



- (51) International Patent Classification:
C07K 7/06 (2006.01)
- (21) International Application Number:
PCT/US2012/033293
- (22) International Filing Date:
12 April 2012 (12.04.2012)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/475,042 13 April 2011 (13.04.2011) US
61/616,341 27 March 2012 (27.03.2012) US
- (71) Applicant (for all designated States except US): **THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK** [US/US]; 412 Low Memorial Library, 535 West 116th Street, New York, NY 10027 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **BANTA, Scott** [US/US]; 252 Old Oaks Road, Fairfield, CT 06825 (US).
- (74) Agents: **LOVE, Janes, M.** et al.; Wilmer Cutler Pickering Hale and Dorr LLP, 7 World Trade Center, 250 Greenwich Street, New York, NY 10007 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,

[Continued on next page]

(54) Title: PRECIPITABLE PEPTIDES

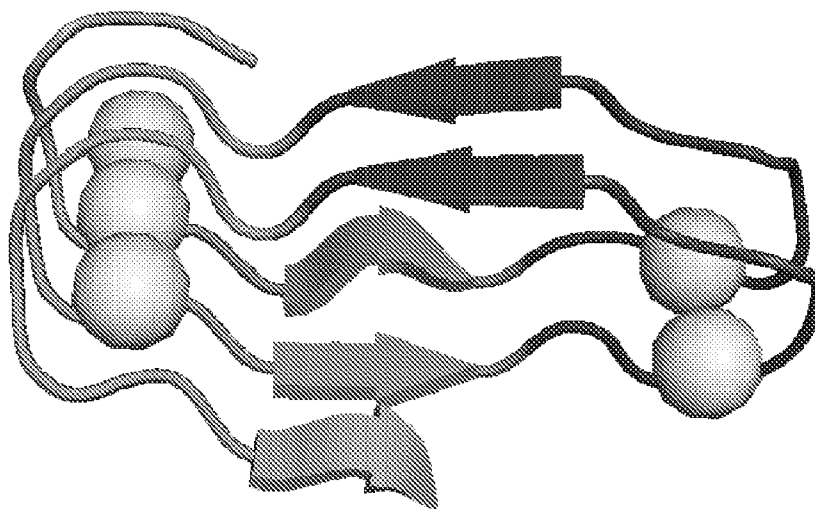


Figure 5

(57) Abstract: Ca²⁺ precipitable polypeptide tags and cassettes useful for purification of molecules from heterogeneous samples are disclosed. Methods for bioseparation of molecules comprising Ca²⁺ precipitable tags and cassettes are also disclosed.



LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

— *before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))*

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

PRECIPITABLE PEPTIDES

[0001] This invention was made with government support under grant number W9132T-08-2-0012 awarded by the DTRA and under grant number W9132T-08-2-0002 awarded by the US Army. The government has certain rights in the invention.

[0002] This application claims the benefit of and priority to U.S. provisional application Ser. No.: 61/475,042 filed April 13, 2011, and also claims the benefit of and priority to U.S. provisional application Ser. No.: 61/616,341 filed March 27, 2012, the disclosures of all of which are hereby incorporated by reference in their entireties for all purposes.

[0003] This patent disclosure contains material that is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure as it appears in the U.S. Patent and Trademark Office patent file or records, but otherwise reserves any and all copyright rights.

[0004] All patents, patent applications and publications cited herein are hereby incorporated by reference in their entirety. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art as known to those skilled therein as of the date of the invention described herein.

BACKGROUND OF THE INVENTION

[0005] Rapid protein purification is an important requirement in many bioengineering applications where significant amounts of time are currently spent purifying proteins from heterogeneous samples. There are currently a number of approaches for performing bioseparation, but these approaches are expensive, time consuming, can require specialized treatments.

[0006] A variety of approaches currently exist for purifying recombinant proteins such as using a poly-histidine tag, glutathione S-transferase (GST) fusions or fusion to an elastin-like peptide (ELP). In the case of ELPs, a fusion protein can be precipitated from solution by increasing the temperature of the sample (Banki, et al., Nat Meth, vol. 2, no. 9, pp. 659-662, 2005; Fong et al, Trends in Biotechnology, vol. 28, no. 5, pp. 272-279, May. 2010). One limitation of ELP technology is that increased temperature can adversely affect the stability of fusion proteins. Another limitation of ELP technology is that inducing temperature changes are difficult in large scale preparations.

[0007] There is a need for improved purification methods for rapid purification of molecules (e.g. exogenously expressed proteins) from heterogeneous samples in a rapid manner and with high levels of recovery. This invention addresses these needs.

SUMMARY OF THE INVENTION

[0008] In certain aspects, the invention relates to a precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs comprise the amino acid sequence of SEQ ID NO: 1

[0009] In certain aspects, the invention relates to a precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of: (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs: 25-1337

[0010] In certain aspects, the invention relates to a precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of: (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337, (c) a polypeptide comprising the amino acid sequence GXXXXXXXX, wherein, (i) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and (ii) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and (iii) the X at position 4 is an amino acid selected from the group consisting of glycine or alanine, and (iv) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and (v) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine, (vi) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and (vii) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and (viii) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine.

[0011] In certain aspects, the invention relates to a precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of: (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337, (c) a

polypeptide comprising the amino acid sequence GXXXXXXXX, wherein, (i) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and (ii) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and (iii) the X at position 4 is an amino acid selected from the group consisting of glycine or alanine, and (iv) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and (v) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine, (vi) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and (vii) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and (viii) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine, or (c) a variant PBRT.

[0012] In certain aspects, the invention relates to a precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of: (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337, (c) a polypeptide comprising the amino acid sequence GXXXXXXXX, wherein, (i) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and (ii) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and (iii) the X at position 4 is an amino acid selected from the group consisting of glycine or alanine, and (iv) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and (v) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine, (vi) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and (vii) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and (viii) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine, or (c) a variant PBRT, wherein the PBRC further comprises a capping sequence.

[0013] In certain aspects, the invention relates to a precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are

independently any of: (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337, (c) a polypeptide comprising the amino acid sequence GXXXXXXXX, wherein, (i) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and (ii) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and (iii) the X at position 4 is an amino acid selected from the group consisting of glycine or alanine, and (iv) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and (v) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine, (vi) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and (vii) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and (viii) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine, or (c) a variant PBRT, wherein the PBRC further comprises a stabilizing polypeptide.

[0014] In certain aspects, the invention relates to a PBRC linked purification moiety comprising a precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of: (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337, (c) a polypeptide comprising the amino acid sequence GXXXXXXXX, wherein, (i) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and (ii) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and (iii) the X at position 4 is an amino acid selected from the group consisting of glycine or alanine, and (iv) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and (v) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine, (vi) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and (vii) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and (viii) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine, or (c) a variant PBRT. In

certain embodiments, the PBRC is linked to the purification moiety by a peptide bond. In certain embodiments, the PBRC is linked to the purification moiety by a chemical bond that is not a peptide bond.

[0015] In certain aspects, the invention relates to a PBRC linked purification moiety comprising a precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of: (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337, (c) a polypeptide comprising the amino acid sequence GXXXXXXXX, wherein, (i) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and (ii) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and (iii) the X at position 4 is an amino acid selected from the group consisting of glycine or alanine, and (iv) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and (v) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine, (vi) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and (vii) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and (viii) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine, or (c) a variant PBRT, wherein the PBRC further comprises a cleavage site located N-terminally or C-terminally to one or more of the one or more PBRTs. In certain embodiments, the cleavage site is selected from the group comprising an intein cleavage site, a Factor Xa cleavage site, a thrombin cleavage site, an enterokinase cleavage site, or a signal peptidase cleavage site.

[0016] In certain aspects, the invention relates to a polypeptide comprising a PBRC linked purification moiety comprising a precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of: (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337, (c) a polypeptide comprising the amino acid sequence GXXXXXXXX, wherein, (i) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and (ii) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and (iii) the X at position 4 is an amino

acid selected from the group consisting of glycine or alanine, and (iv) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and (v) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine, (vi) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and (vii) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and (viii) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine, or (c) a variant PBRT; and a purification moiety.

[0017] In certain aspects, the invention relates to a nucleic acid encoding any of the polypeptides described herein.

[0018] In certain embodiments, the invention relates to a method for purifying a PBRC linked purification moiety, the method comprising (a) expressing the PBRC linked purification moiety in an expression system, (b) collecting the PBRC linked purification moiety in a first medium, (c) adding Ca^{2+} to the first medium so as to induce precipitation of PBRC linked purification moiety, (d) removing unprecipitated material from the medium from the precipitated PBRC linked purification moiety, (e) resuspending the PBRC linked purification moiety in a second medium having a lower than the free Ca^{2+} concentration than the free Ca^{2+} concentration obtained after step (c). In certain embodiments, a calcium chelator is added to the second medium of step (e). In certain embodiments, steps (c) to (e) are repeated one or more times. In certain embodiments, the method further comprises a step of removing precipitated material between step (b) and step (c). In certain embodiments, the PBRC comprises a cleavage site between the PBRC and the purification moiety. In certain embodiments, the method further comprises steps of: (i) cleaving the PBRC linked purification moiety so as to separate the purification moiety from the PBRC, (ii) adding Ca^{2+} to the medium so as to induce precipitation of the PBRC, and (iii) isolating the unprecipitated purification moiety. In certain embodiments, the cleavage site is an intein cleavage site.

[0019] In certain aspects, the invention relates to an expression vector comprising, as arranged from 5' to 3', a promoter, a nucleic acid sequence encoding the PBRC of any of claims 1-4, and at least one cloning site.

[0020] In certain aspects, the invention relates to an expression vector comprising, as arranged from 5' to 3', a promoter, at least one cloning site, and a nucleic acid sequence encoding the PBRC of any of claims 1-4.

[0021] In certain aspects, the invention relates to an expression vector comprising, as arranged from 5' to 3', a promoter, at least one cloning site, a nucleic acid sequence encoding the PBRC of any of claims 1-4 and at least a second cloning site.

BRIEF DESCRIPTION OF THE FIGURES

[0022] **Figure 1** shows a SDS-PAGE gel showing purification of a molecule comprising a precipitable beta-roll tag. Figure 1A shows purification of precipitable maltose binding protein comprising a precipitable beta roll tag and an enterokinase cleavage site (MBP-PBRT). Total lysate is shown in lane 1 and lanes 2-7 are precipitation/ wash cycles. After two cycles, the sample consists nearly only of the MBP fusion protein with the precipitating tag attached (MBP-PBRT). Recovery is nearly 100% of the expressed protein. Figure 1B shows SDS-PAGE analysis of purified MBP-PBRT and MBP-PBRT subjected to digestion of the enterokinase digestion site (Figure 1B). Lane 1 shows purified MBP-PBRT, lane 2 shows supernatant after overnight digest and lane 3 shows the pellet.

[0023] **Figure 2** shows a SDS-PAGE gel showing a successful purification of a polypeptide comprising a 5 or 17 repeat C-capped precipitable beta-roll tags. Lanes (from left to right): 1. MBP-5cap lysate, 2. MBP-5cap supernatant, 3. MBP-5C resuspended precipitate. The 4-6 lanes are the same expect with the capped 17 repeat construct.

[0024] **Figure 3** is circular dichroism data (CD) showing precipitation of a polypeptide comprising a 17 repeat C-capped precipitable beta-roll tag (GGAGNDTLY)17INAGADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEII HAANQAVDQAGIEKLVEAMAQYPD (SEQ ID NO: 1344) out of solution with increasing calcium concentrations from 0 to 100 mM calcium. Loss of spectra indicates that the peptide is precipitating out of solution and no longer visible via CD.

[0025] **Figure 4** shows a GGXGXDXXX (SEQ ID NO: 2) sequence heat map. The heat map was determined by using BLAST to find beta roll sequences similar to the metalloprotease of *S. marcescens* and then quantifying the frequency of amino acids at each of the nine positions after beta roll sequences were identified. See SEQ ID NO: 1343.

[0026] **Figure 5** shows a schematic illustration of the corkscrew configuration of tandem Ca²⁺ binding sequences. The figure shows a crystal structure image of the beta roll domain from the metalloprotease of *S. marcescens* (1SAT in the Protein Databank). Alternating 9-amino acid repeats are highlighted in green and red. Coordinated calcium ions are in white. The image represents 5 repeats of beta roll sequence. While there is no crystal structure for the adenylate cyclase beta roll domain, the high degree of sequence similarity to the consensus beta roll indicates that the adenylate cyclase beta roll domain is similar.

[0027] **Figure 6** shows a schematic illustration of the corkscrew configuration of tandem Ca²⁺ binding sequences from a different angle than shown in Figure 5. The 6th residue binds the calcium ion. The 7th and 9th residues of each repeat are those that face outwards. The 8th residue is buried in the hydrophobic core. These residues are threonine and tyrosine, respectively in SEQ ID NO: 1.

[0028] **Figure 7** shows the full crystal structure of the metalloprotease of *S. marcescens*. The black spheres indicate the position of the calcium ions within the beta roll domain.

[0029] **Figure 8** shows a characterization of beta roll distribution, sequence deviation and number of repeats. Figure 8A shows a distribution of beta roll lengths as frequency plotted against the number of beta roll repeats. Figure 8B shows beta roll sequence deviation from consensus by position plotted as proportion deviation as a function of distance from terminus. Figure 8C shows a probability of deviation from consensus plotted as probability as a function of the number of beta roll repeats. Figure 8D shows amount of beta roll in overall protein plotted as the number of beta roll residues as a function of the number of total residues.

[0030] **Figure 9** shows an exemplary protocol for purification of a polypeptide comprising a precipitable beta roll tag. The images depict precipitation of maltose binding protein fused to seventeen repeats of a PBRT. The PBRC comprises 17 repeats of the amino acid sequence of SEQ ID NO: 1. The PBRC does not comprise a capping sequence.

[0031] **Figure 10** shows purification of a maltose binding protein/PBRT/green fluorescent protein fusion (MBP-PBRT-GFP). Figure 10A shows a precipitation and resuspension of the MBP-PBRT-GFP polypeptide. Figure 10B shows SDS-PAGE of multiple precipitation wash cycles. The lanes of the SDS gel are as follows: lane 1 – ladder; lane 2 - clarified lysate; lane 3 - precipitation in 25 mM calcium followed by three washes; lane 4 - precipitation in 50 mM calcium followed by three washes; lane 5 - precipitation in 75 mM calcium followed by three

washes; lane - 6 precipitation in 100 mM calcium followed by three washes. All lanes are normalized in terms of concentration. Recovery percentage can be estimated by comparing the band intensity in lane 2 to subsequent lanes.

[0032] Figure 11 shows non-limiting examples of polypeptides comprising a PBRT suitable for use with the methods described herein.

DETAILED DESCRIPTION OF THE INVENTION

[0033] The issued patents, applications, and other publications that are cited herein are hereby incorporated by reference to the same extent as if each was specifically and individually indicated to be incorporated by reference.

[0034] Purification is a major requirement in many bioengineering applications where significant amounts of time are currently spent purifying proteins from heterogeneous samples. The invention described herein relates to methods for rapidly purifying a purification moiety (e.g. a target polypeptide) from a heterogeneous medium using a PBRC. For example, in certain embodiments, a target polypeptide can be produced as a fusion protein in frame with a PBRC. In certain embodiments, the fusion protein comprising the target protein and the PBRC can further comprise a specific cleavage site (e.g. an intein cleavage site or an enterokinase cleavage site) between the target protein sequence and the PBRC sequence. In such embodiments, cleavage at the cleavage site can be used to separate the PBRC from the target polypeptide.

[0035] The beta-roll domain is a right-handed beta helix found in a number of proteins. The consensus sequence for beta-roll peptides is tandem repeats of the 9 amino acid sequence GGXGXDX(L/F/I)X (SEQ ID NO: 24). In the presence of calcium, the conformation aligns to adopt the helical turns. Two repeats of the sequence are required to make a complete helical turn and each of these turns binds a calcium atom. In the absence of calcium, the peptide exists in a disordered conformation. Therefore the β -roll domain exhibits natural allosteric regulation. A synthetic version of the β -roll peptide has been produced with 8 repeats of GGSGNDNLS (SEQ ID NO: 1338) and this peptide was found to bind calcium and fold into the β -roll structure (Lilie et al., FEBS Lett 470 (2), 173 (2000)). The domain is capable of reversibly unfolding upon removal of the calcium. Beta roll sequences are known to play a role in secretion as part of the bacterial Type I secretion system (Davidson, et al., Microbiol. Mol. Biol. Rev. 72 (2008), pp. 317–364; Holland et al., Mol. Membr. Biol. 22 (2005), pp. 29–39; Chenal, et al., J. Biol. Chem. 284(2009), pp. 1781–1789; Welch, Pore-

Forming Toxins 257 (2001), pp. 85–111; Rose et al., J. Biol. Chem. 270 (1995), pp. 26370–26376; Baumann, J. Mol. Biol. 242 (1994), pp. 244–251; Angkawidjaja, et al., FEBS Lett. 581(2007), pp. 5060–5064; Meier et al., J. Biol. Chem. 282 (2007), pp. 31477–31483; Bauche et al., J. Biol. Chem. 281 (2006), pp. 16914–16926; Baumann et al., EMBO J. 12 (1993), pp. 3357–3364; Angkawidjaja et al., FEBS Lett. 579(2005), pp. 4707–4712).

[0036] The precipitable-beta roll tags and precipitable-beta roll cassettes described herein are class of designed peptides which possess the ability to reversibly precipitate in response to calcium ions. In one aspect, the invention described herein relates to the surprising finding that PBRCs (e.g. PBRTs repeats of sequence GGAGNDTLY (SEQ ID NO: 1)) undergo reversible precipitation upon calcium binding. In another aspect, the invention described herein relates to the surprising finding that attachment of a PBRC to a second molecule (e.g. attachment to a protein as a fusion protein comprising an in-frame) can be used to purify the second molecule through reversible precipitation. In another aspect, the invention described herein relates to the use of calcium concentration changes at room temperature to induce precipitation of recombinant molecules comprising a precipitable beta-roll tag.

[0037] In addition to target polypeptides, the PBRCs describe herein are also suitable for purifying non-peptide purification moieties of widely varying types, including, for example, lipids, oligonucleotides and carbohydrates, small organic or inorganic molecules, proteins, single-stranded or double-stranded oligonucleotides, polynucleotides. In certain aspects, applications for the methods and compositions described herein include, but are not limited to, the purification of recombinant proteins the removal of target proteins from a sample, and detection of compounds for diagnostic purposes. The invention also extends to the antibodies that specifically bind to a PBRT or a PBRC and the methods for using the PBRTs and PBRCs described herein.

[0038] Without wishing to be bound to theory, in certain embodiments, the PBRTs and PBRCs described herein can undergo a reversible Ca^{2+} binding dependent transition wherein they are structurally disordered and highly soluble in a medium below a Ca^{2+} concentration (or free Ca^{2+}) transition concentration, but exhibit a disorder to order phase transition when the Ca^{2+} or free Ca^{2+} concentration is raised above the Ca^{2+} (or free Ca^{2+}) transition concentration. Again, without wishing to be bound by theory, in some embodiments, the disorder to order phase transition leads to precipitation of the PBRTs or PBRCs. Precipitation of PBRC can be used to remove and isolated them from solution (e.g. by centrifugation). In one embodiment, the invention described herein relates to a PBRC which functions reversible

Ca²⁺ precipitable tag when linked to a purification moiety of interest. In embodiments where the PBRC is linked to a purification moiety of interest, the methods described herein can be used to induce precipitation of the PBRC linked purification moiety. Because the transition concentration dependent phase transition is reversible, the PBRT and PBRC can be resolubilized in a medium having a Ca²⁺ concentration (or free Ca²⁺) below the transition concentration. In certain embodiments, this can be accomplished by introducing medium having reduced, or no Ca²⁺, or by removing, or chelating Ca²⁺ from the medium. When the precipitate is resuspended in calcium-free buffer or in a buffer comprising a calcium ion chelator (e.g. EGTA or EDTA), the precipitate resuspends into solution.

[0039] The singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise.

[0040] The term "about" is used herein to mean approximately, in the region of, roughly, or around. When the term "about" is used in conjunction with a numerical range, it modifies that range by extending the boundaries above and below the numerical values set forth. In general, the term "about" is used herein to modify a numerical value above and below the stated value by a variance of 20%.

[0041] In certain embodiments, the term "precipitable-beta roll tag" (PBRT) refers to an amino acid sequence having the sequence GGAGNDTLY (SEQ ID NO: 1). In certain embodiments, a PBRT refers to an amino acid sequence having the amino acid sequence GXXXXXXXX (SEQ ID NO: 1343), wherein (a) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and (b) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and (c) the X at position 4 is an amino acid selected from the group consisting of glycine or alanine, and (d) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and (e) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine, (f) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and (g) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and (h) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine, or a nucleic acid encoding the same. A PBRT refers to an amino acid sequence having the sequence set forth in any of SEQ ID NOs: 25-1337.

[0042] Table 1. SEQ ID NOs: 25-1337

Sequence (SEQ ID NO)		Sequence (SEQ ID NO)		Sequence (SEQ ID NO)	
GDNASDLFS	SEQ ID NO: 25	GGGGNDTLI	SEQ ID NO: 463	GGAGDDTLV	SEQ ID NO: 901
GNAGNDTLY	SEQ ID NO: 26	GNGGNDTLI	SEQ ID NO: 464	GNAGDDTLV	SEQ ID NO: 902
GDAGNDTLY	SEQ ID NO: 27	GDGGNDTLI	SEQ ID NO: 465	GDAGDDTLV	SEQ ID NO: 903
GGSGNDTLY	SEQ ID NO: 28	GGDGNDTLI	SEQ ID NO: 466	GGSGDDTLV	SEQ ID NO: 904
GNSGNDTLY	SEQ ID NO: 29	GNDGNDTLI	SEQ ID NO: 467	GNSGDDTLV	SEQ ID NO: 905
GDSGNDTLY	SEQ ID NO: 30	GDDGNDTLI	SEQ ID NO: 468	GDSGDDTLV	SEQ ID NO: 906
GGGGNDTLY	SEQ ID NO: 31	GGAGDDTLI	SEQ ID NO: 469	GGGGDDTLV	SEQ ID NO: 907
GNGGNDTLY	SEQ ID NO: 32	GNAGDDTLI	SEQ ID NO: 470	GNGGDDTLV	SEQ ID NO: 908
GDGGNDTLY	SEQ ID NO: 33	GDAGDDTLI	SEQ ID NO: 471	GDGGDDTLV	SEQ ID NO: 909
GGDGNDTLY	SEQ ID NO: 34	GGSGDDTLI	SEQ ID NO: 472	GGDGDDTLV	SEQ ID NO: 910
GNDGNDTLY	SEQ ID NO: 35	GNSGDDTLI	SEQ ID NO: 473	GNDGDDTLV	SEQ ID NO: 911
GDDGNDTLY	SEQ ID NO: 36	GDSGDDTLI	SEQ ID NO: 474	GDDGDDTLV	SEQ ID NO: 912
GGAGDDTLY	SEQ ID NO: 37	GGGGDDTLI	SEQ ID NO: 475	GGAGADTLV	SEQ ID NO: 913
GNAGDDTLY	SEQ ID NO: 38	GNGGDDTLI	SEQ ID NO: 476	GNAGADTLV	SEQ ID NO: 914
GDAGDDTLY	SEQ ID NO: 39	GDGGDDTLI	SEQ ID NO: 477	GDAGADTLV	SEQ ID NO: 915
GGSGDDTLY	SEQ ID NO: 40	GGDGDDTLI	SEQ ID NO: 478	GGSGADTLV	SEQ ID NO: 916
GNSGDDTLY	SEQ ID NO: 41	GNDGDDTLI	SEQ ID NO: 479	GNSGADTLV	SEQ ID NO: 917
GDSGDDTLY	SEQ ID NO: 42	GDDGDDTLI	SEQ ID NO: 480	GDSGADTLV	SEQ ID NO: 918
GGGGDDTLY	SEQ ID NO: 43	GGAGADTLI	SEQ ID NO: 481	GGGGADTLV	SEQ ID NO: 919
GNGGDDTLY	SEQ ID NO: 44	GNAGADTLI	SEQ ID NO: 482	GNGGADTLV	SEQ ID NO: 920
GDGGDDTLY	SEQ ID NO: 45	GDAGADTLI	SEQ ID NO: 483	GDGGADTLV	SEQ ID NO: 921
GGDGDDTLY	SEQ ID NO: 46	GGSGADTLI	SEQ ID NO: 484	GGDGADTLV	SEQ ID NO: 922
GNDGDDTLY	SEQ ID NO: 47	GNSGADTLI	SEQ ID NO: 485	GNDGADTLV	SEQ ID NO: 923
GDDGDDTLY	SEQ ID NO: 48	GDSGADTLI	SEQ ID NO: 486	GDDGADTLV	SEQ ID NO: 924
GGAGADTLY	SEQ ID NO: 49	GGGGADTLI	SEQ ID NO: 487	GGAGNNTLV	SEQ ID NO: 925
GNAGADTLY	SEQ ID NO: 50	GNGGADTLI	SEQ ID NO: 488	GNAGNNTLV	SEQ ID NO: 926
GDAGADTLY	SEQ ID NO: 51	GDGGADTLI	SEQ ID NO: 489	GDAGNNTLV	SEQ ID NO: 927
GGSGADTLY	SEQ ID NO: 52	GGDGADTLI	SEQ ID NO: 490	GGSGNNTLV	SEQ ID NO: 928
GNSGADTLY	SEQ ID NO: 53	GNDGADTLI	SEQ ID NO: 491	GNSGNNTLV	SEQ ID NO: 929
GDSGADTLY	SEQ ID NO: 54	GDDGADTLI	SEQ ID NO: 492	GDSGNNTLV	SEQ ID NO: 930
GGGGADTLY	SEQ ID NO: 55	GGAGNNTLI	SEQ ID NO: 493	GGGGNNTLV	SEQ ID NO: 931
GNGGADTLY	SEQ ID NO: 56	GNAGNNTLI	SEQ ID NO: 494	GNGGNNTLV	SEQ ID NO: 932
GDGGADTLY	SEQ ID NO: 57	GDAGNNTLI	SEQ ID NO: 495	GDGGNNTLV	SEQ ID NO: 933
GGDGADTLY	SEQ ID NO: 58	GGSGNNTLI	SEQ ID NO: 496	GGDGNNTLV	SEQ ID NO: 934
GNDGADTLY	SEQ ID NO: 59	GNSGNNTLI	SEQ ID NO: 497	GNDGNNTLV	SEQ ID NO: 935
GDDGADTLY	SEQ ID NO: 60	GDSGNNTLI	SEQ ID NO: 498	GDDGNNTLV	SEQ ID NO: 936
GGAGNNTLY	SEQ ID NO: 61	GGGGNNTLI	SEQ ID NO: 499	GGAGDNTLV	SEQ ID NO: 937
GNAGNNTLY	SEQ ID NO: 62	GNGGNNTLI	SEQ ID NO: 500	GNAGDNTLV	SEQ ID NO: 938
GDAGNNTLY	SEQ ID NO: 63	GDGGNNTLI	SEQ ID NO: 501	GDAGDNTLV	SEQ ID NO: 939
GGSGNNTLY	SEQ ID NO: 64	GGDGNNTLI	SEQ ID NO: 502	GGSGDNTLV	SEQ ID NO: 940
GNSGNNTLY	SEQ ID NO: 65	GNDGNNTLI	SEQ ID NO: 503	GNSGDNTLV	SEQ ID NO: 941
GDSGNNTLY	SEQ ID NO: 66	GDDGNNTLI	SEQ ID NO: 504	GDSGDNTLV	SEQ ID NO: 942
GGGGNNTLY	SEQ ID NO: 67	GGAGDNTLI	SEQ ID NO: 505	GGGGDNTLV	SEQ ID NO: 943
GNGGNNTLY	SEQ ID NO: 68	GNAGDNTLI	SEQ ID NO: 506	GNGGDNTLV	SEQ ID NO: 944
GDGGNNTLY	SEQ ID NO: 69	GDAGDNTLI	SEQ ID NO: 507	GDGGDNTLV	SEQ ID NO: 945
GGDGNTLY	SEQ ID NO: 70	GGSGDNTLI	SEQ ID NO: 508	GGDGDNTLV	SEQ ID NO: 946
GNDGNNTLY	SEQ ID NO: 71	GNSGDNTLI	SEQ ID NO: 509	GNDGDNTLV	SEQ ID NO: 947
GDDGNNTLY	SEQ ID NO: 72	GDSGDNTLI	SEQ ID NO: 510	GDDGDNTLV	SEQ ID NO: 948
GGAGDNTLY	SEQ ID NO: 73	GGGGDNTLI	SEQ ID NO: 511	GGAGANTLV	SEQ ID NO: 949
GNAGDNTLY	SEQ ID NO: 74	GNGGDNTLI	SEQ ID NO: 512	GNAGANTLV	SEQ ID NO: 950
GDAGDNTLY	SEQ ID NO: 75	GDGGDNTLI	SEQ ID NO: 513	GDAGANTLV	SEQ ID NO: 951

