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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 January 2023 (19.01.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/284822 A1

(51) International Patent Classification:

C07K 19/00 (2006.01)	*A61P 1/16* (2006.01)
C07K 14/435 (2006.01)	*A61P 3/04* (2006.01)
A61K 39/395 (2006.01)	*A61P 3/06* (2006.01)
C12N 15/62 (2006.01)	*A61P 3/10* (2006.01)
C12N 15/63 (2006.01)	*A61P 13/12* (2006.01)

(21) International Application Number:
PCT/CN2022/105686

(22) International Filing Date:
14 July 2022 (14.07.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PCT/CN2021/106316
14 July 2021 (14.07.2021) CN

(71) Applicant: **BEIJING QL BIOPHARMACEUTICAL CO., LTD.** [CN/CN]; #101, Building 7, 20 Life Science Park Road, Changping District, Beijing 102206 (CN).

(72) Inventors: **ZHANG, Yuanyuan**; #101, Building 7, 20 Life Science Park Road, Changping District, Beijing 102206 (CN). **WU, Bo**; #101, Building 7, 20 Life Science Park Road, Changping District, Beijing 102206 (CN). **GUO, Wei**; #101, Building 7, 20 Life Science Park Road, Changping District, Beijing 102206 (CN). **DONG, Xiaona**; #101, Building 7, 20 Life Science Park Road, Changping District, Beijing 102206 (CN). **ZHANG, Yuying**; #101, Building 7, 20 Life Science Park Road, Changping District, Beijing 102206 (CN).

(74) Agent: **JUN HE LAW OFFICES**; 20/F, China Resources Building 8 Jianguomenbei Avenue, Beijing 100005 (CN).

(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, ME, 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, 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, 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:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: FUSION POLYPEPTIDES FOR METABOLIC DISORDERS

(57) Abstract: Provided are conjugates of polypeptides comprising nanobody and FGF21 linked by a polypeptide linker. Further provided are conjugates of polypeptides comprising GLP-1, nanobody, and FGF21 linked by two polypeptide linkers respectively. Pharmaceutical compositions comprising the same and methods of treating diseases are also provided.

WO 2023/284822 A1

FUSION POLYPEPTIDES FOR METABOLIC DISORDERS

FIELD OF THE INVENTION

[0001] The present invention relates to polypeptides and producing method thereof, pharmaceutical compositions, and methods of using such to prevent and/or treat diseases, especially metabolic diseases.

BACKGROUND

[0002] Fibroblast growth factor 21 (FGF 21), a member of the fibroblast growth factor(FGF) family, is a hormone synthesized in several metabolically active organs and regulates glucose and lipid homeostasis. The biology of FGF21 is intrinsically complicated owing to its diverse metabolic functions in multiple target organs. FGF21 has been reported to function in organs such as liver, adipocytes, pancreas, hypothalamus and muscles tissues (Fisher FM, Annu Rev Physiol, 2016, 78:223).

[0003] Major biologically active fragment of Glucagon-Like Peptide-1 (GLP-1) is a 30 or 31 amino acid peptide fragment (amino acid 7-36 or 7-37 of GLP-1) deriving from the posttranslational processing of the proglucagon peptide. The initial GLP-1 product GLP-1 stimulates insulin synthesis and secretion and has been shown to prevent hyperglycemia in diabetics, especially type 2 diabetes. However, endogenous GLP-1 only has a half-life of approximately 2 minutes, which results in fasting plasma levels of GLP-1 of only 0-15 pmol/L.

[0004] Metabolic disorders are commonly associated with insulin resistance, visceral adiposity, atherogenic dyslipidemia, etc., which pose major and escalating public health and clinical challenge worldwide. However, existing treatment for metabolic diseases faces problems such as short half-life and/or low efficacy.

[0005] Therefore, there is a need for an improved therapeutic solution for treating metabolic diseases.

SUMMARY OF THE INVENTION

[0006] On a first aspect, the present disclosure provides a polypeptide, which is substantially a fusion polypeptide or fusion protein comprising two or three functional domains that are connected via one or two linkers.

[0007] Specifically, in an N-terminus to C-terminus direction, the polypeptide comprises a first fragment and a second fragment, which are connected via a first linker. The first fragment comprises a nanobody domain capable of binding to serum albumin (e.g. human serum albumin), and the second fragment comprises a biologically active FGF21 domain.

[0008] Herein, the FGF21 domain comprises an amino acid sequence having at least 90% sequence identity to, while substantially retaining biological activity of, SEQ ID NO: 1. As such, the FGF21 domain may comprise no more than 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, or 2 amino acid residue mutations relative to SEQ ID NO: 1, as long as such mutation-containing FGF21 variant still retains the biological activity of FGF21. Herein, an amino acid residue mutation can a substitution, an insertion, or a deletion.

[0009] According to some embodiments of the polypeptide, the FGF21 domain may comprise one or more amino acid residue mutations, each at a position selected from positions 121, 168, 171, and 180 relative to SEQ ID NO: 1. Optionally, the one or more amino acid residue mutations in the FGF21 domain may include one, two, three or four of the following three substitutions: N121Q, M168L, P171G and A180E.

[0010] According to some embodiments of the polypeptide, the FGF21 domain may comprise all three substitutions of N121Q, M168L, and A180E, and the amino acid sequence of the FGF21 domain is set forth in SEQ ID NO: 5. Other amino acid residues in the FGF21 domain may be additionally mutated on top of the three mutations. In certain embodiments, the FGF21 domain further comprises a substitution P171G. According to some embodiments of the polypeptide, the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 14.

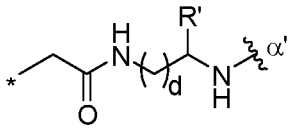
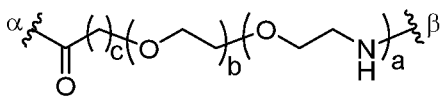
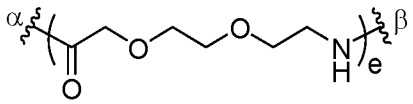
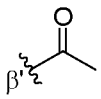
[0011] According to some embodiments of the polypeptide, the FGF21 domain further comprises a conjugatable residue. Herein, the conjugatable residue can optionally be at a position within a C-terminal fragment spanning from position 169 to position 181 relative to SEQ ID NO: 1. According to certain embodiments, the conjugatable residue is at a position selected from the group consisting of positions 169, 170, 171, 172, 173, 174, 180 and 181 relative to SEQ ID NO: 1. According to some embodiments of the polypeptide, the FGF21 domain comprises the amino acid sequence of SEQ ID NOs: 2-5, 89-91, 14, and 102-105, except for a mutation to a conjugatable residue at a position within a C-terminal fragment spanning from position 169 to position 181 relative to SEQ ID NO: 1.

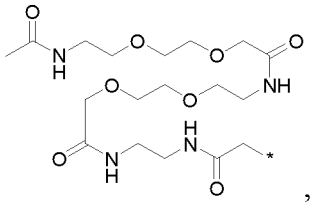
[0012] According to some embodiments of the polypeptide, the FGF21 domain comprises the amino acid sequence of SEQ ID NOs: 6-13, 16-19, and 92.

[0013] According to some embodiments, the polypeptide is conjugated with a functional moiety at the conjugatable residue in the FGF21 domain-containing second fragment.

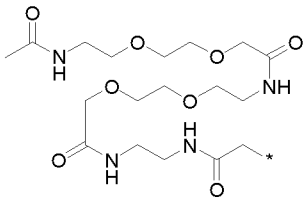
[0014] Herein, the functional moiety that is conjugated to the FGF2 domain may optionally comprise a glycosyl moiety or a synthetic chemical moiety.

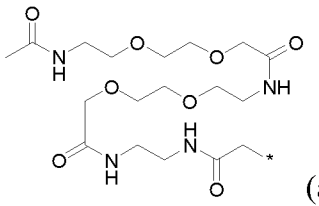
[0015] According to some embodiments of the polypeptide, the functional moiety comprises a glycosyl moiety, and the conjugatable residue can be an introduced residue that is glycosylatable, such as Threonine (T) or Asparagine (N) residues. More specifically, the conjugatable residue may optionally be an introduced T at position 172 or position 173, or may be an introduced N residue at position 170 or position 174, relative to SEQ ID NO: 1. According to some embodiments of the polypeptide, the FGF21 domain comprises the amino acid sequence selected from the group consisting of SEQ ID NOs: 16-19.

[0016] According to some other embodiments of the polypeptide, the functional moiety comprises a synthetic chemical moiety. As such, the conjugatable residue may optionally comprise an introduced cysteine (C), and the synthetic chemical moiety may comprise a structure of *-X-Y-Z. Herein X, Y, and Z are interconnected via bonds, and the * end of X is connected to the conjugatable residue on the polypeptide. X can be ; Y can be , or ; and Z can be , wherein position α is linked to position α', position β is linked to position β'. Herein, R1 can be hydrogen or –COOH; d can be 1, 2, or 3; a can be 1, 2 or 3; b can be 1, 2 or 3; c can be 1 or 2; d can be 1, 2, or 3; and e can be 1, 2, or 3.

[0017] According to some more specific embodiments of the polypeptide, the synthetic chemical moiety has the below structure: , also referred herein as Ac-2XADO-EDA-CO-CH2-*.

[0018] Optionally, the introduced cysteine can be at a position selected from a group consisting of 169, 170, 171, 172, 173, 174, 180 and 181 relative to SEQ ID NO: 1. According to certain embodiments, the conjugatable residue is at the position 171 or 174 relative to SEQ ID NO: 1. According to some of these embodiments of the polypeptide, the FGF21 domain comprises the amino acid sequence of SEQ ID NOs: 6-13 and 92.

[0019] In some embodiments of the polypeptide, the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 8, with the introduced cysteine at position 171, conjugated at the introduced cysteine residue with the synthetic chemical moiety having the below structure:  (also referred herein as Ac-2XADO-EDA-CO-CH2*).

[0020] In some embodiments of the polypeptide, the FGF21 domain can comprise the amino acid sequence of SEQ ID NO: 92, with the introduced cysteine at position 174, conjugated at the introduced cysteine residue with the synthetic chemical moiety having the below structure:  (also referred herein as Ac-2XADO-EDA-CO-CH2*).

[0021] Other embodiments of the polypeptide may alternatively include:

i) the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 6, and the introduced cysteine is at position 169;

ii) the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 7, and the introduced cysteine is at position 170;

iii) the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 8, and the introduced cysteine is at position 171;

iv) the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 9, and the introduced cysteine is at position 172;

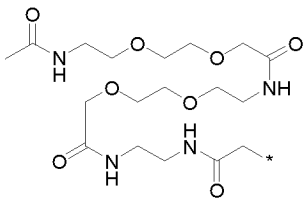
v) the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 10, and the introduced cysteine is at position 173;

vi) the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 11, and the introduced cysteine is at position 174;

vii) the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 12, and the introduced cysteine is at position 180; or

viii) the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 13, and the introduced cysteine is at position 181,

ix) the FGF21 domain comprises the amino acid sequence of SEQ ID NO: 92, and the introduced cysteine is at position 174,

and optionally is conjugated at the introduced cysteine residue with the synthetic chemical moiety having the below structure:  (also referred herein as Ac-2XADO-EDA-CO-CH2*).

[0022] Additionally and alternatively, in some other embodiments of the polypeptide, the FGF21 domain further comprises a substitution of P171G relative to SEQ ID NO: 1. In such embodiments, the FGF21 domain comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 14 and 92.

[0023] In any of the above embodiments of the polypeptide, the nanobody domain may optionally comprise a VHH domain, optionally a VHH domain capable of binding to human serum albumin (HSA).

[0024] Further optionally, the VHH domain can be humanized.

[0025] According to some embodiments, the VHH domain may comprise a complementarity determining region 1 (CDR1), a complementarity determining region 2 (CDR2), and a complementarity determining region 3 (CDR3). Herein, the CDR1 may comprise the sequence of SEQ ID NO: 20 or a variant thereof having up to 3, 2, or 1 amino acid mutation; the CDR2 may comprise the sequence of SEQ ID NO: 21 or a variant thereof having up to 3, 2, or 1 amino acid mutation; and/or the CDR3 may comprise the sequence of SEQ ID NO: 22 or a variant thereof having up to 3, 2, or 1 amino acid mutation, and the VHH domain substantially retains the binding specificity to serum albumin, optionally to human serum albumin.

[0026] In certain specific embodiments of the polypeptide, the VHH domain comprises a complementarity determining region 1 (CDR1) comprising the sequence of SEQ ID NO: 20, a CDR2 comprising the sequence of SEQ ID NO: 21, and a CDR3 comprising the sequence of SEQ ID NO: 22.

[0027] In certain specific embodiments of the polypeptide, the VHH domain comprises an amino acid sequence of SEQ ID NO: 23, or a variant thereof having at least 70% (e.g. at least 75%, 80%, 85%, 90%, 95%, 99% identity to SEQ ID NO: 23, and the variant substantially retains binding specificity and/or affinity to serum albumin.

[0028] Herein, optionally, the variant of SEQ ID NO: 23 may have up to 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid mutation relative to SEQ ID NO: 23.

[0029] In any of the above embodiments, the nanobody domain may optionally further comprise an N-terminal extension attached to a N-terminus of the VHH domain, the N-terminal extension may optionally comprise amino acid residues of SG, AG, S or A.

[0030] According to some specific embodiment of the polypeptide, the nanobody domain comprises an amino acid sequence selected from SEQ ID NOs: 24-27.

[0031] In any embodiments of the polypeptide described above, the first linker can have a length of at least four amino acid residues.

[0032] According to some embodiments, the first linker may comprise no acidic amino acid residue, and more specifically, the first linker may not comprise any of D residue or E residue.

[0033] Optionally, the first linker may comprise one or more units of a first repeating sequence, and according to some embodiments, the first repeating sequence may consist of no more than 4 or 6 types of amino acid residues selected from the group consisting of: G, Q, A, P, T and S.

[0034] According to some specific embodiments of the polypeptide, the first repeating sequence may comprise or consist of an amino acid sequences selected from a group consisting of G_fS_g (herein each of f and g is independently an integer selected from 1 to 5), SEQ ID NO: 35 (GAQP), SEQ ID NO: 36 (GQAP), SEQ ID NO: 37 (GPAQ), SEQ ID NO: 38 (GPQA), SEQ ID NO: 39 (GSQP), SEQ ID NO: 40 (GASP), SEQ ID NO: 41 (GPAS), SEQ ID NO: 42 (GPSA), SEQ ID NO: 43 (GGGS), SEQ ID NO: 44 (GSGS), SEQ ID NO: 45 (GGGGS), SEQ ID NO: 46 (GSAPGSPAGSPTGSAPGSPA), and GS.

[0035] In certain embodiments, the first repeating sequence may have an amino acid sequence set forth in SEQ ID NO: 35 (GAQP), and the number of units may be an integer between 1 and 10.

[0036] According to certain specific embodiment of the polypeptide, the first linker comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 35 (GAQP), SEQ ID NO: 49 $((GAQP)_2)$, SEQ ID NO: 50 $((GAQP)_5)$, SEQ ID NO: 51 $((GAQP)_{10})$, and SEQ ID NO: 48 (GGGGSGGGS).

[0037] In any of the embodiments of the polypeptide described above, the polypeptide may further comprise, over an N-terminus of the first fragment, a third fragment that comprises another functional domain. Herein, the first fragment and the third fragment are connected via a second linker.

[0038] Herein, the another functional domain of the third fragment may optionally comprise a biologically active protein, or a fragment thereof, of one selected from

the group consisting of glucagon-like peptide- 1 (GLP-1), insulin, leptin, glucagon, gastrin, gastric inhibitory polypeptide (GIP), amylin, calcitonin, cholecystokinin, peptide YY, neuropeptide Y, bone morphogenetic protein-6 (BMP-6), bone morphogenetic protein-9 (BMP-9), oxyntomodulin, oxytocin, glucagon-like peptide-2 (GLP-2), irisin, fibronectin type III domain containing protein 5 (FNDC5), apelin, adiponectin, Clq and tumor necrosis factor related protein (CTRP family), resistin, visfatin, omentin, retinol binding protein-4 (RBP-4), glicentin, angiopoietin, interleukin-22 (IL-22), exendin-4, and growth hormone.

[0039] According to some embodiments of the polypeptide, the another functional domain of the third fragment comprises a biologically active peptide, or a fragment thereof, of GLP-1, and may comprise an amino acid sequence having at least 70% sequence identity to, while retaining substantial biological activity of, SEQ ID NO: 28.

[0040] Herein optionally, the another functional domain may comprise one or more mutations at positions 8, 22, 26, 34, and 36, or any combination thereof, relative to SEQ ID NO: 28.

[0041] According to some embodiments of the polypeptide, the one or more mutations in the another functional domain of the third fragment of the polypeptide may comprise A8G, G22E, K26R, K34R, and R36G, or any combination thereof. According to some embodiments of the polypeptide, the one or more mutations in the another functional domain of the third fragment of the polypeptide may comprise A8G, G22E, and R36G, or any combination thereof.

[0042] According to some specific embodiment of the polypeptide, the another functional domain of the third fragment comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 29, and SEQ ID NOs: 31-34.

[0043] In any of the above mentioned embodiments of the polypeptide comprising a first fragment which connects with an N-terminus of the first fragment via a second linker.

[0044] In certain embodiments, the second linker may have a length of at least four amino acid residues. In certain embodiments, the second linker may have a length of at least 8, 12, 16, or 20 amino acid residues.

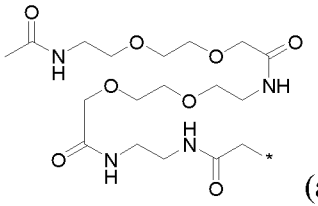
[0045] Herein the second linker may comprise one or more units of a second repeating sequence, and the second repeating sequence may consist of no more than 4 or 6 types of amino acid residues selected from the group consisting of: G, Q, A, E, P, T and S.

[0046] According to some specific embodiments of the polypeptide, the second repeating sequence may comprise or consist of an amino acid sequence that is selected from a group consisting of G_hS_i (herein, each of h and i is independently an integer selected from 1 to 5), SEQ ID NO: 35 (GAQP), SEQ ID NO:55 (GQEP), SEQ ID NO: 56 (GEQP), SEQ ID NO: 57 (GPQE), SEQ ID NO: 58 (GPEQ), SEQ ID NO: 59 (GSEP), SEQ ID NO: 60 (GESP), SEQ ID NO: 61 (GPSE), SEQ ID NO: 62 (GPES), SEQ ID NO: 36 (GQAP), SEQ ID NO: 37 (GPAQ), SEQ ID NO: 38 (GPQA), SEQ ID NO: 39 (GSQP), SEQ ID NO: 40 (GASP), SEQ ID NO: 41 (GPAS), SEQ ID NO: 42 (GPSA), SEQ ID NO: 43 (GGGS), SEQ ID NO: 44 (GSGS), SEQ ID NO: 45 (GGGGS), SEQ ID NO: 46 (GSAPGSPAGSPTGSAPGSPA) and GS.

[0047] Further according to some embodiments, the second repeating sequence may have an amino acid sequence set forth in SEQ ID NO: 35 (GAQP), and a number of the one or more units may be an integer between 1 and 15.

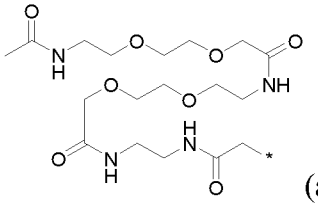
[0048] According to certain specific embodiments of the polypeptide, the second linker comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 49 ((GAQP)2), SEQ ID NO: 50 ((GAQP)5), SEQ ID NO: 51 ((GAQP)10), and SEQ ID NO: 52 ((GAQP)14), and SEQ ID NO: 47 ((GGGGS)4).

[0049] According to certain specific embodiments of the polypeptide which consisting of the first and second fragments, the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 63-68, 93, 99, 100, 101 and 107. According to certain specific embodiments of the polypeptide, the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 63, 68, and 100, and is conjugated at the introduced cysteine residue at position 171 relative to SEQ ID NO: 1 with the synthetic chemical moiety provided herein. According to certain specific embodiments of the polypeptide, the polypeptide comprises an amino acid sequence of SEQ ID NO: 93, and is conjugated at the introduced cysteine residue at position 174 relative to SEQ ID NO: 1 with the synthetic chemical moiety provided herein. According to certain of these specific embodiments of the polypeptide, the synthetic chemical moiety has the below structure:



(also referred herein as Ac-2XADO-EDA-CO-CH2*). According to certain specific embodiments of the polypeptide, the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 64-67, 93 and 107, and is not conjugated.

[0050] According to certain specific embodiments of the polypeptide which consisting of the first, second and third fragments, the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 70, 74, 75, 79-83, 85, 94-98, and 108-109. According to certain specific embodiments of the polypeptide, the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 74, 81-83, 85, and 96-97, and is conjugated at the introduced cysteine residue at position 171 relative to SEQ ID NO: 1 with the synthetic chemical moiety provided herein. According to certain specific embodiments of the polypeptide, the polypeptide comprises an amino acid sequence of SEQ ID NO: 94, and is conjugated at the introduced cysteine residue at position 174 relative to SEQ ID NO: 1 with the synthetic chemical moiety. According to certain of these specific embodiments of the polypeptide, the synthetic chemical moiety has the below structure:



(also referred herein as Ac-2XADO-EDA-CO-CH2*). According to certain specific embodiments of the polypeptide, the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 70, 75, 79-80, 85, 94, 95, 98 and 108-109, and is not conjugated.

[0051] On a second aspect of the present disclosure, a pharmaceutical composition is further provided. The pharmaceutical composition comprises the polypeptide according to any embodiments as described above, and a pharmaceutically acceptable carrier.

[0052] On a third aspect of the present disclosure, a method of preventing or treating a metabolic disorder in a subject in need thereof is also provided, comprising the step of administering a therapeutically effective amount of the polypeptide according to any embodiments as described above, or the pharmaceutical composition described above.

[0053] Herein, the metabolic disorder can be diabetes, obesity, non-alcoholic steatohepatitis (NASH), cardiovascular like dyslipidaemia, arteriosclerosis, alcoholic steatohepatitis (ASH), diabeticnephropathy, gestational diabetes, metabolic syndrome such as metabolic syndrome X, nonalcoholic fatty liver disease (NAFLD), end-stage liver disease, hepatic steatosis (fatty liver), liver cirrhosis, primary biliary cirrhosis (PBC) or severe hypertriglyceridemia (SHTG).

[0054] On a fourth aspect of the present disclosure, a polynucleotide is further provided, which encodes the polypeptide as defined according to any embodiments as described above.

[0055] On a fifth aspect of the present disclosure, a vector is further provided, which comprises the polynucleotide as described above.

[0056] On a sixth aspect of the present disclosure, a host cell is further provided, which comprises the vector as described above.

[0057] On a seventh aspect of the present disclosure, a process for producing the polypeptide as described above is further provided. The process comprises the steps of:

S100: culturing the host cell as described above under a condition that allows expression of the polynucleotide as defined above or a precursor thereof further comprising a removable tag; and

S200: collecting and purifying the polypeptide or the precursor thereof from the host cell.

[0058] According to some embodiments of the process, the polypeptide is conjugated with a functional moiety at a conjugatable residue in the second fragment, and the process further comprises the following step after step S200 of collecting and purifying the polypeptide from the host cell:

S300: conjugating the functional moiety to the polynucleotide.

[0059] According to some embodiments of the process, the host cell is *E. coli*, the vector comprises an *E. coli*-compatible vector, and the polypeptide encoded by the polynucleotide in the vector is codon optimized for *E. coli* expression.

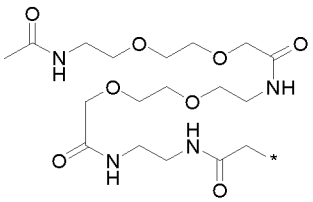
[0060] According to some embodiments of the process, step S200 of collecting and purifying the polypeptide from the host cell may comprise the following sub-steps:

S210: collecting the precursor of the polypeptide;

S220: allowing the precursor of the polypeptide to refold;

S230: treating the refolded precursor of the polypeptide to remove the tag thereby obtain the polypeptide; and

S240: purifying the polypeptide.

[0061] According to some embodiments of the polypeptide producing process, the process further comprises conjugating the purified polypeptide with the functional moiety. In certain embodiments, the functional moiety to be conjugated to the polypeptide is  (Ac-2XADO-EDA-CO-CH2-*). Throughout the present disclosure, the articles "a," "an," and "the" are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, "a polypeptide" means one polypeptide or more than one polypeptides.

[0062] In all occurrences in this application where there are a series of recited numerical values, it is to be understood that any of the recited numerical values may be the upper limit or lower limit of a numerical range. It is to be further understood that the invention encompasses all such numerical ranges, i.e., a range having a combination of an upper numerical limit and a lower numerical limit, wherein the numerical value for each of the upper limit and the lower limit can be any numerical value recited herein. Ranges provided herein are understood to include all values within the range. For example, 1-10 is understood to include all of the values 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10, and fractional values as appropriate. Similarly, ranges delimited by "at least" are understood to include the lower value provided and all higher numbers.

[0063] As used herein, the term "approximately", "about", "around" or alike, is understood to include within three standard deviations of the mean or within standard ranges of tolerance in the specific art. In certain embodiments, about is understood a variation of no more than 0.5.

[0064] The term "or" is used inclusively herein to mean, and is used interchangeably with, the term "and/or," unless context clearly indicates otherwise.

[0065] The terms such as "comprise", "comprising", "contain", "containing" or alike, are to be construed as open-ended terms (i.e., meaning "including, but not limited to,") unless otherwise noted.

[0066] The term "including" is used herein to mean, and is used interchangeably with, the phrase "including but not limited to". Similarly, "such as" is used herein to mean, and is used interchangeably, with the phrase "such as but not limited to".

[0067] The term "consist essentially of" implies the exclusion of all, most or all but a negligible amount of other elements, or it allows for elements not explicitly recited, but excludes elements that are found in the prior art or that affect a basic or novel property. When a polypeptide is said to "consist essentially of" an amino acid sequence, this means that the polypeptide is composed mostly of the amino acid sequence, with no more than 10 (for example, no more than 6, 5, 4, 3, 2, or 1) additional amino acid residues at one end or both ends of the polypeptide.

BRIEF DESCRIPTION OF THE FIGURES

[0068] FIGS 1A and 1B illustrate a polypeptide according to two different embodiments of the present disclosure.

[0069] FIGS 2A and 2B show *in vitro* FGF21 activity of the Molecules MLC#9, MLC#10, MLC#14, Control #6 and a reference compound YH-dual.

[0070] FIGS 3A and 3B show efficacy of Molecules Control#2, Control#6, MLC#9 and MLC#10 on body weight reduction (Figure 3A) and food intake suppression (Figure 3B) in DIO rats.

[0071] FIGS 4A to 4L show efficacy of Molecules MLC#9, MLC#14, MLC#16 and MLC#17, and reference compounds semaglutide, YH-dual, and Tirzepatide on body weight reduction (Figure 4A), glucose control (Figure 4B), reduction of triglyceride (Figure 4C), LDL-C (Figure 4D), total cholesterol (Figure 4E) and ALT (Figure 4F) concentration in plasma and reduction of adipose weight (Figure 4G) and liver weight (Figure 4H), along with the improvement of liver TG (Figure 4I), liver TC (Figure 4J) and insulin sensitivity (Figure 4K), and increase in adiponectin level (Figure 4L) in DIO mice.

[0072] FIGS 5A to 5M show efficacy of Molecules MLC#14 and MLC#16 at different dosages, in comparison to reference compounds semaglutide, YH-dual, and Tirzepatide on body weight reduction (Figure 5A), glucose control (Figure 5B), reduction of

triglyceride (Figure 5C), LDL-C (Figure 5D), total cholesterol (Figure 5E), ALT (Figure 5F) and AST (Figure 5G) concentration in plasma and reduction of adipose weight (Figure 5H) and liver weight (Figure 5I) along with the improvement of liver TG (Figure 5J), liver TC (Figure 5K) and insulin sensitivity (Figure 5L), and increase in adiponectin level (Figure 5M) in DIO mice.

[0073] FIGS 6A to 6H show efficacy of Molecules MLC#16, MLC#19, MLC#6, and MLC#23 at different dosages, in comparison to reference compounds semaglutide, serum triglyceride (Figure 6C), LDL-C (Figure 6D) and total cholesterol (Figure 6E) concentration in plasma and reduction of adipose weight (Figure 6F) and liver weight (Figure 6G), along with the improvement of liver TG (Figure 6H).

[0074] FIG 7 show all the amino acid sequences disclosed in the present disclosure.

DETAILED DESCRIPTION

[0075] The following description of the invention is merely intended to illustrate various embodiments of the invention. As such, the specific modifications discussed are not to be construed as limitations on the scope of the invention. It will be apparent to one skilled in the art that various equivalents, changes, and modifications may be made without departing from the scope of the invention, and it is understood that such equivalent embodiments are to be included herein. All references cited herein, including publications, patents and patent applications are incorporated herein by reference in their entirety.

Definitions

[0076] The term “protein”, “peptide” and "polypeptide" are used interchangeably herein and refer a polymer of amino acid residues linked by covalent bonds such as peptide bonds. A protein or polypeptide as provided herein can comprise naturally occurring or non-natural amino acid residues, or both. Polypeptides, peptides and proteins provided herein can comprise any suitable length of amino acid residues, for example, from at least 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, or more amino acid residues in length.

[0077] The term "naturally occurring" amino acid residue, as used herein, refers to an amino acid residue found in native proteins or peptides, all in their D and L stereoisomers

if their structure allows such stereoisomeric forms. Examples of naturally occurring amino acid residues include, but not limited to, 20 standard amino acids, including, glycine (Gly or G), alanine (Ala or A), valine (Val or V), leucine (Leu or L), isoleucine (Ile or I), serine (Ser or S), cysteine (Cys or C), threonine (Thr or T), methionine (Met or M), proline (Pro or P), phenylalanine (Phe or F), tyrosine (Tyr or Y), tryptophan (Trp or W), histidine (His or H), lysine (Lys or K), arginine (Arg or R), aspartate (Asp or D), glutamate (Glu or E), asparagine (Asn or N), and glutamine (Gln or Q), and their natural analogs, such as canavanine, pyrrolysine (PYL), selenocysteine, pyrroline-carboxy-lysine (PCL), Sarcosine, beta-Alanine, phosphoserine, γ-carboxyglutamate, and ornithine. Examples of naturally occurring amino acid residues in their D stereoisomer include, for example, D-aspartate, D-Serine, D-Cysteine, D-Alanine, D-glutamate and so on.

