



(12) 发明专利申请

(10) 申请公布号 CN 101712722 A

(43) 申请公布日 2010.05.26

(21) 申请号 200910173888.8

(22) 申请日 2001.11.29

(30) 优先权数据

60/251954 2000.12.07 US

(62) 分案原申请数据

01820232.2 2001.11.29

(71) 申请人 伊莱利利公司

地址 美国印第安纳州

(72) 发明人 W·格莱斯纳 R·米卡诺维克

张盛虹

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 郭文洁

(51) Int. Cl.

C07K 19/00 (2006.01)

A61K 38/26 (2006.01)

A61K 47/48 (2006.01)

A61P 3/10 (2006.01)

A61P 3/04 (2006.01)

权利要求书 6 页 说明书 97 页 附图 15 页

(54) 发明名称

GLP-1 融合蛋白

(57) 摘要

本发明涉及与一些蛋白融合的胰高血糖素样肽-1 (GLP-1) 化合物, 所述蛋白具有延长胰高血糖素样肽-1 化合物的体内半衰期的作用。这些 GLP-1 融合蛋白可以用于治疗非胰岛素依赖型糖尿病和各种状况。

1. 一种异源融合蛋白,其具有第一个多肽,该多肽具有N末端和与具有N末端和C末端的第二个多肽融合的C末端,其中第一个多肽是具有式I,即SEQ ID NO:2所示序列的GLP-1化合物

7 8 9 10 11 12 13 14 15 16 17
His-Xaa-Xaa-Gly-Xaa-Phe-Thr-Xaa-Asp-Xaa-Xaa-
18 19 20 21 22 23 24 25 26 27 28
Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Phe-
29 30 31 32 33 34 35 36 37 38 39
Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-
40 41 42 43 44 45
Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

式I (SEQID NO:2)

其中:

第8位的Xaa是Ala, Gly, Ser或Val;

第9位的Xaa是Glu或Asp;

第11位的Xaa是Thr;

第14位的Xaa是Ser;

第16位的Xaa是Val, Ala, Leu, Ile, Trp或Phe;

第17位的Xaa是Ser;

第18位的Xaa是Ser, Tyr,或Lys;

第19位的Xaa是Tyr或Gln;

第20位的Xaa是Leu, Met,或Phe;

第21位的Xaa是Glu;

第22位的Xaa是Gly, Ser, Glu, Asp或Gln;

第23位的Xaa是Gln或Glu;

第24位的Xaa是Ala;

第25位的Xaa是Ala;

第26位的Xaa是Lys, Arg或His;

第27位的Xaa是Leu, Glu或Ile;

第30位的Xaa是Ala, Glu或Asp;

第31位的Xaa是Trp;

第32位的Xaa是Leu;

第33位的Xaa是Val, Ala, Ile或Lys;

第34位的Xaa是Asn, Lys, Arg或Gly;

第35位的Xaa是Gly;

第36位的Xaa是Gly, Arg或His;

第37位的Xaa是Pro, Gly或被去除;

第38位的Xaa是Ser或被去除;

第39位的Xaa是Ser或被去除;

第 40 位的 Xaa 是 Gly, Ser 或被去除 ;
第 41 位的 Xaa 是 Ala 或被去除 ;
第 42 位的 Xaa 是 Ser, Pro 或被去除 ;
第 43 位的 Xaa 是 Ser, Pro 或被去除 ;
第 44 位的 Xaa 是 Gly, Pro 或被去除 ; 并且
第 45 位的 Xaa 是 Ala, Ser 或被去除 ;

前提是当第 37, 38, 39, 40, 41, 42, 43 或 44 位的氨基酸被去除时, 该氨基酸下游的所有氨基酸也被去除 ; 第二个多肽是人白蛋白, 其中, 第一个多肽的 C 末端与第二个多肽的 N 末端融合。

2. 一种异源融合蛋白, 其具有第一个多肽, 该多肽具有 N 末端和与具有 N 末端和 C 末端的第二个多肽融合的 C 末端, 其中第一个多肽是具有式 I, 即 SEQ ID NO : 2 所示序列的 GLP-1 化合物

7 8 9 10 11 12 13 14 15 16 17
His-Xaa-Xaa-Gly-Xaa-Phe-Thr-Xaa-Asp-Xaa-Xaa-
18 19 20 21 22 23 24 25 26 27 28
Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Phe-
29 30 31 32 33 34 35 36 37 38 39
Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-
40 41 42 43 44 45
Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

式 I (SEQID NO : 2)

其中 :

第 8 位的 Xaa 是 Ala, Gly, Ser 或 Val ;
第 9 位的 Xaa 是 Glu 或 Asp ;
第 11 位的 Xaa 是 Thr ;
第 14 位的 Xaa 是 Ser ;
第 16 位的 Xaa 是 Val, Ala, Leu, Ile, Trp 或 Phe ;
第 17 位的 Xaa 是 Ser ;
第 18 位的 Xaa 是 Ser, Tyr, 或 Lys ;
第 19 位的 Xaa 是 Tyr 或 Gln ;
第 20 位的 Xaa 是 Leu, Met, 或 Phe ;
第 21 位的 Xaa 是 Glu ;
第 22 位的 Xaa 是 Gly, Ser, Glu, Asp 或 Gln ;
第 23 位的 Xaa 是 Gln 或 Glu ;
第 24 位的 Xaa 是 Ala ;
第 25 位的 Xaa 是 Ala ;
第 26 位的 Xaa 是 Lys, Arg 或 His ;
第 27 位的 Xaa 是 Leu, Glu 或 Ile ;
第 30 位的 Xaa 是 Ala, Glu 或 Asp ;

第 31 位的 Xaa 是 Trp ;
第 32 位的 Xaa 是 Leu ;
第 33 位的 Xaa 是 Val, Ala, Ile 或 Lys ;
第 34 位的 Xaa 是 Asn, Lys, Arg 或 Gly ;
第 35 位的 Xaa 是 Gly ;
第 36 位的 Xaa 是 Gly, Arg 或 His ;
第 37 位的 Xaa 是 Pro, Gly 或被去除 ;
第 38 位的 Xaa 是 Ser 或被去除 ;
第 39 位的 Xaa 是 Ser 或被去除 ;
第 40 位的 Xaa 是 Gly, Ser 或被去除 ;
第 41 位的 Xaa 是 Ala 或被去除 ;
第 42 位的 Xaa 是 Ser, Pro 或被去除 ;
第 43 位的 Xaa 是 Ser, Pro 或被去除 ;
第 44 位的 Xaa 是 Gly, Pro 或被去除 ;并且
第 45 位的 Xaa 是 Ala, Ser 或被去除 ;

前提是当第 37, 38, 39, 40, 41, 42, 43 或 44 位的氨基酸被去除时, 该氨基酸下游的所有氨基酸也被去除,

第二个多肽是人白蛋白, 其中第一个多肽的 C 末端通过肽接头与第二个多肽的 N 末端融合。

3. 如权利要求 2 所述的异源融合蛋白, 其中所述的肽接头是具有序列 [Gly-Gly-Gly-Gly-Ser]_n 的肽, 其中 n 是 1, 2, 3, 4, 5 或 6。

4. 权利要求 1-3 中任一项所述的异源融合蛋白, 其中, GLP-1 化合物具有不超过 6 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

5. 权利要求 4 的异源融合蛋白, 其中 GLP-1 化合物具有不超过 5 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

6. 权利要求 5 的异源融合蛋白, 其中 GLP-1 化合物具有不超过 4 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

7. 权利要求 6 的异源融合蛋白, 其中 GLP-1 化合物具有不超过 3 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

8. 权利要求 7 的异源融合蛋白, 其中 GLP-1 化合物具有不超过 2 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

9. 一种异源融合蛋白, 其具有第一个多肽, 该多肽具有 N 末端和与具有 N 末端和 C 末端的第二个多肽融合的 C 末端, 其中第一个多肽是具有式 I, 即 SEQ ID NO : 2 所示序列的 GLP-1 化合物

7	8	9	10	11	12	13	14	15	16	17
His	Xaa	Xaa	Gly	Xaa	Phe	Thr	Xaa	Asp	Xaa	Xaa
18	19	20	21	22	23	24	25	26	27	28
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Phe
29	30	31	32	33	34	35	36	37	38	39

Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-

40 41 42 43 44 45

Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

式 I (SEQID NO :2)

其中：

第 8 位的 Xaa 是 Ala, Gly, Ser 或 Val；

第 9 位的 Xaa 是 Glu 或 Asp；

第 11 位的 Xaa 是 Thr；

第 14 位的 Xaa 是 Ser；

第 16 位的 Xaa 是 Val, Ala, Leu, Ile, Trp 或 Phe；

第 17 位的 Xaa 是 Ser；

第 18 位的 Xaa 是 Ser, Tyr, 或 Lys；

第 19 位的 Xaa 是 Tyr 或 Gln；

第 20 位的 Xaa 是 Leu, Met, 或 Phe；

第 21 位的 Xaa 是 Glu；

第 22 位的 Xaa 是 Gly, Ser, Glu, Asp 或 Gln；

第 23 位的 Xaa 是 Gln 或 Glu；

第 24 位的 Xaa 是 Ala；

第 25 位的 Xaa 是 Ala；

第 26 位的 Xaa 是 Lys, Arg 或 His；

第 27 位的 Xaa 是 Leu, Glu 或 Ile；

第 30 位的 Xaa 是 Ala, Glu 或 Asp；

第 31 位的 Xaa 是 Trp；

第 32 位的 Xaa 是 Leu；

第 33 位的 Xaa 是 Val, Ala, Ile 或 Lys；

第 34 位的 Xaa 是 Asn, Lys, Arg 或 Gly；

第 35 位的 Xaa 是 Gly；

第 36 位的 Xaa 是 Gly, Arg 或 His；

第 37 位的 Xaa 是 Pro, Gly 或被去除；

第 38 位的 Xaa 是 Ser 或被去除；

第 39 位的 Xaa 是 Ser 或被去除；

第 40 位的 Xaa 是 Gly, Ser 或被去除；

第 41 位的 Xaa 是 Ala 或被去除；

第 42 位的 Xaa 是 Ser, Pro 或被去除；

第 43 位的 Xaa 是 Ser, Pro 或被去除；

第 44 位的 Xaa 是 Gly, Pro 或被去除；并且

第 45 位的 Xaa 是 Ala, Ser 或被去除；

前提是当第 37, 38, 39, 40, 41, 42, 43 或 44 位的氨基酸被去除时, 该氨基酸下游的所有氨基酸也被去除, 所述第二个多肽是免疫球蛋白的 Fc 部分, 其中第一个多肽的 C 末端与第

二个多肽的 N 末端融合。

10. 一种异源融合蛋白,其具有第一个多肽,该多肽具有 N 末端和与具有 N 末端和 C 末端的第二个多肽融合的 C 末端,其中第一个多肽是具有式 I,即 SEQ ID NO:2 所示序列的 GLP-1 化合物

7 8 9 10 11 12 13 14 15 16 17
His-Xaa-Xaa-Gly-Xaa-Phe-Thr-Xaa-Asp-Xaa-Xaa-
18 19 20 21 22 23 24 25 26 27 28
Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Phe-
29 30 31 32 33 34 35 36 37 38 39
Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-
40 41 42 43 44 45
Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

式 I (SEQID NO:2)

其中:

第 8 位的 Xaa 是 Ala, Gly, Ser 或 Val ;

第 9 位的 Xaa 是 Glu 或 Asp ;

第 11 位的 Xaa 是 Thr ;

第 14 位的 Xaa 是 Ser ;

第 16 位的 Xaa 是 Val, Ala, Leu, Ile, Trp 或 Phe ;

第 17 位的 Xaa 是 Ser ;

第 18 位的 Xaa 是 Ser, Tyr, 或 Lys ;

第 19 位的 Xaa 是 Tyr 或 Gln ;

第 20 位的 Xaa 是 Leu, Met, 或 Phe ;

第 21 位的 Xaa 是 Glu ;

第 22 位的 Xaa 是 Gly, Ser, Glu, Asp 或 Gln ;

第 23 位的 Xaa 是 Gln 或 Glu ;

第 24 位的 Xaa 是 Ala ;

第 25 位的 Xaa 是 Ala ;

第 26 位的 Xaa 是 lys, Arg 或 His ;

第 27 位的 Xaa 是 Leu, Glu 或 Ile ;

第 30 位的 Xaa 是 Ala, Glu 或 Asp ;

第 31 位的 Xaa 是 Trp ;

第 32 位的 Xaa 是 Leu ;

第 33 位的 Xaa 是 Val, Ala, Ile 或 Lys ;

第 34 位的 Xaa 是 Asn, Lys, Arg 或 Gly ;

第 35 位的 Xaa 是 Gly ;

第 36 位的 Xaa 是 Gly, Arg 或 His ;

第 37 位的 Xaa 是 Pro, Gly 或被去除 ;

第 38 位的 Xaa 是 Ser 或被去除 ;

- 第 39 位的 Xaa 是 Ser 或被去除；
- 第 40 位的 Xaa 是 Gly, Ser 或被去除；
- 第 41 位的 Xaa 是 Ala 或被去除；
- 第 42 位的 Xaa 是 Ser, Pro 或被去除；
- 第 43 位的 Xaa 是 Ser, Pro 或被去除；
- 第 44 位的 Xaa 是 Gly, Pro 或被去除；并且
- 第 45 位的 Xaa 是 Ala, Ser 或被去除；

前提是当第 37, 38, 39, 40, 41, 42, 43 或 44 位的氨基酸被去除时, 该氨基酸下游的所有氨基酸也被去除, 所述第二个多肽是免疫球蛋白的 Fc 部分, 其中第一个多肽的 C 末端通过肽接头与第二个多肽的 N 末端融合。

11. 如权利要求 10 所述的异源融合蛋白, 其中所述的肽接头是具有序列 [Gly-Gly-Gly-Gly-Ser]_n 的肽, 其中 n 是 1, 2, 3, 4, 5 或 6。

12. 权利要求 9-11 中任一项所述的异源融合蛋白, 其中, GLP-1 化合物具有不超过 6 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

13. 权利要求 12 的异源融合蛋白, 其中 GLP-1 化合物具有不超过 5 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

14. 权利要求 13 的异源融合蛋白, 其中 GLP-1 化合物具有不超过 4 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

15. 权利要求 14 的异源融合蛋白, 其中 GLP-1 化合物具有不超过 3 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

16. 权利要求 15 的异源融合蛋白, 其中 GLP-1 化合物具有不超过 2 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。

17. 用于治疗非胰岛素依赖型糖尿病患者的药物组合物, 其包含权利要求 1-3, 9-11 任一项的异源融合蛋白或其可药用盐, 以及一种或多种可药用赋形剂。

18. 用于治疗肥胖症患者的药物组合物, 其包含权利要求 1-3, 9-11 任一项的异源融合蛋白或其可药用盐, 以及一种或多种可药用赋形剂。

GLP-1 融合蛋白

[0001] 本申请是申请日为 2001 年 11 月 29 日,申请号为 01620232.2 的、发明名称和本发明相同的发明专利申请的分案申请。

[0002] 本发明涉及与一些蛋白融合的胰高血糖素样肽,包括其类似物和衍生物,所述蛋白具有延长胰高血糖素样肽的体内半衰期的作用。这些融合蛋白可以用于治疗非胰岛素依赖型糖尿病和各种状况。

[0003] 胰高血糖素样肽 -1 (GLP-1) 是一个含有 37 个氨基酸的、可对食物摄入作出应答的肽,是由肠 L- 细胞分泌的。已经发现胰高血糖素样肽 -1 可刺激胰岛素分泌 (促胰岛素活性),从而引起细胞对葡萄糖的吸收和血清葡萄糖水平下降 (见如 Mojsov, S., Int. J. Peptide Protein Research, 40 :333-343 (1992))。然而, GLP-1 活性很低。随后在第 6 位和第 7 位之间的内源性裂解产生了生物活性更强的 GLP-1 (7-37) OH 肽。已知有很多 GLP-1 类似物和衍生物,在此称作“GLP-1 化合物”。这些 GLP-1 类似物包括 GILA 毒蜥 (monster) 毒液中的 Exendins。Exendins 具有与天然 GLP-1 同源的序列,结合 GLP-1 受体并起始由 GLP-1 (7-37) OH 引起的负责多种活性的信号传导级联。

[0004] GLP-1 化合物具有各种重要的生理活性。例如,已经证明 GLP-1 刺激胰岛素释放,降低胰高血糖素分泌,抑制胃排空,并且增强葡萄糖利用。[Nauck, M. A., et al. (1993) Diabetologia 36 :741-744 ;Gutniak, M., et al. (1992) New England J. of Med. 326 : 13161322 ;Nauck, M. A., et al., (1993) J. Clin. Invest. 91 :301307]。

[0005] GLP-1 表现出作为非胰岛素依赖型糖尿病 (NIDDM) 的治疗方法的巨大潜能。市场上已经有许多口服药物,用于治疗与 NIDDM 相关的胰岛素抵抗。然而,当疾病进展时,患者必须开始刺激胰岛素释放,并最终进行涉及胰岛素注射的治疗。然而,当前刺激胰岛素释放的药物和胰岛素的实际给药也可导致低血糖症。GLP-1 活性受血糖水平的控制。当该水平降低到某一阈值水平时, GLP-1 无活性。因此。没有发生与 GLP-1 治疗相关的低血糖症的危险。

[0006] 然而,涉及 GLP-1 肽的治疗的用途受到了它们的迅速清除和短半衰期的限制。例如, GLP-1 (7-37) 的血清半衰期仅仅为 3-5 分钟。当皮下给药时, GLP-1 (7-36) 酰胺的作用时间为约 50 分钟。甚至抗内源性蛋白酶裂解的类似物和衍生物也没有足够长的半衰期以避免在 24 小时的时间段中重复给药。当需要在长时间内维持治疗剂的高血液水平时,治疗剂的迅速清除是不方便的,因为必须重复给药。此外,长作用时间的化合物对于过去的治疗方案只包括口服药物的糖尿病患者特别重要。这些患者通常对于包括多次药物注射的方案具有极为困难的时间过渡。

[0007] 本发明克服了与递送具有短血清半衰期的化合物相关的问题。本发明包括与另一种具有长的循环半衰期的蛋白,如免疫球蛋白或白蛋白的 Fc 部分融合的 GLP-1 化合物。

[0008] 一般地,小的治疗性肽难以操作,因为即使是它们的结构中的轻微改变也可以影响稳定性和 / 或生物活性。对于目前开发中的 GLP-1 化合物来说更是如此。例如,例如, GLP-1 (7-37) OH 具有进行构象改变的倾向,从主要为 α 螺旋结构改变成主要为 β 折叠结构。 β 折叠的形成导致被认为是无活性的物质的聚集。因此,能够开发半衰期增加的生物

活性 GLP-1 融合蛋白是出乎意料的。因此,由于对 GLP-1 (737) OH 本身和相对于所连接的 GLP-1 肽来说很大的融合配偶体进行工作的困难,这尤其是预料不到的。

[0009] 本发明的化合物包括一种异源融合蛋白,它包含第一个多肽,该多肽具有 N 末端和与具有 N 末端和 C 末端的第二个多肽融合的 C 末端,其中第一个多肽是 GLP-1 化合物,并且第二个多肽选自

[0010] a) 人白蛋白;

[0011] b) 人白蛋白类似物;和

[0012] c) 人白蛋白的片段,

[0013] 并且其中第一个多肽的 C 末端与第二个多肽的 N 末端融合。

[0014] 本发明的化合物还包括一种异源融合蛋白,它包含第一个多肽,该多肽具有 N 末端和与具有 N 末端和 C 末端的第二个多肽融合的 C 末端,其中第一个多肽是 GLP-1 化合物,并且第二个多肽选自

[0015] a) 人白蛋白;

[0016] b) 人白蛋白类似物;和

[0017] c) 人白蛋白的片段,

[0018] 并且其中第一个多肽的 C 末端通过肽接头与第二个多肽的 N 末端融合。肽接头优选选自:

[0019] a) 富含甘氨酸的肽;

[0020] b) 具有序列 $[\text{Gly-Gly-Gly-Gly-Ser}]_n$ 的肽,其中 n 是 1,2,3,4,5 或 6;和

[0021] c) 具有序列 $[\text{Gly-Gly-Gly-Gly-Ser}]_3$ 的肽。

[0022] 本发明的其它化合物包括一种异源融合蛋白,包含第一个多肽,该多肽具有 N 末端和与具有 N 末端和 C 末端的第二个多肽融合的 C 末端,其中第一个多肽是 GLP-1 化合物,并且第二个多肽选自

[0023] a) 免疫球蛋白的 Fc 部分;

[0024] b) 免疫球蛋白的 Fc 部分的类似物;和

[0025] c) 免疫球蛋白的 Fc 部分的片段,

[0026] 并且其中第一个多肽的 C 末端与第二个多肽的 N 末端融合。GLP-1 化合物可以通过肽接头与第二个多肽融合。肽接头优选选自:

[0027] a) 富含甘氨酸的肽;

[0028] b) 具有序列 $[\text{Gly-Gly-Gly-Gly-Ser}]_n$ 的肽,其中 n 是 1,2,3,4,5 或 6;和

[0029] c) 具有序列 $[\text{Gly-Gly-Gly-Gly-Ser}]_3$ 的肽。

[0030] 作为异源融合蛋白的一部分的 GLP-1 化合物一般优选具有不超过 6 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。GLP-1 化合物甚至更优选具有不超过 5 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。GLP-1 化合物最优选具有不超过 4,3 或 2 个不同于 GLP-1 (7-37) OH, GLP-1 (7-36) OH 或 Exendin-4 中的相应氨基酸的氨基酸。GLP-1 化合物优选为第 8 位具有甘氨酸或缬氨酸的异源融合蛋白的一部分。

[0031] 本发明还包括编码此处描述的异源融合蛋白的多核苷酸,包含这些多核苷酸的载体,和用此处描述的载体转染或转化的宿主细胞。还包括一种产生异源融合蛋白的方法,该

方法包括在以可检测量表达异源融合蛋白的条件下转录和翻译此处描述的多肽的步骤。

[0032] 本发明还包括一种使有需要的哺乳动物中的血糖水平正常化的方法,包括给予治疗有效量的此处描述的异源融合蛋白。

[0033] 参考以下附图对本发明进行进一步描述:

[0034] 图 1:包括铰链区,CH2 和 CH3 结构域的 IgG1 Fc 氨基酸序列。

[0035] 图 2:人血清白蛋白氨基酸序列。

[0036] 图 3:A. SDS-PAGE 凝胶和同一凝胶的免疫印迹,描述 IgG1-Fc 和 GLP-1-Fc 融合蛋白的分子量(第 1 道,分子量标准;第 2 道,纯化的 Fc;第 3 道,伪转化的培养基;第 4 道,Val⁸-GLP-1-Fc;第 5 道,Exendin-4-Fc)。B. SDS-PAGE 凝胶和同一凝胶的免疫印迹,描述人 HSA 和 GLP-1-HSA 融合蛋白的分子量(第 1 道,分子量标准;第 2 道,纯化的 HSA;第 3 道,伪转化的培养基;第 4 道,Val⁸-GLP-1-HSA;第 5 道,Val₈-GLP-1-[Gly-Gly-Gly-Gly-Ser]₃-HSA;第 6 道,Exendin-4-HSA;第 7 道,Exendin-4-[Gly-Gly-Gly-Gly-Ser]₃-HSA)。

[0037] 图 4:纯化 Fc,白蛋白和 GLP-1 融合蛋白的 SDS-PAGE 凝胶(第 1 道,分子量标准;第 2 道,纯化的 Fc;第 3 道,Val⁸-GLP-1-Fc;第 4 道,Exendin-4-Fc;第 5 道,分子量标准;第 6 道,Val⁸-GLP-1-HSA;第 7 道 Exendin-4-HSA;第 8 道,Exendin-4-[Gly-Gly-Gly-Gly-Ser]₃-HSA)。

[0038] 图 5:含有图 1 所示 Fc 区的表达克隆载体。

[0039] 图 6:含有图 2 所示 Fc 区的表达克隆载体。

[0040] 图 7:含有与图 2 所示白蛋白序列的 5' 框内融合的含 15 个氨基酸的接头的编码 DNA 的表达克隆载体。

[0041] 图 8:GLP-1 融合蛋白的体外剂量反应活性。

[0042] 图 9:GLP-1Fc 和 HSA 融合蛋白的药代动力学。

[0043] 图 10:两只正常禁食狗中对 Exendin-Fc 的葡萄糖动力学(Glucodynamic)反应。

[0044] 图 11:两只正常禁食狗中对 Exendin-Fc 的促胰岛素反应。

[0045] 图 12:编码人 IgG1 Fc 区的 DNA 序列。

[0046] 图 13:编码人白蛋白的 DNA 序列。

[0047] 本发明的融合蛋白包含与人白蛋白,人白蛋白类似物,人白蛋白的片段,免疫球蛋白的 Fc 部分,免疫球蛋白的 Fc 部分的类似物,或免疫球蛋白的 Fc 部分的片段融合的 GLP-1 化合物。GLP-1 化合物的 C 末端可以直接或通过肽接头与白蛋白或 Fc 蛋白的 N 末端融合。这些异源融合蛋白具有生物学活性并且与天然 GLP-1 相比,半衰期增加。

[0048] 组成异源融合蛋白的一部分的 GLP-1 化合物优选包括具有约 25- 约 39 个天然存在的或非天然存在的与天然 GLP-1 (7-37)OH 具有充分同源性的氨基酸的多肽,使得它们可以通过与胰腺 β 细胞上的 GLP-1 受体结合而表现出促胰岛素活性。GLP-1 化合物一般包括具有 GLP-1 (7-37)OH, GLP-1 (7-37)OH 的类似物, GLP-1 (7-37)OH 的片段或 GLP-1 (7-37)OH 的片段的类似物的氨基酸序列的多肽。GLP-1 (7-37)OH 具有 SEQ ID NO:1 的氨基酸序列:

[0049] 7 8 9 10 11 12 13 14 15 16 17

[0050] His-Ala-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-Ser-

[0051] 18 19 20 21 22 23 24 25 26 27 28

[0052] Ser-Tyr-Leu-Glu-Gly-Gln-Ala-Ala-Lys-Glu-Phe-

[0053] 29 30 31 32 33 34 35 36 37

[0054] Ile-Ala-Trp-Leu-Val-Lys-Gly-Arg-Gly

[0055] (SEQ ID NO :1)

[0056] 按照本领域中的习惯, GLP-1 (7-37) OH 的氨基末端定为第 7 位残基, 而羧基末端定为第 37 位残基。对多肽中的其他氨基酸依次进行编号, 如 SEQ ID NO :1 中所示。例如, 第 12 位为苯丙氨酸而第 22 位为甘氨酸。

[0057] GLP-1 化合物也包括“GLP-1 片段”。GLP-1 片段是从 GLP-1 (7-37) OH 或其类似物或衍生物的 N- 末端或 C- 末端截掉一个或多个氨基酸后得到的多肽。用于描述 GLP-1 (7-37) OH 的术语也适用于 GLP-1 片段。例如, GLP-1 (9-36) OH 表示从 N- 末端截掉 2 个氨基酸和从 C- 末端截掉 1 个氨基酸后得到的 GLP-1。此片段中的氨基酸是用与 GLP-1 (7-37) OH 中相应氨基酸的数字来表示的。例如, GLP-1 (9-36) OH 中的 N- 末端谷氨酸在第 9 位 ; 第 12 位是苯丙氨酸 ; 而第 22 位是甘氨酸, 这与 GLP-1 (7-37) OH 相同。对于 GLP-1 (7-36) OH, GLP-1 (7-37) OH 第 37 位的甘氨酸被去除。

[0058] GLP-1 化合物也包括在 GLP-1 (7-37) OH 或其片段或类似物的 N- 末端和 / 或 C- 末端添加一个或多个氨基酸后得到的多肽。这类 GLP-1 化合物含有的氨基酸优选最多可达约 39 个。“延长的” GLP-1 化合物中的氨基酸是用与 GLP-1 (7-37) OH 中相应氨基酸的数字来表示的。例如, 通过在 GLP-1 (7-37) OH 的 N- 末端添加两个氨基酸得到的 GLP-1 化合物的 N- 末端氨基酸在第 5 位 ; 而通过在 GLP-1 (7-37) OH 的 C- 末端添加一个氨基酸得到的 GLP-1 化合物的 C- 末端氨基酸在第 38 位 ; 因此, 与 GLP-1 (7-37) OH 中的序号相同, 在这两类“延长的” GLP-1 化合物中第 12 位是苯丙氨酸 ; 而第 22 位是甘氨酸。“延长的” GLP-1 化合物的 1-6 位氨基酸优选与 GLP-1 (7-37) OH 中相应的氨基酸相同, 或是对 GLP-1 (7-37) OH 中相应氨基酸的保守取代。“延长的” GLP-1 化合物的第 38-45 位氨基酸优选的与胰高血糖素或 Exendin-4 中相应的氨基酸相同, 或是对胰高血糖素或 Exendin-4 中相应氨基酸的保守取代。

[0059] 本发明的 GLP-1 化合物也包括“GLP-1 类似物”。GLP-1 类似物与 GLP-1 (7-37) OH 或其片段具有充分的同源性, 从而这些类似物也具有促胰岛素活性。优选地, GLP-1 类似物包含有 GLP-1 (7-37) OH 或其片段的氨基酸序列, 对其进行修饰, 这样其氨基酸序列中有 1、2、3、4 或 5 个氨基酸与 GLP-1 (7-37) OH 或 GLP-1 (7-37) OH 片段中相应位置上的氨基酸不同。在此处用来命名 GLP-1 化合物的命名法中, 取代氨基酸及其位置与母结构相同。例如, Glu²²-GLP-1 (7-37) OH 表示这样一个 GLP-1 化合物, 其中 GLP-1 (7-37) OH 正常的第 22 位甘氨酸由谷氨酸取代 ; Val⁸-Glu²²-GLP-1 (7-37) OH 表示这样一个 GLP-1 化合物, 其中 GLP-1 (7-37) OH 正常的第 8 位丙氨酸和第 22 位甘氨酸分别由缬氨酸和谷氨酸取代。

[0060] 本发明的 GLP-1 化合物也包括“GLP-1 衍生物”。GLP-1 衍生物也定义为具有 GLP-1 或 GLP-1 类似物的氨基酸序列的分子, 但在其一个或多个氨基酸侧基、 α 碳原子、末端氨基或末端羧基还额外具有化学修饰。化学修饰包括, 但不限于, 添加化学部分, 产生新的键, 和去除化学部分。氨基酸侧基的修饰包括, 但不限于, 赖氨酸 ϵ 氨基的酰化, 精氨酸, 组氨酸, 或赖氨酸的 N 烷基化, 谷氨酸或天冬氨酸的羧基的烷基化, 以及谷氨酰胺或天冬酰胺的脱酰胺。末端氨基的修饰包括, 但不限于, 脱氨基、N- 低级烷基、N- 二 - 低级烷基和 N- 酰基修饰。末端羧基的修饰包括, 但不限于, 酰胺、低级烷基酰基、二烷基酰胺和低级烷基酯修饰。

低级烷基为 C₁-C₄ 烷基。此外,可以用化学领域普通技术人员公知的保护基对一个或多个侧基或端基进行保护。氨基酸的 α 碳可以被单或二甲基化。

[0061] 任何 GLP-1 化合物都可以是本发明的异源融合蛋白的一部分,只要该 GLP-1 化合物本身可以结合 GLP-1 受体并通过 GLP-1 受体诱导信号传递。GLP-1 受体结合和信号传导可以分别用 EP 619,322 和美国专利 5,120,712 中描述的体外测定方法进行评估。

[0062] 许多有活性的 GLP-1 片段、类似物和衍生物是本领域公知的,这些类似物和衍生物的任意一个都可以是本发明的异源融合蛋白的一部分。此处提供了新的 GLP-1 类似物和本领域公知的 GLP-1 类似物和衍生物的一些例子。

[0063] 本领域公知的一些 GLP-1 类似物和 GLP-1 片段包括,例如, GLP-1(7-34) 和 GLP-1(7-35), GLP-1(7-36), Gln⁹-GLP-1(7-37), D-Gln⁹-GLP-1(7-37), Thr¹⁶-Lys¹⁸-GLP-1(7-37), 以及 Lys¹⁸-GLP-1(7-37)。GLP-1 类似物如 GLP-1(7-34) 和 GLP-1(7-35) 公开于美国专利 5,118,666。具有促胰岛素特征的 GLP-1 的生物加工形式,例如 GLP-1(7-36) 是已知的。其它已知的生物活性 GLP-1 化合物公开于 Buckley 等人的美国专利 5,545,618,以及 Adelhorst 等, J. Biol. Chem. 269:6275(1994)。

[0064] 一组优选的 GLP-1 类似物由式 I (SEQ ID NO:2) 的 GLP-1 类似物组成

[0065] 7 8 9 10 11 12 13 14 15 16 17

[0066] His-Xaa-Xaa-Gly-Xaa-Phe-Thr-Xaa-Asp-Xaa-Xaa-

[0067] 18 19 20 21 22 23 24 25 26 27 28

[0068] Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Phe-

[0069] 29 30 31 32 33 34 35 36 37 38 39

[0070] Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-

[0071] 40 41 42 43 44 45

[0072] Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

[0073] 式 I (SEQ ID NO:2)

[0074] 其中:

[0075] 第 8 位的 Xaa 是 Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys;

[0076] 第 9 位的 Xaa 是 Glu, Asp, 或 Lys;

[0077] 第 11 位的 Xaa 是 Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp, 或 Lys;

[0078] 第 14 位的 Xaa 是 Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys;

[0079] 第 16 位的 Xaa 是 Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp, Trp, 或 Lys;

[0080] 第 17 位的 Xaa 是 Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys;

[0081] 第 18 位的 Xaa 是 Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Trp, Tyr, 或 Lys;

[0082] 第 19 位的 Xaa 是 Tyr, Phe, Trp, Glu, Asp, Gln, 或 Lys;

[0083] 第 20 位的 Xaa 是 Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Trp, Tyr, 或 Lys;

[0084] 第 21 位的 Xaa 是 Glu, Asp, 或 Lys;

[0085] 第 22 位的 Xaa 是 Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys;

[0086] 第 23 位的 Xaa 是 Gln, Asn, Arg, Glu, Asp, 或 Lys;

[0087] 第 24 位的 Xaa 是 Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp, 或 Lys;

- [0088] 第 25 位的 Xaa 是 Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys ;
- [0089] 第 26 位的 Xaa 是 Lys, Arg, Gln, Glu, Asp, 或 His ;
- [0090] 第 27 位的 Xaa 是 Leu, Glu, Asp, 或 Lys ;
- [0091] 第 30 位的 Xaa 是 Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys ;
- [0092] 第 31 位的 Xaa 是 Trp, Phe, Tyr, Glu, Asp, 或 Lys ;
- [0093] 第 32 位的 Xaa 是 Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp, 或 Lys ;
- [0094] 第 33 位的 Xaa 是 Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp, 或 Lys ;
- [0095] 第 34 位的 Xaa 是 Asn, Lys, Arg, Glu, Asp, 或 His ;
- [0096] 第 35 位的 Xaa 是 Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys ;
- [0097] 第 36 位的 Xaa 是 Gly, Arg, Lys, Glu, Asp, 或 His ;
- [0098] 第 37 位的 Xaa 是 Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys, 或被去除 ;
- [0099] 第 38 位的 Xaa 是 Ser, Arg, Lys, Glu, Asp, 或 His, 或被去除 ;
- [0100] 第 39 位的 Xaa 是 Ser, Arg, Lys, Glu, Asp, 或 His, 或被去除 ;
- [0101] 第 40 位的 Xaa 是 Gly, Asp, Glu, 或 Lys, 或被去除 ;
- [0102] 第 41 位的 Xaa 是 Ala, Phe, Trp, Tyr, Glu, Asp, 或 Lys, 或被去除 ;
- [0103] 第 42 位的 Xaa 是 Ser, Pro, Lys, Glu, 或 Asp, 或被去除 ;
- [0104] 第 43 位的 Xaa 是 Ser, Pro, Glu, Asp, 或 Lys, 或被去除 ;
- [0105] 第 44 位的 Xaa 是 Gly, Pro, Glu, Asp, 或 Lys, 或被去除 ; 并且
- [0106] 第 45 位的 Xaa 是 Ala, Ser, Val, Glu, Asp, 或 Lys, 或被去除 ;
- [0107] 前提是当第 37, 38, 39, 40, 41, 42, 43 或 44 位的氨基酸被去除时, 该氨基酸下游的每个氨基酸也被去除。
- [0108] 式 I 的 GLP-1 化合物优选具有不超过 6 个不同于 GLP-1 (7-37) OH 或 Exendin-4 中的相应氨基酸的氨基酸。更优选具有不超过 5 个不同于 GLP-1 (7-37) OH 或 Exendin-4 中的相应氨基酸的氨基酸。甚至更优选具有不超过 4 个不同于 GLP-1 (7-37) OH 或 Exendin-4 中的相应氨基酸的氨基酸。
- [0109] 本发明的 GLP-1 化合物包括式 I 的衍生物, 如其 C-1-6- 酯或酰胺或 C-1-6- 烷基酰胺, 或 C-1-6- 二烷基酰胺。W099/43706 描述了式 I 的 GLP-1 化合物的衍生物, 在此引入作为参考。W099/43706 描述的衍生的式 I 化合物和未衍生的式 I 化合物都包括在本发明的范围内。
- [0110] 另一组优选的 GLP-1 化合物由式 II (SEQ ID NO :3) 的 GLP-1 类似物组成
- [0111] 7 8 9 10 11 12 13 14 15 16 17
- [0112] Xaa-Xaa-Xaa-Gly-Xaa-Xaa-Thr-Ser-Asp-Xaa-Ser-
- [0113] 18 19 20 21 22 23 24 25 26 27 28
- [0114] Xaa-Xaa-Leu-Glu-Gly-Xaa-Xaa-Ala-Xaa-Xaa-Phe-
- [0115] 29 30 31 32 33 34 35 36 37
- [0116] Ile-Xaa-Xaa-Leu-Xaa-Xaa-Xaa-Xaa-R
- [0117] 式 II (SEQ ID NO :3)
- [0118] 其中 :

[0119] 第 7 位的 Xaa 是 :L- 组氨酸, D- 组氨酸, 脱氨基 - 组氨酸, 2- 氨基 - 组氨酸, β - 羟基 - 组氨酸, 高组氨酸, α - 氟甲基 - 组氨酸或 α - 甲基 - 组氨酸 ;

[0120] 第 8 位的 Xaa 是 :Gly, Ala, Val, Leu, Ile, Ser, 或 Thr ;

[0121] 第 9 位的 Xaa 是 :Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu, 或 His ;

[0122] 第 11 位的 Xaa 是 :Asp, Glu, Arg, Thr, Ala, Lys, 或 His ;

[0123] 第 12 位的 Xaa 是 :His, Trp, Phe, 或 Tyr ;

[0124] 第 16 位的 Xaa 是 :Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu, 或 Ala ;

