SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 14-064524- -00000-0000

Fecha: 2014-03-26 16:48:40 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 6

Industria y Comercio

SUPERINTENDENCIA

DELEGATURA DE PROPIEDAD INDUSTRIAL

DIRECCIÓN DE NUEVAS CREACIONES

SOLICITUD DE PATENTE DE INVENCIÓN

TITULO: PROTEÍNAS DE FUNCIÓN DUAL PARA TRATAR TRASTORNOS METABÓLICOS

SOLICITANTE: NOVARTIS AG.

DIRECCIÓN: Lichtstrasse 35, CH-4056 Basel, Suiza

APODERADO: CARLOS R. OLARTE

BOGOTÁ: 26 de marzo de 2014

P2014/000040

OLARTEM
OLARTE MOURE & ASO(

Industria y Comercio
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 14-064524- -00000-0000

Fecha: 2014-03-26 16:48:40 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 6

1

PETITORIO - FORMULARIO ÚNICO DE SOLICITUD
PATENTE PCT
ENTRADA EN FASE NACIONAL

P2014/000040

SOLICITUD DE:

(1)

[X] Patente de Invención. [ ] Patente de Modelo de Utilidad.

[X] Capítulo I. [ ] Capítulo II.

(2)

SOLICITANTE (71)

Nombre: **NOVARTIS AG y IRM LLC**
Dirección: **(1)** Lichtstrasse 35, CH-4056 Basel **(2)** 131 Front Street, PO Box HM 2899, Bermuda, HM LX (BM).
Nacionalidad o Domicilio: **(1)**Suiza **(2)** Bermuda
Lugar de Constitución: Suiza
Teléfono: **Fax:**
E-mail:

IDENTIFICACIÓN
C.C. [ ] **NIT** [ ]
C.E. [ ] **Otro** [ ]
Cual
Número

REPRESENTANTE O APODERADO

Nombre: CARLOS R. OLARTE
Dirección: Carrera 5 No 34-03 Bogotá D.C.
Teléfono: 601-7700 **Fax:** 601-7799
E-mail: office@olartemoure.com

IDENTIFICACIÓN
C.C. [X] **NIT** [ ]
C.E. [ ] **Otro** [ ]
Cual
Número 79782747Bta
TP 74.295

INVENTOR (ES) (72)

Nombre: VER PAGINA ADJUNTA
Dirección: VER PAGINA ADJUNTA
Nacionalidad o Domicilio: VER PAGINA ADJUNTA

(5)

PCT (86)

Solicitud Internacional No. PCT/US2012/057371 Fecha Septiembre 26, 2012

Publicación Internacional No. WO 2013/049234 Fecha Abril 04, 2013

(6)

Título (54) PROTEÍNAS DE FUNCIÓN DUAL PARA TRATAR TRASTORNOS METABÓLICOS

Clasificación Internacional (51) C07K 14/435, C07K 14/50, C07K 14/605.

Prioridad

SI [X] **NO** [ ]

(33) País de Origen	(32) Fecha	(31) Número de Solicitud
US	Septiembre 26, 2011	61/539,290

(9)

Comprobante de pago No. 14- **Fecha**

(10) **ANEXOS**

- [x] Comprobante de pago de la tasa de presentación de la solicitud.
- [ ] Documento que acredita la existencia y representación legal cuando el solicitante sea persona jurídica.
- [x] Poderes, si fuere el caso.
- [ ] Documento de cesión del inventor al solicitante o a su causante.
- [x] Declaraciones. Declaración bajo la Regla 4.17 (ii) del PCT.
- [x] Copia de la solicitud internacional. (Resumen, descripción, reivindicaciones, dibujos).
- [x] Traducción de la solicitud internacional al castellano. (Resumen, descripción, reivindicaciones, dibujos).
- [ ] Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen Preliminar (IPEA).
- [ ] Traducción del examen preliminar internacional efectuado por la IPEA.
- [ ] De ser el caso, copia del contrato de acceso.
- [ ] De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas.
- [ ] De ser el caso, certificado de depósito del material biológico.
- [ ] Arte final 12 x 12.
- [ ] De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- [ ] De ser el caso, modificaciones efectuadas a la solicitud internacional.
- [x] Otros. Búsqueda Internacional

(11) **FIGURA CARACTERÍSTICA:**

(12) Solicito la concesión de la patente,

NOMBRE: CARLOS R. OLARTE FIRMA:

C.C. 79.782.747 Btá T.P. 74.295

LISTA DE INVENTORES

PCT/US2012/057371

Brian R. BOETTCHER
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Estadounidense

Shari L. CAPLAN
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Estadounidense

Susan E. CELLITTI
10675 John Jay Hopkins Drive,
San Diego, California 92121
Estados Unidos de América
Nacionalidad: Estadounidense

Douglas S. DANIELS
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Estadounidense

Norio HAMAMATSU
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Japonesa

Bernhard Hubert GEIERSTANGER
10675 John Jay Hopkins Drive,
San Diego, California 92121
Estados Unidos de América
Nacionalidad: Alemania

Stuart LICHT
100 Memorial Drive,
Cambridge, Massachusetts 02142
Estados Unidos de América
Nacionalidad: Estadounidense

Andreas LOEW
250 Massachusetts Avenue,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Alemania

Stephen Craig WELDON
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de America
Nacionalidad: Estadounidense

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
NIT : 800.176.089-2
- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 14 - 0029535

Bogotá D.C., Marzo 17 de 2014 - 13:53:00

RECIBIDO DE : OLARTE MOURE & ASOCIADOS NI 830.122.870

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	609953573	14/03/2014	17.818.250.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr UNDITARIO	Vr CONCEPTO
1	50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE INVENCION	515.000.00	515.000.00
				$515.000.00

SON: **QUINIENTOSQUINCE MIL PESOS MONEDA CORRIENTE***

Responsable: _____
-****-

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 14-064524- -00000-0000

Fecha: 2014-03-26 16:48:40 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 6

5

Industria y Comercio
SUPERINTENDENCIA

PLANILLA DE RADICACIÓN - DIGITALIZACIÓN DIARIA
RELACIÓN DE ANEXOS

Dependencia:	Fecha: 2014 03 26 (AAAA MM DD)	Recorrido:
2020		4

Tipo de Radicación: Entrada: X Salida: ☐ Traslado: ☐ Convenio: ☐

Radicador: | Planilla:

ID	*No. Radicación*	*Consecutivo*	*Item*
1	14-645'24	0	7
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			

Item	*Unidad de Conservación*
1	Bolsa
2	Caja
3	Libro
4	Sobre
5	Tomo
6	**Otro**
Item	***Soporte Documental***
X	CD
8	Disquete
9	DVD
10	Folleto
11	Foto
12	Periódico
13	Plano
14	Publicación Periódica
15	USB
16	**Otro**

Observaciones:

EL CD ANEXO CONTIENE UN TOTAL DE 9 ARCHIVOS:

9 EN FORMATO PDF, LOS CUALES FUERON PROCESADOS

EL CD SE ENCUENTRA EN EL FOLIO (5)

JP

Industria y Comercio
SUPERINTENDENCIA

LISTA DE CHEQUEO
ADMISIÓN A TRÁMITE – NUEVAS CREACIONES

PATENTE DE INVENCIÓN ☐ MODELO DE UTILIDAD ☐ Art 33 Decisión 486/00

- ☐ Indicación que se solicita una patente.
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Descripción de la invención
- ☐ Dibujos de ser estos pertinentes
- ☐ Comprobante de pago de las tasas establecidas (De ser el caso formato de descuento)

Completa ☐ Incompleta ☐

PATENTE DE INVENCIÓN PCT ☒ MODELO DE UTILIDAD PCT ☐ Art.33 Decisión 486/00, Circular Única

- ☒ Indicación que se solicita una PCT
- ☒ Copia de la solicitud en español, tal como fue presentada Inicialmente (capítulo descriptivo, reivindicatorio, resumen)
- ☒ Dibujos de ser estos pertinentes
- ☒ Comprobante de pago de las tasas establecidas (de ser el caso formato de descuento)

Completa ☒ Incompleta ☐

DISEÑO INDUSTRIAL ☐ (Art. 119 Decisión 486/00)

- ☐ Indicación que se solicita Diseño industrial
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Representación gráfica y fotográfica del Diseño industrial o muestra del material que incorpora el diseño
- ☐ Comprobante de pago de las tasas establecidas

Completa ☐ Incompleta ☐

ESQUEMA DE TRAZADO ☐ (Art. 92 Decisión 486/00)

- ☐ Indicación que se solicita un esquema de trazado
- ☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
- ☐ Representación gráfica de un esquema de trazado
- ☐ Comprobante de pago de las tasas establecidas

Completa ☐ Incompleta ☐

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 14-064524- -00000-0000

Fecha: 2014-03-26 16:48:40 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 6

OLARTEMOURE

OLARTE MOURE & ASOCIADOS LTDA.

Industria y Comercio
SUPERINTENDENCIA

Espacio reservado para el adhesivo de radicación.

PETITORIO - FORMULARIO ÚNICO DE SOLICITUD
PATENTE PCT
ENTRADA EN FASE NACIONAL **P2014/000040**

SOLICITUD DE:

1

[X] Patente de Invención. [ ] Patente de Modelo de Utilidad.

[X] Capítulo I. [ ] Capítulo II.

2

SOLICITANTE (71)

Nombre: **NOVARTIS AG y IRM LLC**
Dirección: **(1)** Lichtstrasse 35, CH-4056 Basel **(2)** 131 Front Street, PO Box HM 2899, Bermuda, HM LX (BM).
Nacionalidad o Domicilio: **(1)**Suiza **(2)** Bermuda

Lugar de Constitución: Suiza

Teléfono: _____ **Fax:** _____
E-mail: _____

IDENTIFICACIÓN

C.C. [ ] **NIT** [ ]
C.E. [ ] **Otro** [ ]
Cual _____
Número _____

REPRESENTANTE O APODERADO

Nombre: CARLOS R. OLARTE
Dirección: Carrera 5 No 34-03 Bogotá D.C.
Teléfono: 601-7700 **Fax:** 601-7799
E-mail: office@olartemoure.com

IDENTIFICACIÓN

C.C. [X] **NIT** [ ]
C.E. [ ] **Otro** [ ]
Cual _____
Número 79782747Bta
TP 74.295

INVENTOR (ES) (72)

Nombre: VER PAGINA ADJUNTA
Dirección: VER PAGINA ADJUNTA
Nacionalidad o Domicilio: VER PAGINA ADJUNTA

5

PCT (86)

Solicitud Internacional No. PCT/US2012/057371 Fecha Septiembre 26, 2012

Publicación Internacional No. WO 2013/049234 Fecha Abril 04, 2013

6

Título (54) PROTEÍNAS DE FUNCIÓN DUAL PARA TRATAR TRASTORNOS METABÓLICOS

Clasificación Internacional (51) C07K 14/435, C07K 14/50, C07K 14/605.

Prioridad

SI [X] **NO** [ ]

(33) País de Origen	(32) Fecha	(31) Número de Solicitud
US	Septiembre 26, 2011	61/539,290

9

Comprobante de pago No. 14- **Fecha** _____

ANEXOS

- [x] Comprobante de pago de la tasa de presentación de la solicitud.
- [ ] Documento que acredita la existencia y representación legal cuando el solicitante sea persona jurídica.
- [x] Poderes, si fuere el caso. Adjunto copia simple
- [ ] Documento de cesión del inventor al solicitante o a su causante.
- [ ] Declaraciones. Declaración bajo la regla 4.17 (ii) del PCT.
- [x] Copia de la solicitud internacional. (Resumen, descripción, reivindicaciones, dibujos).
- [x] Traducción de la solicitud internacional al castellano. (Resumen, descripción, reivindicaciones, dibujos).
- [ ] Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen Preliminar (IPEA).
- [ ] Traducción del examen preliminar internacional efectuado por la IPEA.
- [ ] De ser el caso, copia del contrato de acceso.
- [ ] De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas.
- [ ] De ser el caso, certificado de depósito del material biológico.
- [ ] Arte final 12 x 12.
- [ ] De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- [ ] De ser el caso, modificaciones efectuadas a la solicitud internacional.
- [x] Otros. Búsqueda Internacional

(11) **FIGURA CARACTERÍSTICA:**

(12) Solicito la concesión de la patente,

NOMBRE: CARLOS R. OLARTE FIRMA: _____

C.C. 79.782.747 Btá T.P. 74.295

LISTA DE INVENTORES

PCT/US2012/057371

Brian R. BOETTCHER
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Estadounidense

Shari L. CAPLAN
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Estadounidense

Susan E. CELLITTI
10675 John Jay Hopkins Drive,
San Diego, California 92121
Estados Unidos de América
Nacionalidad: Estadounidense

Douglas S. DANIELS
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Estadounidense

Norio HAMAMATSU
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Japonesa

Bernhard Hubert GEIERSTANGER
10675 John Jay Hopkins Drive,
San Diego, California 92121
Estados Unidos de América
Nacionalidad: Alemania

Stuart LICHT
100 Memorial Drive,
Cambridge, Massachusetts 02142
Estados Unidos de América
Nacionalidad: Estadounidense

Andreas LOEW
250 Massachusetts Avenue,
Cambridge, Massachusetts 02139
Estados Unidos de América
Nacionalidad: Alemania

Stephen Craig WELDON
100 Technology Square,
Cambridge, Massachusetts 02139
Estados Unidos de America
Nacionalidad: Estadounidense

TRADUCCIÓN OFICIAL DE UN DOCUMENTO QUE ESTÁ ORIGINALMENTE EN INGLÉS

Nota del traductor:

El documento adjunto de poder de Novartis AG a Carlos R. Olarte y/o J. Ian Raisbeck y/o Ana Maria Frieri y/o Juan Guillermo Moure y/o Natalia Castro lleva los siguientes datos en inglés.

CERTIFICACIÓN NOTARIAL

1. En esta fecha, día 30 de Agosto de 2006, yo, el suscrito Notario Público certifico por medio del presente la autenticidad de las firmas del señor Arthur Canonica, ciudadano de Corticiasca, Suiza, residente en Wisen/SO, Suiza, y del Sr. Pascal Traub, ciudadano de Riehen, Suiza, residente en Basilea, Suiza, ambos personas mayores de edad y con dirección comercial en Lichstrasse 35, 4056, Basilea, Suiza. La autenticidad de la firmas fue establecida por medio de la comparación.
2. Además, yo, el Notario Público, certifico por el presente que las anteriores personas han ejecutado el anterior documento como apoderados de Novartis AG.
3. Además, yo, el Notario Público certifico por el presente que Novartis AG, con su principal lugar de negocios en Basilea, Suiza, está debidamente constituida, está actualmente vigente y que los propósitos para los cuales se otorgó el Documento de Poder están dentro del alcance de sus propósitos corporativos.

Traductor Oficial
JORGE E. SIERRA REYES
Licencia 376 - 1993 Miniusticia

4. Además, yo, el Notario Público, certifico por el presente que las anteriores personas tienen la autoridad para ejecutar el documento de Poder y que por este instrumento, su representación es legal.

La base para mis certificaciones de los puntos 2, 3 y 4 fue la exhibición de los siguientes documentos auténticos:

- Documento de Poder de Novartis AG fechado el 19 de junio de 2006 y
- Extracto del Registro de Comercio del Cantón de la ciudad de Basilea para Novartis AG, en Basilea, fechado el 25 de enero de 2006.

Basilea, Suiza, este día 30 de Agosto de 2006.

(Firmado) Dr. Matthias Staehelin, Notario Público del Cantón de la ciudad de Basilea, Suiza.

Mi licencia vence el 11 de Marzo de 2009.

Hay una Apostilla de Suiza, firmada por Heidi Fischer, bajo el No. 8049/13584.

Traductor Oficial
JORGE E. SIERRA REYES
Licencia 376 - 1993 Minjusticia

Certificación:

Certifico que la presente es una traducción verdadera y completa del documento adjunto: "Novartis"

Fecha: 12 de Septiembre de 2006.

Traductor: Jorge Enrique Sierra Reyes.

Licencia No. 376 de 1993, del Ministerio de Justicia.

Dirección: Carrera 15 No. 99-13 Oficina 411. Bogotá, Colombia.

Teléfono: (57-1) 2182308, (57-1) 6232925.

E-mail: treseser1@yahoo.com. No. 8346

Traductor Oficial
JORGE E. SIERRA REYES
Licencia 376 - 1993 Minjusticia

OLARTERAISBECK
OLARTE RAISBECK & FRIERI LTDA.

Carlos R. Olarte
J. Ian Raisbeck
Ana María Frieri
Juan Guillermo Moure
Natalia Castro

Zayde Chahin
Jacqueline Chaparro
José Luis Londoño
Felipe Mendoza
Simón Salinas

Of Counsel
Guillermo Chahin
Graciela Melo

Power of Attorney

Poder

The undersigned,

El suscrito,

NOVARTIS AG

domiciled in

domiciliado en

Lichtstrasse 35, CH-4056 Basel / Switzerland

by means of these presents hereby grants to **Carlos R. Olarte and/or J. Ian Raisbeck and/or Ana María Frieri and/or Juan Guillermo Moure and/or Natalia Castro** full and sufficient Power of Attorney to file, prosecute, obtain, defend and enforce any and all of the company's **Patents for Inventions, Utility Models, Layout-Designs (topographies) of Integrated Circuits, Industrial Designs, Trademarks, Advertising Slogans, Trade Names, Commercial Emblems, Geographical Indications, Protection of Trade Secrets, Obtainer's Certificates, Copyrights, Registration of Domain Names, Health Registrations, or any other intellectual property right** in the Republic of Colombia, to which end it empowers them to take all steps necessary before any administrative or judicial authority of any class for the objects stated, to file petitions, recourses, replies and oppositions, pay taxes and annuities, request testimonies, receive documents and securities, abandon applications and collect refunds. Sufficient power is hereby granted to receive notifications on behalf of the Company, to confer powers on those who are to represent the company judicially and extrajudicially, to answer in court in the matter of all claims and demands that may be presented in connection with the objects stated, giving them likewise power to substitute these presents and revoke such substitutions.

por medio del presente otorga a **Carlos R. Olarte y/o J. Ian Raisbeck y/o Ana María Frieri y/o Juan Guillermo Moure y/o Natalia Castro** poder especial, amplio y suficiente para presentar, tramitar, obtener, defender y hacer respetar todas y cualquiera de las **Patentes de Invención, Modelos de Utilidad, Esquemas de Trazados de Circuitos Integrados, Diseños Industriales, Marcas, Lemas Comerciales, Nombres Comerciales, Enseñas Comerciales, Denominaciones de Origen, Protección de Secretos Empresariales, Certificados de Obtentor, Derechos de Autor, Registro de Nombres de Dominio, Registros Sanitarios, o cualquier otro bien de propiedad intelectual** de la compañía en la República de Colombia, a cuyo efecto los faculta para dar todos los pasos necesarios ante cualquier autoridad administrativa o judicial de cualquier jerarquía para cumplir con lo encomendado, elevar solicitudes, recursos, respuestas y oposiciones, pagar impuestos y anualidades, solicitar testimonios, recibir documentos y valores, desistir y percibir. Se concede poder bastante para recibir notificaciones a nombre de la compañía, otorgar poderes a quienes vayan a representar a la compañía judicial y extrajudicialmente, responder en juicio a todas las reclamaciones o demandas que por motivo de lo encomendado se presentaren, dándoles asimismo facultad para sustituir el presente y revocar dichas sustituciones.

Signature/Firma:

Arthur Canonica

Pascal Traub

Date/Fecha: June 27, 2006

City,Country/Ciudad,País: Basel, Switzerland

Attestation

1 On this date, the 30th (thirtieth) day of August 2006 (two thousand and six), I, the undersigned Notary Public certify herewith the authenticity of the signatures of **Mr. Arthur Canonica**, citizen of Corticiasca, Switzerland, residing in Wisen/SO, Switzerland and of **Mr. Pascal Traub**, citizen of Riehen, Switzerland, residing in Basel, Switzerland, both of legal age and with business address at, Lichtstrasse 35, 4056 Basel, Switzerland. The authenticity of the signatures was established by means of comparison.

2 Further, I, the Notary Public certify herewith that the foregoing persons have executed the foregoing document as proxy holders for Novartis AG.

3 Further, I, the Notary Public certify herewith that Novartis AG, having its principal place of business located in Basel / Switzerland, is duly incorporated, is currently in existence, and that the purposes for which the Power of Attorney is granted are within the scope of its corporate purposes.

4 Further, I, the Notary Public certify herewith that the foregoing persons have the authority to execute the Power of Attorney and that for this instrument their representation is legal.

The foundation for my certifications contained in points 2 (two), 3 (three) and 4 (four) was the exhibition of the following authentic documents:

- Power of Attorney by Novartis AG dated June 19 (nineteen), 2006 (two thousand and six); and
- Extract from the Register of Commerce of the Canton of Basel-City for Novartis AG, in Basel, dated January 25 (twenty-five), 2006 (two thousand and six).

BASEL, Switzerland, this 30th (thirtieth) day of August 2006 (two thousand and six)

M. Staehelin Notar

Dr. Matthias Staehelin, Notary Public
in the Canton of Basel-City, Switzerland

My license expires on March 11 (eleven), 2009 (two thousand and nine).

Leg.Prot.Nr. 975 /2006

APOSTILLE

(Hague Convention of October 5th, 1961)

1. Country: Switzerland (Schweiz / Suisse)
   This official document
2. is signed by .....
   .....
3. in his/her function as .....
4. and certified by the seal of the .....
   .....

Certified 24. May. 2006

5. in Basel (Bâle) 6. on .....
7. by the State Chancery of the Canton of Basel-Stadt .....
8. no. 2'049'13'389
9. Seal / stamp: 10. Signature:

Heidi Fischer

STATE CHANCERY * BASLE-CITY *

ESCRITURA PÚBLICA NÚMERO:

SEIS MIL VEINTICUATRO (6024)- - - - - - - -

FECHA DE OTORGAMIENTO:

TRECE (13) DE JULIO DE DOS MIL ONCE (2011)- - - -

NOTARIA TREINTA Y OCHO (38) DEL CÍRCULO DE BOGOTA D.C. ----------

SUPERINTENDENCIA DE NOTARIADO Y REGISTRO.----------

CÓDIGO NOTARIAL 1100100038. ----------

FORMULARIO DE CALIFICACIÓN. ----------

CLASE DE ACTO: PROTOCOLIZACIÓN DE DOCUMENTOS.----------

PERSONAS QUE INTERVIENEN: CARLOS REINALDO OLARTE, identificado con la cédula de ciudadanía número 79.782.747 expedida en Bogotá.----------

En la ciudad de Bogotá, Distrito Capital, Departamento de Cundinamarca, República de Colombia, a los TRECE ------(13-) días del mes de JULIO- - - de dos mil once (2011), ante mí EDUARDO DURAN GOMEZ- - - - - - - - - - - - -

NOTARIO(A) TREINTA Y OCHO (38) DEL CIRCULO DE BOGOTA D.C.--------

Compareció: **CARLOS REINALDO OLARTE,** mayor de edad, domiciliado y residente en la ciudad de Bogotá D.C., identificado con la cédula de ciudadanía número 79.782.747 expedida en Bogotá, de estado civil casado con sociedad conyugal vigente, quien obra en nombre propio y manifestó:----------

PRIMERO.- Presenta para su protocolización los siguientes documentos: ----------

1. Diez (10) folios útiles, en los que consta El **Poder General** otorgado por IRM LLC. al señor CARLOS R. OLARTE, de fecha 15 de Junio de 2011 ante el Notario Público en San Diego, California, US, texto en ingles y español y su respectiva Apostilla.----------

SEGUNDO.- Que de conformidad con lo expuesto y atendiendo a las prescripciones del artículo 56, 57 y 58 del Decreto 960 de 1970, los documentos mencionados, contenidos en diez (10) folios útiles, quedan legalmente protocolizados ante el Notario Treinta y Ocho (38) del Círculo Bogotá, de manera que desde ahora quedan bajo su custodia y guarda para que surtan los efectos legales a que haya lugar y los interesados puedan consultarlos y obtener las copias que de el solicite. ----------

IMPRESO EN JUNIO DE 2011 POR POLYPRINT EDITORIAL LTDA. - NIT 830 029 959-5

--------------------------------**HASTA AQUÍ LA MINUTA**---------------------------------------

SE ADVIRTIÓ al (a los) otorgante (s) de esta escritura de la **obligación que tiene (n) de leer la totalidad de su texto,** a fin de verificar la exactitud de todos los datos en ella consignados con el fin de aclarar, modificar o corregir lo que le(s) pareciere; la firma de la misma demuestra su aprobación total del texto.--------------

Se advierte igualmente la necesidad de diligenciar los espacios en blanco correspondientes a la información personal y de trabajo consignados en el espacio destinado para la firma de los suscriptores del instrumento publico, con el objeto de confrontar la información solicitada con el contenido de la escritura previo a la autorización de la misma. --

En consecuencia, la notaria no asume ninguna responsabilidad por error o inexactitudes establecidas con posterioridad a la firma del (los) otorgante (s) y del notario. ---

En tal caso, de la existencia de estos, deben ser corregidos mediante el otorgamiento de una nueva escritura, suscrita por el (los) que intervino (eron) en la inicial y sufragada por el (ellos) mismo (s). (Articulo 35 decreto ley 960 de 1.970).---

ADVERTENCIA PARA TODOS LOS COMPARECIENTES QUE EXTIENDAN Y OTORGUEN INSTRUMENTOS, SEGÚN ART. 34 C.N., LEY 190 DE 1995, LEY 333 DE 1996 Y LEY 365 DE 1997. Los comparecientes bajo la gravedad del juramento manifiestan clara y expresamente que todos los dineros, bienes muebles e inmuebles contenidos en este instrumento fueron adquiridos por medios y actividades licitas. Así como el dinero con el que se efectúa el pago de los gastos notariales.--

OTORGAMIENTO Y AUTORIZACIÓN. Leído el presente instrumento por el otorgante y advertido de la formalidad de Ley, estuvo de acuerdo con él lo aceptó en la forma como está redactado y en testimonio de que le da su asentimiento y aprobación lo firma por ante mí y conmigo el Notario que de todo lo expuesto doy fe lo autorizo.--

CONSTANCIA: El presente instrumento se extiende en la siguiente hoja de papel Notarial No. **

7 700133 020793/ 020762/ -

- -

Este documento público es una traducción al español de un documento en idioma inglés presentado ante mí.
ROBERTO M. OLARTE GARCÍA
Traductor e Intérprete Oficial
Ministerio de Justicia
Licencia 0041/1996

ESTADO DE CALIFORNIA

SECRETARIO DE ESTADO

La presente Apostilla solamente certifica la autenticidad de la firma y del cargo del funcionario público quien suscribió este documento público, y, donde fuera el caso, la identidad del sello que llevara el documento público.
La presente Apostilla no certifica la autenticidad del documento anexo para lo cual fue otorgada.
La presente Apostilla no es válida en ningún sitio dentro de los Estados Unidos de América, sus territorios o posesiones.
Para verificar la veracidad de esta Apostilla, diríjase a www.sos.ca.gov/business/notary/apostille-search/

APOSTILLA

(Convención de la Haya del 5 de octubre de 1961)

1. País: Estados Unidos de América
2. Este documento público ha sido firmado por **Pamela M. Stone**
3. actuando en calidad de **Notario Público, Estado de California**
4. lleva el sello del **Pamela M. Stone, Notario Público, Estado de California**

   CERTIFICA
5. en Sacramento, California
6. el **5 de julio de 2011**
7. por Secretario de Estado encargado, Estado de California
8. **No. 30327**
9. Sello: (ilegible)
10. Firma: (ilegible)
    Secretario de Estado
    POR (firma ilegible)

MINISTERIO DE RELACIONES EXTERIORES
JEFE DE LEGALIZACIONES
BOGOTA D.C.
2011 JUL -8 A 8:01

NO SE ASUME RESPONSABILIDAD DEL TEXTO

CUYA FIRMA APARECE EN EL PRESENTE DOCUMENTO DESCARGADA

Poder General

Firma: (ilegible)

Nombre: Timothy L. Smith, J.D., Ph.D.
Firma autorizada y con poder para IRM LLC

Fecha: 15 de junio de 2011

Ciudad, País: San Diego, California, Estados Unidos de América

RECONOCIMIENTO TODO-PROPÓSITO – CALIFORNIA

Estado de California
Condado de San Diego

El 15 de junio de 2011 ante mi, Pamela M. Stone, Notario Público, compareció Timothy L. Smith, quien con base en evidencia satisfactoria demostró ser la persona quien suscribió el anterior documento y reconoció que él firmó el mismo debidamente autorizado, y que mediante su firma sobre el instrumento, la persona o sociedad en nombre de quien actuó, suscribió dicho instrumento.

Certifico bajo sanción por perjurio según las leyes del Estado de California, que el anterior párrafo es verdadero y correcto.

En testimonio de lo cual, me suscribo y coloco el sello oficial.

Firma: (ilegible)

Sello:

PAMELA M. STONE
Comisión No. 1870231
Notario Público – California
Condado de San Diego
Mi comisión vence el 2 de noviembre de 2013

OPCIONAL

Aunque la información a continuación no es legalmente requerida, podrá ser útil para aquellas personas que dependan del documento y podría prevenir la remoción o re-adición fraudulenta de este formulario a otro documento.

Descripción del documento adjunto
Título o Tipo de Documento: Poder General – OlarteMoure y Asociados, Ltda.
Fecha del documento: 15 de junio de 2011
Número de páginas: 2

ESTADO DE CALIFORNIA

SECRETARIO DE ESTADO

La presente Apostilla solamente certifica la autenticidad de la firma y del cargo del funcionario público quien suscribió este documento público, y, donde fuera el caso, la identidad del sello que llevara el documento público.
La presente Apostilla no certifica la autenticidad del documento anexo para lo cual fue otorgada.
La presente Apostilla no es válida en ningún sitio dentro de los Estados Unidos de América, sus territorios o posesiones.
Para verificar la veracidad de esta Apostilla, diríjase a www.sos.ca.gov/business/notary/apostille-search/

APOSTILLA

(Convención de la Haya del 5 de octubre de 1961)

1. País: Estados Unidos de América
2. Este documento público ha sido firmado por **Pamela M. Stone**
3. actuando en calidad de **Notario Público, Estado de California**
4. lleva el sello del **Pamela M. Stone, Notario Público, Estado de California**

   CERTIFICA
5. en Sacramento, California
6. el **5 de julio de 2011**
7. por Secretario de Estado encargado, Estado de California
8. **No. 30123**
9. Sello: (ilegible)
10. Firma: (ilegible)
    Secretario de Estado
    POR (firma ilegible)

6024

OLARTEMOURE

OLARTE MOURE & ASOCIADOS LTDA.

General Power of Attorney

The undersigned,

IRM LLC

domiciled in

131 Front Street, P.O. Box HM 2899, Hamilton HM LX, Bermuda.

hereby appoint to **Carlos R. Olarte** as legal representative with authority to appoint judicial or extrajudicial agents for matters related to **Patents for Inventions and Industrial Designs** in the Republics of Argentina, Belize, Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Nicaragua, Panama, Paraguay, Peru, Oriental Republic of Uruguay and the Bolivarian Republic of Venezuela, in order to initiate before any authorities all the necessary actions to accomplish these purposes and to intervene in all instances, steps and procedural levels; to file and prosecute applications, receive notifications, declarations, complaints, to claim, counterclaim, answer claims and counterclaims; to make amendments; oppositions and all types of objections and writings of any nature; cancellations and administrative injunction proceedings; to pay all installments and effect all other payments prescribed by law; to receive documents and monies; act as complainants as well as defendants, before the competent judges, with full powers to request preliminary injunctions of any nature, settle, make arrangements, to submit to arbitrators, to desist from claims or procedures, to accept claims, to collect, to ratify legal proceedings subscribed before the appointment and/or presentation of this Power, substitute or delegate the legal representation and other acts prescribed by law, in short, with all such further powers as may be required; to comply with all other requirements and, to take all those measures which he may deem fit for the protection of our interests.

Signature/Firma: T. L. Smith

Name/Nombre: Timothy L. Smith, J.D., Ph.D.
Authorized Signatory and
Power of Attorney for IRM LLC

Date/Fecha: 15 June 2011

City,Country/
Ciudad,País: San Diego, California, US

Poder General

Los suscritos,

IRM LLC

domiciliados en

131 Front Street, P.O. Box HM 2899, Hamilton HM LX, Bermuda.

nombramos por el presente a **Carlos R. Olarte** como nuestro representante legal con facultades para nombrar apoderados judiciales o extrajudiciales para asuntos de **Patentes de Invención y Diseños Industriales** en las Repúblicas de Argentina, Belice, Bolivia, Chile, Colombia, Costa Rica, Cuba, República Dominicana, Ecuador, El Salvador, Guatemala, Honduras, Nicaragua, Panamá, Paraguay, Perú, República Oriental del Uruguay y la República Bolivariana de Venezuela, para iniciar ante cualquier autoridad todas las acciones necesarias para cumplir con el objeto indicado e intervenir en todas las instancias, grados o etapas procesales; elevar y tramitar solicitudes, recibir notificaciones, declaraciones y reclamos, para demandar, reconvenir, contestar demandas y reconvenciones; formular enmiendas, oposiciones y todo tipo de impugnaciones y escritos de cualquier naturaleza; instaurar cancelaciones y amparos administrativos; abonar todos los impuestos y cuotas y efectuar cualquier otro pago determinado por la ley; recibir todos los documentos y valores; intervenir como demandantes o demandados ante los jueces que sean competentes, pudiendo solicitar medidas cautelares de cualquier naturaleza, conciliar, transar, someter a árbitros, desistirse de la pretensión y/o el procedimiento, allanarse a la pretensión, percibir, ratificar los actos jurídicos procesales suscritos con anterioridad al otorgamiento y/o presentación del presente poder, sustituir o delegar la representación procesal y los demás actos que exprese la ley, con todas las demás facultades que resulten necesarias; llenar cualesquiera otros requisitos y tomar, en fin, todas las medidas que crea conducentes al resguardo de nuestros intereses.

Signature/Firma:

Name/Nombre:

6024

Date/Fecha:

City,Country/
Ciudad,País:

Diagonal 34 No. 5 - 28 Bogotá D.C., Colombia.
Tel: + 57 (1) 601-7700 Fax: + 57 (1) 601-7799 e-mail: office@olartemoure.com
www.olartemoure.com



Certificación Notarial

En esta fecha
15 de junio de 2011

el Notario Público suscrito certifica:

1. Que
Timothy L. Smith

mayor de edad y domiciliado en
10675 John Jay Hopkins Drive
San Diego, California 92121, Estados Unidos de América

suscribió el documento que antecede.

2. Que el otorgante ha suscrito el referido documento en su carácter de
Firma Autorizada y con poder para IRM LLC

3. Que IRM LLC,
con sede principal ubicada en
131 Front Street, P.O. Box HM 2899,
Hamilton HM LX, Bermuda

está legalmente constituida, que tiene existencia legal vigente y que los propósitos para los cuales se otorga el poder se encuentra dentro de los que comprenden su objeto social.

4. Que el otorgante tiene la representación para suscribir el poder, y que está es legítima.

5. El fundamento de mis certificaciones contenidas en los puntos 2, 3 y 4 fue la exhibición de los siguientes documentos auténticos:

Poder
de IRM LLC a Timothy Smith
suscrito el 10 de noviembre de 2010

Certificado de constitución de IRM LLC
suscrito el 2 de junio de 1998

NOTARIO PÚBLICO (firma): ilegible
Pamela M. Stone, Notario Público

SELLO (opcional):

PAMELA M. STONE
Comisión No. 1870231
Notario Público – California
Condado de San Diego

ROBERTO OLARTE
Traductor e Intérprete Oficial
Resolución No. 0041
Minijusticia 1996

State of California

SECRETARY OF STATE

This Apostille only certifies the authenticity of the signature and the title of the public official who has signed the public document, and, where appropriate, the identity of the seal or stamp which the public document bears. This Apostille does not certify the authenticity of the underlying document for which it was issued. This Apostille is not valid for use anywhere within the United States of America, its territories or possessions. This Apostille is not valid for use anywhere within the United States of America, its territories or possessions. To verify the issuance of this Apostille, go to www.sos.ca.gov/business/notary/apostille-search/

APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. *Country: United States of America*
   This public document
2. *has been signed by Pamela M. Stone*
3. *acting in the capacity of Notary Public, State of California*
4. *bears the seal/stamp of Pamela M. Stone, Notary Public, State of California*

CERTIFIED

5. *At Sacramento, California*
6. *the 5th day of July 2011*
7. *by Deputy Secretary of State, State of California*
8. *No. 30327*
9. *Seal/Stamp:*

10. *Signature*

Debra Bowen
Secretary of State

BY

OSP 07 104060

NP-24 A (REV 1-07)

CALIFORNIA ALL-PURPOSE ACKNOWLEDGMENT

State of California

County of San Diego

On June 15, 2011 (Date) before me, Pamela M. Stone, Notary Public (Here Insert Name and Title of the Officer),

personally appeared Timothy K. Smith (Name(s) of Signer(s)),

who proved to me on the basis of satisfactory evidence to be the person(s) whose name(s) is/are subscribed to the within instrument and acknowledged to me that he/she/they executed the same in his/her/their authorized capacity(ies), and that by his/her/their signature(s) on the instrument, the person(s), or the entity upon behalf of which the person(s) acted, executed the instrument.

I certify under PENALTY OF PERJURY under the laws of the State of California that the foregoing paragraph is true and correct.

WITNESS my hand and official seal.

PAMELA M. STONE
Commission # 1870231
Notary Public - California
San Diego County
My Comm. Expires Nov 2, 2013

Signature: Pamela M. Stone
Signature of Notary Public

Place Notary Seal and/or Stamp Above

OPTIONAL

Though the information below is not required by law, it may prove valuable to persons relying on the document and could prevent fraudulent removal and reattachment of this form to another document.

Description of Attached Document

Title or Type of Document: General Power of Attorney - Olarte Moure + Asociados Ltda.

Document Date: June 15, 2011 Number of Pages: 2

Signer(s) Other Than Named Above: _____

Capacity(ies) Claimed by Signer(s)

Signer's Name: _____

☐ Corporate Officer — Title(s): _____

☐ Individual

☐ Partner — ☐ Limited ☐ General

☐ Attorney in Fact

☐ Trustee

☐ Guardian or Conservator

☐ Other: _____

Signer Is Representing: _____

RIGHT THUMBPRINT OF SIGNER
Top of thumb here

Signer's Name: _____

☐ Corporate Officer — Title(s): _____

☐ Individual

☐ Partner — ☐ Limited ☐ General

☐ Attorney in Fact

☐ Trustee

☐ Guardian or Conservator

☐ Other: _____

Signer Is Representing: _____

RIGHT THUMBPRINT OF SIGNER
Top of thumb here

© 2008 National Notary Association • 9350 De Soto Ave., P O Box 2402 • Chatsworth, CA 91313-2402 • www NationalNotary org Item #5907 Reorder Call Toll-Free 1-800-876-6827

State of California

SECRETARY OF STATE

This Apostille only certifies the authenticity of the signature and the title of the public official who has signed the public document, and, where appropriate, the identity of the seal or stamp which the public document bears. This Apostille does not certify the authenticity of the underlying document for which it was issued. This Apostille is not valid for use anywhere within the United States of America, its territories or possessions. To verify the issuance of this Apostille, go to www.sos.ca.gov/business/notary/apostille-search/

APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. *Country: United States of America*
   This public document
2. *has been signed by Pamela M. Stone*
3. *acting in the capacity of Notary Public, State of California*
4. *bears the seal/stamp of Pamela M. Stone, Notary Public, State of California*
   CERTIFIED
5. *At Sacramento, California*
6. *the 5th day of July 2011*
7. *by Deputy Secretary of State, State of California*
8. *No. 30123*
9. *Seal/Stamp:*

10. *Signature*

Debra Bowen
Secretary of State

BY _____

NP-24 A (REV 1-07)

OSP 07 104060

OLARTEMOURE
OLARTE MOURE & ASOCIADOS LTDA.

Notarial Certification

On this date,
15 June 2011
The undersigned Notary Public certifies:

1. That
Timothy L. Smith
of legal age and domiciled in
10675 John Jay Hopkins Drive
San Diego, California 92121, US
signed the foregoing document.

2. That the grantor has executed the foregoing document as
Authorized Signatory and Power of Attorney
for **IRM LLC**

3. That **IRM LLC,**
having its principal place of business located in
131 Front Street, P.O. Box HM 2899,
Hamilton HM LX, Bermuda
is duly incorporated, is currently in existence, and that the purposes for which the General Power is granted are within the scope of its corporate purposes.

4. That the grantor has the authority to execute the General Power and that his/her representation is legal.

5. The foundation for my certifications contained in points 2, 3 and 4 was the exhibition of the following authentic document(s):

Power of Attorney
of IRM LLC to Timothy Smith
executed on 10 November 2010

Certificate of Formation of IRM LLC
executed on 02 June 1998

Certificación Notarial

En esta fecha,

el Notario Público suscrito certifica:

1. Que
Timothy L. Smith
mayor de edad y domiciliado
10675 John Jay Hopkins Drive
San Diego, California 92121, US
suscribió el documento que antecede.

2. Que el otorgante ha suscrito el referido documento en su carácter de

de

4. Que
con sede principal ubicada en

está legalmente constituida, que tiene existencia legal vigente y que los propósitos para los cuales se otorga el poder se encuentra dentro de los que comprenden su objeto social

4. Que el otorgante tiene la representación para suscribir el poder, y que ésta es legitima.

5. El fundamento de mis certificaciones contenidas en los puntos 2, 3 y 4 fue la exhibición de los (del) siguiente(s) documento(s) auténtico(s):

NOTARY PUBLIC (Signature): Pamela M. Stone

Pamela M. Stone, Notary Public

SEAL (optional)

6024

PAMELA M. STONE
Commission # 1870231
Notary Public - California
San Diego County
My Comm. Expires Nov 2, 2013

(See attached All-Purpose Acknowledgement)

6024

FOLIO ANTERIOR

7 700133 020793--- - - - - - - - - - - - - - - - - -

ESCRITURA PÚBLICA NÚMERO:

SEIS MIL VEINTICUATRO (6024)- - - - - - - - - - -

FECHA DE OTORGAMIENTO:

TRECE (13) DE JULIO DE DOS MIL ONCE (2011)--- - -

NOTARIA TREINTA Y OCHO (38) DEL CÍRCULO DE BOGOTA D.C. ----------

RETENCION EN LA FUENTE. ARTICULO 20 Y 64 LEY 0075 DE 1986.....$ EXENTO -- - -

DERECHOS NOTARIALES.....$ 44.000 - - - - - - - - - - - - - - - - -

SUPERINTENDENCIA.....$3.700.oo.----------

FONDO NACIONAL DEL NOTARIADO.....$3.700.oo.----------

IVA .....$ 14.400- - - - - - - - - - - - - - - - -

DECRETO 1681 DE SEPTIEMBRE 16 DE 1996 Y RESOLUCION 11621 DE DICIEMBRE 22 DE 2010 Y MODIFICADA POR LA RESOLUCION 11903 DE DICIEMBRE 30 DE 2010. ---------- --------

CARLOS REINALDO OLARTE

C.C.

DIRECCIÓN RESIDENCIA:

DIRECCIÓN OFICINA:

TELEFONO RESIDENCIA:

TELEFONO OFICINA:

CELULAR:

ESTADO CIVIL:

EL NOTARIO TREINTA Y OCHO (38)

DEL CIRCULO DE BOGOTA, D.C.

EDUARDO DURAN GOMEZ

REPUBLICA DE COLOMBIA
EDUARDO DURAN GOMEZ
NOTARIO
NOTARIA 38
NOTARIA TREINTA Y OCHO DE BOGOTA D.C.

JOHANNA V.

Z. HDEZ

201106136-PROTOCOLIZACIONES

PAPEL DE USO EXCLUSIVO DEL PROTOCOLO NOTARIAL NO TIENE COSTO PARA EL USUARIO

IMPRESO EN JUNIO DE 2011 POR POLYPRINT EDITORIAL LTDA - NIT 830.026.959-5

OCR TAB

ES LA COPIA **(FOTOCOPIA)** NÙMERO =4= DE LA ESCRITURA PÚBLICA =6024= FECHA =13= DE =JULIO= DEL AÑO =2011= TOMADA DE SU ORIGINAL QUE SE EXPIDE CONFORME AL ARTÍCULO DEL DECRETO 2148 DE 1983, EN =7= HOJAS ÚTILES CON DESTINO AL *= INTERESADO=*

DADO EN BOGOTA, D.C. 02 DE SEPTIEMBRE DE 2011

MONICA ZULMA BAUTISTA NAVARRO

NOTARIA (38E) DEL CIRCULO DE BOGOTA

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 April 2013 (04.04.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/049234 A2

(51) International Patent Classification:
C07K 14/435 (2006.01)

(21) International Application Number:
PCT/US2012/057371

(22) International Filing Date:
26 September 2012 (26.09.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/539,290 26 September 2011 (26.09.2011) US

(71) Applicants *(for all designated States except US)*: **NOVARTIS AG** [CH/CH]; Lichtstrasse 35, CH-4056 Basel (CH). **IRM LLC** [—/US]; a Delaware Limited, Liability Company, 131 Front Street, PO Box HM 2899, Bermuda, HM LX (BM).

(72) Inventors; and

(71) Applicants *(for US only)*: **BOETTCHER, Brian R.** [US/US]; Novartis Institutes for BioMedical, Research, Inc., 100 Technology Square, Cambridge, Massachusetts 02139 (US). **CAPLAN, Shari L.** [US/US]; Novartis Institutes for BioMedical Research, Inc., 100 Technology Square, Cambridge, Massachusetts 02139 (US). **CELLITTI, Susan E.** [US/US]; Novartis Institute for Functional Genomics, Inc., 10675 John Jay Hopkins Drive, San Diego, California 92121 (US). **DANIELS, Douglas S.** [US/US]; Novartis Institutes for BioMedical Research, Inc., 100 Technology Square, Cambridge, Massachusetts 02139 (US). **HAMAMATSU, Norio** [JP/US]; Novartis Institutes for BioMedical Research, Inc, 100 Technology Square, Cambridge, Massachusetts 02139 (US). **GEIER-STANGER, Bernhard Hubert** [DE/US]; Novartis Institute for Functional Genomics, Inc., 10675 John Jay Hopkins Drive, San Diego, California 92121 (US). **LICHT, Stuart** [US/US]; 100 Memorial Drive, Cambridge, Massachusetts 02142 (US). **LOEW, Andreas** [DE/US]; Novartis Institutes for BioMedical, Research, Inc., 250 Massachusetts Avenue, Cambridge, Massachusetts 02139 (US). **WELDON, Stephen Craig** [US/US]; Novartis Institutes for BioMedical Research, Inc., 100 Technology Square, Cambridge, Massachusetts 02139 (US).

(74) Agent: **PAGLIERANI, Paul J.**; Novartis Institutes for BioMedical, Research, Inc., 220 Massachusetts Avenue, Cambridge, Massachusetts 02139 (US).

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

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

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2013/049234 A2

(54) Title: DUAL FUNCTION PROTEINS FOR TREATING METABOLIC DISORDERS

(57) Abstract: The present invention relates to the identification of new proteins comprising fibroblast growth factor 21 (FGF21) and other metabolic regulators, including variants thereof, known to improve metabolic profiles in subjects to whom they are administered. Also disclosed are methods for treating FGF21-associated disorders, GLP-1-associated disorders, and Exendin-4-associated disorders, including metabolic conditions.

DUAL FUNCTION PROTEINS FOR TREATING METABOLIC DISORDERS

FIELD OF THE INVENTION

[0001] The present invention relates to new proteins comprising fibroblast growth factor 21 (FGF21) and other metabolic regulators known to improve metabolic profiles in subjects to whom they are administered.

BACKGROUND OF THE INVENTION

[0002] The fibroblast growth factor (FGF) family is characterized by 22 genetically distinct, homologous ligands, which are grouped into seven subfamilies. FGF-21 is most closely related to, and forms a subfamily with, FGF-19 and FGF-23. This FGF subfamily regulates diverse physiological processes uncommon to classical FGFs, namely energy and bile acid homeostasis, glucose and lipid metabolism, and phosphate as well as vitamin D homeostasis. Moreover, unlike other FGFs, this subfamily acts in an endocrine fashion (Moore, D.D. (2007) Science 316, 1436-8)(Beenken et al. (2009) Nature Reviews Drug Discovery 8, 235).

[0003] FGF21 is a 209 amino acid polypeptide containing a 28 amino acid leader sequence (SEQ ID NO:132). Human FGF21 has about 79% amino acid identity to mouse FGF21 and about 80% amino acid identity to rat FGF21. Fibroblast growth factor 21 (FGF21) has been described as a treatment for ischemic vascular disease, wound healing, and diseases associated with loss of pulmonary, bronchia or alveolar cell function (Nishimura et al. (2000) Biochimica et Biophysica Acta, 1492:203-206; patent publication WO01/36640; and patent publication WO01/18172). Although FGF-21 activates FGF receptors and downstream signaling molecules, including FRS2a and ERK, direct interaction of FGFRs and FGF-21 has not been detected. Studies have identified β-klotho, which is highly expressed in liver, adipocytes and pancreas, as a determinant of the cellular response to FGF-21 and a cofactor which mediates FGF-21 signaling through FGFRs (Kurosu, H. et al. (2007) J Biol Chem 282, 26687-95). FGF21 is a potent agonist of the FGFR1(IIIc), FGFR2(IIIc) and FGFR3(IIIc) β-klotho signaling complexes.

[0004] FGF-21 has been shown to induce insulin-independent glucose uptake. FGF-21 has also been shown to ameliorate hyperglycemia in a range of diabetic rodent models. In addition, transgenic mice over-expressing FGF-21 were found to be resistant to diet-induced metabolic abnormalities, and demonstrated decreased body weight and fat mass, and enhancements in insulin sensitivity (Badman, M.K. et al. (2007) Cell Metab 5, 426-37). Administration of FGF-21 to diabetic non-human primates

caused a decline in fasting plasma glucose, triglycerides, insulin and glucagon levels, and led to significant improvements in lipoprotein profiles, including a nearly 80% increase in HDL cholesterol (Kharitonenkov, A. et al. (2007) Endocrinology 148, 774-81). Recent studies investigating the molecular mechanisms of FGF21 action have identified FGF21 as an important endocrine hormone that helps to control adaptation to the fasting state (Badman et al. (2009) Endocrinology 150, 4931)(Inagaki et al. (2007) Cell Metabolism 5, 415). This provides a previously missing link downstream of PPARα, by which the liver communicates with the rest of the body in regulating the biology of energy homeostasis (Galman et al. (2008) Cell Metabolism 8, 169)(Lundasen et al. (2007) Biochemical and Biophysical Research Communications 360, 437).

[0005] FGF21 regulates adipocyte homeostasis through activation of an AMPK/SIRT1/PGC1α pathway to inhibit PPARγ expression and increase mitochondrial function (Chau et al. (2010) PNAS 107, 12553). FGF21 also increases glucose uptake by skeletal muscle as measured in cultured human myotubes and isolated mouse tissue (Mashili et al. (2011) Diabetes Metab Res Rev 27, 286-97). FGF21 treatment of rodent islet cells leads to improved function and survival through activation of ERK1/2 and Akt pathways (Wente et al. (2006) Diabetes 55, 2470). FGF21 treatment also results in altered gene expression for lipogenesis and fatty acid oxidation enzymes in rodent livers, likely through HNF4α and Foxa2 signaling. However, recent studies (Wei et al. (2012) PNAS 109, 3143-48) indicate that treatment of diet-induced obese mice with FGF21 induces bone loss, due to a diminished inactivation of PPARγ (via reduced sumoylation); a shift of mesenchymal stem cell differentiation from osteoblasts to adipocytes is seen in the presence of increased PPARγ activity in the bone following FGF21 treatment.

[0006] A difficulty associated with using FGF-21 directly as a biotherapeutic is that its half-life is very short (Kharitonenkov, A. et al. (2005) Journal of Clinical Investigation 115:1627-1635). In mice, the half-life of human FGF21 is 0.5 to 1 hours, and in cynomolgus monkeys, the half-life is 2 to 3 hours. FGF21 may be utilized as a multi-use, sterile pharmaceutical formulation. However, it has been determined that preservatives, e.g., m-cresol, have an adverse effect on its stability under these conditions.

[0007] Another potent metabolic regulator already represented in the clinic is Glucagon-Like Peptide-1 (GLP-1) (Knudsen et al. (2004) Journal of Medicinal Chemistry 47, 4128). GLP-1 is a 36 amino acid incretin secreted by L-cells of the mammalian gut, acting on both alpha and beta cells to stimulate insulin secretion and inhibit glucagon release in a glucose-dependent manner (Hare et al. (2010) Diabetes 59, 1765; Meier et

al. (2005) Diabetes-Metabolism Research and Reviews 21, 91). GLP-1 binds to and activates the GLP-1 receptor (GLP-1R), a seven-transmembrane helix protein of the class II family of G-protein coupled receptors (GPCRs) (Mayo et al. (2003) Pharmacological Reviews 55:167). As a GLP-1 receptor agonist, GLP-1 has an important role in decreasing post-prandial blood glucose levels by stimulating insulin secretion from the pancreas in order to increase glucose absorption in the peripheral tissues and inhibiting glucagon secretion, resulting in reduced hepatic glucose release.

[0008] A second clinically important GLP-1 receptor agonist is Exendin-4. Exendin-4 is a 39 residue polypeptide produced in the salivary glands of the Gila Monster lizard (Goke et al. (1993) Diabetes 46:433-439; Fehmann et al. (1995) Endocrine Rev. 16:390-410). Although it is the product of a uniquely non-mammalian gene and appears to be expressed only in the salivary gland, Exendin-4 shares a 52% amino acid sequence homology with GLP-1, and in mammals interacts with the GLP-1 receptor (Goke, et al.; Thorens et al. (1993) Diabetes 42:1678-1682). In vitro, Exendin-4 has been shown to promote insulin secretion by insulin producing cells and, given in equimolar quantities, is more potent than GLP-1 at causing insulin release from insulin producing cells. Furthermore, Exendin-4 potently stimulates insulin release to reduce plasma glucose levels in both rodents and humans and is longer acting than GLP-1; however, because it does not occur naturally in mammals, Exendin-4 has certain potential antigenic properties in mammals that GLP-1 lacks.

[0009] The ability of GLP-1 and Exendin-4 analogues (e.g., Liraglutide and Byetta) to improve glucose control in humans is established in the clinic (Idris (2010) Diabetes Obesity & Metabolism 12, 89; Monami et al (2009) European Journal of Endocrinology 160, 909). GLP-1 has also been reported to increase beta cell mass both through induced proliferation and inhibition of apoptosis (Egan, A et al (2003) Diabetes-Metabolism Research and Reviews 19, 115;Farilla, L. et al. (2003) Endocrinology 144, 5149;Xu, G. et al. (1999) Diabetes 48, 2270). It also acts as an intestinal hormone to inhibit acid secretion and gastric emptying in the stomach while providing a satiety signal that decreases appetite (Vilsboll et al. (2009) Best Practice & Research Clinical Endocrinology & Metabolism 23, 453). These effects likely account for beneficial weight loss observed with administration of GLP-1 analogues to type 2 diabetes patients. GLP-1 has also been shown to be cardioprotective in postischemic rodent hearts (Ossum et al. (2009) Pharmacological Research 60, 411; Sonne, D.P. et al. (2008) Regulatory Peptides 146, 243;Nikolaidis, L. A. et al. (2004) Circulation 109, 962).

[00010] Additionally, GLP-1 can reduce the differentiation of human mesenchymal stem cells (hMSCs) to adipocytes by reducing the expression of PPARγ, and GLP-1

promotes cellular proliferation and cytoprotection of hMSCs (Sanz et al. (2010) Am J Physiol Endocrinol Metab 298, E634-E643).

[00011] In developing an FGF21 protein, including a variant or analogue thereof, for use as a therapeutic in the treatment of type 1 and type 2 diabetes mellitus and other metabolic conditions, an increase in half-life and stability would be desirable. FGF21 proteins having enhanced half-life and stability would allow for less frequent dosing of patients being administered the protein. Clearly, there is a need to develop a stable aqueous protein formulation for the therapeutic protein FGF21.

[00012] Furthermore, a significant challenge in the development of protein pharmaceuticals, such as metabolic regulators FGF21, GLP-1, and Exendin-4, is to cope with their physical and chemical instabilities. The compositional variety and characteristics of proteins define specific behaviors such as folding, conformational stability, and unfolding/denaturation. Such characteristics should be addressed when aiming to stabilize proteins in the course of developing pharmaceutical formulation conditions utilizing aqueous protein solutions (Wang, W., Int. J. of Pharmaceutics, 18, (1999)). A desired effect of stabilizing therapeutic proteins of interest, e.g., the proteins of the present invention, is increasing resistance to proteolysis and enzymatic degradation, thereby improving protein stability and reducing protein aggregation.

SUMMARY OF THE INVENTION

[00013] The invention relates to the identification of new proteins, e.g., fusion proteins, which comprise fibroblast growth factor 21 (FGF21) and other metabolic regulators, e.g., GLP-1 and Exendin-4, and which have improved pharmaceutical properties over the constituent agents under pharmaceutical formulation conditions, e.g., are more stable, possess the ability to improve metabolic parameters for subjects to whom they are administered, are less susceptible to proteolysis and enzymatic degradation, are less likely to aggregate and form complexes and are less likely to be immunogenic. The proteins of the invention possess both FGF21 receptor agonist and GLP-1 receptor agonist activity; they comprise truncations and variants of FGF21, and further comprise one or more of, e.g., glucagon-like peptide-1 (GLP-1), Exendin-4, or other metabolic regulators or variants thereof.

[00014] Also disclosed are methods for treating FGF21-associated and GLP-1 associated disorders, as well as other metabolic, endocrine, and cardiovascular disorders, such as obesity, type 1 and type 2 diabetes mellitus, pancreatitis, dyslipidemia, nonalcoholic fatty liver disease (NAFLD), nonalcoholic steatohepatitis (NASH), insulin resistance, hyperinsulinemia, glucose intolerance, hyperglycemia, metabolic syndrome, acute myocardial infarction, hypertension, cardiovascular disease,

atherosclerosis, peripheral arterial disease, stroke, heart failure, coronary heart disease, kidney disease, diabetic complications, neuropathy, gastroparesis, disorders associated with severe inactivating mutations in the insulin receptor, lipodystrophies including HIV-associated lypodistrophy and other metabolic disorders, and in reducing the mortality and morbidity of critically ill patients.

[00015] The proteins of the present invention may be used as a regularly administered (e.g., daily, more preferably weekly, biweekly, or monthly) injectable, either alone or in combination with oral anti-diabetic agents, which will improve the glycemic control, body weight and lipid profile of type 1 and type 2 diabetes mellitus patients. The proteins may also be used for the treatment of obesity or other FGF21- or GLP-1-associated conditions.

[00016] The proteins of the invention, e.g., GLP-1-FGF21 variant and Exendin-4-FGF21 variant fusion proteins of the invention, overcome the significant hurdles of physical instabilities associated with protein therapeutics, including, for instance, with the administration of the wild-type FGF21, as they are more stable, less susceptible to proteolysis and enzymatic degradation, less likely to aggregate and form complexes and less likely to be immunogenic than wild-type FGF21 under pharmaceutical formulation conditions.

[00017] In a first aspect, the invention provides Fibroblast Growth Factor 21 (FGF21) proteins, e.g., fusion proteins, comprising one or more of the sequences listed in Table 1, and further described herein. The proteins of the invention can further comprise GLP-1 and/or Exendin-4 proteins, whether wild-type, truncated, or mutated versions, or variants thereof. The FGF21 sequences listed in Table 1 are variants of the wild-type FGF21 sequence, e.g., the wild-type FGF21 sequence with NCBI reference number NP_061986.1, and found in such issued patents as, e.g., US 6,716,626B1, assigned to Chiron Corporation. The GLP-1 and Exendin-4 sequences listed in Table 1 are variants of the wild-type GLP-1 and Exendin-4 sequences, e.g., those sequences with NCBI reference numbers NP_002045 and AAB22006.1, respectively, and can be found in such patent publications as, e.g., WO98/19698 and WO87/06941A, assigned to Eli Lilly and Co. and the General Hospital Corp., respectively (GLP-1) and US 5,424,286, assigned to Amylin (Exendin-4).

[00018] Other embodiments are drawn to polynucleotides encoding the dual function proteins of the invention, a vector containing said polynucleotides and a host cell carrying said vector.

[00019] Provided herein are methods used to generate the proteins of the invention, wherein such methods involve modification of the wild-type FGF21 protein, via, e.g., the site-specific incorporation of amino acids at positions of interest within the wild-type

FGF21 protein, as well as the fusion between the FGF21 portion of the molecule to other metabolic regulators, such as glucagon-like peptide-1 (GLP-1) and Exendin-4, or conjugates with polymers modified with GLP-1 and/or Exendin-4. Said modifications and fusions enhance the biological properties of the proteins of the invention relative to the wild-type versions of the proteins (e.g., FGF21, GLP-1, and Exendin-4), as well as, in some cases, serving as points of attachment for, e.g., labels and protein half-life extension agents, and for purposes of affixing said variants to the surface of a solid support. Related embodiments of the invention are methods to produce cells capable of producing said proteins of the invention, and of producing vectors containing DNA encoding said variants and fusions.

[00020] In various embodiments, the proteins of the invention can comprise one or more fragments of the FGF21, Exendin-4, and/or GLP-1 sequences, including fragments as small as 8-12 amino acid residues in length, wherein the polypeptide is capable of lowering blood glucose in a mammal. In various embodiments, the proteins of the invention can comprise one or more variants of the FGF21, Exendin-4, and/or GLP-1 sequences, e.g., with one or more amino acid deletions, insertions, additions, or substitutions relative to the wild-type sequences thereof.

[00021] In some embodiments, the proteins of the invention can be covalently linked to one or more polymers, such as polyethylene glycol (PEG) or polysialic acid, whether at the position of site-specific amino acid modifications made relative to the wild-type FGF21, GLP-1, or Exendin-4, or at the position of amino acids commonly shared with the wild-type versions of those proteins. The PEG group is attached in such a way so as enhance, and/or not to interfere with, the biological function of the constituent portions of the fusion proteins of the invention, e.g., the GLP-1 protein variants or FGF21 protein variants. In other embodiments, the polypeptides of the invention can be fused to a heterologous amino acid sequence, optionally via a linker, such as GS, GGGGSGGGGSGGGGS (SEQ ID NO:8), or SGGGGSGGG (SEQ ID NO:128). The heterologous amino acid sequence can be an IgG constant domain or fragment thereof (e.g., the Fc region), Human Serum Albumin (HSA), or albumin-binding polypeptides. Such fusion proteins disclosed herein can also form multimers.

[00022] In some embodiments, a heterologous amino acid sequence (e.g., HSA, Fc, etc.) is fused to the amino-terminal of the proteins of the invention. In other embodiments, the fusion heterologous amino acid sequence (e.g., HSA, Fc, etc.) is fused to the carboxyl-terminal of the proteins of the invention. In still other, more preferred embodiments, the heterologous amino acid sequence (e.g., HSA, Fc, etc.) is situated in the middle of the dual function proteins of the invention, i.e., between the C-terminal residue of the GLP-1 or Exendin-4 sequence and the N-terminal residue of the

FGF21 sequence. Said preferred embodiment, e.g., leaves a free N-terminus for maximum GLP-1 (Exendin-4) activity and a free, intact C-terminus for maximum FGF21 activity.

[00023] In some embodiments, the GLP-1 receptor agonist is fused to the N-terminus of heavy and light chain of an antibody and FGF21 is simultaneously fused to the C-terminus of heavy and light chain of the same antibody (i.e., a fusobody, as described herein). Said preferred embodiment leaves a free N-terminus for maximum GLP-1 (Exendin-4) activity and a free, intact C-terminus for maximum FGF21 activity. A preferred embodiment uses the antibody sequence described in PCT publication WO2011/076781.

[00024] In some embodiments, the GLP-1 or Exendin-4 peptide is chemically attached to FGF21. In some embodiments, said peptides are attached to an FGF21 amino acid residue side chain. In other embodiments, said peptides are attached to the N-terminus of FGF21 through native chemical ligation or other methods known to the art. The preferred embodiment leaves a free N-terminus of GLP-1 (Exendin-4) for maximal activity and a free, intact C-terminus of FGF21 for maximal activity.

[00025] In some embodiments, the GLP-1 or Exendin-4 peptide is covalently attached to a polymer molecule that in turn is attached to the FGF21 protein variant. In a preferred embodiment, the GLP-1 or Exendin-4 peptide is attached to a PEG polymer that is simultaneously attached to a FGF21 protein variant. Said preferred embodiment leaves a free N-terminus for maximum GLP-1 (Exendin-4) activity and a free, intact C-terminus for maximum FGF21 activity. In some embodiments of the invention, a GLP-1 receptor agonist peptide is connected to a FGF21 variant through a PEG linker or other polymer linker that simultaneously provides half-life extension as well as a covalent connection for the two receptor agonists.

[00026] Yet another embodiment is drawn to methods of treating a patient exhibiting one or more FGF21-associated disorders or GLP-1-associated disorders, such as obesity, type 2 diabetes mellitus, type 1 diabetes mellitus, pancreatitis, dyslipidemia, nonalcoholic fatty liver disease (NAFLD), nonalcoholic steatohepatitis (NASH), insulin resistance, hyperinsulinemia, glucose intolerance, hyperglycemia, metabolic syndrome, acute myocardial infarction, hypertension, cardiovascular disease, atherosclerosis, peripheral arterial disease, stroke, heart failure, coronary heart disease, kidney disease, diabetic complications, neuropathy, gastroparesis, disorders associated with inactivating mutations in the insulin receptor, lipodystrophies including HIV-associated lipodystrophy and other metabolic disorders, comprising administering to said patient in need of such treatment a therapeutically effective amount of one or more proteins of the invention or a pharmaceutical composition thereof.

[00027] The invention also provides pharmaceutical compositions comprising the dual function proteins of the invention disclosed herein and a pharmaceutically acceptable formulation agent. Such pharmaceutical compositions can be used in a method for treating a metabolic disorder, and the method comprises administering to a human patient in need thereof a pharmaceutical composition of the invention. Non-limiting examples of metabolic disorders that can be treated include type 1 and type 2 diabetes mellitus and obesity.

[00028] These and other aspects of the invention will be elucidated in the following detailed description of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[00029] Figures 1A and 1B show the activity of GLP-1-FGF21-PEG fusion proteins with different N-terminal mutations, and Exendin-4-FGF21-PEG fusion proteins. Figure 1A shows the mutations added to slow processing by DPP-4 protease (GLP-1 peptide (circles) or dual function proteins with wild-type (V231; open squares), G0 (V251; open circles), A8G (V258; open triangles), A8S (V232; triangles), or E9P (V271; squares) GLP-1). Figure 1B shows Exendin-4 fusions (GLP-1 peptide (squares) or dual function proteins with Exendin-4 1-39 (V234; triangles), 1-30 (V267; circles) or 1-30/E16G/E17Q (V268; open triangles)).

[00030] Figure 2 shows pharmacokinetic properties (PK) of the GLP-1-FGF21-PEG fusion proteins (FGF21-PEG (V294; open squares), GLP-1(A8S)-PEG (V253; open circles) or dual function proteins with wild-type GLP-1 (V237; circles) or GLP-1(A8S) (V235; triangles) GLP-1). The PK of wild-type non-PEGylated FGF21 injected at 0.25 mg/kg is shown for comparison (squares).

[00031] Figures 3A-3C show results from using the oral glucose tolerance test (OGTT) to measure the efficacy of half-life extended GLP-1-FGF21-PEG fusion proteins. Eight week C57BL/6J mice (n=5) were dosed by intraperitoneal (i.p.) injection with 1 mg/kg compound or vehicle (solid white bars; triangles). Blood glucose was measured at 1 and 24 hours post dose to assess acute effects of the GLP-1, which were similar for the wild-type GLP-1 (V239; hashed bars; squares) and GLP-1(A8S) versions (V232; black bars; circles) at 1 hour and retained slightly better by the GLP-1(A8S) version at 24 hours. On the third night, the mice were fasted before challenge with 1.5 g/kg oral glucose at 72 hours post-dose. The mice dosed with the A8S version showed significantly improved control of blood glucose compared to those with the wild-type GLP-1 version, suggesting that the A8S mutation increased long-term levels of active GLP-1.

[00032] Figure 4 shows efficacy of GLP-1-FGF21-PEG fusion proteins in ob/ob mice. Ten week old male ob/ob mice (n=8) were dosed i.p. twice weekly with the indicated compounds for two weeks. Equivalent dosing of FGF21-PEG (V238; diagonal hashed bars; squares) and GLP-1-FGF21(R154C)-PEG (V239; horizontal hashed bars; open squares) showed very similar improvements in blood glucose levels, body weight, and liver health (as measured by serum ALT levels and liver weight) when compared to the vehicle-treated group (circles). The DPP-4-resistant fusion, 0.2 mg/kg GLP-1(A8S)-FGF21(R154C)-PEG (V235; vertical hashed bars; open triangles) showed similar efficacy to other compounds, but equal dosing at 1 mg/kg GLP-1(A8S)-FGF21(R154C)-PEG (V235; black bars; triangles) gave additional lowering of blood glucose, body weight, alanine aminotransferase (ALT), and liver weight.

[00033] Figures 5A-5G show a comparison of the GLP-1(A8S)-FGF21(R154C)-PEG (V235; black bars; open triangles) fusion protein to co-administration of FGF21(R154C)-PEG + GLP-1(A8S)-PEG (vertical hashed bars; open circles), single administration of FGF21(R154C)-PEG (V238; 0.2 mg/kg (light diagonal hashed bars); 1 mg/kg (dark diagonal hashed bars; squares)), and single administration of GLP-1(A8S)-PEG (V253; 0.2 mg/kg (light horizontal hashed bars); 1 mg/kg (dark horizontal hashed bars; open squares)) in ob/ob mice. Nine week old male ob/ob mice (n=8) were dosed i.p. twice weekly with the indicated compounds or vehicle (solid white bars; circles) for four weeks. The fusion showed significantly improved efficacy compared to co-administration of FGF21(R154C)-PEG + GLP-1(A8S)-PEG. Although only moderate weight loss was observed with FGF21(R154C)-PEG and GLP-1(A8S)-FGF21(R154C)-PEG, other groups gained a substantial amount of weight, only in part due to differences in food intake.

[00034] Figures 6A-6F show efficacy of GLP-1(A8S)-FGF21-PEG fusion proteins in ob/ob mice, comparing FGF21(R154C)-PEG (V235; dark bars) to FGF21(V76)-PEG (V272; light bars). Nine week old male ob/ob mice (n=8) were dosed i.p. twice weekly with the indicated compounds or vehicle (solid white bars) for two weeks. Both the GLP-1(A8S)-FGF21(R154C)-PEG and the GLP-1(A8S)-FGF21(V76)-PEG fusions showed a dose response from 0.05 to 0.2 mg/kg (0.05 diagonal hashed bars; 0.1 horizontal hashed bars; 0.2 vertical hashed bars) for glucose control and body weight. At 0.2 mg/kg, GLP-1(A8S)-FGF21(V76)-PEG was more efficacious than GLP-1(A8S)-FGF21FGF21(R154C)-PEG particularly for lowering of glucose levels, body weight, serum triglycerides and cholesterol. Figures 6A-6C show levels of basal glucose (AUC), alanine aminotransferase (ALT), and serum total cholesterol, respectively. Figures 6D-6F show levels of D12 body weight, hepatic lipids, and serum triglycerides, respectively.

[00035] Figures 7A-7B show the ability of GLP-1(A8S)-FGF21(V76)-PEG fusion proteins to improve pancreatic function and increase islet insulin content, relative to the combination of the individual agents. Db/db mice were dosed twice per week, for four weeks, with a combination of individual GLP-1(A8S)-PEG +FGF21(V76)-PEG (V76+V253; vertical hashed bars), as well as with the GLP-1(A8S)-FGF21(V76)-PEG dual function fusion protein of the invention (V272; diagonal hashed bars; light for 0.5 mg/kg and dark for 1 mg/kg), or vehicle (white bars).

[00036] Figures 8A-8B show the ability of GLP-1-FGF21-PEG fusion proteins to improve glucose lowering and body weight, relative to the combination of the individual agents. Db/db mice were dosed twice per week, for two weeks, with vehicle (white bars; open circles), a combination of individual GLP-1(A8S)-PEG (V253; 3 mg/kg (light horizontal hashed bars; open diamonds) and 5 mg/kg (dark horizontal bars; diamonds)) and FGF21(V76)-PEG (3 mg/kg (V76; light diagonal hashed bars; open squares) and 5 mg/kg (dark diagonal hashed bars; squares)), as well as with the GLP-1(A8S)-FGF21(V76)-PEG dual function fusion protein of the invention (V272; 0.2 mg/kg (smalled checked bars; open triangles); 1 mg/kg (large checked bars; triangles)). As seen in the figure, 0.2 mg/kg of the GLP-1(7-35; A8S)-FGF21(V76)-PEG dual function fusion protein (V272) is as effective as 5 mg/kg of FGF21(V76)-PEG. Also as seen in the figure, 1.0 mg/kg of the GLP-1(7-35; A8S)-FGF21(V76)-PEG dual function fusion protein (V272) is more effective that the maximal effective combination doses of FGF21(V76)-PEG + GLP-1(7-35; A8S)-PEG (1+1 mg/kg (light horizontal hashed bars; exes) and 3 + 3 mg/kg (dark horizontal bars; stars), for both glycemic and body weight endpoints.

[00037] Figure 9 shows the results of an assay for FGF21 activity. Phosphorylation of ERK is measured after treatment of HEK293-beta-klotho cells with the indicated compounds. The activity of GLP-1(A8S)-V76-PEG (V272) is significantly higher in magnitude of signal compared to the same FGF21 variant FGF21(V76)-PEG in the presence or absence of equimolar Exenatide.

[00038] Figure 10 presents results of a receptor pharmacology assay used to measure the activity of fusion proteinscompared to matched single agonist proteins or peptides. HEK293 cells transfected with GLP-1R and FGFR1c/beta-klotho (10a and 10b) or with GLP-1R alone (10c and 10d) were assayed for beta-arrestin recruitment to GLP-1R after treatment with compounds for 1 hour. Molecules tested in the assay were: Exendin-4 (squares), GLP-1(A8S)-FGF21(V76)-PEG (V272; circles), GLP-1(A8S)-PEG (V253; triangles), V253 +FGF21(V76)-PEG (open diamonds), Exendin-4(1-39)-Fc-FGF21(V103) (V211; open circles), Exendin-4(1-39)-Fc (V201; open triangles), and V201+ FGF21(V101) (diamonds).

DETAILED DESCRIPTION OF THE INVENTION

[00039] The proteins of the present invention represent modified versions of the full-length, wild-type FGF21 polypeptide, as known in the art. FGF21 wild-type sequence will serve as a reference sequence (SEQ ID NO:1), for instance, when comparisons between the FGF21 wild-type sequence and the protein variants are necessary. The FGF21 wild-type sequence has NCBI reference sequence number NP_061986.1, and can be found in such issued patents as, e.g., US 6,716,626B1, assigned to Chiron Corporation (SEQ ID NO:1).

```
Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Ser
 1               5                  10                  15
Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
            20                  25                  30
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
        35                  40                  45
Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
    50                  55                  60
Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
65                  70                  75                  80
Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
                85                  90                  95
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
            100                 105                 110
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
        115                 120                 125
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
    130                 135                 140
His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly
145                 150                 155                 160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
                165                 170                 175
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
            180                 185                 190
Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
        195                 200                 205
Ser
209
```

[00040] The corresponding mRNA sequence coding for the full-length FGF21 polypeptide (NCBI reference sequence number NM_019113.2) is shown below (SEQ ID NO:2):

```
  1 ctgtcagctg aggatccagc cgaaagagga gccaggcact caggccacct gagtctactc
 61 acctggacaa ctggaatctg gcaccaattc taaaccactc agcttctccg agctcacacc
121 ccggagatca cctgaggacc cgagccattg atggactcgg acgagaccgg gttcgagcac
181 tcaggactgt gggtttctgt gctggctggt cttctgctgg gagcctgcca ggcacacccc
```

```
241 atccctgact ccagtcctct cctgcaattc gggggccaag tccggcagcg gtacctctac
301 acagatgatg cccagcagac agaagcccac ctggagatca gggaggatgg gacggtgggg
361 ggcgctgctg accagagccc cgaaagtctc ctgcagctga aagccttgaa gccgggagtt
421 attcaaatct tgggagtcaa gacatccagg ttcctgtgcc agcggccaga tggggccctg
481 tatggatcgc tccactttga ccctgaggcc tgcagcttcc gggagctgct tcttgaggac
541 ggatacaatg tttaccagtc cgaagcccac ggcctcccgc tgcacctgcc agggaacaag
601 tccccacacc gggaccctgc accccgagga ccagctcgct tcctgccact accaggcctg
661 ccccccgcac tcccggagcc acccggaatc ctggcccccc agccccccga tgtgggctcc
721 tcggaccctc tgagcatggt gggaccttcc cagggccgaa gccccagcta cgcttcctga
781 agccagaggc tgtttactat gacatctcct ctttatttat taggttattt atcttattta
841 ttttttttatt tttcttactt gagataataa agagttccag aggagaaaaa aaaaaaaaaa
901 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
```

[00041] The mature FGF21 sequence lacks a leader sequence and may also include other modifications of a polypeptide such as proteolytic processing of the amino terminus (with or without a leader sequence) and/or the carboxyl terminus, cleavage of a smaller polypeptide from a larger precursor, N-linked and/or O-linked glycosylation, and other post-translational modifications understood by those with skill in the art. A representative example of a mature FGF21 sequence has the following sequence (SEQ ID NO:3, which represents amino acid positions 29-209 of full-length FGF21 protein sequence (NCBI reference sequence number NP_061986.1)):

```
His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
                  5                  10                  15
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
             20                  25                  30
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
         35                  40                  45
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
     50                  55                  60
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly
 65                  70                  75                  80
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
                 85                  90                  95
Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
            100                 105                 110
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
        115                 120                 125
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
    130                 135                 140
Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
145                 150                 155                 160
Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser
                165                 170                 175
Pro Ser Tyr Ala Ser
            180
```

[00042] The corresponding cDNA sequence coding for the mature FGF21 polypeptide (SEQ ID NO:3) is shown below (SEQ ID NO:4):

```
   1 caccccatcc ctgactccag tcctctcctg caattcgggg gccaagtccg gcagcggtac
  61 ctctacacag atgatgccca gcagacagaa gcccacctgg agatcaggga ggatgggacg
 121 gtgggggggcg ctgctgacca gagccccgaa agtctcctgc agctgaaagc cttgaagccg
 181 ggagttattc aaatcttggg agtcaagaca tccaggttcc tgtgccagcg gccagatggg
 240 gccctgtatg gatcgctcca ctttgaccct gaggcctgca gcttccggga gctgcttctt
 301 gaggacggat acaatgttta ccagtccgaa gcccacggcc tcccgctgca cctgccaggg
 360 aacaagtccc cacaccggga ccctgcaccc cgaggaccag ctcgcttcct gccactacca
 421 ggcctgcccc ccgcactccc ggagccaccc ggaatcctgg ccccccagcc ccccgatgtg
 481 ggctcctcgg accctctgag catggtggga ccttcccagg gccgaagccc cagctacgct
 541 tcctga
```

[00043] The proteins of the invention of the present invention represent modified versions of the full-length, wild-type GLP-1 polypeptide, as known in the art. GLP-1 wild-type sequence will serve as a reference sequence (SEQ ID NO:5), for instance, when comparisons between the GLP-1 wild-type sequence and the protein variants are necessary.

[00044] The GLP-1 wild-type sequence is post-translationally modified, and otherwise derived from, preproglucagon wild-type sequence. Preproglucagon sequence (SEQ ID NO:5) has NCBI reference sequence number NP_002045, and can be found in such patent publications as, e.g., WO98/19698 and WO87/06941A, assigned to Eli Lilly and Co. and the General Hospital Corp., respectively. As described, for example, in Goke, et al. (1991) European Journal of Clinical Investigation 21, 135, GLP-1 is a 37mer that is derived from preproglucagon (GLP-1 constitutes residues 92-138 of preproglucagon, and is underlined in SEQ ID NO:5, below). GLP-1 is further processed into an active 31mer, by cleavage of the six N-term residues.

```
   1 MKSIYFVAGL FVMLVQGSWQ RSLQDTEEKS RSFSASQADP LSDPDQMNED KRHSQGTFTS
  61 DYSKYLDSRR AQDFVQWLMN TKRNRNNIAK RHDEFERHAE GTFTSDVSSY LEGQAAKEFI
 121 AWLVKGRGRR DFPEEVAIVE ELGRRHADGS FSDEMNTILD NLAARDFINW LIQTKITDRK
```

[00045] An example of the processed, wild-type GLP-1 sequence is as follows (SEQ ID NO:129):

```
1  HDEFERHAEG TFTSDVSSYL EGQAAKEFIA WLVKGRG
```

[00046] An example of the active GLP-1 sequence is as follows (SEQ ID NO:30):

1 HAEGTFTSDV SSYLEGQAAK EFIAWLVKGR G

[00047] The corresponding mRNA sequence coding for the preproglucagon GLP-1 wild-type sequence (SEQ ID NO:5) is shown below (SEQ ID NO:6):

```
   1 gcatagaatg cagatgagca aagtgagtgg gagagggaag tcatttgtaa caaaaactca
  61 ttatttacag atgagaaatt tatattgtca gcgtaatatc tgtgaggcta aacagagctg
 121 gagagtatat aaaagcagtg cgccttggtg cagaagtaca gagcttagga cacagagcac
 181 atcaaaagtt cccaaagagg gcttgctctc tcttcacctg ctctgttcta cagcacacta
 241 ccagaagaca gcagaaatga aaagcattta ctttgtggct ggattatttg taatgctggt
 301 acaaggcagc tggcaacgtt cccttcaaga cacagaggag aaatccagat cattctcagc
 361 ttcccaggca gacccactca gtgatcctga tcagatgaac gaggacaagc gccattcaca
 421 gggcacattc accagtgact acagcaagta tctggactcc aggcgtgccc aagattttgt
 481 gcagtggttg atgaatacca agaggaacag gaataacatt gccaaacgtc acgatgaatt
 541 tgagagacat gctgaaggga cctttaccag tgatgtaagt tcttatttgg aaggccaagc
 601 tgccaaggaa ttcattgctt ggctggtgaa aggccgagga aggcgagatt tcccagaaga
 661 ggtcgccatt gttgaagaac ttggccgcag acatgctgat ggttctttct ctgatgagat
 721 gaacaccatt cttgataatc ttgccgccag ggactttata aactggttga ttcagaccaa
 781 aatcactgac aggaaataac tatatcacta ttcaagatca tcttcacaac atcacctgct
 841 agccacgtgg gatgtttgaa atgttaagtc ctgtaaattt aagaggtgta ttctgaggcc
 901 acattgcttt gcatgccaat aaataaattt tcttttagtg ttgtgtagcc aaaaattaca
 961 aatggaataa agttttatca aaatattgct aaaatatcag ctttaaaata tgaaagtgct
1021 agattctgtt attttcttct tattttggat gaagtacccc aacctgttta catttagcga
1081 taaaattatt tttctatgat ataatttgta aatgtaaatt attccgatct gacatatctg
1141 cattataata ataggagaat agaagaactg gtagccacag tggtgaaatt ggaaagagaa
1201 ctttcttcct gaaacctttg tcttaaaaat actcagcttt caatgtatca aagatacaat
1261 taaataaaat tttcaagctt ctttaccaaa aaaaaaaa
```

[00048] The proteins of the invention of the present invention represent modified versions of the full-length, wild-type Exendin-4 polypeptide, as known in the art. Exendin-4 wild-type sequence will serve as a reference sequence (SEQ ID NO:7), for instance, when comparisons between the Exendin-4 wild-type sequence and the protein variants are necessary.

[00049] The Exendin-4 wild-type sequence has NCBI reference sequence number GenBank: AAB22006.1, and can be found in such issued patents as, e.g., US 5,424,286, assigned to Amylin Pharmaceuticals, Inc. and Eli Lilly and Co. An example of the wild-type sequence is as follows (SEQ ID NO:7):

1 HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPS

[00050] The proteins of the invention may comprise protein variants or mutants of the wild-type proteins listed herein, e.g., FGF21 variants, GLP-1 variants, and/or Exendin-4 variants. As used herein, the terms "protein variant," "human variant," "polypeptide or protein variant," "variant," "mutant," as well as any like terms or specific versions thereof (e.g., "FGF21 protein variant," "human GLP-1 variant," "Exendin-4 polypeptide or protein variant," "variant," "FGF21 mutant," etc.) define protein or polypeptide sequences that comprise modifications, truncations, other variants of naturally occurring (i.e., wild-type) protein or polypeptide counterparts or corresponding native sequences. "Variant FGF21" or "FGF21 mutant," for instance, is described relative to the wild-type (i.e., naturally occurring) FGF21 protein as described herein.

Representative dual function protein sequences of the invention are listed in Table 1. The descriptions of said agonists include the individual constituent agonists and, where applicable a linker. If a variant is used as a constituent agonist, the changes or substitutions made are numbered relative to their wild-type counterparts. By way of example, "Dual Function 1- Protein" (SEQ ID NO:9) contains residues 7-35 of the wild-type GLP-1 sequence as described herein (i.e., a GLP-1 receptor agonist), a linker sequence, and an FGF21 variant (i.e., an FGF21 receptor agonist) with a number of listed changes to the FGF21 wild-type sequence (SEQ ID NO:1) as described herein.

