



US012383623B2

(12) **United States Patent**
Zhang et al.

(10) **Patent No.:** US 12,383,623 B2
(45) **Date of Patent:** Aug. 12, 2025

(54) **LIQUID PHARMACEUTICAL COMPOSITIONS OF POLYPEPTIDE CONJUGATES AND METHODS OF USES THEREOF**

(71) Applicant: **BEIJING QL BIOPHARMACEUTICAL CO., LTD.**, Beijing (CN)

(72) Inventors: **Yuanyuan Zhang**, Beijing (CN); **Ting Chen**, Beijing (CN); **Bo Wu**, Beijing (CN)

(73) Assignee: **BEIJING QL BIOPHARMACEUTICAL CO., LTD.**, Beijing (CN)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **18/331,197**

(22) Filed: **Jun. 8, 2023**

(65) **Prior Publication Data**

US 2023/0310616 A1 Oct. 5, 2023

Related U.S. Application Data

(63) Continuation of application No. PCT/CN2023/085045, filed on Mar. 30, 2023.

(30) **Foreign Application Priority Data**

Mar. 30, 2022 (WO) PCT/CN2022/083923
Mar. 27, 2023 (WO) PCT/CN2023/084208

(51) **Int. Cl.**

A61K 38/26 (2006.01)
A61K 47/54 (2017.01)
A61K 47/65 (2017.01)
A61P 3/04 (2006.01)
A61P 3/10 (2006.01)
A61P 25/28 (2006.01)

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

CPC *A61K 47/543* (2017.08); *A61K 38/26* (2013.01); *A61K 47/65* (2017.08); *A61P 3/04* (2018.01); *A61P 3/10* (2018.01); *A61P 25/28* (2018.01)

(58) **Field of Classification Search**

CPC A61K 47/543; A61K 38/26; A61K 47/65;
A61P 3/10

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,208,020 A	5/1993	Chari et al.
6,191,102 B1	2/2001	Dimarchi et al.
6,268,343 B1	7/2001	Knudsen et al.
6,441,163 B1	8/2002	Chari et al.
6,458,924 B2	10/2002	Knudsen et al.
6,737,236 B1	5/2004	Pieken et al.
7,083,970 B2	8/2006	Schultz et al.

7,211,557 B2	5/2007	Dimarchi et al.
7,226,990 B2	6/2007	Knudsen et al.
7,235,627 B2	6/2007	Knudson et al.
7,576,059 B2	8/2009	Jonassen et al.
7,893,017 B2	2/2011	Lau et al.
8,097,698 B2	1/2012	Knudsen et al.
8,129,343 B2	3/2012	Lau et al.
8,445,433 B2	5/2013	Werbitzky et al.
8,536,122 B2	9/2013	Lau et al.
8,563,521 B2	10/2013	Skerra et al.
8,603,972 B2	12/2013	Lau et al.
8,648,041 B2	2/2014	Garibay et al.
8,673,860 B2	3/2014	Schellenberger et al.
8,895,694 B2	11/2014	Spetzler et al.
8,957,021 B2	2/2015	Schellenberger et al.
9,006,178 B2	4/2015	Kofoed et al.
9,067,977 B2	6/2015	Spetzler et al.
9,498,534 B2	11/2016	Caggiano et al.
9,527,900 B2	12/2016	Linderoth et al.
9,550,819 B2	1/2017	Lindhout et al.
9,657,079 B2	5/2017	Spetzler et al.
9,708,383 B2	7/2017	Madsen et al.
9,732,137 B2	8/2017	Lau et al.
9,758,560 B2	9/2017	Lau et al.
10,000,542 B2	6/2018	Kofoed et al.
10,005,827 B2	6/2018	Spetzler et al.
10,010,614 B2	7/2018	Reedtz-Runge et al.
10,308,700 B2	6/2019	Lau et al.
10,370,426 B2	8/2019	Oh et al.
10,392,428 B2	8/2019	Kofoed
10,398,782 B2	9/2019	Gao et al.
10,400,020 B2	9/2019	Oh et al.

(Continued)

FOREIGN PATENT DOCUMENTS

CN	1232470 A	10/1999
CN	100444898 C	12/2008

(Continued)

OTHER PUBLICATIONS

Wang et al., "Expanding the Genetic Code of *Escherichia coli*", Science 292: 498-500, Apr. 20, 2001.
Wang and Schultz et al., "An Expanded Eukaryotic Genetic Code", Science 301: 964-967, Aug. 15, 2003.
Altschul et al., "Basic Local Alignment Search Tool", J. Mol. Biol., 215:403-410 (1990).
Altschul et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Research, 1997, vol. 25, No. 17, 3389-3402.
Higgins et al., "Using CLUSTAL for Multiple Sequence Alignments", Methods in Enzymology, 266:383-402 (1996).
Larkin et al., "Clustal W and Clustal X version 2.0", Bioinformatics (Oxford, England), 23(21): 2947-8 (2007).
Lau et al., "The discovery of the once weekly glucagon like peptide 1 (GLP-1) analog semaglutide", J. Med. Chem. 2015, 58, 7370-7380.

(Continued)

Primary Examiner — Gyan Chandra

(74) Attorney, Agent, or Firm — Junhe Law Office P.C.; James J . Zhu

(57) **ABSTRACT**

The present disclosure provides polypeptide conjugates comprising GLP-1 receptor agonist and a peptide linker, and liquid pharmaceutical compositions comprising the same. Methods of using such for treating diseases are also provided.

16 Claims, 11 Drawing Sheets

Specification includes a Sequence Listing.

(56)

References Cited

U.S. PATENT DOCUMENTS

10,604,554 B2	3/2020	Kofoed et al.
10,689,429 B2	6/2020	Linderoth et al.
10,946,074 B2	3/2021	Kofoed et al.
11,130,794 B2	9/2021	Tornoe et al.
11,529,394 B2*	12/2022	Zhang
2003/0027996 A1	2/2003	Staby
2003/0199672 A1	10/2003	Knudsen et al.
2004/0265952 A1	12/2004	Deiters et al.
2007/0042956 A1	2/2007	Johansen et al.
2007/0093417 A1	4/2007	Hansen et al.
2007/0203058 A1	8/2007	Lau et al.
2007/0203068 A1	8/2007	Nielsen
2008/0081038 A1	4/2008	Cho et al.
2008/0214439 A1	9/2008	Grabstein et al.
2009/0005312 A1	1/2009	Hansen et al.
2009/0092582 A1	4/2009	Bogin et al.
2009/0156478 A1	6/2009	Lau et al.
2010/0239554 A1	9/2010	Schellenberger et al.
2010/0260706 A1	10/2010	Bogin et al.
2010/0317057 A1	12/2010	Lau et al.
2011/0213131 A1	9/2011	Christensen et al.
2011/0312881 A1	12/2011	Silverman et al.
2012/0000418 A1	1/2012	Huang
2013/0244931 A1	9/2013	Lau et al.
2015/0017188 A1	1/2015	Eigenbrot, Jr. et al.
2015/0152157 A1	6/2015	Kofoed et al.
2015/0210745 A1	7/2015	Kofoed et al.
2015/0266943 A1	9/2015	Chhabra et al.
2015/0273069 A1	10/2015	Bjerregaard et al.
2016/0143998 A1	5/2016	Reedtz-Runge et al.
2016/0158321 A1	6/2016	Cleland et al.
2016/0199454 A1	7/2016	Liu et al.
2017/0037088 A1	2/2017	Schellenberger et al.
2017/0058014 A1	3/2017	Wieczorek et al.
2017/0145069 A1	5/2017	Lau et al.
2017/0240614 A1	8/2017	Baldwin et al.
2017/0320927 A1	11/2017	Sauerberg et al.
2018/0051063 A1	2/2018	Cleland et al.
2018/0125988 A1	5/2018	Yang et al.
2018/0127512 A1	5/2018	Doronina et al.
2018/0291076 A1	10/2018	Kjeldsen et al.
2019/0083577 A1	3/2019	Schellenberger et al.
2019/0153115 A1	5/2019	Schellenberger et al.
2020/0087379 A1	3/2020	Schellenberger et al.
2020/0231645 A1	7/2020	Tornoe et al.
2020/0392195 A1	12/2020	Schellenberger et al.
2022/0009989 A1	1/2022	Muenzel et al.
2022/0202704 A1*	6/2022	Giehm
		A61K 47/02

FOREIGN PATENT DOCUMENTS

CN	101712722 A	5/2010
CN	103619175 A	3/2014
WO	1991011457 A1	8/1991
WO	1996029342 A1	9/1996
WO	1998008871 A1	3/1998
WO	1999043341 A1	9/1999
WO	1999043705 A1	9/1999
WO	1999043706 A1	9/1999
WO	1999043707 A1	9/1999
WO	02/46227 A2	6/2002
WO	2007103515 A2	9/2007
WO	2011123813 A2	10/2011

WO	2016131893 A1	8/2016
WO	2020207477 A1	10/2020
WO	2021/139744 A1	7/2021
WO	2022/068920 A1	4/2022
WO	2023/179796 A1	9/2023

OTHER PUBLICATIONS

- Dennis et al., "Albumin Binding as a General Strategy for Improving the Pharmacokinetics of Proteins", *J. Biol Chem.* 277, 38 (2002) 35035-35043.
- Batzer et al., "Enhanced evolutionary PCR using oligonucleotides with inosine at the 3'-terminus", *Nucleic Acids Research* (1991), vol. 19, No. 18, 5081.
- Ohtsuka et al., "An Alternative Approach to Deoxyoligonucleotides as Hybridization Probes by Insertion of Deoxyinosine at Ambiguous Codon Positions**", *J. Biol. Chem.* 260: 2605-2608 (1985).
- Rossolini et al., "Use of deoxyinosine-containing primers vs degenerate primers for polymerase chain reaction based on ambiguous sequence information", *Mol. Cell. Probes* 8:91-98 (1994).
- Carter et al., "High Level *Escherichia coli* Expression and Production of a Bivalent Humanized Antibody Fragment", *Bio/Technology* 10:163-167 (1992).
- Carlsson et al., "Protein Thiolation and Reversible Protein-Protein Conjugation N-Succinimidyl 3-(2-Pyridyl)dithio)Propionate, A New Heterobifunctional Reagent", *Biochem. J.* (1978) 173, 723-737.
- Kim, Y. et al., "Novel AGLP-1 albumin fusion protein as a long-lasting agent for type 2 diabetes", *BMB Reports*, Dec. 31, 2013 (Dec. 31, 2013) No. 12, vol. 46, 606-610.
- Jiang, Y.N. et al., "Application of new fusion sequence in long-term GLP-1 drugs development", *Chinese Doctoral Dissertations & Master's Theses Full-text Database (Master) Medicine and Health Sciences*, Feb. 15, 2018 (Feb. 15, 2018) No. 02.
- McBrayer, D.N. et al., "Recent Advances in GLP-1 Receptor Agonists for Use in Diabetes Mellitus", *Drug Dev Res.*, Sep. 30, 2017 (Sep. 30, 2017) No. 6, vol. 78, abstract, pp. 292-299.
- Deiters et al., "Adding Amino Acids with Novel Reactivity to the Genetic Code of *Saccharomyces cerevisiae*", *J. Am. Chem. Soc.* 2003, 125, 11782-11783.
- Zhang et al., "A New Strategy for the Synthesis of Glycoproteins", *Science* 303:371-373, 2004.
- The first office action of the corresponding Chinese application 202180006277.3, mailed on Aug. 26, 2022.
- Park Junyoung et al.: "Effect of C-terminus Conjugation via Different Conjugation Chemistries on In Vivo Activity of Albumin-Conjugated Recombinant GLP-1", *Pharmaceutics*, vol. 13, No. 2, Feb. 15, 2021 (Feb. 15, 2021), p. 263, XP093244774, Switzerland ISSN: 1999-4923, DOI: 10.3390/pharmaceutics/3020263.
- Tang Daoqi et al.: "C-terminal site-specific PEGylated Exendin-4 analog: A long-acting glucagon like Peptide-1 receptor agonist, on glycemic control and beta cell function in diabetic db/db mice", *Journal of Pharmacological Sciences*, vol. 138, No. 1, Sep. 1, 2018 (Sep. 1, 2018), pp. 23-30, XP093244771, JP ISSN: 1347-8613, DOI: 10.1016/j.jphs.2018.08.009.
- Gao Mingming et al.: "A site-specific PEGylated analog of exendin-4 with improved pharmacokinetics and pharmacodynamics in vivo", *Journal of Pharmacy and Pharmacology: JPP*, vol. 64, No. 11, Jun. 4, 2012 (Jun. 4, 2012), pp. 1646-1653, XP093182980, GB ISSN: 0022-3573, DOI: 10.1111/j.2042-7158.2012.01545.x.
- The partial supplementary European search report for counterpart EP application 23778368.3, mailed on Feb. 11, 2025.