[0125] 第 18 位的 Xaa 是 :His, Pro, Asp, Glu, Arg, Ser, Ala, 或 Lys ;

[0126] 第 19 位的 Xaa 是 :Gly, Asp, Glu, Gln, Asn, Lys, Arg, 或 Cys ;

[0127] 第 23 位的 Xaa 是 :His, Asp, Lys, Glu, Gln, 或 Arg ;

[0128] 第 24 位的 Xaa 是 :Glu, Arg, Ala, 或 Lys ;

[0129] 第 26 位的 Xaa 是 :Trp, Tyr, Phe, Asp, Lys, Glu, 或 His ;

[0130] 第 27 位的 Xaa 是 :Ala, Glu, His, Phe, Tyr, Trp, Arg, 或 Lys ;

[0131] 第 30 位的 Xaa 是 :Ala, Glu, Asp, Ser, 或 His ;

[0132] 第 31 位的 Xaa 是 :Asp, Glu, Ser, Thr, Arg, Trp, 或 Lys ;

[0133] 第 33 位的 Xaa 是 :Asp, Arg, Val, Lys, Ala, Gly, 或 Glu ;

[0134] 第 34 位的 Xaa 是 :Glu, Lys, 或 Asp ;

[0135] 第 35 位的 Xaa 是 :Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, 或 Glu ;

[0136] 第 36 位的 Xaa 是 :Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu, 或 His ;

[0137] 第 37 位的 R 是 :Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, 或被去除。

[0138] 另一组优选的 GLP-1 化合物由式 III (SEQ ID NO :4) 的 GLP-1 类似物组成 :

[0139] 7 8 9 10 11 12 13 14 15 16 17

[0140] Xaa-Xaa-Glu-Gly-Xaa-Xaa-Thr-Ser-Asp-Xaa-Ser-

[0141] 18 19 20 21 22 23 24 25 26 27 28

[0142] Ser-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Xaa-Lys-Xaa-Phe-

[0143] 29 30 31 32 33 34 35 36 37

[0144] Ile-Xaa-Trp-Leu-Xaa-Xaa-Xaa-Xaa-R

[0145] 式 III (SEQ ID NO :4)

[0146] 其中 :

[0147] 第 7 位的 Xaa 是 :L- 组氨酸, D- 组氨酸, 脱氨基 - 组氨酸, 2- 氨基 - 组氨酸, β - 羟基 - 组氨酸, 高组氨酸, α - 氟甲基 - 组氨酸或 α - 甲基 - 组氨酸 ;

[0148] 第 8 位的 Xaa 是 :Gly, Ala, Val, Leu, Ile, Ser, 或 Thr ;

[0149] 第 11 位的 Xaa 是 :Asp, Glu, Arg, Thr, Ala, Lys, 或 His ;

[0150] 第 12 位的 Xaa 是 :His, Trp, Phe, 或 Tyr ;

[0151] 第 16 位的 Xaa 是 :Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, 或 Ala ;

[0152] 第 22 位的 Xaa 是 :Gly, Asp, Glu, Gln, Asn, Lys, Arg, 或 Cys ;

[0153] 第 23 位的 Xaa 是 :His, Asp, Lys, Glu, 或 Gln ;

[0154] 第 24 位的 Xaa 是 :Glu, His, Ala, 或 Lys ;

- [0155] 第 25 位的 Xaa 是 :Asp, Lys, Glu, 或 His ;
[0156] 第 27 位的 Xaa 是 :Ala, Glu, His, Phe, Tyr, Trp, Arg, 或 Lys ;
[0157] 第 30 位的 Xaa 是 :Ala, Glu, Asp, Ser, 或 His ;
[0158] 第 33 位的 Xaa 是 :Asp, Arg, Val, Lys, Ala, Gly, 或 Glu ;
[0159] 第 34 位的 Xaa 是 :Glu, Lys, 或 Asp ;
[0160] 第 35 位的 Xaa 是 :Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, 或 Glu ;
[0161] 第 36 位的 Xaa 是 :Arg, Glu, 或 His ;
[0162] 第 37 位的 R 是 :Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, 或被去除。

[0163] 另一组优选的 GLP-1 化合物由式 IV (SEQ ID NO :5) 的 GLP-1 类似物组成 :

[0164] 7 8 9 10 11 12 13 14 15 16 17

[0165] Xaa-Xaa-Glu-Gly-Thr-Xaa-Thr-Ser-Asp-Xaa-Ser-

[0166] 18 19 20 21 22 23 24 25 26 27 28

[0167] Ser-Tyr-Leu-Glu-Xaa-Xaa-Ala-Ala-Xaa-Glu-Phe-

[0168] 29 30 31 32 33 34 35 36 37

[0169] Ile-Xaa-Trp-Leu-Val-Lys-Xaa-Arg-R

[0170] 式 IV (SEQ ID NO :5)

[0171] 其中 :

[0172] 第 7 位的 Xaa 是 :L- 组氨酸, D- 组氨酸, 脱氨基 - 组氨酸, 2- 氨基 - 组氨酸, β - 羟基 - 组氨酸, 高组氨酸, α - 氟甲基 - 组氨酸或 α - 甲基组氨酸 ;

[0173] 第 8 位的 Xaa 是 :Gly, Ala, Val, Leu, Ile, Ser, 或 Thr ;

[0174] 第 12 位的 Xaa 是 :His, Trp, Phe, 或 Tyr ;

[0175] 第 16 位的 Xaa 是 :Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, 或 Ala ;

[0176] 第 22 位的 Xaa 是 :Gly, Asp, Glu, Gln, Asn, Lys, Arg, 或 Cys ;

[0177] 第 23 位的 Xaa 是 :His, Asp, Lys, Glu, 或 Gln ;

[0178] 第 26 位的 Xaa 是 :Asp, Lys, Glu, 或 His ;

[0179] 第 30 位的 Xaa 是 :Ala, Glu, Asp, Ser, 或 His ;

[0180] 第 35 位的 Xaa 是 :Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, 或 Glu ;

[0181] 第 37 位的 R 是 :Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, $-\text{NH}_2$, Gly, Gly-Pro, 或 Gly-Pro- NH_2 , 或被去除。

[0182] 另一组优选的 GLP-1 化合物由式 V (SEQ ID NO :6) 的 GLP-1 类似物组成 :

[0183] 7 8 9 10 11 12 13 14 15 16 17

[0184] Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-Ser-

[0185] 18 19 20 21 22 23 24 25 26 27 28

[0186] Ser-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Ala-Lys-Glu-Phe-

[0187] 29 30 31 32 33 34 35 36 37

[0188] Ile-Xaa-Trp-Leu-Val-Lys-Gly-Arg-R

[0189] 式 V (SEQ ID NO :6)

[0190] 其中 :

[0191] 第 7 位的 Xaa 是 :L- 组氨酸, D- 组氨酸, 脱氨基 - 组氨酸, 2- 氨基 - 组氨酸, β - 羟基 - 组氨酸, 高组氨酸, α - 氟甲基 - 组氨酸或 α - 甲基 - 组氨酸 ;

[0192] 第 8 位的 Xaa 是 :Gly, Ala, Val, Leu, Ile, Ser, 或 Thr ;

[0193] 第 22 位的 Xaa 是 :Gly, Asp, Glu, Gln, Asn, Lys, Arg, 或 Cys ;

[0194] 第 23 位的 Xaa 是 :His, Asp, Lys, Glu, 或 Gln ;

[0195] 第 24 位的 Xaa 是 :Ala, Glu, His, Phe, Tyr, Trp, Arg, 或 Lys ;

[0196] 第 30 位的 Xaa 是 :Ala, Glu, Asp, Ser, 或 His ;

[0197] 第 37 位的 R 是 :Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, 或被去除。

[0198] 式 I, II, III, IV 和 V 的优选 GLP-1 化合物包含 GLP-1 类似物或 GLP-1 类似物的片段, 其中该类似物或片段在第 8 位含有除丙氨酸以外的氨基酸 (第 8 位类似)。与天然 GLP-1 (7-37) OH 中的相应氨基酸相比, 这些第 8 位类似物优选含有在 9、11、12、16、18、22、23、24、26、27、30、31、33、34、35、36 和 37 位上进行的一或多个额外改变。与天然 GLP-1 (7-37) OH 中的相应氨基酸相比, 这些类似物也优选含有 6 个或少于 6 个改变。与天然 GLP-1 (7-37) OH 或 GLP-1 (7-36) OH 中的相应氨基酸相比, 这些类似物更优选含有 5 个或少于 5 个改变或含有 4 个或少于 4 个改变。与天然 GLP-1 (7-37) OH 或 GLP-1 (7-36) OH 中的相应氨基酸相比, 这些类似物甚至更优选的含有 3 个或少于 3 个改变。与天然 GLP-1 (7-37) OH 或 GLP-1 (7-36) OH 中的相应氨基酸相比, 这些类似物最优选含有 2 个或少于 2 个改变。

[0199] 已经发现, 式 II, III, IV 和 V 的化合物降低了聚集并产生不溶形式的倾向。这在融合蛋白的范围中也是重要的, 其中相对小的 GLP-1 肽尽管融合于大得多的蛋白, 也必须保持活性构象。本发明的融合蛋白包含的式 II, III, IV 和 V 的优选 GLP-1 化合物包括 GLP-1 类似物或 GLP-1 类似物的片段, 其中第 22 位的甘氨酸, 优选第 8 位的丙氨酸被其它氨基酸取代。

[0200] 当第 22 位氨基酸是天冬氨酸、精氨酸或赖氨酸时, 第 8 位氨基酸优选为甘氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸或甲硫氨酸, 而更优选是缬氨酸或甘氨酸。当第 22 位是磺酸如半胱磺酸时, 第 8 位优选为甘氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸或甲硫氨酸, 而更优选为缬氨酸或甘氨酸。

[0201] 其它优选的 GLP-1 化合物包括式 IV (SEQ ID NO :5) 的 GLP-1 类似物, 其中该类似物具有 GLP-1 (7-37) OH 的序列, 但第 8 位的氨基酸优选是甘氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸或甲硫氨酸, 更优选是缬氨酸或甘氨酸, 而第 30 位是谷氨酸、天冬氨酸、丝氨酸或组氨酸, 更优选为谷氨酸。

[0202] 其它优选的 GLP-1 化合物包括式 IV (SEQ ID NO :5) 的 GLP-1 类似物, 其中该类似物具有 GLP-1 (7-37) OH 的序列, 但第 8 位氨基酸优选是甘氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸或甲硫氨酸, 更优选为缬氨酸或甘氨酸, 而第 37 位是组氨酸、赖氨酸、精氨酸、苏氨酸、丝氨酸、谷氨酸、天冬氨酸、色氨酸、酪氨酸、苯丙氨酸, 而更优选是组氨酸。

[0203] 其它优选的 GLP-1 化合物包括式 IV (SEQ ID NO :5) 的 GLP-1 类似物, 其中该类似物具有 GLP-1 (7-37) OH 的序列, 但第 8 位氨基酸优选是甘氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸或甲硫氨酸, 更优选的为缬氨酸或甘氨酸, 而第 22 位是谷氨酸、赖氨酸、天冬氨酸或精氨酸, 更优选的是谷氨酸或赖氨酸, 而第 23 位是赖氨酸、精氨酸、谷氨酸、天冬

氨酸和组氨酸,更优选的是赖氨酸或谷氨酸。

[0204] 其它优选的 GLP-1 化合物包括式 V (SEQ ID NO :6) 的 GLP-1 类似物,其中该类似物具有 GLP-1 (7-37) OH 的序列,但第 8 位氨基酸优选是甘氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸或甲硫氨酸,更优选是缬氨酸或甘氨酸,而第 22 位是谷氨酸、赖氨酸、天冬氨酸或精氨酸,更优选是谷氨酸或赖氨酸,而第 27 位是丙氨酸、赖氨酸、精氨酸、色氨酸、酪氨酸、苯丙氨酸或组氨酸,更优选是丙氨酸。

[0205] 其它优选的 GLP-1 化合物包括式 I I 的 GLP-1 类似物,其中该类似物具有 GLP-1 (7-37) OH 的序列,但第 8 位氨基酸和选自第 9 位、11 位、12 位、16 位、18 位、22 位、23 位、24 位、26 位、27 位、30 位、31 位、33 位、34 位、35 位、36 位及 37 位氨基酸的 1、2 或 3 个氨基酸与天然的 GLP-1 (7-37) OH 中相应位置的氨基酸不同。

[0206] 其它优选的式 II 的 GLP-1 化合物包括:

[0207] Val⁸-GLP-1 (7-37) OH, Gly⁸-GLP-1 (7-37) OH, Glu²²-GLP-1 (7-37) OH, Asp²²-GLP-1 (7-37) OH, Arg²²-GLP-1 (7-37) OH, Lys²²-GLP-1 (7-37) OH, Cys²²-GLP-1 (7-37) OH, Val⁸-Glu²²-GLP-1 (7-37) OH, Val⁸-Asp²²-GLP-1 (7-37) OH, Val⁸-Arg²²-GLP-1 (7-37) OH, Val⁸-Lys²²-GLP-1 (7-37) OH, Val⁸-Cys²²-GLP-1 (7-37) OH, Gly⁸-Glu²²-GLP-1 (7-37) OH, Gly⁸-Asp²²-GLP-1 (7-37) OH, Gly⁸-Arg²²-GLP-1 (7-37) OH, Gly⁸-Lys²²-GLP-1 (7-37) OH, Gly⁸-Cys²²-GLP-1 (7-37) OH, Glu²²-GLP-1 (7-36) OH, Asp²²-GLP-1 (7-36) OH, Arg²²-GLP-1 (7-36) OH, Lys²²-GLP-1 (7-36) OH, Cys²²-GLP-1 (7-36) OH, Val⁸-Glu²²-GLP-1 (7-36) OH, Val⁸-Asp²²-GLP-1 (7-36) OH, Val⁸-Arg²²-GLP-1 (7-36) OH, Val⁸-Lys²²-GLP-1 (7-36) OH, Val⁸-Cys²²-GLP-1 (7-36) OH, Gly⁸-Glu²²-GLP-1 (7-36) OH, Gly⁸-Asp²²-GLP-1 (7-36) OH, Gly⁸-Arg²²-GLP-1 (7-36) OH, Gly⁸-Lys²²-GLP-1 (7-36) OH, Gly⁸-Cys²²-GLP-1 (7-36) OH, Lys²³-GLP-1 (7-37) OH, Val⁸-Lys²³-GLP-1 (7-37) OH, Gly⁸-Lys²³-GLP-1 (7-37) OH, His²⁴-GLP-1 (7-37) OH, Val⁸-His²⁴-GLP-1 (7-37) OH, Gly⁸-His²⁴-GLP-1 (7-37) OH, Lys²⁴-GLP-1 (7-37) OH, Val⁸-Lys²⁴-GLP-1 (7-37) OH, Gly⁸-Lys²³-GLP-1 (7-37) OH, Glu³⁰-GLP-1 (7-37) OH, Val⁸-Glu³⁰-GLP-1 (7-37) OH, Gly⁸-Glu³⁰-GLP-1 (7-37) OH, Asp³⁰-GLP-1 (7-37) OH, Val⁸-Asp³⁰-GLP-1 (7-37) OH, Gly⁸-Asp³⁰-GLP-1 (7-37) OH, Gln³⁰-GLP-1 (7-37) OH, Val⁸-Gln³⁰-GLP-1 (7-37) OH, Gly⁸-Gln³⁰-GLP-1 (7-37) OH, Tyr³⁰-GLP-1 (7-37) OH, Val⁸-Tyr³⁰-GLP-1 (7-37) OH, Gly⁸-Tyr³⁰-GLP-1 (7-37) OH, Ser³⁰-GLP-1 (7-37) OH, Val⁸-Ser³⁰-GLP-1 (7-37) OH, Gly⁸-Ser³⁰-GLP-1 (7-37) OH, His³⁰-GLP-1 (7-37) OH, Val⁸-His³⁰-GLP-1 (7-37) OH, Gly⁸-His³⁰-GLP-1 (7-37) OH, Glu³⁴-GLP-1 (7-37) OH, Val⁸-Glu³⁴-GLP-1 (7-37) OH, Gly⁸-Glu³⁴-GLP-1 (7-37) OH, Ala³⁴-GLP-1 (7-37) OH, Val⁸-Ala³⁴-GLP-1 (7-37) OH, Gly⁸-Ala³⁴-GLP-1 (7-37) OH, Gly³⁴-GLP-1 (7-37) OH, Val⁸-Gly³⁴-GLP-1 (7-37) OH, Gly⁸-Gly³⁴-GLP-1 (7-37) OH, Ala³⁵-GLP-1 (7-37) OH, Val⁸-Ala³⁵-GLP-1 (7-37) OH, Gly⁸-Ala³⁵-GLP-1 (7-37) OH, Lys³⁵-GLP-1 (7-37) OH, Val⁸-Lys³⁵-GLP-1 (7-37) OH, Gly⁸-Lys³⁵-GLP-1 (7-37) OH, His³⁵-GLP-1 (7-37) OH, Val⁸-His³⁵-GLP-1 (7-37) OH, Gly⁸-His³⁵-GLP-1 (7-37) OH, Pro³⁵-GLP-1 (7-37) OH, Val⁸-Pro³⁵-GLP-1 (7-37) OH, Gly⁸-Pro³⁵-GLP-1 (7-37) OH, Glu³⁵-GLP-1 (7-37) OH, Val⁸-Glu³⁵-GLP-1 (7-37) OH, Gly⁸-Glu³⁵-GLP-1 (7-37) OH, Val⁸-Ala²⁷-GLP-1 (7-37) OH, Val⁸-His³⁷-GLP-1 (7-37)

OH, Val⁸-Glu²²-Lys²³-GLP-1 (7-37) OH, Val⁸-Glu²²-Glu²³-GLP-1 (7-37) OH, Val⁸-Glu²²-Ala²⁷-GLP-1 (7-37) OH, Val⁸-Gly³⁴-Lys³⁵-GLP-1 (7-37) OH, Val⁸-His³⁷-GLP-1 (7-37) OH, Gly⁸-His³⁷-GLP-1 (7-37) OH, Val⁸-Glu²²-Ala²⁷-GLP-1 (7-37) OH, Gly⁸-Glu²²-Ala²⁷-GLP-1 (7-37) OH, Val⁸-Lys²²-Glu²³-GLP-1 (7-37) OH, 和 Gly⁸-Lys²²-Glu²³-GLP-1 (7-37) OH.

[0208] 用于本发明的另一组优选 GLP-1 类似物和衍生物由式 VI (SEQ ID NO :7) 的分子组成

[0209] R₁-X-Glu-Gly¹⁰-Thr-Phe-Thr-Ser-Asp¹⁵-Val-Ser-

[0210] Ser-Tyr-Leu²⁰-Y-Gly-Gln-Ala-Ala²⁵-Lys-Z-

[0211] Phe-Ile-Ala³⁰-Trp-Leu-Val-Lys-Gly³⁵-Arg-R₂

[0212] 式 VI (SEQ ID NO :7)

[0213] 其中 :R₁ 选自 L-组氨酸, D-组氨酸, 脱氨基-组氨酸, 2-氨基组氨酸, β-羟基-组氨酸, 高组氨酸, α-氟甲基-组氨酸, 和 α-甲基-组氨酸 ;X 选自 Ala, Gly, Val, Thr, Ile 和 α-甲基-Ala ;Y 选自 Glu, Gln, Ala, Thr, Ser 和 Gly ;Z 选自 Glu, Gln, Ala, Thr, Ser 和 Gly, 并且 R₂ 是 Gly-OH.

[0214] 用于本发明的另一组优选的 GLP-1 化合物公开于 WO 91/11457, 并且基本由以下化合物组成 :GLP-1 (7-34), GLP-1 (7-35), GLP-1 (7-36), 或 GLP-1 (7-37), 或其酰胺形式, 及其可药用盐, 它们具有选自下组的至少一种修饰 :

[0215] (a) 用甘氨酸, 丝氨酸, 半胱氨酸, 苏氨酸, 天冬酰胺, 谷氨酰胺, 酪氨酸, 丙氨酸, 缬氨酸, 异亮氨酸, 亮氨酸, 甲硫氨酸, 苯丙氨酸, 精氨酸, 或 D-赖氨酸取代第 26 位和 / 或 34 位的赖氨酸 ;或用甘氨酸, 丝氨酸, 半胱氨酸, 苏氨酸, 天冬酰胺, 谷氨酰胺, 酪氨酸, 丙氨酸, 缬氨酸, 异亮氨酸, 亮氨酸, 甲硫氨酸, 苯丙氨酸, 赖氨酸, 或 D-精氨酸取代第 36 位的精氨酸。

[0216] (b) 用抗氧化氨基酸取代第 31 位的色氨酸 ;

[0217] (c) 以下取代的至少一种 :用酪氨酸取代第 16 位的缬氨酸 ;用赖氨酸取代第 18 位的丝氨酸 ;用天冬氨酸取代第 21 位的谷氨酸 ;用丝氨酸取代第 22 位的甘氨酸 ;用精氨酸取代第 23 位的谷氨酰胺 ;用精氨酸取代第 24 位的丙氨酸 ;并且用谷氨酰胺取代第 26 位的赖氨酸, 和

[0218] (d) 以下取代的至少一种 :用甘氨酸, 丝氨酸, 或半胱氨酸取代第 8 位的丙氨酸 ;用天冬氨酸, 甘氨酸, 丝氨酸, 半胱氨酸, 苏氨酸, 天冬酰胺, 谷氨酰胺, 酪氨酸, 丙氨酸, 缬氨酸, 异亮氨酸, 亮氨酸, 甲硫氨酸, 或苯丙氨酸取代第 9 位的谷氨酸 ;用丝氨酸, 半胱氨酸, 苏氨酸, 天冬酰胺, 谷氨酰胺, 酪氨酸, 丙氨酸, 缬氨酸, 异亮氨酸, 亮氨酸, 甲硫氨酸, 或苯丙氨酸取代第 10 位的甘氨酸 ;用谷氨酸取代第 15 位的天冬氨酸 ;以及

[0219] (e) 用甘氨酸, 丝氨酸, 半胱氨酸, 苏氨酸, 天冬酰胺, 谷氨酰胺, 酪氨酸, 丙氨酸, 缬氨酸, 异亮氨酸, 亮氨酸, 甲硫氨酸, 或苯丙氨酸, 或组氨酸的 D 形式或 N-酰化或烷基化形式取代第 7 位的组氨酸 ;其中, 当取代为 (a), (b), (d), 和 (e) 时, 被取代的氨基酸可以任选为 D 形式, 并且在第 7 位被取代的氨基酸可以任选为 N-酰化或 N-烷基化形式。

[0220] 由于二肽基-肽酶 IV (DPP IV) 可以负责观察到的所给予的 GLP-1 的迅速体内灭活 [见, 例如, Mentlein, R., et al., Eur. J. Biochem., 214 :829-835 (1993)], 在融合蛋白范围内防止了 DPP IV 活性的 GLP-1 类似物和衍生物是优选的, 其中的 GLP-1 化

合物是 Gly⁸-GLP-1(7-37)OH, Val⁸-GLP-1(7-37)OH, α-甲基-Ala⁸-GLP-1(7-37)OH, 或 Gly⁸-Gln²¹-GLP-1(7-37)OH 的融合蛋白更优选。

[0221] 用于本发明的另一组优选 GLP-1 化合物由美国专利 5,512,549 所要求保护的式 VII (SEQ ID NO :8) 化合物组成, 在此引入该文献作为参考。

[0222]
$$\text{R}^1\text{-Ala-Glu-Gly}^{10}\text{-}$$

[0223] Thr-Phe-Thr-Ser-Asp¹⁵-Val-Ser-Ser-Tyr-Leu²⁰-

[0224] Glu-Gly-Gln-Ala-Ala²⁵-Xaa-Glu-Phe-Ile-Ala³⁰-

[0225] Trp-Leu-Val-Lys-Gly³⁵-Arg-R³

[0226] |

[0227] R²

[0228] 式 VII (SEQ ID NO :8)

[0229] 其中 R¹ 选自 4-咪唑丙基, 4-咪唑乙基, 或 4-咪唑-α, α-二甲基-乙基, R² 选自 C₆-C₁₀ 未分支的酰基, 或不存在; R³ 选自 Gly-OH 或 NH₂; 并且 Xaa 是 Lys 或 Arg。

[0230] 用于本发明的更优选的式 IV 化合物的 Xaa 是 Arg, 并且 R² 是 C₆-C₁₀ 未分支的酰基。用于本发明的甚至更优选的式 IV 化合物的 Xaa 是 Arg, R² 是 C₆-C₁₀ 未分支的酰基, 并且 R³ 是 Gly-OH。用于本发明的高度优选的式 IV 化合物的 Xaa 是 Arg, R² 是 C₆-C₁₀ 未分支的酰基, R³ 是 Gly-OH, 并且 R¹ 是 4-咪唑丙基。用于本发明的尤其优选的式 IV 化合物的 Xaa 是 Arg, R² 是 C₈ 未分支的酰基, R³ 是 Gly-OH, 并且 R¹ 是 4-咪唑丙基。

[0231] 优选的 GLP-1 化合物包括 GLP-1 类似物, 其中这些类似物或片段的主链在第 8 位含有除丙氨酸外的其他氨基酸 (8 位类似物)。主链的第 7 位也可包括 L-组氨酸、D-组氨酸或修饰形式的组氨酸如脱氨基组氨酸、2-氨基-组氨酸、β-羟基-组氨酸、高组氨酸、α-氟甲基-组氨酸或 α-甲基组氨酸。与天然 GLP-1(7-37)OH 中的相应氨基酸相比, 这些第 8 位类似物优选含有在 12、16、18、19、20、22、25、27、30、33 和 37 位上进行的一个或多个额外改变。与天然 GLP-1(7-37)OH 中的相应氨基酸相比, 这些第 8 位类似物更优选含有在 16、18、22、25 和 33 位上进行的一个或多个额外改变。

[0232] 在一个优选的实施方案中, GLP-1 类似物是 GLP-1(7-37)OH, 其中第 12 位的氨基酸选自色氨酸或酪氨酸。更优选地, 除了第 12 位的取代, 第 8 位的氨基酸被甘氨酸, 缬氨酸, 亮氨酸, 异亮氨酸, 丝氨酸, 苏氨酸, 或甲硫氨酸取代, 更优选被缬氨酸或甘氨酸取代。甚至更优选地, 除了第 12 位和第 8 位的取代, 第 22 位的氨基酸被谷氨酸取代。

[0233] 在另一种优选实施方案中, GLP-1 类似物是 GLP-1(7-37)OH, 其中第 16 位的氨基酸被色氨酸, 异亮氨酸, 亮氨酸, 苯丙氨酸, 或酪氨酸取代。更优选地, 除了第 16 位的取代, 第 8 位的氨基酸被甘氨酸, 缬氨酸, 亮氨酸, 异亮氨酸, 丝氨酸, 苏氨酸, 或甲硫氨酸或甲硫氨酸取代, 更优选被缬氨酸或甘氨酸取代。甚至更优选地, 除了第 12 位和第 8 位的取代, 第 22 位的氨基酸被谷氨酸取代。除了第 16 位和第 8 位的取代, 第 30 位的氨基酸也优选被谷氨酸取代。除了第 16 位和第 8 位的取代, 第 37 位的氨基酸也优选被组氨酸取代。

[0234] 在另一种优选实施方案中, GLP-1 类似物是 GLP-1(7-37)OH, 其中第 18 位的氨基酸选自色氨酸, 酪氨酸, 苯丙氨酸, 赖氨酸, 亮氨酸, 或异亮氨酸, 优选色氨酸, 酪氨酸, 和异亮氨酸。更优选地, 除了第 18 位的取代, 第 8 位的氨基酸被甘氨酸, 缬氨酸, 亮氨酸, 异亮氨酸, 丝氨酸, 苏氨酸, 或甲硫氨酸取代, 更优选被缬氨酸或甘氨酸取代。甚至更优选地, 除了

第 18 位和第 8 位的取代,第 22 位的氨基酸被谷氨酸取代。除了第 18 位和第 8 位的取代,第 30 位的氨基酸也优选被谷氨酸取代。除了第 18 位和第 8 位的取代,第 37 位的氨基酸也优选被组氨酸取代。

[0235] 在另一种优选实施方案中,GLP-1 类似物是 GLP-1 (7-37) OH,其中第 19 位的氨基酸选自色氨酸或苯丙氨酸,优选色氨酸。更优选地,除了第 19 位的取代,第 8 位的氨基酸被甘氨酸,缬氨酸,亮氨酸,异亮氨酸,丝氨酸,苏氨酸,或甲硫氨酸取代,更优选被缬氨酸或甘氨酸取代。甚至更优选地,除了第 19 位和第 8 位的取代,第 22 位的氨基酸被谷氨酸取代。除了第 19 位和第 8 位的取代,第 30 位的氨基酸也优选被谷氨酸取代。除了第 19 位和第 8 位的取代,第 37 位的氨基酸也优选被组氨酸取代。

[0236] 在另一种优选实施方案中,GLP-1 类似物是 GLP-1 (7-37) OH,其中第 20 位的氨基酸选自苯丙氨酸,酪氨酸,或色氨酸。更优选地,除了第 20 位的取代,第 8 位的氨基酸被甘氨酸,缬氨酸,亮氨酸,异亮氨酸,丝氨酸,苏氨酸,或甲硫氨酸取代,更优选被缬氨酸或甘氨酸取代。甚至更优选地,除了第 20 位和第 8 位的取代,第 22 位的氨基酸被谷氨酸取代。除了第 20 位和第 8 位的取代,第 30 位的氨基酸也优选被谷氨酸取代。除了第 20 位和第 8 位的取代,第 37 位的氨基酸也优选被组氨酸取代。

[0237] 在另一种优选实施方案中,GLP-1 类似物是 GLP-1 (7-37) OH,其中第 25 位的氨基酸选自缬氨酸,异亮氨酸,和亮氨酸,优选缬氨酸。更优选地,除了第 25 位的取代,第 8 位的氨基酸被甘氨酸,缬氨酸,亮氨酸,异亮氨酸,丝氨酸,苏氨酸,或甲硫氨酸取代,更优选被缬氨酸或甘氨酸取代。甚至更优选地,除了第 25 位和第 8 位的取代,第 22 位的氨基酸被谷氨酸取代。除了第 25 位和第 8 位的取代,第 30 位的氨基酸也优选被谷氨酸取代。除了第 25 位和第 8 位的取代,第 37 位的氨基酸也优选被组氨酸取代。

[0238] 在另一种优选实施方案中, GLP-1 类似物是 GLP-1 (7-37) OH,其中第 27 位的氨基酸选自异亮氨酸或丙氨酸。更优选地,除了第 26 位的取代,第 8 位的氨基酸被甘氨酸,缬氨酸,亮氨酸,异亮氨酸,丝氨酸,苏氨酸,或甲硫氨酸取代,更优选被缬氨酸或甘氨酸取代。甚至更优选地,除了第 27 位和第 8 位的取代,第 22 位的氨基酸被谷氨酸取代。除了第 27 位和第 8 位的取代,第 30 位的氨基酸也优选被谷氨酸取代。除了第 27 位和第 8 位的取代,第 37 位的氨基酸也优选被组氨酸取代。

[0239] 在另一种优选实施方案中,GLP-1 类似物是 GLP-1 (7-37) OH,其中第 33 位的氨基酸是异亮氨酸。更优选地,除了第 33 位的取代,第 8 位的氨基酸被甘氨酸,缬氨酸,亮氨酸,异亮氨酸,丝氨酸,苏氨酸,或甲硫氨酸取代,更优选被缬氨酸或甘氨酸取代。甚至更优选地,除了第 33 位和第 8 位的取代,第 22 位的氨基酸被谷氨酸取代。除了第 33 位和第 8 位的取代,第 30 位的氨基酸也优选被谷氨酸取代。除了第 33 位和第 8 位的取代,第 37 位的氨基酸也优选被组氨酸取代。

[0240] GLP-1 化合物在以下一个或多个位置具有修饰 :8、12、16、18、19、20、22、25、27、30、33、和 37。这些 GLP-1 化合物与 GLP-1 (7-37) OH 相比效力增加,并且包含式 IX (SEQ ID NO :12) 的氨基酸序列

[0241] Xaa₇-Xaa₈-Glu-Gly-Thr-Xaa₁₂-Thr-Ser-Asp-Xaa₁₆-Ser-

[0242] Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Gln-Ala-Xaa₂₅-Lys-Xaa₂₇-

[0243] Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Lys-Gly-Arg-Xaa₃₇

[0244] 式 IX (SEQ ID NO :12)

[0245] 其中：

[0246] Xaa₇ 是：L- 组氨酸，D- 组氨酸，脱氨基 - 组氨酸，2- 氨基 - 组氨酸，β - 羟基 - 组氨酸，高组氨酸，α - 氟甲基 - 组氨酸，或 α - 甲基 - 组氨酸；

[0247] Xaa₈ 是：Ala, Gly, Val, Leu, Ile, Ser, 或 Thr；

[0248] Xaa₁₂ 是：Phe, Trp, 或 Tyr；

[0249] Xaa₁₆ 是：Val, Trp, Ile, Leu, Phe, 或 Tyr；

[0250] Xaa₁₈ 是：Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val；

[0251] Xaa₁₉ 是：Tyr, Trp, 或 Phe；

[0252] Xaa₂₀ 是：Leu, Phe, Tyr, 或 Trp；

[0253] Xaa₂₂ 是：Gly, Glu, Asp, 或 Lys；

[0254] Xaa₂₅ 是：Ala, Val, Ile, 或 Leu；

[0255] Xaa₂₇ 是：Glu, Ile, 或 Ala；

[0256] Xaa₃₀ 是：Ala 或 Glu；

[0257] Xaa₃₃ 是：Val, 或 Ile；并且

[0258] Xaa₃₇ 是：Gly, His, NH₂, 或不存在。

[0259] 一些优选的式 IX 的 GLP-1 化合物包括

[0260] GLP-1 (7-37) OH, GLP-1 (7-36) -NH₂, Gly⁸-GLP-1 (7-37) OH, Gly⁸-GLP-1 (7-36) NH₂, Val⁸-GLP-1 (7-37) OH, Val⁸-GLP-1 (7-36) NH₂, Leu⁸-GLP-1 (7-37) OH, Leu⁸-GLP-1 (7-36) NH₂, Ile⁸-GLP-1 (7-37) OH, Ile⁸-GLP-1 (7-36) NH₂, Ser⁸-GLP-1 (7-37) OH, Ser⁸-GLP-1 (7-36) NH₂, Thr⁸-GLP-1 (7-37) OH, Thr⁸-GLP-1 (7-36) NH₂, Val⁸-Tyr¹²-GLP-1 (7-37) OH, Val⁸-Tyr¹²-GLP-1 (7-36) NH₂, Val⁸-Tyr¹⁶-GLP-1 (7-37) OH, Val⁸-Tyr¹⁶-GLP-1 (7-36) NH₂, Val⁸-Glu²²-GLP-1 (7-37) OH, Val⁸-Glu²²-GLP-1 (7-36) NH₂, Gly⁸-Glu²²-GLP-1 (7-37) OH, Gly⁸-Glu²²-GLP-1 (7-36) NH₂, Val⁸-Asp²²-GLP-1 (7-37) OH, Val⁸-Asp²²-GLP-1 (7-36) NH₂, Gly⁸-Asp²²-GLP-1 (7-37) OH, Gly⁸-Asp²²-GLP-1 (7-36) NH₂, Val⁸-Lys²²-GLP-1 (7-37) OH, Val⁸-Lys²²-GLP-1 (7-36) NH₂, Gly⁸-Lys²²-GLP-1 (7-37) OH, Gly⁸-Lys²²-GLP-1 (7-36) NH₂, Leu⁸-Glu²²-GLP-1 (7-37) OH, Leu⁸-Glu²²-GLP-1 (7-36) NH₂, Ile⁸-Glu²²-GLP-1 (7-37) OH, Ile⁸-Glu²²-GLP-1 (7-36) NH₂, Leu⁸-Asp²²-GLP-1 (7-37) OH, Leu⁸-Asp²²-GLP-1 (7-36) NH₂, Ile⁸-Asp²²-GLP-1 (7-37) OH, Ile⁸-Asp²²-GLP-1 (7-36) NH₂, Leu⁸-Lys²²-GLP-1 (7-37) OH, Leu⁸-Lys²²-GLP-1 (7-36) NH₂, Ile⁸-Lys²²-GLP-1 (7-37) OH, Ile⁸-Lys²²-GLP-1 (7-36) NH₂, Ser⁸-Glu²²-GLP-1 (7-37) OH, Ser⁸-Glu²²-GLP-1 (7-36) NH₂, Thr⁸-Glu²²-GLP-1 (7-37) OH, Thr⁸-Glu²²-GLP-1 (7-36) NH₂, Ser⁸-Asp²²-GLP-1 (7-37) OH, Ser⁸-Asp²²-GLP-1 (7-36) NH₂, Thr⁸-Asp²²-GLP-1 (7-37) OH, Thr⁸-Asp²²-GLP-1 (7-36) NH₂, Ser⁸-Lys²²-GLP-1 (7-37) OH, Ser⁸-Lys²²-GLP-1 (7-36) NH₂, Thr⁸-Lys²²-GLP-1 (7-37) OH, Thr⁸-Lys²²-GLP-1 (7-36) NH₂, Glu²²-GLP-1 (7-37) OH, Glu²²-GLP-1 (7-36) NH₂, Asp²²-GLP-1 (7-37) OH, Asp²²-GLP-1 (7-36) NH₂, Lys²²-GLP-1 (7-37) OH, Lys²²-GLP-1 (7-36) NH₂, Val⁸-Ala²⁷-GLP-1 (7-37) OH, Val⁸-Glu²²-Ala²⁷-GLP-1 (7-37) OH, Val⁸-Glu³⁰-GLP-1 (7-37) OH, Val⁸-Glu³⁰-GLP-1 (7-36) NH₂, Gly⁸-Glu³⁰-GLP-1 (7-37) OH, Gly⁸-Glu³⁰-GLP-1 (7-36) NH₂, Leu⁸-Glu³⁰-GLP-1 (7-37) OH, Leu⁸-Glu³⁰-GLP-1 (7-36) NH₂, Ile⁸-Glu³⁰-GLP-1 (7-37) OH, Ile⁸-Glu³⁰-GLP-1 (7-36)

NH₂, Ser⁸-Glu³⁰-GLP-1 (7-37) OH, Ser⁸-Glu³⁰-GLP-1 (7-36) NH₂, Thr⁸-Glu³⁰-GLP-1 (7-37) OH, Thr⁸-Glu³⁰-GLP-1 (7-36) NH₂, Val⁸-His³⁷-GLP-1 (7-37) OH, Val⁸-His³⁷-GLP-1 (7-36) NH₂, Gly⁸-His³⁷-GLP-1 (7-37) OH, Gly⁸-His³⁷-GLP-1 (7-36) NH₂, Leu⁸-His³⁷-GLP-1 (7-37) OH, Leu⁸-His³⁷-GLP-1 (7-36) NH₂, Ile⁸-His³⁷-GLP-1 (7-37) OH, Ile⁸-His³⁷-GLP-1 (7-36) NH₂, Ser⁸-His³⁷-GLP-1 (7-37) OH, Ser⁸-His³⁷-GLP-1 (7-36) NH₂, Thr⁸-His³⁷-GLP-1 (7-37) OH, Thr⁸-His³⁷-GLP-1 (7-36) NH₂.