Table 1. Dual Function Protein and Nucleotides of the invention

SEQ ID NO:	Variant NO:	Sequence	Long Name	Short Name
9	V272	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76)-40KPEGb
10		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAT AGCAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT TGCGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76)
11	V277	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGG SGGGGSGGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS	Exendin4(1-30)-SGGGGSGGGGS GGGGSGGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A)	Ex(1-30)-L20-FGF21(V76)-40KPEGb

12		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGGT TCTGGCGGCG GTGGTAGCGG TGGCGGCGGT GATAGCAGCC CGCTGCTGCA GTTTGGCGGC CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGG AAACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGCATCAGA GCCCGGAAAG CCTGCTGGAA CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGAAAC CGGATGGCGC GCTGTATGGC AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CCGTAGCCCG CATTGCGATC CGGCACCGCA GGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGGCGA TGGTGGGTCC GAGCCAGGGT CGTAGCCCGA GCTATGCGAG CTAA	Exendin4(1-30)-SGGGGSGGGGS GGGGSGGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	Ex(1-30)-L20-FGF21(V76)
13	V220	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGDSSPLLQF GGQVRQRYLY TDDAQQTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	GLP-1(7-35;A8S)-GSG-FGF21(33-209)	GLP-1(A8S)-L3-FGF21
14		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGATA GCAGCCCGCT GCTGCAGTTT GGCGGCCAGG TGCGTCAGCG TTATCTGTAT ACCGATGATG CGCAGCAGAC CGAAGCGCAT CTGGAAATTC GTGAAGATGG CACCGTGGGC GGTGCGGCGG ATCAGAGCCC GGAAAGCCTG CTGCAGCTGA AAGCGCTGAA ACCGGGCGTG ATTCAGATTC TGGGCGTGAA AACCAGCCGT TTTCTGTGCC AGCGTCCGGA TGGCGCGCTG TATGGCAGCC TGCATTTTGA TCCGGAAGCG TGCAGCTTTC GTGAACTGCT GCTGGAAGAT GGCTATAACG TGTATCAGAG CGAAGCGCAT GGCCTGCCGC TGCATCTGCC GGGCAACAAA AGCCCGCATC GTGATCCGGC ACCGCGTGGT CCGGCGCGTT TTCTGCCGCT GCCGGGTCTG CCGCCGGCAC TGCCGGAACC GCCGGGTATT CTGGCCCCGC AGCCGCCGGA TGTTGGTAGC AGCGATCCGC TGTCTATGGT GGGTCCGAGC CAGGGTCGTA GCCCGAGCTA TGCGAGCTAA	GLP-1(7-35;A8S)-GSG-FGF21(33-209)	GLP-1(A8S)-L3-FGF21

15	V219	HAEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGDSSPLLQF GGQVRQRYLY TDDAQQTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	GLP-1(7-35)-GSG-FGF21(33-209)	GLP-1-L3-FGF21
16		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGATA GCAGCCCGCT GCTGCAGTTT GGCGGCCAGG TGCGTCAGCG TTATCTGTAT ACCGATGATG CGCAGCAGAC CGAAGCGCAT CTGGAAATTC GTGAAGATGG CACCGTGGGC GGTGCGGCGG ATCAGAGCCC GGAAAGCCTG CTGCAGCTGA AAGCGCTGAA ACCGGGCGTG ATTCAGATTC TGGGCGTGAA AACCAGCCGT TTTCTGTGCC AGCGTCCGGA TGGCGCGCTG TATGGCAGCC TGCATTTTGA TCCGGAAGCG TGCAGCTTTC GTGAACTGCT GCTGGAAGAT GGCTATAACG TGTATCAGAG CGAAGCGCAT GGCCTGCCGC TGCATCTGCC GGGCAACAAA AGCCCGCATC GTGATCCGGC ACCGCGTGGT CCGGCGCGTT TTCTGCCGCT GCCGGGTCTG CCGCCGGCAC TGCCGGAACC GCCGGGTATT CTGGCCCCGC AGCCGCCGGA TGTTGGTAGC AGCGATCCGC TGTCTATGGT GGGTCCGAGC CAGGGTCGTA GCCCGAGCTA TGCGAGCTAA	GLP-1(7-35)-GSG-FGF21(33-209)	GLP-1-L3-FGF21
17	V295	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGDSSPLLQF GGQVRQRYLY TDDAQQTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	Exendin4(1-39)-GGG-FGF21(33-209)	Ex(1-39)-L3-FGF21
18		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT GGTGGCGACT CGAGCCCGCT GCTGCAGTTT GGCGGCCAGG TGCGTCAGCG TTATCTGTAT ACCGATGATG CGCAGCAGAC CGAAGCGCAT CTGGAAATTC GTGAAGATGG CACCGTGGGC GGTGCGGCGG ATCAGAGCCC GGAAAGCCTG CTGCAGCTGA AAGCGCTGAA ACCGGGCGTG ATTCAGATTC TGGGCGTGAA AACCAGCCGT TTTCTGTGCC AGCGTCCGGA TGGCGCGCTG TATGGCAGCC TGCATTTTGA TCCGGAAGCG TGCAGCTTTC GTGAACTGCT GCTGGAAGAT GGCTATAACG TGTATCAGAG CGAAGCGCAT GGCCTGCCGC TGCATCTGCC GGGCAACAAA AGCCCGCATC GTGATCCGGC ACCGCGTGGT CCGGCGCGTT TTCTGCCGCT GCCGGGTCTG CCGCCGGCAC TGCCGGAACC GCCGGGTATT	Exendin4(1-39)-GGG-FGF21(33-209)	Ex(1-39)-L3-FGF21

		CTGGCCCCGC AGCCGCCGGA TGTTGGTAGC AGCGATCCGC TGTCTATGGT GGGTCCGAGC CAGGGTCGTA GCCCGAGCTA TGCGAGCTAA		
19	V224	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S)-L10-FGF21
20		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S)-L10-FGF21
21	V223	HAEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35)-GSGGGGSGGG-FGF21(33-209)	GLP-1-L10-FGF21
22		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT	GLP-1(7-35)-GSGGGGSGGG-FGF21(33-209)	GLP-1-L10-FGF21

		CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
23	V234	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGDSS PLLQFGGQVR QRYLYTDDAQ QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQR PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPGNKSPHRD PAPRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLS MVGPSQGRSP SYAS	Exendin4 (1-39)-GGGGSGGG-FGF21(33-209)	Ex(1-39)-L8-FGF21
24		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT GGTGGTGGTT CTGGCGGTGG CGACTCGAGC CCGCTGCTGC AGTTTGGCGG CCAGGTGCGT CAGCGTTATC TGTATACCGA TGATGCGCAG CAGACCGAAG CGCATCTGGA AATTCGTGAA GATGGCACCG TGGGCGGTGC GGCGGATCAG AGCCCGGAAA GCCTGCTGCA GCTGAAAGCG CTGAAACCGG GCGTGATTCA GATTCTGGGC GTGAAAACCA GCCGTTTTCT GTGCCAGCGT CCGGATGGCG CGCTGTATGG CAGCCTGCAT TTTGATCCGG AAGCGTGCAG CTTTCGTGAA CTGCTGCTGG AAGATGGCTA TAACGTGTAT CAGAGCGAAG CGCATGGCCT GCCGCTGCAT CTGCCGGGCA ACAAAAGCCC GCATCGTGAT CCGGCACCGC GTGGTCCGGC GCGTTTTCTG CCGCTGCCGG GTCTGCCGCC GGCACTGCCG GAACCGCCGG GTATTCTGGC CCCGCAGCCG CCGGATGTTG GTAGCAGCGA TCCGCTGTCT ATGGTGGGTC CGAGCCAGGG TCGTAGCCCG AGCTATGCGA GCTAA	Exendin4 (1-39)-GGGGSGGG-FGF21(33-209)	Ex(1-39)-L8-FGF21
25	V193	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGG SGGGGSADKT HTCPPCPAPE AAGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGKGGGGSG GGGSGGGGSD SSPLLQFGGQ VRQRYLYTDD ACQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPCNRSPH RDPASRGPAR	GLP-1(7-35;A8S)-S(GGGGS)$_3$A-Fc-(GGGGS)$_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L17-Fc-L15-FGF21(V103)

		FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGGSQAR SPSYAS		
26	V194	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGG SGGGGSADKT HTCPPCPAPE AAGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGKGGGGSG GGGSGGGGSD SSPLLQFGGQ VRQRYLYTDD ACQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPCNRSPH RDPASRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYES	GLP-1(7-35;A8S)-S(GGGGS)$_3$A-Fc-(GGGGS)$_3$-FGF21(33-209; Q55C-R105K-G148C-K150R-P158S-S195A-G202A-A208E)	GLP-1(A8S)-L17-Fc-L15-FGF21(V194)
27	V195	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGG SGGGGSADKT HTCPPCPAPE AAGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGKGGGGSG GGGSGGGGSD SSPLLQFGGQ VRQRYLYTDD ACQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPCNRSPH RDPASRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYDS	GLP-1(7-35;A8S)-S(GGGGS)$_3$A-Fc-(GGGGS)$_3$-FGF21(33-209;Q55C-R105K-G148C-P158S-K150R-S195A-G202A-A208D)	GLP-1(A8S)-L17-Fc-L15-FGF21(V195)
28	V296	HGEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8G)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A)	GLP-1(A8G)-L10-FGF21(V76)-40KPEGb
29		CATGGTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAT AGCAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT	GLP-1(7-35;A8G)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-	GLP-1(A8G)-L10-FGF21(V76)

		GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT TGCGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	R105K-K150R-R154C-R159Q-S195A)	
34	V294	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH UDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209; R154Pcl-40 kDa branched PEG)	GLP-1(A8S)-L10-FGF21-154Pcl-40KPEGb
35		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TAGGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209; R154TAG)	GLP-1(A8S)-L10-FGF21-154TAG
38	V298	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGDSS PLLQFGGQVR QRYLYTDDAQ QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQR PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPGNKSPHUD PAPRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLS MVGPSQGRSP SYAS	Exendin4 (1-39)-GGGGSGGG-FGF21(33-209; R154TAG)-40 kDa branched	Ex(1-39)-L8-FGF21(154Pcl)-40KPEGb

			PEG	
39		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT GGTGGTGGTT CTGGCGGTGG CGACTCGAGC CCGCTGCTGC AGTTTGGCGG CCAGGTGCGT CAGCGTTATC TGTATACCGA TGATGCGCAG CAGACCGAAG CGCATCTGGA AATTCGTGAA GATGGCACCG TGGGCGGTGC GGCGGATCAG AGCCCGGAAA GCCTGCTGCA GCTGAAAGCG CTGAAACCGG GCGTGATTCA GATTCTGGGC GTGAAAACCA GCCGTTTTCT GTGCCAGCGT CCGGATGGCG CGCTGTATGG CAGCCTGCAT TTTGATCCGG AAGCGTGCAG CTTTCGTGAA CTGCTGCTGG AAGATGGCTA TAACGTGTAT CAGAGCGAAG CGCATGGCCT GCCGCTGCAT CTGCCGGGCA ACAAAAGCCC GCATTAGGAT CCGGCACCGC GTGGTCCGGC GCGTTTTCTG CCGCTGCCGG GTCTGCCGCC GGCACTGCCG GAACCGCCGG GTATTCTGGC CCCGCAGCCG CCGGATGTTG GTAGCAGCGA TCCGCTGTCT ATGGTGGGTC CGAGCCAGGG TCGTAGCCCG AGCTATGCGA GCTAA	Exendin4 (1-39)-GGGGSGGG-FGF21(33-209; R154TAG)	Ex(1-39)-L8-FGF21(154TAG)
40	V235	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(A8S)-L8-FGF21(154C)-40KPEGb

41		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S)-L8-FGF21(154C)
42	V239	HAEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1-L8-FGF21(154C)-40KPEGb
43		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1-L8-FGF21(154C)

44	V299	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGDSS PLLQFGGQVR QRYLYTDDAQ QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQR PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPGNKSPHCD PAPRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLS MVGPSQGRSP SYAS	Exendin4 (1-39)-GGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	Ex(1-39)-L8-FGF21(154C)-40KPEGb
45		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT GGTGGTGGTT CTGGCGGTGG CGACTCGAGC CCGCTGCTGC AGTTTGGCGG CCAGGTGCGT CAGCGTTATC TGTATACCGA TGATGCGCAG CAGACCGAAG CGCATCTGGA AATTCGTGAA GATGGCACCG TGGGCGGTGC GGCGGATCAG AGCCCGGAAA GCCTGCTGCA GCTGAAAGCG CTGAAACCGG GCGTGATTCA GATTCTGGGC GTGAAAACCA GCCGTTTTCT GTGCCAGCGT CCGGATGGCG CGCTGTATGG CAGCCTGCAT TTTGATCCGG AAGCGTGCAG CTTTCGTGAA CTGCTGCTGG AAGATGGCTA TAACGTGTAT CAGAGCGAAG CGCATGGCCT GCCGCTGCAT CTGCCGGGCA ACAAAAGCCC GCATTGCGAT CCGGCACCGC GTGGTCCGGC GCGTTTTCTG CCGCTGCCGG GTCTGCCGCC GGCACTGCCG GAACCGCCGG GTATTCTGGC CCCGCAGCCG CCGGATGTTG GTAGCAGCGA TCCGCTGTCT ATGGTGGGTC CGAGCCAGGG TCGTAGCCCG AGCTATGCGA GCTAA	Exendin4 (1-39)-GGGGSGGG-FGF21(33-209;R154C)	Ex(1-39)-L8-FGF21 (154C)
46	V271	HAPGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; E9P)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(E9P)-L10-FGF21(154C)-40KPEGb
47		CATGCGCCGG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT	GLP-1(7-35; E9P)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(E9P)-L10-FGF21(154C)

		TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
48	V251	GHAEGTFTSD VSSYLEGQAA KEFIAWLVKG GSGGGGSGGG DSSPLLQFGG QVRQRYLYTD DAQQTEAHLE IREDGTVGGA ADQSPESLLQ LKALKPGVIQ ILGVKTSRFL CQRPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNKSP HCDPAPRGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLSMVGPSQG RSPSYAS	GLP-1(6-35; 6G)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	G-GLP-1-L10-FGF21(154C)-40KPEGb
49		GGTCATGCGG AAGGCACTTT TACTAGCGAT GTTTCTAGCT ACCTGGAAGG CCAGGCTGCG AAAGAATTCA TCGCGTGGCT GGTTAAAGGC GGTTCTGGTG GTGGTGGTTC TGGCGGTGGC GACTCGAGCC CGCTGCTGCA GTTTGGCGGC CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGC AGACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGGATCAGA GCCCGGAAAG CCTGCTGCAG CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGCGTC CGGATGGCGC GCTGTATGGC AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CAAAAGCCCG CATTGCGATC CGGCACCGCG TGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGTCTA TGGTGGGTCC GAGCCAGGGT CGTAGCCCGA GCTATGCGAG CTAA	GLP-1(6-35; 6G)-GSGGGGSGGG-FGF21(33-209)	G-GLP-1-L10-FGF21(154C)
50	V265	HSEGTFTADA SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-S14A-V16A)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(A8S; 14/16)-L10-FGF21(154C)-40KPEGb
51		CATTCTGAAG GCACTTTTAC TGCTGATGCT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT	GLP-1(7-35; A8S-S14A-V16A)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S; 14/16)-L10-FGF21(154C)

		CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
52	V270	HSEGTFTSDA AAYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-V16A-S17A-S18A)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(A8S; 16/17/18)-L10-FGF21(154C)-40KPEGb
53		CATTCTGAAG GCACTTTTAC TAGCGATGCT GCTGCTTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-V16A-S17A-S18A)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S)-16/17/18)-L10-FGF21(154C)
54	V266	HSEGTFTSDV SSYAEGAAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-L20A-Q23A)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(A8S; 20/23)-L10-FGF21(154C)-40KPEGb

55		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACG CTGAAGGCGC TGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-L20A-Q23A)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S; 20/23)-L10-FGF21(154C)
58	V300	HSEGTFTSDS SPLLQFGGQV RQRYLYTDDA QQTEAHLEIR EDGTVGGAAD QSPESLLQLK ALKPGVIQIL GVKTSRFLCQ RPDGALYGSL HFDPEACSFR ELLLEDGYNV YQSEAHGLPL HLPGNKSPHC DPAPRGPARF LPLPGLPPAL PEPPGILAPQ PPDVGSSDPL SMVGPSQGRS PSYAS	GLP-1(7-14; A8S)-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(A8S; 7-14)-L0-FGF21(154C)-40KPEGb
59		CATTCTGAAG GCACTTTTAC TAGCGATAGC AGCCCGCTGC TGCAGTTTGG CGGCCAGGTG CGTCAGCGTT ATCTGTATAC CGATGATGCG CAGCAGACCG AAGCGCATCT GGAAATTCGT GAAGATGGCA CCGTGGGCGG TGCGGCGGAT CAGAGCCCGG AAAGCCTGCT GCAGCTGAAA GCGCTGAAAC CGGGCGTGAT TCAGATTCTG GGCGTGAAAA CCAGCCGTTT TCTGTGCCAG CGTCCGGATG GCGCGCTGTA TGGCAGCCTG CATTTTGATC CGGAAGCGTG CAGCTTTCGT GAACTGCTGC TGGAAGATGG CTATAACGTG TATCAGAGCG AAGCGCATGG CCTGCCGCTG CATCTGCCGG GCAACAAAAG CCCGCATTGC GATCCGGCAC CGCGTGGTCC GGCGCGTTTT CTGCCGCTGC CGGGTCTGCC GCCGGCACTG CCGGAACCGC CGGGTATTCT GGCCCCGCAG CCGCCGGATG TTGGTAGCAG CGATCCGCTG TCTATGGTGG GTCCGAGCCA GGGTCGTAGC CCGAGCTATG CGAGCTAA	GLP-1(7-14; A8S)-FGF21(33-209;R154C)	GLP-1(A8S;7-14)-L0-FGF21(154C)
60	V262	HSEGTFTSDV SSGGGGSGGG DSSPLLQFGG QVRQRYLYTD DAQQTEAHLE IREDGTVGGA ADQSPESLLQ LKALKPGVIQ ILGVKTSRFL CQRPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNKSP HCDPAPRGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD	GLP-1(7-18; A8S)-GGGGSGGG-FGF21(33-209;R154C-40 kDa	GLP-1(A8S;7-18)-L8-FGF21(154C)-

		PLSMVGPSQG RSPSYAS	branched PEG)	40KPEGb
61		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTTCTGGTG GTGGTGGTTC TGGCGGTGGC GACTCGAGCC CGCTGCTGCA GTTTGGCGGC CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGC AGACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGGATCAGA GCCCGGAAAG CCTGCTGCAG CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGCGTC CGGATGGCGC GCTGTATGGC AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CAAAAGCCCG CATTGCGATC CGGCACCGCG TGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGTCTA TGGTGGGTCC GAGCCAGGGT CGTAGCCCGA GCTATGCGAG CTAA	GLP-1(7-18; A8S)-GGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S;7-18)-L8-FGF21(154C)
62	V261	HSEGTFTSDV SSYLESGGGG SGGGDSSPLL QFGGQVRQRY LYTDDAQQTE AHLEIREDGT VGGAADQSPE SLLQLKALKP GVIQILGVKT SRFLCQRPDG ALYGSLHFDP EACSFRELLL EDGYNVYQSE AHGLPLHLPG NKSPHCDPAP RGPARFLPLP GLPPALPEPP GILAPQPPDV GSSDPLSMVG PSQGRSPSYA S	GLP-1(7-21; A8S)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(A8S;7-21)-L9-FGF21(154C)-40KPEGb
63		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAATCTGG TGGTGGTGGT TCTGGCGGTG GCGACTCGAG CCCGCTGCTG CAGTTTGGCG GCCAGGTGCG TCAGCGTTAT CTGTATACCG ATGATGCGCA GCAGACCGAA GCGCATCTGG AAATTCGTGA AGATGGCACC GTGGGCGGTG CGGCGGATCA GAGCCCGGAA AGCCTGCTGC AGCTGAAAGC GCTGAAACCG GGCGTGATTC AGATTCTGGG CGTGAAAACC AGCCGTTTTC TGTGCCAGCG TCCGGATGGC GCGCTGTATG GCAGCCTGCA TTTTGATCCG GAAGCGTGCA GCTTTCGTGA ACTGCTGCTG GAAGATGGCT ATAACGTGTA TCAGAGCGAA GCGCATGGCC TGCCGCTGCA TCTGCCGGGC AACAAAAGCC CGCATTGCGA TCCGGCACCG CGTGGTCCGG CGCGTTTTCT GCCGCTGCCG GGTCTGCCGC CGGCACTGCC GGAACCGCCG GGTATTCTGG CCCCGCAGCC GCCGGATGTT GGTAGCAGCG ATCCGCTGTC TATGGTGGGT CCGAGCCAGG GTCGTAGCCC GAGCTATGCG AGCTAA	GLP-1(7-21; A8S)-SGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S;7-21)-L9-FGF21(154C)

66	V260	HSEGTFTSDV SSYLEGQAAK EFISGGGGSG GGDSSPLLQF GGQVRQRYLY TDDAQQTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHCDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	GLP-1(7-29; A8S)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(A8S;7-29)-L9-FGF21(154C)-40KPEGb
67		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCT CTGGTGGTGG TGGTTCTGGC GGTGGCGACT CGAGCCCGCT GCTGCAGTTT GGCGGCCAGG TGCGTCAGCG TTATCTGTAT ACCGATGATG CGCAGCAGAC CGAAGCGCAT CTGGAAATTC GTGAAGATGG CACCGTGGGC GGTGCGGCGG ATCAGAGCCC GGAAAGCCTG CTGCAGCTGA AAGCGCTGAA ACCGGGCGTG ATTCAGATTC TGGGCGTGAA AACCAGCCGT TTTCTGTGCC AGCGTCCGGA TGGCGCGCTG TATGGCAGCC TGCATTTTGA TCCGGAAGCG TGCAGCTTTC GTGAACTGCT GCTGGAAGAT GGCTATAACG TGTATCAGAG CGAAGCGCAT GGCCTGCCGC TGCATCTGCC GGGCAACAAA AGCCCGCATT GCGATCCGGC ACCGCGTGGT CCGGCGCGTT TTCTGCCGCT GCCGGGTCTG CCGCCGGCAC TGCCGGAACC GCCGGGTATT CTGGCCCCGC AGCCGCCGGA TGTTGGTAGC AGCGATCCGC TGTCTATGGT GGGTCCGAGC CAGGGTCGTA GCCCGAGCTA TGCGAGCTAA	GLP-1(7-29; A8S)-SGGGGSGGG-FGF21(33-209)	GLP-1(A8S;7-29)-L9-FGF21(154C)
68	V259	HSEGTFTSDV SSYLEGQAAK EFIAWLSGGG GSGGGDSSPL LQFGGQVRQR YLYTDDAQQT EAHLEIREDG TVGGAADQSP ESLLQLKALK PGVIQILGVK TSRFLCQRPD GALYGSLHFD PEACSFRELL LEDGYNVYQS EAHGLPLHLP GNKSPHCDPA PRGPARFLPL PGLPPALPEP PGILAPQPPD VGSSDPLSMV GPSQGRSPSY AS	GLP-1(7-32; A8S)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(A8S;7-32)-L9-FGF21(154C)-40KPEGb
69		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGTC TGGTGGTGGT GGTTCTGGCG GTGGCGACTC GAGCCCGCTG CTGCAGTTTG GCGGCCAGGT GCGTCAGCGT TATCTGTATA CCGATGATGC GCAGCAGACC GAAGCGCATC TGGAAATTCG TGAAGATGGC ACCGTGGGCG GTGCGGCGGA TCAGAGCCCG GAAAGCCTGC TGCAGCTGAA AGCGCTGAAA CCGGGCGTGA TTCAGATTCT GGGCGTGAAA ACCAGCCGTT TTCTGTGCCA GCGTCCGGAT GGCGCGCTGT ATGGCAGCCT GCATTTTGAT CCGGAAGCGT GCAGCTTTCG TGAACTGCTG CTGGAAGATG GCTATAACGT GTATCAGAGC GAAGCGCATG GCCTGCCGCT GCATCTGCCG GGCAACAAAA GCCCGCATTG CGATCCGGCA CCGCGTGGTC CGGCGCGTTT TCTGCCGCTG CCGGGTCTGC CGCCGGCACT GCCGGAACCG	GLP-1(7-32; A8S)-SGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S;7-32)-L9-FGF21(154C)

		CCGGGTATTC TGGCCCCGCA GCCGCCGGAT GTTGGTAGCA GCGATCCGCT GTCTATGGTG GGTCCGAGCC AGGGTCGTAG CCCGAGCTAT GCGAGCTAA		
70	V263	HSEGTFTSDV CSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-S17C-40 kDa branched PEG)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-S17C)-40KPEGb-L10-FGF21
71		CATTCTGAAG GCACTTTTAC TAGCGATGTT TGTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-S17C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-S17C)-L10-FGF21
72	V269	HSEGTFTSDV SSYLEGCAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-Q23C-40 kDa branched PEG)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-Q23C)-40KPEGb-L10-FGF21
73		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCTG TGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT	GLP-1(7-35; A8S-Q23C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-Q23C)-L10-FGF21

		CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
74	V243	HSEGTFTSDV SSYLEGQAAC EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-K26C-40 kDa branched PEG)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-K26C)-L10-FGF21-40KPEGb
75		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGTGT GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-K26C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-K26C)-L10-FGF21
76	V264	HSEGTFTSDV SSYLEGQAAK CFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-E27C-40 kDa branched PEG)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-E27C)-40KPEGb-L10-FGF21
77		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA TGTTTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG	GLP-1(7-35; A8S-E27C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-E27C)-L10-FGF21

		GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
78	V244	HSEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-K34C-40 kDa branched PEG)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-K34C)-L10-FGF21-40KPEGb
79		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-K34C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-K34C)-L10-FGF21
80	V250	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGCGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S)-GSGGGGC(40 kDa branched PEG)GGG-FGF21(33-209)	GLP-1(A8S)-L10-C-FGF2140 KPEGb

81		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTGTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S)-GSGGGGCGGG-FGF21(33-209)	GLP-1(A8S)-L10-C-FGF21
82	V240	HAEGTFTSDV SSYLEGQAAC EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; K26C-40 kDa branched PEG)-GSGGGGSGGG-FGF21(33-209)	GLP-1(K26C)-L10-FGF21-40KPEGb
83		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGTGT GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; K26C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(K26C)-L10-FGF21

84	V241	HAEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; K34C-40 kDa branched PEG)-GSGGGGSGGG-FGF21(33-209)	GLP-1-(K34C)-L10-FGF21-40KPEGb
85		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; K34C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(K34C)-L10-FGF21
86	V242	HAEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGCGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35)-GSGGGGC(40 kDa branched PEG)GGG-FGF21(33-209)	GLP-1-L10-C-FGF21-40KPEGb
87		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTGTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT	GLP-1(7-35)-GSGGGGCGGG-FGF21(33-209)	GLP-1-L10-C-FGF21

		TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
88	V267	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	Exendin4(1-30)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	Ex(1-30)-L9-FGF21(154C)-40KPEGb
89		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	Exendin4(1-30)-SGGGGSGGG-FGF21(33-209;R154C)	Ex(1-30)-L9-FGF21(154C)
90	V268	HGEGTFTSDL SKQMEGQAVR LFIEWLKNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	Exendin4(1-30;E16G-E17Q)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	Ex(1-30;GQ)-L9-FGF21(154C)-40KPEGb
91		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGGGCA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC	Exendin4(1-30;E16G-E17Q)-SGGGGSGGG-FGF21(33-209;R154C)	Ex(1-30;GQ)-L9-FGF21(154C)

		CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
92	V301	HSEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH UDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-K34C)-GSGGGGSGGG-FGF21(33-209;R154Pcl)	GLP-1(A8S-K34C)-FGF21(154Pcl)-
93		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TAGGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-K34C)-GSGGGGSGGG-FGF21(33-209;R154TAG)	GLP-1(A8S-K34C)-L10-FGF21(154TAG)
94	V302	HAEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH UDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; K34C)-GSGGGGSGGG-FGF21(33-209;R154Pcl)	GLP-1(K34C)-L10-FGF21(154Pcl)
95		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT	GLP-1(7-35; K34C)-GSGGGGSGGG-FGF21(33-209;R154TAG)	GLP-1(K34C)-L10-FGF21(154TAG)

		GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TAGGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
96	V303	HSEGTFTSDV SSYLEGQAAK EFIAWLVUGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-K34Pcl)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S-K34Pcl)-L10-FGF21(154C)
97		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTAGGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-K34TAG)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S-K34TAG)-L10-FGF21(154C)
98	V281	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-	GLP-1(A8S)-L10-FGF21(V76-154R)

			S195A)	
99		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76-154R)
100	V304	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76-154R-130E)
101		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76-154R-130E)

		CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
102	V273	HSEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S-K34C-40 kDa branched PEG)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	GLP-1(A8S-K34C)-L10-FGF21(V76-154R)-40KPEGb
103		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35;A8S-K34C)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	GLP-1(A8S-K34C)-L10-FGF21(V76-154R)
104	V305	HSEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYAS	GLP-1(7-35;A8S-K34C-40 kDa branched PEG)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	GLP-1(A8S-K34C)-L10-FGF21(V76-154R-202A)-40KPEGb

105		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGCGCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35;A8S-K34C)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	GLP-1(A8S-K34C)-L10-FGF21(V76-154R-202A)
108	V306	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	Exendin4(1-30)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30)-L9-FGF21(V76-154R)
109		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	Exendin4(1-30)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30)-L9-FGF21(V76-154R)

110	V274	HGEGTFTSDL SKQMEEEAVR LFIEWLCNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	Exendin4(1-30;K27C-40 kDa branched PEG)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30;K27C)-L9-FGF21(V76-154R-130E)-40KPEGb
111		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGTG TAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	Exendin4(1-30;K27C)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30;K27C)-L9-FGF21(V76-154R-130E)
112	V307	HGEGTFTSDL SKQMEEEAVR LFIEWLKCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	Exendin4(1-30;N28C-40 kDa branched PEG)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30;N28C) -L9-FGF21(V76-154R-130E)-40KPEGb
113		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA ATGTGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG	Exendin4(1-30;N28C)-SGGGGSGGG-FGF21(33-209;Q56E-	Ex(1-30;N28C)-L9-FGF21(V76-

		GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	154R-130E)
114	V308	HGEGTFTSDL SKQMEEEAVR LFIEWLCNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYAS	Exendin4(1-30;K27C-40 kDa branched PEG)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	Ex(1-30;K27C)--L9-FGF21(V76-154R-202A)-40KPEGb
115		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGTG TAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGCGCGT AGCCCGAGCT ATGCGAGCTA A	Exendin4(1-30;K27C)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	Ex(1-30;K27C)-L9-FGF21(V76-154R-202A)

116	V309	HGEGTFTSDL SKQMEEEAVR LFIEWLKCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYAS	Exendin4(1-30;N28C-40 kDa branched PEG)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	Ex(1-30;N28C)-L9-FGF21(V76-154R-202A) -40KPEGb
117		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA ATGTGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGCGCGT AGCCCGAGCT ATGCGAGCTA A	Exendin4(1-30;N28C)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	Ex(1-30;N28C)-L9-FGF21(V76-154R-202A)
118	V276	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GSGGGGSGGG GSGGGGSGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS	GLP-1(7-35;A8S)-GGS$(GGGGS)_3$-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A)	GLP-1(A8S)-L20-FGF21(V76)-40KPEGb
119		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT GGTAGCGGTG GCGGCGGTTC TGGTGGTGGT GGTTCTGGCG GTGGCGGTAG CGGTGGCGGC GATAGCAGCC CGCTGCTGCA GTTTGGCGGC	GLP-1(7-35;A8S)-GGS$(GGGGS)_3$-FGF21(33-209;Q56E-D74H-Q82E-	GLP-1(A8S)-L20-FGF21(V76)

		CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGG AAACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGCATCAGA GCCCGGAAAG CCTGCTGGAA CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGAAAC CGGATGGCGC GCTGTATGGC AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CCGTAGCCCG CATTGCGATC CGGCACCGCA GGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGGCGA TGGTGGGTCC GAGCCAGGGT CGTAGCCCGA GCTATGCGAG CTAA	R105K-K150R-R154C-R159Q-S195A)	
120	V197	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG GGGGSDKTHT CPPCPAPEAA GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GKGGGGSGGG GSGGGGSDSS PLLQFGGQVR QRYLYTDDAC QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQK PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPCNRSPHRD PASRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLA MVGGSQARSP SYAS	Exendin4(1-30)-(GGGGS)-Fc-(GGGGS)$_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Exendin (1-30)-L5-Fc-L15-FGF21(V103))
121	V196	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG GGGGSGGGGS GGGGSDKTHT CPPCPAPEAA GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GKGGGGSGGG GSGGGGSDSS PLLQFGGQVR QRYLYTDDAC QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQK PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPCNRSPHRD PASRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLA MVGGSQARSP SYAS	Exendin4(1-30)-(GGGGS)$_3$-Fc-(GGGGS)$_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex (1-30)-L15-Fc-L15-FGF21(V103)
122	V199	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE	Exendin4(1-39)-(GGGGS)-Fc-(GGGGS)$_3$-FGF21(33-209;Q55C-R105K-	Ex (1-39)-L5-Fc-L15-FGF21(V103)

		EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGGGGSGGGG SGGGGSDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP DVGSSDPLAM VGGSQARSPS YAS	G148C-K150R-P158S-S195A-P199G-G202A)	
123	V198	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGGGGSGGGG SGGGGSDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP DVGSSDPLAM VGGSQARSPS YAS	Exendin4(1-39)-(GGGGS)$_3$-Fc-(GGGGS)$_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex (1-39)-L15-Fc-L15-FGF21(V103)
124	V203	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGGGGSGGGG SGGGGSDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP DVGSSDPLAM VGGSQARSPS YAS	GLP-1(7-35;A8S)-(GGGGS)-Fc-(GGGGS)$_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L5-Fc-L15-FGF21(V103)
125	V202	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGGGGSGGGG SGGGGSDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ	GLP-1(7-35;A8S)-(GGGGS)$_3$-Fc-(GGGGS)$_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L15-FGF21(V103)

		SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP DVGSSDPLAM VGGSQARSPS YAS		
126	V310	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSGGGGDS SPLLQFGGQV RQRYLYTDDA QETEAHLEIR EDGTVGGAAH QSPESLLELK ALKPGVIQIL GVKTSRFLCQ KPDGALYGSL HFDPEACSFR ELLLEDGYNV YQSEAHGLPL HLPGNRSPHC DPAPQGPARF LPLPGLPPAL PEPPGILAPQ PPDVGSSDPL AMVGPSQGRS PSYAS	Exendin4(1-39)-$(GGGGS)_3GGGG$-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG- R159Q-S195A)	Ex(1-39)-L19-FGF21(V76)-40KPEGb
127		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT GGTGGTGGTT CTGGCGGTGG CGGTTCTGGC GGCGGTGGTA GCGGTGGCGG CGGTGATAGC AGCCCGCTGC TGCAGTTTGG CGGCCAGGTG CGTCAGCGTT ATCTGTATAC CGATGATGCG CAGGAAACCG AAGCGCATCT GGAAATTCGT GAAGATGGCA CCGTGGGCGG TGCGGCGCAT CAGAGCCCGG AAAGCCTGCT GGAACTGAAA GCGCTGAAAC CGGGCGTGAT TCAGATTCTG GGCGTGAAAA CCAGCCGTTT TCTGTGCCAG AAACCGGATG GCGCGCTGTA TGGCAGCCTG CATTTTGATC CGGAAGCGTG CAGCTTTCGT GAACTGCTGC TGGAAGATGG CTATAACGTG TATCAGAGCG AAGCGCATGG CCTGCCGCTG CATCTGCCGG GCAACCGTAG CCCGCATTGC GATCCGGCAC CGCAGGGTCC GGCGCGTTTT CTGCCGCTGC CGGGTCTGCC GCCGGCACTG CCGGAACCGC CGGGTATTCT GGCCCCGCAG CCGCCGGATG TTGGTAGCAG CGATCCGCTG GCGATGGTGG GTCCGAGCCA GGGTCGTAGC CCGAGCTATG CGAGCTAA	Exendin4(1-39)-$(GGGGS)_3GGGG$-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	Ex(1-39)-L19-FGF21(V76)
31	V206	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSD KTHTCPPCPA PEAAGGPSVF LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR VVSVLTVLHQ DWLNGKEYKC KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGKDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP	Exendin4(1-39)-Fc-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L0-FGF21(V103)

		DVGSSDPLAM VGGSQARSPS YAS		
32	V208	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSD KTHTCPPCPA PEAAGGPSVF LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR VVSVLTVLHQ DWLNGKEYKC KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGKGSDS SPLLQFGGQV RQRYLYTDDA CQTEAHLEIR EDGTVGGAAD QSPESLLQLK ALKPGVIQIL GVKTSRFLCQ RPDGTLYGSL HFDPEACSFR ELLLEDGYNV YQSEAHGLPL HLPCNRSPHR DPASRGPARF LPLPGLPPAL PEPPGILAPQ PPDVGSSDPL AMVGGSQARS PSYAS	Exendin4(1-39)-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L2-FGF21(V101)
33		CACGGAGAAG GCACCTTTAC ATCGGACTTG TCGAAGCAGA TGGAGGAAGA AGCGGTGAGG CTCTTCATCG AGTGGCTCAA GAATGGAGGA CCCTCAAGCG GAGCGCCTCC TCCTTCCGAC AAAACCCATA CATGTCCGCC TTGTCCCGCA CCAGAAGCAG CGGGTGGGCC CTCGGTGTTC CTGTTCCCGC CAAAACCGAA GGACACACTT ATGATTTCAC GCACACCGGA AGTGACTTGC GTCGTGGTGG ATGTATCGCA CGAGGACCCC GAGGTCAAAT TCAACTGGTA TGTCGATGGA GTGGAGGTGC ACAATGCAAA GACCAAGCCG AGGGAAGAAC AATACAATAG CACGTACCGA GTCGTGTCCG TCTTGACGGT CCTTCACCAG GACTGGCTGA ACGGAAAGGA GTACAAGTGC AAAGTGAGCA ATAAGGCCCT CCCTGCCCCG ATTGAGAAAA CCATTTCCAA GGCCAAAGGT CAGCCTAGAG AACCTCAAGT GTATACTCTT CCGCCCTCAC GCGAAGAGAT GACGAAAAAC CAAGTGTCGC TTACGTGTCT TGTCAAAGGT TTCTACCCCT CGGACATCGC CGTAGAGTGG GAGTCGAACG GCCAGCCGGA GAACAACTAC AAGACCACGC CCCCTGTCTT GGATAGCGAC GGATCGTTTT TCCTCTACTC GAAACTCACA GTAGATAAGT CCCGATGGCA ACAGGGTAAT GTCTTTAGCT GCAGCGTGAT GCACGAGGCG CTTCACAATC ATTACACACA AAAATCACTG TCGCTTAGCC CGGGAAAGGG TTCAGATTCG TCGCCCCTGT TGCAGTTTGG TGGACAGGTC AGACAGCGCT ACCTTTACAC GGATGACGCC TGCCAGACAG AGGCACACCT CGAAATCAGA GAGGACGGTA CGGTCGGGGG TGCGGCCGAT CAGAGCCCCG AGTCGCTTCT CCAGTTGAAG GCCCTTAAGC CAGGAGTCAT CCAGATTTTG GGAGTAAAGA CCTCACGGTT TCTCTGTCAG CGTCCAGATG GGACACTGTA CGGCTCATTG	Exendin4(1-39)-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L2-FGF21(V101)

		CATTTCGATC CCGAAGCGTG CTCGTTCCGG GAGTTGCTGC TTGAGGACGG ATATAACGTC TATCAGAGCG AAGCGCATGG CCTCCCCCTT CACCTCCCGT GTAACAGGTC GCCGCATCGG GATCCGGCCT CGAGGGGTCC CGCGAGATTT CTTCCGTTGC CCGGGTTGCC TCCCGCGCTG CCCGAGCCTC CCGGGATCCT CGCGCCACAG CCTCCTGATG TAGGGTCCTC GGACCCTTTG GCGATGGTAG GTGGATCACA AGCACGGTCC CCGAGCTATG CATCA		
36	V209	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQRPDGT LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS	Exendin4(1-39)-$(GGGGS)_3$-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V101)
37		CACGGAGAAG GCACCTTTAC ATCGGACTTG TCGAAGCAGA TGGAGGAAGA AGCGGTGAGG CTCTTCATCG AGTGGCTCAA GAATGGAGGA CCCTCAAGCG GAGCGCCTCC TCCTTCCGGA GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGGTTCAG ATTCGTCGCC CCTGTTGCAG	Exendin4(1-39)-$(GGGGS)_3$-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V101)

		TTTGGTGGAC AGGTCAGACA GCGCTACCTT TACACGGATG ACGCCTGCCA GACAGAGGCA CACCTCGAAA TCAGAGAGGA CGGTACGGTC GGGGGTGCGG CCGATCAGAG CCCCGAGTCG CTTCTCCAGT TGAAGGCCCT TAAGCCAGGA GTCATCCAGA TTTTGGGAGT AAAGACCTCA CGGTTTCTCT GTCAGCGTCC AGATGGGACA CTGTACGGCT CATTGCATTT CGATCCCGAA GCGTGCTCGT TCCGGGAGTT GCTGCTTGAG GACGGATATA ACGTCTATCA GAGCGAAGCG CATGGCCTCC CCCTTCACCT CCCGTGTAAC AGGTCGCCGC ATCGGGATCC GGCCTCGAGG GGTCCCGCGA GATTTCTTCC GTTGCCCGGG TTGCCTCCCG CGCTGCCCGA GCCTCCCGGG ATCCTCGCGC CACAGCCTCC TGATGTAGGG TCCTCGGACC CTTTGGCGAT GGTAGGTGGA TCACAAGCAC GGTCCCCGAG CTATGCATCA		
133	V210	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSD KTHTCPPCPA PEAAGGPSVF LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR VVSVLTVLHQ DWLNGKEYKC KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGKGSDS SPLLQFGGQV RQRYLYTDDA CQTEAHLEIR EDGTVGGAAD QSPESLLQLK ALKPGVIQIL GVKTSRFLCQ KPDGALYGSL HFDPEACSFR ELLLEDGYNV YQSEAHGLPL HLPCNRSPHR DPASRGPARF LPLPGLPPAL PEPPGILAPQ PPDVGSSDPL AMVGGSQARS PSYAS	Exendin4(1-39-Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L2-FGF21(V103)
56		CACGGAGAAG GCACCTTTAC ATCGGACTTG TCGAAGCAGA TGGAGGAAGA AGCGGTGAGG CTCTTCATCG AGTGGCTCAA GAATGGAGGA CCCTCAAGCG GAGCGCCTCC TCCTTCCGAC AAAACCCATA CATGTCCGCC TTGTCCCGCA CCAGAAGCAG CGGGTGGGCC CTCGGTGTTC CTGTTCCCGC CAAAACCGAA GGACACACTT ATGATTTCAC GCACACCGGA AGTGACTTGC GTCGTGGTGG ATGTATCGCA CGAGGACCCC GAGGTCAAAT TCAACTGGTA TGTCGATGGA GTGGAGGTGC ACAATGCAAA GACCAAGCCG AGGGAAGAAC AATACAATAG CACGTACCGA GTCGTGTCCG TCTTGACGGT CCTTCACCAG GACTGGCTGA ACGGAAAGGA GTACAAGTGC AAAGTGAGCA ATAAGGCCCT CCCTGCCCCG ATTGAGAAAA CCATTTCCAA GGCCAAAGGT CAGCCTAGAG AACCTCAAGT GTATACTCTT CCGCCCTCAC GCGAAGAGAT GACGAAAAAC CAAGTGTCGC TTACGTGTCT TGTCAAAGGT TTCTACCCCT CGGACATCGC CGTAGAGTGG GAGTCGAACG GCCAGCCGGA GAACAACTAC AAGACCACGC CCCCTGTCTT GGATAGCGAC GGATCGTTTT TCCTCTACTC GAAACTCACA	Exendin4(1-39-Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L2-FGF21(V103)

		GTAGATAAGT CCCGATGGCA ACAGGGTAAT GTCTTTAGCT GCAGCGTGAT GCACGAGGCG CTTCACAATC ATTACACACA AAAATCACTG TCGCTTAGCC CGGGAAAGGG TTCAGATTCG TCGCCCCTGT TGCAGTTTGG TGGACAGGTC AGACAGCGCT ACCTTTACAC GGATGACGCC TGCCAGACAG AGGCACACCT CGAAATCAGA GAGGACGGTA CGGTCGGGGG TGCGGCCGAT CAGAGCCCCG AGTCGCTTCT CCAGTTGAAG GCCCTTAAGC CAGGAGTCAT CCAGATTTTG GGAGTAAAGA CCTCACGGTT TCTCTGTCAG AAACCAGATG GGGCACTGTA CGGCTCATTG CATTTCGATC CCGAAGCGTG CTCGTTCCGG GAGTTGCTGC TTGAGGACGG ATATAACGTC TATCAGAGCG AAGCGCATGG CCTCCCCCTT CACCTCCCGT GTAACAGGTC GCCGCATCGG GATCCGGCCT CGAGGGGTCC CGCGAGATTT CTTCCGTTGC CCGGGTTGCC TCCCGCGCTG CCCGAGCCTC CCGGGATCCT CGCGCCACAG CCTCCTGATG TAGGGTCCTC GGACCCTTTG GCGATGGTAG GTGGATCACA AGCACGGTCC CCGAGCTATG CATCA		
134	V211	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS	Exendin4(1-39)-$(GGGGS)_3$Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103)
57		CACGGAGAAG GCACCTTTAC ATCGGACTTG TCGAAGCAGA TGGAGGAAGA AGCGGTGAGG CTCTTCATCG AGTGGCTCAA GAATGGAGGA CCCTCAAGCG GAGCGCCTCC TCCTTCCGGA GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT	Exendin4(1-39)-$(GGGGS)_3$Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103)

		TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGGTTCAG ATTCGTCGCC CCTGTTGCAG TTTGGTGGAC AGGTCAGACA GCGCTACCTT TACACGGATG ACGCCTGCCA GACAGAGGCA CACCTCGAAA TCAGAGAGGA CGGTACGGTC GGGGGTGCGG CCGATCAGAG CCCCGAGTCG CTTCTCCAGT TGAAGGCCCT TAAGCCAGGA GTCATCCAGA TTTTGGGAGT AAAGACCTCA CGGTTTCTCT GTCAGAAACC AGATGGGGCA CTGTACGGCT CATTGCATTT CGATCCCGAA GCGTGCTCGT TCCGGGAGTT GCTGCTTGAG GACGGATATA ACGTCTATCA GAGCGAAGCG CATGGCCTCC CCCTTCACCT CCCGTGTAAC AGGTCGCCGC ATCGGGATCC GGCCTCGAGG GGTCCCGCGA GATTTCTTCC GTTGCCCGGG TTGCCTCCCG CGCTGCCCGA GCCTCCCGGG ATCCTCGCGC CACAGCCTCC TGATGTAGGG TCCTCGGACC CTTTGGCGAT GGTAGGTGGA TCACAAGCAC GGTCCCCGAG CTATGCATCA		
135	V214	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQRPDGT LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS	GLP-1(7-35;A8S)-$(GGGGS)_3$-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L2-FGF21(V101)
64		CACTCCGAAG GAACATTCAC TTCCGATGTA AGCTCGTATT TGGAAGGGCA GGCGGCTAAG GAGTTTATCG CATGGTTGGT CAAAGGTGGT GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC	GLP-1(7-35;A8S)-$(GGGGS)_3$-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-	GLP-1(A8S)-L15-Fc-L2-FGF21(V101)

		TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGGTTCAG ATTCGTCGCC CCTGTTGCAG TTTGGTGGAC AGGTCAGACA GCGCTACCTT TACACGGATG ACGCCTGCCA GACAGAGGCA CACCTCGAAA TCAGAGAGGA CGGTACGGTC GGGGGTGCGG CCGATCAGAG CCCCGAGTCG CTTCTCCAGT TGAAGGCCCT TAAGCCAGGA GTCATCCAGA TTTTGGGAGT AAAGACCTCA CGGTTTCTCT GTCAGCGTCC AGATGGGACA CTGTACGGCT CATTGCATTT CGATCCCGAA GCGTGCTCGT TCCGGGAGTT GCTGCTTGAG GACGGATATA ACGTCTATCA GAGCGAAGCG CATGGCCTCC CCCTTCACCT CCCGTGTAAC AGGTCGCCGC ATCGGGATCC GGCCTCGAGG GGTCCCGCGA GATTTCTTCC GTTGCCCGGG TTGCCTCCCG CGCTGCCCGA GCCTCCCGGG ATCCTCGCGC CACAGCCTCC TGATGTAGGG TCCTCGGACC CTTTGGCGAT GGTAGGTGGA TCACAAGCAC GGTCCCCGAG CTATGCATCA	S195A- P199G- G202A)	
136	V216	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS	GLP-1(7-35;A8S)-(GGGGS)$_3$-Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L2-FGF21(V103)
65		CACTCCGAAG GAACATTCAC TTCCGATGTA AGCTCGTATT TGGAAGGGCA GGCGGCTAAG GAGTTTATCG CATGGTTGGT CAAAGGTGGT GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT	GLP-1(7-35;A8S)-(GGGGS)$_3$-Fc-GS-FGF21(33-	GLP-1(A8S)-L15-Fc-L2-FGF21(V

		CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGGTTCAG ATTCGTCGCC CCTGTTGCAG TTTGGTGGAC AGGTCAGACA GCGCTACCTT TACACGGATG ACGCCTGCCA GACAGAGGCA CACCTCGAAA TCAGAGAGGA CGGTACGGTC GGGGGTGCGG CCGATCAGAG CCCCGAGTCG CTTCTCCAGT TGAAGGCCCT TAAGCCAGGA GTCATCCAGA TTTTGGGAGT AAAGACCTCA CGGTTTCTCT GTCAGAAACC AGATGGGGCA CTGTACGGCT CATTGCATTT CGATCCCGAA GCGTGCTCGT TCCGGGAGTT GCTGCTTGAG GACGGATATA ACGTCTATCA GAGCGAAGCG CATGGCCTCC CCCTTCACCT CCCGTGTAAC AGGTCGCCGC ATCGGGATCC GGCCTCGAGG GGTCCCGCGA GATTTCTTCC GTTGCCCGGG TTGCCTCCCG CGCTGCCCGA GCCTCCCGGG ATCCTCGCGC CACAGCCTCC TGATGTAGGG TCCTCGGACC CTTTGGCGAT GGTAGGTGGA TCACAAGCAC GGTCCCCGAG CTATGCATCA	209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	103)
137	V218	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KDSSPLLQFG GQVRQRYLYT DDACQTEAHL EIREDGTVGG AADQSPESLL QLKALKPGVI QILGVKTSRF LCQKPDGALY GSLHFDPEAC SFRELLLEDG YNVYQSEAHG LPLHLPCNRS PHRDPASRGP ARFLPLPGLP PALPEPPGIL APQPPDVGSS DPLAMVGGSQ ARSPSYAS	GLP-1(7-35;A8S)-$(GGGGS)_3$-Fc-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L0-FGF21(V103)

106		CACTCCGAAG GAACATTCAC TTCCGATGTA AGCTCGTATT TGGAAGGGCA GGCGGCTAAG GAGTTTATCG CATGGTTGGT CAAAGGTGGT GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGATTCGT CGCCCCTGTT GCAGTTTGGT GGACAGGTCA GACAGCGCTA CCTTTACACG GATGACGCCT GCCAGACAGA GGCACACCTC GAAATCAGAG AGGACGGTAC GGTCGGGGGT GCGGCCGATC AGAGCCCCGA GTCGCTTCTC CAGTTGAAGG CCCTTAAGCC AGGAGTCATC CAGATTTTGG GAGTAAAGAC CTCACGGTTT CTCTGTCAGA AACCAGATGG GGCACTGTAC GGCTCATTGC ATTTCGATCC CGAAGCGTGC TCGTTCCGGG AGTTGCTGCT TGAGGACGGA TATAACGTCT ATCAGAGCGA AGCGCATGGC CTCCCCCTTC ACCTCCCGTG TAACAGGTCG CCGCATCGGG ATCCGGCCTC GAGGGGTCCC GCGAGATTTC TTCCGTTGCC CGGGTTGCCT CCCGCGCTGC CCGAGCCTCC CGGGATCCTC GCGCCACAGC CTCCTGATGT AGGGTCCTCG GACCCTTTGG CGATGGTAGG TGGATCACAA GCACGGTCCC CGAGCTATGC ATCA	GLP-1(7-35;A8S)-(GGGGS)$_3$-Fc-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L0-FGF21(V103)
107	V200	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG K	Exendin4(1-39)-(GGGGS)$_3$-Fc	Ex(1-39)-L15-Fc

139	V201	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG K	Exendin4(1-39)-GGGGS-Fc	Ex(1-39)-L5-Fc
140	V207	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSD KTHTCPPCPA PEAAGGPSVF LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR VVSVLTVLHQ DWLNGKEYKC KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGKGGGG SGGGGSGGGG SDSSPLLQFG GQVRQRYLYT DDACQTEAHL EIREDGTVGG AADQSPESLL QLKALKPGVI QILGVKTSRF LCQKPDGALY GSLHFDPEAC SFRELLLEDG YNVYQSEAHG LPLHLPCNRS PHRDPASRGP ARFLPLPGLP PALPEPPGIL APQPPDVGSS DPLAMVGGSQ ARSPSYAS	Exendin4(1-39)-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L15-FGF21(V103)
141	V212	HSEGTFTSDV SSYLEGQAAK EFIAWLVKDK THTCPPCPAP EAAGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE VKFNWYVDGV EVHNAKTKPR EEQYNSTYRV VSVLTVLHQD WLNGKEYKCK VSNKALPAPI EKTISKAKGQ PREPQVYTLP PSREEMTKNQ VSLTCLVKGF YPSDIAVEWE SNGQPENNYK TTPPVLDSDG SFFLYSKLTV DKSRWQQGNV FSCSVMHEAL HNHYTQKSLS LSPGKGGGGS GGGGSGGGGS DSSPLLQFGG QVRQRYLYTD DACQTEAHLE IREDGTVGGA ADQSPESLLQ LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPCNRSP HRDPASRGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGGSQA RSPSYAS	GLP-1(7-35;A8S)-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L0-Fc-L15-FGF21(V103)
142	V213	HSEGTFTSDV SSYLEGQAAK EFIAWLVKDK THTCPPCPAP EAAGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE VKFNWYVDGV EVHNAKTKPR EEQYNSTYRV VSVLTVLHQD WLNGKEYKCK VSNKALPAPI EKTISKAKGQ PREPQVYTLP PSREEMTKNQ VSLTCLVKGF YPSDIAVEWE SNGQPENNYK TTPPVLDSDG SFFLYSKLTV DKSRWQQGNV FSCSVMHEAL HNHYTQKSLS LSPGKGSDSS PLLQFGGQVR QRYLYTDDAC QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQR PDGTLYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPCNRSPHRD PASRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLA	GLP-1(7-35;A8S)-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L0-Fc-L2-FGF21(V101)

		MVGGSQARSP SYAS		
143	V215	HSEGTFTSDV SSYLEGQAAK EFIAWLVKDK THTCPPCPAP EAAGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE VKFNWYVDGV EVHNAKTKPR EEQYNSTYRV VSVLTVLHQD WLNGKEYKCK VSNKALPAPI EKTISKAKGQ PREPQVYTLP PSREEMTKNQ VSLTCLVKGF YPSDIAVEWE SNGQPENNYK TTPPVLDSDG SFFLYSKLTV DKSRWQQGNV FSCSVMHEAL HNHYTQKSLS LSPGKGSDSS PLLQFGGQVR QRYLYTDDAC QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQK PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPCNRSPHRD PASRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLA MVGGSQARSP SYAS	GLP-1(7-35;A8S)-Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L0-Fc-L2-FGF21(V103)
144	V217	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSTDTLLL WVLLLWVPGS TGHGEGTFTS DLSKQMEEEA VRLFIEWLKN GGPSSGAPPP SGGGGSGGGG SGGGGSDKTH TCPPCPAPEA AGGPSVFLFP PKPKDTLMIS RTPEVTCVVV DVSHEDPEVK FNWYVDGVEV HNAKTKPREE QYNSTYRVVS VLTVLHQDWL NGKEYKCKVS NKALPAPIEK TISKAKGQPR EPQVYTLPPS REEMTKNQVS LTCLVKGFYP SDIAVEWESN GQPENNYKTT PPVLDSDGSF FLYSKLTVDK SRWQQGNVFS CSVMHEALHN HYTQKSLSLS PGKGSDSSPL LQFGGQVRQR YLYTDDACQT EAHLEIREDG TVGGAADQSP ESLLQLKALK PGVIQILGVK TSRFLCQKPD GALYGSLHFD PEACSFRELL LEDGYNVYQS EAHGLPLHLP CNRSPHRDPA SRGPARFLPL PGLPPALPEP PGILAPQPPD VGSSDPLAMV GGSQARSPSY AS	Exendin4(1-39)-$(GGGGS)_3$-Exendin4(1-39)-$(GGGGS)_3$-Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Ex(1-39)-L15-Fc-L2-FGF21(V103)
145	V278	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-bis-maleimide dimer 40 kDa branched PEG-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76) Dimer-40KPEGb

146	V279	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-bis-maleimide dimer 20 kDa branched PEG-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76) Dimer-20KPEGb
147	V280	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76)
148	V283	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGG SGGGGSGGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS	Exendin4(1-30)-S(GGGGS)$_3$GGG G-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	Ex(1-30_-L20-FGF21(V76)
149	V284	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGG SGGGGSGGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLS	Exendin4(1-30)-S(GGGGS)$_3$GGG G-FGF21(33-195;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q)	Ex(1-30)-L20-FGF21(V76;CΔ14)-40KPEGb
150	V285	DLSKQMEEEA VRLFIEWLKN GGSGGGGSGG GGSGGGGSGG GGDSSPLLQF GGQVRQRYLY TDDAQETEAH LEIREDGTVG GAAHQSPESL LELKALKPGV IQILGVKTSR FLCQKPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNR SPHCDPAPQG PARFLPLPGL	Exendin4(9-30)-S(GGGGS)$_3$GGG G-FGF21(33-209;Q56E-D74H-Q82E-	Ex(9-30)-L20-FGF21(V76)-40KPEGb

		PPALPEPPGI LAPQPPDVGS SDPLAMVGPS QGRSPSYAS	R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A)	
151	V289	GGQVRQRYLY TDDAQETEAH LEIREDGTVG GAAHQSPESL LELKALKPGV IQILGVKTSR FLCQKPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNR SPHCDPAPQG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLAMVGPS QGRSPSYAS	FGF21(42-209; Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A)	FGF21(V76;NΔ9)-40KPEGb
152	V290	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(42-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76;NΔ9)-40KPEGb
153	V291	GVSTSEAKFE QDSAILWYGV EFAKLHTSGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S scramble)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A)	GLP-1(A8S;scramble)-L10-FGF21(V76)-40KPEGb
154	V311	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGGSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched	GLP-1(A8S)-L10-FGF21(V76-P199G)-40KPEGb

			PEG-R159Q-S195A-P199G)	
155		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAATT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT TGCGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTGGTAG CCAGGGTCGT AGCCCGAGCT ATGCGAGC	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A-P199G)	GLP-1(A8S)-L10-FGF21(V76-P199G)-40KPEGb
156	V312	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGG SGGGGSGGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGGSQG RSPSYAS	Exendin4(1-30)-S(GGGGS)$_3$GGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-S195A-P199G)	Ex(1-30)-L20-FGF21(V76-P199G)-40KPEGb
157		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGGT TCTGGCGGCG GTGGTAGCGG TGGCGGCGGT GACTCGAGCC CGCTGCTGCA GTTTGGCGGC CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGG AAACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGCATCAGA GCCCGGAAAG CCTGCTGGAA CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGAAAC CGGATGGCGC GCTGTATGGC	Exendin4(1-30)-S(GGGGS)$_3$GGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa branched PEG-R159Q-	Ex(1-30)-L20-FGF21(V76-P199G)-40KPEGb

		AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CCGTAGCCCG CATTGCGATC CGGCACCGCA GGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGGCGA TGGTGGGTGG TAGCCAGGGT CGTAGCCCGA GCTATGCGAG C	S195A-P199G)	
158	V313	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYA	Exendin4(1-39)-$(GGGGS)_3$-Fc-GS-FGF21 (33-208; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103;CΔ1)
159	V314	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSY	Exendin4(1-39)-$(GGGGS)_3$-Fc-GS-FGF21 (33-207; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103;CΔ2)
160	V315	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA	Exendin4(1-39)-$(GGGGS)_3$-Fc-GS-FGF21 (33-206; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103;CΔ3)

		HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPS		
161	V316	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS P	Exendin4(1-39)-$(GGGGS)_3$-Fc-GS-FGF21 (33-209; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A)-P	Ex(1-39)-L15-Fc-L2-FGF21(V103)+P210
162	V317	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAA P	Exendin4(1-39)-$(GGGGS)_3$-Fc-GS-FGF21 (33-209; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A-S209A)-p	Ex(1-39)-L15-Fc-L2-FGF21(V103-S209A)+P210
163	V225	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGGSG DSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIR EDGTVGGAADQSPESLLQLKALKPGVIQILGV KTSRFLCQRPDGALYGSLHFDPEACSFRELLL EDGYNVYQSEAHGLPLHLPGNKSPHCDPAPRG PARFLPLPGLPPALPEPPGILAPQPPDVGSSD PLSMVGPSQGRSPSYAS	GLP-1(7-35)-SG-FGF21(33-209;R154C)	GLP-1-L2-FGF21(154C)
164	V226	HSEGTFTSDVSSYLEGQAAKEFIAWLVKGGSG DSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIR EDGTVGGAADQSPESLLQLKALKPGVIQILGV KTSRFLCQRPDGALYGSLHFDPEACSFRELLL EDGYNVYQSEAHGLPLHLPGNKSPHCDPAPRG PARFLPLPGLPPALPEPPGILAPQPPDVGSSD PLSMVGPSQGRSPSYAS	GLP-1(7-35;A8S)-SG-FGF21(33-209;R154C)	GLP-1(A8S)-L2-FGF21(154C)
165	V229	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGGSG DSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIR EDGTVGGAADQSPESLLQLKALKPGVIQILGV KTSRFLCQRPDGALYGSLHFDPEACSFRELLL	GLP-1(7-35)-SG-FGF21(33-209;R154C-	GLP-1-L2-FGF21(154C)-

		EDGYNVYQSEAHGLPLHLPGNKSPHCDPAPRG PARFLPLPGLPPALPEPPGILAPQPPDVGSSD PLSMVGPSQGRSPSYAS	40 kDa linear PEG)	40KPEGl
166	V230	HSEGTFTSDVSSYLEGQAAKEFIAWLVKGGSG DSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIR EDGTVGGAADQSPESLLQLKALKPGVIQILGV KTSRFLCQRPDGALYGSLHFDPEACSFRELLL EDGYNVYQSEAHGLPLHLPGNKSPHCDPAPRG PARFLPLPGLPPALPEPPGILAPQPPDVGSSD PLSMVGPSQGRSPSYAS	GLP-1(7-35; A8S)-SG-FGF21(33-209;R154C-40 kDa linear PEG)	GLP-1(A8S)-L2-FGF21(154C)-40KPEGl
167	V231	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGGSG GGGSGGGDSSPLLQFGGQVRQRYLYTDDAQQT EAHLEIREDGTVGGAADQSPESLLQLKALKPG VIQILGVKTSRFLCQRPDGALYGSLHFDPEAC SFRELLLEDGYNVYQSEAHGLPLHLPGNKSPH CDPAPRGPARFLPLPGLPPALPEPPGILAPQP PDVGSSDPLSMVGPSQGRSPSYAS	GLP-1(7-35)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa linear PEG)	GLP-1-L8-FGF21(154C)-40KPEGl
168	V232	HSEGTFTSDVSSYLEGQAAKEFIAWLVKGGSG GGGSGGGDSSPLLQFGGQVRQRYLYTDDAQQT EAHLEIREDGTVGGAADQSPESLLQLKALKPG VIQILGVKTSRFLCQRPDGALYGSLHFDPEAC SFRELLLEDGYNVYQSEAHGLPLHLPGNKSPH CDPAPRGPARFLPLPGLPPALPEPPGILAPQP PDVGSSDPLSMVGPSQGRSPSYAS	GLP-1(7-35; A8S)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa linear PEG)	GLP-1(A8S)-L8-FGF21(154C)-40KPEGl
169	V237	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGGSG DSSPLLQFGGQVRQRYLYTDDAQQTEAHLEIR EDGTVGGAADQSPESLLQLKALKPGVIQILGV KTSRFLCQRPDGALYGSLHFDPEACSFRELLL EDGYNVYQSEAHGLPLHLPGNKSPHCDPAPRG PARFLPLPGLPPALPEPPGILAPQPPDVGSSD PLSMVGPSQGRSPSYAS	GLP-1(7-35)-SG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1-L2-FGF21(154C)-40KPEGb
170	V238	GDSSPLLQFGGQVRQRYLYTDDAQQTEAHLEI REDGTVGGAADQSPESLLQLKALKPGVIQILG VKTSRFLCQRPDGALYGSLHFDPEACSFRELL LEDGYNVYQSEAHGLPLHLPGNKSPHCDPAPR GPARFLPLPGLPPALPEPPGILAPQPPDVGSS DPLSMVGPSQGRSPSYAS	G-FGF21(33-209;R154C-40 kDa branched PEG)	FGF21(R154C)-40KPEGb
171	V253	HSEGTFTSDVSSYLEGQAAKEFIAWLVKGGSG GGGSGGGDSSPLLQFGGQVRQRYLYTDDAQQT EAHLEIREDGTVGGAADQSPESLLQLKALKPG VIQILGVKTSRFLCQRPDGALYGSLHFDPEAC SFRELLLEDGYNVYQSEAHGLPLHLPGNKSPH CDPAPRGPARFLPLPGLPPALPEPPGILAPQP PDVGSSDPLS	GLP-1(7-35; A8S)-SGGGGSGGG-FGF21(33-194;R154C-40 kDa branched PEG)	GLP-1(A8S)-L8-FGF21-(CΔ14;154C)-40KPEGb
172	V258	HGEGTFTSDVSSYLEGQAAKEFIAWLVKGGSG GGGSGGGDSSPLLQFGGQVRQRYLYTDDAQQT EAHLEIREDGTVGGAADQSPESLLQLKALKPG VIQILGVKTSRFLCQRPDGALYGSLHFDPEAC SFRELLLEDGYNVYQSEAHGLPLHLPGNKSPH CDPAPRGPARFLPLPGLPPALPEPPGILAPQP PDVGSSDPLSMVGPSQGRSPSYAS	GLP-1(7-35; A8G)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa branched PEG)	GLP-1(A8G)-L8-FGF21(154C)-40KPEGb

[00051] The variants or mutants used in the proteins of the invention, e.g., variants of wild-type FGF21, GLP-1, and/or Exendin-4 feature at least one substituted, added, and/or removed amino acid relative to the wild-type protein. Additionally, the variants may include N-and/or C-terminal truncations relative to the wild-type proteins. Generally speaking, a variant possesses some modified property, structural or functional, of the wild-type protein. For example, the variant may have enhanced or improved physical stability in concentrated solutions (e.g., less hydrophobic mediated aggregation), enhanced or improved plasma stability when incubated with blood plasma or reduced risk for immunogenicity or enhanced or improved bioactivity while maintaining a favorable bioactivity profile.

[00052] Acceptable amino acid substitutions and modifications which constitute differences between the portions of the proteins of the invention and their wild-type comparator proteins include, but are not limited to, one or more amino acid substitutions, including substitutions with pyrrolysine, pyrroline-carbyoxy-lysine (Pcl) and non-naturally occurring amino acid analogs, and truncations. Thus, the proteins of the invention (e.g., the fusion proteins of the invention) include, but are not limited to, site-directed mutants, truncated polypeptides, proteolysis-resistant mutants, aggregation-reducing mutants, combination mutants, and fusion proteins, as described herein.

[00053] One skilled in the art of expression of proteins will recognize that methionine or methionine-arginine sequence can be introduced at the N-terminus of any of the proteins of the invention, for expression in E. coli, and are contemplated within the context of this invention.

[00054] One skilled in the art of expression of proteins will recognize that additional tags or fusion domains for the purposes of modulating expression levels, purification, or stabilization can be introduced to the N-terminus of any of the proteins of the invention, with or without an additional peptide to target digestion by a specific protease for later removal of that tag or fusion domain, for expression in any host cell, and are contemplated within the context of this invention.

[00055] One skilled in the art of expression of proteins will recognize that leader peptides targeting the expressed protein to the periplasm or extracellular space can be introduced at the N-terminus of any of the proteins of the invention, for expression in E. coli or other bacterial hosts, and are contemplated within the context of this invention.

[00056] One skilled in the art of expression of proteins will recognize that leader peptides targeting the expressed protein to the ER, secretory vesicles, or extracellular space can be introduced at the N-terminus of any of the proteins of the invention, for

expression in eukaryotic host cells, and are contemplated within the context of this invention.

[00057] The proteins of the invention may possess increased compatibility with pharmaceutical preservatives (e.g., m-cresol, phenol, benzyl alcohol), thus enabling the preparation of a preserved pharmaceutical formulation that maintains the physiochemical properties and biological activity of the protein during storage. Accordingly, variants with enhanced pharmaceutical stability relative to wild-type, have improved physical stability in concentrated solutions under both physiological and preserved pharmaceutical formulation conditions, while maintaining biological potency. By way of non-limiting example, the proteins of the invention may be more resistant to proteolysis and enzymatic degradation; may have improved stability; and may be less likely to aggregate, than their wild-type counterparts or corresponding native sequence. As used herein, these terms are not mutually exclusive or limiting, it being entirely possible that a given variant has one or more modified properties of the wild-type protein.

[00058] The invention also encompasses nucleic acid molecules encoding the proteins of the invention, comprising, for example, an FGF21 amino acid sequence that is at least about 95% identical to the amino acid sequence of SEQ ID NO:3, but wherein specific residues conferring a desirable property to the FGF21 protein variant, e.g., improved potency to FGF21-receptors, proteolysis-resistance, increased half-life or aggregation-reducing properties and combinations thereof have not been further modified. In other words, with the exception of residues in the FGF21 mutant sequence that have been modified in order to confer proteolysis-resistance, aggregation-reducing, or other properties, about 5% (alternately 4%, alternately 3%, alternately 2%, alternately 1%) of all other amino acid residues in the FGF21 mutant sequence can be modified. Such FGF21 mutants possess at least one activity of the wild-type FGF21 polypeptide.

[00059] Similarly, the invention also comprises nucleic acid molecules encoding the GLP-1 and Exendin-4 portions of the molecule, whose amino acid sequences are at least about 85%, identical, and more preferably at least about 90 to 95% identical, to the amino acid sequence of SEQ ID NO:30 and 7, respectively, but wherein specific residues conferring a desirable property to the dual function protein variant, e.g., proteolysis-resistance, increased half-life or aggregation-reducing properties and combinations thereof have not been further modified.

[00060] The invention also encompasses a nucleic acid molecule comprising a nucleotide sequence that is at least about 85%, identical, and more preferably, at least about 90 to 95% identical to the nucleotide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, and the cDNA sequence encoding wild-type Exendin-4, but wherein the

nucleotides encoding amino acid residues conferring the encoded protein's proteolysis-resistance, aggregation-reducing or other properties have not been further modified. In other words, with the exception of nucleotides that encode residues in the FGF21, GLP-1, or Exendin-4 mutant sequences that have been modified in order to confer proteolysis-resistance, aggregation-reducing, or other properties, about 15%, and more preferably about 10 to 5% of all other nucleotides in the mutant sequence can be modified. Such nucleic acid molecules encode proteins possessing at least one activity of their wild-type counterparts.

[00061] Provided herein are methods used to generate the proteins of the invention, wherein such methods involve site-specific modification and non-site-specific modification of the wild-type versions of the proteins (e.g., the FGF21 wild-type protein as described herein), e.g., truncations of the wild-type proteins, and the site-specific incorporation of amino acids at positions of interest within the wild-type proteins. Said modifications enhance the biological properties of the proteins of the invention relative to the wild-type proteins, as well as, in some cases, serving as points of attachment for, e.g., labels and protein half-life extension agents, and for purposes of affixing said variants to the surface of a solid support. Related embodiments of the invention are methods of producing cells capable of producing said dual function proteins of the invention, and of producing vectors containing DNA encoding said variants.

[00062] In certain embodiments, such modifications, e.g., site-specific modifications, are used to attach conjugates, e.g., PEG groups to proteins, polypeptides, and/or peptides of the invention, for purposes of, e.g., extending half-life or otherwise improving the biological properties of said proteins, polypeptides, and/or peptides. Said techniques are described further herein.

[00063] In other embodiments, such modifications, e.g., site-specific modifications are used to attach other polymers, small molecules and recombinant protein sequences that extend half-life of the protein of the invention. One such embodiment includes the attachment of fatty acids or specific albumin binding compounds to proteins, polypeptides, and/or peptides. In other embodiments, the modifications are made at a particular amino acid type and may be attached at one or more sites on the protein.

[00064] In other embodiments, such modifications, e.g., site-specific modifications are used as means of attachment for the production of wild-type and/or variant multimers, e.g., dimers (homodimers or heterodimers) or trimers or tetramers. These multimeric protein molecules may additionally have groups such as PEG, sugars, and/or PEG-cholesterol conjugates attached or be fused either amino-terminally or carboxy-terminally to other proteins such as Fc, Human Serum Albumin (HSA), etc.

[00065] In other embodiments, such site-specific modifications are used to produce proteins, polypeptides and/or peptides wherein the position of the site-specifically incorporated pyrrolysine, pyrroline-carboxy-lysine, or pyrrolysine analogue or non-naturally occurring amino acids (para-acetyl-Phe, para-azido-Phe) allows for controlled orientation and attachment of such proteins, polypeptides and/or peptides onto a surface of a solid support or to have groups such as PEG, sugars and/or PEG-cholesterol conjugates attached.

[00066] In other embodiments, such site-specific modifications are used to site-specifically cross-link proteins, polypeptides and/or peptides thereby forming hetero-oligomers including, but not limited to, heterodimers and heterotrimers. In other embodiments, such site-specific modifications are used to site-specifically cross-link proteins, polypeptides and/or peptides thereby forming protein-protein conjugates, protein-polypeptide conjugates, protein-peptide conjugates, polypeptide-polypeptide conjugates, polypeptide-peptide conjugates or peptide-peptide conjugates. In other embodiments, a site specific modification may include a branching point to allow more than one type of molecule to be attached at a single site of a protein, polypeptide or peptide.

[00067] In other embodiments, the modifications listed herein can be done in a non-site-specific manner and result in protein-protein conjugates, protein-polypeptide conjugates, protein-peptide conjugates, polypeptide-polypeptide conjugates, polypeptide-peptide conjugates or peptide-peptide conjugates of the invention.

Definitions

[00068] Various definitions are used throughout this document. Most words have the meaning that would be attributed to those words by one skilled in the art. Words specifically defined either below or elsewhere in this document have the meaning provided in the context of the present invention as a whole and as are typically understood by those skilled in the art.

[00069] As used herein, the term “FGF21” refers to a member of the fibroblast growth factor (FGF) protein family. An amino acid sequence of FGF21 (GenBank Accession No. NP_061986.1) is set forth as SEQ ID NO:1, the corresponding polynucleotide sequence of which is set forth as SEQ ID NO:2 (NCBI reference sequence number NM_019113.2). “FGF21 variant,” “FGF21 mutant,” and similar terms describe modified version of the FGF21 protein, e.g., with constituent amino acid residues deleted, added, modified, or substituted.

[00070] As used herein, the term "FGF21 receptor" refers to a receptor for FGF21 (Kharitonenkov,A, et al. (2008) Journal of Cellular Physiology 215:1-7; Kurosu,Het al. (2007) JBC 282:26687-26695; Ogawa, Yet al. (2007) PNAS 104:7432-7437).

[00071] The term "FGF21 polypeptide" refers to a naturally-occurring polypeptide expressed in humans. For purposes of this disclosure, the term "FGF21 polypeptide" can be used interchangeably to refer to any full-length FGF21 polypeptide, e.g., SEQ ID NO:1, which consists of 209 amino acid residues and which is encoded by the nucleotide sequence of SEQ ID NO:2; any mature form of the polypeptide, which consists of 181 amino acid residues, and in which the 28 amino acid residues at the amino-terminal end of the full-length FGF21 polypeptide (i.e., which constitute the signal peptide) have been removed.

[00072] "Variant 76," or "V76," as used herein, is an FGF21 protein variant, featuring a 40 kDa branched PEG linked through Cys154, and eight point mutations relative to the 177 amino acid wild-type protein. Synthesis of the variant is described in greater detail herein.

[00073] "Variant 101," or "V101," as used herein, is an FGF21 protein variant, featuring an engineered disulfide bridge, and eight point mutations relative to the 177 amino acid wild-type protein, expressed as a fusion to human IgG1 Fc-domain with a GS linker. Synthesis of the variant is described in greater detail herein.

[00074] "Variant 103," or "V103," as used herein, is an FGF21 protein variant, featuring an engineered disulfide bridge, and eight point mutations relative to the 177 amino acid wild-type protein, expressed as a fusion to human IgG1 Fc-domain with a GS linker. Synthesis of the variant is described in greater detail herein.

[00075] "Variant 188," or "V188," as used herein, is an FGF21 protein variant, featuring an engineered disulfide bridge, and eight point mutations relative to the 177 amino acid wild-type protein, expressed as a fusion to human IgG1 Fc-domain with a $(SGGGG)_3$ linker. Synthesis of the variant is described in greater detail herein

[00076] "GLP-1-FGF21-PEG dual agonists," "dual function proteins," "dual function fusion proteins," "dual activity proteins," "fusion products," "dual FGF21 receptor agonist and GLP-1 receptor agonist," "dual FGF21 receptor agonist and GLP-1 receptor agonist proteins of the invention," "GLP-1-FGF21 fusion proteins," "fusion proteins of the invention," "fusions of the invention," and similar terms define protein or polypeptide fusions comprising at least an FGF21 polypeptide or protein variant, mutant, or truncated version, fused or linked to another metabolic regulator such as GLP-1 or Exendin-4. They comprise a single molecule with dual activity or dual function vis-à-vis the receptors of their respective constituents, i.e., they show the ability to agonize the FGF21 receptor and the GLP-1 receptor. The constituent sequences of said fusion

proteins may comprise modifications, truncations, other variants of naturally occurring (i.e., wild-type) protein or polypeptide counterparts, and may employ any number of various other modifications, e.g., PEG groups for half-life extension.

[00077] A particularly preferred embodiment of the GLP-1-FGF21-PEG fusion protein of the invention incorporates V76 (as defined herein) as the FGF21 variant. Said preferred embodiment is also referred to herein as "GLP-1(A8S)-FGF21-PEG" and features a substitution from alanine to serine at position 8 relative to the wild-type GLP-1 sequence (SEQ ID NO:5129) and substitution from arginine to cysteine at position 154 relative to the wild-type FGF21 sequence (SEQ ID:1). The sequence of GLP-1(A8S)-FGF21(V76)-PEG is as follows (SEQ ID NO:9).

```
1   HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD
61  AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS
121 LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA
181 LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS
```

[00078] The term "isolated nucleic acid molecule" refers to a nucleic acid molecule of the present invention that (1) has been separated from at least about 50 percent of proteins, lipids, carbohydrates, or other materials with which it is naturally found when total nucleic acid is isolated from the source cells, (2) is not linked to all or a portion of a polynucleotide to which the "isolated nucleic acid molecule" is linked in nature, (3) is operably linked to a polynucleotide which it is not linked to in nature, or (4) does not occur in nature as part of a larger polynucleotide sequence. Preferably, the isolated nucleic acid molecule of the present invention is substantially free from any other contaminating nucleic acid molecules or other contaminants that are found in its natural environment that would interfere with its use in polypeptide production or its therapeutic, diagnostic, prophylactic or research use.

[00079] The term "vector" is used to refer to any molecule (e.g., nucleic acid, plasmid, or virus) used to transfer coding information to a host cell.

[00080] The term "expression vector" refers to a vector that is suitable for transformation of a host cell and contains nucleic acid sequences that direct and/or control the expression of inserted heterologous nucleic acid sequences. Expression includes, but is not limited to, processes such as transcription, translation, and RNA splicing, if introns are present.

[00081] The term "operably linked" is used herein to refer to an arrangement of flanking sequences wherein the flanking sequences so described are configured or assembled so as to perform their usual function. Elements of fusions proteins may be

operably linked to one another so as to allow the fusion protein to function as if it were a naturally occurring, endogenous protein, and/or to combine disparate elements of said fusion proteins in a synergistic fashion.

[00082] On a nucleotide level, a flanking sequence operably linked to a coding sequence may be capable of effecting the replication, transcription and/or translation of the coding sequence. For example, a coding sequence is operably linked to a promoter when the promoter is capable of directing transcription of that coding sequence. A flanking sequence need not be contiguous with the coding sequence, so long as it functions correctly. Thus, for example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

[00083] The term "host cell" is used to refer to a cell which has been transformed, or is capable of being transformed with a nucleic acid sequence and then of expressing a selected gene of interest. The term includes the progeny of the parent cell, whether or not the progeny is identical in morphology or in genetic make-up to the original parent, so long as the selected gene is present.

[00084] The term "amino acid," as used herein, refers to naturally occurring amino acids, unnatural amino acids, amino acid analogues and amino acid mimetics that function in a manner similar to the naturally occurring amino acids, all in their D and L stereoisomers if their structure allows such stereoisomeric forms. Amino acids are referred to herein by either their name, their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission.

[00085] The term "naturally occurring" when used in connection with biological materials such as nucleic acid molecules, polypeptides, host cells, and the like, refers to materials which are found in nature and are not manipulated by man. Similarly, "non-naturally occurring" as used herein refers to a material that is not found in nature or that has been structurally modified or synthesized by man. When used in connection with nucleotides, the term "naturally occurring" refers to the bases adenine (A), cytosine (C), guanine (G), thymine (T), and uracil (U). When used in connection with amino acids, the term "naturally occurring" refers to the 20 conventional amino acids (i.e., alanine (A), cysteine (C), aspartic acid (D), glutamic acid (E), phenylalanine (F), glycine (G), histidine (H), isoleucine (I), lysine (K), leucine (L), methionine (M), asparagine (N), proline (P), glutamine (Q), arginine (R), serine (S), threonine (T), valine (V), tryptophan (W), and tyrosine (Y)), as well as selenocysteine, pyrrolysine (Pyl, or O), and pyrroline-carboxy-lysine (Pcl, or Z).

[00086] Pyrrolysine (Pyl) is an amino acid naturally found within methylamine methyltransferases of methanogenic archaea of the family *Methanosarcina*. Pyrrolysine is a lysine analogue co-translationally incorporated at in-frame UAG codons in the respective mRNA, and it is considered the 22nd natural amino acid.

[00087] As described at least in PCT patent publication WO2010/48582 (applicant IRM, LLC), attempts to biosynthesize pyrrolysine (Pyl) in *E. coli* resulted in the formation of a "demethylated pyrrolysine," referred to herein as pyrroline-carboxy-lysine, or Pcl. "Pcl," as used herein, refers to either Pcl-A or Pcl-B.

[00088] The terms "non-natural amino acid" and "unnatural amino acid," as used herein, are interchangeably intended to represent amino acid structures that cannot be generated biosynthetically in any organism using unmodified or modified genes from any organism, whether the same or different. The terms refer to an amino acid residue that is not present in the naturally occurring (wild-type) FGF21 protein sequence or the sequences of the present invention. These include, but are not limited to, modified amino acids and/or amino acid analogues that are not one of the 20 naturally occurring amino acids, selenocysteine, pyrrolysine (Pyl), or pyrroline-carboxy-lysine (Pcl, e.g., as described in PCT patent publication WO2010/48582). Such non-natural amino acid residues can be introduced by substitution of naturally occurring amino acids, and/or by insertion of non-natural amino acids into the naturally occurring (wild-type) FGF21 protein sequence or the sequences of the invention. The non-natural amino acid residue also can be incorporated such that a desired functionality is imparted to the FGF21 molecule, for example, the ability to link a functional moiety (e.g., PEG). When used in connection with amino acids, the symbol "U" shall mean "non-natural amino acid" and "unnatural amino acid," as used herein.

[00089] In addition, it is understood that such "unnatural amino acids" typically require a modified tRNA and a modified tRNA synthetase (RS) for incorporation into a protein. These "selected" orthogonal tRNA/RS pairs are generated by a selection process as developed by Schultz et al. or by random or targeted mutation. As way of example, pyrroline-carboxy-lysine is a "natural amino acid" as it is generated biosynthetically by genes transferred from one organism into the host cells and as it is incorporated into proteins by using natural tRNA and tRNA synthetase genes, while *p*-aminophenylalanine (See, Generation of a bacterium with a 21 amino acid genetic code, Mehl RA, Anderson JC, Santoro SW, Wang L, Martin AB, King DS, Horn DM, Schultz PG. J Am Chem Soc. 2003 Jan 29;125(4):935-9) is an "unnatural amino acid" because, although generated biosynthetically, it is incorporated into proteins by a "selected" (i.e. not "naturally-occurring") orthogonal tRNA/tRNA synthetase pair.

[00090] Modified encoded amino acids include, but are not limited to, hydroxyproline, γ-carboxyglutamate, O-phosphoserine, azetidinecarboxylic acid, 2-aminoadipic acid, 3-aminoadipic acid, beta-alanine, aminopropionic acid, 2-aminobutyric acid, 4-aminobutyric acid, 6-aminocaproic acid, 2-aminoheptanoic acid, 2-aminoisobutyric acid, 3-aminoisobutyric acid, 2-aminopimelic acid, tertiary-butylglycine, 2,4-diaminoisobutyric acid, desmosine, 2,2'-diaminopimelic acid, 2,3-diaminoproprionic acid, N-ethylglycine, N-methylglycine, N-ethylasparagine, homoproline, hydroxylysine, allo-hydroxylysine, 3-hydroxyproline, 4-hydroxyproline, isodesmosine, allo-isoleucine, N-methylalanine, N-methylglycine, N-methylisoleucine, N-methylpentylglycine, N-methylvaline, naphthalanine, norvaline, norleucine, ornithine, pentylglycine, pipecolic acid and thioproline. The term "amino acid" also includes naturally occurring amino acids that are metabolites in certain organisms but are not encoded by the genetic code for incorporation into proteins. Such amino acids include, but are not limited to, ornithine, D-ornithine, and D-arginine.

[00091] The term "amino acid analogue," as used herein, refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, by way of example only, an α-carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group. Amino acid analogues include the natural and unnatural amino acids which are chemically blocked, reversibly or irreversibly, or their C-terminal carboxy group, their N-terminal amino group and/or their side-chain functional groups are chemically modified. Such analogues include, but are not limited to, methionine sulfoxide, methionine sulfone, S-(carboxymethyl)-cysteine, S-(carboxymethyl)-cysteine sulfoxide, S-(carboxymethyl)-cysteine sulfone, aspartic acid-(beta-methyl ester), N-ethylglycine, alanine carboxamide, homoserine, norleucine, and methionine methyl sulfonium.

[00092] The term "amino acid mimetics," as used herein, refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but functions in a manner similar to a naturally occurring amino acid.

[00093] The term "biologically active variant" refers to any polypeptide variant used in the dual function proteins of the invention, e.g., as a constituent protein of the fusions, that possesses an activity of its wild-type (e.g., naturally-occurring) protein or polypeptide counterpart, such as the ability to modulate blood glucose, HbA1c, insulin, triglyceride, or cholesterol levels; increase pancreatic function; reduce lipid levels in liver; reduce body weight; and to improve glucose tolerance, energy expenditure, or insulin sensitivity, regardless of the type or number of modifications that have been introduced into the polypeptide variant. Polypeptide variants possessing a somewhat

decreased level of activity relative to their wild-type versions can nonetheless be considered to be biologically active polypeptide variants. A non-limiting representative example of a biologically active polypeptide variant of the invention is an FGF21 variant, which is modified after, and possesses similar or enhanced biological properties relative to, wild-type FGF21.

[00094] The terms "effective amount" and "therapeutically effective amount" each refer to the amount of a fusion protein of the invention used to support an observable level of one or more biological activities of the wild-type polypeptide or protein counterparts, such as the ability to lower blood glucose, insulin, triglyceride or cholesterol levels; reduce liver triglyceride or lipid levels; reduce body weight; or improve glucose tolerance, energy expenditure, or insulin sensitivity. For example, a "therapeutically-effective amount" administered to a patient exhibiting, suffering, or prone to suffer from FGF21-associated disorders or GLP-1-associated disorders (such as type 1 or type 2 diabetes mellitus, obesity, or metabolic syndrome), is such an amount which induces, ameliorates or otherwise causes an improvement in the pathological symptoms, disease progression, physiological conditions associated with or resistance to succumbing to the afore mentioned disorders. For the purposes of the present invention a "subject" or "patient" is preferably a human, but can also be an animal, more specifically, a companion animal (e.g., dogs, cats, and the like), farm animals (e.g., cows, sheep, pigs, horses, and the like) and laboratory animals (e.g., rats, mice, guinea pigs, and the like).

[00095] The term "pharmaceutically acceptable carrier" or "physiologically acceptable carrier" as used herein refers to one or more formulation materials suitable for accomplishing or enhancing the delivery of a fusion protein of the invention.

[00096] The term "antigen" refers to a molecule or a portion of a molecule that is capable of being bound by an antibody, and additionally that is capable of being used in an animal to produce antibodies that are capable of binding to an epitope of that antigen. An antigen may have one or more epitopes.

[00097] The term "native Fc" refers to molecule or sequence comprising the sequence of a non-antigen-binding fragment resulting from digestion of whole antibody or produced by other means, whether in monomeric or multimeric form, and can contain the hinge region. The original immunoglobulin source of the native Fc is preferably of human origin and can be any of the immunoglobulins, although IgG1 and IgG2 are preferred. Native Fc molecules are made up of monomeric polypeptides that can be linked into dimeric or multimeric forms by covalent (i.e., disulfide bonds) and non-covalent association. The number of intermolecular disulfide bonds between monomeric subunits of native Fc molecules ranges from 1 to 4 depending on class (e.g., IgG, IgA,

and IgE) or subclass (e.g., IgG1, IgG2, IgG3, IgA1, and IgGA2). One example of a native Fc is a disulfide-bonded dimer resulting from papain digestion of an IgG (see Ellison et al., 1982, Nucleic Acids Res. 10: 4071-9). The term "native Fc" as used herein is generic to the monomeric, dimeric, and multimeric forms. The term "Fc variant" refers to a molecule or sequence that is modified from a native Fc but still comprises a binding site for the salvage receptor, FcRn (neonatal Fc receptor). International Publication Nos. WO 97/34631 and WO 96/32478 describe exemplary Fc variants, as well as interaction with the salvage receptor, and are hereby incorporated by reference. Thus, the term "Fc variant" can comprise a molecule or sequence that is humanized from a non-human native Fc. Furthermore, a native Fc comprises regions that can be removed because they provide structural features or biological activity that are not required for the fusion molecules of the Proteins of the invention. Thus, the term "Fc variant" comprises a molecule or sequence that lacks one or more native Fc sites or residues, or in which one or more Fc sites or residues has be modified, that affect or are involved in: (1) disulfide bond formation, (2) incompatibility with a selected host cell, (3) N-terminal heterogeneity upon expression in a selected host cell, (4) glycosylation, (5) interaction with complement, (6) binding to an Fc receptor other than a salvage receptor, or (7) antibody-dependent cellular cytotoxicity (ADCC). Fc variants are described in further detail hereinafter.

[00098] The term "Fc domain" encompasses native Fc and Fc variants and sequences as defined above. As with Fc variants and native Fc molecules, the term "Fc domain" includes molecules in monomeric or multimeric form, whether digested from whole antibody or produced by other means. In some embodiments of the present invention, an Fc domain can be fused to FGF21 or a FGF21 mutant (including a truncated form of FGF21 or a FGF21 mutant) via, for example, a covalent bond between the Fc domain and the FGF21 sequence. Such fusion proteins can form multimers via the association of the Fc domains and both these fusion proteins and their multimers are an aspect of the present invention.

[00099] As used in the present text, the term "fusobody" refers to an antibody-like soluble protein comprising two heterodimers, each heterodimer consisting of one heavy and one light chains of amino acids, stably associated together, for example, via one or more disulfide bond(s). Each heavy or light chain comprises constant regions of an antibody, referred hereafter respectively as the heavy and light chain constant regions of the fusobody. The heavy chain constant region comprises at least C_H1 region of an antibody and may further comprise C_H2 and C_H3 regions, including the hinge region. The light chain constant region comprises C_L region of an antibody. In a fusobody, the variable regions of an antibody are replaced by heterologous soluble binding domains.

By way of non-limiting example, a fusobody of the invention can comprise a dual function fusion protein of the invention, wherein the GLP-1 receptor agonist is fused to the N-terminus of heavy and light chain of an antibody and FGF21 is simultaneously fused to the C-terminus of heavy and light chain of the same antibody.

[000100] The term "heterologous" means that these domains are not naturally found associated with constant regions of an antibody. In particular, such heterologous binding domains do not have the typical structure of an antibody variable domain consisting of 4 framework regions, FR1, FR2, FR3 and FR4 and the 3 complementarity determining regions (CDRs) in-between. Each arm of the fusobody therefore comprises a first single chain polypeptides comprising a first binding domain covalently linked at the N-terminal part of a constant C_H1 heavy chain region of an antibody, and a second single chain polypeptide comprising a second binding domain covalently linked at the N-terminal part of a constant C_L light chain of an antibody. The covalent linkage may be direct, for example via peptidic bound or indirect, via a linker, for example a peptidic linker. The two heterodimers of the fusobody are covalently linked, for example, by at least one disulfide bridge at their hinge region, like an antibody structure. Examples of molecules with a fusobody structure have been described in the art, in particular, fusobodies comprising ligand binding region of heterodimeric receptor (see for example international patent publications WO01/46261 and WO11/076781).

[000101] The term "polyethylene glycol" or "PEG" refers to a polyalkylene glycol compound or a derivative thereof, with or without coupling agents or derviatization with coupling or activating moieties.

[000102] The term "Exenatide" indicates a synthetic version of exendin-4, Exenatide, marketed as Byetta and Bydureon, is a glucagon-like peptide-1 agonist (GLP-1 agonist) medication approved in April 2005 for the treatment of diabetes mellitus type 2. It belongs to the group of incretin mimetics and is manufactured by Amylin Pharmaceuticals.

[000103] Exendin-4 is a 39 residue polypeptide produced in the salivary glands of the Gila Monster lizard (Goke et al. (1993) Diabetes 46:433-439; Fehmann et al. (1995) Endocrine Rev. 16:390-410). Although it is the product of a uniquely non-mammalian gene and appears to be expressed only in the salivary gland, Exendin-4 shares a 52% amino acid sequence homology with GLP-1 and in mammals interacts with the GLP-1 receptor (Goke, et al.; Thorens et al. (1993) Diabetes 42:1678-1682).

[000104] The term "FGF21-associated disorders," "GLP-1-associated disorders," "Exendin-4-associated disorders," and terms similarly used herein, includes obesity, type 1 and type 2 diabetes mellitus, pancreatitis, dyslipidemia, nonalcoholic fatty liver disease (NAFLD), nonalcoholic steatohepatitis (NASH), insulin resistance,

hyperinsulinemia, glucose intolerance, hyperglycemia, metabolic syndrome, acute myocardial infarction, hypertension, cardiovascular disease, atherosclerosis, peripheral arterial disease, stroke, heart failure, coronary heart disease, kidney disease, diabetic complications, neuropathy, gastroparesis, disorders associated with severe inactivating mutations in the insulin receptor, lipodystrophies including HIV-associated lipodystrophy, and other metabolic disorders.

[000105] The term "disorders associated with severe inactivating mutations in the insulin receptor," and terms similarly used herein, describe conditions in subjects afflicted with mutations in the insulin receptor (or possible proteins directly downstream from it) which cause severe insulin resistance but are often (though not always) seen without the obesity common in Type 2 diabetes mellitus. In many ways, subjects afflicted with these conditions manifest hybrid symptoms of Type 1 diabetes mellitus and Type 2 diabetes mellitus. Subjects thereby afflicted fall into several categories of roughly increasing severity, including: Type A Insulin Resistance, Type C Insulin Resistance (AKA HAIR-AN Syndrome), Rabson-Mendenhall Syndrome and finally Donohue's Syndrome or Leprechaunism. These disorders are associated with very high endogenous insulin levels, and very often, hyperglycemia. Subjects thereby afflicted also present with various clinical features associated with "insulin toxicity," including hyperandrogenism, polycystic ovarian syndrome (PCOS), hirsuitism, and acanthosis nigricans (excessive growth and pigmentation) in the folds of the skin.

[000106] "Lipodystrophies, including HIV-associated lipodystrophy" are disorders of adipose tissue characterized by a selective loss of body fat. Patients with lipodystrophy have a tendency to develop insulin resistance, hyperinsulinemia, hyperglycemia, hypertriglyceridemia and fatty liver. There are numerous forms of lipodystrophy that are inherited (genetic) or acquired. The genetic forms of lipodystrophy include congenital generalized lipodystrophy (the Berardinelli-Seip syndrome) and several types of familial partial lipodystrophy (the Dunnigan type, the Kobberling type, the mandibuloacral dysplasis type). The acquired forms of lipodystrophy include acquired generalized lipodystrophy (the Lawrence syndrome), acquired partial lipodystrophy (the Barraquer-Simons syndrome), and lipodystrophy induced by protease inhibitors and nucleoside reverse transcriptase inhibitors used to treat HIV.

[000107] "Type 2 diabetes mellitus" is a condition characterized by excess glucose production in spite of the availability of insulin, and circulating glucose levels remain excessively high as a result of inadequate glucose clearance.

[000108] "Type 1 diabetes mellitus" is a condition characterized by high blood glucose levels caused by total lack of insulin. This occurs when the body's immune system

attacks the insulin-producing beta cells in the pancreas and destroys them. The pancreas then produces little or no insulin.

[000109] "Glucose intolerance" or Impaired Glucose Tolerance (IGT) is a pre-diabetic state of dysglycemia that is associated with increased risk of cardiovascular pathology. The pre-diabetic condition prevents a subject from moving glucose into cells efficiently and utilizing it as an efficient fuel source, leading to elevated glucose levels in blood and some degree of insulin resistance.

[000110] "Hyperglycemia" is defined as an excess of sugar (glucose) in the blood.

[000111] "Hypoglycemia", also called low blood sugar, occurs when your blood glucose level drops too low to provide enough energy for your body's activities.

[000112] "Hyperinsulinemia" is defined as a higher-than-normal level of insulin in the blood.

[000113] "Insulin resistance" is defined as a state in which a normal amount of insulin produces a subnormal biologic response.

[000114] "Obesity," in terms of the human subject, can be defined as that body weight over 20 percent above the ideal body weight for a given population (R. H. Williams, Textbook of Endocrinology, 1974, p. 904-916).

[000115] "Diabetic complications" are problems, caused by high blood glucose levels, with other body functions such as kidneys, nerves (neuropathies), feet (foot ulcers and poor circulation) and eyes (e.g. retinopathies). Diabetes also increases the risk for heart disease and bone and joint disorders. Other long-term complications of diabetes include skin problems, digestive problems, sexual dysfunction and problems with teeth and gums.

[000116] "Metabolic syndrome" can be defined as a cluster of at least three of the following signs: abdominal fat--in most men, a 40-inch waist or greater; high blood sugar--at least 110 milligrams per deciliter (mg/dl) after fasting; high triglycerides--at least 150 mg/dL in the bloodstream; low HDL--less than 40 mg/dl; and, blood pressure of 130/85 mmHg or higher.

[000117] "Pancreatitis" is inflammation of the pancreas.

[000118] "Dyslipidemia" is a disorder of lipoprotein metabolism, including lipoprotein overproduction or deficiency. Dyslipidemias may be manifested by elevation of the total cholesterol, low-density lipoprotein (LDL) cholesterol and triglyceride concentrations, and a decrease in high-density lipoprotein (HDL) cholesterol concentration in the blood.

[000119] "Nonalcoholic fatty liver disease (NAFLD)" is a liver disease, not associated with alcohol consumption, characterized by fatty change of hepatocytes.

[000120] "Nonalcoholic steatohepatitis (NASH)" is a liver disease, not associated with alcohol consumption, characterized by fatty change of hepatocytes, accompanied by intralobular inflammation and fibrosis.

[000121] "Hypertension" or high blood pressure that is a transitory or sustained elevation of systemic arterial blood pressure to a level likely to induce cardiovascular damage or other adverse consequences. Hypertension has been arbitrarily defined as a systolic blood pressure above 140 mmHg or a diastolic blood pressure above 90 mmHg.

[000122] "Cardiovascular diseases" are diseases related to the heart or blood vessels.

[000123] "Acute myocardial infarction" occurs when there is interruption of the blood supply to a part of the heart. The resulting ischemia and oxygen shortage, if left untreated for a sufficient period of time, can cause damage or death (infarction) of the heart muscle tissue (myocardium).

[000124] "Peripheral arterial disease" occurs when plaque builds up in the arteries that carry blood to the head, organs and limbs. Over time, plaque can harden and narrow the arteries which limits the flow of oxygen-rich blood to organs and other parts of the body.

[000125] "Atherosclerosis" is a vascular disease characterized by irregularly distributed lipid deposits in the intima of large and medium-sized arteries, causing narrowing of arterial lumens and proceeding eventually to fibrosis and calcification. Lesions are usually focal and progress slowly and intermittently. Limitation of blood flow accounts for most clinical manifestations, which vary with the distribution and severity of lesions.

[000126] "Stroke" is any acute clinical event, related to impairment of cerebral circulation, that lasts longer than 24 hours. A stroke involves irreversible brain damage, the type and severity of symptoms depending on the location and extent of brain tissue whose circulation has been compromised.

[000127] "Heart failure", also called congestive heart failure, is a condition in which the heart can no longer pump enough blood to the rest of the body.

[000128] "Coronary heart disease", also called coronary artery disease, is a narrowing of the small blood vessels that supply blood and oxygen to the heart.

[000129] "Kidney disease" or nephropathy is any disease of the kidney. Diabetic nephropathy is a major cause of morbidity and mortality in people with type 1 or type 2 diabetes mellitus.

[000130] "Neuropathies" are any diseases involving the cranial nerves or the peripheral or autonomic nervous system.

[000131] "Gastroparesis" is weakness of gastric peristalsis, which results in delayed emptying of the bowels.

[000132] "Click chemistry" is a term that was introduced by K. B. Sharpless in 2001 to describe reactions that are high yielding, wide in scope, create only byproducts that can be removed without chromatography, are stereospecific, simple to perform, and can be conducted in easily removable or benign solvents.

[000133] The critically ill patients encompassed by the present invention generally experience an unstable hypermetabolic state. This unstable metabolic state is due to changes in substrate metabolism, which may lead to relative deficiencies in some nutrients. Generally there is an increased oxidation of both fat and muscle.

[000134] Moreover, critically ill patients are preferably patients that experience systemic inflammatory response syndrome or respiratory distress. A reduction in morbidity means reducing the likelihood that a critically ill patient will develop additional illnesses, conditions, or symptoms or reducing the severity of additional illnesses, conditions, or symptoms. For example reducing morbidity may correspond to a decrease in the incidence of bacteremia or sepsis or complications associated with multiple organ failure.

[000135] As used herein, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise. Thus, for example, reference to "an antibody" includes a mixture of two or more such antibodies.

[000136] As used herein, the term "about" refers to +/- 20%, more preferably, +/- 10%, or still more preferably, +/- 5% of a value.

[000137] The terms "polypeptide" and "protein", are used interchangeably and refer to a polymeric form of amino acids of any length, which can include coded and non-coded amino acids, naturally and non-naturally occurring amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones. The term includes fusion proteins, including, but not limited to, fusion proteins with a heterologous amino acid sequence, fusions with heterologous and homologous leader sequences, with or without N-terminal methionine residues; immunologically tagged proteins; and the like.

[000138] The terms "individual", "subject", "host" and "patient" are used interchangeably and refer to any subject for whom diagnosis, treatment, or therapy is desired, particularly humans. Other subjects may include cattle, dogs, cats, guinea pigs, rabbits, rats, mice, horses, and the like. In some preferred embodiments the subject is a human.

[000139] As used herein, the term "sample" refers to biological material from a patient. The sample assayed by the present invention is not limited to any particular type. Samples include, as non-limiting examples, single cells, multiple cells, tissues, tumors, biological fluids, biological molecules, or supernatants or extracts of any of the

foregoing. Examples include tissue removed for biopsy, tissue removed during resection, blood, urine, lymph tissue, lymph fluid, cerebrospinal fluid, mucous, and stool samples. The sample used will vary based on the assay format, the detection method and the nature of the tumors, tissues, cells or extracts to be assayed. Methods for preparing samples are well known in the art and can be readily adapted in order to obtain a sample that is compatible with the method utilized.

[000140] As used herein, the term "biological molecule" includes, but is not limited to, polypeptides, nucleic acids, and saccharides.

[000141] As used herein, the term "modulating" refers to a change in the quality or quantity of a gene, protein, or any molecule that is inside, outside, or on the surface of a cell. The change can be an increase or decrease in expression or level of the molecule. The term "modulates" also includes changing the quality or quantity of a biological function/activity including, without limitation, the ability to lower blood glucose, insulin, triglyceride, or cholesterol levels; to reduce liver lipid or liver triglyceride levels; to reduce body weight; and to improve glucose tolerance, energy expenditure, or insulin sensitivity.

[000142] As used herein, the term "modulator" refers to a composition that modulates one or more physiological or biochemical events associated with an FGF21-associated disorder, such as type 1 or type 2 diabetes mellitus or a metabolic condition like obesity. Said events include but are not limited to the ability to lower blood glucose, insulin, triglyceride, or cholesterol levels; to reduce liver lipid or liver triglyceride levels; to reduce body weight; and to improve glucose tolerance, energy expenditure, or insulin sensitivity.

[000143] A "gene product" is a biopolymeric product that is expressed or produced by a gene. A gene product may be, for example, an unspliced RNA, an mRNA, a splice variant mRNA, a polypeptide, a post-translationally modified polypeptide, a splice variant polypeptide etc. Also encompassed by this term are biopolymeric products that are made using an RNA gene product as a template (i.e. cDNA of the RNA). A gene product may be made enzymatically, recombinantly, chemically, or within a cell to which the gene is native. In some embodiments, if the gene product is proteinaceous, it exhibits a biological activity. In some embodiments, if the gene product is a nucleic acid, it can be translated into a proteinaceous gene product that exhibits a biological activity.

[000144] "Modulation of FGF21 activity," "Modulation of GLP-1 activity," and "Modulation of Exendin-4 activity," as used herein, refers to an increase or decrease in FGF21, GLP-1, or Exendin-4 activity, respectively, that can be a result of, for example, interaction of an agent with an FGF21, GLP-1, or Exendin-4 polynucleotide or polypeptide, inhibition of FGF21, GLP-1, or Exendin-4 transcription and/or translation

(e.g., through antisense or siRNA interaction with the FGF21, GLP-1, or Exendin-4 gene or FGF21, GLP-1, or Exendin-4 transcript, through modulation of transcription factors that facilitate FGF21, GLP-1, or Exendin-4 expression), and the like. For example, modulation of a biological activity refers to an increase or a decrease in a biological activity. FGF21, GLP-1, or Exendin-4 activity can be assessed by means including, without limitation, assaying blood glucose, insulin, triglyceride, or cholesterol levels in a subject, assessing FGF21, GLP-1, or Exendin-4 polypeptide levels, or by assessing FGF21, GLP-1, or Exendin-4 transcription levels.

