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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 June 2008 (05.06.2008)

PCT

(10) International Publication Number
WO 2008/065415 A2

(51) International Patent Classification:
*C07K 1/22* (2006.01) *C12N 15/62* (2006.01)
*C07K 1/32* (2006.01) *G01N 33/50* (2006.01)
*C07K 14/00* (2006.01)

(21) International Application Number:
PCT/GB2007/004597

(22) International Filing Date:
30 November 2007 (30.11.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0624071.7 1 December 2006 (01.12.2006) GB

(71) Applicant *(for all designated States except US)*: **BIOLINE RESEARCH LIMITED** [GB/GB]; 16 The Edge Business Centre, Humber Road, London Greater London NW2 6EW (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **IGNATOV, Konstantin** [RU/RU]; Obrucheva Str. 35-1-82, Moscow, 117246 (RU). **KRAMAROV, Vladimir** [RU/RU]; Yushunskaya M. Str. 8-2-80, Moscow, 113303 (RU). **GUBANOV, Sergey Ivanovich** [RU/DE]; Louis-Pasteur Str. 23, 14943 Luckenwalde (DE).

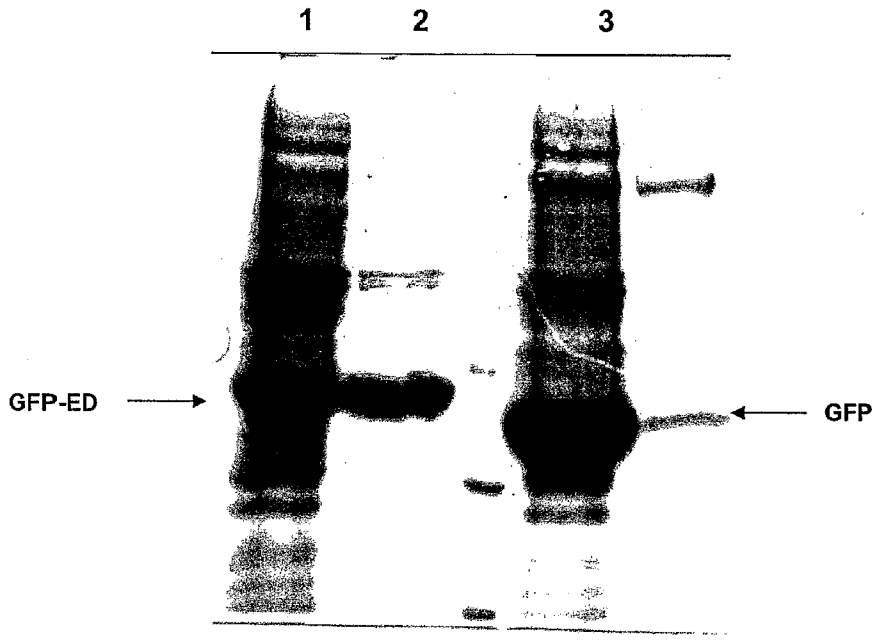
(74) Agents: **KREMER, Simon** et al.; Mewburn Ellis LLP, York House, 23 Kingsway, London, Greater London WC2B 6HP (GB).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL,

*[Continued on next page]*

(54) Title: PURIFICATION METHODS AND MATERIALS




(57) Abstract: The invention provides methods and materials for producing and purifying soluble targets, such as recombinant proteins, from a sample. The invention employs "tags" which can be appended to the target. By addition of a tag-binding entity (or "ligand", the terms are used interchangeably) to the tagged target, the solubility of the target can be reduced and it can be selectively precipitated. Preferred tags are metal binding peptide tags (which may share homology with Ca2+ binding EGF-like domains) wherein addition of appropriate sparingly soluble metal salts to the tagged target causes precipitation.

WO 2008/065415 A2

PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**

— *of inventorship (Rule 4.17(iv))*

**Published:**

— *without international search report and to be republished upon receipt of that report*

# Purification methods and materials

## Technical field

The present invention relates generally to methods and materials for use in recovering or enriching target entities from a sample, for example by reduction of solubility and precipitation. The invention further provides novel tags which may be used to produce and/or purify recombinant polypeptides or proteins. This invention also relates to fusion proteins containing affinity peptide tags, genes coding for such proteins, expression vectors and transformed microorganisms containing such genes.

## Background art

Typically, proteins are purified using conventional chromatography or high-performance liquid chromatography methods.

Purification of proteins by conventional chromatography is usually achieved using a combination of chromatographic methods including gel-filtration, ion-exchange, hydrophobic-interaction, dye-interaction, affinity and immunoaffinity chromatography. If affinity or immunoaffinity columns are not available for purifying a particular protein target, then it is typically necessary to utilize sequential chromatographic steps and to analyze (such as by one-dimensional gel electrophoresis) the protein after each purification step, to determine if the protein is homogeneous.

The possibility of preparing hybrid genes by gene technology has opened up new routes for the workup of recombinant proteins. Because the highest affinity interactions in biology occur between macromolecules it has been common to utilise a protein:protein interaction as an affinity step. By linking the coding gene sequence of a desired protein to the coding gene sequence of a protein fragment having a high affinity for a ligand (affinity peptide or affinity tag), it is possible to purify desired recombinant proteins in the form of fusion proteins in one step using the affinity tag.

A variety of fusion tags have been used to date for recombinant proteins:

(i) a peptide which recognizes a specific antibody e.g. the myc tag, (Munro, S., and Pelham, H. R. B. (1986) Cell, 46, 291-300; Ward, E. S., Gussow, D., Griffiths, A. D., Jones, P. T. and

Winter, G. (1989) Nature, 341, 544-546); the Flag-peptide (Hopp, T. P. et al (1988) BioTechnology, 6, 1204-1210); the KT3 epitope (Martin, G. A. et al (1990) Cell, 63, 843-849; Martin, G. A., et al (1992) Science, 255, 192-194); an alpha-tubulin epitope (Skinner, R. H. et al (1991) J.Biol.Chem., 266, 14163-14166); the T7 gene 10 protein peptide tag (Lutz-Freyermuth, C., Query, C. C. and Keene, J. D. (1990) Proc. Natl. Acad. Sci. U.S.A., 87, 6393-6397); the strep-tag, a peptide tag binding to streptavidin (Schmidt, T. G. M. and Skerra, A. (1993) Protein Engineering, 6, 109-122). The process of capturing this complex from the solution on a solid support providing an antibody is sometimes referred to as "immunoprecipitation".

(ii) a protein domain which forms a complex with a second (macro)molecule e.g. glutathione-S-transferase, GST, (Smith, D. B. and Johnson, K. S. (1988) Gene 67, 31-40); bovine pancreatic trypsin inhibitor, BPTI, (Borijin, M. and Nathans, J. (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 337-341); maltose binding protein, MBP, (Bedouelle, H. & Duplay, P. (1988) Eur. J. Biochem. 171 541-549; Maina, C. V. et al. (1988) Gene 74 365-373), or (iii) a polypeptide sequence that can be biotinylated and thus made to interact with avidin or streptavidin (Schatz, P. J. (1993) Bio/Technology 11, 1138-1143).

The above mentioned tags have a high affinity for their respective ligands. Thus, harsh conditions are necessary to elute the bound protein from column (e.g. extreme pH values or chaotropic agents), which may impair the activity of the purified protein. For example, the high affinity of BPTI for trypsin ($K_d = 6 \times 10^{-14}$ M), makes elution of BPTI-tagged proteins from trypsin very difficult. Furthermore, the tags recognized by specific monoclonal antibodies require columns which are expensive and/or difficult to prepare. Moreover, when the tag is a protein or a large part of protein, such as MBP, BPTI or GST, properties of fusion recombinant proteins place limits on the possible applications of the obtained proteins.

Another strategy for the purification of recombinant proteins is purification by immobilized metal ion affinity chromatography (IMAC). In 1975, Porath introduced IMAC for fractionating proteins [J. Porath, J. Carlsson, I. Olsson, and G. Belfrage, Nature (London) 258, 598-599 (1975)]. In Porath's work, IMAC consists of derivatizing a resin with iminodiacetic acid (IDA) and chelating metal ions to the IDA-derivatized resin. Proteins bind to the metal ions through unoccupied coordination sites and are immobilized on the column. Since then, workers have used ligands other than IDA to chelate metal ions to resins. Studies with serum proteins report IMAC to be a specific and selective separation technique [J. Porath and B. Olin, Biochemistry 22, 1621-1630 (1983)]. For IMAC purification of a recombinant protein, a poly-histidine tag is

fused with a recombinant protein. Poly-histidine tag allows the recombinant protein to bind to nickel- (or cobalt-) chelating resin (Skerra, A., Pfitzinger, I. and Pluckthun, A. (1991) BioTechnology, 9, 273-278;. Unfortunately, commercially available nickel- and cobalt-chelating resins are relatively expensive.

An additional problem with column based chromatography, particularly in respect of delicate or multi-component targets, is that the physical journey through the resin bed has the potential to disrupt the protein or protein assembly. This may present a problem in the biochemical analysis of activity or of protein complexes. Even quite stable protein assemblies partition between different molecular weight fractions on size exclusion columns, and affinity chromatography often produces heterogeneous elution profiles. Thus purification of targets such as multi-protein complexes; labile (protease or thermally susceptible) polypeptides, and unusually high or low molecular weight proteins can be particularly labour intensive.

One biochemical approach for the identification of protein networks is disclosed by Dziembowski and Seraphin (FEBS Letters 556 (2004) 1-6)) and Rigaut *et al.* (Nature Biotechnology 17 (1999) pp 1030-1032). In this strategy, a protein is genetically tagged twice and the wild type gene replaced selectively by homologous recombination (referred to as Tandem Affinity Purification or "TAP" tagging for short).In this way, populations of cells are generated that express only the tagged version of a given protein target at natural levels. The tags provide a means of purifying the target protein in conjunction with its network partners. The identity of these partners is provided by mass spectrometry and bioinformatics. However the downstream processing (purification of the protein network) utilizes column chromatography and is therefore in principle invasive.

One solution to the problem of column fouling has been proposed by Lilius *et al.* (1991) Eur. J. Biochem *198*, 499-504. That paper discusses a particular combination of polyhistidine tags in conjunction with a metal chelators (such as ethylene glycol-bis(beta-aminoethyl ether) N,N,N',N" -tetra-acetic acid, EGTA(Zn)2) to selectively precipitate proteins. The precipitation process is, in mechanistic terms, very similar to selective immunoprecipitation, in which a (bivalent) polypeptide specific antibody forms an extended non-covalent aggregate which, beyond a threshold concentration, becomes insoluble. The system disclosed in Lilius is referred to therein as “affinity cross linking” and requires that each protein has at least two tags, and that the proteins are cross linked into “linear polymeric chains” which can in turn form a large three dimensional complex, which complex is precipitated at a threshold concentration. The system is suggested as being particularly suitable for bis-chelate cross

linking agents when used with oligomeric proteins (dimeric, or even higher) so that many cross links are formed and low molecular-mass aggregates can be avoided.

Thus it can be seen that novel materials for, and methods of, target enrichment, for example utilising different columns or which were column free, while having the capacity to also be selective and operate physiological or near physiological conditions, would provide a contribution to the art.

## Disclosure of the invention

The present invention describes an improved method for producing and purifying soluble targets, such as recombinant proteins, from a sample. The invention employs "tags" which can be appended to the target. By addition of a tag-binding entity (or "ligand", the terms are used interchangeably) to the tagged target, the solubility of the target can be reduced and it can be selectively precipitated. This can be performed in the absence of chromatographic resin or other separation media and represents a potentially powerful tool for enrichment and purification. The use of a non-protein binding entity (as opposed to, say, an antibody) avoids the requirement for having to disrupt strong protein:protein interactions to recover targets, or to remove free antibody from the resulting product.

Preferred tags are metal binding peptide tags, wherein addition of appropriate metal salts to the tagged target causes the reduction in solubility and "co-precipitation" with the metal salts. The process can be reversed by extracting or desorbing the metal from the tag using, for example, other anions or chelators. Where desired, the tags of the present invention may also be used in chromatographic methods. As described below, the most preferred tags are calcium-binding affinity polypeptide tags and the methods for purifying the tagged proteins.

It was known in the art for a very long time that some calcium-binding proteins (such as blood coagulation factors II, VII, IX, and X, thrombomodulin and anticoagulant proteins C and S) could be co-precipitated with calcium or barium poorly soluble salts (Stenflo, G., et al. (1976) J. Biol. Chem., v.251, pp.355-363; DiScipio,R.G., et al. (1977) Biochemistry, v.16, pp.698-706; Dahlback, B., et al. (1990) J. Biol. Chem., v.265, pp.18481-89). This property was known to arise out of EGF-like calcium-binding domains of the proteins and was used for purification of them (Broze, G.J., et al. (1980) J. Biol. Chem., v.255, pp.1242-1247). However no suggestion was made in the papers to that such domains could be used as heterologous tags for the purification of proteins which do not naturally possess them and for which they are not

physiologically relevant.

Unlike the method of Lilius *et al.* (1991, supra) the precipitation methods of the present invention do not require the cross linking of multivalent (tagged) proteins with bis-chelate ligands in order to reduce solubility. Rather it is the binding of the metal ion or salt to the tagged protein which itself causes the reduction in solubility, and selective co-precipitation without the necessity to form large 3 dimensional cross-linked complexes. The metal ion or salt is thus desirably not a bis-chelate (each of which would bind a tag entity in a 1:2 ratio and form a linear polymer of oligomeric proteins) – indeed such may undesirable, for example where the intention is to co-precipitate a target entity with an unknown binding partner.

The present invention further provides methods for producing and purifying recombinant proteins in a form of a fusion with a polyvalent/divalent metal ion-binding artificial polypeptide tag (affinity tag). The invention further provides *inter alia* affinity tags, expression vectors, transformed microorganisms, methods, reagents and kits for production, isolation and purification of recombinant proteins.

The invention particularly provides for the linking of a target coding gene sequence of a desired protein to the coding DNA sequence of the tag, expression of the hybrid gene, and purification of the fused protein by selective co-precipitation of the tagged protein with poorly soluble polyvalent\divalent metal (e.g. calcium or barium) salts. The invention further provides fusion proteins which contain one or two affinity polypeptide tags, which tags have the high affinity for polyvalent, in particular, divalent metal ions, such as $Ca^{2+}$ or $Ba^{2+}$, and a desired polypeptide or protein linked directly or indirectly to this/these affinity tag(s); a process for their preparation by means of recombinant DNA technology and a process for their purification by means of co-precipitation with slightly soluble salts. Without wishing to be bound by theory, a schematic representation of one embodiment of the present invention utilising calcium phosphate is shown in Figure 7.

These and other aspects of the present invention will now be described in more detail.

In one aspect the invention provides for a method of reducing the solubility of a soluble target entity from a sample, such as to facilitate recovery, enrichment, or purification thereof, by: (i) appending a heterologous tag entity to the target entity to form a tagged target, (ii) contacting the heterologous tag with a non-protein tag-binding entity which binds to the tag-entity causing reduction in solubility of the tagged target.

As discussed in more detail below, the non-protein tag-binding entity is a soluble or partially soluble metal salt and the method can be performed in the absence of chromatographic resin.

The invention further provides a method of reducing the solubility of a tagged target from a sample, such as to facilitate recovery, enrichment, or purification thereof, wherein the tagged target is a soluble target comprising a heterologous tag entity, the method comprising contacting the heterologous tag with the non-protein tag-binding entity which binds to the tag-entity causing reduction in solubility of the tagged target.

Preferably the methods of the present invention actually cause precipitation of the tagged target i.e. the condensation of a solid from a solution when the solution is supersaturated, whereupon the solid forms from the solute phase, and usually sinks to the bottom of the solution.

For example the methods may cause at least 10, 20, 30, 40, 40, 60, 70, 80, 90% or more of the tagged target to precipitate.

