



(51) International Patent Classification:

C12N 15/67 (2006.01) C12N 15/74 (2006.01)
C12N 15/70 (2006.01)

(21) International Application Number:

PCT/IB2015/055189

(22) International Filing Date:

9 July 2015 (09.07.2015)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2245/MUM/2014 9 July 2014 (09.07.2014) IN

(71) Applicant: **LUPIN LIMITED** [IN/IN]; 159 CST Road, Kalina, Santacruz (East), Mumbai 400 098, State of Maharashtra (IN).

(72) Inventors: **SALUNKHE, Shardul**; Lupin Limited (Biotechnology Division), Gate No. 1156, Ghotawade Village, Mulshi Taluka, Pune 412 115 (IN). **VARSHNEY, Brajesh**; Lupin Limited (Biotechnology Division), Gate No. 1156, Ghotawade Village, Mulshi Taluka, Pune 412 115 (IN). **SOORAPANENI, Sudheerbabu**; Lupin Limited (Biotechnology Division), Gate No. 1156, Ghotawade Village, Mulshi Taluka, Pune 412 115 (IN).

(74) Agents: **MAJUMDAR, Subhatosh** et al.; S. Majumdar & Co., 5 Harish Mukherjee Road, Kolkata 700 025, State of West Bengal (IN).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

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

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

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: DUAL CISTRONIC BACTERIAL EXPRESSION SYSTEM

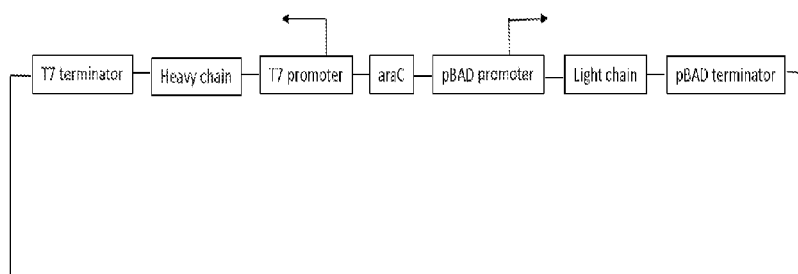


Figure 1

(57) Abstract: The present invention relates to the dual, independent cistron expression system in a single vector for the production of protein of interest proteins and peptides expressed as insoluble inclusion bodies formed in the bacteria *E. coli*. The present invention also provides the process for the expression of protein of interest using said bicistronic vector.

DUAL CISTRONIC BACTERIAL EXPRESSION SYSTEM

Field Of The Invention

The present invention relates to the dual, independent cistron expression system in a single
5 vector for the production of protein of interest comprises recombinant Fab fragments of
antibodies or other antibody fragment, peptides and protein expressed as insoluble inclusion
bodies formed in the bacteria *E. coli*. The present invention also provides the process for the
expression of protein of interest using said bicistronic vector.

Background Of The Invention

Recombinant DNA technology (rDNA) has revolutionized the way therapeutic agents are
prepared. The required proteins are now made inside a foreign cell and purified.

Proteins having post-translational modifications (PTMs) are generally expressed as
recombinant molecules in mammalian or yeast system. The yeast expression systems like
15 *Pichia* and *Saccharomyces* are closer to mammalian systems in terms of PTMs but still differ
in the types of glycosylations like high mannose glycans in case of *Pichia* make them
unsuitable for expression of recombinant proteins for human use.

Monoclonal antibodies (mAbs), antibodies, fusion proteins, Fab fragments of mAbs are used
20 as therapeutic agents. The rDNA technology uses specialized vectors and expression systems
for production of therapeutic proteins. The expression systems mainly consist of bacterial,
yeast, insect or mammalian expression systems. Initially, most of the recombinant proteins
were expressed in bacterial expression system using *E. coli* as host. There are several
advantages of using *E. coli* as expression host such as ease of cloning, ease of expression,
25 shorter timelines, shorter incubation periods and very high yields. Thus, proteins which do
not need any PTMs can be safely expressed in *E. coli*.

Fabs, which are antigen binding fragment part of mAbs, need not to be expressed in
mammalian systems as they do not contain glycosylation sites present in the Fc portion of the
antibody. Hence, Fabs are usually expressed in *E. coli* system. During 1980-90s several

researchers attempted the expression of Fabs in *E. coli*. Plückthun *et. al.*, 1990 *Behring Inst. Mitt.* (87):48-55 are some of the earlier workers who reported secretion of Fab antibody from *E. coli*. Williamson R.A. *et. al.*, 1991 *Biochem J.* 277 (Pt 2):561-3 reported use of bacteriophage lambda vectors for expression of Fab molecules in *E. coli*. Phage display system for production of Fab, bivalent antibody or chimeric antibody fragments in *E. coli*.
5 Moreover, Fab was also produced in *E. coli* as misfolded, inclusion bodies and then refolded them to get the functional molecule and thereby 40% increase in the yields of antibody was obtained.

- 10 Most of the studies mentioned above used single promoter, i.e., *phoA* to drive the expression of both heavy and light chains. The ribosome binding site (rbs) present in between heavy and light chains drives the transcription and translation of second gene.

US5648237 also used similar single promoter (*phoA*) strategy to express Fab genes in *E. coli* to get secreted product. The major drawback of the above strategy is that the expression levels of the second gene are usually lower than first gene, thus limiting the yields of the functional Fab.
15

Patent No. WO03018771 discloses a process for producing an antibody by two separate translational units, respectively encoding the light and heavy chains of said antibody or fragment, wherein both the chains are expressed in a sequential fashion, thereby specifically separating the production of the light and heavy chains and allowing the assembly of the light and heavy chains.
20

25 Patent No. EP1356052B1 discloses a method to produce full antibodies in prokaryotic cells. There is a presence of a first promoter and a first cistron to produce immunoglobulin light chain and a second promoter and a second cistron to produce immunoglobulin heavy chain, wherein both the chains are folded and assembled to form a biologically active immunoglobulin.

Summary Of The Invention

In an embodiment, the invention is related to a dual, independent cistron expression system in a single vector for the production of recombinant proteins and peptides expressed as insoluble inclusion bodies in bacterial cells.

In another embodiment, the invention is related to a process of preparation of dual, independent cistron expression system in a single vector having two different promoters for production of recombinant proteins and peptides expressed as insoluble inclusion bodies in bacterial cells.

In another embodiment, the invention is related to a dual, independent cistron expression system in a single vector having two different promoters for production of antibody fragments expressed as insoluble inclusion bodies in the bacterial cells.

In another embodiment, the invention is related to a dual, independent cistron expression system in a single vector having two different promoters for production of recombinant Fab fragment of antibodies expressed as insoluble inclusion bodies in the bacterial cells.

In another embodiment, the invention is related to a dual, independent cistron expression system in a single vector having two different promoters for production of recombinant peptides expressed as insoluble inclusion bodies in the bacterial cells.

In another embodiment the dual cistronic expression system comprises:

- a) first cistron comprising a promoter operably linked with polynucleotide sequence encoding protein of interest;
- b) second cistron comprising a promoter operably linked with polynucleotide sequence encoding protein of interest;

wherein the first and second cistrons are positioned in single vector and express a polynucleotide sequence encoding the protein of interest as inclusion bodies in bacterial cell.

5 In another embodiment the invention relates to dual cistronic vector comprising promoter operably linked to multiple cloning site contain gene of interest, ribosome binding site and terminator.

In yet another embodiment the invention relates to process for producing a protein of interest using dual cistronic expression system.

10

The details of one or more embodiments of the invention set forth below are illustrative in nature only and not intended to limit the scope of the invention. Other features, objects and advantages of the inventions will be apparent from the description.

15 **Brief description of accompanying figures**

Figure 1; illustrates the formula of bicistronic vector

Figure 2; illustrates the vector map of clone pET21a-HC-LC

Figure 3; illustrates the SDS PAGE analysis of insoluble pellet fraction of *E. coli* BL21A1 clone along with controls and the reference product

20 Figure 4; illustrates the RP-HPLC analysis of solubilized IB samples with LC and HC peaks seen in clone compared to reduced Fab molecule

Figure 5; illustrates the HPLC runs to separate heavy chain peak from other proteins

Figure 6; illustrates significant increase in expression of SAK-Lira clone in dual cistron construct as compared with single cistron clone in *E. coli* BL21 A1 cell line.

25 Figure 7; illustrates the vector map pBAD24M-LC

Detailed Description of the Invention

Definations:

As used herein, the term, "Protein of interest" refers herein to any polypeptide including protein and peptides used in biotherapeutic industry or for diagnostic or research purpose.

As used herein, the term "polynucleotide sequence encoding a protein of interest" as used
5 herein includes DNA coding for a gene, preferably a heterologous gene expressing the polypeptide.

As used herein, the terms "recombinant protein and peptide" refers to a protein or peptide that is produced by the expression of recombinant DNA within living cells.

10

As used herein, the terms "Fab" and "antibody" are used interchangeably because antibody comprises two parts, i.e., Fab and Fc region.

As used herein, the term "vector" refers to a DNA molecule used as a vehicle to artificially
15 carry foreign genetic material into bacterial cell, where it can be replicated and expressed.

As used herein, the term "cistron" refers to a section of DNA that contains the genetic code for a single polypeptide and functions as a hereditary unit.

20 As used herein, the term "dual independent cistron expression" refers to two separate cistrons which are used to express two same or different proteins independently.

As used herein, the term "same" is interchanble with identical or similar.

25 The term "Dual cistronic expression system" as used herein includes a polynucleotide sequence encoding a polypeptide to be expressed and sequences controlling its expression such as a promoter and optionally an enhancer sequence. The promoter of the invention is either operably linked to the gene to be expressed, i.e. transcription unit, or is separated therefrom by intervening DNA such as for example by the 5'-untranslated region of the

heterologous gene. Preferably the expression system is flanked by one or more suitable restriction sites in order to enable the insertion of the expression cassette into a vector and/or its excision from a vector. Thus, the expression system according to the present invention may be used for the construction of an expression vector, in particular a bacterial expression
5 vector.

