(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 July 2023 (13.07.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/131811 A2

(51) International Patent Classification:
Not classified

(21) International Application Number:
PCT/IB2022/000815

(22) International Filing Date:
15 December 2022 (15.12.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/361,399 15 December 2021 (15.12.2021) US

(71) Applicant: **MEIRAGTX UK II LIMITED** [GB/GB]; 92 Britannia Walk, London N1 7NQ (GB).

(72) Inventors: **GUO, Xuecui**; 62 Woodland Dr., Oyster Bay, NY 11771 (US). **FORBES, Alexandria**; 14 East 10th Street, New York, NY 10003 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

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

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: POLYCISTRONIC EXPRESSION OF GUT PEPTIDES

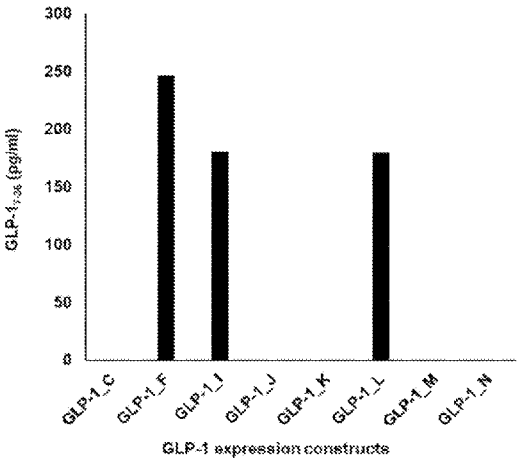


Fig. 1A

(57) Abstract: Provided herein are monocistronic, bicistronic, and polycistronic expression constructs for the expression of gut peptides and pharmaceutical compositions comprising such expression constructs. Also provided are methods of using such expression constructs, for example, for inducing satiation in a subject in need thereof or for treating obesity in a subject in need thereof.

WO 2023/131811 A2

POLYCISTRONIC EXPRESSION OF GUT PEPTIDES

FIELD

[0001] The disclosure relates to compositions and methods in the field of molecular biology. Specifically, the disclosure relates to polycistronic expression constructs for the expression of peptides as well as methods of using these polycistronic expression constructs.

BACKGROUND

[0002] Current treatments for obesity involve stimulant medications which are mildly effective and can have detrimental side effects, particularly if used long-term. Other current treatment methods involve invasive bariatric surgery which, while sometimes effective, can involve a variety of serious complications. More recently, satiation gut peptides (also referred to as satiation peptides or gut peptides) have been investigated as potential treatments for obesity.

[0003] Satiation gut peptides are chemical messengers that regulate gastrointestinal (GI) functions such as secretion, motility, absorption, digestion, and cell proliferation. These polypeptides are produced by endocrine cells in the stomach, pancreas, or intestine and act locally through autocrine or paracrine mechanisms, or at distant sites in a classical endocrine manner. Penetrating from plasma through the blood-brain barrier, they act by activating specific receptors in the satiety center of the hypothalamus, thus inducing satiation.

[0004] Acute supplemental therapy with satiation gut peptides reduces food intake and body weight in obese animal models as well as in lean and obese human subjects.

[0005] It is widely acknowledged that satiation gut peptides would not be effective through ingested oral administration since enzymes and acids in the gut would degrade them prior to reaching the blood. Accordingly, novel mechanisms for the expression of satiation gut peptides are urgently needed.

SUMMARY

[0006] Provided herein are expression constructs for the expression of gut peptides and methods of using such expression constructs.

[0007] In one aspect, provided is a bicistronic expression construct encoding a polyprotein, wherein:

a. the polyprotein comprises a signal peptide, a first gut peptide, and a second gut peptide; and
b. the polyprotein encoding sequence comprises:
    i. a sequence encoding the signal peptide;
    ii. a sequence encoding the first gut peptide; and
    iii. a sequence encoding the second gut peptide.

[0008] In some embodiments, the first gut peptide and/or the second gut peptide comprises a sequence selected from human glucagon like peptide 1 (hGLP-1) peptide, human glucose dependent insulinotropic (hGIP) peptide, human oxyntomodulin (hOXM) peptide, peptide YY (PYY), human glucagon (hGlucagon) peptide, and amlyn peptide. In embodiments, the hGLP-1 peptide is the hGLP-1_{7-37} peptide. In embodiments, the hGIP peptide is the $hGIP_{1-42}$ peptide. In some embodiments, the first gut peptide and/or the second gut peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:1-5. In some embodiments, the first gut peptide and/or the second gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS: 1-5. In some embodiments, the first gut peptide gut peptide and/or the second gut peptide comprises a sequence selected from SEQ ID NOS: 1-5.

[0009] In some embodiments, the sequence encoding the first gut peptide and/or the second gut peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:6-12. In some embodiments, the sequence encoding the first gut peptide gut peptide and/or the second gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:6-12. In some embodiments, the sequence encoding the first gut peptide gut peptide and/or the second gut peptide comprises a sequence that is selected from SEQ ID NOS:6-12.

[0010] In some embodiments, the first gut peptide and the second gut peptide are the same gut peptide. In some embodiments, the sequence encoding the first gut peptide and the sequence encoding the second gut peptide are different. In some embodiments, at least one of the sequence encoding the first gut peptide and the sequence encoding the second gut peptide is codon-optimized. In some embodiments, the sequence encoding the first gut peptide and the sequence encoding the second gut peptide are codon-optimized. In some embodiments, the first gut peptide and the second gut peptide is hGLP-1. In some embodiments, the first gut peptide and the second gut peptide each comprise a sequence that is at least 80% identical to SEQ ID NO:1. In some embodiments, the first gut peptide and the second gut peptide each comprise a sequence that is at least 90% identical to SEQ ID NO:1. In some embodiments, the first gut peptide and the second gut peptide each comprise SEQ ID NO:1.

[0011] In some embodiments, the sequence encoding the first gut peptide and the sequence encoding the second gut peptide each comprise a sequence that is at least 80% identical to a sequence selected from SEQ ID NOS:6-8. In some embodiments, the sequence encoding the first gut peptide and the sequence encoding the second gut peptide each comprise a sequence that is at least 90% identical to a sequence selected from SEQ ID NOS:6-8. In some embodiments, the sequences encoding the first and the second gut peptide are selected from SEQ ID NOS:6-8.

[0012] In some embodiments, the bicistronic expression construct comprises a sequence encoding a polypeptide that is at least 80% identical to SEQ ID NO:45 or SEQ ID NO:55. In some embodiments, the bicistronic expression construct comprises a sequence encoding a polypeptide that is at least 90% identical to SEQ ID NO:45 or SEQ ID NO:55. In some embodiments, the bicistronic expression construct encodes a polypeptide comprising SEQ ID NO:45 or SEQ ID NO:55.

[0013] In some embodiments, the bicistronic expression construct comprises a sequence that is at least 80% identical to SEQ ID NO:50 or SEQ ID NO:57. In some embodiments, the bicistronic expression construct comprises a sequence that is at least 90% identical to SEQ ID NO:50 or SEQ ID NO:57. In some embodiments, the bicistronic expression construct comprises SEQ ID NO:50 or SEQ ID NO:57.

[0014] In some embodiments, the first gut peptide and the second gut peptide are different gut peptides. In some embodiments, the first gut peptide and the second gut peptide are selected from the group consisting of hGLP-1 and hGIP. In embodiments, the hGLP-1 peptide is the hGLP-1_{7-37} peptide. In embodiments, the hGIP peptide is the $hGIP_{1-42}$ peptide.

[0015] In some embodiments, the bicistronic expression construct encodes a sequence comprising a sequence that is at least 80% identical to any one of SEQ ID NOS:46-49 or SEQ ID NO:56. In some embodiments, the bicistronic expression construct encodes a sequence comprising a sequence that is at least 90% identical to any one of SEQ ID NOS: 46-49 or SEQ ID NO:56. In some embodiments, the bicistronic expression construct encodes a sequence comprising any one of SEQ ID NOS:46-49 or SEQ ID NO:56.

[0016] In some embodiments, the bicistronic expression construct comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:51-54 or SEQ ID NO:58. In some embodiments, the bicistronic expression construct comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:51-54 or SEQ ID NO:58. In some embodiments, the bicistronic expression construct comprises a sequence selected from SEQ ID NOS:51-54 or SEQ ID NO:58.

[0017] In one aspect, provided is a tricistronic expression construct encoding a polyprotein, wherein:

a. the polyprotein comprises a signal peptide, a first gut peptide, a second gut peptide, and a third gut peptide; and
b. the polyprotein encoding sequence comprises:
    i. a sequence encoding the signal peptide;
    ii. a sequence encoding the first gut peptide;
    iii. a sequence encoding the second gut peptide; and
    iv. a sequence encoding the third gut peptide.

[0018] In some embodiments, the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence selected from the group consisting of hGLP-1 peptide, hGIP peptide, hOXM peptide, peptide YY (PYY), hGlucagon peptide, and amlyn peptide. In embodiments, the hGLP-1 peptide is the hGLP-1_{7-37} peptide. In embodiments, the hGIP peptide is the $hGIP_{1-42}$ peptide. In some embodiments, the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:1-5. In some embodiments, the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:1-4. In some embodiments, the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence selected from SEQ ID NOS:1-5.

[0019] In some embodiments, the sequence encoding first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:6-12. In some embodiments, the sequence encoding the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:6-12. In some embodiments, the sequence encoding first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is selected from SEQ ID NOS:6-12.

[0020] In some embodiments, the first gut peptide, the second gut peptide, and the third gut peptide are the same gut peptide. In some embodiments, the sequence encoding the first gut peptide, the sequence encoding the second gut peptide, and the sequence encoding the third gut peptide are different. In some embodiments, at least one of the sequence encoding the first gut peptide, the sequence encoding the second gut peptide, and the sequence encoding the third gut peptide is codon-optimized. In some embodiments, the sequence encoding the sequence encoding the first gut peptide, the sequence encoding the second gut peptide, and the sequence encoding the third gut peptide are codon-optimized.

[0021] In some embodiments, the first gut peptide, the second gut peptide and the third gut peptide is hGLP-1.

[0022] In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence that is at least 80% identical to any one of SEQ ID NOS:59-61 or SEQ ID NO:75. In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence that is at least 90% identical to any one of SEQ ID NOS:59-61 or SEQ ID NO:75. In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence selected from SEQ ID NOS:59-61 or SEQ ID NO:75.

[0023] In some embodiments, the tricistronic expression construct comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:67-69 or SEQ ID NO:78. In some embodiments, the tricistronic expression construct comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:67-69 or SEQ ID NO:78. In some embodiments, the tricistronic expression construct comprises a sequence selected from SEQ ID NOS:67-69 or SEQ ID NO:78.

[0024] In some embodiments, the first gut peptide, and the second gut peptide are different gut peptides. In some embodiments, the first gut peptide, the second gut peptide, and the third gut peptide are different gut peptides.

[0025] In some embodiments, the first gut peptide, the second gut peptide, and the third gut peptide are selected from the group consisting of (1) hGLP-1 peptide, hOXM peptide, and PYY or (2) hGLP-1 peptide, hGlucagon peptide, and hGIP peptide. In embodiments, the hGLP-1 peptide is the hGLP-1_{7-37} peptide. In embodiments, the hGIP peptide is the $hGIP_{1-42}$ peptide.

[0026] In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence that is at least 80% identical to any one of SEQ ID NOS:62-66 or SEQ ID NOS:76-77. In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence that is at least 90% identical to any one of SEQ ID NOS:62-66 or SEQ ID NOS:76-77. In some embodiments, the tricistronic expression construct encodes a sequence comprising any one of SEQ ID NOS:62-66 or SEQ ID NOS:76-77.

[0027] In some embodiments, the tricistronic expression construct comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:70-74 or SEQ ID NOS:79-80. In some embodiments, the tricistronic expression construct comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:70-74 or SEQ ID NOS:79-80. In some embodiments, the tricistronic expression construct comprises any one of SEQ ID NOS:70-74 or SEQ ID NOS:79-80.

[0028] In some embodiments, the bicistronic or the tricistronic expression construct encodes a polyprotein, wherein the polyprotein comprises a signal peptide. In some embodiments, the signal peptide is selected from the group consisting of an immunoglobulin M (IgM) signal peptide, human insulin (hInsul) signal peptide, murine Igh protein (mIgh) protein signal peptide, human growth hormone (hGH) signal peptide, murine erythropoietin (mEpo) signal peptide, murine growth hormone-releasing hormone (mGHRH) signal peptide, human albumin signal peptide, and human factor IX (FIX) signal peptide. In some embodiments, the signal peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:13-20. In some embodiments, the signal peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:13-20. In some embodiments, the signal peptide comprises a sequence selected from SEQ ID NOS:13-20. In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:21-28. In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:21-28. In some embodiments, the sequence encoding the signal peptide comprises a sequence selected from SEQ ID NOS:21-28.

[0029] In some embodiments, the bicistronic or the tricistronic expression construct further comprises a promoter sequence. In some embodiments, the promoter is a CMV or a CASI promoter.

[0030] In some embodiments, the bicistronic or the tricistronic expression construct encodes a polyprotein comprising a protease cleavage site positioned between the first gut peptide and the second gut peptide. In some embodiments, the tricistronic expression construct encodes a polyprotein wherein the polyprotein further comprises a protease cleavage site that allows release of the first gut peptide, second gut peptide, and/or the third peptide from the polyprotein.

[0031] In some embodiments, at least one of the protease cleavage sites is a furin cleavage site.

[0032] In some embodiments, the bicistronic expression construct or the tricistronic expression construct comprises a riboswitch comprising an aptamer, wherein the aptamer binds to a small molecule.

[0033] In some embodiments, the bicistronic expression construct or the tricistronic expression construct comprises a gene regulation cassette comprising an aptamer, wherein the aptamer binds to a small molecule.

[0034] Provided herein is a vector comprising a bicistronic expression or a tricistronic expression construct disclosed herein. In some embodiments, the vector is an adeno-associated virus (AAV) vector.

[0035] Provided herein is a cell comprising a vector disclosed herein. In some embodiments, the cell is isolated.

[0036] Provided herein is a pharmaceutical composition comprising a vector disclosed herein and a pharmaceutically acceptable excipient.

[0037] In one aspect, provided is a method of inducing satiation in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0038] In one aspect, provided is a method of treating obesity in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0039] In one aspect, provided is a method of suppressing appetite in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0040] In one aspect, provided is a method of reducing of reducing weight or reducing weight gain in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0041] In one aspect, provided is a method of improving glucose tolerance in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0042] In one aspect, provided is a method of inducing insulin release in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

BRIEF DESCRIPTION OF THE FIGURES

[0043] **Figs. 1A, 1B,** and **1C** illustrate the expression of gut peptides using monocistronic expression constructs. **Fig. 1A.** Expression of hGLP-1_{7-37} peptide, a gut peptide, as determined by ELISA. See **Tables 5** and **6** for the nomenclature of hGLP-1 expression constructs. **Fig. 1B.** Expression of hGLP-1_{7-37} peptide, a gut peptide, as determined by ELISA. See **Tables 5** and **6** for the nomenclature of hGLP-1 expression constructs. **Fig. 1C.** Expression of $hGIP_{1-42}$ peptide,

a gut peptide, as determined by ELISA. See **Tables 5** and **6** for the nomenclature of hGIP expression constructs.

[0044] **Figs. 2A, 2B, and 2C** illustrate the expression of gut peptides using mono-, bi-, and tricistronic expression constructs. **Fig. 2A.** Exemplary bi- and tricistronic expression constructs. **Fig. 2B.** Comparison of monocistronic (GLP-1_M), bicistronic (2xGLP-1_2xB) or tricistronic (3xGLP-1_3xB) expression of the GLP-1_{7-37} peptide as determined by ELISA. The ELISA kit used was designed to detect GLP-1_{7-36}. See Tables 5-10 for the nomenclature of expression constructs. **Fig. 2C.** Comparison of certain monocistronic and tricistronic constructs encoding the GLP1_{7-37} peptide. Expression was determined by ELISA. See **Tables 5-6** and **9-10** for the nomenclature of expression constructs.

[0045] **Figs. 3A, 3B, 3C, 3D,** and **3E** illustrate the expression of gut peptides using mono- and tricistronic expression constructs. **Fig. 3A.** Expression of the GLP-1_{7-37} peptide from bicistronic expression constructs encoding for a polyprotein comprising the GLP-1_{7-37} peptide and the hGIP$_{1-42}$ peptide. See **Tables 7** and **8** for the nomenclature of expression constructs. **Fig. 3B.** Expression of the hGIP$_{1-42}$ peptide from bicistronic expression constructs encoding for a polyprotein comprising the GLP-1_{7-37} peptide and the hGIP$_{1-42}$ peptide. See **Tables 7** and **8** for the nomenclature of expression constructs. **Fig. 3C.** Expression of the GLP-1_{7-37} peptide from monocistronic expression constructs encoding the GLP-1_{7-37} peptide (GLP1_J and GLP-1_L) or the hGIP$_{1-42}$ peptide (GIP_G) and from a tricistronic expression construct expressing the GLP-1_{7-36} peptide, the hGIP$_{1-42}$ peptide, and a hGlucagon peptide (GGG_A). See **Tables 5, 6, 9**, and **10** for the nomenclature of expression constructs. **Fig. 3D.** Expression of the hGIP$_{1-42}$ peptide from monocistronic expression constructs encoding the GLP-1_{7-37} peptide (GLP1_J and GLP-1_L) or the hGIP$_{1-42}$ peptide (GIP_G) and from a tricistronic expression construct expressing the GLP-1_{7-37} peptide, the hGIP$_{1-42}$ peptide, and a hGlucagon peptide (GGG_A). See **Tables 5, 6, 9**, and **10** for the nomenclature of expression constructs. **Fig. 3E.** Expression of the GLP-1_{7-37} peptide by indicated tricistronic expression constructs (expressing GLP-1_{7-37} peptide, OXM peptide, and PYY). See **Tables 9**, and **10** for the nomenclature of expression constructs. The ELISA kit used was designed to detect GLP-1_{7-36}.

[0046] **Figs**. **4A, 4B, 4C, 4D,** and **4E** illustrate the riboswitch-regulated expression of gut peptides. **Fig. 4A.** Expression of the GLP-1 peptide by the indicated, regulatable bicistronic expression constructs based on GG_L (expressing the hGLP-1_{7-37} peptide and the hGIP$_{1-42}$ peptide) described in Example 4. The concentration of the small molecule inducer is shown in μM. **Fig. 4B.** Expression of the hGIP$_{1-42}$ peptide by the indicated, regulatable bicistronic expression constructs (expressing the hGLP-1_{7-37} peptide and the hGIP$_{1-42}$ peptide) described

in Example 4. No significant expression was observed for 0 mM of the small molecule inducer. The concentration of the small molecule inducer is shown in μM. **Fig. 4C.** Expression of the hGLP-1_{7-37} peptide by the indicated, regulatable tricistronic expression construct 3xGLP-1_3xC described in Example 4 (comprising three hGLP-1_{7-37} peptide encoding sequences). **Fig. 4C.** Expression of the hGLP-1_{7-37} peptide by indicated regulatable, bicistronic and tricistronic expression constructs described in Example 4. MX-001 is the small molecule inducer. **Fig. 4D.** Expression of the hGLP-1_{7-37} peptide by the indicated, regulatable bicistronic expression constructs based on GG_F described in Example 4 (expressing the hGLP-1_{7-36} peptide and the $hGIP_{1-42}$ peptide. **Fig. 4E.** Expression of PYY from a regulatable tricistronic expression construct expressing a polyprotein comprising GLP-1, hOXM, and PYY expressed PYY. The ELISA kit used was designed to detect GLP-1_{7-36}.

[0047] **Figs. 5A** and **5B** illustrate that gut peptides expressed from the polycistronic expression constructs disclosed herein are biologically active. **Fig. 5A.** Biological activity of the hGLP-1_{7-37} peptide expressed by the indicated mono-, bi-, and tricistronic expression constructs. **Fig. 5B.** Biological activity of the $hGIP_{1-42}$ peptide expressed by the indicated mono-, bi-, and tricistronic expression constructs. **Fig. 5C.** Male C57Bl/6 mice were fed with high fat diet (HFD) starting at 6 weeks of age. At week eight, the mice were injected with either PBS or AAV8 vectors containing GG_F. Mice on a low fat diet (LFD) were injected PBS and served as a control group. Animal body weight was monitored before and after AAV injection weekly.

[0048] **Figs. 6A and 6B** illustrate that GLP-1 and GIP peptides expressed from a AAV8.GG_F_7-GLP-1 vectors improve glucose tolerance *in vivo*. **Fig. 6A.** Experimental setup. **Fig. 6B.** GLP-1 and GIP peptides expressed from a AAV8.GG_F_7-GLP-1 vectors improve glucose tolerance *in vivo*.

DETAILED DESCRIPTION OF THE INVENTION

[0049] Provided herein are expression constructs encoding for one or more gut peptides as well as methods of using these expression constructs. In embodiments, the gut peptides are expressed as a polyprotein, which is cleaved to produce the desired gut peptides. As used herein, a polyprotein is a protein which is destined for processing to produce two or more polypeptide products.

[0050] Expression constructs

[0051] Provided herein are monocistronic, bicistronic, tricistronic and other polycistronic expression constructs for the expression of gut peptides.

[0052] In embodiments, the expression construct is a monocistronic expression construct for the expression of a single polypeptide.

[0053] In embodiments, the expression construct is a bicistronic expression construct for the expression of two polypeptides. The two polypeptides may be expressed as a polyprotein and the individual polypeptides maybe be released from the polyprotein after proteolytic cleavage.

[0054] In embodiments, the expression construct is a tricistronic expression construct for the expression of three polypeptides. The three polypeptides may be expressed as a polyprotein and the individual polypeptides maybe be released from the polyprotein after proteolytic cleavage.

[0055] In embodiments, the expression construct is a polycistronic expression construct for the expression of two or more polypeptides. The two or more polypeptides may be expressed as a polyprotein and the individual polypeptides maybe be released from the polyprotein after proteolytic cleavage. In some embodiments, the polycistronic expression construct expresses two, three, four, five, six, seven, eight, nine, or ten polypeptides. The two or more polypeptides may be the same or different polypeptides.

[0056] In one aspect, the expression constructs provided herein encode one or more gut peptides.

[0057] *Monocistronic expression constructs*

[0058] In some embodiments, the gut peptide is human glucagon-like peptide 1 (hGLP-1) peptide, human gastric inhibitory peptide (hGIP) peptide, human oxyntomodulin (hOXM) peptide, peptide YY or peptide tyrosine tyrosine (PYY), human glucagon (hGlucagon) peptide, or amlyn peptide (also called insulinoma amyloid polypeptide (IAPP)). In embodiments, the hGLP-1 peptide is the hGLP-1_{7-36} peptide. In embodiments, the hGIP peptide is the $hGIP_{1-42}$ peptide. In some embodiments, the gut peptide is a gut peptide disclosed in **Table 1** or a portion of one of the gut peptides disclosed in **Table 1**.

[0059] In some embodiments, provided is an expression construct that encodes for a polypeptide comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:1. In some embodiments, the expression construct encodes for a polypeptide comprising SEQ ID NO:1.

[0060] In some embodiments, provided is an expression construct that encodes for a polypeptide comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:2. In some embodiments, the expression construct encodes for a polypeptide comprising SEQ ID NO:2.

[0061] In some embodiments, provided is an expression construct that encodes for a polypeptide comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:3. In some embodiments, the expression construct encodes for a polypeptide comprising SEQ ID NO:3.

[0062] In some embodiments, provided is an expression construct that encodes for a polypeptide comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:4. In some embodiments, the expression construct encodes for a polypeptide comprising SEQ ID NO:4.

[0063] In some embodiments, provided is an expression construct that encodes for a polypeptide comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:5. In some embodiments, the expression construct encodes for a polypeptide comprising SEQ ID NO:5.

Table 1. Selected gut peptide amino acid sequences.

SEQ ID NO	Gut peptide	Sequence
1	hGLP-1_{7-37}	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG
2	hGIP$_{1-42}$	YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ
3	hOXM	HSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIA
4	PYY$_{3-36}$	IKPEAPREDASPEELNRYYASLRHYLNLVTRQRY
5	hGlucagon	HSQGTFTSDYSKYLDSRRAQDFVQWLMNT

[0064] In some embodiments, the expression construct comprises a sequence disclosed in **Table 2** or a portion of a sequence disclosed in **Table 2.**

[0065] In some embodiments, the expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least

95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:6. In some embodiments, the expression construct comprises SEQ ID NO:6.

[0066] In some embodiments, the expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:7. In some embodiments, the expression construct comprises SEQ ID NO:7.

[0067] In some embodiments, the expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:8. In some embodiments, the expression construct comprises SEQ ID NO:8.

[0068] In some embodiments, the expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:9. In some embodiments, the expression construct comprises SEQ ID NO:9.

[0069] In some embodiments, the expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:10. In some embodiments, the expression construct comprises SEQ ID NO:10.