[000145] Comparisons of FGF21, GLP-1, or Exendin-4 activity can also be accomplished by, e.g., measuring levels of an FGF21, GLP-1, or Exendin-4 downstream biomarker, and measuring increases in FGF21, GLP-1, or Exendin-4 signaling. FGF21, GLP-1, or Exendin-4 activity can also be assessed by measuring: cell signaling; kinase activity; glucose uptake into adipocytes; blood insulin, triglyceride, or cholesterol level fluctuations; liver lipid or liver triglyceride level changes; interactions between FGF21, GLP-1, or Exendin-4 and a receptor; or phosphorylation of an FGF21, GLP-1, or Exendin-4 receptor. In some embodiments phosphorylation of an FGF21, GLP-1, or Exendin-4 receptor can be tyrosine phosphorylation. In some embodiments modulation of FGF21, GLP-1, or Exendin-4 activity can cause modulation of an FGF21, GLP-1, or Exendin-4-related phenotype.

[000146] A "FGF21, GLP-1, or Exendin-4 downstream biomarker," as used herein, is a gene or gene product, or measurable indicia of a gene or gene product. In some embodiments, a gene or activity that is a downstream marker of FGF21, GLP-1, or Exendin-4 exhibits an altered level of expression. In some embodiments, an activity of the downstream marker is altered in the presence of an FGF21, GLP-1, or Exendin-4 modulator. In some embodiments, the downstream markers exhibit altered levels of expression when FGF21, GLP-1, or Exendin-4 signaling is perturbed with dual function protein of the present invention. For example, FGF21 downstream markers include, without limitation, glucose or 2-deoxy-glucose uptake, pERK and other phosphorylated or acetylated proteins or NAD levels.

[000147] As used herein, the term "up-regulates" refers to an increase, activation or stimulation of an activity or quantity. For example, FGF21, GLP-1, or Exendin-4 modulators, such as the dual function proteins of the invention, may increase or agonize the activity of an FGF21, GLP-1, or Exendin-4 receptor. In one embodiment, one or more of FGFR-1(IIIc), FGFR-2(IIIc), or FGFR-3(IIIc) and/or β-klotho may be up-regulated in response to a dual function protein of the invention. Up-regulation can also refer to an FGF21, GLP-1, or Exendin-4-related activity, such as e.g., the ability to lower blood glucose, insulin, triglyceride, or cholesterol levels; to reduce liver lipid or

triglyceride levels; to reduce body weight; to improve glucose tolerance, energy expenditure, or insulin sensitivity; or to cause phosphorylation of an FGF21, GLP-1, or Exendin-4 receptor; or to increase an FGF21, GLP-1, or Exendin-4 downstream marker. For example, the FGFR21 receptor can be one or more of FGFR-1(IIIc), FGFR-2(IIIc), or FGFR-3(IIIc) and/or β-klotho. Up-regulation may range anywhere from 25% to 500% as compared to a control.

[000148] As used herein, the term "N-terminus" refers to at least the first 20 amino acids of a protein.

[000149] As used herein, the terms "N-terminal domain" and "N-terminal region" are used interchangeably and refer to a fragment of a protein that begins at the first amino acid of the protein and ends at any amino acid in the N-terminal half of the protein. For example, the N-terminal domain of FGF21 is from amino acid 1 of SEQ ID NO:1 to any amino acid between about amino acids 10 and 105 of SEQ ID NO:1.

[000150] As used herein, the term "C-terminus" refers to at least the last 20 amino acids of a protein.

[000151] As used herein, the terms "C-terminal domain" and "C-terminal region" are used interchangeably and refer to a fragment of a protein that begins at any amino acid in the C-terminal half of the protein and ends at the last amino acid of the protein. For example, the C-terminal domain of FGF21 begins at any amino acid from amino acid 105 to about amino acid 200 of SEQ ID NO:1 and ends at amino acid 209 of SEQ ID NO:1.

[000152] The term "domain" as used herein refers to a structural part of a biomolecule that contributes to a known or suspected function of the biomolecule. Domains may be co-extensive with regions or portions thereof and may also incorporate a portion of a biomolecule that is distinct from a particular region, in addition to all or part of that region.

[000153] As used herein, the term "signal domain" (also called "signal sequence" or "signal peptide") refers to a peptide domain that resides in a continuous stretch of amino acid sequence at the N-terminal region of a precursor protein (often a membrane-bound or secreted protein) and is involved in post-translational protein transport. In many cases the signal domain is removed from the full-length protein by specialized signal peptidases after the sorting process has been completed. Each signal domain specifies a particular destination in the cell for the precursor protein. The signal domain of FGF21 is amino acids 1-28 of SEQ ID NO:1.

[000154] As used herein, the term "receptor binding domain" refers to any portion or region of a protein that contacts a membrane-bound receptor protein, resulting in a cellular response, such as a signaling event.

[000155] As used herein, the term "ligand binding domain" refers to any portion or region of a fusion protein of the invention retaining at least one qualitative binding activity of a corresponding native sequence.

[000156] The term "region" refers to a physically contiguous portion of the primary structure of a biomolecule. In the case of proteins, a region is defined by a contiguous portion of the amino acid sequence of that protein. In some embodiments a "region" is associated with a function of the biomolecule.

[000157] The term "fragment" as used herein refers to a physically contiguous portion of the primary structure of a biomolecule. In the case of proteins, a portion is defined by a contiguous portion of the amino acid sequence of that protein and refers to at least 3-5 amino acids, at least 8-10 amino acids, at least 11-15 amino acids, at least 17-24 amino acids, at least 25-30 amino acids, and at least 30-45 amino acids. In the case of oligonucleotides, a portion is defined by a contiguous portion of the nucleic acid sequence of that oligonucleotide and refers to at least 9-15 nucleotides, at least 18-30 nucleotides, at least 33-45 nucleotides, at least 48-72 nucleotides, at least 75-90 nucleotides, and at least 90-130 nucleotides. In some embodiments, portions of biomolecules have a biological activity. In the context of the present invention, FGF21 polypeptide fragments do not comprise the entire FGF21 polypeptide sequence set forth in SEQ ID NO:1.

[000158] A "native sequence" polypeptide is one that has the same amino acid sequence as a polypeptide derived from nature. Such native sequence polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. Thus, a native sequence polypeptide can have the amino acid sequence of naturally occurring human polypeptide, murine polypeptide, or polypeptide from any other mammalian species.

[000159] As used herein, the phrase "homologous nucleotide sequence," or "homologous amino acid sequence," or variations thereof, refers to sequences characterized by a homology, at the nucleotide level or amino acid level, of at least a specified percentage and is used interchangeably with "sequence identity." Homologous nucleotide sequences include those sequences coding for isoforms of proteins. Such isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. Homologous nucleotide sequences include nucleotide sequences encoding for a protein of a species other than humans, including, but not limited to, mammals. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. Homologous amino acid sequences include those amino

acid sequences which contain conservative amino acid substitutions and which polypeptides have the same binding and/or activity. In some embodiments, a nucleotide or amino acid sequence is homologous if it has at least 60% or greater, up to 99%, identity with a comparator sequence. In some embodiments, a nucleotide or amino acid sequence is homologous if it shares one or more, up to 60, nucleotide/amino acid substitutions, additions, or deletions with a comparator sequence. In some embodiments, the homologous amino acid sequences have no more than 5 or no more than 3 conservative amino acid substitutions.

[000160] Percent homology or identity can be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison WI), using default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489). In some embodiments, homology between the probe and target is between about 75% to about 85%. In some embodiments, nucleic acids have nucleotides that are at least about 95%, about 97%, about 98%, about 99% and about 100% homologous to SEQ ID NO:2, or a portion thereof.

[000161] Homology may also be at the polypeptide level. In some embodiments, constituent polypeptides of the dual function proteins of the invention may be at least 95% homologous to their full-length wild-type counterparts or corresponding native sequences, or to portions thereof. The degree or percentage identity of dual function proteins of the invention, or portions thereof, and different amino acid sequences is calculated as the number of exact matches in an alignment of the two sequences divided by the length of the “invention sequence” or the “foreign sequence”, whichever is shortest. The result is expressed as percent identity.

[000162] As used herein, the term “mixing” refers to the process of combining one or more compounds, cells, molecules, and the like together in the same area. This may be performed, for example, in a test tube, petri dish, or any container that allows the one or more compounds, cells, or molecules, to be mixed.

[000163] As used herein, the term “substantially purified” refers to a compound (e.g., either a polynucleotide or a polypeptide or an antibody) that is removed from its natural environment and is at least 60% free, at least 75% free, and at least 90% free from other components with which it is naturally associated.

[000164] The term “pharmaceutically acceptable carrier” refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which can be administered without undue toxicity. Suitable carriers

can be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Pharmaceutically acceptable carriers in therapeutic compositions can include liquids such as water, saline, glycerol and ethanol. Auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, can also be present in such vehicles.

[000165] Naturally occurring disulfide bonds, as provided by cysteine residues, generally increase thermodynamic stability of proteins. Successful examples of increased thermodynamic stability, as measured in increase of the melting temperature, are multiple disulfide-bonded mutants of the enzymes T4 lysozyme (Matsumuraet al., PNAS 86:6562-6566 (1989)) and barnase (Johnson et al., J. Mol. Biol. 268:198-208 (1997)). An aspect of the present invention is an enhancement of the physical stability of FGF21 in the presence of a preservative, achieved by the presence of disulfide bonds within the variants, which constrain the flexibility of wild-type FGF21 and thereby limit access of the preservative to the hydrophobic core of the protein. Enhancement of the physical stability of FGF21 within the proteins of the present invention, due to the presence of disulfide bonds within the variants, can confer additional protection to said proteins, whether or not in the presence of a preservative, including but not limited to protection against fluctuations in environmental conditions such as pH and temperature.

Improvements of the Dual Function Proteins of the Invention Over Wild Type Protein Comparators and Combinations Thereof

[000166] It is well known in the art that a significant challenge in the development of protein pharmaceuticals is to deal with the physical and chemical instabilities of proteins. This is even more apparent when a protein pharmaceutical formulation is intended to be a multiple use, injectable formulation requiring a stable, concentrated and preserved solution, while maintaining a favorable bioactivity profile. Biophysical characterization of wild-type FGF21 in the literature established that a concentrated protein solution (>5 mg/ml), when exposed to stress conditions, such as high temperature or low pH, lead to accelerated association and aggregation (i.e., poor physical stability and biopharmaceutical properties). Exposure of a concentrated protein solution of FGF21 to pharmaceutical preservatives (e.g., m-cresol) also had a negative impact on physical stability.

[000167] Therefore, an embodiment of the present invention is to enhance physical stability of concentrated solutions, while maintaining chemical stability and biological potency, under both physiological and preserved formulation conditions. It is thought

that association and aggregation may result from hydrophobic interactions, since, at a given protein concentration, temperature, and ionic strength have considerable impact on physical stability.

[000168] For the most part, non-conserved, presumed surface exposed amino acid residues are targeted. The local environment of these residues is analyzed and, those not deemed structurally important are selected for mutagenesis. One method to initiate specific changes is to further decrease the pI of the protein by introducing glutamic acid residues ("glutamic acid scan"). The introduction of charged substitutes can inhibit hydrophobic-mediated aggregation via charge-charge repulsion and potentially improve preservative compatibility. In addition, one skilled in the art would also recognize that with sufficient mutagenesis, the pI could be shifted into a basic pH range by the introduction of positive charge, with or without concomitant decrease in negative charge, thus allowing for charge-charge repulsion.

[000169] An additional difficulty associated with therapeutic applications of wild-type FGF21 as a biotherapeutic, for instance, is that its half-life is very short *in vivo* (on the order of 0.5 and 2 h, respectively, in mouse and primate). There is hence a need to develop follow-up compounds that are more efficacious either through higher potency or longer half-life. Similarly, various mechanisms have been employed to enhance serum half-life for GLP-1 (1-2 minutes due to rapid cleavage by endogenous proteases, particularly dipeptidyl peptidase-4 (DPP4)) through use of Exendin-4 or other analogues that are resistant to cleavage by DPP4 as well as further modifications and formulations for extended half-life or slow release of compound.

[000170] The proteins of the invention, e.g., dual FGF21 receptor agonist and GLP-1 receptor agonist proteins of the invention, were developed as a way to achieve the desirable effects of FGF21 and GLP-1 treatment at a higher potency and in a half-life-extended formulation. The co-administration of GLP-1 and FGF21 has been described in the literature, e.g., in patent publications WO2010/068735 and WO2009/020802, with data suggesting additive or synergistic effects of co-administration of GLP-1 and FGF21 for treatment of obesity and type 2 diabetes. However, co-localization of the two receptor agonists in a single molecule, e.g., in the form of the proteins of the invention, provides better access of both GLP-1 and FGF21 to tissues or cells and provides increased benefit over simple co-administration as separate entities. This advantage of combining two agonists into a single molecule compared to simple co-administration of each is especially advantageous for tissues in which both the GLP-1 and the FGF21 receptors are expressed in the same tissues, such as adipose, pancreatic β-cells, hepatic and hypothalamic cells. Dual FGF21 receptor agonist and GLP-1 receptor agonist proteins of the invention and other single dual activity entities have improved

biological properties due to altered receptor trafficking, altered signal transduction effects and/or entropic (avidity) effects.

[000171] Because FGF21 and GLP-1 (or Exendin-4) act through different mechanisms, it is expected that they will have additive or synergistic effects when, for example, administered at the same time in the form of the dual function proteins of the invention. GLP-1 and Exendin-4 act primarily to increase insulin secretion in response to food intake while FGF21 sensitizes the body to respond better to insulin, which may provide benefit for both type 1 and type 2 diabetes in managing blood glucose levels. Beta-cell protective effects combined with the improved beta-cell function and insulin sensitivity has the potential to provide benefit even in type 1 diabetics.

[000172] Another example is weight loss: The satiety signal and slowed gastric emptying by GLP-1 or Exendin-4 are expected to lower appetite while FGF21 increases metabolism in adipose and other tissues which could increase loss of fat. These two effects could result in an additive or even synergistic weight loss with combined dosing. Expression of receptors for both FGF21 and GLP-1 in metabolically active tissues suggests that simultaneous delivery of a FGF21 receptor agonist and a GLP-1 receptor agonist to such tissues may be beneficial for treating metabolic diseases. A combination of GLP-1 and Exendin-4's cardioprotective effects and improved lipid profiles seen with FGF21 could also result in an additive benefit for cardiovascular disease associated with obesity, type 2 and type 1 diabetes.

[000173] Another likely benefit of GLP-1 and FGF21 dual agonism is an improved, i.e., reduced, side effect profile. Recent studies (Wei et al. (2012) PNAS 109, 3143-48) indicate that treatment of diet-induced obese mice with FGF21 induces bone loss, due to a diminished inactivation of PPARγ (via reduced sumoylation); a shift of mesenchymal stem cell differentiation from osteoblasts to adipocytes is seen in the presence of increased PPARγ activity in the bone following FGF21 treatment. However, it has also been reported (Sanz et al. (2010) Am J Physiol Endocrinol Metab 298, E634-E643) that GLP-1 can reduce the differentiation of human mesenchymal stem cells to adipocytes by reducing the expression of PPARγ. Therefore, as GLP-1 has the potential to be bone protective, GLP-1-FGF21 fusion proteins are likely to reduce the risk of bone loss compared to FGF21-only treatments.

[000174] Although the embodiments of the present invention concern the physical and chemical stability under both physiological and preserved pharmaceutical formulation conditions, maintaining the biological potency of the proteins of the invention as compared to, e.g., wild-type FGF21 is an important factor of consideration as well. Therefore, the biological potency of the proteins of the present invention is defined by

the ability of the proteins to affect glucose uptake and/or the lowering of plasma glucose levels, as shown herein in the examples.

[000175] The proteins, polypeptides, and/or peptides of the invention administered according to this invention may be generated and/or isolated by any means known in the art. The most preferred method for producing the variant is through recombinant DNA methodologies and is well known to those skilled in the art. Such methods are described in Current Protocols in Molecular Biology (John Wiley & Sons, Inc.), which is incorporated herein by reference.

[000176] Additionally, the preferred embodiments include a biologically active peptide derived from the variant described herein. Such a peptide will contain at least one of the substitutions described and the variant will possess biological activity. The peptide may be produced by any and all means known to those skilled in the art, examples of which included but are not limited to enzymatic digestion, chemical synthesis or recombinant DNA methodologies.

[000177] It is established in the art that fragments of peptides of certain fibroblast growth factors are biologically active. See for example, Baird et al., Proc. Natl. Acad. Sci (USA) 85:2324-2328 (1988), and J. Cell. Phys. Suppl. 5:101-106 (1987). Therefore, the selection of fragments or peptides of the variant is based on criteria known in the art. For example, it is known that dipeptidyl peptidase IV (DPP-IV, or DPP-4) is a serine type protease involved in inactivation of neuropeptides, endocrine peptides, and cytokines (Damme et al. Chem. Immunol. 72: 42-56, (1999)). The N-terminus of FGF21 (HisProIlePro) contains two dipeptides that could potentially be substrates to DPP-IV, resulting in a fragment of FGF21 truncated at the N-terminus by 4 amino acids. Unexpectedly, this fragment of wild-type FGF21 has been demonstrated to retain biological activity, thus, proteins of the present invention truncated at the N-terminus by up to 4 amino acids, is an embodiment of the present invention.

[000178] The invention also encompasses polynucleotides encoding the above-described variants that may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded. The coding sequences that encode the proteins of the present invention may vary as a result of the redundancy or degeneracy of the genetic code.

[000179] The polynucleotides that encode for the proteins of the invention may include the following: only the coding sequence for the variant, the coding sequence for the variant and additional coding sequence such as a functional polypeptide, or a leader or secretory sequence or a pro-protein sequence; the coding sequence for the variant and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the variant. Thus the term "polynucleotide encoding a variant"

encompasses a polynucleotide that may include not only coding sequence for the variant but also a polynucleotide, which includes additional coding and/or non-coding sequence.

[000180] The invention further relates to variants of the described polynucleotides that encode for fragments, analogs and derivatives of the polypeptide that contain the indicated substitutions. The variant of the polynucleotide may be a naturally occurring allelic variant of the human FGF21 sequence, a non-naturally occurring variant, or a truncated variant as described above. Thus, the present invention also includes polynucleotides encoding the variants described above, as well as variants of such polynucleotides, which variants encode for a fragment, derivative or analog of the disclosed variant. Such nucleotide variants include deletion variants, substitution variants, truncated variants, and addition or insertion variants as long as at least one of the indicated amino acid substitutions of the first or second embodiments is present.

[000181] The polynucleotides of the invention will be expressed in hosts after the sequences have been operably linked to (i.e., positioned to ensure the functioning of) an expression control sequence. These expression vectors are typically replicable in the host organisms either as episomes or as an integral part of the host chromosomal DNA. Commonly, expression vectors will contain selection markers, e.g., tetracycline, neomycin, and dihydrofolate reductase, to permit detection of those cells transformed with the desired DNA sequences. The dual function proteins or fragments thereof can be expressed in mammalian cells, insect, yeast, bacterial or other cells under the control of appropriate promoters. Cell free translation systems can also be employed to produce such proteins using RNAs derived from DNA constructs of the present invention.

[000182] E. coli is a prokaryotic host useful particularly for cloning the polynucleotides of the present invention. Other microbial hosts suitable for use include Bacillus subtilus, Salmonella typhimurium, and various species of Serratia, Pseudomonas, Streptococcus, and Staphylococcus, although others may also be employed as a matter of choice. In these prokaryotic hosts, one can also make expression vectors, which will typically contain expression control sequences compatible with the host cell (e.g., an origin of replication). In addition, any of a number of well-known promoters may be present, such as the lactose promoter system, a tryptophan (Trp) promoter system, a beta-lactamase promoter system, or a promoter system from phages lambda or T7. The promoters will typically control expression, optionally with an operator sequence, and have ribosome binding site sequences and the like, for initiating and completing transcription and translation.

[000183] One skilled in the art of expression of proteins will recognize that methionine or methionine-arginine sequence can be introduced at the N-terminus of the mature

sequence (SEQ ID NO: 3) for expression in E. coli and are contemplated within the context of this invention. Thus, unless otherwise noted, proteins of the present invention expressed in E. coli have a methionine sequence introduced at the N-terminus.

[000184] Other microbes, such as yeast or fungi, may also be used for expression. Pichia pastoris, Saccharomyces cerevisiae, Schizosaccharomyces pombe, and Pichia angusta are examples of preferred yeast hosts, with suitable vectors having expression control sequences, such as promoters, including 3-phosphoglycerate kinase or other glycolytic enzymes, and an origin of replication, termination sequences and the like as desired. Aspergillus niger, Trichoderma reesei; and Schizophyllum commune, are examples of fungi hosts, although others may also be employed as a matter of choice.

[000185] Mammalian tissue cell culture may also be used to express and produce the polypeptides of the present invention. Eukaryotic cells are actually preferred, because a number of suitable host cell lines capable of secreting intact variants have been developed in the art, and include the CHO cell lines, various COS cell lines, NSO cells, Syrian Hamster Ovary cell lines, HeLa cells, or human embryonic kidney cell lines (i.e. HEK293, HEK293EBNA).

[000186] Expression vectors for these cells can include expression control sequences, such as an origin of replication, a promoter, an enhancer, and necessary processing information sites, such as ribosome binding sites, RNA splice sites, polyadenylation sites, and transcriptional terminator sequences. Preferred expression control sequences are promoters derived from SV40, adenovirus, bovine papilloma virus, cytomegalovirus, Raus sarcoma virus, and the like. Preferred polyadenylation sites include sequences derived from SV40 and bovine growth hormone.

[000187] The vectors containing the polynucleotide sequences of interest (e.g., the Proteins of the invention and expression control sequences) can be transferred into the host cell by well-known methods, which vary depending on the type of cellular host. For example, calcium chloride transfection is commonly utilized for prokaryotic cells, whereas calcium phosphate treatment or electroporation may be used for other cellular hosts.

[000188] Various methods of protein purification may be employed and such methods are known in the art and described, for example, in Deutscher, Methods in Enzymology 182: 83-9 (1990) and Scopes, Protein Purification: Principles and Practice, Springer-Verlag, NY (1982). The purification step(s) selected will depend, for example, on the nature of the production process used for the Proteins of the invention.

[000189] The proteins, polypeptides, and/or peptides of the invention, e.g., the dual activity fusion proteins of the invention, should be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the

patient, the site of delivery of the protein compositions, the method of administration, the scheduling of administration, and other factors known to practitioners. The "therapeutically effective amount" of the proteins of the invention for purposes herein is thus determined by such considerations.

[000190] The pharmaceutical compositions of the proteins of the present invention may be administered by any means that achieve the generally intended purpose: to treat type 1 and type 2 diabetes mellitus, obesity, metabolic syndrome, or critically ill patients. Non-limiting permissible means of administration include, for example, by inhalation or suppository or to mucosal tissue such as by lavage to vaginal, rectal, urethral, buccal and sublingual tissue, orally, nasally, topically, intranasally, intraperitoneally, parenterally, intravenously, intramuscularly, intrasternally, by intraarticular injection, intralymphatically, interstitially, intra-arterially, subcutaneously, intrasynovial, transepithelial, and transdermally. In some embodiments, the pharmaceutical compositions are administered by lavage, orally or inter-arterially. Other suitable methods of introduction can also include rechargeable or biodegradable devices and slow or sustained release polymeric devices. The pharmaceutical compositions of this invention can also be administered as part of a combinatorial therapy with other known metabolic agents.

[000191] The dosage administered will be dependent upon the age, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired. Compositions within the scope of the invention include all compositions wherein an FGF21 variant is present in an amount that is effective to achieve the desired medical effect for treatment type 1 or type 2 diabetes mellitus, obesity, or metabolic syndrome. While individual needs may vary from one patient to another, the determination of the optimal ranges of effective amounts of all of the components is within the ability of the clinician of ordinary skill.

[000192] The proteins of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions. A desired formulation would be one that is a stable lyophilized product that is reconstituted with an appropriate diluent or an aqueous solution of high purity with optional pharmaceutically acceptable carriers, preservatives, excipients or stabilizers [Remington's Pharmaceutical Sciences 16th edition (1980)]. The proteins of the present invention may be combined with a pharmaceutically acceptable buffer, and the pH adjusted to provide acceptable stability, and a pH acceptable for administration.

[000193] For parenteral administration, in one embodiment, the proteins of the invention are formulated generally by mixing one or more of them at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a

pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. Preferably, one or more pharmaceutically acceptable anti-microbial agents may be added. Phenol, m-cresol, and benzyl alcohol are preferred pharmaceutically acceptable anti-microbial agents.

[000194] Optionally, one or more pharmaceutically acceptable salts may be added to adjust the ionic strength or tonicity. One or more excipients may be added to further adjust the isotonicity of the formulation. Glycerin, sodium chloride, and mannitol are examples of an isotonicity adjusting excipient.

[000195] Those skilled in the art can readily optimize pharmaceutically effective dosages and administration regimens for therapeutic compositions comprising Proteins of the invention, as determined by good medical practice and the clinical condition of the individual patient. A typical dose range for the proteins of the present invention will range from about 0.01 mg per day to about 1000 mg per day (or about 0.05 mg per week to about 5000 mg per week administered once per week) for an adult. Preferably, the dosage ranges from about 0.1 mg per day to about 100 mg per day (or about 0.5 mg per week to about 500 mg per week administered once per week), more preferably from about 1.0 mg/day to about 10 mg/day (or about 5 mg per week to about 50 mg per week administered once per week). Most preferably, the dosage is about 1-5 mg/day (or about 5 mg per week to about 25 mg per week administered once per week). The appropriate dose of an FGF21 variant administered will result in lowering blood glucose levels and increasing energy expenditure by faster and more efficient glucose utilization, and thus is useful for treating type 1 and type 2 diabetes mellitus, obesity and metabolic syndrome.

[000196] In addition, because hyperglycemia and insulin resistance are common in critically ill patients given nutritional support, some intensive care units (ICUs) administer insulin to treat excessive hyperglycemia in fed critically ill patients. In fact, recent studies document the use of exogenous insulin to maintain blood glucose at a level no higher than 110 mg per deciliter reduced morbidity and mortality among critically ill patients in the surgical intensive care unit, regardless of whether they had a history of diabetes (Van den Bergheet al. N Engl J Med., 345(19):1359, (2001)). Thus, proteins of the present invention are uniquely suited to help restore metabolic stability in metabolically unstable critically ill patients. Proteins of the invention such as those containing variants of FGF21 are unique in that they stimulate glucose uptake and enhances insulin sensitivity but do not induce hypoglycemia.

[000197] In another aspect of the present invention, proteins of the invention for use as a medicament for the treatment of obesity, type 1 and type 2 diabetes mellitus,

pancreatitis, dyslipidemia, nonalcoholic fatty liver disease (NAFLD), nonalcoholic steatohepatitis (NASH), insulin resistance, hyperinsulinemia, glucose intolerance, hyperglycemia, metabolic syndrome, acute myocardial infarction, conditions associated with severe inactivating mutations in the insulin receptor, lipodystrophies including HIV-associated lipodystrophy, and other metabolic disorders is contemplated.

[000198] Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

[000199] The practice of the present invention will employ, unless otherwise indicated, conventional methods of chemistry, biochemistry, molecular biology, immunology and pharmacology, within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Remington's Pharmaceutical Sciences, 18th Edition (Easton, Pennsylvania: Mack Publishing Company, 1990); Methods In Enzymology (S. Colowick and N. Kaplan, eds., Academic Press, Inc.); and Handbook of Experimental Immunology, Vols. I-IV (D.M. Weir and C.C. Blackwell, eds., 1986, Blackwell Scientific Publications); and Sambrook et al., Molecular Cloning: A Laboratory Manual (2nd Edition, 1989).

Site-Specific FGF21 Mutants

[000200] In some embodiments, the fusion proteins of the invention include additional mutants of GLP-1 or Exendin-4, GLP-1/Glucagon hybrid peptides, GLP-1 analogues with unnatural amino acids (to convey DPP-4 resistance, for PEGylation, or for other purposes).

[000201] In some embodiments, the proteins of the invention comprise FGF21 agonists with one or more of the following additional modifications of wild-type FGF21:

[000202] (i) additional disulfides, unnatural amino acids, or modifications to promote dimerization such as formation of a disulfide at R154C or introduction of a cysteine at another site, or dimerization through a fused Fc domain, or dimer formation through a cross-linker such as a bifunctional PEG;

[000203] (ii) fragments of FGF21;

[000204] (iii) proteins selected to have FGF21 activity (binding to beta-klotho and binding and activation of the FGFR's); and

[000205] (iv) an FGF21 mimetic antibody (of various formats such as Fab, unibody, svFc etc.).

[000206] In some embodiments, the dual activity proteins of the invention comprise one or more of the following linkers: a simple amide bond, short peptides (particularly Ser/Gly repeats), additional residues from the FGF21 translated sequence, or a larger

linker up to an entire protein (such as an Fc domain, an HSA-binding helix bundle, HSA, etc.). The two moieties can also be linked by other chemical means, such as through unnatural amino acids or standard chemical linkers (maleimide-Cys, NHS-Lys, "click chemistry", etc.) "Linker" for the FGF21 mimetic antibody approach could include those already listed and also an insertion into a loop with subsequent cleavage to release the GLP-1 N-terminus.

[000207] In some embodiments, the fusion protein of the invention comprises PEGylation occurring at one, two, or more specific sites. In preferred embodiments, the PEGylation occurs within the FGF21 molecule or the linker. In some embodiments, the PEGylation is not within ~10 amino acids of the N-terminus of the dual function protein, preferably not within ~10 amino acids of the start of FGF21. PEGylation attachment chemistries can include NHS-Lys, maleimide-Cys, unnatural amino acids (aldehyde, "click chemistry", Pcl, etc.) and can be combined with suitable protein variants to control the stoichiometry of the reaction.

[000208] The PEG group of the fusion proteins of the invention can be of any size (e.g., 1, 2, 3.4, 5, 10, 20, 24, 29, 30, 40 kDa), and can be linear, branched, or comb structured, with a preference for a total PEGylation of greater than or equal to 40 kDa. Optimal PEGylation achieves half-life extension sufficient for once weekly dosing. PEGylation of protein dimers, trimers, tetramers etc. may result in adequate serum half-life extension using shorter PEG polymers. Branched and comb PEG structures may be beneficial in terms of lower viscosity.

[000209] Preferred half-life extension methodologies for the fusion proteins of the invention include integrating an IgG Fc domain or HSA into the linker and may not require PEGylation. Additionally, using Fc domain fusions will result in dimerization and may result in enhanced potency in addition to half-life extension.

[000210] Other embodiments of the invention include but are not limited to the following attachments, for half-life extension: HSA-binding lipid or small molecule or micelle to either the monomeric or a dimeric version of the fusion.

[000211] In certain embodiments of the invention, other attachments may be made to proteins, polypeptides, and/or peptides of the invention, to achieve half-life extension and other improved biological properties. They can include attaching PEG-cholesterol conjugates (including micelles and liposomes) to the proteins, polypeptides, and/or peptides of the invention, and/or attaching sugars (glycosylate) to the proteins, polypeptides, and/or peptides of the invention. In still other embodiments, similar techniques are employed to add conjugates of , e.g., polysialic acid (PSA), hydroxyethyl starch (HES), albumin-binding ligands, or carbohydrate shields to proteins, polypeptides, and/or peptides.

[000212] The HESylation technique, for example, couples branched hydroxyethylstarch (HES) chains (60 kDa or 100 kDa, highly branched amylopectin fragments from corn starch) to a protein, polypeptides, and/or peptides via reductive alkylation. Polysialation conjugates proteins, polypeptides, and/or peptides of interest with polysialic acid (PSA) polymers in a manner similar to PEGylation. PSA polymers are negatively charged, non-immunogenic polymers that occur naturally in the body and are available in molecular weights of 10-50kD.

[000213] In still other embodiments of the invention, other attachments or modifications may be made to proteins, polypeptides, and/or peptides of the invention, to achieve half-life extension and other improved biological properties. These include the creation of recombinant PEG (rPEG) groups, and their attachment to the proteins, polypeptides, and/or peptides of the invention. As developed by the company Amunix, Inc. The rPEG technology is based on protein sequences with PEG-like properties that are genetically fused to biopharmaceuticals, avoiding the extra chemical conjugation step. rPEGs are extended half-life Exenatide constructs that contain a long unstructured tail of hydrophilic amino acids, and which are capable of both increasing a protein or peptide's serum half-life and slowing its rate of absorption, thus reducing the peak-trough ratio significantly. rPEGs have an increased hydrodynamic radius and show an apparent molecular weight that is about 15-fold their actual molecular weight, mimicking the way PEGylation achieves a long serum half-life. Similar recombinant polypeptide sequences are also developed by XL-protein GmbH for the "PASylation" of proteins.

Chemically-Modified Dual Function Protein Mutants

[000214] Chemically modified forms of the fusion proteins described herein, including, e.g., truncated and variant forms of the dual function fusion proteins described herein, can be prepared by one skilled in the art, given the disclosures described herein. Such chemically modified dual function proteins are altered such that the chemically modified mutant is different from the unmodified mutant, either in the type or location of the molecules naturally attached to the mutant. Chemically modified mutants can include molecules formed by the deletion of one or more naturally-attached chemical groups.

[000215] In one embodiment, proteins of the present invention can be modified by the covalent attachment of one or more polymers. For example, the polymer selected is typically water-soluble so that the protein to which it is attached does not precipitate in an aqueous environment, such as a physiological environment. Included within the scope of suitable polymers is a mixture of polymers. Preferably, for therapeutic use of the end-product preparation, the polymer will be pharmaceutically acceptable. Non-

water soluble polymers conjugated to proteins of the present invention also form an aspect of the invention.

[000216] Exemplary polymers each can be of any molecular weight and can be branched or unbranched. The polymers each typically have an average molecular weight of between about 2 kDa to about 100 kDa (the term "about" indicating that in preparations of a water-soluble polymer, some molecules will weigh more and some less than the stated molecular weight). The average molecular weight of each polymer is preferably between about 5 kDa and about 50 kDa, more preferably between about 12 kDa and about 40 kDa, and most preferably between about 20 kDa and about 35 kDa.

[000217] Suitable water-soluble polymers or mixtures thereof include, but are not limited to, N-linked or O-linked carbohydrates, sugars, phosphates, polyethylene glycol (PEG) (including the forms of PEG that have been used to derivatize proteins, including mono-(C1-C10), alkoxy-, or aryloxy-polyethylene glycol), monomethoxy-polyethylene glycol, dextran (such as low molecular weight dextran of, for example, about 6 kD), cellulose, or other carbohydrate based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, polypropylene oxide/ethylene oxide co-polymers, polyoxyethylated polyols (e.g., glycerol), and polyvinyl alcohol. Also encompassed by the present invention are bifunctional crosslinking molecules that can be used to prepare covalently attached dual function protein variant multimers. Also encompassed by the present invention are dual function protein variants covalently attached to polysialic acid.

[000218] In some embodiments of the present invention, a dual function protein variant is covalently, or chemically, modified to include one or more water-soluble polymers, including, but not limited to, polyethylene glycol (PEG), polyoxyethylene glycol, or polypropylene glycol. See, e.g., U.S. Pat. Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192; and 4,179,337. In some embodiments of the present invention, adual function protein comprises one or more polymers, including, but not limited to, monomethoxy-polyethylene glycol, dextran, cellulose, another carbohydrate-based polymer, poly-(N-vinyl pyrrolidone)-polyethylene glycol, propylene glycol homopolymers, a polypropylene oxide/ethylene oxide co-polymer, polyoxyethylated polyols (e.g., glycerol), polyvinyl alcohol, or mixtures of such polymers.

[000219] In some embodiments of the present invention, a dual function protein is covalently-modified with PEG subunits. In some embodiments, one or more water-soluble polymers are bonded at one or more specific positions (for example, at the N-terminus) of the dual function protein mutant. In some embodiments, one or more water-soluble polymers are randomly attached to one or more side chains of a dual function

protein mutant. In some embodiments, PEG is used to improve the therapeutic capacity of a dual function protein mutant. Certain such methods are discussed, for example, in U.S. Pat. No. 6,133,426, which is hereby incorporated by reference for any purpose.

[000220] In embodiments of the present invention wherein the polymer is PEG, the PEG group can be of any convenient molecular weight, and can be linear or branched. The average molecular weight of the PEG group will preferably range from about 2 kD to about 100 kDa, and more preferably from about 5 kDa to about 50 kDa, e.g., 10, 20, 30, 40, or 50 kDa. The PEG groups will generally be attached to the dual function protein mutant via acylation or reductive alkylation through a reactive group on the PEG moiety (e.g., an aldehyde, amino, thiol, or ester group) to a reactive group on the dual function protein mutant (e.g., an aldehyde, amino, or ester group).

[000221] Branched PEG derivatives, also known as "Y-shaped" PEG derivatives, contain two linear methoxy PEG chain attached to a central core. The sterically bulky structure of these "Y-shaped" PEG derivatives will facilitate the single point attachment of the modified molecules. By way of example, three kinds of "Y-shaped" PEG derivatives are Y-NHS-40K (useful for amine PEGylation); Y-MAL-40K (useful for thiol PEGylation); and Y-ALD-40K (e.g., Y-AALD-40K and Y-PALD-40K)(useful for N-terminal PEGylation). For amine PEGylation, the "Y-shape" NHS ester will react with the amino group of lysine(s) or the N-terminal amine in biological active molecules to produce a stable amide linkage(s). This NHS ester will couple with the targeted molecules at pH 7-8.5. For thiol PEGylation, the "Y-shape" maleimide will react with the thiol groups in biological active molecules to generate a stable 3-thiosuccinimidyl ether linkage. This maleimide will couple with the targeted molecules at an approximate pH of 7.4 in the presence of other functional groups. For N-terminal PEGylation, the "Y-shape" aldehyde will preferably react with the N-terminal amine in biological active molecules to produce a stable amine linkage in the presence of a reducing reagent such as sodium cyanoborohydride. This aldehyde will couple with the N-terminal amine of the targeted molecules at pH 5-8. Reagents for performing branched PEGylation are available through, e.g., JenKem Technology.

[000222] The PEGylation of a polypeptide, including the proteins of the invention, can be specifically carried out using any of the PEGylation reactions known in the art. Such reactions are described, for example, in the following references: Francis et al., 1992, Focus on Growth Factors 3: 4-10; European Patent Nos. 0 154 316 and 0 401 384; and U.S. Pat. No. 4,179,337. For example, PEGylation can be carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer) as described herein. For the acylation reactions, a selected polymer should have a single reactive ester group. For reductive

alkylation, a selected polymer should have a single reactive aldehyde group. A reactive aldehyde is, for example, polyethylene glycol propionaldehyde, which is water stable, or mono C1-C10 alkoxy or aryloxy derivatives thereof (see, e.g., U.S. Pat. No. 5,252,714).

[000223] In some embodiments of the present invention, a useful strategy for the attachment of the PEG group to a polypeptide involves combining, through the formation of a conjugate linkage in solution, a peptide and a PEG moiety, each bearing a special functionality that is mutually reactive toward the other. The peptides can be easily prepared with conventional solid phase synthesis. The peptides are "preactivated" with an appropriate functional group at a specific site. The precursors are purified and fully characterized prior to reacting with the PEG moiety. Ligation of the peptide with PEG usually takes place in aqueous phase and can be easily monitored by reverse phase analytical HPLC. The PEGylated peptides can be easily purified by preparative HPLC and characterized by analytical HPLC, amino acid analysis and laser desorption mass spectrometry.

[000224] Polysaccharide polymers are another type of water-soluble polymer that can be used for protein modification. Therefore, the Proteins of the invention fused to a polysaccharide polymer form embodiments of the present invention. Dextrans are polysaccharide polymers comprised of individual subunits of glucose predominantly linked by alpha 1-6 linkages. The dextran itself is available in many molecular weight ranges, and is readily available in molecular weights from about 1 kD to about 70 kD. Dextran is a suitable water-soluble polymer for use as a vehicle by itself or in combination with another vehicle (e.g., Fc). See, e.g., International Publication No. WO 96/11953. The use of dextran conjugated to therapeutic or diagnostic immunoglobulins has been reported. See, e.g., European Patent Publication No. 0 315 456, which is hereby incorporated by reference. The present invention also encompasses the use of dextran of about 1 kD to about 20 kD.

[000225] In general, chemical modification can be performed under any suitable condition used to react a protein with an activated polymer molecule. Methods for preparing chemically modified polypeptides will generally comprise the steps of: (a) reacting the polypeptide with the activated polymer molecule (such as a reactive ester or aldehyde derivative of the polymer molecule) under conditions whereby a FGF21 protein variant becomes attached to one or more polymer molecules, and (b) obtaining the reaction products. The optimal reaction conditions will be determined based on known parameters and the desired result. For example, the larger the ratio of polymer molecules to protein, the greater the percentage of attached polymer molecule. In one embodiment of the present invention, chemically modified FGF21 mutants can have a

single polymer molecule moiety at the amino-terminus (see, e.g., U.S. Pat. No. 5,234,784)

[000226] In another embodiment of the present invention, proteins of the invention can be chemically coupled to biotin. The biotinylated proteins of the invention are then allowed to bind to avidin, resulting in tetravalent avidin/biotin/proteins of the invention. Proteins of the invention can also be covalently coupled to dinitrophenol (DNP) or trinitrophenol (TNP) and the resulting conjugates precipitated with anti-DNP or anti-TNP-IgM to form decameric conjugates with a valency of 10.

[000227] Generally, conditions that can be alleviated or modulated by the administration of the present chemically modified dual function proteins mutants include those described herein for proteins of the invention. However, the chemically modified dual function proteins mutants disclosed herein can have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to unmodified dual function proteins mutants.

Therapeutic Compositions of Dual Function Proteins and Administration Thereof

[000228] Therapeutic compositions comprising the dual function proteins of the invention are within the scope of the present invention, and are specifically contemplated in light of, e.g., the identification of several mutant dual function proteins sequences exhibiting enhanced properties. Such dual function proteins mutant pharmaceutical compositions can comprise a therapeutically effective amount of a dual function proteins protein variant in admixture with a pharmaceutically or physiologically acceptable formulation agent selected for suitability with the mode of administration.

[000229] Acceptable formulation materials preferably are nontoxic to recipients at the dosages and concentrations employed.

[000230] The pharmaceutical composition can contain formulation materials for modifying, maintaining, or preserving, for example, the pH, osmolarity, viscosity, clarity, color, isotonicity, odor, sterility, stability, rate of dissolution or release, adsorption, or penetration of the composition. Suitable formulation materials include, but are not limited to, amino acids (such as glycine, glutamine, asparagine, arginine, or lysine), antimicrobials, antioxidants (such as ascorbic acid, sodium sulfite, or sodium hydrogen-sulfite), buffers (such as borate, bicarbonate, Tris-HCl, citrates, phosphates, or other organic acids), bulking agents (such as mannitol or glycine), chelating agents (such as ethylenediamine tetraacetic acid (EDTA)), complexing agents (such as caffeine, polyvinylpyrrolidone, beta-cyclodextrin, or hydroxypropyl-beta-cyclodextrin), fillers, monosaccharides, disaccharides, and other carbohydrates (such as glucose, mannose, or dextrins), proteins (such as serum albumin, gelatin, or immunoglobulins), coloring,

flavoring and diluting agents, emulsifying agents, hydrophilic polymers (such as polyvinylpyrrolidone), low molecular weight polypeptides, salt-forming counterions (such as sodium), preservatives (such as benzalkonium chloride, benzoic acid, salicylic acid, thimerosal, phenethyl alcohol, methylparaben, propylparaben, chlorhexidine, sorbic acid, or hydrogen peroxide), solvents (such as glycerin, propylene glycol, or polyethylene glycol), sugar alcohols (such as mannitol or sorbitol), suspending agents, surfactants or wetting agents (such as pluronics; PEG; sorbitan esters; polysorbates such as polysorbate 20 or polysorbate 80; triton; tromethamine; lecithin; cholesterol or tyloxapal), stability enhancing agents (such as sucrose or sorbitol), tonicity enhancing agents (such as alkali metal halides; preferably sodium or potassium chloride; or mannitol sorbitol), delivery vehicles, diluents, excipients and/or pharmaceutical adjuvants (see, e.g., Remington's Pharmaceutical Sciences (18th Ed., A. R. Gennaro, ed., Mack Publishing Company 1990), and subsequent editions of the same, incorporated herein by reference for any purpose).

[000231] The optimal pharmaceutical composition will be determined by a skilled artisan depending upon, for example, the intended route of administration, delivery format, and desired dosage (see, e.g., Remington's Pharmaceutical Sciences, supra). Such compositions can influence the physical state, stability, rate of in vivo release, and rate of in vivo clearance of the fusion protein of the invention.

[000232] The primary vehicle or carrier in a pharmaceutical composition can be either aqueous or non-aqueous in nature. For example, a suitable vehicle or carrier for injection can be water, physiological saline solution, or artificial cerebrospinal fluid, possibly supplemented with other materials common in compositions for parenteral administration. Neutral buffered saline or saline mixed with serum albumin are further exemplary vehicles. Other exemplary pharmaceutical compositions comprise Tris buffer of about pH 7.0-8.5, or acetate buffer of about pH 4.0-5.5, which can further include sorbitol or a suitable substitute. In one embodiment of the present invention, dual function pharmaceutical compositions can be prepared for storage by mixing the selected composition having the desired degree of purity with optional formulation agents (Remington's Pharmaceutical Sciences, supra) in the form of a lyophilized cake or an aqueous solution. Further, the dual function protein product can be formulated as a lyophilizate using appropriate excipients such as sucrose.

[000233] The pharmaceutical compositions containing the fusion proteins of the invention can be selected for parenteral delivery. Alternatively, the compositions can be selected for inhalation or for delivery through the digestive tract, such as orally. The preparation of such pharmaceutically acceptable compositions is within the skill of the art.

[000234] The formulation components are present in concentrations that are acceptable to the site of administration. For example, buffers are used to maintain the composition at physiological pH or at a slightly lower pH, typically within a pH range of from about 5 to about 8.

[000235] When parenteral administration is contemplated, the therapeutic compositions for use in this invention can be in the form of a pyrogen-free, parenterally acceptable, aqueous solution comprising the desired dual function protein in a pharmaceutically acceptable vehicle. A particularly suitable vehicle for parenteral injection is sterile distilled water in which a dual function protein is formulated as a sterile, isotonic solution, properly preserved. Yet another preparation can involve the formulation of the desired molecule with an agent, such as injectable microspheres, bio-erodible particles, polymeric compounds (such as polylactic acid or polyglycolic acid), beads, or liposomes, that provides for the controlled or sustained release of the product which can then be delivered via a depot injection. Hyaluronic acid can also be used, and this can have the effect of promoting sustained duration in the circulation. Other suitable means for the introduction of the desired molecule include implantable drug delivery devices.

[000236] In one embodiment, a pharmaceutical composition can be formulated for inhalation. For example, a dual function protein of the invention can be formulated as a dry powder for inhalation. Dual function protein inhalation solutions can also be formulated with a propellant for aerosol delivery. In yet another embodiment, solutions can be nebulized. Pulmonary administration is further described in International Publication No. WO 94/20069, which describes the pulmonary delivery of chemically modified proteins.

[000237] It is also contemplated that certain formulations can be administered orally. In one embodiment of the present invention, dual function proteins of the invention that are administered in this fashion can be formulated with or without those carriers customarily used in the compounding of solid dosage forms such as tablets and capsules. For example, a capsule can be designed to release the active portion of the formulation at the point in the gastrointestinal tract when bioavailability is maximized and pre-systemic degradation is minimized. Additional agents can be included to facilitate absorption of the dual function proteins of the invention. Diluents, flavorings, low melting point waxes, vegetable oils, lubricants, suspending agents, tablet disintegrating agents, and binders can also be employed.

[000238] Another pharmaceutical composition can involve an effective quantity of the proteins of the invention in a mixture with non-toxic excipients that are suitable for the manufacture of tablets. By dissolving the tablets in sterile water, or another appropriate

vehicle, solutions can be prepared in unit-dose form. Suitable excipients include, but are not limited to, inert diluents, such as calcium carbonate, sodium carbonate or bicarbonate, lactose, or calcium phosphate; or binding agents, such as starch, gelatin, or acacia; or lubricating agents such as magnesium stearate, stearic acid, or talc.

[000239] Additional pharmaceutical compositions comprising dual function proteins of the invention will be evident to those skilled in the art, including formulations involving dual function proteins of the invention in sustained- or controlled-delivery formulations. Techniques for formulating a variety of other sustained- or controlled-delivery means, such as liposome carriers, bio-erodible microparticles or porous beads and depot injections, are also known to those skilled in the art (see, e.g., International Publication No. WO 93/15722, which describes the controlled release of porous polymeric microparticles for the delivery of pharmaceutical compositions, and Wischke & Schwendeman, 2008, Int. J Pharm. 364: 298-327, and Freiberg & Zhu, 2004, Int. J Pharm. 282: 1-18, which discuss microsphere/microparticle preparation and use).

[000240] Additional examples of sustained-release preparations include semipermeable polymer matrices in the form of shaped articles, e.g. films, or microcapsules. Sustained release matrices can include polyesters, hydrogels, polylactides (U.S. Pat. No. 3,773,919 and European Patent No. 0 058 481), copolymers of L-glutamic acid and gamma ethyl-L-glutamate (Sidman et al., 1983, Biopolymers 22: 547-56), poly(2-hydroxyethyl-methacrylate) (Langer et al., 1981, J. Biomed. Mater. Res. 15: 167-277 and Langer, 1982, Chem. Tech. 12: 98-105), ethylene vinyl acetate (Langer et al., supra) or poly-D-3-hydroxybutyric acid (European Patent No. 0 133 988). Sustained-release compositions can also include liposomes, which can be prepared by any of several methods known in the art. See, e.g., Epstein et al., 1985, Proc. Natl. Acad. Sci. U.S.A. 82: 3688-92; and European Patent Nos. 0 036 676, 0 088 046, and 0 143 949.

[000241] The pharmaceutical compositions of the invention to be used for in vivo administration typically must be sterile. This can be accomplished by filtration through sterile filtration membranes. Where the composition is lyophilized, sterilization using this method can be conducted either prior to, or following, lyophilization and reconstitution. The composition for parenteral administration can be stored in lyophilized form or in a solution. In addition, parenteral compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

[000242] Once the pharmaceutical composition has been formulated, it can be stored in sterile vials as a solution, suspension, gel, emulsion, solid, or as a dehydrated or

lyophilized powder. Such formulations can be stored either in a ready-to-use form or in a form (e.g., lyophilized) requiring reconstitution prior to administration.

[000243] In a specific embodiment, the present invention is directed to kits for producing a single-dose administration unit. The kits can each contain both a first container having a dried protein and a second container having an aqueous formulation. Also included within the scope of this invention are kits containing single and multi-chambered pre-filled syringes (e.g., liquid syringes and lyosyringes).

Dosages of Dual Function Proteins and Administration Thereof

[000244] The effective amount of a pharmaceutical composition of the invention to be employed therapeutically will depend, for example, upon the therapeutic context and objectives. One skilled in the art will appreciate that the appropriate dosage levels for treatment will thus vary depending, in part, upon the molecule delivered, the indication for which the fusion protein variant is being used, the route of administration, and the size (body weight, body surface, or organ size) and condition (the age and general health) of the patient. Accordingly, the clinician can titer the dosage and modify the route of administration to obtain the optimal therapeutic effect. A typical dosage can range from about 0.1 µg/kg to up to about 100 mg/kg or more, depending on the factors mentioned above. In other embodiments, the dosage can range from 0.1 µg/kg up to about 100 mg/kg; or 1 µg/kg up to about 100 mg/kg.

[000245] The frequency of dosing will depend upon the pharmacokinetic parameters of the dual function protein in the formulation being used. Typically, a clinician will administer the composition until a dosage is reached that achieves the desired effect. The composition can therefore be administered as a single dose, as two or more doses (which may or may not contain the same amount of the desired molecule) over time, or as a continuous infusion via an implantation device or catheter. Further refinement of the appropriate dosage is routinely made by those of ordinary skill in the art and is within the ambit of tasks routinely performed by them. Appropriate dosages can be ascertained through use of appropriate dose-response data.

[000246] The route of administration of the pharmaceutical composition is in accord with known methods, e.g., orally; through injection by intravenous, intraperitoneal, intracerebral (intraparenchymal), intracerebroventricular, intramuscular, intraarterial, intraportal, or intralesional routes; by sustained release systems (which may also be injected); or by implantation devices. Where desired, the compositions can be administered by bolus injection or continuously by infusion, or by implantation device.

[000247] Alternatively or additionally, the composition can be administered locally via implantation of a membrane, sponge, or other appropriate material onto which the

desired molecule has been absorbed or encapsulated. Where an implantation device is used, the device can be implanted into any suitable tissue or organ, and delivery of the desired molecule can be via diffusion, timed-release bolus, or continuous administration.

Therapeutic Uses of Dual Function Proteins

[000248] Proteins of the invention can be used to treat, diagnose, ameliorate, or prevent a number of diseases, disorders, or conditions, including, but not limited to metabolic disorders. In one embodiment, the metabolic disorder to be treated is diabetes, e.g., type 2 diabetes mellitus. In another embodiment, the metabolic disorder is obesity. Other embodiments include metabolic conditions or disorders such as type 1 diabetes mellitus, pancreatitis, dyslipidemia, nonalcoholic fatty liver disease (NAFLD), nonalcoholic steatohepatitis (NASH), insulin resistance, hyperinsulinemia, glucose intolerance, hyperglycemia, metabolic syndrome, hypertension, cardiovascular disease, acute myocardial infarction, atherosclerosis, peripheral arterial disease, stroke, heart failure, coronary heart disease, kidney disease, diabetic complications, neuropathy, disorders associated with severe inactivating mutations in the insulin receptor, lipodystrophies including HIV-associated lipodystrophy, gastroparesis and other metabolic disorders.

[000249] In application, a disorder or condition such as type 1 or type 2 diabetes mellitus or obesity can be treated by administering a dual function protein variant as described herein to a patient in need thereof in the amount of a therapeutically effective dose. The administration can be performed as described herein, such as by IV injection, intraperitoneal injection, intramuscular injection, or orally in the form of a tablet or liquid formation. In most situations, a desired dosage can be determined by a clinician, as described herein, and can represent a therapeutically effective dose of the dual function protein polypeptide. It will be apparent to those of skill in the art that a therapeutically effective dose of dual function protein polypeptide will depend, inter alia, upon the administration schedule, the unit dose of antigen administered, whether the nucleic acid molecule or polypeptide is administered in combination with other therapeutic agents, the immune status and the health of the recipient. The term “therapeutically effective dose,” as used herein, means that amount of dual function protein polypeptide that elicits the biological or medicinal response in a tissue system, animal, or human being sought by a researcher, medical doctor, or other clinician, which includes alleviation of the symptoms of the disease or disorder being treated.

Pharmaceutical Compositions

[000250] The present invention also provides pharmaceutical compositions comprising one or more of the dual function proteins of the invention or mutants described herein and a pharmaceutically acceptable carrier. In some embodiments the pharmaceutical compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier. Pharmaceutically acceptable salts can also be present in the pharmaceutical composition, e.g., mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in Remington: The Science and Practice of Pharmacy (1995) Alfonso Gennaro, Lippincott, Williams, & Wilkins.

Fusion Proteins and Peptidic Compounds

[000251] In another embodiment, the proteins of the present invention can be made into a fusion protein or peptidic compound derived from the dual function proteins of the invention amino acid sequences. Such fusion proteins and peptidic compounds can be made using standard techniques known in the art. For example, peptidic compounds can be made by chemical synthesis using standard peptide synthesis techniques and then introduced into cells by a variety of means known in the art for introducing peptides into cells (e.g., liposome and the like).

[000252] The in vivo half-life of the fusion protein or peptidic compounds of the invention can be improved by making peptide modifications, such as the addition of N-linked glycosylation sites into the dual function proteins of the invention, or conjugating dual function proteins of the invention to poly(ethylene glycol)(PEG; pegylation), e.g., via lysine-monopegylation or cysteine-monopegylation. Such techniques have proven to be beneficial in prolonging the half-life of therapeutic protein drugs. It is expected that pegylation of the proteins of the invention of the invention may result in similar pharmaceutical advantages.

[000253] In addition, PEGylation can be achieved in any part of a polypeptide of the invention by the introduction of a non-natural amino acid. Certain non-natural amino acids can be introduced by the technology described in Deiters et al., J Am Chem Soc 125:11782-11783, 2003; Wang and Schultz, Science 301:964-967, 2003; Wang et al., Science 292:498-500, 2001; Zhang et al., Science 303:371-373, 2004 or in US Patent No. 7,083,970. Briefly, some of these expression systems involve site-directed mutagenesis to introduce a nonsense codon, such as an amber TAG, into the open reading frame encoding a polypeptide of the invention. Such expression vectors are

then introduced into a host that can utilize a tRNA specific for the introduced nonsense codon and charged with the non-natural amino acid of choice. Particular non-natural amino acids that are beneficial for purpose of conjugating moieties to the polypeptides of the invention include those with acetylene and azido side chains. The proteins of the invention containing these novel amino acids can then be PEGylated at these chosen sites in the protein.

[000254] Similarly PEGylation can be achieved in any part of a protein of the invention by the introduction of pyrrolysine or pyrroline-carboxy-lysine as described by Ou et al. (Proc Natl Acad Sci U S A. 2011 Jun 28;108(26):10437-42. Epub 2011 Jun 13).

EXAMPLES

Example 1: Design of GLP-1/Proteins of the invention.

[000255] In order to test the efficacy of GLP-1 and FGF21 together, a fusion molecule was designed. Because GLP-1 requires a free N-terminus and FGF21 requires a free C-terminus for receptor binding and activation, the two were cloned in the order of N-GLP-1-linker-FGF21-C. Initial constructs were made with GLP-1 residues 7-35 and with or without several modifications to the N-terminus for DPP-4 protection (discussed below). Further modifications of GLP-1 (point mutants or deletions) were tested and scored by in vitro potency.

[000256] Linkers of 3, 8, 10, or 20 amino acids were cloned. For FGF21, residues 33-209 of the wild-type human protein were used. Constructs with and without PEGylation sites were produced. PEGylation was accomplished using maleimide-PEG reagents (40 kDa linear and branched PEGs) reactive toward an introduced Cys. The Cys was placed at position R154 of FGF21 and at several sites in the GLP-1 or Exendin-4 and linker. Constructs have also been designed for modification through incorporation of a Pcl, Pyl, Pyl analog or a reactive non-naturally occurring amino acid by introducing a TAG (amber) codon at position R154 of FGF21 or K34 of GLP-1. Additional constructs were designed for two different modifications by inclusion of the Cys at R154 or K34 and a TAG codon for incorporation of a Pcl, Pyl, Pyl analog or a reactive non-naturally occurring amino acid at the other site. Constructs were made fusing GLP-1 and FGF21 to an Fc domain of a human antibody. Variant 76 (V76) sequence of FGF21 was introduced into the fusion format.

[000257] Expression and purification of the fusions: Fusion proteins were typically expressed in E. coli by cloning the DNA sequence encoding the particular dual function variant into a vector containing coding sequences for one of several removable domains (for example, NPro or NPro variants with or without His6 tag, or His6 tag with a TEV-

protease recognition peptide, or His6-Ubiquitin, or His6-Smt3) The vectors further included an inducible promoter to initiate mRNA transcription for protein expression such as the lac, T7 or arabinose promoters). Briefly, the vectors were transformed to DH10b-derived E. coli cells or BL21(DE3)-derived cells, grown under standard conditions, and induced to express protein with 0.2% arabinose or IPTG added to the culture media. Cells were harvested 3-4 hours post-induction, spun, and frozen. Pellets were thawed and resuspended for lysis by sonication and insoluble protein was isolated by centrifugation. The pellet was then solubilized in 6M guanidine, and the dissolved, clarified proteins were loaded to Ni-NTA columns. The columns were washed with denaturing and then non-denaturing buffers and eluted in a non-denaturing buffer according to standard practice. The eluates were buffer exchanged to remove imidazole, digested with specific proteases for removal of tags, (for example TEV protease, ubquitinase Usp2, or Ulp1 for specific removal of smt3), and purified by Ni-NTA. The fusion protein-containing fractions were further purified by size exclusion chromatography before PEGylation (typically with NOF Sunbright GL2-400MA 40kDa branched PEG maleimide). PEGylated proteins were isolated by anion exchange chromatography, placed in PBS, and concentrated or stored for various assays.

[000258] Protection of GLP-1 by mutation of the N-terminus: Because processing of the N-terminus of GLP-1 by DPP-4 is known to inactivate the peptide toward its primary receptor, several mutations reported in the literature were explored to slow this process. Constructs with the addition of an extra glycine to the N-terminus, mutation of Glu9 to Pro, or mutation of Ala8 to Gly or Ser, were all tested in a cell-based assay of GLP-1 activity. In the context of the fusion to FGF21 with a 40 kDa PEG attached to FGF21, the mutations at Ala8 retained superior activity to the other N-terminal modifications (Figure 1a), and the A8S mutation was used in subsequent designs. In addition, the DPP-4-resistant analogue, Exendin-4 was also tested in the fusions as an alternative moiety with extended half-life *in vivo* (Figure 1b). [See, e.g., J. Y. Ohet al. (2009) Bulletin of the Korean Chemical Society 30, 2471-2474; K. Adelhorst, (1994) JBC 269, 6275; B. D. Green et al. (2003) Biological Chemistry 384, 1543; J. C. Parkeret al. (1998) Journal of Peptide Research 52, 398; C. F. Deaconet al. (1998) Diabetologia 41, 271; R. Burcelinet al. (1999) Metabolism-Clinical and Experimental 48, 252; U. Ritzelet al. (1998) Journal of Endocrinology 159, 93.]

[000259] The A8S mutation was also compared to wild-type GLP-1 for pharmacokinetic properties *in vivo*. A fusion protein with wild-type GLP-1 (V329) and one with the GLP-1(A8S) mutation (V232) were injected intravenously into rats at 1 mg/kg, and the serum levels were measured against standard curves produced with the dosing solutions. When measured with a human FGF21-reactive ELISA kit, the

molecules appeared to behave similarly to each other and to FGF21 alone with the same size (40 kDa) PEG attached (Figure 2). To measure the half-life extension of GLP-1 activity, C57BL/6J mice were injected with the same two dual function protein variants with either wild-type GLP-1 or GLP-1(A8S). While both acutely lowered the fed glucose level after dosing, the GLP-1(A8S) variant retained greater activity in an oral glucose tolerance test (OGTT) measurement 3 days after dosing (Figure 3). These data demonstrate that the A8S mutation increases the effective half-life of the GLP-1 moiety of the fusion while the PK of the FGF21 moiety is not much affected by the addition of GLP-1.

[000260] Design of the peptide linker between GLP-1 and FGF21: Linkers of 3, 8, 10, or 20 amino acids were tested. No significant difference in activity was observed for the variants with different linkers, although there were differences in expression yields in some fusion contexts. Additionally, constructs were tested with various lengths of GLP-1 (8 to 31 residues) or Exendin-4 (30 or more residues) and tested for in vitro activity in the GLP-1R cell-based assay.

Example 2: Test of fusions in the ob/ob and db/db diabetic mouse models.

[000261] FGF21 has been shown to improve blood glucose levels, liver lipid levels, and body weight in the *ob/ob* mouse model of type-2 diabetes (T2D; see, e.g., A. Kharitonenkovet al. (2005) Journal of Clinical Investigation 115, 1627; T. Coskun et al. (2008) Endocrinology 149, 6018; and E. D. Berglund et al. (2009) Endocrinology 150, 4084). Likewise, GLP-1 analogues have been shown to improve glucose control, beta cell function and liver health in this genetic mouse diabetes model (Gallwitz B., Glucagon-like peptide-1 as a treatment option for type 2 diabetes and its role in restoring beta-cell mass. *Diabetes Technol Ther*. 2005, 7:651-7).

[000262] To determine if the fusion of GLP-1 to FGF21 would lead to additional efficacy, the WT GLP-1 and GLP-1(A8S) versions were tested against an equivalent FGF21 molecule each with the same 40 kDa branched PEG attached. In a two week study with twice weekly dosing, the FGF21(R154C)-PEG (V238) and GLP-1-FGF21(R154C)-PEG (V239) showed similar effects on blood glucose, body weight, and liver health (assessed by alanine aminotransferase (ALT) serum levels, weight, and appearance) (Figure 4). Glucose measurements showed a faster improvement for the fusion than for FGF21 alone after the first dose. The two converged at time points more than a day after dosing, suggesting that the additional benefit due to GLP-1 was shorter lived than the FGF21 efficacy.

[000263] The GLP-1(A8S)-FGF21(R154C)-PEG (V235) fusion also showed similar effects at 0.2 mg/kg, rather than the 1 mg/kg used for the other compounds. This

compound also generally gave more consistent results between dosing, suggesting that the DPP-4-resistance was critical to the overall improvement in efficacy as compared to the up-and-down behavior of GLP-1-FGF21(R154C)-PEG (V239). It was also observed that the GLP-1(A8S)-FGF21(R154C)-PEG (V235) at 1 mg/kg showed additional lowering of glucose, body weight, ALT/AST, and liver weight.

[000264] Additional studies were conducted to see if the improved efficacy could be replicated with co-administration of separate FGF21 and GLP-1 compounds. To make dosing of proteins similar between groups, a tool compound was made in which 14 amino acids were removed from the FGF21 C-terminus. This molecule (V253) was at least 1000x less potent in the FGF21 receptor assay but retained equal potency in the GLP-1R assay as compared to the other fusions and was unable to inhibit wild-type FGF21 activity. This compound had pharmacokinetics in rat similar to the other fusions when assessed by FGF21 levels (Figure 2).

[000265] Results of a four week study in which *ob/ob* mice were treated with FGF21(V76)-PEG, GLP-1(A8S)-PEG (V253), both together, or GLP-1(A8S)-FGF21(V76)-PEG fusion (V272) are shown in Figure 5. The fusion was significantly more efficacious than co-treatment of the individual FGF21(V76)-PEG and GLP-1(A8S)-PEG molecules on fed glucose AUC, OGTT glucose AUC, body weight at the end of the study, hepatic lipid content, and ALT measurement (p-values < 0.05 for all) .

[000266] Serum exposure of human FGF21 and GLP-1 were checked in the terminal serum samples to confirm exposure in the various groups. Human FGF21 was detected in all treated groups, and a significant increase in active GLP-1 was measured in all groups treated with GLP-1 (A8S)-PEG (V253) or GLP-1(A8S)-FGF21(V76)-PEG (V272) . The GLP-1(A8S)-FGF21(V76)-PEG (V272) group showed similar levels of human FGF21 to the 0.2 mg/kg FGF21(V76)-PEG group, suggesting that the improved efficacy is not due to a higher systemic accumulation of the fusion in the animals. The GLP-1(A8S)-PEG (V253) molecule was fully active *in vitro* (Table 8). These data demonstrate that the efficacy of the fusion was dependent on both moieties being in a single molecule, and, e.g., benefitted from a synergy at the cellular level.

[000267] To further test GLP-1(A8S)-FGF21(V76)-PEG (V272) vs. co-administration, three studies were conducted in db/db mice. In each study, male db/db mice (8-11 weeks old were dosed i.p. twice weekly. In an initial three week study, the dual function fusion protein, GLP-1(A8S)-FGF21(V76)-PEG (V272), matched the glucose lowering of the co-administration at equal doses and surpassed the efficacy of single entity dosing. The dual function fusion protein also showed superior body weight loss.

[000268] In a second two week study, GLP-1(A8S)-FGF21(V76)-PEG (V272) showed similar glucose lowering compared to FGF21(V76)-PEG at a 25-fold lower dose and

GLP-1(A8S)-PEG (V253) at a 5-fold lower dosing. In addition, GLP-1(A8S)-FGF21(V76)-PEG (V272) lowered HbA1c significantly more and lowered body weight significantly more than the mono-therapies or the maximally efficacious dose of the co-administration groups (Figure 8, Table 2). Unexpectedly, altering the dosing ratio of the co-treatments higher than 1:1 did not improve the observed efficacy (Table 2). This suggests that dual function proteins including equal numbers of GLP-1 moieties and FGF21 moieties can achieve optimal efficacy during in vivo treatments.

Table 2. Therapeutic Dosages

Treatment	Dose (mg/kg 2x/wk)	HbA1c, change from initial (%)	Body weight, day 12 (g)
Vehicle	N/A	1.6±0.9	42.3±2.1
V76	1	2.2±1.0	43.3±2.1
V76	3	1.0±0.9	42.3±1.3
V76	5	0.7±0.4[1]	42.1±1.9
V253	1	0.8±0.7[1]	42.6±2.1
V253	3	0.6±0.8[1]	40.9±2.6
V253	5	-0.1±0.6[1]	40.4±2.1
V76+V253	1+1	-0.3±0.6[1]	37.7±2.2[1]
V76+V253	1+3	0.5±0.6[1]	37.7±1.7[1]
V76+V253	3+1	-0.1±0.5[1]	38.9±1.2[1]
V76+V253	3+3	-0.3±0.4[1]	36.0±1.7[1]
V272	0.2	0.4±1.1[1]	41.1±2.7
V272	1	-1.1±0.5[1,2]	33.8±2.1[1,2]

[000269] [1] p-value vs. vehicle < 0.05. [2] p-value vs. co-treatment groups < 0.05

[000270] Table 2 shows efficacy of mono-therapies and co-treatments at different doses compared to dual function proteins in a db/db mouse study with 2 weeks of dosing. The reported values are averages of measurements made in eight animals with standard deviation. In a four week study, male db/db mice (n=8) were dosed twice weekly for four weeks with mono-therapies, co-treatments, or dual function fusion proteins. On day 27, the mice were fasted and the blood insulin levels were measured for all animals. At the end of the study, pancreatic insulin contents were extracted and measured (n=3). GLP-1(A8S)-FGF21(V76)-PEG (V272) showed a robust dose response from 0.1-1 mg/kg with efficacy equal to or better than FGF21(V76)-PEG or GLP-1(A8S)-PEG (V253) alone on glucose levels, HbA1c, and body weight at more than 10-fold lower dose, and improved efficacy over the maximally efficacious dosing of the co-administration. GLP-1(A8S)-FGF21(V76)-PEG (V272) also lowered fasting

insulin levels in the serum while increasing insulin content of the pancreas to a greater extent than the single or combination treatments of GLP-1(A8S)-PEG (V253) and FGF21(V76)-PEG (Figure 7). GLP-1(A8S)-FGF21(V76)-PEG (V272) -treated animals were close to lean control mice in fed glucose, HbA1c, and body weight numbers.

[000271] To estimate metabolic rate and substrate utilization, we measured oxygen consumption and carbon dioxide production using an Oxymax indirect calorimetry system (Columbus Instruments, Columbus, OH). Mice were housed in the chamber with a 12-h light/12-h dark cycle in an ambient temperature of 22–24°C. VO_2 and VCO_2 rates were determined under Oxymax system settings and protocol. The system was calibrated against a standard gas mixture to measure O_2 consumed (VO_2, ml/kg/h) and CO_2 generated (VCO_2, ml/kg/h). Metabolic rate (VO_2) and respiratory exchange ratio (RER) (ratio of VCO_2/VO_2) were evaluated over a 72-h period. Calculated RER values indicated that animals treated with fusion protein relied more on lipid substrates (RER values close to 0.75-0.85) particularly during the 24-72 hours after the second dose when compared to vehicle or combination treated groups. During this period, vehicle and combination treated groups showed RER values were close to 0.9 suggesting more of carbohydrate substrate utilization for energy and to a lesser extent lipid substrate utilization for energy expenditure. The data suggest that fusion protein treatment caused an increase in fatty acid oxidation that may be contributing to its body weight reduction effect in db/db mice

[000272] A dual function fusion protein, when compared to combination, showed improved beta-cell function in vivo. After four doses of either dual function protein GLP-1(A8S)-FGF21(V76)-PEG (V272) or GLP-1(A8S)-PEG (V253) plus FGF21(V76)-PEG, db/db mice were dosed orally with glucose and arginine. The glucose excursion of the dual function protein-treated group (72% lower than vehicle) was significantly lower than the combination-treated group (56% lower than vehicle). During the experiment, the amount of insulin secreted was also higher for the dual function protein-treated group (279% higher than vehicle) than for the combination-treated group (122% higher than vehicle).

[000273] Overall, the figures and tables, as well as data not shown herein, demonstrate the ability of the dual function fusion proteins of the invention, e.g., GLP-1(A8S)-FGF21(V76)-PEG (V272), to improve metabolic parameters for in vitro and in vivo rodent models of diabetes, when compared to a combination of individual GLP-1 and FGF21 (e.g.,FGF21(V76)-PEG and GLP-1(A8S)-PEG (V253)). Said improved parameters (as used herein, “metabolic parameters”) include but are not limited to fed glucose (AUC), body weight, liver triglycerides, plasma HbA1c, serum triglyceride levels

total cholesterol levels, oral glucose tolerance test serum glucose measurement (AUC), fasting serum insulin, pancreatic insulin content, and body fat percentage.

Example 3: Design of GLP-1/Proteins of the invention from FGF21 Variants.

[000274] A fusion was made with the following FGF21 variant sequence, which we refer to herein as "Variant 76" or "V76":

```
1    DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ
61   ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP
121  HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS
(SEQ ID NO:130)
```

[000275] FGF21(V76)-PEG features a 40 kDa branched PEG linked through Cys154, and eight point mutations relative to the 177 amino acid wild-type protein (Q56E, D74H, Q82E, R105K, K150R, R154C, R159Q, S195A, all made relative to full-length FGF21 protein sequence (NCBI reference sequence number NP_061986.1)). The risk of clinical immunogenicity toward this molecule is considered low based on an EpiScreen time course assay of human T-cell responses. The molecule has a serum half-life time of more than 30 h in mouse and rat and significantly lowers glucose AUC in *ob/ob* mice with twice-weekly injections of 1 mg/kg.

[000276] The GLP-1(A8S)-FGF21(V76)-PEG (V272) molecule was tested against the GLP-1(A8S)-FGF21(R154C)-PEG (V235) fusion at three doses (0.05, 0.1 and 0.2 mg/kg). Both fusions showed a dose response in lowering glucose AUC, body weight, food intake, and ALT(Figure 6). GLP-1(A8S)-FGF21(V76)-PEG (V272) was generally equal or better and showed higher efficacy for all parameters at the 0.2 mg/kg dose. GLP-1(A8S)-FGF21(V76)-PEG (V272) also showed lowering of serum triglycerides and cholesterol at both 0.1 and 0.2 mg/kg. Based on these data, the GLP-1(A8S)-FGF21(V76)-PEG (V272) molecule shows similar or improved properties to the initial fusions and is suitable for further study and development.

[000277] Ex4(1-30)-L20-FGF21(V76)-PEG (V277) also showed similar efficacy to GLP-1(A8S)-FGF21(V76)-PEG (V272) in an ob/ob 2 week study for glucose control, body weight and lipid levels.

[000278] Choice of PEGylation site for half-life extension. As seen for example in Figure 2, PEGylation was used to extend the half-life of the molecule from a few minutes for GLP-1 or less than an hour for FGF21 to longer than 30 hours as measured in rats. To see if the two moieties of the fusion could be modulated by placement of the PEG, a series of constructs was made with the extra cysteine not in the FGF21 sequence but in

that of GLP-1 or the linker. These constructs were tested for both GLP-1 and FGF21 activity in cell-based assays. Experiments reveal a stretch of positions within the fusion protein sequences at which PEGylation does not greatly impact either activity. Careful placement of the PEG at the N-terminal end of this stretch could be used to make a molecule with knocked-down GLP-1 activity while placement in the linker could be used to knock-down FGF21 activity. Such a molecule may be useful in tuning the potency of the fusion (for example, if the high potency of GLP-1 or Exendin-4 resulted in a poor therapeutic window when dosed at the efficacious level of a lower potency FGF21).

[000279] Based on these data, fusions with PEGylation at K34C of GLP-1 (V273) or K27C of Exendin-4 (V274) were prepared with FGF21(V76-C154R) whose sequence is as follows:

DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEEGY NVYQSEAHGL PLHLPGNRSP HRDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS (SEQ ID NO:131)

[000280] A two week study in *ob/ob* mice demonstrates that, although the new fusions (V273 and V274) exhibited potency *in vitro* similar to the previous fusions and similar pharmacokinetics in rat, they were not as efficacious *in vivo*. At 0.2 mg/kg, the GLP-1 fusion variant showed no significant efficacy for glucose or body weight. The Exendin-4 fusion variant showed significant lowering of ALT and body weight and a trend toward lower AST compared to vehicle but no significant effect on blood glucose. Although these fusions may show efficacy at higher doses, the original format with PEGylation at R154C in FGF21 was chosen for further study. It is unclear if this unexpected difference indicates a special property of the GLP-1(A8S)-FGF21(V76-154R)-PEG version of the fusion that is blocked by PEGylation at the alternate site or if there is another explanation such as the PEG blocking the GLP-1R interaction more on the native cells with native receptor levels than on the transfected cells used for the *in vitro* assay.

Example 4: Further Characterization of GLP-1/FGF21 Proteins of the invention.

[000281] In order to further optimize and rank the GLP-1-FGF21-PEG dual activity proteins of the invention, and to gain a fuller appreciation of their improved efficacy over co-administration models, the fusions are tested in the *ob/ob* model of efficacy. These include but are not limited to molecules with longer or shorter linkers, additional proteins of the invention (which may include but are not limited to mutants based on Variant 76

or additional variants in development such as additional or different point mutations, insertions or deletions, dimerized molecules, alternative PEGylation strategies, molecules with additional disulfide bridges, molecules produced in different expression systems), and variants of Exendin-4 or GLP-1 to improve serum stability or activity. Candidates are also filtered by in silico immunogenicity prediction, expression levels, and quality of final products (related to solubility, aggregation and stability).

[000282] In order to further characterize the synergy of FGF21 and GLP-1 for treatment of diabetes and obesity, interaction dynamics of the two molecules are mapped, and the fusion proteins are compared to the maximally efficacious co-administration of the individual parent molecules. These experiments are conducted in models that focus on obesity and weight loss (Diet Induced Obesity in rat or mouse) and those that better model aspects of diabetes such as hyperglycemia, dislipidemia, and impaired beta cell function (*ob/ob* or *db/db* mice). The extent of synergy achieved by the dual activity proteins of the invention can be further determined by mapping the efficacy of each individual molecule, the combination of the two at different ratios, and the fusion.

[000283] Testing in cell-based assays with primary or immortalized rat islets and primary human islets is conducted to assess the efficacy of the fusion for cell proliferation, protection from apoptosis, and function, e.g., to validate the dual activity proteins of the invention in accepted models of Type 1 Diabetes (see further Van Belle, T.L. et al. Drug Discovery today: disease models. 209 pp. 41-45). Preferred *in vivo* models are those featuring pancreatic ablation by streptozotocin (STZ); targeted beta-cell ablation in a genetically modified and inducible mouse strain (RIP-DTA) in which the beta-cells are destroyed by supplementing the diet with doxycycline to induce targeted expression of diphtheria toxin; or the non-obese diabetic (NOD) mouse model of Type-1 diabetes with an autoimmunity mechanism. In these models, testing is possible in both prophylactic dosing to assess beta-cell protection (particularly in NOD) as well as post-challenge dosing to assess beta-cell stimulation and proliferation (particularly in STZ and RIP-DTA).

[000284] The following experiments are conducted to more fully realize the mechanism of efficacy synergy achieved by the dual activity proteins of the invention over the co-administration regimens: Cells co-transfected with both receptors (GLP-1R; FGFR1(IIIc), 2(IIIc), 3(IIIc), or 4; and beta-klotho) can be used to determine if the receptors are able to potentiate the opposing signal at the cell surface. Crosstalk between downstream signals can be detected in cells naturally expressing both receptors (e.g., beta cells). In animals, the contribution of food intake (investigated with pair-fed cohorts), increased metabolic rate (investigated in clamp and metabolic

chamber experiments), and other mechanisms could be more thoroughly investigated to elucidate a mode of action for the synergistic effects of the two signals. Gene expression profiling of key tissues (liver, pancreas, adipose, intestine, heart, aorta, brain, etc.) may also be conducted in order to elucidate the unique signaling of the fusion proteins, potentially accounting for their improved efficacy.

[000285] The following experiments are conducted to understand the activity of dual function proteins of the invention to control or modify differentiation of mesenchymal stem cells (MSC): MSC can be treated with proteins of the invention alone or with compounds or mixtures of compounds known to induce osteogenesis or with compounds or mixtures of compounds known to induce adipogenesis and the resulting rates of differentiation to osteoblast-like cells or adipocyte-like cells can be measured.

Example 5: Design of GLP-1 and Exendin-4-containing proteins of the invention from Fc-fusion FGF21 variants.

[000286] Constructs were made with alternative linkers (for example, SGGGGSGGGGSGGGGSA (SEQ ID NO:138), GGGGS (SEQ ID NO:173), GG, and GGGGSGGGGSGGGGS (SEQ ID NO:174)), FGF21 variants, and GLP-1-related peptides. Typically the dual function proteins containing an Fc domain were expressed in HEK293 cells. The DNA coding sequences of these fusions were cloned into vectors containg sequence encoding a leader peptide to direct the proteins for secretion and further containing sequences necessary to promote mRNA syntehsis and protein expression of the desired products. HEK293 cells were transiently transfected with the vectors. The media from these cell cultures were collected, filtered, and purified by protein A affinity chromatography. The eluates were brought to netural pH and further purified by size exclusion chromatography. The products were then concentrated and stored for various assays. Variants were tested for activity in vitro in four assays: CRE-luciferase expression induction in HEK293 transfected with GLP-1R (Table 3) and glucose stimulated insulin secretion in INS1E cells (Table 6) to measure activity through GLP-1R; ERK phosphorylation induction in HEK293 transfected with beta-klotho (Table 4) and 2-deoxyglucose uptake by 3T3L1 cells (Table 5) to measure FGF21 activity. Variants with GLP-1(A8S) and a 15 amino acid linker were more active than variants with GLP-1 and shorter linkers. Exendin-4 (1-39) variants with all linkers tested were active. Exendin-4 (1-30)-containing variants had similar in vitro potency to Exendin-4 (1-39)-containing constructs. A variant with a tandem repeat of Exendin-4(1-39) was not as potent as variants with a single copy of a GLP-1-related peptide.

Table 3. GLP-1R activation in HEK293-GLP-1R-CRE-Luciferase Reporter Cells.

Variant#	GLP-1R EC_{50} (nM)
V193	0.024
V194	0.042
V195	0.041
V196	0.027
V197	0.085
V198	0.024
V199	0.046
V202	0.025
V203	0.252
V206	0.047
V207	0.088
V208	0.044
V209	0.033
V210	0.038
V211	0.032
V212	5.90
V213	18.4
V214	0.023
V215	8.99
V216	0.025
V217	4.53

Table 4. Activity measured in an ERK phosphorylation assay of FGF21 activity in HEK293-KLB cells.

Variant#	βklotho pERK EC_{50} (nM)	% y-max vs. WT FGF21	Potency (Fold over WT FGF21)
V193	2.1	101	1.2
V194	9.8	97	5.8
V195	13	76	7.6
V196	0.87	58	0.18
V197	0.92	63	0.19
V198	0.94	67	0.24
V199	1.04	58	0.27
V202	1.78	67	0.37
V203	1.01	70	0.21
V206	0.51	83	0.15
V207	1.10	83	0.34
V208	0.53	54	0.20
V209	0.70	83	0.21
V210	0.84	88	0.24
V211	0.58	73	0.17
V212	1.24	89	0.37
V213	1.00	61	0.37
V214	0.55	50	0.21
V215	0.92	51	0.34

V216	0.85	56	0.32
V217	2.23	46	0.83

Table 5. Activity measured in a 2-deoxyglucose uptake in 3T3L1 mouse adipocytes.

Variant#	2DOG EC_{50} (nM)	% y-max vs. WT FGF21	Potency (Fold over WT FGF21)
V194	2.8	118	1.6
V196	0.055	90	0.14
V197	0.077	100	0.18
V198	0.079	105	0.21
V199	0.090	101	0.37
V202	0.45	86	2.5
V203	0.078	100	0.31
V206	0.093	85	0.34
V208	0.067	87	0.25
V209	0.157	91	0.59
V210	0.067	98	0.27
V211	0.073	99	0.29
V214	0.213	87	0.81
V216	0.137	80	0.56

Table 6. Activity measured by glucose-stimulated insulin secretion (GSIS) in INS1E rat insulinoma cells.

Variant#	GSIS EC50 (nM) (avg of 2)
V196	9.1
V197	46
V198	4.4
V199	5.0
V202	33
V203	128
V206	2.4
V208	1.6
V209	1.9
V210	2.7
V211	1.5
V214	4.6
V216	6.8

[000287] Eight ob/ob mice were dosed with variants twice per week at 0.5 mg/kg for two weeks. Blood glucose and body weight were measured during the studies, and hepatic lipids were measured at the end of the study. As seen in the following table (Table 7), treated animals had reduced blood glucose, lower body weight, and lower

hepatic lipids throughout the study, when compared to vehicle-treated age-matched animals.

Table 7. Summary of results of treating ob/ob mice twice per week at 0.5 mg/kg for two weeks.

Variant#	BW loss, D12 (% of Veh)	Fed AUC lowering (% of Veh)	TG FASTED (8h, or O/N) (% of Veh)	Hepatic Lipids (% of Veh)
V193	-20*	-55*	39	-66*
V195	-7*	-30*	N/A	-10
V196	-23*	-63*	8	-65*
V197	-16*	-53*	52	-85*
V198	-24*	-64*	-7	-69*
V199	-24*	-64*	-7	-69*
V200	-14*	-39*	-5	-62*
V201	-22*	-60*	-10	-58*
V202	-18*	-59*	22	-63*
V203	-8*	-33*	68*	-73*
V206	-16*	-58*	-19	-56*
V207	-22*	-64*	-10	-55*
V208	-21*	-60*	10	-57*
V209	-28*	-66*	-11	-62*
V210	-23*	-67*	-14	-57*
V211	-28*	-68*	0	-52*
V214	-19*	-58*	110*	-41*
V216	-21*	-62*	29	-51*

* p-value vs. vehicle < 0.05

Example 6: Further characterization of proteins of the invention in receptor pharmacology assays.

[000288] As seen, for example, in Figure 9, dual agonist proteins of the invention were compared to various forms of FGF21 and Exenatide for their ability to induce ERK phosphorylation in HEK293 transfected with beta-klotho. At 10 minutes, cells treated with GLP-1(A8S)-FGF21(V76)-PEG (V272) showed a higher potency and higher maximal proportion of phospho-ERK to total ERK when compared to cells treated with FGF21(V76)-PEG plus Exenatide. The higher maximal phospho-ERK signal shows that the dual function protein V272 is a superior agonist for signaling through FGFR1c/beta-klotho when compared to the FGF21(V76)-PEG molecule, possibly due to better

interaction with the receptor (for example, the N-terminal extension causing the dual function protein to attain a more favorable structure for receptor binding or activation).

[000289] Select dual function proteins were tested for their ability to induce phospho-ERK signaling in human adipocytes. When compared to a single function FGF21 molecule (V101), the dual function proteins V208, V209, V211, V14, and V272 were more potent for stimulating phosphorylation of ERK. Because these are human cells, it may also suggest that the dual function proteins of the invention will be highly active in treatment of human disease when compared to treatments including other FGF21 mimetics, similar to what has been demonstrated for activity in rodents to date.

[000290] As seen, for example, in Table 8 and Figure 10, dual agonist proteins of the invention were compared to Exenatide, GLP-1 peptide, Exendin-4(1-39)-L5-Fc (V201), and GLP-1(A8S)-PEG (V253) for their ability to signal through GLP-1R in HEK293 cells transfected with GLP-1R or cells co-transfected with GLP-1R, beta-klotho and FGFR1c. Cells were treated with compound for 30 minutes, and cAMP levels were measured. Table 8 shows that in cells with both the GLP-1R and the FGF21 receptor complex, the dual agonist proteins (V272 and V211) showed higher potency than the single agonists (V253 and V101) alone or in combination with Exendin-4. In cells with only GLP-1R, the potency of the single agonists alone or in combination with Exendin-4 was equal to that of the dual agonist proteins. FGF21 variants alone were inactive in the assay.

[000291] Table 8. Assay of cAMP induction in cells treated with dual function proteins of the invention.

	GLP-1R/FGFR1c/beta-klotho			GLP-1R only		
Compound	EC50 (nM)	EC50 (fold-Ex-4)	y-max (%Ex-4)	EC50 (nM)	EC50 (fold-Ex-4)	y-max (%Ex-4)
Ex-4	6.0	1.0	100	11	1.0	100
V253	24	4.1	92	34	3.2	81
V253+V76	12	2.1	91	38	3.5	82
V272	2.4	0.4	97	64	6.0	76
V201	26	4.4	93	32	3.0	100
V201+V101	15	2.4	98	18	1.7	82
V211	6.6	1.1	92	39	3.6	98

[000292] Two additional HEK293 cell lines were generated in Discoverx's PathHunder assay format which utilizes complementation of beta-galactosidase fragments on beta-arrestin and the C-terminus of GLP-1R to measure recruitment of arrestin upon receptor activation. Figures 10a-10d show that with a 1 hour incubation of cells with protein or peptide, GLP-1(A8S)-FGF21(V76)-PEG (V272), Exendin-4(1-39)-L15-Fc-L15-

FGF21(V103) (V211), GLP-1(A8S)-PEG (V253), and Exendin-4(1-39)-L5-Fc (V201), were less potent than Exenatide for recruitment of beta-arrestin and that no difference in behavior was noted whether cells contained GLP-1R/FGFR1c/beta-klotho or GLP-1R alone.

[000293] The receptor pharmacology studies presented here suggest that the dual function proteins of the invention will be superior in activity and efficacy when compared to combinations or co-treatments of conventional FGF21 and GLP-1/Exenatide molecules. The PEGylated dual function protein GLP-1(A8S)-FGF21(V76)-PEG (V272) unexpectedly showed higher potency and maximal signal for phospho-ERK induction when compared to the FGF21(V76)-PEG molecule. These data show that the dual function proteins are equivalent or superior for FGF21 signaling when compared to other FGF21 variants. Likewise, the dual function proteins exhibit higher potency for cAMP signaling downstream of GLP-1R when cells express both the GLP-1 and FGF21 receptors. These properties may reflect a direct interaction between the receptors or an indirect effect such as a boost in local concentration through avidity toward the two receptors. These observations suggest a mechanism for improved activity of the dual function proteins that may explain why co-treatments of individual GLP-1 and FGF21 variants have not been sufficient to match the in vivo efficacy of the dual function proteins, as presented in other examples of this invention.

Example 7: Characterization of proteins of the invention in models of Type-1 Diabetes (T1D)

[000294] RIP-DTA transgenic mice (9-10 weeks old) were treated with doxycycline for 3 days to induce controlled beta-cell ablation (similar to Thorel et al., Nature 2010(464)1149-1154). The mice were then treated for three weeks, twice per week with FGF21(V76)-PEG, FGF21(V76)-PEG + GLP-1(A8S)-PEG (V253), or GLP-1(A8S)-FGF21(V76)-PEG (V272). Treated animals in all three groups showed lower basal fed glucose levels during the study. The dual function protein treatment resulted in significantly lower glucose AUC (64% reduced from vehicle) compared to combination treatment (55% reduced from vehicle). After fasting, the mice were given a glucose bolus, and glucose excursion was measured. All three treatments lowered blood glucose levels compared to vehicle-treated animals, with the dual function protein treatment resulting in a 76% reduction compared to 66% for the combination treatment. All treated groups had increased pancreatic insulin content compared to vehicle, with the dual function protein treatment resulting in a 76% greater insulin content than combination treatment.

[000295] NOD mice were treated twice per week for three weeks with GLP-1(A8S)-FGF21(154C)-PEG (V235) or GLP-1(A8S)-PEG (V253) + FGF21(154C)-PEG (V238). After two weeks, the mice were fasted and given a glucose bolus for an OGTT. The dual function protein-treated mice showed significantly lower fasted glucose (~60% of vehicle) and glucose excursion during challenge (~45%) compared to vehicle while combination-treated animals had fasted glucose levels that were not significantly different from vehicle (~70%) and similar glucose excursion during challenge compared to vehicle-treated animals. At the end of the study, the dual function protein-treated animals had significantly lower basal glucagon levels (~50%) than vehicle, while combination-treated animals had levels similar to vehicle.

Example 8. Predictive immunogenicity results – MHC-associated Peptide Proteomics (MAPPs) and T-cell assay results.

[000296] The formation of anti-drug antibodies (ADA) to mAbs and other therapeutic proteins could potentially lead to severe immunotoxicological reactions, such as IgE-mediated anaphylactic reactions (Chung et al. (2008) N Eng J Med, 358, 1109-17) or immune complex disease, e.g., vasculitis, glomerulonephritis (Descotes and Gouraud (2008) Expert Opin Drug Metab Toxicol 4, 1537-49) as well as to a loss of clinical exposure and efficacy. Some patients treated with the therapeutic proteins PEGylated megakaryocyte growth and development factor (PEG-MGDF) and erythropoietin (EPO; Eprex) developed neutralizing ADA that were cross-reactive to their respective endogenous counterparts, leading to severe thrombocytopenia with PEG-MGDF and pure red-cell aplasia with Eprex (Li et al. (2007) Blood 98, 3241-8; Casadevall et al. (2002) N Engl J Med 346, 469-75). Hence, it is important to assess the immunogenicity risk prior to human testing. Evaluating the immunogenicity risk and putting in place a solid immunogenicity risk mitigation plan for the clinical development phase will be especially important for therapeutic proteins containing modified, non-human sequences such as the FGF21 variants that are the subject of this invention.

[000297] Formation of ADA can be induced in at least two different ways. T cell-dependent and -independent pathways have been described for B cell activation. A strong, high affinity IgG response is T cell-dependent and requires involvement of CD4+ T helper cells (TH cells). The immune response is analogous to a response against foreign antigens: naive TH cells are specifically activated by professional antigen presenting cells (APCs), such as dendritic cells (DCs), and in turn induce activation of drug-specific B cells.

[000298] The MHC-associated Peptide Proteomics (MAPPs) assay involves *in vitro* identification of HLA class II associated peptides, which are processed by

professional antigen presenting cells (APCs) such as dendritic cells. Antigen uptake, processing and presentation processes are taken into account. In this approach, immature human monocyte-derived DCs of different healthy blood donors, are incubated with different biotherapeutic drug candidates in the presence of an activation stimulus. The naturally processed HLA class II-associated peptides, which are derived from the biotherapeutic protein, are identified by liquid chromatography-mass spectrometry (Kropshofer and Spindeldreher (2005) in Antigen Presenting Cells: From Mechanisms to Drug Development, eds. Kropshofer and Vogt, Wiley-VCH, Weinheim, 159-98).

[000299] T cell assays provide a format in which the potential risk for immunogenicity of whole protein therapeutics can be assessed. T cell assays evaluate the capacity of a therapeutic protein to induce a $CD4^+$ T cell response. Using a cohort of healthy blood donors covering a broad panel of HLA class II haplotypes, purified therapeutic proteins are tested for T cell proliferation and / or cytokine secretion *in vitro*. This technology been used successfully to compare protein variants for the potential to induce an immune response *in vivo* (Jones et al. (2004) J Interferon Cytokine Res. 24, 560-72.; Jones et al. (2005) J Thromb Haemost. 3, 991-1000) and this assessment of currently approved monoclonal antibodies does show some degree of correlation between the activation of T cells observed *in vitro* and immunogenicity in the clinic (Perry et al. (2008) Drugs R D 9, 385-96). In the context of this invention, peripheral blood mononuclear cells (PBMCs) from a cohort of 50 healthy donors representing the world population (based on HLA allotypes) are incubated with the FGF21 variants and T cell responses are measured using proliferation assays ($[^3H]$-Thymidine uptake) and IL-2 cytokine secretion (ELISpot). Subsequently, analysis of the frequency and magnitude of the $CD4^+$ T cell responses are carried out to assess the risk of clinical immunogenicity.

[000300] The combined use of the T cell and MAPPs assays provides an effective process for evaluation of immunogenicity risk in the clinic using human cells. Modifications to a therapeutic protein candidate that results in a reduced number of peptides presented by dendritic cells and a reduced number of responding donors in the T-cell assay will be advantageous as these proteins bear a lower risk to develop immunogenicity in the clinic. An example of a protein modification that has been described to reduce immunogenicity is PEGylation. However, reduced immunogenicity does not always occur with PEGylation (Li et al. (2007) Blood 98, 3241-8).