As used herein, the term “promoters” refers to a regulatory region of DNA usually located upstream of a gene, providing a control point for regulated gene transcription.

10 As used herein, the terms "operably linked" refer to a functional relationship between two or more DNA segments, in particular gene sequences to be expressed and those sequences controlling their expression.

As used herein, the term “small peptides” or “peptides” refers to peptides ranging from 2 to
15 10 kDa used in biotherapeutic industry and diagnostic and research purposes like Liraglutide, exanetide, PTH, etc..

The present invention provides a dual cistronic expression system for the production of variety of recombinant protein of interest. In certain embodiment the dual cistronic
20 expression system comprises two cistrons having promoter operably linked with polynucleotide sequence encoding protein of interest and terminator.

In certain embodiment the dual cistronic expression system comprises two cistrons express polynucleotide sequence encoding protein of interest are positioned in single vector.

25

In an embodiment the dual cistronic expression system comprises;

a) first cistron comprising a promoter operably linked with polynucleotide sequence encoding protein of interest;

b) second cistron comprising a promoter operably linked with polynucleotide sequence encoding protein of interest;

wherein the first and second cistrons are positioned in single vector and express a polynucleotide sequence encoding the protein of interest as inclusion bodies formed in host cell.

In an embodiment the promoter may be selected from T7 promoter, arabinose promoter phoA, tac, lpp, lac-lpp, lac, trp, trc, preferably T7 promoter and arabinose promoter. In certain embodiment the dual cistronic expression system comprises two cistrons express polynucleotide sequence encoding protein of interest under the control of two promoters. In one embodiment both promoters control the expression of polynucleotide sequence encoding the same protein of interest. In another embodiment both promoters control the expression of polynucleotide sequence encoding protein of interest different in length of amino acid or physio-chemical properties.

In certain embodiment the protein of interest may be selected from peptides and proteins.

In some embodiment the proteins may be expressed in bicistronic vector. The protein comprises antibody or fragment thereof. Antibody fragment may be expressed in bicistronic expression system. The antibody fragment may be selected from Fab heavy chain and light chains of antibodies or other antibody fragments such as scFv, Diabodies, Triabodies, Tetrabodies, Bis-scFv, Minibodies Fab₂ (bispecific), Fab3 (trispecific). In preferred embodiment the bicistronic expression system express polynucleotide sequence encoding heavy chain and light chain of antibody which forms a Fab antibody. In such embodiment the Fab antibody shows affinity to VEGF receptor and said Fab antibody is Ranibizumab.

In another embodiment the protein may be selected from but not limiting to G-CSF, IFN, erythropoietin, insulin and its variants, PTH (1-84aa), FSH, LH, GH and Protein disulfide isomerase (PDI).

- 5 In some embodiment the peptides may be expressed in bicistronic vector. The peptides comprise amino acid sequence are selected from at least less than 40 amino acid or preferably less than 31 amino acid or more preferably less than 10 amino acid. In certain embodiment the peptide molecular weight is selected from about 2 to about 10 kDa. The peptide may be selected from but not limiting to GLP-1 peptide analogues such as Liraglutide
10 or Exendin or GLP-2 peptide like teduglutide and PTH (1-34aa) and insulin. In another preferred embodiment the bicistronic expression system express polynucleotide sequence encoding GLP-1 agonist peptide. In such embodiment the GLP-1 peptide is Liraglutide.

- 15 In another embodiment both promoters independently control the expression of different protein of interest such as heavy chain or light chain of antibody which are different in length of amino acid and physio-chemical property.

In an embodiment the dual cistronic expression system comprises:

- 20 a) first cistron comprises T7 promoter operably linked with polynucleotide sequence encoding heavy chain of antibody;
b) second cistron comprises arabinose promoter operably linked with polynucleotide sequence encoding light chain of antibody;

Wherein the first and second cistrons are positioned in single vector and express the heavy chain and light chain of the antibody as inclusion bodies formed in host cell.

25

In such embodiment the antibody heavy chain and light chain of antibody comprise the nucleotide sequence sequence ID nos. 1 and sequence ID nos. 2 or amino acid sequence sequence ID nos. 3 and sequence ID nos. 4. In some embodiment the position of first and second cistron is interchangeable wherein the second cistron may be cloned in vector at the

position of first cistron and first cistron may be positioned at second cistron. The heavy chain and light chain of antibody independently express as inclusion bodies and may be further treated to obtain Fab antibody which shows affinity to VEGF receptor and said Fab antibody is Ranibizumab.

5

In certain embodiment the heavy chain and light chain of antibody are optionally expressed in combination with signal peptide, preferably pelB. Signal peptide directs the expression of protein in periplasmic space of the host cell.

10 In embodiment, the dual cistronic expression system in a single vector having two different promoters, arabinose and T7 promoters regulating the production of heavy and light chains of recombinant Fab fragments, respectively and both having a pelB tag produced as insoluble inclusion bodies in the periplasmic space of *E. coli*.

15 In certain embodiment the heavy chain or light chain of antibody is optionally expressed in combination with regulator, preferably AraC gene in order to further increase the expression of protein.

The dual cistronic expression system provides equimolar expression of protein of interest.

20 Equimolar expression is highly desirable in order to obtain protein of interest in suitable quality and quantity. It depends over the ratio of heavy and light chain or the ratio of subunit of polypeptide cloned in to the vector. In certain embodiment the heavy chain and light chain are cloned in suitable ratio comprises the heavy chain is at least equal or higher than light chain to obtain equimolar expression of heavy and light chain. The heavy chain and
25 light chain are cloned in a ratio selected from 1:5:0.7 to 1:1 which includes 1:3:0.8, 1:2:0.9, 1:2:1 1:1.

In embodiment the dual cistronic expression system comprises nucleotide sequence as set forth in sequence ID nos 19.

In another embodiment the dual cistronic expression system comprising:

a) firstcistron comprises T7 promoter operably linked with polynucleotide sequence encoding peptide;

5 b) secondcistron comprises arabinose promoter operably linked with polynucleotide sequence encoding peptide;

wherein the first and second cistrons are positioned in single vector and express the peptide as inclusion bodies formed in host cell.

10 In such embodiment, the peptide is a GLP-1 analogue comprise nucleotide sequence as set forth in sequence ID nos 6 encoding the GLP-1 agonist peptide which is Liraglutide having amino acid sequence of sequence ID nos. 7.

In certain embodiemt the peptide may be optionally expressed with signal peptide or regulator/enhancer known to the skilled person.

15

In certain embodiemt the peptide may be optionally expressed with fusion partner or fusion tag in order to prevent the degradation of peptide. The fusion partner comprises amino acid sequence from 30 amino acid to 300 amino acid. The fusion partner comprises amino acid sequence selected from about 50 amino acid, 100 amino acid, about 136 amino acid, about 20 175 amino acid, about 250 amino acid, 300 amino acid, preferably about 136 amino acid. Fusion tag may be selected from but not limiting to Histidine-tag, glutathione-s-transferase (GST), Maltose binding protein, NusA, thioredoxin (TRX), polyhistidine (HIS), small ubiquitin-like modifier (SUMO) and ubiquitin (Ub) and staphylokinase (SAK) gene. In preferred embodiment the fusion tag is SAK gene. The detail use of SAK gene as fusion tag 25 with protein of interest is disclosed in US8853380 which is incorporated herein as reference.

In some embodiment the bicistronic expression system further comprises selection marker which is selected from ampicillin, kanamycine, preferably ampicillin.

In another embodiment the present invention provides a process for producing a protein of interest comprising the steps of:

(i) transforming the host cell with a single vector essentially consisting of dual cistronic expression system;

5 (ii) culturing the transformed cell in suitable medium to express protein of interest, wherein the first and second cistron expresses the protein of interest in inclusion bodies;

(iii) performing solubilization of the inclusion bodies;

(iv) performing refolding of the protein of interest.

10 In embodiment the bicistronic expression system is transfected in to the suitable bacterial host cell in order to express the protein of interest. The suitable bacterial host cell is *E.coli* in which the protein of interest is expressed in the form of inclusion bodies. Inclusion bodies are the insoluble substance formed in the periplasm or cytoplasm of *E.coli*. Inclusion bodies may be isolated, solubilized and protein of interest may be recovered in active form by the
15 techniques well known in the art.

In an embodiment, Fab heavy and light chains of antibodies or other antibody fragments such as scFv, Diabodies, Triabodies, Tetrabodies, Bis-scFv, Minibodies Fab₂ (bispecific), Fab3 (trispecific) were expressed as insoluble inclusion bodies in the periplasmic space of *E. coli* by
20 constructing two independent cistrons in a single vector having two different promoters, arabinose and T7 promoters. The two different promoters, i.e., T7 promoter and arabinose promoter helped in the expression of heavy and light chains of Fab molecule, respectively. The antibody heavy and light chains were produced as non-functional inclusion bodies in the bacterial cell, i.e, *E. coli* which are subsequently extracted, refolded and purified.

25

In an embodiment of the invention, the cistron comprises as such that each gene (heavy and light chain) would have its own promoter and terminator in a single vector. The heavy chain was cloned under the control of T7 promoter while light chain was cloned under the control

of arabinose promoter. Both the chains were preceded by signal sequence pelB tag for obtaining the product in the periplasmic space of the bacterial membrane.

5 The advantage of the bicistronic expression system is that both arabinose and T7 being strong promoters, high expression of both light and heavy chains are obtained from a single fermentation run instead of separate fermentations with light and heavy chain clones. The dual-cistron expression system makes it simpler to characterize and maintain a single cell bank instead of separate cell banks for light and heavy chain clones. Moreover, the inclusion bodies thus obtained are relatively pure when extracted from the periplasmic space of the bacterial cells. The high level of expression and much purer forms of the light and heavy chains obtained as inclusion bodies are relatively easier to fold into functional Fab *ex vivo*, thereby significantly increasing the yield of the product.

15 Another advantage of the system is that a protein of interest may be cloned and expressed under arabinose and T7 promoters and expression level of the protein may be increased significantly.

The examples disclosed below are only for illustrative purpose of the invention and are not intended to be limiting.