[0261] 一些优选的式 IX 的 GLP-1 化合物具有多个取代, 包括 GLP-1 (7-37) OH, 其中第 8 位是缬氨酸或甘氨酸, 第 22 位是谷氨酸, 第 16 位是酪氨酸, 亮氨酸或色氨酸, 第 18 位是酪氨酸, 色氨酸, 或异亮氨酸, 第 25 位是缬氨酸, 并且第 33 位是异亮氨酸。其它优选的 GLP-1 化合物包括以下化合物:

[0262] Val⁸-Tyr¹⁶-GLP-1 (7-37) OH, Val⁸-Tyr¹²-Glu²²-GLP-1 (7-37) OH, Val⁸-Tyr¹⁶-Phe¹⁹-GLP-1 (7-37) OH, Val⁸-Tyr¹⁶-Glu²²-GLP-1 (7-37) OH, Val⁸-Trp¹⁶-Glu²²-GLP-1 (7-37) OH, Val⁸-Leu¹⁶-Glu²²-GLP-1 (7-37) OH, Val⁸-Ile¹⁶-Glu²²-GLP-1 (7-37) OH, Val⁸-Phe¹⁶-Glu²²-GLP-1 (7-37) OH, Val⁸-Trp¹⁸-Glu²²-GLP-1 (7-37) OH, Val⁸-Tyr¹⁸-Glu²²-GLP-1 (7-37) OH, Val⁸-Phe¹⁸-Glu²²-GLP-1 (7-37) OH, 和 Val⁸-Ile¹⁸-Glu²²-GLP-1 (7-37) OH.

[0263] 本发明的 GLP-1 化合物也包括 Exendin 化合物。Exendin-3 和 Exendin-4 是首先从 Helodermatidae 蜥蜴毒液中分离的生物活性肽, 并且证明其结合 GLP-1 受体并且在哺乳动物体壁细胞中刺激 cAMP- 依赖性 H⁺ 产生。Exendin-3 和 Exendin-4 都是 39 个氨基酸的肽, 大约 53% 同源于 GLP-1。它们作为 GLP-1 活性的潜在激动剂。引人注意的是, 称作 Exendin (9-39 个氨基酸) 的 Exendin 的 N 端截短衍生物是 Exendin-3, Exendin-4 和 GLP-1 的抑制剂。

[0264] Exendin 化合物通常包括具有 Exendin-3, Exendin-4 或其类似物或片段的氨基酸序列的多肽。Exendin-3 和 Exendin-4 公开于美国专利 5, 424, 286。

[0265] Exendin-3 具有 SEQ ID NO:9 的氨基酸序列:

[0266] 7 8 9 10 11 12 13 14 15 16 17

[0267] His-Ser-Asp-Gly-Thr-Phe-Thr-Ser-Asp-Leu-Ser-

[0268] 18 19 20 21 22 23 24 25 26 27 28

[0269] Lys-Gln-Met-Glu-Glu-Glu-Ala-Val-Arg-Leu-Phe-

[0270] 29 30 31 32 33 34 35 36 37 38 39

[0271] Ile-Glu-Trp-Leu-Lys-Asn-Gly-Gly-Pro-Ser-Ser-

[0272] 40 41 42 43 44 45

[0273] Gly-Ala-Pro-Pro-Pro-Ser

[0274] (SEQ ID NO:9)

[0275] Exendin-3 具有 SEQ ID NO:10 的氨基酸序列:

[0276] 7 8 9 10 11 12 13 14 15 16 17

[0277] His-Gly-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Leu-Ser-

[0278] 18 19 20 21 22 23 24 25 26 27 28

[0279] Lys-Gln-Met-Glu-Glu-Glu-Ala-Val-Arg-Leu-Phe-

[0280] 29 30 31 32 33 34 35 36 37 38 39

[0281] Ile-Glu-Trp-Leu-Lys-Asn-Gly-Gly-Pro-Ser-Ser-

[0282] 40 41 42 43 44 45

[0283] Gly-Ala-Pro-Pro-Pro-Ser

[0284] (SEQ ID NO:10)

[0285] GLP-1 化合物也包括作为从 Exendin 或 Exendin 类似物的 N 末端和 / 或 C 末端截短一个或多个氨基酸而得到的多肽的 Exendin 片段。此外, GLP-1 化合物包括将一个或多个氨基酸添加到 Exendin 或其片段的 N 末端和 / 或 C 末端得到的 Exendin 多肽。这种类型的 Exendin 化合物最多具有约 45 个氨基酸。

[0286] GLP-1 化合物也称作“Exendin 类似物”。Exendin 类似物与 Exendin-4, Exendin-3 或其片段具有足够同源性, 使得该类似物具有促胰岛素活性。可以用例如 EP 619, 322 和美国专利 5, 120, 712 中描述的体外测定方法评估 Exendin 片段和 / 或类似物的活性。

[0287] Exendin 类似物优选具有修饰的 Exendin-4 或其片段的氨基酸序列, 使得 1、2、3、4 或 5 个氨基酸不同于 Exendin-4 或 Exendin-4 片段中相应位置的氨基酸。在此处使用的命名系统中, 取代氨基酸及其位置表示于母体结构前。例如 Val⁸-Exendin-4 表示通常存在于 Exendin-4 第 8 位的甘氨酸被缬氨酸缬氨酸取代的 Exendin 化合物。

[0288] 另一组优选的 GLP-1 化合物由式 VIII (SEQ ID NO:11) 的 GLP-1/Exendin-4 类似物组成。

[0289] 7 8 9 10 11 12 13 14 15 16 17

[0290] Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Xaa-Ser-

[0291] 18 19 20 21 22 23 24 25 26 27 28

[0292] Xaa-Xaa-Xaa-Glu-Xaa-Xaa-Ala-Xaa-Xaa-Xaa-Phe-

[0293] 29 30 31 32 33 34 35 36 37

[0294] Ile-Xaa-Trp-Leu-Xaa-Xaa-Gly-Xaa-R

[0295] 式 VIII (SEQ ID NO:11)

[0296] 其中:

[0297] 第 7 位的 Xaa 是:L- 组氨酸, D- 组氨酸, 脱氨基- 组氨酸, 2- 氨基- 组氨酸, β - 羟基- 组氨酸, 高组氨酸, α - 氟甲基- 组氨酸或 α - 甲基- 组氨酸;

[0298] 第 8 位的 Xaa 是:Gly, Ala, 或 Val;

[0299] 第 16 位的 Xaa 是:Leu 或 Val;

[0300] 第 18 位的 Xaa 是:Lys 或 Ser;

[0301] 第 19 位的 Xaa 是:Gln 或 Tyr;

[0302] 第 20 位的 Xaa 是:Met 或 Leu;

[0303] 第 22 位的 Xaa 是:Glu 或 Gln;

[0304] 第 23 位的 Xaa 是:Glu, 或 Gln;

[0305] 第 25 位的 Xaa 是:Val 或 Ala;

[0306] 第 26 位的 Xaa 是:Arg 或 Lys;

[0307] 第 27 位的 Xaa 是:Leu 或 Glu;

[0308] 第 30 位的 Xaa 是:Glu 或 Ala;

[0309] 第 33 位的 Xaa 是 :Val 或 Lys ;

[0310] 第 34 位的 Xaa 是 :Asn 或 Lys ;

[0311] 第 36 位的 Xaa 是 :Gly 或 Arg ;并且

[0312] 第 37 位的 R 是 :Gly,Pro,Pro-Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser,或不存在。该类中的 18 种不同的化合物见表 6。

[0313] 其它用于本发明的 Exend in 类似物描述于 PCT 专利申请 W099/25728(Beeley et al.),W0 99/25727 Beeley et al.),W0 98/05351(Young et al.),W0 99/40788(Young et al.),W0 99/07404(Beeley et al),和 W0 99/43708(Knudsen et al)。

[0314] 本发明的 GLP-1 融合蛋白可以包含糖基化位点。糖基化是一种化学修饰,其中将糖部分添加到蛋白的特定位点。蛋白的糖基化在保证成熟蛋白的正确电荷、构象和稳定性以及将蛋白导向于细胞表面和蛋白的最终分泌中起作用。最重要的是,糖基化影响了许多蛋白的体内清除速度。糖可以是 O- 连接的或 N- 连接的。一般地,O- 连接的糖添加到丝氨酸和苏氨酸的羟基氧,而 N- 连接的糖添加到天冬酰胺的酰胺氮。N 糖基化的共有位点是 Asn X1 X2 其中 X1 是除 Pro 外的任何氨基酸,X2 是 Ser 或 Thr。

[0315] GLP-1 化合物在体内一般是不糖基化的;然而,令人感兴趣的是,包含具有与 Fc 序列融合的 C 末端延伸片段的本发明的 GLP-1 化合物在 C 末端延伸片段的最后一个丝氨酸被糖基化(SSGAPPS*),并且在 Fc 的 N 末端区的第 11 位的苏氨酸被糖基化(AEPKSCDKTHT*CPPC...)。

[0316] 异源 Fc 融合蛋白 :

[0317] 此处描述的 GLP-1 化合物可以直接或通过肽接头与免疫球蛋白的 Fc 部分融合。

[0318] 免疫球蛋白是含有由二硫键连接在一起的多肽链的分子,一般具有两条轻链和两条重链。在每条链中,一个结构域(V)具有依赖于分子的抗体特异性的可变氨基酸序列。另一结构域(C)具有同一类分子所共有的恒定序列。

[0319] 此处用到的免疫球蛋白的 Fc 部分具有免疫学领域术语的常用含义。具体地,该术语是指通过从抗体去除两个抗原结合区(Fab 片段)而得到的抗体片段。去除 Fab 片段的一种方法是用木瓜蛋白酶消化免疫球蛋白。因此,Fc 部分是由两条重链的恒定区的几乎相等大小的片段形成的,这两条链是通过非共价相互作用和二硫键而缔合。Fc 部分可以包括铰链区并且延伸通过 CH2 和 CH3 结构域到达抗体 C 末端。人和小鼠免疫球蛋白的代表性铰链区可以见 Antibody Engineering, A Practical Guide, Borrebaeck, C. A. K., ed., W. H. Freeman and Co., 1992, 在此引入其教导作为参考。Fc 部分可以进一步包括一个或多个糖基化位点。包含铰链区、CH2 和 CH3 结构域,以及第 82 位的一个 N- 糖基化位点的代表性 Fc 部分的氨基酸序列表示于图 1。

[0320] 有 5 种具有不同效应物特征和药代动力学特征的人免疫球蛋白 Fc 区:IgG, IgA, IgM, IgD, 和 IgE。IgG 是血清中含量最高的免疫球蛋白。IgG 在所有免疫球蛋白中也是血清半衰期最长的(23 天)。与其它免疫球蛋白不同,IgG 在与 Fc 受体结合后有效再循环。有四种 IgG 亚类 G1, G2, G3, 和 G4, 每一种都有不同的效应物功能。G1, G2 和 G3 可以结合 C1q 并且固定补体,而 G4 则不能。尽管 G3 能够比 G1 更有效地结合 C1q, G1 在介导补体导致的细胞裂解中最有效。IgG 中的 C1q 结合位点位于 CH2 结构域的羧基末端区。

[0321] 所有的 IgG 亚类都能够结合 Fc 受体(CD16, CD32, CD64), G1 和 G3 比 G2 和 G4 更

有效。IgG 的 Fc 受体结合区由位于 CH2 结构域的铰链区和羧基末端区的残基形成。

[0322] IgA 可以以单体和由 J 链连接的二聚体形式存在。IgA 是血清中其次多的 Ig，但其半衰期只有 6 天。IgA 具有三种效应物功能。它结合巨噬细胞和嗜酸性细胞上的 IgA 特异性受体，分别驱动吞噬和脱颗粒。它也可通过未知的其它途径固定补体。

[0323] IgM 作为 5 聚体或 6 聚体表达，这两者都由 J 链连接。IgM 的血清半衰期为 5 天。它通过位于其 CH3 结构域上的结合位点与 C1q 微弱结合。IgD 的血清半衰期为 3 天。还不清楚这种 Ig 的效应物功能。IgG 是单体 Ig，血清半衰期为 2.5 天。IgE 与两个 Fc 受体结合，驱动脱颗粒并导致促炎剂的释放。

[0324] 根据需要的体内作用，本发明的异源融合蛋白可以任何上述同种型或可以含有突变的 Fc 区，其中补体和 / 或 Fc 受体结合功能改变。这样，本发明的异源融合蛋白可以含有与 GLP-1 化合物融合的免疫球蛋白的完整 Fc 部分、免疫球蛋白 Fc 部分的片段或其类似物。

[0325] 本发明的融合蛋白可以由单链蛋白组成或作为多链多肽。可以产生两个或多个 Fc 融合蛋白，使得它们通过在 Fc 区之间天然形成的二硫键而相互作用。这些多聚体对于 GLP-1 化合物来说可以是同源的，或者它们可以含有不同的 GLP-1 化合物，所述化合物融合在融合蛋白的 Fc 部分 N 末端。

[0326] 不论融合蛋白的最终结构如何，Fc 或 Fc 样区域必须延长在 N 末端融合的 GLP-1 化合物的体内血浆半衰期。此外，融合的 GLP-1 化合物必须保持一定的生物学活性。可以用实施例 7 描述的方法证明半衰期的增加，其中比较了融合蛋白和单独的 GLP-1 化合物的半衰期。可以通过本领域公知的体外和体内方法确定生物学活性。代表性的生物测定描述于实施例 6、8 和 9。

[0327] 由于通过蛋白水解产生的 IgG 的 Fc 区与完整的 IgG 分子具有相同的体内半衰期，并且 Fab 片段迅速降解，人们认为延迟半衰期的相关序列存在于 CH2 和 / 或 CH3 结构域。此外，文献中证明，不结合高亲和性 Fc 受体或 C1q 的 IgG 变体的分解代谢速度与母体野生型抗体的清除速度相似，表明分解代谢位点不同于参与 Fc 受体或 C1q 结合的位点。[Wawrzynczak et al., (1992) Molecular Immunology 29 :221]。采用鼠 IgG1Fc 区进行的定点诱变表明，控制分解代谢速度的 IgG1Fc 区的位点位于 CH2-CH3 结构域的界面。

[0328] 基于这些研究，可以在分解代谢位点修饰 Fc 区，以便优化融合蛋白的半衰期。用于本发明的异源融合蛋白的 Fc 区优选来源于 IgG1 或 IgG4 的 Fc 区。Fc 区甚至更优选是 IgG4 或来源于 IgG4。IgG Fc 区优选含有 CH2 和 CH3 区，包括铰链区。

[0329] 异源白蛋白融合蛋白：

[0330] 上述 GLP-1 化合物可以直接或通过肽接头与白蛋白或其类似物、片段或衍生物融合。

[0331] 构成本发明的融合蛋白的一部分的白蛋白一般可以来源于从任何物种克隆的白蛋白。然而，人白蛋白及其片段和类似物优选减少融合蛋白在人体内产生免疫原性的危险。人血清白蛋白 (HSA) 由单个未糖基化的 585 个氨基酸的多肽链组成，式分子量为 66,500。人 HSA 的氨基酸序列见图 2。[见 Meloun, et al. (1975) FEBS Letters 58 :136; Behrens, et al. (1975) Fed. Proc. 34 :591; Lawn, et al. (1981) Nucleic Acids Research 9 :6102-6114; Minghetti, et al. (1986) J. Biol. Chem. 261 :6747]。已经描述了白蛋白的多种多形性变体以及类似物和片段。[见 Weitkamp, et al., (1973) Ann. Hum. Genet. 37 :219]。例如，在

EP 322,094 中,发明人公开了 HSA 的各种较短形式。这些片段中的一些包括 HSA(1-373), HSA(1-388), HSA(1-389), HSA(1-369), 和 HSA(1-419) 以及 1-369 和 1-419 之间的片段。EP 399,666 公开了包括 HSA(1-177) 和 HSA(1-200) 的片段以及 HSA(1-177) 和 HSA(1-200) 之间的片段。

[0332] 可以理解,本发明的异源融合蛋白包含与任何白蛋白,包括其片段、类似物和衍生物偶联的 GLP-1 化合物,其中该融合蛋白具有生物学活性并且比单独的 GLP-1 化合物具有更长的血浆半衰期。因此,融合蛋白的白蛋白部分不必须具有等于天然人白蛋白的血浆半衰期。半衰期比天然人白蛋白和感兴趣的 GLP-1 化合物长,或半衰期介于两者之间的片段、类似物和衍生物是公知的,或可以产生。

[0333] 本发明的异源融合蛋白包括在融合蛋白的 GLP-1 化合物和 / 或白蛋白部分中具有保守氨基酸取代的蛋白。“保守取代”是指利用另一个与被取代氨基酸具有相同净电荷和具有大约相同大小和形状的氨基酸来取代此氨基酸。当它们侧链上的碳原子和杂原子总数相差不超过 4 个时,具有脂肪族或取代的脂肪族氨基酸侧链的氨基酸大致上是大小相同的。当它们侧链上的分支数目相差不超过 1 个时,它们大致具有相同的形状。侧链上具有苯基或取代苯基的氨基酸可认为是大小和形状大约相同的。除非特别指明,保守取代优选通过天然氨基酸来进行。

[0334] 然而,此处的“氨基酸”是在其最广泛的含义中使用的,包括天然存在的氨基酸和非天然存在的氨基酸,包括氨基酸类似物和衍生物。后者包括含氨基酸部分的分子。本领域技术人员将意识到,根据该广义的定义,此处的氨基酸包括,例如天然存在的形成蛋白的 L-氨基酸;D-氨基酸;化学修饰的氨基酸,例如氨基酸类似物和衍生物;天然存在的氨基酸不形成蛋白的氨基酸,例如正亮氨酸、 β -丙氨酸、鸟氨酸、GABA 等;以及具有本领域公知的氨基酸特征的化学合成的化合物。此处用到的术语“形成蛋白的”表示氨基酸可以通过代谢途径掺入细胞中的肽、多肽或蛋白。

[0335] 非天然氨基酸,包括合成的非天然氨基酸、取代的氨基酸或一个或多个 D-氨基酸插入本发明的异源融合蛋白可以通过许多途径带来益处。含有 D-氨基酸的肽等与含有 L-氨基酸的对应物相比,表现出体外或体内稳定性增加。因此,当需要或要求更大的细胞内稳定性时,肽的构建,例如掺入 D-氨基酸是特别有用的。更具体地, D-肽等抗内源性肽酶和蛋白酶活性,从而在需要时改进分子的生物利用度和延长体内寿命。此外, D-肽等不能进行有效加工以用于向 T 辅助细胞进行 I 类主要组织相容性复合物限制的呈递,因此,不容易在整个生物体内诱导体液免疫应答。

[0336] 本发明除了包括各种多肽的结构 / 功能分析,选择用于取代的氨基酸还可以考虑许多因素,进行这种改变的一个可以考虑的因素时氨基酸的亲水指数。亲水性氨基酸指数在赋予蛋白上相互作用的生物功能方面的重要性由 Kyte 和 Doolittle 进行了讨论 (1982, J. Mol. Biol., 157:105-132)。已知氨基酸的相对亲水特征是由于得到的蛋白的二级结构。这随后影响了蛋白与酶、底物、受体、配体、DNA、抗体、抗原等分子的相互作用。根据其疏水性和电荷特征,赋予每个氨基酸以下亲水指数:异亮氨酸 (+4.5);缬氨酸 (+4.2);亮氨酸 (+3.8);苯丙氨酸 (+2.8);半胱氨酸 / 胱氨酸 (+2.5);甲硫氨酸 (+1.9);丙氨酸 (+1.8);甘氨酸 (-0.4);苏氨酸 (-0.7);丝氨酸 (-0.8);色氨酸 (-0.9);酪氨酸 (-1.3);脯氨酸 (-1.6);组氨酸 (-3.2);谷氨酸 / 谷氨酰胺 / 天冬氨酸 / 天冬酰胺 (-3.5);赖氨酸 (-3.9);

和精氨酸 (-4.5)。

[0337] 如本领域所公知,可以用具有相似亲水指数或评分的其它氨基酸取代肽、多肽或蛋白中的某些氨基酸,以产生例如具有相似或甚至改进的生物活性的肽。在进行这种改变时,优选具有 ± 2 以内的亲水指数的氨基酸互相取代。更优选的取代是其中的氨基酸具有 ± 1 以内的亲水指数。最优选的取代是其中的氨基酸具有 ± 0.5 以内的亲水指数。

[0338] 氨基酸也可以根据亲水性进行取代。美国专利 4,554,101 公开了蛋白的最大局部平均亲水性,这是由其邻近的氨基酸的亲水性决定的,并且与蛋白的生物特征相关。赋予氨基酸以下亲水性值:精氨酸/赖氨酸 (+3.0);天冬氨酸/谷氨酸 (+3.01);丝氨酸 (+0.3);天冬酰胺/谷氨酰胺 (+0.2);甘氨酸 (0);苏氨酸 (-0.4);脯氨酸 (-0.51);丙氨酸/组氨酸 (-0.5);半胱氨酸 (-1.0);甲硫氨酸 (-1.3);缬氨酸 (-1.5);亮氨酸/异亮氨酸 (-1.8);酪氨酸 (-2.3);苯丙氨酸 (-2.5);和色氨酸 (-3.4)。因此,可以用具有相似亲水性评分的其它氨基酸取代肽、多肽或蛋白中的氨基酸,以产生例如具有相似的生物活性,即仍然保持正确的生物功能的肽。在进行这种改变时,优选具有 ± 2 以内的亲水指数的氨基酸互相取代。更优选 ± 1 以内,最优选 ± 0.5 以内。

[0339] 如前面所概括,本发明的融合蛋白中的氨基酸取代可以基于氨基酸侧链取代基的相对相似性,例如,它们的疏水性、亲水性、电荷、大小等。此外,可以基于二级机构特征进行取代。例如,可以用保持螺旋结构的氨基酸取代螺旋状氨基酸。例如,考虑前面的特征以产生在肽中导致沉默改变的保守氨基酸改变的示例性的取代可以选自该天然存在的氨基酸所属类中的其它成员。氨基酸可以分成以下四组:(1) 酸性氨基酸;(2) 碱性氨基酸;(3) 中性极性氨基酸;和 (4) 中性非极性氨基酸。

[0340] 制备本发明的异源融合蛋白的一般方法

[0341] 尽管可以用各种不同的方法制备本发明的异源融合蛋白,重组方法是优选的。对于此处公开和要求保护的本发明,如下定义了以下通用的分子生物学术语和缩写。除非特别指定,本文中使用的术语和缩写具有其通常的含义。例如,“ $^{\circ}\text{C}$ ”表示摄氏度;“ mmol ”表示毫摩尔;“ mg ”表示毫克;“ μg ”表示微克;“ ml 或 mL ”表示毫升;“ μl 或 μL ”表示微升。氨基酸缩写阐述于 37C.F.R. $\S$ 1.822(b)(2)(1994)。

[0342] 此处用到的“碱基对”或“ bp ”是指 DNA 或 RNA。当存在于 DNA 分子中时,缩写 A, C, G 和 T 分别对应于脱氧核糖核苷(脱氧)腺苷、(脱氧)胞苷、(脱氧)鸟苷和胸苷的 5' - 单磷酸酯形式。当存在于 RNA 分子中时,缩写 U, C, G, 和 A 分别对应于核糖核苷尿苷、胞苷、鸟苷和腺苷的 5' - 单磷酸酯形式。在双链 DNA 中,碱基对可以指 A 与 T 或 C 与 G 的配对关系。在 DNA/RNA 中,异双链体碱基对可以指 A 与 U 或 C 与 G 的配对关系。(见下文中“互补”的定义)。

[0343] DNA 的“消化”或“限制”是指用仅仅作用于 DNA 中的某些序列的限制酶(“序列特异性内切核酸酶”)对 DNA 进行催化性裂解。此处用到的各种限制酶是可买到的,并且按照本领域技术人员所公知的内容使用它们的反应条件、辅因子和其它要求。用于特定限制酶的适当缓冲液和底物量是由生产商规定的,或可以很容易在文献中找到。

[0344] “连接”是指在两个双链核酸片段之间形成磷酸二酯键的过程。除非特别指出,可以用 DNA 裂解酶,如 T4 DNA 裂解酶,采用已知的缓冲液和条件完成连接。

[0345] “质粒”是指染色体外(通常)自我复制的遗传元件。质粒一般由小写字母“p”和

其后的字母和 / 或数字表示。此处的起始质粒是购买的、可以有限地公共获得的,或者可根据公开的程序用可得到的质粒构建。此外,与描述的质粒等同的质粒是公知的,对本领域技术人员是显而易见的。

[0346] 此处用到的“重组 DNA 克隆载体”是指任何自主复制的因子,包括但不限于质粒和噬菌体,包含可以或已经添加的一个或多个额外的 DNA 片段的 DNA 分子。

[0347] 此处用到的“重组 DNA 表达载体”是指任何 DNA 克隆载体,其中掺入了控制插入的 DNA 转录的启动子。

[0348] “转录”是指一种使 DNA 的核苷酸序列中含有的信息转移到互补的 RNA 序列的过程。

[0349] “转染”是指用宿主细胞摄取表达载体,而不论实际上是否表达任何编码序列。许多转染方法是本领域技术人员公知的,例如,磷酸钙共沉淀、脂质体转染和电穿孔。成功的转染一般是当指示该载体的操作在宿主细胞发生时被识别。

[0350] “转化”是指将 DNA 导入生物体,使得 DNA 可以作为染色体外元件或通过染色体整合而复制。转化细菌和真核宿主的方法是本领域公知的,例如核注射、原生质体融合或用氯化钙进行钙处理,这些方法概括于 J. Sambrook, et al., *Molecular Cloning :A Laboratory Manual*, (1989)。一般地,当将 DNA 导入酵母时,术语转化与术语转染的含义相反。

[0351] 此处用到的“翻译”是指用信使 RNA (mRNA) 的遗传信息指定和指导多肽链合成的过程。

[0352] “载体”是指用于基因操作中的细胞转染和 / 或转化的核酸复合物,其携带对应于适当的蛋白分子的多核苷酸序列,该序列适当的控制序列结合时,则赋予被转染和 / 或转化的宿主细胞特定的属性。质粒、病毒和噬菌体是适当的载体。通过使用限制酶和连接酶从不同来源切割和连接 DNA 分子而构建人工载体。此处用到的术语“载体”包括重组 DNA 克隆载体和重组 DNA 表达载体。

[0353] 此处用到的“互补的”或“互补”是指在双链核酸中通过氢键缔合的碱基对(嘌呤和嘧啶)。以下碱基对是互补的:鸟嘌呤和胞嘧啶;腺嘌呤和胸腺嘧啶;以及腺嘌呤和尿嘧啶。

[0354] 这里用到的“杂交”是指一条核酸链通过碱基配对与互补链连接。两条不相同但非常相似的互补核酸的杂交所使用的条件根据两条链的互补程度和链长而改变。这种技术和条件是本领域技术人员公知的。

[0355] “分离的氨基酸序列”是指任意的位置上不同于天然存在的氨基酸的构建的或合成的氨基酸序列。

[0356] “分离的 DNA 复合物”是指任意的位置上不同于其在基因组中的天然位置的构建的或合成的 DNA 序列。

[0357] “分离的核酸复合物”是指任意的位置上不同于其天然位置的构建的或合成的 RNA 或 DNA 序列。

[0358] “引物”是指作为用于酶促或合成延伸的起始底物的核酸片段。

[0359] “启动子”是指直到从 DNA 转录为 RNA 的 DNA 序列。

[0360] “探针”是指与另一种核酸复合物杂交的核酸复合物或其片段。

[0361] 杂交反应的“严格性”可以由本领域技术人员容易地判断,并且一般是根据探针长度、洗涤温度和盐浓度而进行的经验计算值。一般地,较长的探针需要较高的温度进行适当退火,而短探针需要较低的温度。杂交一般取决于互补链存在于低于其解链温度的环境中时变性的 DNA 重新退火的能力。探针和可杂交序列间需要的同源性程度越高,可以使用的相对温度越高。因此,越高的相对温度将使反应更严格,而较低的温度使严格性更低。对于杂交反应严格性的其它细节和解释,参见 Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, 1995。

[0362] 此处定义的“严格条件”或“高严格性条件”可以通过以下条件鉴定:(1) 使用低离子强度和高温进行洗涤,例如 15mM 氯化钠 /1.5mM 柠檬酸钠 /0.1% 十二烷基硫酸钠, 50°C ;(2) 在杂交中使用变性剂,例如甲酰胺,例如 50% (v/v) 甲酰胺,0.1% 牛血清白蛋白 /0.1% ficoll/0.1% 聚乙烯吡咯烷酮 /50mM 磷酸钠缓冲液, pH6.5, 750mM 氯化钠 /75mM 柠檬酸钠, 42°C ;或 (3) 使用 50% 甲酰胺, 5×SSC (750mM 氯化钠, 75mM 柠檬酸钠), 50mM 磷酸钠 (pH 6.8), 0.1% 焦磷酸钠, 5×Denhardt' s 溶液, 超声波处理的鲑鱼精子 DNA (50 μg/ml), 0.1% SDS, 和 10% 硫酸右旋糖苷, 42°C, 在 42°C, 0.2×SSC (30mM 氯化钠 /3mM 柠檬酸钠), 和 55°C, 50% 甲酰胺中洗涤,然后在 55°C 用含 EDTA 的 0.1×SSC 进行高严格性洗涤。

[0363] “中等严格条件”可以通过 ambrook et al. [*Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Press, (1989)] 中的描述进行鉴定,并且包括使用比上述严格性低的洗涤溶液和杂交条件(例如温度、离子强度和 % SDS)。中等严格条件的一个例子是在 37°C, 在含有 20% 甲酰胺、5×SSC (750mM 氯化钠, 75mM 柠檬酸钠), 50mM pH7.6 的磷酸钠, 5×Denhardt' s 溶液, 10% 硫酸右旋糖苷和 20mg/mL 变性的剪切的鲑鱼精子 DNA 的溶液中孵育过夜,然后在 37–50°C, 在 1×SSC 中洗涤滤器。本领域技术人员将理解如何调节温度、离子强度等,以及必要时调节探针长度等因子。

[0364] “PCR”是指公知的使用热稳定性 DNA 聚合酶的聚合酶链式反应。

[0365] “前导序列”是指可以酶促或化学去除以产生需要的感兴趣多肽的氨基酸序列。

[0366] “分泌信号序列”是指一般存在于较大的多肽 N 末端区的氨基酸序列,其功能是起始该多肽与内质网等细胞膜腔室的缔合和该多肽通过浆膜的分泌。

[0367] 构建编码本发明的异源融合蛋白的 DNA:

[0368] 可以从各种来源获得野生型白蛋白和免疫球蛋白。例如,这些蛋白可以从由以可检测水平表达感兴趣的 mRNA 的组织或细胞制备的 cDNA 文库获得。可以用特定的感兴趣蛋白的公开 DNA 或蛋白序列设计探针,对文库进行筛选。例如,免疫球蛋白轻链或重链恒定区描述于 Adams, et al. (1980) *Biochemistry* 19:2711–2719; Goughet, et al. (1980) *Biochemistry* 19:2702–2710; Dolby, et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:6027–6031; Rice et al. (1982) *Proc. Natl. Acad. Sci. USA* 79:7862–7862; Falkner, et al. (1982) *Nature* 298:286–288; 和 Morrison, et al. (1984) *Ann. Rev. Immunol.* 2:239–256。一些参考文献公开了白蛋白的蛋白和 DNA 序列,包括 Meloun, et al. (1975) *FEBS Letters* 58:136; Behrens, et al. (1975) *Fed. Proc.* 34:591; Lawn, et al. (1981) *Nucleic Acids Research* 9:6102–6114; 和 Minghetti, et al. (1986) *J. Biol. Chem.* 261:6747。

[0369] 可以采用标准程序,用选择的探针筛选 cDNA 或基因组文库,例如描述于 Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory

Press, NY (1989) 中的程序。分离编码白蛋白或免疫球蛋白的基因的另一种方法是采用 PCR 方法 [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY (1995)]。可以根据公开的序列设计 PCR 引物。

[0370] 从特定物种克隆的全长野生型序列一般可以作为模板, 用于产生保留赋予作为融合蛋白的一部分的 GLP-1 化合物更长的血浆半衰期的能力的类似物、片段和衍生物。本发明的异源融合蛋白的 Fc 和白蛋白部分优选来源于天然人序列, 以减少融合蛋白在人体内潜在的免疫原性危险。

[0371] 具体地, 本发明包括的融合蛋白的免疫球蛋白部分优选仅包含免疫球蛋白的 Fc 片段。根据融合蛋白的所需特定效应物功能以及结构特征, Fc 片段可以含有铰链区以及 CH2 和 CH3 结构域或一些其它组合。可以使用设计为与相应于片段的所需末端的序列杂交的引物, 采用 PCR 技术产生这些 Fc 片段。

[0372] 可以通过各种不同的方法, 包括上述克隆方法和化学合成 DNA 的方法制备编码本发明的 GLP-1 化合物的 DNA。如果所编码的肽较短, 则化学合成更好。GLP-1 的氨基酸序列和前胰高血糖素原基因已经公开 [Lopez, et al. (1983) Proc. Natl. Acad. Sci., USA 80:5485-5489; Bell, et al. (1983) Nature, 302:716-718; Heinrich, G., et al. (1984) Endocrinol, 115:2176-2181; Ghiglierone, M., et al. (1984) Diabetologia 27:599-600]。因此, 可以设计引物以便对天然 GLP-1 化合物及其片段进行 PCR。

[0373] 然后通过框内连接编码 GLP-1 化合物的 DNA 与编码白蛋白或 Fc 部分的 DNA 而构建编码融合蛋白的基因。也可以通过编码接头肽的基因框内连接编码 GLP-1 化合物的基因与编码白蛋白或 Fc 部分的基因。

[0374] 可以通过添加小肽接头以防止潜在的不希望的结构域相互作用而对本发明的异源融合蛋白的体内功能和稳定性进行优化。尽管这些接头可以是任意长度并且由任意氨基酸组合组成, 其长度优选不超过防止不希望的结构域相互作用和 / 或优化生物功能和 / 或稳定性所必须的长度。一般地, 接头不应该含有具有极大侧链的氨基酸或容易诱导显著的二级结构的氨基酸。接头优选富含丝氨酸 - 甘氨酸, 并且其长度优选不超过 30 个氨基酸。接头的长度更优选不超过 20 个氨基酸。接头的长度甚至更优选不超过 15 个氨基酸。优选的接头含有序列 Gly-Gly-Gly-Gly-Ser 的重复。优选有 2-6 个该序列的重复。甚至更优选有 3-4 个该序列的重复。

[0375] 编码野生型 GLP-1、白蛋白及其 Fc 多肽和片段的 DNA 可以在连接之前或在编码整个融合蛋白的 cDNA 的范围内进行突变。多种诱变技术是本领域公知的。例如, 利用链重叠延伸以制备用于改变相应蛋白中的特定氨基酸序列的特定碱基突变的诱变 PCR 方法。该 PCR 诱变要求使用四种引物, 两个前向 (引物 A 和 C) 和两个反向 (引物 B 和 D)。在两个不同阶段中从野生型模板扩增突变的基因。第一个反应通过进行 A 到 B 反应和独立的 C 到 D 反应将基因分成两半进行扩增, 其中 B 和 C 引物定向于待突变的基因区域。当将这些引物与靶区域比对时, 它们含有被定向进行改变的碱基的错配。一旦完成了 A 到 B 和 C 到 D 反应, 分离反应产物并且进行混合, 用作 A 到 D 反应的模板。该反应然后产生完整的突变产物。

[0376] 一旦产生了编码完整融合蛋白的基因, 它可以被克隆到适当的表达载体。可以用于制备本发明的 GLP-1 融合蛋白的特定策略描述于实施例 1。

[0377] 用于重组表达本发明的异源融合蛋白的一般方法

[0378] 用此处描述的表达或克隆载体转染或转化宿主细胞,用于产生异源融合蛋白,并且在适于诱导启动子的改良的常规营养培养基中培养,选择转化体或扩增编码所需序列的基因。本领域技术人员不需要过多实验就可以选择培养条件,如培养基、温度、pH等。概言之,用于增加细胞培养物产生的原则、方案和实用技术可以参见 *Mammalian Cell Biotechnology: A Practical Approach*, M. Butler, ed. (IRL Press, 1991) 和 Sambrook, et al., *supra*。转染方法是本领域技术人员公知的,例如 CaPO_4 和电穿孔。宿主系统转化的一般方面描述于美国专利 4,399,216。一般根据 van Solingen et al., *J. Bact.* 130(2): 946-7(1977) 和 Hsiao et al., *Proc. Natl. Acad. Sci. USA* 76(8): 3829-33(1979) 的方法转化至酵母。然而,也可以使用其它用于将 DNA 导入细胞的方法,例如核微量注射、电穿孔、与完整细胞进行细菌原生质体融合或多阳离子,如 polybrene 或 polyomithine。对于转化哺乳动物细胞的各种技术,参见 Keown, et al., *Methods in Enzymology* 185: 527-37(1990) 和 Mansour, et al., *Nature* 336(6197): 348-52(1988)。

[0379] 用于在此处的载体中克隆或表达核酸(如 DNA)的适当宿主细胞包括原核细胞、酵母或高等真核细胞。适当的原核细胞包括,但不限于真细菌,如革兰氏阴性或革兰氏阳性生物,例如肠细菌,如大肠杆菌。可以公共得到各种大肠杆菌株,例如大肠杆菌 K12 株 MM294(ATCC 31.446);大肠杆菌 X1 776(ATCC 31.537);大肠杆菌株 W3110(ATCC 27.325)和 K5 772(ATCC 53.635)。其它适当的真核宿主细胞包括肠细菌,如埃希氏菌,如大肠杆菌,肠杆菌,欧文氏菌,克雷白氏菌,变形菌,沙门氏菌如伤寒沙门氏菌,沙雷氏菌如粘质沙雷氏菌,和志贺氏菌,以及芽孢杆菌如枯草芽孢杆菌和地衣芽孢杆菌(如公开于 1989 年 4 月 12 日发表的 DD266, 710 的地衣芽孢杆菌 41P),假单胞菌如铜绿假单胞菌和链霉菌。这些例子是说明性的,而不是限制性的。菌株 W3110 是特别优选的宿主或亲代宿主,因为它是用于重组 DNA 产物发酵的常用宿主株。优选地,宿主细胞分泌最小量的蛋白水解酶。例如,可以对菌株 W3110 进行修饰以在编码宿主的内源蛋白的基因中实现基因突变,所述宿主的例子包括具有完整的基因型 *ronA* 的大肠杆菌 W3110 菌株 1A2;具有完整的基因型 *tonA ptr3* 的大肠杆菌 W3110 菌株 9E4;具有完整的基因型 *tonA ptr3 phoA E15(argF-lac) I69 degPompT/can'* 的大肠杆菌 W3110 菌株 27C7(ATCC 55,244);大肠杆菌 W3110 菌株 40B4,它是具有非卡那霉素抗性 *degP* 缺失突变的菌株 37D6;以及具有 1990 年 8 月 7 日授权的美国专利 4,946,783 公开的突变周质蛋白酶的大肠杆菌株。或者,PCR 或其它核酸聚合酶反应等体内克隆方法是适用的。