Sequence (SEQ ID NO)		Sequence (SEQ ID NO)		Sequence (SEQ ID NO)	
GGSGDNTLY	SEQ ID NO: 76	GGDGDNTLI	SEQ ID NO: 514	GGSGANTLV	SEQ ID NO: 952
GNSGDNTLY	SEQ ID NO: 77	GNDGDNTLI	SEQ ID NO: 515	GNSGANTLV	SEQ ID NO: 953
GDSGDNTLY	SEQ ID NO: 78	GDDGDNTLI	SEQ ID NO: 516	GDSGANTLV	SEQ ID NO: 954
GGGGDNTLY	SEQ ID NO: 79	GGAGANTLI	SEQ ID NO: 517	GGGGANTLV	SEQ ID NO: 955
GNGGDNTLY	SEQ ID NO: 80	GNAGANTLI	SEQ ID NO: 518	GNGGANTLV	SEQ ID NO: 956
GDGGDNTLY	SEQ ID NO: 81	GDAGANTLI	SEQ ID NO: 519	GDGGANTLV	SEQ ID NO: 957
GGDGDNTLY	SEQ ID NO: 82	GGSGANTLI	SEQ ID NO: 520	GGDGANTLV	SEQ ID NO: 958
GNDGDNTLY	SEQ ID NO: 83	GNSGANTLI	SEQ ID NO: 521	GNDGANTLV	SEQ ID NO: 959
GDDGDNTLY	SEQ ID NO: 84	GDSGANTLI	SEQ ID NO: 522	GDDGANTLV	SEQ ID NO: 960
GGAGANTLY	SEQ ID NO: 85	GGGGANTLI	SEQ ID NO: 523	GGAGNDILV	SEQ ID NO: 961
GNAGANTLY	SEQ ID NO: 86	GNGGANTLI	SEQ ID NO: 524	GNAGNDILV	SEQ ID NO: 962
GDAGANTLY	SEQ ID NO: 87	GDGGANTLI	SEQ ID NO: 525	GDAGNDILV	SEQ ID NO: 963
GGSGANTLY	SEQ ID NO: 88	GGDGANTLI	SEQ ID NO: 526	GGSGNDILV	SEQ ID NO: 964
GNSGANTLY	SEQ ID NO: 89	GNDGANTLI	SEQ ID NO: 527	GNSGNDILV	SEQ ID NO: 965
GDSGANTLY	SEQ ID NO: 90	GDDGANTLI	SEQ ID NO: 528	GDSGNDILV	SEQ ID NO: 966
GGGGANTLY	SEQ ID NO: 91	GGAGNDILI	SEQ ID NO: 529	GGGGNDILV	SEQ ID NO: 967
GNGGANTLY	SEQ ID NO: 92	GNAGNDILI	SEQ ID NO: 530	GNGGNDILV	SEQ ID NO: 968
GDGGANTLY	SEQ ID NO: 93	GDAGNDILI	SEQ ID NO: 531	GDGGNDILV	SEQ ID NO: 969
GGDGANTLY	SEQ ID NO: 94	GGSGNDILI	SEQ ID NO: 532	GGDGNDILV	SEQ ID NO: 970
GNDGANTLY	SEQ ID NO: 95	GNSGNDILI	SEQ ID NO: 533	GNDGNDILV	SEQ ID NO: 971
GDDGANTLY	SEQ ID NO: 96	GDSGNDILI	SEQ ID NO: 534	GDDGNDILV	SEQ ID NO: 972
GGAGNDILY	SEQ ID NO: 97	GGGGNDILI	SEQ ID NO: 535	GGAGDDILV	SEQ ID NO: 973
GNAGNDILY	SEQ ID NO: 98	GNGGNDILI	SEQ ID NO: 536	GNAGDDILV	SEQ ID NO: 974
GDAGNDILY	SEQ ID NO: 99	GDGGNDILI	SEQ ID NO: 537	GDAGDDILV	SEQ ID NO: 975
GGSGNDILY	SEQ ID NO: 100	GGDGNDILI	SEQ ID NO: 538	GGSGDDILV	SEQ ID NO: 976
GNSGNDILY	SEQ ID NO: 101	GNDGNDILI	SEQ ID NO: 539	GNSGDDILV	SEQ ID NO: 977
GDSGNDILY	SEQ ID NO: 102	GDDGNDILI	SEQ ID NO: 540	GDSGDDILV	SEQ ID NO: 978
GGGGNDILY	SEQ ID NO: 103	GGAGDDILI	SEQ ID NO: 541	GGGGDDILV	SEQ ID NO: 979
GNGGNDILY	SEQ ID NO: 104	GNAGDDILI	SEQ ID NO: 542	GNGGDDILV	SEQ ID NO: 980
GDGGNDILY	SEQ ID NO: 105	GDAGDDILI	SEQ ID NO: 543	GDGGDDILV	SEQ ID NO: 981
GGDGNDILY	SEQ ID NO: 106	GGSGDDILI	SEQ ID NO: 544	GGDGDDILV	SEQ ID NO: 982
GNDGNDILY	SEQ ID NO: 107	GNSGDDILI	SEQ ID NO: 545	GNDGDDILV	SEQ ID NO: 983
GDDGNDILY	SEQ ID NO: 108	GDSGDDILI	SEQ ID NO: 546	GDDGDDILV	SEQ ID NO: 984
GGAGDDILY	SEQ ID NO: 109	GGGGDDILI	SEQ ID NO: 547	GGAGADILV	SEQ ID NO: 985
GNAGDDILY	SEQ ID NO: 110	GNGGDDILI	SEQ ID NO: 548	GNAGADILV	SEQ ID NO: 986
GDAGDDILY	SEQ ID NO: 111	GDGGDDILI	SEQ ID NO: 549	GDAGADILV	SEQ ID NO: 987
GGSGDDILY	SEQ ID NO: 112	GGDGDDILI	SEQ ID NO: 550	GGSGADILV	SEQ ID NO: 988
GNSGDDILY	SEQ ID NO: 113	GNDGDDILI	SEQ ID NO: 551	GNSGADILV	SEQ ID NO: 989
GDSGDDILY	SEQ ID NO: 114	GDDGDDILI	SEQ ID NO: 552	GDSGADILV	SEQ ID NO: 990
GGGGDDILY	SEQ ID NO: 115	GGAGADILI	SEQ ID NO: 553	GGGGADILV	SEQ ID NO: 991
GNGGDDILY	SEQ ID NO: 116	GNAGADILI	SEQ ID NO: 554	GNGGADILV	SEQ ID NO: 992
GDGGDDILY	SEQ ID NO: 117	GDAGADILI	SEQ ID NO: 555	GDGGADILV	SEQ ID NO: 993
GGDGDDILY	SEQ ID NO: 118	GGSGADILI	SEQ ID NO: 556	GGDGADILV	SEQ ID NO: 994
GNDGDDILY	SEQ ID NO: 119	GNSGADILI	SEQ ID NO: 557	GNDGADILV	SEQ ID NO: 995
GDDGDDILY	SEQ ID NO: 120	GDSGADILI	SEQ ID NO: 558	GDDGADILV	SEQ ID NO: 996
GGAGADILY	SEQ ID NO: 121	GGGGADILI	SEQ ID NO: 559	GGAGNNILV	SEQ ID NO: 997
GNAGADILY	SEQ ID NO: 122	GNGGADILI	SEQ ID NO: 560	GNAGNNILV	SEQ ID NO: 998
GDAGADILY	SEQ ID NO: 123	GDGGADILI	SEQ ID NO: 561	GDAGNNILV	SEQ ID NO: 999
GGSGADILY	SEQ ID NO: 124	GGDGADILI	SEQ ID NO: 562	GGSGNNILV	SEQ ID NO: 1000
GNSGADILY	SEQ ID NO: 125	GNDGADILI	SEQ ID NO: 563	GNSGNNILV	SEQ ID NO: 1001
GDSGADILY	SEQ ID NO: 126	GDDGADILI	SEQ ID NO: 564	GDSGNNILV	SEQ ID NO: 1002
GGGGADILY	SEQ ID NO: 127	GGAGNNILI	SEQ ID NO: 565	GGGGNNILV	SEQ ID NO: 1003
GNGGADILY	SEQ ID NO: 128	GNAGNNILI	SEQ ID NO: 566	GNGGNNILV	SEQ ID NO: 1004

Sequence (SEQ ID NO)		Sequence (SEQ ID NO)		Sequence (SEQ ID NO)	
GDGGADILY	SEQ ID NO: 129	GDAGNNILI	SEQ ID NO: 567	GDGGNNILV	SEQ ID NO: 1005
GGDGADILY	SEQ ID NO: 130	GGSGNNILI	SEQ ID NO: 568	GGDGNNILV	SEQ ID NO: 1006
GNDGADILY	SEQ ID NO: 131	GNSGNNILI	SEQ ID NO: 569	GNDGNNILV	SEQ ID NO: 1007
GDDGADILY	SEQ ID NO: 132	GDSGNNILI	SEQ ID NO: 570	GDDGNNILV	SEQ ID NO: 1008
GGAGNNILY	SEQ ID NO: 133	GGGGNNILI	SEQ ID NO: 571	GGAGDNILV	SEQ ID NO: 1009
GNAGNNILY	SEQ ID NO: 134	GNGGNNILI	SEQ ID NO: 572	GNAGDNILV	SEQ ID NO: 1010
GDAGNNILY	SEQ ID NO: 135	GDGGNNILI	SEQ ID NO: 573	GDAGDNILV	SEQ ID NO: 1011
GGSGNNILY	SEQ ID NO: 136	GGDGNNILI	SEQ ID NO: 574	GGSGDNILV	SEQ ID NO: 1012
GNSGNNILY	SEQ ID NO: 137	GNDGNNILI	SEQ ID NO: 575	GNSGDNLV	SEQ ID NO: 1013
GDSGNNILY	SEQ ID NO: 138	GDDGNNILI	SEQ ID NO: 576	GDSGDNLV	SEQ ID NO: 1014
GGGGNNILY	SEQ ID NO: 139	GGAGDNILI	SEQ ID NO: 577	GGGGDNILV	SEQ ID NO: 1015
GNGGNNILY	SEQ ID NO: 140	GNAGDNILI	SEQ ID NO: 578	GNGGDNLV	SEQ ID NO: 1016
GDGGNNILY	SEQ ID NO: 141	GDAGDNILI	SEQ ID NO: 579	GDGGDNILV	SEQ ID NO: 1017
GGDGNNILY	SEQ ID NO: 142	GGSGDNILI	SEQ ID NO: 580	GGDGDNILV	SEQ ID NO: 1018
GNDGNNILY	SEQ ID NO: 143	GNSGDNLV	SEQ ID NO: 581	GNDGDNLV	SEQ ID NO: 1019
GDDGNNILY	SEQ ID NO: 144	GDSGDNLV	SEQ ID NO: 582	GDDGDNLV	SEQ ID NO: 1020
GGAGDNILY	SEQ ID NO: 145	GGGGDNILI	SEQ ID NO: 583	GGAGANILV	SEQ ID NO: 1021
GNAGDNILY	SEQ ID NO: 146	GNGGDNILI	SEQ ID NO: 584	GNAGANILV	SEQ ID NO: 1022
GDAGDNILY	SEQ ID NO: 147	GDGGDNILI	SEQ ID NO: 585	GDAGANILV	SEQ ID NO: 1023
GGSGDNILY	SEQ ID NO: 148	GGDGDNILI	SEQ ID NO: 586	GGSGANILV	SEQ ID NO: 1024
GNSGDNLV	SEQ ID NO: 149	GNDGDNLV	SEQ ID NO: 587	GNSGANILV	SEQ ID NO: 1025
GDSGDNLV	SEQ ID NO: 150	GDDGDNLV	SEQ ID NO: 588	GDSGANILV	SEQ ID NO: 1026
GGGGDNILY	SEQ ID NO: 151	GGAGANILI	SEQ ID NO: 589	GGGGANILV	SEQ ID NO: 1027
GNGGDNLV	SEQ ID NO: 152	GNAGANILI	SEQ ID NO: 590	GNGGANILV	SEQ ID NO: 1028
GDGGDNILY	SEQ ID NO: 153	GDAGANILI	SEQ ID NO: 591	GDGGANILV	SEQ ID NO: 1029
GGDGDNILY	SEQ ID NO: 154	GGSGANILI	SEQ ID NO: 592	GGDGANILV	SEQ ID NO: 1030
GNDGDNLV	SEQ ID NO: 155	GNSGANILI	SEQ ID NO: 593	GNDGANILV	SEQ ID NO: 1031
GDDGDNLV	SEQ ID NO: 156	GDSGANILI	SEQ ID NO: 594	GDDGANILV	SEQ ID NO: 1032
GGAGANILY	SEQ ID NO: 157	GGGGANILI	SEQ ID NO: 595	GGAGNDVLV	SEQ ID NO: 1033
GNAGANILY	SEQ ID NO: 158	GNGGANILI	SEQ ID NO: 596	GNAGNDVLV	SEQ ID NO: 1034
GDAGANILY	SEQ ID NO: 159	GDGGANILI	SEQ ID NO: 597	GDAGNDVLV	SEQ ID NO: 1035
GGSGANILY	SEQ ID NO: 160	GGDGANILI	SEQ ID NO: 598	GGSGNDVLV	SEQ ID NO: 1036
GNSGANILY	SEQ ID NO: 161	GNDGANILI	SEQ ID NO: 599	GNSGNDVLV	SEQ ID NO: 1037
GDSGANILY	SEQ ID NO: 162	GDDGANILI	SEQ ID NO: 600	GDSGNDVLV	SEQ ID NO: 1038
GGGGANILY	SEQ ID NO: 163	GGAGNDVLI	SEQ ID NO: 601	GGGGNDVLV	SEQ ID NO: 1039
GNGGANILY	SEQ ID NO: 164	GNAGNDVLI	SEQ ID NO: 602	GNGGNDVLV	SEQ ID NO: 1040
GDGGANILY	SEQ ID NO: 165	GDAGNDVLI	SEQ ID NO: 603	GDGGNDVLV	SEQ ID NO: 1041
GGDGANILY	SEQ ID NO: 166	GGSGNDVLI	SEQ ID NO: 604	GGDGNDVLV	SEQ ID NO: 1042
GNDGANILY	SEQ ID NO: 167	GNSGNDVLI	SEQ ID NO: 605	GNDGNDVLV	SEQ ID NO: 1043
GDDGANILY	SEQ ID NO: 168	GDSGNDVLI	SEQ ID NO: 606	GDDGNDVLV	SEQ ID NO: 1044
GGAGNDVLY	SEQ ID NO: 169	GGGGNDVLI	SEQ ID NO: 607	GGAGDDVLV	SEQ ID NO: 1045
GNAGNDVLY	SEQ ID NO: 170	GNGGNDVLI	SEQ ID NO: 608	GNAGDDVLV	SEQ ID NO: 1046
GDAGNDVLY	SEQ ID NO: 171	GDGGNDVLI	SEQ ID NO: 609	GDAGDDVLV	SEQ ID NO: 1047
GGSGNDVLY	SEQ ID NO: 172	GGDGNDVLI	SEQ ID NO: 610	GGSGDDVLV	SEQ ID NO: 1048
GNSGNDVLY	SEQ ID NO: 173	GNDGNDVLI	SEQ ID NO: 611	GNSGDDVLV	SEQ ID NO: 1049
GDSGNDVLY	SEQ ID NO: 174	GDDGNDVLI	SEQ ID NO: 612	GDSGDDVLV	SEQ ID NO: 1050
GGGGNDVLY	SEQ ID NO: 175	GGAGDDVLI	SEQ ID NO: 613	GGGGDDVLV	SEQ ID NO: 1051
GNGGNDVLY	SEQ ID NO: 176	GNAGDDVLI	SEQ ID NO: 614	GNGGDDVLV	SEQ ID NO: 1052
GDGGNDVLY	SEQ ID NO: 177	GDAGDDVLI	SEQ ID NO: 615	GDGGDDVLV	SEQ ID NO: 1053
GGDGNDVLY	SEQ ID NO: 178	GGSGDDVLI	SEQ ID NO: 616	GGDGDDVLV	SEQ ID NO: 1054
GNDGNDVLY	SEQ ID NO: 179	GNSGDDVLI	SEQ ID NO: 617	GNDGDDVLV	SEQ ID NO: 1055
GDDGNDVLY	SEQ ID NO: 180	GDSGDDVLI	SEQ ID NO: 618	GDDGDDVLV	SEQ ID NO: 1056
GGAGDDVLY	SEQ ID NO: 181	GGGGDDVLI	SEQ ID NO: 619	GGAGADVLV	SEQ ID NO: 1057

Sequence (SEQ ID NO)		Sequence (SEQ ID NO)		Sequence (SEQ ID NO)	
GNAGDDVLY	SEQ ID NO: 182	GNGGDDVLI	SEQ ID NO: 620	GNAGADVLY	SEQ ID NO: 1058
GDAGDDVLY	SEQ ID NO: 183	GDGGDDVLI	SEQ ID NO: 621	GDAGADVLY	SEQ ID NO: 1059
GGSGDDVLY	SEQ ID NO: 184	GGDGDDVLI	SEQ ID NO: 622	GGSGADVLY	SEQ ID NO: 1060
GNSGDDVLY	SEQ ID NO: 185	GNDGDDVLI	SEQ ID NO: 623	GNSGADVLY	SEQ ID NO: 1061
GDSGDDVLY	SEQ ID NO: 186	GDDGDDVLI	SEQ ID NO: 624	GDSGADVLY	SEQ ID NO: 1062
GGGGDDVLY	SEQ ID NO: 187	GGAGADVLI	SEQ ID NO: 625	GGGGADVLY	SEQ ID NO: 1063
GNGGDDVLY	SEQ ID NO: 188	GNAGADVLI	SEQ ID NO: 626	GNGGADVLY	SEQ ID NO: 1064
GDGGDDVLY	SEQ ID NO: 189	GDAGADVLI	SEQ ID NO: 627	GDGGADVLY	SEQ ID NO: 1065
GGDGDDVLY	SEQ ID NO: 190	GGSGADVLI	SEQ ID NO: 628	GGDGADVLY	SEQ ID NO: 1066
GNDGDDVLY	SEQ ID NO: 191	GNSGADVLI	SEQ ID NO: 629	GNDGADVLY	SEQ ID NO: 1067
GDDGDDVLY	SEQ ID NO: 192	GDSGADVLI	SEQ ID NO: 630	GDDGADVLY	SEQ ID NO: 1068
GGAGADVLY	SEQ ID NO: 193	GGGGADVLI	SEQ ID NO: 631	GGAGNNVLY	SEQ ID NO: 1069
GNAGADVLY	SEQ ID NO: 194	GNGGADVLI	SEQ ID NO: 632	GNAGNNVLY	SEQ ID NO: 1070
GDAGADVLY	SEQ ID NO: 195	GDGGADVLI	SEQ ID NO: 633	GDAGNNVLY	SEQ ID NO: 1071
GGSGADVLY	SEQ ID NO: 196	GGDGADVLI	SEQ ID NO: 634	GGSGNNVLY	SEQ ID NO: 1072
GNSGADVLY	SEQ ID NO: 197	GNDGADVLI	SEQ ID NO: 635	GNSGNNVLY	SEQ ID NO: 1073
GDSGADVLY	SEQ ID NO: 198	GDDGADVLI	SEQ ID NO: 636	GDSGNNVLY	SEQ ID NO: 1074
GGGGADVLY	SEQ ID NO: 199	GGAGNNVLI	SEQ ID NO: 637	GGGGNNVLY	SEQ ID NO: 1075
GNGGADVLY	SEQ ID NO: 200	GNAGNNVLI	SEQ ID NO: 638	GNGGNNVLY	SEQ ID NO: 1076
GDGGADVLY	SEQ ID NO: 201	GDAGNNVLI	SEQ ID NO: 639	GDGGNNVLY	SEQ ID NO: 1077
GGDGADVLY	SEQ ID NO: 202	GGSGNNVLI	SEQ ID NO: 640	GGDGNNVLY	SEQ ID NO: 1078
GNDGADVLY	SEQ ID NO: 203	GNSGNNVLI	SEQ ID NO: 641	GNDGNNVLY	SEQ ID NO: 1079
GDDGADVLY	SEQ ID NO: 204	GDSGNNVLI	SEQ ID NO: 642	GDDGNNVLY	SEQ ID NO: 1080
GGAGNNVLY	SEQ ID NO: 205	GGGGNNVLI	SEQ ID NO: 643	GGAGDNVLY	SEQ ID NO: 1081
GNAGNNVLY	SEQ ID NO: 206	GNGGNNVLI	SEQ ID NO: 644	GNAGDNVLY	SEQ ID NO: 1082
GDAGNNVLY	SEQ ID NO: 207	GDGGNNVLI	SEQ ID NO: 645	GDAGDNVLY	SEQ ID NO: 1083
GGSGNNVLY	SEQ ID NO: 208	GGDGNNVLI	SEQ ID NO: 646	GGSGDNVLY	SEQ ID NO: 1084
GNSGNNVLY	SEQ ID NO: 209	GNDGNNVLI	SEQ ID NO: 647	GNSGDNVLY	SEQ ID NO: 1085
GDSGNNVLY	SEQ ID NO: 210	GDDGNNVLI	SEQ ID NO: 648	GDSGDNVLY	SEQ ID NO: 1086
GGGGNNVLY	SEQ ID NO: 211	GGAGDNVLI	SEQ ID NO: 649	GGGGDNVLY	SEQ ID NO: 1087
GNGGNNVLY	SEQ ID NO: 212	GNAGDNVLI	SEQ ID NO: 650	GNGGDNVLY	SEQ ID NO: 1088
GDGGNNVLY	SEQ ID NO: 213	GDAGDNVLI	SEQ ID NO: 651	GDGGDNVLY	SEQ ID NO: 1089
GGDGNNVLY	SEQ ID NO: 214	GGSGDNVLI	SEQ ID NO: 652	GGDGDNVLY	SEQ ID NO: 1090
GNDGNNVLY	SEQ ID NO: 215	GNSGDNVLI	SEQ ID NO: 653	GNDGDNVLY	SEQ ID NO: 1091
GDDGNNVLY	SEQ ID NO: 216	GDSGDNVLI	SEQ ID NO: 654	GDDGDNVLY	SEQ ID NO: 1092
GGAGDNVLY	SEQ ID NO: 217	GGGGDNVLI	SEQ ID NO: 655	GGAGANVLY	SEQ ID NO: 1093
GNAGDNVLY	SEQ ID NO: 218	GNGGDNVLI	SEQ ID NO: 656	GNAGANVLY	SEQ ID NO: 1094
GDAGDNVLY	SEQ ID NO: 219	GDGGDNVLI	SEQ ID NO: 657	GDAGANVLY	SEQ ID NO: 1095
GGSGDNVLY	SEQ ID NO: 220	GGDGDNVLI	SEQ ID NO: 658	GGSGANVLY	SEQ ID NO: 1096
GNSGDNVLY	SEQ ID NO: 221	GNDGDNVLI	SEQ ID NO: 659	GNSGANVLY	SEQ ID NO: 1097
GDSGDNVLY	SEQ ID NO: 222	GDDGDNVLI	SEQ ID NO: 660	GDSGANVLY	SEQ ID NO: 1098
GGGGDNVLY	SEQ ID NO: 223	GGAGANVLI	SEQ ID NO: 661	GGGGANVLY	SEQ ID NO: 1099
GNGGDNVLY	SEQ ID NO: 224	GNAGANVLI	SEQ ID NO: 662	GNGGANVLY	SEQ ID NO: 1100
GDGGDNVLY	SEQ ID NO: 225	GDAGANVLI	SEQ ID NO: 663	GDGGANVLY	SEQ ID NO: 1101
GGDGDNVLY	SEQ ID NO: 226	GGSGANVLI	SEQ ID NO: 664	GGDGANVLY	SEQ ID NO: 1102
GNDGDNVLY	SEQ ID NO: 227	GNSGANVLI	SEQ ID NO: 665	GNDGANVLY	SEQ ID NO: 1103
GDDGDNVLY	SEQ ID NO: 228	GDSGANVLI	SEQ ID NO: 666	GDDGANVLY	SEQ ID NO: 1104
GGAGANVLY	SEQ ID NO: 229	GGGGANVLI	SEQ ID NO: 667	GGAGNDTIV	SEQ ID NO: 1105
GNAGANVLY	SEQ ID NO: 230	GNGGANVLI	SEQ ID NO: 668	GNAGNDTIV	SEQ ID NO: 1106
GDAGANVLY	SEQ ID NO: 231	GDGGANVLI	SEQ ID NO: 669	GDAGNDTIV	SEQ ID NO: 1107
GGSGANVLY	SEQ ID NO: 232	GGDGANVLI	SEQ ID NO: 670	GGSGNDTIV	SEQ ID NO: 1108
GNSGANVLY	SEQ ID NO: 233	GNDGANVLI	SEQ ID NO: 671	GNSGNDTIV	SEQ ID NO: 1109
GDSGANVLY	SEQ ID NO: 234	GDDGANVLI	SEQ ID NO: 672	GDSGNDTIV	SEQ ID NO: 1110

