

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2007292295 B2**

(54) Title
Fusion peptide therapeutic compositions

(51) International Patent Classification(s)
C12P 21/06 (2006.01) **C07K 14/00** (2006.01)
A61K 38/00 (2006.01)

(21) Application No: **2007292295** (22) Date of Filing: **2007.09.06**

(87) WIPO No: **WO08/030968**

(30) Priority Data

(31) Number	(32) Date	(33) Country
60/842,464	2006.09.06	US

(43) Publication Date: **2008.03.13**

(44) Accepted Journal Date: **2013.06.06**

(71) Applicant(s)
PhaseBio Pharmaceuticals, Inc.

(72) Inventor(s)
Chilkoti, Ashutosh

(74) Agent / Attorney
Allens Patent & Trade Mark Attorneys, Deutsche Bank Place Corner Hunter and Phillip Streets, SYDNEY, NSW, 2000

(56) Related Art
Meyer, D.E and Chikolti, A., Nat. Biotech. 1999, Vol. 17, No. 11, pages 1112-1115
Meyer, D.E. et al., Biotechnol. Prog. 2001, Vol. 17, No. 4, pages 720-728
Chow, D.C. et al., Biotechnol. Prog. 2006, Vol. 22, No. 3, pages 638-646.
US 6852834
US 20050255554

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
13 March 2008 (13.03.2008)

PCT

(10) International Publication Number
WO 2008/030968 A3

(51) International Patent Classification:
C12P 21/06 (2006.01) **C07K 14/00** (2006.01)
A61K 38/00 (2006.01)

(21) International Application Number:
PCT/US2007/077767

(22) International Filing Date:
6 September 2007 (06.09.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/842,464 6 September 2006 (06.09.2006) US

(71) Applicant (for all designated States except US): **PHASE BIOSCIENCE, INC.** [US/US]; 4134 S. Alston Ave., Suite 101, Durham, NC 27713-1870 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **CHILKOTI, Ashutosh** [US/US]; 915 Englewood Ave., Durham, NC 27701 (US).

(74) Agent: **REYNOLDS, Kelly, K.**; Intellectual Property/Technology Law, P.O. Box 14329, Research Triangle Park, NC 27709 (US).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:
13 November 2008

(54) Title: FUSION PEPTIDE THERAPEUTIC COMPOSITIONS

(57) Abstract: Therapeutic compositions containing fusion proteins (FPs) including elastin-like peptides (ELPs) and peptide active therapeutic agents, and methods of making and using such compositions and fusion proteins. Therapeutic compositions of such type enable improved efficacy of the peptide active therapeutic agent to be achieved, in relation to the peptide active therapeutic agent alone. Enhanced efficacy of the peptide active therapeutic agent in the therapeutic composition may include improved solubility, bioavailability, bio-unavailability, half-life, etc., as compared to corresponding compositions containing the same peptide active therapeutic agent without associated ELPs.



WO 2008/030968 A3

FUSION PEPTIDE THERAPEUTIC COMPOSITIONS

RELATED APPLICATION DATA

[0001] The application claims priority under 35 U.S.C. § 119(e) to U.S. Patent application Serial No. 60/842,464, filed September 6, 2006, incorporated by reference herein in its entirety.

FIELD OF THE INVENTION

[0002] The invention provides a new generation of therapeutic compositions, incorporating fusion proteins (FPs) including elastin-like peptides (ELPs) and peptide active therapeutic agents. The therapeutic compositions of the invention enable improved solubility, bioavailability or bio-unavailability, and half-life of the administered peptide active therapeutic agents to be achieved, as compared to corresponding compositions containing the same peptide active therapeutic agents without associated ELPs.

BACKGROUND OF THE INVENTION

[0003] The disclosures of US Patent 6,852,834, issued February 8, 2005 in the name of Ashutosh Chilkoti for "FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION," and US Patent Application No. 11/053,100 filed February 8, 2005 in the name of Ashutosh Chilkoti for "FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION," are hereby incorporated herein by reference, in their respective entireties, for all purposes.

[0004] The aforementioned Chilkoti patent and patent application disclose genetically-encodable, environmentally-responsive fusion proteins comprising ELP peptides. Such fusion proteins exhibit unique physico-chemical and functional properties that can be modulated as a function of solution environment.

SUMMARY OF THE INVENTION

[0005] The present invention relates broadly to fusion protein (FP) therapeutic compositions including elastin-like peptides (ELPs) and peptide active therapeutic agents.

[0005a] In a first aspect the invention provides a therapeutic composition comprising a fusion protein and a pharmaceutically-acceptable carrier, wherein the fusion protein comprises a therapeutic protein and at least one elastin-like protein (ELP) component, and wherein the therapeutic protein exhibits an extended half-life in circulation as compared to its unfused counterpart and wherein the therapeutic protein is glucagon-like peptide-1 (GLP-1).

[0005b] In a second aspect the invention provides a method of treating a subject in need of a therapeutic protein, the method comprising administering to the subject a therapeutic composition according to the first aspect.

[0005c] In a third aspect the invention provides use of a therapeutic composition according to the first aspect in the manufacture of a medicament for the treatment of a subject in need of a therapeutic protein.

[0006] In the FP therapeutic compositions of the invention, at least one peptide active therapeutic agent is coupled to one or more ELPs, e.g., being covalently bonded at an N- or C-terminus thereof, to achieve enhancement of the efficacy of the peptide active therapeutic agent(s), in relation to the corresponding therapeutic agent(s) alone. The peptide active therapeutic agent-ELP construct has enhanced efficacy in respect of any of various properties, such as solubility, bioavailability, bio-unavailability, therapeutic dose, resistance to proteolysis, half-life of the administered peptide active therapeutic agent, etc.

[0007] In another aspect, the invention relates to fusion gene constructs, including heterologous nucleotide sequences operably linked to an expression control element, e.g., a promoter of appropriate type, wherein the heterologous nucleotide sequences encode a fusion protein including at least one peptide active therapeutic agent coupled to at least one ELP.

[0008] In a further aspect, the invention relates to a method of enhancing efficacy of a peptide active therapeutic agent. The method includes coupling the peptide active therapeutic agent with at least one ELP to form a FP therapeutic composition, wherein the peptide active therapeutic agent in such FP therapeutic composition has enhanced efficacy, in relation to the peptide active therapeutic agent alone. In one aspect the enhanced efficacy is *in vivo* efficacy.

[0009] Another aspect of the invention relates to a method of treating a subject in need of a

2a

peptide active therapeutic agent, including administering to the patient a therapeutic composition including: (i) the peptide active therapeutic agent to coupled with at least one ELP, or (ii) a nucleotide sequence encoding a fusion protein including the peptide active therapeutic agent and at least one ELP, operably linked to an expression control element
5 therefore.

[0010] In still another aspect, the invention relates to a therapeutic agent dose form, in which the therapeutic agent is conjugated with an ELP.

[0011] Various other aspects, features and embodiments of the invention will be more fully apparent from the ensuing disclosure and appended claims.

10 [0011a] It is to be noted that, throughout the description and claims of this specification, the word 'comprise' and variations of the word, such as 'comprising' and 'comprises', is not intended to exclude other variants or additional components, integers or steps. Modifications and improvements to the invention will be readily apparent to those skilled in the art. Such modifications and improvements are intended to be within the scope of this invention.

15 [0011b] Any reference to or discussion of any document, act or item of knowledge in this specification is included solely for the purpose of providing a context for the present invention. It is not suggested or represented that any of these matters or any combination thereof formed at the priority date part of the common general knowledge, or was known to be relevant to an attempt to solve any problem with which this specification is concerned.

20

BRIEF DESCRIPTION OF THE FIGURES

[0012] Figure 1 is an SDS-PAGE gel showing expression of a 37 amino acid peptide, using the expression and purification methods of Example 1.

[0013] Figure 2 is a graph confirming the purification of the peptides resulting from the methods of Example 1.

[0014] Figure 3 is an SDS-PAGE gel showing the results of ITC purification of BFP, CAT and K1-3, as set forth in Example 6.

[0015] Figures 4A and 4B are graphs of the increase in turbidity as a function of temperature of each of the fusion constructs of Example 8 in PBS buffer.

[0016] Figure 5 is graph illustrating the blood concentration time-course for ^{14}C labeled ELP, as set forth in Example 9.

[0017] Figure 6 is a graph showing biodistribution of ^{14}C labeled ELP1-150 and ELP 2-160 in nude mice, as described in Example 10.

[0018] Figure 7 is a graph showing biodistribution of ^{14}C labeled ELP2-[V₁A₈G₇-160] in nude mice, as described in Example 10.

DETAILED DESCRIPTION OF THE INVENTION

[0019] The invention provides therapeutic compositions incorporating fusion proteins (FPs) including elastin-like peptides (ELPs) and peptide active therapeutic agents.

[0020] The therapeutic compositions of the invention enable increased efficacy of the peptide active therapeutic agent, e.g., improved solubility, bioavailability, bio-unavailability (where desired to avoid build up and/or toxicity, for example cardiotoxicity, etc.), half-life of the administered peptide active therapeutic agent, etc., to be achieved, as compared to corresponding compositions containing the same peptide active therapeutic agents without associated ELPs.

[0021] For ease of reference in the ensuing discussion, set out below are definitions of specific terms appearing in such discussion.

[0022] The term "protein" is used herein in a generic sense to include polypeptides of any length.

[0023] The term "peptide" as used herein is intended to be broadly construed as inclusive of polypeptides per se having molecular weights of up to about 10,000, as well as proteins having molecular weights of greater than about 10,000, wherein the molecular weights are number average

molecular weights. In a specific aspect, peptides having from about 2 to about 100 amino acid residues are particularly preferred as peptide therapeutic active agents of the invention.

[0024] As used herein, the term “coupled” means that the specified moieties are either directly covalently bonded to one another, or indirectly covalently joined to one another through an intervening moiety or moieties, such as a bridge, spacer, or linkage moiety or moieties, or they are non-covalently coupled to one another, e.g., by hydrogen bonding, ionic bonding, Van der Waals forces, etc.

[0025] As used herein, the term “half-life” means the period of time that is required for a 50% diminution of bioactivity of the active agent to occur. Such term is to be contrasted with “persistence,” which is the overall temporal duration of the active agent in the body, and “rate of clearance” as being a dynamically changing variable that may or may not be correlative with the numerical values of half-life and persistence.

[0026] The word “transform” is broadly used herein to refer to introduction of an exogenous polynucleotide sequence into a prokaryotic or eukaryotic cell by any means known in the art (including, for example, direct transmission of a polynucleotide sequence from a cell or virus particle as well as transmission by infective virus particles), resulting in a permanent or temporary alteration of genotype in an immortal or non-immortal cell line.

[0027] The term “functional equivalent” is used herein to refer to a protein that is an active analog, derivative, fragment, truncation isoform or the like of a native protein. A polypeptide is active when it retains some or all of the biological activity of the corresponding native polypeptide.

[0028] As used herein, “pharmaceutically acceptable” component (such as a salt, carrier, excipient or diluent) of a formulation according to the present invention is a component which (1) is compatible with the other ingredients of the formulation in that it can be combined with the FPs of the present invention without eliminating the biological activity of the FPs; and (2) is suitable for use with animals (including humans) without undue adverse side effects (such as toxicity, irritation, and allergic response). Side effects are “undue” when their risk outweighs the benefit provided by the pharmaceutical composition. Examples of pharmaceutically acceptable components include, without limitation, any of the standard pharmaceutical carriers such as phosphate buffered saline solutions, water, emulsions such as oil/water emulsions, microemulsions and various types of wetting agents.

[0029] As used herein, the term “native” used in reference to a protein indicates that the protein has the amino acid sequence of the corresponding protein as found in nature.

[0030] As used herein, the term “spacer” refers to any moiety that may be interposed between the ELP and the peptide active therapeutic agent in a given ELP/peptide active therapeutic agent construct. For example, the spacer may be a divalent group that is covalently bonded at one terminus to the ELP, and covalently bonded at the other terminus to the peptide active therapeutic agent. The ELP/peptide active therapeutic agent construct therefore is open to the inclusion of any additional chemical structure that does not preclude the efficacy of the ELP/peptide active therapeutic agent construct for its intended purpose. The spacer may for example be a protease-sensitive spacer moiety that is provided to control the pharmacokinetics of the ELP/peptide active therapeutic agent construct, or it may be a protease-insensitive ELP/ peptide active therapeutic agent construct.

[0031] Fusion protein (FP) therapeutic compositions of the invention at least one elastin-like peptide (ELP) coupled with at least one peptide active therapeutic agent. The ELP and peptide active therapeutic agent components of the composition may be coupled with one another in any suitable manner, including covalent bonding, ionic bonding, associative bonding, complexation, or any other coupling modality that is effective to aggregate the ELP and peptide active therapeutic agent components, so that the peptide active therapeutic agent is efficacious for its intended purpose, and so that the presence of the coupled ELP enhances the peptide active therapeutic agent in the composition in some functional, therapeutic or physiological aspect, so that it is more efficacious than the peptide active therapeutic agent alone.

[0032] Thus, the ELP-coupled peptide active therapeutic agent in the FP therapeutic composition may be enhanced in any other properties, e.g., its bioavailability, bio-unavailability, therapeutic dose, formulation compatibility, resistance to proteolysis or other degradative modalities, solubility, half-life or other measure of persistence in the body subsequent to administration, rate of clearance from the body subsequent to administration, etc.

[0033] In the FP therapeutic compositions of the invention, at least one peptide active therapeutic agent is coupled to one or more ELPs, e.g., being covalently bonded at an N- or C-terminus thereof, to achieve enhancement of the efficacy of the peptide active therapeutic agent(s), in relation to the corresponding therapeutic agent(s) alone.

[0034] The FP therapeutic compositions of the invention may be therapeutically administered directly, or otherwise be produced *in vivo* from corresponding fusion gene constructs, including heterologous nucleotide sequences operably linked to an expression control element, e.g., a promoter of appropriate type, wherein the heterologous nucleotide sequences encode a fusion protein including at least one peptide active therapeutic agent coupled to at least one ELP.

[0035] The invention enables the enhancement of the efficacy of a peptide active therapeutic agent, e.g., by coupling the peptide active therapeutic agent with at least one ELP to form a FP therapeutic composition, wherein the peptide active therapeutic agent in such FP therapeutic composition has enhanced efficacy in relation to the peptide active therapeutic agent alone.

[0036] The invention may be practiced using any suitable therapeutic dose form including at least one peptide active therapeutic agent, coupled with at least one ELP.

[0037] The invention enables stabilization of a peptide active therapeutic agent against proteolytic degradation, by coupling such agent with at least one ELP to form a FP therapeutic composition.

[0038] The FP therapeutic composition of the invention may include one or more ELP species, and one or more peptide active therapeutic agents. As indicated hereinabove, the ELP species and peptide active therapeutic agents may be coupled directly with one another, or alternatively such coupling may be effected in a construct including a spacer moiety intermediate the ELP and the peptide active therapeutic agent.

[0039] The ELP species used in the FP therapeutic composition of the invention may be of any suitable type. ELPs are repeating peptide sequences that have been found to exist in the elastin protein. Among these repeating peptide sequences are polytetra-, polypenta-, polyhexa-, polyhepta-, polyocta, and polynona-peptides.

[0040] ELPs undergo a reversible inverse temperature transition. They are structurally disordered and highly soluble in water below a transition temperature (T_t), but exhibit a sharp (2-3°C range) disorder-to-order phase transition when the temperature is raised above T_t , leading to desolvation and aggregation of the polypeptides. The ELP aggregates, when reaching sufficient size, can be readily removed and isolated from solution by centrifugation. Such phase transition is reversible, and isolated ELP aggregates can be completely resolubilized in buffer solution when the temperature is returned below the T_t of the ELPs.

[0041] In the practice of the present invention, the ELPs species functions to stabilize or otherwise improve the peptide active therapeutic agent in the therapeutic composition. Subsequent to administration of the coupled peptide active therapeutic agent-ELP construct to the patient in need of the peptide therapeutic agent, the peptide active therapeutic agent and the ELP remain coupled with one another while the peptide active therapeutic agent is therapeutically active, e.g., for treatment or prophylaxis of a disease state or physiological condition, or other therapeutic intervention.

[0042] For example, the ELPs in therapeutic compositions of the present invention may comprise ELPs formed of polymeric or oligomeric repeats of various characteristic tetra-, penta-, hexa-, hepta-, octa-, and nonapeptides, including but not limited to:

- (a) tetrapeptide Val-Pro-Gly-Gly, or VPGG (SEQ ID NO: 1);
- (b) tetrapeptide Ile-Pro-Gly-Gly, or IPGG (SEQ ID NO: 2);
- (c) pentapeptide Val-Pro-Gly-X-Gly (SEQ ID NO: 3), or VPGXG, wherein X is any natural or non-natural amino acid residue, and wherein X optionally varies among polymeric or oligomeric repeats;
- (d) pentapeptide Ala-Val-Gly-Val-Pro, or AVGVP (SEQ ID NO: 4);
- (e) pentapeptide Ile-Pro-Gly-Val-Gly, or IPGVG (SEQ ID NO: 5);
- (f) pentapeptide Leu-Pro-Gly-Val-Gly, or LPGVG (SEQ ID NO: 6);
- (g) hexapeptide Val-Ala-Pro-Gly-Val-Gly, or VAPGVG (SEQ ID NO: 7);
- (h) octapeptide Gly-Val-Gly-Val-Pro-Gly-Val-Gly, or GVGVPGVG (SEQ ID NO: 8);
- (i) nonapeptide Val-Pro-Gly-Phe-Gly-Val-Gly-Ala-Gly, or VPGFGVGAG (SEQ ID NO: 9); and
- (j) nonapeptides Val-Pro-Gly-Val-Gly-Val-Pro-Gly-Gly, or VPGVGVPGG (SEQ ID NO: 10).

[0043] Any other polymeric or oligomeric repeat units of other sizes and constitutions can be usefully employed in the broad practice of the present invention.

[0044] In one embodiment, the ELP in the peptide active therapeutic agent-ELP construct includes repeat units of the pentapeptide Val-Pro-Gly-X-Gly, wherein X is as defined above, and wherein the ratio of Val-Pro-Gly-X-Gly pentapeptide units to other amino acid residues of the ELP is greater than about 75%, more preferably greater than about 85%, still more preferably greater than about 95%.

[0045] The peptide active therapeutic agent-ELP constructs of the invention may be synthetically, e.g., recombinantly, produced.

[0046] In the peptide active therapeutic agent-ELP construct, the ELP may be joined at a C- and/or N-terminus of the peptide active therapeutic agent, and optionally, a spacer sequence may be present separating the ELP from the peptide active therapeutic agent.

[0047] In one aspect, the invention contemplates a polynucleotide comprising a nucleotide sequence encoding a peptide active therapeutic agent-ELP fusion protein, optionally including a spacer sequence as above described, separating the ELP from the peptide active therapeutic agent. The polynucleotide may be provided as a component of an expression vector. The invention also contemplates a host cell (prokaryotic or eukaryotic) transformed by such expression vector to express the fusion protein.

[0048] The peptide active therapeutic agent-ELP construct subsequent to its synthesis or expression can be isolated by a method involving effecting a phase transition, e.g., by raising temperature, or in other manner, producing a phase transition of the fusion protein in the medium in which is contained in non-isolated form.

[0049] For example, the peptide active therapeutic agent-ELP construct may be synthesized and recovered, by steps including forming a polynucleotide comprising a nucleotide sequence encoding a peptide active therapeutic agent-ELP fusion protein exhibiting a phase transition, expressing the fusion protein in culture, and subjecting fusion protein-containing material from the culture to processing involving separation (e.g., by centrifugation, membrane separation, etc.) and inverse transition cycling to recover the peptide active therapeutic agent-ELP fusion protein.

[0050] In one specific embodiment, the peptide active therapeutic agent-ELP fusion protein includes an ELP moiety including polymeric or oligomeric repeats of a polypeptide selected from the group consisting of VPGG, IPGG, AVGVG, IPGVG, LPGVG, VAPGVG, GVGVPGVG, VPGFGVGAG, and VPGVGVPGG.

[0051] In another specific embodiment, the peptide active therapeutic agent-ELP fusion protein includes an ELP moiety including polymeric or oligomeric repeat units selected from the group consisting of LPGXG (SEQ ID NO: 11), IPGXG (SEQ ID NO: 12), and combinations thereof, wherein X is an amino acid residue that does not preclude phase transition of the ELP fusion protein.

[0052] The peptide active therapeutic agent-ELP construct of the invention comprises an amino acid sequence endowing the construct with phase transition characteristics.

[0053] The ELP in the peptide active therapeutic agent-ELP construct can include a β -turn component. Examples of polypeptides suitable for use as the β -turn component are described in Urry, et al. International Patent Application PCT/US96/05186. Alternatively, the ELP in the peptide active therapeutic agent-ELP construct can be a component lacking a β -turn component, or otherwise having a different conformation and/or folding character.

[0054] The ELPs, as mentioned, can include polymeric or oligomeric repeats of various tetra-, penta-, hexa-, hepta-, octa-, and nonapeptides, including but not limited to VPGG, IPGG, VPGXG, AVGVP, IPGVG, LPGVG, VAPGVG, GVGVPVG, VPGFGVGAG, and VPGVGVPGG (SEQ NO: 1 to SEQ NO: 10). It will be appreciated by those of skill in the art that the ELPs need not consist of only polymeric or oligomeric sequences as listed hereinabove, in order to exhibit a phase transition or otherwise constitute a suitable ELPs species for use in the peptide active therapeutic agent-ELP constructs of the invention.

[0055] In one embodiment, the peptide active therapeutic agent-ELP construct includes ELPs that are polymeric or oligomeric repeats of the pentapeptide VPGXG (SEQ ID NO: 3), where the guest residue X is any amino acid that does not eliminate the phase transition characteristics of the ELP. X may be a naturally occurring or non-naturally occurring amino acid. For example, X may be selected from the group consisting of: alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine and valine. In a specific embodiment, X is not proline.

[0056] X may be a non-classical amino acid. Examples of non-classical amino acids include: D-isomers of the common amino acids, 2,4-diaminobutyric acid, α -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, γ -Abu, ϵ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β -alanine, fluoro-amino acids, designer amino acids such as β -methyl amino acids, C α -methyl amino acids, N α -methyl amino acids, and amino acid analogs in general.

[0057] Selection of the identity of X is independent in each ELP repetition. Selection may be based on any desired characteristic, such as consideration of positively charged or negatively

charged residues in the X position. It may be considered that ELPs with neutral values in the X position may have longer half-lives.

[0058] In another embodiment, the peptide active therapeutic agent-ELP construct includes ELPs that are polymeric or oligomeric repeats of the pentapeptide IPGXG (SEQ ID NO: 11) or LPGXG (SEQ ID NO: 12), where X is as defined hereinabove.

[0059] The polymeric or oligomeric repeats of the ELP sequences may be separated by one or more amino acid residues that do not eliminate the overall phase transition characteristic of the peptide active therapeutic agent-ELP construct. In one specific embodiment, when the ELP component of the peptide active therapeutic agent-ELP construct comprises polymeric or oligomeric repeats of the pentapeptide VPGXG, the ratio of VPGXG repeats to other amino acid residues of the ELP is greater than about 75%, more preferably greater than about 85%, still more preferably greater than about 95%, and most preferably greater than about 99%.

[0060] In each repeat, X is independently selected. Different resulting ELP species are distinguished here using the notation ELP_k [X_iY_j-n], where k designates the specific type of ELP repeat unit, the bracketed capital letters are single letter amino acid codes and their corresponding subscripts designate the relative ratio of each guest residue X in the repeat units, and n describes the total length of the ELP in number of the pentapeptide repeats. For example, ELP1 [V₅A₂G₃-10] designates a polypeptide containing 10 repeating units of the pentapeptide VPGXG, where X is valine, alanine, and glycine at a relative ratio of 5:2:3; ELP1 [K₁V₂F₁-4] designates a polypeptide containing 4 repeating units of the pentapeptide VPGXG, where X is lysine, valine, and phenylalanine at a relative ratio of 1:2:1; ELP1 [K₁V₇F₁-9] designates a polypeptide containing 4 repeating units of the pentapeptide VPGXG, where X is lysine, valine, and phenylalanine at a relative ratio of 1:7:1; ELP1 [V-5] designates a polypeptide containing 5 repeating units of the pentapeptide VPGXG, where X is exclusively valine; ELP1 [V-20] designates a polypeptide containing 20 repeating units of the pentapeptide VPGXG, where X is exclusively valine; ELP2 [5] designates a polypeptide containing 5 repeating units of the pentapeptide AVGVP; ELP3 [V-5] designates a polypeptide containing 5 repeating units of the pentapeptide IPGXG, where X is exclusively valine; ELP4 [V-5] designates a polypeptide containing 5 repeating units of the pentapeptide LPGXG, where X is exclusively valine.

[0061] Previous studies by Urry and colleagues have shown that the fourth residue (X) in the elastin pentapeptide sequence, VPGXG, can be altered without eliminating the formation of the β -turn. These studies also showed that the T_i is a function of the hydrophobicity of the guest residue.

By varying the identity of the guest residue(s) and their mole fraction(s), ELPs can be synthesized that exhibit an inverse transition over a 0-100°C range.

[0062] The T_i at a given ELP length can be decreased by incorporating a larger fraction of hydrophobic guest residues in the ELP sequence. Examples of suitable hydrophobic guest residues include valine, leucine, isoleucine, phenylalanine, tryptophan and methionine. Tyrosine, which is moderately hydrophobic, may also be used. Conversely, the T_i can be increased by incorporating residues, such as those selected from the group consisting of: glutamic acid, cysteine, lysine, aspartate, alanine, asparagine, serine, threonine, glycine, arginine, and glutamine; preferably selected from alanine, serine, threonine and glutamic acid.

[0063] The ELP in one embodiment is selected to provide a T_i ranging from about 10 to about 80°C, more preferably from about 35 to about 60°C, most preferably from about 38 to about 45°C.

[0064] The T_i can also be varied by varying ELP chain length. The T_i increases with decreasing MW. For polypeptides having a molecular weight >100,000, the hydrophobicity scale developed by Urry et al. (PCT/US96/05186) is preferred for predicting the approximate T_i of a specific ELP sequence.

[0065] For polypeptides having a molecular weight <100,000, the T_i is preferably determined by the following quadratic function:

$$T_i = M_0 + M_1X + M_2X^2$$

where X is the MW of the FP, and $M_0 = 116.21$; $M_1 = -1.7499$; $M_2 = 0.010349$.

[0066] While the T_i of the ELP and, therefore of a construct of an ELP linked to a peptide active therapeutic agent, is affected by the identity and hydrophobicity of the guest residue, X, additional properties of the construct may also be affected. Such properties include, but are not limited to solubility, bioavailability or bio-unavailability, and half-life of the ELP itself and the construct.

[0067] In the Examples section below, it is seen that the ELP-coupled active therapeutic agent retains a significant amount of the therapeutic agent's biological activity, as compared to free protein forms of such therapeutic agent. Additionally, it is shown that ELPs exhibit long half-lives. Correspondingly, ELPs can be used in accordance with the invention to substantially increase (e.g. by greater than 10%, 20%, 30%, 50%, 100%, 200% or more, in specific embodiments) the half-life of the therapeutic agent, as conjugated with an ELP, in comparison to the half-life of the free (unconjugated) form of the therapeutic agent. Furthermore, ELPs are shown

to target high blood content organs, when administered *in vivo*, and thus, can partition in the body, to provide a predetermined desired corporeal distribution among various organs or regions of the body, or a desired selectivity or targeting of a therapeutic agent. In sum, active ELP-therapeutic agent conjugates contemplated by the invention are administered or generated *in vivo* as active, site-specific compositions having extended half-lives.

[0068] In one embodiment of the invention, the ELP length is from 5 to about 500 amino acid residues, more preferably from about 10 to about 450 amino acid residues, and still more preferably from about 15 to about 150 amino acid residues. ELP length can be reduced while maintaining a target T_i by incorporating a larger fraction of hydrophobic guest residues in the ELP sequence.

[0069] The active therapeutic agent in the peptide active therapeutic agent-ELP construct can be of any suitable type. Suitable peptides include those of interest in medicine, agriculture and other scientific and industrial fields, particularly including therapeutic proteins such as erythropoietins, magainins, beta-defensins, interferons, insulin, monoclonal antibodies, blood factors, colony stimulating factors, growth hormones, interleukins, growth factors, therapeutic vaccines, calcitonins, tumor necrosis factors (TNF), receptor antagonists, corticosteroids, and enzymes. Specific examples of such peptides include, without limitation, enzymes utilized in replacement therapy; antibacterial peptides; hormones for promoting growth; and active proteinaceous substances used in various applications. Specific examples include, but are not limited to: superoxide dismutase, interferon, asparaginase, glutamase, arginase, arginine deaminase, adenosine deaminase ribonuclease, trypsin, chymotrypsin, papain, insulin, calcitonin, ACTH, glucagon, glucagon-like peptide-1 (GLP-1), somatostatin, somatropin, somatomedin, parathyroid hormone, erythropoietin, hypothalamic releasing factors, prolactin, thyroid stimulating hormones, endorphins, enkephalins, and vasopressin.

[0070] In one embodiment of the invention, the peptide active therapeutic agent is thioredoxin.

[0071] In another embodiment, the peptide active therapeutic agent is tendamistat. The tendamistat-ELP fusion protein provides a readily-isolated, active version of tendamistat for use as an α -amylase inhibitor, e.g., in the treatment of pancreatitis. This fusion protein is suitably provided as a component of a pharmaceutical formulation in association with a pharmaceutically acceptable carrier. The tendamistat-ELP fusion protein retains most of the α -amylase inhibition activity of the free tendamistat, and is a stable construct.

[0072] In one specific embodiment, the peptide active therapeutic agent includes a physiologically active peptide selected from the group consisting of insulin, calcitonin, ACTH, glucagon, somatostatin, somatotropin, somatomedin, parathyroid hormone, erythropoietin, hypothalamic releasing factors, prolactin, thyroid stimulating hormones, endorphins, enkephalins, vasopressin, non-naturally occurring opioids, superoxide dismutase, interferon, asparaginase, arginase, arginine deaminase, adenosine deaminase ribonuclease, trypsin, chymotrypsin, and papain.

[0073] The invention thus comprehends various compositions for therapeutic (*in vivo*) application, wherein the peptide component of the peptide active therapeutic agent-ELP construct is a physiologically active, or bioactive, peptide. In preferred forms of such peptide-containing compositions, the coupling of the peptide component to ELP species is effected by direct covalent bonding or indirect (through appropriate spacer groups) bonding, and the peptide and ELP moieties can be structurally arranged in any suitable manner involving such direct or indirect covalent bonding, relative to one another. Thus, a wide variety of peptide species can be accommodated in the broad practice of the present invention, as necessary or desirable in a given therapeutic application.

[0074] The peptides utilized as peptide active therapeutic agents in the peptide active therapeutic agent-ELP constructs of the invention in one embodiment include enzymes utilized in replacement therapy and hormones for promoting growth. Among such enzymes are superoxide dismutase, interferon, asparaginase, glutamase, arginase, arginine deaminase, adenosine deaminase ribonuclease, cytosine deaminase, trypsin, chromotrypsin, and papin. Among the peptide hormones, specific species amenable to use in the peptide active therapeutic agent-ELP constructs of the invention include, without limitation, insulin, calcitonin, ACTH, glucagon, somatostatin, somatotropin, somatomedin, parathyroid hormone, erythropoietin, hypothalamic releasing factors, prolactin, thyroid stimulating hormones, endorphins, enkephalins, and vasopressin.

[0075] In another specific aspect, the peptide active therapeutic agent in the ELPs/peptide active therapeutic agent construct is selected from among the following species, and all variants, fragments and derivatives of such species: agouti related peptide, amylin, angiotensin, cecropin, bombesin, gastrin, including gastrin releasing peptide, lactoferrin, antimicrobial peptides including but not limited to magainin, urodilatin, nuclear localization signal (NLS), collagen peptide, survivin, amyloid peptides, including β -amyloid, natriuretic peptides, peptide YY, neuroregenerative peptides and neuropeptides, including but not limited to neuropeptide Y, dynorphin, endomorphin, endothelin, enkaphalin, exendin, fibronectin, neuropeptide W and neuropeptide S, peptide T, melanocortin, amyloid precursor protein, sheet breaker peptide, CART

peptide, amyloid inhibitory peptide, prion inhibitory peptide, chlorotoxin, corticotropin releasing factor, oxytocin, vasopressin, cholecystokinin, secretin, thymosin, epidermal growth factor (EGF), vascular endothelial cell growth factor (VEGF), platelet-derived growth factor (PDGF), Insulin-like growth factor (IGF), fibroblast growth factors (aFGF, bFGF), pancreastatin, melanocyte stimulating hormone, osteocalcin, bradykinin, adrenomedullin, perinerin, metastatin, aprotinin, galanins, including galanin-like peptide, leptin, defensins, including but not limited to α -defensin and β -defensin, salusin, and various venoms, including but not limited to conotoxin, decorsin, kurtosin, anenomae venom, tarantula venom; natriuretic peptides including brain natriuretic peptide (B-type natriuretic peptide, or BNP), atrial natriuretic peptide, and vasonatin; neurokinin A, neurokinin B; neuromedin; neurotensin; orexin, pancreatic polypeptide, pituitary adenylate cyclase activating peptide (PACAP), prolactin releasing peptide, proteolipid protein (PLP), somatostatin, TNF- α ; Grehlin, Protein C (Xigris), SS1(dsFv)-PE38 and pseudomonas exotoxin protein, clotting factors, including antithrombin III and Coagulation Factor VIIA, Factor VIII, Factor IX, streptokinase, tissue plasminogen activators, urokinase, beta glucocerebrosidase and alpha-D-galactosidase, alpha L-iduronidase, alpha-1, 4-glucosidase, arylsulfatase B, iduronate-2-sulfatase, deoxyribonuclease I, human activated protein, follicle-stimulating hormone, chorionic gonadotropin, luteinizing hormone, somatropin, bone morphogenetic protein, nesiritide, parathyroid hormone, erythropoietin, keratinocyte growth factor, human granulocyte colony-stimulating factor (G-CSF), human granulocyte-macrophage colony stimulating factor (GM-CSF), alpha interferon, beta interferon, gamma interferon, interleukins, including IL-1, IL-1Ra, IL-2, IL-4, IL-5, IL-6, IL-10, IL-11, IL-12, glycoprotein IIB/IIIA, immune globulins, including hepatitis B, gamma globulin, venoglobulin, hirudin, aprotinin, antithrombin III, alpha-1-proteinase inhibitor, filgrastim, and etanercept.

[0076] In another embodiment, the peptide component of the peptide active therapeutic agent-ELP constructs of the present invention may be an antibody or antigen, in connection with immunotherapy, or other therapeutic intervention.

[0077] Various other proteins and peptides, such as insulin A peptide, T20 peptide, interferon alpha 2B peptide, tobacco etch virus protease, small heterodimer partner orphan receptor, androgen receptor ligand binding domain, glucocorticoid receptor ligand binding domain, estrogen receptor ligand binding domain, G protein alpha Q, 1-deoxy-D-xylulose 5-phosphate reductoisomerase peptide, G protein alpha S, angiostatin (K1-3), blue fluorescent protein (BFP), calmodulin (CaM), chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), interleukin 1 receptor antagonist (IL-1Ra), luciferase, tissue transglutaminase (tTg), morphine modulating neuropeptide

(MMN), neuropeptide Y (NPY), orexin-B, leptin, ACTH, calcitonin, adrenomedullin (AM), parathyroid hormone (PTH), defensin and growth hormone have been fused with different ELP polypeptides to form FPs that exhibit inverse phase transition behavior.

[0078] The proteins and peptides employed as active therapeutic agents can be significantly different in their primary, secondary, and tertiary structures, sizes, molecular weights, solubility, electric charge distribution, viscosity, and biological functions.

[0079] Also included within the scope of the invention are derivatives comprising FPs, which have been differentially modified during or after synthesis, e.g., by benzylation, glycosylation, acetylation, phosphorylation, amidation, PEGylation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. In one embodiment, the FPs are acetylated at the N-terminus and/or amidated at the C-terminus. In another embodiment, the FPs are conjugated to polymers, e.g., polymers known in the art to facilitate oral delivery, decrease enzymatic degradation, increase solubility of the polypeptides, or otherwise improve the chemical properties of the therapeutic polypeptides for administration to humans or other animals.

[0080] The peptide active therapeutic agent-ELP constructs of the invention can be obtained by known recombinant expression techniques. To recombinantly produce the peptide active therapeutic agent-ELP construct, a nucleic acid sequence encoding the construct is operatively linked to a suitable promoter sequence such that the nucleic acid sequence encoding such fusion peptide will be transcribed and/or translated into the desired fusion peptide in the host cells. Preferred promoters are those useful for expression in *E. coli*, such as the T7 promoter.

[0081] Any commonly used expression system may be used, e.g., eukaryotic or prokaryotic systems. Specific examples include yeast, pichia, baculovirus, mammalian, and bacterial systems, such as *E. coli*, and *Caulobacter*.

