LIPOCALIN FUSION PARTNERS

STATEMENT AS TO FEDERALLY SPONSORED RESEARCH

[0001] This invention was made with the support of the United States government under Contract numbers R01CA135491, NIH ROI AI059543, AI094419, and AI097786 by the National Institutes of Health.

SEQUENCE LISTING

[0001.1] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on February 24, 2014, is named 44189-701.602_SL.txt and is 277,591 bytes in size.

BACKGROUND OF THE INVENTION

[0002] Efforts toward drug discovery continue to use vast technical and financial resources to identify and develop new and useful drugs. Unfortunately, finding new drugs has continued to be difficult. For example, development of less damaging, more precisely targeted cancer therapies is essential. But even after decades of research, scientists still struggle to identify therapeutic compounds with the right mix of medicinal and cancer-targeting properties which not only reduce the likelihood that a therapeutic compound could serve as a treatment strategy but creates a need for successful methods of surgical resection.

[0003] A wide variety of types of compounds have been studied and pursued for a large breadth of therapeutic purposes. For example, small chemical molecules and larger biologics (e.g., antibodies) have been used for a plethora of therapeutic applications with varied success. Some smaller peptides have also been shown to be useful as drugs, e.g., by virtue of their natural potency.

[0004] A lack of methods for rapid and efficient production of peptides and proteins for clinical applications has limited the discovery of peptides and proteins that might serve as therapeutic compounds. Creating fusions of peptides, protein domains, or proteins, to a different protein to enhance production of peptides and proteins has been explored in bacterial expression systems, e.g., *E. coli* gene expression. However, bacterial protein expression systems are generally limited due to errors in protein folding efficiency. Thus, there is still a need for methods enabling the rapid and efficient production of peptides, protein domains, and proteins.

SUMMARY OF THE INVENTION

[0005] The present invention relates to methods for the production of siderocalin fusion proteins. In various aspects, the present invention relates to the fusion proteins produced according to those methods. In certain apsects, the fusion protein is cleaved, thereby producing a peptide according to the present disclosure.The present invention further relates to methods for producing a fusion protein, the methods comprising expressing, in a cell, a fusion protein, the fusion protein comprising a peptide or protein domain and a lipocalin protein, thereby producing the fusion

[0015] FIG. 7 provides a schematic of pooled library production, in accordance with an embodiment of the present invention. FIG. 7B shows an expanded view of FIG. 7A. FIG. 7A discloses SEQ ID NOS 121-124, respectively, in order of appearance. FIG. 7B discloses SEQ ID NOS 125-126, respectively, in order of appearance.

[0016] FIG. 8 describes representative sequencing data from a cloned knottin library, in accordance with an embodiment of the present invention. FIG. 8 discloses SEQ ID NOS 127-155, respectively, in order of appearance.

[0017] FIG. 9 shows SDS-PAGE analysis of 3000-member knottin libraries, in accordance with embodiments of the present invention.

[0018] FIG. 10 shows a generalized process for manufacturing knottins, in accordance with an aspect of the present disclosure. FIG. 10 discloses "6xHis" as SEQ ID NO: 93.

[0019] FIG. 11 shows a generalized process for manufacturing knottins, in accordance with an aspect of the present disclosure.

[0020] FIG. 12 a generalized process for manufacturing knottins, in accordance with an aspect of the present disclosure. According to this aspect, cleavage occurs at the furin-knottin interface.

[0021] FIG. 13 shows examples of knottins made with the Daedalus system and corresponding SDS PAGE analyses in accordance with an aspect of the present disclosure. The Daedalus system is desribed in Bandaranayake A.D., et al., *Nucleic Acids Res.* (2011) 39(21):e143.

[0022] FIG. 14 shows the alignment of Scn sequences from 18 species (SEQ ID NOS 39, 40, 47, 51, 45, 49, 43, 44, 52, 54, 42, 156, 46, 157, 50, 53, 48, and 41, respectively, in order of appearance). The alignment shows a high level of sequence conservation between the 18 species with recognizable orthologs. Positions with an asterisk are possible sites for ligand binding. Correnti, C. & Strong, R. K. (2013) 'Iron sequestration in immunity' In Metals in Cells; Encyclopedia of Inorganic and Bioinorganic Chemistry. (Culotta, V. & Scott, R. A., eds.) John Wiley & Sons, pp. 349-59

[0023] FIG. 15 depicts a 3D model of super-stable Scn in accordance with an aspect of the present disclosure. In this form of Scn, a second disulfide bond was engineered in order to secure the N-terminus and increase thermal stability.

[0024] FIG. 16 shows a schematic of a generic Scn fusion that can be used in various aspects of the present disclosure. The native signal peptide was been removed and an exogenous sFLAG and HIS tag were added to facilitate purification. These modifications are optionally present in various aspects of the disclosure.

[0025] FIG. 17 depicts a schematic of a light chain antibody fusion with Scn and corresponding SDS PAGE analysis according to an aspect of the present disclosure. FIG. 17 discloses SEQ ID NO: 158.

[0026] FIG. 18 shows the SDS PAGE analysis comparing the expression of a Scn fusion with a cytoplasmic enzyme, HMOX1, both before and after cleavage according to an aspect of the present dislcosure.

[0027] FIG. 19 depicts the expression of a Scn fusion with a cytoplasmic viral protein Adv2 and the corresponding SDS PAGE analysis according to one aspect of the present disclosure.

protein is siderocalin. In certain embodiments, the lipocalin protein is human siderocalin, murine siderocalin, chicken Ex-FABP, or quail Q86.

[0060] In some embodiments, the expression of peptides (e.g., knotted peptides) as fusions with SCN can be utilized with an endogenously cleaving SCN, with RARYKR (SEQ ID NO: 98) right after the CIDG (SEQ ID NO: 101), and an exogenously cleaved one, with ENLYFQ (SEQ ID NO: 95) in that position. The former can be cleaved by the mammalian cells during protein export (e.g., by furin), and the free SCN and knotted peptide can be secreted into surrounding media. ENLYFQ (SEQ ID NO: 95) is a tobacco etch virus (TEV) protease site, which is not found endogenously in mammalian cells. The constructs in this system can be secreted as fusions, allowing for the knotted peptide to be cleaved off later by adding exogenous TEV protease. This can be useful for recovering the knottins. In some embodiments, purification "handles" such as poly-histidine or poly-arginine can be added to the SCN and subsequently removed by proteolysis. In addition to the knotted peptides, these fusion systems can also used for difficult-to-express proteins of medical interest such as chemokines, interleukins, and peptide hormones.

[0061] In some embodiments, the peptide comprises a knotted-peptide. In certain embodiments, the knotted-peptide is selected from the group consisting of chymotrypsin inhibitor, hefutoxin, bubble protein, the C-terminal domain of midkine, and epiregulin. In some embodiments, the lipocalin protein comprises siderocalin.

