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

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
15 September 2011 (15.09.2011)

PCT

(10) **International Publication Number**
WO 2011/113027 A2

(51) **International Patent Classification**:
C07K 19/00 (2006.01) *C12N 15/85* (2006.01)
C12N 15/62 (2006.01) *A61P 9/00* (2006.01)
C12N 15/55 (2006.01)

(21) **International Application Number**:
PCT/US2011/028233

(22) **International Filing Date**:
11 March 2011 (11.03.2011)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/340,066 12 March 2010 (12.03.2010) US

(71) **Applicant** *(for all designated States except US)*: **SYNAGEVA BIOPHARMA CORP** [US/US]; Synageva BioPharma Corp., 128 Spring Street, Suite 520, Lexington, Massachusetts 02421 (US).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **QUINN, Anthony** [GB/US]; 107 Manet Road, Chestnut HIll, Massachusetts 02467 (US). **HARVEY, Alex J.** [US/US]; 106 Fox Trace, Athens, Georgia 30606 (US). **XIA, Zhinan** [US/US]; 29 Burke Lane, Wellesley, Massachusetts 02481 (US).

(74) **Agent**: **CHANG, Hak J.**; Synageva BioPharma Corp., 128 Spring Street, Suite 520, Lexington, Massachusetts 02421 (US).

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

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

[Continued on next page]

(54) Title: NPP1 FUSION PROTEINS

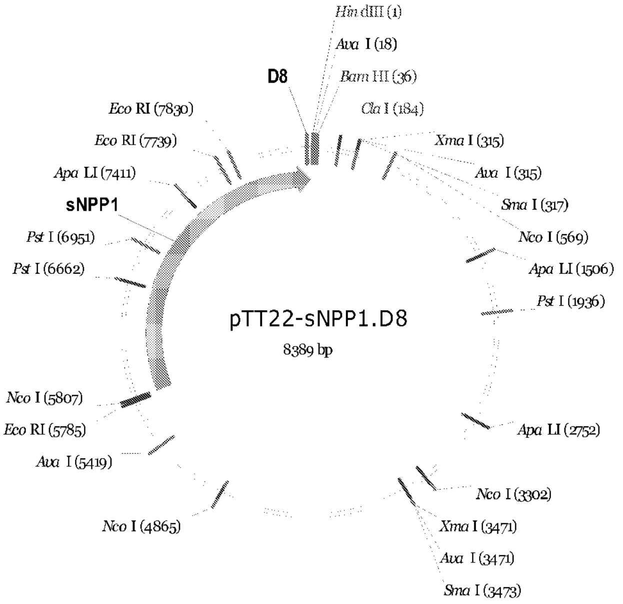

Fig. 19

(57) **Abstract**: The present invention provides a novel fusion polypeptide containing a catalytic portion of NPP1 fused to a targeting moiety, nucleic acids encoding the fusion polypeptide, a vector containing the nucleic acid integrated thereinto, a host cell transformed with the vector and pharmaceutical compositions comprising the fusion polypeptide.

WO 2011/113027 A2

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

NPP1 FUSION PROTEINS

RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application No. 61/340,066, filed on March 12, 2010. The entire teachings of the above application are incorporated herein by reference.

BACKGROUND OF THE INVENTION

Ectonucleotide pyrophosphatase/phosphodiesterase 1 (NPP1/ENPP1/PC-1) is a type II transmembrane glycoprotein that forms a homodimer. The protein cleaves a variety of substrates, including phosphodiester bonds of nucleotides and nucleotide sugars and pyrophosphate bonds of nucleotides and nucleotide sugars. NPP1 protein functions to hydrolyze nucleoside 5' triphosphtase to either corresponding monophosphates and also hydrolyzes diadenosine polyphosphates. Mutations in the NPP1 gene have been associated with idiopathic infantile arterial calcification (IIAC), insulin resistance, hypophosphatemic rickets, and ossification of the posterior longitudinal ligament of the spine.

IIAC, a rare autosomal recessive and nearly always fatal disorder, is characterized by calcification of the internal elastic lamina of muscular arteries and stenosis due to myointimal proliferation. There are more than 160 cases of IIAC that have been reported world-wide. The symptoms of the disease most often appear by early infancy, and the disease is lethal by 6 months of age, generally because of ischemic cardiomyopathy, and other complications of obstructive arteriopathy including renal artery stenosis. In more than a dozen reported cases of IIAC, peri-articular calcifications of large joints also developed in infancy. Rutsch *et al.* (2003) reported that mutations in ENPP1 are associated with IIAC characterized by the spontaneous periarticular and aortic calcifications in early life and systemic lowering of nucleotide pyrophosphatase/phosphodiesterase activity in the affected individuals.

Although defects in the NPP1 protein have been implicated in such serious disease as IIAC, there is no treatment available in the art for those who are affected by the disease. Therefore, a dire need exists for an effective and safe composition, formulation and medicament for the treatment of IIAC, insulin resistance, hypophosphatemic rickets, and ossification of the posterior longitudinal ligament of the spine.

SUMMARY OF THE INVENTION

The present invention includes fusion proteins of truncated domains of NPP1 (*i.e.*, an NPP1 component) fused to a targeting moiety. The targeting moiety functions to enhance the efficiency in targeting the NPP1 fusion protein to a site clinical or biological importance (*e.g.*, site of calcification in a subject that needs lowering of calcification). Without wishing to limit the invention to any particular theory or mechanism of operation it is believed that the NPP1 component function to inhibit calcification by enhancing the formation of pyrophosphate (PPi) and/or by cleaving pyrophosphate to produce soluble phosphate (Pi) and/or by increasing the availability of adenosine monophosphate (AMP) and/or adenosine. It is contemplated that the targeting moiety can be attached to the N-terminus and/or the C-terminus of the NPP1 component at any useful position. Additionally, the NPP1 fusion protein disclosed herein can also include one or more of Fc fragment, PEG, polypeptide linker or other additional polypeptides to enhance the enzymatic activity, stability or targeting.

The fusion proteins of the invention can be used to treat a wide variety of conditions in a subject. Any condition that can be beneficially treated by the administering of a fusion protein of the invention is included within the scope of the invention. For example, treatment of conditions that can be improved by reducing and/or eliminating one or more calcification structures and/or preventing calcification structures from forming in a subject such as a mammal, for example, a human patient is within the scope of the invention. Conditions such as arterial blockage are contemplated for treatment by employing fusion proteins of the invention. In one particularly useful embodiment, the condition to be treated is generalized arterial calcification of infancy also referred to as idiopathic arterial calcification of infancy and arterial media calcification of infancy. Conditions such as insulin resistance, hypophosphatemic rickets, and ossification of the posterior longitudinal ligament of the spine are also contemplated for treatment.

Fusion proteins of the invention can be produced in any useful protein expression system including, without limitation, cell culture (*e.g.*, CHO cells, COS cells, HEK203), bacteria such as *Escherichia coli* (*E. coli*) and transgenic animals, including, but no limited to, mammals and avians (*e.g.,* chickens, quail, duck and turkey).

The manufacture of pharmaceutical compositions (or pharmaceutical formulations) is well known in the art and such pharmaceutical compositions are contemplated for use in accordance with fusion proteins of the invention.

Generally, the dosage of fusion protein administered to a subject will vary depending upon known factors such as age, health and weight of the recipient, type of concurrent treatment, frequency of treatment, and the like. Usually, a dosage of active ingredient (*i.e.*, fusion protein) can be between about 0.0001 and about 50 milligrams per kilogram of body weight. Precise dosage, frequency of administration and time span of treatment can be determined by a physician skilled in the art of administration of therapeutic proteins.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates the amino acid sequence of wild-type NPP1 protein. The cytosolic and transmembrane regions are underlined. The potential N-glycosylation sites are in bold. PSCAKE in italics is the start of soluble NPP1 which includes the cysteine rich region.

Fig. 2 illustrates the amino acid sequence of the catalytic domain(s) of NPP1 protein having no target moiety attached (“sssNPP1”).

Fig. 3 illustrates the amino acid sequence of TAGsssNPP1. The targeting moiety of eight consecutive aspartic acid residues is fused to the N-terminus of the sssNPP1. The signal peptide is underlined and the targeting moiety is indicated in bold.

Fig. 4 illustrates the amino acid sequence of TAGsssNPP1 that contains the targeting moiety of ten consecutive aspartic acid residues fused to the C-terminus of the sssNPP1. The signal peptide is underlined and the targeting moiety is indicated in bold.

Fig. 5 illustrates the nucleic acid sequence of TAGssNPP1 fusion protein. The targeting moiety of eight consecutive aspartic acid residues is fused to the N-terminus of the ssNPP1.

Fig. 6 illustrates the amino acid sequence of TAGssNPP1 fusion protein. The targeting moiety of eight consecutive aspartic acid residues is fused to the N-terminus of the ssNPP1. The signal peptide is underlined and the targeting moiety is indicated in bold.

Fig. 7 illustrates the nucleic acid sequence of ssNPP1.

Fig. 8 illustrates the amino acid sequence of ssNPP1. The signal peptide is underlined.

Fig. 9 illustrates the nucleic acid sequence of sNPP1.

Fig. 10 illustrates the amino acid sequence of sNPP1. The signal peptide is underlined.

Fig. 11 illustrates the nucleic acid sequence of TAGsNPP1. The targeting moiety of eight consecutive aspartic acid residues is fused to the N-terminus of the sNPP1.

Fig. 12 illustrates the amino acid sequence of TAGsNPP1. The targeting moiety of eight consecutive aspartic acid residues is fused to the N-terminus of the ssNPP1. The signal peptide is underlined and the targeting moiety is indicated in bold.

Fig. 13 illustrates the nucleic acid sequence of TAGsNPP1. The targeting moiety of eight consecutive aspartic acid residues is fused to the C-terminus of the sNPP1.

Fig. 14 illustrates the amino acid sequence of TAGsNPP1. The targeting moiety of eight consecutive aspartic acid residues is fused to the N-terminus of the sNPP1. The signal peptide is underlined and the targeting moiety is indicated in bold.

Fig. 15 illustrates the amino acid sequence of a linker peptide.

Fig. 16 illustrates the amino acid sequence of an immunoglobulin Fc segment.

Fig. 17 illustrates the amino acid sequence of TAGsssNPP1 which contains the targeting moiety of eight consecutive aspartic acid residues fused to the N-terminus of the sssNPP1 via a peptide linker. The Fc segment is fused to N-terminus of the target moiety. The signal peptide is underlined and the targeting moiety is indicated in bold. The peptide linker is in italics.

Fig. 18 illustrates the amino acid sequence of TAGsssNPP1 which contains the targeting moiety of eight consecutive aspartic acid residues fused to the C-terminus of the sssNPP1 via a peptide linker. The Fc segment is fused to N-terminus of the sssNPP1. The signal peptide is underlined and the targeting moiety is indicated in bold. The peptide linker is in italics.

Fig. 19 is a schematic representation of an expression vector (*i.e.*, pTT22) containing a TAGsNPP1 construct. The targeting moiety of eight consecutive aspartic acid residues fused to the C-terminus of the sNPP1.

Fig. 20 illustrates Western blot analysis of TAGsNPP1. R, reducing condition; NR, non-reducing condition.

Fig. 21 demonstrates the enzymatic activity of TAGsNPP1 produced and isolated from HEK293 cells.

Figs. 22A-22C illustrate schematics of TAGNPP1 fusion protein constructs described herein.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides novel human NPP1 fusion proteins that are soluble and contain truncated and biologically active domain(s) of NPP1 (*i.e.*, NPP1 components that contain at least one extracellular catalytic domain of naturally occurring NPP1 for the pyrophosphatase and/or phosphodiesterase activity) and one or more targeting moieties (*i.e.*, “TAG”). The NPP1 fusion proteins of the present invention comprise at least the NPP1 domain essential to carry out the pyrophosphatase and/or phosphodiesterase activity. Accordingly, the invention features isolated fusion proteins comprising the amino acid residues A205 through L591 of SEQ ID NO:1 fused to one or more targeting moieties. The targeting moiety can be recombinantly fused or chemically bonded (*e.g.*, covalent bond, ionic bond, hydrophobic bond and Van der Waals force) to the NPP1 component by methods well known in the art and direct the NPP1 component to a certain target site where the attached NPP1 component will have a desirable effect (*e.g.*, catalysis of a reaction such as solubilizing a substrate such as PPi or preventing the formation of a substrate such as PPi) in a subject to which the fusion protein of the present invention is administered.

TAGNPP1s

All NPP1 fusion proteins ("TAGNPP1s") of the present invention have the N-terminal cytosolic and the transmembrane domains of the naturally occurring human NPP1 removed. Optionally, TAGNPP1s fusion proteins of the present invention can also contain C-terminal truncation of wild-type NPP1 in various lengths. The amino acid sequence of full-length wild-type NPP1 is set forth in SEQ ID NO: 1.

In one embodiment, the fusion protein contains a polypeptide comprising the amino acid residues A205 through L591 of SEQ ID NO:1 ("sssNPP1") fused a TAG on either the N- or C-terminus of the polypeptide ("TAGsssNPP1"). In one embodiment, the fusion protein comprises a polypeptide comprising the amino acid residues A205 through D925 of SEQ ID NO:1 ("ssNPP1") fused to a TAG on either the N- or C-terminus ("TAGssNPP1"). In one embodiment, the fusion protein comprises a polypeptide comprising the amino acid residues P99 through D925 of SEQ ID NO:1 ("sNPP1") fused to a TAG on either the N- or C- terminus of the polypeptide ("TAGsNPP1"). Also contemplated is any consecutive fragment of sNPP1 that comprises at least the amino acid resides A205 through L591 of SEQ ID NO:1 and the polypeptide fragment is fused to a TAG on either the N- or C-terminus.