The method differs from conventional precipitation using metal salts or ammonium sulphate because the tag-binding entity preferentially or selectively binds to the tag-entity. For example, certain tags described herein capable of binding metals are based on modified "Ca-binding EGF-like" motifs. These Ca-binding EGF-like domains are known to bind metal (e.g. $Ca^{2+}$ ions) with an affinity of around 1 μM. Preferred tags of the invention will have an affinity for the tag-binding entity of this order, or an even higher affinity, under the conditions used to effect the precipitation.

The invention further provides for methods of purifying or enriching the tagged target or target from a sample for example by washing or recovering the precipitate, and optionally resolubilising it.

More specifically, the invention provides a process for producing tagged-target in enriched form from a sample, or detecting it in a sample, which process comprises:
(i) precipitating the target entity by performing a method as described above,
(ii) separating precipitate from the sample,
(iii) optionally resolubilising it e.g. by removal of the tag or tag-binding species,
(iv) optionally further enriching the target.

The invention further provides methods producing or detecting the binding partner of a target entity from a sample, the methods comprising (i) enriching or recovering the target entity from the sample as described above; (ii) determining if one or more binding partners are bound thereto, (iii) optionally isolating and\or identifying the binding partner(s).

The invention further provides a method of screening an agent for its ability to modulate a biological activity of a target, the method comprising: (i) enriching or recovering the target entity from the sample by performing the method of as described above; (ii) comparing the biological activity of the target in the presence and absence of the agent. The method may further comprise providing the modulator (inhibitor or stimulator) in isolated form.

In one aspect, the present invention provides the tags described herein (e.g. Calcium-binding affinity polypeptide artificial tags) and the expression vectors *for use* in obtaining the tagged recombinant proteins/polypeptides. In another aspect, the present invention provides the methods and the kits of reagents *for use* in purifying the tagged recombinant proteins/polypeptides.

***Definitions***

In order to provide a clear and consistent understanding of the specification and claims, the following definitions are provided.

The term "heterologous" means that the tag and target are non-naturally occurring together in a single entity (e.g. polypeptide chain) in nature but have been combined artificially by human intervention. For example where the tag and target are part of a single polypeptide chain, this may have been produced by recombinant engineering, i.e. by human intervention.

As used herein the terms "purify", "purifying" or "purified" mean making substantially free of other compounds.

The terms "thermostable enzyme" and "thermostable protein", as used herein, refer to an enzyme or protein, which is stable to heating up to 90 °C.

The term "recombinant" or "artificial" protein refers to a protein obtained by means of recombinant DNA techniques. Recombinant nucleic acid sequences are not found together in

nature (do not run contiguously) have been ligated or otherwise combined artificially. Alternatively they may have been synthesised directly e.g. using an automated synthesiser.

The terms "fusion", "fusion protein", "fused protein" or "fused polypeptide" are used herein in the context of joining of two heterologous polypeptides/proteins to form a molecule. The preferred mode of joining is via a peptide bond, so that a "fusion protein" may be produce by expression from coding nucleic acid thereof.

The term "nucleic acids", as used herein may include DNA, RNA, modified nucleic acids or nucleic acid analogs (e.g. peptide nucleic acid). Where a DNA sequence is specified, e.g. with reference to a figure, unless context requires otherwise the RNA equivalent, with U substituted for T where it occurs, is encompassed. Nucleic acid molecules according to the present invention may be provided isolated and/or purified from their natural environment or source, in substantially pure or homogeneous form. The nucleic acid molecules may be wholly or partially synthetic.

The term "DNA sequence" refers to a single- or double-stranded polymer of deoxyribonucleotide bases read from the 5' to the 3' end.

The terms “oligonucleotide primer”, “oligonucleotide” or “primer” refer to a single-stranded polymer of deoxyribonucleotides.

The terms "artificial peptide/polypeptide tag" and “artificial tag” as used herein are meant a polypeptide, which is artificially linked or can be linked to N-terminal or C-terminal end of a target protein/polypeptide, for example, by means of recombinant DNA techniques.

The terms “affinity peptide/polypeptide”, "affinity peptide/polypeptide tag" and “affinity tag”, as used herein, are meant an amino acid sequence that affinity binds with a ligand. Affinity polypeptide tags of this invention are calcium-binding tags.

The term " calcium-binding tag", maybe used for brevity herein but refers to an amino acid sequence that affinity-chelates calcium ions and optionally other polyvalent/divalent metal ions of metals selected from the group, which can include, but is not limited to, barium, calcium, cadmium, cobalt, chrome, copper, iron, mercury, magnesium, manganese, nickel, lead, zinc and others.

The term "tagged protein" or "tagged polypeptide ", as used herein, refers to protein/polypeptide fused to the artificial tag.

***Preferred tags***

Preferably the tag-entity is a protein, polypeptide or peptide (the terms are used interchangeably).

As described below, preferably the tag-*binding* entity or ligand is a metal ion or salt and the tag entity is capable of coordinating, complexing, chelating or otherwise bonding a metal ion or salt. Preferred metals and salts are discussed in more detail hereinafter.

In one embodiment the tag-entity is a "Ca-binding EGF-like" peptide or derivative thereof. In some embodiments these may share homology with "EF-hand" motifs, which are are so named owing to the presence of two alpha helices: one containing a key Glu (E) residue and the other helix a unique Phe (F) residue, spatially arranged via a short linker segment into a "hand-like" configuration. The cleft of the out-stretched thumb and finger binds $Ca^{2+}$ with high affinity. This motif can be appended to any target molecules and, in conjunction with a bio-inorganic phenomenon in which $Ca^{2+}$ ions that interact with the EF hand, can be reversibly precipitated and resolubilised in the presence of a range of co-precipitants.

Without wishing to be bound by theory, in naturally occurring proteins (such as Calmodulin) the coordination by $Ca^{2+}$ by EF hand proteins is believed to trigger a significant conformational change. Thus the coordination of the Calcium may stabilise the structure of the tag and lead to the aggregation of the recombinant protein in a Calcium dependent manner, analogous to the aggregation of blood clotting factors.

A derivative Ca-binding EGF-like sequence may include within any of the sequences shown herein a single amino acid or 2, 3, 4, 5 or more changes. Preferred changes will be conservative changes (i.e. substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as arginine for lysine, glutamic for aspartic acid, or glutamine for asparagine) although non-conservative changes may also be tolerated in certain regions of the peptide not required for metal coordination.

Thus in one embodiment the tag entity comprises, consists, or consists essentially of the following sequence of amino acids residues (a tag "consists essentially of" a sequence is that sequence is 'of the essence of' the molecule, generally making up more than 50% of it).

$X_m$-X1-X2-X3-X4-X5-cys-$X_z$-cys-X8-X9-$X_n$

wherein:

$\mathbf{X_m}$ = a terminal sequence of m amino acids which can be any whole number from 0 to 100 (more preferably is in range selected from: between 0 and 90, 80, 70, 60, 50, 40, 30, 20, 10 or 5)

X1 = an acidic amino acid

X2 = an acidic amino acid

X3 = a hydrophobic amino acid

X4 = an acidic amino acid

X5 = an acidic amino acid

$\mathbf{X_z}$ = a spacer sequence of z amino acids which can be any whole number from 1 to 25 (more preferably is in range selected from: between 1 and 20, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3 or 2).

X8 = any amino acid, preferably hydrophobic or basic, more preferably basic

X9 = any amino acid, preferably polar or acidic, more preferably acidic

$\mathbf{X_n}$ = a terminal sequence of n amino acids which can be any whole number from 0 to 100 (more preferably is in range selected from: between 0 and 90, 80, 70, 60, 50, 40, 30, 20, 10 or 5)

Preferred "acidic" amino acids are D or E or artificial acidic amino acids.

Preferred "hydrophobic" amino acids are G, A, V, L, or I, or artificial hydrophobic amino acids. Preferred are V,L, or I .

Preferred "basic" amino acids are K, R, or H or artificial basic amino acids.

Preferred "polar" amino acids are N or Q or artificial polar amino acids.

Preferably $\mathbf{X_m}$ and $\mathbf{X_n}$ are absent (i.e. m and n are zero).

In one embodiment, $X_z$ may comprise a F residue.

The tags having these sequences may be termed herein "affinity polypeptide tags".

Preferred affinity polypeptide tags of the invention are defined by the general amino acid sequence (SEQ ID No 1):

$X_m$-(E or D)-(D or E)-(V or I or L)-(D or E)-(E or D)-C-$X_z$-C-(K or R or H)-(D or E)-$X_n$;
wherein $X_m$, $X_z$ and $X_n$ can be any amino acid sequences comprising ***m, z*** and ***n*** amino acid residuals, respectively, and wherein ***m*** and ***n*** can be any whole numbers from 0 to 100 (more preferably is in range selected from: between 0 and 90, 80, 70, 60, 50, 40, 30, 20, 10 or 5), and ***z*** can be any whole number from 1 to 25 (more preferably is in range selected from: between 1 and 20, 15, 14,13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3 or 2).

In a preferred embodiment, ***z*** is 7, 8 or 9. Most preferably it is 8.

In a preferred embodiment, the affinity tags of the invention have the following amino acid sequence (SEQ ID No 2):

$X_m$-E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D-$X_n$;

Other exemplary embodiments of affinity tags of the present invention may be as follows:

$X_m$ -(E or D)-(D or E)-(V or I or L)-(D or E)-(E or D)-C-Xz-C-(K or R or H)-(D or E) -$X_n$

$X_m$ -E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D-$X_n$

$X_m$ -D-D-V-D-E-C-S-E-N-S-S-A-Q-L-C-K-D-$X_n$

$X_m$ -E-D-V-E-E-C-S-E-S-G-S-G-Q-L-C-R-D-$X_n$

$X_m$ -D-D-I-D-E-C-S-E-S-M-S-A-Q-I-C-R-E-$X_n$

$X_m$ -E-E-L-D-E-C-S-D-G-M--Q-L-C-R-D-$X_n$

$X_m$ -E-D-V-E-D-C-S-E-G-S-G-S-M-S-A-A-L-C-R-E-$X_n$

$X_m$ -E-D-L-D-E-C-S-E-N-M S-G-S-A-S-V-K-C-H-D-$X_n$

$X_m$ -E-D-V-D-E-C-S-E-N-M-S-A-Q-H-C-K-D-G-E-$X_n$

$X_m$ -D-D-V-D-E-C-A-E-N-L-S-A-Q-L-C-R-D-$X_n$

($X_m$ -D-D-V-D-E-C-A-E-N-L-C-A-Q-L-C-R-D)-YPGGYSCYCDGKKGFKLAQDQKSCE

wherein $X_m$ and $X_n$ can be any amino acid sequences comprising ***m*** and ***n*** amino acid residuals, respectively; and wherein ***m*** and ***n*** can be any whole numbers as described above.

***Preferred tag-binding entities or ligands***

The above described polypeptide tags of the invention can facilitate the binding of tagged proteins to suitable tag binding entities (also referred to herein as ligands) such as polyvalent and/or divalent metal ions. In particular, the ligand may be selected from the group of divalent ions: $Ba^{2+}$, $Ca^{2+}$, $Cd^{2+}$, $Co^{2+}$, $Cr^{2+}$, $Cu^{2+}$, $Fe^{2+}$, $Hg^{2+}$, $Hg_2^{2+}$, $Mg^{2+}$, $Mn^{2+}$, $Ni^{2+}$, $Pb^{2+}$, $Ra^{2+}$, $Sr^{2+}$, $Zn^{2+}$ and others, or from the group of polyvalent ions: $Sc^{3+}$, $Y^{3+}$, $La^{3+}$ and other lanthanides, $Ac^{3+}$ and other actinides, $In^{3+}$ ,$Bi^{3+}$ and others.

In a preferred embodiment, the ligand is $Ca^{2+}$, $Sr^{2+}$, $Ba^{2+}$, $Y^{3+}$, $La^{3+}$ and $In^{3+}$.

Thus the present invention provides for the use of the affinity tags disclosed herein for achieving co-precipitation of a tagged protein with a slightly soluble salt containing the ligand.

A slightly soluble salt for selective co-precipitation of the tagged protein can be a slightly soluble salt of a polyvalent or a divalent metal. In particular, the slightly soluble salt may be selected from salts of barium, calcium, cadmium, cobalt, chrome, copper, iron, mercury, magnesium, manganese, nickel, lead, zinc, strontium, yttrium, lanthanum and other lanthanides, indium, bismuth and others.

In a preferred embodiment, the slightly soluble salt can be a salt of barium, strontium, calcium, yttrium or lanthanum such as calcium or barium citrate, oxalate, sulphate and, particularly, phosphate and hydrogen phosphate.

Selective co-precipitation of the tagged protein(s) can be performed by means of generating a poorly soluble salt of calcium or barium or other polyvalent/divalent metals in the presence of the tagged protein(s). Generation of the slightly soluble salt can be achieved by mixing (or adding, in either order):

1) a solution containing a soluble salt of the metal selected from the group, which includes, but are not limited to barium, calcium, cadmium, cobalt, chrome, copper, iron, mercury,

magnesium, manganese, nickel, lead and zinc, with

2) another solution containing a soluble salt selected from the group, which includes, but are not limited to phosphates, hydrogen phosphates, citrates, oxalates, and sulphates.

The mixing of the tagged protein and the solution in step 1) may be performed under conditions and concentrations to permit the selective binding of the metal ion to the tag e.g. by use of relatively soluble chloride salt. Without wishing to be bound by theory, it is understood that this process is driven by entropy. Those skilled in the art, in the light of the present disclosure, will be readily able to arrive at suitable conditions to achieve the initial mixing and coordination without undue burden.

The mixing of the metal-ion tagged protein and the solution in step 2) may be performed under conditions and concentrations to reduce the solubility of the metal ion and tag. For example a counter ion may be added (or the concentration may be increased) such as to form a metal salt or coordinate complex which is less soluble than the metal species formed in step 1). For example a sparingly or slightly soluble salt such as a phosphate or sulphate may be used (where the $K_{sp}$ for the calcium salts is small i.e. less than $10^{-20}$). This will cause precipitation of the metal-ion tagged protein. Those skilled in the art will appreciate the absolute $K_{sp}$ may not be critical, since the point of precipitation will be dictated by a number of factors affecting solubility. However in the light of the present disclosure, those skilled in the art will be readily able to arrive at suitable conditions to achieve the precipitation by reference to known $K_{sp}$ values of the metal salts in question, and experimental assessment of concentrations. However preferred salts may have a $K_{sp}$ of around $10^{-18}$ or lower, e.g. less than $10^{-19}$ or $10^{-20}$.

If desired, the precipitate may be washed in suitable concentration of the metal-ion salt so as to remove impurities without desorbing the metal ion. This concentration will generally be the same as that used for the precipitation, although may be slightly higher or lower. Those skilled in the art, in the light of the present disclosure, will be readily able to arrive at suitable conditions and concentrations to wash the precipitate without resolublising the target.

In other embodiments, the mixing and co-precipitation may be performed in a single step using the sparingly soluble salt at appropriate concentration, although this is less preferred.

In a preferred embodiment, selective co-precipitation of the tagged protein(s) can be performed in the presence of the tagged protein(s) by mixing a solution containing a soluble

salt of calcium or barium with a solution containing a soluble hydrogen phosphate(s) or citrate(s) or oxalate(s). However other salts or counterions (e.g. having a $K_{sp}$ between the soluble salt such as the chloride, and the phosphate, citrate, or oxalate salt) may be used.

Novel tagged-targets of the invention may also, if desired according to circumstance, be used in immobilized metal ion affinity chromatography (IMAC) or adsorption chromatography. This is discussed hereinafter.

***Preferred targets***

Preferably the target is a biomolecule. Most preferably the target is a polypeptide.

The polypeptide may form part of a larger soluble complex, aggregate, sub-cellular structure, or virus. Thus in one embodiment the target is a protein complex comprising a least 2 binding partners, where the heterologous tag entity is appended to 1 of the binding partners.