20

Example 1: Cloning of heavy chain in pET21a vector

25 The DNA sequence used for cloning of heavy and light chain of Fab fragments is given in sequence IDnos 1 and 2, respectively. The heavy chain insert was amplified from synthetic DNA using gene specific primers. Primers are designed according to methods well known in the art. The heavy chain PCR product was then digested with *NdeI-HindIII* enzymes and ligated to pET21a vector digested with the same enzymes.. The clones were screened by colony PCR and confirmed by restriction analysis. The resultant clone was designated as

pET21a-HC. The recombinant vector was introduced into BL21A1 cell line and checked for expression of heavy chain.

Example 2: Cloning of light chain in pBAD24M vector

5

The light chain insert was amplified from synthetic DNA using gene specific primers. Primers are designed according to methods well known in the art .The amplified light chain was digested with *NdeI-HindIII* enzymes and ligated to digested pBAD24M vector (available inthe laboratory) at same sites. The clones were screened by colony PCR and confirmed by
10 restriction analysis. The resultant clone was designated as pBAD24M-LC. The recombinant vector was introduced into BL21A1 cell line and checked for expression of light chain.

Example 3: Construction of two independent cistrons in same vector

15

Primers were designed to amplify light chain along with arabinose promoter, terminator and araCgene. Primers are designed according to methods well known in the art.The primers added *BglII* linker to the amplified product. The pET21a vector had single *BglII* site upstream of the T7 promoter. The light chain expression cassette was amplified from the template pBAD24M with the vector specific primers and cloned into pET21a-HC clone at *BglII* site. The clone was confirmed by restriction digestion and sequencing. The final clone
20 was designated as pET21a-HC-LC and suitable clone short listed based on expression. The clone map of the pET21a-HC-LC is presented in Figure 2.

The clone thus generated contains all the segments required for independent regulation and expression of both heavy and light chains.

25

Example 4: Expression analysis

*E. coli*BL21 A1 cell line was used as expression host. Apart from BL21 A1, BL21 DE3 or any other cell line containing T7 promoter in the genome is used. BL21 A1 cells were transformed using the above selected clone along with pET21a-HC and pBAD24M-LC as

controls. The Heavy chain was induced by IPTG while the light chain was induced by arabinose. The inducer concentration was 13 mM arabinose and 1 mM IPTG and the induction was done when the culture OD₆₀₀ was ~1. The cells were harvested 4 hr post induction. The study was done in shake flasks. The harvest obtained was bead lysed and
5 centrifuged to separate soluble and insoluble fractions. The samples were loaded on 12% SDS PAGE gels to check the expression. The SDS PAGE gel analysis is shown in Figure 3. Reduced ranibizumab was loaded in Lane 5 offFigure 3 to confirm expression of reduced light and heavy chains.

10 The SDS PAGE analysis showed expression of both chains in the insoluble pellet fraction and the same was confirmed by RP-HPLC analysis wherein the retention times of light and heavy chains of the reference product corresponded to the retention times of the in-house product. The controls used were reduced Fab molecule (reference product), and products of pET21a-HC clone and pBAD24M clone. Thus, the expression of both the heavy and the light
15 chain from a single clone was confirmed. The RP-HPLC analysis is shown in Figures 4 and 5. In RP-HPLC, solubilized and reduced IBs of the dual cistron clone were compared with reduced ranibizumab (RMP) and clones separately expressing heavy and light chains. i. e. pET21-HC and pBAD24MLC.

20 The retention time (RT) of principal peak of solubilized IB of pBAD24MLC expressing only light chain matches with the RT of light chain of reduced RMP. An impurity peak at RT 13 mins matches the retention time of heavy chain, which was indicative of similar hydrophobicity. The impurities therefore were characterized by LC-MS/MS and were finally annotated as host cell protein OMP C and light chain with uncleaved leader sequence at RT
25 17 mins. The heavy chain expressed by pET21a-HC matches with the reference standard heavy chain. The profile also indicated a post peak at RT 19 mins which was characterized as heavy chain with uncleaved leader sequence. The dual cistronic clone pET21a_HC_LC that expresses both LC and HC have 2 main peaks that have equivalent retention times as that of

the LC and HC of reference standard. But, since the OMP C co-eluted with heavy chain, the reversed phase method of testing had to be resolved better and this is presented in the Fig. 4.

5 The existing method on Zorbax C8 RP column was modified to Aeriswidepore C8 and the co-eluting species were resolved. The solubilized IB of pET21a_HC_LC on Aeriswidepore C8 exhibited a distinct LC, HC and OMP C peaks enabling the identity and precise quantification of individual subunits in IB as evident in Figure 5.

10 Although certain embodiments and examples have been described in detail above, those having ordinary skill in the art will clearly understand that many modifications are possible in the embodiments and examples without departing from the teachings thereof.

Example no 5: Cloning of small peptide (Liraglutide) with staphylokinase (SAK) fusion tag in pET24a vector

15 The SAK and Liraglutide genes were amplified from synthetic DNA using gene specific primers. Primers are designed according to methods well known in the art and PCR products were digested with *NdeI-BamHI* and *BamHI-HindIII* enzymes and ligated to digested pET24a vector at *NdeI-HindIII* sites. The clones were screened by colony PCR and confirmed by restriction analysis. The resultant clone was designated as pET24a-SAK-Lira.

20

Example 6: Cloning of small peptide (Liraglutide) with staphylokinase fusion tag in pBAD24M vector

As given in example no 6, Liraglutide with SAK tag was cloned into pBAD24M vector. The clone was designated as pBAD24M-SAK-Lira.

25

Example no 7: Construction of two independent cistrons in same vector with both cistrons expressing SAK-Lira fusion peptide

The clone design strategy used in example no. 3 was used to construct dual cistron clone of Liraglutide, wherein SAK-Lira fusion gene alongwith arabinose expression cassette was

amplified from pBAD24M-SAK-Lira clone and cloned into pET24a-SAK-Lira clone to construct dual cistron construct. The clone was labelled as pET-ara-SAK-Lira

Example no 8: Expression analysis of dual cistron clone with SAK-Lira fusion protein.

5 *E. coli* BL21 A1 cell line was used as expression host. Apart from BL21 A1, BL21 DE3 or any other cell line containing T7 promoter in genome is used. BL21 A1 cells were transformed using the above single and dual cistron constructs. The clones were induced by IPTG and arabinose. The inducer concentration was 13 mM arabinose and 1 mM IPTG and the induction was done when the culture OD₆₀₀ was ~1. The cells were harvested 4 hr post
10 induction. The study was done in shake flasks. The harvest obtained was bead lysed and centrifuged to separate soluble and insoluble fractions. The samples were loaded on 12% SDS PAGE gels to check the expression.

The SDS PAGE gel analysis clearly shows increased expression of SAK-Lira fusion protein in dual cistron clone (Figure 6 lane 2) as compared with single cistron pET24a-SAK-Lira
15 clone (Figure 6 Lane 3).

20

25

CLAIMS

1. A dual cistronic expression system comprising:
 - a) first cistron comprising a promoter operably linked with polynucleotide sequence
5 encoding protein of interest;
 - b) second cistron comprising a promoter operably linked with polynucleotide sequence
encoding protein of interest;
wherein the first and second cistrons are positioned in single vector and express a
polynucleotide sequence encoding the protein of interest as inclusion bodies formed
10 in host cell.
2. The dual cistronic expression system as claimed in claim 1 wherein the first cistron
and second cistron comprise a promoter selected from T7 promoter, arabinose
promoter phoA, tac, lpp, lac-lpp, lac, trp and trc.
3. The dual cistronic expression system as claimed in claim 1 or claim 2 wherein the
15 first and second cistron comprises a different promoter.
4. The dual cistronic expression system as claimed in claim 2 wherein the first cistron
comprises a T7 promoter or arabinose promoter.
5. The dual cistronic expression system as claimed in claim 2 wherein the second cistron
comprises a T7 promoter or arabinose promoter.
- 20 6. The dual cistronic expression system as claimed in claim 1 wherein the protein of
interest is selected from proteins and peptides.
7. The dual cistronic expression system as claimed in claim 6 wherein the proteins are
selected from GCSF, EPO, FSH, IFN, PTH (1-84 aa), Insulin, LH, RH, antibodies or
fragments thereof.
- 25 8. The dual cistronic expression system as claimed in claim 7 wherein the antibody
fragments are selected from light chain of antibody, heavy chain of antibody, scFv,
FAB, Diabodies, Triabodies, Tetrabodies, Bis-scFv, Minibodies, Fab2 (bispecific)
antibodies and Fab3 (trispecific) antibodies.

9. The dual cistronic expression system as claimed in claim 8 wherein the antibody fragment is Fab.
10. The dual cistronic expression system as claimed in claim 9 wherein the Fab fragment of antibody is Ranibizumab.
- 5 11. The dual cistronic expression system as claimed in claim 6 wherein the peptides are selected from Liraglutide, exanetide, PTH (1-34aa) and teduglutide.
12. The dual cistronic expression system as claimed in any one of the preceding claim comprising:
- 10 a) firstcistron comprising T7 promoter operably linked with polynucleotide sequence encoding heavy chain of antibody;
- b) secondcistron comprising arabinose promoter operably linked with polynucleotide sequence encoding light chain of antibody;
- wherein the first and second cistrons are positioned in single vector and express the heavy chain and light chain of the antibody as inclusion bodies formed in host cell.
- 15 13. The dual cistronic expression system as claimed in claim 12 wherein the heavy chain of antibody comprises the nucleotide sequence as set forth in sequence ID no 1 or the amino acid sequence as set forth in sequence ID no 3.
14. The dual cistronic expression system as claimed in claim 12 wherein the light chain of antibody comprises the nucleotide sequence as set forth in sequence ID no 2 or the amino acid sequence as set forth in sequence ID no 4.
- 20 15. The dual cistronic expression system as claimed in claim 12 wherein the heavy chain and light chain are cloned in suitable ratio having the heavy chain is equal or higher than light chain to obtain equimolar expression of heavy and light chain.
16. The dual cistronic expression system as claimed in any one of the preceding claim wherein the heavy and light chain of antibody have affinity to VEGF.
- 25 17. The dual cistronic expression system as claimed in any one of the preceding claim comprising:
- a) firstcistron comprising T7 promoter operably linked with polynucleotide sequence encoding peptide

b) second cistron comprising arabinose promoter operably linked with polynucleotide sequence encoding peptide;

wherein the first and second cistrons are positioned in single vector and express the peptide as inclusion bodies formed in host cell.