[0380] 除原核生物外,真核微生物如丝状真菌或酵母是融合蛋白载体的适当克隆或表达宿主。啤酒糖酵母是常用的低等真核宿主微生物。其它包括粟酒裂殖糖酵母 [Beach and Nurse, *Nature* 290: 140-3(1981); EP 139,383 published 2 May 1995]; *Muyveromyces* 宿主 [美国专利 4,943,529; Fleer, et al., *Biol. Technology* 9(10): 968-75(1991)], 如克鲁维氏酵母 (MW98-8C, CBS683, CBS 4574) [de Louvencourt et al., *J. Bacteriol.* 154(2): 737-42(1983)]; 脆壁克鲁维氏酵母 (ATCC 12,424), 保加利亚克鲁维氏酵母 (ATCC 16,045), 威克曼氏克鲁维氏酵母 (ATCC 24,178), *waltii* 克鲁维氏酵母 (ATCC 56,500), 果蝇克鲁维氏酵母 (ATCC 36.906) [Van den Berg et al., *Biol. Technology* 8(2): 135-9(1990)]; *thermophilus* 克鲁维氏酵母和马克斯克鲁维氏酵母; *yarrowia* (EP 402,226); 巴斯德毕赤氏酵母 (EP 183,070) [Sreekrishna et

al., J. Basic Microbiol. 28(4) :265-78(1988)] ; 假丝酵母; *Trichoderma reesia* (EP 244, 234) ; 粗糙链孢霉 [Case, et al., Proc. Natl. Acad. Sci. USA 76(10) :5259-63(1979)] ; 许旺氏酵母如西方许旺氏酵母 (1990 年 10 月 31 日公开的 EP 394, 538) 和丝状真菌如链孢霉、青霉、*Tolypocladium* (1991 年 1 月 10 日公开的 WO 91/00357) ; 和曲霉宿主如构巢曲霉 [Ballance et al., Biochem. Biophys. Res. Comm. 112(1) :284-9(1983)] ; Tilburn, et al., Gene 26(2-3) :205-21(1983) ; Yelton, et al., Proc. Natl. Acad. Sci. USA 81(5) :1470-4(1984)] 和黑曲霉 [Kelly and Hynes, EMBJ. 4(2) :475-9(1985)]。甲基营养酵母选自以下属: 汉逊氏酵母属、假丝酵母属、克氏酵母属、毕赤氏酵母属、啤酒糖酵母属、球拟酵母属和红酵母属。这种类型的酵母的示例性具体物种的列表可以见 C. Antony, The Biochemistry of Methylotrophs 269(1982)。

[0381] 用于表达本发明的融合蛋白的适当宿主细胞来源于多细胞生物。无脊椎细胞的例子包括昆虫细胞如果蝇 S2 和夜蛾, 夜蛾 high5 和植物细胞。有用的哺乳动物宿主细胞系包括中国仓鼠卵巢 (CHO) 和 COS 细胞。更具体的例子包括 SV40 转化的猴肾 CV1 细胞系 (COS-7, ATCC CRL1651) ; 人胚肾细胞系 [亚克隆的用于在悬浮培养基中生长的 293 或 293 细胞, Graham, et al., J. Gen. Tirol., 36(1) :59-74(1977)] ; 中国仓鼠卵巢细胞 /-DHFR [CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77(7) :4216-20(1980)] ; 小鼠支持细胞 [TM4, Mather, Biol. Reprod. 23(1) :243-52(1980)] ; 人肺细胞 (W138, ATCC CCL75) ; 人肝细胞 (Hep G2, HB 8065) 和小鼠乳腺肿瘤 (MMT 060562, ATCC CCL51)。适当宿主细胞的选择被认为在本领域的技术范围内。

[0382] 本发明的融合蛋白可以直接重组产生, 或作为具有信号序列或其它序列的蛋白产生, 所述序列在成熟融合蛋白的 N 末端产生特异性裂解位点。概言之, 信号序列可以是载体的一种成分, 或者它可以是插入载体的融合蛋白的编码 DNA 的一部分。信号序列可以是例如碱性磷酸酶、青霉素酶 1pp, 或耐热肠毒素 II 前导序列中选择原核信号序列。对于酵母分泌, 信号序列可以是, 例如, 酵母转化酶前导序列, α 因子前导序列 (包括啤酒糖酵母和克鲁维氏酵母 α 因子前导序列, 后者描述于美国专利 5, 010, 182), 或酸性磷酸酶前导序列, *C. albicans* 葡糖淀粉酶前导序列 (EP 362, 179) 或描述于 WO 90/13646 中的信号。在哺乳动物细胞表达中, 可以用哺乳动物信号序列指导蛋白分泌, 例如来自于相同或相关物种的分泌多肽的信号序列和病毒的分泌性前导序列。

[0383] 表达和克隆载体都含有使载体能够在一个或多个选择的宿主细胞中复制的核酸序列。这种序列在多种细菌、酵母和病毒中都是公知的。来自于质粒 pBR322 的复制起点适于大多数革兰氏阴性细菌, 2 μ 质粒起点适于酵母, 各种病毒 (SV40, 多瘤病毒, 腺病毒, VSV 或 BPV) 起点适用于哺乳动物细胞中的克隆载体。

[0384] 表达和克隆载体一般含有选择基因, 也称作可选择标记。典型的选择基因编码具有以下特征的蛋白: (a) 赋予抗生素或其它毒素, 如氨苄青霉素、新霉素、氨甲喋呤或四环素抗性, (b) 补充自养的营养缺陷, 或 (c) 补充复合培养基所不能得到的关键营养, 如编码芽孢杆菌的 D- 丙氨酸消旋酶的基因。

[0385] 用于哺乳动物细胞的适当可选择标记的例子是使得能够鉴定可以摄取融合蛋白的编码核酸的细胞的标记, 如 DHFR 或胸苷激酶。使用野生型 DHFR 时的适当宿主细胞是缺乏 DHFR 活性的 CHO 细胞系, 其制备和增殖方法如 [Urlaub and Chasin, Proc. Natl. Acad. Sci.

USA, 77(7):4216-20(1980)] 的描述。用于酵母的适当选择基因是存在于酵母质粒 YRp7 中的 *trp1* 基因 [Stinchcomb, et al., Nature 282(5734):39-43(1979);Kingsman, et al., Gene 7(2):141-52(1979);Tschumper, et al., Gene 10(2):157-66(1980)]。*trp1* 基因提供了用于缺乏在色氨酸中生长的能力的突变酵母株,例如 ATCC No. 44076 或 PEPC1 [Jones, Genetics 85:23-33(1977)] 的选择标记。

[0386] 表达和克隆载体通常含有与编码融合蛋白的核酸序列可操作性连接的启动子,以指导 mRNA 合成。通过各种可能的宿主细胞识别启动子是公知的。适用于原核宿主的启动子包括内酰胺酶和乳糖启动子系统 [Chang, et al., Nature 275(5681):617-24(1978);Goeddel, et al., Nature 281(5732):544-8(1979)],碱性磷酸酶,色氨酸(上游)启动子系统 [Goeddel, Nucleic Acids Res. 8(18):4057-74(1980);1981 年 9 月 30 日公开的 EP 36,776],以及杂交启动子如 *tat* 启动子 [deBoer, et al., Proc. Natl. Acad. Sci. USA 80(1):21-5(1983)]。用于细菌系统的启动子也将包含与编码融合蛋白的 DNA 可操作性连接的 Shine-Dalgarno (S. D.) 序列。

[0387] 用于酵母宿主的适当启动子序列的例子包括 3-磷酸甘油酯激酶的启动子 [Hitzeman, et al., J. Biol. Chem. 255(24):12073-80(1980)] 或其它糖裂解酶 [Hess et al., J. Adv. Enzyme Reg. 7:149(1968);Holland, Biochemistry 17(23):49007(1978)],例如烯醇酶、甘油醛-3-磷酸脱氢酶、己糖激酶、丙酮酸脱羧酶、磷酸果糖激酶、葡萄糖-6-磷酸异构酶、3-磷酸甘油酯变位酶、丙酮酸激酶、丙糖磷酸异构酶、磷酸葡萄糖异构酶和葡萄糖激酶。

[0388] 其它具有由生长条件控制的额外转录优点的诱导型酵母启动子是醇脱氢酶 2、异细胞色素 C、酸磷酸酶、与氮代谢相关的降解酶、金属硫蛋白酶、甘油醛-3-磷酸脱氢酶和负责麦芽糖和半乳糖利用的酶的启动子区。用于酵母表达的适当载体和启动子进一步描述于 EP73,657。可以通过从多瘤病毒、禽痘病毒、腺病毒(例如腺病毒 2)、牛乳头瘤病毒、鸟类肉瘤病毒、巨细胞病毒、逆转录病毒、乙型肝炎病毒和猿猴病毒 40(SV40)等病毒基因组获得的启动子、异源哺乳动物启动子如肌动蛋白启动子或免疫球蛋白启动子和热激启动子控制从哺乳动物宿主细胞中的载体转录编码融合蛋白的 mRNA,前提是这些启动子与宿主细胞系统相容。

[0389] 可以通过在载体中插入增强子序列增加高等真核细胞转录编码融合蛋白的多核苷酸。增强子是 DNA 的顺式作用元件,通常约 10-300bp,作用于启动子以增加其转录。已知许多来自于哺乳动物基因(球蛋白、弹性蛋白酶、白蛋白、 α -酮蛋白和胰岛素)的增强子序列。然而,一般地,将使用来自于真核细胞病毒的增强子。其例子包括复制起点近侧的 SV40 增强子(bp 100-270),巨细胞病毒早期启动子增强子,复制起点近侧的多瘤病毒增强子和腺病毒增强子。增强子可以剪切为位于融合蛋白编码序列 5' 或 3' 的载体中,但优选位于启动子 5' 方向的位点。

[0390] 用于真核宿主细胞(酵母、真菌、昆虫、植物、动物、人或其它多细胞生物的有核细胞)的表达载体也将含有转录终止和稳定 mRNA 所必须的序列。这些序列通常可以从真核细胞或病毒 DNAs 或 cDNAs 的 5' 区和偶然从其 3' 非翻译区获得。这些区在编码融合蛋白的 mRNA 的非翻译部分中含有转录为聚腺苷酸化片段的核苷酸片段。

[0391] 可以从培养基或宿主细胞裂解物中回收各种形式的融合蛋白。如果是膜包围的,

可以使用适当的去污剂溶液（例如 Triton-X 100）或通过酶裂解而从膜中释放。在融合蛋白的表达中使用的细胞可以通过各种物理或化学方法破坏，例如冻融循环、超声处理、机械破坏或细胞裂解剂。

[0392] 纯化本发明的异源融合蛋白：

[0393] 一旦本发明的异源融合蛋白在适当的宿主细胞中表达，可以分离和纯化类似物。以下程序是示例性的适当纯化程序：在羧甲基纤维素上分级；凝胶过滤如 Sephadex G-75；阴离子交换树脂，如 DEAE 或 Mono-Q；阳离子交换树脂如 CM 或 Mono-S；去除 IgG 等污染物的蛋白 A 琼脂糖；用金属螯合柱结合表位标记形式的多肽；反相 HPLC；层析聚焦；硅胶；乙醇沉淀；和硫酸铵沉淀。

[0394] 可以使用蛋白纯化的各种方法，并且这些方法是本领域公知的，并且描述于例如，Deutscher, Methods in Enzymology 182:83-9(1990) 和 Scopes, Protein Purification: Principles and Practice, Springer-Verlag, NY(1982)。所选择的纯化步骤将取决于使用的生产过程的性质和产生的特定融合蛋白。例如，可以用蛋白 A 或蛋白 G 亲和基质有效纯化包含 Fc 片段的融合蛋白。可以使用低或高 pH 缓冲液从亲和基质上洗脱融合蛋白。温和的洗脱条件将有助于防止融合蛋白的不可逆变性。也可以使用含咪唑的缓冲液。实施例 3 描述了用于本发明的融合蛋白的一些成功的纯化方案。

[0395] 本发明的融合蛋白的表征：

[0396] 可以用许多方法表征本发明的融合蛋白。这些方法中的一些包括：与蛋白染色方法偶联的 SDS-PAGE 或使用抗 IgG 或抗 HSA 抗体的免疫印迹。其它方法包括基质辅助的激光解吸 / 电离质谱 (MALDI-MS)、液相色谱 / 质谱、等电聚焦、分析性阴离子交换、层析聚焦和圆二色散。一些代表性的异源融合蛋白是用免疫印迹偶联的 SDS-PAGE 以及质谱表征的（见实施例 4 和 5 以及图 3 和 4）。

[0397] 例如表 3（见实施例 5）说明了代表性的融合蛋白的计算分子量和由质谱仪测定的质量。此外，图 3 和 4 说明了由 SDS PAGE 测定的代表性的融合蛋白的分子量。所有检测的异源融合蛋白都是瞬时表达并且分泌的。此外，切除了 Ig κ 信号序列以产生具有正确 N 端的蛋白。

[0398] 此外，表 3 说明，在一些情况下，由质谱测定的质量高于所预期的。这是因为 Fc 部分的糖基化和 C 末端延伸片段。用酶消化融合蛋白，然后通过反相 HPLC 和质谱可以表征含有糖部分的肽级分。可以对这些级分进行 N 末端氨基酸测序，以鉴定潜在的糖基化位点。例如，Exendin-4-Fc (SEQ ID NO:29) 的表征表明，第 39 位的丝氨酸和第 50 位的苏氨酸是 O- 连接糖基化的，第 122 位的天冬酰胺是 N- 连接糖基化的。

[0399] 也检测了代表性的 GLP-1 融合蛋白的活性，可以用许多方法体外和体内检测 GLP-1 活性（见，例如实施例 6, 7, 8 和 9）。表 4（实施例 6）说明，GLP-1 受体活性与一些 GLP-1 融合相关。其数目是相对于与 Val⁸-GLP-1 (7-37)OH 相关的活性。所有检测的融合蛋白具有 GLP-1 受体活性。低水平的体外活性不一定表示体外作用弱。由于这些融合蛋白半衰期的很大增加，体外活性弱一般不是体内活性弱的指标。图 7 和实施例说明，延长的半衰期与本发明的融合蛋白相关。例如，Val⁸-GLP-1-Fc 在猴中的半衰期为大约 45 小时，Val⁸-GLP-1-HSA 在猴中的半衰期为大约 87 小时，Gly⁸-Glu²²-GLP-1-CEX- 接头-IgG1 在狗中静脉给药后的半衰期为大约 55 小时，Gly⁸-Glu²²-GLP-1-CEX- 接头-IgG1 在狗中皮下给

药后的半衰期为大约 38 小时。

[0400] 本发明的组合物：

[0401] 物理稳定性也是治疗性蛋白制剂的关键特征。GLP-1 化合物特别难以制造和制剂化,因为在加工中存在结构改变。例如,一些 GLP-1 化合物具有聚集的倾向。此外,已经证明一些 GLP-1 化合物从可溶的和活性 α 螺旋形式转化为不溶的和潜在无活性的 β 折叠形式。GLP-1 化合物与大的蛋白如 IgG 的 Fc 区或白蛋白的融合不仅增加 GLP-1 化合物的半衰期,也有助于 GLP-1 化合物的物理和构象稳定性。例如,Val⁸-GLP-1-接头-HSA 在 PBS 中可以在 37°C 稳定约 30 天。

[0402] 本发明的异源融合蛋白可以与一种或多种赋形剂进行制剂化。本发明的活性融合蛋白可以与可药用缓冲液组合,并且调节 pH 以提供可接受的稳定性和适于肠胃外给药等给药的 pH。

[0403] 可以添加一种或多种可药用的抗微生物剂。间甲酚和苯酚是优选的可药用抗微生物剂。可以加入一种或多种可药用盐以调节离子强度或渗透压。可以加入一种或多种赋形剂以进一步调节制剂的等渗性。甘油是调节等渗性的赋形剂的一个例子。适于人或其它动物给药的可药用手段不含毒性元素或不需要的污染物,并且不干扰其活性化合物的活性。

[0404] 本发明的异源融合蛋白的可药用盐形式可以用于本发明。通常用于形成酸加成盐的酸是无机酸,如盐酸、氢溴酸、氢碘酸、硫酸、磷酸等,和有机酸如对甲苯磺酸、甲烷磺酸、草酸、对溴苯磺酸、碳酸、琥珀酸、柠檬酸、苯甲酸、乙酸等。优选的酸加成盐是与无机酸如盐酸和氢溴酸形成的盐。

[0405] 碱加成盐包括来源于无机碱,如铵或碱或碱土金属羟化物的盐、碳酸盐、碳酸氢盐等。因此用于制备本发明的盐的这些碱包括氢氧化钠、氢氧化钾、氢氧化铵、碳酸钾等。

[0406] 组合物的给药：

[0407] 可以通过医生和普通技术人员公知的有效途径进行给药。外周肠胃外给药是这样的一种方法。医学文献中通常将肠胃外给药理解为用注射器或一些其它的机械设备如输注泵将剂型注射到体内。外周肠胃外给药可以包括静脉内、肌内、皮下和腹膜内给药途径。

[0408] 本发明的异源融合蛋白也可以用于通过口服、直肠、鼻内、下呼吸道途径给药,这些都是非肠胃外途径。在这些非肠胃外途径中,优选呼吸道途径和口服途径。

[0409] 本发明的融合蛋白可用于治疗多种疾病和状况。本发明的融合蛋白主要通过作用于称为“GLP-1 受体”的受体来发挥它们的生物学作用。因此,可利用本发明中的 GLP-1 化合物来治疗患有可对 GLP-1 受体刺激和 / 或对 GLP-1 化合物给药作出有益应答的疾病或状况的患者。这些患者被称作“需要利用 GLP-1 化合物进行治疗”或“需要 GLP-1 受体刺激”。这样的患者包括患有胰岛素非依赖型糖尿病、胰岛素依赖型糖尿病、卒中 (见 WO 00/16979)、心肌梗塞 (见 WO 98/08531)、肥胖症 (见 WO 98/19698)、手术后分解代谢改变 (见美国专利 . NO 6,006,753)、功能性消化不良和肠激惹综合征 (见 WO 99/64060) 的患者。也包括需要利用 GLP-1 化合物进行预防性处理的患者,如具有发生胰岛素非依赖型糖尿病危险性的患者 (见 WO 00/07617)。葡萄糖耐受异常或空腹葡萄糖异常的受试者、体重超出相对于患者身高和体格来讲的正常体重 25% 的受试者、进行了部分胰切除手术的受试者、双亲中有一或多个患有胰岛素非依赖型糖尿病的受试者、患有妊娠期糖尿病的受试者及患有急或慢性胰腺炎的受试者具有患胰岛素非依赖型糖尿病的危险性。

[0410] GLP-1 化合物的“有效量”是指对需要 GLP-1 受体刺激的患者给药时,在不造成不需要的副作用的条件下,能够产生所需治疗和 / 或预防效果的量。“所需的治疗效果”包括下列的一或多个方面:1) 与疾病和状况相关的症状的改善;2) 延缓与疾病或状况相关的症状的出现;3) 与没有进行治疗的情况相比寿命增加;以及4) 相对于没有进行治疗的情况具有更高的生活质量。例如,用于治疗糖尿病的 GLP-1 化合物的“有效量”指与没有进行治疗的情况相比,能对血糖浓度进行更大程度上的控制,从而延缓糖尿病并发症如视网膜病变、神经病或肾病出现的量。用于预防糖尿病的 GLP-1 化合物的“有效量”指与没有进行治疗的情况相比,能延缓升高的血糖水平出现的量,而血糖水平升高是需要用抗高血糖药物如磺脲类、噻唑烷二酮、胰岛素和 / 或氯苯呱嗪来进行治疗的。

[0411] 有效使患者的血糖正常化的融合蛋白的剂量将取决于许多因素,其中包括但不限于受试者的性别、体重和年龄、不能调节血糖的严重性、给药途径和生物利用度、融合蛋白的药代动力学曲线、效力和制剂。

[0412] 本发明包括具有通过与白蛋白、白蛋白片段、白蛋白类似物 Fc 部分、Fc 片段或 Fc 类似物融合而具有改进的生化和生物物理特征的 GLP-1 化合物。这些异源蛋白可以在宿主细胞中成功表达,保留与 GLP-1 受体的激活相关的信号传递活性,并且具有延长的半衰期。

[0413] 提供了以下实施例,以便进一步描述本发明。本发明的范围不是仅仅由以下实施例组成。本领域技术人员将理解,此处描述的特定试剂、设备和程序仅仅是说明性的,并没有以任何方式限制本发明。

[0414] 实施例 1:构建编码异源融合蛋白的 DNA

[0415] 实施例 1a 构建编码 Val⁸-GLP-1 (7-37)-Fc 的 DNA:

[0416] 从 cDNA 文库分离人 IgG1 的 Fc 部分,并且含有完整的铰链区和 CH2 和 CH3 结构域。将含有人 IgG1 的该 Fc 部分的 696 个碱基对的片段亚克隆到哺乳动物表达载体 pJB02 的 NheI 和 Eco47III 位点,以产生 pJB02/Fc (见图 5)。通过 4 个重叠和互补的寡核苷酸的体外杂交产生编码与 Val⁸-GLP-1 (7-37) 融合的 Ig κ 分泌信号的 DNA:

[0417] 5' -CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCAC TGGTGACCAAGT-3' [SEQ ID NO:12]

[0418] 5' -GAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTGGAGGGCCAGGCCGCCAAGGAGTTCATCGC CTGGCTGGTGAAGGGAAGAGGC-3' [SEQ ID NO:13]

[0419] 5' -TGAAGGTGCCCTCCACGTGGTCACCAAGTGAACCTGGAACCCAGAGCAGCAGTACCCATAGCAGG AGTGTGTCTGTCTCCATGGTGG-3' [SEQ ID NO:14]

[0420] 5' -GCCTCTTCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGCCTGGCCCTCCAGATAGGAGGACA CGTCGGAGG-3' [SEQ ID NO:15]

[0421] 采用等量的每种寡核苷酸 (每种寡核苷酸的终浓度为 1 μ M) 进行杂交反应。将寡核苷酸的化合物在连接缓冲液中 100°C 下加热 5 分钟 (50mM Tris-HCl, pH 7.5, 10mM MgCl₂, 10mM DTT, 1mM ATP, 25 μ g/ml 牛血清白蛋白),然后经至少 2 小时冷却到 30°C。

[0422] 得到的杂交产物与用 NheI 和 Eco47III 消化的 pJB02/Fc 载体主链在室温下连接 2 小时或在 16°C 连接过夜。用连接产物转化感受态的 XL-1 Blue 细胞 (Stratagene)。通过用 NcoI 消化克隆筛选重组质粒中编码肽的插入序列的存在 (编码 Kozak 序列和信号肽的第一个 Met) 并且测序。得到的用于转染测定的表达质粒称作 pJB02-V8-GLP-1-Fc (图 5)。

[0423] 实施例 1b 构建编码 Val⁸-GLP-1 (7-37)-HSA 的 DNA

[0424] 质粒 HSA/pcDNA3.1GS 购自 Invitrogen (目录号 H-M12523M-pcDNA3.1/GS) 并且作为模板分离编码人血清白蛋白 (HSA) 的 cDNA。用 PCR 制备 HSA cDNA, 其中从 5' 末端去除编码前导序列和 6 个氨基酸的前肽的 DNA。此外, 在 HSA 编码序列的 3' 末端直接添加终止密码子。最后, 将限制酶位点工程化在 5' 和 3' 末端以促进克隆。与天然人序列相比, 存在于购自 Invitrogen 的原始载体中的 HSA DNA 序列在基因的 3' 区 (第 667 位) 含有单碱基改变。该改变将产生 Asn, 而不是 Asp 的密码子。因此, 使用上述链重叠 PCR 诱变方法, 将密码子改变为在此位置编码 Asp。将得到的 HSA 编码 DNA 克隆到 pJB02 的 NheI 和 HindIII 位点以产生 pJB02-HSA (图 6)。

[0425] 按实施例 1a 的讨论产生与 Val⁸-GLP-1 (7-37) 序列融合的 Igκ 前导序列。将该 DNA 连接到 pJB02-HSA 的 NheI 和 FspI 位点以产生 pJB02-Val⁸-GLP-1-HSA。

[0426] 实施例 1c 构建编码 Val⁸-GLP-1 (7-37)-接头-HSA 的 DNA :

[0427] 按实施例 1b 的描述制备载体 pJB02-HSA。将编码接头序列 [GGGGS]₃ 的 DNA 框内连接于 HAS 的编码 DNA 的 5' 末端以产生 pJB02-接头-HSA (图 7), 按实施例 1a 的描述制备编码融合于 Val⁸-GLP-1 (7-37) 序列和接头序列的 5' 部分的 IgGκ 前导序列的 DNA。将该 DNA 连接于 pJB02 的 NheI 和 BspEI 位点以产生 pJB02-Val⁸-GLP-1-接头-HSA。

[0428] 实施例 1d 构建编码 Exendin-4-Fc 的 DNA :

[0429] 按实施例 1a 的描述制备质粒 pJB02/Fc。通过以下重叠和互补寡核苷酸的体外杂交产生编码融合于 Exendin-4 的 Igκ 信号序列的 DNA :

[0430] 5' -CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCACCGGTCAC-3' [SEQ ID NO :16]

[0431] 5' -GGAGAGGGAACCTTCACCAGCGACCTGAGCAAGCAGATGGAGGAGGAGGCCGTGAGACTG-3' [SEQ ID NO :17]

[0432] 5' -TTCATCGAGTGGCTGAAGAACGGAGGACCAAGCAGCGGAGCCCCCTCCTCCTAGC-3' [SEQ ID NO :18]

[0433] 5' -GAACCTGGAACCCAGAGCAGCAGTACCCATAGCAGGAGTGTGTCTGTCTCCATGGTGG-3' [SEQ ID NO :19]

[0434] 5' -CTCCTCCTCCATCTGCTTGCTCAGGTCGCTGGTGAAGGTTCCCTCTCCGTGACCGGTG-3' [SEQ ID NO :20]

[0435] 5' -GCTAGGAGGAGGGGCTCCGCTGCTTGGTCTCCGTTCTTCAGCCACTCGATGAACAGTCTCACGGC-3' [SEQ ID NO :21]

[0436] 按实施例 1a 的描述进行杂交反应。将杂交的产物连接于用 NheI 和 Eco47III 消化的 pJB02 载体, 以产生 pJB02-Exendin-4-Fc。

[0437] 实施例 1e 构建编码 Exendin-4-HSA 的 DNA :

[0438] 按实施例 1b 的描述制备质粒 pJB02-HSA。通过描述于实施例 1d 中的重叠和互补寡核苷酸的体外杂交产生编码融合于 Exendin-4 的 Igκ 信号序列的 DNA。也按上述内容进行杂交反应。将 DNA 克隆至 pJB02-HSA 中的独特 NheI 和 FspI 位点, 以产生 pJB02-Exendin-4-HSA。

[0439] 实施例 1f 构建编码 Exendin-4-接头-HSA 的 DNA :

[0440] 按实施例 1c 的描述制备质粒 pJB02-接头-HSA。按实施例 1d 的描述制备编码融合于 Exendin-4 和接头序列的 5' 部分的 IgG κ 前导序列的 DNA。将该 DNA 克隆至 pJB02-接头-HSA 中的独特 NheI 和 BspEI 位点,以产生 pJB02-Exendin-4-接头-HSA。

[0441] 实施例 1g 构建编码 Val⁸-GLP-1/C-Ex-Fc 的 DNA:

[0442] 按实施例 1d 的描述制备质粒 pJB02-Exendin-4-Fc。用 AgeI 和 Eco47III 从载体切除 Exendin-4 的编码 DNA。通过以下重叠和互补寡核苷酸的体外杂交产生编码 Val⁸-GLP-1/C-Ex 的 DNA:

[0443] 5' -CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTGGAGGGCCAGGCCGCCA-3' [SEQ ID NO:22]

[0444] 5' -AGGAATTCATCGCCTGGCTGGTGAAGGGCCGGGGCAGCAGCGGAGCCCCCTCCTCCTAGC-3' [SEQ ID NO:23]

[0445] 5' -CTCCAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCTCCACGTGA-3' [SEQ ID NO:24]

[0446] 5' -GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGCCCTTCACCAGCCAGGCGATGAATTCCTTGGCGGCCTGGCC-3' [SEQ ID NO:25]

[0447] 按实施例 1a 的描述进行杂交反应。将杂交的产物代替 Exendin-4 连接于 pJB02-Exendin-4-Fc 表达载体,以产生 pJB02-Val⁸-GLP-1/C-Ex-Fc。

[0448] 实施例 1h 构建编码 Val⁸-Glu²²-GLP-1-Fc 的 DNA:

[0449] 按实施例 1d 的描述制备质粒 pJB02-Exendin-4-Fc。用 AgeI 和 Eco47III 从载体切除 Exendin-4 的编码 DNA。通过以下重叠和互补寡核苷酸的体外杂交产生编码 Val⁸-Glu²²-GLP-1 的 DNA:

[0450] 5' -CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTCGAGGAGCAGGCCGCCA-3' [SEQ ID NO:26]

[0451] 5' -AGGAGTTCATCGCCTGGCTGGTGAAGGGCCGGGGC-3' [SEQ ID NO:27]

[0452] 5' -GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGCCTGCTC-3' [SEQ ID NO:28]

[0453] 5' -CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCTCCACGTGA-3' [SEQ ID NO:29]

[0454] 按实施例 1a 的描述进行杂交反应。将杂交的产物代替 Exendin-4 连接于 pJB02-Exendin-4-Fc 表达载体,以产生 pJB02-Val⁸-Glu²²-GLP-1-Fc。

[0455] 实施例 1i 构建编码 Val⁸-Glu²²-GLP-1/C-Ex-Fc 的 DNA:

[0456] 按实施例 1d 的描述制备质粒 pJB02-Exendin-4-Fc。用 AgeI 和 Eco47III 从载体切除 Exendin-4 的编码 DNA。通过以下重叠和互补寡核苷酸的体外杂交产生编码 Val⁸-Glu²²-GLP-1/C-Ex 的 DNA:

[0457] 5' -CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTCGAGGAGCAGGCCGCCA-3' [SEQ ID NO:30]

[0458] 5' -AGGAATTCATCGCCTGGCTGGTGAAGGGCCGGGGCAGCAGCGGAGCCCCCTCCTCCTAGC-3' [SEQ ID NO:31]

[0459] 5' -CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCTCCACGTGA-3' [SEQ ID NO:32]

[0460] 5' -GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGCCCTTCACCAGCCAGGCGATGAATTCCTTGGCGGCCTGCTC-3' [SEQ ID NO:33]

[0461] 按实施例 1a 的描述进行杂交反应。将杂交的产物代替 Exendin-4 连接于

pJB02-Exendin-4-Fc 表达载体,以产生 pJB02-Val⁸-Glu²²-GLP-1/C-Ex-Fc。

[0462] 实施例 1j 构建编码 Gly⁸-GLP-1-Fc 的 DNA :

[0463] 按实施例 1d 的描述制备质粒 pJB02-Exendin-4-Fc。用 AgeI 和 Eco47III 从载体切除 Exendin-4 的编码 DNA。通过以下重叠和互补寡核苷酸的体外杂交产生编码 Gly⁸-GLP-1 的 DNA :

[0464] 5' -CCGGTCACGGCGAGGGCACCTTCACTAGTGACGTGTCCTCCTATCTGGAGGGCCAGGCCGCCA-3' [SEQ ID NO :34]

[0465] 5' -AGGAGTTCATCGCCTGGCTGGTGAAGGGCCGGGGC -3' [SEQ ID NO :35]

[0466] 5' -CTCCAGATAGGAGGACACGTCAGTGAAGGTGCCCTCGCCGTGA-3' [SEQ ID NO :36]

[0467] 5' -GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGCCTGGCC-3' [SEQ ID NO :37]

[0468] 按实施例 1a 的描述进行杂交反应。将杂交的产物代替 Exendin-4 连接于 pJB02-Exendin-4-Fc 表达载体,以产生 pJB02-Gly⁸-GLP-1-Fc。

[0469] 实施例 2 :表达异源融合蛋白

[0470] 通过瞬时转染 HEK 293EBNA 细胞(贴壁的和悬浮的)表达实施例 1 中的 DNA 构建体编码的融合蛋白。在转染前对细胞进行计数和种植。通过将 FuGeneTM6 转染试剂(Roche Molecular Biochemicals,目录号 1814443)与 OptiMEM(Gibco/BRL)混合而制备转染混合物,并且在室温下孵育 5 分钟,此时加入 DNA,将混合物再孵育 15 分钟。转染前,向平板中加入新鲜的生长培养基。表 1 和 2 提供了进一步的转染细节。

[0471] 表 1 :HEK 293EBNA 细胞的瞬时转染中使用的试剂

[0472]

组织培养容器	种植的细胞数	DNA (μg)	FuGene (μl)	OptiMEM 培养基 (ml)	Vol. of 生长培养基 (ml)
35mm	5 × 10 ⁵	1.5	9	0.102	2
100mm	2 × 10 ⁶	12	73	0.820	10
700cm ² (RB)	2 × 10 ⁷	65	400	4.0	100

[0473] 表 2 :培养基组成

[0474]

生长和转染培养基	收获培养基
DMEMF123 : 1	Hybritech 碱
5% FBS	1mM Ca ²⁺

生长和转染培养基	收获培养基
20mMHEPES	20mMHEPES
2mM L- 谷氨酰胺	1 μ g/ml Nuselin(人胰岛素)
50 μ g/ml 遗传霉素 (G418 NEO)	1 μ g/ml 人转铁蛋白
50 μ g/ml 妥布拉霉素	50 μ g/ml 妥布拉霉素

[0475] 对于小规模转染 (35mm-10mm 容器), 用 PBS 冲洗细胞, 并且在转染后 24 小时转移到收获培养基, 收集培养基并且每 24 小时更换一次, 持续几天。对于大规模转染 (700cm² 滚瓶), 转染后 48 小时用 PBS 冲洗滚瓶并且改变为收获培养基。在至少连续 10 天中收集培养基并且每 24 小时更换。一般地, 只有 10 次的收获物用于其后的蛋白纯化。

[0476] 实施例 3: 异源融合蛋白的纯化

[0477] 实施例 3a Val⁸-GLP-1-Fc 的纯化

[0478] 用 CUNO 过滤系统对来自于大规模转染物的约 4.5 升条件培养基 (融合蛋白表达水平为大约 20 μ g/ml) 进行过滤, 并且使用具有 10K 滤膜的 ProFlux 切向流动式过滤系统浓缩至 250ml。用 HiTrap 蛋白 A 柱, 采用 pH 7.4 的 1×PBS, 以 2ml/min 的流速捕获 Val⁸-GLP-1-Fc, 并且用 pH 3.3 的 50mM 柠檬酸洗脱。将级分 (1ml) 收集在含 4ml 1×PBS 和 1000 μ l 1M Tris pH 8 的试管中。

[0479] 将通过 SDS-PAGE 和 Zorbax C8 上的反相 HPLC 确定的含有融合蛋白的级分进行合并, 并且加入 Superdex 7560/60 柱, 1×PBS pH 7.4 中, 流速为 10ml/min。收集阳性级分 (20ml/管) 并且合并。然后将合并的级分在 0.1% TFA 水中以 3ml/min 的流速进行 C4 反相层析。在 70 分钟内用 5% B (溶于乙腈的 0.1% TFA)-100% B 的梯度洗脱 Val⁸-GLP-1-Fc。收集洗脱的级分 (3ml/管)。通过真空干燥去除乙腈, 并且加入 1ml H₂O。用 4 升 1×PBS pH7.4 透析纯化样品 (大约 32ml) 两次。

[0480] 然后用 MILLEX-GV 0.22 μ m 过滤单位过滤透析过的样品, 用 280nm 处的吸光度确定浓度。

[0481] 实施例 3b Val⁸-GLP-1-HSA 或 Val⁸-GLP-1-接头-HSA 的纯化

[0482] 用 CUNO 过滤系统对约 6.5 升条件培养基 (融合蛋白表达水平为大约 10 μ g/ml) 进行过滤, 并且使用具有 10K 滤膜的 ProFlux 切向流动式过滤系统浓缩至 380ml。

[0483] 用 50ml Fast Flow Q 柱 (Pharmacia), 在流速为 5ml/min 的 20mM Tris pH 7.4 中捕获融合蛋白。采用以下梯度洗脱蛋白: 0% -50% 20mM Tris pH 7.4, 10CV 中的 1M NaCl, 然后 2CV 中的 100% B。

[0484] 合并含融合蛋白的级分, 在 0.1% TFA 水中以 5ml/min 的流速进行 C4 反相层析。在 120 分钟内用 20% B (溶于乙腈的 0.1% TFA)-90% B 的梯度洗脱融合蛋白。收集洗脱的级分 (3.5ml/管)。通过真空干燥去除乙腈。

[0485] 用 1×PBS 将大约 9ml 合并的样品稀释至 40ml 并且用 4 升 1×PBS pH7.4 透析过夜。然后过滤样品, 用 280nm 处的吸光度确定浓度。

[0486] 实施例 3c Exendin-4-Fc 的纯化：

[0487] 用 CUNO 过滤系统对约 4 升条件培养基（融合蛋白表达水平为大约 $8\mu\text{g/ml}$ ）进行过滤，并且使用具有 30K 滤膜的 ProFlux 切向流动式过滤系统浓缩至 250ml。

[0488] 用 5ml HiTrap 蛋白 A 柱，在流速为 2ml/min 的 $1\times\text{PBS}$, pH 7.4 中捕获 Exendin-4-Fc，并且用 pH3.3 的 50mM 柠檬酸洗脱。合并含融合蛋白的级分，过滤，并且用 4 升 $1\times\text{PBS}$ pH7.4 透析过夜。然后将透析过的样品加入 Superdex 7560/60 柱， $1\times\text{PBS}$ pH7.4, 0.5M NaCl 中，流速为 10ml/min。收集含融合蛋白的级分（20ml/管），合并，并且浓缩至大约 1mg/ml 。然后用 MILLEX-GV 0.22um 滤器过滤浓缩后的样品。