[0062] The lipocalin fusions (e.g., siderocalin and/or Lcn2 fused with a knotted peptide) can be used in several ways different ways. It could be used to increase the size of the target protein (for example a potential therapeutic) in order to increase its half-life. It could be used to secrete the target protein where the target protein is naturally expressed in the cytoplasm. SCN also has unique ligand specificity and tightly binds catecholate siderophores (bacterial iron chelators). This opens the possibility of loading the SCN fusion with specific ligands, such as a chemotherapeutic or radioactive reagent or some type or a compound that has beneficial properties. SCN, when loaded with siderophores and iron, has a deep red color that can aid in chromatography or other purification steps.

[0063] The lipocalin fusions of the present disclosure (e.g., SCN fused with HO-1) can be used in a variety of ways. For example, fusions can be used to increase the size of the target protein (e.g., a potential therapeutic) in order to increase its biological half-life. Fusions can be used to secrete the target protein where the target protein is naturally expressed in the cytoplasm. Fusions can also be used to target the fusion partner protein to specific locations to maximize therapeutic effects. SCN also has unique ligand specificity, tightly binding catecholate

dimerization at a cysteine residue (see Goetz, D. H., et al., 'The Neutrophil Lipocalin NGAL is a Bacteriostatic Agent that Interferes with Siderophore-mediated Iron Acquisition' *Molecular Cell* (2002) 10: 1033-43).

[0075] In some aspects, at least one of the amino acids in the native sequence of the non-human lipocalin protein orthologous to the human SCN may be substituted for a non-native amino acid. For example, non-human orthologs of the human lipocalin protein that may be used with the methods and compositions described herein include, but are not limited to, murine Lcn2 (e.g., 24p3), Lcn1, Lcn6, Lcn8, Lcn9, Lcn10, Lcn12 and Lcn15 (see FIG. 14). In some aspects, the mutations may be generated to prevent non-human lipocalin protein orthologous to the human SCN from dimerizing. For example, one non-human lipocalin protein orthologous to the human SCN protein may dimerize with another non-human lipocalin protein orthologous to the human SCN protein at cysteine residues or one non-human lipocalin protein orthologous to the human SCN protein may dimerize with a different protein at cysteine residues. For example, generating a mutation at a site orthologous to the C87S mutation in human lipocalin protein may prevent dimeraizaiton at a cysteine residue (see Goetz, D. H., et al., 'The Neutrophil Lipocalin NGAL is a Bacteriostatic Agent that Interferes with Siderophore-mediated Iron Acquisition' *Molecular Cell* (2002) 10: 1033-43).

[0076] In an exemplary aspect, a lipocalin fusion protein may contain the following protein sequence: MPLGLLWLGLALLGALHAQAQDSTSDLIPAPPLSKVPLQQNFQDNQFQGKWYVVGLAGNAILRE DKDPQKMYATIYELKEDKSYNVTSVLFRKKKCDYWIRTFVPGCQPGEFTLGNIKSYPGLTSYLVR VVSTNYNQHAMVFFKKVSQNREYFKITLYGRTKELTSELKENFIRFSKSLGLPENHIVFPVPIDQCI DG (SEQ ID NO: 102)

In this sequence, the annotations indicate, a signal peptide, **glycosylation site**, **disulfide bond** and the C87S mutation.

[0077] In some aspects, fusion of a protein or peptide of interest to lipocalin may improve the biological properties of the target protein (e.g., a potential therapeutic). For example, fusion of a protein or peptide of interest to lipocalin may increase the half-life of the protein or peptide of interest by increasing the size of the overall protein. For example, the increased size of the protein may prevent glomerular filtration of the protein or peptide of interest. In some aspects, fusions of small proteins, such as antibody fragments, exhibit decreased glomerular filtration. These effects are observed with both enzymes and Fab antibody fragments. In some embodiments, the fusion protein further comprises a peptide selected from the group consisting of: an IgK starter sequence, a sFLAG, a HIS, siderocalin, a TEV and the knotted peptide

detection and localization could be performed either *in vitro* or *in vivo*. In some aspects, the addition of SCN ligands to the fusion proteins could be used to purify fusion proteins from mixtures. For example, an SCN ligand (e.g., a siderophore/metal complex) could be contacted with at least one purification resin metal other than, including but not limited to, aluminum, gadolinium, indium, vanadium, plutonium or thorium, and any related isotopes.

[0084] In some apects, the addition of SCN ligands to the fusion proteins could be used for the delivery of radionuclides to a target tissue. In some aspects, the addition of SCN ligands to the fusion proteins could be used to deliver iron to a target. In addition, SCN bound to siderophores and iron can be a dark red color. For example, the dark red color can be combined with methods and compositions of chromatography or other steps in a method of purification.

[0085] In an exemplary apect, a self-cleaving SCN isoform may be added to a peptide or protein of interest to generate a fusion protein. For example, the self-cleaving SCN isoform may contain the RARYKR (SEQ ID NO: 98) amino acid sequence immediately following the CIDG (SEQ ID NO: 101) amino acid sequence. In this case, the RARYKR (SEQ ID NO: 98) sequence may be cleaved by an endogenous enzyme to the cells of the protein expression system (e.g., the mammalian cells) during export of the fusion protein. For example, furin may cleave the RARYKR (SEQ ID NO: 98) sequence. In this case, SCN and the peptide or protein of interest may be free and located in the extracellular space.

[0086] In an exemplary apect, an exogenously cleaved SCN isoform may be added to a peptide or protein of interest to generate a fusion protein. For example, the exogenously cleaved SCN isoform may contain the ENLYFQ (SEQ ID NO: 95) amino acid sequence immediately following the CIDG (SEQ ID NO: 101) amino acid sequence. In this case, the ENLYFQ (SEQ ID NO: 95) sequence may be cleaved by an exogenous enzyme to the cells of the protein expression system (e.g., the mammalian cells) during export or after export of the fusion protein. For example, a tobacco etch protease may cleave the ENLYFQ (SEQ ID NO: 95) site.

[0087] In some apects, the fusion proteins may be secreted from the cells of the protein expression system (e.g., the mammalian cells) as fusion proteins. In this case, the peptide or protein of interest may be cleaved from the SCN protein by TEV protease. For example, the TEV protease may be added to the cells of the protein expression system (e.g., the mammalian cells) or added after removal of the fusion protein from the cells.

[0088] In some apects, the SCN peptide may be modified to include compounds for purification or isolation. In some aspects, the compounds may be an amino acid or more than one amino acid. For example, the compounds may be poly-histidine or poly-arginine and may be located between Lcn2 and a signal peptide in the fusion protein. In some aspects, the compounds may be removed from Lcn2 using an enzyme or proteolysis.

shown, chymotrypsin inhibitor (CTI), epiregulin (EPI), hefutoxin (HTX), bubble protein (BUB), potato carboxypeptidase inhibitor (PCI) were properly folded.

[00120] FIG. 9 shows a schematic describing production of a pooled library of knottins. In this example, sequences of thousands of knottins can be encoded in an oligonucleotide pool (1) and selectively amplified using unique primer pairs (2). DNA sublibraries can be cloned into the expression vector, which results in the knottin variants that can, e.g., have unique parental mass signatures and unique tryptic fragment mass signatures that can be resolvable using current techniques, such as mass spectroscopy.