When expressed in a cell culture or transgenic animal, the TAGNPP1 fusion proteins can further comprise a signal peptide (or leader sequence) at its N-terminus. The signal peptide co-translationally or post-translationally directs transport of the TAGNPP1 fusion proteins through the subcellular organelles of the cell expressing the TAGNPP1 fusion proteins and, thereby determining the post-translational modification of the TAGNPP1 fusion proteins. It is to be understood that because the signal peptide is cleaved at the co-translational or post-translational stage of the fusion protein, the TAGNNP1 fusion proteins are generally devoid of the signal peptide when once secreted and isolated. Accordingly, in the embodiments that are directed to the nucleic acid sequences encoding the TAGNPP1 fusion proteins are described, the leader sequences are also contemplated as used in the present invention. For example, the nucleotide sequence set forth in SEQ ID NO:2 contains an example of the leader sequence for TAGNNP1 at its 5' end.

Each of the fusion proteins disclosed herein is contemplated with one or more targeting moieties ("TAG"). The TAG component according to the present invention comprises four or more negatively charged amino acids such as aspartic acid and glutamic acid. The TAG component can be a stretch of negatively charged amino acid residues, for example, aspartic acids and/or glutamic acids that are between about 4 and about 20 amino acid residues in length. The TAG can be fused to either the N- or C-terminus of the NPP1 component. The TAG can be also fused to both N- and C-termini of the NPP1 component. Accordingly, the amino acid sequence of the TAGsNPP1 fusion protein, for example, includes PSCAKE through the C-terminus end of the NPP1 component and one or more targeting moieties (*e.g.*, polyglutamic acid tag or polyaspartic acid tag) at the N- and/or C-terminus end of the NPP1 component. The fusion protein comprising the NPP1 component having the TAG fused to the C-terminus is a particularly useful embodiment. In one very specific embodiment, TAG having a stretch of eight aspartic acids is employed as can be seen in the exemplary sequences for TAGsNPP1 and TAGssNPP1 in Figures, though any useful number of negatively charged amino acid residues (*e.g.*, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18) can be used in accordance with the invention. The TAG component is indicated as "A" in Figures.

The invention also encompasses polynucleotides which encode various TAGNPP1 fusion proteins described herein. Accordingly, any nucleic acid sequence which encodes the amino acid sequence of any TAGNPP1 fusion protein can be used to generate recombinant molecules which express the corresponding TAGNPP1 fusion protein. In a particular embodiment, the invention encompasses the polynucleotide comprising the nucleic acid sequence of SEQ ID NO:2 as shown in FIG. 2.

Within certain specific embodiments, the fusion protein comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence. Certain preferred polypeptide can, for example, assist in dimerization and stability or minimize aggregation of the fusion proteins. For example, the additional polypeptide can be the Fc region of the immunoglobulin G1 to increase stability in serum. The use of the Fc segment is well known in the art and described in U.S. Pat. No. 7,902,151; and U.S. Pat. No. 7,858,297, the entire teachings of which are incorporated herein by reference in their entirety. The cysteine rich region of wild-type NPP1 (*i.e.*, PSCAKE through NEPQCP; the amino acid sequence from P99 to P204 of SEQ ID NO:1) can be employed to facilitates dimerization of the TAGNPP1 fusion proteins.

In another embodiment, the polyethylene glycol (PEG) can be conjugated to the TAGNPP1 fusion proteins. Other polypeptides may be selected so as to minimize aggregation and immunogenicity, to increase the solubility of the protein, or to enable the protein to be targeted to desired sites of clinical or biological importance.

TAGNPP1 can be also fused or conjugated to an appropriate polypeptide linker or other sequence for ease of identification, synthesis, or purification of the fusion protein, or to better preserve the native structure of the NPP1 component which can enhance the activity and targeting of the TAGNPP1. Specifically, a peptide linker sequence can be employed to separate the first and the second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein between the NPP1 component and the TAG component using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on their ability to adopt a flexible extended conformation and their inability to adopt a secondary structure that could interact with functional portion on the NPP1, TAG or other secondary polypeptides described herein (*e.g.*, Fc). Preferred peptide linker sequences contain Gly, His, Asn and Ser residues. The useful peptide linkers include, without limitation, poly-Gly, poly-His, poly-Asn, or poly-Ser. Other near neutral amino acids, such as Thr and Ala can be also used in the linker sequence. Amino acid sequences which can be usefully employed as linkers include those disclosed in Maratea *et al.*, Gene 40:39-46, 1985; Murphy *et al.*, *Proc. Natl. Acad Sci.* USA 83:8258-8262, 1986; U.S. Pat. No. 4,935,233 and U.S. Pat. No. 4,751,180. The linker sequence may be from 1 to about 20 amino acid residues in length. Preferably, the polypeptide linker is between about 8 and about 12 amino acids in length. In a preferred embodiment, the peptide linker used in the invention is GGGGSGGGGS (SEQ ID NO:15), although any functional combination of Gly, Ser, His, or Asn can be employed.

Fusion proteins can also comprise a TAGNPP1 of the present invention together with an unrelated polypeptide. Preferably the unrelated polypeptide is capable of enhancing the targeting of the fusion protein to the site of clinical or biological importance (*e.g.*, site of calcification). For example, peptides that have high affinity to the bone are described in U.S. Pat. No. 7,323,542, the entire teachings of which are incorporated herein by reference.

TAGNPP1 can be prepared using standard methods, including recombinant techniques or chemical conjugation well known in the art. Techniques useful for isolating and characterizing the nucleic acids and proteins of the present invention are well known to those of skill in the art and standard molecular biology and biochemical manuals can be consulted to select suitable protocols for use without undue experimentation. See, for example, Sambrook et al, 1989, "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor, the content of which is herein incorporated by reference in its entirety. Briefly, DNA sequences encoding the polypeptide components can be assembled separately, and ligated into an appropriate expression vector. For example, the 3' end of the DNA sequence encoding the NPP1 component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component such as TAG PEG, or Fc so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides. The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements including a promoter. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptide such as the leader sequence encoding a signal peptide. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

The invention also encompasses TAGNPP1 variants. A preferred TAGNPP1 variant is one having 80%, 85%, 90%, 95% and more preferably 96% amino acid sequence identity to the amino acid sequence A205 through L591 of SEQ ID NO:1. A most preferred TAGNPP1 variant is one having at least 97% amino acid sequence identity to amino acid sequence A205 through L591 of SEQ ID NO:1.

The invention also relates to a polynucleotide sequence comprising the complement of SEQ ID NO:2 or variants thereof. In addition, the present invention also features polynucleotide sequences which hybridize under stringent conditions to SEQ ID NO:2 and whose the antisense sequence is 85%, 90%, 95%, 97%, 98%, or 99% identical to SEQ ID NO:2. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Wahl, G. M. and S. L. Berger (1987; *Methods Enzymol.* 152:399-407) and Kimmel, A. R. (1987; *Methods Enzymol.* 152:507-511), and can be used at a defined stringency.

The invention additionally contemplates nucleic acid sequences encoding polypeptides, oligonucleotides, peptide nucleic acids (PNA), fragments, portions or antisense molecules thereof.

Although nucleotide sequences which encode TAGNPP1 and its variants are preferably capable of hybridizing to the nucleotide sequence of the TAGNPP1 under appropriately selected conditions of stringency, it can be advantageous to produce nucleotide sequences encoding TAGNPP1 or its derivatives possessing a substantially different codon usage. Codons can be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding TAGNPP1 and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life.

Altered nucleic acid sequences encoding TAGNPP1 which are encompassed by the invention include deletions, insertions, or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent of TAGNPP1. The encoded protein can also contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent TAGNPP1. Deliberate amino acid substitutions can be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of TAGNPP1 is retained. For example, positively charged amino acid residues include Lys and Arg; negatively charged amino acid residues include Asp and Glu; and amino acids with uncharged polar head groups having similar hydrophilicity can include Leu, Ile, and Val; Gly and Ala; Asp and Gln; Ser and Thr; Phe and Tyr.

Expression vector

Methods which are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding TAGNPP1 and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. *et al.* (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y., the teachings of which are incorporated herein by reference in its entirety.

A variety of expression vector/host systems can be utilized to contain and express sequences encoding TAGNPP1. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (*e.g.*, baculovirus) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems (*e.g.*, pTT22 vector).

The control elements or regulatory sequences can includes those non-translated regions of the vector-enhancers, promoters, 5' and 3' untranslated regions-which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including, tissue-specific, constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript^{TM} phagemid (Stratagene, LaJolla, California) or pSport1^{TM} plasmid (Gibco BRL) and the like can be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding TAGNPP1, vectors based on SV40 or EBV can be used with an appropriate selectable marker. When an avian expression system is used, suitable vectors for expression various TAGNPP1 constructs are described in U.S. Pat. No. 6,730,822; U.S. Pat. No. 6,825,396; U.S. Pat. No. 6,875,588; U.S. Pat. No. 7,294,507; U.S. Pat. No. 7,521,591; U.S. Pat. No. 7,534,929; and U.S. patent application Ser. No. 11/376,023, the entire teachings of which are incorporated herein by reference in their entirety. Briefly, when an avian expression system is employed to express TAGNPP1, suitable oviduct-specific promoters, for example, and without limitation, ovomucoid promoters, ovalbumin promoters, lysozyme promoters, conalbumin promoters, ovomucin promoters, ovotransferrin promoters and functional portions of each of these promoters are contemplated. Suitable non-specific promoters can include, for example and without limitation, cytomegalovirus (CMV) promoters, MDOT promoters and rous-sarcoma virus (RSV) promoters, murine leukemia virus (MLV) promoters, mouse mammary tumor virus (MMTV) promoters and SV40 promoters and functional portions of each of these promoters. Non-limiting examples of other promoters which can be useful in the present invention include, without limitation, Pol III promoters (for example, type 1, type 2 and type 3 Pol III promoters) such as H1 promoters, U6 promoters, tRNA promoters, RNase MPR promoters and functional portions of each of these promoters. Typically, functional terminator sequences are selected for use in the present invention in accordance with the promoter that is employed.

Host Cells

The present invention includes the production of soluble TAGNPP1 in a transgenic avian (*e.g.*, transgenic chicken) system as is well known in the art, for example, in U.S. Patent No. 7,534,929, the disclosure of which is incorporated in its entirety herein by reference. Production in the avian system (*e.g.*, in the avian oviduct) of an NPP1 component with or without to a targeting moiety (*e.g.*, ssNPP1, sNPP1, TAGsNPP1 and TAGssNPP1) is within the scope of the invention. Furthermore, TAGNPP1 produced in any useful protein expression system including, without limitation, transgenic avians, transgenic mammal, cell culture (*e.g.*, CHO cells, HEK293 cells, and COS cells), bacteria such as *E. coli*, transgenic animals such as mammals and avians (*e.g.*, chickens, quail, duck and turkey) and in plant systems including duck weed, is contemplated herein.

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed TAGNPP1s in the desired fashion. Such modifications of the polypeptide of TAGNNP1 include, without limitation, acetylation, carboxylation, sialylation, glycosylation, phosphorylation, lipidation, and acylation. Different host cells such as CHO, COS, HeLa, MDCK, HEK293, and W138, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, can be chosen to ensure the correct modification and processing of the fusion protein of the present invention. Avian tumor cell line is also contemplated as a host cell for expressing the polypeptide of the present invention. Examples of useful avian cell lines (*e.g.*, an avian oviduct tumor cell line) which can be employed in the present invention are described in U.S. Pat. Publication No. 2009/0253176, the entire teachings of which are incorporated herein by reference.

Production of TAGNPP1

TAGNPP1 can be produced using any of a variety of well-known techniques. TAGNPP1 encoded by DNA sequences as described above can be readily prepared from the DNA sequences using any of a variety of expression vectors described herein or well known to those of ordinary skill in the art. Expression can be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide of the present invention. Supernatants from suitable host/vector systems which secrete recombinant fusion protein or polypeptide into culture media can be first concentrated using a commercially available filter. Following concentration, the concentrate can be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. One or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

For high-yield production of recombinant proteins, stable expression is preferred. Cell lines stably expressing TAGNPP1 can be transformed using expression vectors which contain viral origins of replication and/or endogenous expression elements and/or a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. Methods of producing exogenous protein in mammalian cell lines are well known in the art. Illustrative examples of this and other aspects and embodiments of the present invention for the production of heterologous polypeptides such as TAGNPP1 fusion proteins in avian cells are fully disclosed in U.S. patent application serial No. 09/877,374, filed Jun. 8, 2001, published as U.S. 2002/0108132-A1 on Aug. 8, 2002, and U.S. patent application serial No. 10/251,364, filed Sep. 18, 2002, each of which are incorporated herein by reference in their entirety. Examples of producing exogenous proteins in avian tumor cell lines are also described in U.S. Pat. Publication No. 2009/0253176, the entire teachings of which are incorporated herein by reference in entirety.

The invention specifically contemplates the production of the TAGNPP1 proteins disclosed herein in a transgenic avian system. In one particularly useful embodiment, the invention is drawn to the production of TAGNPP1 which can be produced in the oviduct of a transgenic avian, such as a chicken, in accordance with the invention. Examples of producing exogenous proteins in transgenic avian expression system are also described in U.S. Pat. No. 6,730,822, the entire teachings of which are incorporated herein by reference in entirety. Briefly, a suitable avian vector described above that contains a nucleic acid sequence encoding a TAGNPP1 fusion protein, operably linked to a tissue-specific or constitutive promoter that drives expression of the encoding sequence in the chicken oviduct are introduced into chicken stage X embryonic cells. The transformed embryonic cells are incubated under conditions conducive to hatching live chicks. Live chicks are nurtured into a mature chimeric chicken which are mated with a non-transgenic chicken naturally or via artificial insemination. A transgenic chicken is identified by screening progeny for germ line incorporation of the protein encoding sequence. The transgenic progeny can be mated with another transgenic or a non-transgenic chicken to produce eggs containing the TAGNPP1 fusion protein. The TAGNPP1 is then isolated and purified by methods well known in the art. Accordingly, the invention provides recombinant TAGNPP1 fusion proteins that have been produced by transgenic avians.