[000301] In the MAPPs assay, all tested PEGylated GLP-1(A8S) and Exendin-4 FGF21-fusion molecules, including V272 (PEGylated) and V277 (PEGylated) and the non-PEGylated Fc-FGF21 variants containing the Q55C, G148C mutations (V101, V103

and V188), show a low number of clusters, and peptide length variants that are comparable to V76 (PEGylated) and lower than wild-type FGF21 or V76 (non-PEGylated). In the T cell assay, the frequency of T-cell responses was <10% of the study cohort and the magnitude of the responses were low for V272 (PEGylated) and V277 (PEGylated) and the non-PEGylated Fc-FGF21 variants containing the Q55C, G148C mutations (V101, V103 and V188) (Table 10). The absence of a PEG moiety or the absence of the additional disulfide bond at Q55C, G148C may contribute to the increased MAPPs and T-cell assay responses seen with wild-type FGF21 and V76 (non-PEGylated) (Tables 9 and 10). Based on the results from the MAPPs and T cell assays, the risk to develop immunogenicity in the clinic can be considered to be low for V76 (PEGylated), V272 (PEGylated), V277 (PEGylated) and the non-PEGylated Fc-FGF21 variants containing the Q55C, G148C mutations (V101, V103, V188). It is known that the addition of disulfide bonds to proteins can enhance their proteolytic stability and this feature of the Fc-FGF21 variants containing the Q55C, G148C mutations may contribute to the reduced number of peptides displayed by antigen-loaded mature dendritic cells in the MAPPs assay.

[000302] In addition, the increased MAPPs response in the assays to non-PEGylated V76 may be explained by the free cysteine leading to dimerization of the molecule and thereby impacting antigen processing and presentation. This is in line with the observation that wild type FGF21 with the introduced free cysteine showed the same tendency to dimerize and resulted in increased presentation of peptides. It is likely that additional PEGylated dual function proteins, and the GLP-1(A8S)-Fc-FGF21 variants (V202, V203, V212, V213, V214, V215, V216, V218 etc.) and the Exendin-4-Fc-FGF21 variants (V196, V197, V198, V199, V206, V207, V208, V209, V210, V211 etc.) containing the Q55C, G148C mutations, will also exhibit MAPPs and T cell assay responses that would be consistent with a low risk of developing immunogenicity in the clinic.

[000303]

Table 9. T-cell assay summary with wild-type FGF21, V76 (non-PEGylated) and V76 (PEGylated) Experiment

Outcome T cell assay	KLH (positive control)	FGF21-WT (non-PEGylated)	FGF21-WT-(PEGylated)	V76 (non-PEGylated)	V76 (PEGylated)
Proliferation %	90	6	4	16	4
ELISpot %	82	10	8	10	4
Proliferation and ELISpot %	78	6	4	10	2

Table 10. T-cell assay summary with V76 (PEGylated), V272 (PEGylated), V277 (PEGylated), V101, V103 and V188 Experiment

Outcome T cell assay	KLH (positive control)	V76 (PEGylated)	V272 (PEGylated)	V277 (PEGylated)	V101	V103	V188
Proliferation %	92	4	6	4	4	6	4
ELISpot %	91	6	6	4	4	6	9
Proliferation and ELISpot %	85	4	6	4	4	6	4

CLAIMS

What is claimed is:

1. A dual function fusion protein comprising a GLP-1 receptor agonist region and an FGF21 receptor agonist region.

2. The dual function fusion protein of claim 1 comprising a GLP-1 receptor agonist peptide fused to a FGF21 variant, a linker, and a PEG group attached in such a way so as enhance the biological function of said GLP-1 receptor agonist or FGF21 variant.

3. The dual function fusion protein of claim 2, wherein the FGF21 variant is variant 76.

4. The dual function fusion protein of claim 3, wherein the linker is an Fc.

5. The dual function fusion protein of claim 3, wherein the linker is an Fc variant.

6. The dual function fusion protein of claim 1, wherein said dual function fusion protein of further comprises variant 208, variant 209, variant 211, variant 214, variant 272, variant 277, or variant 311.

7. A method of treating a metabolic disorder by administering to a subject in need a dual function fusion protein comprising a GLP-1 receptor agonist and an FGF21 receptor agonist.

8. The method of claim 7, wherein said dual function fusion protein improves metabolic parameters in subjects over the administration of individual GLP-1 receptor agonists and FGF21 receptor agonists in combination.

9. The method of claim 7, wherein said dual function fusion protein further comprises variant 208, variant 209, variant 211, variant 214, variant 272, variant 277, or variant 311.

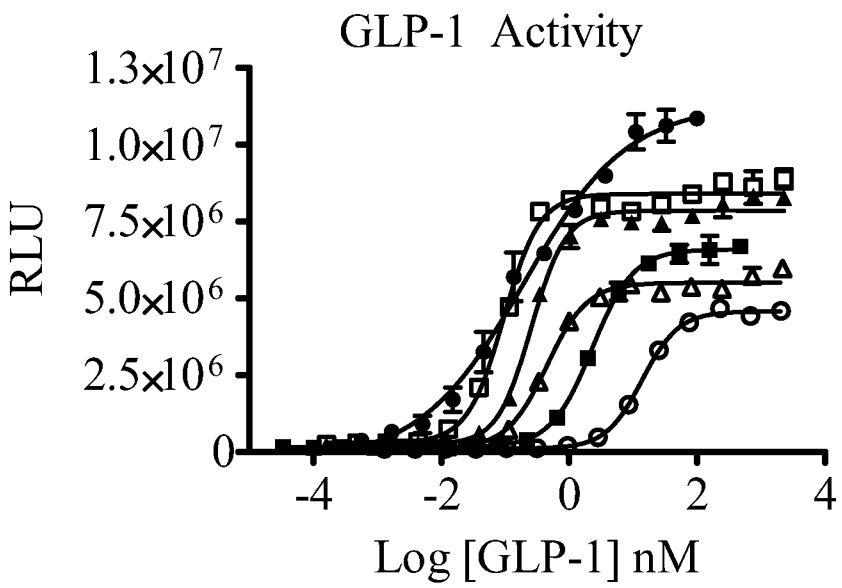

GLP-1 Activity
1.3x10^7
1.0x10^7
7.5x10^6
5.0x10^6
2.5x10^6
0
RLU
-4
-2
0
2
4
Log [GLP-1] nM

FIG 1A

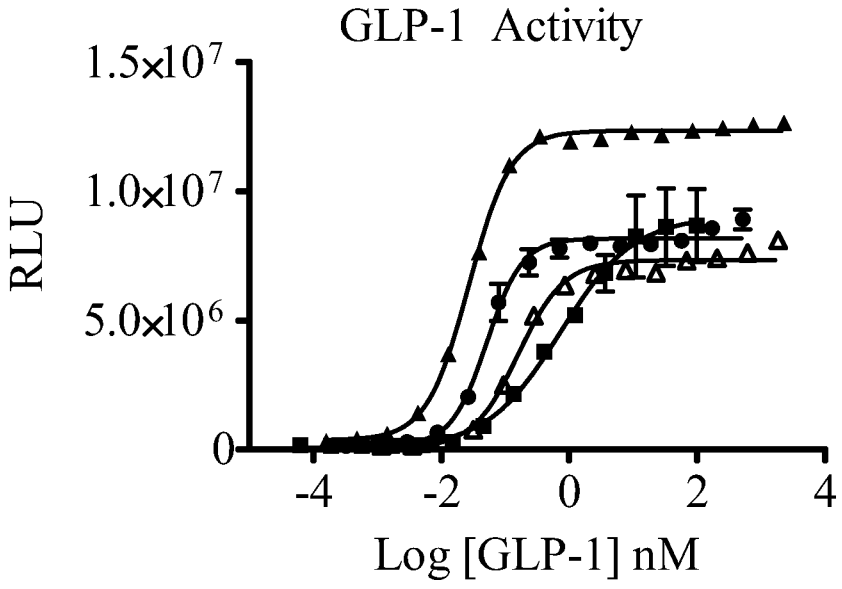

GLP-1 Activity
1.5x10^7
1.0x10^7
5.0x10^6
0
RLU
-4
-2
0
2
4
Log [GLP-1] nM

FIG 1B

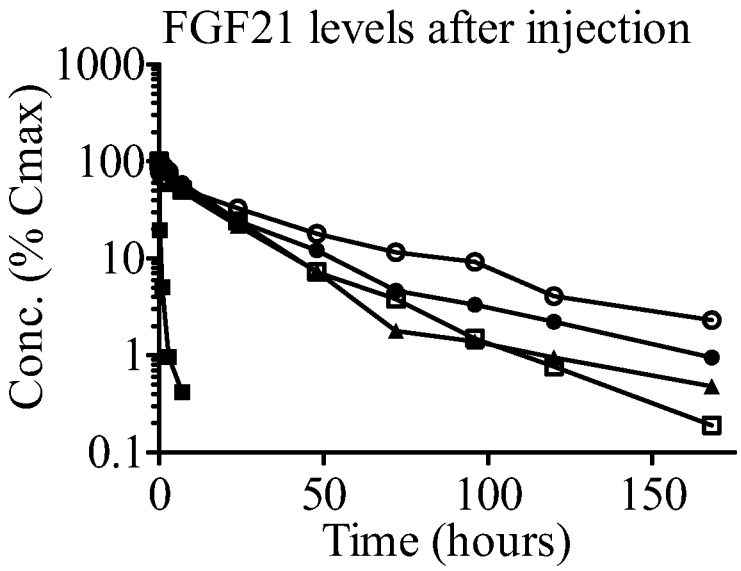

FGF21 levels after injection
Conc. (% Cmax)
1000
100
10
1
0.1
0
50
100
150
Time (hours)

FIG 2

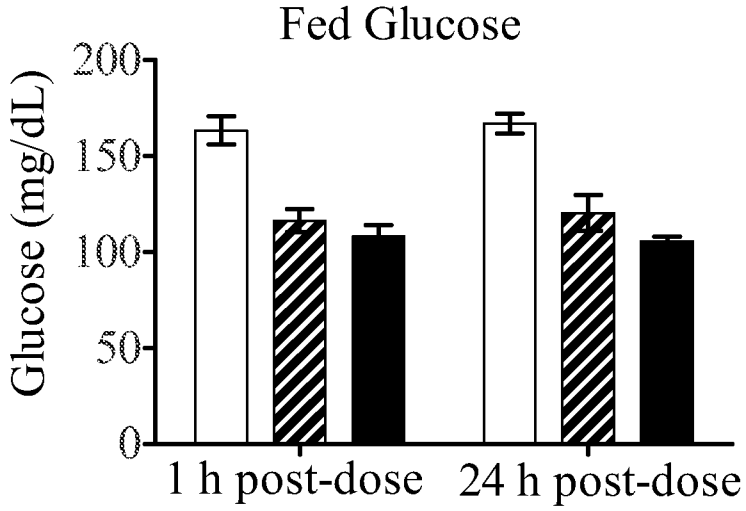

Fed Glucose
Glucose (mg/dL)
200
150
100
50
0
1 h post-dose
24 h post-dose

FIG 3A

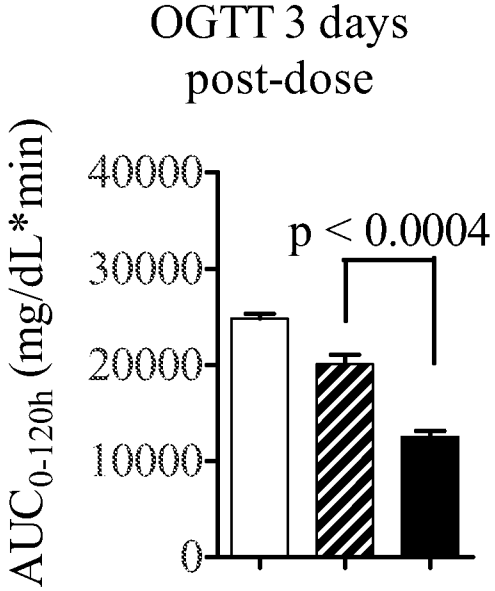

OGTT 3 days
post-dose
AUC0-120h (mg/dL*min)
40000
30000
20000
10000
0
p < 0.0004

FIG 3B

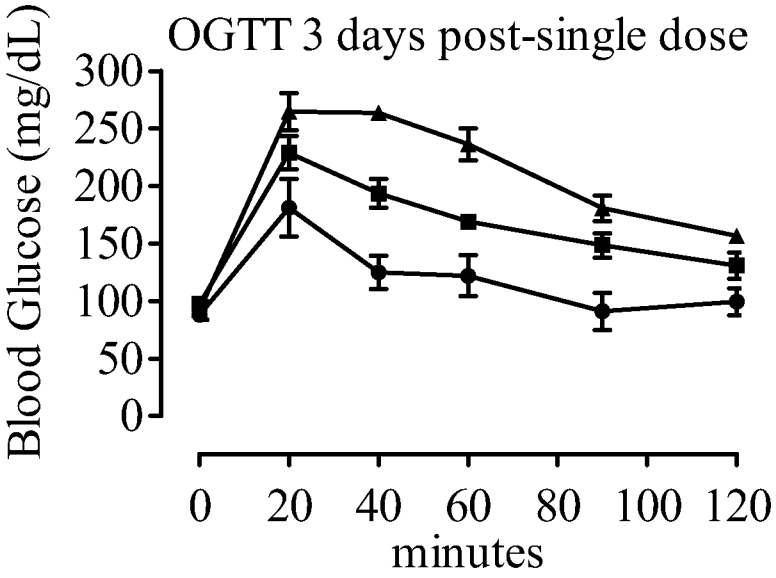

OGTT 3 days post-single dose
Blood Glucose (mg/dL)
300
250
200
150
100
50
0
0
20
40
60
80
100
120
minutes

FIG 3C

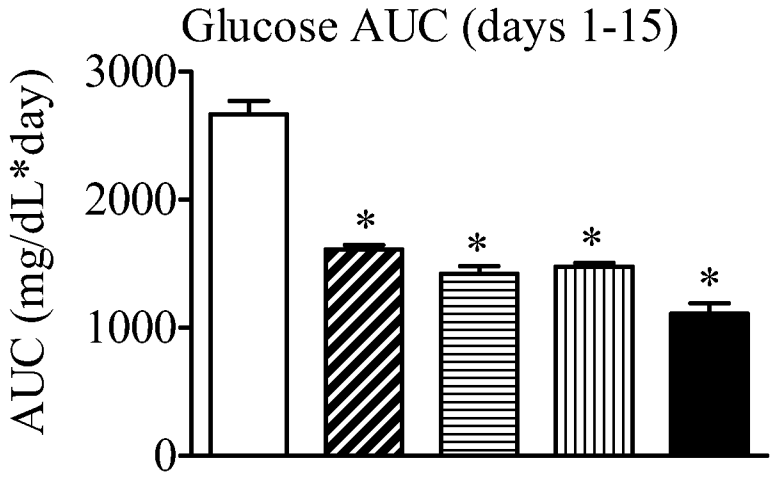


FIG 4

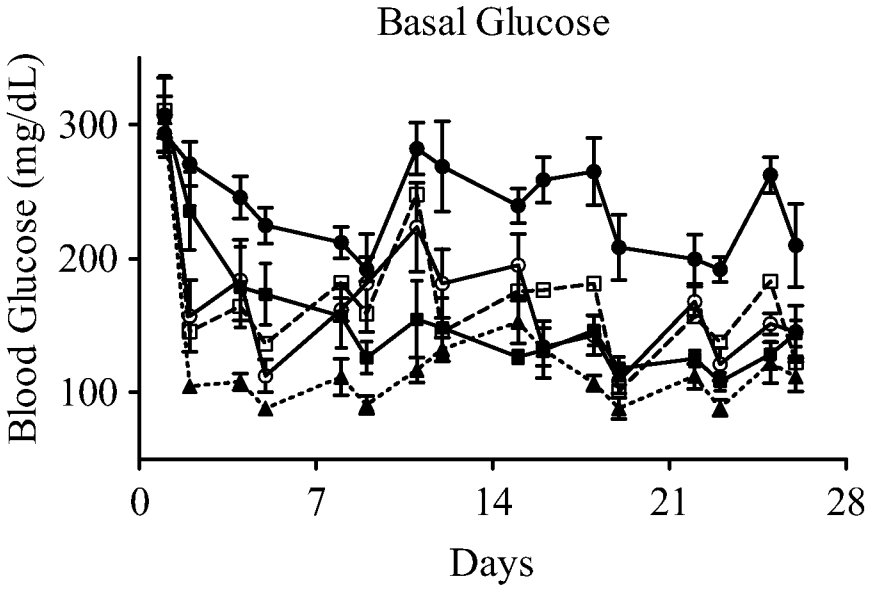

Basal Glucose
Blood Glucose (mg/dL)
300
200
100
0
7
14
21
28
Days

FIG 5A

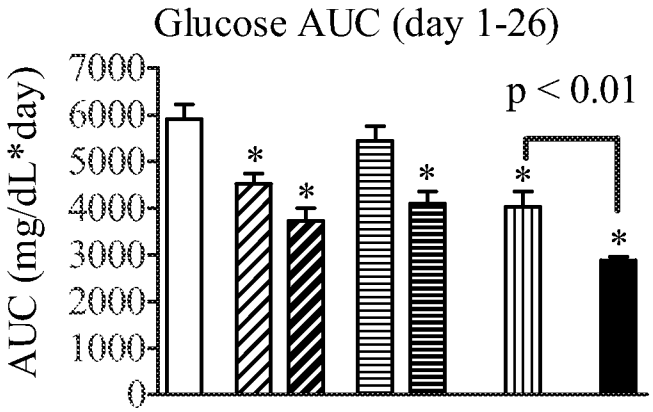

Glucose AUC (day 1-26)
AUC (mg/dL*day)
7000
6000
5000
4000
3000
2000
1000
0
p < 0.01
*

FIG 5B

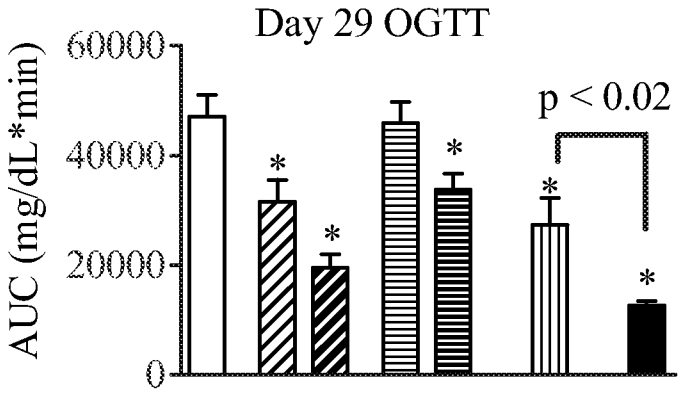

Day 29 OGTT
AUC (mg/dL*min)
60000
40000
20000
0
p < 0.02
*
*
*
*
*

FIG 5C

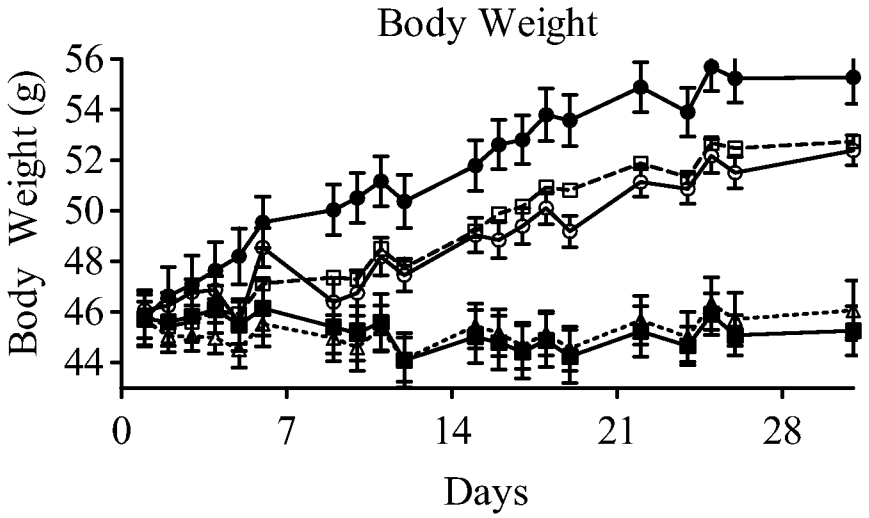

Body Weight
Body Weight (g)
56
54
52
50
48
46
44
0
7
14
21
28
Days

FIG 5D

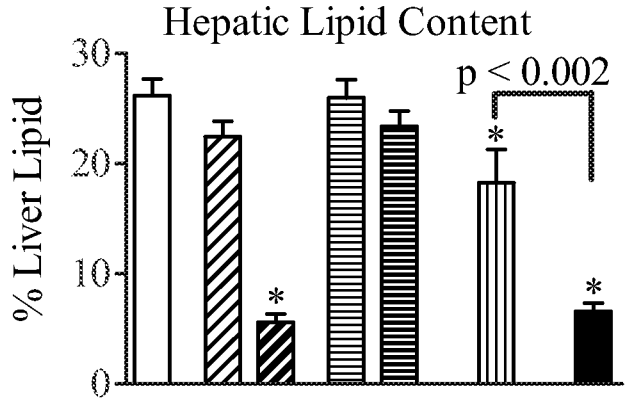

Hepatic Lipid Content
% Liver Lipid
30
20
10
0
p < 0.002
*
*
*

FIG 5E

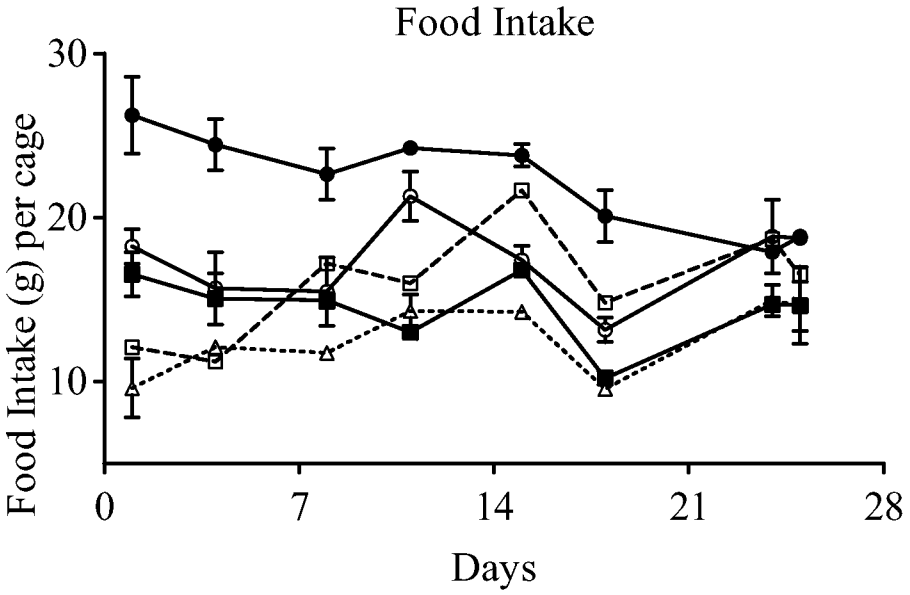

Food Intake
Food Intake (g) per cage
30
20
10
0
7
14
21
28
Days

FIG 5F

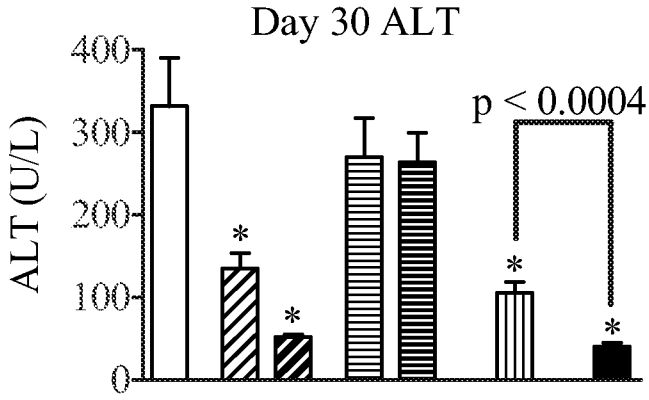


FIG 5G

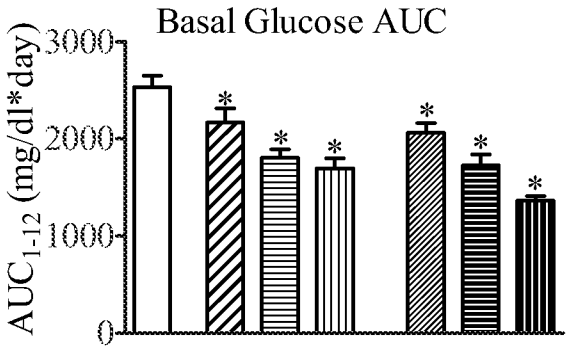


FIG 6A

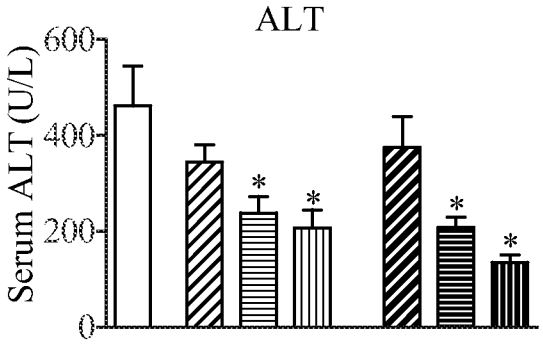


FIG 6B

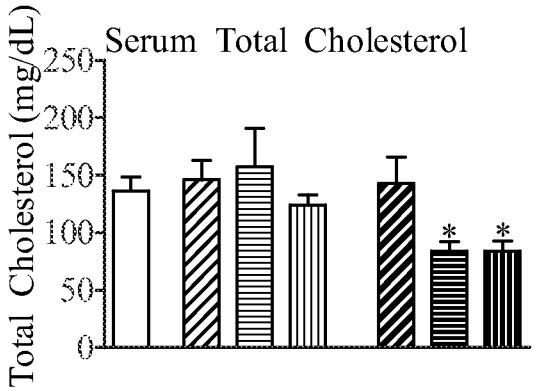


FIG 6C

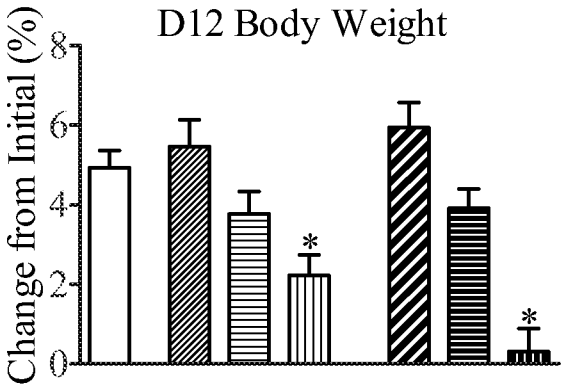


FIG 6D

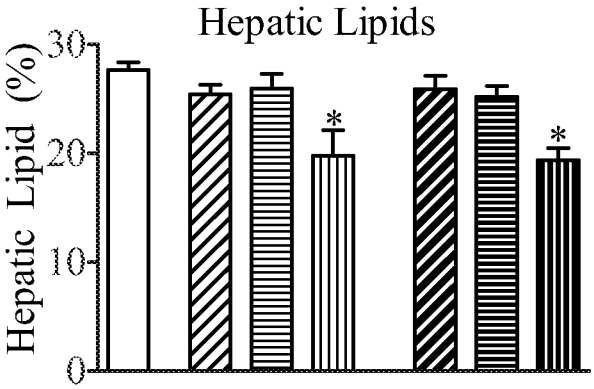


FIG 6E

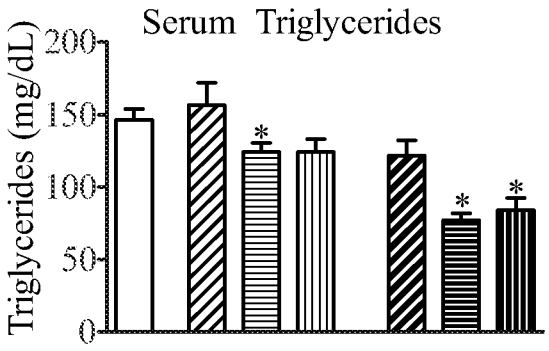


FIG 6F

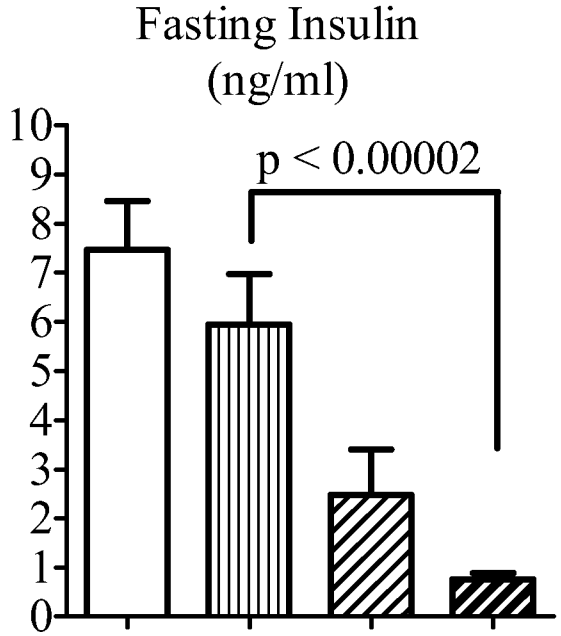

Fasting Insulin
(ng/ml)
10
9
8
7
6
5
4
3
2
1
0
p < 0.00002

FIG 7A

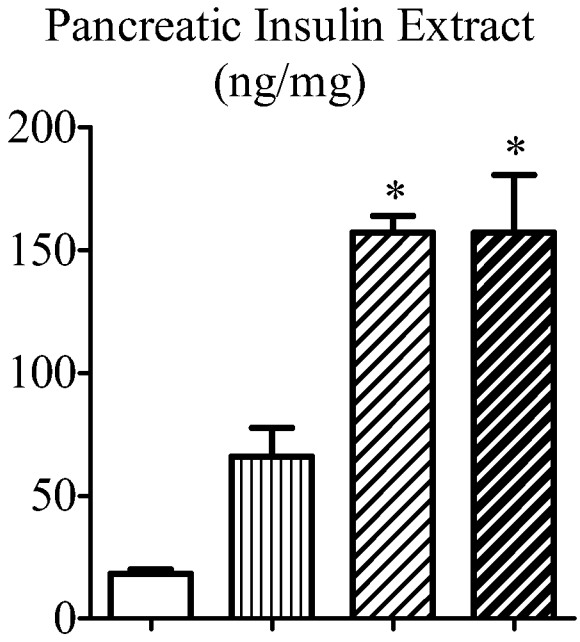

Pancreatic Insulin Extract
(ng/mg)
200
150
100
50
0
*
*

FIG 7B

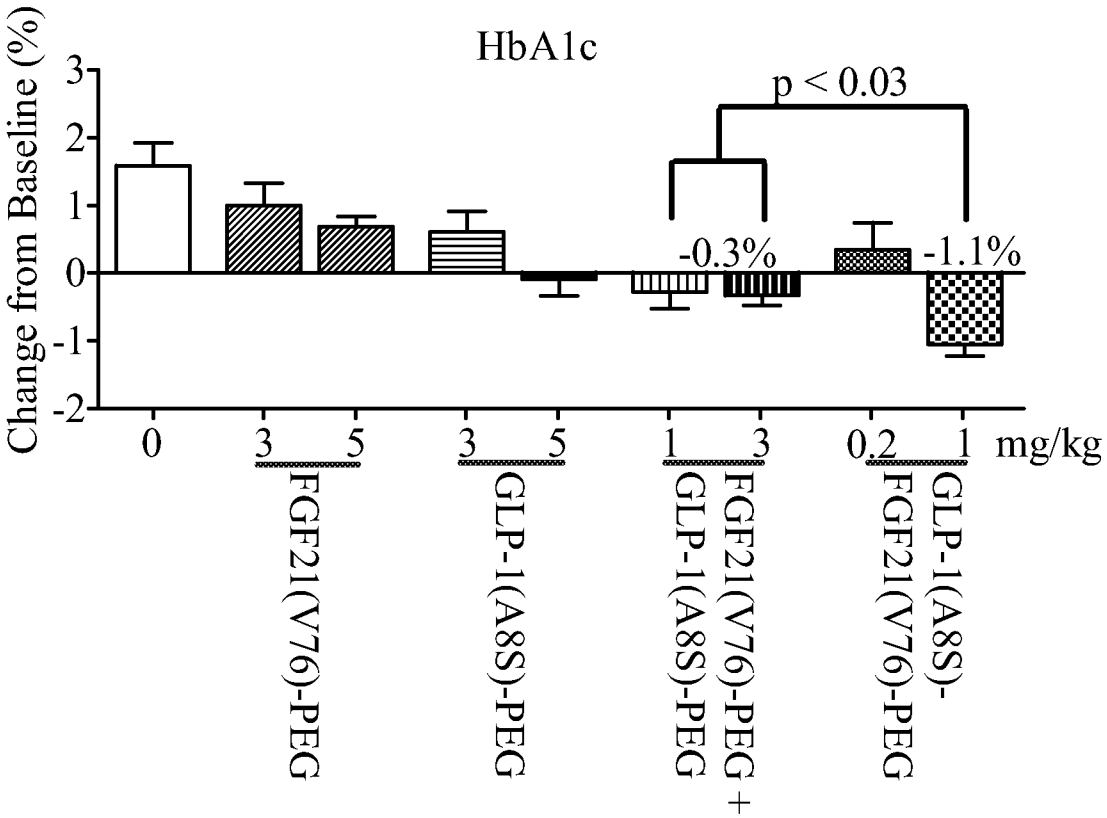

HbA1c
Change from Baseline (%)
3
2
1
0
-1
-2
p < 0.03
-0.3%
-1.1%
0
3
5
3
5
1
3
0.2
1
mg/kg
FGF21(V76)-PEG
GLP-1(A8S)-PEG
GLP-1(A8S)-PEG
FGF21(V76)-PEG +
FGF21(V76)-PEG
GLP-1(A8S)-

FIG 8A

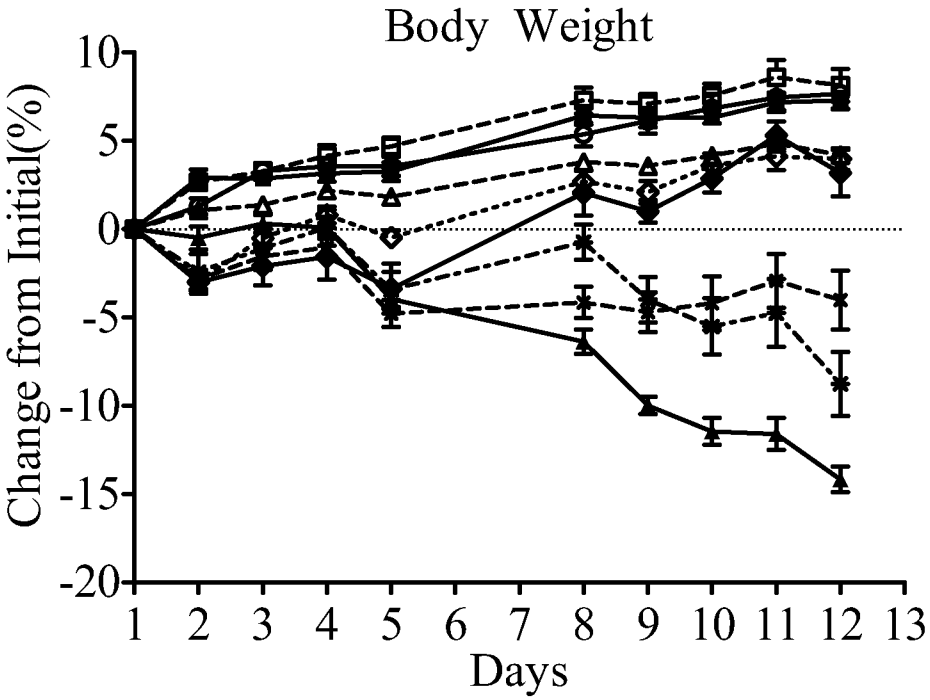

Body Weight
Change from Initial(%)
10
5
0
-5
-10
-15
-20
1 2 3 4 5 6 7 8 9 10 11 12 13
Days

FIG 8B

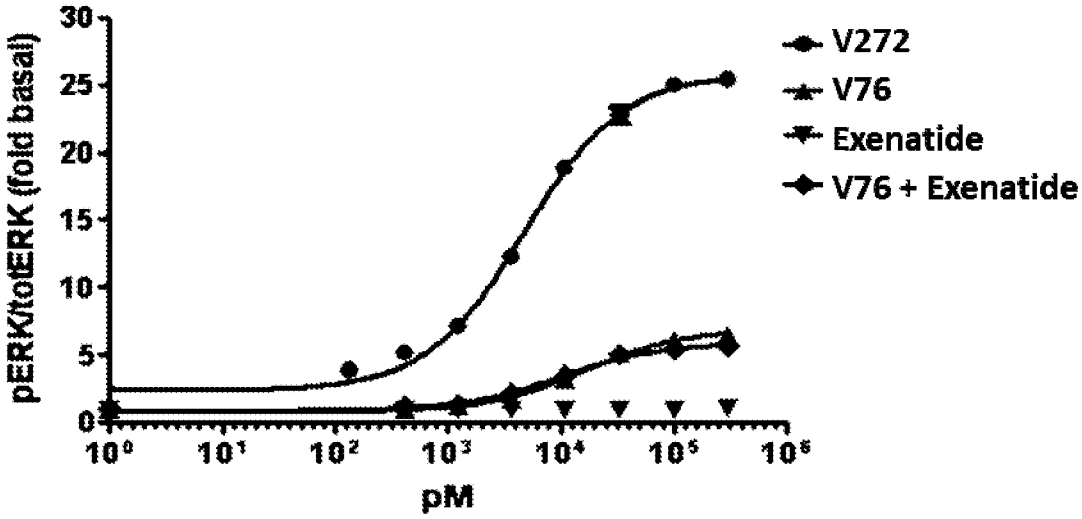

pERK/totERK (fold basal)
30
25
20
15
10
5
0
10⁰
10¹
10²
10³
10⁴
10⁵
10⁶
pM
V272
V76
Exenatide
V76 + Exenatide

FIG 9

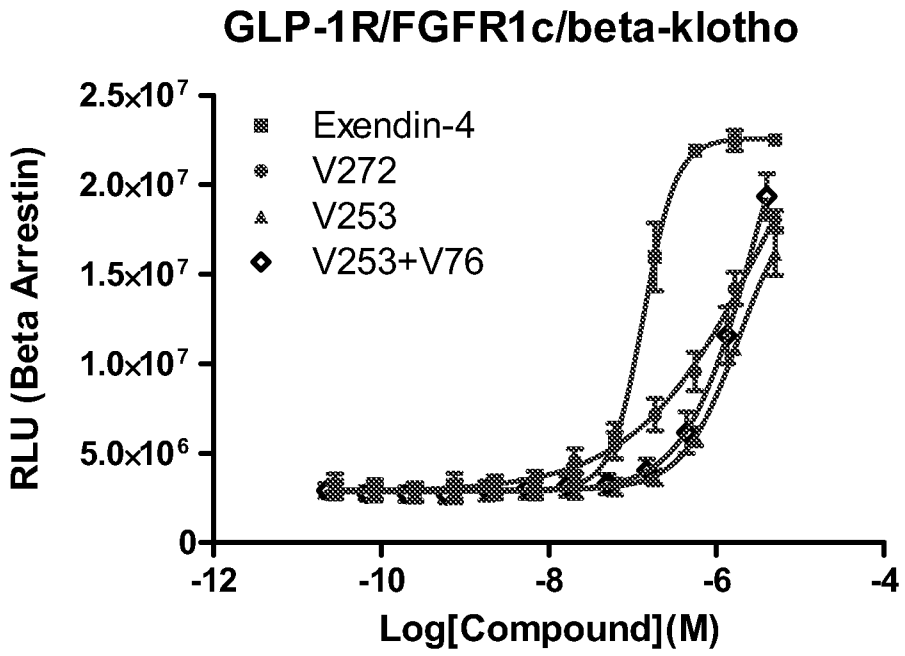


FIG 10A

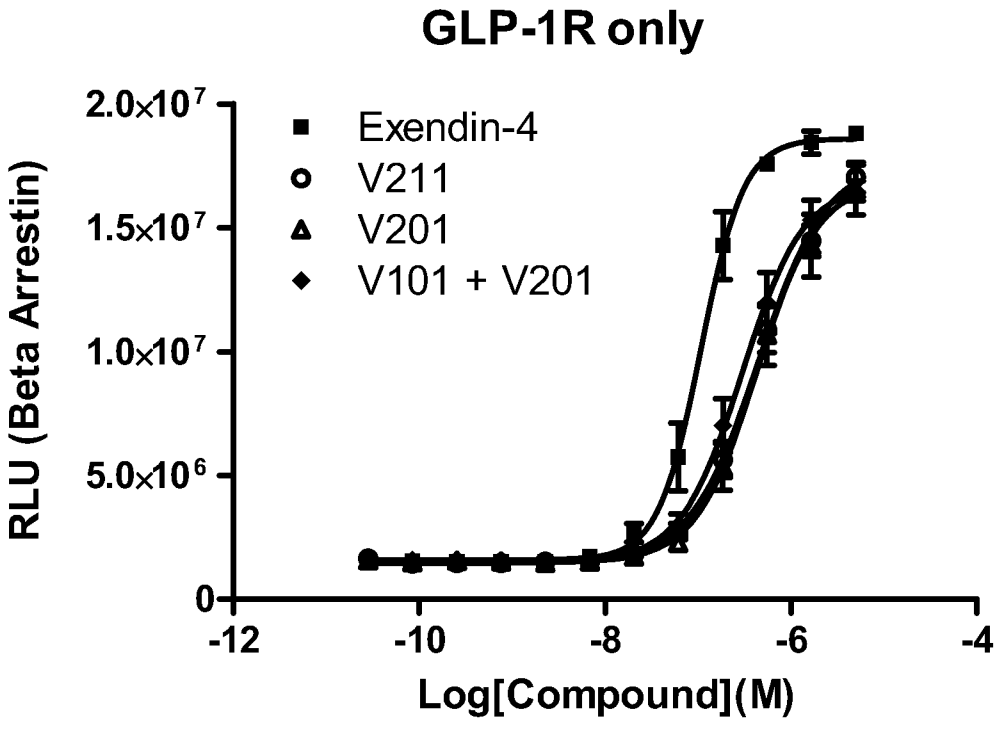


FIG 10B

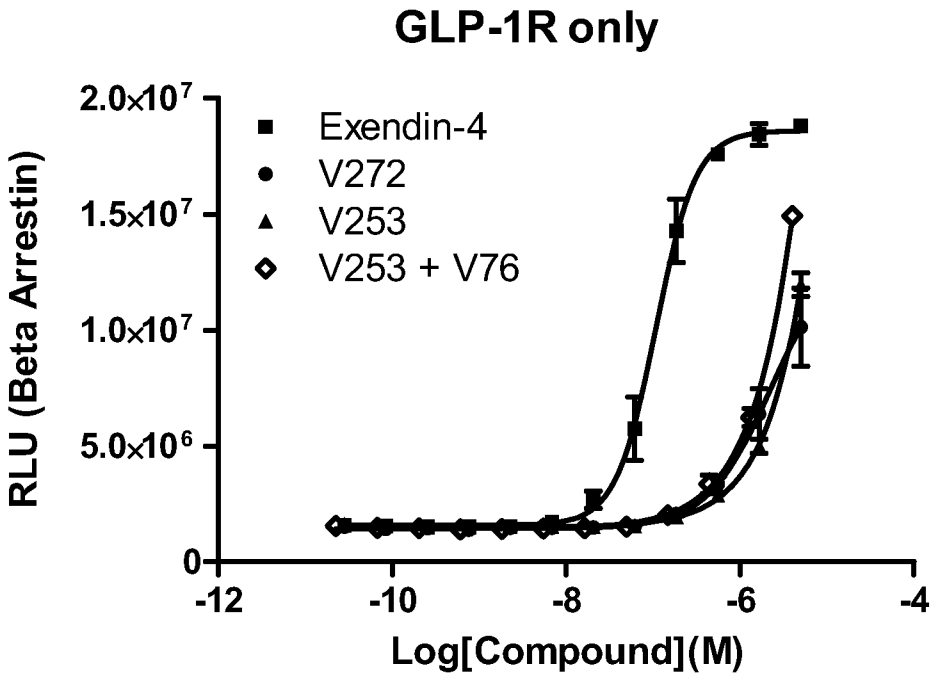


FIG 10C

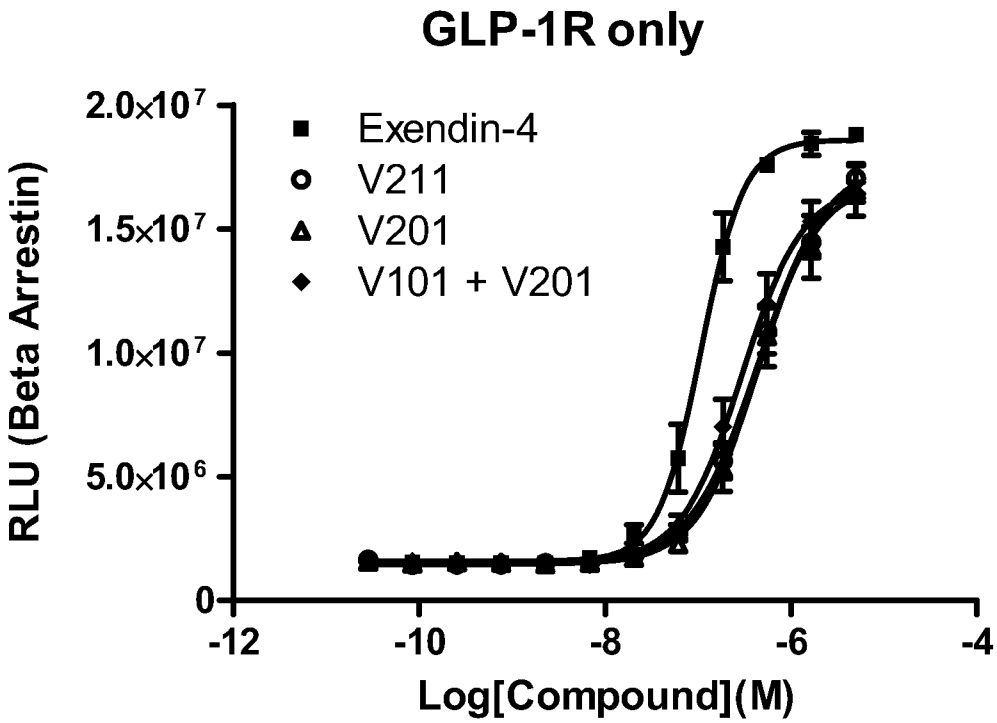


FIG 10D

PROTEÍNAS DE FUNCIÓN DUAL PARA TRATAR TRASTORNOS METABÓLICOS

CAMPO DEL INVENTO

[0001] La presente invención se refiere a nuevas proteínas que comprenden el factor de crecimiento de fibroblastos 21 (FGF21) y otros reguladores metabólicos conocidos por mejorar los perfiles metabólicos en sujetos en quienes se administren.

FUDAMENTO DEL INVENTO

[0002] La familia de factores de crecimiento de fibroblastos (FGF) se caracteriza por 22 ligandos homólogos genéticamente distintos, que se agrupan en siete subfamilias. El FGF-21 está más estrechamente relacionado con, y forma una subfamilia con el FGF-19 y el FGF-23. Esta subfamilia de FGF regula diversos procesos fisiológicos no comunes a los FGFs clásicos, a saber, la energía y la homeostasis de los ácidos biliares, el metabolismo de la glucosa y de los lípidos y homeostasis del fosfato así como la de la vitamina D. Además, a diferencia de los FGFs clásicos, esta subfamilia actúa de manera endocrina. (Moore, D.D. (2007) Science 316, 1436-8) (Beenken et al. (2009) Nature Reviews Drug Discovery 8, 235).

[0003] FGF21 es un polipéptido de 209 aminoácidos que contiene una secuencia líder de 28 aminoácidos (SEQ ID NO: 132). El FGF21 humano tiene una identidad aminoácida del 79% con el FGF21 de ratón y una identidad aminoácida del 80% con el FGF21 de rata. El factor de crecimiento de fibroblastos 21 (FG21) se ha descrito como un tratamiento para la enfermedad isquémica vascular, cicatrización de heridas y enfermedades asociadas con la pérdida de la función celular pulmonar, bronquial o alveolar. (Nishimura et al. (2000) Biochimica et Biophysica Acta, 1492:203-206; publicación de patente WO01/36640; y publicación de patente WO01/18172). A pesar de que el FGF-21 activa los receptores del FGF y moléculas de señalización corriente abajo, incluyendo FRS2a y ERK, no se ha detectado la interacción directa de FGFRs y el FGF-21. Estudios recientes han identificado β-klotho, que es altamente expresado en el hígado, adipocitos y en el páncreas, como factor determinante de la respuesta celular al FGF-21 y un cofactor que media señalización de FGF-21 a través FGFRs (Kurosu, H. (2007) J Biol Chem 282, 26687-95). FGF21 es un agonista potente de los complejos de señalización FGFR1(IIIc), FGFR2(IIIc) y FGFR3(IIIc) β-klotho .

[0004] Se ha mostrado que FGF-21 induce la absorción de glucosa independiente de insulina. También se ha mostrado que el FGF-21 mejora la hiperglucemia en una amplia gama de modelos de roedores diabéticos. Además, los ratones transgénicos que sobre-expresan FGF-21 resultaron ser resistentes a anormalidades metabólicas inducidas por la dieta, incluyendo la disminución del peso corporal y la masa grasa, y aumentos en la sensibilidad a la insulina (Badman, MK . (2007) Cell Metab 5, 426-37). La administración del FGF-21 en primates no humanos diabéticos causó una disminución de los niveles de glucosa plasmática, triglicéridos, insulina y glucagón en ayunas, y dió lugar a mejoras significativas en los perfiles lipoproteicos incluyendo un aumento de cerca del 80% en colesterol HDL (Kharitonenkov, A. (2007) Endocrinology 148, 774 a 81). Estudios recientes que investigan los mecanismos moleculares de acción de FGF21 han identificado a FGF21 como una hormona endocrina importante que ayuda a controlar la adaptación al estado en ayudas: (Cadigan et al. (2009) Endocrinology 150, 4931)(Inagaki et al. (2007) Cell Metabolism 5, 415). Esto proporciona un vínculo que faltaba, corriente abajo de PPARα, mediante el cual el hígado se comunica con el resto del cuerpo en la regulación de la biología de la homeostasis energética. (Cadigan et al. (2008) Cell Metabolism 8, 169)(Lundasen et al. (2007) Biochemical and Biophysical Research Communications 360, 437).

[0005] FGF21 regula la homeostasis de adipocitos a través de la activación de una ruta AMPK/SIRT1/PGC1α para inhibir la expresión de PPARγ y aumentar la función mitocondrial. (Chau et al. (2010) PNAS 107, 12553) FGF21 también aumenta la absorción de glucosa por musculo esquelético en miotúbulos humanos cultivados y tejido de ratón aislado (Mashili et al. (2011) Diabetes Metab Res Rev 27, 286-97). El tratamiento con FGF21 de células de islotes de roedores conduce a unión y supervivencia mejoradas a través de la activación de las rutas ERK1/2 y Akt. (Wente et al. (2006) Diabetes 55, 2470) el tratamiento con FGF21 también resulta en expresión génica alterada para lipogénesis y enzimas de oxidación de ácidos grasos en hígados de roedores, probablemente a través de señalización de HNF4α y Foxa2. Sin embargo, estudios recientes (Wei, [0007] (2012) PNAS 109, 3143-48) indican que el tratamiento de ratones obesos inducidos por dieta con FGF21 induce la pérdida de hueso, debido a una disminución de la inactivación de PPAR□ (a través de la reducción de sumoilación); un cambio de la diferenciación de células madre mesenquimales de osteoblastos a adipocitos se ve en la presencia de aumento de la actividad de PPAR□ en el hueso después del tratamiento con FGF21.

[0006] Una dificultad asociada con el uso del FGF-21 directamente como un bioterapéutico es que su vida media es muy corta. (Kharitonenkov, A. et al. (2005) Journal of Clinical Investigation 115:1627-1635). En ratones, la vida media del FGF21 humano es de 0,5 a 1 hora, y en monos cyno, la vida media es de 2 a 3 horas. El FGF21 puede utilizarse como una formulación farmacéutica estéril multi-uso. Sin embargo, se ha determinado que los preservativos, es decir, m-cresol, tienen un efecto adverso en la estabilidad en estas condiciones.

[0007] Otro regulador metabólico potente ya representado en la clínica es el Péptido similar al Glucagón 1 (GLP-1). (Knudsen et al. (2004) Journal of Medicinal Chemistry 47, 4128) GLP-1 es una incretina de 36 aminoácidos secretada por las células L del intestino de mamíferos, que actúa sobre las células alfa y beta para estimular la secreción de insulina e inhibir la liberación de glucagón de una manera dependiente de la glucosa. (Hare et al. (2010) Diabetes 59, 1765) (Meier (2005) Diabetes-Metabolism Research and Reviews 21, 91). El GLP-1 se une con y activa el receptor GLP-1 (GLP-1R), una proteína helicoidal de siete-transmembrana de la familia clase II de receptores acoplados con la proteína G (GPCRs) (Mayo (2003) Pharmacological Reviews 55:167.) Como un agonista del receptor de GLP-1, GLP-1 tiene un papel importante en la disminución de los niveles pos-prandiales de glucosa en sangre mediante la estimulación de secreción de insulina desde el páncreas con el fin de aumentar la absorción de glucosa en los tejidos periféricos y la inhibición de la secreción de glucagón, lo que resulta en reducción de la liberación de glucosa hepática.

[0008] Un segundo agonista del receptor de GLP-1 clínicamente importante es exendina-4. La exendina-4 es un polipéptido de 39 residuos producido en las glándulas salivales del lagarto monstruo de Gila (Goke et al. (1993) Diabetes 46:433-439) (Fehmann et al. (1995) Endocrine Rev. 16:390-410). Aunque es el producto de un gen excepcionalmente no mamífero y parece que se expresa solamente en la glándula salival, la exendina-4 comparte una homología de la secuencia aminoácida del 52% con GLP-1 y en mamíferos interactúa con el receptor de GLP-1 (Goke) (Thorens et al. (1993) Diabetes 42:1678-1682). In vitro, la exendina-4 ha sido mostrada por promover la secreción de insulina por las células productoras de insulina y, administrada en cantidades equimolares, es más potente que el GLP-1 para provocar la secreción de insulina a partir de células productoras de insulina. Además, la exendina-4 potentemente estimula la liberación de insulina para reducir los niveles de glucosa en plasma tanto en roedores como en humanos y de acción más prolongada que el GLP-1, sin embargo, debido a que no se

produce naturalmente en los mamíferos, la exendina-4 tiene ciertas propiedades antigénicas potenciales en los mamíferos que no tiene el GLP-1.

[0009] La capacidad de GLP-1 y análogos de la exendina-4 (ej., liraglutida y Byetta) para mejorar el control de la glucosa en humanos se establece en la clínica. (Idris (2010) Diabetes Obesity & Metabolism 12, 89)(Monami et al (2009) European Journal of Endocrinology 160, 909). También se ha reportado que GLP-1 aumenta la masa de células beta a través de la proliferación inducida y la inhibición de la apoptosis. (Egan, A (2003) Diabetes-Metabolism Research and Reviews 19, 115)(Farilla, L. et al. (2003) Endocrinology 144, 5149) (Xu, G. (1999). Diabetes 48, 2270). También actúa como una hormona intestinal para inhibir la secreción de ácido y el vaciado gástrico en el estómago mientras que proporciona una señal de saciedad que disminuye el apetito. (Vilsboll et al. (2009) Best Practice & Research Clinical Endocrinology & Metabolism 23, 453). Estos efectos probablemente explican la pérdida de peso beneficiosa observada con la administración de análogos de GLP-1 en pacientes con diabetes tipo 2. También se ha mostrado que GLP-1 es cardioprotector en corazones de roedores pos-isquémicos. (Ossum et al. (2009) Pharmacological Research 60, 411) (Sonne, DP (2008) Regulatory Peptides 146, 243; Nikolaidis, LA (2004) Circulation 109, 962).

[00010] Además, el GLP-1 puede reducir la diferenciación de células madre mesenquimales humanas (hMSC) a adipocitos mediante la reducción de la expresión de PPARγ, y GLP-1 promueve la proliferación celular y la citoprotección de hMSC (Sanz, (2010) Am J Physiol Endocrinol Metab 298, E634-E643).

[00011] En el desarrollo de una proteína de FGF21, incluyendo una variante o análogo de la misma, para su uso como agente terapéutico en el tratamiento de la diabetes mellitus tipo 1 y tipo 2 y otros trastornos metabólicos, un aumento en la vida media y en la estabilidad sería deseable. Las proteínas del FGF21 que tienen una vida media y una estabilidad aumentada permitirían una dosificación menos frecuente en los pacientes a quienes se administró la proteína. Claramente, hay una necesidad de desarrollar una formulación proteica estable acuosa para el FGF21 proteico terapéutico.

[00012] Además, un reto significativo en el desarrollo de farmacéuticos de proteínas, tales como reguladores metabólicos FGF21, GLP-1, y Exendina-4, es hacer frente a sus inestabilidades físicas y químicas. La variedad de composición y características de las proteínas definen los comportamientos específicos tales como dobleces, estabilidad conformacional, y desdoblamientos/desnaturalizaciones. Estas características se deben abordar para estabilizar las proteínas en el desarrollo de las condiciones de formulación

farmacéutica que utilizan soluciones acuosas de proteínas (Wang, W., int. J. of Pharmaceutics, 18, (1999)). Un efecto deseado de la estabilización de proteínas terapéuticas de interés, por ejemplo, las proteínas de la presente invención, es la resistencia creciente a proteólisis y la degradación enzimática, mejorando así la estabilidad de las proteínas y reduciendo la agregación de proteínas.

RESUMEN DE LA INVENCION

[00013] La invención se refiere a la identificación de nuevas proteínas, por ejemplo proteínas de fusión, que comprenden el factor de crecimiento de fibroblastos 21 (FGF21) y otros reguladores metabólicos, por ejemplo GLP-1 y Exendina-4, y que tienen propiedades farmacéuticas mejoradas sobre los agentes constituyentes bajo condiciones de formulación farmacéutica ej., son más estables, poseen la capacidad de mejorar los parámetros metabólicos de sujetos a quienes se administran, y son menos susceptibles a proteolisis y degradación enzimática, son menos propensas a agregar y formar complejos y tienen menos probabilidad de ser inmunogénicos. Las proteínas de la invención poseen tanto actividad de agonistas de receptores de FGF21 como agonistas de receptores de GLP-1, y comprenden truncamientos y variantes de FGF21, y comprenden además uno o más reguladores del peptido-1 tipo glucagón-1 (GLP-1), exendina- 4, u otros reguladores o variantes metabólicas de los mismos.

[00014] También se divulgan métodos para el tratamiento de trastornos asociados con FGF21 GLP-1, así como otros trastornos metabólicos, endocrinos y cardiovasculares, tales como la obesidad, diabetes mellitus tipo 1 y tipo 2, pancreatitis, dislipidemia, enfermedad de hígado graso no alcohólico (NAFLD), esteatohepatitis no alcohólica (NASH), resistencia a la insulina, hiperinsulinemia, intolerancia a la glucosa, hiperglucemia, síndrome metabólico, infarto agudo de miocardio, hipertensión, enfermedades cardiovasculares, ateroesclerosis, enfermedad arterial periférica, accidente cerebro-vascular, insuficiencia cardiaca, enfermedad cardiaca coronaria, enfermedad renal, complicaciones diabéticas, neuropatía, gastroparesia, trastornos asociados con severas mutaciones inactivadoras en el receptor de la insulina, lipodistrofias incluyendo lipodistrofia asociada a VIH y otros trastornos metabólicos, y en la reducción de la mortalidad y la morbilidad de los pacientes críticos.

[00015] Las proteínas de la presente invención pueden ser utilizadas como un inyectable administrado de manera regular (ej., a diario, mas preferiblemente una vez a la semana, una vez cada dos semanas, o una vez al mes) ya sea solo o en combinación con

agentes antidiabéticos orales que mejoran el control de la glucemia, el peso corporal y el perfil lipídico de pacientes con diabetes mellitus tipo 1 y tipo 2. Las proteínas también pueden ser utilizadas para el tratamiento de la obesidad o de otras condiciones asociadas con FGF21-o GLP1.

[00016] Las proteínas de la invención, por ejemplo proteínas de fusión de variante GLP-1-FGF21 y de variante Exendina-FGF21 de la invención, superan los obstáculos significativos de inestabilidades físicas asociadas con terapéuticos de proteína, incluyendo, por ejemplo, con la administración del FGF21 tipo silvestre, puesto que son más estables, menos susceptibles a la proteólisis y la degradación enzimática, menos probables de agregarse y formar complejos y menos probables de ser inmunogénicos que el FGF21 tipo silvestre en condiciones de la formulación farmacéutica.

[00017] En un primer aspecto, la invención proporciona proteínas del Factor de Crecimiento Fibroblástico 21 (FGF21), por ejemplo proteínas de fusión, comprenden una o más de las secuencias enumeradas en la Tabla 1, y se describe adicionalmente en la presente. Las proteínas de la invención pueden además comprender proteínas de GLP-1 y/o Exendina-4, ya sea versiones tipo silvestre, truncadas o mutadas, o variantes de las mismas. Las secuencias de FGF21 enumeradas en la Tabla 1 son variantes de la secuencia de FGF21 tipo silvestre, por ejemplo, la secuencia de FGF21 tipo silvestre con número de referencia NCBI NP_061986.1, y encontrada en tales patentes emitidas como, por ejemplo, EE.UU. 6.716.626 B1, asignada a Chiron Corporation. Las secuencias de GLP-1 y exendina-4 presentadas en la Tabla 1 son variantes de las secuencias tipo salvaje de GLP-1 y exendina-4, por ejemplo, aquellas secuencias con números de referencia NCBI NP_002045 y AAB22006.1, respectivamente, y se pueden encontrar en publicaciones de patentes tales como, por ejemplo, WO98/19698 y WO87/06941A, asignadas a Eli Lilly y Co. y el General Hospital Corp., respectivamente (GLP-1) y EE.UU. 5.424.286, asignada a Amylin (exendina-4).

[00018] Otras realizaciones están relacionadas con polinucleótidos que codifican las proteínas de función dual de la invención, un vector que contiene dichos polinucleótidos y una célula huésped que lleva dicho vector.

[00019] Se proporcionan en la presente, métodos utilizados para generar las proteínas de la invención, en donde tales métodos implican la modificación de la proteína FGF21 tipo silvestre, por ejemplo, la incorporación específica de sitio de los aminoácidos en las posiciones de interés dentro de la proteína FGF21 tipo silvestre, así como también la fusión entre la porción de FGF21 de la molécula a otros reguladores metabólicos, tales

como el péptido-1 tipo glucagón-1 (GLP-1) y la exendina-4, o conjugados con polímeros modificados con GLP-1 y/o exendina-4. Dichas modificaciones y fusiones mejoran las propiedades biológicas de las proteínas de la invención en relación con las versiones tipo silvestre de las proteínas (por ejemplo, FGF21, GLP-1, y la exendina-4), así como, en algunos casos, sirven como puntos de enlace para, por ejemplo, etiquetas y agentes de extensión de vida media de proteínas, y para los propósitos de fijar dichas variantes a la superficie de un soporte sólido. Las realizaciones afines de la invención son los métodos de producir células capaces de producir dichas proteínas de la invención, y de la producción de vectores que contengan ADN que codifique dichos variantes y fusiones.

[00020] En varias realizaciones, las proteínas de la invención pueden comprender uno o más fragmentos de las secuencias de FGF21, Exendina-4 y/o GLP-1, incluyendo fragmentos tan pequeños como residuos de 8-12 aminoácidos de longitud, en donde el polipéptido es capaz de reducir la glucosa en sangre en un mamífero. En varias realizaciones, las proteínas de la invención pueden comprender una o más variantes de las secuencias de FGF21, Exendina-4 y/o GLP-1, ej., con una o más deleciones, inserciones, adiciones o sustituciones de aminoácidos con relación a las secuencias tipo silvestre de los mismos.

[00021] En algunas realizaciones, las proteínas de la invención pueden unirse covalentemente a uno o más polímeros, tales como el polietilenglicol (PEG) o ácido polisiálico, ya sea en la posición de modificaciones aminoácidas específicas a sitio hechas en relación con el FGF21 tipo silvestre, GLP-1, o Exendina, o en la posición de los aminoácidos comúnmente compartidos con las versiones tipo silvestre de estas proteínas. El grupo PEG es fijado en tal forma para aumentar, y/o no interferir con, la función biológica de las pociones constituyentes de las proteínas de fusión de la invención, por ejemplo, las variantes proteicas de GLP-1 o las variantes proteicas de FGF21. En otras realizaciones, los polipéptidos del invento pueden ser fusionados a una secuencia aminoácida heteróloga, opcionalmente mediante un enlazador, tal como GS o GGGGSGGGGSGGGGS (SEC ID NO: 8), o SGGGGSGGG (SEC ID NO:128). La secuencia aminoácida heteróloga puede ser un dominio constante de IgG o fragmento del mismo (por ejemplo, la región Fc), Albúmina de Suero Humano (ASH), o polipéptidos fijadores de albúmina. Tales proteínas de fusión descritas en este documento también pueden formar multímeros.

[00022] En algunas realizaciones, la secuencia aminoácida heteróloga (por ejemplo, ASH, Fc, etc) se fusiona con el amino-terminal de las proteínas de fusión de la

invención. En otras realizaciones, la secuencia aminoácida heteróloga de fusión (por ejemplo, ASH, Fc, etc) se fusiona con el carboxilo-terminal de las proteínas de la invención. En aún otras realizaciones más preferidas, la secuencia heteróloga de aminoácidos (por ejemplo, HSA, Fc, etc) está situada en el medio de las proteínas de función dual de la invención, es decir, entre el residuo C-terminal del péptido GLP-1 o secuencia de exendina-4 y el residuo N-terminal de la secuencia de FGF21. Dicha realización preferida, por ejemplo, deja libre un extremo-N para máxima actividad de GLP-1 (exendina-4) y un extremo-C intacto, libre para máxima actividad de FGF21.

[00023] En algunas realizaciones, el agonista del receptor de GLP-1 está unido con el extremo-N de la cadena pesada y ligera de un anticuerpo y FGF21 se fusiona simultáneamente al extremo-C de la cadena pesada y ligera del mismo anticuerpo (es decir, un cuerpo fusionado, como se describe en la presente). Dicha realización preferida deja un extremo-C libre para una máxima actividad de GLP-1 (exendina-4) y un extremo C intacto, libre para máxima actividad de FGF21. Una realización preferida utiliza la secuencia de anticuerpos descrita en la publicación PCT WO2011/076781.

[00024] En algunas realizaciones, el péptido GLP-1 o Exendina-4 está unido químicamente a FGF21. En algunas realizaciones, dichos péptidos se unen a una cadena lateral de residuos aminoácidos de FGF21. En otras realizaciones, dichos péptidos se unen al extremo-N de FGF21 a través de la ligadura química nativa u otros métodos conocidos en la técnica. La realización preferida deja un extremo-N libre de GLP-1 (exendina-4) para una actividad máxima y un extremo-C intacto, libre de FGF21 para actividad máxima.

[00025] En algunas realizaciones, el péptido GLP-1 o Exendina-4 se une covalentemente a una molécula de polímero que a su vez está unida a la variante de la proteína FGF21. En una realización preferida, el péptido GLP-1 o exendina 4- está unido a un polímero de PEG que está simultáneamente unido a una variante de la proteína FGF21. Dicha realización preferida deja libre un extremo N para máxima actividad de GLP-1 (exendina-4) y un extremo-C libre e intacto para máxima actividad de FGF21. En algunas formas de realización de la invención, un péptido agonista del receptor GLP-1 está conectado a una variante de FGF21 a través de un enlazador PEG u otro enlazador polimérico que proporciona simultáneamente la extensión de media-vida, así como una conexión covalente para los dos agonistas de receptores.

[00026] Sin embargo, otra realización se relaciona con métodos de tratar a un paciente que exhibe uno o más trastornos asociados con FGF21 o trastornos asociados con GLP-

1, como la obesidad, diabetes mellitus tipo 2, diabetes mellitus tipo 1, pancreatitis, dislipidemia, enfermedad hepática grasa no alcohólica, esteatohepatitis no alcohólica (EHNA), resistencia a la insulina, hiperinsulinemia, intolerancia a la glucosa, hiperglucemia, síndrome metabólico, infarto agudo del miocardio, hipertensión, enfermedad cardiovascular, arteriosclerosis, enfermedad arterial periférica, accidente cerebro vascular, insuficiencia cardiaca, enfermedad coronaria, enfermedad renal, complicaciones de la diabetes, neuropatía, gastroparesia, trastornos asociados con mutaciones inactivadoras en el receptor de insulina, lipodistrofias incluyendo lipodistrofia asociada a VIH y otros trastornos metabólicos, los cuales comprenden administrar a dicho paciente en necesidad de tal tratamiento una cantidad terapéuticamente efectiva de una o más proteínas de la invención o una composición farmacéutica de las mismas.

[00027] La invención también proporciona composiciones farmacéuticas que comprenden las proteínas de función dual de la invención divulgadas en la presente y un agente de formulación farmacéuticamente aceptable. Estas composiciones farmacéuticas se pueden utilizar en un método para tratar un trastorno metabólico, y el método comprende administrar a un paciente humano que lo necesite, una composición farmacéutica del invento. Los ejemplos no limitantes de trastornos metabólicos que pueden ser tratados incluyen la diabetes mellitus tipo 1 y tipo 2 y la obesidad.

[00028] Estos y otros aspectos del invento se explicarán en la siguiente descripción detallada del invento.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[00029] Las figuras 1A y 1B muestran la actividad de las proteínas de fusión: GLP-1-FGF21-PEG con diferentes mutaciones N-terminal y proteínas de fusión de la Exendina-4-FGF21-PEG. La figura 1A muestra mutaciones agregadas para retardar el procesamiento por la proteasa DPP-4 (péptido GLP-1 (círculos) o proteínas de función dual con tipo silvestre (V231; cuadros vacíos), G0 (V251; círculos vacíos), A8G (V258; triángulos vacíos), A8S (V232; triángulos), o E9P (V271; cuadrados) GLP-1). En la figura 1B se observan las fusiones de la Exendina-4 (péptido GLP-1 (cuadrados) o proteínas de fusión dual con la Excendina-4 1-39 (V234; triángulos), 1-30 (V267; círculos) o 1-30/E16G/E17Q (V268; triángulos vacíos)).

[00030] La figura 2 indica las propiedades farmacocinéticas (PK) de la fusión de las proteínas de fusión GLP-1-FGF21-PEG (FGF21-PEG (V294; cuadrados vacíos), GLP-1(A8S)-PEG (V253; círculos vacíos) o proteínas de función dual con con GLP-1 tipo silvestre (V237; círculos) o GLP-1 (A8S) (V235; triángulos) GLP-1). La PK de FGF21 no-PEGilado tipo silvestre inyectado a 0,25 mg / kg se muestra para fines comparativos (cuadrados).

[00031] Las figuras 3A-3C muestran los resultados de la prueba de tolerancia oral de glucosa (PTGO), para medir la eficacia de la vida media prolongada de la fusión de proteínas GLP-1-FGF21-PEG. A los ratones de 8 semanas C57BL/6J (n = 5) se les administró por vía intraperitonea (ip) una inyección de 1 mg / kg de compuesto o vehículo (barras blancas sólidas; triangulos). La glucosa en sangre fue medida a la hora 1 y hora 24 después de inyectada la dosis para evaluar los efectos agudos de GLP-1, los cuales fueron similares para el GLP-1 tipo silvestre (barras rayadas; cuadrados) y versiones de GLP-1 A8S (V232; barras negras; círculos) en 1 hora y retenido ligeramente mejor por la versión GLP-1 (A8S) en 24 horas. En la tercera noche, los ratones se mantuvieron en ayunas antes de la exposición a 1,5 g / kg de glucosa por vía oral, 72 horas después de la dosis. Los ratones dosificados con la versión A8S manifestaron significativamente mejoría en el control de la glucosa en sangre en comparación de aquellos con la versión GLP-1 tipo silvestre, lo que sugiere que la mutación A8S aumentó a largo plazo los niveles activos de GLP-1.

[00032] La figura 4 muestra la eficacia de la proteína de fusión GLP-1-FGF21-PEG en ratones ob/ob. Ratones machos de diez semanas (n=8) fueron dosificados vía intraperitonea (ip) dos veces a la semana con los compuestos indicados durante dos semanas. La dosis equivalente de FGF21-PEG (V238; barras diagonales; cuadrados) y GLP-1FGF21R(R154C)-PEG (V239; barras horizantales rayadas; cuadrados vacíos), mostró mejorías muy similares en los niveles de glucosa en sangre, peso corporal y salud hepática (medida por los niveles séricos de ALT y el peso del hígado) cuando se compara con el grupo tratado con vehículo (circulos). La enzima resistente DPP-4 de fusión, 0,2 mg / kg de GLP-1 (A8S)-FGF21(R154C)-PEG (V235; barras verticales rayadas; triángulos vacíos) mostró una eficacia similar a otros compuestos, pero la dosificación igual a 1 mg / kg de GLP-1 (A8S)-FGF21(R154C)-PEG (V235; barras negras; triángulos) tuvo un descenso adicional de la glucosa en sangre, el peso corporal, la alanina aminotransferasa (ALT), y el peso del hígado. La fusión resistente DPP-4, 0,2 mg/kg de GLP-1 (A8S)-FGF21-PEG (barras verticales rayadas; triángulos vacíos)

expresó una eficacia similar a otros compuestos, pero la dosificación igual a 1 mg/kg de GLP-1 (A8S)-FGF21-PEG (barras negras; triángulos) dio un descenso adicional de la glucosa en sangre, el peso corporal, la alanina aminotransferasa (ALT) y el peso del hígado.

[00033] Las figuras 5A-5G muestran una comparación de la proteína de fusión GLP -1(A8S)-FGF21-(R154C)-PEG (V235; barras negras; triángulos vacíos) con co-administración de FGF21(R154C)-PEG + GLP-1 (A8S)-PEG (barras verticales de hash; círculos vacíos), la administración individual de FGF21(R154C)-PEG (V238; 0,2 mg / kg (barras rayadas diagonales); 1 mg /kg (barras diagonales puntadas; cuadrados), y finalmente la administración de GLP-1 (A8S)-PEG (V253; 0,2 mg / kg (barras rayadas horizontales); 1 mg / kg (barras horizontales, oscuras rayadas; cuadrados vacíos)) en ratones ob / ob. Los ratones machos ob/ob de nueve semanas de edad (n = 8) se dosificaron (ip) dos veces por semana con los compuestos indicados o vehículo (barras blancas sólidas; circulos) durante cuatro semanas. La fusión de proteínas mostró una significativa mejoría en la eficacia en comparación con la co-administración de FGF21(R154C)-PEG + GLP-1 (A8S)-PEG. Aunque sólo la pérdida de peso moderada se observó con FGF21(R154C)-PEG y GLP-1 (A8S)-FGF21(R154C)-PEG, otros grupos subieron sustancialmente de peso, sólo en parte debido a diferencias en la ingesta de alimentos.

[00034] Las figuras 6A-6F muestran una eficacia de las proteínas de fusión GLP-1(A8S)-FGF21-PEG en ratones ob/ob, comparando FGF21(R154C)-PEG (V235; barras oscuras) con FGF21(V76)-PEG (V272; barras claras). Los ratones ob/ob machos de nueve semanas de edad fueron dosificados por vía intraperitonea dos veces a la semana por dos semanas con los compuestos indicados o vehículo (barras blancas sólidas). Tanto la proteína de fusión GLP-1(A8S)-FGF21(R154C)-PEG como la GLP-1(A8S)-FGF21(V76)-PEG mostraron una repuesta a la dosis de 0.05 a 0.2 mg/kg (0.05 barras rayadas diagonales; 0.1 barras horizontales rayadas; 0.2 barras rayadas, horizontales) para control de glucosa y peso corporal. A una dosis de 0.2 mg/kg, GLP-1(A8S)-FGF21(V76)-PEG fue mas eficaz que GLP-1(A8S)-FGF21 FGF21(R154C)-PEG, particularmente por la reducción de niveles de glucosa, peso corporal, triglicéridos séricos y colesterol. La figuras 6A-6C, demuestran niveles de glucosa basal (ABC), alanina aminotransferasa (ALT) y colesterol total en suero, respectivamente. Las figuras 6C-6F evidencian los niveles de D12 en peso corporal, líquidos hapáticos y triglicéridos en suero respectivamente.

[00035] Las figuras 7A-78B indican la habilidad de las proteína de fusión GLP-1(A8S)-FGF21(V76)-PEG para mejorar la función pacreática e incrementar el contenido de insulina en el islote, relacionado con la combinación de agentes individuales. Los ratones Db/db fueron dosificados dos veces por semana y por un tiempo de 4 semanas con una combinacón individual de GLP-1(A8S)-PEG +FGF21(variante 76)-PEG (V76+V253; barras rayadas, así como la proteína de la función dual GLP-1(A8S)-FGF21(V76)-PEG de la invención (V272; barras diagonales rayadas; barras claras para 0.5 mg/kg y barras oscuras para 1 mg/kg), o vehículo (barras claras).

[00036] Las figuras 8A-8B muestran la habilidad de la proteína de fusión GLP-1-FGF21-PEG para mejorar la disminución de la glucosa y el peso corporal relacionados con la combinación de agentes individuales. Los ratones Db/db fueron dosificados dos veces por semana, durante dos semanas, con vehículo (barras blancas, circulos abiertos), una combinación individual de GLP-1(A8S)-PEG (V253; 3 mg/kg (barras claras horizontales rayadas; rombos vacíos) and 5 mg/kg (barras oscuras horizontales; diamantes) y FGF21 (variante 76)-PEG (3 mg/kg (barra clara, diagonal rayada; cuadrados vacíos) y 5 mg/kg (barras oscuras, diagonales, rayadas; cuadrados)), así como con la proteína de fusión de la invención con función dual GLP-1(A8S)-FGF21(V76)-PEG (0.2 mg/kg (barras cortas con cuadrados; triángulos vacios); 1 mg/kg (barras largas con cuadrados; triángulos)). Como se ha visto en la figura, 0.02 mg/kg de la proteína de fusión con función dual GLP-1(7-35; A8S)-FGF21(V76)-PEG (V272) es tan eficaz como 5 mg / kg de FGF21 (V76)-PEG. Así mismo, se puede observar en la figura que 1.0 mg/kg de la proteína de fusión con función dual GLP-1(7-35; A8S)-FGF21(V76)-PEG (V272) es más efectiva que las dosis máximas eficaces de combinación de FGF21 (variante 76)-PEG + GLP-1(7-35; A8S)-PEG (1+1 mg/kg (barras horizonatales claras rayadas; x) y 3 + 3 mg/kg (barras oscuras horizontales; asteriscos), para los dos criterios de valoración, glicemia y peso corporal.

[00037] La Figura 9 muestra los resultados de un ensayo para la actividad de FGF21. La fosforilación de ERK se mide después del tratamiento de células HEK293-beta-Klotho con los compuestos indicados. La actividad del GLP-1 (A8S)-V76-PEG (V272) es significativamente mayor en magnitud de la señal en comparación con la misma variante FGF21 FGF21 (V76)-PEG en la presencia o ausencia de exenatida equimolar.

[00038] Figura 10 presenta los resultados de un ensayo de farmacología de receptores utilizados para medir la actividad de proteínas de fusión en comparación con las proteínas agonistas individuales coincidentes o péptidos. Células HEK293 transfectadas

con GLP-1R y FGFR1c/beta-klotho (10a y 10b) o con solo el GLP-1R (10c y 10d) se ensayaron para el reclutamiento de beta-arrestina a GLP-1R después del tratamiento con compuestos durante 1 hora. Las moléculas probadas en el ensayo fueron: Exendina-4 (cuadrados), GLP-1(A8S)-FGF21(V76)-PEG (V272; circulos), GLP-1(A8S)-PEG (V253; triangulos), V253 +FGF21(V76)-PEG (rombos abiertos), Exendina-4(1-39)-Fc-FGF21(V103) (V211; circulos abiertos), Exendina-4(1-39)-Fc (V201; triangulos abiertos), y V201+ FGF21(V101) (rombos).

DESCRIPCIÓN DETALLADA DEL INVENTO

[00039] Las proteínas de la presente invención representan versiones modificadas del polipéptido FGF21 tipo silvestre de longitud completa, como se conoce en el estado de la técnica. La secuencia FGF21 tipo silvestre servirá como una secuencia de referencia (SEQ ID NO: 1), por ejemplo, cuando las comparaciones entre la secuencia FGF21 tipo silvestre y los variantes proteicos son necesarios. La secuencia FGF21 tipo silvestre tiene un número NCBI de secuencia de referencia NP_061986.1, y se puede encontrar en tales patentes concedidas como, por ejemplo, EE.UU. 6,716,626 B1, asignada Chiron Corporation (SEQ ID NO:1).

```
Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Ser
 1               5                  10                  15
Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
            20                  25                  30
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
        35                  40                  45
Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
    50                  55                  60
Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
65                  70                  75                  80
Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
                85                  90                  95
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
            100                 105                 110
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
        115                 120                 125
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
    130                 135                 140
His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly
145                 150                 155                 160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
                165                 170                 175
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
```

180 185 190
Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
195 200 205
Ser
209

[00040] La secuencia correspondiente de ARNc que codifica el polipéptido de FGF21 de longitud completa (número NCBI de referencia de secuencia NM_019113.2) se muestra a continuación (SEQ ID NO:2)

1 ctgtcagctg aggatccagc cgaaagagga gccaggcact caggccacct gagtctactc
61 acctggacaa ctggaatctg gcaccaattc taaaccactc agcttctccg agctcacacc
121 ccggagatca cctgaggacc cgagccattg atggactcgg acgagaccgg gttcgagcac
181 tcaggactgt gggtttctgt gctggctggt cttctgctgg gagcctgcca ggcacacccc
241 atccctgact ccagtcctct cctgcaattc gggggccaag tccggcagcg gtacctctac
301 acagatgatg cccagcagac agaagcccac ctggagatca gggaggatgg gacggtgggg
361 ggcgctgctg accagagccc cgaaagtctc ctgcagctga aagccttgaa gccgggagtt
421 attcaaatct tgggagtcaa gacatccagg ttcctgtgcc agcggccaga tggggccctg
481 tatggatcgc tccactttga ccctgaggcc tgcagcttcc gggagctgct tcttgaggac
541 ggatacaatg tttaccagtc cgaagcccac ggcctcccgc tgcacctgcc agggaacaag
601 tccccacacc gggaccctgc accccgagga ccagctcgct tcctgccact accaggcctg
661 ccccccgcac tcccggagcc acccggaatc ctggcccccc agccccccga tgtgggctcc
721 tcggaccctc tgagcatggt gggaccttcc cagggccgaa gccccagcta cgcttcctga
781 agccagaggc tgtttactat gacatctcct ctttatttat taggttattt atcttattta
841 ttttttatt tttcttactt gagataataa agagttccag aggagaaaaa aaaaaaaaaa
901 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa

[00041] La secuencia madura FGF21 carece de una secuencia líder y también puede incluir otras modificaciones de un polipéptido como el procesamiento proteolítico del extremo amino (con o sin una secuencia líder) y/o el extremo carboxilo, la escisión de un polipéptido más pequeño desde un precursor más grande, glicosilación ligada a N y/o ligada a O, y otras modificaciones post-translacionales entendidas por los expertos en el estado de la técnica. Un ejemplo representativo de una secuencia madura FGF21 tiene la siguiente secuencia (SEQ ID NO: 3, la cual representa las posiciones aminoácidas 29-209 de la secuencia proteica de FGF21 de longitud completa (número NCBI de referencia de secuencia NP_061986.1)):

His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
5 10 15
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
20 25 30
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser

35 40 45
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
50 55 60
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly
65 70 75 80
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
85 90 95
Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
100 105 110
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
115 120 125
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
130 135 140
Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
145 150 155 160
Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser
165 170 175
Pro Ser Tyr Ala Ser
180

[00042] La secuencia de ADNc que codifica para el polipéptido maduro FGF21 (SEQ ID NO: 3) se muestra a continuación (SEQ ID NO: 4):

1 cacccatcc ctgactccag tcctctcctg caattcgggg gccaagtccg gcagcggtac
61 ctctacacag atgatgccca gcagacagaa gcccacctgg agatcaggga ggatgggacg
121 gtgggggggcg ctgctgacca gagccccgaa agtctcctgc agctgaaagc cttgaagccg
181 ggagttattc aaatcttggg agtcaagaca tccaggttcc tgtgccagcg gccagatggg
240 gccctgtatg gatcgctcca ctttgaccct gaggcctgca gcttccggga gctgcttctt
301 gaggacggat acaatgttta ccagtccgaa gcccacggcc tcccgctgca cctgccaggg
360 aacaagtccc cacaccggga ccctgcaccc cgaggaccag ctcgcttcct gccactacca
421 ggcctgcccc ccgcactccc ggagccaccc ggaatcctgg ccccccagcc ccccgatgtg
481 ggctcctcgg accctctgag catggtggga ccttcccagg gccgaagccc cagctacgct
541 tcctga

[00043] Las proteínas de la presente invención representan versiones modificadas del polipéptido GLP-1 tipo silvestre de longitud completa, como se conoce en el estado de la técnica. La secuencia GLP-1 tipo silvestre servirá como una secuencia de referencia (SEQ ID NO: 5), por ejemplo, cuando las comparaciones entre la secuencia GLP-1 tipo silvestre y los variantes proteicos son necesarias.

[00044] La secuencia tipo silvestre de GLP-1 es pos-traduccionalmente modificada, y de otra manera derivada a partir de la secuencia tipo silvestre de preproglucagón. La secuencia de preproglucagon (SEQ ID NO: 5) tiene número de secuencia de referencia NCBI, NP_002045, y puede encontrarse en publicaciones de patentes tales como, por ejemplo, WO98/19698 y WO87/06941A, asignadas a Eli Lilly y Co. y General Hospital Corp., respectivamente . Como se describe, por ejemplo, en Goke, (1991) European Journal of Clinical Investigation 21, 135, el GLP-1 es un péptido 37mer que se deriva de preproglucagón (GLP-1 constituye residuos 92-138 de preproglucagón, y se subraya en la SEQ ID NO: 5, a continuación). GLP-1 se procesa adicionalmente en un péptido activo 31mer, por escisión de los seis residuos N-terminales.

1 MKSIYFVAGL FVMLVQGSWQ RSLQDTEEKS RSFSASQADP LSDPDQMNED KRHSQGTFTS

61 DYSKYLDSRR AQDFVQWLMN TKRNRNNIAK RHDEFERHAE GTFTSDVSSY LEGQAAKEFI

121 AWLVKGRGRR DFPEEVAIVE ELGRRHADGS FSDEMNTILD NLAARDFINW LIQTKITDRK

[00045] Un ejemplo de la secuencia tipo silvestre procesada GLP-1 es como sigue (SEC ID NO:129).

1 HDEFERHAEG TFTSDVSSYL EGQAAKEFIA WLVKGRG

[00046] Un ejemplo de la secuencia GLP-1 activa es como sigue (SEQ ID: NO:30):

1 HAEGTFTSDV SSYLEGQAAK EFIAWLVKGR G

[00047] La secuencia de ARNc correspondiente que codifica la secuencia tipo silvestre de preproglucagón GLP-1 (SEQ ID NO: 5) se muestra a continuación (SEQ ID NO: 6):

1 gcatagaatg cagatgagca aagtgagtgg gagagggaag tcatttgtaa caaaaactca

61 ttatttacag atgagaaatt tatattgtca gcgtaatatc tgtgaggcta aacagagctg

121 gagagtatat aaaagcagtg cgccttggtg cagaagtaca gagcttagga cacagagcac

181 atcaaaagtt cccaaagagg gcttgctctc tcttcacctg ctctgttcta cagcacacta
241 ccagaagaca gcagaaatga aaagcattta ctttgtggct ggattatttg taatgctggt
301 acaaggcagc tggcaacgtt cccttcaaga cacagaggag aaatccagat cattctcagc
361 ttcccaggca gacccactca gtgatcctga tcagatgaac gaggacaagc gccattcaca
421 gggcacattc accagtgact acagcaagta tctggactcc aggcgtgccc aagattttgt
481 gcagtggttg atgaatacca agaggaacag gaataacatt gccaaacgtc acgatgaatt
541 tgagagacat gctgaaggga cctttaccag tgatgtaagt tcttatttgg aaggccaagc
601 tgccaaggaa ttcattgctt ggctggtgaa aggccgagga aggcgagatt tcccagaaga
661 ggtcgccatt gttgaagaac ttggccgcag acatgctgat ggttctttct ctgatgagat
721 gaacaccatt cttgataatc ttgccgccag ggactttata aactggttga ttcagaccaa
781 aatcactgac aggaaataac tatatcacta ttcaagatca tcttcacaac atcacctgct
841 agccacgtgg gatgtttgaa atgttaagtc ctgtaaattt aagaggtgta ttctgaggcc
901 acattgcttt gcatgccaat aaataaattt tcttttagtg ttgtgtagcc aaaaattaca
961 aatggaataa agttttatca aaatattgct aaaatatcag ctttaaaata tgaaagtgct
1021 agattctgtt attttcttct tattttggat gaagtacccc aacctgttta catttagcga
1081 taaaattatt tttctatgat ataatttgta aatgtaaatt attccgatct gacatatctg
1141 cattataata ataggagaat agaagaactg gtagccacag tggtgaaatt ggaaagagaa
1201 ctttcttcct gaaacctttg tcttaaaaat actcagcttt caatgtatca aagatacaat
1261 taaataaaat tttcaagctt ctttaccaaa aaaaaaaa

[00048] Las proteínas de la presente invención representan versiones modificadas del polipéptido de Exendina-4 tipo silvestre de longitud completa, como se conoce en el estado de la técnica. La secuencia tipo silvestre de Exendina-4 servirá como una secuencia de referencia (SEQ ID NO: 7), por ejemplo, cuando las comparaciones entre la secuencia tipo silvestre de Exendina-4 y los variantes proteicos son necesarias.