[00121] FIG. 8 includes example knottin variants that describe representative sequencing from a cloned knottin library. The sequences show raw sequencing data from a single round of library cloning. The sequence portions highlighted in grey are full length knottin variants, and the errors in oligonucleotide synthesis can explain the truncated and extended peptide sequences.

[00122] Using the methods described in this example, variants of several knottin scaffolds were generated and analyzed. FIG. 9 shows an SDS-PAGE analysis of 3000 member knottin libraries for, e.g., hefutoxin, knottin and chymotrypsin inhibitor. Each column of the SDS-PAGE gel shows a purified sample of a pool of 3000 knottin protein variants run under native and reducing conditions. The migration shift between the paired bands indicates disulfide formation.

[00123] Four scaffolds were selected for the generation of defined libraries: hefutoxin, CTI, knottin, and epiregulin. A list of target amino acid sequences was generated in silico such that every member of each library would have a tryptic fragment with a unique mass; mutations were selected to be structurally adjacent in order to generate binding epitopes. The cysteines were not mutated, and lysine was specifically avoided in order to make N-terminal conjugation unambiguous. 3000 variants of each scaffold were generated, and each scaffold was flanked by a unique set of PCR primer sites so that each of the four sublibraries could be amplified independently. All constructs had an N-terminal BamHI site and a C-terminal NotI site, and following PCR amplification of each sublibrary from the pool of 12000 oligonucleotides, each sublibrary was restriction digested and cloned into cut parental vector (both the furin-cleaved and TEV-cleaved versions) as an SCN fusion protein using standard techniques. HEK293 cells were transfected with this plasmid library as well as the accessory plasmids needed for Daedalus expression, and the virus in the media harvested 3-4 days later. Virus was concentrated by centrifugation and used to infect HEK293 cells for protein production using standard procedures. We have found that the TEV-cleavable construct is technically easier to handle when producing libraries because it allows for facile recovery of the fusion by IMAC on nickel resin. Following IMAC, the fusion protein was dialyzed into PBS and allowed to cleave overnight with 6xHis (SEQ ID NO: 93)

[00128] Siderocalin can be used according to the present disclosure for construction of a Scn-peptide fusion, which can be secreted and cleaved for the efficient production of excreted peptides. FIG. 24 depicts the crystal structure of Scn with a Th ligand.

[00129] FIG. 14 shows the alignment of Scn sequences from 18 species. The alignment shows a high level of sequence conservation between the 18 species with recognizable orthologs. Positions with an asterisk are possible cites for ligand binding.

[00130] FIG. 15 depicts a 3D model of super-stable Scn in accordance with an aspect of the present disclosure. In this form of Scn, a second disulfide bond was engineered in order to secure the N-terminus and increase thermal stability.

[00131] FIGS. 10-12 depict various manufacturing methods according to the present disclosure, which enable the efficient production of knottin peptides. FIG. 10 depicts the elements of the siderocalin fusion protein and use thereof for the production of knottin peptides. In this example, the knottin is secreted as a fusion protein to the siderocalin. The IgK signal peptide is derived from the mouse Light Chain IgG. The sFLAG is the short FLAG (DYKDE) (SEQ ID NO: 91) for enhanced cleavage of the signal peptide relative to the long FLAG signal peptide. The His is a histidine tag (HHHHHH) (SEQ ID NO: 93) and TEV is the tobacco etch virus protease site (ENLYFQ) (SEQ ID NO: 95). According to this method, and as depicted in FIG 12., the fusion protein is secreted, followed by cleavage and isolation. IgK SP is Murine IgK light chain signal peptide, sFLAG: Shortened FLAG epitope, HIS: 6xhistidine tag (SEQ ID NO: 93), TEV: Tobacco Etch Virus Protease recognition site and Furin – furin cleavage site with BamHI site.

[00132] FIG. 10 depicts the elements of the siderocalin fusion protein and the use thereof for the production of knottin peptides. In this example, the knottin is secreted as a fusion protein to HIS-tagged siderocalin. The HIS tag is six histidines (HHHHHH) (SEQ ID NO: 93) that will reversibly bind nickel and is therefore a useful purification tag for Immobilized Metal Chromatography (IMAC). Following isolation of the fusion protein by IMAC, the fusion is cleaved with TEV protease and the cleaved siderocalin removed by another round of IMAC, leaving pure knottin.

[00133] FIG. 11 depicts the elements of the siderocalin fusion protein and use thereof for the production of knottin peptides. This method includes the secretion of a siderocalin fusion with concomitant, intracellular cleavage at the Furin-His junction and purification by IMAC. The HIS tag can then be removed from the knottin with TEV protease. The secretion method depicted in FIG. 11 utilizes the same restriction sites as the method depicted in FIG. 10.

[00134] FIG. 12 depicts the elements of a system for the secretion of cleaved knottin peptides. The siderocalin in this case is present during translation but is cleaved off by

SEQUENCES

[00146] The following are DNA and/or amino acid sequences of genes of interest and constructs identified herein.

Construction of parental construct for seamless cloning:

SEQ ID NO: 1 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-PARENTAL *for Xho/Bam cut from pUC57 and ligation into pCVL*

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAGGGA**GGCGGCCGC**TAAGGATCCCGGACCGCCTCTCC

NotI cut is AACCTGTATTTTCAGGGA**GGC** (SEQ ID NO: 103)

GCTAAGGATCCCGGACCGCCTCTCC (SEQ ID NO: 104)

Fusion protein Sequences – original set of 10 – cloned into NotI cut parent above by seamless cloning:

SEQ ID NO: 2 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-BubbleProtein

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAGGGA**GGCGATACCTGCGGCAGCGGCTATAATGTGGATCAGCGTCGTACCAATAG**
CGGCTGCAAAGCGGGCAATGGCGATCGTCATTTTTGCGGCTGCGATCGTACCGGCGTG