In another embodiment the target is a viral protein e.g. as prepared or cultured for use in a vaccine.

The affinity binding of tagged proteins with the ligands can provide selective immobilization, adsorption and/or (in preferred forms) precipitation of tagged targets. The use of the precipitation methods of the present invention may be particularly preferred for the enrichment of the target plus its binding partners, since this avoids the relatively harsh fractionating environment of a resin-based approach.

Selective co-precipitation of the tagged proteins of the invention can be performed both in denaturing and non-denaturing conditions for purification of recombinant proteins in denatured and native forms accordingly. Thus the tags of the invention provide the purification by selective co-precipitation of denatured tagged proteins dissolved, for example, in solutions containing guanidine hydrochloride or urea. In a preferred embodiment of the invention, the denatured tagged proteins can be dissolved in solutions containing about 3 M – 6 M guanidine hydrochloride or/and 4 M – 8 M urea.

The tags of the invention allow purifying thermostable recombinant enzymes/proteins after preheating these proteins, e.g. for denaturation of host cell proteins. In a preferred embodiment of the invention, the preheating can be performed at 60-90°C for 5-30 min.

The preferred thermostable recombinant enzymes, which can be produced and purified by means of the methods of the present invention, include, but are not limited to thermostable DNA polymerases, such as Taq, Tth, Tfl, Tli, Pfu and other thermostable DNA polymerases.

In a preferred embodiment of the invention, the tagged thermostable DNA polymerases expressed in E.coli can be purified by means of the selective co-precipitation following preheating at 70-80°C for denaturation of host cell proteins.

***Preferred methods and materials for tagging***

The tagging of the target may be achieved by directly appending the tag to the target, optionally via a linker, by methods conventional in the art. A linker region may be included which may facilitate the folding of the domains into their natural conformation by reducing steric hindrance between the domains, or facilitating removal of the tag.

The link between the tag and target may be covalent (e.g. peptide, disulphide bridge), ionic or via any other interaction which maintains the association through the methods of the invention for reducing solubility, precipitating, enriching or otherwise purifying as described herein.

For example peptides may be joined to nucleic acid targets by means of Peptide Nucleic Acid (PNA) chemistry – see e.g. "Peptide Nucleic Acid: Protocols and Applications (Second Edition)" Publisher: Horizon Bioscience; Editor: Peter E. Nielsen. Other conventional means of joining peptides and other moieties is via the avidin and biotin interaction, which is well known to those skilled in the art.

Thus a polypeptide linker used for the link of target protein/polypeptide to the affinity tag may contain one or more site(s) which can be recognized and cleaved by a proteolytic agent(s). This site can be used for cutting the fused protein by the proteolytic agent.

In a preferred embodiment, the above site has the amino acid sequence (SEQ ID No 4): Ile-Glu-Gly-Arg, and can be cleaved at the carboxyl side of the arginine residue by using blood coagulation factor Xa.

In another embodiment, the above site has the amino acid sequence (SEQ ID No 17): Glu-X-X-Tyr-X-Gln/Ser, and can be cleaved by using the tomato etch virus (TEV) protease. This

enzyme is obtained commercially (e.g. from Invitrogen).

Preferably, where the target and tag are peptides, it is preferred that the tagging is by expression of a fusion protein comprising said target incorporating the tag

The invention further provides a process for making a tagged target comprising: (i) providing a target entity (ii) appending a heterologous tag entity as described above to it.

In various aspects the present invention further provides: a tagged target comprising: (i) a target entity, and (ii) a heterologous tag entity as described above (e.g. a native or modified Ca-binding EGF-like motif), optionally also comprising a tag binding species as described above (e.g. a partially soluble). This may be provided as a precipitate in a sample substantially free of chromatographic resin.

This invention provides fusion proteins/polypeptides which contain a target protein or polypeptide linked to at least one affinity polypeptide tag of the invention. Fusion polypeptide or protein can be linked directly or indirectly to this/these affinity tag(s).

In the fusion proteins/polypeptides of this invention, the N-terminal amino acid of the affinity polypeptide tag can be linked to the C-terminal amino acid of a target protein/polypeptide. In another embodiment, the C-terminal amino acid of the tag can be linked to the N-terminal amino acid of a target protein/polypeptide.

In a further embodiment, the fusion proteins/polypeptides of the invention can comprise two polypeptide tags linked to the N- and C-terminal amino acids of a target protein/polypeptide, wherein at least one polypeptide tag is the affinity tag of the invention, and another tag can be the affinity tag of the invention or can be different. For example, the fusion protein may comprise the affinity tag of the invention and a poly-His tag linked to a target protein.

In another embodiment, the fusion proteins/polypeptides of the invention can comprise two or more interconnected polypeptide tags linked to the N- or C-terminal amino acids of a target protein/polypeptide, wherein at least one polypeptide tag is the affinity tag of the invention.

A target protein/polypeptide, in the fusion proteins/polypeptides of the invention, can be linked to the affinity tag(s) directly, via a peptide bond, or indirectly, for example, through a peptide/polypeptide linker.

***Optional further treatments***

All methods of the invention may optionally be combined with further upstream or downstream steps depending on the required application (for example in respect of desired purity, yield, and so on). For example, if the target is required in ng quantities for a highly sensitive assay, then control over the concentration of the eluted material is perhaps not as critical as purity. For experiments requiring $\mu$g to mg quantities for mass spectrometry concentrated salt free samples are desirable. Where a more concentrated protein target is required, this can be achieved by ammonium sulphate precipitation, centrifugation and re-solubilisation at a suitable concentration.

The methods of the invention may further comprise the step of resolubilising the tagged target by either:
(i) removing the tag-entity,
(ii) removing the tag-binding entity,
(iii) dilution.

For example where the tagged target is a polypeptide the tag entity may be removed by proteolysis as described above. In certain embodiments removal by a selective protease may not be preferred if it is desired to avoid contaminating material with a protease preparation. Alternatively the protease itself maybe tagged to facilitate removal.

In another embodiment, the tag-binding entity is a metal ion and this is removed from the tagged target by addition of counter ion, wherein the metal salt of the counter ion is largely insoluble.

By the introduction of a higher concentration (than was used in the precipitation step) of a counter ion such as phosphate or sulphate (where the Ksp is small i.e. less than $10^{-20}$) the bound metal ion may be displaced (or extracted or desorped) into an insoluble form which is no longer available for complexation with the tag and the recombinant protein re-solubilises. The precipitated calcium salt may then be removed by simple centrifugation.

***Nucleic acids, vectors and host cells***

The fusion proteins/polypeptides comprising the affinity tag(s) of the invention can be obtained

by means of recombinant DNA technology or can be produced by means of classical methodology of solution or solid phase synthesis.

The fusions of this invention preferably are obtained by expression of the hybrid genes coding for desired polypeptide/protein fused to the polypeptide tag(s) of the invention.

Thus the present invention provides a novel approach to production and purification of desired polypeptides or proteins obtained by recombinant DNA techniques. The protein of interest is produced by constructing an expression vector containing a gene encoding the target protein molecule fused to the affinity polypeptide tag of the invention and expressing the fusion in an appropriate host. The tagged recombinant protein can be purified by contact with the appropriate tag-binding ligand – for example by selective co-precipitation, which is described herein.

Thus the invention also provides expression vectors and transformed microorganisms for production of recombinant proteins fused to the affinity polypeptide tags, which are described above.

Suitable vectors for expression of genes encoding the fused proteins can be chosen or constructed, containing appropriate regulatory DNA sequences, such as promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate. The vectors of the invention also contain DNA sequence(s) coding for at least one polypeptide tag of the affinity polypeptide artificial tags claimed in the present invention or DNA sequence(s) coding for protein(s)/polypeptide(s) fused to the tag(s) of the invention.

Thus in particular aspects the invention provides hybrid genes comprising DNA sequences coding for the desired polypeptide or protein and the affinity artificial tag(s) of the invention. The hybrid genes of the invention can also comprise appropriate expression regulatory DNA sequences, such as promoter sequences, 5'- and 3'- untranslated regions, terminator fragments and others.

The hybrid genes can be prepared by means of standard recombinant DNA methods, which are well known (see, for example, Molecular Cloning: a Laboratory Manual: 2nd edition, Sambrook et al, 1989, Cold Spring Harbor Laboratory Press). For example, a DNA sequence coding for the desired affinity peptide can be first synthesized and then linked to a DNA

sequence coding for the target polypeptide or protein. The hybrid gene thus obtained can be incorporated into an expression vector, such as plasmids pQE, pET or other commercial or generally accessible plasmids, using standard methods.

For the alternative obtaining hybrid genes coding for the fusions of the invention, a DNA sequence coding for the target polypeptide or protein can be incorporated into a specially constructed expression vector containing pre-inserted DNA sequence coding for the affinity peptide of the invention. In this case, the DNA sequence coding for the target polypeptide/protein is first obtained and then linked to DNA sequence coding for the tag of the invention by means of inserting it into the expression vector.

Methods for the expression of the hybrid genes coding for the fusion proteins of this invention are also described by Sambrook et al., supra. They comprise the following procedures: (a) transformation of a suitable host organism, preferably E. coli, with an expression vector in which the hybrid gene is operatively linked to an expression control sequence, and (b) cultivation of the transformed host organism under suitable growth conditions.

Also provided are processes for making the recombinant vectors comprising the step of ligating or otherwise linking a nucleic acid sequence encoding the tag-entity to a sequence encoding a target polypeptide.

The transformed microorganisms of the invention contain an expression vector comprising the hybrid gene of the invention, wherein DNA sequence coding for the target protein/polypeptide linked to the tag(s) claimed in the invention. The transformed microorganisms can be obtained by means of transformation procedures depending on the host used, which are well known. For details see, for example, Molecular Cloning: a Laboratory Manual: 2nd edition, Sambrook et al, 1989, Cold Spring Harbor Laboratory Press. A variety of different host cells are frequently employed by those skilled in the art, including bacteria, mammalian cells, yeast and baculovirus systems. Mammalian cell lines available in the art for expression of a heterologous polypeptide include Chinese hamster ovary cells, HeLa cells, COS cells and others. A common, preferred bacterial host is E. coli.

As mentioned above, in a preferred embodiment of the invention, recombinant protein(s)/polypeptide(s) fused to the tag(s) of the invention can be purified by means of the selective co-precipitation method described herein. This invention also provides kits of reagents for recovery of proteins from host cells and/or purification of tagged recombinant

proteins by means of the selective co-precipitation method.

***Kits***

This invention provides kits of reagents for recovery of proteins from host cells and/or purification of tagged recombinant proteins by means of the selective co-precipitation method. Kits of the invention comprise separate containers with reaction components that are necessary to generate the slightly soluble salt for selective co-precipitation of the tagged proteins.

For example, a kit can comprise:

1) a solution containing a soluble salt of the metal selected from the group, which includes, but are not limited to barium, calcium, cadmium, cobalt, chrome, copper, iron, mercury, magnesium, manganese, nickel, lead, strontium, zinc, yttrium, lanthanum and other lanthanides, indium, bismuth and others; and

2) another solution containing a soluble salt selected from the group, which includes, but are not limited to phosphate(s), hydrogen phosphate(s), citrate(s), oxalate(s), sulphate(s) and others. Mixing of this solutions generate precipitation of the slightly soluble salt and co-precipitation of the tagged proteins of the invention.

In a preferred embodiment, the kit can comprise a solution containing a soluble salt of calcium or barium and another solution containing a soluble hydrogen phosphate(s), or citrate(s), or oxalate(s).

The kit may contain a plurality of additional components. Among the additional components, one may include a reagent for lysis of host cells, such as lysozyme, lysostaphin or/and detergent(s). Other components may be nucleic acids or vectors e.g. encoding the tags of the present invention, and optionally other materials to facilitate the production of fusion proteins. Kits may also include instructions for use in the methods of the present invention.

Preferred detergents, which can be included into the kit, are non-ionic detergents. In particular, these detergents can be selected from the group, which include, but is not limited to alkyl glycosides or alkythioglycosides. For example, the kit can include hexyl- ,heptyl-, octyl-,

nonyl- or decyl-glucopyranoside (or -thioglucopyranoside).

***Use of novel tagged proteins in IMAC***

In an alternative embodiment of the invention the precipitation methods described above, the use of heterologous affinity tags as described above can provide purification of tagged proteins by affinity interaction with the ligand during immobilized metal ion affinity chromatography (IMAC) or adsorption chromatography. This therefore provides for the use of additional immobilised ligands in these processes compared to those in the art. For example, in preferred embodiments of this aspect, immobilized metal ion affinity chromatography (IMAC) can be performed by means of metal chelate affinity chromatography, wherein a metal may be selected from divalent or polyvalent metals, such as barium, calcium, cadmium, cobalt, chrome, copper, iron, mercury, magnesium, manganese, nickel, lead, zinc, yttrium, lanthanum and other lanthanides, indium, bismuth and others. In a preferred embodiment, IMAC may be performed by means of calcium, cadmium, strontium, yttrium, lanthanum or barium chelate affinity chromatography.

Adsorption chromatography of the tagged proteins can be performed by chromatography on a column of hydroxyapatite.

The chromatography purification of the tagged proteins may be performed in denaturing or non-denaturing conditions. For performing chromatography in denaturing conditions, chromatography buffers may contain guanidine hydrochloride or urea.

The following Examples and non-limiting Figures illustrate aspects of the invention. The Examples are illustrative of, but not binding on, the present invention. Other embodiments of the invention will occur to those skilled in the art in the light of these.

Any methods, preparations, solutions and such like, which are not specifically defined, may be found in Sambrook et al. All solutions are aqueous and made up in sterile, deionised water, unless otherwise specified. All enzymes were obtained from Promega Corporation.

The disclosure of all references cited herein, inasmuch as it may be used by those skilled in the art to carry out the invention, is hereby specifically incorporated herein by cross-reference.

All subtitles herein are for convenience only and are not to be construed as limiting the

disclosure in any way.

**Figures**

**FIG. 1**. Electrophoretic analysis of fractions I (bacterial lysates, lanes 1 and 3) and fractions II (proteins which were purified by the selective co-precipitation, lanes 2 and 4) obtained in Example 4 and containing GFP-ED protein of the invention (lanes 1 and 2) and ordinary GFP (lanes 3 and 4). Fig. 1 compares the yields of GFP-ED recombinant protein comprising the tag of the invention (lane 2) and the ordinary GFP non-comprising the tag of the invention (lane 4) obtainable by the selective co-precipitation method of purification.
The fractions were analysed via SDS-PAGE on a 10% gel.

**FIG. 2** Electrophoretic analysis of fraction I (bacterial lysate containing PPase-ED, lane 1) and fraction II (purified PPase-ED protein, lane 2) obtained in Example 5. The PPase-ED recombinant protein comprising the tag of the invention was purified by the selective co-precipitation method of the invention. The fractions were analysed by SDS-PAGE on a 10% gel. Lane **M**: protein molecular weight markers.

**FIG. 3** Electrophoretic analysis of fraction I (bacterial lysate, lane 1), fraction II (heated-up lysate, lane 2), and fraction III (purified Taq-ED protein, lane 3) obtained in Example 6 during purification of the Taq-ED recombinant protein. Taq-ED comprising the tag of the invention was purified from the cell lysate by denaturing E.coli host proteins at 75°C with subsequent selective co-precipitation of the tagged protein. The fractions were analysed via SDS-PAGE on a 10% gel. Lane **M**: protein molecular weight markers.

**FIG. 4** Electrophoretic analysis of the PCR product obtained using the purified Taq-ED DNA polymerase.
Lane 1: a 500-bp DNA fragment was amplified with Taq-ED DNA polymerase. Lane M: a DNA molecular weight marker.

