<table>
<thead>
<tr>
<th>(54)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCSF fusion protein systems suitable for high expression of peptides</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(51)</th>
<th>International Patent Classification(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C07K 14/435 8BHEP C12N (2006.01) 15/62 C12N 15/62 (2006.01) 20060101AL 2009092 C07K 14/435 4BHEP 20060101AF 2009052 PCT/IN2008/000599</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(21)</th>
<th>Application No:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008327513</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(22)</th>
<th>Application Date:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008.09.22</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(87)</th>
<th>WIPO No:</th>
</tr>
</thead>
<tbody>
<tr>
<td>WO09/066320</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(30)</th>
<th>Priority Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1848/MUM/2007 2007.09.21 IN</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(43)</th>
<th>Publication Date:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009.05.28</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(71)</th>
<th>Applicant(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cadila Healthcare Limited</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(72)</th>
<th>Inventor(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saraswat, Vibhor; Patel, Pankaj R.; Singh, Arun Kumar; Bandyopadhyay, Sanjay; Mendiretta, Sanjeev Kumar</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(74)</th>
<th>Agent/Attorney</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesica Perrin, 58 Rundle Street (cnr The Parade West), Kent Town, SA, 5067</td>
<td></td>
</tr>
</tbody>
</table>
(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)
(19) World Intellectual Property Organization
International Bureau
(43) International Publication Date
28 May 2009 (28.05.2009)
(41) International Publication Number
WO 2009/063620 A3

(51) International Patent Classification:
C12N 15/82 (2006.01)
(21) International Application Number:
PCT/IN2008/000599
(22) International Filing Date:
22 September 2008 (22.09.2008)
(25) Filing Language:
English
(26) Publication Language:
English
(30) Priority Data:
(71) Applicant (for all designated States except US):
CAHAL HEATHCARE LIMITED [IN/IN], Zydus Research Centre, Zydus Tower, Satellite Cross Roads, Ahmedabad 380 015, Gujarat (IN).
(72) Inventors:
Sanjeev, Kumar [IN/IN]; Cadila Healthcare Limited, Zydus Research Centre, Zydus Tower, Satellite Cross Roads, Ahmedabad 380 015, Gujarat (IN);
PATEL, Pankaj, R. [IN/IN]; Cadila Healthcare Limited, Zydus Tower, Satellite Cross Roads, Ahmedabad 380 015, Gujarat (IN);
BANDIVADHYAY, Sanjay [IN/IN]; Cadila Healthcare Limited, Zydus Tower, Satellite Cross Roads, Ahmedabad 380 015, Gujarat (IN);
SARASWAT, Vibhav [IN/IN]; Cadila Healthcare Limited, Zydus Research Centre, Zydus Tower, Satellite Cross Roads, Ahmedabad 380 015, Gujarat (IN);
SINGH, Arun, Kumar [IN/IN]; Cadila Healthcare Limited, Zydus Research Centre, Zydus Tower, Satellite Cross Roads, Ahmedabad 380 015, Gujarat (IN).
(74) Agent:

(55) Title: GCFS Fusion Protein Systems Suitable for High Expression of Peptides

(57) Abstract: The present invention discloses an improved process for the production of desired recombinant peptides from bacterial cells by using r-GCFS as a novel fusion partner for their high level expression in these cells. The invention further provides an expression system comprising the fusion protein wherein the GCFS is operatively linked to the peptide of interest via an enzymatic or chemical cleavage site which can be used to separate the fusion partner from the said peptide.
FUSION PROTEIN SYSTEMS SUITABLE FOR HIGH EXPRESSION OF PEPTIDES

Field of the invention

The present invention relates to the field of production of recombinant peptides. More specifically, the present invention discloses an improved method for the production of desired recombinant peptides from bacterial cells by using G-CSF as a novel fusion partner for their high level expression in these cells. The invention further provides an expression system comprising the fusion protein wherein the G-CSF is operatively linked to the peptide of interest via an enzymatic or chemical cleavage site which can be used to separate the fusion partner from the said peptide.

Background of the invention

Peptides are heteropolymers of amino acids that are linked via their carboxyl and amino groups by amide bonds. Since the above definition holds true for proteins as well, they are often differentiated from proteins based upon their chain length, being usually described as those heteropolymers that are ranging in chain length from two to a few to several dozen amino acid residues.

Some examples of these peptides that vary vastly in their molecular weight and their functions are described here [Geysen, H. M. et al., J. Immunol. Methods, 102:259-274 (1987); Milich, D. R. Semin. Immunol., 2(5):307-315 (1990); Cochran, A. G. Chemistry & Biology, 7:R85-R94 (2000)]. The active peptide, insulin (51 residues, 5773 Da), formed after processing of the larger but inactive peptide, pro-insulin (86 residues), plays a critical role in glucose metabolism in the body and lack of it leads to type I diabetes. Parathyroid hormone (84 residues) is the most important endocrine regulator of the levels of calcium and phosphorus in the blood. A much smaller version of the parathyroid hormone, teriparatide (34 residues), is sold by Eli Lilly for the treatment of osteoporosis. Calcitonin (32 residues) also plays a critical role in calcium metabolism and has opposite effects with respect to the parathyroid hormone. It is also used in the treatment for osteoporosis. Angiotensin I, II, III and IV having 10, 8, 7 and 6 residues respectively, have diverse roles as endocrine, paracrine, and intracrine hormones, causing vasoconstriction, increasing blood pressure and causing release of aldosterone from the adrenal cortex. The non-caloric sweetener, aspartame (a dipeptide of aspartic acid and ~esterified phenylalanine) is used as a sugar substitute. These examples illustrate the vast diversity of functions, chain length and utility amongst
peptides. Their role as mediators of key biological functions e.g., as hormones, enzyme substrates or inhibitors, neurotransmitters, immunomodulators and, antiviral agents make them particularly attractive as therapeutic agents. More than 40 therapeutic peptides are available in the world market today and more than 400 peptides are in advanced pre-clinical phases of drug development worldwide [Parnar, H. Therapeutic Peptides in Europe: Finding the Opportunities, Frost and Sullivan Market Insight, November 26, 2004]. Therefore the therapeutic application of peptides has an enormous potential.

Originally peptides of therapeutic significance used to be isolated from biological tissues. For example, insulin was produced from the ox pancreas [Collip, J. J. Physiol., 66:416-430 (1928)], and Calcitonin from the ultimobronchial glands of fish [(Parkes, C.O. et al., Fed. Proc., 28:413 (1969)]. The tissue origin of these peptides made the isolation methods difficult and cumbersome, yielded non-ideal product purity, carried the risk of transmitting infections and generally affected the commercial scalability, ultimately limiting the commercialization of peptides as therapeutics.

Chemical method is one of the solutions to the above problem of commercial production - at least for the production of small and medium sized peptides ranging from about 5 to 80 residues [Kimmerlin, T. and Seebach, D. J. Pept. Res., 65(2):229-260 (2005)]. But this method also has many disadvantages that make it inefficient in cost, such as, the possibility of racemization, poor solubility of protected peptide fragments, limitation on the length of peptide and serial side reactions. The undesirable side reactions associated with this method decrease the yield and render necessary, difficult and lengthy purification procedures. Therefore, even though chemical synthesis is the most mature technology for peptide synthesis today, it is fraught with problems that add to the overall cost of their production signifying the need for developing other distinctly different methods of their production.

