Lys Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 721
<211> 522
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 721
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg 60

agcggtttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 120

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 180

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 240

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 300

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 360

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 420

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 480

gacgagcgcc tgatcacccc cgaccatcac catcaccatc at 522

<210> 722
<211> 174
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 722

Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp
20 25 30

Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly
35 40 45

Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met
50 55 60

Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val
65 70 75 80

Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu
85 90 95

Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val
100 105 110

Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys
115 120 125

Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly

130 135 140

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
145 150 155 160

Asp Glu Arg Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 723
<211> 522
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 723
atgaacgtga ccggctaccg gctgttcaag aagattagca acggctcgag cggtggctcg 60

agcggtttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 120

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 180

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 240

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 300

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 360

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 420

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 480

gacgagcgcc tgatcacccc cgaccatcac catcaccatc at 522

<210> 724
<211> 174
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 724

Met Asn Val Thr Gly Tyr Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp

20 25 30

Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly
35 40 45

Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met
50 55 60

Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val
65 70 75 80

Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu
85 90 95

Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val
100 105 110

Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys
115 120 125

Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly
130 135 140

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
145 150 155 160

Asp Glu Arg Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 725
<211> 516
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 725
atggtgaccg gctaccggct gttcgagaag attagcggct cgagcggtgg ctcgagcggt 60

ttcacactcg acgatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa 120

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg 180

atcatgagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc 240

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg 300

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac 360

ggggttacgc cgaacaagct gaactatttc ggacggccgt atgaaggcat cgccgtgttc 420

gacggcaaaa agatcactac cacagggacc ctgtggaacg gcaacaaaat tatcgacgag 480

cgcctgatca cccccgacca tcaccatcac catcat 516

<210> 726
<211> 172
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 726

Met Val Thr Gly Tyr Arg Leu Phe Glu Lys Ile Ser Gly Ser Ser Gly
1 5 10 15

Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln
20 25 30

Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser
35 40 45

Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile
50 55 60

Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile
65 70 75 80

Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val
85 90 95

Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu
100 105 110

Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn
115 120 125

Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys
130 135 140

Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu
145 150 155 160

Arg Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 727
<211> 516
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 727
atggtgaccg gctaccggct gttcgagaag attctgggct cgagcggtgg ctcgagcggt 60

ttcacactcg acgatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa 120

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg 180

atcatgagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc 240

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg 300

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac 360

ggggttacgc cgaacaagct gaactatttc ggacggccgt atgaaggcat cgccgtgttc 420

gacggcaaaa agatcactac cacagggacc ctgtggaacg gcaacaaaat tatcgacgag 480

cgcctgatca cccccgacca tcaccatcac catcat 516

<210> 728
<211> 172
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 728

Met Val Thr Gly Tyr Arg Leu Phe Glu Lys Ile Leu Gly Ser Ser Gly
1 5 10 15

```
Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln
            20                  25                  30

Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser
        35                  40                  45

Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile
    50                  55                  60

Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile
65                  70                  75                  80

Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val
                85                  90                  95

Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu
            100                 105                 110

Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn
        115                 120                 125

Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys
    130                 135                 140

Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu
145                 150                 155                 160

Arg Leu Ile Thr Pro Asp His His His His His His
                165                 170
```

<210> 729
<211> 516
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 729
atggtgagcg gctggcggct gttcaagaag attagcggct cgagcggtgg ctcgagcggt 60

ttcacactcg acgatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa 120

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg 180

atcatgagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc 240

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg 300

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac 360

ggggttacgc cgaacaagct gaactatttc ggacggccgt atgaaggcat cgccgtgttc 420

gacggcaaaa agatcactac cacagggacc ctgtggaacg gcaacaaaat tatcgacgag 480

cgcctgatca cccccgacca tcaccatcac catcat 516

<210> 730
<211> 172
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 730

Met Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser Ser Gly
1 5 10 15

Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln
20 25 30

Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser
35 40 45

Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile
50 55 60

Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile
65 70 75 80

Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val
85 90 95

Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu
100 105 110

Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn

115 120 125

Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys
130 135 140

Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu
145 150 155 160

Arg Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 731
<211> 516
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 731
atggtgagcg gctggcggct gttcaagaag attggcggct cgagcggtgg ctcgagcggt 60

ttcacactcg acgatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa 120

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg 180

atcatgagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc 240

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg 300

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac 360

ggggttacgc cgaacaagct gaactatttc ggacggccgt atgaaggcat cgccgtgttc 420

gacggcaaaa agatcactac cacagggacc ctgtggaacg gcaacaaaat tatcgacgag 480

cgcctgatca cccccgacca tcaccatcac catcat 516

<210> 732
<211> 172
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 732

Met Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Gly Gly Ser Ser Gly

1 5 10 15

Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln
20 25 30

Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser
35 40 45

Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile
50 55 60

Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile
65 70 75 80

Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val
85 90 95

Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu
100 105 110

Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn
115 120 125

Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys
130 135 140

Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu
145 150 155 160

Arg Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 733
<211> 516
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 733
atggtgagcg gctggcggct gttcaagaag tcgggcggct cgagcggtgg ctcgagcggt 60

ttcacactcg acgatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa 120

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg 180

atcatgagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc 240

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg 300

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac 360

ggggttacgc cgaacaagct gaactatttc ggacggccgt atgaaggcat cgccgtgttc 420

gacggcaaaa agatcactac cacagggacc ctgtggaacg gcaacaaaat tatcgacgag 480

cgcctgatca cccccgacca tcaccatcac catcat 516

<210> 734
<211> 172
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 734

Met Val Ser Gly Trp Arg Leu Phe Lys Lys Ser Gly Gly Ser Ser Gly
1 5 10 15

Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln
20 25 30

Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser
35 40 45

Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile
50 55 60

Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile
65 70 75 80

Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val
85 90 95

Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu
100 105 110

Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn
115 120 125

Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys
130 135 140

Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu
145 150 155 160

Arg Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 735
<211> 510
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 735
atggtgagcg gctggcggct gttcaagaag ggctcgagcg gtggctcgag cggtttcaca 60

ctcgacgatt tcgttgggga ctgggaacag acagccgcct acaacctgga ccaagtcctt 120

gaacagggag gtgtgtccag tttgctgcag aatctcgccg tgtccgtaac tccgatcatg 180

aggattgtcc ggagcggtga aaatgccctg aagatcgaca tccatgtcat catcccgtat 240

gaaggtctga gcgccgacca aatggcccag atcgaagagg tgtttaaggt ggtgtaccct 300

gtggatgatc atcactttaa ggtgatcctg ccctatggca cactggtaat cgacggggtt 360

acgccgaaca agctgaacta tttcggacgg ccgtatgaag gcatcgccgt gttcgacggc 420

aaaaagatca ctaccacagg gaccctgtgg aacggcaaca aaattatcga cgagcgcctg 480

atcacccccg accatcacca tcaccatcat 510

<210> 736
<211> 170
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 736

```
Met Val Ser Gly Trp Arg Leu Phe Lys Lys Gly Ser Ser Gly Gly Ser
1               5                  10                  15

Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
            20                  25                  30

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
        35                  40                  45

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
    50                  55                  60

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
65                  70                  75                  80

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
                85                  90                  95

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
            100                 105                 110

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
        115                 120                 125

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
    130                 135                 140

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
145                 150                 155                 160

Ile Thr Pro Asp His His His His His His
                165                 170
```

<210> 737
<211> 513
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 737

atggtgagcg gctggcggct gttcaagaag attggctcga gcggtggctc gagcggtttc 60

acactcgacg atttcgttgg ggactgggaa cagacagccg cctacaacct ggaccaagtc 120

cttgaacagg gaggtgtgtc cagtttgctg cagaatctcg ccgtgtccgt aactccgatc 180

atgaggattg tccggagcgg tgaaaatgcc ctgaagatcg acatccatgt catcatcccg 240

tatgaaggtc tgagcgccga ccaaatggcc cagatcgaag aggtgtttaa ggtggtgtac 300

cctgtggatg atcatcactt taaggtgatc ctgccctatg gcacactggt aatcgacggg 360

gttacgccga acaagctgaa ctatttcgga cggccgtatg aaggcatcgc cgtgttcgac 420

ggcaaaaaga tcactaccac agggaccctg tggaacggca acaaaattat cgacgagcgc 480

ctgatcaccc ccgaccatca ccatcaccat cat 513

<210> 738
<211> 171
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 738

Met Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Gly Ser Ser Gly Gly
1 5 10 15

Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr
20 25 30

Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser
35 40 45

Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val
50 55 60

Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro
65 70 75 80

Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe
85 90 95

Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro

100 105 110

Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr
115 120 125

Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile
130 135 140

Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg
145 150 155 160

Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 739
<211> 522
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 739
atggtttccg tgagcggctg gcggctgttc aagaagtcgg gcggctcgag cggtggctcg 60

agcggtttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 120

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 180

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 240

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 300

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 360

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 420

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 480

gacgagcgcc tgatcacccc cgaccatcac catcaccatc at 522

<210> 740
<211> 174
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 740

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ser Gly Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp
20 25 30

Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly
35 40 45

Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met
50 55 60

Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val
65 70 75 80

Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu
85 90 95

Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val
100 105 110

Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys
115 120 125

Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly
130 135 140

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
145 150 155 160

Asp Glu Arg Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 741
<211> 522
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 741
atggtttccg tgagcggctg gcggctgttc aagaagattg gcggctcgag cggtggctcg 60

agcggtttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 120

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 180

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 240

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 300

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 360

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 420

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 480

gacgagcgcc tgatcacccc cgaccatcac catcaccatc at 522

<210> 742
<211> 174
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 742

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Gly Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp
20 25 30

Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly
35 40 45

Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met
50 55 60

Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val
65 70 75 80

Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu
85 90 95

```
Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val
            100                 105                 110

Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys
        115                 120                 125

Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly
    130                 135                 140

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
145                 150                 155                 160

Asp Glu Arg Leu Ile Thr Pro Asp His His His His His His
                165                 170
```

<210> 743
<211> 516
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 743
atggtttccg tgagcggctg gcggctgttc aagaagggct cgagcggtgg ctcgagcggt 60

ttcacactcg acgatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa 120

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg 180

atcatgagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc 240

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg 300

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac 360

ggggttacgc cgaacaagct gaactatttc ggacggccgt atgaaggcat cgccgtgttc 420

gacggcaaaa agatcactac cacagggacc ctgtggaacg gcaacaaaat tatcgacgag 480

cgcctgatca cccccgacca tcaccatcac catcat 516

<210> 744
<211> 172
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 744

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Gly Ser Ser Gly
1   5   10   15

Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln
            20   25   30

Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser
        35   40   45

Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile
    50   55   60

Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile
65   70   75   80

Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val
                85   90   95

Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu
            100   105   110

Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn
        115   120   125

Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys
    130   135   140

Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu
145   150   155   160

Arg Leu Ile Thr Pro Asp His His His His His His
                165   170

<210> 745
<211> 518
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 745
atggtttccg tgagcggctg gcggctgttc aagaagattg gctcgagcgg tggctcgagc 60

ggtttcacac tcgacgattt cgttggggac tgggaacaga cagccgccta caacctggac 120

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 180

ccgatcatga ggattgtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 240

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 300

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 360

gacggggtta cgccgaacaa gctgaactat ttcggacggc cgtatgaagg catcgccgtg 420

ttcgacggca aaaagatcac taccacaggg accctgtgga acggcaacaa aattatcgac 480

gagcgcctga tcacccccga ccatcaccat caccatca 518

<210> 746
<211> 173
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 746

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Gly Ser Ser
1   5   10   15

Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu
            20   25   30

Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val
        35   40   45

Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg
    50   55   60

Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile
65   70   75   80

Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu

```
                85                  90                  95

Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile
            100                 105                 110

Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu
        115                 120                 125

Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys
    130                 135                 140

Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp
145                 150                 155                 160

Glu Arg Leu Ile Thr Pro Asp His His His His His His
                165                 170
```

<210> 747
<211> 513
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 747

```
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggttcgttc       60

acactcgacg atttcgttgg ggactgggaa cagacagccg cctacaacct ggaccaagtc      120

cttgaacagg gaggtgtgtc cagtttgctg cagaatctcg ccgtgtccgt aactccgatc      180

atgaggattg tccggagcgg tgaaaatgcc ctgaagatcg acatccatgt catcatcccg      240

tatgaaggtc tgagcgccga ccaaatggcc cagatcgaag aggtgtttaa ggtggtgtac      300

cctgtggatg atcatcactt taaggtgatc ctgccctatg gcacactggt aatcgacggg      360

gttacgccga acaagctgaa ctatttcgga cggccgtatg aaggcatcgc cgtgttcgac      420

ggcaaaaaga tcactaccac agggaccctg tggaacggca acaaaattat cgacgagcgc      480

ctgatcaccc ccgaccatca ccatcaccat cat                                   513
```

<210> 748
<211> 171
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 748

```
Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1               5                  10                  15


Ser Gly Ser Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr
            20                  25                  30


Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser
        35                  40                  45


Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val
    50                  55                  60


Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro
65                  70                  75                  80


Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe
                85                  90                  95


Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro
            100                 105                 110


Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr
        115                 120                 125


Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile
    130                 135                 140


Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg
145                 150                 155                 160


Leu Ile Thr Pro Asp His His His His His His
                165                 170
```

<210> 749
<211> 516
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic

<400> 749
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg 60

ttcacactcg acgatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa 120

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg 180

atcatgagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc 240

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg 300

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac 360

ggggttacgc cgaacaagct gaactatttc ggacggccgt atgaaggcat cgccgtgttc 420

gacggcaaaa agatcactac cacagggacc ctgtggaacg gcaacaaaat tatcgacgag 480

cgcctgatca cccccgacca tcaccatcac catcat 516

<210> 750
<211> 172
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 750

Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

Ser Gly Gly Ser Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln
20 25 30

Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser
35 40 45

Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile
50 55 60

Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile
65 70 75 80

```
Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val
                85                  90                  95

Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu
            100                 105                 110

Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn
        115                 120                 125

Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys
    130                 135                 140

Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu
145                 150                 155                 160

Arg Leu Ile Thr Pro Asp His His His His His His
                165                 170
```

<210> 751
<211> 519
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 751

```
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg       60

ggtttcacac tcgacgattt cgttggggac tgggaacaga cagccgccta caacctggac      120

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact      180

ccgatcatga ggattgtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc      240

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg      300

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc      360

gacggggtta cgccgaacaa gctgaactat ttcggacggc cgtatgaagg catcgccgtg      420

ttcgacggca aaaagatcac taccacaggg accctgtgga acggcaacaa aattatcgac      480

gagcgcctga tcacccccga ccatcaccat caccatcat                             519
```

<210> 752

<211> 173
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 752

Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

Ser Gly Gly Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu
20 25 30

Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val
35 40 45

Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg
50 55 60

Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile
65 70 75 80

Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu
85 90 95

Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile
100 105 110

Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu
115 120 125

Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys
130 135 140

Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp
145 150 155 160

Glu Arg Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 753

<211> 525
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 753
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg 60

agcggtggtt tcacactcga cgatttcgtt ggggactggg aacagacagc cgcctacaac 120

ctggaccaag tccttgaaca gggaggtgtg tccagtttgc tgcagaatct cgccgtgtcc 180

gtaactccga tcatgaggat tgtccggagc ggtgaaaatg ccctgaagat cgacatccat 240

gtcatcatcc cgtatgaagg tctgagcgcc gaccaaatgg cccagatcga agaggtgttt 300

aaggtggtgt accctgtgga tgatcatcac tttaaggtga tcctgcccta tggcacactg 360

gtaatcgacg gggttacgcc gaacaagctg aactatttcg gacggccgta tgaaggcatc 420

gccgtgttcg acggcaaaaa gatcactacc acagggaccc tgtggaacgg caacaaaatt 480

atcgacgagc gcctgatcac ccccgaccat caccatcacc atcat 525

<210> 754
<211> 175
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 754

Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Gly Phe Thr Leu Asp Asp Phe Val Gly Asp
20 25 30

Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly
35 40 45

Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile
50 55 60

Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His

65 70 75 80

Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile
85 90 95

Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys
100 105 110

Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn
115 120 125

Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp
130 135 140

Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile
145 150 155 160

Ile Asp Glu Arg Leu Ile Thr Pro Asp His His His His His His
165 170 175

<210> 755
<211> 528
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 755
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg 60

agcggtggta gcttcacact cgacgatttc gttggggact gggaacagac agccgcctac 120

aacctggacc aagtccttga acagggaggt gtgtccagtt tgctgcagaa tctcgccgtg 180

tccgtaactc cgatcatgag gattgtccgg agcggtgaaa atgccctgaa gatcgacatc 240

catgtcatca tcccgtatga aggtctgagc gccgaccaaa tggcccagat cgaagaggtg 300

tttaaggtgg tgtaccctgt ggatgatcat cactttaagg tgatcctgcc ctatggcaca 360

ctggtaatcg acggggttac gccgaacaag ctgaactatt tcggacggcc gtatgaaggc 420

atcgccgtgt tcgacggcaa aaagatcact accacaggga ccctgtggaa cggcaacaaa 480

attatcgacg agcgcctgat caccccgac catcaccatc accatcat 528

<210> 756
<211> 176
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 756

```
Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1               5                  10                  15


Ser Gly Gly Ser Ser Gly Gly Ser Phe Thr Leu Asp Asp Phe Val Gly
            20                  25                  30


Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln
        35                  40                  45


Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro
    50                  55                  60


Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile
65                  70                  75                  80


His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln
                85                  90                  95


Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe
            100                 105                 110


Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro
        115                 120                 125


Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe
    130                 135                 140


Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys
145                 150                 155                 160


Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp His His His His His His
                165                 170                 175
```

<210> 757
<211> 531
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 757
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg 60

agcggtggta gctcgttcac actcgacgat ttcgttgggg actgggaaca gacagccgcc 120

tacaacctgg accaagtcct tgaacaggga ggtgtgtcca gtttgctgca gaatctcgcc 180

gtgtccgtaa ctccgatcat gaggattgtc cggagcggtg aaaatgccct gaagatcgac 240

atccatgtca tcatcccgta tgaaggtctg agcgccgacc aaatggccca gatcgaagag 300

gtgtttaagg tggtgtaccc tgtggatgat catcacttta aggtgatcct gccctatggc 360

acactggtaa tcgacggggt tacgccgaac aagctgaact atttcggacg gccgtatgaa 420

ggcatcgccg tgttcgacgg caaaaagatc actaccacag ggaccctgtg gaacggcaac 480

aaaattatcg acgagcgcct gatcaccccc gaccatcacc atcaccatca t 531

<210> 758
<211> 177
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 758

Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Gly Ser Ser Phe Thr Leu Asp Asp Phe Val
20 25 30

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
35 40 45

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
50 55 60

```
Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
65                  70                  75                  80


Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
                85                  90                  95


Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
            100                 105                 110


Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
        115                 120                 125


Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
    130                 135                 140


Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
145                 150                 155                 160


Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp His His His His His
                165                 170                 175


His
```

<210> 759
<211> 522
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 759

```
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg       60

agcggtttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg      120

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta      180

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc      240

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag      300

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta      360
```

atcgacgggg ttacgccgaa caagctgaac tatttcggac acccgtatga aggcatcgcc 420

gtgttcgacg gcgagaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 480

gacgagcgcc tgatcacccc cgaccatcac catcaccatc at 522

<210> 760
<211> 174
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 760

Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp
20 25 30

Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly
35 40 45

Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met
50 55 60

Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val
65 70 75 80

Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu
85 90 95

Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val
100 105 110

Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys
115 120 125

Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly
130 135 140

Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile

145 150 155 160

Asp Glu Arg Leu Ile Thr Pro Asp His His His His His His
165 170

<210> 761
<211> 522
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 761
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg 60

agcggtttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 120

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 180

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 240

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 300

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 360

atcgacgggg ttacgccgaa caagctgaac tatttcggac acccgtatga aggcatcgcc 420

gtgttcgacg gcgagaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 480

gacgagcgcc tgatcgatcc cgaccatcac catcaccatc at 522

<210> 762
<211> 174
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 762

Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp
20 25 30

Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly

35 40 45

Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met
50 55 60

Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val
65 70 75 80

Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu
85 90 95

Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val
100 105 110

Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys
115 120 125

Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly
130 135 140

Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
145 150 155 160

Asp Glu Arg Leu Ile Asp Pro Asp His His His His His His
165 170

<210> 763
<211> 531
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 763
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg 60

agcggtggta gctcgttcac actcgacgat ttcgttgggg actgggaaca gacagccgcc 120

tacaacctgg accaagtcct tgaacaggga ggtgtgtcca gtttgctgca gaatctcgcc 180

gtgtccgtaa ctccgatcat gaggattgtc cggagcggtg aaaatgccct gaagatcgac 240

atccatgtca tcatcccgta tgaaggtctg agcgccgacc aaatggccca gatcgaagag 300

gtgtttaagg tggtgtaccc tgtggatgat catcacttta aggtgatcct gccctatggc 360

acactggtaa tcgacggggt tacgccgaac aagctgaact atttcggaca cccgtatgaa 420

ggcatcgccg tgttcgacgg cgagaagatc actaccacag ggaccctgtg gaacggcaac 480

aaaattatcg acgagcgcct gatcaccccc gaccatcacc atcaccatca t 531

<210> 764
<211> 177
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 764

Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Gly Ser Ser Phe Thr Leu Asp Asp Phe Val
20 25 30

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
35 40 45

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
50 55 60

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
65 70 75 80

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
85 90 95

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
100 105 110

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
115 120 125

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val
130 135 140

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
145 150 155 160

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp His His His His His
165 170 175

His

<210> 765
<211> 531
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 765
atgaacgtga gcggctggcg gctgttcaag aagattagca acggctcgag cggtggctcg 60

agcggtggta gctcgttcac actcgacgat ttcgttgggg actgggaaca gacagccgcc 120

tacaacctgg accaagtcct tgaacaggga ggtgtgtcca gtttgctgca gaatctcgcc 180

gtgtccgtaa ctccgatcat gaggattgtc cggagcggtg aaaatgccct gaagatcgac 240

atccatgtca tcatcccgta tgaaggtctg agcgccgacc aaatggccca gatcgaagag 300

gtgtttaagg tggtgtaccc tgtggatgat catcacttta aggtgatcct gccctatggc 360

acactggtaa tcgacggggt tacgccgaac aagctgaact atttcggaca cccgtatgaa 420

ggcatcgccg tgttcgacgg cgagaagatc actaccacag ggaccctgtg gaacggcaac 480

aaaattatcg acgagcgcct gatcgatccc gaccatcacc atcaccatca t 531

<210> 766
<211> 177
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 766

Met Asn Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Asn Gly Ser
1 5 10 15

```
Ser Gly Gly Ser Ser Gly Gly Ser Ser Phe Thr Leu Asp Asp Phe Val
            20                  25                  30

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
        35                  40                  45

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
    50                  55                  60

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
65                  70                  75                  80

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
                85                  90                  95

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
            100                 105                 110

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
        115                 120                 125

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val
    130                 135                 140

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
145                 150                 155                 160

Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp His His His His His
                165                 170                 175

His
```

<210> 767
<211> 477
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 767

atggtcttca cactcgaaga tttcgttggg gactggaagc agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggatggt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgcactatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcgc caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgtccttcc aggtaaccat caacagc 477

<210> 768
<211> 159
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 768

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 769
<211> 477
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 769
atggtcttca cactcgaaga tttcgttggg gactggaagc agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggatggt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgcactatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcgc caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgtccttcc aggtaaccat caacagc 477

<210> 770
<211> 159
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 770

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu

20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 771
<211> 477
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 771
atggtcttca cactcgaaga tttcgttggg gactggaagc agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggatagt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgcactatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcgc caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgtccttcc aggtaaccat caacagc 477

<210> 772
<211> 159
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 772

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser

145 150 155

<210> 773
<211> 477
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 773
atggtcttca cactcgaaga tttcgttggg gactggaagc agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaagggtggt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgcactatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcgc caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgtccttcc aggtaaccat caacagc 477

<210> 774
<211> 159
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 774

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Val Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

```
Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65                  70                  75                  80


Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr
                85                  90                  95


Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
            100                 105                 110


Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125


Val Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140


Ile Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145                 150                 155


<210>  775
<211>  477
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  775
atggtcttca cactcgaaga tttcgttggg gactggaagc agacagccgc ctacaacctg       60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta      120

actccgatcc aaaggtttgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc      180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag      240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgcactatgg cacactggta      300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc      360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcgc caaaattatc      420

gacgagcgcc tgatcacccc cgacggctcc atgtccttcc aggtaaccat caacagc         477


<210>  776
<211>  159
<212>  PRT
```

<213> Artificial sequence

<220>
<223> Synthetic

<400> 776

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Phe Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 777
<211> 477
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 777
atggtcttca cactcgaaga tttcgttggg gactggaagc agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggctggt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgcactatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcgc caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgtccttcc aggtaaccat caacagc 477

<210> 778
<211> 159
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 778

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Leu Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe

100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 779
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 779
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga cgatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacaagctg 360

aactatttcg gacacccgta tgaaggcatc gccgtgttcg acggcgagaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 780
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 780

```
Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1               5                  10                  15

Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
            20                  25                  30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Glu Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 781
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 781
atgaaacatc accatcacca tcatgcgatc gccatggatt tcacactcga cgatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacaagctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcgagaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 782
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 782

Met Lys His His His His His His Ala Ile Ala Met Asp Phe Thr Leu
1   5   10   15

Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
            20   25   30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35   40   45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50   55   60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65   70   75   80

```
Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Glu Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 783
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 783

```
atgaaacatc accatcacca tcatgcgatc gccatggatt tcacactcga cgatttcgtt       60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg      120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc      180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc      240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac      300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacaagctg      360

aactatttcg gacacccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta      420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc      480
```

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 784
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 784

Met Lys His His His His His His Ala Ile Ala Met Asp Phe Thr Leu
1 5 10 15

Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 785
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 785
atgaaacatc accatcacca tcatgcgatc gccatggatt tcacactcga cgatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacacccgta tgaaggcatc gccgtgttcg acggcgagaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 786
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 786

Met Lys His His His His His His Ala Ile Ala Met Asp Phe Thr Leu
1 5 10 15

Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp

```
            20                  25                  30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly His Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Glu Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 787
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 787
atgaaacatc accatcacca tcatgcgatc gccatggatt tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacaagctg 360

aactatttcg gacacccgta tgaaggcatc gccgtgttcg acggcgagaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 788
<211> 170
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 788

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe

```
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val
145                 150                 155                 160

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
                165                 170
```

<210> 789
<211> 170
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 789

```
Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1               5                   10                  15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
            20                  25                  30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
        35                  40                  45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
    50                  55                  60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65                  70                  75                  80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
```

```
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val
145                 150                 155                 160

Thr Gly Tyr Arg Leu Phe Glu Glu Ile Leu
                165                 170
```

<210> 790
<211> 102
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 790

```
Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1               5                   10                  15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
            20                  25                  30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
        35                  40                  45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
    50                  55                  60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65                  70                  75                  80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp
```

100

<210> 791
<211> 124
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 791

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly
115 120

<210> 792
<211> 133
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 792

```
Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1               5                  10                  15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
            20                  25                  30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
        35                  40                  45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
    50                  55                  60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65                  70                  75                  80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Thr Thr Gly Thr Leu
    130
```

<210> 793
<211> 148
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 793

```
Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1               5                  10                  15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
            20                  25                  30
```

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp
145

<210> 794
<211> 68
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 794

Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly
1 5 10 15

Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp
20 25 30

Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser
35 40 45

Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Arg Leu Phe
    50   55   60

Lys Lys Ile Ser
65

<210> 795
<211> 46
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 795

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
1   5   10   15

Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr
            20   25   30

Ile Asn Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
        35   40   45

<210> 796
<211> 37
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 796

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
1   5   10   15

Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Arg Leu
            20   25   30

Phe Lys Lys Ile Ser
        35

<210> 797

<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 797

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Ser Gly Trp Arg
1 5 10 15

Leu Phe Lys Lys Ile Ser
20

<210> 798
<211> 68
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 798

Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly
1 5 10 15

Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp
20 25 30

Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser
35 40 45

Met Leu Phe Arg Val Thr Ile Asn Ser Val Thr Gly Tyr Arg Leu Phe
50 55 60

Glu Glu Ile Leu
65

<210> 799
<211> 46
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 799

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
1 5 10 15

Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr
20 25 30

Ile Asn Ser Val Thr Gly Tyr Arg Leu Phe Glu Glu Ile Leu
35 40 45

<210> 800
<211> 37
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 800

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
1 5 10 15

Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Thr Gly Tyr Arg Leu
20 25 30

Phe Glu Glu Ile Leu
35

<210> 801
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 801

Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Val Thr Gly Tyr Arg
1 5 10 15

Leu Phe Glu Glu Ile Leu
20

<210> 802
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 802

Met Lys His His His His His His Ala Ile Ala Met Asp Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Glu Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 803
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 803
atgaaacatc accatcacca tcatgcgatc gccatggatt tcacactcga cgatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacaagctg 360

aactatttcg gacacccgta tgaaggcatc gccgtgttcg acggcgagaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 804
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 804

Met Lys His His His His His His Ala Ile Ala Met Asp Phe Thr Leu
1 5 10 15

Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala

```
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Glu Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180


<210>  805
<211>  999
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  805
atgaaacatc atcaccatca ccacgcagaa atcggtactg gctttccatt cgacccccat       60

tatgtggaag tcctgggcga gcgcatgcac tacgtcgatg ttggtccgcg cgatggcacc      120

cctgtgctgt tcctgcacgg taacccgacc tcctcctacg tgtggcgcaa catcatcccg      180
```

catgttgcac cgacccatcg ctgcattgct ccagacctga tcggtatggg caaatccgac 240

aaaccagacc tgggttattt cttcgacgac cacgtccgct tcatggatgc cttcatcgaa 300

gccctgggtc tggaagaggt cgtcctggtc attcacgact ggggctccgc tctgggtttc 360

cactgggcca agcgcaatcc agagcgcgtc aaaggtattg catttatgga gttcatccgc 420

cctatcccga cctgggacga atggccagaa tttgcccgcg agaccttcca ggccttccgc 480

accaccgacg tcggccgcaa gctgatcatc gatcagaacg tttttatcga gggtacgctg 540

ccgatgggtg tcgtccgccc gctgactgaa gtcgagatgg accattaccg cgagccgttc 600

ctgaatcctg ttgaccgcga gccactgtgg cgcttcccaa acgagctgcc aatcgccggt 660

gagccagcga acatcgtcgc gctggtcgaa gaatacatgg actggctgca ccagtcccct 720

gtcccgaagc tgctgttctg gggcacccca ggcgttctga tcccaccggc cgaagccgct 780

cgcctggcca aaagcctgcc taactgcaag gctgtggaca tcggcccggg tctgaatctg 840

ctgcaagaag acaacccgga cctgatcggc agcgagatcg cgcgctggct gtcgacgctc 900

gagatttccg gcgagccaac cactagcgat aacggctcca tgctgttccg agtaaccatc 960

aacagcgtta gcggctggcg cctgttcaag aagatcagc 999

<210> 806
<211> 333
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 806

Met Lys His His His His His His Ala Glu Ile Gly Thr Gly Phe Pro
1 5 10 15

Phe Asp Pro His Tyr Val Glu Val Leu Gly Glu Arg Met His Tyr Val
20 25 30

Asp Val Gly Pro Arg Asp Gly Thr Pro Val Leu Phe Leu His Gly Asn
35 40 45

Pro Thr Ser Ser Tyr Val Trp Arg Asn Ile Ile Pro His Val Ala Pro
50 55 60

```
Thr His Arg Cys Ile Ala Pro Asp Leu Ile Gly Met Gly Lys Ser Asp
65                  70                  75                  80

Lys Pro Asp Leu Gly Tyr Phe Phe Asp Asp His Val Arg Phe Met Asp
                85                  90                  95

Ala Phe Ile Glu Ala Leu Gly Leu Glu Glu Val Val Leu Val Ile His
            100                 105                 110

Asp Trp Gly Ser Ala Leu Gly Phe His Trp Ala Lys Arg Asn Pro Glu
        115                 120                 125

Arg Val Lys Gly Ile Ala Phe Met Glu Phe Ile Arg Pro Ile Pro Thr
    130                 135                 140

Trp Asp Glu Trp Pro Glu Phe Ala Arg Glu Thr Phe Gln Ala Phe Arg
145                 150                 155                 160

Thr Thr Asp Val Gly Arg Lys Leu Ile Ile Asp Gln Asn Val Phe Ile
                165                 170                 175

Glu Gly Thr Leu Pro Met Gly Val Val Arg Pro Leu Thr Glu Val Glu
            180                 185                 190

Met Asp His Tyr Arg Glu Pro Phe Leu Asn Pro Val Asp Arg Glu Pro
        195                 200                 205

Leu Trp Arg Phe Pro Asn Glu Leu Pro Ile Ala Gly Glu Pro Ala Asn
    210                 215                 220

Ile Val Ala Leu Val Glu Glu Tyr Met Asp Trp Leu His Gln Ser Pro
225                 230                 235                 240