[0070] In some embodiments, the expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:11. In some embodiments, the expression construct comprises SEQ ID NO:11.

[0071] In some embodiments, the expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:12. In some embodiments, the expression construct comprises SEQ ID NO:12.

Table 2. Selected gut peptide nucleic acid sequences.

SEQ ID NO	Encoded gut peptide	Sequence
6	hGLP-1 (encoding sequence variant 1)	CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA
7	hGLP-1 variant (encoding sequence variant 2)	CATGCAGAGGGAACATTTACTAGTGATGTCAGTTCATATCTTGAGGGACAAGCTGCTAAAGAATTTATTGCTTGGCTTGTGAAGGGAAGAGGA
8	hGLP-1 variant (encoding sequence variant 3)	CATGCTGAAGGGACATTTACCTCAGATGTTTCTTCATACCTGGAAGGACAGGCTGCCAAGGAATTTATTGCATGGCTTGTGAAAGGCAGGGGC
9	hGIP	TACGCGGAAGGGACTTTCATCAGTGACTACAGTATTGCCATGGACAAGATTCACCAACAAGACTTTGTGAACTGGCTGCTGGCCCAAAAGGGGAAGAAGAATGACTGGAAACACAACATCACCCAG
10	hOXM	CATTCACAGGGCACATTCACCAGTGACTACAGCAAGTATCTGGACTCCAGGCGTGCCCAAGATTTTGTGCAGTGGTTGATGAATACCAAGAGGAACAGGAATAACATTGCC
11	PYY	ATCAAACCCGAGGCTCCCCGCGAAGACGCCTCGCCGGAGGAGCTGAACCGCTACTACGCCTCCCTGCGCCACTACCTCAACCTGGTCACCCGGCAGCGGTAT
12	hGlucagon	CATTCACAGGGCACATTCACCAGTGACTACAGCAAGTATCTGGACTCCAGGCGTGCCCAAGATTTTGTGCAGTGGTTGATGAATACC

[0072] In some embodiments, the expression construct encodes for a gut peptide, wherein the gut peptide is fused to a signal peptide. In some embodiment, the signal peptide is immunoglobulin M (IgM) signal peptide, human insulin (hInsul) signal peptide, murine Igh protein (mIgh) signal peptide, human growth hormone (hGH) signal peptide, murine erythropoietin (mEpo) signal peptide, murine growth hormone-releasing hormone (mGHRH) signal peptide, human albumin (hAlbumin) signal peptide, or human factor IX (hFIX) signal peptide. In some embodiments, the signal peptide is a signal peptide disclosed in **Table 3** or a portion of a signal peptide disclosed in **Table 3.**

[0073] In some embodiments, the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:13. In some embodiments, the signal peptide comprises SEQ ID NO:13.

[0074] In some embodiments, the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%,

at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:14. In some embodiments, the signal peptide comprises SEQ ID NO:14.

[0075] In some embodiments, the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:15. In some embodiments, the signal peptide comprises SEQ ID NO:15.

[0076] In some embodiments, the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:16. In some embodiments, the signal peptide comprises SEQ ID NO:16.

[0077] In some embodiments, the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:17. In some embodiments, the signal peptide comprises SEQ ID NO:17.

[0078] In some embodiments, the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:18. In some embodiments, the signal peptide comprises SEQ ID NO:18.

[0079] In some embodiments, the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:19. In some embodiments, the signal peptide comprises SEQ ID NO:19.

[0080] In some embodiments, the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:20. In some embodiments, the signal peptide comprises SEQ ID NO:20.

Table 3. Selected signal peptide amino acid sequences.

SEQ ID NO	Signal peptide	Sequence
13	IgM	MGWSCIILFLVATATGAHSA
14	hInsul	MALWMRLLPLLALLALWGPDPAAA
15	mIgh	MAWVWTLLFLMAAAQSIQA
16	hGH	MATGSRTSLLLAFGLLCLPWLQEGSA
17	mEpo	MGVPERPTLLLLLSLLLIPLGLPVLC
18	mGHRH	MLLWVLFVILILTSGSHCS
19	hAlbumin	MKWVTFISLLFLFSSAYS
20	hFIX	MQRVNMIMAESPGLITICLLGYLLSAEC

[0081] In some embodiments, the expression construct comprises a sequence encoding a signal peptide, wherein the signal peptide is fused to the gut peptide. In some embodiments, the sequence encoding the signal peptide comprises a sequence disclosed in **Table 4** or a portion of a sequence disclosed in **Table 4.**

[0082] In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:21. In some embodiments, the sequence encoding the signal peptide comprises SEQ ID NO:21.

[0083] In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:22. In some embodiments, the sequence encoding the signal peptide comprises SEQ ID NO:22.

[0084] In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:23. In some embodiments, the sequence encoding the signal peptide comprises SEQ ID NO:23.

[0085] In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:24. In some embodiments, the sequence encoding the signal peptide comprises SEQ ID NO:24.

[0086] In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:25. In some embodiments, the sequence encoding the signal peptide comprises SEQ ID NO:25.

[0087] In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:26. In some embodiments, the sequence encoding the signal peptide comprises SEQ ID NO:26.

[0088] In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:27. In some embodiments, the sequence encoding the signal peptide comprises SEQ ID NO:27.

[0089] In some embodiments, the sequence encoding the signal peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:28. In some embodiments, the sequence encoding the signal peptide comprises SEQ ID NO:28.

Table 4. Selected signal peptide nucleic acid sequences.

SEQ ID NO	Encoded signal peptide	Sequence
21	IgM	ATGGGATGGAGCTGTATCATCCTCTTCTTGGTAGCAACAGCTACAGGCGCGCACAGTGCA
22	hInsul	ATGGCCCTGTGGATGCGCCTCCTGCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCAGCCGCAGCC
23	mIgh	ATGGCCTGGGTCTGGACACTCCTGTTTCTGATGGCTGCTGCCCAGTCCATTCAGGCC
24	hGH	ATGGCTACAGGCTCCCGGACGTCCCTGCTCCTGGCTTTTGGCCTGCTCTGCCTGCCCTGGCTTCAAGAGGGCAGTGCC
25	mEpo	ATGGGGGTGCCCGAACGTCCCACCCTGCTGCTTTTACTCTCCTTGCTACTGATTCCTCTGGGCCTCCCAGTCCTCTGT
26	mGHRH	ATGCTGCTCTGGGTGCTCTTTGTGATCCTCATCCTCACCAGTGGCTCCCACTGCTCA
27	hAlbumin	ATGAAGTGGGTAACCTTTATTTCCCTTCTTTTTCTCTTTAGCTCGGCTTATTCC
28	hFIX	ATGCAGCGCGTGAACATGATCATGGCAGAATCACCAGGCCTCATCACCATCTGCCTTTTAGGATATCTACTCAGTGCTGAATGT

[0090] Provided is an expression construct encoding a polypeptide comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any of the sequences SEQ ID NOS:29-36. Provided is an expression construct encoding a polypeptide comprising any one of SEQ ID NOS:29-36.

[0091] Provided is an expression construct encoding a polypeptide comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any of the sequences SEQ ID NOs:30, 21 or 34. Provided is an expression construct encoding a polypeptide comprising any one of SEQ ID NOs:30, 21 or 34.

[0092] Provided is an expression construct encoding a polypeptide comprising any one of the sequences disclosed in **Table 5** or a portion of a sequence disclosed in **Table 5.**

Table 5. Amino acid sequences of exemplary monocistronic expression constructs. hGLP-1 (SEQ ID NO:1) is shown in bold. hGIP (SEQ ID NO:2) is underlined. Signal peptide encoding sequence is shown in italic letters. The position where furin cleavage occurs is marked with *. Constructs GLP-1_C, N, M and J, respectively, have the same leader sequence as constructs GIP_C, E, F and G, respectively.

SEQ ID NO	Construct name	Signal peptide	Encoded gut peptide	Sequence
29	GLP-1_C	IgM SP	hGLP-1	*MGWSCIILFLVATATGAHSA*YPYDVPDYARKKR* **HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
30	GLP-1_F	hInsul	hGLP-1	*MALWMRLLPLLALLALWGPDPAAA*YPYDVPDYARKKR* **HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
31	GLP-1_I	mIgh	hGLP-1	*MAWVWTLLFLMAAAQSIQA*YPYDVPDYARKKR* **HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
32	GLP-1_J	hGH	hGLP-1	*MATGSRTSLLLAFGLLCLPWLQEGSA*FPTIPLSRLFDNAMLRARKKR* **HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
33	GLP-1_K	mEpo	hGLP-1	*MGVPERPTLLLLLSLLLIPLGLPVLC*APPRLICDSRVLERYRKKR* **HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
34	GLP-1_L	mGHRH	hGLP-1	*MLLWVLFVILILTSGSHCS*LPPSPPFRMQR* **HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
35	GLP-1_M	hAlbumin	hGLP-1	*MKWVTFISLLFLFSSAYS*RGVFRR* **HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
36	GLP-1_N	hFIX9	hGLP-1	*MQRVNMIMAESPGLITICLLGYLLSAEC*TVFLDHENANKILNRPKR* **HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
111	GIP_C	IgM SP	hGIP	*MGWSCIILFLVATATGAHSA*YPYDVPDYARKKR* YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ
112	GIP_E	hFIX9	hGIP	*MQRVNMIMAESPGLITICLLGYLLSAEC*TVFLDHENANKILNRPKR* YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ
113	GIP_F	hAlbumin	hGIP	*MKWVTFISLLFLFSSAYS*RGVFRR* YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ
114	GIP_G	hGH	hGIP	*MATGSRTSLLLAFGLLCLPWLQEGSA*FPTIPLSRLFDNAMLRARKKR* YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ

[0093] Provided is an expression construct comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any of the

sequences SEQ ID NOS:37-44. Provided is an expression construct encoding a polypeptide comprising any one of SEQ ID NOS:37-44.

[0094] Provided is an expression construct comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any of the sequences SEQ ID NOS:38, 39, or 42. Provided is an expression construct encoding a polypeptide comprising any one of SEQ ID NOS: 38, 39, or 42.

[0095] Provided is an expression comprising any one of the sequences disclosed in **Table 6** or a portion of a sequence disclosed in **Table 6.**

Table 6. Nucleic acid sequences of exemplary monocistronic expression constructs. All sequences comprise hGLP-1 encoding sequence variant 1 (SEQ ID NO:6). Signal peptide encoding sequence is shown in italic upper-case letters. Furin cleavage site encoding sequence is shown in lower case letters. Gut peptide encoding sequence is shown in upper case letters (not italic).

SEQ ID NO	Construct name	Signal peptide	Encoded gut peptide	Sequence
37	GLP-1_C	IgM	hGLP-1	*ATGGGATGGAGCTGTATCATCCTCTTCTTGGTAGCAACAGCTACAGGCGCGCACAGTGCA*tacccatacgatgttccagattacgctagaaaaaagagaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGATGA
38	GLP-1_F	hInsul	hGLP-1	*ATGGCCCTGTGGATGCGCCTCCTGCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCAGCCGCAGCC*tacccatacgatgttccagattacgctagaaaaaagagaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGATGA
39	GLP-1_I	mIgh	hGLP-1	*ATGGCCTGGGTCTGGACACTCCTGTTTCTGATGGCTGCTGCCCAGTCCATTCAGGCC*tacccatacgatgttccagattacgctagaaaaaagagaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGATGA
40	GLP-1_J	hGH	hGLP-1	*ATGGCTACAGGCTCCCGGACGTCCCTGCTCCTGGCTTTTGGCCTGCTCTGCCTGCCCTGGCTTCAAGAGGGCAGTGCC*ttcccaaccattcccttatccaggctttttgacaacgctatgctccgcgccagaaaaaagagaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAA

SEQ ID NO	Construct name	Signal peptide	Encoded gut peptide	Sequence
				GGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGATGA
41	GLP-1_K	mEpo	hGLP-1	*ATGGGGGTGCCCGAACGTCCCACCCTGCTGCTTTTACTCTCCTTGCTACTGATTCCTCTGGGCCTCCCAGTCCTCTGT*gctcccccacgcctcatctgcgacagtcgagttctggagaggtacagaaaaaagagaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGATGA
42	GLP-1_L	mGHRH	hGLP-1	*ATGCTGCTCTGGGTGCTCTTTGTGATCCTCATCCTCACCAGTGGCTCCCACTGCTCA*actgccccccctcacctcccttcaggatgcagcgaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGATGA
43	GLP-1_M	hAlbumin	hGLP-1	*ATGAAGTGGGTAACCTTTATTTCCCTTCTTTTTCTCTTTAGCTCGGCTTATTCC*aggggtgtgtttcgtcgaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGATGA
44	GLP-1_N	hFIX9	hGLP-1	*ATGCAGCGCGTGAACATGATCATGGCAGAATCACCAGGCCTCATCACCATCTGCCTTTTAGGATATCTACTCAGTGCTGAATGT*acagttttttcttgatcatgaaaacgccaacaaaattctgaatcggccaaagaggCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGATGA

[0096] <u>*Bicistronic expression constructs*</u>

[0097] In one aspect, provided is a bicistronic expression construct encoding a polyprotein, wherein:

a. the polyprotein comprises a signal peptide, a first gut peptide, and a second gut peptide; and
b. the polyprotein encoding sequence comprises:
    i. a sequence encoding the signal peptide;
    ii. a sequence encoding the first gut peptide; and
    iii. a sequence encoding the second gut peptide.

[0098] In some embodiments, the first gut peptide and/or the second gut peptide comprises a sequence selected from the group consisting of hGLP-1 peptide, hGIP peptide, hOXM

peptide, PYY, hGlucagon peptide, and amlyn peptide. In embodiments, the hGLP-1 peptide is the hGLP-1_{7-37} peptide. In embodiments, the hGIP peptide is the $hGIP_{1-42}$ peptide. In some embodiments, the first gut peptide and/or the second gut peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:1-5. In some embodiments, the first gut peptide and/or the second gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS: 1-5. In some embodiments, the first gut peptide gut peptide and/or the second gut peptide comprises a sequence selected from SEQ ID NOS: 1-5.

[0099] In some embodiments, the sequence encoding the first gut peptide gut peptide and/or the second gut peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:6-12. In some embodiments, the sequence encoding the first gut peptide gut peptide and/or the second gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:6-12. In some embodiments, the sequence encoding the first gut peptide gut peptide and/or the second gut peptide comprises a sequence that is selected from SEQ ID NOS:6-12.

[0100] In some embodiments, the first gut peptide and the second gut peptide are the same gut peptide. In some embodiments, the first gut peptide and the second gut peptide are the same gut peptide, but the sequence encoding the first gut peptide and the sequence encoding the second gut peptide are different. In some embodiments, at least one of the sequence encoding the first gut peptide and the sequence encoding the second gut peptide is codon-optimized. In some embodiments, the sequence encoding the first gut peptide and the sequence encoding the second gut peptide are codon-optimized. In some embodiments, the first gut peptide and the second gut peptide is hGLP-1.

[0101] In some embodiments, the first gut peptide and the second gut peptide comprise a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:1. In some embodiments, the first gut peptide and the second gut peptide comprise a sequence that is at least 90% identical to SEQ ID NO:1. In some embodiments, the first gut peptide and the second gut peptide comprise SEQ ID NO:1.

[0102] In some embodiments, the sequence encoding the first gut peptide and the sequence encoding the second gut peptide comprise a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least

96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to a sequence selected from SEQ ID NOS:6-8. In some embodiments, the sequence encoding the first gut peptide and the sequence encoding the second gut peptide comprise a sequence that is at least 90% identical to a sequence selected from SEQ ID NOS:6-8. In some embodiments, the sequences encoding the first and the second gut peptide are selected from SEQ ID NOS:6-8.

[0103] In some embodiments, the bicistronic expression construct comprises a sequence disclosed in **Table 7** or a portion of a sequence disclosed in **Table 7.**

[0104] In some embodiments, the bicistronic expression construct comprises a sequence encoding a polypeptide that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:45 or SEQ ID NO:55. In some embodiments, the bicistronic expression construct comprises a sequence encoding a polypeptide that is at least 90% identical to SEQ ID NO:45 or SEQ ID NO:55 In some embodiments, the bicistronic expression construct encodes a polypeptide comprising SEQ ID NO:45 or SEQ ID NO:55.

[0105] In some embodiments, the bicistronic expression construct encodes a sequence comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:46-49 or SEQ ID NO:56. In some embodiments, the bicistronic expression construct encodes a sequence comprising a sequence that is at least 90% identical to any one of SEQ ID NOS:46-49 or SEQ ID NO:56. In some embodiments, the bicistronic expression construct encodes a sequence comprising any one of SEQ ID NOS:46-49 or SEQ ID NO:56.

Table 7. Amino acid sequences of exemplary bicistronic expression constructs. hGLP-1 (SEQ ID NO:1) is shown in bold. hGIP (SEQ ID NO:2) is underlined. The position where furin cleavage occurs is marked with *.

SEQ ID NO	Construct name	Signal peptide	Gut peptides	Sequence
45	2xGLP-1_2xB	hAlbumin	hGLP-1, hGLP-1	MKWVTFISLLFLFSSAYSRGVFRR***HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR***HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
46	GG_J	hGH	hGLP-1, hGIP	MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNAMLRARKKR***HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR*<u>YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ</u>
47	GG_F	hInsul	hGLP-1, hGIP	MALWMRLLPLLALLALWGPDPAAAYPYDVPDYARKKR***HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR*<u>YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ</u>
48	GG_L	mGHRH	hGLP-1, hGIP	MLLWVLFVILILTSGSHCSLPPSPPFRMQR***HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR*<u>YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ</u>
49	GG_M	hAlbumin	hGLP-1, hGIP	MKWVTFISLLFLFSSAYSRGVFRR***HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR*<u>YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ</u>
55	2xGLP-1 without signal peptide	n/a	hGLP-1, hGLP-1	**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR***HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
56	GG without signal peptide	n/a	hGLP-1, hGIP	**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR*<u>YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ</u>

[0106] In some embodiments, the bicistronic expression comprises a sequence disclosed in **Table 8** or a portion of a sequence disclosed in **Table 8.**

[0107] In some embodiments, the bicistronic expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO:50 or SEQ ID NO:57. In some embodiments, the bicistronic expression construct comprises a sequence that is at least 90% identical to SEQ ID NO:50 or SEQ ID NO:57 In some embodiments, the bicistronic expression construct comprises SEQ ID NO:50 or SEQ ID NO:57.

[0108] In some embodiments, the bicistronic expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:51-54 or SEQ ID NO:58. n some embodiments, the bicistronic expression construct comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:51-54 or SEQ ID NO:58. In some embodiments, the bicistronic expression construct comprises a sequence selected from SEQ ID NOS:51-54 or SEQ ID NO:58.

Table 8. Nucleic acid sequences of exemplary bicistronic expression constructs. All sequences comprise hGLP-1 encoding sequence variant 1 (SEQ ID NO:6). Signal peptide encoding sequence is shown in italic upper-case letters. Furin clevage site encoding sequence is shown in lower case letters. Gut peptide encoding sequence is shown in upper case letters (not italic).

SEQ ID NO	Construct name	Signal peptide	Gut peptides	Sequence
50	2xGLP-1_2xB	hAlbumin	hGLP-1, hGLP-1	*ATGAAGTGGGTAACCTTTATTTCCCTTCTTTTTCTCTTTAGCTCGGCTTATTCC*aggggtgtgtttcgtcgaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaCATGCAGAGGGAACATTTACTAGTGATGTCAGTTCATATCTTGAGGGACAAGCTGCTAAAGAATTTATTGCTTGGCTTGTGAAGGGAAGAGGAtga
51	GG_J	hGH	hGLP-1, hGIP	*ATGGCTACAGGCTCCCGGACGTCCCTGCTCCTGGCTTTTGGCCTGCTCTGCCTGCCCTGGCTTCAAGAGGGCAGTGCC*ttcccaaccattcccttatccaggctttttgacaacgctatgctccgcgccagaaaaaagagaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaTACGCGGAAGGGACTTTCATCAGTGACTACAGTATTGCCATGGACAAGATTCACCAACAAGACTTTGTGAACTGGCTGCTGGCCCAAAAGGGGAAGAAGAATGACTGGAAACACAACATCACCCAGTGA
52	GG_F	hInsul	hGLP-1, hGIP	*ATGGCCCTGTGGATGCGCCTCCTGCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCAGCCGCAGCC*tacccatacgatgttccagattacgctagaaaaaagagaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaTGCGGAAGGGACTTTCATCAGTGACTACAG

SEQ ID NO	Construct name	Signal peptide	Gut peptides	Sequence
				TATTGCCATGGACAAGATTCACCAACAAGACTTTGTGAACTGGCTGCTGGCCCAAAAGGGGAAGAAGAATGACTGGAAACACAACATCACCCAGTGA
53	GG_L	mGHRH	hGLP-1, hGIP	*ATGCTGCTCTGGGTGCTCTTTGTGATCCTCATCCTCACCAGTGGCTCCCACTGCTCA*ctgccccccctcacctcccttcaggatgcagcgaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaTACGCGGAAGGGACTTTCATCAGTGACTACAGTATTGCCATGGACAAGATTCACCAACAAGACTTTGTGAACTGGCTGCTGGCCCAAAAGGGGAAGAAGAATGACTGGAAACACAACATCACCCAGTGA
54	GG_M	hAlbumin	hGLP-1, hGIP	*ATGAAGTGGGTAACCTTTATTTCCCTTCTTTTTCTCTTTAGCTCGGCTTATTCC*aggggtgtgtttcgtcgaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaTACGCGGAAGGGACTTTCATCAGTGACTACAGTATTGCCATGGACAAGATTCACCAACAAGACTTTGTGAACTGGCTGCTGGCCCAAAAGGGGAAGAAGAATGACTGGAAACACAACATCACCCAGTGA
57	2xGLP-1 without signal peptide encoding sequence	n/a	hGLP-1, hGLP-1	CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaCATGCAGAGGGAACATTTACTAGTGATGTCAGTTCATATCTTGAGGGACAAGCTGCTAAAGAATTTATTGCTTGGCTTGTGAAGGGAAGAGGAtga
58	GG without signal peptide encoding sequence	n/a	hGLP-1, hGIP	CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaTACGCGGAAGGGACTTTCATCAGTGACTACAGTATTGCCATGGACAAGATTCACCAACAAGACTTTGTGAACTGGCTGCTGGCCCAAAAGGGGAAGAAGAATGACTGGAAACACAACATCACCCAGTGA

[0109] *Tricistronic expression constructs*

[0110] In one aspect, provided is a tricistronic expression construct encoding a polyprotein, wherein:

a. the polyprotein comprises a signal peptide, a first gut peptide, a second gut peptide, and a third gut peptide; and
b. the polyprotein encoding sequence comprises:
    i. a sequence encoding the signal peptide
    ii. a sequence encoding the first gut peptide;
    iii. a sequence encoding the second gut peptide; and
    iv. a sequence encoding the third gut peptide.

[0111] In some embodiments, the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence selected from the group consisting of hGLP-1 peptide, hGIP peptide, hOXM peptide, PYY, hGlucagon peptie, and amlyn peptide. In embodiments, the hGLP-1 peptide is the hGLP-1_{7-37} peptide. In embodiments, the hGIP peptide is the $hGIP_{1-42}$ peptide. In some embodiments, the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:1-5. In some embodiments, the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:1-5. In some embodiments, the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence selected from SEQ ID NOS:1-5.

[0112] In some embodiments, the sequence encoding first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:6-12. In some embodiments, the sequence encoding the first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:6-12. In some embodiments, the sequence encoding first gut peptide, the second gut peptide, and/or the third gut peptide comprises a sequence that is selected from SEQ ID NOS:6-12.

[0113] In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence disclosed in **Table 9** or a portion of a sequence disclosed in **Table 9.** In

some embodiments, the tricistronic expression construct comprises a sequence disclosed in **Table 10** or a portion of a sequence disclosed in **Table 10.**

[0114] In some embodiments, the first gut peptide, the second gut peptide, and the third gut peptide are the same gut peptide. In some embodiments, the first gut peptide, the second gut peptide, and the third gut peptide are the same gut peptide, but the sequence encoding the first gut peptide, the sequence encoding the second gut peptide, and the sequence encoding the third gut peptide are different. In some embodiments, at least one of the sequence encoding the first gut peptide, the sequence encoding the second gut peptide, and the sequence encoding the third gut peptide is codon-optimized. In some embodiments, the sequence encoding the sequence encoding the first gut peptide, the sequence encoding the second gut peptide, and the sequence encoding the third gut peptide are codon-optimized.

[0115] In some embodiments, the first gut peptide, the second gut peptide and the third gut peptide are hGLP-1.

[0116] In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:59-61 or SEQ ID NO:75. In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence that is at least 90% identical to any one of SEQ ID NOS:59-61 or SEQ ID NO:75. In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence selected from SEQ ID NOS:59-61 or SEQ ID NO:75.

[0117] In some embodiments, the tricistronic expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:67-69 or SEQ ID NO:78. In some embodiments, the tricistronic expression construct comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:67-69 or SEQ ID NO:78. In some embodiments, the tricistronic expression construct comprises a sequence selected from SEQ ID NOS:67-69 or SEQ ID NO:78.