[00049] La secuencia tipo silvestre de exendina-4 tiene número de secuencia de referencia NCBI, GenBank: AAB22006.1, y puede encontrarse en tales patentes emitidas como, por ejemplo, EE.UU. 5.424.286, asignada a Amylin Pharmaceuticals, Inc. y Eli Lilly y Co. Un ejemplo de la secuencia tipo silvestre es la siguiente (SEQ ID NO: 7) :

1 HGEGTFTSDL SKQME**EE**AVR LFIEWLKNGG PSSGAPPPS

[00050] Las proteínas de la invención pueden comprender variantes proteicos o mutantes de las proteínas tipo silvestre listadas en la presente, por ejemplo, variantes de FGF21, variantes de GLP-1, y/o variantes de Exendina-4. Como se utilizan en la presente, los términos “variante proteico”, “variante humano”, “polipéptido o variante proteico,” “variante”, ”mutante,” así como cualquier término similar o versiones

específicas de los mismos (por ejemplo, "variante proteico de FGF21,"variante de GLP humano," "polipéptido de Exendina-4 o variante proteico," “variante”, "mutante FGF21," etc.) definen secuencias de proteínas o polipéptidos que comprenden modificaciones, truncamientos, otros variantes de las contrapartes de proteínas o polipéptidos de origen natural *es decir, tipo silvestre) o secuencias nativas correspondientes. “FGF21 variante” o “mutante de FGF21,” por ejemplo, se describe con relación a la proteína FGF21 tipo silvestre (es decir, de ocurrencia natural) como se describe en la presente. Las secuencias de proteínas de función dual representativas de la invención se listan en la Tabla 1. Las descripciones de dichos agonistas incluyen agonistas constituyentes individuales y, donde es aplicable, un enlazador. Si un variante se utiliza como un agonista constituyente, los cambios o sustituciones hechos se enumeran con relación a sus contrapartes tipo silvestre. A modo de ejemplo, la "proteína de función dual 1" (SEQ ID NO: 9) contiene los residuos 7-35 de la secuencia GLP1 tipo silvestre como se describe en la presente (es decir, un agonista del receptor GLP1), una secuencia de unión, y una variante de FGF21 (es decir, un agonista del receptor de FGF21) con un número de cambios listados a la secuencia tipo silvestre de FGF21 como se describe en la presente. **Tabla 1. Proteínas y Nucleótidos de Función Dual de la Invención**

SEC ID NO:	Variante NO:	Secuencia	Nombre Largo	Nombre Corto
9	V272	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76)-40KPEGb
10		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAT AGCAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76)

		GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT TGCGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
11	V277	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGG SGGGGSGGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS	Exendina4(1-30)-SGGGGSGGGGSGGGGSGGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A)	Ex(1-30)-L20-FGF21(V76)-40KPEGb
12		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGGT TCTGGCGGCG GTGGTAGCGG TGGCGGCGGT GATAGCAGCC CGCTGCTGCA GTTTGGCGGC CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGG AAACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGCATCAGA	Exendina4(1-30)-SGGGGSGGGGSGGGGSGGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	Ex(1-30)-L20-FGF21(V76)

		GCCCGGAAAG CCTGCTGGAA CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGAAAC CGGATGGCGC GCTGTATGGC AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CCGTAGCCCG CATTGCGATC CGGCACCGCA GGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGGCGA TGGTGGGTCC GAGCCAGGGT CGTAGCCCGA GCTATGCGAG CTAA		
13	V220	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGDSSPLLQF GGQVRQRYLY TDDAQQTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	GLP-1(7-35;A8S)-GSG-FGF21(33-209)	GLP-1(A8S)-L3-FGF21
14		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGATA GCAGCCCGCT GCTGCAGTTT GGCGGCCAGG TGCGTCAGCG TTATCTGTAT ACCGATGATG CGCAGCAGAC CGAAGCGCAT CTGGAAATTC GTGAAGATGG CACCGTGGGC GGTGCGGCGG ATCAGAGCCC GGAAAGCCTG CTGCAGCTGA AAGCGCTGAA ACCGGGCGTG ATTCAGATTC TGGGCGTGAA AACCAGCCGT TTTCTGTGCC AGCGTCCGGA TGGCGCGCTG TATGGCAGCC TGCATTTTGA TCCGGAAGCG	GLP-1(7-35;A8S)-GSG-FGF21(33-209)	GLP-1(A8S)-L3-FGF21

		TGCAGCTTTC GTGAACTGCT GCTGGAAGAT GGCTATAACG TGTATCAGAG CGAAGCGCAT GGCCTGCCGC TGCATCTGCC GGGCAACAAA AGCCCGCATC GTGATCCGGC ACCGCGTGGT CCGGCGCGTT TTCTGCCGCT GCCGGGTCTG CCGCCGGCAC TGCCGGAACC GCCGGGTATT CTGGCCCCGC AGCCGCCGGA TGTTGGTAGC AGCGATCCGC TGTCTATGGT GGGTCCGAGC CAGGGTCGTA GCCCGAGCTA TGCGAGCTAA		
15	V219	HAEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGDSSPLLQF GGQVRQRYLY TDDAQQTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	GLP-1(7-35)-GSG-FGF21(33-209)	GLP-1-L3-FGF21
16		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGATA GCAGCCCGCT GCTGCAGTTT GGCGGCCAGG TGCGTCAGCG TTATCTGTAT ACCGATGATG CGCAGCAGAC CGAAGCGCAT CTGGAAATTC GTGAAGATGG CACCGTGGGC GGTGCGGCGG ATCAGAGCCC GGAAAGCCTG CTGCAGCTGA AAGCGCTGAA ACCGGGCGTG ATTCAGATTC TGGGCGTGAA AACCAGCCGT TTTCTGTGCC AGCGTCCGGA TGGCGCGCTG TATGGCAGCC TGCATTTTGA TCCGGAAGCG TGCAGCTTTC GTGAACTGCT GCTGGAAGAT GGCTATAACG TGTATCAGAG CGAAGCGCAT GGCCTGCCGC TGCATCTGCC GGGCAACAAA AGCCCGCATC GTGATCCGGC ACCGCGTGGT CCGGCGCGTT TTCTGCCGCT	GLP-1(7-35)-GSG-FGF21(33-209)	GLP-1-L3-FGF21

		GCCGGGTCTG CCGCCGGCAC TGCCGGAACC GCCGGGTATT CTGGCCCCGC AGCCGCCGGA TGTTGGTAGC AGCGATCCGC TGTCTATGGT GGGTCCGAGC CAGGGTCGTA GCCCGAGCTA TGCGAGCTAA		
17	V295	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGDSSPLLQF GGQVRQRYLY TDDAQQTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHRDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	Exendina4(1-39)-GGG-FGF21(33-209)	Ex(1-39)-L3-FGF21
18		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT GGTGGCGACT CGAGCCCGCT GCTGCAGTTT GGCGGCCAGG TGCGTCAGCG TTATCTGTAT ACCGATGATG CGCAGCAGAC CGAAGCGCAT CTGGAAATTC GTGAAGATGG CACCGTGGGC GGTGCGGCGG ATCAGAGCCC GGAAAGCCTG CTGCAGCTGA AAGCGCTGAA ACCGGGCGTG ATTCAGATTC TGGGCGTGAA AACCAGCCGT TTTCTGTGCC AGCGTCCGGA TGGCGCGCTG TATGGCAGCC TGCATTTTGA TCCGGAAGCG TGCAGCTTTC GTGAACTGCT GCTGGAAGAT GGCTATAACG TGTATCAGAG CGAAGCGCAT GGCCTGCCGC TGCATCTGCC GGGCAACAAA AGCCCGCATC GTGATCCGGC ACCGCGTGGT CCGGCGCGTT TTCTGCCGCT GCCGGGTCTG CCGCCGGCAC TGCCGGAACC GCCGGGTATT CTGGCCCCGC AGCCGCCGGA TGTTGGTAGC AGCGATCCGC TGTCTATGGT GGGTCCGAGC CAGGGTCGTA	Exendina4(1-39)-GGG-FGF21(33-209)	Ex(1-39)-L3-FGF21

		GCCCGAGCTA TGCGAGCTAA		
19	V224	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S)-L10-FGF21
20		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S)-L10-FGF21

21	V223	HAEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35)-GSGGGGSGGG-FGF21(33-209)	GLP-1-L10-FGF21
22		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35)-GSGGGGSGGG-FGF21(33-209)	GLP-1-L10-FGF21
23	V234	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGDSS PLLQFGGQVR QRYLYTDDAQ QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQR	Exendina4 (1-39)-GGGGSGGG-FGF21(33-209)	Ex(1-39)-L8-FGF21

		PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPGNKSPHRD PAPRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLS MVGPSQGRSP SYAS		
24		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT GGTGGTGGTT CTGGCGGTGG CGACTCGAGC CCGCTGCTGC AGTTTGGCGG CCAGGTGCGT CAGCGTTATC TGTATACCGA TGATGCGCAG CAGACCGAAG CGCATCTGGA AATTCGTGAA GATGGCACCG TGGGCGGTGC GGCGGATCAG AGCCCGGAAA GCCTGCTGCA GCTGAAAGCG CTGAAACCGG GCGTGATTCA GATTCTGGGC GTGAAAACCA GCCGTTTTCT GTGCCAGCGT CCGGATGGCG CGCTGTATGG CAGCCTGCAT TTTGATCCGG AAGCGTGCAG CTTTCGTGAA CTGCTGCTGG AAGATGGCTA TAACGTGTAT CAGAGCGAAG CGCATGGCCT GCCGCTGCAT CTGCCGGGCA ACAAAAGCCC GCATCGTGAT CCGGCACCGC GTGGTCCGGC GCGTTTTCTG CCGCTGCCGG GTCTGCCGCC GGCACTGCCG GAACCGCCGG GTATTCTGGC CCCGCAGCCG CCGGATGTTG GTAGCAGCGA TCCGCTGTCT ATGGTGGGTC CGAGCCAGGG TCGTAGCCCG AGCTATGCGA GCTAA	Exendina4 (1-39)-GGGGSGGG-FGF21(33-209)	Ex(1-39)-L8-FGF21
25	V193	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGG SGGGGSADKT HTCPPCPAPE AAGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY	GLP-1(7-35;A8S)-S$(GGGGS)_3$A-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L17-Fc-L15-FGF21(V103)

		PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGKGGGGSG GGGSGGGGSD SSPLLQFGGQ VRQRYLYTDD ACQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPCNRSPH RDPASRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGGSQAR SPSYAS		
26	V194	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGG SGGGGSADKT HTCPPCPAPE AAGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGKGGGGSG GGGSGGGGSD SSPLLQFGGQ VRQRYLYTDD ACQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPCNRSPH RDPASRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYES	GLP-1(7-35;A8S)-S$(GGGGS)_3$A-Fc-$(GGGGS)_3$-FGF21(33-209; Q55C-R105K-G148C-K150R-P158S- S195A-G202A-A208E)	GLP-1(A8S)-L17-Fc-L15-FGF21(V194)
27	V195	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGG SGGGGSADKT HTCPPCPAPE AAGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGKGGGGSG	GLP-1(7-35;A8S)-S$(GGGGS)_3$A-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-P158S- K150R-S195A-G202A-A208D)	GLP-1(A8S)-L17-Fc-L15-FGF21(V195)

		GGGSGGGGSD SSPLLQFGGQ VRQRYLYTDD ACQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPCNRSPH RDPASRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYDS		
28	V296	HGEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8G)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A)	GLP-1(A8G)-L10-FGF21(V76)-40KPEGb
29		CATGGTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAT AGCAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT TGCGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG	GLP-1(7-35;A8G)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	GLP-1(A8G)-L10-FGF21(V76)

		CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
34	V294	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH UDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209; R154Pcl-40 kDa PEG ramificado)	GLP-1(A8S)-L10-FGF21-154Pcl-40KPEGb
35		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TAGGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209; R154TAG)	GLP-1(A8S)-L10-FGF21-154TAG

38	V298	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGDSS PLLQFGGQVR QRYLYTDDAQ QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQR PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPGNKSPHUD PAPRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLS MVGPSQGRSP SYAS	Exendina4 (1-39)-GGGGSGGG-FGF21(33-209; R154TAG)-40 kDa PEG ramificado	Ex(1-39)-L8-FGF21(154Pcl)-40KPEGb
39		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT GGTGGTGGTT CTGGCGGTGG CGACTCGAGC CCGCTGCTGC AGTTTGGCGG CCAGGTGCGT CAGCGTTATC TGTATACCGA TGATGCGCAG CAGACCGAAG CGCATCTGGA AATTCGTGAA GATGGCACCG TGGGCGGTGC GGCGGATCAG AGCCCGGAAA GCCTGCTGCA GCTGAAAGCG CTGAAACCGG GCGTGATTCA GATTCTGGGC GTGAAAACCA GCCGTTTTCT GTGCCAGCGT CCGGATGGCG CGCTGTATGG CAGCCTGCAT TTTGATCCGG AAGCGTGCAG CTTTCGTGAA CTGCTGCTGG AAGATGGCTA TAACGTGTAT CAGAGCGAAG CGCATGGCCT GCCGCTGCAT CTGCCGGGCA ACAAAAGCCC GCATTAGGAT CCGGCACCGC GTGGTCCGGC GCGTTTTCTG CCGCTGCCGG GTCTGCCGCC GGCACTGCCG GAACCGCCGG GTATTCTGGC CCCGCAGCCG CCGGATGTTG GTAGCAGCGA TCCGCTGTCT ATGGTGGGTC CGAGCCAGGG TCGTAGCCCG AGCTATGCGA GCTAA	Exendina4 (1-39)-GGGGSGGG-FGF21(33-209; R154TAG)	Ex(1-39)-L8-FGF21(154TAG)
40	V235	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI	GLP-1(7-35; A8S)-GSGGGGSGGG-FGF21(33-209;R154C-40	GLP-1(A8S)-L8-FGF21(154C)-

		LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	kDa PEG ramificado)	40KPEGb
41		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S)-L8-FGF21(154C)
42	V239	HAEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1-L8-FGF21(154C)-40KPEGb

43		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1-L8-FGF21(154C)
44	V299	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGDSS PLLQFGGQVR QRYLYTDDAQ QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQR PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPGNKSPHCD PAPRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLS MVGPSQGRSP SYAS	Exendina4 (1-39)-GGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	Ex(1-39)-L8-FGF21(154C)-40KPEGb
45		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT	Exendina4 (1-39)-GGGGSGGG-FGF21(33-209;R154C)	Ex(1-39)-L8-FGF21(154C)

		GGTGGTGGTT CTGGCGGTGG CGACTCGAGC CCGCTGCTGC AGTTTGGCGG CCAGGTGCGT CAGCGTTATC TGTATACCGA TGATGCGCAG CAGACCGAAG CGCATCTGGA AATTCGTGAA GATGGCACCG TGGGCGGTGC GGCGGATCAG AGCCCGGAAA GCCTGCTGCA GCTGAAAGCG CTGAAACCGG GCGTGATTCA GATTCTGGGC GTGAAAACCA GCCGTTTTCT GTGCCAGCGT CCGGATGGCG CGCTGTATGG CAGCCTGCAT TTTGATCCGG AAGCGTGCAG CTTTCGTGAA CTGCTGCTGG AAGATGGCTA TAACGTGTAT CAGAGCGAAG CGCATGGCCT GCCGCTGCAT CTGCCGGGCA ACAAAAGCCC GCATTGCGAT CCGGCACCGC GTGGTCCGGC GCGTTTTCTG CCGCTGCCGG GTCTGCCGCC GGCACTGCCG GAACCGCCGG GTATTCTGGC CCCGCAGCCG CCGGATGTTG GTAGCAGCGA TCCGCTGTCT ATGGTGGGTC CGAGCCAGGG TCGTAGCCCG AGCTATGCGA GCTAA		
46	V271	HAPGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; E9P)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(E9P)-L10-FGF21(154C)-40KPEGb
47		CATGCGCCGG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG	GLP-1(7-35; E9P)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(E9P)-L10-FGF21(154C)

		GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
48	V251	GHAEGTFTSD VSSYLEGQAA KEFIAWLVKG GSGGGGSGGG DSSPLLQFGG QVRQRYLYTD DAQQTEAHLE IREDGTVGGA ADQSPESLLQ LKALKPGVIQ ILGVKTSRFL CQRPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNKSP HCDPAPRGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLSMVGPSQG RSPSYAS	GLP-1(6-35; 6G)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	G-GLP-1-L10-FGF21(154C)-40KPEGb
49		GGTCATGCGG AAGGCACTTT TACTAGCGAT GTTTCTAGCT ACCTGGAAGG CCAGGCTGCG AAAGAATTCA TCGCGTGGCT GGTTAAAGGC GGTTCTGGTG GTGGTGGTTC TGGCGGTGGC GACTCGAGCC CGCTGCTGCA GTTTGGCGGC CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGC AGACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGGATCAGA GCCCGGAAAG CCTGCTGCAG CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGCGTC	GLP-1(6-35; 6G)-GSGGGGSGGG-FGF21(33-209)	G-GLP-1-L10-FGF21(154C)

		CGGATGGCGC GCTGTATGGC AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CAAAAGCCCG CATTGCGATC CGGCACCGCG TGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGTCTA TGGTGGGTCC GAGCCAGGGT CGTAGCCCGA GCTATGCGAG CTAA		
50	V265	HSEGTFTADA SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-S14A-V16A)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(A8S; 14/16)-L10-FGF21(154C)-40KPEGb
51		CATTCTGAAG GCACTTTTAC TGCTGATGCT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC	GLP-1(7-35; A8S-S14A-V16A)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S; 14/16)-L10-FGF21(154C)

		CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
52	V270	HSEGTFTSDA AAYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-V16A-S17A-S18A)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(A8S; 16/17/18)-L10-FGF21(154C)-40KPEGb
53		CATTCTGAAG GCACTTTTAC TAGCGATGCT GCTGCTTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG	GLP-1(7-35; A8S-V16A-S17A-S18A)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S)-16/17/18)-L10-FGF21(154C)

		ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
54	V266	HSEGTFTSDV SSYAEGAAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-L20A-Q23A)-GSGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(A8S; 20/23)-L10-FGF21(154C)-40KPEGb
55		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACG CTGAAGGCGC TGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-L20A-Q23A)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S; 20/23)-L10-FGF21(154C)

58	V300	HSEGTFTSDS SPLLQFGGQV RQRYLYTDDA QQTEAHLEIR EDGTVGGAAD QSPESLLQLK ALKPGVIQIL GVKTSRFLCQ RPDGALYGSL HFDPEACSFR ELLLEDGYNV YQSEAHGLPL HLPGNKSPHC DPAPRGPARF LPLPGLPPAL PEPPGILAPQ PPDVGSSDPL SMVGPSQGRS PSYAS	GLP-1(7-14; A8S)-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(A8S; 7-14)-L0-FGF21(154C)-40KPEGb
59		CATTCTGAAG GCACTTTTAC TAGCGATAGC AGCCCGCTGC TGCAGTTTGG CGGCCAGGTG CGTCAGCGTT ATCTGTATAC CGATGATGCG CAGCAGACCG AAGCGCATCT GGAAATTCGT GAAGATGGCA CCGTGGGCGG TGCGGCGGAT CAGAGCCCGG AAAGCCTGCT GCAGCTGAAA GCGCTGAAAC CGGGCGTGAT TCAGATTCTG GGCGTGAAAA CCAGCCGTTT TCTGTGCCAG CGTCCGGATG GCGCGCTGTA TGGCAGCCTG CATTTTGATC CGGAAGCGTG CAGCTTTCGT GAACTGCTGC TGGAAGATGG CTATAACGTG TATCAGAGCG AAGCGCATGG CCTGCCGCTG CATCTGCCGG GCAACAAAAG CCCGCATTGC GATCCGGCAC CGCGTGGTCC GGCGCGTTTT CTGCCGCTGC CGGGTCTGCC GCCGGCACTG CCGGAACCGC CGGGTATTCT GGCCCCGCAG CCGCCGGATG TTGGTAGCAG CGATCCGCTG TCTATGGTGG GTCCGAGCCA GGGTCGTAGC CCGAGCTATG CGAGCTAA	GLP-1(7-14; A8S)-FGF21(33-209;R154C)	GLP-1(A8S;7-14)-L0-FGF21(154C)
60	V262	HSEGTFTSDV SSGGGGSGGG DSSPLLQFGG QVRQRYLYTD DAQQTEAHLE IREDGTVGGA ADQSPESLLQ LKALKPGVIQ ILGVKTSRFL CQRPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNKSP HCDPAPRGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLSMVGPSQG RSPSYAS	GLP-1(7-18; A8S)-GGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(A8S;7-18)-L8-FGF21(154C)-40KPEGb

61		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTTCTGGTG GTGGTGGTTC TGGCGGTGGC GACTCGAGCC CGCTGCTGCA GTTTGGCGGC CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGC AGACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGGATCAGA GCCCGGAAAG CCTGCTGCAG CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGCGTC CGGATGGCGC GCTGTATGGC AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CAAAAGCCCG CATTGCGATC CGGCACCGCG TGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGTCTA TGGTGGGTCC GAGCCAGGGT CGTAGCCCGA GCTATGCGAG CTAA	GLP-1(7-18; A8S)-GGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S;7-18)-L8-FGF21(154C)
62	V261	HSEGTFTSDV SSYLESGGGG SGGGDSSPLL QFGGQVRQRY LYTDDAQQTE AHLEIREDGT VGGAADQSPE SLLQLKALKP GVIQILGVKT SRFLCQRPDG ALYGSLHFDP EACSFRELLL EDGYNVYQSE AHGLPLHLPG NKSPHCDPAP RGPARFLPLP GLPPALPEPP GILAPQPPDV GSSDPLSMVG PSQGRSPSYA S	GLP-1(7-21; A8S)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(A8S;7-21)-L9-FGF21(154C)-40KPEGb
63		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAATCTGG TGGTGGTGGT TCTGGCGGTG GCGACTCGAG CCCGCTGCTG CAGTTTGGCG GCCAGGTGCG TCAGCGTTAT CTGTATACCG ATGATGCGCA GCAGACCGAA GCGCATCTGG AAATTCGTGA AGATGGCACC GTGGGCGGTG CGGCGGATCA	GLP-1(7-21; A8S)-SGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S;7-21)-L9-FGF21(154C)

		GAGCCCGGAA AGCCTGCTGC AGCTGAAAGC GCTGAAACCG GGCGTGATTC AGATTCTGGG CGTGAAAACC AGCCGTTTTC TGTGCCAGCG TCCGGATGGC GCGCTGTATG GCAGCCTGCA TTTTGATCCG GAAGCGTGCA GCTTTCGTGA ACTGCTGCTG GAAGATGGCT ATAACGTGTA TCAGAGCGAA GCGCATGGCC TGCCGCTGCA TCTGCCGGGC AACAAAAGCC CGCATTGCGA TCCGGCACCG CGTGGTCCGG CGCGTTTTCT GCCGCTGCCG GGTCTGCCGC CGGCACTGCC GGAACCGCCG GGTATTCTGG CCCCGCAGCC GCCGGATGTT GGTAGCAGCG ATCCGCTGTC TATGGTGGGT CCGAGCCAGG GTCGTAGCCC GAGCTATGCG AGCTAA		
66	V260	HSEGTFTSDV SSYLEGQAAK EFISGGGGSG GGDSSPLLQF GGQVRQRYLY TDDAQQTEAH LEIREDGTVG GAADQSPESL LQLKALKPGV IQILGVKTSR FLCQRPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNK SPHCDPAPRG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLSMVGPS QGRSPSYAS	GLP-1(7-29; A8S)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(A8S;7-29)-L9-FGF21(154C)-40KPEGb
67		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCT CTGGTGGTGG TGGTTCTGGC GGTGGCGACT CGAGCCCGCT GCTGCAGTTT GGCGGCCAGG TGCGTCAGCG TTATCTGTAT ACCGATGATG CGCAGCAGAC CGAAGCGCAT CTGGAAATTC GTGAAGATGG CACCGTGGGC GGTGCGGCGG ATCAGAGCCC GGAAAGCCTG CTGCAGCTGA AAGCGCTGAA ACCGGGCGTG ATTCAGATTC TGGGCGTGAA AACCAGCCGT TTTCTGTGCC AGCGTCCGGA TGGCGCGCTG TATGGCAGCC TGCATTTTGA TCCGGAAGCG	GLP-1(7-29; A8S)-SGGGGSGGG-FGF21(33-209)	GLP-1(A8S;7-29)-L9-FGF21(154C)

		TGCAGCTTTC GTGAACTGCT GCTGGAAGAT GGCTATAACG TGTATCAGAG CGAAGCGCAT GGCCTGCCGC TGCATCTGCC GGGCAACAAA AGCCCGCATT GCGATCCGGC ACCGCGTGGT CCGGCGCGTT TTCTGCCGCT GCCGGGTCTG CCGCCGGCAC TGCCGGAACC GCCGGGTATT CTGGCCCCGC AGCCGCCGGA TGTTGGTAGC AGCGATCCGC TGTCTATGGT GGGTCCGAGC CAGGGTCGTA GCCCGAGCTA TGCGAGCTAA		
68	V259	HSEGTFTSDV SSYLEGQAAK EFIAWLSGGG GSGGGDSSPL LQFGGQVRQR YLYTDDAQQT EAHLEIREDG TVGGAADQSP ESLLQLKALK PGVIQILGVK TSRFLCQRPD GALYGSLHFD PEACSFRELL LEDGYNVYQS EAHGLPLHLP GNKSPHCDPA PRGPARFLPL PGLPPALPEP PGILAPQPPD VGSSDPLSMV GPSQGRSPSY AS	GLP-1(7-32; A8S)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(A8S;7-32)-L9-FGF21(154C)-40KPEGb
69		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGTC TGGTGGTGGT GGTTCTGGCG GTGGCGACTC GAGCCCGCTG CTGCAGTTTG GCGGCCAGGT GCGTCAGCGT TATCTGTATA CCGATGATGC GCAGCAGACC GAAGCGCATC TGGAAATTCG TGAAGATGGC ACCGTGGGCG GTGCGGCGGA TCAGAGCCCG GAAAGCCTGC TGCAGCTGAA AGCGCTGAAA CCGGGCGTGA TTCAGATTCT GGGCGTGAAA ACCAGCCGTT TTCTGTGCCA GCGTCCGGAT GGCGCGCTGT ATGGCAGCCT GCATTTTGAT CCGGAAGCGT GCAGCTTTCG TGAACTGCTG CTGGAAGATG GCTATAACGT GTATCAGAGC GAAGCGCATG GCCTGCCGCT GCATCTGCCG GGCAACAAAA GCCCGCATTG CGATCCGGCA CCGCGTGGTC CGGCGCGTTT	GLP-1(7-32; A8S)-SGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S;7-32)-L9-FGF21(154C)

		TCTGCCGCTG CCGGGTCTGC CGCCGGCACT GCCGGAACCG CCGGGTATTC TGGCCCCGCA GCCGCCGGAT GTTGGTAGCA GCGATCCGCT GTCTATGGTG GGTCCGAGCC AGGGTCGTAG CCCGAGCTAT GCGAGCTAA		
70	V263	HSEGTFTSDV CSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-S17C-40 kDa PEG ramificado)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-S17C)-40KPEGb-L10-FGF21
71		CATTCTGAAG GCACTTTTAC TAGCGATGTT TGTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT	GLP-1(7-35; A8S-S17C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-S17C)-L10-FGF21

		ATGCGAGCTA A		
72	V269	HSEGTFTSDV SSYLEGCAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-Q23C-40 kDa PEG ramificado)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-Q23C)-40KPEGb-L10-FGF21
73		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCTG TGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-Q23C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-Q23C)-L10-FGF21

74	V243	HSEGTFTSDV SSYLEGQAAC EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-K26C-40 kDa PEG ramificado)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-K26C)-L10-FGF21-40KPEGb
75		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGTGT GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-K26C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-K26C)-L10-FGF21
76	V264	HSEGTFTSDV SSYLEGQAAK CFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS	GLP-1(7-35; A8S-E27C-40 kDa PEG ramificado)-GSGGGGSGGG-FGF21(33-	GLP-1(A8S-E27C)-40KPEGb-L10-FGF21

		LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	209)	
77		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA TGTTTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-E27C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-E27C)-L10-FGF21
78	V244	HSEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-K34C-40 kDa PEG ramificado)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-K34C)-L10-FGF21-40KPEGb

79		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-K34C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(A8S-K34C)-L10-FGF21
80	V250	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGCGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S)-GSGGGGC(40 kDa PEG ramificado)GGG-FGF21(33-209)	GLP-1(A8S)-L10-C-FGF2140 KPEGb
81		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTGTGG CGGTGGCGAC	GLP-1(7-35; A8S)-GSGGGGCGGG-FGF21(33-209)	GLP-1(A8S)-L10-C-FGF21

		TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
82	V240	HAEGTFTSDV SSYLEGQAAC EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; K26C-40 kDa PEG ramificado)-GSGGGGSGGG-FGF21(33-209)	GLP-1(K26C)-L10-FGF21-40KPEGb
83		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGTGT GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG	GLP-1(7-35; K26C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(K26C)-L10-FGF21

		GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
84	V241	HAEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; K34C-40 kDa PEG ramificado)-GSGGGGSGGG-FGF21(33-209)	GLP-1-(K34C)-L10-FGF21-40KPEGb
85		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC	GLP-1(7-35; K34C)-GSGGGGSGGG-FGF21(33-209)	GLP-1(K34C)-L10-FGF21

		CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
86	V242	HAEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGCGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH RDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35)-GSGGGGC(40 kDa PEG ramificado)GGG-FGF21(33-209)	GLP-1-L10-C-FGF21-40KPEGb
87		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTGTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT	GLP-1(7-35)-GSGGGGCGGG-FGF21(33-209)	GLP-1-L10-C-FGF21

		CGTGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
88	V267	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	Exendina4(1-30)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	Ex(1-30)-L9-FGF21(154C)-40KPEGb
89		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG	Exendina4(1-30)-SGGGGSGGG-FGF21(33-209;R154C)	Ex(1-30)-L9-FGF21(154C)

		CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
90	V268	HGEGTFTSDL SKQMEGQAVR LFIEWLKNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	Exendina4(1-30;E16G-E17Q)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	Ex(1-30;GQ)-L9-FGF21(154C)-40KPEGb
91		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGGGCA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	Exendina4(1-30;E16G-E17Q)-SGGGGSGGG-FGF21(33-209;R154C)	Ex(1-30;GQ)-L9-FGF21(154C)

92	V301	HSEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH UDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S-K34C)-GSGGGGSGGG-FGF21(33-209;R154Pcl)	GLP-1(A8S-K34C)-FGF21(154Pcl)-
93		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TAGGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-K34C)-GSGGGGSGGG-FGF21(33-209;R154TAG)	GLP-1(A8S-K34C)-L10-FGF21(154TAG)
94	V302	HAEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS	GLP-1(7-35; K34C)-GSGGGGSGGG-FGF21(33-209;R154Pcl)	GLP-1(K34C)-L10-FGF21(154Pcl)

		LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH UDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS		
95		CATGCGGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TAGGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; K34C)-GSGGGGSGGG-FGF21(33-209;R154TAG)	GLP-1(K34C)-L10-FGF21(154TAG)
96	V303	HSEGTFTSDV SSYLEGQAAK EFIAWLVUGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQQTEAHLEI REDGTVGGAA DQSPESLLQL KALKPGVIQI LGVKTSRFLC QRPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNKSPH CDPAPRGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LSMVGPSQGR SPSYAS	GLP-1(7-35; A8S- K34Pcl)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S-K34Pcl)-L10-FGF21(154C)

97		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTAGGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGCAGA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG GATCAGAGCC CGGAAAGCCT GCTGCAGCTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGCGTCCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACAA AAGCCCGCAT TGCGATCCGG CACCGCGTGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGTCTATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	GLP-1(7-35; A8S-K34TAG)-GSGGGGSGGG-FGF21(33-209;R154C)	GLP-1(A8S-K34TAG)-L10-FGF21(154C)
98	V281	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76-154R)
99		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-	GLP-1(A8S)-L10-FGF21(V76-154R)

		TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	R105K-K150R-R159Q-S195A)	
100	V304	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76-154R-130E)
101		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76-154R-130E)

		CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
102	V273	HSEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S-K34C-40 kDa PEG ramificado)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	GLP-1(A8S-K34C)-L10-FGF21(V76-154R)-40KPEGb
103		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC	GLP-1(7-35;A8S-K34C)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	GLP-1(A8S-K34C)-L10-FGF21(V76-154R)

		CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
104	V305	HSEGTFTSDV SSYLEGQAAK EFIAWLVCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYAS	GLP-1(7-35;A8S-K34C-40 kDa PEG ramificado)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	GLP-1(A8S-K34C)-L10-FGF21(V76-154R-202A)-40KPEGb
105		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TTGTGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT	GLP-1(7-35;A8S-K34C)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	GLP-1(A8S-K34C)-L10-FGF21(V76-154R-202A)

		CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGCGCGT AGCCCGAGCT ATGCGAGCTA A		
108	V306	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	Exendina4(1-30)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30)-L9-FGF21(V76-154R)
109		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG	Exendina4(1-30)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30)-L9-FGF21(V76-154R)

		CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A		
110	V274	HGEGTFTSDL SKQMEEEAVR LFIEWLCNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	Exendina4(1-30;K27C-40 kDa PEG ramificado)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30;K27C)-L9-FGF21(V76-154R-130E)-40KPEGb
111		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGTG TAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	Exendina4(1-30;K27C)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30;K27C)-L9-FGF21(V76-154R-130E)

112	V307	HGEGTFTSDL SKQMEEEAVR LFIEWLKCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEEGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	Exendina4(1-30;N28C-40 kDa PEG ramificado)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30;N28C)-L9-FGF21(V76-154R-130E)-40KPEGb
113		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA ATGTGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA AGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGGTCGT AGCCCGAGCT ATGCGAGCTA A	Exendina4(1-30;N28C)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-D130E-K150R-R159Q-S195A)	Ex(1-30;N28C)-L9-FGF21(V76-154R-130E)
114	V308	HGEGTFTSDL SKQMEEEAVR LFIEWLCNGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS	Exendina4(1-30;K27C-40 kDa PEG ramificado)-SGGGGSGGG-FGF21(33-	Ex(1-30;K27C)--L9-FGF21(V76-154R-202A)-

		LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYAS	209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	40KPEGb
115		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGTG TAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGCGCGT AGCCCGAGCT ATGCGAGCTA A	Exendina4(1-30;K27C)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	Ex(1-30;K27C)-L9-FGF21(V76-154R-202A)
116	V309	HGEGTFTSDL SKQMEEEAVR LFIEWLKCGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH RDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQAR SPSYAS	Exendina4(1-30;N28C-40 kDa PEG ramificado)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	Ex(1-30;N28C)-L9-FGF21(V76-154R-202A) -40KPEGb

117		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA ATGTGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC TCGAGCCCGC TGCTGCAGTT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT CGTGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTCCGAG CCAGGCGCGT AGCCCGAGCT ATGCGAGCTA A	Exendina4(1-30;N28C)-SGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R159Q-S195A-G202A)	Ex(1-30;N28C)-L9-FGF21(V76-154R-202A)
118	V276	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GSGGGGSGGG GSGGGGSGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS	GLP-1(7-35;A8S)-GGS$(GGGGS)_3$-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A)	GLP-1(A8S)-L20-FGF21(V76)-40KPEGb
119		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT GGTAGCGGTG	GLP-1(7-35;A8S)-GGS$(GGGGS)_3$-FGF21(33-209;Q56E-	GLP-1(A8S)-L20-FGF21(V76)

		GCGGCGGTTC TGGTGGTGGT GGTTCTGGCG GTGGCGGTAG CGGTGGCGGC GATAGCAGCC CGCTGCTGCA GTTTGGCGGC CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGG AAACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGCATCAGA GCCCGGAAAG CCTGCTGGAA CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGAAAC CGGATGGCGC GCTGTATGGC AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CCGTAGCCCG CATTGCGATC CGGCACCGCA GGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGGCGA TGGTGGGTCC GAGCCAGGGT CGTAGCCCGA GCTATGCGAG CTAA	D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	
120	V197	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG GGGGSDKTHT CPPCPAPEAA GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GKGGGGSGGG GSGGGGSDSS PLLQFGGQVR QRYLYTDDAC QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQK PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPCNRSPHRD PASRGPARFL PLPGLPPALP	Exendina4(1-30)-(GGGGS) -Fc-$(GGGGS)_3$ -FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Exendin (1-30)-L5-Fc-L15-FGF21(V103))

		EPPGILAPQP PDVGSSDPLA MVGGSQARSP SYAS		
121	V196	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG GGGGSGGGGS GGGGSDKTHT CPPCPAPEAA GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GKGGGGSGGG GSGGGGSDSS PLLQFGGQVR QRYLYTDDAC QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQK PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPCNRSPHRD PASRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLA MVGGSQARSP SYAS	Exendina4(1-30)-$(GGGGS)_3$-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex (1-30)-L15-Fc-L15-FGF21(V103)
122	V199	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGGGGSGGGG SGGGGSDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP DVGSSDPLAM VGGSQARSPS YAS	Exendina4(1-39)-(GGGGS)-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex (1-39)-L5-Fc-L15-FGF21(V103)

123	V198	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGGGGSGGGG SGGGGSDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP DVGSSDPLAM VGGSQARSPS YAS	Exendina4(1-39)-$(GGGGS)_3$-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex (1-39)-L15-Fc-L15-FGF21(V103)
124	V203	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGGGGSGGGG SGGGGSDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP DVGSSDPLAM VGGSQARSPS YAS	GLP-1(7-35; A8S)-(GGGGS)-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L5-Fc-L15-FGF21(V103)

125	V202	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGGGGSGGGG SGGGGSDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP DVGSSDPLAM VGGSQARSPS YAS	GLP-1(7-35; A8S)-$(GGGGS)_3$-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L15-FGF21(V103)
126	V310	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSGGGGDS SPLLQFGGQV RQRYLYTDDA QETEAHLEIR EDGTVGGAAH QSPESLLELK ALKPGVIQIL GVKTSRFLCQ KPDGALYGSL HFDPEACSFR ELLLEDGYNV YQSEAHGLPL HLPGNRSPHC DPAPQGPARF LPLPGLPPAL PEPPGILAPQ PPDVGSSDPL AMVGPSQGRS PSYAS	Exendina4(1-39)-$(GGGGS)_3$GGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A)	Ex(1-39)-L19-FGF21(V76)-40KPEGb
127		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT CCGTCCTCCG GCGCTCCTCC GCCTTCTGGT GGTGGTGGTT CTGGCGGTGG CGGTTCTGGC GGCGGTGGTA GCGGTGGCGG CGGTGATAGC AGCCCGCTGC TGCAGTTTGG CGGCCAGGTG CGTCAGCGTT ATCTGTATAC CGATGATGCG CAGGAAACCG AAGCGCATCT GGAAATTCGT GAAGATGGCA CCGTGGGCGG TGCGGCGCAT	Exendina4(1-39)-$(GGGGS)_3$GGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	Ex(1-39)-L19-FGF21(V76)

		CAGAGCCCGG AAAGCCTGCT GGAACTGAAA GCGCTGAAAC CGGGCGTGAT TCAGATTCTG GGCGTGAAAA CCAGCCGTTT TCTGTGCCAG AAACCGGATG GCGCGCTGTA TGGCAGCCTG CATTTTGATC CGGAAGCGTG CAGCTTTCGT GAACTGCTGC TGGAAGATGG CTATAACGTG TATCAGAGCG AAGCGCATGG CCTGCCGCTG CATCTGCCGG GCAACCGTAG CCCGCATTGC GATCCGGCAC CGCAGGGTCC GGCGCGTTTT CTGCCGCTGC CGGGTCTGCC GCCGGCACTG CCGGAACCGC CGGGTATTCT GGCCCCGCAG CCGCCGGATG TTGGTAGCAG CGATCCGCTG GCGATGGTGG GTCCGAGCCA GGGTCGTAGC CCGAGCTATG CGAGCTAA		
31	V206	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSD KTHTCPPCPA PEAAGGPSVF LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR VVSVLTVLHQ DWLNGKEYKC KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGKDSSP LLQFGGQVRQ RYLYTDDACQ TEAHLEIRED GTVGGAADQS PESLLQLKAL KPGVIQILGV KTSRFLCQKP DGALYGSLHF DPEACSFREL LLEDGYNVYQ SEAHGLPLHL PCNRSPHRDP ASRGPARFLP LPGLPPALPE PPGILAPQPP DVGSSDPLAM VGGSQARSPS YAS	Exendina4(1-39)-Fc-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L0-FGF21(V103)
32	V208	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSD KTHTCPPCPA PEAAGGPSVF LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR	Exendina4(1-39)-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-	Ex(1-39)-L0-Fc-L2-FGF21(V101)

		VVSVLTVLHQ DWLNGKEYKC KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGKGSDS SPLLQFGGQV RQRYLYTDDA CQTEAHLEIR EDGTVGGAAD QSPESLLQLK ALKPGVIQIL GVKTSRFLCQ RPDGTLYGSL HFDPEACSFR ELLLEDGYNV YQSEAHGLPL HLPCNRSPHR DPASRGPARF LPLPGLPPAL PEPPGILAPQ PPDVGSSDPL AMVGGSQARS PSYAS	S195A-P199G-G202A)	
33		CACGGAGAAG GCACCTTTAC ATCGGACTTG TCGAAGCAGA TGGAGGAAGA AGCGGTGAGG CTCTTCATCG AGTGGCTCAA GAATGGAGGA CCCTCAAGCG GAGCGCCTCC TCCTTCCGAC AAAACCCATA CATGTCCGCC TTGTCCCGCA CCAGAAGCAG CGGGTGGGCC CTCGGTGTTC CTGTTCCCGC CAAAACCGAA GGACACACTT ATGATTTCAC GCACACCGGA AGTGACTTGC GTCGTGGTGG ATGTATCGCA CGAGGACCCC GAGGTCAAAT TCAACTGGTA TGTCGATGGA GTGGAGGTGC ACAATGCAAA GACCAAGCCG AGGGAAGAAC AATACAATAG CACGTACCGA GTCGTGTCCG TCTTGACGGT CCTTCACCAG GACTGGCTGA ACGGAAAGGA GTACAAGTGC AAAGTGAGCA ATAAGGCCCT CCCTGCCCCG ATTGAGAAAA CCATTTCCAA GGCCAAAGGT CAGCCTAGAG AACCTCAAGT GTATACTCTT CCGCCCTCAC GCGAAGAGAT GACGAAAAAC CAAGTGTCGC TTACGTGTCT TGTCAAAGGT TTCTACCCCT CGGACATCGC CGTAGAGTGG GAGTCGAACG GCCAGCCGGA GAACAACTAC AAGACCACGC CCCCTGTCTT GGATAGCGAC	Exendina4(1-39)-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L2-FGF21(V101)

		GGATCGTTTT TCCTCTACTC GAAACTCACA GTAGATAAGT CCCGATGGCA ACAGGGTAAT GTCTTTAGCT GCAGCGTGAT GCACGAGGCG CTTCACAATC ATTACACACA AAAATCACTG TCGCTTAGCC CGGGAAAGGG TTCAGATTCG TCGCCCCTGT TGCAGTTTGG TGGACAGGTC AGACAGCGCT ACCTTTACAC GGATGACGCC TGCCAGACAG AGGCACACCT CGAAATCAGA GAGGACGGTA CGGTCGGGGG TGCGGCCGAT CAGAGCCCCG AGTCGCTTCT CCAGTTGAAG GCCCTTAAGC CAGGAGTCAT CCAGATTTTG GGAGTAAAGA CCTCACGGTT TCTCTGTCAG CGTCCAGATG GGACACTGTA CGGCTCATTG CATTTCGATC CCGAAGCGTG CTCGTTCCGG GAGTTGCTGC TTGAGGACGG ATATAACGTC TATCAGAGCG AAGCGCATGG CCTCCCCCTT CACCTCCCGT GTAACAGGTC GCCGCATCGG GATCCGGCCT CGAGGGGTCC CGCGAGATTT CTTCCGTTGC CCGGGTTGCC TCCCGCGCTG CCCGAGCCTC CCGGGATCCT CGCGCCACAG CCTCCTGATG TAGGGTCCTC GGACCCTTTG GCGATGGTAG GTGGATCACA AGCACGGTCC CCGAGCTATG CATCA		
36	V209	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV	Exendina4(1-39)-$(GGGGS)_3$–Fc–GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V101)

		GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQRPDGT LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS		
37		CACGGAGAAG GCACCTTTAC ATCGGACTTG TCGAAGCAGA TGGAGGAAGA AGCGGTGAGG CTCTTCATCG AGTGGCTCAA GAATGGAGGA CCCTCAAGCG GAGCGCCTCC TCCTTCCGGA GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGGTTCAG ATTCGTCGCC	Exendina4(1-39)-$(GGGGS)_3$–Fc–GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V101)

		CCTGTTGCAG TTTGGTGGAC AGGTCAGACA GCGCTACCTT TACACGGATG ACGCCTGCCA GACAGAGGCA CACCTCGAAA TCAGAGAGGA CGGTACGGTC GGGGGTGCGG CCGATCAGAG CCCCGAGTCG CTTCTCCAGT TGAAGGCCCT TAAGCCAGGA GTCATCCAGA TTTTGGGAGT AAAGACCTCA CGGTTTCTCT GTCAGCGTCC AGATGGGACA CTGTACGGCT CATTGCATTT CGATCCCGAA GCGTGCTCGT TCCGGGAGTT GCTGCTTGAG GACGGATATA ACGTCTATCA GAGCGAAGCG CATGGCCTCC CCCTTCACCT CCCGTGTAAC AGGTCGCCGC ATCGGGATCC GGCCTCGAGG GGTCCCGCGA GATTTCTTCC GTTGCCCGGG TTGCCTCCCG CGCTGCCCGA GCCTCCCGGG ATCCTCGCGC CACAGCCTCC TGATGTAGGG TCCTCGGACC CTTTGGCGAT GGTAGGTGGA TCACAAGCAC GGTCCCCGAG CTATGCATCA		
133	V210	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSD KTHTCPPCPA PEAAGGPSVF LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR VVSVLTVLHQ DWLNGKEYKC KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGKGSDS SPLLQFGGQV RQRYLYTDDA CQTEAHLEIR EDGTVGGAAD QSPESLLQLK ALKPGVIQIL GVKTSRFLCQ KPDGALYGSL HFDPEACSFR ELLLEDGYNV YQSEAHGLPL HLPCNRSPHR DPASRGPARF LPLPGLPPAL PEPPGILAPQ PPDVGSSDPL AMVGGSQARS PSYAS	Exendina4(1-39-Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L2-FGF21(V103)

56		CACGGAGAAG GCACCTTTAC ATCGGACTTG TCGAAGCAGA TGGAGGAAGA AGCGGTGAGG CTCTTCATCG AGTGGCTCAA GAATGGAGGA CCCTCAAGCG GAGCGCCTCC TCCTTCCGAC AAAACCCATA CATGTCCGCC TTGTCCCGCA CCAGAAGCAG CGGGTGGGCC CTCGGTGTTC CTGTTCCCGC CAAAACCGAA GGACACACTT ATGATTTCAC GCACACCGGA AGTGACTTGC GTCGTGGTGG ATGTATCGCA CGAGGACCCC GAGGTCAAAT TCAACTGGTA TGTCGATGGA GTGGAGGTGC ACAATGCAAA GACCAAGCCG AGGGAAGAAC AATACAATAG CACGTACCGA GTCGTGTCCG TCTTGACGGT CCTTCACCAG GACTGGCTGA ACGGAAAGGA GTACAAGTGC AAAGTGAGCA ATAAGGCCCT CCCTGCCCCG ATTGAGAAAA CCATTTCCAA GGCCAAAGGT CAGCCTAGAG AACCTCAAGT GTATACTCTT CCGCCCTCAC GCGAAGAGAT GACGAAAAAC CAAGTGTCGC TTACGTGTCT TGTCAAAGGT TTCTACCCCT CGGACATCGC CGTAGAGTGG GAGTCGAACG GCCAGCCGGA GAACAACTAC AAGACCACGC CCCCTGTCTT GGATAGCGAC GGATCGTTTT TCCTCTACTC GAAACTCACA GTAGATAAGT CCCGATGGCA ACAGGGTAAT GTCTTTAGCT GCAGCGTGAT GCACGAGGCG CTTCACAATC ATTACACACA AAAATCACTG TCGCTTAGCC CGGGAAAGGG TTCAGATTCG TCGCCCCTGT TGCAGTTTGG TGGACAGGTC AGACAGCGCT ACCTTTACAC GGATGACGCC TGCCAGACAG AGGCACACCT CGAAATCAGA GAGGACGGTA CGGTCGGGGG TGCGGCCGAT CAGAGCCCCG AGTCGCTTCT CCAGTTGAAG GCCCTTAAGC CAGGAGTCAT CCAGATTTTG GGAGTAAAGA CCTCACGGTT TCTCTGTCAG	Exendina4(1-39-Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L2-FGF21(V103)

		AAACCAGATG GGGCACTGTA CGGCTCATTG CATTTCGATC CCGAAGCGTG CTCGTTCCGG GAGTTGCTGC TTGAGGACGG ATATAACGTC TATCAGAGCG AAGCGCATGG CCTCCCCCTT CACCTCCCGT GTAACAGGTC GCCGCATCGG GATCCGGCCT CGAGGGGTCC CGCGAGATTT CTTCCGTTGC CCGGGTTGCC TCCCGCGCTG CCCGAGCCTC CCGGGATCCT CGCGCCACAG CCTCCTGATG TAGGGTCCTC GGACCCTTTG GCGATGGTAG GTGGATCACA AGCACGGTCC CCGAGCTATG CATCA		
134	V211	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS	Exendina4(1-39)-$(GGGGS)_3$Fc–GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103)
57		CACGGAGAAG GCACCTTTAC ATCGGACTTG TCGAAGCAGA TGGAGGAAGA AGCGGTGAGG CTCTTCATCG AGTGGCTCAA GAATGGAGGA CCCTCAAGCG GAGCGCCTCC TCCTTCCGGA GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG	Exendina4(1-39)-$(GGGGS)_3$Fc–GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103)

		TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGGTTCAG ATTCGTCGCC CCTGTTGCAG TTTGGTGGAC AGGTCAGACA GCGCTACCTT TACACGGATG ACGCCTGCCA GACAGAGGCA CACCTCGAAA TCAGAGAGGA CGGTACGGTC GGGGGTGCGG CCGATCAGAG CCCCGAGTCG CTTCTCCAGT TGAAGGCCCT TAAGCCAGGA GTCATCCAGA TTTTGGGAGT AAAGACCTCA CGGTTTCTCT GTCAGAAACC AGATGGGGCA CTGTACGGCT CATTGCATTT CGATCCCGAA GCGTGCTCGT TCCGGGAGTT GCTGCTTGAG GACGGATATA ACGTCTATCA GAGCGAAGCG CATGGCCTCC CCCTTCACCT CCCGTGTAAC AGGTCGCCGC ATCGGGATCC GGCCTCGAGG GGTCCCGCGA		

		GATTTCTTCC GTTGCCCGGG TTGCCTCCCG CGCTGCCCGA GCCTCCCGGG ATCCTCGCGC CACAGCCTCC TGATGTAGGG TCCTCGGACC CTTTGGCGAT GGTAGGTGGA TCACAAGCAC GGTCCCCGAG CTATGCATCA		
135	V214	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQRPDGT LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS	GLP-1(7-35;A8S)-$(GGGGS)_3$–Fc–GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L2-FGF21(V101)
64		CACTCCGAAG GAACATTCAC TTCCGATGTA AGCTCGTATT TGGAAGGGCA GGCGGCTAAG GAGTTTATCG CATGGTTGGT CAAAGGTGGT GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT	GLP-1(7-35;A8S)-$(GGGGS)_3$–Fc–GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L2-FGF21(V101)

		GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGGTTCAG ATTCGTCGCC CCTGTTGCAG TTTGGTGGAC AGGTCAGACA GCGCTACCTT TACACGGATG ACGCCTGCCA GACAGAGGCA CACCTCGAAA TCAGAGAGGA CGGTACGGTC GGGGGTGCGG CCGATCAGAG CCCCGAGTCG CTTCTCCAGT TGAAGGCCCT TAAGCCAGGA GTCATCCAGA TTTTGGGAGT AAAGACCTCA CGGTTTCTCT GTCAGCGTCC AGATGGGACA CTGTACGGCT CATTGCATTT CGATCCCGAA GCGTGCTCGT TCCGGGAGTT GCTGCTTGAG GACGGATATA ACGTCTATCA GAGCGAAGCG CATGGCCTCC CCCTTCACCT CCCGTGTAAC AGGTCGCCGC ATCGGGATCC GGCCTCGAGG GGTCCCGCGA GATTTCTTCC GTTGCCCGGG TTGCCTCCCG CGCTGCCCGA GCCTCCCGGG ATCCTCGCGC CACAGCCTCC TGATGTAGGG TCCTCGGACC CTTTGGCGAT GGTAGGTGGA TCACAAGCAC GGTCCCCGAG CTATGCATCA		

136	V216	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS	GLP-1(7-35;A8S)-$(GGGGS)_3$–Fc–GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L2-FGF21(V103)
65		CACTCCGAAG GAACATTCAC TTCCGATGTA AGCTCGTATT TGGAAGGGCA GGCGGCTAAG GAGTTTATCG CATGGTTGGT CAAAGGTGGT GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA	GLP-1(7-35;A8S)-$(GGGGS)_3$–Fc–GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L2-FGF21(V103)

		GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGGTTCAG ATTCGTCGCC CCTGTTGCAG TTTGGTGGAC AGGTCAGACA GCGCTACCTT TACACGGATG ACGCCTGCCA GACAGAGGCA CACCTCGAAA TCAGAGAGGA CGGTACGGTC GGGGGTGCGG CCGATCAGAG CCCCGAGTCG CTTCTCCAGT TGAAGGCCCT TAAGCCAGGA GTCATCCAGA TTTTGGGAGT AAAGACCTCA CGGTTTCTCT GTCAGAAACC AGATGGGGCA CTGTACGGCT CATTGCATTT CGATCCCGAA GCGTGCTCGT TCCGGGAGTT GCTGCTTGAG GACGGATATA ACGTCTATCA GAGCGAAGCG CATGGCCTCC CCCTTCACCT CCCGTGTAAC AGGTCGCCGC ATCGGGATCC GGCCTCGAGG GGTCCCGCGA GATTTCTTCC GTTGCCCGGG TTGCCTCCCG CGCTGCCCGA GCCTCCCGGG ATCCTCGCGC CACAGCCTCC TGATGTAGGG TCCTCGGACC CTTTGGCGAT GGTAGGTGGA TCACAAGCAC GGTCCCCGAG CTATGCATCA		
137	V218	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE	GLP-1(7-35;A8S)-$(GGGGS)_3$–Fc-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L0-FGF21(V103)

		EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KDSSPLLQFG GQVRQRYLYT DDACQTEAHL EIREDGTVGG AADQSPESLL QLKALKPGVI QILGVKTSRF LCQKPDGALY GSLHFDPEAC SFRELLLEDG YNVYQSEAHG LPLHLPCNRS PHRDPASRGP ARFLPLPGLP PALPEPPGIL APQPPDVGSS DPLAMVGGSQ ARSPSYAS		
106		CACTCCGAAG GAACATTCAC TTCCGATGTA AGCTCGTATT TGGAAGGGCA GGCGGCTAAG GAGTTTATCG CATGGTTGGT CAAAGGTGGT GGAGGTGGGT CGGGCGGTGG AGGCTCCGGA GGGGGAGGGA GCGACAAAAC CCATACATGT CCGCCTTGTC CCGCACCAGA AGCAGCGGGT GGGCCCTCGG TGTTCCTGTT CCCGCCAAAA CCGAAGGACA CACTTATGAT TTCACGCACA CCGGAAGTGA CTTGCGTCGT GGTGGATGTA TCGCACGAGG ACCCCGAGGT CAAATTCAAC TGGTATGTCG ATGGAGTGGA GGTGCACAAT GCAAAGACCA AGCCGAGGGA AGAACAATAC AATAGCACGT ACCGAGTCGT GTCCGTCTTG ACGGTCCTTC ACCAGGACTG GCTGAACGGA AAGGAGTACA AGTGCAAAGT GAGCAATAAG GCCCTCCCTG CCCCGATTGA GAAAACCATT TCCAAGGCCA AAGGTCAGCC TAGAGAACCT CAAGTGTATA CTCTTCCGCC CTCACGCGAA GAGATGACGA AAAACCAAGT GTCGCTTACG TGTCTTGTCA AAGGTTTCTA CCCCTCGGAC ATCGCCGTAG AGTGGGAGTC GAACGGCCAG CCGGAGAACA ACTACAAGAC CACGCCCCCT GTCTTGGATA GCGACGGATC GTTTTTCCTC TACTCGAAAC TCACAGTAGA TAAGTCCCGA	GLP-1(7-35;A8S)-$(GGGGS)_3$–Fc-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L15-Fc-L0-FGF21(V103)

		TGGCAACAGG GTAATGTCTT TAGCTGCAGC GTGATGCACG AGGCGCTTCA CAATCATTAC ACACAAAAAT CACTGTCGCT TAGCCCGGGA AAGGATTCGT CGCCCCTGTT GCAGTTTGGT GGACAGGTCA GACAGCGCTA CCTTTACACG GATGACGCCT GCCAGACAGA GGCACACCTC GAAATCAGAG AGGACGGTAC GGTCGGGGGT GCGGCCGATC AGAGCCCCGA GTCGCTTCTC CAGTTGAAGG CCCTTAAGCC AGGAGTCATC CAGATTTTGG GAGTAAAGAC CTCACGGTTT CTCTGTCAGA AACCAGATGG GGCACTGTAC GGCTCATTGC ATTTCGATCC CGAAGCGTGC TCGTTCCGGG AGTTGCTGCT TGAGGACGGA TATAACGTCT ATCAGAGCGA AGCGCATGGC CTCCCCCTTC ACCTCCCGTG TAACAGGTCG CCGCATCGGG ATCCGGCCTC GAGGGGTCCC GCGAGATTTC TTCCGTTGCC CGGGTTGCCT CCCGCGCTGC CCGAGCCTCC CGGGATCCTC GCGCCACAGC CTCCTGATGT AGGGTCCTCG GACCCTTTGG CGATGGTAGG TGGATCACAA GCACGGTCCC CGAGCTATGC ATCA		
107	V200	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG K	Exendina4(1-39)-$(GGGGS)_3$-Fc	Ex(1-39)-L15-Fc
139	V201	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN	Exendina4(1-39)-GGGGS-Fc	Ex(1-39)-L5-Fc

		WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG K		
140	V207	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSD KTHTCPPCPA PEAAGGPSVF LFPPKPKDTL MISRTPEVTC VVVDVSHEDP EVKFNWYVDG VEVHNAKTKP REEQYNSTYR VVSVLTVLHQ DWLNGKEYKC KVSNKALPAP IEKTISKAKG QPREPQVYTL PPSREEMTKN QVSLTCLVKG FYPSDIAVEW ESNGQPENNY KTTPPVLDSD GSFFLYSKLT VDKSRWQQGN VFSCSVMHEA LHNHYTQKSL SLSPGKGGGG SGGGGSGGGG SDSSPLLQFG GQVRQRYLYT DDACQTEAHL EIREDGTVGG AADQSPESLL QLKALKPGVI QILGVKTSRF LCQKPDGALY GSLHFDPEAC SFRELLLEDG YNVYQSEAHG LPLHLPCNRS PHRDPASRGP ARFLPLPGLP PALPEPPGIL APQPPDVGSS DPLAMVGGSQ ARSPSYAS	Exendina4(1-39)-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L0-Fc-L15-FGF21(V103)
141	V212	HSEGTFTSDV SSYLEGQAAK EFIAWLVKDK THTCPPCPAP EAAGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE VKFNWYVDGV EVHNAKTKPR EEQYNSTYRV VSVLTVLHQD WLNGKEYKCK VSNKALPAPI EKTISKAKGQ PREPQVYTLP PSREEMTKNQ VSLTCLVKGF YPSDIAVEWE SNGQPENNYK TTPPVLDSDG SFFLYSKLTV DKSRWQQGNV FSCSVMHEAL HNHYTQKSLS LSPGKGGGGS GGGGSGGGGS DSSPLLQFGG QVRQRYLYTD DACQTEAHLE IREDGTVGGA ADQSPESLLQ LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS	GLP-1(7-35;A8S)-Fc-$(GGGGS)_3$-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L0-Fc-L15-FGF21(V103)

		FRELLLEDGY NVYQSEAHGL PLHLPCNRSP HRDPASRGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGGSQA RSPSYAS		
142	V213	HSEGTFTSDV SSYLEGQAAK EFIAWLVKDK THTCPPCPAP EAAGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE VKFNWYVDGV EVHNAKTKPR EEQYNSTYRV VSVLTVLHQD WLNGKEYKCK VSNKALPAPI EKTISKAKGQ PREPQVYTLP PSREEMTKNQ VSLTCLVKGF YPSDIAVEWE SNGQPENNYK TTPPVLDSDG SFFLYSKLTV DKSRWQQGNV FSCSVMHEAL HNHYTQKSLS LSPGKGSDSS PLLQFGGQVR QRYLYTDDAC QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQR PDGTLYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPCNRSPHRD PASRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLA MVGGSQARSP SYAS	GLP-1(7-35;A8S)-Fc-GS-FGF21(33-209;Q55C-A109T-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L0-Fc-L2-FGF21(V101)
143	V215	HSEGTFTSDV SSYLEGQAAK EFIAWLVKDK THTCPPCPAP EAAGGPSVFL FPPKPKDTLM ISRTPEVTCV VVDVSHEDPE VKFNWYVDGV EVHNAKTKPR EEQYNSTYRV VSVLTVLHQD WLNGKEYKCK VSNKALPAPI EKTISKAKGQ PREPQVYTLP PSREEMTKNQ VSLTCLVKGF YPSDIAVEWE SNGQPENNYK TTPPVLDSDG SFFLYSKLTV DKSRWQQGNV FSCSVMHEAL HNHYTQKSLS LSPGKGSDSS PLLQFGGQVR QRYLYTDDAC QTEAHLEIRE DGTVGGAADQ SPESLLQLKA LKPGVIQILG VKTSRFLCQK PDGALYGSLH FDPEACSFRE LLLEDGYNVY QSEAHGLPLH LPCNRSPHRD PASRGPARFL PLPGLPPALP EPPGILAPQP PDVGSSDPLA MVGGSQARSP SYAS	GLP-1(7-35;A8S)-Fc-GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	GLP-1(A8S)-L0-Fc-L2-FGF21(V103)

144	V217	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSTDTLLL WVLLLWVPGS TGHGEGTFTS DLSKQMEEEA VRLFIEWLKN GGPSSGAPPP SGGGGSGGGG SGGGGSDKTH TCPPCPAPEA AGGPSVFLFP PKPKDTLMIS RTPEVTCVVV DVSHEDPEVK FNWYVDGVEV HNAKTKPREE QYNSTYRVVS VLTVLHQDWL NGKEYKCKVS NKALPAPIEK TISKAKGQPR EPQVYTLPPS REEMTKNQVS LTCLVKGFYP SDIAVEWESN GQPENNYKTT PPVLDSDGSF FLYSKLTVDK SRWQQGNVFS CSVMHEALHN HYTQKSLSLS PGKGSDSSPL LQFGGQVRQR YLYTDDACQT EAHLEIREDG TVGGAADQSP ESLLQLKALK PGVIQILGVK TSRFLCQKPD GALYGSLHFD PEACSFRELL LEDGYNVYQS EAHGLPLHLP CNRSPHRDPA SRGPARFLPL PGLPPALPEP PGILAPQPPD VGSSDPLAMV GGSQARSPSY AS	Exendina4(1-39)-$(GGGGS)_3$–Exendina4(1-39)-$(GGGGS)_3$-Fc–GS-FGF21(33-209;Q55C-R105K-G148C-K150R-P158S-S195A-P199G-G202A)	Ex(1-39)-L15-Ex(1-39)-L15-Fc-L2-FGF21(V103)
145	V278	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-bis-maleimide dimer 40 kDa PEG ramificado-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76) Dimer-40KPEGb
146	V279	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-bis-maleimide dimer 20 kDa	GLP-1(A8S)-L10-FGF21(V76) Dimer-20KPEGb

		LAMVGPSQGR SPSYAS	PEG ramificado-R159Q-S195A)	
147	V280	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76)
148	V283	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGG SGGGGSGGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS	Exendina4(1-30)-S(GGGGS)$_3$GGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-R159Q-S195A)	Ex(1-30_-L20-FGF21(V76)
149	V284	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGG SGGGGSGGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLS	Exendina4(1-30)-S(GGGGS)$_3$GGGG-FGF21(33-195;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q)	Ex(1-30)-L20-FGF21(V76;CΔ14)-40KPEGb
150	V285	DLSKQMEEEA VRLFIEWLKN GGSGGGGSGG GGSGGGGSGG GGDSSPLLQF GGQVRQRYLY TDDAQETEAH LEIREDGTVG GAAHQSPESL LELKALKPGV IQILGVKTSR FLCQKPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNR SPHCDPAPQG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLAMVGPS QGRSPSYAS	Exendina4(9-30)-S(GGGGS)$_3$GGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A)	Ex(9-30)-L20-FGF21(V76)-40KPEGb

151	V289	GGQVRQRYLY TDDAQETEAH LEIREDGTVG GAAHQSPESL LELKALKPGV IQILGVKTSR FLCQKPDGAL YGSLHFDPEA CSFRELLLED GYNVYQSEAH GLPLHLPGNR SPHCDPAPQG PARFLPLPGL PPALPEPPGI LAPQPPDVGS SDPLAMVGPS QGRSPSYAS	FGF21(42-209; Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A)	FGF21(V76;NΔ9)-40KPEGb
152	V290	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(42-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A)	GLP-1(A8S)-L10-FGF21(V76;NΔ9)-40KPEGb
153	V291	GVSTSEAKFE QDSAILWYGV EFAKLHTSGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS	GLP-1(7-35;A8S scramble)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A)	GLP-1(A8S;scramble)-L10-FGF21(V76)-40KPEGb
154	V311	HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA LPEPPGILAP QPPDVGSSDP LAMVGGSQGR SPSYAS	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A-P199G)	GLP-1(A8S)-L10-FGF21(V76-P199G)-40KPEGb
155		CATTCTGAAG GCACTTTTAC TAGCGATGTT TCTAGCTACC TGGAAGGCCA GGCTGCGAAA GAATTCATCG CGTGGCTGGT TAAAGGCGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGAC	GLP-1(7-35;A8S)-GSGGGGSGGG-FGF21(33-209;Q56E-D74H-Q82E-	GLP-1(A8S)-L10-FGF21(V76-P199G)-

		TCGAGCCCGC TGCTGCAATT TGGCGGCCAG GTGCGTCAGC GTTATCTGTA TACCGATGAT GCGCAGGAAA CCGAAGCGCA TCTGGAAATT CGTGAAGATG GCACCGTGGG CGGTGCGGCG CATCAGAGCC CGGAAAGCCT GCTGGAACTG AAAGCGCTGA AACCGGGCGT GATTCAGATT CTGGGCGTGA AAACCAGCCG TTTTCTGTGC CAGAAACCGG ATGGCGCGCT GTATGGCAGC CTGCATTTTG ATCCGGAAGC GTGCAGCTTT CGTGAACTGC TGCTGGAAGA TGGCTATAAC GTGTATCAGA GCGAAGCGCA TGGCCTGCCG CTGCATCTGC CGGGCAACCG TAGCCCGCAT TGCGATCCGG CACCGCAGGG TCCGGCGCGT TTTCTGCCGC TGCCGGGTCT GCCGCCGGCA CTGCCGGAAC CGCCGGGTAT TCTGGCCCCG CAGCCGCCGG ATGTTGGTAG CAGCGATCCG CTGGCGATGG TGGGTGGTAG CCAGGGTCGT AGCCCGAGCT ATGCGAGC	R105K-K150R-R154C-40 kDa PEG ramificado-R159Q-S195A-P199G)	40KPEGb
156	V312	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG SGGGGSGGGG SGGGGSGGGG DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGGSQG RSPSYAS	Exendina4(1-30)-$S(GGGGS)_3$ GGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-K150R-R154C-40 kDa PEG ramificado-R159Q- S195A-P199G)	Ex(1-30)-L20-FGF21(V76-P199G)-40KPEGb
157		CATGGTGAGG GTACGTTTAC TTCTGATCTG TCTAAACAGA TGGAAGAAGA AGCTGTTCGC CTGTTCATTG AATGGCTGAA AAATGGTGGT TCTGGTGGTG GTGGTTCTGG CGGTGGCGGT TCTGGCGGCG GTGGTAGCGG TGGCGGCGGT GACTCGAGCC	Exendina4(1-30)-$S(GGGGS)_3$ GGGG-FGF21(33-209;Q56E-D74H-Q82E-R105K-	Ex(1-30)-L20-FGF21(V76-P199G)-40KPEGb

		CGCTGCTGCA GTTTGGCGGC CAGGTGCGTC AGCGTTATCT GTATACCGAT GATGCGCAGG AAACCGAAGC GCATCTGGAA ATTCGTGAAG ATGGCACCGT GGGCGGTGCG GCGCATCAGA GCCCGGAAAG CCTGCTGGAA CTGAAAGCGC TGAAACCGGG CGTGATTCAG ATTCTGGGCG TGAAAACCAG CCGTTTTCTG TGCCAGAAAC CGGATGGCGC GCTGTATGGC AGCCTGCATT TTGATCCGGA AGCGTGCAGC TTTCGTGAAC TGCTGCTGGA AGATGGCTAT AACGTGTATC AGAGCGAAGC GCATGGCCTG CCGCTGCATC TGCCGGGCAA CCGTAGCCCG CATTGCGATC CGGCACCGCA GGGTCCGGCG CGTTTTCTGC CGCTGCCGGG TCTGCCGCCG GCACTGCCGG AACCGCCGGG TATTCTGGCC CCGCAGCCGC CGGATGTTGG TAGCAGCGAT CCGCTGGCGA TGGTGGGTGG TAGCCAGGGT CGTAGCCCGA GCTATGCGAG C	K150R-R154C-40 kDa PEG ramificado-R159Q- S195A-P199G)	
158	V313	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYA	Exendina4(1-39)-$(GGGGS)_3$–Fc–GS-FGF21 (33-208; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103;CΔ1)

159	V314	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSY	Exendina4(1-39)-$(GGGGS)_3$–Fc–GS-FGF21 (33-207; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103;CΔ2)
160	V315	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPS	Exendina4(1-39)-$(GGGGS)_3$–Fc–GS-FGF21 (33-206; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A)	Ex(1-39)-L15-Fc-L2-FGF21(V103;CΔ3)
161	V316	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK	Exendina4(1-39)-$(GGGGS)_3$–Fc–GS-FGF21	Ex(1-39)-L15-Fc-L2-FGF21(V

		PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAS P	(33-209; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A)-P	103)+P210
162	V317	HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSDKTHTC PPCPAPEAAG GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRE EMTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG KGSDSSPLLQ FGGQVRQRYL YTDDACQTEA HLEIREDGTV GGAADQSPES LLQLKALKPG VIQILGVKTS RFLCQKPDGA LYGSLHFDPE ACSFRELLLE DGYNVYQSEA HGLPLHLPCN RSPHRDPASR GPARFLPLPG LPPALPEPPG ILAPQPPDVG SSDPLAMVGG SQARSPSYAA P	Exendina4(1-39)-$(GGGGS)_3$–Fc–GS-FGF21 (33-209; Q55C-R105K-G148C-K150R-P158S-P174L-S195A-P199G-G202A-S209A)-p	Ex(1-39)-L15-Fc-L2-FGF21(V103-S209A)+P210
163	V225	HAEGTFTSDVSSYLEGQAAKEFIAWLV KGGSGDSSPLLQFGGQVRQRYLYTDD AQQTEAHLEIREDGTVGGAADQSPESL LQLKALKPGVIQILGVKTSRFLCQRPD GALYGSLHFDPEACSFRELLLEDGYNV YQSEAHGLPLHLPGNKSPHCDPAPRGP ARFLPLPGLPPALPEPPGILAPQPPDVGS SDPLSMVGPSQGRSPSYAS	GLP-1(7-35)-SG-FGF21(33-209;R154C)	GLP-1-L2-FGF21(154C)

164	V226	HSEGTFTSDVSSYLEGQAAKEFIAWLV KGGSGDSSPLLQFGGQVRQRYLYTDD AQQTEAHLEIREDGTVGGAADQSPESL LQLKALKPGVIQILGVKTSRFLCQRPD GALYGSLHFDPEACSFRELLLEDGYNV YQSEAHGLPLHLPGNKSPHCDPAPRGP ARFLPLPGLPPALPEPPGILAPQPPDVGS SDPLSMVGPSQGRSPSYAS	GLP-1(7-35; A8S)-SG-FGF21(33-209;R154C)	GLP-1(A8S)-L2-FGF21(154C)
165	V229	HAEGTFTSDVSSYLEGQAAKEFIAWLV KGGSGDSSPLLQFGGQVRQRYLYTDD AQQTEAHLEIREDGTVGGAADQSPESL LQLKALKPGVIQILGVKTSRFLCQRPD GALYGSLHFDPEACSFRELLLEDGYNV YQSEAHGLPLHLPGNKSPHCDPAPRGP ARFLPLPGLPPALPEPPGILAPQPPDVGS SDPLSMVGPSQGRSPSYAS	GLP-1(7-35)-SG-FGF21(33-209;R154C-40 kDa linear PEG)	GLP-1-L2-FGF21(154C)-40KPEGl
166	V230	HSEGTFTSDVSSYLEGQAAKEFIAWLV KGGSGDSSPLLQFGGQVRQRYLYTDD AQQTEAHLEIREDGTVGGAADQSPESL LQLKALKPGVIQILGVKTSRFLCQRPD GALYGSLHFDPEACSFRELLLEDGYNV YQSEAHGLPLHLPGNKSPHCDPAPRGP ARFLPLPGLPPALPEPPGILAPQPPDVGS SDPLSMVGPSQGRSPSYAS	GLP-1(7-35; A8S)-SG-FGF21(33-209;R154C-40 kDa linear PEG)	GLP-1(A8S)-L2-FGF21(154C)-40KPEGl
167	V231	HAEGTFTSDVSSYLEGQAAKEFIAWLV KGGSGGGGSGGGDSSPLLQFGGQVRQ RYLYTDDAQQTEAHLEIREDGTVGGA ADQSPESLLQLKALKPGVIQILGVKTSR FLCQRPDGALYGSLHFDPEACSFRELL LEDGYNVYQSEAHGLPLHLPGNKSPH CDPAPRGPARFLPLPGLPPALPEPPGIL APQPPDVGSSDPLSMVGPSQGRSPSYA S	GLP-1(7-35)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa linear PEG)	GLP-1-L8-FGF21(154C)-40KPEGl
168	V232	HSEGTFTSDVSSYLEGQAAKEFIAWLV KGGSGGGGSGGGDSSPLLQFGGQVRQ RYLYTDDAQQTEAHLEIREDGTVGGA ADQSPESLLQLKALKPGVIQILGVKTSR FLCQRPDGALYGSLHFDPEACSFRELL LEDGYNVYQSEAHGLPLHLPGNKSPH CDPAPRGPARFLPLPGLPPALPEPPGIL APQPPDVGSSDPLSMVGPSQGRSPSYA S	GLP-1(7-35; A8S)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa linear PEG)	GLP-1(A8S)-L8-FGF21(154C)-40KPEGl
169	V237	HAEGTFTSDVSSYLEGQAAKEFIAWLV KGGSGDSSPLLQFGGQVRQRYLYTDD AQQTEAHLEIREDGTVGGAADQSPESL LQLKALKPGVIQILGVKTSRFLCQRPD GALYGSLHFDPEACSFRELLLEDGYNV YQSEAHGLPLHLPGNKSPHCDPAPRGP ARFLPLPGLPPALPEPPGILAPQPPDVGS SDPLSMVGPSQGRSPSYAS	GLP-1(7-35)-SG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1-L2-FGF21(154C)-40KPEGb

170	V238	GDSSPLLQFGGQVRQRYLYTDDAQQT EAHLEIREDGTVGGAADQSPESLLQLK ALKPGVIQILGVKTSRFLCQRPDGALY GSLHFDPEACSFRELLLEDGYNVYQSE AHGLPLHLPGNKSPHCDPAPRGPARFL PLPGLPPALPEPPGILAPQPPDVGSSDPL SMVGPSQGRSPSYAS	G-FGF21(33-209;R154C-40 kDa PEG ramificado)	FGF21(R154C)-40KPEGb
171	V253	HSEGTFTSDVSSYLEGQAAKEFIAWLV KGGSGGGGSGGGDSSPLLQFGGQVRQ RYLYTDDAQQTEAHLEIREDGTVGGA ADQSPESLLQLKALKPGVIQILGVKTSR FLCQRPDGALYGSLHFDPEACSFRELL LEDGYNVYQSEAHGLPLHLPGNKSPH CDPAPRGPARFLPLPGLPPALPEPPGIL APQPPDVGSSDPLS	GLP-1(7-35; A8S)-SGGGGSGGG-FGF21(33-194;R154C-40 kDa PEG ramificado)	GLP-1(A8S)-L8-FGF21-(CΔ14;154C)-40KPEGb
172	V258	HGEGTFTSDVSSYLEGQAAKEFIAWLV KGGSGGGGSGGGDSSPLLQFGGQVRQ RYLYTDDAQQTEAHLEIREDGTVGGA ADQSPESLLQLKALKPGVIQILGVKTSR FLCQRPDGALYGSLHFDPEACSFRELL LEDGYNVYQSEAHGLPLHLPGNKSPH CDPAPRGPARFLPLPGLPPALPEPPGIL APQPPDVGSSDPLSMVGPSQGRSPSYA S	GLP-1(7-35; A8G)-SGGGGSGGG-FGF21(33-209;R154C-40 kDa PEG ramificado)	GLP-1(A8G)-L8-FGF21(154C)-40KPEGb

[00051] Los variantes o mutantes utilizados en las proteínas de la invención, por ejemplo variantes de FGF21 tipo silvestre, GLP-1, y/o Exendina-4 presentan al menos un aminoácido sustituido, agregado, y/o retirado con relación a la proteína tipo silvestre. Además, los variantes pueden incluir truncamientos N- y/o C-terminales en relación con las proteínas tipo silvestre. En términos generales, un variante posee alguna propiedad modificada, estructural o funcional de la proteína tipo silvestre. Por ejemplo, el variante puede tener estabilidad física aumentada o mejorada en soluciones concentradas (por ejemplo, agregación mediada menos hidrofóbica), estabilidad plasmática aumentada o mejorada cuando se incuba con el plasma de la sangre o riesgo reducido de inmunogenicidad o bioactividad mejorada o aumentada mientras se mantiene un perfil de bioactividad favorable.

[00052] Sustituciones y modificaciones aminoácidas aceptables que constituyen diferencias entre las porciones de las proteínas de la invención y sus proteínas comparadoras tipo silvestre incluyen, pero no se limitan a, una o más sustituciones aminoácidas, incluyendo las sustituciones con pirrolisina, pirrolina-carboxi-lisina (Pcl) y

análogos aminoácidos de origen no natural, y truncamientos. Por lo tanto, las proteínas de la invención (ej., las proteínas de fusión de la invención) incluyen, pero no se limitan a, mutantes dirigidos a sitio, polipéptidos truncados, mutantes resistentes a proteolisis, mutantes reductores de agregación, mutantes de combinación, y proteínas de fusión, como se describen en la presente.

[00053] El experto en el estado de la técnica de expresión de proteínas reconocerá que la secuencia de metionina o metionina-arginina puede ser introducida en el extremo N de cualquiera de las proteínas de la invención, para expresión en E. coli, y se contemplan dentro del contexto de este invento.

[00054] Un experto en la técnica de expresión de proteínas reconocerá que las etiquetas adicionales o dominios de fusión para los propósitos de la modulación de los niveles de expresión, purificación, o estabilización pueden ser introducidos al extremo N de cualquiera de las proteínas de la invención, con o sin un péptido adicional para dirigir la digestión por una proteasa específica para la eliminación posterior de esa etiqueta o dominio de fusión, para la expresión en cualquier célula huésped, y se contemplan dentro del contexto de esta invención.

[00055] Un experto en el estado de la técnica de la expresión de proteínas reconocerá que los péptidos líder que dirigen la proteína expresada al periplasma o el espacio extracelular se pueden introducir en el extremo N-terminal de cualquiera de las proteínas de la invención, para la expresión en E. coli u otros huéspedes bacterianos, y se contemplan dentro del contexto de esta invención.

[00056] Un experto en el estado de la técnica de la expresión de proteínas reconocerá que los péptidos líder que dirigen la proteína expresada a ER, vesículas secretoras o el espacio extracelular se pueden introducir en el extremo N-terminal de cualquiera de las proteínas de la invención, para la expresión en células hospederas eucariotas, y se contemplan dentro del contexto de esta invención.

[00057] Las proteínas de la invención pueden poseer una mayor compatibilidad con conservantes farmacéuticos (por ejemplo, m-cresol, fenol, alcohol bencílico), lo que permite la preparación de una formulación farmacéutica conservada que mantiene las propiedades físico-químicas y actividad biológica de la proteína durante el almacenamiento. En consecuencia, los variantes con estabilidad farmacéutica aumentada con relación al tipo silvestre, tienen estabilidad física mejorada en soluciones concentradas, bajo condiciones fisiológicas y conservadas de la formulación farmacéutica, mientras se mantiene la potencia biológica. A modo de ejemplo no

limitante, las proteínas de la invención pueden ser más resistentes a la proteólisis y la degradación enzimática, pueden tener una mayor estabilidad, y pueden ser menos propensas a agregarse, que sus homólogos de tipo silvestre o la correspondiente secuencia nativa. Como se utilizan en la presente, estos términos no son mutuamente excluyentes o limitantes, siendo muy posible que un variante determinado tenga una o más propiedades modificadas de la proteína tipo silvestre.

[00058] La invención también incluye moléculas de ácidos nucleicos que codifican las proteínas de la invención, que comprenden, por ejemplo, una secuencia aminoácida de FGF21 que es al menos 95% idéntica a la secuencia aminoácida de SEQ ID NO:3, pero en donde los residuos específicos que confieren una propiedad deseable a la variante de proteina FGF21, por ejemplo potencia mejorada a receptores FGF21, resistencia a proteólisis, vida media aumentada ó propiedades reductoras de agregación y combinaciones de las mismas, no se han modificado adicionalmente. En otras palabras, con la excepción de los residuos en la secuencia mutante de FGF21 que han sido modificados con el fin de conferir resistencia a la proteólisis-, reducción de la agregación, u otras propiedades, alrededor del 5% (alternativamente 4%, alternativamente 3%, alternativamente 2%, alternativamente 1%) de todos los demás residuos aminoácidos en la secuencia mutante FGF21 pueden ser modificados. Tales mutantes de FGF21 poseen al menos una actividad del polipéptido FGF21 de tipo silvestre.

[00059] De manera similar, la invención también comprende moléculas de ácido nucleico que codifican las porciones de GLP-1 y Exendina-4 de la molécula, cuyas secuencias de aminoácidos son al menos aproximadamente 85%, idénticas, y más preferiblemente al menos aproximadamente 90 a 95% idénticas, a la secuencia de aminoácidos de SEQ ID NO: 30 y 7, respectivamente, pero en donde los residuos específicos que confieren una propiedad deseable a la variante de la proteína de función dual, por ejemplo, resistencia a proteolisis, vida media aumentada o propiedades reductoras de agregación y combinaciones de los mismas no se han modificado adicionalmente .

[00060] El invento también comprende una molécula de ácido nucleico que comprende una secuencia nucleótida que es al menos 85% idéntica, y más preferiblemente, al menos 90 a 95% idéntica a la secuencia nucleótida de SEQ ID NO: 2 o SEQ ID NO:4, SEC ID NO:6, y la secuencia de ADNc que codifica Exendina-4 tipo silvestre, pero en donde los nucleótidos que codifican los residuos aminoácidos que

confieren a la proteína la resistencia a proteólisis, reducción de agregación u otras propiedades, no se han modificado adicionalmente. En otras palabras, con la excepción de los nucleótidos que codifican los residuos en las secuencias mutantes de FGF21, GLP-1, o Exendina-4 que han sido modificados con el fin de conferir las propiedades de resistencia a la proteólisis, reducción de la agregación, u otras propiedades, alrededor de 15%, y más preferiblemente entre 10 y 5% aproximadamente de todos los nucleótidos en la secuencia mutante, pueden ser modificados. Tales moléculas de ácido nucleico codifican proteínas que poseen al menos una actividad de sus contrapartes tipo silvestre.

[00061] Se proporcionan en la presente métodos para generar las proteínas de la invención, en donde tales métodos implican la modificación específica a sitio y modificación no específica a sitio de las versiones tipo silvestre de las proteínas (ej., la proteína tipo silvestre FGF21 descrita en la presente), por ejemplo, truncamientos de las proteínas tipo silvestre, y la incorporación específica a sitio de los aminoácidos en las posiciones de interés dentro de las proteínas tipo silvestre. Dichas modificaciones mejoran las propiedades biológicas de las proteínas de la invención en relación con las proteínas tipo silvestre, así como, en algunos casos, sirven como puntos de fijación para, por ejemplo, las etiquetas y los agentes de extensión de vida media proteica, y para fines de fijar dichos variantes a la superficie de un soporte sólido. Las realizaciones afines del invento son métodos para producir células capaces de producir dichas proteínas de función dual de la invención, y de la producción de vectores que contienen el ADN que codifica dichos variantes.

[00062] En ciertas realizaciones, tales modificaciones, por ejemplo modificaciones específicas a sitio, se utilizan para fijar conjugados, ej., grupos PEG a proteínas, polipéptidos, y/o péptidos de la invención, para fines de por ejemplos extender la vida media o de otra manera mejorar las propiedades biológicas de dichas proteínas, polipéptidos, y/o péptidos. Dichas técnicas se describen adicionalmente en la presente.

[00063] En otras realizaciones, tales modificaciones, ej., las modificaciones específicas a sitios se utilizan para fijar otros polímeros, moléculas pequeñas y secuencias de proteínas recombinantes que extienden la vida media de la proteína de la invención. Tal realización incluye la fijación de ácidos grados o compuestos de fijación de albúmina a proteínas, polipéptidos, y/o péptidos. En otras realizaciones, las modificaciones se hacen en un tipo de aminoácido particular y pueden fijarse en uno o más sitios sobre la proteína.

[00064] En otras realizaciones, tales modificaciones, por ejemplo, modificaciones específicas a sitio se utilizan como elementos de fijación para la producción de multímeros de tipo silvestre y/o variantes, por ejemplo, dímeros (homodímeros o heterodímeros) o trímeros o tetradímeros.. Estas moléculas de proteínas multiméricas, pueden además tener grupos tales como PEG, azúcares y/o conjugados PEG -colesterol fijados o a ser fusionados ya sea de manera amino-terminal o carboxi-terminal con otras proteínas, tales como Fc, HSA, etc.

[00065] En otras realizaciones, tales modificaciones específicas a sitio se utilizan para producir proteínas, polipéptidos y/o péptidos en donde la posición de la pirrolisina o análogo de pirrolisina específicamente incorporada a sitio, pirrolina-carboxi-lisina, o análogo de pirrolisina o aminoácidos de origen no natural (para-acetil-Phe, para-azido-Phe) permite la orientación controlada y fijación de tales proteínas, polipéptidos y/o péptidos en una superficie de un soporte sólido o tener grupos tales como PEG, azúcares y/o conjugados PEG-colesterol fijados.

[00066] En otras realizaciones, tales modificaciones específicas a sitio se utilizan para entrecruzar específicamente a sitio proteínas, polipéptidos y/o péptidos formando hetero-oligómeros, incluyendo pero sin limitar a heterodímeros y heterotrímeros. En otras realizaciones, tales modificaciones específicas a sitio se utilizan para entrecruzar específicamente a sitio proteínas, polipéptidos y/o péptidos formando conjugados proteína-proteína, conjugados proteína-polipéptido, conjugados proteína-péptido, conjugados polipéptido-polipéptido, conjugados polipéptido-péptido o conjugados péptido-péptido. En otras realizaciones, una modificación específica a sitio puede incluir una punto de ramificación para permitir que más de un tipo de molécula sea fijado en un único sitio de una proteína, polipéptido o péptido.

[00067] En otras realizaciones, las modificaciones listadas en la presente pueden realizarse en una forma no específica a sitio y resultar en conjugados proteína-proteína, conjugados proteína-polipéptido, conjugados proteína-péptido, conjugados polipéptido-polipéptido, conjugados polipéptido-péptido o conjugados péptido-péptido de la invención.

Definiciones

[00068] Varias definiciones se utilizan en este documento. La mayoría de las palabras tienen el significado que se atribuye a esas palabras por un experto en el estado de la técnica. Las palabras definidas específicamente a continuación o en otra parte de este

documento tienen el significado proporcionado en el contexto del presente invento en su conjunto y como normalmente se entiende por los expertos en el estado de la técnica.

[00069] Como se utiliza en la presente, el término "FGF21" se refiere a un miembro de la familia de proteínas del factor de crecimiento de fibroblastos (FGF). Una secuencia aminoácida de FGF21 (No. de Acceso GenBank NP_061986.1) se establece como SEQ ID NO: 1, la correspondiente secuencia polinucleótida la cual se establece como SEQ ID NO: 2 (Número NCBI de referencia de secuencia NM_019113.2). "variante FGF21," "mutante FGF21", y términos similares describen la versión modificada de la proteína FGF21, por ejemplo, con residuos aminoácidos constituyentes suprimidos, agregados, modificados, o sustituidos.

[00070] Como se utiliza en la presente, el término "receptor FGF21 " se refiere a un receptor para FGF21 (Kharitonenkov,A. (2008) Journal of Cellular Physiology 215:1-7; Kurosu,Het al. (2007) JBC 282:26687-26695; Ogawa, Yet al. (2007) PNAS 104:7432-7437).

[00071] El término "Polipéptido FGF21" se refiere a un polipéptido de origen natural expresado en los seres humanos. Para fines de esta divulgación, el término "polipéptido FGF21" se puede utilizar indistintamente para referirse a cualquier polipéptido FGF21 de longitud completa, por ejemplo, SEQ ID NO: 1, que consta de 209 residuos aminoácidos y que es codificado por la secuencia nucleótida SEQ ID NO: 2; cualquier forma madura del polipéptido, que consiste en 181 residuos aminoácidos, y en la que los 28 residuos aminoácidos en el extremo amino-terminal del polipéptido FGF21 de longitud completa (es decir, que constituyen el péptido señal) han sido eliminados.

[00072] "Variante 76," o "V76," como se utiliza en la presente, es una variante de proteína de FGF21, que presenta un PEG ramificado de 40 kDa ligado a través de Cys154, y mutaciones de ocho puntos en relación con la proteína tipo silvestre de 177 aminoácidos. La síntesis del variante se describe en detalla en la presente.

[00073] "Variante 101," o "V101," tal como se usa en el presente documento, es una variante de la proteína FGF21, que ofrece un puente de disulfuro modificado, y ocho mutaciones puntuales en relación con la proteína tipo silvestre de 177 aminoácidos, expresada como una fusión con el dominio Fc de IgG1 humana con un enlazador GS. La síntesis del variante se describe en detalla en la presente.

[00074] "Variante 103," o "V103," tal como se usa en el presente documento, es una variante de la proteína FGF21, que ofrece un puente de disulfuro modificado, y ocho mutaciones puntuales en relación con la proteína tipo silvestre de 177 aminoácidos,

expresada como una fusión con el dominio Fc de IgG1 humana con un enlazador GS. La síntesis del variante se describe en detalla en la presente.

[00075] "Variante 188," o "V188," tal como se usa en el presente documento, es una variante de la proteína FGF21, que ofrece un puente de disulfuro modificado, y ocho mutaciones puntuales en relación con la proteína tipo silvestre de 177 aminoácidos, expresada como una fusión con el dominio Fc de IgG1 humana con un enlazador $(SGGGG)_3$. La síntesis del variante se describe en detalla en la presente.

[00076] " Agonistas duales de GLP-1-FGF21-PEG ", "proteínas de función dual", "proteínas de fusión de función dual", "proteínas de actividad dual", "productos de fusión", "agonista del receptor de GLP-1 dual y agonista del receptor de FGF21 dual", "agonista del receptor FGF21 dual y proteínas de agonista del receptor GLP-1 de la invención", "proteínas de fusión de GLP-1-FGF21", "proteínas de fusión de la invención", " fusiones de la invención ", y términos similares definen fusiones proteicas o polipépticas que comprenden al menos una versión variante, mutante o truncada proteica o polipeptídica de FGF21, fusionada o ligada a otro regulador metabólico, tal como GLP-1 o Exendina-4. Comprenden una única molécula con actividad dual o función dual vis a vis los receptores de sus componentes respectivos, es decir, muestran la capacidad de agonizar el receptor de FGF21 y el receptor de GLP-1. Las secuencias constituyentes de dichas proteínas de fusión pueden comprender modificaciones, truncamientos, otras variantes de homólogos proteicos o polipeptídicos de origen natural (es decir, de tipo silvestre), y pueden emplear cualquier número de otras diversas modificaciones, por ejemplo, los grupos PEG para la extensión de vida media.

[00077] Una realización particularmente preferida de la proteína de fusión GLP-1-FGF21-PEG de la invención, incorpora V76 (tal como se define en la presente) como la variante de FGF21. Dicha realización preferida también se conoce en la presente como "GLP-1 (A8S)-FGF21-PEG" y dispone de una sustitución de alanina a serina en la posición 8 en relación con la secuencia GLP-1 tipo silvestre (SEQ ID NO: 5129) y una sustitución de arginina a cisteína en la posición 154 en relación con la secuencia FGF21 tipo silvestre (SEQ ID: 1). La secuencia de GLP-1 (A8S)-FGF21(V76)-PEG es la siguiente (SEQ ID NO: 9).

1 HSEGTFTSDV SSYLEGQAAK EFIAWLVKGG SGGGGSGGGD SSPLLQFGGQ VRQRYLYTDD

61 AQETEAHLEI REDGTVGGAA HQSPESLLEL KALKPGVIQI LGVKTSRFLC QKPDGALYGS

121 LHFDPEACSF RELLLEDGYN VYQSEAHGLP LHLPGNRSPH CDPAPQGPAR FLPLPGLPPA

181 LPEPPGILAP QPPDVGSSDP LAMVGPSQGR SPSYAS

[00078] El término "molécula de ácido nucleico aislada" se refiere a una molécula de ácido nucleico del presente invento que (1) se ha separado de por lo menos un 50 por ciento de proteínas, lípidos, hidratos de carbono, u otros materiales con los que se encuentra de forma natural cuando el ácido nucleico es aislado desde las células fuente, (2) no está enlazada a la totalidad o parte de un polinucleótido al cual la "molécula de ácido nucleico aislada" está enlazada en la naturaleza, (3) está enlazada operativamente a un polinucleótido que no está enlazado en la naturaleza, o (4) no se da en la naturaleza como parte de una secuencia polinucleótida más grande. Preferiblemente, la molécula de ácido nucleico aislada de la presente invención está sustancialmente libre de cualquier otra molécula de ácida nucleico contaminante u otros contaminantes que se encuentran en su entorno natural que pudieran interferir con su uso en la producción de polipéptidos o su uso terapéutico, diagnóstico, profiláctico o de investigación.

[00079] El término "vector" se utiliza para referirse a cualquier molécula (por ejemplo, ácido nucleico, plásmido o virus) que se utiliza para transferir la información de codificación a una célula huésped.

[00080] El término "vector de expresión" se refiere a un vector que es adecuado para la transformación de una célula huésped y contiene secuencias de ácidos nucleicos que dirigen y/o controlan la expresión de secuencias heterólogas de ácidos nucleicos. La expresión incluye, pero no se limita a, los procesos como la transcripción, traducción y empalme de ARN, siempre y cuando haya intrones.

[00081] El término "enlazado operativamente" se utiliza aquí para referirse a un arreglo de las secuencias de flanqueo en donde las secuencias de flanqueo así descritas se configuran o montan para llevar a cabo su función habitual. Elementos de proteínas de fusión pueden enlazarse operativamente entre si para permitir que la proteína de fusión funcione como si ésta fuera una proteína endógena natural, y/o para combinar elementos diferentes de dichas proteínas de fusión en una forma sinérgica.

[00082] En un nivel nucleótido, una secuencia de flanqueo operativamente enlazada a una secuencia de codificación puede ser capaz de realizar la replicación, transcripción

y/o traducción de la secuencia codificadora. Por ejemplo, una secuencia codificadora está enlazada operativamente a un promotor cuando el promotor es capaz de dirigir la transcripción de la secuencia codificadora. Una secuencia de flanqueo no tiene por qué ser contigua a la secuencia codificadora, siempre y cuando funcione correctamente. Por lo tanto, por ejemplo, secuencias sin traducir que intervienen pero que si fueron transcritas pueden estar presentes entre una secuencia promotora y la secuencia codificadora y la secuencia promotora puede seguir siendo considerada "operativamente enlazada" a la secuencia codificadora.

[00083] El término "célula huésped" se utiliza para referirse a una célula que se ha transformado, o es capaz de ser transformada con una secuencia de ácidos nucleicos y luego de expresar un gen de interés seleccionada. El término incluye la progenie de la célula precursora, bien sea que la progenie sea idéntica en la morfología o en la composición genética al precursor original, siempre y cuando el gen seleccionado esté presente.

[00084] El término "aminoácido", como se utilizan en la presente, se refiere a aminoácidos de origen natural, aminoácidos no naturales, análogos de aminoácidos y miméticos de aminoácidos que funcionan de manera similar a los aminoácidos naturales, todos en sus estereoisómeros D y L, si su estructura permite tales formas estereoisoméricas. Los aminoácidos se denominan en adelante ya sea por su nombre, sus símbolos de tres letras comúnmente conocidos o por los símbolos de una letra recomendados por la Comisión de Nomenclatura de Bioquímicos IUPAC-IUB.

[00085] El término "natural" cuando se utiliza en relación con los materiales biológicos, tales como moléculas de ácido nucleico, polipéptidos, células huésped, y similares, se refiere a los materiales que se encuentran en la naturaleza y no son manipulados por el hombre. Del mismo modo, "no natural" como se utiliza en la presente se refiere a un material que no se encuentra en la naturaleza o que ha sido modificado estructuralmente o sintetizado por el hombre. Cuando se utiliza en relación con los nucleótidos, el término "natural" se refiere a las bases adenina (A), citosina (C), guanina (G), timina (T) y uracilo (U). Cuando se utiliza en relación con los aminoácidos, el término "natural" se refiere a los 20 amino ácidos convencionales (es decir, la alanina (A), cisteína (C), ácido aspártico (D), ácido glutámico (E), fenilalanina (F) , glicina (G), histidina (H), isoleucina (I), lisina (K), leucina (L), metionina (M), asparagina (N), prolina (P), glutamina (Q), arginina (R) , serina (S), treonina (T), valina

(V), triptófano (W) y tirosina (Y)), así como selenocisteína, pirrolisina (PYL), y pirrolina-carboxi-lisina (Pcl, o Z).

[00086] La pirrolisina (Pyl) es un aminoácido que se encuentra naturalmente dentro de metilamina metiltransferasas de arqueobacterias (archaea) metanogénicas de la familia Methanosarcina. La pirrolisina es un análogo de lisina incorporado de manera co-traduccional en codones UAG dentro del mismo marco de lectura correspondiente, y se considera el aminoácido natural número 22.

[00087] Como se ha descrito al menos en la publicación de patente PCT WO2010/48582 (solicitante IRM, LLC), los intentos de biosintetizar la pirrolisina (PYL) en E. coli dieron lugar a la formación de una "pirrolisina desmetilada," en adelante, pirrolina-carboxi-lisina, o Pcl. "Pc", como se utiliza en la presente, se refiere tanto a PCL-A como a PCL-B.

[00088] Los términos "aminoácido no natural" y “aminoácido innatural” como se utilizan en la presente para representar las estructuras de aminoácidos que no se pueden generar biosintéticamente en cualquier organismo usando genes sin modificar o modificados de cualquier organismo, ya sea el mismo o diferente. Los términos se refieren a un residuo de aminoácido que no está presente en la secuencia proteica de FGF21 de origen natural (tipo silvestre) o las secuencias de de la presente invento. Estos incluyen, pero no se limitan a, aminoácidos modificados y/o análogos de aminoácidos que no son uno de los 20 aminoácidos naturales, selenocisteína, pirrolisina (PYL), o pirrolina-carboxi-lisina (PCL, por ejemplo, como se describe en la publicación de patente PCT WO2010/48582). Tales residuos aminoácidos no naturales pueden introducirse por la sustitución de aminoácidos de origen natural y/o mediante la inserción de aminoácidos no naturales en la secuencia proteica FGF21 natural (tipo silvestre) o las secuencias de la invención. El residuo aminoácido no natural también puede incorporarse de tal manera que una funcionalidad deseada se imparte a la molécula de FGF21, por ejemplo, la capacidad para enlazar una fracción funcional (por ejemplo, PEG). Cuando se utiliza en conexión con aminoácidos, el símbolo “U” significa “aminoácido no natural” y aminoácido innatural,” como se utiliza en la presente.

[00089] Además, se entiende que tales "aminoácidos no naturales" típicamente requieren un ARNt modificado y una sintetasa de ARNt modificada (RS) para su incorporación en una proteína. Estos pares ARNt/RS ortogonales "seleccionados" pares son generados por un proceso de selección desarrollado por Schultz o por mutación aleatoria o específica. A manera de ejemplo, la pirrolina-carboxi-lisina es un

"aminoácido natural" generado biosintéticamente por genes transferidos desde un organismo en las células huésped e incorporado en proteínas utilizando genes de sintetasa de ARNt natural y ARNt, mientras p-aminofenillalanina (Ver, Generation of a bacterium with a 21 amino acid genetic code, Mehl RA, Anderson JC, Santoro SW, Wang L, Martin AB, King DS, Horn DM, Schultz PG. J Am Chem Soc. 2003 Jan 29;125(4):935-9) es un "aminoácido no natural" puesto que, aaunque se genera biosintéticamente, se incorpora en las proteínas por un par tRNA/tRNA sintetasa ortogonal "seleccionado" (es decir, de origen no natural).