GTGGAATGCAAAGGCGGCAAATGGACCGAAGTGCAGGATTGCGGCAGCAGCAGCTGC AAAGGCACCAGCAATGGCGGCGCGACCTGCTAATGCTAAGGATCCCGGA

SEQ ID NOS: 3 and 105 -

atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt

M E T D T L L L W V L L L W V P G S T G

gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac

D Y K D E H H H H H H G G S Q D S T S D

ctgatcccagccccacctctgagcaaggtcctctgcagcagaacttccaggacaaccaa

L I P A P P L S K V P L Q Q N F Q D N Q

ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa

F Q G K W Y V V G L A G N A I L R E D K

gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc

D P Q K M Y A T I Y E L K E D K S Y N V

Acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt

T S V L F R K K K C D Y W I R T F V P G

tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac

C Q P G E F T L G N I K S Y P G L T S Y

ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt

L V R V V S T N Y N Q H A M V F F K K V

tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg

S Q N R E Y F K I T L Y G R T K E L T S

gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc

E L K E N F I R F S K S L G L P E N H I

gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt

V F P V P I D Q C I D G G G S E N L Y F

cagggaggcgatacctgcggcagcggctataatgtggatcagcgtcgtaccaatagcggc

Q G G D T C G S G Y N V D Q R R T N S G

tgcaaagcgggcaatggcgatcgtcatttttgcggctgcgatcgtaccggcgtggtggaa

C K A G N G D R H F C G C D R T G V V E

tgcaaaggcggcaaatggaccgaagtgcaggattgcggcagcagcagctgcaaaggcacc

C K G G K W T E V Q D C G S S S C K G T

agcaatggcggcgcgacctgc

S N G G A T C

SEQ ID NO: 4 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-Attractin

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC

TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG

CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA

ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT

CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA

AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT

TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT

AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA

AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT

GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC

ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT

TTTCAGGGA**GGCGATCAGAATTGCGATATTGGCAATATTACCAGCCAGTGCCAGATGCA**

GCATAAAAATTGCGAAGATGCGAATGGCTGCGATACCATTATTGAAGAATGCAAAACC

AGCATGGTGGAACGTTGCCAGAATCAGGAATTTGAAAGCGCGGCGGGCAGCACCACCC

TGGGCCCGCAGTAATGCTAAGGATCCCGGA

SEQ ID NOS: 5 and 106-

atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt

M E T D T L L L W V L L L W V P G S T G

gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac

D Y K D E H H H H H H G G S Q D S T S D

ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa

L I P A P P L S K V P L Q Q N F Q D N Q

ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa

F Q G K W Y V V G L A G N A I L R E D K

gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc

D P Q K M Y A T I Y E L K E D K S Y N V

acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt

T S V L F R K K K C D Y W I R T F V P G

tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac

C Q P G E F T L G N I K S Y P G L T S Y

ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt

L V R V V S T N Y N Q H A M V F F K K V

tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg

S Q N R E Y F K I T L Y G R T K E L T S

gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc

E L K E N F I R F S K S L G L P E N H I

```
gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt
 V  F  P  V  P  I  D  Q  C  I  D  G  G  G  S  E  N  L  Y  F
cagggaggcgatcagaattgcgatattggcaatattaccagccagtgccagatgcagcat
 Q  G  G  D  Q  N  C  D  I  G  N  I  T  S  Q  C  Q  M  Q  H
aaaaattgcgaagatgcgaatggctgcgataccattattgaagaatgcaaaaccagcatg
 K  N  C  E  D  A  N  G  C  D  T  I  I  E  E  C  K  T  S  M
gtggaacgttgccagaatcaggaatttgaaagcgcggcgggcagcaccaccctgggcccg
 V  E  R  C  Q  N  Q  E  F  E  S  A  A  G  S  T  T  L  G  P
cag
 Q
```

SEQ ID NO: 6 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-Hefutoxin

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAGGGA**GGCGGCCATGCGTGCTATCGTAATTGCTGGCGTGAAGGCAATGATGAAGA**
AACCTGCAAAGAACGTTGCTAATGCTAAGGATCCCGGA*ACCGCC*

SEQ ID NOS: 7 and 107-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
 D  Y  K  D  E  H  H  H  H  H  H  G  G  S  Q  D  S  T  S  D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
 L  I  P  A  P  P  L  S  K  V  P  L  Q  Q  N  F  Q  D  N  Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
 F  Q  G  K  W  Y  V  V  G  L  A  G  N  A  I  L  R  E  D  K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
 D  P  Q  K  M  Y  A  T  I  Y  E  L  K  E  D  K  S  Y  N  V
acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt
 T  S  V  L  F  R  K  K  K  C  D  Y  W  I  R  T  F  V  P  G
```

tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac

C Q P G E F T L G N I K S Y P G L T S Y

ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt

L V R V V S T N Y N Q H A M V F F K K V

tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg

S Q N R E Y F K I T L Y G R T K E L T S

gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc

E L K E N F I R F S K S L G L P E N H I

gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt

V F P V P I D Q C I D G G G S E N L Y F

cagggaggcggccatgcgtgctatcgtaattgctggcgtgaaggcaatgatgaagaaacc

Q G G G H A C Y R N C W R E G N D E E T

tgcaaagaacgttgc

C K E R C

SEQ ID NO: 8 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-Hanatoxin

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGC

TGCTCTGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACC

ATGGTGGAAGCCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGG

TCCCTCTGCAGCAGAACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTA
GGCCTGGCAGGGAATGCAATTCTCAGAGAAGACAAAGACCCGCAAAAGATGTATGC
CACCATCTATGAGCTGAAAGAAGACAAGAGCTACAATGTCACCTCCGTCCTGTTTAG
GAAAAAGAAGTGTGACTACTGGATCAGGACTTTTGTTCCAGGTTGCCAGCCCGGCGA
GTTCACGCTGGGCAACATTAAGAGTTACCCTGGATTAACGAGTTACCTCGTCCGAGT
GGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCAAGAAAGTTTCTCAAAA
CAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCTGACTTCGGAAC
TAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACCACATCG
TCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTATT
TTCAGGGA**GGC**<u>**GAATGCCGTTATCTGTTTGGCGGCTGCAAAACCACCAGCGATT**
GCTGCAAACATCTGGGCTGCAAATTTCGTGATAAATATTGCGCGTGGGATTTTA
CCTTTAGC</u>**TAATGCTAAGGAT**CCCGGA

SEQ ID NOS: 9 and 108-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
 D  Y  K  D  E  H  H  H  H  H  H  G  G  S  Q  D  S  T  S  D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
 L  I  P  A  P  P  L  S  K  V  P  L  Q  Q  N  F  Q  D  N  Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
 F  Q  G  K  W  Y  V  V  G  L  A  G  N  A  I  L  R  E  D  K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
 D  P  Q  K  M  Y  A  T  I  Y  E  L  K  E  D  K  S  Y  N  V
acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt
 T  S  V  L  F  R  K  K  K  C  D  Y  W  I  R  T  F  V  P  G
tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac
 C  Q  P  G  E  F  T  L  G  N  I  K  S  Y  P  G  L  T  S  Y
ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt
 L  V  R  V  V  S  T  N  Y  N  Q  H  A  M  V  F  F  K  K  V
tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg
 S  Q  N  R  E  Y  F  K  I  T  L  Y  G  R  T  K  E  L  T  S
gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc
 E  L  K  E  N  F  I  R  F  S  K  S  L  G  L  P  E  N  H  I
gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt
 V  F  P  V  P  I  D  Q  C  I  D  G  G  G  S  E  N  L  Y  F
cagggaggcgaatgccgttatctgtttggcggctgcaaaaccaccagcgattgctgcaaa
 Q  G  G  E  C  R  Y  L  F  G  G  C  K  T  T  S  D  C  C  K
catctgggctgcaaatttcgtgataaatattgcgcgtgggattttacctttagc
```

H L G C K F R D K Y C A W D F T F S

SEQ ID NO: 10 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-ChymotrypsinInhibitor