* cited by examiner

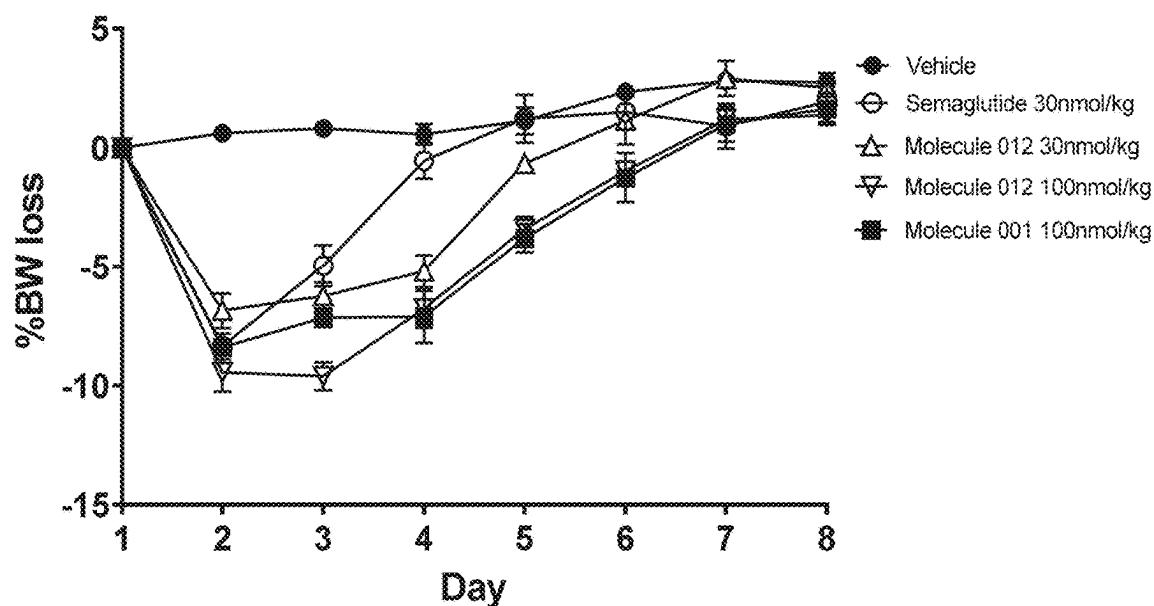


Figure 1A

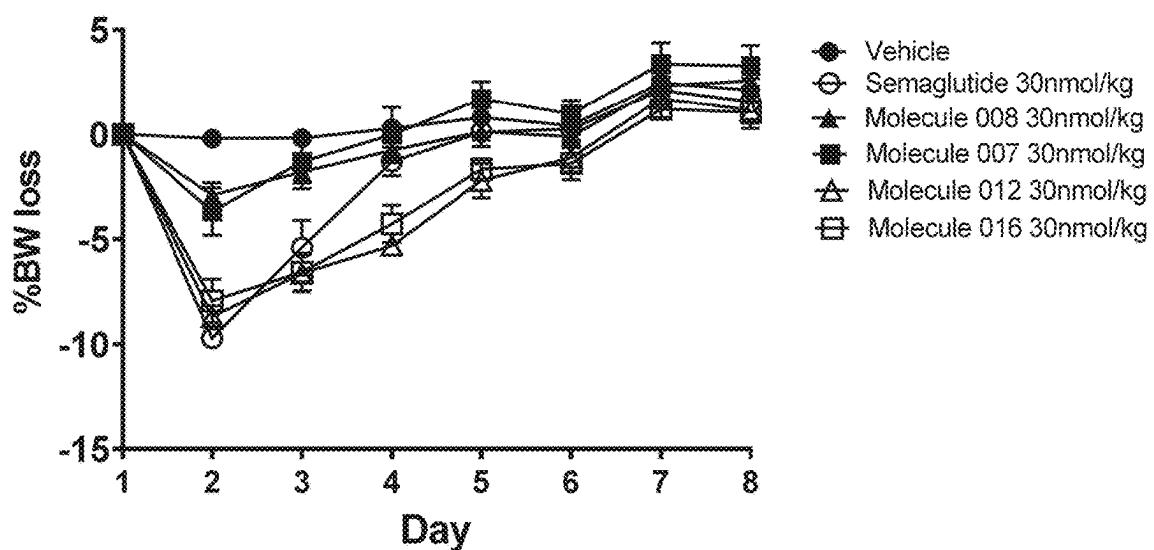


Figure 1B

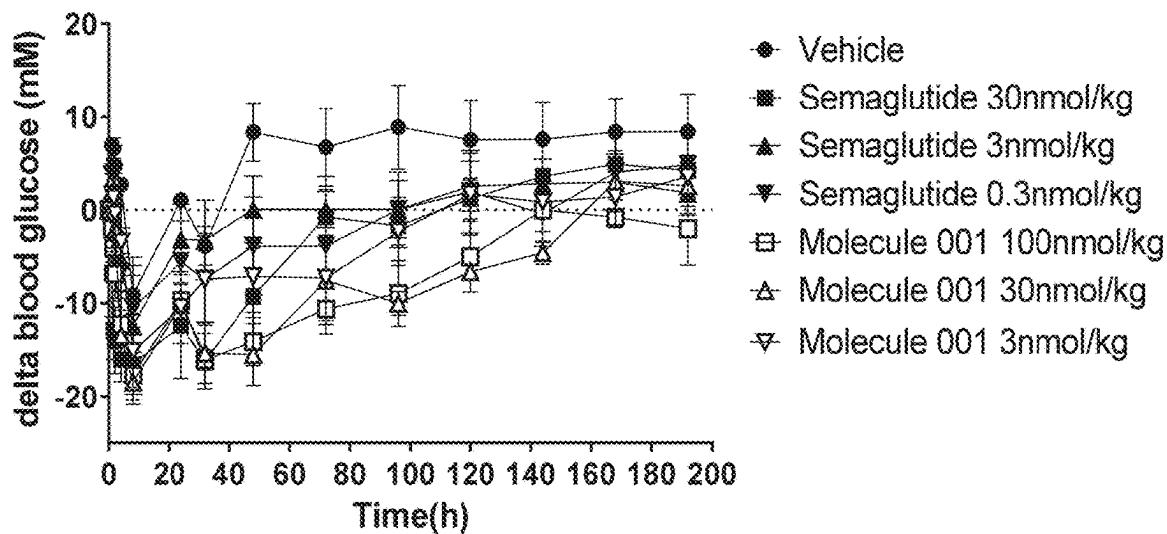


Figure 2A

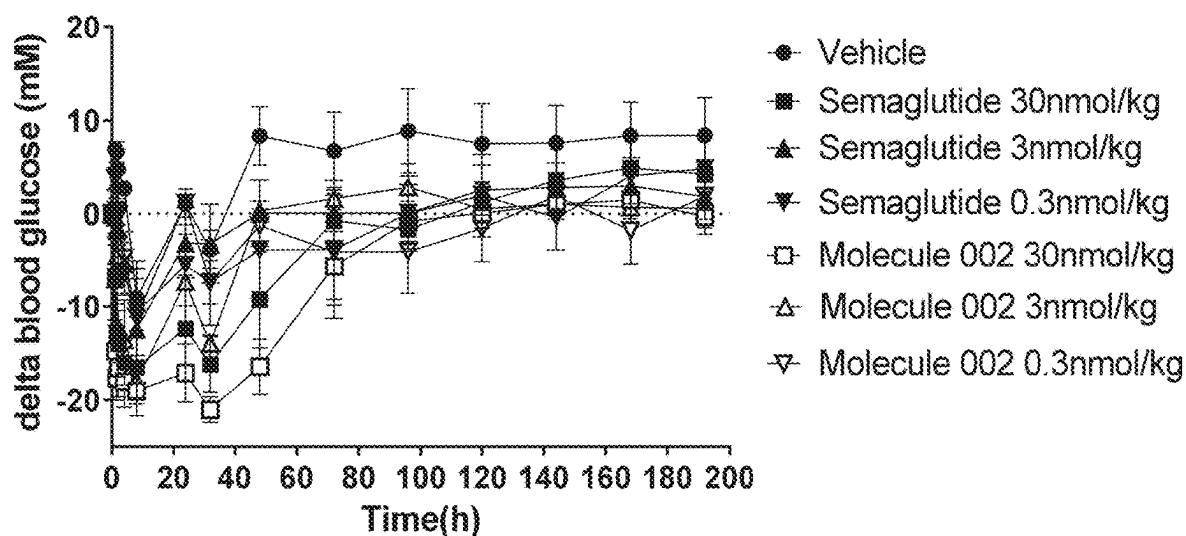


Figure 2B

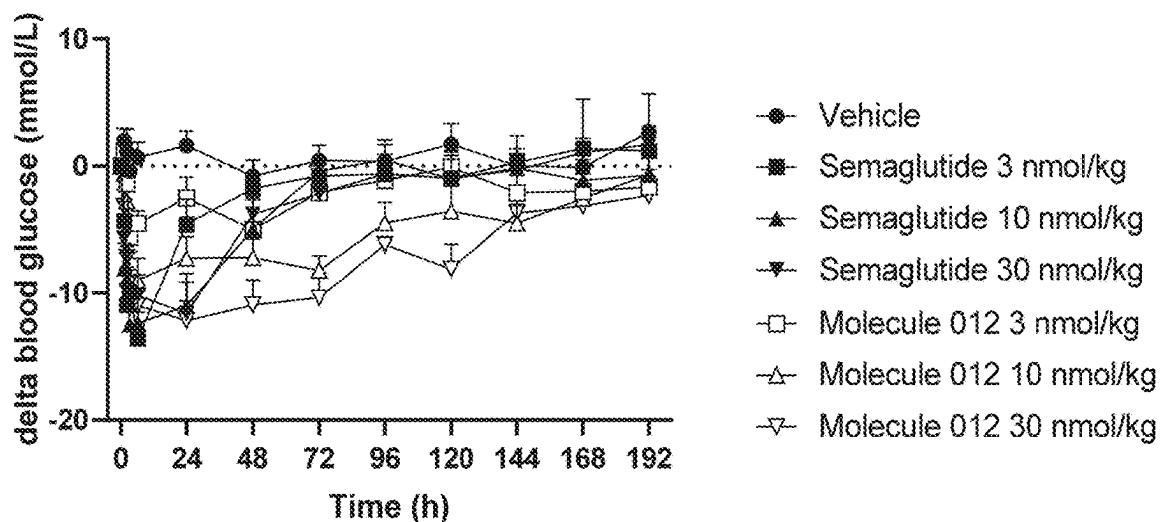


Figure 2C

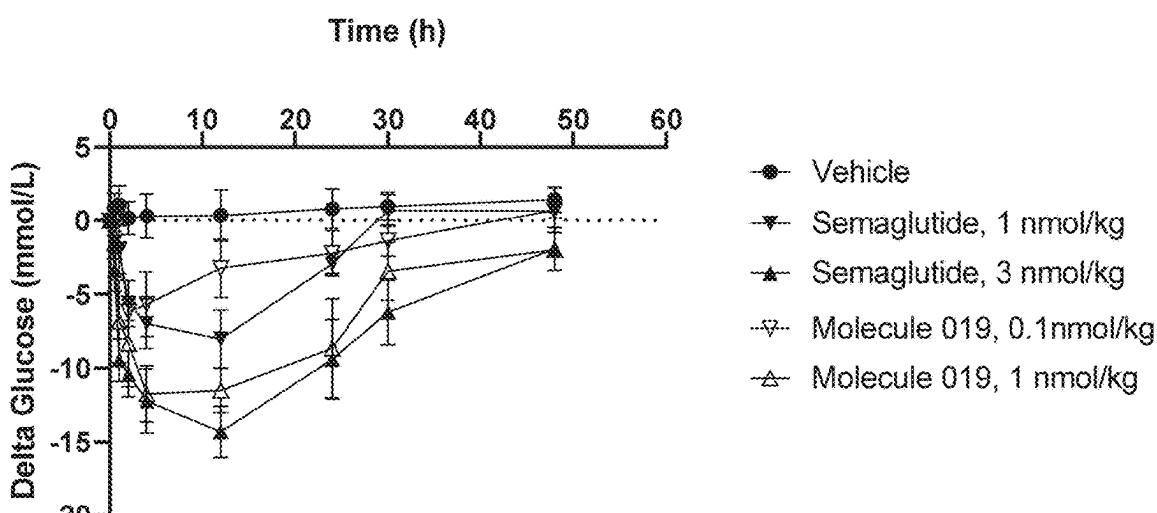


Figure 2D

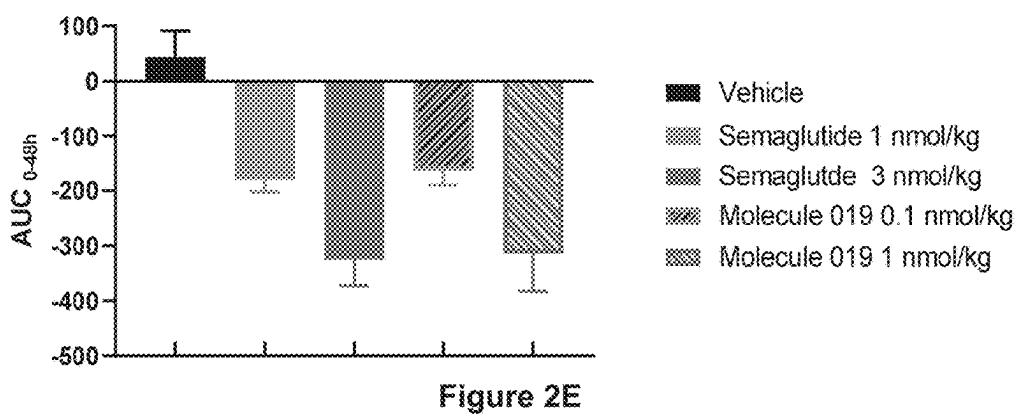


Figure 2E

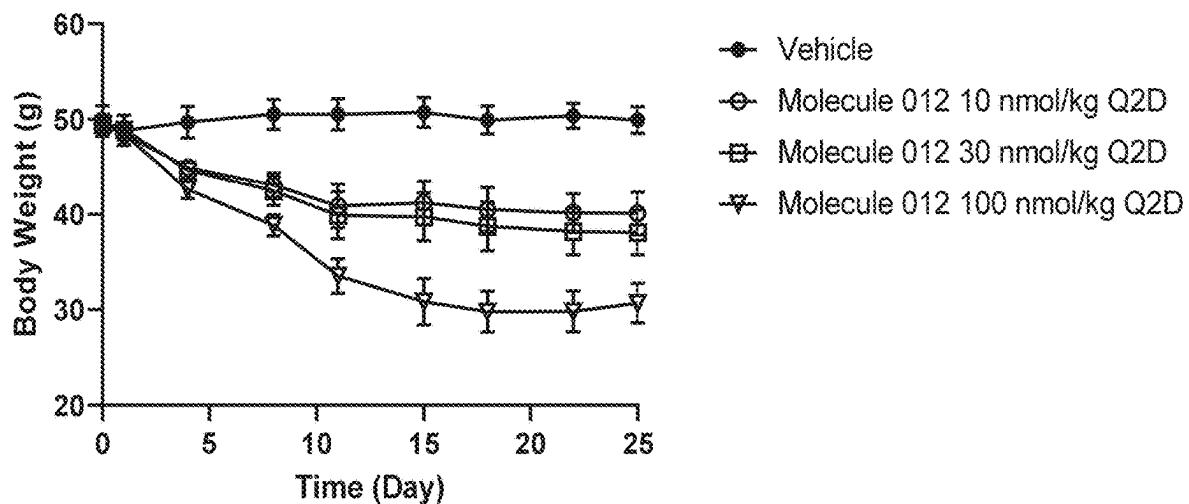


Figure 3A

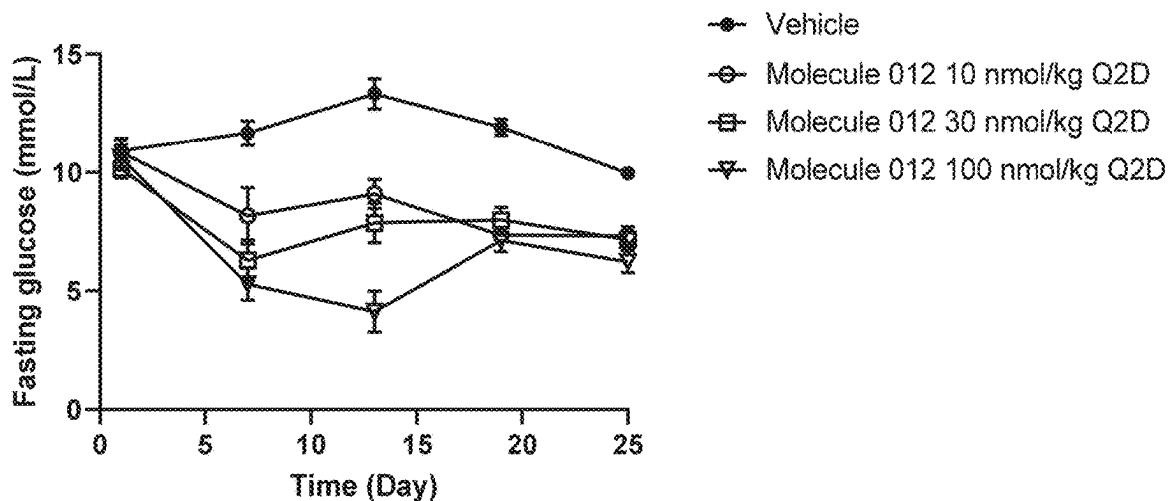


Figure 3B

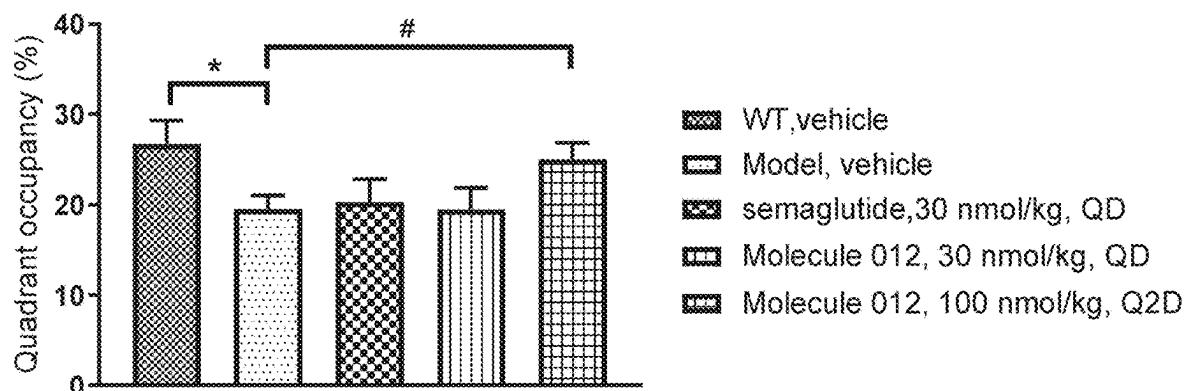


Figure 4A

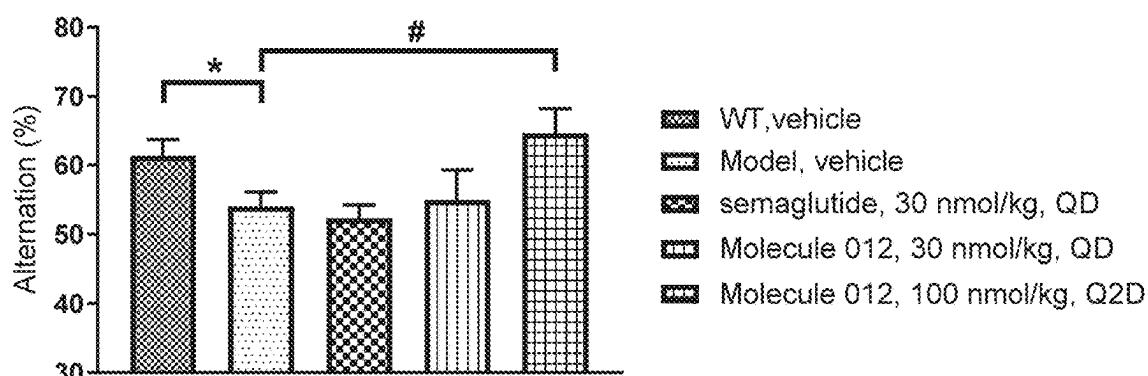


Figure 4B

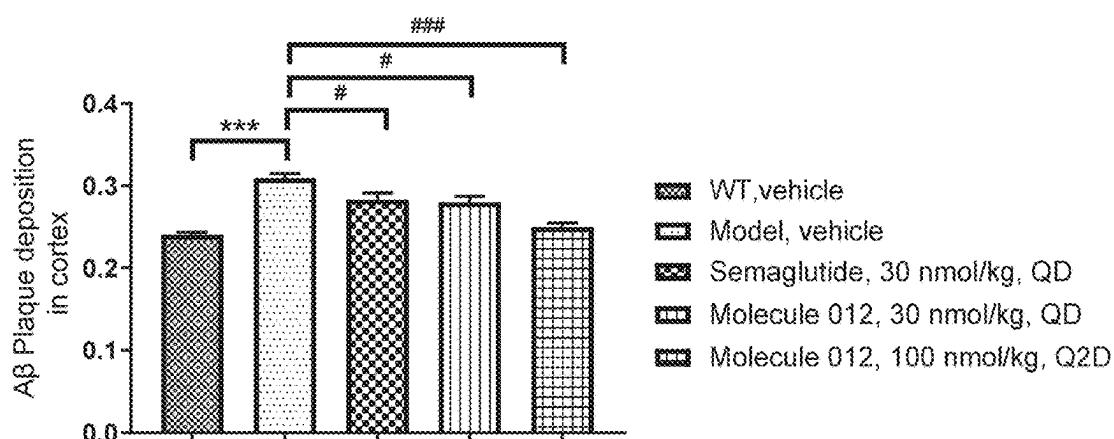


Figure 4C

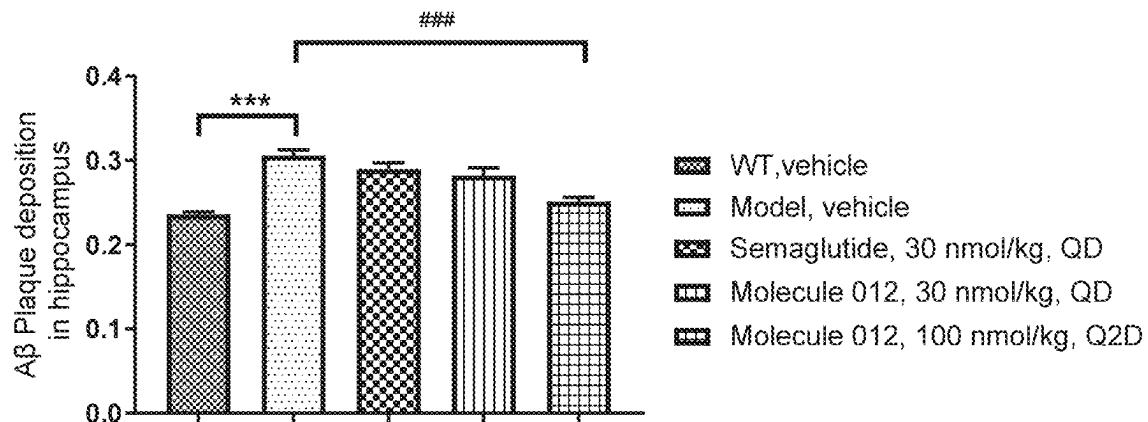


Figure 4D

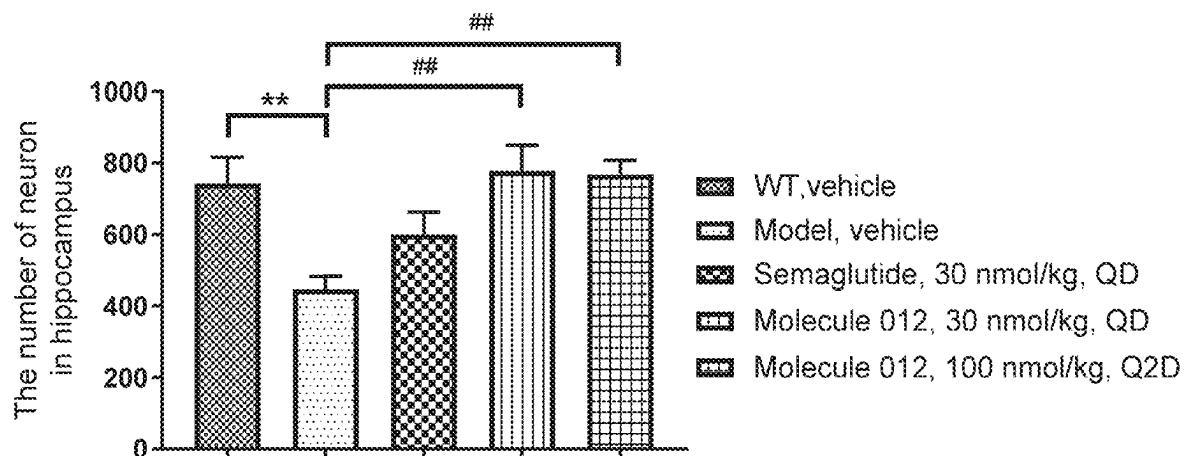


Figure 4E

GLP-1 Sequences in the Polypeptide Portion of the Polypeptide Conjugate

SEQ ID NO	GLP-1 Sequence
1	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG
2	X ₇ X ₈ EGTFTSDVSSYLEX ₂₂ X ₂₃ AAX ₂₆ X ₂₇ FIX ₃₀ WLVX ₃₄ GX ₃₆ G
3	HAib-EGTFTSDVSSYLEGQAAKEFIAWLVRGGG
4	HAib-EGTFTSDVSSYLEGQAACEFIAWLVRGGG
5	HAib-EGTFTSDVSSYLEGQAAREFIAWLVRGGG
6	HAib-EGTFTSDVSSYLEEQAAKEFIAWLVRGGG
7	HAib-EGTFTSDVSSYLEEQAAACEFIAWLVRGGG
8	HAib-EGTFTSDVSSYLEEQAAAREFIAWLVRGGG
9	HGEQTFTSDVSSYLEEQAAAREFIAWLVRGGG

Figure 5A

Repeating Sequences in the Peptide Linker (i.e. Linker Repeats)
in the Polypeptide Portion of the Polypeptide Conjugate

SEQ ID NO	Linker Repeat Sequence
10	GQEPGAQP
11	GAQPGAQP
12	GQEP
13	GAQP
14	GAQPGQEPGAQP
15	GAQPGQEP
16	GEQP
17	GPQE
18	GPEQ
19	GSEP
20	GESP
21	GPSE
22	GPES
23	GQAP
24	GPAQ

Figure 5B

25	GPQA
26	GSQP
27	GASP
28	GPAS
29	GPSA
30	GGGS
31	GSGS
32	GGGGS
33	GQEPGQAP
34	GQAPGQEP
35	SEPATSGSETPGTSESATPESGPGTSTEPSEG
36	SEPATS
37	GSETPG
38	TSESAT
39	PESGPG
40	TSTEPS

Figure 5B (Continue)

Other Sequences in the Peptide Linker in the Polypeptide Portion of the Polypeptide Conjugate (underlined residues representing the CRM residue)

SEQ ID NO	Sequence
41	G <u>Q</u> KP
42	G <u>Q</u> CP

Figure 5C

Peptide Linker Sequences (i.e. Linker Sequences) in the Polypeptide Portion of the Polypeptide Conjugate

SEQ ID NO	Linker Sequence
43	GAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGQKP
44	GAQPGAQPGAQPGAQPGAQPGAQPGAQPGAQPGQKP
45	GAQPGAQPGAQPGAQPGAQPGAQPGAQPGQKP

Figure 5D

46	GAQPGAAQPGAQPGAAQPGAAQPGQKP
47	GAQPGAAQPGAQPGAAQPGAAQPGAAQPGAAQPGQCP
48	GAQPGAAQPGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQCP
49	GAQPGAAQPGAQPGAAQPGAAQPGAAQPGAAQPGQCP
50	GAQPGAAQPGAQPGAAQPGAAQPGQCP
82	GAQPGAAQPGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQKP
83	GAQPGAAQPGAQPGAAQPGAAQPGQCP
84	GAQPGAAQPGQKP

Figure 5D (Continue)

**Polypeptide Portion Sequences (i.e. Fusion Polypeptide Sequences)
of the Polypeptide Conjugate**

SEQ ID NO	Fusion Polypeptide Sequence
51	HAib- EGTFTSDVSSYLEGQAAKEFIAWLVRGGGGAQPGAAQPGAAQPGAAQPGAAQPGQKP
52	HAib- EGTFTSDVSSYLEGQAAREFIAWLVRGGGGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQKP
53	HAib- EGTFTSDVSSYLEGQAAKEFIAWLVRGGGGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQKP
54	HAib- EGTFTSDVSSYLEGQAAKEFIAWLVRGGGGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQKP
55	HAib- EGTFTSDVSSYLEGQAAREFIAWLVRGGGGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQKP
56	HAib- EGTFTSDVSSYLEGQAAREFIAWLVRGGGGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQKP
57	HAib- EGTFTSDVSSYLEEQAQAEFIAWLVRGGGGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQKP
58	HAib- EGTFTSDVSSYLEEQAQAEFIAWLVRGGGGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQKP
59	HAib- EGTFTSDVSSYLEEQAQAEFIAWLVRGGGGAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGAAQPGQKP

Figure 5E

60	HAib- EGTFTSDVSSYLEEQAAREFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQKP
61	HAib- EGTFTSDVSSYLEEQAAREFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQKP
62	HAib- EGTFTSDVSSYLEEQAAREFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGQKP
63	HAib- EGTFTSDVSSYLEGQAACEFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
64	HAib- EGTFTSDVSSYLEGQAACEFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
65	HAib- EGTFTSDVSSYLEGQAACEFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
66	HAib- EGTFTSDVSSYLEGQAAREFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
67	HAib- EGTFTSDVSSYLEGQAAREFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
68	HAib- EGTFTSDVSSYLEGQAAREFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGQCP
69	HGEGTFTSDVSSYLEEQAAREFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGAQPGQCP
70	HGEGTFTSDVSSYLEEQAAREFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
71	HGEGTFTSDVSSYLEEQAAREFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
72	HAib- EGTFTSDVSSYLEEQAACEFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
73	HAib- EGTFTSDVSSYLEEQAACEFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
74	HAib- EGTFTSDVSSYLEEQAACEFIAWLVRGGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP

Figure 5E (Continue)

75	HAib- EGTFTSDVSSYLEGQAAKEFIAWLRGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
76	HAib- EGTFTSDVSSYLEGQAAKEFIAWLRGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
77	HAib- EGTFTSDVSSYLEGQAAKEFIAWLRGGGAQPGAQPGAQPGAQPGAQPGQCP
78	HAib- EGTFTSDVSSYLEEQAQEFIAWLRGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
79	HAib- EGTFTSDVSSYLEEQAQEFIAWLRGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
80	HAib- EGTFTSDVSSYLEEQAQEFIAWLRGGGAQPGAQPGAQPGAQPGAQPGAQPGQCP
85	HAib- EGTFTSDVSSYLEGQAAKEFIAWLRGGGAQPGAQPGAQPGAQPGAQPGAQPGAQPGQCP
86	HAib- EGTFTSDVSSYLEGQAAKEFIAWLRGGGAQPGAQPGAQPGAQPGAQPGAQPGQKP
87	HAib-EGTFTSDVSSYLEGQAAKEFIAWLRGGGAQPGAQPGQKP
88	HAib-EGTFTSDVSSYLEGQAAREFIAWLRGGGAQPGAQPGQKP

Figure 5E (Continue)

1

**LIQUID PHARMACEUTICAL
COMPOSITIONS OF POLYPEPTIDE
CONJUGATES AND METHODS OF USES
THEREOF**

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted electronically in XML format and is hereby incorporated by reference in its entirety. Said XML copy, created on Jun. 6, 2023, is named 074585-8010US01.xml and is 127,753 bytes in size.

FIELD OF THE INVENTION

The present disclosure relates generally to the field of therapeutic peptides, and more specifically relates to polypeptide conjugates, i.e. glucagon-like peptide 1 (GLP-1) compounds, their pharmaceutical compositions, and methods of using such to prevent and/or treat metabolic disorders such as diabetes.

BACKGROUND

Metabolic disorders or metabolic disease are a group of conditions characterized by the inability of the body to properly transform foods into energy, and to utilize and/or store energy, with the most prominent metabolic disorder being diabetes. Metabolic disorders are commonly associated with insulin resistance, visceral adiposity, atherogenic dyslipidemia, etc., which pose major and escalating public health and clinical challenge worldwide.

Glucagon-like-peptide-1 (GLP-1) is a proglucagon-derived peptide that is secreted from intestinal L-cells in response to nutrient ingestion. GLP-1 primarily acts as an incretin, i.e. an endocrine hormone, to increase the insulin response following oral intake of food, by generally regulating the concentrations of glucagons, slowing down gastric emptying, stimulating the biosynthesis of (Pro-)insulin, increasing the sensitivity toward insulin, and stimulating the insulin-independent biosynthesis of glycogen, etc. Because GLP-1 can rapidly lower glucose levels in both normal and diabetic subjects, there has been considerable interest in developing GLP-1 based pharmaceutical agents (i.e. referred to as GLP-1 compounds hereinafter) for preventing and/or treating type 2 diabetes.

Currently, it is known that human GLP-1 in its native form (having 37 amino acid residues) is poorly active, and the two major naturally-occurring and biologically active truncated versions of the native GLP-1 include a 30- and 31-amino acid peptide fragment, i.e. GLP-1(7-36) or GLP-1 (7-37), which derive from the posttranslational processing of the proglucagon peptide, have extremely short in vivo half-lives due primarily to the N-terminal cleavage and inactivation by the dipeptidyl peptidase DPP-IV. Although tremendous efforts have been invested in developing newer and better GLP-1 compounds with more favorable administration regimens, existing GLP-1 compounds and their pharmaceutical compositions still have far from ideal half-lives, efficacy, and/or dosing frequency.

Therefore, there is a need for an improved design for GLP-1 compounds and for improved pharmaceutical compositions thereof.

SUMMARY OF THE INVENTION

Provided herein include pharmaceutical compositions and methods of use thereof for treating/preventing metabolic disorders.

2

In a first aspect, the present disclosure provides a pharmaceutical composition, which comprises a polypeptide conjugate and a pharmaceutically acceptable excipient.

The polypeptide conjugate substantially comprises a polypeptide portion and a conjugate portion. The polypeptide portion comprises a single biologically active peptide and a peptide linker. The biologically active peptide is attached to N-terminus of the peptide linker and comprises a GLP-1 receptor agonist. The conjugate portion comprises a first clearance-reducing moiety (CRM) conjugated to a first CRM residue in the peptide linker.

In some embodiments, the first CRM residue is at least 5 amino acid residues (exclusive of the CRM residue), e.g. 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, 33, 35, 38, 40, 43, 45, 48, 50, 53, 55, 58, 60, 63, 65, 68, 70, 73, 75, or 78 amino acid residues, away from the C-terminal amino acid residue of the biologically active peptide (e.g. the GLP-1 receptor agonist).

In some embodiments, the GLP-1 receptor agonist comprises GLP-1. As used herein, the term "GLP-1" refers to a molecule having the following characteristics: the molecule has substantially the biological activity of a wildtype human GLP-1 (SEQ ID NO: 1), and the molecule comprises a polypeptide fragment having an amino acid sequence of at least 50% identity to the wildtype human GLP-1.

In some embodiments, the GLP-1 comprises an amino acid sequence having at least 70% sequence identity to SEQ ID NO: 1 while retaining substantial biological activity of SEQ ID NO: 1.

In some embodiments, the GLP-1 comprises an amino acid sequence of $X_7X_8EGTFTSDVSS-YLEX_{22}X_{23}AAX_{26}X_{27}FIX_{30}WLVX_{34}GX_{36}G$ (SEQ ID NO: 2), where the X_7 is H, imidazole-4-acetate (IA), or imidazolepropionic acid (IPA); the X_8 is A, G, S, V, Aib, T, I, or L; the X_{22} is G, or E; the X_{23} is Q, C or K; the X_{26} is K, R, or C; the X_{27} is E, K, or C; the X_{30} is A, C or K; the X_{34} is R, K, or C; and the X_{36} is R or G.

In some embodiments of the GLP-1, the X_7 is H; and X_8 is G or Aib.

In some embodiments, the GLP-1 comprises or consists of one or more mutations at a position selected from the group consisting of: A8, G22, K26, K34, and R36, or any combination thereof, relative to SEQ ID NO: 1.

In some embodiments, the GLP-1 comprises or consists of one or more substitutions selected from the group consisting of: A8Aib, G22E, K26R, K34R and R36G, or any combination thereof.

In some embodiments, the GLP-1 comprises an amino acid sequence selected from the group consisting of: SEQ ID NOS: 3-9.

In any embodiment of the polypeptide conjugate as mentioned above, the peptide linker can have a length of at least 10 amino acid residues, at least 12, 24, 32, 40, 48, 50, 60, 70, or 80 amino acid residues.

In some embodiments, the peptide linker consists of amino acid residues selected from the group consisting of G, Q, A, E, P, S, and T, except for the first CRM residue.

In some embodiments, the peptide linker comprises a first sequence and a second sequence. The first sequence consists of one or more repeats of a repeating sequence, and is connected to N-terminus of the second sequence. The first CRM residue is in the second sequence, and is lysine residue or cysteine residue.

In some embodiments, the repeating sequence in the first sequence of the peptide linker has an amino acid sequence selected from a group consisting of SEQ ID NO: 10 (GQEP-

3

GAQP), SEQ ID NO: 11 (GAQPGAQP), SEQ ID NO: 12 (GQEP), SEQ ID NO: 13 (GAQP), SEQ ID NO: 14 (GAQPGQEPGAQP), SEQ ID NO: 15 (GAQPGQEP), SEQ ID NO: 16 (GEQP), SEQ ID NO: 17 (GPQE), SEQ ID NO: 18 (GPEQ), SEQ ID NO: 19 (GSEP), SEQ ID NO: 20 (GESP), SEQ ID NO: 21 (GPSE), SEQ ID NO: 22 (GPES), SEQ ID NO: 23 (GQAP), SEQ ID NO: 24 (GPAQ), SEQ ID NO: 25 (GPQA), SEQ ID NO: 26 (GSQP), SEQ ID NO: 27 (GASP), SEQ ID NO: 28 (GPAS), SEQ ID NO: 29 (GPSA), SEQ ID NO: 30 (GGGS), SEQ ID NO: 31 (GSGS), SEQ ID NO: 32 (GGGGS), SEQ ID NO: 33 (GQEPGQAP), SEQ ID NO: 34 (GQAPGQEP), SEQ ID NO: 35 (SEPATSG-SETPGTSESATPESGPGTSTEPSEG), SEQ ID NO: 36 (SEPATS), SEQ ID NO: 37 (GSETPG), SEQ ID NO: 38 (TSESAT), SEQ ID NO: 39 (PESGPG), SEQ ID NO: 40 (TSTEPS) and GS.

In some embodiments, the repeating sequence in the first sequence of the peptide linker has a sequence of SEQ ID NO: 13 (GAQP). Herein optionally, the number of the one or more repeats of the repeating sequence is an integer between 1 and 30. In certain embodiments, the number of the one or more repeats of the repeating sequence is selected from a group consisting of 5, 7, 9 and 11.

In certain embodiments, the second sequence of the peptide linker has a sequence of SEQ ID NO: 41 (GQKP) or SEQ ID NO: 42 (GQCP).

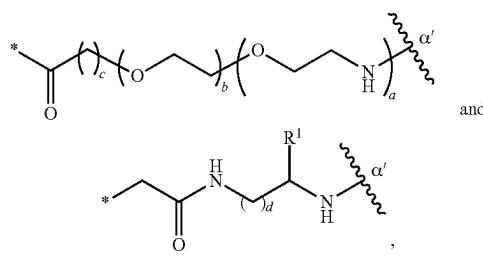
In certain embodiments, the polypeptide linker comprises an amino acid sequence selected from the group consisting of: SEQ ID NOS: 43-50, and 82-84.

In any embodiment of the polypeptide conjugate as mentioned above, the first CRM may comprise a plasma protein-binding moiety, a polymer, Fc, HSA (albumin), Xten sequence, or PAS sequence.

In some embodiments, the first CRM comprises an albumin-binding moiety.

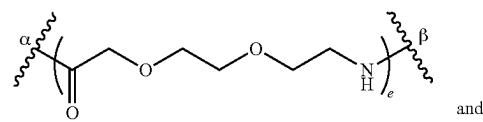
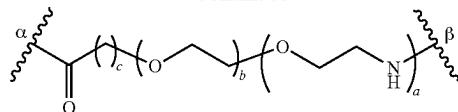
In some embodiments, the albumin-binding moiety comprising a structure of: *-A-B—C-D-E, wherein A, B, C, D and E are interconnected via amide bonds, and the * end of A is connected to a reactive group of the conjugatable residue on the polypeptide complex.

Herein, A is selected from a bond,



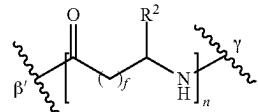
where a, b, c and d are independently an integer from 0 to 4, R¹ is hydrogen or —COOH.

Herein, B is selected from a bond,

**4****-continued**

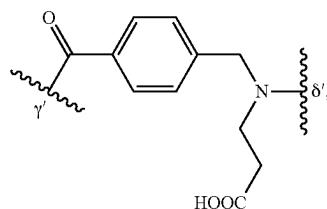
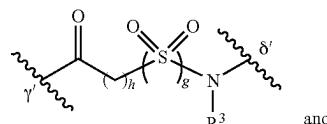
where e is an integer from 1 to 4, wherein position α is linked to position α' .

Herein, C is a bond or



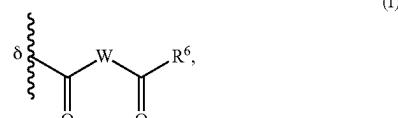
where R² is —CH₃SO₃H or —COOH, f is an integer from 1 to 4, n is an integer from 1 to 25, wherein when B is not bond, then position β' is linked to position β , or when B is bond, then position β' is linked to position α' .

Herein, D is selected from a bond,



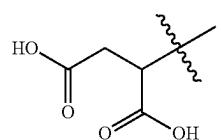
where g and h are independently 0 or 1, and R³ is H or —CH₂COOH. Herein, when B is not a bond and C is a bond, then position γ' is linked to position β ; when C is not a bond, then position γ' is linked to position γ ; and when B is a bond and C is a bond, then position γ' is linked to position α' .

Herein, E is an acidic group having a formula:



where W represents —(CR⁴R⁵)₁—, where R⁴ and R⁵ are independently selected from the group consisting of hydrogen, halogen, cyano, hydroxyl, alkyl, alkenyl, alkynyl, cycloalkyl, heterocyclyl, aryl, heteroaryl, hydroxyalkyl, amino, aminoalkyl, carboxyl, carboxylalkyl, alkoxy, aryloxy, and carboxamide, R⁶ is selected from hydroxyl or NR⁷R⁸, where R⁷ and R⁸ are independently selected from the group consisting of hydrogen, alkyl, hydroxyl, and

5



and I is an integer from 10 to 20. Herein, when D is not a bond, then position δ is linked to position δ' ; when C is not a bond and D is a bond, then position δ is linked to position γ ; when B is not a bond, C is a bond and D is a bond, then position δ is linked to position β ; when A is not a bond, and all of B, C, and D are bond, then position δ is linked to position α' .

In some embodiments of the first CRM, A is a bond or

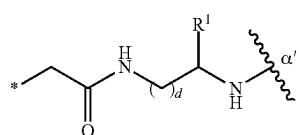
6

R^4 and R^5 is hydrogen, R^6 is hydroxyl, I is an integer from 10 to 20, and position δ is linked to position γ .

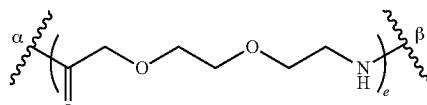
In some embodiments of the first CRM, d is 1, e is 2, f is 1, n is 1, and I is an integer from 14 to 20.

In some embodiments of the first CRM, I is 16.

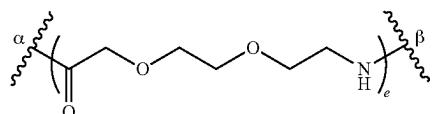
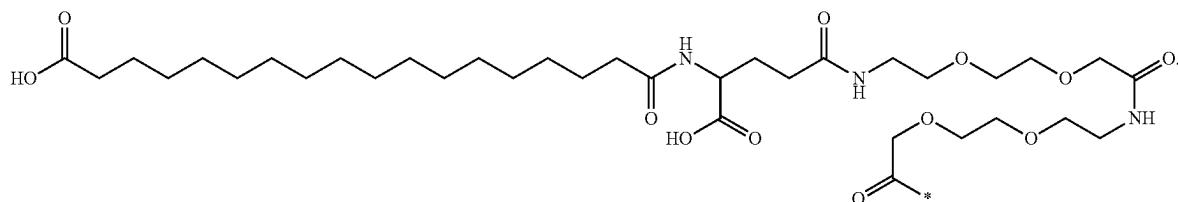
In some embodiments of the polypeptide conjugate, the first CRM residue is lysine residue, wherein: A is a bond and B is



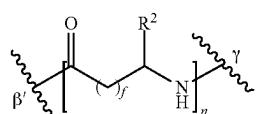
where d is an integer from 1 to 4, R^1 is hydrogen; B is a bond or



In some embodiments, the first CRM comprises Moiety A ($HOOC-(CH_2)^{16}-CO-gGlu-2XADO$), the Moiety A having the structure of below formula:



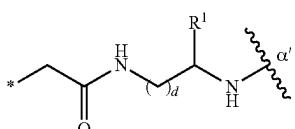
where e is an integer from 1 to 4, wherein position α is linked to position α' ; C is



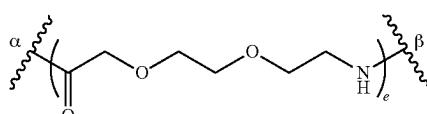
where R^2 is $-COOH$, f is an integer from 1 to 4, n is an integer from 1 to 25, and when B is not bond, then position β' is linked to position β , or when B is bond, then position β' is linked to position α' ; D is a bond; regarding E, each of

45

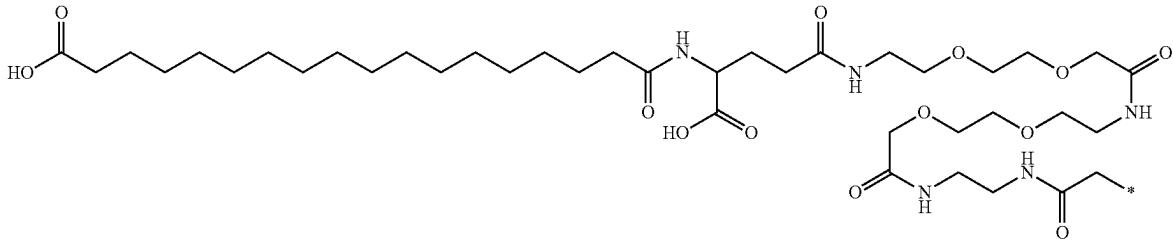
In some embodiments of the polypeptide conjugate, the first CRM residue is cysteine residue, wherein A is



and B is



In some embodiments, the first CRM comprises Moiety B ($HOOC-(CH_2)^{16}-CO-gGlu-2XADO-EDA-CO-CH_2$), the Moiety B having the structure of below formula:



In certain embodiments, the polypeptide conjugate provided herein is mono-conjugated, and has the CRM conjugated to the peptide linker but not to the biologically active peptide (e.g. GLP-1).
25

In some embodiments of the mono-conjugated polypeptide conjugate, the first CRM residue is lysine residue, and the polypeptide conjugate comprises only one lysine residue due.
30

In some embodiments of the mono-conjugated polypeptide conjugate, the first CRM residue is cysteine residue, and the polypeptide conjugate comprises only one cysteine residue.
35

In some embodiments, the GLP-1 comprises an amino acid sequence selected from the group consisting of: SEQ ID NOS: 3, 5, 6, 8, and 9.