Pharmaceutical Composition

The present invention also features pharmaceutical compositions comprising isolated and substantially purified TAGNPP1 or a pharmaceutically acceptatable salt thereof. Pharmaceutical composition of the present invention can also include a pharmaceutically acceptable carrier or expient therefor. Compositions comprising such carriers, including composite molecules, are formulated by well-known conventional methods (*see*, for example, Remington's Pharmaceutical Sciences, 14th Ed., Mack Publishing Co., Easton, Pa.), the entire teachings of which are incorporated herein by reference. The carrier may comprise a diluent. In one embodiment, the pharmaceutical carrier can be a liquid and the fusion protein may be in the form of a solution. The pharmaceutical carrier can be wax, fat, or alcohol. In another embodiment, the pharmaceutically acceptable carrier may be a solid in the form of a powder, a lyophilized powder, or a tablet. In one embodiment, the carrier may comprise a liposome or a microcapsule.

The pharmaceutical compositions can be in the form of a sterile lyophilized powder for injection upon reconstitution with a diluent. The diluent can be water for injection, bacteriostatic water for injection, or sterile saline. The lyophilized powder may be produced by freeze drying a solution of the fusion protein to produce the protein in dry form. As is known in the art, the lyophilized protein generally has increased stability and a longer shelf life than a liquid solution of the protein.

Definitions:

As used herein, the term “acceptable” with respect to a formulation, composition or ingredient, as used herein, means having no persistent detrimental effect on the general health of the subject being treated.

As used herein, the term “administration” or “administering” refers to providing a fusion protein of the invention to a subject in need of treatment.

“Alterations,” as used herein, comprise any alteration in the sequence of polynucleotides encoding TAGNPP1 including deletions, insertions, and point mutations that may be detected using hybridization assays.

The term “animal” is used herein to include all vertebrate animals, including avians and mammals such as rat, mouse and human. It also includes an individual animal in all stages of development, including embryonic and fetal stages.

Where “amino acid sequence” is recited herein to refer to an amino acid sequence of a fusion protein molecule, “amino acid sequence” and like terms, such as “polypeptide” or “protein” are not meant to limit the amino acid sequence to the complete amino acid sequence associated with the recited protein or polypeptide molecule.

The term “avian” as used herein refers to any species, subspecies or strain of organism of the taxonomic class ava, such as, but not limited to, such organisms as chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary. The term includes the various known strains of Gallus gallus, or chickens, (for example, White Leghorn, Brown Leghorn, Barred-Rock, Sussex, New Hampshire, Rhode Island, Ausstralorp, Minorca, Amrox, California Gray, Italian Partridge-colored), as well as strains of turkeys, pheasants, quails, duck, ostriches and other poultry commonly bred in commercial quantities.

The phrase “based on” or “derived from” as in a retroviral vector being based on or derived from a particular retrovirus or based on a nucleotide sequence of a particular retrovirus mean that the genome of the retroviral vector contains a substantial portion of the nucleotide sequence of the genome of the particular retrovirus. The substantial portion may be a particular gene or nucleotide sequence such as the nucleotide sequence encoding the gag, pol and/or env proteins or other structural or functional nucleotide sequence of the virus genome such as sequences encoding the LTRs or may be substantially the complete retrovirus genome, for example, most (*e.g.*, more than 60% or more than 70% or more than 80% or more than 90%) or all of the retrovirus genome, as will be apparent from the context in the specification as the knowledge of one skilled in the art. Examples of retroviral vectors that are based on or derived from a retrovirus are the NL retroviral vectors (*e.g.*, NLB) which are based on the ALV retrovirus as disclosed in Cosset *et al.*, *Journal of Virology* (1991) vol 65, p 3388-3394.

The term “biologically active,” as used herein, refers to a fusion protein having structural, regulatory, or biochemical functions of pyrophosphatase/phosphodiesterase of a naturally occurring NPP1 protein.

The term “construct,” as used herein, refers to a linear or circular nucleotide sequence such as DNA that has been assembled from more than one segments of nucleotide sequence which have been isolated from a natural source or have been chemically synthesized, or combinations thereof.

The term “complementary,” as used herein, refers to two nucleic acid molecules that can form specific interactions with one another. In the specific interactions, an adenine base within one strand of a nucleic acid can form two hydrogen bonds with thymine within a second nucleic acid strand when the two nucleic acid strands are in opposing polarities. Also in the specific interactions, a guanine base within one strand of a nucleic acid can form three hydrogen bonds with cytosine within a second nucleic acid strand when the two nucleic acid strands are in opposing polarities. Complementary nucleic acids as referred to herein, may further comprise modified bases wherein a modified adenine may form hydrogen bonds with a thymine or modified thymine, and a modified cytosine may form hydrogen bonds with a guanine or a modified guanine.

A “deletion,” as used herein, refers to a change in either amino acid or nucleotide sequence in which one or more amino acid or nucleotide residues, respectively, are absent.

The term “expressed” or “expression” as used herein refers to the transcription from a gene to give an RNA nucleic acid molecule at least complementary in part to a region of one of the two nucleic acid strands of the gene. The term “expressed” or “expression” as used herein can also refer to the translation of RNA to produce a protein or peptide.

The term “expression vector” as used herein refers to a nucleic acid vector that comprises a gene expression controlling region, such as a promoter or promoter component, operably linked to a nucleotide sequence coding at least one polypeptide.

“Functional portion” or “functional fragment” are used interchangeably and as used herein means a portion or fragment of a whole capable of performing, in whole or in part, a function of the whole. For example, a biologically functional portion of a molecule means a portion of the molecule that performs a biological function of the whole or intact molecule. For example, a functional portion of a gene expression controlling region is a fragment or portion of the specified gene expression controlling region that, in whole or in part, regulates or controls gene expression (*e.g.*, facilitates either in whole or in part) in a biological system (*e.g.*, a promoter). Functional portions may be of any useful size.

The term “gene expression controlling region” as used herein refers to nucleotide sequences that are associated with a coding sequence and which regulate, in whole or in part, expression of the coding sequence, for example, regulate, in whole or in part, the transcription of the coding sequence. Gene expression controlling regions may be isolated from a naturally occurring source or may be chemically synthesized and can be incorporated into a nucleic acid vector to enable regulated transcription in appropriate cells. The “gene expression controlling regions” may precede, but is not limited to preceding, the region of a nucleic acid sequence that is in the region 5' of the end of a coding sequence that may be transcribed into mRNA.

The terms “heterologous,” “exogenous” and “foreign” are used interchangeably herein and in general refer to a biomolecule such as a nucleic acid or a protein that is not normally found in a certain organism or in a certain cell, tissue or other component contained in or produced by an organism. For example, a protein that is heterologous or exogenous to an egg is a protein that is not normally found in the egg. As used herein, the terms “heterologous,” “exogenous” and “foreign” with reference to nucleic acids, such as DNA and RNA, are used interchangeably and refer to nucleic acid that does not occur naturally as part of a chromosome, a genome or cell in which it is present or which is found in a location(s) and/or in amounts that differ from the location(s) and/or amounts in which it occurs in nature. It can be nucleic acid that is not endogenous to the genome, chromosome or cell and has been exogenously introduced into the genome, chromosome or cell. Examples of heterologous DNA include, but are not limited to, a DNA comprising a gene expression control region and DNA that encodes a product or products, for example, RNA or protein product. Examples of heterologous DNA include, but are not limited to, gene expression controlling regions or promoters disclosed herein once isolated from the avian and as used thereafter, *e.g.*, after re-introduction into an avian genome.

The term “isolated nucleic acid” as used herein covers, for example, (a) a DNA which has the sequence of part of a naturally occurring genomic molecule but is not flanked by at least one of the sequences that flank that part of the molecule in the genome of the species in which it naturally occurs; (b) a nucleic acid which has been incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting vector or genomic DNA is not identical to naturally occurring DNA from which the nucleic acid was obtained; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), ligase chain reaction (LCR) or chemical synthesis, or a restriction fragment; (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein, and (e) a recombinant nucleotide sequence that is part of a hybrid sequence that is not naturally occurring. Isolated nucleic acid molecules of the present invention can include, for example, natural allelic variants as well as nucleic acid molecules modified by nucleotide deletions, insertions, inversions, or substitutions.

An “insertion” or “addition,” as used herein, refers to a change in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid or nucleotide residues, respectively, as compared to the TAGNPP1 molecule.

The term "nucleic acid" as used herein refers to any linear or sequential array of nucleotides and nucleosides, for example cDNA, genomic DNA, mRNA, tRNA, oligonucleotides, oligonucleosides and derivatives thereof. For ease of discussion, non-naturally occurring nucleic acids may be referred to herein as constructs. Nucleic acids can include bacterial plasmid vectors including expression, cloning, cosmid and transformation vectors such as, animal viral vectors such as, but not limited to, modified adenovirus, influenza virus, polio virus, pox virus, retroviruses such as avian leukosis virus (ALV) retroviral vector, a murine leukemia virus (MLV) retroviral vector, and a lentivirus vector, and the like and fragments thereof. In addition, the nucleic acid can be an LTR of an avian leukosis virus (ALV) retroviral vector, a murine leukemia virus (MLV) retroviral vector, or a lentivirus vector and fragments thereof. Nuclic acids can also include NL vectors such as NLB, NLD and NLA and fragments thereof and synthetic oligonucleotides such as chemically synthesized DNA or RNA. Nucleic acids can include modified or derivatised nucleotides and nucleosides such as, but not limited to, halogenated nucleotides such as, but not only, 5-bromouracil, and derivatised nucleotides such as biotin-labeled nucleotides.

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand.

The term "operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Gene expression controlling regions or promoters (*e.g.*, promoter components) operably linked to a coding sequence are capable of effecting the expression of the coding sequence. The controlling sequences need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

The term "oviduct specific promoter" as used herein refers to promoters and promoter components which are functional, *i.e.*, provide for transcription of a coding sequence, to a large extent, for example, primarily (*i.e.*, more than 50% of the transcription product produced in the animal by a particular promoter type being produced in oviduct cells) or exclusively in oviduct cells of a bird. Examples of useful oviduct specific promoters include, without limitation, ovalbumin promoter, ovomucoid promoter, ovoinhibitor promoter, lysozyme promoter and ovotransferrin promoter and functional portions of these promoters, *e.g.*, promoter components.

The terms “polynucleotide,” “oligonucleotide,” “nucleotide sequence” and “nucleic acid sequence” can be used interchangeably herein and include, but are not limited to, coding sequences, *i.e.*, polynucleotide(s) or nucleic acid sequence(s) which are transcribed and translated into polypeptide in vitro or in vivo when placed under the control of appropriate regulatory or control sequences; controlling sequences, *e.g.*, translational start and stop codons, promoter sequences, ribosome binding sites, polyadenylation signals, transcription factor binding sites, transcription termination sequences, upstream and downstream regulatory domains, enhancers, silencers, DNA sequences to which a transcription factor(s) binds and alters the activity of a gene's promoter either positively (induction) or negatively (repression) and the like. No limitation as to length or to synthetic origin are suggested by the terms described herein.

As used herein, the terms “polypeptide” and “protein” can be used interchangealy and refer to a polymer of amino acids of three or more amino acids in a serial array, linked through peptide bonds. The term “polypeptide” includes proteins such as fusion proteins, protein fragments, protein analogues, oligopeptides and the like. The term “polypeptides” includes polypeptides as defined above that are encoded by nucleic acids, produced through recombinant technology (*e.g.*, isolated from a transgenic bird), or chemically synthesized.

As used herein, the term “promoter” refers to a DNA sequence useful to initiate transcription initiation by an RNA polymerase in an avian cell. A “promoter component” is a DNA sequence that can, by itself or, in combination with other DNA sequences effect or facilitate transcription. Specific promoter components such as ovalbumin promoter components, ovomucoid promoter components and lysozyme promoter components and other promoters and promoter components disclosed and claimed herein do not describe a specific promoter sequence. Rather, they encompass any sequence or sequence fragment of the respective promoter that is useful to effect or facilitate transcription of a coding sequence. For example, an ovomucoid promoter component includes, without limitation, the about 1.8 kb, the about 3.9 kb and the about 10 kb ovomucoid promoters disclosed in U.S. Publication No. 11/649,543, published May 17 2007, which is incorporated in its entirety herein by reference. "Promoter components" can also encompass rearranged gene expression controlling regions which function to initiate RNA transcription and hybrid DNA molecules composed of naturally occurring DNA sequences and/or synthetic DNA sequences which function to initiate RNA transcription.

The terms "recombinant nucleic acid" and "recombinant DNA" as used herein refer to combinations of at least two nucleic acid sequences that are not naturally found in a eukaryotic or prokaryotic cell. The nucleic acid sequences may include, but are not limited to, nucleic acid vectors, gene expression regulatory elements, origins of replication, suitable gene sequences that when expressed confer antibiotic resistance, protein-encoding sequences and the like. The term "recombinant polypeptide" or "recombinant protein" is meant to include a polypeptide produced by recombinant DNA techniques such that it is distinct from a naturally occurring polypeptide either in its location, purity or structure. Generally, such a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

The term "stringent conditions," as used herein, is the "stringency" which occurs within a range from about Tm - 5°C (5 °C below the melting temperature (Tm) of the probe) to about 20°C to 25°C below Tm. As will be understood by those of skill in the art, the stringency of hybridization may be altered in order to identify or detect identical or related polynucleotide sequences.

As used herein, the term "subject" or "patient" encompasses mammals and non-mammals. Examples of mammals include, but are not limited to, humans, chimpanzees, apes monkeys, cattle, horses, sheep, goats, swine; rabbits, dogs, cats, rats, mice, guinea pigs, and the like. Examples of non-mammals include, but are not limited to, birds, fish and the like.

A "substitution," as used herein, refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

As used herein, the term "therapeutically effective amount" refers to any amount of a compound which, as compared to a corresponding subject who has not received such amount, results in improved treatment, healing, prevention, or amelioration of a disease, disorder, or side effect, or a decrease in the rate of advancement of a disease or disorder. The term also includes within its scope amounts effective to enhance normal physiological function.

