US 20200140511A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2020/0140511 A1**

Porat et al. (43) **Pub. Date: May 7, 2020**

(54) **DELIVERY CONSTRUCTS FOR TRANSCYTOSIS AND RELATED METHODS**

(71) Applicant: **Applied Molecular Transport Inc.**, South San Francisco, CA (US)

(72) Inventors: **Amir Porat**, Belmont, CA (US); **Randall J. Mrsny**, Los Altos Hills, CA (US); **Elbert Seto**, San Francisco, CA (US); **Tahir Mahmood**, Burlingame, CA (US); **Charles Olson**, South San Francisco, CA (US); **Weijun Feng**, Danville, CA (US)

(21) Appl. No.: **16/686,671**

(22) Filed: **Nov. 18, 2019**

Related U.S. Application Data

(63) Continuation of application No. PCT/US2019/060356, filed on Nov. 7, 2019, Continuation-in-part of application No. PCT/US2019/050708, filed on Sep. 11, 2019, Continuation-in-part of application No. PCT/US2019/021474, filed on Mar. 8, 2019.

(60) Provisional application No. 62/756,889, filed on Nov. 7, 2018, provisional application No. 62/888,133, filed on Aug. 16, 2019, provisional application No. 62/888,238, filed on Aug. 16, 2019.

Publication Classification

(51) **Int. Cl.**

C07K 14/54	(2006.01)
C07K 14/28	(2006.01)
A61K 47/64	(2006.01)
A61K 47/65	(2006.01)
C07K 1/18	(2006.01)
A61P 1/00	(2006.01)

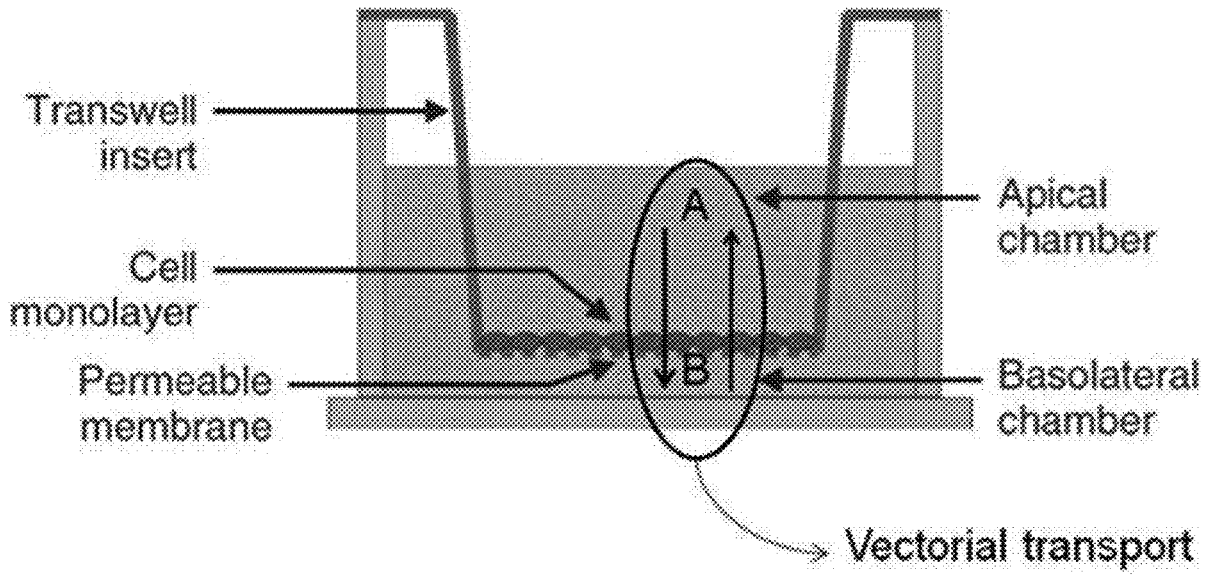
(52) **U.S. Cl.**

CPC ***C07K 14/54*** (2013.01); ***C07K 14/28*** (2013.01); ***A61P 1/00*** (2018.01); ***A61K 47/65*** (2017.08); ***C07K 1/18*** (2013.01); ***A61K 47/6415*** (2017.08)

(57) **ABSTRACT**

The present disclosure provides non-naturally occurring fusion molecules comprising therapeutic cargo moieties, such as IL-22 with a carrier. The disclosure also provides methods and compositions for the production, purification, refolding, formulation, and administration of fusion molecules. Methods and for using the purified molecules to treat and prevent diseases or disorders are also provided herein.

Specification includes a Sequence Listing.



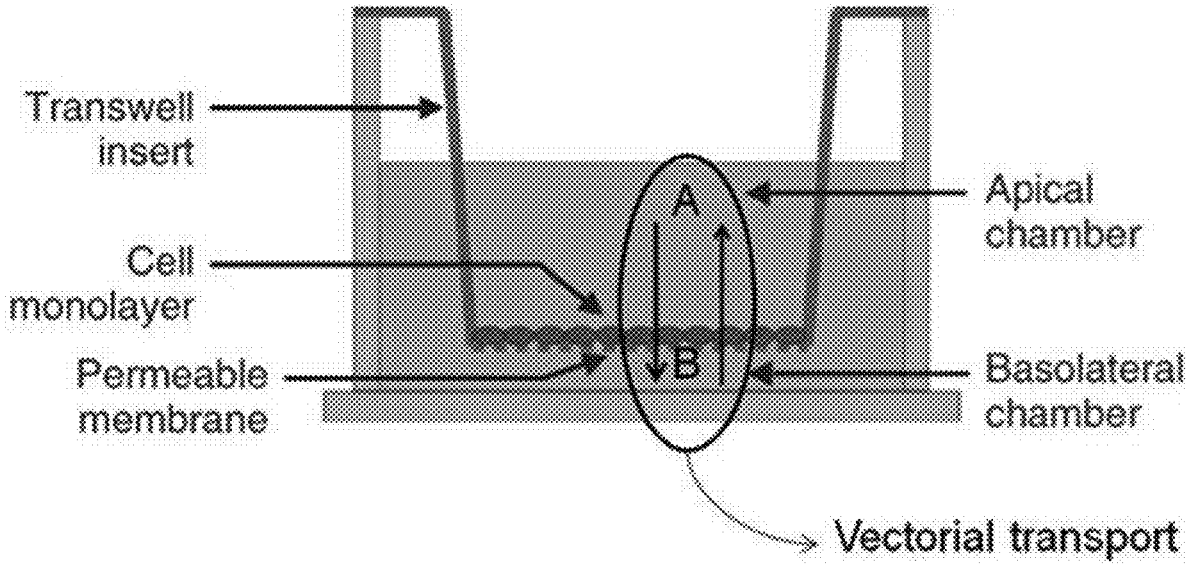

Transwell insert
Cell monolayer
Permeable membrane
A
B
Apical chamber
Basolateral chamber
Vectorial transport

FIG. 1

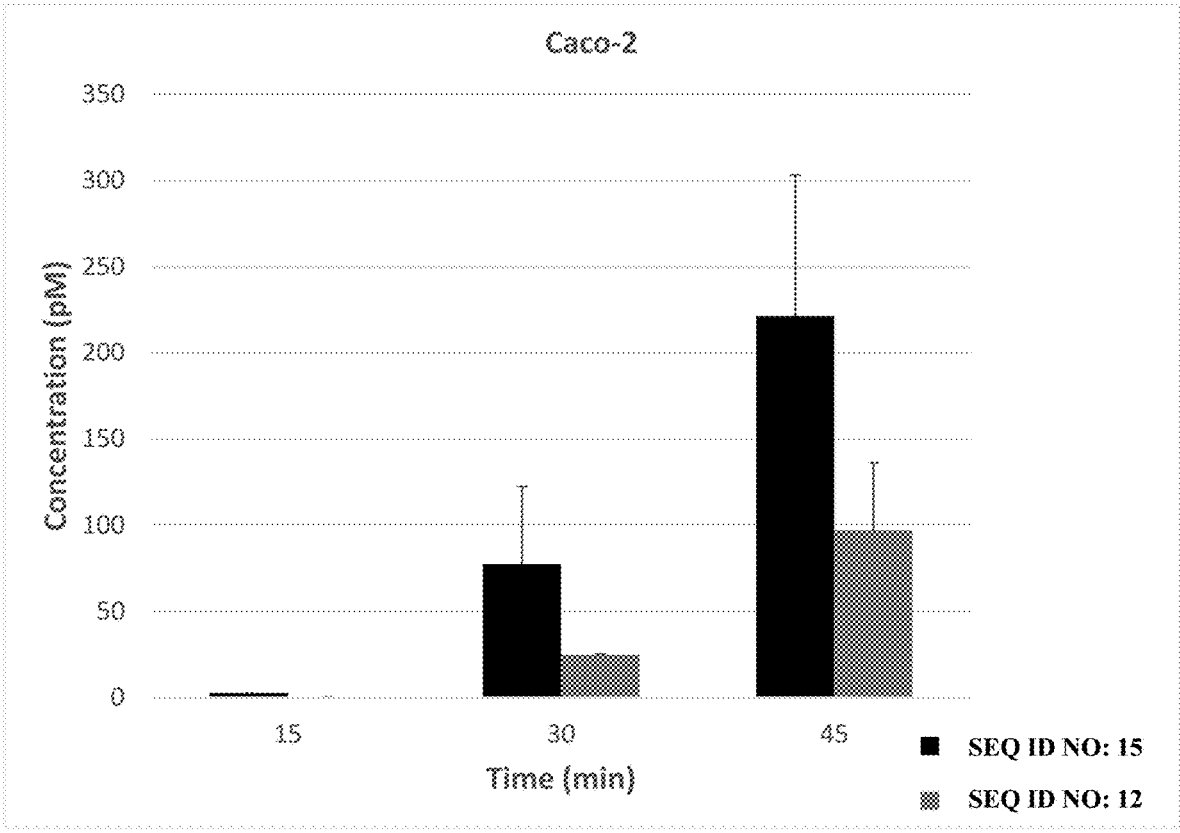

Caco-2
350
300
250
200
150
100
50
0
Concentration (pM)
15
30
45
Time (min)
SEQ ID NO: 15
SEQ ID NO: 12

FIG. 2

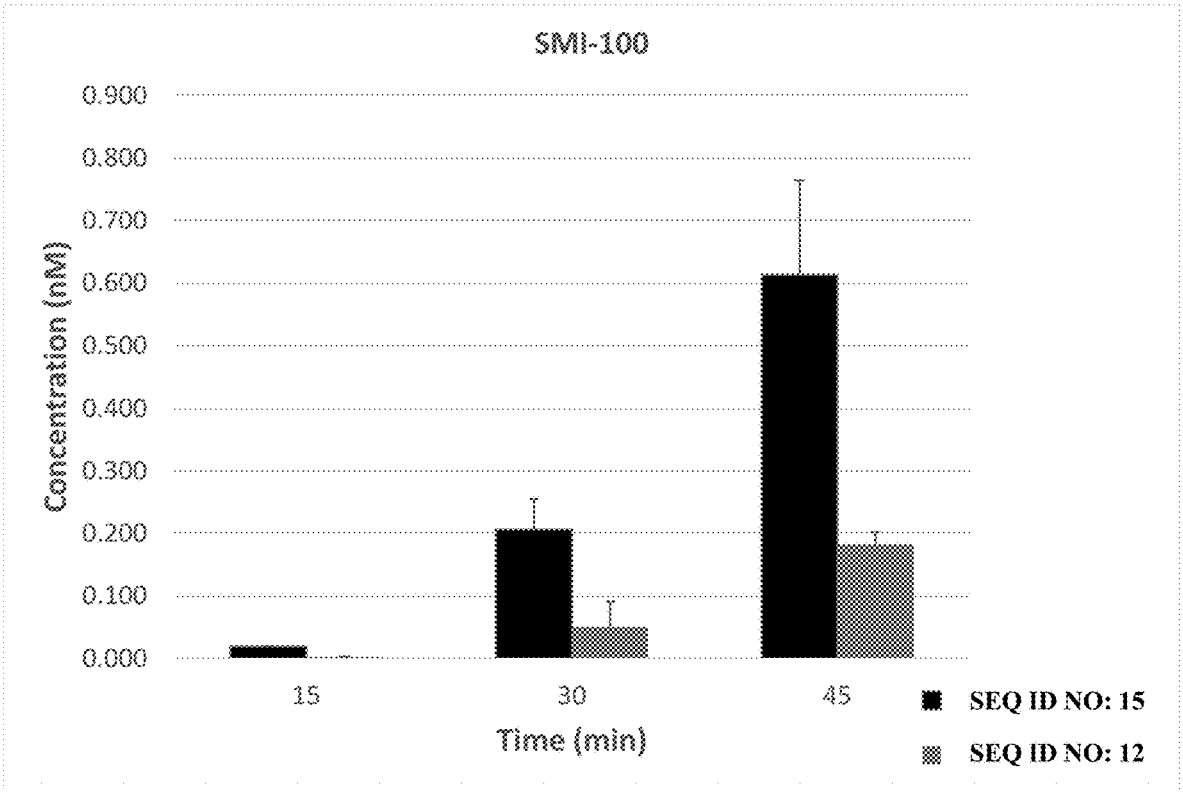

SMI-100
0.900
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
Concentration (nM)
15
30
45
Time (min)
SEQ ID NO: 15
SEQ ID NO: 12

FIG. 3

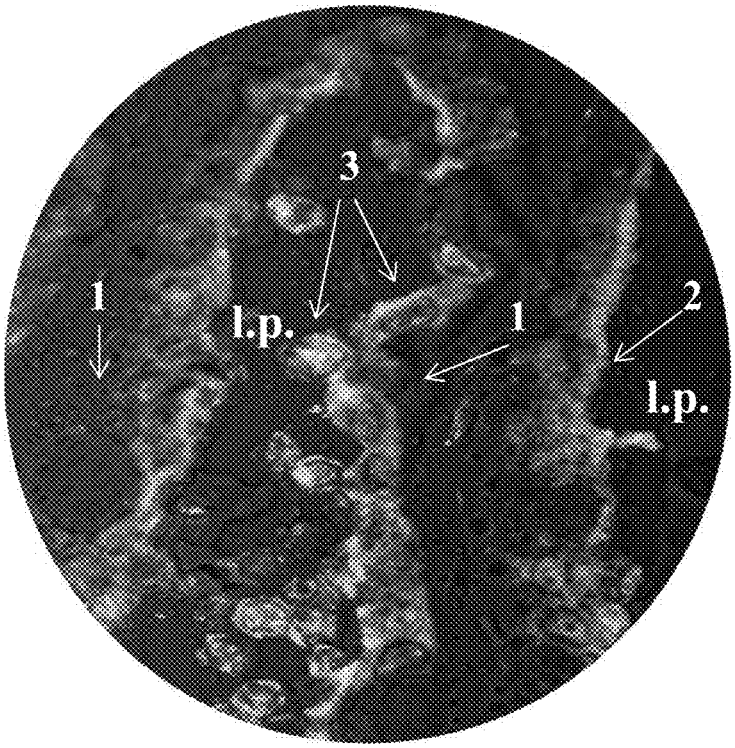


FIG. 4

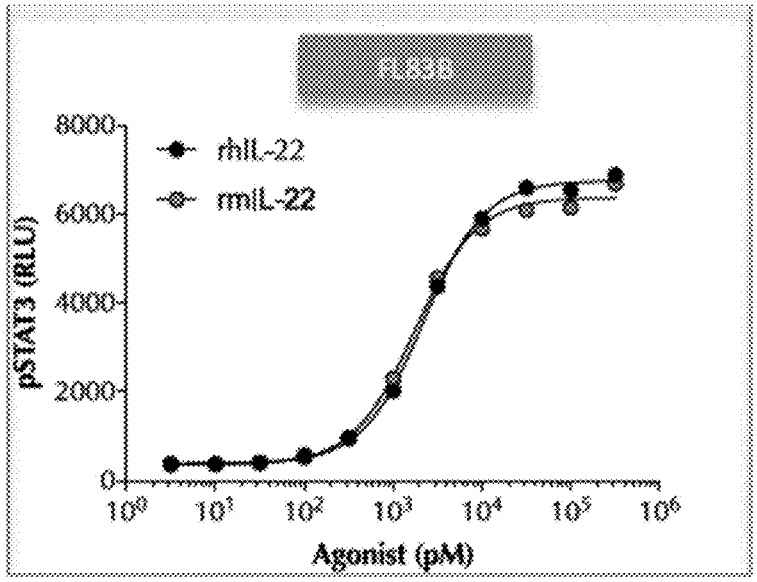

8000
6000
4000
2000
0
pSTAT3 (RLU)
rhIL-22
rmIL-22
10^0 10^1 10^2 10^3 10^4 10^5 10^6
Agonist (pM)

FIG. 5A

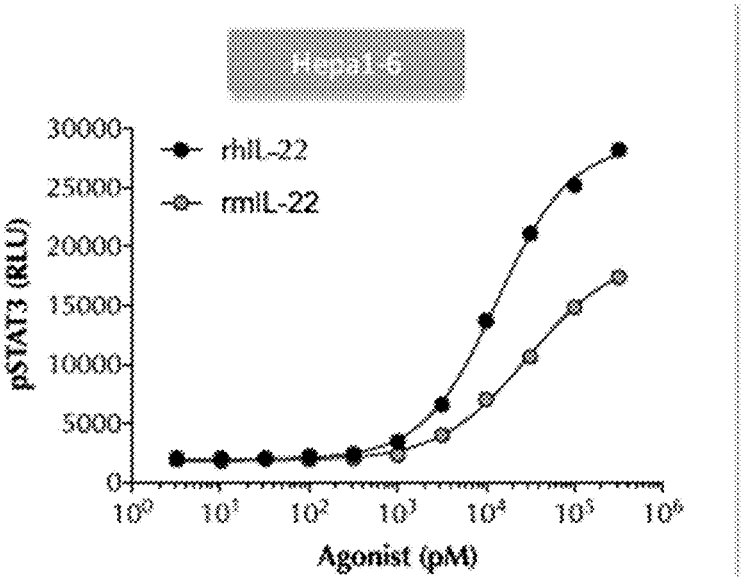

30000
25000
20000
15000
10000
5000
0
pSTAT3 (RLU)
rhIL-22
rmIL-22
10^0 10^1 10^2 10^3 10^4 10^5 10^6
Agonist (pM)

FIG. 5B

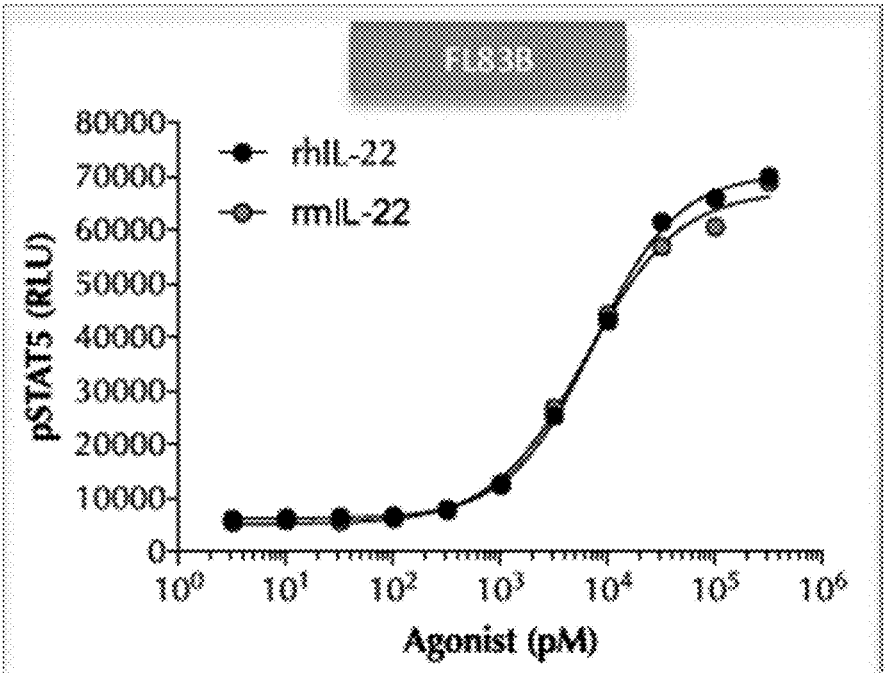

80000
70000
60000
50000
40000
30000
20000
10000
0
pSTAT5 (RLU)
rhIL-22
rmIL-22
10^0 10^1 10^2 10^3 10^4 10^5 10^6
Agonist (pM)

FIG. 6A

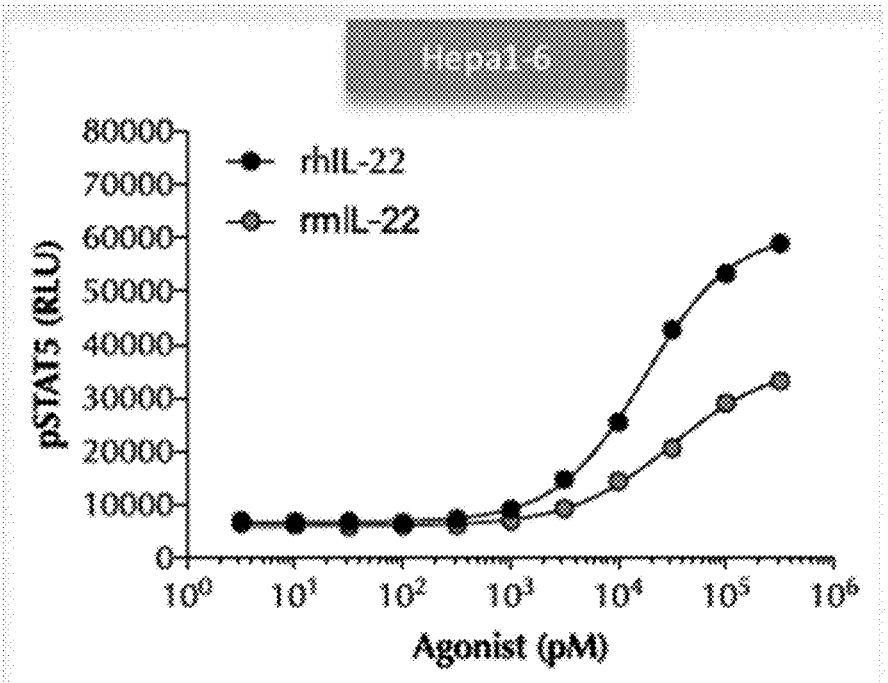

80000
70000
60000
50000
40000
30000
20000
10000
0
pSTAT5 (RLU)
rhIL-22
rmIL-22
10^0 10^1 10^2 10^3 10^4 10^5 10^6
Agonist (pM)

FIG. 6B

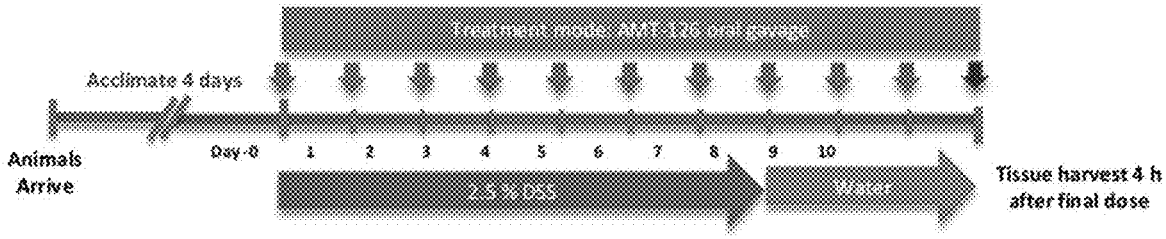


FIG. 7A

Group	Route, frequency	Dose
Naive	p.o., q.d.	NA
Vehicle	p.o., q.d.	10 mg/mL SBTI in bicarb
Cyclosporine A (model positive control)	p.o., q.d.	75 mg/kg
rhIL-22	i.p., q.d.	4 mg/kg
SEQ ID NO: 15	p.o., q.d.	1 mg/kg
SEQ ID NO: 15	p.o., q.d.	30 mg/kg

FIG. 7B

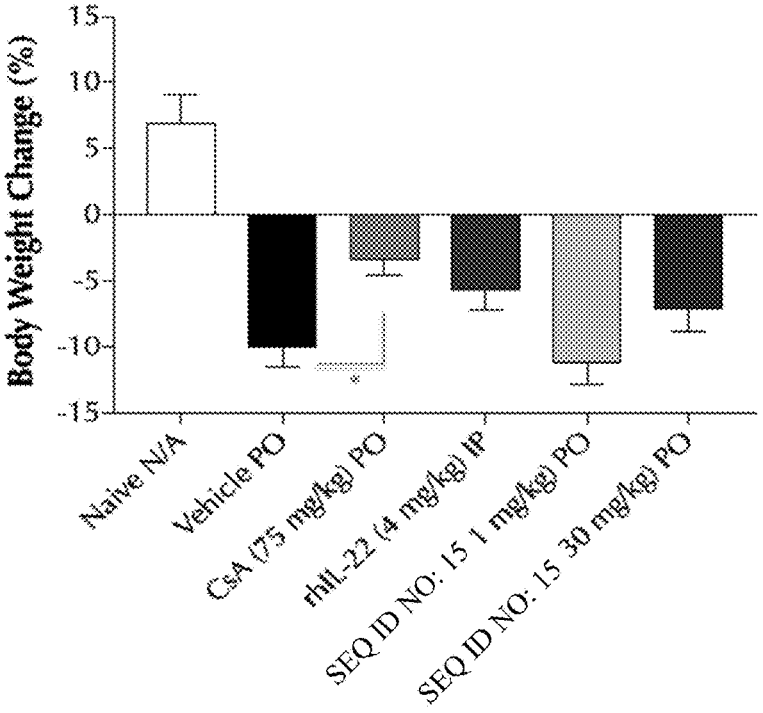

Body Weight Change (%)
15
10
5
0
-5
-10
-15
Naive N/A
Vehicle PO
CsA (75 mg/kg) PO
rhIL-22 (4 mg/kg) IP
SEQ ID NO: 15 1 mg/kg) PO
SEQ ID NO: 15 30 mg/kg) PO

FIG. 8A

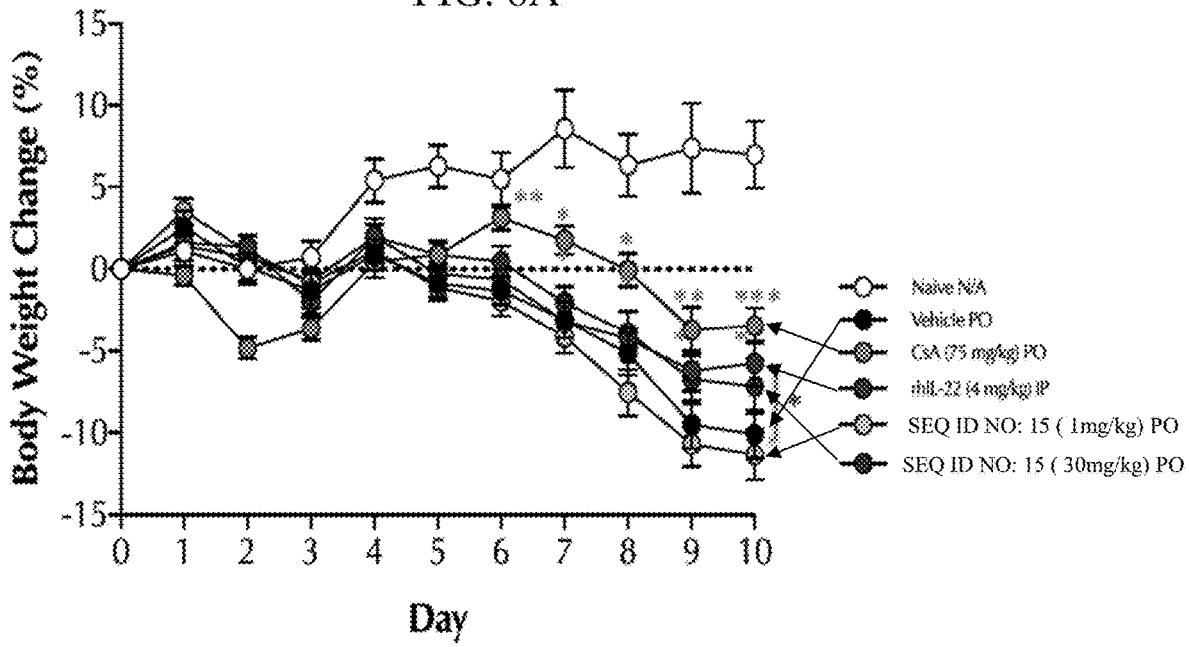

Body Weight Change (%)
15
10
5
0
-5
-10
-15
0 1 2 3 4 5 6 7 8 9 10
Day
Naive N/A
Vehicle PO
CsA (75 mg/kg) PO
rhIL-22 (4 mg/kg) IP
SEQ ID NO: 15 (1mg/kg) PO
SEQ ID NO: 15 (30mg/kg) PO

FIG. 8B

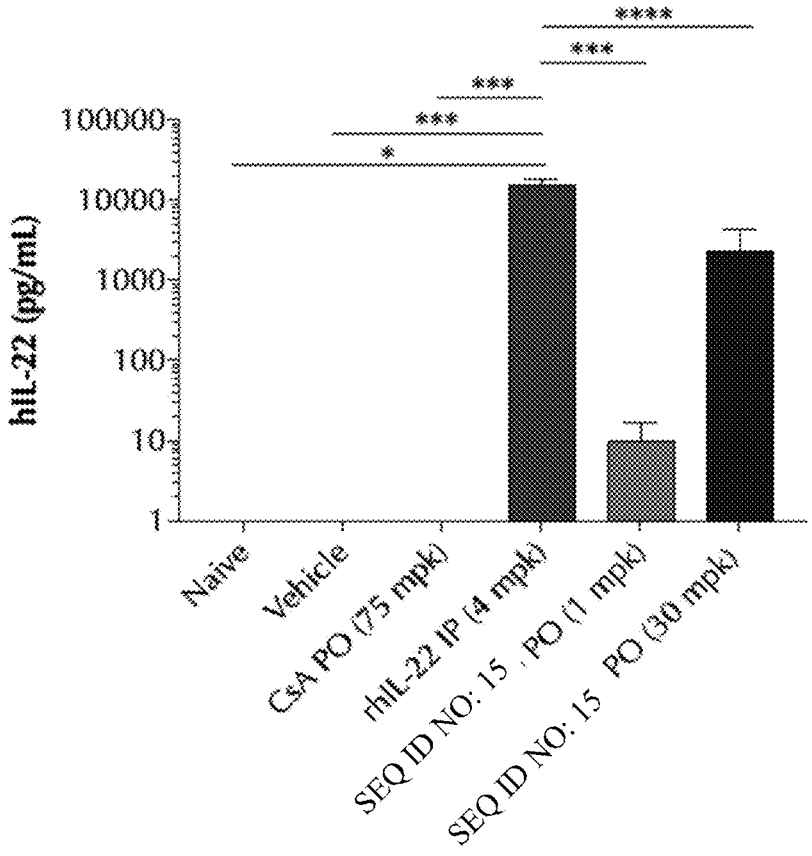
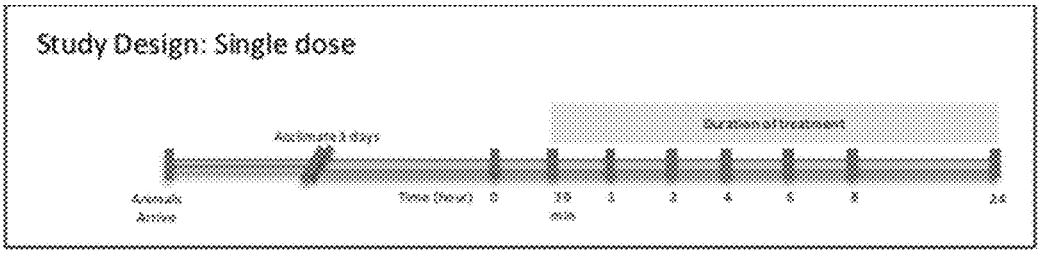
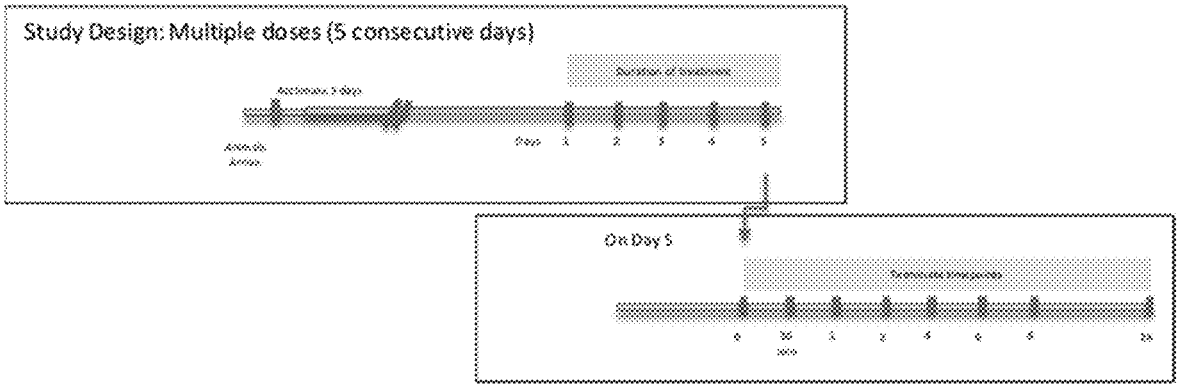

hIL-22 (pg/mL)
100000
10000
1000
100
10
1
*
***
***
***
****
Naive
Vehicle
CsA PO (75 mpk)
rhIL-22 IP (4 mpk)
SEQ ID NO: 15 PO (1 mpk)
SEQ ID NO: 15 PO (30 mpk)

FIG. 9



Study Details		Study Endpoints for Plasma
Animals	Mice, male CD-1, non-fasted	• Total IL-22 in plasma and colon tissue • PD biomarkers in plasma and colon tissue
Treatment groups	• Time Points: 0, 30min, 1, 2, 4, 6, 8 and 24h • rhIL-22 (s.c.): 2 mg/kg, q.d.	
N per group	4 Total animal number: 32	• Plasma • Colon Tissue=Proximal and distal segments

FIG. 10A



Study Details		Study Endpoints for Plasma
Animals	Mice, male CD-1, non-fasted	• Total IL-22 in plasma and colon tissue • PD biomarkers in plasma and colon tissue
Treatment groups	• rhIL-22 (s.c.): 2 mg/kg ; q.d., 5 days dose • Day 5 time Points: 0, 30min, 1, 2, 4, 6, 8 and 24h	
N per group	4 Total animal number: 32	• Plasma • Colon Tissue=Proximal and distal segments