**FIG. 5** Electrophoretic analysis of fractions I (bacterial lysates containing PPase-ED (lane 1) and GFP-ED (lane 3)) and fractions II (purified PPase-ED (lane 2) and GFP-ED (lane 4)) obtained in Example 8. GFP-ED and PPase-ED proteins comprising the tag of the invention were purified under denaturing conditions by the selective co-precipitation method. The fractions were analysed by SDS-PAGE on a 10% gel. Lane **M**: protein molecular weight markers.

**Fig. 6** Electrophoretic analysis of fractions obtained in Example 9 during two-step purification of GFP-ED protein by IMAC and selective co-precipitation methods.
Lane 1: fraction I, bacterial lysate containing GFP-ED. Lane 2: fraction II obtained after IMAC chromatography on $Ni^{2+}$ chelating resin. Lane 3: fraction III obtained after selective co-precipitation of GFP-ED from fraction II. Lane M: protein weight marker.

**FIG. 7** Schematic representation of one embodiment of the present invention utilising calcium phosphate.

**Examples**

**Example 1 - Creation of pQE-GFP and pQE-GFP-ED vectors.**

pQE-GFP-ED and pQE-GFP vectors were constructed on the basis of expression vector pQE-30 (Qiagen). pQE-GFP and pQE-GFP-ED contain recombinant DNA coding for the Green Fluorescent Protein (GFP) fused at its N-terminus to a 6xHis-tag. pQE-GFP-ED additionally contains a DNA fragment coding for a calcium-binding peptide tag of the invention (ED-peptide) linked to the C-terminus of GFP. The amino acid sequence of this calcium-binding peptide tag of the invention is given in SEQ ID No 3: -E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D

For the creation of pQE-GFP, the DNA fragment coding for GFP was amplified by means of PCR with primers: GFP-Bam 5'-ACATGGATCCGTGAGCAAGGGCGAGGAGCT and GFP-Sac 5'-TCTAGAGCTCGCATGCCTTGTACAGCTCGTCCATG. The nucleic acid sequence of the amplified DNA fragment is given in SEQ ID No 5. The obtained DNA fragment was inserted into pQE-30 vector between Bam H1 and Sac 1 restriction sites. The resulting plasmid pQE-GFP contains and can express recombinant DNA coding for GFP fused at its N-terminus to a 6xHis-tag. The nucleic acid sequence of pQE-GFP is given in SEQ ID No 6.

For the creation of pQE-GFP-ED, a DNA fragment coding for the calcium-binding peptide tag of the invention (ED-peptide, amino acid SEQ ID No 3) was synthesized and inserted into pQE-GFP vector between the Pst 1 and Hind 3 restriction sites. The nucleic acid sequence of the synthesized DNA fragment is given in SEQ ID No 7. The resulting plasmid pQE-GFP-ED contains and can express the recombinant DNA coding for GFP fused to N-terminal 6xHis-tag and C-terminal calcium-binding tag of the invention. The nucleic acid sequence of pQE-GFP-ED is given in SEQ ID No 8.

**Example 2 - Creation of pQE-PPase-ED vector.**

pQE-PPase-ED contains and can express recombinant DNA coding for Inorganic Pyrophosphatase of Escherichia coli (PPase) fused to N-terminal 6xHis-tag and to C-terminal calcium-binding tag (ED-peptide) of the invention. The amino acid sequence of ED-peptide is given in SEQ ID No 3.

For the creation of pQE-PPase-ED, DNA fragment coding for PPase was amplified by means of PCR with primers: PPase-Bam 5'-ACATGGATCCAGCTTACTCAACGTCCCTGCGGGT and PPase-Sal 5'- GACCGTCGACCTTTATTCTTTGCGCGCTCGAAG. The nucleic acid sequence of the amplified DNA fragment is given in SEQ ID No 9. The obtained DNA fragment was inserted into the pQE-GFP-ED vector between Bam H1 and Sal 1 restriction sites instead of the DNA fragment encoding GFP. The nucleic acid sequence of pQE-PPase-ED is given in SEQ ID No10.

**Example 3 - Creation of pQE-Taq-ED vector.**

pQE-Taq-ED contains and can express recombinant DNA coding for Taq DNA polymerase fused N-terminally to 6xHis-tag and to C-terminal calcium-binding tag (ED-peptide) of the invention. The amino acid sequence of ED-peptide is given in SEQ ID No 3.

For the creation of pQE-Taq-ED, DNA fragment coding for Taq polymerase was amplified by means of PCR with primers: Taq-Bgl 5'- ACATAGATCTAGGGGGATGCTGCCCCTCTTTG and Taq-Sal 5'- GACCGTCGACCCTCCTTGGCGGAGAGCCAGTC. The nucleic acid sequence of the amplified DNA fragment is given in SEQ ID No 11. The obtained DNA fragment was inserted into pQE-GFP-ED vector between Bam H1 and Sal 1 restriction sites instead of DNA fragment encoding GFP. The nucleic acid sequence of pQE-Taq-ED is given in SEQ ID No 12.

**Example 4 - Purification of recombinant Green Fluorescing Protein fused to the calcium-binding tag of the invention (GFP-ED protein) by means of the selective co-precipitation method.**

Fusion GFP-ED protein of the invention comprising GFP and C-terminally calcium-binding tag (ED-peptide) of the invention was obtained by expression of pQE-GFP-ED (see Example 1) in

E. coli strain JM 109. The amino acid sequence of GFP-ED is given in SEQ ID No 13.

Used as a negative control, recombinant GFP protein not comprising the calcium-binding tag of the invention was obtained by expression of pQE-GFP (see Example 1) in E. coli strain JM 109. The amino acid sequence of the recombinant GFP is given in SEQ ID No 14.

*Expression of Plasmid pQE-GFP-ED*

E. coli JM 109 cells were transformed with the plasmid pQE-GFP-ED according to the method of Dower et al. [1988, Nucl. Acid. Res., V.16, P.6127]. The transformed cells were grown to an optical density of $A_{600}$ = 0.4 in 15 ml of LB medium containing ampicillin (75 mkg/ml) at 37°C. Expression of the gene encoding the tagged GFP-ED protein was induced by adding IPTG to concentration 0.75 mM. The cells were further incubated for 8 h at 35°C. Cells were harvested by centrifugation.

*Expression of Plasmid pQE-GFP*

This was performed in the same manner as expression of pQE-GFP-ED.

*Purification of the GFP-ED recombinant protein*

The GFP-ED protein comprising the calcium-binding tag of the invention was purified from cells by means of the selective co-precipitation method.

The bacterial cells containing the recombinant protein of the invention were collected by centrifugation from 15 ml of medium. The lysis of the cells was performed by resuspending the bacterial pellet in 0.5 ml of lysis buffer and incubating on ice for 3-5 minutes. The lysis buffer contained: 150 mM $K_2HPO_4$; 300 mM KCl; 1.5% Octyl-D-glucopyranoside; 0.2% Poly(ethyleneimine). The bacterial lysate was centrifuged at 12,000g for 5 minutes in a microfuge. The supernatant was transferred to a fresh tube (Fraction I). 100 $\mu$l of 1 M $CaCl_2$ solution were added to lysate and mixed by vortexing. The tube was incubated at 0°C on ice for 10 minutes to mature the precipitate. The precipitate was harvested by centrifugation for 5 minutes at 12,000g in a microfuge. The supernatant was removed by aspiration. The precipitate was washed by resuspending in 0.5 ml of a washing buffer containing: 1 M KCl; 0.18% Octyl-D-glucopyranoside; 100 mM $Na_2SO_4$; 20 mM Tris-HCl (pH=7.0) whilst maintaining the $CaCl_2$ concentration to that in the previous step, of around 180 mM.

After washing the precipitate was harvested by centrifugation and the supernatant was removed. This washing procedure was repeated four times. The target recombinant protein

(GFP-ED protein) was eluted from the washed precipitate by means of resuspending in 0.5 ml of elution buffer containing 250 mM $K_2HPO_4$. The suspension was stored on ice for 10 minutes and then centrifuged at 12,000g for 5 minutes. The supernatant containing the eluted protein was transferred to a fresh tube (Fraction II).

The recombinant GFP protein not comprising the calcium-binding tag of the invention was used as a negative control of the purification. The bacterial cells containing the recombinant GFP were treated with the same procedures as the bacterial cells containing GFP-ED protein during purification.

FIG. 1 depicts an electrophoretic analysis of bacterial lysate fractions (fractions I, lanes 1 and 3) and purified protein fractions (fractions II, lanes 2 and 4) obtained during purification of GFP-ED (lanes 1 and 2) and GFP (lanes 3 and 4) recombinant proteins by the selective co-precipitation method of the invention. The fractions were analysed by SDS-PAGE on a 10% gel. Compared to the ordinary recombinant GFP non-comprising the tag of the invention, selective recovering and purification of target protein were achieved only for GFP-ED protein containing the calcium-binding tag of the invention (compare the yields of ordinary GFP, lane 4 and GFP-ED, lane 2). Thus, the calcium-binding tag of the invention provides selective recovering and purification of the tagged protein.

The purified GFP-ED protein of the invention (Fraction II) had the characteristic GFP fluorescence at 511 nm.

**Example 5 - Purification of recombinant Inorganic Pyrophosphatase fused to the calcium-binding tag of the invention (PPase-ED protein) by means of the selective co-precipitation method.**

Fusion PPase-ED protein of the invention comprising PPase and C-terminally calcium-binding tag (ED-peptide) of the invention was obtained by expression of the pQE- PPase-ED plasmid vector (see Example 2) in E. coli strain JM 109. The amino acid sequence of PPase-ED is given in SEQ ID No 15.

Expression of plasmid pQE-PPase-ED and purification of PPase-ED recombinant protein of the invention were performed as described in Example 4.

FIG. 2 depicts an electrophoretic analysis of fractions I (bacterial lysate, lane 1) and II (purified

target protein, lane 2) obtained during purification of the PPase-ED recombinant protein by the selective co-precipitation method of the invention. The fractions were analysed by SDS-PAGE on a 10% gel. Fig. 2 demonstrates that calcium-binding tag of the invention provides selective recovery and purification of the tagged recombinant protein.

**Example 6 - Purification of recombinant Taq DNA polymerase fused to the calcium-binding tag of the invention (Taq-ED DNA polymerase) by means of the selective co-precipitation method.**

Fusion Taq-ED protein of the invention comprising Taq DNA polymerase and C-terminally calcium-binding tag (ED-peptide) of the invention was obtained by expression of pQE-Taq-ED plasmid vector (see Example 3) in E. coli strain JM 109. The amino acid sequence of Taq-ED is given in SEQ ID No 16.

Expression of plasmid pQE-Taq-ED was performed as described in Example 4.

*Purification of Taq-ED recombinant protein*

Taq-ED protein comprising the calcium-binding tag of the invention was purified from cells by means of selective co-precipitation after denaturing E.coli host proteins at 75°C.

Bacterial cells containing the recombinant protein of the invention were collected by centrifugation from 15 ml of medium. The lysis of the cells was performed by resuspending the bacterial pellet in 0.5 ml of a lysis buffer and incubating on ice for 3-5 minutes.

The lysis-buffer contained: 150 mM $K_2HPO_4$; 300 mM KCl; 1.5% Octyl-D-glucopyranoside; 0.2% Poly(ethyleneimine).

The bacterial lysate was centrifuged at 12,000g for 5 minutes in a microfuge. The supernatant was transferred to a fresh tube (Fraction I).

The supernatant, Fraction I, was incubated at 75°C for 15 min to denature E.coli host proteins. After the heat treatment, the lysate was centrifuged at 12,000g for 5 minutes in a microfuge and the supernatant was transferred to a fresh tube (Fraction II).

100 $\mu$l of 1 M $CaCl_2$ solution were added to the lysate and mixed by vortexing. The tube was incubated at 0°C on ice for 10 minutes to mature the precipitate.

The precipitate was harvested by centrifugation for 5 minutes at 12,000g in a microfuge. The supernatant was removed by aspiration. The precipitate was washed by resuspending in 0.5 ml of washing-buffer containing: 1 M KCl; 0.18% Octyl-D-glucopyranoside; 100 mM $Na_2SO_4$; 20 mM Tris-HCl (pH=7.0), whilst maintaining the $CaCl_2$ concentration to that in the previous step, of around 180 mM.

After washing the precipitate was harvested by centrifugation and the supernatant was removed. This washing procedure was repeated four times. The target recombinant protein (Taq-ED protein) was eluted from the washed precipitate by means of resuspending in 0.5 ml of eluting-buffer containing 250 mM $K_2HPO_4$. The suspension was stored on ice for 10 minutes and then centrifuged at 12,000g for 5 minutes. The supernatant containing the eluted protein was transferred to a fresh tube (Fraction III). FIG. 3 depicts an electrophoretic analysis of fraction I (bacterial lysate, lane 1), fraction II (heated-up lysate, lane 2), and fraction III (purified target protein, lane 3) obtained during the purification of Taq-ED recombinant protein by the selective co-precipitation method of the invention. The fractions were analysed via SDS-PAGE on a 10% gel. Fig. 3 demonstrates that after the heat treatment, the calcium-binding tag of the invention provides selective recovering and purification of the tagged thermostable protein.

The purified Taq-ED protein of the invention (Fraction III) had DNA polymerase activity and provided DNA amplification by PCR (see Example 7).

**Example 7 - Performing PCR with purified Taq-ED DNA polymerase.**

The purified Taq-ED DNA polymerase from 0.5 ml Fraction III (Example 6) was precipitated by ammonium sulphate at 75% saturation. The precipitated polymerase was collected by centrifugation at 14,000 g and dissolved in 0.1 ml of storage buffer containing 100 mM NaCl; 10 mM Tris HCl, pH 7.5; 1 mM DTT; 0.2% Tween 20 and 50% (v/v) glycerol. The enzyme was stored at –20°C.

PCR amplification of a 500-bp bacteriophage lambda DNA fragment was performed with the following primers: Pr.lambda.1 (5'-GATGAGTTCGTGTCCGTACAACTGG) and Pr.lambda.2 (5'-GGTTATCGAAATCAGCCACAGCGCC). The reaction mixture (50 $\mu$l ) contained 67 mM Tris-HCl (pH 8.8 at 25°C.), 16.6 mM $(NH_4)_2SO_4$, 0.01% Tween-20, 0.2 mM each dNTP, 1.5 mM $MgCl_2$, 10 pmol each primer, 50 ng template lambda DNA and 1 $\mu$l of the above storage enzyme solution. PCR was performed for 30 cycles: 94°C for 30 seconds, 58°C for 40

seconds and 70°C for 30 seconds.

The results of the PCR are shown in FIG. 4, and indicate that recombinant Taq-ED DNA polymerase comprising the calcium-binding polypeptide tag of the invention and purified by the selective co-precipitation method of the invention provides for DNA amplification of the target DNA fragment by PCR.

**Example 8 - Purification of recombinant PPase-ED and GFP-ED proteins comprising the calcium-binding tag of the invention performed under denaturing conditions by the selective co-precipitation method.**

The fusion proteins of the invention comprising the C-terminal calcium-binding tag (ED-peptide) of the invention were obtained by expression of plasmids pQE-GFP-ED and pQE-PPase-ED in E. coli strain JM 109 as described in Example 4.

PPase-ED and GFP-ED proteins comprising the calcium-binding tag of the invention were purified from E.coli cells under denaturing conditions by means of the selective co-precipitation method.

Bacterial cells containing the recombinant proteins of the invention were collected by centrifugation from 15 ml of medium. The lysis of the cells was performed by resuspending the bacterial pellets in 0.5 ml of the lysis buffer and incubating at room temperature for 60 minutes. The lysis buffer contained: 6 M guanidine hydrochloride; 150 mM $K_2HPO_4$; 300 mM KCl; 1.5% Octyl-D-glucopyranoside; 0.2% Poly(ethyleneimine). The bacterial lysates were centrifuged at 14,000g for 25 minutes in a microfuge. The supernatants were transferred to fresh tubes (Fraction I). 100 $\mu$l of 1 M $CaCl_2$ solution were added to each lysate and mixed by vortexing. The tubes were incubated at room temperature for 15 minutes to mature the precipitates. The precipitates were harvested by centrifugation for 10 minutes at 12,000g in a microfuge. The supernatants were removed by aspiration. The precipitates were washed by resuspending in 0.5 ml of washing-buffer containing: 8 M urea; 1 M KCl; 0.18% Octyl-D-glucopyranoside; 100 mM $Na_2SO_4$; 20 mM Tris-HCl (pH=7.0) whilst maintaining the $CaCl_2$ concentration to that in the previous step, of around 180 mM. After washing, the precipitates were harvested by centrifugation and supernatants were removed. This washing procedure was repeated four times.