5 18. The dual cistronic expression system as claimed in claim 17 wherein the peptide is selected from GLP-1 analogue, GLP-2 analogue.

19. The dual cistronic expression system as claimed in claim 17 wherein the peptide comprise nucleotide sequence as set forth in sequence ID no 6 or the amino acid sequence as set forth in sequence ID no 7.

10 20. The dual cistronic expression system as claimed in claim 18 wherein the GLP-1 agonist peptide is Liraglutide.

21. The dual cistronic expression system as claimed in any of preceding claim wherein peptide or protein is expressed in combination with fusion tag.

15 22. The dual cistronic expression system as claimed in claim 21 wherein the fusion tag is staphylokinase (SAK) gene comprises nucleotide sequence as set forth in sequence ID 8.

23. The dual cistronic expression system as claimed in any one of the preceding claim further comprises an optional one or more elements selected from the group consisting of polynucleotide sequence encoding regulator to enhance the expression
20 of protein of interest, polynucleotide sequence encoding signal peptide to regulate the expression of protein of interest in periplasmic space of host cell and selection marker.

24. The dual cistronic expression system as claimed in claim 23 wherein the regulator enhance the expression of light chain of antibody.

25 25. The dual cistronic expression system as claimed in claim 23 wherein the regulator is AraC gene.

26. The dual cistronic expression system as claimed in claim 23 wherein the signal peptide is pelB.

27. The dual cistronic expression system as claimed in claim 23 wherein the selection marker is ampicillin.

28. The dual cistronic expression system as claimed in any one of the preceding claim wherein the host cell is *E.coli*.

5 29. A process for producing a protein of interest comprising the steps of:

(i) transforming the host cell with a single vector essentially consisting of dual cistronic expression system as claimed in any of the preceding claims;

(ii) culturing the transformed cell in suitable medium to express protein of interest, wherein the first and second cistron expresses the protein of interest in inclusion
10 bodies;

(iii) performing solubilization of the inclusion bodies;

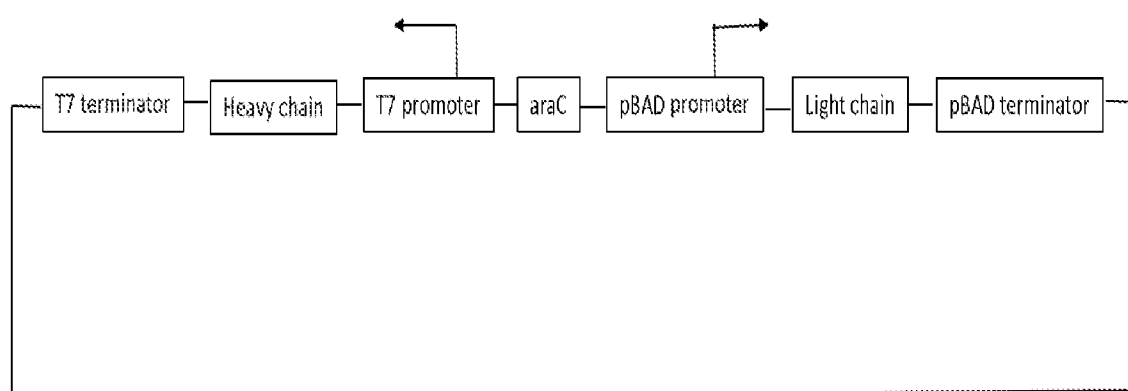
(iv) performing refolding of the protein of interest.

30. The dual cistronic expression system as claimed in any one of the preceding claim comprises nucleotide sequence as set forth in sequence id no 19.

15

20

25

**Figure 1**

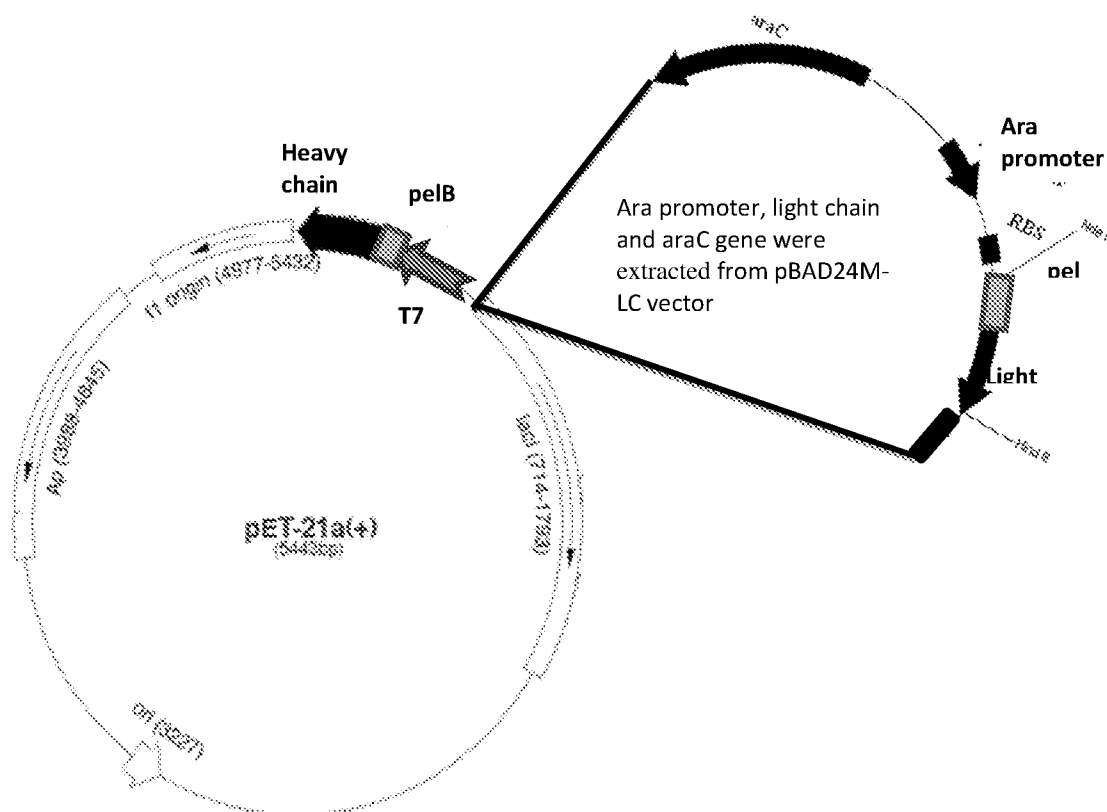
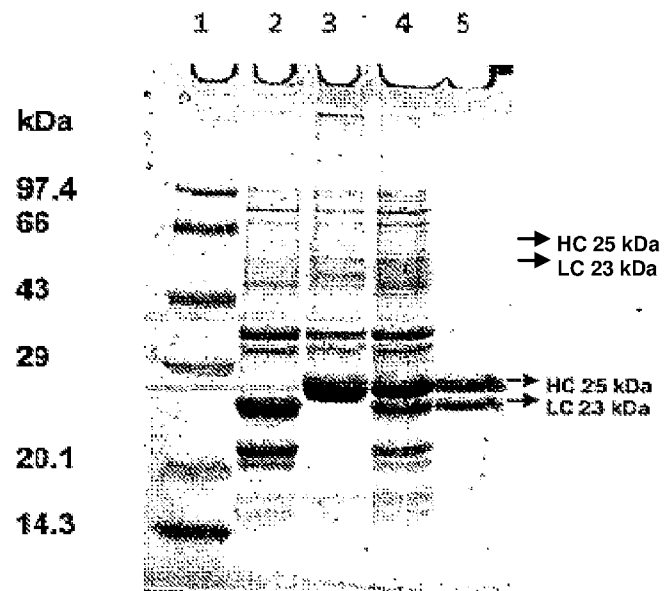


Figure 2



Lane. No	Lane description
1	Protein molecular weight marker
2	Light chain pBAD/BL21 A1
3	Heavy chain pET21a/BL21A1
4	Dual cistron clone pET21a-HC-LC /BL21 A1
5	Ranibizumab standard