[0082] A vector comprising the nucleic acid sequence can be introduced into a cell for expression of the peptide active therapeutic agent-ELP construct. The vector can remain episomal or become chromosomally integrated, as long as the gene carried by it can be transcribed to produce the desired RNA. Vectors can be constructed by standard recombinant DNA technology methods. Vectors can be plasmids, phages, cosmids, phagemids, viruses, or any other types known in the art, which are used for replication and expression in prokaryotic or eukaryotic cells. It will be appreciated by one of skill in the art that a wide variety of components known in the art may be

included in such vectors, including a wide variety of transcription signals, such as promoters and other sequences that regulate the binding of RNA polymerase onto the promoter. Any promoter known to be effective in the cells in which the vector will be expressed can be used to initiate expression of the peptide active therapeutic agent-ELP construct. Suitable promoters may be inducible or constitutive. Examples of suitable promoters include the SV40 early promoter region, the promoter contained in the 3' long terminal repeat of Rous sarcoma virus, the HSV-1 (herpes simplex virus-1) thymidine kinase promoter, the regulatory sequences of the metallothionein gene, etc., as well as the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells; insulin gene control region which is active in pancreatic beta cells, immunoglobulin gene control region which is active in lymphoid cells, mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells, albumin gene control region which is active in liver, alpha-fetoprotein gene control region which is active in liver, alpha 1-antitrypsin gene control region which is active in the liver, beta-globin gene control region which is active in erythroid cells, myelin basic protein gene control region which is active in oligodendrocyte cells in the brain, myosin light chain-2 gene control region which is active in skeletal muscle, and gonadotropin releasing hormone gene control region which is active in the hypothalamus.

[0083] In one embodiment, a mammal is genetically modified to produce the peptide active therapeutic agent-ELP construct in its milk. Techniques for performing such genetic modifications are described in U.S. Patent 6,013,857, issued January 11, 2000, for "Transgenic Bovines and Milk from Transgenic Bovines." The genome of the transgenic animal is modified to comprise a transgene comprising a DNA sequence encoding a peptide active therapeutic agent-ELP construct operably linked to a mammary gland promoter. Expression of the DNA sequence results in the production of the peptide active therapeutic agent-ELP construct in the milk. The peptide active therapeutic agent-ELP construct can then be isolated by phase transition from milk obtained from the transgenic mammal. The transgenic mammal is preferably a bovine.

[0084] The peptide active therapeutic agent-ELP constructs of the invention can be separated from other contaminating proteins to high purity using inverse transition cycling procedures, e.g., utilizing the temperature-dependent solubility of the peptide active therapeutic agent-ELP construct, or salt addition to the medium containing the construct. Successive inverse phase transition cycles can be used to obtain a high degree of purity.

[0085] In addition to temperature and ionic strength, other environmental variables useful for modulating the inverse transition of peptide active therapeutic agent-ELP constructs include pH, the addition of inorganic and organic solutes and solvents, side-chain ionization or chemical modification, and pressure.

[0086] In one specific illustrative embodiment of the invention, a 10 polypentapeptide ELP (an ELP 10-mer) is constructed. The ELP 10-mer may be oligomerized or polymerized up to 18 times to create a library of ELPs with precisely specified molecular masses (10-, 20-, 30-, 60-, 90-, 120-, 150-, and 180- mers). The ELP polymers or oligomers may then be fused to the C- or N-terminus of the peptide active therapeutic agent, to form the peptide active therapeutic agent-ELP construct. A second peptide active therapeutic agent may be fused to the ELP component of the fusion protein construct, providing a ternary fusion. Optionally, one or more spacers may be used to separate the ELP component from the peptide active therapeutic agent(s).

[0087] The invention thus affords a peptide active therapeutic agent-ELP construct in which the peptide active therapeutic agent may be a natural or synthetic version of any of a wide variety of endogenous molecules, or alternatively a non-naturally-occurring peptide species, or a functional equivalent of any of the foregoing.

[0088] The peptide active therapeutic agent-ELP constructs of the invention overcome the major deficiency of peptide active therapeutic agents when given parenterally, namely, that such peptides are easily metabolized by plasma proteases. The oral route of administration of peptide active therapeutic agents is even more problematic because in addition to proteolysis in the stomach, the high acidity of the stomach destroys such peptide active therapeutic agents before they reach their intended target tissue. Peptides and peptide fragments produced by the action of gastric and pancreatic enzymes are cleaved by exo and endopeptidases in the intestinal brush border membrane to yield di- and tripeptides, and even if proteolysis by pancreatic enzymes is avoided, polypeptides are subject to degradation by brush border peptidases. Any of the peptide active therapeutic agents that survive passage through the stomach are further subjected to metabolism in the intestinal mucosa where a penetration barrier prevents entry into the cells. The peptide active therapeutic agent-ELP constructs of the invention overcome such deficiencies, and provide compositional forms of the peptide active therapeutic agent having enhanced efficacy, in bioavailability, bio-unavailability, therapeutic half-life, degradation assistance, etc.

[0089] The peptide active therapeutic agent-ELP constructs of the invention thus enable oral and parenteral dose forms, as well as various other dose forms, by which peptide active therapeutic

agents can be utilized in a highly effective manner. For example, such constructs enable dose forms that achieve high mucosal absorption, and the concomitant ability to use lower doses to elicit an optimum therapeutic effect.

[0090] The ELP/peptide active therapeutic agent construct may also include a spacer as a moiety in the construct. The spacer may be of any suitable type, and may be a peptide spacer, or alternatively a non-peptide chemical moiety.

[0091] Peptide spacers may be protease-cleavable or non-cleavable. By way of example, cleavable peptide spacer species include, without limitation, in a peptide sequences recognized by proteases of varying type, such as thrombin, factor Xa, plasmin (blood proteases), metalloproteases, cathepsins (e.g., GFLG, etc.), and proteases found in other corporeal compartments. The non-cleavable spacer may likewise be of any suitable type, including, for example, non-cleavable spacer moieties having the formula $[(\text{Gly})_n\text{-Ser}]_m$ where n is from 1 to 4, inclusive, and m is from 1 to 4, inclusive.

[0092] Non-peptide chemical spacers can additionally be of any suitable type, including for example, by functional linkers described in Bioconjugate Techniques, Greg T. Hermanson, published by Academic Press, Inc., 1995, and those specified in the Cross-Linking Reagents Technical Handbook, available from Pierce Biotechnology, Inc. (Rockford, Illinois), the disclosures of which are hereby incorporated by reference, in their respective entireties. Illustrative chemical spacers include homobifunctional linkers that can attach to amine groups of Lys, as well as heterobifunctional linkers that can attach to Cys at one terminus, and to Lys at the other terminus, and other bifunctional linkers that can link proteins to the Fc region of antibodies, in which the antibody's carbohydrate is first oxidized to a diol or aldehyde.

[0093] The peptide active therapeutic agent-ELP constructs of the invention have application in prophylaxis or treatment of condition(s) or disease state(s). Although such constructs are described herein with reference to peptide active therapeutic agents having utility for animal subjects, the invention also contemplates peptide active therapeutic agent-ELP constructs having utility for prophylaxis or treatment of condition(s) or disease state(s) in plant systems. By way of example, the peptide component of the peptide active therapeutic agent-ELP construct having such plant utility may have insecticidal, herbicidal, fungicidal, and/or pesticidal efficacy.

[0094] A further aspect of the invention relates to gene therapy utilizing fusion gene therapeutic compositions of the invention, in conjunction with vectors of any suitable type, e.g., AAV, vaccinia, pox virus, HSV, retrovirus, lipofection, RNA transfer, etc.

[0095] In therapeutic usage, the present invention contemplates a method of treating an animal subject having or latently susceptible to such condition(s) or disease state(s) and in need of such treatment, including administering to such animal an effective amount of a peptide active therapeutic agent-ELP construct of the present invention which is therapeutically effective for said condition or disease state.

[0096] Animal subjects to be treated by the peptide active therapeutic agent-ELP constructs of the present invention include both human and non-human animal (e.g., bird, dog, cat, cow, horse) subjects, and preferably are mammalian subjects, and most preferably human subjects.

[0097] Depending on the specific condition or disease state to be combated, animal subjects may be administered peptide active therapeutic agent-ELP constructs of the invention at any suitable therapeutically effective and safe dosage, as may readily be determined within the skill of the art, without undue experimentation, based on the disclosure herein.

[0098] In general, suitable doses of the peptide active therapeutic agent in the peptide active therapeutic agent-ELP construct for achievement of therapeutic benefit, can for example be in a range of 1 microgram (μg) to 100 milligrams (mg) per kilogram body weight of the recipient per day, preferably in a range of 10 μg to 50 mg per kilogram body weight per day and most preferably in a range of 10 μg to 50 mg per kilogram body weight per day. The desired dose can be presented as two, three, four, five, six, or more sub-doses administered at appropriate intervals throughout the day. These sub-doses can be administered in unit dosage forms, for example, containing from 10 μg to 1000 mg, preferably from 50 μg to 500 mg, and most preferably from 50 μg to 250 mg of active ingredient per unit dosage form. Alternatively, if the condition of the recipient so requires, the doses may be administered as a continuous infusion.

[0099] The mode of administration and dosage forms will of course affect the therapeutic amount of the peptide active therapeutic agent that is desirable and efficacious for a given treatment application.

[00100] For example, orally administered dosages can be at least twice, e.g., 2-10 times, the dosage levels used in parenteral administration methods, for the same peptide active therapeutic agent.

[0100] The peptide active therapeutic agent-ELP constructs of the invention may be administered per se as well as in forms of such constructs including pharmaceutically acceptable esters, salts, and other physiologically functional derivatives thereof.

[0101] The present invention also contemplates pharmaceutical formulations, both for veterinary and for human medical use, which include peptide active therapeutic agent-ELP constructs of the invention.

[0102] In such pharmaceutical and medicament formulations, the peptide active therapeutic agent-ELP construct can be utilized together with one or more pharmaceutically acceptable carrier(s) therefore and optionally any other therapeutic ingredients. The carrier(s) must be pharmaceutically acceptable in the sense of being compatible with the other ingredients of the formulation and not unduly deleterious to the recipient thereof. The peptide active therapeutic agent-ELP construct is provided in an amount effective to achieve the desired pharmacological effect, as described above, and in a quantity appropriate to achieve the desired daily dose.

[0103] The formulations of the peptide active therapeutic agent-ELP constructs include those suitable for parenteral as well as non-parenteral administration, and specific administration modalities include oral, rectal, buccal, topical, nasal, ophthalmic, subcutaneous, intramuscular, intravenous, transdermal, intrathecal, intra-articular, intra-arterial, sub-arachnoid, bronchial, lymphatic, vaginal, and intra-uterine administration. Formulations suitable for oral and parenteral administration are preferred.

[0104] When the peptide active therapeutic agent-ELP construct is utilized in a formulation including a liquid solution, the formulation advantageously can be administered orally or parenterally. When the peptide active therapeutic agent-ELP construct is employed in a liquid suspension formulation or as a powder in a biocompatible carrier formulation, the formulation may be advantageously administered orally, rectally, or bronchially.

[0105] When the peptide active therapeutic agent-ELP construct is utilized directly in the form of a powdered solid, the active agent can be advantageously administered orally. Alternatively, it may be administered bronchially, via nebulization of the powder in a carrier gas, to form a gaseous dispersion of the powder which is inspired by the patient from a breathing circuit comprising a suitable nebulizer device.

[0106] The formulations comprising the peptide active therapeutic agent-ELP constructs of the present invention may conveniently be presented in unit dosage forms and may be prepared by any

of the methods well known in the art of pharmacy. Such methods generally include the step of bringing the peptide active therapeutic agent-ELP construct(s) into association with a carrier which constitutes one or more accessory ingredients. Typically, the formulations are prepared by uniformly and intimately bringing the peptide active therapeutic agent-ELP construct(s) into association with a liquid carrier, a finely divided solid carrier, or both, and then, if necessary, shaping the product into dosage forms of the desired formulation.

[0107] Formulations of the present invention suitable for oral administration may be presented as discrete units such as capsules, cachets, tablets, or lozenges, each containing a predetermined amount of the active ingredient as a powder or granules; or a suspension in an aqueous liquor or a non-aqueous liquid, such as a syrup, an elixir, an emulsion, or a draught.

[0108] A tablet may be made by compression or molding, optionally with one or more accessory ingredients. Compressed tablets may be prepared by compressing in a suitable machine, with the peptide active therapeutic agent-ELP construct(s) being in a free-flowing form such as a powder or granules which optionally is mixed with a binder, disintegrant, lubricant, inert diluent, surface active agent, or discharging agent. Molded tablets comprised of a mixture of the powdered peptide active therapeutic agent-ELP construct(s) with a suitable carrier may be made by molding in a suitable machine.

[0109] A syrup may be made by adding the peptide active therapeutic agent-ELP construct(s) to a concentrated aqueous solution of a sugar, for example sucrose, to which may also be added any accessory ingredient(s). Such accessory ingredient(s) may include flavorings, suitable preservative, agents to retard crystallization of the sugar, and agents to increase the solubility of any other ingredient, such as a polyhydroxy alcohol, for example glycerol or sorbitol.

[0110] Formulations suitable for parenteral administration conveniently comprise a sterile aqueous preparation of the peptide active therapeutic agent-ELP construct(s), which preferably is isotonic with the blood of the recipient (e.g., physiological saline solution). Such formulations may include suspending agents and thickening agents or other microparticulate systems which are designed to target the peptide active therapeutic agent to blood components or one or more organs. The formulations may be presented in unit-dose or multi-dose form.

[0111] Nasal spray formulations comprise purified aqueous solutions of the peptide active therapeutic agent-ELP construct(s) with preservative agents and isotonic agents. Such formulations are preferably adjusted to a pH and isotonic state compatible with the nasal mucus membranes.

[0112] Formulations for rectal administration may be presented as a suppository with a suitable carrier such as cocoa butter, hydrogenated fats, or hydrogenated fatty carboxylic acid.

[0113] Ophthalmic formulations are prepared by a similar method to the nasal spray, except that the pH and isotonic factors are preferably adjusted to match that of the eye.

[0114] Topical formulations comprise the peptide active therapeutic agent-ELP construct(s) dissolved or suspended in one or more media, such as mineral oil, petroleum, polyhydroxy alcohols, or other bases used for topical pharmaceutical formulations.

[0115] In addition to the aforementioned ingredients, the formulations of this invention may further include one or more accessory ingredient(s) selected from diluents, buffers, flavoring agents, disintegrants, surface active agents, thickeners, lubricants, preservatives (including antioxidants), and the like.

[0116] The features and advantages of the present invention are more fully shown with respect to the following non-limiting examples.

EXAMPLES

[0117] Features of the invention are more fully shown with illustrative reference to experiments involving the expression of fusion proteins containing various different recombinant proteins, such as thioredoxin, tendamistat, insulin, T20 protein, interferon, tobacco etch virus protease, small heterodimer partner orphan receptor, androgen receptor ligand binding protein, glucocorticoid receptor ligand binding protein, estrogen receptor ligand binding protein, G proteins, 1-deoxy-D xylulose 5-phosphate reductoisomerase, angiostatin (K1-3), blue fluorescent protein (BFP), calmodulin (CaM), chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), interleukin 1 receptor antagonist (IL-1Ra), luciferase, tissue transglutaminase (tTg), morphine modulating neuropeptide (MMN), neuropeptide Y (NPY), orexin-B, leptin, ACTH, calcitonin, adrenomedullin (AM), parathyroid hormone (PTH), defensin and growth hormone that are fused to various different ELP sequences.

EXAMPLE 1:**PRODUCTION AND PURIFICATION OF PROTEINS AND LONG PEPTIDES**

[0118] In the case studies presented, *E. coli* strain BL21 star (Invitrogen) containing ELP-(TEV)-peptide/protein constructs were grown in media supplemented with antibiotic at 37°C for 24 hrs without induction. The culture was harvested and resuspended in 50 mM Tris-HCL pH 8.0 and 1 mM EDTA. Cells were lysed by ultrasonic disruption on ice. Cell debris was removed by centrifugation at 20,000g at 4°C for 30 minutes. Inverse temperature transition was induced by adding NaCl to a final concentration of 1.5 M to the lysate at 25°C, followed by centrifugation at 20,000g for 15 minutes at 25°C. The resulting pellet contained ELP-(TEV)-peptide/protein fusion and non-specifically NaCl precipitated proteins. The pellet was resuspended in 40 ml ice-cold buffer and centrifuged at 20,000 g, 4°C for 15 minutes to remove non-specific insoluble proteins. The temperature transition cycle was repeated three additional times to increase the purity of ELP-TEV fusion protein and to reduce the final volume to less than 5 ml.

[0119] Separation of the peptide/protein from ELP was achieved by adding ELP-TEV protease and incubating at 25°C for 18 hrs. Cleaved peptide/protein was further purified from ELP and ELP-TEV protease using a final temperature transition in the presence of 0.5 M NaCl followed by centrifugation at 10,000g at room temperature. NaCl transitioned ELP, ELP-TEV protease and non-cleaved ELP-peptide/protein are found in the insoluble fraction while the peptide/protein remained in the soluble fraction. HPLC and liquid chromatography mass spectrum (LC-MS) analysis was carried out to test how accurately TEV cleaved ELP-(TEV)-peptide/protein and final purity of the peptide/protein. The concentration of ELP-(TEV)-peptide/protein, ELP and purified peptide/protein was determined spectrophotometrically using extinction coefficients calculated by ExPASy tools ProtParam. (19th Annual American Peptide Symposium, June 2005; poster presentation.)

Production of a 37 amino acid peptide

[0120] A 37 amino acid peptide was expressed and purified using the above (deltaPhase™) system. The expressed ELP-peptide fusion was purified through several rounds of transitions. The purified fusion was incubated with TEV protease to cleave the peptide. The TEV protease was prepared as an ELP fusion in a separate experiment which allowed removal from solution along with the cleaved ELP after incubation. Results are shown in Fig. 1, where M is the molecular weight marker, S is the lysate after sonication, P is the pellet from centrifugation (pre-transition), L is the soluble lysate, and T_n is the pellet from the nth transition.

[0121] The resulting peptide had greater than 90% purity with a minor deamidated impurity, as is seen in Fig. 2, the graph results of confirmation of molecular weight and purity by LC-MS.

Rapid Production of a Series of Peptide Variants

[0122] The throughput and purity possible for a series of peptides was then determined. The results, shown in Table 1, demonstrate the ability to produce consistent results across a series of peptides. Previously, the limits of chemical synthesis limited peptide production to one peptide every 3 to 6 weeks, which limited the rate of peptide optimization. Using the deltaPhase™ System, as set forth above, the following six peptides could be produced in less than two weeks. Given the ability to parallel process this system, the throughput could have easily been increased to several hundred in several weeks.

Table 1: Yield and Purity for a Series of Peptide Variants.

Peptide	Final Yield		Peptide Purity Level (LC-MS)
	ELP-Peptide (mg/L)	Peptide (mg/L)	
Core	280	18	94%
Variant 1	389	32	93%
Variant 2	194	20	90%
Variant 3	195	21	98%
Variant 4	267	32	92%
Variant 5	195	20	92%

EXAMPLE 2:

FUSION PROTEINS CONTAINING THIOREDOXIN AND/OR TENDAMISTAT

[0123] Thioredoxin and tendamistat exemplify two limiting scenarios of protein expression: (1) the peptide active therapeutic agent over-expresses at high levels and is highly soluble (thioredoxin), and (2) the peptide active therapeutic agent is expressed largely as insoluble inclusion bodies (tendamistat).

[0124] The thioredoxin-ELP fusion protein exhibited only a small increase in T_i (1-2 °C) compared to free ELP, while the tendamistat fusion displayed a more dramatic 15°C reduction in T_i . This shift was identical for both the ternary (thioredoxin-ELP-tendamistat) and binary (ELP-tendamistat) constructs, indicating that the T_i shift was associated specifically with tendamistat. These observations are consistent with the conclusion that the decreased T_i was due to interactions

between the ELP chain and solvent-exposed hydrophobic regions in tendamistat, whereas, for the highly soluble thioredoxin, these hydrophobic interactions were negligible. Moreover, with highly soluble proteins only a small perturbation of T_i relative to the free ELP is likely to be introduced upon fusion with an ELP tag.

[0125] In order to demonstrate fundamental concepts, a gene encoding an ELP sequence was synthesized and ligated into two fusion protein constructs. In the first construct, an ELP sequence was fused to the C-terminus of *E. coli* thioredoxin, a 109 residue protein that is commonly used as a carrier to increase the solubility of target recombinant proteins. In the second, more complex construct, tendamistat, a 77 residue protein inhibitor of α -amylase, was fused to the C-terminus of a thioredoxin-ELP fusion, forming a ternary fusion.

[0126] Previous studies by Urry and colleagues have shown that two ELP-specific variables, guest residue(s) composition (i.e., identity and mole fraction of X in the VPGXG monomer) and chain length of the ELP profoundly affect the transition temperature, and thereby permit the peptide active therapeutic agent-ELP construct to be characterized by the T_i .

[0127] A gene was synthesized encoding an ELP sequence (SEQ ID NO: 13) with guest residues valine, alanine, and glycine in the ratio 5:2:3, with a predicted T_i of $\sim 40^\circ\text{C}$ in water. The synthetic gene, which encoded 10 VPGXG pentapeptide repeats (the "10-mer"), was oligomerized up to 18 times to create a library of genes encoding ELPs with precisely-specified molecular weights (MWs) ranging from 3.9 to 70.5 kDa. Thioredoxin was expressed as an N-terminal fusion with the 10-, 20-, 30-, 60-, 90-, 120-, 150-, and 180-mer ELP sequences, and tendamistat was expressed as a C-terminal fusion to thioredoxin/90-mer ELP.

[0128] The FPs were expressed in *E. coli* and purified from cell lysate either by immobilized metal affinity chromatography (IMAC) using a (histidine)₆ tag present in the fusion protein or by inverse transition cycling (described below). The purified FP was cleaved with thrombin to liberate the target protein from the ELP. The ELP was then separated from the target protein by another round of inverse transition cycling, resulting in pure target protein. For each construct, the purified FP, target protein, and ELP were characterized by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), which confirmed protein purity, verified completeness of thrombin cleavage, and showed that the migration of each protein was consistent with its predicted size (results not shown).

[0129] The inverse transition of the fusion protein so formed can be spectrophotometrically characterized by monitoring solution turbidity as a function of temperature, due to aggregation of the ELP-containing fusion protein as it undergoes the transition. As the temperature is raised up to a critical temperature, the solution remains clear. Further increase in temperature results in a sharp increase in turbidity over a ~2 °C range to a maximum value ($OD_{350} \sim 2.0$). The T_i , defined as the temperature at the midpoint of the spectrophotometrically-observed transition, is a convenient parameter to describe this process.

[0130] The inverse transition of free ELP, thioredoxin-ELP fusion, ELP-tendamistat fusion, and ternary thioredoxin-ELP-tendamistat fusion in PBS were studied. The T_i was 51°C for free ELP and 54°C for the thioredoxin fusion, showing that the T_i is only slightly affected by fusion to thioredoxin. Thioredoxin-ELP produced by cleavage from the ternary tendamistat fusion had a higher T_i compared to thioredoxin-ELP produced directly (60°C vs. 54°C), presumably due to differences in the leader and trailer amino acid sequences immediately adjacent to the ELP sequence. The transition profiles of ELP-tendamistat and the thioredoxin-ELP-tendamistat were nearly identical, with a T_i of 34 °C. Aggregation of the FPs was reversible, and the aggregates were resolubilized completely upon lowering the temperature below the T_i . However, resolubilization kinetics were slower for ELP-tendamistat and thioredoxin-ELP-tendamistat fusions, typically requiring 5 to 10 minutes versus only a few seconds for free ELP and thioredoxin-ELP. Thioredoxin and tendamistat controls exhibited no change in absorbance with increasing temperature, indicating that the thermally-induced aggregation observed for the fusion proteins was due to the inverse transition of the ELP carrier. Typically, the inverse transition of the fusion proteins was also slightly broader than that of free ELP, and small upper and lower shoulders were observed in their turbidity profiles.

[0131] In studies by Urry and colleagues, a decrease in T_i was observed with increasing chain length, and the effect of ELP MW on the inverse transition of FPs was also investigated. The T_i of a set of thioredoxin-FPs was determined as a function of the MW of the ELP carrier, which ranged from 12.6 to 71.0 kDa. The T_i 's of the higher MW fusion proteins approached the design target temperature of 40°C (42°C for the 71 kDa ELP), while the T_i 's for the lower MW fusions were significantly greater (e.g., 77°C for the 12.6 kDa ELP).

[0132] In addition to ELP-specific variables that affect the T_i (i.e., guest residue composition and MW), the T_i can be further modulated for a given ELP by several extrinsic factors, such as the choice of solvent, ELP concentration, and ionic strength. Controlling the ionic strength, in particular, allows the T_i to be tuned over a 50°C range, and thereby provides a convenient method

to optimize the T_i of a given ELP for a specific application. Manipulating the solution temperature and ionic strength also provides experimental flexibility in inducing the inverse transition for a specific ELP by several methods: (1) by increasing the solution temperature above the T_i at a given ionic strength, (2) by increasing the ionic strength isothermally to reduce the T_i below solution temperature, or (3) by simultaneously changing the solution temperature and ionic strength.

[0133] The specific activity of the thioredoxin/60-mer FP, determined by an insulin reduction assay, was identical to that of commercially-available *E. coli* thioredoxin (results not shown), indicating that below the T_i , the ELP had no effect on thioredoxin activity. For the ternary thioredoxin-ELP-tendamistat fusion, an α -amylase inhibition assay showed that the thioredoxin/90-mer ELP carrier reduced the α -amylase inhibition activity of tendamistat by 2-fold (results not shown). However, after thrombin cleavage and purification of tendamistat from the thioredoxin-ELP carrier, the activity of purified tendamistat was indistinguishable from recombinant tendamistat, which was independently purified by IMAC.

[0134] The application of inverse transition cycling for protein purification requires that the phase transition of the ELP does not denature the target protein. The aggregation, resolubilization, and functional activity of the thioredoxin/60-mer ELP fusion upon thermally cycling in 1.5 M NaCl were therefore monitored. 1.5 M NaCl was added to the buffer simply to lower the T_i (from 62°C in water to 27°C) so that the FP would undergo its inverse transition in each thermal cycle between the experimentally-convenient temperatures of 24 and 35°C. Before commencing thermal cycling, the solution temperature of 24°C was below the T_i of the thioredoxin-FP, and the protein solution exhibited no detectable turbidity. The thioredoxin activity of the fusion protein was initially assayed at this temperature to establish a baseline. Upon increasing the temperature to 35 °C, the fusion protein aggregated, resulting in increased turbidity ($OD_{350} \sim 2.0$). After lowering the temperature to 24°C, the solution cleared completely, indicating that the fusion protein had resolubilized. An aliquot was removed and assayed for thioredoxin activity, which was found to be identical to the initial value. This thermal cycling process was repeated twice. No change in activity was observed at 24°C after each thermal cycle, which confirmed that the small temperature change and the resulting aggregation/resolubilization had no effect on protein stability and function. In addition, resolubilization and recovery of the aggregated fusion protein was quantitative and complete after lowering the temperature to 24°C.

[0135] Six thioredoxin-FPs, where each fusion protein contained a C-terminal 30-, 60-, 90-, 120-, 150-, or 180-mer ELP tag, and the thioredoxin/90-mer ELP/tendamistat fusion protein were

purified from cell lysate by inverse transition cycling, achieved by repeated centrifugation at conditions (i.e., NaCl concentration and temperature) alternating above and below the transition temperature.

[0136] Before purification, the induced *E. coli* were harvested from culture media by centrifugation, resolubilized in a low salt buffer (typically PBS), and lysed by ultrasonic disruption. After high-speed centrifugation to remove insoluble matter, polyethylenimine was added to the lysate to precipitate DNA, yielding soluble lysate. Inverse transition cycling was then initiated by adding NaCl and/or increasing the solution temperature to induce the inverse transition of the FP, causing the solution to become turbid as a result of aggregation of the FP. The aggregated fusion protein was separated from solution by centrifugation at a temperature greater than the T_i , and a translucent pellet formed at the bottom of the centrifuge tube. The supernatant, containing contaminating *E. coli* proteins, was decanted and discarded. The pellet was redissolved in a low ionic strength buffer at a temperature below the T_i of the ELP, and centrifuged at low temperature to remove any remaining insoluble matter. Although additional rounds of inverse transition cycling were undertaken, the level of contaminating proteins was below the detection limit of SDS-PAGE after a single round of inverse transition cycling.

[0137] A study of thioredoxin specific activity at each stage of purification of the thioredoxin/ELP fusion protein, as well as a determination of the total protein as estimated by BCA assay, showed that approximately 20% of the total protein in the soluble lysate (1) was precipitated in the first round of inverse transition purification (3), and the remaining soluble protein was decanted and discarded (2). The low thioredoxin activity measured in the supernatant, a portion of which is contributed by native *E. coli* thioredoxin, confirmed that this fraction primarily contained contaminating host proteins. The thioredoxin specific activity of the resolubilized protein approached that of commercially-available thioredoxin (data not shown), which confirmed that one round of inverse transition cycling resulted in complete purification. A second round of purification resulted in no detectable increase in thioredoxin specific activity (data not shown). The total thioredoxin activity after several rounds of inverse transition purification was experimentally-indistinguishable from that of the cell lysate, indicating negligible loss of target protein in the discarded supernatant. These results quantitatively confirmed the high purity and efficient recovery of the thioredoxin-FP, and further demonstrated that functional activity of thioredoxin is fully retained after undergoing several rounds of inverse transition cycling.

[0138] Protein yields for the thioredoxin fusion constructs were typically greater than 50 milligrams of purified fusion protein per liter culture. It was found that the total gravimetric yield

of fusion protein decreased with increasing ELP length, with the 30-mer (MW = 12.6 kDa) averaging ~70 mg/L and the 180-mer (MW = 71.0 kDa) averaging ~50 mg/L. Expression levels of soluble tendamistat were slightly larger for the ternary thioredoxin-ELP-tendamistat fusion (45 mg/L ternary fusion, or 7 mg/L tendamistat) compared to its fusion with thioredoxin only (10 mg/L thioredoxin-tendamistat fusion, 4 mg/L tendamistat).

[0139] As described hereinabove, two recombinant proteins, thioredoxin and tendamistat, fused to an environmentally-responsive ELP sequence, were expressed and a gentle, one-step separation of these fusion proteins from other soluble *E. coli* proteins was achieved by exploiting the inverse transition of the ELP sequence. Thioredoxin and tendamistat were selected as target proteins because they exemplify two limiting scenarios of soluble protein expression: (1) the target protein over-expresses at high levels and is highly soluble (thioredoxin), and (2) the protein is expressed largely as insoluble inclusion bodies (tendamistat). However, proteins representative of this latter class must exhibit some level of expression as soluble protein to be purified by inverse transition cycling.

[0140] Thioredoxin is expressed as soluble protein at high levels in *E. coli*, and is therefore a good candidate for determining whether the reversible, soluble-insoluble inverse transition of the ELP tag would be retained in a fusion protein. In contrast, tendamistat was selected as the other test protein because it is largely expressed as insoluble protein in inclusion bodies. Although fusion with thioredoxin promotes the soluble expression of target proteins, only 5-10% of over-expressed thioredoxin-tendamistat fusion protein was recovered as soluble and functionally-active protein.

[0141] The ELP polypeptide tag used for thermally-induced, phase separation of the target recombinant protein was derived from polypeptide repeats found in mammalian elastin. Because the phase transition of ELPs is the fundamental basis of protein purification by inverse transition cycling, specifying the transition temperature is the primary objective in the design of an ELP tag.

[0142] Previous studies by Urry and colleagues have shown that the fourth residue (X) in the polypentapeptide sequence, VPGXG, can be altered without eliminating the formation of the β -turn, a structure that is advantageous to the inverse transition. These studies also showed that the T_i is a function of the hydrophobicity of the guest residue. Therefore, by varying the identity of the guest residue(s) and their mole fraction(s), ELP copolymers can be synthesized that exhibit an inverse transition over a 0-100°C range. Based on these results, an amino acid sequence was selected to result in a predicted T_i of ~40 °C in water, so that the ELP carrier would remain soluble in *E. coli* during culture but could be aggregated by a small increase in temperature after cell lysis.

[0143] In addition to the amino acid sequence, it is known that T_i also varies with ELP chain length. The design therefore incorporated precise control of molecular weight by a gene oligomerization strategy so that a library of ELPs with systematically varied molecular weight could be synthesized. The T_i 's of the higher molecular weight ELPs approached the target temperature, with an experimentally-observed T_i of 42°C for the thioredoxin/180-mer fusion (at 25 μ M in PBS). However, the T_i increased dramatically with decreasing MW. In low ionic strength buffers, the T_i 's of the lower molecular weight ELPs are too high for protein purification, and would consequently require a high concentration of NaCl to decrease the T_i to a useful temperature. ELP chain length is also important with respect to protein yields. In addition to the decreased total yield of expressed fusion protein observed with increasing ELP MW, the weight percent of target protein versus the ELP also decreases as the MW of the ELP carrier increases. Therefore, the design of the ELP tags for purification preferably maximizes target protein expression by minimizing the ELP molecular weight, while retaining a target T_i near 40°C through the incorporation of a larger fraction of hydrophobic guest residues in the ELP sequence.

[0144] The thioredoxin-ELP fusion as described hereinabove exhibited only a small increase in T_i (1-2°C) compared to free ELP, while the tendamistat-ELP fusion displayed a more dramatic 15°C reduction in T_i . This shift was identical for both the ternary (thioredoxin-ELP-tendamistat) and binary (ELP-tendamistat) constructs, indicating that the T_i shift is associated specifically with tendamistat. These observations suggested that the decreased T_i was due to interactions between the ELP chain and solvent-exposed hydrophobic regions in tendamistat, whereas, for the highly soluble thioredoxin, these hydrophobic interactions were negligible. Although this shift in T_i added complexity to the design of ELP carriers for inverse transition purification of proteins containing a significant fraction of exposed hydrophobic area, for highly soluble proteins only a small perturbation of T_i relative to the free ELP is likely to be introduced upon fusion with an ELP tag.

[0145] Standard molecular biology protocols were used for gene synthesis and oligomerization of the ELP tags. The synthetic gene for the 10-mer polypentapeptide VPGXG ELP was constructed from four 5'-phosphorylated, PAGE-purified synthetic oligonucleotides (Integrated DNA Technologies, Inc.), ranging in size from 86 to 97 bases. The oligonucleotides were annealed to form double-stranded DNA spanning the ELP gene with EcoRI and HindIII compatible ends. The annealed oligonucleotides were then ligated, using T4 DNA ligase, into EcoRI/HindIII linearized and dephosphorylated pUC-19 (NEB, Inc.). Chemically competent *E. coli* cells (XL1-Blue) were transformed with the ligation mixture, and incubated on ampicillin-containing agar plates. Colonies were initially screened by blue-white screening, and subsequently by colony PCR to

verify the presence of an insert. The DNA sequence of a putative insert was verified by dye terminator DNA sequencing (ABI 370 DNA sequencer).

[0146] First, a 20-mer ELP gene was created by ligating a 10-mer ELP gene into a vector containing the same 10-mer ELP gene. The 20-mer gene was similarly combined with the original 10-mer gene to form a 30-mer gene. This combinatorial process was repeated to create a library of genes encoding ELPs ranging from 10-mer to 180-mer polypentapeptides. For a typical polymerization or oligomerization, the vector was linearized with PflMI and enzymatically dephosphorylated. The insert was doubly digested with PflMI and BglII, purified by agarose gel electrophoresis (Qiaex II Gel Extraction Kit, Qiagen Inc.), ligated into the linearized vector with T4 DNA ligase, and transformed into chemically competent *E. coli* cells. Transformants were screened by colony PCR, and further confirmed by DNA sequencing.

[0147] For the thioredoxin fusion proteins, pET-32b expression vector (Novagen Inc.) was modified to include an SfiI restriction site and a transcriptional stop codon downstream of the thioredoxin gene. For the ternary tendamistat fusion, a previously constructed pET-32a based plasmid containing a gene for a thioredoxin-tendamistat fusion was modified to contain an SfiI restriction site in two alternate locations, upstream or downstream of the thrombin recognition site. ELP gene segments, produced by digestion with PflMI and BglII, were then ligated into the SfiI site of each modified expression vector. Cloning was confirmed by colony PCR and DNA sequencing.

[0148] The expression vectors were transformed into the expression strains BLR(DE3) (for thioredoxin fusions) or BL21-trxB(DE3) (for tendamistat fusion) (Novagen, Inc.). Shaker flasks with 2xYT media, supplemented with 100 µg/ml ampicillin, were inoculated with transformed cells, incubated at 37°C with shaking (250 rpm), and induced at an OD₆₀₀ of 0.8 by the addition of isopropyl α-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The cultures were incubated an additional 3 hours, harvested by centrifugation at 4°C, resolubilized in low ionic strength buffer (~1/30 culture volume), and lysed by ultrasonic disruption at 4°C. The lysate was centrifuged at ~20,000 x g at 4°C for 15 minutes to remove insoluble matter. Nucleic acids were precipitated by the addition of polyethylenimine (0.5% final concentration), followed by centrifugation at ~20,000 x g at 4°C for 15 minutes. Soluble and insoluble fractions of the cell lysate were then characterized by sodium-dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

[0149] The thioredoxin-ELP fusions, which contained a (His)₆ tag, were purified by immobilized metal ion affinity chromatography (IMAC) using a nickel-chelating nitrilotriacetic derivatized resin

(Novagen Inc.) or alternatively by inverse transition cycling. The tendamistat-ELP fusion was purified exclusively by inverse transition cycling. For purification by inverse transition cycling, FPs were aggregated by increasing the temperature of the cell lysate to ~ 45°C and/or by adding NaCl to a concentration ~ 2 M. The aggregated fusion protein was separated from solution by centrifugation at 35-45 °C at 10-15,000 x g for 15 minutes. The supernatant was decanted and discarded, and the pellet containing the fusion protein was resolubilized in cold, low ionic strength buffer. The resolubilized pellet was then centrifuged at 4°C to remove any remaining insoluble matter.