FIG. 10B

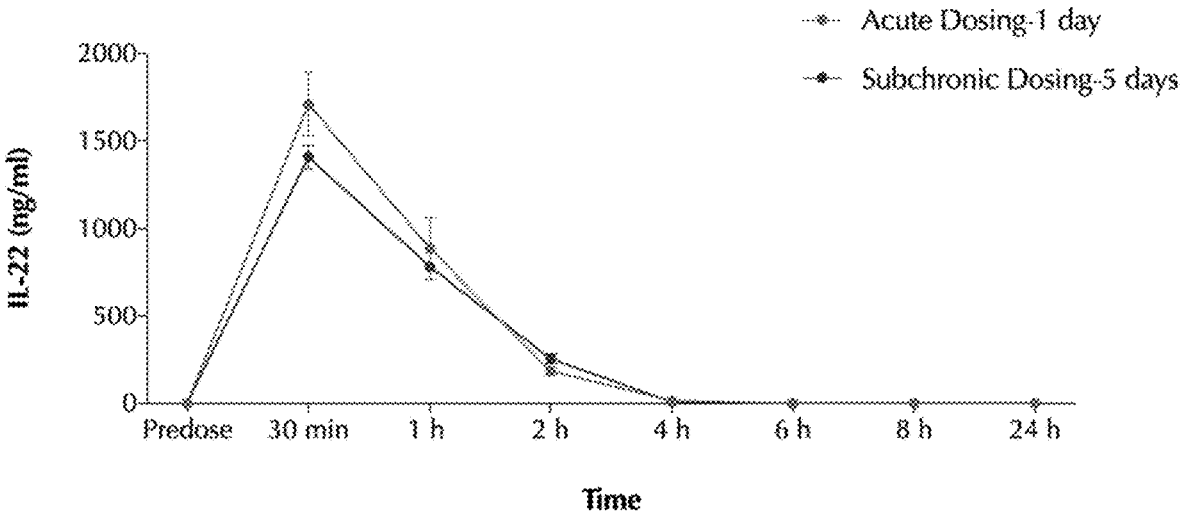


FIG. 11A

Parameter	Acute Dosing	Subchronic Dosing
Cmax	1711 ng/ml	1409 ng/ml
Tmax	30 min	30 min
AUC_{0-24}	2796 (ng/ml)xh	2454 (ng/ml)xh

FIG. 11B

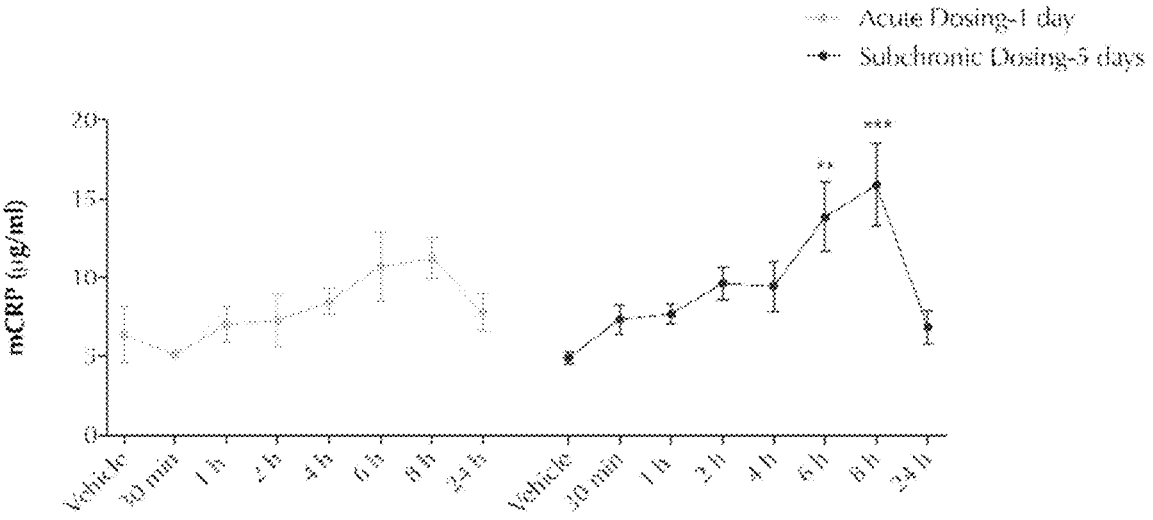


FIG. 12A

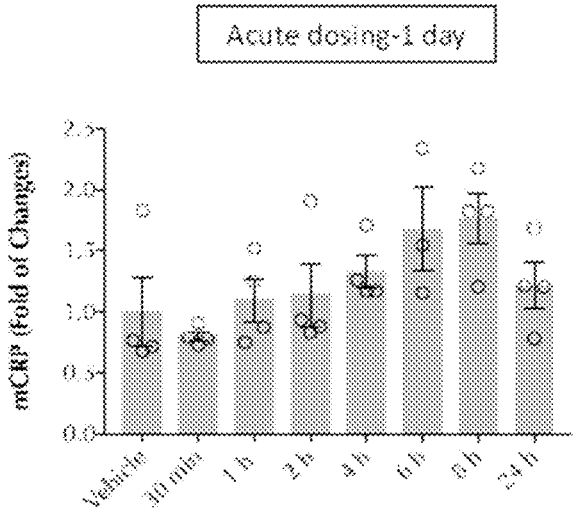


FIG. 12B

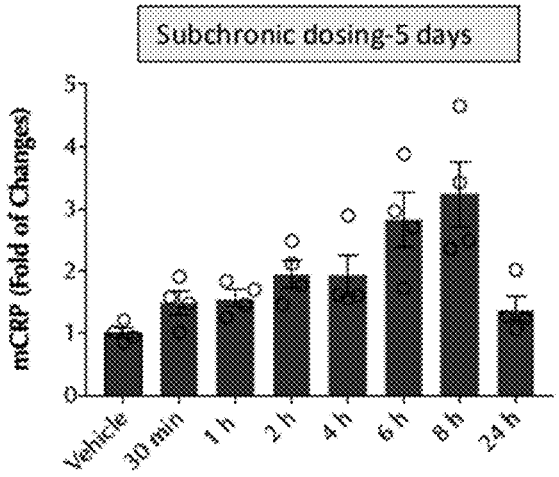


FIG. 12C

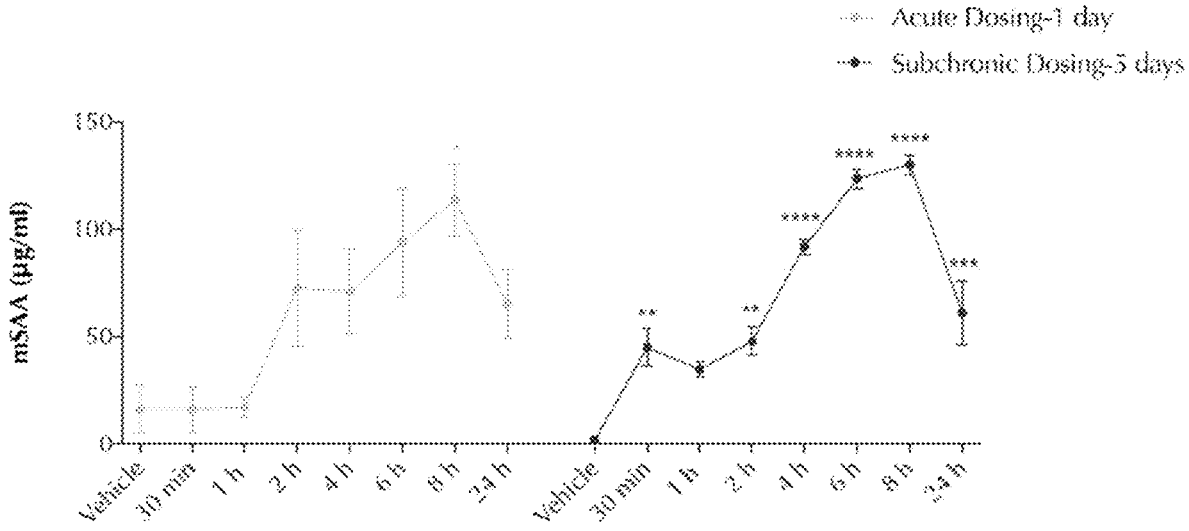


FIG. 13A

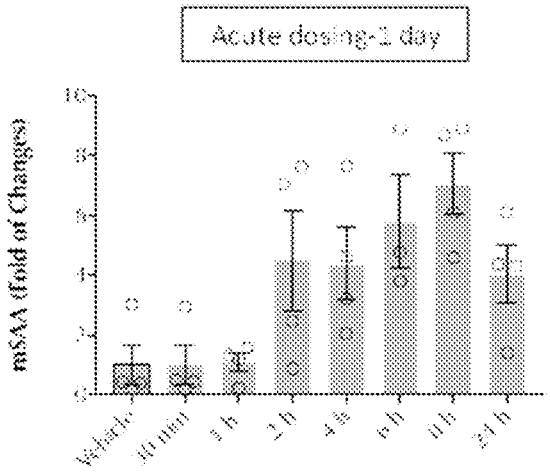


FIG. 13B

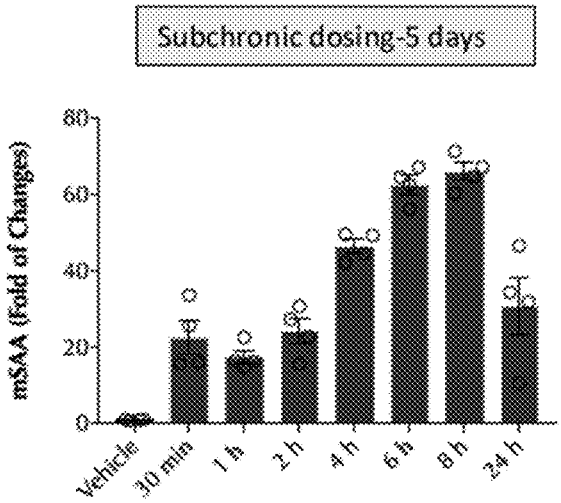


FIG. 13C

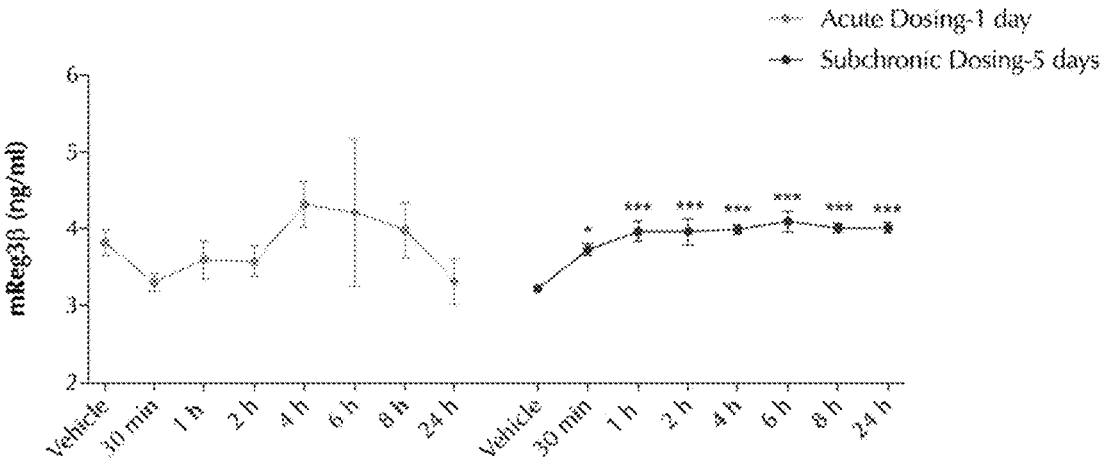


FIG. 14A

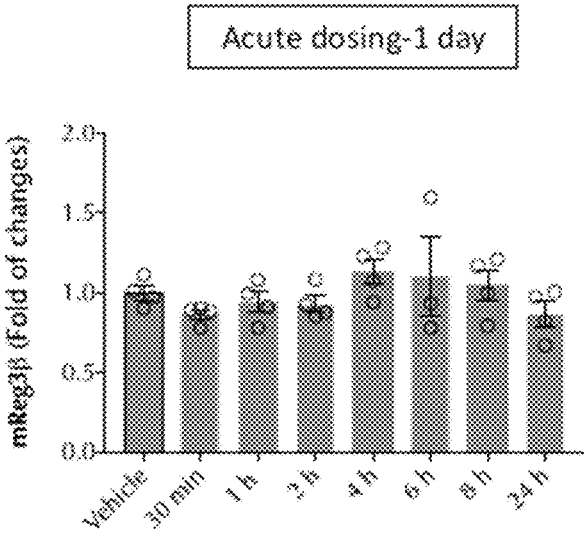


FIG. 14B

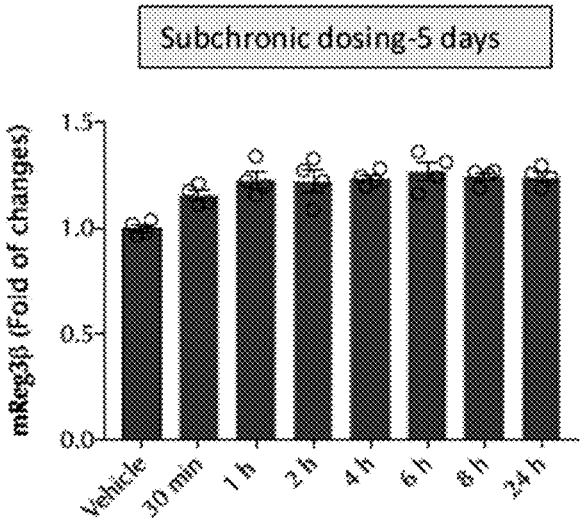


FIG. 14C

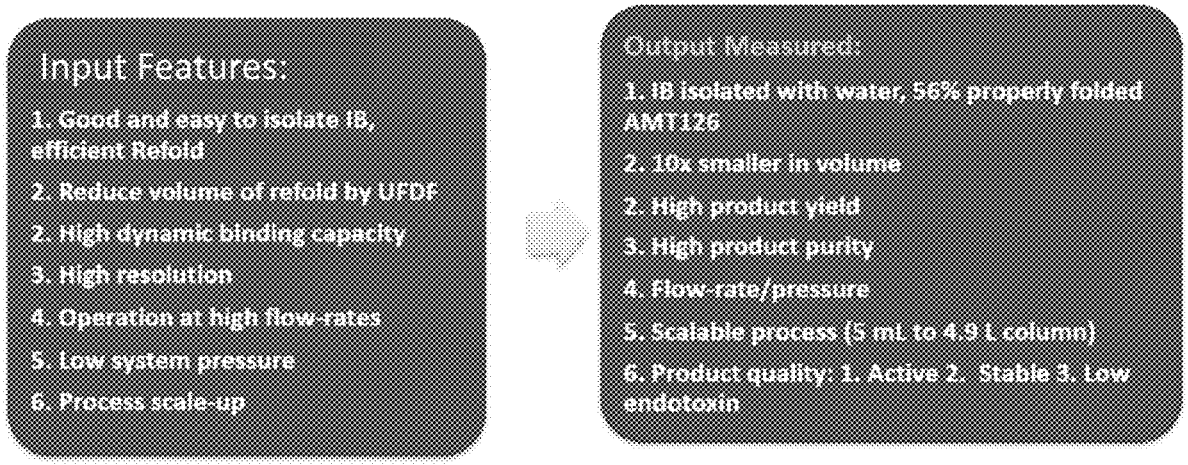
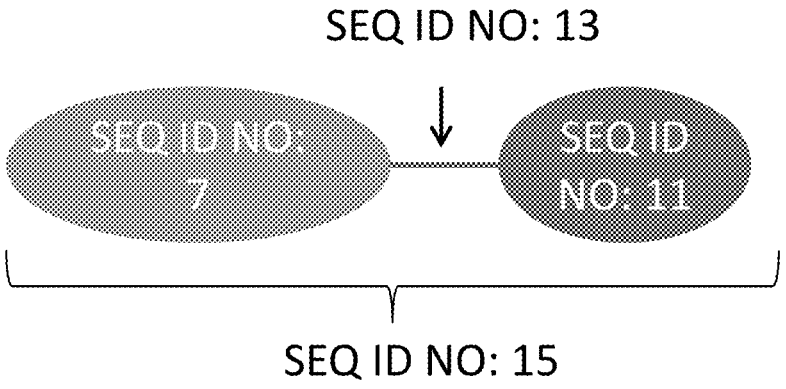

Input Features:
1. Good and easy to isolate IB, efficient Refold
2. Reduce volume of refold by UFDF
2. High dynamic binding capacity
3. High resolution
4. Operation at high flow-rates
5. Low system pressure
6. Process scale-up
Output Measured:
1. IB isolated with water, 56% properly folded AMT126
2. 10x smaller in volume
2. High product yield
3. High product purity
4. Flow-rate/pressure
5. Scalable process (5 mL to 4.9 L column)
6. Product quality: 1. Active 2. Stable 3. Low endotoxin

FIG. 15



MW: 48 kDa
Theoretical pI: 5.5
Disulfide bonds: 4

FIG. 16

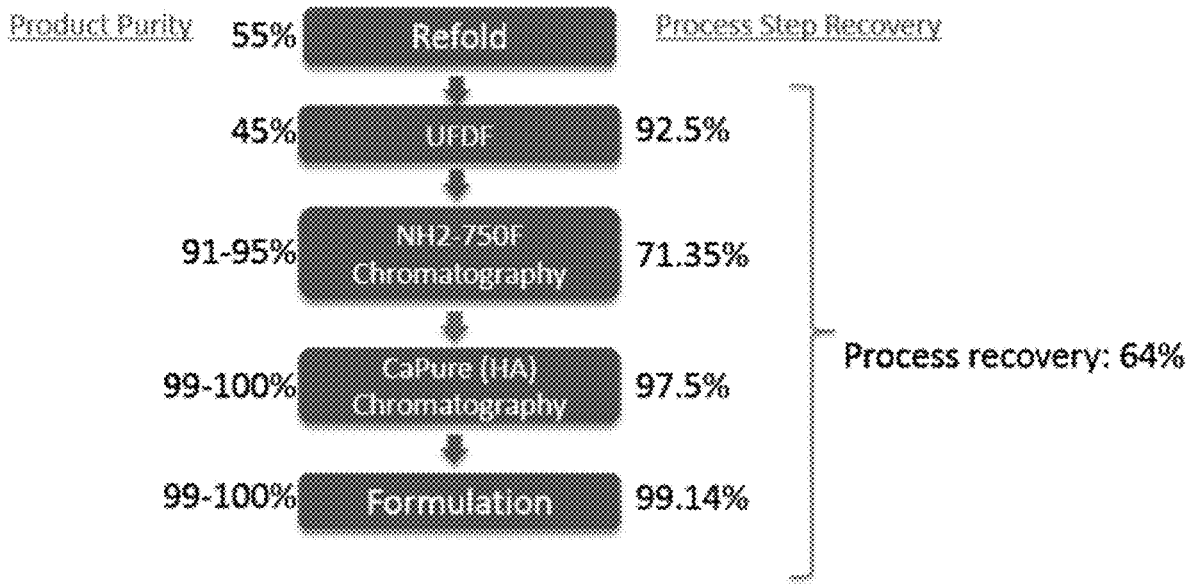

Product Purity
Process Step Recovery
55%
Refold
45%
UFDF
92.5%
91-95%
NH2-750F Chromatography
71.35%
99-100%
CaPure (HA) Chromatography
97.5%
99-100%
Formulation
99.14%
Process recovery: 64%

FIG. 17

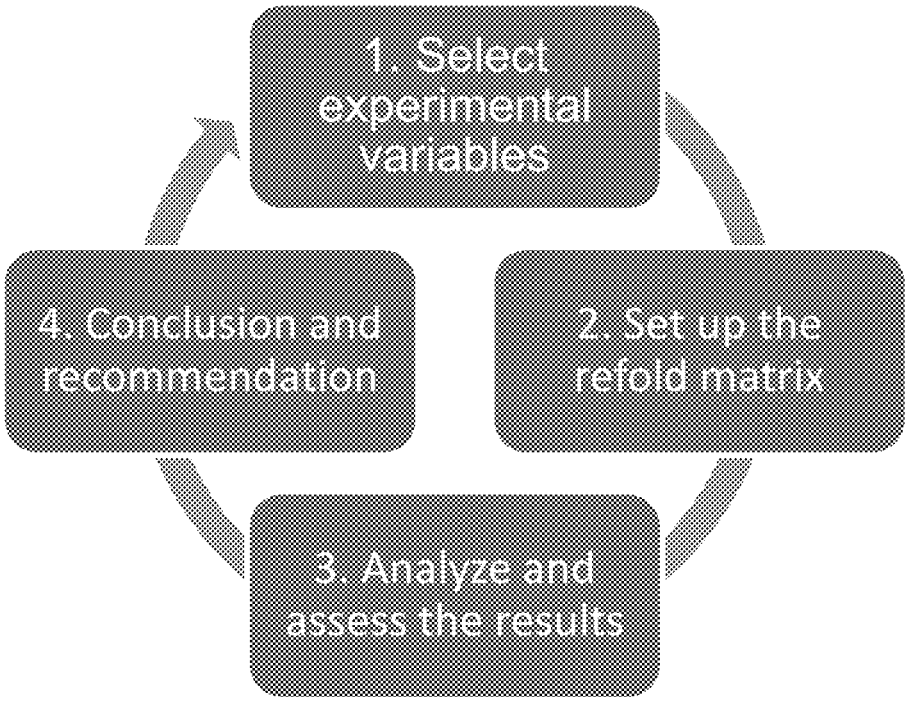

1. Select experimental variables
2. Set up the refold matrix
3. Analyze and assess the results
4. Conclusion and recommendation

FIG. 18

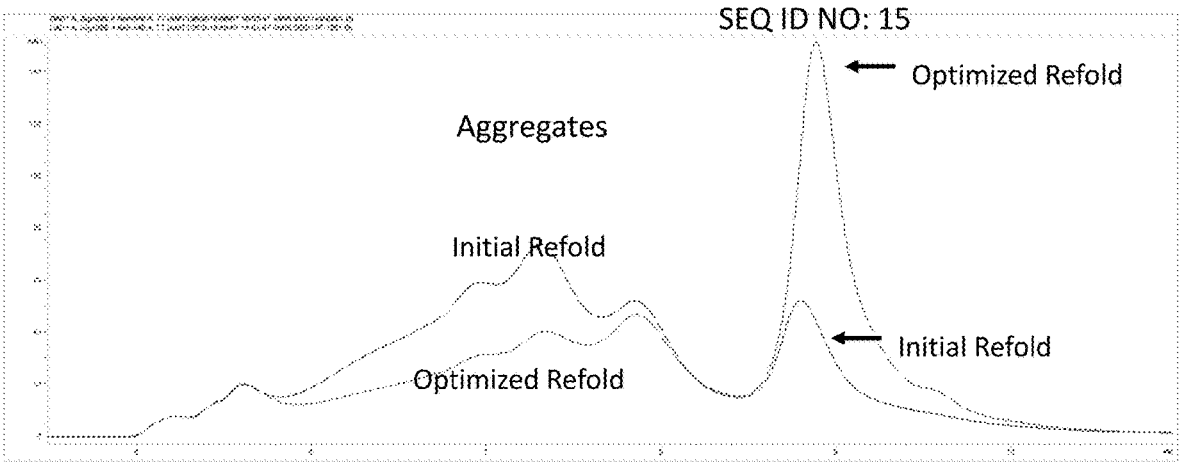

SEQ ID NO: 15
Optimized Refold
Aggregates
Initial Refold
Initial Refold
Optimized Refold

FIG. 19

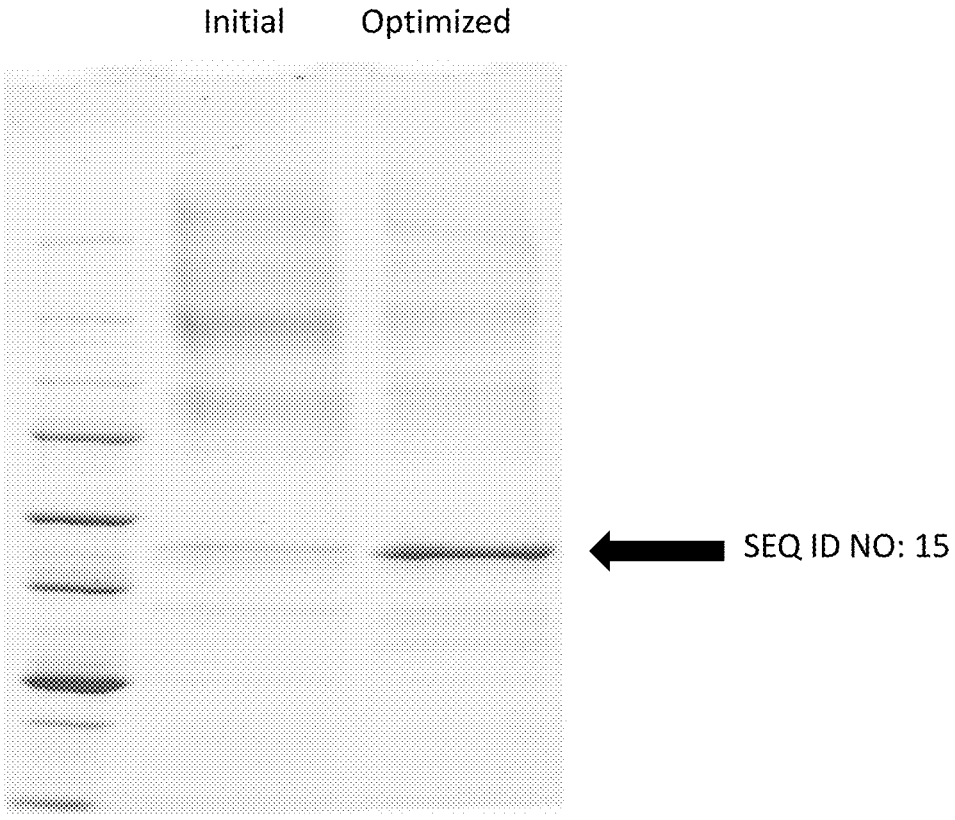


FIG. 20

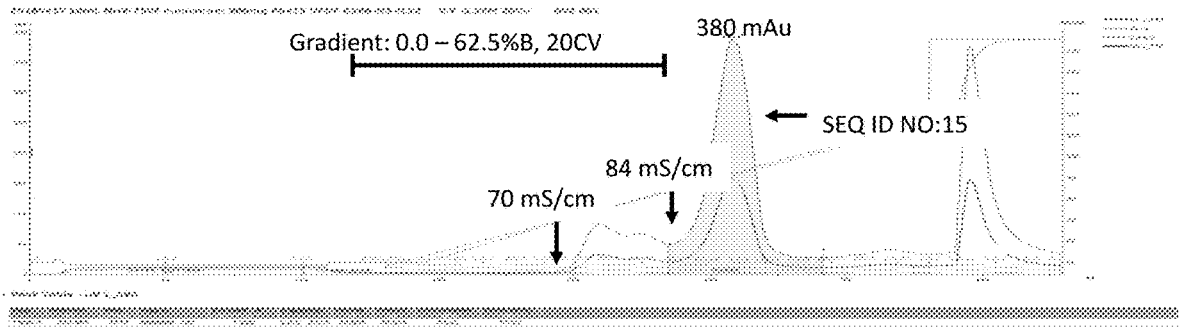


FIG. 21A

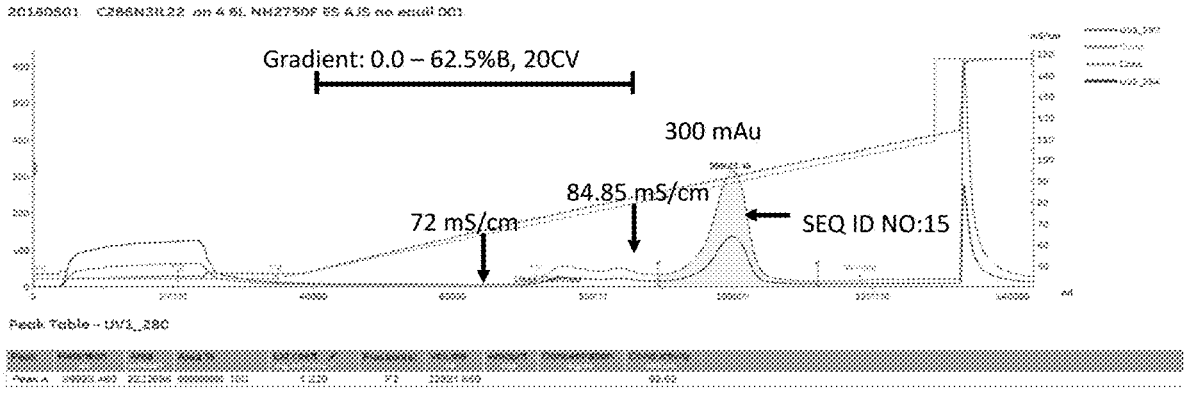

Gradient: 0.0 – 62.5%B, 20CV
300 mAu
84.85 mS/cm
72 mS/cm
SEQ ID NO:15
Peak Table - UV1_280

FIG. 21B

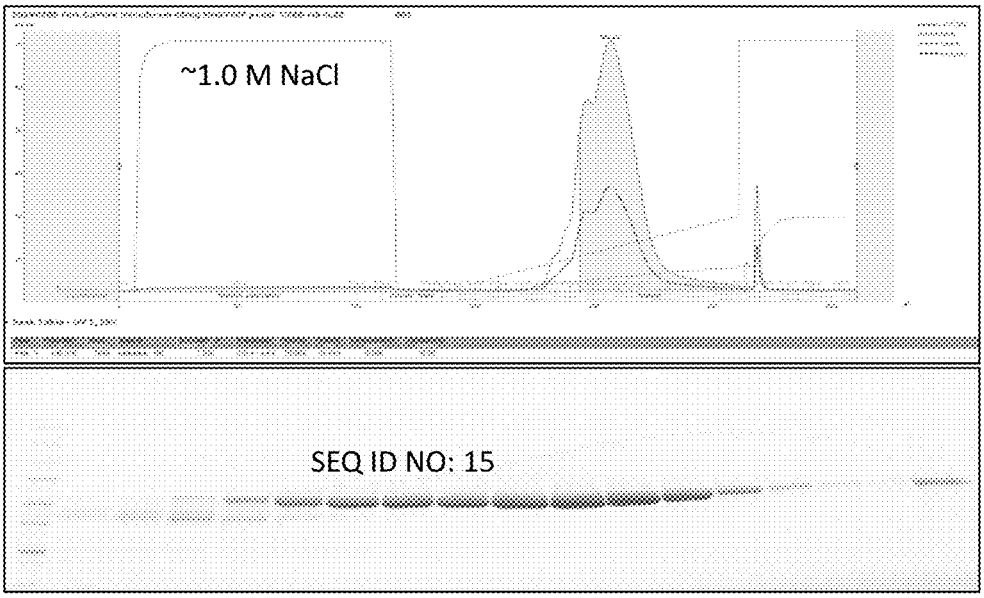


FIG. 22A

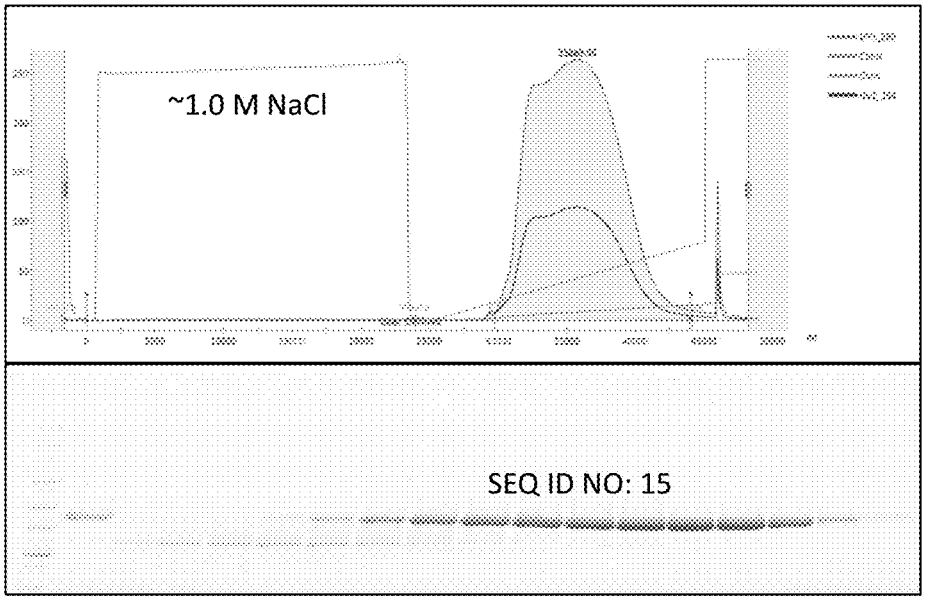


FIG. 22B

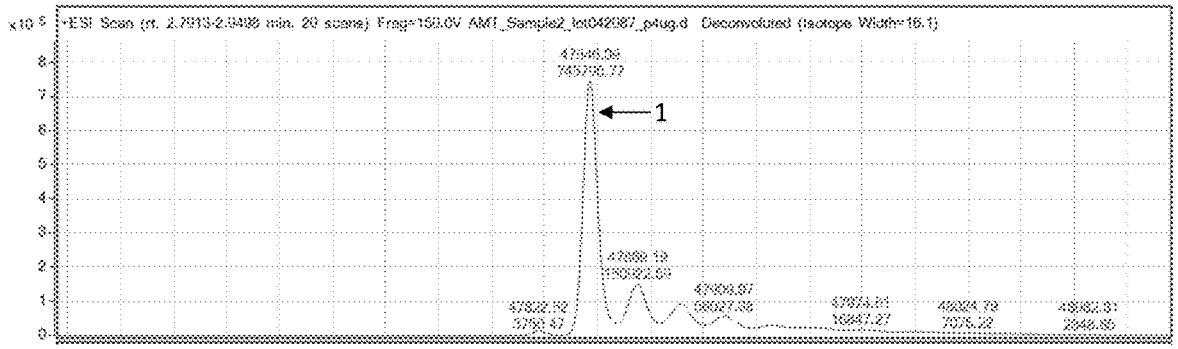


FIG. 23A

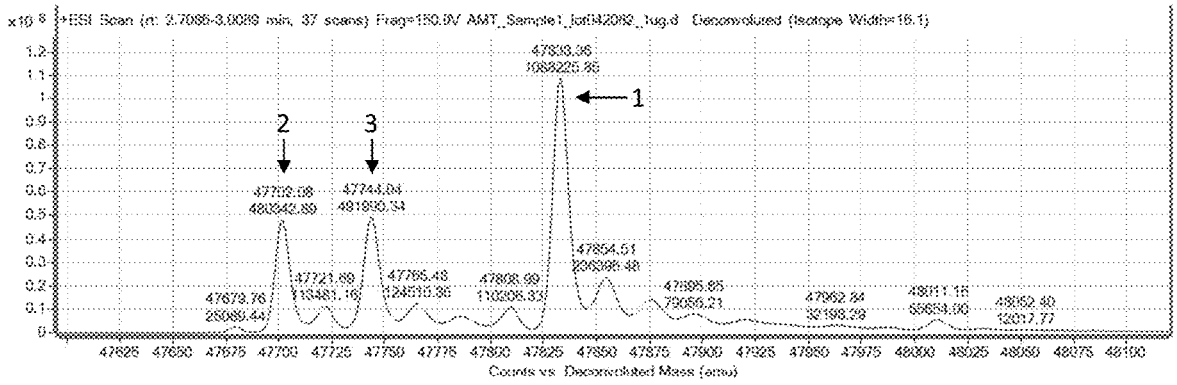


FIG. 23B

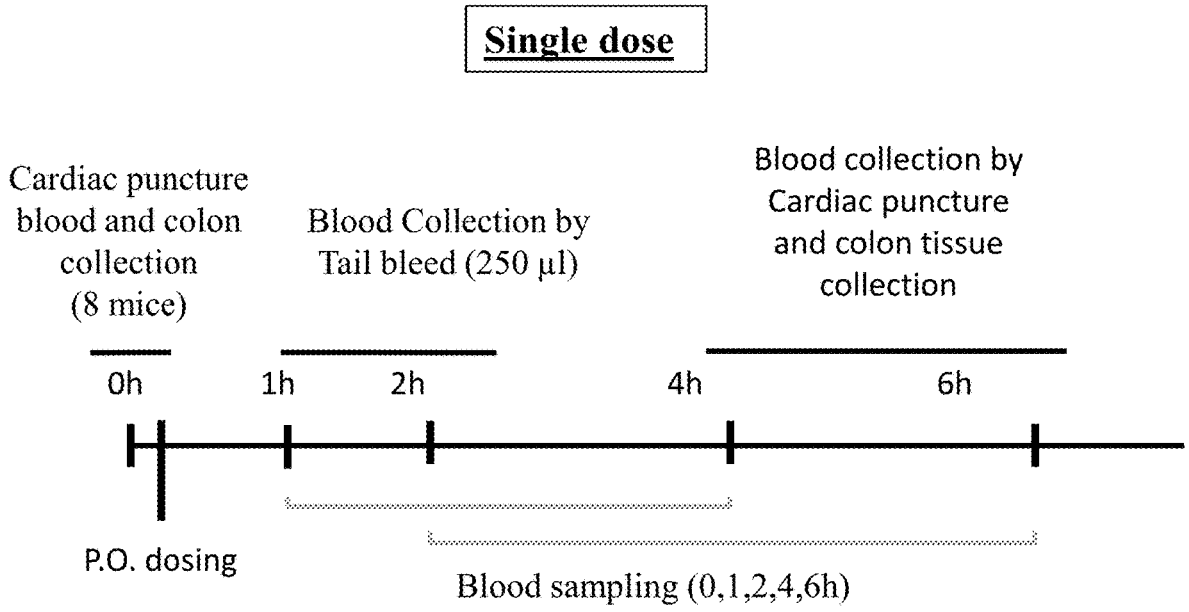

Single dose
Cardiac puncture blood and colon collection (8 mice)
Blood Collection by Tail bleed (250 µl)
Blood collection by Cardiac puncture and colon tissue collection
0h
1h
2h
4h
6h
P.O. dosing
Blood sampling (0,1,2,4,6h)