Figure 3

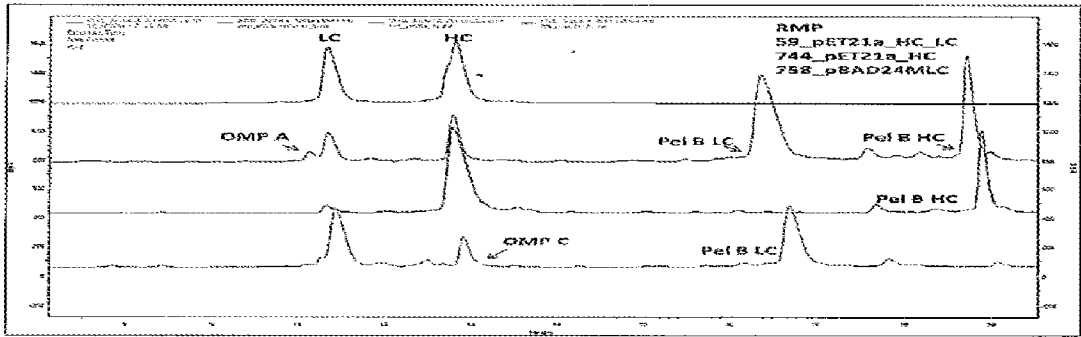


Figure 4

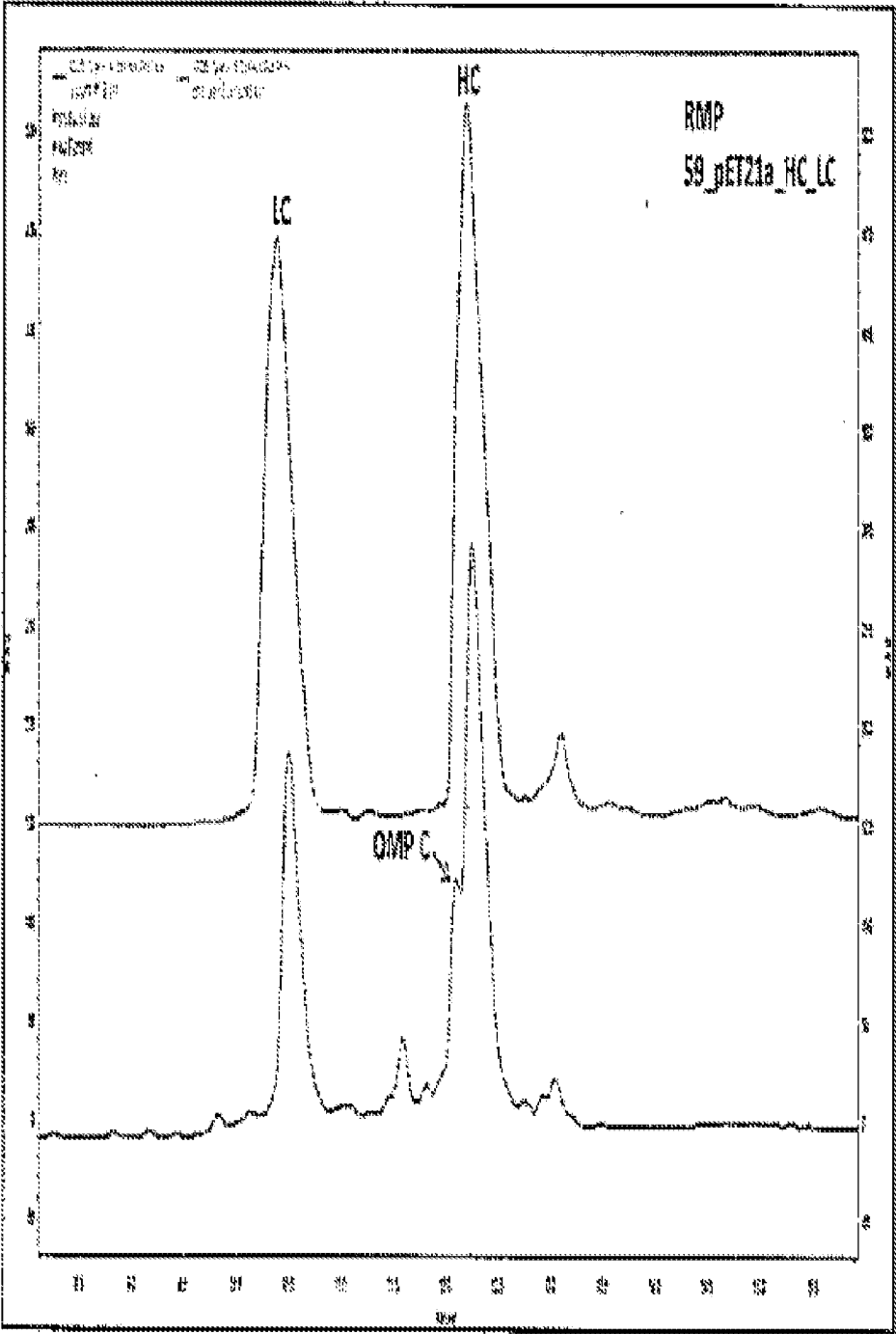
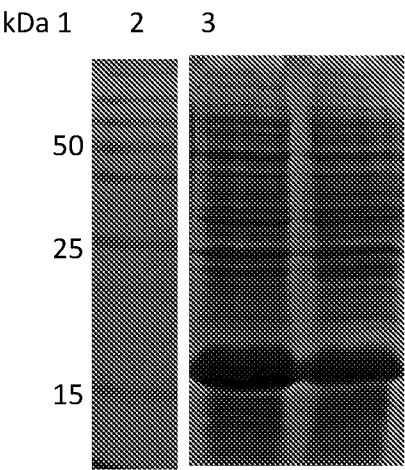


Figure 5



Lane. No	Lane description
1	Protein molecular weight marker
2	Dual cistron clone pET-ara-SAK-Lira /BL21 A1
3	Single cistron clone pET-SAK-Lira

Figure 6

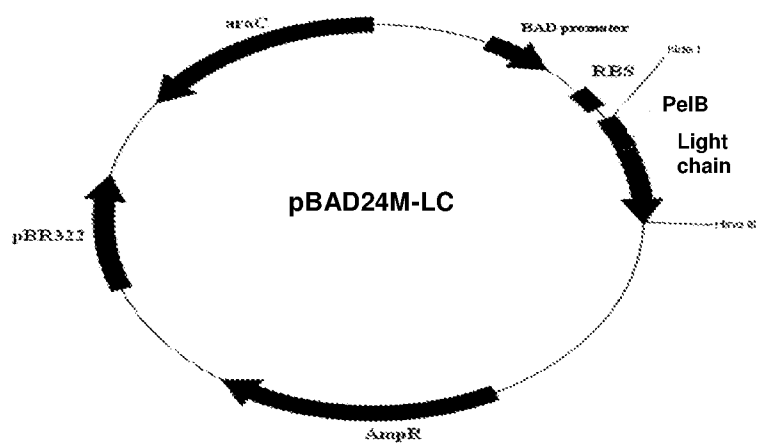


Figure 7

FPAA4613PCT-seq1-000001
SEQUENCE LISTING

<110> LUPIN LIMITED

<120> Dual cistronic bacterial expression system

<130> FPAA4613PCT

<150> 2245/MUM/2014

<151> 2014-07-09

<160> 19

<170> PatentIn version 3.5

<210> 1

<211> 762

<212> DNA

<213> artificial

<220>

<223> Ranibizumab heavy chain sequence

<400> 1

atgaaatacc tgctgccgac agctgctgct ggtctgctgc tcctcgctgc ccagccggcg	60
atggccgaag tccaactggt cgaatcgggt ggtggtctgg tccaaccggg tggctctctg	120
cgtctgtcgt gtgctgcctc gggctatgat ttaccatt atggtatgaa ctgggtccgt	180
caggccccgg gtaaaggctt ggaatgggtg ggctggatta atacctacac gggatgaaccg	240
acctatgcgg ccgattttta acgtcgcttt acgttctctc tggacacctc gaagagcacg	300
gcatactctg agatgaacag tctgcgcgcg gaagataccg ccgtgtatta ctgcgcgaag	360
taccctgatt actatggcac gtccactgg tattttgacg tttggggcca aggtaccctg	420
gtcaccgtga gcagcgcgag caccaaaggc ccgagcgtgt tcccgctggc cccgagttcc	480
aagtctacca gtggcggtag ggcagctctg ggttgtctgg ttaaagatta ttttccggaa	540
ccggttaccg tctcctggaa cagcggcgca ctgacctctg gtgtgcatac gttcccggt	600
gttctgcagt catcgggcct gtacagcctg agcagcgtgg ttaccgttcc gagttcctca	660
ctgggtaccc aaacgtatat ctgcaacgtc aatcacaaac cgagcaatac caaagtggac	720
aaaaaagtgg aaccgaaatc gtgtgataaa acgcatctgt aa	762

<210> 2

<211> 711

<212> DNA

<213> artificial

<220>

<223> Ranibizumab Light chain: 2295-3005Nt seq

<400> 2

atgaaatacc tgctgccgac agctgctgct ggtctgctgc tcctcgctgc ccagccggcg	60
atggccgaca ttcaactgac gcaaagtccg agcagcctga gcgcatccgt gggcgaccgt	120
gtgacgatta cctgttccgc aagccaagac atctctaact atctgaattg gtaccagcaa	180
aaaccgggca aggcaccgaa agtcctgatt tattttacca gctctctgca ttccggcggt	240

FPAA4613PCT-seq1-000001

ccgtcacgtt ttagcggctc tggtagtggc accgatttca ccctgacgat cagttccctg 300
cagccggaag actttgctac gtattactgc cagcaatata gcaccgtgcc gtggacgttc 360
ggtcagggca ccaaggttga aattaaacgt acggttgccg ccccgctctgt ctttatcttc 420
ccgccgagtg atgaacagct gaaatcgggt accgcaagcg tggtttgtct gctgaacaat 480
ttctatccgc gcgaagcaaa ggtccagtgg aaagtggaca acgctctgca gtccggcaat 540
tcacaagaat cggtgaccga acaagatagc aaggactcta cgtacagtct gtcacgacc 600
ctgacgctgt ccaaagcgga ttatgaaaaa cacaaggttt acgcctgcga agtcacccat 660
caaggtctgt cgtctccggt taccaagagt ttcaatcgtg gcgaatgtta a 711

<210> 3
<211> 253
<212> PRT
<213> artificial

<220>
<223> Ranibizumab heavy chain amino acid sequence

<400> 3

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

Ala Gln Pro Ala Met Ala Glu Val Gln Leu Val Glu Ser Gly Gly Gly
20 25 30

Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
35 40 45

Tyr Asp Phe Thr His Tyr Gly Met Asn Trp Val Arg Gln Ala Pro Gly
50 55 60

Lys Gly Leu Glu Trp Val Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro
65 70 75 80

Thr Tyr Ala Ala Asp Phe Lys Arg Arg Phe Thr Phe Ser Leu Asp Thr
85 90 95

Ser Lys Ser Thr Ala Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp
100 105 110

Thr Ala Val Tyr Tyr Cys Ala Lys Tyr Pro Tyr Tyr Tyr Gly Thr Ser
115 120 125

His Trp Tyr Phe Asp Val Trp Gly Gln Gly Thr Leu Val Thr Val Ser
130 135 140

Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser
145 150 155 160

Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp

FPAA4613PCT-seql-000001

165

170

175

Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr
180 185 190

Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr
195 200 205

Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln
210 215 220

Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp
225 230 235 240

Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Leu
245 250

<210> 4
<211> 236
<212> PRT
<213> artificial

<220>
<223> Ranibizumab Light chain Amino acid sequence

<400> 4

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

Ala Gln Pro Ala Met Ala Asp Ile Gln Leu Thr Gln Ser Pro Ser Ser
20 25 30

Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Ser Ala Ser
35 40 45

Gln Asp Ile Ser Asn Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys
50 55 60