Sequence (SEQ ID NO)		Sequence (SEQ ID NO)		Sequence (SEQ ID NO)	
GGGGANVLY	SEQ ID NO: 235	GGAGNDTII	SEQ ID NO: 673	GGGGNDTIV	SEQ ID NO: 1111
GNGGANVLY	SEQ ID NO: 236	GNAGNDTII	SEQ ID NO: 674	GNGGNDTIV	SEQ ID NO: 1112
GDGGANVLY	SEQ ID NO: 237	GDAGNDTII	SEQ ID NO: 675	GDGGNDTIV	SEQ ID NO: 1113
GGDGANVLY	SEQ ID NO: 238	GGSGNDTII	SEQ ID NO: 676	GGDGNDTIV	SEQ ID NO: 1114
GNDGANVLY	SEQ ID NO: 239	GNSGNDTII	SEQ ID NO: 677	GNDGNDTIV	SEQ ID NO: 1115
GDDGANVLY	SEQ ID NO: 240	GDSGNDTII	SEQ ID NO: 678	GDDGNDTIV	SEQ ID NO: 1116
GGAGNDTIY	SEQ ID NO: 241	GGGGNDTII	SEQ ID NO: 679	GGAGDDTIV	SEQ ID NO: 1117
GNAGNDTIY	SEQ ID NO: 242	GNGGNDTII	SEQ ID NO: 680	GNAGDDTIV	SEQ ID NO: 1118
GDAGNDTIY	SEQ ID NO: 243	GDGGNDTII	SEQ ID NO: 681	GDAGDDTIV	SEQ ID NO: 1119
GGSGNDTIY	SEQ ID NO: 244	GGDGNDTII	SEQ ID NO: 682	GGSGDDTIV	SEQ ID NO: 1120
GNSGNDTIY	SEQ ID NO: 245	GNDGNDTII	SEQ ID NO: 683	GNSGDDTIV	SEQ ID NO: 1121
GDSGNDTIY	SEQ ID NO: 246	GDDGNDTII	SEQ ID NO: 684	GDSGDDTIV	SEQ ID NO: 1122
GGGGNDTIY	SEQ ID NO: 247	GGAGDDTII	SEQ ID NO: 685	GGGGDDTIV	SEQ ID NO: 1123
GNGGNDTIY	SEQ ID NO: 248	GNAGDDTII	SEQ ID NO: 686	GNGGDDTIV	SEQ ID NO: 1124
GDGGNDTIY	SEQ ID NO: 249	GDAGDDTII	SEQ ID NO: 687	GDGGDDTIV	SEQ ID NO: 1125
GGDGNDTIY	SEQ ID NO: 250	GGSGDDTII	SEQ ID NO: 688	GGDGDDTIV	SEQ ID NO: 1126
GNDGNDTIY	SEQ ID NO: 251	GNSGDDTII	SEQ ID NO: 689	GNDGDDTIV	SEQ ID NO: 1127
GDDGNDTIY	SEQ ID NO: 252	GDSGDDTII	SEQ ID NO: 690	GDDGDDTIV	SEQ ID NO: 1128
GGAGDDTIY	SEQ ID NO: 253	GGGGDDTII	SEQ ID NO: 691	GGAGADTIV	SEQ ID NO: 1129
GNAGDDTIY	SEQ ID NO: 254	GNGGDDTII	SEQ ID NO: 692	GNAGADTIV	SEQ ID NO: 1130
GDAGDDTIY	SEQ ID NO: 255	GDGGDDTII	SEQ ID NO: 693	GDAGADTIV	SEQ ID NO: 1131
GGSGDDTIY	SEQ ID NO: 256	GGDGDDTII	SEQ ID NO: 694	GGSGADTIV	SEQ ID NO: 1132
GNSGDDTIY	SEQ ID NO: 257	GNDGDDTII	SEQ ID NO: 695	GNSGADTIV	SEQ ID NO: 1133
GDSGDDTIY	SEQ ID NO: 258	GDDGDDTII	SEQ ID NO: 696	GDSGADTIV	SEQ ID NO: 1134
GGGGDDTIY	SEQ ID NO: 259	GGAGADTII	SEQ ID NO: 697	GGGGADTIV	SEQ ID NO: 1135
GNGGDDTIY	SEQ ID NO: 260	GNAGADTII	SEQ ID NO: 698	GNGGADTIV	SEQ ID NO: 1136
GDGGDDTIY	SEQ ID NO: 261	GDAGADTII	SEQ ID NO: 699	GDGGADTIV	SEQ ID NO: 1137
GGDGDDTIY	SEQ ID NO: 262	GGSGADTII	SEQ ID NO: 700	GGDGADTIV	SEQ ID NO: 1138
GNDGDDTIY	SEQ ID NO: 263	GNSGADTII	SEQ ID NO: 701	GNDGADTIV	SEQ ID NO: 1139
GDDGDDTIY	SEQ ID NO: 264	GDSGADTII	SEQ ID NO: 702	GDDGADTIV	SEQ ID NO: 1140
GGAGADTIY	SEQ ID NO: 265	GGGGADTII	SEQ ID NO: 703	GGAGNNTIV	SEQ ID NO: 1141
GNAGADTIY	SEQ ID NO: 266	GNGGADTII	SEQ ID NO: 704	GNAGNNTIV	SEQ ID NO: 1142
GDAGADTIY	SEQ ID NO: 267	GDGGADTII	SEQ ID NO: 705	GDAGNNTIV	SEQ ID NO: 1143
GGSGADTIY	SEQ ID NO: 268	GGDGADTII	SEQ ID NO: 706	GGSGNNTIV	SEQ ID NO: 1144
GNSGADTIY	SEQ ID NO: 269	GNDGADTII	SEQ ID NO: 707	GNSGNNTIV	SEQ ID NO: 1145
GDSGADTIY	SEQ ID NO: 270	GDDGADTII	SEQ ID NO: 708	GDSGNNTIV	SEQ ID NO: 1146
GGGGADTIY	SEQ ID NO: 271	GGAGNNTII	SEQ ID NO: 709	GGGGNNTIV	SEQ ID NO: 1147
GNGGADTIY	SEQ ID NO: 272	GNAGNNTII	SEQ ID NO: 710	GNGGNNTIV	SEQ ID NO: 1148
GDGGADTIY	SEQ ID NO: 273	GDAGNNTII	SEQ ID NO: 711	GDGGNNTIV	SEQ ID NO: 1149
GGDGADTIY	SEQ ID NO: 274	GGSGNNTII	SEQ ID NO: 712	GGDGNTIV	SEQ ID NO: 1150
GNDGADTIY	SEQ ID NO: 275	GNSGNNTII	SEQ ID NO: 713	GNDGNNTIV	SEQ ID NO: 1151
GDDGADTIY	SEQ ID NO: 276	GDSGNNTII	SEQ ID NO: 714	GDDGNNTIV	SEQ ID NO: 1152
GGAGNNTIY	SEQ ID NO: 277	GGGGNNTII	SEQ ID NO: 715	GGAGDNTIV	SEQ ID NO: 1153
GNAGNNTIY	SEQ ID NO: 278	GNGGNNTII	SEQ ID NO: 716	GNAGDNTIV	SEQ ID NO: 1154
GDAGNNTIY	SEQ ID NO: 279	GDGGNNTII	SEQ ID NO: 717	GDAGDNTIV	SEQ ID NO: 1155
GGSGNNTIY	SEQ ID NO: 280	GGDGNTII	SEQ ID NO: 718	GGSGDNTIV	SEQ ID NO: 1156
GNSGNNTIY	SEQ ID NO: 281	GNDGNNTII	SEQ ID NO: 719	GNSGDNTIV	SEQ ID NO: 1157
GDSGNNTIY	SEQ ID NO: 282	GDDGNNTII	SEQ ID NO: 720	GDSGDNTIV	SEQ ID NO: 1158
GGGGNNTIY	SEQ ID NO: 283	GGAGDNTII	SEQ ID NO: 721	GGGGDNTIV	SEQ ID NO: 1159
GNGGNNTIY	SEQ ID NO: 284	GNAGDNTII	SEQ ID NO: 722	GNGGDNTIV	SEQ ID NO: 1160
GDGGNNTIY	SEQ ID NO: 285	GDAGDNTII	SEQ ID NO: 723	GDGGDNTIV	SEQ ID NO: 1161
GGDGNTIY	SEQ ID NO: 286	GGSGDNTII	SEQ ID NO: 724	GGDGNTIV	SEQ ID NO: 1162
GNDGNNTIY	SEQ ID NO: 287	GNSGDNTII	SEQ ID NO: 725	GNDGDNTIV	SEQ ID NO: 1163

Sequence (SEQ ID NO)		Sequence (SEQ ID NO)		Sequence (SEQ ID NO)	
GDDGNNTIY	SEQ ID NO: 288	GDSGDNTII	SEQ ID NO: 726	GDDGDNTIV	SEQ ID NO: 1164
GGAGDNTIY	SEQ ID NO: 289	GGGGDNTII	SEQ ID NO: 727	GGAGANTIV	SEQ ID NO: 1165
GNAGDNTIY	SEQ ID NO: 290	GNGGDNTII	SEQ ID NO: 728	GNAGANTIV	SEQ ID NO: 1166
GDAGDNTIY	SEQ ID NO: 291	GDGGDNTII	SEQ ID NO: 729	GDAGANTIV	SEQ ID NO: 1167
GGSGDNTIY	SEQ ID NO: 292	GGDGDNTII	SEQ ID NO: 730	GGSGANTIV	SEQ ID NO: 1168
GNSGDNTIY	SEQ ID NO: 293	GNDGDNTII	SEQ ID NO: 731	GNSGANTIV	SEQ ID NO: 1169
GDSGDNTIY	SEQ ID NO: 294	GDDGDNTII	SEQ ID NO: 732	GDSGANTIV	SEQ ID NO: 1170
GGGGDNTIY	SEQ ID NO: 295	GGAGANTII	SEQ ID NO: 733	GGGGANTIV	SEQ ID NO: 1171
GNGGDNTIY	SEQ ID NO: 296	GNAGANTII	SEQ ID NO: 734	GNGGANTIV	SEQ ID NO: 1172
GDGGDNTIY	SEQ ID NO: 297	GDAGANTII	SEQ ID NO: 735	GDGGANTIV	SEQ ID NO: 1173
GGDGDNTIY	SEQ ID NO: 298	GGSGANTII	SEQ ID NO: 736	GGDGANTIV	SEQ ID NO: 1174
GNDGDNTIY	SEQ ID NO: 299	GNSGANTII	SEQ ID NO: 737	GNDGANTIV	SEQ ID NO: 1175
GDDGDNTIY	SEQ ID NO: 300	GDSGANTII	SEQ ID NO: 738	GDDGANTIV	SEQ ID NO: 1176
GGAGANTIY	SEQ ID NO: 301	GGGGANTII	SEQ ID NO: 739	GGAGNDIIV	SEQ ID NO: 1177
GNAGANTIY	SEQ ID NO: 302	GNGGANTII	SEQ ID NO: 740	GNAGNDIIV	SEQ ID NO: 1178
GDAGANTIY	SEQ ID NO: 303	GDGGANTII	SEQ ID NO: 741	GDAGNDIIV	SEQ ID NO: 1179
GGSGANTIY	SEQ ID NO: 304	GGDGANTII	SEQ ID NO: 742	GGSGNDIIV	SEQ ID NO: 1180
GNSGANTIY	SEQ ID NO: 305	GNDGANTII	SEQ ID NO: 743	GNSGNDIIV	SEQ ID NO: 1181
GDSGANTIY	SEQ ID NO: 306	GDDGANTII	SEQ ID NO: 744	GDSGNDIIV	SEQ ID NO: 1182
GGGGANTIY	SEQ ID NO: 307	GGAGNDIII	SEQ ID NO: 745	GGGGNDIIV	SEQ ID NO: 1183
GNGGANTIY	SEQ ID NO: 308	GNAGNDIII	SEQ ID NO: 746	GNGGNDIIV	SEQ ID NO: 1184
GDGGANTIY	SEQ ID NO: 309	GDAGNDIII	SEQ ID NO: 747	GDGGNDIIV	SEQ ID NO: 1185
GGDGANTIY	SEQ ID NO: 310	GGSGNDIII	SEQ ID NO: 748	GGDGNDIIV	SEQ ID NO: 1186
GNDGANTIY	SEQ ID NO: 311	GNSGNDIII	SEQ ID NO: 749	GNDGNDIIV	SEQ ID NO: 1187
GDDGANTIY	SEQ ID NO: 312	GDSGNDIII	SEQ ID NO: 750	GDDGNDIIV	SEQ ID NO: 1188
GGAGNDIY	SEQ ID NO: 313	GGGGNDIII	SEQ ID NO: 751	GGAGDDIIV	SEQ ID NO: 1189
GNAGNDIY	SEQ ID NO: 314	GNGGNDIII	SEQ ID NO: 752	GNAGDDIIV	SEQ ID NO: 1190
GDAGNDIY	SEQ ID NO: 315	GDGGNDIII	SEQ ID NO: 753	GDAGDDIIV	SEQ ID NO: 1191
GGSGNDIY	SEQ ID NO: 316	GGDGDNDIII	SEQ ID NO: 754	GGSGDDIIV	SEQ ID NO: 1192
GNSGNDIY	SEQ ID NO: 317	GNDGNDIII	SEQ ID NO: 755	GNSGDDIIV	SEQ ID NO: 1193
GDSGNDIY	SEQ ID NO: 318	GDDGNDIII	SEQ ID NO: 756	GDSGDDIIV	SEQ ID NO: 1194
GGGGNDIY	SEQ ID NO: 319	GGAGDDIII	SEQ ID NO: 757	GGGGDDIIV	SEQ ID NO: 1195
GNGGNDIY	SEQ ID NO: 320	GNAGDDIII	SEQ ID NO: 758	GNGGDDIIV	SEQ ID NO: 1196
GDGGNDIY	SEQ ID NO: 321	GDAGDDIII	SEQ ID NO: 759	GDGGDDIIV	SEQ ID NO: 1197
GGDGDNDIY	SEQ ID NO: 322	GGSGDDIII	SEQ ID NO: 760	GGDGDDIIV	SEQ ID NO: 1198
GNDGDNDIY	SEQ ID NO: 323	GNSGDDIII	SEQ ID NO: 761	GNDGDDIIV	SEQ ID NO: 1199
GDDGDNDIY	SEQ ID NO: 324	GDSGDDIII	SEQ ID NO: 762	GDDGDDIIV	SEQ ID NO: 1200
GGAGDDIY	SEQ ID NO: 325	GGGGDDIII	SEQ ID NO: 763	GGAGADIIV	SEQ ID NO: 1201
GNAGDDIY	SEQ ID NO: 326	GNGGDDIII	SEQ ID NO: 764	GNAGADIIV	SEQ ID NO: 1202
GDAGDDIY	SEQ ID NO: 327	GDGGDDIII	SEQ ID NO: 765	GDAGADIIV	SEQ ID NO: 1203
GGSGDDIY	SEQ ID NO: 328	GGDGDDIII	SEQ ID NO: 766	GGSGADIIV	SEQ ID NO: 1204
GNSGDDIY	SEQ ID NO: 329	GNDGDDIII	SEQ ID NO: 767	GNSGADIIV	SEQ ID NO: 1205
GDSGDDIY	SEQ ID NO: 330	GDDGDDIII	SEQ ID NO: 768	GDSGADIIV	SEQ ID NO: 1206
GGGGDDIY	SEQ ID NO: 331	GGAGADIII	SEQ ID NO: 769	GGGGADIIV	SEQ ID NO: 1207
GNGGDDIY	SEQ ID NO: 332	GNAGADIII	SEQ ID NO: 770	GNGGADIIV	SEQ ID NO: 1208
GDGGDDIY	SEQ ID NO: 333	GDAGADIII	SEQ ID NO: 771	GDGGADIIV	SEQ ID NO: 1209
GGDGDDIY	SEQ ID NO: 334	GGSGADIII	SEQ ID NO: 772	GGDGADIIV	SEQ ID NO: 1210
GNDGDDIY	SEQ ID NO: 335	GNSGADIII	SEQ ID NO: 773	GNDGADIIV	SEQ ID NO: 1211
GDDGDDIY	SEQ ID NO: 336	GDSGADIII	SEQ ID NO: 774	GDDGADIIV	SEQ ID NO: 1212
GGAGADIY	SEQ ID NO: 337	GGGGADIII	SEQ ID NO: 775	GGAGNNIIV	SEQ ID NO: 1213
GNAGADIY	SEQ ID NO: 338	GNGGADIII	SEQ ID NO: 776	GNAGNNIIV	SEQ ID NO: 1214
GDAGADIY	SEQ ID NO: 339	GDGGADIII	SEQ ID NO: 777	GDAGNNIIV	SEQ ID NO: 1215
GGSGADIY	SEQ ID NO: 340	GGDGADIII	SEQ ID NO: 778	GGSGNNIIV	SEQ ID NO: 1216

Sequence (SEQ ID NO)		Sequence (SEQ ID NO)		Sequence (SEQ ID NO)	
GNSGADIY	SEQ ID NO: 341	GNDGADIII	SEQ ID NO: 779	GNSGNIIIV	SEQ ID NO: 1217
GDSGADIY	SEQ ID NO: 342	GDDGADIII	SEQ ID NO: 780	GDSGNIIIV	SEQ ID NO: 1218
GGGGADIY	SEQ ID NO: 343	GGAGNNIII	SEQ ID NO: 781	GGGGNIIIV	SEQ ID NO: 1219
GNGGADIY	SEQ ID NO: 344	GNAGNNIII	SEQ ID NO: 782	GNGGNIIIV	SEQ ID NO: 1220
GDGGADIY	SEQ ID NO: 345	GDAGNNIII	SEQ ID NO: 783	GDGGNIIIV	SEQ ID NO: 1221
GGDGADIY	SEQ ID NO: 346	GGSGNNIII	SEQ ID NO: 784	GGDGNNIIV	SEQ ID NO: 1222
GNDGADIY	SEQ ID NO: 347	GNSGNNIII	SEQ ID NO: 785	GNDGNNIIV	SEQ ID NO: 1223
GDDGADIY	SEQ ID NO: 348	GDSGNNIII	SEQ ID NO: 786	GDDGNNIIV	SEQ ID NO: 1224
GGAGNNIY	SEQ ID NO: 349	GGGGNNIII	SEQ ID NO: 787	GGAGDNIIV	SEQ ID NO: 1225
GNAGNNIY	SEQ ID NO: 350	GNGGNNIII	SEQ ID NO: 788	GNAGDNIIV	SEQ ID NO: 1226
GDAGNNIY	SEQ ID NO: 351	GDGGNNIII	SEQ ID NO: 789	GDAGDNIIV	SEQ ID NO: 1227
GGSGNNIY	SEQ ID NO: 352	GGDGNNIII	SEQ ID NO: 790	GGSGDNIIV	SEQ ID NO: 1228
GNSGNNIY	SEQ ID NO: 353	GNDGNNIII	SEQ ID NO: 791	GNSGDNIIV	SEQ ID NO: 1229
GDSGNNIY	SEQ ID NO: 354	GDDGNNIII	SEQ ID NO: 792	GDSGDNIIV	SEQ ID NO: 1230
GGGGNNIY	SEQ ID NO: 355	GGAGDNI	SEQ ID NO: 793	GGGGDNIIV	SEQ ID NO: 1231
GNGGNNIY	SEQ ID NO: 356	GNAGDNI	SEQ ID NO: 794	GNGGDNIIV	SEQ ID NO: 1232
GDGGNNIY	SEQ ID NO: 357	GDAGDNI	SEQ ID NO: 795	GDGGDNIIV	SEQ ID NO: 1233
GGDGNNIY	SEQ ID NO: 358	GGSGDNI	SEQ ID NO: 796	GGDGdNIIV	SEQ ID NO: 1234
GNDGNNIY	SEQ ID NO: 359	GNSGDNI	SEQ ID NO: 797	GNDGDNIIV	SEQ ID NO: 1235
GDDGNNIY	SEQ ID NO: 360	GDSGDNI	SEQ ID NO: 798	GDDGDNIIV	SEQ ID NO: 1236
GGAGDNIY	SEQ ID NO: 361	GGGGDNI	SEQ ID NO: 799	GGAGANIIV	SEQ ID NO: 1237
GNAGDNIY	SEQ ID NO: 362	GNGGDNI	SEQ ID NO: 800	GNAGANIIV	SEQ ID NO: 1238
GDAGDNIY	SEQ ID NO: 363	GDGGDNI	SEQ ID NO: 801	GDAGANIIV	SEQ ID NO: 1239
GGSGDNIY	SEQ ID NO: 364	GGDGdNI	SEQ ID NO: 802	GGSGANIIV	SEQ ID NO: 1240
GNSGDNIY	SEQ ID NO: 365	GNDGDNI	SEQ ID NO: 803	GNSGANIIV	SEQ ID NO: 1241
GDSGDNIY	SEQ ID NO: 366	GDDGDNI	SEQ ID NO: 804	GDSGANIIV	SEQ ID NO: 1242
GGGGDNIY	SEQ ID NO: 367	GGAGANIII	SEQ ID NO: 805	GGGGANIIV	SEQ ID NO: 1243
GNGGDNIY	SEQ ID NO: 368	GNAGANIII	SEQ ID NO: 806	GNGGANIIV	SEQ ID NO: 1244
GDGGDNIY	SEQ ID NO: 369	GDAGANIII	SEQ ID NO: 807	GDGGANIIV	SEQ ID NO: 1245
GGDGdNIY	SEQ ID NO: 370	GGSGANIII	SEQ ID NO: 808	GGDGANIIV	SEQ ID NO: 1246
GNDGDNIY	SEQ ID NO: 371	GNSGANIII	SEQ ID NO: 809	GNDGANIIV	SEQ ID NO: 1247
GDDGDNIY	SEQ ID NO: 372	GDSGANIII	SEQ ID NO: 810	GDDGANIIV	SEQ ID NO: 1248
GGAGANIY	SEQ ID NO: 373	GGGGANIII	SEQ ID NO: 811	GGAGNDVIV	SEQ ID NO: 1249
GNAGANIY	SEQ ID NO: 374	GNGGANIII	SEQ ID NO: 812	GNAGNDVIV	SEQ ID NO: 1250
GDAGANIY	SEQ ID NO: 375	GDGGANIII	SEQ ID NO: 813	GDAGNDVIV	SEQ ID NO: 1251
GGSGANIY	SEQ ID NO: 376	GGDGANIII	SEQ ID NO: 814	GGSGNDVIV	SEQ ID NO: 1252
GNSGANIY	SEQ ID NO: 377	GNDGANIII	SEQ ID NO: 815	GNSGNDVIV	SEQ ID NO: 1253
GDSGANIY	SEQ ID NO: 378	GDDGANIII	SEQ ID NO: 816	GDSGNDVIV	SEQ ID NO: 1254
GGGGANIY	SEQ ID NO: 379	GGAGNDVII	SEQ ID NO: 817	GGGGNDVIV	SEQ ID NO: 1255
GNGGANIY	SEQ ID NO: 380	GNAGNDVII	SEQ ID NO: 818	GNGGNDVIV	SEQ ID NO: 1256
GDGGANIY	SEQ ID NO: 381	GDAGNDVII	SEQ ID NO: 819	GDGGNDVIV	SEQ ID NO: 1257
GGDGANIY	SEQ ID NO: 382	GGSGNDVII	SEQ ID NO: 820	GGDGNDVIV	SEQ ID NO: 1258
GNDGANIY	SEQ ID NO: 383	GNSGNDVII	SEQ ID NO: 821	GNDGNDVIV	SEQ ID NO: 1259
GDDGANIY	SEQ ID NO: 384	GDSGNDVII	SEQ ID NO: 822	GDDGNDVIV	SEQ ID NO: 1260
GGAGNDVIY	SEQ ID NO: 385	GGGGNDVII	SEQ ID NO: 823	GGAGDDVIV	SEQ ID NO: 1261
GNAGNDVIY	SEQ ID NO: 386	GNGGNDVII	SEQ ID NO: 824	GNAGDDVIV	SEQ ID NO: 1262
GDAGNDVIY	SEQ ID NO: 387	GDGGNDVII	SEQ ID NO: 825	GDAGDDVIV	SEQ ID NO: 1263
GGSGNDVIY	SEQ ID NO: 388	GGDGNDVII	SEQ ID NO: 826	GGSGDDVIV	SEQ ID NO: 1264
GNSGNDVIY	SEQ ID NO: 389	GNDGNDVII	SEQ ID NO: 827	GNSGDDVIV	SEQ ID NO: 1265
GDSGNDVIY	SEQ ID NO: 390	GDDGNDVII	SEQ ID NO: 828	GDSGDDVIV	SEQ ID NO: 1266
GGGGNDVIY	SEQ ID NO: 391	GGAGDDVII	SEQ ID NO: 829	GGGGDDVIV	SEQ ID NO: 1267
GNGGNDVIY	SEQ ID NO: 392	GNAGDDVII	SEQ ID NO: 830	GNGGDDVIV	SEQ ID NO: 1268
GDGGNDVIY	SEQ ID NO: 393	GDAGDDVII	SEQ ID NO: 831	GDGGDDVIV	SEQ ID NO: 1269