As used herein, the terms "TAGNPP1," "fusion protein," "TAGNPP1 polypeptide" and "NPP1 component fused to a targeting moiety" are used interchangeably.

As used herein, the term "treat," "treating" or "treatment" refers to methods of alleviating, abating or ameliorating a disease or condition symptoms, preventing additional symptoms, ameliorating or preventing the underlying causes of symptoms, inhibiting the disease or condition, arresting the development of the disease or condition, relieving the disease or condition, causing regression of the disease or condition, relieving a condition caused by the disease or condition, or stopping the symptoms of the disease or condition either prophylactically and/or therapeutically.

A "variant" of TAGNPP1, as used herein, refers to an amino acid sequence that is altered by one or more amino acids. Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include Asp and Glu; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include Leu, Ile and Val; Gly and Ala; Asp and Gln; and Ser, Thr, Phe and Tyr. Other groups of amino acids that may represent conservative changes include: Ala, Pro, Gly, Glu, Asp, Gln, Asn, Ser, Thr; (2) Cys, Ser, Tyr, Thr; (3) Val, Ile, Leu, Met, Ala, Phe; (4) Lys, Arg, His; and (5) Phe, Tyr, Trp, His. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer, or by, for example, replacement of a Gly with a Trp. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the NPP1 component. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity can be found using computer programs well known in the art.

The term “vector” and “nucleic acid vector” as used herein refers to a natural or synthetic single or double stranded plasmid or viral nucleic acid molecule that can be transfected or transformed into cells and replicate independently of, or within, the host cell genome. A circular double stranded vector can be linearized by treatment with an appropriate restriction enzyme based on the nucleotide sequence of the vector. A nucleic acid can be inserted into a vector by cutting the vector with restriction enzymes and ligating the desired pieces together.

The term “portion,” as used herein, with regard to a fusion protein refers to fragments of that protein. The fragments may range in size from four amino acid residues to the entire amino acid sequence minus one amino acid. Thus, a protein “comprising at least a portion of the amino acid sequence of SEQ ID NO:1” encompasses the full-length TAGNPP1 and fragments thereof.

“Transformation” or “transfection,” as used herein, describes a process by which exogenous DNA enters and changes a recipient cell using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the host cell being transformed and may include, but is not limited to, electroporation, particle bombardment, viral infection, and lipofection. Such “transformed” cells include stably transformed cells in which the inserted DNA is capable of replicating either as an autonomously replicating plasmid or as part of the host chromosome. They also include cells which transiently express the inserted DNA or RNA for limited periods of time.

EXAMPLES

The present invention is further exemplified by the following examples. The examples are for illustrative purpose only and are not intended, nor should they be construed as limiting the invention in any manner.

Example I

The TAGsNNP1 construct containing the targeting moiety having eight consecutive aspartic acids fused to sNPP1 was ligated into pTT22 vector using EcoRI and HindIII sites (pTT22-sNPP1.D8; Fig. 19). pTT22-sNPP1.D8 was transfected into HEK203E cells and the transformants were cultured to express TAGsNNP1. TAGsNNP1 was isolated from the culture media and partially purified as well known in the art. Following the purification, the pyrophosphase/phosphodiesterase activity of TAGsNPP1 was measured for its ability to hydrolyze thymmidine 5' monophosphate p-nitrophenyl ester. Briefly, TAGsNPP1 was diluted to 1 ng/μL in 50 mM Tris, 250 mM NaCl, pH 9.5. In a plate containing 50 μL of 1 ng/μL TAGsNPP1, 50 μL of 10 mM thymmidine 5' monophosphate p-nitrophenyl ester (Sigma™, Catalog #T4510) substrate was added. The enzyme activity of TAGsNPP1 was measured at 405 nm (absobance) in kinetic mode for 5 minutes. As shown in Fig. 21, the activity of TAGsNPP1 was detected above the level observed in control containing no TAGsNPP1. Particularly, TAGsNPP1 produced from HEK203D6 exhibited the highest level of the enzymatic activity. This results strongly suggested that the truncated NPP1 fused to a targeting moiety (*i.e.*, D8) sufficiently maintained its normal function as nuclease.

Example II

This non-limiting prophetic example describes how to treat idiopathic infantile arterial calcification by administering a formulation comprising a TAGNPP1 fusion protein.

A clinician uses a diagnostic test to verify that a patient has high levels of calcification in the artery. A genetic test can be also performed for NPP1 defects as described in Rutsch *et al.* (2003), *Nature Genetics* 34:379-81.

The pharmaceutical compositions of the present invention are preferably administered intravenously, although interadermal, intramuscular or oral administration are employed in certain circumstances.

The clinician determines a dose which may vary depending on the gender, age, health, and weight of the patient. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician.

The formulation containing TAGNPP1 can be infused, between about 10 mg/kg and about 1000 mg/kg per week weekly. 10-30 mg/kg can be administered once. During the infusion period, patients are monitored closely and appropriate clinical intervention is taken in the event of an adverse event. Treatment lasts at least 1 month or for the life of the patient. A window of 48 hours may be allowed for each infusion. An infusion schedule in which the rate of infusion increases with time reduces or eliminates adverse events. Infusions for infants can be administered according to the following schedule: 5-10 cc/hr for 60 minutes in each interval.

On the other hand, when continuous intravenous administration is desired, typical example of the slow release systems comprises that 1-100 mg/kg of effective TAGNPP1 proteins can be continuously released for more than 1 day.

While this invention has been particularly shown and described with references to example embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

CLAIMS

What is claimed is:

1. An isolated polypeptide comprising a targeting moiety and a C-terminus catalytic domain of NPP1.

2. The polypeptide of Claim 1, wherein said catalytic domain is a pyrophosphatase or phosphodiesterase domain of NPP1.

3. The polypeptide of Claim 1, wherein said polypeptide is a recombinant protein.

4. The polypeptide of Claim 1, wherein said targeting moiety is chemically linked to said catalytic domain of NPP1.

5. The polypeptide of Claim 1, wherein said catalytic domain of NPP1 comprises A205 through L591 of SEQ ID NO:1.

6. The polypeptide of Claim 5, wherein said catalytic domain of NPP1 comprises P99 through L591 of SEQ ID NO:1.

7. The polypeptide of Claim 5, wherein said catalytic domain of NPP1 comprises A205 through D925 of SEQ ID NO: 1.

8. The polypeptide of Claim 5, wherein said catalytic domain of NPP1 comprises P99 through D925 of SEQ ID NO: 1.

9. The polypeptide of Claim 8, wherein said polypeptide is a consecutive fragment of the polypeptide having the amino acid sequence of P99 through D925 of SEQ ID NO:1 and containing at least the amino acid resides A205 through L591 of SEQ ID NO:1.

10. The polypeptide of Claim 5, wherein said targeting moiety is a peptide comprising at least four negatively charged amino acid residues.

11. The polypeptide of Claim 5, wherein said targeting moiety is a peptide comprising between about five and about fifteen negatively charged amino acid residues.

12. The polypeptide of Claim 5, wherein said negatively charged amino acid residues comprise at least one aspartic acid or glutamic acid.

13. The polypeptide of Claim 5, wherein said negatively charged amino acid residues comprise at least four aspartic acid residues.

14. The polypeptide of Claim 5, wherein said negatively charged amino acid residues comprise at least four glutamic acid residues.

15. The polypeptide of Claim 5, wherein said negatively charged amino acid residues comprise eight consecutive aspartic acid residues.

16. The polypeptide of Claim 5, wherein said targeting moiety is fused to either N-terminus or C-terminus of said catalytic domain of NPP1.

17. The polypeptide of Claim 5, wherein said targeting moiety is fused to N-terminus and C-terminus of said catalytic domain of NPP1.

18. The polypeptide of Claim 12, wherein said polypeptide further comprises a polypeptide linker between said targeting moiety and said catalytic domain of NPP1.

19. The polypeptide of Claim 1, wherein said polypeptide further comprises an Fc region of an immunoglobulin.

20. The polypeptide of Claim 1, wherein said polypeptide further comprises a signal peptide.

21. The polypeptide of Claim 1, wherein said polypeptide forms a homodimer.

22. The polypeptide of Claim 1, wherein said polypeptide protein is a monomer.

23. An isolated nucleic acid encoding the polypeptide of Claim 1.

24. A replication or expression vector carrying the isolated nucleic acid of Claim 23.

25. A host cell transformed with the replication or expression vector according to Claim 23.

26. The host cell of Claim 25, wherein said host cell is selected from the group consisting of CHO cell, HEK293 cell, or COS cell.

27. The host cell of Claim 20, wherein said host cell is an avian tumor cell.

28. A transgenic animal producing the polypeptide of Claim 1.

29. The transgenic animal of Claim 28, wherein said transgenic animal is mammalian or avian.

30. A method producing said polypeptide of Claim 1 by an avian cell comprising culturing an avian cell transfected with at least one expression vector comprising a transcription unit having a nucleotide sequence encoding said polypeptide of Claim 1 operably linked to a promoter and a transcriptional terminator, and wherein the cultured avian cell produces said polypeptide.

31. A method for producing said polypeptide of Claim 1 in an egg of a chicken, the method comprising:

(a) providing an avian leukosis viral vector comprising a nucleic acid sequence encoding said polypeptide, and a promoter operably linked to said sequence, wherein said promoter drives expression of the encoding sequence in the chicken oviduct;
(b) introducing said vector into chicken stage X embryonic cells;
(c) incubating said embryonic cells under conditions conducive to hatching live chicks;

(d) nurturing growth of a mature chimeric chicken from said chicks;

(e) mating said chimeric chicken, either naturally or via artificial insemination, with a non-transgenic chicken;

(f) identifying a transgenic chicken by screening the progeny of step (e) for germ line incorporation of the protein encoding sequence; and

(g) mating the transgenic progeny with non-transgenic chickens to produce eggs containing the exogenous protein.

32. A method of treating a disorder in a subject in need thereof comprising administering an effective amount of a pharmaceutical composition comprising the polypeptide of Claim 1.

33. The method of claim 32, wherein said subject is a mammal.

34. The method of claim 33, wherein said mammal is a human.

35. The method of Claim 32, wherein said disorder is arterial calcification, insulin resistance, hypophosphatemic rickets, or ossification of posterior longitudinal ligament of spine.

36. The method of Claim 35, wherein said disorder is arterial calcification.

37. The method of Claim 36, wherein the treatment reduces calcification in the artery.

38. The method of Claim 36, wherein the arterial calcification is generalized arterial calcification of infancy.

NPP1 (wild-type)

MERDGCAGGGSRGGEGGRAPREGPAGNGRDRGRSHAAEAPGDPQAAASLLAPMDVGEEPLEKAARA
RTAKDPNTYKVLSLVLSVCVLTTILGCIFGLK**PSCAKE**VKSCKGRCFERTFGNCRCDAACVELGNCCLDYQET
CIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDKGDCCINYSSVCQGEKSWVEEPCESINEPQCP**AGFETP**
PTLLFSLDGFRAEYLHTWGGLLPVISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDP
KM**N**ASFSLKSKEKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMY**N**GSVPFEERILAV
LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLHRCLNLILISD
HGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEGIAR**N**LSCREPNQHFKPYLKHF
LPKRLHFAKSDRIEPLTFYLDPQWQLAL**N**PSERKYCGSGFHGSDNVFSNMQALFVGYGPGFKHGIEADTF
ENIEVYNLMCDLL**N**LTPAPN**N**GTHGSLNHLLKNPVYTPKHPKEVHPLVQCPFTRNPRDNLGCSC**N**PSILPI
EDFQTQFNLTVAEEKIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR**N**DSFSTE
DFSNCLYQDFRIPLSPVHKCSFYK**N**NTKVSYGFLSPPQLNK**N**SSGIYSEALLTTNIVPMYQSFQVIWRYFHD
TLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILIPTHFFIVLTSCKDTSQTPLHCENL
DTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHRARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTFSQ
ED (SEQ ID NO:1)

Fig. 1

sssNPP1 (NPP1 catalytic domain with no TAG)

AGFETPPTLLFSLDGFRAEYLHTWGGLLPVISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDN
KMYDPKM**N**ASFSLKSKEKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMY**N**GSVPF
EERILAVLQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLHRCL
NLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEGIAR**N**LSCREPNQHFK
PYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLAL**N**PSERKYCGSGFHGSDNVFSNMQALFVGYGPGFKHG
IEADTFENIEVYNLMCDLL**N**LTPAPN**N**GTHGSL (SEQ ID NO:2)

Fig. 2

TAGsssNPP1 (N-terminus D8)

IGVLLTQRTLLSLVLALLFPSMASM**DDDDDDDD**AGFETPPTLLFSLDGFRAEYLHTWGGLLPVISKLKKCG
TYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSKEKFNPEWYKGEPIWVTAKY
QGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAVLQWLQLPKDERPHFYTLYLEEPDSSGHSYG
PVSSEVIKALQRVDGMVGMLMDGLKELNLHRCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPA
ARLRPSDVPDKYYSFNYEGIARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSE
RKYCGSGFHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSL (SEQ
ID NO:3)

Fig. 3

TAGsssNPP1 (C-terminus D10)

IGVLLTQRTLLSLVLALLFPSMASMAGFETPPTLLFSLDGFRAEYLHTWGGLLPVISKLKKCGTYTKNMRPV
YPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSKEKFNPEWYKGEPIWVTAKYQGLKSGTFF
WPGSDVEINGIFPDIYKMYNGSVPFEERILAVLQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKAL
QRVDGMVGMLMDGLKELNLHRCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVP
DKYYSFNYEGIARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSGFH
GSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSL**DDDDDDDDDD**
(SEQ ID NO:4)

Fig. 4

TAGssNPP1 (N-terminus D8)