Ala Pro Lys Val Leu Ile Tyr Phe Thr Ser Ser Leu His Ser Gly Val
65 70 75 80

Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr
85 90 95

Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln
100 105 110

Tyr Ser Thr Val Pro Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
115 120 125

Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp
130 135 140

FPAA4613PCT-seql-000001

Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn
145 150 155 160

Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu
165 170 175

Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp
180 185 190

Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr
195 200 205

Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser
210 215 220

Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
225 230 235

<210> 5
<211> 762
<212> DNA
<213> artificial

<220>
<223> Heavy chain 940-178 complement

<400> 5
ttacagatgc gttttatcac acgatttcgg ttccactttt ttgtccactt tggatttgct 60
cggtttgtga ttgacgttgc agatatacgt ttgggtaccc agtgaggaaac tcggaacggt 120
aaccacgctg ctcaggctgt acaggcccga tgactgcaga acagccggga acgtatgcac 180
accagaggtc agtgcgccgc tgttccagga gacggtaacc gggtccggaa aataatcttt 240
aaccagacaa cccagagctg ccgtaccgcc actggtagac ttggaactcg gggccagcgg 300
gaacacgctc gggccttttg tgctcgcgct gtcacgggtg accaggggtac cttggcccca 360
aacgtcaaaa taccagtggg acgtgccata gtaatacggg tacttcgcgc agtaatacac 420
ggcggtatct tccgcgcgca gactgttcat ctgcagatat gccgtgctct tcgaggtgtc 480
cagagagaac gtaaagcgac gtttaaaatc ggccgcatag gtcggttcac ccgtgtaggt 540
attaatccag cccacccatt ccagaccttt acccggggcc tgacggaccc agttcatacc 600
ataatgggta aaatcatagc ccgaggcagc acacgacaga cgcagagagc cacccggttg 660
gaccagacca ccacccgatt cgaccagttg gacttcggcc atcgccggct gggcagcgag 720
gagcagcaga ccagcagcag ctgtcggcag caggatatttc at 762

<210> 6
<211> 96
<212> DNA
<213> artificial

<220>
<223> DNA sequence of Liraglutide

FPAA4613PCT-seql-000001

<400> 6
catgcagaag gcacctttac gagtgatgtg agctcttatac tggaaggcca ggcggccaaa 60
gaatttattg cgtggctggt tcgtggccgt ggttaa 96

<210> 7
<211> 31
<212> PRT
<213> artificial

<220>
<223> Amino acid sequence of Liraglutide

<400> 7
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Arg Gly Arg Gly
20 25 30

<210> 8
<211> 456
<212> DNA
<213> artificial

<220>
<223> SAK nucleotide sequence (with EK site)

<400> 8
catatgtcaa gttcattcga caaaggaaaa tataaaaaag gcgatgacgc gagttatttt 60
gaaccaacag gcccgtattt gatggtaa atgtgactggag ttgatggtaa aggaaatgaa 120
ttgctatccc ctcatattgt cgagtttcct attaaacctg ggactacact tacaaaagaa 180
aaaattgaat acctgcagga tgatgatgat aaatacgtag aatgggcatt agatgcgaca 240
gcatataaag agtttagagt agttgaatta gatccaagcg caaagatcga agtcacttat 300
tatgataaga ataagaaaaa agaagaaacg aagtctttcc ctataacaga aaaagggttt 360
gttgtcccag atttatcaga gcatattaaa aaccctggat tcaacttaat tacaaagggt 420
gttatagaaa agaaaggatc cgatgatgat gataaa 456

<210> 9
<211> 151
<212> PRT
<213> artificial

<220>
<223> SAK amino acid sequence (with EK tag)

<400> 9
Met Ser Ser Ser Phe Asp Lys Gly Lys Tyr Lys Lys Gly Asp Asp Ala
1 5 10 15

Ser Tyr Phe Glu Pro Thr Gly Pro Tyr Leu Met Val Asn Val Thr Gly
20 25 30

FPAA4613PCT-seql-000001

Val Asp Gly Lys Gly Asn Glu Leu Leu Ser Pro His Tyr Val Glu Phe
35 40 45

Pro Ile Lys Pro Gly Thr Thr Leu Thr Lys Glu Lys Ile Glu Tyr Leu
50 55 60

Gln Asp Asp Asp Asp Lys Tyr Val Glu Trp Ala Leu Asp Ala Thr Ala
65 70 75 80

Tyr Lys Glu Phe Arg Val Val Glu Leu Asp Pro Ser Ala Lys Ile Glu
85 90 95

Val Thr Tyr Tyr Asp Lys Asn Lys Lys Lys Glu Glu Thr Lys Ser Phe
100 105 110

Pro Ile Thr Glu Lys Gly Phe Val Val Pro Asp Leu Ser Glu His Ile
115 120 125

Lys Asn Pro Gly Phe Asn Leu Ile Thr Lys Val Val Ile Glu Lys Lys
130 135 140

Gly Ser Asp Asp Asp Asp Lys
145 150

<210> 10
<211> 17
<212> DNA
<213> artificial

<220>
<223> T7 promoter: 1011-1027

<400> 10
ctatagtgag tcgtatt

17

<210> 11
<211> 52
<212> DNA
<213> artificial

<220>
<223> T7 terminator:26-72

<400> 11
ttcagcaaaa aaccctcaa gaccgttta gaggcccaa ggggttatgc ta

52

<210> 12
<211> 10
<212> DNA
<213> artificial

<220>
<223> RBS of pET21a: 947-957

<400> 12
tctccttctt

10

FPAA4613PCT-seq1-000001

<210> 13
<211> 879
<212> DNA
<213> artificial

<220>
<223> AraC gene: 1048-1927

<400> 13
ttatgacaac ttgacggcta catcattcac tttttcttca caaccggcac ggaactcgct 60
cgggctggcc ccggtgcatt ttttaaatac ccgcgagaaa tagagttgat cgtcaaaacc 120
aacattgcga ccgacggtgg cgatagggcat ccgggtggtg ctcaaaagca gcttcgcctg 180
gctgatacgt tggtcctcgc gccagcttaa gacgctaate cctaactgct ggcggaaaag 240
atgtgacaga cgcgacggcg acaagcaaac atgctgtgcg acgctggcga tatcaaaatt 300
gctgtctgcc aggtgatcgc tgatgtactg acaagcctcg cgtaccgat tatccatcgg 360
tggatggagc gactcgtaa tgcgttccat gcgccgcagt aacaattgct caagcagatt 420
tatcgccagc agctccgaat agcgcccttc cccttgcccg gcgttaatga tttgccc aaa 480
caggtcgctg aaatgcggct ggtgcgcttc atccggggcga aagaaccccg tattggcaaa 540
tattgacggc cagttaagcc attcatgcca gtaggcgcgc ggacgaaagt aaaccactg 600
gtgataccat tcgcgagcct ccggatgacg accgtagtga tgaatctctc ctggcgggaa 660
cagcaaaata tcacccggtc ggcaaaaaa ttctcgtccc tgatttttca ccacccctg 720
accgcaatg gtgagattga gaataaacc tttcattccc agcggtcggt cgataaaaaa 780
atcgagataa ccgttggcct caatcggcgt taaacccgcc accagatggg cattaaacga 840
gtatcccggc agcaggggat cattttgcgc ttcagccat 879

<210> 14
<211> 27
<212> DNA
<213> artificial

<220>
<223> Arabinose promoter: 2203-2230

<400> 14
acgcttttta tcgcaactct ctactgt 27

<210> 15
<211> 425
<212> DNA
<213> artificial

<220>
<223> Arabinose terminator: 3011-3438

<400> 15
gctgttttgg cggatgagag aagattttca gcctgataca gattaaatca gaacgcagaa 60
gcggtctgat aaaacagaat ttgcctggcg gcagtagcgc ggtgggtcca cctgacccca 120
tgccgaactc agaagtga aa cgccgtagcg ccgatggtag tgtggggtct ccccatgcga 180

FPAA4613PCT-seq1-000001

gagtagggaa	ctgccaggca	tcaaataaaa	cgaaaggctc	agtcgaaaga	ctgggccttt	240
cgttttatct	gttgtttgtc	ggtgaacgct	ctcctgagta	ggacaaatcc	gccgggagcg	300
gatttgaacg	ttgcgaagca	acggccccga	gggtggcggg	caggacgccc	gccataaact	360
gccaggcatc	aaattaagca	gaaggccatc	ctgacggatg	gcctttttgc	gtttctacaa	420
actct						425

<210> 16
 <211> 1084
 <212> DNA
 <213> artificial

<220>
 <223> LacI sequence: 3810-4890

<400> 16	
tgtgaaacca	gtaacgttat
acgatgtcgc	agagtatgcc
ggtgtctctt	atcagaccgt
	60
ttcccgctg	gtgaaccagg
ccagccacgt	ttctgcgaaa
acgcgggaaa	aagtggaagc
	120
ggcgatggcg	gagctgaatt
acattcccaa	ccgcgtggca
caacaactgg	cgggcaaaca
	180
gtcgttgctg	attggcgttg
ccacctccag	tctggccctg
cacgcgccgt	cgcaaattgt
	240
cgcggcgatt	aaatctcgcg
ccgatcaact	gggtgccagc
gtggtggtgt	cgatggtaga
	300
acgaagcggc	gtcgaagcct
gtaaagcggc	ggtgcacaat
cttctcgcgc	aacgcgtcag
	360
tgggctgata	attaactatc
cgctggatga	ccaggatgcc
attgctgtgg	aagctgcctg
	420
cactaatgtt	ccggcggtat
ttcttgatgt	ctctgaccag
acacccatca	acagtattat
	480
tttctcccat	gaagacggta
cgcgactggg	cgtggagcat
ctggtcgcgc	tgggtcacca
	540
gcaaatacgc	ctgttagcgg
gcccattaag	ttctgtctcg
gcgcgtctgc	gtctggctgg
	600
ctggcataaa	tatctcactc
gcaatcaa	tcagccgata
gcggaacggg	aaggcgactg
	660
gagtgccatg	tccggttttc
aacaaacat	gcaaatacgt
aatgagggca	tcgttccac
	720
tgcgatgctg	gttgccaacg
atcagatggc	gctgggcgca
atgcgcgcca	ttaccgagtc
	780
cgggctgcgc	gttggtgcgg
atatctcggt	agtgggatac
gacgataccg	aagacagctc
	840
atgttatatc	ccgccgttaa
ccaccatcaa	acaggatttt
cgctgtctgg	ggcaaaccag
	900
cgtggaccgc	ttgctgcaac
tctctcaggg	ccaggcggtg
aagggaatc	agctgttgcc
	960
cgtctcactg	gtgaaaagaa
aaaccaccct	ggcgcccaat
acgcaaaccg	cctctccccg
	1020
cgcgttggcc	gattcattaa
tgacagctgg	acgacaggtt
tcccactgg	aaagcgggca
	1080
gtga	1084