Herein, in some embodiments, the polypeptide portion comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 52, 55, 56, 60-62, 66-71, and 75-80; and the first CRM residue is lysine residue or cysteine residue at a position selected from the group consisting of 40
60, 68, and 76.
45

The following lists certain embodiments of the mono-conjugated polypeptide conjugate disclosed herein:

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 52, and is conjugated with the 50
first CRM at 76K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 55, and is conjugated with the first CRM at 68K;
55

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 56, and is conjugated with the first CRM at 60K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 60, and is conjugated with the 60
first CRM at 76K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 61, and is conjugated with the first CRM at 68K;
65

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 62, and is conjugated with the first CRM at 60K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 66, and is conjugated with the first CRM at 76C;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 67, and is conjugated with the first CRM at 68C;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 68, and is conjugated with the first CRM at 60C;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 69, and is conjugated with the first CRM at 76C;
35

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 70, and is conjugated with the first CRM at 68C;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 71, and the first CRM is conjugated with the polypeptide portion at 60C;
40

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 75, and is conjugated with the first CRM at 76C;
45

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 76, and is conjugated with the first CRM at 68C;
50

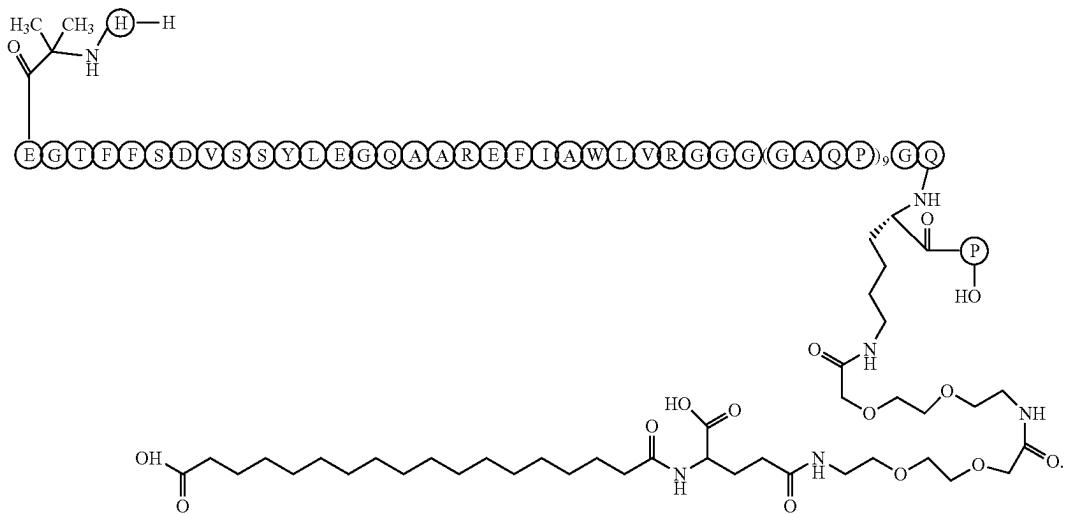
the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 77, and is conjugated with the first CRM at 60C;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 78, and is conjugated with the first CRM at 76C;
55

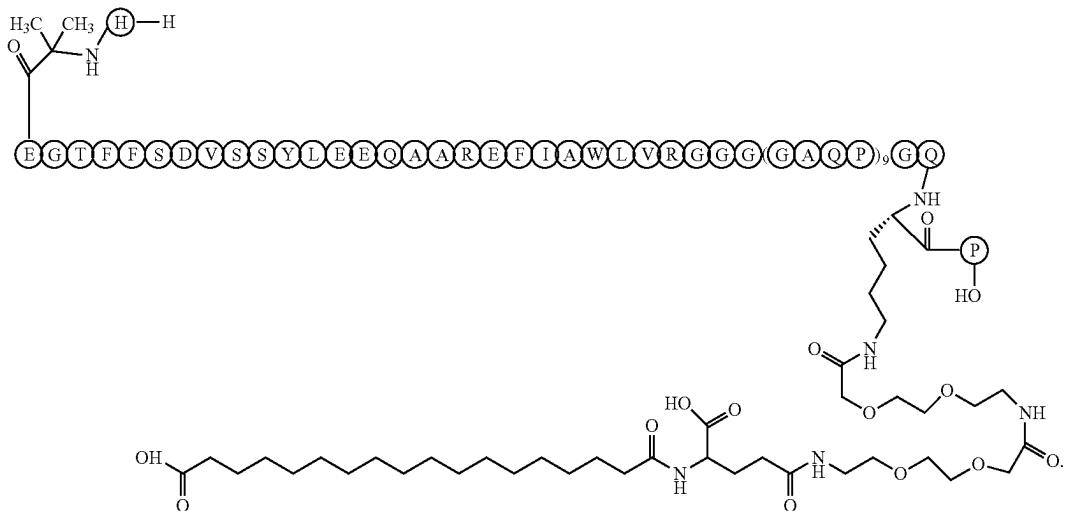
the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 79, and is conjugated with the first CRM at 68C; or
60

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 80, and is conjugated with the first CRM at 60C.
65

In certain embodiment of the polypeptide conjugate, the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 52, and the first CRM residue is 76K. More specifically, certain embodiment of the polypeptide conjugate has the structure shown below:



In certain embodiment of the polypeptide conjugate, the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 60, and the first CRM residue is 76K. More 25 specifically, certain embodiment of the polypeptide conjugate has the structure shown below:



In some embodiments, the polypeptide conjugate is a 50 double-conjugated polypeptide conjugate and further comprises a second CRM conjugated to a second CRM residue.

In some embodiments of the double-conjugated polypeptide conjugate, the first CRM residue and the second CRM residue are both lysine residues, and the polypeptide conjugate comprises only two lysine residues.

Herein optionally, the second CRM residue is in the GLP-1, and is optionally selected from the group consisting of K23, K26, K27, K30 and K34.

In some embodiments, the second CRM residue is K26.

In some embodiments of the double-conjugated polypeptide conjugate, the first CRM residue and the second CRM residues are both cysteine residues, and the polypeptide conjugate comprises only two cysteine residues.

Herein optionally, the second CRM residue is in the GLP-1, and is optionally selected from the group consisting of C23, C26, C27, C30 and C34.

In some embodiments, the second CRM residue is C26. Herein, the positions as described above (i.e. 23, 26, 27, 30 and 34, either as K or C) are referenced relative to the wildtype human GLP-1 (SEQ ID NO: 1) and starts from the N-terminus thereof.

Herein, according to some embodiments, the second CRM is identical to the first CRM. In some embodiments, the second CRM residue and the first CRM residue are both lysine residues or both cysteine residues.

In some embodiments of the double-conjugated polypeptide conjugate, the GLP-1 comprises an amino acid sequence selected from the group consisting of: SEQ ID NOS: 3, 4, 6 and 7.

In some embodiments, the polypeptide portion comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 51, 53, 54, 57-59, 63-65, and 72-74, 85 and 86; the first CRM residue is lysine residue or cysteine residue at a position selected from the group consisting of 68, 76, and 84; and the second CRM residue is lysine residue or cysteine residue at position 26.

The following list certain embodiments of double-conjugated polypeptide conjugate as disclosed herein:

11

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 51, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 53, and is conjugated with the first CRM and the second CRM respectively at 26K and 84K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 54, and is conjugated with the first CRM and the second CRM respectively at 26K and 68K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 58, and is conjugated with the first CRM and the second CRM respectively at 26K and 84K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 59, and is conjugated with the first CRM and the second CRM respectively at 26K and 68K;

5

10

15

20

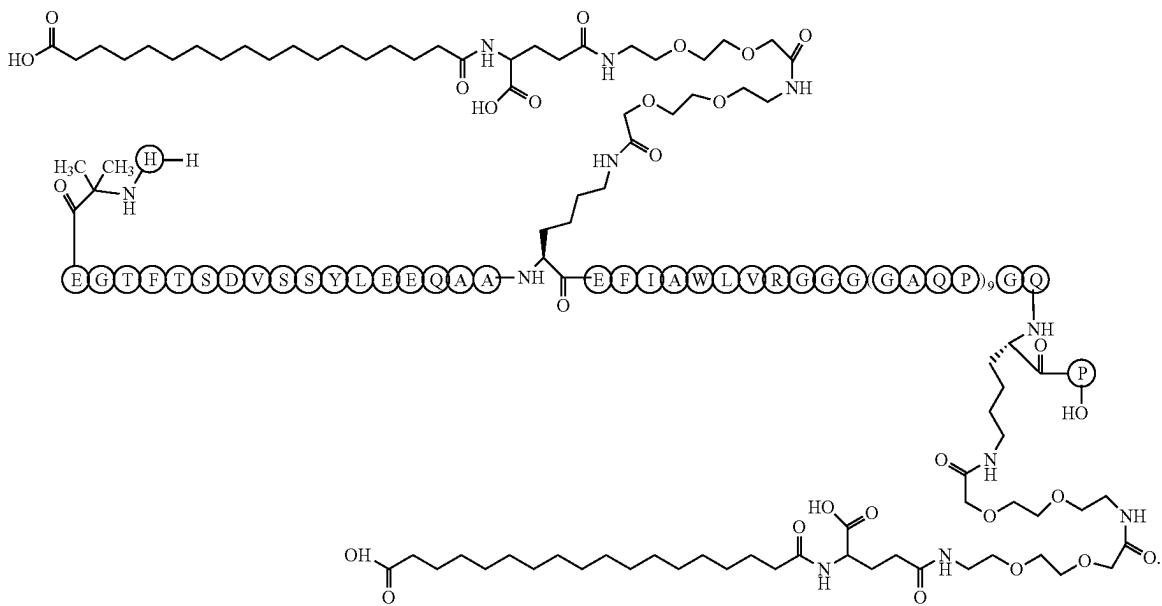
the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 65, and is conjugated with the first CRM and the second CRM respectively at 26C and 68C;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 72, and is conjugated with the first CRM and the second CRM respectively at 26C and 84C;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 73, and is conjugated with the first CRM and the second CRM respectively at 26C and 76C; or

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 74, and is conjugated with the first CRM and the second CRM respectively at 26C and 68C.

In certain embodiments, the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K. More specifically, certain embodiment of the polypeptide conjugate has the structure shown below:



the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 85, and is conjugated with the first CRM and the second CRM respectively at 26K and 96K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 86, and is conjugated with the first CRM and the second CRM respectively at 26K and 60K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 63, and is conjugated with the first CRM and the second CRM respectively at 26C and 84C;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 64, and is conjugated with the first CRM and the second CRM respectively at 26C and 76C;

Herein, in order to obtain the polypeptide conjugate in the pharmaceutical composition provided in the first aspect, a polynucleotide encoding the polypeptide portion (or fragment thereof) of the polypeptide conjugate can be designed, which can be operably arranged in a vector. The vector can be transferred in a host cell, such as a prokaryotic cell or eukaryotic cell, and the host cell can be cultured under a condition that allows expression of the polynucleotide, thereby obtaining the polypeptide portion of the polypeptide conjugate. Herein optionally, the polypeptide portion is expressed as soluble proteins. After obtaining the polypeptide portion, the CRM can be further conjugated to the polypeptide portion to thereby obtain the single-conjugated or double-conjugated polypeptide conjugate.

In another aspect, the pharmaceutical composition provided herein is a liquid formulation.

13

In certain embodiments, the pharmaceutically acceptable excipient comprises a buffer and an isotonic agent.

In certain embodiments, the buffer is selected from the group consisting of a phosphate buffer, citrate buffer, acetate buffer, histidine buffer, glycine buffer, carbonate buffer, borate buffer, glutamate buffer, glycylglycine buffer, lysine buffer, and arginine buffer.

In certain embodiments, the buffer is a phosphate buffer.

In certain embodiments, the phosphate buffer is present in a concentration of 0.01-50 mM of the pharmaceutical composition.

In certain embodiments, the phosphate buffer is present in a concentration of 5-20 mM or 5-10 mM of the pharmaceutical composition, optionally about 10 mM, 9 mM, 8 mM, 7 mM, 6 mM or 5 mM.

In certain embodiments, the phosphate buffer is selected from the group consisting of sodium dihydrogen phosphate, disodiumhydrogen phosphate, sodium phosphate, or a hydrate thereof.

In certain embodiments, the hydrate is dodecahydrate or dihydrate.

In certain embodiments, the phosphate buffer is disodiumhydrogen phosphate dodecahydrate or disodium phosphate dihydrate.

In certain embodiments, the disodiumhydrogen phosphate dodecahydrate is present in a concentration of about 0.1 to 15 mg/mL, 0.5 to 15 mg/mL, 1 to 15 mg/mL, 0.5 to 12 mg/mL, 0.5 to 10 mg/mL, 0.5 to 8 mg/mL, 0.5 to 7 mg/mL, or 0.5 to 5 mg/mL of the pharmaceutical composition.

In certain embodiments, the disodiumhydrogen phosphate dodecahydrate is present in a concentration of about 2.87 mg/mL of the pharmaceutical composition.

In certain embodiments, the disodium phosphate dihydrate is present in a concentration of about 0.1 to 15 mg/mL, 0.5 to 15 mg/mL, 1 to 15 mg/mL, 0.5 to 12 mg/mL, 0.5 to 10 mg/mL, 0.5 to 8 mg/mL, 0.5 to 7 mg/mL, or 0.5 to 5 mg/mL of the pharmaceutical composition.

In certain embodiments, the disodium phosphate dihydrate is present in a concentration of about 1.42 mg/mL.

In certain embodiments, the buffer is a citrate buffer.

In certain embodiments, the citrate buffer is present in a concentration of about 1-50 mM.

In certain embodiments, the citrate buffer comprises a mixture of citric acid anhydrous and trisodium citrate (or a hydrate thereof, e.g. trisodium citrate dihydrate).

In certain embodiments, the citrate buffer comprises a mixture of 0.14 mg/mL citric acid anhydrous and 2.74 mg/mL trisodium citrate dihydrate.

In certain embodiments, the buffer is a histidine buffer.

In certain embodiments, the histidine buffer is present in a concentration of about 1-70 mM of the pharmaceutical composition, optionally about 5-50 mM, 5-20 mM or 5-10 mM, or optionally about 10 mM, 9 mM, 8 mM, 7 mM, 6 mM or 5 mM.

In certain embodiments, the histidine buffer is present in a concentration of about 0.5-10 mg/mL, 0.5-5 mg/mL, 1-5 mg/mL, or 1-3 mg/mL of the pharmaceutical composition.

In certain embodiments, the histidine buffer is present in a concentration of about 1.24 mg/mL of the pharmaceutical composition.

In certain embodiments, the isotonic agent is selected from the group consisting of sodium chloride, propylene glycol, sorbitol, sucrose, glycine, mannitol, lactose monohydrate, arginine, myoinositol and dimethylsulfone.

In certain embodiments, the isotonic agent is sodium chloride.

14

In certain embodiments, the sodium chloride is about 5-15 mg/m L.

In certain embodiments, the isotonic agent is propylene glycol.

In certain embodiments, the propylene glycol is about 1 mg/mL to about 50 mg/mL (e.g. about 5 mg/mL to about 25 mg/mL, about 8 mg/mL to about 16 mg/mL).

In certain embodiments, the isotonic agent is mannitol.

In certain embodiments, the mannitol is about 20 mg/mL to about 60 mg/mL (e.g. about 25 mg/mL to about 50 mg/mL, about 30 mg/mL to about 50 mg/mL, about 35 mg/mL to about 50 mg/mL).

In certain embodiments, the pharmaceutical excipient further comprises a preservative, a chelating agent, and/or a stabilizer.

In certain embodiments, the pharmaceutical composition has a pH of about 6.5 to about 8.3 (e.g. about 6.5 to 7.4, or from 7.4 to 8.3).

In certain embodiments, the pharmaceutical composition has a pH of about 7.4.

In certain embodiments, the polypeptide conjugate is a double-conjugated polypeptide conjugate as provided herein. In certain embodiments, the pharmaceutical composition has about 1-80 mg/ml, 1-100 mg/ml, or 1-120 mg/ml, or optionally, 5-40 mg/ml, 5-80 mg/ml, 5-100 mg/ml, or 5-120 mg/ml of the polypeptide conjugate (e.g. about 5-90 mg/mL, about 5-70 mg/mL, about 5-60 mg/mL, about 5-50 mg/mL, about 5-30 mg/mL, about 5-20 mg/mL, about 5-10 mg/mL). In some embodiments of the pharmaceutical composition, the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K.

In certain embodiments, the pharmaceutical compositions comprise:

(a) about 1-80 mg/mL, 1-100 mg/ml or 1-120 mg/mL of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;

(b) a buffer selected from the group consisting of phosphate buffer, citrate buffer, and histidine buffer;

(c) an isotonic agent selected from the group consisting of sodium chloride, propylene glycol, mannitol; and

(d) a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

(a) about 1-80 mg/mL, 1-100 mg/ml or 1-120 mg/mL of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;

(b) a buffer selected from the group consisting of phosphate buffer, citrate buffer, and histidine buffer;

(c) an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and

(d) a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

(a) about 1-80 mg/mL, 1-100 mg/ml or 1-120 mg/mL of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is

15

- conjugated with the first CRM and the second CRM respectively at 26K and 76K;
- about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
 - an isotonic agent selected from the group consisting of sodium chloride, propylene glycol, mannitol; and
 - a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- about 1-80 mg/mL, 1-100 mg/ml or 1-120 mg/mL of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;
- about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
- an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and
- a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- about 5-80 mg/mL, 5-100 mg/ml, or 5-120 mg/mL of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;
- about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
- an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and
- a pH of about 6.5 to about 8.3.

In certain embodiments, the polypeptide conjugate is a mono-conjugated polypeptide conjugate as provided herein. In certain embodiments, the pharmaceutical composition has about 0.5-20 mg/ml, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate. In certain embodiments of the pharmaceutical composition, the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60, and is conjugated with the first CRM at 76K. In certain embodiments of the pharmaceutical composition, the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 52, and is conjugated with the first CRM at 76K.

In certain embodiments, the pharmaceutical compositions comprise:

- about 0.5-20 mg/mL, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60 or SEQ ID NO: 52, and is conjugated with the first CRM at 76K;
- a buffer selected from the group consisting of phosphate buffer, citrate buffer, and histidine buffer;
- an isotonic agent selected from the group consisting of sodium chloride, propylene glycol, mannitol; and
- a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- about 0.5-20 mg/mL, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60 or SEQ ID NO: 52, and is conjugated with the first CRM at 76K;

16

- about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
- an isotonic agent selected from the group consisting of sodium chloride, propylene glycol, mannitol; and
- a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- about 0.5-20 mg/mL, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60 or SEQ ID NO: 52, and is conjugated with the first CRM at 76K;
- a buffer selected from the group consisting of phosphate buffer, citrate buffer, and histidine buffer;
- an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and
- a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- about 0.5-20 mg/mL, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60 or SEQ ID NO: 52, and is conjugated with the first CRM at 76K;
- about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
- an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and
- a pH of about 6.5 to about 8.3.

In a second aspect, the present disclosure provides a method of preventing or treating a metabolic disorder in a subject in need thereof, comprising administering a therapeutically effective amount of the polypeptide conjugate of the first aspect.

In certain embodiments, the metabolic disorder is diabetes, obesity, overweight, non-alcoholic steatohepatitis (NASH), cardiovascular like dyslipidemia, arteriosclerosis, alcoholic steatohepatitis (ASH), diabetic nephropathy, gestational diabetes, metabolic syndrome such as metabolic syndrome X, nonalcoholic fatty liver disease (NAFLD), end-stage liver disease, hepatic steatosis (fatty liver), liver cirrhosis, or primary biliary cirrhosis (PBC), or Alzheimer's disease.

In certain embodiments, the diabetes can be any form of diabetes, including without limitation, hyperglycemia, type 2 diabetes, impaired glucose tolerance, type 1 diabetes, non-insulin dependent diabetes, MODY (maturity onset diabetes of the young), gestational diabetes, and elevated level of HbA1C.

In certain embodiments, the metabolic disorder is Alzheimer's disease. In certain embodiments, the present disclosure provides a method of preventing or treating Alzheimer's disease in a subject in need thereof, comprising administering a therapeutically effective amount of the polypeptide conjugate or a pharmaceutical composition provided herein.

In certain embodiments, the present disclosure provides a method of preventing or treating Alzheimer's disease in a subject in need thereof, comprising administering a pharmaceutical composition provided herein comprising the polypeptide conjugate provided herein, and wherein the pharmaceutical composition is a liquid formulation.

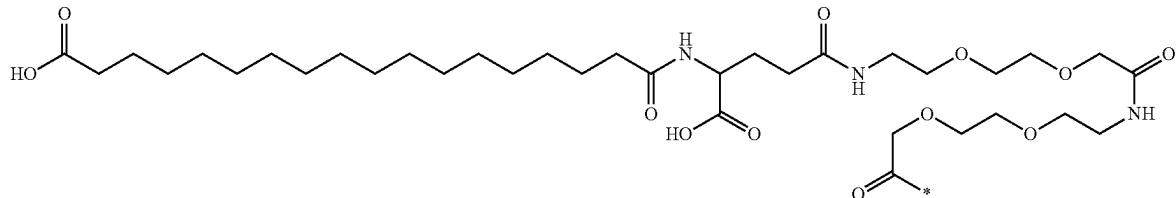
In certain embodiments, the polypeptide conjugate comprises a polypeptide portion and a conjugate portion, wherein the polypeptide portion comprises a single biologi-

17

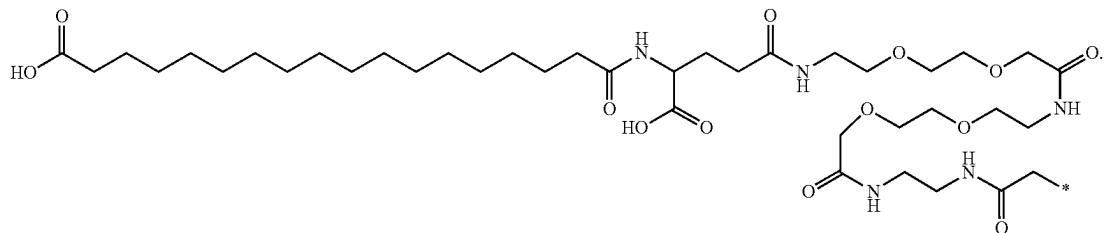
cally active peptide and a peptide linker, and the conjugate portion comprises a first clearance-reducing moiety (CRM) conjugated to a first CRM residue in the peptide linker. In certain embodiments, the polypeptide portion comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 52, 55, 56, 60-62, 66-71, and 75-80; and the first CRM residue is lysine residue or cysteine residue at a position selected from the group consisting of 60, 68, and 76.

In some embodiments, the polypeptide conjugate is a double-conjugated polypeptide conjugate and further comprises a second CRM conjugated to a second CRM residue. In some embodiments of the double-conjugated polypeptide conjugate, the first CRM residue and the second CRM residue are both lysine residues, and the polypeptide conjugate comprises only two lysine residues. In some embodiments, the polypeptide portion comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 51, 53, 54, 57-59, 63-65, and 72-74, 85 and 86; the first CRM residue is lysine residue or cysteine residue at a position selected from the group consisting of 68, 76, and 84; and the second CRM residue is lysine residue or cysteine residue at position 26. In some embodiments, the polypeptide portion comprises the amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K.

In some embodiments, the CRM residue is lysine residue and the CRM comprises Moiety A (HOOC—(CH₂)₁₆CO-gGlu-2XADO), the Moiety A having the structure of below formula:



In some embodiments, the CRM residue is cysteine residue and the CRM comprises Moiety B (HOOC—(CH₂)₁₆CO-gGlu-2XADO-EDA-CO—CH₂), the Moiety B having the structure of below formula:



In a third aspect, the present disclosure provides a method of managing body weight in a subject in need thereof, comprising administering a therapeutically effective amount of the polypeptide conjugate provided herein, thereby managing body weight of the subject.

In a fourth aspect, the present disclosure provides a method of reducing food intake in a subject in need thereof, comprising administering a therapeutically effective amount

18

of the polypeptide conjugate provided herein, thereby reducing food intake of the subject.

In a fifth aspect, the present disclosure provides a method of reducing body weight in a subject in need thereof, comprising administering a therapeutically effective amount of the polypeptide conjugate provided herein, thereby reducing body weight of the subject.

In any of the method as mentioned above in the second through fifth aspect, the subject can be human.

In certain embodiments, the subject has a fasting blood glucose level of 125 mg/dL or greater.

In certain embodiments, the subject has a body mass index (BMI) of at least or higher than 25.

In certain embodiments, the pharmaceutical composition is administered at a dosing regimen that is no more frequently than twice daily, once daily, once every 3 days, or once weekly, once every two weeks, once every three weeks, or once monthly.

In certain embodiments, the pharmaceutical composition is administered twice-weekly, once-weekly, once bi-weekly, once every three weeks, once monthly, or once every two months.

In certain embodiments, the dosing regimen has a dosing interval ranging from about once every 3 days to about once per month, or from about once weekly to about once per month.

In certain embodiments, the pharmaceutical composition is administered orally.

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 fusion polypeptide” means one fusion polypeptide or more than one fusion polypeptides.

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.

As used herein, “about” or “approximately” 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.

The articles “a” and “an” are used herein to refer to one or more than one (i.e. to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

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”.

The term “or” is used inclusively herein to mean, and is used interchangeably with, the term “and/or,” unless context clearly indicates otherwise.

BRIEF DESCRIPTION OF THE FIGURES

FIGS. 1A to 1B show in vivo activities of test Molecules in C57BL/6 mice. To evaluate the body weight, 10-week old male C57BL/6 mice were dosed subcutaneously with testing drug substances once. FIG. 1A shows body weight loss after Molecule 012 or Molecule 001 treatment from Day1 to Day 8. FIG. 1B shows body weight loss after Molecules 007, 008, 012 and 016 treatment from Day1 to Day 8. Data are indicated as mean values and standard error (SEM). Semaglutide was tested in parallel as a comparative control.

FIGS. 2A, 2B, 2C, 2D and 2E show in vivo activities of Molecules 001, 002, 012 and 019 in db/db mice. FIG. 2A shows fasting glucose after Molecule 001 single dose treatment. FIG. 2B shows fasting glucose after Molecule 002 single dose treatment. Data are indicated as mean values and standard error (SEM). FIG. 2C shows non-fasting glucose after Molecule 012 single dose treatment. FIG. 2D shows non-fasting glucose after Molecule 019 single dose treatment, and FIG. 2E shows area under the curve for 0-48 hours for non-fasting glucose after Molecule 019 single dose treatment. To evaluate the glucose, 10-week old male db/db mice were administered subcutaneously with the designated GLP-1 polypeptide conjugates at the indicated dosage. Fasting glucose (FIGS. 2A and 2B) or non-fasting glucose (FIGS. 2C and 2D) was measured on different times and five animals were used for each group. Delta blood glucose is glucose subtracted by baseline level. Data are indicated as mean values and standard error (SEM).

FIGS. 3A and 3B show in vivo efficacy in Diet Induced Obese (DIO) mice. To evaluate the effects of the respective GLP-1 polypeptide conjugates on body weight, food intake and glucose levels, 22-week old DIO mice (C57BL/6 mice on high fat diet for 13 weeks) were administered every other day subcutaneously with different concentrations of the designated GLP-1 polypeptide conjugates for 25 days. FIG. 3A shows body weight change with dose titrations of Molecule 012 at 10, 30 and 100 nmol/kg.

FIG. 3B shows fasting glucose change upon treatment. Data are expressed as mean values and standard error (SEM).

FIGS. 4A-4E show in vivo efficacy in Alzheimer’ disease (APP/PS1, APPswe/PSEN1dE9) mice. To evaluate the effects of the respective GLP-1 polypeptide conjugates on

cognitive function (Morris water maze and Y-maze), brain A β plaque deposition, and hippocampal pyramidal neuron number, 6 months old APP/PS1 male mice were administered once day (QD) or every other day (Q2D) subcutaneously with different concentrations of the designated GLP-1 polypeptide conjugate (Molecule 012) for 10 weeks. FIG. 4A shows quadrant occupancy change with Molecule 012 at 30 and 100 nmol/kg, respectively. FIG. 4B shows alternation percentage change with Molecule 012 at 30 and 100 nmol/kg, respectively. FIG. 4C shows A β plaque deposition in cortex change with Molecule 012 at 30 and 100 nmol/kg, respectively. FIG. 4D shows A β plaque deposition in hippocampus change with Molecule 012 at 30 and 100 nmol/kg, respectively. FIG. 4E shows the number of neuron in hippocampus change with Molecule 012 at 30 and 100 nmol/kg, respectively. Data are expressed as mean values and standard error (SEM). Wild type mice treated with vehicle was used as control. Alzheimer’ disease (APP/PS1) mice treated with vehicle was used as control. Alzheimer’ disease (APP/PS1) mice treated with semaglutide was also used as control.

FIGS. 5A-5E show all the sequences disclosed in the present disclosure (except for SEQ ID NO: 81).

DETAILED DESCRIPTION

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

The term “amino acid” as used herein refers to an organic compound containing amine ($-\text{NH}_2$) and carboxyl ($-\text{COOH}$) functional groups, along with a side chain specific to each amino acid.

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, pyrrolidine (PYL), selenocysteine, pyrrolidine-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.

An “amino acid analog” is a compound that has the same basic chemical structure as a naturally occurring amino acid, i.e., a carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleu-

cine, 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.

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 isofarnesyl 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), b-alanine, 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, N α -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 polypeptide 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 U.S. Pat. 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. For another example, non-natural amino acid can be chemically synthesized and inserted into or attached to a polypeptide by chemical reaction such as acylation.

“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 sub-

stitution 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.

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.

The term “functional form” as used herein, refers to different forms (such as variants, fragments, 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.

The term “variant” as used herein refers to a polypeptide having at least 70% sequence identity to the parent polypeptide. A variant may differ from the parent peptide by one or more amino acid residues. For example, a variant may have substitutions, additions, deletions, insertions, or truncations of one or more amino acid residue of the parent polypeptide.

The term “fragment” as used herein refers to partial sequence of the parent polypeptide of any length. A fragment can still retain at least partial function of the parent polypeptide.

The term “derivative” as used herein refers to a chemically modified polypeptide or fusion polypeptide, in which one or more well-defined number of 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.

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.

"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.

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. 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-1 α) 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®, pCDM8, pCDNA1.1/amp, pcDNA3.1, pRc/RSV, PCR 2.1, pEF-1, pFB, pSG5, pXT1, pCDEF3, pSVSPORT, pEF-Bos etc.

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

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.

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.

Polypeptide Conjugates

In one aspect, the present disclosure provides a polypeptide conjugate, which comprises a polypeptide portion and a conjugate portion. The polypeptide portion comprises a single biologically active peptide and a peptide linker. The biologically active peptide is attached to N-terminus of the peptide linker and comprises a GLP-1 receptor agonist. The conjugate portion comprises a first clearance-reducing moiety (CRM) conjugated to a first CRM residue in the peptide linker. Herein, the first CRM residue is at least 5 amino acid residues (exclusive of the CRM residue) away from the C-terminal amino acid residue of the biologically active peptide (e.g. the GLP-1 receptor agonist).

The term "peptide" and "polypeptide" are used interchangeably herein and refer a polymer of amino acid residues linked by covalent bonds such as peptide bonds. A peptide or polypeptide as provided herein can comprise naturally occurring or non-natural amino acid residues, or both. Polypeptides and peptides provided herein can comprise any suitable length of amino acid residues, for example, from at least 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.

The polypeptide conjugate comprises a single, i.e., only one, biologically active peptide. "Single" with respect to biologically active peptide is intended to mean that the polypeptide conjugate does not contain two or more different biologically active peptides which are attached respectively to the peptide linker. The single biologically active peptide, however, can include fragments or portions from different biologically active peptides fused together, for example, as a hybrid or chimera, as long as these portions are integrated as one peptide and are not respectively attached to the linker. If, however, two different biologically active peptides are attached respectively to the N-terminus and the C-terminus of the linker, then that is not a single biologically active peptide as used in the present disclosure.

The term "biologically active peptide" as used herein means a peptide having a biological function or activity, for example, a physiological function, or a therapeutic function. In certain embodiments, the biologically active peptide is therapeutically active. A peptide which, when stand alone, does not have biological function is not a biologically active peptide. For example, a peptide linker is not a biologically active peptide unless it has its own biological function or activity, for example when used alone.

The biologically active peptide comprises a Glucagon-like peptide-1 (GLP-1) receptor agonist.

The term "Glucagon-like peptide-1 (GLP-1) receptor" (also referred to as GLP1R) is a receptor protein found on beta cells on the pancreas and on neurons of the brain, comprising one extracellular domain and one transmembrane domain. The extracellular domain can bind to the C-terminal helix of GLP-1, and the transmembrane domain can bind to the N-terminal regions of GLP-1. The GLP-1 receptor is involved in the control of blood sugar level by enhancing insulin secretion. When expressed in the brain, the GLP-1 receptor can also be involved in the control of appetite.

The term "Glucagon-like peptide-1 (GLP-1) receptor agonist" or "GLP-1 receptor agonist" as used herein refers to a

molecule which is capable of binding to and activating the GLP-1 receptor. A GLP-1 receptor agonist may elicit a magnitude of GLP-1 receptor response that is similar to or partial of a natural ligand.

The term "clearance-modifying moiety" or "CRM" as used herein refers to moiety that can alter one or more pharmacokinetic (PK) properties (for example, to increase the half-life *in vivo*). Examples of CRMs can include, without limitation, fatty acid, polyethylene glycol (PEG), glucuronic acid or other sugar based linkers, polar, positively or negatively charged groups that can increase the rates of hydrolysis of a succinimidyl ring and reduce or minimize the rate of reverse Michael reaction, therefore reduce or minimize the rate of loss of drug and the linker group from the biologically active peptide to other thiol-containing proteins and small molecules.