Sequence (SEQ ID NO)		Sequence (SEQ ID NO)		Sequence (SEQ ID NO)	
GGDGNDVIY	SEQ ID NO: 394	GGSGDDVII	SEQ ID NO: 832	GGDGDDVIV	SEQ ID NO: 1270
GNDGNDVIY	SEQ ID NO: 395	GNSGDDVII	SEQ ID NO: 833	GNDGDDVIV	SEQ ID NO: 1271
GDDGNDVIY	SEQ ID NO: 396	GDSGDDVII	SEQ ID NO: 834	GDDGDDVIV	SEQ ID NO: 1272
GGAGDDVIY	SEQ ID NO: 397	GGGGDDVII	SEQ ID NO: 835	GGAGADVIV	SEQ ID NO: 1273
GNAGDDVIY	SEQ ID NO: 398	GNGGDDVII	SEQ ID NO: 836	GNAGADVIV	SEQ ID NO: 1274
GDAGDDVIY	SEQ ID NO: 399	GDGGDDVII	SEQ ID NO: 837	GDAGADVIV	SEQ ID NO: 1275
GGSGDDVIY	SEQ ID NO: 400	GGDGDDVII	SEQ ID NO: 838	GGSGADVIV	SEQ ID NO: 1276
GNSGDDVIY	SEQ ID NO: 401	GNDGDDVII	SEQ ID NO: 839	GNSGADVIV	SEQ ID NO: 1277
GDSGDDVIY	SEQ ID NO: 402	GDDGDDVII	SEQ ID NO: 840	GDSGADVIV	SEQ ID NO: 1278
GGGGDDVIY	SEQ ID NO: 403	GGAGADVII	SEQ ID NO: 841	GGGGADVIV	SEQ ID NO: 1279
GNGGDDVIY	SEQ ID NO: 404	GNAGADVII	SEQ ID NO: 842	GNGGADVIV	SEQ ID NO: 1280
GDGGDDVIY	SEQ ID NO: 405	GDAGADVII	SEQ ID NO: 843	GDGGADVIV	SEQ ID NO: 1281
GGDGDDVIY	SEQ ID NO: 406	GGSGADVII	SEQ ID NO: 844	GGDGADVIV	SEQ ID NO: 1282
GNDGDDVIY	SEQ ID NO: 407	GNSGADVII	SEQ ID NO: 845	GNDGADVIV	SEQ ID NO: 1283
GDDGDDVIY	SEQ ID NO: 408	GDSGADVII	SEQ ID NO: 846	GDDGADVIV	SEQ ID NO: 1284
GGAGADVIY	SEQ ID NO: 409	GGGGADVII	SEQ ID NO: 847	GGAGNNVIV	SEQ ID NO: 1285
GNAGADVIY	SEQ ID NO: 410	GNGGADVII	SEQ ID NO: 848	GNAGNNVIV	SEQ ID NO: 1286
GDAGADVIY	SEQ ID NO: 411	GDGGADVII	SEQ ID NO: 849	GDAGNNVIV	SEQ ID NO: 1287
GGSGADVIY	SEQ ID NO: 412	GGDGADVII	SEQ ID NO: 850	GGSGNNVIV	SEQ ID NO: 1288
GNSGADVIY	SEQ ID NO: 413	GNDGADVII	SEQ ID NO: 851	GNSGNNVIV	SEQ ID NO: 1289
GDSGADVIY	SEQ ID NO: 414	GDDGADVII	SEQ ID NO: 852	GDSGNNVIV	SEQ ID NO: 1290
GGGGADVIY	SEQ ID NO: 415	GGAGNNVII	SEQ ID NO: 853	GGGGNNVIV	SEQ ID NO: 1291
GNGGADVIY	SEQ ID NO: 416	GNAGNNVII	SEQ ID NO: 854	GNGGNNVIV	SEQ ID NO: 1292
GDGGADVIY	SEQ ID NO: 417	GDAGNNVII	SEQ ID NO: 855	GDGGNNVIV	SEQ ID NO: 1293
GGDGADVIY	SEQ ID NO: 418	GGSGNNVII	SEQ ID NO: 856	GGDGNNVIV	SEQ ID NO: 1294
GNDGADVIY	SEQ ID NO: 419	GNSGNNVII	SEQ ID NO: 857	GNDGNNVIV	SEQ ID NO: 1295
GDDGADVIY	SEQ ID NO: 420	GDSGNNVII	SEQ ID NO: 858	GDDGNNVIV	SEQ ID NO: 1296
GGAGNNVIY	SEQ ID NO: 421	GGGGNNVII	SEQ ID NO: 859	GGAGDNVIV	SEQ ID NO: 1297
GNAGNNVIY	SEQ ID NO: 422	GNGGNNVII	SEQ ID NO: 860	GNAGDNVIV	SEQ ID NO: 1298
GDAGNNVIY	SEQ ID NO: 423	GDGGNNVII	SEQ ID NO: 861	GDAGDNVIV	SEQ ID NO: 1299
GGSGNNVIY	SEQ ID NO: 424	GGDGNNVII	SEQ ID NO: 862	GGSGDNVIV	SEQ ID NO: 1300
GNSGNNVIY	SEQ ID NO: 425	GNDGNNVII	SEQ ID NO: 863	GNSGDNVIV	SEQ ID NO: 1301
GDSGNNVIY	SEQ ID NO: 426	GDDGNNVII	SEQ ID NO: 864	GDSGDNVIV	SEQ ID NO: 1302
GGGGNNVIY	SEQ ID NO: 427	GGAGDNVII	SEQ ID NO: 865	GGGGDNVIV	SEQ ID NO: 1303
GNGGNNVIY	SEQ ID NO: 428	GNAGDNVII	SEQ ID NO: 866	GNGGDNVIV	SEQ ID NO: 1304
GDGGNNVIY	SEQ ID NO: 429	GDAGDNVII	SEQ ID NO: 867	GDGGDNVIV	SEQ ID NO: 1305
GGDGNNVIY	SEQ ID NO: 430	GGSGDNVII	SEQ ID NO: 868	GGDGDNVIV	SEQ ID NO: 1306
GNDGNNVIY	SEQ ID NO: 431	GNSGDNVII	SEQ ID NO: 869	GNDGDNVIV	SEQ ID NO: 1307
GDDGNNVIY	SEQ ID NO: 432	GDSGDNVII	SEQ ID NO: 870	GDDGDNVIV	SEQ ID NO: 1308
GGAGDNVIY	SEQ ID NO: 433	GGGGDNVII	SEQ ID NO: 871	GGAGANVIV	SEQ ID NO: 1309
GNAGDNVIY	SEQ ID NO: 434	GNGGDNVII	SEQ ID NO: 872	GNAGANVIV	SEQ ID NO: 1310
GDAGDNVIY	SEQ ID NO: 435	GDGGDNVII	SEQ ID NO: 873	GDAGANVIV	SEQ ID NO: 1311
GGSGDNVIY	SEQ ID NO: 436	GGDGDNVII	SEQ ID NO: 874	GGSGANVIV	SEQ ID NO: 1312
GNSGDNVIY	SEQ ID NO: 437	GNDGDNVII	SEQ ID NO: 875	GNSGANVIV	SEQ ID NO: 1313
GDSGDNVIY	SEQ ID NO: 438	GDDGDNVII	SEQ ID NO: 876	GDSGANVIV	SEQ ID NO: 1314
GGGGDNVIY	SEQ ID NO: 439	GGAGANVII	SEQ ID NO: 877	GGGGANVIV	SEQ ID NO: 1315
GNGGDNVIY	SEQ ID NO: 440	GNAGANVII	SEQ ID NO: 878	GNGGANVIV	SEQ ID NO: 1316
GDGGDNVIY	SEQ ID NO: 441	GDAGANVII	SEQ ID NO: 879	GDGGANVIV	SEQ ID NO: 1317
GGDGDNVIY	SEQ ID NO: 442	GGSGANVII	SEQ ID NO: 880	GGDGANVIV	SEQ ID NO: 1318
GNDGDNVIY	SEQ ID NO: 443	GNSGANVII	SEQ ID NO: 881	GNDGANVIV	SEQ ID NO: 1319
GDDGDNVIY	SEQ ID NO: 444	GDSGANVII	SEQ ID NO: 882	GDDGANVIV	SEQ ID NO: 1320
GGAGANVIY	SEQ ID NO: 445	GGGGANVII	SEQ ID NO: 883	GDEASDLFF	SEQ ID NO: 1321
GNAGANVIY	SEQ ID NO: 446	GNGGANVII	SEQ ID NO: 884	GDLASDLFF	SEQ ID NO: 1322

Sequence (SEQ ID NO)		Sequence (SEQ ID NO)		Sequence (SEQ ID NO)	
GDAGANVIY	SEQ ID NO: 447	GDGGANVII	SEQ ID NO: 885	GDNASDLFF	SEQ ID NO: 1323
GGSGANVIY	SEQ ID NO: 448	GGDGANVII	SEQ ID NO: 886	GDEASDLFT	SEQ ID NO: 1324
GNSGANVIY	SEQ ID NO: 449	GNDGANVII	SEQ ID NO: 887	GDLASDLFT	SEQ ID NO: 1325
GDSGANVIY	SEQ ID NO: 450	GDDGANVII	SEQ ID NO: 888	GDNASDLFT	SEQ ID NO: 1326
GGGGANVIY	SEQ ID NO: 451	GGAGNDTLV	SEQ ID NO: 889	GDEASDLFN	SEQ ID NO: 1327
GNGGANVIY	SEQ ID NO: 452	GNAGNDTLV	SEQ ID NO: 890	GDLASDLFN	SEQ ID NO: 1328
GDGGANVIY	SEQ ID NO: 453	GDAGNDTLV	SEQ ID NO: 891	GDNASDLFN	SEQ ID NO: 1329
GGDGANVIY	SEQ ID NO: 454	GGSGNDTLV	SEQ ID NO: 892	GDEASDLFD	SEQ ID NO: 1330
GNDGANVIY	SEQ ID NO: 455	GNSGNDTLV	SEQ ID NO: 893	GDLASDLFD	SEQ ID NO: 1331
GDDGANVIY	SEQ ID NO: 456	GDSGNDTLV	SEQ ID NO: 894	GDNASDLFD	SEQ ID NO: 1332
GGAGNDTLI	SEQ ID NO: 457	GGGGNDTLV	SEQ ID NO: 895	GDEASDLFK	SEQ ID NO: 1333
GNAGNDTLI	SEQ ID NO: 458	GNGGNDTLV	SEQ ID NO: 896	GDLASDLFK	SEQ ID NO: 1334
GDAGNDTLI	SEQ ID NO: 459	GDGGNDTLV	SEQ ID NO: 897	GDNASDLFK	SEQ ID NO: 1335
GGSGNDTLI	SEQ ID NO: 460	GGDGNDTLV	SEQ ID NO: 898	GDEASDLFS	SEQ ID NO: 1336
GNSGNDTLI	SEQ ID NO: 461	GNDGNDTLV	SEQ ID NO: 899	GDLASDLFS	SEQ ID NO: 1337
GDSGNDTLI	SEQ ID NO: 462	GDDGNDTLV	SEQ ID NO: 900		

[0043] As used herein, the term “precipatable-beta roll cassette” (PBRC) refers to an amino acid sequence comprising at least one PBRT. In certain embodiments, a PBRC will comprise at least two PBRTs. In certain embodiments, a PBRC will comprise at least 3 PBRTs, at least 4 PBRTs, at least 5 PBRTs, at least 6 PBRTs, at least 7 PBRTs, at least 8 PBRTs, at least 9 PBRTs, at least 10 PBRTs, at least 11 PBRTs, at least 12 PBRTs, at least 13 PBRTs, at least 14 PBRTs, at least 15 PBRTs, at least 16 PBRTs, at least 17 PBRTs, at least 18 PBRTs, at least 19 PBRTs, at least 20 PBRTs, or 20 or more PBRTs. In certain embodiments, the PBRCs described herein will comprise a plurality of precipatable beta roll tags arranged in a tandem repeat. For example, in certain embodiments, the PBRCs described herein can comprise at least 2 PBRTs, at least 3 PBRTs, at least 4 PBRTs, at least 5 PBRTs, at least 6 PBRTs, at least 7 PBRTs, at least 8 PBRTs, at least 9 PBRTs, at least 10 PBRTs, at least 11 PBRTs, at least 12 PBRTs, at least 13 PBRTs, at least 14 PBRTs, at least 15 PBRTs, at least 16 PBRTs, at least 17 PBRTs, at least 18 PBRTs, at least 19 PBRTs, at least 20 PBRTs, or 20 or more PBRTs in tandem repeat. In certain embodiments, a PBRC can comprise at least two PBRCs separated by a linking amino acid sequence. Where a linking amino acid sequence is present between two PBRTs, a PBRT located at either end of the linking sequence can be an individual PBRT or it can be a PBRT that is part of a tandem arrangement of two or more PBRTs.

[0044] The PBRCs can comprise polymeric or oligomeric repeats of a PBRT. In certain embodiments, the PBRCs described herein can comprise one or more different PBRTs. In one embodiment, all of the PBRTs comprised in a PBRC are identical in amino acid

sequence. In one embodiment, all of the PBRTs comprised in a PBRC have different amino acid sequences. In one embodiment, at least one PBRT comprised in a PBRC has a different amino acid sequence as compared to another PBRT in the PBRC. Thus, in certain embodiments, the PBRCs described herein can comprise at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, or at least 30 different PBRTs.

[0045] In certain embodiments, a PBRC can also comprise a capping sequence (“CS”) refers to the an amino acid sequence comprising the amino acid sequence of SEQ ID NO: 3. The term “capping sequence” also refers to a capping sequence having the amino acid sequence of any of SEQ ID NO: 4-23.

[0046] Without wishing to be bound to theory, in some embodiments of the invention, the ability of polypeptide comprising one or more PBRT to undergo reversible Ca^{2+} precipitation, can require that the one or more PBRTs be located N-terminally or C-terminally to a capping sequence. Thus, in certain embodiments, the capping sequence is an amino acid sequence, which, when located C-terminally or N-terminally to one or more PBRTs, allows the one or more PBRTs bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the PBRC.

[0047] Thus in certain embodiments, a PBRC will comprise at least 2 PBRTs, at least 3 PBRTs, at least 4 PBRTs, at least 5 PBRTs, at least 6 PBRTs, at least 7 PBRTs, at least 8 PBRTs, at least 9 PBRTs, at least 10 PBRTs, at least 11 PBRTs, at least 12 PBRTs, at least 13 PBRTs, at least 14 PBRTs, at least 15 PBRTs, at least 16 PBRTs, at least 17 PBRTs, at least 18 PBRTs, at least 19 PBRTs, at least 20 PBRTs, or 20 or more PBRTs, all of which are located N-terminally to a CS. In certain embodiments, a PBRC will comprise 2, at least 3 PBRTs, at least 4 PBRTs, at least 5 PBRTs, at least 6 PBRTs, at least 7 PBRTs, at least 8 PBRTs, at least 9 PBRTs, at least 10 PBRTs, at least 11 PBRTs, at least 12 PBRTs, at least 13 PBRTs, at least 14 PBRTs, at least 15 PBRTs, at least 16 PBRTs, at least 17 PBRTs, at least 18 PBRTs, at least 19 PBRTs, at least 20 PBRTs, or 20 or more PBRTs, all of which are located C-terminally to a CS.

[0048] In one embodiment, a PBRC is an amino acid sequence comprising at least five tandem PBRTs situated N-terminally to a capping sequence.

[0049] In one embodiment, a PBRC is an amino acid sequence comprising at least six to about 16 tandem PBRTs situated N-terminally to a capping sequence.

[0050] In one embodiment, a precipitable beta roll cassette is an amino acid sequence comprising 17 or more tandem PBRTs situated N-terminally to a capping sequence.

[0051] In one embodiment, the capping sequence comprises the sequence INAGADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIHAANQAVDQAGIEKLVEAMAQYPD (SEQ ID NO: 3) which is a C-terminal sequence on the block V beta roll domain of the adenylate cyclase toxin of *B. pertussis*.

[0052] In another embodiment, a PBRC comprises the amino acid sequence GGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYINAGADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIHAANQAVDQAGIEKLVEAMAQYPD (SEQ ID NO: 4). In another embodiment, a PBRC comprises the amino acid sequence GGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYINAGADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIHAANQAVDQAGIEKLVEAMAQYPD (SEQ ID NO: 5). In another embodiment, the capping sequence comprises the sequence

INAGADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIHAANQAIDPAGIEKLVEAMAQYPD (SEQ ID NO: 6) (adenylate cyclase-hemolysin [*Bordetella bronchiseptica*]). In another embodiment, the capping sequence comprises the sequence INAGADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIHAANQAIDPAGIEKLVEAMAQYPD (SEQ ID NO: 7) (adenylate cyclase-hemolysin [*Bordetella bronchiseptica*]). In another embodiment, the capping sequence comprises the sequence INAGADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIHAANQTVDPAGIEKLVEAMAQYPD (SEQ ID NO: 8) (adenylate cyclase hemolysin [*Bordetella bronchiseptica*]). In another embodiment, the capping sequence comprises the sequence QLWFSKSGSDLEVRVVGTDDAVTVAGWYSGAEHHMDSIETADGTVLLDSMVDRLVQAMAGF (SEQ ID NO: 9) (*Azospirillum* sp. B510 calcium binding hemolysin protein). In another embodiment, the capping sequence comprises the sequence

ADQLWFRHVGNDLEISILGTGDTATVRDWYLGSR YQIEQIRVDDGRTL VNADVEKL
VQAMA (SEQ ID NO: 10) (hemolysin-type calcium-binding region *Burkholderia*
cenocepacia MC0-3). In another embodiment, the capping sequence comprises the sequence
ADQLWFRHVGNDLEISILGSSDTATVRDWYSGSR YQIEQIRLDDGRTL VNADVEKL
VQAMA (SEQ ID NO: 11) (hemolysin-type calcium-binding region [*Burkholderia ambifaria*
MC40-6]). In another embodiment, the capping sequence comprises the sequence
DARQTNLWFSQVGKDLQIDVLGSTDQVTVKDWYAGADNRVERIKTADGKTL YDSD
VDKLVQAMASF (SEQ ID NO: 12) (calcium binding secreted hemolysin
protein)[*Herbaspirillum seropedicae* SmR1]). In another embodiment, the capping sequence
comprises the sequence
DARQTNLWFSQVGKDLQIDVLGSTDQVTVKDWYAGADNRVERIKTADGKTL YDSD
VDKLVQAMASF (SEQ ID NO: 13) (calcium binding secreted hemolysin protein
[*Herbaspirillum seropedicae* SmR1]). In another embodiment, the capping sequence
comprises the sequence
ELWFSRENNDLIKSLLSEDKVTVQNWYSHQDHKIENIRLSNEQTLVSTQVEKMVES
MAGF (SEQ ID NO: 14) (RTX toxin protein [*Actinobacillus pleuropneumoniae* serovar]). In
another embodiment, the capping sequence comprises the sequence
EELWFSRDGNDLQINVIGTDNQVEISDWYSGVNYQLDKVQVGDSVLLNTQLEQLVS
AMASF (SEQ ID NO: 15) (hemolysin-type calcium binding protein [*Shewanella*
piezotolerans WP3]). In another embodiment, the capping sequence comprises the sequence
GLSELWFSRENNDLIKSLLSEDKVTVQNWYSHQDHKIENIRLSNEQMLVSTQVEKM
VESMAGF (SEQ ID NO: 16) (RTX toxin protein [*Actinobacillus pleuropneumoniae* serovar
10]). In another embodiment, the capping sequence comprises the sequence
EDLWFSRDGNNLQINIIGTDDQVEVNNWYNDTNYQLDQIQVGGSVLLNNQLEQLVS
AMASF (SEQ ID NO: 17) (RTX toxin [*Shewanella violacea* DSS12]). In another
embodiment, the capping sequence comprises the sequence
ELWFSRENNDLIKSLLSEDKVTVQNWYSHQDHKIENIRLSNEQTLVSTQVEKMVES
MASF (SEQ ID NO: 18) (RTX toxin protein [*Actinobacillus pleuropneumoniae* serovar 6]).
In another embodiment, the capping sequence comprises the sequence
ADNFWFVKSGNDLEIDILGTHQQVTVADWFLGGSYQLQEIKAGGLELDTQVTQLVQ
AMATY (SEQ ID NO: 19) (protein BRADO6535 [*Bradyrhizobium* sp. ORS278]). In
another embodiment, the capping sequence comprises the sequence
ELWFSRENNDLIKSLLSEDKVTVQNWYSHQDHKIENIRLSNEQTLVSTQVEKMVES

MAGF (SEQ ID NO: 20) ([*Actinobacillus pleuropneumoniae* L20]). In another embodiment, the capping sequence comprises the sequence

LWFSRENNDLIKSLLSEDKVTVQNWYSHQDHKIENIRLSNEQTLVSTQVEKMVESM

AGF (SEQ ID NO: 21) (ApxIVA [*Actinobacillus pleuropneumoniae*]). In another embodiment, the capping sequence comprises the sequence

LWFRKSGNNLEVSIIGTSDKLVMSNWyAGSQYQVERFQAGDGKALQANQVQSLVQ

AMASF (SEQ ID NO: 22) (hemolysin-type calcium-binding protein [*Xanthomonas axonopodis* pv. *citri* str. 306]). In another embodiment, the capping sequence comprises the sequence

ELWFSRENNDLIKSLLSEDKVTVQNWYSHQDHKIENIRLSNEQTLVSTQVEKMVES

MAGF (SEQ ID NO: 23) (RTX toxin IVA [*Actinobacillus pleuropneumoniae*])

[0053] Without wishing to be bound to theory, in some embodiments of the invention, the ability of polypeptide comprising one or more PBRT to undergo reversible Ca^{2+} precipitation, can require that the one or more PBRTs be located N-terminally or C-terminally to a stabilizing polypeptide. It is known that when a certain stabilizing polypeptides (e.g. GFP, maltose binding protein) are attached to the C-terminus of the beta roll, calcium-induced folding can occur (Blenner et al., *Journal of Molecular Biology*. 400 (2010), pp 244-256; Szilvay et al., *Biochemistry*, 48 (2009), pp 11273-11282).

[0054] Thus, in one embodiment, a PBRC is an amino acid sequence comprising one or more PBRTs located N-terminally or C-terminally to a stabilizing polypeptide, wherein the stabilizing polypeptide can be, but is not limited to, glutathione S-transferase (GST), maltose E binding protein (MBP), Green Fluorescent Protein (GFP), and variants thereof. In still a further embodiment, the stabilizing polypeptide is an amino acid sequence of any amino acid composition wherein the sequence comprises at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, at least 45, at least 46, at least 47, at least 48, at least 49, at least 50, at least 51, at least 52, at least 53, at least 54, at least 55, at least 56, at least 57, at least 58, at least 59, at least 60, at least 61, at least 62, at least 63, at least 64, at least 65, at least 66, at least 67, or at least 68 amino acids.

[0055] In another aspect of the invention, the PBRCs described herein can further comprise one or more cleavage sites. In certain embodiments, the cleavage site can be positioned C-terminally or N-terminally to a PBRC so as to allow for cleavage of a PBRC from a linked purification moiety (e.g. a polypeptide purification moiety linked to a PBRC as part of a fusion protein). In embodiments where the PBRC comprises more than one cleavage site, a first cleavage site can be positioned C-terminally or N-terminally to a PBRC so as to allow for cleavage of a PBRC from a linked purification moiety (e.g. a polypeptide purification moiety linked to a PBRC as part of a fusion protein) and a second cleavage site can be positioned between a PBRT and a capping sequence or a PBRT and a stabilizing polypeptide so as to allow so as to allow for cleavage of the capping sequence or the stabilizing polypeptide from the one or more PBRTs in the PBRC. In certain embodiments, cleavage at such cleavage sites can be useful for purification of a purification moiety of interest.

[0056] In one embodiment, the cleavage site is a proteolytic cleavage site. Exemplary proteolytic cleavage sites, include, but are not limited to Factor Xa, thrombin, or enterokinase. In another embodiment, the cleavage site is a signal peptidase cleavage site. In another embodiment, the cleavage site is a self cleaving intein cleavage site (Amitai et al., *Proceedings of the National Academy of Sciences*, vol. 106, no. 27, pp. 11005 -11010, Jul. 2009; Hiraga et al., *Journal of Molecular Biology*, vol. 393, no. 5, pp. 1106-1117, Nov. 2009). Any other specific cleavage sites known in the art can be used in connection with the methods described herein.

[0057] The PBRTs or PBRCs described herein may be linked to a purification moiety by any means known in the art. The PBRTs or PBRCs described herein can be located at any site in a polypeptide comprising a purification moiety of interest, including a location that is N-terminal, a location that is C-terminal or a location within the sequence of the purification moiety of interest.

[0058] In addition to fusion proteins comprising the PBRTs or PBRCs described herein, the PBRTs or PBRCs described herein can also be chemically linked to purification moieties other than by means of a fusion protein. Thus, reference to a PBRC linked purification moiety, or to a PBRT linked purification moiety encompasses for purification moieties linked to a PBRC or PBRT by peptide linkage (e.g. as a fusion protein) or by non-peptide bond chemical linkage.

[0059] The chemical modification of PBRTs or PBRCs described herein can be performed according to any method known in the art.. For example, amides of the PBRTs or PBRCs described herein can be prepared by techniques well known in the art for converting a carboxylic acid group or precursor, to an amide. One method for amide formation at the C-terminal carboxyl group is to cleave the polypeptide, or fusion thereof from a solid support with an appropriate amine, or to cleave in the presence of an alcohol, yielding an ester, followed by aminolysis with the desired amine.

[0060] Salts of carboxyl groups of the PBRTs or PBRCs described herein can be prepared by contacting the polypeptide, or fusion thereof with one or more equivalents of a desired base such as, for example, a metallic hydroxide base, e.g., sodium hydroxide; a metal carbonate or bicarbonate base such as, for example, sodium carbonate or sodium bicarbonate; or an amine base such as, for example, triethylamine, triethanolamine, and the like.

[0061] N-acyl derivatives of an amino group of the PBRTs or PBRCs described herein can be prepared by utilizing an N-acyl protected amino acid for the final condensation, or by acylating a protected or unprotected polypeptide, or fusion thereof. O-acyl derivatives can be prepared, for example, by acylation of a free hydroxy polypeptide or polypeptide resin. Either acylation can be carried out using standard acylating reagents such as acyl halides, anhydrides, acyl imidazoles, and the like. Both N- and O-acylation can be carried out together, if desired.

[0062] Formyl-methionine, pyroglutamine and trimethyl-alanine can be substituted at the N-terminal residue of PBRTs or PBRCs described herein. Other amino-terminal modifications include aminooxypentane modifications.

[0063] Such chemical linkages can be useful for purifying non-peptide molecules such as lipids, oligonucleotides and carbohydrates, small organic or inorganic molecules, proteins, single-stranded or double-stranded oligonucleotides, polynucleotides, metals (e.g. cobalt, zinc, nickel or copper) and the like. The chemically modified PBRTs or PBRCs described herein can be assayed for the ability to bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the PBRT or PBRC using methods known to those skilled in the art.

[0064] The PBRTs and PBRCs described herein can also be coupled with a radioisotope or enzymatic label to facilitate their detection. For example, the PBRTs or PBRCs described herein can be isotopically-labeled where one or more atoms are replaced or substituted by an

atom having an atomic mass or mass number different from the atomic mass or mass number typically found in nature (i.e., naturally occurring). Suitable radionuclides that may be incorporated in compounds of the present invention include but are not limited to ^2H (also written as D for deuterium), ^3H (also written as T for tritium), ^{11}C , ^{13}C , ^{14}C , ^{13}N , ^{15}N , ^{15}O , ^{17}O , ^{18}O , ^{18}F , ^{35}S , ^{36}Cl , ^{82}Br , ^{75}Br , ^{76}Br , ^{77}Br , ^{123}I , ^{124}I , ^{125}I and ^{131}I . The radionuclide that is incorporated in the instant radio-labeled compounds can depend on the specific application of that radio-labeled compound.

[0065] Alternatively, the PBRTs or PBRCs described herein can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In another embodiment, the PBRTs or PBRCs described herein can be labeled with a fluorescent dye, spin label, heavy metal or radio-labeled peptides.

[0066] Esters of carboxyl groups of the PBRTs or PBRCs described herein can also be prepared by any of the usual methods known in the art.

[0067] The methods and compositions described herein are useful in a broad range of bioseparation applications. The methods and compositions described herein can be used for rapid expression and purification of a purification moiety linked to a PBRC. Such PBRC linked purification moieties can be expressed in any number of expression systems, including in vitro and in vivo expression systems. Exemplary in vivo expression systems suitable for expressing the PBRC linked purification moieties described herein, include, but are not limited to, bacterial systems, yeast systems, and mammalian systems.

[0068] In one aspect, the invention relates to a method for purifying purification moieties (e.g. a PBRC or a purification moiety linked to a PBRC). In one aspect, the methods described herein can be used for purifying one or more purification moieties from a heterogeneous mixture of biomaterials in a sample. In another aspect, the methods described herein can be used to purify chemically synthesized purification moieties or in-vitro synthesized purification moieties.

[0069] In certain embodiments, the bioseparation methods described herein can comprise expressing a PBRT linked purification moiety (e.g. a PBRC fusion protein) in a cellular expression system (e.g. a bacterial cell). The PBRC linked purification moiety can then be released into a medium by cell lysis. Any method of cell lysis known in the art can be used in conjunction with the methods described herein, including, but not limited to chemical lysis

(e.g. detergents) or physical methods (e.g. sonication or French press). In certain embodiments, the PBRC or PBRC linked purification moiety can be expressed in an in-vitro expression system (e.g. a rabbit reticulocyte system) such that the purification moiety is expressed into the expression system medium.

[0070] After expression of the PBRC linked purification moiety, bioseparation can be achieved by increasing the free Ca^{2+} concentration in the medium comprising the PBRC linked purification moiety to induce precipitation of the PBRC linked purification moiety, followed by removing material that does not precipitate from the medium and then resuspending the precipitated material in a medium having a reduced Ca^{2+} concentration or in a medium having a reduced concentration of free Ca^{2+} (e.g. a medium comprising a Ca^{2+} chelators such as EDTA). These steps can be repeated until the desired level of purity is reached.