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAGGGA**GGCGAAATTAGCTGCGAACCGGGCAAAACCTTTAAAGATAAATGCAATAC**
CTGCCGTTGCGGCGCGGATGGCAAAAGCGCGGCGTGCACCCTGAAAGCGTGCCCGAAT
CAGTAATGCTAAGGATCCCGGA

SEQ ID NOS: 11 and 109-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
 D  Y  K  D  E  H  H  H  H  H  H  G  G  S  Q  D  S  T  S  D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
 L  I  P  A  P  P  L  S  K  V  P  L  Q  Q  N  F  Q  D  N  Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
 F  Q  G  K  W  Y  V  V  G  L  A  G  N  A  I  L  R  E  D  K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
 D  P  Q  K  M  Y  A  T  I  Y  E  L  K  E  D  K  S  Y  N  V
acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt
 T  S  V  L  F  R  K  K  K  C  D  Y  W  I  R  T  F  V  P  G
tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac
 C  Q  P  G  E  F  T  L  G  N  I  K  S  Y  P  G  L  T  S  Y
ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt
 L  V  R  V  V  S  T  N  Y  N  Q  H  A  M  V  F  F  K  K  V
tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg
 S  Q  N  R  E  Y  F  K  I  T  L  Y  G  R  T  K  E  L  T  S
gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc
 E  L  K  E  N  F  I  R  F  S  K  S  L  G  L  P  E  N  H  I
gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt
 V  F  P  V  P  I  D  Q  C  I  D  G  G  G  S  E  N  L  Y  F
cagggaggcgaaattagctgcgaaccgggcaaaacctttaaagataaatgcaatacctgc
 Q  G  G  E  I  S  C  E  P  G  K  T  F  K  D  K  C  N  T  C
cgttgcggcgcggatggcaaaagcgcggcgtgcaccctgaaagcgtgcccgaatcag
 R  C  G  A  D  G  K  S  A  A  C  T  L  K  A  C  P  N  Q
```

SEQ ID NO: 12- IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-ToxinK

```
GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGC
TGCTCTGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACC
ATGGTGGAAGCCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAG
GTCCCTCTGCAGCAGAACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTA
GGCCTGGCAGGGAATGCAATTCTCAGAGAAGACAAAGACCCGCAAAAGATGTATGC
CACCATCTATGAGCTGAAAGAAGACAAGAGCTACAATGTCACCTCCGTCCTGTTTAG
GAAAAAGAAGTGTGACTACTGGATCAGGACTTTTGTTCCAGGTTGCCAGCCCGGCGA
GTTCACGCTGGGCAACATTAAGAGTTACCCTGGATTAACGAGTTACCTCGTCCGAGT
GGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCAAGAAAGTTTCTCAAAA
CAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCTGACTTCGGAAC
```

TAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACCACATCG
TCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTATT
TTCAGGGA**GGCGTGTGCCGTGATTGGTTTAAAGAAACCGCGTGCCGTCATGCGA**
AAAGCCTGGGCAATTGCCGTACCAGCCAGAAATATCGTGCGAATTGCGCGAAA
ACCTGCGAACTGTGCTAATGCTAAGGATCCCGGA

SEQ ID NOS: 13 and 110-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
 D  Y  K  D  E  H  H  H  H  H  H  G  G  S  Q  D  S  T  S  D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
 L  I  P  A  P  P  L  S  K  V  P  L  Q  Q  N  F  Q  D  N  Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
 F  Q  G  K  W  Y  V  V  G  L  A  G  N  A  I  L  R  E  D  K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
 D  P  Q  K  M  Y  A  T  I  Y  E  L  K  E  D  K  S  Y  N  V
acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt
 T  S  V  L  F  R  K  K  K  C  D  Y  W  I  R  T  F  V  P  G
tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac
 C  Q  P  G  E  F  T  L  G  N  I  K  S  Y  P  G  L  T  S  Y
```

```
ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt
 L  V  R  V  V  S  T  N  Y  N  Q  H  A  M  V  F  F  K  K  V
tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg
 S  Q  N  R  E  Y  F  K  I  T  L  Y  G  R  T  K  E  L  T  S
gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc
 E  L  K  E  N  F  I  R  F  S  K  S  L  G  L  P  E  N  H  I
gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt
 V  F  P  V  P  I  D  Q  C  I  D  G  G  G  S  E  N  L  Y  F
cagggaggcgtgtgccgtgattggtttaaagaaaccgcgtgccgtcatgcgaaaagcctg
 Q  G  G  V  C  R  D  W  F  K  E  T  A  C  R  H  A  K  S  L
ggcaattgccgtaccagccagaaatatcgtgcgaattgcgcgaaaacctgcgaactgtgc
 G  N  C  R  T  S  Q  K  Y  R  A  N  C  A  K  T  C  E  L  C
```

SEQ ID NO: 14 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-EGFepiregulinCore

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAGGGA**GGC**<u>**GTGAGCATTACCAAATGCAGCAGCGATATGAATGGCTATTGCCTGCA**
TGGCCAGTGCATTTATCTGGTGGATATGAGCCAGAATTATTGCCGTTGCGAAGTGGGC
TATACCGGCGTGCGTTGCGAACATTTTTTTCTG</u>**TAATGCTAAGGAT**CCCGGA