[0078] An "amino acid analog" is a compound that has the same basic chemical structure as a naturally occurring amino acid, i.e., an a carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs can have modified R groups {e.g., norleucine) or modified peptide backbones, but will retain the same basic chemical structure as a naturally occurring amino acid.

[0079] A "non-natural" amino acid residue, as used herein, refers to any amino acid residues that are not found in nature, including without limitation, a modified amino acid residue, and/or an amino acid mimetic, which is not one of the known naturally occurring amino acids, yet functions in a manner similar to the naturally occurring amino acids. Modified amino acid or a mimetic can be generated by addition of a chemical entity such as a carbohydrate group, a phosphate group, a farnesyl group, an isofamesyl group, a fatty acid group, a linker for conjugation, functionalization, or other modification, etc. A non-natural amino acid can also refer to an amino acid manufactured by chemical synthesis. Exemplary non-natural amino acids include, but not limited to, 2-Aminoisobutyric acid (Aib), imidazole-4-acetate (IA), imidazolepropionic acid (IPA), a-aminobutyric acid (Abu), tert-butylglycine (Tle), 3-aminomethyl benzoic acid, anthranilic acid, des-amino-histidine (abbreviated DesaminoHis, alternative name imidazopropionic acid, abbreviated Impr), the beta analogues of amino acids such as β-alanine, 2-amino-histidine, β-hydroxy-histidine, homohistidine, Nα-acetyl-histidine, α-fluoro-methyl-histidine, α-methyl-histidine, α,α-dimethyl-glutamic acid, m-CF3-phenylalanine, α,β-diaminopropionic acid (abbreviated Dap), 3-pyridylalanine, 2-pyridylalanine or 4-pyridylalanine, (1-aminocyclopropyl)carboxylic acid, (1-

aminocyclobutyl)-carboxylic acid, (1-aminocyclopentyl)carboxylic acid, (1-aminocyclohexyl)carboxylic acid, (1-aminocycloheptyl)carboxylic acid, and (1-aminocyclooctyl)carboxylic acid. Introduction of non-natural amino acids into a fusion polypeptide, polypeptide fragment, and/or polypeptide complex may be realized by the technology described in Wang et al., Science 292:498-500, 2001; Deiters et al., J Am Chem Soc 125:1 1782-1 1783, 2003; Wang and Schultz, Science 301 :964- 967, 2003; Zhang et al., Science 303:371-373, 2004 or in US Patent No. 7,083,970. Briefly, some of these expression systems involve site-directed mutagenesis to introduce a stop codon, such as an amber (UAG), ochre (UAA), and opal (UGA) codons) into the open reading frame encoding a fusion polypeptide of the present disclosure. Other codons, such as a four-base codon (e.g. AGGA, AGGU, CGGU, CGCU, CGAU, CCCU, CUCU, CUAU, and GGGU), a five- base codon, a six-base codon, etc. can also be introduced into the expression systems for non-natural amino acids. Such expression vectors are then introduced into a host that can utilize a tRNA specific for the introduced stop codon or other codons and carried with the non-natural amino acid of choice.

[0080] The term "fusion" or "fused" when used with respect to amino acid sequences (e.g. peptide, polypeptide or protein) refers to combination of two or more amino acid sequences, for example by chemical bonding or recombinant means, into a single amino acid sequence which does not exist naturally. A fusion amino acid sequence may be produced by genetic recombination of two encoding polynucleotide sequences, and can be expressed by a method of introducing a construct containing the recombinant polynucleotides into a host cell.

[0081] "Percent (%) sequence identity" is defined as the percentage of amino acid (or nucleic acid) residues in a candidate sequence that are identical to the amino acid (or nucleic acid) residues in a reference sequence, after aligning the sequences and, if necessary, introducing gaps, to achieve the maximum number of identical amino acids (or nucleic acids). In other words, percent (%) sequence identity of an amino acid sequence (or nucleic acid sequence) can be calculated by dividing the number of amino acid residues (or bases) that are identical relative to the reference sequence to which it is being compared by the total number of the amino acid residues (or bases) in the candidate sequence or in the reference sequence, whichever is shorter. Conservative substitution of the amino acid residues is not considered as identical residues. Alignment for purposes of determining percent amino acid (or nucleic acid) sequence identity can be achieved, for example, using publicly available tools such as BLASTN, BLASTp (available on the website of U.S. National Center for Biotechnology

Information (NCBI), see also, Altschul S.F. et al, J. Mol. Biol., 215:403–410 (1990); Stephen F. et al, Nucleic Acids Res., 25:3389–3402 (1997)), ClustalW2 (available on the website of European Bioinformatics Institute, see also, Higgins D.G. et al, Methods in Enzymology, 266:383-402 (1996); Larkin M.A. et al, Bioinformatics (Oxford, England), 23(21): 2947-8 (2007)), and ALIGN or Megalign (DNASTAR) software. Those skilled in the art may use the default parameters provided by the tool, or may customize the parameters as appropriate for the alignment, such as for example, by selecting a suitable algorithm.

[0082] In one aspect, the present disclosure provides a polypeptide which is substantially a fusion polypeptide/protein comprising two or three functional domains, and every two neighboring functional domains is operably connected via a linker. Optionally, the fusion polypeptide/protein can be conjugated.

[0083] FIG. 1A and FIG. 1B illustrate two major embodiments of the polypeptide provided in this present disclosure. As shown in FIG. 1A, the first embodiment of the polypeptide 001A comprises, in an N-terminus to C-terminus direction, a first fragment 100 and a second fragment 200, which are connected via a first linker 10. The first fragment 100 substantially comprises a nanobody domain that is capable of binding to serum albumin, such as human serum albumin (HSA). The second fragment 200 comprises a biologically active FGF21 domain, which substantially comprises a functional form of FGF21, and according to certain specific embodiments of the disclosure, comprises an amino acid sequence having at least 90% sequence identity to, while substantially retaining biological activity of, SEQ ID NO: 1.

[0084] Optionally, the polypeptide may further comprise yet one additional functional domain besides the nanobody domain in the first fragment 100 and the FGF21 domain in the second fragment 200. As shown in FIG. 1B, in addition to the first fragment 100 and the second fragment 200 as in the first embodiment illustrated in FIG. 1A, this second embodiment of the polypeptide 001B further includes a third fragment 300 over an N-terminus of the first fragment 100. The third fragment 300 comprises an additional functional domain that can exert a biological activity that is additive to, or in synergy with, the FGF21 domain in the second fragment 200. The first fragment 100 and the third fragment 300 are connected via a second linker 20. According to some embodiments of the present disclosure, this additional functional domain in the third fragment 300 is a biologically active glucagon-like peptide-1 (GLP-1) domain, which substantially comprises a functional form of GLP-1. According to

some embodiments, comprises an amino acid sequence having at least 70% sequence identity to, while substantially retaining the biological activity of, SEQ ID NO: 28.

[0085] The term “functional form” as used herein, refers to different forms (such as variants, fragments/portions, fusions, derivatives and mimetics) of the parent molecule, which, despite of having difference in amino acid sequences or in chemical structures, still retains substantial biological activity of the parent molecule. The expression “retain substantial biological activity”, as used herein, means exhibiting at least part of (for example, no less than about 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%) or all of the biological activity of the parent molecule. A functional form of a parent polypeptide may include both naturally-occurring variant forms and non-naturally occurring forms such as those obtained by recombinant methods or chemical synthesis. The functional forms may contain non-natural amino acid residues.

[0086] The term "variant" as used herein refers to a polypeptide having at least 70% sequence identity to the parent polypeptide and retain at least partial function of the parent polypeptide. A variant may differ from the parent polypeptide by one or more amino acid residues. For example, a variant may have substitutions, additions, deletions, insertions, or truncations of one more amino acid residue of the parent polypeptide.

[0087] The term "fragment" or “portion” as used herein refers to partial sequence of the parent polypeptide of any length. A fragment may still retain at least partial function of the parent polypeptide.

[0088] The term "derivative" as used herein refers to a chemically modified polypeptide or fusion polypeptide, in which one or more well-defined substituent groups have been covalently attached to one or more specific amino acid residues of the polypeptide or fusion polypeptide. Exemplary chemical modification can be, e.g. alkylation, acylation, esterification, amidation, phosphorylation, glycosylation, labeling, methylation of one or more amino acids, or conjugation with one or more moieties.

[0089] The term "mimetics" as used herein refers to molecular structures that serve as substitutes for amino acids, peptides, polypeptides, or fusion polypeptide. For example, amino acid mimetics, as used herein, can be synthetic structures (either known or yet unknown), which may or may not be an amino acid, but retain the functional features of the parent amino acids while the structure of the amino acid mimetic is different from the structure of the parent

amino acid. Examples include a methacryloyl or acryloyl derivative of an amide, β-, γ-, δ-imino acids (such as piperidine-4-carboxylic acid) and the like.

[0090] **FGF21 Domain**

[0091] The polypeptides disclosed herein comprises the FGF21 domain. For example, each of the first and second embodiments of the polypeptide disclosed herein and illustrated in FIGS. 1A and 1B comprises the FGF21 domain.

[0092] The term "FGF21" as used herein refers to, and is short for, "Fibroblast Growth Factor 21", and is intended to broadly encompasses all functional forms including the native human FGF21, as well as functional variants, fragments, fusions, derivatives or mimetics thereof. The native human FGF21 consists of 209 amino acid residues (Uniprot datebase with access no. Q9NSA1), with amino acid residues 1-28 being a signal polypeptide and amino acid residues 29-209 being a mature polypeptide of 181 residues. The expression "biologically active FGF21 domain" as provided in the present disclosure refers to a functional form of FGF21, which can be the mature polypeptide, or functional variants, fragments, fusions, derivatives or mimetics thereof. The mature polypeptide of human FGF21 is included herein as SEQ ID NO: 1. As used herein, the numbering of the residues in FGF21 is referred to the sequence of SEQ ID NO: 1, which begins with His at position 1 and Ser at position 181.

[0093] A functional form of the mature polypeptide of FGF21 is capable of activating the FGF21 receptor at a level comparable to, or no less than about 20% (or no less than 30%, 40%, 50%, 60%, 70%, 80%, 90%) of, that of the mature polypeptide of the native human FGF21. Activation of the FGF21 receptor can lead to biological activities such as, for example, the ability to activate glucose uptake in adipocytes, the ability to lower blood glucose and triglyceride levels, or the ability to lower bodyweight (Tezze C et al, Front Physiol. 2019, 10:419) . Many functional forms of the mature polypeptide of FGF21 are known in the art, for example, without limitation, those disclosed in, WO2019043457A2, WO2018088838A1, WO2018039081A1, WO2017220706A1, WO2017180988A2, WO2017116207A1, WO2017093465, WO2017059371A1, WO2016102562A1, WO2016065326A, WO2013173158A1, WO2013052311A1, WO2013033452A2, WO2012066075A1, WO2012059873A2, WO2012010553A1, WO2011140086A2, WO2010084169A2, WO2010065439A1, WO2008121563A2, WO2006028595A2, WO2006028714A1, WO2005113606A2, WO2016102562A, disclosure of which are incorporated herein by their entirety.

[0094] In certain embodiments, the FGF21 domain provided herein comprises an amino acid sequence having at least 90% sequence identity to, while retaining substantial biological activity of, SEQ ID NO: 1.

[0095] In certain embodiments, the FGF21 domain comprises no more than 12, 11, 10, 9, 8, 7 amino acid mutations (e.g. substitutions, insertions, or deletions) relative to SEQ ID NO: 1 while retaining substantial biological activity of SEQ ID NO: 1. In certain embodiments, the FGF21 domain comprises the amino acid sequence of SEQ ID NOs: 2-14, 16-19, 89-92, and 102-105.

[0096] In certain embodiments, the FGF21 comprises one or more mutations. In certain embodiments, the one or more mutations comprise a conservative substitution. In certain embodiments, the one or more mutations is at a position selected from positions consisting of 121, 168, 171 and 180 relative to SEQ ID NO: 1.

[0097] In certain embodiments, the one or more mutations in the FGF21 domain are selected from N121Q, M168L, P171G and A180E, or any combination thereof. For example, the FGF21 domain in the polypeptide disclosed herein may optionally contain one of the four mutations, two of the four mutations, three of the four mutations or four of the mutations.

[0098] In certain embodiments, the FGF21 domain comprises of a combination of mutations relative to SEQ ID NO: 1 selected from the group consisting of:

1) N121Q,
2) M168L;
3) A180E;
4) N121Q and M168L;
5) N121Q and A180E;
6) M168L and A180E;
7) N121Q, P171G and A180E;
8) N121Q, M168L, and P171G;
9) M168L, P171G and A180E; and
10) N121Q, M168L, P171G and A180E.

[0099] In certain embodiments, the FGF21 domain may comprise an amino acid sequence set forth in SEQ ID NO: 2 (N121Q), SEQ ID NO: 3 (M168L), SEQ ID NO: 4 (A180E), SEQ ID NO: 5 (N121Q, M168L, and A180E), SEQ ID NO: 89 (N121Q and M168L), SEQ ID NO: 90 (N121Q and A180E), SEQ ID NO: 91 (M168L and A180E), or SEQ ID NO: 14 (N121Q, M168L, P171G and A180E).

[00100] In certain embodiments, the FGF21 domain in the polypeptide further comprises a conjugatable residue or up to one conjugatable residue, to which the functional moiety is conjugated.

[00101] The term "conjugatable residue" as used herein, refers to an amino acid residue at a particular position within the polypeptide, which not only has a functional group capable of being chemically or enzymatically conjugated to a chemical moiety, but its position in the polypeptide allows it to be prone to such conjugation in a conjugating reaction for that particular functional group. "Conjugation" as used herein refers to a reaction that joins two molecules together to form one physical entity. For example, a covalent bond linking the two molecules can be formed in conjugation. A conjugatable residue can be a natural amino acid residue, a non-natural amino acid residue, modified amino acid residue or an amino acid mimetic. Examples of conjugatable residues include, without limitation, lysine, cysteine, or a non-natural amino acid residue.

[00102] A skilled person would understand that, whether a residue is conjugatable residue would depend on the given conjugating reaction, and/or the functional group to be conjugated. For example, if the function group to be conjugated is a thiol group and/or the conjugation reaction is specific or selective for a thiol group, then lysine (i.e. having no functional thiol group in its side chain) would not be the conjugatable residue regardless of its position, whereas an unpaired cysteine residue that does not form a disulfide bond (whether intrachain bond or interchain bond), could be a conjugatable residue.

[00103] In certain embodiments, the polypeptide comprises a conjugatable residue for a conjugation reaction specific or selective for thiol group (for example, a maleimide reaction, or a reaction for disulfide bond formation). The expression "conjugatable" when referred to an amino acid residue is intended to mean that, the residue at a particular position in the polypeptide is sufficiently accessible for conjugation. For example, a cysteine residue as part of a disulfide bridge is not a conjugatable residue in the present disclosure.

[00104] In certain embodiments, the conjugatable residue is cysteine residue, preferably a free cysteine residue which is not part of a disulfide bridge.

[00105] In certain embodiments, the conjugatable residue CR in the FGF21 domain of the polypeptide provided herein can be at a position within a C-terminal fragment spanning of the FGF21 domain, e.g. from position 169 to position 181 relative to SEQ ID NO: 1. Further optionally, the conjugatable residue can be at a position selected from the group consisting of positions 169, 170, 171, 172, 173, 174, 180 and 181 relative to SEQ ID NO: 1. Without wishing to be bound by any theory, it is found by the present inventors that conjugation at C-terminal fragment of the FGF21 domain reduces the C-terminal degradation of the FGF21 domain.

[00106] In certain embodiments, the FGF21 domain in the polypeptide has a combination of mutations of: 1) N121Q and M168L, 2) N121Q, M168L and P171G, 3) N121Q, M168L and A180E, or 4) N121Q, M168L, P171G and A180E and further comprises a conjugatable residue (e.g. a cysteine residue) at position 169, 170, 171, 172, 173, 174, or 180, 181 relative to SEQ ID NO: 1.

[00107] In certain embodiments, the FGF21 domain in the polypeptide has a combination of mutations of N121Q and M168L, or a combination mutations of N121Q, M168L and P171G, and further comprises a conjugatable residue (e.g. a cysteine residue) at position 180 relative to SEQ ID NO: 1.

[00108] In certain embodiments, the FGF21 domain in the polypeptide has a combination of mutations of N121Q, P171G and M168L, or a combination mutations of N121Q, M168L, P171G and A180E, and further comprises a conjugatable residue at a position relative to SEQ ID NO: 1, listed as follows: introduced cysteine residue (e.g. via substitution) at position 169, position 170, position 172, position 173, or position 174.

[00109] In certain embodiments, the FGF21 domain in the polypeptide has an amino acid sequence comprising has an introduced cysteine residue at position 169 (e.g. SEQ ID NO: 6), position 170 (e.g. SEQ ID NO: 7), position 171 (e.g. SEQ ID NO: 8), position 172 (e.g. SEQ ID NO: 9), position 173 (e.g. SEQ ID NO: 10), position 174 (e.g. SEQ ID NO: 11, or SEQ ID NO: 92), position 180 (e.g. SEQ ID NO: 12), or position 181 (e.g. SEQ ID NO: 13). In certain embodiments, the FGF21 comprises the amino acid sequence of SEQ ID NOs: 2-5, 89-91, 14, and 102-105, except for one or more amino acid residue mutations, each at a position selected from positions 121, 168, 171, and 180 relative to SEQ ID NO: 1.

[00110] In certain embodiments, the FGF21 domain in the polypeptide has an introduced G at position 171 (e.g. SEQ ID NO: 14). Without wishing to be bound by any theory, it is found by the present inventors that introduction of 171G in the FGF21 domain could also be useful to reduce the C-terminal degradation of the FGF21 domain.

[00111] In certain embodiments, the FGF21 domain in the polypeptide have a combination of mutations of N121Q, P171G and M168L, or a combination mutations of N121Q, M168L, P171G and A180E, and further comprises a conjugatable residue at a position relative to SEQ ID NO: 1, listed as follows: an introduced T at position 172, position 173, position 174, wherein the position is relative to SEQ ID NO: 1; or introduced N at position 170 or position 174, wherein the position is relative to SEQ ID NO:1.

[00112] In certain embodiments, the FGF21 domain in the polypeptide has an introduced T at position 172 (e.g. SEQ ID NO: 16) or at position 173 (e.g. SEQ ID NO: 17); or an introduced N at position 170 (e.g. SEQ ID NO: 18) or at position 174 (e.g. SEQ ID NO: 19), relative to SEQ ID NO: 1. In certain embodiments, the FGF21 domain in the polypeptide has the amino acid sequence selected from the group consisting of SEQ ID NOs: 16-19.

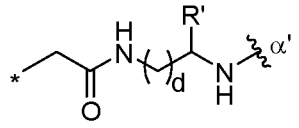
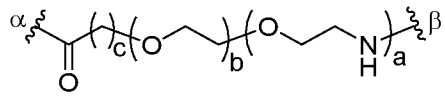
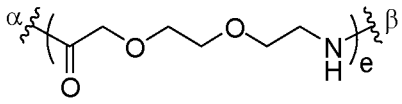
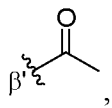
[00113] **Functional moiety on the FGF21 domain**

[00114] In the first and the second embodiments of the polypeptide shown in FIG. 1A and FIG. 1B, either or both of the polypeptide 001A and 001B can optionally be conjugated with a functional moiety 800 at a conjugatable residue CR (as illustrated by the "*" in both figures) in the second fragment 200 (i.e. the FGF21 domain) to thereby form a polypeptide conjugate. In certain other embodiments, the polypeptide is not conjugated with the functional moiety, for example, when the polypeptide has an introduced G at position 171.

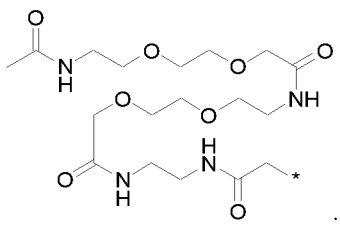
[00115] The term "conjugate" as used herein refers to a compound as a result of two or more molecules joined together to form one physical entity. For example, the polypeptide of the present disclosure can, according to certain embodiments, form a polypeptide conjugate, which is substantially a compound resulting from the polypeptide and a functional moiety being joined together. The molecules (e.g. the functional moiety and the polypeptide) may attach together by covalent bonds, non-covalent bonds, linkers, chemical modification, or protein fusion or by any means known to one skilled in the art. Preferably, the molecules may attach together by covalent bonds. The attachment may be permanent or reversible. In some embodiments, certain cleavable or non-cleavable linkers may be included.

[00116] The term "functional moiety" as used herein refers to a moiety that can functionally alter the biological, pharmacokinetic (PK), pharmacodynamic (PD) properties (e.g. to enhance the biological activity, to increase the stability *in vitro*, to increase the half-life *in vivo*, or to enhance the binding to target receptor, etc.). The functional moiety that is conjugated to the FGF2 domain may optionally comprise a glycosyl moiety or a synthetic chemical moiety.

[00117] Optionally, the functional moiety 800 can comprise a synthetic chemical moiety, and the conjugatable residue CR can be an introduced residue that can be conjugated with the synthetic chemical moiety.

[00118] According to some embodiments, the conjugatable residue CR can be an introduced cysteine residue, in any of the above mentioned positions 169, 170, 171, 172, 173, 174, 180 and 181 relative to SEQ ID NO: 1. Herein, the synthetic chemical moiety can optionally comprise a structure of *-X-Y-Z, wherein X, Y, and Z are interconnected via bonds, and the * end of X is connected to the conjugatable residue on the polypeptide. Herein, X can be ; Y can be , or ; and Z can be , wherein position α is linked to position α', position β is linked to position β'. Herein, R1 can be hydrogen or –COOH; d can be 1, 2, or 3; a can be 1, 2 or 3; b can be 1, 2 or 3; c can be 1 or 2; and d can be 1, 2, or 3, and e can be 1, 2, or 3.

[00119] According to some embodiments, the synthetic chemical moiety can be Ac-2XADO-EDA-CO-CH_2*, where the * end is connected to the conjugatable cysteine residue on the polypeptide, having a structure of:

.

[00120] In each of the above embodiments of the FGF21 domain comprising an introduced cysteine residue at their respective positions as indicated above whose sequences are set forth in SEQ ID NOs: 6-13 and 92, a synthetic chemical moiety, such as Ac-2XADO-EDA-CO-CH_2-* can be conjugated at the introduced cysteine residue (e.g. via substitution) to thereby obtain a polypeptide conjugate carrying the synthetic chemical moiety.

[00121] In the polypeptide carrying a conjugated functional moiety, the functional moiety 800 can optionally comprise a glycosyl moiety, and correspondingly, the conjugatable residue CR can be an introduced residue that is glycosylatable (i.e. can be glycosylated), which can be an introduced T at position 172 or position 173, or an introduced N residue at position 170 or position 174, relative to SEQ ID NO: 1.

[00122] In each of the above embodiments of the FGF21 domain comprising an introduced T or N at their respective positions as indicated above whose sequences are set forth in SEQ ID NOs: 16-19, a glycosyl moiety can be conjugated at the introduced cysteine residue to thereby obtain a polypeptide conjugate carrying the glycosyl moiety.

[00123] In certain embodiments of the FGF21 domain comprise an introduced G at position 171 relative to SEQ ID NO: 1, and is not conjugated with the functional moiety. In certain embodiments, such FGF21 domain can comprise an amino acid sequence of SEQ ID NO: 14.

[00124] **<u>Nanobody Domain</u>**

[00125] Each of the first and second embodiments of the polypeptide disclosed herein and illustrated in FIGS. 1A and 1B comprises the nanobody domain.

[00126] The term “nanobody” as used herein is considered to be exchangeable with “single-domain antibody”, which refers to an antibody fragment containing a single variable domain of a heavy chain or a single variable domain of a light chain. A nanobody contains three complementarity determining regions (CDRs). In certain embodiments, the nanobody is capable of binding to a specific antigen (e.g., serum albumin).

[00127] The term “serum albumin” as used herein refers to an albumin (a type of globular protein) found in vertebrate blood. Serum albumin is produced by the liver, occurs dissolved in blood plasma and is the most abundant blood protein in mammals. Serum albumin typically has a half-life of around three weeks, which is mainly regulated by neonatal Fc receptor (FcRn). FcRn protects serum albumin from intracellular degradation by binding it with high affinity and diverting it from a lysosomal pathway, returning it to the extracellular

compartment. In some embodiments, the serum albumin is selected from human serum albumin (HSA), cynomolgus monkeys serum albumin and mouse serum albumin. In some embodiments, the serum albumin provided herein is HSA.

[00128] The term "antibody" as used herein includes any immunoglobulin that binds to a specific antigen. Conventional antibody (e.g., antibodies from human or mice) comprises two heavy (H) chains and two light (L) chains. The heavy chains are classified as α, δ, ε, γ, and μ, each heavy chain consists of a variable domain (VH domain) and a first, second, third, and optionally fourth constant region (CH1, CH2, CH3, CH4 respectively); The light chains are classified as λ or κ, while each light chain consists of a variable domain (VL domain) and a constant domain. The antibody has a "Y" shape, with the stem of the Y consisting of the second and third constant domain of two heavy chains bound together via disulfide bonding. Each arm of the Y includes the variable domain and first constant domain of a single heavy chain bound to the variable and constant domains of a single light chain. The variable domains of the light and heavy chains are responsible for antigen binding. The variable domains in both chains generally contain three hypervariable regions called the complementarity determining regions (CDRs, namely CDR1, CDR2, and CDR3 of light chain or heavy chain). CDR boundaries for the antibodies and antigen-binding fragments disclosed herein may be defined or identified by the conventions of Kabat, IMGT, Chothia, or Al-Lazikani (Al-Lazikani, B., Chothia, C., Lesk, A. M., *J. Mol. Biol.*, 273(4), 927 (1997); Chothia, C. *et al.*, *J Mol Biol.* Dec 5;186(3):651-63 (1985); Chothia, C. and Lesk, A.M., *J. Mol. Biol.*, 196,901 (1987); Chothia, C. *et al.*, *Nature*. Dec 21-28;342(6252):877-83 (1989); Kabat E.A. *et al.*, Sequences of Proteins of immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, Md. (1991); Marie-Paule Lefranc *et al.*, *Developmental and Comparative Immunology*, 27: 55-77 (2003); Marie-Paule Lefranc *et al.*, *Immunome Research*, 1(3), (2005); Marie-Paule Lefranc, Molecular Biology of B cells (second edition), chapter 26, 481-514, (2015)). The three CDRs are interposed between flanking stretches known as framework regions (FRs, namely FR1, FR2, FR3 and FR4 of light chain or heavy chain), which are more highly conserved than the CDRs and form a scaffold to support the highly variable loops. The constant domains of the heavy and light chains are not involved in antigen-binding, but exhibit various effector functions. Conventional antibodies are assigned to classes based on the amino acid sequences of the constant domains of their heavy chains.

[00129] The single variable domain can be derived from the variable domain of a camelid antibody (VHH domain), or the variable domain of cartilaginous fish antibody (VNAR

domain). Both camelid antibodies and cartilaginous fish antibodies naturally lack light chains and consist of a pair of heavy chains. Alternatively, the single variable domain can be derived from the variable domain of a conventional antibody (e.g., from humans or mice) heavy chain (VH domain) or the variable domain of a common antibody light chain (VL domain). It is contemplated that a single domain antibody is fairly small in size, for example, has a molecular weight of no more than 25 kD, no more than 20 kD, no more than 15 kD, or no more than 10 kD.

[00130] It is to be noted that the term "nanobody" or "single-domain antibody" is used herein in its broadest sense and is not limited to a specific biological source or to a specific method of preparation. For example, a single domain antibody can be obtained, for example (1) by isolating the VHH domain or VNAR domain of a naturally occurring heavy chain antibody; (2) by expression of a nucleotide sequence encoding a naturally occurring VHH domain or VNAR domain; (3) by "humanization" (as described below) of a naturally occurring VHH domain or VNAR domain or by expression of a nucleic acid encoding a such humanized VHH domain or VNAR domain; (4) by "camelization" of a naturally occurring VH domain from any animal species, in particular a species of mammal, such as from a human being, or by expression of a nucleic acid encoding such a camelized VH domain; 5) using synthetic or semi-synthetic techniques for preparing proteins, polypeptides or other amino acid sequences; and/or (6) by any combination of the foregoing. Suitable methods and techniques for performing the foregoing will be clear to the skilled person.

[00131] In some embodiments, the single domain antibody described herein comprises a VHH domain derived from antibodies raised in Camelidae species, for example in camel, dromedary, alpaca and guanaco. A single domain antibody comprising a VHH domain are highly soluble and highly stable to heat, pH, proteases and other denaturing agents or conditions.

[00132] In some embodiments, the first polypeptide fragment comprises one or more single domain antibody that is capable of specifically binding to serum albumin.

[00133] The term "binding specificity", "specific binding" or "specifically binds" in reference to the interaction of a binding molecule, e.g., an antibody, and its binding partner, e.g., an antigen, means that the interaction is dependent upon the presence of a particular structure, e.g., an antigenic determinant or epitope, on the binding partner. In other words, the antibody preferentially binds or recognizes the binding partner even when the binding partner

is present in a mixture of other molecules or organisms. Antibodies or fragments thereof that immunospecifically bind to an antigen may be cross-reactive with related antigens, carrying the same epitope. The "binding specificity" is generally measured against nonspecific background binding. Typically, an antibody is considered specific when it binds to the target antigen at least 10 times above background binding.

[00134] The serum binding single domain antibody contemplated herein can bind to or otherwise associate with serum albumin in such a way that the binding of said serum albumin molecule to FcRn is not (significantly) reduced or inhibited (i.e. compared to the binding of said serum albumin molecule to FcRn when the single domain antibody is not bound thereto). In this aspect of the invention, by "not significantly reduced or inhibited" is meant that the binding affinity for serum albumin to FcRn (as measured using a suitable assay, such as SPR) is not reduced by more than 50%, preferably not reduced by more than 30%, even more preferably not reduced by more than 10%, such as not reduced by more than 5%, or essentially not reduced at all. In this aspect "not significantly reduced or inhibited" may also mean that the half-life of the serum albumin molecule is not significantly reduced (e.g., is not reduced by more than 50%, preferably not reduced by more than 30%, even more preferably not reduced by more than 10%, such as not reduced by more than 5%, or essentially not reduced at all, as measured using a suitable technique known per se). In some embodiments, the single domain antibody are capable of binding to amino acid residues on serum albumin that are not involved in binding of serum albumin to FcRn.