The target recombinant proteins were eluted from the washed precipitates by means of resuspending in 0.5 ml of eluting-buffer containing: 8 M urea; 0.5 M KCl; 100 mM EDTA. The

suspensions were incubated at room temperature for 15 minutes and then centrifuged at 12,000g for 5 minutes. Supernatants containing the eluted proteins were transferred to fresh tubes (Fraction II).

FIG. 5 depicts an electrophoretic analysis of fractions I (bacterial lysate) and II (purified target protein) obtained during the purification of GFP-ED and PPase-ED proteins in denaturing conditions by the selective co-precipitation method of the invention. The fractions were analysed via SDS-PAGE on a 10% gel. The results shown in FIG. 5 indicate that the calcium-binding tag of the invention provides selective recovering and purification of the tagged protein under denaturing conditions.

**Example 9 - Two-step purification of GFP-ED protein by IMAC and the selective co-precipitation method of the invention**

The fusion protein of the invention comprising the C-terminal calcium-binding tag of the invention and N-terminal 6xHis-tag was obtained by expression of plasmids pQE-GFP-ED in E. coli strain JM 109 as described in Example 4.

GFP-ED protein was purified from E.coli cells by means of IMAC chromatography on $Ni^{2+}$ chelating resin and the selective co-precipitation method of the invention.
Bacterial cells containing the recombinant proteins of the invention were collected by centrifugation from 40 ml of medium.

FIG. 6 depicts an electrophoretic analysis of fractions obtained during two-step purification of GFP-ED protein by means of IMAC and the selective co-precipitation methods. The obtained results demonstrate that the sequential purification of a recombinant protein linked to an N-terminal poly-Histidine-tag and a C-terminal calcium-binding tag can be purified to homogeneity in only two simple steps

**Example 10 – discussion of the ED peptide tag of the Examples**

Dahlback *et al.* (J Biol Che, 1990: 265,30: 1841-1849) showed that Protein S contains 4 distinct $Ca^{2+}$ binding sites. A modified $Ca^{2+}$ binding EGF-like domain of the invention (termed herein "Ed-peptide" is aligned below with the 4 binding sites from Protein S.

**EDVDECSENMSAQL----CKD ED-peptide (Ed-peptide)**

**AIPDQCNPIPCNEDGFMTCKD Domain1-ProtS**

**SDINECKDPVNINGGSQICEN Domain2-ProtS**

**KDVDECVLKPSICGTAV-CKN Domain3-ProtS**

**DDVDECAENLCAQL----CVN Domain4-ProtS**

Domain 4 is the most similar:

**EDVDECSENMSAQLCKD ED-peptide (Ed-peptide)**

**DDVDECAENLCAQLCVN Domain4-ProtS**

Dahlback *et al.* measured a Kd for Calcium of 1 $\mu$M for Domain4, which is described by the authors as "very high affinity" compared with all other Calcium binding domains.

The authors also noted The Protein S-C4BP interaction and has a 100-fold higher affinity in the presence of $Ca^{2+}$ showing that the short domain appears to facilitate polypeptide aggregation.

Without wishing to be bound by theory, it is envisaged that the Ed-peptide may form an aggregation-competent structure in the presence of Calcium. Upon removal of the initial, high concentration of Calcium (in the form of $CaCl_2$), the Calcium can be displaced by the addition of a soluble salt of e.g. phosphate that eventually forms an insoluble Calcium salt and draws the peptide bound calcium out of the aqueous equilibrium. This maturation process removes active calcium species from solution and the consequent switch in conformation of the Ed-peptide leads to disaggregation of the recombinant protein and separation can be achieved by removal of the calcium phosphate precipitate.

## SEQUENCE LISTING

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 1**
<211> LENGTH: can possess the value from 12 to 234
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<223> OTHER INFORMATION: Description of Artificial Sequence – Calcium-binding Artificial Polypeptide, wherein $X_m$, $X_z$ and $X_n$ can be any amino acid sequences comprising ***m, z*** and ***n*** amino acid residuals correspondingly; and wherein ***m*** and ***n*** can be any whole numbers from 0 to 100, and ***z*** can be any whole number from 3 to 25.
<400> SEQUENCE 1:

$\mathbf{X_m}$-(E or D)-(D or E)-(V or I or L)-(D or E)-(E or D)-C-$\mathbf{X_z}$-C-(K or R or H)-(D or E)-$\mathbf{X_n}$

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 2**
<211> LENGTH: can possess the value from 17 to 217
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<223> OTHER INFORMATION: Description of Artificial Sequence – Calcium-binding Artificial Polypeptide, wherein $X_m$, $X_z$ and $X_n$ can be any amino acid sequences comprising ***m, z*** and ***n*** amino acid residuals correspondingly; and wherein ***m*** and ***n*** can be any whole numbers from 0 to 100.
<400> SEQUENCE 2:

$\mathbf{X_m}$-E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D-$\mathbf{X_n}$

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 3**
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: ED-peptide
<223> OTHER INFORMATION: Description of Artificial Sequence – Calcium-binding Artificial Polypeptide
<400> SEQUENCE 3:

E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 4**
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<223> OTHER INFORMATION: Description of Artificial Sequence – Amino Acid Sequence, which can be recognized and cleaved by blood coagulation factor Xa
<400> SEQUENCE :

Ile-Glu-Gly-Arg

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 5**
<211> LENGTH: 740
<212> TYPE: DNA
<213> ORGANISM: Aequorea victoria / Artificial Sequence
<223> OTHER INFORMATION: Description of Sequence – DNA sequence encoding Green Fluorescing Protein (GFP). Encoding part of the sequence is marked grey.
<400> SEQUENCE 5:

ACATGGATCCGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTC
GAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC
GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCG
TGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTA
CCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCC
AGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAA
GTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAG
GACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATAT
CATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCG
AGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACG
GCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGA
CCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATC
ACTCTCGGCATGGACGAGCTGTACAAGGCATGCGAGCTCTAGA

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 6**
<211> LENGTH: 4176
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: pQE-GFP
<222> LOCATION: (1) – (4176)
<223> OTHER INFORMATION: Description of Artificial Sequence – DNA sequence of pQE-GFP plasmid. DNA sequence encoding recombinant GFP is marked grey. Sequence encoding 6xHis-tag of the recombinant protein is marked by slanted characters
<400> SEQUENCE 6:

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCAA
TTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGAG
GATCG*CATCACCATCACCATCAC*GGATCCGTGAGCAAGGGCGAGGAGCTGTTCACCGG
GGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTG
TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCA
CCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGT
GCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCA
TGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAA
GACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAA
GGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTAC
AACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTT
CAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAG
AACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCC
AGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGCATGCGAGCTC

GGTACCCCGGGTCGACCTGCAGCCAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATA
GATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCG
GGCGTTTTTTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAA
GCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGT
AAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAG
CTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCC
TTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGGCAATGAAA
GACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCA
AACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACA
CATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTT
TATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTA
AACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACG
CAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTCTGTGATG
GCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGC
GGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCCTGGGGTAATGACTCT
CTAGCTTGAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTA
TCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGCTCTAGAGCTGC
CTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGG
TCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGC
GGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGT
GTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCG
GTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTT
CCTCGCTCACTGACTCGCTGCGCTCGGTCTGTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT
GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTT
CCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGG
CGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG
CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA
GCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGC
TCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCG
GTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC
ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG
GTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC
CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGT
AGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA
AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAG
GGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATG
AAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTA
ATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGCTGCCTGACTC
CCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAAT
GATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCG
GAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAAT
TGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGC
CATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCG
GTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGC
TCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGT
TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC
TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT
GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATC
ATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAG
TTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGT
TTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA
CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTT
ATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCC

GCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATT
AACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAC

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 7**
<211> LENGTH: 63
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<223> OTHER INFORMATION: Description of Artificial Sequence – DNA fragment encoding calcium-binding ED-peptide. DNA sequence encoding ED-peptide is marked grey.
<400> SEQUENCE 7:

CTGCAGAAGATGTGGATGAATGCAGCGAAAACATGAGCGCGCAGCTGTGCAAAGATTAA
GCTT

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 8**
<211> LENGTH: 4225
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: pQE-GFP-ED
<222> LOCATION: (1) – (4225)
<223> OTHER INFORMATION: Description of Artificial Sequence – DNA sequence of pQE-GFP-ED plasmid. DNA sequence encoding recombinant GFP-ED protein is marked grey. Sequences encoding N-terminally 6xHis-tag and C-terminally calcium-binding tag (ED-peptide) of the recombinant protein are marked by slanted characters
<400> SEQUENCE 8:

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCAA
TTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGAG
GATCG***CATCACCATCACCATCAC***GGATCCGTGAGCAAGGGCGAGGAGCTGTTCACCGG
GGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTG
TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCA
CCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGT
GCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCA
TGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAA
GACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAA
GGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTAC
AACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTT
CAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAG
AACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCC
AGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGCATGCGAGCTC
GGTACCCCGGGTCGACCTGCA***GAAGATGTGGATGAATGCAGCGAAAACATGAGCGCG***
***CAGCTGTGCAAAGAT***TAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGT
AATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTT
TTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAAT
GGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACA
TTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATAT
TACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCA
CATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGGCAATGAAAGACGGTG

AGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAA
CGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATT
CGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAG
AATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTG
GCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGC
GACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTCTGTGATGGCTTCCA
TGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCG
TAATTTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCCTGGGGTAATGACTCTCTAGCTT
GAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTG
TTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGCTCTAGAGCTGCCTCGCGC
GTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCT
TGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTT
GGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTG
GCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAA
TACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTC
ACTGACTCGCTGCGCTCGGTCTGTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA
GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGC
TCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCC
GACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG
TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC
GCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC
TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA
TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA
ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCC
TAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA
CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGT
GGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCT
TTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTG
GTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTA
AATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA
GGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGCTGCCTGACTCCCCGTCGT
GTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGC
GAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC
CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCC
GGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCA
ACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCG
GTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCA
GCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAG
TACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGC
GTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAA
ACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGT
AACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGT
GAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATG
TTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCA
TGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACAT
TTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATA
AAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAC

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 9**
<211> LENGTH: 546

<212> TYPE: DNA
<213> ORGANISM: Escherichia coli / Artificial Sequence
<223> OTHER INFORMATION: Description of Sequence – DNA fragment encoding Inorganic Pyrophosphatase (PPase). Encoding part of the sequence is marked grey.
<400> SEQUENCE 9:

ACATGGATCCAGCTTACTCAACGTCCCTGCGGGTAAAGATCTGCCGGAAGACATCTACG
TTGTTATTGAGATCCCGGCTAACGCAGATCCGATCAAATACGAAATCGACAAAGAGAGC
GGCGCACTGTTCGTTGACCGCTTCATGTCCACCGCGATGTTCTATCCATGCAACTACGG
TTACATCAACCACACCCTGTCTCTGGACGGTGACCCGGTTGACGTACTGGTCCCAACTC
CGTACCCGCTGCAGCCGGGTTCTGTGATCCGTTGCCGTCCGGTTGGCGTTCTGAAAAT
GACCGACGAAGCCGGTGAAGATGCGAAACTGGTTGCGGTTCCGCACAGCAAGCTGAGC
AAAGAATACGATCACATTAAAGACGTTAACGATCTGCCTGAACTGCTGAAAGCGCAAATC
GCTCACTTCTTCGAGCACTACAAAGACCTCGAAAAAGGCAAGTGGGTGAAAGTTGAAGC
TTGGGAAAACGCAGAAGCCGCTAAAGCTGAAATCGTTGCCTCCTTCGAGCGCGCAAAGA
ATAAAGGTCGACGGTC

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 10**
<211> LENGTH: 4015
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: pQE-PPase-ED
<222> LOCATION: (1) – (4015)
<223> OTHER INFORMATION: Description of Artificial Sequence – DNA sequence of pQE-PPase-ED plasmid. DNA sequence encoding recombinant PPase-ED protein is marked grey. Sequences encoding N-terminally 6xHis-tag and to C-terminally calcium-binding tag (ED-peptide) of the recombinant protein are marked by slanted characters
<400> SEQUENCE 10:

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCAA
TTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGAG
GATCG***CATCACCATCACCATCAC***GGATCCAGCTTACTCAACGTCCCTGCGGGTAAAGAT
CTGCCGGAAGACATCTACGTTGTTATTGAGATCCCGGCTAACGCAGATCCGATCAAATA
CGAAATCGACAAAGAGAGCGGCGCACTGTTCGTTGACCGCTTCATGTCCACCGCGATGT
TCTATCCATGCAACTACGGTTACATCAACCACACCCTGTCTCTGGACGGTGACCCGGTT
GACGTACTGGTCCCAACTCCGTACCCGCTGCAGCCGGGTTCTGTGATCCGTTGCCGTC
CGGTTGGCGTTCTGAAAATGACCGACGAAGCCGGTGAAGATGCGAAACTGGTTGCGGT
TCCGCACAGCAAGCTGAGCAAAGAATACGATCACATTAAAGACGTTAACGATCTGCCTG
AACTGCTGAAAGCGCAAATCGCTCACTTCTTCGAGCACTACAAAGACCTCGAAAAAGGC
AAGTGGGTGAAAGTTGAAGGTTGGGAAAACGCAGAAGCCGCTAAAGCTGAAATCGTTGC
CTCCTTCGAGCGCGCAAAGAATAAAGGTCGACCTGCA***GAAGATGTGGATGAATGCAGC***
***GAAAACATGAGCGCGCAGCTGTGCAAAGAT***TAAGCTTAATTAGCTGAGCTTGGACTCC
TGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTT
GCCGCCGGGCGTTTTTTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAGC
TAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATG
GCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGAC
CGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTA
TCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGGC
AATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCA
TGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGT
TTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTA

AAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTT
TTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAAT
ATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTC
TGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTG
GCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCCTGGGGTAA
TGACTCTCTAGCTTGAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTT
TCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGCTCTA
GAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCG
GAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCG
CGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATA
GCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACC
ATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCT
TCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTGTCGGCTGCGGCGAGCGGTAT
CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAG
AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGG
CGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG
AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT
CGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTC
GTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCT
TATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA
GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGC
TGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC
TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC
GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA
AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAA
TGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGCTGCC
TGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC
TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC
CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTC
TATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGT
TGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCA
GCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCG
GTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACT
CATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTC
TGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTT
GCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTG
CTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAG
ATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCAC
CAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGG
GCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATC
AGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGG
GGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCAT
GACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAC

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 11**
<211> LENGTH:
<212> TYPE: DNA
<213> ORGANISM: Thermus aquaticus / Artificial Sequence
<223> OTHER INFORMATION: Description of Sequence – DNA fragment encoding Taq DNA

polymerase. Encoding part of the sequence is marked grey.