[0489] 实施例 3d 纯化 Exendin-4-HSA 和 Exendin-4-接头-HSA：

[0490] 用 CUNO 过滤系统对约 1.1 升条件培养基（融合蛋白表达水平为大约 $6\mu\text{g/ml}$ ）进行过滤，并且使用具有 30K 滤膜的 ProFlux 切向流动式过滤系统浓缩至 175ml。

[0491] 然后用 50ml HiTrap Q- 琼脂糖柱 (Pharmacia)，在流速为 2ml/min 的 20mM Tris pH 7.4 中捕获融合蛋白。采用以下梯度洗脱蛋白：0% -50% 20mM Tris pH 7.4, 12CV 中的 1M NaCl，然后 4CV 中的 100% B。

[0492] 合并含融合蛋白的级分，在 0.1% TFA 水中以 5ml/min 的流速进行 C4 反相层析。在 70 分钟内用 10% B（溶于乙腈的 0.1% TFA）-100% B 的梯度洗脱融合蛋白。收集洗脱的级分（10ml/管）。通过真空干燥去除乙腈。

[0493] 用 4 升 $1\times\text{PBS}$ pH7.4 透析大约 8ml 合并的样品过夜。然后过滤样品，用 280nm 处的吸光度确定浓度。然后将透析过的样品加入 Superdex 20026/60 柱， $\times\text{PBS}$ pH 7.4, 0.5M NaCl 中，流速为 2ml/min。收集、合并、浓缩和过滤含融合蛋白的级分（3ml/管）。

[0494] 实施例 4：通过 SDS PAGE 表征融合蛋白：

[0495] 使用 SDS-PAGE 和其后的免疫印迹分析来自于用各种融合蛋白表达载体转染的纯化融合蛋白和条件培养基。用 Novex 16% Tris- 甘氨酸 Precast 凝胶 (EC6498)、电泳缓冲液 (10x, LC2675) 和样品缓冲液 (L2676) 在 Novex Powerase 500 系统上进行 SDS-PAGE。用 50mM DTT 还原样品，加样前在 95°C 加热 3-5 分钟。

[0496] 进行 SDS-PAGE 凝胶电泳后，用水和转移缓冲液 (1X Tris- 甘氨酸 Seprabuff (Owl Scientific 目录号 ER26-S)，含 20% 甲醇) 从凝胶上冲洗 SDS。使用具有 PVDF (BioRad, 目录号 162-0174) 和硝酸纤维素膜 (BioRad, 目录号 1703965 或 1703932) 的 Novex 转移装置。在室温下 30-35V 转移 90min。用含有 0.1% Tween-20 (Sigma, 目录号 P7949) 和 5% 乳汁 (BioRad, 目录号 170-6404) 的 $1\times\text{PBS}$ 将膜在 4°C 封闭 1-12 小时。将抗体稀释至 $1\times\text{PBS}+5\%$ 乳汁中并且在这些溶液中 4°C 孵育印迹 1-2 小时。在孵育之间，用 $1\times\text{PBS}$ 和 0.2% Tween-20 在室温下洗涤印迹 5 分钟，共 4 次。在 25°C 用 GIBCO $10\times\text{PBS}$ (目录号 70011) 制备 PBS 以得到 1mM 磷酸二氢钾、3mM 氢二钠、153mM 氯化钠，pH 7.4 的最终组分，或用 Sigma 的 PBS 袋 (目录号 10003) 制备 PBS 以得到 120mM NaCl, 2.7mM KCl 和 10mM 磷酸盐，pH 7.4。

[0497] 第一抗体为单克隆山羊抗 IgG1 或兔抗 HSA。第二抗体为抗山羊 IgGHRP 或抗兔 IgGHRP。对第二抗体进行 1 : 5000 稀释。用 ECL 系统 (Amersham Pharmacia Biotech, 目录号 RN2108 和目录号 RPN1674H) 对印迹进行显影。

[0498] 图 3A 比较了纯化 Fc 蛋白与来自 pJB02-Val⁸-GLP-1-Fc 和 pJB02-Exendin-4-Fc 转染的细胞的条件培养基。迁移率的减少与由于融合蛋白的 GLP-1 部分导致的大小增加相一

致。图 3B 相似地比较了纯化 HSA 和来自于 pJB02-Val⁸-GLP-1-HSA, pJB02-Val⁸-GLP-1-接头-HSA, pJB02-Exendin-4-HSA, 或 pJB02-Exendin-4-接头-HSA 转染的细胞的条件培养基。图 4 鉴定了纯化的融合蛋白制剂。

[0499] 实施例 5 :用质谱法表征融合蛋白

[0500] 所有实验都是在装备有延时聚焦电子设备、Reflectron(用于分析 0-8000Da 的肽范围)、线性检测器(用于高质量/良好信号分析)、和后加速检测器(或 P. A. D., 用于高质量/极低信号分析)的 MicromassTofSpec-2E 质谱仪上进行的。线性模式中仪器的有效径长是 1.2 米,反射(Reflectron)模式中是 2.3 米。固定两个双微通道板检测器用于线性和反式模式检测。使用的激光是 Laser Science Inc. VSL-337i 氮激光,在 337nm 下,每秒 5 次激光发射。用 2GHz,8 比特内部数字化仪获得所有数据,每个谱平均最多 50 次激光发射。

[0501] 以线性方式操作仪器,用于分析所考虑的 GLP-1 融合蛋白。线性检测器是检测沿 MALDI-ToF-MS 仪器的飞行管向下运行的离子的设备。它检测随时间的离子丰度,并且相数字化仪发送信号用于转换。数字化仪是一种模拟-数字转换器,允许质谱仪的细胞转移到计算机,在其中重建为可用的 m/z 谱。

[0502] 用重结晶的饱和芥子酸溶液(在 50/50Acn/H₂O 和 0.1% TFA 中稀释)作为离子化基质。芥子酸是大于约 10Kda 的蛋白的适当基质。将适当的参考蛋白作为内部和外部校准标准,以获得所分析的样品的准确质量测定值。采用样品与基质稀释液 1 : 2 的比例分析样品。起初在以下的线性检测器条件下设定线性检测器条件:

[0503] 源电压 :20.0keV 脉冲电压 :3.0keV

[0504] 萃取电压 :20.0keV 粗激光 :50

[0505] 聚焦电压 :16.0keV 细激光 :50

[0506] 线性检测器 :3.7keV

[0507] P. A. D. : (离线)

[0508] (根据需要)对这些设置进行调整,以得到最佳信/噪比和最高的分辨率。表 3 提供了不同 GLP-1 融合蛋白的表征。

[0509] 表 3

[0510]

融合蛋白	预期的质量 (KDa)	由质谱测定的质量 (kDa)
Val ⁸ -GLP-1-IgG1	59.08	61.94
Val ⁸ -Glu ²² -GLP-1-IgG1	59.23	63.61
Gly ⁸ -GLP-1-IgG1	59.00	62.93
Val ⁸ -GLP-1-CEX-IgG1	60.45	65.1-65.6

融合蛋白	预期的质量 (KDa)	由质谱测定的质量 (kDa)
Val ⁸ -Glu ²² -GLP-1-CEx-IgG1	60.69	65.86
Exendin-4-IgG1	60.69	65.86
Val ⁸ -GLP-1-接头-HSA	70.70	69.89, 70.74
Exendin-4-HSA	70.56	70.62
Exendin-4-接头-HSA	71.56	71.62

[0511] Cex 称作 C 末端延伸片段, 包含序列 Ser-Ser--Gly-Ala-Pro-Pro-Pro-Ser。接头是 Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser。

[0512] 实施例 6: 异源融合蛋白的活性:

[0513] 用分别描述于 Gelfand 等的 EP 619,322 和美国专利 5,120,712 中的方法等体外测定方法评估本发明的融合蛋白激活 GLP-1 受体的能力。相对于 Val⁸-GLP-1 (7-37)OH 活性的这些化合物的活性报道于表 4。图 8 表示 Val⁸-GLP-1 和 Exendin-4 融合蛋白的体外剂量反应曲线。此外, 表 5a 和 5b 提供了可以与 Fc 或白蛋白融合以制备生物活性融合蛋白的一大组 GLP-1 类似物的体外活性。这些活性是与 GLP-1 (7-37)OH 相比的结果。

[0514] 表 4: GLP-1 融合蛋白的体外活性

[0515]

融合蛋白	体外活性 (Val ⁸ -GLP-1 的百分比)
Val ⁸ -GLP-1-IgG1	1
Exendin-4-IgG1	240
Val ⁸ -GLP-1-接头-HSA	0.2
Exendin-4-HSA	20
Exendin-4-接头-HSA	90
Exendin-4	500
Val ⁸ -Glu ²² -GLP-1-IgG1	3.7

融合蛋白	体外活性 (Val ⁸ -GLP-1 的百分比)
Gly ⁸ -GLP-1-IgG1	3.3
Val ⁸ -GLP-1-CEX-IgG1	3.3
Val ⁸ -Glu ²² -GLP-1-CEX-IgG1	29
Gly ⁸ -Glu ²² -GLP-1-C2-IgG1	75
Gly ⁸ -Glu ²² -GLP-1-CEX-接头-IgG1	150
Exendin-4-C2-IgG1	250
Exendin-4-接头-IgG1	330
Gly ⁸ -Glu ²² -GLP-1-CEX-接头-HSA	4
Gly ⁸ -Glu ²² -GLP-1-CEX-接头-IgG4	80

[0516] Cex 称作 C 末端延伸片段, 包含序列 Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser。接头是 Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Ser。

[0517] C2 是 Ser-Ser-Gly-Ala-Ser-Ser-Gly-Ala。

[0518] 描述于表 3 和 4 的融合蛋白的氨基酸序列表示于 SEQ ID NO:13 至 SEQ ID NO:31。

[0519] Val⁸-GLP-1- 人血清白蛋白氨基酸序列由 SEQ ID NO:13 表示。

[0520] 1 HVEGTFTSDV SSYLEGQAAK EFIAWLVKGR GDAHKSEVAH RFKDLGEENF KALVLIAFAQ

[0521] 61 YLQQCPFEDH VKLVNEVTEF AKTCVADESA ENCDKSLHTL FGDKLCTVAT LRETYGEMAD

[0522] 121 CCAKQEPERN ECFLQHKDDN PNLRLVRPE VDMCTAFHD NEETFLKKYL YEIARRHPYF

[0523] 181 YAPELLFFAK RYKAAFTECC QAADKAAACL PKLDELDEG KASSAKQRLK CASLQKFGER

[0524] 241 AFKAWAVARL SQRFPKAEFA EVSKLVTDLT KVHTECCHGD LLECADDRAD LAKYICENQD

[0525] 301 SISSKLKECC EKPLLEKSHC IAEVENDEMP ADLPSLAADF VESKDVCCKNY AEAKDVFLGM

[0526] 361 FLYEYARRHP DYSVVLLLRL AKTYETTLEK CCAAADPHEC YAKVFDEFKP LVEEPQNLK

[0527] 421 QNCELFEQLG EYKFQNALLV RYTKKVPQVS TPTLVEVSRN LGKVGSKCCK HPEAKRMPCA

[0528] 481 EDYLSVVLNQ LCVLHEKTPV SDRVTKCCTE SLVNRPCFS ALEVDETYVP KEFNAETFTF

[0529] 541 HADICTLSEK ERQIKKTAL VELVKHKPKA TKEQLKAVMD DFAAFVEKCC KADDKETCFA

[0530] 601 EEGKKLVAAS QAALGL [SEQ ID NO:13]

[0531] Val⁸-GLP-1-接头-人血清白蛋白氨基酸序列由 SEQ ID NO:14 表示。

[0532] 1 HVEGTFTSDV SSYLEGQAAK EFIAWLVKGR GGGGSGGGG SGGGSDAHK SEVAHRFKDL

[0533] 61 GEENFKALVL IAFQYLQQC PFEDHVKLNV EVTEFAKTCV ADESAENCDK SLHTLFGDKL

[0534] 121 CTVATLRETY GEMADCCAKQ EPERNECFLQ HKDDNP NLPR LVRPEVDVMC TAFHDNEETF

- [0535] 181 LKKYLYEIAR RHPYFYAPEL LFFAKRYKAA FTECCQAADK AACLLPKLDE LRDEGKASSA
- [0536] 241 KQRLKCASLQ KGERAFKAW AVARLSQRFP KAFAEVSKL VTDLT KVHTE CCHGDLLECA
- [0537] 301 DDRADLAKYI CENQDSISSK LKECCEKPLL EKSHCIAEVE NDEMPADLPS LAADFVESKD
- [0538] 361 VCKNYAEAKD VFLGMFLY EY ARRHPDYSVV LLLRLAKTYE TTLEKCCAAA DPHECYAKVF
- [0539] 421 DEFKPLVEEP QNLIKQNC EL FEQLGEYKFQ NALLVRYTKK VPQVSTPTLV EVSRNLGKVG
- [0540] 481 SKCKKHPEAK RMPCAEDYLS VVLNQLCVLH EKTPVSDRVT KCCTESLVNR RPCFSALEVD
- [0541] 541 ETYVPKEFNA ETFTFHADIC TLSEKERQIK KQTALVELVK HKPKATKEQL KAVMDDFAAF
- [0542] 601 VEKCKADDDK ETCFAEEGKK LVAASQAALG L [SEQ ID NO :14]
- [0543] Gly8-Glu²²-GLP-1-CEx- 接头 - 人血清白蛋白氨基酸序列由 SEQ ID NO :15 表示。
- [0544] 1 HEGTFTSDV SSYLEEQA AK EFIAWLKGR GSSGAPPSG GGGSGGGGS GGGSDAHS
- [0545] 61 EVAHRFKDLG EENFKALVLI AFAQYLQQCP FEDHVKLVNE VTEFAKTCVA DESAENCDKS
- [0546] 121 LHTLFGDKLC TVATLRETYG EMADCCAKQE PERNECFLQH KDDNPNLRL VRPEVDMCT
- [0547] 181 AFHDNEETFL KKYLYEIARR HPYFYAPELL FFAKRYKAAF TECCQAADKA ACLLPKLDEL
- [0548] 241 RDEGKASSAK QRLKCASLQK FGERAFKAWA VARLSQRFPK AEFAEVSKLV TDLTKVHTEC
- [0549] 301 CHGDLLECAD DRADLAKYIC ENQDSISSKL KFCCEKPLLE KSHCIAEVEN DEMPADLPSL
- [0550] 361 AADFVESKDV CKNYAEAKDV FLGMFLYEYA RRHPDYSVVL LLRLAKTYET TLEKCCAAAD
- [0551] 421 PHECYAKVFD EFKPLVEEPQ NLIKQNC EL EQLG EYKFQ N ALLVRYTKKV PQVSTPTLVE
- [0552] 481 VSRNLGKVG S KCKKHPEAKR MPACEDYLSV VLNQLCVLHE KTPVSDRVTK CCTESLVNRR
- [0553] 541 PCFSALEVDE TYVPKEFNAE TFTFHADICT LSEKERQIKK QTALVELVKH KPKATKEQLK
- [0554] 601 AVMDDFAA FV EKCKADDKE TCFAEEGKKL VAASQAALGL [SEQ ID NO :15]
- [0555] Exendin-4- 人血清白蛋白氨基酸序列由 SEQ ID NO :16 表示。
- [0556] 1 HEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPSD AHKSEVAHRF KDLGEENFKA
- [0557] 61 LVLI AFAQYL QQCPFEDHVK LVNEVTEFAK TCVADESAEN CDKSLHTLFG DKLCTVATLR
- [0558] 121 ETYGE MADCC AKQEPERNEC FLQHKDDNPN LPRLVRPEVD VMCTAFHDNE ETFLKKYLYE
- [0559] 181 IARRHPYFYA PELLFFAKRY KAAFTCCQA ADKAAACLLPK LDEL RDEGKA SSAKQRLKCA
- [0560] 241 SLQKFGERA F KAWAVARLSQ RFPKAFAEV SKLVTDLT KV HTECCHGDL ECADDRADLA
- [0561] 301 KYICENQDSI SSKLKECCEK PLLEKSHCIA EVENDEMPAD LPSLAADFVE SKDVCKNYAE
- [0562] 361 AKDVFLGMFL YEYARRHPDY SVVLLLRLAK TYETTLEKCC AAADPHECYA KVFDEFKPLV
- [0563] 421 EEPQNLIKQN CELFEQLGEY KFQNALLVRY TTKVPQVSTP TLVEVSRNLG KVGSKCKHP
- [0564] 481 EAKRMPCAED YLSVVLNQLC VLHEKTPVSD RVT KCCTESL VNRRPCFSAL EVDETYVPKE
- [0565] 541 FNAETFTFHA DICTLSEKER QIKKQTALVE LVKHKPKATK EQLKAVMDDF AAFVEKCKA
- [0566] 601 DDKETCF AEE GKKLVAASQA ALGL [SEQ ID NO :16]
- [0567] Exendin-4- 接头 - 人血清白蛋白氨基酸序列由 SEQ ID NO :17 表示。
- [0568] 1 HEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPSG GGGSGGGGS GGGSDAHS
- [0569] 61 EVAHRFKDLG EENFKALVLI AFAQYLQQCP FEDHVKLVNE VTEFAKTCVA DESAENCDKS
- [0570] 121 LHTLFGDKLC TVATLRETYG EMADCCAKQE PERNECFLQH KDDNPNLRL VRPEVDMCT
- [0571] 181 AFHDNEETFL KKYLYEIARR HPYFYAPELL FFAKRYKAAF TECCQAADKA ACLLPKLDEL
- [0572] 241 RDEGKASSAK QRLKCASLQK FGERAFKAWA VARLSQRFPK AEFAEVSKLV TDLTKVHTEC
- [0573] 301 CHGDLLECAD DRADLAKYIC ENQDSISSKL KECCEKPLLE KSHCIAEVEN DEMPADLPSL

- [0574] 361 AADFVESKDV CKNYAEAKDV FLGMFLYEYA RRHPDYSVVL LLRLAKTYET TLEKCCAAAD
- [0575] 421 PHECYAKVFD EFKPLVEEPQ NLIKQNCSELF EQLGEYKFQN ALLVRYTKKV PQVSTPTLVE
- [0576] 481 VSRNLGKVGs KCKKHPEAKR MPACEDYLSV VLNQLCVLHE KTPVSDRVTK CCTESLVNRR
- [0577] 541 PCFSALEVDE TYVPKEFNAE TFTFHADICT LSEKERQIKK QTALVELVKH KPKATKEQLK
- [0578] 601 AVMDDFAAAFV EKCKKADDKE TCFAEEGKKL VAASQAALGL [SEQ ID NO :17]
- [0579] Val⁸-GLP-1-IgG1 氨基酸序列由 SEQ ID NO :18 表示。
- [0580] 1 HVEGTFTSDV SSYLEGQAAK EFIAWLKGR GAEPKSCDKT HTCPPCPAPE LLGGPSVFLF
- [0581] 61 PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
- [0582] 121 SVLTVLHQDW LNGKEYKCKV SNKALPAIE KTISKAKGQP REPQVYTLPP SREEMTKNQV
- [0583] 181 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSGDS FFLYSKLTVD KSRWQQGNVF
- [0584] 241 SCSVMHEALH NHYTQKSLSL SPGK [SEQ ID NO :18]
- [0585] Val⁸-GLP-1-Cex-IgG1 氨基酸序列由 SEQ ID NO :19 表示。
- [0586] 1 HVEGTFTSDV SSYLEGQAAK EFIAWLKGR GSSGAPPPSA EPKSCDKTHT CPPCPAPELL
- [0587] 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
- [0588] 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAIEKT ISKAKGQPRE PQVYTLPPSR
- [0589] 181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTP PVLDSGDSFF LYSKLTVDKS
- [0590] 241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEQ ID NO :19]
- [0591] Val⁸-Glu²²-GLP-1-IgG1 氨基酸序列由 SEQ ID NO :20 表示。
- [0592] 1 HVEGTFTSDV SSYLEEQAAK EFIAWLKGR GAEPKSCDKT HTCPPCPAPE LLGGPSVFLF
- [0593] 61 PPKPKDTLWI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
- [0594] 121 SVLTVLHQDW LNGKEYKCKV SNKALPAIE KTISKAKGQP REPQVYTLPP SREEMTKNQV
- [0595] 181 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSGDS FFLYSKLTVD KSRWQQGNVF
- [0596] 241 SCSVMHEALH NHYTQKSLSL SPGK [SEQ ID NO :20]
- [0597] Val⁸-Glu²²-GLP-1-CEX-IgG1 氨基酸序列由 SEQ ID NO :21 表示。
- [0598] 1 HVEGTFTSDV SSYLEEQAAK EFIAWLKGR GSSGAPPPSA EPKSCDKTHT CPPCPAPELL
- [0599] 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
- [0600] 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAIEKT ISKAKGQPRE PQVYTLPPSR
- [0601] 181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTP PVLDSGDSFF LYSKLTVDKS
- [0602] 241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEQ ID NO :21]
- [0603] Gly⁸-Glu²²-GLP-1-C2-IgG1 氨基酸序列由 SEQ ID NO :22 表示。
- [0604] 1 HVEGTFTSDV SSYLEEQAAK EFIAWLKGR GSSGASSGAA EPKSCDKTHT CPPCPAPELL
- [0605] 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
- [0606] 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAIEKT ISKAKGQPRE PQVYTLPPSR
- [0607] 181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTP PVLDSGDSFF LYSKLTVDKS
- [0608] 241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEQ ID NO :22]
- [0609] Gly⁸-Glu²²-GLP-1-CEX- 接头 -IgG1 氨基酸序列由 SEQ ID NO :23 表示。
- [0610] 1 HVEGTFTSDV SSYLEEQAAK EFIAWLKGR GSSGAPPPSG GGGSGGGGSG GGGSAEPKSC
- [0611] 61 DKTHTCPPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
- [0612] 121 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDWLNKEYK CKVSNKALPA PIEKTISKAK

- [0613] 181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTPPVLDLDS
- [0614] 241 DGSFFLYSKL TVDKSRWQQG NVFSCSVME ALHNHYTQKS LSLSPGK [SEQ ID NO :23]
- [0615] Gly⁸-Glu²²GLP-1-CEx- 接头 -IgG4 氨基酸序列由 SEQ ID NO :24 表示。
- [0616] 1 HEGTFTSDV SSYLEEQA AK EFIAWLKGR GSSGAPPSG GGS GGGGSG GGS AESKYG
- [0617] 61 PPCSPCAPE FLGGSVFLF PPKPKDTLMI SRTPEVTCVV VDVSDPEV QFNWYVDGVE
- [0618] 121 VHNATKPRE EQFNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKGLPSSIE KTISKAKGQP
- [0619] 181 REPQVYTLPP SQEEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDS DGS
- [0620] 241 FFLYSRLTVD KSRWQEGNVF SCSVMHEALH NHYTQKSLSL SLGK [SEQ ID NO :24]
- [0621] Gly⁸-Glu²²GLP-1-CEx-2 接头 -IgG1 氨基酸序列由 SEQ ID NO :25 表示。
- [0622] 1 HEGTFTSDV SSYLEEQA AK EFIAWLKGR GSSGAPPSG GGS GGGGSG GGS GGGGSG
- [0623] 61 GGS GGGGSA EPKCDKTHT CPPCAPELL GGS VFLFPP KPKDTLMISR TPEVTCVVVD
- [0624] 121 VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN
- [0625] 181 KALPAIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG
- [0626] 241 QPENNYKTP PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP
- [0627] 301 GK [SEQ ID NO :25]
- [0628] Gly⁸-Glu²²GLP-1-2 接头 -IgG1 氨基酸序列由 SEQ ID NO :26 表示。
- [0629] 1 HEGTFTSDV SSYLEEQA AK EFIAWLKGR GGS GGGGSG GGS SGGGSGGGG SGGGSGGGG
- [0630] 61 SAEPKCDKT HTCPPCAPE LLGGSVFLF PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV
- [0631] 121 KFNWYVDGVE VHNATKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKALPAIE
- [0632] 181 KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT
- [0633] 241 TPPVLDS DGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK
- [0634] [SEQ ID NO :26]
- [0635] Gly⁸-Glu²²GLP-1-2CEx-IgG1 氨基酸序列由 SEQ ID NO :27 表示。
- [0636] 1 HEGTFTSDV SSYLEEQA AK EFIAWLKGR GSSGAPPSG SGAPPSAEP KCDKTHTCP
- [0637] 61 PCAPELLGG PSVFLFPPK KDTLMISRTP EVTCVVVDVS HEDPEVKFNW YVDGVEVHNA
- [0638] 121 KTKPREEQYN STYRVVSVLT VLHQDWLNGK EYKCKVSNKA LPAPIEKTIS KAKGQPREPQ
- [0639] 181 VYTLPPSREE MTKNQVSLTC LVKGFYPSDI AVEWESNGQP ENNYKTPPV LDSDGSFFLY
- [0640] 241 SKLTVDKSRW QQGNVFSCSV MHEALHNHYT QKSLSLSPGK [SEQ ID NO :27]
- [0641] Gly⁸-Glu²²-Val²⁵-Ile³³GLP-1-CEx- 接头 -IgG1 氨基酸序列由 SEQ ID NO :28 表示。
- [0642] 1 HEGTFTSDV SSYLEEQA VK EFIAWLKGR GSSGAPPSG GGS GGGGSG GGS SAEPKSC
- [0643] 61 DKHTHTCPPC APELLGGSV FLFPPKPKDT LMISRTP EVT CVVDVSHED PEVKFNWYVD
- [0644] 121 GVEVHNATK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
- [0645] 181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTPPVLDLDS
- [0646] 241 DGSFFLYSKL TVDKSRWQQG NVFSCSVME ALHNHYTQKS LSLSPGK [SEQ ID NO :28]
- [0647] Exendin-4-IgG1 氨基酸序列由 SEQ ID NO :29 表示。
- [0648] 1 HEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPSA EPKCDKTHT CPPCAPELL
- [0649] 61 GGS VFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
- [0650] 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAIEKT ISKAKGQPRE PQVYTLPPSR
- [0651] 181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTP PVLDSDGSFF LYSKLTVDKS

- [0652] 241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEQID NO :29]
 [0653] Exendin-4-C2-IgG1 氨基酸序列由 SEQ ID NO :30 表示。
 [0654] 1 HEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGASSGAA EPKSCDKTHT CPPCPAPELL
 [0655] 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NQYVDGVEVH NAKTKPREEQ
 [0656] 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
 [0657] 181 EEMTKNQVSL TCLVKGFYPS DIAVEWEENG QPENNYKTP PVLDSDGSFF LYSKLTVDKS
 [0658] 241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEQ ID NO :30]
 [0659] Exendin-4- 接头 -IgG1 氨基酸序列由 SEQ ID NO :31 表示。
 [0660] 1 HEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSAEPKSC
 [0661] 61 DKHTCPPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
 [0662] 121 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
 [0663] 181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTPPVLD
 [0664] 241 DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK [SEQ ID NO :31]
 [0665] 表 5a :GLP-1 类似物的体外活性

	<u>GLP-1 化合物</u>	<u>GLP-1 受体 激活</u>
	GLP-1 (7-37)OH	1.0
	Val ⁸ -GLP-1 (7-37)OH	0.47 (n = 6)
[0666]	Gly ⁸ -His ¹¹ -GLP-1 (7-37)OH	0.282
	Val ⁸ -Ala ¹¹ -GLP-1 (7-37)OH	0.021
	Val ⁸ -Lys ¹¹ -GLP-1 (7-37)OH	0.001
	Val ⁸ -Tyr ¹² -GLP-1 (7-37)OH	0.81
	Val ⁸ -Glu ¹⁶ -GLP-1 (7-37)OH	0.047

[0667]

Val ⁸ -Ala ¹⁶ -GLP-1(7-37)OH	0.112
Val ⁸ -Tyr ¹⁶ -GLP-1(7-37)OH	1.175
Val ⁸ -Lys ²⁰ -GLP-1(7-37)OH	0.33
Gln ²² -GLP-1(7-37)OH	0.42
Val ⁸ -Ala ²² -GLP-1(7-37)OH	0.56
Val ⁸ -Ser ²² -GLP-1(7-37)OH	0.50
Val ⁸ -Asp ²² -GLP-1(7-37)OH	0.40
Val ⁸ -Glu ²² -GLP-1(7-37)OH	1.29
Val ⁸ -Lys ²² -GLP-1(7-37)OH	0.58
Val ⁸ -Pro ²² -GLP-1(7-37)OH	0.01
Val ⁸ -His ²² -GLP-1(7-37)OH	0.14
Val ⁸ -Lys ²² -GLP-1(7-36)NH ₂	0.53
Val ⁸ -Glu ²² -GLP-1(7-36)NH ₂	1.0
Gly ⁸ -Glu ²² -GLP-1(7-37)OH	1.07
Val ⁸ -Lys ²³ -GLP-1(7-37)OH	0.18
Val ⁸ -His ²⁴ -GLP-1(7-37)OH	0.007

[0668]

Val ⁸ -Lys ²⁴ -GLP-1(7-37)OH	0.02
Val ⁸ -His ²⁶ -GLP-1(7-37)OH	1.6
Val ⁸ -Glu ²⁶ -GLP-1(7-37)OH	1.5
Val ⁸ -His ²⁷ -GLP-1(7-37)OH	0.37
Val ⁸ -Ala ²⁷ -GLP-1(7-37)OH	0.47
Gly ⁸ -Glu ³⁰ -GLP-1(7-37)OH	0.29
Val ⁸ -Glu ³⁰ -GLP-1(7-37)OH	0.29
Val ⁸ -Asp ³⁰ -GLP-1(7-37)OH	0.15
Val ⁸ -Ser ³⁰ -GLP-1(7-37)OH	0.19
Val ⁸ -His ³⁰ -GLP-1(7-37)OH	0.19
Val ⁸ -Glu ³³ -GLP-1(7-37)OH	0.039
Val ⁸ -Ala ³³ -GLP-1(7-37)OH	0.1
Val ⁸ -Gly ³³ -GLP-1(7-37)OH	0.01
Val ⁸ -Glu ³⁴ -GLP-1(7-37)OH	0.17
Val ⁸ -Pro ³⁵ -GLP-1(7-37)OH	0.094
Val ⁸ -His ³⁵ -GLP-1(7-37)OH	0.41
Val ⁸ -Glu ³⁵ -GLP-1(7-37)OH	0.15

[0669]

Val ⁸ -Glu ³⁶ -GLP-1 (7-37)OH	0.11
Val ⁸ -His ³⁶ -GLP-1 (7-37)OH	0.22
Val ⁸ -His ³⁷ -GLP-1 (7-37)OH	0.33
Val ⁸ -Leu ¹⁶ -Glu ²⁶ -GLP-1 (7-37)OH	0.23
Val ⁸ -Lys ²² -Glu ³⁰ -GLP-1 (7-37)OH	0.37
Val ⁸ -Lys ²² -Glu ²³ -GLP-1 (7-37)OH	0.35
Val ⁸ -Glu ²² -Ala ²⁷ -GLP-1 (7-37)OH	1.02
Val ⁸ -Glu ²² -Lys ²³ -GLP-1 (7-37)OH	1.43
Val ⁸ -Lys ³³ -Val ³⁴ -GLP-1 (7-37)OH	0.08
Val ⁸ -Lys ³³ -Asn ³⁴ -GLP-1 (7-37)OH	0.09
Val ⁸ -Gly ³⁴ -Lys ³⁵ -GLP-1 (7-37)OH	0.34
Val ⁸ -Gly ³⁶ -Pro ³⁷ -GLP-1 (7-37)NH ₂	0.53

[0670] 表 5b :GLP-1 类似物的体外活性

[0671]

GLP-1 化合物	GLP-1 受体 激活
GLP-1 (7-37) OH	1.0
Val ⁸ -GLP-1 (7-37) OH	0.47
Gly ⁸ -GLP-1 (7-37) OH	0.80
Val ⁸ -Tyr ¹² -GLP-1 (7-37) OH	0.80
Val ⁸ -Tyr ¹² -GLP-1 (7-36) NH ₂	0.52

GLP-1 化合物	GLP-1 受体 激活
Val ⁸ -Trp ¹² -GLP-1 (7-37) OH	0.52
Val ⁸ -Leu ¹⁶ -GLP-1 (7-37) OH	0.52
Val ⁸ -Val ¹⁶ -GLP-1 (7-37) OH	0.52
Val ⁸ -Tyr ¹⁶ -GLP-1 (7-37) OH	1.18
Gly ⁸ -Glu ²² -GLP-1 (7-37) OH	1.03
Val ⁸ -Leu ²⁵ -GLP-1 (7-37) OH	0.24
Val ⁸ -Tyr ¹² -Tyr ¹⁶ -GLP-1 (7-37) OH	0.70
Val ⁸ -Trp ¹² -Glu ²² -GLP-1 (7-37) OH	0.80
Val ⁸ -Tyr ¹² -Glu ²² -GLP-1 (7-37) OH	1.27
Val ⁸ -Tyr ¹⁶ -Phe ¹⁹ -GLP-1 (7-37) OH	1.32
Val ⁸ -Tyr ¹⁶ -Glu ²² -GLP-1 (7-37) OH	1.69, 1.79
Val ⁸ -Trp ¹⁶ -Glu ²² -GLP-1 (7-37) OH	2.30, 2.16
Val ⁸ -Leu ¹⁶ -Glu ²² -GLP-1 (7-37) OH	2.02
Val ⁸ -Ile ¹⁶ -Glu ²² -GLP-1 (7-37) OH	1.55
Val ⁸ -Phe ¹⁶ -Glu ²² -GLP-1 (7-37) OH	1.08
Val ⁸ -Trp ¹⁸ -Glu ²² -GLP-1 (7-37) OH	1.50, 3.10
Val ⁸ -Tyr ¹⁸ -Glu ²² -GLP-1 (7-37) OH	2.40, 2.77
Val ⁸ -Phe ¹⁸ -Glu ²² -GLP-1 (7-37) OH	0.94
Val ⁸ -Ile ¹⁸ -Glu ²² -GLP-1 (7-37) OH	1.88

GLP-1 化合物	GLP-1 受体 激活
Val ⁸ -Lys ¹⁸ -Glu ²² -GLP-1 (7-37) OH	1. 18
Val ⁸ -Trp ¹⁹ -Glu ²² -GLP-1 (7-37) OH	1. 50
Val ⁸ -Phe ¹⁹ -Glu ²² -GLP-1 (7-37) OH	0. 70
Val ⁸ -Phe ²⁰ -Glu ²² -GLP-1 (7-37) OH	1. 27
Val ⁸ -Glu ²² -Leu ²⁵ -GLP-1 (7-37) OH	1. 32
Val ⁸ -Glu ²² -Ile ²⁵ -GLP-1 (7-37) OH	1. 46
Val ⁸ -Glu ²² -Val ²⁵ -GLP-1 (7-37) OH	2. 21, 1. 36
Val ⁸ -Glu ²² -Ile ²⁷ -GLP-1 (7-37) OH	0. 94
Val ⁸ -Glu ²² -Ala ²⁷ -GLP-1 (7-37) OH	1. 03
Val ⁸ -Glu ²² -Ile ³³ -GLP-1 (7-37) OH	2. 21, 1. 79, 1. 60
Val ⁸ -Asp ⁹ -Ile ¹¹ -Tyr ¹⁶ -Glu ²² - GLP-1 (7-37) OH	2. 02
Val ⁸ -Tyr ¹⁶ -Trp ¹⁹ -Glu ²² -GLP-1 (7- 37) OH	1. 64
Val ⁸ -Trp ¹⁶ -Glu ²² -Val ²⁵ -Ile ³³ - GLP-1 (7-37) OH	2. 35
Val ⁸ -Trp ¹⁶ -Glu ²² -Ile ³³ -GLP-1 (7- 37) OH	1. 93
Val ⁸ -Glu ²² -Val ²⁵ -Ile ³³ -GLP-1 (7- 37) OH	2. 30, 2. 73, 3. 15
Val ⁸ -Trp ¹⁶ -Glu ²² -Val ²⁵ -GLP-1 (7- 37) OH	2. 07

GLP-1 化合物	GLP-1 受体 激活
Val ⁸ -Cys ¹⁶ -Lys ²⁶ -GLP-1 (7-37) OH	1. 97
Val ⁸ -Cys ¹⁶ -Lys ²⁶ -Arg ³⁴ -GLP-1 (7-37) OH	2. 4, 1. 9

[0672]

[0673] 表 6 :GLP-1/Exendin 类似物的体外活性

[0674]

肽序列	体外活性 (Val ⁸ -GLP-1 (7-37) OH 的百分比)
HGEGTFTSDLSKQMEEEEAVRLFIEWLKNGGP-NH ₂	6. 21
HGEGTFTSDLSKQMEEEEAVRLFIEWLKNGGPSSGAPP PS-NH ₂	6. 75, 3. 25
HVEGTFTSDLSKQMEEEEAVRLFIAWLKGRG	2. 86
HVEGTFTSDVSSYLEEEAVRLFIAWLKGRG	1. 47
HVEGTFTSDLSKQMEGQAAKEFIAWLKGRG	0. 11
HVEGTFTSDVSKQMEGQAAKEFIAWLKGRG	0. 04
HGEGTFTSDLSKQMEGQAAKEFIEWLKNGGP-NH ₂	1. 44
HGEGTFTSDLSKQMEEEAKEFIEWLKNGGP-NH ₂	2. 80
HGEGTFTSDVSSYLEEEAVRLFIEWLKNGGP-NH ₂	5. 40
HGEGTFTSDLSSYLEEEAVRLFIEWLKNGGP-NH ₂	5. 07
HAEGTFTSDVSSYLEGQAAKEFIAWLKGRPSSGAPPPS -NH ₂	3. 30

肽序列	体外活性 (Val ⁸ -GLP-1 (7-37)OH 的百分比)
HAEGTFTSDVSKQLEEEAAKEFIAWLVKGRG	2.15
HVEGTFTSDVSSYLEGQAAKEFIEWLKNGGP-NH ₂	2.36
HGEGTFTSDLSKQMEEEAVRLFIAWLVKGRG	3.25
HVEGTFTSDVSSYLEEEAAKEFIAWLVKGRG	1.00
HVEGTFTSDVSSYLEGQAAKEFIAWLKNGRG	0.20
HVEGTFTSDVSSYLEGQAAKEFIAWLVKGRG	1.00
HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG	2.12

[0675]

[0676] 实施例 7:Val⁸-GLP-1-IgG1 和 Val⁸-GLP-1-HSA 的体内药代动力学:

[0677] 在 cynomolgus 猴中进行 Val⁸-GLP-1-IgG1 和 Val⁸-GLP-1-HSA 的药代动力学研究。给予猴 5.6nmol/kg 的 Val⁸-GLP-1-IgG1 或 Val⁸-GLP-1-HSA。通过静脉内给药而给予化合物。在给药前和给药后 0.083、0.25、0.5、1、4、8、12、24、48、72、96、120、144、168 和 216 小时采集血液至含 EDTA 的试管。用放射免疫测定进行具有免疫反应性的 Val⁸-GLP-1 的血浆浓度测定,该测定使用主要的特异性是针对 Val⁸-GLP-1 (7-37) 的 N 末端 (7-16) 区的多克隆抗血清。图 9 描述了对两只 cynomolgus 猴进行单次静脉内给药后 Val⁸-GLP-1-Fc 和 Val⁸-GLP-1-接头-HSA 的血浆浓度。Fc 融合蛋白的半衰期大约为 45 小时,白蛋白融合体的半衰期大约为 87 小时。