Val Pro Lys Leu Leu Phe Trp Gly Thr Pro Gly Val Leu Ile Pro Pro
                245                 250                 255

Ala Glu Ala Ala Arg Leu Ala Lys Ser Leu Pro Asn Cys Lys Ala Val
            260                 265                 270

Asp Ile Gly Pro Gly Leu Asn Leu Leu Gln Glu Asp Asn Pro Asp Leu
```

275 280 285

Ile Gly Ser Glu Ile Ala Arg Trp Leu Ser Thr Leu Glu Ile Ser Gly
290 295 300

Glu Pro Thr Thr Ser Asp Asn Gly Ser Met Leu Phe Arg Val Thr Ile
305 310 315 320

Asn Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
325 330

<210> 807
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 807
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactgga agcagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 808
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 808

```
Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1               5                  10                  15

Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr Asn Leu Asp
            20                  25                  30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 809
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 809
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggc gccagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgta tgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 810
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 810

Glu Lys Met Leu Phe Arg Val Thr Ile Glu Ser Trp Lys
1   5   10

<210> 811
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 811

Glu Lys Leu Leu Phe Arg Val Thr Ile Glu Ser Trp Lys
1   5   10

<210> 812

<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 812

Glu Lys Leu Leu Phe Arg Val Thr Ile Glu Ser Tyr Lys
1 5 10

<210> 813
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 813

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Arg Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 814
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 814
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgtt cgaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 815
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 815

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
20 25 30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
35 40 45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
50 55 60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65 70 75 80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
85 90 95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
100 105 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Phe Glu
115 120 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
130 135 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145 150 155 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
165 170 175

Phe Lys Lys Ile Ser
180

<210> 816
<211> 13
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 816

Asp Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Tyr Lys
1 5 10

<210> 817
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 817

Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly
1 5 10 15

Ile Ala Val Phe Asp Gly
20

<210> 818
<211> 9
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 818

Lys Lys Ile Thr Thr Thr Gly Thr Leu
1 5

<210> 819
<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 819

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp

1 5 10 15

<210> 820
<211> 396
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 820
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggctaa 396

<210> 821
<211> 131
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 821

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala

65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly
130

<210> 822
<211> 654
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 822
gggagctccg gtggtggcgg gagcggaggt ggaggctcga gcggtatgac gtataagtta 60

atccttaatg gtaaaacatt gaaaggcgag acaactactg aagctgttga tgctgctact 120

gcagaaaaag tcttcaaaca atacgctaac gacaacggtg ttgacggtga atggacttac 180

gacgatgcga cgaaaacctt tacggtcacc gaaaaaccag aagtgatcga tgcgtctgaa 240

ttaacaccag ccgtgacaac ttacaaactt gttattaatg gtaaaacatt gaaaggcgaa 300

acaactactg aggctgttga tgctgctact gcagagaagg tgttcaaaca atatgcgaat 360

gacaacggtg ttgacggtga gtggacttac gacgatgcga ctaagacctt tacagttact 420

gaaaaaccag aagtgatcga tgcgtctgag ttaacaccag ccgtgacaac ttacaaactt 480

gttattaatg gtaaaacatt gaaaggcgaa acaactacta aagcagtaga cgcagaaact 540

gcggagaagg ccttcaaaca atacgctaac gacaacggtg ttgatggtgt ttggacttat 600

gatgatgcca caaaaacctt tacggtaact gagcatcatc accatcacca ctaa 654

<210> 823
<211> 217
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 823

```
Gly Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser Gly Met
1               5                   10                  15


Thr Tyr Lys Leu Ile Leu Asn Gly Lys Thr Leu Lys Gly Glu Thr Thr
            20                  25                  30


Thr Glu Ala Val Asp Ala Ala Thr Ala Glu Lys Val Phe Lys Gln Tyr
        35                  40                  45


Ala Asn Asp Asn Gly Val Asp Gly Glu Trp Thr Tyr Asp Asp Ala Thr
    50                  55                  60


Lys Thr Phe Thr Val Thr Glu Lys Pro Glu Val Ile Asp Ala Ser Glu
65                  70                  75                  80


Leu Thr Pro Ala Val Thr Thr Tyr Lys Leu Val Ile Asn Gly Lys Thr
                85                  90                  95


Leu Lys Gly Glu Thr Thr Thr Glu Ala Val Asp Ala Ala Thr Ala Glu
            100                 105                 110


Lys Val Phe Lys Gln Tyr Ala Asn Asp Asn Gly Val Asp Gly Glu Trp
        115                 120                 125


Thr Tyr Asp Asp Ala Thr Lys Thr Phe Thr Val Thr Glu Lys Pro Glu
    130                 135                 140


Val Ile Asp Ala Ser Glu Leu Thr Pro Ala Val Thr Thr Tyr Lys Leu
145                 150                 155                 160


Val Ile Asn Gly Lys Thr Leu Lys Gly Glu Thr Thr Thr Lys Ala Val
                165                 170                 175


Asp Ala Glu Thr Ala Glu Lys Ala Phe Lys Gln Tyr Ala Asn Asp Asn
            180                 185                 190
```

Gly Val Asp Gly Val Trp Thr Tyr Asp Asp Ala Thr Lys Thr Phe Thr
195 200 205

Val Thr Glu His His His His His His
210 215

<210> 824
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 824

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 825
<211> 10
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 825

Val Ser Gly Trp Arg Leu Phe Lys Lys Ile
1 5 10

<210> 826
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 826

Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
1 5 10

<210> 827
<211> 13
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 827

Gly Lys Met Leu Phe Arg Val Thr Ile Asn Ser Trp Lys
1 5 10

<210> 828
<211> 543
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 828
atgaaacatc accatcacca tcatgcgatc gccatggtct tcacactcga agatttcgtt 60

ggggactggg aacagacagc cgcctacaac ctggaccaag tccttgaaca gggaggtgtg 120

tccagtttgc tgcagaatct cgccgtgtcc gtaactccga tccaaaggat tgtccggagc 180

ggtgaaaatg ccctgaagat cgacatccat gtcatcatcc cgtatgaagg tctgagcgcc 240

gaccaaatgg cccagatcga agaggtgttt aaggtggtgt accctgtgga tgatcatcac 300

tttaaggtga tcctgcccta tggcacactg gtaatcgacg gggttacgcc gaacatgctg 360

aactatttcg gacggccgtg ggaaggcatc gccgtgttcg acggcaaaaa gatcactgta 420

acagggaccc tgtggaacgg caacaaaatt atcgacgagc gcctgatcac ccccgacggc 480

tccatgctgt tccgagtaac catcaacggc gttagcggct ggcgcctgtt caagaagatc 540

agc 543

<210> 829
<211> 181
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 829

Met Lys His His His His His His Ala Ile Ala Met Val Phe Thr Leu
1 5 10 15

```
Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp
            20                  25                  30

Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala
        35                  40                  45

Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg Ser Gly Glu Asn Ala
    50                  55                  60

Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala
65                  70                  75                  80

Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val
                85                  90                  95

Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile
            100                 105                 110

Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Trp Glu
        115                 120                 125

Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr Leu
    130                 135                 140

Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly
145                 150                 155                 160

Ser Met Leu Phe Arg Val Thr Ile Asn Gly Val Ser Gly Trp Arg Leu
                165                 170                 175

Phe Lys Lys Ile Ser
            180
```

<210> 830
<211> 363
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 830

atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

taa 363

<210> 831
<211> 120
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 831

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly

115 120

<210> 832
<211> 396
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 832
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggctaa 396

<210> 833
<211> 131
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 833

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala

65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly
130

<210> 834
<211> 423
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 834
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

taa 423

<210> 835
<211> 140
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 835

```
Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1               5                  10                  15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
            20                  25                  30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
        35                  40                  45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
    50                  55                  60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65                  70                  75                  80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
                85                  90                  95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
            100                 105                 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
        115                 120                 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu
    130                 135                 140


<210>  836
<211>  432
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  836
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa       60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg      120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc      180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc      240
```

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggct aa 432

<210> 837
<211> 143
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 837

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly
130 135 140

<210> 838
<211> 483
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 838
atggtttccg tgagcggctg gcggctgttc aagaagatta gcttcacact cgacgatttc 60

gttggggact gggaacagac agccgcctac aacctggacc aagtccttga acagggaggt 120

gtgtccagtt tgctgcagaa tctcgccgtg tccgtaactc cgatcatgag gattgtccgg 180

agcggtgaaa atgccctgaa gatcgacatc catgtcatca tcccgtatga aggtctgagc 240

gccgaccaaa tggcccagat cgaagaggtg tttaaggtgg tgtaccctgt ggatgatcat 300

cactttaagg tgatcctgcc ctatggcaca ctggtaatcg acggggttac gccgaacaag 360

ctgaactatt tcggacggcc gtatgaaggc atcgccgtgt tcgacggcaa aaagatcact 420

accacaggga ccctgtggaa cggcaacaaa attatcgacg agcgcctgat caccccgac 480

taa 483

<210> 839
<211> 160
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 839

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Phe Thr
1 5 10 15

Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu
20 25 30

Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu
35 40 45

Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn
50 55 60

```
Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser
65                  70                  75                  80

Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro
                85                  90                  95

Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val
            100                 105                 110

Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr
        115                 120                 125

Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr
    130                 135                 140

Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145                 150                 155                 160
```

<210> 840
<211> 495
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 840
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggcagctc cggtttcaca 60

ctcgacgatt tcgttgggga ctgggaacag acagccgcct acaacctgga ccaagtcctt 120

gaacagggag gtgtgtccag tttgctgcag aatctcgccg tgtccgtaac tccgatcatg 180

aggattgtcc ggagcggtga aaatgccctg aagatcgaca tccatgtcat catcccgtat 240

gaaggtctga gcgccgacca aatggcccag atcgaagagg tgtttaaggt ggtgtaccct 300

gtggatgatc atcactttaa ggtgatcctg ccctatggca cactggtaat cgacggggtt 360

acgccgaaca agctgaacta tttcggacgg ccgtatgaag gcatcgccgt gttcgacggc 420

aaaaagatca ctaccacagg gaccctgtgg aacggcaaca aaattatcga cgagcgcctg 480

atcacccccg actaa 495

<210> 841

<211> 164
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 841

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser
1 5 10 15

Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
20 25 30

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
35 40 45

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
50 55 60

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
65 70 75 80

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
85 90 95

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
100 105 110

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
115 120 125

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
130 135 140

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
145 150 155 160

Ile Thr Pro Asp

<210> 842

<211> 507
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 842
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggctcgag cggtggctcg 60

agcggtttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 120

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 180

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 240

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 300

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 360

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 420

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 480

gacgagcgcc tgatcacccc cgactaa 507

<210> 843
<211> 168
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 843

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp
20 25 30

Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly
35 40 45

Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met
50 55 60

Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val

65 70 75 80

Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu
85 90 95

Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val
100 105 110

Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys
115 120 125

Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly
130 135 140

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
145 150 155 160

Asp Glu Arg Leu Ile Thr Pro Asp
165

<210> 844
<211> 519
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 844
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggctcgag cggtggctcg 60

agcggtggct cgagcggttt cacactcgac gatttcgttg gggactggga acagacagcc 120

gcctacaacc tggaccaagt ccttgaacag ggaggtgtgt ccagtttgct gcagaatctc 180

gccgtgtccg taactccgat catgaggatt gtccggagcg gtgaaaatgc cctgaagatc 240

gacatccatg tcatcatccc gtatgaaggt ctgagcgccg accaaatggc ccagatcgaa 300

gaggtgttta aggtggtgta ccctgtggat gatcatcact ttaaggtgat cctgccctat 360

ggcacactgg taatcgacgg ggttacgccg aacaagctga actatttcgg acggccgtat 420

gaaggcatcg ccgtgttcga cggcaaaaag atcactacca cagggaccct gtggaacggc 480

aacaaaatta tcgacgagcg cctgatcacc cccgactaa 519

<210> 845
<211> 172
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 845

```
Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser
1               5                  10                  15


Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe
            20                  25                  30


Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu
        35                  40                  45


Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val
    50                  55                  60


Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile
65                  70                  75                  80


Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met
                85                  90                  95


Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His
            100                 105                 110


His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val
        115                 120                 125


Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala
    130                 135                 140


Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly
145                 150                 155                 160


Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
                165                 170
```

<210> 846
<211> 531
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 846
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggctcgag cggtggctcg 60

agcggtggct cgagcggtgg ctcgagcggt ttcacactcg acgatttcgt tggggactgg 120

gaacagacag ccgcctacaa cctggaccaa gtccttgaac agggaggtgt gtccagtttg 180

ctgcagaatc tcgccgtgtc cgtaactccg atcatgagga ttgtccggag cggtgaaaat 240

gccctgaaga tcgacatcca tgtcatcatc ccgtatgaag gtctgagcgc cgaccaaatg 300

gcccagatcg aagaggtgtt taaggtggtg taccctgtgg atgatcatca ctttaaggtg 360

atcctgccct atggcacact ggtaatcgac ggggttacgc cgaacaagct gaactatttc 420

ggacggccgt atgaaggcat cgccgtgttc gacggcaaaa agatcactac cacagggacc 480

ctgtggaacg gcaacaaaat tatcgacgag cgcctgatca cccccgacta a 531

<210> 847
<211> 176
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 847

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Phe Thr
20 25 30

Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu
35 40 45

Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu
50 55 60

```
Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn
65                  70                  75                  80


Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser
                85                  90                  95


Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro
            100                 105                 110


Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val
        115                 120                 125


Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr
    130                 135                 140


Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr
145                 150                 155                 160


Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
                165                 170                 175


<210>  848
<211>  543
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  848
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggctcgag cggtggctcg       60

agcggtggct cgagcggtgg ctcgagcggt ggctcgagcg gtttcacact cgacgatttc      120

gttggggact gggaacagac agccgcctac aacctggacc aagtccttga acagggaggt      180

gtgtccagtt tgctgcagaa tctcgccgtg tccgtaactc cgatcatgag gattgtccgg      240

agcggtgaaa atgccctgaa gatcgacatc catgtcatca tcccgtatga aggtctgagc      300

gccgaccaaa tggcccagat cgaagaggtg tttaaggtgg tgtaccctgt ggatgatcat      360

cactttaagg tgatcctgcc ctatggcaca ctggtaatcg acggggttac gccgaacaag      420

ctgaactatt tcggacggcc gtatgaaggc atcgccgtgt tcgacggcaa aaagatcact      480
```

accacaggga ccctgtggaa cggcaacaaa attatcgacg agcgcctgat caccccgac 540

taa 543

<210> 849
<211> 180
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 849

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser
20 25 30

Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
35 40 45

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
50 55 60

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
65 70 75 80

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
85 90 95

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
100 105 110

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
115 120 125

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
130 135 140

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
145 150 155 160

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
165 170 175

Ile Thr Pro Asp
180

<210> 850
<211> 537
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 850
atggtgagcg gctggcggct gttcaagaag attagcggct cgagcggtgg ctcgagcggt 60

ggctcgagcg gtggctcgag cggtggctcg agcggtttca cactcgacga tttcgttggg 120

gactgggaac agacagccgc ctacaacctg gaccaagtcc ttgaacaggg aggtgtgtcc 180

agtttgctgc agaatctcgc cgtgtccgta actccgatca tgaggattgt ccggagcggt 240

gaaaatgccc tgaagatcga catccatgtc atcatcccgt atgaaggtct gagcgccgac 300

caaatggccc agatcgaaga ggtgtttaag gtggtgtacc ctgtggatga tcatcacttt 360

aaggtgatcc tgccctatgg cacactggta atcgacgggg ttacgccgaa caagctgaac 420

tatttcggac ggccgtatga aggcatcgcc gtgttcgacg gcaaaaagat cactaccaca 480

gggaccctgt ggaacggcaa caaaattatc gacgagcgcc tgatcacccc cgactaa 537

<210> 851
<211> 178
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 851

Met Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser Ser Gly
1 5 10 15

Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly
20 25 30

Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr
35 40 45

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
50 55 60

Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly
65 70 75 80

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
85 90 95

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
100 105 110

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr
115 120 125

Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg
130 135 140

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr
145 150 155 160

Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr
165 170 175

Pro Asp

<210> 852
<211> 486
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 852
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacgtttcc gtgagcggct ggcggctgtt caagaagatt 480

agctaa 486

<210> 853
<211> 161
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 853

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr

115 120 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile
145 150 155 160

Ser

<210> 854
<211> 498
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 854
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcg agcggtgttt ccgtgagcgg ctggcggctg 480

ttcaagaaga ttagctaa 498

<210> 855
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 855

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala

1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Ser Gly Val Ser Val Ser Gly Trp Arg Leu
145 150 155 160