<400> SEQUENCE 11:

ACATAGATCTAGGGGGATGCTGCCCCTCTTTGAGCCCAAGGGCCGGGTCCTCCTGGTG
GACGGCCACCACCTGGCCTACCGCACCTTCCACGCCCTGAAGGGCCTCACCACCAGCC
GGGGGGAGCCGGTGCAGGCGGTCTACGGCTTCGCCAAGAGCCTCCTCAAGGCCCTCA
AGGAGGACGGGGACGCGGTGATCGTGGTCTTTGACGCCAAGGCCCCCTCCTTCCGCCA
CGAGGCCTACGGGGGGTACAAGGCGGGCCGGGCCCCCACGCCGGAGGACTTTCCCCG
GCAACTCGCCCTCATCAAGGAGCTGGTGGACCTCCTGGGGCTGGCGCGCCTCGAGGTC
CCGGGCTACGAGGCGGACGACGTCCTGGCCAGCCTGGCCAAGAAGGCGGAAAAGGAG
GGCTACGAGGTCCGCATCCTCACCGCCGACAAAGACCTTTACCAGCTCCTTTCCGACCG
CATCCACGCCCTCCACCCCGAGGGGTACCTCATCACCCCGGCCTGGCTTTGGGAAAAG
TACGGCCTGAGGCCCGACCAGTGGCCCGACTACCGGGCCCTGACCGGGGACGAGTCC
GACAACCTTCCCGGGGTCAAGGGCATCGGGGAGAAGACGGCGAGGAAGCTTCTGGAG
GAGTGGGGGAGCCTGGAAGCCCTCCTCAAGAACCTGGACCGGCTGAAGCCCGCCATCG
CGGAGAAGATCCTGGCCCACATGGACGATCTGAAGCTCTCCTGGGACCTGGCCAAGGT
GCGCACCGACCTGCCCCTGGAGGTGGACTTCGCCAAAAGGCGGGAGCCCGACCGGGA
GAGGCTTAGGGCCTTTCTGGAGAGGCTTGAGTTTGGCAGCCTCCTCCACGAGTTCGGC
CTTCTGGAAAGCCCCAAGGCCCTGGAGGAGGCCCCCTGGCCCCCGCCGGAAGGGGCC
TTCGTGGGCTTTGTGCTTTCCCGCAAGGAGCCCATGTGGGCCGATCTTCTGGCCCTGG
CCGCCGCCAGGGGGGGCCGGGTCCACCGGGCCCCCGAGCCTTATAAAGCCCTCAGGG
ACCTGAAGGAGGCGCGGGGGCTTCTCGCCAAAGACCTGAGCGTTCTGGCCCTGAGGGA
AGGCCTTGGCCTCCCGCCCGGCGACGACCCCATGCTCCTCGCCTACCTCCTGGACCCT
TCCAACACCACCCCCGAGGGGGTGGCCCGGCGCTACGGCGGGGAGTGGACGGAGGAG
GCGGGGGAGCGGGCCGCCCTTTCCGAGAGGCTCTTCGCCAACCTGTGGGGGAGGCTT
GAGGGGGAGGAGAGGCTCCTTTGGCTTTACCGGGAGGTGGAGAGGCCCCTTTCCGCT
GTCCTGGCCCACATGGAGGCCACGGGGGGTGCGCCTGGACGTGGCCTATCTCAGGGCC
TTGTCCCTGGAGGTGGCCGAGGAGATCGCCCGCCTCGAGGCCGAGGTCTTCCGCCTG
GCCGGCCACCCCTTCAACCTCAACTCCCGGGACCAGCTGGAAAGGGTCCTCTTTGACG
AGCTAGGGCTTCCCGCCATCGGCAAGACGGAGAAGACCGGCAAGCGCTCCACCAGCGC
CGCCGTCCTGGAGGCCCTCCGCGAGGCCCACCCCATCGTGGAGAAGATCCTGCAGTAC
CGGGAGCTCACCAAGCTGAAGAGCACCTACATTGACCCCTTGCCGGACCTCATCCACCC
CAGGACGGGCCGCCTCCACACCCGCTTCAACCAGACGGCCACGGCCACGGGCAGGCT
AAGTAGCTCCGATCCCAACCTCCAGAACATCCCCGTCCGCACCCCGCTTGGGCAGAGG
ATCCGCCGGGCCTTCATCGCCGAGGAGGGGTGGCTATTGGTGGCCCTGGACTATAGCC
AGATAGAGCTCAGGGTGCTGGCCCACCTCTCCGGCGACGAGAACCTGATCCGGGTCTT
CCAGGAGGGGCGGGACATCCACACGGAGACCGCCAGCTGGATGTTCGGCGTCCCCCG
GGAGGCCGTGGACCCCCTGATGCGCCGGGCGGCCAAGACCATCAACTTCGGGGTCCT
CTACGGCATGTCGGCCCACCGCCTCTCCCAGGAGCTAGCCATCCCTTACGAGGAGGCC
CAGGCCTTCATTGAGCGCTACTTTCAGAGCTTCCCCAAGGTGCGGGCCTGGATTGAGAA
GACCCTGGAGGAGGGCAGGAGGCGGGGGTACGTGGAGACCCTCTTCGGCCGCCGCCG
CTACGTGCCAGACCTAGAGGCCCGGGTGAAGAGCGTGCGGGAGGCGGCCGAGCGCAT
GGCCTTCAACATGCCCGTCCAGGGCACCGCCGCCGACCTCATGAAGCTGGCTATGGTG
AAGCTCTTCCCCAGGCTGGAGGAAATGGGGGCCAGGATGCTCCTTCAGGTCCACGACG
AGCTGGTCCTCGAGGCCCCAAAAGAGAGGGCGGAGGCCGTGGCCCGGCTGGCCAAGG
AGGTCATGGAGGGGGTGTATCCCCTGGCCGTGCCCCTGGAGGTGGAGGTGGGGATAG
GGGAGGACTGGCTCTCCGCCAAGGAGGGTCGACGGTC

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 12**
<211> LENGTH: 5983
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

<220> FEATURE:
<221> NAME/KEY: pQE-Taq-ED
<222> LOCATION: (1) – (5983)
<223> OTHER INFORMATION: Description of Artificial Sequence – DNA sequence of pQE-Taq-ED plasmid. DNA sequence encoding recombinant Taq-ED protein is marked grey. Sequences encoding N-terminally 6xHis-tag and to C-terminally calcium-binding tag (ED-peptide) of the recombinant protein are marked by slanted characters
<400> SEQUENCE 12:

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCAA
TTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGAG
GATCG*CATCACCATCACCATCAC*GGATCTAGGGGGATGCTGCCCCTCTTTGAGCCCAA
GGGCCGGGTCCTCCTGGTGGACGGCCACCACCTGGCCTACCGCACCTTCCACGCCCTG
AAGGGCCTCACCACCAGCCGGGGGGAGCCGGTGCAGGCCGTCTACGGCTTCGCCAAG
AGCCTCCTCAAGGCCCTCAAGGAGGACGGGGACGCCGTGATCGTGGTCTTTGACGCCA
AGGCCCCCTCCTTCCGCCACGAGGCCTACGGGGGGTACAAGGCGGGCCGGGCCCCCA
CGCCGGAGGACTTTCCCCGGCAACTCGCCCTCATCAAGGAGCTGGTGGACCTCCTGGG
GCTGGCGCGCCTCGAGGTCCCGGGCTACGAGGCGGACGACGTCCTGGCCAGCCTGGC
CAAGAAGGCGGAAAAGGAGGGCTACGAGGTCCGCATCCTCACCGCCGACAAAGACCTT
TACCAGCTCCTTTCCGACCGCATCCACGCCCTCCACCCCGAGGGGTACCTCATCACCCC
GGCCTGGCTTTGGGAAAAGTACGGCCTGAGGCCCGACCAGTGGGCCGACTACCGGGC
CCTGACCGGGGACGAGTCCGACAACCTTCCCGGGGTCAAGGGCATCGGGGAGAAGAC
GGCGAGGAAGCTTCTGGAGGAGTGGGGGAGCCTGGAAGCCCTCCTCAAGAACCTGGA
CCGGCTGAAGCCGGCCATCCGGGAGAAGATCCTGGCCCACATGGACGATCTGAAGCTC
TCCTGGGACCTGGCCAAGGTGCGCACCGACCTGCCCCTGGAGGTGGACTTCGCCAAAA
GGCGGGAGCCCGACCGGGAGAGGCTTAGGGCCTTTCTGGAGAGGCTTGAGTTTGGCA
GCCTCCTCCACGAGTTCGGCCTTCTGGAAAGCCCCAAGGCCCTGGAGGAGGCCCCCTG
GCCCCCGCCGGAAGGGGCCTTCGTGGGCTTTGTGCTTTCCCGCAAGGAGCCCATGTGG
GCCGATCTTCTGGCCCTGGCCGCCGCCAGGGGGGGCCGCGGTCCACCGGGCCCCCGAG
CCTTATAAAGCCCTCAGGGACCTGAAGGAGGCGCGGGGGCTTCTCGCCAAAGACCTGA
GCGTTCTGGCCCTGAGGGAAGGCCTTGGCCTCCCGCCCGGCGACGACCCCATGCTCCT
CGCCTACCTCCTGGACCCTTCCAACACCACCCCCGAGGGGGTGGCCCGGCGCTACGG
CGGGGAGTGGACGGAGGAGGCGGGGGAGCGGGCCGCCCTTTCCGAGAGGCTCTTCG
CCAACCTGTGGGGGAGGCTTGAGGGGGAGGAGAGGCTCCTTTGGCTTTACCGGGAGG
TGGAGAGGCCCCTTTCCGCTGTCCTGGCCCACATGGAGGCCACGGCGGGTGCGCCTGG
ACGTGGCCTATCTCAGGGCCTTGTCCCTGGAGGTGGCCGAGGAGATCGCCCGCCTCGA
GGCCGAGGTCTTCCGCCTGGCCGGCCACCCCTTCAACCTCAACTCCCGGGACCAGCTG
GAAAGGGTCCTCTTTGACGAGCTAGGGCTTCCCGCCATCGGCAAGACGGAGAAGACCG
GCAAGCGCTCCACCAGCGCCGCCGTCCTGGAGGCCCTCCGCGAGGCCCACCCCATCG
TGGAGAAGATCCTGCAGTACCGGGAGCTCACCAAGCTGAAGAGCACCTACATTGACCCC
TTGCCGGACCTCATCCACCCCAGGACGGGCCGCCTCCACACCCGCTTCAACCAGACGG
CCACGGCCACGGGCAGGCTAAGTAGCTCCGATCCCAACCTCCAGAACATCCCCGTCCG
CACCCCGCTTGGGCAGAGGATCCGCCGGGCCTTCATCGCCGAGGAGGGGTGGCTATT
GGTGGCCCTGGACTATAGCCAGATAGAGCTCAGGGTGCTGGCCCACCTCTCCGGCGAC
GAGAACCTGATCCGGGTCTTCCAGGAGGGGCGGGACATCCACACGGAGACCGCCAGCT
GGATGTTCGGCGTCCCCCGGGAGGCCGTGGACCCCCTGATGCGCGCGGGCGGCCAAGA
CCATCAACTTCGGGGTCCTCTACGGCATGTCGGCCCACCGCCTCTCCCAGGAGCTAGC
CATCCCTTACGAGGAGGCCCAGGCCTTCATTGAGCGCTACTTTCAGAGCTTCCCCAAGG
TGCGGGCCTGGATTGAGAAGACCCTGGAGGAGGGCAGGAGGCGGGGGTACGTGGAGA
CCCTCTTCGGCCGCCGCCGCTACGTGCCAGACCTAGAGGCCCGGGTGAAGAGCGTGC
GGGAGGCGGCCGAGCGCATGGCCTTCAACATGCCCGTCCAGGGCACCGCCGCCGACC
TCATGAAGCTGGCTATGGTGAAGCTCTTCCCCAGGCTGGAGGAAATGGGGGCCAGGAT
GCTCCTTCAGGTCCACGACGAGCTGGTCCTCGAGGCCCCAAAAGAGAGGCGGAGGC
CGTGGCCCGGCTGGCCAAGGAGGTCATGGAGGGGGTGTATCCCCTGGCCGTGCCCCT

GGAGGTGGAGGTGGGGATAGGGGAGGACTGGCTCTCCGCCAAGGAGGGTCGACCTGC
AGAAGATGTGGATGAATGCAGCGAAAACATGAGCGCGCAGCTGTGCAAAGATTAAGC
TTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCT
GGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTATTGGTGAGAATCCAAGCT
AGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTGGATAT
ACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTT
GCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTA
AAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAAT
GCTCATCCGGAATTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGT
TCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGA
ATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGG
TGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAAT
CCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCC
CCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGC
GATTCAGGTTCATCATGCCGTCTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATT
ACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGT
GCCCTTAAACGCCTGGGGTAATGACTCTCTAGCTTGAGGCATCAAATAAAACGAAAGGC
TCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAG
TAGGACAAATCCGCCGCTCTAGAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAAC
CTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGA
GCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCA
TGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGC
AGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGA
AAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTG
TCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT
CAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG
TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACA
AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGT
ATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTA
GGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGT
ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTT
GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATT
ACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGC
TCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTT
CACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA
CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTAT
TTCGTTCATCCATAGCTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGC
TTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG
ATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACT
TTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA
GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTC
GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCC
CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAG
TTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATG
CCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAG
TGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC
ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAA
GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTT
CAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCC
GCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAA

TATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTT
AGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC
TAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTC
GTCTTCAC

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 13**
<211> LENGTH:
<212> TYPE: PRT
<213> ORGANISM: Aequorea victoria / Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: GFP-ED protein
<223> OTHER INFORMATION: Description of Sequence – Amino acid sequence of recombinant GFP-ED protein. Sequences of N-terminally 6xHis-tag and C-terminally calcium-binding tag (ED-peptide) of the recombinant protein are marked by slanted characters
<400> SEQUENCE 13:

MRGS***HHHHHH***GSVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAE
VKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGS
VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELY
KACELGTPGRPA***EDVDECSENMSAQLCKD***

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 14**
<211> LENGTH:
<212> TYPE: PRT
<213> ORGANISM: Aequorea victoria / Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: GFP
<223> OTHER INFORMATION: Description of Sequence – Amino acid sequence of recombinant Green Fluorescing Protein (GFP). Sequences of N-terminally 6xHis-tag of the recombinant protein is marked by slanted characters
<400> SEQUENCE 14:

MRGS***HHHHHH***GSVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAE
VKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGS
VQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELY
KACELGTPGRPAAK

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 15**
<211> LENGTH:
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli / Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: PPase-ED protein
<223> OTHER INFORMATION: Description of Sequence – Amino acid sequence of recombinant PPase-ED protein. Sequences of N-terminally 6xHis-tag and C-terminally calcium-binding tag (ED-peptide) of the recombinant protein are marked by slanted characters

<400> SEQUENCE 15:

MRGS***HHHHHH***GSSLLNVPAGKDLPEDIYVVIEIPANADPIKYEIDKESGALFVDRFMSTAMFY
PCNYGYINHTLSLDGDPVDVLVPTPYPLQPGSVIRCRPVGVLKMTDEAGEDAKLVAVPHSKL
SKEYDHIKDVNDLPELLKAQIAHFFEHYKDLEKGKWVKVEGWENAEAAKAEIVASFERAKNK
GRPA***EDVDECSENMSAQLCKD***

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 16**
<211> LENGTH:
<212> TYPE: PRT
<213> ORGANISM: Thermus aquaticus / Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: Taq-ED DNA polymerase
<223> OTHER INFORMATION:Description of Sequence – Amino acid sequence of recombinant Taq-ED DNA polymerase. Sequences of N-terminally 6xHis-tag and C-terminally calcium-binding tag (ED-peptide) of the recombinant protein are marked by slanted characters
<400> SEQUENCE 16:

MRGS***HHHHHH***GSRGMLPLFEPKGRVLLVDGHHLAYRTFHALKGLTTSRGEPVQAVYGFAK
SLLKALKEDGDAVIVVFDAKAPSFRHEAYGGYKAGRAPTPEDFPRQLALIKELVDLLGLARLE
VPGYEADDVLASLAKKAEKEGYEVRILTADKDLYQLLSDRIHALHPEGYLITPAWLWEKYGLR
PDQWADYRALTGDESDNLPGVKGIGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMD
DLKLSWDLAKVRTDLPLEVDFAKRREPDRERLRAFLERLEFGSLLHEFGLLESPKALEEAPW
PPPEGAFVGFVLSRKEPMWADLLALAAARGGRVHRAPEPYKALRDLKEARGLLAKDLSVLA
LREGLGLPPGDDPMLLAYLLDPSNTTPEGVARRYGGEWTEEAGERAALSERLFANLWGRLE
GEERLLWLYREVERPLSAVLAHMEATGVRLDVAYLRALSLEVAEEIARLEAEVFRLAGHPFNL
NSRDQLERVLFDELGLPAIGKTEKTGKRSTSAAVLEALREAHPIVEKILQYRELTKLKSTYIDPL
PDLIHPRTGRLHTRFNQTATATGRLSSSDPNLQNIPVRTPLGQRIRRAFIAEEGWLLVALDYS
QIELRVLAHLSGDENLIRVFQEGRDIHTETASWMFGVPREAVDPLMRRAAKTINFGVLYGMS
AHRLSQELAIPYEEAQAFIERYFQSFPKVRAWIEKTLEEGRRRGYVETLFGRRRYVPDLEARV
KSVREAAERMAFNMPVQGTAADLMKLAMVKLFPRLEEMGARMLLQVHDELVLEAPKERAEA
VARLAKEVMEGVYPLAVPLEVEVGIGEDWLSAKEGRPA***EDVDECSENMSAQLCKD***

<200> SEQUENCE CHARACTERISTICS:
<210> **SEQ ID NO 17**
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<223> OTHER INFORMATION: Description of Artificial Sequence – Amino Acid Sequence, which can be recognized and cleaved by the tomato etch virus (TEV) protease
<400> SEQUENCE :17

Glu-X-X-Tyr-X-Gln/Ser

Claims

1. A method of reducing the solubility of a soluble target-entity from a sample, such as to facilitate recovery thereof, the method comprising the steps of:

(i) appending a heterologous tag-entity to the target entity to form a tagged target,

(ii) contacting the heterologous tag with a ligand which is a metal ion or salt which binds to the tag-entity in a 1:1 ratio causing reduction in solubility of the tagged target.