[00135] In some embodiments, the single domain antibody described herein binds to serum albumin selected from HSA, cynomolgus monkeys serum albumin and mouse serum albumin. In some embodiments, the binding affinity towards mouse serum albumin is about weaker than that towards human or cynomolgus serum albumin. In some embodiments, the single domain antibody specifically binds to HSA.

[00136] In some embodiments, the single domain antibody described herein binds to serum albumin with sufficient binding affinity. The term "affinity" as used herein refers to the strength of non-covalent interaction between an immunoglobulin molecule (i.e. antibody) or fragment thereof and an antigen. Affinity can be expressed numerically using "Kd" values. In general, a lower Kd value corresponds to a stronger binding. Kd may be determined by using any conventional method known in the art, including but are not limited to radioimmunoassays (RIA), enzyme-linked immunosorbent assays (ELISA), surface plasmon resonance (SPR) method, microscale thermophoresis method, HPLC-MS method and flow cytometry (such as

FACS) method. In some embodiment, the antibody disclosed here has a K_d value of $\leqslant 10^{-6}$ M (*e.g.* $\leqslant 5x10^{-7}$ M, $\leqslant 2x10^{-7}$ M, $\leqslant 10^{-7}$ M, $\leqslant 5x10^{-8}$ M, $\leqslant 2x10^{-8}$ M, $\leqslant 10^{-8}$ M, $\leqslant 5x10^{-9}$ M, $\leqslant 4x10^{-9}$M, $\leqslant 3x10^{-9}$M, $\leqslant 2x10^{-9}$ M, or $\leqslant 10^{-9}$ M) with the specific antigen.

[00137] In certain embodiments, the single domain antibody provided herein are capable of binding to HSA at a Kd value of 10^{-5}M to $1x10^{-12}$M or less, of 10^{-7}M to $1x10^{-12}$M or less, or of 10^{-8}M to $1x10^{-12}$M or less. In some embodiments, the Kd is no more than $1x10^{-7}$M (*e.g.* no more than $5x10^{-7}$ M, no more than $2x10^{-7}$ M, no more than 10^{-7} M, no more than $5x10^{-8}$ M, no more than $2x10^{-8}$ M, no more than 10^{-8} M, no more than $5x10^{-9}$ M, no more than $4x10^{-9}$M, no more than $3x10^{-9}$M, no more than $2x10^{-9}$ M, or no more than 10^{-9} M).

[00138] In some embodiments, the single domain antibody provided herein is a humanized antibody. The term "humanized" as used herein means that the single domain antibody comprises CDRs derived from a non-human animal, and FR regions derived from human. A humanized antibody polypeptide is desirable in its reduced immunogenicity in human. A humanized antibody polypeptide is chimeric in its variable regions, as non-human CDR sequences are grafted to human or substantially human FR sequences. Humanization of an antibody polypeptide can be essentially performed by substituting the non-human (such as camelid) CDR genes for the corresponding human CDR genes in a human immunoglobulin gene (see, for example, Jones et al. (1986) Nature 321:522-525; Riechmann et al. (1988) Nature 332:323-327; Verhoeyen et al. (1988) Science 239:1534-1536).

[00139] Various single domain antibodies that are capable of specifically binding to HSA with a high affinity are known in the art, such as fully human domain antibodies isolated using phase display, VHH antibodies developed form camelidae family, and VNAR antibodies developed from cartilaginous fish, see Zorzi, A et al, Med Chem Commun, 2019,10, 1068. Exemplary HSA-binding single domain antibodies are disclosed in US8188223B2, US9067991B2, US9321832B2, PCT application WO2008028977A2, WO2008043822A2, WO2020099871A1, G. Winter, et al, Annu. Rev. Immunol., 1994, 12, 433–455., L. J. Holt, et al., Protein Eng., Des. Sel., 2008, 21(5), 283–288, A. Walker, et al., Protein Eng., Des. Sel., 2010, 23(4),271–278, L. J. Goodall, et al.,PLoS One, 2015, 10(9), e0137065, R. L. O'Connor-Semmes, et al., Clin. Pharmacol. Ther., 2014, 96(6), 704–712, C. Read, et al., Basic Clin. Pharmacol.Toxicol., 2019, 1–8, R. Adams, et al., mAbs, 2016, 8(7),1336–1346, E. Dave, et al., mAbs, 2016, 8(7), 1319–1335, S. Steeland, et al., Drug Discovery Today, 2016, 21(7), 1076–1113, K. Coppieters, et al., Arthritis Rheum., 2006, 54(6), 1856–1866, M. Van Roy, , et

al., Arthritis Res. Ther., 2015, 17, 135, C. McMahon, et al., Nat. Struct. Mol. Biol., 2018, 25(3), 289–296, M. R. Muller, et al., mAbs, 2012, 4(6), 673–685, all of which are contemplated within the scope of the disclosure and incorporated by reference.

[00140] In some embodiments, the single domain antibody comprises a VHH domain. In some embodiments, the VHH domain is humanized.

[00141] In some embodiments, the VHH domain comprises a complementarity determining region 1 (CDR1), CDR2 and CDR3, wherein the CDR1 comprises the sequence of **SEQ ID NO: 20 (SFGMS)** or a variant thereof having up to 3, 2, or 1 amino acid mutation, a CDR2 comprising the sequence of **SEQ ID NO: 21 (SISGSGSDTLYADSVKG)** or a variant thereof having up to 3, 2, or 1 amino acid mutation, and/or a CDR3 comprising the sequence of **SEQ ID NO: 22 (GGSLSR)** or a variant thereof having up to 3, 2, or 1 amino acid mutation, wherein the VHH domain retains the binding specificity to serum albumin, optionally to human serum albumin.

[00142] In some embodiments, the VHH domain comprises a complementarity determining region 1 (CDR1) comprising the sequence of **SEQ ID NO: 20**, a CDR2 comprising the sequence of **SEQ ID NO: 21**, and a CDR3 comprising the sequence of **SEQ ID NO: 22.**

[00143] In some embodiments, the VHH domain comprises a complementarity determining region 1 (CDR1) consisting of the sequence of **SEQ ID NO: 20**, a CDR2 consisting of the sequence of **SEQ ID NO: 21**, and a CDR3 consisting of the sequence of **SEQ ID NO: 22.**

[00144] In some embodiments, the VHH domain comprises the sequence of **SEQ ID NO: 23**, or a variant thereof having at least 70% (e.g. at least 75%, 80%, 85%, 90%, 95%, 99% identity to **SEQ ID NO: 23**, wherein the variant retains the binding specificity and/or affinity to serum albumin.

[00145] In some embodiments, the variant of **SEQ ID NO: 23** has up to 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid mutation relative to **SEQ ID NO: 23**.

[00146] In certain embodiments of the polypeptide illustrated in FIG. 1A, the serum albumin binding nanobody domain further comprises an N-terminal extension attached to the VHH domain. In certain embodiments, the N-terminal extension comprises amino acid residues of SG, AG, S or A, and thus may comprise an amino acid sequence selected from SEQ ID NOs: 24-27.

[00147] In certain embodiments, the N-terminal extension comprises a tag, optionally a cleavable tag. Without wishing to be bound by any theory, it is believed that certain N-terminal extension can be useful for expression and post-translational processing.

[00148] In certain embodiments, the serum albumin-binding nanobody domain does not comprise an N-terminal extension attached to the VHH domain. For example, the VHH domain can be attached to a cleavable tag which, after cleavage, is no longer present in the final product.

[00149] In some embodiments, the first fragment may comprise one or more serum albumin binding nanobody domains.

[00150] **GLP-1 domain**

[00151] The second embodiments of the polypeptide disclosed herein and illustrated in FIG. 1B further comprises, in addition to the above described FGF21 domain in the second fragment 200 and the serum albumin binding nanobody domain in the first fragment 100, a third fragment 300 comprising an additional functional domain.

[00152] According to some embodiments, the additional functional domain comprises a biologically active GLP-1 domain.

[00153] The term "Glucagon-like peptide-1" or "GLP-1" as used herein is intended to broadly encompasses native GLP-1 peptide and all its functional forms such as its functional variants, fragments, fusions, derivatives and mimetics.

[00154] The term "native GLP-1 peptide" as used herein refers to the native human Glucagon-Like Peptide-1 (GLP-1 (7-37)), the sequence of which is set forth in SEQ ID NO: 28. The numbering of the residues in GLP-1 is referred to the sequence of SEQ ID NO: 28, which begins with the H residue at position 7 and ends with the G residue at position 37.

[00155] A functional form of the native GLP-1 peptide is capable of activating the GLP-1 receptor at a level comparable to, or no less than about 20% (or no less than 30%, 40%, 50%, 60%, 70%, 80%, 90%) of, that of the native GLP-1 peptide. Activation of the GLP-1 receptor typically initiates signal transduction pathway resulting in insulinotropic action or other physiological effects as is known in the art. Many functional forms of GLP-1 peptide are known in the art, for example, without limitation, liraglutide, semaglutide, dulaglutide, albiglutide, and those disclosed in WO2000055203A1, WO 98/08871, WO 2006/097537, WO2007139589A1, WO1998019698A1, WO2001098331A2, WO2003040309A2,

WO2005000892A2, WO2015000942A1, WO2016083499A1 the disclosure of which is incorporated herein to its entirety.

[00156] In certain embodiments, the GLP-1 domain provided herein comprises an amino acid sequence having at least 70% (e.g. at least 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%) sequence identity to SEQ ID NO: 28 while retaining substantial biological activity of SEQ ID NO: 28.

[00157] In certain embodiments, the GLP-1 domain comprises no more than 9, 8, 7, 6, 5, 4, 3, or 2 substitutions relative to SEQ ID NO: 28 while retaining substantial biological activity of SEQ ID NO: 28. In certain embodiments, the GLP-1 domain comprises at least 2, 3, 4, 5, 6, 7, 8, or 9 substitutions relative to SEQ ID NO: 28 while retaining substantial biological activity of SEQ ID NO: 28.

[00158] In certain embodiments, the GLP-1 domain further comprises one or more mutations. One of ordinary skill in the art will appreciate that various amino acid substitutions, e.g., conservative amino acid substitutions, may be made in the sequence of any of the polypeptide described herein, without necessarily decreasing its activity. Examples of amino acid substitutions include substituting an L-amino acid for its corresponding D-amino acid, substituting cysteine for homocysteine or other non-natural amino acids having a thiol-containing side chain, substituting a lysine for homolysine, diaminobutyric acid, diaminopropionic acid, ornithine or other non-national amino acids having an amino containing side chain, or substituting an alanine for norvaline or the like.

[00159] Various substitutions have been introduced to native GLP-1 peptide, and have been shown to be capable of retaining or even improving its biological activities. In certain embodiments, the GLP-1 comprises one or more mutations at a position selected from the group consisting of: A8, G22, K34, R36, and H7, and or any combination thereof, relative to SEQ ID NO: 28. For example, it is believed that substitution at A8 is useful to prevent DPP4 enzymatic cleavage at the residue, substitution at G22 is desirable to improve activity and solubility, and substitution at R36 is useful to reduce immunogenicity. In certain embodiments, the GLP-1 comprises one or more mutations at a position selected from the group consisting of: A8, G22, K26, K34, and R36, and or any combination thereof, relative to SEQ ID NO: 28. Examples of substitutions at these positions include, without limitation, A8G, A8S, A8V, A8Aib, A8T, A8I, A8L, G22E, K26R, K34R, R36G, or any combination thereof, as well as

the substitutions described in U.S. Patent Nos. 8,273,854, which are incorporated herein by its entirety. In certain embodiments, the one or more additional substitutions comprises a conservative substitution.

[00160] In certain embodiments, the GLP-1 domain comprises a substitution of A8 which is selected from the group consisting of: A8G, A8S, A8V, A8Aib, A8T, and A8L. In certain embodiments, the GLP-1 domain comprises a substitution of G22E. In certain embodiments, the GLP-1 domain comprises n substitution of R36G. In certain embodiments, the GLP-1 domain comprises a substitution of K26 which is K26R. In certain embodiments, the GLP-1 domain comprises a substitution of K34 which is K34R.

[00161] In certain embodiments, the GLP-1 comprises or consists of one or more substitutions selected from the group consisting of A8G, K26R, K34R, G22E, and R36G. In certain embodiments, the GLP-1 comprises or consists of one or more substitutions selected from the group consisting of A8G, G22E, and R36G.

[00162] In the present disclosure, at least the following embodiments of the biologically active GLP-1 domain in the third fragment 300 in FIG. 1B are provided:

[00163] (1) SEQ ID NO: 28, representing a wildtype functional form of GLP-1;

[00164] (2) SEQ ID NO: 29, representing a three-substitution functional form of GLP-1 comprising substitutions A8G, G22E, and R36G;

[00165] (3) SEQ ID NO: 31, representing a four-substitution functional form of GLP-1 comprising substitutions A8G, G22E, K34R, and R36G;

[00166] (4) SEQ ID NO: 32, representing a five-substitution functional form of GLP-1 comprising substitutions A8G, G22E, K26R, K34R, and R36G;

[00167] (4) SEQ ID NO: 33, representing a five-substitution functional form of GLP-1 comprising substitutions A8G, G22E, K26R, K34R, and R36K; and

[00168] (5) SEQ ID NO: 34, representing a four-substitution functional form of GLP-1 comprising substitutions A8G, G22E, K26R, and R36G.

[00169] It is to be noted that in addition to GLP-1, the additional functional domain of the third fragment 300 as illustrated in FIG. 1B may alternatively comprise a biologically active form (i.e. functional form) of insulin, C-peptide, leptin, glucagon, gastrin, gastric inhibitory polypeptide (GIP), amylin, calcitonin, cholecystokinin, peptide YY, neuropeptide Y,

bone morphogenetic protein-6 (BMP-6), bone morphogenetic protein-9 (BMP-9), oxyntomodulin, oxytocin, glucagon-like peptide-2 (GLP-2), irisin, fibronectin type III domain containing protein 5 (FNDC5), apelin, adiponectin, Clq and tumor necrosis factor related protein (CTRP family), resistin, visfatin, omentin, retinol binding protein-4 (RBP-4), glicentin, angiopoietin, interleukin-22 (IL-22), exendin-4, or growth hormone.

[00170] **The First and/or Second Linker**

[00171] Each of the first and second embodiments of the polypeptide disclosed herein and illustrated in FIGS. 1A and 1B utilizes the first linker 10 to connect the first fragment 100 (comprising the nanobody domain capable of binding to serum albumin) and the second fragment 200 (comprising the biologically active FGF21 domain), and the second embodiments of the polypeptide illustrated in FIG. 1B utilizes the second linker 20 to connect the first fragment 100 and the third fragment 300 (comprising a biologically active additional functional domain that has additive or synergistic effect on FGF21 domain; e.g. GLP-1 domain).

[00172] The term "linker" as used herein, such as the first linker 10 and/or the second linker 20, refers generally to a "polypeptide linker", which can be any suitable polypeptide capable of bonding two entities to thereby form one molecule, or maintaining association of the two entities in sufficiently close proximity, yet without substantially interference to the respective biological activities of the two entities.

[00173] The linker can be integrated in the resulting linked molecule or structure. Herein, the first linker 10 operably separates the nanobody domain 100 and the FGF21 domain 200 without substantial interference to the respective biological activities of the two functional domains, and the second linker 20 operably separates the GLP-1 domain and the nanobody domain 100 without substantial interference to the respective biological activities of the two functional domains. The linker can be made up of amino acid residues linked together by peptide bonds, yet may optionally further comprise one or more non-natural amino acids.

[00174] Generally, each of the first linker 10 and the second linker 20 has a length of at least four amino acid residues. As such, in certain embodiments, each linker has a length of at least 4, 8, 10, 20, 24, 28, 30, 40, 48, 50, 60, 70, 80, 90, 100, 110, 120 or more amino acid resides. Without wishing to be bound by any theory, it is believed that a suitable length of a linker can further improve the biological activity, stability, or pharmacokinetic parameters of the each of the two functional domains such linked thereby in the whole polypeptide molecule.

[00175] Any suitable polypeptide can be used as a linker. For example, the polypeptide linker may comprise or consist of amino acid residues selected from the amino acids glycine(G), serine(S), alanine(A), methionine(M), asparagine(N), glutamine(Q), cysteine(C), Proline (P), Glutamate (E), Threonine (T) and lysine (K). In some embodiments, the polypeptide linker can be made up of a majority of amino acids that are sterically unhindered, such as glycine and alanine. In some embodiments, linkers are polyglycines, polyalanines, combinations of glycine and alanine (such as poly(Gly-Ala)), or combinations of glycine and serine (such as poly (Gly-Ser)).

[00176] In certain embodiments, each of the first linker and the second linker comprises or consists of one or more repeats of a repeating sequence. In certain embodiments, the polypeptide linker comprises or consists of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 repeats of a repeating sequence, or within any numerical range defined by any two numbers listed above.

[00177] Herein, based on the experimental data shown in the Examples below, the following configurations are provided which are nonetheless optional and non-limiting in scope.

[00178] Regarding the first linker, it may comprise no acidic amino acid residue (e.g. D or E) according to some embodiments. The first linker may optionally comprise one or more units of a first repeating sequence, and the first repeating sequence may consist of no more than 4 or 6 types of amino acid residues, which can be selected from the group consisting of: G, Q, A, P, T and S. According to some embodiments of the polypeptide, the first repeating sequence comprises or consists of an amino acid sequences selected from a group consisting of G_fS_g (each of f and g is independently an integer selected from 1 to 5), SEQ ID NO: 35 (GAQP), SEQ ID NO: 36 (GQAP), SEQ ID NO: 37 (GPAQ), SEQ ID NO: 38 (GPQA), SEQ ID NO: 39 (GSQP), SEQ ID NO: 40 (GASP), SEQ ID NO: 41 (GPAS), SEQ ID NO: 42 (GPSA), SEQ ID NO: 43 (GGGS), SEQ ID NO: 44 (GSGS), SEQ ID NO: 45 (GGGGS), SEQ ID NO: 46 (GSAPGSPAGSPTGSAPGSPA) and GS. In certain embodiments of the polypeptide, the first repeating sequence has an amino acid sequence set forth in SEQ ID NO: 35 (GAQP), and a number of the one or more units is an integer between 1 and 10. In certain embodiments, the first linker comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 35 (GAQP), SEQ ID NO: 49 $((GAQP)_2)$, SEQ ID NO: 50 $((GAQP)_5)$, SEQ ID NO: 51 $((GAQP)_{10})$, and SEQ ID NO: 48 (GGGGSGGGS).

[00179] Regarding the second linker, it may similarly comprise one or more units of a second repeating sequence, and the second repeating sequence may consist of no more than 4 or 6 types of amino acid residues selected from the group consisting of: G, Q, A, E, P, T and S. In certain embodiments of the polypeptide, the second repeating sequence comprises or consists of an amino acid sequences selected from a group consisting of G_hS_i (each of h and i is independently an integer selected from 1 to 5), SEQ ID NO: 35 (GAQP), SEQ ID NO:55 (GQEP), SEQ ID NO: 56 (GEQP), SEQ ID NO: 57 (GPQE), SEQ ID NO: 58 (GPEQ), SEQ ID NO: 59 (GSEP), SEQ ID NO: 60 (GESP), SEQ ID NO: 61 (GPSE), SEQ ID NO: 62 (GPES), SEQ ID NO: 36 (GQAP), SEQ ID NO: 37 (GPAQ), SEQ ID NO: 38 (GPQA), SEQ ID NO: 39 (GSQP), SEQ ID NO: 40 (GASP), SEQ ID NO: 41 (GPAS), SEQ ID NO: 42 (GPSA), SEQ ID NO: 43 (GGGS), SEQ ID NO: 44 (GSGS), SEQ ID NO: 45 (GGGGS), SEQ ID NO: 46 (GSAPGSPAGSPTGSAPGSPA) and GS. In certain embodiments, the second repeating sequence has an amino acid sequence set forth in SEQ ID NO: 35 (GAQP), and a number of the one or more units is an integer between 1 and 15, optionally 1, 2, 5, 10, or 14. In certain embodiments, the second linker comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 49 $((GAQP)_2)$, SEQ ID NO: 50 $((GAQP)_5)$, SEQ ID NO: 51 $((GAQP)_{10})$, and SEQ ID NO: 52 $((GAQP)_{14})$, and SEQ ID NO: 47 $((GGGGS)_4)$.

[00180] In certain embodiments, each of the first linker and the second linker may comprise or consist of more than one repeating sequence. For example, one such linker may comprise or consist of 2, 3, or 4 different repeating sequences. In certain embodiments, one such linker may comprise or consist of sequential or tandem repeats of the different repeating sequences. The number of repeats of each repeating sequence more be independently selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more.

[00181] **Other Notes**

[00182] The following is noted for each of the FGF21 domain, the nanobody domain, the GLP-1 domain, and the first/second linker as described above. One of ordinary skill in the art will appreciate that various amino acid substitutions, e.g., conservative amino acid substitutions, may be made in the sequence of any of the polypeptide fragment described herein, without necessarily decreasing its activity. Examples of amino acid substitutions include substituting an L-amino acid for its corresponding D-amino acid, substituting cysteine for homocysteine or other non-natural amino acids having a thiol-containing side chain, substituting a lysine for homolysine, diaminobutyric acid, diaminopropionic acid, ornithine or

other non-national amino acids having an amino containing side chain, or substituting an alanine for norvaline or the like.

[00183] A "conservative substitution" with reference to amino acid sequence refers to replacing an amino acid residue with a different amino acid residue having a side chain with similar physiochemical properties. For example, conservative substitutions can be made among amino acid residues with hydrophobic side chains (e.g. Met, Ala, Val, Leu, and Ile), among residues with neutral hydrophilic side chains (e.g. Cys, Ser, Thr, Asn and Gln), among residues with acidic side chains (e.g. Asp, Glu), among amino acids with basic side chains (e.g. His, Lys, and Arg), or among residues with aromatic side chains (e.g. Trp, Tyr, and Phe). As known in the art, conservative substitution usually does not cause significant change in the protein conformational structure, and therefore could retain the biological activity of a protein.

[00184] The following full-length sequences are provided for certain embodiments of the first embodiments (illustrated in FIG. 1A) and of the second embodiments (illustrated in FIG. 1B).

[00185] Certain embodiments of the polypeptide is substantially a fusion polypeptide between a serum albumin-binding nanobody domain and an FGF21 domain in an N-terminus-to-C-terminus direction, and may have an amino acid sequence that is selected from SEQ ID NOS: 63-68, 93, 99-101 and 107. In certain of these embodiments, the polypeptide is conjugated to a synthetic chemical moiety provided herein (see, e.g. based on FIG. 1A). The synthetic chemical moiety, for example, can be Ac-2XADO-EDA-CO-CH2-*, which is conjugated with an introduced cysteine residue at position 171 or position 174 relative to SEQ ID NO: 1 for a polypeptide with a sequence selected from the group consisting of SEQ ID NOS: 63, 68, 93 and 100. In certain of the alternative embodiments, the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 64-67, 69, 99 and 108, and is not conjugated.

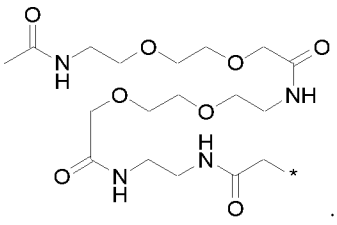
[00186] Table 1A below shows the SEQ ID NOs of the exemplary fusion polypeptide sequences, and the SEQ ID NOs for the nanobody, the FGF21, and the first polypeptide linker and the second polypeptide linker contained in the fusion polypeptides.

[00187] Table 1A. Exemplary fusion polypeptide sequences

Molecule Code. (MLC)	Fusion Polypeptide	Nanobody	First Linker	FGF21 (Mutations#)	Conjugation site; functional moiety
MLC#1	SEQ ID NO:63	SEQ ID NO: 27	(GAQP)10 (SEQ ID NO: 51)	121Q,168L,171C,180E (SEQ ID NO: 8)	171C; Moiety A##
MLC#2	SEQ ID NO:64	SEQ ID NO: 27	(GAQP)10 (SEQ ID NO: 51)	121Q,168L,171G,180E (SEQ ID NO:14)	No
MLC#3	SEQ ID NO:65	SEQ ID NO: 27	(GAQP)5 (SEQ ID NO: 50)	121Q,168L,171G,180E (SEQ ID NO:14)	No
MLC#4	SEQ ID NO:66	SEQ ID NO: 27	(GAQP)2 (SEQ ID NO: 49)	121Q,168L,171G,180E (SEQ ID NO:14)	No
MLC#5	SEQ ID NO:67	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171G,180E (SEQ ID NO:14)	No
MLC#6	SEQ ID NO:68	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171C,180E (SEQ ID NO: 8)	171C; Moiety A##
MLC#7	SEQ ID NO:93	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171G,174C, 180E (SEQ ID NO: 92)	174C; Moiety A##
MLC#22	SEQ ID NO: 99	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,180E (SEQ ID NO: 5)	No
MLC#23	SEQ ID NO: 100	SEQ ID NO: 27	G4S (SEQ ID NO: 45)	121Q,168L,171C,180E (SEQ ID NO: 8)	171C; Moiety A##
MLC#25	SEQ ID NO: 101	SEQ ID NO: 27	G4S (SEQ ID NO: 45)	121Q,168L,180E (SEQ ID NO: 5)	No
MLC#26	SEQ ID NO: 107	SEQ ID NO: 27	G4S (SEQ ID NO: 45)	121Q,168L,171G,180E (SEQ ID NO:14)	No

#: the mutation positions are relative to SEQ ID NO: 1.

Moiety A refers to Ac-2XADO-EDA-CO-CH_2*, where the * end is connected to the conjugatable cysteine residue on the polypeptide, having a structure of:



.

[00188] Certain embodiments of the polypeptide is substantially a fusion polypeptide among a GLP-1 domain, a serum albumin-binding nanobody domain, and an FGF21 domain in an N-terminus-to-C-terminus direction, which may have an amino acid sequence that is selected from SEQ ID NOS: 70, 74, 75, 79-83, 85, 94-98 and 108-109. In certain of these embodiments, the polypeptide is conjugated to a synthetic chemical moiety as provided herein (see, e.g., based on FIG. 1B). The synthetic chemical moiety, for example, can be Ac-2XADO-EDA-CO-CH2-*, which is conjugated with an introduced cysteine residue at position 171 or position 174 relative to SEQ ID NO: 1 for the amino acid sequences selected from the group consisting of SEQ ID NOs: 74, 81-83, 85, 94, 96 and 97. In certain of the alternative embodiments, the polypeptide comprises an amino acid sequence selected from SEQ ID NOs: 70, 75, 79, 80, 95, 98, and 108-109, and is not conjugated.

[00189] Table 1B below shows the SEQ ID NOs of the exemplary fusion polypeptide sequences, and the SEQ ID NOs for the GLP-1 domain, the first polypeptide linker, nanobody, the FGF21 domain, and the first polypeptide linker and the second polypeptide linker contained in the fusion polypeptides.

[00190] Table 1B. Exemplary fusion polypeptide sequences

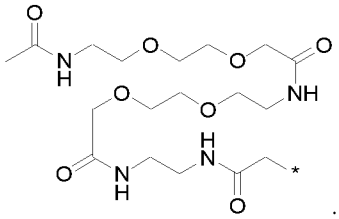
Molecule Code. (MLC); Fusion Polypeptide (SEQ ID NO)	GLP1 domain (Mutations&)	Second linker	Nanobody	First Linker	FGF21 domain (Mutations#)	Conjugation site; functional moiety
MLC#8 (SEQ ID NO:70)	8G, 22E, 36G (SEQ ID NO:29)	(G4S)4 (SEQ ID NO: 47)	SEQ ID NO: 27	G4SG3S (SEQ ID NO: 48)	121Q,168L,171G, 180E (SEQ ID NO:14)	No
MLC#9 (SEQ ID NO:74)	8G, 22E, 36G (SEQ ID NO:29)	(GAQP)5 (SEQ ID NO: 50)	SEQ ID NO: 27	(GAQP)10 (SEQ ID NO: 51)	121Q,168L,171C,180E (SEQ ID NO: 8)	171C; Moiety A##
MLC#10 (SEQ ID NO:75)	8G, 22E, 36G (SEQ ID NO:29)	(GAQP)5 (SEQ ID NO: 50)	SEQ ID NO: 27	(GAQP)10 (SEQ ID NO: 51)	121Q,168L,171G, 180E (SEQ ID NO:14)	No
MLC#11 (SEQ ID NO:79)	8G, 22E, 36G (SEQ ID NO:29)	(GAQP)10 (SEQ ID NO: 51)	SEQ ID NO: 27	(GAQP)10 (SEQ ID NO: 51)	121Q,168L,171G, 180E (SEQ ID NO:14)	No
MLC#12 (SEQ ID NO:80)	8G, 22E, 36G (SEQ ID NO:29)	(GAQP)5 (SEQ ID NO: 50)	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171G, 180E (SEQ ID NO:14)	No
MLC#13 (SEQ ID NO:81)	8G, 22E, 36G (SEQ ID NO:29)	(GAQP)5 (SEQ ID NO: 50)	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171C,180E(SEQ ID NO: 8)	171C; Moiety A##
MLC#14 (SEQ ID NO:82)	8G, 22E, 36G (SEQ ID NO:29)	(GAQP)10 (SEQ ID NO: 51)	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171C,180E (SEQ ID NO: 8)	171C; Moiety A##
MLC#15 (SEQ ID NO:83)	8G, 22E, 36G (SEQ ID NO:29)	(GAQP)2 (SEQ ID NO: 49)	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171C,180E (SEQ ID NO: 8)	171C; Moiety A##
MLC#16 (SEQ ID NO:85)	8G, 22E, 36G (SEQ ID NO:29)	(GAQP)14 (SEQ ID NO: 52)	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171C,180E (SEQ ID NO: 8)	171C; Moiety A##
MLC#17 (SEQ ID NO:94)	8G, 22E, 36G (SEQ ID NO:29)	(GAQP)14 (SEQ ID NO: 52)	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171G, 174C,180E (SEQ ID NO: 92)	174C; Moiety A##
MLC#18 (SEQ ID NO: 95)	8G,22E, 36G (SEQ ID NO:29)	(GAQP)14 (SEQ ID NO: 52)	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,180E (SEQ ID NO: 5)	No
MLC#19 (SEQ ID NO: 96)	8G,22E, 36G (SEQ ID NO:29)	(GAQP)14 (SEQ ID NO: 52)	SEQ ID NO: 27	(G4S) (SEQ ID NO: 45)	121Q,168L,171C,180E (SEQ ID NO: 8)	171C; Moiety A##
MLC#20 (SEQ ID NO: 97)	8G,22E, 36G (SEQ ID NO:29)	(GAQP)13(G4S) (SEQ ID NO: 106)	SEQ ID NO: 27	(G4S) (SEQ ID NO: 45)	121Q,168L,171C,180E (SEQ ID NO: 8)	171C; Moiety A##
MLC#21 (SEQ ID NO: 98)	8G,22E, 36G (SEQ ID NO:29)	(GAQP)14 (SEQ ID NO: 52)	SEQ ID NO: 27	(G4S) (SEQ ID NO: 45)	121Q,168L,180E (SEQ ID NO: 5)	No

MLC#27 (SEQ ID NO: 108)	8G,22E, 36G (SEQ ID NO:29)	(GAQP)14 (SEQ ID NO: 52)	SEQ ID NO: 27	(GAQP) (SEQ ID NO: 35)	121Q,168L,171G, 180E (SEQ ID NO:14)	No
MLC#28 (SEQ ID NO: 109)	8G,22E, 36G (SEQ ID NO:29)	(GAQP)14 (SEQ ID NO: 52)	SEQ ID NO: 27	(G4S) (SEQ ID NO: 45)	121Q,168L,171G, 180E (SEQ ID NO:14)	No

&: the mutation positions are relative to SEQ ID NO: 28.