[0678] 实施例 8:Exendin-4-IgG1 的体内药代动力学:

[0679] 在禁食过夜后研究两只长期插入导管的正常雄性猎犬。通过动脉和静脉血管入口,将导管皮下插入头静脉并且缝合。将动物置于笼子中,将其导管与旋转/束缚系统连接。通过头静脉导管静脉内注射含融合蛋白 Exendin-4-IgG1 (11.8WM) 的溶液 (1.0nmol/kg)。然后用 10ml 盐水清洗导管。2 小时后,启动高血糖 (150mg/dl) 夹,持续 3 小时。在此 5 小时的时间段中取出动脉血样品,用于测定融合蛋白、葡萄糖和胰岛素的血浆浓度。

[0680] 比较此研究的结果和以前的相似研究的结果,在以前的研究中两只动物接受盐水皮下注射,3 小时后用 3 小时高血糖 (150mg/dl) 夹进行研究。

[0681] 在两个系列的研究中,使用 Beckman 血糖分析仪测定血浆葡萄糖浓度。由 Linco Research 公司的雇员采用他们的实验室开发的 RIA 试剂盒测定血浆胰岛素浓度。数据表示于图 10 和 11。

[0682] 实施例 9 :Gly⁸-Glu²²-GLP-1-CEx- 接头-IgG1 的体内药代动力学：

[0683] 通过皮下 (SC) 或静脉内 (IV) 给药使两组各三只正常雄性猎犬接受 0.1mg/kg Gly⁸-Glu²²-GLP-1-CEx- 接头-IgG1。在 IV 和 SC 组中,通过免疫测定在给药前 30 分钟至给药后 216 小时中采集的样品中测定 Gly⁸-Glu²²-GLP-1-CEx- 接头-IgG1 的血浆浓度。然后用这些浓度确定报道的药代动力学参数。IV 给予 Gly⁸-Glu²²-GLP-1-CEx- 接头-IgG1 的平均清除半衰期为大约 55 小时,总清除速度是 1.5mL/h/kg。SC 给予 Gly⁸-Glu²²-GLP-1-CEx- 接头-IgG1 的平均清除半衰期为大约 38 小时。

[0684] 序列表

[0685] <110> 伊莱利利公司

[0686] <120>GLP-1 融合蛋白

[0687] <130>X-13991

[0688] <150>US 60/251,954

[0689] <151>2000-06-12

[0690] <160>35

[0691] <170>PatentIn version 3.1

[0692] <210>1

[0693] <211>31

[0694] <212>PRT

[0695] <213> 人 (Homo sapiens)

[0696] <400>1

[0697] His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly

[0698] 1 5 10 15

[0699] Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly

[0700] 20 25 30

[0701] <210>2

[0702] <211>39

[0703] <212>PRT

[0704] <213> 人工序列

[0705] <220>

[0706] <223> 合成构建体

[0707] <220>

[0708] <221>MISC_ 特征

[0709] <222>(2).. (2)

[0710] <223> 第 2 位的 Xaa 是 Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys ;

[0711] <220>

[0712] <221>MISC_ 特征

[0713] <222>(3).. (3)

[0714] <223> 第 3 位的 Xaa 是 Glu, Asp, 或 Lys ;

[0715] <220>

- [0716] <221>MISC_特征
- [0717] <222>(5)..(5)
- [0718] <223>第5位的Xaa是Thr,Ala,Gly,Ser,Leu,Ile,Val,Glu,Asp,或Lys;
- [0719] <220>
- [0720] <221>MISC_特征
- [0721] <222>(8)..(8)
- [0722] <223>第8位的Xaa是Ser,Ala,Gly,Thr,Leu,Ile,Val,Glu,Asp,或Lys;
- [0723] <220>
- [0724] <221>MISC_特征
- [0725] <222>(10)..(10)
- [0726] <223>第10位的Xaa是Val,Ala,Gly,Ser,Thr,Leu,Ile,Tyr,Glu,Asp,或Lys;
- [0727] <220>
- [0728] <221>MISC_特征
- [0729] <222>(11)..(11)
- [0730] <223>第11位的Xaa是Ser,Ala,Gly,Thr,Leu,Ile,Val,Glu,Asp,或Lys;
- [0731] <220>
- [0732] <221>MISC_特征
- [0733] <222>(12)..(12)
- [0734] <223>第12位的Xaa是Ser,Ala,Gly,Thr,Leu,Ile,Val,Glu,Asp,Lys,Trp或Tyr;
- [0735] <220>
- [0736] <221>MISC_特征
- [0737] <222>(13)..(13)
- [0738] <223>第13位的Xaa是Tyr,Phe,Trp,Glu,Asp,Gln,或Lys;
- [0739] <220>
- [0740] <221>MISC_特征
- [0741] <222>(14)..(14)
- [0742] <223>第14位的Xaa是Leu,Ala,Gly,Ser,Thr,Ile,Val,Glu,Asp,Met,Lys,Trp或Tyr;
- [0743] <220>
- [0744] <221>MISC_特征
- [0745] <222>(15)..(15)
- [0746] <223>第15位的Xaa是Glu,Asp,或Lys;
- [0747] <220>
- [0748] <221>MISC_特征
- [0749] <222>(16)..(16)
- [0750] <223>第16位的Xaa是Gly,Ala,Ser,Thr,Leu,Ile,Val,Glu,Asp,Trp或Lys;
- [0751] <220>
- [0752] <221>MISC_特征

- [0753] <222>(17)..(17)
- [0754] <223> 第 17 位的 Xaa 是 Gln, Asn, Arg, Glu, Asp, 或 Lys ;
- [0755] <220>
- [0756] <221>MISC_ 特征
- [0757] <222>(18)..(18)
- [0758] <223> 第 18 位的 Xaa 是 Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp, 或 Lys ;
- [0759] <220>
- [0760] <221>MISC_ 特征
- [0761] <222>(19)..(19)
- [0762] <223> 第 19 位的 Xaa 是 Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys ;
- [0763] <220>
- [0764] <221>MISC_ 特征
- [0765] <222>(20)..(20)
- [0766] <223> 第 20 位的 Xaa 是 Lys, Arg, Gln, Glu, Asp, 或 His ;
- [0767] <220>
- [0768] <221>MISC_ 特征
- [0769] <222>(21)..(21)
- [0770] <223> 第 21 位的 Xaa 是 Leu, Glu, Asp, 或 Lys ;
- [0771] <220>
- [0772] <221>MISC_ 特征
- [0773] <222>(24)..(24)
- [0774] <223> 第 24 位的 Xaa 是 Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys ;
- [0775] <220>
- [0776] <221>MISC_ 特征
- [0777] <222>(25)..(25)
- [0778] <223> 第 25 位的 Xaa 是 Trp, Phe, Tyr, Glu, Asp, 或 Lys ;
- [0779] <220>
- [0780] <221>MISC_ 特征
- [0781] <222>(26)..(26)
- [0782] <223> 第 26 位的 Xaa 是 Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp, 或 Lys ;
- [0783] <220>
- [0784] <221>MISC_ 特征
- [0785] <222>(27)..(27)
- [0786] <223> 第 27 位的 Xaa 是 Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp, 或 Lys ;
- [0787] <220>
- [0788] <221>MISC_ 特征
- [0789] <222>(28)..(28)
- [0790] <223> 第 28 位的 Xaa 是 Asn, Lys, Arg, Glu, Asp, 或 His ;
- [0791] <220>

- [0792] <221>MISC_特征
- [0793] <222>(29)..(29)
- [0794] <223>第 29 位的 Xaa 是 Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys ;
- [0795] <220>
- [0796] <221>MISC_特征
- [0797] <222>(30)..(30)
- [0798] <223>第 30 位的 Xaa 是 Gly, Arg, Lys, Glu, Asp, 或 His ;
- [0799] <220>
- [0800] <221>MISC_特征
- [0801] <222>(31)..(31)
- [0802] <223>第 31 位的 Xaa 是 Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, 或 Lys, 或去除 ;
- [0803] <220>
- [0804] <221>MISC_特征
- [0805] <222>(32)..(32)
- [0806] <223>第 32 位的 Xaa 是 Ser, Arg, Lys, Glu, Asp, 或 His, 或 is deleted ;
- [0807] <220>
- [0808] <221>MISC_特征
- [0809] <222>(33)..(33)
- [0810] <223>第 33 位的 Xaa 是 Ser, Arg, Lys, Glu, Asp, 或 His, 或 is deleted ;
- [0811] <220>
- [0812] <221>MISC_特征
- [0813] <222>(34)..(34)
- [0814] <223>第 34 位的 Xaa 是 Gly, Asp, Glu, 或 Lys, 或去除 ;
- [0815] <220>
- [0816] <221>MISC_特征
- [0817] <222>(35)..(35)
- [0818] <223>第 35 位的 Xaa 是 Ala, Phe, Trp, Tyr, Glu, Asp, 或 Lys, 或 isdeleted ;
- [0819] <220>
- [0820] <221>MISC_特征
- [0821] <222>(36)..(36)
- [0822] <223>第 36 位的 Xaa 是 Ser, Pro, Lys, Glu, 或 Asp, 或去除 ;
- [0823] <220>
- [0824] <221>MISC_特征
- [0825] <222>(37)..(37)
- [0826] <223>第 37 位的 Xaa 是 Ser, Pro, Glu, Asp, 或 Lys, 或去除 ;
- [0827] <220>
- [0828] <221>MISC_特征
- [0829] <222>(38)..(38)

[0830]	<223> 第 38 位的 Xaa 是 Gly, Pro, Glu, Asp, 或 Lys, 或去除 ;
[0831]	<220>
[0832]	<221>MISC_特征
[0833]	<222>(39)..(39)
[0834]	<223> 第 39 位的 Xaa 是 Ala, Ser, Val, Glu, Asp, 或 Lys, 或 is deleted ;
[0835]	<400>2
[0836]	His Xaa Xaa Gly Xaa Phe Thr Xaa Asp Xaa Xaa Xaa Xaa Xaa Xaa Xaa
[0837]	1 5 10 15
[0838]	Xaa Xaa Xaa Xaa Xaa Phe Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
[0839]	20 25 30
[0840]	Xaa Xaa Xaa Xaa Xaa Xaa Xaa
[0841]	35
[0842]	<210>3
[0843]	<211>32
[0844]	<212>PRT
[0845]	<213>人工序列
[0846]	<220>
[0847]	<223>合成构建体
[0848]	<220>
[0849]	<221>MISC_特征
[0850]	<222>(1)..(1)
[0851]	<223> 第 1 位的 Xaa 是 L-组氨酸, D-组氨酸, 或去除 .
[0852]	<220>
[0853]	<221>MISC_特征
[0854]	<222>(2)..(2)
[0855]	<223> 第 2 位的 Xaa 是 Gly, Ala, Val, Leu, Ile, Ser, 或 Thr ;
[0856]	<220>
[0857]	<221>MISC_特征
[0858]	<222>(3)..(3)
[0859]	<223> 第 3 位的 Xaa 是 Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu, 或 His ;
[0860]	<220>
[0861]	<221>MISC_特征
[0862]	<222>(5)..(5)
[0863]	<223> 第 5 位的 Xaa 是 Asp, Glu, Arg, Thr, Ala, Lys, 或 His ;
[0864]	<220>
[0865]	<221>MISC_特征
[0866]	<222>(6)..(6)
[0867]	<223> 第 6 位的 Xaa 是 His, Trp, Phe, 或 Tyr ;
[0868]	<220>

- [0869] <221>MISC_特征
[0870] <222>(10)..(10)
[0871] <223>第 10 位的 Xaa 是 Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu, 或 Ala ;
[0872] <220>
[0873] <221>MISC_特征
[0874] <222>(12)..(12)
[0875] <223>第 12 位的 Xaa 是 His, Pro, Asp, Glu, Arg, Ser, Ala, 或 Lys ;
[0876] <220>
[0877] <221>MISC_特征
[0878] <222>(16)..(16)
[0879] <223>第 13 位的 Xaa 是 Gly, Asp, Glu, Gln, Asn, Lys, Arg, 或 Cys ;
[0880] <220>
[0881] <221>MISC_特征
[0882] <222>(17)..(17)
[0883] <223>第 17 位的 Xaa 是 His, Asp, Lys, Glu, Gln, 或 Arg ;
[0884] <220>
[0885] <221>MISC_特征
[0886] <222>(18)..(18)
[0887] <223>第 18 位的 Xaa 是 Glu, Arg, Ala, 或 Lys ;
[0888] <220>
[0889] <221>MISC_特征
[0890] <222>(20)..(20)
[0891] <223>第 20 位的 Xaa 是 Trp, Tyr, Phe, Asp, Lys, Glu, 或 His ;
[0892] <220>
[0893] <221>MISC_特征
[0894] <222>(21)..(21)
[0895] <223>第 21 位的 Xaa 是 Ala, Glu, His, Phe, Tyr, Trp, Arg, 或 Lys ;
[0896] <220>
[0897] <221>MISC_特征
[0898] <222>(24)..(24)
[0899] <223>第 24 位的 Xaa 是 Ala, Glu, Asp, Ser, 或 His ;
[0900] <220>
[0901] <221>MISC_特征
[0902] <222>(25)..(25)
[0903] <223>第 25 位的 Xaa 是 Asp, Glu, Ser, Thr, Arg, Trp, 或 Lys ;
[0904] <220>
[0905] <221>MISC_特征
[0906] <222>(27)..(27)
[0907] <223>第 27 位的 Xaa 是 Asp, Arg, Val, Lys, Ala, Gly, 或 Glu ;

- [0908] <220>
- [0909] <221>MISC_特征
- [0910] <222>(28)..(28)
- [0911] <223>第 28 位的 Xaa 是 Glu, Lys, 或 Asp ;
- [0912] <220>
- [0913] <221>MISC_特征
- [0914] <222>(29)..(29)
- [0915] <223>第 29 位的 Xaa 是 Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, 或 Glu ;
- [0916] <220>
- [0917] <221>MISC_特征
- [0918] <222>(30)..(30)
- [0919] <223>第 30 位的 Xaa 是 Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu, 或 His ;
- [0920] <220>
- [0921] <221>MISC_特征
- [0922] <222>(31)..(31)
- [0923] <223>第 31 位的 Xaa 是 Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, 或去除 .
- [0924] <220>
- [0925] <221>MISC_特征
- [0926] <222>(32)..(32)
- [0927] <223>第 31 位的 Xaa 是 Pro 或去除 .
- [0928] <400>3
- [0929] Xaa Xaa Xaa Gly Xaa Xaa Thr Ser Asp Xaa Ser Xaa Tyr Leu Glu Xaa
- [0930] 1 5 10 15
- [0931] Xaa Xaa Ala Xaa Xaa Phe Ile Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa
- [0932] 20 25 30
- [0933] <210>4
- [0934] <211>32
- [0935] <212>PRT
- [0936] <213>人工序列
- [0937] <220>
- [0938] <223>合成构建体
- [0939] <220>
- [0940] <221>MISC_特征
- [0941] <222>(1)..(1)
- [0942] <223>第 1 位的 Xaa 是 L- 组氨酸, D- 组氨酸, 或去除 .
- [0943] <220>
- [0944] <221>MISC_特征

- [0945] <222>(2)..(2)
- [0946] <223> 第 2 位的 Xaa 是 Gly, Ala, Val, Leu, Ile, Ser, 或 Thr ;
- [0947] <220>
- [0948] <221>MISC_ 特征
- [0949] <222>(5)..(5)
- [0950] <223> 第 5 位的 Xaa 是 Asp, Glu, Arg, Thr, Ala, Lys, 或 His ;
- [0951] <220>
- [0952] <221>MISC_ 特征
- [0953] <222>(6)..(6)
- [0954] <223> 第 6 位的 Xaa 是 His, Trp, Phe, 或 Tyr ;
- [0955] <220>
- [0956] <221>MISC_ 特征
- [0957] <222>(10)..(10)
- [0958] <223> 第 10 位的 Xaa 是 Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, 或 Ala ;
- [0959] <220>
- [0960] <221>MISC_ 特征
- [0961] <222>(16)..(16)
- [0962] <223> 第 16 位的 Xaa 是 Gly, Asp, Glu, Gln, Asn, Lys, Arg, 或 Cys ;
- [0963] <220>
- [0964] <221>MISC_ 特征
- [0965] <222>(17)..(17)
- [0966] <223> 第 17 位的 Xaa 是 His, Asp, Lys, Glu, 或 Gln ;
- [0967] <220>
- [0968] <221>MISC 特征
- [0969] <222>(18)..(18)
- [0970] <223> 第 18 位的 Xaa 是 Glu, His, Ala, 或 Lys ;
- [0971] <220>
- [0972] <221>MISC_ 特征
- [0973] <222>(19)..(19)
- [0974] <223> 第 19 位的 Xaa 是 Asp, Lys, Glu, 或 His ;
- [0975] <220>
- [0976] <221>MISC_ 特征
- [0977] <222>(21)..(21)
- [0978] <223> 第 21 位的 Xaa 是 Ala, Glu, His, Phe, Tyr, Trp, Arg, 或 Lys ;
- [0979] <220>
- [0980] <221>MISC_ 特征
- [0981] <222>(24)..(24)
- [0982] <223> 第 24 位的 Xaa 是 Ala, Glu, Asp, Ser, 或 His ;
- [0983] <220>

- [0984] <221>MISC_特征
- [0985] <222>(27)..(27)
- [0986] <223>第 27 位的 Xaa 是 Asp, Arg, Val, Lys, Ala, Gly, 或 Glu ;
- [0987] <220>
- [0988] <221>MISC_特征
- [0989] <222>(28)..(28)
- [0990] <223>第 28 位的 Xaa 是 Glu, Lys, 或 Asp ;
- [0991] <220>
- [0992] <221>MISC_特征
- [0993] <222>(29)..(29)
- [0994] <223>第 29 位的 Xaa 是 Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, 或 Glu ;
- [0995] <220>
- [0996] <221>MISC_特征
- [0997] <222>(30)..(30)
- [0998] <223>第 30 位的 Xaa 是 Arg, Glu, 或 His ;
- [0999] <220>
- [1000] <221>MISC_特征
- [1001] <222>(31)..(31)
- [1002] <223>第 31 位的 Xaa 是 Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, 或去除 .
- [1003] <220>
- [1004] <221>MISC_特征
- [1005] <222>(32)..(32)
- [1006] <223>第 32 位的 Xaa 是 Pro, 或去除 .
- [1007] <400>4
- [1008] Xaa Xaa Glu Gly Xaa Xaa Thr Ser Asp Xaa Ser Ser Tyr Leu Glu Xaa
- [1009] 1 5 10 15
- [1010] Xaa Xaa Xaa Lys Xaa Phe Ile Xaa Trp Leu Xaa Xaa Xaa Xaa Xaa Xaa
- [1011] 20 25 30
- [1012] <210>5
- [1013] <211>32
- [1014] <212>PRT
- [1015] <213>人工序列
- [1016] <220>
- [1017] <223>合成构建体
- [1018] <220>
- [1019] <221>MISC_特征
- [1020] <222>(1)..(1)

- [1021] <223> 第 1 位的 Xaa 是 L- 组氨酸, D- 组氨酸, 或去除 .
- [1022] <220>
- [1023] <221>MISC_ 特征
- [1024] <222>(2).. (2)
- [1025] <223> 第 2 位的 Xaa 是 Gly, Ala, Val, Leu, Ile, Ser, Met, 或 Thr ;
- [1026] <220>
- [1027] <221>MISC_ 特征
- [1028] <222>(6).. (6)
- [1029] <223> 第 6 位的 Xaa 是 His, Trp, Phe, 或 Tyr ;
- [1030] <220>
- [1031] <221>MISC_ 特征
- [1032] <222>(10).. (10)
- [1033] <223> 第 10 位的 Xaa 是 Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, 或 Ala ;
- [1034] <220>
- [1035] <221>MISC_ 特征
- [1036] <222>(16).. (16)
- [1037] <223> 第 16 位的 Xaa 是 Gly, Asp, Glu, Gln, Asn, Lys, Arg, 或 Cys ;
- [1038] <220>
- [1039] <221>MISC_ 特征
- [1040] <222>(17).. (17)
- [1041] <223> 第 17 位的 Xaa 是 His, Asp, Lys, Glu, 或 Gln ;
- [1042] <220>
- [1043] <221>MISC_ 特征
- [1044] <222>(20).. (20)
- [1045] <223> 第 20 位的 Xaa 是 Asp, Lys, Glu, 或 His ;
- [1046] <220>
- [1047] <221>MISC_ 特征
- [1048] <222>(24).. (24)
- [1049] <223> 第 24 位的 Xaa 是 Ala, Glu, Asp, Ser, 或 His ;
- [1050] <220>
- [1051] <221>MISC_ 特征
- [1052] <222>(29).. (29)
- [1053] <223> 第 29 位的 Xaa 是 Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, 或 Glu ;
- [1054] <220>
- [1055] <221>MISC_ 特征
- [1056] <222>(31).. (31)
- [1057] <223> 第 31 位的 Xaa 是 Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, 或去除 .

- [1058] <220>
- [1059] <221>MISC_特征
- [1060] <222>(32)..(32)
- [1061] <223>第 32 位的 Xaa 是 Pro 或去除。
- [1062] <400>5
- [1063] Xaa Xaa Glu Gly Thr Xaa Thr Ser Asp Xaa Ser Ser Tyr Leu Glu Xaa
- [1064] 1 5 10 15
- [1065] Xaa Ala Ala Xaa Glu Phe Ile Xaa Trp Leu Val Lys XaaArg Xaa Xaa
- [1066] 20 25 30
- [1067] <210>6
- [1068] <211>32
- [1069] <212>PRT
- [1070] <213>人工序列
- [1071] <220>
- [1072] <223>合成构建体
- [1073] <220>
- [1074] <221>MISC_特征
- [1075] <222>(1)..(1)
- [1076] <223>第 1 位的 Xaa 是 L-组氨酸, D-组氨酸, 或去除。
- [1077] <220>
- [1078] <221>MISC_特征
- [1079] <222>(2)..(2)
- [1080] <223>第 2 位的 Xaa 是 Gly, Ala, Val, Leu, Ile, Ser, 或 Thr ;
- [1081] <220>
- [1082] <221>MISC_特征
- [1083] <222>(16)..(16)
- [1084] <223>第 16 位的 Xaa 是 Gly, Asp, Glu, Gln, Asn, Lys, Arg, 或 Cys ;
- [1085] <220>
- [1086] <221>MISC_特征
- [1087] <222>(17)..(17)
- [1088] <223>第 17 位的 Xaa 是 His, Asp, Lys, Glu, 或 Gln ;
- [1089] <220>
- [1090] <221>MISC_特征
- [1091] <222>(18)..(18)
- [1092] <223>第 18 位的 Xaa 是 Ala, Glu, His, Phe, Tyr, Trp, Arg, 或 Lys ;
- [1093] <220>
- [1094] <221>MISC_特征
- [1095] <222>(24)..(24)
- [1096] <223>第 24 位的 Xaa 是 Ala, Glu, Asp, Ser, 或 His ;

- [1097] <220>
- [1098] <221>MISC_ 特征
- [1099] <222>(31).. (31)
- [1100] <223> 第 31 位的 Xaa 是 Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, 或去除 .
- [1101] <220>
- [1102] <221>MISC_ 特征
- [1103] <222>(32).. (32)
- [1104] <223> 第 32 位的 Xaa 是 Pro 或去除 .
- [1105] <400>6
- [1106] Xaa Xaa Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Xaa
- [1107] 1 5 10 15
- [1108] Xaa Xaa Ala Lys Glu Phe Ile Xaa Trp Leu Val Lys Gly Arg Xaa Xaa
- [1109] 20 25 30
- [1110] <210>7
- [1111] <211>31
- [1112] <212>PRT
- [1113] <213> 人工序列
- [1114] <220>
- [1115] <223> 合成构建体
- [1116] <220>
- [1117] <221>MISC_ 特征
- [1118] <222>(1).. (1)
- [1119] <223> 第 1 位的 Xaa 是 L- 组氨酸, D- 组氨酸, 或去除 .
- [1120] <220>
- [1121] <221>MISC_ 特征
- [1122] <222>(2).. (2)
- [1123] <223> 第 2 位的 Xaa 是 Ala, Gly, Val, Thr, Ile, 和 (- 甲基 -Ala ;
- [1124] <220>
- [1125] <221>MISC_ 特征
- [1126] <222>(15).. (15)
- [1127] <223> 第 15 位的 Xaa 是 Glu, Gln, Ala, Thr, Ser, 和 Gly ;
- [1128] <220>
- [1129] <221>MISC_ 特征
- [1130] <222>(21).. (21)
- [1131] <223> 第 21 位的 Xaa 是 Glu, Gln, Ala, Thr, Ser, 和 Gly ;
- [1132] <400>7
- [1133] Xaa Xaa Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Xaa Gly
- [1134] 1 5 10 15

[1135] Gln Ala Ala Lys Xaa Phe Ile Ala Trp Leu Val Lys GlyArg Gly
 [1136] 20 25 30
 [1137] <210>8
 [1138] <211>30
 [1139] <212>PRT
 [1140] <213> 人工序列
 [1141] <220>
 [1142] <223> 合成构建体
 [1143] <220>
 [1144] <221>MISC_ 特征
 [1145] <222>(19).. (19)
 [1146] <223> 第 19 位的 Xaa 是 Lys 或 Arg ;
 [1147] <220>
 [1148] <221>MOD_RES
 [1149] <222>(27).. (27)
 [1150] <223> 乙酰化
 [1151] <220>
 [1152] <221>MISC_ 特征
 [1153] <222>(30).. (30)
 [1154] <223> 第 30 位的 Xaa 是 Gly ;
 [1155] <220>
 [1156] <221>MOD_RES
 [1157] <222>(30).. (30)
 [1158] <223> 酰胺化
 [1159] <400>8
 [1160] Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly Gln
 [1161] 1 5 10 15
 [1162] Ala Ala Xaa Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Xaa
 [1163] 20 25 30
 [1164] <210>9
 [1165] <211>39
 [1166] <212>PRT
 [1167] <213> 人工序列
 [1168] <220>
 [1169] <223> 合成构建体
 [1170] <400>9
 [1171] His Ser Asp Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 [1172] 1 5 10 15
 [1173] Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser

[1174]	20	25	30
[1175]	Ser Gly Ala Pro Pro Pro Ser		
[1176]	35		
[1177]	<210>10		
[1178]	<211>39		
[1179]	<212>PRT		
[1180]	<213> 人工序列		
[1181]	<220>		
[1182]	<223> 合成构建体		
[1183]	<400>10		
[1184]	His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu		
[1185]	1	5	10
[1186]	Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser		
[1187]	20	25	30
[1188]	Ser Gly Ala Pro Pro Pro Ser		
[1189]	35		
[1190]	<210>11		
[1191]	<211>39		
[1192]	<212>PRT		
[1193]	<213> 人工序列		
[1194]	<220>		
[1195]	<223> 合成构建体		
[1196]	<220>		
[1197]	<221>MISC_ 特征		
[1198]	<222>(1).. (1)		
[1199]	<223> 第 1 位的 Xaa 是 L- 组氨酸, D- 组氨酸, 或去除 .		
[1200]	<220>		
[1201]	<221>MISC_ 特征		
[1202]	<222>(2).. (2)		
[1203]	<223> 第 2 位的 Xaa 是 Gly, Ala, 或 Val ;		
[1204]	<220>		
[1205]	<221>MISC_ 特征		
[1206]	<222>(10).. (10)		
[1207]	<223> 第 10 位的 Xaa 是 Leu 或 Val ;		
[1208]	<220>		
[1209]	<221>MISC_ 特征		
[1210]	<222>(12).. (12)		
[1211]	<223> 第 12 位的 Xaa 是 Lys 或 Ser ;		
[1212]	<220>		

- [1213] <221>MISC_特征
[1214] <222>(13)..(13)
[1215] <223>第 13 位的 Xaa 是 Gln 或 Tyr ;
[1216] <220>
[1217] <221>MISC_特征
[1218] <222>(14)..(14)
[1219] <223>第 14 位的 Xaa 是 Met 或 Leu ;
[1220] <220>
[1221] <221>MISC_特征
[1222] <222>(16)..(16)
[1223] <223>第 16 位的 Xaa 是 Glu 或 Gln ;
[1224] <220>
[1225] <221>MISC_特征
[1226] <222>(17)..(17)
[1227] <223>第 17 位的 Xaa 是 Glu 或 Gln ;
[1228] <220>
[1229] <221>MISC_特征
[1230] <222>(19)..(19)
[1231] <223>第 19 位的 Xaa 是 Val 或 Ala ;
[1232] <220>
[1233] <221>MISC_特征
[1234] <222>(20)..(20)
[1235] <223>第 20 位的 Xaa 是 Arg 或 Lys ;
[1236] <220>
[1237] <221>MISC_特征
[1238] <222>(21)..(21)
[1239] <223>第 21 位的 Xaa 是 Leu 或 Glu ;
[1240] <220>
[1241] <221>MISC_特征
[1242] <222>(24)..(24)
[1243] <223>第 24 位的 Xaa 是 Glu 或 Ala ;
[1244] <220>
[1245] <221>MISC_特征
[1246] <222>(27)..(27)
[1247] <223>第 27 位的 Xaa 是 Val 或 Lys ;
[1248] <220>
[1249] <221>MISC_特征
[1250] <222>(28)..(28)
[1251] <223>第 28 位的 Xaa 是 Asn 或 Lys ;

- [1252] <220>
[1253] <221>MISC_特征
[1254] <222>(30)..(30)
[1255] <223>第 30 位的 Xaa 是 Gly 或 Arg ;和
[1256] <220>
[1257] <221>MISC_特征
[1258] <222>(31)..(31)
[1259] <223>第 31 位的 Xaa 是 Gly,或 Pro ;
[1260] <220>
[1261] <221>MISC_特征
[1262] <222>(32)..(32)
[1263] <223>第 32 位的 Xaa 是 Ser,或缺乏 .
[1264] <220>
[1265] <221>MISC_特征
[1266] <222>(33)..(33)
[1267] <223>第 33 位的 Xaa 是 Ser,或缺乏 .
[1268] <220>
[1269] <221>MISC_特征
[1270] <222>(34)..(34)
[1271] <223>第 34 位的 Xaa 是 Gly,或缺乏 .
[1272] <220>
[1273] <221>MISC_特征
[1274] <222>(35)..(35)
[1275] <223>第 35 位的 Xaa 是 Ala,或缺乏 .
[1276] <220>
[1277] <221>MISC_特征
[1278] <222>(36)..(36)
[1279] <223>第 36 位的 Xaa 是 Pro,或缺乏 .
[1280] <220>
[1281] <221>MISC_特征
[1282] <222>(37)..(37)
[1283] <223>第 37 位的 Xaa 是 Pro,或缺乏 .
[1284] <220>
[1285] <221>MISC_特征
[1286] <222>(38)..(38)
[1287] <223>第 38 位的 Xaa 是 Pro,或缺乏 .
[1288] <220>
[1289] <221>MISC_特征
[1290] <222>(39)..(39)

- [1291] <223> 第 39 位的 Xaa 是 Pro,或缺乏 .
- [1292] <220>
- [1293] <221>MISC_ 特征
- [1294] <222>(39).. (39)
- [1295] <223> 第 39 位的 Xaa 是 Ser,或缺乏 .
- [1296] <400>11
- [1297] Xaa Xaa Glu Glu Thr Phe Thr Ser Asp Xaa Ser Xaa Xaa Xaa Glu Xaa
- [1298] 1 5 10 15
- [1299] Xaa Ala Xaa Xaa Xaa Phe Ile Xaa Trp Leu Xaa Xaa Gly Xaa Xaa Xaa
- [1300] 20 25 30
- [1301] Xaa Xaa Xaa Xaa Xaa Xaa Xaa
- [1302] 35
- [1303] <210>12
- [1304] <211>31
- [1305] <212>PRT
- [1306] <213> 人工序列
- [1307] <220>
- [1308] <223> 合成构建体
- [1309] <220>
- [1310] <221>MISC_ 特征
- [1311] <222>(1).. (1)
- [1312] <223> 第 1 位的 Xaa 是 L- 组氨酸, D- 组氨酸,或去除 .
- [1313] <220>
- [1314] <221>MISC_ 特征
- [1315] <222>(2).. (2)
- [1316] <223> 第 2 位的 Xaa 是 Ala, Gly, Val, Leu, Ile, Ser,或 Thr ;
- [1317] <220>
- [1318] <221>MISC_ 特征
- [1319] <222>(6).. (6)
- [1320] <223> 第 6 位的 Xaa 是 Phe, Trp,或 Tyr ;
- [1321] <220>
- [1322] <221>MISC_ 特征
- [1323] <222>(10).. (10)
- [1324] <223> 第 10 位的 Xaa 是 Val, Trp, Ile, Leu, Phe,或 Tyr ;
- [1325] <220>
- [1326] <221>MISC_ 特征
- [1327] <222>(12).. (12)
- [1328] <223> 第 12 位的 Xaa 是 Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val ;
- [1329] <220>

- [1330] <221>MISC_特征
[1331] <222>(13)..(13)
[1332] <223>第 13 位的 Xaa 是 Tyr, Trp, 或 Phe ;
[1333] <220>
[1334] <221>MISC_特征
[1335] <222>(14)..(14)
[1336] <223>第 14 位的 Xaa 是 Leu, Phe, Tyr, 或 Trp ;
[1337] <220>
[1338] <221>MISC_特征
[1339] <222>(16)..(16)
[1340] <223>第 16 位的 Xaa 是 Gly, Glu, Asp, 或 Lys ;
[1341] <220>
[1342] <221>MISC_特征
[1343] <222>(19)..(19)
[1344] <223>第 19 位的 Xaa 是 Ala, Val, Ile, 或 Leu ;
[1345] <220>
[1346] <221>MISC_特征
[1347] <222>(21)..(21)
[1348] <223>第 21 位的 Xaa 是 Glu, Ile, 或 Ala ;
[1349] <220>
[1350] <221>MISC_特征
[1351] <222>(24)..(24)
[1352] <223>第 24 位的 Xaa 是 Ala 或 Glu ;
[1353] <220>
[1354] <221>MISC_特征
[1355] <222>(27)..(27)
[1356] <223>第 27 位的 Xaa 是 Val 或 Ile ;和
[1357] <220>
[1358] <221>MOD_RES
[1359] <222>(30)..(30)
[1360] <223>酰胺化
[1361] <220>
[1362] <221>MISC_特征
[1363] <222>(31)..(31)
[1364] <223>第 31 位的 Xaa 是 Gly, His 或缺乏 .
[1365] <400>12
[1366] Xaa Xaa Glu Gly Thr Xaa Thr Ser Asp Xaa Ser Xaa Xaa Xaa Glu Xaa
[1367] 1 5 10 15
[1368] Gln Ala Xaa Lys Xaa Phe Ile Xaa Trp Leu Xaa Lys Gly Arg Xaa

[1369]	20	25	30
[1370]	<210>13		
[1371]	<211>616		
[1372]	<212>PRT		
[1373]	<213> 人工序列		
[1374]	<220>		
[1375]	<223> 合成构建体		
[1376]	<400>13		
[1377]	His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly		
[1378]	1	5	10 15
[1379]	Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Asp		
[1380]	20	25	30
[1381]	Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu		
[1382]	35	40	45
[1383]	Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln		
[1384]	50	55	60
[1385]	Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe		
[1386]	65	70	75 80
[1387]	Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser		
[1388]	85	90	95
[1389]	Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg		
[1390]	100	105	110
[1391]	Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu		
[1392]	115	120	125
[1393]	Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro		
[1394]	130	135	140
[1395]	Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp		
[1396]	145	150	155 160
[1397]	Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg		
[1398]	165	170	175
[1399]	His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr		
[1400]	180	185	190
[1401]	Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys		
[1402]	195	200	205
[1403]	Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser		
[1404]	210	215	220
[1405]	Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg		
[1406]	225	230	235 240
[1407]	Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys		

[1408]	245	250	255
[1409]	Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val		
[1410]	260	265	270
[1411]	His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg		
[1412]	275	280	285
[1413]	Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser		
[1414]	290	295	300
[1415]	Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys		
[1416]	305	310	315
[1417]	Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu		
[1418]	325	330	335
[1419]	Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu		
[1420]	340	345	350
[1421]	Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg		
[1422]	355	360	365
[1423]	His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr		
[1424]	370	375	380
[1425]	Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys		
[1426]	385	390	395
[1427]	Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln		
[1428]	405	410	415
[1429]	Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr		
[1430]	420	425	430
[1431]	Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln		
[1432]	435	440	445
[1433]	Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val		
[1434]	450	455	460
[1435]	Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala		
[1436]	4654	704	75
[1437]	Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu		
[1438]	485	490	495
[1439]	Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu		
[1440]	500	505	510
[1441]	Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr		
[1442]	515	520	525
[1443]	Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile		
[1444]	530	535	540
[1445]	Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu		
[1446]	545	550	555
			560

[1447] Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys
 [1448] 565 570 575
 [1449] Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala
 [1450] 580 585 590
 [1451] Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala
 [1452] 595 600 605
 [1453] Ala Ser Gln Ala Ala Leu Gly Leu
 [1454] 610 615
 [1455] <210>14
 [1456] <211>631
 [1457] <212>PRT
 [1458] <213>人工序列
 [1459] <220>
 [1460] <223>合成构建体
 [1461] <400>14
 [1462] His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 [1463] 1 5 10 15
 [1464] Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Gly
 [1465] 20 25 30
 [1466] Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ala
 [1467] 35 40 45
 [1468] His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn
 [1469] 50 55 60
 [1470] Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Cys
 [1471] 65 70 75 80
 [1472] Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala
 [1473] 85 90 95
 [1474] Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu
 [1475] 100 105 110
 [1476] His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu
 [1477] 115 120 125
 [1478] Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg
 [1479] 130 135 140
 [1480] Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg
 [1481] 145 150 155 160
 [1482] Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn
 [1483] 165 170 175
 [1484] Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His
 [1485] 180 185 190