[0150] The optical absorbance at 350 nm of ELP fusion solutions were monitored in the 4-80°C range on a Cary 300 UV-visible spectrophotometer equipped with a multi-cell thermoelectric temperature controller. The T_i was determined from the midpoint of the change in optical absorbance at 350 nm due to aggregation of FPs as a function of temperature at a heating or cooling rate of 1.5 °C min⁻¹.

[0151] SDS-PAGE analysis used precast Mini-Protein 10-20% gradient gels (BioRad Inc.) with a discontinuous buffer system. The concentration of the fusion proteins was determined spectrophotometrically using calculated extinction coefficients. Total protein concentrations were determined by BCA assay (Pierce). Thioredoxin activity was determined by a colorimetric insulin reduction assay. Tendamistat activity was determined by a colorimetric α -amylase inhibition assay (Sigma).

[0152] ELP-GFP fusion proteins were also synthesized, wherein the ELP 90-mer and 180-mer were fused either N-terminal or C-terminal to green fluorescent protein (GFP) or its variant - blue fluorescent protein (BFP). All fusion polypeptides exhibited a reversible inverse transition as characterized by UV-vis spectrophotometric measurement of turbidity as a function of temperature, as well as temperature dependent fluorescence measurement. The inverse transition of the GFP-ELP and BFP-ELP fusions, was used to purify these fusion proteins to homogeneity by ITC, and was verified by SDS-PAGE and Coomassie staining.

[0153] Standard molecular biology protocols were further used for synthesis and polymerization/oligomerization of the ELP genes with reduced ELP molecular weight (Ausubel, et al.). Monomer genes for two ELP sequences were utilized in this example.

[0154] The first, ELP1 [V₅A₂G₃-10] encoded ten Val-Pro-Gly-Xaa-Gly repeats where Xaa was Val, Ala, and Gly in a 5:2:3 ratio (SEQ ID NO: 13), respectively. The second monomer, ELP1 [V-

5] (SEQ ID NO: 14), encoded five Val-Pro-Gly-Val-Gly pentapeptides (i.e., Xaa was exclusively Val). The coding sequence for the ELP1 [V-5] monomer gene was: 5'-GTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGC-3' (SEQ ID NO: 15). The monomer genes were assembled from chemically synthesized, 5'-phosphorylated oligonucleotides (Integrated DNA Technologies, Coralville, IA), and ligated into a pUC19-based cloning vector. A detailed description of the monomer gene synthesis is presented elsewhere.

[0155] The monomer genes for both ELP sequences, ELP1 [V₅A₂G₃-10] and ELP1 [V-5], were seamlessly oligomerized by tandem repetition to encode libraries of increasing ELP molecular weight. A detailed description of the gene oligomerization, using a methodology termed "recursive directional ligation," is presented elsewhere. Briefly, an ELP gene segment (the monomer gene initially and larger multiples of the monomer in later rounds) is excised by restriction digest from its vector, purified, and ligated into a second cloning vector containing the same or a different ELP gene segment, thereby concatenating the two gene segments. This process can be repeated recursively, doubling the gene length with each round.

[0156] Different ELP constructs are distinguished here using the notation ELP_k [X_iY_j-n], where k designates the specific type of ELP repeat unit, the bracketed capital letters are single letter amino acid codes and their corresponding subscripts designate the relative ratio of each guest residue X in the repeat units, and n describes the total length of the ELP in number of the pentapeptide repeats. The two ELP constructs central to the present example are ELP1 [V₅A₂G₃-90] (35.9 kDa) (SEQ ID NO: 16) and ELP1 [V-20] (9.0 kDa) (SEQ ID NO: 17).

[0157] To produce the thioredoxin fusion proteins, genes encoding ELP1 [V₅A₂G₃-90] and ELP1 [V-20] were excised from their respective cloning vectors and separately ligated into a pET-32b expression vector (Novagen, Madison, WI), which had been previously modified to introduce a unique Sfi I site located 3' to the thioredoxin gene, a (His)₆ tag, and a thrombin protease cleavage site. The modified pET32b vector encoding free thioredoxin with no ELP tag ("thioredoxin(His₆)") and the two expression vectors encoding each fusion protein ("thioredoxin-ELP1 [V₅A₂G₃-90]" and "thioredoxin-ELP1 [V-20]") were transformed into the BLR(DE3) *E. coli* strain (Novagen).

[0158] For quantitative comparison of the protein expression levels and purification yields, the three constructs were each expressed and purified in parallel. For each sample (four samples each of thioredoxin(His₆), thioredoxin-ELP1 [V-20], and thioredoxin-ELP1 [V₅A₂G₃-90]), a 2 ml starter culture (CircleGrow media, Qbiogene, Carlsbad, CA, supplemented with 100 µg/ml ampicillin)

was inoculated with a stab taken from a single colony on a freshly streaked agar plate, and incubated overnight at 37°C with shaking at 300 rpm. To remove β -lactamase from the media, the cells were then pelleted from 500 μ l of the confluent overnight culture by centrifugation (2000 x g, 4°C, 15 min), resuspended in fresh media wash, and repelleted. After a second resuspension in fresh media, the cells were used to inoculate 50 ml expression cultures in 250 ml flasks (CircleGrow media with 100 μ g/ml ampicillin).

[0159] The culture flasks were incubated at 37°C with shaking at 300 rpm. Growth was monitored by the optical density at 600 nm, and protein expression was induced at $OD_{600} = 1.0$ by the addition of isopropyl β -thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After a further 3 hours of culture, the cells were harvested from 40 ml by centrifugation (2,000 x g, 4°C, 15 min), resuspended in 2 ml of IMAC binding buffer (5 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl, pH 7.9) for thioredoxin(His₆) or PBS (137 mM NaCl, 2.7 mM KCl, 4.2 mM Na₂HPO₄, 1.4 mM KH₂PO₄, pH 7.3) for thioredoxin-ELP1 [V-20] and thioredoxin-ELP1 [V₅A₂G₃-90], and stored frozen at -20°C until purified. The culture density at harvest was measured by OD_{600} , after 1:10 dilution in fresh buffer. The amount of plasmid DNA at harvest was quantified by UV-visible spectrophotometry following plasmid isolation (plasmid miniprep spin kit, Qiagen, Valencia, CA).

[0160] As a control for ITC purification of the thioredoxin-ELP fusion proteins, free thioredoxin was purified using standard IMAC protocols. Briefly, the thawed cells were transferred to iced 15 ml centrifuge tubes and lysed by ultrasonic disruption (Fisher Scientific 550 Sonic Dismembrator using a microtip). After transferring to 1.5 ml micro centrifuge tubes, the *E. coli* lysate was centrifuged (16,000 x g, 4°C, 30 min) to remove the insoluble cellular debris. 1 ml of the soluble cell lysate was loaded by gravity flow onto a column packed a 1 ml bed of nitrilotriacetic acid resin that had been charged with 5 ml of 50 mM NiSO₄.

[0161] After the column was washed with 15 ml of IMAC binding buffer, thioredoxin(His₆) was eluted in 6 ml of IMAC binding buffer supplemented with 250 mM imidazole. Imidazole was removed from the eluent by dialysis against a low salt buffer (25 mM NaCl, 20 mM Tris-HCl, pH 7.4) overnight using a 3,500 MWCO membrane. The IMAC purification was monitored by SDS-PAGE using precast 10-20% gradient gels (BioRad Inc., Hercules, CA) with a discontinuous buffer system.

[0162] The yield of the purified thioredoxin(His₆) was determined by spectrophotometry, using a molar extinction coefficient of thioredoxin modified to include the absorption of the single Trp

residue present in the C-terminal tag ($\epsilon_{280} = 19870 \text{ M}^{-1}\text{cm}^{-1}$ for thioredoxin(His₆) and all thioredoxin-ELP fusion proteins, independent of ELP molecular weight.

[0163] In a typical purification by ITC, the thawed cells were transferred to iced 15 ml centrifuge tubes and lysed by ultrasonic disruption (Fisher Scientific 550 Sonic Dismembrator with a microtip). After transferring to 1.5 ml micro centrifuge tubes, the *E. coli* lysate was centrifuged at 4°C for 30 min to remove the insoluble cellular debris. (All centrifugation steps during purification by ITC were performed at 16,000 x g in Eppendorf 5415C microcentrifuges.)

[0164] Polyethylenimine was added (to 0.5% w/v) to the decanted supernatant of the cell lysate to precipitate nucleic acids, which were removed by an additional 20 min centrifugation at 4°C. The supernatant was retained, and the ELP phase transition was induced by increasing the NaCl concentration by 1.3 M. The aggregated fusion protein was separated from solution by centrifugation at 33°C for 5 min, which resulted in the formation of translucent pellet at the bottom of the tube.

[0165] The supernatant was decanted and discarded, and the pellet containing the fusion protein was redissolved in an equal volume of PBS at 4°C. Any remaining insoluble matter was removed by a final centrifugation step at 4°C for 15 min, and the supernatant containing the purified fusion protein was retained. The progression of fusion protein purification was monitored by SDS-PAGE, and the protein concentrations were determined by spectrophotometry, as described above for IMAC purification.

[0166] Thioredoxin was liberated from its ELP fusion partner using thrombin protease (Novagen), which cleaved the fusion protein at a recognition site located between thioredoxin and the ELP tag. The thrombin proteolysis reaction was allowed to proceed overnight at room temperature in PBS using ~10 units of thrombin per μmol of fusion protein, which was typically at a concentration of ~100 μM . Free ELP was then separated from the cleaved thioredoxin by another round of ITC, this time retaining the supernatant that contained the product thioredoxin.

[0167] The inverse transition can be monitored by assaying solution turbidity photometrically as a function of temperature, taking advantage of the fact that increase in temperature beyond a critical point results in a sharp increase in turbidity over an approximately 2°C range to a maximum value (OD_{350} approximately 2.0), because of aggregation of the ELP. The temperature at 50% maximal turbidity, T_b , is a convenient parameter for quantitatively monitoring the aggregation process.

[0168] The temperature-dependent aggregation behaviors of the thioredoxin-ELP fusion proteins were characterized by measuring the optical density at 350 nm as a function of temperature. Fusion proteins at concentrations typical of those found in the *E. coli* lysate during protein purification (160 μ M for thioredoxin-ELP1 [V-20] and 40 μ M for thioredoxin-ELP1 [V₅A₂G₃-90]) were heated or cooled at a constant rate of 1 °C min⁻¹ in a Cary Bio-300 UV-visible spectrophotometer (Varian Instruments, Walnut Creek, CA), which was equipped with a thermoelectric temperature-controlled multicell holder. The experiments were performed in PBS variously supplemented with additional NaCl. The ELP T_i was defined as the temperature at which the optical density reached 5% of the maximum optical density at 350 nm.

[0169] Dynamic light scattering (DLS) was used to monitor the particle size distribution of the thioredoxin-ELP fusion proteins as a function of temperature and NaCl concentration. Samples were prepared to reflect the protein and solvent compositions used in the turbidity measurements described above, and were centrifuged at 4°C and 16,000 x g for 10 minutes to remove air bubbles and insoluble debris. Prior to particle size measurement, samples were filtered through a 20 nm Whatman Anodisc filter at a temperature below the T_i.

[0170] Autocorrelation functions were collected using a DynaPro-LSR dynamic light scattering instrument (Protein Solutions, Charlottesville, VA) equipped with a Peltier temperature control unit. Analysis was performed using Protein Solutions' Dynamics software version 5.26.37 using its regularization analysis for spherical particles. Light scattering data were collected at regular temperature intervals (either 1 or 2°C) as solutions were heated from 20° to 60°C. Data were collected at each temperature by ramping the cell up to the temperature of interest, allowing the sample temperature equilibrate for at least 1 minute, and collecting 10 measurements, each with a 5 second collection time.

[0171] The inverse transition of each thioredoxin-ELP fusion protein in solution was characterized by monitoring the optical density at 350 nm as a function of temperature. Because different NaCl solutions are routinely used during ITC purification to depress the T_i or isothermally trigger the inverse transition, turbidity profiles were obtained for 40 μ M thioredoxin-ELP1 [V₅A₂G₃-90] and 160 μ M thioredoxin-ELP1 [V-20] in PBS and in PBS with an additional 1M, 2M, and 3M NaCl.

[0172] Optical density at 350 nm as a function of temperature was assessed for solutions of the thioredoxin-ELP fusion proteins. The turbidity profiles were obtained for thioredoxin-ELP1 [V-20] (solid lines) and thioredoxin-ELP1 [V₅A₂G₃-90] (dashed lines) in PBS, and in PBS supplemented with 1, 2, and 3 M NaCl, while heating at a rate of 1 °C min⁻¹. The concentration of thioredoxin-

ELP1 [V₅A₂G₃-90] was 40 μ M in each of the four PBS solutions, and that of thioredoxin-ELP1 [V-20] was 160 μ M, which matched the typical concentration of each protein in the soluble cell lysate during ITC purification. All solutions showed a rapid rise in turbidity as they were heated through the T_i, but with continued heating beyond the T_i, the thioredoxin-ELP1 [V-20] solutions eventually became less turbid while the thioredoxin-ELP1 [V₅A₂G₃-90] solutions remained consistently turbid. All solutions of thioredoxin-ELP1 [V₅A₂G₃-90] cleared fully upon cooling the solution to below the T_i. However, solutions of ELP1 [V-20] cleared reversibly only if the solutions were not heated to above ~55°C, suggesting thermal denaturation of the thioredoxin fusion protein occurred above this temperature.

[0173] The protein concentrations were chosen as typical of the concentrations obtained for each fusion protein in the soluble fraction of *E. coli* lysate, the stage at which the ELP inverse transition is first induced during ITC purification. Turbidity profiles obtained directly in the *E. coli* soluble cell lysate, supplemented with 1 and 2 M NaCl, were indistinguishable from the corresponding profiles determined for the thioredoxin fusion protein as described in the preceding paragraph. (Turbidity profiles were not routinely obtained in *E. coli* lysate because of the potential for turbidity arising from thermal denaturation of *E. coli* proteins, which could not be differentiated from turbidity arising from aggregation of the ELP fusion protein.) Turbidity profiles were also obtained for each fusion protein in PBS with 1.3 M salt, which matched the conditions used for the ITC purification described below.

[0174] The heating and cooling turbidity profiles for the solution conditions used in ITC purification were determined for solutions of thioredoxin-ELP1 [V-20] (solid lines) and thioredoxin-ELP1 [V₅A₂G₃-90] (dashed lines) at lysate protein concentrations in PBS with 1.3 M NaCl, corresponding to ITC conditions used for the quantitative comparison of expression and purification. These conditions were chosen so that the maximum turbidity of the thioredoxin-ELP1 [V-20] solution occurred at the centrifugation temperature of 33°C. The solutions were heated and cooled at 1 °C min⁻¹. The slight path differences between the heating and cooling curves that were observed were primarily due to slow settling of the aggregates over time at temperatures above T_i, and to the slower kinetics of disaggregation versus aggregation as the solutions are cooled to below T_i.

[0175] The thermally induced aggregation behavior of thioredoxin-ELP1 [V₅A₂G₃-90] was similar to that of free ELPs. All four salt concentrations, as the temperature of the thioredoxin-ELP1 [V₅A₂G₃-90] solutions was increased, remain clear until they reach the ELP T_i, at which point the turbidity sharply increased. This occurred at 51, 31, 15, and 4°C in PBS with 0, 1, 2, and 3 M

added NaCl, respectively. A free thioredoxin control solution exhibited no change in turbidity with increasing temperature over this temperature range, indicating that the thermally induced aggregation observed was due to the inverse transition of the ELP tag (results not shown). As these solutions were heated further beyond the T_i , the turbidity level remained essentially constant, and was only slightly reduced by settling of the aggregates over time. Upon cooling to below the T_i , the aggregates resolubilize and the optical density returned to zero, showing that the inverse transition of the ELP1 [V₅A₂G₃-90] fusion protein was completely reversible. While increasing the NaCl concentration markedly decreases the T_i , salt has no measurable effect on the maximum optical density, on the general shape of the turbidity profiles, or on the reversibility of the aggregation.

[0176] In contrast, the phase transition behavior of thioredoxin-ELP1 [V-20] was considerably more complex than for the thioredoxin-ELP1 [V₅A₂G₃-90] fusion protein and free ELPs. Although the initial rapid rise in turbidity at the T_i (33, 17, and 4°C in PBS supplemented with 1, 2, and 3 M NaCl, respectively) was similar to the other ELP constructs, the maximum turbidity observed with each of the thioredoxin-ELP1 [V-20] solutions increased with increasing salt concentration. Furthermore, increases in temperature beyond the T_i eventually resulted in a significant decrease in turbidity. This decrease was reversible; if the solution was cooled after heating to the point of decreased turbidity, the turbidity again increased. Because the clearing phenomenon is a reversible function of temperature, it was concluded that a second, thermodynamically driven molecular rearrangement occurs with increasing temperature after the initial ELP aggregation event at T_i .

[0177] Another unique feature of the thioredoxin-ELP1 [V-20] turbidity profiles was a second increase in turbidity beginning at ~55°C, which may have been due to aggregation arising from the irreversible thermal denaturation of thioredoxin. Samples heated to less than 55°C reversibly cleared upon cooling to below the T_i , whereas samples that are heated to above 55°C, for salt concentrations of 1 M and greater, remained turbid even upon cooling to below the T_i (not shown). This phenomenon appeared to be unique to the thioredoxin-ELP1 [V-20] fusion protein, as solutions of free thioredoxin and of its fusion proteins to larger ELPs were stable to much higher temperatures (results not shown). No inverse transition was observed for thioredoxin-ELP1 [V-20] in PBS below 60°C, however, with added salt the T_i was depressed so that it occurred below the denaturation temperature in the PBS + 1, 2, and 3 M NaCl solutions.

[0178] The sizes of the fusion protein particles were measured using DLS as a function of temperature to determine the effect of temperature and salt on the particle size distribution (radius of hydration, R_h) of 40 μ M thioredoxin-ELP1 [V₅A₂G₃-90] in PBS, PBS + 1 M NaCl, and PBS+ 2

M NaCl. The sizes of thioredoxin-ELP1 [V₅A₂G₃-90] particles in PBS, PBS with 1M added NaCl, and PBS with 2M added NaCl indicated that the sharp increase in turbidity at the T_t resulted from the conversion of monomers with hydrodynamic radii (R_h) of 5.9 ± 3.9 nm to aggregates with R_h of 180 ± 62 nm. These aggregates grew with temperature until reaching a stable R_h of 2.2 ± 3.8 μm approximately 6°C above the onset of the transition. Although the T_t was depressed by the addition of NaCl, the sizes of both monomers and fully formed aggregates were not significantly affected by either the salt concentration or the temperature (outside the range immediately adjacent to the T_t), providing a rationale for the steady-state turbidity above the inverse T_t. The temperature at the onset of large aggregate formation closely matched the T_t determined by the turbidity measurements for corresponding solution conditions.

[0179] The corresponding quantitative breakdown of scattered intensity attributed to each type of particle was also studied for each of the salt concentrations investigated. When two or more phases coexist over a given temperature range, these data show shifts in the relative particle populations. It should be noted that the intensity attributed to a particular population was not linearly correlated with the mass of that population, and that calculating the relative masses of multiple particles was complicated by changes in packing density that would likely accompany the inverse phase transition. Without a more detailed understanding of how temperature affects the packing density of ELPs and ELP fusion proteins, it was not possible to make a reasonable estimate for the mass attributed to each type of particle. Given these quantitative limitations, this data nonetheless showed that at the T_t the amount of scattered light attributed to the aggregate dramatically increased at the expense of the monomer.

[0180] The data also reflected the occasional presence of both an unidentified small particle (with apparent R_h = 17 ± 31 nm, albeit highly variable) and an extremely large aggregate (with apparent R_h = 74 ± 55 μm) coexisting with the 2 μm aggregates. It is unlikely that the small particle is a true component of the aggregate suspension; rather, its presence reflects an artifact in the regularization algorithm resulting from noise in the autocorrelation function. Assignment as an analysis artifact is supported by the small particle's highly variable size and by its inconsistent presence at temperatures above the transition. Likewise, because its apparent size is much larger than can be discerned by the DLS instrument, it is also unlikely that the extremely large aggregate predicted from the data analysis represented a true species in suspension. Rather, the scattering attributed to this species may result from the coordinated slow movements of a network of smaller particles.

[0181] In contrast to thioredoxin-ELP1 [V₅A₂G₃-90], the smaller thioredoxin-ELP1 [V-20] fusion protein showed a more complicated temperature-dependent particle size distribution, which was consistent with its more complex turbidity profile.

[0182] The effect of temperature on the particle size distribution of ELP1 [V-20] in PBS + 1 M NaCl and PBS + 2 M NaCl was studied. The clearing in turbidity when the temperature was increased beyond T_t coincided with the shifting of mass from large aggregates to a new, smaller particle ($R_h = 12$ nm).

[0183] The effects of salt and temperature on the distribution of the particle R_h and the corresponding contribution of each particle population to scattered intensity of 160 μ M thioredoxin-ELP1 [V-20] in PBS with 1M and 2M added NaCl, was likewise studied. For thioredoxin-ELP1 [V-20] with 1M added salt monomers with R_h of 5.9 ± 5.1 nm were converted to aggregates with R_h of 140 ± 79 nm at 30°C, corresponding to a small shoulder that preceded a rapid increase in turbidity at T_t . Above 30°C, aggregates grew with increasing temperature (up to $R_h = 1.5 \pm 0.98$ μ m at 40°C), which was consistent with the rapid increase in turbidity observed starting at 33°C. Similar to the aggregation behavior of the large fusion protein, at temperatures greater than 40°C thioredoxin-ELP1 [V-20] in PBS with 1M added NaCl showed the presence of very large aggregates (apparent $R_h = 64 \pm 67$ μ m) that may reflect the coordinated slow movements of a network of smaller particles.

[0184] However, unlike the larger fusion protein, thioredoxin-ELP1 [V-20] also showed the consistent presence of a previously unobserved small particle at temperatures above 40°C. This particle had a R_h of 12 ± 4.9 nm, which was roughly twice that of the monomer. Yet, relative to its mean R_h , its variability was only one half that of the monomer. The size, consistency, and continuous presence of this particle above 40°C indicated that it was neither an analysis artifact resulting from noise in the autocorrelation function nor was it resolved monomer. The 12 nm particle appeared to form at the expense of mass in the aggregates initially present above T_t , as evidenced by the reduction in size and scattering intensity of the larger aggregates ($R_h = 200 \pm 210$ nm) when the 12 nm particles were present.

[0185] A similar 12 nm particle was observed when the NaCl concentration was increased to 2 M. At this NaCl concentration, the T_t was lowered to 17°C as determined by the turbidity measurements. This temperature range was limited at lower temperatures by the condensation of water vapor on the sample cuvette. Therefore, between 20°C and 30°C, the thioredoxin-ELP1 [V-20] had already transitioned into stable aggregates with average R_h of 2.4 ± 1.7 μ m. As the

samples was heated beyond $\sim 36^\circ\text{C}$, the R_h of the aggregates gradually decreased in size to 230 ± 170 nm and 12 nm particles ($R_h = 12 \pm 4.7$ nm) appeared. The percentage of scattered light attributable to the 12 nm particles also gradually increased at the expense of the shrinking larger aggregates.

[0186] Thioredoxin-ELP1 [V-20] and thioredoxin-ELP1 [V₅A₂G₃-90] were each purified by ITC from the soluble fraction of lysed *E. coli* cultures, and thioredoxin(His₆) was purified by IMAC as a control having no ELP tag. The inverse transition was induced by the addition of 1.3 M NaCl, and the centrifugation was carried out at 33°C . The smaller ELP1 [V-20] tag was successfully used to purify the fusion protein by ITC to homogeneity, with a yield and purity similar to that of the free thioredoxin purified by a conventional affinity chromatography method.

[0187] Note that the ELP tag was not stained by Coomassie, and therefore only the thioredoxin portion of the fusion protein was visible in stained gels. Qualitative comparison of the expression levels in the soluble cell lysate for thioredoxin-ELP1 [V-20] and thioredoxin-ELP1 [V₅A₂G₃-90] clearly showed that truncating the size of the ELP tag from 36 kDa to 9 kDa greatly enhanced the expression yield of the thioredoxin. Furthermore, thioredoxin-ELP1 [V-20] was expressed to a level qualitatively comparable to that of free thioredoxin. SDS-PAGE analysis also showed that there was no detectable loss to the insoluble fraction of the cell lysate for any the target proteins (results not shown).

[0188] For the ITC purifications, the ELP phase transition was triggered by adding 1.3 M additional NaCl and increasing the solution temperature to above $\sim 33^\circ\text{C}$. The cell lysates became turbid as a result of aggregation of the thioredoxin-ELP fusion proteins, which were then separated from solution by centrifugation at $\sim 33^\circ\text{C}$ to form a translucent pellet at the bottom of the centrifuge tube. SDS-PAGE showed that most contaminating *E. coli* proteins were retained in the decanted supernatant. The pellets were dissolved in PBS at $\sim 4^\circ\text{C}$, and centrifuged at low temperature ($\sim 12^\circ\text{C}$) to remove any remaining insoluble matter. The supernatants containing purified thioredoxin-ELP fusion proteins were retained. Finally, purified, free thioredoxin was obtained after cleavage of each fusion protein by thrombin at the encoded recognition site located between thioredoxin and the ELP tag, followed by a second round of ITC to remove the ELP tag from solution. Here, thrombin was retained with the target thioredoxin in the supernatant (although it was below the detection limit of Coomassie staining), however a thrombin-ELP fusion could be developed that would be removed after cleavage along with the free ELP.

[0189] These SDS-PAGE results clearly showed that thioredoxin can be purified by ITC to homogeneity, as ascertained by Coomassie staining, using the shorter, 9 kDa ELP1 [V-20]. However, differences were observed in the purification efficiency of the two ELP fusion proteins under these conditions, as qualitatively ascertained by SDS-PAGE. Recovery of thioredoxin-ELP1 [V₅A₂G₃-90] by ITC from the soluble cell lysate was essentially complete, whereas a small but significant fraction of thioredoxin-ELP1 [V-20] remained in the discarded supernatant. The level of purity obtained by ITC with the ELP1 [V-20] tag was qualitatively as good or better than that obtained by IMAC purification of the free thioredoxin, although with IMAC purification there was no detectable loss of the target protein in the column flow-through.

[0190] Using UV-visible spectrophotometry, the yield of each protein recovered by ITC or IMAC purification was quantified. Although these data described the amount of protein recovered after purification, SDS-PAGE results suggested that this quantity was nearly equal to expression yield in the soluble lysate. For this analysis, four cultures were grown in parallel under identical conditions for each of the three protein constructs. For experimental convenience, these data were obtained for 50 ml cultures, and extrapolated to yield per liter of culture. Purification of separate 1 liter cultures confirmed that the actual yields closely matched the extrapolated values (data not shown).

[0191] The total yields of thioredoxin(His₆), thioredoxin-ELP1 [V-20], and thioredoxin-ELP1 [V₅A₂G₃-90] from the 50 ml test cultures were determined, extrapolated to milligrams per liter of culture (mean $\pm$ SD, n = 4). The separate contributions of the ELP tag and thioredoxin to the yield, as calculated using their respective mass fractions of the fusion protein, were also determined for comparison. With all other experimental conditions identical, reducing the ELP tag from 36 (thioredoxin-ELP1 [V₅A₂G₃-90]) to 9 kDa (thioredoxin-ELP1 [V-20]) resulted in a near four-fold increase in the yield of the target thioredoxin.

[0192] The data showed that decreasing the molecular weight of the ELP tag can dramatically increase the yield of thioredoxin. Under experimentally identical conditions of *E. coli* culture, decreasing the ELP tag size from 36 kDa in thioredoxin-ELP1 [V₅A₂G₃-90] to 9 kDa in thioredoxin-ELP1 [V-20] increased the yield of fusion protein by 70% (82 ± 12 mg/L versus 137 ± 21 mg/L, respectively; $P < 0.005$, unpaired t test). Furthermore, since truncating the size of the ELP tag reduced its mass fraction in the fusion protein, the target protein thioredoxin (i.e., if separated from the fusion protein at the thrombin cleavage site) constituted a larger fraction of the mass in the fusion protein yield. Thus, the yield of thioredoxin was 365% greater using the smaller tag (23 ± 3.3 mg/L versus 83 ± 12 mg/L for the larger and smaller tags, respectively; $P < 0.0001$).

This yield of thioredoxin obtained by ITC using the 9 kDa tag was statistically indistinguishable from that obtained for thioredoxin expressed without an ELP tag and purified using IMAC (93 ± 13 mg/L; $P > 0.25$).

[0193] These results corroborated the SDS-PAGE results since the relative yields of thioredoxin correlated with the expression levels observed in the cell lysate. The yield of the ELP tag was the same for both fusion proteins (59 ± 8.6 mg/L for thioredoxin-ELP1 [V₅A₂G₃-90] and 54 ± 8.1 mg/L for thioredoxin-ELP1 [V-20]; $P > 0.4$). This was consistent with previous observations that the gravimetric yield of the ELP tag in thioredoxin fusion proteins was essentially constant with respect to ELP molecular weight within the ELP1 [V₅A₂G₃-90]] family of polypeptides ranging from 24 to 72 kDa.

[0194] To demonstrate the relationship between purification efficiency and ITC solution conditions, ITC purification of the thioredoxin-ELP1 [V-20] fusion protein was repeated using different combinations of salt concentration and centrifugation temperature.

[0195] SDS-PAGE analysis of the effect of NaCl concentration and centrifugation temperature on purification of thioredoxin-ELP[V-20] by ITC was carried out (SL = soluble cell lysate; S = supernatant after inverse transition of fusion protein and centrifugation to remove aggregated target protein; and P = redissolved pellet containing the purified fusion protein, after dissolution in PBS). The molar NaCl concentration and centrifugation temperature for each purification was noted. Although a high level of purity was achieved in each case, selection of an appropriate NaCl concentration and centrifugation temperature is critical to achieve complete purification efficiency.

[0196] When PBS with 1 M NaCl combined with centrifugation at 49°C was used for ITC purification, the majority of the target fusion protein was lost in the discarded supernatant. When PBS plus 2 M NaCl and a centrifugation temperature of 33 °C was used, more than half of the target protein was captured by centrifugation. Finally, using PBS with 3 M NaCl and centrifugation at 12°C, the vast majority of the target protein was successfully purified. Although the target protein was purified to homogeneity in each of these examples, these results showed that selection of salt concentration and temperature was an important factor influencing the efficiency of ITC purification.

[0197] The objective of this example was to produce an ELP tag for ITC purification that was reduced in size relative to those previously reported, and to characterize the effect of this reduction on expression levels and on purification efficiency. In a prior effort, a first generation of ELP

purification tags was developed based on a ELP1 [V₅A₂G₃-10] monomer sequence. This sequence was recursively oligomerized to create a library of synthetic genes encoding ELPs with molecular weights ranging from 4 kDa (ELP1 [V₅A₂G₃-10]) to 71 kDa (ELP1 [V₅A₂G₃-180]). This particular guest residue composition was selected based on previous studies of Urry et al., and ELPs with this composition were predicted to exhibit a T_i of ~40°C for molecular weights of ~100 kDa in water. A 40°C T_i was targeted so that the fusion proteins would remain soluble during culture at 37°C, but could be induced to reversibly aggregate through the ELP phase transition by a modest increase in salt concentration or solution temperature.

[0198] Although the T_i 's of the higher molecular weight constructs approached 40°C (T_i = 42 °C for the thioredoxin-ELP1 [V₅A₂G₃-180], with MW_{ELP} = 71 kDa, in PBS at 25 µM), the T_i of the thioredoxin-ELP1 [V₅A₂G₃] fusion proteins increased dramatically with decreasing molecular weight (T_i = 77 °C for thioredoxin-ELP1 [V₅A₂G₃-30], with MW_{ELP} = 13 kDa, under the same conditions). The high T_i 's of the lower molecular weight ELPs required the addition of a very high concentration of NaCl (> 3 M) to reduce their T_i to a useful temperature (e.g., 20-40 °C), which precluded their general use for purification by ITC because of the potential for salt-induced denaturation of target proteins. Although the larger ELP1 [V₅A₂G₃] polypeptides were successfully used to purify thioredoxin and second model target protein, tendamistat, it was observed that the yield of the fusion protein was significantly decreased as the ELP1 [V₅A₂G₃] chain length was increased.

[0199] These observations motivated the redesign of the ELP expression tag in the above experiment to reduce the size of the ELP expression tag while also depressing its T_i , so that lower molecular weight ELP tags would exhibit a T_i near 40°C at more moderate NaCl concentrations. The second monomer gene, which was newly synthesized for this study, encoded a five pentamer ELP sequence where the fourth guest residue was exclusively Val (ELP1 [V-5]). Because the Val present in ELP1 [V] was more hydrophobic than the Ala and Gly present in ELP1 [V₅A₂G₃], thioredoxin-ELP1 [V] fusion proteins were predicted to have a T_i of 40°C at smaller ELP molecular weights than for thioredoxin-ELP1 [V₅A₂G₃] fusions.

[0200] The ELP1 [V-20] sequence (four tandem repeats of the ELP1 [V-5] gene) was selected from a library of ELP1 [V-5] oligomers for further characterization at a ITC purification tag due to the empirical observation of its T_i near 40°C at lysate protein concentration with moderate (1 M) NaCl. In the present example, the thioredoxin-ELP1 [V-20] construct (MW_{ELP} = 9 kDa) was compared to the previously described thioredoxin-ELP1 [V₅A₂G₃-90] construct (MW_{ELP} = 36 kDa) because the two fusion proteins had very similar T_i 's in lysate conditions for varying NaCl

concentrations. That is, they are thermal analogs from each of the two libraries that meet the above-described desired T_1 characteristics for ITC purification tags.

[0201] Although previous observations suggested that decreasing the size of the ELP was likely to enhance the overall expression level of the fusion protein, it was not obvious, a priori, whether the decreased size of the tag would adversely affect purification of ELP fusion proteins by ITC. Therefore, in addition to its effect on the expression level of the target protein, the effect of the ELP tag length on the purification efficiency (i.e., degree of recovery) and on the purity of the target protein after ITC purification was explored.

[0202] The SDS-PAGE and spectrophotometry results showed that decreasing the ELP molecular weight from 36 kDa to 9 kDa enhanced expression of the fusion protein by nearly four-fold, and did not adversely affect the purity of the final protein under any of the solution conditions (i.e., NaCl concentration and temperature) used to induce the inverse transition. The level of expression with the ELP[V-20] tag was comparable to that of free thioredoxin, indicating that further reduction of the ELP tag would not be expected to increase the thioredoxin yield.

[0203] One possible explanation for the observed increase in thioredoxin yield as the ELP tag length was reduced is that, for given culture conditions, the mass of ELP that can be expressed by the cells is limited independent of ELP chain length. This was supported by the results, as well as by observations with other ELPs of various molecular weight. Such a limitation would likely be engendered by a metabolic factor, perhaps by an insufficient tRNA pool and/or by amino acid depletion due to the highly repetitious ELP sequence. If the mass yield of ELP is a limiting factor, then this provides a rationale for the increased thioredoxin yields with the ELP[V-20] tag. For a given gravimetric yield of ELP, decreasing the ELP chain length increases the molar yield of the fusion protein, and hence, the target protein. Furthermore, this also suggests that increasing the gravimetric yield of ELP, e.g., through supplementation of specific, ELP-related amino acids during culture, offers another potential route for improvement of the fusion protein yield.

[0204] Although the yield of the target protein was increased with the shorter ELP1 [V-20] tag, this benefit entailed a more complicated transition behavior. The efficiency of recovery with this tag depends on the solution conditions used for ITC, whereas, with the larger ELP1 [V₅A₂G₃-90] tag, recovery of the fusion protein was complete under all solution conditions (results not shown). Thus, although the truncated ELP1 [V-20] tag enabled thioredoxin to be purified to homogeneity by ITC, the efficiency of purification was sensitive to the specific conditions chosen to induce the inverse transition.

[0205] The turbidity and DLS data provided insights into the sensitivity of purification efficiency for the smaller ELP1 [V-20] tag on solution conditions. While solutions of thioredoxin-ELP1 [V₅A₂G₃-90] remained turbid at all temperatures above T_i , the turbidity profiles for thioredoxin-ELP1 [V-20], after an initial rapid rise at T_i , began to clear with further heating at a temperature above T_i . This phenomenon of clearing with increasing temperature has not been previously observed, to my knowledge, with other ELPs or ELP fusion proteins. To study this complex aggregation behavior, the sizes of the fusion protein particles were measured using dynamic light scattering as a function of temperature to determine the structural basis for the markedly different turbidity profiles of the two fusion proteins.