ATGGGTGTACTGCTCACACAGAGGACGCTGCTCAGTCTGGTCCTTGCACTCCTGTTTCCAAGCATGGCGAGCA
TGGATGACGATGATGACGACGATGACGCAGGGTTTGAAACGCCTCCTACACTCTTGTTTTCTTTGGATGGATT
CAGGGCAGAATATTTGCACACTTGGGGTGGACTTCTTCCTGTTATTAGCAAACTCAAAAAATGTGGAACATAT
ACTAAAAACATGAGACCGGTGTATCCAACAAAAACTTTCCCCAATCACTACAGCATTGTCACCGGATTGTATCC
AGAATCTCATGGCATAATCGACAATAAGATGTATGATCCCAAAATGAATGCTTCCTTTTCACTTAAAAGTAAAG
AGAAATTTAATCCGGAGTGGTACAAAGGAGAACCAATTTGGGTCACAGCTAAGTATCAAGGCCTCAAGTCTG
GCACATTTTTCTGGCCAGGATCAGATGTGGAAATTAACGGAATTTTCCCAGACATCTATAAAATGTATAATGGT
TCAGTGCCATTTGAAGAAAGGATTTTGGCTGTTCTTCAGTGGCTGCAGCTTCCAAAAGATGAAAGACCACACT
TTTACACTTTGTATTTGGAAGAACCAGATTCTTCAGGTCATTCATATGGACCAGTCAGCAGTGAAGTCATCAAA
GCCTTGCAGAGGGTTGATGGTATGGTTGGTATGCTGATGGATGGTCTGAAAGAGCTGAACTTGCACAGATGC
CTGAACCTCATCCTTATTTCAGATCATGGCATGGAACAAGGCAGTTGTAAGAAATACATATATCTGAATAAGTA
TTTGGGGGATGTTAAAAATATTAAAGTTATCTATGGACCTGCAGCTCGATTGAGACCCTCTGATGTCCCAGATA
AATACTATTCATTTAACTATGAAGGCATTGCCCGAAATCTTTCTTGCCGGGAACCAAACCAGCACTTCAAACCT
TATCTGAAACATTTCTTGCCTAAGCGTTTGCACTTTGCTAAGAGTGATAGAATTGAGCCCTTGACATTCTATTTG
GACCCTCAGTGGCAACTTGCATTGAATCCCTCAGAAAGGAAATATTGTGGAAGTGGATTTCATGGCTCTGACA
ATGTGTTTTCAAATATGCAAGCCCTCTTTGTTGGCTATGGACCTGGATTCAAGCATGGCATTGAGGCTGACACC
TTTGAAAACATTGAAGTCTATAACTTGATGTGTGATTTGCTGAATTTGACACCGGCTCCTAATAACGGAACTCA
TGGAAGTCTTAACCACCTTCTGAAGAATCCTGTTTATACGCCAAAGCATCCCAAAGAAGTGCACCCCCTGGTGC
AGTGCCCCTTCACAAGAAACCCCAGAGATAACCTTGGCTGCTCATGTAACCCTTCCATTTTGCCGATTGAGGAT
TTTCAAACACAGTTCAATCTGACCGTGGCAGAAGAGAAGATTATTAAGCATGAAACTTTGCCCTATGGAAGAC
CTAGAGTTCTCCAGAAGGAAAACACCATCTGTCTTCTTTCCCAGCACCAGTTTATGAGTGGATACAGCCAAGAC
ATCTTGATGCCCCTTTGGACATCCTATACCGTGGACAGAAATGACAGTTTCTCTACGGAAGACTTCTCCAACTG
TCTGTACCAGGACTTTAGAATTCCTCTTAGTCCTGTCCATAAATGTTCATTTTATAAAAATAACACCAAAGTGAG
TTACGGGTTCCTCTCCCCACCACAACTGAATAAGAATTCAAGTGGAATATATTCTGAAGCCTTGCTTACTACAA
ATATAGTGCCAATGTACCAGAGTTTTCAAGTTATATGGCGCTACTTTCATGACACCCTCTTGCGAAAGTATGCA
GAAGAAAGAAATGGTGTCAATGTCGTCAGTGGTCCTGTGTTTGACTTTGATTATGATGGACGTTGTGATTCCT
TGGAGAATTTGAGGCAAAAAAGAAGAGTCATCCGTAACCAAGAAATTTTGATTCCAACTCATTTCTTCATTGTG
CTGACAAGCTGTAAAGATACATCTCAGACGCCTTTGCACTGTGAAAACCTGGACACCTTGGCTTTCATTTTGCC
TCACAGGACTGATAACAGCGAGAGCTGTGTGCATGGGAAGCATGACTCCTCATGGGTTGAAGAATTGTTGAT
GTTGCACAGAGCACGGATCACAGACGTCGAGCACATCACTGGACTCAGCTTTTATCAACAAAGAAAAGAGCC
AGTTTCAGACATTTTGAAGTTGAAAACACATTTGCCAACCTTTAGCCAAGAAGATTGA (SEQ ID NO:5)

Fig. 5

TAGssNPP1 (N-terminus D8; signal peptide sequence underlined and the targeting moiety in bold)

IGVLLTQRTLLSLVLALLFPSMASM**DDDDDDDD**AGFETPPTLLFSLDGFRAEYLHTWGGLLPVISKLKKCG
TYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSKEKFNPEWYKGEPIWVTAKY
QGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAVLQWLQLPKDERPHFYTLYLEEPDSSGHSYG
PVSSEVIKALQRVDGMVGMLMDGLKELNLHRCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGP
AARLRPSDVPDKYYSFNYEGIARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNP
SERKYCGSGFHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLN
HLLKNPVYTPKHPKEVHPLVQCPFTRNPRDNLGCSCNPSILPIEDFQTQFNLTVAEEKIIKHETLPYGRPRV
LQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDRNDSFSTEDFSNCLYQDFRIPLSPVHKCSFYKNNTK
VSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQSFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYD
GRCDSLENLRQKRRVIRNQEILIPTHFFIVLTSCKDTSQTPLHCENLDTLAFILPHRTDNSESCVHGKHDSS
WVEELLMLHRARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTFSQED (SEQ ID NO:6)

Fig. 6

ssNPP1 (2238bp)

ATGGGTGTACTGCTCACACAGAGGACGCTGCTCAGTCTGGTCCTTGCACTCCTGTTTCCAAGCATGGCGAGCA
TGGCAGGGTTTGAAACGCCTCCTACACTCTTGTTTTCTTTGGATGGATTCAGGGCAGAATATTTGCACACTTGG
GGTGGACTTCTTCCTGTTATTAGCAAACTCAAAAAATGTGGAACATATACTAAAAACATGAGACCGGTGTATC
CAACAAAAACTTTCCCCAATCACTACAGCATTGTCACCGGATTGTATCCAGAATCTCATGGCATAATCGACAAT
AAGATGTATGATCCCAAAATGAATGCTTCCTTTTCACTTAAAAGTAAAGAGAAATTTAATCCGGAGTGGTACA
AAGGAGAACCAATTTGGGTCACAGCTAAGTATCAAGGCCTCAAGTCTGGCACATTTTTCTGGCCAGGATCAGA
TGTGGAAATTAACGGAATTTTCCCAGACATCTATAAAATGTATAATGGTTCAGTGCCATTTGAAGAAAGGATTT
TGGCTGTTCTTCAGTGGCTGCAGCTTCCAAAAGATGAAAGACCACACTTTTACACTTTGTATTTGGAAGAACCA
GATTCTTCAGGTCATTCATATGGACCAGTCAGCAGTGAAGTCATCAAAGCCTTGCAGAGGGTTGATGGTATGG
TTGGTATGCTGATGGATGGTCTGAAAGAGCTGAACTTGCACAGATGCCTGAACCTCATCCTTATTTCAGATCAT
GGCATGGAACAAGGCAGTTGTAAGAAATACATATATCTGAATAAGTATTTGGGGGATGTTAAAAATATTAAA
GTTATCTATGGACCTGCAGCTCGATTGAGACCCTCTGATGTCCCAGATAAATACTATTCATTTAACTATGAAGG
CATTGCCCGAAATCTTTCTTGCCGGGAACCAAACCAGCACTTCAAACCTTATCTGAAACATTTCTTGCCTAAGC
GTTTGCACTTTGCTAAGAGTGATAGAATTGAGCCCTTGACATTCTATTTGGACCCTCAGTGGCAACTTGCATTG
AATCCCTCAGAAAGGAAATATTGTGGAAGTGGATTTCATGGCTCTGACAATGTGTTTTCAAATATGCAAGCCC
TCTTTGTTGGCTATGGACCTGGATTCAAGCATGGCATTGAGGCTGACACCTTTGAAAACATTGAAGTCTATAAC
TTGATGTGTGATTTGCTGAATTTGACACCGGCTCCTAATAACGGAACTCATGGAAGTCTTAACCACCTTCTGAA
GAATCCTGTTTATACGCCAAAGCATCCCAAAGAAGTGCACCCCCTGGTGCAGTGCCCCTTCACAAGAAACCCC
AGAGATAACCTTGGCTGCTCATGTAACCCTTCCATTTTGCCGATTGAGGATTTTCAAACACAGTTCAATCTGAC
CGTGGCAGAAGAGAAGATTATTAAGCATGAAACTTTGCCCTATGGAAGACCTAGAGTTCTCCAGAAGGAAAA
CACCATCTGTCTTCTTTCCCAGCACCAGTTTATGAGTGGATACAGCCAAGACATCTTGATGCCCCTTTGGACATC
CTATACCGTGGACAGAAATGACAGTTTCTCTACGGAAGACTTCTCCAACTGTCTGTACCAGGACTTTAGAATTC
CTCTTAGTCCTGTCCATAAATGTTCATTTTATAAAAATAACACCAAAGTGAGTTACGGGTTCCTCTCCCCACCAC
AACTGAATAAGAATTCAAGTGGAATATATTCTGAAGCCTTGCTTACTACAAATATAGTGCCAATGTACCAGAG
TTTTCAAGTTATATGGCGCTACTTTCATGACACCCTCTTGCGAAAGTATGCAGAAGAAAGAAATGGTGTCAAT
GTCGTCAGTGGTCCTGTGTTTGACTTTGATTATGATGGACGTTGTGATTCCTTGGAGAATTTGAGGCAAAAAA
GAAGAGTCATCCGTAACCAAGAAATTTTGATTCCAACTCATTTCTTCATTGTGCTGACAAGCTGTAAAGATACA
TCTCAGACGCCTTTGCACTGTGAAAACCTGGACACCTTGGCTTTCATTTTGCCTCACAGGACTGATAACAGCGA
GAGCTGTGTGCATGGGAAGCATGACTCCTCATGGGTTGAAGAATTGTTGATGTTGCACAGAGCACGGATCAC
AGACGTCGAGCACATCACTGGACTCAGCTTTTATCAACAAAGAAAAGAGCCAGTTTCAGACATTTTGAAGTTG
AAAACACATTTGCCAACCTTTAGCCAAGAAGAT (SEQ ID NO:7)

Fig. 7

ssNPP1 (signal peptide sequence underlined)

IGVLLTQRTLLSLVLALLFPSMASMAGFETPPTLLFSLDGFRAEYLHTWGGLLPVISKLKKCGTYTKNMRPV
YPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSKEKFNPEWYKGEPIWVTAKYQGLKSGTFF
WPGSDVEINGIFPDIYKMYNGSVPFEERILAVLQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKAL
QRVDGMVGMLMDGLKELNLHRCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDV
PDKYYSFNYEGIARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG
FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNHLLKNPVYTP
KHPKEVHPLVQCPFTRNPRDNLGCSCNPSILPIEDFQTQFNLTVAEEKIIKHETLPYGRPRVLQKENTICLLS
QHQFMSGYSQDILMPLWTSYTVDRNDSFSTEDFSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQL
NKNSSGIYSEALLTTNIVPMYQSFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQ
KRRVIRNQEILIPTHFFIVLTSCKDTSQTPLHCENLDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHRA
RITDVEHITGLSFYQQRKEPVSDILKLKTHLPTFSQED (SEQ ID NO:8)