FIG. 24

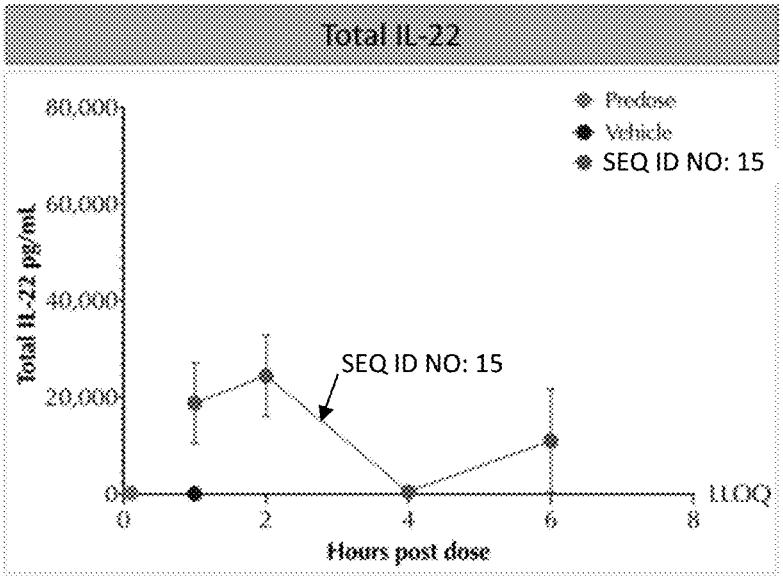


FIG. 25A

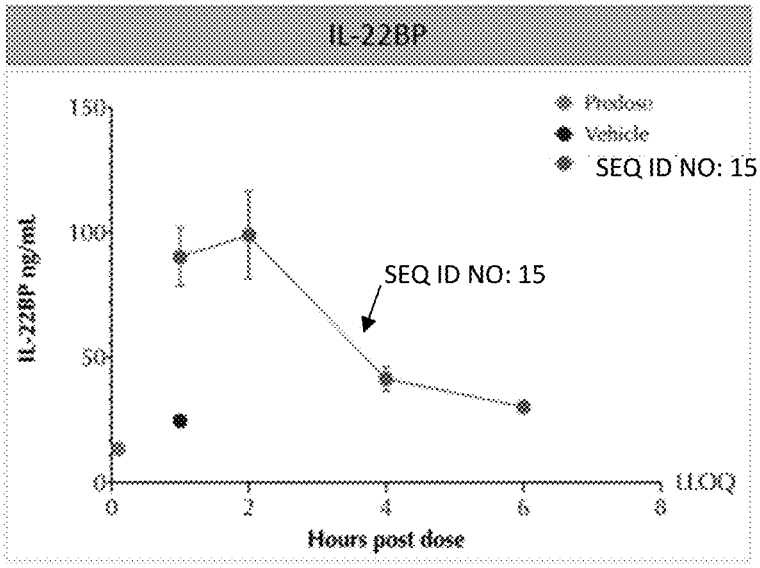


FIG. 25B

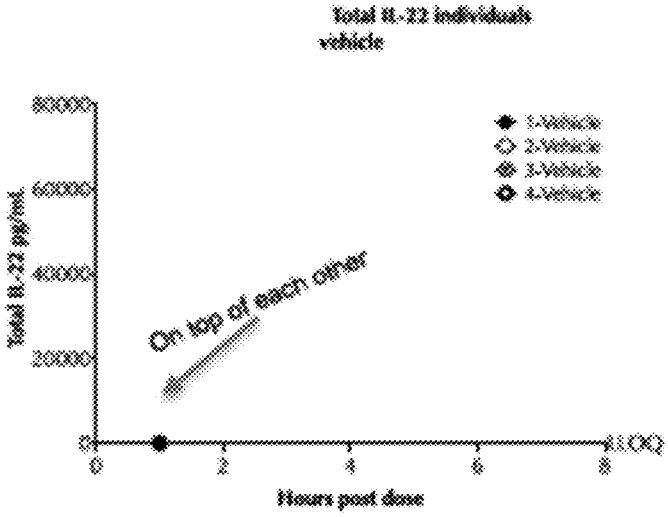


FIG. 26A

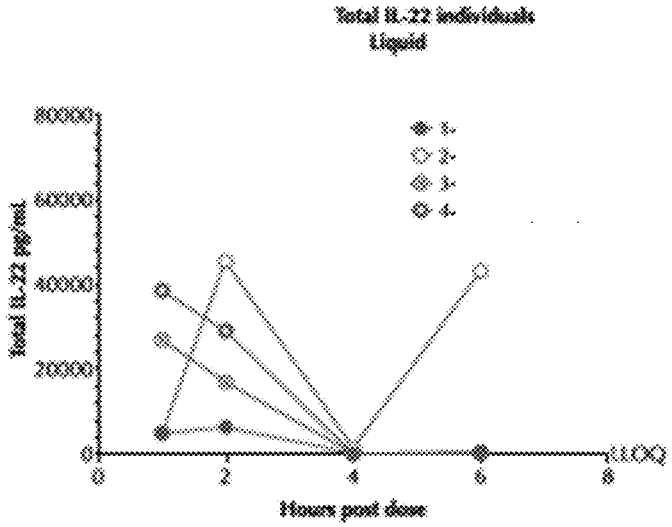


FIG. 26B

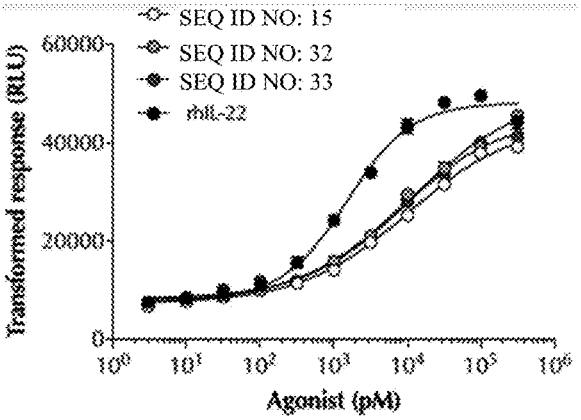


FIG. 27A

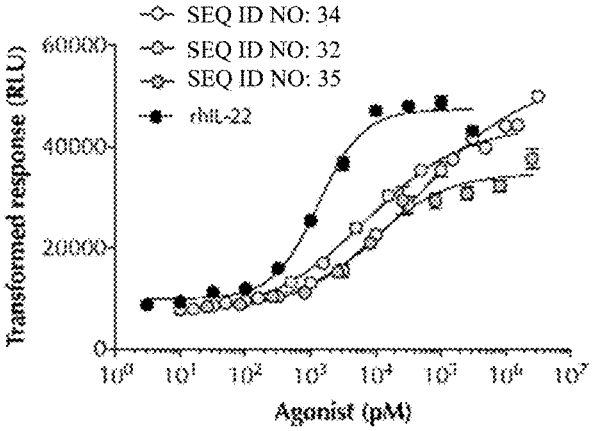


FIG. 27B

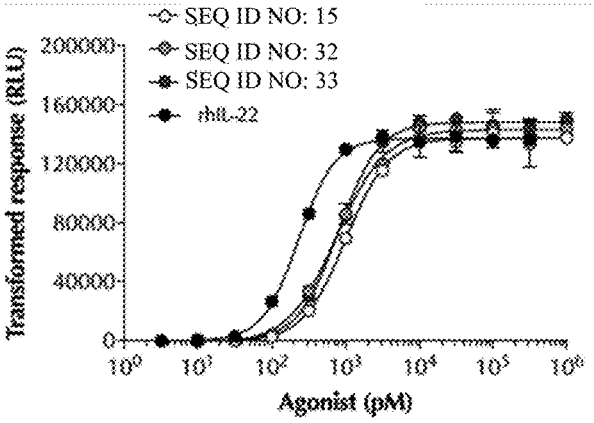


FIG. 27C

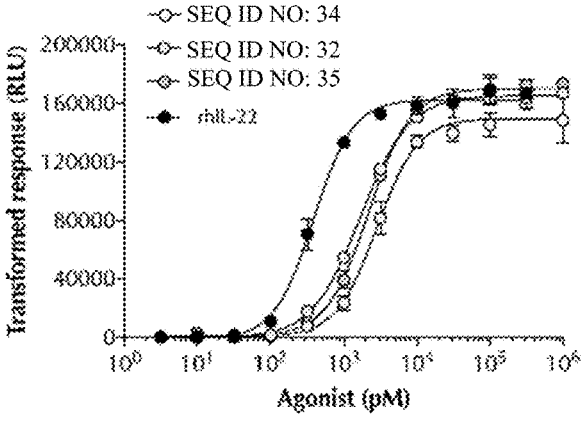


FIG. 27D

DELIVERY CONSTRUCTS FOR TRANSCYTOSIS AND RELATED METHODS

CROSS-REFERENCE

[0001] This application is a Continuation Application of International Application No. PCT/US2019/060356, filed Nov. 7, 2019, which claims the benefit of U.S. Provisional Application No. 62/756,889, filed Nov. 7, 2018, U.S. Provisional Application No. 62/888,133, filed Aug. 16, 2019, and U.S. Provisional Application No. 62/888,238 filed Aug. 16, 2019, International Application No. PCT/US2019/050708 filed Sep. 11, 2019, and International Application No. PCT/US2019/021474 filed Mar. 8, 2019, which applications are incorporated herein by reference in their entirety.

BACKGROUND

[0002] The gut epithelium has thwarted efforts to orally administer large therapeutic molecules such as proteins because proteins cannot diffuse across the intact epithelial barrier or cross the barrier through the tight junctions. Once taken up by an epithelial cell, a therapeutic protein can enter the destructive lysosomal trafficking pathway, or can be released back into the intestinal lumen. This inability to be readily transported across the intestinal epithelium can be a limiting factor in developing commercially viable oral formulations, particularly for polypeptide-based therapeutics.

[0003] Parenteral administration such as intravenous or subcutaneous administration can be a solution, but these administration routes can often create considerable side effects, lower the therapeutic efficacy, and reduce patient convenience that can negatively affect compliance. There is a need for improved compositions and methods for transporting therapeutics across an epithelium, e.g., a gut epithelium.

[0004] Additionally, purification and refolding of biologically active polypeptides in order to obtain correctly folded, biologically active, and stable polypeptides in high yields and with low endotoxin levels is still considered one of the most challenging aspects for a cost- and resource-effective production of biological therapeutics (e.g., polypeptides). Thus, there is also a need for improved methods for the production (fermentation, refolding, purification, and formulation) of such biologically active molecules.

SUMMARY

[0005] In the various aspects, the present disclosure provides a delivery construct comprising an amino acid sequence set forth in SEQ ID NO: 15 or SEQ ID NO: 17, or an amino acid sequence having at least 90%, 95% or 99% sequence identity thereto.

[0006] In the various aspects, the present disclosure provides a delivery construct comprising a carrier consisting of an amino acid sequence set forth in SEQ ID NO: 7 or SEQ ID NO: 9. In some instances, the carrier is coupled to a heterologous payload. In some instances, the heterologous payload is a human IL-22. In some instances, the human IL-22 consists of an amino acid sequence set forth in SEQ ID NO: 11. In some instances, the carrier is coupled covalently or non-covalently to the IL-22. In some instances, the carrier is coupled covalently to the IL-22 via a spacer. In some instances, the spacer consists of an amino acid sequence set forth in SEQ ID NO: 13. In some instances, the delivery construct consists of an amino acid sequence set forth in SEQ ID NO: 15 or SEQ ID NO: 17.

[0007] Provided herein is a method of treating an inflammatory disease in a subject, the method comprising administering to the subject an effective amount of a delivery construct described herein (e.g., SEQ ID NO: 15 or SEQ ID NO: 17). In some instances, the inflammatory disease is hepatitis, obesity, fatty liver disease, liver inflammation, or pancreatitis, Crohn's disease, ulcerative colitis, pouchitis, proctitis, multiple sclerosis, systemic lupus erythematosus, graft versus host disease, rheumatoid arthritis, or psoriasis. In some instances, the disease is Crohn's disease or ulcerative colitis.

[0008] Described herein, in certain embodiments, are methods for obtaining a purified non-naturally occurring fusion protein, the method comprising: performing anion exchange chromatography on a mixture comprising the non-naturally occurring fusion protein to obtain a first fraction comprising the non-naturally occurring fusion protein; wherein the non-naturally occurring fusion protein comprises IL-22 and a carrier. In some embodiments, wherein performing anion exchange chromatography comprises binding the non-naturally occurring fusion protein to anionic exchange resin and providing an increasing salt gradient for subsequent elution of the non-naturally occurring fusion protein to obtain the first fraction. In some embodiments, performing anion exchange chromatography comprises contacting the mixture with a resin comprising amine-functionalized polymethacrylate beads. In some embodiments, the resin is an NH_2-750F resin.

[0009] In some embodiments, the method further comprises refolding the non-naturally occurring fusion protein prior to performing anion exchange chromatography. In some embodiments, refolding the non-naturally occurring fusion protein comprises contacting chaotrope-solubilized protein from inclusion bodies with a refolding solution, wherein the refolding solution comprises: arginine (0.75 M to 1.25 M); glycerol (2% to 20% v/v); cysteine (0.5 mM to 10 mm); and cystamine (0.2 mM to 10 mM); wherein the refolding solution has a pH of between 7.5 and 8.5. In some embodiments, the arginine is present in the refolding solution at a concentration of between 0.9 M and 1.1 M; glycerol is present in the refolding solution at a concentration of between 7% and 13% (w/w); cysteine is present in the refolding solution at a concentration of between 1.5 mM and 6 mM; cystamine is present in the refolding solution at a concentration of between 0.6 mM and 3 mM; and the refolding solution has a pH of between 7.8 and 8.2.

[0010] In some embodiments, the method further comprises subjecting a sample comprising the first fraction to a hydroxyapatite resin to obtain a second fraction comprising the non-naturally occurring fusion protein. In some embodiments, the hydroxyapatite resin is a CaPure-hydroxyapatite resin. In some embodiments, the method further comprises performing cation exchange chromatography on a sample comprising the first fraction. In some embodiments, performing cation exchange chromatography comprises contacting the sample comprising the first fraction with a resin comprising sulfate-functionalized polymethacrylate beads. In some embodiments, the resin is a TOYOPEARL Sulfate-650F resin.

[0011] In some embodiments, upon contact with a cell, the carrier promotes endocytosis or transcytosis of the non-naturally occurring fusion protein. In some embodiments,

upon contact with the cell, the carrier promotes transcytosis of the non-naturally occurring fusion protein. In some embodiments, the cell is a gut epithelial cell. In some embodiments, the gut epithelial cell is a polarized gut epithelial cell. In some embodiments, the carrier is a truncated variant of a naturally occurring or non-naturally occurring cholix polypeptide. In some embodiments, the carrier has at least 70%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% sequence identity to any one of SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7, 8, 9, 22, or 23. In some embodiments, the non-naturally occurring fusion protein has at least 70%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% sequence identity to the sequence of any one of SEQ ID NOS: 14-21. In some embodiments, IL-22 has at least 85%, 90%, 95%, 98%, 99% or 100% sequence identity to any one of SEQ ID NOS: 10, 11, or 12.

[0012] In some embodiments, the carrier, by itself, has a first isoelectric point (pI) and the IL-22, by itself, has a second isoelectric point, wherein the first isoelectric point is at least 1 pH unit, at least 1.5 pH units, at least 1.7 pH units, or at least 2 pH units lower than the second isoelectric point. For instance, in some embodiments, the carrier has a pI of between 4.8 and 5.4, between 4.9 and 5.3, between 5.0 and 5.2, such as a pI of about 5.1. In some embodiments, the pI of the IL-22 is between about 6.8 and 7.4, such as between about 6.9 and 7.3, between about 7.0 and 7.2, such as about 7.1. In some embodiments, the non-naturally occurring fusion protein is obtained by any of the methods described herein.

[0013] Described herein, in certain embodiments, are methods of refolding a non-naturally occurring fusion protein comprising a carrier and IL-22, the method comprising: (i) contacting inclusion bodies comprising the non-naturally occurring fusion protein with a solubilization solution comprising a chaotropic agent to produce a soluble non-naturally occurring fusion protein; (iii) contacting the non-naturally occurring fusion protein with a refolding solution, wherein the refolding solution comprises: arginine (0.75 M to 1.25 M); glycerol (2% to 20% v/v); cysteine (0.5 mM to 10 mm); and cystamine (0.2 mM to 10 mM); wherein the refolding solution has a pH of between 7.5 and 8.5.

[0014] In some embodiments, arginine is present in the refolding solution at a concentration of between 0.9 M and 1.1 M; glycerol is present in the refolding solution at a concentration of between 7% and 13% (w/w); cysteine is present in the refolding solution at a concentration of between 1.5 mM and 6 mM; cystamine is present in the refolding solution at a concentration of between 0.6 mM and 3 mM; and the refolding solution has a pH of between 7.8 and 8.2.

INCORPORATION BY REFERENCE

[0015] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] The novel features of the disclosure are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure are utilized, and the accompanying drawings of which:

[0017] FIG. **1** schematically shows a setup comprising an apical chamber above an epithelial cell monolayer and a basal chamber below such epithelial cell monolayer. For apical to basolateral permeability, test articles (e.g., delivery constructs, payloads, etc.) were applied to the apical (A) side and the amount of permeated (e.g., transcytosed) material was determined on the basolateral (B) side.

[0018] FIG. **2** shows that a delivery construct (SEQ ID NO: 15), which includes a carrier (SEQ ID NO: 7) coupled to an IL-22 (SEQ ID NO: 11) via a spacer (SEQ ID NO: 13), transported the IL-22 payload across intact, polarized, Caco-2 gut epithelial cell monolayers in a time-dependent manner (the amount of protein on the basolateral site was measured 15, 30, and 45 minutes after the delivery construct was applied to the basal membrane of the monolayer as described above in FIG. **1**). The data further shows that when the delivery construct with a carrier of SEQ ID NO: 7 and IL-22 (SEQ ID NO: 11) is applied to the to the Caco-2 epithelial cells, about 2-3 fold more IL-22 crossed the Caco-2 epithelial cell monolayer as compared to when an IL-22 (SEQ ID NO: 12) was not coupled to a carrier.

[0019] FIG. **3** shows that a delivery construct (SEQ ID NO: 15) resulted in IL-22 (SEQ ID NO: 11) being transported across intact and polarized SMI-100 gut epithelial cell monolayers in a time-dependent manner (the amount of protein in the basolateral chamber was measured at 15, 30, and 45 minutes after the delivery construct was applied to the basal membrane of the monolayer). The data further shows that when the delivery construct including the carrier with SEQ ID NO: 7 coupled to IL-22 (SEQ ID NO: 11) is applied to the SMI-100 epithelial cells, about 2-3 fold more IL-22 crossed the SMI-100 epithelial cell monolayer as compared to when an IL-22 (SEQ ID NO: 12) was not coupled to a carrier.

[0020] FIG. **4** demonstrates that a delivery construct (SEQ ID NO: 15) including a carrier (SEQ ID NO: 7) coupled to an IL-22 (SEQ ID NO: 11) via a spacer (SEQ ID NO: 13) results in IL-22 being transported in significant amounts across an intact and polarized gut epithelium in vivo. The apical site of the gut epithelium is highlighted by white arrow #1. The *lamina propria* is abbreviated as "l.p." The outer basal membrane of the polarized epithelium is highlighted by white arrow #2. IL-22 localization is indicated white arrow #3).

[0021] FIG. **5**A demonstrates STAT3 phosphorylation in murine FL83 B cells as a function of agonist concentration (in pM), wherein the agonist is recombinant human IL-22 (rhIL-22, SEQ ID NO: 12) or recombinant murine IL-22 (rmIL-22; MAVLQKSMSFSLMGTLAASCLLLIALWAQE ANALPVNTRCKLEVSNFQQPYIVNRTFMLAKEAS-LADNNTDVRLIGEKLFRGVSAKDQCYL MKQVLN-FTLEDVLLPQSDRFQPYMQEVVPFLTKLSNQLSS-CHISGDDQNIQKNVRRLKETV KKLGESGEIKAIGELDLLFMSLRNACV (SEQ ID NO: 30)). The data show that rhIL-22 and rmIL-22 induced STAT3 phosphorylation in a concentration-dependent manner, demonstrating that the human IL-22 protein can bind and induce signal transduction through mouse IL-22 receptors as effectively as mouse IL-22.

[0022] FIG. **5**B demonstrates that rhIL-22 (SEQ ID NO: 12) and rmIL-22 also induced STAT3 phosphorylation in murine Hepa1-6 cells.

[0023] FIG. **6**A demonstrates that rhIL-22 (SEQ ID NO: 12) and rmIL-22 induced STAT5 phosphorylation in murine FL83 B cells in a dose-dependent manner, demonstrating that the human IL-22 protein can bind and induce signal transduction through mouse IL-22 receptors as effectively as mouse IL-22.

[0024] FIG. **6**B demonstrates that rhIL-22 (SEQ ID NO: 12) and rmIL-22 also induced STAT5 phosphorylation in murine Hepa1-6 cells.

[0025] FIG. **7**A shows an overview of an in vivo study design and timeline for comparing peroral (p.o.) administration of the delivery construct with SEQ ID NO: 15 at two once daily dose levels, 1 mg/kg and 30 mg/kg, and one once daily administration (intraperitoneal (i.p.)) of rhIL-22 at 4 mg/kg. The in vivo study was performed in Normal chow-fed, 8-10 week old, ~23 g, female C57BL/6 mice. Colitis was chemically induced with 2.5% dextran sulfate sodium (DSS) in drinking water, ad libitum. Study endpoints included percent change in body weight, disease activity index, colon length and weight, and histopathology.

[0026] FIG. **7**B shows characteristics of the experimental and control groups used in the study described in FIG. **7**A.

[0027] FIG. **8**A shows the percent (%) change in body weight of animals at study day 10 relative to day 0 of the various experimental and control groups. Baseline (day 0 vs. 10) body weight change by rhIL-22 with SEQ ID NO: 12 or the delivery construct with SEQ ID NO: 15 relative to vehicle was not significant as assessed by 1-way ANOVA. Positive model control (cyclosporine) CsA was significantly different relative to Vehicle as assessed by a 1-way ANOVA.

[0028] FIG. **8**B shows the change in body weight of experimental and control animals over the first 10 study days. CsA group mean body weight over the study period was significantly improved relative to vehicle control day 6 through day 10 as assessed by 2-way ANOVA(*, **, ***) rhIL-22 (i.p.) body weight was significantly improved relative to vehicle control at day 10 as assessed by 2-way ANOVA (*), *p<0.05, **p<0.01, ***p<0.001.

[0029] FIG. **9** shows results of plasma ELISA experiments to determine IL-22 plasma levels. The results show that plasma rhIL-22 concentrations trended towards increased levels in plasma after oral gavage of both 1 and 30 mg/kg doses of the delivery construct with SEQ ID NO: 15 in a dose-dependent manner (*p<0.05, **p<0.01, ***p<0.001. ****p<0.0001; Mean+SEM; n=2 Naïve, 5 Vehicle, 5 CsA, 10 others)

[0030] FIG. **10**A shows an overview of a single dose in vivo study designed to identify biomarker(s) of target engagement after single (acute dosing) of rhIL-22 in healthy CD-1 mice.

[0031] FIG. **10**B shows an overview of a multiple dose (sub-chronic dosing) in vivo study designed to identify biomarker(s) of target engagement after single and multiple doses of rhIL-22 in healthy CD-1 mice.

[0032] FIG. **11**A shows that acute and sub-chronic administration of rhIL-22 increases IL-22 concentration consistently.

[0033] FIG. **11**B shows some pharmacokinetic parameters measured during the acute and sub-chronic dosing experiments described in FIG. **11**A.

[0034] FIG. **12**A shows the induced circulating murine C-reactive protein (mCRP) concentration in response to 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12) and after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12). The results show a time-dependent increase in circulating mCRP in both study groups.

[0035] FIG. **12**B shows fold change of plasma mCRP concentration 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12).

[0036] FIG. **12**C shows fold change of plasma mCRP concentration after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12).

[0037] FIG. **13**A shows the induced circulating murine serum amyloid protein A (mSAA) concentration in response to 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12) and after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12). The results show a time-dependent increase in circulating mSAA in both study groups, with approximate plasma peak concentrations about 8 hours after administration

[0038] FIG. **13**B shows fold change of plasma mSAA concentration 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12).

[0039] FIG. **13**C shows fold change of plasma mSAA concentration after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12).

[0040] FIG. **14**A shows the induced circulating regenerating islet-derived protein 3β (Reg3β) in response to 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12) and after 5 days of sub-chronic dosing rhIL-22 (SEQ ID NO: 12).

[0041] FIG. **14**B shows fold change of plasma Reg3β concentration 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12).

[0042] FIG. **14**C shows fold change of plasma Reg3/3 concentration after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12).

[0043] FIG. **15** illustrates a flow chart showing how improved manufacturability may enable clinical translation.

[0044] FIG. **16** illustrates that SEQ ID NO: 15 is a heterologous fusion protein comprising a carrier domain (SEQ ID NO: 7) linked via a spacer (SEQ ID NO: 13) to interleukin-22 (IL-22, SEQ ID NO: 11). The heterologous fusion proteins may be sequestered in *E. coli* inclusion bodies (IB) and thus require refolding and/or purification.

[0045] FIG. **17** illustrates a flow chart summarizing the refolding, purification, and formulation of heterologous fusion proteins described herein, such as, for example SEQ ID NOS: 14-21, and fusion proteins comprising a carrier of SEQ ID NO: 1-9, 22, or 23.

[0046] FIG. **18** illustrates iterative optimization for the improved refolding efficiency of the construct of SEQ ID NO: 15.

[0047] FIG. **19** is a size-exclusion chromatograph (SEC) showing the signal of the optimized refolded protein (SEQ ID NO: 15) and protein aggregates having higher molecular masses and thus lower retention times.

[0048] FIG. **20** illustrates an exemplary SDS-PAGE of the initial and optimized refolded protein (SEQ ID NO: 15).

[0049] FIGS. **21**A-**21**B illustrates a size-exclusion chromatograph following a first column purification of SEQ ID NO: 15 using anion exchange resin NH_2-750F. FIG. **21**A illustrates a size-exclusion chromatograph following a first column purification of SEQ ID NO: 15 using anion exchange resin NH_2-750F with a column volume of 10 mL and a bed height of 20 cm. FIG. **21**B illustrates a size-exclusion chromatograph following a first column purifica-

tion of SEQ ID NO: 15 using anion exchange resin NH_2-750F with column volume of 4.6 L and a bed height of 30 cm.

[0050] FIGS. **22**A-**22**B illustrate a size exclusion chromatograph following a second column purification of SEQ ID NO: 15 using hydroxyapatite resin CaPure®. FIG. **22**A illustrates a size exclusion chromatograph following a second column purification of SEQ ID NO: 15 using hydroxyapatite resin Ca^{++}Pure with a column volume of 5 mL, a bed height of 10 cm, and a gradient of 0-25% B, 25C. FIG. **22**B illustrates a size exclusion chromatograph following a second column purification of SEQ ID NO: 15 using hydroxyapatite resin CaPure® with a column volume of 800 mL, a bed height of 21 cm, and a gradient of 0-25% B, 25CV. For each figure a corresponding SDS-PAGE of samples from the different fractions eluted from the CaPure® resin.

[0051] FIGS. **23**A-**23**B illustrate a liquid chromatography-mass spectrophotometry (LC-MS) chromatograph following purification of heterologous fusion proteins. FIG. **23**A illustrates an LC-MS chromatograph of SEQ ID NO: 17. Peak 1 illustrates correctly refolded SEQ ID NO: 17. FIG. **23**B illustrates an LC-MS chromatograph of SEQ ID NO: 15. Peak 1 illustrates corrected refolded SEQ ID NO: 15. Peak 2 illustrates SEQ ID NO: 15 with the terminal methionine cleaved (illustrated by SEQ ID NO: 20). Peak 3 illustrates SEQ ID NO: 15 with the terminal methionine cleaved (illustrates by SEQ ID NO: 20) and acetylation of the resulting N-terminal amino acid.

[0052] FIG. **24** illustrates the study overview of EXAMPLE 16 for assessing transport of the delivery construct having the sequence of SEQ ID NO: 15 across the gut epithelium and into circulation after oral (abbreviated herein as P.O.) delivery of the construct.

[0053] FIGS. **25**A-**25**B illustrate graphs showing IL-22 plasma concentration as a function of time after administration of the delivery construct of SEQ ID NO: 15. FIG. **25**A illustrates total IL-22 plasma concentration as a function of time after administration of the delivery construct of SEQ ID NO: 15. FIG. **25**B illustrates IL-22BP plasma concentration as a function of time after administration of the delivery construct of SEQ ID NO: 15.

[0054] FIG. **26**A (vehicle) and FIG. **26**B (SEQ ID NO: 15) are graphs showing IL-22 plasma concentration as a function of time for individual mice.

[0055] FIGS. **27**A-**27**D illustrate that spacer length and coupling of payload to the N- or C-terminus of a carrier does not significantly affect a payload's biological activity. FIG. **27**A illustrates the length of various amino acid spacers did not affect the induction of IL-22 receptor dimerization.

[0056] FIG. **27**B illustrates that coupling the IL-22 payload to the N- or C-terminus of a carrier did not affect induction of IL-22 dimerization. FIG. **27**C illustrates the length of various amino acid spacers did not affect induction of pSTAT3 activation. FIG. **27**D illustrates that coupling the IL-22 payload to the N- or C-terminus of a carrier did not affect induction of pSTAT3 activation.

DETAILED DESCRIPTION

[0057] The present disclosure describes non-naturally occurring fusion proteins (e.g. delivery constructs capable of transporting one or more heterologous payload molecules) and methods for the refolding and purification of non-naturally occurring fusion proteins. For example, the non-naturally occurring fusion protein can be an IL-22 delivery construct described herein.

[0058] The below terms are discussed to illustrate meanings of the terms as used in this specification, in addition to the understanding of these terms by those of skill in the art. As used herein and in the appended claims, the singular forms "a," "an," and, "the" include plural referents unless the context clearly dictates otherwise. It is further noted that the claims can be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely," "only," and the like in connection with the recitation of claim elements, or use of a "negative" limitation.

[0059] Certain ranges or numbers are presented herein with numerical values being preceded by the term "about." The term "about" is used herein shall mean plus or minus 1%, 2%, 3%, 4%, or 5% of the number that the term refers to. As used herein, the terms "subject" and "individual," are used interchangeably and can be any animal, including mammals (e.g., a human or non-human animal).

[0060] As used herein, the terms "treat," "treating," or "treatment," and other grammatical equivalents, include alleviating, abating or ameliorating one or more symptoms of a disease or condition, ameliorating, preventing or reducing the appearance, severity or frequency of one or more additional symptoms of a disease or condition, ameliorating or preventing the underlying causes of one or more symptoms of a disease or condition, inhibiting the disease or condition, such as, for example, arresting the development of the disease or condition, relieving the disease or condition, causing regression of the disease or condition, relieving a condition caused by the disease or condition, or inhibiting the symptoms of the disease or condition either prophylactically and/or therapeutically.

[0061] As described herein, the term "percent (%) sequence identity," and terms related thereto, in the context of amino acid sequences, is the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a selected sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as Clustal Omega, BLAST, BLAST-2, ALIGN, ALIGN-2 or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full-length of the sequences being compared.

[0062] The terms "polypeptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. Peptides also include essentially any polyamino acid including, but not limited to, peptide mimetics such as amino acids joined by an ether as opposed to an amide bond.

Payloads and Delivery Constructs

[0063] Provided herein are delivery constructs capable of transporting one or more heterologous payload molecules (e.g., one or more therapeutic payloads) across epithelial cells (e.g., polarized gut epithelial cells) and into the *lamina propria* via transcytosis. The delivery constructs can com-

prise a carrier coupled to a heterologous payload. The delivery construct can be a non-naturally occurring fusion protein. The carrier can be capable of transporting the heterologous payload across polarized epithelial cells (e.g., polarized gut epithelial cells) using endogenous trafficking pathways. Utilization of endogenous trafficking pathways, as opposed to use of passive diffusion, can allow the carrier to shuttle the heterologous payload rapidly and efficiently across epithelial cells without impairing the barrier function of these cells or the biological activity of the heterologous payload. Further provided herein are methods for purification and refolding of the delivery constructs described herein.