#: the mutation positions are relative to SEQ ID NO: 1.

Moiety A refers to Ac-2XADO-EDA-CO-CH_2*, where the * end is connected to the conjugatable cysteine residue on the polypeptide, having a structure of:



[00191] **Pharmaceutical Composition**

[00192] In another aspect, the present disclosure further provides a pharmaceutical composition, which comprises the polypeptide according to any of the embodiments as described above, and a pharmaceutically acceptable carrier.

[00193] The term "pharmaceutically acceptable" indicates that the designated carrier, vehicle, diluent, excipient(s), and/or salt is generally chemically and/or physically compatible with the other ingredients comprising the formulation, and physiologically compatible with the recipient thereof.

[00194] A "pharmaceutically acceptable carrier" refers to an ingredient in a pharmaceutical formulation, other than an active ingredient, which is bioactivity acceptable and nontoxic to a subject. Pharmaceutical acceptable carriers for use in the pharmaceutical compositions disclosed herein may include, for example, pharmaceutically acceptable liquid, gel, or solid carriers, aqueous vehicles, nonaqueous vehicles, antimicrobial agents, isotonic agents, buffers, antioxidants, anesthetics, suspending/dispending agents, sequestering or chelating agents, diluents, adjuvants, excipients, or non-toxic auxiliary substances, other components known in the art, or various combinations thereof.

[00195] Suitable components may include, for example, antioxidants, fillers, binders, disintegrants, buffers, preservatives, lubricants, flavorings, thickeners, coloring agents,

emulsifiers or stabilizers such as sugars and cyclodextrins. Suitable antioxidants may include, for example, methionine, ascorbic acid, EDTA, sodium thiosulfate, platinum, catalase, citric acid, cysteine, thioglycerol, thioglycolic acid, thiosorbitol, butylated hydroxanisol, butylated hydroxytoluene, and/or propyl gallate. As disclosed herein, inclusion of one or more antioxidants such as methionine in a pharmaceutical composition provided herein decreases oxidation of the polypeptide complex or the bispecific polypeptide complex. This reduction in oxidation prevents or reduces loss of binding affinity, thereby improving protein stability and maximizing shelf-life. Therefore, in certain embodiments, compositions are provided that comprise the polypeptide, the polypeptide complex or the conjugate disclosed herein and one or more antioxidants such as methionine.

[00196] To further illustrate, pharmaceutical acceptable carriers may include, for example, aqueous vehicles such as sodium chloride injection, Ringer's injection, isotonic dextrose injection, sterile water injection, or dextrose and lactated Ringer's injection, nonaqueous vehicles such as fixed oils of vegetable origin, cottonseed oil, corn oil, sesame oil, or peanut oil, antimicrobial agents at bacteriostatic or fungistatic concentrations, isotonic agents such as sodium chloride or dextrose, buffers such as phosphate or citrate buffers, antioxidants such as sodium bisulfate, local anesthetics such as procaine hydrochloride, suspending and dispersing agents such as sodium carboxymethylcelluose, hydroxypropyl methylcellulose, or polyvinylpyrrolidone, emulsifying agents such as Polysorbate 80 (TWEEN-80), sequestering or chelating agents such as EDTA (ethylenediaminetetraacetic acid) or EGTA (ethylene glycol tetraacetic acid), ethyl alcohol, polyethylene glycol, propylene glycol, sodium hydroxide, hydrochloric acid, citric acid, or lactic acid. Antimicrobial agents utilized as carriers may be added to pharmaceutical compositions in multiple-dose containers that include phenols or cresols, mercurials, benzyl alcohol, chlorobutanol, methyl and propyl p-hydroxybenzoic acid esters, thimerosal, benzalkonium chloride and benzethonium chloride. Suitable excipients may include, for example, water, saline, dextrose, glycerol, or ethanol. Suitable non-toxic auxiliary substances may include, for example, wetting or emulsifying agents, pH buffering agents, stabilizers, solubility enhancers, or agents such as sodium acetate, sorbitan monolaurate, triethanolamine oleate, or cyclodextrin.

[00197] The pharmaceutical compositions can be a liquid solution, suspension, emulsion, pill, capsule, tablet, sustained release formulation, or powder. Oral formulations can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch,

magnesium stearate, polyvinyl pyrollidone, sodium saccharine, cellulose, magnesium carbonate, etc.

[00198] In embodiments, the pharmaceutical compositions are formulated into an injectable composition. The injectable pharmaceutical compositions may be prepared in any conventional form, such as for example liquid solution, suspension, emulsion, or solid forms suitable for generating liquid solution, suspension, or emulsion. Preparations for injection may include sterile and/or non-pyretic solutions ready for injection, sterile dry soluble products, such as lyophilized powders, ready to be combined with a solvent just prior to use, including hypodermic tablets, sterile suspensions ready for injection, sterile dry insoluble products ready to be combined with a vehicle just prior to use, and sterile and/or non-pyretic emulsions. The solutions may be either aqueous or nonaqueous.

[00199] In certain embodiments, unit-dose parenteral preparations are packaged in an ampoule, a vial or a syringe with a needle. All preparations for parenteral administration should be sterile and not pyretic, as is known and practiced in the art.

[00200] In certain embodiments, a sterile, lyophilized powder is prepared by dissolving the polypeptide, the polypeptide complex or the conjugate as disclosed herein in a suitable solvent. The solvent may contain an excipient which improves the stability or other pharmacological components of the powder or reconstituted solution, prepared from the powder. Excipients that may be used include, but are not limited to, water, dextrose, sorbital, fructose, corn syrup, xylitol, glycerin, glucose, sucrose or other suitable agent. The solvent may contain a buffer, such as citrate, sodium or potassium phosphate or other such buffer known to those of skill in the art at, in one embodiment, about neutral pH. Subsequent sterile filtration of the solution followed by lyophilization under standard conditions known to those of skill in the art provides a desirable formulation. In one embodiment, the resulting solution will be apportioned into vials for lyophilization. Each vial can contain a single dosage or multiple dosages of the polypeptide, the polypeptide complex, or the conjugate provided herein or composition thereof. Overfilling vials with a small amount above that needed for a dose or set of doses (e.g., about 10%) is acceptable so as to facilitate accurate sample withdrawal and accurate dosing. The lyophilized powder can be stored under appropriate conditions, such as at about 4 °C to room temperature.

[00201] Reconstitution of a lyophilized powder with water for injection provides a formulation for use in parenteral administration. In one embodiment, for reconstitution the

sterile and/or non-pyretic water or other liquid suitable carrier is added to lyophilized powder. The precise amount depends upon the selected therapy being given, and can be empirically determined.

[00202] Administration of the pharmaceutical composition as described herein may be via any route known to be effective by the physician of ordinary skill. One example is peripheral parenteral administration by a sterile syringe or some other mechanical device such as an infusion pump. In certain embodiments, peripheral parenteral route is intravenous, intramuscular, subcutaneous, or intraperitoneal routes of administration.

[00203] In certain embodiments, the polypeptide, the polypeptide complex, or the conjugate described herein is formulated in a form suitable for non-parenteral routes administration, such as oral, rectal, nasal, or lower respiratory routes administration.

[00204] In certain embodiments, the polypeptide, the polypeptide complex, or the conjugate described herein is formulated in a solid formulation such as lyophilization or spray drying, which is then reconstituted in a suitable diluent solution prior to administration. Standard pharmaceutical formulation techniques, such as those described in Remington: The Science and Practice of Pharmacy (D.B. Troy, Editor, 21st Edition, Lippincott, Williams & Wilkins, 2006), may be employed. Alternatively, the polypeptide, the polypeptide complex, or the conjugate described herein can be formulated for administration through the lingual, sublingual, buccal, in the mouth, oral, in the stomach and intestine, nasal, pulmonary, for example, through the bronchioles and alveoli or a combination thereof, epidermal, dermal, transdermal, vaginal, rectal, ocular, for examples through the conjunctiva, uretal, transdermal, or pulmonary route. As a still further option, the polypeptide, the polypeptide complex, or the conjugate described herein can be formulated for administration through transdermal administration, for example, by needle-free injection or from a patch, optionally an iontophoretic patch, or transmucosal, for example, buccal, administration.

[00205] **Methods of Treatment**

[00206] In yet another aspect, the present disclosure further provides a method and a kit of preventing or treating a metabolic disorder in a subject in need thereof.

[00207] The method substantially comprises: administering a therapeutically effective amount of the polypeptide or the pharmaceutical composition as described above to the subject.

[00208] "Treating" or "treatment" of a condition as used herein includes preventing or alleviating a condition, slowing the onset or rate of development of a condition, reducing the risk of developing a condition, preventing or delaying the development of symptoms associated with a condition, reducing or ending symptoms associated with a condition, generating a complete or partial regression of a condition, curing a condition, or some combination thereof.

[00209] The term "subject" or "individual" or "animal" or "patient" as used herein refers to human or non-human animal, including a mammal or a primate, in need of diagnosis, prognosis, amelioration, prevention and/or treatment of a disease or disorder. Mammalian subjects include humans, domestic animals, farm animals, and zoo, sports, or pet animals such as dogs, cats, guinea pigs, rabbits, rats, mice, horses, swine, cows, bears, and so on.

[00210] In certain embodiments, the subject has been identified as having a disorder or condition likely to respond to the polypeptide or the pharmaceutical composition provided herein.

[00211] In certain embodiments, the metabolic disorder is diabetes, obesity, non-alcoholic steatohepatitis (NASH), cardiovascular like dyslipidaemia, artheroscelerosis, alcoholic steatohepatitis (ASH), diabeticnephropathy, gestational diabetes, metabolic syndrome such as metabolic syndrome X, nonalcoholic fatty liver disease (NAFLD), end-stage liver disease, hepatic steatosis (fatty liver), liver cirrhosis, primary biliary cirrhosis (PBC), or severe hypertriglyceridemia (SHTG).

[00212] For example, a metabolic condition or disorder that can be treated or ameliorated using the polypeptide or the pharmaceutical composition provided herein, includes a condition where a human subject has a fasting blood glucose level of 125 mg/dL or greater, for example 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200 or greater than 200 mg/dL. Blood glucose levels can be determined in the fed or fasted state, or at random. The metabolic condition or disorder can also comprise a condition in which a subject is at increased risk of developing a metabolic condition. For a human subject, such conditions include a fasting blood glucose level of 100 mg/dL.

[00213] The therapeutically effective amount of the polypeptide or the pharmaceutical composition provided herein will depend on various factors known in the art, such as for example body weight, age, past medical history, present medications, state of health of the subject and potential for cross-reaction, allergies, sensitivities and adverse side-effects,

as well as the administration route and extent of disease development. Dosages may be proportionally reduced or increased by one of ordinary skill in the art (e.g., physician or veterinarian) as indicated by these and other circumstances or requirements. The therapeutically effective amount can be an amount of the polypeptide conjugate or the pharmaceutical composition provided herein, that elicits a biological or medicinal response in a tissue system, animal, or human being sought by a researcher, medical doctor, or other clinician, which includes alleviation or amelioration of the symptoms of the disease or disorder being treated, i.e., an amount that supports an observable level of one or more desired biological or medicinal response, for example lowering blood glucose, insulin, triglyceride, or cholesterol levels; reducing body weight; or improving glucose tolerance, energy expenditure, or insulin sensitivity.

[00214] In certain embodiments, the polypeptide or the pharmaceutical composition provided herein may be administered at a therapeutically effective dosage of about 0.01 mg/kg to about 100 mg/kg (e.g., about 0.01 mg/kg, about 0.5 mg/kg, about 1 mg/kg, about 2 mg/kg, about 5 mg/kg, about 10 mg/kg, about 15 mg/kg, about 20 mg/kg, about 25 mg/kg, about 30 mg/kg, about 35 mg/kg, about 40 mg/kg, about 45 mg/kg, about 50 mg/kg, about 55 mg/kg, about 60 mg/kg, about 65 mg/kg, about 70 mg/kg, about 75 mg/kg, about 80 mg/kg, about 85 mg/kg, about 90 mg/kg, about 95 mg/kg, or about 100 mg/kg). In certain of these embodiments, the polypeptide or the pharmaceutical composition provided herein is administered at a dosage of about 50 mg/kg or less, and in certain of these embodiments the dosage is 10 mg/kg or less, 5 mg/kg or less, 1 mg/kg or less, 0.5 mg/kg or less, or 0.1 mg/kg or less. In certain embodiments, the administration dosage may change over the course of treatment. For example, in certain embodiments the initial administration dosage may be higher than subsequent administration dosages. In certain embodiments, the administration dosage may vary over the course of treatment depending on the reaction of the subject.

[00215] Dosage regimens may be adjusted to provide the optimum desired response (e.g., a therapeutic response). For example, a single dose may be administered, or several divided doses may be administered over time.

[00216] The polypeptide or the pharmaceutical composition provided herein may be administered by any route known in the art, such as for example parenteral (e.g., subcutaneous, intraperitoneal, intravenous, including intravenous infusion, intramuscular, or intradermal injection) or non-parenteral (e.g., oral, intranasal, intraocular, sublingual, rectal, or topical) routes.

[00217] The polypeptide or the pharmaceutical composition may be administered alone or in combination with one or more additional therapeutic means or agents.

[00218] In certain embodiments, when used for treating a metabolic disease, the polypeptide or the pharmaceutical composition provided herein may be administered in combination with any other therapeutic agent for use in the treatment of a metabolic disease or any medical disorder that related. "Administered in combination" as used herein includes administration simultaneously as part of the same pharmaceutical composition, simultaneously as separate compositions, or at different timings as separate compositions. A composition administered prior to or after another agent is considered to be administered "in combination" with that agent as the phrase is used herein, even if the composition and the second agent are administered via different routes. Where possible, additional therapeutic agents administered in combination with the fusion polypeptide, the polypeptide complex or the conjugate provided herein are administered according to the schedule listed in the product information sheet of the additional therapeutic agent, or according to the Physicians' Desk Reference (Physicians' Desk Reference, 70th Ed (2016)) or protocols well known in the art.

[00219] **Kits**

[00220] A kit for practicing the methods for administering the pharmaceutical composition as described above is also provided. Such kits can comprise a pharmaceutical composition such as those described herein, which can be provided in a sterile container. Optionally, instructions on how to employ the provided pharmaceutical composition in the treatment of a metabolic disorder can also be included or be made available to a patient or a medical service provider.

[00221] Such a kit may comprise: (a) a pharmaceutical composition comprising a therapeutically effective amount of a fusion polypeptide conjugate; and (b) one or more containers for the pharmaceutical composition. Such a kit can also comprise instructions for the use thereof; the instructions can be tailored to the precise metabolic disorder being treated. The instructions can describe the use and nature of the materials provided in the kit. In certain embodiments, kits include instructions for a patient to carry out administration to treat a metabolic disorder, such as elevated glucose levels, elevated insulin levels, diabetes, obesity, non-alcoholic steatohepatitis (NASH), cardiovascular like dyslipidaemia, artherosclerosis, alcoholic steatohepatitis (ASH), diabeticnephropathy, metabolic syndrome such as metabolic syndrome X, nonalcoholic fatty liver disease (NAFLD), end-stage liver disease, hepatic

steatosis (fatty liver), liver cirrhosis, primary biliary cirrhosis (PBC) or severe hypertriglyceridemia (SHTG).

[00222] Instructions can be printed on a substrate, such as paper or plastic, etc, and can be present in the kits as a package insert, in the labeling of the container of the kit or components thereof (e.g., associated with the packaging), etc. In other embodiments, the instructions are present as an electronic storage data file present on a suitable computer readable storage medium, e.g. CD-ROM, diskette, etc. In yet other embodiments, the actual instructions are not present in the kit, but means for obtaining the instructions from a remote source, such as over the internet, are provided. An example of this embodiment is a kit that includes a web address where the instructions can be viewed and/or from which the instructions can be downloaded. Often it will be desirable that some or all components of a kit are packaged in suitable packaging to maintain sterility. The components of a kit can be packaged in a kit containment element to make a single, easily handled unit, where the kit containment element, e.g. , box or analogous structure, may or may not be an airtight container, e.g., to further preserve the sterility of some or all of the components of the kit.

[00223] **Polynucleotides Encoding the Polypeptide**

[00224] In yet other aspects, the present disclosure further provides methods for preparing the polypeptide (or optionally the polypeptide conjugate) as described above. To this end, the following are provided in the present disclosure.

[00225] Firstly, the present disclosure provides isolated nucleic acids or polynucleotides that encode the fusion polypeptide as described herein.

[00226] The term “nucleic acid” or “polynucleotide” as used herein refers to deoxyribonucleic acids (DNA) or ribonucleic acids (RNA) and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses polynucleotides containing known analogues of natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular polynucleotide sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions), alleles, orthologs, SNPs, and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (see Batzer et al., Nucleic Acid Res.

19:5081 (1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); and Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)).

[00227] The nucleic acids or polynucleotides encoding the fusion polypeptide described herein can be constructed using recombinant techniques. To this end, DNA encoding the serum albumin-binding nanobody, DNA encoding the FGF21 domain, and optionally DNA encoding the GLP-1 domain, can be obtained and operably linked to allow transcription and expression in a host cell to produce the polypeptide. Polynucleotide sequences encoding for the polypeptide linkers are also operably linked to allow expression of the desired product.

[00228] The encoding polynucleotide sequences can be further operably linked to one or more regulatory sequences, optionally in an expression vector, such that the expression or production of the fusion polypeptides is feasible and under proper control.

[00229] The encoding polynucleotide sequence(s) can be inserted into a vector for further cloning (amplification of the DNA) or for expression, using recombinant techniques known in the art. Many vectors are available. The vector components generally include, but are not limited to, one or more of the following: a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter (e.g. prokaryotic promoters such as T7, T7lac, Sp6, araBAD, trp, lac, tac, pLm, A3, lac, lpp, npr, pac, syn, trc and T3, or eukaryotic promoters such as SV40, CMV, and EF-1α), and a transcription termination sequence.

[00230] **Vectors and Host Cells**

[00231] Secondly and thirdly, the present disclosure further provides a vector comprising the polynucleotide of provided above, and a host cell that comprises the vector described herein.

[00232] The term "vector" as used herein refers to a vehicle into which a polynucleotide encoding a protein may be operably inserted so as to bring about the expression of that protein. A vector may be used to transform, transduce, or transfect a host cell so as to bring about expression of the genetic element it carries within the host cell. Non-limiting examples of vectors include plasmids, phagemids, cosmids, artificial chromosomes such as yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), or P1-derived artificial chromosome (PAC), bacteriophages such as lambda phage or M13 phage, and animal viruses. Categories of animal viruses used as vectors include retrovirus (including lentivirus), adenovirus, adeno-associated virus, herpesvirus (*e.g.*, herpes simplex virus), poxvirus, baculovirus, papillomavirus, and papovavirus (*e.g.*, SV40). A vector may contain a variety of

elements for controlling expression, including promoter sequences, transcription initiation sequences, enhancer sequences, selectable elements, and reporter genes. In addition, the vector may contain an origin of replication. A vector may also include materials to aid in its entry into the cell, including but not limited to a viral particle, a liposome, or a protein coating. A vector can be an expression vector or a cloning vector. The present disclosure provides vectors (e.g., expression vectors) containing the nucleic acid sequence provided herein encoding the fusion polypeptide, at least one promoter (e.g., SV40, CMV, EF-1α) operably linked to the nucleic acid sequence, and at least one selection marker. Examples of vectors include, but are not limited to, retrovirus (including lentivirus), adenovirus, adeno-associated virus, herpesvirus (e.g., herpes simplex virus), poxvirus, baculovirus, papillomavirus, papovavirus (e.g., SV40), lambda phage, and M13 phage, plasmid pcDNA3.3, pMD18-T, pOptivec, pCMV, pEGFP, pIRES, pQD-Hyg-GSeu, pALTER, pBAD, pcDNA, pCal, pL, pET, pGEMEX, pGEX, pCI, pEGFT, pSV2, pFUSE, pVITRO, pVIVO, pMAL, pMONO, pSELECT, pUNO, pDUO, Psg5L, pBABE, pWPXL, pBI, p15TV-L, pPro18, pTD, pRS10, pLexA, pACT2.2, pCMV-SCRIPT.RTM., pCDM8, pCDNA1.1/amp, pcDNA3.1, pRc/RSV, PCR 2.1, pEF-1, pFB, pSG5, pXT1, pCDEF3, pSVSPORT, pEF-Bos etc.

[00233] The phrase "host cell" as used herein refers to a cell into which an exogenous polynucleotide and/or a vector has been introduced.

[00234] Vectors comprising the polynucleotide sequence(s) provided herein can be introduced to a host cell for cloning or gene expression. The phrase "host cell" as used herein refers to a cell into which an exogenous polynucleotide and/or a vector has been introduced. In other embodiments, the vectors are extra-chromosomal. The host cells can be isolated if desired. In certain embodiments, the host cell is a prokaryotic cell, and in some other embodiments, the host cell is an eukaryotic cell.

[00235] Suitable host cells for cloning or expressing the DNA in the vectors herein are mainly prokaryotes. Suitable prokaryotes for this purpose include eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as Escherichia, e.g., *E. coli*, Enterobacter, Erwinia, Klebsiella, Proteus, Salmonella, e.g., *Salmonella typhimurium*, Serratia, e.g., *Serratia marcescans*, and Shigella, as well as Bacilli such as *B. subtilis* and *B. licheniformis*, Pseudomonas such as *P. aeruginosa*, and Streptomyces. In some embodiments, host cells are eukaryotes, such as yeast and mammalian cells (e.g., immortalized mammalian cells).

[00236] A vector comprising the polynucleotide sequence(s) provided herein can be introduced into a host cell using any suitable method known to a skilled person in the art, e.g., transformation, transfection or transduction. In one example, the polynucleotide sequence encoding the fusion polypeptide can be subcloned into an expression vector, which is expressed as inclusion bodies in the host cells. The vector can be a viral vector, and any suitable viral vector can be used in this capacity.

[00237] Herein, the host cell can be a prokaryotic cell or an eukaryotic cell. Host cells transformed with the above-described expression or cloning vectors can be cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the cloning vectors.

[00238] In another aspect, the present disclosure provides a method of producing the fusion polypeptide as described herein, comprising culturing the host cell provided herein under a condition that allows expression of the fusion polypeptide as described herein.

[00239] For production of the fusion polypeptide as described herein, the host cells transformed with the expression vector may be cultured in a variety of media. Commercially available bacteria growth media such as Terrific Broth, LB Broth, LB Agar, M9 minimal media, MagiaMedia Medium, and ImMedia Medium (ThermoFisher) are suitable for culturing the bacterial host cells. Commercially available media such as Ham's F10 (Sigma), Minimal Essential Medium (MEM), (Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium (DMEM), Sigma) are suitable for culturing the eukaryotic host cells. Any of these media may be supplemented as necessary with hormones and/or other growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleotides (such as adenosine and thymidine), antibiotics (such as GENTAMYCIN^{TM} drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

[00240] **Methods for Producing the Polypeptide**

[00241] In one aspect, the present disclosure provides a method of expressing the fusion polypeptide as described herein, comprising culturing the host cell provided herein under the condition at which the fusion polypeptide as described herein is expressed.

[00242] In certain embodiments, the fusion polypeptide is expressed as inclusion bodies. In certain embodiments, the method further comprises renaturing the fusion polypeptide from the inclusion bodies.

[00243] When using recombinant techniques, the fusion polypeptide as described herein can be produced intracellularly, in the periplasmic space, or directly secreted into the medium. If the product is produced intracellularly, as a first step, the particulate debris, either host cells or lysed fragments, is removed, for example, by centrifugation or ultrafiltration. Carter et al., Bio/Technology 10:163-167 (1992) describe a procedure for isolating proteins which are secreted to the periplasmic space of E. coli. Briefly, cell paste is thawed in the presence of sodium acetate (pH 3.5), EDTA, and phenylmethylsulfonylfluoride (PMSF) over about 30 min. Cell debris can be removed by centrifugation. Where the product is secreted into the medium, supernatants from such expression systems are generally first concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. A protease inhibitor such as PMSF may be included in any of the foregoing steps to inhibit proteolysis and antibiotics may be included to prevent the growth of adventitious contaminants.

[00244] In certain embodiments, the method further comprises isolating the fusion polypeptide.

[00245] The fusion polypeptide as described herein prepared from the cells can be purified using, for example, hydroxylapatite chromatography, gel electrophoresis, dialysis, DEAE-cellulose ion exchange chromatography, ammonium sulfate precipitation, salting out, and affinity chromatography.

[00246] Other techniques for protein purification such as fractionation on an ion-exchange column, ethanol precipitation, Reverse Phase HPLC, chromatography on silica, chromatography on heparin SEPHAROSE^{TM} chromatography on an anion or cation exchange resin (such as a polyaspartic acid column), chromatofocusing, SDS-PAGE, and ammonium sulfate precipitation are also available depending on the protein to be recovered.

[00247] A method or process for producing the polypeptide as described above is also provided in the present disclosure, which substantially comprises the following two steps:

[00248] S100: culturing the host cell as described above under a condition that allows expression of the polynucleotide as defined above or a precursor thereof further comprising a removable tag; and

[00249] S200: collecting and purifying the polypeptide or the precursor thereof from the host cell.

[00250] In certain embodiments, the polypeptide is expressed as inclusion bodies. In certain embodiments, the method further comprises renaturing the polypeptide from the inclusion bodies to allow a refolding thereof. Further according to some embodiments, the host cell is *E. coli*, the vector comprises an *E. coli*-compatible vector, and the polypeptide encoded by the polynucleotide in the vector is codon optimized for *E. coli* expression.

[00251] When using recombinant techniques, the polypeptide as described herein can be produced intracellularly, in the periplasmic space, or directly secreted into the medium. If the product is produced intracellularly, as a first step, the particulate debris, either host cells or lysed fragments, is removed, for example, by centrifugation or ultrafiltration.

[00252] According to some embodiments, step S200 of collecting and purifying the polypeptide from the host cell may comprise the following sub-steps:

[00253] S210: collecting the precursor of the polypeptide;

[00254] S220: allowing the precursor of the polypeptide to refold;

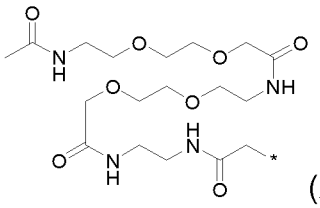
[00255] S230: treating the refolded precursor of the polypeptide to remove the tag and thereby obtain the polypeptide; and

[00256] S240: purifying the polypeptide.

[00257] In certain embodiments, the host cell is lysed and the polypeptide or the precursor of the polypeptide is obtained from the insoluble fraction containing the polypeptide or the precursor of the polypeptide.

[00258] According to some embodiments of the polypeptide producing process, the process further comprises conjugating the purified polypeptide with the functional moiety.

In certain embodiments, the functional moiety to be conjugated to the polypeptide is



(Ac-2XADO-EDA-CO-CH2-*).

[00259] EXAMPLES

[00260] Example 1: Recombinant expression of nanobody-FGF21 proteins and GLP-1-nanobody-FGF21

[00261] The GLP-1-nanobody-FGF21 proteins or nanobody-FGF21 fusion proteins listed in **Tables 1A and 1B, and Table 1C** were produced from bacterial *E. coli* expression system, using BL21(DE3) derivative strain. The DNA coding for the GLP-1-nanobody-FGF21 fusion precursor or nanobody-FGF21 fusion proteins was codon optimized for *E. coli* expression, de novo synthesized and subcloned into PET derivative expression vectors (Novagen). Amino acid substitutions were accomplished by modification of the corresponding genetic codes. Overexpression of GLP-1-nanobody-FGF21 fusion precursor or nanobody-FGF21 fusion proteins was induced with 0.5 mM isopropyl b-d-thiogalactoside (IPTG) when the cell density reached an OD600 of 2.0 in Terrific Broth (TB) medium. The cells were harvested after protein induction at 37 °C for 20-22 hours.