In principle a large number of the above problems can be solved by employing the recombinant DNA technology that allows the selection, amplification and manipulation of expression of endogenous and foreign genes in microbial cells. This technology is better than the chemical technology in that it naturally produces non-racemic and fully correct a.a. sequences, as nature has evolved living organisms to do so. Further this technology also allows more environmentally friendly processes of production and purification. However, this technology is limited by the fact that in general, it is not possible to get high levels of expression of peptides in a microbial

A common strategy to overcome the problem of degradation of peptides in the expression host is to express them as fusion proteins in conjunction with another larger peptide or protein, which acts as a fusion partner. After expression, the peptide and the fusion partner are separated from each other by chemical or proteolytic cleavage at a site which was pre-designed into the fusion product. In this regard many proteins have been used as fusion partners that produce the fusion product in both soluble and insoluble form inside the cell. Ray et al., [Ray, M. V. L., et al., BioTechnology, 11:64-70 (1993)] describe the use of glutathione-S-transferase (GST) as a fusion partner to get soluble intracellular expression in E. coli of salmon calcitonin which was cleaved away from the fusion protein with cyanogen bromide. Dykes et al., [Dykes, C. W. et al., Eur. J. Biochem., 174:411-416 (1988)] describes the soluble intracellular expression in E. coli of a fusion protein consisting of a human atrial natriuretic peptide and chloramphenicol acetyltransferase, where the fusion protein was either proteolytically cleaved by thrombin or enterokinase or chemically cleaved with 2-(2-nitrophenylsulphenyl)-methyl-3'-bromoindolene to release the peptide of interest. Maltose-binding protein and thioredoxin have also been used as solubilizing fusion partners [Banex, B. Curr. Opin. Biotechnol., 10:411-421 (1999)].

Another example is that of US5223404 which describes the use of ompA as a fusion partner with PTH (1-34) in order to get the expression of soluble protein in periplasmic space.


In fact, many major biotech companies have been using fusion proteins to express therapeutic peptides in the form of inclusion bodies. In EP 0055945, Genentech reports the use of TrpE and B-galactosidase (lacZ) as a fusion partner to express proinsulin, which is later cleaved off using cyanogen bromide. In EP0211299 A, Hoechst AG, describes the use of D peptide of E. coli Trp gene as a fusion partner to express fusion peptides proinsulin and hirudin. In US 5670340, Suntory Ltd, Osaka, Japan, reports the use of a fragment of beta galactosidase as a fusion partner along with
a leader peptide linked with human calcitonin to express the fusion product as inclusion bodies, where calcitonin is cleaved off using V8 protease after expression. In US 6506647, Mogam Biotechnology Institute, exemplified the production of human PTH as inclusion bodies by fusing it to a phosphoribulokinase gene fragment of Rhodobacter sphaeroides or its mutated gene as a fusion partner where the fusion product is also containing a urokinase-specific cleavage site. [Wingender, E. et al., J. Biol. Chem. 264(8):4367-4373 (1989)] describes the use of variable sizes of cro-beta galactosidase as a fusion partner with PTH which is later cleaved off using acid hydrolysis. Despite the above described successes associated with the use of fusion partners for the production of therapeutic peptides, literature sites a number of problems that still exist with this technology. Some of the examples below illustrate these problems and support the need for inventing better fusion partners.

Not all fusion partners form stable inclusion bodies. Shen [Shen, S-H. PNAS, 81:4627-4631 (1984)], describes the bacterial expression of proinsulin using an 80 a.a. long, amino-terminus fragment of ${\beta}$-galactosidase as a fusion partner. This fusion partner resulted in extremely unstable expression of the fusion product not detectable in SDS-PAGE. Stable expression in the form of inclusion bodies was obtained only when two tandem copies of pro-insulin gene were used in conjunction with the above described fusion partner. Therefore an ideal size of the fusion partner appears to be an important criteria for the formation of inclusion bodies.

While often the formation of inclusion bodies is a consequence of high expression rates [Kane, J. F. and Hartley, D. L. Trends Biotechnol. 6:95-101 (1988); Fahmert, B. et al., Adv. Biochem Engg./Biotechnol. 89:93-142 (2004)], the literature teaches us that such a correlation is not always correct. Mac [Mac, T-T. Towards solid-state NMR spectroscopic studies of the ETBR/ET-1 complex. Ph.D. Thesis (2007), Freie Universität Berlin] used thioredoxin as a fusion partner for the expression of BigET-1 peptide with the intention that the small size of thioredoxin (12 kDa) and its solubility in E.coli will ensure a soluble fusion product. In contrast to this expectation, the fusion product formed inclusion bodies but at very poor expression levels (45 mg/L). When GST was used a fusion protein for the expression of ET-1 in the form of inclusion bodies, expression levels of the fusion product increased but still remained in mg/L levels only. Use of beta-galactosidase fragments as a fusion partner also led to the expression of ET-1 in the form of inclusion bodies but the expression levels remained low, yielding only few mg/L quantities of ET-1 at the end of the process [Ohashi, H. et

The vast numbers of above references indicate that the production of small peptides from bacteria has been problematic for a variety of reasons. While the problem of proteolysis of peptides in microbial cells is usually taken care of quite successfully by expressing them as fusion proteins but making these processes commercially feasible has been fraught with numerous challenges. While expressing peptides as insoluble fusion proteins is a preferred method of commercial production of peptides by recombinant methods, the choice of a fusion partner is not straight forward. The previously reported fusion proteins do not always behave in a predictable fashion as far as formation of inclusion bodies and/or high and stable expression is concerned. Further, many of them are either not very easily available through commercial sources or their use for commercial production of a desired peptide has not been fully established. Hence there has been a long felt need in the art for a suitable expression system that comprises a fusion partner that would consistently give stable and high expression with a variety of peptides and at a low cost. With this view, the object of the present invention is to develop an expression system comprising such a fusion partner and establish a production method for the production of peptides at a high yield and production efficiency which is equal to or better than those of the existing processes.

The present invention discloses a novel expression system that utilizes a novel fusion partner, G-CSF, which can be consistently used for the high expression of peptides in the form of inclusion bodies. The peptides of interest that can be expressed with this fusion partner, vary not only in their amino acid content but also in their chain length, and are separated from the fusion partner after cleaving the cleavage site, which is pre-designed into the fusion product. The fusion peptide obtained as such, may then be purified using the standard downstream purification methods. The process of production utilizing such an expression system was found to be highly scalable and facilitated in the stable high expression for three peptides that it was tested with.

Description of Figures:

Fig 1 depicts the expression of three fusion products namely, proinsulin, angiotensin and PTH in both uninduced and IPTG-induced cultures of E. coli transformed with the expression vectors encoding the said fusion products.
Objects of Invention:

The present invention, in one aspect, provides a process for the production of peptides comprising the use of GCSF or its suitable variants as a fusion partner. In an embodiment is provided a process for the expression of peptides as a fusion product using a novel fusion partner, G-CSF or its suitable variants. In another aspect of the invention is provided a novel fusion partner, G-CSF, for the expression of peptides as fusion products.

In another aspect it provides novel fusion products comprising G-CSF linked to the peptides of interest through a cleavage site which can optionally, where appropriate, be through suitable linker that acts as the cleavage site for chemical and/or enzymatic cleavage agents. In another aspect it provides nucleotide constructs encoding the above novel fusion products. In yet another aspect it provides the expression vectors comprising the above nucleotide constructs. In yet another aspect it provides the expression systems comprising the above expression vectors.

In another aspect, it provides methods of construction of such expression vectors encoding such fusion product constructs.

In yet another aspect, it provides the methods of construction of expression systems for the high level expression of peptides as fusion products.

In a still further aspect it provides the use of such expression systems comprising the novel fusion product constructs for the high level expression of peptides.

In a specific embodiment is provided method of expressing desired peptides using the fusion partner and expression system of the present invention. Suitable peptides of interest may include but are not limited to peptides of about 10-90 a.a. residues in length. Examples of such peptides include but are not limited to PTH and its analogues, Insulin and its analogues, Calcitonin and its analogues, peptides of the Angiotensin class, and the like, and such peptides including their short chain peptide variants which act similarly.

Summary

The present invention provides a process for producing recombinant peptides of a large range of sizes & therapeutic classes by employing a novel fusion partner, G-CSF. Preferably, these peptides are linked to G-CSF at its C terminus via a chemical or enzymatic cleavage site. This novel expression system utilizes the biosynthetic machinery of the E.coli host cell and provides expression of such fusion peptides
preferably in the form of insoluble aggregates of the fusion product in the transformed host cells. The inclusion bodies can be easily isolated from the host cells, and from which, the peptide purified using standard protein purification methodology after cleaving it from the fusion partner, G-CSF by known techniques.