[1486]	Pro	Tyr	Phe	Tyr	Ala	Pro	Glu	Leu	Leu	Phe	Phe	Ala	Lys	Arg	Tyr	Lys
[1487]																
[1488]	Ala	Ala	Phe	Thr	Glu	Cys	Cys	Gln	Ala	Ala	Asp	Lys	Ala	Ala	Cys	Leu
[1489]																
[1490]	Leu	Pro	Lys	Leu	Asp	Glu	Leu	Arg	Asp	Glu	Gly	Lys	Ala	Ser	Ser	Ala
[1491]																
[1492]	Lys	Gln	Arg	Leu	Lys	Cys	Ala	Ser	Leu	Gln	Lys	Phe	Gly	Glu	Arg	Ala
[1493]																
[1494]	Phe	Lys	Ala	Trp	Ala	Val	Ala	Arg	Leu	Ser	Gln	Arg	Phe	Pro	Lys	Ala
[1495]																
[1496]	Glu	Phe	Ala	Glu	Val	Ser	Lys	Leu	Val	Thr	Asp	Leu	Thr	Lys	Val	His
[1497]																
[1498]	Thr	Glu	Cys	Cys	His	Gly	Asp	Leu	Leu	Glu	Cys	Ala	Asp	Asp	Arg	Ala
[1499]																
[1500]	Asp	Leu	Ala	Lys	Tyr	Ile	Cys	Glu	Asn	Gln	Asp	Ser	Ile	Ser	Ser	Lys
[1501]																
[1502]	Leu	Lys	Glu	Cys	Cys	Glu	Lys	Pro	Leu	Leu	Glu	Lys	Ser	His	Cys	Ile
[1503]																
[1504]	Ala	Glu	Val	Glu	Asn	Asp	Glu	Met	Pro	Ala	Asp	Leu	Pro	Ser	Leu	Ala
[1505]																
[1506]	Ala	Asp	Phe	Val	Glu	Ser	Lys	Asp	Val	Cys	Lys	Asn	Tyr	Ala	Glu	Ala
[1507]																
[1508]	Lys	Asp	Val	Phe	Leu	Gly	Met	Phe	Leu	Tyr	Glu	Tyr	Ala	Arg	Arg	His
[1509]																
[1510]	Pro	Asp	Tyr	Ser	Val	Val	Leu	Leu	Leu	Arg	Leu	Ala	Lys	Thr	Tyr	Glu
[1511]																
[1512]	Thr	Thr	Leu	Glu	Lys	Cys	Cys	Ala	Ala	Ala	Asp	Pro	His	Glu	Cys	Tyr
[1513]																
[1514]	Ala	Lys	Val	Phe	Asp	Glu	Phe	Lys	Pro	Leu	Val	Glu	Glu	Pro	Gln	Asn
[1515]																
[1516]	Leu	Ile	Lys	Gln	Asn	Cys	Glu	Leu	Phe	Glu	Gln	Leu	Gly	Glu	Tyr	Lys
[1517]																
[1518]	Phe	Gln	Asn	Ala	Leu	Leu	Val	Arg	Tyr	Thr	Lys	Lys	Val	Pro	Gln	Val
[1519]																
[1520]	Ser	Thr	Pro	Thr	Leu	Val	Glu	Val	Ser	Arg	Asn	Leu	Gly	Lys	Val	Gly
[1521]																
[1522]	Ser	Lys	Cys	Cys	Lys	His	Pro	Glu	Ala	Lys	Arg	Met	Pro	Cys	Ala	Glu
[1523]																
[1524]	Asp	Tyr	Leu	Ser	Val	Val	Leu	Asn	Gln	Leu	Cys	Val	Leu	His	Glu	Lys

[1525]	500	505	510
[1526]	Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val		
[1527]	515	520	525
[1528]	Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val		
[1529]	530	535	540
[1530]	Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys		
[1531]	545	550	555
[1532]	Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val		
[1533]	565	570	575
[1534]	Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala		
[1535]	580	585	590
[1536]	Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp		
[1537]	595	600	605
[1538]	Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala		
[1539]	610	615	620
[1540]	Ser Gln Ala Ala Leu Gly Leu		
[1541]	625	630	
[1542]	<210>15		
[1543]	<211>640		
[1544]	<212>PRT		
[1545]	<213> 人工序列		
[1546]	<220>		
[1547]	<223> 合成构建体		
[1548]	<400>15		
[1549]	His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu		
[1550]	1	5	10
[1551]	Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser		
[1552]	20	25	30
[1553]	Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly		
[1554]	35	40	45
[1555]	Gly Ser Gly Gly Gly Gly Ser Asp Ala His Lys Ser Glu Val Ala His		
[1556]	50	55	60
[1557]	Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile		
[1558]	65	70	75
[1559]	Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys		
[1560]	85	90	95
[1561]	Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu		
[1562]	100	105	110
[1563]	Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys		

[1564]	115	120	125
[1565]	Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp		
[1566]	130	135	140
[1567]	Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His		
[1568]	145	150	155
[1569]	Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp		
[1570]	165	170	175
[1571]	Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys		
[1572]	180	185	190
[1573]	Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu		
[1574]	195	200	205
[1575]	Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys		
[1576]	210	215	220
[1577]	Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu		
[1578]	225	230	235
[1579]	Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala		
[1580]	245	250	255
[1581]	Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala		
[1582]	260	265	270
[1583]	Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys		
[1584]	275	280	285
[1585]	Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp		
[1586]	290	295	300
[1587]	Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys		
[1588]	305	310	315
[1589]	Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys		
[1590]	325	330	335
[1591]	Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu		
[1592]	340	345	350
[1593]	Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys		
[1594]	355	360	365
[1595]	Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met		
[1596]	370	375	380
[1597]	Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu		
[1598]	385	390	395
[1599]	Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys		
[1600]	405	410	415
[1601]	Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe		
[1602]	420	425	430

[1603]	Lys	Pro	Leu	Val	Glu	Glu	Pro	Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	Glu
[1604]				435					440						445	
[1605]	Leu	Phe	Glu	Gln	Leu	Gly	Glu	Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	Val
[1606]				450					455						460	
[1607]	Arg	Tyr	Thr	Lys	Lys	Val	Pro	Gln	Val	Ser	Thr	Pro	Thr	Leu	Val	Glu
[1608]				465					470						475	480
[1609]	Val	Ser	Arg	Asn	Leu	Gly	Lys	Val	Gly	Ser	Lys	Cys	Cys	Lys	His	Pro
[1610]					485					490					495	
[1611]	Glu	Ala	Lys	Arg	Met	Pro	Cys	Ala	Glu	Asp	Tyr	Leu	Ser	Val	Val	Leu
[1612]				500						505					510	
[1613]	Asn	Gln	Leu	Cys	Val	Leu	His	Glu	Lys	Thr	Pro	Val	Ser	Asp	Arg	Val
[1614]				515						520					525	
[1615]	Thr	Lys	Cys	Cys	Thr	Glu	Ser	Leu	Val	Asn	Arg	Arg	Pro	Cys	Phe	Ser
[1616]				530						535					540	
[1617]	Ala	Leu	Glu	Val	Asp	Glu	Thr	Tyr	Val	Pro	Lys	Glu	Phe	Asn	Ala	Glu
[1618]				545								550		555		560
[1619]	Thr	Phe	Thr	Phe	His	Ala	Asp	Ile	Cys	Thr	Leu	Ser	Glu	Lys	Glu	Arg
[1620]					565						570				575	
[1621]	Gln	Ile	Lys	Lys	Gln	Thr	Ala	Leu	Val	Glu	Leu	Val	Lys	His	Lys	Pro
[1622]				580							585				590	
[1623]	Lys	Ala	Thr	Lys	Glu	Gln	Leu	Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala
[1624]				595						600					605	
[1625]	Phe	Val	Glu	Lys	Cys	Cys	Lys	Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala
[1626]				610						615					620	
[1627]	Glu	Glu	Gly	Lys	Lys	Leu	Val	Ala	Ala	Ser	Gln	Ala	Ala	Leu	Gly	Leu
[1628]				625						630				635		640
[1629]	<210>16															
[1630]	<211>624															
[1631]	<212>PRT															
[1632]	<213> 人工序列															
[1633]	<220>															
[1634]	<223> 合成构建体															
[1635]	<400>16															
[1636]	His	Gly	Glu	Gly	Thr	Phe	Thr	Ser	Asp	Leu	Ser	Lys	Gln	Met	Glu	Glu
[1637]				1			5				10				15	
[1638]	Glu	Ala	Val	Arg	Leu	Phe	Ile	Glu	Trp	Leu	Lys	Asn	Gly	Gly	Pro	Ser
[1639]					20					25					30	
[1640]	Ser	Gly	Ala	Pro	Pro	Pro	Ser	Asp	Ala	His	Lys	Ser	Glu	Val	Ala	His
[1641]				35						40					45	

[1642]	Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile
[1643]	50 55 60
[1644]	Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys
[1645]	65 70 75 80
[1646]	Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu
[1647]	85 90 95
[1648]	Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys
[1649]	100 105 110
[1650]	Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp
[1651]	115 120 125
[1652]	Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His
[1653]	130 135 140
[1654]	Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp
[1655]	145 150 155 160
[1656]	Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys
[1657]	165 170 175
[1658]	Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu
[1659]	180 185 190
[1660]	Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys
[1661]	195 200 205
[1662]	Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu
[1663]	210 215 220
[1664]	Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala
[1665]	225 230 235 240
[1666]	Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala
[1667]	245 250 255
[1668]	Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys
[1669]	260 265 270
[1670]	Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp
[1671]	275 280 285
[1672]	Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys
[1673]	290 295 300
[1674]	Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
[1675]	305 310 315 320
[1676]	Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
[1677]	325 330 335
[1678]	Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
[1679]	340 345 350
[1680]	Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met

[1681]	355	360	365
[1682]	Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu		
[1683]	370	375	380
[1684]	Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys		
[1685]	385	390	395
[1686]	Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe		
[1687]	405	410	415
[1688]	Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu		
[1689]	420	425	430
[1690]	Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val		
[1691]	435	440	445
[1692]	Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu		
[1693]	450	455	460
[1694]	Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro		
[1695]	465	470	475
[1696]	Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu		
[1697]	485	490	495
[1698]	Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val		
[1699]	500	505	510
[1700]	Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser		
[1701]	515	520	525
[1702]	Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu		
[1703]	530	535	540
[1704]	Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg		
[1705]	545	550	555
[1706]	Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro		
[1707]	565	570	575
[1708]	Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala		
[1709]	580	585	590
[1710]	Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala		
[1711]	595	600	605
[1712]	Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu		
[1713]	610	615	620
[1714]	<210>17		
[1715]	<211>640		
[1716]	<212>PRT		
[1717]	<213> 人工序列		
[1718]	<220>		
[1719]	<223> 合成构建体		

[1720] <400>17
 [1721] His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 [1722] 1 5 10 15
 [1723] Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
 [1724] 20 25 30
 [1725] Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly
 [1726] 35 40 45
 [1727] Gly Ser Gly Gly Gly Gly Ser Asp Ala His Lys Ser Glu Val Ala His
 [1728] 50 55 60
 [1729] Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile
 [1730] 65 70 75 80
 [1731] Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys
 [1732] 85 90 95
 [1733] Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu
 [1734] 100 105 110
 [1735] Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys
 [1736] 115 120 125
 [1737] Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp
 [1738] 130 135 140
 [1739] Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His
 [1740] 145 150 155 160
 [1741] Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp
 [1742] 165 170 175
 [1743] Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys
 [1744] 180 185 190
 [1745] Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu
 [1746] 195 200 205
 [1747] Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys
 [1748] 210 215 220
 [1749] Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu
 [1750] 225 230 235 240
 [1751] Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala
 [1752] 245 250 255
 [1753] Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala
 [1754] 260 265 270
 [1755] Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys
 [1756] 275 280 285
 [1757] Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp
 [1758] 290 295 300

[1759]	Leu	Leu	Glu	Cys	Ala	Asp	Asp	Arg	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	Cys
[1760]	305					310					315					320
[1761]	Glu	Asn	Gln	Asp	Ser	Ile	Ser	Ser	Lys	Leu	Lys	Glu	Cys	Cys	Glu	Lys
[1762]					325					330						335
[1763]	Pro	Leu	Leu	Glu	Lys	Ser	His	Cys	Ile	Ala	Glu	Val	Glu	Asn	Asp	Glu
[1764]					340					345						350
[1765]	Met	Pro	Ala	Asp	Leu	Pro	Ser	Leu	Ala	Ala	Asp	Phe	Val	Glu	Ser	Lys
[1766]			355						360				365			
[1767]	Asp	Val	Cys	Lys	Asn	Tyr	Ala	Glu	Ala	Lys	Asp	Val	Phe	Leu	Gly	Met
[1768]	370						375					380				
[1769]	Phe	Leu	Tyr	Glu	Tyr	Ala	Arg	Arg	His	Pro	Asp	Tyr	Ser	Val	Val	Leu
[1770]	385					390					395					400
[1771]	Leu	Leu	Arg	Leu	Ala	Lys	Thr	Tyr	Glu	Thr	Thr	Leu	Glu	Lys	Cys	Cys
[1772]					405					410						415
[1773]	Ala	Ala	Ala	Asp	Pro	His	Glu	Cys	Tyr	Ala	Lys	Val	Phe	Asp	Glu	Phe
[1774]					420					425					430	
[1775]	Lys	Pro	Leu	Val	Glu	Glu	Pro	Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	Glu
[1776]			435						440						445	
[1777]	Leu	Phe	Glu	Gln	Leu	Gly	Glu	Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	Val
[1778]	450								455					460		
[1779]	Arg	Tyr	Thr	Lys	Lys	Val	Pro	Gln	Val	Ser	Thr	Pro	Thr	Leu	Val	Glu
[1780]	465					470						475				480
[1781]	Val	Ser	Arg	Asn	Leu	Gly	Lys	Val	Gly	Ser	Lys	Cys	Cys	Lys	His	Pro
[1782]					485					490						495
[1783]	Glu	Ala	Lys	Arg	Met	Pro	Cys	Ala	Glu	Asp	Tyr	Leu	Ser	Val	Val	Leu
[1784]					500					505						510
[1785]	Asn	Gln	Leu	Cys	Val	Leu	His	Glu	Lys	Thr	Pro	Val	Ser	Asp	Arg	Val
[1786]			515							520						525
[1787]	Thr	Lys	Cys	Cys	Thr	Glu	Ser	Leu	Val	Asn	Arg	Arg	Pro	Cys	Phe	Ser
[1788]	530								535						540	
[1789]	Ala	Leu	Glu	Val	Asp	Glu	Thr	Tyr	Val	Pro	Lys	Glu	Phe	Asn	Ala	Glu
[1790]	545								550				555			560
[1791]	Thr	Phe	Thr	Phe	His	Ala	Asp	Ile	Cys	Thr	Leu	Ser	Glu	Lys	Glu	Arg
[1792]					565						570					575
[1793]	Gln	Ile	Lys	Lys	Gln	Thr	Ala	Leu	Val	Glu	Leu	Val	Lys	His	Lys	Pro
[1794]					580						585					590
[1795]	Lys	Ala	Thr	Lys	Glu	Gln	Leu	Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala
[1796]					595									600		605
[1797]	Phe	Val	Glu	Lys	Cys	Cys	Lys	Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala

[1798]	610	615	620
[1799]	Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu		
[1800]	625	630	635
[1801]	<210>18		640
[1802]	<211>264		
[1803]	<212>PRT		
[1804]	<213> 人工序列		
[1805]	<220>		
[1806]	<223> 合成构建体		
[1807]	<400>18		
[1808]	His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly		
[1809]	1	5	10
[1810]	Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ala		15
[1811]	20	25	30
[1812]	Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala		
[1813]	35	40	45
[1814]	Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro		
[1815]	50	55	60
[1816]	Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val		
[1817]	65	70	75
[1818]	Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val		80
[1819]	85	90	95
[1820]	Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln		
[1821]	100	105	110
[1822]	Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln		
[1823]	115	120	125
[1824]	Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala		
[1825]	130	135	140
[1826]	Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro		
[1827]	145	150	155
[1828]	Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr		160
[1829]	165	170	175
[1830]	Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser		
[1831]	180	185	190
[1832]	Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr		
[1833]	195	200	205
[1834]	Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr		
[1835]	210	215	220
[1836]	Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe		

[1837]	225	230	235	240
[1838]	Ser Cys Ser Val Met His Glu Ala Leu Hi s Asn His Tyr Thr Gln Lys			
[1839]		245	250	255
[1840]	Ser Leu Ser Leu Ser Pro Gly Lys			
[1841]		260		
[1842]	<210>19			
[1843]	<211>272			
[1844]	<212>PRT			
[1845]	<213> 人工序列			
[1846]	<220>			
[1847]	<223> 合成构建体			
[1848]	<400>19			
[1849]	His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly			
[1850]	1	5	10	15
[1851]	Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser			
[1852]		20	25	30
[1853]	Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr			
[1854]		35	40	45
[1855]	His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser			
[1856]		50	55	60
[1857]	Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg			
[1858]	65	70	75	80
[1859]	Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro			
[1860]		85	90	95
[1861]	Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala			
[1862]		100	105	110
[1863]	Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val			
[1864]		115	120	125
[1865]	Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr			
[1866]		130	135	140
[1867]	Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr			
[1868]	145	150	155	160
[1869]	Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu			
[1870]		165	170	175
[1871]	Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys			
[1872]		180	185	190
[1873]	Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser			
[1874]		195	200	205
[1875]	Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp			

[1876]	210	215	220
[1877]	Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser		
[1878]	225	230	235
[1879]	Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala		
[1880]	245	250	255
[1881]	Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys		
[1882]	260	265	270
[1883]	<210>20		
[1884]	<211>264		
[1885]	<212>PRT		
[1886]	<213>人工序列		
[1887]	<220>		
[1888]	<223>合成构建体		
[1889]	<400>20		
[1890]	His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu		
[1891]	1	5	10
[1892]	Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ala		
[1893]	20	25	30
[1894]	Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala		
[1895]	35	40	45
[1896]	Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro		
[1897]	50	55	60
[1898]	Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val		
[1899]	65	70	75
[1900]	Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val		
[1901]	85	90	95
[1902]	Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln		
[1903]	100	105	110
[1904]	Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln		
[1905]	115	120	125
[1906]	Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala		
[1907]	130	135	140
[1908]	Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro		
[1909]	145	150	155
[1910]	Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr		
[1911]	165	170	175
[1912]	Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser		
[1913]	180	185	190
[1914]	Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr		

[1915]	195					200					205					
[1916]	Lys	Thr	Thr	Pro	Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	Phe	Phe	Leu	Tyr
[1917]	210					215					220					
[1918]	Ser	Lys	Leu	Thr	Val	Asp	Lys	Ser	Arg	Trp	Gln	Gln	Gly	Asn	Val	Phe
[1919]	225		230					235					240			
[1920]	Ser	Cys	Ser	Val	Met	His	Glu	Ala	Leu	His	Asn	His	Tyr	Thr	Gln	Lys
[1921]	245					250					255					
[1922]	Ser	Leu	Ser	Leu	Ser	Pro	Gly	Lys								
[1923]	260															
[1924]	<210>21															
[1925]	<211>272															
[1926]	<212>PRT															
[1927]	<213> 人工序列															
[1928]	<220>															
[1929]	<223> 合成构建体															
[1930]	<400>21															
[1931]	His	Val	Glu	Gly	Thr	Phe	Thr	Ser	Asp	Val	Ser	Ser	Tyr	Leu	Glu	Glu
[1932]	1	5				10					15					
[1933]	Gln	Ala	Ala	Lys	Glu	Phe	Ile	Ala	Trp	Leu	Val	Lys	Gly	Arg	Gly	Ser
[1934]	20					25					30					
[1935]	Ser	Gly	Ala	Pro	Pro	Pro	Ser	Ala	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr
[1936]	35				40					45						
[1937]	His	Thr	Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser
[1938]	50			55					60							
[1939]	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg
[1940]	65		70					75					80			
[1941]	Thr	Pro	Glu	Val	Thr	Cys	Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro
[1942]	85					90					95					
[1943]	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala
[1944]	100					105					110					
[1945]	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val
[1946]	115					120					125					
[1947]	Ser	Val	Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr
[1948]	130					135					140					
[1949]	Lys	Cys	Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr
[1950]	145		150					155					160			
[1951]	Ile	Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu
[1952]	165					170					175					
[1953]	Pro	Pro	Ser	Arg	Glu	Glu	Met	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys

[1954]	180	185	190
[1955]	Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser		
[1956]	195	200	205
[1957]	Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp		
[1958]	210	215	220
[1959]	Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser		
[1960]	225	230	235
[1961]	Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala		
[1962]	245	250	255
[1963]	Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys		
[1964]	260	265	270
[1965]	<210>22		
[1966]	<211>272		
[1967]	<212>PRT		
[1968]	<213> 人工序列		
[1969]	<220>		
[1970]	<223> 合成构建体		
[1971]	<400>22		
[1972]	His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu		
[1973]	1	5	10
[1974]	Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser		
[1975]	20	25	30
[1976]	Ser Gly Ala Ser Ser Gly Ala Ala Glu Pro Lys Ser Cys Asp Lys Thr		
[1977]	35	40	45
[1978]	His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser		
[1979]	50	55	60
[1980]	Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg		
[1981]	65	70	75
[1982]	Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro		
[1983]	85	90	95
[1984]	Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala		
[1985]	100	105	110
[1986]	Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val		
[1987]	115	120	125
[1988]	Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr		
[1989]	130	135	140
[1990]	Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr		
[1991]	145	150	155
[1992]	Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu		

[1993]					165				170				175									
[1994]	Pro	Pro	Ser	Arg	Glu	Glu	Met	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys						
[1995]					180				185				190									
[1996]	Leu	Val	Lys	Gly	Phe	Tyr	Pro	Ser	Asp	Ile	Ala	Val	Glu	Trp	Glu	Ser						
[1997]					195				200				205									
[1998]	Asn	Gly	Gln	Pro	Glu	Asn	Asn	Tyr	Lys	Thr	Thr	Pro	Pro	Val	Leu	Asp						
[1999]					210				215				220									
[2000]	Ser	Asp	Gly	Ser	Phe	Phe	Leu	Tyr	Ser	Lys	Leu	Thr	Val	Asp	Lys	Ser						
[2001]	225								230								240					
[2002]	Arg	Trp	Gln	Gln	Gly	Asn	Val	Phe	Ser	Cys	Ser	Val	Met	His	Glu	Ala						
[2003]					245								250				255					
[2004]	Leu	His	Asn	His	Tyr	Thr	Gln	Lys	Ser	Leu	Ser	Leu	Ser	Pro	Gly	Lys						
[2005]					260								265				270					
[2006]	<210>23																					
[2007]	<211>287																					
[2008]	<212>PRT																					
[2009]	<213> 人工序列																					
[2010]	<220>																					
[2011]	<223> 合成构建体																					
[2012]	<400>23																					
[2013]	His	Gly	Glu	Gly	Thr	Phe	Thr	Ser	Asp	Val	Ser	Ser	Tyr	Leu	Glu	Glu						
[2014]	1								5								10		15			
[2015]	Gln	Ala	Ala	Lys	Glu	Phe	Ile	Ala	Trp	Leu	Val	Lys	Gly	Arg	Gly	Ser						
[2016]					20								25				30					
[2017]	Ser	Gly	Ala	Pro	Pro	Pro	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly						
[2018]					35								40				45					
[2019]	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His						
[2020]	50								55								60					
[2021]	Thr	Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val						
[2022]	65								70								75				80	
[2023]	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr						
[2024]					85								90				95					
[2025]	Pro	Glu	Val	Thr	Cys	Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu						
[2026]					100								105				110					
[2027]	Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys						
[2028]					115								120				125					
[2029]	Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser						
[2030]	130								135								140					
[2031]	Val	Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys						

[2032]	145	150	155	160
[2033]	Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile			
[2034]	165	170	175	
[2035]	Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro			
[2036]	180	185	190	
[2037]	Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu			
[2038]	195	200	205	
[2039]	Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn			
[2040]	210	215	220	
[2041]	Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser			
[2042]	225	230	235	240
[2043]	Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg			
[2044]	245	250	255	
[2045]	Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu			
[2046]	260	265	270	
[2047]	His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys			
[2048]	275	280	285	
[2049]	<210>24			
[2050]	<211>284			
[2051]	<212>PRT			
[2052]	<213> 人工序列			
[2053]	<220>			
[2054]	<223> 合成构建体			
[2055]	<400>24			
[2056]	His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu			
[2057]	1	5	10	15
[2058]	Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser			
[2059]	20	25	30	
[2060]	Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly			
[2061]	35	40	45	
[2062]	Ser Gly Gly Gly Gly Ser Ala Glu Ser Lys Tyr Gly Pro Pro Cys Pro			
[2063]	50	55	60	
[2064]	Ser Cys Pro Ala Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe			
[2065]	65	70	75	80
[2066]	Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val			
[2067]	85	90	95	
[2068]	Thr Cys Val Val Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe			
[2069]	100	105	110	
[2070]	Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro			

[2071]	115	120	125
[2072]	Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr		
[2073]	130	135	140
[2074]	Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val		
[2075]	145	150	155
[2076]	Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala		
[2077]	165	170	175
[2078]	Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln		
[2079]	180	185	190
[2080]	Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly		
[2081]	195	200	205
[2082]	Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro		
[2083]	210	215	220
[2084]	Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser		
[2085]	225	230	235
[2086]	Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu		
[2087]	245	250	255
[2088]	Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His		
[2089]	260	265	270
[2090]	Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly Lys		
[2091]	275	280	
[2092]	<210>25		
[2093]	<211>302		
[2094]	<212>PRT		
[2095]	<213> 人工序列		
[2096]	<220>		
[2097]	<223> 合成构建体		
[2098]	<400>25		
[2099]	His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu		
[2100]	1	5	10
[2101]	Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser		15
[2102]	20	25	30
[2103]	Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly		
[2104]	35	40	45
[2105]	Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser		
[2106]	50	55	60
[2107]	Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr		
[2108]	65	70	75
[2109]	Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe		80

[2110]	85	90	95
[2111]	Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro		
[2112]	100	105	110
[2113]	Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val		
[2114]	115	120	125
[2115]	Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr		
[2116]	130	135	140
[2117]	Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val		
[2118]	145	150	155
[2119]	Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys		
[2120]	165	170	175
[2121]	Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser		
[2122]	180	185	190
[2123]	Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro		
[2124]	195	200	205
[2125]	Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val		
[2126]	210	215	220
[2127]	Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly		
[2128]	225	230	235
[2129]	Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp		
[2130]	245	250	255
[2131]	Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp		
[2132]	260	265	270
[2133]	Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His		
[2134]	275	280	285
[2135]	Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys		
[2136]	290	295	300
[2137]	<210>26		
[2138]	<211>294		
[2139]	<212>PRT		
[2140]	<213> 人工序列		
[2141]	<220>		
[2142]	<223> 合成构建体		
[2143]	<400>26		
[2144]	His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu		
[2145]	1	5	10
[2146]	Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Gly		
[2147]	20	25	30
[2148]	Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly		

[2149]	35	40	45
[2150]	Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Glu Pro		
[2151]	50	55	60
[2152]	Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu		
[2153]	65	70	75
[2154]	80		
[2154]	Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp		
[2155]	85	90	95
[2156]	Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp		
[2157]	100	105	110
[2158]	Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly		
[2159]	115	120	125
[2160]	Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn		
[2161]	130	135	140
[2162]	Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp		
[2163]	145	150	155
[2164]	160		
[2164]	Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro		
[2165]	165	170	175
[2166]	Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu		
[2167]	180	185	190
[2168]	Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn		
[2169]	195	200	205
[2170]	Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile		
[2171]	210	215	220
[2172]	Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr		
[2173]	225	230	235
[2174]	240		
[2174]	Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys		
[2175]	245	250	255
[2176]	Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys		
[2177]	260	265	270
[2178]	Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu		
[2179]	275	280	285
[2180]	Ser Leu Ser Pro Gly Lys		
[2181]	290		
[2182]	<210>27		
[2183]	<211>280		
[2184]	<212>PRT		
[2185]	<213> 人工序列		
[2186]	<220>		
[2187]	<223> 合成构建体		

[2188] <400>27
 [2189] His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
 [2190] 1 5 10 15
 [2191] Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
 [2192] 20 25 30
 [2193] Ser Gly Ala Pro Pro Pro Ser Ser Ser Gly Ala Pro Pro Pro Ser Ala
 [2194] 35 40 45
 [2195] Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
 [2196] 50 55 60
 [2197] Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
 [2198] 65 70 75 80
 [2199] Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
 [2200] 85 90 95
 [2201] Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
 [2202] 100 105 110
 [2203] Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
 [2204] 115 120 125
 [2205] Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr ValLeu His Gln
 [2206] 130 135 140
 [2207] Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
 [2208] 145 150 155 160
 [2209] Leu Pro Ala ProIle Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
 [2210] 165 170 175
 [2211] Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
 [2212] 180 185 190
 [2213] Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
 [2214] 195 200 205
 [2215] Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
 [2216] 210 215 220
 [2217] Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
 [2218] 225 230 235 240
 [2219] Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
 [2220] 245 250 255
 [2221] Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
 [2222] 260 265 270
 [2223] Ser Leu Ser Leu Ser Pro Gly Lys
 [2224] 275 280
 [2225] <210>28
 [2226] <211>287

[2227]	<212>PRT
[2228]	<213> 人工序列
[2229]	<220>
[2230]	<223> 合成构建体
[2231]	<400>28
[2232]	His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
[2233]	1 5 10 15
[2234]	Gln Ala Val Lys Glu Phe Ile Ala Trp Leu Ile Lys Gly Arg Gly Ser
[2235]	20 25 30
[2236]	Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
[2237]	35 40 45
[2238]	Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
[2239]	50 55 60
[2240]	Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
[2241]	65 70 75 80
[2242]	Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
[2243]	85 90 95
[2244]	Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
[2245]	100 105 110
[2246]	Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
[2247]	115 120 125
[2248]	Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
[2249]	130 135 140
[2250]	Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
[2251]	145 150 155 160
[2252]	Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
[2253]	165 170 175
[2254]	Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
[2255]	180 185 190
[2256]	Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
[2257]	195 200 205
[2258]	Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
[2259]	210 215 220
[2260]	Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
[2261]	225 230 235 240
[2262]	Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
[2263]	245 250 255
[2264]	Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
[2265]	260 265 270

[2266]	His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
[2267]	275 280 285
[2268]	<210>29
[2269]	<211>272
[2270]	<212>PRT
[2271]	<213> 人工序列
[2272]	<220>
[2273]	<223> 合成构建体
[2274]	<400>29
[2275]	His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
[2276]	1 5 10 15
[2277]	Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
[2278]	20 25 30
[2279]	Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
[2280]	35 40 45
[2281]	His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
[2282]	50 55 60
[2283]	Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
[2284]	65 70 75 80
[2285]	Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
[2286]	85 90 95
[2287]	Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
[2288]	100 105 110
[2289]	Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
[2290]	115 120 125
[2291]	Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
[2292]	130 135 140
[2293]	Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
[2294]	145 150 155 160
[2295]	Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
[2296]	165 170 175
[2297]	Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
[2298]	180 185 190
[2299]	Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
[2300]	195 200 205
[2301]	Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
[2302]	210 215 220
[2303]	Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
[2304]	225 230 235 240

[2305]	Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
[2306]	245 250 255
[2307]	Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
[2308]	260 265 270
[2309]	<210>30
[2310]	<211>272
[2311]	<212>PRT
[2312]	<213> 人工序列
[2313]	<220>
[2314]	<223> 合成构建体
[2315]	<400>30
[2316]	His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
[2317]	1 5 10 15
[2318]	Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
[2319]	20 25 30
[2320]	Ser Gly Ala Ser Ser Gly Ala Ala Glu Pro Lys Ser Cys Asp Lys Thr
[2321]	35 40 45
[2322]	His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
[2323]	50 55 60
[2324]	Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
[2325]	65 70 75 80
[2326]	Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
[2327]	85 90 95
[2328]	Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
[2329]	100 105 110
[2330]	Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
[2331]	115 120 125
[2332]	Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
[2333]	130 135 140
[2334]	Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
[2335]	145 150 155 160
[2336]	Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
[2337]	165 170 175
[2338]	Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
[2339]	180 185 190
[2340]	Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
[2341]	195 200 205
[2342]	Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
[2343]	210 215 220

[2344]	Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
[2345]	225 230 235 240
[2346]	Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
[2347]	245 250 255
[2348]	Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
[2349]	260 265 270
[2350]	<210>31
[2351]	<211>287
[2352]	<212>PRT
[2353]	<213> 人工序列
[2354]	<220>
[2355]	<223> 合成构建体
[2356]	<400>31
[2357]	His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
[2358]	1 5 10 15
[2359]	Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
[2360]	20 25 30
[2361]	Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
[2362]	35 40 45
[2363]	Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
[2364]	50 55 60
[2365]	Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
[2366]	65 70 75 80
[2367]	Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
[2368]	85 90 95
[2369]	Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
[2370]	100 105 110
[2371]	Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
[2372]	115 120 125
[2373]	Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
[2374]	130 135 140
[2375]	Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
[2376]	145 150 155 160
[2377]	Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
[2378]	165 170 175
[2379]	Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
[2380]	180 185 190
[2381]	Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
[2382]	195 200 205

[2383] Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
 [2384] 210 215 220
 [2385] Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
 [2386] 225 230 235 240
 [2387] Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
 [2388] 245 250 255
 [2389] Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
 [2390] 260 265 270
 [2391] His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 [2392] 275 280 285
 [2393] <210>32
 [2394] <211>232
 [2395] <212>PRT
 [2396] <213>人 (Homo sapiens)
 [2397] <400>32
 [2398] Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
 [2399] 1 5 10 15
 [2400] Ala Pro Glu Lys Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
 [2401] 20 25 30
 [2402] Lys Asp Thr Lys Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
 [2403] 35 40 45
 [2404] Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
 [2405] 50 55 60
 [2406] Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
 [2407] 65 70 75 80
 [2408] Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
 [2409] 85 90 95
 [2410] Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
 [2411] 100 105 110
 [2412] Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
 [2413] 115 120 125
 [2414] Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
 [2415] 130 135 140
 [2416] Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
 [2417] 145 150 155 160
 [2418] Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
 [2419] 165 170 175
 [2420] Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
 [2421] 180 185 190

[2422]	Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe	
[2423]	195	200 205
[2424]	Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys	
[2425]	210	215 220
[2426]	Ser Leu Ser Leu Ser Pro Gly Lys	
[2427]	225	230
[2428]	<210>33	
[2429]	<211>703	
[2430]	<212>DNA	
[2431]	<213> 人 (Homo sapiens)	
[2432]	<400>33	
[2433]	gagcccaaat cttgtgacaa aactcacaca tgcccaccgt gccagcacc tgaactcctg	60
[2434]	gggggaccgt cagtcttcct ctcccccca aaaccaagg acaccctcat gatctcccgg	120
[2435]	acccttgagg tcacatgcgt ggtggtggac gtgagccacg aagaccctga ggtcaagtgc	180
[2436]	aactggtacg tggacggcgt ggaggtgcat aatgccaaga caaagccgcg ggaggagcag	240
[2437]	tacaacagca cgtaccgtgt ggtcagcgtc ctaccgtcc tgcaccagga ctggctgaat	300
[2438]	ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc cagcccccat cgagaaaacc	360
[2439]	atctccaaag ccaaagggca gccccgagaa ccacaggtgt acaccctgcc cccatcccgg	420
[2440]	gaggagatga ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatcccagc	480
[2441]	gacatcgccg tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct	540
[2442]	cccgctgctgg actccgacgg ctcccttcttc ctctatagca agtcaccgt ggacaagagc	600
[2443]	aggtggcagc aggggaacgt cttctcatgc tccgtgatgc atgaggctct gcacaaccac	660
[2444]	tacacgcaga agagcctctc cctgtctccg ggtaaattgat agt	703
[2445]	<210>34	
[2446]	<211>585	
[2447]	<212>PRT	
[2448]	<213> 人 (Homo sapiens)	
[2449]	<400>34	
[2450]	Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu	
[2451]	1	5 10 15
[2452]	Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln	
[2453]	20	25 30
[2454]	Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu	
[2455]	35	40 45
[2456]	Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys	
[2457]	50	55 60
[2458]	Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu	
[2459]	65	70 75 80
[2460]	Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro	

[2461]		85		90		95
[2462]	Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu					
[2463]		100		105		110
[2464]	Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His					
[2465]		115		120		125
[2466]	Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg					
[2467]		130		135		140
[2468]	Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg					
[2469]	145		150		155	160
[2470]	Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala					
[2471]		165		170		175
[2472]	Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser					
[2473]		180		185		190
[2474]	Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu					
[2475]		195		200		205
[2476]	Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro					
[2477]		210		215		220
[2478]	Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys					
[2479]	225		230		235	240
[2480]	Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp					
[2481]		245		250		255
[2482]	Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser					
[2483]		260		265		270
[2484]	Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His					
[2485]		275		280		285
[2486]	Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser					
[2487]		290		295		300
[2488]	Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala					
[2489]	305		310		315	320
[2490]	Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg					
[2491]		325		330		335
[2492]	Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr					
[2493]		340		345		350
[2494]	Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu					
[2495]		355		360		365
[2496]	Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro					
[2497]		370		375		380
[2498]	Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Asn Leu Gly Glu					
[2499]	385		390		395	400

[2500]	Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro	
[2501]	405	410 415
[2502]	Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys	
[2503]	420	425 430
[2504]	Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys	
[2505]	435	440 445
[2506]	Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His	
[2507]	450	455 460
[2508]	Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser	
[2509]	465	470 475 480
[2510]	Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr	
[2511]	485	490 495
[2512]	Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp	
[2513]	500	505 510
[2514]	Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala	
[2515]	515	520 525
[2516]	Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu	
[2517]	530	535 540
[2518]	Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys	
[2519]	545	550 555 560
[2520]	Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val	
[2521]	565	570 575
[2522]	Ala Ala Ser Gln Ala Ala Leu Gly Leu	
[2523]	580	585
[2524]	<210>35	
[2525]	<211>1762	
[2526]	<212>DNA	
[2527]	<213> 人 (Homo sapiens)	
[2528]	<400>35	
[2529]	gatgcgcaca agagtgaggt tgctcatcgg tttaaagatt tgggagaaga aaattttcaaa	60
[2530]	gccttggtgt tgattgcctt tgctcagtat cttcagcagt gtccatttga agatcatgta	120
[2531]	aaattagtga atgaagtaac tgaatttgca aaaacatgtg ttgctgatga gtcagctgaa	180
[2532]	aattgtgaca aatcacttca tacccttttt ggagacaaat tatgcacagt tgcaactctt	240
[2533]	cgtgaaacct atggtgaaat ggctgactgc tgtgcaaaac aagaacctga gagaaatgaa	300
[2534]	tgcttcttgc aacacaaaga tgacaacca aacctcccc gattgggtgag accagaggtt	360
[2535]	gatgtgatgt gcaactgctt tcatgacaat gaagagacat ttttgaaaaa atacttatat	420
[2536]	gaaattgcca gaagacatcc ttacttttat gccccggaac tccttttctt tgctaaaagg	480
[2537]	tataaagctg cttttacaga atgttgccaa gctgctgata aagctgcctg cctgttgcca	540
[2538]	aagctcgatg aacttcggga tgaagggaag gcttcgtctg ccaaacagag actcaagtgt	600