[00262] Table 1C Control molecules of GLP-1/FGF21 proteins

Molecule Code. (MLC); Fusion Polypeptide (SEQ ID NO)	Note	Conjugation on FGF21 domain
Control#1 (SEQ ID NO: 71)	GLP1(del7-8, 22E, 36G)-(G4S)4-ALB8-G4SG3S-FGF21 (121Q,168L,171G,180E) Control to MLC#8, GLP1 dead	no
Control#2 (SEQ ID NO: 72)	GLP1(del7-8, 22E, 36G)-(GAQP)5-ALB8-(GAQP)10-FGF21 (121Q,168L,171G,180E) GLP1 dead	no
Control#3 (SEQ ID NO: 73)	GLP1(del7-8, 22E, 36G)-(GAQP)5-ALB8-(GQEPGAQP)6-FGF21 (121Q,168L,171G,180E) GLP1 dead	no
Control#4 (SEQ ID NO: 76)	GLP1(del7-8, 22E, 36G)-(GAQP)5-ALB8-(GAQP)5-FGF21 (121Q,168L,171G,180E) GLP1 dead	no
Control#5 (SEQ ID NO: 77)	GLP1(del7-8, 22E, 36G)-(GAQP)5-ALB8-(GAQP)2-FGF21 (121Q,168L,171G,180E) GLP1 dead	no
Control#6 (SEQ ID NO: 78)	GLP1(del7-8, 22E, 36G)-(GAQP)5-ALB8-(GAQP)-FGF21 (121Q,168L,171G,180E) GLP1 dead	no
Control#7 (SEQ ID NO: 84)	GLP1(del7-8, 22E, 36G)-(GAQP)5-ALB8-FGF21 (121Q,168L,171G,180E) GLP1 dead	no
Control#8 (SEQ ID NO: 86)	GLP1(del 7-8, 22E, 36G)-(GAQP)14-ALB8-(GAQP)-FGF21 (121Q,168L,171C,180E) Ac-2XADO-EDA-CO-CH2 at 171C (0025 GLP1 dead)	Conjugated
Control#9 (SEQ ID NO: 87)	GLP1(8G,22E, 36G)-(GAQP)14-ALB8-(GAQP)-FGF21 (L137P,121Q,168L,171C,180E) Ac-2XADO-EDA-CO-CH2 at 171C (0025 FGF21 dead)	Conjugated
Control#10 (SEQ ID NO: 88)	GLP1(del7-8, 22E, 36G)-(GAQP)14-ALB8-(GAQP)-FGF21 (L137P,121Q,168L,171C,180E) Ac-2XADO-EDA-CO-CH2 at 171C (0025 GLP1 dead, FGF21 dead)	Conjugated

[00263] **Example 2: Purification of nanobody-FGF21 proteins**

[00264] Cells were harvested and lysed in 20mM Tris pH8.0, 0.15M NaCl buffer by cell disruptor (900 bar, for twice). The insoluble fractions, containing nanobody-FGF21 fusion proteins were collected and by centrifugation (8,000 X g, for 30min). After refolding, the fusion proteins were purified by anion exchange chromatography. The samples in each step were characterized by LC/MS to confirm the correct molecular weight.

[00265] **Example 3: Purification of GLP-1-nanobody-FGF21**

[00266] Cells were harvested and lysed in 20mM Tris pH8.0, 0.15M NaCl buffer by cell disruptor (900 bar, for twice). The insoluble fractions, containing the GLP-1-nanobody-FGF21 fusion precursor were collected and by centrifugation (8,000 X g, for 30min). After refolding, the fusion protein precursors were captured by anion exchange chromatography. After removing tag by protease, the proteins were purified by hydrophobic interaction chromatography. The samples in each step were characterized by LC/MS to confirm the correct molecular weight.

[00267] **Example 4: Preparation of nanobody-FGF21 and GLP-1-nanobody-FGF21 fusion protein conjugate**

[00268] To a solution of a nanobody-FGF21 or GLP-1-nanobody-FGF21 fusion protein in Tris buffer was added with Ac-2XADO-EDA-CO-CH2-Br in organic solvent dropwise. The reaction was stirred at room temperature for 1 h. Then the product was applied to aion exchange chromatography. This provided the compounds as listed in Table 1 as shown above.

[00269] The conjugated fusion proteins were detected and characterized by LC-MS method with Waters BioAccord LC-MS system, or by UPLC with Waters Acquity UPLC system, using conditions optimized for different conjugates, following the supplier's manuals.

[00270] **Example 5: In vitro activities**

[00271] Method: The in vitro GLP-1 activities of the fusion protein were measured using a BHK cell line overexpressing human GLP-1 receptor and CRE luciferase reporter with or without 1% human serum albumin (HSA). Tested fusion proteins were measured at 100 nM as top concentration in the presence of 1% HSA with 3-fold serial dilutions. After cells were

treated with molecules for 4 hours, luciferase activities were measured by Steadylite plus kit (Perkin Elmer, 6066751).

[00272] The activity of each fusion protein was represented by EC50, derived from non-linear regression analysis.

[00273] The *in vitro* FGF21 activities of the fusion protein were assessed using a HEK293 cell line overexpressing human beta-Klotho. Tested fusion protein conjugates were measured at 400 nM top concentration with 4-fold serial dilutions in the presence of 1% HSA. After cells were treated with fusion protein conjugates for 12 mins, p-ERK levels were measured by p-ERK kit (Cisbio, 64ERKPEH).

[00274] The activity of each fusion protein was represented by $EC_{50,}$ derived from non-linear regression analysis.

[00275] Conclusion: As shown in table 2 and 3, all fusion proteins showed comparable potency with native FGF21. However, the fusion molecules (MLC#9, MLC#10, MLC#12, MLC#13, MLC#14, MLC#15, MLC#16, MLC#17, MLC#19 and MLC#21) exhibited different GLP-1 activities. The molecule MLC#14, MLC#16, MLC#17, MLC#19 and MLC#21 showed significantly higher GLP1 activity than MLC#9, MLC#12, MLC#13 and MLC#15. Compared with other fusion molecules tested, MLC#15 has relatively shorter second linker and the GLP-1 activity is also lower. Taken together, this may indicate that the fusion molecule with longer second linker can have higher GLP1 activity. As shown in Figure 2A and 2B, the fusion molecules (MLC#9, MLC#10, control#6 and MLC#14) showed significantly higher FGF21 efficacy than YH-dual molecule (i.e. the molecule disclosed as SEQ ID NO. 66 in WO 2017/074123).

[00276] **Table 2.** In vitro GLP1 activities and FGF21 activity of GLP-1-Nanobody-FGF21 fusion proteins conjugates.

Molecule Code.	**Second linker**	**First linker**	**Conjugation on FGF21 domain**	**GLP-1 activity (EC_{50}, pM)**	**FGF21 activity (EC_{50}, nM)**
Semaglutide	no	no	no	282	NA
Native FGF21	no	no	no	NA	0.96

Control#1	(G4S)4	G4SG3S	no	>50000	1.04
Control#2	(GAQP)5	(GAQP)10	no	>50000	2.10
Control#13	(GAQP)5	(GQEPGAQP)6	no	>50000	3.47
Control#4	(GAQP)5	(GAQP)5	no	>50000	1.47
Control#5	(GAQP)5	(GAQP)2	no	>50000	1,35
Control#6	(GAQP)5	(GAQP)	no	>50000	2.26
MLC#10	(GAQP)5	(GAQP)10	no	654	1.83
MLC#9	(GAQP)5	(GAQP)10	171C	2184	4.26
MLC#12	(GAQP)5	(GAQP)	no	2220	1.78
MLC#13	(GAQP)5	(GAQP)	171C	3854	5.06
MLC#14	(GAQP)10	(GAQP)	171C	416	4.01
MLC#15	(GAQP)2	(GAQP)	171C	9711	1.01
MLC#16	(GAQP)14	(GAQP)	171C	118	2.31
MLC#17	(GAQP)14	(GAQP)	174C	207	3.75
MLC#21	(GAQP)14	G4S	no	128	5.28
MLC#19	(GAQP)14	G4S	171C	192	4.98

NA: not available.

[00277] **Table 3.** In vitro FGF21 activity of Nanobody-FGF21 fusion proteins conjugates.

Molecule Code.	**First linker**	**Conjugation on FGF21 domain**	**FGF21 activity (EC_{50}, nM)**
MLC#5	(GAQP)	no	7.65
MLC#6	(GAQP)	171C	6.65
MLC#25	G4S	no	7.45
MLC#23	G4S	171C	5.83

[00278] **Example 6: Efficacy study in disease models**

[00279] Selected molecules are assessed in disease animal models (such db/db mice, Diet-Induced-Obesity (DIO) mice) to determine body weight, food intake, glucose efficacy with dose responses in chronic studies. Some biomarkers are also measured, including

plasma insulin, plasma triglyceride, plasma cholesterol, plasma LDL-c, plasma adiponectin, liver triglyceride, liver cholesterol and liver function indexes (ALT, AST).

[00280] A) Food Intake and Body Weight Reduction

[00281] Method:

[00282] 22 week old DIO male C57BL/6 mice (~50g) were injected once every other day (Q2D) subcutaneously with designated GLP-1 polypeptide conjugates (i.e., Molecule 012) for 25 days. Food intake and body weight were measured twice a week and fasting blood glucose was measured once a week. Five animals were used for each treatment group. Body weight and fasting blood glucose was monitored for each individual animal, but food intake for each group animals was measured together. Day 1 and Day 25 are first day and last day of molecule dosage. Data are indicated as mean values and standard error (SEM) or pooled values. Statistical analysis was performed by One-way ANOVA. Body weight reduction on Day 25 is calculated by -1*(%BW loss -%BW loss of vehicle group); Cumulative food intake reduction is calculated by -100*(cumulative food intake-cumulative food intake of vehicle)/cumulative food intake of vehicle.

[00283] Conclusion: in DIO study, as shown in Figure 3A, 3B and Table 8, Molecule Control#2, Control#6, MLC#9 and MLC#10 have significant effect on body weight reduction and food intake suppression.

Table 8 Food intake and body weight reduction in DIO mice at Day 11

Group	**Cumulative food intake reduction (% reduction compared to vehicle) (Day11)**	**Body weight reduction (% reduction compared to vehicle) (Day11)**
Control#2 10nmol/kg QD	19.1	21.7
Control#6 10nmol/kg QD	13.1	24.6
MLC#9 10nmol/kg QD	27.1	30.3
MLC#10 10nmol/kg QD	28.0	30.9

[00284] B) Metabolic Parameters in DIO animal model

[00285] Method:

[00286] 16 week old DIO male C57BL/6 mice (35~40g) were injected once daily (QD) subcutaneously with designated GLP-1/FGF21 conjugates (i.e., MLC#9, MLC#14, MLC#16 and MLC#17) for 21 days. Food intake and body weight were measured once everyday and non-fasting blood glucose was measured once every three days. Five animals were used for each treatment group. Body weight and blood glucose was monitored for each individual animal, but food intake for each group of animals was measured together. Day 1 and Day 21 are first day and last day respectively of the treatment. Terminal blood was collected and EDTA-K3 plasma was prepared and frozen at -80 °C for biomarker measurement (LDL-C, TC, TG, ALT, insulin, adiponectin). Liver and adipose tissue were also collected and frozen in liquid nitrogen and store at -80 °C. Data are indicated as mean values and standard error (SEM) or pooled values. Statistical analysis was performed by One-way ANOVA. Body weight reduction on Day 21 is calculated by -1*(%BW loss -%BW loss of vehicle group); Cumulative food intake reduction is calculated by -100*(cumulative food intake-cumulative food intake of vehicle)/cumulative food intake of vehicle.

[00287] Conclusion:

[00288] In DIO mice study, as shown in Figure 4A, Molecule MLC#9, MLC#14, MLC#16 and MLC#17 showed much better efficacy on body weight reduction compared to Semaglutide, Tirzepatide and YH-dual (SEQ ID 66 in WO2017/074123). MLC#9, MLC#14, MLC#16 and MLC#17 induced body weight loss by about 30-35% in the study. In contrast, Semaglutide, Tirzepatide and YH-dual induced body weight loss by about 20% to 25%.

[00289] In Figure 4B, the MLC#9, MLC#14, MLC#16 and MLC#17 groups showed better glucose control than Semaglutide, Tirzepatide and YH-dual. MLC#9, MLC#14, MLC#16 and MLC#17 decreased non-fasting glucose level to lower than 7 mmol/L. In contrast, Semaglutide decreased non-fasting glucose to about 7.5-10 mmol/L.

[00290] MLC#9, MLC#14, MLC#16 and MLC#17 also induced the reduction of triglyceride (Figure 4C), LDL-C (Figure 4D), total cholesterol (Figure 4E) and ALT (Figure 4F) concentration in plasma and reduction of adipose weight (Figure 4G) and liver weight (Figure 4H), along with the improvement of liver TG (Figure 4I), liver TC (Figure 4J) and insulin sensitivity (Figure 4K).

[00291] In vivo, FGF21 can induce secretion of adiponectin which is reported to be an insulin sensitizer. The fusion molecules MLC#9, MLC#14, MLC#16 and MLC#17 showed better effect on increase of adiponectin level than YH-dual indicating the better FGF21 activity (Figure 4L).

[00292] C) Metabolic Parameters in ob/ob animal model at different dosages

[00293] Method: 10 week old ob/ob male mice (42~55g) were injected once daily (QD) subcutaneously with designated GLP-1/FGF21conjugates (i.e., MLC#14 and MLC#16) for 14 days. Food intake and body weight were measured once everyday. Non-fasting blood glucose was measured once every three days. Plasma triglyceride (TG) levels was measured once every week. Five or six animals were used for each treatment group. Body weight and blood glucose was monitored for each individual animal, but food intake for each group animals was measured together. Day 1 and Day 14 are first day and last day of molecule dosage. On the Day 14 after treatment, mice were sacrificed via cardiac puncture under anesthesia. Terminal blood was collected and frozen at -80 °C for biomarker measurement (LDL-C, TC, TG, ALT/AST, insulin, adiponectin). Liver and adipose tissue were also collected and frozen in liquid nitrogen and store at -80 °C. Data are indicated as mean values and standard error (SEM) or pooled values. Statistical analysis was performed by One-way ANOVA. Body weight reduction on Day 14 is calculated by -1*(%BW loss -%BW loss of vehicle group); Cumulative food intake reduction is calculated by -100*(cumulative food intake-cumulative food intake of vehicle)/cumulative food intake of vehicle.

[00294] Conclusion: in ob/ob mice study, as shown in Figure 5A-5M, the fusion molecule MLC#14 and MLC#16 showed dose-dependent efficacy on body weight reduction (Figure 5A), glucose lowering and biomarkers changes. In Figure 5A, molecule MLC#14 and MLC#16 showed much better efficacy on body weight reduction compared to Semaglutide, Tirzepatide and YH-dual (SEQ ID 66 in WO2017/074123) at the same dosage. In Figure 5B, MLC#14 and MLC#16 groups reached similar effect on glucose control as Semaglutide at the same dosage. As shown in Figure 5C, MLC#14 and MLC#16 induced better reduction of plasma triglyceride compared to Semaglutide, Tirzepatide and YH-dual at the same dosage. MLC#14 and MLC#16 also reduced the level of LDL-C (Figure 5D), total cholesterol (Figure 5E), ALT (Figure 5F) and AST (Figure 5G) concentration in plasma and reduction of adipose weight (Figure 5H) and liver weight (Figure 5I) along with the improvement of liver TG (Figure 5J), liver TC (Figure 5K) and insulin sensitivity (Figure 5L). The fusion molecule

MLC#14 and MLC#16 showed significantly better effect on increase of adiponectin level than YH-dual at the same dosage indicating the better FGF21 activity (Figure 5M).

[00295] D) Metabolic Parameters in DIO animal model

[00296] Method: 16 week old DIO male C57BL/6 mice (35~50g) were injected once daily (QD) subcutaneously with designated GLP-1/FGF21 conjugates and FGF21 conjugates (i.e., MLC#16, MLC#19, MLC#6 and MLC#23) for 22 days. Food intake and body weight were measured once every three days and fasting blood glucose was measured every weeks. Five animals were used for each treatment group. Body weight and blood glucose was monitored for each individual animal, but food intake for each group of animals was measured together. Day 1 and Day 22 are first day and last day respectively of the treatment. Terminal blood was collected and EDTA-K3 plasma was prepared and frozen at -80 °C for biomarker measurement (LDL-C, TC, TG). Liver and adipose tissue were also collected and frozen in liquid nitrogen and store at -80 °C. Data are indicated as mean values and standard error (SEM) or pooled values. Statistical analysis was performed by One-way ANOVA.

[00297] Conclusion

[00298] In DIO mice study, as shown in Figure 6A, the GLP-1-Nanobody-FGF21 fusion proteins conjugates (MLC#16 and MLC#19) and the Nanobody-FGF21 fusion conjugates (MLC#6 and MLC#23) showed good efficacy on body weight reduction. In Figure 6B, MLC#16, MLC#19, MLC#6 and MLC#23 groups showed better glucose control than Semaglutide, and also induced the reduction of serum triglyceride (Figure 6C), LDL-C (Figure 6D) and total cholesterol (Figure 6E) concentration in plasma and reduction of adipose weight (Figure 6F) and liver weight (Figure 6G), along with the improvement of liver TG (Figure 6H).

[00299] **Example 7: Pharmacokinetic study in rat**

[00300] Method: 6-8 weeks male SD rats were administrated in a single subcutaneous dose of 15 nmol/kg protein MLC#9 (conjugated) or MLC#10 (non-conjugated) (n=3/group). Plasma samples were collected pre-dose (-5 min), 0.5 hr, 1 hr, 2 hr, 4 hr, 6 hr, 8 hr, 12 hr, 24 hr, 32 hr, 48 hr, 72 hr, 96 hr, after subcutaneous administration. The concentrations of the polypeptide conjugates in the plasma were measured by ELISA method. Based on the graph showing plasma concentration of each polypeptide conjugate versus time after subcutaneous injection, the pharmacokinetic parameters were calculated by WinNonlin.

[00301] Conclusion: The MLC#9 has half-life $T_{1/2}$ of 14.3 hrs in rat which shows longer half life than the MLC#10 with half-life $T_{1/2}$ of 9.5 hrs.

[00302] **Example 8: PK study in minipig.**

[00303] Pharmacokinetics of selected molecules are assessed in minipig. Both subcutaneous and intravenous injections are performed.

[00304] **Example 9: PK study in non-human primates.**

[00305] Pharmacokinetics of selected molecules are assessed in monkeys. Both subcutaneous and intravenous injections are performed.

[00306] Methods: 4~5 years male Cynomolgus monkeys were administrated by a single subcutaneous dose of 5 mg/kg (123 nmol/kg) MLC#16 (n=2/group). Plasma samples were collected pre-dose (-5 min), 0.5 hr, 1 hr, 2 hr, 4 hr, 6 hr, 8 hr, 12 hr, 24 hr, 36 hr, 48 hr, 72 hr, 96 hr, 120 hr, 144 hr and 168 hr after subcutaneous administration. The concentrations of MLC#16 in the plasma were measured by LC-MS/MS method. Based on the graph showing plasma concentration of MLC#16 versus time after subcutaneous injection, the pharmacokinetic parameters were calculated by WinNonlin.

[00307] Conclusion: MLC#16 showed a half-life of 54.3 hours in monkey, which potentially supports a once weekly dose frequency in human (**Table 9**).

Table 9: Pharmacokinetic parameters of MLC#16 in monkeys. Pharmacokinetic data were analyzed by WinNonlin software. T_{max}, C_{max}, $T_{1/2}$, AUC were calculated.

PK parameters	Unit	MLC#16
T_{max}	hr	12
C_{max}	nmol/L	671
Terminal $t_{1/2}$	hr	54.3
AUC_{0-t}	hr*nmol/L	47851 (t=168h)

[00308] **Example 10: Immunogenicity assessment.**

[00309] Selected GLP-1 polypeptide conjugates are also assessed for immunogenicity by in silico (iTope and TCED methods) and ex vivo (EpiScreen) methods.

[00310] **Example 11: Stability assessment.**

[00311] To test the stability, the different GLP-1 polypeptide conjugates are formulated in buffers with different compositions (pH6, 7, 7.4 and 8.0) and stored at different temperatures (such as 4 °C and 25 °C) for 2-4 weeks. The %HMWP and %LMW are analyzed by size exclusion chromatography (SEC)-HPLC. The concentration and modifications were analyzed by reverse phase (RP)-UPLC and LC/MS.

[00312] **Example 12: Human serum albumin binding**

[00313] Method: Binding of molecules to serum albumin was characterized by surface plasmon resonance in a Biacore 8K instrument. Human serum albumin was covalently bound to CM5 sensor chips surface until 4000 RU was reached. The chip was blocked by 1M ethanolamine with flowrate of 10 μL/min for 420 s. Each molecule sample was diluted and injected at a flow rate of 30 μL/min to allow for binding to chip-bound albumin for 120 s and for dissociation for 300 s. Binding buffer without molecule was sent over the chip at the flow rate of 20 seconds to allow spontaneous dissociation of bound molecule for 30 seconds.

[00314] Conclusion: All fusion molecules (MLC#9, MLC#10, Control#4, Control#5, Control#6, MLC#12 and MLC#16) show similar binding affinity to human serum albumin (see **Table 10**).

[00315] Table 10: Human serum albumin binding affinity

Molecule	KD (μM)	Rmax (RU)
MLC#9 (conjugated)	**0.28**	1507.8
MLC#10 (non-conjugated)	**0.28**	1391.0
Control#4	**0.27**	1252.0
Control#5	**0.22**	1604.0
Control#6	**0.22**	1289.8
MLC#12 (non-conjugated)	**0.25**	1248.6
MLC#13 (conjugated)	**0.27**	1393.3

MLC#14 (conjugated)	**0.22**	1513.5
MLC#15 (conjugated)	**0.31**	1341.8
MLC#16 (conjugated)	**0.27**	1625.6
MLC#19(conjugated)	**0.22**	450.0

[00316] **Example 13: FAP enzyme cleavage on the fusion proteins**

[00317] Method: Fibroblast activation protein (FAP) is a serine protease. This enzyme has been reported to regulate degradation of FGF21, In order to test the C terminal degradation of the fusion protein by FAP enzyme, the fusion protein (MLC#23, MLC#25) and FAP enzyme were incubated with ration 200:1 at 37 °C for 20 hours. LC-MS were performed to analyze the degradation percentage of the fusion protein.

[00318] Conclusion: The fusion protein without conjugation (MLC#25, MLC#18 and MLC#21) showed 50.6-65.2% C-terminal degradation when incubation with FAP enzyme at 37 °C for 20 hours. However, the fusion protein conjugate (MLC#23, MLC#16 and MLC#19) and the fusion protein with 171G substitution (MLC#10) showed resistance to FAP enzyme.

[00319] Table 11: Result of fusion proteins by FAP enzyme cleavage

Molecule	C-terminal degradation
MLC#23 (conjugated)	**Not detected**
MLC#25 (non-conjugated, 171P)	**58.9%**
MLC#16 (conjugated)	**Not detected**
MLC#18 (non-conjugated, 171P)	**65.2%**
MLC#19 (conjugated)	**Not detected**
MLC#21 (non-conjugated, 171P)	**50.6%**
MLC#10 (non-conjugated, 171G)	**Not detected**

Claims

1. A polypeptide, comprising, in a direction from N terminus to C terminus:

a first fragment comprising a nanobody domain capable of binding to serum albumin; and

a second fragment comprising a biologically active FGF21 domain;

wherein:

the first fragment and the second fragment are connected via a first linker.

2. The polypeptide of claim 1, wherein the FGF21 domain comprises one or more amino acid residue mutations, each at a position selected from positions selected from the group consisting of positions 121, 168, 171, and 180 relative to SEQ ID NO: 1.

3. The polypeptide of claim 2, wherein the one or more amino acid residue mutations in the FGF21 domain comprise N121Q, M168L, P171G and A180E, or any combination thereof.

4. The polypeptide of any one of the preceding claims, wherein the FGF21 domain further comprises a conjugatable residue.

5. The polypeptide of claim 4, wherein:

the conjugatable residue is at a position within a C-terminal fragment spanning from position 169 to position 181 relative to SEQ ID NO: 1; and optionally

the conjugatable residue is at a position selected from the group consisting of positions 169, 170, 171, 172, 173, 174, 180 and 181 relative to SEQ ID NO: 1.

6. The polypeptide of claim 4 or 5, wherein the FGF21 domain comprises the amino acid sequence of SEQ ID NOs: 6-13, 16-19, and 92.

7. The polypeptide of any of claims 4-6, wherein the polypeptide is conjugated with a functional moiety at the conjugatable residue in the second fragment.

8. The polypeptide of claim 7, wherein the functional moiety comprises a glycosyl moiety or a synthetic chemical moiety.

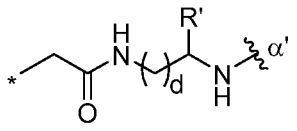
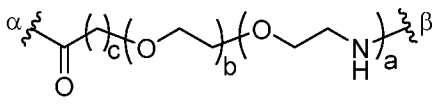
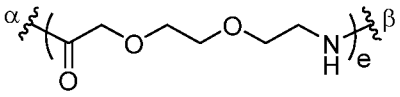
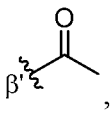
9. The polypeptide of any one of claims 7-8, wherein the functional moiety comprises a glycosyl moiety, and the conjugatable residue is an introduced residue that is glycosylatable.

10. The polypeptide of claim 9, wherein the conjugatable residue comprises an introduced T at position 172 or position 173, or an introduced N residue at position 170 or position 174, relative to SEQ ID NO: 1.

11. The polypeptide of any one of claims 7-8, wherein the functional moiety comprises a synthetic chemical moiety.

12. The polypeptide of claim 11, wherein the conjugatable residue comprises an introduced cysteine residue.

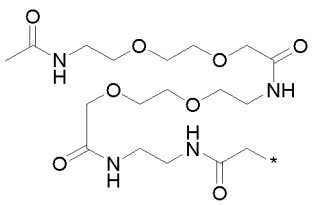
13. The polypeptide of claim 12, wherein the synthetic chemical moiety comprises a structure of *-X-Y-Z, wherein X, Y, and Z are interconnected via bonds, and the * end of X is connected to the conjugatable residue on the polypeptide, wherein:

X can be ; Y can be , or ; and Z can be ,

wherein:

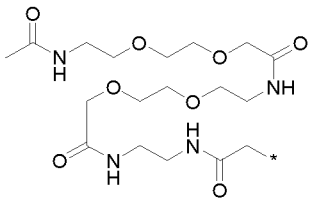
position α is linked to position α', position β is linked to position β', R1 is hydrogen or –COOH; d is 1, 2, or 3; a is 1, 2 or 3; b is 1, 2 or 3; c is 1 or 2; and d is 1, 2, or 3, and e is 1, 2, or 3.

14. The polypeptide of claim 13, wherein the synthetic chemical moiety has the below structure:



15. The polypeptide of any of claims 12-14, wherein the introduced cysteine is at a position selected from a group consisting of 169, 170, 171, 172, 173, 174, 180 and 181 relative to SEQ ID NO: 1.

16. The polypeptide of claim 15, wherein the introduced cysteine is at position 171, and the synthetic chemical moiety has the below structure:



17. The polypeptide of claim 16, wherein the FGF21 domain comprises an amino acid sequence of SEQ ID NO: 8, and optionally the introduced cysteine is at position 171.

18. The polypeptide of claim 15, wherein:

i) the FGF21 domain comprises an amino acid sequence of SEQ ID NO: 6, and the introduced cysteine is at position 169;

ii) the FGF21 domain comprises an amino acid sequence of SEQ ID NO: 7, and the introduced cysteine is at position 170;

iii) the FGF21 domain comprises an amino acid sequence of SEQ ID NO: 9, and the introduced cysteine is at position 172;

iv) the FGF21 domain comprises an amino acid sequence of SEQ ID NO: 10, and the introduced cysteine is at position 173;

v) the FGF21 domain comprises an amino acid sequence of SEQ ID NO: 11, and the introduced cysteine is at position 174;

vi) the FGF21 domain comprises an amino acid sequence of SEQ ID NO: 12, and the introduced cysteine is at position 180;

vii) the FGF21 domain comprises an amino acid sequence of SEQ ID NO: 13, and the introduced cysteine is at position 181, or

viii) the FGF21 domain comprises an amino acid sequence of SEQ ID NO: 92, and the introduced cysteine is at position 174.

19. The polypeptide of any one of the claims 1-15, wherein the FGF21 domain further comprises a substitution P171G relative to SEQ ID NO: 1.

20. The polypeptide of any of preceding claims, wherein the nanobody domain comprises a VHH domain.

21. The polypeptide of claim 20, wherein the VHH domain is humanized.

22. The polypeptide of claim 20 or 21, wherein the VHH domain comprises a complementarity determining region 1 (CDR1), a complementarity determining region 2 (CDR2) and a complementarity determining region 3 (CDR3), wherein:

the CDR1 comprises the sequence of SEQ ID NO: 20 or a variant thereof having up to 3, 2, or 1 amino acid mutation;

the CDR2 comprises the sequence of SEQ ID NO: 21 or a variant thereof having up to 3, 2, or 1 amino acid mutation; and/or

the CDR3 comprises the sequence of SEQ ID NO: 22 or a variant thereof having up to 3, 2, or 1 amino acid mutation;

wherein:

the VHH domain substantially retains the binding specificity to serum albumin, optionally to human serum albumin.

23. The polypeptide of claim 22, wherein the VHH domain comprises a complementarity determining region 1 (CDR1) comprising the sequence of SEQ ID NO: 20, a CDR2 comprising the sequence of SEQ ID NO: 21, and a CDR3 comprising the sequence of SEQ ID NO: 22.

24. The polypeptide of any of claims 20-23, wherein the VHH domain comprises an amino acid sequence of SEQ ID NO: 23, or a variant thereof having at least 70% (e.g. at least 75%, 80%, 85%, 90%, 95%, 99% identity to SEQ ID NO: 23, wherein the variant substantially retains binding specificity and/or affinity to serum albumin.

25. The polypeptide of claim 24, wherein the variant of SEQ ID NO: 23 has up to 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acid mutation relative to SEQ ID NO: 23.

26. The polypeptide of any one of claims 20-25, wherein the nanobody domain further comprises an N-terminal extension attached to a N-terminus of the VHH domain.

27. The polypeptide of claim 26, wherein the N-terminal extension comprises amino acid residues of SG, AG, S or A.

28. The polypeptide of claim 27, wherein the nanobody domain comprises an amino acid sequence selected from SEQ ID NOs: 24-27.

29. The polypeptide of any of the preceding claims, wherein the first linker has a length of at least four amino acid residues.

30. The polypeptide of claim 29, wherein the first linker comprises no acidic amino acid residue.

31. The polypeptide of claim 30, wherein the first linker does not comprise any of D residue or E residue.

32. The polypeptide of claim 29, wherein the first linker comprises one or more units of a first repeating sequence.

33. The polypeptide of claim 32, wherein the first repeating sequence consists of no more than 4 or 6 types of amino acid residues selected from the group consisting of: G, Q, A, P, T and S.

34. The polypeptide of claim 33, wherein the first repeating sequence comprises or consists of an amino acid sequences selected from a group consisting of G_fS_g, SEQ ID NO: 35 (GAQP), SEQ ID NO: 36 (GQAP), SEQ ID NO: 37 (GPAQ), SEQ ID NO: 38 (GPQA), SEQ ID NO: 39 (GSQP), SEQ ID NO: 40 (GASP), SEQ ID NO: 41 (GPAS), SEQ ID NO: 42 (GPSA), SEQ ID NO: 43 (GGGS), SEQ ID NO: 44 (GSGS), SEQ ID NO: 45 (GGGGS), SEQ ID NO: 46 (GSAPGSPAGSPTGSAPGSPA) and GS, wherein each of f and g is independently an integer selected from 1 to 5.