"CRM residue" as used herein refers to the amino acid residue that is conjugated to a CRM.

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 conjugate of the present disclosure means a compound as a result of the polypeptide and one or more clearance-modifying moieties joined together. The molecules may attach together by covalent, 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 joining may be permanent or reversible. In some embodiments, certain cleavable or non-cleavable linkages may be included.

While conjugation of a CRM to a GLP-1 receptor agonist may extend half-life of the GLP-1 receptor agonist, it could also adversely affect the biological activity of the GLP-1 receptor agonist, rendering it less active than the unconjugated counterpart. For example, semaglutide, a conjugated GLP-1 derivative, has a significantly extended half-life, but suffers from a 939-fold loss in GLP-1 receptor binding in presence of human serum albumin (I) relative to its unconjugated counterpart (*J. Med. Chem.* 2015, 58, 7370-7380).

It is unexpectedly found by the inventors that, certain CRM-conjugated polypeptide comprising GLP-1 receptor agonist can have both an extended half-life and retained biological activity, if the CRM is conjugated outside of the GLP-1 receptor agonist on a peptide linker attached to its C-terminus. Interestingly, the distance between the C-terminus of the GLP-1 receptor agonist and the conjugation site has been found by the inventors to be of particular importance. If the CRM conjugation on the peptide linker is at a position close to the C-terminus of the GLP-1 receptor agonist, it would significantly reduce biological activity, which could be even much lower than conjugation on the GLP-1 receptor agonist itself. However, if the conjugation is at a position sufficiently farther away from the C-terminus of the GLP-1 receptor agonist, then the CRM-conjugated GLP-1 receptor agonist could have both extended half-life and retained biological activity. Such unexpected effects are found with certain GLP-1 receptor agonists and certain CRMs.

Herein, according to some embodiments of the polypeptide conjugate, the first CRM residue is at least 5 amino acid residues (exclusive of the CRM residue), e.g. 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, 33, 35, 38, 40, 43, 45, 48, 50, 53, 55, 58, 60, 63, 65, 68, 70, 73, 75, or 78, etc. amino acid residues, away from the C-terminal amino acid residue of the biologically active peptide (e.g. the GLP-1 receptor agonist). In some embodiments of the polypeptide conjugate, the peptide

linker can have a length of at least 10 amino acid residues, e.g. at least 12, 24, 32, 40, 48, 50, 60, 70, or 80 amino acid residues.

In certain embodiments, the first CRM residue is arranged in the peptide linker at a position close to the C-terminus of the peptide linker, which can be between 20 and 50 amino acid residues away from the C-terminal amino acid residue of the biologically active peptide (e.g. the GLP-1 receptor agonist). More details will be provided below.

According to some embodiments, the polypeptide conjugate is mono-conjugated, i.e., the first CRM is the only conjugate and the polypeptide conjugate is conjugated only at the first CRM residue (in the peptide linker) with the first CRM.

According to some other embodiments, the polypeptide conjugate is double-conjugated, i.e., the polypeptide conjugate is conjugated with two conjugates. One such conjugate is the first CRM, which is conjugated to the polypeptide conjugate at the first CRM residue (in the peptide linker). The other such conjugate is a second CRM, which is conjugated to the polypeptide conjugate at a second CRM residue, which is arranged in the biologically active peptide (i.e. GLP-1 receptor agonist such as GLP).

More details for the GLP-1 receptor agonist, the CRM and CRM residues, and polypeptide conjugates are provided below.

GLP-1 Receptor Agonist

In certain embodiments, the GLP-1 receptor agonist comprises GLP-1. The term "Glucagon-like peptide-1" or "GLP-1" as used herein is intended to broadly encompass native GLP-1 peptide and all its functional forms such as its functional variants, fragments, fusions, derivatives and mimetics.

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:1. As used herein, when referring to a particular amino acid residue in SEQ ID NO: 1 (i.e. GLP-1 (7-37)), the numbering of the GLP-1 (1-37) is followed. In other words, SEQ ID NO: 1 corresponds to GLP-1 (7-37), and therefore the 1st residue (which is Histidine (H)) in SEQ ID NO: 1 is referred to as 7H, meaning that it corresponds to the 7th residue in GLP(1-37); and the 31st residue (which is Glycine (G)) in SEQ ID NO: 1 is referred to as 37G, meaning corresponds to the 37th residue in GLP(1-37).

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. Functional forms of the native GLP-1 peptide can contain one or more substitutions, additions, or deletions relative to SEQ ID NO: 1. Many functional forms of native 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, the disclosure of which is incorporated herein to its entirety.

In certain embodiments, the GLP-1 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: 1 while retaining substantial biological activity of SEQ ID NO: 1.

In certain embodiments, the GLP-1 comprises no more than 9, 8, 7, 6, 5, 4, 3, or 2 mutations (e.g. addition, deletion, substitution) relative to SEQ ID NO: 1 while retaining substantial biological activity of SEQ ID NO: 1. In certain embodiments, the GLP-1 comprises at least 2, 3, 4, 5, 6, 7, 8, or 9 mutations (e.g. addition, deletion, substitution) relative to SEQ ID NO: 1 while retaining substantial biological activity of SEQ ID NO: 1.

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.

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 some embodiments of the polypeptide conjugate provided herein, the GLP-1 comprises or consists of one or more mutations at a position selected from the group consisting of: A8, G22, K26, K34, and R³⁶, or any combination thereof, relative to SEQ ID NO: 1. 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. Examples of substitutions at these positions include, without limitation, A8G, A8Aib, A8T, G22E, K34R, R³⁶G, as well as the substitutions described in U.S. Pat. Nos. 8,273,854, which are incorporated herein by its entirety.

In some embodiments of the polypeptide conjugate, the GLP-1 comprises or consists of one or more substitutions selected from the group consisting of: A8Aib, G22E, K26R, K34R and R³⁶G, or any combination thereof.

In some embodiments of the polypeptide conjugate, the GLP-1 is a molecule which has substantially the biological activity of a wildtype human GLP-1 (SEQ ID NO: 1), and comprises a polypeptide fragment having an amino acid sequence of at least 50% (e.g. at least 60%, 65%, 70%, 75%, 80%, 85%, 90% or 95%) sequence identity to the wildtype human GLP-1.

In some embodiments, the GLP-1 comprises an amino acid sequence having at least 70% (e.g. at least 70%, 75%, 80%, 85%, 90% or 95%) sequence identity to SEQ ID NO: 1 while retaining substantial biological activity of SEQ ID NO: 1.

In some embodiments, the GLP-1 comprises an amino acid sequence of X₇X₈EGTFTS-DVSSYLEX₂₂X₂₃AAX₂₆X₂₇FIX₃₀WLVX₃₄GX₃₆G (SEQ ID NO: 2), where the X₇ is H, imidazole-4-acetate (IA), or imidazolepropionic acid (IPA); the X₈ is A, G, S, V, Aib, T, I, or L; the X₂₂ is G, or E; the X₂₃ is Q, C or K; the X₂₆ is K, R, or C; the X₂₇ is E, K, or C; the X₃₀ is A, C or K; the X₃₄ is R, K, or C; and the X₃₆ is R or G.

In some embodiments of the GLP-1, the X₇ is H; and X₈ is G or Aib.

In certain embodiments, the GLP-1 comprises an amino acid sequence selected from the group consisting of SEQ ID Nos: 1, and 3-9, which are listed in the following (genotypes are indicated following the Sequence Identification number):

SEQ ID NO: 1 (WT GLP-1)
 SEQ ID NO: 3 (8Aib, 34R, 36G),
 SEQ ID NO: 4 (8Aib, 26C, 34R, 36G),
 SEQ ID NO: 5 (8Aib, 26R, 34R, 36G),
 SEQ ID NO: 6 (8Aib, 22E, 34R, 36G),
 SEQ ID NO: 7 (8Aib, 22E, 26C, 34R, 36G),
 SEQ ID NO: 8 (8Aib, 22E, 26R, 34R, 36G),
 SEQ ID NO: 9 (8G, 22E, 26R, 34R, 36G).

In certain embodiments, the GLP-1 receptor agonist in the 10 polypeptide conjugate comprises or is GLP-1 as provided in the present disclosure. In some embodiments, the GLP-1 in the polypeptide conjugates provided herein comprises an amino acid sequence selected from the group consisting of: SEQ ID Nos: 3-9.

15 Peptide Linker

In the polypeptide conjugate disclosed herein, the biologically active peptide, i.e. GLP-1 receptor agonist (e.g. GLP-1), is attached to the N-terminus of the peptide linker.

Herein, the GLP-1 receptor agonist can optionally be 20 attached to the peptide linker via a direct linkage (e.g., a covalent bond such as a peptide bond), and the peptide linker can optionally be made up of amino acid residues linked together by peptide bonds. Optionally, the peptide linker can further comprise one or more non-natural amino acids.

25 In some embodiments of the polypeptide conjugate, the peptide linker can have a length of at least 10 amino acid residues, e.g. at least 12, 24, 32, 40, 48, 50, 60, 70, or 80 amino acid residues.

Any suitable polypeptide linkers can be used. For 30 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) 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)).

35 In some embodiments, the peptide linker consists of amino acid residues selected from the group consisting of G, Q, A, E, P, S, and T, except for the first CRM residue.

In some embodiments, the peptide linker comprises a first 40 sequence and a second sequence. The first sequence consists of one or more repeats of a repeating sequence, and is connected to N-terminus of the second sequence. The first CRM residue is in the second sequence, and can optionally be lysine residue or cysteine residue.

In certain embodiments, the repeating sequence comprises or consists of a sequence selected from the group 45 consisting of: SEQ ID Nos: 10-40, and GS.

In some embodiments, the repeating sequence has a sequence of SEQ ID NO: 13 (GAQP).

Herein optionally, the number of the one or more repeats 50 of the repeating sequence is an integer between 1 and 30. In certain embodiments, the number of the one or more repeats of the repeating sequence is selected from a group consisting of 5, 7, 9 and 11.

In certain embodiments, the second sequence of the 55 peptide linker has a sequence of SEQ ID NO: 41 (GQKP) or SEQ ID NO: 42 (GQCP).

In certain embodiments, the polypeptide linker comprises an amino acid sequence selected from the group consisting of: SEQ ID NOS: 43-50, and 82-84.

60 It is noted that in addition to the above arrangement where the peptide linker comprises a first sequence and a second sequence, where the first sequence consists of one or more

repeats of a repeating sequence, there can be other arrangements as well. In certain instances, the peptide linker may comprise or consist of more than one repeating sequence. For example, the polypeptide linker comprises or consists of 2, 3, or 4 different repeating sequences. In certain embodiments, the polypeptide linker comprises or consists of sequential or tandem repeats of the different repeating sequences.

CRM Residue

Without wishing to be bound by any theory, it is believed that CRM (or fatty acid) conjugation to GLP-1 can reduce its activity in the presence of I. For example, the reduction in GLP-1 activity has been reported for semaglutide, which has a fatty acid conjugated to K26 of GLP-1 (8Aib, 36R) and showed significant reduction in activity compared to its non-conjugated counterpart in the presence of I. Since I is present in human blood and is unavoidable under physiological conditions, reduction in GLP-1 receptor agonist activity in the presence of I is believed to compromise therapeutic activity of the protein conjugates. However, some of the polypeptide conjugates provided herein can retain most of the GLP-1 activity in the presence of I despite of CRM conjugation. In such embodiments, the polypeptide conjugates provided herein is mono-conjugated, and has the CRM conjugated to the peptide linker but not to the biologically active peptide (e.g. GLP-1).

In some of these embodiments, such mono-conjugated polypeptide conjugates provided herein has increased GLP-1 receptor agonist activity in the presence of human serum albumin (I), compared to a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1. In certain embodiments, the GLP-1 receptor agonist activity in the presence of I can be determined in an in vitro assay for GLP-1 receptor activation, either in a cell free assay such as cAMP assay or in a cell-based assay such as reporter cell assay, as known in the art. In certain embodiments, the comparative polypeptide conjugate is semaglutide. In certain embodiments, the mono-conjugated polypeptide conjugates provided herein activates GLP-1 receptor in the presence of human serum albumin (I) at an EC₅₀ no more than 50% (or no more than 40%, 30%, 20%, 10%, 5%, or 3%) than that of semaglutide in the presence of I, as determined in the same or comparable assay. In certain embodiments, the mono-conjugated polypeptide conjugates provided herein activates GLP-1 receptor in the absence of I at an EC₅₀ comparable to (e.g. from 20% to 300% of) that of semaglutide in the absence of I, as determined in the same or comparable assay. In certain embodiments, the amount of human serum albumin (I) allows assessment of impact of I binding to the GLP-1 receptor activation. In certain embodiments, the in vitro assay for GLP-1 receptor activation is conducted in the presence of at least 0.5%, 1%, 1.2%, 1.5%, 1.8% or 2% I.

In certain embodiments, the mono-conjugated polypeptide conjugates provided herein binds to I at a binding affinity (KD) comparable to (e.g. from about 70% to about 500% of) that of semaglutide, as determined in the same or comparable assay.

In certain embodiments, the mono-conjugated polypeptide conjugates provided herein has at least comparable terminal half life compared to a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1, such as semaglutide, as determined under comparable in vivo study conditions and in the same type of experimental animals. Suitable animals for terminal half life determination include for example mice, rats, minipigs, or monkeys. In certain embodiments, the

terminal half-life is determined in a suitable animal after a single dose intravenous administration or subcutaneous administration or oral administration at a dose suitable for providing therapeutic efficacy.

In certain embodiments, the polypeptide conjugates further comprise a second CRM conjugated to a second CRM residue. In other words, the present disclosure also provides double-conjugated polypeptide conjugate. It is unexpectedly found by the inventors that a second CRM conjugation can further extend the half-life of the polypeptide conjugate. In certain embodiments, the double-conjugated polypeptide conjugate having both the first and the second CRM conjugations can have an extended half-life that is at least 1-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold or even longer than a corresponding polypeptide conjugate (e.g. semaglutide) having only one CRM conjugation.

The second CRM residue can be at any suitable position. In certain embodiments, the second CRM residue is in the GLP-1 receptor agonist or in the peptide linker.

In certain embodiments, the second CRM residue is at K26 of the GLP-1 peptide. It has been reported that conjugation with two fatty acid moieties on a GLP-1 peptide can significantly reduce the activity or even render the conjugated GLP-1 peptide ineffective. It is therefore unexpectedly found by the present inventors that, conjugation both on GLP-1 (e.g. K26) and on the peptide linker at a position sufficiently away from the C-terminus of the GLP-1 peptide, can minimize the negative impact of fatty acid conjugation to the biological activity of GLP-1 polypeptide in the presence of I.

In some of these embodiments, such double-conjugated polypeptide conjugates provided herein has a comparable or acceptable GLP-1 receptor agonist activity in the presence of I, compared to a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1. In certain embodiments, the double-conjugated polypeptide conjugates provided herein activates GLP-1 receptor in the presence of I at an EC₅₀ of no more than 2000% (or no more than 1500%, 1000%, 900%, 800%, 700%, 600%, or 500%) than that of a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1, such as semaglutide in the presence of I, as determined in the same or comparable assay. In certain embodiments, the double-conjugated polypeptide conjugates provided herein activates GLP-1 receptor in the absence of I at an EC₅₀ comparable to (e.g. no more than 500%, 400%, 300%, or 200% of, or even less than) that of a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1, such as semaglutide, as determined in the same or comparable assay. In certain embodiments, the double-conjugated polypeptide conjugates provided herein activates GLP-1 receptor at a first EC₅₀ in the presence I, and at a second EC₅₀ in the absence of I, and the first EC₅₀ is higher than the second EC₅₀ but is no more than 1000-fold of that of the second EC₅₀, for example, the first EC₅₀ (in the presence of I) is no more than 900-fold, 800-fold, 700-fold, 600-fold, or 500-fold of the second EC₅₀ (in the absence of I). In certain embodiments, I is present in the in vitro assay for GLP-1 receptor activation at a suitable amount that allows assessment of impact of I binding to the GLP-1 receptor activation. In certain embodiments, the in vitro assay for GLP-1 receptor activation is conducted in the presence of a suitable amount of human serum albumin (I) that allows assessment of impact of I binding to the GLP-1 receptor activation. In

31

certain embodiments, the in vitro assay for GLP-1 receptor activation is conducted in the presence of at least 0.5%, 1%, 1.2%, 1.5%, 1.8% or 2% I.

In certain embodiments, the double-conjugated polypeptide conjugates provided herein binds to I at a higher binding affinity, represented by a KD value significantly lower than (e.g. no more than 50%, 40%, 30%, or 20% of) that of a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1, such as semaglutide, as determined in the same or comparable assay.

In certain embodiments, double-conjugated polypeptide conjugates provided herein has enhanced pharmacokinetic properties in comparison to a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the biologically active peptide (e.g. GLP-1), wherein the pharmacokinetic properties are determined by measuring blood concentration of the polypeptide conjugate after administration of a therapeutically effective dose to a subject.

In certain embodiments, the double-conjugated polypeptide conjugates provided herein has an increase in terminal half-life by at least 50%, 60%, 70%, 80%, 90%, 100%, 120%, 150%, 180%, 200%, 300%, or 400% than that of a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1, such as semaglutide, as determined under comparable in vivo study conditions and in the same type of experimental animals. Suitable animals for terminal half life determination include for example mice, rats, minipigs, or monkeys. In certain embodiments, the terminal half-life is determined in a suitable animal after a single dose intravenous administration or subcutaneous administration at a dose suitable for providing therapeutic efficacy.

In certain embodiments, the double-conjugated polypeptide conjugates provided herein has blood or plasma or serum concentrations that remain within therapeutic window for the polypeptide conjugate for a period at least about 50% longer (60%, 70%, 80%, 90%, 100%, 120%, 150%, 180%, 200%, 300%, or 400% longer) than that of a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1, such as semaglutide, as determined under comparable in vivo study conditions and in the same type of experimental animals. The term "therapeutic window" as used herein means that the range of concentration level of the polypeptide conjugates in the blood or plasma or serum that provides for therapeutic benefit or efficacy for the condition to be treated without unacceptable toxicity. The range of concentration level can be from the minimal concentration that results in a therapeutic response to the maximal concentration that provides for therapeutic response yet without inducing unacceptable toxicity.

In certain embodiments, the double-conjugated polypeptide conjugates provided herein provides for an extended duration of therapeutic efficacy in comparison with a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1, such as semaglutide, as determined under comparable in vivo study conditions and in the same type of experimental animals. Such an extended therapeutic efficacy can be characterized by area under curve (AUC) for time-response curve, which can be plotted after a single dose or after repeated doses that are suitable for providing an intended therapeutic effect in a subject having a metabolic condition (e.g. a disease model animal). The extended therapeutic efficacy duration can also be characterized by the duration of the therapeutic response.

32

In certain embodiments, the therapeutic response comprises reduction in body weight, reduction in food intake, or reduction in glucose level (fasting glucose level or non-fasting glucose level). In certain embodiments, the double-conjugated polypeptide conjugates provided herein has an increase in the AUC for the time-response curve by at least 50%, 60%, 70%, 80%, 90%, 100%, 120%, 150%, 180%, 200% than that of a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1, such as semaglutide, as determined under comparable in vivo study conditions and in the same type of experimental animals. In certain embodiments, the double-conjugated polypeptide conjugates provided herein has an increase in duration of the therapeutic response by at least 50%, 60%, 70%, 80%, 90%, 100%, 120%, 150%, 180%, 200%, 300%, or 400% than that of a comparative polypeptide conjugate without the peptide linker and having a CRM conjugated on the GLP-1, such as semaglutide, as determined under comparable in vivo study conditions and in the same type of experimental animals. In certain embodiments, the experimental animals for therapeutic response are disease model animals such as db/db mice, or diet-induced obese (DIO) animals.

CRM Residue

Within the scope of the present disclosure provided herein, different CRM residues may be used in the polypeptide conjugate depending on the different CRM used, and there can be one or two CRM residues in the polypeptide conjugate that mediate(s) the conjugation of the CRM to the polypeptide conjugate.

In certain embodiments, the polypeptide conjugate can be mono-conjugated with a CRM.

In certain embodiments, the peptide linker-residing first CRM residue is lysine residue, and the polypeptide conjugate comprises only one lysine residue. In such embodiments, the peptide linker comprises only one lysine residue, and the GLP-1 receptor agonist comprises no lysine residue, and in accordance, the naturally occurring residue found in the native GLP-1 (1-37) sequence, namely K26 and K34, can optionally be substituted to non-lysine residues (e.g. R, Q, A, G, H, S, and T, etc.). In certain embodiments, the GLP-1 comprises substitutions of K26R and K34R.

It is noted that throughout the specification, when the CRM residue (e.g. K or C) is present in GLP-1, the position of the CRM residue is identified in reference to the amino acid sequence of GLP-1 (1-37). For example, K26 indicates that the 26th position in reference to the amino acid sequence of GLP-1 (1-37) (which corresponds to the 20th position in reference to the amino acid sequence of GLP-1 (7-37), i.e. SEQ ID NO: 1) is K. Similarly, K34 indicates that the 34th position in reference to the amino acid sequence of GLP-1 (1-37) (which corresponds to the 28th position in reference to the amino acid sequence of GLP-1 (7-37), i.e. SEQ ID NO: 1) is K.

In certain embodiments, the peptide linker-residing first CRM residue is cysteine residue, and the polypeptide conjugate comprises only one cysteine residue. In such embodiments, the peptide linker comprises only one cysteine residue, and the GLP-1 receptor agonist comprises no cysteine residue. In certain embodiments, the single biologically active peptide polypeptide in the polypeptide conjugate comprises GLP-1. Native GLP-1 does not contain any cysteine residue, and therefore any GLP-1 derivative (including those provided herein) may be used as long as it does not contain a cysteine residue.

In certain embodiments, the first CRM residue is a non-natural amino acid residue, and the polypeptide conjugate

comprises only one non-natural amino acid residue as the CRM residue. Non-natural amino acid can contain a variety of functional groups or reactive groups, which can provide for additional functions and/or reactivity. Particular non-natural amino acids that are beneficial for purpose of conjugating moieties to the fusion polypeptides of the present disclosure include those with a side chain having azide, alkyne, alkene, cycloalkyne or halide.

In the mono-conjugated polypeptide conjugate, when the peptide linker-residing first CRM residue is lysine residue, the GLP-1 can comprise an amino acid sequence of SEQ ID NO: 5 or SEQ ID NO: 8. Herein, the peptide linker can optionally comprise an amino acid sequence selected from the group consisting of: SEQ ID NO: 43, SEQ ID NO: 45, and SEQ ID NO: 46. In certain embodiments, the polypeptide conjugate comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 52, 55, 56, and 60-62, which are listed in the following (the first CRM residues are indicated following the Sequence Identification number):

- SEQ ID NO: 52 (76K),
- SEQ ID NO: 55 (68K),
- SEQ ID NO: 56 (60K),
- SEQ ID NO: 60 (76K),
- SEQ ID NO: 61 (68K), or
- SEQ ID NO: 62 (60K).

In the mono-conjugated polypeptide conjugate, when the peptide linker-residing first CRM residue is cysteine residue, the GLP-1 can comprise an amino acid sequence selected from the group consisting of: SEQ ID NOS: 3, 5, 6, and 9. Herein, the peptide linker can optionally comprise an amino acid sequence selected from the group consisting of: SEQ ID NO: 47, SEQ ID NO: 49, and SEQ ID NO: 50. In certain embodiments, the polypeptide conjugate comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 66-71, and 75-80, which are listed in the following (the first CRM residues are indicated following the Sequence Identification number):

- SEQ ID NO: 66 (76C),
- SEQ ID NO: 67 (68C),
- SEQ ID NO: 68 (60C),
- SEQ ID NO: 69 (76C),
- SEQ ID NO: 70 (68C),
- SEQ ID NO: 71 (60C),
- SEQ ID NO: 75 (76C),
- SEQ ID NO: 76 (68C),
- SEQ ID NO: 77 (60C),
- SEQ ID NO: 78 (76C),
- SEQ ID NO: 79 (68C),
- SEQ ID NO: 80 (60C).

In certain embodiments, the polypeptide conjugate further comprises a second CRM residue, i.e. the polypeptide conjugate is double-conjugated, having two CRMs (i.e. the first CRM and the second CRM) respectively conjugated at the two CRM residues (i.e. the first CRM residue and second CRM residue).

In certain embodiments, the second CRM residue is in the peptide linker, just like the first CRM residue. In such embodiments, the second CRM residues can be at any suitable distance from the first CRM residue, as long as both CRM residues can be properly conjugated.

In certain embodiments, the second CRM residue is in the GLP-1 receptor agonist, for example, GLP-1. In certain embodiments, the second CRM residue is a naturally occurring residue found in the native GLP-1 sequence, or an introduced residue, for example, by substitution of a naturally occurring residue, or by insertion of a new residue.

In certain embodiments, the first and the second CRM residues are both lysine residues, and the polypeptide conjugate comprises only two lysine residues. In certain embodiments, the second lysine residue is in the GLP-1, and is selected from the group consisting of K23, K26, K27, K30, and K34. In certain embodiments, the second CRM residue is a naturally occurring residue found in the native GLP-1 sequence, for example, K26 or K34. In certain embodiments, the second CRM residue is an introduced residue. The CRM residue can be introduced into GLP-1 sequence at any suitable position, for example by substitution, as long as such substitution does not substantially diminish the GLP-1R agonist activity of the GLP-1. The second CRM residue can be introduced, for example, by substitution of Q23K, E27K, or A30K. In certain embodiments, all lysine residue(s) other than the CRM residue(s) in the GLP-1 are substituted to a non-lysine residue, so that the polypeptide conjugate comprises no additional lysine residue except for the CRM residues. For example, if K26 is the second CRM residue, then K34 is substituted to a non-lysine residue, or vice versa. Non-lysine residue can be selected by a skilled person in the art, and examples include Arginine I, Glutamine (Q), Alanine (A), Glycine (G), Histidine (H), Serine (S), or Threonine (T).

In the double-conjugated polypeptide conjugate, when both the first and the second CRM residues are lysine residue, the GLP-1 can comprise an amino acid sequence of SEQ ID NO: 3 or SEQ ID NO: 6. Herein, the peptide linker can optionally comprise an amino acid sequence selected from the group consisting of: SEQ ID NOS: 43-45, or 83. In certain embodiments, the polypeptide conjugate comprises an amino acid sequence selected from the group consisting of SEQ ID Nos: 51, 53, 54, 57-59, 85 and 86, which are listed in the following (the first and second CRM residues are indicated following the Sequence Identification number):

- SEQ ID NO: 51 (26K, 76K),
- SEQ ID NO: 53 (26K, 84K),
- SEQ ID NO: 54 (26K, 68K),
- SEQ ID NO: 57 (26K, 76K),
- SEQ ID NO: 58 (26K, 84K),
- SEQ ID NO: 59 (26K, 68K),
- SEQ ID NO: 85 (26K, 96K) or
- SEQ ID NO: 86 (26K, 60K).

In certain embodiments, the first and the second CRM residues are both cysteine residues, and the polypeptide conjugate comprises only two cysteine residues. In certain embodiments, the second cysteine residue is in the GLP-1. In certain embodiments, the second cysteine residue is in the peptide linker.

In certain embodiments, the first and the second CRM residues are both non-natural amino acid residues, and the polypeptide conjugate comprises only two non-natural amino acid residues as the CRM residues.

In certain embodiments, the cysteine residue or the non-natural amino acid residue in the GLP-1 is introduced by a substitution at the position selected from the group consisting of: Q23, K26, E27, A30, and K34, relative to SEQ ID NO: 1. In certain embodiments, the cysteine residue or the NNAA in the GLP-1 is introduced by a substitution at the position of K26 or E27, relative to SEQ ID NO: 1. In certain embodiments, the second CRM residue in the GLP-1 is cysteine and is introduced by a substitution at a position selected from the group consisting of: Q23C, K26C, E27C, A30C and K34C.

In the double-conjugated polypeptide conjugate, when both the first and the second CRM residues are cysteine

35

residue, the GLP-1 can comprise an amino acid sequence of SEQ ID NO: 4 or SEQ ID NO: 7. Herein, the peptide linker can optionally comprise an amino acid sequence selected from the group consisting of: SEQ ID NOS: 47-49. In certain embodiments, the polypeptide conjugate comprises an amino acid sequence selected from the group consisting of SEQ ID Nos: 63-65 and 72-74, which are listed in the following (the first and second CRM residues are indicated following the Sequence Identification number):

SEQ ID NO: 63 (26C, 84C),
 SEQ ID NO: 64 (26C, 76C),
 SEQ ID NO: 65 (26C, 68C),
 SEQ ID NO: 72 (26C, 84C),
 SEQ ID NO: 73 (26C, 76C),
 SEQ ID NO: 74 (26C, 68C).

CRM

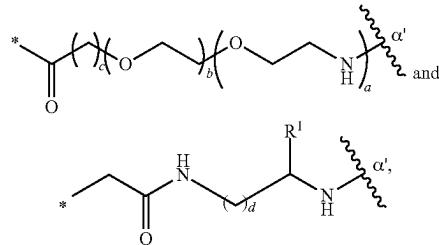
In certain embodiments, each of the first CRM and/or the second CRM (or generally CRM) comprises a plasma protein-binding moiety, a polymer, Fc, human serum albumin (HSA) and functional fragments thereof, Xten sequence, or PAS sequence. In certain embodiments, the Xten sequence is an extended recombinant polypeptide sequence with an amino acid sequence described in WO2007103515, WO2009023270, WO2010091122, WO2011123813, 20 WO2013130683, WO2017146979, WO2011084808, WO2013040093, WO2013122617, WO2014011819, WO2013184216, WO2014164568, WO2015023891, WO2016077505 and WO2017040344, disclosures of which have been incorporated by their entirety. In certain embodiments, the term "PAS", which can also be used interchangeable with the term "APS", refers to an amino acid repeats consisting of Ala, Ser, and Pro residues, as described in U.S. Pat. No. 8,563,521B2, disclosure of which has been incorporated by its entirety.

In certain embodiments, the CRM comprises an albumin-binding moiety. The term "albumin-binding moiety" refers to any functional moiety that is capable of binding albumin (e.g. human serum albumin) or any functional fragment thereof with sufficient specificity, preferably non-covalently. The albumin-binding moiety attached to a therapeutic fusion polypeptide, polypeptide, or polypeptide complex typically has an affinity below 10 μM to human serum albumin and preferably below 1 pM. The albumin-binding moiety can include, without limitation, an albumin-binding domain, an albumin-binding sequences from synthetic peptides, and an albumin-binding chemical moiety. For example, the albumin-binding moiety is selected from an albumin-binding domain from streptococcal protein G, an albumin-binding domain from *Peptostreptococcus magnus* protein PAB, an albumin-binding peptide having the core sequence DICL-PRWGCLW (SEQ ID NO: 81). A number of small peptides which are albumin binding moieties have been described in J. Biol. Chem. 277, 38 (2002) 35035-35043. For another example, the albumin-binding moiety is selected from linear and branched lipophilic moieties containing 4-40 carbon atoms, compounds with a cyclopentanophenanthrene skeleton etc. For example, the albumin-binding moiety is a group of the formula $\text{CH}_3(\text{CH}_2)_v\text{CO}-\text{NHCH}(\text{COOH})(\text{CH}_2)_2\text{CO}-$, wherein v is an integer of from 10 to 24.

