A BACTERIAL EXPRESSION VECTOR FOR ENHANCED PROTEIN SECRETION

FIELD OF THE INVENTION

The invention relates to a bacterial expression vector(s) capable of enhanced secretion of recombinant protein in the periplasm space or extracellularly. More specifically, the invention relates to expression vector for expressing and secreting recombinant protein from *Escherichia coli*.

BACKGROUND OF THE INVENTION

Heterologous expression and purification of recombinant proteins *in vitro* and *in vivo* represent routine applications of modern molecular biology. Expression of recombinant proteins is typically carried out in prokaryotic host cells, and the microbe *Escherichia coli* have been used regularly for mass production of recombinant proteins, relevant to pharmaceuticals and industrial industries. In view of this, there have been quite a few modifications in host cells to enhance the production of protein of interest in these host cells, using several known methods and mechanisms, to maximize its utility in the industrial production.

The industrial production requires the protein of interest to be secreted out of the host cell for simplifying the purification and down-stream processing of the protein of interest, in a considerable volume. Several efforts are still being made to enhance the extracellular production of target proteins in *Escherichia coli*, by adapting the target cell of expression. One such important adaptation is the use of selected secretory signal peptides, and co-expression of protein of interest along with the secretory signal peptides which assists the translocation of recombinant protein of interest to periplasmic space or extracellular culture media.

Several prior art citations provide novel expression systems for generating protein of interest which are secreted out of host cells:

US5583038A describes expression vectors for expressing and secreting proteins which are heterologous to the bacterium which expresses such proteins wherein such vectors further include DNA encoding at least the secretion signals of lipoproteins designed to achieve lipid acylation and surface expression of heterologous proteins. The bacterial system used is the mycobacteria specifically, and the secretory signal described is mycobacterial lipoprotein secretion signal sequence of Outer Surface Protein A. This expression vector was specifically designed for expressing and secreting an antigen which elicits antibodies against Borrelia burgdorferi to be used for treating Lyme disease;

US5432082A describes an expression vector for yeasts which is useful for preparing heterologous proteins, comprising a synthetic oligonucleotide which directs the secretion of the heterologous protein wherein the synthetic oligonucleotide is positioned between the inducible hybrid promoter GAL-CYC and a multiple-site polylinker followed by the signals of transcription termination recognized by the RNA polymerase of the yeasts;

CN101687910A describes a mammalian cell based expression and secretion system. The DNA encoding the signal peptide sequence is selected from: MMRP encoding amino acid sequence [hydrophobic amino] nTSALA DNA sequences or encoded amino acid sequence MKT [hydrophobic amino] nCATVHC DNA sequence, wherein n is an integer between 4-16, the hydrophobic amino acid is A, I, L, M, F or V; and

CN107082801A discloses a pelB signal peptide mutant capable of improving the protein secretion efficiency.

Although numerous proteins have been successfully produced by manipulating the secretary method in *Escherichia coli*, still the volume of production is a limiting factor due to proteins not being exported correctly or not exported in a functional state due to

aggregation in the cytoplasm; lysis of the cells; incorrect folding; limitations to translocation or proteolytic degradation.

The present invention takes into account the drawbacks of the prior art and provides a novel bacterial expression vector for secretion of recombinant protein extracellularly.

OBJECT OF THE INVENTION

The main object of the invention is to provide a bacterial expression vector carrying a novel secretory signal sequence to direct the secretion of recombinant protein into extracellular media, wherein, the secretory signal sequence is a novel combination of DNA sequence of a signal peptide and DNA sequence of a carrier protein in tandem, and DNA sequence encoding recombinant protein operably linked thereto.

Accordingly, another object of the invention is provide a novel secretory signal sequence for construction of a bacterial expression vector, for enhancing secretion of recombinant protein from the host cells, wherein the secretory signal sequence is a combination of a) at least one DNA sequence encoding a signal sequence of genes selected from the group consisting of *pelB* represented by Seq. ID 1, *ompA* represented by Seq. ID 2, *yebF* represented by Seq. ID 3, and *ompF* represented by Seq. ID 4, and b) a DNA sequence encoding a carrier protein, preferably, DNA sequence encoding yebf truncated peptide represented by Seq. ID 5 and Seq. ID 6, wherein Seq. ID 6 is generated by mutation of Seq. ID 5.

Yet another object of the invention is provide a system to enhance secretion of recombinant protein from the host cells, more specifically *Escherichia coli*, into the extracellular media.