**Description of the Invention:**

The present invention provides an expression system for the production of peptides in the host cell, where the desired peptide is operatively linked with the fusion partner. The linking or connection is done through an creation of an appropriate cleavage site, thereby forming a complete fusion product. Preferably, the fusion products are insoluble aggregates.

**Fusion partner:**

The novel expression systems described in this invention comprise of a fusion partner, G-CSF (Genbank Accession number, gi:27437050), which is fused with a suitable peptide to form the fusion protein. Suitable mutants and variants of G-CSF are also contemplated & included as fusion partners as long as they perform the same function of G-CSF. A preferred advantage of the fusion protein is that it provides high level of expression of the said peptide. More specifically, the fusion partner of the present invention is methionyl human GCSF (Sequence ID No 1).

**Fused Peptide:**

The fused peptide is the peptide which is expressed along with the fusion partner in the form of a single fusion product construct. The fused peptide may have several forms. It includes a natural peptide of interest and/or any of its mutants or variants, and/or may include new peptides not found in nature and which is completely designed by man.

Preferably, the peptides are selected from those having from about 10-90 amino acid length. In one embodiment, a fused peptide may be selected from, calcitonin, insulin, angiotensin, tissue plasminogen activator, growth hormones, growth factors, growth hormone releasing factors, erythropoietin, interferons, interleukins, oxytocin, vasopressin, ACTH, collagen binding protein, parathyroid hormones, glucagon like peptides, glucagon, proinsulin, tumor necrosis factors, substance P, brain naturetic peptide, individual heavy and light antibody chains, fluzon, octreotide, somatostatin and the variants including mutants, synthetic analogs and mimetics of these peptides. In a preferred embodiment of the present invention are provided fused peptides that include, PTH, angiotensin and proinsulin. As examples of preferred
embodiments are provided fused peptides which comprise the amino acid sequence of PTH (1-34) (Sequence ID Nos. 6 and 7), Angiotension (Seq. ID Nos. 16 and 17) and Proinsulin (Seq. ID Nos. 12 and 13).

Fusion product or Fusion product construct-

The fusion product construct of the present invention is represented by the formula

Fusion partner ——— CS ——— Fusion peptide

wherein a fusion peptide of interest is a physiologically active peptide, CS is a cleavage site either present within the fusion partner or the fusion peptide amino acid (a.a.) sequences or is a suitable linker of a suitable length of a.a. residues, and the fusion partner is GCSF or its variants.

The preferred fusion product or the fusion product construct contemplated in this invention may have the following formulae:

N-terminal Fusion Partner ——— CS ——— Peptide of interest C-terminal
N-terminal Peptide of interest ——— CS ——— Fusion Partner C-terminal

In a more preferred embodiment of the present invention a fusion product of the following formula is contemplated:

N-terminal Peptide of interest ——— CS ——— Fusion Partner C-terminal

wherein CS represents an enzymatic or chemical cleavage site. As preferred embodiments of the present invention fusion proteins of the following formulae are contemplated:

G-CSF ——— Entero kinase cleavage site ——— PTH (G-CSF-PTH), Sequence ID Nos. 6 and 7
G-CSF ——— CNBr cleavage site ——— Proinsulin (G-CSF-Proinsulin), Seq. ID Nos. 12 and 13
G-CSF ——— Entero kinase cleavage site ——— Angiotensin (G-CSF-Angiotensin), Seq. ID Nos. 16 and 17

Cleavage Site-

The cleavage sites contemplated in the above invention comprise chemical cleavage sites such as those cleavable by cyanogen bromide or 2-(2-nitrophenylsulphonyl)-methyl-3-bromoindolene, BNPS-skatole (3-bromo-3-methyl-2-(2-nitrophenyl)thiol-3H-indole), N-bromosuccinimide, O-iodosobenzoic acid, HBr/DMSO, NTCB (2-Nitro-5-thiocyanobenzoate), Sodium metal in liquid ammonia,
Hydroxylamine, dilute acid for acid hydrolysis, and the like; or enzymatic cleavage sites such as those recognized by Enterokinase, Trypsin, Chymotrypsin, Elastase, Pepsin, Papain, Subtilisin, Thermolysin, V8 Protease, Endoprotease Arg C (submaxillaris protease), Clostridin, Thrombin, Collagenase, Lysobacter enzymogenes (Lys C), Mycobacterium A1-1 Protease, Factor Xa, and the like. The cleavage site is inserted by using standard molecular biology techniques such as the polymerase chain reaction where the reaction is carried out using one or both of the above genes (peptide and fusion partner) as templates and mutated oligos capable of introducing the said sites as primers of the reaction. Alternatively, the cleavage site may also exist naturally within the peptide or fusion partner.

Expression Vector-

- An expression vector is a suitable plasmid vector that comprises an expression cassette encoding the gene for such a fusion product or fusion product construct as the one described above wherein the expression cassette is operably linked to a suitable promoter and other suitable regulatory elements functional in the expression host into which the expression vector has been transformed or transfected for the expression of such fusion product. Examples of expression vectors include those vectors which can express the fusion product in E. coli. Examples of such expression vectors may be pET27, pET11, pET3, pET32 and the like which use the T7 promoter system, pBAD or pARA (Invitrogen) series of vectors based on the arabinose operon and AraC activator. Other examples include the vectors containing lambda PL promoters such as pGWI7, a pBR322 based vector containing the lambda PL promoter, PL vector series (Invitrogen) inducible by temperature shift to 42 deg C, or other plasmid vectors designed to contain the lac, lacUV5, tac, trc (IPTG inducible), trp(Tryptophan), araBAD (Arabinose), phoA (phosphate starvation), est-1 (glucose starvation), cspA (cold inducible), SP6, T3 promoter to give expression in E.coli.

Expression System-

The expression of a foreign protein by interaction of the recombinant expression vectors encoding the gene for the said protein with the biosynthetic machinery of the host cells is a well known technique for the biochemical synthesis of proteins. By the term, “Expression system”, is meant an expression vector that encodes the gene for such a fusion product or fusion product construct as the one described above, and the transformed E.coli host cell, in which the fusion product is produced.

In the present invention, the E.coli expression system utilized for the expression
of the peptide of interest comprises an expression vector, in which the desired peptide is operatively linked with the fusion partner, G-CSF, in any order, with either the peptide, or the fusion partner, at the C-terminus, with a cleavage site in between the two units. Preferentially, the peptide is encoded at the C-terminus of G-CSF. An example of the preferred expression vector is pET27B-GCSF-PTH, which encodes for a fusion product, G-CSF-PTH (1-34). E.coli DH5α transformed with the above vector have been deposited as pET-GFL-PTH1-34 cells, at Microbial Type Culture Collection and Gene Bank, Institute of Microbial Technology, India, under the accession number MTCC 5425. Another preferred example of such an expression vector is pET27B-GCSF-Proinsulin, which encodes for a fusion product, G-CSF-Proinsulin. E.coli DH5α transformed with the above vector have been deposited as pET-GFL-Proinsulin cells, at Microbial Type Culture Collection and Gene Bank, Institute of Microbial Technology, India, under the accession number MTCC 5424.

An E.coli cell which has been produced after transformation by an expression vector such as the one described above, and that expresses the peptide of interest as a fusion product, is also a part of the present invention. In a preferred embodiment of this aspect of the invention, the transformed cell has been obtained by transforming BL21 (DE3) strain of E.coli.

The expression system described in this invention may be used for the expression of any peptide in form of a fusion product with G-CSF as a fusion partner. Preferentially, the length of the peptide varies from 10-86 a.a. residues.

Production of Peptide in the host cell:

For the production of the peptide in the form of a fusion product the host cells are transformed with an expression vector encoding the previously contemplated fusion product construct. The expression vector may be derived from any base vector that is compatible with the host of interest by cloning the DNA segment encoding the fusion product into the base vector under the control of an appropriate host-compatible promoter. Besides the promoter, the base vector may also consist of other host-compatible regulatory elements such as those involved in the regulation of transcription and translation of the gene product. In general the choice of the base vector and the regulatory elements therein is based upon the compatibility of the vector with the host to obtain high expression of the fusion product. The host cell contemplated in the invention may be any microbial host cell that tends to form inclusion bodies upon high
expression of protein inside the cell. In one preferred embodiment of the invention, the
microbial host cell is E. coli.