[0064] A carrier herein may be derived from a polypeptide secreted by a bacterium. Such carrier may be derived from a polypeptide secreted from *Vibrio Cholerae* or *Pseudomonas aeruginosa*. The polypeptide secreted by *Vibrio Cholerae* can be a Cholix polypeptide. A carrier derived from a Cholix polypeptide can be naturally occurring or non-naturally occurring. For example, a non-naturally occurring Cholix polypeptide can consist of the amino acid sequence set forth in SEQ ID NO: 1 (an example of a Cholix$^{1-634}$) (TABLE 1). A carrier derived from a Cholix polypeptide can be a truncated and/or mutated variant of a polypeptide derived from Cholix. For example, the carrier can comprise or consist of an amino acid sequence of those with amino acid residues 1-206, 1-245, 1-251, 1-266, and 1-386 of SEQ ID NO: 1 or SEQ ID NO: 26. In some instances, such carriers have an amino acid sequence of those with amino acid residues 1-206, 1-245, 1-251, 1-266, and 1-386 of SEQ ID NO: 4. Mutation(s) can include one or more substitution (s), deletion(s), and/or addition(s). For example, a carrier herein can comprise a V1L substitution. Stated differently, in some embodiments, the cholix-related carrier has a leucine amino acid at position "1." (Position 1 refers to the first amino acid of variants that do not have an N-terminal methionine or the second position in variants that include an N-terminal methionine. In other words, in determining the length of a carrier, an N-terminal methionine, if present, is ignored.) In some embodiments, carriers comprising the V1L substitution experience reduced or eliminated cleavage of the N-terminal amino acid. In some embodiments, carriers comprising the V1L substitution experience reduced or eliminated acetylation of the N-terminal amino acid. A carrier provided herein can have a reduced (e.g., at least 50% reduced) or ablated ADP ribosylation activity (e.g., ribosylation of elongation factor 2) relative to a naturally-occurring Cholix variant. In some embodiments, the carrier can comprise an N-terminal methionine. In other embodiments, no N-terminal methionine is present.

[0065] A truncated Cholix carrier can consist of, consist essentially of, or comprise amino acid residues 1-386 of a sequence set forth in SEQ ID NO: 26 (FORMULA I). A truncated Cholix carrier can consist of, consist essentially of, or comprise amino acid residues 1-266 of a sequence set forth in SEQ ID NO: 26 (FORMULA I). In such instances, a carrier can consist of, consist essentially of, or comprise amino acid residues 1-266 of a sequence set forth in SEQ ID NO: 1. Thus, in some instances, the carrier consists of the amino acid sequence set forth in SEQ ID NO: 2 (an example of Cholix$^{1-386}$) or SEQ ID NO: 3 (an example of Cholix$^{1-266}$). In some instances, a carrier has the amino acid sequence represented by SEQ ID NO: 2 with a V1L substitution. Thus, in some instances, the carrier consists of the amino acid sequence set forth in SEQ ID NO: 4 (an example of V1L-Cholix$^{1-386}$) or SEQ ID NO: 5 (an example of V1L Cholix$^{1-266}$). Any of these carriers can include one or more amino acids at its N-terminus for expression in various microorganisms (e.g., bacteria), e.g., an N-terminal methionine. Such carrier can have an amino acid sequence set forth in SEQ ID NOS: 6-9.

[0066] The Cholix polypeptide can be a protein comprising, consisting essentially of, or consisting of an amino acid sequence having at least 80%, 85%. 90%, 95%, 98%, or 99% sequence identity, or having 100% sequence identity, to an amino acid sequence set forth in any one of SEQ ID NOS: 1-9 or SEQ ID NOS: 22-23. An example of a Cholix polypeptide is provided herein as SEQ ID NO: 1. or SEQ ID NO: 22. Also contemplated herein are truncated Cholix polypeptide variants that are able to transport a payload across polarized epithelia cells (e.g., polarized gut epithelial cells). Such Cholix polypeptides can be truncated at any one of the amino acid positions from 206 to 633 as compared to a reference sequence, e.g. SEQ ID NO: 1, SEQ ID NO: 22, or SEQ ID NO: 26, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 98% or 99% sequence identity thereto.

[0067] Also contemplated herein are delivery constructs comprising carriers having high sequence identity to the sequences above. Such high sequence identity can include, at least 90%, 95%, 96%, 97%, 98% or 99% sequence identity. Thus, in some instances, the carrier comprises a sequence identify of at least 90%, 95%, 96%, 97%, 98% or 99% sequence identity to an amino acid sequence set forth in SEQ ID NO: 6-9.

[0068] A carrier contemplated herein can be coupled to a payload, such as a heterologous payload. Such payload can be a therapeutic payload. A therapeutic payload can be a cytokine, a hormone, a growth factor, a therapeutic antibody, an antigen, a functional fragment of any of the above, or any other protein that has biological, therapeutic activity, or a protein that may be deficient in a subject (e.g., a genetic/inherited deficiency of a certain protein). Cytokines contemplated herein include monomeric chemokines and interleukins (also abbreviated herein as "ILs"). The interleukin can be IL-22. The interleukin may be from any species (e.g., from a human or a rodent). The interleukin may be a human interleukin. Human IL-22 can have the amino acid sequence set forth in SEQ ID NO: 10 (IL-22$^{1-179}$) or SEQ ID NO: 11 (IL-22$^{34-179}$). An IL-22 herein can further include a methionine at its N-terminus, e.g., when such IL-22 protein is bacterially expressed. In one instance, an IL-22 has an amino acid sequence set forth in SEQ ID NO: 12 (M+IL-22$^{34-179}$). An IL-22 herein can further include a methionine at its N-terminus, e.g., when such IL-22 protein is bacterially expressed. In one instance, an IL-22 has an amino acid sequence set forth in SEQ ID NO: 12 (M+IL-22$^{34-179}$). The IL-22 can comprise, consist essentially of, or consist of SEQ ID NO: 10, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 98% or 99%% sequence identity thereto or fragment thereof. The IL-22 can comprise, consists essentially of, or consist of SEQ ID NO: 11 (IL-22$^{34-179}$), which is a secreted form of IL-22, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 98% or 99%% sequence identity thereto or fragment thereof. An IL-22 can include a methionine at its N-terminus, e.g., when such IL-22 protein is bacterially expressed. Such IL-22 can

consist of, consist essentially of, or comprise an amino acid sequence set forth in SEQ ID NO: 12 (M+IL-22^{34-179}).

[0069] In some instances, a carrier used in any of the delivery constructs herein can be a protein or another type of molecule capable of transporting the therapeutic payload across an epithelium (e.g., a polarized gut epithelium of a subject). Such transport can include transcytosis. As referred to herein, "transcytosis" refers to the trafficking of the fusion molecule through a polarized epithelial cell. Such trafficking permits the release of the biologically active cargo from the basolateral membrane of the polarized epithelial cell. The transcytosis process may involve interaction(s) of the carrier with one or more receptor(s) and/or protein(s) on the apical and/or basal surface(s) as well as inside a cell of the epithelium (e.g., a polarized gut epithelial cell). The carrier can be capable of transporting the therapeutic payload across an epithelium without impairing the epithelium, the carrier, and/or the biological and/or therapeutic function of the payload.

[0070] A carrier can be coupled to a therapeutic payload covalently or non-covalently and directly or indirectly. The therapeutic payload can be directly coupled to the N-terminus or C-terminus of the carrier. In instances where the carrier is covalently coupled to the payload, the carrier can be coupled to such payload via a spacer (also referred to herein as a linker). The spacer can comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or more amino acid residues. Spacer amino acid residues can be e.g., glycine, serine, and/or tryptophan. A spacer may be a cleavable or a non-cleavable spacer. A cleavable spacer may be cleavable by an enzyme or in a pH-dependent manner.

[0071] Examples of spacers contemplated herein include oligopeptide sequences such as S, (GS)x, (GGS)x, (GGGS)x (SEQ ID NO: 27), (GGGGS)x (SEQ ID NO: 28), or (GGGGGS)x (SEQ ID NO: 29), wherein x=1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15. In some cases, a spacer consists of, consists essentially of, or comprises a sequence set forth in SEQ ID NO: 13 ($(G_4S)_3$). (GGGGSGGGGSGGGGS). In some cases, a spacer consists of, consists essentially of, or comprises a sequence set forth in SEQ ID NO: 31 ($(G_4S)_5$). (GGGGSGGGGSGGGGSGGGGSGGGGS). In some instances, a delivery construct comprises a therapeutic payload and a carrier that are non-covalently linked (e.g., via ionic interactions, van der Waals interactions, π-π interactions, etc.). A carrier can further comprise one or more features, elements, amino acids, or modifications on its N-terminus and/or C-terminus. For instance, some embodiments include a N-terminal methionine at the N-terminus of the carrier. Other modifications (e.g., acetylation) may also be present, including modifications for expression in a heterologous system.

[0072] A delivery construct herein can have an amino acid sequence set forth in SEQ ID NO: 15 or 17 (examples of M+Cholix$^{1-266}$-$(G_4S)_3$—IL-22^{34-179}) (TABLE 1). Other delivery constructs can have the amino acid sequence set forth in SEQ ID NOs: 24 or 25 (e.g., when expressed in a mammalian cell such as CHO cell). Such exemplary delivery constructs transport IL-22 across intact, polarized gut epithelial cells and into the *lamina propria*.

[0073] In various embodiments, the non-naturally occurring fusion protein has an amino acid sequence set forth in SEQ ID NO: 14-21, 24, or 25, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 98% or 99% sequence identity thereto. A schematic of the carrier, spacer, and payload of a non-naturally occurring fusion protein comprising an amino acid sequence set forth in SEQ ID NO: 15 is illustrated in FIG. **16**.

[0074] TABLE 1 shows exemplary amino acid sequences of biologically active molecules used in combination with the herein disclosed methods.

TABLE 1

Exemplary Amino Acid Sequences

Category	SEQ ID NO	Amino Acid Sequence
Cholix$^{1-634}$	SEQ ID NO: 1	VEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLAR SRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLDS VFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQAADILSLFC PDADKSCVASNNDQANINIESRSGRSYLPENRAVIT PQGVTNWTYQELEATHQALTREGYVFVGYHGTNH VAAQTIVNRIAPVPRGNNTENEEKWGGLYVATHAE VAHGYARIKEGTGEYGLPTRAERDARGVMLRVYIP RASLERFYRTNTPLENAEEHITQVIGHSLPLRNEAFT GPESAGGEDETVIGWDMAIHAVAIPSTIPGNAYEEL AIDEEAVAKEQSISTKPPYKERKDELK
Cholix$^{1-386}$	SEQ ID NO: 2	VEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLAR SRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLDS

TABLE 1-continued

Exemplary Amino Acid Sequences

Category	SEQ ID NO	Amino Acid Sequence
		VFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQA
Cholix$^{1-266}$	SEQ ID NO: 3	VEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKG
(V1L)-Cholix$^{1-386}$	SEQ ID NO: 4	LEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLAR SRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLDS VFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQA
(V1L)-Cholix$^{1-266}$	SEQ ID NO: 5	LEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKG
M + Cholix$^{1-386}$	SEQ ID NO: 6	MVEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLA RSRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLD SVFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQA
M + Cholix$^{1-266}$	SEQ ID NO: 7	MVEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKG
M + (V1L)-Cholix$^{1-386}$	SEQ ID NO: 8	MLEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLA RSRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLD SVFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQA
M + (V1L)-Cholix$^{1-266}$	SEQ ID NO: 9	MLEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI

TABLE 1-continued

Exemplary Amino Acid Sequences

Category	SEQ ID NO	Amino Acid Sequence
		YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKG
IL-22$^{1-179}$	SEQ ID NO: 10	MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAP ISSHCRLDKSNFQQPYITNRTFMLAKEASLADNNTD VRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFP QSDRFQPYMQEVVPFLARLSNRLSTCHIEGDDLHIQ RNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNA CI
IL-22$^{34-179}$	SEQ ID NO: 11	APISSHCRLDKSNFQQPYITNRTFMLAKEASLADNN TDVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVL FPQSDRFQPYMQEVVPFLARLSNRLSTCHIEGDDLH IQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRN ACI
M + IL-22$^{34-179}$	SEQ ID NO: 12	MAPISSHCRLDKSNFQQPYITNRTFMLAKEASLADN NTDVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEV LFPQSDRFQPYMQEVVPFLARLSNRLSTCHIEGDDL HIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLR NACI
$(G_4S)_3$	SEQ ID NO: 13	GGGGSGGGGSGGGGS
M + Cholix$^{1-386}$-$(G_4S)_3$-IL-22$^{34-179}$	SEQ ID NO: 14	MLEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLA RSRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLD SVFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQAGGGGSG GGGSGGGGSAPISSHCRLDKSNFQQPYITNRTFMLA KEASLADNNTDVRLIGEKLFHGVSMSERCYLMKQ VLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGE LDLLFMSLRNACI
M + Cholix$^{1-266}$-$(G_4S)_3$-IL-22$^{34-179}$	SEQ ID NO: 15	MVEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGGGGGSGGGGSGGGGSAPISSH CRLDKSNFQQPYITNRTFMLAKEASLADNNTDVRLI GEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDR FQPYMQEVVPFLARLSNRLSTCHIEGDDLHIQRNVQ KLKDTVKKLGESGEIKAIGELDLLFMSLRNACI
M + (V1L)-Cholix$^{1-386}$-$(G_4S)_3$-IL-22$^{34-179}$	SEQ ID NO: 16	MVEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLA RSRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLD SVFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQAGGGGSG GGGSGGGGSAPISSHCRLDKSNFQQPYITNRTFMLA KEASLADNNTDVRLIGEKLFHGVSMSERCYLMKQ VLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGE LDLLFMSLRNACI
M + (V1L)-Cholix$^{1-266}$-	SEQ ID NO: 17	MLEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR

TABLE 1-continued

Exemplary Amino Acid Sequences

Category	SEQ ID NO	Amino Acid Sequence
$(G_4S)_3$-IL-22^{34-179}		ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGGGGGSGGGGSGGGGSAPISSH CRLDKSNFQQPYITNRTFMLAKEASLADNNTDVRLI GEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDR FQPYMQEVVPFLARLSNRLSTCHIEGDDLHIQRNVQ KLKDTVKKLGESGEIKAIGELDLLFMSLRNACI
Cholix$^{1-386}$- $(G_4S)_3$-IL-22^{34-179}	SEQ ID NO: 18	VEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLAR SRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLDS VFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQAGGGGSG GGGSGGGGSAPISSHCRLDKSNFQQPYITNRTFMLA KEASLADNNTDVRLIGEKLFHGVSMSERCYLMKQ VLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGE LDLLFMSLRNACI
(V1L)-Cholix$^{1-386}$- $(G_4S)_3$-IL- 22^{34-179}	SEQ ID NO: 19	LEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLAR SRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLDS VFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQAGGGGSG GGGSGGGGSAPISSHCRLDKSNFQQPYITNRTFMLA KEASLADNNTDVRLIGEKLFHGVSMSERCYLMKQ VLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGE LDLLFMSLRNACI
Cholix$^{1-266}$- $(G_4S)_3$-IL-22^{34-179}	SEQ ID NO: 20	VEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGGGGGSGGGGSGGGGSAPISSHC RLDKSNFQQPYITNRTFMLAKEASLADNNTDVRLI GEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDR FQPYMQEVVPFLARLSNRLSTCHIEGDDLHIQRNVQ KLKDTVKKLGESGEIKAIGELDLLFMSLRNACI
(V1L)-Cholix$^{1-266}$- $(G_4S)_3$-IL- 22^{34-179}	SEQ ID NO: 21	LEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGGGGGSGGGGSGGGGSAPISSHC RLDKSNFQQPYITNRTFMLAKEASLADNNTDVRLI GEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDR FQPYMQEVVPFLARLSNRLSTCHIEGDDLHIQRNVQ KLKDTVKKLGESGEIKAIGELDLLFMSLRNACI
(V1L)-Cholix$^{1-634}$	SEQ ID NO: 22	LEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA

TABLE 1-continued

Exemplary Amino Acid Sequences

Category	SEQ ID NO	Amino Acid Sequence
		TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLAR SRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLDS VFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQAADILSLFC PDADKSCVASNNDQANINIESRSGRSYLPENRAVIT PQGVTNWTYQELEATHQALTREGYVFVGYHGTNH VAAQTIVNRIAPVPRGNNTENEEKWGGLYVATHAE VAHGYARIKEGTGEYGLPTRAERDARGVMLRVYIP RASLERFYRTNTPLENAEEHITQVIGHSLPLRNEAFT GPESAGGEDETVIGWDMAIHAVAIPSTIPGNAYEEL AIDEEAVAKEQSISTKPPYKERKDELK
M + (V1L)-Cholix$^{1-634}$	SEQ ID NO: 23	MLEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLA RSRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLD SVFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQAADILSLFC PDADKSCVASNNDQANINIESRSGRSYLPENRAVIT PQGVTNWTYQELEATHQALTREGYVFVGYHGTNH VAAQTIVNRIAPVPRGNNTENEEKWGGLYVATHAE VAHGYARIKEGTGEYGLPTRAERDARGVMLRVYIP RASLERFYRTNTPLENAEEHITQVIGHSLPLRNEAFT GPESAGGEDETVIGWDMAIHAVAIPSTIPGNAYEEL AIDEEAVAKEQSISTKPPYKERKDELK
Cholix$^{1-266}$-$(G_4S)_3$-IL-22$^{34-179}$	SSEQ ID NO: 24	VEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGGGGGSGGGGSGGGGSAPISSHC RLDKSNFQQPYITNRTFMLAKEASLADNNTDVRLI GEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDR FQPYMQEVVPFLARLSNRLSTCHIEGDDLHIQRNVQ KLKDTVKKLGESGEIKAIGELDLLFMSLRNACI
(V1L)-Cholix$^{1-266}$-$(G_4S)_3$-IL-22$^{34-179}$	SSEQ ID NO: 25	LEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVLD EGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVRA TRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEGEF AINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLYSI DLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSVS YKAAQKEGSRHKRWAHWHTGLALCWLVPMDAIY NYITQQNCTLGDNWFGGSYETVAGTPKVITVKQGI EQKPVEQRIHFSKGGGGGSGGGGSGGGGSAPISSHC RLDKSNFQQPYITNRTFMLAKEASLADNNTDVRLI GEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDR FQPYMQEVVPFLARLSNRLSTCHIEGDDLHIQRNVQ KLKDTVKKLGESGEIKAIGELDLLFMSLRNACI
M + Cholix$^{1-266}$-$(G_4S)_1$-IL-22$^{34-179}$	SEQ ID NO: 32	MVEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGGGGGSAPISSHCRLDKSNFQQP YITNRTFMLAKEASLADNNTDVRLIGEKLFHGVSM SERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVP

TABLE 1-continued

Exemplary Amino Acid Sequences

Category	SEQ ID NO	Amino Acid Sequence
		FLARLSNRLSTCHIEGDDLHIQRNVQKLKDTVKKL GESGEIKAIGELDLLFMSLRNACI
M + Cholix$^{1-266}$-$(G_4S)_5$-IL-22$^{34-179}$	SEQ ID NO: 33	MVEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGGGGGSGGGGSGGGGSGGGGS GGGGSAPISSHCRLDKSNFQQPYITNRTFMLAKEAS LADNNTDVRLIGEKLFHGVSMSERCYLMKQVLNFT LEEVLFPQSDRFQPYMQEVVPFLARLSNRLSTCHIE GDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLF MSLRNACI
M + IL-22$^{34-179}$-$(G_4S)_1$-Cholix$^{1-266}$	SEQ ID NO: 34	MAPISSHCRLDKSNFQQPYITNRTFMLAKEASLADN NTDVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEV LFPQSDRFQPYMQEVVPFLARLSNRLSTCHIEGDDL HIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLR NACIGGGGSVEEALNIFDECRSPCSLTPEPGKPIQSK LSIPSDVVLDEGVLYYSMTINDEQNDIKDEDKGESII TIGEFATVRATRHYVNQDAPFGVIHLDITTENGTKT YSYNRKEGEFAINWLVPIGEDSPASIKISVDELDQQR NIIEVPKLYSIDLDNQTLEQWKTQGNVSFSVTRPEH NIAISWPSVSYKAAQKEGSRHKRWAHWHTGLALC WLVPMDAIYNYITQQNCTLGDNWFGGSYETVAGT PKVITVKQGIEQKPVEQRIHFSKG
M + Cholix$^{1-386}$-$(G_4S)_1$-IL-22$^{34-179}$	SEQ ID NO: 35	MVEEALNIFDECRSPCSLTPEPGKPIQSKLSIPSDVVL DEGVLYYSMTINDEQNDIKDEDKGESIITIGEFATVR ATRHYVNQDAPFGVIHLDITTENGTKTYSYNRKEG EFAINWLVPIGEDSPASIKISVDELDQQRNIIEVPKLY SIDLDNQTLEQWKTQGNVSFSVTRPEHNIAISWPSV SYKAAQKEGSRHKRWAHWHTGLALCWLVPMDAI YNYITQQNCTLGDNWFGGSYETVAGTPKVITVKQG IEQKPVEQRIHFSKGNAMSALAAHRVCGVPLETLA RSRKPRDLTDDLSCAYQAQNIVSLFVATRILFSHLD SVFTLNLDEQEPEVAERLSDLRRINENNPGMVTQVL TVARQIYNDYVTHHPGLTPEQTSAGAQAGGGGSAP ISSHCRLDKSNFQQPYITNRTFMLAKEASLADNNTD VRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFP QSDRFQPYMQEVVPFLARLSNRLSTCHIEGDDLHIQ RNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNA CI
Cholix Consensus Sequence	SEQ ID NO: 26	X1-E-X3-X4-L-X6-I-F-D-E-C-R-S-P-C-X16-L-T-P-E- X21-G-K-X24-I-Q-S-K-L-X30-I-P-X33-D-V-V-L-D-E- G-V-L-Y-Y-S-M-T-I-N-D-E-Q-N-D-I-X56-D-E-X59-K- G-E-S-I-I-T-X67-G-E-F-A-T-X73-R-A-T-R-H-Y-V-X81- Q-D-A-P-F-G-V-I-X90-L-D-I-T-T-E-N-G-T-K-X101-Y- S-X104-N-R-K-X108-X109-E-F-X112-I-X114-W-L-V- X118-X119-G-E-D-S-P-A-S-I-K-I-S-X131-D-E-X134-D- Q-X137-R-N-I-I-E-V-P-K-L-Y-S-I-D-L-D-N-Q-T-L-E- Q-W-X160-X161-Q-G-N-V-X166-F-X168-V-T-R-P-E- X174-X175-I-A-I-S-W-P-S-V-S-Y-X186-A-A-X189-K- X191-G-X193-R-H-K-R-W-A-X200-W-X202-T-X204- X205-X206-X207-X208-X209-L-X211-X212-X213- X214-X215-X216-X217-X218-X219-X220-X221-X222- X223-X224-C-T-X227-G-X229-X230-W-X232-G-G- X235-Y-X237-T-V-A-G-X242-P-X244-X245-I-X247-V- K-Q-G-X252-E-Q-K-X256-V-E-Q-R-I-H-F-S-X265- X266-N-A-X269-X270-X271-L-A-A-H-R-V-C-G-V-P-L- E-T-L-A-R-X288-R-K-P-R-X293-L-X295-D-D-L-X299- C-X301-Y-X303-A-Q-X306-I-V-S-L-F-X312-A-T-R- X316-L-F-X319-H-X321-D-S-X324-F-T-L-N-L-X330- X331-Q-X333-P-X335-V-X337-E-R-L-X341-X342- X343-R-X345-I-N-E-X349-N-P-G-X353-V-X355-Q-V- L-T-X360-A-R-Q-I-Y-N-D-Y-V-T-X371-H-P-X374-L- X376-P-E-Q-T-S-A-X383-A-Q-A-A-D-I-L-S-L-X393- X394-P-D-X397-D-X399-X400-C-V-A-X404-X405- X406-D-Q-A-N-I-N-X413-E-S-R-S-G-R-S-Y-L-X423-E- N-R-A-V-I-T-X431-Q-G-V-T-N-W-T-Y-Q-E-L-X443-

TABLE 1-continued

Exemplary Amino Acid Sequences

Category	SEQ ID NO	Amino Acid Sequence
		X444-X445-H-Q-X448-L-T-X451-E-X453-Y-V-F-V-G- Y-H-G-T-N-H-X465-A-A-Q-X469-I-V-N-R-I-X475-P-V- P-R-G-X481-X482-T-E-X485-E-X487-X488-W-G-G- X492-Y-V-X495-T-X497-A-X499-X500-X501-X502- X503-Y-X505-R-X507-X508-X509-G-T-X512-X513- X514-X515-X516-X517-T-X519-X520-X521-X522- X523-X524-R-G-V-M-L-X530-V-Y-X533-X534-X535- A-S-L-E-R-F-Y-R-X544-N-X546-X547-L-E-X550- X551-X552-X553-X554-X555-X556-X557-V-I-G-H- X562-L-P-L-R-N-E-A-F-T-G-X573-X574-X575-X576- X577-G-X579-X580-E-T-X583-I-G-W-D-X588-A-I- X591-X592-V-A-I-P-S-T-I-P-G-N-X603-Y-X605-X606- L-X608-X609-X610-E-E-A-X614-A-X616-E-Q-S-I-S- X622-K-P-P-Y-K-E-X629-X630-D-E-L-K; wherein X1 is selected from the group consisting of V and L; X3 is selected from the group consisting of E and D; X4 is selected from the group consisting of A and E; X6 is selected from the group consisting of N and K; X16 is selected from the group consisting of S and L; X21 is selected from the group consisting of P and L; X24 is selected from the group consisting of P and Q; X30 is selected from the group consisting of S and F; X33 is selected from the group consisting of S and G; X56 is selected from the group consisting of K and M; X59 is selected from the group consisting of D and G; X67 is selected from the group consisting of I and F; X73 is selected from the group consisting of V and I; X81 is selected from the group consisting of N and S; X90 is selected from the group consisting of H and N; X101 is selected from the group consisting of T and M; X104 is selected from the group consisting of Y and F; X108 is selected from the group consisting of E and D; X109 is selected from the group consisting of G and S; X112 is selected from the group consisting of A and T; X114 is selected from the group consisting of N and H; X118 is selected from the group consisting of P and I; X119 is selected from the group consisting of I and P; X131 is selected from the group consisting of V and I; X134 is selected from the group consisting L and I; X137 is selected from the group consisting Q and K; X160 is selected from the group consisting K and E; X161 is selected from the group consisting T and N; X166 is selected from the group consisting S and F; X168 is selected from the group consisting S and A; X174 is selected from the group consisting H and Q; X175 is selected from the group consisting N, S, SIAKQS, and SIAKQSIAKQS; X186 is selected from the group consisting of K and N; X189 is selected from the group consisting of Q, E, and H; X191 is selected from the group consisting of E, N, and D; X193 is selected from the group consisting of S and A; X200 is selected from the group consisting of H and N; X202 is selected from the group consisting of H, L, F, and R; X204 is selected from the group consisting of G and T; X205 is selected from the group consisting of L and S; X206 is selected from the group consisting of A and P; X207 is selected from the group consisting of L, E, and K; X208 is selected from the group consisting of C and V; X209 is selected from the group consisting of W, V, and T; X211 is selected from the group consisting of V and no amino acid; X212 is selected from the group consisting of P and no amino acid; X213 is selected from the group consisting of M, I, L, and no amino acid; X214 is selected from the group consisting of D and no amino acid; X215 is selected from the group consisting of A and no amino acid; X216 is selected from the group consisting of I and no amino acid; X217 is selected from the group consisting of Y and C; X218 is selected from the group consisting of N and F; X219 is selected from the group consisting of Y and F; X220 is selected from the group consisting of I and E; X221 is selected from the group consisting of T and D; X222 is selected from the group consisting of Q and P;

TABLE 1-continued

Exemplary Amino Acid Sequences

Category	SEQ ID NO	Amino Acid Sequence
		X223 is selected from the group consisting of Q, E, and A; X224 is selected from the group consisting of N, L, and Q; X227 is selected from the group consisting of L and Y; X229 is selected from the group consisting of D and E; X230 is selected from the group consisting of N and D; X232 is selected from the group consisting of F, H, and Y; X235 is selected from the group consisting of S and A; X237 is selected from the group consisting of E and K; X242 is selected from the group consisting of T and I; X244 is selected from the group consisting of K, E, and G; X245 is selected from the group consisting of V and A; X247 is selected from the group consisting of T and M; X252 is selected from the group consisting of I and M; X256 is selected from the group consisting of P, T, and A; X265 is selected from the group consisting of K, Q, and N; X266 is selected from the group consisting of G and K; X269 is selected from the group consisting of M and I; X270 is selected from the group consisting of S and E; X271 is selected from the group consisting of A and T; X288 is selected from the group consisting of S and G; X293 is selected from the group consisting of D and Y; X295 is selected from the group consisting of T, P, and Q; X299 is selected from the group consisting of S and Q; X301 is selected from the group consisting of A and V; X303 is selected from the group consisting of Q and N; X306 is selected from the group consisting of N and Q; X312 is selected from the group consisting of V and L; X316 is selected from the group consisting of I and M; X319 is selected from the group consisting of S and T; X321 is selected from the group consisting of L and I; X324 is selected from the group consisting of V and I; X330 is selected from the group consisting of D, E, and H; X331 is selected from the group consisting of E and G; X333 is selected from the group consisting of E and A; X335 is selected from the group consisting of E and A; X337 is selected from the group consisting of A and T; X341 is selected from the group consisting of S, D, and T; X342 is selected from the group consisting of D and A; X343 is selected from the group consisting of L and I; X345 is selected from the group consisting of R and Q; X349 is selected from the group consisting of N and D; X353 is selected from the group consisting of M and V; X355 is selected from the group consisting of T and I; X360 is selected from the group consisting of V and I; X371 is selected from the group consisting of H and E; X374 is selected from the group consisting of G and L; X376 is selected from the group consisting of T and I; X383 is selected from the group consisting of G and S; X393 is selected from the group consisting of F and L; X394 is selected from the group consisting of C and Y; X397 is selected from the group consisting of A and T; X399 is selected from the group consisting of K, E, and G; X400 is selected from the group consisting of S, P, and H; X404 is selected from the group consisting of S and L; X405 is selected from the group consisting of N and D; X406 is selected from the group consisting of N and S; X413 is selected from the group consisting of I and V; X423 is selected from the group consisting of P and L; X431 is selected from the group consisting of P and Q; X443 is selected from the group consisting of E and D; X444 is selected from the group consisting of A and T; X445 is selected from the group consisting of T and K; X448 is selected from the group consisting of A and T; X451 is selected from the group consisting of R and Q; X453 is selected from the group consisting of G and D; X465 is selected from the group consisting of V and A; X469 is selected from the group consisting of T, S, and N; X475 is selected from the group consisting of A, S, and T; X481 is selected from the group consisting of N and S; X482 is selected from the group consisting of N and D; X485 is selected from the group consisting of N, S, and K;

TABLE 1-continued

Exemplary Amino Acid Sequences		
Category	SEQ ID NO	Amino Acid Sequence
		X487 is selected from the group consisting of E, R, and K; X488 is selected from the group consisting of K, A, and E; X492 is selected from the group consisting of L and V; X495 is selected from the group consisting of A and S; X497 is selected from the group consisting of H and D; X499 is selected from the group consisting of E and S; X500 is selected from the group consisting of V and L; X501 is selected from the group consisting of A and N; X502 is selected from the group consisting of H and Y; X503 is selected from the group consisting of G and R; X505 is selected from the group consisting of A and T; X507 is selected from the group consisting of I and L; X508 is selected from the group consisting of K and Q; X509 is selected from the group consisting of E and K; X512 is selected from the group consisting of G and A; X513 is selected from the group consisting of E, D, and N; X514 is selected from the group consisting of Y, G, A, and N; X515 is selected from the group consisting of G and E; X516 is selected from the group consisting of L and G; X517 is selected from the group consisting of P and L; X519 is selected from the group consisting of R, P, and T; X520 is selected from the group consisting of A and E; X521 is selected from the group consisting of E and K; X522 is selected from the group consisting of R, Q, and K; X523 is selected from the group consisting of D, K, and E; X524 is selected from the group consisting of A, T, and S; X530 is selected from the group consisting of R and K; X533 is selected from the group consisting of I and L; X534 is selected from the group consisting of P and H; X535 is selected from the group consisting of R and Q; X544 is selected from the group consisting of T and I; X546 is selected from the group consisting of T, A, and I; X547 is selected from the group consisting of P and D; X550 is selected from the group consisting of N and K; X551 is selected from the group consisting of A and E; X552 is selected from the group consisting of E, R, and D; X553 is selected from the group consisting of E, N, and R; X554 is selected from the group consisting of H and L; X555 is selected from the group consisting of I and V; X556 is selected from the group consisting of T and E; X557 is selected from the group consisting of Q, R, H, and D; X562 is selected from the group consisting of S and P; X573 is selected from the group consisting of P and T; X574 is selected from the group consisting of E and D; X575 is selected from the group consisting of S, A, and R; X576 is selected from the group consisting of A, E, and V; X577 is selected from the group consisting of G, E, and D; X579 is selected from the group consisting of E and S; X580 is selected from the group consisting of D and N; X583 is selected from the group consisting of V and A; X588 is selected from the group consisting of M and I; X591 is selected from the group consisting of H and Y; X592 is selected from the group consisting of A and G; X603 is selected from the group consisting of A and S; X605 is selected from the group consisting of E and A; X606 is selected from the group consisting of E, A, Q, G, V, and R; X608 is selected from the group consisting of A, P, and T; X609 is selected from the group consisting of I, T, and P; X610 is selected from the group consisting of D and A; X614 is selected from the group consisting of V and VVKEAI; X616 is selected from the group consisting of K and E; X622 is selected from the group consisting of T, A, and P; and X629 is selected from the group consisting of R, Q, and H; and X630 is selected from the group consisting of K and no amino acid.

[0075] In some embodiments, the non-naturally occurring fusion protein comprises or consists of: a carrier selected from the group consisting of any one of the sequences set forth in SEQ ID NO: 1-9, 22-23, and 26 and a payload selected from the group consisting of any one of the sequences set forth in SEQ ID NO: 10-12, and wherein the carrier and the payload are optionally coupled by a spacer selected from the group consisting of any one of the sequences set forth in SEQ ID NO: 13 and 27-29.

[0076] Further provided herein are pharmaceutical compositions comprising a delivery construct and one or more pharmaceutically acceptable carriers. In some cases, the delivery construct consists of the amino acid sequence set forth in SEQ ID NO: 15. In some cases, the delivery construct consists of the amino acid sequence set forth in SEQ ID NO: 17. Such pharmaceutical compositions can be formulated for administration to a subject. In some instances, a pharmaceutical composition is formulated for oral administration to a subject.