Phe Lys Lys Ile Ser
165

<210> 856
<211> 510
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 856
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcg agcggtggct cgagcggtgt ttccgtgagc 480

ggctggcggc tgttcaagaa gattagctaa 510

<210> 857
<211> 169
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 857

Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Ser Gly Gly Ser Ser Gly Val Ser Val Ser
145 150 155 160

Gly Trp Arg Leu Phe Lys Lys Ile Ser
165

<210> 858
<211> 504
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 858
atggtcttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcg agcggtggct cgagcggtgt gagcggctgg 480

cggctgttca agaagattag ctaa 504

<210> 859
<211> 167
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 859

```
Met Val Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1               5                  10                  15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
            20                  25                  30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
        35                  40                  45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
    50                  55                  60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65                  70                  75                  80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140

Ile Thr Pro Asp Gly Ser Ser Gly Gly Ser Ser Gly Val Ser Gly Trp
145                 150                 155                 160

Arg Leu Phe Lys Lys Ile Ser
                165
```

<210> 860
<211> 501
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 860

atggtttccg tgagcggctg gcggctgttc aagaagatta gcttcacact cgacgatttc 60

gttggggact gggaacagac agccgcctac aacctggacc aagtccttga acagggaggt 120

gtgtccagtt tgctgcagaa tctcgccgtg tccgtaactc cgatcatgag gattgtccgg 180

agcggtgaaa atgccctgaa gatcgacatc catgtcatca tcccgtatga aggtctgagc 240

gccgaccaaa tggcccagat cgaagaggtg tttaaggtgg tgtaccctgt ggatgatcat 300

cactttaagg tgatcctgcc ctatggcaca ctggtaatcg acggggttac gccgaacaag 360

ctgaactatt tcggacggcc gtatgaaggc atcgccgtgt tcgacggcaa aaagatcact 420

accacaggga ccctgtggaa cggcaacaaa attatcgacg agcgcctgat caccccgac 480

catcaccatc accatcatta a 501

<210> 861
<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 861

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Phe Thr
1 5 10 15

Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu
20 25 30

Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu
35 40 45

Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn
50 55 60

Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser
65 70 75 80

Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro
85 90 95

Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val

100 105 110

Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr
115 120 125

Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr
130 135 140

Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155 160

His His His His His His
165

<210> 862
<211> 513
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 862
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggcagctc cggtttcaca 60

ctcgacgatt tcgttgggga ctgggaacag acagccgcct acaacctgga ccaagtcctt 120

gaacagggag gtgtgtccag tttgctgcag aatctcgccg tgtccgtaac tccgatcatg 180

aggattgtcc ggagcggtga aaatgccctg aagatcgaca tccatgtcat catcccgtat 240

gaaggtctga gcgccgacca aatggcccag atcgaagagg tgtttaaggt ggtgtaccct 300

gtggatgatc atcactttaa ggtgatcctg ccctatggca cactggtaat cgacggggtt 360

acgccgaaca agctgaacta tttcggacgg ccgtatgaag gcatcgccgt gttcgacggc 420

aaaaagatca ctaccacagg gaccctgtgg aacggcaaca aaattatcga cgagcgcctg 480

atcacccccg accatcacca tcaccatcat taa 513

<210> 863
<211> 170
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 863

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser
1 5 10 15

Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
20 25 30

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
35 40 45

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
50 55 60

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
65 70 75 80

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
85 90 95

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
100 105 110

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
115 120 125

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
130 135 140

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
145 150 155 160

Ile Thr Pro Asp His His His His His His
165 170

<210> 864
<211> 525
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 864
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggctcgag cggtggctcg 60

agcggtttca cactcgacga tttcgttggg gactgggaac agacagccgc ctacaacctg 120

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 180

actccgatca tgaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 240

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 300

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 360

atcgacgggg ttacgccgaa caagctgaac tatttcggac ggccgtatga aggcatcgcc 420

gtgttcgacg gcaaaaagat cactaccaca gggaccctgt ggaacggcaa caaaattatc 480

gacgagcgcc tgatcacccc cgaccatcac catcaccatc attaa 525

<210> 865
<211> 174
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 865

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp
20 25 30

Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly
35 40 45

Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met
50 55 60

Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val
65 70 75 80

Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu
85 90 95

```
Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val
            100                 105                 110

Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys
        115                 120                 125

Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly
    130                 135                 140

Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile
145                 150                 155                 160

Asp Glu Arg Leu Ile Thr Pro Asp His His His His His His
                165                 170
```

<210> 866
<211> 537
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 866
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggctcgag cggtggctcg 60

agcggtggct cgagcggttt cacactcgac gatttcgttg gggactggga acagacagcc 120

gcctacaacc tggaccaagt ccttgaacag ggaggtgtgt ccagtttgct gcagaatctc 180

gccgtgtccg taactccgat catgaggatt gtccggagcg gtgaaaatgc cctgaagatc 240

gacatccatg tcatcatccc gtatgaaggt ctgagcgccg accaaatggc ccagatcgaa 300

gaggtgttta aggtggtgta ccctgtggat gatcatcact ttaaggtgat cctgccctat 360

ggcacactgg taatcgacgg ggttacgccg aacaagctga actatttcgg acggccgtat 420

gaaggcatcg ccgtgttcga cggcaaaaag atcactacca cagggaccct gtggaacggc 480

aacaaaatta tcgacgagcg cctgatcacc cccgaccatc accatcacca tcattaa 537

<210> 867
<211> 178
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 867

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser
1   5   10   15

Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe
            20   25   30

Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu
        35   40   45

Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val
    50   55   60

Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile
65   70   75   80

Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met
                85   90   95

Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His
            100   105   110

His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val
        115   120   125

Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala
    130   135   140

Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly
145   150   155   160

Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp His His His His
                165   170   175

His His

<210> 868

<211> 549
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 868
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggctcgag cggtggctcg 60

agcggtggct cgagcggtgg ctcgagcggt ttcacactcg acgatttcgt tggggactgg 120

gaacagacag ccgcctacaa cctggaccaa gtccttgaac agggaggtgt gtccagtttg 180

ctgcagaatc tcgccgtgtc cgtaactccg atcatgagga ttgtccggag cggtgaaaat 240

gccctgaaga tcgacatcca tgtcatcatc ccgtatgaag gtctgagcgc cgaccaaatg 300

gcccagatcg aagaggtgtt taaggtggtg taccctgtgg atgatcatca ctttaaggtg 360

atcctgccct atggcacact ggtaatcgac ggggttacgc cgaacaagct gaactatttc 420

ggacggccgt atgaaggcat cgccgtgttc gacggcaaaa agatcactac cacagggacc 480

ctgtggaacg gcaacaaaat tatcgacgag cgcctgatca cccccgacca tcaccatcac 540

catcattaa 549

<210> 869
<211> 182
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 869

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser
1 5 10 15

Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Phe Thr
20 25 30

Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu
35 40 45

Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu
50 55 60

```
Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn
65                  70                  75                  80


Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser
                85                  90                  95


Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro
            100                 105                 110


Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val
        115                 120                 125


Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr
    130                 135                 140


Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr
145                 150                 155                 160


Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
                165                 170                 175


His His His His His His
            180


<210>  870
<211>  555
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  870
atggtgagcg gctggcggct gttcaagaag attagcggct cgagcggtgg ctcgagcggt       60

ggctcgagcg gtggctcgag cggtggctcg agcggtttca cactcgacga tttcgttggg      120

gactgggaac agacagccgc ctacaacctg gaccaagtcc ttgaacaggg aggtgtgtcc      180

agtttgctgc agaatctcgc cgtgtccgta actccgatca tgaggattgt ccggagcggt      240

gaaaatgccc tgaagatcga catccatgtc atcatcccgt atgaaggtct gagcgccgac      300

caaatggccc agatcgaaga ggtgtttaag gtggtgtacc ctgtggatga tcatcacttt      360
```

aaggtgatcc tgccctatgg cacactggta atcgacgggg ttacgccgaa caagctgaac 420

tatttcggac ggccgtatga aggcatcgcc gtgttcgacg gcaaaaagat cactaccaca 480

gggaccctgt ggaacggcaa caaaattatc gacgagcgcc tgatcacccc cgaccatcac 540

catcaccatc attaa 555

<210> 871
<211> 184
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 871

Met Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser Ser Gly
1 5 10 15

Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly
20 25 30

Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr
35 40 45

Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
50 55 60

Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly
65 70 75 80

Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
85 90 95

Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
100 105 110

Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr
115 120 125

Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg
130 135 140

Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr
145 150 155 160

Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr
165 170 175

Pro Asp His His His His His His
180

<210> 872
<211> 561
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 872
atggtttccg tgagcggctg gcggctgttc aagaagatta gcggctcgag cggtggctcg 60

agcggtggct cgagcggtgg ctcgagcggt ggctcgagcg gtttcacact cgacgatttc 120

gttggggact gggaacagac agccgcctac aacctggacc aagtccttga acagggaggt 180

gtgtccagtt tgctgcagaa tctcgccgtg tccgtaactc cgatcatgag gattgtccgg 240

agcggtgaaa atgccctgaa gatcgacatc catgtcatca tcccgtatga aggtctgagc 300

gccgaccaaa tggcccagat cgaagaggtg tttaaggtgg tgtaccctgt ggatgatcat 360

cactttaagg tgatcctgcc ctatggcaca ctggtaatcg acggggttac gccgaacaag 420

ctgaactatt tcggacggcc gtatgaaggc atcgccgtgt tcgacggcaa aaagatcact 480

accacaggga ccctgtggaa cggcaacaaa attatcgacg agcgcctgat caccccgac 540

catcaccatc accatcatta a 561

<210> 873
<211> 186
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 873

Met Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser

1 5 10 15

Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Gly Ser
20 25 30

Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
35 40 45

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
50 55 60

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg
65 70 75 80

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
85 90 95

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
100 105 110

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
115 120 125

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe
130 135 140

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
145 150 155 160

Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
165 170 175

Ile Thr Pro Asp His His His His His His
180 185

<210> 874
<211> 507
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 874
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgacgtttc cgtgagcggc 480

tggcggctgt tcaagaagat tagctaa 507

<210> 875
<211> 168
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 875

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Val Ser Val Ser Gly
145 150 155 160

Trp Arg Leu Phe Lys Lys Ile Ser
165

<210> 876
<211> 519
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 876
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgacggctc gagcggtgtt 480

tccgtgagcg gctggcggct gttcaagaag attagctaa 519

<210> 877
<211> 172
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 877

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1   5   10   15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
            20   25   30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
        35   40   45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
    50   55   60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65   70   75   80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
                85   90   95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
            100   105   110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
        115   120   125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
    130   135   140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Ser Gly Val
145   150   155   160

Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
                165   170

<210> 878
<211> 525
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 878
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgacggctc gagcggtggc 480

tcgagcggtg tgagcggctg gcggctgttc aagaagatta gctaa 525

<210> 879
<211> 174
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 879

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
            20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
        35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
    50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His

85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Ser Gly Gly
145 150 155 160

Ser Ser Gly Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
165 170

<210> 880
<211> 531
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 880
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgacggctc gagcggtggc 480

tcgagcggtg tttccgtgag cggctggcgg ctgttcaaga agattagcta a 531

<210> 881
<211> 176
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 881

```
Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1               5                  10                  15


Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
            20                  25                  30


Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
        35                  40                  45


Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
    50                  55                  60


Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65                  70                  75                  80


Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
                85                  90                  95


Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
            100                 105                 110


Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
        115                 120                 125


Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
    130                 135                 140


Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Ser Gly Gly
145                 150                 155                 160


Ser Ser Gly Val Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
                165                 170                 175
```

<210> 882
<211> 498
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic

<400> 882
atggtcttca cactcgaaga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgctgttcc gagtaaccat caacagccat 480

catcaccatc accactaa 498

<210> 883
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 883

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

```
Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His
145                 150                 155                 160

His His His His His
                165
```

<210> 884
<211> 501
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 884

```
atggtgagcg gctggcggct gttcaagaag attagccacc atcaccatca ccatcatcac       60

ttcacactcg acgatttcgt tggggactgg gaacagacag ccgcctacaa cctggaccaa      120

gtccttgaac agggaggtgt gtccagtttg ctgcagaatc tcgccgtgtc cgtaactccg      180

atcatgagga ttgtccggag cggtgaaaat gccctgaaga tcgacatcca tgtcatcatc      240

ccgtatgaag gtctgagcgc cgaccaaatg gcccagatcg aagaggtgtt taaggtggtg      300

taccctgtgg atgatcatca ctttaaggtg atcctgccct atggcacact ggtaatcgac      360

ggggttacgc cgaacaagct gaactatttc ggacggccgt atgaaggcat cgccgtgttc      420

gacggcaaaa agatcactac cacagggacc ctgtggaacg gcaacaaaat tatcgacgag      480

cgcctgatca cccccgacta a                                                501
```

<210> 885

<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 885

Met Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser His His His His
1 5 10 15

His His His His Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln
20 25 30

Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser
35 40 45

Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile
50 55 60

Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile
65 70 75 80

Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val
85 90 95

Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu
100 105 110

Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn
115 120 125

Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys
130 135 140

Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu
145 150 155 160

Arg Leu Ile Thr Pro Asp
165

<210> 886

<211> 510
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 886
atgaaacatc accatcacca tcatgtgagc ggctggcggc tgttcaagaa gattagcggc 60

agctccggtt tcacactcga cgatttcgtt ggggactggg aacagacagc cgcctacaac 120

ctggaccaag tccttgaaca gggaggtgtg tccagtttgc tgcagaatct cgccgtgtcc 180

gtaactccga tcatgaggat tgtccggagc ggtgaaaatg ccctgaagat cgacatccat 240

gtcatcatcc cgtatgaagg tctgagcgcc gaccaaatgg cccagatcga agaggtgttt 300

aaggtggtgt accctgtgga tgatcatcac tttaaggtga tcctgcccta tggcacactg 360

gtaatcgacg gggttacgcc gaacaagctg aactatttcg gacggccgta tgaaggcatc 420

gccgtgttcg acggcaaaaa gatcactacc acagggaccc tgtggaacgg caacaaaatt 480

atcgacgagc gcctgatcac ccccgactaa 510

<210> 887
<211> 169
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 887

Met Lys His His His His His His Val Ser Gly Trp Arg Leu Phe Lys
1 5 10 15

Lys Ile Ser Gly Ser Ser Gly Phe Thr Leu Asp Asp Phe Val Gly Asp
20 25 30

Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly
35 40 45

Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile
50 55 60

Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His

65 70 75 80

Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile
85 90 95

Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys
100 105 110

Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn
115 120 125

Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp
130 135 140

Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile
145 150 155 160

Ile Asp Glu Arg Leu Ile Thr Pro Asp
165

<210> 888
<211> 507
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 888
atggtgagcg gctggcggct gttcaagaag attagcggca gctccggttt cacactcgac 60

gatttcgttg gggactggga acagacagcc gcctacaacc tggaccaagt ccttgaacag 120

ggaggtgtgt ccagtttgct gcagaatctc gccgtgtccg taactccgat catgaggatt 180

gtccggagcg gtgaaaatgc cctgaagatc gacatccatg tcatcatccc gtatgaaggt 240

ctgagcgccg accaaatggc ccagatcgaa gaggtgttta aggtggtgta ccctgtggat 300

gatcatcact ttaaggtgat cctgccctat ggcacactgg taatcgacgg ggttacgccg 360

aacaagctga actatttcgg acggccgtat gaaggcatcg ccgtgttcga cggcaaaaag 420

atcactacca cagggaccct gtggaacggc aacaaaatta tcgacgagcg cctgatcacc 480

cccgaccatc accatcacca tcattaa 507

<210> 889
<211> 168
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 889

```
Met Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser Ser Gly
1               5                  10                  15


Phe Thr Leu Asp Asp Phe Val Gly Asp Trp Glu Gln Thr Ala Ala Tyr
            20                  25                  30


Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln
        35                  40                  45


Asn Leu Ala Val Ser Val Thr Pro Ile Met Arg Ile Val Arg Ser Gly
    50                  55                  60


Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly
65                  70                  75                  80


Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val
                85                  90                  95


Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly Thr
            100                 105                 110


Leu Val Ile Asp Gly Val Thr Pro Asn Lys Leu Asn Tyr Phe Gly Arg
        115                 120                 125


Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Thr Thr
    130                 135                 140


Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu Ile Thr
145                 150                 155                 160


Pro Asp His His His His His His
                165
```