Fig. 8

sNPP1

ATGGGTGTACTGCTCACACAGAGGACGCTGCTCAGTCTGGTCCTTGCACTCCTGTTTCCAAGCATGGCGAGCA
TGCCAAGTTGTGCCAAAGAAGTTAAAAGTTGCAAAGGTCGCTGTTTCGAGAGAACATTTGGGAACTGTCGCT
GTGATGCTGCCTGTGTTGAGCTTGGAAACTGCTGTTTGGATTACCAGGAGACGTGCATAGAACCAGAACATAT
ATGGACTTGCAACAAATTCAGGTGTGGTGAGAAAAGATTGACCAGAAGCCTCTGTGCCTGTTCAGATGATTGC
AAGGACAAGGGCGACTGCTGCATCAACTACAGTTCAGTGTGTCAAGGTGAGAAAAGTTGGGTGGAAGAACC
ATGTGAGAGCATTAATGAGCCACAGTGCCCAGCAGGGTTTGAAACGCCTCCTACACTCTTGTTTTCTTTGGATG
GATTCAGGGCAGAATATTTGCACACTTGGGGTGGACTTCTTCCTGTTATTAGCAAACTCAAAAAATGTGGAAC
ATATACTAAAAACATGAGACCGGTGTATCCAACAAAAACTTTCCCCAATCACTACAGCATTGTCACCGGATTGT
ATCCAGAATCTCATGGCATAATCGACAATAAGATGTATGATCCCAAAATGAATGCTTCCTTTTCACTTAAAAGT
AAAGAGAAATTTAATCCGGAGTGGTACAAAGGAGAACCAATTTGGGTCACAGCTAAGTATCAAGGCCTCAAG
TCTGGCACATTTTTCTGGCCAGGATCAGATGTGGAAATTAACGGAATTTTCCCAGACATCTATAAAATGTATAA
TGGTTCAGTGCCATTTGAAGAAAGGATTTTGGCTGTTCTTCAGTGGCTGCAGCTTCCAAAAGATGAAAGACCA
CACTTTTACACTTTGTATTTGGAAGAACCAGATTCTTCAGGTCATTCATATGGACCAGTCAGCAGTGAAGTCAT
CAAAGCCTTGCAGAGGGTTGATGGTATGGTTGGTATGCTGATGGATGGTCTGAAAGAGCTGAACTTGCACAG
ATGCCTGAACCTCATCCTTATTTCAGATCATGGCATGGAACAAGGCAGTTGTAAGAAATACATATATCTGAATA
AGTATTTGGGGGATGTTAAAAATATTAAAGTTATCTATGGACCTGCAGCTCGATTGAGACCCTCTGATGTCCCA
GATAAATACTATTCATTTAACTATGAAGGCATTGCCCGAAATCTTTCTTGCCGGGAACCAAACCAGCACTTCAA
ACCTTATCTGAAACATTTCTTGCCTAAGCGTTTGCACTTTGCTAAGAGTGATAGAATTGAGCCCTTGACATTCTA
TTTGGACCCTCAGTGGCAACTTGCATTGAATCCCTCAGAAAGGAAATATTGTGGAAGTGGATTTCATGGCTCT
GACAATGTGTTTTCAAATATGCAAGCCCTCTTTGTTGGCTATGGACCTGGATTCAAGCATGGCATTGAGGCTG
ACACCTTTGAAAACATTGAAGTCTATAACTTGATGTGTGATTTGCTGAATTTGACACCGGCTCCTAATAACGGA
ACTCATGGAAGTCTTAACCACCTTCTGAAGAATCCTGTTTATACGCCAAAGCATCCCAAAGAAGTGCACCCCCT
GGTGCAGTGCCCCTTCACAAGAAACCCCAGAGATAACCTTGGCTGCTCATGTAACCCTTCCATTTTGCCGATTG
AGGATTTTCAAACACAGTTCAATCTGACCGTGGCAGAAGAGAAGATTATTAAGCATGAAACTTTGCCCTATGG
AAGACCTAGAGTTCTCCAGAAGGAAAACACCATCTGTCTTCTTTCCCAGCACCAGTTTATGAGTGGATACAGCC
AAGACATCTTGATGCCCCTTTGGACATCCTATACCGTGGACAGAAATGACAGTTTCTCTACGGAAGACTTCTCC
AACTGTCTGTACCAGGACTTTAGAATTCCTCTTAGTCCTGTCCATAAATGTTCATTTTATAAAAATAACACCAAA
GTGAGTTACGGGTTCCTCTCCCCACCACAACTGAATAAGAATTCAAGTGGAATATATTCTGAAGCCTTGCTTAC
TACAAATATAGTGCCAATGTACCAGAGTTTTCAAGTTATATGGCGCTACTTTCATGACACCCTCTTGCGAAAGT
ATGCAGAAGAAAGAAATGGTGTCAATGTCGTCAGTGGTCCTGTGTTTGACTTTGATTATGATGGACGTTGTGA
TTCCTTGGAGAATTTGAGGCAAAAAAGAAGAGTCATCCGTAACCAAGAAATTTTGATTCCAACTCATTTCTTCA
TTGTGCTGACAAGCTGTAAAGATACATCTCAGACGCCTTTGCACTGTGAAAACCTGGACACCTTGGCTTTCATT
TTGCCTCACAGGACTGATAACAGCGAGAGCTGTGTGCATGGGAAGCATGACTCCTCATGGGTTGAAGAATTG
TTGATGTTGCACAGAGCACGGATCACAGACGTCGAGCACATCACTGGACTCAGCTTTTATCAACAAAGAAAAG
AGCCAGTTTCAGACATTTTGAAGTTGAAAACACATTTGCCAACCTTTAGCCAAGAAGAT (SEQ ID NO:9)

Fig. 9

sNPP1

IGVLLTQRTLLSLVLALLFPSMASMPSCAKEVKSCKGRCFERTFGNCRCDAACVELGNCCLDYQETCIEPE
HIWTCNKFRCGEKRLTRSLCACSDDCKDKGDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLF
SLDGFRAEYLHTWGGLLPVISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMN
ASFSLKSKEKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAVLQ
WLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLHRCLNLILISDHG
MEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEGIARNLSCREPNQHFKPYLKHFLPK
RLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSGFHGSDNVFSNMQALFVGYGPGFKHGIEADTFENI
EVYNLMCDLLNLTPAPNNGTHGSLNHLLKNPVYTPKHPKEVHPLVQCPFTRNPRDNLGCSCNPSILPIED
FQTQFNLTVAEEKIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDRNDSFSTEDF
SNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQSFQVIWRYFHDTL
LRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILIPTHFFIVLTSCKDTSQTPLHCENLD
TLAFILPHRTDNSESCVHGKHDSSWVEELLMLHRARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTFSQE
D (SEQ ID NO:10)

Fig. 10

TAGsNPP1 (N-terminus D8)

ATGGGTGTACTGCTCACACAGAGGACGCTGCTCAGTCTGGTCCTTGCACTCCTGTTTCCAAGCATGGCGAGCA
TGGATGACGATGATGACGACGATGACCCAAGTTGTGCCAAAGAAGTTAAAAGTTGCAAAGGTCGCTGTTTCG
AGAGAACATTTGGGAACTGTCGCTGTGATGCTGCCTGTGTTGAGCTTGGAAACTGCTGTTTGGATTACCAGGA
GACGTGCATAGAACCAGAACATATATGGACTTGCAACAAATTCAGGTGTGGTGAGAAAAGATTGACCAGAAG
CCTCTGTGCCTGTTCAGATGATTGCAAGGACAAGGGCGACTGCTGCATCAACTACAGTTCAGTGTGTCAAGGT
GAGAAAAGTTGGGTGGAAGAACCATGTGAGAGCATTAATGAGCCACAGTGCCCAGCAGGGTTTGAAACGCC
TCCTACACTCTTGTTTTCTTTGGATGGATTCAGGGCAGAATATTTGCACACTTGGGGTGGACTTCTTCCTGTTAT
TAGCAAACTCAAAAAATGTGGAACATATACTAAAAACATGAGACCGGTGTATCCAACAAAAACTTTCCCCAAT
CACTACAGCATTGTCACCGGATTGTATCCAGAATCTCATGGCATAATCGACAATAAGATGTATGATCCCAAAAT
GAATGCTTCCTTTTCACTTAAAAGTAAAGAGAAATTTAATCCGGAGTGGTACAAAGGAGAACCAATTTGGGTC
ACAGCTAAGTATCAAGGCCTCAAGTCTGGCACATTTTTCTGGCCAGGATCAGATGTGGAAATTAACGGAATTT
TCCCAGACATCTATAAAATGTATAATGGTTCAGTGCCATTTGAAGAAAGGATTTTGGCTGTTCTTCAGTGGCTG
CAGCTTCCAAAAGATGAAAGACCACACTTTTACACTTTGTATTTGGAAGAACCAGATTCTTCAGGTCATTCATA
TGGACCAGTCAGCAGTGAAGTCATCAAAGCCTTGCAGAGGGTTGATGGTATGGTTGGTATGCTGATGGATGG
TCTGAAAGAGCTGAACTTGCACAGATGCCTGAACCTCATCCTTATTTCAGATCATGGCATGGAACAAGGCAGT
TGTAAGAAATACATATATCTGAATAAGTATTTGGGGGATGTTAAAAATATTAAAGTTATCTATGGACCTGCAG
CTCGATTGAGACCCTCTGATGTCCCAGATAAATACTATTCATTTAACTATGAAGGCATTGCCCGAAATCTTTCTT
GCCGGGAACCAAACCAGCACTTCAAACCTTATCTGAAACATTTCTTGCCTAAGCGTTTGCACTTTGCTAAGAGT
GATAGAATTGAGCCCTTGACATTCTATTTGGACCCTCAGTGGCAACTTGCATTGAATCCCTCAGAAAGGAAAT
ATTGTGGAAGTGGATTTCATGGCTCTGACAATGTGTTTTCAAATATGCAAGCCCTCTTTGTTGGCTATGGACCT
GGATTCAAGCATGGCATTGAGGCTGACACCTTTGAAAACATTGAAGTCTATAACTTGATGTGTGATTTGCTGA
ATTTGACACCGGCTCCTAATAACGGAACTCATGGAAGTCTTAACCACCTTCTGAAGAATCCTGTTTATACGCCA
AAGCATCCCAAAGAAGTGCACCCCCTGGTGCAGTGCCCCTTCACAAGAAACCCCAGAGATAACCTTGGCTGCT
CATGTAACCCTTCCATTTTGCCGATTGAGGATTTTCAAACACAGTTCAATCTGACCGTGGCAGAAGAGAAGATT
ATTAAGCATGAAACTTTGCCCTATGGAAGACCTAGAGTTCTCCAGAAGGAAAACACCATCTGTCTTCTTTCCCA
GCACCAGTTTATGAGTGGATACAGCCAAGACATCTTGATGCCCCTTTGGACATCCTATACCGTGGACAGAAAT
GACAGTTTCTCTACGGAAGACTTCTCCAACTGTCTGTACCAGGACTTTAGAATTCCTCTTAGTCCTGTCCATAAA
TGTTCATTTTATAAAAATAACACCAAAGTGAGTTACGGGTTCCTCTCCCCACCACAACTGAATAAGAATTCAAG
TGGAATATATTCTGAAGCCTTGCTTACTACAAATATAGTGCCAATGTACCAGAGTTTTCAAGTTATATGGCGCT
ACTTTCATGACACCCTCTTGCGAAAGTATGCAGAAGAAAGAAATGGTGTCAATGTCGTCAGTGGTCCTGTGTT
TGACTTTGATTATGATGGACGTTGTGATTCCTTGGAGAATTTGAGGCAAAAAAGAAGAGTCATCCGTAACCAA
GAAATTTTGATTCCAACTCATTTCTTCATTGTGCTGACAAGCTGTAAAGATACATCTCAGACGCCTTTGCACTGT
GAAAACCTGGACACCTTGGCTTTCATTTTGCCTCACAGGACTGATAACAGCGAGAGCTGTGTGCATGGGAAGC
ATGACTCCTCATGGGTTGAAGAATTGTTGATGTTGCACAGAGCACGGATCACAGACGTCGAGCACATCACTGG
ACTCAGCTTTTATCAACAAAGAAAAGAGCCAGTTTCAGACATTTTGAAGTTGAAAACACATTTGCCAACCTTTA
GCCAAGAAGAT (SEQ ID NO:11)

Fig. 11

TAGsNPP1 (N-terminus D8)

IGVLLTQRTLLSLVLALLFPSMASM**DDDDDDDD**PSCAKEVKSCKGRCFERTFGNCRCDAACVELGNCCL
DYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDKGDCCINYSSVCQGEKSWVEEPCESINEPQCP
AGFETPPTLLFSLDGFRAEYLHTWGGLLPVISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIID
NKMYDPKMNASFSLKSKEKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSV
PFEERILAVLQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH
RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEGIARNLSCREPNQ
HFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSGFHGSDNVFSNMQALFVGYGPGF
KHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNHLLKNPVYTPKHPKEVHPLVQCPFTRNPRDNL
GCSCNPSILPIEDFQTQFNLTVAEEKIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYT
VDRNDSFSTEDFSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQS
FQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILIPTHFFIVLTSCK
DTSQTPLHCENLDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHRARITDVEHITGLSFYQQRKEPVSD
ILKLKTHLPTFSQED (SEQ ID NO:12)

Fig. 12

TAGsNPP1 (C-terminus D8)