[0077] A pharmaceutical composition comprising a delivery construct can be administered to a subject (e.g., a human) in need thereof to treat a disease. Diseases that can be treated using the delivery constructs of this disclosure include autoimmune diseases and inflammatory diseases. In some instances, the disease is an epithelial cell injury or damage of epithelial cell membranes (e.g., in the GI tract). In some instances, the disease is hepatitis, obesity, fatty liver disease, liver inflammation, or pancreatitis, Crohn's disease (e.g., fistulizing Crohn's disease), ulcerative colitis (e.g., mild-to-moderate or moderate-to-severe), pouchitis, proctitis, multiple sclerosis, systemic lupus erythematosus, graft versus host disease, rheumatoid arthritis, or psoriasis.

Purification Methods and Compositions

[0078] The present disclosure contemplates methods for obtaining a purified non-naturally occurring fusion protein. The non-naturally occurring fusion protein can comprise IL-22 and a carrier, and optionally a spacer coupling the IL-22 to the carrier. The non-naturally occurring fusion protein can be a protein comprising, consisting essentially of, or consisting of an amino acid sequence having at least 80%, 85%. 90%, 95%, 98%, or 99% sequence identity, or having 100% sequence identity, to an amino acid sequence set forth in SEQ ID NOS: 14-21, 24, or 25. Advantages of the methods and compositions described herein include maintaining high purity and biological activity of the purified polypeptides or proteins.

[0079] Specific advantages of the herein disclosed methods include: a) high biological activity of the non-naturally occurring fusion proteins due to correct folding using specifically developed folding buffers; b) high chemical purity of the purified non-naturally occurring fusion proteins paired with low toxicity; c) high recovery yield of the material subjected to purification; d) reproducibility of the results enables reliable production and supply; e) manufacturability and scalability to multi-gram and multi-kilogram scale for clinical and commercial applications; f) sustainable use of materials and resources provide a cost- and logistically effective purification method for clinically relevant molecules. Additionally, improved manufacturing methods can increase the speed at which these purified proteins can be developed for therapeutic use (FIG. **22**).

[0080] An exemplary process for purification of the non-naturally occurring fusion proteins described herein is shown in FIG. **17**, which further illustrates exemplary purity and recovery amounts for the non-naturally occurring fusion protein at each stage in the purification process.

[0081] As used herein, "purity," indicates a level of the desired protein (e.g. non-naturally occurring fusion protein) or desired form of the protein (e.g. a monomer of the non-naturally occurring fusion protein, a corrected folded non-naturally occurring fusion protein, or a combination thereof) in a composition which can also comprise non-desired proteins or non-desired forms of the protein (e.g. an aggregate of the non-naturally occurring fusion protein, an incorrectly folded protein, or a combination thereof). The level can be represented by a percentage (%) of the desired protein or desired form of the protein. For example, the purity can be greater than 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 99%, or can be 100%. The term "purified" can be used to indicate a composition which has experienced an increase in purity. An increase in purity can be due to purification procedures described herein, such as anion exchange chromatography, hydroxyapatite chromatography, cation exchange chromatography, or a combination thereof.

[0082] In some cases, the methods and compositions of the present disclosure comprise the use of at least two chromatography columns that are performed in tandem, follow by a concentration/buffer exchange step (ultrafiltration/diafiltration (UF/DF) to concentrate and buffer exchange the solution). The chromatography columns can be low pressure chromatography systems, such as the AKTA Avant 150 column or an AKTA Pilot column from General Electric (GE). In some cases, at least one column contains an anion exchange resin (e.g. NH_2-750F resin) and at least one column contains a hydroxyapatite resin (e.g., CaPure® resin). In some case, at least one column containing the anion exchange resin is used for the capture step, whereas the at least one column containing the hydroxyapatite resin is used for the polishing step. In some cases, instead (or in addition to) the hydroxyapatite column, a cation exchange column may be used. For instance, in some embodiments, the cation exchange column with a sulfate-functionalized polymethacrylate resin, such as TOYOPEARL Sulfate-650F. The purity of the solubilized protein can increase from 44% to 47% following UF/DF to from 93% to 97% following anion exchange chromatography to at least 99% following hydroxyapatite chromatography. The recovery of the solubilized protein can increase from 70% to 73% following the anion exchange chromatography to at least 97% following the hydroxyapatite chromatography.

[0083] Isolation of Non-Naturally Occurring Fusion Proteins from Inclusion Bodies

[0084] In some embodiments, the non-naturally occurring fusion proteins to be purified are isolated from inclusion bodies (IBs). IBs as described herein may be nuclear or cytoplasmic aggregates of stable substances, such as proteins and polypeptides. In some embodiments, double-washed inclusion bodies (DWIB) comprising the non-naturally occurring fusion proteins (e.g., SEQ ID NOs: 14-21) from fermentation is resuspended in a specific buffer system (see EXAMPLE 6).

[0085] In some cases, the concentration of DWIB is from about 0.5 g DWIB/100 mL buffer to about 5 g DWIB/10 mL buffer. In some cases, the concentration of DWIB is from about 2 g DWIB/50 mL buffer to about 5 g DWIB/50 mL buffer. In some cases, the concentration of DWIB is from about 1 g DWIB/20 mL buffer to about 1 g DWIB/10 mL

buffer. In some cases, the concentration of DWIB is at least 1 g DWIB/100 mL buffer. In some cases, the concentration of DWIB is at least 1 g DWIB/10 mL buffer.

[0086] The buffer system for resuspension may comprise a variety of buffer agents and ingredients. In some cases, the buffer system for resuspension comprises Guanidine/HCl (Gu-HCl) at a concentration of at least 1M, at least 2 M, at least 4M, at least 6 M, at least 8M. In some cases, the concentration of Gu-HCl is from about 4 M to about 8 M. In some cases, the buffer system for resuspension comprises Tris buffer, wherein the concentration of Tris is at least 5 mM, at least 10 mM, at least 20 mM, at least 50 mM, or at least 100 mM. In some cases, the Tris buffer has a concentration from about 40 mM to about 60 mM. The Tris buffer can be Tris-HCl. The Tris-HCl can have a pH from 8.0 to 8.5. The Tris-HCl can have a pH of 8.2. In some cases, the buffer system for resuspension comprises dithiothreitol (DTT). The DTT can have a concentration of at least 7 mM, 8 mM, 9 mM, 10 mM, 20 mM, 30 mM, 40 mM, 50 mM, 60 mM, or 70 mM. The DDT can have a concentration from 7 mM to 11 mM, from 8 mM to 10 mM, or from 9 mM to 10 mM. In some cases, the buffer system for resuspension comprises ethylenediaminetetraacetic acid (EDTA). The EDTA can have a concentration of at least 1 mM, 1.5 mM, 2 mM, or 2.5 mM. The EDTA can have a concentration from 1 mM to 2.5 mM or from 1.9 mM to 2.1 mM. In some embodiments, the buffer system for resuspension comprises, consists essentially of, or consists of a Tris buffer, Gu-HCl, and DTT. The buffer system used for resuspension may have a pH from about 6 to about 10. In some cases, the buffer system has a pH from about 7.5 to about 8.5. In some cases, the pH of the buffer system is at least 7. In some cases, the pH of the buffer system is at least 8. In some cases, the pH of the buffer system is at least 8.5.

[0087] Reduction of Solubilized Non-Naturally Occurring Fusion Proteins

[0088] A reducing agent may be added to the resuspension solution. The reducing agent may be dithiothreitol (DTT). The amount of reducing agent used may depend on the volume of the resuspension solution and the fusion protein present in that solution. In some cases, the amount of reducing agent used is from about 0.025 mM to about 50 mM, from 0.5 mM to 30 mM, from 1 mM to 10 mM. In some cases, the amount of reducing agent used is at least (or between) 2 mM, 3 mM, 4 mM, 5 mM, 6 mM, 7 mM, 8 mM, 9 mM, or 10 mM.

[0089] The solubilized non-naturally occurring fusion protein (e.g., SEQ ID NOs: 14-21) solution may be incubated at various temperatures ranging from about 2° C. to about 40° C. In some cases, the protein solution may be incubated at room temperature (RT). Incubation may be performed while the solution is stirring. The solution may be stirred on a magnetic stir-plate and then centrifuged. Centrifugation may be performed at about 1,000×g to about 20,000×g. In some cases, centrifugation is performed at 15,970×g for 90 minutes at 4° C. The incubation time may vary from about 20 minutes to about 180 minutes, depending on the concentration of ingredients, the concentration of the fusion protein (e.g., SEQ ID NOs: 14-21).

[0090] Protein concentration in various solutions throughout the methods and procedures as described herein may be performed using Bradford assay. Concentration of the solubilized non-naturally occurring fusion protein may be from about 1 mg/mL to about 50 mg/mL. In some cases, the final concentration of the solubilized non-naturally occurring fusion protein may be from about 3 mg/mL to about 30 mg/mL. In some cases, the final concentration of the solubilized non-naturally occurring fusion protein may be from about 5 mg/mL to about 20 mg/mL. In some cases, the final concentration of the solubilized non-naturally occurring fusion protein may be at least 15 mg/mL.

[0091] Refolding of Non-Naturally Occurring Fusion Proteins

[0092] The non-naturally occurring fusion proteins of the present disclosure may exert their biological activity due to their specific three-dimensional structure or folding. Therefore, refolding of these polypeptides can be important in developing these polypeptides for pharmaceutical applications. Desirable refolding of a fusion protein comprising at least a carrier and an IL-22 can comprise refolding of the carrier or the IL-22 into a tertiary structure similar to a tertiary structure of a homologous naturally occurring carrier or IL-22 sequence or a tertiary structure which results in maintenance of the desired activity of the carrier (e.g. transcytosis of the fusion protein) and IL-22. The methods described herein can comprise refolding the solubilized non-naturally occurring fusion protein to produce refolded non-naturally occurring fusion protein. The refolding can occur prior to performing the anion exchange chromatography.

[0093] The solubilized protein (e.g., SEQ ID NOs: 14-21) may be added to a refolding solution, also referred to as a refold buffer solution. The amount of solubilized protein used may be from 0.1 mg/mL to 1 mg/mL, 1 mg/mL to 100 mg/mL, from 1 mg/mL to 50 mg/mL, from 1 mg/mL to 10 mg/mL, from 5 mg/mL to 20 mg/mL, or about 15 mg/mL. The amount of solubilized protein can be at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mg/mL. The amount of solubilized protein can be no more than 0.1, 0.5, 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100 mg/mL. The refolding solution can have a pH of between 7.5 and 8.5, or about 7.0.

[0094] The refold buffer solution can comprise an amino acid, a polyol, a salt, a sugar, a redox reagent, a chaotrope, or a combination thereof. The amino acid can be proline, glycine, arginine, or alanine. The polyol can be glycerol. The salt can be sodium chloride (NaCl), potassium chloride (KCl), or magnesium chloride ($MgCl_2$). The sugar can be glucose or sucrose. The redox reagent can be cysteine, cystamine, glutathione, dithiothreitol (DTT), or copper sulfate ($CuSO_4$). In some embodiments, the refold buffer comprises Tris-base, Arginine-HCl, urea, EDTA, glycerol, L-Cysteine, Cystamine-2HCl, DTT, Gu-HCl, or a combination thereof. In some embodiments, the refold buffer comprises, consists essentially of, or consists of Tris pH 8.5, arginine glycerol, cysteine, and cystamine. Concentrations of Tris buffer may range from about 20 mM to about 200 mM with pH ranging from about 6 to about 9. Concentrations of Arginine-HCl may range from about 0.1 M to about 1.5 M, from about 0.75 M to 1.25M, or about 1.0 M. Concentrations of urea may range from about 0.5 M to 1.5 M. Concentrations of EDTA may range from 1 mM to 3 mM. Concentrations of glycerol may range from about 2% v/v to about 20% v/v, from about 8% v/v to about 12% v/v, or about 10% v/v. Concentrations of cysteine can be from 1 mM to 5 mM, from 2 mM to 4 mM, or about 3 mM. The cysteine can be L-cysteine. Concentrations of cystamine may range from about 0.5 mM to about 5 mM, from about 1 mM to about 4 mM, or about 3 mM. The cystamine can

be cystamine-2HCl. Concentrations of DTT may range from about 0.1 mM to about 1 mM. Concentrations of Gu-HCl may range from about 50 mM to about 500 mM. The refold buffer can comprise, consist essentially of, or consist of Tris, L-arginine, urea, EDTA, cysteine, and cystamine. The refold buffer can comprise, consist essentially of, or consist of 100 mM Tris pH 8.5, 1M arginine, 10% (v/v) glycerol, 3 mM L-Cysteine, and 1 mM Cystamine-2HCl. The solubilized non-naturally occurring fusion protein can be added to the refold buffer to produce a refold mixture. The concentration of the non-naturally occurring fusion protein, after addition to the refold mixture, can be from about 0.1 mg/mL to 1.0 mg/mL, 0.5/mL mg to 1.5 mg/mL, from 0.75 mg/mL to 1.25 mg/mL, or about 1 mg/mL. The concentration of the non-naturally occurring fusion protein in the refold mixture can be less than 0.5 mg/mL, 0.6 mg/mL, 0.7 mg/mL, 0.8 mg/mL, 0.9 mg/mL or 1.0 mg/mL.

[0095] In various embodiments, a volume of about 100 mL to about 10,000 L refold buffer is used. In some cases, the volume is from about 50 mL to about 1500 L. In some cases, the volume is from about 10 L to about 300 L. In some cases, the volume is at least 200 L.

[0096] The refold solution may have a certain temperature. For example, the refold solution may be pre-chilled to about 2-12° C. In some cases, the refold solution is pre-chilled to at least about 3° C. This refolding process may be optimized depending on the protein used and the application (FIG. **18**).

[0097] In some embodiments of the present disclosure, the refolding mixture is incubated for a certain period of time. In some cases, the process may be about one hour, two hours, four hours, five hours, or 20 hours. In some cases, the refolding process takes about 15 to about 25 hours. In some cases, a peristaltic pump set at a certain flow rate depending on the constitution of the solution (e.g., concentration or volume) to deliver the solubilized material into the refold solution. In some cases, a peristaltic pump set at a flow rate of 60-80 ml/min. The optimization process utilized a design of experiments (DOE) of 15 different matrices, 12 variables and 200 refold reactions. Decisions were made by a process of elimination.

[0098] In various embodiments, quantitative analysis may be performed to determine the amount or percentage of correctly folded protein in the solution comprising the refolded non-naturally occurring fusion proteins. In some instances, SEC-HPLC is performed to determine the amount of the properly folded non-naturally occurring fusion protein present in the refold samples.

[0099] The purity of the refolded non-naturally occurring fusion protein can be from 45% to 65%, from 40% to 60%, or from 45% to 55%. The purity of the refolded non-naturally occurring fusion protein can be at least 40%, 45%, 50%, or 55%.

[0100] Tangential Flow Filtration (TFF) of Solubilized IL-22 Delivery Constructs

[0101] In various embodiments, protein refolding is performed in combination with or prior to tangential flow filtration (TFF) systems (e.g., Millipore) and ultrafiltration/diafiltration (UF/DF) systems with certain molecular weight cut-offs (MWCOs).

[0102] Protein refolding mixture may be subsequently processed by TFF to concentrate and buffer exchange the solution. In some cases, the process is performed by ultrafiltration/diafiltration (UF/DF). During ultrafiltration, the solution can be concentrated from 8-fold to 12-fold, or about 10-fold. Ultrafiltration can be followed by a 5-fold buffer exchange during diafiltration with a diafiltration buffer. The UF/DF can comprise the use of 10-20 kDa MWCO Millipore Ultracell Pellican3 filters, with 1-2 m^2 TFF flat sheet cassettes. The specific parameters (i.e. TMP) may be varied depending on the molecular characteristics of the protein. The diafiltration buffer can comprise from 10-25 mM Tris-base at pH from about 7 to about 8.5, and from 50-200 mM NaCl. The specific parameters may be varied depending on the molecular characteristics of the protein. The final volume at the end of the UF/DF process may be from about 5 L to about 100 L. In some cases, the final volume ranges from about 10 to about 50 L. In some cases, the final volume ranges from about 80 mL to about 10 L.

[0103] Subsequent filtration may be carried out using a flow filtration system. In some cases, the flow filtration system is an AkroPac filter system and Cole-Parmer peristaltic pump and tubing. Bradford assay may be performed to determine the protein concentration, yield, and step recovery.

[0104] In various embodiments, quantitative analysis may be performed to determine the amount or percentage of correctly folded protein in the sample solutions. In some instances, SEC-HPLC is performed to determine the amount of the properly folded non-naturally occurring fusion protein present in the refold and UF/DF samples.

[0105] The purity of a solution comprising the refolded non-naturally occurring fusion protein following UF/DF can be from 35% to 55%, from 40% to 50%, from 43% to 47%, or about 45%. The purity of a solution comprising the refolded non-naturally occurring fusion protein following UF/DF can be at least 35%, 40%, or 45%. The recovery of the solubilized, refolded non-naturally occurring fusion protein following UF/DF can be from 87% to 97%, from 90% to 94%, from 92% to 93%, or about 92.5%. (Other concentration and buffer exchange systems can also be used.)

[0106] Purification and Recovery of Non-Naturally Occurring Fusion Proteins

[0107] As previously described, the use of two chromatography methods in tandem, such as an anion exchange chromatography and a hydroxyapatite chromatography (or alternatively cation exchange chromatography), can result in increased purity and recovery of the non-naturally occurring fusion protein, such as the IL-22 delivery constructs described herein, from solution. In various embodiments of the present disclosure, the chromatography specifics (e.g., column type, resin, flow rate, buffer systems, gradient, and conductivity) are determined and optimized depending on various parameters, e.g., the protein to be purified. A variety of purification columns may be used for protein purification as described herein.

[0108] Anion Exchange Chromatography

[0109] The methods described herein can comprise performing anion exchange chromatography (AEX) on a mixture comprising the non-naturally occurring fusion protein. The mixture can be a solution comprising the refolded non-naturally fusion protein. Performing anion exchange chromatography can produce a first fraction comprising the non-naturally occurring fusion protein.

[0110] A variety of resins may be used in combination with the methods and compositions of the present disclosure. In some cases, the resin comprises amine-functionalized polymethylacrylate beads. In some cases, anion

exchange resin NH_2-750F is used for protein purification. A column with a bed-height of at least 15 cm, 20 cm, 25 cm, or 30 cm can be filled with the resin. A column with a bed-height from 10 to 50 cm may be filled with resin in a way to ensure a column volume that may facilitate an appropriate dynamic binding capacity (e.g., >20 g/L) to enable production of the desired quantity of protein. In some cases, the dynamic binding capacity of the column is from 5 g/L to 100 g/L, from 20 g/L to about 50 g/L, from 15 g/L to 30 g/L, or from 20 g/L to 25 g/L.

[0111] Buffer systems used for elution (e.g., gradient elution) in anion exchange chromatography may comprise one, two, three, or four different buffer solutions (e.g., Buffers A through D). In some embodiments, two buffers are used in the anion exchange chromatography and are referred to herein as Buffer A and Buffer B. In some cases, a buffer solution comprises Tris (e.g., 10-50 mM, pH 7-9), and/or NaCl (e.g., 0.1-5 M). In some cases, the buffer solution further comprises glycerol. In some embodiments, the buffer solution(s) used in anion exchange chromatography, such as for example Buffer A or Buffer B, comprises, consists essentially of, or consists of Tris pH 7.5 and NaCl, and optionally, glycerol.

[0112] Buffer A can comprise from 10 mM to 30 mM, from 15 mM to 25 mM, from 19 mM to 21 mM, or about 20 mM Tris pH 7.5. Buffer A can comprise from 0.25 M to 0.95 M, from 0.35 M to 0.75M, from 0.45 M to 0.55 M, about 0.5 M, from 0.65 M to 0.75 M, or about 0.7 M NaCl. Buffer B can comprise from 10 mM to 30 mM, from 15 mM to 25 mM, from 19 mM to 21 mM, or about 20 mM Tris pH 7.5. Buffer B can comprise from 1 M to 3 M, from 1.5 M to 2.5 M, from 1.9 M to 2.1 M, or about 2 M NaCl. Buffer A can further comprise from 8% to 12%, from 9% to 11%, or about 10% (v/v) glycerol. Buffer B can further comprise from 8% to 12%, from 9% to 11%, or about 10% (v/v) glycerol.

[0113] For anion exchange chromatography of SEQ ID NO: 15, Buffer A can comprise 20 mM Tris pH 7.5 and 0.5 M NaCl and Buffer B can comprise 20 mM Tris pH 7.5 and 2.0 M NaCl (TABLE 4). For anion exchange chromatography of SEQ ID NO: 17, Buffer A can comprise 20 mM Tris pH 7.5 and 0.5 M NaCl and Buffer B can comprise 20 mM Tris pH 7.5 and 2.0 M NaCl. For anion exchange chromatography of SEQ ID NO: 14, Buffer A can comprise 20 mM Tris pH 7.5, 0.7 M NaCl, and 10% (v/v) glycerol, and Buffer B can comprise 20 mM Tris pH 7.5, 2.0 M NaCl, and 10% (v/v) glycerol (TABLE 5).

[0114] In some embodiments, a certain volume of a salt solution is added to the protein solution. In some cases, NaCl solution with a concentration raging from about 0.1 to about 5 M is added to the protein solution.

[0115] Protein solutions may be loaded onto the column such that a specific flow-rate and a specific column residence time may be observed to ensure appropriate interaction of the protein and the solid phase (e.g., resin). In some cases, the flow rate is controlled such that a column residence time is from about 30 seconds to about 6 minutes, or about 5 minutes. The column residence time can be at least 30 seconds, or 1, 2, 3, 4, or 5 minutes. The column residence time can be no more than 3, 4, 5, 6, 7, 8, 9, or 10 minutes. The flow-through may be collected and analyzed to determine if there is unbound protein, which is considered as a loss. For protein analysis and purification, the absorbance at about 280 nm is determined for protein concentration and SEC-HPLC for protein purity.

[0116] In some embodiments, a percentage (e.g., 0-100%) or flow rate (e.g., 1-10 mL/min) of a first buffer is combined with a second buffer to achieve a specific final flow rate (e.g., 1-10 mL/min) of the column purification system. For example, a linear gradient of 20 column volumes (CV) from 0.0-62.5% B (100-37.5% A), followed by a step gradient to 100% B (0% A) for additional 5 CV is performed for protein purification. In some cases, a gradient of 0-75% B (100-25% A) 40 CV, Step 100% B, 10 CV may be performed as an elution gradient indicating 91-93% pure protein and a recovery of >90%.

[0117] The purity of the non-naturally occurring fusion protein in the first fraction can be from 90% to 99% or from 91% to 95%. The purity of the solubilized non-naturally occurring fusion protein in the first fraction can be at least 90%, 91%, 92%, 93%, 94%, or 95%. The recovery of the solubilized non-naturally occurring fusion protein in the first fraction can be from 61% to 81%, from 66% to 76%, or from 71% to 72%.

[0118] Cation Exchange Chromatography or Hydroxyapatite Chromatography

[0119] The methods described herein can further comprise subjecting the first fraction obtained following the anion exchange chromatography to a cation exchange resin to obtain a second fraction comprising the non-naturally occurring fusion protein. The cation exchange resin can be a sulfate-functionalized methacrylate resin. The cation exchange resin can be a TOYOPEARL® Sulfate-650F resin.

[0120] The methods described herein can additionally or alternatively comprise subjecting the first fraction obtained following the anion exchange chromatography to a hydroxyapatite resin to obtain a second fraction comprising the non-naturally occurring fusion protein. The hydroxyapatite resin can be a cation exchange resin further comprising a calcium affinity. The hydroxyapaptite resin can comprise calcium phosphate. The hydroxyapaptite resin can comprise a chemical formula of: $Ca_{10}(PO_4)_6(OH)_2$. The hydroxyapaptite resin can comprise a particle size from 30 μm to 50 μm, from 35 μm to 45 μm, or about 39 μm. The hydroxyapaptite resin can be the CaPure® resin.

[0121] Advantages of the CaPure® resin can include a) salt tolerance, allowing protein adsorption at high conductivity in aqueous solutions; b) no preparation of proteins (e.g., fusion proteins such as SEQ ID NO: 14-21) required prior to load; c) results in a high recovery (>90%); d) high binding capacity (>20 mg/mL resin); e) increases purity by removal of low molecular weight (LMW) impurities; and f) ensures endotoxin clearance (<1.0 EU/mg).

[0122] The cation exchange resin or hydroxyapatite resin can be used to pack a chromatograph column. The column can comprise a bed height from 10 cm to 50 cm, or about 20 cm. The column can comprise a bed height of at least 10 cm, 15 cm, 20 cm, 25 cm, or 30 cm. The column can comprise a bed height of no more than 20 cm, 30 cm, 40 cm, or 50 cm. In some cases, the hydroxyapatite mixed-mode resin CaPure® is used to pack a column with a bed-height of 10-50 cm, ensuring a column volume that facilitates a dynamic binding capacity of maximum 10-100 g/L to enable production of the desired quantity of protein.

[0123] Buffer systems used for elution (e.g., gradient elution) in hydroxyapatite chromatography may comprise one, two, three, or four different buffer solutions (e.g., Buffers A through D). In some embodiments, two buffers are used in the hydroxyapatite chromatography and are referred

to herein as Buffer A and Buffer B. In some cases, a buffer solution comprises Tris (e.g., 10-50 mM, pH 7-9), and/or NaCl (e.g., 0.1-5 M). In some cases, the buffer solution further comprises glycerol. In some embodiments, a first buffer solution used in hydroxyapatite chromatography, such as for example Buffer A, comprises, consists essentially of, or consists of Tris pH 7.5, NaCl, and $CaCl_2$. In some embodiments, a second buffer solution used in hydroxyapatite chromatography, such as for example Buffer B, comprises, consists essentially of, or consists of sodium phosphate pH 7.0, NaCl, and $CaCl_2$.

[0124] Buffer A can comprise from 10 mM to 30 mM, from 15 mM to 25 mM, from 19 mM to 21 mM, or about 20 mM Tris pH 7.5. Buffer A can comprise from 80 mM to 120 mM, from 90 mM to 110 mM, or about 100 mM NaCl. Buffer A can comprise from 0.5 mM to 1.5 mM, from 0.75 mM to 1.25 mM, from 0.9 mM to 1.1 mM, or about 1 mM of $CaCl_2$. Buffer B can comprise from 150 mM to 250 mM, from 175 mM to 225 mM, from 190 mM to 210 mM, or about 200 mM sodium phosphate pH 7.0. Buffer B can comprise from 50 mM to 150 mM, from 75 mM to 125 mM, from 90 mM to 110 mM, or about 100 mM NaCl. Buffer B can comprise from 0.5 mM to 1.5 mM, from 0.75 mM to 1.25 mM, from 0.9 mM to 1.1 mM, or about 1 mM of $CaCl_2$.

[0125] For hydroxyapatite chromatography of SEQ ID NO: 15, Buffer A can comprise 20 mM Tris pH 7.5, 100 mM NaCl, and 1 mM $CaCl_2$ and Buffer B can comprise 200 mM sodium phosphate pH 7.0, 100 mM NaCl, and 1 mM $CaCl_2$ (TABLE 6). For hydroxyapatite chromatography of SEQ ID NO: 17, Buffer A can comprise 20 mM Tris pH 7.5, 100 mM NaCl, and 1 mM $CaCl_2$ and Buffer B can comprise 200 mM sodium phosphate pH 7.0, 100 mM NaCl, and 1 mM $CaCl_2$. For hydroxyapatite chromatography of SEQ ID NO: 14, Buffer A can comprise 20 mM Tris pH 7.5, 100 mM NaCl, and 1 mM $CaCl_2$ and Buffer B can comprise 200 mM sodium phosphate pH 7.0, 100 mM NaCl, and 1 mM $CaCl_2$ (TABLE 7).

[0126] The purity of the solubilized protein following the use of cation exchange resin or hydroxyapatite resin can be from 99% to 100%. The purity of the solubilized protein following the use of hydroxyapatite resin can be at least 95%, 96%, 97%, 98%, 99%, or 100%. The recovery of the solubilized protein following the use of hydroxyapatite resin can be from 85% to 100%, from 96% to 99%, or from 97% to 98%. The hydroxyapatite resin can be CaPure®.

[0127] Fractions collected during column chromatography that contain the purified compound (e.g., fusion protein) may be concentrated and formulated for administration using buffer exchange or diafiltration. For example, fractions collected during CaPure® can be concentrated and formulated using a TFF system (e.g., Pall corporation) to concentrate the protein. The protein can be concentrated to a final concentration of 20 mg/mL followed by 5-fold buffer exchange. The filtration process may be performed using ultrafiltration/diafiltration (UF/DF) and filters with a 10 kDa MWCO Millipore Pellican3, 0.114 m^2 TFF flat sheet cassettes. The MWCO of the filtration system may be varied depending on the protein (e.g., molecular weight) to be purified. The diafiltration buffer may consist of 10-20 mM sodium phosphate at about pH 7.0, 50-100 mM NaCl. Formulated SEQ ID NOS: 14-21 may subsequently be filtered using a flow filtration system employing an AkroPac 0.8/0.2 um filter and a Cole-Parmer peristaltic pump and tubing. Purified and formulated protein may be stored in aliquots at −80° C. until further use.

[0128] In specific embodiments, proteins analyzed by SEC-HPLC may show above 97% purity (typically >98%) using TSKgel GW3000SWXL, 5 μm, 7.8 mm IDx30.0 cm L column (Tosoh Bioscience, 8541).

Protein Analysis

[0129] Samples from different fractions may be analyzed by SDS-PAGE using the Bio-Rad ChemiDoc™ MP imaging system, and the fractions collected during column chromatography may be analyzed for protein content using Thermo Fisher Nanodrop One™. General techniques for protein detection or analysis include gel electrophoresis, fluorescence microscopy, capillary electrophoresis, mass spectrometry, electrophoretic mobility-shift assay, or nuclear magnetic resonance.

Methods of Treatment

[0130] In various embodiments of the present disclosure, pharmaceutical compositions comprising the fusion molecules of the disclosure are provided for use in treating and/or preventing inflammatory diseases. These pharmaceutical compositions can be formulated for oral delivery. "Inflammatory diseases" may include all diseases associated with acute or chronic inflammation. Acute inflammation is the initial response of the body to harmful stimuli and results from an increased movement of plasma and leukocytes (such as e.g. granulocytes) from the blood into the injured tissues. A number of biochemical events propagates and matures the inflammatory response, involving the local vascular system, the immune system, and various cells within the injured tissue. Prolonged inflammation is referred to as chronic inflammation, which leads to a progressive shift in the type of cells present at the site of inflammation and is characterized by simultaneous destruction and healing of the tissue from the inflammatory process. Inflammatory diseases can be caused by e.g. burns, chemical irritants, frostbite, toxins, infections by pathogens, physical injury, immune reactions due to hypersensitivity, ionizing radiation, or foreign bodies, such as e.g. splinters, dirt and debris. In some embodiments, the inflammatory disease is epithelial cell injury, hepatitis, obesity, fatty liver disease, liver inflammation, pancreatitis, Crohn's disease, fistulizing Crohn's disease, ulcerative colitis, mild-to-moderate ulcerative colitis, moderate-to-severe ulcerative colitis, pouchitis, proctitis, multiple sclerosis, systemic lupus erythematosus, graft versus host disease, rheumatoid arthritis, or psoriasis.

[0131] Further described herein are methods for treating a disease or condition in a subject, comprising administering to the subject the non-naturally occurring fusion protein. In some embodiments, the disease or condition is epithelial cell injury, hepatitis, obesity, fatty liver disease, liver inflammation, pancreatitis, Crohn's disease, fistulizing Crohn's disease, ulcerative colitis, mild-to-moderate ulcerative colitis, moderate-to-severe ulcerative colitis, pouchitis, proctitis, multiple sclerosis, systemic lupus erythematosus, graft versus host disease, rheumatoid arthritis, or psoriasis. The non-naturally occurring fusion protein can be oral administered to the subject. The subject can be human.

EXAMPLES

[0132] These examples are provided for illustrative purposes only and not to limit the scope of the claims provided herein.

Example 1

Expression of a Delivery Construct

[0133] In this example, the preparation of a delivery construct as a single amino acid sequence comprising a carrier sequence derived from Cholix, a spacer sequence, and a therapeutic payload, is generally described.

[0134] First, a nucleic acid sequence encoding a delivery construct of SEQ ID NO: 15 was amplified by PCR, incorporating restriction enzymes pairs of NdeI and EcoRI, PstI and PstI, AgeI and EcoRI, or PstI and EcoRI sites at two ends of the PCR products. After restriction enzyme digestion, the PCR products were cloned into an appropriate plasmid for cellular expression, which was digested with the corresponding restriction enzyme pairs. The plasmid encoded a delivery construct comprising the amino acid sequence set forth in SEQ ID NO: 15.

[0135] The delivery construct was expressed as follows: *E. coli* BL21(DE3) pLysS competent cells (Novagen, Madison, Wis.) were transformed using a standard heat-shock method in the presence of the appropriate plasmid to generate delivery construct expression cells, selected on ampicillin-containing media, and isolated and grown in Luria-Bertani broth (Difco; Becton Dickinson, Franklin Lakes, N.J.) with antibiotic, then induced for protein expression by the addition of 1 mM isopropyl-D-thiogalactopyranoside (IPTG) at OD 0.6. Two hours following IPTG induction, cells were harvested by centrifugation at 5,000 rpm for 10 min. Inclusion bodies were isolated following cell lysis and washed twice with water at a ratio of 1 g/10 mL. The pellet from the cell lysis was resuspended with water followed by one hour centrifugation at 10,000 rpm at 4° C. This wash was then repeated once. Double-washed IB (DWIB) therein were solubilized in a buffer containing 50 mM Tris-HCl (pH 8.2), 6 M guanidine HCl, and 10 mM dithiothreitol (DTT). Solubilized material was then diluted into a refold buffer containing 0.1 M Tris (pH=8.5 at 4° C.), 1.0 M L-arginine, 10% glycerol, 3 mM L-cysteine, 1 mM cystamine-2HCl. The refolded protein with SEQ ID NO: 15 was purified by anion exchange chromatography (Q sepharose Ion Exchange) and Superdex 200 Gel Filtration chromatography (Amersham Biosciences, Inc., Sweden). The purity of the protein was assessed by SDS-PAGE and analytic HPLC (Agilent, Inc. Palo Alto, Calif.).

[0136] The delivery construct was evaluated to verify the proper folding with regard to its anticipated molecular size. Following induction, expressed protein was collected from inclusion bodies. Using the inclusion bodies, the extent of expression of the delivery construct was verified by gel electrophoresis, and the apparent molecular weight was compared to the calculated mass.

[0137] The results demonstrated stable and efficient production of a functional delivery construct in high yield and purity.

Example 2

In Vitro Model Assessing Transport of Payload Across Epithelial Cell Monolayers

[0138] This example demonstrates an in vitro model designed to evaluate the transport properties of payloads or delivery constructs described herein.

[0139] FIG. **1** schematically shows a setup comprising an apical chamber above the epithelial cell monolayer and a basal chamber below such epithelial cell monolayer.