[00090] Los aminoácidos codificados modificados incluyen, pero no se limitan a, hidroxiprolina, γ-carboxiglutamato, O-fosfoserina, ácido azetidinacarboxílico, ácido 2-aminoadípico, ácido 3-aminoadípico, beta-alanina, ácido aminopropiónico, ácido 2-amino-butírico, ácido 4-aminobutírico, ácido 6-aminocaproico, ácido 2-aminoheptanoico, ácido 2-aminoisobutírico, ácido 3-aminoisobutírico ácido 2-aminopimélico, butilglicina terciaria, ácido 2,4-diaminoisobutírico, desmosina, ácido 2,2'-diaminopimélico, ácido 2,3-diaminopropriónico, N-etilglicina, N-metilglicina, N-etilasparagina, homoprolina, hidroxilisina, alo-hidroxilisina, 3-hidroxiprolina, 4-hidroxiprolina, isodesmosina, alo-isoleucina, N-metilalanina, N-metilglicina, N-metilisoleucina, N-metilpentilglicina, N-metilvalina, naftalanina, norvalina, norleucina, ornitina, pentilglicina, ácido pipecólico y tioprolina. El término "aminoácido" también incluye aminoácidos naturales que son metabolitos de ciertos organismos, pero no están codificados por el código genético para su incorporación en las proteínas. Tales aminoácidos incluyen, pero no se limitan a, ornitina, D-ornitina y D-arginina.

[00091] El término "análogo aminoácido," como se utiliza en la presente, se refiere a compuestos que tienen la misma estructura básica que un aminoácido natural, a manera de ejemplo solamente, un carbono-α que está unido a un hidrógeno, un grupo carboxilo, un grupo amino, y un grupo R Análogos aminoácidos incluyen los aminoácidos naturales y no naturales que son químicamente bloqueados, de manera reversible o irriversible, o su grupo carboxi C-terminal, su grupo amino N-terminal y/o sus grupos funcionales de cadena lateral son químicamente modificados. Tales análogos incluyen, pero no se limitan a, sulfóxido de metionina, metionina sulfona, S-(carboximetil)-cisteína, S-(carboximetil)-cisteína sulfóxido, S-(carboximetil)-cisteína sulfona, ácido aspártico-(beta-éster metílico), N-etilglicina, carboxamida de alanina, homoserina, norleucina, y metionina metil-sulfonio.

[00092] El término "miméticos de aminoácidos," como se utiliza en la presente, se refiere a los compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero funciona de manera similar que un aminoácido natural.

[00093] El término "variante biológicamente activo" se refiere a cualquier variante polipéptido usado en las proteínas de función dual de la invención, por ejemplo, como una proteína constituyente de las fusiones, que posee una actividad de su proteína tipo silvestre (por ejemplo, de origen natural) o contraparte polipéptida, tal como la capacidad de modular niveles de glucosa en sangre, HbA1c, niveles de insulina, triglicéridos, o colesterol; aumentar la función pancreática; reducir los niveles de lípidos en el hígado; reducir el peso corporal, y mejorar la tolerancia a la glucosa, el gasto de energía, o sensibilidad a la insulina, independientemente del tipo o la cantidad de modificaciones que se han introducido en la variante polipeptídica. Los variantes polipeptídicos que poseen un nivel algo disminuido de la actividad relativa a sus versiones tipo silvestre no obstante, pueden considerarse como variantes polipeptídicos biológicamente activos. Un ejemplo no limitante representativo de un variante polipeptídico biológicamente activo de la invención es una variante de FGF21, que se modifica después, y posee propiedades biológicas similares o mejoradas con respecto a FGF21 de tipo silvestre.

[00094] Los términos "cantidad efectiva" y "cantidad terapéuticamente efectiva" se refieren cada uno a la cantidad de una proteína de fusión de la invención utilizada para para apoyar un nivel observable de una o más actividades biológicas del polipéptido tipo silvestre o contrapartes de proteínas, tales como la capacidad para reducir los niveles de glucosa en sangre, insulina, triglicéridos o colesterol, reducir los triglicéridos en el hígado o los niveles de lípidos, reducir el peso corporal, o mejorar la tolerancia a la glucosa, el gasto energético, o la sensibilidad a la insulina. Por ejemplo, una "cantidad terapéuticamente efectiva" se administra a un paciente que exhibe, padece, o está propenso a sufrir de trastornos asociados con FGF21 o trastornos asociados con GLP-1 (tales como diabetes mellitus tipo 1 o tipo 2, obesidad o síndrome metabólico), es una cantidad tal que induce, mejora o de lo contrario provoca una mejoría de los síntomas patológicos, progresión de la enfermedad, condiciones fisiológicas asociadas con los trastornos antes mencionados o resistencia a sucumbir a los mismos. Para efectos del presente invento, un "sujeto" o "paciente" es preferentemente un ser humano, pero también puede ser un animal, más específicamente, un animal de compañía (por ejemplo,

perros, gatos, etc), animales de granja (por ejemplo, vacas, ovejas, cerdos, caballos, y similares) y animales de laboratorio (por ejemplo, ratas, ratones, cobayas, y similares).

[00095] El término "vehículo farmacéuticamente aceptable" o "vehículo fisiológicamente aceptable" como se utiliza en la presente se refiere a uno o más materiales de formulación adecuados para lograr o mejorar la liberación de una proteína de fusión de la invención.

[00096] El término "antígeno" se refiere a una molécula o una porción de una molécula que es capaz de ser fijada por un anticuerpo, y, además, que es capaz de ser utilizada en un animal para producir anticuerpos que son capaces de unirse a un epítopo de ese antígeno. Un antígeno puede tener uno o más epítopos.

[00097] El término "Fc nativo" se refiere a una molécula o secuencia que comprende la secuencia de un fragmento no fijador de antígenos resultante de la digestión del anticuerpo entero o producido por otros medios, ya sea en forma monomérica o multimérica, y puede contener la región de bisagra. La fuente original de la inmunoglobulina del Fc nativo es preferentemente de origen humano y puede ser cualquiera de las inmunoglobulinas, aunque IgG1 e IgG2 son las preferidas. Las moléculas Fc nativas están formadas por polipéptidos monoméricos que pueden conectarse a formas diméricas o multiméricas mediante asociación covalente (es decir, enlaces de disulfuro) y asociación no covalente. El número de enlaces de disulfuro intermoleculares entre las subunidades monoméricas de moléculas Fc nativas oscila de 1 a 4 dependiendo de la clase (por ejemplo, IgG, IgA e IgE) o la subclase (por ejemplo, IgG1, IgG2, IgG3, IgA1 y IgGA2). Un ejemplo de un Fc nativo es un dímero unido por puentes de disulfuro resultante de la digestión de papaína de una IgG (véase Ellison, 1982, Nucleic Acids Res. 10: 4071-9). El término "Fc nativo" como se utiliza en la presente es genérico a las formas monomérica, dimérica, y multimérica. El término “variante Fc” se refiere a una molécula o secuencia que es modificada desde un Fc nativo pero aún comprende un sitio de unión para el receptor de salvamento, FcRn (receptor Fc neonatal). Las Publicaciones Internacionales No. WO 97/34631 y WO 96/32478 describen variantes Fc de ejemplo, así como la interacción con el receptor de salvamento, y se incorporan en la presente como referencia. Por lo tanto, el término "variante Fc" puede comprender una molécula o secuencia que se humaniza a partir de un Fc nativo no humano. Además, un Fc nativo comprende regiones que se pueden quitar, ya que proporcionan características estructurales o actividad biológica que no son necesarias para las moléculas de fusión de las proteínas de fusión de la invención. Por lo tanto, el

término "variante Fc" comprende una molécula o secuencia que carece de uno o más sitios Fc nativos o residuos, o en los que uno o más sitios Fc o residuos se han modificado, que afectan o están implicados en: (1) formación de enlaces de disulfuro, (2) incompatibilidad con una célula huésped seleccionada, (3) heterogeneidad N-terminal en la expresión en una célula huésped seleccionada, (4) glicosilación (5), interacción con el complemento, (6) la unión a un receptor Fc diferente de un receptor de rescate, o (7) citotoxicidad celular dependiente de anticuerpos (ADCC). Los variantes Fc se describen más adelante con más detalle.

[00098] El término "dominio Fc" comprende variantes y secuencias Fc y Fc nativas como se definió anteriormente. Al igual que con los variantes Fc y moléculas Fc nativas, el término "dominio Fc" incluye moléculas en forma monomérica o multimérica, ya sea digeridas desde el anticuerpo completo o producidas por otros medios. En algunas realizaciones del presente invento, un dominio Fc puede fusionarse con FGF21 o un mutante de FGF21 (incluyendo una forma truncada de FGF21 o un mutante de FGF21) a través de, por ejemplo, un enlace covalente entre el dominio Fc y la secuencia de FGF21. Tales proteínas de fusión pueden formar multímeros a través de la asociación de los dominios Fc y estas proteínas de fusión y sus multímeros son un aspecto del presente invento.

[00099] Tal como se usa en el presente texto, el término "cuerpo fusionado" se refiere a una molecula soluble tipo anticuerpo que comprende dos heterodímeros, cada heterodímero consta de una cadena pesada y una cadena ligera de aminoácidos, asociados juntos de manera estable, por ejemplo, a través de uno o más enlaces de disulfuro. Cada cadena pesada o ligera comprende regiones constantes de un anticuerpo, denominadas en lo sucesivo, respectivamente, como regiones constantes de cadena pesada y ligera del cuerpo fusionado. La región constante de cadena pesada comprende al menos la región CH_1 de un anticuerpo y puede comprender además regiones CH_2 y CH_3, incluyendo la región de bisagra. La región constante de cadena ligera comprende una región C_L de un anticuerpo. En un cuerpo fusionado, las regiones variables de un anticuerpo se sustituyen por dominios de unión heterólogos solubles. A modo de ejemplo no limitativo, un cuerpo fusionado de la invención puede comprender una proteína de fusión de función dual de la invención, en donde el agonista del receptor GLP-1 se fusiona con el extremo-N de la cadena pesada y ligera de un anticuerpo y FGF21 es simultáneamente fusionado al extremo-C de la cadena pesada y ligera del mismo anticuerpo.

[000100] El término “heterólogo” significa que estos dominios no se encuentran de manera natural asociados con las regiones constantes de un anticuerpo. En particular, dichos dominios de unión heterólogos no tienen la estructura típica de un dominio variable de anticuerpo que consiste en 4 regiones marco, FR1, FR2, FR3 y FR4 y las 3 regiones determinantes de complementariedad (CDRs) en el medio. Cada brazo del cuerpo fusionado por lo tanto comprende un primer polipéptido de cadena única que comprende un primer dominio de unión unido covalentemente en la parte N-terminal de una región constante de cadena pesada C_H1 de un anticuerpo, y un segundo polipéptido de cadena única que comprende un segundo dominio de unión unido covalentemente a la parte N-terminal de una cadena constante ligera C_L de un anticuerpo. El enlace covalente puede ser directo, por ejemplo a través de un enlace peptídico o indirecto, a través de un enlazador, por ejemplo, un enlazador peptídico. Los dos heterodímeros del cuerpo fusionado se unen covalentemente, por ejemplo, por al menos un puente de disulfuro en su región de bisagra, como una estructura de anticuerpo. Ejemplos de moléculas con una estructura de cuerpo fusionado se han descrito en la técnica, en particular, cuerpos fusionados que comprenden la región de unión a ligando del receptor heterodimérico (véase, por ejemplo las publicaciones de patente internacionales WO01/46261 y WO11/076781).

[000101] El término "polietilenglicol" o "PEG" se refiere a un compuesto de polialquilenglicol o un derivado del mismo, con o sin agentes de acoplamiento o derivatización con fracciones acopladoras o activadoras.

[000102] El término "exenatida" indica una versión sintética de la exendina-4, exenatida, comercializado como Byetta y Bydureon, es un medicamento agonista del péptido-1 tipo glucagón (agonista GLP-1) aprobado en abril de 2005 para el tratamiento de la diabetes mellitus tipo 2. Pertenece al grupo de los miméticos de incretina y es fabricado por Amylin Pharmaceuticals.

[000103] La exendina-4 es un polipéptido de 39 residuos producido en las glándulas salivales del lagarto monstruo de Gila (Goke et al. (1993) Diabetes 46:433-439); (Fehmann et al. (1995) Endocrine Rev. 16:390-410). Aunque es el producto de un gen excepcionalmente no mamífero y parece que se expresa solamente en la glándula salival, la exendina-4 comparte una homología de secuencia aminoácida del 52% con GLP-1 y en mamíferos interactúa con el receptor GLP-1 (Goke) (Thorens et al. (1993) Diabetes 42:1678-1682).

[000104] El término "trastornos asociados con FGF21", “trastornos asociados con GLP-1,” “trastornos asociados con Exendina-4”y los términos utilizados de manera similar en la presente, incluyen pero no se limitan a la obesidad, diabetes mellitus tipo 1 y tipo 2, pancreatitis, dislipidemia, enfermedad hepática grasa no alcohólica, esteatohepatitis no alcohólica (EHNA), resistencia a la insulina, hiperinsulinemia, intolerancia a la glucosa, hiperglucemia , síndrome metabólico, infarto agudo del miocardio, hipertensión, enfermedades cardiovasculares, arteriosclerosis, enfermedad arterial periférica, accidente cerebro vascular, insuficiencia cardiaca, enfermedad coronaria, enfermedad renal, complicaciones de la diabetes, neuropatía, gastroparesia, trastornos asociados con severas mutaciones inactivadoras en el receptor de insulina, lipodistrofias incluyendo lipodistrofia asociada a VIH y otros trastornos metabólicos.

[000105] El término "trastornos asociados con severas mutaciones inactivadoras en el receptor de la insulina", y términos similares utilizados en la presente, describen las condiciones en los sujetos afectados con mutaciones en el receptor de la insulina (o posibles proteínas directamente aguas abajo del mismo) que causan severa resistencia a la insulina, pero a menudo (aunque no siempre) se observan sin la obesidad común en la diabetes mellitus tipo 2. En muchos sentidos, los sujetos que padecen estas enfermedades manifiestan síntomas híbridos de la diabetes mellitus tipo 1 y diabetes mellitus tipo 2. Los sujetos de este modo afligidos caen en varias categorías de severidad más o menos creciente, incluyendo: Resistencia a la insulina Tipo A, resistencia a la insulina Tipo C (síndrome AKA HAIR-AN), Síndrome Rabson-Mendenhall y, finalmente, el Síndrome de Donohue o Leprechaunismo. Estos trastornos están asociados con niveles muy altos de insulina endógena, y muy a menudo hiperglucemia,. Los sujetos así afectados también presentan diversas características clínicas asociadas con la "toxicidad por insulina", incluyendo hiperandrogenismo, síndrome de ovario poliquístico (SOP), hirsutismo, y acantosis nigricans (crecimiento y pigmentación excesivos) en los pliegues de la piel.

[000106] "Lipodistrofias, incluyendo lipodistrofia asociada al VIH" son trastornos de tejido adiposo que se caracteriza por una pérdida selectiva de grasa corporal Los pacientes con lipodistrofia tienen una tendencia a desarrollar resistencia a la insulina, hiperinsulinemia, hiperglucemia, hipertrigliceridemia e hígado graso. Existen numerosas formas de lipodistrofia que son heredados (genéticos) o adquiridos. Las formas genéticas de la lipodistrofia incluyen lipodistrofia congénita generalizada (el síndrome de Berardinelli-Seip) y varios tipos de lipodistrofia familiar parcial (el tipo Dunnigan, el

tipo de Kobberling, el tipo displasia mandibuloacral). Las formas adquiridas de lipodistrofia incluyen lipodistrofia generalizada adquirida (síndrome de Lawrence), lipodistrofia parcial adquirida (síndrome de Barraquer-Simons), y lipodistrofia inducida por los inhibidores de la proteasa e inhibidores de la nucleosido transcriptasa inversa utilizados para tratar el VIH.

[000107] "La diabetes mellitas tipo 2" es una condición caracterizada por la producción de exceso de glucosa a pesar de la disponibilidad de insulina y los niveles de glucosa circulante siguen siendo excesivamente altos como consecuencia de la eliminación inadecuada de glucosa.

[000108] "La diabetes mellitus tipo 1" es una condición caracterizada por altos niveles de glucosa en la sangre causada por la falta total de insulina. Esto ocurre cuando el sistema inmunológico del cuerpo ataca las células beta productoras de insulina en el páncreas y las destruye. Entonces, el páncreas produce poca o ninguna insulina.

[000109] La "intolerancia a la glucosa", o intolerancia a la glucosa deteriorada (IGT) es un estado pre-diabético de la disglucemia que se asocia con un mayor riesgo de patología cardiovascular. La condición pre-diabética impide que un sujeto mueva la glucosa en las células de manera eficiente y la utilice como fuente de combustible eficiente, llevando a niveles elevados de glucosa en la sangre y algún grado de resistencia a la insulina.

[000110] "La hiperglucemia" se define como un exceso de azúcar (glucosa) en la sangre.

[000111] "Hipoglucemia", también conocida como azúcar baja en la sangre, se produce cuando el nivel de glucosa en sangre baja demasiado, para proporcionar la energía suficiente para las actividades del cuerpo.

[000112] "La hiperinsulinemia" se define como un nivel más alto de lo normal de la insulina en la sangre.

[000113] "Resistencia a la insulina" se define como un estado en el que una cantidad normal de insulina produce una respuesta biológica subnormal.

[000114] "La obesidad," en términos del sujeto humano, se puede definir como que el peso corporal de más de 20 por ciento por encima del peso ideal para una población dada (R.H. Williams, Textbook of Endocrinology, 1974, p. 904-916).

[000115] Las "complicaciones de la diabetes" son problemas causados por altos niveles de glucosa en la sangre, con otras funciones del cuerpo como los riñones, los nervios (neuropatía), pies (úlceras en los pies y la mala circulación) y los ojos (retinopatías, por ejemplo). La diabetes también aumenta el riesgo de enfermedad cardiaca y trastornos

óseos y articulares. Otras complicaciones a largo plazo de la diabetes incluyen problemas de la piel, problemas digestivos, disfunción sexual y problemas con los dientes y las encías.

[000116] "El síndrome metabólico" se puede definir como un grupo de al menos tres de los siguientes signos: la grasa abdominal - en la mayoría de los hombres, una cintura de 40 pulgadas o más; azúcar en la sangre - por lo menos 110 miligramos por decilitro (mg/dl) después de un ayuno; triglicéridos altos - por lo menos 150 mg/dL en la sangre, HDL bajo- menos de 40 mg/dl, y, presión arterial de 130/85 mmHg o superior.

[000117] "Pancreatitis" es la inflamación del páncreas.

[000118] "Dislipidemia" es un trastorno del metabolismo de las lipoproteínas, incluyendo la sobreproducción o deficiencia de lipoproteínas. Las dislipidemias se pueden manifestar por la elevación de las concentraciones de colesterol total, colesterol de lipoproteínas de baja densidad (LDL) y de triglicéridos y una disminución en la concentración de colesterol de lipoproteínas de alta densidad (HDL) en la sangre.

[000119] “Enfermedad hepática grasa no alcohólica (NAFLD)”, es una enfermedad hepática, no asociada con el consumo de alcohol, caracterizada por el cambio graso de los hepatocitos.

[000120] "La esteatohepatitis no alcohólica (EHNA)" es una enfermedad del hígado, no se asocia con el consumo de alcohol, se caracteriza por el cambio graso de los hepatocitos, acompañada por la inflamación y la fibrosis intralobular.

[000121] "La hipertensión" o presión arterial alta es una elevación transitoria o sostenida de la presión arterial sistémica a un nivel que pueda provocar daños cardiovasculares u otras consecuencias adversas. La hipertensión se ha definido arbitrariamente como una presión arterial sistólica por encima de 140 mmHg o una presión arterial diastólica por encima de 90 mmHg.

[000122] "Las enfermedades cardiovasculares" son las enfermedades relacionadas con el corazón o los vasos sanguíneos.

[000123] “Infarto agudo del miocardio” ocurre cuando existe la interrupción del suministro de sangre a una parte del corazón. La isquemia resultante y la falta de oxígeno, si se deja sin tratamiento durante un período de tiempo suficiente, pueden causar daño o muerte (infarto) del tejido muscular cardiaco (miocardio)

[000124] "La enfermedad arterial periférica" ocurre cuando se acumula placa en las arterias que llevan sangre a la cabeza, órganos y extremidades. Con el tiempo, la placa se

endurece y estrecha las arterias, lo cual limita el flujo de sangre rica en oxígeno a los órganos y otras partes del cuerpo.

[000125] "Aterosclerosis" es una enfermedad vascular caracterizada por depósitos de lípidos distribuidos irregularmente en la íntima de las arterias grandes y medianas, provocando un estrechamiento de las luces arteriales y procediendo eventualmente a la fibrosis y calcificación. Las lesiones suelen ser focales y progresar lentamente y de forma intermitente. La limitación del flujo de sangre representa la mayoría de las manifestaciones clínicas que varían con la distribución y gravedad de las lesiones.

[000126] "Accidente cerebro vascular" es cualquier evento clínico agudo, relacionado con el deterioro de la circulación cerebral, que dura más de 24 horas. Un accidente cerebro vascular implica un daño cerebral irreversible, el tipo y la severidad de los síntomas están en función de la localización y extensión del tejido cerebral cuya circulación se ha visto comprometida.

[000127] "La insuficiencia cardiaca", también denominada insuficiencia cardíaca congestiva, es una condición en la cual el corazón ya no puede bombear suficiente sangre al resto del cuerpo.

[000128] "La enfermedad coronaria", también llamada enfermedad de las arterias coronarias, es un estrechamiento de los pequeños vasos sanguíneos que suministran sangre y oxígeno al corazón.

[000129] "La enfermedad renal" o nefropatía es una enfermedad del riñón. La nefropatía diabética es una causa importante de morbilidad y mortalidad en personas con diabetes mellitus tipo 1 o tipo 2.

[000130] "Neuropatías" son enfermedades que afectan a los nervios craneales o el sistema nervioso periférico o autónomo.

[000131] "Gastroparesia" es la debilidad de la peristalsis gástrica, que se traduce en retraso en el vaciado de los intestinos.

[000132] "Química click" es un término que fue introducido por K.B. Sharpless en 2001 para describir las reacciones que son de alto rendimiento, amplias en su alcance, que crean solos subproductos que se pueden eliminar sin cromatografía, son estereoespecíficas, fáciles de realizar, y pueden llevarse a cabo en disolventes fácilmente extraíbles o benignos.

[000133] Los pacientes en estado crítico contemplados por el presente invento generalmente experimentan un estado hipermetabólico inestable. Este estado metabólico inestable se debe a los cambios en el metabolismo del sustrato, que puede llevar a

deficiencias relativas en algunos nutrientes. Por lo general hay un aumento de la oxidación de grasa y músculo.

[000134] Además, los pacientes críticamente enfermos son preferentemente pacientes que experimentan el síndrome de respuesta inflamatoria sistémica o dificultad respiratoria. Una reducción de la morbilidad significa reducir la probabilidad de que un paciente críticamente enfermo desarrolle enfermedades, condiciones, o síntomas adicionales o la reducción de la severidad de enfermedades, condiciones, o síntomas adicionales. Por ejemplo la reducción de la morbilidad puede corresponder a una disminución en la incidencia de bacteremia o sepsis o complicaciones asociadas con la falla de múltiples órganos.

[000135] Como se utilizan en la presente, las formas singulares "un", "una" y "la" incluyen referencias plurales a menos que el contenido claramente dicte lo contrario. Así, por ejemplo, la referencia a "un anticuerpo" incluye una mezcla de dos o más anticuerpos.

[000136] Como se utiliza en la presente, el término "aproximadamente" se refiere a +/- 20%, +/- 10%, o +/- 5% de un valor.

[000137] Los términos "polipéptido" y "proteína", se utilizan indistintamente y se refieren a una forma polimérica de aminoácidos de cualquier longitud, que puede incluir aminoácidos codificados y sin codificar, aminoácidos naturales y no naturales, aminoácidos química o bioquímicamente modificados o derivatizados, y polipéptidos con esqueletos peptídicos modificados. El término incluye las proteínas de fusión, incluyendo pero sin limitar a, las proteínas de fusión con una secuencia aminoácida heteróloga, fusiones con secuencias líderes heterólogas y homólogas, con o sin residuos de metionina N-terminales; proteínas inmunológicamente etiquetadas; y similares.

[000138] Los términos "individuo", "sujeto", "huésped" y "paciente" se usan indistintamente y se refieren a cualquier sujeto para el cual se desea el diagnóstico, tratamiento o terapia, en particular los seres humanos. Otros sujetos pueden incluir ganado, perros, gatos, cobayas, conejos, ratas, ratones, caballos, y similares. En algunas realizaciones preferidas el sujeto es un ser humano.

[000139] Como se utilizan en la presente, el término "muestra" se refiere a material biológico de un paciente. La muestra ensayada por el presente invento no se limita a un determinado tipo. Las muestras incluyen como ejemplos no limitativos, células individuales, múltiples células, tejidos, tumores, fluidos biológicos, moléculas biológicas, o sobrenadantes o extractos de cualquiera de los anteriores. Los ejemplos

incluyen tejido para biopsia, tejido extraído durante la resección, muestras de sangre, orina, tejido linfático, fluido linfático, fluido cefalorraquídeo, moco, y muestras coprológicas. La muestra utilizada puede variar en función del formato de ensayo, el método de detección y la naturaleza de los tumores, tejidos, células o extractos a ensayar. Los métodos para preparar las muestras son bien conocidos en la técnica y pueden ser fácilmente adaptados con el fin de obtener una muestra de que es compatible con el método utilizado.

[000140] Como se utiliza en la presente, el término "molécula biológica" incluye, pero no se limita a, polipéptidos, ácidos nucleicos, y sacáridos.

[000141] Como se utilizan en la presente, el término "modulación" se refiere a un cambio en la calidad o cantidad de un gen, proteína, o cualquier otra molécula que se encuentra dentro, fuera o sobre la superficie de una célula. El cambio puede ser un aumento o una disminución en la expresión o el nivel de la molécula. El término "modula" también incluye el cambio de calidad o cantidad de una función/actividad biológica, incluyendo, sin limitación, la capacidad para reducir los niveles de glucosa en sangre, insulina, triglicéridos o colesterol, reducir los niveles de lípidos del hígado o triglicéridos en el hígado, reducir el peso corporal y mejorar la tolerancia a la glucosa, el gasto energético, o sensibilidad a la insulina.

[000142] Tal como se utiliza en la presente, el término "modulador" se refiere a una composición que modula uno o más eventos fisiológicos y bioquímicos relacionados con un trastorno asociado de FGF21, tales como la diabetes mellitus tipo 1 o tipo 2 o una afección metabólica como la obesidad. Dichos eventos incluyen, pero no se limitan a la capacidad de reducir los niveles de glucosa en sangre, insulina, triglicéridos o colesterol, reducir los niveles de lípidos del hígado o de triglicéridos en el hígado; reducir el peso corporal, y mejorar la tolerancia a la glucosa, el gasto energético, o la sensibilidad a la insulina.

[000143] Un "producto génico" es un producto biopolimérico que se expresa o se produce por un gen. Un producto génico puede ser, por ejemplo, ARN sin cortar ni empalmar, un ARNm, un variante cortado y empalmado de ARNm, un polipéptido, un polipéptido modificado pos-traducción, un polipéptido variante de corte y empalme, etc. También incluidos por este término están los productos biopoliméricos que se elaboran utilizando un producto génico de ARN como un modelo (es decir, ADNc del RNA). Un producto génico puede ser hecho enzimáticamente, recombinantemente, químicamente, o dentro de una célula a la que el gen es nativo. En algunas realizaciones, si el producto

génico es proteinaceo, éste presenta una actividad biológica. En algunas realizaciones, si el producto génico es un ácido nucleico, éste puede traducirse en un producto génico proteinaceo que presenta una actividad biológica.

[000144] "Modulación de la actividad de FGF21," "Modulación de la actividad de GLP-1", y "Modulación de la actividad de Exendina-4", como se utilizan en la presente, se refieren a un aumento o disminución en actividad de FGF21, GLP-1, o Exendina-4, respectivamente, que puede ser el resultado de, por ejemplo, la interacción de un agente con un polinucleótido o polipéptido de FGF21, GLP-1 o Exendina-4, inhibición de la transcripción y/o traducción de FGF21, GLP-1 o exendina-4 (por ejemplo, a través de la interacción antisentido o ARNsi con el gen de FGF21, GLP-1 o exendina-4 o el transcripto de FGF21, GLP-1 o exendina-4, mediante la modulación de factores de transcripción que facilitan la expresión de FGF21, GLP-1, o exendina-4), y similares. Por ejemplo, la modulación de la actividad biológica se refiere a un aumento o una disminución en la actividad biológica. La actividad de FGF21, GLP-1 o exendina-4 puede evaluarse por medios que incluyen, sin limitación, el ensayo de glucosa en sangre, los niveles de insulina, triglicéridos, o colesterol en un sujeto, la evaluación de los niveles polipeptídicos de FGF21, GLP-1 o exendina-4 o evaluación de los niveles de transcripción de FGF21, GLP-1 o Exendina-4.

[000145] Las comparaciones de actividad de FGF21, GLP-1, o Exendina-4 también pueden realizarse mediante, por ejemplo, la medición de los niveles de un biomarcador corriente abajo de FGF21, GLP-1 o Exendina-4, y la medición de los aumentos en señalización de FGF21, GLP-1 o Exendina -4. La actividad de FGF21, GLP-1, o Exendina-4 también se puede evaluar mediante la medición de: la señalización celular, la actividad de quinasa, la absorción de glucosa en adipocitos; midiendo fluctuaciones de los niveles de insulina en sangre, triglicéridos, o colesterol, cambios en los niveles de lípidos en el hígado o triglicéridos en el hígado, interacciones entre el FGF21, GLP-1, o Exendina-4 y un receptor; o la fosforilación de un receptor de FGF21, GLP-1, ó Exendina-4. En algunas realizaciones la fosforilación de un receptor de FGF21, GLP-1 ó Exendina-4 puede ser la fosforilación de la tirosina. En algunas realizaciones, la modulación de actividad de FGF21, GLP-1 ó Exendina-4 puede producir la modulación de un fenotipo relacionado con FGF21, GLP-1, ó Exendina-4.

[000146] A "biomarcador de FGF21, GLP-1, o Exendina-4 corriente abajo", como se utiliza en la presente, es un gen o producto génico, o indicios medibles de un producto génico o gen. En algunas realizaciones, un gen o una actividad que es un marcador

corriente abajo de FGF21, GLP-1, o Exendina-4 exhibe un nivel alterado de expresión. En algunas realizaciones, una actividad del marcador corriente abajo se altera en presencia de un modulador de FGF21, GLP-1, o Exendina-4. En algunas realizaciones, los marcadores corriente abajo muestran niveles alterados de expresión cuando la señalización de FGF21, GLP-1, o Exendina-4 es perturbada con una proteína de función dual de la presente invención. Los marcadores corriente abajo de FGF21 incluyen, sin limitación, proteínas de absorción de glucosa o 2-desoxi-glucosa, pERK y otras proteínas fosforiladas o acetiladas o niveles NAD.

[000147] Tal como se utiliza en la presente, el término "regula de manera positiva" se refiere a un aumento, activación o estimulación de una actividad o cantidad. Por ejemplo, los moduladores de FGF21, GLP-1, o Exendina-4, tales como las proteínas de función dual de la invención, puede aumentar o agonizar la actividad de un receptor de FGF21, GLP-1 o Exendina-4 En una realización, uno o más de FGFR-1(IIIc), FGFR-2(IIIc), o FGFR-3(IIIc) y/o β-klotho pueden regularse de manera positiva en respuesta a una proteína de función dual de la invención. La regulación positiva también puede referirse a una actividad relacionada con FGF21, GLP-1, ó Exendina-4, tal como por ejemplo, la capacidad para reducir los niveles de la glucosa en sangre, insulina, triglicéridos o colesterol, reducir los niveles de lípidos o triglicéridos en el hígado, reducir el peso corporal, mejorar la tolerancia a la glucosa, el gasto energético, o la sensibilidad a la insulina, o producir la fosforilación de un receptor de FGF21, GLP-1, ó Exendina-4, o incrementar un marcador corriente abajo de FGF21, GLP-1, ó Exendina-4. Por ejemplo, el receptor FGFR21 puede ser uno o más de FGFR-1(IIIc), FGFR-2(IIIc), o FGFR-3(IIIc) y/o β-klotho. La regulación positiva puede oscilar entre 25% y 500% en comparación con un control.

[000148] Como se utiliza en la presente, el término "extremo N " se refiere a por lo menos los primeros 20 aminoácidos de una proteína.

[000149] Como se utilizan en la presente, los términos "dominio N-terminal" y "región N-terminal" se usan indistintamente y se refieren a un fragmento de una proteína que se inicia en el primer aminoácido de la proteína y termina en cualquier aminoácido en la mitad N- terminal de la proteína. Por ejemplo, el dominio N-terminal de FGF21 es desde el aminoácido 1 de la SEQ ID NO: 1 a cualquier aminoácido entre los aminoácidos 10 y 105 de la SEQ ID NO: 1

[000150] Como se utiliza en la presente, el término "extremo C " se refiere a por lo menos los últimos 20 aminoácidos de una proteína.

[000151] Como se utilizan en la presente, los términos "dominio C-terminal" y "región C-terminal " se usan indistintamente y se refieren a un fragmento de una proteína que se inicia en cualquier aminoácido en el medio C-terminal de la proteína y termina en el último aminoácido de la proteína. Por ejemplo, el dominio C-terminal de FGF21 se inicia en cualquier aminoácido desde el aminoácido 105 hasta aproximadamente el aminoácido 200 de la SEQ ID NO: 1 y termina en el aminoácido 209 de la SEQ ID NO: 1.

[000152] El término "dominio" como se utilizan en la presente se refiere a una parte estructural de una biomolécula que contribuye a una función conocida o sospechada de la biomolécula. Los dominios pueden ser co-extensivos con las regiones o partes del mismo y también pueden incorporar una parte de una biomolécula que es distinta de una región particular, además de toda o parte de esa región.

[000153] Como se utiliza en la presente, el término "dominio señal" (también llamado "secuencia señal" o "péptido señal") se refiere a un dominio peptídico que se encuentra en un tramo continuo de la secuencia aminoácida en la región N-terminal de una proteína precursora (a menudo unido por membrana o proteína secretada) y está involucrado en el transporte proteico pos-traducción. En muchos casos el dominio señal se elimina de la proteína de longitud completa por peptidasas señal especializadas después de que el proceso de clasificación se ha completado. Cada dominio señal especifica un destino particular en la célula para la proteína precursora. El dominio señal de FGF21 es los aminoácidos 1-28 de la SEQ ID NO: 1.

[000154] Como se utiliza en la presente, el término "dominio fijador de receptores" se refiere a cualquier parte o región de una proteína que entra en contacto con una proteína receptora fijada a la membrana, lo que resulta en una respuesta celular, tal como un evento de señalización.

[000155] Como se utiliza en la presente, el término "dominio fijador de ligandos" se refiere a cualquier parte o región de una proteína de fusión de la invención que retiene al menos una actividad cualitativa de unión de una secuencia nativa correspondiente.

[000156] El término "región" se refiere a una porción físicamente contigua de la estructura primaria de una biomolécula. En el caso de las proteínas, la región se define por una parte contigua de la secuencia aminoácida de esta proteína. En algunas realizaciones, una "región" se asocia con una función de la biomolécula.

[000157] El término "fragmento" como se utilizan en la presente se refiere a una porción físicamente contigua de la estructura primaria de una biomolécula. En el caso de

las proteínas, una parte se define por una parte contigua de la secuencia aminoácida de esa proteína y se refiere a por lo menos 3-5 aminoácidos, por lo menos 8-10 aminoácidos, por lo menos 11 a 15 aminoácidos, en por lo menos 17 a 24 aminoácidos, por lo menos 25 a 30 aminoácidos, y por lo menos 30 a 45 aminoácidos. En el caso de oligonucleótidos, una parte se define por una parte contigua de la secuencia de ácidos nucleicos de ese oligonucleótido y se refiere a por lo menos 9-15 nucleótidos, al menos 18 a 30 nucleótidos, al menos 33 a 45 nucleótidos, al menos 48 - 72 nucleótidos, al menos 75-90 nucleótidos, y al menos 90-130 nucleótidos. En algunas realizaciones, algunas partes de las biomoléculas tienen una actividad biológica. En el contexto del presente invento, fragmentos polipeptídicos de FGF21 no comprenden toda la secuencia polipéptidica de FGF21 divulgada en SEQ ID NO: 1.

[000158] Un polipéptido de "secuencia nativa" es aquel que tiene la misma secuencia aminoácida que un polipéptido derivado de la naturaleza. Tales polipéptidos de secuencia nativa pueden aislarse de la naturaleza o se pueden producir por vía recombinante o sintética. Por lo tanto, un polipéptido de secuencia nativa puede tener la secuencia aminoácida de polipéptido humano natural, polipéptido murino, o polipéptido de cualquier otra especie mamífera.

[000159] Como se utiliza en la presente, la frase "secuencia nucleótida homóloga," o "secuencia aminoácida homóloga", o variaciones de la misma, se refiere a las secuencias caracterizadas por una homología, a nivel nucleótido o a nivel aminoácido, de por lo menos un porcentaje determinado y se utiliza de manera intercambiable con "identidad de secuencia". Secuencias nucleótidas homólogas incluyen aquellas secuencias que codifican las isoformas de las proteínas. Tales isoformas pueden expresarse en diferentes tejidos del mismo organismo como resultado de, por ejemplo, un corte y empalme alternativo del ARN. Por otra parte, las isoformas pueden ser codificadas por genes diferentes. Las secuencias nucleótidas homólogas incluyen secuencias nucleótidas que codifican una proteína de una especie diferente a los seres humanos, incluyendo pero sin limitar a, mamíferos. Secuencias nucleótidas homólogas también incluyen, pero no se limitan a, las variaciones alélicas de origen natural y las mutaciones de las secuencias nucleótidas divulgadas en el presente. Secuencias aminoácidas homólogas incluyen las secuencias aminoácidas que contienen sustituciones aminoácidas conservativas y cuyos polipéptidos tengan la misma y/o actividad. En algunas realizaciones, una secuencia nucleótida o aminoácido es homóloga si tiene al menos 60% o más, hasta 99%, de identidad con una secuencia comparadora. En algunas realizaciones, una secuencia

nucleótida o aminoácido es homóloga si comparte uno o más, hasta 60 sustituciones, adicionales, o supresiones nucleótidas/aminoácidos con una secuencia de comparación. En algunas realizaciones, las secuencias aminoácidas homólogas no tienen más de cinco o no más de tres sustituciones aminoácidas conservadoras.

[000160] El porcentaje de homología o identidad puede determinarse, por ejemplo, mediante el programa Gap (Paquete de Análisis de Secuencias de Wisconsin, versión 8 para UNIX, Genetics Computer Group, Universidad De investigación Park, Madison WI), con la configuración predeterminada, que utiliza el algoritmo de Smith y Waterman (Adv. Appl. Math., 1981, 2, 482 a 489). En algunas realizaciones, la homología entre la sonda y la diana está alrededor del 75% al 85%. En algunas realizaciones, los ácidos nucleicos tienen nucleótidos que son al menos 95%, aproximadamente 97%, aproximadamente 98%, aproximadamente 99% y aproximadamente 100% de homología con la SEC ID NO: 2, o una parte de la misma.

[000161] La homología también puede ser a nivel de polipéptidos. En algunas realizaciones, los polipéptidos constituyentes de las proteínas de función dual de la invención pueden ser al menos 95% homólogos con sus contrapartes tipo silvestre de longitud completa o secuencias nativas correspondientes, o porciones de las mismas. El grado o porcentaje de identidad de las Proteínas de Fusión de la invención, o porciones de las mismas, y secuencias aminoácidas diferentes se calcula como el número de equivalencias exactas en una alineación de dos secuencias, dividido por la longitud de la “secuencia de la invención” o la “secuencia extraña”, la que sea más corta. El resultado se expresa como el porcentaje de identidad.

[000162] Como se utiliza en la presente, el término "mezclar" se refiere al proceso de combinar de uno o más compuestos, células, moléculas, y similares juntos en la misma zona. Esto se puede realizar, por ejemplo, en un tubo de ensayo, placa de Petri, o cualquier otro recipiente que permite a los compuestos, células o moléculas, mezclarse.

[000163] Como se utiliza en la presente, el término "sustancialmente purificado" se refiere a un compuesto (por ejemplo, un polinucleótido o un polipéptido o un anticuerpo) que se extrae de su entorno natural y es por lo menos 60% libre, por lo menos 75% libre, y por lo menos 90% libre de otros componentes con los que se asocia naturalmente.

[000164] El término "vehículo farmacéuticamente aceptable" se refiere a un vehículo para la administración de un agente terapéutico, tales como los anticuerpos o un polipéptido, genes y otros agentes terapéuticos. El término se refiere a cualquier vehículo farmacéutico que no induce por sí mismo la producción de anticuerpos dañinos

para el individuo que recibe la composición, y que se puede administrar sin toxicidad indebida. Los vehículos adecuados pueden ser grandes macromoléculas, metabolizadas lentamente tales como las proteínas, polisacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros aminoácidos, agregados lípidos y partículas de virus inactivas. Dichos vehículos son bien conocidos por los expertos en el estado de la técnica. Los vehículos farmacéuticamente aceptables en composiciones terapéuticas pueden incluir líquidos tales como agua, solución salina, glicerol y etanol. Sustancias auxiliares, tales como agentes humectantes o emulsionantes, sustancias amortiguadoras del pH, y similares, también pueden estar presentes en dichos vehículos.

[000165] Los enlaces de disulfuro de origen natural, como se proporcionan mediante residuos de cisteína, generalmente aumentan la estabilidad termodinámica de las proteínas. Ejemplos exitosos de la estabilidad termodinámica aumentada, medida en aumento de la temperatura de fusión, son múltiples mutantes unidos por disulfuro de las enzimas lisozima T4 (Matsumura, PNAS 86:6562-6566 (1989)) y barnasa (Johnson ., J. Mol. Biol. 268:198-208 (1997)). Un aspecto del presente invento es una mejora de la estabilidad física de FGF21 en la presencia de un conservante, lograda por la presencia de enlaces de disulfuro dentro de los variantes, lo cual restringe la flexibilidad del FGF21 tipo silvestre y así limita el acceso del presertativo al núcleo hidrofóbico de la proteína. La mejora de la estabilidad física de FGF21 dentro de las proteínas de la presente invención, debido a la presencia de enlaces disulfuro dentro de las variantes, puede conferir protección adicional a dichas proteínas, ya sea o no en la presencia de un conservante, incluyendo, pero no limitado a la protección contra las fluctuaciones en condiciones ambientales tales como pH y la temperatura.

Mejoras de las Proteínas de Función Dual de la Invención Sobre Comparadores de Proteínas Tipo Silvestre y Combinaciones de los Mismos

[000166] Es bien sabido en el estado de la técnica que un reto importante en el desarrollo de productos farmacéuticos de proteínas es hacer frente a las inestabilidades físicas y químicas de las proteínas. Esto es aún más evidente cuando una formulación farmacéutica proteica está destinada a ser una formulación inyectable multi-uso, que requiere una solución estable, concentrada y conservada, mientras mantiene un perfil de bioactividad favorable. La caracterización biofísica detallada de FGF21 tipo silvestre establece que una solución proteica concentrada (> 5 mg/ml), cuando se expone a condiciones de estrés, tales como altas temperaturas o bajo pH, lleva a asociación y

agregación aceleradas (es decir, poca estabilidad física y propiedades biofarmacéuticas). La exposición de una solución proteica concentrada de FGF21 a los preservativos farmacéuticos (por ejemplo, m-cresol) también tuvo un impacto negativo sobre la estabilidad física.

[000167] Por lo tanto, una realización del presente invento pretende mejorar la estabilidad física de las soluciones concentradas, manteniendo la estabilidad química y la potencia biológica, tanto en condiciones fisiológicas y preservadas de la formulación. Se cree que la asociación y agregación pueden resultar de las interacciones hidrofóbicas, ya que, a una concentración proteica determinada, la temperatura y fuerza iónica tienen un impacto considerable sobre la estabilidad física.

[000168] Principalmente, se atacaron residuos aminoácidos no conservados, que se presumían expuestos en superficie. El ambiente local de los residuos se analizó y, aquellos no considerados estructuralmente importantes fueron seleccionados para la mutagénesis. Un método para iniciar los cambios específicos es disminuir aún más el pI de la proteína mediante la introducción de residuos de ácido glutámico ("barrido con ácido glutámico"). La introducción de sustitutos cargados puede inhibir la agregación hidrofóbica mediada a través de una repulsión de carga contra carga y, potencialmente, mejorar la compatibilidad preservativa. Además, el experto en el estado de la técnica también reconocería que con suficiente mutagénesis el pI podría ser cambiado en un rango de pH básico mediante la introducción de carga positiva con o sin disminución concomitante en carga negativa, lo que permite la repulsión de carga contra carga.

[000169] Una dificultad adicional asociada con las aplicaciones terapéuticas de FGF21 tipo silvestre como un bioterapéutico, por ejemplo, es que su vida media es muy corta in vivo (en el orden de 0,5 y 2 h, respectivamente, en ratón y primate). Existe por lo tanto una necesidad de desarrollar compuestos de seguimiento que sean más eficaces ya sea a través de una potencia más alta o vida media mayor. Del mismo modo, se han empleado diversos mecanismos para mejorar la vida media sérica de GLP-1 (1-2 minutos debido a la rápida escisión por proteasas endógenas, particularmente dipeptidil peptidasa 4- (DPP4)) mediante el uso de Exendina-4, u otros análogos que son resistentes a la escisión por DPP-4 así como otras modificaciones y formulaciones para extender la vida media o retardar la liberación del compuesto.

[000170] Las proteínas de la invención, por ejemplo, las proteínas duales de agonistas de receptores de FGF21 y de GLP-1 de la invención, fueron desarrolladas como una forma para lograr los efectos deseables de tratamiento con FGF21 y GLP-1 una potencia

más alta y en una formulación con vida media extendida. La co-administración de GLP-1 y FGF21 se ha descrito en la literatura, por ejemplo, en las publicaciones de patentes WO2010/068735 y WO2009/020802), con datos que indican efectos aditivos o efectos sinérgicos de la co-administración de GLP-1 y FGF21 para el tratamiento de la obesidad y la diabetes tipo 2. Sin embargo, la co-localización de los dos agonistas de los receptores en una sola molécula, por ejemplo, en la forma de las proteínas de la invención, proporciona un mejor acceso de ambos GLP-1 y FGF21 a tejidos o células y proporciona un beneficio aumentado durante la co-administración sencilla como entidades separadas. Esta ventaja de combinar dos agonistas en una sola molécula en comparación con la co-administración de cada uno es especialmente ventajoso para tejidos en los que tanto el GLP-1 y los receptores de FGF21 se expresan en los mismos tejidos, tales como tejido adiposo, células □- pancreáticas, hepáticas e hipotalámicas. Proteínas de agonista del receptor de FGF21 Dual y agonista del receptor GLP-1 de la invención y otras entidades de actividad única dual han mejorado las propiedades biológicas debido al tráfico alterado de receptores, efectos de transducción de señales alteradas y/o efectos entrópicos (avidez).

[000171] Debido a que FGF21 y GLP-1 (o Exendina-4) actúan a través de diferentes mecanismos, se espera que tengan efectos aditivos o sinérgicos cuando, por ejemplo, se administran al mismo tiempo en la forma de las proteínas de función dual de la invención. GLP-1 y Exendina-4 actúan principalmente para aumentar la secreción de insulina en respuesta a la ingesta de alimentos mientras que FGF21 sensibiliza el cuerpo para responder mejor a la insulina, lo que puede proporcionar beneficios tanto para diabetes tipo 1 y diabetes tipo 2 en el manejo de los niveles de glucosa en sangre. Los efectos protectores de células beta combinados con la mejora de la función de células beta y la sensibilidad a la insulina tienen el potencial de proporcionar un beneficio incluso en diabéticos tipo 1.

[000172] Otro ejemplo es la pérdida de peso: La señal de saciedad y el vaciado gástrico lento por GLP-1 o Exendina-4 se espera reduzcan el apetito, mientras FGF21 aumenta el metabolismo en tejidos adiposos y otros tejidos que podrían aumentar la pérdida de grasa. Estos dos efectos podrían resultar en una pérdida de peso aditiva o aún sinérgica con la dosificación combinada. La expresión de los receptores tanto para FGF21 como para GLP-1 en tejidos metabólicamente activos sugiere que la liberación simultánea de un agonista del receptor de FGF21 y un agonista del receptor de GLP-1 a tales tejidos puede ser beneficiosa para el tratamiento de enfermedades metabólicas. Una

combinación de los efectos cardioprotectores de GLP-1 y Exendina-4 y perfiles lípidos mejorados observados con FGF21 también podría resultar en un beneficio aditivo para la enfermedad cardiovascular asociada con obesidad, diabetes tipo 2 y tipo 1.

[000173] Otro beneficio probable de agonismo dual de GLP-1 y FGF21 es un mejor perfil de efectos secundarios, es decir reducidos. Estudios recientes (Wei, (2012) PNAS 109, 3143-48) indican que el tratamiento de ratones obesos inducida por la dieta con FGF21 induce la pérdida de hueso, debido a una disminución de la inactivación de PPAR□ (a través de la sumoilación reducida); un cambio de la diferenciación de células madre mesenquimales de osteoblastos a adipocitos se ve en la presencia de aumento de la actividad de PPAR □ en el hueso después del tratamiento con FGF21. Sin embargo, también se ha reportado (Sanz, (2010) Am J Physiol Endocrinol Metab 298, E634-E643) que el GLP-1 puede reducir la diferenciación de las células madre mesenquimales humanas a adipocitos mediante la reducción de la expresión de PPAR□. Por lo tanto, como de GLP-1 tiene el potencial de ser protector de hueso, las proteínas de fusión GLP-1-FGF21 tienen probabilidad para reducir el riesgo de pérdida de masa ósea en comparación con tratamientos con solamente FGF21.

[000174] Aunque las realizaciones de la presente invención se preocupan por la estabilidad física y química, tanto en condiciones fisiológicas como preservadas de la formulación farmacéutica, el mantenimiento de la potencia biológica de las proteínas de fusión de la invención en comparación con el FGF21 tipo silvestre también es un factor importante de consideración también. Por lo tanto, la potencia biológica de las proteínas de la presente invención se define por la capacidad de las proteínas para afectar la absorción de glucosa y/o reducir los niveles de glucosa plasmática, como se muestra en la presente en los ejemplos.

[000175] Las proteínas, polipéptidos, y/o péptidos de la invención administrados de acuerdo con esta invención pueden ser generados y/o aislados mediante cualquiera de los medios conocidos en el estado de la técnica. El método más preferido para la producción del variante es a través de metodologías de ADN recombinantes y es bien conocido por los expertos en el estado de el estado de la técnica. Estos métodos se describen en Current Protocols in Molecular Biology (John Wiley & Sons, Inc.), que se incorpora en la presente como referencia.

[000176] Además, las realizaciones preferidas incluyen un péptido biológicamente activo derivado del variante descrito en este documento. Dicho péptido contiene al menos una de las sustituciones descritas y el variante posee actividad biológica. El

péptido puede ser producido por cualquiera y todos los medios conocidos por los expertos en el estado de la técnica, ejemplos incluyen pero no se limitan a la digestión enzimática, la síntesis química o metodologías de ADN recombinante.

[000177] Se establece en el estado de la técnica que los fragmentos de péptidos de ciertos factores de crecimiento de fibroblastos son biológicamente activos. Véase, por ejemplo, Baird, Proc. Natl. Acad. Ciencia (EE.UU.) 85:2324-2328 (1988), y J. Cell. Phys. Supl. 5:101-106 (1987). Por lo tanto, la selección de fragmentos o péptidos del variante se basa en criterios conocidos en el estado de la técnica. Por ejemplo, se sabe que la dipeptidil peptidasa IV (DPP-IV) es una proteasa de tipo serina involucrada en la inactivación de neuropéptidos, péptidos endocrinos, y citoquinas (Damme Chem. Immunol 72: 42-56, (1999)). El extremo N de FGF21 (HisProIlePro) contiene dos dipéptidos que podrían ser sustratos potencialmente a la DPP-IV, lo que resulta en un fragmento de FGF21 truncado en el extremo N por 4 aminoácidos. Inesperadamente, se ha demostrado que este fragmento de FGF21 tipo silvestre mantiene la actividad biológica, por lo tanto, las proteínas de la presente invención truncadas en el extremo N por un máximo de 4 aminoácidos, es una realización de la presente invención.

[000178] El invento también contempla polinucleótidos que codifican los variantes descritos anteriormente que pueden estar en forma de ARN o en la forma de ADN, tal ADN incluye ADNc, ADN genómico, y ADN sintético. El ADN puede ser de doble hebra o de hebra sencilla. Las secuencias codificadoras que codifican las proteínas de la presente invención pueden variar como consecuencia de la redundancia o degeneración del código genético.

[000179] Los polinucleótidos que codifican las proteínas de la invención pueden incluir lo siguiente: sólo la secuencia codificadora para el variante, la secuencia codificadora para el variante y la secuencia codificadora adicional tal como un polipéptido funcional, o una secuencia líder o secretora o una secuencia pro-proteica, la secuencia codificadora del variante y la secuencia no codificadora, tales como intrones o la secuencia no codificadora 5 'y/o 3' de la secuencia codificadora para el variante. Así, el término "polinucleótido que codifica un variante" abarca un polinucleótido que puede incluir no sólo la secuencia codificadora del variante, sino también un polinucleótido, que incluye la secuencia adicional codificadora y/o no codificadora.

[000180] El invento se refiere además a los variantes de los polinucleótidos descritos que codifican los fragmentos, análogos y derivados del polipéptido que contienen las sustituciones indicadas. El variante del polinucleótido puede ser un variante alélico de

origen natural de la secuencia FGF21 humana, un variante de origen no natural, o un variante truncado como se describió anteriormente. Por lo tanto, el presente invento también incluye polinucleótidos que codifican los variantes descritas anteriormente, así como los variantes de tales polinucleótidos, que codifican un fragmento, derivado o análogo del variante divulgado. Tales variantes nucleótidos incluyen variantes de supresión, variantes de sustitución, variantes truncados, y variantes de adición o inserción, siempre que al menos una de las sustituciones aminoácidas indicas de la primera o segunda realizaciones esté presente.

[000181] Los polinucleótidos de la invención se expresarán en huéspedes después que las secuencias se han enlazado de manera operativa a (es decir, posicionado para garantizar el funcionamiento de) una secuencia de control de la expresión. Estos vectores de expresión son típicamente replicables en los organismos hospederos ya sea como episomas o como una parte integral del ADN cromosómico del huésped. Por lo general, los vectores de expresión contendrán marcadores de selección, por ejemplo, tetraciclina, neomicina y la dihidrofolato reductasa, para permitir la detección de las células transformadas con las secuencias de ADN deseadas. Las proteínas de función dual o fragmentos de las mismas puede expresarse en células de mamíferos, insectos, levaduras, bacterias u otras células bajo el control de promotores apropiados. Los sistemas de traducción libres de células también pueden ser empleados para producir tales proteínas, usando ARNs derivados de constructos de ADN del presente invento.

[000182] E. coli es un huésped procariótico particularmente útil para la clonación de los polinucleótidos de la presente invención. Otros huéspedes microbianos adecuados para utilizar incluyen Bacillus subtilus, Salmonella typhimurium, y varias especies de Serratia, Pseudomonas, Streptococcus y Staphylococcus, aunque otros también pueden ser empleados como una cuestión de elección. En estos huéspedes procarióticos, también se pueden hacer vectores de expresión, que normalmente contendrán las secuencias de control de expresión compatibles con la célula huésped (por ejemplo, un origen de replicación). Además, cualquiera de una serie de promotores bien conocidos pueden estar presentes, tales como el sistema promotor de lactosa, un sistema promotor de triptófano (trp), un sistema promotor de la beta-lactamasa, o un sistema promotor a partir de fagos lambda o T7. Los promotores normalmente controlan la expresión, opcionalmente con una secuencia operadora, y tienen secuencias de sitios de unión de ribosomales y similares, para iniciar y completar la transcripción y traducción.

[000183] Un experto en el estado de la técnica de la expresión de proteínas reconocerá que la secuencia de metionina o metionina-arginina puede ser introducida en el extremo N de la secuencia madura (SEQ ID NO: 3) para la expresión en E. coli y se contemplan dentro del contexto de esta invención. Por lo tanto, a menos que se indique lo contrario, las proteínas de la presente invención expresados en E. coli tienen una secuencia de metionina introducida en el extremo N.

[000184] Otros microbios, como las levaduras o los hongos, también pueden ser utilizados para la expresión. Pichia pastoris,Saccharomyces cerevisiae, Schizosaccharomyces pombe, y Pichia angusta son ejemplos de los huéspedes de levadura preferidos, con vectores adecuados que tienen secuencias de control de expresión, tales como promotores, incluyendo 3-fosfoglicerato quinasa u otras enzimas glicolíticas, y un origen de replicación, secuencias de terminación y similares como se desee. Aspergillus niger, Trichoderma reesei y Schizophyllum commune, son ejemplos de los huéspedes de hongos, aunque otros también pueden ser empleados como una cuestión de elección.

[000185] El cultivo de células de tejido mamífero también puede utilizarse para expresar y producir los polipéptidos de la presente invención. Las células eucarioticas son en realidad preferidas, debido a que un número de líneas celulares hospederas adecuadas capaces de secretar variantes intactos se han desarrollado en el estado de la técnica, e incluyen las líneas de células CHO, varias líneas de células COS, células NSO, líneas celulares de ovario de hámster sirio, células HeLa, o líneas celulares de riñón embriónico humano (es decir, HEK293, HEK293EBNA).

[000186] Los vectores de expresión para estas células pueden incluir secuencias de control de la expresión, tales como un origen de replicación, un promotor, un potenciador, y sitios de procesamiento de información necesarios, tales como sitios de unión del ribosomas, sitios de corte y empalme de ARN, sitios de poliadenilación, secuencias terminales de transcripción. Las secuencias de control de expresión son promotores derivados de SV40, adenovirus, virus del papiloma bovino, citomegalovirus, virus del sarcoma de Raus, y similares. Los sitios preferidos de poliadenilación incluyen secuencias derivadas de la hormona de crecimiento bovino y SV40.

[000187] Los vectores que contienen las secuencias polinucleótidas de interés (por ejemplo, las proteínas de la invención y las secuencias de control de expresión) pueden transferirse a la célula huésped mediante métodos bien conocidos, los cuales varían dependiendo del tipo de huésped celular. Por ejemplo, la transfección de cloruro de

calcio se utiliza comúnmente para las células procariotas, mientras que el tratamiento con fosfato de calcio o electroporación puede utilizarse para otros huéspedes celulares.

[000188] Varios métodos de purificación de proteínas pueden emplearse y tales métodos son conocidos en el estado de la técnica y se describen, por ejemplo, en Deutscher, Methods in Enzymology 182: 83-9 (1990) y Scopes, Protein Purification: Principles and Practice, Springer-Verlag, NY (1982). Los pasos de purificación seleccionadas dependerán, por ejemplo, de la naturaleza de los procesos de producción utilizados para las proteínas de la invención.

[000189] Las proteínas, polipéptidos y/o péptidos de la invención, por ejemplo, las proteínas de fusión de actividad dual de la invención, deben formularse y dosificarse de una forma consistente con la buena práctica médica, teniendo en cuenta el estado clínico del paciente, el sitio de suministro de las composiciones de proteína, el método de administración, el programa de administración y otros factores conocidos por los médicos. La "cantidad terapéuticamente efectiva" de las proteínas de la invención para propósitos en la presente está por lo tanto determinada por tales consideraciones.

[000190] Las composiciones farmacéuticas de las proteínas de la presente invención se pueden administrar por cualquier medio que alcance el objetivo general previsto: para tratar diabetes mellitus tipo 1 y tipo 2, obesidad, síndrome metabólico, o pacientes en estado crítico. Medios permisibles no limitativos de administración incluyen, por ejemplo, por inhalación o supositorio o al tejido mucosal tal como por lavado a tejido vaginal, rectal, uretral, bucal y sublingual, de manera oral, nasal, tópica, intranasal, intraperitonea, parenteral, intravenosa, intramuscular, intraesternal, mediante inyección intraarticular, de manera intralinfática, intersticial, intraarterial, subcutánea, intrasinovial, transepitelial y transdérmica. En algunas realizaciones, las composiciones farmacéuticas se administran mediante lavado, por vía oral o inter-arterial. Otros métodos adecuados de introducción también pueden incluir dispositivos recargables o biodegradables o dispositivos poliméricos de liberación lenta o sostenida. Las composiciones farmacéuticas de esta invención también pueden administrarse como parte de una terapia combinatoria con otros agentes metabólicos conocidos.

[000191] La dosificación administrada dependerá de la edad, la salud y peso del receptor, el tipo de tratamiento concurrente, en su caso, la frecuencia del tratamiento, y la naturaleza del efecto deseado. Composiciones en el ámbito del invento incluyen todas las composiciones en las que un variante del FGF21 está presente en una cantidad que es eficaz para lograr el efecto médico deseado para el tratamiento de diabetes mellitus tipo 1

o tipo 2, obesidad o síndrome metabólico. Si bien las necesidades individuales pueden variar de un paciente a otro, la determinación de los rangos óptimos de cantidades efectivas de todos los componentes se encuentra dentro de la capacidad del médico experto.

[000192] Las proteínas de la presente invención pueden formularse de acuerdo con los métodos conocidos para preparar composiciones farmacéuticamente útiles. Una formulación deseada sería una que es un producto estable liofilizado que se reconstituye con un diluyente apropiado o una solución acuosa de alta pureza con vehículos farmacéuticamente aceptables opcionales, conservantes, excipientes o estabilizadores [Remington Pharmaceutical Sciences, 16 ª edición (1980)]. Las proteínas de la presente invención se pueden combinar con un buffer farmacéuticamente aceptable, y el pH se puede ajustar para proporcionar estabilidad aceptable, y un pH aceptable para administración.

[000193] Para administración parenteral, en una realización, las proteínas de la invención se formulan generalmente mediante la mezcla de una o más de ellas en el grado de pureza deseado, en una forma inyectable de dosificación unitaria (solución, suspensión o emulsión), con un vehículo farmacéuticamente aceptable, es decir, uno que no sea tóxico para los receptores en dosis y concentraciones empleadas y es compatible con otros ingredientes de la formulación. Preferentemente, uno o más agentes anti-microbianos farmacéuticamente aceptables se pueden añadir. Fenol, m-cresol y alcohol bencílico son los agentes anti-microbianos farmacéuticamente aceptables preferidos.

[000194] Opcionalmente, una o más sales farmacéuticamente aceptables se pueden añadir para ajustar la fuerza iónica o la tonicidad. Uno o más excipientes se pueden añadir para ajustar aún más la isotonicidad de la formulación. Glicerina, cloruro de sodio y manitol, son ejemplos de un excipiente de ajuste de isotonicidad.

[000195] Los expertos en el estado de la técnica fácilmente pueden optimizar las dosificaciones farmacéuticamente eficaces y pautas de administración de composiciones terapéuticas que comprenden las Proteínas de la invención, según lo determinado por la buena práctica médica y la condición clínica de cada paciente. Un intervalo de dosis típico para las proteínas de la presente invención variará de aproximadamente 0,01 mg por día a 1000 mg por día (o alrededor de 0,05 mg por semana a unos 5000 mg por semana administrada una vez a la semana) para un adulto. Preferiblemente, la dosis varía de 0,1 mg por día a 100 mg por día (o 0,5 mg por semana a 500 mg por semana administrada una vez por semana), más preferiblemente de aproximadamente 1,0 mg/día

a 10 mg/día (o alrededor de 5 mg por semana a 50 mg por semana administrados una vez por semana). Más preferiblemente, la dosis es de 1-5 mg/día (aproximadamente 5 mg por semana a 25 mg por semana administrados una vez por semana). La dosis adecuada de un variante del FGF21 administrado se traducirá en reducción de los niveles de glucosa en sangre y el aumento del gasto energético por la utilización de glucosa más rápido y eficiente, y por lo tanto es útil para el tratamiento de diabetes mellitus tipo 1 y tipo 2, obesidad y síndrome metabólico.

[000196] Además, debido a que la hiperglucemia y la resistencia a la insulina son comunes en pacientes críticamente enfermos dado el soporte nutricional, algunas unidades de cuidados intensivos (UCIs) administran insulina para el tratamiento de la hiperglucemia excesiva en pacientes alimentados y en estado crítico. De hecho, estudios recientes documentan que el uso de insulina exógena para mantener la glucosa sanguínea en un nivel no superior a 110 mg por decilitro, redujo la morbilidad y mortalidad entre los pacientes críticamente enfermos en la unidad de cuidados intensivos quirúrgicos, independientemente de si tenían antecedentes de diabetes (Van den Berghe N Engl J Med., 345(19):1359, (2001)). Por lo tanto, las proteínas de la presente invención son las más apropiadas para ayudar a restaurar la estabilidad metabólica en pacientes metabólicamente inestables en estado crítico. Las proteínas de la invención tales como aquellas que contienen variantes de FGF21 son únicas puesto que estimulan la absorción de glucosa y aumentan la sensibilidad a la insulina pero no inducen hipoglucemia.

[000197] En otro aspecto de la presente invención, se contemplan las proteínas de la invención para uso como un medicamento para el tratamiento de la obesidad, diabetes mellitus tipo 1 y tipo 2, pancreatitis, dislipidemia, enfermedad de hígado graso no alcohólico (NAFLD), esteatohepatitis no alcohólica (EHNA), resistencia a la insulina, hiperinsulinemia, intolerancia a la glucosa, hiperglucemia, síndrome metabólico, infarto agudo de miocardio, afecciones asociadas con severas mutaciones inactivadoras en el receptor de la insulina, lipodistrofias incluyendo lipodistrofia asociada con VIH y otros trastornos metabólicos.

[000198] Habiendo descrito el presente invento en detalle, el mismo se entiende con mayor claridad en relación con los siguientes ejemplos, que se incluyen para fines de ilustración solamente y no pretenden ser limitativos del invento.

[000199] La práctica del presente invento empleará, a menos que se indique lo contrario, métodos convencionales de química, bioquímica, biología molecular, inmunología y farmacología, en el conocimiento del estado de la técnica. Tales técnicas

se explican con detalle en la literatura. Véase, por ejemplo, Remington 's Pharmaceutical Sciences, 18 ª edición (Easton, Pennsylvania: Mack Publishing Company, 1990), Methods In Enzymology (Colowick S. y N. Kaplan, editores, Academic Press, Inc.), y Handbook of Experimental Immunology, Vols. IIV (DM Weir y CC Blackwell, eds, 1986, Blackwell Scientific Publications.), Y Sambrook, Molecular Cloning: A Laboratory Manual (2 ª edición, 1989).

Mutantes FGF21 Específicos a Sitio

[000200] En algunas realizaciones, las proteínas de fusión de la invención incluyen mutantes adicionales de GLP-1 o Exendina-4, péptidos híbridos de GLP-1/Glucagón, análogos de GLP-1 con aminoácidos no naturales (para transmitir resistencia a DPP-4, para la PEGilación, o para otros fines).

[000201] En algunas realizaciones, las proteínas de la invención comprenden agonistas de FGF21 con una o más de las siguientes modificaciones adicionales de FGF21 tipo silvestre:

[000202] (i) disulfuros adicionales, aminoácidos no naturales, o modificaciones para promover la dimerización tal como la formación de un disulfuro en R154C o la introducción de una cisteína en otro sitio, o la dimerización a través de un dominio Fc fusionado, o la formación de dímeros a través de un entrecruzador, tal como un PEG bifuncional;

[000203] (ii) fragmentos de FGF21;

[000204] (iii) proteínas seleccionadas por tener actividad FGF21 (unión a beta-Klotho y unión y activación de las FGFRs); y

[000205] (iv) un anticuerpo FGF21 mimético (de diversos formatos tales como Fab, UniBody®, svFc etc.).

[000206] En algunas realizaciones, las proteínas de actividad dual de la invención comprenden uno o más de los siguientes enlazadores: un enlace de amida simple, péptidos cortos (particularmente repeticiones Ser/Gly), residuos adicionales desde la secuencia traducida FGF21, o un enlazador más grande hasta una proteína entera (tal como un dominio Fc, un fardo helicoidal fijador de HSA, HSA, etc.). Las dos fracciones también pueden ser unidas por medios químicos, tales como a través de aminoácidos no naturales o enlaces químicos estándar (maleimida-Cys, NHS-Lys, clic, etc.). "Enlazador" para el enfoque de anticuerpos miméticos de FGF21 pueden incluir los ya enumerados y también una inserción en un bucle con la posterior escisión para liberar el GLP-1 N-terminal.

[000207] En algunas realizaciones, la proteína de fusión de la invención comprende PEGilación que ocurre en uno, dos o más sitios específicos. En realizaciones preferidas, la PEGilación se produce dentro de la molécula de FGF21 o el enlazador. En algunas realizaciones, la PEGilación no está dentro de ~ 10 aminoácidos del N-término de la proteína de función dual, preferiblemente no dentro de ~ 10 aminoácidos del comienzo de FGF21. Las químicas de enlace de PEGilación pueden incluir NHS-Lys, maleimida-Cys, aminoácidos no naturales (aldehído, "química click", Pcl, etc.) y se pueden combinar con variantes de proteínas adecuadas para controlar la estequiometría de la reacción.

[000208] El grupo PEG de las proteínas de fusión de la invención pueden ser de cualquier tamaño (por ejemplo, 1, 2, 3,4, 5, 10, 20, 24, 29, 30, 40 kDa), y pueden tener estructura lineal, ramificada o tipo peine, con una preferencia por una PEGilación total mayor o igual a 40 kDa. La PEGilación óptima alcanza una extensión de vida media suficiente para dosificación una vez por semana. La PEGilación de dímeros, trímeros, tetrámeros, etc. de proteínas puede resultar en una extensión de vida media en suero adecuada usando polímeros PEG más cortos. Estructuras PEG ramificadas y combinadas pueden ser benéficas en términos de más baja viscosidad.

[000209] Metodologías preferidas para extensión de vida media para las proteínas de fusión de la invención incluyen la integración de un dominio Fc de IgG o HSA en el enlazador y pueden no requerir PEGilación. Además, el uso de fusiones de dominio Fc resultará en la dimerización y puede dar lugar a mayor potencia, además de la extensión de vida media.

[000210] Otras realizaciones de la invención incluyen, pero no se limitan a los siguientes enlaces, para la extensión de vida media: unión de HSA a lípidos o moléculas pequeñas o micelios a bien sea la versión monomérica o la versión dimérica de la fusión.

[000211] En ciertas realizaciones de la invención, otros enlaces se pueden hacer a las proteínas, polipéptidos, y/o péptidos de la invención, para lograr extensión de la vida media y otras propiedades biológicas mejoradas. Estas pueden incluir fijación de conjugados de PEG-colesterol (incluyendo micelios y liposomas) a las proteínas, polipéptidos, y/o péptidos de la invención, y/o fijación de azúcares (glicosilato) a las proteínas, polipéptidos, y/o péptidos de la invención. En aún otras realizaciones, se emplean técnicas similares para añadir conjugados de, por ejemplo, ácido polisiálico (PSA), hidroxietil almidón (HES), ligandos de unión de albúmina, o escudos de carbohidratos a proteínas, polipéptidos, y/o péptidos.

[000212] La técnica de hesilación, por ejemplo, acopla cadenas de hidroxietilalmidón (HES) ramificado (60 kDa o 100 kDa, fragmentos de la amilopectina altamente ramificados de almidón de maíz) a una proteína, polipéptidos, y/o péptidos a través de alquilación reductora. Conjugados por polisialilación de proteínas, polipéptidos y/o péptidos de interés con polímeros de ácido polisiálico de manera similar a la PEGilación. Los polímeros de PSA son polímeros inmunogénicos con carga negativa que se producen naturalmente en el cuerpo y están disponibles en pesos moleculares de 10-50kD.

[000213] En aún otras realizaciones de la invención, otros enlaces o modificaciones se pueden hacer a las proteínas, polipéptidos, y/o péptidos de la invención, para lograr la vida media extendida y otras propiedades biológicas mejoradas. Estas incluyen la creación de grupos PEG recombinantes (rPEG), y su fijación a las proteínas, polipéptidos, y/o péptidos de la invención. Desarrollada por la empresa Amunix, Inc., tecnología rPEG se basa en secuencias de proteínas con propiedades tipo PEG que son genéticamente fusionadas a los productos biofarmacéuticos, evitando el paso de la conjugación química adicional. Los rPEGs son constructos de exenatida de vida media extendida que contienen una larga cola no estructurada de aminoácidos hidrófilos, y que son capaces de aumentar la vida media sérica de la proteína o péptido y retardar su tasa de absorción, reduciendo así la relación pico-valle significativamente. rPEGs tienen un radio hidrodinámico mayor y muestran un peso molecular aparente que es aproximadamente 15 veces su peso molecular actual, imitando la forma en que la pegilación logra una larga vida media sérica. Secuencias polipeptídicas recombinantes similares también son desarrolladas por XL-proteína GmbH para la "PASilación" de proteínas.

Mutantes de Proteínas de Función Dual Químicamente Modificadas

[000214] Formas químicamente modificadas de las proteínas de fusión descritas en este documento, incluyendo, ej., formas truncadas o variantes de las proteínas de fusión dual descritas en la presente, pueden ser preparadas por un experto en el estado de la técnica, dadas las divulgaciones descritas en la presente. Tales Proteínas de función dual químicamente modificadas se alteran de tal manera que el mutante modificado químicamente se difiere del mutante no modificado, ya sea en el tipo o la localización de las moléculas fijadas de manera natural al mutante. Los mutantes químicamente modificados pueden incluir moléculas formadas por la supresión de uno o más grupos químicos naturalmente fijados.

[000215] En una realización, las proteínas de la presente invención pueden modificarse por la unión covalente de uno o más polímeros. Por ejemplo, el polímero seleccionado es típicamente soluble en agua para que la proteína a la que se fija no se precipite en un medio acuoso, tal como un medio fisiológico. Incluidos en el ámbito de los polímeros adecuados está una mezcla de polímeros. Preferentemente, para el uso terapéutico de la preparación del producto final, el polímero será farmacéuticamente aceptable. Polímeros no solubles en agua conjugados a proteínas de la presente invento también constituyen un aspecto de la invención.

[000216] Los polímeros de ejemplo cada uno puede ser de cualquier peso molecular y pueden ser ramificados o no. Los polímeros cada uno suelen tener un peso molecular medio de entre aproximadamente 2 kDa a 100 kDa (el término "aproximadamente" que indica que en las preparaciones de un polímero soluble en agua, algunas moléculas pesarán más y algunas menos que el peso molecular indicado). El peso molecular medio de cada polímero es preferiblemente entre aproximadamente 5 kDa y 50kDa, más preferiblemente entre aproximadamente 12 kDa y aproximadamente 40 kDa, y más preferiblemente entre aproximadamente 20 kDa y alrededor de 35 kDa.

[000217] Polímeros solubles en agua apropiados o mezclas de los mismos incluyen, pero no están limitados a, carbohidratos N-enlazados o O enlazados, azúcares, fosfatos, glicol polietileno (PEG) (incluyendo las formas de PEG que se han utilizado para derivatizar proteínas, incluyendo mono - (C1-C10), alcoxi, ariloxi-polietilenglicol), monometoxi-polietilenglicol, dextrano (por ejemplo, dextrano de bajo peso molecular, por ejemplo, alrededor de 6 kD), celulosa, y otros polímeros basados en carbohidratos, poli-(N-vinil pirrolidona) polietilenglicol, homopolímeros de glicol de propileno, óxido de propileno/co-polímeros de óxido de etileno, polioles polioxietilados (por ejemplo, glicerol), y alcohol polivinílico. El presente invento también contempla moléculas bifuncionales de entrecruzamiento que se pueden utilizar para preparar multímeros variantes proteicos de función dual fijados de manera covalente. El presente invento también contempla variantes de proteínas de función dual covalentemente fijados al ácido polisiálico.

[000218] En algunas realizaciones de la presente invención, una variante de proteína de función dual está covalentemente, o químicamente, modificado para incluir uno o más polímeros solubles en agua, incluyendo pero sin limitar a, polietilenglicol (PEG), polioxietilenglicol, o glicol de propileno. Véase, por ejemplo, Patentes EEUU N º

4.640.835, 4.496.689, 4.301.144, 4.670.417, 4.791.192, y 4.179.337. En algunas realizaciones del presente invento, una proteína de función dual comprende uno o más polímeros, incluyendo pero sin limitar a, monometoxi-polietilenglicol, dextrano, celulosa, otro polímero a base de carbohidratos, poli-(N-vinil pirrolidona)-polietilenglicol, homopolímeros de glicol de propileno, un óxido de propileno/co-polímero de óxido de etileno, polioles polioxietilados (por ejemplo, glicerol), alcohol polivinílico, o mezclas de estos polímeros.

[000219] En algunas realizaciones del presente invento, una proteína de función dual es covalentemente modificada con subunidades PEG. En algunas realizaciones, uno o más polímeros solubles en agua se unen en una o más posiciones específicas (por ejemplo, en el extremo N-terminal) del mutante de proteína de función dual. En algunas realizaciones, uno o más polímeros solubles en agua son aleatoriamente unidos a una o más cadenas laterales de un mutante de proteína de función dual. En algunas realizaciones, el PEG se utiliza para mejorar la capacidad terapéutica de un mutante de proteína de función dual. Ciertos métodos se discuten, por ejemplo, en EE.UU. Pat. N º 6.133.426, que se incorpora como referencia para cualquier propósito.

[000220] En las realizaciones del presente invento en donde el polímero es PEG, el grupo PEG pueden ser de cualquier peso molecular conveniente, y puede ser lineal o ramificado. El peso molecular promedio del grupo PEG preferiblemente variará desde aproximadamente 2 kDa a aproximadamente 100 kDa, y más preferiblemente de aproximadamente 5 kDa a aproximadamente 50 kDa, por ejemplo, 10, 20, 30, 40 o 50 kDa. Los grupos PEG generalmente se fijarán al mutante de proteína de función dual a través de acilación o alquilación reductora, a través de un grupo reactivo en la fracción PEG (por ejemplo, un aldehído, amino, tiol o grupo éster) a un grupo reactivo en el mutante de proteína de función dual (por ejemplo, un aldehído, amino, o un grupo éster).

[000221] Derivados de PEG ramificado, también conocidos como derivados PEG "en forma de Y", contienen dos cadenas lineales metoxi -PEG unidas a un núcleo central. La estructura estéricamente voluminosa de estos derivados PEG "en forma de Y" facilitará la unión en un solo punto de las moléculas modificadas. A modo de ejemplo, tres tipos de derivados PEG "en forma de Y" son Y-NHS-40K (útil para PEGilación de amina), Y-MAL-40K (útil para PEGilación de tiol), y Y-ALD-40K (por ejemplo, Y -AALD-40K y Y- PALD-40K) (útil para PEGilación N-terminal). Para PEGilación de amina, el éster NHS en "forma de Y" reaccionará con el grupo amino de lisina(s) o la amina N-terminal en moléculas biológicas activas para producir uno o varios enlaces de amida estables.

Este ester NHS se acoplará con las moléculas sobre las cuales se puede actuar a pH 7-8.5. Para PEGilación de tiol, la maleimida en "forma de Y" reaccionará con los grupos tiol en moléculas biológicas activas para generar un enlace estable de 3-tiosuccinimidil éter. Esta maleimida se acoplará con las moléculas objetivo a un pH de 7.4 en presencia de otros grupos funcionales. Para PEGilación N-terminal, el aldehído "en forma de Y" preferiblemente reaccionará con la amina N-terminal en moléculas biológicas activas para producir un enlace estable de amina en presencia de un reactivo reductor tal como cianoborohidruro de sodio. Este aldehído se acoplará con la amina N-terminal de las moléculas objetivo a pH 5.8. Los reactivos para la realización de PEGilación ramificada están disponibles a través de, por ejemplo JenKem Technology.

[000222] La PEGilación de un polipéptido, incluyendo las proteínas de la invención, puede ser específicamente realizada mediante cualquiera de las reacciones de PEGilación conocidas en el estado de la técnica. Estas reacciones se describen, por ejemplo, en las siguientes referencias: Francis , 1992, Focus on Growth Factors 3: 4-10; Patentes Europeas No. 0 154 316 y 0 401384 , y Patente EE.UU. N º 4.179.337. Por ejemplo, la PEGilación puede llevarse a cabo a través de una reacción de acilación o una reacción de alquilación con una molécula de polietilenglicol reactivo (o un polímero soluble en agua reactivo análogo) como se describe en la presente. Para las reacciones de acilación, un polímero seleccionado debe tener un único grupo éster reactivo. Para la alquilación reductiva, un polímero seleccionado debe tener un único grupo aldehído reactivo. Un aldehído reactivo, por ejemplo, propionaldehído de polietilenglicol, que es estable en agua, o derivados mono C1-C10 alcoxilo o ariloxilo de los mismos (véase, por ejemplo, Pat. EE.UU N º 5.252.714).

[000223] En algunas realizaciones del presente invento, una estrategia útil para la unión del grupo PEG a un polipéptido involucra combinar, a través de la formación de un enlace de conjugado en solución, un péptido y una fracción PEG, cada uno con una funcionalidad especial que es mutuamente reactiva hacia el otro. Los péptidos se pueden preparar fácilmente con la síntesis de fase sólida convencional. Los péptidos son "pre-activado" con un grupo funcional apropiado en un sitio específico. Los precursores son purificados y completamente caracterizados antes de reaccionar con la fracción PEG. La ligadura del péptido con PEG usualmente se lleva a cabo en la fase acuosa y puede ser fácilmente monitoreada por HPLC analítica de fase inversa. Los péptidos PEGilados pueden ser fácilmente purificados mediante HPLC preparativa y se caracterizan por

HPLC analítica, análisis de aminoácidos y espectrometría de masa con desorción inducida por láser.

[000224] Los polímeros polisacáridos son otro tipo de polímero soluble en agua que puede utilizarse para la modificación de proteínas. Por lo tanto, las proteínas de la presente invención fusionadas a un polímero polisacárido forman realizaciones de la presente invención. Los dextranos son polímeros de polisacaridos compuestos de subunidades individuales de glucosa predominantemente enlazadas por 1-6 enlaces alfa. El dextrano por si mismo está disponible en muchos rangos de peso molecular, y está fácilmente disponible en pesos moleculares entre 1 kD y 70 kD aproximadamente. El dextrano es un polímero soluble en agua adecuado para utilizar como un vehículo por sí solo o en combinación con otro vehículo (por ejemplo, Fc). Véase, por ejemplo, la Publicación Internacional N º WO 96/11953. El uso de dextrano conjugado con inmunoglobulinas terapéuticas o diagnósticas, se ha informado. Véase, por ejemplo, la Publicación de Patente Europea No. 0 315 456, la cual se incorpora en la presente como referencia. El presente invento también contempla el uso de dextrano de alrededor de 1 kD a alrededor de 20 kD.

[000225] En general, la modificación química se puede realizar bajo cualquier condición adecuada utilizada para reaccionar una proteína con una molécula polimérica activada. Los métodos para preparar polipéptidos modificados químicamente por lo general comprenden las etapas de: (a) hacer reaccionar el polipéptido con la molécula polimérica activada (tal como un éster reactivo o un derivado aldehído de la molécula polimérica), bajo condiciones en las que un variante proteico de FGF21 resulta unido a una o más moléculas poliméricas, y (b) la obtención de los productos de reacción. Las condiciones óptimas de reacción se determinan en función de parámetros conocidos y el resultado deseado. Por ejemplo, cuanto mayor sea la proporción de las moléculas poliméricas a la proteína, mayor es el porcentaje de molécula polimérica fijada. En una realización del presente invento, los mutantes FGF21 químicamente modificados pueden tener una sola fracción de molécula polimérica en el extremo amino (véase, por ejemplo, Pat. EE.UU N º 5234784)

[000226] En otra realización de la presente invención, las proteínas de la invención pueden ser químicamente acopladas a biotina. Las proteínas biotiniladas de la invención luego se dejan fijar a avidina, dando como resultado avidina/biotina/proteínas tetravalentes de la invención. Las proteínas de la invención también pueden ser covalentemente acopladas a dinitrofenol (DNP) o trinitrofenol (TNP) y los conjugados

resultantes precipitados con anti-DNP o anti-TNP-IgM para formar conjugados decaméricos con una valencia de 10.

[000227] En general, las condiciones que pueden ser aliviadas o moduladas por la administración de los presentes mutantes de proteínas de función dual modificados químicamente incluyen aquellas descritas en este documento para proteínas de la invención. Sin embargo, los mutantes de proteínas de función dual químicamente modificados divulgados en la presente pueden tener actividades adicionales, actividad biológica aumentada o reducida, u otras características, tales como vida media aumentada o disminuida, en comparación con mutantes de proteínas de función no modificados.

Composiciones Terapéuticas de Proteínas de Función Dual y Su Administración

[000228] Las composiciones terapéuticas que comprenden las proteínas de función dual de la invención están dentro del alcance de la presente invención, y se contemplan específicamente a la luz de la identificación de varias secuencias de proteínas de función dual que exhiben propiedades mejoradas. Tales composiciones farmacéuticas de mutantes de proteínas de función dual pueden comprender una cantidad terapéuticamente efectiva de un variante proteico de proteínas de función dual en mezcla con un agente de la formulación farmacéutica o fisiológicamente aceptable seleccionado para determinar su idoneidad con el modo de administración.

[000229] Los materiales de la formulación aceptable preferiblemente no son tóxicos para los receptores en las dosis y concentraciones empleadas.

[000230] La composición farmacéutica puede contener materiales de formulación para modificar, mantener o preservar, por ejemplo, el pH, la osmolaridad, la viscosidad, la claridad, el color, la isotonicidad, el olor, la esterilidad, estabilidad, velocidad de disolución o liberación, absorción, o penetración de la composición. Materiales adecuados de formulación incluyen, pero no se limitan a, aminoácidos (tales como la glicina, glutamina, asparagina, arginina, o lisina), antimicrobianos, antioxidantes (tales como el ácido ascórbico, sulfito de sodio, o hidrógeno-sulfito de sodio), buffers (tales como borato, bicarbonato, Tris-HCl, citratos, fosfatos u otros ácidos orgánicos), agentes de dilución (como manitol o glicina), agentes quelantes (como el ácido etilendiaminotetraacético (EDTA)), agentes formadores de complejos (como la cafeína, la polivinilpirrolidona, beta-ciclodextrina o hidroxipropil-beta-ciclodextrina), rellenos, monosacáridos, disacáridos y otros hidratos de carbono (como la glucosa, manosa, o

dextrinas), proteínas (como la albúmina sérica, gelatina o inmunoglobulinas), agentes colorantes, aromatizantes y de dilución, agentes emulsionantes, polímeros hidrofílicos (tales como polivinilpirrolidona), polipéptidos de bajo peso molecular, contraiones que forman sales (como el sodio), conservantes (como el cloruro de benzalconio, ácido benzoico, ácido salicílico, el timerosal, alcohol fenetílico, metilparabeno, propilparabeno, clorhexidina, ácido sórbico, o peróxido de hidrógeno), solventes (como glicerina, glicol de propileno, o polietilenglicol), alcoholes de azúcar (como manitol o sorbitol), agentes de suspensión, surfactantes o agentes humectantes (como Pluronics; PEG; ésteres de sorbitan, polisorbatos tales como polisorbato 20 o polisorbato 80; tritón, trometamina, lecitina, colesterol o tiloxapal), agentes mejoradores de estabilidad (por ejemplo, sacarosa o sorbitol), agentes mejoradores de la tonicidad (por ejemplo, haluros de metales alcalinos, preferiblemente cloruro de sodio o potasio; sorbitol manitol), vehículos de liberación, diluyentes, excipientes y/o adyuvantes farmacéuticos (véase, por ejemplo, Remington 's Pharmaceutical Sciences (18 ª ed., AR Gennaro, ed., Mack Publishing Company, 1990), y las ediciones posteriores de la misma, incorporada en la presente como referencia para cualquier propósito).

[000231] La composición farmacéutica óptima será determinada por un experto en el estado de la técnica dependiendo de, por ejemplo, la vía de administración, formato de liberación, y la dosis deseada (véase, por ejemplo, Remington's Pharmaceutical Sciences, arriba). Estas composiciones pueden influir en el estado físico, estabilidad, velocidad de liberación in vivo, y la tasa de depuración in vivo de la proteína de fusión de la invención.

[000232] El principal vehículo en una composición farmacéutica puede ser de naturaleza acuosa o no acuosa. Por ejemplo, un vehículo adecuado para la inyección puede ser agua, solución salina fisiológica, o líquido cefalorraquídeo artificial, posiblemente complementado con otros materiales comunes en composiciones para administración parenteral. La solución salina amortiguada neutral o solución salina mezclada con albúmina sérica son vehículos de ejemplo adicionales. Otras composiciones farmacéuticas de ejemplo comprenden el buffer Tris con pH de alrededor de 7.0-8.5, o buffer de acetato con pH 4.0 a 5.5, que puede incluir además sorbitol o un sustituto adecuado. En una realización de la presente invención, las composiciones farmacéuticas de función dual pueden prepararse para el almacenamiento mediante la mezcla de la composición seleccionada con el grado de pureza deseado con agentes de formulación opcionales (Remington's Pharmaceutical Sciences, arriba) en la forma de

una torta liofilizada o una solución acuosa. Además, la proteína de función dual puede formularse como un liofilizado con excipientes adecuados, tales como la sacarosa.

[000233] Las composiciones farmacéuticas que contienen las proteínas de fusión de la invención pueden seleccionarse para administración parenteral. Por otra parte, las composiciones pueden seleccionarse para la inhalación o para la liberación a través del tracto digestivo, por vía oral por ejemplo. La preparación de tales composiciones farmacéuticamente aceptable está en el conocimiento del estado de la técnica.

[000234] Los componentes de la formulación están presentes en concentraciones que son aceptables para el sitio de administración. Por ejemplo, se utilizan buffers para mantener la composición a un pH fisiológico o a un pH ligeramente inferior, por lo general dentro de un rango de pH de 5 a 8.

[000235] Cuando se contempla la administración parenteral, las composiciones terapéuticas para utilizar en este invento pueden estar en la forma de una solución acuosa parenteralmente aceptable, libre de pirógenos que comprende la proteína de función dual deseada en un vehículo farmacéuticamente aceptable. Un vehículo especialmente adecuado para la inyección parenteral es agua destilada estéril en donde se formula una proteína de función dual como una solución estéril, isotónica y preservada de manera adecuada. Sin embargo, otra preparación puede involucrar la formulación de la molécula deseada con un agente, tal como microesferas inyectables, partículas bio-erosionables, compuestos poliméricos (como el ácido poliláctico o ácido poliglicólico), esferas, o liposomas, que proporciona la liberación controlada o sostenida del producto, que puede ser liberado a través de una inyección de depósito. El ácido hialurónico también puede utilizarse, y este puede tener el efecto de promover la duración sostenida en la circulación. Otros medios apropiados para la introducción de la molécula deseada incluyen los dispositivos implantables de administración de fármacos.

[000236] En una realización, una composición farmacéutica puede formularse para la inhalación. Por ejemplo, una proteína de función dual de la invención puede formularse como un polvo seco para inhalación. Las soluciones de inhalación de proteínas de función dual también se pueden formular con un propulsor para liberación en aerosol. En aún otra realización, las soluciones pueden ser nebulizadas. La administración pulmonar se describe con más detalle en la Publicación Internacional N º WO 94/20069, que describe la administración pulmonar de las proteínas modificadas químicamente.

[000237] También se contempla que ciertas formulaciones pueden ser administradas por vía oral. En una realización de la presente invención, las Proteínas de Fusión de la

invención que se administran de esta manera se pueden formular con o sin los vehículos habitualmente utilizados en la composición de formas de dosificación sólidas tales como tabletas y cápsulas. Por ejemplo, una cápsula puede ser diseñada para liberar la porción activa de la formulación en el punto en el tracto gastrointestinal cuando se maximiza la biodisponibilidad y la degradación pre-sistémica se reduce al mínimo. Agentes adicionales pueden ser incluidos para facilitar la absorción de las proteínas de función dual de la invención. Diluyentes, aromatizantes, ceras de bajo punto de fusión, aceites vegetales, lubricantes, agentes de suspensión, agentes desintegrantes de tabletas, y aglutinantes también se pueden emplear.

[000238] Otra composición farmacéutica puede implicar una cantidad eficaz de las proteínas de fusión de la invención en una mezcla con excipientes no tóxicos que son adecuados para la fabricación de comprimidos. Mediante la disolución de las tabletas en agua estéril u otro medio adecuado, las soluciones pueden ser preparadas en forma de dosis unitarias. Excipientes adecuados incluyen, pero no se limitan a, diluyentes inertes, tales como carbonato de calcio, carbonato de sodio o bicarbonato, lactosa o fosfato de calcio, o aglomerantes, tales como almidón, gelatina o goma arábiga; o agentes lubricantes tales como estearato de magnesio, ácido esteárico, o talco.

[000239] Composiciones farmacéuticas adicionales que comprenden Proteínas de Función dual de la invención será evidentes para los expertos en la técnica, incluyendo formulaciones que implican Proteínas de Función dual de la invención en formulaciones de liberación sostenida o controlada. Técnicas para la formulación de una variedad de medios de liberación sostenida o controlada, tales como vehículos de liposoma, micropartículas bio-erosionables o esferas porosas e inyecciones de depósito, también son conocidos por los expertos en el estado de la técnica (véase, por ejemplo, la Publicación Internacional N º WO.93/15722, que describe la liberación controlada de micropartículas poliméricas porosas para la liberación de composiciones farmacéuticas, y Wischke & Schwendeman, 2008, Int J Pharm 364: 298-327, y Freiberg y Zhu, 2004, Int. J. Pharm. 1-18, donde se trata la preparación y uso de microesferas/micropartículas). J Pharm. 364: 298-327, y Freiberg & Zhu, 2004, Int. J Pharm. 282: 1-18, que discuten la preparación y uso de microesferas/microparticulas).

[000240] Otros ejemplos de preparaciones de liberación sostenida incluyen matrices poliméricas semipermeables en forma de articulos formados, por ejemplo, películas, o microcápsulas. Las matrices de liberación sostenida pueden incluir poliésteres, hidrogeles, polilactidas (Pat. EE.UU. No. 3,773,919 y la patente europea No. 0 058 481),

copolímeros de ácido L-glutámico y gamma etil-L-glutamato (Sidman, 1983, 22 Biopolímeros 22: 547-56), poli(2-hidroxietil-metacrilato) (Langer , 1981, J. Biomed Mater Res 15: 167-277 y Langer, 1982, Chem Tech 12. 98-105) , vinil acetato de etileno (Langer, arriba) o ácido poli-D-3-hidroxibutírico (Patente Europea No. 0 133 988). Las composiciones de liberación sostenida pueden incluir también los liposomas, que pueden prepararse por uno de varios métodos conocidos en el estado de la técnica. Véase, por ejemplo, Epstein, 1985, Proc. Natl. Acad. Ciencia. U.S.A. 82: 3688-92, y la Patente Europea N º 0 036 676 0 088 046 y 0 949 143. 3688-92; Patentes Europeas No. 0 036 676, 0 088 046, y 0 143 949.

[000241] Las composiciones farmacéuticas de la invención que va a utilizarse para la administración in vivo normalmente deben ser estériles. Esto se puede lograr mediante filtración a través de membranas de filtración estériles. Donde la composición es liofilizada, la esterilización que utiliza este método se puede realizar ya sea antes o después de liofilización y reconstitución. La composición para la administración parenteral puede ser almacenada en forma liofilizada o en solución. Además, las composiciones parenterales generalmente se colocan en un recipiente que tenga un puerto de acceso estéril, por ejemplo, una bolsa de solución intravenosa o vial que tiene un tapón bloqueador que pueda ser perforable por una aguja de inyección hipodérmica.

[000242] Una vez que la composición farmacéutica ha sido formulada, puede ser almacenada en viales estériles como una solución, suspensión, gel, emulsión, sólido o en forma de polvo deshidratado o liofilizado. Estas formulaciones se pueden almacenar en una forma fácil de usar, o en una forma (por ejemplo, liofilizada) que requiere de una reconstitución antes de su administración.

[000243] En una realización concreta, el presente invento se refiere a los kits para la producción de una unidad de administración de dosis única. Los kits pueden contener tanto un primer recipiente que tiene una proteína seca y un segundo recipiente con una formulación acuosa. También se incluyen dentro del alcance de este invento los kits que contienen una o varias jeringas pre-cargadas de cámara única o múltiple (por ejemplo, jeringas líquidas o liojeringas).