SUMMARY OF THE INVENTION

In the main embodiment of the present invention, the invention provides a novel bacterial expression vector to enhance secretion of recombinant protein from a host bacterial cells, preferably, *E. coli*, for easy and efficient purification of recombinant protein. Said bacterial expression vector for enhanced secretion of recombinant protein from host bacterial cells, comprising of

at least one secretory signal sequence,

at least one gene expression cassette comprising of at least one inducible promoter, an RBS**,** DNA sequence encoding the recombinant protein, DNA sequence encoding an affinity tag, and at least one gene terminator, with the secretory signal sequence operably linked to the gene expression cassette, and the DNA sequence of the affinity tag operably linked to the DNA sequence of the recombinant protein,

at least one bacterial *ori* gene sequence for replication of the vector in the host bacterial cell, and

at least one DNA sequence for coding a selectable marker with a suitable promoter and a gene terminator sequence flanking the DNA sequence of the selectable marker.

Further, the secretory signal sequence is a combination of:

a) at least one DNA sequence of signal peptide of genes selected from the group consisting of *pelB* represented by Seq. ID 1 encoding amino acid sequence Seq. ID 9, the DNA sequence of signal peptide *ompA* represented by Seq. ID 2 encoding amino acid sequence Seq. ID 11, the DNA sequence of signal peptide *yebF* represented by Seq. ID 3 encoding amino acid sequence Seq. ID 10, and the DNA sequence of signal peptide *ompF* represented by Seq. ID 4 encoding amino acid sequence Seq. ID 12; and

b) at least one DNA sequence encoding a carrier peptide, preferably, DNA sequence encoding truncated yebF, and the DNA sequence encoding the truncated yebF belongs to the group consisting of Seq. ID 5 and Seq. ID 6, wherein, Seq. ID 5 encodes truncated yebF represented by Seq. ID 7; Seq. ID 6 encodes truncated yebF represented by Seq. ID 8; and the DNA sequence Seq. ID 6 is synthesized by mutating Seq. ID 5, wherein, TGC codon at position 40 in Seq. ID 5 is mutated to GCG codon, to mutate Cys at position 14 of amino acid sequence represented by Seq. ID 7 to Ala at position 14 of amino acid sequence represented by Seq. ID 8.

The bacterial expression vector represented by Seq. ID 13 is 6793 basepair vector comprises of secretory signal sequence, wherein, the secretory signal sequence is a combination of DNA sequence which encodes signal peptide of *ompF* of Seq. ID 12; and truncated yebF peptide of Seq. ID 8.

In yet another embodiment the invention provides a novel bacterial expression vector comprising of at least one affinity tag sequence operably linked to DNA sequence encoding recombinant protein enabling purification of recombinant protein.

In yet another embodiment, the secretory signal sequence is used in conjugation with additional elements in a basic vector that are necessary for its function as a vector including at least one antibiotic selectable marker, and at least one additional selection marker; wherein, the selectable marker is inducible lac operon, and the vector is an inducible vector under lac operon, and is induced by Lactose or Lactose analogues including IPTG.

BRIEF DESCRIPTION OF THE DRAWING

The object of the invention may be understood in more details and more particularly description of the invention briefly summarized above by reference to certain embodiments thereof which are illustrated in the appended drawings, which drawings form a part of this specification. It is to be noted, however, that the appended drawings illustrate preferred embodiments of the invention and are therefore not to be considered limiting of its scope, for the invention may admit to other equally effective equivalent embodiments.

Figure 1 is a schematic representation of a novel bacterial expression vector comprising of combination of secretory signal sequences for enhancing secretion of protein/peptide of interest outside the host cells;

Figure 2 is a schematic representation of pBacSec-LC vector;

Figure 3 is a graphical representation of comparative efficiencies of Seq. ID 1, Seq. ID 2, Seq. ID 3, and Seq. ID 4 using Gaussian luciferase assay;

Figure 4 is pictorial representation of SDS-PAGE gel with lysates or media component of bacterial cell culture expressing pBacSec-LC vector having secretory signal sequence encoding Seq. ID 12 and Seq. ID 8;

Figure 5 is a graphical representation of yield of recombinant protein after time points of induction by IPTG; and

Figure 6 is pictorial representation of SDS-PAGE gel with media component of bacterial cell culture expressing pBacSec-LC vector having secretory signal sequence encoding Seq. ID 12 in combination with Seq. ID 7, or Seq. ID 12 in combination with Seq. ID 8 under reducing or non-reducing conditions.