[0071] In another embodiment, the PBRC linked purification moiety can be expressed in a cellular expression system, and bioseparation can be achieved by increasing the free Ca^{2+} concentration within the cell prior to cellular lysis. Many methods for increasing intracellular Ca^{2+} concentrations are known in the art, including, but not limited to adding Ca^{2+} to a cellular medium, with or without presence of ionophores or cell permeabilization agents. In such embodiments, the cells can then be subjected to lysis conditions (e.g. chemical lysis or physical lysis) and the resulting precipitate can be recovered. Precipitated PBRC or the PBRC linked purification moieties can then be recovered by reducing the free Ca^{2+} concentration (e.g. through the addition of a Ca^{2+} chelator) to induce solubilization of the purification moieties from the precipitate and bioseparation can be achieved by eliminating the precipitate. Precipitation and solubilization steps can be repeated until a desired level of purity is reached.

[0072] In still other embodiments, the PBRC linked purification moieties described herein can further comprise a peptide sequence to induce secretion into the periplasm of a cell (e.g. an E. coli cell) or to the medium outside of a cell. Where the PBRC linked purification moiety is secreted to the periplasm of a cell, cell lysis may be required for further purification of the PBRC linked purification moiety. Where the PBRC linked purification moiety is secreted into the medium outside of the cell, purification can be achieved without cell lysis by eliminating intact cells (e.g. by centrifugation) and purification of the PBRC linked purification moiety from the extracellular medium by increasing the free Ca^{2+} concentration of the supernatant.

[0073] The adjustment of conditions during the purification process can be achieved by numerous methods, including, but not limited to, adjusting the temperature, pH or salt concentration of the aqueous media.

[0074] The methods described herein can be used to purify purification moieties of any size. A purified PBRT linked purification moiety can contain less than about 50%, less than about 75%, or less than about 90%, of the materials with which it was originally associated.

[0075] In one embodiment, a purification moiety of interest can be linked to PBRC comprising a cleavable peptide sequence (e.g. a self-cleaving peptide sequence such as an intein) positioned between the PBRC and the purification moiety. Once the PBRC linked purification moiety is expressed in an expression system, the PBRC linked purification moiety can be recovered using standard techniques as either a homogenous mixture or as a heterogeneous sample. The mixture can then be exposed to calcium to induce precipitation of the PBRC linked purification moiety. The PBRC linked purification moiety can then be resuspended in buffer that has reduced Ca^{2+} , that has reduced free Ca^{2+} or that contains a calcium chelator (e.g. EDTA). The PBRC linked purification moiety can then be subjected to a condition that causes cleavage to separate the PBRC from the purification moiety and calcium can be once again added to the mixture. This will precipitate out the PBRC moiety and thereby leaving behind a sample of purified purification moiety of interest.

[0076] In another embodiment, PBRC linked purification moiety can be a moiety which binds a second molecule and the second molecule can be used to remove the purification moiety from the sample (e.g. a resin or beads coated with the second molecule) after induced cleavage at a site between the purification moiety and the PBRC.

[0077] In some embodiments, immobilization of the PBRC linked purification moieties described herein or its binding proteins can be used to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Immobilization of the PBRC linked purification moieties described herein can be by linking to a solid support, including a plastic or glass plate or bead, a chromatographic resin, a filter or a membrane. Methods of attachment of proteins, or membranes containing same, to such supports are well known in the art. Immobilization of the PBRC linked purification moieties described herein can also be accomplished in any vessel suitable for containing the reactants. Examples include microtiter plates, test tubes, and micro-centrifuge tubes.

[0078] In one embodiment, a fusion protein can be provided which adds a domain that allows the PBRT linked purification moiety described herein to be bound to a matrix. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads or glutathione derivatized microtiter plates, which are then combined with the cell lysates, and the mixture incubated under conditions conducive to complex formation.. Following incubation, the beads can be washed to remove any unbound fraction,. Alternatively, the complexes can be dissociated from the matrix using standard electrophoretic techniques.

[0079] The methods described herein depend, in part on the finding that PBRCs undergo a reversible Ca^{2+} binding dependent transition. PBRCs or PBRC linked purification moieties undergo reversible precipitation at a Ca^{2+} concentration (or free Ca^{2+}) phase transition concentration.

[0080] The transition concentrations reversible and the isolated precipitable beta-roll tags or purification moieties comprising a precipitable beta-roll tag can be completely resolubilized in a medium below a certain Ca^{2+} concentration (or free Ca^{2+}) transition concentration, through, for example the addition of a calcium chelator into the medium comprising the PBRC or PBRC linked purification moiety.

[0081] The concentration of Ca^{2+} required to induce reversible precipitation of a PBRC or PBRC linked purification moiety can be readily determined by adding increasing amounts of Ca^{2+} until such time as the PBRC or PBRC linked purification moiety begins to precipitate from a the medium. One can readily determine the extent of precipitation by centrifuging the medium. In one embodiment, the amount of Ca^{2+} required to induce reversible precipitation a PBRC of PBRC linked purification moiety will be about 1mM Ca^{2+} , more than about 1mM Ca^{2+} , more than about 5mM Ca^{2+} , more than about 10mM Ca^{2+} , more than about 20mM Ca^{2+} , more than about 30mM Ca^{2+} , more than about 50mM Ca^{2+} , more than about 75mM Ca^{2+} , more than about 100 mM Ca^{2+} , more than about 150mM Ca^{2+} , more than about 200mM Ca^{2+} , or more than about 500mM Ca^{2+} . In certain embodiments, the amount of Ca^{2+} required to induce precipitation of a PBRC or a PBRC linked purification moiety can increase as a function of the number of PBRTs in the PBRC. For example, a PBRC linked purification moiety comprising 8 PBRTs may precipitate in 150mM Ca^{2+} wherein a PBRC linked purification moiety comprising 17 PBRTs may precipitate in 25mM Ca^{2+} . One of skill in the art will readily be capable of determining the amount of Ca^{2+} required to precipitate a particular PBRC or a particular PBRC linked purification moiety simply by titrating increasing concentrations of Ca^{2+} .

[0082] The concentration of Ca^{2+} required to reverse precipitation of a PBRT or PBRC or of a PBRT or PBRC linked purification moiety can be readily determined by reducing the concentration of free Ca^{2+} in a medium until such time as a precipitated PBRC or PBRC linked purification moiety begins to solubilize into the medium. One can readily determine the extent of precipitation by centrifuging the medium. In one embodiment, the amount of free Ca^{2+} in the medium required to solubilize a precipitated PBRT or PBRC or of a PBRT or PBRC linked purification moiety will be less than about 1mM Ca^{2+} , less than about 1mM Ca^{2+} , less than about 5mM Ca^{2+} , less than about 10mM Ca^{2+} , less than about 20mM Ca^{2+} , less than about 30mM Ca^{2+} , less than about 50mM Ca^{2+} , less than about 75mM Ca^{2+} , less than about 100 mM Ca^{2+} , less than about 150mM Ca^{2+} , less than about 200mM Ca^{2+} , or less than about 500mM Ca^{2+} . In certain embodiments, the free Ca^{2+} concentration required to reverse precipitation of a PBRC or a PBRC linked purification moiety can correlated to the number of PBRTs in the PBRC. For example, a PBRC linked purification moiety comprising 8 PBRTs may become soluble in a higher free Ca^{2+} concentration as compared to a PBRC linked purification moiety comprising 17 PBRTs. One of skill in the art will readily be capable of determining the Ca^{2+} concentration required to reverse precipitation a particular PBRC or a particular PBRC linked purification moiety simply by decreasing free Ca^{2+} concentrations.

[0083] The free Ca^{2+} concentration of a medium comprising a PBRC or a PBRC linked purification moiety can be reduced by adding one or more calcium chelators into the medium. Any number of calcium chelators can be used in the connection with the methods described herein. Examples of suitable calcium chelators include, but are not limited to EDTA, EGTA, and BAPTA. In one embodiment, the amount of a calcium chelator required to solubilize a precipitated PBRT or PBRC or of a PBRT or PBRC linked purification moiety will be about 1mM Ca^{2+} , more than about 1mM Ca^{2+} , more than about 5mM Ca^{2+} , more than about 10mM Ca^{2+} , more than about 20mM Ca^{2+} , more than about 30mM Ca^{2+} , more than about 50mM Ca^{2+} , more than about 75mM Ca^{2+} , more than about 100 mM Ca^{2+} , more than about 150mM Ca^{2+} , more than about 200mM Ca^{2+} , or more than about 500mM Ca^{2+} .

[0084] In addition to temperature and ionic strength, other environmental variables useful for modulating the solubility of PBRT or PBRC or of the PBRT linked purification moieties described herein include pH, the addition of organic solutes and solvents, side-chain ionization or chemical modification, and pressure.

[0085] The PBRC linked purification moieties described herein can be further purified or isolated according to any method of protein purification or isolation known in the art. For example, PBRCs or PBRC linked purification moieties can be purified by various methods including, without limitation, preparative disc-gel electrophoresis, isoelectric focusing, HPLC, reversed-phase HPLC, gel filtration, ion exchange and partition chromatography, precipitation and salting-out chromatography, extraction, and countercurrent distribution. For some purposes, the PBRCs or the PBRC linked purification moieties can be produced in a recombinant system in which the protein contains an additional sequence tag that facilitates purification, such as, but not limited to, a polyhistidine sequence, or a sequence that specifically binds to an antibody, such as FLAG and GST. In one embodiment, the PBRCs or the PBRC linked purification moieties can be purified from a crude lysate of the host cell by chromatography on an appropriate solid-phase matrix. Alternatively, antibodies produced against the PBRTs or PBRCs, or a PBRC linked purification moiety or against polypeptides derived therefrom can be used as purification reagents.

[0086] The methods and compositions described herein can be useful for the detection of a broad range of purification moieties in biosensing applications. For example, the methods and compositions described herein can be used for the separation of protein of interest from a sample for detection of bimolecular interactions. In one embodiment, if a PBRC is linked to an antibody, the PBRC linked antibody can be added to a sample. Ca^{2+} can then be added to the sample to induce precipitation of the antibody such that antigen that interact with, or form a complex with, the antibody also precipitate upon the addition of Ca^{2+} . The precipitate can then be collected and resuspended and the sample can be characterized. The presence and quantity of the target in the original sample, as well as any associated purification moieties, can be characterized and determined. Any antibodies, antibody fragments, antibody configurations, classes, or subclasses known in the art can be used in connection with the methods described herein. In another embodiment, if a PBRC is fused to a polypeptide capable of binding to a second purification moiety, it can be added to a sample. Ca^{2+} can then be added to the sample to induce precipitation of the polypeptide such that other purification moieties that interact with, or form a complex with, the polypeptide also precipitate upon the addition of Ca^{2+} . The precipitate can then be collected and resuspended and the sample can be characterized. The presence and quantity of the target in the original sample, as well as any associated additional purification moieties, can be characterized and determined.

[0087] The PBRC linked purification moieties described herein can be produced in prokaryotic or eukaryotic host cells by expression of nucleic acids encoding a polypeptide of this invention. The production of these polypeptides can also be done as part of a larger polypeptide.

[0088] The PBRC linked purification moieties described herein can also be synthesized in vitro, e.g., by the solid phase polypeptide synthetic method or by recombinant DNA approaches described herein. The solid phase polypeptide synthetic method is an established and widely used method. These PBRC or PBRC linked purification moieties described herein can be further purified by fractionation on immunoaffinity or ion-exchange columns; ethanol precipitation; reverse phase HPLC; chromatography on silica or on an anion-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; or ligand affinity chromatography.

[0089] The PBRC linked purification moieties described herein can also be produced using any in-vitro expression system known in the art or can be synthesized by chemical methods. Methods for expression of heterologous proteins in recombinant hosts, chemical synthesis of polypeptides, and in vitro translation are well known in the art and are described further in Sambrook J et al.; Berger and Kimmel, *Methods in Enzymology*, Volume 152, *Guide to Molecular Cloning Techniques* (1987), Academic Press, Inc., San Diego, Calif.; Gutte B and Merrifield R B, *J. Am. Chem. Soc.* 91:501-02 (1969); Chaiken I M, *CRC Crit. Rev. Biochem.* 11:255-301 (1981); Kaiser E T et al., *Science* 243:187-92 (1989); Merrifield B, *Science* 232:341-47 (1986); Kent S B H, *Ann. Rev. Biochem.* 57:957-89 (1988); Offord, R. E. (1980) *Semisynthetic Proteins*, Wiley Publishing. Exemplary peptide synthesis methods known in the art include, but are not limited to those described in Stewart et al., *Solid Phase Peptide Synthesis*, Pierce Biotechnology, Inc., Rockford, Ill., 1984; Bodanszky, *Principles of Peptide Synthesis*, Springer-Verlag, New York, 1984; and Pennington et al., *Peptide Synthesis Protocols*, Humana Press, Totowa, N.J., 1994). Additionally, many companies offer custom peptide synthesis services.

[0090] The PBRC linked purification moieties described herein can also be produced by direct chemical synthesis. For example, the PBRC linked purification moieties described herein can be produced as modified polypeptides, with nonpeptide moieties attached by covalent linkage to the N-terminus and/or C-terminus. In certain embodiments, either the carboxy-terminus or the amino-terminus, or both, are chemically modified. Common modifications of the terminal amino and carboxyl groups, include, but are not limited to

acetylation and amidation, respectively. Amino-terminal modifications such as acylation (e.g., acetylation) or alkylation (e.g., methylation) and carboxy-terminal-modifications such as amidation, as well as other terminal modifications, including cyclization, can be incorporated into various embodiments. Certain amino-terminal and/or carboxy-terminal modifications and/or polypeptide extensions to the core sequence can provide advantageous physical, chemical, biochemical, and pharmacological properties, such as: enhanced stability, increased potency and/or efficacy, resistance to serum proteases, desirable pharmacokinetic properties, and others.

[0091] The PBRC linked purification moieties can be prepared using recombinant DNA and molecular cloning techniques. Genes encoding the PBRC linked purification moieties may be produced in heterologous host cells, particularly in the cells of microbial hosts. Any techniques for transfecting host cells and purifying proteins and polypeptides known in the art can be used in connection with the methods described herein. Exemplary epitope tags suitable for use with the methods described herein include, but are not limited to FLAG, HA, Myc and T7 epitope tags. The PBRTs, PBRCs or PBRC linked purification moieties described herein can be synthesized chemically using standard polypeptide synthesis techniques.

[0092] The invention also extends to the DNA expression vector comprising DNA coding for the PBRTs or PBRCs described herein, whether or not the encoded products further comprise a linked purification moiety. The invention also provides the expression vector comprising sequences coding for a PBRT or a PBRC configured to allow insertion of a DNA sequence downstream of the sequence coding for the PBRT or the PBRC so as to facilitate production of a fusion protein comprising a PBRT or a PBRC. For example, such vectors can comprise one or more cloning sites between the sequence coding for the PBRT or the PBRC to enable generation of an in-frame translation product. Such vectors may comprise multiple cloning sites in any of three reading frames. Methods for generating such expression vectors are well known in the art.

[0093] A variety of expression systems can be used to produce the PBRCs and PBRC linked purification moieties described herein. Such expression systems include vector based expression systems. Exemplary vector base expression systems suitable for use with the methods described herein include, but are not limited to, chromosomal, episomal and virus-derived vectors, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from insertion elements, from yeast episomes, from viruses such as

baculoviruses, retroviruses and vectors derived from combinations thereof such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids.

[0094] The expression system vectors may contain regulatory regions that regulate as well as engender expression. In general, any system or vector suitable to maintain, propagate or express polynucleotide or polypeptide in a host cell may be used for expression in this regard. Expression systems and expression vectors can contain regulatory sequences that direct high level expression of foreign proteins relative to the growth of the host cell. Regulatory sequences are well known to those skilled in the art and examples include, but are not limited to, those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of regulatory elements in the vector, for example, enhancer sequences. Any of these could be used to construct chimeric genes for production of the any of the binding peptides of the present invention. These chimeric genes could then be introduced into appropriate microorganisms via transformation to provide high level expression of the peptides.

[0095] A number of recombinant expression vectors can be used for expression of the PBRCs and PBRC linked purification moieties described herein. For example, the PBRT linked purification moieties described herein can be expressed in bacterial cells such as *E. coli*, insect cells (e.g., using baculovirus expression vectors), yeast cells, amphibian cells, or mammalian cells. Suitable host cells are well known to one skilled in the art. Alternatively, the recombinant expression vector can be transcribed and translated in vitro, using, for example T7 promoter regulatory sequences and T7 polymerase.

[0096] Examples of *E. coli* expression vectors include pTrc (Amann E et al., *Gene* 69:301-15 (1988)) and pET 11d (Studier et al., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990) pp. 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

[0097] One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman S, *Gene Expression Technology: Methods in Enzymology* 185, Academic

Press, San Diego, Calif. (1990) pp. 119-28). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada K et al., *Nucleic Acids Res.* 20(Suppl.):2111-18 (1992)). Such alteration of nucleic acid sequences can be carried out by standard DNA synthesis techniques.

[0098] In another approach, a nucleic acid can be expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed B, *Nature* 329:840-41 (1987)) and pMT2PC (Kaufman R J et al., *EMBO J.* 6:187-95 (1987)). When used in mammalian cells, the expression vector's control functions can be provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus, and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells, see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

[0099] A number of these methodologies can also be applied in vivo, systemically or locally, in a complex biological system such as a human. For example, increased copy number of nucleic acids PBRC or PBRC linked purification moieties described herein in expressible from (by DNA transfection), can be employed.

[00100] Nucleic acid purification moieties encoding PBRT or PBRC linked purification moieties described herein can be administered to cells by a variety of methods known to those of skill in the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as biodegradable polymers, hydrogels, cyclodextrins (see for example, Gonzalez et al., *Bioconjugate Chem.* 10:1068-1074, 1999; Wang et al., International PCT Publication Nos. WO 03/47518 and WO 03/46185), poly(lactic-co-glycolic)ac-id (PLGA) and PLCA microspheres (see for example, U.S. Pat. No. 6,447,796 and U.S. Patent Application Publication No. US 2002130430), biodegradable nanocapsules, and bioadhesive microspheres, or by proteinaceous vectors (O'Hare and Normand, International PCT Publication No. WO 00/53722).

[00101] This invention may also be of use in the pharmaceutical/biotechnology industry where therapeutic compounds need to be purified in large quantities. This approach provides very pure product in a very quick manner.

[00102] Purification moieties that can be linked to the PBRTs or PBRCs described herein can be any purification moiety, including a biologically active protein (e.g., a therapeutic peptide, protein or an enzyme useful in industrial biocatalysis).

[00103] The purification moieties suitable for use with the methods described herein can be of widely varying types, including, for example, peptides, non-peptide proteins, lipids, oligonucleotides and carbohydrates, or alternatively a ligand-binding protein or an active fragment thereof having binding affinity to a molecule selected from the group consisting of small organic or inorganic molecules, proteins, peptides, single-stranded or double-stranded oligonucleotides, polynucleotides, lipids, and carbohydrates.

[00104] Suitable purification moieties include, but are not limited to, molecules useful in medicine, agriculture and other scientific and industrial fields. For example, suitable molecules include those of interest in medicine, agriculture or other scientific or industrial fields. Examples of suitable proteins include enzymes utilized in replacement therapy; hormones for promoting growth in animals, or cell growth in cell culture; and active proteinaceous substances used in various applications, e.g., in biotechnology or in medical diagnostics. One of skill in the art will recognize that many types of recombinant polypeptides can be produced using the methods described herein. The present invention is not limited to any specific types of recombinant polypeptide described herein. Instead, it encompasses any and all recombinant polypeptides.

[00105] The PBRTs or PBRCs described herein can be joined to a purification moiety from any source or origin and can include a polypeptide found in prokaryotes, viruses, and eukaryotes, including fungi, plants, yeasts, insects, and animals, including mammals (e.g. humans). Purification moieties suitable for use with the methods described herein include, but are not limited to any polypeptide sequences, known or hypothetical or unknown, which can be identified using common sequence repositories. Examples of such sequence repositories, include, but are not limited to GenBank EMBL, DDBJ and the NCBI. Other repositories can easily be identified by searching on the internet. Polypeptides that can be produced using the methods described herein also include polypeptides have at least about 60%, 70%, 75%, 80%, 90%, 95%, or at least about 99% or more identity to any known or available polypeptide (e.g., a therapeutic polypeptide, a diagnostic polypeptide, an industrial enzyme, or portion thereof, and the like).

[00106] Purification moieties suitable for use with the methods described herein include, but are not limited to, polypeptides comprising one or more non-natural amino acids.

[00107] Purification moieties suitable for use with the methods described herein include, but are not limited to, cytokines, inflammatory molecules, growth factors, their receptors, and oncogene products or portions thereof. Examples of cytokines, inflammatory molecules, growth factors, their receptors, and oncogene products include, but are not limited to e.g., alpha-1 antitrypsin, Angiostatin, Antihemolytic factor, antibodies (including an antibody or a functional fragment or derivative thereof selected from: Fab, Fab', F(ab)2, Fd, Fv, ScFv, diabody, tribody, tetrabody, dimer, trimer or minibody), angiogenic molecules, angiostatic molecules, Apolipopolypeptide, Apopolypeptide, Asparaginase, Adenosine deaminase, Atrial natriuretic factor, Atrial natriuretic polypeptide, Atrial peptides, Angiotensin family members, Bone Morphogenic Polypeptide (BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6, BMP-7, BMP-8a, BMP-8b, BMP-10, BMP-15, etc.); C--X--C chemokines (e.g., T39765, NAP-2, ENA-78, Gro-a, Gro-b, Gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG), Calcitonin, CC chemokines (e.g., Monocyte chemoattractant polypeptide-1, Monocyte chemoattractant polypeptide-2, Monocyte chemoattractant polypeptide-3, Monocyte inflammatory polypeptide-1 alpha, Monocyte inflammatory polypeptide-1 beta, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262), CD40 ligand, C-kit Ligand, Ciliary Neurotrophic Factor, Collagen, Colony stimulating factor (CSF), Complement factor 5a, Complement inhibitor, Complement receptor 1, cytokines, (e.g., epithelial Neutrophil Activating Peptide-78, GRO alpha/MGSA, GRO beta, GRO gamma, MIP-1 alpha, MIP-1 delta, MCP-1), deoxyribonucleic acids, Epidermal Growth Factor (EGF), Erythropoietin ("EPO", representing a preferred target for modification by the incorporation of one or more non-natural amino acid), Exfoliating toxins A and B, Factor IX, Factor VII, Factor VIII, Factor X, Fibroblast Growth Factor (FGF), Fibrinogen, Fibronectin, G-CSF, GM-CSF, Glucocerebrosidase, Gonadotropin, growth factors, Hedgehog polypeptides (e.g., Sonic, Indian, Desert), Hemoglobin, Hepatocyte Growth Factor (HGF), Hepatitis viruses, Hirudin, Human serum albumin, Hyalurin-CD44, Insulin, Insulin-like Growth Factor (IGF-I, IGF-II), interferons (e.g., interferon-alpha, interferon-beta, interferon-gamma, interferon-epsilon, interferon-zeta, interferon-eta, interferon-kappa, interferon-lambda, interferon-T, interferon-zeta, interferon-omega), glucagon-like peptide (GLP-1), GLP-2, GLP receptors, glucagon, other agonists of the GLP-1R, natriuretic peptides (ANP, BNP, and CNP), Fuzeon and other inhibitors of HIV fusion, Hirudin and related anticoagulant peptides, Prokineticins and

related agonists including analogs of black mamba snake venom, TRAIL, RANK ligand and its antagonists, calcitonin, amylin and other glucoregulatory peptide hormones, and Fc fragments, exendins (including exendin-4), exendin receptors, interleukins (e.g., IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, etc.), I-CAM-1/LFA-1, Keratinocyte Growth Factor (KGF), Lactoferrin, leukemia inhibitory factor, Luciferase, Neurturin, Neutrophil inhibitory factor (NIF), oncostatin M, Osteogenic polypeptide, Parathyroid hormone, PD-ECSF, PDGF, peptide hormones (e.g., Human Growth Hormone), Oncogene products (Mos, Rel, Ras, Raf, Met, etc.), Pleiotropin, Polypeptide A, Polypeptide G, Pyrogenic exotoxins A, B, and C, Relaxin, Renin, ribonucleic acids, SCF/c-kit, Signal transcriptional activators and suppressors (p53, Tat, Fos, Myc, Jun, Myb, etc.), Soluble complement receptor 1, Soluble I-CAM 1, Soluble interleukin receptors (IL-1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15), soluble adhesion molecules, Soluble TNF receptor, Somatomedin, Somatostatin, Somatotropin, Streptokinase, Superantigens, i.e., Staphylococcal enterotoxins (SEA, SEB, SEC1, SEC2, SEC3, SED, SEE), Steroid hormone receptors (such as those for estrogen, progesterone, testosterone, aldosterone, LDL receptor ligand and corticosterone), Superoxide dismutase (SOD), Toll-like receptors (such as Flagellin), Toxic shock syndrome toxin (TSST-1), Thymosin a 1, Tissue plasminogen activator, transforming growth factor (TGF- alpha, TGF- beta), Tumor necrosis factor beta (TNF beta), Tumor necrosis factor receptor (TNFR), Tumor necrosis factor- alpha (TNF alpha), transcriptional modulators (for example, genes and transcriptional modular polypeptides that regulate cell growth, differentiation and/or cell regulation), Vascular Endothelial Growth Factor (VEGF), virus-like particle, VLA-4/VCAM-1, Urokinase, signal transduction molecules, estrogen, progesterone, testosterone, aldosterone, LDL, corticosterone.

[00108] Additional purification moieties suitable for use with the methods described herein include, but are not limited to, enzymes (e.g., industrial enzymes) or portions thereof. Examples of enzymes include, but are not limited to amidases, amino acid racemases, acylases, dehalogenases, dioxygenases, diarylpropane peroxidases, epimerases, epoxide hydrolases, esterases, isomerases, kinases, glucose isomerases, glycosidases, glycosyl transferases, haloperoxidases, monooxygenases (e.g., p450s), lipases, lignin peroxidases, nitrile hydratases, nitrilases, proteases, phosphatases, subtilisins, transaminase, and nucleases. In certain embodiments, such enzymes comprising a PBRT or PBRC can be used as immobilized enzymes in industrial biocatalysis. The enzymes comprising a PBRTs or a

PBRC can also be added to a solution to facilitate biocatalysis and then reisolated from the solution.

[00109] Additional purification moieties suitable for use with the methods described herein include, but are not limited to, agriculturally related polypeptides such as insect resistance polypeptides (e.g., Cry polypeptides), starch and lipid production enzymes, plant and insect toxins, toxin-resistance polypeptides, Mycotoxin detoxification polypeptides, plant growth enzymes (e.g., Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase), lipoxygenase, and Phosphoenolpyruvate carboxylase.

[00110] Additional purification moieties suitable for use with the methods described herein include, but are not limited to, antibodies, immunoglobulin domains of antibodies and their fragments. Examples of antibodies include, but are not limited to antibodies, antibody fragments, antibody derivatives, Fab fragments, Fab' fragments, F(ab)₂ fragments, Fd fragments, Fv fragments, single-chain Fv fragments (scFv), diabodies, tribodies, tetrabodies, dimers, trimers, and minibodies.