In certain embodiments, the albumin-binding moiety comprises a structure of: *-A-B—C-D-E, wherein A, B, C, D and E are interconnected via amide bonds, and the * end 65 of A is connected to a reactive group of the conjugatable residue on the polypeptide complex, and wherein:

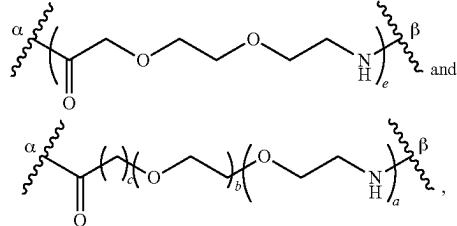
36

A is selected from a bond,



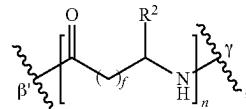
where a, b, c and d are independently an integer from 0 to 4, R¹ is hydrogen or —COOH;

B is selected from a bond,



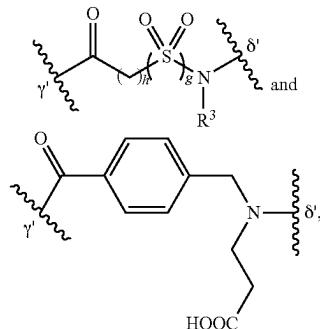
where e is an integer from 1 to 4, wherein position α is linked to position α' ;

C is a bond or



where R² is —CH₂SO₃H or —COOH, f is an integer from 1 to 4, n is an integer from 1 to 25, wherein when B is not bond, then position β' is linked to position β , or when B is bond, then position β' is linked to position α' ;

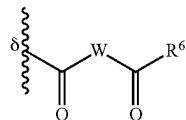
D is selected from a bond,



where g and h are independently 0 or 1, and R³ is H or —CH₂COOH, wherein:
 when B is not a bond and C is a bond, then position γ' is linked to position β ;
 when C is not a bond, then position γ' is linked to position γ ; and

37

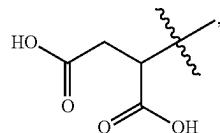
when B is a bond and C is a bond, then position γ' is linked to position α' ;
E is an acidic group having a formula:



(I)
5
10

wherein W represents $-(CR^4R^5)_I-$, where R^4 and R^5 are independently selected from the group consisting of hydrogen, halogen, cyano, hydroxyl, alkyl, alkenyl, alkynyl, cycloalkyl, heterocyclyl, aryl, heteroaryl, hydroxyalkyl, amino, aminoalkyl, carboxyl, carboxylalkyl, alkoxy, aryloxy, and carboxamide, R^6 is selected from hydroxyl or NR^7R^8 , where R^7 and R^8 are independently selected from the group consisting of hydroxyl, alkyl, hydroxyl, and

15
20



25
30

and I is an integer from 10 to 20, and wherein:

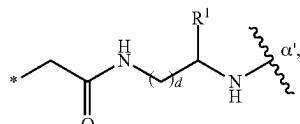
when D is not a bond, then position δ is linked to position δ' ,

when C is not a bond and D is a bond, then position δ is linked to position γ ,

when B is not a bond, C is a bond and D is a bond, then position δ is linked to position β ,

when A is not a bond, and all of B, C, and D are bond, then position δ is linked to position α' .

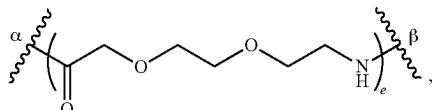
In certain embodiments,
A is a bond or



40

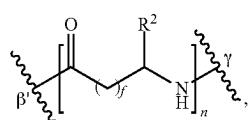
38

where d is an integer from 1 to 4, R^1 is hydrogen; B is a bond or



where e is an integer from 1 to 4, wherein position α is linked to position α' ;

C is



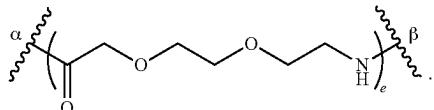
where R^2 is $-COOH$, f is an integer from 1 to 4, and n is an integer from 1 to 25, wherein: when B is not bond, then position β' is linked to position β , or when B is bond, then position β' is linked to position α' ;

D is a bond;

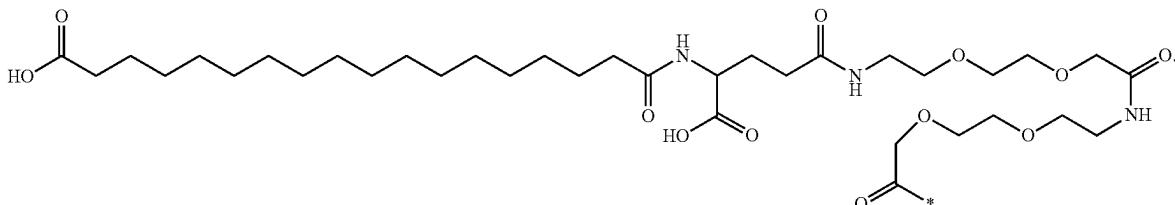
regarding E, each of R^4 and R^5 is hydrogen, R^6 is hydroxyl, I is an integer from 10 to 20, and position δ is linked to position γ .

Further optionally, d is 1, e is 2, f is 1, n is 1, and I is an integer from 14 to 20.

In certain embodiments, the CRM is conjugated to a lysine residue, optionally the lysine residue is in the peptide linker or in the GLP-1. Herein, A can be a bond and B can be

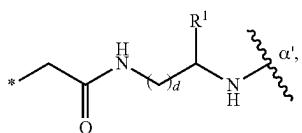


In such embodiments, the CRM comprises the structure of below formula (also referred to as $-HOOC-(CH_2)^{16}-CO-gGlu-2XADO$, where 2XADO means two consecutive ADO moieties, and ADO is short for 8-amino-3,6-dioxaoctanoic acid):

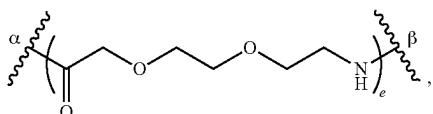


39

In certain embodiments, the CRM is conjugated to a cysteine residue, optionally the cysteine residue is in the peptide linker or in the GLP-1. Herein, A can be

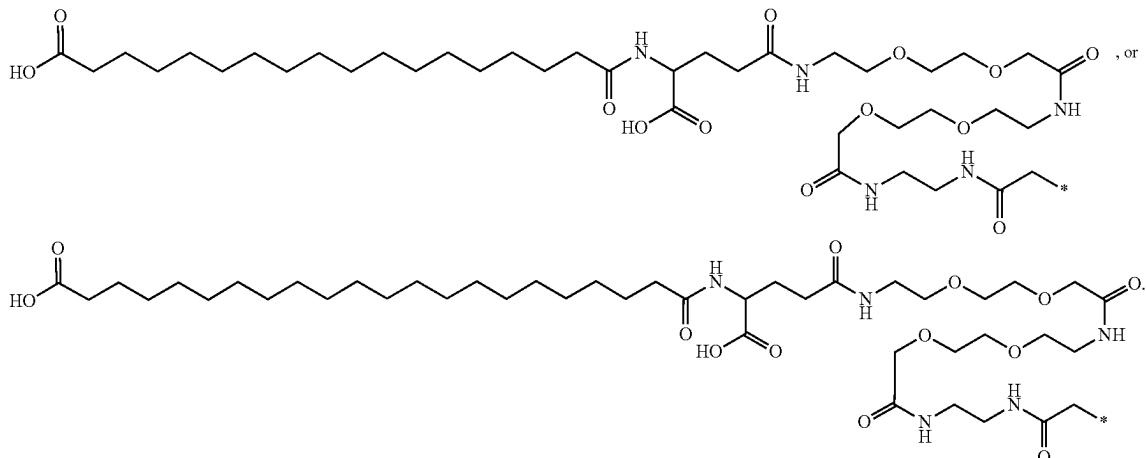


and B can be

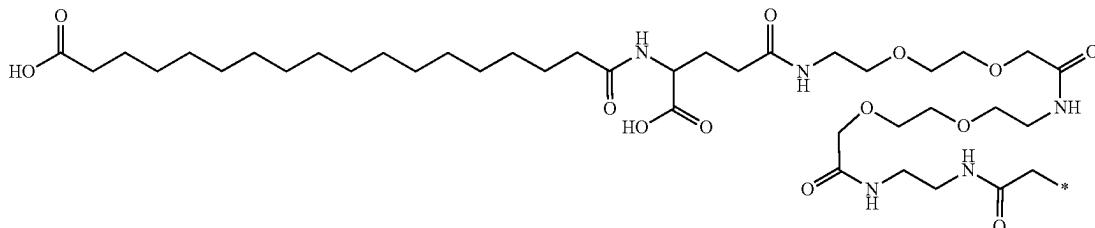


wherein position α is linked to position α' .

In such embodiments, the CRM comprises the structure of below formula (also referred to as HOOC—(CH₂)₁₆-CO-gGlu-2XADO-EDA-CO—CH₂) or HOOC—(CH₂)₂₀-CO-gGlu-2XADO-EDA-CO—CH₂):



Further according to some embodiments, the CRM comprises the structure of below formula (also referred to as HOOC—(CH₂)₁₆-CO-gGlu-2XADO-EDA-CO—CH₂):



Polypeptide Conjugates

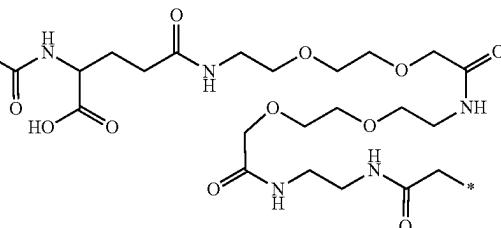
In certain embodiments, the polypeptide conjugate provided herein comprises or consists of an amino acid sequence selected from the group consisting of SEQ ID NOs: 51-80, and 85-88.

40

In certain embodiments, the polypeptide conjugate provided herein is mono-conjugated, i.e. a single CRM is conjugated at a single CRM residue in the polypeptide portion of the polypeptide conjugate. Herein, the GLP-1 comprises an amino acid sequence selected from the group consisting of: SEQ ID NOS: 3, 5, 6, 8, and 9. Herein, the polypeptide portion of the polypeptide conjugate can comprise or consist of an amino acid sequence selected from the group consisting of SEQ ID NOs: 52, 55, 56, 60-62, 66-71, and 75-80, the CRM residue can be lysine residue or cysteine residue in the peptide linker of the polypeptide portion at a position selected from a group consisting of 60, 68, and 76, and the CRM can comprise the structure of Moiety A (HOOC—(CH₂)₁₆-CO-gGlu-2XADO) for lysine conjugation, or of Moiety B (HOOC—(CH₂)₁₆-CO-gGlu-2XADO-EDA-CO—CH₂) for cysteine conjugation.

In the mono-conjugated polypeptide conjugate where the CRM residue is lysine residue, the various embodiments of the polypeptide conjugate are listed below, each indicated 20 for the amino acid sequence for the polypeptide portion and the position for the CRM residue: SEQ ID NO: 52 (76K); SEQ ID NO: 55 (68K); SEQ ID NO: 56 (60K); SEQ ID NO: 60 (76K); SEQ ID NO: 61 (68K); and SEQ ID NO: 62 (60K). In certain embodiments, the polypeptide conjugate comprises the polypeptide portion comprising or consisting

of an amino acid sequence of SEQ ID NO: 52, and the first CRM residue is 76K. In certain embodiments, the polypeptide conjugate comprises the polypeptide portion comprising



or consisting of an amino acid sequence of SEQ ID NO: 60, and the first CRM residue is 76K. Each polypeptide conjugate has Moiety A (HOOC—(CH₂)₁₆-CO-gGlu-2XADO) as the CRM.

In the mono-conjugated polypeptide conjugate where the CRM residue is cysteine residue, the various embodiments of the polypeptide conjugate are listed below, each indicated for the amino acid sequence for the polypeptide portion and the position for the CRM residue: SEQ ID NO: 66 (76C); SEQ ID NO: 67 (68C); SEQ ID NO: 68 (60C); SEQ ID NO: 69 (76C); SEQ ID NO: 70 (68C); SEQ ID NO: 71 (60C); SEQ ID NO: 75 (76C); SEQ ID NO: 76 (68C); SEQ ID NO: 77 (60C); SEQ ID NO: 78 (76C); SEQ ID NO: 79 (68C); and SEQ ID NO: 80 (60C). In certain embodiments, the polypeptide conjugate comprises the polypeptide portion comprising or consisting of an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K. Each polypeptide conjugate has Moiety B (HOOC—(CH₂)₁₆-CO-gGlu-2XADO-EDA-00—CH₂) as the CRM.

In certain embodiments, the polypeptide conjugate provided herein is double-conjugated, i.e. two CRMs are respectively conjugated at two CRM residues, one in the GLP-1 and another in the peptide linker, of the polypeptide portion of the polypeptide conjugate. Herein, the GLP-1 comprises an amino acid sequence selected from the group consisting of: SEQ ID NOS: 3, 4, 6, and 7.

Herein, the polypeptide portion of the polypeptide conjugate can comprise or consist of an amino acid sequence selected from the group consisting of SEQ ID NOS: 51, 53, 54, 57-59 and 63-65, 72-74, 85, and 86. The two CRM residues can be lysine residues or cysteine residues, with one at position 26, and another at a position selected from a group consisting of 68, 76, and 84. The two CRMs are identical, comprising the structure of Moiety A (HOOC—(CH₂)₁₆-CO-gGlu-2XADO) for lysine conjugation, or of

Moiety B (HOOC—(CH₂)₁₆-CO-gGlu-2XADO-EDA-CO—CH₂) for cysteine conjugation.

In the double-conjugated polypeptide conjugate where the two CRM residues are both lysine residues, the various embodiments of the polypeptide conjugate are listed below, each indicated for the amino acid sequence for the polypeptide portion and the position for the two CRM residues: SEQ ID NO: 51 (26K, 76K); SEQ ID NO: 53 (26K, 84K); SEQ ID NO: 54 (26K, 68K); SEQ ID NO: 57 (26K, 76K); SEQ ID NO: 58 (26K, 84K); SEQ ID NO: 59 (26K, 68K); SEQ ID NO: 85 (26K, 96K); and SEQ ID NO: 86 (26K, 60K). Each polypeptide conjugate has Moiety A (HOOC—(CH₂)₁₆-CO-gGlu-2XADO) as the two CRMs.

In the double-conjugated polypeptide conjugate where the two CRM residues are both cysteine residues, the various embodiments of the polypeptide conjugate are listed below, each indicated for the amino acid sequence for the polypeptide portion and the position for the two CRM residues: SEQ ID NO: 63 (26C, 84C); SEQ ID NO: 64 (26C, 76C); SEQ ID NO: 65 (26C, 68C); SEQ ID NO: 72 (26C, 84C); SEQ ID NO: 73 (26C, 76C); and SEQ ID NO: 74 (26C, 68C). Each polypeptide conjugate has Moiety B (HOOC—(CH₂)₁₆-CO-gGlu-2XADO-EDA-CO—CH₂) as the two CRMs.

In certain embodiments, the CRM is conjugated to a non-natural amino acid residue in the fusion polypeptide. In certain embodiments, the CRM is conjugated to a non-natural amino acid residue in the polypeptide.

Table 1 below shows the detailed information of certain exemplary polypeptide conjugates, including SEQ ID NOS of the polypeptide portion, the GLP-1, and the peptide linker, and the CRM residue. Mutations in the GLP-1, as well as the repeating sequences and number of repeats in the peptide linker sequences are also shown.

TABLE 1

Exemplary polypeptide conjugate sequences				
Mole- cule Code (MLC)	Mutations in GLP-1 ** (SEQ ID NO of the GLP-1)	Repeating sequences & No. of repeats (SEQ ID NO of the peptide linker)	CRM residue(s) [#] (SEQ ID NO of the polypeptide portion)	CRM
002	8Aib, 26R, 34R, 36G (SEQ ID NO: 5)	(GAQP)9(GQKP) (SEQ ID NO: 43)	76K (SEQ ID NO: 52)	Moiety A ^{&}
009	8Aib, 26R, 34R, 36G (SEQ ID NO: 5)	(GAQP)7(GQKP) (SEQ ID NO: 45)	68K (SEQ ID NO: 55)	Moiety A ^{&}
010	8Aib, 26R, 34R, 36G (SEQ ID NO: 5)	(GAQP)5(GQKP) (SEQ ID NO: 46)	60K (SEQ ID NO: 56)	Moiety A ^{&}
011	8Aib, 26R, 34R, 36G (SEQ ID NO: 5)	(GAQP)2(GQKP) (SEQ ID NO: 84)	48K (SEQ ID NO: 88)	Moiety A ^{&}
019	8Aib, 22E, 26R, 34R, 36G (SEQ ID NO: 8)	(GAQP)9(GQKP) (SEQ ID NO: 43)	76K (SEQ ID NO: 60)	Moiety A ^{&}
020	8Aib, 22E, 26R, 34R, 36G (SEQ ID NO: 8)	(GAQP)7(GQKP) (SEQ ID NO: 45)	68K (SEQ ID NO: 61)	Moiety A ^{&}
021	8Aib, 22E, 26R, 34R, 36G (SEQ ID NO: 8)	(GAQP)5(GQKP) (SEQ ID NO: 46)	60K (SEQ ID NO: 62)	Moiety A ^{&}
041	8Aib, 26R 34R, 36G (SEQ ID NO: 5)	(GAQP)9(GQCP) (SEQ ID NO: 47)	76C (SEQ ID NO: 66)	Moiety B ^{&&}
042	8Aib, 26R, 34R, 36G (SEQ ID NO: 5)	(GAQP)7(GQCP) (SEQ ID NO: 49)	68C (SEQ ID NO: 67)	Moiety B ^{&&}
043	8Aib, 26R, 34R, 36G (SEQ ID NO: 5)	(GAQP)5(GQCP) (SEQ ID NO: 50)	60C (SEQ ID NO: 68)	Moiety B ^{&&}
052	8G, 22E, 26R, 34R, 36G (SEQ ID NO: 9)	(GAQP)9(GQCP) (SEQ ID NO: 47)	76C (SEQ ID NO: 69)	Moiety B ^{&&}
053	8G, 22E, 26R, 34R, 36G (SEQ ID NO: 9)	(GAQP)7(GQCP) (SEQ ID NO: 49)	68C (SEQ ID NO: 70)	Moiety B ^{&&}
054	8G, 22E, 26R, 34R, 36G (SEQ ID NO: 9)	(GAQP)5(GQCP) (SEQ ID NO: 50)	60C (SEQ ID NO: 71)	Moiety B ^{&&}

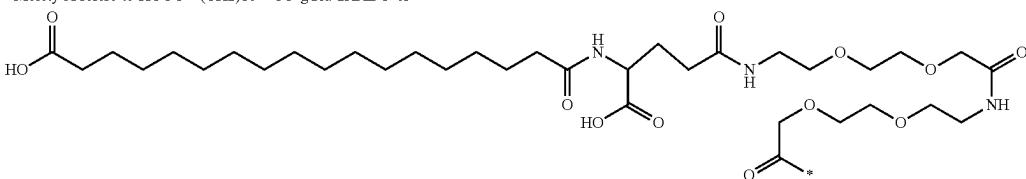
TABLE 1-continued

Exemplary polypeptide conjugate sequences				
Mole-cule Code (MLC)	Mutations in GLP-1 ** (SEQ ID NO of the GLP-1)	Repeating sequences & No. of repeats (SEQ ID NO of the peptide linker)	CRM residue(s) [#] (SEQ ID NO of the poly peptide portion)	CRM
065	8Aib, 34R, 36G (SEQ ID NO: 3)	(GAQP)9(GQCP) (SEQ ID NO: 47)	76C (SEQ ID NO: 75)	Moiety B&&
066	8Aib, 34R, 36G (SEQ ID NO: 3)	(GAQP)7(GQCP) (SEQ ID NO: 49)	68C (SEQ ID NO: 76)	Moiety B&&
067	8Aib, 34R, 36G (SEQ ID NO: 3)	(GAQP)5(GQCP) (SEQ ID NO: 50)	60C (SEQ ID NO: 77)	Moiety B&&
068	8Aib, 22E, 34R, 36G (SEQ ID NO: 6)	(GAQP)9(GQCP) (SEQ ID NO: 47)	76C (SEQ ID NO: 78)	Moiety B&&
069	8Aib, 22E, 34R, 36G (SEQ ID NO: 6)	(GAQP)7(GQCP) (SEQ ID NO: 49)	68C (SEQ ID NO: 79)	Moiety B&&
070	8Aib, 22E, 34R, 36G (SEQ ID NO: 6)	(GAQP)5(GQCP) (SEQ ID NO: 50)	60C (SEQ ID NO: 80)	Moiety B&&
001	8Aib, 34R, 36G (SEQ ID NO: 3)	(GAQP)9(GQKP) (SEQ ID NO: 43)	26K, 76K (SEQ ID NO: 51)	Moiety A&
004	8Aib, 34R, 36G (SEQ ID NO: 3)	(GAQP)14(GQKP) (SEQ ID NO: 82)	26K, 96K (SEQ ID NO: 85)	Moiety A&
005	8Aib, 34R, 36G (SEQ ID NO: 3)	(GAQP)11(GQKP) (SEQ ID NO: 44)	26K, 84K (SEQ ID NO: 53)	Moiety A&
006	8Aib, 34R, 36G (SEQ ID NO: 3)	(GAQP)7(GQKP) (SEQ ID NO: 45)	26K, 68K (SEQ ID NO: 54)	Moiety A&
007	8Aib, 34R, 36G (SEQ ID NO: 3)	(GAQP)5(GQKP) (SEQ ID NO: 83)	26K, 60K (SEQ ID NO: 86)	Moiety A&
008	8Aib, 34R, 36G (SEQ ID NO: 3)	(GAQP)2(GQKP) (SEQ ID NO: 84)	26K, 48K (SEQ ID NO: 87)	Moiety A&
012	8Aib, 22E, 34R, 36G (SEQ ID NO: 6)	(GAQP)9(GQKP) (SEQ ID NO: 43)	26K, 76K (SEQ ID NO: 57)	Moiety A&
015	8Aib, 22E, 34R, 36G (SEQ ID NO: 6)	(GAQP)11(GQKP) (SEQ ID NO: 44)	26K, 84K (SEQ ID NO: 58)	Moiety A&
016	8Aib, 22E, 34R, 36G (SEQ ID NO: 6)	(GAQP)7(GQKP) (SEQ ID NO: 45)	26K, 68K (SEQ ID NO: 59)	Moiety A&
036	8Aib, 26C, 34R, 36G (SEQ ID NO: 4)	(GAQP)11(GQCP) (SEQ ID NO: 48)	26C, 84C (SEQ ID NO: 63)	Moiety B&&
037	8Aib, 26C, 34R, 36G (SEQ ID NO: 4)	(GAQP)9(GQCP) (SEQ ID NO: 47)	26C, 76C (SEQ ID NO: 64)	Moiety B&&
038	8Aib, 26C, 34R, 36G (SEQ ID NO: 4)	(GAQP)7(GQCP) (SEQ ID NO: 49)	26C, 68C (SEQ ID NO: 65)	Moiety B&&
060	8Aib, 22E, 26C, 34R, 36G (SEQ ID NO: 7)	(GAQP)11(GQCP) (SEQ ID NO: 48)	26C, 84C (SEQ ID NO: 72)	Moiety B&&
061	8Aib, 22E, 26C, 34R, 36G (SEQ ID NO: 7)	(GAQP)9(GQCP) (SEQ ID NO: 47)	26C, 76C (SEQ ID NO: 73)	Moiety B&&
062	8Aib, 22E, 26C, 34R, 36G (SEQ ID NO: 7)	(GAQP)7(GQCP) (SEQ ID NO: 49)	26C, 68C (SEQ ID NO: 74)	Moiety B&&

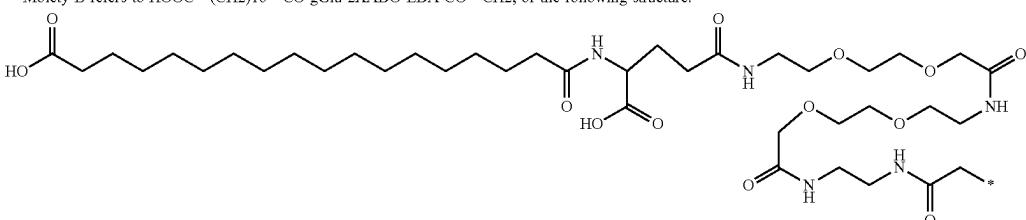
** Mutations in GLP-1 means mutations relative to SEQ ID NO: 1, wherein the first residue is 7H, and the last residue is 37G ;

#CRM residue(s) means position of the CRM residue counted from N to C in the polypeptide sequence containing GLP-1 (wherein the first residue is 7H) with its C-terminus attached to the peptide linker.

& Moiety A refers to HOOC—(CH₂)₁₆—CO-g-Glu-2XADO or



&&Moiety B refers to HOOC—(CH₂)₁₆—CO-g-Glu-2XADO-EDA-CO—CH₂, or the following structure:

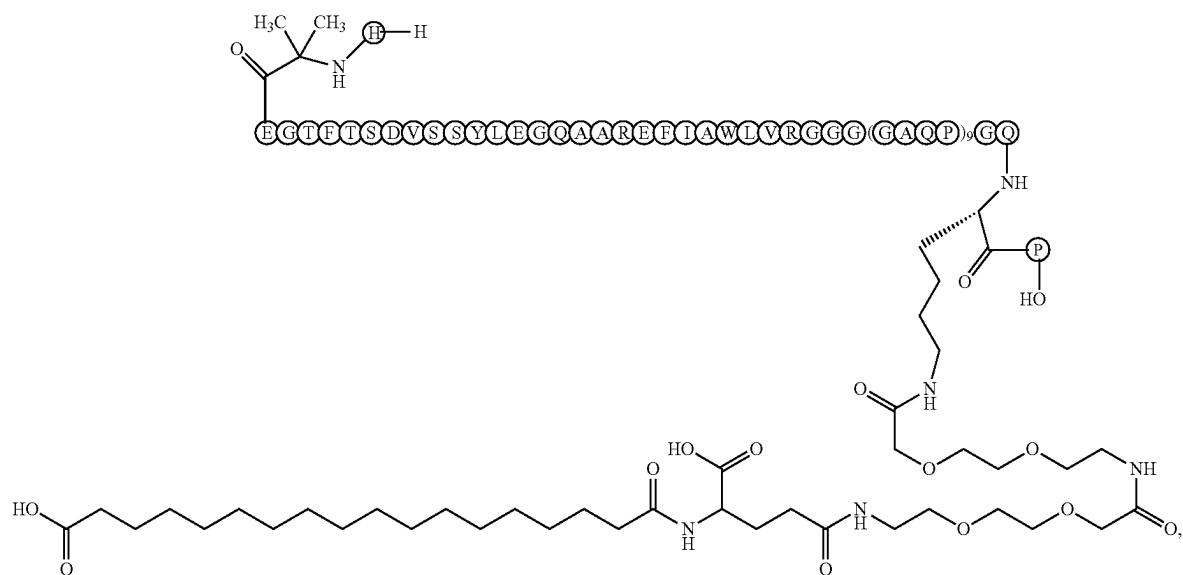


45

In certain embodiments, the polypeptide conjugate comprises a structure shown below (including the mono-conjugated Molecule 002 and Molecule 019, and the double-conjugated Molecule 012), where the amino acid residues are represented as one letter abbreviations in circles:

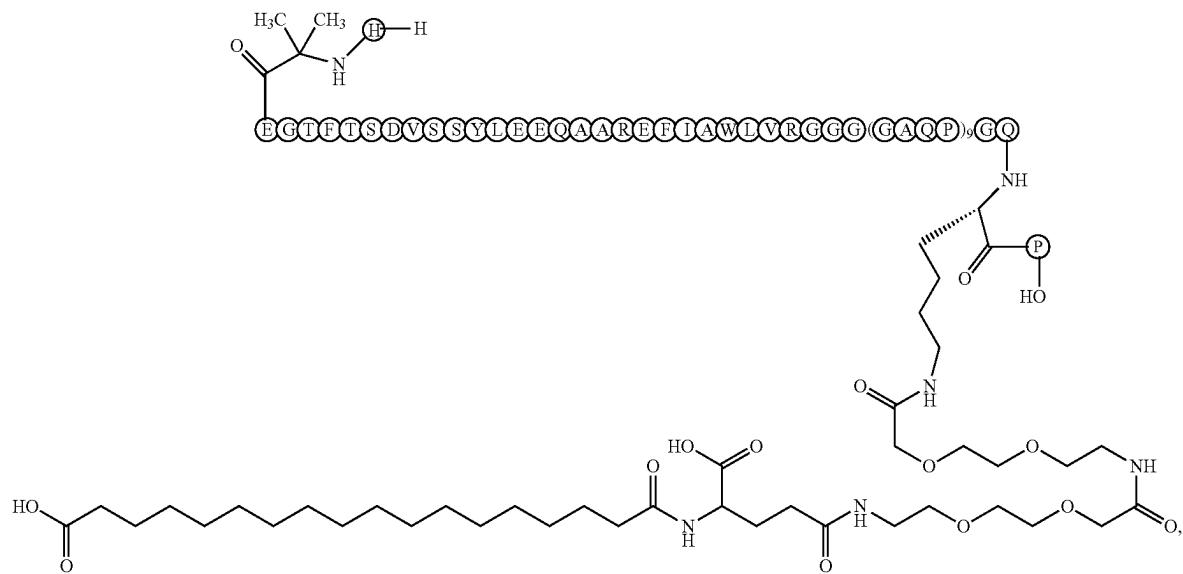
46

Molecule 002



or

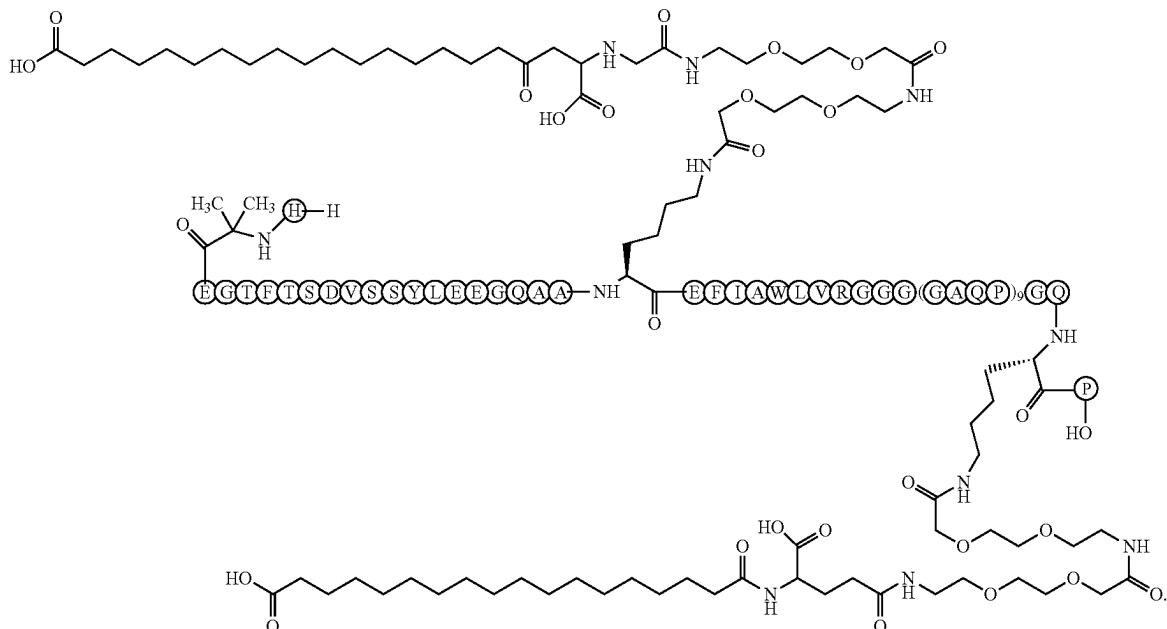
Molecule 019



-continued

or

Molecule 012



30

Method of Preparation

The present disclosure provides isolated nucleic acids or polynucleotides that encode the polypeptide portion (or a fragment thereof) of the polypeptide conjugates provided herein.