<210> 17
 <211> 861
 <212> DNA
 <213> artificial

<220>
 <223> bla (ampicillin resistance): 7085-7942

<400> 17

FPAA4613PCT-seq1-000001

ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat	60
agttgcctga ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc	120
cagtgtcgca atgataccgc gagaccacg ctcaccggct ccagatttat cagcaataaa	180
ccagccagcc ggaagggccg agcgcagaag tggtcctgca actttatccg cctccatcca	240
gtctattaat tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa	300
cgttgttgcc attgctgcag gcatcggtgt gtcacgctcg tcgtttggtg tggcttcatt	360
cagctccggt tcccaacgat caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc	420
ggtagctcc ttcggtcctc cgatcgttgt cagaagtaag ttggccgcag tgttatcact	480
catggttatg gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgcttttc	540
tgtgactggt gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg	600
ctcttgcccg gcgtcaatac gggataatac cgcgccacat agcagaactt taaaagtgtc	660
catcattgga aaacgttctt cggggcgaaa actctcaagg atcttaccgc tgttgagatc	720
cagttcgatg taaccactc gtgcaccaa ctgatcttca gcattcttta ctttcaccag	780
cgtttctggg tgagcaaaaa caggaaggca aaatgccgca aaaaagggaa taagggcgac	840
acggaaatgt tgaatactca t	861

<210> 18
 <211> 1
 <212> DNA
 <213> artificial

<220>
 <223> pBR322 origin 6323 sequence

<400> 18
 t

1

<210> 19
 <211> 8540
 <212> DNA
 <213> artificial

<220>
 <223> Complete clone sequence of dual cistron vector including heavy
 and light chainfab

<400> 19	
atccggatat agttcctcct ttcagcaaaa aaccctcaa gaccgttta gaggcccaa	60
ggggttatgc tagttattgc tcagcgggtg cagcagccaa ctcagcttcc tttcgggctt	120
tgtagcagc cggatctcag tgggtggtgt ggtggtgctc gagtgcggcc gcaagctttt	180
acagatgcgt tttatcacac gatttcggtt ccactttttt gtccactttg gtattgctcg	240
gtttgtgatt gacgttcgag atatacgttt gggtagccag tgaggaactc ggaacggtaa	300
ccacgctgct caggctgtac aggcccgatg actgcagaac agccgggaac gtatgcacac	360
cagaggtcag tgcgccgctg ttccaggaga cggtaacggg ttccggaaaa taatcttta	420
ccagacaacc cagagctgcc gtaccgccac tggtagactt ggaactcggg gccagcggga	480

FPAA4613PCT-seq1-000001

acacgctcgg	gcctttggtg	ctcgcgctgc	tcacggtgac	caggggtacct	tggcccaaaa	540
cgtcaaaata	ccagtgggac	gtgccatagt	aatacgggta	cttcgcgcag	taatacacgg	600
cggatatctt	cgcgcgcaga	ctgttcatct	gcagatatgc	cgtgctcttc	gaggtgtcca	660
gagagaacgt	aaagcgacgt	ttaaaatcgg	ccgcataggt	cggttcaccc	gtgtaggtat	720
taatccagcc	cacccattcc	agacctttac	ccggggcctg	acggaccacg	ttcataccat	780
aatgggtaaa	atcatagccc	gaggcagcac	acgacagacg	cagagagcca	cccggttgga	840
ccagaccacc	acccgattcg	accagttgga	cttcggccat	cgccggctgg	gcagcgagga	900
gcagcagacc	agcagcagct	gtcggcagca	ggtatttcat	atgtatatct	ccttcttaaa	960
gttaaacaaa	attatttcta	gaggggaatt	gttatccgct	cacaattccc	ctatagttag	1020
tcgtattaat	ttcgcgggat	cgagatcttt	atgacaactt	gacggctaca	tcattcactt	1080
tttcttcaca	accggcacgg	aactcgctcg	ggctggcccc	ggtgcatttt	ttaaataccc	1140
gcgagaaaata	gagttgatcg	tcaaaaccaa	cattgcgacc	gacgggtggcg	ataggcatcc	1200
gggtggtgct	caaaagcagc	ttcgcctggc	tgatacgttg	gtcctcgcgc	cagcttaaga	1260
cgctaataccc	taactgctgg	cggaaaagat	gtgacagacg	cgacggcgac	aagcaaacat	1320
gctgtgcgac	gctggcgata	tcaaaattgc	tgtctgccag	gtgatcgctg	atgtactgac	1380
aagcctcgcg	tacccgatta	tccatcggtg	gatggagcga	ctcgttaatc	gcttccatgc	1440
gccgcagtaa	caattgctca	agcagattta	tcgccagcag	ctccgaatag	cgcccttccc	1500
cttgcccggc	gttaatgatt	tgcccaaaca	ggtcgctgaa	atgcggctgg	tgcgcttcat	1560
ccgggcgaaa	gaaccccgtg	ttggcaaata	ttgacggcca	gttaagccat	tcatgccagt	1620
aggcgcgcg	acgaaagtaa	accactgggt	gataccattc	gcgagcctcc	ggatgacgac	1680
cgtagtgatg	aatctctcct	ggcgggaaca	gcaaaatata	acccggctcg	caaacaaatt	1740
ctcgtccctg	atttttcacc	accccctgac	cgcgaatggt	gagattgaga	atataacctt	1800
tcattcccag	cggtcggtcg	ataaaaaaat	cgagataacc	gttggcctca	atcggcgtta	1860
aacccgccac	cagatgggca	ttaaacgagt	atcccggcag	caggggatca	ttttgcgctt	1920
cagccatact	tttcatactc	ccgccattca	gagaagaaac	caattgtcca	tattgcatca	1980
gacattgccg	tactgcgtc	ttttactggc	tcttctcgct	aaccaaaccg	gtaaccccg	2040
ttattaaaag	cattctgtaa	caaagcgggg	caaagccat	gacaaaaacg	cgtaacaaaa	2100
gtgtctataa	tcacggcaga	aaagtccaca	ttgattattt	gcacggcgtc	acactttgct	2160
atgccatagc	atttttatcc	ataagattag	cggatcctac	ctgacgcttt	ttatcgcaac	2220
tctctactgt	ttctccatac	ccgttttttt	gggctagaaa	taattttggt	taactttaag	2280
aaggagatat	acatatgaaa	tacctgctgc	cgacagctgc	tgctggtctg	ctgctcctcg	2340
ctgcccagcc	ggcgatggcc	gacattcaac	tgacgcaaag	tccgagcagc	ctgagcgcat	2400
ccgtgggcga	ccgtgtgacg	attacctgtt	ccgcaagcca	agacatctct	aactatctga	2460
attggtacca	gcaaaaaccg	ggcaaggcac	cgaaagtcct	gatttatattt	accagctctc	2520