DETAILED DESCRIPTION OF THE INVENTION

Abbreviations used:

pelB refers to: leader DNA sequence encoding the N-terminal amino acid residues of pectatelyase B

ompA refers to: leader DNA sequence encoding the amino acid residues of Outer membrane protein A

ompF refers to: leader DNA sequence encoding the amino acid residues of Outer membrane protein F

yebF refers to: leader DNA sequence encoding the amino acid residues of protein yebF

Amp^r/Kan^r refers to: DNA sequence encoding Ampcillin/Kanamycin resistance gene

F1 Ori refers to: Origin of replication

ptac promoter refers to: Promoter for binding RNA polymerase or T7 polymerase

Lac or Lac1 refers to: DNA sequence encodes lac repressor/operon

IPTG refers to: Isopropyl β- d-1-thiogalactopyranoside, inducer of lac operon

Lactose refers to: disaccharide and inducer of lac operon

RBS refers to: Ribosomal binding site

Recombinant protein: Protein of interest encoded by gene cloned in expression vector and expressed in bacterial cells

The present invention now will be described hereinafter with reference to the detailed description, in which some, but not all embodiments of the invention are indicated. Indeed, the invention may be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments

are provided so that this disclosure will satisfy applicable legal requirements. The present invention is described fully herein with non-limiting embodiments and exemplary experimentation.

The present invention relates to a novel bacterial expression vector to enhance secretion of recombinant protein from a host bacterial cells, preferably, *E. coli,* for easy and efficient purification of the recombinant protein.

In the main embodiment, the invention provides a bacterial expression vector (**100),** as depicted **Figure 1,** comprising of a secretory signal sequence in tandem with DNA sequence encoding recombinant protein (**103)**, wherein, the secretory signal sequence is a combination comprising of:

a) at least one DNA sequence encoding a signal sequence (**101)** of genes selected from the group consisting of *pelB* represented by Seq. ID 1 encoding amino acid sequence Seq. ID 9, *ompA* represented by Seq. ID 2 encoding amino acid sequence Seq. ID 11, *yebF* represented by Seq. ID 3 encoding amino acid sequence Seq. ID 10, and *ompF* represented by Seq. ID 4 encoding amino acid sequence Seq. ID 12, and b) at least one DNA sequence encoding a carrier peptide (**102),** preferably, DNA sequence encoding truncated yebF represented by Seq. ID 5 and Seq. ID 6.

Seq. ID 5 encodes for 33 amino acid carrier peptide represented by Seq. ID 7, and Seq. ID 6 encodes for peptide represented by Seq. 8.

the DNA sequence Seq. ID 6 is synthesized by mutating Seq. ID 5, wherein, TGC codon at position 40 in Seq. ID 5 is mutated to GCG codon, to mutate Cys at position 14 of Seq. ID 7 to Ala at position 14 of Seq. ID 8.

Table 1 provides the DNA sequence encoding the signal sequence or the carrier peptides

SEQ. ID No.	DNA sequence
1	ATGAAATACCTGTTACCTACCGCGGCTGCGGGGCTGCTGCTG TTAGCAGCTCAGCCGGCAATGGCT
2	ATGAAGAAGACCGCGATTGCGATTGCGGTGGCGCTGGCGGG TTTTGCGACCGTGGCGCAGGCG
3	ATGAAAAAGCGTGGTGCGTTCCTGGGCCTGCTGCTGGTTAGC GCGTGCGCGAGCGTGTTTGCG
4	ATGATGAAGCGCAATATTCTGGCAGTGATCGTCCCTGCTCTG TTAGTAGCAGGTACTGCAAACGCT
5	GCGAACAACGAAACCAGCAAGAGCGTGACCTTTCCGAAATG CGAAGATCTGGATGCGGCGGGTATTGCGGCGAGCGTTAAGC GTGACTACCAGCAAAAC
6	GCGAATAATGAGACCAGCAAAAGCGTGACCTTTCCGAAGGC GGAGGACCTGGATGCGGCGGGTATTGCGGCGAGCGTTAAAC GTGACTACCAGCAAAAC
	Peptide Sequence
7	ANNETSKSVTFPKCEDLDAAGIAASVKRDYQQN
8	ANNETSKSVTFPKAEDLDAAGIAASVKRDYQQN
9	MKYLLPTAAAGLLLLAAQPAMA
10	MKKRGAFLGLLLVSACASVFA
11	MKKTAIAIAVALAGFATVAQA
12	MMKRNILAVIVPALLVAGTANA
13	TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGG CGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCC AGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTC TCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGG GGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCG ACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGG CCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAG TCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACA ACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGG ATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATT TAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTT

	ACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAAC CCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCG CTCATGAATTAATTCTTAGAAAAACTCATCGAGCATCAAATG AAACTGCAATTTATTCATATCAGGATTATCAATACCATATTT TTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGA GGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCG ATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCC TCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGT GACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTT CTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCAT CAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATT GCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGA CAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACAC TGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTC TTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGT GAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGA TGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTG ACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCA TGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAT CGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCC CATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAAT CGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATA ACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATT GTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTG AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAG ATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAA AACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAG CTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCG CAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGC CACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT CTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAG TCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGAT AAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACA GCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACC TACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGG AGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAA CAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTG GTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGA GCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATG GAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTT TTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCT GATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT ACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGT GAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCC TTACGCATCTGTGCGGTATTTCACACCGCATATATGGTGCAC TCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTA

	TACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCC CGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTG TCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTC CGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAA ACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGT GAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCT CGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAA AGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTG ATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGA TACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACT GATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAA ACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCA CTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTG TTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGG AACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTT TACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCA GGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTC GCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCG CCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGC GCACCCGTGGGGCCGCCATGCCGGCGATAATGGCCTGCTTCT CGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGA GCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCC GATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAA TGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATG ATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCC CCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCA AGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAA CTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCG GGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACG CGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTT TTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTC ACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCT GGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAA CGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCC CACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGG TAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAA CCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCA TGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCC GTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTAT GCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAAT GGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGAC CAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGA AAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGA AATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAAT GGCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACT GACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGG

	CTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGG CACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTT GCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCA ATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGG TTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTT TCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACG CGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGAC ATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTG ACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTT GCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTAT GCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGG CCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGA GATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCA TACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGA GCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCA GCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCG TCCGGCGTAGAGGATCGAGATCTATACGAAACGGGAATGCG GTAATTACGCTTTGTTTTTATAAGTCAGATTTTAATTTTTATT GGTTAACATAACGAAAGGTAAAATACATAAGGCTTACTAAA AGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGG TATAAGAATATATACTGATATGTATACCCGAAGTATGTCAAA AAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGA CAGCGACAGCTATCAGTTGCTCAAGGCATATGATGTCAATAT CTCCGGTCTGGTAAGCACAACCATGCAGAATGAAGCCCGTC GTCTGCGTGCCGAACGCTGGAAAGCGGAAAATCAGGAAGGG ATGGCTGAGGTCGCCCGGTTTATTGAAATGAACGGCTCTTTT GCTGACGAGAACAGGGACTGGTGAAATGCAGTTTAAGGTTT ACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATG TACAGAGTGATATTATTGACACGCCCGGGCGACGGATGGTG ATCCCCCTGGCCAGTGCACGTCTGCTGTCAGATAAAGTCTCC CGTGAACTTTACCCGGTGGTGCATATCGGGGATGAAAGCTG GCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGT TATCGGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAAATG ACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATATAA ATGTCAGGCTCCGTTATACACAGCCAGTCTGCAGCGATCCCG CGAAATTTGACAATTAATCATCGGCTCGTATAATGTGTGGAA TTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTT AACTTTAAGAAGGAGATATACATATGATGAAACGTAATATC CTGGCGGTGATTGTTCCGGCGCTGCTGGTTGCGGGCACCGCG AATGCGGCGAATAATGAGACCAGCAAAAGCGTGACCTTTCC GAAGGCGGAGGACCTGGATGCGGCGGGTATTGCGGCGAGCG TTAAACGTGACTACCAGCAAAACGGTGGCAGCGGTGGCAGC GGTAGCCACCATCATCATCACCACAGCAGCGGTGGCAGCGG TACCGACTATAAGGACGATGACGATAAACACGCGGAAGGCA CCTTTACCAGCGATGTGAGCAGCTACCTGGAGGGTCAAGCG GCGAAGGAGTTCATTGCGTGGCTGGTGCGTGGTCGTGGCTA