[0206] With increasing temperature, monomers of thioredoxin-ELP1 [V₅A₂G₃-90] went through an abrupt, discontinuous phase transition to form aggregates that persisted at all temperatures above T_i with a steady state R_h of $\sim 2 \mu\text{m}$. Because the aggregates were stable above the T_i , the aggregated protein was able to be completely recovered by centrifugation at any temperature above its T_i (or at any NaCl concentration for which the T_i was depressed to below the solution temperature).

[0207] Although thioredoxin-ELP1 [V-20] also exhibited an abrupt phase transition to form aggregates, these aggregates were not stable at all temperatures above its phase transition. As the temperature was increased beyond the T_i , small aggregates with R_h of $\sim 12 \text{ nm}$ formed at the expense of mass in the larger aggregates, which also showed a decrease in size with increasing temperature. This provides a structural rationale for the decrease in turbidity observed above the T_i of thioredoxin-ELP1 [V-20]. Upon heating to temperatures greater than T_i (beginning $\sim 10^\circ\text{C}$ above T_i for PBS with 1 M NaCl, and $\sim 15^\circ\text{C}$ above T_i for PBS with 2 M NaCl), larger scattering centers were converted to small particles that scatter light less effectively. The formation of these 12 nm particles at the expense of the larger aggregates resulted in incomplete recovery by centrifugation of the fusion protein from the soluble lysate. Thus, when ELP1 [V-20] (and potentially other small ELP tags) were used for purification of fusion proteins, it was imperative for complete protein recovery that a NaCl concentration and complimentary solution temperature be chosen such that only the larger aggregates, which are easily separable by centrifugation, were present in suspension.

[0208] On the basis of size alone, the precise structure of the 12 nm particle was not able to be predicted. However, the particle may be a micelle-like structure containing a small number of fusion protein molecules (perhaps on the order of 40 to 60) that are aggregated such that solvated thioredoxin domains encase the collapsed, hydrophobic ELP domains in the particle's core. The size of the observed particle ($R_h \approx 12 \text{ nm}$) would be consistent with such a structure, as the

hydrophilic thioredoxin “head” was ~3 nm in diameter and the hydrophobic 20 pentamer ELP “tail” was ~7 nm in length.

[0209] The proximity of the thioredoxin molecules required in such a micellular structure may also explain the irreversible aggregation that is observed at temperatures greater than ~55°C. Denaturation at this low temperature was only observed for thioredoxin fused to ELP1 [V-20], and only for NaCl concentrations of 1 M and greater. And, it is only for these conditions that the 12 nm particle was observed. An extremely high effective concentration of thioredoxin in the solvated, hydrophilic shell of the micelle, with little ELP buffering between thioredoxin molecules, is consistent with the observed decrease in thermal stability.

[0210] Appropriate selection of NaCl concentration and solution temperature is appropriate to efficiently carrying out ITC. Three centrifugation temperatures were selected for experimental convenience: 12°C when a microcentrifuge was placed in a 4°C refrigerated laboratory cabinet, 33°C when placed on a laboratory bench top at 22 °C, and 49°C when placed in a 37°C static incubator (all sample temperatures were measured directly by thermocouple after a 10 minute centrifugation). The NaCl concentrations were selected in 1 M increments to depress the T_i to some point below each centrifugation temperature.

[0211] For the first two examples, recovery was incomplete because at these combinations of centrifugation temperature and NaCl concentration, thioredoxin-ELP1 [V-20] showed a two phase behavior where larger aggregates coexisted with the 12 nm particles. Because of their small mass, these particles remained suspended during centrifugation, and only the fraction of fusion protein contained in the larger aggregate phase was removed by centrifugation and recovered in the resolubilized pellet. At 49°C, the thioredoxin-ELP1 [V-20] turbidity profile in PBS with 1 M NaCl was significantly decreased from its maximum value, and data showed that a majority of the scattering intensity came from the 12 nm particles. Correspondingly, the SDS-PAGE data showed that only a small fraction of the fusion protein present was captured by centrifugation during ITC purification. At 33°C in PBS with 2 M NaCl, although still below its maximum value, the turbidity of thioredoxin-ELP1 [V-20] was closer to its peak value, and the data showed that the scattering intensity attributed to the 12 nm particle was much smaller. Consistent with these observations, a majority of fusion protein was captured by ITC purification as ascertained by SDS-PAGE, although loss in the supernatant due to the 12 nm particles was still significant.

[0212] Using a centrifugation temperature of 12°C in PBS with 3 M NaCl, recovery of the fusion protein in the resolubilized pellet was nearly complete. Under these conditions, the solution

turbidity was very near its maximum value. The degree of turbidity, combined with the trends in particle size distribution established for lower salt concentrations, suggest that the complete recovery obtained by ITC with these conditions is explained by the presence of only the larger aggregates for these solution conditions.

[0213] These examples illustrate that for efficient ITC purification of thioredoxin-ELP1 [V-20], and potentially for other soluble fusion proteins with small ELP tags, the NaCl concentration and centrifugation temperature should be selected to achieve the maximum point in the turbidity profile. For microcentrifuges without temperature control, this is most practically achieved by determining the centrifuge sample temperature, and then adjusting the T_i of the fusion protein by the precise addition of salt. For larger centrifuges that are equipped with refrigeration systems, recovery efficiency can be maximized by the combined alteration of NaCl concentration and centrifugation temperature. The required precision in controlling solution conditions during ITC for thioredoxin-ELP1 [V-20] versus that for thioredoxin-ELP1 [V₅A₂G₃-90], which can be fully recovered using any combination of temperature and salt concentration that induces the inverse transition, is the price paid for the four-fold increase in yield of the target protein.

[0214] Decreasing the length of the ELP purification tag from 36 to 9 kDa produced a four-fold increase in the expression levels of *E. coli* thioredoxin, a model target protein. The expression level with the 9 kDa tag was similar to that of free thioredoxin expressed without an ELP tag, and therefore further reduction of the ELP tag size is not likely to provide any additional benefit. Although truncation of the ELP did not adversely affect the purity of the final protein product, it is important to select an appropriate combination of salt concentration and solution temperature to favor the formation of larger aggregates during ITC purification.

EXAMPLE 3:

HIGH-THROUGHPUT PURIFICATION OF RECOMBINANT PROTEINS USING ELP TAGS

[0215] The gene for the 5-polypentapeptide VPGVG ELP sequence was constructed by annealing two 5'-phosphorylated synthetic oligonucleotides (Integrated DNA Technologies, Coralville, IA) to yield double stranded DNA with PflMI and HindIII compatible ends. This gene was inserted into a PflMI / HindIII linearized and dephosphorylated modified pUC-19 (New England Biolabs, Beverly, MA) vector and polymerized using recursive directional ligation with PflMI and BglII (Meyer, 1999; Meyer, 2000) to generate the gene for the 20-polypentapeptide ELP sequence. This

ELP gene was then excised with PflMI and BglI, gel purified (QIAquick Gel Extraction Kit, Qiagen, Valencia, CA), and inserted into a SfiI linearized and dephosphorylated modified pET32b vector (Novagen, Madison, WI; Meyer, 1999). This expression vector was then transformed into the BLR(DE3) (Novagen) *E. Coli* expression strain.

[0216] The aforementioned cells were taken from frozen (DMSO) stock and streaked onto agar plates supplanted with 100 µg/ml ampicillin and allowed to grow overnight. Two hundred microliters of growth media (100 µg/ml ampicillin in CircleGrow media; Qbiogene, Inc., Carlsbad, CA) were injected into each well of a standard 96 well microplate (Costar, Corning Inc., Corning, NY) using a multichannel pipetter. Using 200 µl pipet tips, each well of the microplate was inoculated with a pinhead-sized aggregation of cells from colonies on the aforementioned agar plates. With the lid on, the microplate was incubated at 37°C and shaken at 275 r.p.m. The microplate was held in place in the shaker using an ad hoc microplate holder. The cultures were induced by adding isopropyl α-thiogalactopyranoside to a final concentration of 1 mM when the OD₆₅₀ reached 0.65 for a majority of the cultures as measured using a microplate reader (Thermomax; Molecular Devices Co., Sunnyvale, CA) - this optical density corresponds to an OD₆₅₀ of 2.0 as measured using an UV-visible spectrophotometer (UV-1601, Shimadzu Scientific Instruments, Inc.). The cultures were incubated and shaken for 4 hours post-induction and then harvested by centrifugation at 1100 g for 40 minutes at 4°C using matched-weight microplate carrier adaptors (Beckman Instruments, Inc., Palo Alto, CA). The media was discarded and the cell pellets were frozen in the microplates at -80°C until they were ready to be purified.

[0217] The ELP1 [V-20]/thioredoxin protein was purified from cell cultures in the microplates as follows. The cells were lysed by adding 1 µl of lysozyme solution (25 mg/ml; Grade VI; Sigma, St. Louis, MO) and 25 µl of lysis buffer (50 mM NaCl, 5% glycerol, 50 mM Tris-HCl, pH 7.5) to each well. The micro plate was then shaken using an orbital shaker at 4°C for 20 minutes. Two µl of 1.35% (by mass) sodium doxycholate solution were added to each well and the microplate was shaken at 4°C for 5 minutes. Two µl of deoxyribonuclease I solution (100 units/ul; Type II; Sigma, St. Louis, MO) were added to each well and the microplate was shaken at 4°C for 10 minutes. The microplate was then centrifuged at 1100 g for 20 minutes at 4°C using matched-weight microplate carrier adaptors (Beckman Instruments, Inc., Palo Alto, CA) to pellet cell particulates and insoluble proteins. Two µl of 10% (by mass) polyethylenimine solution was added to each well and the microplate was shaken at 4°C for 15 minutes. The microplate was then centrifuged at 1100 g for 20 minutes at 4°C to pellet DNA. The supernatants were transferred to wells on a new microplate

and the old microplate was discarded. To induce ELP1 [V-20]/thioredoxin aggregation, 20 μ l of saturated NaCl solution was added to each well; a marked increase in turbidity indicated aggregation of the target protein. To pellet the aggregated proteins, the microplate was centrifuged at 1100 g for 40 minutes at 30°C. The protein pellets were resolubilized in 30 μ l of phosphate buffer solution after which the microplate was centrifuged at 1100 g for 20 minutes at 4°C to remove insoluble lipids. Finally, the purified protein supernatants were transferred to wells of a new microplate and stored at 4°C. SDS-PAGE gel analysis for the ELP1 [V-20]/thioredoxin fusion protein purified by ITC was carried out.

[0218] Alternatively, ELPs/ELP-fusion proteins can be purified using a commercially available extraction reagent in accordance with the following protocol. Lyse cells by adding 25 microliters of Novagen BugBuster Protein Extraction Reagent to each microplate well. The microplate is placed on a Fisher Vortex Genie at shaker speed 2 (alternatively on an orbital shaker at maximum speed) for fifteen minutes at room temperature. Using the microplate adaptors, centrifugation is conducted (2300 rpm, 1700 x g for Beckman adaptor for the JS4.2 rotor) for 20 minutes at 4 degrees Celsius to form a pellet. Add 2 microliters polyethylenimine (to 0.66%) to the wells and shake using Vortex Genie or shaker for 5 minutes. Incubate on ice 10 minutes, shaking occasionally. Using the microplate adaptors, centrifuge at maximum speed for 25 minutes at 4 degrees Celsius. Transfer the supernatant to the new microplate and discard the old microplate with the pellet. Add NaCl (crystals) and/or increase the solution temperature to induce ELP aggregation. Mix by shaking only - pipeting will aggregate the ELP on the pipet tip. Solution should turn turbid to some extent. Centrifuge at a temperature above the transition temperature (2300 rpm, 1700 g, 35-40 degrees Celsius, 45 minutes). Discard supernatant and resuspend the pellet (typically non-visible or a tiny pellet) in 30 microliters of cold buffer of choice (PBS) by repeatedly pipeting around the bottom and walls of the well. Centrifuge (2300 rpm, 1700 x g, 4 degrees Celsius, 20 minutes) to spin out insoluble impurities such as lipids. Transfer the supernatant to another microplate. The purified ELP may be stored frozen at -80 degrees Celsius in the microplate until ready for use. (For fusions, it must be ensured that freezing is suitable for the fusion protein.) The appropriate NaCl concentration and temperature employed in this technique depends on the ELP, fusion partner, and ELP concentration. The objective is to lower the effective ELP transition temperature at least 3 to 5 degrees below the solution temperature. An effective transition temperature of 25-30 degrees Celsius and warm centrifugation at 35-40 degrees Celsius has been usefully employed, although higher temperatures may be used if tolerated by the fusion protein.

[0219] Protein concentration was determined by measuring A_{280} (UV-1601, Shimadzu Scientific Instruments, Inc.) and using the molar extinction coefficient for ELP1 [V-20]/Thioredoxin ($\epsilon=19,870$); this assumes that the ELP1 [V-20]/Thioredoxin protein samples are pure of protein and DNA impurities. Thioredoxin activity was determined using an insulin reduction assay (Holmgren, 1984).

[0220] For the construction of the fusion protein, a small ELP tag was designed with a T_i of around 70°C, using previously published theoretical T_i data (Urry, 1991). Characterization of the ELP tag showed that the T_i was 76.2°C, confirming that it is possible to rationally design ELP tags with specified T_i . For the ELP/thioredoxin fusion protein, the T_i in low salt buffer, 1 M, and 2 M salt solutions were 68°C, 37°C and 18°C, respectively, confirming that fusion of a soluble protein to an ELP tag minimally affects its T_i and showing that the T_i can be manipulated over a wide range by adjusting the salt concentration.

[0221] Based on the foregoing, the creation of a family of plasmid expression vectors that contain an ELP sequence and a polylinker region (into which the target protein is inserted) joined by a cleavage site can be employed to facilitate the expression of a variety of proteins. The ELP sequences embedded in such family of plasmids can have different transition temperatures (by varying the identity of the guest residue). The expression vector for a particular target protein is desirably selected based on the protein's surface hydrophobicity characteristics. The salt concentration of the solution then is adjusted during purification to obtain the desired T_i .

[0222] For protein expression involving growth of cell cultures in microplate wells, the cell cultures can be desirably induced at $OD_{600} \approx 2$ and grown for 4 hours post-induction. The cell density at induction for the microplate growths is two to three times that achieved by conventional protein expression protocols. Even at these high cell densities, rapid and healthy cell growth can be maintained in the microplate wells by aeration of the cultures, which as grown in the wells are characterized by a high surface area to volume ratio. Cell cultures that are grown longer post-induction yielded minimally more target protein, and growth using a hyper expression protocol (Guda, 1995) had much more contaminant protein (around tenfold) with minimally more fusion protein. In order to avoid evaporation of the cell media in the high surface area to volume ratio cell growth in the microplate wells, it was necessary to cover the microplate with an appropriate lid during growth and to infuse the cell growth with additional media during induction. On a per liter basis, cultures grown in the microplate wells had a higher level of fusion protein expression than cultures grown with conventional protocols.

[0223] High throughput protein purification utilizing ITC was successful when cells were lysed with commercial nonionic protein extraction formulations. After cell lysis, addition of polyethylenimine removed nucleic acids and high molecular mass proteins from the soluble fraction of the crude lysate upon centrifugation. At the fusion protein and salt concentrations of the soluble lysate, the T_i of the fusion protein was approximately 65°C. Heating the soluble lysate above this temperature to induce fusion protein aggregation denatures and precipitates soluble contaminant proteins as well as the target protein itself. Furthermore, this temperature could not be maintained within the centrifuge chamber during centrifugation. Therefore, salt was added to the soluble lysate to approximately 2 M; this depressed the T_i of the fusion protein to approximately 18°C, allowing for aggregation of the fusion protein at room temperature. This salt concentration did not precipitate any contaminant proteins nor did it alter the functionality of the final purified protein product.

[0224] High throughput protein purification using ITC was both effective and efficient. About 15% of the expressed fusion protein was lost in the insoluble protein fraction of the cell lysate. Centrifugation of the sample after fusion protein aggregation effectively separated the proteins: 90% of the fusion protein was pelleted while 10% of the fusion protein remained in the supernatant along with all soluble contaminant proteins. Overall, about 75% of the expressed protein was abstracted using ITC purification and *E. coli* contaminant protein levels in the purified products were below those detectable by SDS-PAGE. The purification process can be expedited and purification efficiency increased by increasing the centrifugation speeds; higher centrifugation speeds allow for reduced centrifugation times and at higher centrifugation speeds (5000 g), all of the fusion protein is pelleted during centrifugation post aggregation. Addition of thrombin completely cleaved the fusion protein and a second round of ITC separated the ELP tag from the thioredoxin target protein with no loss of thioredoxin.

[0225] The average amount of fusion protein purified per well determined using absorbance measurements (A_{280} , $\epsilon=19,870$) was 33 ug with a standard deviation of 8.5 ug. Values were dispersed evenly between 19.7 and 48.3 ug per well. The large variation in yield of purified protein was due more to the different amounts of protein expressed in the different wells than to a variation in the purification efficiency of the ITC process. Varying amounts of protein were expressed in the different cell cultures because 1) the imprecision of the inoculation meant that cell cultures had varying amounts of cells to begin with and 2) due in all likelihood to more abundant aeration, the cell cultures in peripheral wells tended to have faster growth and reach a higher stationary phase

cell density. For simplicity of effort, all of the cell cultures were induced and then harvested at the same times as opposed to induction and harvesting of individual cell cultures.

[0226] The enzymatic activity of the thioredoxin target protein was measured using an insulin reduction assay. The average amount of fusion protein per well, determined on the basis of such enzymatic activity, was 35.7 ug with a standard deviation of 8.0 ug. Again, values were dispersed evenly, between a minimum of 24.6 and a maximum of 50.8 ug per well. It is important to note that thioredoxin was enzymatically active though still attached to the ELP tag. The thioredoxin expressed and purified using this high throughput ITC technique had, on average, 10.3% greater enzymatic activity per unit mass than that of commercial thioredoxin (Sigma), a testament to the gentleness of and purity achieved by the ITC process.

[0227] On average, high throughput ELP/thioredoxin protein expression and purification produced around 160 mg of protein per liter of growth. This is comparable to ELP/thioredoxin yields obtained using conventional protein expression and ITC purification methods (140-200 mg protein/L of growth).

[0228] An SDS-PAGE gel of the stages of high throughput protein purification using microplates and inverse transition cycling was carried out according to the above-described procedure, in which ELP/thioredoxin fusion protein was purified using the documented protocol. Gel samples were denatured with SDS, reduced with beta-mercaptoethanol, and run at 200 V for 45 minutes on a 10-20% gradient Tris-HCl gel.

[0229] Histograms were employed for quantization of purified protein samples, including a histogram of total fusion protein per well determined using absorbance measurements (A_{280} , $\epsilon=19,870$) ($n=20$, $\mu=32.97$, $\sigma=8.48$) and a histogram of fusion protein functionality/purity for each sample compared to commercial thioredoxin (from Sigma) ($n=20$, $\mu=110.37\%$, $\sigma=16.54\%$).

[0230] Considering such high throughput protein expression and purification method, it is noted that whereas nickel-chelated multiwell plates can purify only 1 ng of His-tagged protein per well, the capacity of high throughput purification using ITC is limited only by the amount of the protein that can be expressed by cultures grown in the well; for ELP tagged proteins, the level of protein expression is in the tens of microgram range.

[0231] High throughput purification using ITC thus provides high yields, producing sufficient fusion protein for purification of the peptide active therapeutic agent-ELP construct to produce active ingredient for therapeutic compositions. Milligram levels of purified fusion protein can be

obtained by growing cell cultures in other vessels and transferring the resuspended cell pellet to the multiwell plate for the purification process. Finally, such high throughput purification technique is technically simpler and less expensive than current conventional commercial high throughput purification methods as it requires only one transfer of purification intermediates to a new multiwell plate.

EXAMPLE 4:

CONSTRUCTION OF VARIOUS ELP GENE EXPRESSION SERIES

Bacterial Strains and Plasmids

[0232] Cloning steps were conducted in *Escherichia coli* strain XL1-Blue (*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17* (r_K^- , m_K^+), *supE44*, *relA1*, *lac[F'*, *proAB*, *lacI^q*ZAM15, Tn10 (Tet^r)] (Stratagene La Jolla, CA). pUC19 (NEB, Beverly, MA) was used as the cloning vector the ELP construction (Meyer and Chilkoti, 1999). Modified forms of pET15b and pET24d vectors (Novagen) were used to express ELP and ELP-fusion proteins in BL21 Star (DE3) strain (*F'*, *ompT*, *hsdS_B* (r_B^- m_B^-), *gal*, *dcm*, *rne131*, (DE3)) (Invitrogen Carlsbed, CA) or BLR(DE3) (*F'*, *ompT*, *hsdS_B* (r_B^- m_B^-), *gal*, *dcm*, Δ (*srl-recA*) 306::Tn10(Tc^R)(DE3)) (Novagen Madison, WI). Synthetic DNA oligos were purchased from Integrated DNA Technologies, Coralville, IA. All vector constructs were made using standard molecular biology protocols (Ausubel, et al., 1995).

Construction of ELP1 [V₅A₂G₃] Gene Series

[0233] The ELP1 [V₅A₂G₃] series designate polypeptides containing multiple repeating units of the pentapeptide VPGXG, where X is valine, alanine, and glycine at a relative ratio of 5:2:3.

[0234] The ELP1 [V₅A₂G₃] series monomer, ELP1 [V₅A₂G₃-10], was created by annealing four 5' phosphorylated, PAGE purified synthetic oligos to form double stranded DNA with *EcoR*I and *Hind*III compatible ends (Meyer and Chilkoti, 1999). The oligos were annealed in a 1 μ M mixture of the four oligos in 50 μ l 1X ligase buffer (Invitrogen) to 95°C in a heating block than the block was allowed to cool slowly to room temperature. The ELP1 [V₅A₂G₃-10]/*EcoR*I-*Hind*III DNA segment was ligated into a pUC19 vector digested with *EcoR*I and *Hind*III and CIAP dephosphorylated (Invitrogen) to form pUC19-ELP1[V₅A₂G₃-10]. Building of the ELP1 [V₅A₂G₃] series library began by inserting ELP1 [V₅A₂G₃-10] *Pf*lM1/*Bgl*I fragment from pUC19-ELP1[V₅A₂G₃-10] into pUC19-ELP1[V₅A₂G₃-10] linearized with *Pf*lM1 and dephosphorylated with CIAP to create pUC19-ELP1[V₅A₂G₃-20]. pUC19-ELP1[V₅A₂G₃-20] was then built up to

pUC19-ELP1[V₅A₂G₃-30] and pUC19-ELP1[V₅A₂G₃-40] by ligating ELP1[V₅A₂G₃-10] or ELP1[V₅A₂G₃-20] *Pfl*MI/*Bgl*I fragments respectively into *Pfl*MI digested pUC19-ELP1[V₅A₂G₃-20]. This procedure was used to expand the ELP1 [V₅A₂G₃] series to create pUC19-ELP1[V₅A₂G₃-60], pUC19-ELP1[V₅A₂G₃-90] and pUC19-ELP1[V₅A₂G₃-180] genes.

Construction of ELP1 [K₁V₂F₁] Gene Series

[0235] The ELP1 [K₁V₂F₁] series designate polypeptides containing multiple repeating units of the pentapeptide VPGXG, where X is lysine, valine, and phenylalanine at a relative ratio of 1:2:1.

[0236] The ELP1 [K₁V₂F₁] series monomer, ELP1 [K₁V₂F₁-4] (SEQ ID NO: 18), was created by annealing two 5' phosphorylated, PAGE purified synthetic oligos to form double stranded DNA with *Eco*R1 and *Hind*III compatible ends (Meyer and Chilkoti, 1999). The oligos were annealed in a 1 μM mixture of the four oligos in 50 μl 1X ligase buffer (Invitrogen) to 95°C in a heating block than the block was allowed to cool slowly to room temperature. The ELP1 [K₁V₂F₁-4]/*Eco*R1-*Hind*III DNA segment was ligated into a pUC19 vector digested with *Eco*R1 and *Hind*III and CIAP dephosphorylated (Invitrogen) to form pUC19-ELP1[K₁V₂F₁-4]. Building of the ELP1 [K₁V₂F₁] series library began by inserting ELP1 [K₁V₂F₁-4] *Pfl*MI/*Bgl*I fragment from pUC19-ELP1[K₁V₂F₁-4] into pUC19-ELP1[K₁V₂F₁-4] linearized with *Pfl*MI and dephosphorylated with CIAP to create pUC19-ELP1[K₁V₂F₁-8]. Using the same procedure the ELP1 [K₁V₂F₁] series was doubled at each ligation to form pUC19-ELP1[K₁V₂F₁-16], pUC19-ELP1[K₁V₂F₁-32], pUC19-ELP1[K₁V₂F₁-64] and pUC19-ELP1[K₁V₂F₁-128].

Construction of ELP1 [K₁V₇F₁] Gene Series

[0237] The ELP1 [K₁V₇F₁] series designate polypeptides containing multiple repeating units of the pentapeptide VPGXG, where X is lysine, valine, and phenylalanine at a relative ratio of 1:7:1.

[0238] The ELP1 [K₁V₇F₁] series monomer, ELP1 [K₁V₇F₁-9] (SEQ ID NO: 19), was created by annealing four 5' phosphorylated, PAGE purified synthetic oligos to form double stranded DNA with *Pfl*MI and *Hind*III compatible ends. The ELP1 [K₁V₇F₁-9] DNA segment was then ligated into *Pfl*MI/*Hind*III dephosphorylated pUC19-ELP1[V₅A₂G₃-180] vector thereby substituting ELP1 [V₅A₂G₃-180] for ELP1 [K₁V₇F₁-9] to create the pUC19-ELP1[K₁V₇F₁-9] monomer. The ELP1 [K₁V₇F₁] series was expanded in the same manner as the ELP1 [K₁V₂F₁] series to create pUC19-ELP1[K₁V₇F₁-18], pUC19-ELP1[K₁V₇F₁-36], pUC19-ELP1[K₁V₇F₁-72] and pUC19-ELP1[K₁V₇F₁-144].

Construction of ELP1 [V] Gene Series

[0239] The ELP1 [V] series designate polypeptides containing multiple repeating units of the pentapeptide VPGXG, where X is exclusively valine.

[0240] The ELP1 [V] series monomer, ELP1 [V-5] (SEQ ID NO: 14), was created by annealing two 5' phosphorylated, PAGE purified synthetic oligos to form double stranded DNA with *EcoRI* and *HindIII* compatible ends. The ELP1 [V-5] DNA segment was then ligated into *EcoRI/HindIII* dephosphorylated pUC19 vector to create the pUC19-ELP1[V-5] monomer. The ELP1 [V] series was created in the same manner as the ELP1 [V₅A₂G₃] series, ultimately expanding pUC19-ELP1[V-5] to pUC19-ELP1[V-60] and pUC19-ELP1[V-120].

Construction of ELP2 Gene Series

[0241] The ELP2 series designate polypeptides containing multiple repeating units of the pentapeptide AVGVV.

[0242] The ELP2 series monomer, ELP2 [5] (SEQ ID NO: 20), was created by annealing two 5' phosphorylated, PAGE purified synthetic oligos to form double stranded DNA with *EcoRI* and *HindIII* compatible ends. The ELP2 [5] DNA segment was then ligated into *EcoRI/HindIII* dephosphorylated pUC19 vector to create the pUC19-ELP2[5] monomer. The ELP2 series was expanded in the same manner as the ELP1 [K₁V₂F₁] series to create pUC19-ELP2[10], pUC19-ELP2[30], pUC19-ELP2[60] and pUC19-ELP2[120].

Construction of ELP3 [V] Gene Series

[0243] The ELP3 [V] series designate polypeptides containing multiple repeating units of the pentapeptide IPGXG, where X is exclusively valine.

[0244] The ELP3 [V] series monomer, ELP3 [V-5] (SEQ ID NO: 21), was created by annealing two 5' phosphorylated, PAGE purified synthetic oligos to form double stranded DNA with *PfLM1* amino terminal and *GGC* carboxyl terminal compatible ends due to the lack of a convenient carboxyl terminal restriction site but still enable seamless addition of the monomer. The ELP3 [V-5] DNA segment was then ligated into *PfLM1/BglII* dephosphorylated pUC19-ELP4[V-5], thereby substituting ELP4 [V-5] for ELP3 [V-5] to create the pUC19-ELP3[V-5] monomer. The ELP3 [V] series was expanded by ligating the annealed ELP3 oligos into pUC19-ELP3[V-5] digested with *PfLM1*. Each ligation expands the ELP3 [V] series by 5 to create ELP3 [V-10], ELP3 [V-15], etc.

Construction of the ELP4 [V] Gene Series

[0245] The ELP4 [V] series designate polypeptides containing multiple repeating units of the pentapeptide LPGXG, where X is exclusively valine.

[0246] The ELP4 [V] series monomer, ELP4 [V-5] (SEQ ID NO: 22), was created by annealing two 5' phosphorylated, PAGE purified synthetic oligos to form double stranded DNA with *EcoRI* and *HindIII* compatible ends. The ELP4 [V-5] DNA segment was then ligated into *EcoRI/HindIII* dephosphorylated pUC19 vector to create the pUC19-ELP4[V-5] monomer. The ELP4 [V] series was expanded in the same manner as the ELP1 [K₁V₂F₁] series to create pUC19-ELP4[V-10], pUC19-ELP4[V-30], pUC19-ELP4[V-60] and pUC19-ELP4[V-120].

[0247] The ELP genes were also inserted into other vectors such as pET15b-SD0, pET15b-SD3, pET15b-SD5, pET15b-SD6, and pET24d-SD21. The pET vector series are available from Novagen, San Diego, CA.

[0248] The pET15b-SD0 vector was formed by modifying the pET15b vector using SD0 double-stranded DNA segment containing the multicloning restriction site (*SacI-NdeI-NcoI-XhoI-SnaBI-BamHI*). The SD0 double-stranded DNA segment had *XbaI* and *BamHI* compatible ends and was ligated into *XbaI/BamHI* linearized and 5'-dephosphorylated pET15b to form the pET15b-SD0 vector.

[0249] The pET15b-SD3 vector was formed by modifying the pET15b-SD0 vector using SD3 double-stranded DNA segment containing a *SfiI* restriction site upstream of a hinge region-thrombin cleavage site followed by the multicloning site (*NdeI-NcoI-XhoI-SnaBI-BamHI*). The SD3 double-stranded DNA segment had *SacI* and *NdeI* compatible ends and was ligated into *SacI/NdeI* linearized and 5'-dephosphorylated pET15b-SD0 to form the pET15b-SD3 vector.

[0250] The pET15b-SD5 vector was formed by modifying the pET15b-SD3 vector using the SD5 double-stranded DNA segment containing a *SfiI* restriction site upstream of a thrombin cleavage site followed by a hinge and the multicloning site (*NdeI-NcoI-XhoI-SnaBI-BamHI*). The SD5 double-stranded DNA segment had *SfiI* and *NdeI* compatible ends and was ligated into *SfiI/NdeI* linearized and 5'-dephosphorylated pET15b-SD3 to form the pET15b-SD5 vector.

[0251] The pET15b-SD6 vector was formed by modifying the pET15b-SD3 vector using the SD6 double-stranded DNA segment containing a *SfiI* restriction site upstream of a linker region-TEV cleavage site followed by the multicloning site (*NdeI-NcoI-XhoI-SnaBI-BamHI*). The SD6

double-stranded DNA segment had *Sfi*I and *Nde*I compatible ends and was ligated into *Sfi*I/*Nde*I linearized and 5'-dephosphorylated pET15b-SD3 to form the pET15b-SD6 vector.

[0252] The pET24d-SD21 vector was formed by modifying the pET24d vector using the SD21 double-stranded DNA segment with *Nco*I and *Nhe*I compatible ends. The SD21 double-stranded DNA segment was ligated into *Nco*I/*Nhe*I linearized and 5' dephosphorylated pET24d to create the pET24d-SD21 vector, which contained a new multi-cloning site *Nco*I-*Sfi*I-*Nhe*I-BamHI-EcoRI-SacI-SalI-HindIII-NotI-XhoI with two stop codons directly after the *Sfi*I site for insertion and expression of ELP with the minimum number of extra amino acids.

[0253] The pUC19-ELP1[V₅A₂G₃-60], pUC19-ELP1[V₅A₂G₃-90], and pUC19-ELP1[V₅A₂G₃-180] plasmids produced in XL1-Blue were digested with *Pf*IM1 and *Bgl*I, and the ELP-containing fragments were ligated into the *Sfi*I site of the pET15b-SD3 expression vector as described hereinabove to create pET15b-SD3-ELP1[V₅A₂G₃-60], pET15b-SD5-ELP1[V₅A₂G₃-90] and pET15b-SD5-ELP1[V₅A₂G₃-180], respectively.

[0254] The pUC19-ELP1[V₅A₂G₃-90], pUC19-ELP1[V₅A₂G₃-180], pUC19-ELP1[V-60] and pUC19-ELP1[V-120] plasmids produced in XL1-Blue were digested with *Pf*IM1 and *Bgl*I, and the ELP-containing fragments were ligated into the *Sfi*I site of the pET15b-SD5 expression vector as described hereinabove to create pET15b-SD5-ELP1[V₅A₂G₃-90], pET15b-SD5-ELP1[V₅A₂G₃-180], pET15b-SD5-ELP1[V-60] and pET15b-SD5-ELP1[V-120], respectively.

[0255] The pUC19-ELP1[V₅A₂G₃-90] plasmid produced in XL1-Blue was digested with *Pf*IM1 and *Bgl*I, and the ELP-containing fragment was ligated into the *Sfi*I site of the pET15b-SD6 expression vector as described hereinabove to create pET15b-SD6-ELP1[V₅A₂G₃-90].

[0256] The pUC19-ELP1[K₁V₂F₁-64], and pUC19-ELP1[K₁V₂F₁-128] plasmids produced in XL1-Blue were digested with *Pf*IM1 and *Bgl*I, and the ELP-containing fragments were ligated into the *Sfi*I site of the pET24d-SD21 expression vector as described hereinabove to create pET24d-SD21-ELP1[K₁V₂F₁-64] and pET24d-SD21-ELP1[K₁V₂F₁-128], respectively.

[0257] The pUC19-ELP1[K₁V₇F₁-72] and pUC19-ELP1[K₁V₇F₁-144] plasmids produced in XL1-Blue were digested with *Pf*IM1 and *Bgl*I, and the ELP-containing fragments were ligated into the *Sfi*I site of the pET24d-SD21 expression vector as described hereinabove to create pET24d-SD21-ELP1[K₁V₇F₁-72] pET24d-SD21-ELP1[K₁V₇F₁-144], respectively.

[0258] The pUC19-ELP2[60] and pUC19-ELP2[120] plasmids produced in XL1-Blue were digested with *Nco*I and *Hind*III, and the ELP-containing fragments were ligated into the *Nco*I and *Hind*III sites of the pET24d-SD21 expression vector as described hereinabove to create pET24d-SD21-ELP2[60], pET24d-SD21-ELP2[120], respectively.

[0259] The pUC19-ELP4[V-60] and pUC19-ELP4[V-120] plasmids produced in XL1-Blue were digested with *Nco*I and *Hind*III, and the ELP-containing fragments were ligated into the *Nco*I and *Hind*III sites of the pET24d-SD21 expression vector as described hereinabove to create pET24d-SD21-ELP4[V-60], pET24d-SD21-ELP4[V-120], respectively.

EXAMPLE 5:

CONSTRUCTION, ISOLATION AND PURIFICATION OF VARIOUS FUSION PROTEINS

[0260] It is to be noted that the following fusion proteins illustrate a variety of peptide active therapeutic agent and ELP species in specific combinations.

[0261] Although these fusion proteins were designed with cleavage sites between the respective peptide active therapeutic agent and ELP moieties, for use in cleaving reactions to produce peptide active therapeutic agent and ELP moieties for further study, corresponding peptide active therapeutic agent-ELP constructs lacking such cleavage sites are readily produced, by the simple expedient of direct bonding of the peptide active therapeutic agent to the ELP, without any interposed cleavage group or moiety that is susceptible to scission by proteases or other degradative agents or conditions that may be encountered by the construct *in vivo* subsequent to its administration.