Typically, the recombinant gene inserted into an appropriate base vector which
is used to transform the host cells leads to the expression inside the transformed host
cell of the fusion product as either a soluble product or as an insoluble product. In the
present invention the preferred form of expression of the fusion product is in the form
of insoluble product that results in the formation of inclusion bodies. The peptides can
be selected from those described above and elsewhere in the specification. In particular,
the fusion product comprises a single copy of the peptide of interest that is selected
from the group comprising of PTH (1-34), Sequence ID Nos. 6 and 7, Angiotension
Seq. ID Nos. 16 and 17 and Proinsulin Seq. ID Nos. 12 and 13 and that is expressed as
inclusion bodies in an E. coli host cell.

One of the examples of high expression of peptide described in the invention is
the use of the expression systems of the invention to express high levels of PTH. The
Human PTH (1-34) gene was synthesised by first annealing with each other, the two
complementary oligonucleotides which were designed on the basis of the sequence of
the Parathyroid hormone, and then by carrying out the polymerase chain reaction to fill
the remaining gaps. Subsequently, the enzyme cleavage site and the 5'- and 3'-end
restriction enzyme sites were incorporated into the gene by another PCR reaction using
specific primers. This modified-PTH gene which now carries the enzyme cleavage site
at its 5'-end and also carried the necessary restriction enzyme sites at both its ends, was
further cloned into the pET vector already containing the GCSF gene. This cloning
was done in such a manner so that the fusion product expressed by the cell transformed
by it would comprise of the G-CSF a.a. sequence, followed by the enzyme cleavage
site, further followed by the PTH (1-34) a.a. sequence as a single fusion product.

One of the preferred expression systems contemplated in the above described
invention comprises the expression of a fusion product consisting of G-CSF and PTH
separated by an enterokinase cleavage site. Another expression system contemplated in
this invention comprises the expression of a fusion product consisting of G-CSF' and
proinsulin separated by a CNBr cleavage site. A still another preferred expression
system contemplated in this invention comprises the expression of a fusion product
consisting of G-CSF and angiotensin separated by the enterokinase cleavage site.

The above peptides of interest may be produced in an industrially feasible
manner as insoluble fusion products forming inclusion bodies in the E.coli host using
standard fermentation methods as described in the literature. A single colony of the recombinant E.coli, harboring an expression vector encoding the desired fusion product is transferred to 5-mL Kanamycin-containing (50 mg/L) Luria Bertani medium. The culture is grown at 37 °C for 12-15 h at 200 rpm. The broth is used to inoculate 100-mL of the Kanamycin-containing (50 mg/L) Luria Bertani medium. The culture is sub-cultured to get about 1L broth, which is used to inoculate 10-L culture medium in a fermenter. The growth phase of fermentation is carried out either in a batch or a fed-batch manner. The composition of batch and fed-batch feed media are as follows:

Composition of batch media:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (1/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KH$_2$PO$_4$</td>
<td>13.3 g</td>
</tr>
<tr>
<td>(NH$_4$)$_2$HPO$_4$</td>
<td>4.0 g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>1.0 g</td>
</tr>
<tr>
<td>Glucose</td>
<td>20.0 g</td>
</tr>
<tr>
<td>Citric acid</td>
<td>1.7 g</td>
</tr>
<tr>
<td>MgSO$_4$ 7H$_2$O</td>
<td>1.2 g</td>
</tr>
<tr>
<td>Trace metal solution</td>
<td>20 ml</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>50 mg</td>
</tr>
</tbody>
</table>

Composition of fed-batch feed media:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (1/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>700 g</td>
</tr>
<tr>
<td>MgSO$_4$ 7H$_2$O</td>
<td>20 g</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>500 mg</td>
</tr>
<tr>
<td>Trace metal solution</td>
<td>20 ml</td>
</tr>
</tbody>
</table>

The fermentation is carried out at 37 °C, pH 6.8-7.0, 30-70% dissolved oxygen, with agitation. The feed-media is added either continuously or discontinuously. After attaining sufficient optical density (30-100 AU at 600 nm) the gene expression is induced by the addition of IPTG (0.1-2 mM) into the culture broth, thereby starting the production phase. The production phase is carried out either in batch, or fed-batch or repeated fed-batch mode. After the sufficient expression of protein in the form of inclusion bodies, the batch is harvested and centrifuged. The cell paste, obtained from the broth, is transferred for downstream processing.

The inclusion bodies containing the fusion product comprising of the peptide...
and the fusion partner, are isolated from the E. coli. cells by disruption using any of the suitable methods reported in the literature. The fusion product is solubilized from the isolated inclusion bodies using any of the standard methods reported in the literature. The desired peptide is separated from the fusion partner by treating it with an appropriate cleavage molecule, i.e., enzyme or chemical agent. The peptide of interest is then obtained from this mixture and brought to a high level of purity and yield by using standard methods of purification such as precipitation, column chromatography etc. reported extensively in the literature. The order of cleavage within the various downstream steps and the number and type of purification methods used in the invention may follow any of the suitable protocols reported in the standard literature.

The expression systems of this invention may be used to produce high level of expression of other peptides as well, e.g., insulin, calcitonin, exendin and the like. The invention is described in further details through the following examples which teach the skilled person to carry out the present invention. It will be appreciated that these examples are illustrative and the skilled person, following the teachings of these examples, which are for specific peptides, replicate the teachings with suitable modifications, alterations etc. as may be necessary, and which are within the scope of a skilled person, for other peptides which have been contemplated to be within the scope of the present invention.

**EXAMPLE 1**

**Construction of GCSF expression vector**

Human placental total RNA (Clontech) was used as a template for reverse transcriptase-polymerase chain reaction (RT-PCR) based cloning of the desired GCSF (174 amino acid variant) gene sequence. PCR was performed using a pair of gene specific oligonucleotide primers (Seq ID No 3 and 4). The oligonucleotides also contained the cloning sites and the start codon for Methionine, and were designed to modify the 5' end of the gene such that the GC content of the 5' region was reduced without affecting the amino acid sequence and the optimized codons were also efficiently used by E.coli. Seven codons were modified at 5' end. Double stranded cDNA was synthesized from Human placental total RNA using MMLV reverse transcriptase (MBI Fermentas, USA) by gene specific priming [Maniatis et al., Molecular cloning; A Laboratory Manual (Cold Spring harbor Laboratory Press, Cold Spring Harbor, N.Y.), (1990)]. This cDNA was then subjected to 40 cycles of PCR amplification using 100 picomoles of gene specific oligonucleotide primers in a volume
of 100 µl containing 50 mM Tris-Cl (pH8.3), 2.5 mM MgCl₂, 200 µM each of the 4 dNTPs and 2.5 units of Pfu Polymerase. Each PCR amplification cycle consisted of incubations at 94°C for 30 sec (denaturation), 61°C for 30 sec (annealing) and 72°C for 1 min (extension). This amplified PCR product was cleaned using Gel extraction Kit (Qiagen) and was digested with Nde I and Bam HI restriction endonucleases. The digested PCR product was further purified using Gel Extraction kit (Qiagen). The plasmid vector pET27b(+), (Novagen) was digested using Nde I and Bam HI restriction endonucleases and subsequently the linearized plasmid vector was gel purified using Gel extraction kit (Qiagen). This purified linearized vector DNA was ligated with the purified GCSF PCR product digested with Nde I and Bam HI. The ligation product was transformed in E. coli DH5α and transformants were scored on the basis of Ampicillin resistance. Plasmid DNA isolated from such 10 colonies was analysed for the presence of GCSF gene by restriction digestion using various restriction enzymes. One such plasmid was sequenced using automated DNA sequencer (ABI) and found to have the correct integration and sequence of the protein gene. This plasmid DNA was named pET27b(+)-GCSF-1. This plasmid DNA was transformed in E.coli BL21 DE3 competent cells by CaCl₂ chemical transformation method to obtain the G-CSF expressing E.coli cell clones. One such E.coli clone was earlier used to obtain very high expression levels of up to 9.5 g/L using a fermentation process described in WO 2007/102174 A2.