GTGGGTGTACTGCTCACACAGAGGACGCTGCTCAGTCTGGTCCTTGCACTCCTGTTTCCAAGCATGGCGAGCA
TGCCAAGTTGTGCCAAAGAAGTTAAAAGTTGCAAAGGTCGCTGTTTCGAGAGAACATTTGGGAACTGTCGCT
GTGATGCTGCCTGTGTTGAGCTTGGAAACTGCTGTTTGGATTACCAGGAGACGTGCATAGAACCAGAACATAT
ATGGACTTGCAACAAATTCAGGTGTGGTGAGAAAAGATTGACCAGAAGCCTCTGTGCCTGTTCAGATGATTGC
AAGGACAAGGGCGACTGCTGCATCAACTACAGTTCAGTGTGTCAAGGTGAGAAAAGTTGGGTGGAAGAACC
ATGTGAGAGCATTAATGAGCCACAGTGCCCAGCAGGGTTTGAAACGCCTCCTACACTCTTGTTTTCTTTGGATG
GATTCAGGGCAGAATATTTGCACACTTGGGGTGGACTTCTTCCTGTTATTAGCAAACTCAAAAAATGTGGAAC
ATATACTAAAAACATGAGACCGGTGTATCCAACAAAAACTTTCCCCAATCACTACAGCATTGTCACCGGATTGT
ATCCAGAATCTCATGGCATAATCGACAATAAGATGTATGATCCCAAAATGAATGCTTCCTTTTCACTTAAAAGT
AAAGAGAAATTTAATCCGGAGTGGTACAAAGGAGAACCAATTTGGGTCACAGCTAAGTATCAAGGCCTCAAG
TCTGGCACATTTTTCTGGCCAGGATCAGATGTGGAAATTAACGGAATTTTCCCAGACATCTATAAAATGTATAA
TGGTTCAGTGCCATTTGAAGAAAGGATTTTGGCTGTTCTTCAGTGGCTGCAGCTTCCAAAAGATGAAAGACCA
CACTTTTACACTTTGTATTTGGAAGAACCAGATTCTTCAGGTCATTCATATGGACCAGTCAGCAGTGAAGTCAT
CAAAGCCTTGCAGAGGGTTGATGGTATGGTTGGTATGCTGATGGATGGTCTGAAAGAGCTGAACTTGCACAG
ATGCCTGAACCTCATCCTTATTTCAGATCATGGCATGGAACAAGGCAGTTGTAAGAAATACATATATCTGAATA
AGTATTTGGGGGATGTTAAAAATATTAAAGTTATCTATGGACCTGCAGCTCGATTGAGACCCTCTGATGTCCCA
GATAAATACTATTCATTTAACTATGAAGGCATTGCCCGAAATCTTTCTTGCCGGGAACCAAACCAGCACTTCAA
ACCTTATCTGAAACATTTCTTGCCTAAGCGTTTGCACTTTGCTAAGAGTGATAGAATTGAGCCCTTGACATTCTA
TTTGGACCCTCAGTGGCAACTTGCATTGAATCCCTCAGAAAGGAAATATTGTGGAAGTGGATTTCATGGCTCT
GACAATGTGTTTTCAAATATGCAAGCCCTCTTTGTTGGCTATGGACCTGGATTCAAGCATGGCATTGAGGCTG
ACACCTTTGAAAACATTGAAGTCTATAACTTGATGTGTGATTTGCTGAATTTGACACCGGCTCCTAATAACGGA
ACTCATGGAAGTCTTAACCACCTTCTGAAGAATCCTGTTTATACGCCAAAGCATCCCAAAGAAGTGCACCCCCT
GGTGCAGTGCCCCTTCACAAGAAACCCCAGAGATAACCTTGGCTGCTCATGTAACCCTTCCATTTTGCCGATTG
AGGATTTTCAAACACAGTTCAATCTGACCGTGGCAGAAGAGAAGATTATTAAGCATGAAACTTTGCCCTATGG
AAGACCTAGAGTTCTCCAGAAGGAAAACACCATCTGTCTTCTTTCCCAGCACCAGTTTATGAGTGGATACAGCC
AAGACATCTTGATGCCCCTTTGGACATCCTATACCGTGGACAGAAATGACAGTTTCTCTACGGAAGACTTCTCC
AACTGTCTGTACCAGGACTTTAGAATTCCTCTTAGTCCTGTCCATAAATGTTCATTTTATAAAAATAACACCAAA
GTGAGTTACGGGTTCCTCTCCCCACCACAACTGAATAAGAATTCAAGTGGAATATATTCTGAAGCCTTGCTTAC
TACAAATATAGTGCCAATGTACCAGAGTTTTCAAGTTATATGGCGCTACTTTCATGACACCCTCTTGCGAAAGT
ATGCAGAAGAAAGAAATGGTGTCAATGTCGTCAGTGGTCCTGTGTTTGACTTTGATTATGATGGACGTTGTGA
TTCCTTGGAGAATTTGAGGCAAAAAAGAAGAGTCATCCGTAACCAAGAAATTTTGATTCCAACTCATTTCTTCA
TTGTGCTGACAAGCTGTAAAGATACATCTCAGACGCCTTTGCACTGTGAAAACCTGGACACCTTGGCTTTCATT
TTGCCTCACAGGACTGATAACAGCGAGAGCTGTGTGCATGGGAAGCATGACTCCTCATGGGTTGAAGAATTG
TTGATGTTGCACAGAGCACGGATCACAGACGTCGAGCACATCACTGGACTCAGCTTTTATCAACAAAGAAAAG
AGCCAGTTTCAGACATTTTGAAGTTGAAAACACATTTGCCAACCTTTAGCCAAGAAGATGATGACGATGATGA
CGACGATTGA (SEQ ID NO:13)

Fig. 13

TAGsNPP1 (C-terminus D8; signal peptide sequence underlined and the targeting moiety in bold)

QGVLLTQRTLLSLVLALLFPSMASMPSCAKEVKSCKGRCFERTFGNCRCDAACVELGNCCLDYQETCIEPEHIWTCN
KFRCGEKRLTRSLCACSDDCKDKGDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHT
WGGLLPVISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSKEKFNPEWYKG
EPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAVLQWLQLPKDERPHFYTLYLEEPDSSGH
SYGPVSSEVIKALQRVDGMVGMLMDGLKELNLHRCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARL
RPSDVPDKYYSFNYEGIARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSGFH
GSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNHLLKNPVYTPKHPKEVHP
LVQCPFTRNPRDNLGCSCNPSILPIEDFQTQFNLTVAEEKIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDIL
MPLWTSYTVDRNDSFSTEDFSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMY
QSFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILIPTHFFIVLTSCKDTSQT
PLHCENLDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHRARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTFS
QE**DDDDDDDD** (SEQ ID NO:14)

Fig. 14

Linker peptide

GGGGSGGGGS (SEQ ID NO:15)

Fig. 15

The amino acid sequence of the Fc region

EPKSCDKTHTCPPCPAPEAAGAPSVFLFPPKPKDTLMISRTPEVTCVVVDVSH EDPEVKFNWY
VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISK
AKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS
KLTVDKSRWQ QGNVFSCSVMHEALHNHYTQKSLSLSPGK (SEQ ID NO:16)

Fig. 16

The amino acid sequence of TAGsssNPP1 (N-terminus D8) + Fc + Linker

IGVLLTQRTLLSLVLALLFPSMASMEPKSCDKTHTCPPCPAPEAAGAPSVFLFPPKPKDTLMISRTPEVTCV
VVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP
APIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS
DGSFFLYS KLTVDKSRWQ QGNVFSCSVM HEALHNHYTQ KSLSLSPGK **DDDDDDDD**
*GGGGSGGGGS*AGFETPPTLLFSLDGFRAEYLHTWGGLLPVISKLKKCGTYTKNMRPVYPTKTFPNHYSIVT
GLYPESHGIIDNKMYDPKMNASFSLKSKEKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDI
YKMYNGSVPFEERILAVLQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMD
GLKELNLHRCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEGIARNL
SCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSGFHGSDNVFSNMQALF
VGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSL (SEQ ID NO:17)

Fig. 17

The amino acid sequence of TAGsssNPP1 (C-terminus D8) + Fc + Linker

IGVLLTQRTLLSLVLALLFPSMASM*EPKSCDKTHTCPPCPAPEAAGAPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK*AGFETPPTLLFSLDGFRAEYLHTWGGLLPVISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSKEKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAVLQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLHRCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEGIARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSGFHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSL*GGGGSGGGGS* **DDDDDDDD** (SEQ ID NO: 18)

Fig. 18

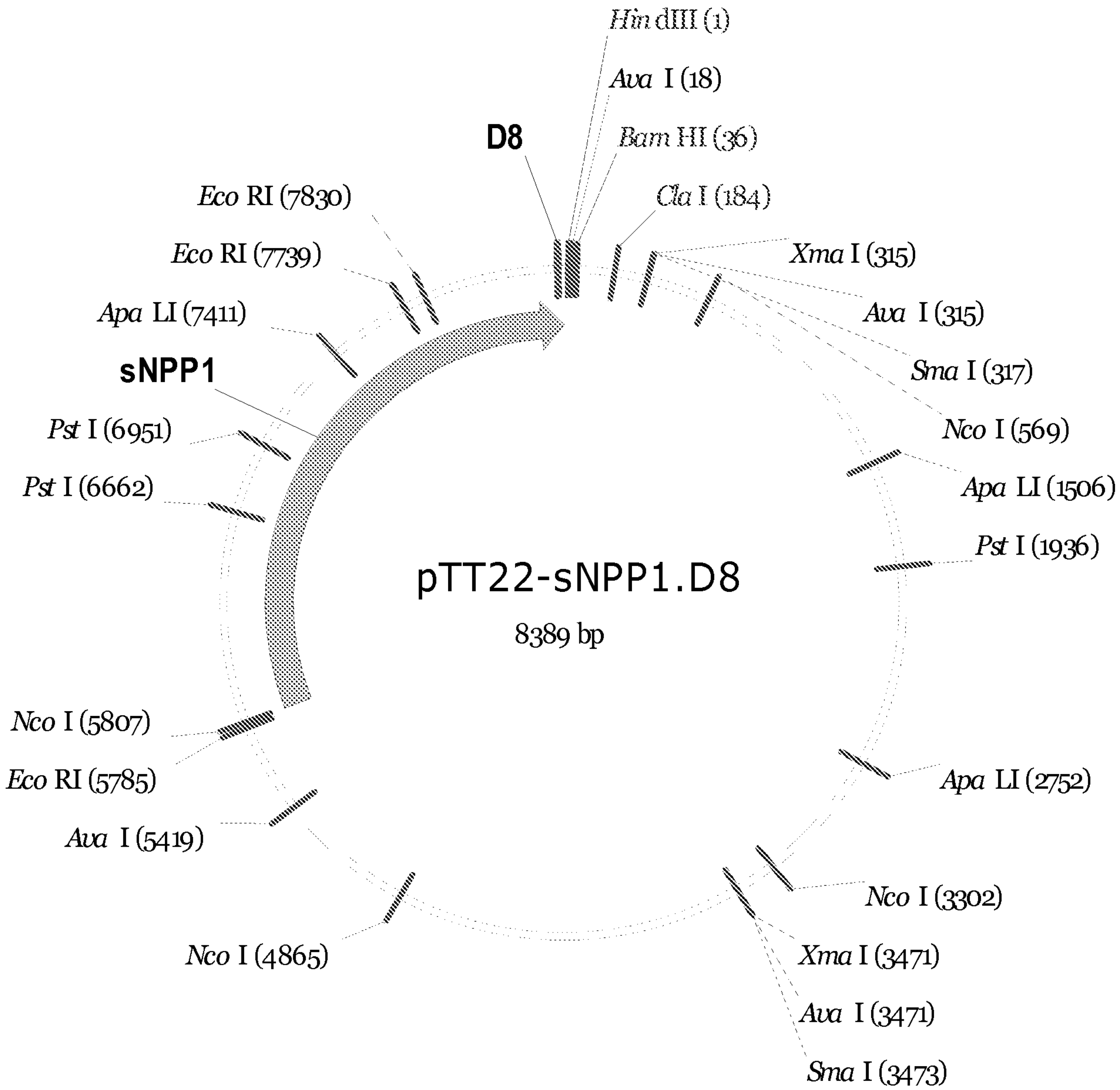
Hin dIII (1)
Ava I (18)
Bam HI (36)
Cla I (184)
Xma I (315)
Ava I (315)
Sma I (317)
Nco I (569)
Apa LI (1506)
Pst I (1936)
Apa LI (2752)
Nco I (3302)
Xma I (3471)
Ava I (3471)
Sma I (3473)
Nco I (4865)
Ava I (5419)
Eco RI (5785)
Nco I (5807)
Pst I (6662)
Pst I (6951)
sNPP1
Apa LI (7411)
Eco RI (7739)
Eco RI (7830)
D8
pTT22-sNPP1.D8
8389 bp

Fig. 19

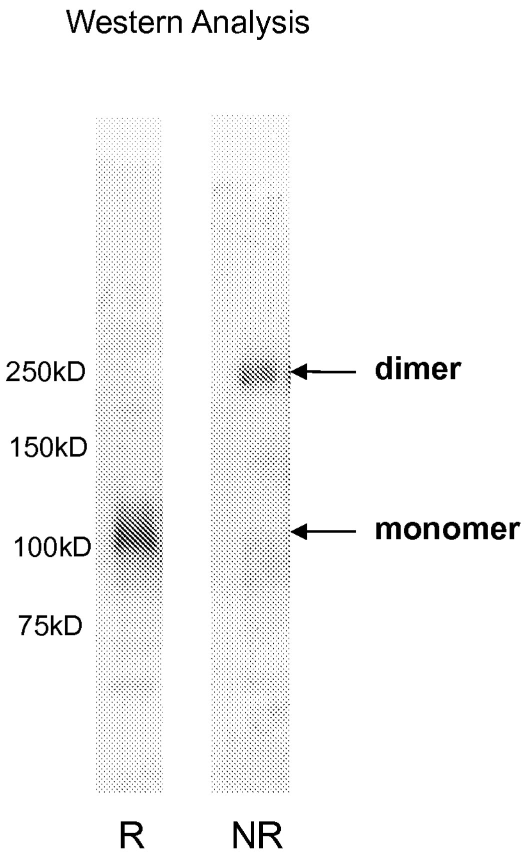

Fig. 20

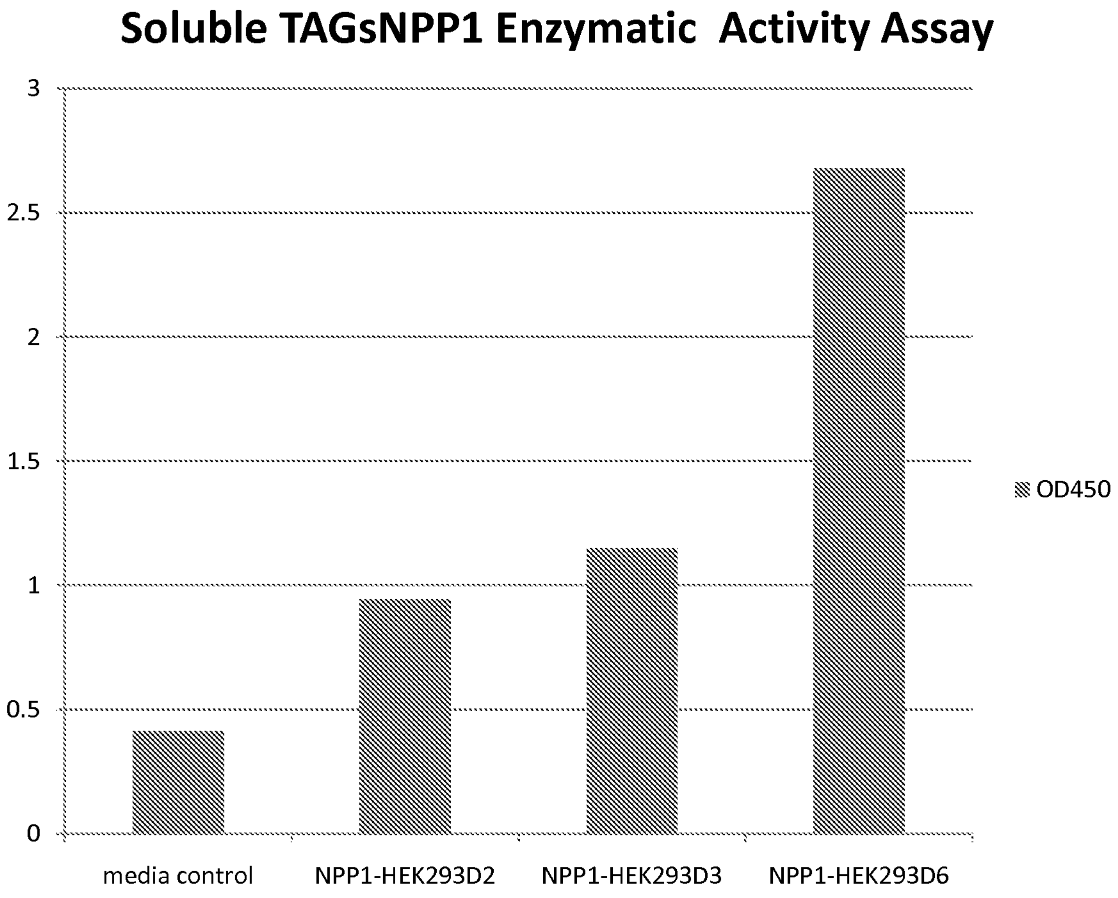
Soluble TAGsNPP1 Enzymatic Activity Assay
3
2.5
2
1.5
1
0.5
0
▨ OD450
media control
NPP1-HEK293D2
NPP1-HEK293D3
NPP1-HEK293D6