SEQ ID NOS: 15 and 111-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
 D  Y  K  D  E  H  H  H  H  H  H  G  G  S  Q  D  S  T  S  D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
 L  I  P  A  P  P  L  S  K  V  P  L  Q  Q  N  F  Q  D  N  Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
 F  Q  G  K  W  Y  V  V  G  L  A  G  N  A  I  L  R  E  D  K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
```

D P Q K M Y A T I Y E L K E D K S Y N V

acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt

T S V L F R K K K C D Y W I R T F V P G

tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac

C Q P G E F T L G N I K S Y P G L T S Y

ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt

L V R V V S T N Y N Q H A M V F F K K V

tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg

S Q N R E Y F K I T L Y G R T K E L T S

gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc

E L K E N F I R F S K S L G L P E N H I

gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt

V F P V P I D Q C I D G G G S E N L Y F

cagggaggcgtgagcattaccaaatgcagcagcgatatgaatggctattgcctgcatggc

Q G G V S I T K C S S D M N G Y C L H G

cagtgcatttatctggtggatatgagccagaattattgccgttgcgaagtgggctatacc

Q C I Y L V D M S Q N Y C R C E V G Y T

Ggcgtgcgttgcgaacatttttttctg

G V R C E H F F L

SEQ ID NO: 16 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-Circulin

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAGGGA**GGCGGCATTCCGTGCGGCGAAAGCTGCGTGTGGATTCCGTGCATTAGCG**
CGGCGCTGGGCTGCAGCTGCAAAAATAAAGTGTGCTATCGTAATTAATGCTAAGGATCC
CGGA

SEQ ID NOS: 17 and 112-

atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
M E T D T L L L W V L L L W V P G S T G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
D Y K D E H H H H H H G G S Q D S T S D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
L I P A P P L S K V P L Q Q N F Q D N Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
F Q G K W Y V V G L A G N A I L R E D K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
D P Q K M Y A T I Y E L K E D K S Y N V
acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt
T S V L F R K K K C D Y W I R T F V P G
tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac
C Q P G E F T L G N I K S Y P G L T S Y
ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt
L V R V V S T N Y N Q H A M V F F K K V
tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg
S Q N R E Y F K I T L Y G R T K E L T S
gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc
E L K E N F I R F S K S L G L P E N H I

gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt

V F P V P I D Q C I D G G G S E N L Y F

cagggaggcggcattccgtgcggcgaaagctgcgtgtggattccgtgcattagcgcggcg

Q G G G I P C G E S C V W I P C I S A A

ctgggctgcagctgcaaaaataaagtgtgctatcgtaat

L G C S C K N K V C Y R N

SEQ ID NO: 18 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-Brazzein

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGC

TGCTCTGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACC

ATGGTGGAAGCCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAG

GTCCCTCTGCAGCAGAACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTA

GGCCTGGCAGGGAATGCAATTCTCAGAGAAGACAAAGACCCGCAAAAGATGTATGC

CACCATCTATGAGCTGAAAGAAGACAAGAGCTACAATGTCACCTCCGTCCTGTTTAG

GAAAAAGAAGTGTGACTACTGGATCAGGACTTTTGTTCCAGGTTGCCAGCCCGGCGA

GTTCACGCTGGGCAACATTAAGAGTTACCCTGGATTAACGAGTTACCTCGTCCGAGT

GGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCAAGAAAGTTTCTCAAAA

CAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCTGACTTCGGAAC

TAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACCACATCG

TCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTATTTTCAG
GGA**GGCCAGGATAAATGCAAAAAAGTGTATGAAAATTATCCGGTGAGCAAATGCCAGC**
TGGCGAATCAGTGCAATTATGATTGCAAACTGGATAAACATGCGCGTAGCGGCGAATG
CTTTTATGATGAAAAACGTAATCTGCAGTGCATTTGCGATTATTGCGAATATTAATGCT
AAGGATCCCGGA

SEQ ID NOS: 19 and 113-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
 D  Y  K  D  E  H  H  H  H  H  H  G  G  S  Q  D  S  T  S  D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
 L  I  P  A  P  P  L  S  K  V  P  L  Q  Q  N  F  Q  D  N  Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
 F  Q  G  K  W  Y  V  V  G  L  A  G  N  A  I  L  R  E  D  K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
 D  P  Q  K  M  Y  A  T  I  Y  E  L  K  E  D  K  S  Y  N  V
acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt
 T  S  V  L  F  R  K  K  K  C  D  Y  W  I  R  T  F  V  P  G
tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac
 C  Q  P  G  E  F  T  L  G  N  I  K  S  Y  P  G  L  T  S  Y
ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt
 L  V  R  V  V  S  T  N  Y  N  Q  H  A  M  V  F  F  K  K  V
tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg
 S  Q  N  R  E  Y  F  K  I  T  L  Y  G  R  T  K  E  L  T  S
gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc
 E  L  K  E  N  F  I  R  F  S  K  S  L  G  L  P  E  N  H  I
gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt
 V  F  P  V  P  I  D  Q  C  I  D  G  G  G  S  E  N  L  Y  F
cagggaggccaggataaatgcaaaaaagtgtatgaaaattatccggtgagcaaatgccag
 Q  G  G  Q  D  K  C  K  K  V  Y  E  N  Y  P  V  S  K  C  Q
ctggcgaatcagtgcaattatgattgcaaactggataaacatgcgcgtagcggcgaatgc
 L  A  N  Q  C  N  Y  D  C  K  L  D  K  H  A  R  S  G  E  C
ttttatgatgaaaaacgtaatctgcagtgcatttgcgattattgcgaatat
 F  Y  D  E  K  R  N  L  Q  C  I  C  D  Y  C  E  Y
```

SEQ ID NOS: 20 and 114-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
```

gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac

D Y K D E H H H H H H G G S Q D S T S D

ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa

L I P A P P L S K V P L Q Q N F Q D N Q

ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa

F Q G K W Y V V G L A G N A I L R E D K

gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc

D P Q K M Y A T I Y E L K E D K S Y N V

acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt

T S V L F R K K K C D Y W I R T F V P G

tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac

C Q P G E F T L G N I K S Y P G L T S Y

ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt

L V R V V S T N Y N Q H A M V F F K K V

tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg

S Q N R E Y F K I T L Y G R T K E L T S

gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc

E L K E N F I R F S K S L G L P E N H I

gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt

V F P V P I D Q C I D G G G S E N L Y F

cagggaggcatgtgcatgccgtgctttaccaccgatcatcagatggcgcgtaaatgcgat

Q G G M C M P C F T T D H Q M A R K C D

gattgctgcggcggcaaaggccgtggcaaatgctatggcccgcagtgcctgtgccgt

D C C G G K G R G K C Y G P Q C L C R

Construction of parental construct for BamH1/NotI cloning:

SEQ ID NO: 21 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GS-PARENTAL

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAG**GGATCC**TAATGTTGGCCATGATGTTAG**GCGGCCGC**TAAGGATCCCGGA

BamHI site: GGATCC
NotI site: GCGGCCGC

A BamHI site adds "GS" before a knottin. This construct can be used for cloning libraries.

Construction of parental construct for furin cleavage, BamHI/NotI cloning can include an idealized furin cut site is RARYKRS (SEQ ID NO: 115) –RARYKRGS (SEQ ID NO: 99) can be used for a Bam HI site.

SEQ ID NO: 22 - IgK-SF-H6-GGS-lcn2C-GGS-furin-GS-PARENTAL

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT

GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCcgcgcgcgctataaacgc
GGATCCTAATGTTGGCCATGATGTTAG**GCGGCCGC**TAAGGATCCCGGA

SEQ ID NO: 23- IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GS-MIDKINE
GACTGAGTCGCCCG*CTCGAG*ACCATGGAGACAGACACACTCCTGCTATGGGTACTGC
TGCTCTGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACC
ATGGTGGAAGCCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAG
GTCCCTCTGCAGCAGAACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTA
GGCCTGGCAGGGAATGCAATTCTCAGAGAAGACAAAGACCCGCAAAAGATGTATGC
CACCATCTATGAGCTGAAAGAAGACAAGAGCTACAATGTCACCTCCGTCCTGTTTAG
GAAAAAGAAGTGTGACTACTGGATCAGGACTTTTGTTCCAGGTTGCCAGCCCGGCGA
GTTCACGCTGGGCAACATTAAGAGTTACCCTGGATTAACGAGTTACCTCGTCCGAGT
GGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCAAGAAAGTTTCTCAAAA
CAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCTGACTTCGGAAC
TAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACCACATCG
TCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTATT
TTCAG**AGC**GATTGCAAATATAAATTTGAAAACTGGGGCGCGTGCGATGGCGGCACC

GGCACCAAAGTGCGCCAGGGCACCCTGAAAAAAGCGCGCTATAACGCGCAGTGCCAGGAAA
CCATTCGCGTGACCAAACCGTGCTAAT**GCT***GGATCC*CGGACCGCCTCTCC

SEQ ID NOS: 24 and 116-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
 D  Y  K  D  E  H  H  H  H  H  H  G  G  S  Q  D  S  T  S  D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
 L  I  P  A  P  P  L  S  K  V  P  L  Q  Q  N  F  Q  D  N  Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
 F  Q  G  K  W  Y  V  V  G  L  A  G  N  A  I  L  R  E  D  K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
 D  P  Q  K  M  Y  A  T  I  Y  E  L  K  E  D  K  S  Y  N  V
acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt
 T  S  V  L  F  R  K  K  K  C  D  Y  W  I  R  T  F  V  P  G
tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac
 C  Q  P  G  E  F  T  L  G  N  I  K  S  Y  P  G  L  T  S  Y
ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt
 L  V  R  V  V  S  T  N  Y  N  Q  H  A  M  V  F  F  K  K  V
tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg
 S  Q  N  R  E  Y  F  K  I  T  L  Y  G  R  T  K  E  L  T  S
gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc
 E  L  K  E  N  F  I  R  F  S  K  S  L  G  L  P  E  N  H  I
gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt
 V  F  P  V  P  I  D  Q  C  I  D  G  G  G  S  E  N  L  Y  F
cagagcgattgcaaatataaatttgaaaactggggcgcgtgcgatggcggcaccggcacc
 Q  S  D  C  K  Y  K  F  E  N  W  G  A  C  D  G  G  T  G  T
aaagtgcgccagggcaccctgaaaaaagcgcgctataacgcgcagtgccaggaaaccatt
 K  V  R  Q  G  T  L  K  K  A  R  Y  N  A  Q  C  Q  E  T  I
cgcgtgaccaaaccgtgc
 R  V  T  K  P  C
```