EXAMPLE 2

Construction of PTH fusion protein expression vector.

Human GCSF gene was PCR amplified from pET27b(+)-GCSF-1 plasmid vector prepared earlier, using a set of oligonucleotide primers where the forward primer contained the restriction site Nde I (Seq ID No 3), and reverse primer contained a BamHI restriction enzyme site (Seq ID No 5). This amplified PCR product was cleaned using Gel extraction Kit (Qiagen) and was digested with Nde I and Bam HI restriction endonucleases. The digested PCR product was further purified using Gel Extraction kit (Qiagen).

Human Parathyroid Hormone (1-34) gene was synthesised using two complementary oligonucleotides (Seq ID Nos. 8 and 9) corresponding to the coding and anti-coding strand of the Human Parathyroid Hormone (1-34) gene. Equimolar concentrations of both the oligonucleotides were heated in a Thermal Cycler at 95
degree celsius for 5 minutes and then allowed to anneal by cooling at room temperature for an hour. The annealed oligonucleotides were then subjected to 40 cycles of PCR amplification using 100 picomoles of gene specific oligonucleotide forward and reverse primers (Seq Id Nos. 10 and 11) in a volume of 100 μl containing 50 mM Tris-Cl (pH 8.3), 2.5 mM MgCl2, 200 μM each of the 4 dNTPs and 2.5 units of Pfu Polymerase. Each PCR amplification cycle consisted of incubations at 94°C for 30 sec (denaturation), 62°C for 30 sec (annealing) and 72°C for 1 min (extension). The forward primer contained the BamHI restriction site along with the Enterokinase Cleavage site. The amplified product of the PCR reaction was resolved on a 2.5% Agarose gel and was purified using Gel Extraction kit (Qiagen) The isolated PTH gene was digested using BamHI and Bpu1102I restriction endonucleases and subsequently purified using gel extraction kit (Qiagen).

The plasmid vector pET27 b (+), (Novagen) was digested using Nde I and Bpu1102I restriction endonucleases and subsequently the linearized plasmid vector was gel purified using Gel extraction kit (Qiagen).

Ligation of the linearized vector with GCSF and PTH (1-34) gene fragments was carried out in a 3 fragment ligation reaction. The ligation mix was transformed in E.coli DH5α-FT and plated on LB containing kanamycin. The colonies were screened using various restriction enzymes. The sequence was further confirmed by sequencing using automated DNA sequencer (ABI) and found to be having the correct integration and sequence of the HuGCSF-PTH (1-34) fusion protein gene (Seq ID Nos. 6 and 7). This plasmid clone was named pET27b-GCSF-PTH (1-34).

**EXAMPLE 3**

**Construction of Proinsulin fusion protein expression vector**

Human pancreatic total RNA (Clontech) was used as a template for reverse transcriptase-polymerase chain reaction (RT-PCR) based cloning of the desired proinsulin gene sequence. PCR was performed using a pair of gene specific oligonucleotide forward and reverse primers (Seq ID Nos. 14 and 15) corresponding to the coding region of the Proinsulin gene (Seq ID Nos. 12 and 13). Double stranded cDNA was synthesized from Human placental total RNA using MMLV reverse transcriptase (MBI Fermentas, USA) by gene specific priming [Maniatis et al., Molecular cloning; A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.), (1990)]. This cDNA was then subjected to 40 cycles of PCR
amplification using 100 pmoles of gene specific oligonucleotide primers in a volume of 100 µl containing 50 mM Tris-Cl (pH8.3), 2.5 mM MgCl₂, 200 µM each of the 4 dNTPs and 2.5 units of Pfu Polymerase. Each PCR amplification cycle consisted of incubations at 94°C for 30 sec (denaturation), 58°C for 30 sec (annealing) and 72°C for 1 min (extension). Amplified product of the PCR reaction was resolved on a 1.5% Agarose gel. The desired fragment of approximately 280 base pairs in size was excised out from the gel and purified using Qiagen Gel extraction kit. This purified DNA fragment was ligated into pET127b-GCSF-PTH (1-34) after restriction digestion of both the vector and the purified PCR product with BamHI and Bpu1102I (MBI Fermentas, USA). The digestion of the vector construct with BamHI and Bpu1102I excised the PTH (1-34) gene from the construct and the digested expression vector was purified using Qiagen Gel extraction kit. Purified vector was used for ligation with the digested and purified Proinsulin PCR product. The ligation product was transformed into E. coli DH5α and transformants were scored on the basis of Kanamycin resistance. Plasmid DNA isolated from such 10 colonies was analysed for the presence of Proinsulin gene by restriction digestion using various restriction enzymes. One such plasmid was sequenced using automated DNA sequencer (ABI) and found to have the correct integration and sequence of the Hu-GCSF-Proinsulin fusion protein (Seq ID Nos. 16 and 17). This plasmid DNA was named pET27B-GCSF-Proinsulin.

EXAMPLE 4

Construction of Angiotensin (1-10) fusion protein expression vector

The gene encoding the Angiotensin fusion protein was constructed by using pET27b-GCSF-PTH vector construct as a template in a PCR reaction using oligonucleotide forward and reverse primers of Seq ID Nos. 18 and 19, where the forward primer had complementarity to the upstream region of the vector and the reverse primer contained the Enterokinase cleavage site and the Angiotensin (1-10) anticoding region and had complementarity to the 3' end of G-CSF, such that the PCR product of the reaction was the coding sequence for the G-CSF-Enterokinase-Angiotensin fusion product (Seq ID No 16). The PCR reaction was carried out for 40 cycles of PCR amplification using 100 pmoles of gene specific oligonucleotide primers in a volume of 100 µl containing 50 mM Tris-Cl (pH8.3), 2.5 mM MgCl₂, 200 µM each of the 4 dNTPs and 2.5 units of Pfu Polymerase. Each PCR amplification cycle consisted of incubations at 94°C for 30 sec (denaturation), 55°C for 30 sec
(annealing) and 72°C for 1 min (extension). Amplified product of the PCR reaction was resolved on a 1% Agarose gel. The desired fragment of approx 600 base pairs in size was excised out from the gel and purified using Qiagen Gel extraction kit. This purified DNA fragment was ligated into pET27b after restriction digestion of both the vector and the purified PCR product with XbaI and BpuI102I (MBI Fermentas, USA). The ligation product was transformed in E. coli DH5α and transformants were scored on the basis of Kanamycin resistance. Plasmid DNA isolated from such 10 colonies was analysed for the presence of GCSF-Angiotensin fusion protein gene by restriction digestion using various restriction enzymes. One such plasmid was sequenced using automated DNA sequencer (ABI) and found to be having the correct integration and sequence of the fusion protein. This plasmid DNA was named pET27b-GCSF-Angiotensin.

EXAMPLE 5
Expression analysis of the fusion protein constructs.

All the three fusion protein expression vector constructs described above were extracted from their respective E. coli DH5α hosts described above and re-transformed into E.coli BL21 (DE3) using CaCl₂ chemical transformation method. Individual colonies from each set were inoculated in LB broth containing kanamycin (50 mg/L). On reaching an OD₆₀₀ of 0.8, the cultures were induced using 2 mM IPTG. The cells were harvested after 16 Hrs of induction. 25 μl cell suspension was lysed in the presence of SDS and beta mercapto ethanol (Gel loading buffer) and it was loaded on 15 % SDS PAGE and the induced protein band was quantitated using densitometry (Image quant 400, image quant TL ver 2005, 1D gel analysis).

Figure 1 shows the picture of gels showing the various fusion products in uninduced and induced cultures. Densitometric analysis of these gels showed that GCSF-PTTI fusion protein accumulated in the cultures at a very high expression level of 55.56 % of total bacterial protein. GCSF-Proinsulin fusion protein was also expressed at high levels of 40 % of total bacterial protein. The small peptide of 10 amino acids, angiotensin, when expressed as a GCSF-Angiotensin fusion product, was also produced at a high expression level of 50.28 % of total bacterial protein. The proinsulin and PTH fusion proteins clones were tested for the presence of the induced protein in soluble or insoluble form. The cells were lysed by high pressure cell homogenizer and the insoluble inclusion bodies were centrifuged and separated and
solubilized. The solubilized inclusion bodies and the supernatant were both analysed by
gel electrophoresis. More than 90 % of the induced protein was found to be present in
the form of insoluble inclusion bodies.