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)).

The nucleic acids or polynucleotides encoding the polypeptide provided herein (or fragment thereof) can be constructed using recombinant techniques. To this end, DNA encoding the GLP-1 receptor agonist (such as GLP-1) and DNA encoding the peptide linker can be obtained and operably linked to allow transcription and expression in a host cell to produce the fusion polypeptide. If needed, polynucleotide sequences encoding for one or more linkers are also operably linked to allow expression of the desired product.

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 or polypeptide complexes is feasible and under proper control.

The encoding polynucleotide sequence(s) can be inserted 35 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 40 more marker genes, an enhancer element, a promoter (e.g. prokaryotic promoters such as T7, T7lac, Sp6, araBAD, trp, lac, tac, pLM, A3, lac, Ipp, npr, pac, syn, trc and T3, or eukaryotic promoters such as SV40, CMV, and EF-1 α), and a transcription termination sequence.

45 Vectors and Host Cells

In another aspect, the present disclosure provides a vector comprising the polynucleotide provided herein.

Vectors comprising the polynucleotide sequence(s) provided herein can be introduced to a host cell for cloning or 50 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 55 prokaryotic cell or eukaryotic cell.

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 marcescens*, 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).

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 GLP-1 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.

In another aspect, the present disclosure provides a host cell comprising the vector provided herein. The host cell is prokaryotic cell or a 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.

In another aspect, the present disclosure provides a method of producing the polypeptide conjugates provided herein, comprising culturing the host cell provided herein under a condition that allows expression of the polynucleotide provided herein to obtain the polypeptide portion of the polypeptide conjugate.

For production of the polypeptide portion provided 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™ 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.

In one aspect, the present disclosure provides a method of expressing the polypeptide portion of the polypeptide conjugates provided herein, comprising culturing the host cell provided herein under the condition at which the polypeptide portion is expressed. In certain embodiments, the polypeptide portion is expressed as a soluble polypeptide.

When using recombinant techniques, the polypeptide provided 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 com-

mercially 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.

In certain embodiments, the method further comprises isolating the polypeptide.

The polypeptide provided 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.

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™ 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.

In certain embodiments, the method further comprises conjugating the CRM to the polypeptide. The polypeptide can be conjugated at, for example, the lysine residue, the cysteine residue, or the non-natural amino acid by a suitable conjugation reaction.

For example, the polypeptide having one or more CRM residues such as lysine may be reacted with an amino-reactive agent. In certain embodiments, the CRM is conjugated to the lysine residue via an acyl group in an acylation reaction. Exemplary methods of acylation reaction is described in, for example, WO2009083549 and WO2010029159, the content of which is incorporated herein to its entirety. The CRM to be conjugated in an acylation reaction may contain a carboxylic acid group, an α,ω -fatty diacid residue, an activated ester, or an activated N-hydroxy imide ester, among others. Examples of activated esters include, O-succinimide reagents like N-hydroxysuccinimidyl (NHS) or sulfo-NHS esters and imido ester compounds like Traut's reagent, which can react with the ϵ -amino group of a conjugatable lysine residue to form amide or amidine bonds. Additional examples of suitable amino-reactive agent include, O-acylisourea, N-hydroxy triazole esters, anhydride, phenyl active esters, P-hydroxamic active esters, acylimidazoles, acylbenzotriazoles, acyl azides, acid hydrides, phosphonium salts, aminium/uronium salts.

For another example, the polypeptide having one or more CRM residues such as cysteine may be linked to a thiol-reactive agent. In certain embodiments, the CRM is conjugated to the cysteine residue in an alkylation reaction. In certain embodiments, the CRM is conjugated to the conjugatable cysteine residue via a maleimide or an iodoacetamide to form a carbon-sulfur bond. In certain embodiments, the CRM is conjugated to the conjugatable cysteine residue via a disulphide to form a disulfide bond. Additional examples of suitable thiol-reactive group include, dienyl sulfone, α -haloacyl, or other thiol-reactive conjugation partner. See, for details, Haugland, 2003, Molecular Probes Handbook of Fluorescent Probes and Research Chemicals, Molecular Probes, Inc.; Brinkley, 1992, Bioconjugate Chem. 3:2; Garman, 1997, Non-Radioactive Labelling: A Practical Approach, Academic Press, London; Means (1990) Bioconjugate Chem. 1:2; Hermanson, G. in Bioconjugate Techniques (1996) Academic Press, San Diego, pp. 40-55, 643-671.

For example, the polypeptide having one or more CRM residues such as non-natural amino acid residue (NNAA) can be conjugated to a CRM such that a stable linkage can

51

be formed between the NNAA of the fusion polypeptide and the CRM. For example, NNAAAs containing keto group or aldehyde or β -diketomoiety can react with a hydrazide- or O-alkylhydroxylamine-, hydroxylamine-containing agents to form a hydrazone or an O-alkylated oxime linkage. For another example, NNAAAs containing an azide group can react with an alkyne derivative to form a stable triazole linker by copper (I) catalyzed [3+2] cycloaddition (and vice versa). For another example, NNAAAs containing an azide group can be ligated with an appropriate water soluble phosphine-containing agent to form an amide linkage by a Staudinger ligation. Further, a thioester moiety in an NNAA can react with an amine-containing agent to form amide linkage. The fusion polypeptides provided herein incorporated with an NNAA can be conjugated with an agent via cycloaddition reactions, such as (4+2) cycloaddition between diene and dienophile (Diels-Alder reaction), (3+2) cycloaddition via 1, 3-dipolar Huisgen cycloaddition, and (3+2) cycloaddition via Nitrene-olefin cycloaddition. Cycloaddition methods suitable for antibody conjugation have been described in, for example, WO05003294, US20120004183, WO06009901, WO07130453 and U.S. Pat. No. 6,737,236.

For another example, the polypeptide may be conjugated to biotin, then indirectly conjugated to CRM that is conjugated to avidin. For still another example, the fusion polypeptide or the polypeptide complex may be linked to a coupling agent which further links to the CRM. Examples of the coupling agents include bifunctional moieties such as N-succinimidyl-3-(2-pyridylidithio) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCl), active esters (such as disuccinimidyl suhereate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as toluene 2,6-diisocyanate), and his-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). Particularly preferred coupling agents include N-succinimidyl-3-(2-pyridylidithio) propionate (SPDP) (Carlsson et al., Biochem. J. 173:723-737 (1978)) and N-succinimidyl-4-(2-pyridylthio)pentanoate (SPP) to provide for a disulphide linkage.

Additional methods for the conjugation of CRM to the polypeptides are found, for example, in U.S. Pat. Nos. 5,208,020; 6,411,163; WO2005037992; WO2005081711; and WO2006/034488, which are incorporated herein by reference to the entirety. Specific examples of methods of preparing the conjugates of the present disclosure are also included in the experimental part of the present disclosure.

Pharmaceutical Composition

In another aspect, the present disclosure also provides a pharmaceutical composition comprising the polypeptide conjugate provided herein. In certain embodiments, the pharmaceutical composition is liquid composition for parenteral administration.

Herein, the polypeptide conjugate substantially serves as the active pharmaceutical ingredient (API) in the pharmaceutical composition. In addition to the polypeptide conjugate, the pharmaceutical composition further comprises one or more pharmaceutically acceptable excipients.

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

52

A "pharmaceutically acceptable excipient" refers to an ingredient in a pharmaceutical formulation, other than an active ingredient, which is bioactivity acceptable and non-toxic to a subject.

Examples of pharmaceutical acceptable excipients for liquid formulation may include, pharmaceutically acceptable liquid, gel, solid carriers, aqueous vehicles, nonaqueous vehicles, antimicrobial agents, isotonic agents, buffers, anti-oxidants, anesthetics, suspending/dispersing agents, sequestering or chelating agents, diluents, adjuvants, excipients, or non-toxic auxiliary substances, other components known in the art, or various combinations thereof.

Suitable pharmaceutically-acceptable excipients for a liquid formulation 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 carboxymethylcellulose, hydroxypropyl methylcellulose, or polyvinylpyrrolidone, emulsifying agents such as Polysorbate 80 (TWEEN-80), sequestering or chelating agents such as EDTA (ethylene-diaminetetraacetic 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.

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.

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.

In certain embodiments, a sterile, lyophilized powder is prepared by dissolving the polypeptide 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 fusion 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.

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.

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.

In certain embodiments, the polypeptide conjugates 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.

Liquid Formulation

In another aspect, the pharmaceutical composition provided herein is a liquid formulation. In certain embodiments, the liquid formulation is an aqueous solution. An aqueous solution can contain, for example, at least 70% w/w, at least 75% w/w, at least 80%, at least 85% w/w, at least 90% w/w, or at least 95% w/w water.

In certain embodiments, the pharmaceutical composition provided herein comprising a polypeptide conjugate and a pharmaceutically acceptable excipient.

In certain embodiments, the pharmaceutically acceptable excipient comprises a buffer and an isotonic agent.

In certain embodiments, the buffer is selected from the group consisting of a phosphate buffer, citrate buffer, acetate buffer, histidine buffer, glycine buffer, carbonate buffer, borate buffer, glutamate buffer, glycylglycine buffer, lysine buffer, and arginine buffer.

In certain embodiments, the buffer is a phosphate buffer.

In certain embodiments, the phosphate buffer is present in a concentration of 0.01-50 mM of the pharmaceutical composition. For example, the phosphate buffer is present in a concentration of 0.01-45 mM, 0.01-40 mM, 0.01-35 mM, 0.01-30 mM, 0.01-25 mM, 0.01-20 mM, 0.01-15 mM, 0.1-30 mM, 1-25 mM, 1-20 mM, 1-15 mM, 5-30 mM, 5-25 mM, 5-20 mM, or 5-15 mM of the pharmaceutical composition.

In certain embodiments, the phosphate buffer is present in a concentration of 5-20 mM or 5-10 mM of the pharmaceutical composition, optionally about 10 mM, 9 mM, 8 mM, 7 mM, 6 mM or 5 mM.

5 In certain embodiments, the phosphate buffer is selected from the group consisting of sodium dihydrogen phosphate, disodiumhydrogen phosphate, sodium phosphate, or a hydrate thereof.

10 In certain embodiments, the hydrate is dodecahydrate or dihydrate.

In certain embodiments, the phosphate buffer is disodiumhydrogen phosphate dodecahydrate or disodium phosphate dihydrate.

15 In certain embodiments, the disodiumhydrogen phosphate dodecahydrate is present in a concentration of about 0.1 to 15 mg/mL, 0.5 to 15 mg/mL, 1 to 15 mg/mL, 0.5 to 12 mg/mL, 0.5 to 10 mg/mL, 0.5 to 8 mg/mL, 0.5 to 7 mg/mL, or 0.5 to 5 mg/mL of the pharmaceutical composition. In certain embodiments, the disodiumhydrogen phosphate dodecahydrate is present in a concentration of about 2.87 mg/mL of the pharmaceutical composition.

20 In certain embodiments, the disodium phosphate dihydrate is present in a concentration of about 0.1 to 15 mg/mL, 0.5 to 15 mg/mL, 1 to 15 mg/mL, 0.5 to 12 mg/mL, 0.5 to 10 mg/mL, 0.5 to 8 mg/mL, 0.5 to 7 mg/mL, or 0.5 to 5 mg/mL of the pharmaceutical composition. In certain embodiments, the disodium phosphate dihydrate is present in a concentration of about 1.42 mg/mL.

25 30 In certain embodiments, the buffer is a citrate buffer.

In certain embodiments, the citrate buffer is present in a concentration of about 1-50 mM.

In certain embodiments, the citrate buffer comprises a mixture of citric acid anhydrous and trisodium citrate (or a hydrate thereof, e.g. trisodium citrate dihydrate).

35 In certain embodiments, the citrate buffer comprises a mixture of 0.14 mg/mL citric acid anhydrous and 2.74 mg/mL trisodium citrate dihydrate.

In certain embodiments, the buffer is a histidine buffer.

40 In certain embodiments, the histidine buffer is present in a concentration of about 1-70 mM of the pharmaceutical composition, optionally about 5-50 mM, 5-20 mM or 5-10 mM, or optionally about 10 mM, 9 mM, 8 mM, 7 mM, 6 mM or 5 mM.

45 In certain embodiments, the histidine buffer is present in a concentration of about 0.5-10 mg/mL, 0.5-5 mg/mL, 1-5 mg/mL, or 1-3 mg/mL of the pharmaceutical composition.

In certain embodiments, the histidine buffer is present in a concentration of about 1.24 mg/mL of the pharmaceutical composition.

50 In certain embodiments, the isotonic agent is selected from the group consisting of sodium chloride, propylene glycol, sorbitol, sucrose, glycine, mannitol, lactose monohydrate, arginine, myoinositol and dimethylsulfone.

55 In certain embodiments, the isotonic agent is sodium chloride.

In certain embodiments, the sodium chloride is about 1-20 mg/mL, 1-18 mg/mL, 1-16 mg/mL, 1-15 mg/mL, 2-20 mg/mL, 3-20 mg/mL, 4-20 mg/mL, or 5-20 mg/mL. In certain embodiments, the sodium chloride is about 5-15 mg/mL.

In certain embodiments, the isotonic agent is propylene glycol.

60 65 In certain embodiments, the propylene glycol is about 1 mg/mL to about 50 mg/mL (e.g. about 5 mg/mL to about 25 mg/mL, about 8 mg/mL to about 16 mg/mL).

In certain embodiments, the isotonic agent is mannitol.

55

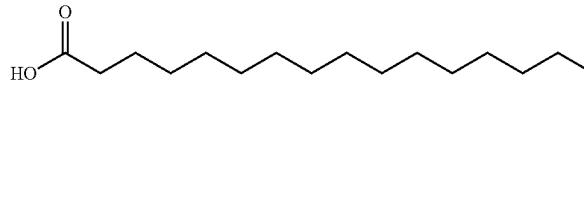
In certain embodiments, the mannitol is about 20 mg/mL to about 60 mg/mL (e.g. about 25 mg/mL to about 50 mg/mL, about 30 mg/mL to about 50 mg/mL, about 35 mg/mL to about 50 mg/mL).

In certain embodiments, the pharmaceutical excipient further comprises a preservative, a chelating agent, and/or a stabilizer.

In certain embodiments, the pharmaceutical composition has a pH of about 6.5 to about 8.3 (e.g. about 6.5 to 7.4, or from 7.4 to 8.3).

In certain embodiments, the pharmaceutical composition has a pH of about 7.4.

In certain embodiments, the polypeptide conjugate is double-conjugated. In certain embodiments of the pharmaceutical composition, the polypeptide conjugate is double-conjugated, and the conjugate portion of the polypeptide conjugate further comprises a second CRM conjugated to a second CRM residue in the polypeptide portion. Optionally, the second CRM residue is lysine residue, and further optionally, each of the first CRM and the second CRM comprises Moiety A (HOOC—(CH₂)₁₆-CO-gGlu-2XADO), the Moiety A having the structure of below formula:



In some embodiments of the pharmaceutical composition, the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K. In certain embodiments, the pharmaceutical composition has about 1-80 mg/ml, 1-100 mg/ml or 1-120 mg/ml, or optionally, 5-40 mg/ml, 5-80 mg/ml, 5-100 mg/ml, or 5-120 mg/ml of the polypeptide conjugate (e.g. about 5-90 mg/mL, about 5-70 mg/mL, about 5-60 mg/mL, about 5-50 mg/mL, about 5-30 mg/mL, about 5-20 mg/mL, about 5-10 mg/mL).

In certain embodiments, the pharmaceutical compositions comprise:

- (a) about 1-80 mg/mL, 1-100 mg/ml or 1-120 mg/mL of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;
- (b) a buffer selected from the group consisting of phosphate buffer, citrate buffer, and histidine buffer;
- (c) an isotonic agent selected from the group consisting of sodium chloride, propylene glycol, mannitol; and
- (d) a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- (a) about 1-80 mg/mL, 1-100 mg/ml or 1-120 mg/mL of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;

56

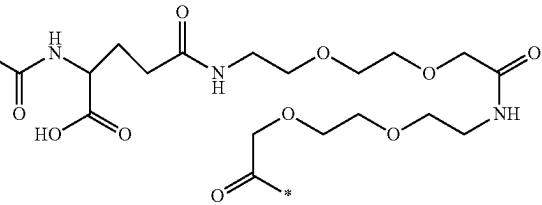
- (b) a buffer selected from the group consisting of phosphate buffer, citrate buffer, and histidine buffer;
- (c) an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and
- (d) a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- (a) about 1-80 mg/mL, 1-100 mg/ml or 1-120 mg/mL of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;
- (b) about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
- (c) an isotonic agent selected from the group consisting of sodium chloride, propylene glycol, mannitol; and
- (d) a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- (a) about 1-80 mg/mL, 1-100 mg/ml or 1-120 mg/mL of the polypeptide conjugate, wherein the polypeptide



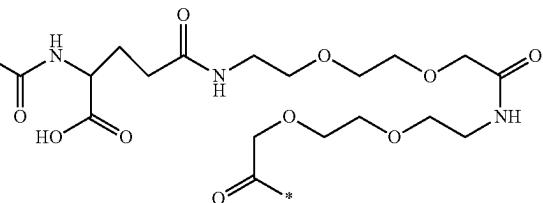
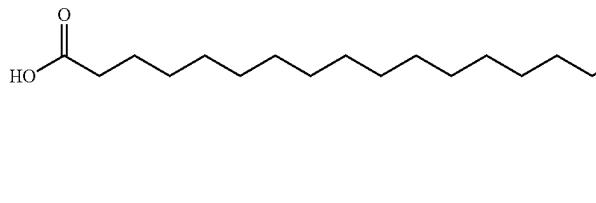
portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;

- (b) about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
- (c) an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and
- (d) a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- (a) about 5-80 mg/mL, 5-100 mg/ml or 5-120 mg/mL of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;
- (b) about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
- (c) an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and
- (d) a pH of about 6.5 to about 8.3.

In certain embodiments, the polypeptide conjugate is single-conjugated, and the first CRM is the only one CRM in the conjugate portion of the polypeptide conjugate. Optionally, the first CRM residue is lysine residue, and further optionally, and the first CRM comprises Moiety A (HOOC—(CH₂)₁₆-CO-gGlu-2XADO), the Moiety A having the structure of below formula:



In certain embodiments of the pharmaceutical composition, the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60, and is conjugated with the first CRM at 76K. In certain embodiments of the pharmaceutical composition, the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 52, and is conjugated with the first CRM at 76K. In certain embodiments, the pharmaceutical composition has about 0.5-20 mg/ml, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate.

In certain embodiments, the pharmaceutical compositions comprise:

- (a) about 0.5-20 mg/mL, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60 or SEQ ID NO: 52, and is conjugated with the first CRM at 76K;
- (b) a buffer selected from the group consisting of phosphate buffer, citrate buffer, and histidine buffer;
- (c) an isotonic agent selected from the group consisting of sodium chloride, propylene glycol, mannitol; and
- (d) a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- (a) about 0.5-20 mg/mL, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60 or SEQ ID NO: 52, and is conjugated with the first CRM at 76K;
- (b) about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
- (c) an isotonic agent selected from the group consisting of sodium chloride, propylene glycol, mannitol; and
- (d) a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- (a) about 0.5-20 mg/mL, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60 or SEQ ID NO: 52, and is conjugated with the first CRM at 76K;
- (b) a buffer selected from the group consisting of phosphate buffer, citrate buffer, and histidine buffer;
- (c) an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and
- (d) a pH of about 6.5 to about 8.3.

In certain embodiments, the pharmaceutical compositions comprise:

- (a) about 0.5-20 mg/mL, 0.5-50 mg/ml or 0.5-80 mg/ml of the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 60 or SEQ ID NO: 52, and is conjugated with the first CRM at 76K;

- (b) about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer;
- (c) an isotonic agent selected from the group consisting of 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and
- (d) a pH of about 6.5 to about 8.3.

Method of Treatment

In another aspect, the present disclosure provides a method of preventing or treating a metabolic disorder in a subject in need thereof, comprising administering a therapeutically effective amount of the polypeptide conjugate or a pharmaceutical composition provided herein.

Therapeutic methods are also provided, comprising: administering a therapeutically effective amount of the polypeptide conjugate provided herein to a subject in need thereof, thereby treating or preventing a condition or a disorder. In certain embodiments, the subject has been identified as having a disorder or condition likely to respond to the polypeptide conjugate provided herein.

In certain embodiments, the metabolic disorder is diabetes, obesity, overweight, non-alcoholic steatohepatitis (NASH), cardiovascular like dyslipidemia, atherosclerosis, 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, or primary biliary cirrhosis (PBC), or Alzheimer's disease.

In certain embodiments, the condition diabetes includes all forms of diabetes, such as hyperglycemia, type 2 diabetes, impaired glucose tolerance, type 1 diabetes, non-insulin dependent diabetes, MODY (maturity onset diabetes of the young), gestational diabetes, and/or an elevated level of HbA1C.

In certain embodiments, the condition diabetes includes diabetic complications such as angiopathy.

In certain embodiments, the present disclosure provides a method of preventing or treating Alzheimer's disease in a subject in need thereof, comprising administering a therapeutically effective amount of the polypeptide conjugate or a pharmaceutical composition provided herein. The term "Alzheimer's disease" as used herein is intended to encompass to all forms of and stages of Alzheimer's disease. Alzheimer's disease is a progressive, degenerative brain disorder that gradually impairs memory, cognitive function, and behavior. There are five stages of Alzheimer's disease, which include preclinical Alzheimer's disease, mild cognitive impairment caused by Alzheimer's disease, mild dementia resulting from Alzheimer's disease, moderate dementia linked to Alzheimer's disease, and severe dementia associated with Alzheimer's disease. The term "dementia" is used to refer to a collection of symptoms that significantly impact intellectual and social capabilities, ultimately interfering with daily functioning. There are the two predominant forms of the Alzheimer's disease: Familial

Alzheimer's disease and Sporadic or Late onset Alzheimer's disease. Familial Alzheimer's disease is typically caused by inherited genetic mutations, including mutations in one of three genes: APP, PSEN1 or PSEN2. This form of the disease is a rare and devastating illness with onset potentially occurring in mid-life. The second and far more common form of the disease is Sporadic or Late onset Alzheimer's disease, which is the most common form and occurs without any known genetic cause. In some embodiments, the subject has or suffer from Alzheimer's disease may have the symptoms selected from the group of cognitive deficits including memory impairment, language dysfunction, and visuospatial skills; functional impairment that may span occupational and social issues (e.g., activities of daily living); and behavioral symptoms including depression, anxiety, aggression and psychosis may also appear as the disease progresses in severity.

In certain embodiments, the subject has preclinical Alzheimer's disease, mild cognitive impairment caused by Alzheimer's disease, mild dementia resulting from Alzheimer's disease, moderate dementia linked to Alzheimer's disease, or severe dementia associated with Alzheimer's disease.

In certain embodiments, the subject has Familial Alzheimer's disease or Sporadic Alzheimer's disease.

It has been reported that conjugated GLP-1 agonist, for example, semaglutide, can be useful for treating Alzheimer's disease, see for example, published U.S. Patent Application US2022/0280612A1, which is incorporated herein to its entirety. It has been reported that GLP-1 agonist semaglutide reduced measures of neuro-inflammation which may affect cognition and function in animal studies relevant to Alzheimer's disease. Furthermore, real-world evidence suggests a possible link between a reduced risk of dementia and treatment with GLP-1 agonists. In subjects who received GLP-1 agonist treatment, the rate of developing dementia was significantly lower.

The present disclosure shows the polypeptide conjugate provided herein and the pharmaceutical composition provided herein are effective in treating Alzheimer's disease. In certain embodiments, the polypeptide conjugate provided herein and the pharmaceutical composition provided herein are effective in improving cognitive function such as memory and attention, and/or delay the onset or progression of Alzheimer's disease. In certain embodiments, the polypeptide conjugate provided herein are more effective than semaglutide in treating Alzheimer's disease.

In another aspect, the present disclosure provides a method of managing body weight in a subject in need thereof, comprising administering a therapeutically effective amount of the polypeptide conjugate provided herein.

In another aspect, the present disclosure provides a method of reducing food intake in a subject in need thereof, comprising administering a therapeutically effective amount of the polypeptide conjugate provided herein.

In certain embodiments, the condition or metabolic disorder that can be treated or ameliorated using the polypeptide conjugate 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.

5 In certain embodiments, the condition or metabolic disorder that can be treated or ameliorated using the polypeptide conjugate provided herein, includes a condition where a human subject has a body mass index (BMI) of at least or higher than 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 10 38, 39 or 40. In certain embodiments, the human subject has an BMI ranging from 25 to 30, 26 to 30, 27 to 30, 28 to 30, 25 to 29, or 25 to 28.

15 The therapeutically effective amount of the polypeptide conjugate 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 fusion 20 polypeptide, the polypeptide complex and the conjugate 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 25 includes alleviation or amelioration of the symptoms of the 30 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, 35 or insulin sensitivity.

In certain embodiments, the polypeptide conjugate provided herein may be administered at a therapeutically effective dosage of about 0.01 mg/kg to about 100 mg/kg (e.g., 40 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 conjugate 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, 45 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 50 administration dosage may vary over the course of treatment depending on the reaction of the subject.

In certain embodiments, the polypeptide conjugate provided herein may be administered to the subject (e.g. human) at a dosing regimen that is no more frequently than 60 once daily, once every 3 days, once weekly, once every two weeks, once every three weeks, or once monthly. In certain embodiments, the polypeptide conjugate provided herein may be administered to the subject (e.g. human) at a dosing interval of twice weekly, once weekly, once every two weeks, once every three weeks, once monthly, or once every 65 two months. Therapeutic efficacy with low dosing frequency have the potential to improve a patient's adherence and

61

long-term treatment success. The currently available treatment semaglutide is dosed once weekly.

Without wishing to be bound by any theory, it is believed that certain polypeptide conjugate provided herein have significantly extended half life and are suitable for less frequent dosing than semaglutide for treating a metabolic condition, for example, less frequent than once weekly dosing (e.g. once every 8 days, once every 9 days, once every 10 days, once every 11 days, once every 12 days, once every 13 days, once every 14 days, once every 18 days, once every three weeks, once every 24 days, once every 4 weeks, or once every month). In certain embodiments, the dosing regimen is a continuous dosing regimen selected from twice-weekly dosing, once-weekly dosing, once bi-weekly dosing, once every three weeks dosing, once monthly dosing, or once every two months dosing. In certain embodiments, the dosing regimen has a dosing interval ranging from about once every 3 days to about once per month, or from about once weekly to about once per month.

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.

The pharmaceutical composition provided herein may be administered orally.

The polypeptide conjugate may be administered alone or in combination with one or more additional therapeutic means or agents.

In certain embodiments, when used for treating a metabolic disease, the polypeptide conjugate 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. A non-limiting list of examples of compounds that can be administered in combination with a pharmaceutical composition comprising the polypeptide conjugates provided herein, include rosiglitazone, pioglitazone, repaglinide, nateglinide, metformin, exenatide, sitagliptin, pramlintide, glipizide, glimepiride-acarbose, and miglitol.

Kit

Also provided are kits for practicing the disclosed methods. Such kits can comprise a pharmaceutical composition such as those described herein, including nucleic acids encoding the polypeptide conjugate provided herein, vectors and cells comprising such nucleic acids, and pharmaceutical compositions comprising such nucleic acid-containing compounds, 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.

62

In one aspect, a kit comprises (a) a pharmaceutical composition comprising a therapeutically effective amount of the polypeptide conjugate provided herein or a mutant form thereof; 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 dyslipidemia, arteriosclerosis, alcoholic steatohepatitis (ASH), diabetic nephropathy, metabolic syndrome such as metabolic syndrome X, nonalcoholic fatty liver disease (NAFLD), end-stage liver disease, hepatic steatosis (fatty liver), liver cirrhosis, or primary biliary cirrhosis (PBC), or Alzheimer's disease.

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.

Example 1: Recombinant Expression and Purification of GLP-1 Proteins

The GLP-1 fusion polypeptide listed in Table 1 (i.e., SEQ ID NOS: 51-80, 85-88) were produced from bacterial *E. coli* expression system, using BL21(DE3) derivative strain. The DNA coding for the tagged GLP-1 fusion polypeptides 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 the tagged GLP-1 fusion polypeptides 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. Cells were harvested and lysed in 20 mM Tris pH 8.0, 0.15M NaCl buffer by cell disruptor (900 bar, for twice). The soluble fractions, containing the tagged GLP-1 fusion polypeptides, were collected and by centrifugation (8,000×g, for 30 min). After removing tag by protease, the GLP-1 fusion polypeptides were purified by reverse phase chromatography. The samples in each step were characterized by LC/MS to confirm the correct molecular weight.

Example 2: Incorporation of Non-Protogenic Amino Acid in Recombinant Protein

The N-terminal His-Aib-Glu-Gly tetrapeptide or His-Aib dipeptide was dissolved in organic solvent and was added

63

into a solution of GLP-1 fusion polypeptides in organic solvent. The reaction was stirred at room temperature for 3 h. Then the piperidine was added into the reaction solution to remove Fmoc-protecting group.

Example 3: Preparation of GLP-1 Polypeptide Conjugates

To a solution of the GLP-1 fusion polypeptides obtain in Example 2 in NaOH was added with CRM reagent (e.g. HOOC—(CH₂)₁₆-CO-gGlu-2XADO, HOOC—(CH₂)₁₆-CO-gGlu-2XADO-EDA-CO—CH₂, etc.) in organic solvent dropwise. The reaction was stirred at room temperature for 1 h. Then the product was applied to reverse phase chromatography. This provided the GLP-1 polypeptide conjugates as listed in Table 1 as shown above.

The GLP-1 polypeptide conjugates 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.

Example 4a: In Vitro Activities

Method: The in vitro GLP-1 activities 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 GLP-1 polypeptide conjugates were measured at 1 nM or 100 nM as top concentration when in the absence or 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).

The activity of each GLP-1 polypeptide conjugate was represented by EC₅₀, derived from non-linear regression analysis.

Conclusion: Almost all molecules (i.e. GLP-1 polypeptide conjugates) show comparable or even better potency than semaglutide in the assay without HSA supplement. However, different molecules exhibited different degrees of reduction in GLP-1 activity (i.e. increased EC₅₀) in the presence of 1% HSA. Data in Table 2a and 2b suggested that linker length, fatty acid position and number of conjugated fatty acid moieties could be relevant to GLP-1 activities. As shown in Table 2, the molecules (whether mono-acylation or double-acylation) having a short linker length (i.e. 12 amino acid residues) had significantly reduced GLP1-activity, as compared with those molecules with longer linker lengths. As the distance increased between the C-terminal residue of GLP-1 and the CRM residue (i.e. lysine in the linker), the GLP-1 activity of the molecules in the presence of HSA seemed to steadily improve.

As shown in Table 2a, mono-acylated Molecules 002, 010 and 011 showed much higher activity than semaglutide in the presence of 1% HSA. For double-acylated Molecules 012, 016, 004, 006, 007, 005, 001 and 061, all showed relatively lower GLP-1 activity in the presence of 1% HSA than semaglutide but were still in an acceptable range.