[0262] Experiments were conducted to show the use of various target proteins (peptide active therapeutic agents) in forming ELP-containing fusion proteins and the inverse phase transition behavior exhibited by such fusion proteins. Specifically, the following thirty-six (36) ELP-containing fusion proteins were formed in *E. coli* by using known recombinant expression techniques consistent with the teachings and disclosures hereinabove:

- Insulin A peptide and ELP1 [V-60] polypeptide with an enterokinase protease cleavage site therebetween (SEQ ID NO: 23);
- Insulin A peptide and ELP1 [V₅A₂G₃-90] polypeptide with an enterokinase protease cleavage site therebetween (SEQ ID NO: 24);

- Insulin A peptide and ELP1 [V-120] polypeptide with an enterokinase protease cleavage site therebetween (SEQ ID NO: 25);
- Insulin A peptide and ELP1 [V₅A₂G₃-180] polypeptide with an enterokinase protease cleavage site therebetween (SEQ ID NO: 26);
- T20 peptide and ELP1 [V-60] polypeptide with an enterokinase protease cleavage site therebetween (SEQ ID NO: 27);
- T20 peptide and ELP1 [V₅A₂G₃-90] polypeptide with an enterokinase protease cleavage site therebetween (SEQ ID NO: 28);
- T20 peptide and ELP1 [V-120] polypeptide with an enterokinase protease cleavage site therebetween (SEQ ID NO: 29);
- T20 peptide and ELP1 [V-60] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 30);
- T20 peptide and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 31);
- T20 peptide and ELP1 [V-120] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 32);
- T20 peptide and ELP1 [V-60] polypeptide with a tobacco etch virus (TEV) protease cleavage site (cleavage between QS residues) therebetween (SEQ ID NO: 33);
- T20 peptide and ELP1 [V₅A₂G₃-90] polypeptide with a TEV protease cleavage site (cleavage between QS residues) therebetween (SEQ ID NO: 34);
- T20 peptide and ELP1 [V-120] polypeptide with a TEV protease cleavage site (cleavage between QS residues) therebetween (SEQ ID NO: 35);
- T20 peptide and ELP1 [V-60] polypeptide with a TEV protease cleavage site (cleavage between QY residues) therebetween (SEQ ID NO: 36);
- T20 peptide and ELP1 [V₅A₂G₃-90] polypeptide with a TEV protease cleavage site (cleavage between QY residues) therebetween (SEQ ID NO: 37);
- T20 peptide and ELP1 [V-120] polypeptide with a TEV protease cleavage site (cleavage between QY residues) therebetween (SEQ ID NO: 38);
- Interferon alpha 2B protein and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 39);
- Tobacco etch virus protease and ELP1 [V-60] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 40);
- Tobacco etch virus protease and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 41);

- Tobacco etch virus protease and ELP1 [V-120] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 42);
- Tobacco etch virus protease and ELP1 [V₅A₂G₃-180] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 43);
- Small heterodimer partner orphan receptor and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 44);
- Androgen receptor ligand binding domain and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 45);
- Androgen receptor ligand binding domain and ELP1 [V₅A₂G₃-180] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 46);
- Glucocorticoid receptor ligand binding domain and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 47);
- Estrogen receptor ligand binding domain and ELP1 [V₅A₂G₃-60] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 48);
- Estrogen receptor ligand binding domain and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 49);
- Estrogen receptor ligand binding domain and ELP1 [V₅A₂G₃-180] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 50);
- Estrogen receptor ligand binding domain and ELP1 [V₅A₂G₃-90] polypeptide with a TEV protease cleavage site (cleavage between QG residues) therebetween (SEQ ID NO: 51);
- G protein alpha Q and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 52);
- G protein alpha Q and ELP1 [V₅A₂G₃-180] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 53);
- 1-Deoxy-D-Xylulose 5-Phosphate reductoisomerase peptide and ELP1 [V₅A₂G₃-60] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 54);
- 1-Deoxy-D-Xylulose 5-Phosphate reductoisomerase peptide and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 55);
- 1-Deoxy-D-Xylulose 5-Phosphate reductoisomerase peptide and ELP1 [V₅A₂G₃-180] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 56);
- 1-Deoxy-D-Xylulose 5-Phosphate reductoisomerase peptide and ELP1 [V₅A₂G₃-90] polypeptide with a TEV protease cleavage site (cleavage between QG residues) therebetween (SEQ ID NO: 57); and

- G protein alpha S and ELP1 [V₅A₂G₃-90] polypeptide with a thrombin protease cleavage site therebetween (SEQ ID NO: 58).

[0263] All of the above-listed thirty-six ELP-containing fusion proteins were found to retain the inverse phase transition behavior of the corresponding ELP tags, and were successfully isolated and purified by using inverse transition cycling (ITC) techniques, according to the following experimental procedure:

Isolation and Purification of Fusion Proteins Containing Insulin A Peptide (InsA)

[0264] A single colony of *E. coli* strain BLR (DE3) (Novagen) containing the respective ELP-InsA fusion protein was inoculated into 5 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 µg/ml ampicillin (Sigma) and grown at 37°C with shaking at 250 rpm for 5 hours. The 5 ml culture was then inoculated into a 500 ml culture and allowed to grow at 25°C for 16 hours before inducing with 1 mM IPTG for 4 hours at 25°C. The culture was harvested and suspended in 40 ml 20 mM Tris-HCL pH 7.4, 50 mM NaCl, 1 mM DTT and 1 Complete EDTA free Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice for 3 minutes, which consisted of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0265] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 1.0 M therein, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the respective ELP-InsA fusion protein and non-specifically NaCl precipitated proteins.

[0266] The pellet was re-suspended in 40 ml ice-cold ml 20 mM Tris-HCL pH 7.4, 50 mM NaCl, 1 mM DTT and re-centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional times to increase the purity of the respective ELP-InsA fusion protein and reduce the final volume to 0.5 ml.

Isolation and Purification of Fusion Proteins Containing T20 Peptide (T20)

[0267] A single colony of *E. coli* strain BLR (DE3) (Novagen) containing the respective ELP-T20 fusion protein was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 µg/ml ampicillin (Sigma) and grown at 37°C with shaking at 250 rpm for 24 hours. The culture was harvested and suspended in 40 ml 50 mM Tris pH 8.0, 0.5 mM EDTA and 1 Complete Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice

for 3 minutes, which consisted of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0268] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 1.0 M therein, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the respective ELP-T20 fusion protein and non-specifically NaCl precipitated proteins.

[0269] The pellet was re-suspended in 40 ml ice-cold 50 mM Tris pH 8.0, 0.5 mM EDTA and re-centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional times to increase the purity of the respective ELP-T20 fusion protein and reduce the final volume to 5 ml.

Isolation and Purification of Fusion Protein Containing Interferon Alpha 2B Peptide (IFNA2)

[0270] A single colony of *E. coli* strain BL21(DE3) TrxB⁻ (Novagen) containing the ELP-IFNA2 fusion protein and Codon Plus-RIL plasmid (Stratagene) was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 µg/ml ampicillin (Sigma), 25 µg/ml Chloramphenicol (Sigma) and incubated at 27°C with shaking at 250 rpm for 48 hours. The culture was harvested and suspended in 50 mM Tris-HCL pH 7.4, 50 mM NaCl and 1 Complete EDTA free Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice for 3 minutes, which consists of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0271] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 1.5 M, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the ELP-IFNA2 fusion protein and non-specifically NaCl precipitated proteins.

[0272] The pellet was re-suspended in 40 ml ice-cold 50 mM Tris-HCL pH 7.4 and 50 mM NaCl and re-centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional times to increase the purity of the ELP-IFNA2 fusion protein and reduce the final volume to 5 ml.

Isolation and Purification of Fusion Proteins Containing Tobacco Etch Virus Protease (TEV)

[0273] A single colony of *E. coli* strain BL21 star or BRL(DE3) containing pET15b-SD5-ELP-TEV constructs and Codon Plus-RIL plasmid (Stratagene) was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 µg/ml ampicillin (Sigma), 25 µg/ml

Chloramphenicol (Sigma) and incubated at 27°C with shaking at 250 rpm for 48 hours. The culture was harvested and suspended in 50 mM Tris-HCL pH 8.0, 1 mM EDTA, 5 mM DTT, 10% glycerol and 1mM PMSF. Cells were lysed by ultrasonic disruption on ice for 3 minutes, consisting of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0274] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 1.5 M, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the respective ELP-TEV fusion protein and non-specifically NaCl precipitated proteins.

[0275] The pellet was re-suspended in 40 ml ice-cold 50 mM Tris-HCL pH 8.0, 1 mM EDTA, 5 mM DTT, 10% glycerol and re-centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional times to increase the purity of the respective ELP-TEV fusion protein and reduce the final volume to 1 ml.

Isolation and Purification of Fusion Protein Containing Small Heterodimer Partner Orphan Receptor (SHP)

[0276] A single colony of *E. coli* strain BL21 Star (DE3) containing the ELP-SHP fusion protein was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 µg/ml ampicillin (Sigma) and 10% sucrose and grown at 27°C with shaking at 250 rpm for 48 hours. The culture was harvested and suspended in 50 mM Tris-HCL pH 8.0, 150 mM KCL, 1 mM DTT 1 mM EDTA and 1 Complete EDTA free Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice for 3 minutes, which consists of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. DNA and RNA in the soluble lysate were further degraded by adding 2 µl Benzonase (Novagen) and incubating at 4°C for 30 minutes. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0277] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 1.5 M, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the ELP-SHP fusion protein and non-specifically NaCl precipitated proteins.

[0278] The pellet was re-suspended in 40 ml ice-cold 50 mM Tris-HCL pH 8.0, 150 mM KCL, 1 mM DTT 1 mM EDTA, and 1% N-Octylglucoside and re-centrifuged at 20,000 g, 4°C for 15

minutes to remove non-specific insoluble proteins. The temperature transition cycle was repeated two additional times to increase the purity of the ELP-SHP fusion protein and reduce the final volume to 2 ml.

Isolation and Purification of Fusion Proteins Containing Androgen Receptor Ligand Binding Domain (AR-LBD)

[0279] A single colony of *E. coli* strain BL21 Star (DE3) containing the respective ELP-AR-LBD fusion protein was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 µg/ml ampicillin (Sigma) and 10 µM DHT and grown at 27°C with shaking at 250 rpm for 48 hours. The culture was harvested and suspended in 40 ml 50mM Hepes pH 7.5, 150 mM NaCl, 0.1% N-Octylglycoside, 10% glycerol, 1 mM DTT, 1 µM DHT and 1 Complete EDTA free Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice for 3 minutes, which consisted of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. DNA and RNA in the soluble sonicate were further degraded by adding 2 µl Benzonase (Novagen) and incubating at 4°C for 30 minutes. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0280] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 2.0 M, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the respective ELP-AR-LBD fusion protein and non-specifically NaCl precipitated proteins.

[0281] The pellet was re-suspended in 40 ml ice-cold 50mM Hepes pH 7.5, 150 mM NaCl, 0.1% N-Octylglycoside, 10% glycerol, 1 mM DTT and 1 µM DHT and re-centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional times to increase the purity of the respective ELP-AR-LBD fusion protein and reduce the final volume to 25 ml.

Isolation and Purification of Fusion Protein Containing Glucocorticoid Receptor Ligand Binding Domain (GR-LBD)

[0282] A single colony of *E. coli* strain BL21 Star (DE3) containing the ELP-GR-LBD fusion protein was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 µg/ml ampicillin (Sigma) and grown at 37°C with shaking at 250 rpm for 24 hours. The culture was harvested and suspended in 50 mM Hepes pH 7.5, 150 mM NaCl, 1 mM DTT, 10% glycerol, 0.1% CHAPS and 1 Complete EDTA free Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice for 3 minutes, which consisted of 10 seconds

bursts at 35% power separated by 30 second cooling down intervals. DNA and RNA in the soluble lysate were further degraded by adding 2 μ l Benzonase (Novagen) and incubating at 4°C for 30 minutes. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0283] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 2.0 M, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the ELP-GR-LBD fusion protein and non-specifically NaCl precipitated proteins.

[0284] The pellet was re-suspended in 40 ml ice-cold in 50 mM Hepes pH 7.5, 150 mM NaCl, 1 mM DTT, 10% glycerol, 0.1% CHAPS and re-centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional times to increase the purity of the ELP-GR-LBD fusion protein and reduce the final volume to 10 ml.

Isolation and Purification of Fusion Proteins Containing Estrogen Receptor Ligand Binding Domain (ER α -LBD)

[0285] A single colony of *E. coli* strain BL21 Star (DE3) containing the respective ELP-ER α -LBD fusion protein was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 μ g/ml ampicillin (Sigma), 10% sucrose (Sigma) and grown at 27°C with shaking at 250 rpm for 48 hours. The culture was harvested and suspended in 40 ml 50mM Tris-HCL pH 8.0, 150 mM KCL, 1 mM EDTA, 1 mM DTT and 1 Complete EDTA free Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice for 3 minutes, which consisted of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. DNA and RNA in the soluble lysate were further degraded by adding 2 μ l Benzonase (Novagen) and incubating at 4°C for 30 minutes. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0286] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 1.5 M, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the respective ELP-ER α -LBD fusion protein and non-specifically NaCl precipitated proteins.

[0287] The pellet was re-suspended in 40 ml ice-cold 50mM Tris-HCL pH 8.0, 150 mM KCL, 1 mM EDTA, 1 mM DTT and re-centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional

times to increase the purity of the respective ELP-ER α -LBD fusion protein and reduce the final volume to 10 ml.

Isolation and Purification of Fusion Proteins Containing G Protein Alpha Q (G α q)

[0288] A single colony of *E. coli* strain BL21 Star (DE3) containing the respective ELP-G α q fusion protein was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 μ g/ml ampicillin (Sigma) and 1 μ M GDP and grown at 37°C with shaking at 250 rpm for 24 hours. The culture was harvested and suspended in 40 ml 50mM Hepes pH 7.5, 150 mM NaCl, 1.0% CHAPS, 10% glycerol, 1 mM DTT, 10 μ M GDP and 1 Complete EDTA free Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice for 3 minutes, which consisted of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. DNA and RNA in the soluble lysate were further degraded by adding 2 μ l Benzonase (Novagen) and incubating at 4°C for 30 minutes. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0289] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 2.0 M, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the respective ELP-G α q fusion protein and non-specifically NaCl precipitated proteins.

[0290] The pellet was re-suspended in 30 ml ice-cold 50mM Hepes pH 7.5, 150 mM NaCl, 1.0% CHAPS, 10% glycerol, 1 mM DTT, 10 μ M GDP and re-centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional times to increase the purity of the respective ELP-G α q fusion protein and reduce the final volume to 5 ml.

Isolation and Purification of Fusion Proteins Containing 1-Deoxy-D-Xylulose 5-Phosphate Reductoisomerase (DXR)

[0291] A single colony of *E. coli* strain BL21 Star (DE3) containing the respective ELP-DXR fusion protein was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 μ g/ml ampicillin (Sigma), 1mM MnCl₂ (VWR) and grown at 37°C with shaking at 250 rpm for 24 hours. The culture was harvested and suspended in 40 ml 0.1M Tris pH 7.6, 1 mM DTT and 1 Complete EDTA free Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice for 3 minutes, which consisted of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. DNA and RNA in the soluble lysate were further

degraded by adding 2 μ l Benzonase (Novagen) and incubating at 4°C for 30 minutes. Cell debris was removed by centrifugation at 20,000g at 4°C for 30 minutes.

[0292] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to achieve a final concentration of 2.0 M, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the respective ELP-DXR fusion protein and non-specifically NaCl precipitated proteins.

[0293] The pellet was re-suspended in 20 ml ice-cold 0.1 M Tris pH7.6, 1mM DTT and centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional times to increase the purity of the respective ELP-DXR fusion protein and reduce the final volume to 5 ml.

Isolation and Purification of Fusion Protein Containing G Protein Alpha S (G α s)

[0294] A single colony of *E. coli* strain BL21 Star (DE3) containing the ELP-G α s fusion protein was inoculated into 500 ml CircleGrow (Q-BIOgene, San Diego, CA) supplemented with 100 μ g/ml ampicillin (Sigma) and grown at 37°C with shaking at 250 rpm for 24 hours. The culture was harvested and suspended in 40 ml PBS, 10% glycerol, 1 mM DTT and 1 Complete EDTA free Protease inhibitor pellet (Roche, Indianapolis, IN). Cells were lysed by ultrasonic disruption on ice for 3 minutes, which consisted of 10 seconds bursts at 35% power separated by 30 second cooling down intervals. DNA and RNA in the soluble lysate were further degraded by adding 2 μ l Benzonase (Novagen) and incubating at 4°C for 30 minutes. Cell debris was removed by centrifugation at 20,000g, 4°C for 30 minutes.

[0295] Inverse phase transition was induced by adding NaCl to the cell lysate at room temperature to a final concentration of 1.5 M, followed by centrifugation at 20,000g for 15 minutes at room temperature. The resulting pellet contained the ELP-G α s fusion protein and non-specifically NaCl precipitated proteins.

[0296] The pellet was re-suspended in 10 ml ice-cold PBS, 10% glycerol, 1 mM DTT and centrifuged at 20,000 g, 4°C for 15 minutes to remove the non-specifically NaCl precipitated proteins. The inverse transition cycle was repeated two additional times to increase the purity of the ELP-G α s fusion protein and reduce the final volume to 1 ml.

EXAMPLE 6:**PRODUCTION OF 10 PROTEINS WITHOUT CHROMATOGRAPHY**

[0297] The deltaPhase™ system technology, as set forth above in Example 1, was successfully tested for expression and purification of ten proteins. The results, presented in Table 2 and in the SDS-PAGE, Fig. 3 for three proteins, clearly show that diverse proteins can be purified to >95% purity. Systematic evaluation of ELP fusion protein expression and purification has been performed by having thoroughly characterized blue fluorescent protein (BFP), thioredoxin (Trx), chloramphenicol acetyltransferase (CAT), calmodulin (CalM), and angiostatin (K1-3) expressed as a fusion protein with ELP1 [V₅A₂G₃-90] and with a tag for purification by immobilized metal affinity chromatography (IMAC). Expression was performed in *E. coli*. Yields obtained for purification of the ELP fusion proteins are listed in Table 2.

Table 2: Applications of deltaPhase™ for Protein Purification

Target Proteins	MW (kDa) ^a	Yield (mg/L) ^b	ELP-Protein Fusion Activity Confirmed
Angiostatin (K1-3)	30.7	27	Yes
Blue fluorescent protein (BFP)	26.9	100	Yes
Calmodulin (CalM)	16.7	75*	Yes
Chloramphenicol acetyltransferase (CAT)	25.7	80	Yes
Green fluorescent protein (GFP)	26.9	78	Yes
Interleukin 1 receptor antagonist (IL1rRa)	17.0	8	Yes
Luciferase	60.8	10	ND**
Tissue transglutaminase (tTg)	77.0	36	Yes
Tendamistat	7.9	22	Yes
Thioredoxin (Trx)	11.7	50	Yes

Table 2. ELP1 Fusion Protein Sequences Synthesized with the deltaPhase™ System.
^a denotes the average molecular weight of the protein. ^bPurified yields indicate the best yield of target protein derived from ELP fusion. **ND=not determined

[0298] Fig. 3 shows an SDS-PAGE gel of ITC purification of BFP, CAT, and K1-3. The figure includes the soluble *E. coli* lysate (L), the supernatant following centrifugation above the T_i of the fusion protein (S), and the purified protein (P). The second gel shows purified ELP[V₅A₂G₃-90] fusions of Trx (A), BFP (B), CAT (C), K1-3 (D), GFP (E).

EXAMPLE 7:**PRODUCTION OF 10 PHARMACEUTICALLY RELEVANT PEPTIDES**

[0299] Using the deltaPhase™ system, as set forth above in Example 1, 10 pharmaceutically relevant peptides ranging in size from 2.0 to 6.2 kDa and ranging in isoelectric points from 4.11-

12.3 were expressed and purified. After extensive work varying multiple expression and purification conditions, 6 of the peptides with greater than 90% purity and yields of 17-23 mg per liter were successfully expressed and purified, as set forth in Table 3 below.

Table 3

Peptide	Amount of Fusion Produced Mg/L	Amount of Peptide Produced Mg/L	Purity
Morphine Modulating Neuropeptide (MMN)	224	17	99%
Neuropeptide Y (NPY)	222	20	98%
Orexin B	320	19	91%
Leptin	415	19	97%
ACTH	133	19	99%
Calcitonin	260	23	98%

[0300] Fusion proteins generated included: ELP4-60-MMN, ELP4-60-NPY, ELP4-60-Orexin B, ELP4-60-Leptin, ELP4-60-ACTH, ELP4-60-GH and ELP1-90-Calcitonin.

[0301] Four of the peptides proved more challenging to produce in substantial quantities. This is not surprising, given the variable nature of peptides, including size, solubility and propensity of proteolysis. Fusion proteins of the challenging peptides, ELP-adrenomedullin (AM), ELP-Parathyroid Hormone (PTH), ELP-Defensin, and ELP-growth hormone were successfully produced. However, after cleavage of the ELP from the peptide of interest, either the cleavage system with tobacco etch virus (TEV) was inadequate, or the peptide was insoluble. Only partial cleavage of ELP-growth hormone was achieved, and no peptide remained after cleavage of ELP-AM, ELP-PTH, and ELP-Defensin. These results prove the flexibility and wide-ranging application of the ELP system for the purification of therapeutically relevant peptides without chromatography.

EXAMPLE 8:

FUSION PROTEIN ACTIVITY

[0302] Fusion peptide therapeutic proteins were generated using the following four proteins: blue fluorescent protein (BFP), chloramphenicol acetyltransferase (CAT), thioredoxin (Trx), and interleukin 1 receptor antagonist (IL-Ra). Each composition was generated in both an ELP/protein and a protein/ELP orientation, utilizing ELP1 [V₅A₂G₃-90].

Linkers of the eight fusion constructs:

CAT/ELP CAT – VENLYFQGGMG (SEQ ID NO: 59) – ELP

ELP/CAT	ELP – VPGWPSSGDYDIPTTENLYFQGAH (SEQ ID NO: 60)– CAT
Trx/ELP	Trx – GSGSGHMHSHHSSGLVPRGSGK (SEQ ID NO: 61) – ELP
ELP/Trx	ELP – VPGWPSSGDYDIPTTENLYFQGAH (SEQ ID NO: 62) – Trx
BFP/ELP	BFP – VDKLAAALDMHHHHHSSGLVPRGSGK (SEQ ID NO: 63)– ELP
ELP/BFP	ELP – VPGWPSSGDYDIPTTENLYFQGAH (SEQ ID NO: 64) – BFP
IL-1Ra/ELP	IL-1Ra – LENLYFQGGMG (SEQ ID NO: 65)– ELP
ELP/IL-1Ra	ELP – VPGWPSSGDYDIPTTENLYFQGAH (SEQ ID NO: 66)– IL-1Ra

[0303] All eight protein fusion constructs have been transformed into BLR(DE3) cells, grown in triplicate in 50 mL TB media, and purified by ITC. During one round of ITC the phase transition is induced by adding NaCl to lower T_i and the large, micron-sized aggregates are collected by centrifugation. The pellets are resuspended in low ionic strength buffer followed by a cold spin to remove insoluble contaminants trapped in the ELP fusion protein pellet. Each fusion construct has been cycled through the phase transitions 3-5 times to obtain pure protein.

[0304] The yields of the protein/ELP fusions was higher than those of the ELP/protein constructs for all constructs, however the ratio between the yields in the two orientations depend on the size of the target protein (Table 4). The yields obtained for the smaller proteins Trx and IL-1Ra are significantly higher than those for the larger proteins CAT and BFP in the ELP/protein direction.

Table 4: Yields, specific activities, and transition temperatures of the eight fusion proteins

Fusion protein	Yield* (mg/L culture)	Specific activity**	T_i (°C) ***
BFP/ELP	79 ± 15	1704 ± 293	62.9 ± 0.3 67.9 ± 0.5
ELP/BFP	0.5 ± 0.06	1620 ± 111	62.4 ± 0.5
CAT/ELP	39 ± 7	8058 ± 1437	46.1 ± 0.3
ELP/CAT	2.2 ± 2.1	2984 ± 1783	47.1 ± 0.2
Trx/ELP	87 ± 4	116.6 ± 9.9	67.3 ± 0.4
ELP/Trx	27 ± 9	68.6 ± 18.0	72.9 ± 0.4
IL-1Ra/ELP	15.8 ± 4.8	2.0 ± 0.4	53.1 ± 0.4
ELP/IL-1Ra	8.2 ± 1.3	0.5 ± 0.2	55.9 ± 0.6

* The yields have been extrapolated from 50 mL cultures to 1 L.

** for Trx and CAT the specific activity is measured in U/mg, one unit corresponds to the conversion of 1 nmole substrate per minute. The specific activity for BFP is reported as the

integrated area obtained by fluorescence per mg protein (A.U./ μ g), and the activity for IL-1Ra is measured as the EC50 value in μ g/mL.

*** All fusion protein concentrations are 2 μ M and the experiments are carried out in PBS buffer.

[0305] No significant changes in activity are observed for Trx/ELP and BFP/ELP compared to the free un-fused target protein (Trabbic-Carlson K, et al. *Protein Eng. Des. Sel.* 2004, 17:57-66; Meyer DE, Chilkoti A, *Nat. Biotechnol.* 1999, 17:1112-1115), whereas the CAT/ELP shows a small decrease in activity of about 15 % compared to free CAT. Previously it is found that the IL-1Ra/ELP activity is decreased more than 100 fold compared to the free IL-1Ra which is the largest difference observed for these ELP fusion proteins (Shamji, Setton et al., accepted, in press).

[0306] The activity of Trx in the two fusion constructs have been measured by the insulin reduction assay as described by Holmgren (11. Holmgren A., *J. Biol. Chem.* 1979, 254:9627-9632; Holmgren A., Bjornstedt M., *Methods Enzymol.* 1984, 107:295-300). In the net enzymatic reaction the disulfide bonds in insulin are reduced while NADPH is oxidized to NADP⁺ which is followed spectroscopically at 340 nm. The initial rates are measured in each experiment at 25°C and converted into specific activities. The assay has been carried out three times for each of the three purified batches. The specific activities in U/mg fusion protein of the two fusion constructs are shown in Table 4 (1 U in the Trx assay is the conversion of 1 nmole substrate per minute). Differences in specific activity between the two Trx constructs have been observed; the specific activity of ELP/Trx is reduced to about 60 % of the Trx/ELP activity.

[0307] The activity of CAT fused to the ELP in the two different orientations has been determined by enzymatic acetylation of the substrate 1-deoxychloramphenicol. The activity has been measured on each of the three purifications in triplicate. The remaining substrate and the formed product are separated by thin layer chromatography before measuring the fluorescence intensity of both. The specific activities of the two CAT constructs are reported in U/mg in Table 4 where 1 U is the conversion of 1 nmole substrate per minute. Here it is seen that the specific activity of the ELP/CAT construct is reduced compared to CAT/ELP. A significant reduction is observed and only about 37 % of the activity remains in the ELP/CAT fusion protein (Table 4).

[0308] I-L1Ra competes with interleukin 1 (IL-1) for the interleukin 1 receptor and the potency of the antagonist is measured by a cell proliferation assay where active IL-1Ra inhibit the growth of the cells. Human peripheral blood leukocytes RPMI 1788 have been grown for 72 hours with and without the presence of IL-1Ra either in the form of ELP fusions or un-fused, commercially available antagonist. The proliferation has been measured by the CellTiter Glo assay. The activities of the two fusion constructs are listed in Table 4. Like CAT and Trx, IL-1Ra also show a

decrease in activity in the ELP/protein orientation and IL-1Ra/ELP is four times more potent than ELP/IL-1Ra. Comparing to un-fused IL-1Ra the free IL-1Ra is about 300 times more active than IL-1Ra/ELP (the EC₅₀ for IL-1Ra is 1.6 ng/ml).

[0309] BFP is not a biologically active protein but fluoresces in the near-UV region. Fluorescence is a sensitive measurement of changes in the tertiary structure of a protein and here it is used to evaluate structural differences between the two BFP fusion constructs. Fluorescence spectra of each BFP construct have been collected from 430 to 600 nm after excitation at 385 nm. The curves were integrated and the area normalized with protein mass. The results are listed in Table 4. The ELP/BFP used in these experiments has been grown up from two 1 L cultures in order to obtain concentrations in the same range as BFP/ELP for the fluorescence measurements. After normalizing with protein mass no significant difference is observed in fluorescence between the two BFP constructs.

[0310] The transition temperature (T_i) for fusion proteins is sensitive to the hydrophobic/hydrophilic ratio of the accessible surface area. The ELP/protein constructs are not as active as in the opposite fusions, except for BFP constructs, and if that decrease in activity is due to major structural changes the transition temperature will shift. The change in optical density of each construct has been followed from 15 to 90°C at 350 nm and T_i was derived as the mid-point of the transition (Fig. 4 and Table 4). The concentration of each fusion protein was 2 μ M, which was chosen due to the very low yields of some of the ELP/protein constructs. Fig. 4 shows the increase in turbidity as a function of temperature of 2 μ M of each of the fusion constructs in PBS buffer: A. Trx/ELP (closed circles), ELP/Trx (open circles), IL-1Ra/ELP (closed down triangles), and ELP/IL-1Ra (open down triangles) and B. BFP/ELP (closed squares), ELP/BFP (open squares), CAT/ELP (closed up triangles), ELP/CAT (open up triangles). T_i is calculated as the mid-point of each transition curve and shown in Table 4.

[0311] The transition temperatures for ELP/Trx and ELP/IL-1Ra are larger than their protein/ELP counterparts. Trx and IL-1Ra constructs differ 5.6 °C and 2.8 °C, respectively, whereas the difference between the two CAT constructs is almost negligible (Figs. 4A and B, Table 4). The ELP/BFP show one transition and form large aggregates at 62.4 °C whereas the BFP/ELP construct show a very different pattern; this fusion protein starts out forming aggregates at almost the same temperature as the ELP/BFP protein but as the temperature increases the aggregates dissociate and instead the BFP/ELP construct forms micelle-like structures. The transition temperature for the micelle-like structure formation is also reported as the mid-point of the curve and shown in Table 4 as the second transition temperature for BFP/ELP.

[0312] All eight constructs were purified by inverse transition cycling where the fusion proteins have been cycled through an aggregated phase induced by adding NaCl followed by a centrifugation step and finally the obtained pellets have been resuspended in buffer. The final yields of ELP/protein fusions after the purification process are lower compared to their respective protein/ELP constructs however smaller target proteins have higher relative yields. The lower yields are not due to significant losses during purification but a result of lower expression levels of the ELP/proteins most likely due to misfolding of the target proteins during translation. The purified ELP/proteins are assumed to fold somewhat differently from the native fold of the target protein; the specific activities are all lower in the ELP/protein orientation, except for BFP. In addition the measured transition temperatures are slightly higher for the ELP/protein constructs compared to the protein/ELP constructs, again except for BFP. The transition temperature depends on the hydrophobic/hydrophilic ratio of the fused protein indicating that the ELP/protein constructs are folded but not in a native fold.

EXAMPLE 9:
HALF-LIFE OF ELP1

[0313] The pharmacokinetics of ELP1 were determined by intravenously administering [^{14}C]ELP1 to nude mice (Balb/c nu/nu) bearing a leg/flank FaDu xenograft and collecting blood samples at various time intervals after administration. The blood concentration time-course and plasma half-lives (initial $t_{1/2\alpha}$ and terminal $t_{1/2\beta}$) are shown in Fig. 5. The blood pharmacokinetics exhibited a characteristic distribution and elimination response for macromolecules, which was well described by a bi-exponential process.

[0314] The plasma concentration time-course curve in Fig. 5 was fit to the analytical solution of a two-compartment model to approximate both an elimination and distribution response (shown as the solid line in Fig. 5) and the relevant pharmacokinetic parameters are shown in Table 5. The distribution volume of the ELP (1.338 μl) was nearly identical to the hypothetical plasma volume of 1.363 μl (Barbee, R.W., et al., Am. J. Physio. 263(3) (1992) R728-R733), indicating that the ELP did not rapidly distribute or bind to specific organs and tissues directly after administration. The AUC is a measure of the cumulative exposure to ELP in the central compartment or the blood plasma. The body clearance is defined as the rate of ELP elimination in the body relative to its plasma concentration and is the summation of clearance through all organs including the kidney, liver and others. These pharmacokinetic parameters, such as a long terminal half-life ($t_{1/2\beta} = 8.37$ hr) and low distribution volume (i.e., nearly equal to the plasma volume), are considered favorable for the delivery of therapeutics to solid tumors and potentially other disease sites. This is because

such values indicate that the ELP has properties suitable for exploiting the EPR effect in a fashion similar to that seen for other successful drug carriers (R. Duncan, Nat. Rev. Drug. Discov. 2(5) (2003) 347-360).

Table 5: Pharmacokinetic parameters calculated for [^{14}C]ELP1

	k_1 (hr^{-1})	k_2 (hr^{-1})	k_e (hr^{-1})	V_d (μL)	AUC (mg ELP hr/mL)	Cl_B ($\mu\text{L/hr}$)
ELP1-150	3.54	1.99	0.24	1,338	7.1	317

[0315] The mass transfer rate constants are from a standard two-compartment model (k_1 , from central to peripheral compartment; k_2 , from peripheral to central compartment; and k_e , elimination from central compartment). The distribution volume (V_d), central compartment concentration time-course area under the curve (AUC) and body clearance (Cl_B) are displayed. Data are shown as the mean values ($n = 5$, except V_d and initial plasma concentration (C_0) was calculated from a similar cohort with $n = 3$).

EXAMPLE 10:

BIODISTRIBUTION OF ELPs IN NUDE MICE

^{14}C labeled ELP1-150 and/or ^{14}C labeled ELP2-160

[0316] ^{14}C labeled ELP1-150 and/or ^{14}C labeled ELP2-160 were administered to nude mice with a FaDu tumor (mean $\pm$ SD, $n=6$). The tumor was heated post administration of the ELP in a water bath at 41.5°C . As can be seen in Fig. 6, the distribution is highest to the organs with the highest blood content: liver, kidneys, spleen, and lungs.

^{14}C labeled ELP2-[V₁A₈G₇-160]

[0317] ^{14}C labeled ELP2-[V₁A₈G₇-160] ($T_t > 60^\circ\text{C}$) was administered to nude mice for a plasma concentration of $15\mu\text{M}$. ELP concentrations were determined following 1 hour of heating (41°C) of an implanted FaDu tumor, located in the right hind leg of the nude mouse. Data are shown as the mean, plus the 95% confidence interval. $N=6$.

[0318] Results are shown in Fig. 7, in the graph of percent injected dose (ID) per gram (g) of tissue vs. tissue type. ELP concentration was measured 1.5 hours following systemic administration of ^{14}C labeled ELP2-[V₁A₈G₇-160]. The highest distribution is seen in organs with the highest blood content: liver, kidneys, spleen, and lungs.

[0319] While the invention has been described herein in reference to specific aspects, features and illustrative embodiments of the invention, it will be appreciated that the utility of the invention is not thus limited, but rather extends to and encompasses numerous other variations, modifications and alternative embodiments, as will suggest themselves to those of ordinary skill in the field of the present invention, based on the disclosure herein.

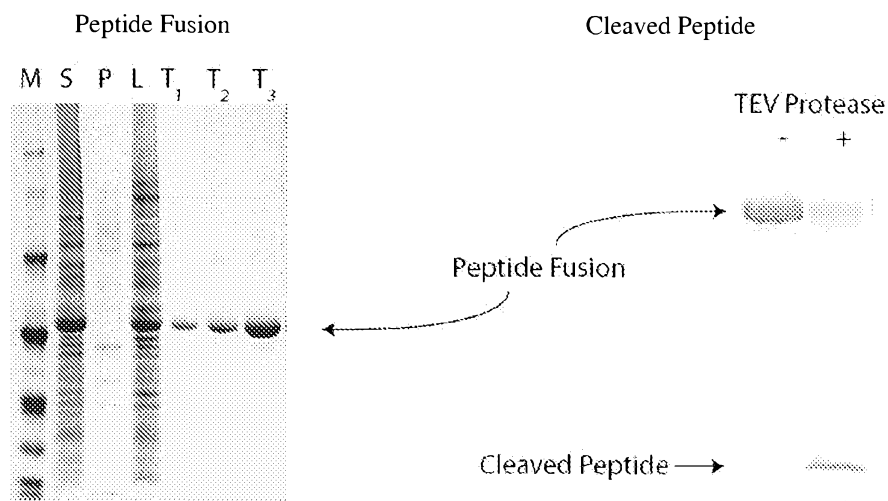
[0320] Correspondingly, the invention as hereinafter claimed is intended to be broadly construed and interpreted, as including all such variations, modifications and alternative embodiments, within its spirit and scope.

The Claims Defining the Invention are as follows:

1. A therapeutic composition comprising a fusion protein and a pharmaceutically-acceptable carrier, wherein the fusion protein comprises a therapeutic protein and at least one elastin-like protein (ELP) component, and wherein the therapeutic protein exhibits an extended half-life in circulation as compared to its unfused counterpart and wherein the therapeutic protein is glucagon-like peptide-1 (GLP-1).
5
2. The therapeutic composition of claim 1, wherein the ELP component is constructed of one or more peptide repeat units defined by SEQ ID NOS: 1-12.
10
3. The therapeutic composition of claim 2, wherein the ELP component comprises repeats of VPGXG, IPGXG, and/or LPGXG, where X is a genetically-encoded amino acid.
15
4. The therapeutic composition of claim 3, wherein the ELP component comprises VPGXG repeats, wherein each X is independently selected from V, A, and G, or is independently selected from K, V, and F.
- 20 5. The therapeutic composition of claim 4, wherein X is V, A, and G in the ratio of about V₅, A₂, and G₃.
6. The therapeutic composition of claim 4, wherein X is K, V, and F in the ratio of about K₁, V₂, and F₁.
25
7. The therapeutic composition of claim 4, wherein each X is V.
8. The therapeutic composition of any one of claims 1 to 7, wherein the ELP component comprises at least 90 repeat units of VPGXG, wherein each X is independently
30 selected from V, A, and G..