2. A method of reducing the solubility of a tagged target from a sample, such as to facilitate recovery thereof,

wherein the tagged target comprises a soluble target-entity and a heterologous tag-entity, the method comprising contacting the heterologous tag-entity with a ligand which is a metal ion or salt which binds to the tag-entity in a 1:1 ratio causing reduction in solubility of the tagged target.

3. A method as claimed in claim 1 or claim 2 wherein the ligand is partially soluble and the method is performed in the absence of chromatographic resin.

4. A method as claimed in any one of claims 1 to 3 wherein the target entity is a protein which is appended to the heterologous tag in a 1:1 ratio. such that addition of the ligand does not cause cross linking of the tagged target.

5. A method as claimed in any one of claims 1 to 4 wherein the reduction in solubility causes precipitation of the tagged target.

6. A method as claimed in any one of claims 1 to 6 wherein the tag-entity is a Ca-binding EGF-like peptide or derivative thereof.

7. A method of precipitating a soluble target-entity from a sample, such as to facilitate recovery thereof, the method comprising the steps of:

(i) appending a heterologous tag-entity to the target entity to form a tagged target,

(ii) contacting the heterologous tag with a ligand which is a metal ion or salt which binds to the tag-entity reduction in solubility of the tagged target,

wherein the tag-entity comprises the following sequence of amino acids residues:

Xm-X1-X2-X3-X4-X5-cys-Xz-cys-X8-X9-Xn

wherein:

Xm = a terminal sequence of m amino acids which can be any whole number from 0 to 100
X1 = an acidic amino acid
X2 = an acidic amino acid
X3 = a hydrophobic amino acid
X4 = an acidic amino acid
X5 = an acidic amino acid
Xz = a spacer sequence of z amino acids which can be any whole number from 1 to 25
X8 = a basic amino acid
X9 = an acidic amino acid
Xn = a terminal sequence of n amino acids which can be any whole number from 0 to 100

8. A method of precipitating a soluble target-entity from a sample, such as to facilitate recovery thereof, the method comprising the steps of:

wherein the tagged target comprises a soluble target-entity and a heterologous tag-entity, the method comprising contacting the heterologous tag-entity with a ligand which is a metal ion or salt which binds to the tag-entity causing reduction in solubility of the tagged target,

wherein the tag-entity comprises the following sequence of amino acids residues:

Xm-X1-X2-X3-X4-X5-cys-Xz-cys-X8-X9-Xn

wherein:
Xm = a terminal sequence of m amino acids which can be any whole number from 0 to 100
X1 = an acidic amino acid
X2 = an acidic amino acid
X3 = a hydrophobic amino acid
X4 = an acidic amino acid
X5 = an acidic amino acid
Xz = a spacer sequence of z amino acids which can be any whole number from 1 to 25
X8 = a basic amino acid
X9 = an acidic amino acid
Xn = a terminal sequence of n amino acids which can be any whole number from 0 to 100

9. A method as claimed in claim 7 or claim 8 wherein the tag-entity comprises an artificial polypeptide which is be defined by the following general amino acid sequence (SEQ ID No 1):

$\mathbf{X_m}$-(E or D)-(D or E)-(V or I or L)-(D or E)-(E or D)-C-$\mathbf{X_z}$-C-(K or R or H)-(D or E)-$\mathbf{X_n}$;

wherein $X_m$, $X_z$ and $X_n$ can be any amino acid sequences comprising ***m, z*** and ***n*** amino

acid residues, respectively; and wherein ***m*** and ***n*** can be any whole numbers from 0 to 100, and ***z*** can be any whole number from 1 to 25.

10. A method as claimed in claim 9, wherein said artificial polypeptide has the following amino acid sequence (SEQ ID No 2):

$\mathbf{X_m}$-E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D-$\mathbf{X_n}$;

wherein $X_m$ and $X_n$ can be any amino acid sequences comprising ***m*** and ***n*** amino acid residues, respectively; and wherein ***m*** and ***n*** can be any whole numbers from 0 to 100.

11. A method as claimed in 10, wherein said artificial polypeptide has the following amino acid sequence (SEQ ID No 3):

E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D.

12. A method as claimed in any one of the preceding claims wherein the ligand is selected from the group consisting of: slightly soluble salt of a polyvalent or divalent metal selected from the group consisting of: slightly soluble salts of barium, calcium, cadmium, cobalt, chrome, copper, iron, lead, mercury, magnesium, manganese, nickel, radium, strontium, zinc, scandium, yttrium, lanthanum and other lanthanides, actinium and other actinides, bismuth and indium.

13. A method as claimed in claim 12 wherein the ligand is provided as a slightly soluble metal salt selected from the group consisting of: calcium, strontium, yttrium, lanthanum or barium citrate, oxalate, sulphate, phosphate and hydrogen phosphate.

14. A method as claimed in any one of claims 1 to 13 wherein the tag-entity is appended to the target via a linker.

15. A method as claimed in any one of claims 1 to 14 wherein the target and the tag-entity are polypeptides and the tag-entity is appended to the target optionally by expression of a fusion protein comprising said target incorporating the tag-entity.

16. A method as claimed in any one of claims 1 to 15 further comprising the step of resolubilising the tagged target by either:

(i) removing the tag-entity, or

(ii) removing the ligand.

17. A method as claimed in claim 16 wherein the tagged target is a polypeptide and the tag-entity is removed by proteolysis.

18. A method as claimed in claim 16 wherein the ligand is a metal ion and this is removed from the tagged target by addition of a counter ion, wherein the metal salt of the counter ion is largely insoluble.

19. An artificial polypeptide which is optionally linked by its amino- or carboxy-terminus to another polypeptide or protein and which is be defined by the following general amino acid sequence (SEQ ID No 1):

$\mathbf{X}_m$-(E or D)-(D or E)-(V or I or L)-(D or E)-(E or D)-C-$\mathbf{X}_z$-C-(K or R or H)-(D or E)-$\mathbf{X}_n$;

wherein $X_m$, $X_z$ and $X_n$ can be any amino acid sequences comprising ***m, z*** and ***n*** amino acid residues, respectively; and wherein ***m*** and ***n*** can be any whole numbers from 0 to 100, and ***z*** can be any whole number from 1 to 25.

20. An artificial polypeptide according to claim 19, wherein said polypeptide has the following amino acid sequence (SEQ ID No 2):

$\mathbf{X}_m$-E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D-$\mathbf{X}_n$;

wherein $X_m$ and $X_n$ can be any amino acid sequences comprising ***m*** and ***n*** amino acid residues, respectively; and wherein ***m*** and ***n*** can be any whole numbers from 0 to 100.

21. An artificial polypeptide according to claim 20, wherein said polypeptide has the following amino acid sequence (SEQ ID No 3):

E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D.

22. A tagged target comprising:

(i) a target entity,

(ii) a heterologous tag-entity which is a polypeptide as defined in any one of claims 6 to 11.

23. A tagged target as claimed in claim 22 wherein the tag-entity is a polypeptide as defined in any one of claims 19 to 21.

24. A tagged target as claimed in claim 22 or claim 23 wherein the tag-entity is linked to a protein target via a covalent bond.

25. A tagged target as claimed in claim 24 wherein the said covalent bond is a disulphide bond.

26. A tagged target as claimed in claim 24 wherein the said covalent bond is a peptide bond and the tagged target is a fusion protein.

27. A fusion protein according to claim 26, wherein said tag-entity is linked to the protein target through a polypeptide linker.

28. A fusion protein according to claim 27, wherein said polypeptide linker comprises a site which can be recognized and cleaved by a proteolytic agent.

29. A fusion protein according to claim 28, wherein said site is Ile-Glu-Gly-Arg (SEQ ID No 4) and the proteolytic agent is blood coagulation factor Xa.

30. A fusion protein according to claim 28, wherein said site is Glu-X-X-Tyr-X-Gln/Ser, (SEQ ID No 17) and proteolytic agent is tomato etch virus (TEV) protease.

31. A nucleic acid sequence coding for an articificial polypeptide, tagged target or fusion protein as claimed in any one of claims 19 to 30.

32. A nucleic acid sequence as claimed in claim 31 which comprises a sequence shown as SEQ ID No 7, 8, 10, or 12.

33. A recombinant vector comprising the nucleic acid of claim 31 or 32.

34. A process for making a recombinant vector encoding a tagged target, the process comprising the steps of:

(i) providing nucleic encoding a target entity

(ii) appending nucleic acid encoding a heterologous tag-entity as defined in any one of claims 5 to 11 to it.

35. A host cell comprising the vector of claim 33 or obtainable by the process of claim 34.

36. A process for making a tagged target comprising:

(i) providing a target entity, and

(ii) appending a heterologous tag-entity as defined in any of claims 6 to 11 to it.

37. A process for producing a recombinant fusion protein according to any one of claims 27 to 30 comprising cultivating the host cell of claim 35 under conditions for expression of the recombinant fusion protein, recovering from the host cell and purifying said protein.

38. A process for producing an enriched or purified tagged target according to claim 1 comprising co-precipitation of the tagged target with a slightly soluble salt of a polyvalent or divalent metal.

39. A process for producing tagged-target in enriched form from a sample, which process comprises:

(i) precipitating the target entity by performing the method of any one of claims 4 to 18, and

(ii) separating precipitate from the sample.

40. A process as claimed in claim 38 and claim 39 wherein the target is a protein complex comprising a least 2 binding partners, where the heterologous tag-entity is appended to 1 of the binding partners.

41. A process as claimed in any one of claims 38 to 40 wherein the target is a viral protein.

42. A process according to claim 38, wherein said slightly soluble salt of polyvalent or divalent metal can is selected from the group consisting of: slightly soluble salts of barium, calcium, cadmium, cobalt, chrome, copper, iron, lead, mercury, magnesium, manganese, nickel, radium, strontium, zinc, scandium, yttrium, lanthanum and other lanthanides, actinium and other actinides, bismuth and indium.

43. A process according to claim 42, wherein said slightly soluble salt is selected from the group consisting of: calcium, strontium, yttrium, lanthanum or barium citrate, oxalate, sulphate, phosphate and hydrogen phosphate.

44. A kit for purifying a tagged target according to claim 23, wherein the kit comprises, in separate containers, reaction components, which are necessary to generate a slightly soluble salt defined in said claims 42 or 43, and optionally a nucleic acid of claim 31.

45. A kit according to claim 44 comprising in separate containers:

a solution containing a soluble salt(s) of the metal selected from the group of barium, calcium, cadmium, cobalt, chrome, copper, iron, lead, mercury, magnesium, manganese, nickel, radium, strontium, zinc, scandium, yttrium, lanthanum and other lanthanides, actinium and other actinides, bismuth and indium;

another solution containing a soluble salt(s) selected from the group of phosphate(s), hydrogen phosphate(s), citrate(s), oxalate(s) and sulphate(s).

46. A method for producing or detecting the binding partner of a target entity from a sample, the method comprising:

(i) enriching or recovering the target entity from the sample by performing the method of any one of claims 4 to 18,

(ii) determining if one or more binding partners are bound thereto, and

(iii) optionally isolating and\or identifying the binding partner(s).

47. A method of screening an agent for its ability to modulate a biological activity of a target, the method comprising:

(i) enriching or recovering the target entity from the sample by performing the method as claimed in any one of claims 1 to 18, and

(ii) comparing the biological activity of the target in the presence and absence of the the agent.

48. A method of isolating a tagged target as claimed in any one of claims 23 to 26 from a sample, the method comprising exposing the tagged target an immobilized metal ion affinity chromatography (IMAC), wherein the immobilized metal ion is a ligand for the tag entity.

49. An artificial polypeptide wherein said polypeptide has the following amino acid sequence (SEQ ID No 3):

E-D-V-D-E-C-S-E-N-M-S-A-Q-L-C-K-D.

50. A tagged polypeptide comprising:

(i) a target polypeptide,

(ii) a heterologous tag-entity which is a polypeptide as defined in claim 49.

51. A tagged polypeptide as claimed in claim 50 wherein the target polypeptide and the heterologous tag-entity are linked by a peptide bond and the tagged target is a fusion protein.

52. A fusion protein according to claim 51, wherein said tag-entity is linked to the protein target through a polypeptide linker.

53. A fusion protein according to claim 52, wherein said polypeptide linker comprises a site which can be recognized and cleaved by a proteolytic agent.

54. A fusion protein according to claim 53, wherein said site is Ile-Glu-Gly-Arg (SEQ ID No 4) and the proteolytic agent is blood coagulation factor Xa.

55. A fusion protein according to claim 53, wherein said site is Glu-X-X-Tyr-X-Gln/Ser, (SEQ ID No 17) and the proteolytic agent is tomato etch virus (TEV) protease.