EXAMPLE 6

Expression of PTH fusion protein in fermenter

Using the expression systems of this invention, the recombinant PTH fusion protein
protein, could be produced in fermenters in an industrially feasible manner. A single
colony of the recombinant E. coli BL21 (DE3), harboring the PTH fusion protein
expression vector was transferred to 5-mL Kanamycin-containing (50 mg/L) Luria
Bertani medium. The culture was grown at 37 °C for 12-15 h at 200 rpm. The broth was
used to inoculate 100-mL of the Kanamycin-containing (50 mg/L) Luria Bertani
medium. The culture was sub-cultured to get about 1L broth, which was used to
inoculate 10-L culture medium in a 30-L fermenter. The growth phase of fermentation
can be carried out in batch or fed-batch manner and the composition of batch and fed-
batch feed media are given below. Specifically in this example we carried out the
growth phase of fermentation in the fed-batch mode.

Composition of batch media:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (1/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KH₂PO₄</td>
<td>13.3 g</td>
</tr>
<tr>
<td>(NH₄)₂HPO₄</td>
<td>4.0 g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>1.0 g</td>
</tr>
<tr>
<td>Glucose</td>
<td>20.0 g</td>
</tr>
<tr>
<td>Citric acid</td>
<td>1.7 g</td>
</tr>
<tr>
<td>MgSO₄·7H₂O</td>
<td>1.2 g</td>
</tr>
<tr>
<td>Thiamine</td>
<td>0.1 g</td>
</tr>
<tr>
<td>Trace metal solution</td>
<td>20 ml</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>50 mg</td>
</tr>
</tbody>
</table>

Composition of fed-batch feed media:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (1/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>700 g</td>
</tr>
<tr>
<td>MgSO₄·7H₂O</td>
<td>20 g</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>500 mg</td>
</tr>
<tr>
<td>Trace metal solution</td>
<td>20 ml</td>
</tr>
</tbody>
</table>

The fermentation was carried out at 37 °C, pH 6.8-7.0, 30-70% dissolved
oxygen, with agitation. The feed-media can be added continuously or discontinuously.
Specifically in this example, the feed-batch media was added continuously in a
substrate-limiting fed-batch manner. After attaining sufficient optical density (30-100
AU at 600 nm) the gene expression was induced by addition of IPTG (0.1-2 mM) into
the culture broth, thereby starting the production phase. The production phase can be carried out in batch, fed-batch or repeated fed-batch mode. Specifically in this example we carried out the production phase in a fed-batch mode and the composition of the fed-batch media used is given below.

Composition of fed-batch feed media for production phase:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (1/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>270 g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>214 g</td>
</tr>
<tr>
<td>MgSO4.7H2O</td>
<td>1 g</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>500 mg</td>
</tr>
</tbody>
</table>

The production phase was carried out for a period of 12 hours using continuous addition of fed-batch media for production phase and was added in a substrate-limiting fed-batch manner. After the sufficient expression of protein the batch is harvested and centrifuged. The cell paste, obtained from the broth, is transferred for downstream processing. The volumetric yield of the PTH fusion protein was measured by densitometric analysis of SDS-PAGE gels, and determined to be 4.62 g/L.

EXAMPLE 7

Expression of Pro-Insulin fusion protein in fermenter

Using the expression systems of this invention, the recombinant Pro-insulin fusion protein, could be produced in fermenters in an industrially feasible manner. A single colony of the recombinant E. coli BL21 (DE3), harboring the Pro-insulin fusion protein expression vector was transferred to 5-mL Kanamycin-containing (50 mg/L) Luria Bertani medium. The culture was grown at 37 °C for 12-15 h at 200 rpm. The broth was used to inoculate 100-mL of the Kanamycin-containing (50 mg/L) Luria Bertani medium. The culture was sub-cultured to get about 1L broth, which was used to inoculate 10-L culture medium in a 30-L fermenter. The growth phase of fermentation can be carried out in batch or fed-batch manner and the composition of batch and fed-batch feed media are given below. Specifically in this example we carried out the growth phase of fermentation in the fed-batch mode.

Composition of batch media:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (1/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KH2PO4</td>
<td>13.3 g</td>
</tr>
<tr>
<td>(NH4)2HPO4</td>
<td>4.0 g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>1.0 g</td>
</tr>
<tr>
<td>Glucose</td>
<td>20.0 g</td>
</tr>
<tr>
<td>Citric acid</td>
<td>1.7 g</td>
</tr>
<tr>
<td>MgSO4.7H2O</td>
<td>1.2 g</td>
</tr>
</tbody>
</table>
Thiamine 0.1 g
Trace metal solution 20 ml
Kanamycin 50 mg

Composition of fed-batch feed media:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (1/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>700 g</td>
</tr>
<tr>
<td>MgSO₄·7H₂O</td>
<td>20 g</td>
</tr>
<tr>
<td>Trace metal solution</td>
<td>20 ml</td>
</tr>
</tbody>
</table>

The fermentation was carried out at 37 °C, pH 6.8-7.0, 30-70% dissolved oxygen, with agitation. The feed-media can be added continuously or discontinuously. After attaining sufficient optical density (50-100 AU at 600 nm) the gene expression was induced by addition of IPTG (0.1-2 mM) into the culture broth, thereby starting the production phase. The production phase can be carried out in batch, fed-batch or repeated fed-batch mode. Specifically in this example we carried out the production phase in a fed-batch mode and the composition of the fed-batch media used is given below.

Composition of fed-batch feed media for production phase:

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (1/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>270 g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>214 g</td>
</tr>
<tr>
<td>MgSO₄·7H₂O</td>
<td>1 g</td>
</tr>
</tbody>
</table>

Kanamycin (20-450 mg/L) was added after a suitable period of time (0-12 h) after induction. The production phase was carried out for a period of 12 hours using continuous addition of fed-batch media for production phase and was added in a substrate-limiting fed-batch manner. After the sufficient expression of protein the batch is harvested and centrifuged. The cell paste, obtained from the broth, is transferred for downstream processing. The volumetric yield of fusion proinsulin, measured by densitometric analysis of SDS-PAGE gels, was determined to be 3.38 g/L.

The sequences of the protein, nucleotides and oligonucleotides (Seq IDs Nos. 1 to 19) are as follows:

**INFORMATION FOR SEQ ID NO 1 AND 2:**

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH : 578 bases
(B) TYPE : Nucleic acid and amino acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) FEATURE
(A) NAME / KEY : CDS (Nathionyl Human SC5F coding sequence and protein)
(B) LOCATION : 1-525
SEQ ID NO 1 and 2

1 AUG ACG GCA TTA GGT CCA GCA ACG TGG TGG CAG AGG TCA CCG CAG AGG TTC CTA 43
5 ANT PFL LGO PAA SSLP OSU L 15

46 CTC AAG TGC TTA GAC GAT CAG AAG ATC CAG GCC GAC CAT GCA 90
16 LKC LEO VRR XIC GUG A 30

91 GCC CTC CAG GAG AAG CTT TTT GCC ACC TAC AAG CTT TCC ACC CAC 135
31 ALQ EKL CAY TKL CCH 45

136 SAG GAG CTG GTG GTG CTG GCA CAC TCT CTT GCC ATC CCC TGG GCT 180
46 BKL VLL LGH STL LGE iP W 60

181 CCC CTG AGC AGC TGC CCC ACC AGC CAG CCC CTG CAG CCC CCA GCC TCC 225
61 PLSSCFSQALQLAGC 75

226 TTAG ACC CAA CTC CAT AGG GCC CTG TCC CTC TAC CAG GGG CTC CTG 270
76 LSQ LHS GLY FLQ GLL 90

271 CAG GCC CTG GAA GGG ATC TCC CCC GAC TGG GGT GCT ACC AGG TGG GAC 315
91 QAELGISPELQPTLD 105

316 ACA CTG CAG CTG GAC GTC GCC GGC TTT GCC ACC AGC ATC TGG CAG 360
106 TLQLD VAD FPATTIQ 120

361 CAG ATG GAA GAA CTT GGG ATG GCC OCT GCC CTG CAG CCC ACC CAG 405
121 QNE ELM NAP LAOFTQ 135

406 GCT GCC ATC CCC GCC TCC GCC TCT GGT TCT CAG GGG GGA GCA GGA 450
136 GMPA FAS AFSFQ RARAG 150

451 GGG GTC CTA GTC CCC TCG CTT CAG AGG TCC CTG GAG GTG TCG 495
151 GVLYV AESHILQS SLEVS 165

496 TAC GCC GTC CTA CCG CAC GTC GCC CAG CCC TCA 538
166 TRVLRHLAQPS 40

22
INFORMATION FOR SEQ ID NO 3:
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH : 25 bases
   (B) TYPE : Nucleic acid (oligonucleotide)
   (C) STRANDEDNESS : single
   (D) TOPOLOGY : linear