[00111] Additional purification moieties suitable for use with the methods described herein include, but are not limited to, prophylactic vaccine or therapeutic vaccine polypeptides. A prophylactic vaccine is one administered to subjects who are not infected with a condition against which the vaccine is designed to protect. In certain embodiments, a preventive vaccine will prevent a virus from establishing an infection in a vaccinated subject. However, even if it does not provide complete protective immunity, a prophylactic vaccine may still confer some protection to a subject. For example, a prophylactic vaccine may decrease the symptoms, severity, and/or duration of the disease. A therapeutic vaccine, is administered to reduce the impact of a viral infection in subjects already infected with that virus. A therapeutic vaccine may decrease the symptoms, severity, and/or duration of the disease. Vaccine polypeptides include polypeptides, or polypeptide fragments from infectious fungi (e.g., *Aspergillus*, *Candida* species) bacteria (e.g. *E. coli*, *Staphylococci aureus*)), or *Streptococci* (e.g., *pneumoniae*); protozoa such as sporozoa (e.g., *Plasmodia*), rhizopods (e.g., *Entamoeba*) and flagellates (*Trypanosoma*, *Leishmania*, *Trichomonas*, *Giardia*, etc.); viruses such as (+) RNA viruses (examples include Poxviruses e.g., *vaccinia*; Picornaviruses, e.g., polio; Togaviruses, e.g., *rubella*; Flaviviruses, e.g., *HCV*; and Coronaviruses), (-) RNA viruses (e.g., *Rhabdoviruses*, e.g., *VSV*; *Paramyxoviruses*, e.g., *RSV*; *Orthomyxoviruses*, e.g., *influenza*; *Bunyaviruses*; and *Arenaviruses*), dsDNA viruses (*Reoviruses*, for example), RNA

to DNA viruses, i.e., Retroviruses, e.g., HIV and HTLV, and certain DNA to RNA viruses such as Hepatitis B.

[00112] Additional purification moieties suitable for use with the methods described herein include, but are not limited to, molecules that comprise a chemical moiety selected from the group consisting of: cytotoxins, pharmaceutical drugs, dyes or fluorescent labels, a nucleophilic or electrophilic group, a ketone or aldehyde, azide or alkyne compounds, photocaged groups, tags, a peptide, a polypeptide, a polypeptide, an oligosaccharide, polyethylene glycol with any molecular weight and in any geometry, polyvinyl alcohol, metals, metal complexes, polyamines, imidizoles, carbohydrates, lipids, biopolymers, particles, solid supports, a polymer, a targeting agent, an affinity group, any agent to which a complementary reactive chemical group can be attached, biophysical or biochemical probes, isotopically-labeled probes, spin-label amino acids, fluorophores, aryl iodides and bromides.

[00113] Reference is also made to a “variant PBRT.” A variant PBRT is a PBRT comprising one or more amino acid substitutions any position in the sequence of SEQ ID NO: 1 wherein the substitution replaces any amino acid in position 1 through 9 with an amino acid having a similar side chain group, an amino acid having a similar side chain configuration, an amino acid having an evolutionary positive relatedness, or an amino acid having an evolutionary neutral relatedness.

[00114] As used herein, the term “variant precipitable-beta roll cassette” (PBRC) refers to an amino acid sequence comprising at least one variant PBRT. In certain embodiments, a variant PBRC will comprise at least two variant PBRTs. In certain embodiments, a variant PBRC will comprise at least 3 variant PBRTs, at least 4 variant PBRTs, at least 5 variant PBRTs, at least 6 variant PBRTs, at least 7 variant PBRTs, at least 8 variant PBRTs, at least 9 variant PBRTs, at least 10 variant PBRTs, at least 11 variant PBRTs, at least 12 variant PBRTs, at least 13 variant PBRTs, at least 14 variant PBRTs, at least 15 variant PBRTs, at least 16 variant PBRTs, at least 17 variant PBRTs, at least 18 variant PBRTs, at least 19 variant PBRTs, at least 20 variant PBRTs, or 20 or more variant PBRTs. In certain embodiments, the PBRCs described herein will comprise a plurality of variant precipitable beta roll tags arranged in a tandem repeat. For example, in certain embodiments, the variant PBRCs described herein can comprise at least 2 variant PBRTs, at least 3 variant PBRTs, at least 4 variant PBRTs, at least 5 variant PBRTs, at least 6 variant PBRTs, at least 7 variant PBRTs, at least 8 variant PBRTs, at least 9 variant PBRTs, at least 10 variant PBRTs, at least 11 variant PBRTs, at least 12 variant PBRTs, at least 13 variant PBRTs, at least 14 variant

PBRTs, at least 15 variant PBRTs, at least 16 variant PBRTs, at least 17 variant PBRTs, at least 18 variant PBRTs, at least 19 variant PBRTs, at least 20 variant PBRTs, or 20 or more variant PBRTs in tandem repeat. In certain embodiments, a PBRC can comprise at least two PBRCs separated by a linking amino acid sequence. Where a linking amino acid sequence is present between two PBRTs, a PBRTs located at either end of the linking sequence can be an individual PBRT or it can be a PBRTs that is part of a tandem arrangement.

[00115] Thus in certain embodiments, a variant PBRC will comprise at least 2 variant PBRCs, at least 3 variant PBRTs, at least 4 variant PBRTs, at least 5 variant PBRTs, at least 6 variant PBRTs, at least 7 variant PBRTs, at least 8 variant PBRTs, at least 9 variant PBRTs, at least 10 variant PBRTs, at least 11 variant PBRTs, at least 12 variant PBRTs, at least 13 variant PBRTs, at least 14 variant PBRTs, at least 15 variant PBRTs, at least 16 variant PBRTs, at least 17 variant PBRTs, at least 18 variant PBRTs, at least 19 variant PBRTs, at least 20 variant PBRTs, or 20 or more variant PBRTs, all of which are located N-terminally to a CS. In certain embodiments, a variant PBRC will comprise at least 2 variant PBRCs, at least 3 variant PBRTs, at least 4 variant PBRTs, at least 5 variant PBRTs, at least 6 variant PBRTs, at least 7 variant PBRTs, at least 8 variant PBRTs, at least 9 variant PBRTs, at least 10 variant PBRTs, at least 11 variant PBRTs, at least 12 variant PBRTs, at least 13 variant PBRTs, at least 14 variant PBRTs, at least 15 variant PBRTs, at least 16 variant PBRTs, at least 17 variant PBRTs, at least 18 variant PBRTs, at least 19 variant PBRTs, at least 20 variant PBRTs, or 20 or more variant PBRTs, all of which are located C-terminally to a CS.

[00116] In certain aspects, the invention relates to a variant PBRT that contains one or more amino acid insertions, deletions or substitutions as compared to the sequence of SEQ ID NO: 1 and wherein the variant PBRT retains an ability to bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the PBRT.

[00117] Changes can be introduced by mutation into nucleic acid sequences, thereby leading to changes in the amino acid sequence of the encoded protein, without altering the functional activity of a PBRT or a PBRC. For example, nucleotide substitutions leading to amino acid substitutions at non-essential amino acid residues can be made in the sequence of a PBRT or a PBRC. A non-essential amino acid residue is a residue that can be altered from the sequence of an amino acid of this invention without altering the ability of the PBRT or PBRC to bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the PBRT or PBRC.

[00118] Exemplary residues which are non-essential and therefore amenable to substitution to generate the variant PBRTs and PBRCs described herein can be identified by one of ordinary skill in the art by performing an amino acid alignment of two more PBRTs or PBRCs and determining residues that are not required for the PBRT or PBRC to bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the PBRT or PBRC.

[00119] Mutations can be introduced randomly along all or part of a nucleic acid sequence encoding a PBRT or PBRC, such as by saturation mutagenesis, and the resultant mutants can be screened, for example, for their ability bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT or variant PBRC. Following mutagenesis, the purification moiety linked to the variant PBRT or variant PBRC can be expressed recombinantly in a host cell and the functional activity of the precipitable beta-roll tag can be determined using assays available in the art for assessing binding to Ca^{2+} , undergoing reversible precipitation in the presence of Ca^{2+} , or inducing reversible precipitation of purification moiety linked to the variant PBRT or variant PBRC. In certain embodiments, the variant PBRTs described herein can comprise one or more amino acid substitutions, insertions or deletions, wherein the variant PBRT is functionally equivalent a PBRT having the sequence GGAGNDTLY (SEQ ID No. 1). In one embodiment, the variant PBRT has an identical ability bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety in a manner similar to, but not necessarily identical to a PBRC comprising only PBRTs of the sequence GGAGNDTLY. In one embodiment, the variant PBRT has a reduced ability bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety in a manner similar to, but not necessarily identical to a PBRC comprising only PBRTs of the sequence GGAGNDTLY. In one embodiment, the variant PBRT has an increased ability bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety in a manner similar to, but not necessarily identical to a PBRC comprising only PBRTs of the sequence GGAGNDTLY.

[00120] The variant PBRCs described herein can also comprise one or more PBRTs in addition to one or more variant PBRTs. The variant PBRCs described herein can also be employed in any embodiments or configuration described herein for a PBRC. Thus, the description of a composition comprising a PBRC, or a method comprising a PBRC applies

equally to a variant PBRT or a variant PBRC so long as the variant PBRT or the variant PBRC can bind to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety in a manner similar to, but not necessarily identical to a PBRC comprising only PBRTs of the sequence GGAGNDTLY.

[00121] In one embodiment, a variant PBRT comprises the sequence GGXGXDXXX (SEQ ID NO: 2) wherein X can 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 one embodiment, the variant PBRT has a sequence of GGXGXDXXX (SEQ ID NO: 2), wherein X is not proline. In one embodiment, the variant PBRT comprises the sequence GGXGXDXXX (SEQ ID NO: 2) wherein X is a natural or non-natural amino acid comprising a modification.

[00122] In one embodiment, a variant PBRT or PBRC comprises an amino acid sequence having at least about 75%, 80%, 85%, 90%, 95%, 98%, 99% identity with an amino acid sequence of SEQ ID NO: 1.

[00123] As used herein, "sequence identity" means the percentage of identical nucleotide or amino acid residues at corresponding positions in two or more sequences when the sequences are aligned to maximize sequence matching, i.e., taking into account gaps and insertions. The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. Techniques for determining sequence identity are well known to one skilled in the art, and include, for example, analysis with a sequence comparison algorithm or FASTA version 3.0t78 using default parameters (Pearson and Lipman, Proc Natl Acad Sci U S A. 1988 Apr; 85(8):2444-8). In another non-limiting example, scoring of amino acid can be calculated using the PAM250 matrix as described in Dayhoff et al., (1978) in Atlas of Protein Sequence and Structure, ed. Dayhoff, M. (Natl. Biomed. Res. Found., Silver Spring, MD), Vol. 5, Suppl. 3, pp. 345-352.

[00124] Percent identity or percent similarity of a DNA or peptide sequence can be determined, for example, by comparing sequence information using the GAP computer program. The GAP program utilizes the alignment method of Needleman et al., 1970, as revised by Smith et al., 1981. Briefly, the GAP program defines similarity as the number of aligned symbols (i.e., nucleotides or amino acids) that are similar, divided by the total number of symbols in the shorter of the two sequences. The preferred parameters for the GAP

program are the default parameters, which do not impose a penalty for end gaps. See e.g., Schwartz et al., 1979; Gribskov et al., 1986. Nucleic acids that differ due to degeneracy of the genetic code, and still encode the PBRTs or PBRCs, described herein are encompassed by the present disclosure.

[00125] Variants can be produced by any number of methods, including but not limited to, error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly, and any combination thereof.

[00126] Variant PBRTs or variant PBRCs falling within the scope of this invention, can, in general, be generated by selecting substitutions that do not differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, (b) the charge or hydrophobicity of the purification moiety at the target site, or (c) the bulk of the side chain.

[00127] In one embodiment, a variant PBRT or a variant PBRC can comprise a conservative amino acid substitution in which an amino acid residue is replaced with an amino acid residue having a similar side chain configuration. Amino acid residues having similar side chain configurations have been defined in the art within in accordance with the following categories: basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine), aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine), aliphatic side chains (e.g., glycine, alanine, valine, leucine, isoleucine), and sulfur-containing side chains (methionine, cysteine). Substitutions can also be made between acidic amino acids and their respective amides (e.g., asparagine and aspartic acid, or glutamine and glutamic acid).

[00128] In one embodiment, a variant PBRT or a variant PBRC can comprise a conservative amino acid substitution in which an amino acid residue is replaced with an amino acid residue having a similar side chain group. Amino acid residues having similar side chain groups have been defined in the art within in accordance with the following categories: a no side chain group (glycine), an aliphatic side chain group (alanine, valine, leucine, isoleucine,

proline), a hydroxyl side chain group (serine, threonine), an acidic side chain group (aspartic acid, glutamic acid), an amide side chain group (asparagine, glutamine), a basic side chain group (lysine, arginine), an imidazole side chain group (histidine), an aromatic side chain group (phenylalanine, tyrosine, tryptophan), and a sulfur containing side chain group (methionine, cysteine) (see Sambrook et al, (2001) *Molecular Cloning: A Laboratory Manual*, Volume 3, Table A7-4).

[00129] In one embodiment, a variant PBRT or a variant PBRC can comprise a conservative amino acid substitution in which an amino acid residue is replaced an amino acid having evolutionarily positive relatedness. Amino acids having evolutionarily positive relatedness have been defined in the art as follows (wherein the amino acid(s) having evolutionarily positive relatedness are indicated in parentheses): Alanine (serine, threonine, proline, glycine); Arginine (glutamine, histidine, lysine, tryptophan); Asparagine (serine, threonine, aspartic acid, glutamic acid, glutamine, histidine, lysine); Aspartic acid (threonine, glycine, asparagine, glutamine, glutamic acid, histidine); Glutamic acid (threonine, asparagine, aspartic acid, glutamine, histidine); Glutamine (asparagine, aspartic acid, glutamic acid, histidine, arginine, lysine); Glycine (serine, threonine, alanine, aspartic acid); Histidine (asparagine, aspartic acid, glutamine, arginine); Isoleucine (threonine, methionine, leucine, valine, phenylalanine); Leucine (methionine, isoleucine, valine, phenylalanine); Lysine (threonine, asparagine, glutamine, arginine); Methionine (isoleucine, leucine, valine); Phenylalanine (isoleucine, leucine, tyrosine); Proline (serine, threonine, alanine); Serine (threonine, proline, alanine, glycine, asparagine); Threonine (serine, proline, alanine, glycine, asparagine, aspartic acid, glutamic acid, lysine, isoleucine, valine); Tryptophan (arginine, tyrosine); Tyrosine (phenylalanine, tryptophan); Valine (threonine, methionine, isoleucine, leucine) (see Dayhoff et al., (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M., Natl. Biomed. Res. Found., Silver Spring, MD), Vol. 5, Suppl. 3, pp. 345-352).

[00130] In one embodiment, variant of a variant PBRT or a variant PBRC can comprise a conservative amino acid substitution in which an amino acid residue is replaced an amino acid having evolutionarily positive relatedness. Amino acids having evolutionarily neutral relatedness have been defined in the art as follows (wherein the amino acid(s) having evolutionarily neutral relatedness are indicated in parentheses): Alanine (asparagine, aspartic acid, glutamine, glutamic acid, valine); Arginine (serine, proline, asparagine, methionine); Asparagine (alanine, glycine, arginine); Aspartic acid (serine, alanine, lysine); Cysteine (serine, tyrosine); Glutamic acid (serine, alanine, glycine, lysine); Glutamine (proline,

alanine); Glycine (asparagine, glutamic acid); Histidine (proline, lysine, tyrosine); Lysine (serine, asparagine, glutamic acid, histidine, methionine); Methionine (arginine, lysine, phenylalanine); Phenylalanine (methionine, tryptophan); Proline (glutamine, histidine, arginine); Serine (cysteine, aspartic acid, glutamic acid, arginine, lysine); Threonine (none); Tryptophan (phenylalanine); Tyrosine (cysteine, histidine); Valine (alanine) (see Dayhoff et al., (1978) in Atlas of Protein Sequence and Structure, ed. Dayhoff, M., Natl. Biomed. Res. Found., Silver Spring, MD), Vol. 5, Suppl. 3, pp. 345-352).

[00131] In one embodiment of a variant PBRT, the glycine at position 1 of SEQ ID NO: 1 is not mutated.

[00132] In another embodiment of a variant PBRT, the glycine at position 1 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the glycine is replaced with an amino acid having an uncharged polar side chain configuration (e.g., asparagine, glutamine, serine, threonine, tyrosine, or cysteine).

[00133] In another embodiment of a variant PBRT, the glycine at position 1 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the glycine is replaced with an amino acid having an aliphatic side chain configuration (e.g., alanine, valine, leucine, isoleucine)

[00134] In another embodiment of a variant PBRT, the glycine at position 1 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily positive relatedness such that the glycine is replaced with any of serine, threonine, alanine, or aspartic acid.

[00135] In another embodiment of a variant PBRT, the glycine at position 1 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily neutral relatedness such that the glycine is replaced with any of asparagine or glutamic acid.

[00136] In one embodiment of a variant PBRT, mutation of the glycine at position 1 of SEQ ID NO: 1 to any of alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine wherein mutation of the glycine at position 1 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00137] In one embodiment of a variant PBRT, mutation of the glycine at position 1 of SEQ ID NO: 1 is with a non-natural or synthetic amino acid wherein mutation of the glycine at position 1 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00138] In one embodiment of a variant PBRT, the glycine at position 2 of SEQ ID NO: 1 can be mutated to an asparagine residue.

[00139] In one embodiment of a variant PBRT, the glycine at position 2 of SEQ ID NO: 1 can be mutated to an aspartic acid residue.

[00140] In another embodiment of a variant PBRT, the glycine at position 2 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the glycine is replaced with an amino acid having an uncharged polar side chain configuration (e.g., asparagine, glutamine, serine, threonine, tyrosine, or cysteine).

[00141] In another embodiment of a variant PBRT, the glycine at position 2 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the glycine is replaced with an amino acid having an aliphatic side chain configuration (e.g., alanine, valine, leucine, isoleucine)

[00142] In another embodiment of a variant PBRT, the glycine at position 2 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily positive relatedness such that the glycine is replaced with any of serine, threonine, alanine, or aspartic acid.

[00143] In another embodiment of a variant PBRT, the glycine at position 2 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily neutral relatedness such that the glycine is replaced with any of asparagine or glutamic acid.

[00144] In one embodiment of a variant PBRT, mutation of the glycine at position 2 of SEQ ID NO: 1 to any of alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine wherein mutation of the glycine at position 2 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00145] In one embodiment of a variant PBRT, mutation of the glycine at position 2 of SEQ ID NO: 1 is with a non-natural or synthetic amino acid wherein mutation of the glycine at position 2 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00146] In one embodiment of a variant PBRT, the alanine at position 3 of SEQ ID NO: 1 can be mutated to a serine, glycine, or aspartic acid residue.

[00147] In one embodiment of a variant PBRT, the alanine at position 3 of SEQ ID NO: 1 can be mutated to a glutamic acid, leucine, or asparagine residue.

[00148] In one embodiment of a variant PBRT, the alanine at position 3 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the alanine is replaced with an amino acid having a nonpolar side chain configuration (e.g., valine, leucine, isoleucine, proline, phenylalanine, methionine, or tryptophan).

[00149] In one embodiment of a variant PBRT, the alanine at position 3 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the alanine is replaced with an amino acid having an aliphatic side chain configuration (e.g., glycine, valine, leucine, isoleucine)

[00150] In another embodiment of a variant PBRT, the alanine at position 3 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain group such that the alanine is replaced with an amino acid having an aliphatic side chain group (e.g., valine, leucine, isoleucine, proline).

[00151] In another embodiment of a variant PBRT, the alanine at position 3 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily positive relatedness such that the alanine is replaced with any of serine, threonine, proline, or glycine.

[00152] In another embodiment of a variant of a precipitable beta-roll tag, the at position 3 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily neutral relatedness such that the alanine is replaced with any of asparagine, aspartic acid, glutamine, glutamic acid, or valine.

[00153] In one embodiment of a variant PBRT, mutation of the alanine at position 3 of SEQ ID NO: 1 to any of glycine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine,

threonine, tryptophan, tyrosine, valine wherein mutation of the alanine at position 3 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00154] In one embodiment of a variant PBRT, mutation of the alanine at position 3 of SEQ ID NO: 1 is with a non-natural or synthetic amino acid wherein mutation of the alanine at position 3 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00155] In one embodiment of a variant PBRT, the glycine at position 4 of SEQ ID NO: 1 can be mutated to an alanine residue.

[00156] In another embodiment of a variant PBRT, the glycine at position 4 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the glycine is replaced with an amino acid having an uncharged polar side chain configuration (e.g., asparagine, glutamine, serine, threonine, tyrosine, or cysteine).

[00157] In another embodiment of a variant PBRT, the glycine at position 4 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the glycine is replaced with an amino acid having an aliphatic side chain configuration (e.g., alanine, valine, leucine, isoleucine)

[00158] In another embodiment of a variant PBRT, the glycine at position 4 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily positive relatedness such that the glycine is replaced with any of serine, threonine, alanine, or aspartic acid.

[00159] In another embodiment of a variant PBRT, the glycine at position 4 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily neutral relatedness such that the glycine is replaced with any of asparagine or glutamic acid.

[00160] In one embodiment of a variant PBRT, mutation of the glycine at position 4 of SEQ ID NO: 1 to any of alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine wherein mutation of the glycine at position 4 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo

reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00161] In one embodiment of a variant PBRT, mutation of the glycine at position 4 of SEQ ID NO: 1 is with a non-natural or synthetic amino acid wherein mutation of the glycine at position 4 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00162] In one embodiment of a variant PBRT, the asparagine at position 5 of SEQ ID NO: 1 can be mutated to an aspartic acid or alanine residue.

[00163] In one embodiment of a variant PBRT, the asparagine at position 5 of SEQ ID NO: 1 can be mutated to a serine residue.

[00164] In one embodiment of a variant PBRT, the asparagine at position 5 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the asparagine is replaced with an amino acid having an uncharged polar side chain configuration (e.g., glycine, glutamine, serine, threonine, tyrosine, cysteine)

[00165] In one embodiment of a variant PBRT, the asparagine at position 5 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the asparagine is replaced with an amino acid having a the side chain configuration of its amide (e.g., aspartic acid).

[00166] In another embodiment of a variant PBRT, the asparagine at position 5 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain group such that the asparagine is replaced with an amino acid having an amide side chain group (e.g., glutamine).

[00167] In another embodiment of a variant PBRT, the asparagine at position 5 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily positive relatedness such that the asparagine is replaced with any of serine, threonine, aspartic acid, glutamic acid, glutamine, histidine, or lysine.

[00168] In another embodiment of a variant PBRT, the asparagine at position 5 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily neutral relatedness such that the asparagine is replaced with any of alanine, glycine, or arginine.

[00169] In one embodiment of a variant PBRT, mutation of the asparagine at position 5 of SEQ ID NO: 1 to any of glycine, alanine, arginine, aspartic acid, cysteine, glutamic acid,

glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine wherein mutation of the asparagine at position 5 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00170] In one embodiment of a variant PBRT, mutation of the asparagine at position 5 of SEQ ID NO: 1 is with a non-natural or synthetic amino acid wherein mutation of the asparagine at position 5 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00171] In one embodiment of a variant PBRT, the aspartic acid at position 6 of SEQ ID NO: 1 can be mutated to an asparagine residue.

[00172] In one embodiment of a variant PBRT, the aspartic acid at position 6 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the aspartic acid is replaced with an amino acid having an acidic side chain configuration (e.g., glutamic acid).

[00173] In one embodiment of a variant PBRT, the aspartic acid at position 6 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the aspartic acid is replaced with an amino acid having a the side chain configuration of its amide (e.g., asparagine).

[00174] In another embodiment of a variant PBRT, the aspartic acid at position 6 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain group such that the aspartic acid is replaced with an amino acid having an acidic side chain group (e.g., glutamic acid).

[00175] In another embodiment of a variant PBRT, the aspartic acid at position 6 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily positive relatedness such that the aspartic acid is replaced with any of threonine, glycine, asparagine, glutamine, glutamic acid, or histidine.

[00176] In another embodiment of a variant PBRT, the aspartic acid at position 6 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily neutral relatedness such that the aspartic acid is replaced with any of serine, alanine, or lysine.

[00177] In one embodiment of a variant PBRT, mutation of the aspartic acid at position 6 of SEQ ID NO: 1 to any of glycine, alanine, arginine, asparagine, cysteine, glutamic acid, glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine wherein mutation of the aspartic acid at position 6 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00178] In one embodiment of a variant PBRT, mutation of the aspartic acid at position 6 of SEQ ID NO: 1 is with a non-natural or synthetic amino acid wherein mutation of the aspartic acid at position 6 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00179] In one embodiment of a variant PBRT, the threonine at position 7 of SEQ ID NO: 1 can be mutated to an isoleucine or valine residue.

[00180] In one embodiment of a variant PBRT, the threonine at position 7 of SEQ ID NO: 1 can be mutated to a leucine residue.

[00181] In one embodiment of a variant PBRT, the threonine at position 7 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the threonine is replaced with an amino acid having an uncharged polar side chain configuration (e.g., glycine, asparagine, glutamine, serine, tyrosine, or cysteine).

[00182] In one embodiment of a variant PBRT, the threonine at position 7 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the threonine is replaced with an amino acid having a beta-branched side chain configuration (e.g., valine, isoleucine).

[00183] In another embodiment of a variant PBRT, the threonine at position 7 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain group such that the threonine is replaced with an amino acid having an hydroxyl side chain group (e.g., serine).

[00184] In another embodiment of a variant PBRT, the threonine at position 7 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily positive relatedness such that the threonine is replaced with any of serine, proline, alanine, glycine, asparagine, aspartic acid, glutamic acid, lysine, isoleucine, or valine

[00185] In one embodiment of a variant PBRT, mutation of the threonine at position 7 of SEQ ID NO: 1 to any of glycine, alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, tryptophan, tyrosine, valine wherein mutation of the threonine at position 7 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00186] In one embodiment of a variant PBRT, mutation of the threonine at position 7 of SEQ ID NO: 1 is with a non-natural or synthetic amino acid wherein mutation of the threonine at position 7 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00187] In one embodiment of a variant PBRT, the leucine at position 8 of SEQ ID NO: 1 can be mutated to an isoleucine residue.

[00188] In one embodiment of a variant PBRT, the leucine at position 8 of SEQ ID NO: 1 can be mutated to a phenylalanine residue.

[00189] In one embodiment of a variant PBRT, the leucine at position 8 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the leucine is replaced with an amino acid having a nonpolar side chain configuration (e.g., alanine, valine, isoleucine, proline, phenylalanine, methionine, or tryptophan).

[00190] In one embodiment of a variant PBRT, the leucine at position 8 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the leucine is replaced with an amino acid having an aliphatic side chain configuration (e.g., glycine, alanine, valine, or isoleucine).

[00191] In another embodiment of a variant PBRT, leucine at position 8 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain group such that the leucine is replaced with an amino acid having an aliphatic side chain group (e.g., alanine, valine, isoleucine, proline).

[00192] In another embodiment of a variant PBRT, the leucine at position 8 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily positive relatedness such that the leucine is replaced with any of methionine, isoleucine, valine, or phenylalanine

[00193] In another embodiment of a variant PBRT, the leucine at position 8 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily neutral relatedness such that the leucine is replaced with any of serine, asparagine, glutamic acid, histidine, or methionine.