9. The therapeutic composition of any one of claims 1 to 8 having an ELP component at the C-terminus of GLP-1.
- 5 10. The therapeutic composition of any one of claims 1 to 9, further comprising a spacer sequence between the therapeutic protein and the ELP component.
11. The therapeutic composition of any one of claims 1 to 10, wherein the composition is formulated for parenteral administration.
- 10 12. The therapeutic composition of claim 11, wherein the composition is formulated for subcutaneous, intramuscular, or intravenous administration.
13. The therapeutic composition of claim 4, wherein the ELP component comprises at least 60 repeat units of VPGXG, wherein each X is independently selected from V, A, and G.
- 15 14. The therapeutic composition of claim 4, wherein the ELP component comprises at least 120 repeat units of VPGXG, wherein each X is independently selected from V, A, and G.
- 20 15. The therapeutic composition of claim 5, wherein the ELP component comprises at least 90 repeat units of VPGXG, wherein X is V, A, and G in the ratio of about V₅, A₂, and G₃.
- 25 16. The therapeutic composition of claim 5, wherein the ELP component comprises at least 60 repeat units of VPGXG, wherein X is V, A, and G in the ratio of about V₅, A₂, and G₃.
- 30 17. The therapeutic composition of claim 5, wherein the ELP component comprises at least 120 repeat units of VPGXG, wherein X is V, A, and G in the ratio of about V₅, A₂, and G₃.

18. A method of treating a subject in need of a therapeutic protein, the method comprising administering to the subject a therapeutic composition according to any one of the claims 1 to 17.
- 5 19. Use of a therapeutic composition according to any one of claims 1 to 17 in the manufacture of a medicament for the treatment of a subject in need of a therapeutic protein.



Key: M = molecular weight marker, S = lysate after sonication, P = pellet from centrifugation (pre-transition), L = soluble lysate, T_n = pellet from the nth transition

FIG. 1

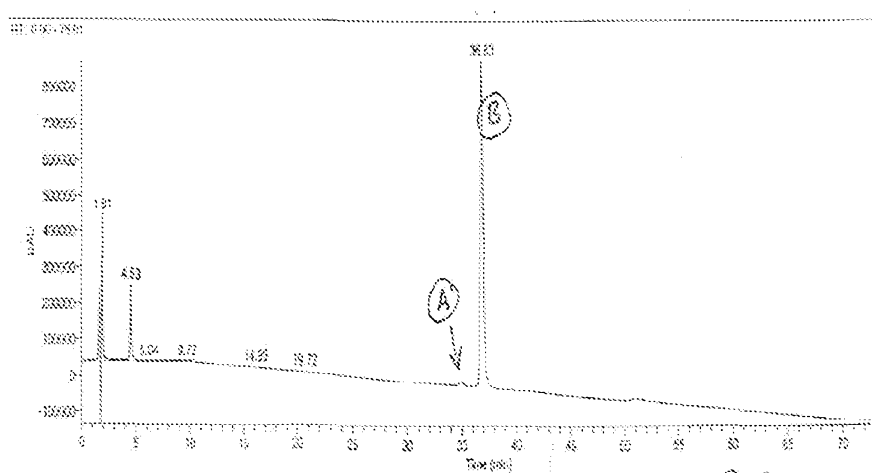


FIG. 2

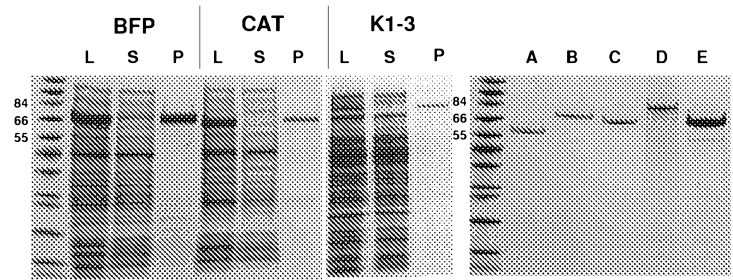


FIG. 3

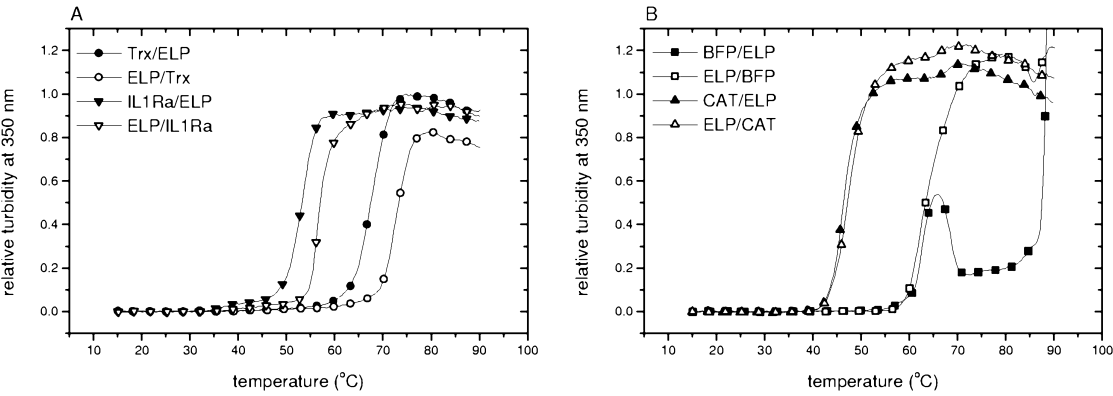
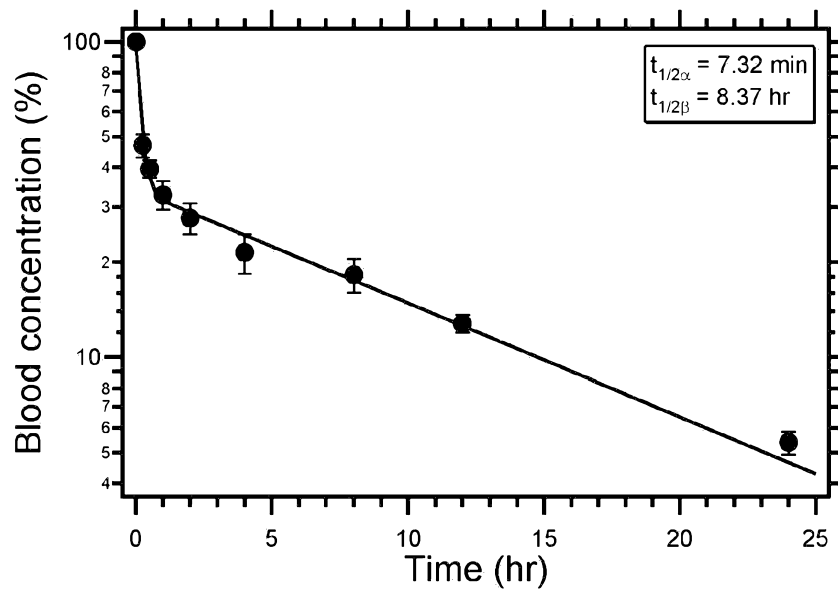


FIG. 4

**FIG. 5**

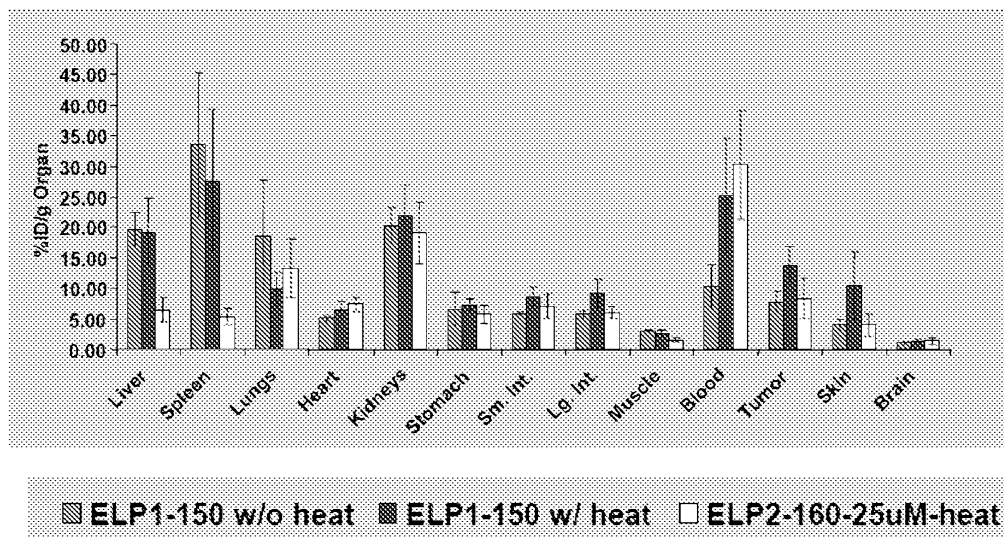


FIG. 6

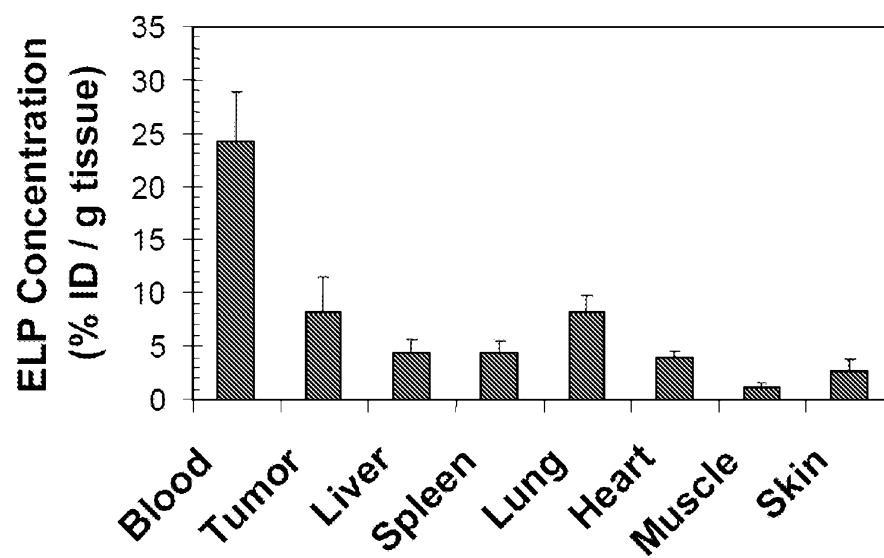


FIG. 7

103 PCT Sequence Listing.ST25
SEQUENCE LISTING

<110> Phase Bioscience, Inc.
CHILKOTI, Ashutosh

<120> FUSION PEPTIDE THERAPEUTIC COMPOSITIONS

<130> 4176-103 PCT

<150> 60/842,464

<151> 2006-09-06

<160> 66

<170> PatentIn version 3.3

<210> 1

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 1

Val Pro Gly Gly

1

<210> 2

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 2

Ile Pro Gly Gly

1

<210> 3

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> MISC_FEATURE

<222> (4)..(4)

<223> Xaa can be any amino acid, natural or non-natural

<400> 3

Val Pro Gly Xaa Gly

1

5

103 PCT Sequence Listing.ST25

<210> 4
 <211> 5
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 4

Ala Val Gly Val Pro
 1 5

<210> 5
 <211> 5
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 5

Ile Pro Gly Val Gly
 1 5

<210> 6
 <211> 5
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 6

Leu Pro Gly Val Gly
 1 5

<210> 7
 <211> 6
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 7

Val Ala Pro Gly Val Gly
 1 5

<210> 8
 <211> 8
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 8

Gly Val Gly Val Pro Gly Val Gly
1 5

<210> 9

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 9

Val Pro Gly Phe Gly Val Gly Ala Gly
1 5

<210> 10

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 10

Val Pro Gly Val Gly Val Pro Gly Gly
1 5

<210> 11

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> MISC_FEATURE

<222> (4)..(4)

<223> Xaa can be any amino acid, natural or non-natural

<400> 11

Ile Pro Gly Xaa Gly
1 5

<210> 12

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

103 PCT Sequence Listing.ST25

<220>
 <221> MISC_FEATURE
 <222> (4)..(4)
 <223> Xaa can be any amino acid, natural or non-natural

<400> 12

Leu Pro Gly Xaa Gly
 1 5

<210> 13
 <211> 50
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> REPEAT
 <222> (1)..(50)
 <223> ELP1 [V5A2G3-10]

<400> 13

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
 1 5 10 15

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro
 20 25 30

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Gly Gly Val Pro Gly
 35 40 45

Gly Gly
 50

<210> 14
 <211> 25
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> REPEAT
 <222> (1)..(25)
 <223> ELP1 [V-5]

<400> 14

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
 1 5 10 15

103 PCT Sequence Listing.ST25

Pro Gly Val Gly Val Pro Gly Val Gly
20 25

<210> 15
<211> 75
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> misc_feature
<222> (1)..(25)
<223> ELP1 [V-5] Coding Region

<400> 15
gtgggtgttc cgggcgtagg tgtcccaggt gtgggcgtac cgggcgttgg tgttcctggt 60
gtcggcgtgc cgggc 75

<210> 16
<211> 450
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> REPEAT
<222> (1)..(450)
<223> ELP1 [V5A2G3-90]

<400> 16

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro
20 25 30

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly
65 70 75 80

Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Gly Gly Val
Page 5

103 PCT Sequence Listing.ST25

85

90

95

Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
115 120 125

Ala Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
165 170 175

Pro Gly Ala Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
210 215 220

Gly Val Pro Gly Ala Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
260 265 270

Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Ala Gly Val Pro Gly
275 280 285

Gly Gly Val Pro Gly Gly Gly Val Pro Gly Gly Gly Val Pro Gly Val
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
305 310 315 320

Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Ala Gly Val
325 330 335

103 PCT Sequence Listing.ST25

Pro Gly Gly Gly Val Pro Gly Gly Gly Val Pro Gly Gly Gly Val Pro
340 345 350

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Ala
370 375 380

Gly Val Pro Gly Gly Gly Val Pro Gly Gly Gly Val Pro Gly Gly Gly
385 390 395 400

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro
420 425 430

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Gly Gly
450

<210> 17
<211> 100
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> REPEAT
<222> (1)..(100)
<223> ELP1 [V-20]

<400> 17

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
35 40 45

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
85 90 95

Pro Gly Val Gly
100

<210> 18
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> REPEAT
<222> (1)..(20)
<223> ELP1 [K1V2F1-4]

<400> 18

Val Pro Gly Lys Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Phe Gly
20

<210> 19
<211> 45
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> REPEAT
<222> (1)..(45)
<223> ELP1 [K1V7F1-9]

<400> 19

Val Pro Gly Lys Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Phe Gly
35 40 45

103 PCT Sequence Listing.ST25

<210> 20
 <211> 25
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> REPEAT
 <222> (1)..(25)
 <223> ELP2 [5]

<400> 20

Ala Val Gly Val Pro Ala Val Gly Val Pro Ala Val Gly Val Pro Ala
 1 5 10 15

Val Gly Val Pro Ala Val Gly Val Pro
 20 25

<210> 21
 <211> 25
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> REPEAT
 <222> (1)..(25)
 <223> ELP3 [V-5]

<400> 21

Ile Pro Gly Val Gly Ile Pro Gly Val Gly Ile Pro Gly Val Gly Ile
 1 5 10 15

Pro Gly Val Gly Ile Pro Gly Val Gly
 20 25

<210> 22
 <211> 25
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> REPEAT
 <222> (1)..(25)
 <223> ELP4 [V-5]

103 PCT Sequence Listing.ST25

<400> 22

Leu Pro Gly Val Gly Leu Pro Gly Val Gly Leu Pro Gly Val Gly Leu
1 5 10 15

Pro Gly Val Gly Leu Pro Gly Val Gly
20 25

<210> 23

<211> 339

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> MISC_FEATURE

<222> (1)..(339)

<223> pET32a-SD15-ELP4-60-EK-Insulin A peptide

<400> 23

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
1 5 10 15

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
130 135 140

103 PCT Sequence Listing.ST25

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
290 295 300

Gly Trp Pro Gly Ala Ser Ser Gly Thr Asp Asp Asp Asp Lys Gly Ile
305 310 315 320

Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn
325 330 335

Tyr Cys Asn

<210> 24
<211> 489
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE

103 PCT Sequence Listing.ST25

<222> (1)..(489)

<223> pET32a-SD15-ELP1-90-EK-Insulin A peptide

<400> 24

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
1 5 10 15

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
35 40 45

Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
85 90 95

Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
130 135 140

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
180 185 190

Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

103 PCT Sequence Listing.ST25

Gly Gly Val Pro Gly₂₄₅ Ala Gly Val Pro Gly₂₅₀ Gly Gly Val Pro Gly₂₅₅ Val

Gly Val Pro Gly₂₆₀ Val Gly Val Pro Gly₂₆₅ Gly Gly Val Pro Gly₂₇₀ Ala Gly

Val Pro Gly₂₇₅ Val Gly Val Pro Gly₂₈₀ Val Gly Val Pro Gly₂₈₅ Val Gly Val

Pro Gly₂₉₀ Gly Gly Val Pro Gly₂₉₅ Ala Gly Val Pro Gly₃₀₀ Gly Gly Val Pro

Gly₃₀₅ Val Gly Val Pro Gly₃₁₀ Val Gly Val Pro Gly₃₁₅ Gly Gly Val Pro Gly₃₂₀

Ala Gly Val Pro Gly₃₂₅ Val Gly Val Pro Gly₃₃₀ Val Gly Val Pro Gly₃₃₅ Val

Gly Val Pro Gly₃₄₀ Gly Gly Val Pro Gly₃₄₅ Ala Gly Val Pro Gly₃₅₀ Gly Gly

Val Pro Gly₃₅₅ Val Gly Val Pro Gly₃₆₀ Val Gly Val Pro Gly₃₆₅ Gly Gly Val

Pro Gly₃₇₀ Ala Gly Val Pro Gly₃₇₅ Val Gly Val Pro Gly₃₈₀ Val Gly Val Pro

Gly₃₈₅ Val Gly Val Pro Gly₃₉₀ Gly Gly Val Pro Gly₃₉₅ Ala Gly Val Pro Gly₄₀₀

Gly Gly Val Pro Gly₄₀₅ Val Gly Val Pro Gly₄₁₀ Val Gly Val Pro Gly₄₁₅ Gly

Gly Val Pro Gly₄₂₀ Ala Gly Val Pro Gly₄₂₅ Val Gly Val Pro Gly₄₃₀ Val Gly

Val Pro Gly₄₃₅ Val Gly Val Pro Gly₄₄₀ Gly Gly Val Pro Gly₄₄₅ Ala Gly Val

Pro Gly₄₅₀ Gly Gly Val Pro Gly₄₅₅ Trp Pro Gly Ala Ser₄₆₀ Ser Gly Thr Asp

Asp₄₆₅ Asp Asp Lys Gly₄₇₀ Ile Val Glu Gln Cys₄₇₅ Cys Thr Ser Ile Cys₄₈₀ Ser

Leu Tyr Gln Leu Glu Asn Tyr Cys Asn

485

<210> 25
 <211> 639
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> MISC_FEATURE
 <222> (1)..(639)
 <223> pET32a-SD15-ELP4-120-EK-Insulin A peptide

<400> 25

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
 1 5 10 15

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
 20 25 30

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
 35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
 50 55 60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
 65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
 85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
 100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
 115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
 130 135 140

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
 145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
 165 170 175

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
290 295 300

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
305 310 315 320

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
325 330 335

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
340 345 350

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
355 360 365

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
370 375 380

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
385 390 395 400

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
405 410 415

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
420 425 430

103 PCT Sequence Listing.ST25

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
435 440 445

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
450 455 460

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
465 470 475 480

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
485 490 495

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
500 505 510

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
515 520 525

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
530 535 540

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
545 550 555 560

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
565 570 575

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
580 585 590

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Trp Pro Gly
595 600 605

Ala Ser Ser Gly Thr Asp Asp Asp Asp Lys Gly Ile Val Glu Gln Cys
610 615 620

Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
625 630 635

<210> 26

<211> 939

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

103 PCT Sequence Listing.ST25

<221> MISC_FEATURE

<222> (1)..(939)

<223> pET32a-SD15-ELP1-180-EK-Insulin A peptide

<400> 26

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
1 5 10 15

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
35 40 45

Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
85 90 95

Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
130 135 140

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
180 185 190

Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
Page 17

103 PCT Sequence Listing.ST25

225		230		235		240
Gly Gly Val Pro	Gly ₂₄₅ Ala Gly Val Pro	Gly ₂₅₀ Gly Gly Val Pro	Gly ₂₅₅ Val			
Gly Val Pro	Gly ₂₆₀ Val Gly Val Pro	Gly ₂₆₅ Gly Gly Val Pro	Gly ₂₇₀ Ala Gly			
Val Pro	Gly ₂₇₅ Val Gly Val Pro	Gly ₂₈₀ Val Gly Val Pro	Gly ₂₈₅ Val Gly Val			
Pro	Gly ₂₉₀ Gly Gly Val Pro	Gly ₂₉₅ Ala Gly Val Pro	Gly ₃₀₀ Gly Gly Val Pro			
Gly ₃₀₅ Val Gly Val Pro	Gly ₃₁₀ Val Gly Val Pro	Gly ₃₁₅ Gly Gly Val Pro	Gly ₃₂₀			
Ala Gly Val Pro	Gly ₃₂₅ Val Gly Val Pro	Gly ₃₃₀ Val Gly Val Pro	Gly ₃₃₅ Val			
Gly Val Pro	Gly ₃₄₀ Gly Gly Val Pro	Gly ₃₄₅ Ala Gly Val Pro	Gly ₃₅₀ Gly Gly			
Val Pro	Gly ₃₅₅ Val Gly Val Pro	Gly ₃₆₀ Val Gly Val Pro	Gly ₃₆₅ Gly Gly Val			
Pro	Gly ₃₇₀ Ala Gly Val Pro	Gly ₃₇₅ Val Gly Val Pro	Gly ₃₈₀ Val Gly Val Pro			
Gly ₃₈₅ Val Gly Val Pro	Gly ₃₉₀ Gly Gly Val Pro	Gly ₃₉₅ Ala Gly Val Pro	Gly ₄₀₀			
Gly Gly Val Pro	Gly ₄₀₅ Val Gly Val Pro	Gly ₄₁₀ Val Gly Val Pro	Gly ₄₁₅ Gly			
Gly Val Pro	Gly ₄₂₀ Ala Gly Val Pro	Gly ₄₂₅ Val Gly Val Pro	Gly ₄₃₀ Val Gly			
Val Pro	Gly ₄₃₅ Val Gly Val Pro	Gly ₄₄₀ Gly Gly Val Pro	Gly ₄₄₅ Ala Gly Val			
Pro	Gly ₄₅₀ Gly Gly Val Pro	Gly ₄₅₅ Val Gly Val Pro	Gly ₄₆₀ Val Gly Val Pro			
Gly ₄₆₅ Gly Gly Val Pro	Gly ₄₇₀ Ala Gly Val Pro	Gly ₄₇₅ Val Gly Val Pro	Gly ₄₈₀			

103 PCT Sequence Listing.ST25

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
485 490 495

Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
500 505 510

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val
515 520 525

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
530 535 540

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly
545 550 555 560

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val
565 570 575

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
580 585 590

Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val
595 600 605

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
610 615 620

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
625 630 635 640

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val
645 650 655

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
660 665 670

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
675 680 685

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro
690 695 700

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
705 710 715 720

Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
725 730 735

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly
740 745 750

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
755 760 765

Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
770 775 780

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
785 790 795 800

Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
805 810 815

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
820 825 830

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
835 840 845

Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
850 855 860

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly
865 870 875 880

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
885 890 895

Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Gly Ala Ser Ser Gly
900 905 910

Thr Asp Asp Asp Asp Lys Gly Ile Val Glu Gln Cys Cys Thr Ser Ile
915 920 925

Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
930 935

<210> 27
<211> 354
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>

103 PCT Sequence Listing.ST25

<221> MISC_FEATURE

<222> (1)..(354)

<223> pET15b-ELP4-60-EK-T20 peptide

<400> 27

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
1 5 10 15

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
130 135 140

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
Page 21

103 PCT Sequence Listing.ST25

225 230 235 240

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
290 295 300

Gly Trp Pro Gly Ala Ser Ser Gly Thr Asp Asp Asp Asp Lys Tyr Thr
305 310 315 320

Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys
325 330 335

Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn
340 345 350

Trp Phe

<210> 28
<211> 504
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(504)
<223> pET17b-ELP1-90-EK-T20 peptide

<400> 28

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
1 5 10 15

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
35 40 45

103 PCT Sequence Listing.ST25

Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
85 90 95

Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
130 135 140

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
180 185 190

Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro
290 295 300

103 PCT Sequence Listing.ST25

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
305 310 315 320

Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
325 330 335

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly
340 345 350

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
355 360 365

Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
370 375 380

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
385 390 395 400

Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
405 410 415

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
420 425 430

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
435 440 445

Pro Gly Gly Gly Val Pro Gly Trp Pro Gly Ala Ser Ser Gly Thr Asp
450 455 460

Asp Asp Asp Lys Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser
465 470 475 480

Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys
485 490 495

Trp Ala Ser Leu Trp Asn Trp Phe
500

<210> 29
<211> 654
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>

103 PCT Sequence Listing.ST25

<221> MISC_FEATURE

<222> (1)..(654)

<223> pET15b-ELP4-120-EK-T20 peptide

<400> 29

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
1 5 10 15

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
130 135 140

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
Page 25

103 PCT Sequence Listing.ST25

225		230		235		240
Val	Gly	Val	Pro	Gly ₂₄₅	Val	Gly
					Val	Pro
				Gly ₂₅₀	Val	Gly
					Val	Pro
					Gly ₂₅₅	Val
Gly	Val	Pro	Gly ₂₆₀	Val	Gly	Val
				Pro	Gly ₂₆₅	Val
					Val	Gly
					Val	Pro
					Gly ₂₇₀	Val
					Val	Gly
Val	Pro	Gly ₂₇₅	Val	Gly	Val	Pro
				Gly ₂₈₀	Val	Gly
					Val	Pro
					Gly ₂₈₅	Val
					Val	Gly
					Val	Gly
Pro	Gly ₂₉₀	Val	Gly	Val	Pro	Gly ₂₉₅
					Val	Gly
					Val	Pro
					Gly ₃₀₀	Val
					Val	Gly
					Val	Pro
Gly	Val	Gly	Val	Pro	Gly ₃₁₀	Val
					Val	Gly
					Val	Pro
					Gly ₃₁₅	Val
					Val	Gly
					Val	Pro
					Gly ₃₂₀	Val
Val	Gly	Val	Pro	Gly ₃₂₅	Val	Gly
					Val	Pro
					Gly ₃₃₀	Val
					Val	Gly
					Val	Pro
					Gly ₃₃₅	Val
Gly	Val	Pro	Gly ₃₄₀	Val	Gly	Val
				Pro	Gly ₃₄₅	Val
					Val	Gly
					Val	Pro
					Gly ₃₅₀	Val
					Val	Gly
Val	Pro	Gly ₃₅₅	Val	Gly	Val	Pro
				Gly ₃₆₀	Val	Gly
					Val	Pro
					Gly ₃₆₅	Val
					Val	Gly
					Val	Gly
Pro	Gly ₃₇₀	Val	Gly	Val	Pro	Gly ₃₇₅
					Val	Gly
					Val	Pro
					Gly ₃₈₀	Val
					Val	Gly
					Val	Pro
Gly	Val	Gly	Val	Pro	Gly ₃₉₀	Val
					Val	Gly
					Val	Pro
					Gly ₃₉₅	Val
					Val	Gly
					Val	Pro
					Gly ₄₀₀	Val
Val	Gly	Val	Pro	Gly ₄₀₅	Val	Gly
					Val	Pro
					Gly ₄₁₀	Val
					Val	Gly
					Val	Pro
					Gly ₄₁₅	Val
Gly	Val	Pro	Gly ₄₂₀	Val	Gly	Val
				Pro	Gly ₄₂₅	Val
					Val	Gly
					Val	Pro
					Gly ₄₃₀	Val
					Val	Gly
Val	Pro	Gly ₄₃₅	Val	Gly	Val	Pro
				Gly ₄₄₀	Val	Gly
					Val	Pro
					Gly ₄₄₅	Val
					Val	Gly
					Val	Gly
Pro	Gly ₄₅₀	Val	Gly	Val	Pro	Gly ₄₅₅
					Val	Gly
					Val	Pro
					Gly ₄₆₀	Val
					Val	Gly
					Val	Pro
Gly	Val	Gly	Val	Pro	Gly ₄₇₀	Val
					Val	Gly
					Val	Pro
					Gly ₄₇₅	Val
					Val	Gly
					Val	Pro
					Gly ₄₈₀	Val

103 PCT Sequence Listing.ST25

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
485 490 495

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
500 505 510

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
515 520 525

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
530 535 540

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
545 550 555 560

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
565 570 575

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
580 585 590

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Trp Pro Gly
595 600 605

Ala Ser Ser Gly Thr Asp Asp Asp Asp Lys Tyr Thr Ser Leu Ile His
610 615 620

Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu
625 630 635 640

Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe
645 650

<210> 30

<211> 357

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> MISC_FEATURE

<222> (1)..(357)

<223> pET17b-ELP4-60-Throm-T20 peptide

<400> 30

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
1 5 10 15

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
130 135 140

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
260 265 270

103 PCT Sequence Listing.ST25

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
290 295 300

Gly Trp Pro Gly Ala Ser Gly Gly Gly Gly Pro Leu Val Pro Arg Gly
305 310 315 320

Ser Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln
325 330 335

Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser
340 345 350

Leu Trp Asn Trp Phe
355

<210> 31
<211> 507
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(507)
<223> pET17b-ELP1-90-Throm-T20 peptide

<400> 31

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
1 5 10 15

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
35 40 45

Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
Page 29

103 PCT Sequence Listing.ST25

85

90

95

Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
130 135 140

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
180 185 190

Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro
290 295 300

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
305 310 315 320

Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
325 330 335

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly
340 345 350

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
355 360 365

Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
370 375 380

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
385 390 395 400

Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
405 410 415

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
420 425 430

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
435 440 445

Pro Gly Gly Gly Val Pro Gly Trp Pro Gly Ala Ser Gly Gly Gly Gly
450 455 460

Pro Leu Val Pro Arg Gly Ser Tyr Thr Ser Leu Ile His Ser Leu Ile
465 470 475 480

Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu
485 490 495

Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe
500 505

<210> 32
<211> 657
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(657)
<223> pET17b-ELP4-120-Throm-T20 peptide

<400> 32

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
1 5 10 15

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
130 135 140

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
260 265 270

103 PCT Sequence Listing.ST25

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
290 295 300

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
305 310 315 320

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
325 330 335

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
340 345 350

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
355 360 365

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
370 375 380

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
385 390 395 400

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
405 410 415

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
420 425 430

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
435 440 445

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
450 455 460

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
465 470 475 480

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
485 490 495

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
500 505 510

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val

103 PCT Sequence Listing.ST25
520 525

515

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
530 535 540

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
545 550 555 560

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
565 570 575

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
580 585 590

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Trp Pro Gly
595 600 605

Ala Ser Gly Gly Gly Gly Pro Leu Val Pro Arg Gly Ser Tyr Thr Ser
610 615 620

Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn
625 630 635 640

Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp
645 650 655

Phe

<210> 33
<211> 357
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(357)
<223> pET17b-ELP4-60-TEV(Q/S)-T20 peptide

<400> 33

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
1 5 10 15

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

103 PCT Sequence Listing.ST25

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
130 135 140

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

103 PCT Sequence Listing.ST25

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
290 295 300

Gly Trp Pro Gly Ala Ser Gly Pro Thr Thr Glu Asn Leu Tyr Phe Gln
305 310 315 320

Ser Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln
325 330 335

Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser
340 345 350

Leu Trp Asn Trp Phe
355

<210> 34
<211> 507
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(507)
<223> pET17b-ELP1-90-TEV(Q/S)-T20 peptide

<400> 34

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
1 5 10 15

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
35 40 45

Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
85 90 95

Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

103 PCT Sequence Listing.ST25

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
130 135 140

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
180 185 190

Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro
290 295 300

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
305 310 315 320

Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
325 330 335

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly
340 345 350

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
Page 37

355

Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
370 375 380

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
385 390 395 400

Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
405 410 415

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
420 425 430

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
435 440 445

Pro Gly Gly Gly Val Pro Gly Trp Pro Gly Ala Ser Gly Pro Thr Thr
450 455 460

Glu Asn Leu Tyr Phe Gln Ser Tyr Thr Ser Leu Ile His Ser Leu Ile
465 470 475 480

Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu
485 490 495

Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe
500 505

<210> 35
<211> 657
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(657)
<223> pET17b-ELP4-120-TEV(Q/S)-T20 peptide

<400> 35

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
1 5 10 15

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

103 PCT Sequence Listing.ST25

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
130 135 140

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

103 PCT Sequence Listing.ST25

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
290 295 300

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
305 310 315 320

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
325 330 335

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
340 345 350

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
355 360 365

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
370 375 380

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
385 390 395 400

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
405 410 415

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
420 425 430

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
435 440 445

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
450 455 460

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
465 470 475 480

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
485 490 495

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
500 505 510

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
515 520 525

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
530 535 540

103 PCT Sequence Listing.ST25

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
545 550 555 560

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
565 570 575

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
580 585 590

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Trp Pro Gly
595 600 605

Ala Ser Gly Pro Thr Thr Glu Asn Leu Tyr Phe Gln Ser Tyr Thr Ser
610 615 620

Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn
625 630 635 640

Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp
645 650 655

Phe

<210> 36
<211> 356
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(356)
<223> pET17b-ELP4-60-TEV(Q/Y)-T20 peptide

<400> 36

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
1 5 10 15

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
Page 41

103 PCT Sequence Listing.ST25

50

55

60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
130 135 140

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
290 295 300

103 PCT Sequence Listing.ST25

Gly Trp Pro Gly Ala Ser Gly Pro Thr Thr Glu Asn Leu Tyr Phe Gln
305 310 315 320

Tyr Thr Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln
325 330 335

Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu
340 345 350

Trp Asn Trp Phe
355

<210> 37
<211> 506
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(506)
<223> pET17b-ELP1-90-TEV(Q/Y)-T20 peptide

<400> 37

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
1 5 10 15

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
35 40 45

Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
50 55 60

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
85 90 95

Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val
115 120 125

103 PCT Sequence Listing.ST25

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
130 135 140

Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
180 185 190

Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro
290 295 300

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
305 310 315 320

Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
325 330 335

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly
340 345 350

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
355 360 365

Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
370 375 380

103 PCT Sequence Listing.ST25

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
385 390 395 400

Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
405 410 415

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
420 425 430

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
435 440 445

Pro Gly Gly Gly Val Pro Gly Trp Pro Gly Ala Ser Gly Pro Thr Thr
450 455 460

Glu Asn Leu Tyr Phe Gln Tyr Thr Ser Leu Ile His Ser Leu Ile Glu
465 470 475 480

Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu
485 490 495

Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe
500 505

<210> 38
<211> 656
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(656)
<223> pET17b-ELP4-120-TEV(Q/Y)-T20 peptide

<400> 38

Met Gly Gly Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
1 5 10 15

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
20 25 30

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
35 40 45

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
Page 45

103 PCT Sequence Listing.ST25

50

55

60

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
65 70 75 80

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
85 90 95

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
100 105 110

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
115 120 125

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
130 135 140

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
145 150 155 160

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
165 170 175

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
180 185 190

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
195 200 205

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
210 215 220

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
225 230 235 240

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
245 250 255

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
260 265 270

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
275 280 285

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
290 295 300

103 PCT Sequence Listing.ST25

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
305 310 315 320

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
325 330 335

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
340 345 350

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
355 360 365

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
370 375 380

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
385 390 395 400

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
405 410 415

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
420 425 430

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
435 440 445

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
450 455 460

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
465 470 475 480

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
485 490 495

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
500 505 510

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
515 520 525

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
530 535 540

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
545 550 555 560

103 PCT Sequence Listing.ST25

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
565 570 575

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
580 585 590

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Trp Pro Gly
595 600 605

Ala Ser Gly Pro Thr Thr Glu Asn Leu Tyr Phe Gln Tyr Thr Ser Leu
610 615 620

Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu
625 630 635 640

Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe
645 650 655

<210> 39
<211> 669
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(669)
<223> pET32a-SD11-ELP1-90-throm-Interferon Alpha 2B

<400> 39

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

103 PCT Sequence Listing.ST25

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro

103 PCT Sequence Listing.ST25
345 350

340

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Leu
450 455 460

Val Pro Arg Gly Ser Pro Gly Ile Ser Gly Gly Gly Gly Gly His Met
465 470 475 480

Pro Met Ala Leu Thr Phe Ala Leu Leu Val Ala Leu Leu Val Leu Ser
485 490 495

Cys Lys Ser Ser Cys Ser Val Gly Cys Asp Leu Pro Gln Thr His Ser
500 505 510

Leu Gly Ser Arg Arg Thr Leu Met Leu Leu Ala Gln Met Arg Arg Ile
515 520 525

Ser Leu Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Gln
530 535 540

Glu Glu Phe Gly Asn Gln Phe Gln Lys Ala Glu Thr Ile Pro Val Leu
545 550 555 560

His Glu Met Ile Gln Gln Ile Phe Asn Leu Phe Ser Thr Lys Asp Ser
565 570 575

Ser Ala Ala Trp Asp Glu Thr Leu Leu Asp Lys Phe Tyr Thr Glu Leu
580 585 590

103 PCT Sequence Listing.ST25
Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Gly Val Gly
595 600 605

Val Thr Glu Thr Pro Leu Met Lys Glu Asp Ser Ile Leu Ala Val Arg
610 615 620

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Lys Glu Lys Lys Tyr Ser
625 630 635 640

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
645 650 655

Leu Ser Thr Asn Leu Gln Glu Ser Leu Arg Ser Lys Glu
660 665

<210> 40
<211> 574
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(574)
<223> pET15b-SD5-ELP4-60-throm-Tobacco etch virus protease