35. The polypeptide of claim 34, wherein the first repeating sequence has an amino acid sequence set forth in SEQ ID NO: 35 (GAQP), and a number of the one or more units is an integer between 1 and 10.

36. The polypeptide of claim 35, wherein the first linker comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 35 (GAQP), SEQ ID NO: 49 ($(GAQP)_2$),

SEQ ID NO: 50 ($(GAQP)_5$), SEQ ID NO: 51 ($(GAQP)_{10}$), and SEQ ID NO: 48 (GGGGSGGGS).

37. The polypeptide of any of the preceding claims, further comprising, over an N-terminus of the first fragment, a third fragment comprising another functional domain, wherein:

the first fragment and the third fragment are connected via a second linker.

38. The polypeptide of claim 37, wherein the another functional domain of the third fragment comprises a biologically active protein, or a fragment thereof, of one selected from the group consisting of glucagon-like peptide- 1 (GLP-1), insulin, leptin, glucagon, gastrin, gastric inhibitory polypeptide (GIP), amylin, calcitonin, cholecystokinin, peptide YY, neuropeptide Y, bone morphogenetic protein-6 (BMP-6), bone morphogenetic protein-9 (BMP-9), oxyntomodulin, oxytocin, glucagon-like peptide-2 (GLP-2), irisin, fibronectin type III domain containing protein 5 (FNDC5), apelin, adiponectin, Clq and tumor necrosis factor related protein (CTRP family), resistin, visfatin, omentin, retinol binding protein-4 (RBP-4), glicentin, angiopoietin, interleukin-22 (IL-22), exendin-4, and growth hormone.

39. The polypeptide of claim 38, wherein the another functional domain of the third fragment comprises a biologically active peptide, or a fragment thereof, of GLP-1, wherein the another functional domain comprises an amino acid sequence having at least 70% sequence identity to, while retaining substantial biological activity of, SEQ ID NO: 28.

40. The polypeptide of claim 39, wherein the another functional domain comprises one or more mutations at positions 8, 22, 26, 34, and 36, or any combination thereof, relative to SEQ ID NO: 28.

41. The polypeptide of claim 40, wherein the one or more mutations comprise A8G, G22E, K26R, K34R, R36G, or any combination thereof.

42. The polypeptide of claim 41, wherein the another functional domain comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 29, and SEQ ID NOs: 31-34.

43. The polypeptide of any of claims 37-42, wherein the second linker has a length of at least eight amino acid residues (e.g. at least 12, 16, or 20 amino acid residues).

44. The polypeptide of claim 43, wherein the second linker comprises one or more units of a second repeating sequence.

45. The polypeptide of claim 41, wherein the second repeating sequence consists of no more than 4 or 6 types of amino acid residues selected from the group consisting of: G, Q, A, E, P, T and S.

46. The polypeptide of claim 45, wherein the second repeating sequence comprises or consists of an amino acid sequences selected from a group consisting of G_hS_i, SEQ ID NO: 35 (GAQP), SEQ ID NO:55 (GQEP), SEQ ID NO: 56 (GEQP), SEQ ID NO: 57 (GPQE), SEQ ID NO: 58 (GPEQ), SEQ ID NO: 59 (GSEP), SEQ ID NO: 60 (GESP), SEQ ID NO: 61 (GPSE), SEQ ID NO: 62 (GPES), SEQ ID NO: 36 (GQAP), SEQ ID NO: 37 (GPAQ), SEQ ID NO: 38 (GPQA), SEQ ID NO: 39 (GSQP), SEQ ID NO: 40 (GASP), SEQ ID NO: 41 (GPAS), SEQ ID NO: 42 (GPSA), SEQ ID NO: 43 (GGGS), SEQ ID NO: 44 (GSGS), SEQ ID NO: 45 (GGGGS), SEQ ID NO: 46 (GSAPGSPAGSPTGSAPGSPA) and GS, wherein each of h and i is independently an integer selected from 1 to 5.

47. The polypeptide of claim 46, wherein the second repeating sequence has an amino acid sequence set forth in SEQ ID NO: 35 (GAQP), and a number of the one or more units is an integer between 1 and 15.

48. The polypeptide of claim 43, wherein the second linker comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 49 $((GAQP)_2)$, SEQ ID NO: 50 $((GAQP)_5)$, SEQ ID NO: 51 $((GAQP)_{10})$, and SEQ ID NO: 52 $((GAQP)_{14})$, and SEQ ID NO: 47 $((GGGGS)_4)$.

49. The polypeptide of any one of preceding claims, comprising an amino acid sequence selected from SEQ ID NOS: 63-68, 93, 99, 100, 101 and 107.

50. The polypeptide of any one of claims 37-48, comprising an amino acid sequence selected from SEQ ID NOS: 70, 74, 75, 79-83, 85, 94-98, and 108-109.

51. A pharmaceutical composition, comprising the polypeptide according to any one of the preceding claims and a pharmaceutically acceptable carrier.

52. A method of preventing or treating a metabolic disorder in a subject in need thereof, comprising:

administering a therapeutically effective amount of the polypeptide according to any one of claims 1-50 or the pharmaceutical composition according to claim 51.

53. The method of claim 52, wherein the metabolic disorder is diabetes, obesity, non-alcoholic steatohepatitis (NASH), cardiovascular like dyslipidaemia, arteriosclerosis, alcoholic steatohepatitis (ASH), diabeticnephropathy, gestational diabetes, metabolic syndrome such as metabolic syndrome X, nonalcoholic fatty liver disease (NAFLD), end-stage liver disease, hepatic steatosis (fatty liver), liver cirrhosis, primary biliary cirrhosis (PBC) or severe hypertriglyceridemia (SHTG).

54. A polynucleotide that encodes the polypeptide as defined in any of claims 1-50.

55. A vector comprising the polynucleotide according to claim 54.

56. A host cell comprising the vector according to claim 55.

57. A process for producing the polypeptide as defined in any of claims 1-50, comprising:

culturing the host cell of claim 56 under a condition that allows expression of the polynucleotide of claim 54 or a precursor thereof further comprising a removable tag; and

collecting and purifying the polypeptide or the precursor thereof from the host cell.

58. The process of claim 57, wherein the polypeptide is conjugated with a functional moiety at a conjugatable residue in the second fragment, wherein the process further comprises, after the collecting and purifying the polypeptide from the host cell:

conjugating the functional moiety to the polynucleotide.

59. The process of claim 57 or claim 58, wherein the host cell is *E. coli,* the vector comprises an *E. coli*-compatible vector, and the polypeptide encoded by the polynucleotide in the vector is codon optimized for *E. coli* expression.

60. The process of any of claims 57-59, wherein the collecting and purifying the polypeptide from the host cell comprises:

collecting the precursor of the polypeptide therefrom;

allowing the precursor of the polypeptide to refold;

treating the refolded precursor of the polypeptide to remove the tag and thereby obtain the polypeptide; and

purifying the polypeptide.

61. The process of any of claims 58-60, further comprising conjugating the purified polypeptide with the functional moiety.