[0118] In some embodiments, the first gut peptide, and the second gut peptide are different gut peptides. In some embodiments, the first gut peptide, the second gut peptide, and the third gut peptide are different gut peptides.

[0119] In some embodiments, the first gut peptide, the second gut peptide, and the third gut peptide are selected from the group consisting of (1) hGLP-1 peptide, hOXM peptide, and PYY

or (2) hGLP-1 peptide, hGlucagon peptide, and hGIP peptide. In embodiments, the hGLP-1 peptide is the hGLP-1_{7-37} peptide. In embodiments, the hGIP peptide is the $hGIP_{1-42}$ peptide.

[0120] In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:62-66 or SEQ ID NOS:76-77. In some embodiments, the tricistronic expression construct encodes a sequence comprising a sequence that is at least 90% identical to any one of SEQ ID NOS:62-66 or SEQ ID NOS:76-77. In some embodiments, the tricistronic expression construct encodes a sequence comprising any one of SEQ ID NOS:62-66 or SEQ ID NOS:76-77.

[0121] In some embodiments, the tricistronic expression construct comprises a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOS:70-74 or SEQ ID NOS:79-80. In some embodiments, the tricistronic expression construct comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:70-74 or SEQ ID NOS:79-80. In some embodiments, the tricistronic expression construct comprises any one of SEQ ID NOS:70-74 or SEQ ID NOS:79-80.

Table 9. Amino acid sequences of exemplary tricistronic expression constructs. hGLP-1 (SEQ ID NO:1) is shown in bold. hOXM (SEQ ID NO:3) is underlined. PYY (SEQ ID NO:4) is bold, underlined, and in italics.

SEQ ID NO	Construct name	Signal peptide	Gut peptides	Sequence
59	3xGLP-1_3xB	hAlbumin	hGLP-1, hGLP-1, hGLP-1,	MKWVTFISLLFLFSSAYSRGVFRR**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR** **G**RKKR**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR**HAEGTFTSDVSSYL** **EGQAAKEFIAWLVKGRG**
60	3xGLP-1_3xC	hInsul	hGLP-1, hGLP-1, hGLP-1,	MALWMRLLPLLALLALWGPDPAAAYPYDVPDYARKKR**HAEGTFTSDVSSYLEGQ** **AAKEFIAWLVKGRG**RKKR**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR**H** **AEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**
61	3xGLP-1_3xD	mGHRH	hGLP-1, hGLP-1, hGLP-1,	MLLWVLFVILILTSGSHCSLPPSPPFRMQR**HAEGTFTSDVSSYLEGQAAKEFIA** **WLVKGRG**RKKR**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR**HAEGTFTS** **DVSSYLEGQAAKEFIAWLVKGRG**
62	GOP_J	hGH	hGLP-1, hOXM, PYY	MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNAMLRARKKR**HAEGTFT** **SDVSSYLEGQAAKEFIAWLVKGRG**RRKRHSQGTFTSDYSKYLDSRRAQDFVQWL MNTKRNRNNIARKKR***IKPEAPREDASPEELNRYYASLRHYLNLVTRQRY***
63	GOP_F	hInsul	hGLP-1, hOXM, PYY	MALWMRLLPLLALLALWGPDPAAAYPYDVPDYARKKR**HAEGTFTSDVSSYLEGQ** **AAKEFIAWLVKGRG**RRKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNI ARKKR***IKPEAPREDASPEELNRYYASLRHYLNLVTRQRY***
64	GOP_L	mGHRH	hGLP-1, hOXM, PYY	MLLWVLFVILILTSGSHCSLPPSPPFRMQR**HAEGTFTSDVSSYLEGQAAKEFIA** **WLVKGRG**RRKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIARKKR***IK*** ***PEAPREDASPEELNRYYASLRHYLNLVTRQRY***
65	GOP_M	hAlbumin	hGLP-1, hOXM, PYY	MKWVTFISLLFLFSSAYSRGVFRR**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR** **G**RRKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIARKKR***IKPEAPRE*** ***DASPEELNRYYASLRHYLNLVTRQRY***
66	GGG_A	hGH	hGLP-1, hGlucagon, hGIP	MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNAMLRARKKR**HAEGTFT** **SDVSSYLEGQAAKEFIAWLVKGRG**RRKRHSQGTFTSDYSKYLDSRRAQDFVQWL MNTRKKRYAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ
75	3xGLP-1 w/o signal peptide	n/a	hGLP-1, hGLP-1, hGLP-1,	**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RKKR**HAEGTFTSDVSSYLEGQAA** **KEFIAWLVKGRG**RKKR**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**

SEQ ID NO	Construct name	Signal peptide	Gut peptides	Sequence
76	hGLP-1, hOXM, PYY w/o signal peptide	n/a	hGLP-1, hOXM, PYY	**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RRKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIARKKR***IKPEAPREDASPEELNRYYASLRHYLNLVTRQRY***
77	hGLP-1, hGlucagon, hGIP w/o signal peptide	n/a	hGLP-1, hGlucagon, hGIP	**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**RRKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTRKKRYAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ

Table 10. Nucleic acid sequences of exemplary tricistronic expression constructs. Signal peptide encoding sequence is shown in italic upper-case letters. Furin cleavage site encoding sequence is shown in lower case letters. Gut peptide encoding sequences are shown in upper case letters (not italic).

SEQ ID NO	Construct name	Signal peptide	Gut peptides	Sequence
67	3xGLP-1_3xB	hAlbumin	hGLP-1, hGLP-1, hGLP-1,	*ATGAAGTGGGTAACCTTTATTTCCCTTCTTTTTCTCTTTAGCTCGGCTTATTCC*aggggtgtgtttcgtcgaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaCATGCAGAGGGAACATTTACTAGTGATGTCAGTTCATATCTTGAGGGACAAGCTGCTAAAGAATTTATTGCTTGGCTTGTGAAGGGAAGAGGAagaaagaagagaCATGCTGAAGGGACATTTACCTCAGATGTTTCTTCATACCTGGAAGGACAGGCTGCCAAGGAATTTATTGCATGGCTTGTGAAAGGCAGGGGCtga
68	3xGLP-1_3xC	hInsul	hGLP-1, hGLP-1, hGLP-1,	*ATGGCCCTGTGGATGCGCCTCCTGCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCAGCCGCAGCC*tacccatacgatgttccagattacgctagaaaaaagagaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaCATGCAGAGGGAACATTTACTAGTGATGTCAGTTCATATCTTGAGGGACAAGCTGCTAAAGAATTTATTGCTTGGCTTGTGAAGGGAAGAGGAagaaagaagagaCAT

SEQ ID NO	Construct name	Signal peptide	Gut peptides	Sequence
				GCTGAAGGGACATTTACCTCAGATGTTTCTTCATACCTGGAAGGACAGGCTGCC AAGGAATTTATTGCATGGCTTGTGAAAGGCAGGGGCtga
69	3xGLP-1_3xD	mGHRH	hGLP-1, hGLP-1, hGLP-1,	*ATGCTGCTCTGGGTGCTCTTTGTGATCCTCATCCTCACCAGTGGCTCCCACTGC TCA*ActgcccccctcacctcccttcaggatgcagcgaCATGCTGAAGGGACCTTT ACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCT TGGCTGGTGAAAGGCCGAGGAagaaaaaagagaCATGCAGAGGGAACATTTACT AGTGATGTCAGTTCATATCTTGAGGGACAAGCTGCTAAAGAATTTATTGCTTGG CTTGTGAAGGGAAGAGGAagaaagaagagaCATGCTGAAGGGACATTTACCTCA GATGTTTCTTCATACCTGGAAGGACAGGCTGCCAAGGAATTTATTGCATGGCTT GTGAAAGGCAGGGGCtga
70	GOP_J	hGH	hGLP-1, hOXM, PYY	*ATGGCTACAGGCTCCCGGACGTCCCTGCTCCTGGCTTTTGGCCTGCTCTGCCTG CCCTGGCTTCAAGAGGGCAGTGCC*ttcccaaccattcccttatccaggcttttt gacaacgctatgctccgcgccagaaaaaagagaCATGCTGAAGGGACCTTTACC AGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGG CTGGTGAAAGGCCGAGGAagacgtaagaggCATTCACAGGGCACATTCACCAGT GACTACAGCAAGTATCTGGACTCCAGGCGTGCCCAAGATTTTGTGCAGTGGTTG ATGAATACCAAGAGGAACAGGAATAACATTGCCagaaaaaagagaATCAAACCC GAGGCTCCCCGCGAAGACGCCTCGCCGGAGGAGCTGAACCGCTACTACGCCTCC CTGCGCCACTACCTCAACCTGGTCACCCGGCAGCGGTATTGA
71	GOP_F	hInsul	hGLP-1, hOXM, PYY	*ATGGCCCTGTGGATGCGCCTCCTGCCCCTGCTGGCGCTGCTGGCCCTCTGGGGA CCTGACCCAGCCGCAGCC*tacccatacgatgttccagattacgctagaaaaaag agaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAA GCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagacgtaagagg CATTCACAGGGCACATTCACCAGTGACTACAGCAAGTATCTGGACTCCAGGCGT GCCCAAGATTTTGTGCAGTGGTTGATGAATACCAAGAGGAACAGGAATAACATT GCCagaaaaaagagaATCAAACCCGAGGCTCCCCGCGAAGACGCCTCGCCGGAG GAGCTGAACCGCTACTACGCCTCCCTGCGCCACTACCTCAACCTGGTCACCCGG CAGCGGTATTGA
72	GOP_L	mGHRH	hGLP-1, hOXM, PYY	*ATGCTGCTCTGGGTGCTCTTTGTGATCCTCATCCTCACCAGTGGCTCCCACTGC TCA*ActgcccccctcacctcccttcaggatgcagcgaCATGCTGAAGGGACCTTT ACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCT

SEQ ID NO	Construct name	Signal peptide	Gut peptides	Sequence
				TGGCTGGTGAAAGGCCGAGGAagacgtaagaggCATTCACAGGGCACATTCACC AGTGACTACAGCAAGTATCTGGACTCCAGGCGTGCCCAAGATTTTGTGCAGTGG TTGATGAATACCAAGAGGAACAGGAATAACATTGCCagaaaaaagagaATCAAA CCCGAGGCTCCCCGCGAAGACGCCTCGCCGGAGGAGCTGAACCGCTACTACGCC TCCCTGCGCCACTACCTCAACCTGGTCACCCGGCAGCGGTATTGA
73	GOP_M	hAlbumin	hGLP-1, hOXM, PYY	*ATGAAGTGGGTAACCTTTATTTCCCTTCTTTTTCTCTTTAGCTCGGCTTATTCC* aggggtgtgtttcgtcgaCATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCT TATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGA GGAagacgtaagaggCATTCACAGGGCACATTCACCAGTGACTACAGCAAGTAT CTGGACTCCAGGCGTGCCCAAGATTTTGTGCAGTGGTTGATGAATACCAAGAGG AACAGGAATAACATTGCCagaaaaaagagaATCAAACCCGAGGCTCCCCGCGAA GACGCCTCGCCGGAGGAGCTGAACCGCTACTACGCCTCCCTGCGCCACTACCTC AACCTGGTCACCCGGCAGCGGTATTGA
74	GGG_A	hGH	hGLP-1, hGlucagon, hGIP	*ATGGCTACAGGCTCCCGGACGTCCCTGCTCCTGGCTTTTGGCCTGCTCTGCCTG* *CCCTGGCTTCAAGAGGGCAGTGCC*ttcccaaccattcccttatccaggcttttt gacaacgctatgctccgcgccagaaaaaagagaCATGCTGAAGGGACCTTTACC AGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGG CTGGTGAAAGGCCGAGGAagacgtaagaggCATTCACAGGGCACATTCACCAGT GACTACAGCAAGTATCTGGACTCCAGGCGTGCCCAAGATTTTGTGCAGTGGTTG ATGAATACCagaaaaaagagaTACGCGGAAGGGACTTTCATCAGTGACTACAGT ATTGCCATGGACAAGATTCACCAACAAGACTTTGTGAACTGGCTGCTGGCCCAA AAGGGGAAGAAGAATGACTGGAAACACAACATCACCCAGTGA
78	3xGLP-1 w/o signal peptide	n/a	hGLP-1, hGLP-1, hGLP-1,	CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCT GCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagaaaaaagagaCAT GCAGAGGGAACATTTACTAGTGATGTCAGTTCATATCTTGAGGGACAAGCTGCT AAAGAATTTATTGCTTGGCTTGTGAAGGGAAGAGGA agaaagaagagaCATGCTGAAGGGACATTTACCTCAGATGTTTCTTCATACCTG GAAGGACAGGCTGCCAAGGAATTTATTGCATGGCTTGTGAAAGGCAGGGGCtga
79	hGLP-1, hOXM, PYY w/o	n/a	hGLP-1, hOXM, PYY	CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCT GCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagacgtaagaggCAT TCACAGGGCACATTCACCAGTGACTACAGCAAGTATCTGGACTCCAGGCGTGCC

SEQ ID NO	Construct name	Signal peptide	Gut peptides	Sequence
	signal peptide			CAAGATTTTGTGCAGTGGTTGATGAATACCAAGAGGAACAGGAATAACATTGCC agaaaaaagagaATCAAACCCGAGGCTCCCCGCGAAGACGCCTCGCCGGAGGAG CTGAACCGCTACTACGCCTCCCTGCGCCACTACCTCAACCTGGTCACCCGGCAG CGGTATTGA
80	hGLP-1, hGlucagon, hGIP w/o signal peptide	n/a	hGLP-1, hGlucagon, hGIP	CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCT GCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGGAagacgtaagaggCAT TCACAGGGCACATTCACCAGTGACTACAGCAAGTATCTGGACTCCAGGCGTGCC CAAGATTTTGTGCAGTGGTTGATGAATACCagaaaaaagagaTACGCGGAAGGG ACTTTCATCAGTGACTACAGTATTGCCATGGACAAGATTCACCAACAAGACTTT GTGAACTGGCTGCTGGCCCAAAAGGGGAAGAAGAATGACTGGAAACACAACATC ACCCAGTGA

162027.52376

[0122] *Furin recognition and cleavage sequences*

[0123] Provided herein are expression constructs that encode one or more sequences that are recognized by a protease. In embodiments, the protease is furin. Furin cleaves proteins just downstream of a basic amino acid minimal furin cleavage site. In embodiments, this minimal furin cleavage site is Arg-X-X-Arg (preferably, Arg-X-(Arg/Lys)-Arg). However, furin may recognize a longer sequence within the target polypeptide in addition to the minimal furin cleavage site. This longer sequence (comprising the minimal furin cleavage site) is referred to herein as a "furin recognition and cleavage sequence." The inclusion of a furin recognition and cleavage sequence can promote the functional N-terminus of expressed polypeptide (such as a gut peptide or a polyprotein comprising one or more gut peptides) to be fully processed and generated in non-endocrine cells.

[0124] In some embodiments, the furin recognition and cleavage sequence comprises (1) any one of SEQ ID NOs:89, 92-96 or (2) a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOs:89, 92-96. In some embodiments, the furin recognition and cleavage sequence comprises a portion of (1) any one of SEQ ID NOs:89, 92-96 or (2) a sequence that is at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOs:89, 92-96.

[0125] In embodiments, provided is an expression construct comprising a sequence encoding any one of SEQ ID NOs:89, 92-96 or a sequence that is least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to any one of SEQ ID NOs:89, 92-96.

[0126] In embodiments, provided is an expression construct comprising a sequence encoding any one of sequences RKKR (SEQ ID NO:97), RMQR (SEQ ID NO:98), VFRR (SEQ ID NO:99), or RKKR (SEQ ID NO:100).

[0127] In embodiments, the monocistronic, bicistronic, or tricistronic expression construct comprises a sequence disclosed in **Table 11** or a portion of a sequence disclosed in **Table 11.** In embodiments, the monocistronic, bicistronic, or tricistronic expression construct comprises a sequence encoding a sequence disclosed in **Table 11** or a portion of a sequence disclosed in **Table 11.**

[0128] Provided herein are expression constructs comprising one or more sequences that are least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to of any of the sequences disclosed herein. Provided herein are expression constructs comprising one or more sequences that comprise a portion of any of the sequences disclosed herein.

[0129] Provided herein are expression constructs encoding for one or more sequences that are least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 91%, at least 97%, at least 98%, or at least 99% identical to of any of the sequences disclosed herein. Provided herein are expression constructs encoding for one or more sequences that comprise a portion of any of the sequences disclosed herein.

Table 11. Selected sequences that facilitate furin cleavage.

SEQ ID NO	Nucleic Acid Sequence	SEQ ID NO	Amino Acid Sequence
81	TACCCATACGATGTTCCAGATTACGCTAGAAAAAAGAGA	89	YPYDVPDYARKKR
84	TTCCCAACCATTCCCTTATCCAGGCTTTTTGACAACGCTATGCTCCGCGCCAGAAAAAAGAGA	92	FPTIPLSRLFDNAMLRARKKR
85	GCTCCCCCACGCCTCATCTGCGACAGTCGAGTTCTGGAGAGGTACAGAAAAAAGAGA	93	APPRLICDSRVLERYRKKR
86	CTGCCCCCCTCACCTCCCTTCAGGATGCAGCGA	94	LPPSPPFRMQR
87	AGGGGTGTGTTTCGTCGA	95	RGVFRR
88	ACAGTTTTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGG	96	TVFLDHENANKILNRPKR
		97	RKKR
		98	RMQR
		99	VFRR
		100	RPKR

[0130] *Leader sequences*

[0131] In embodiments, the expression constructs disclosed herein comprise a leader sequence. As used herein, a “leader sequence” is a sequence that comprises (1) a signal peptide and a protease recognition and cleavage sequence and/or (2) a signal peptide and a minimal protease cleavage site. In embodiments, a leader sequence comprises (1) a signal peptide and a furin recognition and cleavage sequence and/or (2) a signal peptide and a minimal furin cleavage site. The inclusion of

a furin recognition and cleavage sequence in the leader sequence can promote the functional N-terminus of expressed polypeptide (such as a gut peptide or a polyprotein comprising one or more gut peptides) to be fully processed and generated in non-endocrine cells.

[0132] The leader sequence may be derived from a naturally occurring, secreted polypeptide or from a variant of a naturally occurring, secreted polypeptide. In embodiments, the leader sequence, or a portion thereof, is derived from influenza virus hemagglutinin, human growth hormone, murine growth hormone-releasing hormone, or human albumin.

[0133] *Promoters*

[0134] Any suitable promoter may be used in the expression constructs disclosed herein. In some embodiments, the promoter is a CMV or a CASI promoter.

[0135] Regulation of gene expression

[0136] In some embodiments, the expression constructs disclosed herein provide for constitutive expression of the polypeptides disclosed herein.

[0137] In some embodiments, the expression constructs disclosed herein provide for regulatable expression of the polypeptides disclosed herein.

[0138] In embodiments, the sequence encoding a polypeptide disclosed herein comprises a riboswitch comprising an aptamer, wherein the riboswitch is operable linked to the sequence encoding the polypeptide. In embodiments the sequence encoding a polypeptide disclosed herein comprises a gene regulation cassette, wherein the gene regulation cassette comprises an aptamer. In embodiments the polypeptide is a polyprotein disclosed herein.

[0139] Aptamers are single-stranded nucleic acid molecules that non-covalently bind to specific ligands with high affinity and specificity by folding into three-dimensional structures. Aptamer ligands include ions, small molecules, proteins, viruses, and cells. Aptamer ligands can be, for example, an organic compound, amino acid, steroid, carbohydrate, or nucleotide. Non-limiting examples of small molecule aptamer ligands include antibiotics, therapeutics, dyes, cofactors, metabolites, molecular markers, neurotransmitters, pollutants, toxins, food adulterants, carcinogens, drugs of abuse. As such, aptamers are useful for the detection of small molecules. Application of small-molecule detection by aptamers include environmental monitoring, food safety, medicine (including diagnostics), microbiology, analytical chemistry, forensic science, agriculture, and basic biology research. The term “aptamer” as used herein refers to an RNA

polynucleotide (or DNA sequence encoding the RNA polynucleotide) that specifically binds to a class of ligands. The term "ligand" refers to a molecule that is specifically bound by an aptamer. Aptamers have binding regions that are capable of forming complexes with an intended target molecule (*i.e.*, the ligand). An aptamer will typically be between about 15 and about 200 nucleotides in length. More commonly, an aptamer will be between about 30 and about 100 nucleotides in length, for example, 70 to 90 nucleotides in length. Aptamers typically comprise multiple paired (P) regions in which the aptamer forms a stem and unpaired regions where the aptamer forms a joining (J) region or a loop (L) region. The paired regions can be numbered sequentially starting at the 5′ end (P1) and numbering each stem sequentially (P2, P3, etc.). The loops (L1, L2, etc.) are numbered based on the adjacent paired region and the joining regions are numbered according to the paired regions that they link. Aptamers are oligonucleotides that bind to a target ligand with high affinity and specificity.

[0140] In embodiments, the presence of a small molecule that binds to an aptamer leads to an increase in expression of a sequence encoding a polypeptide disclosed herein as compared to the expression of the sequence encoding a polypeptide disclosed herein in absence of the small molecule. In such an embodiment, the aptamer constitutes an "on" switch. In embodiments, the expression of a sequence encoding a polypeptide disclosed herein is increased by at least 3-fold, by at least 5-fold, by at least 10-fold, by at least 15-fold, by at least 20-fold, by at least 25-fold, by at least 30-fold, by at least 40-fold, by at least 50-fold, by at least 100-fold, by at least 1000-fold, or by at least 10,000-fold in presence of the small molecule that binds to an aptamer as compared to in absence of the small molecule. In embodiments, the expression of a sequence encoding a polypeptide disclosed herein is increased by between 2-fold and 10-fold, between 5-fold and 10-fold, between 5-fold and 15-fold, between 5-fold and 20-fold, between 5-fold and 25-fold, between 5-fold and 30-fold, between 10-fold and 20-fold, between 10-fold and 30-fold, between 10-fold and 40-fold, between 10-fold and 50-fold, between 10-fold and 100-fold, between 10-fold and 500-fold, between 10-fold and 1,000-fold, between 50-fold and 100-fold, between 50-fold and 500-fold, between 50-fold and 100-fold, between 50-fold and 1,000-fold, between 100-fold and 1,000-fold, or between 100-fold and 10,000-fold in presence of the small molecule that binds to an aptamer as compared to in absence of the small molecule.

[0141] In embodiments, the presence of a small molecule that binds to an aptamer leads to a decrease in expression of a sequence encoding a polypeptide disclosed herein as compared to the

expression of the sequence encoding a polypeptide disclosed herein in the absence of the small molecule. In such embodiments, the aptamer constitutes an "off" switch. In embodiments, the expression of the sequence encoding a polypeptide disclosed herein is decreased by at least 3-fold, by at least 5-fold, by at least 10-fold, by at least 15-fold, by at least 20-fold, by at least 25-fold, by at least 30-fold, by at least 40-fold, by at least 50-fold, by at least 100-fold, by at least 1000-fold, or by at least 10,000-fold in presence of the small molecule that binds to an aptamer as compared to in absence of the small molecule. In one embodiment, the expression of the sequence encoding a polypeptide disclosed herein is decreased by between 2-fold and 10-fold, between 5-fold and 10-fold, between 5-fold and 15-fold, between 5-fold and 20-fold, between 5-fold and 25-fold, between 5-fold and 30-fold, between 10-fold and 20-fold, between 10-fold and 30-fold, between 10-fold and 40-fold, between 10-fold and 50-fold, between 10-fold and 100-fold, between 10-fold and 500-fold, between 10-fold and 1,000-fold, between 50-fold and 100-fold, between 50-fold and 500-fold, between 50-fold and 100-fold, between 50-fold and 1,000-fold, between 100-fold and 1,000-fold, or between 100-fold and 10,000-fold in presence of the small molecule that binds to an aptamer as compared to in absence of the small molecule.

[0142] In embodiments, the aptamer is part of a riboswitch. Riboswitches are regulatory segments of an RNA polynucleotide that regulate the stability of the RNA polynucleotide and/or regulate the production of a protein from the RNA polynucleotide in response to the presence or absence of aptamer-specific ligand molecules. In embodiments, the riboswitch comprises a sensor region (*e.g.*, the aptamer region) and an effector region that together are responsible for sensing the presence of a ligand (*e.g.*, a small molecule) and causing an effect that leads to increased or decreased expression of the sequence encoding a polypeptide disclosed herein. The riboswitches described herein are recombinant, utilizing polynucleotides from two or more sources. In embodiments, the sensor and effector regions are joined by a polynucleotide linker. In embodiments, the polynucleotide linker forms an RNA stem or paired region (*i.e.*, a region of the RNA polynucleotide that is double-stranded). In embodiments, the paired region linking the aptamer to the effector region comprises all, or some of an aptamer stem (*e.g.*, for example all, or some of the aptamer P1 stem.).