<400> 40

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
35 40 45

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
85 90 95

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
100 105 110

103 PCT Sequence Listing.ST25

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
130 135 140

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
260 265 270

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
290 295 300

Gly Val Pro Gly Trp Pro Ser Ser Gly Leu Val Pro Arg Gly Ser Pro
305 310 315 320

Gly Ile Ser Gly Gly Gly Gly His Met Pro Met Gly Glu Ser Leu
325 330 335

Phe Lys Gly Pro Arg Asp Tyr Asn Pro Ile Ser Ser Thr Ile Cys His
340 345 350

Leu Thr Asn Glu Ser Asp Gly His Thr Thr Ser Leu Tyr Gly Ile Gly
355 360 365

103 PCT Sequence Listing.ST25

Phe Gly Pro Phe Ile Ile Thr Asn Lys His Leu Phe Arg Arg Asn Asn
370 375 380

Gly Thr Leu Leu Val Gln Ser Leu His Gly Val Phe Lys Val Lys Asn
385 390 395 400

Thr Thr Thr Leu Gln Gln His Leu Ile Asp Gly Arg Asp Met Ile Ile
405 410 415

Ile Arg Met Pro Lys Asp Phe Pro Pro Phe Pro Gln Lys Leu Lys Phe
420 425 430

Arg Glu Pro Gln Arg Glu Glu Arg Ile Cys Leu Val Thr Thr Asn Phe
435 440 445

Gln Thr Lys Ser Met Ser Ser Met Val Ser Asp Thr Ser Cys Thr Phe
450 455 460

Pro Ser Ser Asp Gly Ile Phe Trp Lys His Trp Ile Gln Thr Lys Asp
465 470 475 480

Gly Gln Cys Gly Ser Pro Leu Val Ser Thr Arg Asp Gly Phe Ile Val
485 490 495

Gly Ile His Ser Ala Ser Asn Phe Thr Asn Thr Asn Asn Tyr Phe Thr
500 505 510

Ser Val Pro Lys Asn Phe Met Glu Leu Leu Thr Asn Gln Glu Ala Gln
515 520 525

Gln Trp Val Ser Gly Trp Arg Leu Asn Ala Asp Ser Val Leu Trp Gly
530 535 540

Gly His Lys Val Phe Met Ser Lys Pro Glu Glu Pro Phe Gln Pro Val
545 550 555 560

Lys Glu Ala Thr Gln Leu Met Asn Glu Leu Val Tyr Ser Gln
565 570

<210> 41
<211> 724
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

103 PCT Sequence Listing.ST25

<220>

<221> MISC_FEATURE

<222> (1)..(724)

<223> pET15b-SD5-ELP1-90-throm-Tobacco etch virus protease

<400> 41

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Leu
450 455 460

Val Pro Arg Gly Ser Pro Gly Ile Ser Gly Gly Gly Gly Gly His Met
465 470 475 480

103 PCT Sequence Listing.ST25

Pro Met Gly Glu Ser₄₈₅ Leu Phe Lys Gly Pro₄₉₀ Arg Asp Tyr Asn₄₉₅ Ile

Ser Ser Thr Ile₅₀₀ Cys His Leu Thr Asn₅₀₅ Glu Ser Asp Gly His₅₁₀ Thr Thr

Ser Leu Tyr₅₁₅ Gly Ile Gly Phe Gly₅₂₀ Pro Phe Ile Ile Thr₅₂₅ Asn Lys His

Leu Phe₅₃₀ Arg Arg Asn Asn Gly₅₃₅ Thr Leu Leu Val Gln₅₄₀ Ser Leu His Gly

Val₅₄₅ Phe Lys Val Lys Asn₅₅₀ Thr Thr Thr Leu Gln₅₅₅ Gln His Leu Ile Asp₅₆₀

Gly Arg Asp Met Ile₅₆₅ Ile Ile Arg Met Pro₅₇₀ Lys Asp Phe Pro₅₇₅ Pro Phe

Pro Gln Lys Leu₅₈₀ Lys Phe Arg Glu Pro₅₈₅ Gln Arg Glu Glu Arg₅₉₀ Ile Cys

Leu Val Thr₅₉₅ Thr Asn Phe Gln Thr₆₀₀ Lys Ser Met Ser Ser₆₀₅ Met Val Ser

Asp Thr₆₁₀ Ser Cys Thr Phe Pro₆₁₅ Ser Ser Asp Gly Ile₆₂₀ Phe Trp Lys His

Trp₆₂₅ Ile Gln Thr Lys Asp₆₃₀ Gly Gln Cys Gly Ser₆₃₅ Pro Leu Val Ser Thr₆₄₀

Arg Asp Gly Phe Ile₆₄₅ Val Gly Ile His Ser₆₅₀ Ala Ser Asn Phe Thr₆₅₅ Asn

Thr Asn Asn Tyr₆₆₀ Phe Thr Ser Val Pro₆₆₅ Lys Asn Phe Met Glu₆₇₀ Leu Leu

Thr Asn Gln₆₇₅ Glu Ala Gln Gln Trp Val Ser Gly Trp Arg₆₈₅ Leu Asn Ala

Asp Ser₆₉₀ Val Leu Trp Gly Gly₆₉₅ His Lys Val Phe Met₇₀₀ Ser Lys Pro Glu

Glu₇₀₅ Pro Phe Gln Pro Val₇₁₀ Lys Glu Ala Thr Gln₇₁₅ Leu Met Asn Glu₇₂₀ Leu

Val Tyr Ser Gln

103 PCT Sequence Listing.ST25

<210> 42
 <211> 874
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> MISC_FEATURE
 <222> (1)..(874)
 <223> pET15b-SD5-ELP4-120-throm-Tobacco etch virus protease

<400> 42

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
 1 5 10 15

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
 20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
 35 40 45

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
 50 55 60

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
 65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
 85 90 95

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
 100 105 110

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
 115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
 130 135 140

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
 145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
 165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
 Page 57

103 PCT Sequence Listing.ST25
185 190

180

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
260 265 270

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
305 310 315 320

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
340 345 350

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
385 390 395 400

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
420 425 430

103 PCT Sequence Listing.ST25

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
435 440 445

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
450 455 460

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
465 470 475 480

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
485 490 495

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
500 505 510

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
515 520 525

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
530 535 540

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
545 550 555 560

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
565 570 575

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
580 585 590

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
595 600 605

Trp Pro Ser Ser Gly Leu Val Pro Arg Gly Ser Pro Gly Ile Ser Gly
610 615 620

Gly Gly Gly Gly His Met Pro Met Gly Glu Ser Leu Phe Lys Gly Pro
625 630 635 640

Arg Asp Tyr Asn Pro Ile Ser Ser Thr Ile Cys His Leu Thr Asn Glu
645 650 655

Ser Asp Gly His Thr Thr Ser Leu Tyr Gly Ile Gly Phe Gly Pro Phe
660 665 670

Ile Ile Thr Asn Lys His Leu Phe Arg Arg Asn Asn Gly Thr Leu Leu
675 680 685

103 PCT Sequence Listing.ST25

Val Gln Ser Leu His Gly Val Phe Lys Val Lys Asn Thr Thr Thr Leu
690 695 700

Gln Gln His Leu Ile Asp Gly Arg Asp Met Ile Ile Ile Arg Met Pro
705 710 715 720

Lys Asp Phe Pro Pro Phe Pro Gln Lys Leu Lys Phe Arg Glu Pro Gln
725 730 735

Arg Glu Glu Arg Ile Cys Leu Val Thr Thr Asn Phe Gln Thr Lys Ser
740 745 750

Met Ser Ser Met Val Ser Asp Thr Ser Cys Thr Phe Pro Ser Ser Asp
755 760 765

Gly Ile Phe Trp Lys His Trp Ile Gln Thr Lys Asp Gly Gln Cys Gly
770 775 780

Ser Pro Leu Val Ser Thr Arg Asp Gly Phe Ile Val Gly Ile His Ser
785 790 795 800

Ala Ser Asn Phe Thr Asn Thr Asn Asn Tyr Phe Thr Ser Val Pro Lys
805 810 815

Asn Phe Met Glu Leu Leu Thr Asn Gln Glu Ala Gln Gln Trp Val Ser
820 825 830

Gly Trp Arg Leu Asn Ala Asp Ser Val Leu Trp Gly Gly His Lys Val
835 840 845

Phe Met Ser Lys Pro Glu Glu Pro Phe Gln Pro Val Lys Glu Ala Thr
850 855 860

Gln Leu Met Asn Glu Leu Val Tyr Ser Gln
865 870

<210> 43
<211> 1174
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(1174)
<223> pET15b-SD5-ELP1-180-throm-Tobacco etch virus protease

103 PCT Sequence Listing.ST25

<400> 43

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val

103 PCT Sequence Listing.ST25

245

250

255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
450 455 460

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
465 470 475 480

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
485 490 495

103 PCT Sequence Listing.ST25

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
500 505 510

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
515 520 525

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
530 535 540

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
545 550 555 560

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
565 570 575

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
580 585 590

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
595 600 605

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
610 615 620

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
625 630 635 640

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
645 650 655

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
660 665 670

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
675 680 685

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
690 695 700

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
705 710 715 720

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
725 730 735

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
740 745 750

103 PCT Sequence Listing.ST25

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
755 760 765

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
770 775 780

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
785 790 795 800

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
805 810 815

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
820 825 830

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
835 840 845

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
850 855 860

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
865 870 875 880

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
885 890 895

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser
900 905 910

Gly Leu Val Pro Arg Gly Ser Pro Gly Ile Ser Gly Gly Gly Gly Gly
915 920 925

His Met Pro Met Gly Glu Ser Leu Phe Lys Gly Pro Arg Asp Tyr Asn
930 935 940

Pro Ile Ser Ser Thr Ile Cys His Leu Thr Asn Glu Ser Asp Gly His
945 950 955 960

Thr Thr Ser Leu Tyr Gly Ile Gly Phe Gly Pro Phe Ile Ile Thr Asn
965 970 975

Lys His Leu Phe Arg Arg Asn Asn Gly Thr Leu Leu Val Gln Ser Leu
980 985 990

His Gly Val Phe Lys Val Lys Asn Thr Thr Thr Leu Gln Gln His Leu
995 1000 1005

103 PCT Sequence Listing.ST25

Ile	Asp	Gly	Arg	Asp	Met	Ile	Ile	Ile	Arg	Met	Pro	Lys	Asp	Phe
	1010					1015					1020			
Pro	Pro	Phe	Pro	Gln	Lys	Leu	Lys	Phe	Arg	Glu	Pro	Gln	Arg	Glu
	1025					1030					1035			
Glu	Arg	Ile	Cys	Leu	Val	Thr	Thr	Asn	Phe	Gln	Thr	Lys	Ser	Met
	1040					1045					1050			
Ser	Ser	Met	Val	Ser	Asp	Thr	Ser	Cys	Thr	Phe	Pro	Ser	Ser	Asp
	1055					1060					1065			
Gly	Ile	Phe	Trp	Lys	His	Trp	Ile	Gln	Thr	Lys	Asp	Gly	Gln	Cys
	1070					1075					1080			
Gly	Ser	Pro	Leu	Val	Ser	Thr	Arg	Asp	Gly	Phe	Ile	Val	Gly	Ile
	1085					1090					1095			
His	Ser	Ala	Ser	Asn	Phe	Thr	Asn	Thr	Asn	Asn	Tyr	Phe	Thr	Ser
	1100					1105					1110			
Val	Pro	Lys	Asn	Phe	Met	Glu	Leu	Leu	Thr	Asn	Gln	Glu	Ala	Gln
	1115					1120					1125			
Gln	Trp	Val	Ser	Gly	Trp	Arg	Leu	Asn	Ala	Asp	Ser	Val	Leu	Trp
	1130					1135					1140			
Gly	Gly	His	Lys	Val	Phe	Met	Ser	Lys	Pro	Glu	Glu	Pro	Phe	Gln
	1145					1150					1155			
Pro	Val	Lys	Glu	Ala	Thr	Gln	Leu	Met	Asn	Glu	Leu	Val	Tyr	Ser
	1160					1165					1170			

Gln

<210> 44
 <211> 735
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> MISC_FEATURE
 <222> (1)..(735)
 <223> pET15b-SD3-ELP1-90-throm-Small Heterodimer partner orphan

103 PCT Sequence Listing.ST25

receptor

<400> 44

```

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
 1      5      10     15
Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
 20     25     30
Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
 35     40     45
Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
 50     55     60
Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
 65     70     75     80
Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
 85     90     95
Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100    105    110
Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115    120    125
Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130    135    140
Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145    150    155    160
Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165    170    175
Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180    185    190
Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195    200    205
Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210    215    220
Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225    230    235    240

```

103 PCT Sequence Listing.ST25

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Gly
450 455 460

Gly Gly Gly Ser Ile Gly Pro Leu Val Pro Arg Gly Ser His Met Ser
465 470 475 480

Thr Ser Gln Pro Gly Ala Cys Pro Cys Gln Gly Ala Ala Ser Arg Pro
485 490 495

103 PCT Sequence Listing.ST25

Ala Ile Leu Tyr Ala Leu Leu Ser Ser Ser Leu Lys Ala Val Pro Arg
500 505 510

Pro Arg Ser Arg Cys Leu Cys Arg Gln His Arg Pro Val Gln Leu Cys
515 520 525

Ala Pro His Arg Thr Cys Arg Glu Ala Leu Asp Val Leu Ala Lys Thr
530 535 540

Val Ala Phe Leu Arg Asn Leu Pro Ser Phe Trp Gln Leu Pro Pro Gln
545 550 555 560

Asp Gln Arg Arg Leu Leu Gln Gly Cys Trp Gly Pro Leu Phe Leu Leu
565 570 575

Gly Leu Ala Gln Asp Ala Val Thr Phe Glu Val Ala Glu Ala Pro Val
580 585 590

Pro Ser Ile Leu Lys Lys Ile Leu Leu Glu Glu Pro Ser Ser Ser Gly
595 600 605

Gly Ser Gly Gln Leu Pro Asp Arg Pro Gln Pro Ser Leu Ala Ala Val
610 615 620

Gln Trp Leu Gln Cys Cys Leu Glu Ser Phe Trp Ser Leu Glu Leu Ser
625 630 635 640

Pro Lys Glu Tyr Ala Cys Leu Lys Gly Thr Ile Leu Phe Asn Pro Asp
645 650 655

Val Pro Gly Leu Gln Ala Ala Ser His Ile Gly His Leu Gln Gln Glu
660 665 670

Ala His Trp Val Leu Cys Glu Val Leu Glu Pro Trp Cys Pro Ala Ala
675 680 685

Gln Gly Arg Leu Thr Arg Val Leu Leu Thr Ala Ser Thr Leu Lys Ser
690 695 700

Ile Pro Thr Ser Leu Leu Gly Asp Leu Phe Phe Arg Pro Ile Ile Gly
705 710 715 720

Asp Val Asp Ile Ala Gly Leu Leu Gly Asp Met Leu Leu Leu Arg
725 730 735

<210> 45

103 PCT Sequence Listing.ST25

<211> 736
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Synthetic Construct
 <220>
 <221> MISC_FEATURE
 <222> (1)..(736)
 <223> pET15b-SD3-ELP1-90-throm-Androgen receptor ligand binding domain
 <400> 45

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
 1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
 20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
 35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
 50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
 65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
 85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
 100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
 115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
 130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
 145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
 165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
 180 185 190

103 PCT Sequence Listing.ST25

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

103 PCT Sequence Listing.ST25

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Gly
450 455 460

Gly Gly Gly Ser Ile Gly Pro Leu Val Pro Arg Gly Ser His Met His
465 470 475 480

Ile Glu Gly Tyr Glu Cys Gln Pro Ile Phe Leu Asn Val Leu Glu Ala
485 490 495

Ile Glu Pro Gly Val Val Cys Ala Gly His Asp Asn Asn Gln Pro Asp
500 505 510

Ser Phe Ala Ala Leu Leu Ser Ser Leu Asn Glu Leu Gly Glu Arg Gln
515 520 525

Leu Val His Val Val Lys Trp Ala Lys Ala Leu Pro Gly Phe Arg Asn
530 535 540

Leu His Val Asp Asp Gln Met Ala Val Ile Gln Tyr Ser Trp Met Gly
545 550 555 560

Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn Val Asn Ser
565 570 575

Arg Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn Glu Tyr Arg Met
580 585 590

His Lys Ser Arg Met Tyr Ser Gln Cys Val Arg Met Arg His Leu Ser
595 600 605

Gln Glu Phe Gly Trp Leu Gln Ile Thr Pro Gln Glu Phe Leu Cys Met
610 615 620

Lys Ala Leu Leu Leu Phe Ser Ile Ile Pro Val Asp Gly Leu Lys Asn
625 630 635 640

Gln Lys Phe Phe Asp Glu Leu Arg Met Asn Tyr Ile Lys Glu Leu Asp
645 650 655

Arg Ile Ile Ala Cys Lys Arg Lys Asn Pro Thr Ser Cys Ser Arg Arg
660 665 670

Phe Tyr Gln Leu Thr Lys Leu Leu Asp Ser Val Gln Pro Ile Ala Arg
675 680 685

Glu Leu His Gln Phe Thr Phe Asp Leu Leu Ile Lys Ser His Met Val
690 695 700

103 PCT Sequence Listing.ST25

Ser Val Asp Phe Pro Glu Met Met Ala Glu Ile Ile Ser Val Gln Val
705 710 715 720

Pro Lys Ile Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln
725 730 735

<210> 46
<211> 1186
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(1186)
<223> pET15b-SD3-ELP1-180-throm-Androgen receptor ligand binding domain
<400> 46

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
Page 72

103 PCT Sequence Listing.ST25

145		150		155		160
Val	Pro	Gly	Val	Gly ₁₆₅	Val	Pro
				Gly ₁₇₀	Val	Pro
				Gly	Ala	Gly ₁₇₅
Pro	Gly	Val	Gly ₁₈₀	Val	Pro	Gly
				Gly ₁₈₅	Val	Pro
				Gly	Val	Pro
Gly	Gly	Gly ₁₉₅	Val	Pro	Gly	Ala
				Gly ₂₀₀	Val	Pro
				Gly	Gly ₂₀₅	Val
				Gly	Val	Pro
Val	Gly ₂₁₀	Val	Pro	Gly	Val	Gly ₂₁₅
				Gly	Val	Pro
				Gly	Gly ₂₂₀	Val
				Gly	Val	Pro
Gly	Val	Pro	Gly	Val	Gly ₂₃₀	Val
				Gly	Val	Pro
				Gly	Gly ₂₃₅	Val
				Gly	Val	Pro
Val	Pro	Gly	Gly	Gly ₂₄₅	Val	Pro
				Gly	Ala	Gly ₂₅₀
				Gly	Val	Pro
				Gly	Gly ₂₅₅	Val
Pro	Gly	Val	Gly ₂₆₀	Val	Pro	Gly
				Gly ₂₆₅	Val	Pro
				Gly	Gly ₂₇₀	Val
				Gly	Val	Pro
Gly	Ala	Gly ₂₇₅	Val	Pro	Gly	Val
				Gly ₂₈₀	Val	Pro
				Gly	Val	Pro
				Gly	Gly ₂₈₅	Val
				Gly	Val	Pro
Val	Gly ₂₉₀	Val	Pro	Gly	Gly	Gly ₂₉₅
				Gly	Val	Pro
				Gly	Ala	Gly ₃₀₀
				Gly	Val	Pro
Gly	Val	Pro	Gly	Val	Gly ₃₁₀	Val
				Gly	Val	Pro
				Gly	Gly ₃₁₅	Val
				Gly	Val	Pro
				Gly	Gly ₃₂₀	Val
Val	Pro	Gly	Ala	Gly ₃₂₅	Val	Pro
				Gly	Val	Pro
				Gly	Gly ₃₃₀	Val
				Gly	Val	Pro
				Gly	Gly ₃₃₅	Val
Pro	Gly	Val	Gly ₃₄₀	Val	Pro	Gly
				Gly	Gly ₃₄₅	Val
				Gly	Ala	Gly ₃₅₀
				Gly	Val	Pro
Gly	Gly	Gly ₃₅₅	Val	Pro	Gly	Val
				Gly ₃₆₀	Val	Pro
				Gly	Val	Pro
				Gly	Gly ₃₆₅	Val
				Gly	Val	Pro
Gly	Gly ₃₇₀	Val	Pro	Gly	Ala	Gly ₃₇₅
				Gly	Val	Pro
				Gly	Val	Pro
				Gly	Gly ₃₈₀	Val
				Gly	Val	Pro
Gly	Val	Pro	Gly	Val	Gly ₃₉₀	Val
				Gly	Val	Pro
				Gly	Gly ₃₉₅	Val
				Gly	Ala	Gly ₄₀₀

103 PCT Sequence Listing.ST25

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
450 455 460

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
465 470 475 480

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
485 490 495

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
500 505 510

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
515 520 525

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
530 535 540

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
545 550 555 560

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
565 570 575

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
580 585 590

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
595 600 605

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
610 615 620

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
625 630 635 640

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
645 650 655

103 PCT Sequence Listing.ST25

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
660 665 670

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
675 680 685

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
690 695 700

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
705 710 715 720

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
725 730 735

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
740 745 750

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
755 760 765

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
770 775 780

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
785 790 795 800

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
805 810 815

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
820 825 830

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
835 840 845

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
850 855 860

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
865 870 875 880

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
885 890 895

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser
900 905 910

103 PCT Sequence Listing.ST25

Gly Gly Gly Gly Gly Ser Ile Gly Pro Leu Val Pro Arg Gly Ser His
915 920 925

Met His Ile Glu Gly Tyr Glu Cys Gln Pro Ile Phe Leu Asn Val Leu
930 935 940

Glu Ala Ile Glu Pro Gly Val Val Cys Ala Gly His Asp Asn Asn Gln
945 950 955 960

Pro Asp Ser Phe Ala Ala Leu Leu Ser Ser Leu Asn Glu Leu Gly Glu
965 970 975

Arg Gln Leu Val His Val Val Lys Trp Ala Lys Ala Leu Pro Gly Phe
980 985 990

Arg Asn Leu His Val Asp Asp Gln Met Ala Val Ile Gln Tyr Ser Trp
995 1000 1005

Met Gly Leu Met Val Phe Ala Met Gly Trp Arg Ser Phe Thr Asn
1010 1015 1020

Val Asn Ser Arg Met Leu Tyr Phe Ala Pro Asp Leu Val Phe Asn
1025 1030 1035

Glu Tyr Arg Met His Lys Ser Arg Met Tyr Ser Gln Cys Val Arg
1040 1045 1050

Met Arg His Leu Ser Gln Glu Phe Gly Trp Leu Gln Ile Thr Pro
1055 1060 1065

Gln Glu Phe Leu Cys Met Lys Ala Leu Leu Leu Phe Ser Ile Ile
1070 1075 1080

Pro Val Asp Gly Leu Lys Asn Gln Lys Phe Phe Asp Glu Leu Arg
1085 1090 1095

Met Asn Tyr Ile Lys Glu Leu Asp Arg Ile Ile Ala Cys Lys Arg
1100 1105 1110

Lys Asn Pro Thr Ser Cys Ser Arg Arg Phe Tyr Gln Leu Thr Lys
1115 1120 1125

Leu Leu Asp Ser Val Gln Pro Ile Ala Arg Glu Leu His Gln Phe
1130 1135 1140

Thr Phe Asp Leu Leu Ile Lys Ser His Met Val Ser Val Asp Phe
Page 76

1145 103 PCT Sequence Listing.ST25
1150 1155

Pro Glu Met Met Ala Glu Ile Ile Ser Val Gln Val Pro Lys Ile
1160 1165 1170

Leu Ser Gly Lys Val Lys Pro Ile Tyr Phe His Thr Gln
1175 1180 1185

<210> 47
<211> 757
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(757)
<223> pET15b-SD3-ELP1-90-throm-Glucocorticoid receptor ligand binding domain

<400> 47

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

103 PCT Sequence Listing.ST25

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Gly
450 455 460

Gly Gly Gly Ser Ile Gly Pro Leu Val Pro Arg Gly Ser His Met Ile
465 470 475 480

Gln Gln Ala Thr Thr Gly Val Ser Gln Glu Thr Ser Glu Asn Pro Gly
485 490 495

Asp Lys Thr Ile Val Pro Ala Thr Leu Pro Gln Leu Thr Pro Thr Leu
500 505 510

Val Ser Leu Leu Glu Val Ile Glu Pro Glu Val Leu Tyr Ala Gly Tyr
515 520 525

Asp Ser Ser Val Pro Asp Ser Thr Trp Arg Ile Met Thr Thr Leu Asn
530 535 540

Met Leu Gly Gly Arg Gln Val Ile Ala Ala Val Lys Trp Ala Lys Ala
545 550 555 560

Ile Pro Gly Phe Arg Asn Leu His Leu Asp Asp Gln Met Thr Leu Leu
565 570 575

Gln Tyr Ser Trp Met Ser Leu Met Ala Phe Ala Leu Gly Trp Arg Ser
580 585 590

Tyr Arg Gln Ser Ser Ala Asn Leu Leu Cys Phe Ala Pro Asp Leu Ile
595 600 605

Ile Asn Glu Gln Arg Met Thr Leu Pro Asp Met Tyr Asp Gln Cys Lys
610 615 620

His Met Leu Tyr Val Ser Ser Glu Leu His Arg Leu Gln Val Ser Tyr
625 630 635 640

Glu Glu Tyr Leu Cys Met Lys Thr Leu Leu Leu Leu Ser Ser Val Pro

103 PCT Sequence Listing.ST25

645

650

655

Lys Asp Gly Leu Lys Ser Gln Glu Leu Phe Asp Glu Ile Arg Met Thr
660 665 670

Tyr Ile Lys Glu Leu Gly Lys Ala Ile Val Lys Arg Glu Gly Asn Ser
675 680 685

Ser Gln Asn Trp Gln Arg Phe Tyr Gln Leu Thr Lys Leu Leu Asp Ser
690 695 700

Met His Glu Val Val Glu Asn Leu Leu Asn Tyr Cys Phe Gln Thr Phe
705 710 715 720

Leu Asp Lys Thr Met Ser Ile Glu Phe Pro Glu Met Leu Ala Glu Ile
725 730 735

Ile Thr Asn Gln Ile Pro Lys Tyr Ser Asn Gly Asn Ile Lys Lys Leu
740 745 750

Leu Phe His Gln Lys
755

<210> 48

<211> 624

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> MISC_FEATURE

<222> (1)..(624)

<223> pET15b-SD3-ELP1-60-throm-Estrogen receptor ligand binding domain

<400> 48

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Trp Pro Ser Ser Gly Gly Gly Gly Gly Ser Ile Gly
305 310 315 320

103 PCT Sequence Listing.ST25

Pro Leu Val Pro Arg Gly Ser His Met Ser Lys Lys Asn Ser Leu Ala
325 330 335

Leu Ser Leu Thr Ala Asp Gln Met Val Ser Ala Leu Leu Asp Ala Glu
340 345 350

Pro Pro Ile Leu Tyr Ser Glu Tyr Asp Pro Thr Arg Pro Phe Ser Glu
355 360 365

Ala Ser Met Met Gly Leu Leu Thr Asn Leu Ala Asp Arg Glu Leu Val
370 375 380

His Met Ile Asn Trp Ala Lys Arg Val Pro Gly Phe Val Asp Leu Thr
385 390 395 400

Leu His Asp Gln Val His Leu Leu Glu Cys Ala Trp Leu Glu Ile Leu
405 410 415

Met Ile Gly Leu Val Trp Arg Ser Met Glu His Pro Gly Lys Leu Leu
420 425 430

Phe Ala Pro Asn Leu Leu Leu Asp Arg Asn Gln Gly Lys Cys Val Glu
435 440 445

Gly Met Val Glu Ile Phe Asp Met Leu Leu Ala Thr Ser Ser Arg Phe
450 455 460

Arg Met Met Asn Leu Gln Gly Glu Glu Phe Val Cys Leu Lys Ser Ile
465 470 475 480

Ile Leu Leu Asn Ser Gly Val Tyr Thr Phe Leu Ser Ser Thr Leu Lys
485 490 495

Ser Leu Glu Glu Lys Asp His Ile His Arg Val Leu Asp Lys Ile Thr
500 505 510

Asp Thr Leu Ile His Leu Met Ala Lys Ala Gly Leu Thr Leu Gln Gln
515 520 525

Gln His Gln Arg Leu Ala Gln Leu Leu Leu Ile Leu Ser His Ile Arg
530 535 540

His Met Ser Asn Lys Gly Met Glu His Leu Tyr Ser Met Lys Cys Lys
545 550 555 560

Asn Val Val Pro Leu Tyr Asp Leu Leu Leu Glu Met Leu Asp Ala His
565 570 575

103 PCT Sequence Listing.ST25

Arg Leu His Ala Pro Thr Ser Arg Gly Gly Ala Ser Val Glu Glu Thr
580 585 590

Asp Gln Ser His Leu Ala Thr Ala Gly Ser Thr Ser Ser His Ser Leu
595 600 605

Gln Lys Tyr Tyr Ile Thr Gly Glu Ala Glu Gly Phe Pro Ala Thr Val
610 615 620

<210> 49

<211> 774

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<220>

<221> MISC_FEATURE

<222> (1)..(774)

<223> pET15b-SD5-ELP1-90-throm-Estrogen receptor ligand binding domain

<400> 49

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
Page 83

103 PCT Sequence Listing.ST25

130

135

140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Gly
450 455 460

Gly Gly Gly Ser Ile Gly Pro Leu Val Pro Arg Gly Ser His Met Ser
465 470 475 480

Lys Lys Asn Ser Leu Ala Leu Ser Leu Thr Ala Asp Gln Met Val Ser
485 490 495

Ala Leu Leu Asp Ala Glu Pro Pro Ile Leu Tyr Ser Glu Tyr Asp Pro
500 505

Thr Arg Pro Phe Ser Glu Ala Ser Met Met Gly Leu Leu Thr Asn Leu
515 520 525

Ala Asp Arg Glu Leu Val His Met Ile Asn Trp Ala Lys Arg Val Pro
530 535 540

Gly Phe Val Asp Leu Thr Leu His Asp Gln Val His Leu Leu Glu Cys
545 550 555 560

Ala Trp Leu Glu Ile Leu Met Ile Gly Leu Val Trp Arg Ser Met Glu
565 570 575

His Pro Gly Lys Leu Leu Phe Ala Pro Asn Leu Leu Leu Asp Arg Asn
580 585 590

Gln Gly Lys Cys Val Glu Gly Met Val Glu Ile Phe Asp Met Leu Leu
595 600 605

Ala Thr Ser Ser Arg Phe Arg Met Met Asn Leu Gln Gly Glu Glu Phe
610 615 620

Val Cys Leu Lys Ser Ile Ile Leu Leu Asn Ser Gly Val Tyr Thr Phe
625 630 635 640

103 PCT Sequence Listing.ST25

Leu Ser Ser Thr Leu Lys Ser Leu Glu Glu Lys Asp His Ile His Arg
645 650 655

Val Leu Asp Lys Ile Thr Asp Thr Leu Ile His Leu Met Ala Lys Ala
660 665 670

Gly Leu Thr Leu Gln Gln Gln His Gln Arg Leu Ala Gln Leu Leu Leu
675 680 685

Ile Leu Ser His Ile Arg His Met Ser Asn Lys Gly Met Glu His Leu
690 695 700

Tyr Ser Met Lys Cys Lys Asn Val Val Pro Leu Tyr Asp Leu Leu Leu
705 710 715 720

Glu Met Leu Asp Ala His Arg Leu His Ala Pro Thr Ser Arg Gly Gly
725 730 735

Ala Ser Val Glu Glu Thr Asp Gln Ser His Leu Ala Thr Ala Gly Ser
740 745 750

Thr Ser Ser His Ser Leu Gln Lys Tyr Tyr Ile Thr Gly Glu Ala Glu
755 760 765

Gly Phe Pro Ala Thr Val
770

<210> 50
<211> 1225
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(1225)
<223> pET15b-SD5-ELP1-180-throm-Estrogen receptor ligand binding domain

<400> 50

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

103 PCT Sequence Listing.ST25

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly

103 PCT Sequence Listing.ST25

290

295

300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
450 455 460

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
465 470 475 480

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
485 490 495

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
500 505 510

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
515 520 525

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
530 535 540

103 PCT Sequence Listing.ST25

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
545 550 555 560

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
565 570 575

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
580 585 590

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
595 600 605

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
610 615 620

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
625 630 635 640

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
645 650 655

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
660 665 670

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
675 680 685

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
690 695 700

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
705 710 715 720

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
725 730 735

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
740 745 750

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
755 760 765

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
770 775 780

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
785 790 795 800

103 PCT Sequence Listing.ST25

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
805 810 815

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
820 825 830

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
835 840 845

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
850 855 860

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
865 870 875 880

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
885 890 895

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser
900 905 910

Gly Leu Val Pro Arg Gly Ser Pro Gly Ile Ser Gly Gly Gly Gly Gly
915 920 925

His Met Ser Lys Lys Asn Ser Leu Ala Leu Ser Leu Thr Ala Asp Gln
930 935 940

Met Val Ser Ala Leu Leu Asp Ala Glu Pro Pro Ile Leu Tyr Ser Glu
945 950 955 960

Tyr Asp Pro Thr Arg Pro Phe Ser Glu Ala Ser Met Met Gly Leu Leu
965 970 975

Thr Asn Leu Ala Asp Arg Glu Leu Val His Met Ile Asn Trp Ala Lys
980 985 990

Arg Val Pro Gly Phe Val Asp Leu Thr Leu His Asp Gln Val His Leu
995 1000 1005

Leu Glu Cys Ala Trp Leu Glu Ile Leu Met Ile Gly Leu Val Trp
1010 1015 1020

Arg Ser Met Glu His Pro Gly Lys Leu Leu Phe Ala Pro Asn Leu
1025 1030 1035

Leu Leu Asp Arg Asn Gln Gly Lys Cys Val Glu Gly Met Val Glu
1040 1045 1050

103 PCT Sequence Listing.ST25

Ile Phe Asp Met Leu Leu Ala Thr Ser Ser Arg Phe Arg Met Met
1055 1060 1065

Asn Leu Gln Gly Glu Glu Phe Val Cys Leu Lys Ser Ile Ile Leu
1070 1075 1080

Leu Asn Ser Gly Val Tyr Thr Phe Leu Ser Ser Thr Leu Lys Ser
1085 1090 1095

Leu Glu Glu Lys Asp His Ile His Arg Val Leu Asp Lys Ile Thr
1100 1105 1110

Asp Thr Leu Ile His Leu Met Ala Lys Ala Gly Leu Thr Leu Gln
1115 1120 1125

Gln Gln His Gln Arg Leu Ala Gln Leu Leu Leu Ile Leu Ser His
1130 1135 1140

Ile Arg His Met Ser Asn Lys Gly Met Glu His Leu Tyr Ser Met
1145 1150 1155

Lys Cys Lys Asn Val Val Pro Leu Tyr Asp Leu Leu Leu Glu Met
1160 1165 1170

Leu Asp Ala His Arg Leu His Ala Pro Thr Ser Arg Gly Gly Ala
1175 1180 1185

Ser Val Glu Glu Thr Asp Gln Ser His Leu Ala Thr Ala Gly Ser
1190 1195 1200

Thr Ser Ser His Ser Leu Gln Lys Tyr Tyr Ile Thr Gly Glu Ala
1205 1210 1215

Glu Gly Phe Pro Ala Thr Val
1220 1225

<210> 51
<211> 775
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(775)
<223> pET15b-SD6-ELP1-90-TEV-Estrogen receptor ligand binding domain

103 PCT Sequence Listing.ST25

<400> 51

```

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
 1      5      10     15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
 20     25     30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
 35     40     45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
 50     55     60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
 65     70     75     80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
 85     90     95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100    105    110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115    120    125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130    135    140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145    150    155    160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165    170    175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180    185    190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195    200    205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210    215    220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225    230    235    240

```

103 PCT Sequence Listing.ST25

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Asp
450 455 460

Tyr Asp Ile Pro Thr Thr Glu Asn Leu Tyr Phe Gln Gly Ala His Met
465 470 475 480

Ser Lys Lys Asn Ser Leu Ala Leu Ser Leu Thr Ala Asp Gln Met Val
485 490 495

103 PCT Sequence Listing.ST25

Ser Ala Leu Leu Asp Ala Glu Pro Pro Ile Leu Tyr Ser Glu Tyr Asp
500 505 510

Pro Thr Arg Pro Phe Ser Glu Ala Ser Met Met Gly Leu Leu Thr Asn
515 520 525

Leu Ala Asp Arg Glu Leu Val His Met Ile Asn Trp Ala Lys Arg Val
530 535 540

Pro Gly Phe Val Asp Leu Thr Leu His Asp Gln Val His Leu Leu Glu
545 550 555 560

Cys Ala Trp Leu Glu Ile Leu Met Ile Gly Leu Val Trp Arg Ser Met
565 570 575

Glu His Pro Gly Lys Leu Leu Phe Ala Pro Asn Leu Leu Leu Asp Arg
580 585 590

Asn Gln Gly Lys Cys Val Glu Gly Met Val Glu Ile Phe Asp Met Leu
595 600 605

Leu Ala Thr Ser Ser Arg Phe Arg Met Met Asn Leu Gln Gly Glu Glu
610 615 620

Phe Val Cys Leu Lys Ser Ile Ile Leu Leu Asn Ser Gly Val Tyr Thr
625 630 635 640

Phe Leu Ser Ser Thr Leu Lys Ser Leu Glu Glu Lys Asp His Ile His
645 650 655

Arg Val Leu Asp Lys Ile Thr Asp Thr Leu Ile His Leu Met Ala Lys
660 665 670

Ala Gly Leu Thr Leu Gln Gln Gln His Gln Arg Leu Ala Gln Leu Leu
675 680 685

Leu Ile Leu Ser His Ile Arg His Met Ser Asn Lys Gly Met Glu His
690 695 700

Leu Tyr Ser Met Lys Cys Lys Asn Val Val Pro Leu Tyr Asp Leu Leu
705 710 715 720

Leu Glu Met Leu Asp Ala His Arg Leu His Ala Pro Thr Ser Arg Gly
725 730 735

Gly Ala Ser Val Glu Glu Thr Asp Gln Ser His Leu Ala Thr Ala Gly
740 745 750

103 PCT Sequence Listing.ST25

Ser Thr Ser Ser His Ser Leu Gln Lys Tyr Tyr Ile Thr Gly Glu Ala
755 760 765

Glu Gly Phe Pro Ala Thr Val
770 775

<210> 52
<211> 859
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(859)
<223> pET15b-SD1-ELP1-90-throm-G protein alpha Q