<210> 890
<211> 498
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 890
atggtcttca cactcgaaga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacgagaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgctgttcc gagtaaccat caacagccat 480

catcaccatc accactaa 498

<210> 891
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 891

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

```
Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65                  70                  75                  80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Val Thr Gly Thr Leu Trp Asn Glu Asn Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His
145                 150                 155                 160

His His His His His
                165
```

<210> 892
<211> 498
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 892

```
atggtcttca cactcgaaga tttcgttggg gactgggaac agacagccgc ctacaacctg       60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta      120

actccgatcc aaaggattgt ccggagcggt gaaaatgccc tgaagatcga catccatgtc      180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag      240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta      300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc      360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcgt taaaattatc      420

gacgagcgcc tgatcacccc cgacggctcc atgctgttcc gagtaaccat caacagccat      480
```

catcaccatc accactaa 498

<210> 893
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 893

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His
145 150 155 160

His His His His His

165

<210> 894
<211> 498
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 894
atggtcttca cactcgaaga tttcgttggg gactgggaac agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggatggt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgctgttcc gagtaaccat caacagccat 480

catcaccatc accactaa 498

<210> 895
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 895

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr

50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His
145 150 155 160

His His His His His
165

<210> 896
<211> 498
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 896
atggtcttca cactcgaaga tttcgttggg gactggaagc agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggatggt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacggcaa caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgctgttcc gagtaaccat caacagccat 480

catcaccatc accactaa 498

<210> 897
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 897

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His
145 150 155 160

His His His His His
165

<210> 898
<211> 498
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 898
atggtcttca cactcgaaga tttcgttggg gactggaagc agacagccgc ctacaacctg 60

gaccaagtcc ttgaacaggg aggtgtgtcc agtttgctgc agaatctcgc cgtgtccgta 120

actccgatcc aaaggatggt ccggagcggt gaaaatgccc tgaagatcga catccatgtc 180

atcatcccgt atgaaggtct gagcgccgac caaatggccc agatcgaaga ggtgtttaag 240

gtggtgtacc ctgtggatga tcatcacttt aaggtgatcc tgccctatgg cacactggta 300

atcgacgggg ttacgccgaa catgctgaac tatttcggac ggccgtatga aggcatcgcc 360

gtgttcgacg gcaaaaagat cactgtaaca gggaccctgt ggaacgacgt caaaattatc 420

gacgagcgcc tgatcacccc cgacggctcc atgctgttcc gagtaaccat caacagccat 480

catcaccatc accactaa 498

<210> 899
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 899

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg
35 40 45

```
Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
    50                  55                  60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65                  70                  75                  80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
                85                  90                  95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
            100                 105                 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
        115                 120                 125

Val Thr Gly Thr Leu Trp Asn Asp Val Lys Ile Ile Asp Glu Arg Leu
    130                 135                 140

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser His
145                 150                 155                 160

His His His His His
                165
```

<210> 900
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 900

```
Gly Arg Met Leu Phe Arg Val Thr Ile Asn Ser Trp Arg
1               5                   10
```

<210> 901
<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 901

Val Ser Gly Trp Arg Leu Phe Arg Arg Ile Ser
1 5 10

<210> 902
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 902

Gly Arg Leu Leu Phe Val Val Val Ile Glu Arg Tyr Arg
1 5 10

<210> 903
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 903

Val Ser Gly Trp Arg Leu Phe Arg Arg Ile Ser Cys
1 5 10

<210> 904
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 904

Gly Arg Met Leu Phe Arg Val Thr Ile Asn Ser Trp Arg Cys
1 5 10

<210> 905
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 905

Gly Arg Leu Leu Phe Thr Val Thr Ile Glu Arg Tyr Arg Cys
1   5   10

<210> 906
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 906

Gly Lys Leu Leu Phe Val Val Val Ile Glu Lys Tyr Lys
1   5   10

<210> 907
<211> 21
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 907

Gly Lys Leu Leu Phe Val Thr Ile Glu Lys Val Ser Gly Trp Arg Leu
1   5   10   15

Phe Lys Lys Ile Ser
            20

<210> 908
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 908
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa   60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg   120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cacccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgactaa 468

<210> 909
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 909

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 910
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 910
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcgagaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgactaa 468

<210> 911
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 911

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr

```
        35                  40                  45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
    50                  55                  60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65                  70                  75                  80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
                85                  90                  95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
            100                 105                 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
        115                 120                 125

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
    130                 135                 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145                 150                 155
```

<210> 912
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 912

```
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa       60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg      120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc      180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc      240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc      300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga      360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg      420
```

cctaacggca acaaaattat cgacgagcgc ctgatcaccc ccgactaa 468

<210> 913
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 913

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Pro Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 914

<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 914
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcgatc ccgactaa 468

<210> 915
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 915

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp
145 150 155

<210> 916
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 916
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cggccgtatg aaggcatcgc cgtgttcgac ggcaaaaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccg atgactaa 468

<210> 917
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 917

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Lys Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Asp Asp
145 150 155

<210> 918
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 918
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cacccgtatg aaggcatcgc cgtgttcgac ggcgagaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcaccc ccgactaa 468

<210> 919
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 919

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val

115 120 125

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
145 150 155

<210> 920
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 920
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cacccgtatg aaggcatcgc cgtgttcgac ggcgagaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcgatc ccgactaa 468

<210> 921
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 921

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

```
Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
        35                  40                  45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
    50                  55                  60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65                  70                  75                  80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
                85                  90                  95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
            100                 105                 110

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val
        115                 120                 125

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
    130                 135                 140

Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp
145                 150                 155


<210>  922
<211>  468
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  922
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa       60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg      120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc      180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc      240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc      300

ctgccctatg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga      360
```

cacccgtatg aaggcatcgc cgtgttcgac ggcgagaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcgatg atgactaa 468

<210> 923
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 923

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Asp Asp Asp
145 150 155

<210> 924
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 924
atgaaacatc accatcacca tcatgatttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgcccatcg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cacccgtatg aaggcatcgc cgtgttcgac ggcgagaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcgatc ccgactaa 468

<210> 925
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 925

Met Lys His His His His His His Asp Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala

65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Ile Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val
115 120 125

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp
145 150 155

<210> 926
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 926
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgcccatcg gcacactggt aatcgacggg gagacgccga acaagctgaa ctatttcgga 360

cacccgtatg aaggcatcgc cgtgttcgac ggcgagaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcgatc ccgactaa 468

<210> 927
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 927

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1   5   10   15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
            20   25   30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
        35   40   45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
    50   55   60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65   70   75   80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
                85   90   95

Phe Lys Val Ile Leu Pro Ile Gly Thr Leu Val Ile Asp Gly Glu Thr
            100   105   110

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Val
        115   120   125

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
    130   135   140

Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp
145   150   155

<210> 928
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 928

atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgcccatcg gcacactggt aatcgacggg gttacgccga acaagctgaa ctatttcgga 360

cacccgtatg aaggcatcgc cgatttcgac ggcgagaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcgatc ccgactaa 468

<210> 929
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 929

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Ile Gly Thr Leu Val Ile Asp Gly Val Thr
100 105 110

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Asp
115 120 125

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
130 135 140

Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp
145 150 155

<210> 930
<211> 468
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 930
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa 60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg 120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc 180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc 240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc 300

ctgcccatcg gcacactggt aatcgacggg gagacgccga acaagctgaa ctatttcgga 360

cacccgtatg aaggcatcgc cgatttcgac ggcgagaaga tcactaccac agggaccctg 420

tggaacggca acaaaattat cgacgagcgc ctgatcgatc ccgactaa 468

<210> 931
<211> 155
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 931

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu

```
            20                  25                  30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
        35                  40                  45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
    50                  55                  60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65                  70                  75                  80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
                85                  90                  95

Phe Lys Val Ile Leu Pro Ile Gly Thr Leu Val Ile Asp Gly Glu Thr
            100                 105                 110

Pro Asn Lys Leu Asn Tyr Phe Gly His Pro Tyr Glu Gly Ile Ala Asp
        115                 120                 125

Phe Asp Gly Glu Lys Ile Thr Thr Thr Gly Thr Leu Trp Asn Gly Asn
    130                 135                 140

Lys Ile Ile Asp Glu Arg Leu Ile Asp Pro Asp
145                 150                 155


<210>  932
<211>  327
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  932
atgaaacatc accatcacca tcatgtcttc acactcgacg atttcgttgg ggactgggaa       60

cagacagccg cctacaacct ggaccaagtc cttgaacagg gaggtgtgtc cagtttgctg      120

cagaatctcg ccgtgtccgt aactccgatc atgaggattg tccggagcgg tgaaaatgcc      180

ctgaagatcg acatccatgt catcatcccg tatgaaggtc tgagcgccga ccaaatggcc      240

cagatcgaag aggtgtttaa ggtggtgtac cctgtggatg atcatcactt taaggtgatc      300
```

ctgccctatg gcacactggt aatcgac 327

<210> 933
<211> 109
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 933

Met Lys His His His His His His Val Phe Thr Leu Asp Asp Phe Val
1 5 10 15

Gly Asp Trp Glu Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu
20 25 30

Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr
35 40 45

Pro Ile Met Arg Ile Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp
50 55 60

Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala
65 70 75 80

Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His
85 90 95

Phe Lys Val Ile Leu Pro Tyr Gly Thr Leu Val Ile Asp
100 105

<210> 934
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 934

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20   25   30

Gln Asn Leu Ala Val Ser Val Thr Pro Thr Gln Arg Met Val Arg Ser
        35   40   45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    50   55   60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65   70   75   80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85   90   95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100   105   110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115   120   125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130   135   140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145   150   155

<210> 935
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 935

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Asn Gln Thr Ala Ala
1   5   10   15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20   25   30

```
Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
        35                  40                  45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    50                  55                  60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65                  70                  75                  80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85                  90                  95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115                 120                 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145                 150                 155


<210>  936
<211>  474
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  936
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac       60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact      120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc      180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg      240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc      300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg      360
```

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccaag taaccatcaa cagc 474

<210> 937
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 937

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 938
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 938
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgccaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 939
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 939

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val

65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 940
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 940
gtcttcacac tcgaagattt cgttggggac tggaaccaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcga cagc 474

<210> 941
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 941

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Asn Gln Thr Ala Ala
1   5   10   15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20   25   30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
        35   40   45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    50   55   60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65   70   75   80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85   90   95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100   105   110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115   120   125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130   135   140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asp Ser
145   150   155

<210> 942
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 942

gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccaagtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 943
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 943

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Asn Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asp Ser
145 150 155

<210> 944
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 944
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccaagtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 945
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 945

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu

20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile Gln Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 946
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 946
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga aactcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccacgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccgtcaa cagc 474

<210> 947
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 947

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Lys Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Val Asn Ser

145 150 155

<210> 948
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 948
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga aggtcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgcaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 949
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 949

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Val Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

```
Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65                  70                  75                  80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85                  90                  95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Ala
        115                 120                 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn
145                 150                 155
```

<210> 950
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 950
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggttggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgccaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 951
<211> 158
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 951

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Leu Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 952
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 952
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgcg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 953
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 953

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly

100 105 110

Arg Pro Tyr Glu Gly Ile Ala Ala Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 954
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 954
gtcttcacac tcgaagattt cgttggggcc tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccaagtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 955
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 955

Val Phe Thr Leu Glu Asp Phe Val Gly Ala Trp Lys Gln Thr Ala Ala
1 5 10 15

```
Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20                  25                  30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
        35                  40                  45

Gly Glu Asn Ala Leu Lys Ile Asp Ile Gln Val Ile Ile Pro Tyr Glu
    50                  55                  60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65                  70                  75                  80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85                  90                  95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115                 120                 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145                 150                 155
```

<210> 956
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 956
gtcttcacac tcgaagattt cgtaggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcagg atctcgccgt gtccgtaact 120

ccgacccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 957
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 957

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asp Leu Ala Val Ser Val Thr Pro Thr Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 958
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 958
gtcttcacac tcgaagattt cgttggggac tggagacaga ctgccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcgga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccaccaa cagc 474

<210> 959
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 959

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Arg Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Arg Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu

50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Thr Asn Ser
145 150 155

<210> 960
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 960
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggtc 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacac gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ctcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 961

<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 961

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Val Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Thr Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Leu Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 962
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 962
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgaaccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg gaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtgacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 963
<211> 157
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 963

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1   5   10   15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20   25   30

Gln Asn Leu Ala Val Ser Val Thr Pro Asn Gln Arg Met Val Arg Ser
        35   40   45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Gly
    50   55   60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65   70   75   80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85   90   95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn
145 150 155

<210> 964
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 964
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga acctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa agtgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 965
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 965

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala

1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Ser Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 966
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 966
gtcttcacac acgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc cctatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 967
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 967

Val Phe Thr His Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
1 5 10 15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
20 25 30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
35 40 45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
50 55 60

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65 70 75 80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
85 90 95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
100 105 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
115 120 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile

130 135 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145 150 155

<210> 968
<211> 474
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 968

Gly Thr Cys Thr Thr Cys Ala Cys Ala Cys Thr Cys Gly Ala Ala Gly
1 5 10 15

Ala Thr Thr Thr Cys Gly Thr Thr Gly Gly Gly Gly Ala Cys Thr Gly
20 25 30

Gly Ala Ala Ala Cys Ala Gly Thr Cys Ala Gly Cys Cys Gly Cys Cys
35 40 45

Thr Ala Cys Ala Ala Cys Cys Thr Gly Gly Ala Cys Cys Ala Ala Gly
50 55 60

Thr Cys Cys Thr Thr Gly Ala Ala Cys Ala Gly Gly Gly Ala Gly Gly
65 70 75 80

Thr Gly Thr Gly Thr Cys Cys Ala Gly Thr Thr Thr Gly Cys Thr Gly
85 90 95

Cys Ala Gly Ala Ala Thr Cys Thr Cys Gly Cys Cys Gly Thr Gly Thr
100 105 110

Cys Cys Gly Thr Ala Ala Cys Thr Cys Cys Gly Ala Thr Cys Cys Ala
115 120 125

Ala Ala Gly Gly Ala Thr Gly Gly Thr Cys Cys Gly Gly Ala Gly Cys
130 135 140

Gly Gly Thr Gly Ala Ala Ala Ala Thr Gly Cys Cys Cys Thr Gly Ala

```
145                 150                 155                 160

Ala Gly Ala Thr Cys Gly Ala Cys Ala Thr Cys Cys Ala Thr Gly Thr
                165                 170                 175

Cys Ala Thr Cys Ala Thr Cys Cys Cys Gly Thr Ala Thr Gly Ala Ala
            180                 185                 190

Gly Gly Thr Cys Thr Gly Ala Gly Cys Gly Cys Cys Gly Ala Cys Cys
        195                 200                 205

Ala Ala Ala Thr Gly Gly Cys Cys Cys Ala Gly Ala Thr Cys Gly Ala
    210                 215                 220

Ala Gly Ala Gly Gly Thr Gly Thr Thr Thr Ala Ala Gly Gly Thr Gly
225                 230                 235                 240

Gly Thr Gly Thr Ala Cys Cys Cys Thr Gly Thr Gly Gly Ala Thr Gly
                245                 250                 255

Ala Thr Cys Ala Thr Cys Ala Cys Thr Thr Thr Ala Ala Gly Gly Thr
            260                 265                 270

Gly Ala Thr Cys Cys Thr Gly Cys Cys Cys Thr Ala Thr Gly Gly Cys
        275                 280                 285

Ala Cys Ala Cys Thr Gly Gly Thr Ala Ala Thr Cys Gly Ala Cys Gly
    290                 295                 300

Gly Gly Gly Thr Thr Ala Cys Gly Cys Cys Gly Ala Ala Cys Ala Thr
305                 310                 315                 320

Gly Cys Thr Gly Ala Ala Cys Thr Ala Thr Thr Thr Cys Gly Gly Ala
                325                 330                 335

Cys Gly Gly Cys Cys Gly Thr Ala Thr Gly Ala Ala Gly Gly Cys Ala
            340                 345                 350

Thr Cys Gly Cys Cys Gly Thr Gly Thr Thr Cys Gly Ala Cys Gly Gly
        355                 360                 365
```

```
Cys Ala Ala Ala Ala Ala Gly Ala Thr Cys Ala Cys Thr Gly Thr Ala
    370                 375                 380

Ala Cys Ala Gly Gly Gly Ala Cys Cys Cys Thr Gly Thr Gly Gly Ala
385                 390                 395                 400

Ala Cys Gly Gly Cys Gly Thr Cys Ala Ala Ala Ala Thr Thr Ala Thr
                405                 410                 415

Cys Gly Ala Cys Gly Ala Gly Cys Gly Cys Cys Thr Gly Ala Thr Cys
            420                 425                 430

Ala Cys Cys Cys Cys Cys Gly Ala Cys Gly Gly Cys Thr Cys Cys Ala
        435                 440                 445

Thr Gly Thr Cys Cys Thr Thr Cys Cys Gly Ala Gly Thr Ala Ala Cys
    450                 455                 460

Cys Ala Thr Cys Ala Ala Cys Ala Gly Cys
465                 470
```