Dosificaciones de Proteínas de Función Dual y Su Administración

[000244] La cantidad efectiva de una composición farmacéutica de la invención a ser empleada terapéuticamente dependerá, por ejemplo, del contexto terapéutico y objetivos. Un experto en el estado de la técnica apreciará que los niveles de dosificación apropiados

para el tratamiento por lo tanto pueden variar dependiendo, en parte, de la molécula liberada, la indicación para la cual el variante de proteína de fusión es utilizado, la vía de administración, y el tamaño (peso corporal, la superficie del cuerpo, o el tamaño del órgano) y la condición (edad y salud general) del paciente. En consecuencia, el médico puede titular la dosis y modificar la vía de administración para obtener el efecto terapéutico óptimo. Una dosis típica puede oscilar entre 0,1 µg/kg hasta aproximadamente100 mg/kg o más, dependiendo de los factores mencionados anteriormente. En otras realizaciones, la dosificación pueden oscilar entre 0.1 µg/kg y 100 mg/kg aproximadamente; o 1 µg/kg hasta 100 mg/kg.

[000245] La frecuencia de dosificación dependerá de los parámetros farmacocinéticos de la proteína de función dual en la formulación que se utiliza. Por lo general, un médico administrará la composición hasta que se alcance una dosificación para lograr el efecto deseado. La composición por lo tanto, se puede administrar como una sola dosis, como dos o más dosis (que pueden o no contener la misma cantidad de la molécula deseada) a través del tiempo, o como una infusión continua a través de un dispositivo de implante o catéter. El refinamiento adicional de la dosificación adecuada es realizado de manera rutinaria por los expertos en el estado de la técnica y se encuentra dentro del ámbito de las tareas de rutina realizadas por ellos. Las dosis apropiadas se pueden determinar a través del uso de datos apropiados dosis-respuesta.

[000246] La vía de administración de la composición farmacéutica está de acuerdo con métodos conocidos, por ejemplo, por vía oral, mediante inyección por vía intravenosa, intraperitoneal, intracerebral (intraparenquimal), intracerebroventricular, intramuscular, intraocular, intraarterial, intraportal, o intralesional, por los sistemas de liberación sostenida (que también se pueden inyectar), o por los dispositivos de implantación. Si así se desea, las composiciones se pueden administrar por inyección en bolo o por infusión continua o mediante un dispositivo de implantación.

[000247] Alternativa o adicionalmente, la composición puede ser administrada localmente a través de la implantación de una membrana, una esponja, u otro material apropiado sobre el cual la molécula deseada se ha absorbido o encapsulado. Cuando un dispositivo de implante se utiliza, el dispositivo puede ser implantado en cualquier tejido u órgano adecuado, y la liberación de la molécula deseada puede ser a través de difusión, bolo de liberación programada, o la administración continua.

Usos Terapéuticos de las Proteínas de Función Dual

[000248] Las proteínas de la invención pueden utilizarse para tratar, diagnosticar, aliviar o prevenir una serie de enfermedades, trastornos o afecciones, incluyendo pero sin limitar a trastornos metabólicos. En una realización, el trastorno metábolico a tratar es diabetes, ej., diabetes mellitas tipo 2. En otra realización, el trastorno metabólico es la obesidad. Otras realizaciones incluyen condiciones o trastornos metabólicos tales como diabetes mellitus tipo 1, pancreatitis, dislipidemia, enfermedad hepática grasa no alcohólica (EHGNA), esteatohepatitis no alcohólica (EHNA), resistencia a la insulina, hiperinsulinemia, intolerancia a la glucosa, hiperglucemia , síndrome metabólico, infarto agudo del miocardio, hipertensión, enfermedades cardiovasculares, arteriosclerosis, enfermedad arterial periférica, accidente cerebro vascular, insuficiencia cardiaca, enfermedad coronaria, enfermedad renal, complicaciones de la diabetes, neuropatía, trastornos asociados con mutaciones inactivadoras severas en el receptor de insulina, lipodistrofias incluyendo lipodistrofia asociada a VIH, gastroparesia y otros trastornos metabólicos.

[000249] En la práctica, un trastorno o enfermedad, tales como diabetes mellitus tipo 1 o tipo 2 o la obesidad se pueden tratar mediante la administración de un variante proteico de función dual como se describe en la presente a un paciente que lo necesita en la cantidad de una dosis terapéuticamente eficaz. La administración puede realizarse como se describe en este documento, por ejemplo mediante inyección intravenosa, inyección intraperitoneal, intramuscular o por vía oral en forma de un comprimido o formación líquida. En la mayoría de los casos, una dosis deseada puede ser determinada por un médico, tal como se describe en este documento, y puede representar una dosis terapéuticamente eficaz del polipéptido proteico de función dual. Será evidente para los expertos en el estado de la técnica que una dosis terapéuticamente eficaz del polipéptido proteico de función dual dependerá, entre otras cosas, del esquema de administración, la dosis unitaria de antígeno administrado, si la molécula de ácido nucleico o polipéptido se administra en combinación con otros agentes terapéuticos, el estado inmune y la salud del receptor. El término "dosis terapéuticamente eficaz", como se utiliza en la presente, significa la cantidad del polipéptido proteico de función dual que provoca la respuesta biológica o medicinal en un sistema de tejido, animal o humano que busca el investigador, médico, u otro clínico, que incluye el alivio de los síntomas de la enfermedad o trastorno que se trata.

Composiciones Farmacéuticas

[000250] El presente invento también proporciona composiciones farmacéuticas que comprenden una o más de las proteínas de función dual de la invención ó mutantes descritos en la presente y un vehículo farmacéuticamente aceptable. En algunas realizaciones, las composiciones farmacéuticas se preparan como inyectables, ya sea como soluciones o suspensiones líquidas, formas sólidas adecuadas para solución en, o suspensión en, vehículos líquidos antes de la inyección también se pueden preparar. Los liposomas son incluidos en la definición de un vehículo farmacéuticamente aceptable. Sales farmacéuticamente aceptables también pueden estar presentes en la composición farmacéutica, por ejemplo, sales de ácidos minerales tales como clorhidratos, bromhidratos, fosfatos, sulfatos, etc., y las sales de ácidos orgánicos tales como acetatos, propionatos, malonatos, benzoatos, etc. Una discusión completa de excipientes farmacéuticamente aceptables está disponible en Remington: The Science and Practice of Pharmacy (1995) Alfonso Gennaro, Lippincott, Williams y Wilkins.

Proteínas de Fusión y Compuestos Pétidicos

[000251] En otra realización, las proteínas de la presente invención pueden ser hechas en una proteína de fusión o compuesto peptídico derivado de las proteínas de función dual de las secuencias aminoácidas de la invención. Estas proteínas de fusión y compuestos peptídicos se pueden hacer usando técnicas estándares conocidas en el estado de la técnica. Por ejemplo, los compuestos peptídicos se pueden hacer mediante síntesis química usando técnicas estándar de síntesis de péptidos y luego se introducen en las células mediante una variedad de medios conocidos en el estado de la técnica para la introducción de los péptidos en las células (por ejemplo, liposomas y similares).

[000252] La vida media in vivo de la proteína de fusión o compuestos peptídicos de la invención se puede mejorar haciendo modificaciones peptídicas, tales como la adición de sitios glicosilación N- enlazados en proteínas de función dual de la invención, o conjugando las proteínas de función dual de la invención con poli (etilenglicol) (PEG; pegilación), por ejemplo, a través de monopegilación de lisina o monopegilación de cisteína. Estas técnicas han demostrado ser beneficiosas en la prolongación de la vida media de fármacos de proteínas terapéuticas. Se espera que la pegilación de las proteínas de la invención puede dar lugar a ventajas farmacéutica similares.

[000253] Además, la pegilación se puede conseguir en cualquier parte de un polipéptido del invento por la introducción de un aminoácido no natural. Ciertos aminoácidos no naturales pueden introducirse por la tecnología descrita en Deiters, J Am

Chem Soc 125:11782-11783, 2003; Wang y Schultz, Science 301:964-967, 2003; Wang , Science 292: 498-500, 2001; Zhang, Science 303:371-373, 2004 o en la Patente EE.UU. No. En pocas palabras, algunos de estos sistemas de expresión implican la mutagénesis dirigida a sitio para introducir un codón sin sentido, como lo es codón TAG ámbar, en el marco de lectura abierto que codifica un polipéptido de la invención. Dichos vectores de expresión luego se introducen en un huésped que puede utilizar un ARNt específico para el codón sin sentido introducido y cargado con el aminoácido no natural de elección. Aminoácidos no naturales particulares que son beneficiosos para el propósito de conjugar fracciones con los polipéptidos de la invención incluyen aquellos con acetileno y cadenas laterales de azido. Las proteínas de la invención que contienen estos aminoácidos novedosos luego se pueden pegilar en estos sitios elegidos en la proteína.

[000254] Del mismo modo la PEGilación se puede lograr en cualquier parte de una proteína de la invención por la introducción de pirrolisina o pirrolina-carboxi-lisina como se describe por Ou et al. (Proc Natl Acad Sci U S A. 2011 Jun 28;108(26):10437-42. Epub 2011 Jun 13).

EJEMPLOS

Ejemplo 1: Diseño de las Proteínas GLP-1/ de la presente invención.

[000255] Con el fin de probar la eficacia del GLP-1 y FGF21 juntos, se diseño una molécula de fusión. Debido a que la GLP-1 requiere un extremo N-terminal libre y la FGF21 requiere un C-terminal libre para la unión y activación del receptor, las dos fueron clonadas en el orden de N-GLP-1-enlazador-FGF21-C. Los constructos iniciales se hicieron con los residuos 7-35 de GLP-1 y con o sin varias modificaciones en el extremo N-terminal para la protección de DPP-4 (discutido a continuación). Se probaron otras modificaciones de la GLP-1 (mutantes puntuales o deleciones) y se registro su potencia in vitro.

[000256] Se clonaron enlazadores de 3, 8, 10, o 20 aminoácidos. Para la FGF21, se utilizaron los residuos 33-209 de la proteína humana tipo silvestre. Se produjeron constructos con y sin sitios de PEGilación. La PEGilación se llevó a cabo utilizando reactivos de maleimida-PEG (40 kDa PEG lineales y ramificados) reactivos frente a una Cys introducida. La Cys se colocó en la posición R154 de la FGF21 y en varios sitios en la GLP-1 o exendina-4 y el enlazador. Las constructos también se han diseñado para la modificación mediante la incorporación de un PCL, Pyl, Pyl analógico o aminoácidos de origen no natural reactivos mediante la introducción de un codón TAG (ámbar) en la posición R154 de FGF21 o K34 de GLP-1. Los constructos adicionales fueron diseñados

para dos modificaciones diferentes mediante la inclusión de la Cys en R154 o K34 y un codón TAG para la incorporación de un PCL, Pyl, Pyl analógico o un aminoácido de origen no natural reactivo en el otro sitio. Los constructos fueron hechas fusionando a GLP-1 y FGF21 a un dominio Fc de un anticuerpo humano. La secuencia de la variante 76 (V76) de FGF21 se introdujo en el formato de fusión.

[000257] Expresión y purificación de las fusiones: Las proteínas de fusión se expresan típicamente en E. coli mediante la clonación de la secuencia de ADN que codifica la variante de doble función particular, en un vector que contiene las secuencias de codificación para uno varios de los dominios extraíbles (por ejemplo, variantes NPRO o NPRO con o sin una etriqueta His6 o etiqueta His6 con un péptido de reconocimiento TEV-proteasa o His6-ubiquitina o His6-Smt3) Los vectores incluyen, además, un promotor inducible para iniciar la transcripción del ARNm para la expresión de proteínas, como el lac, promotores T7 o arabinosa). Brevemente, los vectores se transformaron en derivados de células DH10b de E. coli o derivados de células BL21 (DE3), cultivadas bajo condiciones estándar, y se indujeron para expresar la proteína con 0,2% de arabinosa o IPTG añadido al medio de cultivo. Se recogieron las células 3-4 horas después de la inducción, se contaron, y se congelaron. Los pellets se descongelaron y se resuspendieron para la lisis por sonicación y la proteína insoluble se aisló por centrifugación. El sedimento se solubilizó en guanidina 6 M, y las proteínas, los aclarados disueltos se cargaron a las columnas de Ni-NTA. Las columnas se lavaron con tampones desnaturalizantes y luego con tampones no desnaturalizantes y se eluyeron en un tampón no desnaturalizante según la práctica estándar. Los eluidos eran intercambiados de tampón para eliminar el imidazol, se digirió con proteasas específicas para la eliminación de las etiquetas, (por ejemplo la proteasa TEV, ubquitinasa USP2, o Ulp1 para la eliminación específica de SMT3), y se purificaron de Ni-NTA. Las fracciones que contienen las proteínas de fusión fueron nuevamente purificadas por cromatografía de exclusión de tamaño antes de la PEGilación (típicamente con PEG maleimida NOF Sunbright GL2-400MA 40 kDa ramificado). Las proteínas pegiladas se aislaron mediante una cromatografía de intercambio aniónico, colocadas en PBS, y se concentraron o almacenaron para los diversos ensayos.

[000258] Protección de GLP-1 por la mutación del extremo N-terminal: Dado que el procesamiento de la N-terminal de GLP-1 por la DPP-4 es conocido para inactivar el péptido hacia su receptor primario, se exploraron varias mutaciones reportadas en la literatura para frenar este proceso. Los constructos con la adición de una glicina adicional

para el extremo N-terminal, la mutación de Glu9 a Pro, o la mutación de Ala8 a Gly o Ser, fueron examinados en un ensayo basado en células de la actividad de GLP-1. En el contexto de la fusión de FGF21 con un PEG unido a FGF21 de 40 kDa, las mutaciones en Ala8 retienen una actividad superior a las otras modificaciones N-terminales (Figura 1A), y la mutación A8S se utiliza en diseños posteriores. Además, del- DPP 4-resistente analógico, se ensayó también la exendina-4 en las fusiones como un resto alternativo con vida media prolongada *in vivo* (Figura 1b). [Véase, e.g., J. Y. Ohet al. (2009) Bulletin of the Korean Chemical Society 30, 2471-2474; K. Adelhorst, (1994) JBC 269, 6275; B. D. Green et al. (2003) Biological Chemistry 384, 1543; J. C. Parkeret al. (1998) Journal of Peptide Research 52, 398; C. F. Deaconet al. (1998) Diabetologia 41, 271; R. Burcelinet al. (1999) Metabolism-Clinical and Experimental 48, 252; U. Ritzelet al. (1998) Journal of Endocrinology 159, 93.]

[000259] La mutación de A8S también se comparó con el tipo silvestre de GLP-1 para las propiedades farmacocinéticas *in vivo*. Una proteína de fusión con el tipo silvestre de GLP-1 (V329) y una con la mutación de GLP-1 (A8S) (V232) se inyectaron por vía intravenosa en ratas a 1 mg / kg, y se midieron los niveles en suero contra curvas estándar producidas con la soluciones de dosificación. Cuando se mide con un kit de ELISA de FGF21- humana reactiva, las moléculas parecen comportarse de manera similar entre sí para FGF21 y solas con el mismo tamaño (40 kDa) de PEG unida (Figura 2). Para medir la extensión de la vida media de la actividad deGLP-1, los ratones C57BL/6J fueron inyectados con las mismas dos variantes de proteínas de doble función ya sea con el tipo silvestre de GLP-1 o GLP-1 (A8S). Mientras tanto se bajó agudamente el nivel de glucosa alimentado después de la dosificación, la variante GLP-1 (A8S) retiene una mayor actividad en una prueba de tolerancia oral a la glucosa (TTOG) Medición de 3 días después de la dosis (Figura 3). Estos datos demuestran que la mutación A8S aumenta la vida media efectiva de la fracción de GLP-1 de la fusión mientras que la PK de la fracción de FGF21 no se ve muy afectada por la adición de GLP-1.

[000260] Diseño del enlazador peptídico entre GLP-1 y FGF21: Se ensayaron los enlazadores de 3, 8, 10, o 20 aminoácidos. No se observó ninguna diferencia significativa en la actividad de las variantes con diferentes enlazadores, aunque hubo diferencias en los rendimientos de expresión en algunos contextos de fusión. Además, los constructos se probaron con diferentes longitudes de GLP-1 (de 8 a 31 residuos) o

exendina-4 (30 o más residuos) y se ensayaron para la actividad in vitro en el ensayo basado en células de GLP-1R.

Ejemplo 2: Ensayo de fusiones en modelos de ratones diabéticos db / db y ob / ob.

[000261] Ee ha demostrado que FGF21 mejora los niveles de glucosa en la sangre, los niveles de lípidos en el hígado, y el peso corporal en el modelo de ratón ob / ob de diabetes tipo 2 (DT2;. Véase, e.g., A. Kharitonenkovet al. (2005) Journal of Clinical Investigation 115, 1627; T. Coskun et al. (2008) Endocrinology 149, 6018; and E. D. Berglund et al. (2009) Endocrinology 150, 4084).. Del mismo modo, los análogos de GLP-1 han demostrado mejorar el control de la glucosa, la función de las células beta y la salud del hígado en este modelo genético de diabetes de ratón (Gallwitz B., Glucagon-like peptide-1 as a treatment option for type 2 diabetes and its role in restoring beta-cell mass. *Diabetes Technol Ther*. 2005, 7:651-7).

[000262] Para determinar si la fusión de GLP-1 a FGF21 daría lugar a una eficacia adicional, se ensayaron las versiones de GLP-1 y GLP-1 (A8S) WT frente a una molécula equivalente de FGF21 cada una con la misma unión de PEG ramificado de 40 kDa. En un estudio de dos semanas con dosis de dos veces por semana, la FGF21 (R154C)-PEG (V238) y la GLP-1-FGF21 (R154C)-PEG (V239) mostraron efectos similares sobre la glucosa en sangre, el peso corporal, y la salud del hígado (evaluada mediante los niveles séricos de la alanina aminotransferasa (ALT), el peso y la apariencia) (Figura 4). Las mediciones de glucosa mostraron una mejora más rápida para la fusión que para FGF21 solo después de la primera dosis. Los dos convergían en puntos de tiempo de más de un día después de la dosis, lo que sugiere que el beneficio adicional debido al GLP-1 fue más corto comparado con la eficacia de FGF21.

[000263] La fusión de GLP-1 (A8S)-FGF21 (R154C)-PEG (V235) también mostró efectos similares a 0.2 mg / kg, en lugar de los 1 mg / kg utilizados para los otros compuestos. Este compuesto por lo general también dio resultados más consistentes entre las dosis, lo que sugiere que la resistencia a DPP-4- fue fundamental para la mejora general de la eficacia en comparación con el comportamiento superior o inferior de GLP-1-FGF21 (R154C)-PEG (V239). También se observó que la GLP-1 (A8S)-FGF21 (R154C)-PEG (V235) a 1 mg / kg mostró una reducción adicional de glucosa, peso corporal, ALT / AST, y del peso del hígado.

[000264] Se llevaron a cabo estudios adicionales para determinar si la eficacia mejorada podría ser replicada con la co-administración de FGF21 separada y compuestos

de GLP-1. Para hacer que la dosificación de las proteínas fueran similares entre los grupos, se hizo un compuesto herramienta en el que se eliminaron 14 aminoácidos desde el FGF21 C-terminal. Esta molécula (V253) era al menos 1000x menos potente en el ensayo de receptor de FGF21 pero conservó igual potencia en el ensayo de GLP-1R en comparación con las otras fusiones y era incapaz de inhibir la actividad de FGF21 tipo silvestre. Este compuesto tenía una farmacocinética similar en ratas frente a las otras fusiones cuando se evaluó mediante los niveles de FGF21 (Figura 2).

[000265] Los resultados de un estudio de cuatro semanas en el que los ratones ob / ob fueron tratados con FGF21 (V76)-PEG, la GLP-1 (A8S)-PEG (V253), los dos juntos, o GLP-1 (A8S)-FGF21 (V76) - fusión de PEG (V272) se muestran en la Figura 5. La fusión fue significativamente más eficaz que el co-tratamiento de la FGF21 individual (V76)-PEG y GLP-1 (A8S) con las moléculas de PEG-en la medición de la AUC de la glucosa proveniente de la dieta, la glucosa OGTT AUC, el peso corporal al final del estudio, el contenido de lípidos hepáticos y de ALT (valores de $p < 0,05$ para todos).

[000266] La exposición sérica de la FGF21 humana y la GLP-1 se verificaron en las muestras de suero terminales para confirmar la exposición de los diversos grupos. Se detectó la FGF21 humana en todos los grupos tratados, y se midió un aumento significativo en el GLP-1 activo en todos los grupos tratados con GLP-1 (A8S)-PEG (V253) o GLP-1 (A8S)-FGF21 (V76)-PEG (V272). El grupo GLP-1 (A8S)-FGF21 (V76)-PEG (V272) mostró niveles similares de FGF21 humana frente a la FGF21 al grupo (V76) –PEGd e 0,2 mg / kg, lo que sugiere que la eficacia mejorada no se debe a una mayor acumulación sistémica de la fusión en los animales. La molécula de GLP-1 (A8S)-PEG (V253) era completamente activa *in vitro* (Tabla 8). Estos datos demuestran que la eficacia de la fusión era dependiente de que ambos restos estuviesen en una sola molécula, y, por ejemplo, se beneficiaban de una sinergia a nivel celular.

[000267] Para prueba adicional de GLP-1 (A8S)-FGF21 (V76)-PEG (V272) vs la coadministración, se llevaron a cabo tres estudios en ratones db / db. En cada estudio, los ratones macho db / db (8-11 semanas de edad fueron dosificados ip dos veces por semana. En un estudio inicial de tres semanas, la proteína de fusión de doble función, el GLP-1 (A8S)-FGF21 (V76)-PEG (V272), coincidía con la reducción de la glucosa de la co-administración en dosis iguales y superaba la eficacia de una dosis de una única entidad. La proteína de función dual de fusión también mostró una pérdida superior de peso corporal.

[000268] En un segundo estudio de dos semanas, la GLP-1 (A8S)-FGF21 (V76)-PEG (V272) mostró una disminución similar de glucosa en comparación con FGF21 (V76)-PEG a una dosis 25 veces menor y frente a GLP-1 (A8S) - PEG (V253) en una dosis de 5 veces menor. Además, la GLP-1 (A8S)-FGF21 (V76)-PEG (V272) rebajaba a HbA1c significativamente más y disminuía el peso corporal significativamente más que las mono-terapias o la dosis máxima eficaz de los grupos de co-administración (Figura 8, Tabla 2). Inesperadamente, alterando la proporción de dosificación de los co-tratamientos superiores a 1:01 no mejoró la eficacia observada (tabla 2). Esto sugiere que las proteínas de función dobles que incluyen un número igual de restos de GLP-1 y de restos de FGF21 pueden conseguir una eficacia óptima durante los tratamientos in vivo.

Tabla 2. Dosis Terapéuticas

Tratamiento	Dosis (mg / kg 2x/semana)	HbA1c, cambio de la inicial (%)	Peso corporal, el día 12 (g)
Vehículo	N / A	1,6 ± 0,9	42,3 ± 2,1
V76	1	2,2 ± 1,0	43.3 ± 2.1
V76	3	1,0 ± 0,9	42,3 ± 1,3
V76	5	0,7 ± 0,41	42,1 ± 1,9
V253	1	0,8 ± 0,71	42.6 ± 2.1
V253	3	0,6 ± 0,81	40.9 ± 2.6
V253	5	-0,1 ± 0,61	40,4 ± 2,1
V76 + V253	1 1	-0,3 ± 0,61	37.7 ± 2.21
V76 + V253	1 3	0,5 ± 0,61	37.7 ± 1.71
V76 + V253	3 1	-0,1 ± 0,51	38.9 ± 1.21
V76 + V253	3 3	-0,3 ± 0,41	36.0 ± 1.71
V272	0.2	0,4 ± 1,11	41,1 ± 2,7
V272	1	-1,1 ± 0.51,2	33,8 ± 2.11,2

[000269] [1] Valor p vs vehículo <0,05. [2] Valor p vs grupos de co-tratamiento <0,05

[000270] La Tabla 2 muestra la eficacia de las terapias de mono-y co-tratamientos a diferentes dosis en comparación con las proteínas de función doble en un estudio de ratón db / db con 2 semanas de dosificación. Los valores reportados son los promedios de las mediciones realizadas en ocho animales con desviación estándar. En un estudio de cuatro semanas, a los ratones macho db / db (n = 8) se les administró dos veces por

semana durante cuatro semanas con mono-terapias, los tratamientos complementarios, o proteínas de función dual de fusión. En el día 27, los ratones se mantuvieron en ayunas y se midieron los niveles de insulina en sangre para todos los animales. Al final del estudio, el contenido de insulina pancreática se extrajo y se midió (n = 3). La GLP-1 (A8S)-FGF21 (V76)-PEG (V272) mostró una respuesta robusta a la dosis 0,1 a 1 mg / kg con una eficacia igual o mejor que la de FGF21 (V76)-PEG o de GLP-1 (A8S)-PEG (V253) solo en los niveles de glucosa, HbA1c, y el peso corporal en más de 10 veces la dosis más baja y una mejor eficacia sobre la dosificación máxima eficaz de la co-administración. La GLP-1 (A8S)-FGF21 (V76)-PEG (V272) también redujo los niveles de insulina en ayunas en el suero al tiempo que aumenta el contenido de insulina del páncreas para una mayor medida que los tratamientos individuales o la combinación de GLP-1 (A8S)-PEG (V253) y FGF21 (V76)-PEG (Figura 7). Los animales tratados con GLP-1 (A8S)-FGF21 (V76)-PEG (V272) estaban cerca de apoyarse en los números de los ratones de control alimentados con glucosa, HbA1c, y del peso corporal.

[000271] Para la estimación de la tasa metabólica y la utilización de sustratos, medimos el consumo de oxígeno y la producción de dióxido de carbono mediante un sistema de calorimetría indirecta Oxymax (Columbus Instruments, Columbus, OH). Los animales fueron alojados en la cámara con un ciclo de 12h oscuridad / 12h luz en una temperatura ambiente de 22-24 ° C. Se determinaron las tasas de VO_2 y VCO_2 en la configuración del sistema Oxymax y el protocolo. El sistema fue calibrado con una mezcla de gas estándar para medir el O_2 consumido (VO_2, ml / kg / h) y el CO_2 generado (VCO_2, ml / kg / h). La tasa metabólica (VO_2) y el índice de intercambio respiratorio (RER) (relación de VCO_2/VO_2) fueron evaluados durante un período de 72h. Los valores de RER calculados indicaron que los animales tratados con la proteína de fusión se basaron más en sustratos lipídicos (valores de RER cerca de 0,75 a 0,85), en particular durante las 24-72 horas después de la segunda dosis cuando se compara con vehículo o grupos de tratamientos combinados. Durante este período, los grupos tratados con vehículo y de la combinación mostraron valores de RER estaban cerca de 0,9 lo que sugiere una mayor utilización del sustrato de carbohidratos para la energía que la utilización de un sustrato lipídico en menor medida por el gasto de energía. Los datos sugieren que el tratamiento con proteína de fusión causó un aumento en la oxidación de ácidos grasos que pueden estar contribuyendo a su efecto de reducción de peso corporal en ratones db / db

[000272] Una proteína de fusión de doble función, en comparación con combinación, mostró una mejora de la función de las células beta in vivo. Después de cuatro dosis de cualquiera de las proteínas de doble función de GLP-1 (A8S)-FGF21 (V76)-PEG (V272) o GLP-1 (A8S)-PEG (V253), además de FGF21 (V76)-PEG, se dosificaron los ratones db / db por vía oral con glucosa y arginina. La excursión de glucosa del grupo de la proteína tratada con doble función (72% inferior a la del vehículo) fue significativamente menor que el grupo de tratamientos combinados (56% inferior a la del vehículo). Durante el experimento, la cantidad de insulina secretada también fue mayor para el grupo tratado con proteína de doble función (279% superior a vehículo) que para el grupo tratado con combinación (122% más alto que el vehículo).

[000273] En general, las figuras y tablas, así como los datos no se muestran en el presente documento, demuestran la capacidad de las proteínas de fusión de doble función de la invención, por ejemplo, la GLP-1 (A8S)-FGF21 (V76)-PEG (V272), para mejorar los parámetros metabólicos in vitro e in vivo en modelos de roedores de la diabetes, en comparación con una combinación de GLP-1 y FGF21 individuales (por ejemplo, FGF21 (V76)-PEG y GLP-1 (A8S)-PEG (V253)). Dichos parámetros mejorados (como se utiliza aquí, "parámetros metabólicos") incluyen, pero no se limitan a glucosa proveniente de la dieta (AUC), peso corporal, los triglicéridos del hígado, la HbA1c de plasma, niveles de triglicéridos en suero, los niveles de colesterol total, la medición de tolerancia oral a la glucosa de glucosa en suero de prueba (AUC), la insulina sérica en ayunas, el contenido de insulina del páncreas, y el porcentaje de grasa corporal.

Ejemplo 3: Diseño de las proteínas GLP-1/ de la presente invención a partir de variantes de FGF21.

[000274] Se hizo una fusión con la siguiente secuencia variante de FGF21, a la cual nos referimos aquí como "Variante 76" o "V76":

1 DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ
61 ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEDGY NVYQSEAHGL PLHLPGNRSP 121 HCDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS
(SEQ ID NO: 130)

[000275] FGF21 (V76)-PEG cuenta con un 40 kDa PEG ramificado unido a través de Cys154, y las mutaciones en ocho puntos relativos a la proteína tipo silvestre de 177 aminoácidos (Q56E, D74H, Q82E, R105K, K150R, R154C, R159Q, S195A, todo hecho en relación de las secuencias de las proteínas de larga duración FGF21 (NCBI referencia secuencia de números NP_061986.1)). El riesgo de inmunogenicidad clínica hacia esta molécula se considera bajo basado en un ensayo de curso de tiempo EpiScreen de respuestas de células T humanas. La molécula tiene una vida media sérica de más de 30 h en ratones y ratas y reduce significativamente la glucosa AUC en ratones ob / ob con dos inyecciones semanales de 1 mg / kg.

[000276] La molécula GLP-1 (A8S)-FGF21 (V76)-PEG (V272) fue probada contra el GLP-1 (A8S)-FGF21 (R154C)-PEG (V235) de fusión en tres dosis (0,05, 0,1 y 0,2 mg / kg). Ambas fusiones mostraron una respuesta a la dosis en la reducción de la AUC de la glucosa, el peso corporal, la ingesta de alimentos, y ALT (Figura 6). La GLP-1 (A8S)-FGF21 (V76)-PEG (V272) era generalmente igual o mejor y mostró una mayor eficacia para todos los parámetros en la dosis de 0,2 mg / kg. La GLP-1 (A8S)-FGF21 (V76)-PEG (V272) también mostró la reducción de los triglicéridos séricos y el colesterol en tanto 0,1 y 0,2 mg / kg. Basándose en estos datos, la molécula de GLP-1 (A8S)-FGF21 (V76)-PEG (V272) muestra unas propiedades similares o mejores a las fusiones iniciales y es adecuada para su posterior estudio y desarrollo.

[000277] La Ex4 (1-30)-L20-FGF21 (V76)-PEG (V277) también mostró una eficacia similar a la de la GLP-1 (A8S)-FGF21 (V76)-PEG (V272) en un estudio de 2 semanas ob / ob para el control de la glucosa , el peso corporal y los niveles de lípidos.

[000278] Elección de sitio de PEGilación para la extensión de la vida media. Como se ve por ejemplo en la Figura 2, la PEGilación se usa para extender la vida media de la molécula desde unos pocos minutos para la GLP-1 o menos de una hora para FGF21 a más de 30 horas como fue medida en ratas. Para ver si las dos mitades de la fusión podrían ser moduladas mediante la colocación de la PEG, se hizo una serie de constructos con la cisteína extra no en la secuencia de FGF21 pero en la de la GLP-1 o del enlazador. Estos constructos se ensayaron para determinar tanto la actividad de GLP-1 y FGF21 en ensayos basados en células. Los experimentos revelan un tramo de posiciones dentro de las secuencias de la proteína de fusión en la que la PEGilación no afecta en gran medida, la actividad. Una cuidadosa colocación del PEG en el extremo N-terminal de este tramo se podría utilizar para hacer una molécula con actividad

disminuida de GLP-1 mientras que la colocación en el enlazador se podría utilizar para eliminar la actividad de FGF21. Tal molécula puede ser útil en el ajuste de la potencia de la fusión (por ejemplo, si hay una alta potencia de GLP-1 o exendina-4 que resulta en una ventana terapéutica pobre cuando se dosifica en el nivel eficaz de una FGF21 de menor potencia).

[000279] Basándose en estos datos, se prepararon las fusiones con PEGilación en K34C de GLP-1 (V273) o K27C de exendina-4 (V274) con FGF21 (V76-C154R) cuya secuencia es la siguiente:

DSSPLLQFGG QVRQRYLYTD DAQETEAHLE IREDGTVGGA AHQSPESLLE LKALKPGVIQ ILGVKTSRFL CQKPDGALYG SLHFDPEACS FRELLLEEGY NVYQSEAHGL PLHLPGNRSP HRDPAPQGPA RFLPLPGLPP ALPEPPGILA PQPPDVGSSD PLAMVGPSQG RSPSYAS
(SEQ ID NO: 131)

[000280] Un estudio de dos semanas en ratones ob / ob demuestra que, a pesar de que las nuevas fusiones (V273 y V274) exhibieron potencia in vitro similar a las fusiones anteriores y una farmacocinética similar en ratas, no eran tan eficaces in vivo. A 0,2 mg / kg, la variante de fusión GLP-1 no mostró una eficacia significativa para la glucosa o el peso corporal. La variante de fusión de exendina-4 mostró una disminución significativa de la ALT y el peso corporal y una tendencia hacia una menor AST en comparación con el vehículo pero no tiene un efecto significativo sobre la glucosa en sangre. Aunque estas fusiones pueden mostrar eficacia en dosis más altas, el formato original con PEGilación en R154C en FGF21 fue elegido para el estudio adicional. No está claro si esta diferencia inesperada indica una característica especial de la versión GLP-1 (A8S)-FGF21 (V76-154R)-PEG de la fusión que está bloqueada mediante una PEGilación en el sitio alternativo o si hay otra explicación, como el bloqueo de PEG de la interacción de GLP-1R más en las células nativascon los niveles de receptores nativos que en las células transfectadas utilizadas para el ensayo *in vitro*.

Ejemplo 4: Caracterización adicional de las proteínas GLP-1/FGF21 de la presente invención.

[000281] Con el fin de optimizar aún más y clasificar las proteínas de doble actividad de GLP-1-FGF21-PEG de la presente invención, y para obtener una apreciación más completa de la mejora de su eficacia respecto a los modelos de co-administración, las fusiones se prueban en el modelo de eficacia ob / ob. Estos incluyen, pero no se limitan a moléculas con enlazadores más largos o más cortos, las proteínas adicionales de la invención (que pueden incluir pero no se limitan a mutantes basados en la variante 76 o variantes adicionales en el desarrollo, tales como mutaciones puntuales, inserciones o deleciones adicionales o diferentes, moléculas dimerizadas, estrategias alternativas de PEGilación, moléculas con puentes disulfuro adicionales, moléculas producidas en diferentes sistemas de expresión), y variantes de exendina-4 o de GLP-1 para mejorar la estabilidad o la actividad sérica. Los candidatos también se filtran mediante la predicción de la inmunogenicidad in silico, los niveles de expresión, y la calidad de los productos finales (relacionados con la solubilidad, la agregación y la estabilidad).

[000282] Con el fin de caracterizar mejor la sinergia de FGF21 y GLP-1 para el tratamiento de la diabetes y la obesidad, se representa la dinámica de la interacción de las dos moléculas, y las proteínas de fusión se comparan con la co-administración máxima eficaz de las moléculas parentales individuales. Estos experimentos se llevan a cabo en los modelos que se centran en la obesidad y la pérdida de peso (dieta inducida por la obesidad en la rata o el ratón) y las que mejoran aspectos del modelo de diabetes, como la hiperglucemia, la dislipidemia y la función alterada de las células beta (ob / ob o db / db). El grado de sinergia logrado por las proteínas de actividad dual de la invención se puede determinar aún más mediante la asignación de la eficacia de cada molécula individual, la combinación de los dos en diferentes proporciones, y la fusión.

[000283] Las pruebas en ensayos basados en células con islotes de rata primarias o inmortalizadas e islotes humanos primarios se llevaron a cabo para evaluar la eficacia de la fusión para la proliferación celular, protección de la apoptosis, y la función, por ejemplo, para validar las proteínas de actividad dual de la invención en modelos aceptados de la diabetes tipo 1 (ver hoy más Van Belle, T.L. et al. Drug Discovery today: disease models. 209 pp. 41-45). La preferencia en modelos in vivo son aquellos con la ablación de páncreas por estreptozotocina (STZ); dirigido a la ablación de las células beta en una cepa de ratón modificado genéticamente e inducible (RIP-DTA) en el que las células beta son destruidas por la suplementación de la dieta con doxiciclina para inducir la expresión dirigida de la toxina de la difteria, o el modelo de ratón diabético no obeso de la diabetes tipo 1 (NOD) con un mecanismo de autoinmunidad. En estos modelos, las

pruebas son posibles tanto en la dosificación profiláctica para evaluar la protección de las células beta (en particular en NOD), así como la dosificación después de la exposición para evaluar la estimulación de las células beta y la proliferación (en particular en STZ y RIP-DTA).

[000284] Los siguientes experimentos se llevan a cabo para comprender más plenamente el mecanismo de la sinergia de eficacia obtenido por las proteínas de actividad dual de la invención en los regímenes de co-administración: Las células co-transfectadas con ambos receptores (GLP-1R; FGFR1 (IIIc), 2 (IIIc), 3 (IIIc), o 4; y beta-Klotho) pueden ser utilizadas para determinar si los receptores son capaces de potenciar la señal de oposición en la superficie celular. La diafonía entre las señales corriente abajo se puede detectar en células que expresan naturalmente ambos receptores (por ejemplo, las células beta). En los animales, la contribución de la ingesta de alimentos (investigada con cohortes par-alimentados), el aumento de la tasa metabólica (investigado en la abrazadera y experimentos en cámaras metabólicas) y otros mecanismos se podría investigar más a fondo para dilucidar un modo de acción de los efectos sinérgicos de la dos señales. Los perfiles de expresión génica de los tejidos clave (hígado, páncreas, tejido adiposo, intestino, corazón, aorta, cerebro, etc) también pueden llevarse a cabo con el fin de dilucidar la señalización única de las proteínas de fusión y potencialmente representan el aumento de su eficacia.

[000285] Los siguientes experimentos se llevan a cabo para entender la actividad de las proteínas de función doble de la presente invención para controlar o modificar la diferenciación de las células madre mesenquimales (MSC): Las MSC pueden ser tratadas con las proteínas de la invención solas o con compuestos o mezclas de compuestos conocidos para inducir la osteogénesis o con compuestos o mezclas de compuestos conocidos para inducir la adipogénesis y se pueden medir las tasas resultantes de la diferenciación a células similares a osteoblastos o células similares a adipocitos.

Ejemplo 5: Diseño de proteínas que contienen GLP-1 y exendina-4- de la presente invención a partir de las variantes de fusión de FGF21 con Fc.

[000286] Los constructos fueron hechos con enlazadores alternativos (por ejemplo, SGGGGSGGGGSGGGGSA (SEQ ID NO: 138), GGGGS (SEC ID N º: 173), GG, y GGGGSGGGGSGGGGS (SEQ ID NO: 174)), variantes de FGF21, y péptidos relacionados GLP-1-. Típicamente las proteínas de función doble que contienen un

dominio Fc se expresaron en células HEK293. Las secuencias de estas fusiones de codificación de ADN se clonaron en vectores que contenían la secuencia que codifica un péptido líder para dirigir la secreción de las proteínas y que contienen además las secuencias necesarias para promover la síntesis de ARNm y la expresión de proteínas de los productos deseados. Las células HEK293 se transfectaron transitoriamente con los vectores. Se recogieron los medios de estos cultivos celulares, se filtraron, y se purificaron por cromatografía de proteína de afinidad A. Los eluatos se llevaron a pH netural y se purificaron por cromatografía de exclusión de tamaño. Los productos se concentraron entonces y se almacenaron durante diversos ensayos. Las variantes se ensayaron para determinar la actividad in vitro en cuatro ensayos: inducción de la expresión de CRE-luciferasa en células HEK293 transfectadas con el GLP-1R (Tabla 3) y la secreción estimulada por glucosa de insulina en las células INS1E (Tabla 6) para medir la actividad a través de GLP-1R; la inducción de la fosforilación de ERK en HEK293 transfectadas con beta-Klotho (Tabla 4) y la captación de 2-desoxiglucosa por células 3T3L1 (Tabla 5) para medir la actividad de FGF21. Las variantes con GLP-1 (A8S) y un enlazador de 15 aminoácidos eran más activos que las variantes con el GLP-1 y enlazadores más cortos. La exendina-4 (1-39) con todas las variantes ensayadas fueron enlazadores activos. Las variantes de exendina-4 (1-30)-que contienen tenían una potencia similar in vitro para constructos que contienen exendina-4 (1-39)-. Una variante con una repetición en tándem de la exendina-4 (1-39) no era tan potente como las variantes con una sola copia de un péptido relacionado con el GLP-1-.

Tabla 3. Activación de GLP-1R en células HEK293 -GLP-1R-CRE- Reportero de luciferasa.

Variante #	**GLP-1R CE50 (nM)**
V193	0.024
V194	0.042
V195	0.041
V196	0.027
V197	0.085
V198	0.024
V199	0.046
V202	0.025
V203	0.252
V206	0.047
V207	0.088

V208	0.044
V209	0.033
V210	0.038
V211	0.032
V212	5.90
V213	18.4
V214	0.023
V215	8.99
V216	0.025
V217	4.53

Tabla 4. Actividad medida en un ensayo de fosforilación de ERK de la actividad de FGF21 en células HEK293-KLB.

Variante #	**βklotho pERK EC50 (nM)**	**% YMáx vs FGF21 PESO**	**Potencia (Dola WT FGF21)**
V193	2.1	101	1.2
V194	9.8	97	5.8
V195	13	76	7.6
V196	0.87	58	0.18
V197	0.92	63	0.19
V198	0.94	67	0.24
V199	1.04	58	0.27
V202	1.78	67	0.37
V203	1.01	70	0.21
V206	0.51	83	0.15
V207	1.10	83	0.34
V208	0.53	54	0.20
V209	0.70	83	0.21
V210	0.84	88	0.24
V211	0.58	73	0.17
V212	1.24	89	0.37
V213	1.00	61	0.37
V214	0.55	50	0.21
V215	0.92	51	0.34
V216	0.85	56	0.32
V217	2.23	46	0.83

Tabla 5. Actividad medida en la captación de 2-desoxiglucosa en adipocitos de ratón 3T3L1.

Variante #	**2DOG EC50 (nM)**	**% YMáx vs FGF21**	**Potencia (Dobla WT FGF21)**

		PESO	
V194	2.8	118	1.6
V196	0.055	90	0.14
V197	0.077	100	0.18
V198	0.079	105	0.21
V199	0.090	101	0.37
V202	0.45	86	2.5
V203	0.078	100	0.31
V206	0.093	85	0.34
V208	0.067	87	0.25
V209	0.157	91	0.59
V210	0.067	98	0.27
V211	0.073	99	0.29
V214	0.213	87	0.81
V216	0.137	80	0.56

Tabla 6. Actividad medida por la secreción de insulina estimulada por glucosa (GSIS) en células de insulinoma de rata INS1E.

Variante #	**GSIS EC50 (nM) (promedio de 2)**
V196	9.1
V197	46
V198	4.4
V199	5.0
V202	33
V203	128
V206	2.4
V208	1.6
V209	1.9
V210	2.7
V211	1.5
V214	4.6
V216	6.8

[000287] Ocho ratones ob / ob fueron dosificados con las variantes dos veces por semana a 0,5 mg / kg durante dos semanas. Se midieron la glucosa en sangre y el peso corporal durante los estudios, y los lípidos hepáticos se midieron al final del estudio. Como se observa en la siguiente tabla (Tabla 7), los animales tratados redujeron la glucosa en sangre, menor peso corporal y los lípidos hepáticos menores durante todo el estudio, en comparación con los animales de la misma edad tratados con vehículo.

Tabla 7. Resumen de los resultados del tratamiento de ratones ob / ob dos veces por semana a 0,5 mg / kg durante dos semanas.

Variante #	Pérdida de PC, D12 (% de Veh)	descenso de AUC de la ingesta (% de Veh)	TG ayuno (8h o O / N) (% de Veh)	lípidos hepáticos (% De Veh)
V193	-20 *	-55 *	39	-66 *
V195	-7 *	-30 *	N / A	-10
V196	-23 *	-63 *	8	-65 *
V197	-16 *	-53 *	52	-85 *
V198	-24 *	-64 *	-7	-69 *
V199	-24 *	-64 *	-7	-69 *
V200	-14 *	-39 *	-5	-62 *
V201	-22 *	-60 *	-10	-58 *
V202	-18 *	-59 *	22	-63 *
V203	-8 *	-33 *	68 *	-73 *
V206	-16 *	-58 *	-19	-56 *
V207	-22 *	-64 *	-10	-55 *
V208	-21 *	-60 *	10	-57 *
V209	-28 *	-66 *	-11	-62 *
V210	-23 *	-67 *	-14	-57 *
V211	-28 *	-68 *	0	-52 *
V214	-19 *	-58 *	110 *	-41 *
V216	-21 *	-62 *	29	-51 *

* Valor de p vs vehículo <0,05

Ejemplo 6: Caracterización adicional de las proteínas de la invención en ensayos de farmacología del receptor.

[000288] Como se ve, por ejemplo, en la Figura 9, las proteínas agonistas duales de la invención se compararon con las diversas formas de FGF21 y exenatida para su capacidad de inducir la fosforilación de ERK en células HEK293 transfectadas con el beta-Klotho. A los 10 minutos, las células tratadas con GLP-1 (A8S)-FGF21 (V76)-PEG (V272) mostraron una mayor potencia y mayor proporción máxima de fosfo-ERK total ERK en comparación con las células tratadas con FGF21 (V76)-PEG además de exenatida. La señal de fosfo-ERK superior máxima muestra que la proteína de función dual V272 es un agonista superior para la señalización a través FGFR1c/beta-klotho en comparación con el FGF21 (V76)-molécula de PEG, posiblemente debido a una mejor

interacción con el receptor (por ejemplo, la extensión N-terminal de la proteína causando doble función para lograr una estructura más favorable para la unión al receptor o de activación).

[000289] Las proteínas seleccionadas de doble función se ensayaron para determinar su capacidad para inducir la señalización de fosfo-ERK en adipocitos humanos. En comparación con una molécula de una sola función FGF21 (V101), las proteínas de función dual V208, V209, V211, V14, y V272 eran más potentes para la estimulación de la fosforilación de ERK. Porque éstas son las células humanas, sino que también puede sugerir que las proteínas de doble función de la invención serán de gran actividad en el tratamiento de las enfermedades humanas en comparación con los tratamientos que incluyen otros miméticos de FGF21, similar a lo que se ha demostrado para la actividad en los roedores a la fecha.

[000290] Como se ve, por ejemplo, en la Tabla 8 y la Figura 10, las proteínas agonistas duales de la invención se compararon con exenatida, péptido GLP-1, exendina 4 (1-39)-L5-Fc (V201), y GLP-1 (A8S)-PEG (V253) para su capacidad de señal a través de GLP-1R en células HEK293 transfectadas con células co-transfectadas con el GLP-1R, beta-Klotho y FGFR1c de GLP-1R o. Las células se trataron con el compuesto durante 30 minutos, y se midieron los niveles de AMPc. La Tabla 8 muestra que en las células tanto con el GLP-1R y el complejo receptor de FGF21, las proteínas agonistas duales (V272 y V211) mostraron mayor potencia que los agonistas individuales (V253 y V101) solos o en combinación con exendina-4. En las células con sólo el GLP-1R, la potencia de los agonistas individuales solos o en combinación con la exendina-4 fue igual a la de las proteínas agonistas duales. Variantes de FGF21 solas que fueron inactivas en el ensayo.

[000291] Tabla 8. Ensayo de inducción de AMPc en las células tratadas con las proteínas de función doble de la presente invención.

	GLP-1R/FGFR1c/beta-klotho			GLP-1R sólo		
Compuesto	CE50 (nM)	EC50 (aumento-Ex-4)	y-max (% Ex-4)	CE50 (nM)	EC50 (aumento-Ex-4)	y-max (% Ex-4)
Ex-4	6.0	1.0	100	11	1.0	100
V253	24	4.1	92	34	3.2	81
V253 + V76	12	2.1	91	38	3.5	82
V272	2.4	0.4	97	64	6.0	76
V201	26	4.4	93	32	3.0	100
V201 +	15	2.4	98	18	1.7	82

V101						
V211	6.6	1.1	92	39	3.6	98

[000292] Dos líneas adicionales de células HEK293 fueron generadas en formato de ensayo PathHunder de DiscoveRx que utiliza la complementación de fragmentos de beta-galactosidasa en beta-arrestina y el C-terminal de GLP-1R para medir el reclutamiento de arrestina una vez se da la activación del receptor. Las figuras 10a-10d muestran que con una incubación de 1 hora de células con proteína o péptido, el GLP-1 (A8S)-FGF21 (V76)-PEG (V272), exendina-4 (1-39)-L15-Fc-L15- FGF21 (V103) (V211), el GLP-1 (A8S)-PEG (V253), y la exendina-4 (1-39)-L5-Fc (V201), era menos potentes que la exenatida para la contratación de la beta-arrestina y que no se observó ninguna diferencia en el comportamiento si las células contenían GLP-1R/FGFR1c/beta-klotho o GLP-1R solo.

[000293] Los estudios de farmacología de receptor presentados aquí sugieren que las proteínas de función doble de la invención serán superiores en actividad y eficacia en comparación con las combinaciones o compañeros de tratamientos de FGF21 convencional y moléculas GLP-1/Exenatida. La proteína de función dual compuestos GLP-1 (A8S)-FGF21 (V76)-PEG (V272) mostró inesperadamente una mayor potencia y una señal máxima para la inducción de fosfo-ERK en comparación con la FGF21 (V76)-molécula de PEG. Estos datos muestran que las proteínas son de doble función equivalente o superior para la señalización de FGF21 en comparación con otras variantes de FGF21. Del mismo modo, las proteínas de función doble exhiben mayor potencia para el campo de señalización corriente abajo de GLP-1R cuando las células expresan tanto los receptores de GLP-1 y FGF21. Estas propiedades pueden reflejar una interacción directa entre los receptores o un efecto indirecto, tales como un aumento en la concentración local a través de la avidez hacia los dos receptores. Estas observaciones sugieren un mecanismo para mejorar la actividad de las proteínas de función dual que pueden explicar por qué los co-tratamientos de GLP-1 y las variantes individuales de FGF21 no han sido suficientes para que coincida con la eficacia in vivo de las proteínas de doble función, como se presenta en otros ejemplos de esta invención.

Ejemplo 7: Caracterización de las proteínas de la invención en modelos de diabetes tipo 1 (T1D)

[000294] Los ratones transgénicos RIP-DTA (9-10 semanas de edad) fueron tratados con doxiciclina durante 3 días para inducir la ablación de las células beta controlada (similar a Thorel et al., Naturaleza 2010 (464) 1149/54). Los ratones se trataron a continuación durante tres semanas, dos veces por semana con FGF21 (V76)-PEG, FGF21 (V76)-PEG + GLP-1 (A8S)-PEG (V253), o GLP-1 (A8S)-FGF21 (V76)-PEG (V272). Los animales tratados en los tres grupos mostraron unos niveles de glucosa después de la ingesta basales más bajos durante el estudio. El tratamiento con la proteína de función dual dio como resultado una AUC de glucosa significativamente menor (64% de reducción de vehículo) en comparación con el tratamiento de combinación (55% de reducción de vehículo). Después del ayuno, los ratones recibieron un bolo de glucosa, y se midió la excursión de la glucosa. Todos los tres tratamientos redujeron los niveles de glucosa en sangre en comparación con los animales tratados con vehículo, el tratamiento con proteína de función dual que resulta en una reducción del 76% en comparación con 66% para el tratamiento de combinación. Todos los grupos tratados habían aumentado el contenido de insulina del páncreas en comparación con el vehículo, con el tratamiento con proteína de función dual que resulta en un 76% más de contenido de insulina que el tratamiento de combinación.

[000295] Los ratones NOD se trataron dos veces por semana durante tres semanas con el GLP-1 (A8S)-FGF21 (154C)-PEG (V235) o GLP-1 (A8S)-PEG (V253) + FGF21 (154C)-PEG (V238). Después de dos semanas, los ratones se mantuvieron en ayunas y se les dio un bolo de glucosa para un OGTT. Los ratones tratados con la proteína de función doble mostraron niveles significativamente más bajos de glucosa en ayunas (~ 60% del vehículo) y el recorrido de glucosa durante la exposición (~ 45%) en comparación con el vehículo mientras que los animales tratados con la combinación en ayuno con los niveles de glucosa que no fueron significativamente diferentes de vehículo (~ 70%) y con excursiones similares de glucosa durante el desafío en comparación con los animales tratados con el vehículo. Al final del estudio, los animales tratados con proteínas de doble función tenían niveles basales significativamente más bajos de glucagón (~ 50%) que en el vehículo, mientras que los animales tratados con la combinación tenían niveles similares al vehículo.

Ejemplo 8. Resultados de inmunogenicidad predictivos - asociados a la Proteómica del péptido MHC (mApps) y resultados del ensayo de células T.

[000296] La formación de anticuerpos anti-fármaco (ADA) a los mAbs y otras proteínas terapéuticas potencialmente podría conducir a reacciones inmunotoxicológicas graves, tales como reacciones anafilácticas mediadas por IgE (Chung et al. (2008) N Eng J Med, 358, 1109-1117) o enfermedad compleja inmune, por ejemplo, vasculitis, glomerulonefritis (Descotes y Gouraud (2008) Expert Opin Drug Metab Toxicol 4, 1537-1549), así como una pérdida de la exposición al fármaco y de la eficacia clínica. Algunos pacientes tratados con las proteínas pegiladas terapéuticas de crecimiento de megacariocitos y factor de desarrollo (PEG-MGDF) y eritropoyetina (EPO; Eprex) desarrollaron la neutralización del ADA que son de reacción cruzada a sus respectivos homólogos endógenos, lo que lleva a trombocitopenia grave con PEG-MGDF y -aplasia de células rojas pura con Eprex (Li et al (2007) Blood 98, 3241-8;.. Casadevall et al (2002) N Engl J Med 346, 469-75). Por lo tanto, es importante para evaluar el riesgo de la inmunogenicidad antes de la prueba humana. La evaluación del riesgo de inmunogenicidad y la puesta en marcha de un plan de mitigación de riesgos de la inmunogenicidad sólida para la fase de desarrollo clínico será especialmente importante para las proteínas terapéuticas que contienen secuencias modificadas, no humanas, tales como las variantes de FGF21 que son objeto de esta invención.

[000297] La formación de la ADA se puede inducir en al menos dos formas diferentes. Vías dependientes e-independientes de células T se han descrito para la activación de células B. Una respuesta fuerte de IgG de alta afinidad es dependiente de células T y requiere la participación de las células T colaboladoras CD4 + (células TH). La respuesta inmune es análoga a una respuesta contra antígenos extraños: Las células TH ingenuas se activan específicamente por el antígeno profesional que de presentación células (APC), tales como las células dendríticas (DCs), y a su vez incuden la activación de células B específicos a los fármacos.

[000298] El ensayo de péptido asociado a la Proteómica-MHC (mApps) implica la identificación in vitro de péptidos asociados HLA de clase II, que son procesados por las células profesionales presentadoras de antígenos (APC), tales como las células dendríticas. Los procesos de captación de antígenos, procesamiento y presentación se tienen en cuenta. En este enfoque, las CD inmaduras derivadas de monocitos humanos de diferentes donantes de sangre sanos, se incubaron con diferentes candidatos de fármacos bioterapéuticos en la presencia de un estímulo de activación. Los péptidos asociados HLA de clase -II procesados naturalmente, que se derivan de la proteína bioterapéutica, se identifican mediante cromatografía líquida-espectrometría de masas (Kropshofer and

Spindeldreher (2005) in Antigen Presenting Cells: From Mechanisms to Drug Development, eds. Kropshofer and Vogt, Wiley-VCH, Weinheim, 159-98)

[000299] Los ensayos de células T proporcionan un formato en el que el riesgo potencial de inmunogenicidad de las proteínas terapéuticas enteras se puede evaluar. Los ensayos de células T evalúan la capacidad de una proteína terapéutica para inducir una respuesta de células T CD4 +. El uso de una cohorte de donantes de sangre sanos que cubren un amplio panel de haplotipos de HLA de clase II, las proteínas terapéuticas purificadas se ensayaron para la proliferación de células T y / o la secreción de citoquinas in vitro. Esta tecnología se ha utilizado con éxito para comparar variantes de proteínas con el potencial para inducir una respuesta inmune in vivo (Jones et al (2004) J Interferon Cytokine Res. 24, 560-72;.... Jones et al (2005) J Thromb Haemost. 3, 991-1000) y esta evaluación de anticuerpos monoclonales actualmente aprobados hace mostrar cierto grado de correlación entre la activación de las células T observadas in vitro y la inmunogenicidad en la clínica (Perry et al. (2008) Drogas RD 9, 385-96). En el contexto de esta invención, las células mononucleares de sangre periférica (PBMCs) de una cohorte de 50 donantes sanos que representan la población mundial (basado en alotipos HLA) se incuban con las variantes de FGF21 y las respuestas de células T se miden usando ensayos de proliferación ([3H] -la captación de timidina) y la IL-2 secreción de citoquinas (ELISPOT). Posteriormente, se llevó a cabo el análisis de la frecuencia y magnitud de las respuestas de las células T CD4 + para evaluar el riesgo de inmunogenicidad clínica.

[000300] El uso combinado de los ensayos de células T y de mApps proporciona un proceso eficaz para la evaluación del riesgo de la inmunogenicidad en la clínica utilizando células humanas. Las modificaciones a un candidato de proteína terapéutica que resulta en un número reducido de péptidos presentados por células dendríticas y un número reducido de donantes que responden al ensayo de células T serán ventajosas como estas proteínas que tienen un menor riesgo de desarrollar la inmunogenicidad en la clínica. Un ejemplo de una modificación de la proteína que se ha descrito para reducir la inmunogenicidad es la PEGilación. Sin embargo, una inmunogenicidad reducida no siempre se da con la PEGilación (Li et al. (2007) Blood 98, 3241-8).

[000301] En el ensayo de mApps, todas las bandas de compuestos GLP-1 (A8S) y moléculas de FGF21-fusión exendina-4, incluyendo V272 (pegilada) y V277 (pegilada) y las variantes de Fc-FGF21 no pegiladas que contienen el Q55C, las mutaciones G148C (V101 , V103 y V188), muestran un bajo número de ramificaciones, y variantes de

longitud del péptido que son comparables a V76 (pegilada) y más bajo que el tipo silvestre FGF21 o V76 (no pegilada). En el ensayo de células T, la frecuencia de las respuestas de las células T fue <10% de la cohorte del estudio y la magnitud de las respuestas eran bajos para V272 (pegilada) y V277 (pegilada) y las variantes de Fc-FGF21 no pegiladas que contienen el Q55C, las mutaciones G148C (V101, V103 y V188) (Cuadro 10). La ausencia de un resto de PEG o la ausencia del enlace disulfuro adicional en Q55C, G148C puede contribuir al aumento de mApps y respuestas del ensayo de células T visto con FGF21 de tipo silvestre y V76 (no pegilada) (Tablas 9 y 10). Basándose en los resultados de los ensayos de células T y mApps, el riesgo de desarrollar la inmunogenicidad en la clínica puede ser considerada a ser baja para V76 (pegilada), V272 (pegilada), V277 (pegilada) y las variantes de Fc-FGF21 no pegiladas que contienen a las mutaciones Q55C, G148C (V101, V103, V188). Se sabe que la adición de enlaces disulfuro a las proteínas puede potenciar su estabilidad proteolítica y esta característica de las variantes de Fc-FGF21 que contienen el Q55C, las mutaciones G148C puede contribuir a la reducción del número de péptidos presentados por las células dendríticas maduras cargadas con antígeno en los ensayos de mApps.

[000302] Además, el aumento de la respuesta en los ensayos mApps a no V76PEGilada puede explicarse por la cisteína libre que conduce a la dimerización de la molécula y afectar de ese modo de procesamiento y presentación de antígenos. Esto está en consonancia con la observación de que el FGF21 de tipo natural con la cisteína libre introducida mostró la misma tendencia a dimerizar y dio como resultado un aumento de la presentación de péptidos. Es probable que las proteínas de función doble PEGiladas adicionales, y las variantes de GLP-1 (A8S)-Fc-FGF21 (V202, V203, V212, V213, V214, V215, V216, V218, etc) y la exendina-4-Fc- las variantes FGF21 (V196, V197, V198, V199, V206, V207, V208, V209, V210, V211, etc) que contienen las mutaciones Q55C, G148C, también exhibirán mApps y respuestas del ensayo de células T que serían consistentes con un bajo riesgo de el desarrollo de la inmunogenicidad en la clínica.

[000303]

Tabla 9. Resumen del ensayo de células T con el tipo silvestre FGF21, Experimento V76 (no pegilado) y V76 (pegilado)

Resultado delEnsayo de células T	KLH (control positivo)	FGF21-WT (no pegilado)	FGF21-WT-(Pegilado)	V76 (no pegilado)	V76 (pegilado)
Proliferación %	90	6	4	16	4

ELISpot %	82	10	8	10	4
Proliferación y ELISPOT %	78	6	4	10	2

Tabla 10. Resumen del ensayo de células T con V76 (pegilado), Experimento V272 (pegilado), V277 (pegilado), V101, V103 y V188

Resultado delEnsayo de células T	KLH (control positivo)	V76 (pegilado)	V272 (pegilado)	V277 (pegilado)	V101	V103	V188
Proliferación %	92	4	6	4	4	6	4
ELISpot %	91	6	6	4	4	6	9
Proliferación y ELISPOT %	85	4	6	4	4	6	4

REIVINDICACIONES:

1. Una proteína de fusión de función dual que comprende una región agonista del receptor de GLP-1 y una región agonista del receptor del FGF21.

2. La proteína de fusión de función dual de la reivindicación 1 que comprende un péptido agonista del receptor de GLP-1 fusionado a una variante de FGF21, un enlazador, y un grupo PEG unido de tal manera con el fin de mejorar la función biológica de dicho agonista del receptor de GLP-1 o variante de FGF21 .

3. La proteína de fusión de función dual de la reivindicación 2, en donde la variante de FGF21 es la variante 76.

4. La proteína de fusión de función dual de la reivindicación 3, en donde el enlazador es un Fc.

5. La proteína de fusión de función dual de la reivindicación 3, en donde el enlazador es una variante de Fc.

6. La proteína de fusión de función dual de la reivindicación 1, en donde dicha proteína de fusión de función dual comprende además la variante 208, variante 209, variante 211, variante 214, variante 272, variante 277, o variante 311.

7. Un método de tratamiento de un trastorno metabólico mediante la administración a un sujeto en necesidad de una proteína de fusión de función dual que comprende un agonista del receptor de GLP-1 y un agonista del receptor de FGF21.

8. El método de la reivindicación 7, en donde dicha proteína de fusión de función dual mejora los parámetros metabólicos en sujetos sobre la administración de agonistas del receptor de GLP-1 por separado y agonistas del receptor de FGF21 en combinación.

9. El método de la reivindicación 7, en el que dicha proteína de fusión de función dual comprende además la variante 208, variante 209, variante 211, variante 214, variante 272, variante 277, o variante 311.

RESUMEN

La presente invención se refiere a la identificación de nuevas proteínas que comprenden el factor de crecimiento de fibroblastos 21 (FGF21) y otros reguladores metabólicos, incluyendo variantes de los mismos, conocidos por mejorar los perfiles metabólicos en sujetos en quienes se administren. También se divulgan métodos para tratar trastornos asociados con el FGF21, trastornos asociados con GLP-1, y trastornos asociados con Exendina-4, incluyendo condiciones metabólicas.

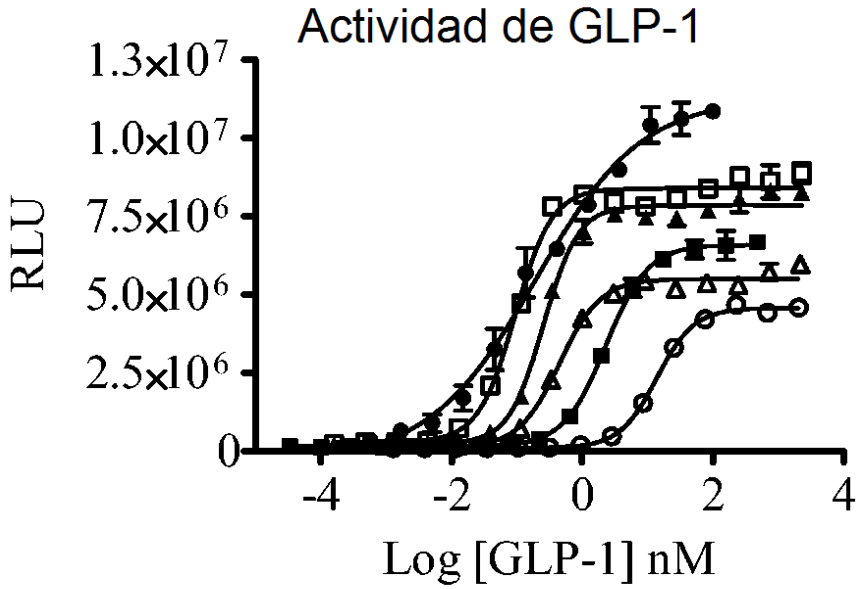

Actividad de GLP-1
1.3x10^7
1.0x10^7
7.5x10^6
5.0x10^6
2.5x10^6
0
RLU
-4
-2
0
2
4
Log [GLP-1] nM

FIG 1A

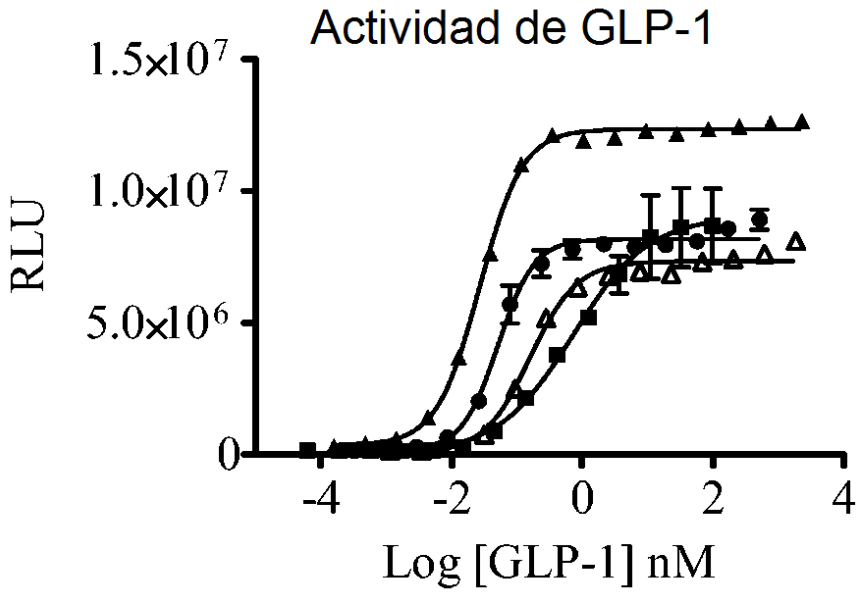

Actividad de GLP-1
1.5x10^7
1.0x10^7
5.0x10^6
0
RLU
-4
-2
0
2
4
Log [GLP-1] nM

FIG 1B

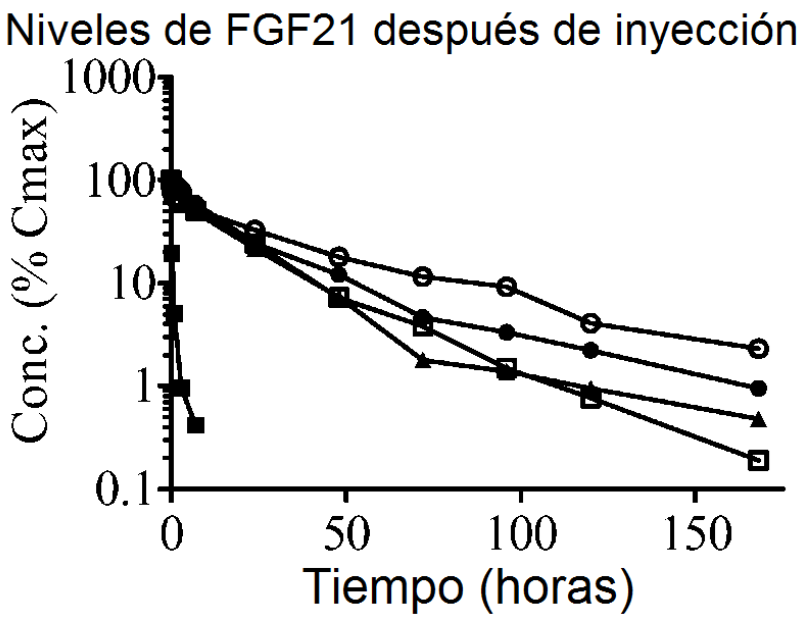

Niveles de FGF21 después de inyección
Conc. (% Cmax)
1000
100
10
1
0.1
0
50
100
150
Tiempo (horas)

FIG 2

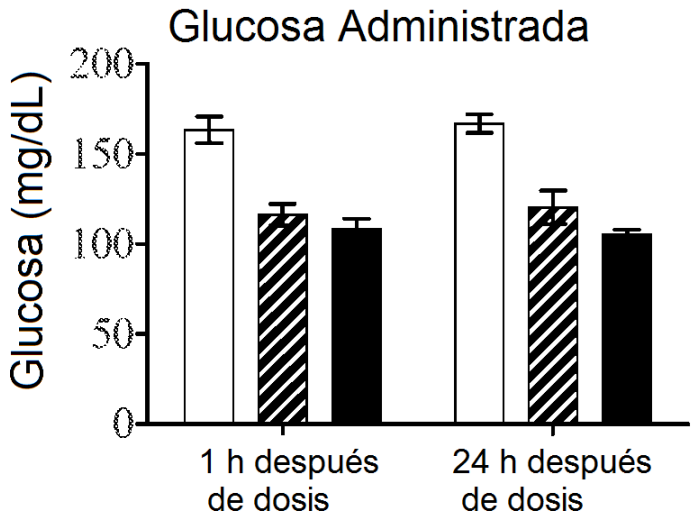

Glucosa Administrada
Glucosa (mg/dL)
200
150
100
50
0
1 h después de dosis
24 h después de dosis

FIG 3A

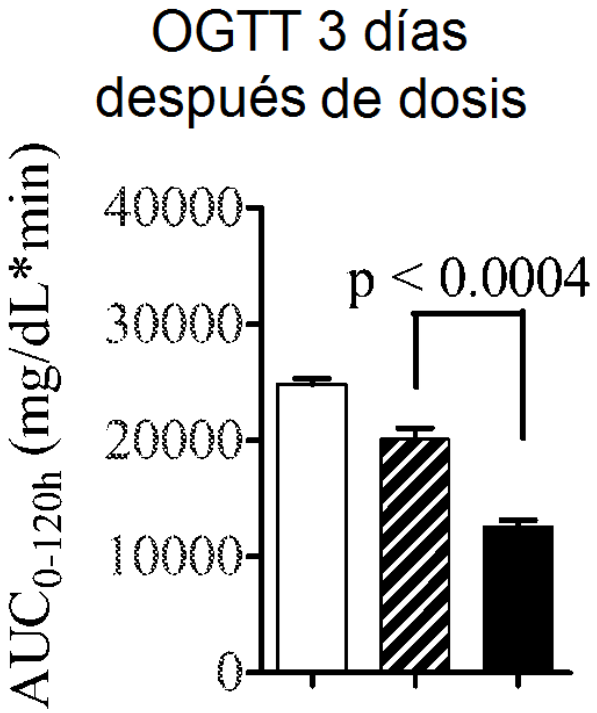

OGTT 3 días después de dosis
AUC0-120h (mg/dL*min)
40000
30000
20000
10000
0
p < 0.0004

FIG 3B

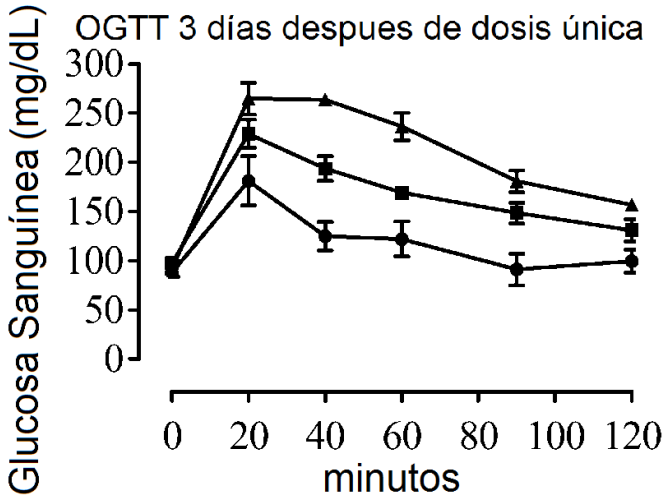

OGTT 3 días despues de dosis única
Glucosa Sanguínea (mg/dL)
300
250
200
150
100
50
0
0
20
40
60
80
100
120
minutos

FIG 3C

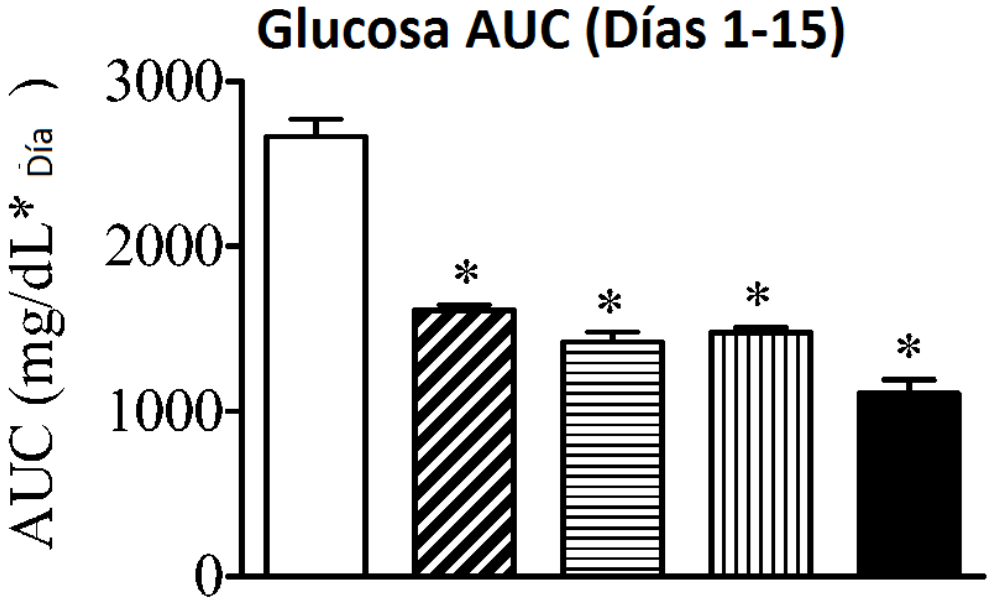


FIG 4

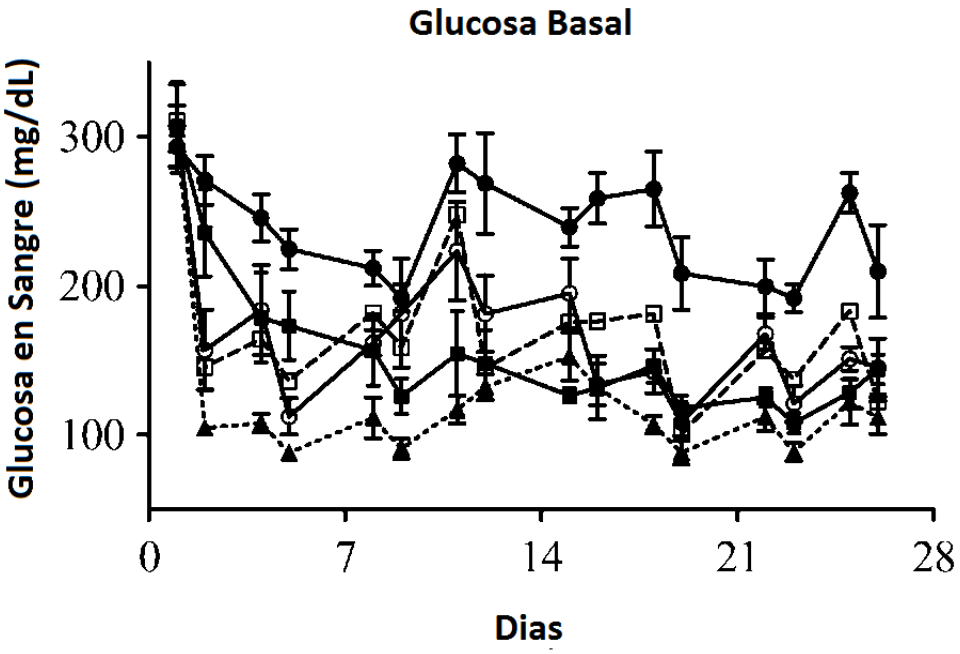

Glucosa Basal
Glucosa en Sangre (mg/dL)
300
200
100
0
7
14
21
28
Dias

FIG 5A

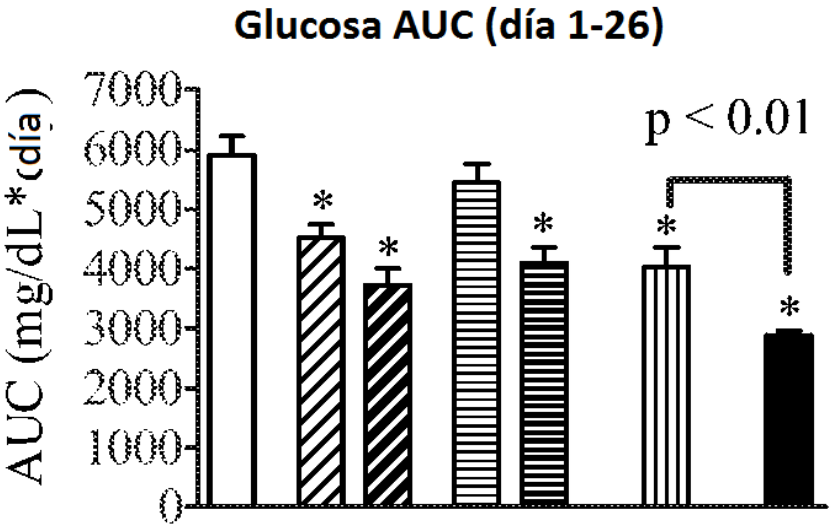
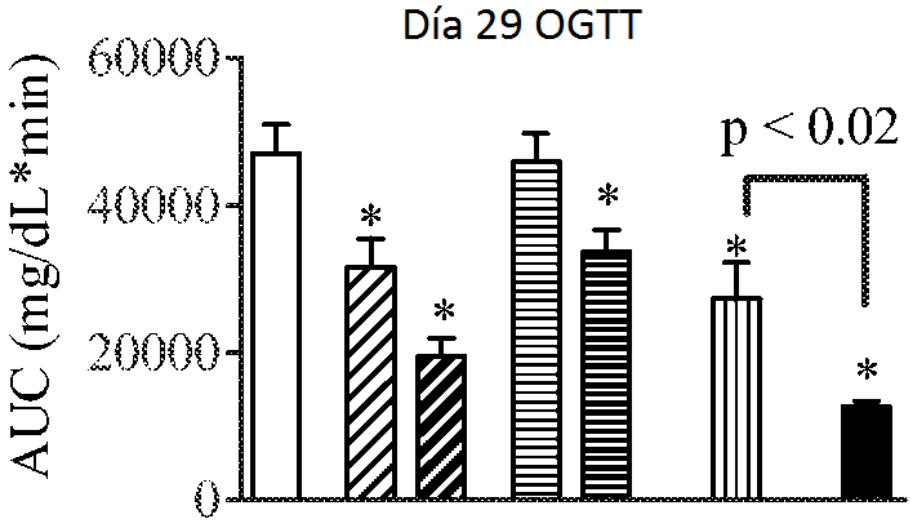
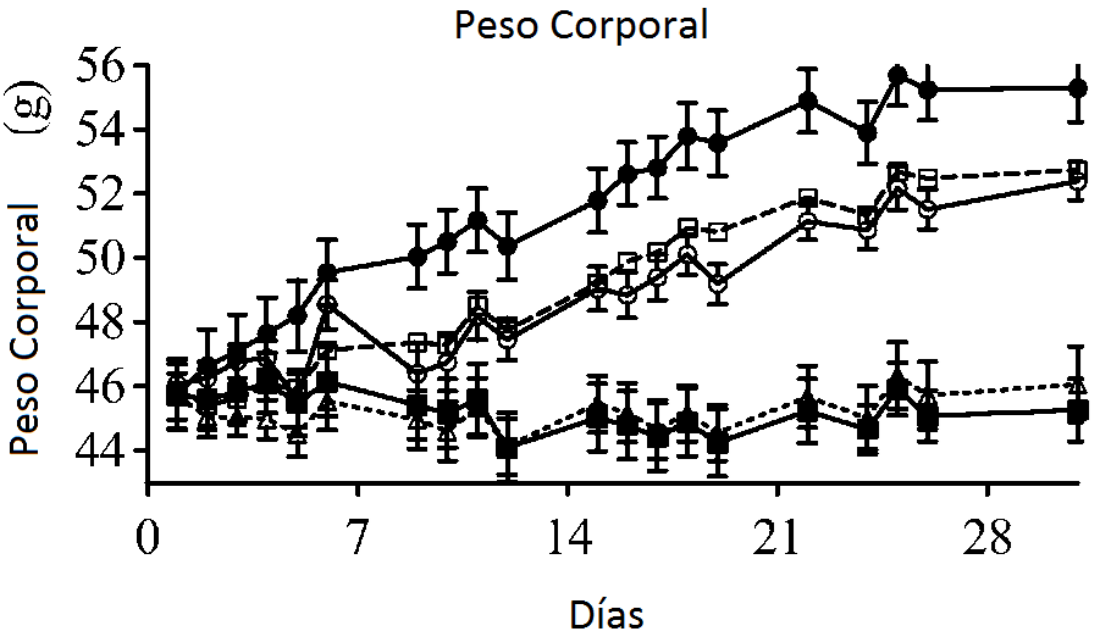
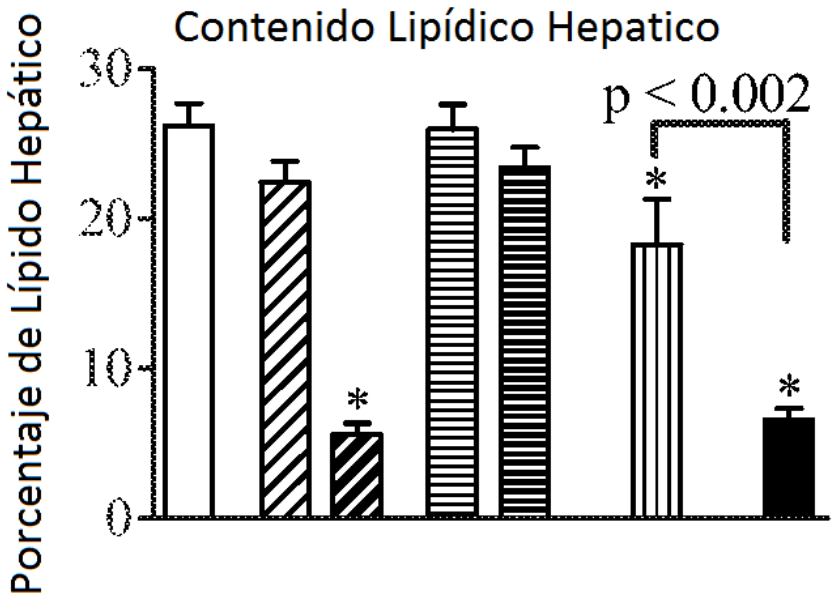

Glucosa AUC (día 1-26)
AUC (mg/dL*(día))
7000
6000
5000
4000
3000
2000
1000
0
p < 0.01
*

FIG 5B


Día 29 OGTT
AUC (mg/dL*min)
60000
40000
20000
0
p < 0.02
*
*
*
*
*


Peso Corporal
Peso Corporal (g)
56
54
52
50
48
46
44
0
7
14
21
28
Días


Contenido Lipídico Hepatico
Porcentaje de Lípido Hepático
30
20
10
0
p < 0.002
*
*
*

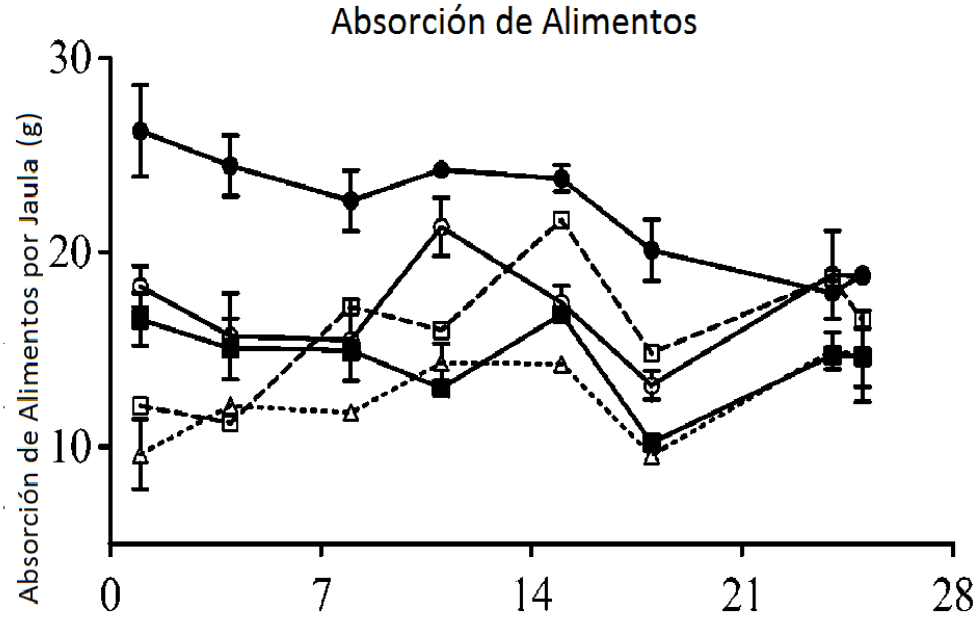

Absorción de Alimentos
Absorción de Alimentos por Jaula (g)
30
20
10
0
7
14
21
28

FIG 5F

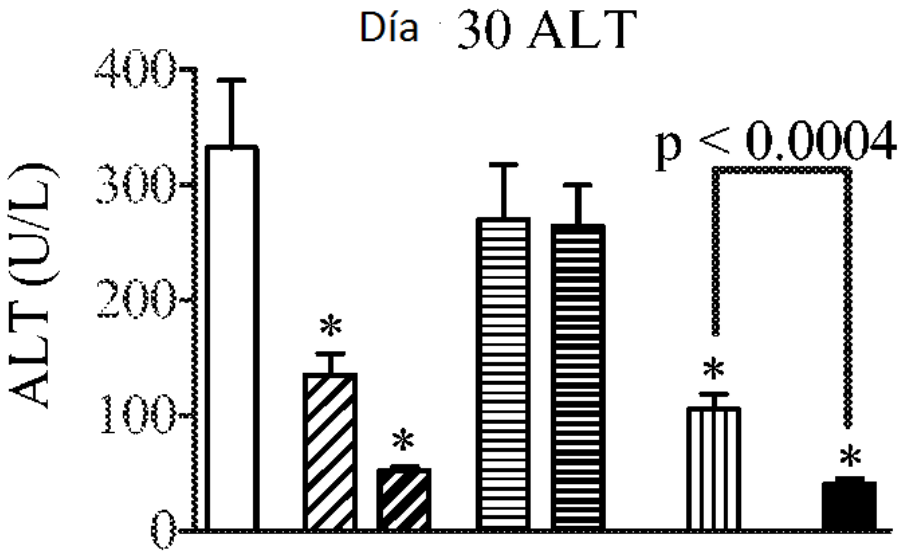


FIG 5G

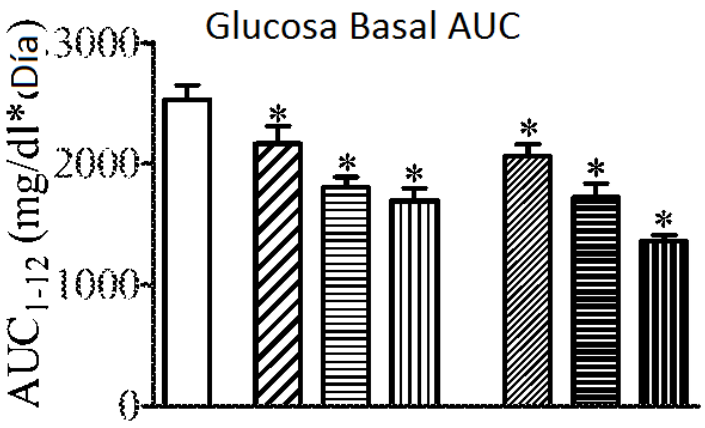

Glucosa Basal AUC
AUC1-12 (mg/dl*(Día)
3000
2000
1000
0
*
*
*
*
*
*

FIG 6A

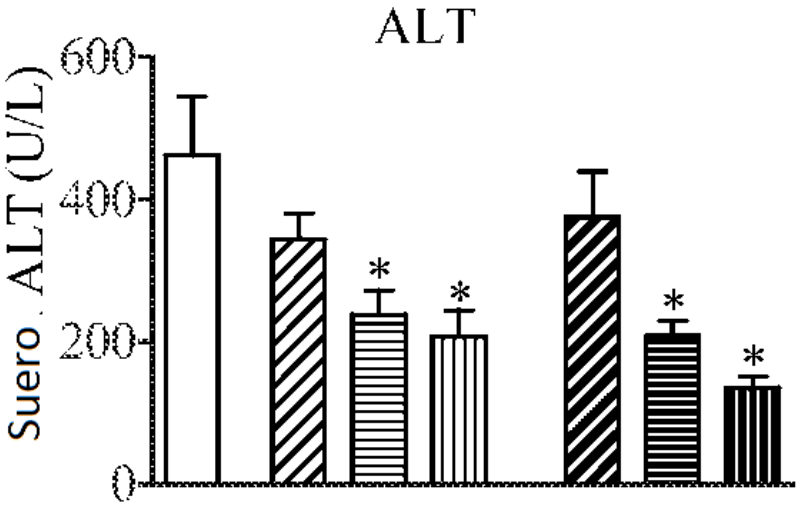

ALT
Suero. ALT (U/L)
600
400
200
0
*
*
*
*

FIG 6B

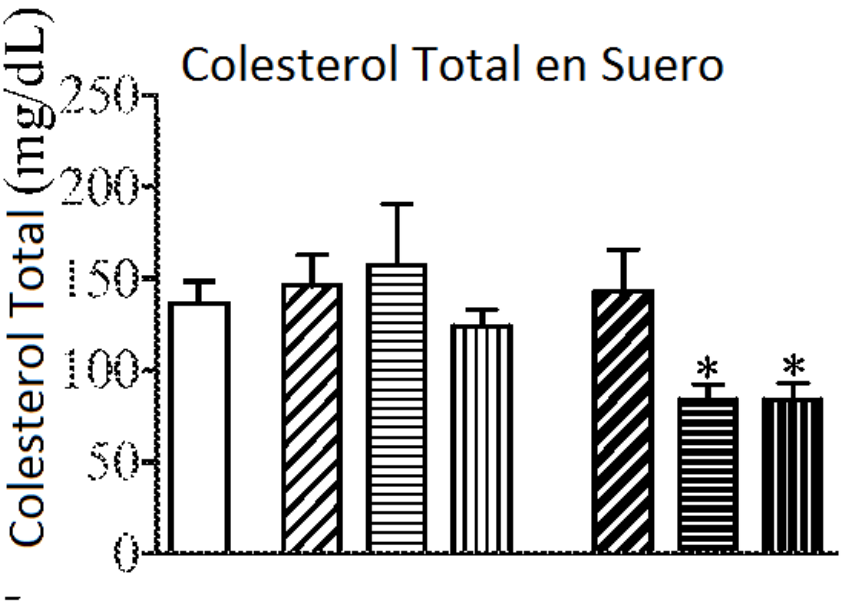

Colesterol Total en Suero
Colesterol Total (mg/dL)
250
200
150
100
50
0
*
*

FIG 6C

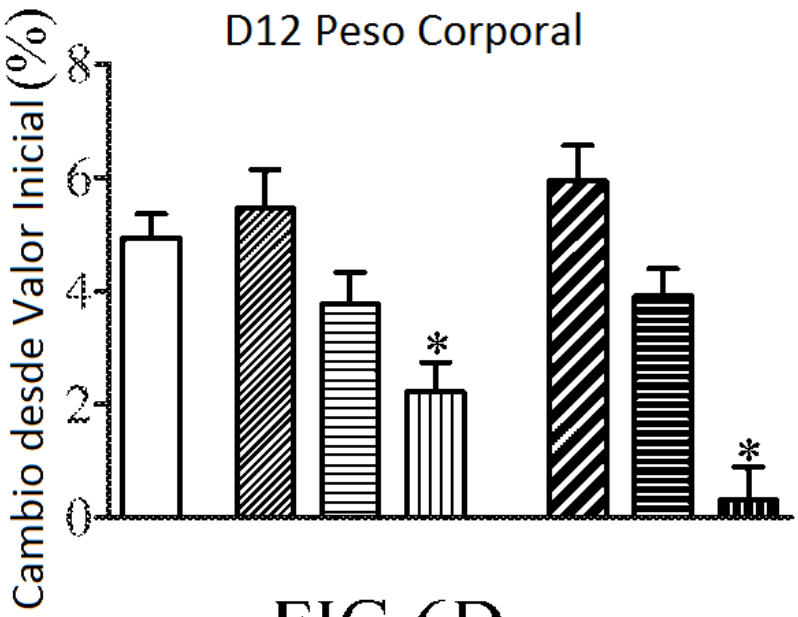

D12 Peso Corporal
Cambio desde Valor Inicial (%)
0
2
4
6
8
*
*

FIG 6D

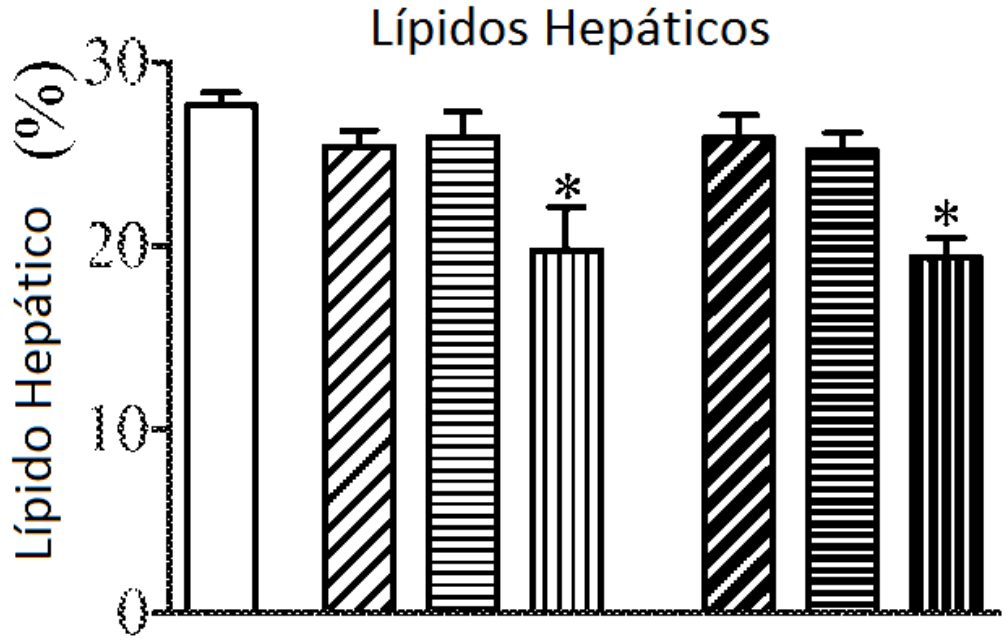

Lípidos Hepáticos
Lípido Hepático (%)
0
20
30
*
*

FIG 6E

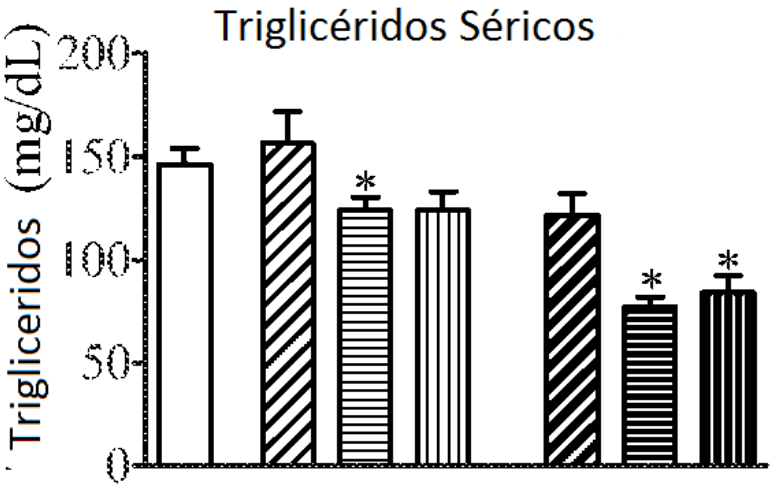


FIG 6F

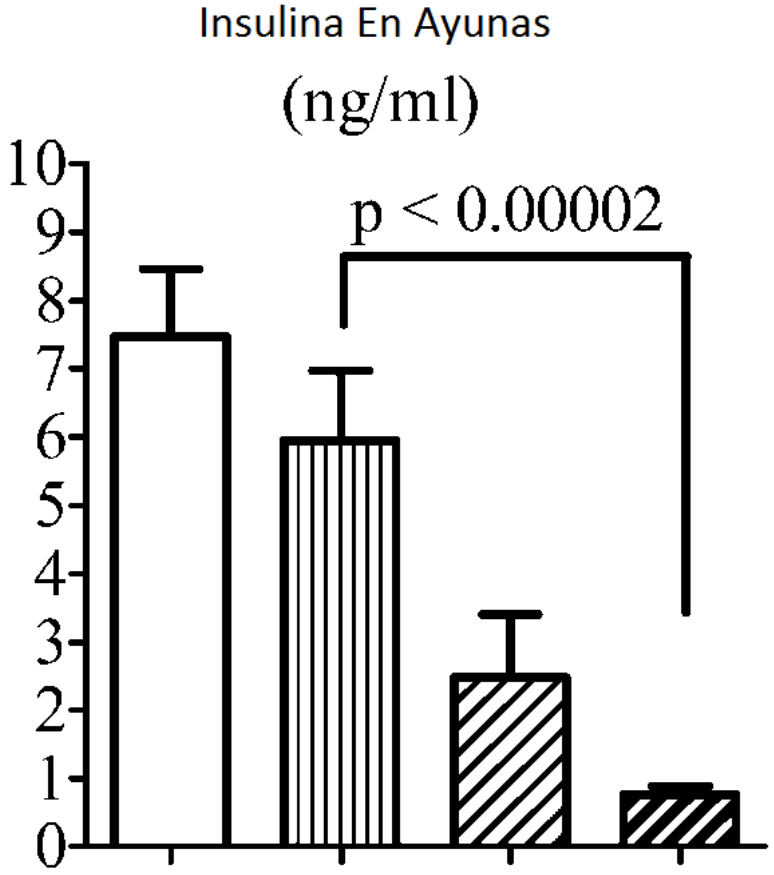


FIG 7A

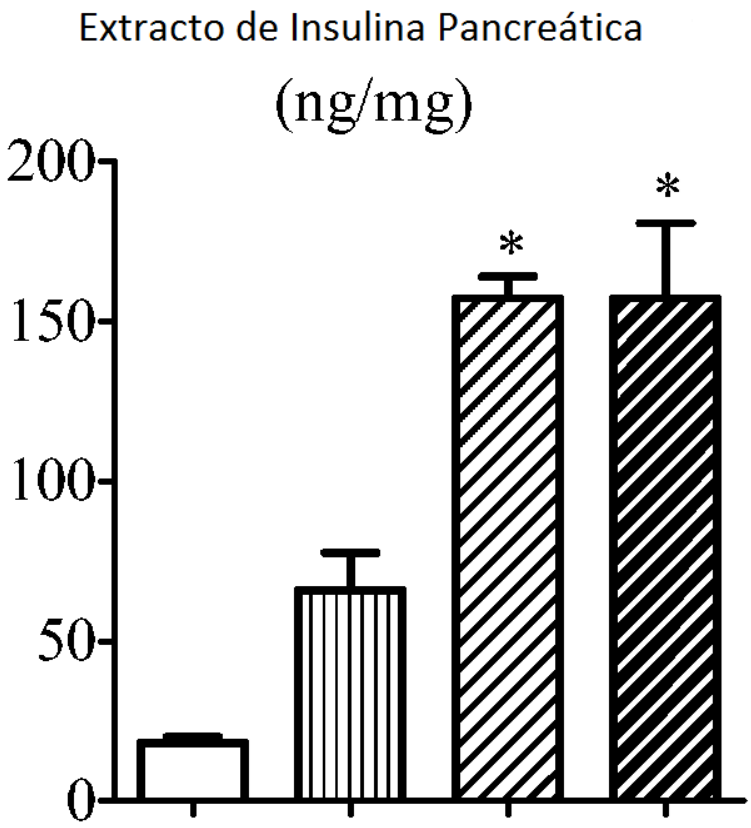


FIG 7B

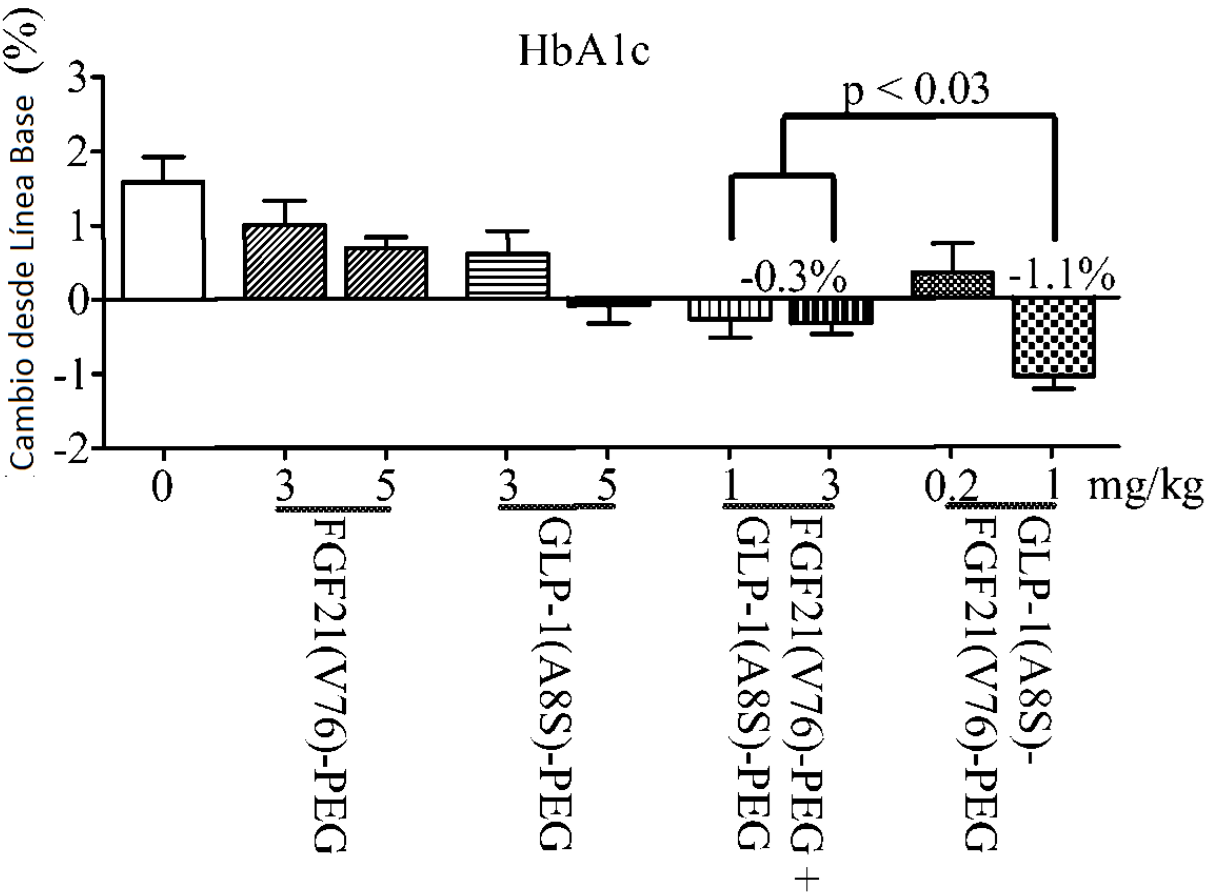

HbA1c
Cambio desde Línea Base (%)
3
2
1
0
-1
-2
p < 0.03
-0.3%
-1.1%
0
3
5
3
5
1
3
0.2
1
mg/kg
FGF21(V76)-PEG
GLP-1(A8S)-PEG
FGF21(V76)-PEG +
GLP-1(A8S)-PEG
FGF21(V76)-PEG
GLP-1(A8S)-