[00194] In one embodiment of a variant PBRT, mutation of the leucine at position 8 of SEQ ID NO: 1 to any of glycine, alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, histidine, isoleucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine wherein mutation of the leucine at position 8 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00195] In one embodiment of a variant PBRT, mutation of the leucine at position 8 of SEQ ID NO: 1 non-natural amino acid wherein mutation of leucine at position 8 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00196] In one embodiment of a variant PBRT, the tyrosine at position 9 of SEQ ID NO: 1 can be mutated to an isoleucine or valine residue.

[00197] In one embodiment of a variant PBRT, the tyrosine at position 9 of SEQ ID NO: 1 can be mutated to a phenylalanine, threonine, asparagine, aspartic acid, lysine, or serine residue.

[00198] In one embodiment of a variant PBRT, the tyrosine at position 9 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the tyrosine is replaced with an amino acid having an uncharged polar side chain configuration (e.g., glycine, asparagine, glutamine, serine, threonine, or cysteine).

[00199] In one embodiment of a variant PBRT, the tyrosine at position 9 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain configuration such that the tyrosine is replaced with an amino acid having an aromatic side chain configuration (e.g., tyrosine, phenylalanine, tryptophan, or histidine).

[00200] In another embodiment of a variant PBRT, the tyrosine at position 9 of SEQ ID NO: 1 can be mutated to an amino acid having a similar side chain group such that the

tyrosine is replaced with an amino acid having an aromatic side chain group (e.g., phenylalanine, tryptophan).

[00201] In another embodiment of a variant PBRT, the tyrosine at position 9 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily positive relatedness such that the tyrosine is replaced with any of phenylalanine or tryptophan.

[00202] In another embodiment of a variant PBRT, the tyrosine at position 9 of SEQ ID NO: 1 can be mutated to an amino acid having evolutionarily neutral relatedness such that the tyrosine is replaced with any of cysteine or histidine.

[00203] In one embodiment of a variant PBRT, mutation of the tyrosine at position 9 of SEQ ID NO: 1 to any of glycine, alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine wherein mutation of tyrosine at position 9 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00204] In one embodiment of a variant PBRT, mutation of the tyrosine at position 9 of SEQ ID NO: 1 is with a non-natural or synthetic amino acid wherein mutation of the tyrosine at position 9 of SEQ ID NO: 1 will result in a precipitable beta-roll tag that is capable of binding to Ca^{2+} , undergo reversible precipitation in the presence of Ca^{2+} , or induce reversible precipitation of a purification moiety linked to the variant PBRT.

[00205] As described herein, a variant PBRC can further comprise a capping sequence. In certain embodiments, the capping sequence in a variant PBRC can be a variant capping sequence. A variant capping sequence can be an amino acid sequence having at least 70%, at least 80%, at least 90%, at least 95% or at least 98% identity to SEQ ID NO: 3. In another embodiment, a variant capping sequence is an amino acid sequence having at least 70%, at least 80%, at least 90%, at least 95% or at least 98% identity to any of SEQ ID NO: 4-23.

[00206] In still a further embodiment, a variant capping sequence is a sequence comprising one or more amino acid substitutions with an amino acid having a similar side chain group. In one embodiment, the variant capping sequence comprises the sequence of SEQ ID NO: 3 or any of SEQ ID NO: 4-23, wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at

least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, at least 45, at least 46, at least 47, at least 48, at least 49, at least 50, at least 51, at least 52, at least 53, at least 54, at least 55, at least 56, at least 57, at least 58, at least 59, at least 60, at least 61, at least 62, at least 63, at least 64, at least 65, at least 66, at least 67, or at least 68, amino acids in the sequence of SEQ ID NO: 3 or any of SEQ ID NO: 4-23 are substituted an amino acid having a similar side chain group.

[00207] In still a further embodiment, a variant capping sequence is a sequence comprising one or more amino acid substitutions with an amino acid having a similar side chain configuration. In one embodiment, the variant capping sequence comprises the sequence of SEQ ID NO: 3 or any of SEQ ID NO: 6-23, wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, at least 45, at least 46, at least 47, at least 48, at least 49, at least 50, at least 51, at least 52, at least 53, at least 54, at least 55, at least 56, at least 57, at least 58, at least 59, at least 60, at least 61, at least 62, at least 63, at least 64, at least 65, at least 66, at least 67, or at least 68, amino acids in the sequence of SEQ ID NO: 3 or any of SEQ ID NO: 6-23 are substituted an amino acid having a similar side chain configuration.

[00208] In still a further embodiment, a variant capping sequence is a sequence comprising one or more amino acid substitutions with an amino acid having evolutionarily positive relatedness. In one embodiment, the variant capping sequence comprises the sequence of SEQ ID NO: 3 or any of SEQ ID NO: 6-23, wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, at least 45, at least 46, at least 47, at least 48, at least 49, at least 50, at least 51, at least 52, at least 53, at least 54, at least 55, at least 56, at least 57, at least 58, at least 59, at least 60, at least

61, at least 62, at least 63, at least 64, at least 65, at least 66, at least 67, or at least 68, amino acids in the sequence of SEQ ID NO: 3 or any of SEQ ID NO: 6-23 are substituted an amino acid having evolutionarily positive relatedness.

[00209] In still a further embodiment, a variant capping sequence is a sequence comprising one or more amino acid substitutions with an amino acid having evolutionarily neutral relatedness. In one embodiment, the variant capping sequence comprises the sequence of SEQ ID NO: 3 or any of SEQ ID NO: 6-23, wherein at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, at least 45, at least 46, at least 47, at least 48, at least 49, at least 50, at least 51, at least 52, at least 53, at least 54, at least 55, at least 56, at least 57, at least 58, at least 59, at least 60, at least 61, at least 62, at least 63, at least 64, at least 65, at least 66, at least 67, or at least 68, amino acids in the sequence of SEQ ID NO: 3 or any of SEQ ID NO: 4-23 are substituted an amino acid having evolutionarily neutral relatedness.

[00210] In another embodiment, the PBRT variants or PBRC variants described herein can also comprise a non-natural amino acid. As used herein, a non-natural amino acid can be, but is not limited to, an amino acid comprising a moiety where a chemical moiety is attached, such as an aldehyde- or keto-derivatized amino acid, or a non-natural amino acid that includes a chemical moiety. A non-natural amino acid can also be an amino acid comprising a moiety where a saccharide moiety can be attached, or an amino acid that includes a saccharide moiety. Examples of non-classical amino acids suitable for use with the methods and compositions described herein include, but are not limited to, D-isomers of the common amino acids, 2,4-diaminobutyric acid, alpha-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, gamma-Abu, epsilon-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, beta-alanine, fluoro-amino acids, designer amino acids such as beta-methyl amino acids, C alpha-methyl amino acids, N alpha-methyl amino acids, and amino acid analogs in general.

[00211] The PBRT variants or PBRC variants described herein can also comprise one or more amino acid analog substitutions, e.g., unnatural amino acids such as alpha alpha-disubstituted amino acids, N-alkyl amino acids, lactic acid, and the like. These analogs include phosphoserine, phosphothreonine, phosphotyrosine, hydroxyproline, gamma-carboxyglutamate; hippuric acid, octahydroindole-2-carboxylic acid, statine, 1,2,3,4,-tetrahydroisoquinoline-3-carboxylic acid, penicillamine, ornithine, citruline, .alpha.-methyl-alanine, para-benzoyl-phenylalanine, phenylglycine, propargylglycine, sarcosine, .epsilon.-N,N,N-trimethyllysine, .epsilon.-N-acetyllysine, N-acetylserine, N-formylmethionine, 3-methylhistidine, 5-hydroxylysine, .omega.-N-methylarginine, and other similar amino acids and imino acids and tert-butylglycine. The ability of PBRTs or PBRCs comprising an analog substitutions to bind to Ca²⁺, undergo reversible precipitation in the presence of Ca²⁺, or induce reversible precipitation of a purification moiety linked to the PBRT or PBRC using methods known to those skilled in the art.

[00212] The PBRT variants or PBRC variants described herein can further comprise polypeptide analogs, such as peptide mimetics (Fauchere J, Adv. Drug Res. 15:29 (1986); Veber D F and Freidinger R M, Trends Neurosci. 8:392-96 (1985); Evans B E et al., J. Med. Chem 30:1229-39 (1987)). Generally, peptidomimetics are structurally similar to a template polypeptide (i.e., a polypeptide that has a biological or pharmacological activity), such as the PBRTs or PBRCs described herein, but have one or more peptide linkages replaced by a linkage selected from the group consisting of: --CH.sub.2NH--, --CH.sub.2S--, --CH.sub.2--CH.sub.2--, --CH.dbd.CH-- (cis and trans), --COCH.sub.2--, --CH(OH)CH.sub.2--, and --CH.sub.2SO--, by methods known in the art and further described in the following references: Spatola A F in "Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins," B. Weinstein, ed., Marcel Dekker, New York, p. 267 (1983); Spatola, A F, Vega Data (March 1983), Vol. 1, Issue 3, "Peptide Backbone Modifications" (general review); Morley J S, Trends Pharmacol. Sci. 1:463-68 (1980) (general review); Hudson D et al., Int. J. Pept. Prot. Res. 14:177-85 (1979) (--CH.sub.2NH--, CH.sub.2CH.sub.2--); Spatola A F et al., Life Sci. 38:1243-49 (1986) (--CH.sub.2--S); Hann M M, J. Chem. Soc. Perkin Trans. 1, 307-314 (1982) (--CH--CH--, cis and trans); Almquist R G et al., J. Med. Chem. 23:1392-98 (1980) (--COCH.sub.2--); Jennings-White C et al., Tetrahedron Lett. 23:2533-34 (1982) (--COCH.sub.2--); EP 0 045 665 (--CH(OH)CH.sub.2--); Holladay M W et al., Tetrahedron Lett., 24:4401-04 (1983) (--C(OH)CH.sub.2--); Hraby V J, Life Sci. 31:189-99 (1982) (--CH.sub.2--S--). One example of a non-peptide linkage is --CH.sub.2NH--.

[00213] Such polypeptide mimetics can have advantages over polypeptide embodiments, including, for example: more economical production, greater chemical stability, enhanced pharmacological properties (half-life, absorption, potency, efficacy, etc.), altered specificity (e.g., a broad-spectrum of biological activities), reduced antigenicity, and others. Labeling of peptidomimetics can involve covalent attachment of one or more labels, directly or through a spacer (e.g., an amide group), to non-interfering position(s) on the peptidomimetic that are predicted by quantitative structure-activity data and/or molecular modeling. Such non-interfering positions can be positions that do not form direct contacts with the macromolecules(s) to which the peptidomimetic binds to produce the therapeutic effect. Derivatization (e.g., labeling) of peptidomimetics can be done without substantially interfering with the desired biological or pharmacological activity of the peptidomimetic. The ability of any peptidomimetics to polypeptides can be assayed for the ability to bind 1,4,-benzothiazepine or derivatives thereof using methods known to those skilled in the art.

[00214] Systematic substitution of one or more amino acids of the PBRTs or PBRCs described herein with a D-amino acid of the same type (e.g., D-lysine in place of L-lysine) can be used to generate additional PBRT and PBRT variants.

[00215] The following methods can be used in connection with the embodiments of the invention.

EXAMPLES

[00216] Example 1: Purification of PBRT or PBRC linked Purification moieties

[00217] 17 tandem repeats of the amino acid sequence GGAGNDTLY followed by a C-terminal "capping" sequence (the sequence is provided below) were fused to maltose binding protein. This construct was expressed in E. coli and the cells are lysed, creating a complex mixture of E. coli proteins and the Beta Roll tagged maltose binding protein. The mixture was exposed to 100 mM calcium chloride solution to form a precipitate form. The precipitate was pelleted and resuspended in calcium-free buffer. One precipitation cycle was sufficient to generate a relatively pure protein (Figure 1). Multiple cycles can be used to achieve better purity. SDS-PAGE samples of this procedure can be found in the supplemental materials.

[00218] The capping sequence used in this example is:

INAGADQLW FARQGNDLEIRILGTDDALTVH DWYRDADHRVEIIHAANQAVDQAGI
EKLVEAMAQYPD (SEQ ID NO: 3)

[00219] Polypeptides comprising a 5 or 17 repeat C-capped precipitable beta-roll tags were expressed as MBP fusion proteins. 50 mM calcium was added to clarified cell lysates to induce precipitation of the polypeptides comprising a 5 or 17 repeat C-capped precipitable beta-roll tags. The precipitate was pelleted by centrifugation and the pellet was washed once and resuspended in buffer with 50 mM EGTA. The lysate, the supernatant after calcium addition, and the resuspended pellet were then run on an SDS-PAGE gel (Figure 2). The results show that the methods described herein can be used to rapidly purify polypeptides comprising a 5 or 17 repeat C-capped precipitable beta-roll tags. Precipitation of purification moieties comprising a precipitable beta-roll tag can be confirmed by circular dichroism analysis (Figure 3).

[00220] An intein domain can be coupled to the construct so that the cleavage reaction and subsequent second precipitation can be examined. Other proteins, in addition to maltose binding protein can be used with the purification protocols described herein.

[00221] Example 2: PBRT sequence heat map.

[00222] The heat map for the precipitable beta-roll tag sequences described herein was determined by using BLAST to find beta roll sequences similar to the metalloprotease of *S. marcescens* and then quantifying the frequency of amino acids at each of the nine positions after beta roll sequences were identified (Figure 4). Certain positions in the precipitable beta-roll tag are not highly variable (e.g. positions 1, 2, 4, 6, and 8), whereas other positions, exhibit moderate conservation. Positions 7, 9 are highly variable and can be substituted with any natural or non-natural amino acid.

[00223] Example 3: Exemplary PBRCs

[00224] Below are the amino acid sequences of two exemplary PBRCs.

[00225] GGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYINA
GADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIHAANQAVDQAGIEKL
VEAMAQYPD (SEQ ID NO: 4)

[00226] GGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGA
GNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGG
AGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYGGAGNDTLYIN
AGADQLWFARQGNDLEIRILGTDDALTVHWDYRDADHRVEIIHAANQAVDQAGIEK
LVEAMAPYD (SEQ ID NO: 5)

[00227] Example 4: Sequences of PBRT and PBRC peptides that do not induce precipitation in response to Ca²⁺

[00228] The following peptides are soluble in the presence of calcium.

GSARDDVLIGDAGANVLNGLADNDVLSGGAGDDVLLGDEGSDLLSGDAGNDDLFG
GQGDDTYLFGVGYGHDTIYESGGGHDTIRGSARDDVLIGDAGANVLNGLADNDVLS
GGAGDDVLLGDEGSDLLSGDAGNDDLFGGQGDDTYLFGVGYGHDTIYESGGGHDTI
RINAGADQLWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIHAANQAVDQAG
IEKLVEAMAQYPD (SEQ ID NO: 1339)

GSARDDVLIGDAGANVLNGLADNDVLSGGAGDDVLLGDEGSDLLSGDAGNDDLFG
GQGDDTYLFGVGYGHDTIYESGGGHDTIRGDAGANVLNGLADNDVLSGGAGDDVL
LGDEGSDLLSGDAGNDDLFGGQGDDTYLFGVGYGHDTIYESGGGHDTIR (SEQ ID
NO: 1340)

GSARDDVLIGDAGANVLNGLADNDVLSGGAGDDVLLGDEGSDLLSGDAGNDDLFG
GQGDDTYLFGVGYGHDTIYESGGGHDTIRGDAGANVLNGLADNDVLSGGAGDDVL
LGDEGSDLLSGDAGNDDLFGGQGDDTYLFGVGYGHDTIYESGGGHDTIRINAGADQ
LWFARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIHAANQAVDQAGIEKLVEAM
AQYPD (SEQ ID NO: 1341)

GGSGNDVIVGNAANNVLKGGAGNDVLFGGGGADELWGGAGKDIFV (SEQ ID NO:
1341)

GGSGNDVIVGNAANNVLKGGAGNDVLFGGGGADELWGGAGKDIFVFSAASDSAPG
ASDWIRDFQKGIDKIDLSFFNKEANSSDFIHFVDHFSGTAGEALLSYNASSNVTDLV
NIGGHQAPDFLVKIVGQVDVATDFIV (SEQ ID NO: 1342)

[00229] References

[00230] Hochuli et al., "Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent", Nature Biotechnology, Issue 6, pp 1321 – 1325 (1988)

[00231] Guana et al., "Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein", Gene, Volume 67, Issue 1, pp 21-30, July 15 1988

[00232] Banki et al., "Simple bioseparations using self-cleaving elastin-like polypeptide tags", Nature Methods, Issue 2, pp 659 – 662, August 2, 2005

[00233] Wood et al., “A genetic system yields self-cleaving inteins for bioseparations.”
Nature Biotechnology, Issue 17, Pages 889 - 892 September 17, 1999.

[00234] U.S. Ser. No.: 11/374403

WHAT IS CLAIMED IS:

1. A precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs comprise the amino acid sequence of SEQ ID NO: 1
2. A precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of:
 - (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or
 - (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337
3. A precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of:
 - (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or
 - (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337
 - (c) a polypeptide comprising the amino acid sequence GXXXXXXXXX, wherein,
 - (i) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and
 - (ii) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and
 - (iii) the X at position 4 is an amino acid selected from the group consisting of glycine or alanine, and
 - (iv) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and

- (v) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine,
 - (vi) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and
 - (vii) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and
 - (viii) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine.
4. A precipitable beta roll cassette (PBRC) comprising one or more beta roll tags (PBRTs) wherein the one or more PBRTs are independently any of:
- (a) a polypeptide having the amino acid sequence of SEQ ID NO: 1, or
 - (b) a polypeptide having the amino acid sequence of any of SEQ ID NOs 25-1337
 - (c) a polypeptide comprising the amino acid sequence GXXXXXXXXX, wherein,
 - (i) the X at position 2 is an amino acid selected from the group consisting of glycine, asparagine or aspartic acid, and
 - (ii) the X at position 3 is an amino acid selected from the group consisting of alanine, serine, glycine, aspartic acid, glutamic acid, leucine or asparagine, and
 - (iii) the X at position 4 is an amino acid selected from the group consisting of glycine or alanine, and
 - (iv) the X at position 5 is an amino acid selected from the group consisting of asparagine, aspartic acid, alanine, or serine, and

(v) the X at position 6 is an amino acid selected from the group consisting of aspartic acid or asparagine,

(vi) the X at position 7 is an amino acid selected from the group consisting of threonine, isoleucine, valine, or leucine, and

(vii) the X at position 8 is an amino acid selected from the group consisting of leucine, isoleucine, or phenylalanine, and

(viii) the X at position 9 is an amino acid selected from the group consisting of tyrosine, isoleucine, valine, phenylalanine, threonine, asparagine, aspartic acid, lysine or serine, or

(c) a variant PBRT.

5. The PBRC of any of claims 1-4 further comprising a capping sequence.
6. The PBRC of any of claims 1-4 further comprising a stabilizing polypeptide.
7. A PBRC linked purification moiety comprising the PBRC of any of claims 1-4.
8. The PBRC linked purification moiety of claim 7 wherein the PBRC is linked to the purification moiety by a peptide bond.
9. The PBRC linked purification moiety of claim 7 wherein the PBRC is linked to the purification moiety by a chemical bond that is not a peptide bond.
10. The PBRC of any of claims 1-4 further comprising a cleavage site located N-terminally or C-terminally to one or more of the one or more PBRTs.
11. The PBRC of claim 10, wherein the cleavage site is selected from the group comprising an intein cleavage site, a Factor Xa cleavage site, a thrombin cleavage site, an enterokinase cleavage site, or a signal peptidase cleavage site.
12. A polypeptide comprising the PBRC of any of claims 1-4 and a purification moiety.
13. A nucleic acid sequence encoding the polypeptide of any of claims 1-12.
14. A method for purifying a PBRC linked purification moiety, the method comprising

- (a) expressing the PBRC linked purification moiety in an expression system,
 - (b) collecting the PBRC linked purification moiety in a first medium,
 - (c) adding Ca^{2+} to the first medium so as to induce precipitation of PBRC linked purification moiety,
 - (d) removing unprecipitated material from the medium from the precipitated PBRC linked purification moiety,
 - (e) resuspending the PBRC linked purification moiety in a second medium having a lower than the free Ca^{2+} concentration than the free Ca^{2+} concentration obtained after step (c).
15. The method of claim 14 wherein a calcium chelator is added to the second medium of step (e).
16. The method of claim 14, wherein steps (c) to (e) are repeated one or more times.
17. The method of claim 14, further comprising a step of removing precipitated material between step (b) and step (c).
18. The method of claim 14, wherein the PBRC comprises a cleavage site between the PBRC and the purification moiety.
19. The method of claim 14, further comprising steps of:
- (i) cleaving the PBRC linked purification moiety so as to separate the purification moiety from the PBRC,
 - (ii) adding Ca^{2+} to the medium so as to induce precipitation of the PBRC, and
 - (iii) isolating the unprecipitated purification moiety.
20. The method of claim 19, wherein the cleavage site is an intein cleavage site.
21. An expression vector comprising, as arranged from 5' to 3', a promoter, a nucleic acid sequence encoding the PBRC of any of claims 1-4, and at least one cloning site.

22. An expression vector comprising, as arranged from 5' to 3', a promoter, at least one cloning site, and a nucleic acid sequence encoding the PBRC of any of claims 1-4.
23. An expression vector comprising, as arranged from 5' to 3', a promoter, at least one cloning site, a nucleic acid sequence encoding the PBRC of any of claims 1-4 and at least a second cloning site.

FIGURES

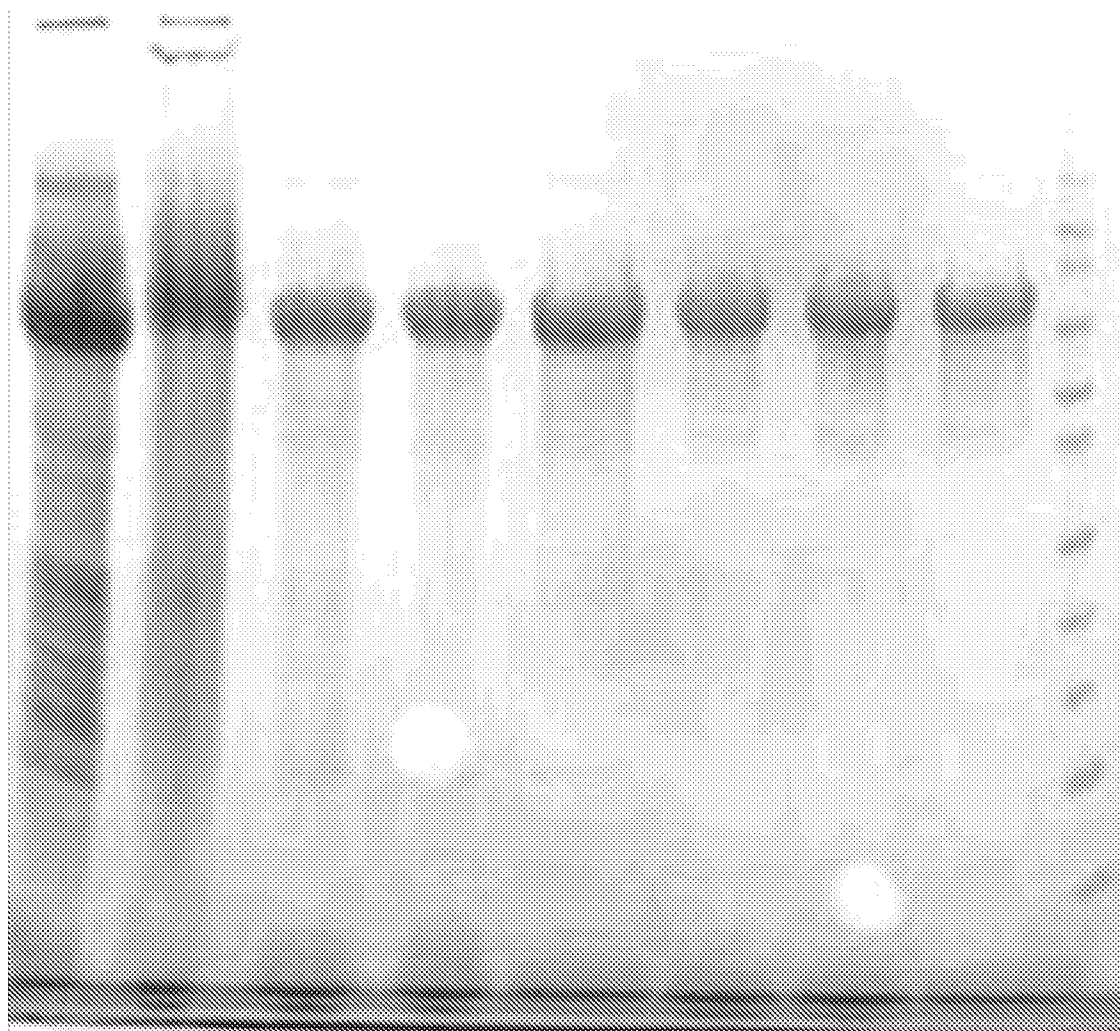
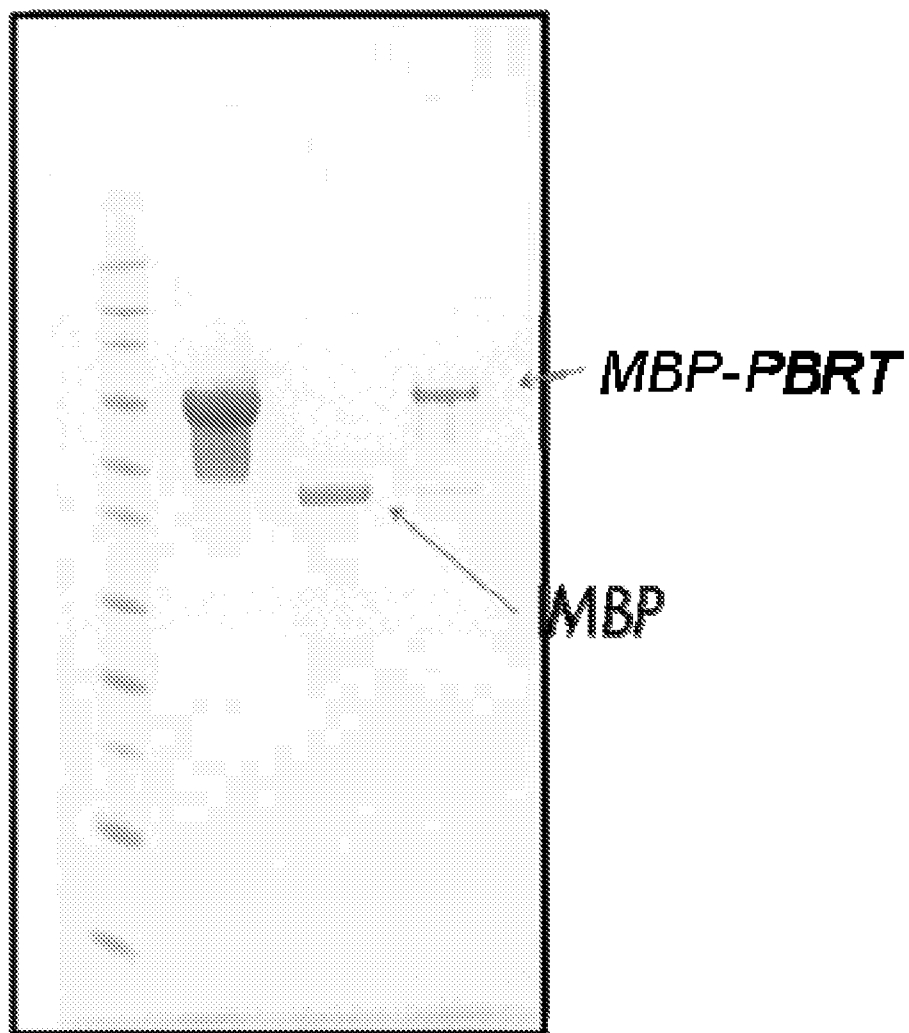