	ATAGTGAGCGGCCGCGGCTGTTTTGGCGGATGAGAGAAGAT TTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTC TGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTC CCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAG CGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGA ACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGA CTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTC CTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGC GAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCA TAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGAC GGATGGCCTTTTTGCGTTTCTACAAACTCTCTCGAGCACCAC CACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAA GGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAG CATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTT TGCTGAAAGGAGGAACTATATCCGGAT

In another embodiment, the bacterial expression vector further comprises of at least one gene expression cassette comprising of at least one inducible promoter, an RBS, DNA sequence encoding the recombinant protein, DNA sequence encoding an affinity tag, and at least one gene terminator, with the secretory signal sequence operably linked to the gene expression cassette, and the DNA sequence of the affinity tag is operably linked to the DNA sequence of the recombinant protein. The bacterial expression vector additionally comprises of at least one multiple cloning site (MSC) to enable cloning of DNA sequence of recombinant protein under the promoter.

In yet another embodiment, the bacterial expression vector further comprises of at least one antibiotic resistance gene and at least one additional selection marker each controlled by a respective gene promoter.

In yet another embodiment, the bacterial expression vector comprises of at least one *ori* sequence for enabling replication of expression vector in the host cells.

In yet another embodiment, the bacterial expression vector, pBacSec-LC, with the vector being inducible under lac operon, and is induced by Lactose or Lactose analogues including IPTG. As depicted in **Figure 2**, the novel bacterial expression vector, pBacSec-LC, is of around 6793 basepairs comprising of secretory signal

sequence for efficient and enhanced secretion of recombinant protein which is in tandem with the secretory signal sequence. The pBacSec-LC vector comprises of:

tac promoter and lac operator as inducible promoter;

an RBS;

a secretory signal sequence which encodes a combination of signal sequence selected from the group consisting of Seq. ID9-12, and Seq. ID 7 or Seq. ID 8;

DNA sequence encoding 6-His tag and FLAG tag which are affinity tags;

DNA sequence encoding recombinant protein;

a gene terminator for transcriptional termination of recombinant protein;

an *ori* sequence to enable replication of vector in *E. coli;*

a lac operon as a selectable marker for blue-white recombinant colony selection; and

a kanamycin resistance gene as an antibiotic selectable marker.

The pBacSec-LC vector represented Seq. ID 13 comprises of secretory signal sequence which encodes a combination of the signal sequence of the gene *ompF*of Seq. ID 12 and the the truncated yebF of Seq. ID 8.

EXAMPLE 1

SECRETORY EFFICIENCY OF DIFFERENT COMBINATIONS OF SIGNAL SEQUENCE AND CARRIER PEPTIDE

A. Luciferase assay:

E. coli strains, NEB 5-alpha and BL21(DE3) were used for transformation and luciferase assay.

Different combinations of secretory signals were constructed such as:

a) signal sequence of Seq. ID 1 and carrier protein represented by Seq. ID 5 encoding combination of Seq. ID 9 and Seq. ID 7 respectively in tandem,
b) signal sequence of Seq. ID 2 and carrier protein represented by Seq. ID 5 encoding combination of Seq. ID 11 and Seq. ID 7 respectively in tandem,
c) signal sequence of Seq. ID 3 and carrier protein represented by Seq. ID 5 encoding combination of Seq. ID 10 and Seq. ID 7 respectively in tandem, and
d) signal sequence of Seq. ID 4 or an optimized sequence of Seq. ID 4, and carrier protein represented by Seq. ID 5 encoding combination of Seq. ID 12 and Seq. ID 7 respectively in tandem.

Gaussia luciferase were used as reporter system for the examination of secretory activity of *E. coli*. Guassian luciferase assays were performed using Pierce Gaussia Luciferase glow assay kit. Media was collected at indicated time intervals from culture after induction and luciferase activity measured from media as described in manufacturer's protocol.

As depicted in **Figure 3**, combination of every signal peptide with Seq. ID 7 shows secretion amongst which combination of signal peptide of ompF represented by Seq. ID 12 with carrier protein Seq. ID 7 showed better secretory efficiency.