[0143] Riboswitches comprising aptamer sequences may be used, for example, to control the formation of rho-independent transcription termination hairpins leading to premature transcription termination. Riboswitches comprising aptamer sequences may also induce structural changes in

the RNA, leading to sequestration for the ribosome binding site and inhibition of translation. Alternative riboswitch structures comprising the aptamer sequences disclosed herein can further affect the splicing of mRNA in response to the presence of the small molecule ligand.

[0144] *Alternative splicing riboswitch*

[0145] In one embodiment, the riboswitches described herein are encoded as part of a gene regulation cassette for the regulation of a sequence encoding a polypeptide disclosed herein by aptamer/ligand mediated alternative splicing of the resulting RNA (*e.g.*, pre-mRNA). In this context, the gene regulation cassette comprises a riboswitch comprising a sensor region (*e.g.*, the aptamers described herein) and an effector region that together are responsible for sensing the presence of a small molecule ligand and altering splicing to an alternative exon. Splicing refers to the process by which an intronic sequence is removed from the nascent pre-messenger RNA (pre-mRNA) and the exons are joined together to form the mRNA. Splice sites are junctions between exons and introns and are defined by different splice site consensus sequences at the 5′ and 3′ ends of the intron (*i.e.*, the splice donor and splice acceptor sites, respectively). Splicing is carried out by a large multi-component structure called the spliceosome, which is a collection of small nuclear ribonucleoproteins (snRNPs) and a diverse array of auxiliary proteins. By recognizing various cis regulatory sequences, the spliceosome defines exon/intron boundaries, removes intronic sequences, and splices together the exons into a final message (*e.g.*, the mRNA). In the case of alternative splicing, certain exons can be included or excluded to vary the final coding message thereby changing the resulting expressed protein.

[0146] In one embodiment, the regulation of a sequence encoding a polypeptide disclosed herein expression is achieved by using any of the DNA constructs disclosed in PCT Patent Publication WO2016/126747, which is hereby incorporated by reference in its entirety. In embodiments of the present disclosure, the riboswitches and polynucleotide cassettes disclosed in PCT Patent Publication WO2016/126747 comprise an aptamer sequence described herein in place of the aptamer sequence disclosed in PCT Patent Publication WO2016/126747.

[0147] In one embodiment, the polynucleotide cassette comprises (a) a riboswitch and (b) an alternatively-spliced exon, flanked by a 5′ intron and a 3′ intron, wherein the riboswitch comprises (i) an effector region comprising a stem forming sequence that includes the 5′ splice site sequence of the 3′ intron and sequence complementary to the 5′ splice site sequence of the 3′ intron, and (ii) an aptamer. In embodiments, the effector region comprises the intronic 5′ splice site (“5′ ss”)

sequence of the intron that is immediately 3′ of the alternative exon, as well as the sequence complimentary to the 5′ ss sequence of the 3′ intron. When the aptamer binds its ligand, the effector region forms a stem and thus prevents splicing to the splice donor site at the 3′ end of the alternative exon. Under certain conditions (for example, when the aptamer is not bound to its ligand), the effector region is in a context that provides access to the splice donor site at the 3′ end of the alternative exon, leading to inclusion of the alternative exon in the mRNA of the sequence encoding a polypeptide disclosed herein. In some embodiments, the polynucleotide cassette is placed in the sequence encoding a polypeptide disclosed herein gene to regulate expression of the sequence encoding a polypeptide disclosed herein in response to a ligand. In one embodiment, the alternatively-spliced exon comprises a stop codon that is in-frame with the sequence encoding a polypeptide disclosed herein when the alternatively-spliced exon is spliced into the mRNA of the sequence encoding a polypeptide disclosed herein.

[0148] In one embodiment, the gene regulation cassette comprises the sequence of SEQ ID NO:101, wherein -X- represents an aptamer sequence. Lower case letters indicate paired stem sequence linking the aptamer to the remainder of the riboswitch. In one embodiment, the alternative exon (underlined in SEQ ID NO:101, below) is replaced with another alternative exon sequence.

[0149] SEQ ID NO:101-

GTGAGTCTATGGGACCCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATG
TCATAGGAAGGGGAGAAGTAACAGGGTACACATATTGACCAAATCAGGGTAATTTT
GCATTTGTAATTTTAAAAAATGCTTTCTTCTTTTAATATACTTTTTTGTTTATCTTATT
TCTAATACTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCATGCC
GAGTAACGCTGTTTCTCTAACTTGTAG<u>GAATGAATTCAGATATTTCCAGAGAATGAA
AAAAAAATCTTCAGTAGAAG</u>gtaatgt-**X**-
acattacGCACCATTCTAAAGAATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAAT
ATTTCTGCATATAAATATTTCTGCATATAAATTGTAACTGATGTAAGAGGTTTCATAT
TGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTTATGGTTGGGATAAG
GCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCATACCTCTTAT
CTTCCTCCCACAG.

[0150] The alternative exon is flanked by 5′ and 3′ intronic sequences. The 5′ and 3′ intronic sequences that can be used in the gene regulation cassettes disclosed herein can be any sequence

that can be spliced out of the sequence encoding a polypeptide disclosed herein creating either the mRNA of the sequence encoding a polypeptide disclosed herein or the sequence encoding a polypeptide disclosed herein comprising the alternative exon in the mRNA, depending upon the presence or absence of a ligand that binds the aptamer. The 5′ and 3′ intronic sequences each have the sequences necessary for splicing to occur, *i.e.*, splice donor, splice acceptor and branch point sequences. In one embodiment, the 5′ and 3′ intronic sequences of the gene regulation cassette are derived from one or more naturally occurring introns or portions thereof. In one embodiment, the 5′ and 3′ intronic sequences are derived from a truncated human beta-globin intron 2 (IVS2Δ), from intron 2 of the human beta-globin gene, from the SV40 mRNA intron (used in pCMV-LacZ vector from Clontech Laboratories, Inc.), from intron 6 of human triose phosphate isomerase (TPI) gene (Nott Ajit, et al. RNA. 2003, 9:6070617), from an intron from human factor IX (Sumiko Kurachi, et al. J. Bio. Chem. 1995, 270(10), 5276), or from any genomic fragment or synthetic introns (Yi Lai, et al. Hum Gene Ther. 2006:17(10): 1036) that contain elements that are sufficient for regulated splicing (Thomas A. Cooper, Methods 2005 (37):331).

[0151] The splice donor and splice acceptor sites in the alternative splicing gene regulation cassette can be modified to be strengthened or weakened. That is, the splice site sequences can be modified to be closer to the consensus for a splice donor or acceptor by standard cloning methods, site directed mutagenesis, and the like. Splice site sequences that are more similar to the splice consensus sequence tend to promote splicing and are thus strengthened. Splice site sequences that are less similar to the splice consensus sequence tend to hinder splicing and are thus weakened. The consensus for the splice donor of the most common class of introns (U2) is A/C A G||G T A/G A G T (SEQ ID NO:102, where || denotes the exon/intron boundary). The consensus for the splice acceptor is C A G||G (where || denotes the exon/intron boundary). The frequency of particular nucleotides at the splice donor and acceptor sites are described in the art (*see, e.g.*, Zhang, M. Q., Hum Mol Genet. 1988. 7(5):919-932). The strength of 5′ and 3′ splice sites can be adjusted to modulate splicing of the alternative exon.

[0152] Additional modifications to 5′ and 3′ introns present in the alternative splicing gene regulation cassette that can be made to modulate splicing include modifying, deleting, and/or adding intronic splicing enhancer elements, intronic splicing suppressor elements and or splice sites, and/or modifying the branch site sequence.

[0153] In one embodiment, the 5′ intron has been modified to contain a stop codon that will be in frame with the sequence encoding a polypeptide disclosed herein. The 5′ and 3′ intronic sequences can also be modified to remove cryptic slice sites, which can be identified with publicly available software (*see, e.g.*, Kapustin, Y. et al. Nucl. Acids Res. 2011. 1-8).

[0154] The lengths of the 5′ and 3′ intronic sequences can be adjusted in order to, for example, meet the size requirements for viral expression constructs. In one embodiment, the 5′ and/or 3′ intronic sequences are about 50 to about 300 nucleotides in length. In one embodiment, the 5′ and/or 3′ intronic sequences are about 125 to about 240 nucleotides in length.

[0155] The stem portion of the effector region should be of a sufficient length (and GC content) to substantially prevent alternative splicing of the alternative exon upon ligand binding the aptamer, while also allowing access to the splice site when the ligand is not present in sufficient quantities. In embodiments, the stem portion of the effector region comprises a stem sequence in addition to the 5′ splice site sequence of the 3′ intron and its complementary sequence of the 5′ splice site sequence. In embodiments, this additional stem sequence comprises a sequence from the aptamer stem. The length and sequence of the stem portion can be modified using known techniques in order to identify stems that allow acceptable background expression of the sequence encoding a polypeptide disclosed herein when no ligand is present and acceptable expression levels of the sequence encoding a polypeptide disclosed herein when the ligand is present. In one embodiment, the effector region stem of the riboswitch is about 7 to about 20 base pairs in length. In one embodiment, the effector region stem is 8 to 11 base pairs in length. In addition to the length of the stem, the GC base pair content of the stem can be altered to modify the stability of the stem.

[0156] In one embodiment, the alternative exon that is part of the alternative splicing gene regulation cassettes disclosed herein is a polynucleotide sequence capable of being transcribed to a pre-mRNA and alternatively spliced into the mRNA of the sequence encoding a polypeptide disclosed herein. In one embodiment, the alternative exon contains at least one sequence that inhibits translation such that when the alternative exon is included in the mRNA of the sequence encoding a polypeptide disclosed herein, expression of the sequence encoding a polypeptide disclosed herein from that mRNA is prevented or reduced. In a preferred embodiment, the alternative exon contains a stop codon (TGA, TAA, TAG) that is in frame with the sequence encoding a polypeptide disclosed herein when the alternative exon is included in the mRNA of the sequence encoding a polypeptide disclosed herein by splicing. In embodiments, the alternative

exon comprises, in addition to a stop codon, or as an alternative to a stop codon, another sequence that reduces or substantially prevents translation when the alternative exon is incorporated by splicing into the mRNA of the sequence encoding a polypeptide disclosed herein including, *e.g.*, a microRNA binding site, which leads to degradation of the mRNA. In one embodiment, the alternative exon comprises a miRNA binding sequence that results in degradation of the mRNA. In one embodiment, the alternative exon encodes a polypeptide sequence which reduces the stability of the protein containing this polypeptide sequence. In one embodiment, the alternative exon encodes a polypeptide sequence which directs the protein containing this polypeptide sequence for degradation.

[0157] The basal or background level of splicing of the alternative exon can be optimized by altering exon splice enhancer (ESE) sequences and exon splice suppressor (ESS) sequences and/or by introducing ESE or ESS sequences into the alternative exon. Such changes to the sequence of the alternative exon can be accomplished using methods known in the art, including, but not limited to site directed mutagenesis. Alternatively, oligonucleotides of a desired sequence (*e.g.*, comprising all or part of the alternative exon) can be obtained from commercial sources and cloned into the gene regulation cassette. Identification of ESS and ESE sequences can be accomplished by methods known in the art, including, for example using ESEfinder 3.0 (Cartegni, L. et al. ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acid Research, 2003, 31(13): 3568-3571) and/or other available resources.

[0158] In one embodiment, the alternative exon is a naturally-occurring exon. In another embodiment, the alternative exon is derived from all or part of a known exon. In this context, “derived” refers to the alternative exon containing sequence that is substantially homologous to a naturally occurring exon, or a portion thereof, but may contain various mutations, such a mutations generated by altering exon splice enhancer (ESE) sequences and exon splice suppressor (ESS) sequences and/or by introducing ESE or ESS sequences into the alternative exon. “Homology” and “homologous” as used herein refer to the percent of identity between two polynucleotide sequences or between two polypeptide sequences. The correspondence between one sequence to another can be determined by techniques known in the art. For example, homology can be determined by a direct comparison of two polypeptide molecules by aligning their sequences and using readily available computer programs. Alternatively, homology can be determined by hybridization of polynucleotides under conditions which form stable duplexes between

homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. Two polynucleotide or two polypeptide sequences are "substantially homologous" to each other when, after optimally aligned with appropriate insertions or deletions, at least about 80%, at least about 85%, at least about 90%, and at least about 95% of the nucleotides or amino acids, respectively, match over a defined length of the molecules, as determined using the methods above.

[0159] In one embodiment, the alternative exon is exogenous to the sequence encoding a polypeptide disclosed herein, although it may be derived from a sequence originating from the organism where the sequence encoding a polypeptide disclosed herein will be expressed. As used herein, "exogenous" means derived from a genotypically distinct entity from that of the rest of the entity to which it is compared or into which it is introduced or incorporated. For example, a polynucleotide introduced by genetic engineering techniques into a different cell type is a heterologous polynucleotide (and, when expressed, can encode a heterologous polypeptide). In one embodiment, the alternatively-spliced exon is derived from exon 2 of the human dihydrofolate reductase gene (DHFR), mutant human Wilms tumor 1 exon 5, mouse calcium/calmodulin-dependent protein kinase II delta exon 16, or SIRT1 exon 6. In embodiments, the alternatively-spliced exon is, or comprises, the modified DHFR exon 2 in SEQ ID NO:103 (GAATGAATTCAGATATTTCCAGAGAATGAAAAAAAAATCTTCAGTAGAAG). In embodiments, the alternatively-spliced exon is, or comprises, the modified DHFR exon 2 in SEQ ID NO:104 (GAATGAATTCAGATATTTCCAGAGAATGAAAAAAAATCTTCAGTAGAAG).

[0160] *Aptamer-mediated cleavage by self-cleaving ribozymes*

[0161] In one embodiment, the aptamer-mediated expression of the sequence encoding a polypeptide disclosed herein is regulated by an aptamer-mediated modulation of small endonucleolytic ribozymes. A ribozyme is an RNA enzyme that catalyzes a chemical reaction. In the nucleic acids and methods disclosed herein, a ribozyme may be any small endonucleolytic ribozyme that will self-cleave in the target cell type including, but not limited to a hammerhead, hairpin, the hepatitis delta virus, the Varkud satellite, twister, twister sister, pistol or hatchet ribozyme. Accordingly, in one embodiment, provided is a riboswitch, and a gene expression cassette comprising the riboswitch that contains a ribozyme linked to an aptamer. WO2017/136608, which is incorporated in its entirety by reference herein, describes such

riboswitches that activate ribozyme self-cleavage in the presence of aptamer ligand ("off" switch) or riboswitches that inhibit ribozyme self-cleavage in the presence of aptamer ("on" switch).

[0162] In an "off" switch scenario, aptamer/ligand binding increases the ribonuclease function of the ribozyme, leading to cleavage of the RNA of the sequence encoding a polypeptide disclosed herein that contains the polynucleotide cassette, thereby reducing expression of the sequence encoding a polypeptide disclosed herein. Examples of such an off switch include a polynucleotide cassette for the regulation of the expression of a sequence encoding a polypeptide disclosed herein comprising a riboswitch that comprises a twister ribozyme linked by a stem to an aptamer, wherein the stem linking the twister ribozyme to the aptamer attaches to the ribozyme at the location of the P3 stem of the twister ribozyme and wherein the sequence encoding a polypeptide disclosed herein is linked to the P1 stem of the twister ribozyme (see, *e.g.* Figs. 1a, 1b, or 3a of WO2017/136608 and the associated text, incorporated herein by reference).

[0163] In an "on" switch scenario, aptamer/ligand binding inhibits the ribonuclease function of the ribozyme, decreasing cleavage of the RNA of the sequence encoding a polypeptide disclosed herein that contains the polynucleotide cassette, thereby increasing expression of the sequence encoding a polypeptide disclosed herein in the presence of ligand. Examples of an on switch include a riboswitch that comprises a twister ribozyme linked to an aptamer, wherein the aptamer is linked to the 3′ or 5' end of the twister ribozyme P1 stem, wherein when the aptamer is linked to the 3' end of the twister ribozyme P1 stem, a portion of the 3' arm of the twister ribozyme P1 stem is alternatively the 5' arm of the aptamer P1 stem, and wherein when the aptamer is linked to the 5' end of the twister ribozyme P1 stem, a portion of the 5' arm of the twister ribozyme P1 stem is alternatively the 3' arm of the aptamer P1 stem (see, *e.g.*, Figs. 6a-6b of WO2017/136608 and the associated text, incorporated herein by reference).

[0164] *Aptamer modulation of polyadenylation*

[0165] In embodiments, the expression of a sequence encoding a polypeptide disclosed herein is regulated by aptamer-modulated polyadenylation. The 3' end of almost all eukaryotic mRNAs comprises a poly(A) tail—a homopolymer of 20 to 250 adenosine residues. Because addition of the poly(A) tail to mRNA protects it from degradation, expression of a gene can be influenced by modulating the polyadenylation the corresponding mRNA.

[0166] In one embodiment, the expression of the sequence encoding a polypeptide disclosed herein is regulated through aptamer-modulated accessibility of polyadenylation signals as

described in and WO2018/156658, which is incorporated in its entirety by reference herein. In such embodiments, the riboswitch comprises an effector stem-loop and an aptamer, wherein the effector stem-loop comprises a polyadenylation signal, and wherein the aptamer and effector stem-loop are linked by an alternatively shared stem arm comprising a sequence that is complementary to the unshared arm of the aptamer stem and to the unshared arm of the effector stem loop (see, *e.g.*, Figs 1a, 1b, 2a, and 5a of WO2018/156658 and the associated text, incorporated herein by reference). In one embodiment, the effector stem-loop is positioned 3′ of the aptamer such that the alternatively shared stem arm comprises all or a portion of the 3′ aptamer stem arm and all or a portion of the 5′ arm of the effector stem. In one embodiment, the effector stem-loop is positioned 5′ of the aptamer such that the alternatively shared stem arm comprises all or a portion of the 5′ aptamer stem arm and all or a portion of the 3′ arm of the effector stem. In one embodiment, the polyadenylation signal is AATAAA (SEQ ID NO:105) or ATTAAA (SEQ ID NO:106). In one embodiment, the polyadenylation signal is a downstream element (DSE). In one embodiment, the polyadenylation signal is an upstream sequence element (USE). In one embodiment, the polynucleotide cassette comprises two riboswitches, wherein the effector stem loop of the first riboswitch comprises all or part of the polyadenylation signal AATAAA (SEQ ID NO:105) or ATTAAA (SEQ ID NO:106) and the effector stem loop of the second riboswitch comprises all or part of the downstream element (DSE). In one embodiment, the two riboswitches each comprise aptamers that bind the same ligand. In one embodiment, the two riboswitches comprise different aptamers that bind different ligands.

[0167] In some embodiments, the riboswitch comprises a sensing region (*e.g.*, an aptamer) and an effector region comprising a binding site for the small nuclear ribonucleoprotein (snRNP) U1, which is part of the spliceosome. WO2017/136591 describes riboswitches wherein the effector region comprises a U1 snRNP binding site, and is incorporated herein by reference in its entirety. When the aptamer binds its ligand, the effector region forms a stem and sequesters the U1 snRNP binding site from binding a U1 snRNP. Under certain conditions (for example, when the aptamer is not bound to its ligand), the effector region is in a context that provides access to the U1 snRNP binding site, allowing U1 snRNP to bind the mRNA and inhibit polyadenylation leading to degradation of the message. The U1 snRNP binding site can be any polynucleotide sequence that is capable of binding the U1 snRNP, thereby recruiting the U1 snRNP to the 3′ UTR of a sequence encoding a polypeptide disclosed herein and suppressing polyadenylation of the mRNA of the

sequence encoding a polypeptide disclosed herein. In one embodiment, the U1 snRNP binding site is the consensus site CAGGTAAGTA (SEQ ID NO:107) (CAGGUAAGUA, SEQ ID NO:108, when in the mRNA). In some embodiments, the U1 snRNP binding site is a variation of this consensus sequence, including for example sequences that are shorter or have one or more nucleotides changed from the consensus sequence. In one embodiment, the U1 snRNP binding site contains the sequence CAGGTAAG (SEQ ID NO:109). In some embodiments, the binding site is encoded by the sequence selected from CAGGTAAGTA (SEQ ID NO:107), CAGGTAAGT (SEQ ID NO:110), and CAGGTAAG (SEQ ID NO:109). The U1 snRNP binding site can be any 5′ splice site from a gene, *e.g.*, the 5′ splice site from human DHFR exon 2.

[0168] *Aptamer-mediated modulation of ribonuclease cleavage*

[0169] In one embodiment, the expression of the sequence encoding a polypeptide disclosed herein is regulated through aptamer-modulated ribonuclease cleavage. Ribonucleases (RNases) recognize and cleave specific ribonuclease substrate sequences. Provided herein are recombinant DNA constructs that, when incorporated into the DNA of a sequence encoding a polypeptide disclosed herein, provide the ability to regulate expression of the sequence encoding a polypeptide disclosed herein by aptamer/ligand mediated ribonuclease cleavage of the resulting RNA. In some embodiments, the aptamer encoding sequence described herein is part of a construct that contains or encodes a ribonuclease substrate sequence and a riboswitch comprising an effector region and the aptamer such that when the aptamer binds a ligand, expression of the sequence encoding a polypeptide disclosed herein occurs (as described in WO2018/161053, which is incorporated in its entirety by reference herein). In embodiments, an RNase P substrate sequence is linked to a riboswitch wherein the riboswitch comprises an effector region and an aptamer, wherein the effector region comprises a sequence complimentary to a portion of the RNase P substrate sequence. Binding of a suitable ligand to the aptamer induces structural changes in the aptamer and effector region, altering the accessibility of the ribonuclease substrate sequence for cleavage by the ribonuclease.

[0170] In one embodiment, the aptamer sequence is located 5′ to the RNase P substrate sequence and the effector region comprises all or part of the leader sequence and all or part of the 5′ acceptor stem sequence of the RNase P substrate sequence. *See, e.g.*, Figs. 1a, 1b, and 3b of WO2018/161053 and the associated text, incorporated herein by reference. In further embodiments, the acceptor stem of the RNase P substrate and the riboswitch effector region are

separated by 0, 1, 2, 3, or 4 nucleotides. In other embodiments, the effector region stem includes, in addition to leader sequence (and its complement), one or more nucleotides of the acceptor stem of the RNase P substrate, and sequence complementary to the one or more nucleotides of the acceptor stem.

[0171] In one embodiment, the aptamer sequence of the polynucleotide cassette is located 3′ to the RNase P substrate sequence and the effector region comprises sequence complimentary to the all or part of the 3′ acceptor stem of the RNase P substrate sequence. *See, e.g.*, Fig. 3a of WO2018/161053 and the associated text, incorporated herein by reference. In further embodiments, the effector region sequence complimentary to the 3′ acceptor stem of the RNase P substrate is 1 to 7 nucleotides. In other words, the effector region stem includes 1 to 7 nucleotides of the acceptor stem and includes sequence that is complementary to this 1 to 7 nucleotides of the acceptor stem. In embodiments, the riboswitch is located 3′ of the RNase P substrate, so the effector region stem and the acceptor stem of the RNase P substrate do not overlap. In embodiments, the effector region and the acceptor stem of the RNase P substrate are immediately adjacent (*i.e.*, not overlapping). In other embodiments, the effector region and the acceptor stem of the RNase P substrate are separated by 1, 2, 3, 4, 5 or more nucleotides.

[0172] Vectors

[0173] In one aspect, provided are recombinant vectors and their use for the introduction of a expression construct disclosed herein. In some embodiments, the vectors disclosed herein include additional DNA elements including DNA segments that provide for the replication of the DNA in a host cell and/or expression of a sequence encoding a polypeptide disclosed herein in target cells at appropriate levels. The ordinarily skilled artisan appreciates that expression control sequences (promoters, enhancers, and the like) are selected based on their ability to promote expression of the sequence encoding a polypeptide disclosed herein in the target cell. “Vector” means a recombinant plasmid, yeast artificial chromosome (YAC), mini chromosome, DNA mini-circle or virus (including virus derived sequences) that comprises a polynucleotide to be delivered into a host cell, either *in vitro* or *in vivo*. In one embodiment, the recombinant vector is a viral vector or a combination of multiple viral vectors.

[0174] Viral vectors for the expression of a sequence encoding a polypeptide disclosed herein in a target cell, tissue, or organism are known in the art and include adenoviral (AV) vectors,

adeno-associated virus (AAV) vectors, retroviral and lentiviral vectors, and Herpes simplex type 1 (HSV1) vectors. Also included are viral particles comprising a nucleic acid encoding a polypeptide disclosed herein. In embodiments the viral particle as an AAV particle.

[0175] Adenoviral vectors include, for example, those based on human adenovirus type 2 and human adenovirus type 5 that have been made replication defective through deletions in the E1 and E3 regions. The transcriptional cassette can be inserted into the E1 region, yielding a recombinant E1/E3-deleted AV vector. Adenoviral vectors also include helper-dependent high-capacity adenoviral vectors (also known as high-capacity, "gutless" or "gutted" vectors), which do not contain viral coding sequences. These vectors contain the cis-acting elements needed for viral DNA replication and packaging, mainly the inverted terminal repeat sequences (ITR) and the packaging signal (CY). These helper-dependent AV vector genomes have the potential to carry from a few hundred base pairs up to approximately 36 kb of foreign DNA.