Figure 1A

**Figure 1B**

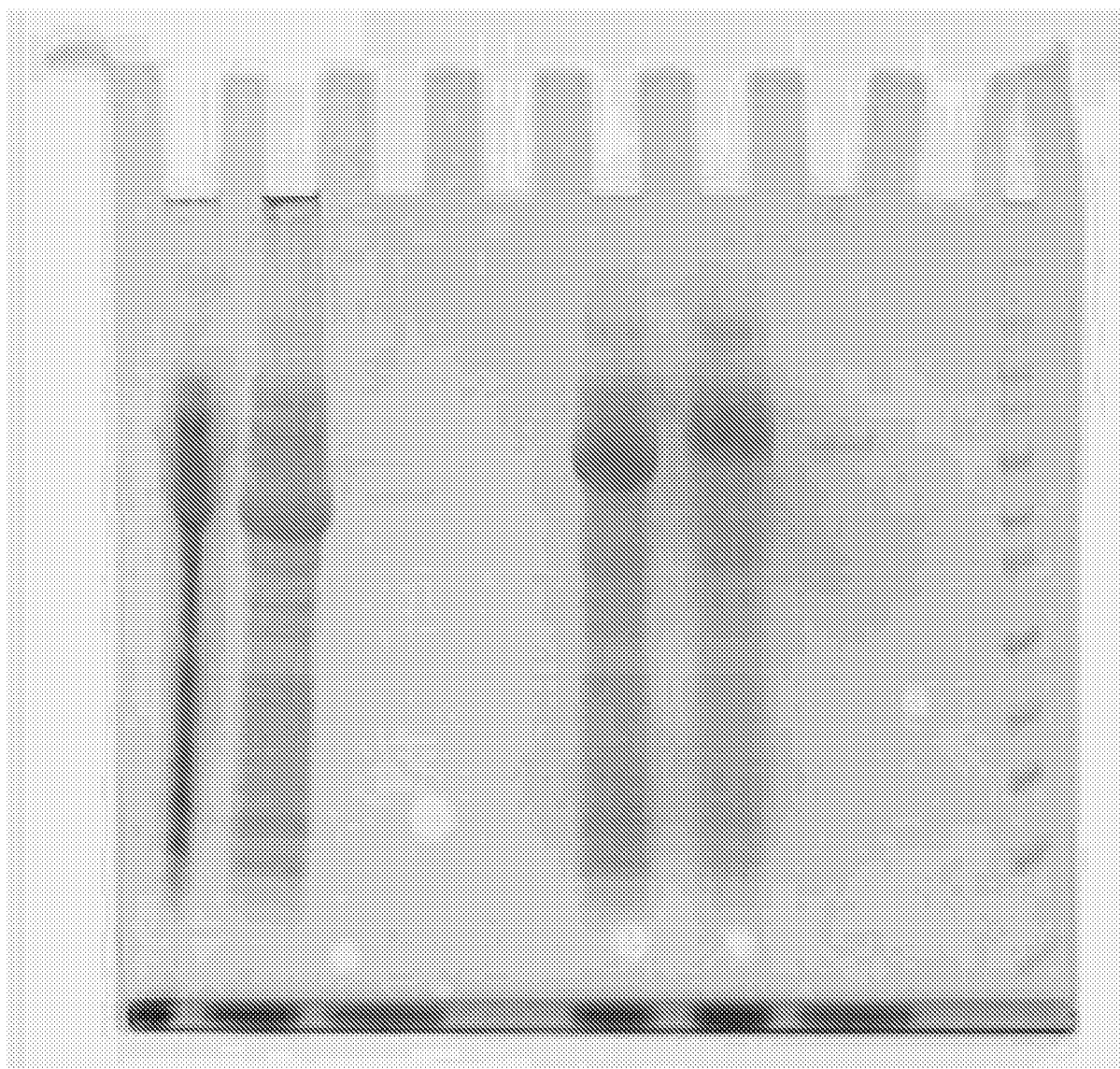
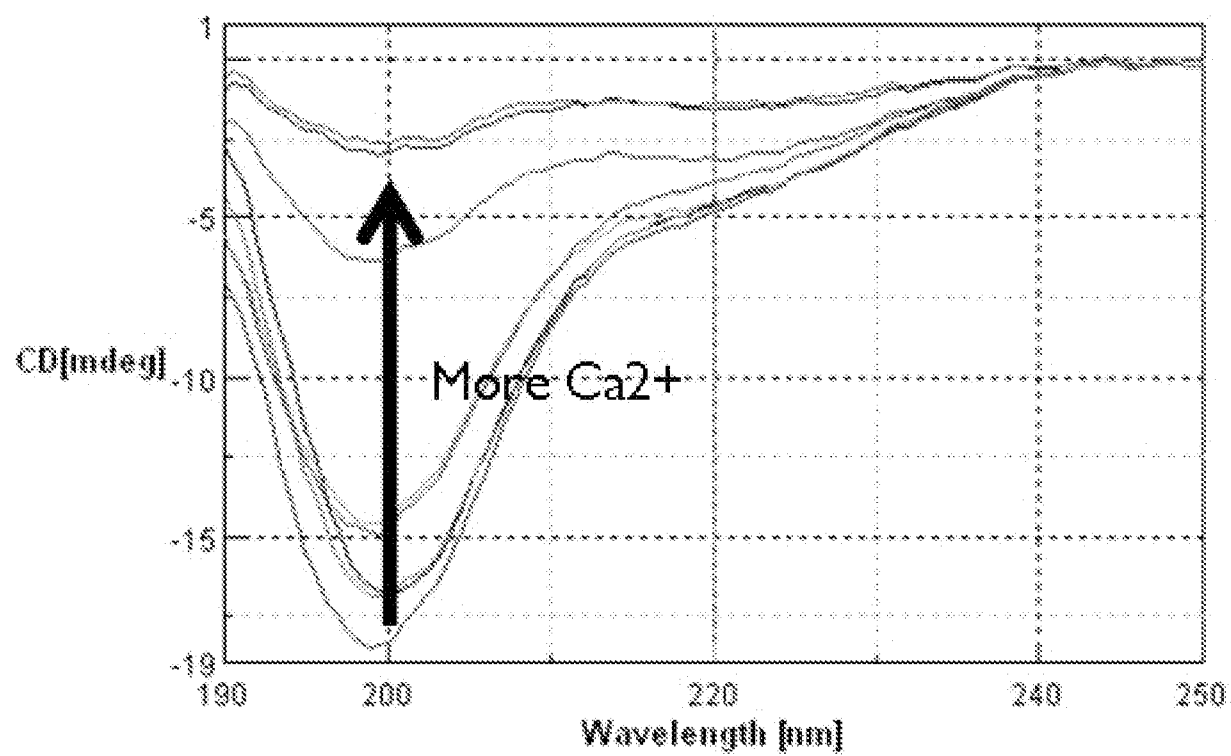


Figure 2

**Figure 3**

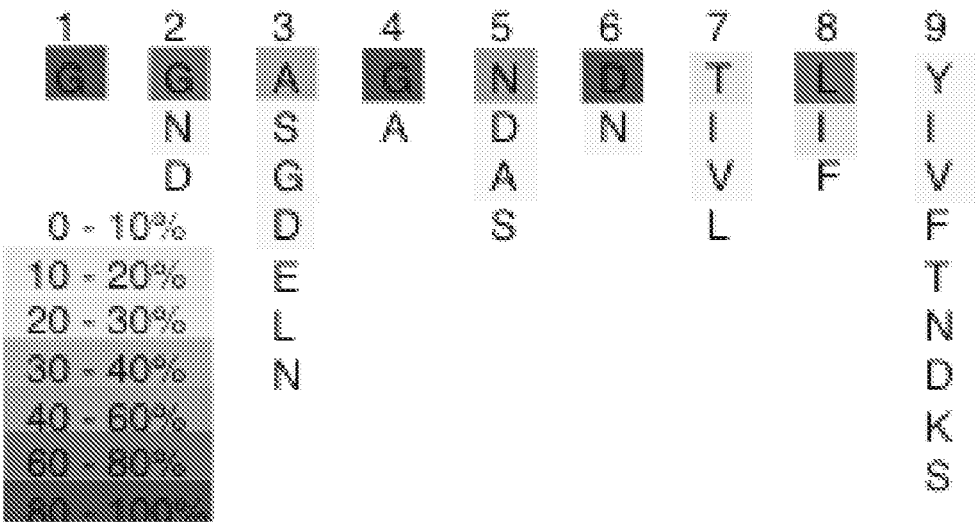


Figure 4

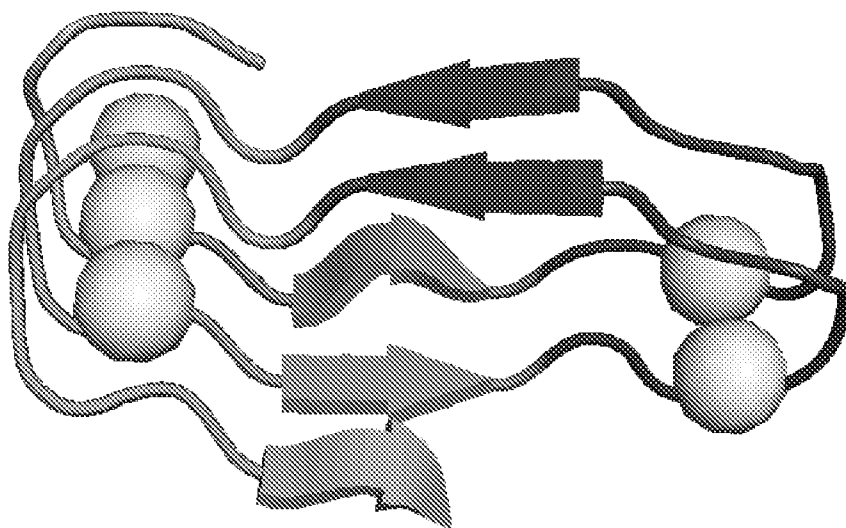


Figure 5

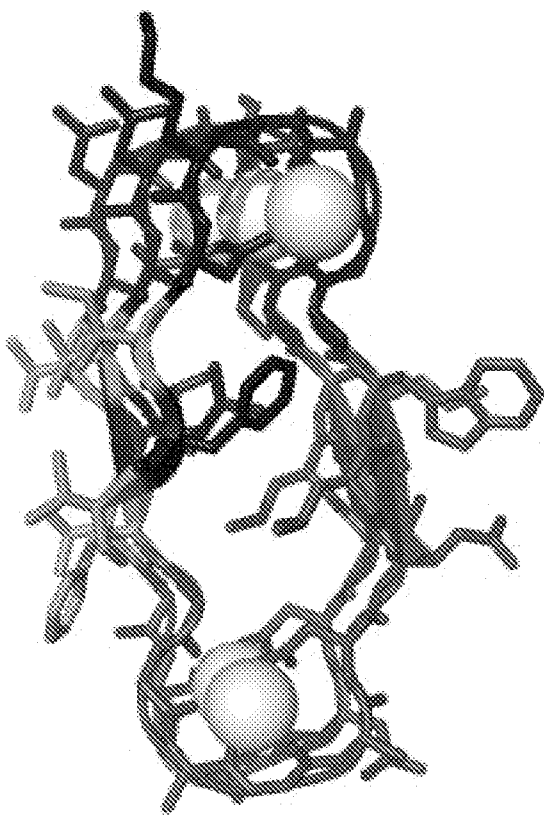


Figure 6



Figure 7

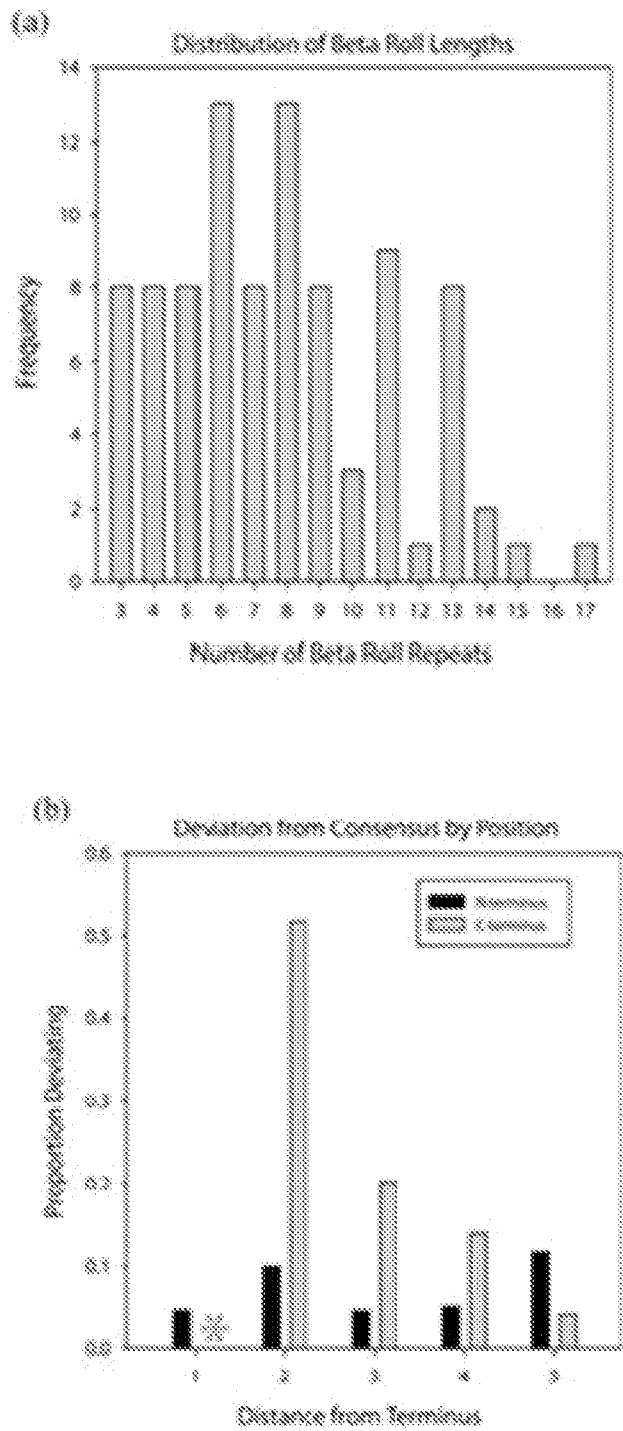


Figure 8A-B

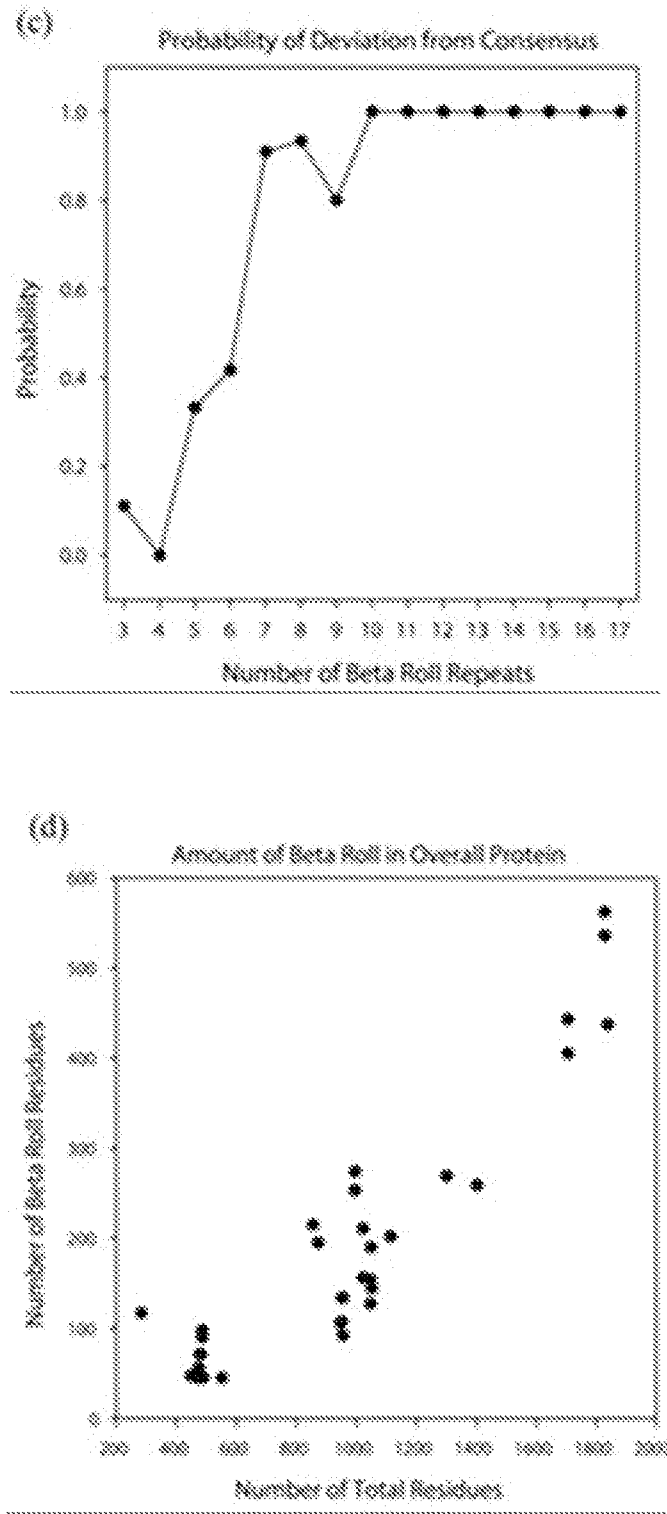
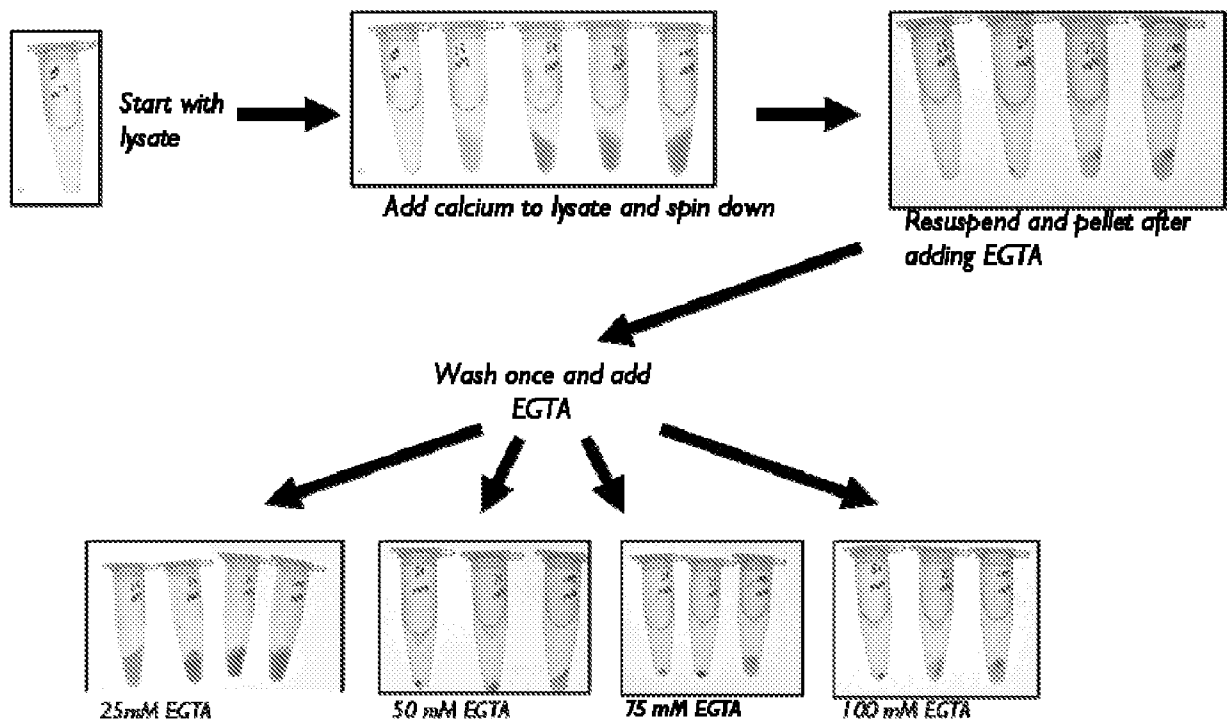
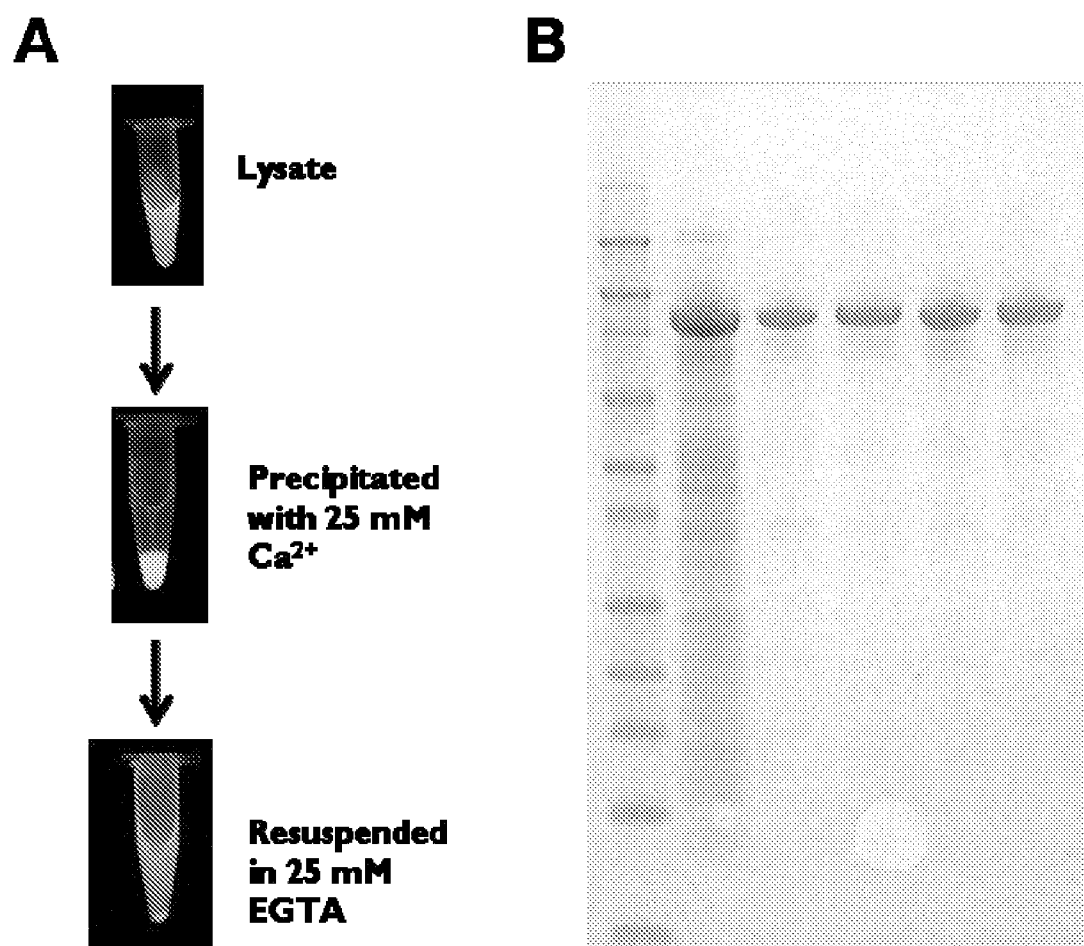
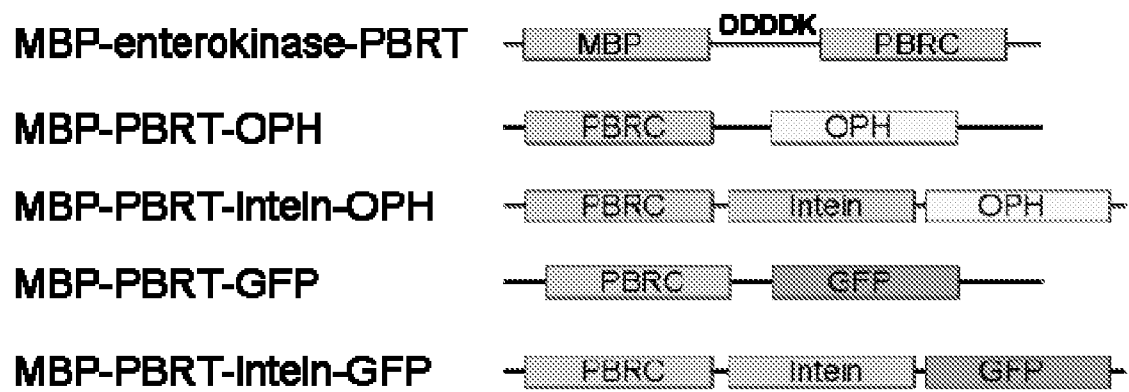


Figure 8C-D

**Figure 9**

**Figure 10**

**Figure 11**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/33293

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C07K 7/06 (2012.01)

USPC - USPC - 530/328

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8) - C07K 1/00; 7/06 (2012.01)

USPC - 530/328, 344, 355

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

IPC(8) - C07K 1/00; 7/06 (2012.01), USPC - 530/328, 344, 355, 930/10: keyword search, as below

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

USPTO PubWest (databases: PGPB,USPT,USOC,EPAB,JPAB), Thompson Innovation (core patent databases), Google Scholar -- Search Terms: Banta, Shur, precipitate, precipitation, beta roll, calcium, hemolysin, repeat, aggregation, agglomeration, purification, affinity, tag, cassette, divalent

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SCOTTER et al. "Metal ion-dependent, reversible, protein filament formation by designed beta-roll polypeptides" BMC Struct. Biol.; 1 October 2007; Vol. 7:63, pages 1-13 [online]; downloaded from http://www.biomedcentral.com/content/pdf/1472-6807-7-63.pdf on 14 September 2012. Entire document, especially abstract, pg. 4, col 1, para 2-para 3; pg. 8, col 2, para 1; pg. 5, col 1, para 1; pg. 2, col 2, para 4; pg. 10, col 2, para 2; pg. 10, col 1, para 3	1-12, 14-23
Y	US 2009/0075335 A1 (BANTA et al.) 19 March 2009 (19.03.2009) para [0014]; [0042]; [0048][0049]	1-12, 14-23
Y	TEBO et al., UniProt Accession A5PAU7; 13 July 2010; [online]; downloaded from http://www.uniprot.org/uniprot/A5PAU7.txt?version=14 on 14 September 2012.	1-12, 21-23
Y	BLENNER et al. "Calcium-Induced Folding of a Beta Roll Motif Requires C-Terminal Entropic Stabilization" JMB; available online 11 May 2010; Vol. 400, pg. 244-256. Entire document, especially the abstract.	6

☐ Further documents are listed in the continuation of Box C.


* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

15 September 2012 (15.09.2012)

Date of mailing of the international search report

01 OCT 2012

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/33293

Box No. 1 Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:

a. (means)

☐

on paper

☒

in electronic form

b. (time)

☐

in the international application as filed

☐

together with the international application in electronic form

☒

subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/33293

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 13
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.