64

TABLE 2a

In vitro activities of GLP-1 polypeptide conjugates. To measure in vitro GLP-1 activity, a BHK cell line overexpressing human GLP-1 receptor was used.						
	Molecule Code.	Linker length (aa)	Fatty acid number	Fatty acid position	GLP-1 activity (EC ₅₀ , pM)	GLP-1 activity (EC ₅₀ , pM, 1% HSA)
10	Semaglutide	0	Mono	GLP-1	4.2	243.8
	001	40	Double	GLP-1 and linker	2.8	1036
	002	40	Mono	Linker	1.1	8.6
	005	48	Double	GLP-1 and linker	5.0	1537
	006	32	Double	GLP-1 and linker	4.0	1387
	007	24	Double	GLP-1 and linker	4.7	3238
	008	12	Double	GLP-1 and linker	4.6	8732
	004	60	Double	GLP-1 and linker	3.0	984.4
	012	40	Double	GLP-1 and linker	2.1	1051
	016	32	Double	GLP-1 and linker	2.1	1459
	010	24	Mono	Linker	1.7	22.9
	011	12	Mono	Linker	3.7	43.9
20	061	40	Double	GLP-1 and linker	4.4	2213

Example 4b: In Vitro Activities

Method: The in vitro GLP-1 activities were measured using a CHO cell line overexpressing human GLP-1 receptor and CRE luciferase reporter with or without 1% human serum albumin (HSA). Tested GLP-1 polypeptide conjugates were measured at 30 nM or 600 nM as top concentration when in the absence or presence of 1% HSA with 3-fold serial dilutions. After cells were treated with molecules for 5 hours, luciferase activities were measured by Bright-Glo Luciferase Assay System (Promega, E2620). The activity of each GLP-1 polypeptide conjugate was represented by EC₅₀, derived from non-linear regression analysis.

Conclusion: As shown in Table 2b, mono-acylated Molecules 002 (1006) and 019 (1023) showed much higher activity than semaglutide in the presence of 1% HSA.

TABLE 2b

In vitro activities of GLP-1 polypeptide conjugates. To measure in vitro GLP-1 activity, a CHO cell line overexpressing human GLP-1 receptor was used.						
	Molecule Code.	Linker length (aa)	Fatty acid number	Fatty acid position	GLP-1 activity (EC ₅₀ , nM, 1% HSA)	
50	Semaglutide	0	Mono	GLP-1	0.33	23.8
	002	40	Mono	Linker	0.12	3.47
	019	40	Mono	Linker	0.07	2.16

Example 5: In Vivo Activities in C57 Lean Mice

Method: 10 week old male C57BL/6 mice were injected subcutaneously with protein once on Day 1. Body weight was measured daily and five animals were used for each treatment group. Body weight was monitored for each individual animal. % BW loss=100*(BW on Day n-BW on Day 1)/(BW on Day 1). Data are indicated as mean values and standard error (SEM). The area under the curve for body weight loss (%) from 0 until 8 days (AUC ΔBW 0-8d) was calculated.

Conclusion: Molecules 001 and 012 showed dose dependent efficacy on body weight control (FIGS. 1A and Table 3).

As shown in FIG. 1D and Table 3, Molecules 007 and 008 have less efficacy than Molecules 012 and 016 which is consistent with the in vitro activity data. Molecules 001, 016 and 012 with double fatty acids exhibited more sustainable efficacy than semaglutide, suggesting Molecules 001, 016 and 012 may have longer half-life (FIGS. 1A and 1B).

TABLE 3

The area under the curve for body weight loss (%) from 0 until 8 days (AUC ΔBW 0-8 d) in C57 lean mice studies.		
Studies	Group	AUC ΔBW 0-8 d (day*%)
Study 1 (FIG. 1A)	Semaglutide 30 nmol/kg	-9.3
	Molecule 012 30 nmol/kg	-13.6
	Molecule 012 100 nmol/kg	-28.3
	Molecule 001 100 nmol/kg	-25.9
Study 2 (FIG. 1B)	Semaglutide 30 nmol/kg	-13.2
	Molecule 008 30 nmol/kg	-1.85
	Molecule 007 30 nmol/kg	-2.63
	Molecule 012 30 nmol/kg	-21.6
	Molecule 016 30 nmol/kg	-20.1

Example 6a: In Vivo Activities in db/db Mice

Method: 10 week old male db/db mice were injected subcutaneously with GLP-1 polypeptide conjugate once. Fasting glucose was measured on different times and three animals were used for each group. Delta blood glucose is glucose subtracted by baseline level. Data are indicated as mean values and standard error (SEM).

Conclusion: Both Molecules 001 and 002 showed dose dependent efficacy on glucose control (FIGS. 2A and 2B). Molecule 001 with double fatty acids exhibited more sustainable efficacy than semaglutide, suggesting Molecule 001 may have longer half-life.

Method: 10 weeks old male db/db mice were injected subcutaneously with GLP-1 polypeptide conjugate once. Non-Fasting glucose was measured on different times and five animals were used for each group. Delta blood glucose is glucose subtracted by baseline level. Data are indicated as mean values and standard error (SEM). The area under the curve for delta blood glucose from 0 until 192 hours (AUCΔBG 0-192 h) was calculated.

Conclusion: Molecule 012 showed dose dependent efficacy on glucose control (FIGS. 2C, Table 4). Molecule 012 with double fatty acids exhibited more sustainable efficacy than semaglutide.

TABLE 4

The area under the curve for non-fasting delta blood glucose (%) from 0 until 192 hours (AUC ABG 0-192 h) in db/db mice for FIG. 2C.		
Group	AUCABG 0-192 h (hr*mmol/L)	
Semaglutide 3 nmol/kg	-317.0	
Semaglutide 10 nmol/kg	-583.4	
Semaglutide 30 nmol/kg	-527.1	
Molecule 012 3 nmol/kg	-424.5	
Molecule 012 10 nmol/kg	-1004.5	
Molecule 012 30 nmol/kg	-1447.0	

Example 6b: In Vivo Activities in db/db Mice

Method: 10 week old male db/db mice were injected subcutaneously with Molecule 019 once. Random blood glucose was measured on different times and five animals were used for each group. Delta blood glucose is glucose subtracted by baseline level. Data are indicated as mean values and standard error (SEM).

Conclusion: Molecule 019 showed dose dependent efficacy on glucose control (FIGS. 2D and 2E) and significantly higher potency than semaglutide.

Example 7: Pharmacokinetic Measurement

Method: 6-8 week old male C57BL/6 mice were administrated in a single subcutaneous dose of 30 nmol/kg protein (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, 48 hr, 72 hr, 96 hr, 120 hr, 144 hr, 168 hr and 192 hr after the injection. The concentrations of GLP-1 polypeptide conjugates in the plasma were measured by ELISA assay. Based on the graph showing plasma concentration of each GLP-1 polypeptide conjugate versus time after subcutaneous injection, the pharmacokinetic parameters were calculated by WinNonlin.

Conclusion: Molecule 001 and 012 showed longer half-life than semaglutide and Molecule 002 in mouse (Table 5), consistent with in vivo efficacies.

TABLE 5

Pharmacokinetic parameters of GLP-1 polypeptide conjugates in mice.					
PK parameters	Unit	Semaglutide	Molecule 001	Molecule 002	Molecule 012
T _{max}	hr	6.7	32	4	24
C _{max}	nmol/L	43.1	172.0	86.5	134
Terminal t _{1/2}	hr	7.4	30.1	7.2	28.2
AUC _{tau}	hr*nmol/L	614.7	12256	1507	7752
		(Tau = 24 h)	(Tau = 168 h)	(Tau = 48 h)	(Tau = 192 h)

US 12,383,623 B2

67

Method: 6-8 weeks male SD rats were administrated in a single subcutaneous dose of 15 nmol/kg GLP-1 polypeptide conjugate (n=3/group) and a single intravenous dose of 15 nmol/kg GLP-1 polypeptide conjugate (n=3/group), respectively. Plasma samples were collected pre-dose (-5 min), 0.5 hr, 1 hr, 2 hr, 4 hr, 6 hr, 8 hr, 24 hr, 32 hr, 48 hr, 72 hr, 96 hr, 120 hr, 144 hr, 168 hr, 192 hr, 216 hr and 240 hr after subcutaneous administration. Plasma samples were collected pre-dose (-5 min), 0.083 hr, 0.25 hr, 0.5 hr, 1 hr, 2 hr, 4 hr, 6 hr, 8 hr, 24 hr, 32 hr, 48 hr, 72 hr, 96 hr and 120 hr

68

after intravenous administration. The concentrations of the polypeptide conjugates in the plasma were measured by LC-MS/MS method. Based on the graph showing plasma concentration of each polypeptide conjugate versus time after subcutaneous injection, the pharmacokinetic parameters were calculated by WinNonlin.

Conclusion: As shown in Table 6, Molecules 001 and 012 showed longer half-life than semaglutide, which was reported to have a $T_{1/2}$ of 12 hrs (s.c. administration, as reported in non-clinical reviews from FDA) in rat.

TABLE 6

Pharmacokinetic parameters of GLP-1 polypeptide conjugates in rats. Pharmacokinetic data were analyzed by WinNonlin software. Tmax, Cmax, $T_{1/2}$, AUC were calculated for each molecule.					
PK parameters	Unit	Molecule 001		Molecule 012	
		i.v.	s.c.	i.v.	s.c.
T_{max}	hr	NA	24	NA	26.7
C_{max} (C0 for iv)	nmol/L	317	56.2	231	52.0
Terminal $t_{1/2}$	hr	26.9	25.8	26.9	26.9
AUC_{tau}	hr*nmol/L	4485 (Tau = 120 h)	2887 (Tau = 192 h)	4140 (Tau = 120 h)	2954 (Tau = 192 h)
F	%	NA	64	NA	71

NA: not available.

Method: 10-month old male Bama minipigs were administered with a single subcutaneous dose of 5 nmol/kg GLP-1 polypeptide conjugate (n=2/group) and a single intravenous dose of 5 nmol/kg protein (n=2/group), respectively. Plasma samples for Molecule 012 and semaglutide group were collected pre-dose (-5 min), 0.5 hr, 1 hr, 3 hr, 8 hr, 24 hr, 48 hr, 72 hr, 96 hr, 168 hr, 336 hr, 504 hr and 672 hr after subcutaneous administration. Plasma samples for Molecule 30 012 and semaglutide group were collected pre-dose (-5 min), 0.083 hr, 0.5 hr, 1 hr, 3 hr, 8 hr, 24 hr, 48 hr, 72 hr, 96 hr, 168 hr, 336 hr and 504 hr after intravenous administration. For semaglutide group, the concentrations of GLP-1 polypeptide conjugates in the plasma were measured by 35 ELISA assay. For Molecule 012 group, the concentrations of proteins in the plasma were measured by LC-MS/MS method. Based on the graph showing plasma concentration of each GLP-1 polypeptide conjugate versus time after administration, the pharmacokinetic parameters were calculated by WinNonlin.

Conclusion: Molecule 012 showed longer half-life than semaglutide in minipigs (Table 7)

TABLE 7

Pharmacokinetic parameters of GLP-1 polypeptide conjugates in minipigs. Pharmacokinetic data were analyzed by WinNonlin software. Tmax, Cmax, $T_{1/2}$, AUC were calculated for each molecule.					
PK parameters	Unit	Semaglutide		Molecule 012	
		i.v.	s.c.	i.v.	s.c.
T_{max}	hr	NA	16	NA	60
C_{max} (C0 for iv)	nmol/L	188	54.9	236	71.4
Terminal $t_{1/2}$	hr	53.0	62.8	136	147
AUC_{tau}	hr*nmol/L	6643 (Tau = 336 h)	5658 (Tau = 336 h)	19617 (Tau = 504 h)	16830 (Tau = 672 h)
F	%	NA	85	NA	86

Example 8a: Efficacy Study in Disease Models

Selected molecules are assessed in disease animal models (such db/db mice) to determine body weight, food intake, glucose efficacy with dose responses in chronic studies. Some biomarkers are also measured, including fasting insulin, plasma triglyceride, cholesterol, liver triglyceride, and inflammatory biomarkers (ALT, AST and CRP).

Method: 22 week old DIO male C57BL/6 mice (~50 g) 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 $-1 * (\% \text{ BW loss} - \% \text{ BW loss of vehicle group})$; Cumulative food intake reduction is calculated by $-100 * (\text{cumulative food intake-cumulative food intake of vehicle}) / \text{cumulative food intake of vehicle}$.

Conclusion: in DIO study, as shown in FIG. 3A, 3B and Table 8, Molecule 012 showed dose dependent efficacy on body weight reduction, food intake reduction and blood glucose control.

TABLE 8

Food intake and body weight reduction in DIO mice at Day 25		
Group	Cumulative food intake reduction (% reduction compared to vehicle) (Day 25)	Body weight reduction (% reduction compared to vehicle) (Day 25)
Molecule 012 10 nmol/kg, Q2D	35	18.42
Molecule 012 30 nmol/kg, Q2D	37	23.0
Molecule 012 100 nmol/kg, Q2D	51	37.6

Example 8b: Efficacy Study in Disease Models

Selected molecules are assessed in Alzheimer' disease animal models (APP/PS1, APPswe/PSEN1dE9 mice) to determine cognitive function (Morris water maze and Y-maze), brain A β plaque deposition, and hippocampal pyramidal neuron number with dose responses in studies.

To assess the cognitive function of the animals, the Morris water maze and Y-maze tests were utilized. The effects of the molecules on beta-amyloid plaque deposition in the cortex and hippocampus, which are regions of the brain affected in Alzheimer's disease, were also measured. Furthermore, the number of pyramidal neurons in the hippocampus was assessed as a measure of the drug's potential neuroprotective effects. These methods allowed for a comprehensive evaluation of the efficacy of the molecules in mitigating the pathological changes associated with Alzheimer's disease in a preclinical setting.

Method: 6 months old APP/PS1 male mice were injected once day (QD) or once every other day (Q2D) subcutaneously with designated GLP-1 polypeptide conjugate (Molecule 012, formulated in 20 mg/mL Molecule 012, 2.87

mg/mL disodiumhydrogen phosphate dodecahydrate and 8.25 mg/ml sodium chloride pH 7.4) for 10 weeks. 15 animals were used for each treatment group. Body weight and food intake was monitored every week. Cognitive function behavior experiments were conducted after dosing including Morris water maze and Y-maze. Brain tissue was taken for immunohistochemical staning to evaluate the A β content in cortex and hippocampus. The number of hippocampal pyramidal neuron was measured by HE staining. Data are indicated as mean values and standard error (SEM). Statistical analysis was performed by One-way ANOVA. Semaglutide was used as a positive control.

Conclusion: in APP/PS1 study, as shown in FIGS. 4A, 4B, 4C, 4D and 4E, Molecule 012 showed dose dependent efficacy on cognitive function improvement, brain A β plaque deposition reduction, and hippocampal pyramidal neuron protection. The Molecule 012 was found to be effective in increasing quadrant occupancy, which refers to the percentage of time spent in the target quadrant during a navigation task, and increasing alternation percentage, which refers to the frequency of alternating between different locations during a maze task. Additionally, the Molecule 012 was observed to reduce A β plaque deposition in both the cortex and hippocampus regions of the brain, which are typically affected in Alzheimer's disease. The number of neurons in the hippocampus, a region critical for memory, was also improved by the Molecule 012. These findings suggest that the molecule 012 has a positive impact on cognitive function and reduce the pathological changes associated with Alzheimer's disease. Molecule 012 showed significantly better efficacy than semaglutide in treating Alzheimer's disease in the animal studies.

Example 8c: Efficacy Study in Disease Models

Selected molecules are assessed in a Alzheimer' disease animal model for sporadic Alzheimer's dementia, which is the senescence-accelerated mouse prone 8 (SAMP8) model. SAMP8 mice exhibit age-related cognitive decline and neuropathological changes similar to those observed in human Alzheimer's disease, including beta-amyloid plaque deposition and neuroinflammation. They also display behavioral abnormalities and reduced synaptic plasticity.

7 months old SAMP8 mice were injected once every other day (Q2D) subcutaneously with designated GLP-1 polypeptide conjugate (Molecule 012) for 16 weeks, at 30 nmol/kg, 100 nmol/kg and 300 nmol/kg, respectively. 15 animals were used for each treatment group. Body weight and food intake was monitored every week. Cognitive function behavior experiments were conducted after dosing including Morris water maze and Y-maze. Brain tissue was taken for immunohistochemical staning to evaluate the A β content in cortex and hippocampus. The number of hippocampal pyramidal neuron was measured by HE staining. Data are indicated as mean values and standard error (SEM). Statistical analysis was performed by One-way ANOVA. Semaglutide and Donepezil were used as positive controls.

Molecule 012 shows dose dependent efficacy on cognitive function improvement, brain A β plaque deposition reduction, and hippocampal pyramidal neuron protection in SAMP8 mice. Molecule 012 also shows significantly better efficacy than the positive controls in treating Alzheimer's disease in SAMP8 mice.

Example 9a: PK Study in Non-Human Primates

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

Method: 3-5 years old male cynomolgus monkeys were administrated in a single subcutaneous dose of 5 nmol/kg GLP-1 polypeptide conjugate (n=2/group). Plasma samples for semaglutide group were collected pre-dose (-5 min), 2 hr, 4 hr, 6 hr, 8 hr, 24 hr, 48 hr, 72 hr, 96 hr, 120 hr, 144 hr, 216 hr, 288 hr, 360 hr, 432 hr and 504 hr after the injection. Plasma samples for Molecule 012 group were collected pre-dose (-5 min), 0.5 hr, 1 hr, 2 hr, 4 hr, 6 hr, 8 hr, 24 hr, 48 hr, 96 hr, 120 hr, 144 hr, 168 hr, 192 hr, 216 hr, 288 hr, 432 hr and 504 hr after the injection. For semaglutide group, the concentrations of proteins in the plasma were measured by ELISA assay. For Molecule 012 group, the concentrations of GLP-1 polypeptide conjugates in the plasma were measured by LC-MS/MS method. Based on the graph showing plasma concentration of each protein versus time after subcutaneous injection, the pharmacokinetic parameters were calculated by WinNonlin.

Conclusion: Molecule 012 showed longer half-life than semaglutide and Molecule 002 in monkey (Table 9a)

TABLE 9a

Pharmacokinetic parameters of GLP-1 polypeptide conjugates in cynomolgus monkeys. Pharmacokinetic data were analyzed by WinNonlin software. Tmax, Cmax, T_{1/2}, AUC were calculated for each molecule.

PK parameters	Unit	Semaglutide	Molecule 012
T _{max}	hr	7	24
C _{max}	nmol/L	48.1	58.0
Terminal t _{1/2}	hr	65	122
AUC _{tau}	hr*nmol/L	2876	10989
	(Tau = 96 h)		(Tau = 504 h)

Example 9b: PK Study in Non-Human Primates

Method: 3-5 years old male cynomolgus monkeys were administrated in a single intravenous dose of 4 nmol/kg and 5 nmol/kg for semaglutide and Molecule 019 (n=2/group), respectively. Plasma samples for semaglutide group were collected pre-dose (-5 min), 0.083 hr, 1 hr, 2 hr, 4 hr, 8 hr, 24 hr, 36 hr, 48 hr, 72 hr, 96 hr, 120 hr, 144 hr, and 168 hr after the injection. Plasma samples for Molecule 012 group were collected pre-dose (-5 min), 0.083 hr, 0.25 hr, 0.5 hr, 1 hr, 2 hr, 4 hr, 8 hr, 12 hr, 24 hr, 36 hr, 48 hr, 72 hr, 96 hr, 120 hr, 144 hr, and 168 hr after the injection. The concentrations of proteins in the plasma were measured by LC-MS/MS method. Based on the graph showing plasma concentration of each protein versus time after intravenous administration, the pharmacokinetic parameters were calculated by WinNonlin.

Conclusion: Molecule 019 showed similar half-life with semaglutide in monkeys (Table 9b), which potentially supports once weekly dosing in human. Further in view of the potent GLP-1 activity (10-fold higher than semaglutide) of Molecule 019, it is expected that Molecule 019 is a promising candidate that not only provides better efficacy than semaglutide but also supports a desirable once-weekly dosing frequency (or potentially even less frequent).

TABLE 9b

Pharmacokinetic parameters of GLP-1 polypeptide conjugates in cynomolgus monkeys. Pharmacokinetic data were analyzed by WinNonlin software. C₀, T_{1/2} and AUC were calculated for each molecule.

PK parameters	Unit	Semaglutide	Molecule 019
C ₀	nmol/L	85.1	176.5
Vz	L/kg	0.09	0.06
Cl	mL/min/kg	0.026	0.015
Terminal t _{1/2}	hr	42.8	49.0
AUC _{tau}	hr*nmol/L	2402	5302
	(t = 168 hr)		(t = 168 hr)

Example 10: Immunogenicity Assessment

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

Example 11: Human Serum Albumin Binding

Method: Binding of molecules to serum albumin was characterized by surface plasmon resonance in a Biacore 8K instrument. Serum albumin from different species was covalently bound to CM5 sensor chips surface until 4000 RU was reached. The chip was blocked by 1M ethanalamine 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.

Conclusion: Molecules 004, 001, 006 and 012 showed higher binding affinity of human serum albumin (see Table 10) than semaglutide, Molecule 019 and Molecule 002 which is consistent with the PK data.

TABLE 10

Human serum albumin binding affinity

	KD (µM)	Rmax (RU)
Semaglutide	2.82	55.2
Molecule 004	0.716	546.0
Molecule 001	0.697	498.3
Molecule 006	0.746	398.6
Molecule 012	0.376	207.3
Molecule 002	9.11	307.8
Molecule 019	10.1	146.3

Example 12: Preparation of Liquid Composition

Unless otherwise specified, Molecule 012 was prepared by diluting Molecule 012 into the formulation buffer, which was composed of buffer agent (e.g. disodiumhydrogen phosphate dodecahydrate or disodium phosphate dihydrate was used for pH 7.4-8.3 and histidine or citric acid anhydrous/trisodium citrate dihydrate was used for pH 6.5) and isotonic agent (sodium chloride, propylene glycol or mannitol). pH was adjusted to the relevant value by sodium hydroxide and/or hydrochloric acid. The composition was filtrated by a 0.22 µm sterile filter.

Compositions comprising Molecule 012 were tested in this experiment. The tested compositions were with propyl-

73

ene glycol (14 mg/ml), sodium chloride (8.25 mg/ml) or mannitol (46.4 mg/mL) as isotonic agent, combined with varieties of pH and protein concentrations as shown in TABLE 11. For compositions number 1-6, 8-11, 13-16 and 18-23, 2.87 mg/mL disodiumhydrogen phosphate dodecahydrate was used as pH buffer. For compositions number 7, 12 and 17, 0.14 mg/mL citric acid anhydrous and 2.74 mg/mL trisodium citrate dihydrate was used as pH buffer. For compositions number 24, 1.24 mg/mL histidine was used as pH buffer.

TABLE 11

Compositions tested in Example 12				
Composition no.	Molecule 012 (mg/mL)	Buffer	Isotonic agent	pH
1	1	phosphate	sodium chloride	7.4
2	5	phosphate	sodium chloride	7.4
3	10	phosphate	sodium chloride	7.4
4	20	phosphate	sodium chloride	7.4
5	40	phosphate	sodium chloride	7.4
6	80	phosphate	sodium chloride	7.4
7	20	citrate	sodium chloride	6.5
8	20	phosphate	sodium chloride	8.3
9	20	phosphate	sodium chloride	7.0
10	20	phosphate	sodium chloride	7.8
11	20	phosphate	propylene glycol	7.4
12	20	citrate	propylene glycol	6.5
13	20	phosphate	propylene glycol	8.3
14	20	phosphate	propylene glycol	7.0
15	20	phosphate	propylene glycol	7.8
16	20	phosphate	mannitol	7.4
17	20	citrate	mannitol	6.5
18	20	phosphate	mannitol	8.3
19	20	phosphate	mannitol	7.0
20	20	phosphate	mannitol	7.8
21	60	phosphate	sodium chloride	7.4
22	100	phosphate	sodium chloride	7.4
23	120	phosphate	sodium chloride	7.4
24	20	histidine	sodium chloride	6.5

Example 13: Stability Test of Molecule 012 Compositions

The 24 compositions listed in Example 12 were prepared and stored at 25° C. for six months. Impurities, high molecular weight products (HMWP) and relative biological activity were investigated with RP-UPLC, SEC-UPLC and hGLP1-R expressed cell assay at 0, 2 and 6 months. The RP-UPLC was performed with Kinetex C18 column (Phenomenex) and UPLC H-class(Waters). SEC-UPLC was performed with TSKgel UP-SW2000 (Tosoh) and UPLC H-class(Waters). relative biological activity were determined using human GLP-1 receptor expressed CHO cell by detecting cAMP activity and calculated using formula (100*

74

(EC50_{standard}/EC50_{sample}))%. The results for compositions number 1-8, 11 and 21-23 were shown in TABLE 12 and 13, and results for compositions number 9,10, 12-20, 24 are not shown. All compositions are tested and results show that they are stable over the tested period. As shown in Table 12 and 13, all Compositions were stable over at least 2 or 6 months at 25° C.

TABLE 12

Composition no.*	Total impurities			HMWP			Relative biological activity(%)		
	0 M	2 M	6 M	0 M	2 M	6 M	0 M	2 M	6 M
1	2.30	4.39	6.89	0.18	0.33	0.56	100	NA	74
2	2.42	3.78	6.48	0.13	0.33	0.56	100	NA	76
3	1.86	3.39	5.70	0.07	0.32	0.51	89	NA	103
4	1.97	3.26	5.82	0.07	0.28	0.65	89	NA	85
5	2.45	4.40	8.15	0.14	0.42	0.68	100	128	78
6	2.46	4.48	7.84	0.14	0.36	1.19	87	129	94
7	1.91	3.39	5.91	0.08	0.29	0.64	89	NA	97
8	2.46	4.44	8.36	0.15	0.42	0.69	100	NA	100
11	2.35	4.39	6.98	0.16	0.35	0.67	100	NA	129
21	2.43	4.33	NA	0.14	0.39	NA	100	129	NA
22	2.00	3.95	NA	0.06	0.48	NA	100	82	NA
23	1.99	4.13	NA	0.04	0.50	NA	100	101	NA

*Compositions 1, 2, 5, 8, 11, 21, 22 and 23 were concentrated/diluted/dialyzed by using Molecule 012 production batch lot 202112001. Compositions 3, 4, 7 were diluted/dialyzed by using Molecule 012 production batch lot 202110001. Composition 6 was concentrated by using Molecule 012 production batch lot 202112002. pH was adjusted with hydrochloric acid or sodium hydroxide. The compositions were filtrated by a 0.22 μm sterile filter. NA = not available.

TABLE 13

Composition no.*	Total impurities Increase (%)		HMWP Increase (%)	
	2 M	6 M	2 M	6 M
1	2.09	4.59	0.15	0.38
2	1.36	4.06	0.20	0.43
3	1.53	3.84	0.25	0.44
4	1.29	3.85	0.21	0.58
5	1.95	5.61	0.28	0.42
6	2.02	5.86	0.21	0.97
7	1.48	4.00	0.21	0.56
8	1.98	5.90	0.27	0.54
11	2.04	4.63	0.19	0.51
21	1.90	NA	0.25	NA
22	1.95	NA	0.42	NA
23	2.14	NA	0.46	NA

*Compositions 1, 2, 5, 8, 11, 21, 22 and 23 were concentrated/diluted/dialyzed by using Molecule 012 production batch lot 202112001. Compositions 3, 4, 7 were diluted/dialyzed by using Molecule 012 production batch lot 202110001. Composition 6 was concentrated by using Molecule 012 production batch lot 202112002. pH was adjusted with hydrochloric acid or sodium hydroxide. The compositions were filtrated by a 0.22 μm sterile filter.

NA = not available.