[0176] Recombinant adeno-associated virus "rAAV" vectors include any vector derived from any adeno-associated virus serotype, including, without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV-7 and AAV-8, AAV-9, AAV-10, AAVrh10, and AAV2-retro (disclosed in PCT Patent Publication WO2017218842A1, which is incorporated herein in its entirety) and the like. rAAV vectors can have one or more of the AAV wild-type genes deleted in whole or in part, preferably the Rep and/or Cap genes, but retain functional flanking ITR sequences. Functional ITR sequences are retained for the rescue, replication, packaging and potential chromosomal integration of the AAV genome. The ITRs need not be the wild-type nucleotide sequences, and may be altered (*e.g.*, by the insertion, deletion or substitution of nucleotides) so long as the sequences provide for functional rescue, replication and packaging.

[0177] Alternatively, other systems such as lentiviral vectors can be used. Lentiviral-based systems can transduce nondividing as well as dividing cells making them useful for applications targeting, for examples, the nondividing cells of the CNS. Lentiviral vectors are derived from the human immunodeficiency virus and, like that virus, integrate into the host genome providing the potential for very long-term gene expression.

[0178] Polynucleotides, including plasmids, YACs, minichromosomes and minicircles, carrying the sequence encoding a polypeptide disclosed herein containing the gene regulation cassette can also be introduced into a cell or organism by nonviral vector systems using, for example, cationic lipids, polymers, or both as carriers. Conjugated poly-L-lysine (PLL) polymer

and polyethylenimine (PEI) polymer systems can also be used to deliver the vector to cells. Other methods for delivering the vector to cells includes hydrodynamic injection and electroporation and use of ultrasound, both for cell culture and for organisms. For a review of viral and non-viral delivery systems for gene delivery see Nayerossadat, N. et al. (Adv Biomed Res. 2012; 1:27) incorporated herein by reference.

[0179] Pharmaceutical compositions

[0180] Provided herein are pharmaceutical composition comprising any of the expression constructs, vectors, or viral particles disclosed herein and a pharmaceutically acceptable excipient. These compositions may comprise, in addition to the expression construct, vector, or viral particle, a pharmaceutically and/or physiologically acceptable excipient, carrier, buffer, stabilizer, antioxidants, preservative, or other additives well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material may be determined by the skilled person according to the route of administration. The pharmaceutical composition is typically in liquid form. Liquid pharmaceutical compositions generally include a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Additional carriers are provided in International Patent Publication No. WO 00/15822, incorporated herein by reference. Physiological saline solution, magnesium chloride, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may be included. In some cases, a surfactant, such as pluronic acid (PF68) 0.001% may be used. In some cases, Ringer's Injection, Lactated Ringer's Injection, or Hartmann's solution is used. Preservatives, stabilizers, buffers, antioxidants and/or other additives may be included, as required.

[0181] For delayed release, the expression construct, vector, or viral particle may be included in a pharmaceutical composition which is formulated for slow release, such as in microcapsules formed from biocompatible polymers or in liposomal carrier systems according to methods known in the art.

[0182] If the expression construct, vector, or viral particle is to be stored long-term, it may be frozen in the presence of glycerol, or other cryopreservative.

[0183] Methods

[0184] In one aspect, provided is a method of inducing satiation in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0185] In one aspect, provided is a method of treating obesity in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0186] In one aspect, provided is a method of suppressing appetite in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0187] In one aspect, provided is a method of reducing weight gain in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0188] In one aspect, provided is a method of improving glucose tolerance in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0189] In one aspect, provided is a method of treating diabetes in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0190] In one aspect, provided is a method of inducing insulin release in a subject in need thereof, the method comprising administering to the subject an expression construct, a vector, or a pharmaceutical composition disclosed herein.

[0191] Provided herein is an expression construct, vector, or pharmaceutical composition disclosed herein for use in a method of inducing satiation in a subject in need thereof.

[0192] Provided herein is an expression construct, vector, or pharmaceutical composition disclosed herein for use in a method of treating obesity in a subject in need thereof.

[0193] Provided herein is an expression construct, vector, or pharmaceutical composition disclosed herein for use in a method of suppressing appetite in a subject in need thereof.

[0194] Provided herein is an expression construct, vector, or pharmaceutical composition disclosed herein for use in a method of reducing weight gain in a subject in need thereof.

[0195] Provided herein is an expression construct, vector, or pharmaceutical composition disclosed herein for use in a method of improving glucose tolerance in a subject in need thereof.

[0196] Provided herein is an expression construct, vector, or pharmaceutical composition disclosed herein for use in a method of treating diabetes in a subject in need thereof.

[0197] Provided herein is an expression construct, vector, or pharmaceutical composition disclosed herein for use in is a method of inducing insulin release in a subject in need thereof.

[0198] Provided herein is the use of an expression construct, vector, or pharmaceutical composition disclosed herein in the manufacture of a medicament for inducing satiation in a subject in need thereof.

[0199] Provided herein is the use of an expression construct, vector, or pharmaceutical composition disclosed herein in the manufacture of a medicament for treating obesity in a subject in need thereof.

[0200] Provided herein is the use of an expression construct, vector, or pharmaceutical composition disclosed herein in the manufacture of a medicament for suppressing appetite in a subject in need thereof.

[0201] Provided herein is the use of an expression construct, vector, or pharmaceutical composition disclosed herein in the manufacture of a medicament for reducing weight gain in a subject in need thereof.

[0202] Provided herein is the use of an expression construct, vector, or pharmaceutical composition disclosed herein in the manufacture of a medicament for improving glucose tolerance in a subject in need thereof.

[0203] Provided herein is the use of an expression construct, vector, or pharmaceutical composition disclosed herein in the manufacture of a medicament for treating diabetes in a subject in need thereof.

[0204] Provided herein is the use of an expression construct, vector, or pharmaceutical composition disclosed herein in the manufacture of a medicament for inducing insulin release in a subject in need thereof.

[0205] Provided herein is a method of treating a subject in in need of increased expression of a polypeptide (including a polyprotein) disclosed herein encoded by a sequence encoding the polypeptide, the method comprising administering to the patient a pharmaceutical composition comprising a ligand, which an aptamer binds to or otherwise responds to, wherein the patient

previously had been administered a recombinant DNA comprising the sequence encoding the polypeptide, and where the sequence encoding the polypeptide contains a gene regulation cassette disclosed herein that provides the ability to regulate expression of the target gene by the ligand of the aptamer.

[0206] The terms "treat," "treated," "treating," or "treatment" as used herein refer to therapeutic treatment, wherein the object is to slow down (lessen) an undesired physiological condition, disorder or disease, or to obtain beneficial or desired clinical results. For the purposes of this invention, beneficial or desired clinical results include, but are not limited to, alleviation of symptoms; diminishment of the extent of the condition, disorder or disease; stabilization (*i.e.*, not worsening) of the state of the condition, disorder or disease; delay in onset or slowing of the progression of the condition, disorder or disease; amelioration of the condition, disorder or disease state; and remission (whether partial or total), or enhancement or improvement of the condition, disorder or disease. Treatment includes eliciting a clinically significant response without excessive levels of side effects.

[0207] In one aspect, in the methods disclosed herein, the expression construct is delivered by gene therapy. The cell specificity of the sequence encoding a polypeptide disclosed herein may be controlled by a promoter and/or other elements within the vector and/or by the capsid of the viral vector. Delivery of the vector construct containing the sequence encoding a polypeptide disclosed herein, and the transfection of the target tissues resulting in stable transfection of the regulated sequence encoding a polypeptide disclosed herein, is the first step in producing the polypeptide.

[0208] In some embodiment, if an aptamer within the sequence encoding the polypeptide disclosed herein is used, the sequence encoding the polypeptide disclosed herein is not expressed at significant levels, *i.e.*, it is in the "off state" in the absence of the specific ligand that binds to the aptamer contained within in the regulatory cassette riboswitch. Only when the aptamer specific ligand is administered is the expression of the sequence encoding the polypeptide disclosed activated. The delivery of the vector construct containing the sequence encoding the polypeptide disclosed herein and the delivery of the activating ligand generally are separated in time. The delivery of the activating ligand will control when the sequence encoding the polypeptide disclosed herein is expressed, as well as the level of protein expression.

[0209] The expression construct, vector, or pharmaceutical composition disclosed herein (and the ligand in case of aptamer-mediated regulation of gene expression) may be delivered by a

number of routes including, but not limited to, intravitreal, intraocular, inhalation, subcutaneous, intramuscular, intradermal, intralesion, topical, intraperitoneal, intravenous (IV), intra-arterial, perivascular, intracerebral, intracerebroventricular, oral, sublingual, sublabial, buccal, nasal, intrathoracic, intracardiac, intrathecal, epidural, intraosseous, or intraarticular.

[0210] If an aptamer is used, the timing of delivery of the ligand can be adjusted as needed. For example, an oral small molecule ligand may be delivered daily, or multiple times a day. Alternatively, the inducing ligand may be dosed less frequently, for example, once a week, every other week, once a month.

[0211] Articles of manufacture and kits

[0212] Also provided are kits or articles of manufacture for use in the methods described herein. In aspects, the kits comprise the compositions described herein (*e.g.*, compositions for delivery of a vector comprising an expression construct disclosed herein) in suitable packaging. Suitable packaging for compositions (such as ocular compositions for injection) described herein are known in the art, and include, for example, vials (such as sealed vials), vessels, ampules, bottles, jars, flexible packaging (*e.g.*, sealed Mylar or plastic bags), and the like. These articles of manufacture may further be sterilized and/or sealed.

[0213] Also provided are kits comprising the compositions described herein. These kits may further comprise instruction(s) on methods of using the composition, such as uses described herein. The kits described herein may further include other materials desirable from a commercial and user standpoint, including buffers, diluents, filters, needles, syringes, and package inserts with instructions for performing the administration of the composition or performing any methods described herein. For example, in some embodiments, the kit comprises an rAAV for the expression of polypeptide disclosed herein, a pharmaceutically acceptable carrier suitable for injection, and one or more of: a buffer, a diluent, a filter, a needle, a syringe, and a package insert with instructions for performing the injections. In some embodiments, the kit is suitable for intraocular injection, intramuscular injection, intravenous injection and the like.

[0214] It is to be understood and expected that variations of the compositions of matter and methods herein disclosed can be made by one skilled in the art and it is intended that such modifications are to be included within the scope of the present disclosure. The following

Examples further illustrate the invention, but should not be construed to limit the scope of the invention in any way.

[0215] All references cited herein are hereby incorporated by reference in their entirety. All nucleotide sequences provided herein are in a 5′ to 3′ orientation unless stated otherwise. A Sequence Listing is filed herewith, the contents of which are incorporated herein by reference in its entirety.

[0216] Overview of sequences disclosed herein

[0217] Selected gut peptides (*see also* Tables 1 and 2)

Gut peptide	SEQ ID NO	SEQ ID NO
	Amino acid	Nucleic acid sequence
hGLP-1	1	6, 7, and 8
hGIP	2	9
hOXM	3	10
PYY	4	11
hGlucagon	5	12

[0218] Selected signal peptides (*see also* Tables 3 and 4)

Signal peptide	SEQ ID NO	SEQ ID NO
	Amino acid	Nucleic acid sequence
IgM	13	21
hInsul	14	22
mIgh	15	23
hGH	16	24
mEpo	17	25
mGHRH	18	26
hAlbumin	19	27
hFIX	20	28

[0219] Exemplary monocistronic constructs (*see also* Tables 5 and 6)

Construct name	Signal peptide	Encoded gut peptide	SEQ ID NO	SEQ ID NO
			Amino acid	Nucleic acid
GLP-1_C	IgM SP	hGLP-1	29	37
GLP-1_F	hInsul	hGLP-1	30	38
GLP-1_I	mIgh	hGLP-1	31	39
GLP-1_J	hGH	hGLP-1	32	40
GLP-1_K	mEpo	hGLP-1	33	41
GLP-1_L	mGHRH	hGLP-1	34	42

Construct name	Signal peptide	Encoded gut peptide	SEQ ID NO	SEQ ID NO
			Amino acid	Nucleic acid
GLP-1_M	hAlbumin	hGLP-1	35	43
GLP-1_N	hFIX9	hGLP-1	36	44
GIP_C	IgM SP	hGIP	111	
GIP_E	hFIX9	hGIP	112	
GIP_F	hAlbumin	hGIP	113	
GIP_G	hGH	hGIP	114	

[0220] Exemplary bicistronic constructs (*see also* Tables 7 and 8)

Construct name	Signal peptide	Gut peptides	SEQ ID NO	SEQ ID NO
			Amino acid	Nucleic acid
2xGLP-1_2xB	hAlbumin	hGLP-1,hGLP-1	45	50
GG_J	hGH	hGLP-1, hGIP	46	51
GG_F	hInsul	hGLP-1, hGIP	47	52
GG_L	mGHRH	hGLP-1, hGIP	48	53
GG_M	hAlbumin	hGLP-1, hGIP	49	54
2xGLP-1 w/o signal peptide	n/a	hGLP-1, hGLP-1	55	57
GG w/o signal peptide	n/a	hGLP-1, hGIP	56	58

[0221] Exemplary tricistronic constructs (*see also* Tables 9 and 10)

Construct name	Signal peptide	Gut peptides	SEQ ID NO	SEQ ID NO
			Amino acid	Nucleic acid
3xGLP-1_3xB	hAlbumin	hGLP-1, hGLP-1, hGLP-1,	59	67
3xGLP-1_3xC	hInsul	hGLP-1, hGLP-1, hGLP-1,	60	68
3xGLP-1_3xD	mGHRH	hGLP-1, hGLP-1, hGLP-1,	61	69
GOP_J	hGH	hGLP-1, hOXM, PYY	62	70
GOP_F	hInsul	hGLP-1, hOXM, PYY	63	71
GOP_L	mGHRH	hGLP-1, hOXM, PYY	64	72
GOP_M	hAlbumin	hGLP-1, hOXM, PYY	65	73
GGG_A	hGH	hGLP-1, hGlucagon, hGIP	66	74
3xGLP-1 w/o signal peptide	n/a	hGLP-1, hGLP-1, hGLP-1,	75	78
hGLP-1, hOXM, PYY w/o signal peptide	n/a	hGLP-1, hOXM, PYY	76	79
hGLP-1, hGlucagon, hGIP w/o signal peptide	n/a	hGLP-1, hGlucagon, hGIP	77	80

[0222] Other sequences

SEQ ID NO	Sequence
81	TACCCATACGATGTTCCAGATTACGCTAGAAAAAAGAGA
84	TTCCCAACCATTCCCTTATCCAGGCTTTTTGACAACGCTATGCTCCGCGCCAGAAAA AAGAGA
85	GCTCCCCCACGCCTCATCTGCGACAGTCGAGTTCTGGAGAGGTACAGAAAAAAGAGA
86	CTGCCCCCCTCACCTCCCTTCAGGATGCAGCGA
87	AGGGGTGTGTTTCGTCGA
88	ACAGTTTTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGG
89	YPYDVPDYARKKR
92	FPTIPLSRLFDNAMLRARKKR
93	APPRLICDSRVLERYRKKR
94	LPPSPPFRMQR
95	RGVFRR
96	TVFLDHENANKILNRPKR
97	RKKR
98	RMQR
99	VFRR
100	RPKR
101	GTGAGTCTATGGGACCCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCA TGTCATAGGAAGGGGAGAAGTAACAGGGTACACATATTGACCAAATCAGGGTAATTT TGCATTTGTAATTTTAAAAAATGCTTTCTTCTTTTAATATACTTTTTTGTTTATCTT ATTTCTAATACTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCA TGCCGAGTAACGCTGTTTCTCTAACTTGTAGGAATGAATTCAGATATTTCCAGAGAA TGAAAAAAAAATCTTCAGTAGAAGgtaatgt-**X**- acattacGCACCATTCTAAAGAATAACAGTGATAATTTCTGGGTTAAGGCAATAGCA ATATTTCTGCATATAAATATTTCTGCATATAAATTGTAACTGATGTAAGAGGTTTCA TATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTTATGGTTGGGA TAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCATACCT CTTATCTTCCTCCCACAG
102	X_1AGGTX_2AGT X_1 = A or C X_2 = A or G
103	GAATGAATTCAGATATTTCCAGAGAATGAAAAAAAAATCTTCAGTAGAAG
104	GAATGAATTCAGATATTTCCAGAGAATGAAAAAAAATCTTCAGTAGAAG
105	AATAAA
106	ATTAAA
107	CAGGTAAGTA
108	CAGGUAAGUA
109	CAGGTAAG
110	CAGGTAAGT

115	agaaagaagagaCATGCTGAAGGGACATTTACCTCAGATGTTTCTTCATACCTGGAA GGACAGGCTGCCAAGGAATTTATTGCATGGCTTGTGAAAGGCAGGGGCTGA

EXAMPLES

[0223] Example 1: Expression of secretable gut peptides using monocistronic expression constructs

[0224] *Experimental Procedures:*

[0225] For the construction of expression vector encoding secretable peptides, gene fragments containing DNA sequences encoding signal peptides, sequences facilitating furin cleavage sites, and human GLP-1_{7-37} or hGIP were synthesized (IDT) and cloned into expression constructs containing a CMV promoter.

[0226] For the transfection, 3.5 x 10^4 human embryonic kidney (HEK) 293 cells were plated in a 96-well flat bottom plate the day before transfection. Plasmid DNA (500 ng) was added to a tube or a 96-well U-bottom plate. Separately, TransIT-293 reagent (Mirus; 1.4 µL) was added to 50 µL Optimum I media (Life Technologies) and allowed to sit for 5 minutes at room temperature (RT). Then, 50 µL of this diluted transfection reagent was added to the DNA, mixed, and incubated at RT for 20 min. Finally, 7 µL of this solution was added to a well of cells in the 96-well plate. The supernatants of the transfected cell were collected 48 hours after transfection and assayed for GLP-1 or GIP peptide.

[0227] For the enzyme-linked immunosorbent assay (ELISA) for GLP-1, HEK 293 cells were transfected using TransIT-293 transfection reagent (Mirus Bio) with the constructs containing coding sequences for the active form of human GLP-1 peptide. 200 nM of Sitagliptin phosphate monohydrate (Sigma) was added to the culture for inhibiting dipeptidyl peptidase DPP-IV in the culture medium. The supernatants from the transfected cells were collected 48 hours after transfection and were subjected to SingleStep ELISA for the detection of human GLP-1_{7-36} in the supernatant following the manufacturer's instruction (Abcam).

[0228] *Results:*

[0229] To generate secretable GLP-1 peptides expression constructs, a sequence encoding a leader sequence (comprising of a signal peptide sequence and a furin recognition and cleavage sequence), was generated and fused to the 5' end of sequence encoding the GLP-1_{7-37} peptide. The signal peptide sequences from various secretory proteins were selected and tested for their ability

to promote secretion of GLP-1. The furin recognition and cleavage sequences comprised sequences with the minimal furin cleavage site (RXXR for consensus furin cleavage site) and a sequence N-terminal of the cleavage site that facilitates furin recognition and cleavage. Inclusion of furin recognition and cleavage sequence in the leader sequence promoted the functional N-terminus of the GLP-1 peptide to be fully processed and generated in non-endocrine cells.

[0230] The expression of the active GLP-1 peptide was assayed using ELISA specific for active GLP-1_{7-36}. As shown in **Fig. 1A**, of all the eight constructs (see **Tables 5** and **6**) generated for expressing GLP-1, only constructs GLP-1_F (with human insulin signal peptide), _I (with mouse Ig heavy chain signal peptide) and _L (with mouse GHRH signal peptide) produced detectable active GLP-1 peptides.

[0231] A second validation experiment confirmed the GLP1-$_{7-37}$ expression from construct GLP-1_F, I and L (**Fig. 1B**). Construct GLP-1_M (with human albumin signal peptide) expressed very low but detectable amount of GLP-1. These expression results suggest that signal peptides from different secretory proteins have different efficiency in promoting secretable GLP-1 expression. In addition, as the ELISA assay detected the furin-cleaved peptide. As such, a low and undetectable expression of GLP-1_{7-37} is likely due to the inefficient furin cleavage.

[0232] The same strategy was used to generate the constructs for the expression of secretable human glucose dependent insulinotropic peptide (GIP_{1-42}). The same leader sequences used in constructs GLP-1_C, N, M and J (see **Tables 5** and **6)** were used to construct GIP_C, E, F and G, respectively. Expression of the GIP peptides was assayed using an ELISA that detects total GIP (including both GIP_{1-42} and GIP_{3-42}). GIP constructs containing the leader sequences that lead to very low or undetectable amount of GLP-1 secretion expressed significant amount of GIP (**Fig. 1C**). However, the assay used here recognized the C-terminus of the GIP, therefore the efficiency of furin cleavage (at the N-terminus of the peptide) was not reflected in the assay.

[0233] Example 2: Expression of secretable gut peptides using bi- and tri-cistronic expression constructs

[0234] To improve the expression level of gut peptides, bi- and tricistronic expression constructs were generated encoding polyproteins comprising two or three GLP-1 peptides. The peptides were separated by a minimal furin cleavage site sequence (RXXR). Posttranslational furin processing of the polyprotein in non-endocrine cells led to the release of individual peptides.

[0235] First, a GLP-1 expression vector was constructed that built on the GLP-1_M construct (encoding for a polypeptide comprising a human albumin signal peptide sequence and a sequence containing a furin cleavage site downstream of the signal peptide, see SEQ ID NO:35). The GLP-1 encoding sequence can be any polynucleotide sequence that encodes the GLP-1_{7-37} peptide.

[0236] Since the GLP-1_M construct (see **Tables 5** and **6**) expresses very low amounts of GLP-1, it was tested whether two copies (bicistronic construct 2xGLP-1_2xB, see **Tables 7** and **8**) or three copies (tricistronic construct 3xGLP-1_3xB, see **Tables 9** and **10**) of the GLP-1 coding sequence would increase the peptide's expression level (**Fig. 2A**). As shown in **Fig. 2B**, two copies of the GLP-1 coding sequence indeed increased active GLP-1 expression, and three copies even substantially increased GLP-1 expression even further.

[0237] Next, tricistronic expression constructs 3xGLP-1_3xC and 3xGLP-1_3xD (see see **Tables 9** and **10)** were generated (containing the leader sequences as in construct GLP-1_F and GLP-1_L, respectively, see **Tables 5** and **6**). As shown in **Fig. 2C**, the tricistronic construct expressed more than 100 times the amount of GLP-1 as compared to monocistronic constructs GL-1_F and GLP-1_L. As only furin-cleaved GLP-1 can be detected in the assay, this enhanced high level of GLP-1_{7-37} expression is also an indication of efficient furin cleavage at the inserted furin sites that link each individual peptide and efficient posttranslational processing.

[0238] These results demonstrated that small peptide expression can be enhanced by using the expression constructs disclosed herein.

[0239] Example 3: Bi- and tri-cistronic expression constructs for the expression of different types of gut peptides

[0240] *Experimental Procedures:*

[0241] To construct the expression vectors encoding secretable multiple peptides, gene fragments containing DNA sequences encoding signal peptides, furin cleavage sequence and human GLP-1_{7-37}, human glucose dependent insulinotropic peptide (GIP_{1-42}) and human peptide tyrosine tyrosine (PYY_{3-36}), and human oxyntomodulin (OXM) were synthesized (IDT) and cloned into expression constructs containing CMV promoter.

[0242] For the ELISA to detect human glucose-dependent insulinotropic polypeptide (GIP), HEK 293 cells were transfected with TransIT-293 transfection reagent (Mirus Bio) with the constructs containing coding sequences for secretable human GLP-1_{7-37} and GIP_{1-42} peptides or

containing coding sequence for GLP-1_{7-37}, Oxyntomodulin and peptide tyrosine tyrosine (PYY_{3-36}). 200 nM of Sitagliptin phosphate monohydrate (Sigma) was added to the culture for inhibiting dipeptidyl peptidase DPP-IV in the culture medium. The supernatants from the transfected cells were collected 48 hours after transfection and were subjected to ELISA assay for active GLP-1_{7-36} (Abcam) and total GIP (EMD Millipore) and total PYY (EMD Millipore) following manufacturer's instruction.

[0243] *Results:*

[0244] Next, expression constructs were designed encoding polyproteins comprising different gut peptides, using the strategy of generating polyprotein sequences discussed in Example 2.

[0245] The leader sequences used in constructs GLP-1_J, F, L and M (see **Tables 5** and **6**) were used to construct GG_J, F, L, M, respectively, for co-expressing the GLP-1_{7-37} peptide and the GIP_{1-42} peptide (GG) (see **Tables 7** and **8**).

[0246] As shown in **Fig. 3A**, all four bicistronic GG expression constructs expressed higher levels of the GLP-1_{7-37} peptide as compared to monocistronic GLP-1 expression constructs. Constructs encoding for polypeptides with a human insulin signal sequence expressed the highest amount of GLP-1_{7-37} peptide (see GG_F in **Fig. 3A** and GLP-1_F in **Figs. 1A** and **1B**).