<210> 969
<211> 158
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 969

```
Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Ser Ala Ala
1               5                   10                  15

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            20                  25                  30

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
        35                  40                  45

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    50                  55                  60
```

```
Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
65                  70                  75                  80

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr Gly
                85                  90                  95

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            100                 105                 110

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        115                 120                 125

Thr Gly Thr Leu Trp Asn Gly Val Lys Ile Ile Asp Glu Arg Leu Ile
    130                 135                 140

Thr Pro Asp Gly Ser Met Ser Phe Arg Val Thr Ile Asn Ser
145                 150                 155
```

<210> 970
<211> 474
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 970
gtcttcacac tcgaagattt cgttggggac tggaaacaga cagccgccta caacctggac 60

caagtccttg aacagggagg tgtgtccagt ttgctgcaga atctcgccgt gtccgtaact 120

ccgatccaaa ggatggtccg gagcggtgaa aatgccctga agatcgacat ccatgtcatc 180

atcccgtatg aaggtctgag cgccgaccaa atggcccaga tcgaagaggt gtttaaggtg 240

gtgtaccctg tggatgatca tcactttaag gtgatcctgc actatggcac actggtaatc 300

gacggggtta cgccgaacat gctgaactat ttcggacggc cgtatgaagg catcgccgtg 360

ttcgacggca aaaagatcac tgtaacaggg accctgtgga acggcgtcaa aattatcgac 420

gagcgcctga tcacccccga cggctccatg tccttccgag taaccatcaa cagc 474

<210> 971
<211> 906
<212> DNA

<213> Artificial sequence

<220>
<223> Synthetic

<400> 971
atgaactcct tctccacaag cgccttcggt ccagttgcct tctccctggg cctgctcctg 60

gtgttgcctg ctgccttccc tgccccagac tacaaagacc atgacggtga ttataaagat 120

catgacatcg attacaagga tgacgatgac aagagagtcc aacctactga atctattgtt 180

agatttccta atattactaa tctttgtcct tttggcgagg tgttcaatgc caccagattc 240

gcctctgtgt acgcctggaa ccggaagcgg atcagcaatt gcgtggccga ctactccgtg 300

ctgtacaact ccgccagctt cagcaccttc aagtgctacg gcgtgtcccc taccaagctg 360

aacgacctgt gcttcacaaa cgtgtacgcc gacagcttcg tgatccgggg agatgaagtg 420

cggcagattg cccctggaca gacaggcaag atcgccgact acaactacaa gctgcccgac 480

gacttcaccg gctgtgtgat tgcctggaac agcaacaacc tggactccaa agtcggcggc 540

aactacaatt acctgtaccg gctgttccgg aagtccaatc tgaagccctt cgagcgggac 600

atctccaccg agatctatca ggccggcagc acccttgta acggcgtgga aggcttcaac 660

tgctacttcc cactgcagtc ctacggcttt cagcccacaa atggcgtggg ctatcagccc 720

tacagagtgg tggtgctgag cttcgaactg ctgcatgccc ctgccacagt gtgcggccct 780

aagaaaagca ccaatctcgt gaagaacaaa tgcgtgaact tcgggagctc cggtggtggc 840

gggagcggag gtggaggctc gagcggtggc aagctcctgt tcacggtaac catcgagaag 900

tataag 906

<210> 972
<211> 302
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 972

Met Asn Ser Phe Ser Thr Ser Ala Phe Gly Pro Val Ala Phe Ser Leu
1 5 10 15

Gly Leu Leu Leu Val Leu Pro Ala Ala Phe Pro Ala Pro Asp Tyr Lys

```
            20                  25                  30

Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp
        35                  40                  45

Asp Asp Lys Arg Val Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn
    50                  55                  60

Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe
65                  70                  75                  80

Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala
                85                  90                  95

Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys
            100                 105                 110

Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val
        115                 120                 125

Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala
    130                 135                 140

Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp
145                 150                 155                 160

Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser
                165                 170                 175

Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser
            180                 185                 190

Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala
        195                 200                 205

Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro
    210                 215                 220

Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro
225                 230                 235                 240
```

```
Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr
                245                 250                 255

Val Cys Gly Pro Lys Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val
            260                 265                 270

Asn Phe Gly Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
        275                 280                 285

Gly Gly Lys Leu Leu Phe Thr Val Thr Ile Glu Lys Tyr Lys
    290                 295                 300
```

<210> 973
<211> 906
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 973

```
atgaactcct tctccacaag cgccttcggt ccagttgcct tctccctggg cctgctcctg       60

gtgttgcctg ctgccttccc tgccccagtc tccgtgagcg gctggcggct gttcaagaag      120

attagcggga gctccggtgg tggcgggagc ggaggtggag gctcgagcgg tagagtccaa      180

cctactgaat ctattgttag atttcctaat attactaatc tttgtccttt tggcgaggtg      240

ttcaatgcca ccagattcgc ctctgtgtac gcctggaacc ggaagcggat cagcaattgc      300

gtggccgact actccgtgct gtacaactcc gccagcttca gcaccttcaa gtgctacggc      360

gtgtccccta ccaagctgaa cgacctgtgc ttcacaaacg tgtacgccga cagcttcgtg      420

atccggggag atgaagtgcg gcagattgcc cctggacaga caggcaagat cgccgactac      480

aactacaagc tgcccgacga cttcaccggc tgtgtgattg cctggaacag caacaacctg      540

gactccaaag tcggcggcaa ctacaattac ctgtaccggc tgttccggaa gtccaatctg      600

aagcccttcg agcgggacat ctccaccgag atctatcagg ccggcagcac cccttgtaac      660

ggcgtggaag gcttcaactg ctacttccca ctgcagtcct acggctttca gcccacaaat      720

ggcgtgggct atcagcccta cagagtggtg gtgctgagct tcgaactgct gcatgccccc      780

gccacagtgt gcggccctaa gaaaagcacc aatctcgtga agaacaaatg cgtgaacttc      840
```

gactacaaag accatgacgg tgattataaa gatcatgaca tcgattacaa ggatgacgat 900

gacaag 906

<210> 974
<211> 302
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 974

Met Asn Ser Phe Ser Thr Ser Ala Phe Gly Pro Val Ala Phe Ser Leu
1 5 10 15

Gly Leu Leu Leu Val Leu Pro Ala Ala Phe Pro Ala Pro Val Ser Val
20 25 30

Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser Gly Ser Ser Gly Gly Gly
35 40 45

Gly Ser Gly Gly Gly Gly Ser Ser Gly Arg Val Gln Pro Thr Glu Ser
50 55 60

Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val
65 70 75 80

Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg
85 90 95

Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser
100 105 110

Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp
115 120 125

Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp
130 135 140

Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr
145 150 155 160

Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn
165 170 175

Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr
180 185 190

Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser
195 200 205

Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly
210 215 220

Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn
225 230 235 240

Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu
245 250 255

Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu
260 265 270

Val Lys Asn Lys Cys Val Asn Phe Asp Tyr Lys Asp His Asp Gly Asp
275 280 285

Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys
290 295 300

<210> 975
<211> 906
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 975
atgaactcct tctccacaag cgccttcggt ccagttgcct tctccctggg cctgctcctg 60

gtgttgcctg ctgccttccc tgccccaggc aagctcctgt tcacggtaac catcgagaag 120

tataagggga gctccggtgg tggcgggagc ggaggtggag gctcgagcgg tagagtccaa 180

cctactgaat ctattgttag atttcctaat attactaatc tttgtccttt tggcgaggtg 240

ttcaatgcca ccagattcgc ctctgtgtac gcctggaacc ggaagcggat cagcaattgc 300

gtggccgact actccgtgct gtacaactcc gccagcttca gcaccttcaa gtgctacggc 360

gtgtccccta ccaagctgaa cgacctgtgc ttcacaaacg tgtacgccga cagcttcgtg 420

atccggggag atgaagtgcg gcagattgcc cctggacaga caggcaagat cgccgactac 480

aactacaagc tgcccgacga cttcaccggc tgtgtgattg cctggaacag caacaacctg 540

gactccaaag tcggcggcaa ctacaattac ctgtaccggc tgttccggaa gtccaatctg 600

aagcccttcg agcgggacat ctccaccgag atctatcagg ccggcagcac cccttgtaac 660

ggcgtggaag gcttcaactg ctacttccca ctgcagtcct acggctttca gcccacaaat 720

ggcgtgggct atcagcccta cagagtggtg gtgctgagct tcgaactgct gcatgcccct 780

gccacagtgt gcggccctaa gaaaagcacc aatctcgtga agaacaaatg cgtgaacttc 840

gactacaaag accatgacgg tgattataaa gatcatgaca tcgattacaa ggatgacgat 900

gacaag 906

<210> 976
<211> 300
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 976

Met Asn Ser Phe Ser Thr Ser Ala Phe Gly Pro Val Ala Phe Ser Leu
1 5 10 15

Gly Leu Leu Leu Val Leu Pro Ala Ala Phe Pro Ala Pro Gly Lys Leu
20 25 30

Leu Phe Thr Val Thr Ile Glu Lys Tyr Lys Gly Ser Ser Gly Gly Gly
35 40 45

Gly Ser Gly Gly Gly Gly Ser Ser Gly Arg Val Gln Pro Thr Glu Ser
50 55 60

Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val
65 70 75 80

```
Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg
                85                  90                  95

Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser
            100                 105                 110

Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp
        115                 120                 125

Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp
    130                 135                 140

Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr
145                 150                 155                 160

Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn
                165                 170                 175

Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr
            180                 185                 190

Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser
        195                 200                 205

Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly
    210                 215                 220

Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn
225                 230                 235                 240

Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu
                245                 250                 255

Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn Leu
            260                 265                 270

Val Lys Asn Lys Cys Val Asn Phe Asp Tyr Lys Asp His Asp Gly Asp
        275                 280                 285

Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp
```

290 295 300

<210> 977
<211> 1914
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 977
atggactcgg gccgggactt cctgaccctg cacggcctac aggatgatga ggatctacag 60

gcgctgctga agggcagcca gctcctgaag gtgaagtcca gctcatggag gagagagcgg 120

ttctacaagt tgcaggagga ctgcaagacc atctggcagg agtcccgcaa ggtcatgcgg 180

accccggagt cccagctgtt ctccatcgag gacattcagg aggtgcgaat ggggcaccgc 240

acggagggtc tggagaagtt cgcccgtgat gtgcccgagg accgctgctt ctccattgtc 300

ttcaaggacc agcgcaatac actagacctc atcgccccat cgccagctga tgcccagcac 360

tgggtgctgg ggctgcacaa gatcatccac cactcaggct ccatggacca gcgtcagaag 420

ctacagcact ggattcactc ctgcttgcga aaagctgaca aaaacaagga caacaagatg 480

agcttcaagg agctgcagaa cttcctgaag gagctcaaca tccagctcgg ctcttcaggc 540

gcagaaatcg gtactggctt tccattcgac ccccattatg tggaagtcct gggcgagcgc 600

atgcactacg tcgatgttgg tccgcgcgat ggcacccctg tgctgttcct gcacggtaac 660

ccgacctcct cctacgtgtg gcgcaacatc atcccgcatg ttgcaccgac ccatcgctgc 720

attgctccag acctgatcgg tatgggcaaa tccgacaaac cagacctggg ttatttcttc 780

gacgaccacg tccgcttcat ggatgccttc atcgaagccc tgggtctgga agaggtcgtc 840

ctggtcattc acgactgggg ctccgctctg ggtttccact gggccaagcg caatccagag 900

cgcgtcaaag gtattgcatt tatggagttc atccgcccta tcccgacctg ggacgaatgg 960

ccagaatttg cccgcgagac cttccaggcc ttccgcacca ccgacgtcgg ccgcaagctg 1020

atcatcgatc agaacgtttt tatcgagggt acgctgccga tgggtgtcgt ccgcccgctg 1080

actgaagtcg agatggacca ttaccgcgag ccgttcctga atcctgttga ccgcgagcca 1140

ctgtggcgct tcccaaacga gctgccaatc gccggtgagc cagcgaacat cgtcgcgctg 1200

gtcgaagaat acatggactg gctgcaccag tcccctgtcc cgaagctgct gttctggggc 1260

accccaggcg ttctgatccc accggccgaa gccgctcgcc tggccaaaag cctgcctaac 1320

tgcaaggctg tggacatcgg cccgggtctg aatctgctgc aagaagacaa cccggacctg 1380

atcggcagcg agatcgcgcg ctggctgtcg acgctggaga tttccggcgg ctcgagcggt 1440

gtcttcactt tggaagattt cgtgggtgac tggaaacaga ctgcagcata caacttagac 1500

caagtgttgg aacagggtgg agttagttcg cttttacaga atttggcagt tagtgttacg 1560

cctatacaac gtatggtgag atcgggagaa aatgcattaa aaatagacat acatgtgata 1620

ataccttatg aaggattatc ggcagaccaa atggcacaga tagaagaggt ttttaaagtt 1680

gtttacccag ttgatgatca tcactttaaa gttatattac actatggaac tttagttata 1740

gacggtgtga ctcctaacat gttaaactat ttcggtagac cttatgaagg aatagcagtt 1800

ttcgacggaa agaaaataac ggttactggt acgttatgga acggagccaa gataatagac 1860

gagagattaa taacgccaga cggaagtatg agtttccagg ttacgataaa ctcg 1914

<210> 978
<211> 638
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 978

Met Asp Ser Gly Arg Asp Phe Leu Thr Leu His Gly Leu Gln Asp Asp
1 5 10 15

Glu Asp Leu Gln Ala Leu Leu Lys Gly Ser Gln Leu Leu Lys Val Lys
20 25 30

Ser Ser Ser Trp Arg Arg Glu Arg Phe Tyr Lys Leu Gln Glu Asp Cys
35 40 45

Lys Thr Ile Trp Gln Glu Ser Arg Lys Val Met Arg Thr Pro Glu Ser
50 55 60

Gln Leu Phe Ser Ile Glu Asp Ile Gln Glu Val Arg Met Gly His Arg
65 70 75 80

Thr Glu Gly Leu Glu Lys Phe Ala Arg Asp Val Pro Glu Asp Arg Cys

```
                85                  90                  95

Phe Ser Ile Val Phe Lys Asp Gln Arg Asn Thr Leu Asp Leu Ile Ala
            100                 105                 110

Pro Ser Pro Ala Asp Ala Gln His Trp Val Leu Gly Leu His Lys Ile
        115                 120                 125

Ile His His Ser Gly Ser Met Asp Gln Arg Gln Lys Leu Gln His Trp
    130                 135                 140

Ile His Ser Cys Leu Arg Lys Ala Asp Lys Asn Lys Asp Asn Lys Met
145                 150                 155                 160

Ser Phe Lys Glu Leu Gln Asn Phe Leu Lys Glu Leu Asn Ile Gln Leu
                165                 170                 175

Gly Ser Ser Gly Ala Glu Ile Gly Thr Gly Phe Pro Phe Asp Pro His
            180                 185                 190

Tyr Val Glu Val Leu Gly Glu Arg Met His Tyr Val Asp Val Gly Pro
        195                 200                 205

Arg Asp Gly Thr Pro Val Leu Phe Leu His Gly Asn Pro Thr Ser Ser
    210                 215                 220

Tyr Val Trp Arg Asn Ile Ile Pro His Val Ala Pro Thr His Arg Cys
225                 230                 235                 240

Ile Ala Pro Asp Leu Ile Gly Met Gly Lys Ser Asp Lys Pro Asp Leu
                245                 250                 255

Gly Tyr Phe Phe Asp Asp His Val Arg Phe Met Asp Ala Phe Ile Glu
            260                 265                 270

Ala Leu Gly Leu Glu Glu Val Val Leu Val Ile His Asp Trp Gly Ser
        275                 280                 285

Ala Leu Gly Phe His Trp Ala Lys Arg Asn Pro Glu Arg Val Lys Gly
    290                 295                 300
```

```
Ile Ala Phe Met Glu Phe Ile Arg Pro Ile Pro Thr Trp Asp Glu Trp
305                 310                 315                 320

Pro Glu Phe Ala Arg Glu Thr Phe Gln Ala Phe Arg Thr Thr Asp Val
                325                 330                 335

Gly Arg Lys Leu Ile Ile Asp Gln Asn Val Phe Ile Glu Gly Thr Leu
            340                 345                 350

Pro Met Gly Val Val Arg Pro Leu Thr Glu Val Glu Met Asp His Tyr
        355                 360                 365

Arg Glu Pro Phe Leu Asn Pro Val Asp Arg Glu Pro Leu Trp Arg Phe
    370                 375                 380

Pro Asn Glu Leu Pro Ile Ala Gly Glu Pro Ala Asn Ile Val Ala Leu
385                 390                 395                 400

Val Glu Glu Tyr Met Asp Trp Leu His Gln Ser Pro Val Pro Lys Leu
                405                 410                 415

Leu Phe Trp Gly Thr Pro Gly Val Leu Ile Pro Pro Ala Glu Ala Ala
            420                 425                 430

Arg Leu Ala Lys Ser Leu Pro Asn Cys Lys Ala Val Asp Ile Gly Pro
        435                 440                 445

Gly Leu Asn Leu Leu Gln Glu Asp Asn Pro Asp Leu Ile Gly Ser Glu
    450                 455                 460

Ile Ala Arg Trp Leu Ser Thr Leu Glu Ile Ser Gly Gly Ser Ser Gly
465                 470                 475                 480

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
                485                 490                 495

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            500                 505                 510

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
```