[0140] For apical to basolateral permeability, test articles (e.g., delivery construct, payload, etc.) were added to the apical (A) side and the amount of permeation was determined on the basolateral (B) side. For basolateral to apical permeability, test articles were added to the basolateral (B) side and amount of permeation was determined on the apical (A) side.

[0141] Data can be expressed as permeability (Papp) according to the following equation: Papp=(dQ/dt)/(CO*A). Q/dt is a rate of permeation, CO is initial concentration of test article, and A is the area of the monolayer. An efflux transport ratio (Re) can be calculated according to the following equation: (Re)=Papp(B-A)/Papp(A-B). Re>2 can indicate a potential substrate for P-gp or other active efflux transporters.

[0142] SMI-100 or Caco-2 cells can be used to assess the transcytosis function of a carrier or delivery construct in vitro.

[0143] For Caco-2 cells, an ELISA assay was performed to evaluate the ability of a carrier or delivery construct to move across Caco-2 cell monolayers via transcytosis. Caco-2 (ATCC HTB-37™) cells were maintained in 5% CO_2 at 37° C. in complete media: Dulbecco's modified Eagle's medium F12 (DMEM F12) supplemented with 10% fetal bovine serum, 2.5 mM glutamine, 100 U of penicillin/ml, and 100 μg of streptomycin/ml (Gibco BRL, Grand Island, N.Y.). Cells were fed every 2 to 3 days with this media (designated complete medium) and passaged every 5 to 7 days. For assays, cells were seeded into 24- or 96-well plates and grown to confluence.

[0144] Caco-2 cells were grown as confluent monolayers on collagen-coated 0.4-μm pore size polycarbonate membrane transwell supports (Corning-Costar, Cambridge, Mass.) and used 18-25 days after attaining a trans-epithelial electrical resistance (TER) of >250 Ω·cm2 as measured using a chopstick Millicell-ERS® voltmeter (Millipore). Apical to basolateral (A→B) transport of a carrier or delivery construct across these monolayer was determined by measuring the amount of transported protein at certain time points (e.g., 15, 30, and 45 minutes) after 4.7 nM, 23.6 nM and 236 nM apical application of delivery construct at 37° C. TER measurements and the extent of 10 kDa fluorescent dextran (measured using an HPLC size exclusion protocol) were used to verify monolayer barrier properties during the course of the study. The extent of transport of the delivery construct (e.g., SEQ ID NO: 15) was determined by titration of collected media in the cell-based cytotoxicity assay. Transported delivery construct was measured by enzyme linked immunosorbant assay (ELISA) using antibodies (e.g., anti-carrier or anti-payload, such as an anti-IL-22 antibody) for capture and detection.

[0145] Confluent monolayers of human small intestinal tissues (SMI-100, MatTek Corporation; Ashland, Mass., USA) established on cell culture inserts were allowed to stabilize for 24 h at 37° C. prior to use. Only inserts having a trans-epithelial electric resistance (TEER) of >400 $\Omega \cdot cm^2$ were considered to have sufficient monolayer integrity for use in studies. A secondary verification of monolayer integrity was performed by assessing suppression of 70 kD dextran transport. The chambers were washed once with transport buffer (PBS). Test molecules, prepared at a con-

centration of 20 μg/mL, were applied to the apical surface of inserts in 100 μL volumes. Basolateral volumes of 500 μL PBS were replaced at each time point for transport studies. Each experimental condition was performed in triplicate.

Example 3

Carrier-Mediated Transport of IL-22 Across Polarized Gut Epithelial Cells

[0146] This example demonstrates that a carrier (SEQ ID NO: 7) can transport an IL-22 payload (SEQ ID NO: 11) across polarized gut epithelial cells in vitro. This example further demonstrates that the carrier with SEQ ID NO: 7 can transport biologically active IL-22 payload across polarized gut epithelial cells and to the *lamina propria* in vivo.

[0147] Transport of delivery construct (SEQ ID NO: 15) across Caco-2 cell monolayers and small intestine epithelial tissue (also referred to herein as SMI-100) was tested by applying the delivery construct to the apical membrane of the epithelial cells, according to EXAMPLE 2 and as illustrated in FIG. **1** for apical (A) to basal (B) transport. Experiments were run in duplicate and samples from the basolateral chamber were collected at 15, 30, and 45 minutes after apical application to determine the amount of transported protein. The amount of transcytosed protein was measured using ELISA assays as described in EXAMPLE 2.

[0148] The data in FIG. **2** and FIG. **3** show that a carrier (SEQ ID NO: 7) when coupled to IL-22 (SEQ ID NO: 11) resulted in the transport of IL-22 payload across both Caco-2 (FIG. **2**) and SMI-100 (FIG. **3**) monolayers in a time-dependent manner, and that the delivery construct resulted in about 2-3 fold more IL-22 crossing the epithelial cells as compared to an IL-22 (SEQ ID NO: 12, M+IL-22^{34-179}) alone that was not coupled to a carrier.

[0149] For in vivo experiments, transcytosis was tested using male Wistar rats. Male Wistar rats were housed 3-5 per cage with a 12/12 h light/dark cycle and were about 225-275 g (approximately 6-8 weeks old) when placed on study. Experiments were conducted during the light phase using a non-recovery protocol that uses continuous isoflurane anesthesia. A 4-5 cm midline abdominal incision that exposed mid-jejunum regions was conducted. Stock solutions at 3.86×10^{-5} M of delivery construct (SEQ ID NO: 15) were prepared in phosphate buffered saline (PBS), with 50 μL (per 250 g rat) being administered by intraluminal injection (ILI) using a 29-gauge needle. The injection site mesentery was then marked with a permanent marker. At study termination, a 3-5 mm region that captured the marked intestine segment was isolated and processed for microscopic assessment.

[0150] The results of the transcytosis activity of the delivery construct (SEQ ID NO: 15) are shown in FIG. **4**, demonstrating that significant amounts of IL-22 payload (SEQ ID NO: 11) crossed an intact and polarized gut epithelium in vivo when provided as part of a delivery construct that includes a cholix-derived carrier (SEQ ID NO: 7) coupled to IL-22. This microscopy image shows transportation of the IL-22 payload (SEQ ID NO: 11) from the apical site of the gut epithelium (highlighted by white arrows #1) to the basal site of the epithelial cells (highlighted by white arrow #2) and into the *lamina propria* (abbreviated as "l.p.") after luminal application of the delivery construct of SEQ ID NO: 15 to the jejunum of Wistar rats. The image further shows that the IL-22 interacted and bound to a significant extent to IL-22 receptors located on cells within the *lamina propria* and on the outer basal membrane of the polarized epithelium (highlighted by white arrows #3). IL-22 localization is indicated by white arrows #2 and #3.

Example 4

Recombinant Human IL-22 Binds to Murine IL-22 Receptors

[0151] This example demonstrates that recombinant human IL-22 (rhIL-22, SEQ ID NO: 12) binds to murine IL-22 receptors in a dose-dependent manner comparable to recombinant murine IL-22 (rmIL-22).

[0152] FIG. **5**A demonstrates STAT3 phosphorylation in murine FL83 B cells as a function of agonist concentration (in pM), wherein the agonist is rhIL-22 (SEQ ID NO: 12) or rmIL-22. The data show that rhIL-22 and rmIL-22 induced STAT3 phosphorylation in a concentration-dependent manner, demonstrating that the human IL-22 protein can bind and induce signal transduction through mouse IL-22 receptors as effectively as mouse IL-22.

[0153] FIG. **5**B demonstrates that rhIL-22 (SEQ ID NO: 12) and rmIL-22 also induced STAT3 phosphorylation in murine Hepa1-6 cells.

[0154] FIG. **6**A demonstrates that rhIL-22 (SEQ ID NO: 12) and rmIL-22 induced STAT5 phosphorylation in murine FL83 B cells in a dose-dependent manner, demonstrating that the human IL-22 protein can bind and induce signal transduction through mouse IL-22 receptors with efficacy comparable to mouse IL-22.

[0155] FIG. **6**B demonstrates that rhIL-22 (SEQ ID NO: 12) and rmIL-22 also induced STAT5 phosphorylation in murine Hepa1-6 cells but additional dose ranging will need to be performed in follow up experiments.

[0156] These data provide strong evidence that rhIL-22 (e.g., as used in the delivery constructs with SEQ ID NOs: 10 or 11) is capable of activating STAT3 and STAT5 in murine cells, thus providing rationale for using such delivery constructs in murine models to assess IL-22 activity and function.

Example 5

In Vivo Efficacy of the Delivery Construct with SEQ ID NO: 15 in Colitis Model

[0157] This example demonstrates the in vivo efficacy of the delivery construct with SEQ ID NO: 15 comprising the Cholix derived carrier of SEQ ID NO: 7 coupled to an IL-22 with SEQ ID NO: 11 via a linker with SEQ ID NO: 13 compared to IL-22 alone (SEQ ID NO: 12) in a dextran sulfate sodium (DSS) mouse model.

[0158] FIG. **7**A shows an overview of an in vivo study design and timeline for comparing peroral (p.o.) administration of the delivery construct of SEQ ID NO: 15 at two once-daily dose levels, 1 mg/kg and 30 mg/kg, and one once daily administration (intraperitoneal (i.p.)) of rhIL-22 at 4 mg/kg. The in vivo study was performed in Normal chow-fed, 8-10 week old, ~23 g, female C57BL/6 mice. Colitis was chemically induced with 2.5% dextran sulfate sodium (DSS) in drinking water, ad libitum. Study endpoints included percent change in body weight, disease activity index, colon length and weight, and histopathology.

[0159] FIG. **7**B shows characteristics of the experimental and control groups used in the study described in FIG. **7**A.

[0160] FIG. **8**A shows the percent (%) change in body weight of animals at study day 10 relative to day 0 of the various experimental and control groups. Baseline (day 0 vs. 10) body weight change by rhIL-22 with SEQ ID NO: 12 or the delivery construct with SEQ ID NO: 15 relative to vehicle was not significant as assessed by 1-way ANOVA. Positive model control (cyclosporine) CsA was significantly different relative to Vehicle as assessed by a 1-way ANOVA.

[0161] FIG. **8**B shows the change in body weight of experimental and control animals over the first 10 study days. CsA group mean body weight over the study period was significantly improved relative to vehicle control day 6 through day 10 as assessed by 2-way ANOVA(*, **, ***) rhIL-22 (i.p.) body weight was significantly improved relative to vehicle control at day 10 as assessed by 2-way ANOVA (*), *p<0.05, **p<0.01, ***p<0.001.

[0162] FIG. **9** shows results of plasma ELISA experiments to determine IL-22 plasma levels. The results show that plasma rhIL-22 concentrations trended towards increased levels in plasma after oral gavage of both 1 and 30 mg/kg doses of the delivery construct with SEQ ID NO: 15 in a dose-dependent manner (*p<0.05, **p<0.01, ***p<0.001. ****p<0.0001; Mean+SEM; n=2 Naïve, 5 Vehicle, 5 CsA, 10 others).

[0163] These data show that the delivery construct with SEQ ID NO: 15 induced a dose-dependent trend towards body weight improvement after administration via oral gavage. IL-22 was measured in increased levels in plasma after oral gavage of both 1 and 30 mg/kg doses of the delivery construct with SEQ ID NO: 15 in a dose-dependent manner. These data suggest that orally administered delivery construct with SEQ ID NO: 15 is capable of reducing symptoms of colitis in a DSS mouse model comparable to i.p. administered rhIL-22.

[0164] In addition, biomarkers for the orally administrable delivery construct with SEQ ID NO: 15 were evaluated by administering rhIL-22 to CD-1 mice as described below.

[0165] FIG. **10**A shows an overview of a single dose in vivo study designed to identify biomarker(s) of target engagement after single (acute dosing) of rhIL-22 in healthy CD-1 mice.

[0166] FIG. **10**B shows an overview of a multiple dose (sub-chronic dosing) in vivo study designed to identify biomarker(s) of target engagement after single and multiple doses of rhIL-22 in healthy CD-1 mice.

[0167] FIG. **11**A shows that acute and sub-chronic administration of rhIL-22 increases IL-22 concentration consistently.

[0168] FIG. **11**B shows some pharmacokinetic parameters measured during the acute and sub-chronic dosing experiments described in FIG. **11**A.

[0169] FIG. **12**A shows the induced circulating murine C-reactive protein (mCRP) concentration in response to 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12) and after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12). The results show a time-dependent increase in circulating mCRP in both study groups.

[0170] FIG. **12**B shows fold change of plasma mCRP concentration 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12).

[0171] FIG. **12**C shows fold change of plasma mCRP concentration after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12).

[0172] FIG. **13**A shows the induced circulating murine serum amyloid protein A (mSAA) concentration in response to 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12) and after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12). The results show a time-dependent increase in circulating mSAA in both study groups, with approximate plasma peak concentrations about 8 hours after administration

[0173] FIG. **13**B shows fold change of plasma mSAA concentration 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12).

[0174] FIG. **13**C shows fold change of plasma mSAA concentration after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12).

[0175] FIG. **14**A shows the induced circulating regenerating islet-derived protein 3/3 (Reg3β) in response to 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12) and after 5 days of sub-chronic dosing rhIL-22 (SEQ ID NO: 12).

[0176] FIG. **14**B shows fold change of plasma Reg3β concentration 1 day after acute dosing of rhIL-22 (SEQ ID NO: 12).

[0177] FIG. **14**C shows fold change of plasma Reg3β concentration after 5 days of sub-chronic dosing of rhIL-22 (SEQ ID NO: 12).

[0178] These results demonstrate that three potential IL-22 PD biomarkers of target engagement: C reactive protein (CRP), serum amyloid protein A (SAA), and regenerating islet-derived protein 3/3 (Reg3β) that may be used in studies evaluating the pharmacodynamics (PD) of delivery constructs such as those with SEQ ID NOs: 15 or 17.

Example 6

Protein Solubilization

[0179] 627 g of double-washed inclusion bodies (DWIB) from SEQ ID NO: 15 fermentation were resuspended in 4500 mL of 8M Guanidine/HCl (Gu-HCl), and 50 mM Tris pH at 8.0. Subsequently, 50 mM Tris pH 8.5 was added to complete the volume to 6.0 L. The solution was stirred gently on a stir-plate and 9.225 g of the reducing agent dithiothreitol (DTT) was added. The final solubilized protein (SEQ ID NO: 15) solution of 6 L consisted of 6M Gu-HCl, 50 mM Tris pH 8.0, 10 mM DTT, and ~1 g DWIB/10 mL buffer. The solubilized protein (SEQ ID NO: 15) solution was incubated at room temperature (RT) while stirred on a magnetic stir-plate, and then centrifuged at 15,970×g for 90 minutes at 4° C. The supernatant comprising the solubilized protein was carefully transferred to a new vessel with a total volume of 5520 mL. The protein concentration determination was performed by Bradford assay and was determined at 15 mg/mL.

[0180] The data demonstrate that protein solubilization can be performed using the above procedure.

Example 7

Protein Refolding

[0181] This example demonstrates protein refolding as part of the purification process as described in the flow chart of FIG. **17**.

[0182] A refolding solution was carefully researched, developed and optimized (FIG. **18**). The optimization process utilized a design of experiments (DOE) of 15 different matrices, 12 variables and 200 refold reactions. Decisions were made by a process of elimination. The initial refold

mixture (containing the initial refold solution and the refolded solubilized protein (SEQ ID NO: 15)) is presented in TABLE 2. An optimized refold mixture (containing the optimized refold solution and the refolded solubilized protein (SEQ ID NO: 15)) is presented in TABLE 3.

TABLE 2

Initial refold mixture
100 mM Tris pH 8.5 (@ 4° C.)
0.5M arginine
1M urea
2 mM EDTA
0.3 mM GSH
1 mM GSSG
0.2 mg/mL refold concentration

TABLE 3

Optimized refold mixture
100 mM Tris pH 8.5 (@ 4° C.)
1.0M arginine
10% glycerol
3 mM L-cysteine
1 mM cystamine-2HCl
>0.1 mg/mL refold concentration

[0183] The 105 L refold solution was incubated at 4° C. for 16 hours and then filtered through a flow filter (AkroPac by Pall corporation or Sartopore 2XLG by Sartorius) using a 0.8/0.2 um membrane and Cole-Parmer peristaltic pump and tubing.

[0184] The solubilized protein (SEQ ID NO: 15) (105 g) from EXAMPLE 6 was added to 100 L of an optimized refold solution prechilled to 4° C.

[0185] This process was programmed to one hour using a Cole-Parmer peristaltic pump set at 73 ml/min. This process took place in a cold room at 4° C.

[0186] The optimized refold buffer solution produced a larger amount of the SEQ ID NO: 15 relative to aggregates of SEQ ID NO: 15 (FIGS. **19-20**). Use of the initial refold solution produced about 10%-15% correctly refolded SEQ ID NO: 15. Use of the optimized refold solution produced about 45%-55% correctly refolded SEQ ID NO: 15.

Example 8

Protein Concentration and Buffer Exchange by TFF UF/DF

[0187] This example demonstrates protein concentration and buffer exchange using tangential flow filtration (TFF) based on ultrafiltration/diafiltration (UF/DF) principles.

[0188] Refolded protein (e.g., of SEQ ID NO: 15) was processed by TFF system (Millipore) to concentrate it 10-fold and buffer exchange 5-fold. The process was performed by ultrafiltration/diafiltration (UF/DF) using four 10 kDa MWCO Millipore Ultracell Pellican3, 1.14 m^2 TFF flat sheet cassettes. The diafiltration buffer consisted of 20 mM Tris-base pH 7.5 and 100 mM NaCl. The final volume at the end of the UF/DF process was 10 L. This material was subsequently filtered using a flow filtration system employing AkroPac 0.8/0.2 um filter by Pall corporation and Cole-Parmer peristaltic pump and tubing. Bradford assay was performed to determine the protein concentration, yield, and step recovery. At this point, a quantitative SEC-HPLC was performed to determine the amount of the properly folded protein with SEQ ID NO: 15 present in the refold and UF/DF samples. A commercially available BSA of a known concentration was used as a reference standard from which the standard curve was generated for the Bradford assay. The Agilent 1100 HPLC system and the TSKgel SuperSW3000, 4 μm, 4.6 mm ID×30.0 cm L (Tosoh Bioscience, 18675) column were used. Based on this quantitative assay, the refold efficiency was determined.

[0189] The data show that protein refolding can be performed using buffer exchange and TFF UF/DF.

Example 9

Protein Chromatography Using the Capture Step NH_2-750F® Resin

[0190] This example demonstrates capture steps in protein anion exchange chromatography using the NH_2-750F® resin.

[0191] The AKTA Avant 150 or AKTA Pilot FPLC systems from General Electric (GE) were used for protein chromatography. The Tosoh anion exchange resin NH_2-750F® was used to pack a column with a minimum bed-height of at least 20 cm and ensuring a column volume that would facilitate a dynamic binding capacity of 20-25 g/L. A buffer of 20 mM sodium acetate, pH 4.5 or 20 mM sodium citrate, pH 4.5 were used to pack the column. The buffers used were as follow: buffer A: 20 mM Tris pH 7.5, 0.5 M NaCl; buffer B: 20 mM Tris pH 7.5, 2.0 M NaCl. The NH_2-750F column was then cleaned with 0.5 M NaOH solution with a contact time of 30 minutes, and then equilibrated with buffer A for at least 3 column volumes (CV), or until pH and conductivity reach stable lines at the expected values (pH 7.5-pH 7.7, ~49 mS/cm+/−1 mS/cm).

[0192] Prior to loading onto the column, the UF/DF protein (SEQ ID NO: 15) solution was supplemented with 0.4 M NaCl by adding a 5 M stock solution and the conductivity was measured to ensure conductivity of 49 mS/cm+/−2 mS/cm. The protein (SEQ ID NO: 15) solution was loaded onto the column with a flow-rate of a minimum of 5 minutes column residence time, and the flow-through was collected. The column was washed with 3 CV of buffer A or until the absorbance at 280 nm returned, and was stable at baseline, near 0.0 mAU at 280 nm. At this point a linear gradient of 20 CV from 0.0-62.5% B, followed by a step gradient to 100% B for additional 5 CV was performed. Fractions were collected throughout and their volumes were no greater than 0.5 of the column volume. Samples from different fractions were analyzed by SDS-PAGE using the Bio-Rad Chemi-Doc™ MP imaging system, and the fractions containing over 90% of SEQ ID NO: 15 were pooled and designated as NH_2-750F-pool. The protein concentration of the NH_2-750F-pool was measured using Thermo Fisher Nanodrop One™, by reading the absorbance at 280 nm (A280) considering extinction coefficient of 1.22 for SEQ ID NO: 15 and a 260/280 nm ratio<0.6. The NH_2-750F-pool contains approximately 1.0 M NaCl.

[0193] An example chromatogram following NH_2-750F purification of SEQ ID NO: 15 is shown in FIG. **21**A for a bed height of 20 cm and a column volume of 10 mL and FIG. **21**B for a bed height of 30 cm and a column volume of 4.6 L.

[0194] A summary of the NH_2-750F purification is show in TABLE 4 for SEQ ID NO: 15 and TABLE 5 for SEQ ID NO: 14, which was generally produced, refolded, and purified in a manner analogous to SEQ ID NO: 15.

TABLE 4

Summary of SEQ ID NO: 15 NH_2-750 F. purification SEQ. ID NO: 15

Buffer A	20 mM Tris pH 7.5, 0.5M NaCl
Buffer B	20 mM Tris pH 7.5, 2.0M NaCl
Bed Height	Minimum 20 cm
Residence Time	5.0 minutes
Gradient	0.0-62.5% B, 20 CV
Binding capacity	20-25 g/L (Breakthrough at 40 g/L)
Purity	93-96%
Recovery	71-76%
Endotoxin	<1.0 EU/mg

TABLE 5

Summary of SEQ ID NO: 14 NH_2-750 F. purification SEQ ID NO: 14

Buffer A	20 mM Tris pH 7.5, 0.7M NaCl, 10% glycerol
Buffer B	20 mM Tris pH 7.5, 2.0M NaCl, 10% glycerol
Bed Height	Minimum 20 cm
Residence Time	5.0 minutes
Gradient	0.0-62.5% B, 20 CV
Binding capacity	20-25 g/L
Purity	≥93-96%
Recovery	80-93%
Endotoxin	<1.0 EU/mg

Example 10

Protein Chromatography Using the CaPure® Procedure

[0195] This example demonstrates a polishing purification step in protein chromatography using the CaPure® procedure.

[0196] The AKTA Avant 150 or AKTA Pilot FPLC systems from General Electric (GE) were used for protein chromatography of proteins of SEQ ID NO: 15 and SEQ ID NO: 14. The Tosoh hydroxyapatite mixed-mode resin CaPure® was used to pack a column for each of SEQ ID NO: 15 and SEQ ID NO: 14 with a minimum bed-height of at least 10 cm and ensuring a column volume that would facilitate a dynamic binding capacity of maximum 20 g/L. The buffers used were as follows for SEQ ID NO: 15: buffer A: 20 mM Tris pH 7.5, 100 mM NaCl, and 1 mM $CaCl_2$; buffer B: 200 mM sodium phosphate pH 7.0, 100 mM NaCl, and 1 mM $CaCl_2$. The buffers used were as follows for SEQ ID NO: 14: buffer A: 20 mM Tris pH 7.5, 100 mM NaCl, and 1 mM $CaCl_2$; buffer B: 200 mM sodium phosphate pH 7.0, 100 mM NaCl, and 1 mM $CaCl_2$. Each CaPure® column was then cleaned with 0.5 M of NaOH with contact time of 30 minutes or more, and then equilibrated with buffer A for at least 3 column volumes (CV), or until pH and conductivity reach stable lines at the expected values (pH 7.5-pH 7.7, ~11 mS/cm+/−1 mS/cm). There was no need to treat the NH_2-750F-pool prior to loading it onto the column. This significantly reduced protein loss, time, and resources. The NH_2-750F pool was loaded onto the column with a flow-rate of a minimum of 5 minutes column residence time, and the flow-through was collected. The column was washed with 3 CV of buffer A or until the absorbance at 280 nm returned, and was stable at baseline, near 0.0 mAu. At this point, a linear gradient of 25 CV from 0-25% B, followed by a step gradient to 100% B for additional 5 CV was performed. Fractions were collected throughout and their volumes were ranging from 0.36 to 1.43 of the column volume, depending on the elution profile. Samples from different fractions were analyzed by SDS-PAGE using the ChemiDoc™ MP imaging system, and the fractions containing over 95% of SEQ ID NO: 15 or SEQ ID NO: 14 were pooled and designated as CaPure-pool. The protein concentration of the CaPure-pool was measured using Thermo Fisher Nanodrop One™, by reading the absorbance at 280 nm (A280) considering extinction coefficient of 1.22 for SEQ ID NO: 15 and a 260/280 nm ratio<0.6.

[0197] An exemplary chromatogram following CaPure® purification of SEQ ID NO: 15 is shown in FIG. **22**A for a bed height of 10 cm and column volume of 5 mL and FIG. **22**B for a bed height of 21 cm and a column volume of 800 mL.

[0198] A summary of the CaPure® purification is show in TABLE 6 for SEQ ID NO: 15 and TABLE 7 for SEQ ID NO: 14.

TABLE 6

Summary of SEQ ID NO: 15 CaPure ® purification SEQ. ID NO: 15

Buffer A	20 mM Tris pH 7.5, 100 mM NaCl, 1 mM $CaCl_2$
Buffer B	200 mM sodium phosphate pH 7.0, 100 mM NaCl, 1 mM $CaCl_2$
Bed Height	20 cm
Residence Time	5.0 minutes
Gradient	0.0-25% B, 25 CV
Binding capacity	20 g/L
Purity	>98%
Recovery	>92%
Endotoxin	<1.0 EU/mg

TABLE 7

Summary of SEQ ID NO: 14 CaPure ® purification SEQ ID NO: 14

Buffer A	20 mM Tris pH 7.5, 100 mM NaCl, 1 mM $CaCl_2$
Buffer B	200 mM sodium phosphate pH 7.0, 100 mM NaCl, 1 mM $CaCl_2$
Bed Height	20 cm
Residence Time	5.0 minutes
Gradient	0.0-25%B, 25 CV
Binding capacity	20 g/L
Purity	>98%
Recovery	>92%
Endotoxin	<1.0 EU/mg

Example 11

Protein Concentration by TFF UF/DF

[0199] This example demonstrates protein formulation by TFF UF/DF procedures.

[0200] The CaPure® pool was concentrated by a TFF system (Pall corporation) to a final concentration of 20

mg/mL followed by 5-fold buffer exchange. The process was performed by ultrafiltration/diafiltration (UF/DF) using three 10 kDa MWCO Millipore Pellican3, 0.114 m^2 TFF flat sheet cassettes. The diafiltration buffer consisted of 10 mM sodium phosphate pH7.0, 100 mM NaCl. The formulated SEQ ID NO: 15 was subsequently filtered using a flow filtration system employing AkroPac 0.8/0.2 um filter by Pall corporation and Cole-Parmer peristaltic pump and tubing. The formulated SEQ ID NO: 15 was then stored in aliquots at −80° C. and the following analyses were performed to ensure its biophysical quality: (1) Protein (e.g., SEQ ID NO: 15) concentration at 20 mg/mL by measuring the absorbance at 280 nm with 260/280 ratio<0.6 using Thermo Fisher Nanodrop One™ and considering protein extinction coefficient of 1.22; (2) LAL endotoxin levels below 1.0 Eu/mg measured by Charles River Laboratories equipment; (3) Purity by SEC-HPLC above 97% purity (typically >98%) using TSKgel GW3000SWXL, 5 μm, 7.8 mm ID×30.0 cm L column (Tosoh Bioscience, 8541) (4) SDS-PAGE analysis using the Bio-Rad gel apparatus and its associated ChemiDoc™ MP imaging system.

Example 12

Evaluation of In Vivo Activity of Purified Fusion Molecules

[0201] This example demonstrates methods for verifying the proper folding of the fusion molecules with regard to their ability to carry a biologically active cargo across an intact epithelium.

[0202] The fusion protein with SEQ ID NO: 15 is expressed by *E. coli* and collected from inclusion bodies and folded using a shuffle exchange buffer system as described in EXAMPLE 8 above. The resulting material is purified according to the methods described in EXAMPLE 9 and EXAMPLE 10, and as summarized for example in TABLES 4-7, depending on the molecular characteristics of the fusion molecule. The preparation has a protein purity of ~98% based upon SDS PAGE. Epithelial cells are treated with the fusion molecule having SEQ ID NO: 15 at concentrations of 25 nM and 250 nM. Compared to untreated matched cells, SEQ ID NO: 15 treated cells produce a dose-dependent decrease in cell number as assessed by flow cytometry of live/dead cells). Values represent n=4±standard deviation.

Example 13

Scale-Up of Fusion Protein Purification

[0203] This example demonstrates a scale-up of a purification method using the fusion protein of SEQ ID NO: 15 (FIGS. **7**A and **7**B).

[0204] The fusion protein with SEQ ID NO: 15 was purified on two different scales.

[0205] The first purification was carried out using a chromatography column with a volume of 10 mL packed with NH_2-750F® resin and the following parameters: Bed height: 20 cm, Residence time: 5 min (flow rate: 2 mL/min), buffer A: 20 mM Tris pH 7.5, 0.5 M NaCl, buffer B: 20 mM Tris pH 7.5, 2 M NaCl, and using the following gradient: 0.0-62.5% B for 20 column volumes (CVs). The fusion protein with SEQ ID NO: 15 was obtained with a purity of 93-96%, a recovery of >71%, and endotoxin levels<1.0 EU/mg.

[0206] The second purification was carried out using a chromatography column with a volume of 4.6 L packed with NH_2-750F® resin and the following parameters: Bed height: 30 cm, Residence time: 11.55 min (flow rate: 400 mL/min), buffer A: 20 mM Tris pH 7.5, 0.5 M NaCl, buffer B: 20 mM Tris pH 7.5, 2 M NaCl, and using the following gradient: 0.0-62.5% B for 20 column volumes (CVs). The fusion protein with SEQ ID NO: 15 was obtained with a purity of 93-96%, a recovery of >71%, and endotoxin levels<1.0 EU/mg.

Example 14

Consistency Between Process Step Recoveries

[0207] The purification process was carried out to purify the fusion protein of SEQ ID NO: 15 four times (lots 1-4), and the recovery of this fusion protein was assessed after each stage of the purification process (TABLE 8). There was a high degree of consistency in the recovery of the fusion protein between these replicates.

TABLE 8

Recovery of the fusion protein of SEQ ID NO: 15 in 4 replicate purification processes

Process step	Process description	Lot 1	Lot 2	Lot 3	Lot 4	Average (%)
2	Refold UF/DF	98.60%	94%	92.40%	85%	92
3	NH_2-750F	68.60%	66.30%	76%	74.50%	71.35%
4	CaPure	99%	96.40%	100%	94.80%	97.55
5	Formulation UF/DF	97.80%	100%	98.75%	100%	99.14

[0208] These data demonstrate that the methods and compositions of the present disclosure allowed the production and purification of therapeutic fusion protein on a large scale, and that the herein disclosed methods and compositions provided high consistency during scale-up.

[0209] All of the methods disclosed and claimed herein may be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this disclosure have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the disclosure. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the disclosure as defined by the appended claims.

Example 15

Comparison of SEQ ID NO: 15 and SEQ ID NO: 17

[0210] Liquid chromatography-mass spectrophotometry (LC-MS) was performed on purified compositions representing SEQ ID NO: 15 and SEQ ID NO: 17, and the results are represented in FIG. **23**B and FIG. **23**A, respectively. The

composition of SEQ ID NO: 17 was produced, refolded, and purified in a manner analogous to that described above in connection with the composition of SEQ ID NO: 15.

Example 16

In Vivo Assessment of Delivery Constructs

[0211] This example describes an in vivo assessment of a delivery construct having the sequence of SEQ ID NO: 15 compared to the vehicle.

[0212] FIG. **24** shows a (single-dose) study overview for assessing transport of the delivery construct having the sequence of SEQ ID NO: 15 across the gut epithelium and into circulation after oral (abbreviated herein as P.O.) delivery of the construct. The animals used in these experiments were CD-1 male mice without fasting. A primary study endpoint was total plasma IL-22 exposure, and a secondary endpoint were plasma biomarkers, including IL-22 binding protein (e.g., IL-22BP).

[0213] TABLE 9 provides details regarding the experimental setup, including information regarding the liquid (unformulated) delivery construct and the vehicle.

TABLE 9

Experimental Setup of the In Vivo Study

Group	Test Article	Vehicle	Formulation	Dose (mpk)	Route	N (4/two TP)	Timepoints (hours)
1	Vehicle	PBS, pH 7	Aq. bolus	na	P.O.	4	1
2	SEQ ID NO: 15 (liquid)	PBS, pH 7	Aq. bolus	90	P.O.	8 (4*2)	1, 2, 4, 6

[0214] TABLE 10 summarizes properties of the SEQ ID NO: 15 construct as delivered in this experiment.

TABLE 10

Properties of Delivery Construct as Delivered in Liquid Form

Groups	Groups	Cmax (pg/mL)	Tmax (hr)	AUC (0-6 hr)
SEQ ID NO: 15-Liquid	n/a	24488	2	58126

[0215] The total plasma IL-22 exposure was measured at the respective time points shown in TABLE 9. FIG. **25**A shows total systemic IL-22 exposure (in pg/mL) and FIG. **25**B shows IL-22BP exposure (in pg/mL) at the various time points where the blood samples and measurements were taken. FIG. **26**A-FIG. **26**B show individual measurements and data points for each group tested in this study.

[0216] It was observed that orally administered delivery construct having the sequence set forth in SEQ ID NO: 15 provided IL-22 in the plasma.

Example 17

Spacer Length and Coupling of Payload to the N- or C-Terminus of a Carrier does not Significantly Affect a Payload's Biological Activity

[0217] This example shows that amino acid linkers of various lengths and the coupling of a heterologous payload to the N-terminus of a carrier does not significantly impact the payload's ability to bind its target when included into a delivery construct. The amino acid linkers examined were SEQ ID NO: 13 (GGGGSGGGGSGGGGS), SEQ ID NO: 28 (GGGGS), and SEQ ID NO: 31 (GGGGSGGGGSGGGGSGGGGSGGGGS).

[0218] The IL-22 receptor dimerization assay was performed by seeding DiscoverX HEK293 cells and incubating the cells for 16 h (5,000 cells per well) using the shown concentrations of agonist (delivery construct containing the IL-22 payload). The endpoint luminescence was read on a plate reader using PathHunter® eXpress IL22RA1/IL10RB Dimerization Assay.

[0219] FIG. **27**A shows that the length of amino acid spacers with SEQ ID NOs: 13, 28, and 31 did not impact the ability of IL-22 (SEQ ID NO: 11) when included in the delivery constructs with SEQ ID NOs: 15, 32, and 33 to induce IL-22 receptor dimerization. The induction of receptor dimerization of control recombinant human IL-22 (rhIL-22, SEQ ID NO: 12) is shown by the black curve.