[0247] Similarly, the GIP expression from GG constructs is higher than that from constructs with single of GIP, and GG_F expressed the highest amount of GIP (**Fig. 3B**).

[0248] Next, a tricistronic expression construct was generated encoding a polyprotein comprising the GLP-1_{7-37} peptide, a glucagon peptide, and the GIP_{1-42} peptide. Tricistronic construct GGG_A (see **Tables 9** and **10**) has the same leader sequence as monocistronic constructs GLP-1_J and GIP_G (see **Tables 5** and **6**), but shows significantly higher expression of both the GLP-1_{7-37} peptide (**Fig. 3C**) and the GIP_{1-42} peptide (**Fig. 3D**) compared to the monocistronic constructs.

[0249] Next, the leader sequences used in constructs GLP-1_J, F, L and M (see **Tables 5** and **6**) were used to construct GOP_J, F, L and M, respectively, (see **Tables 9** and **10**) for co-expressing the GLP-1_{7-37} peptide, OXM peptide and PYY (GOP). As shown in **Fig. 3E**, all the four tricistronic GOP constructs expressed higher levels of the GLP-1 peptide as compared to the monocistronic constructs. Again, GOP_F expressed the highest amount of GLP-1_{7-37} peptide. Further, the GOP_F construct expressed approximately 27,158 pg/ml of PYY_{3-36}.

[0250] These results demonstrate that tricistronic expression constructs disclosed herein provide for significantly increased expression as compared to monocistronic expression constructs and constitute an efficient strategy for co-expressing multiple small peptides.

[0251] Example 4: Riboswitch-regulated expression of gut peptides

[0252] Next, it was investigated whether the controlled expression of gut peptides can be used in gene therapy for treating diabetes, obesity, and other metabolic indications. To that end, riboswitch N5-12G6 cassette (GTGAGTCTATGGGACCCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCAT GTCATAGGAAGGGGAGAAGTAACAGGGTACACATATTGACCAAATCAGGGTAATTT TGCATTTGTAATTTTAAAAAATGCTTTCTTCTTTTAATATACTTTTTTGTTTATCTTAT TTCTAATACTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATCATGC CGAGTAACGCTGTTTCTCTAACTTGTAGGAATGAATTCAGATATTTCCAGAGAATGA AAAAAAAATCTTCAGTAGAAGGTAATGTCTGGGGAGTCCTTCATGCGGGGCTGAGA GGATGGAAGCAATCGACCATCGACCCATTGCACCTGATCCGGTATGTCCCGGCGCA GGGAGACATTACGCACCATTCTAAAGAATAACAGTGATAATTTCTGGGTTAAGGCA ATAGCAATATTTCTGCATATAAATATTTCTGCATATAAATTGTAACTGATGTAAGAG GTTTCATATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTTATGGT TGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCA TACCTCTTATCTTCCTCCCACAG), was inserted into the encoding polypeptide sequence in bicistronic expression construct GG_L (expressing the hGLP-1_{7-37} peptide and the hGIP$_{1-42}$ peptide, see **Tables 7** and **8**) between nucleotide 81 and 82, or between 101 and 102, or between 137 and 138, or between 151 and 152, or between 176 and 177, or between 206 and 207, or between 286 and 287 (count starting with the start codon), respectively, generating regulatable, riboswitch-containing bicistronic expression constructs GG_L_1 through 7.

[0253] As shown in **Fig. 4A**, after transfection into HEK 293 cells, these constructs expressed the hGLP-1_{7-37} peptide in response to the small molecule inducer treatment in a dose dependent manner. The regulated expression of the other peptide, hGIP$_{1-42}$, expressed from the same construct, was also determined. Constructs GG_1 and GG_L_4 expressed the hGIP$_{1-42}$ peptide in a dose dependent manner (**Fig. 4B**).

[0254] Next, the riboswitch cassette was inserted in construct tricistronic expression construct 3xGLP-1_3xC (see **Tables 9** and **10**) at nucleotide position (counts from start codon in the polypeptide encoding sequence) between position 172 and 173, resulting in the regulatable GLP-1_3xC_4 construct. As shown in **Fig. 4C**, the GLP-1_3xC_4 construct expressed the GLP-1_{7-37} peptide in response to the small molecule inducer treatment in a dose dependent manner.

[0255] Next, the riboswitch cassette was inserted into the bicistronic expression construct GG_F (expressing the hGLP-1_{7-37} peptide and the hGIP$_{1-42}$ peptide, see **Tables 7** and **8**) at nucleotide position (counts from the start codon) between nucleotide 122 and 123, or between 158 and 159, or between 172 and 173, or between 197 and 198, or between 227 and 228, or between 307 and 308, respectively, generating regulatable bicistronic constructs GG_F_2 through 7. As shown in **Fig. 4D**, after transfection into HEK 293 cells, these constructs expressed the GLP-1_{7-37} peptide in response to the small molecule inducer treatment in a dose dependent manner. Particularly, regulatable, bicistronic expression construct GG_F_7 expressed the highest level of GLP-1_{7-37} peptide at each indicated concentration of small molecule inducer.

[0256] Next, a sequence encoding an additional furin cleavage site and an additional copy of the GLP-1_{7-37} peptide (agaaagaagagaCATGCTGAAGGGACATTTACCTCAGATGTTTCTTCATACCTGGAAGGA CAGGCTGCCAAGGAATTTATTGCATGGCTTGTGAAAGGCAGGGGCTGA, SEQ ID NO:115) was added 3′ of the nucleic acids encoding the last amino acid codon in the GG_F_7 construct, resulting in the regulatable, tricistronic GG_F_7-GLP-1 expression construct. This GG_F_GLP-1 construct (encoding a polyprotein comprising two copies of the hGLP-1_{7-37} peptide and one copy of the hGIP$_{1-42}$ peptide) expressed further enhanced level of the hGLP-1_{7-37} peptide in a dose responsive manner (**Fig. 4D**).

[0257] Finally, a regulatable tricistronic expression construct expressing a polyprotein comprising GLP-1, hOXM, and PYY expressed PYY in a dose dependent manner in response to the inducer (**Fig. 4E**).

[0258] Example 5: GLP-1 and GIP peptides are biological active *in vitro* and *in vivo*

[0259] *Experimental Procedure:*

[0260] For the transfection, 3.5 x 10^4 human embryonic kidney (HEK) 293 cells were plated in a 96-well flat bottom plate the day before transfection. Plasmid DNA (500 ng) was added to a tube

or a 96-well U-bottom plate. Separately, TransIT-293 reagent (Mirus; 1.4 μL) was added to 50 μL Optimum I media (Life Technologies) and allowed to sit for 5 minutes at room temperature (RT). Then, 50 μL of this diluted transfection reagent was added to the DNA, mixed, and incubated at RT for 20 min. Finally, 7 μL of this solution was added to a well of cells in the 96-well plate. The supernatants of the transfected cell were collected 48 hours after transfection and used as conditioned medium for the source of expressed GLP-1 or GIP peptide.

[0261] For the GLP-1 and GIP bioactivity assay, HEK 293 cells stably expressing the human GLP-1 receptor (HEK-293-hGLP-1R) or expressing the human GIP receptor (HEK-293-hGIPR) were generated by stably transfecting HEK293 cells with pCMV3 plasmid containing hGLP-1R cDNA or hGIPR cDNA (SinoBiological). The established stable cell lines were transfected with pCRE Tluc16-DD (Thermo Scientific) that contains TurboLuc luciferase gene driven under cAMP response element (CRE) promoter. 5 hours after transfection, the transfected cells were plated in 96-well plate at a 2 x 10^4 cells per well the day before the addition of the conditioned medium containing GLP-1 and/or GIP peptides. 1 hour after the addition of 100 μl conditioned medium, a luciferase assay was performed using TurboLuc Luciferase One-step glow assay kit (Thermo Scientific) following manufacturer's instruction, and luminescence was measure using Tecan microplate reader.

[0262] For AAV2/8 (AAV2 genome, AAV8 capsid) viral particle production, expression construct GG_F encoding hGLP-1 and hGIP was cloned into an AAV2 plasmid vector. Expression of the hGLP-1 and hGIP genes was driven by CASI promoter, which includes CMV and ubiquitin C enhancer elements and the chicken β-actin promoter. The AAV plasmid vector was packaged into an AAV8 capsid, generating AAV viral vector AAV8.GG_F.

[0263] For the animal studies, male C57Bl/6 mice (Jackson Laboratory) were fed with high fat diet (HFD) starting at 6 weeks of age. At week eight, the mice were injected with either PBS or 2.5 x 10^{11} genome copies (GC) of AAV8 vectors containing GG_F gene into hind limb both quadricep and gastrocnemius. Eight weeks old, male mice (Jackson Laboratory), which had been fed a low-fat diet (LFD), were injected with PBS as control group. Animal body weight was monitored before and after AAV injection weekly.

[0264] *Results:*

[0265] To further determine the biological activity of the GLP-1 and GIP peptides expressed by the polycistronic expression constructs herein, HEK 293 cells were used that stably express the

GLP-1 receptor or GIP receptor. The activity of GLP-1 and GIP peptide was assayed as described in the Experimental procedure.

[0266] As shown in **Fig. 5A**, hGLP-1_{7-37} peptide expressed from monocistronic expression construct GLP-1_F (expressing the hGLP-1_{7-37} peptide, see **Tables 5** and **6**), bicistronic expression construct GG_F (expressing the hGLP-1_{7-37} peptide and the hGIP$_{1-42}$ peptide, see **Tables 7** and **8**), and tricistronic expression construct GLP-1_3xC (expressing a polyprotein comprising three copies of the hGLP-1_{7-37} peptide, see **Tables 9** and **10**) showed activities in activating CRE promoter-driven luciferase. Supernatant from a GIP construct-transfected culture that does not express the GLP-1 peptide did not show activity in HEK 293 cells expressing GLP-1 receptor (negative control).

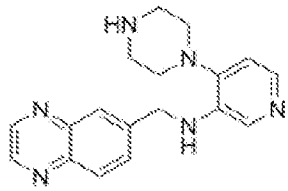
[0267] As shown in **Fig. 5B**, hGIP$_{1-42}$ peptide expressed from bicistronic expression construct GG_F (expressing the hGLP-1_{7-37} peptide and the hGIP$_{1-42}$ peptide, see **Tables 7** and **8**) and monocistronic expression construct GIP_F (expressing the hGIP$_{1-42}$ peptide, see **Tables 5** and **6**) activated CRE promoter driven luciferase gene. Supernatant from monocistronic expression construct GLP-1_F (expressing the hGLP-1_{7-37} peptide, see **Tables 5** and **6**) or tricistronic expression construct GLP-1_3xC (expressing a polyprotein comprising three copies of the hGLP-1_{7-37} peptide, see **Tables 9** and **10**) that does not express GIP did not activate CRE promoter driven luciferase gene (negative control). These results indicate that the gut peptides expressed from the polycistronic expression constructs disclosed herein are biologically active.

[0268] To further test the biological activities of the incretin peptides hGLP-1 and hGIP and their potential of inhibiting body weight gain from high fat diet (HFD), an AAV vector (comprising expression construct GG_F) was used to express the hGLP-1 and hGIP in mice fed with a high fat diet. Mice injected with AAV vector expressing hGLP-1 and hGIP gene had decreased body weight gain compared to mice without AAV.GG_F (**Fig. 5C**). This data indicates that hGLP-1 and hGIP expressed from expression constructs herein are biologically active and are useful in gene therapy for treating obesity.

[0269] Example 6: Riboswitch regulated expression of GLP-1 and GIP enhances glucose tolerance

[0270] *Experimental procedure:*

[0271] Peptide encoding sequences were cloned into AAV plasmid backbone that contains AAV2 ITRs, CSAI promoter and human beta globin polyA sequence and packaged into AAV8 capsid, generating AAV8.GG_F_7-GLP-1 vectors (see Example 4).

[0272] For the animal study (**Fig. 6A**), Balb/c mice received intramuscular injection of total 2.5 x 10^{11} viral genome (VG) per mouse of the receptive AAV8 viral particle into both quadriceps and both gastrocnemii. Compound 004 having the chemical structure  was formulated in 0.5% methylcellulose (MC): 0.25% Tween® 80 in deionized (DI) water for oral administration. 30 days after AAV vector delivery, mice were treated orally via oral gavage with 300 mg/kg compound 004 for 4 days.

[0273] Nor the non-fasting intraperitoneal glucose tolerance test (IPGTT), mice (N=5) injected with AAV vectors were treated with 300 mg/kg compound 004 for 4 days and was given 300 mg/kg 5 hours before IPGTT. Blood glucose in tail vein blood was measured with hand-held glucose meter (CVS Health). The glucose was measured before glucose injection (0 min), then 15 min, 30 min, 60 min and 120 min after glucose (2 g/kg) was injected peritoneally.

[0274] *Results:*

[0275] As shown in **Fig. 6B**, mice treated with the small molecule inducer showed better glucose tolerance than mice that received the dosing vehicle.

[0276] This shows that the GLP-1 and GIP peptides expressed by the injected vector increased the insulin release, thus improving glucose tolerance.

We claim:

1. A bicistronic expression construct encoding a polyprotein, wherein:
    a. the polyprotein comprises a signal peptide, a first gut peptide, and a second gut peptide; and
    b. the polyprotein encoding sequence comprises:
        i. a sequence encoding the signal peptide;
        ii. a sequence encoding the first gut peptide; and
        iii. a sequence encoding the second gut peptide.

2. A tricistronic expression construct encoding a polyprotein, wherein:
    a. the polyprotein comprises a signal peptide, a first gut peptide, a second gut peptide, and a third gut peptide; and
    b. the polyprotein encoding sequence comprises:
        i. a sequence encoding the signal peptide;
        ii. a sequence encoding the first gut peptide;
        iii. a sequence encoding the second gut peptide; and
        iv. a sequence encoding the third gut peptide.

3. The bicistronic expression construct of claim 1 or the tricistronic expression construct of claim 2, wherein the signal peptide is selected from the group consisting of an immunoglobulin M (IgM) signal peptide, human insulin (hInsul) signal peptide, murine Igh protein (Igh) signal peptide, human growth hormone (hGH) signal peptide, murine erythropoietin (mEpo) signal peptide, murine growth hormone-releasing hormone (mGHRH) signal peptide, human albumin signal peptide, and human factor IX (FIX) signal peptide.

4. The bicistronic expression or the tricistronic expression construct of claim 3, wherein the signal peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:13-20.

5. The bicistronic expression or the tricistronic expression construct of claim 4, wherein the signal peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:13-20.

6. The bicistronic expression or the tricistronic expression construct of claim 5, wherein the signal peptide comprises a sequence selected from SEQ ID NOS:13-20.

7. The bicistronic expression or the tricistronic expression construct of claims 3-6, wherein the sequence encoding the signal peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:21-28.

8. The bicistronic expression or the tricistronic expression construct of claim 7, wherein the sequence encoding the signal peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:21-28.

9. The bicistronic expression or the tricistronic expression construct of claim 8, wherein the sequence encoding the signal peptide comprises a sequence selected from SEQ ID NOS:21-28.

10. The bicistronic expression construct of any one of claims 1 or 3-9 or the tricistronic expression construct of any one of claims 2-9, wherein the first gut peptide and/or the second gut peptide comprises a sequence selected from the group consisting of human protein glp-1 (hGLP-1), human glucose dependent insulinotropic peptide (hGIP), human oxyntomodulin (hOXM), peptide YY (PYY), human glucagon, and amlyn.

11. The bicistronic expression construct or the tricistronic expression construct of claim 10, wherein the first gut peptide and/or the second gut peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:1-5.

12. The bicistronic expression construct or the tricistronic expression construct of claim 11, wherein the first gut peptide and/or the second gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:1-5.

13. The bicistronic expression construct or the tricistronic expression construct of claim 12, wherein the first gut peptide gut peptide and/or the second gut peptide comprises a sequence selected from SEQ ID NOS:1-5.

14. The bicistronic expression construct or the tricistronic expression construct of any one of claims 10-13, wherein the sequence encoding the first gut peptide gut peptide and/or the second gut peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:6-12.

15. The bicistronic expression construct or the tricistronic expression construct of claim 14, wherein the sequence encoding the first gut peptide gut peptide and/or the second gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:6-12.

16. The bicistronic expression construct or the tricistronic expression construct of claim 15, wherein the sequence encoding the first gut peptide gut peptide and/or the second gut peptide comprises a sequence that is selected from SEQ ID NOS:6-12.

17. The tricistronic expression construct of any one of claims 2-16, wherein the third gut peptide comprises a sequence selected from the group consisting of human protein glp-1 (hGLP-1), human glucose dependent insulinotropic peptide (hGIP), human oxyntomodulin (hOXM), peptide YY (PYY), human glucagon, and amlyn.

18. The tricistronic expression construct of claim 17, wherein the third gut peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:1-5.

19. The tricistronic expression construct of claim 18, wherein the third gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:1-5.

20. The tricistronic expression construct of claim 19, wherein the third gut peptide comprises a sequence selected from SEQ ID NOS:1-5.

21. The tricistronic expression construct of any one of claims 17-20, wherein the sequence encoding the third gut peptide comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:6-12.

22. The tricistronic expression construct of claim 21, wherein the sequence encoding the third gut peptide comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:6-12.

23. The tricistronic expression construct of claim 22, wherein the sequence encoding the third gut peptide comprises a sequence that selected from SEQ ID NOS:6-12.

24. The bicistronic expression construct of any one of claims 1 or 3-16, wherein the first gut peptide and the second gut peptide are the same gut peptide.

25. The bicistronic expression construct of claim 24, wherein the sequence encoding the first gut peptide and the sequence encoding the second gut peptide are different.

26. The bicistronic expression construct of claim 25, wherein at least one of the sequence encoding the first gut peptide and the sequence encoding the second gut peptide is codon-optimized.

27. The bicistronic expression construct of claim 26, wherein the sequence encoding the first gut peptide and the sequence encoding the second gut peptide are codon-optimized.

28. The bicistronic expression construct of any one of claims 24-27, wherein the first gut peptide and the second gut peptide are hGLP-1.

29. The bicistronic expression construct of claim 28, wherein the first gut peptide and the second gut peptide each comprise a sequence that is at least 80% identical to SEQ ID NO:1.

30. The bicistronic expression construct of claim 29, wherein the first gut peptide and the second gut peptide each comprise a sequence that is at least 90% identical to SEQ ID NO:1.

31. The bicistronic expression construct of claim 30, herein the first gut peptide and the second gut peptide each comprise SEQ ID NO:1.

32. The bicistronic expression construct of claim 28, wherein the sequences encoding the first and the second gut peptide each comprise a sequence that is at least 80% identical to a sequence selected from SEQ ID NOS:6-8.

33. The bicistronic expression construct of claim 32, wherein the sequences encoding the first and the second gut peptide each comprise a sequence that is at least 90% identical to a sequence selected from SEQ ID NOS:6-8.

34. The bicistronic expression construct of claim 33, wherein the sequences encoding the first and the second gut peptide are selected from SEQ ID NOS:6-8.

35. The bicistronic expression construct of claim 28, wherein the bicistronic expression construct comprises a sequence encoding a polypeptide that is at least 80% identical to SEQ ID NO:45 or SEQ ID NO:55.

36. The bicistronic expression construct of claim 35, wherein the bicistronic expression construct comprises a sequence encoding a polypeptide that is at least 90% identical to SEQ ID NO:45 or SEQ ID NO:55.

37. The bicistronic expression construct of claim 36, wherein the bicistronic expression construct encodes a sequence comprising SEQ ID NO:45 or SEQ ID NO:55.

38. The bicistronic expression construct of claim 28, wherein the bicistronic expression construct comprises a sequence that is at least 80% identical to SEQ ID NO:50 or SEQ ID NO:57.

39. The bicistronic expression construct of claim 38, wherein the bicistronic expression construct comprises a sequence that is at least 90% identical to SEQ ID NO:50 or SEQ ID NO:57.

40. The bicistronic expression construct of claim 39, wherein the bicistronic expression construct comprises SEQ ID NO:50 or SEQ ID NO:57.

41. The bicistronic expression construct of any one of claims 1 or 3-16, wherein the first gut peptide and the second gut peptide are different gut peptides.

42. The bicistronic expression construct of claim 41, wherein the first gut peptide and the second gut peptide are selected from the group consisting of hGLP-1 and hGIP.

43. The bicistronic expression construct of claim 42, wherein the bicistronic expression construct encodes a sequence comprising a sequence that is at least 80% identical to any one of SEQ ID NOS:46-49 or SEQ ID NO:56.

44. The bicistronic expression construct of claim 43, wherein the bicistronic expression construct encodes a sequence comprising a sequence that is at least 90% identical to any one of SEQ ID NOS: 46-49 or SEQ ID NO:56.

45. The bicistronic expression construct of claim 44, wherein the bicistronic expression construct encodes a sequence comprising any one of SEQ ID NOS:46-49 or SEQ ID NO:56.

46. The bicistronic expression construct of claim 42, wherein the bicistronic expression construct comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:51-54 or SEQ ID NO:58.

47. The bicistronic expression construct of claim 46, wherein the bicistronic expression construct comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:51-54 or SEQ ID NO:58.

48. The bicistronic expression construct of claim 47, wherein the bicistronic expression construct comprises a sequence selected from SEQ ID NOS:51-54 or SEQ ID NO:58.

49. The tricistronic expression construct of any one of claims 2-23, wherein the first gut peptide, and the second gut peptide are different gut peptides.

50. The tricistronic expression construct of claim 49, wherein the first gut peptide, the second gut peptide, and the third gut peptide are different gut peptides.

51. The tricistronic expression construct of claim 50, wherein the first gut peptide, the second gut peptide, and the third gut peptide are selected from the group consisting of:
    a. hGLP-1, hOXM, and PYY; or
    b. hGLP-1, hGlucagon, and hGIP.

52. The tricistronic expression construct of claim 51, wherein the tricistronic expression construct encodes a sequence comprising a sequence that is at least 80% identical to any one of SEQ ID NOS:62-66 or SEQ ID NOS:76-77.

53. The tricistronic expression construct of claim 52, wherein the tricistronic expression construct encodes a sequence comprising a sequence that is at least 90% identical to any one of SEQ ID NOS:62-66 or SEQ ID NOS:76-77.

54. The tricistronic expression construct of claim 53, wherein the tricistronic expression construct encodes a sequence comprising any one of SEQ ID NOS:62-66 or SEQ ID NOS:76-77.

55. The tricistronic expression construct of claim 51, wherein the tricistronic expression construct comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:70-74 or SEQ ID NOS:79-80.

56. The tricistronic expression construct of claim 55, wherein the tricistronic expression construct comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:70-74 or SEQ ID NOS:79-80.

57. The tricistronic expression construct of claim 56, wherein the tricistronic expression construct comprises any one of SEQ ID NOS:70-74 or SEQ ID NOS:79-80.

58. The tricistronic expression construct of any one of claims 2-23, wherein the first gut peptide, the second gut peptide, and the third gut peptide are the same gut peptide.

59. The tricistronic expression construct of claim 58, wherein the sequence encoding the first gut peptide, the sequence encoding the second gut peptide, and the sequence encoding the third gut peptide are different.

60. The tricistronic expression construct of claim 59, wherein at least one of the sequence encoding the first gut peptide, the sequence encoding the second gut peptide, and the sequence encoding the third gut peptide is codon-optimized.

61. The tricistronic expression construct of claim 60, wherein the sequence encoding the sequence encoding the first gut peptide, the sequence encoding the second gut peptide, and the sequence encoding the third gut peptide are codon-optimized.

62. The tricistronic expression construct of any one of claims 58-61, wherein the first gut peptide, the second gut peptide and the third gut peptide are hGLP-1.

63. The tricistronic expression construct of claim 62, wherein tricistronic expression construct encodes a sequence comprising a sequence that is at least 80% identical to any one of SEQ ID NOS:59-61 or SEQ ID NO:75.

64. The tricistronic expression construct of claim 63, wherein the tricistronic expression construct encodes a sequence comprising a sequence that is at least 90% identical to any one of SEQ ID NOS:59-61 or SEQ ID NO:75.

65. The tricistronic expression construct of claim 64, wherein the tricistronic expression construct encodes a sequence comprising a sequence selected from SEQ ID NOS:59-61 or SEQ ID NO:75.

66. The tricistronic expression construct of claim 62, wherein the tricistronic expression construct comprises a sequence that is at least 80% identical to any one of SEQ ID NOS:67-69 or SEQ ID NO:78.

67. The tricistronic expression construct of claim 66, wherein the tricistronic expression construct comprises a sequence that is at least 90% identical to any one of SEQ ID NOS:67-69 or SEQ ID NO:78.

68. The tricistronic expression construct of claim 67, wherein the tricistronic expression construct comprises a sequence selected from SEQ ID NOS:67-69 or SEQ ID NO:78.

69. The bicistronic expression construct of any one of claims 1, 3-16, or 24-48 or the tricistronic expression construct of any one of claims 2-23 or 49-68, wherein the bicistronic expression construct or the tricistronic expression construct further comprises a promoter sequence.