SEQ ID NO: 25 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-Violacin A

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA

AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAGGGA**GGC**AGCGCCATCAGCTGCGGCGAGACCTGCTTCAAGTTCAAGTGCTACACCCC
CAGATGCAGCTGCAGCTACCCCGTGTGCAAGTAAGCTAAGGATCCCGGACCGCC

SEQ ID NOS: 26 and 117-

atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt

M E T D T L L L W V L L L W V P G S T G

gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac

D Y K D E H H H H H H G G S Q D S T S D

ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa

L I P A P P L S K V P L Q Q N F Q D N Q

ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa

F Q G K W Y V V G L A G N A I L R E D K

gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc

D P Q K M Y A T I Y E L K E D K S Y N V

acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt

T S V L F R K K K C D Y W I R T F V P G

tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac

C Q P G E F T L G N I K S Y P G L T S Y

ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt

L V R V V S T N Y N Q H A M V F F K K V

tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg

S Q N R E Y F K I T L Y G R T K E L T S

gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc

E L K E N F I R F S K S L G L P E N H I

gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt

V F P V P I D Q C I D G G G S E N L Y F

cagggaggcagcgccatcagctgcggcgagacctgcttcaagttcaagtgctacaccccc

Q G G S A I S C G E T C F K F K C Y T P

agatgcagctgcagctaccccgtgtgcaag

R C S C S Y P V C K

SEQ ID NO: 27 -

IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-Lambda Toxin

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAGGGA**GGC**GTGTGCTGCGGCTACAAGCTGTGCCACCCCTGCTAAGCTAAGGATCCCGG
ACC

SEQ ID NOS: 28 and 118-

atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt

M E T D T L L L W V L L L W V P G S T G

gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac

D Y K D E H H H H H H G G S Q D S T S D

ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa

L I P A P P L S K V P L Q Q N F Q D N Q

ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa

F Q G K W Y V V G L A G N A I L R E D K

gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc

D P Q K M Y A T I Y E L K E D K S Y N V

acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt

T S V L F R K K K C D Y W I R T F V P G

tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac

C Q P G E F T L G N I K S Y P G L T S Y

ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt

L V R V V S T N Y N Q H A M V F F K K V

tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg

S Q N R E Y F K I T L Y G R T K E L T S

gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc

E L K E N F I R F S K S L G L P E N H I

gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt

V F P V P I D Q C I D G G G S E N L Y F

```
cagggaggcgtgtgctgcggctacaagctgtgccacccctgc
 Q  G  G  V  C  C  G  Y  K  L  C  H  P  C
```

SEQ ID NO: 29 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-Lambda Toxin NG

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTC
TGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACCATGGTGGAAG
CCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAGGTCCCTCTGCAGCAGA
ACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTAGGCCTGGCAGGGAATGCAATT
CTCAGAGAAGACAAAGACCCGCAAAAGATGTATGCCACCATCTATGAGCTGAAAGAAGACA
AGAGCTACAATGTCACCTCCGTCCTGTTTAGGAAAAAGAAGTGTGACTACTGGATCAGGACT
TTTGTTCCAGGTTGCCAGCCCGGCGAGTTCACGCTGGGCAACATTAAGAGTTACCCTGGATT
AACGAGTTACCTCGTCCGAGTGGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCA
AGAAAGTTTCTCAAAACAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCT
GACTTCGGAACTAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACC
ACATCGTCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTAT
TTTCAGGGA**GGC**AACGGCGTGTGCTGCGGCTACAAGCTGTGCCACCCCTGCTAAGCTAAGGA
TCCCGGACC