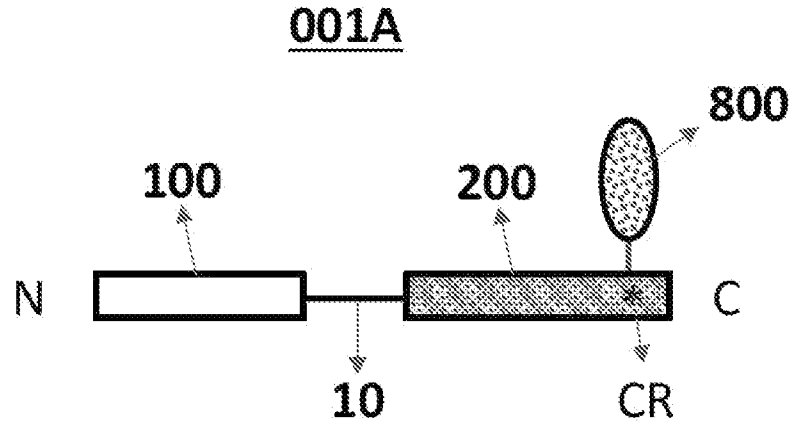


FIG. 1A

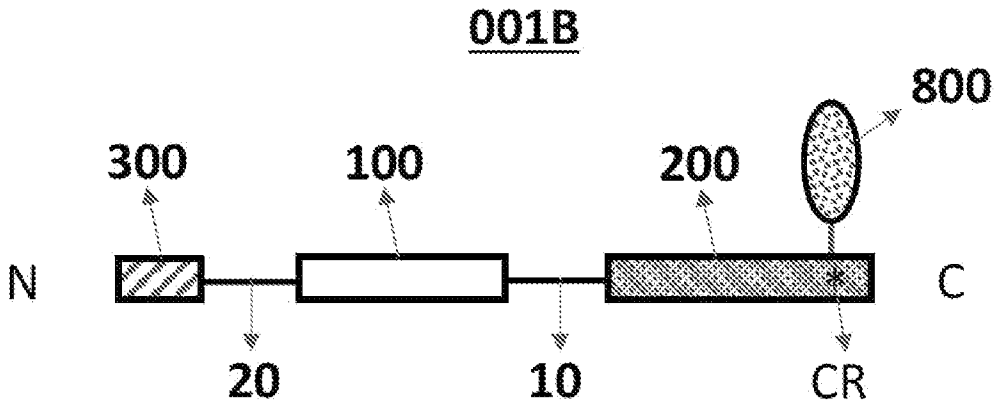


FIG. 1B

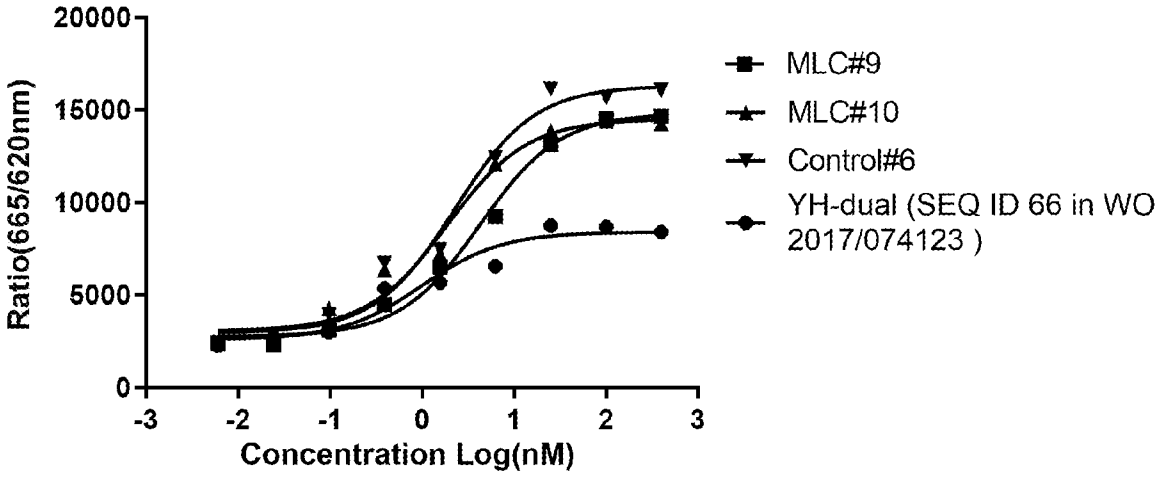


Figure 2A

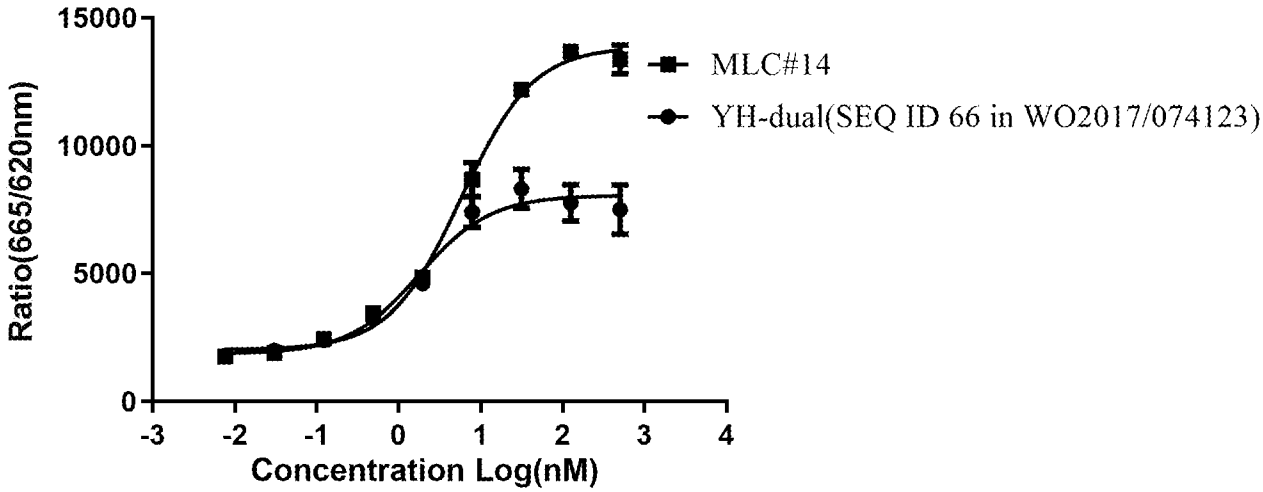


Figure 2B

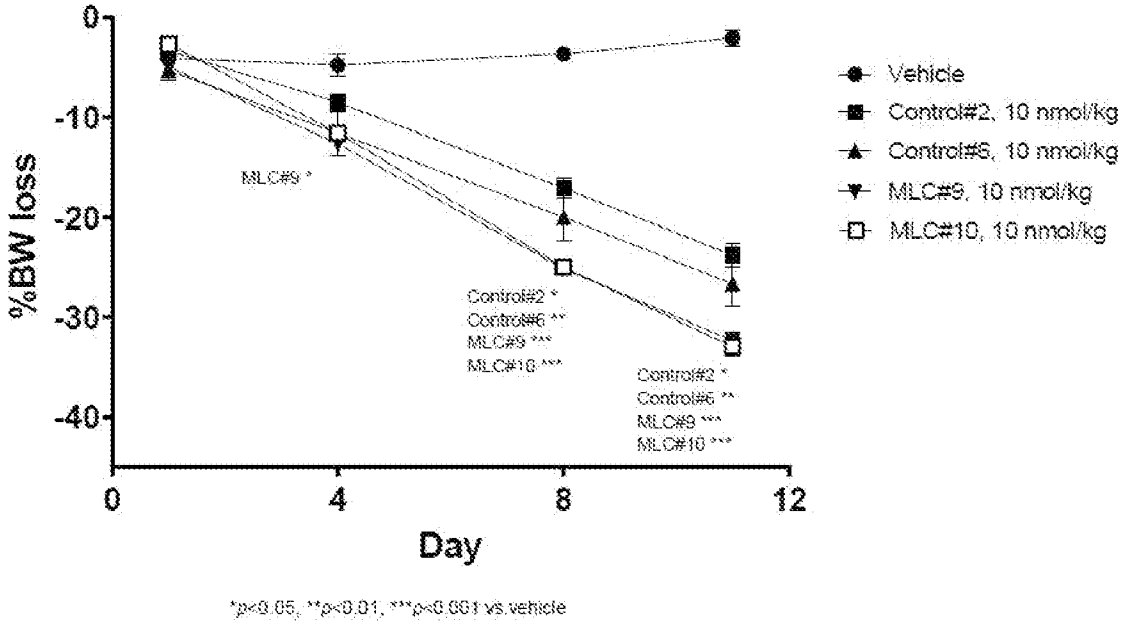


Figure 3A

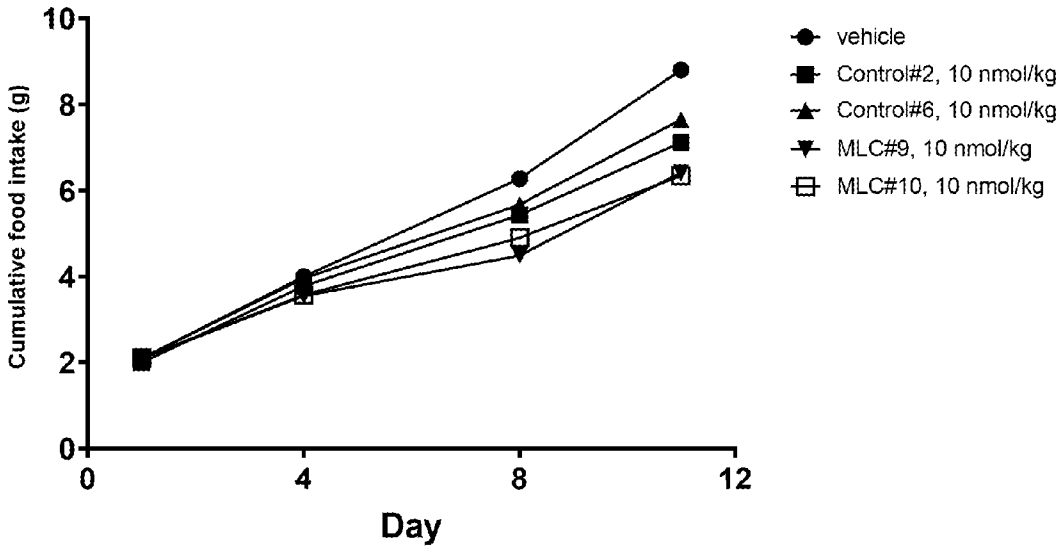


Figure 3B

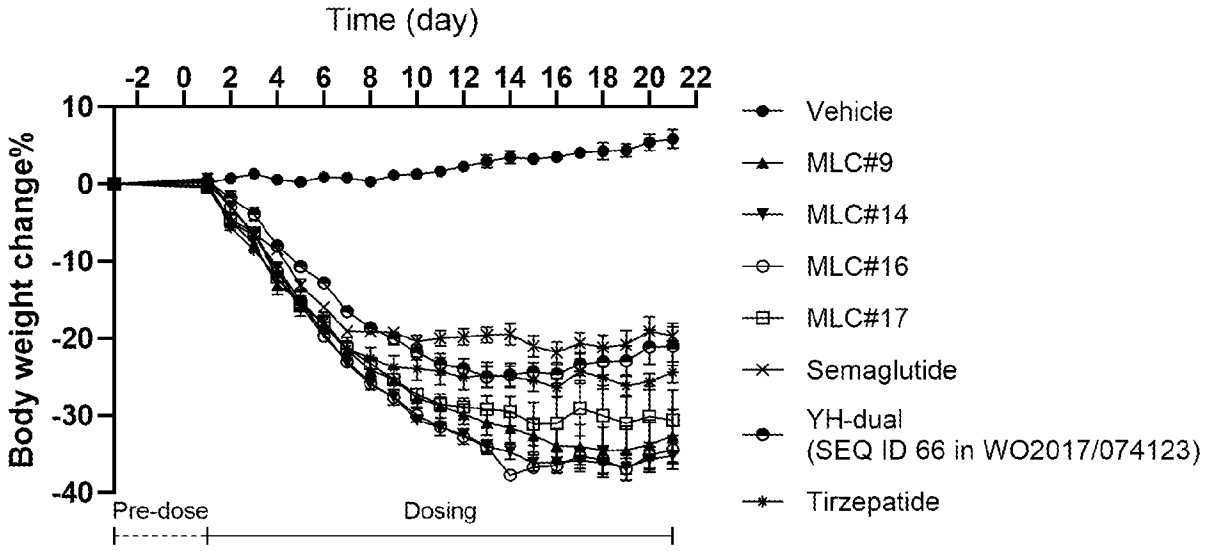


Figure 4A

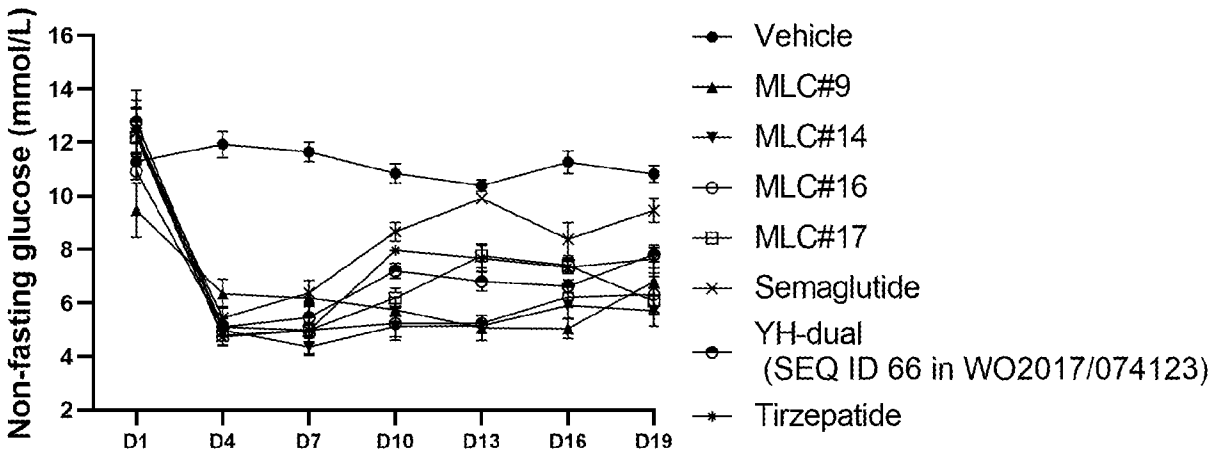


Figure 4B

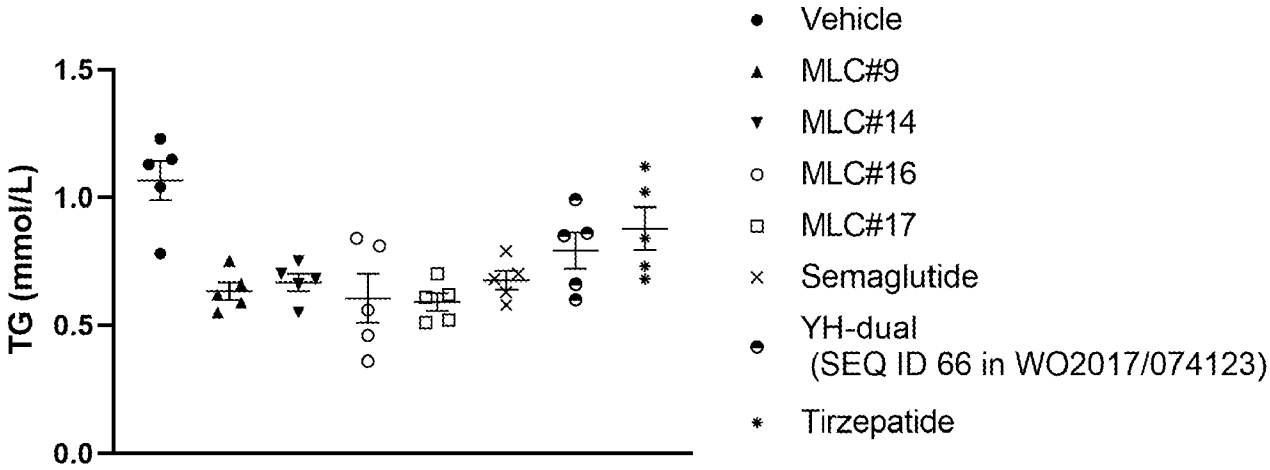


Figure 4C

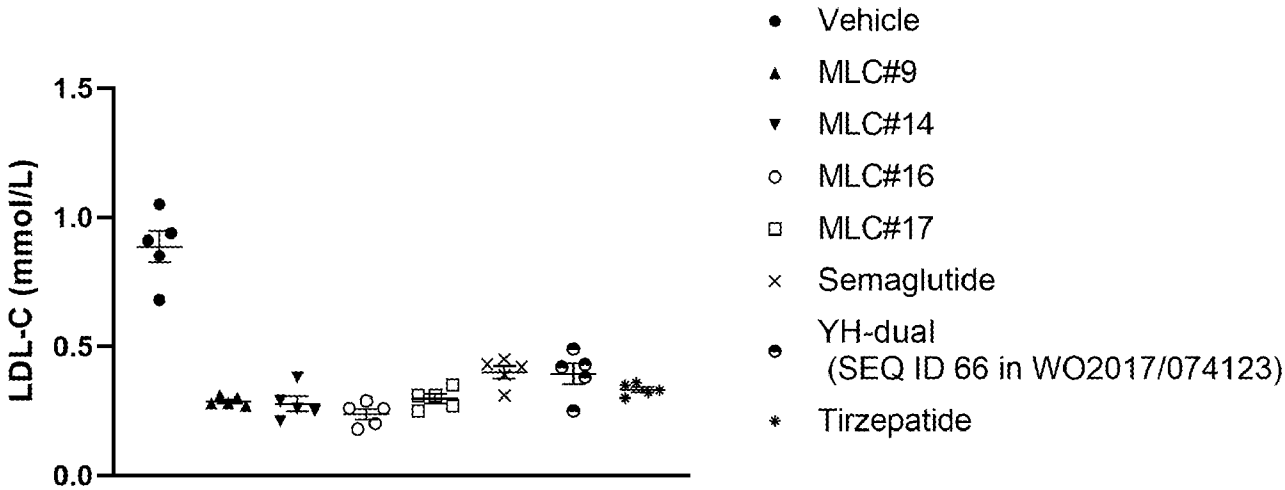


Figure 4D

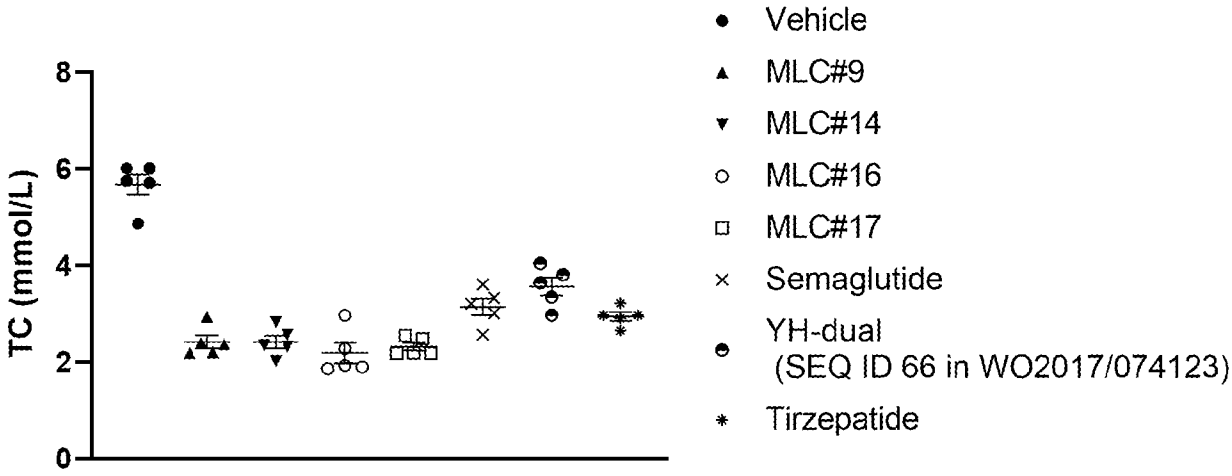


Figure 4E

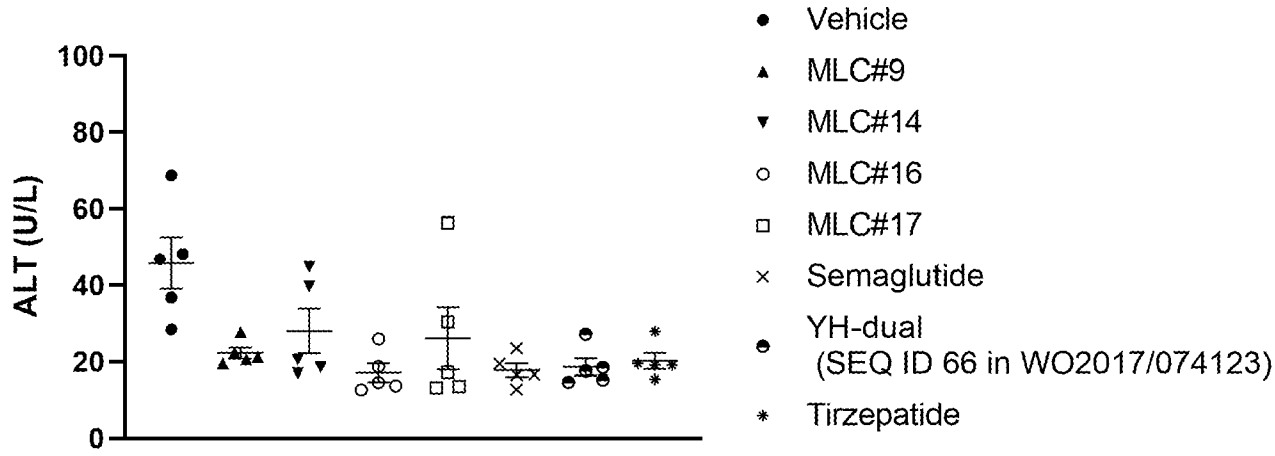


Figure 4F

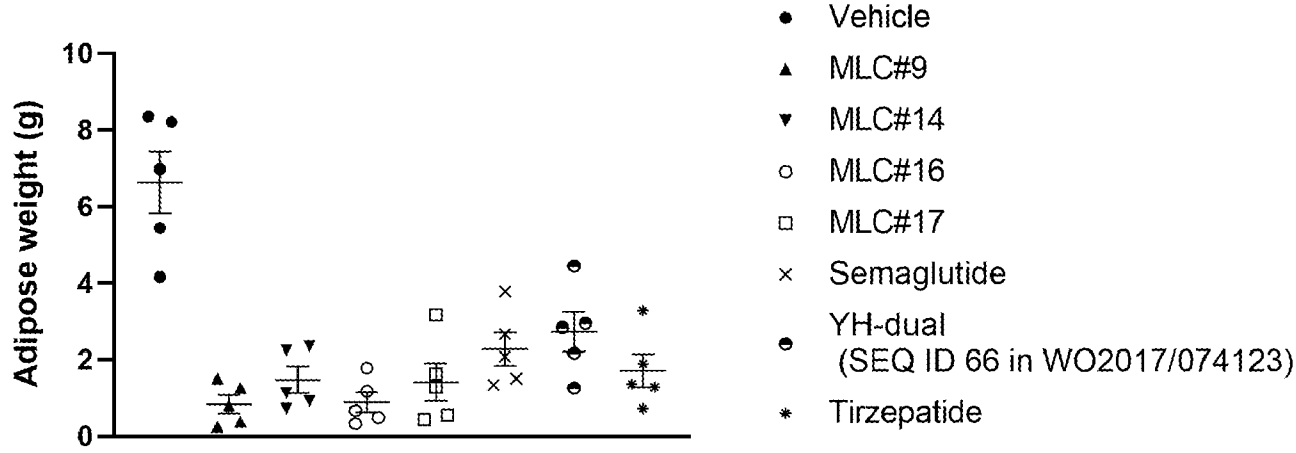


Figure 4G

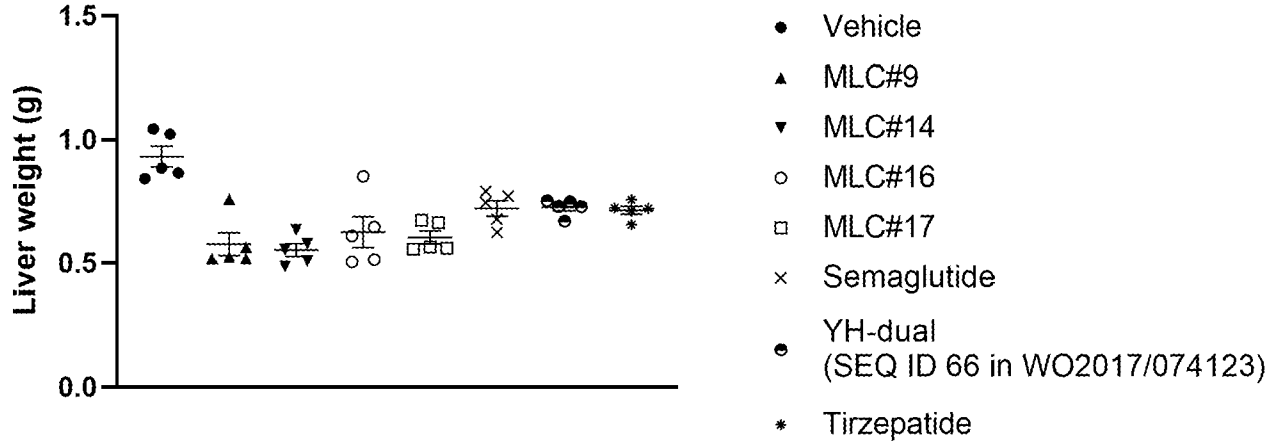


Figure 4H

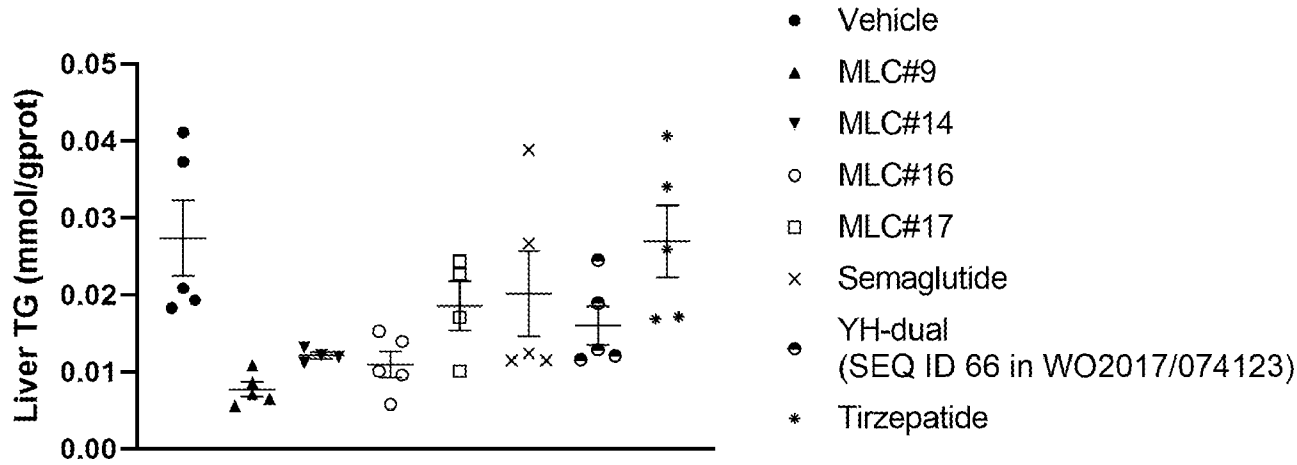


Figure 4I

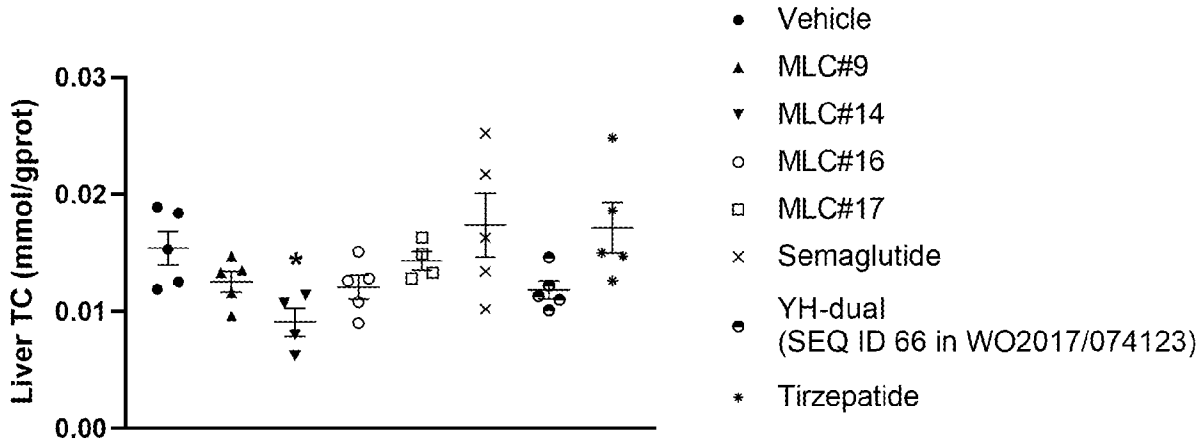


Figure 4J

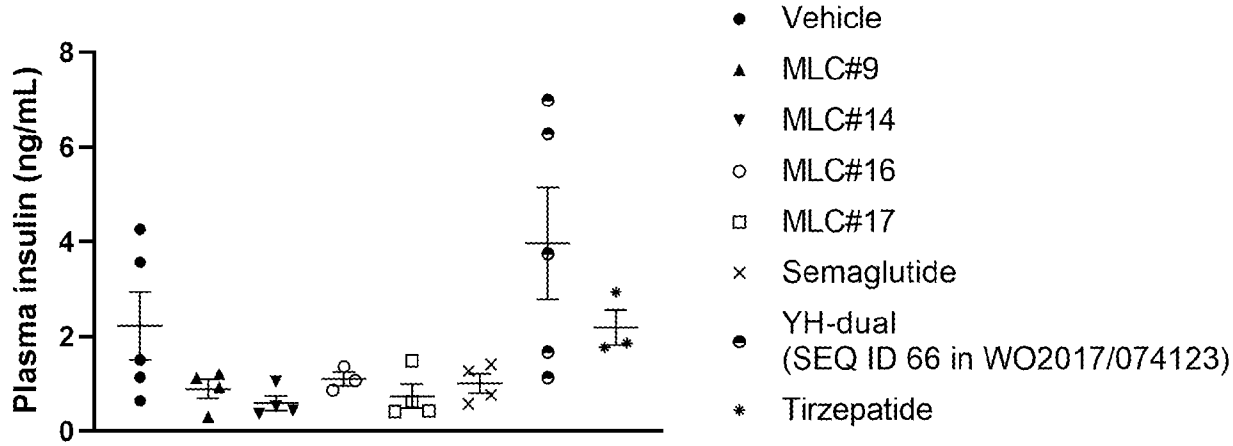


Figure 4K

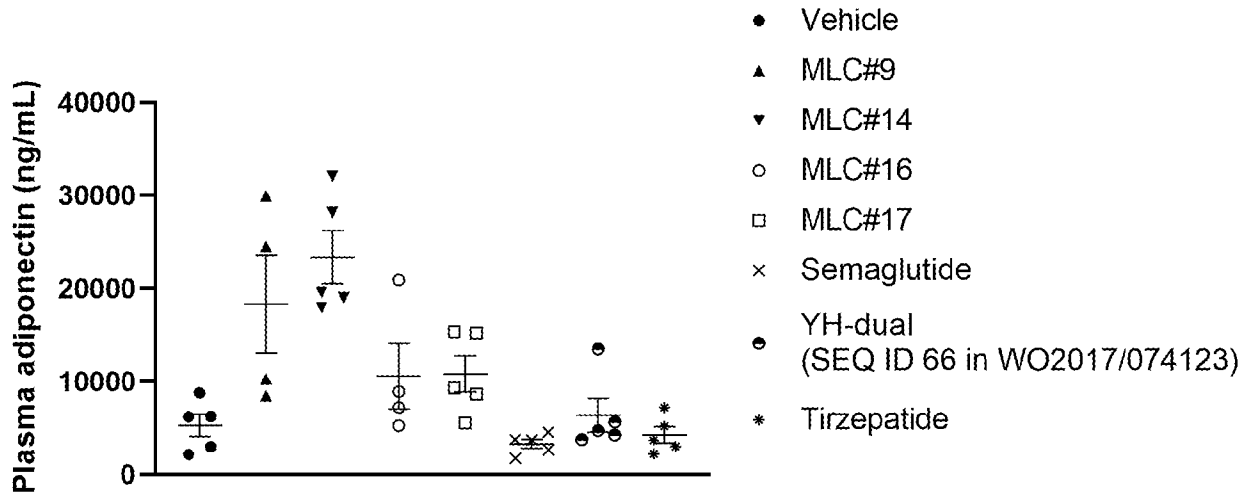


Figure 4L

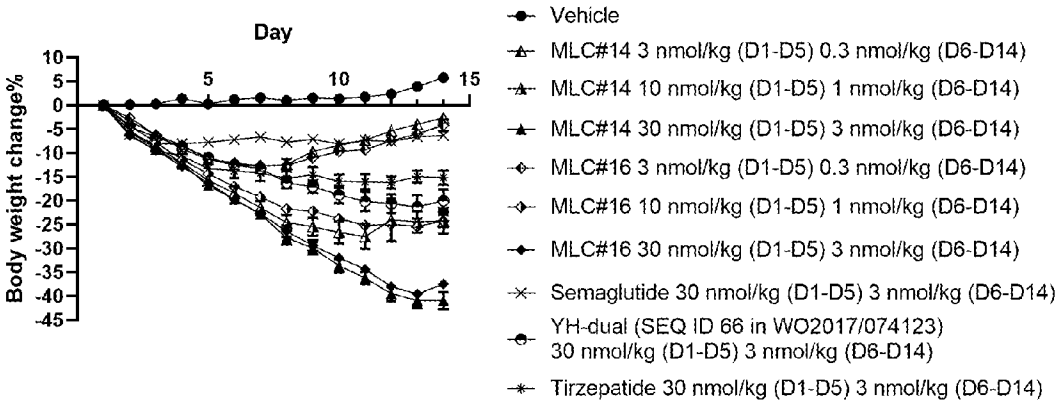


Figure 5A

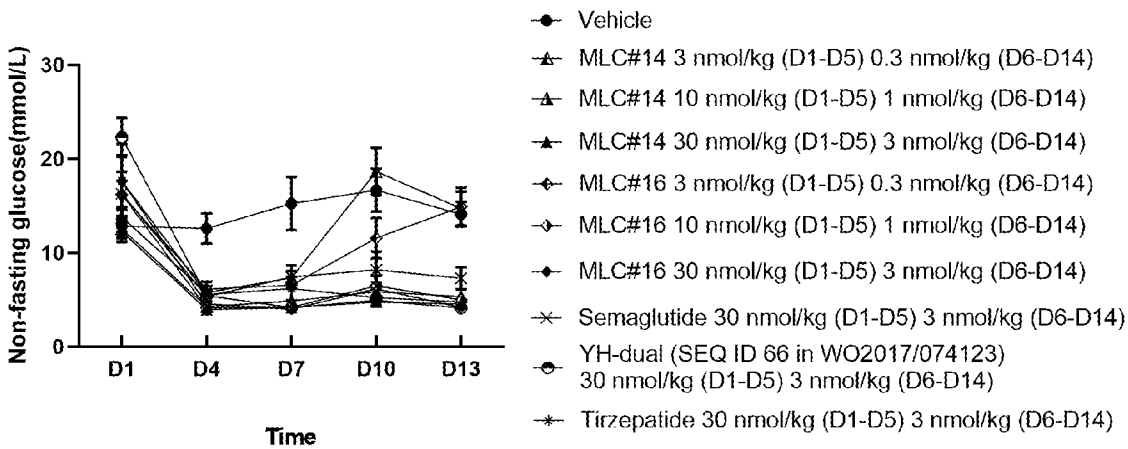


Figure 5B

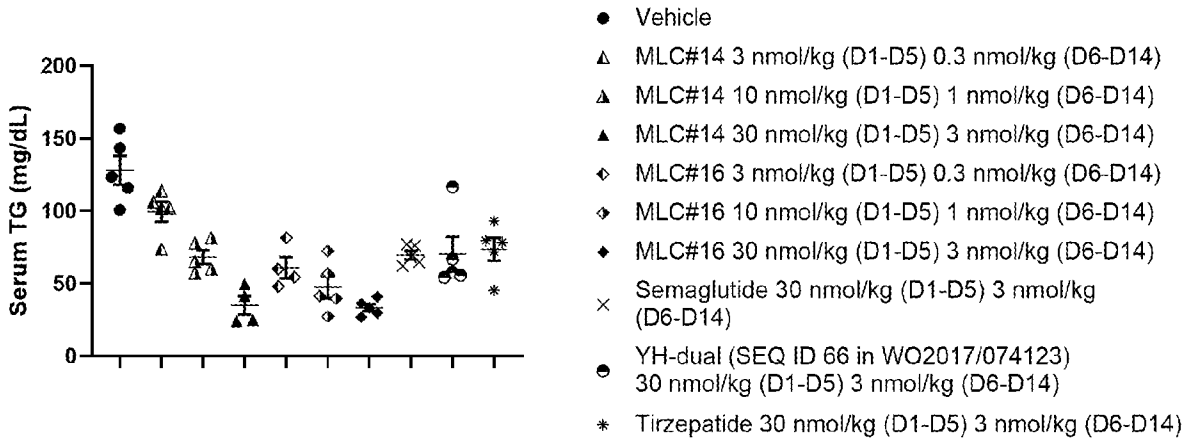


Figure 5C

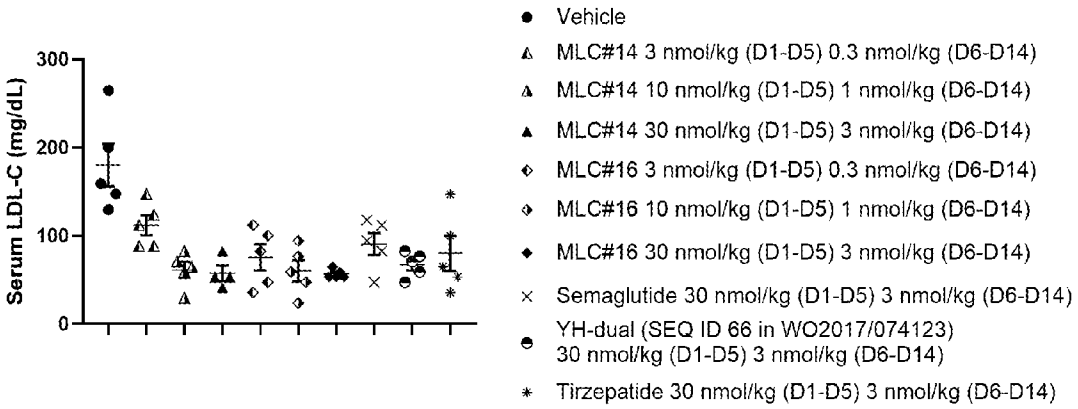


Figure 5D

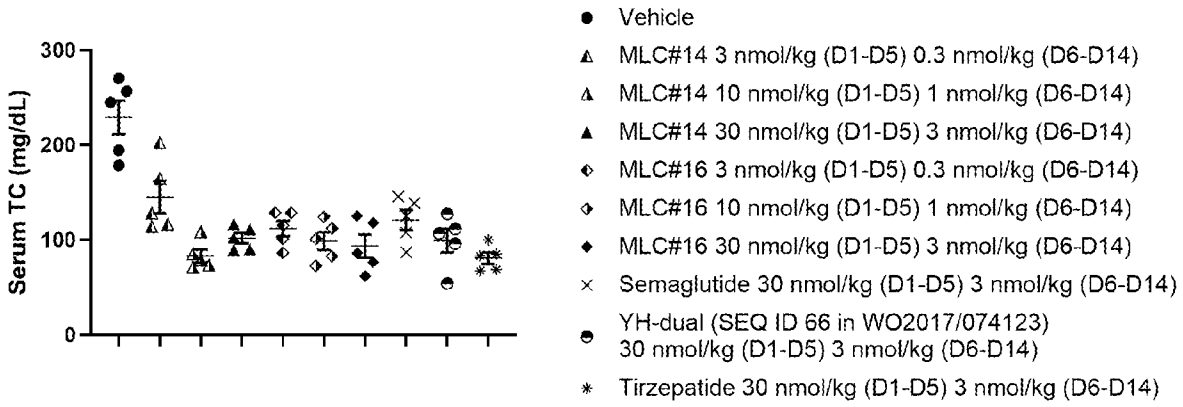


Figure 5E

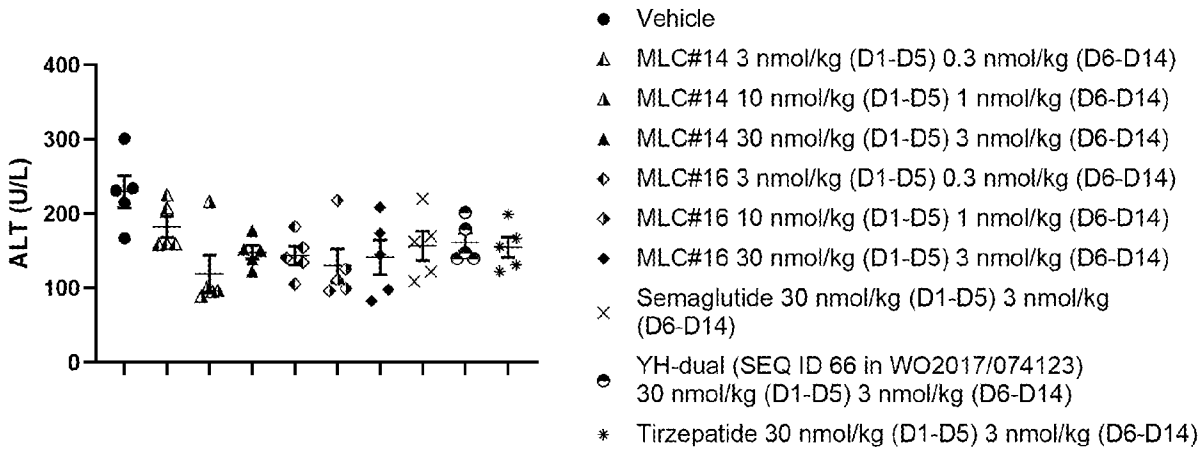


Figure 5F

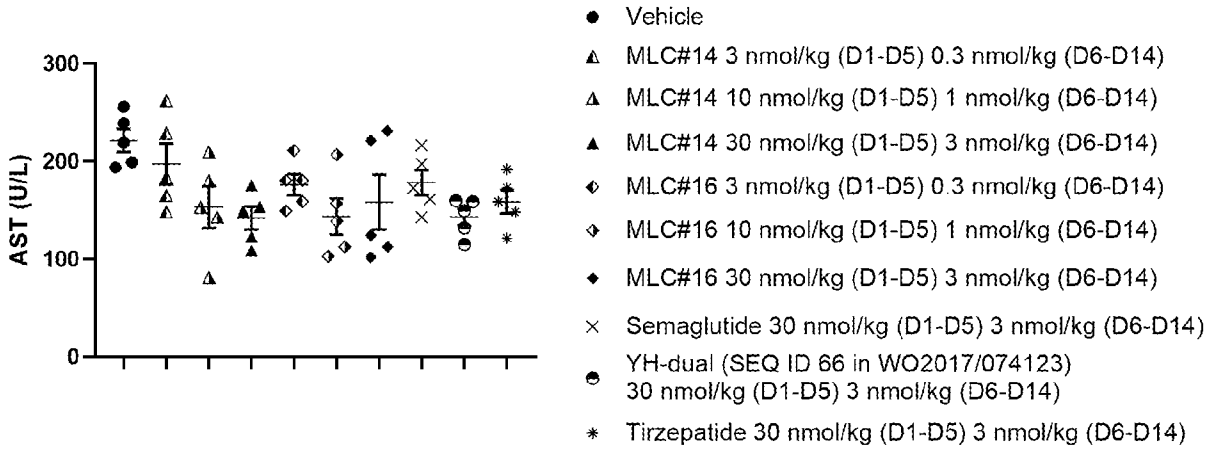


Figure 5G

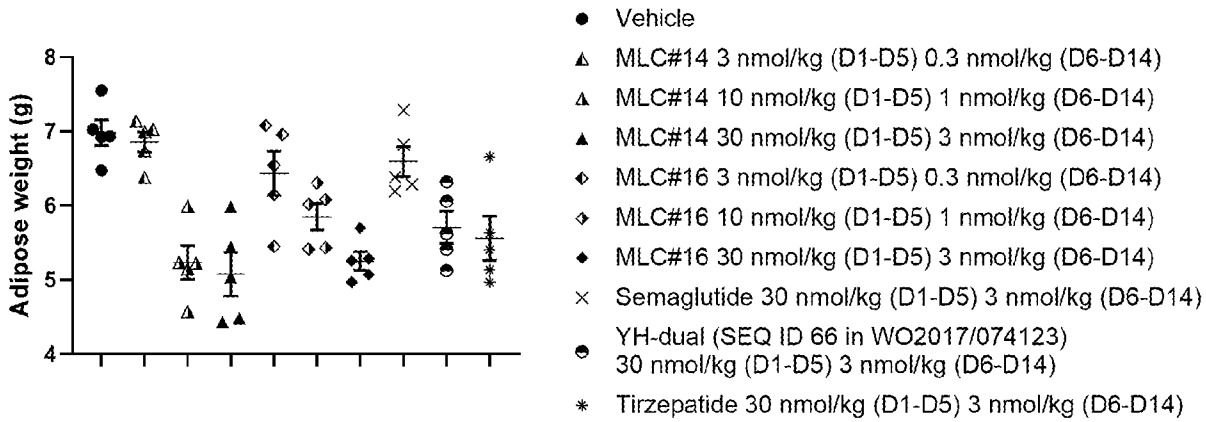


Figure 5H

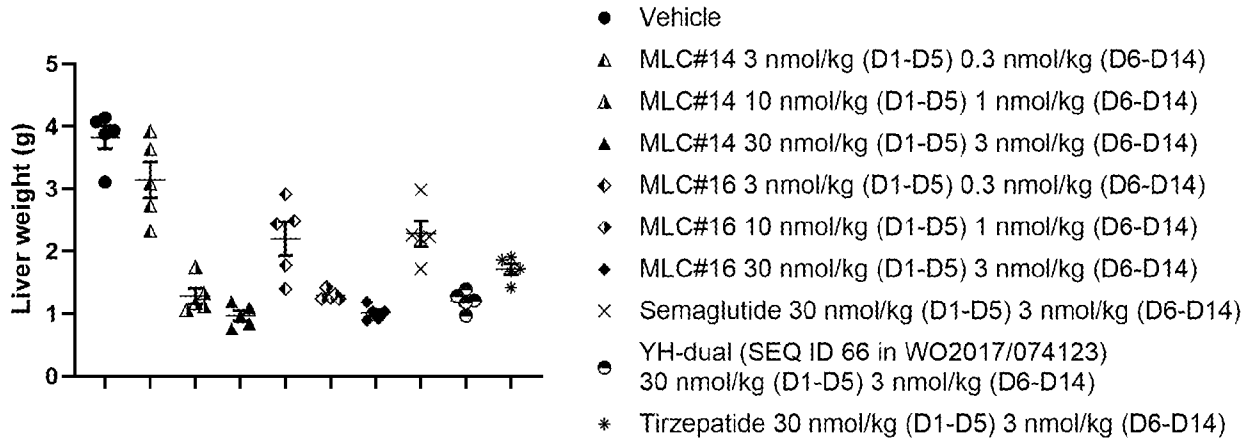


Figure 5I

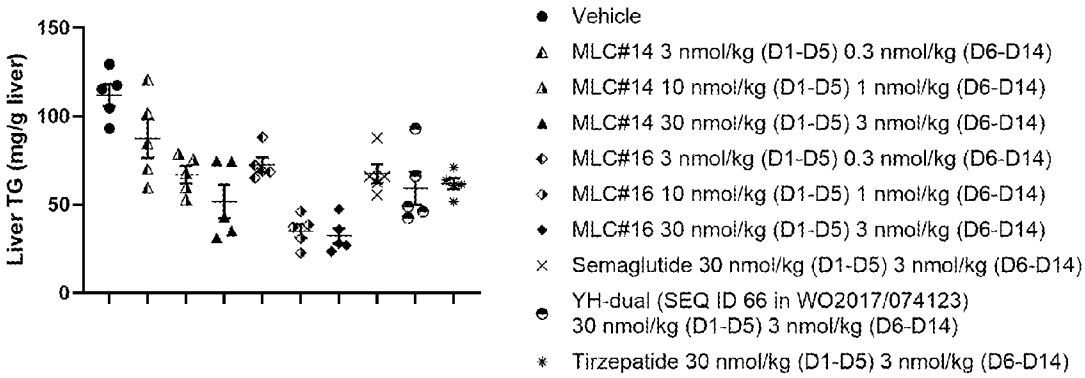


Figure 5J

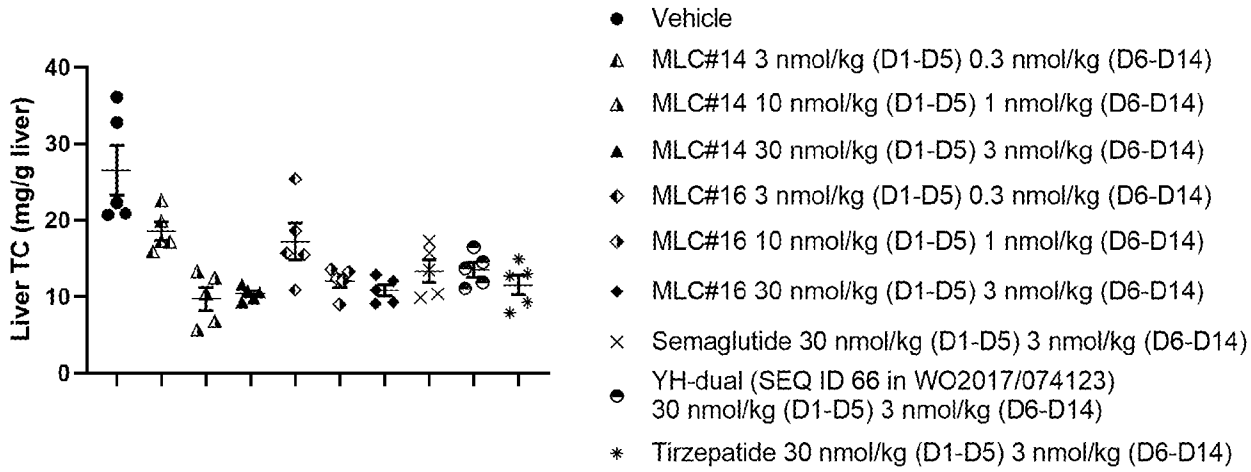


Figure 5K

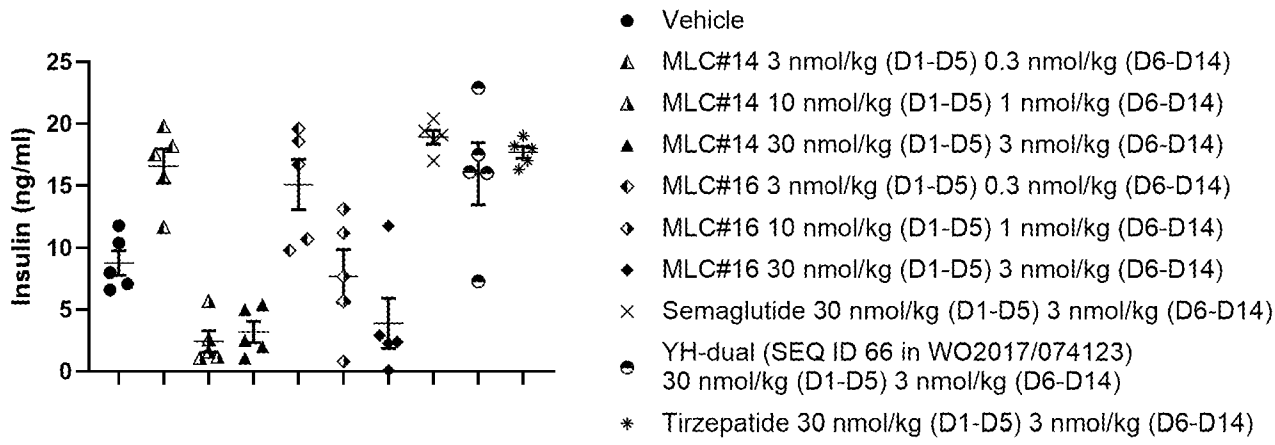


Figure 5L

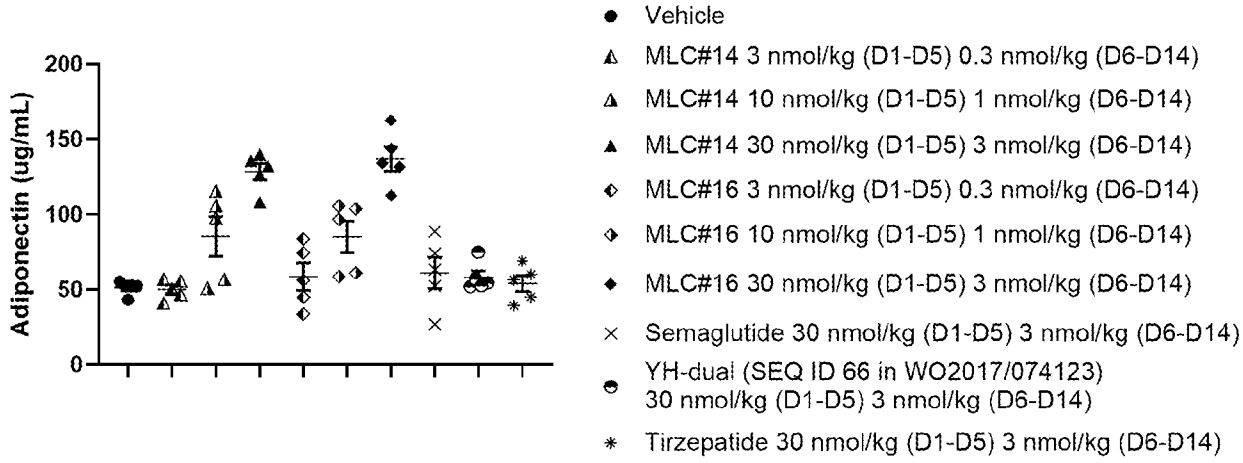


Figure 5M

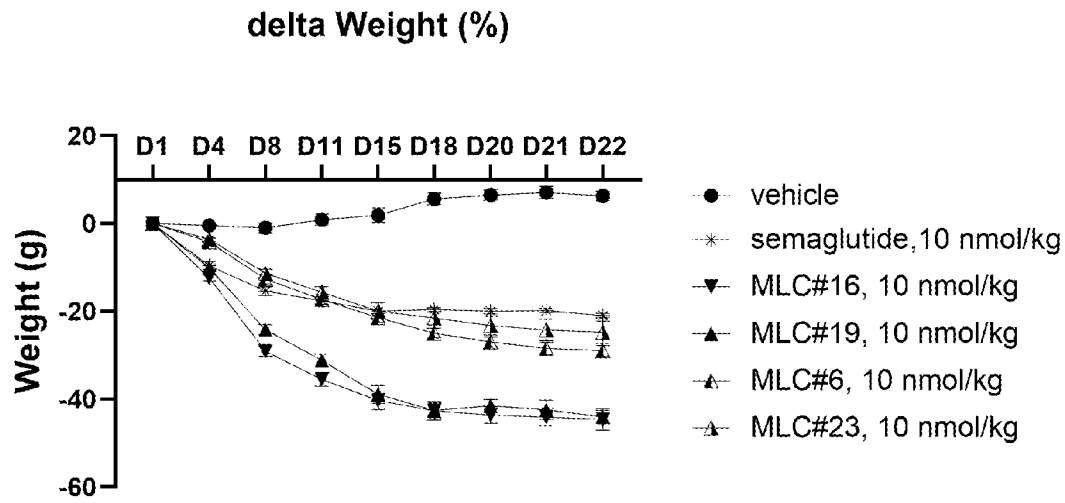


Figure 6A

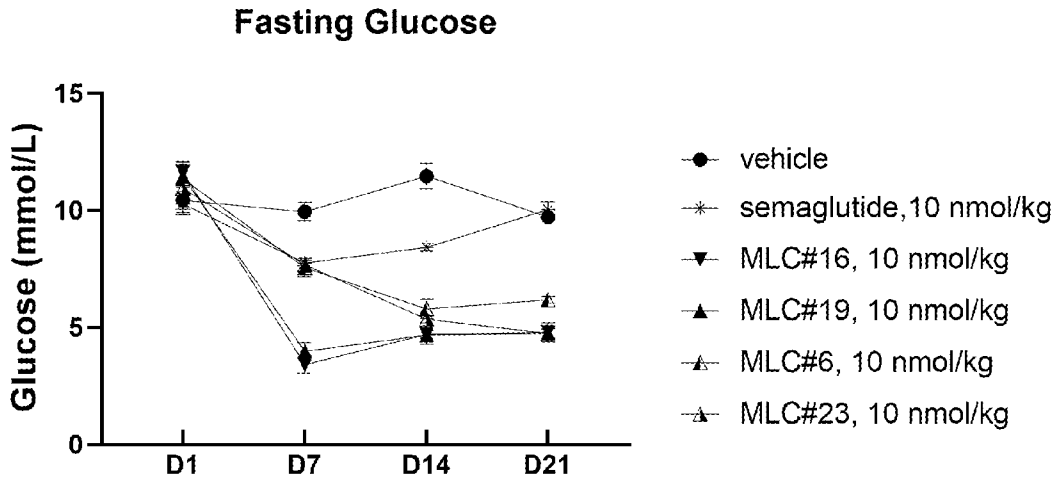


Figure 6B

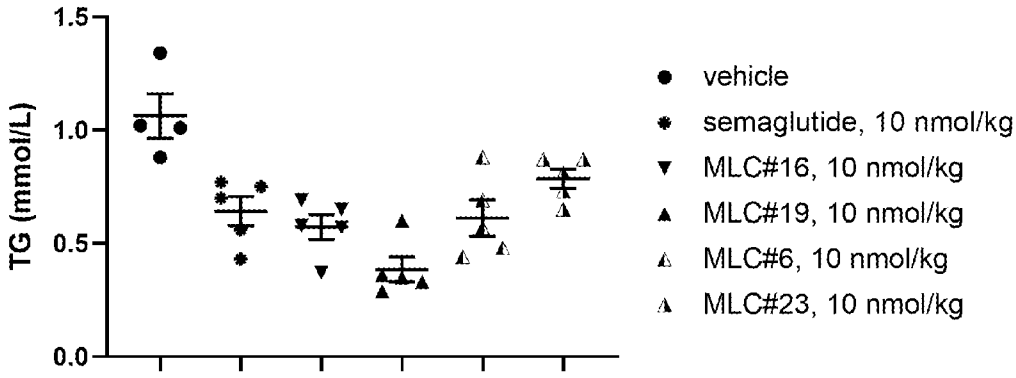


Figure 6C

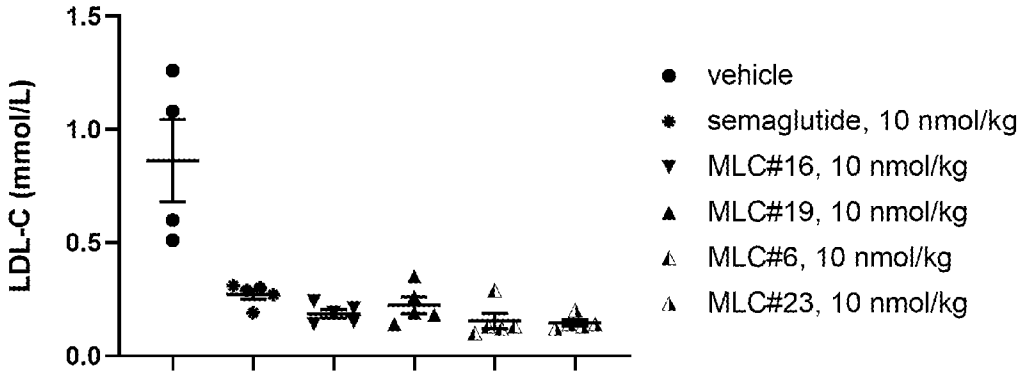


Figure 6D

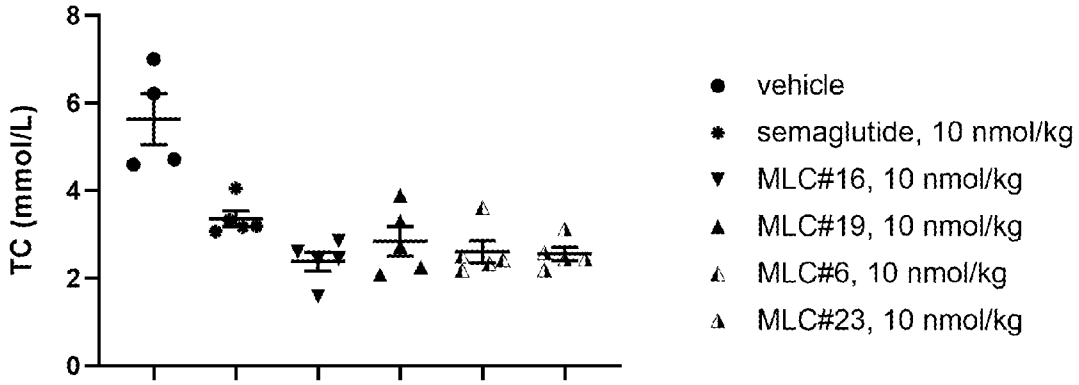


Figure 6E

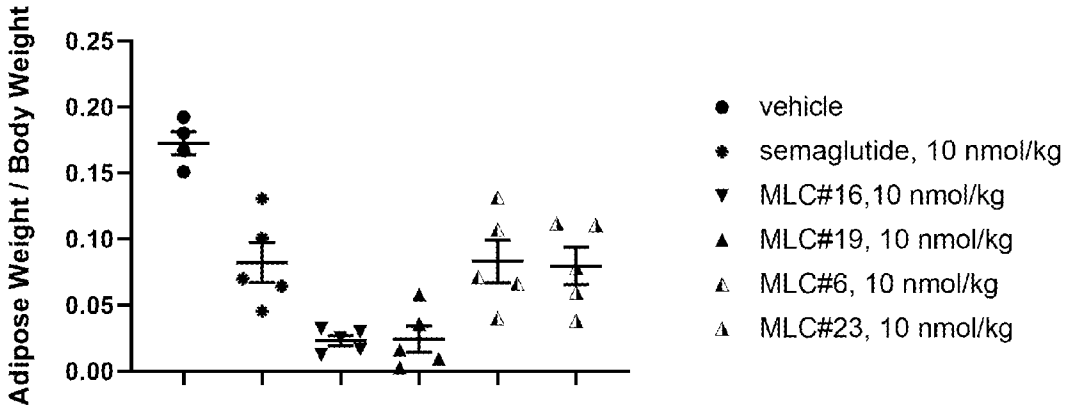


Figure 6F

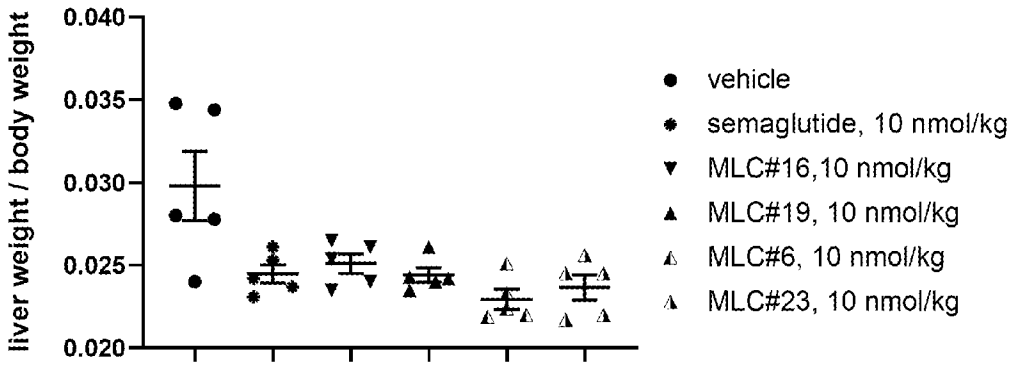


Figure 6G

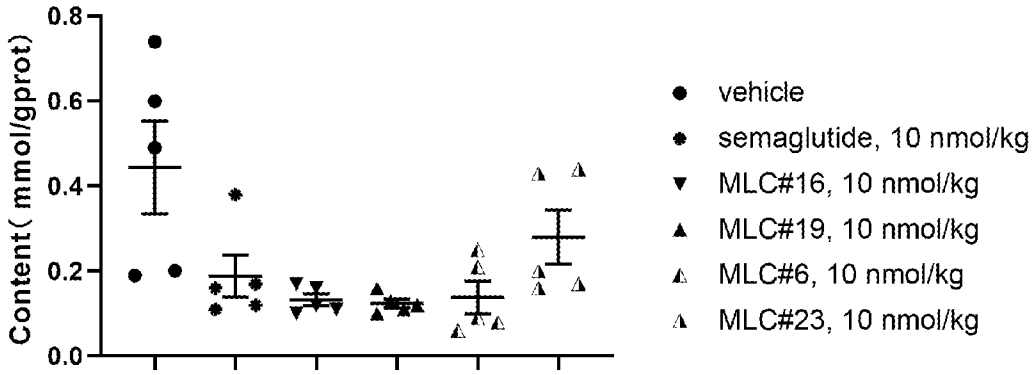

Content（mmol/gprot）
0.8
0.6
0.4
0.2
0.0
vehicle
semaglutide, 10 nmol/kg
MLC#16, 10 nmol/kg
MLC#19, 10 nmol/kg
MLC#6, 10 nmol/kg
MLC#23, 10 nmol/kg

Figure 6H

FGF21

SEQ ID NO:	FGF21 Sequence	FGF21 Features
1	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPGNKSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSMVGPSQGR SPSYAS	**WT**
2	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSMVGPSQGR SPSYAS	121Q
3	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPGNKSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPSQGRS PSYAS	168L
4	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPGNKSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSMVGPSQGR SPSY**E**S	180E
5	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPSQGRS PSY**E**S	N121Q, M168L, A180E
6	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**LC**GPSQGR SPSY**E**S	N121Q, M168L, 169C, A180E
7	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**V**C**PSQGRS PSY**E**S	N121Q, M168L, 170C, A180E
8	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VG**C**SQGR	N121Q, M168L, 171C, A180E

	SPSY**E**S	
9	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGP**C**QGR SPSY**E**S	N121Q, M168L, 172C, A180E
10	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPS**C**GRS PSY**E**S	N121Q, M168L, 173C, A180E
11	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPSQ**C**RS PSY**E**S	N121Q, M168L, 174C, A180E
12	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPSQGRS PSY**C**S	N121Q, M168L, 180C, (NOTE: A180E gone)
13	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPSQGRS PSY**EC**	N121Q, M168L, A180E, 181C
14	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VG**G**SQGR SPSY**E**S	N121Q, M168L, **171G**, A180E
15	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARF**P**PLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VG**C**SQGR SPSY**E**S	N121Q, **L137P**, M168L, 171C, A180E

16	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGP**T**QGR SPSY**E**S	N121Q, M168L, 172T, A180E
17	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPS**T**GRS PSY**E**S	N121Q, M168L, 173T, A180E
18	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**V**N**PSQGRS PSY**E**S	N121Q, M168L, 170N, A180E
19	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPSQ**N**RS PSY**E**S	N121Q, M168L, 174N, A180E
89	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPSQGRS PSYAS	121Q and 168L
90	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSMVGPSQGR SPSY**E**S	121Q and 180E
91	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPGNKSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VGPSQGRS PSY**E**S	168L and 180E
92	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VG**G**SQ**C**R SPSY**E**S	121Q, 168L, P171G, 174C, and 180E

102	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPGNKSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VG**G**SQGR SPSY**E**S	168L, P171G, and 180E
103	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VG**G**SQGR SPSYAS	121Q, P171G, and 168L
104	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSMVG**G**SQGR SPSY**E**S	121Q, P171G and 181E
105	HPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGG AADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHF DPEACSFRELLLEDGYNVYQSEAHGLPLHLPG**Q**KSPHRDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLS**L**VG**G**SQGR SPSY**E**S	N121Q, M168L, P171G, A180E

Nanobody sequences

SEQ ID	Sequence
20	SFGMS
21	SISGSGSDTLYADSVKG
22	GGSLSR
23	EVQLVESGGGLVQPGNSLRLSCAASGFTFS**SFGMS**W VRQAPGKGLEWVS**SISGSGSDTLYADSVKG**RFTISRD NAKTTLYLQMNSLRPEDTAVYYCTI**GGSLSR**SSQGTL VTVSS
24	**A**EVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMS WVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR

	DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGT LVTVSS
25	**S**EVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMS WVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGT LVTVSS
26	**AG**EVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMS WVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGT LVTVSS
27	**SG**EVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMS WVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGT LVTVSS

GLP-1 sequences

SEQ ID NO:	GLP-1 Sequence	GLP-1 Features
28	**HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG**	GLP1-WT
29	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGG	8G, 22E, 36G
30	EGTFTSDVSSYLEEQAAKEFIAWLVKGGG	del7-8 (inactive), 22E, 36G
31	HGEGTFTSDVSSYLEEQAA**K**EFIAWLVRGGG	8G, 22E, 34R,36G
32	HGEGTFTSDVSSYLEEQAAREFIAWLVRGGG	8G, 22E, 26R, 34R,36G
33	HGEGTFTSDVSSYLEEQAAREFIAWLVRG**K**G	8G, 22E, 26R, 34R,36K
34	HGEGTFTSDVSSYLEEQAAREFIAWLVRGGG	8G, 22E, 26R,36G

First and second linker

SEQ ID NO:	1st/2nd Linker Sequence	Features
35	GAQP	(GAQP)1
36	GQAP	
37	GPAQ	
38	GPQA	
39	GSQP	
40	GASP	
41	GPAS	
42	GPSA	
43	GGGS	
44	GSGS	
45	GGGGS	
46	GSAPGSPAGSPTGSAPGSPA	
47	GGGGSGGGGSGGGGSGGGGS	(G4S)4
48	GGGGSGGGS	G4SG3S
49	GAQPGAQP	(GAQP)2
50	GAQPGAQPGAQPGAQPGAQP	(GAQP)5
51	GAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQP	(GAQP)10
52	GAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQP GAQPGAQPGAQP	(GAQP)14
53	GQEPGAQPGQEPGAQPGQEPGAQPGQEPGAQPGQEPGAQPGQEPG AQP	(GQEPGAQP)6
54	GGGGGSGGGSGGGGS	G5SG3SG4S
55	GQEP	
56	GEQP	
57	GPQE	
58	GPEQ	
59	GSEP	
60	GESP	
61	GPSE	
62	GPES	
106	GAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQP GAQPGAQPGGGGS	(GAQP)13(G4S)

Nanobody-FGF21 sequences

SEQ ID NO:	Sequence
63	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPHPIPDSSPLLQFGGQ VRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLC QRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGP ARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRSPSYES
64	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPHPIPDSSPLLQFGGQ VRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLC QRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGP ARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQGRSPSYES
65	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPGAQPGAQPGAQPGAQPGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIRED GTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELL LEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPP DVGSSDPLSLVGGSQGRSPSYES
66	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPES LLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAH GLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGS QGRSPSYES
67	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQL KALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPL HLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQGR SPSYES
68	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQL KALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPL HLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRS PSYES
69	MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIS

	KAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP VLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGS GGGSGGGGSHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQS PESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRERLLEDGYNVYQS EAHGLPLHLPGNKSPHRDPAPRGPARFLPLPGLPPAPPEPPGILAPQPPDVGSSDPLSM VGGSQGRSPSYES
93	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQL KALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPL HLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQCRS PSYES
99	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQL KALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPL HLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGPSQGRS PSYES
100	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG GGGSHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQ LKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLP LHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGR SPSYES
101	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG GGGSHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQ LKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLP LHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGPSQGR SPSYES
107	SGEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG GGGSHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQ LKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLP LHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQGR SPSYES

GLP-1-Nanobody-FGF21 sequences

SEQ ID NO:	Sequence
70	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGGGGSGGGGSGGGGS GGGGSEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAP GKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLQMNSLRPE DTAVYYCTIGGSLSRSSQGTLVTVSSGGGGSGGGSHPIPDSSPLLQFG GQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPG VIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSE AHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPP DVGSSDPLSLVGGSQGRSPSYES
71	EGTFTSDVSSYLEEQAAKEFIAWLVKGGGGGGGSGGGGSGGGGSGG GGSEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGK GLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTA VYYCTIGGSLSRSSQGTLVTVSSGGGGSGGGSHPIPDSSPLLQFGGQV RQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQI LGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHG LPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGS SDPLSLVGGSQGRSPSYES
72	EGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPEVQL VESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTL YADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPHPIPDSSPLLQF GGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILG VKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKS PHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQGRSPSYE S
73	EGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPEVQL VESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTL YADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG QEPGAQPGQEPGAQPGQEPGAQPGQEPGAQPGQEPGAQPGQEPGAQPHPIP DSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALK PGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPL HLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGS QGRSPSYES
74	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGAQPEV QLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTV SSGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPHPIPDSSPLL QFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQI LGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQ KSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRSPS YES
75	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPEV QLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS

	DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTV SSGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPHPIPDSSPLL QFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQI LGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQ KSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQGRSP SYES
76	EGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPEVQL VESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTL YADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPGAQPGAQPGAQPGAQPGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLE IREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPE ACSFRELLLEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALP EPPGILAPQPPDVGSSDPLSLVGGSQGRSPSYES
77	EGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGAQPEVQL VESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTL YADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQ SPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYN VYQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVG SSDPLSLVGGSQGRSPSYES
78	EGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGAQPEVQL VESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTL YADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSG AQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESL LQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQS EAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDP LSLVGGSQGRSPSYES
79	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGA QPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMS WVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLQMNSLRPED TAVYYCTIGGSLSRSSQGTLVTVSSGAQPGAQPGAQPGAQPGAQPGAQPGAQP GAQPGAQPGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVG GAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELL LEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAP QPPDVGSSDPLSLVGGSQGRSPSYES
80	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPEV QLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGS DTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTV SSGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQS PESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNV YQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGS SDPLSLVGGSQGRSPSYES

81	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRSPSYES
82	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRSPSYES
83	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRSPSYES
84	EGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQGRSPSYES
85	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRSPSYES
86	EGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGNSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKS

	PHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRSPSYE S
87	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGA QPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGN SLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLL QFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQI LGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQ KSPHRDPAPRGPARFPPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRSP SYES
88	EGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGAQP GAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGNSLR LSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISRDN AKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLLQF GGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQILG VKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQKS PHRDPAPRGPARFPPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGRSPSY ES
94	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGA QPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGN SLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLL QFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQI LGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQ KSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQCRSPS YES
95	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGA QPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGN SLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLL QFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQI LGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQ KSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGPSQGRSPS YES
96	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGA QPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGN SLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGGGGSHPIPDSSP LLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVI QILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLP GQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGR SPSYES

97	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGA QPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGGGGSEVQLVESGGGLVQPG NSLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTIS RDNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGGGGSHPIPDSS PLLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGV IQILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLP GQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGCSQGR SPSYES
98	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGA QPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGN SLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGGGGSHPIPDSSP LLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVI QILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLP GQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGPSQGR SPSYES
108	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGA QPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGN SLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGAQPHPIPDSSPLL QFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVIQI LGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLPGQ KSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQGRSP SYES
109	HGEGTFTSDVSSYLEEQAAKEFIAWLVKGGGGAQPGAQPGAQPGAQPGAQPGA QPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPEVQLVESGGGLVQPGN SLRLSCAASGFTFSSFGMSWVRQAPGKGLEWVSSISGSGSDTLYADSVKGRFTISR DNAKTTLYLQMNSLRPEDTAVYYCTIGGSLSRSSQGTLVTVSSGGGGSHPIPDSSP LLQFGGQVRQRYLYTDDAQQTEAHLEIREDGTVGGAADQSPESLLQLKALKPGVI QILGVKTSRFLCQRPDGALYGSLHFDPEACSFRELLLEDGYNVYQSEAHGLPLHLP GQKSPHRDPAPRGPARFLPLPGLPPALPEPPGILAPQPPDVGSSDPLSLVGGSQGR SPSYES

Figure 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2022/105686

A. CLASSIFICATION OF SUBJECT MATTER

C07K 19/00(2006.01)i; C07K 14/435(2006.01)i; A61K 39/395(2006.01)i; C12N 15/62(2006.01)i; C12N 15/63(2006.01)i; A61P 1/16(2006.01)i; A61P 3/04(2006.01)i; A61P 3/06(2006.01)i; A61P 3/10(2006.01)i; A61P 13/12(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K, A61K, C12N, A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, VEN, CNKI, CNTXT, USTXT, EPTXT, PUBMED, ISI WEB, NCBI, EMBL-EBI, STNext, BEIJING QL BIOPHARMACEUTICAL, nanobody, serum albumin, HSA, VHH, FGF21, fibroblast growth factor 21, glucagon-like peptide 1, GLP-1, mutation, variant, fusion, fused, linker, half-life, conjugate, SEQ ID NOs:1, 6-13, 16-29,31-52, 55-68,70,74-75,79-83,85,92-101,107-109

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	AU 2016203044 A1 (BRISTOL-MYERS SQUIBB COMPANY) 16 June 2016 (2016-06-16) description, page 2 lines 1-7, page 5 lines 8-27, page 12 lines 26-32, page 40 line 19 - page 41 line 35	1-61
Y	US 2009028880 A1 (ABLYNX N. V.) 29 January 2009 (2009-01-29) claims 1-24, description, paragraph 21	1-61
Y	WO 2017093465 A1 (SANOFI) 08 June 2017 (2017-06-08) claims 1-27	1-61
Y	US 2014228546 A1 (Ablynx N.V.) 14 August 2014 (2014-08-14) claims 1-21	1-61
Y	CN 110028587 A (ANYUAN BIOTECHNOLOGY SHANGHAI CO., LTD.) 19 July 2019 (2019-07-19) claims 1-20	1-61
Y	AU 2013202856 A1 (Ablynx N.V.) 02 May 2013 (2013-05-02) claims 1-34	1-61

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 September 2022	**10 October 2022**
Name and mailing address of the ISA/CN **National Intellectual Property Administration, PRC** **6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088, China** Facsimile No. **(86-10)62019451**	Authorized officer **WANG,Yao** Telephone No. **86-(10)-53961926**

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	OMAR, B.A. et al. "Fibroblast Growth Factor 21 (FGF21) and Glucagon-Like Peptide 1 Contribute to Diabetes Resistance in Glucagon Receptor-Deficient Mice." *DIABETES*, Vol. 63, No. 1, 23 September 2013 (2013-09-23), pages 101-110	1-61

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2022/105686

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
   a. ☑ forming part of the international application as filed:
      ☑ in the form of an Annex C/ST.25 text file.
      ☐ on paper or in the form of an image file.
   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

   [1] The sequence listing actually submitted is in the form of an Annex C/ST.26 text file.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2022/105686

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☑ Claims Nos.: **52-53**
because they relate to subject matter not required to be searched by this Authority, namely:

[1] Claims 52-53 relate to a method of preventing or treating a metabolic disorder. The subject-matter of claims 52-53 relates to the treatment of human or animal bodies, and therefore does not warrant an international search according to the criteria set out in PCT Rule 39.1(iv). An international search has been carried out and based on the manufacturing of a medicament.

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2022/105686

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
AU	2016203044	A1	16 June 2016	None			
US	2009028880	A1	29 January 2009	JP	2017186318	A	12 October 2017
				LT	2949668	T	11 November 2019
				SI	1888641	T1	31 May 2012
				JP	2008539772	A	20 November 2008
				ZA	201100385	B	26 September 2012
				ZA	200709867	B	29 October 2008
				EP	3613767	A1	26 February 2020
				AT	537188	T	15 December 2011
				MX	2007014359	A	06 February 2008
				RU	2017133794	A	28 March 2019
				HU	E045710	T2	28 January 2020
				ES	2384164	T3	02 July 2012
				JP	2012167112	A	06 September 2012
				HK	1113685	A1	10 October 2008
				HK	1215953	A1	30 September 2016
				EP	1888640	A2	20 February 2008
				US	2020079844	A1	12 March 2020
				PL	2949668	T3	31 March 2020
				CR	9519	A	17 October 2008
				AU	2006249144	A1	23 November 2006
				PT	1888641	E	19 March 2012
				CA	2608770	A1	23 November 2006
				IL	186941	D0	09 February 2008
				ES	2778123	T3	07 August 2020
				EP	2365000	A2	14 September 2011
				RU	2007142444	A	27 June 2009
				KR	20080011395	A	04 February 2008
				IL	250946	D0	30 April 2017
				US	2015353635	A1	10 December 2015
				EP	2949668	A1	02 December 2015
				NZ	562695	A	27 May 2011
				SI	1888640	T1	31 July 2012
				BR	PI0609224	A2	18 October 2011
				DK	1888640	T3	18 June 2012
				US	2014335101	A1	13 November 2014
				RU	2012126722	A	10 January 2014
				PL	1888640	T3	31 August 2012
				CN	103254309	A	21 August 2013
				EP	2479191	A2	25 July 2012
				NO	20161028	A1	19 November 2007
				NI	200700291	A	12 May 2008
				TW	200716671	A	01 May 2007
				TW	201321410	A	01 June 2013
				BR	122013001996	B1	01 February 2022
				MX	342271	B	21 September 2016
				CN	101248087	A	20 August 2008
				DK	2949668	T3	11 November 2019
				WO	2006122787	A1	23 November 2006
				SI	2949668	T1	29 November 2019

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2022/105686

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2017093465	A1	08 June 2017	MX	2018005387	A	14 August 2018
				KR	20180083938	A	23 July 2018
				EP	3383890	A1	10 October 2018
				TW	201731867	A	16 September 2017
				IL	259675	A	31 July 2018
				US	2022227825	A1	21 July 2022
				EP	3693382	A1	12 August 2020
				CN	108602869	A	28 September 2018
				RU	2018123704	A	14 January 2020
				JP	2022068275	A	09 May 2022
				BR	112018011002	A2	04 December 2018
				US	2018371041	A1	27 December 2018
				SG	11201804577R	A	28 June 2018
				JP	2019506138	A	07 March 2019
				AU	2016361912	A1	19 July 2018
				CA	3006689	A1	08 June 2017
US	2014228546	A1	14 August 2014	IN	414CHN2014	A	03 April 2015
				US	2017210789	A1	27 July 2017
				AU	2012271974	A1	12 December 2013
				LT	2723771	T	10 December 2019
				PL	2723771	T3	30 April 2020
				DK	2723771	T3	02 December 2019
				HU	E047238	T2	28 April 2020
				JP	2021088585	A	10 June 2021
				HR	P20192160	T1	21 February 2020
				CN	103619878	A	05 March 2014
				ES	2759936	T3	12 May 2020
				SI	2723771	T1	31 December 2019
				EP	2723771	A1	30 April 2014
				PT	2723771	T	11 December 2019
				EP	3466972	A1	10 April 2019
				HR	P20170535	T1	16 June 2017
				JP	2014520129	A	21 August 2014
				JP	2018127478	A	16 August 2018
				WO	2012175400	A1	27 December 2012
				CA	2839779	A1	27 December 2012
				CN	106046168	A	26 October 2016
CN	110028587	A	19 July 2019		None		
AU	2013202856	A1	02 May 2013		None		