Fig. 21

Recombinant NPP1 fusion proteins

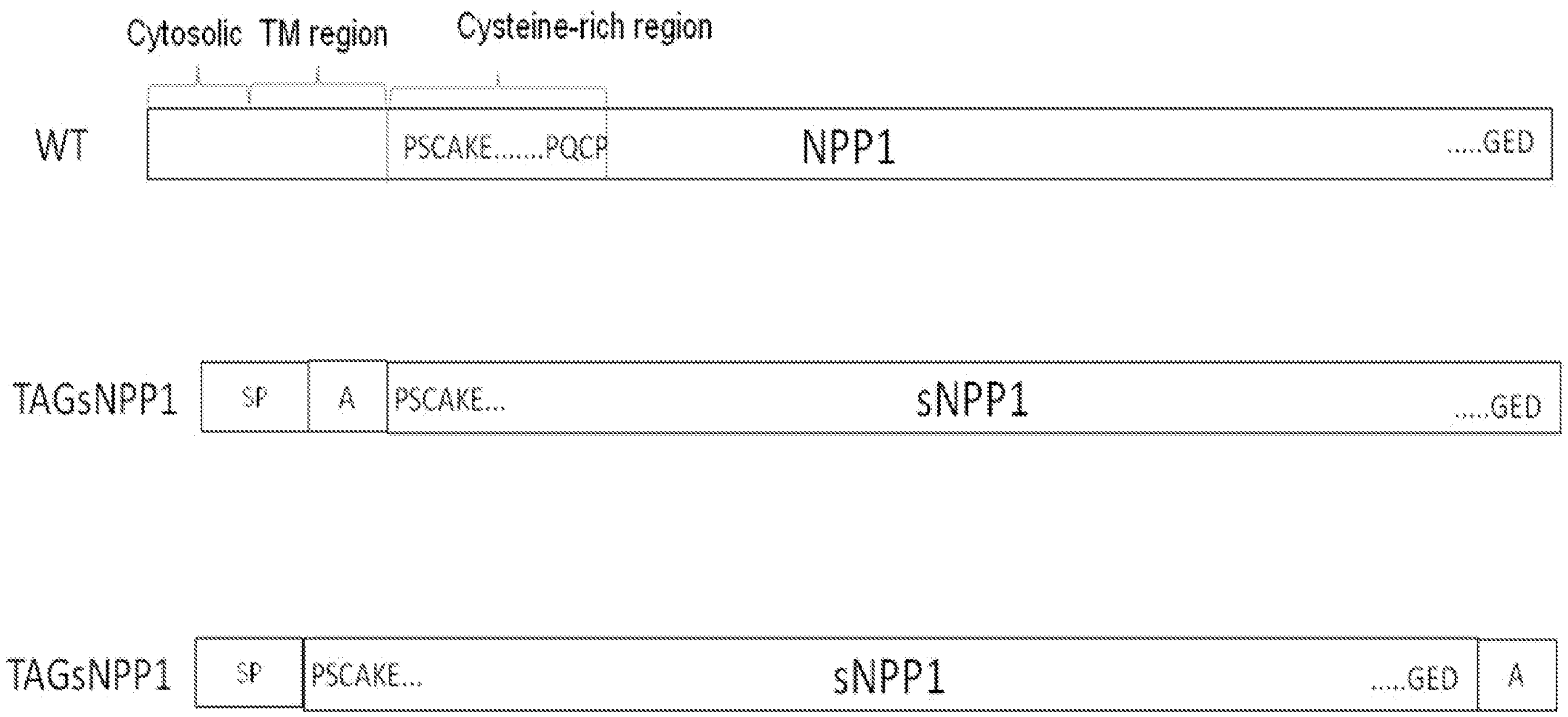

Fig. 22A

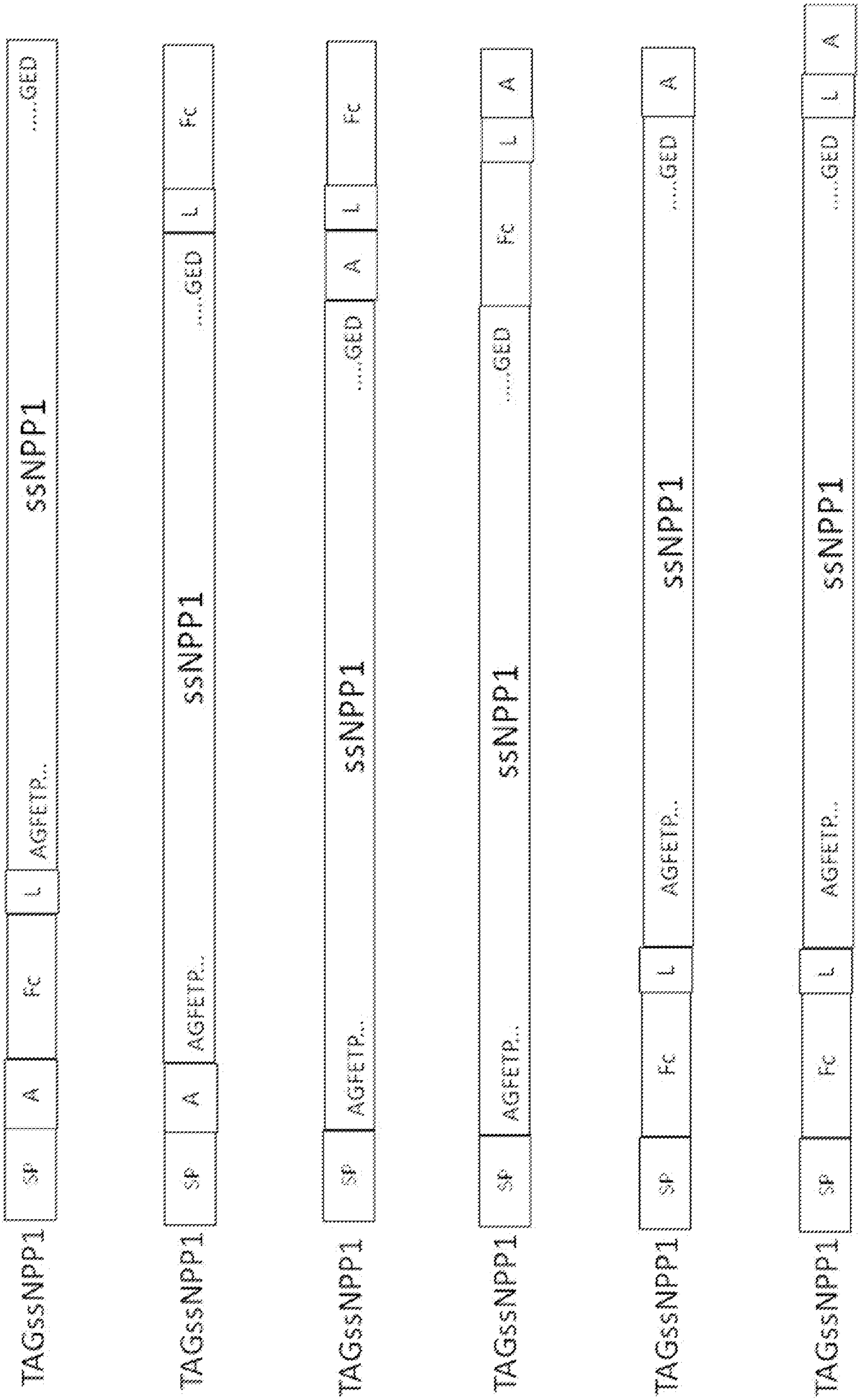
TAGssNPP1
SP
A
Fc
L
AGFETP...
ssNPP1
....GED
TAGssNPP1
SP
A
AGFETP...
ssNPP1
....GED
L
Fc
TAGssNPP1
SP
AGFETP...
ssNPP1
....GED
A
L
Fc
TAGssNPP1
SP
AGFETP...
ssNPP1
....GED
Fc
L
A
TAGssNPP1
SP
Fc
L
AGFETP...
ssNPP1
....GED
A
TAGssNPP1
SP
Fc
L
AGFETP...
ssNPP1
....GED
L
A

Fig. 22B

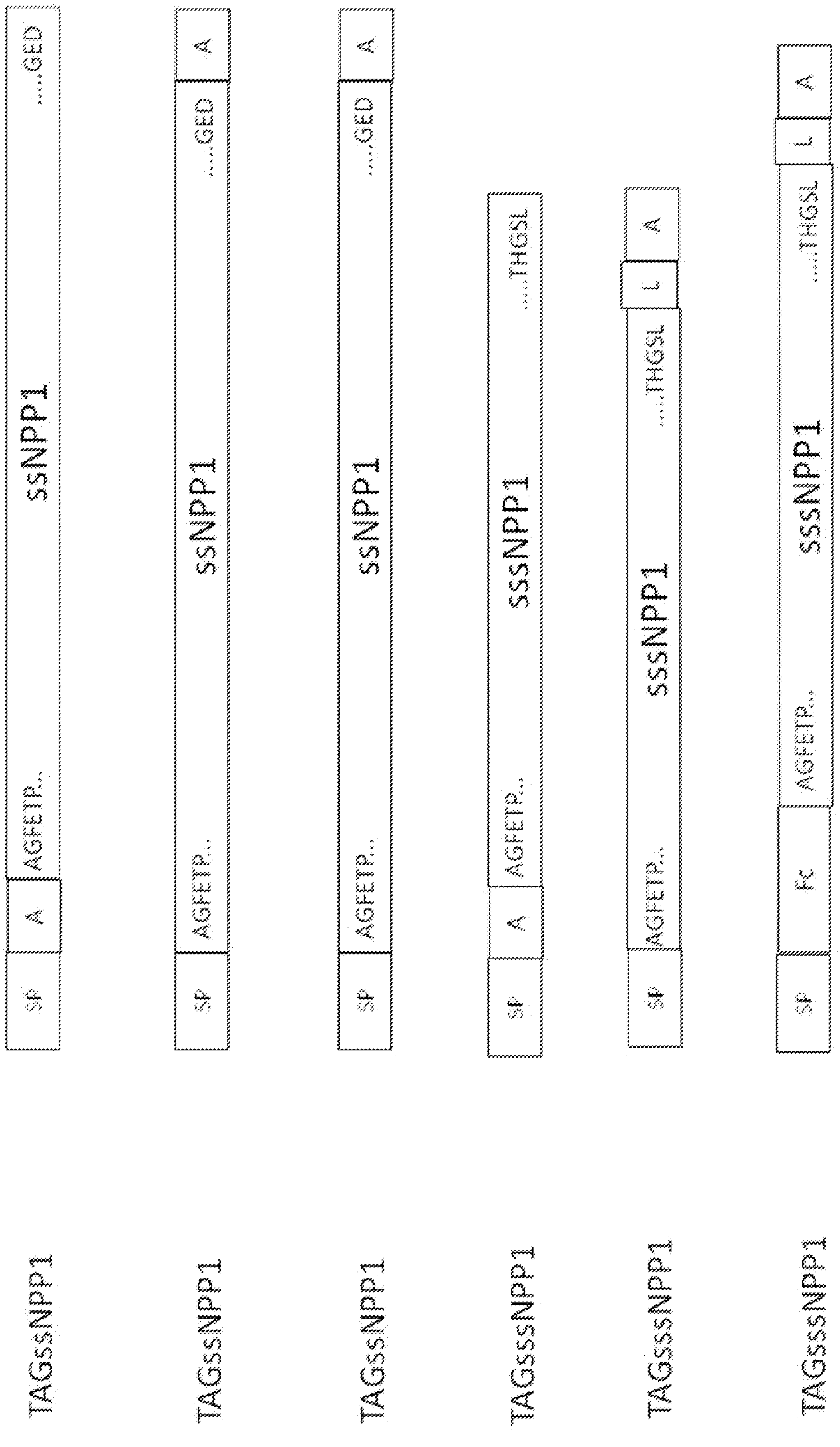
TAGssNPP1
SP
A
AGFETP...
SSNPP1
.....GED
TAGssNPP1
SP
AGFETP...
SSNPP1
.....GED
A
TAGssNPP1
SP
AGFETP...
SSNPP1
.....GED
A
TAGsssNPP1
SP
A
AGFETP...
SSSNPP1
.....THGSL
TAGsssNPP1
SP
AGFETP...
SSSNPP1
.....THGSL
L
A
TAGsssNPP1
SP
Fc
AGFETP...
SSSNPP1
.....THGSL
L
A

Fig. 22C