<400> 52

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
Page 95

103 PCT Sequence Listing.ST25

145		150		155		160
Val	Pro	Gly	Val	Gly ₁₆₅	Val	Pro
				Gly ₁₇₀	Val	Pro
				Gly	Ala	Gly ₁₇₅
				Val		
Pro	Gly	Val	Gly ₁₈₀	Val	Pro	Gly
				Val	Gly ₁₈₅	Val
				Pro	Gly	Val
				Gly ₁₉₀	Val	Pro
Gly	Gly	Gly ₁₉₅	Val	Pro	Gly	Ala
				Gly ₂₀₀	Val	Pro
				Gly	Gly	Gly ₂₀₅
				Val	Pro	Gly
Val	Gly ₂₁₀	Val	Pro	Gly	Val	Gly ₂₁₅
				Val	Pro	Gly
				Gly	Gly	Gly ₂₂₀
				Val	Pro	Gly
				Ala		
Gly	Val	Pro	Gly	Val	Gly ₂₃₀	Val
				Pro	Gly	Val
				Gly ₂₃₅	Val	Pro
				Gly	Val	Gly ₂₄₀
Val	Pro	Gly	Gly	Gly ₂₄₅	Val	Pro
				Gly	Ala	Gly ₂₅₀
				Val	Pro	Gly
				Gly	Gly	Gly ₂₅₅
				Val		
Pro	Gly	Val	Gly ₂₆₀	Val	Pro	Gly
				Val	Gly ₂₆₅	Val
				Pro	Gly	Gly
				Gly ₂₇₀	Val	Pro
Gly	Ala	Gly ₂₇₅	Val	Pro	Gly	Val
				Gly ₂₈₀	Val	Pro
				Gly	Val	Gly ₂₈₅
Val	Gly ₂₉₀	Val	Pro	Gly	Gly	Gly ₂₉₅
				Val	Pro	Gly
				Ala	Gly ₃₀₀	Val
				Pro	Gly	Gly
Gly	Val	Pro	Gly	Val	Gly ₃₁₀	Val
				Pro	Gly	Val
				Gly ₃₁₅	Val	Pro
				Gly	Gly	Gly ₃₂₀
Val	Pro	Gly	Ala	Gly ₃₂₅	Val	Pro
				Gly	Val	Gly ₃₃₀
				Val	Pro	Gly
				Gly ₃₃₅	Val	
Pro	Gly	Val	Gly ₃₄₀	Val	Pro	Gly
				Gly	Gly	Gly ₃₄₅
				Val	Pro	Gly
				Ala	Gly ₃₅₀	Val
				Pro		
Gly	Gly	Gly ₃₅₅	Val	Pro	Gly	Val
				Gly ₃₆₀	Val	Pro
				Gly	Val	Gly ₃₆₅
Gly	Gly ₃₇₀	Val	Pro	Gly	Ala	Gly ₃₇₅
				Val	Pro	Gly
				Val	Gly ₃₈₀	Val
				Pro	Gly	Val
Gly	Val	Pro	Gly	Val	Gly ₃₉₀	Val
				Pro	Gly	Gly
				Gly ₃₉₅	Val	Pro
				Gly	Ala	Gly ₄₀₀

103 PCT Sequence Listing.ST25

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Gly
450 455 460

Gly Ser Ile Gly Pro Leu Val Pro Arg Gly Ser His Ser Met Gly Leu
465 470 475 480

Asn Asp Ile Phe Glu Ala Gln Lys Ile Glu Trp His Glu His Met Pro
485 490 495

Met Ala Leu Glu Met Thr Leu Glu Ser Ile Met Ala Cys Cys Leu Ser
500 505 510

Glu Glu Ala Lys Glu Ala Arg Arg Ile Asn Asp Glu Ile Glu Arg Gln
515 520 525

Leu Arg Arg Asp Lys Arg Asp Ala Arg Arg Glu Leu Lys Leu Leu Leu
530 535 540

Leu Gly Thr Gly Glu Ser Gly Lys Ser Thr Phe Ile Lys Gln Met Arg
545 550 555 560

Ile Ile His Gly Ser Gly Tyr Ser Asp Glu Asp Lys Arg Gly Phe Thr
565 570 575

Lys Leu Val Tyr Gln Asn Ile Phe Thr Ala Met Gln Ala Met Ile Arg
580 585 590

Ala Met Asp Thr Leu Lys Ile Pro Tyr Lys Tyr Glu His Asn Lys Ala
595 600 605

His Ala Gln Leu Val Arg Glu Val Asp Val Glu Lys Val Ser Ala Phe
610 615 620

Glu Asn Pro Tyr Val Asp Ala Ile Lys Ser Leu Trp Asn Asp Pro Gly
625 630 635 640

Ile Gln Glu Cys Tyr Asp Arg Arg Arg Glu Tyr Gln Leu Ser Asp Ser
645 650 655

103 PCT Sequence Listing.ST25

Thr Lys Tyr Tyr Leu Asn Asp Leu Asp Arg Val Ala Asp Pro Ala Tyr
660 665 670

Leu Pro Thr Gln Gln Asp Val Leu Arg Val Arg Val Pro Thr Thr Gly
675 680 685

Ile Ile Glu Tyr Pro Phe Asp Leu Gln Ser Val Ile Phe Arg Met Val
690 695 700

Asp Val Gly Gly Gln Arg Ser Glu Arg Arg Lys Trp Ile His Cys Phe
705 710 715 720

Glu Asn Val Thr Ser Ile Met Phe Leu Val Ala Leu Ser Glu Tyr Asp
725 730 735

Gln Val Leu Val Glu Ser Asp Asn Glu Asn Arg Met Glu Glu Ser Lys
740 745 750

Ala Leu Phe Arg Thr Ile Ile Thr Tyr Pro Trp Phe Gln Asn Ser Ser
755 760 765

Val Ile Leu Phe Leu Asn Lys Lys Asp Leu Leu Glu Glu Lys Ile Met
770 775 780

Tyr Ser His Leu Val Asp Tyr Phe Pro Glu Tyr Asp Gly Pro Gln Arg
785 790 795 800

Asp Ala Gln Ala Ala Arg Glu Phe Ile Leu Lys Met Phe Val Asp Leu
805 810 815

Asn Pro Asp Ser Asp Lys Ile Asn Tyr Ser His Phe Thr Cys Ala Thr
820 825 830

Asp Thr Glu Asn Ile Arg Phe Val Phe Ala Ala Val Lys Asp Thr Ile
835 840 845

Leu Gln Leu Asn Leu Lys Glu Tyr Asn Leu Val
850 855

<210> 53
<211> 1309
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>

103 PCT Sequence Listing.ST25

<221> MISC_FEATURE

<222> (1)..(1309)

<223> pET15b-SD1-ELP1-180-throm-G protein alpha Q

<400> 53

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
Page 99

225 230 235 240

Page 100

103 PCT Sequence Listing.ST25

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
485 490 495

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
500 505 510

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
515 520 525

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
530 535 540

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
545 550 555 560

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
565 570 575

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
580 585 590

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
595 600 605

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
610 615 620

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
625 630 635 640

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
645 650 655

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
660 665 670

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
675 680 685

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
690 695 700

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
705 710 715 720

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
725 730 735

103 PCT Sequence Listing.ST25

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
740 745 750

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
755 760 765

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
770 775 780

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
785 790 795 800

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
805 810 815

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
820 825 830

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
835 840 845

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
850 855 860

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
865 870 875 880

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
885 890 895

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser
900 905 910

Gly Gly Gly Ser Ile Gly Pro Leu Val Pro Arg Gly Ser His Ser Met
915 920 925

Gly Leu Asn Asp Ile Phe Glu Ala Gln Lys Ile Glu Trp His Glu His
930 935 940

Met Pro Met Ala Leu Glu Met Thr Leu Glu Ser Ile Met Ala Cys Cys
945 950 955 960

Leu Ser Glu Glu Ala Lys Glu Ala Arg Arg Ile Asn Asp Glu Ile Glu
965 970 975

Arg Gln Leu Arg Arg Asp Lys Arg Asp Ala Arg Arg Glu Leu Lys Leu
980 985 990

103 PCT Sequence Listing.ST25

Leu Leu Leu Gly Thr Gly Glu Ser Gly Lys Ser Thr Phe Ile Lys Gln
995 1000 1005

Met Arg Ile Ile His Gly Ser Gly Tyr Ser Asp Glu Asp Lys Arg
1010 1015 1020

Gly Phe Thr Lys Leu Val Tyr Gln Asn Ile Phe Thr Ala Met Gln
1025 1030 1035

Ala Met Ile Arg Ala Met Asp Thr Leu Lys Ile Pro Tyr Lys Tyr
1040 1045 1050

Glu His Asn Lys Ala His Ala Gln Leu Val Arg Glu Val Asp Val
1055 1060 1065

Glu Lys Val Ser Ala Phe Glu Asn Pro Tyr Val Asp Ala Ile Lys
1070 1075 1080

Ser Leu Trp Asn Asp Pro Gly Ile Gln Glu Cys Tyr Asp Arg Arg
1085 1090 1095

Arg Glu Tyr Gln Leu Ser Asp Ser Thr Lys Tyr Tyr Leu Asn Asp
1100 1105 1110

Leu Asp Arg Val Ala Asp Pro Ala Tyr Leu Pro Thr Gln Gln Asp
1115 1120 1125

Val Leu Arg Val Arg Val Pro Thr Thr Gly Ile Ile Glu Tyr Pro
1130 1135 1140

Phe Asp Leu Gln Ser Val Ile Phe Arg Met Val Asp Val Gly Gly
1145 1150 1155

Gln Arg Ser Glu Arg Arg Lys Trp Ile His Cys Phe Glu Asn Val
1160 1165 1170

Thr Ser Ile Met Phe Leu Val Ala Leu Ser Glu Tyr Asp Gln Val
1175 1180 1185

Leu Val Glu Ser Asp Asn Glu Asn Arg Met Glu Glu Ser Lys Ala
1190 1195 1200

Leu Phe Arg Thr Ile Ile Thr Tyr Pro Trp Phe Gln Asn Ser Ser
1205 1210 1215

Val Ile Leu Phe Leu Asn Lys Lys Asp Leu Leu Glu Glu Lys Ile
Page 103

103 PCT Sequence Listing.ST25

1220

1225

1230

Met Tyr Ser His Leu Val Asp Tyr Phe Pro Glu Tyr Asp Gly Pro
1235 1240 1245

Gln Arg Asp Ala Gln Ala Ala Arg Glu Phe Ile Leu Lys Met Phe
1250 1255 1260

Val Asp Leu Asn Pro Asp Ser Asp Lys Ile Asn Tyr Ser His Phe
1265 1270 1275

Thr Cys Ala Thr Asp Thr Glu Asn Ile Arg Phe Val Phe Ala Ala
1280 1285 1290

Val Lys Asp Thr Ile Leu Gln Leu Asn Leu Lys Glu Tyr Asn Leu
1295 1300 1305

Val

<210> 54
<211> 728
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(728)
<223> pET15b-SD3-ELP1-60-throm-1-Deoxy-D-Xylulose 5-Phosphate
Reductoisomerase Peptide

<400> 54

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

103 PCT Sequence Listing.ST25

Val Pro Gly Val Gly₈₅ Val Pro Gly Val Gly₉₀ Val Pro Gly Gly₉₅ Val
Pro Gly Ala Gly₁₀₀ Val Pro Gly Gly₁₀₅ Val Pro Gly Val Gly₁₁₀ Val Pro
Gly Val Gly₁₁₅ Val Pro Gly Gly₁₂₀ Val Pro Gly Ala Gly₁₂₅ Val Pro Gly
Val Gly₁₃₀ Val Pro Gly Val Gly₁₃₅ Val Pro Gly Val Gly₁₄₀ Val Pro Gly Gly
Gly₁₄₅ Val Pro Gly Ala Gly₁₅₀ Val Pro Gly Gly₁₅₅ Val Pro Gly Val Gly₁₆₀
Val Pro Gly Val Gly₁₆₅ Val Pro Gly Gly₁₇₀ Val Pro Gly Ala Gly₁₇₅ Val
Pro Gly Val Gly₁₈₀ Val Pro Gly Val Gly₁₈₅ Val Pro Gly Val Gly₁₉₀ Val Pro
Gly Gly Gly₁₉₅ Val Pro Gly Ala Gly₂₀₀ Val Pro Gly Gly₂₀₅ Val Pro Gly
Val Gly₂₁₀ Val Pro Gly Val Gly₂₁₅ Val Pro Gly Gly₂₂₀ Val Pro Gly Ala
Gly₂₂₅ Val Pro Gly Val Gly₂₃₀ Val Pro Gly Val Gly₂₃₅ Val Pro Gly Val Gly₂₄₀
Val Pro Gly Gly₂₄₅ Val Pro Gly Ala Gly₂₅₀ Val Pro Gly Gly₂₅₅ Val
Pro Gly Val Gly₂₆₀ Val Pro Gly Val Gly₂₆₅ Val Pro Gly Gly₂₇₀ Val Pro
Gly Ala Gly₂₇₅ Val Pro Gly Val Gly₂₈₀ Val Pro Gly Val Gly₂₈₅ Val Pro Gly
Val Gly₂₉₀ Val Pro Gly Gly₂₉₅ Val Pro Gly Ala Gly₃₀₀ Val Pro Gly Gly
Gly₃₀₅ Val Pro Gly Trp₃₁₀ Ser Ser Gly Gly₃₁₅ Gly Gly Ser Ile Gly₃₂₀
Pro Leu Val Pro Arg₃₂₅ Gly Ser His Met₃₃₀ Lys Gln Leu Thr Ile₃₃₅ Leu Gly

103 PCT Sequence Listing.ST25

Ser Thr Gly Ser Ile Gly Cys Ser Thr Leu Asp Val Val Arg His Asn
340 345 350

Pro Glu His Phe Arg Val Val Ala Leu Val Ala Gly Lys Asn Val Thr
355 360 365

Arg Met Val Glu Gln Cys Leu Glu Phe Ser Pro Arg Tyr Ala Val Met
370 375 380

Asp Asp Glu Ala Ser Ala Lys Leu Leu Lys Thr Met Leu Gln Gln Gln
385 390 395 400

Gly Ser Arg Thr Glu Val Leu Ser Gly Gln Gln Ala Ala Cys Asp Met
405 410 415

Ala Ala Leu Glu Asp Val Asp Gln Val Met Ala Ala Ile Val Gly Ala
420 425 430

Ala Gly Leu Leu Pro Thr Leu Ala Ala Ile Arg Ala Gly Lys Thr Ile
435 440 445

Leu Leu Ala Asn Lys Glu Ser Leu Val Thr Cys Gly Arg Leu Phe Met
450 455 460

Asp Ala Val Lys Gln Ser Lys Ala Gln Leu Leu Pro Val Asp Ser Glu
465 470 475 480

His Asn Ala Ile Phe Gln Ser Leu Pro Gln Pro Ile Gln His Asn Leu
485 490 495

Gly Tyr Ala Asp Leu Glu Gln Asn Gly Val Val Ser Ile Leu Leu Thr
500 505 510

Gly Ser Gly Gly Pro Phe Arg Glu Thr Pro Leu Arg Asp Leu Ala Thr
515 520 525

Met Thr Pro Asp Gln Ala Cys Arg His Pro Asn Trp Ser Met Gly Arg
530 535 540

Lys Ile Ser Val Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Tyr
545 550 555 560

Ile Glu Ala Arg Trp Leu Phe Asn Ala Ser Ala Ser Gln Met Glu Val
565 570 575

Leu Ile His Pro Gln Ser Val Ile His Ser Met Val Arg Tyr Gln Asp

580 103 PCT Sequence Listing.ST25
585 590

Gly Ser Val Leu Ala Gln Leu Gly Glu Pro Asp Met Arg Thr Pro Ile
595 600 605

Ala His Thr Met Ala Trp Pro Asn Arg Val Asn Ser Gly Val Lys Pro
610 615 620

Leu Asp Phe Cys Lys Leu Ser Ala Leu Thr Phe Ala Ala Pro Asp Tyr
625 630 635 640

Asp Arg Tyr Pro Cys Leu Lys Leu Ala Met Glu Ala Phe Glu Gln Gly
645 650 655

Gln Ala Ala Thr Thr Ala Leu Asn Ala Ala Asn Glu Ile Thr Val Ala
660 665 670

Ala Phe Leu Ala Gln Gln Ile Arg Phe Thr Asp Ile Ala Ala Leu Asn
675 680 685

Leu Ser Val Leu Glu Lys Met Asp Met Arg Glu Pro Gln Cys Val Asp
690 695 700

Asp Val Leu Ser Val Asp Ala Ser Ala Arg Glu Val Ala Arg Lys Glu
705 710 715 720

Val Met Arg Leu Ala Ser Pro Val
725

<210> 55
<211> 879
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(879)
<223> pET15b-SD5-ELP1-90-throm-1-Deoxy-D-Xylulose 5-Phosphate
Reductoisomerase Peptide

<400> 55

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

103 PCT Sequence Listing.ST25

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

103 PCT Sequence Listing.ST25

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Leu
450 455 460

Val Pro Arg Gly Ser Pro Gly Ile Ser Gly Gly Gly Gly Gly His Met
465 470 475 480

Lys Gln Leu Thr Ile Leu Gly Ser Thr Gly Ser Ile Gly Cys Ser Thr
485 490 495

Leu Asp Val Val Arg His Asn Pro Glu His Phe Arg Val Val Ala Leu
500 505 510

Val Ala Gly Lys Asn Val Thr Arg Met Val Glu Gln Cys Leu Glu Phe
515 520 525

Ser Pro Arg Tyr Ala Val Met Asp Asp Glu Ala Ser Ala Lys Leu Leu

103 PCT Sequence Listing.ST25

530

535

540

Lys Thr Met Leu Gln Gln Gln Gly Ser Arg Thr Glu Val Leu Ser Gly
545 550 555 560

Gln Gln Ala Ala Cys Asp Met Ala Ala Leu Glu Asp Val Asp Gln Val
565 570 575

Met Ala Ala Ile Val Gly Ala Ala Gly Leu Leu Pro Thr Leu Ala Ala
580 585 590

Ile Arg Ala Gly Lys Thr Ile Leu Leu Ala Asn Lys Glu Ser Leu Val
595 600 605

Thr Cys Gly Arg Leu Phe Met Asp Ala Val Lys Gln Ser Lys Ala Gln
610 615 620

Leu Leu Pro Val Asp Ser Glu His Asn Ala Ile Phe Gln Ser Leu Pro
625 630 635 640

Gln Pro Ile Gln His Asn Leu Gly Tyr Ala Asp Leu Glu Gln Asn Gly
645 650 655

Val Val Ser Ile Leu Leu Thr Gly Ser Gly Gly Pro Phe Arg Glu Thr
660 665 670

Pro Leu Arg Asp Leu Ala Thr Met Thr Pro Asp Gln Ala Cys Arg His
675 680 685

Pro Asn Trp Ser Met Gly Arg Lys Ile Ser Val Asp Ser Ala Thr Met
690 695 700

Met Asn Lys Gly Leu Glu Tyr Ile Glu Ala Arg Trp Leu Phe Asn Ala
705 710 715 720

Ser Ala Ser Gln Met Glu Val Leu Ile His Pro Gln Ser Val Ile His
725 730 735

Ser Met Val Arg Tyr Gln Asp Gly Ser Val Leu Ala Gln Leu Gly Glu
740 745 750

Pro Asp Met Arg Thr Pro Ile Ala His Thr Met Ala Trp Pro Asn Arg
755 760 765

Val Asn Ser Gly Val Lys Pro Leu Asp Phe Cys Lys Leu Ser Ala Leu
770 775 780

103 PCT Sequence Listing.ST25

Thr Phe Ala Ala Pro Asp Tyr Asp Arg Tyr Pro Cys Leu Lys Leu Ala
785 790 795 800

Met Glu Ala Phe Glu Gln Gly Gln Ala Ala Thr Thr Ala Leu Asn Ala
805 810 815

Ala Asn Glu Ile Thr Val Ala Ala Phe Leu Ala Gln Gln Ile Arg Phe
820 825 830

Thr Asp Ile Ala Ala Leu Asn Leu Ser Val Leu Glu Lys Met Asp Met
835 840 845

Arg Glu Pro Gln Cys Val Asp Asp Val Leu Ser Val Asp Ala Ser Ala
850 855 860

Arg Glu Val Ala Arg Lys Glu Val Met Arg Leu Ala Ser Pro Val
865 870 875

<210> 56
<211> 1329
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(1329)
<223> pET15b-SD5-ELP1-180-throm-1-Deoxy-D-Xylulose 5-Phosphate
Reductoisomerase Peptide

<400> 56

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
85 90 95

103 PCT Sequence Listing.ST25

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
165 170 175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
Page 112

103 PCT Sequence Listing.ST25
345 350

340

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
450 455 460

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
465 470 475 480

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
485 490 495

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
500 505 510

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
515 520 525

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
530 535 540

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
545 550 555 560

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
565 570 575

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
580 585 590

103 PCT Sequence Listing.ST25

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
595 600 605

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
610 615 620

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
625 630 635 640

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
645 650 655

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
660 665 670

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
675 680 685

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
690 695 700

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
705 710 715 720

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
725 730 735

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
740 745 750

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
755 760 765

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
770 775 780

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
785 790 795 800

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
805 810 815

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
820 825 830

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
835 840 845

103 PCT Sequence Listing.ST25

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
850 855 860

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
865 870 875 880

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
885 890 895

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser
900 905 910

Gly Leu Val Pro Arg Gly Ser Pro Gly Ile Ser Gly Gly Gly Gly Gly
915 920 925

His Met Lys Gln Leu Thr Ile Leu Gly Ser Thr Gly Ser Ile Gly Cys
930 935 940

Ser Thr Leu Asp Val Val Arg His Asn Pro Glu His Phe Arg Val Val
945 950 955 960

Ala Leu Val Ala Gly Lys Asn Val Thr Arg Met Val Glu Gln Cys Leu
965 970 975

Glu Phe Ser Pro Arg Tyr Ala Val Met Asp Asp Glu Ala Ser Ala Lys
980 985 990

Leu Leu Lys Thr Met Leu Gln Gln Gln Gly Ser Arg Thr Glu Val Leu
995 1000 1005

Ser Gly Gln Gln Ala Ala Cys Asp Met Ala Ala Leu Glu Asp Val
1010 1015 1020

Asp Gln Val Met Ala Ala Ile Val Gly Ala Ala Gly Leu Leu Pro
1025 1030 1035

Thr Leu Ala Ala Ile Arg Ala Gly Lys Thr Ile Leu Leu Ala Asn
1040 1045 1050

Lys Glu Ser Leu Val Thr Cys Gly Arg Leu Phe Met Asp Ala Val
1055 1060 1065

Lys Gln Ser Lys Ala Gln Leu Leu Pro Val Asp Ser Glu His Asn
1070 1075 1080

Ala Ile Phe Gln Ser Leu Pro Gln Pro Ile Gln His Asn Leu Gly
1085 1090 1095

103 PCT Sequence Listing.ST25

Tyr	Ala	Asp	Leu	Glu	Gln	Asn	Gly	Val	Val	Ser	Ile	Leu	Leu	Thr
	1100					1105					1110			
Gly	Ser	Gly	Gly	Pro	Phe	Arg	Glu	Thr	Pro	Leu	Arg	Asp	Leu	Ala
	1115					1120					1125			
Thr	Met	Thr	Pro	Asp	Gln	Ala	Cys	Arg	His	Pro	Asn	Trp	Ser	Met
	1130					1135					1140			
Gly	Arg	Lys	Ile	Ser	Val	Asp	Ser	Ala	Thr	Met	Met	Asn	Lys	Gly
	1145					1150					1155			
Leu	Glu	Tyr	Ile	Glu	Ala	Arg	Trp	Leu	Phe	Asn	Ala	Ser	Ala	Ser
	1160					1165					1170			
Gln	Met	Glu	Val	Leu	Ile	His	Pro	Gln	Ser	Val	Ile	His	Ser	Met
	1175					1180					1185			
Val	Arg	Tyr	Gln	Asp	Gly	Ser	Val	Leu	Ala	Gln	Leu	Gly	Glu	Pro
	1190					1195					1200			
Asp	Met	Arg	Thr	Pro	Ile	Ala	His	Thr	Met	Ala	Trp	Pro	Asn	Arg
	1205					1210					1215			
Val	Asn	Ser	Gly	Val	Lys	Pro	Leu	Asp	Phe	Cys	Lys	Leu	Ser	Ala
	1220					1225					1230			
Leu	Thr	Phe	Ala	Ala	Pro	Asp	Tyr	Asp	Arg	Tyr	Pro	Cys	Leu	Lys
	1235					1240					1245			
Leu	Ala	Met	Glu	Ala	Phe	Glu	Gln	Gly	Gln	Ala	Ala	Thr	Thr	Ala
	1250					1255					1260			
Leu	Asn	Ala	Ala	Asn	Glu	Ile	Thr	Val	Ala	Ala	Phe	Leu	Ala	Gln
	1265					1270					1275			
Gln	Ile	Arg	Phe	Thr	Asp	Ile	Ala	Ala	Leu	Asn	Leu	Ser	Val	Leu
	1280					1285					1290			
Glu	Lys	Met	Asp	Met	Arg	Glu	Pro	Gln	Cys	Val	Asp	Asp	Val	Leu
	1295					1300					1305			
Ser	Val	Asp	Ala	Ser	Ala	Arg	Glu	Val	Ala	Arg	Lys	Glu	Val	Met
	1310					1315					1320			
Arg	Leu	Ala	Ser	Pro	Val									

1325

<210> 57
 <211> 879
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> MISC_FEATURE
 <222> (1)..(879)
 <223> pET15b-SD6-ELP1-90-TEV-1-Deoxy-D-Xylulose 5-Phosphate
 Reductoisomerase Peptide

<400> 57

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
 1 5 10 15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
 20 25 30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
 35 40 45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
 50 55 60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
 65 70 75 80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
 85 90 95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
 100 105 110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
 115 120 125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
 130 135 140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
 145 150 155 160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
 165 170 175

103 PCT Sequence Listing.ST25

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
180 185 190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
195 200 205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
210 215 220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
225 230 235 240

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

103 PCT Sequence Listing.ST25

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Asp
450 455 460

Tyr Asp Ile Pro Thr Thr Glu Asn Leu Tyr Phe Gln Gly Ala His Met
465 470 475 480

Lys Gln Leu Thr Ile Leu Gly Ser Thr Gly Ser Ile Gly Cys Ser Thr
485 490 495

Leu Asp Val Val Arg His Asn Pro Glu His Phe Arg Val Val Ala Leu
500 505 510

Val Ala Gly Lys Asn Val Thr Arg Met Val Glu Gln Cys Leu Glu Phe
515 520 525

Ser Pro Arg Tyr Ala Val Met Asp Asp Glu Ala Ser Ala Lys Leu Leu
530 535 540

Lys Thr Met Leu Gln Gln Gln Gly Ser Arg Thr Glu Val Leu Ser Gly
545 550 555 560

Gln Gln Ala Ala Cys Asp Met Ala Ala Leu Glu Asp Val Asp Gln Val
565 570 575

Met Ala Ala Ile Val Gly Ala Ala Gly Leu Leu Pro Thr Leu Ala Ala
580 585 590

Ile Arg Ala Gly Lys Thr Ile Leu Leu Ala Asn Lys Glu Ser Leu Val
595 600 605

Thr Cys Gly Arg Leu Phe Met Asp Ala Val Lys Gln Ser Lys Ala Gln
610 615 620

Leu Leu Pro Val Asp Ser Glu His Asn Ala Ile Phe Gln Ser Leu Pro
625 630 635 640

Gln Pro Ile Gln His Asn Leu Gly Tyr Ala Asp Leu Glu Gln Asn Gly
645 650 655

Val Val Ser Ile Leu Leu Thr Gly Ser Gly Gly Pro Phe Arg Glu Thr
660 665 670

Pro Leu Arg Asp Leu Ala Thr Met Thr Pro Asp Gln Ala Cys Arg His

103 PCT Sequence Listing.ST25

675

680

685

Pro Asn Trp Ser Met Gly Arg Lys Ile Ser Val Asp Ser Ala Thr Met
690 695 700

Met Asn Lys Gly Leu Glu Tyr Ile Glu Ala Arg Trp Leu Phe Asn Ala
705 710 715 720

Ser Ala Ser Gln Met Glu Val Leu Ile His Pro Gln Ser Val Ile His
725 730 735

Ser Met Val Arg Tyr Gln Asp Gly Ser Val Leu Ala Gln Leu Gly Glu
740 745 750

Pro Asp Met Arg Thr Pro Ile Ala His Thr Met Ala Trp Pro Asn Arg
755 760 765

Val Asn Ser Gly Val Lys Pro Leu Asp Phe Cys Lys Leu Ser Ala Leu
770 775 780

Thr Phe Ala Ala Pro Asp Tyr Asp Arg Tyr Pro Cys Leu Lys Leu Ala
785 790 795 800

Met Glu Ala Phe Glu Gln Gly Gln Ala Ala Thr Thr Ala Leu Asn Ala
805 810 815

Ala Asn Glu Ile Thr Val Ala Ala Phe Leu Ala Gln Gln Ile Arg Phe
820 825 830

Thr Asp Ile Ala Ala Leu Asn Leu Ser Val Leu Glu Lys Met Asp Met
835 840 845

Arg Glu Pro Gln Cys Val Asp Asp Val Leu Ser Val Asp Ala Ser Ala
850 855 860

Arg Glu Val Ala Arg Lys Glu Val Met Arg Leu Ala Ser Pro Val
865 870 875

<210> 58
<211> 864
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> MISC_FEATURE
<222> (1)..(864)

103 PCT Sequence Listing.ST25

<223> pET15b-SD3-ELP1-90-throm-G protein alpha S

<400> 58

```

Met Arg Ala Leu Met Gly Pro Gly Val Gly Val Pro Gly Val Gly Val
1      5      10     15

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
      20     25     30

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
      35     40     45

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val
      50     55     60

Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
      65     70     75     80

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val
      85     90     95

Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro
      100    105    110

Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly
      115    120    125

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly
      130    135    140

Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly Val Gly
      145    150    155    160

Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val
      165    170    175

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
      180    185    190

Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val Pro Gly
      195    200    205

Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala
      210    215    220

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
      225    230    235    240

```

103 PCT Sequence Listing.ST25

Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly Gly Val
245 250 255

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro
260 265 270

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
275 280 285

Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Gly
290 295 300

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly
305 310 315 320

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
325 330 335

Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro
340 345 350

Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
355 360 365

Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
370 375 380

Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly Ala Gly
385 390 395 400

Val Pro Gly Gly Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
405 410 415

Pro Gly Gly Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
420 425 430

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Gly Gly Val Pro Gly
435 440 445

Ala Gly Val Pro Gly Gly Gly Val Pro Gly Trp Pro Ser Ser Gly Gly
450 455 460

Gly Gly Gly Ser Ile Gly Pro Leu Val Pro Arg Gly Ser His Met Pro
465 470 475 480

Met Ala Leu Glu Met Gly Cys Leu Gly Asn Ser Lys Thr Glu Asp Gln
485 490 495

103 PCT Sequence Listing.ST25

Arg Asn Glu Glu Lys Ala Gln Arg Glu Ala Asn Lys Lys Ile Glu Lys
500 505 510

Gln Leu Gln Lys Asp Lys Gln Val Tyr Arg Ala Thr His Arg Leu Leu
515 520 525

Leu Leu Gly Ala Gly Glu Ser Gly Lys Ser Thr Ile Val Lys Gln Met
530 535 540

Arg Ile Leu His Val Asn Gly Phe Asn Gly Asp Ser Glu Lys Ala Thr
545 550 555 560

Lys Val Gln Asp Ile Lys Asn Asn Leu Lys Glu Ala Ile Glu Thr Ile
565 570 575

Val Ala Ala Met Ser Asn Leu Val Pro Pro Val Glu Leu Ala Asn Pro
580 585 590

Glu Asn Gln Phe Arg Val Asp Tyr Ile Leu Ser Val Met Asn Val Pro
595 600 605

Asp Phe Asp Phe Pro Pro Glu Phe Tyr Glu His Ala Lys Ala Leu Trp
610 615 620

Glu Asp Glu Gly Val Arg Ala Cys Tyr Glu Arg Ser Asn Glu Tyr Gln
625 630 635 640

Leu Ile Asp Cys Ala Gln Tyr Phe Leu Asp Lys Ile Asp Val Ile Lys
645 650 655

Gln Ala Asp Tyr Val Pro Ser Asp Gln Asp Leu Leu Arg Cys Arg Val
660 665 670

Leu Thr Ser Gly Ile Phe Glu Thr Lys Phe Gln Val Asp Lys Val Asn
675 680 685

Phe His Met Phe Asp Val Gly Gly Gln Arg Asp Glu Arg Arg Lys Trp
690 695 700

Ile Gln Cys Phe Asn Asp Val Thr Ala Ile Ile Phe Val Val Ala Ser
705 710 715 720

Ser Ser Tyr Asn Met Val Ile Arg Glu Asp Asn Gln Thr Asn Arg Leu
725 730 735

Gln Glu Ala Leu Asn Leu Phe Lys Ser Ile Trp Asn Asn Arg Trp Leu
Page 123

740

Arg Thr Ile Ser Val Ile Leu Phe Leu Asn Lys Gln Asp Leu Leu Ala
755 760 765

Glu Lys Val Leu Ala Gly Lys Ser Lys Ile Glu Asp Tyr Phe Pro Glu
770 775 780

Phe Ala Arg Tyr Thr Thr Pro Glu Asp Ala Thr Pro Glu Pro Gly Glu
785 790 795 800

Asp Pro Arg Val Thr Arg Ala Lys Tyr Phe Ile Arg Asp Glu Phe Leu
805 810 815

Arg Ile Ser Thr Ala Ser Gly Asp Gly Arg His Tyr Cys Tyr Pro His
820 825 830

Phe Thr Cys Ala Val Asp Thr Glu Asn Ile Arg Arg Val Phe Asn Asp
835 840 845

Cys Arg Asp Ile Ile Gln Arg Met His Leu Arg Gln Tyr Glu Leu Leu
850 855 860

<210> 59
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic construct

<400> 59

Val Glu Asn Leu Tyr Phe Gln Gly Gly Met Gly
1 5 10

<210> 60
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic construct

<400> 60

Val Pro Gly Trp Pro Ser Ser Gly Asp Tyr Asp Ile Pro Thr Thr Glu
1 5 10 15

Asn Leu Tyr Phe Gln Gly Ala His
20

103 PCT Sequence Listing.ST25

<210> 61
 <211> 24
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic construct

<400> 61

Gly Ser Gly Ser Gly His Met His His His His His His Ser Ser Gly
 1 5 10 15

Leu Val Pro Arg Gly Ser Gly Lys
 20

<210> 62
 <211> 24
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic construct

<400> 62

Val Pro Gly Trp Pro Ser Ser Gly Asp Tyr Asp Ile Pro Thr Thr Glu
 1 5 10 15

Asn Leu Tyr Phe Gln Gly Ala His
 20

<210> 63
 <211> 27
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic construct

<400> 63

Val Asp Lys Leu Ala Ala Ala Leu Asp Met His His His His His His
 1 5 10 15

Ser Ser Gly Leu Val Pro Arg Gly Ser Gly Lys
 20 25

<210> 64
 <211> 24
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic construct

<400> 64

103 PCT Sequence Listing.ST25

Val Pro Gly Trp Pro Ser Ser Gly Asp Tyr Asp Ile Pro Thr Thr Glu
1 5 10 15

Asn Leu Tyr Phe Gln Gly Ala His
20

<210> 65
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic construct

<400> 65

Leu Glu Asn Leu Tyr Phe Gln Gly Gly Met Gly
1 5 10

<210> 66
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic construct

<400> 66

Val Pro Gly Trp Pro Ser Ser Gly Asp Tyr Asp Ile Pro Thr Thr Glu
1 5 10 15

Asn Leu Tyr Phe Gln Gly Ala His
20