[0220] FIG. **27**B shows that coupling of the IL-22 payload (SEQ ID NO: 11) to the N- or to the C-terminus of a carrier comprising amino acid residues 1-266 of SEQ ID NO: 1 via the spacer with SEQ ID NO: 28 did not significantly change the ability of the delivery constructs with SEQ ID NOs: 32, 34, and 35 to induce IL-22 receptor dimerization. The induction of receptor dimerization of control recombinant human IL-22 (rhIL-22) is shown by the black curve.

[0221] The pSTAT3 activation assay was conducted using Colo205 cells incubated with 10 μL of agonist (the respective delivery construct or IL-22 control) having the various concentrations for 15 min. The extent of pSTAT3 activation was then read using MSD STAT3 plates (Cat. No. N450SMA-1).

[0222] FIG. **27**C shows that the length of amino acid spacers with SEQ ID NOs: 13, 28, and 31 did not impact the ability of IL-22 (SEQ ID NO: 11) when included in the delivery constructs with SEQ ID NOs: 15, 32, and 33 to induce pSTAT3 activation. The pSTAT3 activation of control recombinant human IL-22 (rhIL-22, SEQ ID NO: 12) is shown by the black curve.

[0223] FIG. **27**D shows that coupling of the IL-22 payload (SEQ ID NO: 11) to the N- or to the C-terminus of a carrier comprising amino acid residues 1-266 of SEQ ID NO: 1 via the spacer with SEQ ID NO: 28 did not significantly change the ability of the delivery constructs with SEQ ID NOs: 32, 34, and 35 to induce pSTAT3 activation. The pSTAT3 activation of control recombinant human IL-22 (rhIL-22) is shown by the black curve.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 41

<210> SEQ ID NO 1
<211> LENGTH: 634
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 1

```
Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1               5                  10                  15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
            20                  25                  30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
        35                  40                  45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
    50                  55                  60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65                  70                  75                  80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
                85                  90                  95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
            100                 105                 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115                 120                 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
    130                 135                 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145                 150                 155                 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
                165                 170                 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
            180                 185                 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
        195                 200                 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
    210                 215                 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225                 230                 235                 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
                245                 250                 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala Leu
            260                 265                 270

Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg Ser
        275                 280                 285

Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln Ala
    290                 295                 300

Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser His
305                 310                 315                 320

Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu Val
                325                 330                 335
```

-continued

```
Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro Gly
            340                 345                 350

Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp Tyr
        355                 360                 365

Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly Ala
    370                 375                 380

Gln Ala Ala Asp Ile Leu Ser Leu Phe Cys Pro Asp Ala Asp Lys Ser
385                 390                 395                 400

Cys Val Ala Ser Asn Asn Asp Gln Ala Asn Ile Asn Ile Glu Ser Arg
                405                 410                 415

Ser Gly Arg Ser Tyr Leu Pro Glu Asn Arg Ala Val Ile Thr Pro Gln
            420                 425                 430

Gly Val Thr Asn Trp Thr Tyr Gln Glu Leu Glu Ala Thr His Gln Ala
        435                 440                 445

Leu Thr Arg Glu Gly Tyr Val Phe Val Gly Tyr His Gly Thr Asn His
    450                 455                 460

Val Ala Ala Gln Thr Ile Val Asn Arg Ile Ala Pro Val Pro Arg Gly
465                 470                 475                 480

Asn Asn Thr Glu Asn Glu Glu Lys Trp Gly Gly Leu Tyr Val Ala Thr
                485                 490                 495

His Ala Glu Val Ala His Gly Tyr Ala Arg Ile Lys Glu Gly Thr Gly
            500                 505                 510

Glu Tyr Gly Leu Pro Thr Arg Ala Glu Arg Asp Ala Arg Gly Val Met
        515                 520                 525

Leu Arg Val Tyr Ile Pro Arg Ala Ser Leu Glu Arg Phe Tyr Arg Thr
    530                 535                 540

Asn Thr Pro Leu Glu Asn Ala Glu Glu His Ile Thr Gln Val Ile Gly
545                 550                 555                 560

His Ser Leu Pro Leu Arg Asn Glu Ala Phe Thr Gly Pro Glu Ser Ala
                565                 570                 575

Gly Gly Glu Asp Glu Thr Val Ile Gly Trp Asp Met Ala Ile His Ala
            580                 585                 590

Val Ala Ile Pro Ser Thr Ile Pro Gly Asn Ala Tyr Glu Glu Leu Ala
        595                 600                 605

Ile Asp Glu Glu Ala Val Ala Lys Glu Gln Ser Ile Ser Thr Lys Pro
    610                 615                 620

Pro Tyr Lys Glu Arg Lys Asp Glu Leu Lys
625                 630


<210> SEQ ID NO 2
<211> LENGTH: 386
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 2

Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1               5                   10                  15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
            20                  25                  30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
        35                  40                  45
```

-continued

```
Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
    50                  55                  60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65                  70                  75                  80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
                85                  90                  95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
            100                 105                 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115                 120                 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
    130                 135                 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145                 150                 155                 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
                165                 170                 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
            180                 185                 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
        195                 200                 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
    210                 215                 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225                 230                 235                 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
                245                 250                 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala Leu
            260                 265                 270

Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg Ser
        275                 280                 285

Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln Ala
    290                 295                 300

Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser His
305                 310                 315                 320

Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu Val
                325                 330                 335

Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro Gly
            340                 345                 350

Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp Tyr
        355                 360                 365

Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly Ala
    370                 375                 380

Gln Ala
385


<210> SEQ ID NO 3
<211> LENGTH: 266
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide
```

-continued

<400> SEQUENCE: 3

Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1 5 10 15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
20 25 30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
35 40 45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
50 55 60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65 70 75 80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
85 90 95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
100 105 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
115 120 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
130 135 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145 150 155 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
165 170 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
180 185 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
195 200 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
210 215 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225 230 235 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
245 250 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly
260 265

<210> SEQ ID NO 4
<211> LENGTH: 386
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 4

Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1 5 10 15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
20 25 30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
35 40 45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
50 55 60

-continued

```
Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65                  70                  75                  80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
                85                  90                  95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
            100                 105                 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115                 120                 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
    130                 135                 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145                 150                 155                 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
                165                 170                 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
            180                 185                 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
        195                 200                 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
    210                 215                 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225                 230                 235                 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
                245                 250                 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala Leu
            260                 265                 270

Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg Ser
        275                 280                 285

Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln Ala
    290                 295                 300

Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser His
305                 310                 315                 320

Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu Val
                325                 330                 335

Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro Gly
            340                 345                 350

Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp Tyr
        355                 360                 365

Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly Ala
    370                 375                 380

Gln Ala
385


<210> SEQ ID NO 5
<211> LENGTH: 266
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 5

Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1               5                   10                  15
```

-continued

```
Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
            20                  25                  30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
        35                  40                  45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
    50                  55                  60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65                  70                  75                  80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
                85                  90                  95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
            100                 105                 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115                 120                 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
    130                 135                 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145                 150                 155                 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
                165                 170                 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
            180                 185                 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
        195                 200                 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
    210                 215                 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225                 230                 235                 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
                245                 250                 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly
            260                 265


<210> SEQ ID NO 6
<211> LENGTH: 387
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 6

Met Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1               5                   10                  15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
            20                  25                  30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
        35                  40                  45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
    50                  55                  60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65                  70                  75                  80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
```

-continued

```
                85                  90                  95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
            100                 105                 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
        115                 120                 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
    130                 135                 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145                 150                 155                 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
                165                 170                 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala
            260                 265                 270

Leu Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg
        275                 280                 285

Ser Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln
    290                 295                 300

Ala Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser
305                 310                 315                 320

His Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu
                325                 330                 335

Val Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro
            340                 345                 350

Gly Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp
        355                 360                 365

Tyr Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly
    370                 375                 380

Ala Gln Ala
385


<210> SEQ ID NO 7
<211> LENGTH: 267
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 7

Met Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1               5                   10                  15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
            20                  25                  30
```

-continued

```
Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
        35                  40                  45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
    50                  55                  60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65                  70                  75                  80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
                85                  90                  95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
            100                 105                 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
        115                 120                 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
    130                 135                 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145                 150                 155                 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
                165                 170                 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly
            260                 265


<210> SEQ ID NO 8
<211> LENGTH: 387
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 8

Met Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1               5                   10                  15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
            20                  25                  30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
        35                  40                  45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
    50                  55                  60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65                  70                  75                  80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
                85                  90                  95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
            100                 105                 110
```

-continued

```
Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
        115                 120                 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
    130                 135                 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145                 150                 155                 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
                165                 170                 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala
            260                 265                 270

Leu Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg
        275                 280                 285

Ser Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln
    290                 295                 300

Ala Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser
305                 310                 315                 320

His Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu
                325                 330                 335

Val Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro
            340                 345                 350

Gly Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp
        355                 360                 365

Tyr Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly
    370                 375                 380

Ala Gln Ala
385


<210> SEQ ID NO 9
<211> LENGTH: 267
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 9

Met Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1               5                   10                  15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
            20                  25                  30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
        35                  40                  45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
```

-continued

```
    50                  55                  60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65                  70                  75                  80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
                85                  90                  95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
            100                 105                 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
        115                 120                 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
    130                 135                 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145                 150                 155                 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
                165                 170                 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly
            260                 265


<210> SEQ ID NO 10
<211> LENGTH: 179
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 10

Met Ala Ala Leu Gln Lys Ser Val Ser Ser Phe Leu Met Gly Thr Leu
1               5                   10                  15

Ala Thr Ser Cys Leu Leu Leu Leu Ala Leu Leu Val Gln Gly Gly Ala
            20                  25                  30

Ala Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser Asn Phe Gln
        35                  40                  45

Gln Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys Glu Ala Ser
    50                  55                  60

Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu Phe
65                  70                  75                  80

His Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys Gln Val Leu
                85                  90                  95

Asn Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp Arg Phe Gln
            100                 105                 110

Pro Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu Ser Asn Arg
        115                 120                 125
```

-continued

```
Leu Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile Gln Arg Asn
    130                 135                 140

Val Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu Ser Gly Glu
145                 150                 155                 160

Ile Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg Asn
                165                 170                 175

Ala Cys Ile
```

<210> SEQ ID NO 11
<211> LENGTH: 146
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 11

```
Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser Asn Phe Gln Gln
1               5                   10                  15

Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys Glu Ala Ser Leu
            20                  25                  30

Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu Phe His
        35                  40                  45

Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys Gln Val Leu Asn
    50                  55                  60

Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp Arg Phe Gln Pro
65                  70                  75                  80

Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu Ser Asn Arg Leu
                85                  90                  95

Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile Gln Arg Asn Val
            100                 105                 110

Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu Ser Gly Glu Ile
        115                 120                 125

Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg Asn Ala
    130                 135                 140

Cys Ile
145
```

<210> SEQ ID NO 12
<211> LENGTH: 147
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 12

```
Met Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser Asn Phe Gln
1               5                   10                  15

Gln Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys Glu Ala Ser
            20                  25                  30

Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu Phe
        35                  40                  45

His Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys Gln Val Leu
    50                  55                  60

Asn Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp Arg Phe Gln
```

-continued

65 70 75 80

Pro Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu Ser Asn Arg
85 90 95

Leu Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile Gln Arg Asn
100 105 110

Val Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu Ser Gly Glu
115 120 125

Ile Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg Asn
130 135 140

Ala Cys Ile
145

<210> SEQ ID NO 13
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 13

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15

<210> SEQ ID NO 14
<211> LENGTH: 548
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 14

Met Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1 5 10 15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
20 25 30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
35 40 45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
50 55 60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65 70 75 80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
85 90 95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
100 105 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
115 120 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
130 135 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145 150 155 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
165 170 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu

-continued

```
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala
            260                 265                 270

Leu Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg
        275                 280                 285

Ser Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln
    290                 295                 300

Ala Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser
305                 310                 315                 320

His Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu
                325                 330                 335

Val Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro
            340                 345                 350

Gly Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp
        355                 360                 365

Tyr Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly
    370                 375                 380

Ala Gln Ala Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
385                 390                 395                 400

Gly Ser Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser Asn Phe
                405                 410                 415

Gln Gln Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys Glu Ala
            420                 425                 430

Ser Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu
        435                 440                 445

Phe His Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys Gln Val
    450                 455                 460

Leu Asn Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp Arg Phe
465                 470                 475                 480

Gln Pro Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu Ser Asn
                485                 490                 495

Arg Leu Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile Gln Arg
            500                 505                 510

Asn Val Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu Ser Gly
        515                 520                 525

Glu Ile Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg
    530                 535                 540

Asn Ala Cys Ile
545
```

<210> SEQ ID NO 15
<211> LENGTH: 428
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 15

Met Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1 5 10 15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
20 25 30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
35 40 45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
50 55 60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65 70 75 80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
85 90 95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
100 105 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
115 120 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
130 135 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145 150 155 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
165 170 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
180 185 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
195 200 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
210 215 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225 230 235 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
245 250 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Gly Gly Gly Gly Ser
260 265 270

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Pro Ile Ser Ser His
275 280 285

Cys Arg Leu Asp Lys Ser Asn Phe Gln Gln Pro Tyr Ile Thr Asn Arg
290 295 300

Thr Phe Met Leu Ala Lys Glu Ala Ser Leu Ala Asp Asn Asn Thr Asp
305 310 315 320

Val Arg Leu Ile Gly Glu Lys Leu Phe His Gly Val Ser Met Ser Glu
325 330 335

Arg Cys Tyr Leu Met Lys Gln Val Leu Asn Phe Thr Leu Glu Glu Val
340 345 350

Leu Phe Pro Gln Ser Asp Arg Phe Gln Pro Tyr Met Gln Glu Val Val
355 360 365

Pro Phe Leu Ala Arg Leu Ser Asn Arg Leu Ser Thr Cys His Ile Glu

-continued

```
    370                 375                 380

Gly Asp Asp Leu His Ile Gln Arg Asn Val Gln Lys Leu Lys Asp Thr
385                 390                 395                 400

Val Lys Lys Leu Gly Glu Ser Gly Glu Ile Lys Ala Ile Gly Glu Leu
                405                 410                 415

Asp Leu Leu Phe Met Ser Leu Arg Asn Ala Cys Ile
            420                 425


<210> SEQ ID NO 16
<211> LENGTH: 548
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 16

Met Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1               5                   10                  15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
            20                  25                  30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
        35                  40                  45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
    50                  55                  60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65                  70                  75                  80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
                85                  90                  95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
            100                 105                 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
        115                 120                 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
    130                 135                 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145                 150                 155                 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
                165                 170                 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala
            260                 265                 270

Leu Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg
        275                 280                 285
```

-continued

```
Ser Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln
    290                 295                 300

Ala Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser
305                 310                 315                 320

His Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu
                325                 330                 335

Val Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro
            340                 345                 350

Gly Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp
        355                 360                 365

Tyr Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly
    370                 375                 380

Ala Gln Ala Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
385                 390                 395                 400

Gly Ser Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser Asn Phe
                405                 410                 415

Gln Gln Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys Glu Ala
            420                 425                 430

Ser Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu
        435                 440                 445

Phe His Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys Gln Val
    450                 455                 460

Leu Asn Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp Arg Phe
465                 470                 475                 480

Gln Pro Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu Ser Asn
                485                 490                 495

Arg Leu Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile Gln Arg
            500                 505                 510

Asn Val Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu Ser Gly
        515                 520                 525

Glu Ile Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg
    530                 535                 540

Asn Ala Cys Ile
545


<210> SEQ ID NO 17
<211> LENGTH: 428
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 17

Met Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1               5                   10                  15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
            20                  25                  30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
        35                  40                  45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
    50                  55                  60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65                  70                  75                  80
```

-continued

```
Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
                85                  90                  95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
            100                 105                 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
        115                 120                 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
    130                 135                 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145                 150                 155                 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
                165                 170                 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Gly Gly Gly Gly Ser
            260                 265                 270

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Pro Ile Ser Ser His
        275                 280                 285

Cys Arg Leu Asp Lys Ser Asn Phe Gln Gln Pro Tyr Ile Thr Asn Arg
    290                 295                 300

Thr Phe Met Leu Ala Lys Glu Ala Ser Leu Ala Asp Asn Asn Thr Asp
305                 310                 315                 320

Val Arg Leu Ile Gly Glu Lys Leu Phe His Gly Val Ser Met Ser Glu
                325                 330                 335

Arg Cys Tyr Leu Met Lys Gln Val Leu Asn Phe Thr Leu Glu Glu Val
            340                 345                 350

Leu Phe Pro Gln Ser Asp Arg Phe Gln Pro Tyr Met Gln Glu Val Val
        355                 360                 365

Pro Phe Leu Ala Arg Leu Ser Asn Arg Leu Ser Thr Cys His Ile Glu
    370                 375                 380

Gly Asp Asp Leu His Ile Gln Arg Asn Val Gln Lys Leu Lys Asp Thr
385                 390                 395                 400

Val Lys Lys Leu Gly Glu Ser Gly Glu Ile Lys Ala Ile Gly Glu Leu
                405                 410                 415

Asp Leu Leu Phe Met Ser Leu Arg Asn Ala Cys Ile
            420                 425


<210> SEQ ID NO 18
<211> LENGTH: 547
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide
```

-continued

<400> SEQUENCE: 18

```
Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1               5                   10                  15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
            20                  25                  30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
        35                  40                  45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
    50                  55                  60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65                  70                  75                  80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
                85                  90                  95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
            100                 105                 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115                 120                 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
    130                 135                 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145                 150                 155                 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
                165                 170                 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
            180                 185                 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
        195                 200                 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
    210                 215                 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225                 230                 235                 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
                245                 250                 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala Leu
            260                 265                 270

Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg Ser
        275                 280                 285

Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln Ala
    290                 295                 300

Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser His
305                 310                 315                 320

Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu Val
                325                 330                 335

Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro Gly
            340                 345                 350

Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp Tyr
        355                 360                 365

Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly Ala
    370                 375                 380

Gln Ala Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
385                 390                 395                 400
```

-continued

```
Ser Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser Asn Phe Gln
                405                 410                 415

Gln Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys Glu Ala Ser
            420                 425                 430

Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu Phe
        435                 440                 445

His Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys Gln Val Leu
    450                 455                 460

Asn Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp Arg Phe Gln
465                 470                 475                 480

Pro Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu Ser Asn Arg
                485                 490                 495

Leu Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile Gln Arg Asn
            500                 505                 510

Val Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu Ser Gly Glu
        515                 520                 525

Ile Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg Asn
    530                 535                 540

Ala Cys Ile
545


<210> SEQ ID NO 19
<211> LENGTH: 547
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 19

Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1               5                   10                  15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
            20                  25                  30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
        35                  40                  45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
    50                  55                  60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65                  70                  75                  80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
                85                  90                  95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
            100                 105                 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115                 120                 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
    130                 135                 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145                 150                 155                 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
                165                 170                 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
```

-continued

```
            180                 185                 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
        195                 200                 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
    210                 215                 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225                 230                 235                 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
                245                 250                 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala Leu
            260                 265                 270

Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg Ser
        275                 280                 285

Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln Ala
    290                 295                 300

Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser His
305                 310                 315                 320

Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu Val
                325                 330                 335

Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro Gly
            340                 345                 350

Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp Tyr
        355                 360                 365

Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly Ala
    370                 375                 380

Gln Ala Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
385                 390                 395                 400

Ser Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser Asn Phe Gln
                405                 410                 415

Gln Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys Glu Ala Ser
            420                 425                 430

Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu Phe
        435                 440                 445

His Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys Gln Val Leu
    450                 455                 460

Asn Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp Arg Phe Gln
465                 470                 475                 480

Pro Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu Ser Asn Arg
                485                 490                 495

Leu Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile Gln Arg Asn
            500                 505                 510

Val Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu Ser Gly Glu
        515                 520                 525

Ile Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg Asn
    530                 535                 540

Ala Cys Ile
545


<210> SEQ ID NO 20
<211> LENGTH: 427
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
```

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 20

Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1 5 10 15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
20 25 30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
35 40 45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
50 55 60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65 70 75 80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
85 90 95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
100 105 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
115 120 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
130 135 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145 150 155 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
165 170 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
180 185 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
195 200 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
210 215 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225 230 235 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
245 250 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Gly Gly Gly Gly Ser Gly
260 265 270

Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Pro Ile Ser Ser His Cys
275 280 285

Arg Leu Asp Lys Ser Asn Phe Gln Gln Pro Tyr Ile Thr Asn Arg Thr
290 295 300

Phe Met Leu Ala Lys Glu Ala Ser Leu Ala Asp Asn Asn Thr Asp Val
305 310 315 320

Arg Leu Ile Gly Glu Lys Leu Phe His Gly Val Ser Met Ser Glu Arg
325 330 335

Cys Tyr Leu Met Lys Gln Val Leu Asn Phe Thr Leu Glu Glu Val Leu
340 345 350

Phe Pro Gln Ser Asp Arg Phe Gln Pro Tyr Met Gln Glu Val Val Pro
355 360 365

Phe Leu Ala Arg Leu Ser Asn Arg Leu Ser Thr Cys His Ile Glu Gly

-continued

```
    370                 375                 380

Asp Asp Leu His Ile Gln Arg Asn Val Gln Lys Leu Lys Asp Thr Val
385                 390                 395                 400

Lys Lys Leu Gly Glu Ser Gly Glu Ile Lys Ala Ile Gly Glu Leu Asp
                405                 410                 415

Leu Leu Phe Met Ser Leu Arg Asn Ala Cys Ile
            420                 425


<210> SEQ ID NO 21
<211> LENGTH: 427
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 21

Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1               5                   10                  15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
            20                  25                  30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
        35                  40                  45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
    50                  55                  60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65                  70                  75                  80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
                85                  90                  95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
            100                 105                 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115                 120                 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
    130                 135                 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145                 150                 155                 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
                165                 170                 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
            180                 185                 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
        195                 200                 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
    210                 215                 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225                 230                 235                 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
                245                 250                 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Gly Gly Gly Gly Ser Gly
            260                 265                 270

Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Pro Ile Ser Ser His Cys
        275                 280                 285
```

-continued

```
Arg Leu Asp Lys Ser Asn Phe Gln Gln Pro Tyr Ile Thr Asn Arg Thr
    290                 295                 300

Phe Met Leu Ala Lys Glu Ala Ser Leu Ala Asp Asn Asn Thr Asp Val
305                 310                 315                 320

Arg Leu Ile Gly Glu Lys Leu Phe His Gly Val Ser Met Ser Glu Arg
                325                 330                 335

Cys Tyr Leu Met Lys Gln Val Leu Asn Phe Thr Leu Glu Glu Val Leu
            340                 345                 350

Phe Pro Gln Ser Asp Arg Phe Gln Pro Tyr Met Gln Glu Val Val Pro
        355                 360                 365

Phe Leu Ala Arg Leu Ser Asn Arg Leu Ser Thr Cys His Ile Glu Gly
    370                 375                 380

Asp Asp Leu His Ile Gln Arg Asn Val Gln Lys Leu Lys Asp Thr Val
385                 390                 395                 400

Lys Lys Leu Gly Glu Ser Gly Glu Ile Lys Ala Ile Gly Glu Leu Asp
                405                 410                 415

Leu Leu Phe Met Ser Leu Arg Asn Ala Cys Ile
            420                 425


<210> SEQ ID NO 22
<211> LENGTH: 634
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 22

Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1               5                   10                  15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
            20                  25                  30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
        35                  40                  45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
    50                  55                  60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65                  70                  75                  80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
                85                  90                  95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
            100                 105                 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115                 120                 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
    130                 135                 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145                 150                 155                 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
                165                 170                 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
            180                 185                 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
        195                 200                 205
```

-continued

```
Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
    210                 215                 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225                 230                 235                 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
                245                 250                 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala Leu
            260                 265                 270

Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg Ser
        275                 280                 285

Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln Ala
    290                 295                 300

Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser His
305                 310                 315                 320

Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu Val
                325                 330                 335

Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro Gly
            340                 345                 350

Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp Tyr
        355                 360                 365

Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly Ala
    370                 375                 380

Gln Ala Ala Asp Ile Leu Ser Leu Phe Cys Pro Asp Ala Asp Lys Ser
385                 390                 395                 400

Cys Val Ala Ser Asn Asn Asp Gln Ala Asn Ile Asn Ile Glu Ser Arg
                405                 410                 415

Ser Gly Arg Ser Tyr Leu Pro Glu Asn Arg Ala Val Ile Thr Pro Gln
            420                 425                 430

Gly Val Thr Asn Trp Thr Tyr Gln Glu Leu Glu Ala Thr His Gln Ala
        435                 440                 445

Leu Thr Arg Glu Gly Tyr Val Phe Val Gly Tyr His Gly Thr Asn His
    450                 455                 460

Val Ala Ala Gln Thr Ile Val Asn Arg Ile Ala Pro Val Pro Arg Gly
465                 470                 475                 480

Asn Asn Thr Glu Asn Glu Glu Lys Trp Gly Gly Leu Tyr Val Ala Thr
                485                 490                 495

His Ala Glu Val Ala His Gly Tyr Ala Arg Ile Lys Glu Gly Thr Gly
            500                 505                 510

Glu Tyr Gly Leu Pro Thr Arg Ala Glu Arg Asp Ala Arg Gly Val Met
        515                 520                 525

Leu Arg Val Tyr Ile Pro Arg Ala Ser Leu Glu Arg Phe Tyr Arg Thr
    530                 535                 540

Asn Thr Pro Leu Glu Asn Ala Glu Glu His Ile Thr Gln Val Ile Gly
545                 550                 555                 560

His Ser Leu Pro Leu Arg Asn Glu Ala Phe Thr Gly Pro Glu Ser Ala
                565                 570                 575

Gly Gly Glu Asp Glu Thr Val Ile Gly Trp Asp Met Ala Ile His Ala
            580                 585                 590

Val Ala Ile Pro Ser Thr Ile Pro Gly Asn Ala Tyr Glu Glu Leu Ala
        595                 600                 605
```

-continued

```
Ile Asp Glu Glu Ala Val Ala Lys Glu Gln Ser Ile Ser Thr Lys Pro
    610                 615                 620

Pro Tyr Lys Glu Arg Lys Asp Glu Leu Lys
625                 630


<210> SEQ ID NO 23
<211> LENGTH: 635
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 23

Met Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1               5                   10                  15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
            20                  25                  30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
        35                  40                  45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
    50                  55                  60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65                  70                  75                  80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
                85                  90                  95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
            100                 105                 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
        115                 120                 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
    130                 135                 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145                 150                 155                 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
                165                 170                 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala
            260                 265                 270

Leu Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg
        275                 280                 285

Ser Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln
    290                 295                 300

Ala Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser
305                 310                 315                 320
```

-continued

```
His Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu
                325                 330                 335

Val Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro
            340                 345                 350

Gly Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp
        355                 360                 365

Tyr Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly
    370                 375                 380

Ala Gln Ala Ala Asp Ile Leu Ser Leu Phe Cys Pro Asp Ala Asp Lys
385                 390                 395                 400

Ser Cys Val Ala Ser Asn Asn Asp Gln Ala Asn Ile Asn Ile Glu Ser
                405                 410                 415

Arg Ser Gly Arg Ser Tyr Leu Pro Glu Asn Arg Ala Val Ile Thr Pro
            420                 425                 430

Gln Gly Val Thr Asn Trp Thr Tyr Gln Glu Leu Glu Ala Thr His Gln
        435                 440                 445

Ala Leu Thr Arg Glu Gly Tyr Val Phe Val Gly Tyr His Gly Thr Asn
    450                 455                 460

His Val Ala Ala Gln Thr Ile Val Asn Arg Ile Ala Pro Val Pro Arg
465                 470                 475                 480

Gly Asn Asn Thr Glu Asn Glu Glu Lys Trp Gly Gly Leu Tyr Val Ala
                485                 490                 495

Thr His Ala Glu Val Ala His Gly Tyr Ala Arg Ile Lys Glu Gly Thr
            500                 505                 510

Gly Glu Tyr Gly Leu Pro Thr Arg Ala Glu Arg Asp Ala Arg Gly Val
        515                 520                 525

Met Leu Arg Val Tyr Ile Pro Arg Ala Ser Leu Glu Arg Phe Tyr Arg
    530                 535                 540

Thr Asn Thr Pro Leu Glu Asn Ala Glu Glu His Ile Thr Gln Val Ile
545                 550                 555                 560

Gly His Ser Leu Pro Leu Arg Asn Glu Ala Phe Thr Gly Pro Glu Ser
                565                 570                 575

Ala Gly Gly Glu Asp Glu Thr Val Ile Gly Trp Asp Met Ala Ile His
            580                 585                 590

Ala Val Ala Ile Pro Ser Thr Ile Pro Gly Asn Ala Tyr Glu Glu Leu
        595                 600                 605

Ala Ile Asp Glu Glu Ala Val Ala Lys Glu Gln Ser Ile Ser Thr Lys
    610                 615                 620

Pro Pro Tyr Lys Glu Arg Lys Asp Glu Leu Lys
625                 630                 635


<210> SEQ ID NO 24
<211> LENGTH: 427
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 24

Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1               5                   10                  15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
```

-continued

```
            20                  25                  30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
        35                  40                  45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
    50                  55                  60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65                  70                  75                  80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
                85                  90                  95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
            100                 105                 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115                 120                 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
    130                 135                 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145                 150                 155                 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
                165                 170                 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
            180                 185                 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
        195                 200                 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
    210                 215                 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225                 230                 235                 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
                245                 250                 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Gly Gly Gly Gly Ser Gly
            260                 265                 270

Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Pro Ile Ser Ser His Cys
        275                 280                 285

Arg Leu Asp Lys Ser Asn Phe Gln Gln Pro Tyr Ile Thr Asn Arg Thr
    290                 295                 300

Phe Met Leu Ala Lys Glu Ala Ser Leu Ala Asp Asn Asn Thr Asp Val
305                 310                 315                 320

Arg Leu Ile Gly Glu Lys Leu Phe His Gly Val Ser Met Ser Glu Arg
                325                 330                 335

Cys Tyr Leu Met Lys Gln Val Leu Asn Phe Thr Leu Glu Glu Val Leu
            340                 345                 350

Phe Pro Gln Ser Asp Arg Phe Gln Pro Tyr Met Gln Glu Val Val Pro
        355                 360                 365

Phe Leu Ala Arg Leu Ser Asn Arg Leu Ser Thr Cys His Ile Glu Gly
    370                 375                 380

Asp Asp Leu His Ile Gln Arg Asn Val Gln Lys Leu Lys Asp Thr Val
385                 390                 395                 400

Lys Lys Leu Gly Glu Ser Gly Glu Ile Lys Ala Ile Gly Glu Leu Asp
                405                 410                 415

Leu Leu Phe Met Ser Leu Arg Asn Ala Cys Ile
            420                 425
```

-continued

<210> SEQ ID NO 25
<211> LENGTH: 427
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 25

Leu Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys Ser
1 5 10 15

Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile Pro
20 25 30

Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
35 40 45

Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser Ile
50 55 60

Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr Val
65 70 75 80

Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr Glu
85 90 95

Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe Ala
100 105 110

Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile Lys
115 120 125

Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val Pro
130 135 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Lys
145 150 155 160

Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn Ile
165 170 175

Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu Gly
180 185 190

Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu Cys
195 200 205

Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln Asn
210 215 220

Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val Ala
225 230 235 240

Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys Pro
245 250 255

Val Glu Gln Arg Ile His Phe Ser Lys Gly Gly Gly Gly Gly Ser Gly
260 265 270

Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Pro Ile Ser Ser His Cys
275 280 285

Arg Leu Asp Lys Ser Asn Phe Gln Gln Pro Tyr Ile Thr Asn Arg Thr
290 295 300

Phe Met Leu Ala Lys Glu Ala Ser Leu Ala Asp Asn Asn Thr Asp Val
305 310 315 320

Arg Leu Ile Gly Glu Lys Leu Phe His Gly Val Ser Met Ser Glu Arg
325 330 335

Cys Tyr Leu Met Lys Gln Val Leu Asn Phe Thr Leu Glu Glu Val Leu

-continued

```
            340                 345                 350

Phe Pro Gln Ser Asp Arg Phe Gln Pro Tyr Met Gln Glu Val Val Pro
        355                 360                 365

Phe Leu Ala Arg Leu Ser Asn Arg Leu Ser Thr Cys His Ile Glu Gly
    370                 375                 380

Asp Asp Leu His Ile Gln Arg Asn Val Gln Lys Leu Lys Asp Thr Val
385                 390                 395                 400

Lys Lys Leu Gly Glu Ser Gly Glu Ile Lys Ala Ile Gly Glu Leu Asp
                405                 410                 415

Leu Leu Phe Met Ser Leu Arg Asn Ala Cys Ile
            420                 425


<210> SEQ ID NO 26
<211> LENGTH: 649
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (1)..(1)
<223> OTHER INFORMATION: V or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)..(3)
<223> OTHER INFORMATION: E or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: A or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)..(6)
<223> OTHER INFORMATION: N or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (16)..(16)
<223> OTHER INFORMATION: S or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (21)..(21)
<223> OTHER INFORMATION: P or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (24)..(24)
<223> OTHER INFORMATION: P or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (30)..(30)
<223> OTHER INFORMATION: S or F
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (33)..(33)
<223> OTHER INFORMATION: S or G
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (56)..(56)
<223> OTHER INFORMATION: K or M
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (59)..(59)
<223> OTHER INFORMATION: D or G
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (67)..(67)
<223> OTHER INFORMATION: I or F
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (73)..(73)
```

-continued

<223> OTHER INFORMATION: V or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (81)..(81)
<223> OTHER INFORMATION: N or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (90)..(90)
<223> OTHER INFORMATION: H or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (101)..(101)
<223> OTHER INFORMATION: T or M
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (104)..(104)
<223> OTHER INFORMATION: Y or F
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (108)..(108)
<223> OTHER INFORMATION: E or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (109)..(109)
<223> OTHER INFORMATION: G or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (112)..(112)
<223> OTHER INFORMATION: A or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (114)..(114)
<223> OTHER INFORMATION: N or H
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (118)..(118)
<223> OTHER INFORMATION: P or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (119)..(119)
<223> OTHER INFORMATION: I or P
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (131)..(131)
<223> OTHER INFORMATION: V or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (134)..(134)
<223> OTHER INFORMATION: L or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (137)..(137)
<223> OTHER INFORMATION: Q or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (160)..(160)
<223> OTHER INFORMATION: K or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (161)..(161)
<223> OTHER INFORMATION: T or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (166)..(166)
<223> OTHER INFORMATION: S or F
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (168)..(168)
<223> OTHER INFORMATION: S or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (174)..(174)
<223> OTHER INFORMATION: H or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (175)..(185)

-continued

<223> OTHER INFORMATION: This region may encompass one of the following
sequences: "N", "S", "SIAKQS", or "SIAKQSIAKQS"
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (196)..(196)
<223> OTHER INFORMATION: K or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (199)..(199)
<223> OTHER INFORMATION: Q, E, or H
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (201)..(201)
<223> OTHER INFORMATION: E, N, or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (203)..(203)
<223> OTHER INFORMATION: S or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (210)..(210)
<223> OTHER INFORMATION: H or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (212)..(212)
<223> OTHER INFORMATION: H, L, F, or R
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (214)..(214)
<223> OTHER INFORMATION: G or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (215)..(215)
<223> OTHER INFORMATION: L or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (216)..(216)
<223> OTHER INFORMATION: A or P
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (217)..(217)
<223> OTHER INFORMATION: L, E, or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (218)..(218)
<223> OTHER INFORMATION: C or V
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (219)..(219)
<223> OTHER INFORMATION: W, V, or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (221)..(221)
<223> OTHER INFORMATION: V or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (222)..(222)
<223> OTHER INFORMATION: P or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (223)..(223)
<223> OTHER INFORMATION: M, I, L, or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (224)..(224)
<223> OTHER INFORMATION: D or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (225)..(225)
<223> OTHER INFORMATION: A or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (226)..(226)
<223> OTHER INFORMATION: I or absent
<220> FEATURE:
<221> NAME/KEY: MOD_RES