70. The bicistronic expression construct or the tricistronic expression construct of claim 69, wherein the promoter is CMV promoter or a CASI promoter.

71. The bicistronic expression construct of any one of claims 1, 3-16, 24-48, or 69-70, wherein the polyprotein further comprises a protease cleavage site allowing for the release of the first gut peptide and the second gut peptide.

72. The tricistronic expression construct of any one of claims 2-23 or 49-70, wherein the polyprotein further comprises two protease cleavage sites allowing for the release of the first gut peptide, the second gut peptide, and the third gut peptide.

73. The bicistronic expression construct of claim 71 or the tricistronic expression construct of claim 72, wherein at least one of the protease cleavage sites is a furin cleavage site.

74. The bicistronic expression construct of any one of claims 1, 3-16, 24-48, 69-71, or 73 or the tricistronic expression construct of any one of claims 2-23, 49-70, or 72-73, wherein the polyprotein encoding sequence comprises a riboswitch comprising an aptamer, wherein the aptamer binds to a small molecule.

75. The bicistronic expression construct of any one of claims 1, 3-16, 24-48, 69-71, or 73 or the tricistronic expression construct of any one of claims 2-23, 49-70, or 72-73, wherein the polyprotein encoding sequence comprises a gene regulation cassette comprising an aptamer, wherein the aptamer binds to a small molecule.

76. A vector comprising the bicistronic expression construct of any one of claims 1, 3-16, 24-48, 69-71, or 73-75 or the tricistronic expression construct of any one of claims 2-23, 49-70, or 72-75.

77. The vector of claim 76, wherein the vector is an AAV vector.

78. A pharmaceutical composition comprising the vector of any one of claims 76-77 and a pharmaceutically acceptable excipient.

79. A method of inducing satiation in a subject in need thereof, the method comprising administering to the subject the pharmaceutical composition of claim 78.

80. A method of treating obesity in a subject in need thereof, the method comprising administering to the subject the pharmaceutical composition of claim 78.

81. A method of suppressing appetite in a subject in need thereof, the method comprising administering to the subject the pharmaceutical composition of claim 78.

82. A method of reducing weight gain in a subject in need thereof, the method comprising administering to the subject the pharmaceutical composition of claim 78.

83. A method of improving glucose tolerance in a subject in need thereof, the method comprising administering to the subject the pharmaceutical composition of claim 78.

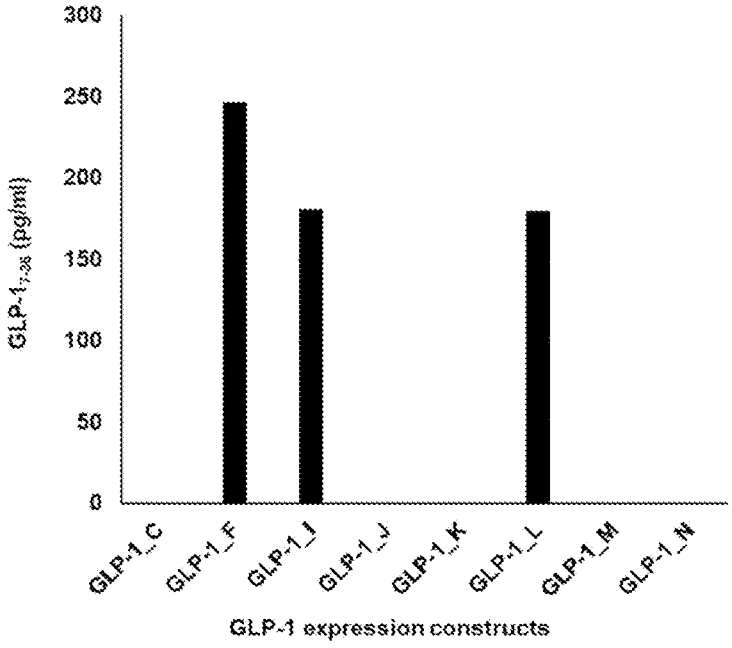

300
250
200
150
100
50
0
GLP-1$_{7-36}$ (pg/ml)
GLP-1_C
GLP-1_F
GLP-1_I
GLP-1_J
GLP-1_K
GLP-1_L
GLP-1_M
GLP-1_N
GLP-1 expression constructs

Fig. 1A

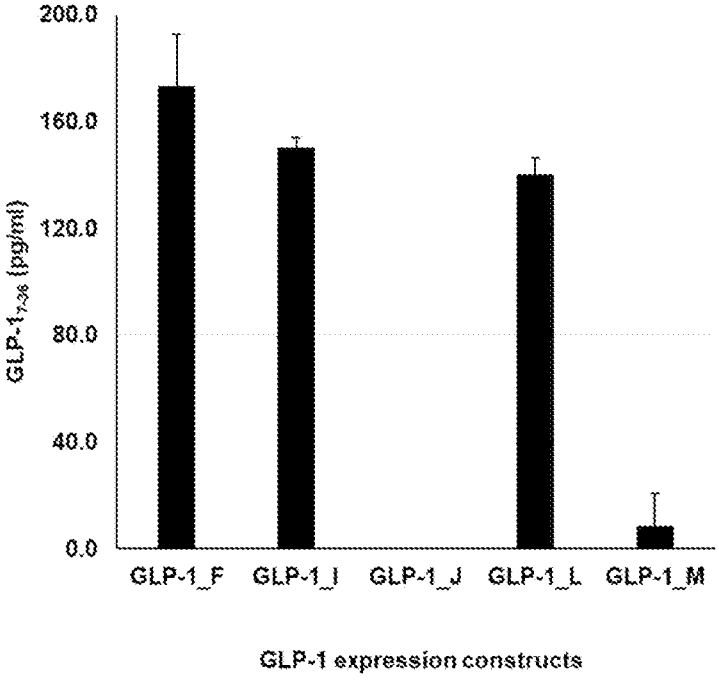

200.0
160.0
120.0
80.0
40.0
0.0
GLP-1$_{7-36}$ (pg/ml)
GLP-1_F
GLP-1_I
GLP-1_J
GLP-1_L
GLP-1_M
GLP-1 expression constructs

Fig. 1B

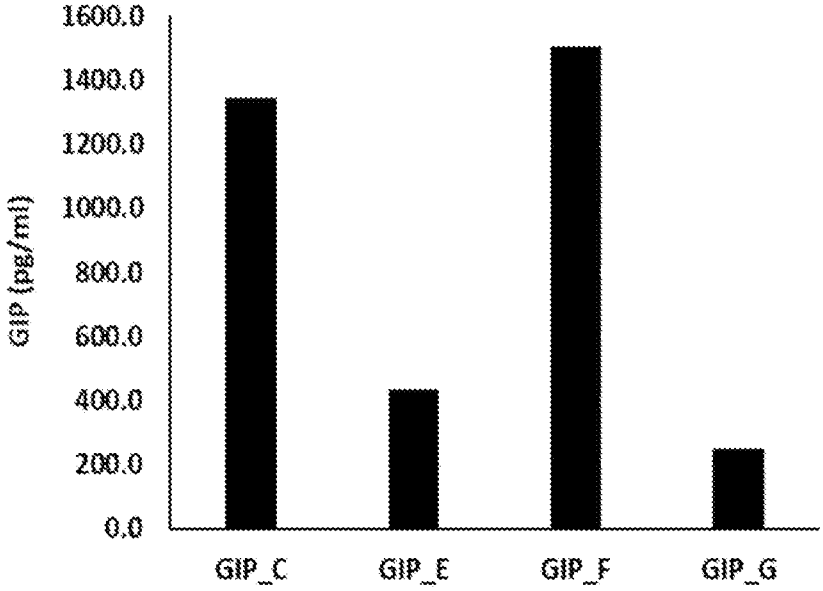

1600.0
1400.0
1200.0
1000.0
800.0
600.0
400.0
200.0
0.0
GIP (pg/ml)
GIP_C
GIP_E
GIP_F
GIP_G

Fig. 1C

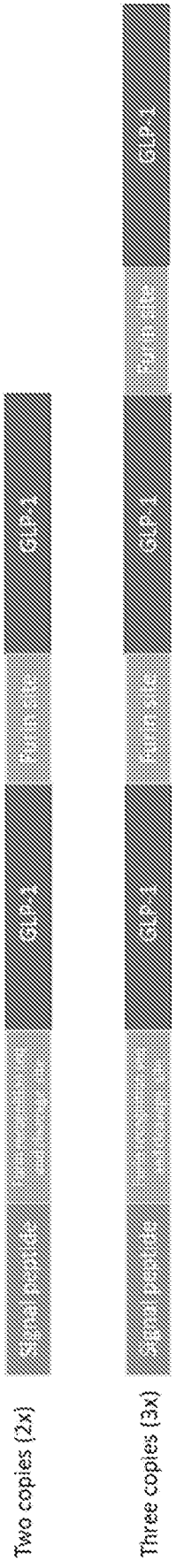

Two copies (2x)
Three copies (3x)
Signal peptide
GLP-1
Furin site
GLP-1
Signal peptide
GLP-1
Furin site
GLP-1
Furin site
GLP-1

Fig. 2A

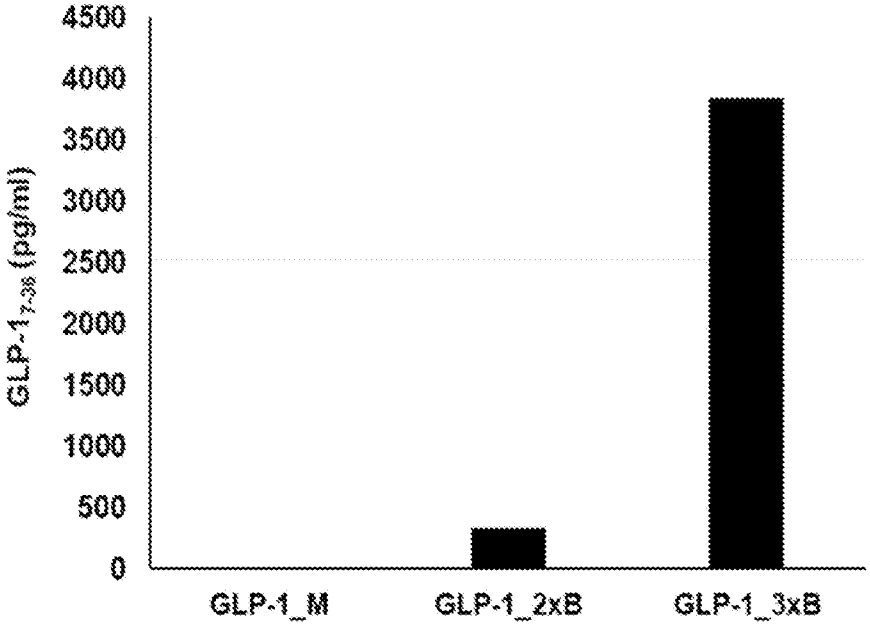

4500
4000
3500
3000
2500
2000
1500
1000
500
0
GLP-1$_{7-36}$ (pg/ml)
GLP-1_M
GLP-1_2xB
GLP-1_3xB

Fig. 2B

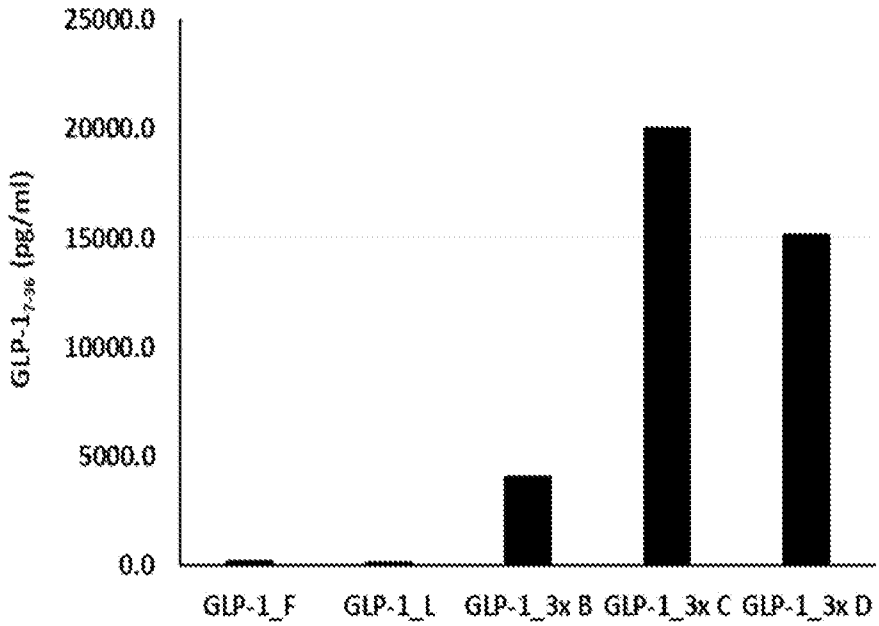

25000.0
20000.0
15000.0
10000.0
5000.0
0.0
GLP-1$_{7-36}$ (pg/ml)
GLP-1_F
GLP-1_L
GLP-1_3x B
GLP-1_3x C
GLP-1_3x D

Fig. 2C

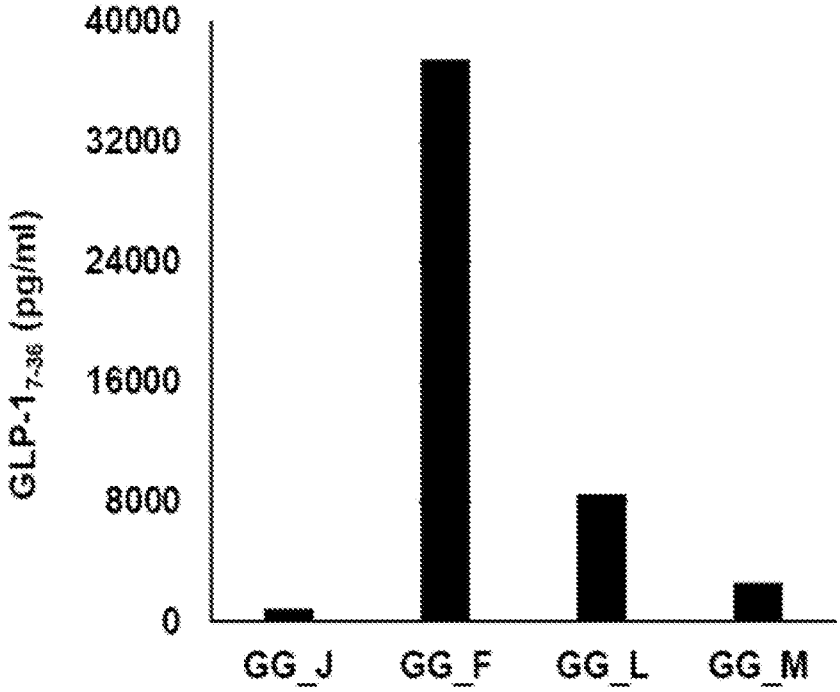

40000
32000
24000
16000
8000
0
GLP-1$_{7-36}$ (pg/ml)
GG_J
GG_F
GG_L
GG_M

Fig. 3A

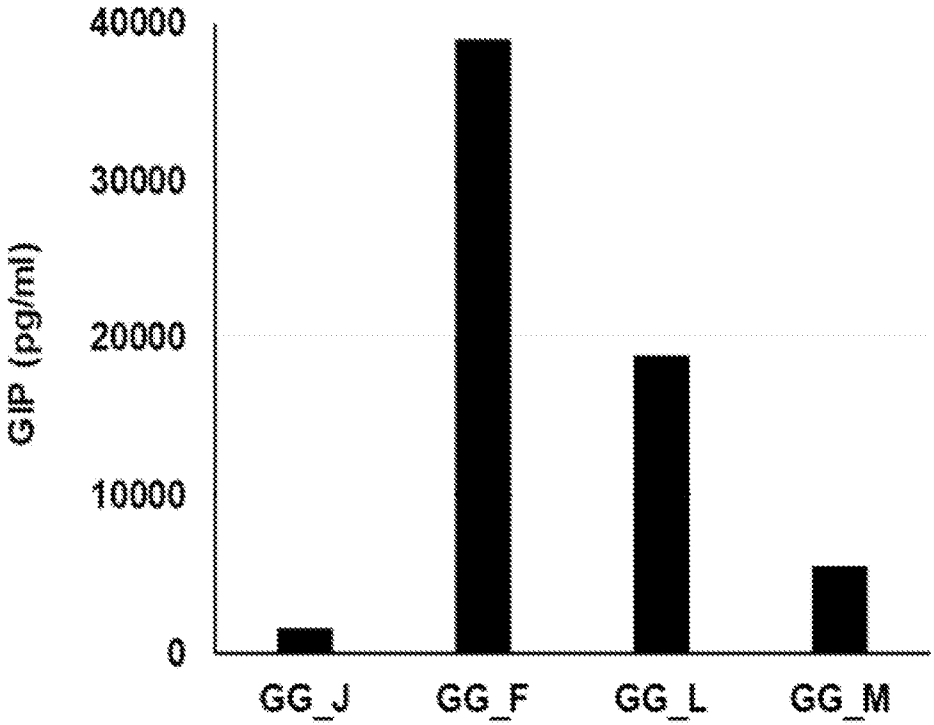

40000
30000
20000
10000
0
GIP (pg/ml)
GG_J
GG_F
GG_L
GG_M

Fig. 3B

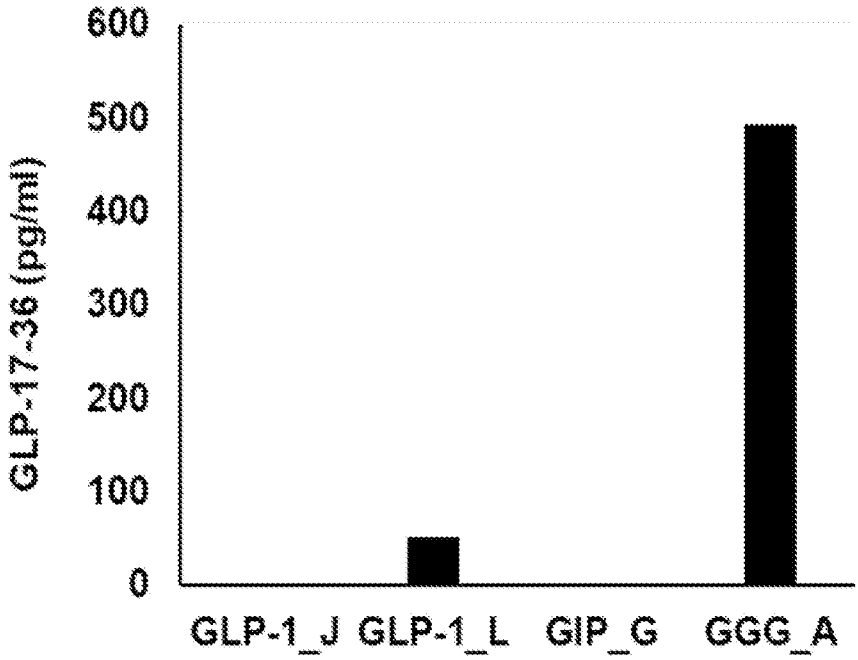

600
500
400
300
200
100
0
GLP-17-36 (pg/ml)
GLP-1_J
GLP-1_L
GIP_G
GGG_A

Fig. 3C

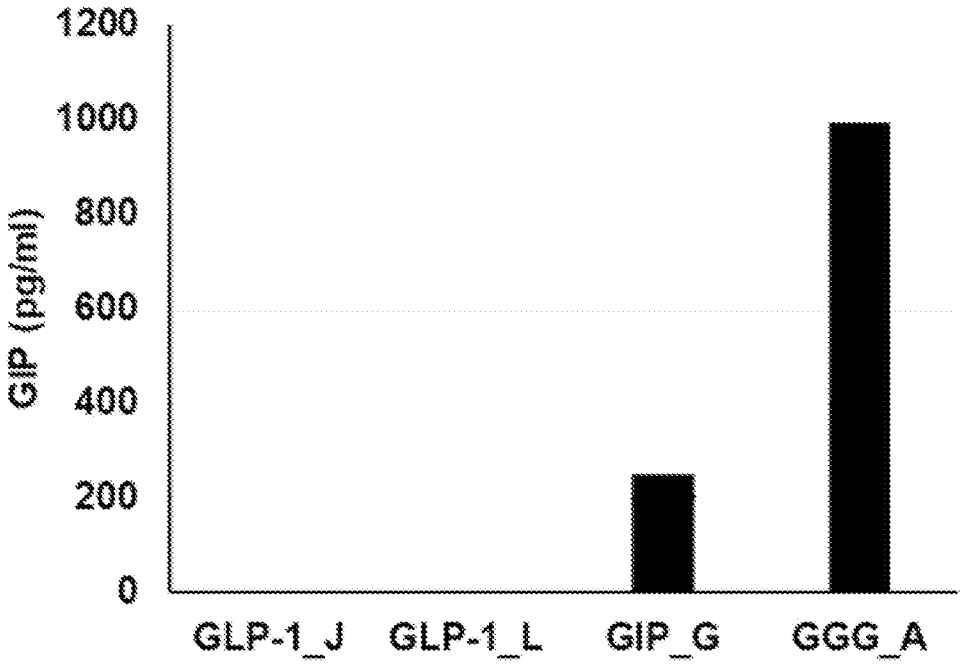

1200
1000
800
600
400
200
0
GIP (pg/ml)
GLP-1_J
GLP-1_L
GIP_G
GGG_A

Fig. 3D

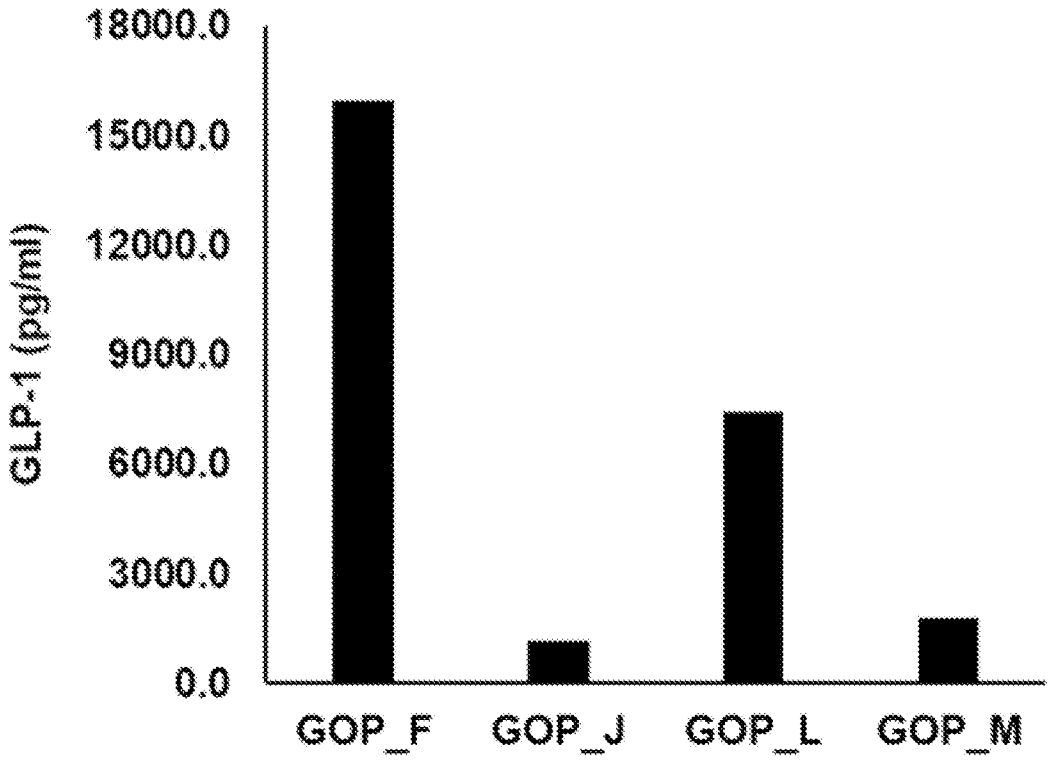

18000.0
15000.0
12000.0
9000.0
6000.0
3000.0
0.0
GLP-1 (pg/ml)
GOP_F
GOP_J
GOP_L
GOP_M

Fig. 3E

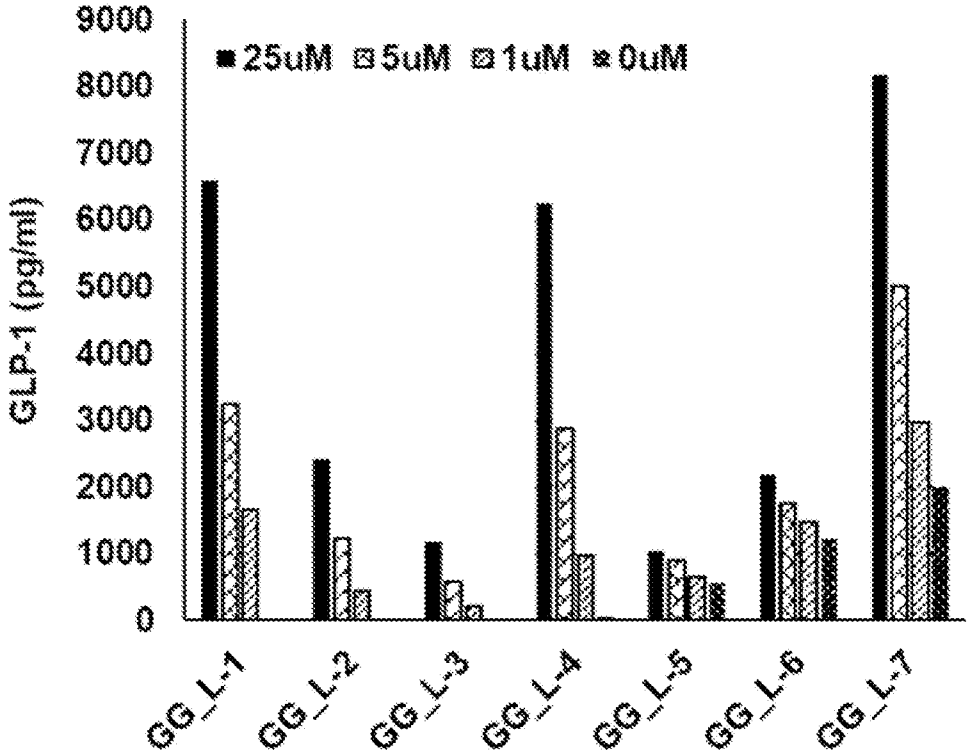

9000
8000
7000
6000
5000
4000
3000
2000
1000
0
GLP-1 (pg/ml)
25uM
5uM
1uM
0uM
GG_L-1
GG_L-2
GG_L-3
GG_L-4
GG_L-5
GG_L-6
GG_L-7