Seq Id no 3 : GGCATGACTCCATTAGCTCC

INFORMATION FOR SEQ ID NO 4:
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH : 32 bases
   (B) TYPE : Nucleic acid (oligonucleotide)
   (C) STRANDEDNESS : single
   (D) TOPOLOGY : linear

Seq Id no 4 : GGGAGTCCTCAGGGGTGGGAAAGGGTGCCGTAG

INFORMATION FOR SEQ ID NO 5:
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH : 25 bases
   (B) TYPE : Nucleic acid (oligonucleotide)
   (C) STRANDEDNESS : single
   (D) TOPOLOGY : linear

Seq Id no 5 : GGGATCGGGGTGGGAAAGGGTGCCGTAG

INFORMATION FOR SEQ ID NO 6 and 7:
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH : 631 bases
   (B) TYPE : Nucleic acid
   (C) STRANDEDNESS : single
   (D) TOPOLOGY : linear
(ii) FEATURE
   (A) NAME / KEY : CDS (GCSP-PTH fusion protein)
   (B) LOCATION : 1-688
(iii) FEATURE
   (A) NAME / KEY : GCSP Fusion partner
   (B) LOCATION : 1-525
(iv) FEATURE
   (A) NAME / KEY : Enterokinase cleavage site
   (B) LOCATION : 532-546
(vi) FEATURE
   (A) NAME / KEY : PTH (1-34)
   (B) LOCATION : 547-648

SEQ ID NO 6 and 7
1   ATG ACT CCA TTA COT CCA CCA ACC TCC CTC CCC CAG AGC TCC CTT
1   MTF LGF PAS SHL PQL FSL   15
55
46  CTC AAG TGC TTA GAG CAA CTC ACC AGC AGC ATC CAG GCC CAT GCC GCA
16  LRC LCE QVRKQ IGDG A   30
91  GCC CTC CAG GAG AAG CTC GCC ACC TAC AAG CTC TGC CAC CCC
135
60  ALQ ERK LCATY KLPCH F   45
136 GAG GAG CTC CTT CTC CTC GGA CAC TCT CTC GGC ATC CCC TGG GCT
   80

23
(D) TOPOLOGY : linear

Seq id no 10 : GCGGATGGATGAGATGATAAAAATCGGCTGCGCAATACG

5

INFORMATION FOR SEQ ID NO 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH : 34 bases
(B) TYPE : Nucleic acid (Synthetic oligonucleotides)
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

Seq id no 11 : TTATGCGCAAGCTCAAAATAAGTGCGACATCTT

15

INFORMATION FOR SEQ ID NO 12 and 13

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH : 785 bases
(B) TYPE : Nucleic acid
(C) STRANDEDNESS : single
(D) TOPOLOGY : linear

(ii) FEATURE
(A) NAME / KEY : CUX (GCSE-FPROINSULIN fusion protein)
(B) LOCATION : 1-32

(iii) FEATURE
(A) NAME / KEY : GDEF Fusion partner
(B) LOCATION : 1-525

(iv) FEATURE
(A) NAME / KEY : Cyanogen bromide cleavage (Metthionine)
(B) LOCATION : 532-534

(v) FEATURE
(A) NAME / KEY : FPROINSULIN (1-34)
(B) LOCATION : 535-792

35

SEQ ID NO 12 and 13

1 ARG ACT CCA TTA GGT CCA GCA AGC TCC CTC CCA CAG AGG TCG CTG 45

1 MTFLGLPASSLPQSFLL15

46 CTC AGG TGC TTA GAS CAA GCT AGG AGG AGC CNG GCC GAT GCA GCA 90

16 LKCNLPQVRFTQQCNSA30

45 GCC CTC CAG GAS AAG CTC TGC GCC ACC TAC AAG CTC TGC TAC CAC CCC 135

31 ALQEKCLCATYKLC45

136 GAS GAG CTG CTG CTG CTC GCA CAC TCT CTT GCC AGC GCC TGG GCT 180

46 BELLGLGLHSLIGFA60

161 CCC CTC AGG AGC TGC CCC AGC CAG CTC CAG CGG CGG TGG CGC 225

61 FLLSCSPQALQLASC75

226 TGT AGC CAA CTC CAT AGG GGC CTY TCC CTC TAC CAG GGC CTC CTG 270

23

COMS ID No: ARCS-274468 Received by IP Australia Time (H:m) 17:31 Date (Y-M-d) 2010-04-21
SEQ ID NO 14 : GCCGATCTCATCTTTTGGGACACGCACCTGTCG

INFORMATION FOR SEQ ID NO 15:

(A) DESCRIPTION: Nucleic acid (Synthetic oligonucleotide)

(B) TYPE: Nucleic acid (Synthetic oligonucleotide)

(C) STRAND: Single

(D) TOPOLOGY: Linear

SEQ ID NO 16 : TTAATGCTCATCTTTTGGGACACGCAATGTGCAGAGCTG

INFORMATION FOR SEQ ID NO 16 and 17:

(A) DESCRIPTION: Nucleic acid

(B) TYPE: Nucleic acid

(C) STRAND: Single

(D) TOPOLOGY: Linear

(FEATURE)

(A) NAME / KEY: CDS (GCSF-ANGIOTENSIN fusion protein)

(B) LOCATION: 1-576

(FEATURE)

(A) NAME / KEY: CDS Fusion partner

(B) LOCATION: 520-546

(FEATURE)

(A) NAME / KEY: Angiogenatin (1-10)

(B) LOCATION: 547-576

SEQ ID NO 16 and 17

1 ATG ACT CCA TTA GCC CCA GCA AGC TCC CTS CCC CAG AGC TTC CGT

1 N T P L G P A S S L P O S V L

45 CTC AAG TGC TTA GAG CCA GTG AGG AAG ATC CAG GGC GAT GGC GCA

15 L K C L F L Q V R K I Q S D G A

90 GGC CTC CAG GAG AAG CTC TGT GCC ACC TAC AAG CTC TGC CAG CCC

135 A L Q E K L C A T Y K L C H F

180 GAG GAG CTA CCG GGC CAC TCT CTS GCC ATC CCC TGG GCT

225 S V L L G H S L G I F W A

275 CCC CTS AGC AGC TGC CCG AGC CAG GCC CTC CAG CTC GCA GCC TGC

320 P L S S C P S Q A L Q L A G C

COMS ID No: ARCS-274468   Received by IP Australia: Time (H:m) 17:31 Date (Y-M-d) 2010-04-21
226 TGG AGG CAA CTG CAC TAG GCC GGT GCC CTC CTA CAG CAG CTC CTA CAG 270
76 L S Q L E S G L F L Y Q G L L
271 CAG GCC GTC GAA GGG ATC TGC CCC GAG TGG GGT CCC ACC TGG CAC 315
91 Q A L E G I S P E L G P I L D
316 ACA CTA GAG CAG CAG GTC GGC GAC TTC GCC ACC ACC ATC TGG CAG 360
106 T L Q L D V A D F A T T I N G
121 Q M E E L G M A D A L Q P T Q 135
406 GGT GCC ATG CCG GCC TCC GCC TCT GCT TCT CAG CGC GCG GCA GGA 450
136 G A M P A F A S A F P Q R R G 150
451 GGG GTC GCT GTA GGC TCC CAT CTG CAG ACC TGC CTC GAG GTG TCG 495
151 G V L V A S H L Q S E L E V S
496 TAC GGC GTT CTA GCA GAC CTT GCC ACC CAG GCC GGA GTC GAT GAT 540
166 Y R V L R H I L A Q P C S D D D 180
541 GAT AAA CAT GGC GGC TAT ATT CAT CCG TTT CAT CTG TAG 579
181 D X R V Y T H P F H L *