SEQUENCE LISTING

```

Sequence total quantity: 88
SEQ ID NO: 1      moltype = AA length = 31
FEATURE          Location/Qualifiers
REGION           1..31
source            note = Synthetic
                  1..31
                  mol_type = protein
                  organism = synthetic construct
SEQUENCE: 1
HAEGTFTSDV SSYLEGQAAK EFIGAWLVKGR G

```

-continued

```

SEQ ID NO: 2      moltype = AA  length = 31
FEATURE          Location/Qualifiers
REGION           1..31
SITE              note = Synthetic
1
note = MISC_FEATURE - X is H, imidazole-4-acetate (IA), or
                  imidazolepropionic acid (IPA)
SITE              2
note = MISC_FEATURE - X is A, G, S, V, Aib, T, I, or L
SITE              16
note = MISC_FEATURE - X is G, or E
SITE              17
note = MISC_FEATURE - X is Q, C or K
SITE              20
note = MISC_FEATURE - X is K, R, or C
SITE              21
note = MISC_FEATURE - X is E, K, or C
SITE              24
note = MISC_FEATURE - X is A, C or K
SITE              28
note = MISC_FEATURE - X is R, K, or C
SITE              30
note = MISC_FEATURE - X is R or G
source            1..31
mol_type = protein
organism = synthetic construct

SEQUENCE: 2
XXEGTFTSDV SSYLEXXAAX XFIXWLWXGX G                               31

SEQ ID NO: 3      moltype = AA  length = 31
FEATURE          Location/Qualifiers
REGION           1..31
SITE              note = Synthetic
2
note = MISC_FEATURE - X is Aib
source            1..31
mol_type = protein
organism = synthetic construct

SEQUENCE: 3
HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG G                               31

SEQ ID NO: 4      moltype = AA  length = 31
FEATURE          Location/Qualifiers
REGION           1..31
SITE              note = Synthetic
2
note = MISC_FEATURE - X is Aib
source            1..31
mol_type = protein
organism = synthetic construct

SEQUENCE: 4
HXEGTFTSDV SSYLEGQAAC EFIAWLVRGG G                               31

SEQ ID NO: 5      moltype = AA  length = 31
FEATURE          Location/Qualifiers
REGION           1..31
SITE              note = Synthetic
2
note = MISC_FEATURE - X is Aib
source            1..31
mol_type = protein
organism = synthetic construct

SEQUENCE: 5
HXEGTFTSDV SSYLEGQAAR EFIAWLVRGG G                               31

SEQ ID NO: 6      moltype = AA  length = 31
FEATURE          Location/Qualifiers
REGION           1..31
SITE              note = Synthetic
2
note = MISC_FEATURE - X is Aib
source            1..31
mol_type = protein
organism = synthetic construct

SEQUENCE: 6
HXEGTFTSDV SSYLEEQAAK EFIAWLVRGG G                               31

SEQ ID NO: 7      moltype = AA  length = 31
FEATURE          Location/Qualifiers

```

-continued

REGION	1..31	
SITE	note = Synthetic	
source	2	
	note = MISC_FEATURE - X is Aib	
	1..31	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 7	HXEGTFTSDV SSYLEEQAAC EFIAWLVRGG G	31
SEQ ID NO: 8	moltype = AA length = 31	
FEATURE	Location/Qualifiers	
REGION	1..31	
SITE	note = Synthetic	
source	2	
	note = MISC_FEATURE - X is Aib	
	1..31	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 8	HXEGTFTSDV SSYLEEQAAAR EFIAWLVRGG G	31
SEQ ID NO: 9	moltype = AA length = 31	
FEATURE	Location/Qualifiers	
REGION	1..31	
source	note = Synthetic	
	1..31	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 9	HGEGTFTSDV SSYLEEQAAAR EFIAWLVRGG G	31
SEQ ID NO: 10	moltype = AA length = 8	
FEATURE	Location/Qualifiers	
REGION	1..8	
source	note = Synthetic	
	1..8	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 10	GQEPGAQP	8
SEQ ID NO: 11	moltype = AA length = 8	
FEATURE	Location/Qualifiers	
REGION	1..8	
source	note = Synthetic	
	1..8	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 11	GAQPGAQP	8
SEQ ID NO: 12	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 12	GQEP	4
SEQ ID NO: 13	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 13	GAQP	4
SEQ ID NO: 14	moltype = AA length = 12	
FEATURE	Location/Qualifiers	
REGION	1..12	
source	note = Synthetic	
	1..12	
	mol_type = protein	
	organism = synthetic construct	

-continued

SEQUENCE: 14	
GAQPGQEPGA QP	12
SEQ ID NO: 15	moltype = AA length = 8
FEATURE	Location/Qualifiers
REGION	1..8
source	note = Synthetic
	1..8
	mol_type = protein
	organism = synthetic construct
SEQUENCE: 15	
GAQPGQEP	8
SEQ ID NO: 16	moltype = AA length = 4
FEATURE	Location/Qualifiers
REGION	1..4
source	note = Synthetic
	1..4
	mol_type = protein
	organism = synthetic construct
SEQUENCE: 16	
GEQP	4
SEQ ID NO: 17	moltype = AA length = 4
FEATURE	Location/Qualifiers
REGION	1..4
source	note = Synthetic
	1..4
	mol_type = protein
	organism = synthetic construct
SEQUENCE: 17	
GPQE	4
SEQ ID NO: 18	moltype = AA length = 4
FEATURE	Location/Qualifiers
REGION	1..4
source	note = Synthetic
	1..4
	mol_type = protein
	organism = synthetic construct
SEQUENCE: 18	
GPEQ	4
SEQ ID NO: 19	moltype = AA length = 4
FEATURE	Location/Qualifiers
REGION	1..4
source	note = Synthetic
	1..4
	mol_type = protein
	organism = synthetic construct
SEQUENCE: 19	
GSEP	4
SEQ ID NO: 20	moltype = AA length = 4
FEATURE	Location/Qualifiers
REGION	1..4
source	note = Synthetic
	1..4
	mol_type = protein
	organism = synthetic construct
SEQUENCE: 20	
GESP	4
SEQ ID NO: 21	moltype = AA length = 4
FEATURE	Location/Qualifiers
REGION	1..4
source	note = Synthetic
	1..4
	mol_type = protein
	organism = synthetic construct
SEQUENCE: 21	
GPSE	4
SEQ ID NO: 22	moltype = AA length = 4
FEATURE	Location/Qualifiers
REGION	1..4
source	note = Synthetic
	1..4
	mol_type = protein

-continued

SEQUENCE: 22	organism = synthetic construct	
GPES		4
SEQ ID NO: 23	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 23		
GQAP		4
SEQ ID NO: 24	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 24		
GPAQ		4
SEQ ID NO: 25	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 25		
GPQA		4
SEQ ID NO: 26	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 26		
GSQP		4
SEQ ID NO: 27	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 27		
GASP		4
SEQ ID NO: 28	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 28		
GPAS		4
SEQ ID NO: 29	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	
	mol_type = protein	
	organism = synthetic construct	
SEQUENCE: 29		
GPSA		4
SEQ ID NO: 30	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
	1..4	

-continued

SEQUENCE: 30	mol_type = protein organism = synthetic construct	
GGGS		4
SEQ ID NO: 31	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
REGION	1..4	
source	note = Synthetic	
SEQUENCE: 31	1..4	
GGGS	mol_type = protein organism = synthetic construct	
		4
SEQ ID NO: 32	moltype = AA length = 5	
FEATURE	Location/Qualifiers	
REGION	1..5	
source	note = Synthetic	
SEQUENCE: 32	1..5	
GGGGS	mol_type = protein organism = synthetic construct	
		5
SEQ ID NO: 33	moltype = AA length = 8	
FEATURE	Location/Qualifiers	
REGION	1..8	
source	note = Synthetic	
SEQUENCE: 33	1..8	
GQEPGQAP	mol_type = protein organism = synthetic construct	
		8
SEQ ID NO: 34	moltype = AA length = 8	
FEATURE	Location/Qualifiers	
REGION	1..8	
source	note = Synthetic	
SEQUENCE: 34	1..8	
GQAPGQEP	mol_type = protein organism = synthetic construct	
		8
SEQ ID NO: 35	moltype = AA length = 32	
FEATURE	Location/Qualifiers	
REGION	1..32	
source	note = Synthetic	
SEQUENCE: 35	1..32	
SEPATSGSET PGTSESATPE SGPGTSTEPS EG	mol_type = protein organism = synthetic construct	
		32
SEQ ID NO: 36	moltype = AA length = 6	
FEATURE	Location/Qualifiers	
REGION	1..6	
source	note = Synthetic	
SEQUENCE: 36	1..6	
SEPATS	mol_type = protein organism = synthetic construct	
		6
SEQ ID NO: 37	moltype = AA length = 6	
FEATURE	Location/Qualifiers	
REGION	1..6	
source	note = Synthetic	
SEQUENCE: 37	1..6	
GSETPG	mol_type = protein organism = synthetic construct	
		6
SEQ ID NO: 38	moltype = AA length = 6	
FEATURE	Location/Qualifiers	
REGION	1..6	
	note = Synthetic	

-continued

source	1..6 mol_type = protein organism = synthetic construct	
SEQUENCE: 38		
TSESAT		6
SEQ ID NO: 39	moltype = AA length = 6	
FEATURE	Location/Qualifiers	
REGION	1..6	
note = Synthetic		
source	1..6 mol_type = protein organism = synthetic construct	
SEQUENCE: 39		
PESGPG		6
SEQ ID NO: 40	moltype = AA length = 6	
FEATURE	Location/Qualifiers	
REGION	1..6	
note = Synthetic		
source	1..6 mol_type = protein organism = synthetic construct	
SEQUENCE: 40		
TSTEPS		6
SEQ ID NO: 41	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
source	1..4 mol_type = protein organism = synthetic construct	
SEQUENCE: 41		
GQKP		4
SEQ ID NO: 42	moltype = AA length = 4	
FEATURE	Location/Qualifiers	
source	1..4 mol_type = protein organism = synthetic construct	
SEQUENCE: 42		
GQCP		4
SEQ ID NO: 43	moltype = AA length = 40	
FEATURE	Location/Qualifiers	
REGION	1..40	
note = Synthetic		
source	1..40 mol_type = protein organism = synthetic construct	
SEQUENCE: 43		
GAQPGAQPGA QPQGAQPGAQP GAQPGAQPGA QPQGAQPGQKP		40
SEQ ID NO: 44	moltype = AA length = 48	
FEATURE	Location/Qualifiers	
REGION	1..48	
note = Synthetic		
source	1..48 mol_type = protein organism = synthetic construct	
SEQUENCE: 44		
GAQPGAQPGA QPQGAQPGAQP GAQPGAQPGA QPQGAQPGQ KP		48
SEQ ID NO: 45	moltype = AA length = 32	
FEATURE	Location/Qualifiers	
REGION	1..32	
note = Synthetic		
source	1..32 mol_type = protein organism = synthetic construct	
SEQUENCE: 45		
GAQPGAQPGA QPQGAQPGAQP GAQPGAQPGQ KP		32
SEQ ID NO: 46	moltype = AA length = 24	
FEATURE	Location/Qualifiers	
source	1..24 mol_type = protein organism = synthetic construct	
SEQUENCE: 46		
GAQPGAQPGA QPQGAQPGAQP GQKP		24

-continued

SEQ ID NO: 47 moltype = AA length = 40
 FEATURE Location/Qualifiers
 REGION 1..40
 note = Synthetic
 source 1..40
 mol_type = protein
 organism = synthetic construct
 SEQUENCE: 47
 GAQPGQAQPGA QPGGAQPGQGP GAQPGQAQPGA QPGGAQPGQCP 40

SEQ ID NO: 48 moltype = AA length = 48
 FEATURE Location/Qualifiers
 REGION 1..48
 note = Synthetic
 source 1..48
 mol_type = protein
 organism = synthetic construct
 SEQUENCE: 48
 GAQPGQAQPGA QPGGAQPGQGP GAQPGQAQPGA QPGGAQPGQGP GAQPGQCP 48

SEQ ID NO: 49 moltype = AA length = 32
 FEATURE Location/Qualifiers
 REGION 1..32
 note = Synthetic
 source 1..32
 mol_type = protein
 organism = synthetic construct
 SEQUENCE: 49
 GAQPGQAQPGA QPGGAQPGQGP GAQPGQAQPGQ CP 32

SEQ ID NO: 50 moltype = AA length = 24
 FEATURE Location/Qualifiers
 REGION 1..24
 note = Synthetic
 source 1..24
 mol_type = protein
 organism = synthetic construct
 SEQUENCE: 50
 GAQPGQAQPGA QPGGAQPGQGP GQCP 24

SEQ ID NO: 51 moltype = AA length = 71
 FEATURE Location/Qualifiers
 REGION 1..71
 note = Synthetic
 SITE 2
 note = MISC_FEATURE - X is Aib
 source 1..71
 mol_type = protein
 organism = synthetic construct
 SEQUENCE: 51
 HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG GGAQPGQAQPG AQPQGAQPGQ PGAQPGQPG 60
 AQPGQAQPGQK P 71

SEQ ID NO: 52 moltype = AA length = 71
 FEATURE Location/Qualifiers
 REGION 1..71
 note = Synthetic
 SITE 2
 note = MISC_FEATURE - X is Aib
 source 1..71
 mol_type = protein
 organism = synthetic construct
 SEQUENCE: 52
 HXEGTFTSDV SSYLEGQAAAR EFIAWLVRGG GGAQPGQAQPG AQPQGAQPGQ PGAQPGQPG 60
 AQPGQAQPGQK P 71

SEQ ID NO: 53 moltype = AA length = 79
 FEATURE Location/Qualifiers
 REGION 1..79
 note = Synthetic
 SITE 2
 note = MISC_FEATURE - X is Aib
 source 1..79
 mol_type = protein
 organism = synthetic construct
 SEQUENCE: 53
 HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG GGAQPGQAQPG AQPQGAQPGQ PGAQPGQPG 60
 AQPGQAQPGQK PGAQPGQCP 79

-continued

```

SEQ ID NO: 54      moltype = AA length = 63
FEATURE          Location/Qualifiers
REGION           1..63
note = Synthetic
SITE              2
note = MISC_FEATURE - X is Aib
source            1..63
mol_type = protein
organism = synthetic construct
SEQUENCE: 54
HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG GGAQPGQAQPG AQPQGAQPGQ PGAQPGQPG 60
QKP                           63

SEQ ID NO: 55      moltype = AA length = 63
FEATURE          Location/Qualifiers
REGION           1..63
note = Synthetic
SITE              2
note = MISC_FEATURE - X is Aib
source            1..63
mol_type = protein
organism = synthetic construct
SEQUENCE: 55
HXEGTFTSDV SSYLEGQAAAR EFIAWLVRGG GGAQPGQAQPG AQPQGAQPGQ PGAQPGQPG 60
QKP                           63

SEQ ID NO: 56      moltype = AA length = 55
FEATURE          Location/Qualifiers
REGION           1..55
note = Synthetic
SITE              2
note = MISC_FEATURE - X is Aib
source            1..55
mol_type = protein
organism = synthetic construct
SEQUENCE: 56
HXEGTFTSDV SSYLEGQAAAR EFIAWLVRGG GGAQPGQAQPG AQPQGAQPGQ PGQKP      55

SEQ ID NO: 57      moltype = AA length = 71
FEATURE          Location/Qualifiers
REGION           1..71
note = Synthetic
SITE              2
note = MISC_FEATURE - X is Aib
source            1..71
mol_type = protein
organism = synthetic construct
SEQUENCE: 57
HXEGTFTSDV SSYLEEQAAK EFIAWLVRGG GGAQPGQAQPG AQPQGAQPGQ PGAQPGQPG 60
AQPQGAQPGQK P      71

SEQ ID NO: 58      moltype = AA length = 79
FEATURE          Location/Qualifiers
REGION           1..79
note = Synthetic
SITE              2
note = MISC_FEATURE - X is Aib
source            1..79
mol_type = protein
organism = synthetic construct
SEQUENCE: 58
HXEGTFTSDV SSYLEEQAAK EFIAWLVRGG GGAQPGQAQPG AQPQGAQPGQ PGAQPGQPG 60
AQPQGAQPGQK P      79

SEQ ID NO: 59      moltype = AA length = 63
FEATURE          Location/Qualifiers
REGION           1..63
note = Synthetic
SITE              2
note = MISC_FEATURE - X is Aib
source            1..63
mol_type = protein
organism = synthetic construct
SEQUENCE: 59
HXEGTFTSDV SSYLEEQAAK EFIAWLVRGG GGAQPGQAQPG AQPQGAQPGQ PGAQPGQPG 60
QKP                           63

SEQ ID NO: 60      moltype = AA length = 71

```

-continued

FEATURE Location/Qualifiers
REGION 1..71
note = Synthetic

SITE 2
note = MISC_FEATURE - X is Aib
source 1..71
mol_type = protein
organism = synthetic construct

SEQUENCE: 60
HXEGTFTSDV SSYLEEQAAR EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
AQPQGAQPGQK P 71

SEQ ID NO: 61 moltype = AA length = 63
FEATURE Location/Qualifiers
REGION 1..63
note = Synthetic

SITE 2
note = MISC_FEATURE - X is Aib
source 1..63
mol_type = protein
organism = synthetic construct

SEQUENCE: 61
HXEGTFTSDV SSYLEEQAAR EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
QKP 63

SEQ ID NO: 62 moltype = AA length = 55
FEATURE Location/Qualifiers
REGION 1..55
note = Synthetic

SITE 2
note = MISC_FEATURE - X is Aib
source 1..55
mol_type = protein
organism = synthetic construct

SEQUENCE: 62
HXEGTFTSDV SSYLEEQAAR EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGQKP 55

SEQ ID NO: 63 moltype = AA length = 79
FEATURE Location/Qualifiers
REGION 1..79
note = Synthetic

SITE 2
note = MISC_FEATURE - X is Aib
source 1..79
mol_type = protein
organism = synthetic construct

SEQUENCE: 63
HXEGTFTSDV SSYLEGQAAC EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
AQPQGAQPGQCP 79

SEQ ID NO: 64 moltype = AA length = 71
FEATURE Location/Qualifiers
REGION 1..71
note = Synthetic

SITE 2
note = MISC_FEATURE - X is Aib
source 1..71
mol_type = protein
organism = synthetic construct

SEQUENCE: 64
HXEGTFTSDV SSYLEGQAAC EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
AQPQGAQPGQC P 71

SEQ ID NO: 65 moltype = AA length = 63
FEATURE Location/Qualifiers
REGION 1..63
note = Synthetic

SITE 2
note = MISC_FEATURE - X is Aib
source 1..63
mol_type = protein
organism = synthetic construct

SEQUENCE: 65
HXEGTFTSDV SSYLEGQAAC EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
QCP 63

SEQ ID NO: 66 moltype = AA length = 71
FEATURE Location/Qualifiers
REGION 1..71

-continued

```

SITE          note = Synthetic
             2
source        note = MISC_FEATURE - X is Aib
             1..71
             mol_type = protein
             organism = synthetic construct
SEQUENCE: 66
HXEGTFTSDV SSYLEGQAAR EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
AQPQGAQPGQC P                               71

SEQ ID NO: 67      moltype = AA  length = 63
FEATURE          Location/Qualifiers
REGION           1..63
SITE            note = Synthetic
               2
source          note = MISC_FEATURE - X is Aib
               1..63
             mol_type = protein
             organism = synthetic construct
SEQUENCE: 67
HXEGTFTSDV SSYLEGQAAR EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
QCP                                         63

SEQ ID NO: 68      moltype = AA  length = 55
FEATURE          Location/Qualifiers
REGION           1..55
SITE            note = Synthetic
               2
source          note = MISC_FEATURE - X is Aib
               1..55
             mol_type = protein
             organism = synthetic construct
SEQUENCE: 68
HXEGTFTSDV SSYLEGQAAR EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGQCP      55

SEQ ID NO: 69      moltype = AA  length = 71
FEATURE          Location/Qualifiers
REGION           1..71
SITE            note = Synthetic
               1..71
source          mol_type = protein
               organism = synthetic construct
SEQUENCE: 69
HGEGTFTSDV SSYLEEQAAAR EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
AQPQGAQPGQC P                               71

SEQ ID NO: 70      moltype = AA  length = 63
FEATURE          Location/Qualifiers
REGION           1..63
SITE            note = Synthetic
               1..63
source          mol_type = protein
               organism = synthetic construct
SEQUENCE: 70
HGEGTFTSDV SSYLEEQAAAR EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
QCP                                         63

SEQ ID NO: 71      moltype = AA  length = 55
FEATURE          Location/Qualifiers
REGION           1..55
SITE            note = Synthetic
               1..55
source          mol_type = protein
               organism = synthetic construct
SEQUENCE: 71
HGEGTFTSDV SSYLEEQAAAR EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGQCP      55

SEQ ID NO: 72      moltype = AA  length = 79
FEATURE          Location/Qualifiers
REGION           1..79
SITE            note = Synthetic
               2
source          note = MISC_FEATURE - X is Aib
               1..79
             mol_type = protein
             organism = synthetic construct
SEQUENCE: 72
HXEGTFTSDV SSYLEEQAAC EFIAWLVRGG GGAQPGAQPG AQPQGAQPGAQ PGAQPGAQPG 60
AQPQGAQPGQAQ PGAQPGQCP                  79

```

-continued

```

SEQ ID NO: 73      moltype = AA length = 71
FEATURE          Location/Qualifiers
REGION           1..71
                  note = Synthetic
SITE              2
                  note = MISC_FEATURE - X is Aib
source             1..71
                  mol_type = protein
                  organism = synthetic construct
SEQUENCE: 73
HXEGTFTSDV SSYLEEQAAC EFIAWLVRGG GGAQPGAQPG AQPQGAQPGQ PGAQPGAQPG 60
AQPQGAQPGQC P                                71

SEQ ID NO: 74      moltype = AA length = 63
FEATURE          Location/Qualifiers
REGION           1..63
                  note = Synthetic
SITE              2
                  note = MISC_FEATURE - X is Aib
source             1..63
                  mol_type = protein
                  organism = synthetic construct
SEQUENCE: 74
HXEGTFTSDV SSYLEEQAAC EFIAWLVRGG GGAQPGAQPG AQPQGAQPGQ PGAQPGAQPG 60
QCP                                         63

SEQ ID NO: 75      moltype = AA length = 71
FEATURE          Location/Qualifiers
REGION           1..71
                  note = Synthetic
SITE              2
                  note = MISC_FEATURE - X is Aib
source             1..71
                  mol_type = protein
                  organism = synthetic construct
SEQUENCE: 75
HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG GGAQPGAQPG AQPQGAQPGQ PGAQPGAQPG 60
AQPQGAQPGQC P                                71

SEQ ID NO: 76      moltype = AA length = 63
FEATURE          Location/Qualifiers
REGION           1..63
                  note = Synthetic
SITE              2
                  note = MISC_FEATURE - X is Aib
source             1..63
                  mol_type = protein
                  organism = synthetic construct
SEQUENCE: 76
HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG GGAQPGAQPG AQPQGAQPGQ PGAQPGAQPG 60
QCP                                         63

SEQ ID NO: 77      moltype = AA length = 55
FEATURE          Location/Qualifiers
REGION           1..55
                  note = Synthetic
SITE              2
                  note = MISC_FEATURE - X is Aib
source             1..55
                  mol_type = protein
                  organism = synthetic construct
SEQUENCE: 77
HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG GGAQPGAQPG AQPQGAQPGQ PGQCP      55

SEQ ID NO: 78      moltype = AA length = 71
FEATURE          Location/Qualifiers
REGION           1..71
                  note = Synthetic
SITE              2
                  note = MISC_FEATURE - X is Aib
source             1..71
                  mol_type = protein
                  organism = synthetic construct
SEQUENCE: 78
HXEGTFTSDV SSYLEEQAAK EFIAWLVRGG GGAQPGAQPG AQPQGAQPGQ PGAQPGAQPG 60
AQPQGAQPGQC P                                71

SEQ ID NO: 79      moltype = AA length = 63

```

-continued

FEATURE Location/Qualifiers
 REGION 1..63
 note = Synthetic
 SITE 2
 note = MISC_FEATURE - X is Aib
 source 1..63
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 79
 HXEGTFTSDV SSYLEEQAAK EFIGLVRGG GGAQPGQAQPG AQPQGAQPGQ PGAQPGQPG 60
 QCP 63

SEQ ID NO: 80 moltype = AA length = 55
 FEATURE Location/Qualifiers
 REGION 1..55
 note = Synthetic
 SITE 2
 note = MISC_FEATURE - X is Aib
 source 1..55
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 80
 HXEGTFTSDV SSYLEEQAAK EFIGLVRGG GGAQPGQAQPG AQPQGAQPGQ PGQCP 55

SEQ ID NO: 81 moltype = AA length = 11
 FEATURE Location/Qualifiers
 source 1..11
 mol_type = protein
 organism = Peptostreptococcus magnus

SEQUENCE: 81 DICLPRWGCL W 11

SEQ ID NO: 82 moltype = AA length = 60
 FEATURE Location/Qualifiers
 REGION 1..60
 note = Synthetic
 source 1..60
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 82 GAQPGAQPGA QPQGAQPGAQP GAQPGAQPGA QPQGAQPGAQP GAQPGAQPGA QPQGAQPGQKP 60

SEQ ID NO: 83 moltype = AA length = 24
 FEATURE Location/Qualifiers
 REGION 1..24
 note = Synthetic
 source 1..24
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 83 GAQPGAQPGA QPQGAQPGAQP GQKP 24

SEQ ID NO: 84 moltype = AA length = 12
 FEATURE Location/Qualifiers
 REGION 1..12
 note = Synthetic
 source 1..12
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 84 GAQPGAQPGQ KP 12

SEQ ID NO: 85 moltype = AA length = 91
 FEATURE Location/Qualifiers
 REGION 1..91
 note = Synthetic
 SITE 2

-continued

```

source          note = MISC_FEATURE - X is Aib
                1..91
                mol_type = protein
                organism = synthetic construct

SEQUENCE: 85
HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG GGAQPGAQPG AQPQGAQPGQ PGAQPGAQPG 60
AQPQGAQPGQ PGAQPGAQPG AQPQGAQPGK P                                91

SEQ ID NO: 86      moltype = AA length = 55
FEATURE          Location/Qualifiers
REGION           1..55
                note = Synthetic
SITE              2
                note = MISC_FEATURE - X is Aib
source            1..55
                mol_type = protein
                organism = synthetic construct

SEQUENCE: 86
HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG GGAQPGAQPG AQPQGAQPGQ PGQKP      55

SEQ ID NO: 87      moltype = AA length = 43
FEATURE          Location/Qualifiers
REGION           1..43
                note = Synthetic
SITE              2
                note = MISC_FEATURE - X is Aib
source            1..43
                mol_type = protein
                organism = synthetic construct

SEQUENCE: 87
HXEGTFTSDV SSYLEGQAAK EFIAWLVRGG GGAQPGAQPG QKP                  43

SEQ ID NO: 88      moltype = AA length = 43
FEATURE          Location/Qualifiers
REGION           1..43
                note = Synthetic
SITE              2
                note = MISC_FEATURE - X is Aib
source            1..43
                mol_type = protein
                organism = synthetic construct

SEQUENCE: 88
HXEGTFTSDV SSYLEGQAR EFIAWLVRGG GGAQPGAQPG QKP                  43

```

The invention claimed is:

1. A liquid pharmaceutical composition, comprising a polypeptide conjugate and a pharmaceutically acceptable excipient, wherein:

the polypeptide conjugate comprises a polypeptide portion and a conjugate portion,

the polypeptide portion comprises a single biologically active peptide and a peptide linker, wherein the biologically active peptide is attached to N-terminus of the peptide linker and comprises GLP-1; and

the conjugate portion comprises a first clearance-reducing moiety (CRM) conjugated to a first CRM residue in the peptide linker, and a second CRM conjugated to a second CRM residue in the polypeptide portion,

wherein the first CRM residue and the second CRM residue are both lysine residues, and the polypeptide conjugate comprises only two lysine residues,

wherein:

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 51, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 53, and is conjugated with the first CRM and the second CRM respectively at 26K and 84K;

40 the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 54, and is conjugated with the first CRM and the second CRM respectively at 26K and 68K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the second CRM respectively at 26K and 76K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 58, and is conjugated with the first CRM and the second CRM respectively at 26K and 84K;

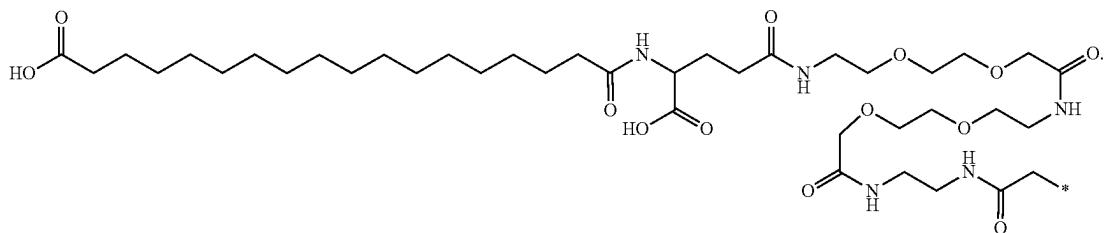
the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 59, and is conjugated with the first CRM and the second CRM respectively at 26K and 68K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 85, and is conjugated with the first CRM and the second CRM respectively at 26K and 96K;

the polypeptide portion comprises an amino acid sequence of SEQ ID NO: 86, and is conjugated with the first CRM and the second CRM respectively at 26K and 60K;

wherein the first clearance-reducing moiety (CRM) and the second CRM both have the structure of below formula:

101

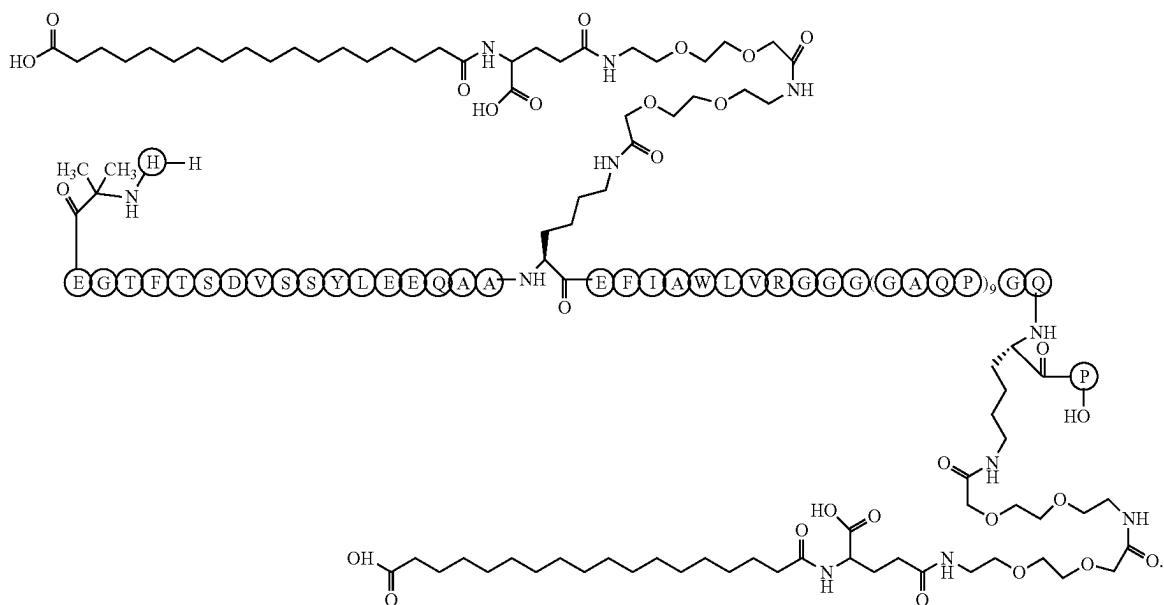


wherein the pharmaceutical composition has about 20-120 mg/ml of the polypeptide conjugate; and wherein the pharmaceutical composition has a pH of 15 about 6.5 to about 7.8.

2. The pharmaceutical composition of claim 1, wherein the polypeptide conjugate has the structure shown below:

11. The pharmaceutical composition of claim 3, wherein the pharmaceutical excipient further comprises a preservative, a chelating agent, and/or a stabilizer.

12. The pharmaceutical composition of claim 3, wherein the pharmaceutical composition has a pH of about 6.5 to 7.4.



3. The pharmaceutical composition of claim 1, wherein the pharmaceutically acceptable excipient comprises a buffer and an isotonic agent.

4. The pharmaceutical composition of claim 3, wherein the buffer is a phosphate buffer.

5. The pharmaceutical composition of claim 4, wherein the phosphate buffer is disodiumhydrogen phosphate dodecahydrate, wherein the disodiumhydrogen phosphate dodecahydrate is present in a concentration of about 0.01-15 mg/mL.

6. The pharmaceutical composition of claim 4, wherein the phosphate buffer is disodium phosphate dihydrate.

7. The pharmaceutical composition of claim 3, wherein the buffer is a citrate buffer.

8. The pharmaceutical composition of claim 3, wherein the buffer is a histidine buffer.

9. The pharmaceutical composition of claim 3, wherein the isotonic agent is sodium chloride, or wherein the isotonic agent is propylene glycol.

10. The pharmaceutical composition of claim 3, wherein the isotonic agent is mannitol.

45 13. The pharmaceutical composition of claim 3, wherein
the pharmaceutical composition has about 20-80 mg/ml, or
20-100 mg/ml of the polypeptide conjugate.

14. The pharmaceutical composition of claim 3, comprising:

about 20-120 mg/ml of the polypeptide conjugate; a buffer selected from the group consisting of phosphate buffer, citrate buffer, acetate buffer, histidine buffer, glycine buffer, carbonate buffer, borate buffer, glutamate buffer, glycylglycine buffer, lysine buffer, and arginine buffer;

an isotonic agent selected from the group consisting of sodium chloride, glycerol, sorbitol, sucrose, propylene glycol, mannitol, glycine, lactose monohydrate, arginine, myoinositol and dimethylsulfone; and a pH of about 6.5 to about 7.8.

15. The pharmaceutical composition of claim 3, comprising:
about 20-80 mg/mL, 20-100 mg/ml or 20-120 mg/mL of
the polypeptide conjugate, wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57 and is

103

conjugated with the first CRM and the second CRM respectively at 26K and 76K;
about 0.5-5 mg/mL phosphate buffer, about 1-50 mM citrate buffer, or about 0.5-10 mg/mL histidine buffer; an isotonic agent selected from the group consisting of 5 5-15 mg/ml sodium chloride, 1-50 mg/mL propylene glycol, and 30-50 mg/mL mannitol; and a pH of about 6.5 to about 7.8.

16. The pharmaceutical composition of claim 1, comprising:
about 20-80 mg/mL, 20-100 mg/ml or 20-120 mg/mL of 10 the polypeptide conjugate,
wherein the polypeptide portion of the polypeptide conjugate comprises an amino acid sequence of SEQ ID NO: 57, and is conjugated with the first CRM and the 15 second CRM respectively at 26K and 76K;
about 0.5-5 mg/mL phosphate buffer;
5-15 mg/ml sodium chloride; and
a pH of about 6.5 to about 7.8.

* * * * *

10

20