SEQ ID NOS: 30 and 119-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
 D  Y  K  D  E  H  H  H  H  H  H  G  G  S  Q  D  S  T  S  D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
 L  I  P  A  P  P  L  S  K  V  P  L  Q  Q  N  F  Q  D  N  Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
 F  Q  G  K  W  Y  V  V  G  L  A  G  N  A  I  L  R  E  D  K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
 D  P  Q  K  M  Y  A  T  I  Y  E  L  K  E  D  K  S  Y  N  V
acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt
 T  S  V  L  F  R  K  K  K  C  D  Y  W  I  R  T  F  V  P  G
tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac
 C  Q  P  G  E  F  T  L  G  N  I  K  S  Y  P  G  L  T  S  Y
ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt
 L  V  R  V  V  S  T  N  Y  N  Q  H  A  M  V  F  F  K  K  V
tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg
 S  Q  N  R  E  Y  F  K  I  T  L  Y  G  R  T  K  E  L  T  S
gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc
 E  L  K  E  N  F  I  R  F  S  K  S  L  G  L  P  E  N  H  I
```

gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt

V F P V P I D Q C I D G G G S E N L Y F

cagggaggcaacggcgtgtgctgcggctacaagctgtgccacccctgc

Q G G N G V C C G Y K L C H P C

SEQ ID NO: 31 - IgK-SF-H6-GGS-lcn2C-GGS-ENLYFQ-GG-Potato Carboxypeptidase Inhibitor

GACTGAGTCGCCCGCTCGAGACCATGGAGACAGACACACTCCTGCTATGGGTACTGC

TGCTCTGGGTTCCAGGTTCCACTGGTGACTACAAGGACGAGCATCACCATCATCACC

ATGGTGGAAGCCAGGACTCCACCTCAGACCTGATCCCAGCCCCACCTCTGAGCAAG

GTCCCTCTGCAGCAGAACTTCCAGGACAACCAATTCCAGGGGAAGTGGTATGTGGTA

GGCCTGGCAGGGAATGCAATTCTCAGAGAAGACAAAGACCCGCAAAAGATGTATGC

CACCATCTATGAGCTGAAAGAAGACAAGAGCTACAATGTCACCTCCGTCCTGTTTAG

GAAAAAGAAGTGTGACTACTGGATCAGGACTTTTGTTCCAGGTTGCCAGCCCGGCGA

GTTCACGCTGGGCAACATTAAGAGTTACCCTGGATTAACGAGTTACCTCGTCCGAGT

GGTGAGCACCAACTACAACCAGCATGCTATGGTGTTCTTCAAGAAAGTTTCTCAAAA

CAGGGAGTACTTCAAGATCACCCTCTACGGGAGAACCAAGGAGCTGACTTCGGAAC

TAAAGGAGAACTTCATCCGCTTCTCCAAATCTCTGGGCCTCCCTGAAAACCACATCG

TCTTCCCTGTCCCAATCGACCAGTGTATCGACGGCGGAGGTAGCGAAAACCTGTATT

TTCAGGGA**GGC**cagcagcatgcggatccgatttgcaacaaaccgtgcaaaacccatgatgattgcagcggcgcgtggttttgccaggcgtgc
tggaacagcgcgcgcacctgcggcccgtatgtgggcTAATGCTAAGGATCCCGGACCG

SEQ ID NOS: 32 and 120-

```
atggagacagacacactcctgctatgggtactgctgctctgggttccaggttccactggt
 M  E  T  D  T  L  L  L  W  V  L  L  L  W  V  P  G  S  T  G
gactacaaggacgagcatcaccatcatcaccatggtggaagccaggactccacctcagac
 D  Y  K  D  E  H  H  H  H  H  H  G  G  S  Q  D  S  T  S  D
ctgatcccagccccacctctgagcaaggtccctctgcagcagaacttccaggacaaccaa
 L  I  P  A  P  P  L  S  K  V  P  L  Q  Q  N  F  Q  D  N  Q
ttccaggggaagtggtatgtggtaggcctggcagggaatgcaattctcagagaagacaaa
 F  Q  G  K  W  Y  V  V  G  L  A  G  N  A  I  L  R  E  D  K
gacccgcaaaagatgtatgccaccatctatgagctgaaagaagacaagagctacaatgtc
 D  P  Q  K  M  Y  A  T  I  Y  E  L  K  E  D  K  S  Y  N  V
acctccgtcctgtttaggaaaaagaagtgtgactactggatcaggacttttgttccaggt
 T  S  V  L  F  R  K  K  K  C  D  Y  W  I  R  T  F  V  P  G
tgccagcccggcgagttcacgctgggcaacattaagagttaccctggattaacgagttac
 C  Q  P  G  E  F  T  L  G  N  I  K  S  Y  P  G  L  T  S  Y
ctcgtccgagtggtgagcaccaactacaaccagcatgctatggtgttcttcaagaaagtt
 L  V  R  V  V  S  T  N  Y  N  Q  H  A  M  V  F  F  K  K  V
tctcaaaacagggagtacttcaagatcaccctctacgggagaaccaaggagctgacttcg
 S  Q  N  R  E  Y  F  K  I  T  L  Y  G  R  T  K  E  L  T  S
gaactaaaggagaacttcatccgcttctccaaatctctgggcctccctgaaaaccacatc
 E  L  K  E  N  F  I  R  F  S  K  S  L  G  L  P  E  N  H  I
gtcttccctgtcccaatcgaccagtgtatcgacggcggaggtagcgaaaacctgtatttt
 V  F  P  V  P  I  D  Q  C  I  D  G  G  G  S  E  N  L  Y  F
cagggaggccagcagcatgcggatccgatttgcaacaaaccgtgcaaaacccatgatgat
 Q  G  G  Q  Q  H  A  D  P  I  C  N  K  P  C  K  T  H  D  D
tgcagcggcgcgtggttttgccaggcgtgctggaacagcgcgcgcacctgcggcccgtat
 C  S  G  A  W  F  C  Q  A  C  W  N  S  A  R  T  C  G  P  Y
gtgggctaa
 V  G
```

Anti-CD3-STa bispecific

The anti-CD3 is an OKT3 variant from the C-terminus of patent 7635462. It is supposed to work as either N- or C- terminal.

SEQ ID NO: 33-

GSDIKLQQSGAELARPGASVKMSCKTSGYTFTRYTMHWVKQRPGQGLEWIGYINPSRGYTNYN
QKFKDKATLTTDKSSSTAYMQLSSLTSEDSAVYYCARYYDDHYCLDYWGQGTTLTVSSVEGGS
GGSGGSGGSGGVDDIQLTQSPAIMSASPGEKVTMTCRASSSVSYMNWYQQKSGTSPKRWIYDTS
KVASGVPYRFSGSGSGTSYSLTISSMEAEDAATYYCQQWSSNPLTFGAGTKLELKGGGGSNSSN
YCCELCCNPACTGCY

This is designed to be dropped into the Bam/Not cut library vector as a TEV-cleavable siderocalin fusion:

SEQ ID NO: 34 -
ACCTGTATTTTCAGGGATCCgatattaaactgcagcagagcggcgcggaactggcgcgcccgggcgcgagcgtgaaaatgagctg
caaaaccagcggctatacctttacccgctataccatgcattgggtgaaacagcgcccgggccagggcctggaatggattggctatattaacccgagccg
cggctataccaactataaccagaaatttaaagataaagcgaccctgaccaccgataaaagcagcagcaccgcgtatatgcagctgagcagcctgaccag
cgaagatagcgcggtgtattattgcgcgcgctattatgatgatcattattgcctggattattggggccagggcaccaccctgaccgtgagcagcgtggaag
gcggcagcggcggcagcggcggcagcggcggcagcggcggcgtggatgatattcagctgacccagagcccggcgattatgagcgcgagcccgggg
cgaaaaagtgaccatgacctgccgcgcgagcagcagcgtgagctatatgaactggtatcagcagaaaagcggcaccagcccgaaacgctggatttatg
ataccagcaaagtggcgagcggcgtgccgtatcgctttagcggcagcggcagcggcaccagctatagcctgaccattagcagcatggaagcggaaga
tgcggcgacctattattgccagcagtggagcagcaacccgctgacctttggcgcgggcaccaaactggaactgaaaggcggcggcggcagcaacag
cagcaactattgctgcgaactgtgctgcaacccggcgtgcaccggctgctatTAATGCGGCCGCTCATCACCATTAATC