INFORMATION FOR SEQ ID NO 18:
(1) SEQUENCE CHARACTERISTICS:
  (A) LENGTHS : 30 bases
  (B) TYPE : Nucleic acid (Synthetic oligonucleotide)
  (C) STRANDEDNESS : single
  (D) TOPOLOGY : linear

SEQ ID NO 18 : CCCCCGTAGAAATATTTTGGTTTACCTTTAAGAAGGAG

INFORMATION FOR SEQ ID NO 19:
(1) SEQUENCE CHARACTERISTICS:
  (A) LENGTHS : 66 bases
  (B) TYPE : Nucleic acid (Synthetic oligonucleotide)
  (C) STRANDEDNESS : single
  (D) TOPOLOGY : linear

SEQ ID NO 19 : GCCCGCACTGGCTCAAGATGAAAGGATGAATATACACGCGAGCTTTACATCA
TCGATCGGATCGG
We Claim

1. A process for the production of peptides comprising the use of GCSF or its suitable variants as a fusion partner.

2. The process as claimed in claim 1 wherein GCSF forms a fusion product with the peptide.

3. The process as claimed in claim 1 or 2 wherein the fusion product construct is represented by the formula:

```
Fusion partner ——— CS ——— Fusion peptide
```

wherein the fusion peptide is a peptide of interest, CS is a suitable cleavage site, and the fusion partner is GCSF or its suitable variants.

4. The process as claimed in claim 3 wherein the fusion peptide is a physiologically active peptide.

5. The process as claimed in any preceding claim wherein the CS is either present within the fusion partner or the fusion peptide amino acid sequences or is a suitable linker.

6. The process as claimed in any of the preceding claims wherein the fusion partner is linked at its C terminal end with a peptide at its N terminal end via a cleavage site.

7. The process as claimed in any of the preceding claims wherein the physiologically active peptide is selected from those polypeptides having a length from about 10 amino acids to about 90 amino acids.

8. The process as claimed in claim 1 wherein the physiologically active peptides comprise peptides selected from the group comprising of calcitonin, calcitonin, insulin, angiotensin, tissue plasminogen activator, growth hormone, growth factors, growth hormone releasing factors, cytokines, erythropoietin, interferons, interleukins, oxytocin, vasopressin, ACTH, collagen binding protein, parathyroid hormone, glucagon like peptide, glucagon, proinsulin, tumor necrosis factor, substance P, brain naturetic peptide, individual heavy and light antibody chains, peptide antibiotics, fuzone, octreotide, somatostatin and the suitable variants of these peptides.

9. The process as claimed in any of the preceding claims wherein the expression cassette encoding the fusion product construct is cloned in a suitable expression vector.
10. The process as claimed in claim 9 wherein the fusion product construct is encoded in a suitable expression vector selected from those which can express the fusion product in E. coli.

11. The process as claimed in claims 7 or 8 wherein the physiologically active peptide is selected from PTH (1-34), Angiotensin and Proinsulin.

12. The process as claimed in claims 7, 8 or 11 wherein the physiologically active peptide is PTH (1-34) of Sequence ID No 6.

13. The process as claimed in claim 12 wherein the physiologically active peptide PTH (1-34) is expressed by expression vector pET27B-GCSF-PTH deposited under MTCC Accession no. 5425.

14. The process as claimed in claims 7, 8 or 11 wherein the physiologically active peptide is Proinsulin of Sequence ID Nos 12 and 13.

15. The process as claimed in claim 14 wherein the physiologically active peptide proinsulin is expressed by expression vector pET27B-GCSF-Proinsulin deposited under MTCC Accession no 5424.

16. The process as claimed in claims 7, 8 or 11 wherein the physiologically active peptide is Angiotensin of Sequence ID Nos 16 and 17.

17. The process as claimed in claim 16 wherein the physiologically active peptide proinsulin is expressed by expression vector pET27B-GCSF-Angiotensin.

18. The process as claimed in any of the claims above wherein the cleavage site is an enzymatic or a chemical cleavage site.

19. The process as claimed in claim 18 wherein the cleavage site is a chemical cleavage site.

20. The process as claimed in claim 18 wherein the cleavage site is an enzymatic cleavage site.

21. The process as claimed in claim 18 or 19 wherein the cleavage sites are those which are cleavable by chemicals selected from cyanogen bromide, 2-(2-nitrophenyl)phenyl)-methyl-3'-bromoindolenine, BNPS-skatole, N-bromosuccinimide, O-iodosobenzonic acid, HBr/DMSO, NTCB, Sodium metal in liquid ammonia, Hydroxylamine or dilute acids.

22. The process as claimed in claim 21 wherein the chemical is cyanogen bromide.

23. The process as claimed in claim 18 or 20 where the enzymatic cleavage sites are those recognized by Enterokinase, Trypsin, Chymotrypsin, Elastase, Pepsin, Papain, Subtilisin, Thermolysin, V8 Protease, Eudoproteinase Arg C.
(submaxillaris protease), Clostripain, Thrombin, Collagenase, Lysobacter enzymogenes (Lys C), Mysobacter A1-1 Protease or Factor Xα.

24. The process as claimed in claim 23 where the enzyme is Enterokinase.

25. The process as claimed in any of preceding claims further comprising the steps of:

(a) cloning the gene encoding the fusion product construct in a suitable expression vector;
(b) transforming suitable host cells with the above expression vector;
(c) expressing the desired polypeptide as a fusion protein in host cells;
(d) disrupting the said host cells and collecting the fusion protein as inclusion bodies;
(e) separating the fusion protein containing inclusion bodies from other host components;
(f) solubilizing the fusion protein with a suitable denaturing agent;
(g) cleaving the cleavage site with a suitable enzyme or chemical to release the fusion peptide; and
(h) purifying the fusion peptide from the reaction mixture.

26. The process as claimed in any of preceding claims wherein the host cell for the expression of the fusion product is E. coli.

27. The process as claimed in claim 1 wherein the said GCSF fusion partner is encoded by the nucleotide sequence set forth in Sequence ID Nos. 1 and 2.

28. A fusion product as claimed in any of preceding claims.

29. A fusion product as claimed in claim 28 having the formula -

Fusion partner ——— CS ——— Fusion peptide

wherein the fusion peptide is a peptide of interest, CS is a suitable cleavage site, and the fusion partner is GCSF or its suitable variants.

30. An expression vector comprising the expression cassette encoding the fusion product construct of any of the preceding claims.

31. An expression vector according to claim 30 wherein the expression cassette is operably linked to a promoter and other suitable regulatory elements functional in the expression host.

32. An expression vector according to claims 30 or 31 which is selected from pET27B-GCSF-PTH, pET27B-GCSF-Proinsulin and pET27B-GCSF-Angiotensin.
Figure 1

<table>
<thead>
<tr>
<th>Lane</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>LMW marker</td>
</tr>
<tr>
<td>1</td>
<td>Proinsulin Induced</td>
</tr>
<tr>
<td>2</td>
<td>Proinsulin Uninduced</td>
</tr>
<tr>
<td>3</td>
<td>Angiotensin uninduced</td>
</tr>
<tr>
<td>4</td>
<td>Angiotensin induced</td>
</tr>
<tr>
<td>5</td>
<td>PTH proinduced</td>
</tr>
<tr>
<td>6</td>
<td>PTH induced</td>
</tr>
</tbody>
</table>

-34-