515 520 525

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
530 535 540

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
545 550 555 560

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly
565 570 575

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
580 585 590

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
595 600 605

Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile
610 615 620

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
625 630 635

<210> 979
<211> 159
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 979

Met Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala
1 5 10 15

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
20 25 30

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg
35 40 45

Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr

50 55 60

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
65 70 75 80

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr
85 90 95

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
100 105 110

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
115 120 125

Val Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu
130 135 140

Ile Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
145 150 155

<210> 980
<211> 459
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 980

Met Ala Glu Ile Gly Thr Gly Phe Pro Phe Asp Pro His Tyr Val Glu
1 5 10 15

Val Leu Gly Glu Arg Met His Tyr Val Asp Val Gly Pro Arg Asp Gly
20 25 30

Thr Pro Val Leu Phe Leu His Gly Asn Pro Thr Ser Ser Tyr Val Trp
35 40 45

Arg Asn Ile Ile Pro His Val Ala Pro Thr His Arg Cys Ile Ala Pro
50 55 60

Asp Leu Ile Gly Met Gly Lys Ser Asp Lys Pro Asp Leu Gly Tyr Phe

```
65                  70                  75                  80

Phe Asp Asp His Val Arg Phe Met Asp Ala Phe Ile Glu Ala Leu Gly
                85                  90                  95

Leu Glu Glu Val Val Leu Val Ile His Asp Trp Gly Ser Ala Leu Gly
            100                 105                 110

Phe His Trp Ala Lys Arg Asn Pro Glu Arg Val Lys Gly Ile Ala Phe
        115                 120                 125

Met Glu Phe Ile Arg Pro Ile Pro Thr Trp Asp Glu Trp Pro Glu Phe
    130                 135                 140

Ala Arg Glu Thr Phe Gln Ala Phe Arg Thr Thr Asp Val Gly Arg Lys
145                 150                 155                 160

Leu Ile Ile Asp Gln Asn Val Phe Ile Glu Gly Thr Leu Pro Met Gly
                165                 170                 175

Val Val Arg Pro Leu Thr Glu Val Glu Met Asp His Tyr Arg Glu Pro
            180                 185                 190

Phe Leu Asn Pro Val Asp Arg Glu Pro Leu Trp Arg Phe Pro Asn Glu
        195                 200                 205

Leu Pro Ile Ala Gly Glu Pro Ala Asn Ile Val Ala Leu Val Glu Glu
    210                 215                 220

Tyr Met Asp Trp Leu His Gln Ser Pro Val Pro Lys Leu Leu Phe Trp
225                 230                 235                 240

Gly Thr Pro Gly Val Leu Ile Pro Pro Ala Glu Ala Ala Arg Leu Ala
                245                 250                 255

Lys Ser Leu Pro Asn Cys Lys Ala Val Asp Ile Gly Pro Gly Leu Asn
            260                 265                 270

Leu Leu Gln Glu Asp Asn Pro Asp Leu Ile Gly Ser Glu Ile Ala Arg
        275                 280                 285
```

```
Trp Leu Ser Thr Leu Glu Ile Ser Gly Gly Ser Ser Gly Val Phe Thr
    290                 295                 300

Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr Asn Leu
305                 310                 315                 320

Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu
                325                 330                 335

Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly Glu Asn
            340                 345                 350

Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser
        355                 360                 365

Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro
    370                 375                 380

Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr Leu Val
385                 390                 395                 400

Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr
                405                 410                 415

Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr
            420                 425                 430

Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
        435                 440                 445

Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
    450                 455


<210>  981
<211>  638
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  981
```

```
Met Asp Ser Gly Arg Asp Phe Leu Thr Leu His Gly Leu Gln Asp Asp
1               5                   10                  15

Glu Asp Leu Gln Ala Leu Leu Lys Gly Ser Gln Leu Leu Lys Val Lys
            20                  25                  30

Ser Ser Ser Trp Arg Arg Glu Arg Phe Tyr Lys Leu Gln Glu Asp Cys
        35                  40                  45

Lys Thr Ile Trp Gln Glu Ser Arg Lys Val Met Arg Thr Pro Glu Ser
    50                  55                  60

Gln Leu Phe Ser Ile Glu Asp Ile Gln Glu Val Arg Met Gly His Arg
65                  70                  75                  80

Thr Glu Gly Leu Glu Lys Phe Ala Arg Asp Val Pro Glu Asp Arg Cys
                85                  90                  95

Phe Ser Ile Val Phe Lys Asp Gln Arg Asn Thr Leu Asp Leu Ile Ala
            100                 105                 110

Pro Ser Pro Ala Asp Ala Gln His Trp Val Leu Gly Leu His Lys Ile
        115                 120                 125

Ile His His Ser Gly Ser Met Asp Gln Arg Gln Lys Leu Gln His Trp
    130                 135                 140

Ile His Ser Cys Leu Arg Lys Ala Asp Lys Asn Lys Asp Asn Lys Met
145                 150                 155                 160

Ser Phe Lys Glu Leu Gln Asn Phe Leu Lys Glu Leu Asn Ile Gln Leu
                165                 170                 175

Gly Ser Ser Gly Ala Glu Ile Gly Thr Gly Phe Pro Phe Asp Pro His
            180                 185                 190

Tyr Val Glu Val Leu Gly Glu Arg Met His Tyr Val Asp Val Gly Pro
        195                 200                 205

Arg Asp Gly Thr Pro Val Leu Phe Leu His Gly Asn Pro Thr Ser Ser
```

```
    210                 215                 220

Tyr Val Trp Arg Asn Ile Ile Pro His Val Ala Pro Thr His Arg Cys
225                 230                 235                 240

Ile Ala Pro Asp Leu Ile Gly Met Gly Lys Ser Asp Lys Pro Asp Leu
                245                 250                 255

Gly Tyr Phe Phe Asp Asp His Val Arg Phe Met Asp Ala Phe Ile Glu
            260                 265                 270

Ala Leu Gly Leu Glu Glu Val Val Leu Val Ile His Asp Trp Gly Ser
        275                 280                 285

Ala Leu Gly Phe His Trp Ala Lys Arg Asn Pro Glu Arg Val Lys Gly
    290                 295                 300

Ile Ala Phe Met Glu Phe Ile Arg Pro Ile Pro Thr Trp Asp Glu Trp
305                 310                 315                 320

Pro Glu Phe Ala Arg Glu Thr Phe Gln Ala Phe Arg Thr Thr Asp Val
                325                 330                 335

Gly Arg Lys Leu Ile Ile Asp Gln Asn Val Phe Ile Glu Gly Thr Leu
            340                 345                 350

Pro Met Gly Val Val Arg Pro Leu Thr Glu Val Glu Met Asp His Tyr
        355                 360                 365

Arg Glu Pro Phe Leu Asn Pro Val Asp Arg Glu Pro Leu Trp Arg Phe
    370                 375                 380

Pro Asn Glu Leu Pro Ile Ala Gly Glu Pro Ala Asn Ile Val Ala Leu
385                 390                 395                 400

Val Glu Glu Tyr Met Asp Trp Leu His Gln Ser Pro Val Pro Lys Leu
                405                 410                 415

Leu Phe Trp Gly Thr Pro Gly Val Leu Ile Pro Pro Ala Glu Ala Ala
            420                 425                 430
```

```
Arg Leu Ala Lys Ser Leu Pro Asn Cys Lys Ala Val Asp Ile Gly Pro
        435                 440                 445

Gly Leu Asn Leu Leu Gln Glu Asp Asn Pro Asp Leu Ile Gly Ser Glu
    450                 455                 460

Ile Ala Arg Trp Leu Ser Thr Leu Glu Ile Ser Gly Gly Ser Ser Gly
465                 470                 475                 480

Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala
                485                 490                 495

Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu
            500                 505                 510

Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser
        515                 520                 525

Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu
    530                 535                 540

Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val
545                 550                 555                 560

Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly
                565                 570                 575

Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly
            580                 585                 590

Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val
        595                 600                 605

Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile
    610                 615                 620

Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
625                 630                 635
```

<210> 982

<211> 510
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 982

Met Pro Lys Lys Lys Arg Lys Val Ala Glu Ile Gly Thr Gly Phe Pro
1 5 10 15

Phe Asp Pro His Tyr Val Glu Val Leu Gly Glu Arg Met His Tyr Val
20 25 30

Asp Val Gly Pro Arg Asp Gly Thr Pro Val Leu Phe Leu His Gly Asn
35 40 45

Pro Thr Ser Ser Tyr Val Trp Arg Asn Ile Ile Pro His Val Ala Pro
50 55 60

Thr His Arg Cys Ile Ala Pro Asp Leu Ile Gly Met Gly Lys Ser Asp
65 70 75 80

Lys Pro Asp Leu Gly Tyr Phe Phe Asp Asp His Val Arg Phe Met Asp
85 90 95

Ala Phe Ile Glu Ala Leu Gly Leu Glu Glu Val Val Leu Val Ile His
100 105 110

Asp Trp Gly Ser Ala Leu Gly Phe His Trp Ala Lys Arg Asn Pro Glu
115 120 125

Arg Val Lys Gly Ile Ala Phe Met Glu Phe Ile Arg Pro Ile Pro Thr
130 135 140

Trp Asp Glu Trp Pro Glu Phe Ala Arg Glu Thr Phe Gln Ala Phe Arg
145 150 155 160

Thr Thr Asp Val Gly Arg Lys Leu Ile Ile Asp Gln Asn Val Phe Ile
165 170 175

Glu Gly Thr Leu Pro Met Gly Val Val Arg Pro Leu Thr Glu Val Glu

```
            180                 185                 190

Met Asp His Tyr Arg Glu Pro Phe Leu Asn Pro Val Asp Arg Glu Pro
        195                 200                 205

Leu Trp Arg Phe Pro Asn Glu Leu Pro Ile Ala Gly Glu Pro Ala Asn
    210                 215                 220

Ile Val Ala Leu Val Glu Glu Tyr Met Asp Trp Leu His Gln Ser Pro
225                 230                 235                 240

Val Pro Lys Leu Leu Phe Trp Gly Thr Pro Gly Val Leu Ile Pro Pro
                245                 250                 255

Ala Glu Ala Ala Arg Leu Ala Lys Ser Leu Pro Asn Cys Lys Ala Val
            260                 265                 270

Asp Ile Gly Pro Gly Leu Asn Leu Leu Gln Glu Asp Asn Pro Asp Leu
        275                 280                 285

Ile Gly Ser Glu Ile Ala Arg Trp Leu Ser Thr Leu Glu Ile Ser Gly
    290                 295                 300

Gly Ser Ser Gly Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Lys
305                 310                 315                 320

Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val
                325                 330                 335

Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg
            340                 345                 350

Met Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile
        355                 360                 365

Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu
    370                 375                 380

Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile
385                 390                 395                 400
```

```
Leu His Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu
                405                 410                 415


Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys
            420                 425                 430


Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp
        435                 440                 445


Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr Ile
    450                 455                 460


Asn Ser Pro Lys Lys Lys Arg Lys Val Pro Lys Lys Lys Arg Lys Val
465                 470                 475                 480


Pro Pro Pro Val Ser Lys Arg Glu Ser Lys Ser Arg Ser Arg Ser Lys
                485                 490                 495


Ser Pro Pro Lys Ser Pro Glu Glu Glu Gly Ala Val Ser Ser
            500                 505                 510


<210>  983
<211>  604
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Synthetic

<400>  983

Met Ser Val Leu Thr Pro Leu Leu Leu Arg Gly Leu Thr Gly Ser Ala
1               5                   10                  15


Arg Arg Leu Pro Val Pro Arg Ala Lys Ile His Ser Leu Gly Asp Pro
            20                  25                  30


Met Ser Val Leu Thr Pro Leu Leu Leu Arg Gly Leu Thr Gly Ser Ala
        35                  40                  45


Arg Arg Leu Pro Val Pro Arg Ala Lys Ile His Ser Leu Gly Lys Leu
    50                  55                  60
```

```
Met Ser Val Leu Thr Pro Leu Leu Leu Arg Gly Leu Thr Gly Ser Ala
65                  70                  75                  80

Arg Arg Leu Pro Val Pro Arg Ala Lys Ile His Ser Leu Gly Asp Pro
                85                  90                  95

Met Ser Val Leu Thr Pro Leu Leu Leu Arg Gly Leu Thr Gly Ser Ala
            100                 105                 110

Arg Arg Leu Pro Val Pro Arg Ala Lys Ile His Ser Leu Gly Lys Leu
        115                 120                 125

Arg Ile Leu Gln Ser Thr Val Pro Arg Ala Arg Asp Pro Pro Val Ala
    130                 135                 140

Thr Met Ala Glu Ile Gly Thr Gly Phe Pro Phe Asp Pro His Tyr Val
145                 150                 155                 160

Glu Val Leu Gly Glu Arg Met His Tyr Val Asp Val Gly Pro Arg Asp
                165                 170                 175

Gly Thr Pro Val Leu Phe Leu His Gly Asn Pro Thr Ser Ser Tyr Val
            180                 185                 190

Trp Arg Asn Ile Ile Pro His Val Ala Pro Thr His Arg Cys Ile Ala
        195                 200                 205

Pro Asp Leu Ile Gly Met Gly Lys Ser Asp Lys Pro Asp Leu Gly Tyr
    210                 215                 220

Phe Phe Asp Asp His Val Arg Phe Met Asp Ala Phe Ile Glu Ala Leu
225                 230                 235                 240

Gly Leu Glu Glu Val Val Leu Val Ile His Asp Trp Gly Ser Ala Leu
                245                 250                 255

Gly Phe His Trp Ala Lys Arg Asn Pro Glu Arg Val Lys Gly Ile Ala
            260                 265                 270

Phe Met Glu Phe Ile Arg Pro Ile Pro Thr Trp Asp Glu Trp Pro Glu
```

```
        275                 280                 285

Phe Ala Arg Glu Thr Phe Gln Ala Phe Arg Thr Thr Asp Val Gly Arg
    290                 295                 300

Lys Leu Ile Ile Asp Gln Asn Val Phe Ile Glu Gly Thr Leu Pro Met
305                 310                 315                 320

Gly Val Val Arg Pro Leu Thr Glu Val Glu Met Asp His Tyr Arg Glu
                325                 330                 335

Pro Phe Leu Asn Pro Val Asp Arg Glu Pro Leu Trp Arg Phe Pro Asn
            340                 345                 350

Glu Leu Pro Ile Ala Gly Glu Pro Ala Asn Ile Val Ala Leu Val Glu
        355                 360                 365

Glu Tyr Met Asp Trp Leu His Gln Ser Pro Val Pro Lys Leu Leu Phe
    370                 375                 380

Trp Gly Thr Pro Gly Val Leu Ile Pro Pro Ala Glu Ala Ala Arg Leu
385                 390                 395                 400

Ala Lys Ser Leu Pro Asn Cys Lys Ala Val Asp Ile Gly Pro Gly Leu
                405                 410                 415

Asn Leu Leu Gln Glu Asp Asn Pro Asp Leu Ile Gly Ser Glu Ile Ala
            420                 425                 430

Arg Trp Leu Ser Thr Leu Glu Ile Ser Gly Gly Ser Ser Gly Val Phe
        435                 440                 445

Thr Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr Asn
    450                 455                 460

Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn
465                 470                 475                 480

Leu Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly Glu
                485                 490                 495
```

```
Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu
            500                 505                 510

Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr
        515                 520                 525

Pro Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr Leu
    530                 535                 540

Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro
545                 550                 555                 560

Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly
                565                 570                 575

Thr Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro
            580                 585                 590

Asp Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser
        595                 600
```