B. Efficiency of secretion of peptides ranging from 5-20 kDa

i. **Construct of expression vector:** DNA sequences of recombinant protein encoding peptides of 5, 10, 15, and 20 kDa were cloned in pBacSec-LC vector having secretory signal sequence comprising of Seq. ID 4 and Seq. ID 6.
ii. **Preparation of starter culture:** 3ml of autoclaved growth media having pH of 6.90 is taken in sterile snap cap tube. Single CFU picked from Luria-Bertani agar plate and inoculated aseptically into growth media and incubated overnight at 37˚C with 225 rpm in rotatory incubator.
iii. **Shake flask culturing:** 1ml of overnight starter culture is inoculated into 25ml of growth media (1:25 dilutions) having pH of 6.90 taken in 250ml baffled flask. Flasks

were incubated in rotatory shaker incubator maintained at 37˚C, 225 rpm. After 4 hours of incubation OD600 reached ~1.8 to 2.0. Cells were induced with 0.2mM Isopropyl β-d-1-thiogalactopyranoside (IPTG), and inducer of lac operon. Post induction, supplements such as glycine, glutamic acid, arginine, reduced glutathione were added. Samples were collected hourly post induction i.e 300µL of culture, centrifuged at 14K rpm for 3 minutes to separate cells and media to check expression and secretion of protein. The cells were lysed to form the lysate. Cell lysate and media components were compared for presence of recombinant peptides.

iv. **SDS-PAGE:** SDS-PAGE was used to analyze the heterologous gene expression of desired recombinant protein from culture media post-induction.

As depicted in **Figure 4**, recombinant peptides of size 5-20 kDa were seen to present in the media component more than the cell lysates clearly indicating secretion of the recombinant peptide. And all the peptides ranging from 5-20 kDa were equally secreted with similar efficiency indicating that the secretory sequence was capable to secrete peptides in the range of 5-20 kDa.

C. The yield of recombinant protein

DNA sequence of recombinant protein was cloned in pBacSec-LC vector with Seq. ID 13, and transformed in *E. coli*. The *E. coli* starter culture was prepared and then upscaled further as described earlier under section B of example 1. The *E. coli* culture was induced with 0.2mM IPTG for recombinant protein expression, and amount of recombinant protein secreted was measured after different time intervals of induction by IPTG.

As depicted in **Figure 5**, the yield of the recombinant protein after 1 hour of induction is around 0.5 g/L, and after 6 hours the yield of the recombinant protein reaches more than 5 g/l. Thus, the average production efficiency is 0.8 g/L/hour. The yield of the

recombinant protein consistently increases after every hour of induction. This suggests that the bacterial secretory technology enables cells to devout maximum resources in protein expression and secretion upon induction after desired cell density is attained. After first hour of induction recombinant protein is secreted into media without stunting cell proliferation mechanism.

EXAMPLE 2

SECRETORY EFFICIENCY OF SEQ. ID 6 COMPARED TO SEQ. ID 5

DNA sequences of recombinant protein were cloned in pBacSec-LC vector having secretory signal sequence comprising of combination of optimized Seq. ID 4 and Seq. ID 5 encoding Seq. ID 12 with Seq. ID 7 in tandem, or optimized Seq. ID 4 and Seq. ID 6 encoding Seq. ID 12 with Seq. ID 8 in tandem.

Seq. ID 6 is synthesized by mutating Seq. ID 5, wherein, TGC codon at position 40 of Seq. ID 5 is mutated to GCG codon in Seq. ID 6, to mutate Cys at position 14 of Seq. ID 7 to Ala in Seq. ID 8. The Cys residue at position 14 of the peptideenables dimerization and increases the chances of development of inclusion bodies. Mutation of Cys to Ala, abrogates the property of dimerization of the peptide thereby reducing the chances of formation of inclusion bodies which in turn should enhance secretion of the peptide and the recombinant peptide.

The *E. coli* cells were transformed with respective vectors and two sets of cultures were prepared under reducing and non-reducing conditions and induction was carried out using IPTG for peptide secretion.

As depicted in **Figure 6**, under reducing conditions Seq. ID 5 encoding Seq. ID 7 showed more secretion as compared to non-reducing conditions, which explained that the dimerization due to Cys residue at position 14 of Seq. ID 7 led to development of inclusion bodies which decreased under reducing conditions. Therefore, mutation of

Cys to Ala at position 14 was bound to be important for secretion, and therefore Seq. ID 6 encoding Seq. ID 8 showed same results under reducing and non-reducing conditions.

While certain exemplary embodiments have been described and shown in the accompanying drawings, it is to be understood that such embodiments are merely illustrative of, and not restrictive on, the broad invention, and that this invention not be limited to the specific constructions and arrangements shown and described, since various other changes, combinations, omissions, modifications and substitutions, in addition to those set forth in the above paragraphs, are possible. Those skilled in the art will appreciate that various adaptations and modifications of the just described embodiments can be configured without departing from the scope and spirit of the invention.