FIG 8A

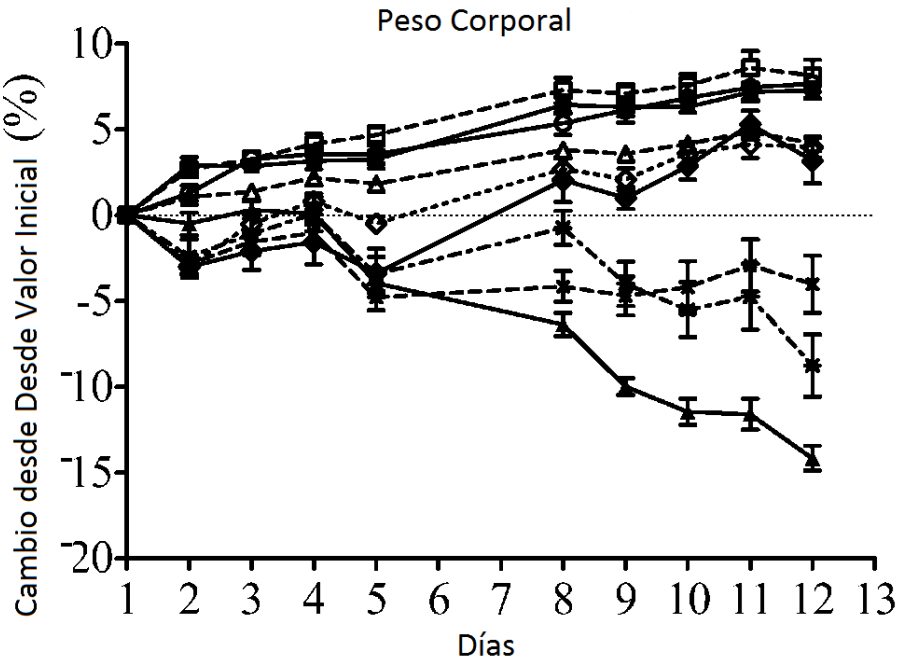

Peso Corporal
Cambio desde Desde Valor Inicial (%)
10
5
0
-5
-10
-15
-20
1 2 3 4 5 6 7 8 9 10 11 12 13
Días

FIG 8B

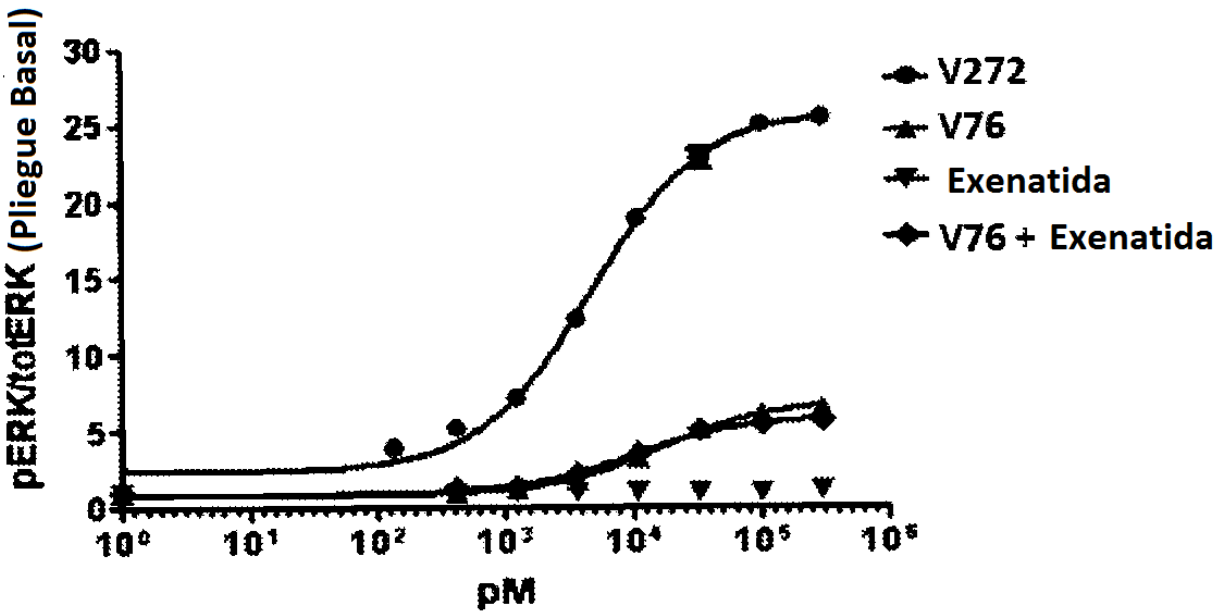

pERK/totERK (Pliegue Basal)
30
25
20
15
10
5
0
10⁰
10¹
10²
10³
10⁴
10⁵
10⁶
pM
V272
V76
Exenatida
V76 + Exenatida

FIG 9

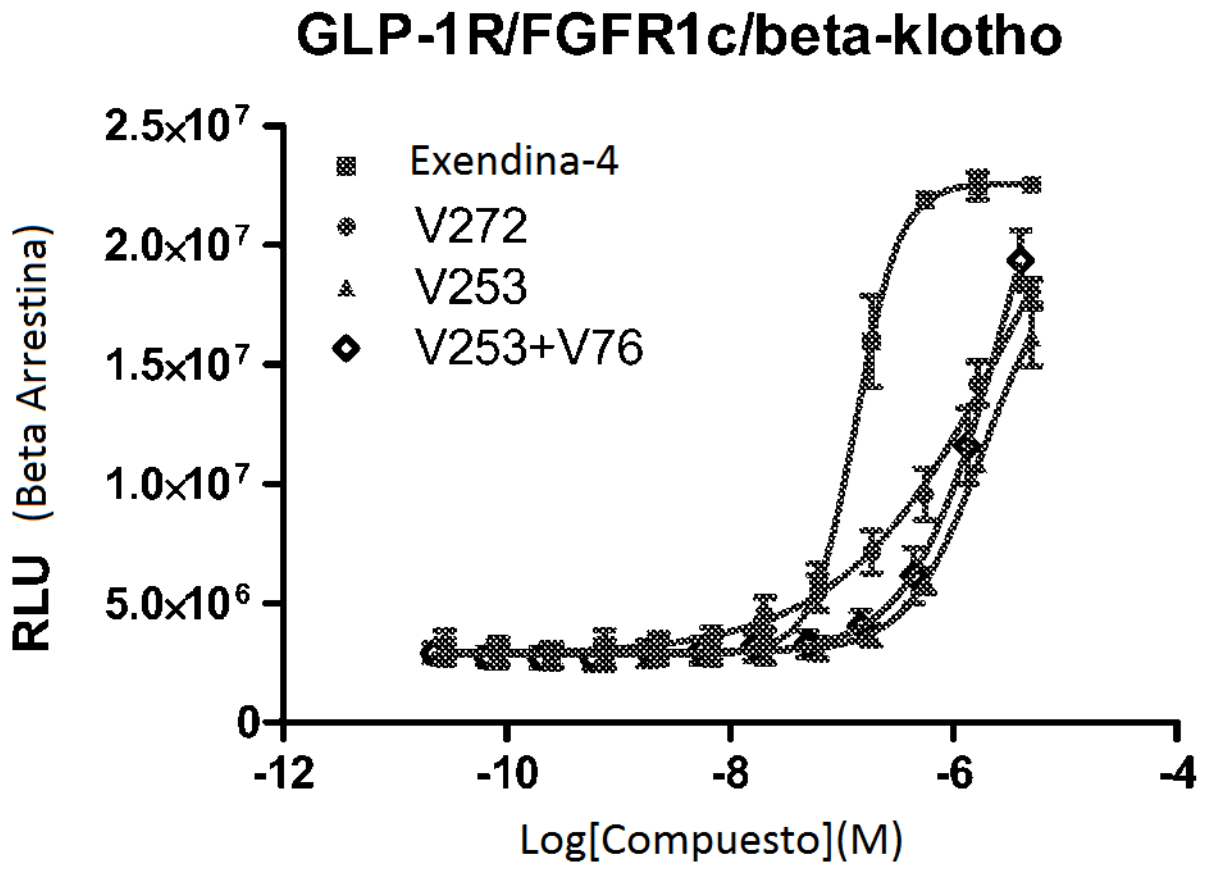

GLP-1R/FGFR1c/beta-klotho
RLU (Beta Arrestina)
2.5x10^7
2.0x10^7
1.5x10^7
1.0x10^7
5.0x10^6
0
Exendina-4
V272
V253
V253+V76
-12
-10
-8
-6
-4
Log[Compuesto](M)

FIG 10A

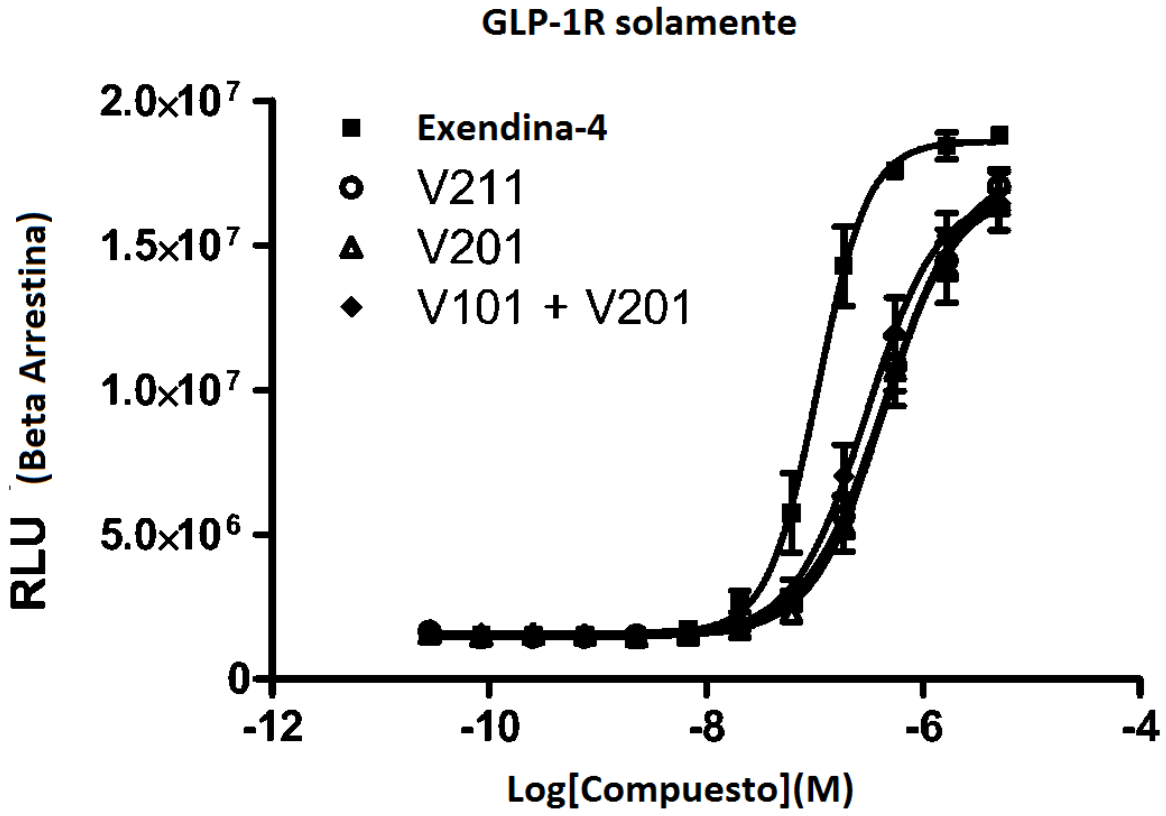

GLP-1R solamente
2.0x10^7
1.5x10^7
1.0x10^7
5.0x10^6
0
RLU (Beta Arrestina)
Exendina-4
V211
V201
V101 + V201
-12
-10
-8
-6
-4
Log[Compuesto](M)

FIG 10B

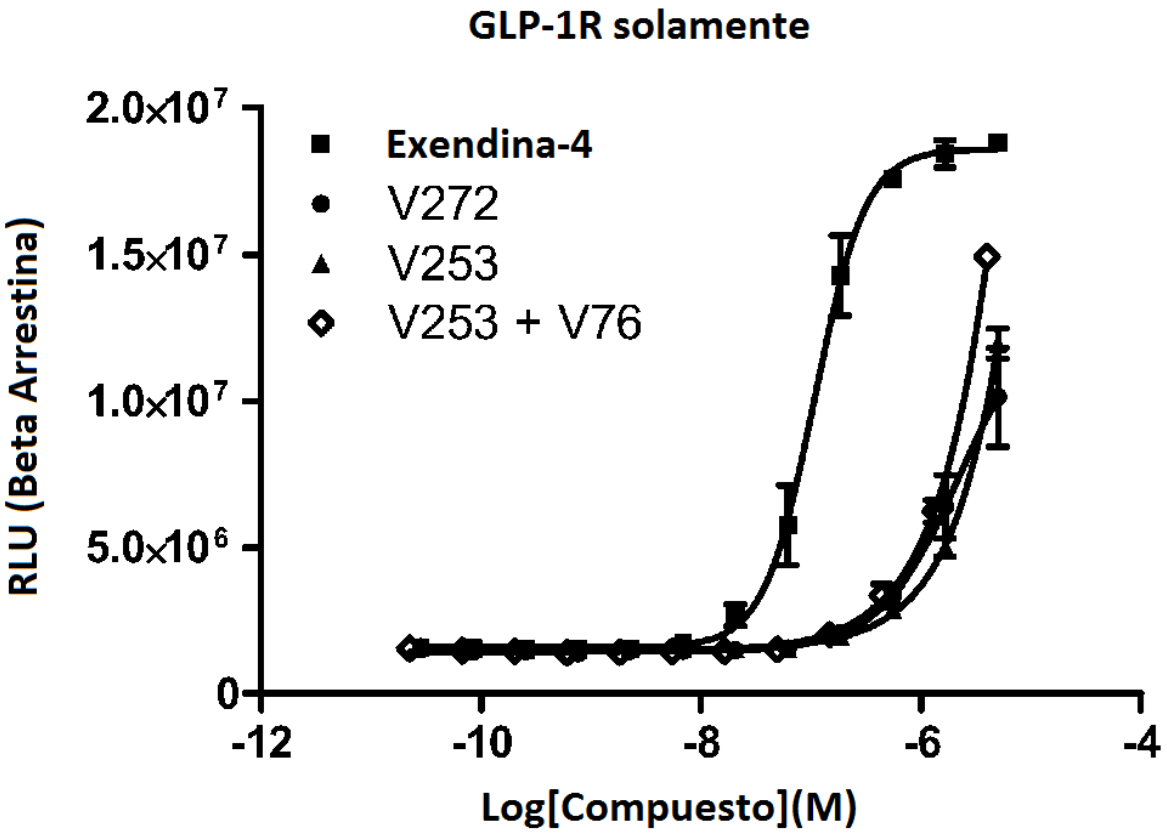

GLP-1R solamente
2.0x10^7
1.5x10^7
1.0x10^7
5.0x10^6
0
RLU (Beta Arrestina)
Exendina-4
V272
V253
V253 + V76
-12
-10
-8
-6
-4
Log[Compuesto](M)

FIG 10C

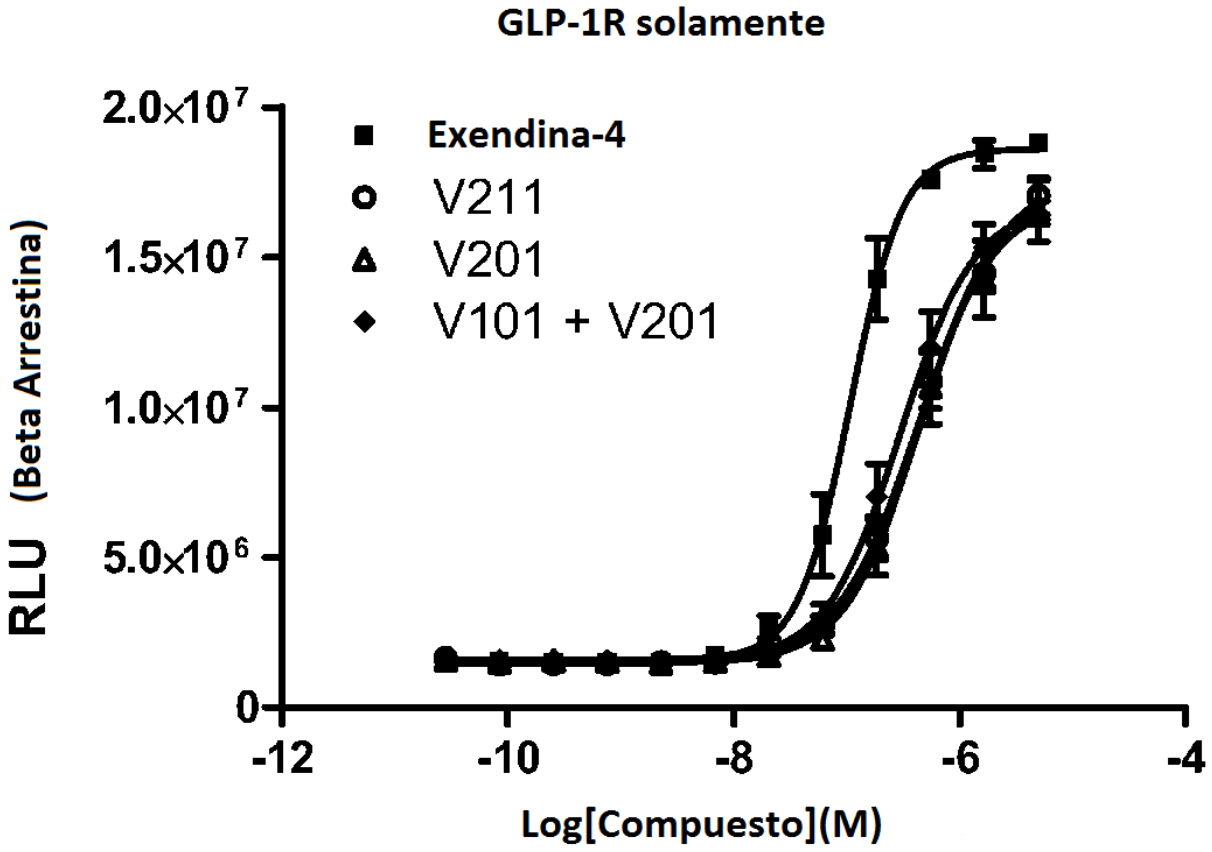

GLP-1R solamente
Exendina-4
V211
V201
V101 + V201
RLU (Beta Arrestina)
2.0×10^7
1.5×10^7
1.0×10^7
5.0×10^6
0
-12
-10
-8
-6
-4
Log[Compuesto](M)

FIG 10D

140325 Spanish translation Sequence listing

LISTADO DE SECUENCIAS

<110> Geierstanger, Bernhard
Cellitti, Susan
Loew, Andreas
Boettcher, Brian
Daniels, Doug
Caplan, Shari
Hamamatsu, Norio
Weldon, Stephen
Licht, Stuart

<120> PROTEINAS DE FUNCIÓN DUAL PARA EL TRATAMIENTO DE TRASTORNOS METABÓLICOS

<130> PAT054621

<160> 132

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 209
<212> PRT
<213> Homo sapiens

<400> 1

```
Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Ser
 1               5                  10                  15
Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala His Pro Ile Pro
            20                  25                  30
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
        35                  40                  45
Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
    50                  55                  60
Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
65                  70                  75                  80
Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
                85                  90                  95
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
            100                 105                 110
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
        115                 120                 125
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
    130                 135                 140
His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly
145                 150                 155                 160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
                165                 170                 175
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
            180                 185                 190
Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
        195                 200                 205
Ser
```

<210> 2
<211> 940
<212> ADN
<213> Homo sapiens

<400> 2

ctgtcagctg aggatccagc cgaaagagga gccaggcact caggccacct gagtctactc 60
acctggacaa ctggaatctg gcaccaattc taaaccactc agcttctccg agctcacacc 120
ccggagatca cctgaggacc cgagccattg atggactcgg acgagaccgg gttcgagcac 180
tcaggactgt gggtttctgt gctggctggt cttctgctgg gagcctgcca ggcacacccc 240
atccctgact ccagtcctct cctgcaattc gggggccaag tccggcagcg gtacctctac 300
acagatgatg cccagcagac agaagcccac ctggagatca gggaggatgg gacggtgggg 360
ggcgctgctg accagagccc cgaaagtctc ctgcagctga aagccttgaa gccgggagtt 420
attcaaatct tgggagtcaa gacatccagg ttcctgtgcc agcggccaga tggggccctg 480
tatggatcgc tccactttga ccctgaggcc tgcagcttcc gggagctgct tcttgaggac 540
ggatacaatg tttaccagtc cgaagcccac ggcctcccgc tgcacctgcc agggaacaag 600
tccccacacc gggaccctgc accccgagga ccagctcgct tcctgccact accaggcctg 660
ccccccgcac tcccggagcc acccggaatc ctggcccccc agccccccga tgtgggctcc 720
tcggaccctc tgagcatggt gggaccttcc cagggccgaa gccccagcta cgcttcctga 780
agccagaggc tgtttactat gacatctcct ctttatttat taggttattt atcttattta 840
ttttttttatt tttcttactt gagataataa agagttccag aggagaaaaa aaaaaaaaaa 900
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa   940

<210> 3
<211> 181
<212> PRT
<213> Homo sapiens

<400> 3
His Pro Ile Pro Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
 1   5   10   15
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
            20   25   30
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
        35   40   45
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
    50   55   60
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly
65   70   75   80
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
                85   90   95
Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
            100   105   110
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro
        115   120   125
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
    130   135   140
Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
145   150   155   160
Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser
                165   170   175
Pro Ser Tyr Ala Ser
            180

<210> 4
<211> 546
<212> ADN
<213> Homo sapiens

<400> 4
caccccatcc ctgactccag tcctctcctg caattcgggg gccaagtccg gcagcggtac 60
ctctacacag atgatgccca gcagacagaa gcccacctgg agatcaggga ggatgggacg 120
gtgggggggcg ctgctgacca gagccccgaa agtctcctgc agctgaaagc cttgaagccg 180
ggagttattc aaatcttggg agtcaagaca tccaggttcc tgtgccagcg gccagatggg 240
gccctgtatg gatcgctcca ctttgaccct gaggcctgca gcttccggga gctgcttctt 300
gaggacggat acaatgttta ccagtccgaa gcccacggcc tcccgctgca cctgccaggg 360
aacaagtccc cacaccggga ccctgcaccc cgaggaccag ctcgcttcct gccactacca 420
ggcctgcccc ccgcactccc ggagccaccc ggaatcctgg ccccccagcc ccccgatgtg 480

ggctcctcgg accctctgag catggtggga ccttcccagg gccgaagccc cagctacgct 540
tcctga 546

<210> 5
<211> 180
<212> PRT
<213> Homo sapiens

<400> 5
Met Lys Ser Ile Tyr Phe Val Ala Gly Leu Phe Val Met Leu Val Gln
 1   5   10   15
Gly Ser Trp Gln Arg Ser Leu Gln Asp Thr Glu Glu Lys Ser Arg Ser
            20   25   30
Phe Ser Ala Ser Gln Ala Asp Pro Leu Ser Asp Pro Asp Gln Met Asn
        35   40   45
Glu Asp Lys Arg His Ser Gln Gly Thr Phe Thr Ser Asp Tyr Ser Lys
    50   55   60
Tyr Leu Asp Ser Arg Arg Ala Gln Asp Phe Val Gln Trp Leu Met Asn
65   70   75   80
Thr Lys Arg Asn Arg Asn Asn Ile Ala Lys Arg His Asp Glu Phe Glu
                85   90   95
Arg His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu
            100   105   110
Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
        115   120   125
Arg Arg Asp Phe Pro Glu Glu Val Ala Ile Val Glu Glu Leu Gly Arg
    130   135   140
Arg His Ala Asp Gly Ser Phe Ser Asp Glu Met Asn Thr Ile Leu Asp
145   150   155   160
Asn Leu Ala Ala Arg Asp Phe Ile Asn Trp Leu Ile Gln Thr Lys Ile
                165   170   175
Thr Asp Arg Lys
            180

<210> 6
<211> 1298
<212> ADN
<213> Homo sapiens

<400> 6
gcatagaatg cagatgagca aagtgagtgg gagagggaag tcatttgtaa caaaaactca 60
ttatttacag atgagaaatt tatattgtca gcgtaatatc tgtgaggcta aacagagctg 120
gagagtatat aaaagcagtg cgccttggtg cagaagtaca gagcttagga cacagagcac 180
atcaaaagtt cccaaagagg gcttgctctc tcttcacctg ctctgttcta cagcacacta 240
ccagaagaca gcagaaatga aaagcattta ctttgtggct ggattatttg taatgctggt 300
acaaggcagc tggcaacgtt cccttcaaga cacagaggag aaatccagat cattctcagc 360
ttcccaggca gacccactca gtgatcctga tcagatgaac gaggacaagc gccattcaca 420
gggcacattc accagtgact acagcaagta tctggactcc aggcgtgccc aagattttgt 480
gcagtggttg atgaatacca agaggaacag gaataacatt gccaaacgtc acgatgaatt 540
tgagagacat gctgaaggga cctttaccag tgatgtaagt tcttatttgg aaggccaagc 600
tgccaaggaa ttcattgctt ggctggtgaa aggccgagga aggcgagatt tcccagaaga 660
ggtcgccatt gttgaagaac ttggccgcag acatgctgat ggttctttct ctgatgagat 720
gaacaccatt cttgataatc ttgccgccag ggactttata aactggttga ttcagaccaa 780
aatcactgac aggaaataac tatatcacta ttcaagatca tcttcacaac atcacctgct 840
agccacgtgg gatgtttgaa atgttaagtc ctgtaaattt aagaggtgta ttctgaggcc 900
acattgcttt gcatgccaat aaataaattt tcttttagtg ttgtgtagcc aaaaattaca 960
aatggaataa agttttatca aaatattgct aaaatatcag ctttaaaata tgaaagtgct 1020
agattctgtt attttcttct tattttggat gaagtacccc aacctgttta catttagcga 1080
taaaattatt tttctatgat ataatttgta aatgtaaatt attccgatct gacatatctg 1140
cattataata ataggagaat agaagaactg gtagccacag tggtgaaatt ggaaagagaa 1200
ctttcttcct gaaacctttg tcttaaaaat actcagcttt caatgtatca aagatacaat 1260
taaataaaat tttcaagctt ctttaccaaa aaaaaaaa 1298

<210> 7
<211> 39
<212> PRT
<213> Homo sapiens

<400> 7
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1   5   10   15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
            20   25   30
Ser Gly Ala Pro Pro Pro Ser
        35

<210> 8
<211> 15
<212> ADN
<213> Homo sapiens

<400> 8
ggggsggggs ggggs 15

<210> 9
<211> 216
<212> PRT
<213> Homo sapiens

<400> 9
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145   150   155   160
Cys Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 10
<211> 651
<212> ADN
<213> Homo sapiens

<400> 10
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgat 120
agcagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
tgcgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 11
<211> 227
<212> PRT
<213> Homo sapiens

<400> 11
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1   5   10   15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
        35   40   45
Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln
    50   55   60
Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu
65   70   75   80
Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala His Gln Ser Pro Glu
                85   90   95
Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu
            100   105   110
Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp Gly Ala Leu
        115   120   125
Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu
    130   135   140
Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu
145   150   155   160
Pro Leu His Leu Pro Gly Asn Arg Ser Pro His Cys Asp Pro Ala Pro
                165   170   175
Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu
            180   185   190
Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser
        195   200   205
Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser
    210   215   220
Tyr Ala Ser
225

<210> 12
<211> 684
<212> ADN
<213> Homo sapiens

<400> 12
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt tctggtggtg gtggttctgg cggtggcggt 120
tctggcggcg gtggtagcgg tggcggcggt gatagcagcc cgctgctgca gtttggcggc 180
caggtgcgtc agcgttatct gtataccgat gatgcgcagg aaaccgaagc gcatctggaa 240
attcgtgaag atggcaccgt gggcggtgcg gcgcatcaga gcccggaaag cctgctggaa 300

ctgaaagcgc tgaaaccggg cgtgattcag attctgggcg tgaaaaccag ccgttttctg 360
tgccagaaac cggatggcgc gctgtatggc agcctgcatt ttgatccgga agcgtgcagc 420
tttcgtgaac tgctgctgga agatggctat aacgtgtatc agagcgaagc gcatggcctg 480
ccgctgcatc tgccgggcaa ccgtagcccg cattgcgatc cggcaccgca gggtccggcg 540
cgttttctgc cgctgccggg tctgccgccg gcactgccgg aaccgccggg tattctggcc 600
ccgcagccgc cggatgttgg tagcagcgat ccgctggcga tggtgggtcc gagccagggt 660
cgtagcccga gctatgcgag ctaa 684

<210> 13
<211> 209
<212> PRT
<213> Homo sapiens

<400> 13
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
        35   40   45
Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
    50   55   60
Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
65   70   75   80
Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
                85   90   95
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
            100   105   110
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
        115   120   125
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
    130   135   140
His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly
145   150   155   160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
                165   170   175
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
            180   185   190
Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
        195   200   205
Ser

<210> 14
<211> 630
<212> ADN
<213> Homo sapiens

<400> 14
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtgata gcagcccgct gctgcagttt 120
ggcggccagg tgcgtcagcg ttatctgtat accgatgatg cgcagcagac cgaagcgcat 180
ctggaaattc gtgaagatgg caccgtgggc ggtgcggcgg atcagagccc ggaaagcctg 240
ctgcagctga aagcgctgaa accgggcgtg attcagattc tgggcgtgaa aaccagccgt 300
tttctgtgcc agcgtccgga tggcgcgctg tatggcagcc tgcattttga tccggaagcg 360
tgcagctttc gtgaactgct gctggaagat ggctataacg tgtatcagag cgaagcgcat 420
ggcctgccgc tgcatctgcc gggcaacaaa agcccgcatc gtgatccggc accgcgtggt 480
ccggcgcgtt ttctgccgct gccgggtctg ccgccggcac tgccggaacc gccgggtatt 540
ctggccccgc agccgccgga tgttggtagc agcgatccgc tgtctatggt gggtccgagc 600
cagggtcgta gcccgagcta tgcgagctaa 630

<210> 15

<211> 209
<212> PRT
<213> Homo sapiens

<400> 15
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
        35   40   45
Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
    50   55   60
Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
65   70   75   80
Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
                85   90   95
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
            100   105   110
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
        115   120   125
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
    130   135   140
His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly
145   150   155   160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
                165   170   175
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
            180   185   190
Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
        195   200   205
Ser

<210> 16
<211> 630
<212> ADN
<213> Homo sapiens

<400> 16
catgcggaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtgata gcagcccgct gctgcagttt 120
ggcggccagg tgcgtcagcg ttatctgtat accgatgatg cgcagcagac cgaagcgcat 180
ctggaaattc gtgaagatgg caccgtgggc ggtgcggcgg atcagagccc ggaaagcctg 240
ctgcagctga aagcgctgaa accgggcgtg attcagattc tgggcgtgaa aaccagccgt 300
tttctgtgcc agcgtccgga tggcgcgctg tatggcagcc tgcattttga tccggaagcg 360
tgcagctttc gtgaactgct gctggaagat ggctataacg tgtatcagag cgaagcgcat 420
ggcctgccgc tgcatctgcc gggcaacaaa agcccgcatc gtgatccggc accgcgtggt 480
ccggcgcgtt ttctgccgct gccgggtctg ccgccggcac tgccggaacc gccgggtatt 540
ctggccccgc agccgccgga tgttggtagc agcgatccgc tgtctatggt gggtccgagc 600
cagggtcgta gcccgagcta tgcgagctaa   630

<210> 17
<211> 219
<212> PRT
<213> Homo sapiens

<400> 17
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1   5   10   15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
            20   25   30

```
Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu
        35                  40                  45
Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala
    50                  55                  60
Gln Gln Thr Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly
65                  70                  75                  80
Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu
                85                  90                  95
Lys Pro Gly Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu
            100                 105                 110
Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro
        115                 120                 125
Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val
    130                 135                 140
Tyr Gln Ser Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys
145                 150                 155                 160
Ser Pro His Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro
                165                 170                 175
Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala
            180                 185                 190
Pro Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly
        195                 200                 205
Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215
```

<210> 18
<211> 660
<212> ADN
<213> Homo sapiens

<400> 18
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt ccgtcctccg gcgctcctcc gccttctggt 120
ggtggcgact cgagcccgct gctgcagttt ggcggccagg tgcgtcagcg ttatctgtat 180
accgatgatg cgcagcagac cgaagcgcat ctggaaattc gtgaagatgg caccgtgggc 240
ggtgcggcgg atcagagccc ggaaagcctg ctgcagctga aagcgctgaa accgggcgtg 300
attcagattc tgggcgtgaa aaccagccgt tttctgtgcc agcgtccgga tggcgcgctg 360
tatggcagcc tgcattttga tccggaagcg tgcagctttc gtgaactgct gctggaagat 420
ggctataacg tgtatcagag cgaagcgcat ggcctgccgc tgcatctgcc gggcaacaaa 480
agcccgcatc gtgatccggc accgcgtggt ccggcgcgtt ttctgccgct gccgggtctg 540
ccgccggcac tgccggaacc gccgggtatt ctggccccgc agccgccgga tgttggtagc 600
agcgatccgc tgtctatggt gggtccgagc cagggtcgta gcccgagcta tgcgagctaa 660

<210> 19
<211> 216
<212> PRT
<213> Homo sapiens

<400> 19
```
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
```

```
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145                 150                 155                 160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165                 170                 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180                 185                 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195                 200                 205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215
```

<210> 20
<211> 651
<212> ADN
<213> Homo sapiens

<400> 20
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a   651

<210> 21
<211> 216
<212> PRT
<213> Homo sapiens

<400> 21
```
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145                 150                 155                 160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
```

```
                165                 170                 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180                 185                 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195                 200                 205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215
```

<210> 22
<211> 651
<212> ADN
<213> Homo sapiens

<400> 22
catgcggaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a   651

<210> 23
<211> 224
<212> PRT
<213> Homo sapiens

<400> 23
```
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1               5                  10                  15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
            20                  25                  30
Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Asp
        35                  40                  45
Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu
    50                  55                  60
Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg Glu
65                  70                  75                  80
Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu Leu
                85                  90                  95
Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val Lys
            100                 105                 110
Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly Ser
        115                 120                 125
Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu
    130                 135                 140
Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu His
145                 150                 155                 160
Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala Pro Arg Gly Pro
                165                 170                 175
Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro
            180                 185                 190
Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp Pro
        195                 200                 205
Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215                 220
```

<210> 24
<211> 675
<212> ADN
<213> Homo sapiens

<400> 24
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt ccgtcctccg gcgctcctcc gccttctggt 120
ggtggtggtt ctggcggtgg cgactcgagc ccgctgctgc agtttggcgg ccaggtgcgt 180
cagcgttatc tgtataccga tgatgcgcag cagaccgaag cgcatctgga aattcgtgaa 240
gatggcaccg tgggcggtgc ggcggatcag agcccggaaa gcctgctgca gctgaaagcg 300
ctgaaaccgg gcgtgattca gattctgggc gtgaaaacca gccgttttct gtgccagcgt 360
ccggatggcg cgctgtatgg cagcctgcat tttgatccgg aagcgtgcag ctttcgtgaa 420
ctgctgctgg aagatggcta taacgtgtat cagagcgaag cgcatggcct gccgctgcat 480
ctgccgggca acaaaagccc gcatcgtgat ccggcaccgc gtggtccggc gcgttttctg 540
ccgctgccgg gtctgccgcc ggcactgccg gaaccgccgg gtattctggc cccgcagccg 600
ccggatgttg gtagcagcga tccgctgtct atggtgggtc cgagccaggg tcgtagcccg 660
agctatgcga gctaa 675

<210> 25
<211> 486
<212> PRT
<213> Homo sapiens

<400> 25
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
 1   5   10   15
Gly Ser Thr Gly His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser
            20   25   30
Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys
        35   40   45
Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
    50   55   60
Gly Ser Ala Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu
65   70   75   80
Ala Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
                85   90   95
Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
            100   105   110
Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly
        115   120   125
Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn
    130   135   140
Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp
145   150   155   160
Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro
                165   170   175
Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu
            180   185   190
Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn
        195   200   205
Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
    210   215   220
Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
225   230   235   240
Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys
                245   250   255
Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
            260   265   270
Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
        275   280   285
Ser Leu Ser Pro Gly Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
    290   295   300

Gly Gly Gly Gly Ser Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln
305 310 315 320
Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Cys Gln Thr Glu Ala
325 330 335
His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln
340 345 350
Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile
355 360 365
Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp
370 375 380
Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe
385 390 395 400
Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala
405 410 415
His Gly Leu Pro Leu His Leu Pro Cys Asn Arg Ser Pro His Arg Asp
420 425 430
Pro Ala Ser Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro
435 440 445
Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp
450 455 460
Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Gly Ser Gln Ala Arg
465 470 475 480
Ser Pro Ser Tyr Ala Ser
485

<210> 26
<211> 486
<212> PRT
<213> Homo sapiens

<400> 26
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15
Gly Ser Thr Gly His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser
20 25 30
Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys
35 40 45
Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
50 55 60
Gly Ser Ala Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu
65 70 75 80
Ala Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
85 90 95
Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
100 105 110
Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly
115 120 125
Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn
130 135 140
Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp
145 150 155 160
Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro
165 170 175
Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu
180 185 190
Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn
195 200 205
Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
210 215 220
Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
225 230 235 240
Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys

245 250 255
Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
260 265 270
Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
275 280 285
Ser Leu Ser Pro Gly Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
290 295 300
Gly Gly Gly Gly Ser Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln
305 310 315 320
Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Cys Gln Thr Glu Ala
325 330 335
His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln
340 345 350
Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile
355 360 365
Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp
370 375 380
Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe
385 390 395 400
Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala
405 410 415
His Gly Leu Pro Leu His Leu Pro Cys Asn Arg Ser Pro His Arg Asp
420 425 430
Pro Ala Ser Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro
435 440 445
Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp
450 455 460
Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln Ala Arg
465 470 475 480
Ser Pro Ser Tyr Glu Ser
485

<210> 27
<211> 486
<212> PRT
<213> Homo sapiens

<400> 27
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15
Gly Ser Thr Gly His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser
20 25 30
Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys
35 40 45
Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
50 55 60
Gly Ser Ala Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu
65 70 75 80
Ala Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
85 90 95
Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
100 105 110
Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly
115 120 125
Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn
130 135 140
Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp
145 150 155 160
Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro
165 170 175
Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu
180 185 190

```
Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn
        195                 200                 205
Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
    210                 215                 220
Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
225                 230                 235                 240
Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys
                245                 250                 255
Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
            260                 265                 270
Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
        275                 280                 285
Ser Leu Ser Pro Gly Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
    290                 295                 300
Gly Gly Gly Gly Ser Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln
305                 310                 315                 320
Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Cys Gln Thr Glu Ala
                325                 330                 335
His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln
            340                 345                 350
Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile
        355                 360                 365
Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp
    370                 375                 380
Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe
385                 390                 395                 400
Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala
                405                 410                 415
His Gly Leu Pro Leu His Leu Pro Cys Asn Arg Ser Pro His Arg Asp
            420                 425                 430
Pro Ala Ser Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro
        435                 440                 445
Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp
    450                 455                 460
Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln Ala Arg
465                 470                 475                 480
Ser Pro Ser Tyr Asp Ser
                485
```

<210> 28
<211> 216
<212> PRT
<213> Homo sapiens

<400> 28

```
His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
```

130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145 150 155 160
Cys Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 29
<211> 651
<212> ADN
<213> Homo sapiens

<400> 29
catggtgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgat 120
agcagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
tgcgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 30
<211> 216
<212> PRT
<213> Homo sapiens

<400> 30
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Gly
1 5 10 15
Gln Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Ser Gly
20 25 30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
35 40 45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
50 55 60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65 70 75 80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
85 90 95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
100 105 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
115 120 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145 150 155 160
Cys Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
195 200 205

Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 31
<211> 651
<212> ADN
<213> Homo sapiens

<400> 31
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagggca agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt tctggtggtg gtggttctgg cggtggcgat 120
agcagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
tgcgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 32
<211> 216
<212> PRT
<213> Homo sapiens

<400> 32
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Ser Gly
20 25 30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
35 40 45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
50 55 60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65 70 75 80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
85 90 95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
100 105 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
115 120 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145 150 155 160
Cys Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 33
<211> 651
<212> ADN
<213> Homo sapiens

<400> 33
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt tctggtggtg gtggttctgg cggtggcgat 120
agcagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
tgcgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 34
<211> 216
<212> PRT
<213> Homo sapiens

<220>
<223> aminoácido no natural

<400> 34
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Xaa Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 35
<211> 651
<212> ADN
<213> Homo sapiens

<400> 35
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300

ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
taggatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 36
<211> 216
<212> PRT
<213> Homo sapiens

<220>
<223> aminoácido no natural

<400> 36
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Xaa Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 37
<211> 651
<212> ADN
<213> Homo sapiens

<400> 37
catgcggaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
taggatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 38
<211> 224
<212> PRT
<213> Homo sapiens

<220>
<223> aminoácido no natural

<400> 38
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1   5   10   15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
            20   25   30
Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Asp
        35   40   45
Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu
    50   55   60
Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg Glu
65   70   75   80
Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu Leu
                85   90   95
Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val Lys
            100   105   110
Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly Ser
        115   120   125
Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu
    130   135   140
Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu His
145   150   155   160
Leu Pro Gly Asn Lys Ser Pro His Xaa Asp Pro Ala Pro Arg Gly Pro
                165   170   175
Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro
            180   185   190
Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp Pro
        195   200   205
Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215   220

<210> 39
<211> 675
<212> ADN
<213> Homo sapiens

<400> 39
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt ccgtcctccg gcgctcctcc gccttctggt 120
ggtggtggtt ctggcggtgg cgactcgagc ccgctgctgc agtttggcgg ccaggtgcgt 180
cagcgttatc tgtataccga tgatgcgcag cagaccgaag cgcatctgga aattcgtgaa 240
gatggcaccg tgggcggtgc ggcggatcag agcccggaaa gcctgctgca gctgaaagcg 300
ctgaaaccgg gcgtgattca gattctgggc gtgaaaacca gccgttttct gtgccagcgt 360
ccggatggcg cgctgtatgg cagcctgcat tttgatccgg aagcgtgcag ctttcgtgaa 420
ctgctgctgg aagatggcta taacgtgtat cagagcgaag cgcatggcct gccgctgcat 480
ctgccgggca acaaaagccc gcattaggat ccggcaccgc gtggtccggc gcgttttctg 540
ccgctgccgg gtctgccgcc ggcactgccg gaaccgccgg gtattctggc cccgcagccg 600
ccggatgttg gtagcagcga tccgctgtct atggtgggtc cgagccaggg tcgtagcccg 660
agctatgcga gctaa   675

<210> 40
<211> 216
<212> PRT
<213> Homo sapiens

<400> 40
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 41
<211> 651
<212> ADN
<213> Homo sapiens

<400> 41
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a   651

<210> 42
<211> 216
<212> PRT
<213> Homo sapiens

<400> 42
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr

50 55 60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65 70 75 80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
85 90 95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
100 105 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
115 120 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145 150 155 160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 43
<211> 651
<212> ADN
<213> Homo sapiens

<400> 43
catgcggaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 44
<211> 224
<212> PRT
<213> Homo sapiens

<400> 44
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30
Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Asp
35 40 45
Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu
50 55 60
Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg Glu
65 70 75 80
Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu Leu
85 90 95
Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val Lys
100 105 110
Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly Ser
115 120 125

```
Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu
    130                 135                 140
Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu His
145                 150                 155                 160
Leu Pro Gly Asn Lys Ser Pro His Cys Asp Pro Ala Pro Arg Gly Pro
                165                 170                 175
Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro
            180                 185                 190
Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp Pro
        195                 200                 205
Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215                 220
```

<210> 45
<211> 675
<212> ADN
<213> Homo sapiens

<400> 45
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt ccgtcctccg gcgctcctcc gccttctggt 120
ggtggtggtt ctggcggtgg cgactcgagc ccgctgctgc agtttggcgg ccaggtgcgt 180
cagcgttatc tgtataccga tgatgcgcag cagaccgaag cgcatctgga aattcgtgaa 240
gatggcaccg tgggcggtgc ggcggatcag agcccggaaa gcctgctgca gctgaaagcg 300
ctgaaaccgg gcgtgattca gattctgggc gtgaaaacca gccgttttct gtgccagcgt 360
ccggatggcg cgctgtatgg cagcctgcat tttgatccgg aagcgtgcag ctttcgtgaa 420
ctgctgctgg aagatggcta taacgtgtat cagagcgaag cgcatggcct gccgctgcat 480
ctgccgggca acaaaagccc gcattgcgat ccggcaccgc gtggtccggc gcgttttctg 540
ccgctgccgg gtctgccgcc ggcactgccg gaaccgccgg gtattctggc cccgcagccg 600
ccggatgttg gtagcagcga tccgctgtct atggtgggtc cgagccaggg tcgtagcccg 660
agctatgcga gctaa   675

<210> 46
<211> 216
<212> PRT
<213> Homo sapiens

<400> 46
```
His Ala Pro Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145                 150                 155                 160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165                 170                 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180                 185                 190
```

Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 47
<211> 651
<212> ADN
<213> Homo sapiens

<400> 47
catgcgccgg gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 48
<211> 217
<212> PRT
<213> Homo sapiens

<400> 48
Gly His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu
1 5 10 15
Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser
20 25 30
Gly Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe
35 40 45
Gly Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln
50 55 60
Thr Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala
65 70 75 80
Ala Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro
85 90 95
Gly Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln
100 105 110
Arg Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala
115 120 125
Cys Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln
130 135 140
Ser Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro
145 150 155 160
His Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro
165 170 175
Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln
180 185 190
Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser
195 200 205
Gln Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 49
<211> 654
<212> ADN

<213> Homo sapiens

<400> 49
ggtcatgcgg aaggcacttt tactagcgat gtttctagct acctggaagg ccaggctgcg 60
aaagaattca tcgcgtggct ggttaaaggc ggttctggtg gtggtggttc tggcggtggc 120
gactcgagcc cgctgctgca gtttggcggc caggtgcgtc agcgttatct gtataccgat 180
gatgcgcagc agaccgaagc gcatctggaa attcgtgaag atggcaccgt gggcggtgcg 240
gcggatcaga gcccggaaag cctgctgcag ctgaaagcgc tgaaaccggg cgtgattcag 300
attctgggcg tgaaaaccag ccgttttctg tgccagcgtc cggatggcgc gctgtatggc 360
agcctgcatt ttgatccgga agcgtgcagc tttcgtgaac tgctgctgga agatggctat 420
aacgtgtatc agagcgaagc gcatggcctg ccgctgcatc tgccgggcaa caaaagcccg 480
cattgcgatc cggcaccgcg tggtccggcg cgttttctgc cgctgccggg tctgccgccg 540
gcactgccgg aaccgccggg tattctggcc ccgcagccgc cggatgttgg tagcagcgat 600
ccgctgtcta tggtgggtcc gagccagggt cgtagcccga gctatgcgag ctaa   654

<210> 50
<211> 216
<212> PRT
<213> Homo sapiens

<400> 50
His Ser Glu Gly Thr Phe Thr Ala Asp Ala Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 51
<211> 651
<212> ADN
<213> Homo sapiens

<400> 51
cattctgaag gcacttttac tgctgatgct tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360

ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 52
<211> 216
<212> PRT
<213> Homo sapiens

<400> 52
His Ser Glu Gly Thr Phe Thr Ser Asp Ala Ala Ala Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 53
<211> 651
<212> ADN
<213> Homo sapiens

<400> 53
cattctgaag gcacttttac tagcgatgct gctgcttacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 54
<211> 216
<212> PRT

<213> Homo sapiens

<400> 54
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Ala Glu Gly
 1   5   10   15
Ala Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 55
<211> 651
<212> ADN
<213> Homo sapiens

<400> 55
cattctgaag gcacttttac tagcgatgtt tctagctacg ctgaaggcgc tgctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a   651

<210> 56
<211> 216
<212> PRT
<213> Homo sapiens

<400> 56
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Ala Ile Ala Ala Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45

Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 57
<211> 651
<212> ADN
<213> Homo sapiens

<400> 57
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaagctatcg cggctctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a   651

<210> 58
<211> 185
<212> PRT
<213> Homo sapiens

<400> 58
His Ser Glu Gly Thr Phe Thr Ser Asp Ser Ser Pro Leu Leu Gln Phe
 1   5   10   15
Gly Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln
            20   25   30
Thr Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala
        35   40   45
Ala Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro
    50   55   60
Gly Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln
65   70   75   80
Arg Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala
                85   90   95
Cys Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln
            100   105   110
Ser Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro

115 120 125
His Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro
130 135 140
Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln
145 150 155 160
Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser
165 170 175
Gln Gly Arg Ser Pro Ser Tyr Ala Ser
180 185

<210> 59
<211> 558
<212> ADN
<213> Homo sapiens

<400> 59
cattctgaag gcacttttac tagcgatagc agcccgctgc tgcagtttgg cggccaggtg 60
cgtcagcgtt atctgtatac cgatgatgcg cagcagaccg aagcgcatct ggaaattcgt 120
gaagatggca ccgtgggcgg tgcggcggat cagagcccgg aaagcctgct gcagctgaaa 180
gcgctgaaac cgggcgtgat tcagattctg ggcgtgaaaa ccagccgttt tctgtgccag 240
cgtccggatg gcgcgctgta tggcagcctg cattttgatc cggaagcgtg cagctttcgt 300
gaactgctgc tggaagatgg ctataacgtg tatcagagcg aagcgcatgg cctgccgctg 360
catctgccgg gcaacaaaag cccgcattgc gatccggcac cgcgtggtcc ggcgcgtttt 420
ctgccgctgc cgggtctgcc gccggcactg ccggaaccgc cgggtattct ggccccgcag 480
ccgccggatg ttggtagcag cgatccgctg tctatggtgg gtccgagcca gggtcgtagc 540
ccgagctatg cgagctaa 558

<210> 60
<211> 197
<212> PRT
<213> Homo sapiens

<400> 60
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Gly Gly Gly Gly
1 5 10 15
Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
20 25 30
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
35 40 45
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
50 55 60
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
65 70 75 80
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly
85 90 95
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
100 105 110
Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
115 120 125
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Cys Asp Pro
130 135 140
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
145 150 155 160
Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
165 170 175
Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser
180 185 190
Pro Ser Tyr Ala Ser
195

<210> 61

<211> 594
<212> ADN
<213> Homo sapiens

<400> 61
cattctgaag gcacttttac tagcgatgtt tcttctggtg gtggtggttc tggcggtggc 60
gactcgagcc cgctgctgca gtttggcggc caggtgcgtc agcgttatct gtataccgat 120
gatgcgcagc agaccgaagc gcatctggaa attcgtgaag atggcaccgt gggcggtgcg 180
gcggatcaga gcccggaaag cctgctgcag ctgaaagcgc tgaaaccggg cgtgattcag 240
attctgggcg tgaaaaccag ccgttttctg tgccagcgtc cggatggcgc gctgtatggc 300
agcctgcatt ttgatccgga agcgtgcagc tttcgtgaac tgctgctgga agatggctat 360
aacgtgtatc agagcgaagc gcatggcctg ccgctgcatc tgccgggcaa caaaagcccg 420
cattgcgatc cggcaccgcg tggtccggcg cgttttctgc cgctgccggg tctgccgccg 480
gcactgccgg aaccgccggg tattctggcc ccgcagccgc cggatgttgg tagcagcgat 540
ccgctgtcta tggtgggtcc gagccagggt cgtagcccga gctatgcgag ctaa 594

<210> 62
<211> 201
<212> PRT
<213> Homo sapiens

<400> 62
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Ser
 1   5   10   15
Gly Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe
            20   25   30
Gly Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln
        35   40   45
Thr Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala
    50   55   60
Ala Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro
65   70   75   80
Gly Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln
                85   90   95
Arg Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala
            100   105   110
Cys Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln
        115   120   125
Ser Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro
    130   135   140
His Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro
145   150   155   160
Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln
                165   170   175
Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser
            180   185   190
Gln Gly Arg Ser Pro Ser Tyr Ala Ser
        195   200

<210> 63
<211> 606
<212> ADN
<213> Homo sapiens

<400> 63
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaatctgg tggtggtggt 60
tctggcggtg gcgactcgag cccgctgctg cagtttggcg gccaggtgcg tcagcgttat 120
ctgtataccg atgatgcgca gcagaccgaa gcgcatctgg aaattcgtga agatggcacc 180
gtgggcggtg cggcggatca gagcccggaa agcctgctgc agctgaaagc gctgaaaccg 240
ggcgtgattc agattctggg cgtgaaaacc agccgttttc tgtgccagcg tccggatggc 300
gcgctgtatg gcagcctgca ttttgatccg gaagcgtgca gctttcgtga actgctgctg 360
gaagatggct ataacgtgta tcagagcgaa gcgcatggcc tgccgctgca tctgccgggc 420

aacaaaagcc cgcattgcga tccggcaccg cgtggtccgg cgcgttttct gccgctgccg 480
ggtctgccgc cggcactgcc ggaaccgccg ggtattctgg ccccgcagcc gccggatgtt 540
ggtagcagcg atccgctgtc tatggtgggt ccgagccagg gtcgtagccc gagctatgcg 600
agctaa 606

<210> 64
<211> 197
<212> PRT
<213> Homo sapiens

<400> 64
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val
            20   25   30
Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His
        35   40   45
Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser
    50   55   60
Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln
65   70   75   80
Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly
                85   90   95
Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg
            100   105   110
Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His
        115   120   125
Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Cys Asp Pro
    130   135   140
Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro
145   150   155   160
Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val
                165   170   175
Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser
            180   185   190
Pro Ser Tyr Ala Ser
        195

<210> 65
<211> 594
<212> ADN
<213> Homo sapiens

<400> 65
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gatagcagcc cgctgctgca gtttggcggc caggtgcgtc agcgttatct gtataccgat 120
gatgcgcagc agaccgaagc gcatctggaa attcgtgaag atggcaccgt gggcggtgcg 180
gcggatcaga gcccggaaag cctgctgcag ctgaaagcgc tgaaaccggg cgtgattcag 240
attctgggcg tgaaaaccag ccgttttctg tgccagcgtc cggatggcgc gctgtatggc 300
agcctgcatt ttgatccgga agcgtgcagc tttcgtgaac tgctgctgga agatggctat 360
aacgtgtatc agagcgaagc gcatggcctg ccgctgcatc tgccgggcaa caaaagcccg 420
cattgcgatc cggcaccgcg tggtccggcg cgttttctgc cgctgccggg tctgccgccg 480
gcactgccgg aaccgccggg tattctggcc ccgcagccgc cggatgttgg tagcagcgat 540
ccgctgtcta tggtgggtcc gagccagggt cgtagcccga gctatgcgag ctaa 594

<210> 66
<211> 209
<212> PRT
<213> Homo sapiens

<400> 66
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly

```
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ser Gly Gly Gly Gly Ser Gly Gly Gly
            20                  25                  30
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
        35                  40                  45
Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu Glu Ile Arg
    50                  55                  60
Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu
65                  70                  75                  80
Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
                85                  90                  95
Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala Leu Tyr Gly
            100                 105                 110
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
        115                 120                 125
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
    130                 135                 140
His Leu Pro Gly Asn Lys Ser Pro His Cys Asp Pro Ala Pro Arg Gly
145                 150                 155                 160
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
                165                 170                 175
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
            180                 185                 190
Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
        195                 200                 205
Ser
```

<210> 67
<211> 630
<212> ADN
<213> Homo sapiens

<400> 67
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatct ctggtggtgg tggttctggc ggtggcgact cgagcccgct gctgcagttt 120
ggcggccagg tgcgtcagcg ttatctgtat accgatgatg cgcagcagac cgaagcgcat 180
ctggaaattc gtgaagatgg caccgtgggc ggtgcggcgg atcagagccc ggaaagcctg 240
ctgcagctga aagcgctgaa accgggcgtg attcagattc tgggcgtgaa aaccagccgt 300
tttctgtgcc agcgtccgga tggcgcgctg tatggcagcc tgcattttga tccggaagcg 360
tgcagctttc gtgaactgct gctggaagat ggctataacg tgtatcagag cgaagcgcat 420
ggcctgccgc tgcatctgcc gggcaacaaa agcccgcatt gcgatccggc accgcgtggt 480
ccggcgcgtt ttctgccgct gccgggtctg ccgccggcac tgccggaacc gccgggtatt 540
ctggccccgc agccgccgga tgttggtagc agcgatccgc tgtctatggt gggtccgagc 600
cagggtcgta gcccgagcta tgcgagctaa   630

<210> 68
<211> 212
<212> PRT
<213> Homo sapiens

<400> 68
```
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Ser Gly Gly Gly Gly Ser
            20                  25                  30
Gly Gly Gly Asp Ser Ser Pro Leu Gln Phe Gly Gly Gln Val Arg
        35                  40                  45
Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu
    50                  55                  60
Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro
65                  70                  75                  80
```

Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile
85 90 95
Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala
100 105 110
Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu
115 120 125
Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly
130 135 140
Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Cys Asp Pro Ala
145 150 155 160
Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala
165 170 175
Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly
180 185 190
Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro
195 200 205
Ser Tyr Ala Ser
210

<210> 69
<211> 639
<212> ADN
<213> Homo sapiens

<400> 69
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctgtc tggtggtggt ggttctggcg gtggcgactc gagcccgctg 120
ctgcagtttg gcggccaggt gcgtcagcgt tatctgtata ccgatgatgc gcagcagacc 180
gaagcgcatc tggaaattcg tgaagatggc accgtgggcg gtgcggcgga tcagagcccg 240
gaaagcctgc tgcagctgaa agcgctgaaa ccgggcgtga ttcagattct gggcgtgaaa 300
accagccgtt ttctgtgcca gcgtccggat ggcgcgctgt atggcagcct gcattttgat 360
ccggaagcgt gcagctttcg tgaactgctg ctggaagatg gctataacgt gtatcagagc 420
gaagcgcatg gcctgccgct gcatctgccg ggcaacaaaa gcccgcattg cgatccggca 480
ccgcgtggtc cggcgcgttt tctgccgctg ccgggtctgc cgccggcact gccggaaccg 540
ccgggtattc tggccccgca gccgccggat gttggtagca gcgatccgct gtctatggtg 600
ggtccgagcc agggtcgtag cccgagctat gcgagctaa 639

<210> 70
<211> 216
<212> PRT
<213> Homo sapiens

<400> 70
His Ser Glu Gly Thr Phe Thr Ser Asp Val Cys Ser Tyr Leu Glu Gly
1 5 10 15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
20 25 30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
35 40 45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
50 55 60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65 70 75 80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
85 90 95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
100 105 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
115 120 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His

145 150 155 160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 71
<211> 651
<212> ADN
<213> Homo sapiens

<400> 71
cattctgaag gcacttttac tagcgatgtt tgtagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 72
<211> 216
<212> PRT
<213> Homo sapiens

<400> 72
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15
Cys Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
20 25 30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
35 40 45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
50 55 60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65 70 75 80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
85 90 95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
100 105 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
115 120 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145 150 155 160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 73
<211> 651
<212> ADN
<213> Homo sapiens

<400> 73
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggctg tgctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 74
<211> 216
<212> PRT
<213> Homo sapiens

<400> 74
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Cys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 75
<211> 651
<212> ADN
<213> Homo sapiens

<400> 75
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgtgt 60

gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 76
<211> 216
<212> PRT
<213> Homo sapiens

<400> 76
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Cys Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 77
<211> 651
<212> ADN
<213> Homo sapiens

<400> 77
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
tgtttcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600

ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 78
<211> 216
<212> PRT
<213> Homo sapiens

<400> 78
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Cys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 79
<211> 651
<212> ADN
<213> Homo sapiens

<400> 79
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt ttgtggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 80
<211> 216
<212> PRT
<213> Homo sapiens

<400> 80
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly

```
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Cys Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145                 150                 155                 160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165                 170                 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180                 185                 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195                 200                 205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215
```

<210> 81
<211> 651
<212> ADN
<213> Homo sapiens

<400> 81
```
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttgtgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a          651
```

<210> 82
<211> 216
<212> PRT
<213> Homo sapiens

<400> 82
```
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Cys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
```

```
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145                 150                 155                 160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165                 170                 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180                 185                 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195                 200                 205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215
```

<210> 83
<211> 651
<212> ADN
<213> Homo sapiens

<400> 83
catgcggaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgtgt 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a   651

<210> 84
<211> 216
<212> PRT
<213> Homo sapiens

<400> 84
```
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Cys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
```

145 150 155 160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 85
<211> 651
<212> ADN
<213> Homo sapiens

<400> 85
catgcggaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt ttgtggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 86
<211> 216
<212> PRT
<213> Homo sapiens

<400> 86
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
20 25 30
Gly Gly Gly Cys Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
35 40 45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
50 55 60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65 70 75 80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
85 90 95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
100 105 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
115 120 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145 150 155 160
Arg Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 87
<211> 651
<212> ADN
<213> Homo sapiens

<400> 87
catgcggaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttgtgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
cgtgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 88
<211> 216
<212> PRT
<213> Homo sapiens

<400> 88
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1   5   10   15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 89
<211> 651
<212> ADN
<213> Homo sapiens

<400> 89
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60

ctgttcattg aatggctgaa aaatggtggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 90
<211> 216
<212> PRT
<213> Homo sapiens

<400> 90
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Gly
 1   5   10   15
Gln Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 91
<211> 651
<212> ADN
<213> Homo sapiens

<400> 91
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagggca agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600

ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 92
<211> 216
<212> PRT
<213> Homo sapiens

<220>
<223> aminoácido no natural

<400> 92
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Cys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Xaa Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 93
<211> 651
<212> ADN
<213> Homo sapiens

<400> 93
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt ttgtggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
taggatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 94
<211> 216
<212> PRT
<213> Homo sapiens

<220>
<223> aminoácido no natural

<400> 94
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Cys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145   150   155   160
Xaa Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 95
<211> 651
<212> ADN
<213> Homo sapiens

<400> 95
catgcggaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt ttgtggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
taggatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a   651

<210> 96
<211> 216
<212> PRT
<213> Homo sapiens

<220>
<223> aminoácido no natural

<400> 96
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly

```
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Xaa Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His
145                 150                 155                 160
Cys Asp Pro Ala Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165                 170                 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180                 185                 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln
        195                 200                 205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215
```

<210> 97
<211> 651
<212> ADN
<213> Homo sapiens

<400> 97

```
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt ttagggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcagcaga ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
gatcagagcc cggaaagcct gctgcagctg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagcgtccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaacaa aagcccgcat 480
tgcgatccgg caccgcgtgg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctgtctatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a          651
```

<210> 98
<211> 216
<212> PRT
<213> Homo sapiens

<400> 98

```
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
```

```
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145                 150                 155                 160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165                 170                 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180                 185                 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
        195                 200                 205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215
```

<210> 99
<211> 651
<212> ADN
<213> Homo sapiens

<400> 99
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a   651

<210> 100
<211> 216
<212> PRT
<213> Homo sapiens

<400> 100
```
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Glu Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
```

145 150 155 160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 101
<211> 651
<212> ADN
<213> Homo sapiens

<400> 101
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga aggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 102
<211> 216
<212> PRT
<213> Homo sapiens

<400> 102
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Cys Gly Gly Ser Gly
20 25 30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
35 40 45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
50 55 60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65 70 75 80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
85 90 95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
100 105 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
115 120 125
Ser Phe Arg Glu Leu Leu Leu Glu Glu Gly Tyr Asn Val Tyr Gln Ser
130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145 150 155 160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 103
<211> 651
<212> ADN
<213> Homo sapiens

<400> 103
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt ttgtggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga aggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 104
<211> 216
<212> PRT
<213> Homo sapiens

<400> 104
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Cys Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145   150   155   160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
        195   200   205
Ala Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 105
<211> 651
<212> ADN
<213> Homo sapiens

<400> 105
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60

gaattcatcg cgtggctggt ttgtggcggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccaggcgcgt agcccgagct atgcgagcta a 651

<210> 106
<211> 216
<212> PRT
<213> Homo sapiens

<400> 106
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1   5   10   15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Ser Gly
            20   25   30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35   40   45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
    50   55   60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65   70   75   80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
                85   90   95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
            100   105   110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115   120   125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130   135   140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145   150   155   160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165   170   175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180   185   190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
        195   200   205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210   215

<210> 107
<211> 651
<212> ADN
<213> Homo sapiens

<400> 107
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600

ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 108
<211> 216
<212> PRT
<213> Homo sapiens

<400> 108
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Ser Gly
20 25 30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
35 40 45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
50 55 60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65 70 75 80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
85 90 95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
100 105 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
115 120 125
Ser Phe Arg Glu Leu Leu Leu Glu Glu Gly Tyr Asn Val Tyr Gln Ser
130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145 150 155 160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
195 200 205
Gly Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 109
<211> 651
<212> ADN
<213> Homo sapiens

<400> 109
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga aggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 110
<211> 216
<212> PRT
<213> Homo sapiens

<400> 110
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu

```
 1               5                  10                  15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Cys Asn Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Glu Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145                 150                 155                 160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165                 170                 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180                 185                 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
        195                 200                 205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215
```

<210> 111
<211> 651
<212> ADN
<213> Homo sapiens

<400> 111
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgtg taatggtggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga aggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a   651

<210> 112
<211> 216
<212> PRT
<213> Homo sapiens

<400> 112
```
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1               5                  10                  15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Cys Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
```

```
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Glu Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145                 150                 155                 160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
                165                 170                 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
            180                 185                 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
        195                 200                 205
Gly Arg Ser Pro Ser Tyr Ala Ser
    210                 215
```

<210> 113
<211> 651
<212> ADN
<213> Homo sapiens

<400> 113
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa atgtggtggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga aggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccagggtcgt agcccgagct atgcgagcta a 651

<210> 114
<211> 216
<212> PRT
<213> Homo sapiens

<400> 114
```
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1               5                  10                  15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Cys Asn Gly Gly Ser Gly
            20                  25                  30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
        35                  40                  45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
    50                  55                  60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65                  70                  75                  80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
                85                  90                  95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
            100                 105                 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
        115                 120                 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
    130                 135                 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
```

145 150 155 160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
195 200 205
Ala Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 115
<211> 651
<212> ADN
<213> Homo sapiens

<400> 115
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgtg taatggtggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccaggcgcgt agcccgagct atgcgagcta a 651

<210> 116
<211> 216
<212> PRT
<213> Homo sapiens

<400> 116
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Cys Gly Gly Ser Gly
20 25 30
Gly Gly Gly Ser Gly Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly
35 40 45
Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr
50 55 60
Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala
65 70 75 80
His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly
85 90 95
Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys
100 105 110
Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys
115 120 125
Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser
130 135 140
Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg Ser Pro His
145 150 155 160
Arg Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly
165 170 175
Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro
180 185 190
Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln
195 200 205
Ala Arg Ser Pro Ser Tyr Ala Ser
210 215

<210> 117
<211> 651
<212> ADN
<213> Homo sapiens

<400> 117
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa atgtggtggt tctggtggtg gtggttctgg cggtggcgac 120
tcgagcccgc tgctgcagtt tggcggccag gtgcgtcagc gttatctgta taccgatgat 180
gcgcaggaaa ccgaagcgca tctggaaatt cgtgaagatg gcaccgtggg cggtgcggcg 240
catcagagcc cggaaagcct gctggaactg aaagcgctga aaccgggcgt gattcagatt 300
ctgggcgtga aaaccagccg ttttctgtgc cagaaaccgg atggcgcgct gtatggcagc 360
ctgcattttg atccggaagc gtgcagcttt cgtgaactgc tgctggaaga tggctataac 420
gtgtatcaga gcgaagcgca tggcctgccg ctgcatctgc cgggcaaccg tagcccgcat 480
cgtgatccgg caccgcaggg tccggcgcgt tttctgccgc tgccgggtct gccgccggca 540
ctgccggaac cgccgggtat tctggccccg cagccgccgg atgttggtag cagcgatccg 600
ctggcgatgg tgggtccgag ccaggcgcgt agcccgagct atgcgagcta a 651

<210> 118
<211> 227
<212> PRT
<213> Homo sapiens

<400> 118
His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1   5   10   15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Gly Gly Ser
            20   25   30
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
        35   40   45
Gly Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln
    50   55   60
Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu
65   70   75   80
Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala His Gln Ser Pro Glu
                85   90   95
Ser Leu Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu
            100   105   110
Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp Gly Ala Leu
        115   120   125
Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu
    130   135   140
Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu
145   150   155   160
Pro Leu His Leu Pro Gly Asn Arg Ser Pro His Cys Asp Pro Ala Pro
                165   170   175
Gln Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu
            180   185   190
Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser
        195   200   205
Ser Asp Pro Leu Ala Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser
    210   215   220
Tyr Ala Ser
225

<210> 119
<211> 684
<212> ADN
<213> Homo sapiens

<400> 119
cattctgaag gcacttttac tagcgatgtt tctagctacc tggaaggcca ggctgcgaaa 60
gaattcatcg cgtggctggt taaaggcggt ggtagcggtg gcggcggttc tggtggtggt 120
ggttctggcg gtggcggtag cggtggcggc gatagcagcc cgctgctgca gtttggcggc 180
caggtgcgtc agcgttatct gtataccgat gatgcgcagg aaaccgaagc gcatctggaa 240
attcgtgaag atggcaccgt gggcggtgcg gcgcatcaga gcccggaaag cctgctggaa 300
ctgaaagcgc tgaaaccggg cgtgattcag attctgggcg tgaaaaccag ccgttttctg 360
tgccagaaac cggatggcgc gctgtatggc agcctgcatt ttgatccgga agcgtgcagc 420
tttcgtgaac tgctgctgga agatggctat aacgtgtatc agagcgaagc gcatggcctg 480
ccgctgcatc tgccgggcaa ccgtagcccg cattgcgatc cggcaccgca gggtccggcg 540
cgttttctgc cgctgccggg tctgccgccg gcactgccgg aaccgccggg tattctggcc 600
ccgcagccgc cggatgttgg tagcagcgat ccgctggcga tggtgggtcc gagccagggt 660
cgtagcccga gctatgcgag ctaa 684

<210> 120
<211> 474
<212> PRT
<213> Homo sapiens

<400> 120
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
 1   5   10   15
Gly Ser Thr Gly His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys
            20   25   30
Gln Met Glu Glu Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn
        35   40   45
Gly Gly Gly Gly Gly Gly Ser Asp Lys Thr His Thr Cys Pro Pro Cys
    50   55   60
Pro Ala Pro Glu Ala Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro
65   70   75   80
Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys
                85   90   95
Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp
            100   105   110
Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu
        115   120   125
Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu
    130   135   140
His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn
145   150   155   160
Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly
                165   170   175
Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu
            180   185   190
Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr
        195   200   205
Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn
    210   215   220
Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe
225   230   235   240
Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
                245   250   255
Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr
            260   265   270
Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Gly Gly Gly Ser Gly
        275   280   285
Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ser Ser Pro Leu Leu Gln
    290   295   300
Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Cys
305   310   315   320
Gln Thr Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly
                325   330   335
Ala Ala Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys

340 345 350
Pro Gly Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys
355 360 365
Gln Lys Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu
370 375 380
Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr
385 390 395 400
Gln Ser Glu Ala His Gly Leu Pro Leu His Leu Pro Cys Asn Arg Ser
405 410 415
Pro His Arg Asp Pro Ala Ser Arg Gly Pro Ala Arg Phe Leu Pro Leu
420 425 430
Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro
435 440 445
Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Gly
450 455 460
Ser Gln Ala Arg Ser Pro Ser Tyr Ala Ser
465 470

<210> 121
<211> 484
<212> PRT
<213> Homo sapiens

<400> 121
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15
Gly Ser Thr Gly His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys
20 25 30
Gln Met Glu Glu Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn
35 40 45
Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
50 55 60
Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Ala Ala
65 70 75 80
Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu
85 90 95
Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser
100 105 110
His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu
115 120 125
Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr
130 135 140
Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn
145 150 155 160
Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro
165 170 175
Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln
180 185 190
Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val
195 200 205
Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val
210 215 220
Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro
225 230 235 240
Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr
245 250 255
Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val
260 265 270
Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
275 280 285
Ser Pro Gly Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
290 295 300

Gly Gly Ser Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg
305 310 315 320
Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Cys Gln Thr Glu Ala His Leu
325 330 335
Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro
340 345 350
Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile
355 360 365
Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp Gly Ala
370 375 380
Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu
385 390 395 400
Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly
405 410 415
Leu Pro Leu His Leu Pro Cys Asn Arg Ser Pro His Arg Asp Pro Ala
420 425 430
Ser Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala
435 440 445
Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly
450 455 460
Ser Ser Asp Pro Leu Ala Met Val Gly Gly Ser Gln Ala Arg Ser Pro
465 470 475 480
Ser Tyr Ala Ser

<210> 122
<211> 483
<212> PRT
<213> Homo sapiens

<400> 122
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15
Gly Ser Thr Gly His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys
20 25 30
Gln Met Glu Glu Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn
35 40 45
Gly Gly Pro Ser Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser
50 55 60
Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Ala Ala Gly
65 70 75 80
Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
85 90 95
Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
100 105 110
Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
115 120 125
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
130 135 140
Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
145 150 155 160
Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
165 170 175
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
180 185 190
Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser
195 200 205
Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
210 215 220
Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
225 230 235 240
Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val

245 250 255
Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
260 265 270
His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
275 280 285
Pro Gly Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
290 295 300
Gly Ser Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln
305 310 315 320
Arg Tyr Leu Tyr Thr Asp Asp Ala Cys Gln Thr Glu Ala His Leu Glu
325 330 335
Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu
340 345 350
Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu
355 360 365
Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp Gly Ala Leu
370 375 380
Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu
385 390 395 400
Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu
405 410 415
Pro Leu His Leu Pro Cys Asn Arg Ser Pro His Arg Asp Pro Ala Ser
420 425 430
Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu
435 440 445
Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser
450 455 460
Ser Asp Pro Leu Ala Met Val Gly Gly Ser Gln Ala Arg Ser Pro Ser
465 470 475 480
Tyr Ala Ser

<210> 123
<211> 493
<212> PRT
<213> Homo sapiens

<400> 123
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15
Gly Ser Thr Gly His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys
20 25 30
Gln Met Glu Glu Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn
35 40 45
Gly Gly Pro Ser Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser
50 55 60
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Lys Thr His Thr Cys
65 70 75 80
Pro Pro Cys Pro Ala Pro Glu Ala Ala Gly Gly Pro Ser Val Phe Leu
85 90 95
Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu
100 105 110
Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys
115 120 125
Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys
130 135 140
Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu
145 150 155 160
Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys
165 170 175
Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys
180 185 190

```
Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser
        195                 200                 205
Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys
    210                 215                 220
Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln
225                 230                 235                 240
Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly
                245                 250                 255
Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln
            260                 265                 270
Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn
        275                 280                 285
His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Gly Gly
    290                 295                 300
Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ser Ser Pro
305                 310                 315                 320
Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp
                325                 330                 335
Asp Ala Cys Gln Thr Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr
            340                 345                 350
Val Gly Gly Ala Ala Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys
        355                 360                 365
Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg
    370                 375                 380
Phe Leu Cys Gln Lys Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe
385                 390                 395                 400
Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr
                405                 410                 415
Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu His Leu Pro Cys
            420                 425                 430
Asn Arg Ser Pro His Arg Asp Pro Ala Ser Arg Gly Pro Ala Arg Phe
        435                 440                 445
Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile
    450                 455                 460
Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met
465                 470                 475                 480
Val Gly Gly Ser Gln Ala Arg Ser Pro Ser Tyr Ala Ser
                485                 490
```

<210> 124
<211> 473
<212> PRT
<213> Homo sapiens

<400> 124

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
 1               5                  10                  15
Gly Ser Thr Gly His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser
            20                  25                  30
Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys
        35                  40                  45
Gly Gly Gly Gly Gly Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro
    50                  55                  60
Ala Pro Glu Ala Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
65                  70                  75                  80
Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
                85                  90                  95
Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
            100                 105                 110
Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
        115                 120                 125
Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
```

```
    130                 135                 140
Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
145                 150                 155                 160
Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
                165                 170                 175
Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met
            180                 185                 190
Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
        195                 200                 205
Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
    210                 215                 220
Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
225                 230                 235                 240
Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
                245                 250                 255
Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
            260                 265                 270
Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Gly Gly Gly Ser Gly Gly
        275                 280                 285
Gly Gly Ser Gly Gly Gly Gly Ser Asp Ser Ser Pro Leu Leu Gln Phe
    290                 295                 300
Gly Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Cys Gln
305                 310                 315                 320
Thr Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala
                325                 330                 335
Ala Asp Gln Ser Pro Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro
            340                 345                 350
Gly Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln
        355                 360                 365
Lys Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala
    370                 375                 380
Cys Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln
385                 390                 395                 400
Ser Glu Ala His Gly Leu Pro Leu His Leu Pro Cys Asn Arg Ser Pro
                405                 410                 415
His Arg Asp Pro Ala Ser Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro
            420                 425                 430
Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln
        435                 440                 445
Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly Gly Ser
    450                 455                 460
Gln Ala Arg Ser Pro Ser Tyr Ala Ser
465                 470


<210> 125
<211> 483
<212> PRT
<213> Homo sapiens

<400> 125
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
 1               5                  10                  15
Gly Ser Thr Gly His Ser Glu Gly Thr Phe Thr Ser Asp Val Ser Ser
            20                  25                  30
Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys
        35                  40                  45
Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
    50                  55                  60
Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Ala Ala Gly
65                  70                  75                  80
Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
                85                  90                  95
```

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
100 105 110
Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
115 120 125
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
130 135 140
Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
145 150 155 160
Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
165 170 175
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
180 185 190
Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser
195 200 205
Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
210 215 220
Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
225 230 235 240
Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
245 250 255
Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
260 265 270
His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
275 280 285
Pro Gly Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
290 295 300
Gly Ser Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln
305 310 315 320
Arg Tyr Leu Tyr Thr Asp Asp Ala Cys Gln Thr Glu Ala His Leu Glu
325 330 335
Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro Glu
340 345 350
Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu
355 360 365
Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp Gly Ala Leu
370 375 380
Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu
385 390 395 400
Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu
405 410 415
Pro Leu His Leu Pro Cys Asn Arg Ser Pro His Arg Asp Pro Ala Ser
420 425 430
Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu
435 440 445
Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser
450 455 460
Ser Asp Pro Leu Ala Met Val Gly Gly Ser Gln Ala Arg Ser Pro Ser
465 470 475 480
Tyr Ala Ser

<210> 126
<211> 235
<212> PRT
<213> Homo sapiens

<400> 126
His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30
Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly

```
        35                  40                  45
Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Asp Ser Ser Pro Leu Leu
    50                  55                  60
Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr Leu Tyr Thr Asp Asp Ala
65                  70                  75                  80
Gln Glu Thr Glu Ala His Leu Glu Ile Arg Glu Asp Gly Thr Val Gly
                85                  90                  95
Gly Ala Ala His Gln Ser Pro Glu Ser Leu Leu Glu Leu Lys Ala Leu
            100                 105                 110
Lys Pro Gly Val Ile Gln Ile Leu Gly Val Lys Thr Ser Arg Phe Leu
        115                 120                 125
Cys Gln Lys Pro Asp Gly Ala Leu Tyr Gly Ser Leu His Phe Asp Pro
    130                 135                 140
Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu Glu Asp Gly Tyr Asn Val
145                 150                 155                 160
Tyr Gln Ser Glu Ala His Gly Leu Pro Leu His Leu Pro Gly Asn Arg
                165                 170                 175
Ser Pro His Cys Asp Pro Ala Pro Gln Gly Pro Ala Arg Phe Leu Pro
            180                 185                 190
Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu Pro Pro Gly Ile Leu Ala
        195                 200                 205
Pro Gln Pro Pro Asp Val Gly Ser Ser Asp Pro Leu Ala Met Val Gly
    210                 215                 220
Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala Ser
225                 230                 235


<210> 127
<211> 708
<212> ADN
<213> Homo sapiens

<400> 127
catggtgagg gtacgtttac ttctgatctg tctaaacaga tggaagaaga agctgttcgc 60
ctgttcattg aatggctgaa aaatggtggt ccgtcctccg gcgctcctcc gccttctggt 120
ggtggtggtt ctggcggtgg cggttctggc ggcggtggta gcggtggcgg cggtgatagc 180
agcccgctgc tgcagtttgg cggccaggtg cgtcagcgtt atctgtatac cgatgatgcg 240
caggaaaccg aagcgcatct ggaaattcgt gaagatggca ccgtgggcgg tgcggcgcat 300
cagagcccgg aaagcctgct ggaactgaaa gcgctgaaac cgggcgtgat tcagattctg 360
ggcgtgaaaa ccagccgttt tctgtgccag aaaccggatg gcgcgctgta tggcagcctg 420
cattttgatc cggaagcgtg cagctttcgt gaactgctgc tggaagatgg ctataacgtg 480
tatcagagcg aagcgcatgg cctgccgctg catctgccgg gcaaccgtag cccgcattgc 540
gatccggcac cgcagggtcc ggcgcgtttt ctgccgctgc cgggtctgcc gccggcactg 600
ccggaaccgc cgggtattct ggccccgcag ccgccggatg ttggtagcag cgatccgctg 660
gcgatggtgg gtccgagcca gggtcgtagc ccgagctatg cgagctaa              708

<210> 128
<211> 9
<212> ADN
<213> Homo sapiens

<400> 128
sggggsggg                                                         9

<210> 129
<211> 37
<212> PRT
<213> Homo sapiens

<400> 129
His Asp Glu Phe Glu Arg His Ala Glu Gly Thr Phe Thr Ser Asp Val
 1               5                  10                  15
Ser Ser Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu
```

20 25 30
Val Lys Gly Arg Gly
35

<210> 130
<211> 177
<212> PRT
<213> Homo sapiens

<400> 130
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
1 5 10 15
Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg
20 25 30
Glu Asp Gly Thr Val Gly Gly Ala Ala His Gln Ser Pro Glu Ser Leu
35 40 45
Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
50 55 60
Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp Gly Ala Leu Tyr Gly
65 70 75 80
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
85 90 95
Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
100 105 110
His Leu Pro Gly Asn Arg Ser Pro His Cys Asp Pro Ala Pro Gln Gly
115 120 125
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
130 135 140
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
145 150 155 160
Pro Leu Ala Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
165 170 175
Ser

<210> 131
<211> 177
<212> PRT
<213> Homo sapiens

<400> 131
Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Gln Val Arg Gln Arg Tyr
1 5 10 15
Leu Tyr Thr Asp Asp Ala Gln Glu Thr Glu Ala His Leu Glu Ile Arg
20 25 30
Glu Asp Gly Thr Val Gly Gly Ala Ala His Gln Ser Pro Glu Ser Leu
35 40 45
Leu Glu Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile Leu Gly Val
50 55 60
Lys Thr Ser Arg Phe Leu Cys Gln Lys Pro Asp Gly Ala Leu Tyr Gly
65 70 75 80
Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu Leu Leu Leu
85 90 95
Glu Glu Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly Leu Pro Leu
100 105 110
His Leu Pro Gly Asn Arg Ser Pro His Arg Asp Pro Ala Pro Gln Gly
115 120 125
Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala Leu Pro Glu
130 135 140
Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly Ser Ser Asp
145 150 155 160

Pro Leu Ala Met Val Gly Pro Ser Gln Gly Arg Ser Pro Ser Tyr Ala
165 170 175
Ser

<210> 132
<211> 28
<212> PRT
<213> Homo sapiens

<400> 132
Met Asp Ser Asp Glu Thr Gly Phe Glu His Ser Gly Leu Trp Val Ser
1 5 10 15
Val Leu Ala Gly Leu Leu Leu Gly Ala Cys Gln Ala
20 25

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	**FOR FURTHER ACTION**	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
54621-WO-PCT		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/US2012/057371	26 September 2012 (26-09-2012)	26 September 2011 (26-09-2011)
Applicant		
NOVARTIS AG		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 6 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**
   a. With regard to the **language,** the international search was carried out on the basis of:
      ☒ the international application in the language in which it was filed
      ☐ a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))
   b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).
   c. ☒ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☐ **Certain claims were found unsearchable** (See Box No. II)

3. ☒ **Unity of invention is lacking** (see Box No III)

4. With regard to the **title,**
   ☒ the text is approved as submitted by the applicant
   ☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**
   ☒ the text is approved as submitted by the applicant
   ☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority

6. With regard to the **drawings,**
   a. the figure of the **drawings** to be published with the abstract is Figure No. _____
      ☐ as suggested by the applicant
      ☐ as selected by this Authority, because the applicant failed to suggest a figure
      ☐ as selected by this Authority, because this figure better characterizes the invention
   b. ☒ none of the figures is to be published with the abstract

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2012/057371

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. (means)

   ☐ on paper

X ☐ in electronic form

   b. (time)

X ☐ in the international application as filed

   ☐ together with the international application in electronic form

   ☐ subsequently to this Authority for the purpose of search

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

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2012/057371

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

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

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

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

see additional sheet(s)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2012/057371

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/50 C07K14/605
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, Sequence Search, WPI Data, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/020319 A1 (CHONGQING FAGEN BIOMEDICAL INC [CN]; FAN KAI [CN]; ZHAO ZHIQUAN [CN];) 24 February 2011 (2011-02-24)	1-5,7,8
A	passages from translation (= EP2468858): page 3, line 45 - line 50 page 4, line 23 - line 29 claims 1-15; figures 3,4	6,9

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
2 July 2013	11/10/2013

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Niebuhr-Ebel, K

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2012/057371

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOYAMA N ET AL: "A novel procedure for the preparation of biologically active recombinant peptides using a cyanylation reaction", JOURNAL OF BIOTECHNOLOGY, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 32, no. 3, 28 February 1994 (1994-02-28), pages 273-281, XP023705313, ISSN: 0168-1656, DOI: 10.1016/0168-1656(94)90213-5 [retrieved on 1994-02-28]	1
A	abstract figure 1	2-9

X,P	EP 2 468 858 A1 (CHONGQING FAGEN BIOMEDICAL INC [CN]) 27 June 2012 (2012-06-27)	1-5,7,8
A,P	page 3, line 45 - line 50 page 4, line 23 - line 29 claims 1-15; figures 3,4	6,9

X,P	WO 2012/066075 A1 (NOVARTIS AG [CH]; IRM LLC [US]; BOETTCHER BRIAN R [US]; CAPLAN SHARI L) 24 May 2012 (2012-05-24)	1-5,7,8
A,P	sequence 39 -& DATABASE Geneseq [Online] 19 July 2012 (2012-07-19), "Human mature fibroblast growth factor 21 polypeptide variant 76, SEQ 39.", XP002699914, retrieved from EBI accession no. GSP:AZW46640 Database accession no. AZW46640 sequence	6,9

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2012/057371

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011020319 A1	24-02-2011	CN 101993485 A	30-03-2011
		EP 2479193 A1	25-07-2012
		US 2012238496 A1	20-09-2012
		US 2012277154 A1	01-11-2012
		WO 2011020319 A1	24-02-2011
		WO 2011020320 A1	24-02-2011
EP 2468858 A1	27-06-2012	CN 101993496 A	30-03-2011
		CN 102770538 A	07-11-2012
		EP 2468858 A1	27-06-2012
WO 2012066075 A1	24-05-2012	AR 083920 A1	10-04-2013
		CA 2817786 A1	24-05-2012
		CN 103328502 A	25-09-2013
		CR 20130234 A	01-07-2013
		EA 201390739 A1	30-09-2013
		EP 2640741 A1	25-09-2013
		SG 190324 A1	28-06-2013
		TW 201302782 A	16-01-2013
		US 2012129766 A1	24-05-2012
		UY 33739 A	29-06-2012
		WO 2012066075 A1	24-05-2012

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 3-5(completely); 1, 2, 6-9(partially)

Fusion protein comprising a GLP-1 part and an FGF21 part, wherein the FGF21 variant is variant 76 (SEQ ID NO: 130), i.e. fusion proteins "variant 272" (SEQ ID NOs: 9 and 10) and "variant 277" (SEQ ID NOs: 11 and 12).

2-98. claims: 1, 2, 6-9(all partially)

Further variants of GLP-1 / FGF21 fusion proteins as listed in Table 1, pharmaceutical compositions and their use

---