56. A nucleic acid sequence coding for the fusion protein as claimed in claim 51.

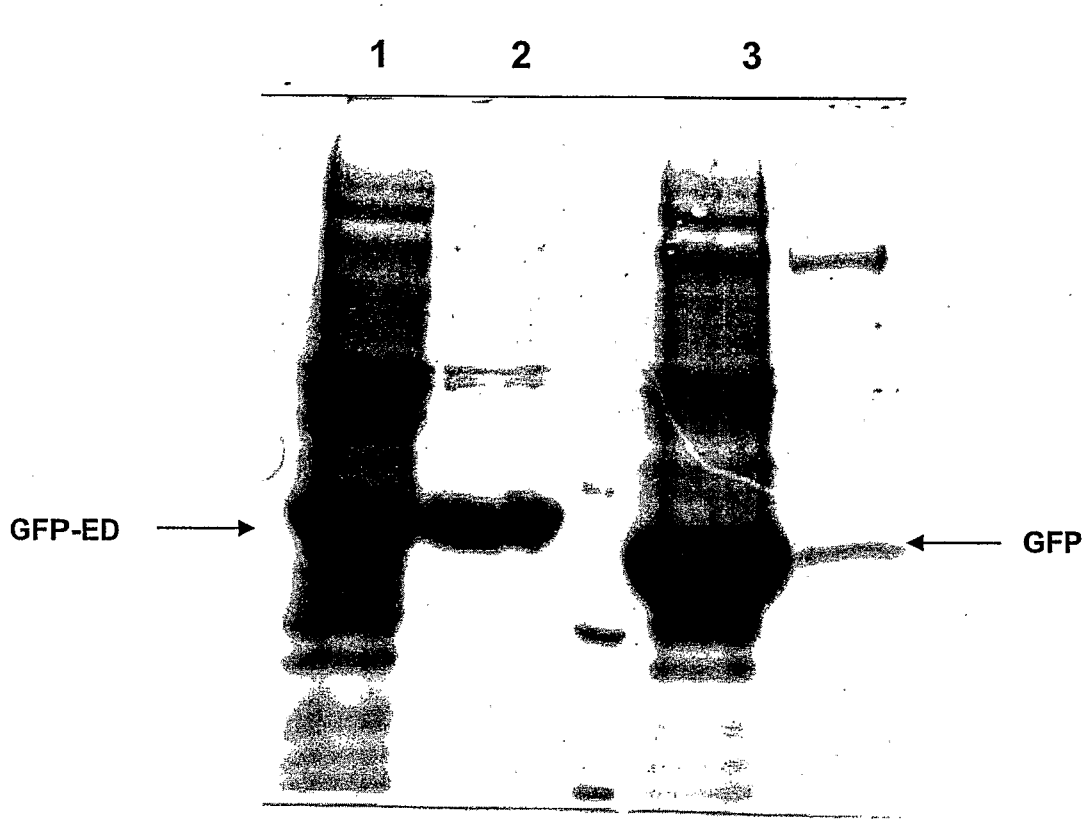
57. A nucleic acid sequence as claimed in claim 56 which comprises a sequence shown as SEQ ID No 7, 8, 10, or 12.

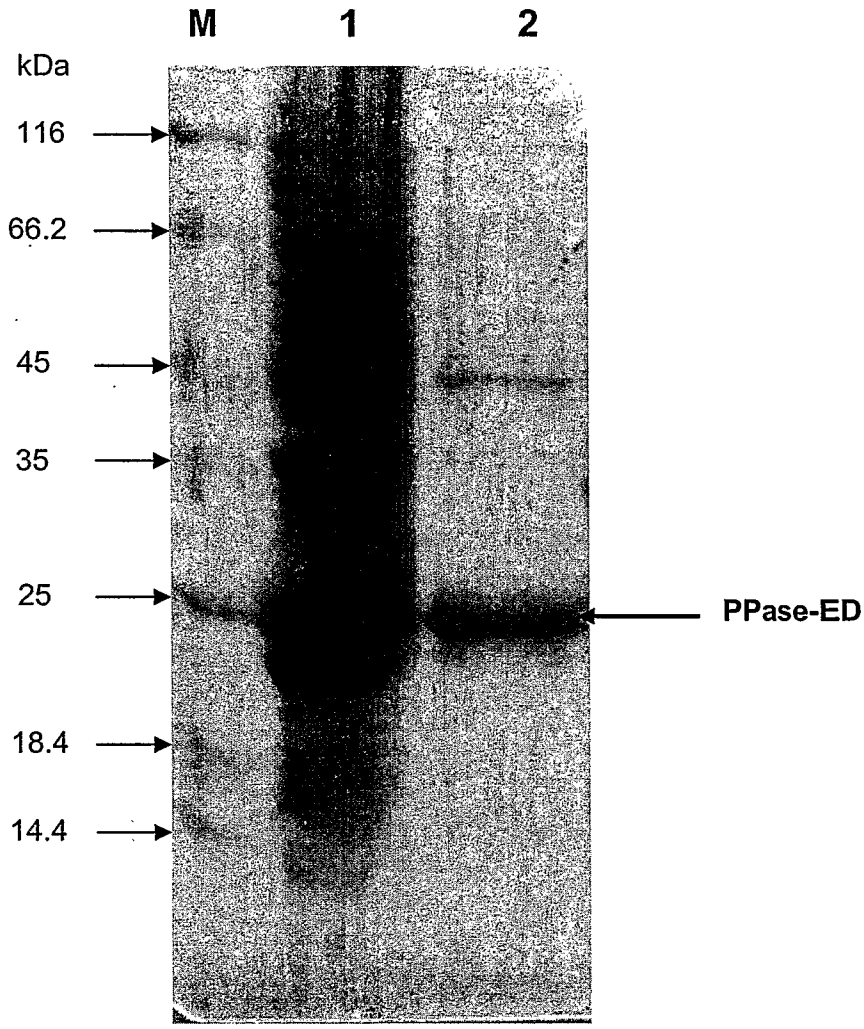
58. A recombinant vector comprising the nucleic acid of claim 56.

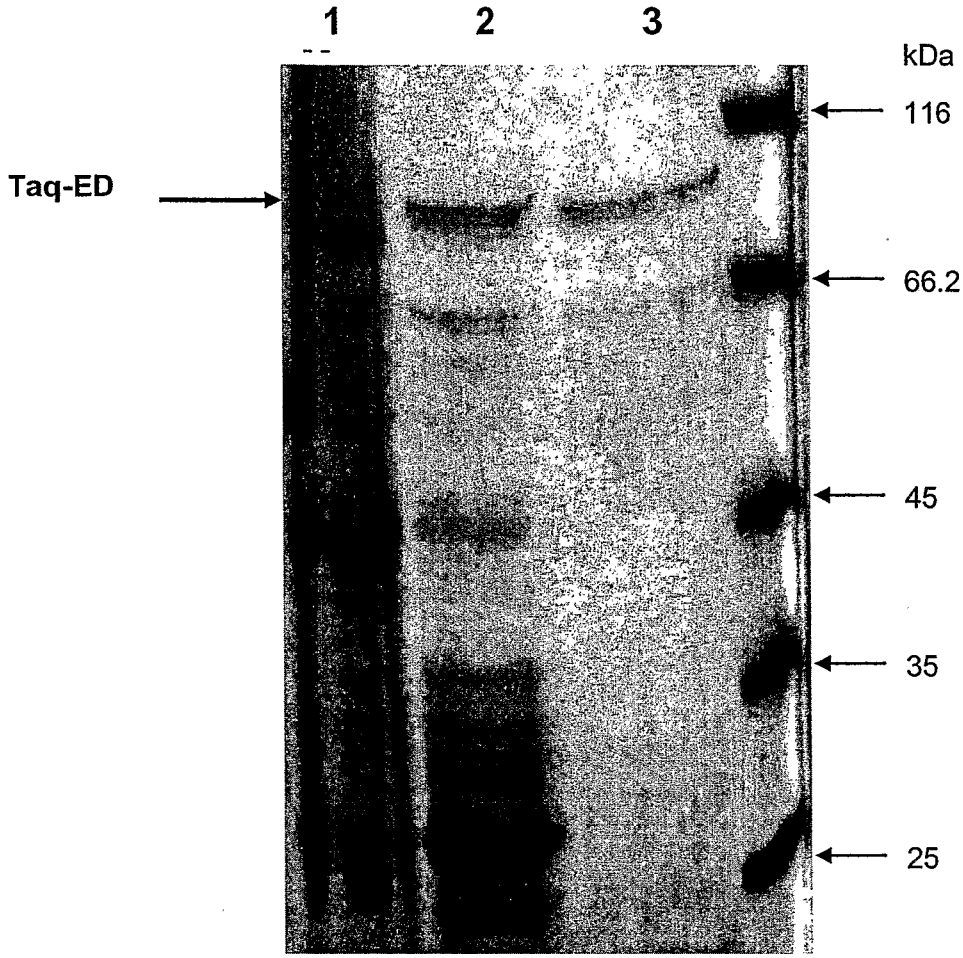
59. A host cell comprising the vector of claim 58.

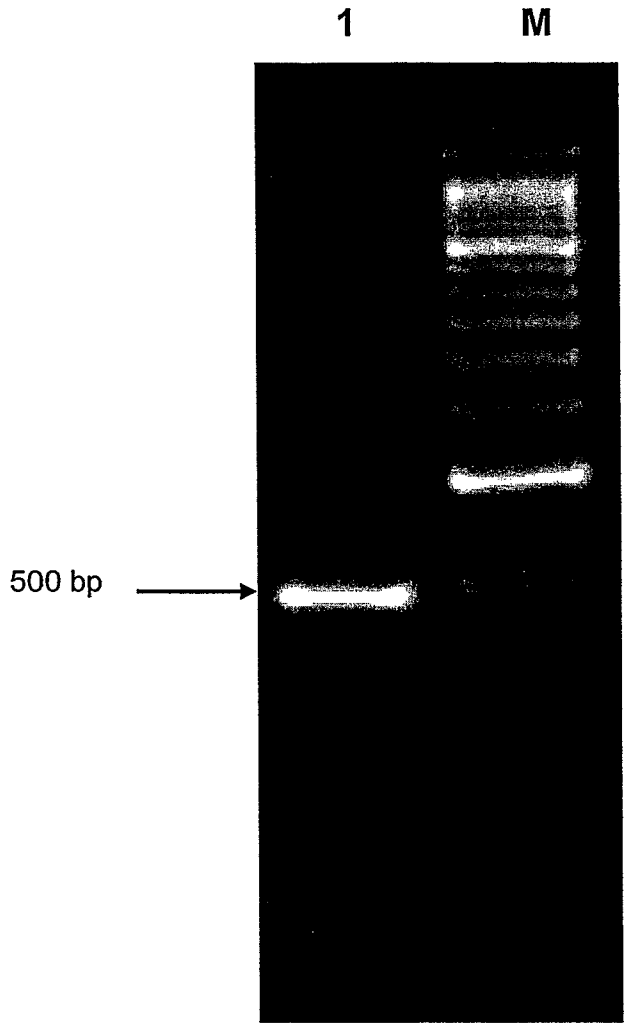
60. A process for producing an enriched or purified tagged polypeptide according to claim 50 comprising co-precipitation of the tagged polypeptide with a slightly soluble salt of a polyvalent or divalent metal.

Fig. 1




M
1
2
kDa
116
66.2
45
35
25
18.4
14.4
PPase-ED


1
2
3
kDa
116
66.2
45
35
25
Taq-ED


1
M
500 bp

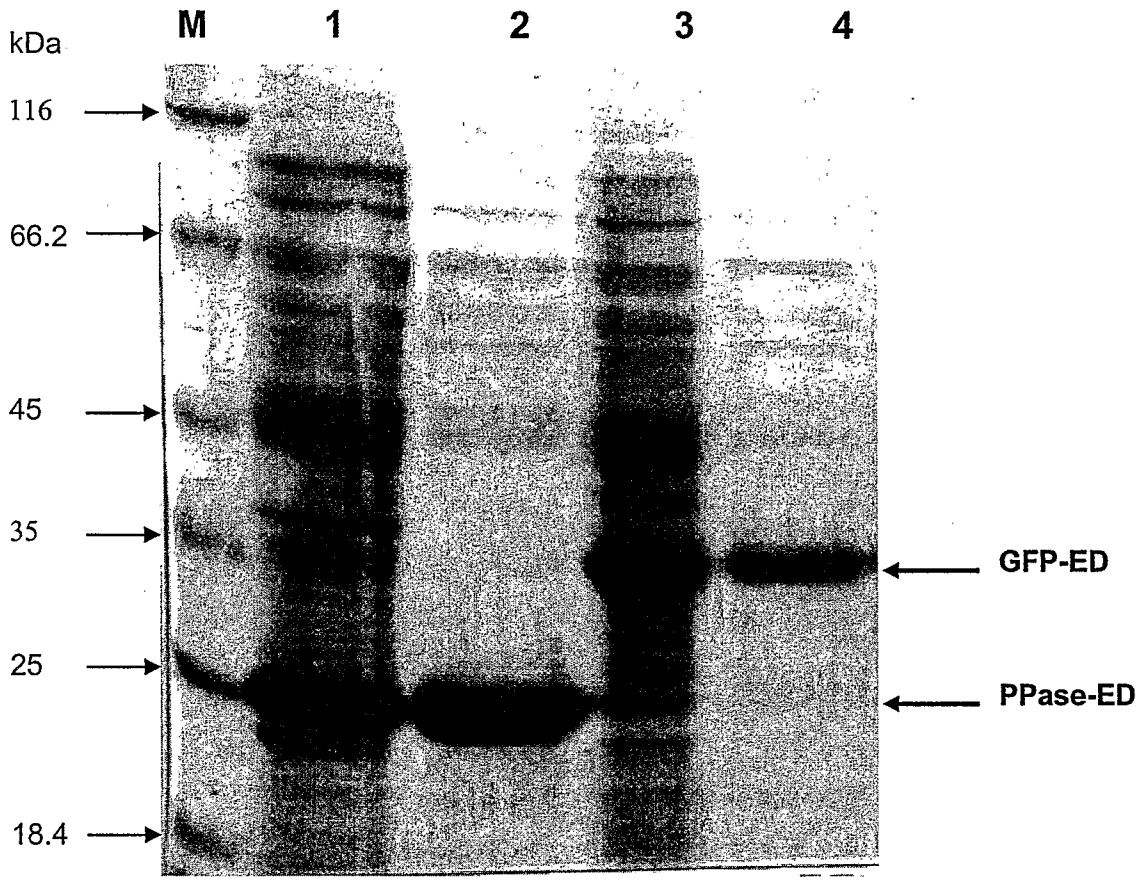
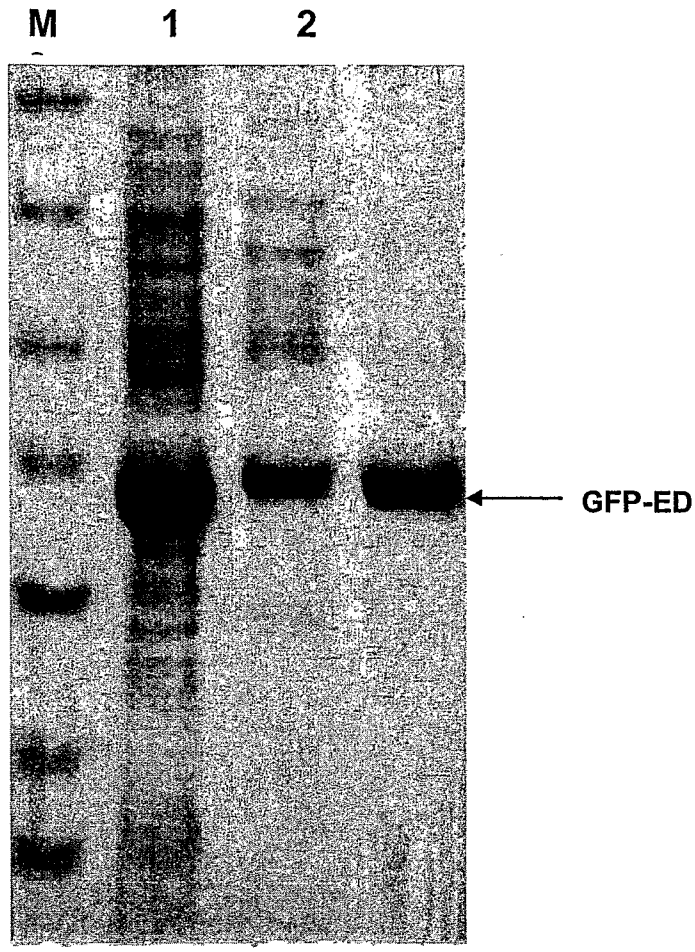

M
1
2
3
4
kDa
116
66.2
45
35
25
18.4
GFP-ED
PPase-ED

Fig. 6



# The mechanism of calcium dependent, selective precipitation of tagged proteins