-continued

<222> LOCATION: (227)..(227)
<223> OTHER INFORMATION: Y or C
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (228)..(228)
<223> OTHER INFORMATION: N or F
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (229)..(229)
<223> OTHER INFORMATION: Y or F
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (230)..(230)
<223> OTHER INFORMATION: I or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (231)..(231)
<223> OTHER INFORMATION: T or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (232)..(232)
<223> OTHER INFORMATION: Q or P
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (233)..(233)
<223> OTHER INFORMATION: Q, E, or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (234)..(234)
<223> OTHER INFORMATION: N, L, or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (237)..(237)
<223> OTHER INFORMATION: L or Y
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (239)..(239)
<223> OTHER INFORMATION: D or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (240)..(240)
<223> OTHER INFORMATION: N or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (242)..(242)
<223> OTHER INFORMATION: F, H, or Y
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (245)..(245)
<223> OTHER INFORMATION: S or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (247)..(247)
<223> OTHER INFORMATION: E or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (252)..(252)
<223> OTHER INFORMATION: T or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (254)..(254)
<223> OTHER INFORMATION: K, E, or G
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (255)..(255)
<223> OTHER INFORMATION: V or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (257)..(257)
<223> OTHER INFORMATION: T or M
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (262)..(262)
<223> OTHER INFORMATION: I or M
<220> FEATURE:
<221> NAME/KEY: MOD_RES

-continued

<222> LOCATION: (266)..(266)
<223> OTHER INFORMATION: P, T, or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (275)..(275)
<223> OTHER INFORMATION: K, Q, or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (276)..(276)
<223> OTHER INFORMATION: G or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (279)..(279)
<223> OTHER INFORMATION: M or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (280)..(280)
<223> OTHER INFORMATION: S or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (281)..(281)
<223> OTHER INFORMATION: A or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (298)..(298)
<223> OTHER INFORMATION: S or G
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (303)..(303)
<223> OTHER INFORMATION: D or Y
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (305)..(305)
<223> OTHER INFORMATION: T, P, or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (309)..(309)
<223> OTHER INFORMATION: S or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (311)..(311)
<223> OTHER INFORMATION: A or V
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (313)..(313)
<223> OTHER INFORMATION: Q or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (316)..(316)
<223> OTHER INFORMATION: N or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (322)..(322)
<223> OTHER INFORMATION: V or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (326)..(326)
<223> OTHER INFORMATION: I or M
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (329)..(329)
<223> OTHER INFORMATION: S or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (331)..(331)
<223> OTHER INFORMATION: L or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (334)..(334)
<223> OTHER INFORMATION: V or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (340)..(340)
<223> OTHER INFORMATION: D, E, or H
<220> FEATURE:
<221> NAME/KEY: MOD_RES

-continued

<222> LOCATION: (341)..(341)
<223> OTHER INFORMATION: E or G
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (343)..(343)
<223> OTHER INFORMATION: E or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (345)..(345)
<223> OTHER INFORMATION: E or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (347)..(347)
<223> OTHER INFORMATION: A or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (351)..(351)
<223> OTHER INFORMATION: S, D, or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (352)..(352)
<223> OTHER INFORMATION: D or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (353)..(353)
<223> OTHER INFORMATION: L or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (355)..(355)
<223> OTHER INFORMATION: R or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (359)..(359)
<223> OTHER INFORMATION: N or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (363)..(363)
<223> OTHER INFORMATION: M or V
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (365)..(365)
<223> OTHER INFORMATION: T or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (370)..(370)
<223> OTHER INFORMATION: V or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (381)..(381)
<223> OTHER INFORMATION: H or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (384)..(384)
<223> OTHER INFORMATION: G or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (386)..(386)
<223> OTHER INFORMATION: T or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (393)..(393)
<223> OTHER INFORMATION: G or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (403)..(403)
<223> OTHER INFORMATION: F or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (404)..(404)
<223> OTHER INFORMATION: C or Y
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (407)..(407)
<223> OTHER INFORMATION: A or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES

-continued

<222> LOCATION: (409)..(409)
<223> OTHER INFORMATION: K, E, or G
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (410)..(410)
<223> OTHER INFORMATION: S, P, or H
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (414)..(414)
<223> OTHER INFORMATION: S or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (415)..(415)
<223> OTHER INFORMATION: N or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (416)..(416)
<223> OTHER INFORMATION: N or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (423)..(423)
<223> OTHER INFORMATION: I or V
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (433)..(433)
<223> OTHER INFORMATION: P or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (441)..(441)
<223> OTHER INFORMATION: P or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (453)..(453)
<223> OTHER INFORMATION: E or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (454)..(454)
<223> OTHER INFORMATION: A or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (455)..(455)
<223> OTHER INFORMATION: T or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (458)..(458)
<223> OTHER INFORMATION: A or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (461)..(461)
<223> OTHER INFORMATION: R or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (463)..(463)
<223> OTHER INFORMATION: G or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (475)..(475)
<223> OTHER INFORMATION: V or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (479)..(479)
<223> OTHER INFORMATION: T, S, or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (485)..(485)
<223> OTHER INFORMATION: A, S, or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (491)..(491)
<223> OTHER INFORMATION: N or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (492)..(492)
<223> OTHER INFORMATION: N or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES

-continued

<222> LOCATION: (495)..(495)
<223> OTHER INFORMATION: N, S, or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (497)..(497)
<223> OTHER INFORMATION: E, R, or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (498)..(498)
<223> OTHER INFORMATION: K, A, or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (502)..(502)
<223> OTHER INFORMATION: L or V
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (505)..(505)
<223> OTHER INFORMATION: A or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (507)..(507)
<223> OTHER INFORMATION: H or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (509)..(509)
<223> OTHER INFORMATION: E or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (510)..(510)
<223> OTHER INFORMATION: V or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (511)..(511)
<223> OTHER INFORMATION: A or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (512)..(512)
<223> OTHER INFORMATION: H or Y
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (513)..(513)
<223> OTHER INFORMATION: G or R
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (515)..(515)
<223> OTHER INFORMATION: A or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (517)..(517)
<223> OTHER INFORMATION: I or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (518)..(518)
<223> OTHER INFORMATION: K or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (519)..(519)
<223> OTHER INFORMATION: E or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (522)..(522)
<223> OTHER INFORMATION: G or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (523)..(523)
<223> OTHER INFORMATION: E, D, or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (524)..(524)
<223> OTHER INFORMATION: Y, G, A, or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (525)..(525)
<223> OTHER INFORMATION: G or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES

-continued

<222> LOCATION: (526)..(526)
<223> OTHER INFORMATION: L or G
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (527)..(527)
<223> OTHER INFORMATION: P or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (529)..(529)
<223> OTHER INFORMATION: R, P, or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (530)..(530)
<223> OTHER INFORMATION: A or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (531)..(531)
<223> OTHER INFORMATION: E or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (532)..(532)
<223> OTHER INFORMATION: R, Q, or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (533)..(533)
<223> OTHER INFORMATION: D, K, or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (534)..(534)
<223> OTHER INFORMATION: A, T, or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (540)..(540)
<223> OTHER INFORMATION: R or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (543)..(543)
<223> OTHER INFORMATION: I or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (544)..(544)
<223> OTHER INFORMATION: P or H
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (545)..(545)
<223> OTHER INFORMATION: R or Q
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (554)..(554)
<223> OTHER INFORMATION: T or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (556)..(556)
<223> OTHER INFORMATION: T, A, or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (557)..(557)
<223> OTHER INFORMATION: P or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (560)..(560)
<223> OTHER INFORMATION: N or K
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (561)..(561)
<223> OTHER INFORMATION: A or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (562)..(562)
<223> OTHER INFORMATION: E, R, or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (563)..(563)
<223> OTHER INFORMATION: E, N, or R
<220> FEATURE:
<221> NAME/KEY: MOD_RES

-continued

<222> LOCATION: (564)..(564)
<223> OTHER INFORMATION: H or L
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (565)..(565)
<223> OTHER INFORMATION: I or V
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (566)..(566)
<223> OTHER INFORMATION: T or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (567)..(567)
<223> OTHER INFORMATION: Q, R, H, or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (572)..(572)
<223> OTHER INFORMATION: S or P
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (583)..(583)
<223> OTHER INFORMATION: P or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (584)..(584)
<223> OTHER INFORMATION: E or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (585)..(585)
<223> OTHER INFORMATION: S, A, or R
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (586)..(586)
<223> OTHER INFORMATION: A, E, or V
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (587)..(587)
<223> OTHER INFORMATION: G, E, or D
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (589)..(589)
<223> OTHER INFORMATION: E or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (590)..(590)
<223> OTHER INFORMATION: D or N
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (593)..(593)
<223> OTHER INFORMATION: V or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (598)..(598)
<223> OTHER INFORMATION: M or I
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (601)..(601)
<223> OTHER INFORMATION: H or Y
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (602)..(602)
<223> OTHER INFORMATION: A or G
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (613)..(613)
<223> OTHER INFORMATION: A or S
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (615)..(615)
<223> OTHER INFORMATION: E or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (616)..(616)
<223> OTHER INFORMATION: E, A, Q, G, V, or R
<220> FEATURE:
<221> NAME/KEY: MOD_RES

-continued

<222> LOCATION: (618)..(618)
<223> OTHER INFORMATION: A, P, or T
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (619)..(619)
<223> OTHER INFORMATION: I, T, or P
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (620)..(620)
<223> OTHER INFORMATION: D or A
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (624)..(629)
<223> OTHER INFORMATION: This region may encompass one of the following sequences: "V" or "VVKEAI"
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (631)..(631)
<223> OTHER INFORMATION: K or E
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (637)..(637)
<223> OTHER INFORMATION: T, A, or P
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (644)..(644)
<223> OTHER INFORMATION: R, Q, or H
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (645)..(645)
<223> OTHER INFORMATION: K or absent
<220> FEATURE:
<223> OTHER INFORMATION: See specification as filed for detailed description of substitutions and preferred embodiments

<400> SEQUENCE: 26

Xaa Glu Xaa Xaa Leu Xaa Ile Phe Asp Glu Cys Arg Ser Pro Cys Xaa
1 5 10 15

Leu Thr Pro Glu Xaa Gly Lys Xaa Ile Gln Ser Lys Leu Xaa Ile Pro
            20 25 30

Xaa Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr Ile
        35 40 45

Asn Asp Glu Gln Asn Asp Ile Xaa Asp Glu Xaa Lys Gly Glu Ser Ile
    50 55 60

Ile Thr Xaa Gly Glu Phe Ala Thr Xaa Arg Ala Thr Arg His Tyr Val
65 70 75 80

Xaa Gln Asp Ala Pro Phe Gly Val Ile Xaa Leu Asp Ile Thr Thr Glu
                85 90 95

Asn Gly Thr Lys Xaa Tyr Ser Xaa Asn Arg Lys Xaa Xaa Glu Phe Xaa
            100 105 110

Ile Xaa Trp Leu Val Xaa Xaa Gly Glu Asp Ser Pro Ala Ser Ile Lys
        115 120 125

Ile Ser Xaa Asp Glu Xaa Asp Gln Xaa Arg Asn Ile Ile Glu Val Pro
    130 135 140

Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp Xaa
145 150 155 160

Xaa Gln Gly Asn Val Xaa Phe Xaa Val Thr Arg Pro Glu Xaa Xaa Xaa
                165 170 175

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ile Ala Ile Ser Trp Pro Ser
            180 185 190

Val Ser Tyr Xaa Ala Ala Xaa Lys Xaa Gly Xaa Arg His Lys Arg Trp
        195 200 205

-continued

```
Ala Xaa Trp Xaa Thr Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa
    210                 215                 220

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Thr Xaa Gly Xaa Xaa
225                 230                 235                 240

Trp Xaa Gly Gly Xaa Tyr Xaa Thr Val Ala Gly Xaa Pro Xaa Xaa Ile
                245                 250                 255

Xaa Val Lys Gln Gly Xaa Glu Gln Lys Xaa Val Glu Gln Arg Ile His
            260                 265                 270

Phe Ser Xaa Xaa Asn Ala Xaa Xaa Xaa Leu Ala Ala His Arg Val Cys
        275                 280                 285

Gly Val Pro Leu Glu Thr Leu Ala Arg Xaa Arg Lys Pro Arg Xaa Leu
    290                 295                 300

Xaa Asp Asp Leu Xaa Cys Xaa Tyr Xaa Ala Gln Xaa Ile Val Ser Leu
305                 310                 315                 320

Phe Xaa Ala Thr Arg Xaa Leu Phe Xaa His Xaa Asp Ser Xaa Phe Thr
                325                 330                 335

Leu Asn Leu Xaa Xaa Gln Xaa Pro Xaa Val Xaa Glu Arg Leu Xaa Xaa
            340                 345                 350

Xaa Arg Xaa Ile Asn Glu Xaa Asn Pro Gly Xaa Val Xaa Gln Val Leu
        355                 360                 365

Thr Xaa Ala Arg Gln Ile Tyr Asn Asp Tyr Val Thr Xaa His Pro Xaa
    370                 375                 380

Leu Xaa Pro Glu Gln Thr Ser Ala Xaa Ala Gln Ala Ala Asp Ile Leu
385                 390                 395                 400

Ser Leu Xaa Xaa Pro Asp Xaa Asp Xaa Xaa Cys Val Ala Xaa Xaa Xaa
                405                 410                 415

Asp Gln Ala Asn Ile Asn Xaa Glu Ser Arg Ser Gly Arg Ser Tyr Leu
            420                 425                 430

Xaa Glu Asn Arg Ala Val Ile Thr Xaa Gln Gly Val Thr Asn Trp Thr
        435                 440                 445

Tyr Gln Glu Leu Xaa Xaa Xaa His Gln Xaa Leu Thr Xaa Glu Xaa Tyr
    450                 455                 460

Val Phe Val Gly Tyr His Gly Thr Asn His Xaa Ala Ala Gln Xaa Ile
465                 470                 475                 480

Val Asn Arg Ile Xaa Pro Val Pro Arg Gly Xaa Xaa Thr Glu Xaa Glu
                485                 490                 495

Xaa Xaa Trp Gly Gly Xaa Tyr Val Xaa Thr Xaa Ala Xaa Xaa Xaa Xaa
            500                 505                 510

Xaa Tyr Xaa Arg Xaa Xaa Xaa Gly Thr Xaa Xaa Xaa Xaa Xaa Xaa Thr
        515                 520                 525

Xaa Xaa Xaa Xaa Xaa Xaa Arg Gly Val Met Leu Xaa Val Tyr Xaa Xaa
    530                 535                 540

Xaa Ala Ser Leu Glu Arg Phe Tyr Arg Xaa Asn Xaa Xaa Leu Glu Xaa
545                 550                 555                 560

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Val Ile Gly His Xaa Leu Pro Leu Arg
                565                 570                 575

Asn Glu Ala Phe Thr Gly Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa Glu Thr
            580                 585                 590

Xaa Ile Gly Trp Asp Xaa Ala Ile Xaa Xaa Val Ala Ile Pro Ser Thr
        595                 600                 605

Ile Pro Gly Asn Xaa Tyr Xaa Xaa Leu Xaa Xaa Xaa Glu Glu Ala Xaa
```

-continued

```
    610                 615                 620

Xaa Xaa Xaa Xaa Xaa Ala Xaa Glu Gln Ser Ile Ser Xaa Lys Pro Pro
625                 630                 635                 640

Tyr Lys Glu Xaa Xaa Asp Glu Leu Lys
                645


<210> SEQ ID NO 27
<211> LENGTH: 60
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (1)..(60)
<223> OTHER INFORMATION: This sequence may encompass 1-15 "Gly Gly Gly
      Ser" repeating units

<400> SEQUENCE: 27

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
1               5                   10                  15

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
            20                  25                  30

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
        35                  40                  45

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser
    50                  55                  60


<210> SEQ ID NO 28
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      peptide

<400> SEQUENCE: 28

Gly Gly Gly Gly Ser
1               5


<210> SEQ ID NO 29
<211> LENGTH: 90
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (1)..(90)
<223> OTHER INFORMATION: This sequence may encompass 1-15 "Gly Gly Gly
      Gly Gly Ser" repeating units

<400> SEQUENCE: 29

Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly
1               5                   10                  15

Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly
            20                  25                  30

Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser
        35                  40                  45

Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly
```

-continued

```
    50                  55                  60

Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly
65                  70                  75                  80

Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser
                85                  90


<210> SEQ ID NO 30
<211> LENGTH: 179
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 30

Met Ala Val Leu Gln Lys Ser Met Ser Phe Ser Leu Met Gly Thr Leu
1               5                   10                  15

Ala Ala Ser Cys Leu Leu Leu Ile Ala Leu Trp Ala Gln Glu Ala Asn
            20                  25                  30

Ala Leu Pro Val Asn Thr Arg Cys Lys Leu Glu Val Ser Asn Phe Gln
        35                  40                  45

Gln Pro Tyr Ile Val Asn Arg Thr Phe Met Leu Ala Lys Glu Ala Ser
    50                  55                  60

Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu Phe
65                  70                  75                  80

Arg Gly Val Ser Ala Lys Asp Gln Cys Tyr Leu Met Lys Gln Val Leu
                85                  90                  95

Asn Phe Thr Leu Glu Asp Val Leu Leu Pro Gln Ser Asp Arg Phe Gln
            100                 105                 110

Pro Tyr Met Gln Glu Val Val Pro Phe Leu Thr Lys Leu Ser Asn Gln
        115                 120                 125

Leu Ser Ser Cys His Ile Ser Gly Asp Asp Gln Asn Ile Gln Lys Asn
    130                 135                 140

Val Arg Arg Leu Lys Glu Thr Val Lys Lys Leu Gly Glu Ser Gly Glu
145                 150                 155                 160

Ile Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg Asn
                165                 170                 175

Ala Cys Val


<210> SEQ ID NO 31
<211> LENGTH: 25
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      peptide

<400> SEQUENCE: 31

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1               5                   10                  15

Gly Gly Gly Ser Gly Gly Gly Gly Ser
            20                  25


<210> SEQ ID NO 32
<211> LENGTH: 418
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
```

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 32

Met Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1 5 10 15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
20 25 30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
35 40 45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
50 55 60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65 70 75 80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
85 90 95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
100 105 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
115 120 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
130 135 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145 150 155 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
165 170 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
180 185 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
195 200 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
210 215 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225 230 235 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
245 250 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Gly Gly Gly Gly Ser
260 265 270

Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser Asn Phe Gln Gln
275 280 285

Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys Glu Ala Ser Leu
290 295 300

Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu Phe His
305 310 315 320

Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys Gln Val Leu Asn
325 330 335

Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp Arg Phe Gln Pro
340 345 350

Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu Ser Asn Arg Leu
355 360 365

Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile Gln Arg Asn Val

-continued

```
    370                 375                 380

Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu Ser Gly Glu Ile
385                 390                 395                 400

Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg Asn Ala
                405                 410                 415

Cys Ile


<210> SEQ ID NO 33
<211> LENGTH: 438
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 33

Met Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1               5                   10                  15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
            20                  25                  30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
        35                  40                  45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
    50                  55                  60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65                  70                  75                  80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
                85                  90                  95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
            100                 105                 110

Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
        115                 120                 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
    130                 135                 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145                 150                 155                 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
                165                 170                 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Gly Gly Gly Gly Ser
            260                 265                 270

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
        275                 280                 285

Gly Gly Gly Ser Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser
```

-continued

```
    290                 295                 300

Asn Phe Gln Gln Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys
305                 310                 315                 320

Glu Ala Ser Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu
                325                 330                 335

Lys Leu Phe His Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys
            340                 345                 350

Gln Val Leu Asn Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp
        355                 360                 365

Arg Phe Gln Pro Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu
    370                 375                 380

Ser Asn Arg Leu Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile
385                 390                 395                 400

Gln Arg Asn Val Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu
                405                 410                 415

Ser Gly Glu Ile Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser
            420                 425                 430

Leu Arg Asn Ala Cys Ile
        435


<210> SEQ ID NO 34
<211> LENGTH: 418
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 34

Met Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser Asn Phe Gln
1               5                   10                  15

Gln Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala Lys Glu Ala Ser
            20                  25                  30

Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile Gly Glu Lys Leu Phe
        35                  40                  45

His Gly Val Ser Met Ser Glu Arg Cys Tyr Leu Met Lys Gln Val Leu
    50                  55                  60

Asn Phe Thr Leu Glu Glu Val Leu Phe Pro Gln Ser Asp Arg Phe Gln
65                  70                  75                  80

Pro Tyr Met Gln Glu Val Val Pro Phe Leu Ala Arg Leu Ser Asn Arg
                85                  90                  95

Leu Ser Thr Cys His Ile Glu Gly Asp Asp Leu His Ile Gln Arg Asn
            100                 105                 110

Val Gln Lys Leu Lys Asp Thr Val Lys Lys Leu Gly Glu Ser Gly Glu
        115                 120                 125

Ile Lys Ala Ile Gly Glu Leu Asp Leu Leu Phe Met Ser Leu Arg Asn
    130                 135                 140

Ala Cys Ile Gly Gly Gly Gly Ser Val Glu Glu Ala Leu Asn Ile Phe
145                 150                 155                 160

Asp Glu Cys Arg Ser Pro Cys Ser Leu Thr Pro Glu Pro Gly Lys Pro
                165                 170                 175

Ile Gln Ser Lys Leu Ser Ile Pro Ser Asp Val Val Leu Asp Glu Gly
            180                 185                 190
```

-continued

```
Val Leu Tyr Tyr Ser Met Thr Ile Asn Asp Glu Gln Asn Asp Ile Lys
        195                 200                 205

Asp Glu Asp Lys Gly Glu Ser Ile Ile Thr Ile Gly Glu Phe Ala Thr
    210                 215                 220

Val Arg Ala Thr Arg His Tyr Val Asn Gln Asp Ala Pro Phe Gly Val
225                 230                 235                 240

Ile His Leu Asp Ile Thr Thr Glu Asn Gly Thr Lys Thr Tyr Ser Tyr
                245                 250                 255

Asn Arg Lys Glu Gly Glu Phe Ala Ile Asn Trp Leu Val Pro Ile Gly
            260                 265                 270

Glu Asp Ser Pro Ala Ser Ile Lys Ile Ser Val Asp Glu Leu Asp Gln
        275                 280                 285

Gln Arg Asn Ile Ile Glu Val Pro Lys Leu Tyr Ser Ile Asp Leu Asp
    290                 295                 300

Asn Gln Thr Leu Glu Gln Trp Lys Thr Gln Gly Asn Val Ser Phe Ser
305                 310                 315                 320

Val Thr Arg Pro Glu His Asn Ile Ala Ile Ser Trp Pro Ser Val Ser
                325                 330                 335

Tyr Lys Ala Ala Gln Lys Glu Gly Ser Arg His Lys Arg Trp Ala His
            340                 345                 350

Trp His Thr Gly Leu Ala Leu Cys Trp Leu Val Pro Met Asp Ala Ile
        355                 360                 365

Tyr Asn Tyr Ile Thr Gln Gln Asn Cys Thr Leu Gly Asp Asn Trp Phe
    370                 375                 380

Gly Gly Ser Tyr Glu Thr Val Ala Gly Thr Pro Lys Val Ile Thr Val
385                 390                 395                 400

Lys Gln Gly Ile Glu Gln Lys Pro Val Glu Gln Arg Ile His Phe Ser
                405                 410                 415

Lys Gly


<210> SEQ ID NO 35
<211> LENGTH: 538
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

<400> SEQUENCE: 35

Met Val Glu Glu Ala Leu Asn Ile Phe Asp Glu Cys Arg Ser Pro Cys
1               5                   10                  15

Ser Leu Thr Pro Glu Pro Gly Lys Pro Ile Gln Ser Lys Leu Ser Ile
            20                  25                  30

Pro Ser Asp Val Val Leu Asp Glu Gly Val Leu Tyr Tyr Ser Met Thr
        35                  40                  45

Ile Asn Asp Glu Gln Asn Asp Ile Lys Asp Glu Asp Lys Gly Glu Ser
    50                  55                  60

Ile Ile Thr Ile Gly Glu Phe Ala Thr Val Arg Ala Thr Arg His Tyr
65                  70                  75                  80

Val Asn Gln Asp Ala Pro Phe Gly Val Ile His Leu Asp Ile Thr Thr
                85                  90                  95

Glu Asn Gly Thr Lys Thr Tyr Ser Tyr Asn Arg Lys Glu Gly Glu Phe
            100                 105                 110
```

-continued

```
Ala Ile Asn Trp Leu Val Pro Ile Gly Glu Asp Ser Pro Ala Ser Ile
        115                 120                 125

Lys Ile Ser Val Asp Glu Leu Asp Gln Gln Arg Asn Ile Ile Glu Val
    130                 135                 140

Pro Lys Leu Tyr Ser Ile Asp Leu Asp Asn Gln Thr Leu Glu Gln Trp
145                 150                 155                 160

Lys Thr Gln Gly Asn Val Ser Phe Ser Val Thr Arg Pro Glu His Asn
                165                 170                 175

Ile Ala Ile Ser Trp Pro Ser Val Ser Tyr Lys Ala Ala Gln Lys Glu
            180                 185                 190

Gly Ser Arg His Lys Arg Trp Ala His Trp His Thr Gly Leu Ala Leu
        195                 200                 205

Cys Trp Leu Val Pro Met Asp Ala Ile Tyr Asn Tyr Ile Thr Gln Gln
    210                 215                 220

Asn Cys Thr Leu Gly Asp Asn Trp Phe Gly Gly Ser Tyr Glu Thr Val
225                 230                 235                 240

Ala Gly Thr Pro Lys Val Ile Thr Val Lys Gln Gly Ile Glu Gln Lys
                245                 250                 255

Pro Val Glu Gln Arg Ile His Phe Ser Lys Gly Asn Ala Met Ser Ala
            260                 265                 270

Leu Ala Ala His Arg Val Cys Gly Val Pro Leu Glu Thr Leu Ala Arg
        275                 280                 285

Ser Arg Lys Pro Arg Asp Leu Thr Asp Asp Leu Ser Cys Ala Tyr Gln
    290                 295                 300

Ala Gln Asn Ile Val Ser Leu Phe Val Ala Thr Arg Ile Leu Phe Ser
305                 310                 315                 320

His Leu Asp Ser Val Phe Thr Leu Asn Leu Asp Glu Gln Glu Pro Glu
                325                 330                 335

Val Ala Glu Arg Leu Ser Asp Leu Arg Arg Ile Asn Glu Asn Asn Pro
            340                 345                 350

Gly Met Val Thr Gln Val Leu Thr Val Ala Arg Gln Ile Tyr Asn Asp
        355                 360                 365

Tyr Val Thr His His Pro Gly Leu Thr Pro Glu Gln Thr Ser Ala Gly
    370                 375                 380

Ala Gln Ala Gly Gly Gly Gly Ser Ala Pro Ile Ser Ser His Cys Arg
385                 390                 395                 400

Leu Asp Lys Ser Asn Phe Gln Gln Pro Tyr Ile Thr Asn Arg Thr Phe
                405                 410                 415

Met Leu Ala Lys Glu Ala Ser Leu Ala Asp Asn Asn Thr Asp Val Arg
            420                 425                 430

Leu Ile Gly Glu Lys Leu Phe His Gly Val Ser Met Ser Glu Arg Cys
        435                 440                 445

Tyr Leu Met Lys Gln Val Leu Asn Phe Thr Leu Glu Glu Val Leu Phe
    450                 455                 460

Pro Gln Ser Asp Arg Phe Gln Pro Tyr Met Gln Glu Val Val Pro Phe
465                 470                 475                 480

Leu Ala Arg Leu Ser Asn Arg Leu Ser Thr Cys His Ile Glu Gly Asp
                485                 490                 495

Asp Leu His Ile Gln Arg Asn Val Gln Lys Leu Lys Asp Thr Val Lys
            500                 505                 510

Lys Leu Gly Glu Ser Gly Glu Ile Lys Ala Ile Gly Glu Leu Asp Leu
```

-continued

```
        515                 520                 525

Leu Phe Met Ser Leu Arg Asn Ala Cys Ile
    530                 535


<210> SEQ ID NO 36
<211> LENGTH: 75
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (1)..(75)
<223> OTHER INFORMATION: This sequence may encompass 1-15 "Gly Gly Gly
      Gly Ser" repeating units

<400> SEQUENCE: 36

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1               5                   10                  15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
            20                  25                  30

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
        35                  40                  45

Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
    50                  55                  60

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
65                  70                  75


<210> SEQ ID NO 37
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      peptide

<400> SEQUENCE: 37

Ser Ile Ala Lys Gln Ser
1               5


<210> SEQ ID NO 38
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      peptide

<400> SEQUENCE: 38

Ser Ile Ala Lys Gln Ser Ile Ala Lys Gln Ser
1               5                   10


<210> SEQ ID NO 39
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      peptide

<400> SEQUENCE: 39

Val Val Lys Glu Ala Ile
```

-continued

```
1               5


<210> SEQ ID NO 40
<211> LENGTH: 30
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (1)..(30)
<223> OTHER INFORMATION: This sequence may encompass 1-15 "Gly Ser"
      repeating units

<400> SEQUENCE: 40

Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser
1               5                   10                  15

Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser
            20                  25                  30


<210> SEQ ID NO 41
<211> LENGTH: 45
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (1)..(45)
<223> OTHER INFORMATION: This sequence may encompass 1-15 "Gly Gly Ser"
      repeating units

<400> SEQUENCE: 41

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly
1               5                   10                  15

Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly
            20                  25                  30

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser
        35                  40                  45
```

What is claimed is:

1. A delivery construct comprising an amino acid sequence set forth in SEQ ID NO: 15 or SEQ ID NO: 17.

2. The delivery construct of claim **1**, wherein the delivery construct comprises the amino acid sequence set forth in SEQ ID NO: 15.

3. The delivery construct of claim **1**, wherein the delivery construct comprises the amino acid sequence set forth in SEQ ID NO: 17.

4. A delivery construct comprising a carrier, wherein the carrier:

has its C-terminus at any one of positions 195 to 347;

has a glutamic acid at position 3 and an alanine at position 4; and

is capable of transcytosing across a polarized epithelial cell.

5. The delivery construct of claim **4**, wherein the carrier is 266 amino acids in length.

6. The delivery construct of claim **4**, wherein the carrier consists of an amino acid sequence set forth in SEQ ID NO: 7 or SEQ ID NO: 9, or an amino acid sequence having at least 90%, 95%, or 99% sequence identity thereto.

7. The delivery construct of claim **6**, wherein the carrier consists of an amino acid sequence set forth in SEQ ID NO: 7.

8. The delivery construct of claim **6**, wherein the carrier consists of the amino acid sequence set forth in SEQ ID NO: 9.

9. The delivery construct of claim **4**, wherein the carrier is coupled to a payload.

10. The delivery construct of claim **9**, wherein the payload is IL-22.

11. The delivery construct of claim **10**, wherein the IL-22 consists of an amino acid sequence set forth in SEQ ID NO: 11.

12. The delivery construct of claim **10**, wherein the carrier is coupled non-covalently to the IL-22.

13. The delivery construct of any one of claim **10**, wherein the carrier is coupled covalently to the IL-22.

14. The delivery construct of claim **13**, wherein the carrier is coupled covalently to the IL-22 via a spacer.

15. The delivery construct of claim **14**, wherein the spacer consists of an amino acid sequence set forth in SEQ ID NO: 13.

16. The delivery construct of claim **4**, wherein the delivery construct consists of an amino acid sequence set forth in SEQ ID NO: 15 or SEQ ID NO: 17.

17. A method of treating an inflammatory disease in a subject, the method comprising administering to the subject an effective amount of the delivery construct of claim **1**.

18. The method of claim **17**, wherein the inflammatory disease is hepatitis, obesity, fatty liver disease, liver inflammation, or pancreatitis, Crohn's disease, ulcerative colitis, pouchitis, proctitis, multiple sclerosis, systemic lupus erythematosus, graft versus host disease, rheumatoid arthritis, or psoriasis.

19. The method of claim **18**, wherein the disease is Crohn's disease or ulcerative colitis.

20. A method for obtaining a purified non-naturally occurring fusion protein, the method comprising:

performing anion exchange chromatography on a mixture comprising the non-naturally occurring fusion protein to obtain a first fraction comprising the non-naturally occurring fusion protein;

wherein the non-naturally occurring fusion protein comprises IL-22 and a carrier; and wherein the carrier has at least 80% sequence identity to any one of SEQ ID NOS: 2, 3, 4, 5, 6, 7, 8, 9, 22, or 23.

21. The method of claim **20**, wherein performing anion exchange chromatography comprises binding the non-naturally occurring fusion protein to anionic exchange resin and providing an increasing salt gradient for subsequent elution of the non-naturally occurring fusion protein to obtain the first fraction.

22. The method of claim **20**, further comprising refolding the non-naturally occurring fusion protein prior to performing anion exchange chromatography.

23. The method of claim **22**, wherein refolding the non-naturally occurring fusion protein comprises contacting chaotrope-solubilized protein from inclusion bodies with a refolding solution, wherein the refolding solution comprises:

arginine (0.75 M to 1.25 M);
glycerol (2% to 20% v/v);
cysteine (0.5 mM to 10 mm); and
cystamine (0.2 mM to 10 mM);
wherein the refolding solution has a pH of between 7.5 and 8.5.

24. The method of claim **20**, further comprising subjecting a sample comprising the first fraction to a hydroxyapatite resin to obtain a second fraction comprising the non-naturally occurring fusion protein.

25. The method of claim **20**, further comprising performing cation exchange chromatography on a sample comprising the first fraction.

26. A method of refolding a non-naturally occurring fusion protein comprising a carrier and IL-22, the method comprising:

(i) contacting inclusion bodies comprising the non-naturally occurring fusion protein with a solubilization solution comprising a chaotropic agent to produce a soluble non-naturally occurring fusion protein; and

(ii) contacting the non-naturally occurring fusion protein with a refolding solution, wherein the refolding solution comprises:

arginine (0.75 M to 1.25 M);
glycerol (2% to 20% v/v);
cysteine (0.5 mM to 10 mm); and
cystamine (0.2 mM to 10 mM);
wherein the refolding solution has a pH of between 7.5 and 8.5.

27. The method of claim **26**, wherein:

arginine is present in the refolding solution at a concentration of between 0.9 M and 1.1 M;
glycerol is present in the refolding solution at a concentration of between 7% and 13% (w/w);
cysteine is present in the refolding solution at a concentration of between 1.5 mM and 6 mM;
cystamine is present in the refolding solution at a concentration of between 0.6 mM and 3 mM; and
the refolding solution has a pH of between 7.8 and 8.2.

* * * * *