Fig. 4A

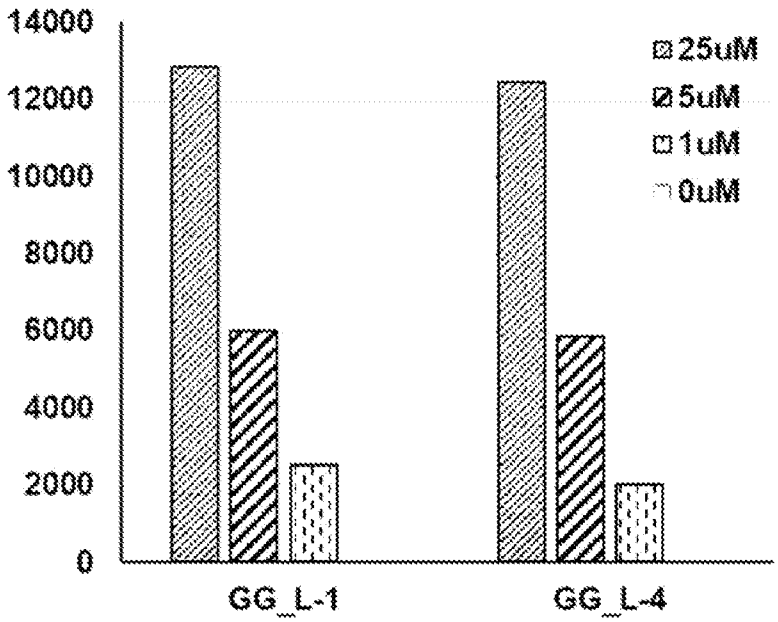

14000
12000
10000
8000
6000
4000
2000
0
25uM
5uM
1uM
0uM
GG_L-1
GG_L-4

Fig. 4B

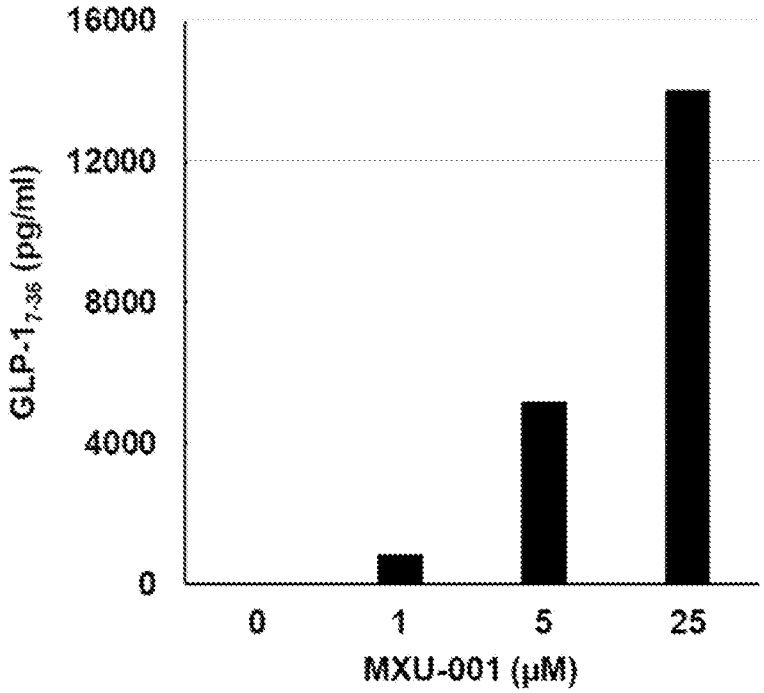

16000
12000
8000
4000
0
GLP-1$_{7-36}$ (pg/ml)
0
1
5
25
MXU-001 (μM)

Fig. 4C

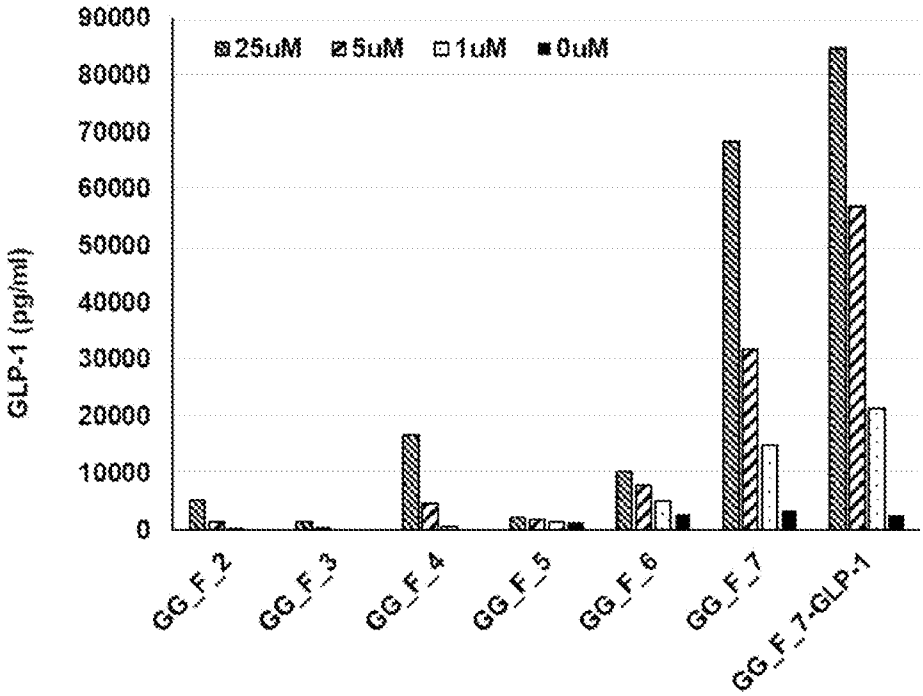
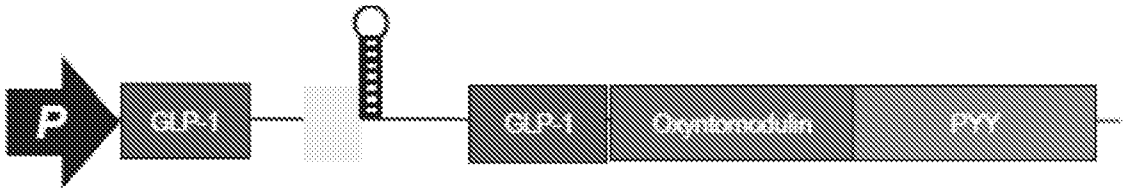

90000
80000
70000
60000
50000
40000
30000
20000
10000
0
GLP-1 (pg/ml)
25uM
5uM
1uM
0uM
GG_F_2
GG_F_3
GG_F_4
GG_F_5
GG_F_6
GG_F_7
GG_F_7-GLP-1

Fig. 4D


P
GLP-1
GLP-1
Oxyntomodulin
PYY

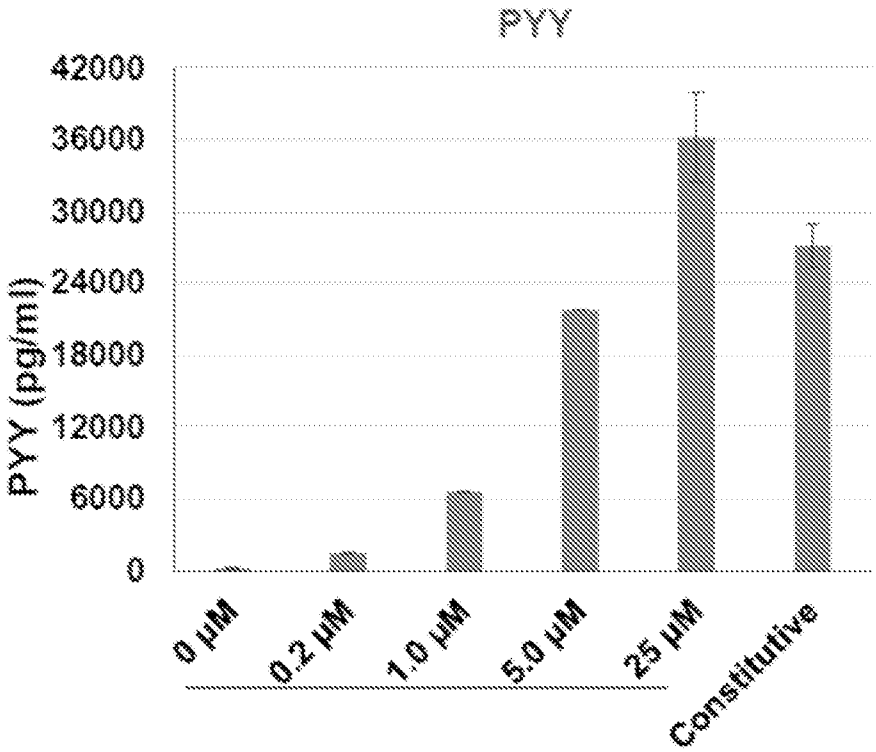

PYY
PYY (pg/ml)
42000
36000
30000
24000
18000
12000
6000
0
0 µM
0.2 µM
1.0 µM
5.0 µM
25 µM
Constitutive

Fig. 4E

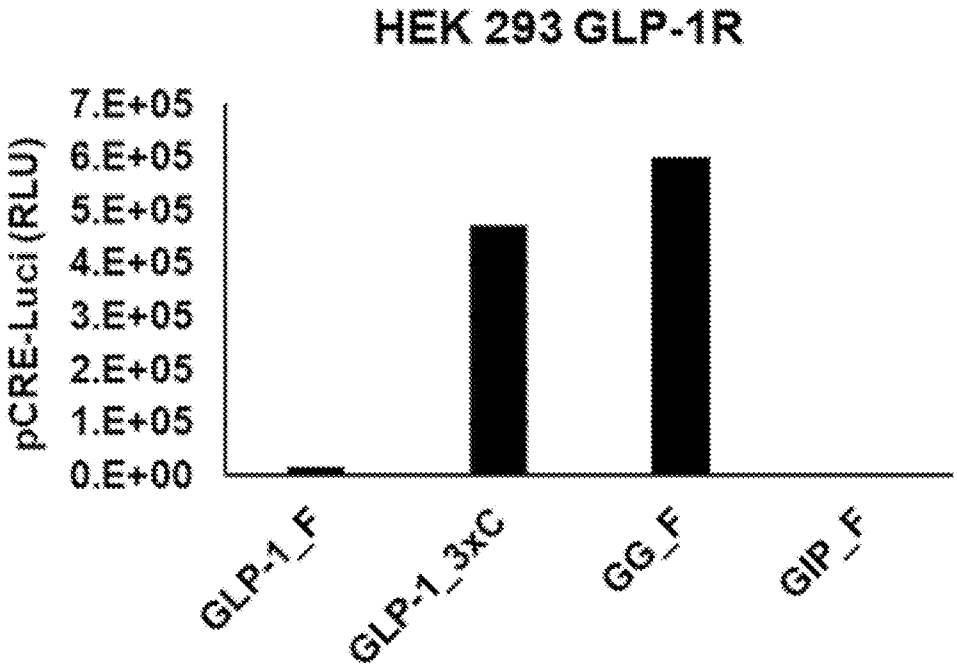

HEK 293 GLP-1R
pCRE-Luci (RLU)
7.E+05
6.E+05
5.E+05
4.E+05
3.E+05
2.E+05
1.E+05
0.E+00
GLP-1_F
GLP-1_3xC
GG_F
GIP_F

Fig. 5A

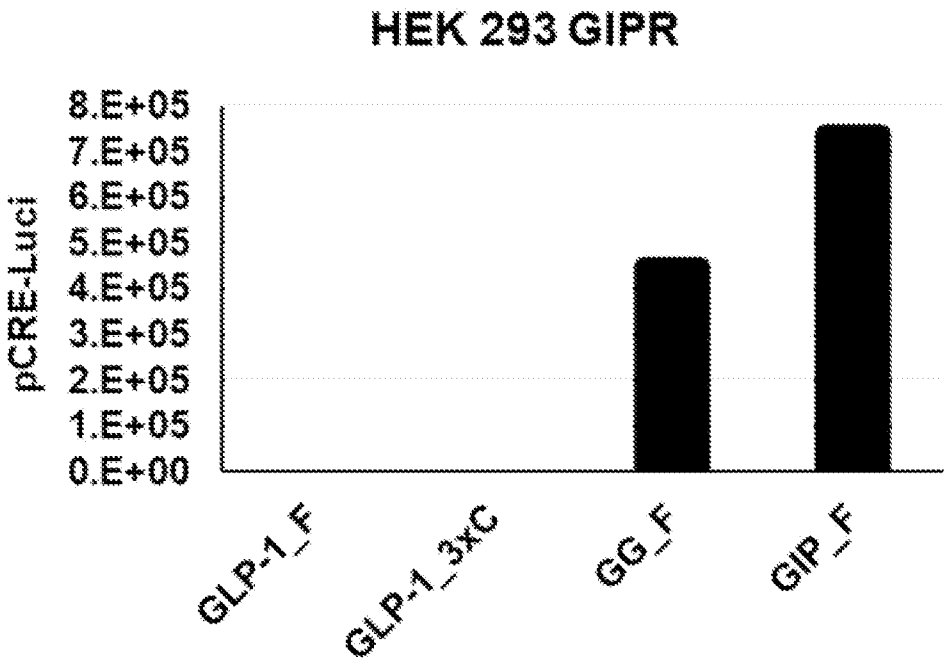

HEK 293 GIPR
pCRE-Luci
8.E+05
7.E+05
6.E+05
5.E+05
4.E+05
3.E+05
2.E+05
1.E+05
0.E+00
GLP-1_F
GLP-1_3xC
GG_F
GIP_F

Fig. 5B

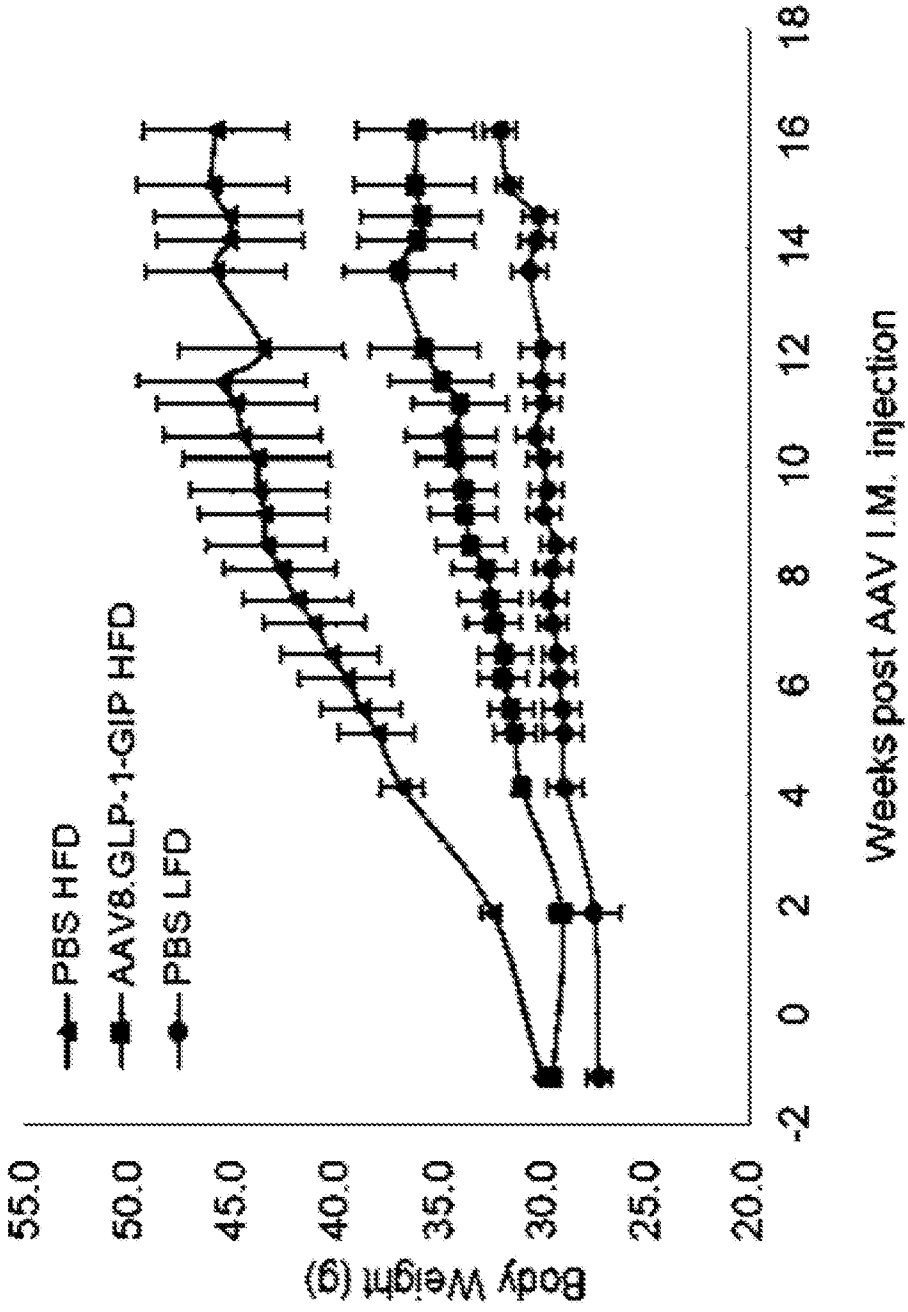

PBS HFD
AAV8.GLP-1-GIP HFD
PBS LFD
Body Weight (g)
55.0
50.0
45.0
40.0
35.0
30.0
25.0
20.0
-2
0
2
4
6
8
10
12
14
16
18
Weeks post AAV I.M. injection

Fig. 5C

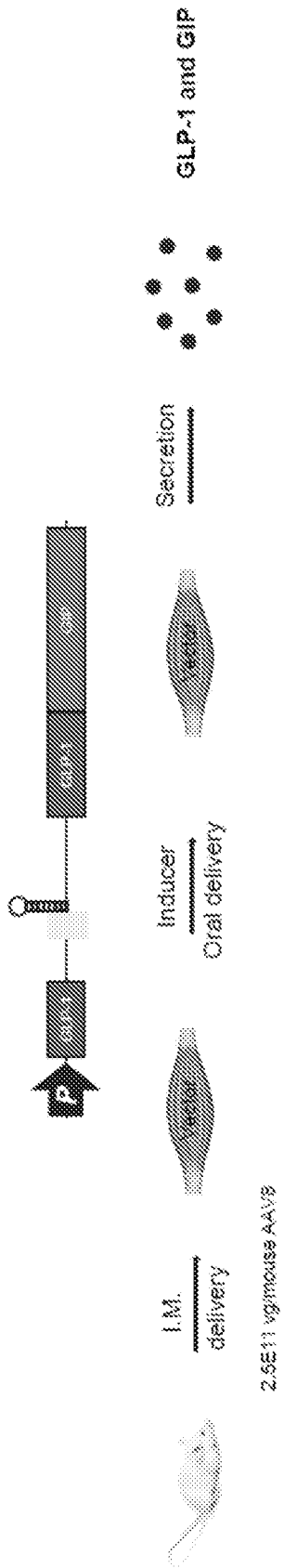

P
I.M.
delivery
2.5E11 vg/mouse AAV8
Inducer
Oral delivery
Secretion
GLP-1 and GIP

Fig. 6A

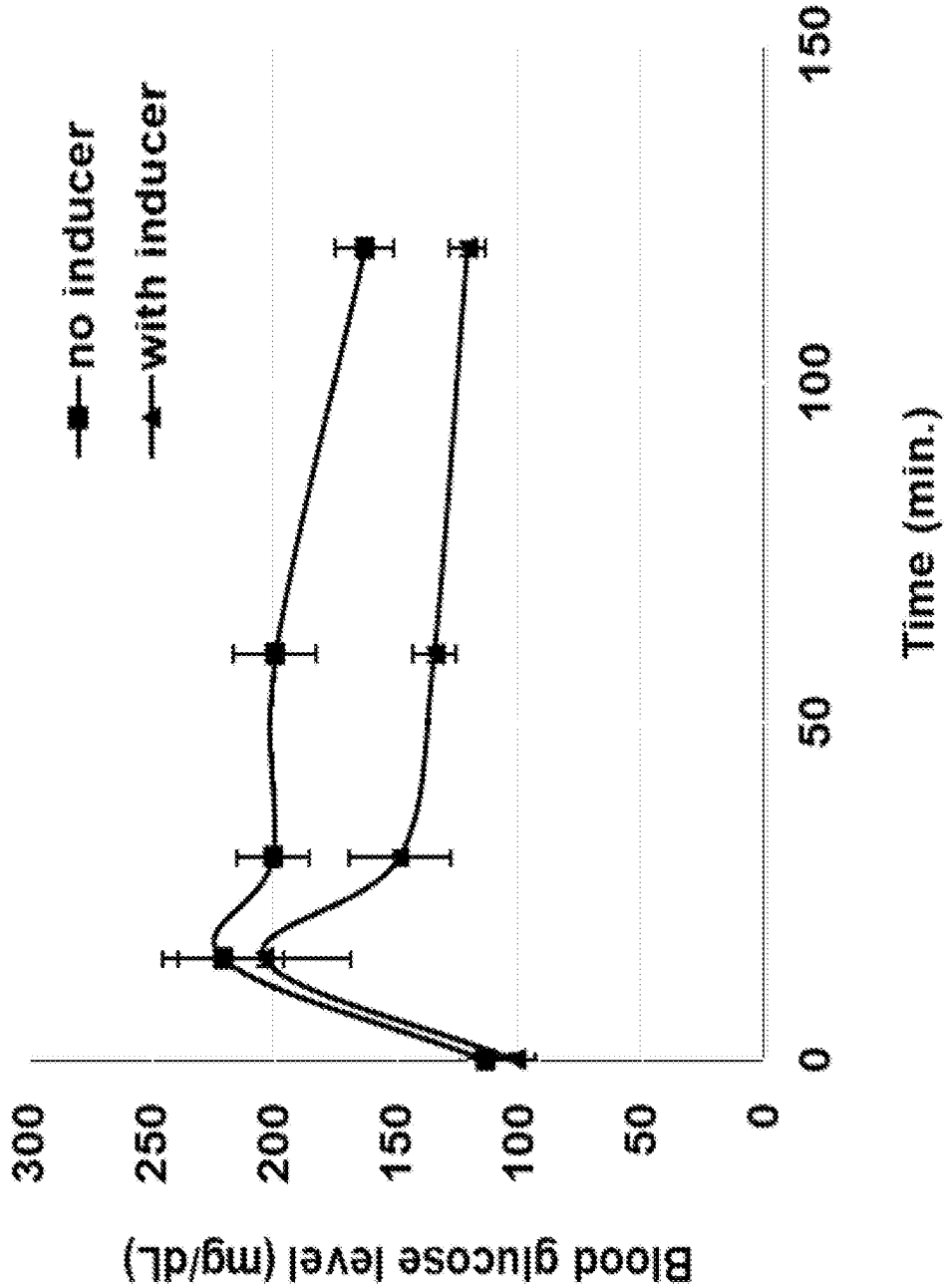

no inducer
with inducer
Blood glucose level (mg/dL)
300
250
200
150
100
50
0
0
50
100
150
Time (min.)

Fig. 6B