<210> 984
<211> 480
<212> PRT
<213> Artificial sequence

<220>
<223> Synthetic

<400> 984

```
Met Leu Leu Ser Val Pro Leu Leu Leu Gly Leu Leu Gly Leu Ala Val
1               5                   10                  15

Ala Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp Glu Gln Thr Ala
            20                  25                  30

Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu
        35                  40                  45

Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln Arg Ile Val Arg
    50                  55                  60
```

```
Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr
65                  70                  75                  80

Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys
                85                  90                  95

Val Val Tyr Pro Val Asp Asp His His Phe Lys Val Ile Leu Pro Tyr
            100                 105                 110

Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe
        115                 120                 125

Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr
    130                 135                 140

Val Thr Gly Thr Leu Trp Asn Gly Asn Lys Ile Ile Asp Glu Arg Leu
145                 150                 155                 160

Ile Thr Pro Asp Gly Ser Met Leu Phe Arg Val Thr Ile Asn Ser Gly
                165                 170                 175

Ser Ser Gly Met Ala Glu Ile Gly Thr Gly Phe Pro Phe Asp Pro His
            180                 185                 190

Tyr Val Glu Val Leu Gly Glu Arg Met His Tyr Val Asp Val Gly Pro
        195                 200                 205

Arg Asp Gly Thr Pro Val Leu Phe Leu His Gly Asn Pro Thr Ser Ser
    210                 215                 220

Tyr Val Trp Arg Asn Ile Ile Pro His Val Ala Pro Thr His Arg Cys
225                 230                 235                 240

Ile Ala Pro Asp Leu Ile Gly Met Gly Lys Ser Asp Lys Pro Asp Leu
                245                 250                 255

Gly Tyr Phe Phe Asp Asp His Val Arg Phe Met Asp Ala Phe Ile Glu
            260                 265                 270

Ala Leu Gly Leu Glu Glu Val Val Leu Val Ile His Asp Trp Gly Ser
```

275 280 285

Ala Leu Gly Phe His Trp Ala Lys Arg Asn Pro Glu Arg Val Lys Gly
290 295 300

Ile Ala Phe Met Glu Phe Ile Arg Pro Ile Pro Thr Trp Asp Glu Trp
305 310 315 320

Pro Glu Phe Ala Arg Glu Thr Phe Gln Ala Phe Arg Thr Thr Asp Val
325 330 335

Gly Arg Lys Leu Ile Ile Asp Gln Asn Val Phe Ile Glu Gly Thr Leu
340 345 350

Pro Met Gly Val Val Arg Pro Leu Thr Glu Val Glu Met Asp His Tyr
355 360 365

Arg Glu Pro Phe Leu Asn Pro Val Asp Arg Glu Pro Leu Trp Arg Phe
370 375 380

Pro Asn Glu Leu Pro Ile Ala Gly Glu Pro Ala Asn Ile Val Ala Leu
385 390 395 400

Val Glu Glu Tyr Met Asp Trp Leu His Gln Ser Pro Val Pro Lys Leu
405 410 415

Leu Phe Trp Gly Thr Pro Gly Val Leu Ile Pro Pro Ala Glu Ala Ala
420 425 430

Arg Leu Ala Lys Ser Leu Pro Asn Cys Lys Ala Val Asp Ile Gly Pro
435 440 445

Gly Leu Asn Leu Leu Gln Glu Asp Asn Pro Asp Leu Ile Gly Ser Glu
450 455 460

Ile Ala Arg Trp Leu Ser Thr Leu Glu Ile Ser Gly Lys Asp Glu Leu
465 470 475 480

<210> 985
<211> 595
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 985

```
Met Ala Glu Ile Gly Thr Gly Phe Pro Phe Asp Pro His Tyr Val Glu
1               5                  10                  15


Val Leu Gly Glu Arg Met His Tyr Val Asp Val Gly Pro Arg Asp Gly
            20                  25                  30


Thr Pro Val Leu Phe Leu His Gly Asn Pro Thr Ser Ser Tyr Val Trp
        35                  40                  45


Arg Asn Ile Ile Pro His Val Ala Pro Thr His Arg Cys Ile Ala Pro
    50                  55                  60


Asp Leu Ile Gly Met Gly Lys Ser Asp Lys Pro Asp Leu Gly Tyr Phe
65                  70                  75                  80


Phe Asp Asp His Val Arg Phe Met Asp Ala Phe Ile Glu Ala Leu Gly
                85                  90                  95


Leu Glu Glu Val Val Leu Val Ile His Asp Trp Gly Ser Ala Leu Gly
            100                 105                 110


Phe His Trp Ala Lys Arg Asn Pro Glu Arg Val Lys Gly Ile Ala Phe
        115                 120                 125


Met Glu Phe Ile Arg Pro Ile Pro Thr Trp Asp Glu Trp Pro Glu Phe
    130                 135                 140


Ala Arg Glu Thr Phe Gln Ala Phe Arg Thr Thr Asp Val Gly Arg Lys
145                 150                 155                 160


Leu Ile Ile Asp Gln Asn Val Phe Ile Glu Gly Thr Leu Pro Met Gly
                165                 170                 175


Val Val Arg Pro Leu Thr Glu Val Glu Met Asp His Tyr Arg Glu Pro
            180                 185                 190
```

```
Phe Leu Asn Pro Val Asp Arg Glu Pro Leu Trp Arg Phe Pro Asn Glu
        195                 200                 205

Leu Pro Ile Ala Gly Glu Pro Ala Asn Ile Val Ala Leu Val Glu Glu
    210                 215                 220

Tyr Met Asp Trp Leu His Gln Ser Pro Val Pro Lys Leu Leu Phe Trp
225                 230                 235                 240

Gly Thr Pro Gly Val Leu Ile Pro Pro Ala Glu Ala Ala Arg Leu Ala
                245                 250                 255

Lys Ser Leu Pro Asn Cys Lys Ala Val Asp Ile Gly Pro Gly Leu Asn
            260                 265                 270

Leu Leu Gln Glu Asp Asn Pro Asp Leu Ile Gly Ser Glu Ile Ala Arg
        275                 280                 285

Trp Leu Ser Thr Leu Glu Ile Ser Gly Gly Ser Ser Gly Val Phe Thr
    290                 295                 300

Leu Glu Asp Phe Val Gly Asp Trp Lys Gln Thr Ala Ala Tyr Asn Leu
305                 310                 315                 320

Asp Gln Val Leu Glu Gln Gly Gly Val Ser Ser Leu Leu Gln Asn Leu
                325                 330                 335

Ala Val Ser Val Thr Pro Ile Gln Arg Met Val Arg Ser Gly Glu Asn
            340                 345                 350

Ala Leu Lys Ile Asp Ile His Val Ile Ile Pro Tyr Glu Gly Leu Ser
        355                 360                 365

Ala Asp Gln Met Ala Gln Ile Glu Glu Val Phe Lys Val Val Tyr Pro
    370                 375                 380

Val Asp Asp His His Phe Lys Val Ile Leu His Tyr Gly Thr Leu Val
385                 390                 395                 400

Ile Asp Gly Val Thr Pro Asn Met Leu Asn Tyr Phe Gly Arg Pro Tyr
```

405 410 415

Glu Gly Ile Ala Val Phe Asp Gly Lys Lys Ile Thr Val Thr Gly Thr
420 425 430

Leu Trp Asn Gly Ala Lys Ile Ile Asp Glu Arg Leu Ile Thr Pro Asp
435 440 445

Gly Ser Met Ser Phe Gln Val Thr Ile Asn Ser Ser Gly Leu Arg Ser
450 455 460

Arg Gly Glu Pro Gln Gln Ser Phe Ser Glu Ala Gln Gln Gln Leu Cys
465 470 475 480

Asn Thr Arg Gln Glu Val Asn Glu Leu Arg Lys Leu Leu Glu Glu Glu
485 490 495

Arg Asp Gln Arg Val Ala Ala Glu Asn Ala Leu Ser Val Ala Glu Glu
500 505 510

Gln Ile Arg Arg Leu Glu His Ser Glu Trp Asp Ser Ser Arg Thr Pro
515 520 525

Ile Ile Gly Ser Cys Gly Thr Gln Glu Gln Ala Leu Leu Ile Asp Leu
530 535 540

Thr Ser Asn Ser Cys Arg Arg Thr Arg Ser Gly Val Gly Trp Lys Arg
545 550 555 560

Val Leu Arg Ser Leu Cys His Ser Arg Thr Arg Val Pro Leu Leu Ala
565 570 575

Ala Ile Tyr Phe Leu Met Ile His Val Leu Leu Ile Leu Cys Phe Thr
580 585 590

Gly His Leu
595

<210> 986
<211> 883
<212> PRT

<213> Artificial sequence

<220>
<223> Synthetic

<400> 986

Met Ala Ala Pro Gly Ser Ala Arg Arg Pro Leu Leu Leu Leu Leu Leu
1 5 10 15

Leu Leu Leu Leu Gly Leu Met His Cys Ala Ser Ala Ala Met Phe Met
20 25 30

Val Lys Asn Gly Asn Gly Thr Ala Cys Ile Met Ala Asn Phe Ser Ala
35 40 45

Ala Phe Ser Val Asn Tyr Asp Thr Lys Ser Gly Pro Lys Asn Met Thr
50 55 60

Phe Asp Leu Pro Ser Asp Ala Thr Val Val Leu Asn Arg Ser Ser Cys
65 70 75 80

Gly Lys Glu Asn Thr Ser Asp Pro Ser Leu Val Ile Ala Phe Gly Arg
85 90 95

Gly His Thr Leu Thr Leu Asn Phe Thr Arg Asn Ala Thr Arg Tyr Ser
100 105 110

Val Gln Leu Met Ser Phe Val Tyr Asn Leu Ser Asp Thr His Leu Phe
115 120 125

Pro Asn Ala Ser Ser Lys Glu Ile Lys Thr Val Glu Ser Ile Thr Asp
130 135 140

Ile Arg Ala Asp Ile Asp Lys Lys Tyr Arg Cys Val Ser Gly Thr Gln
145 150 155 160

Val His Met Asn Asn Val Thr Val Thr Leu His Asp Ala Thr Ile Gln
165 170 175

Ala Tyr Leu Ser Asn Ser Ser Phe Ser Arg Gly Glu Thr Arg Cys Glu
180 185 190

```
Gln Asp Arg Pro Ser Pro Thr Thr Ala Pro Pro Ala Pro Pro Ser Pro
        195                 200                 205

Ser Pro Ser Pro Val Pro Lys Ser Pro Ser Val Asp Lys Tyr Asn Val
    210                 215                 220

Ser Gly Thr Asn Gly Thr Cys Leu Leu Ala Ser Met Gly Leu Gln Leu
225                 230                 235                 240

Asn Leu Thr Tyr Glu Arg Lys Asp Asn Thr Thr Val Thr Arg Leu Leu
                245                 250                 255

Asn Ile Asn Pro Asn Lys Thr Ser Ala Ser Gly Ser Cys Gly Ala His
            260                 265                 270

Leu Val Thr Leu Glu Leu His Ser Glu Gly Thr Thr Val Leu Leu Phe
        275                 280                 285

Gln Phe Gly Met Asn Ala Ser Ser Ser Arg Phe Phe Leu Gln Gly Ile
    290                 295                 300

Gln Leu Asn Thr Ile Leu Pro Asp Ala Arg Asp Pro Ala Phe Lys Ala
305                 310                 315                 320

Ala Asn Gly Ser Leu Arg Ala Leu Gln Ala Thr Val Gly Asn Ser Tyr
                325                 330                 335

Lys Cys Asn Ala Glu Glu His Val Arg Val Thr Lys Ala Phe Ser Val
            340                 345                 350

Asn Ile Phe Lys Val Trp Val Gln Ala Phe Lys Val Glu Gly Gly Gln
        355                 360                 365

Phe Gly Ser Val Glu Glu Cys Leu Leu Asp Glu Asn Ser Met Leu Ile
    370                 375                 380

Pro Ile Ala Val Gly Gly Ala Leu Ala Gly Leu Val Leu Ile Val Leu
385                 390                 395                 400

Ile Ala Tyr Leu Val Gly Arg Lys Arg Ser His Ala Gly Tyr Gln Thr
```

```
               405                 410                 415

Ile Lys Asp Pro Pro Val Ala Thr Met Ala Glu Ile Gly Thr Gly Phe
            420                 425                 430

Pro Phe Asp Pro His Tyr Val Glu Val Leu Gly Glu Arg Met His Tyr
        435                 440                 445

Val Asp Val Gly Pro Arg Asp Gly Thr Pro Val Leu Phe Leu His Gly
    450                 455                 460

Asn Pro Thr Ser Ser Tyr Val Trp Arg Asn Ile Ile Pro His Val Ala
465                 470                 475                 480

Pro Thr His Arg Cys Ile Ala Pro Asp Leu Ile Gly Met Gly Lys Ser
               485                 490                 495

Asp Lys Pro Asp Leu Gly Tyr Phe Phe Asp Asp His Val Arg Phe Met
            500                 505                 510

Asp Ala Phe Ile Glu Ala Leu Gly Leu Glu Glu Val Val Leu Val Ile
        515                 520                 525

His Asp Trp Gly Ser Ala Leu Gly Phe His Trp Ala Lys Arg Asn Pro
    530                 535                 540

Glu Arg Val Lys Gly Ile Ala Phe Met Glu Phe Ile Arg Pro Ile Pro
545                 550                 555                 560

Thr Trp Asp Glu Trp Pro Glu Phe Ala Arg Glu Thr Phe Gln Ala Phe
               565                 570                 575

Arg Thr Thr Asp Val Gly Arg Lys Leu Ile Ile Asp Gln Asn Val Phe
            580                 585                 590

Ile Glu Gly Thr Leu Pro Met Gly Val Val Arg Pro Leu Thr Glu Val
        595                 600                 605

Glu Met Asp His Tyr Arg Glu Pro Phe Leu Asn Pro Val Asp Arg Glu
    610                 615                 620
```

```
Pro Leu Trp Arg Phe Pro Asn Glu Leu Pro Ile Ala Gly Glu Pro Ala
625                 630                 635                 640

Asn Ile Val Ala Leu Val Glu Glu Tyr Met Asp Trp Leu His Gln Ser
                645                 650                 655

Pro Val Pro Lys Leu Leu Phe Trp Gly Thr Pro Gly Val Leu Ile Pro
            660                 665                 670

Pro Ala Glu Ala Ala Arg Leu Ala Lys Ser Leu Pro Asn Cys Lys Ala
        675                 680                 685

Val Asp Ile Gly Pro Gly Leu Asn Leu Leu Gln Glu Asp Asn Pro Asp
    690                 695                 700

Leu Ile Gly Ser Glu Ile Ala Arg Trp Leu Ser Thr Leu Glu Ile Ser
705                 710                 715                 720

Gly Gly Ser Ser Gly Val Phe Thr Leu Glu Asp Phe Val Gly Asp Trp
                725                 730                 735

Lys Gln Thr Ala Ala Tyr Asn Leu Asp Gln Val Leu Glu Gln Gly Gly
            740                 745                 750

Val Ser Ser Leu Leu Gln Asn Leu Ala Val Ser Val Thr Pro Ile Gln
        755                 760                 765

Arg Met Val Arg Ser Gly Glu Asn Ala Leu Lys Ile Asp Ile His Val
    770                 775                 780

Ile Ile Pro Tyr Glu Gly Leu Ser Ala Asp Gln Met Ala Gln Ile Glu
785                 790                 795                 800

Glu Val Phe Lys Val Val Tyr Pro Val Asp Asp His His Phe Lys Val
                805                 810                 815

Ile Leu His Tyr Gly Thr Leu Val Ile Asp Gly Val Thr Pro Asn Met
            820                 825                 830

Leu Asn Tyr Phe Gly Arg Pro Tyr Glu Gly Ile Ala Val Phe Asp Gly
```

```
        835                 840                 845

Lys Lys Ile Thr Val Thr Gly Thr Leu Trp Asn Gly Ala Lys Ile Ile
    850                 855                 860

Asp Glu Arg Leu Ile Thr Pro Asp Gly Ser Met Ser Phe Gln Val Thr
865                 870                 875                 880

Ile Asn Ser
```