FPAA4613PCT-seq1-000001

tgattccgg	cgttccgtca	cgttttagcg	gctctggtag	tggcaccgat	ttcaccctga	2580
cgatcagttc	cctgcagccg	gaagactttg	ctacgtatta	ctgccagcaa	tacagcaccg	2640
tgccgtggac	gttcgggtcag	ggcaccaagg	ttgaaattaa	acgtacgggt	gcggccccgt	2700
ctgtctttat	cttcccgcg	agtgatgaac	agctgaaatc	gggtaccgca	agcgtggttt	2760
gtctgtgtaa	caatttctat	ccgcgcgaag	caaaggtcca	gtggaaagtg	gacaacgctc	2820
tgagtcagg	caattcacia	gaatcgggtga	ccgaacaaga	tagcaaggac	tctacgtaca	2880
gtctgtcatc	gaccctgacg	ctgtccaaag	cggattatga	aaaacacaag	gtttacgcct	2940
gcgaagtcac	ccatcaaggt	ctgtcgtctc	cggttacca	gagtttcaat	cgtggcgaat	3000
gttaaaagct	tggctgtttt	ggcggatgag	agaagatttt	cagcctgata	cagattaaat	3060
cagaacgcag	aagcggctctg	ataaaacaga	atttgcctgg	cggcagtagc	gcggtggtcc	3120
cacctgaccc	catgccgaac	tcagaagtga	aacgccgtag	cgccgatggt	agtgtggggt	3180
ctcccatgc	gagagtaggg	aactgccagg	catcaaataa	aacgaaaggc	tcagtcgaaa	3240
gactgggcct	ttcgttttat	ctgttggttg	tcggtgaacg	ctctcctgag	taggacaaat	3300
ccgccgggag	cggatttgaa	cgttgccaag	caacggcccc	gaggggtggc	ggcaggacgc	3360
ccgccataaa	ctgccaggca	tcaaattaag	cagaaggcca	tcctgacgga	tggccttttt	3420
gcgtttctac	aaactcttag	atctcgatcc	tctacgccgg	acgcatcgtg	gccggcatca	3480
ccggcgccac	aggtgcgggt	gctggcgctt	atatcgccga	catcaccgat	ggggaagatc	3540
gggctcgcca	cttcgggctc	atgagcgctt	gtttcggcgt	gggtatggtg	gcaggccccg	3600
tggccggggg	actgttgggc	gccatctcct	tgcatgcacc	attccttgcg	gcggcggtgc	3660
tcaacggcct	caacctacta	ctgggctgct	tcctaattga	ggagtcgcat	aaggagagagc	3720
gtcgagatcc	cggacaccat	cgaatggcgc	aaaacctttc	gcggtatggc	atgatagcgc	3780
ccggaagaga	gtcaattcag	ggtggtgaat	gtgaaaccag	taacgttata	cgatgtcgca	3840
gagtatgccg	gtgtctctta	tcagaccgtt	tcgcgcgtgg	tgaaccaggc	cagccacgtt	3900
tctgcgaaaa	cgcgggaaaa	agtggaagcg	gcgatggcgg	agctgaatta	cattcccaac	3960
cgctggcac	aacaactggc	gggcaaacag	tcgttgctga	ttggcgttgc	cacctccagt	4020
ctggccctgc	acgcgccgtc	gcaaattgtc	gcggcgatta	aatctcgcg	cgatcaactg	4080
ggtgccagcg	tgggtggtgc	gatggtagaa	cgaagcggcg	tcgaagcctg	taaagcggcg	4140
gtgcacaatc	ttctcgcgca	acgcgtcagt	gggctgatca	ttactatcc	gctggatgac	4200
caggatgcca	ttgctgtgga	agctgcctgc	actaatgttc	cggcgttatt	tcttgatgtc	4260
tctgaccaga	caccatcaa	cagtattatt	ttctcccatg	aagacgggtac	gcgactgggc	4320
gtggagcatc	tggtcgcatt	gggtcaccag	caaatcgcg	tgttagcggg	cccattaagt	4380
tctgtctcgg	cgctctcgc	tctggctggc	tggcataaat	atctcactcg	caatcaaatt	4440
cagccgatag	cggaacggga	aggcgactgg	agtgccatgt	ccggttttca	acaacccatg	4500
caaagtctga	atgagggcat	cgttcccact	gcgatgctgg	ttgccaacga	tcagatggcg	4560

FPAA4613PCT-seq1-000001

ctgggcgcaa	tgcgcgccat	taccgagtcc	gggctgcgcg	ttggtgcgga	tatctcggta	4620
gtgggatacg	acgataccga	agacagctca	tgttatatcc	cgccgttaac	caccatcaaa	4680
caggattttc	gcctgctggg	gcaaaccagc	gtggaccgct	tgctgcaact	ctctcagggc	4740
caggcggtga	agggaatca	gctgttgccc	gtctcactgg	tgaaaagaaa	aaccaccctg	4800
gcgcccata	cgcaaaccgc	ctctccccgc	gcgttgggcg	attcattaat	gcagctggca	4860
cgacaggttt	cccgaactga	aagcgggcag	tgagcgcaac	gcaattaatg	taagttagct	4920
cactcattag	gcaccgggat	ctcgaccgat	gcccttgaga	gccttcaacc	cagtcagctc	4980
cttccggtgg	gcgcggggca	tgactatcgt	cgccgcactt	atgactgtct	tctttatcat	5040
gcaactcgta	ggacaggtag	cggcagcgct	ctgggtcatt	ttcggcgagg	accgctttcg	5100
ctggagcgcg	acgatgatcg	gcctgtcgct	tgcggtattc	ggaatcttgc	acgccctcgc	5160
tcaagccttc	gtcactggtc	ccgccaccaa	acgtttcggc	gagaagcagg	ccattatcgc	5220
cggcatggcg	gccccacggg	tgcgcatgat	cgtgctcctg	tcgttgagga	ccgggctagg	5280
ctggcggggt	tgcttactg	gtagcagaa	tgaatcaccg	atacgcgagc	gaacgtgaag	5340
cgactgctgc	tgcaaaacgt	ctgcgacctg	agcaacaaca	tgaatggtct	tcggtttccg	5400
tgtttcgtaa	agtctggaaa	cgcggaagtc	agcgccctgc	accattatgt	tccggatctg	5460
catcgcagga	tgctgctggc	taccctgtgg	aacacctaca	tctgtattaa	cgaagcgctg	5520
gcattgaccc	tgagtgattt	ttctctggtc	ccgccgcata	cataccgcca	gttgtttacc	5580
ctcacaacgt	tccagtaacc	gggcatgttc	atcatcagta	accggtatcg	tgagcatcct	5640
ctctcgtttc	atcggtatca	ttacccccat	gaacagaaat	cccccttaca	cggaggcatc	5700
agtgacaaa	caggaaaaaa	ccgcccttaa	catggcccg	tttatcagaa	gccagacatt	5760
aacgcttctg	gagaaactca	acgagctgga	cgcggtatga	caggcagaca	tctgtgaatc	5820
gcttcacgac	cacgctgatg	agctttaccg	cagctgcctc	gcgcgtttcg	gtgatgacgg	5880
tgaaaacctc	tgacacatgc	agctcccggg	gacggtcaca	gcttgtctgt	aagcggatgc	5940
cgggagcaga	caagcccgtc	agggcgcgtc	agcgggtgtt	ggcgggtgtc	ggggcgcgagc	6000
catgaccag	tcacgtagcg	atagcggagt	gtatactggc	ttaactatgc	ggcatcagag	6060
cagattgtac	tgagagtga	ccatatatgc	ggtgtgaaat	accgcacaga	tgcgtaagga	6120
gaaaataccg	catcaggcgc	tcttccgctt	cctcgctcac	tgactcgctg	cgctcggctg	6180
ttcggctgcg	gcgagcggta	tcagctcact	caaaggcgg	aatacgggta	tccacagaat	6240
caggggataa	cgcaggaaag	aacatgtgag	caaaaggcca	gcaaaaggcc	aggaaccgta	6300
aaaaggccgc	gttgctggcg	ttttccata	ggctccgccc	ccctgacgag	catcacaaaa	6360
atcgacgctc	aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	caggcgtttc	6420
cccctggaag	ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	ggatacctgt	6480
ccgcctttct	cccttcggga	agcgtggcgc	tttctcatag	ctcacgctgt	aggtatctca	6540
gttcgggtgta	ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaaccccc	gttcagcccc	6600

FPAA4613PCT-seq1-000001

accgctgcg	cttatccggt	aactatcgtc	ttgagtc	cccggt	cacgacttat	6660
cgccactgg	agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	ggcgggtgcta	6720
cagagttctt	gaagtgggtg	cctaactacg	gctacactag	aaggacagta	tttgggtatct	6780
gcgctctgct	gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	tccggcaaac	6840
aaaccaccgc	tggtagcgg	ggttttttt	tttgcaagca	gcagattacg	cgcagaaaaa	6900
aaggatctca	agaagatcct	ttgatctttt	ctacgggggtc	tgacgctcag	tggaacgaaa	6960
actcacgtta	agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	7020
taaattaaaa	atgaagtttt	aatcaatct	aaagtatata	tgagtaaact	tgggtctgaca	7080
gttaccaatg	cttaatcagt	gaggcaccta	tctcagcgat	ctgtctattt	cgttcatcca	7140
tagttgcctg	actccccgtc	gtgtagataa	ctacgatacg	ggagggctta	ccatctggcc	7200
ccagtgtctg	aatgataccg	cgagaccac	gctcaccggc	tccagattta	tcagcaataa	7260
accagccagc	cgaagggcc	gagcgcagaa	gtggctctgc	aactttatcc	gcctccatcc	7320
agtctattaa	ttgttgccgg	gaagctagag	taagtagttc	gccagttaat	agtttgcgca	7380
acgttggtgc	cattgctgca	ggcatcgtgg	tgtcacgctc	gtcgtttgg	atggcttcat	7440
tcagctccgg	ttcccaacga	tcaaggcgag	ttacatgatc	ccccatgttg	tgcaaaaaag	7500
cggtagctc	cttcggtcct	ccgatcgttg	tcagaagtaa	gttgccgca	gtgttatcac	7560
tcatggttat	ggcagcactg	cataattctc	ttactgtcat	gccatccgta	agatgctttt	7620
ctgtgactgg	tgagtactca	accaagtc	tctgagaata	gtgtatgcgg	cgaccgagtt	7680
gctcttgccc	ggcgtcaata	cgggataata	ccgcgccaca	tagcagaact	ttaaaagtgc	7740
tcacattgg	aaaacgttct	tcggggcgaa	aactctcaag	gatcttaccg	ctgttgagat	7800
ccagttcgat	gtaaccact	cgtgcacca	actgatcttc	agcatctttt	actttcacca	7860
gcgtttctgg	gtgagcaaaa	acaggaaggc	aaaatgccgc	aaaaaaggga	ataagggcga	7920
cacggaaatg	ttgaatactc	atactcttcc	tttttcaata	ttattgaagc	atttatcagg	7980
gttattgtct	catgagcgga	tacatatttg	aatgtattta	gaaaaataaa	caaatagggg	8040
ttccgcgcac	atttccccga	aaagtgccac	ctgaaattgt	aaacgttaat	attttgttaa	8100
aattcgcgtt	aaatttttgt	taaatcagct	cattttttta	ccaataggcc	gaaatcggca	8160
aaatccctta	taaatcaaaa	gaatagaccg	agatagggtt	gagtgttggt	ccagtttgga	8220
acaagagtcc	actattaaag	aacgtggact	ccaacgtcaa	agggcgaaaa	accgtctatc	8280
agggcgatgg	cccactacgt	gaaccatcac	cctaatacaag	ttttttgggg	tcgaggtgcc	8340
gtaaagcact	aaatcggaac	cctaaaggga	gccccgatt	tagagcttga	cggggaaagc	8400
cggcgaacgt	ggcgagaaag	gaagggaaga	aagcgaaagg	agcgggcgct	agggcgctgg	8460
caagtgtagc	ggtcacgctg	cgcgtaacca	ccacaccgc	cgcgcttaat	gcgccgctac	8520
agggcgctgc	ccattcgcca					8540