[2539]	gccagtctcc aaaaatttgg agaaagagct ttcaaagcat gggcagtagc tcgcctgagc	660
[2540]	cagagatttc ccaaagctga gtttgcagaa gtttccaagt tagtgacaga tcttaccaaa	720
[2541]	gtccacacgg aatgctgcca tggagatctg cttgaatgtg ctgatgacag ggcggacctt	780
[2542]	gccaaagtata tctgtgaaaa tcaagattcg atctccagta aactgaagga atgctgtgaa	840
[2543]	aaacctctgt tggaaaaatc ccactgcatt gccgaagtgg aaaatgatga gatgcctgct	900
[2544]	gacttgcctt cattagctgc tgattttgtt gaaagtaagg atgtttgcaa aaactatgct	960
[2545]	gaggcaaagg atgtcttcct gggcatgttt ttgtatgaat atgcaagaag gcacccctgat	1020
[2546]	tactctgtcg tgctgctgct gagacttgcc aagacatatg aaaccactct agagaagtgc	1080
[2547]	tgtgccgctg cagatccctca tgaatgctat gccaaagtgt tcgatgaatt taaacctctt	1140
[2548]	gtggaagagc ctcagaattt aatcaaacaa aattgtgagc tttttgagca gcttggagag	1200
[2549]	tacaaattcc agaatgcgct attagtctctg tacaccaaga aagtaccca agtgtcaact	1260
[2550]	ccaactcttg tagaggtctc aagaaacctt ggaaaagtgg gcagcaaattg ttgtaaact	1320
[2551]	cctgaagcaa aaagaatgcc ctgtgcagaa gactatctat ccgtggctct gaaccagtta	1380
[2552]	tgtgtgttgc atgagaaaac gccagtaagt gacagagtca ccaaattgctg cacagaatcc	1440
[2553]	ttggtgaaca ggcgaccatg cttttcagct ctggaagtcg atgaaacata cgttcccaaa	1500
[2554]	gagtttaatg ctgaaacatt caccttccat gcagatatat gcacactttc tgagaaggag	1560
[2555]	agacaaatca agaaacaaac tgcacttggt gagctcgtga aacacaagcc caaggcaaca	1620
[2556]	aaagagcaac tgaaagctgt tatggatgat ttcgcagctt ttgtagagaa gtgctgcaag	1680
[2557]	gctgacgata aggagacctg ctttgccgag gagggtaaaa aacttggtgc tgcaagtcaa	1740
[2558]	gctgccttag gcttataatg ac	1762

	5		10		15
Ala	Glu	Pro	Lys	Ser	Cys
	20		25		30
Pro	Ala	Pro	Glu	Lys	Gly
	35		40		45
Lys	Pro	Lys	Asp	Thr	Lys
	50		55		60
Cys	Val	Val	Val	Asp	Val
	65		70		75
Asn	Trp	Tyr	Val	Asp	Gly
	80		85		90
Pro	Arg	Glu	Glu	Gln	Tyr
	95		100		105
Leu	Thr	Val	Leu	His	Gln
	110		115		120
Cys	Lys	Val	Ser	Asn	Lys
	125		130		135
Ile	Ser	Lys	Ala	Lys	Gly
	140		145		150
Leu	Pro	Pro	Ser	Arg	Glu
	155		160		165
Thr	Cys	Leu	Val	Lys	Gly
	170		175		180
Trp	Glu	Ser	Asn	Gly	Gln
	185		190		195
Pro	Val	Leu	Asp	Ser	Asp
	200		205		210
Thr	Val	Asp	Lys	Ser	Arg
	215		220		225
Ser	Val	Met	His	Glu	Ala
	230				
Leu	Ser	Leu	Ser	Pro	Gly

[SEQ ID NO: 32]

图 1

				5					10					15					
Asp	Ala	His	Lys	Ser	Glu	Val	Ala	His	Arg	Phe	Lys	Asp	Leu	Gly	Glu	Glu	Asn		
20						25					30					35			
Phe	Lys	Ala	Leu	Val	Leu	Ile	Ala	Phe	Ala	Gln	Tyr	Leu	Gln	Gln	Cys	Pro	Phe		
			40				45						50						
Glu	Asp	His	Val	Lys	Leu	Val	Asn	Glu	Val	Thr	Glu	Phe	Ala	Lys	Thr	Cys	Val		
55					60				65					70					
Ala	Asp	Glu	Ser	Ala	Glu	Asn	Cys	Asp	Lys	Ser	Leu	His	Thr	Leu	Phe	Gly	Asp		
		75				80						85				90			
Lys	Leu	Cys	Thr	Val	Ala	Thr	Leu	Arg	Glu	Thr	Tyr	Gly	Glu	Met	Ala	Asp	Cys		
			95					100					105						
Cys	Ala	Lys	Gln	Glu	Pro	Glu	Arg	Asn	Glu	Cys	Phe	Leu	Gln	His	Lys	Asp	Asp		
110					115						120					125			
Asn	Pro	Asn	Leu	Pro	Arg	Leu	Val	Arg	Pro	Glu	Val	Asp	Val	Met	Cys	Thr	Ala		
			130				135					140							
Phe	His	Asp	Asn	Glu	Glu	Thr	Phe	Leu	Lys	Lys	Tyr	Leu	Tyr	Glu	Ile	Ala	Arg		
145				150						155					160				
Arg	His	Pro	Tyr	Phe	Tyr	Ala	Pro	Glu	Leu	Leu	Phe	Phe	Ala	Lys	Arg	Tyr	Lys		
		165				170					175				180				
Ala	Ala	Phe	Thr	Glu	Cys	Cys	Gln	Ala	Ala	Asp	Lys	Ala	Ala	Cys	Leu	Leu	Pro		
			185					190						195					
Lys	Leu	Asp	Glu	Leu	Arg	Asp	Glu	Gly	Lys	Ala	Ser	Ser	Ala	Lys	Gln	Arg	Leu		
200					205						210				215				
Lys	Cys	Ala	Ser	Leu	Gln	Lys	Phe	Gly	Glu	Arg	Ala	Phe	Lys	Ala	Trp	Ala	Val		
			220				225						230						
Ala	Arg	Leu	Ser	Gln	Arg	Phe	Pro	Lys	Ala	Glu	Phe	Ala	Glu	Val	Ser	Lys	Leu		
235				240						245					250				
Val	Thr	Asp	Leu	Thr	Lys	Val	His	Thr	Glu	Cys	Cys	His	Gly	Asp	Leu	Leu	Glu		
		255				260						265					270		
Cys	Ala	Asp	Asp	Arg	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	Cys	Glu	Asn	Gln	Asp	Ser		
			275						280					285					
Ile	Ser	Ser	Lys	Leu	Lys	Glu	Cys	Cys	Glu	Lys	Pro	Leu	Leu	Glu	Lys	Ser	His		
290				295							300				305				
Cys	Ile	Ala	Glu	Val	Glu	Asn	Asp	Glu	Met	Pro	Ala	Asp	Leu	Pro	Ser	Leu	Ala		
			310				315						320						
Ala	Asp	Phe	Val	Glu	Ser	Lys	Asp	Val	Cys	Lys	Asn	Tyr	Ala	Glu	Ala	Lys	Asp		
325				330						335					340				
Val	Phe	Leu	Gly	Met	Phe	Leu	Tyr	Glu	Tyr	Ala	Arg	Arg	His	Pro	Asp	Tyr	Ser		
		345				350					355					360			
Val	Val	Leu	Leu	Leu	Arg	Leu	Ala	Lys	Thr	Tyr	Glu	Thr	Thr	Leu	Glu	Lys	Cys		
			365					370						375					
Cys	Ala	Ala	Ala	Asp	Pro	His	Glu	Cys	Tyr	Ala	Lys	Val	Phe	Asp	Glu	Phe	Lys		
380				385							390				395				
Pro	Leu	Val	Glu	Glu	Pro	Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	Glu	Leu	Phe	Glu		
			400				405						410						
Asn	Leu	Gly	Glu	Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	Val	Arg	Tyr	Thr	Lys	Lys		
415				420						425					430				
Val	Pro	Gln	Val	Ser	Thr	Pro	Thr	Leu	Val	Glu	Val	Ser	Arg	Asn	Leu	Gly	Lys		
		435				440					445				450				
Val	Gly	Ser	Lys	Cys	Cys	Lys	His	Pro	Glu	Ala	Lys	Arg	Met	Pro	Cys	Ala	Glu		
			455					460						465					
Asp	Tyr	Leu	Ser	Val	Val	Leu	Asn	Gln	Leu	Cys	Val	Leu	His	Glu	Lys	Thr	Pro		
470				475							480				485				
Val	Ser	Asp	Arg	Val	Thr	Lys	Cys	Cys	Thr	Glu	Ser	Leu	Val	Asn	Arg	Arg	Pro		

图 2

```

          490          495          500
Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala
505          510          515          520
Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln
          525          530          535          540
Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr
          545          550          555
Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys
560          565          570          575
Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
          580          585
Ala Ala Ser Gln Ala Ala Leu Gly Leu

```

[SEQ ID NO: 34]

图 2 续

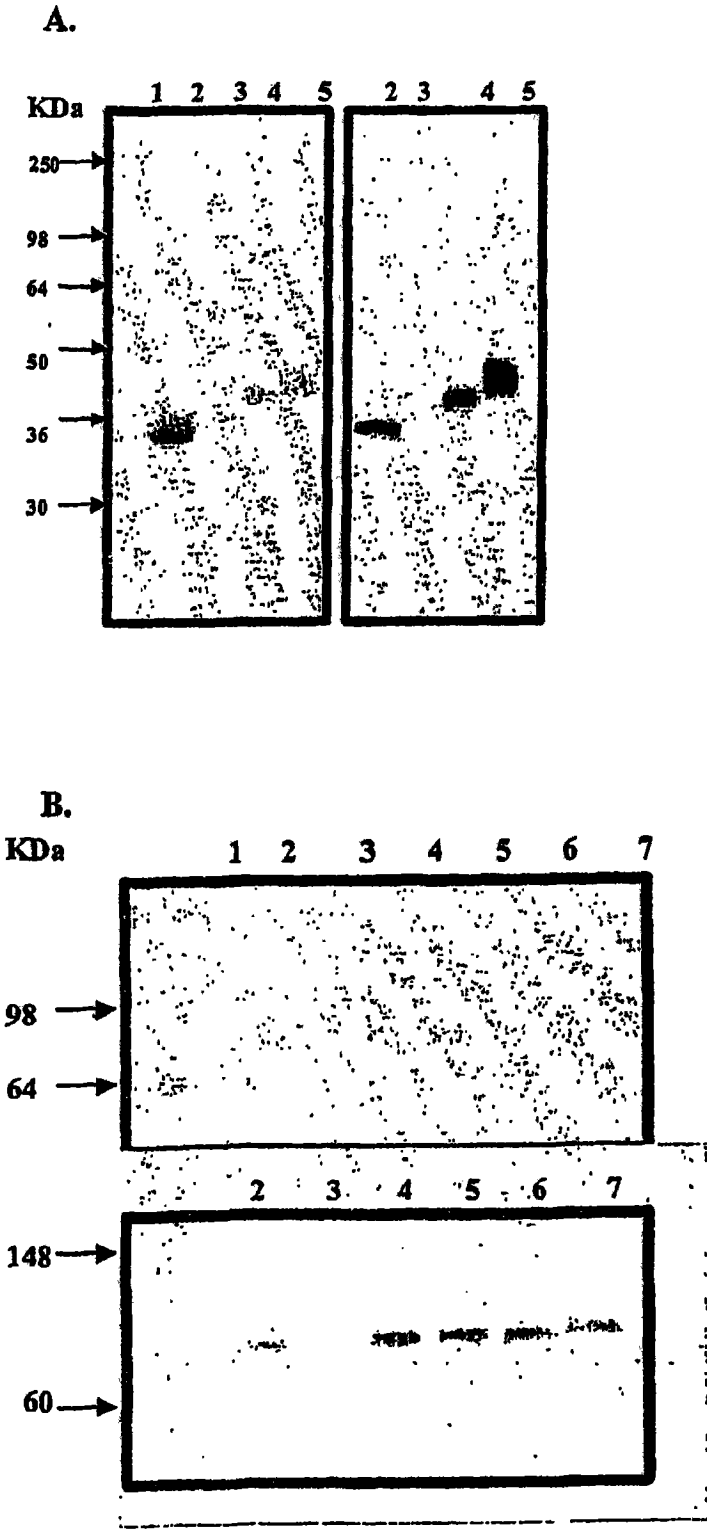


图 3

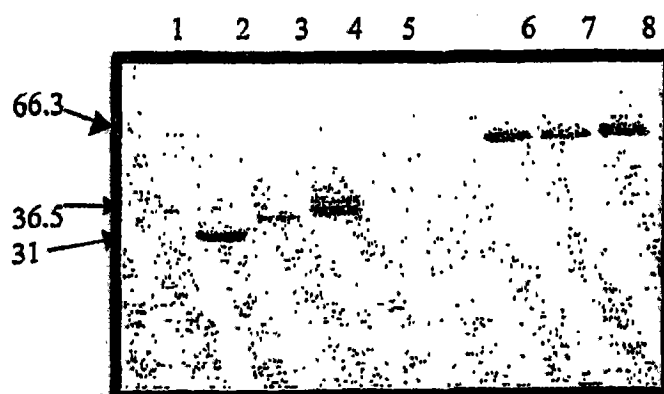


图 4

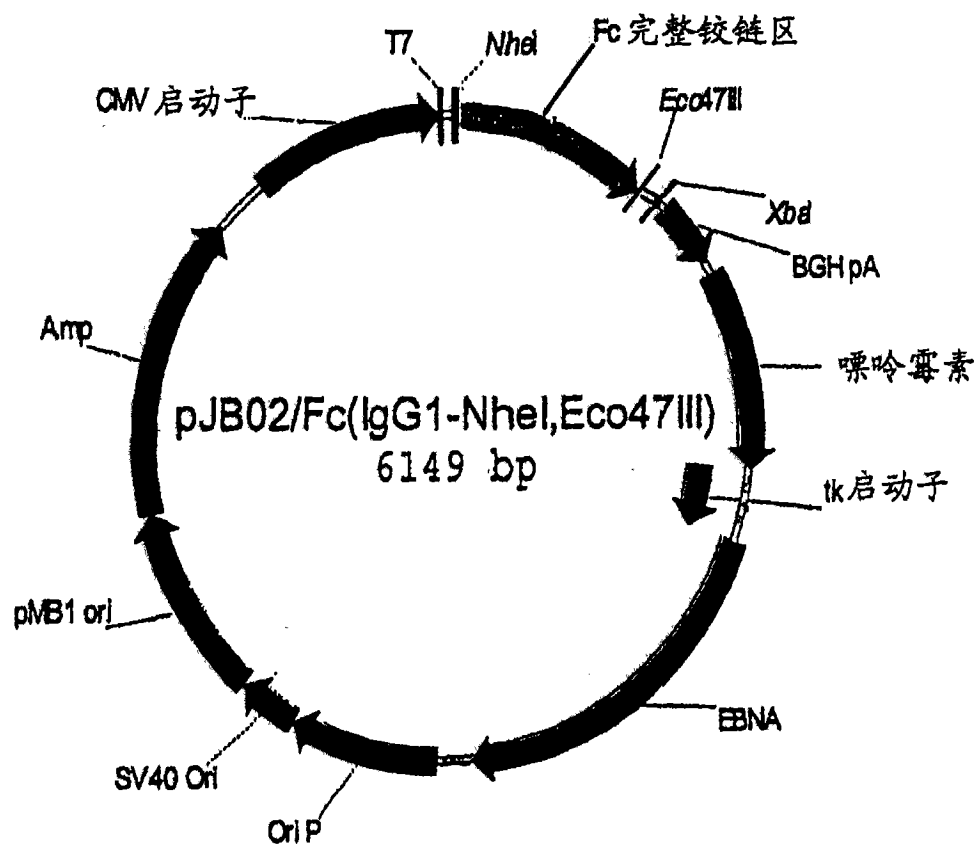


图 5

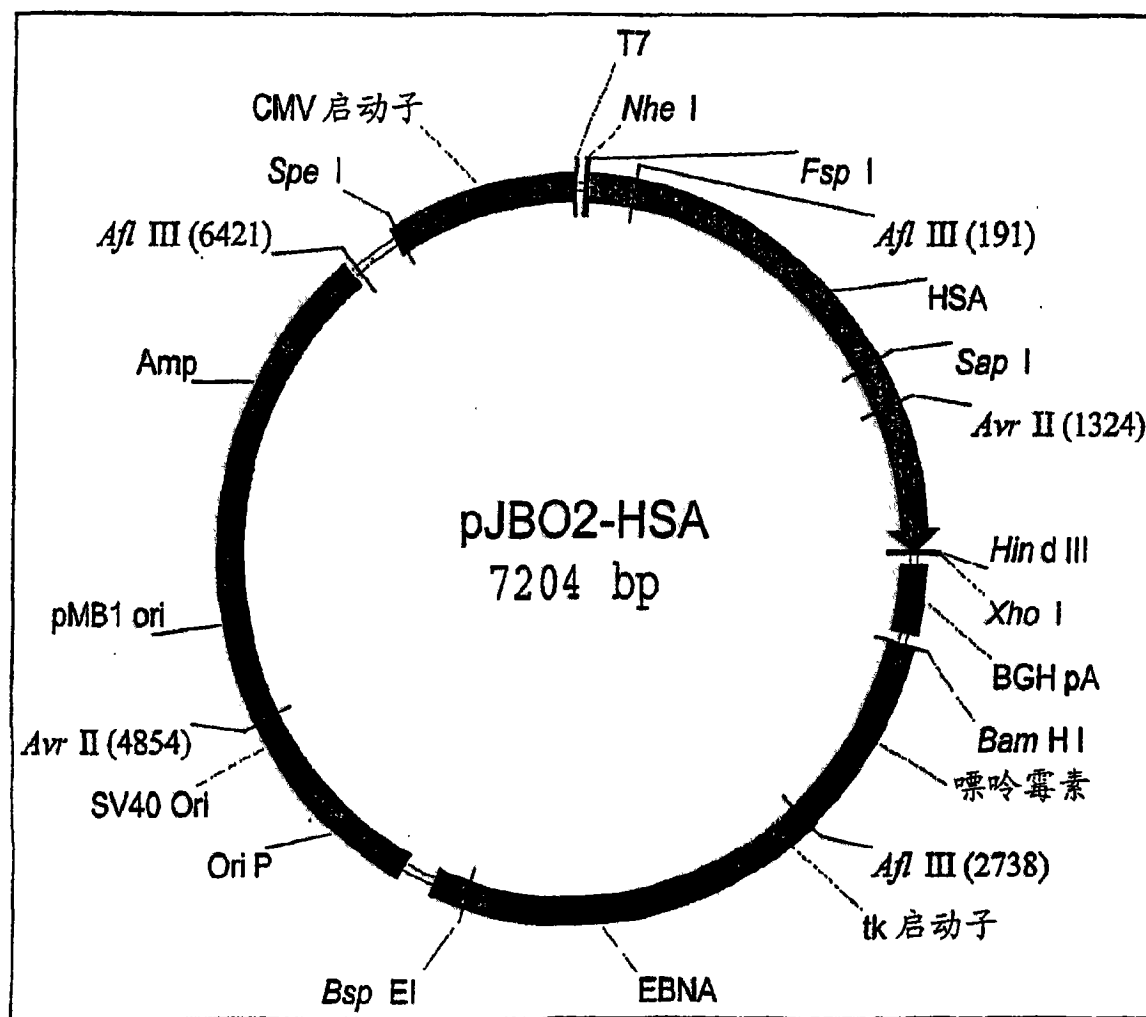


图 6

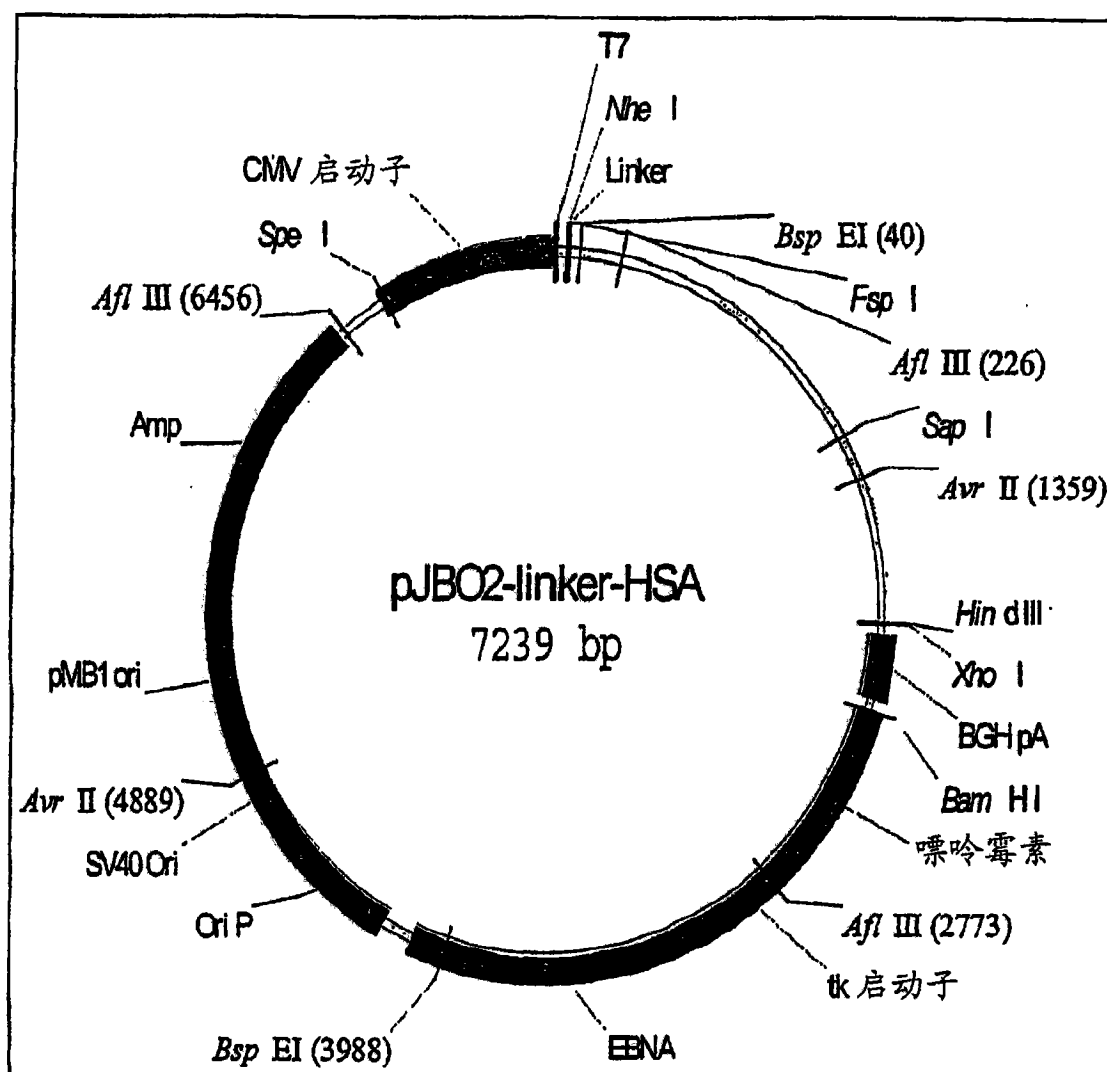


图 7

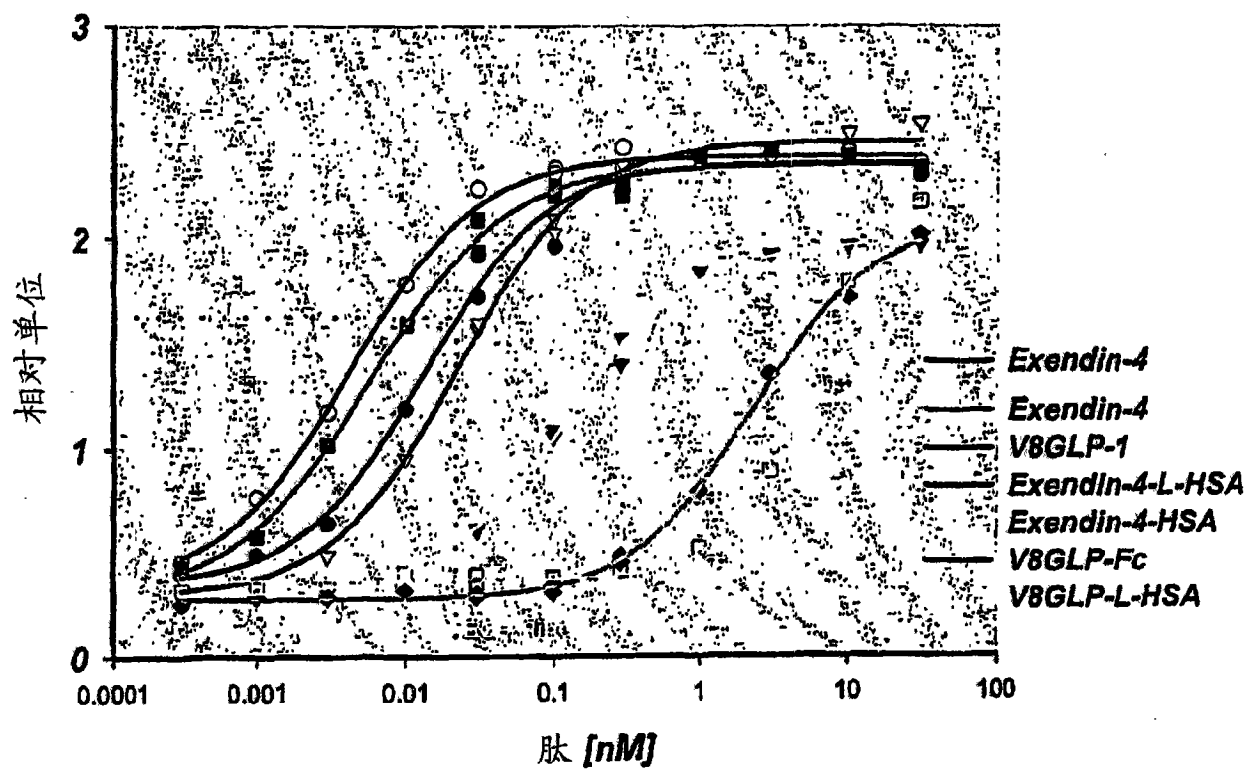


图 8

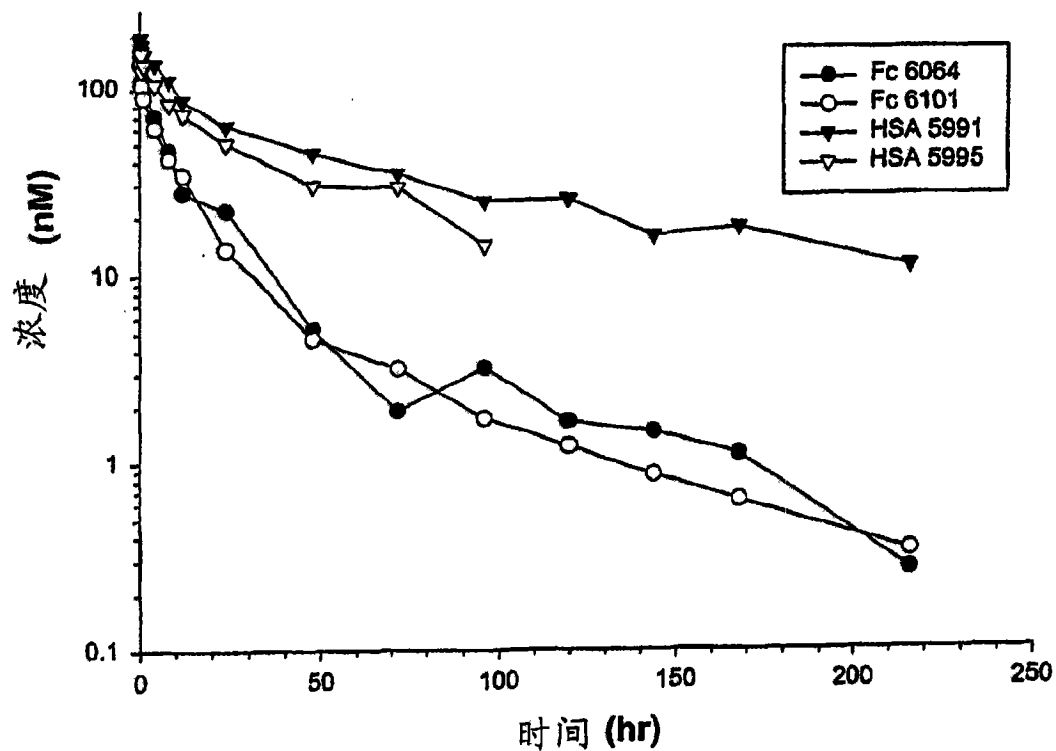


图 9

A.

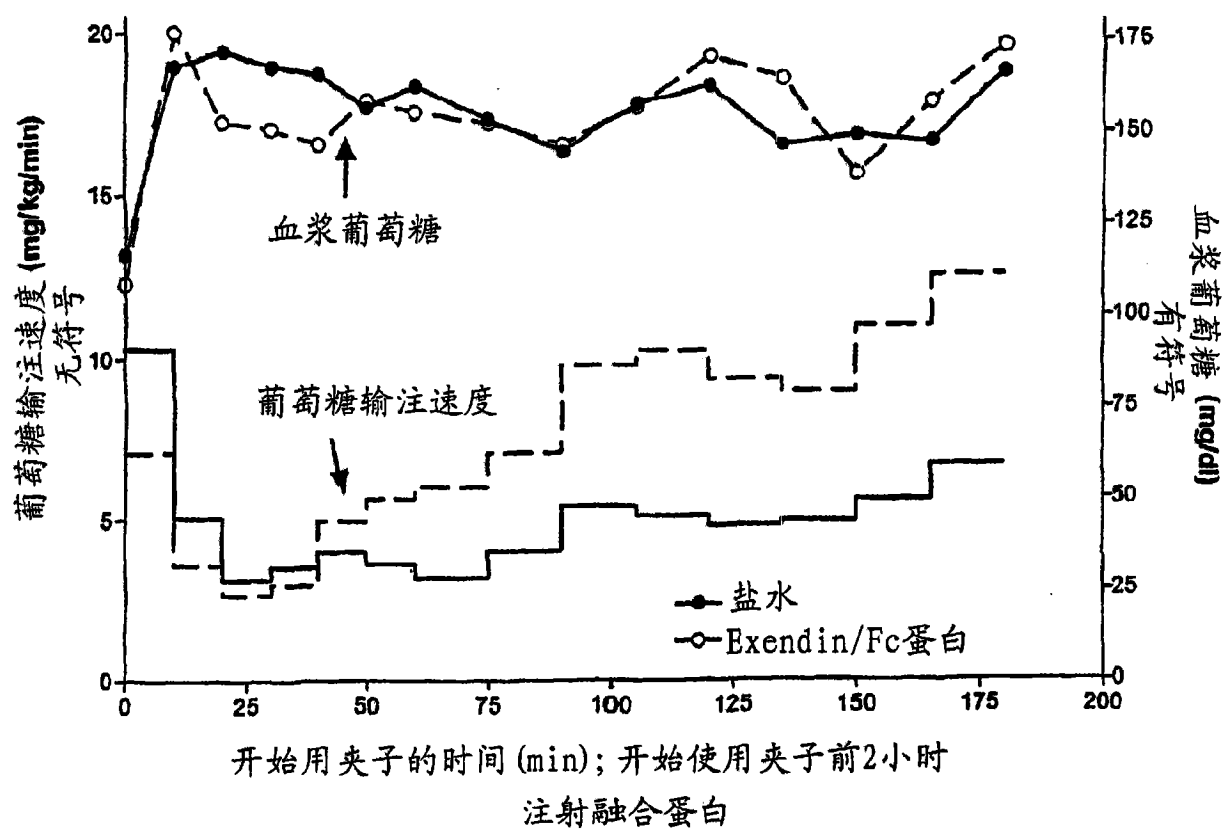


图 10

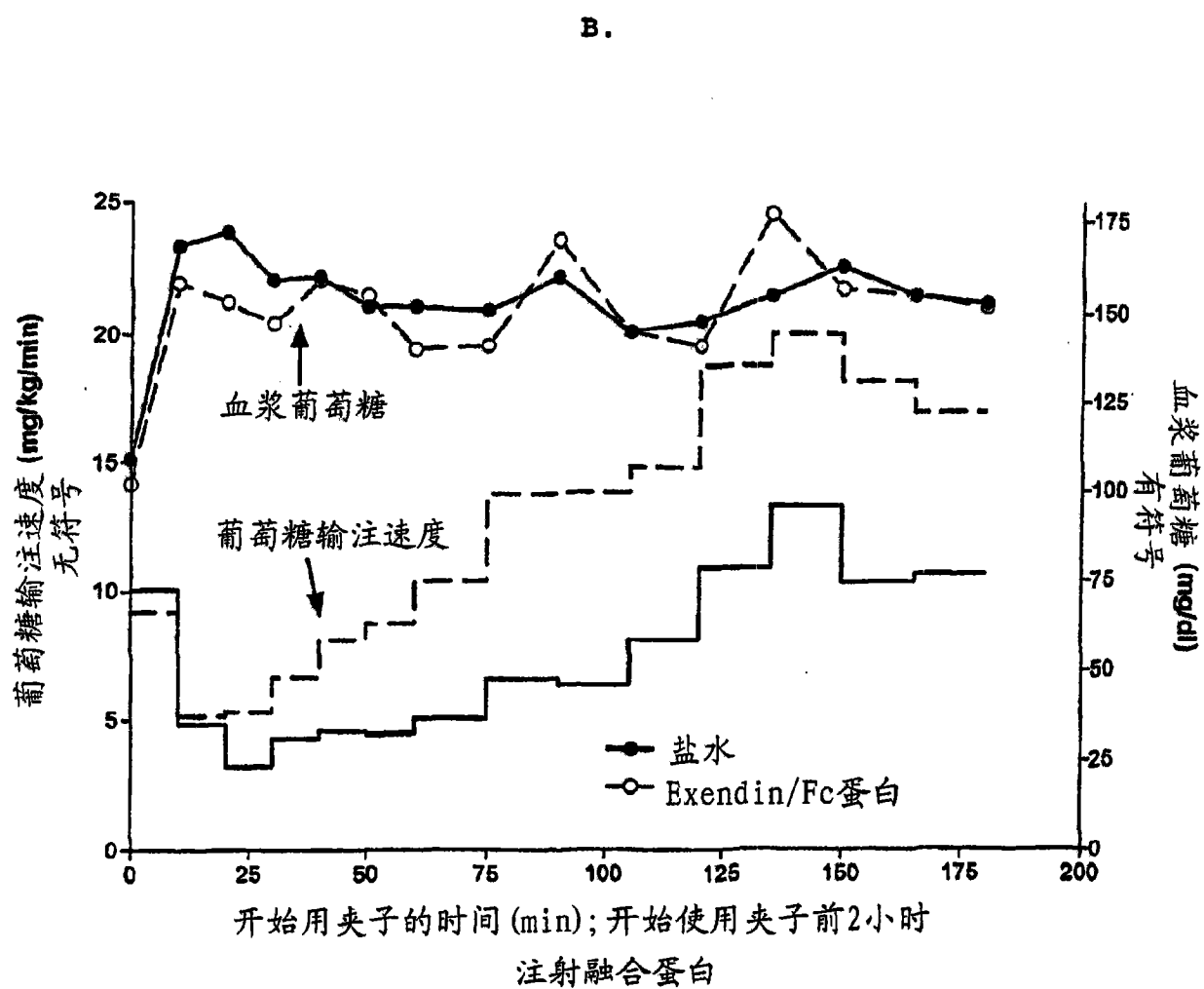


图 10

A.

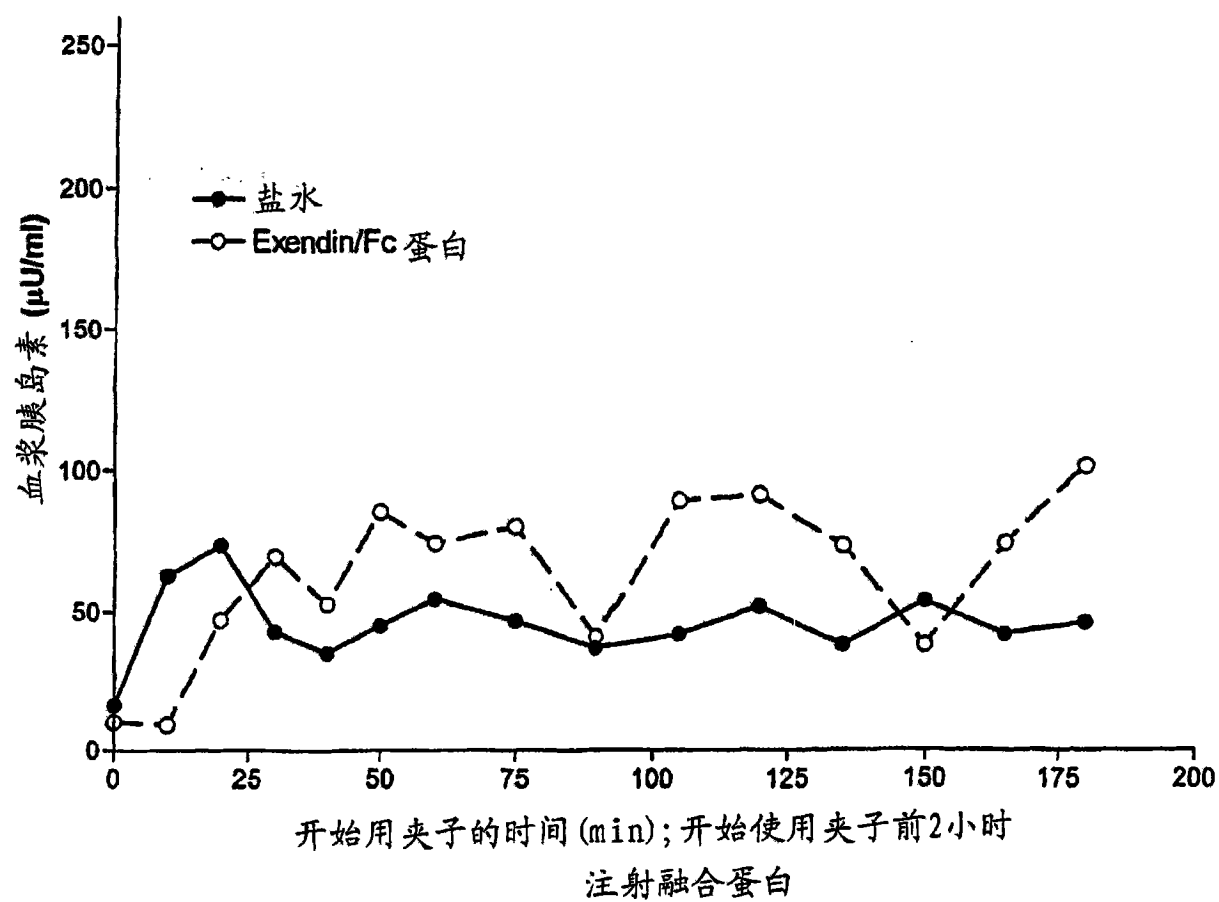


图 11

B.

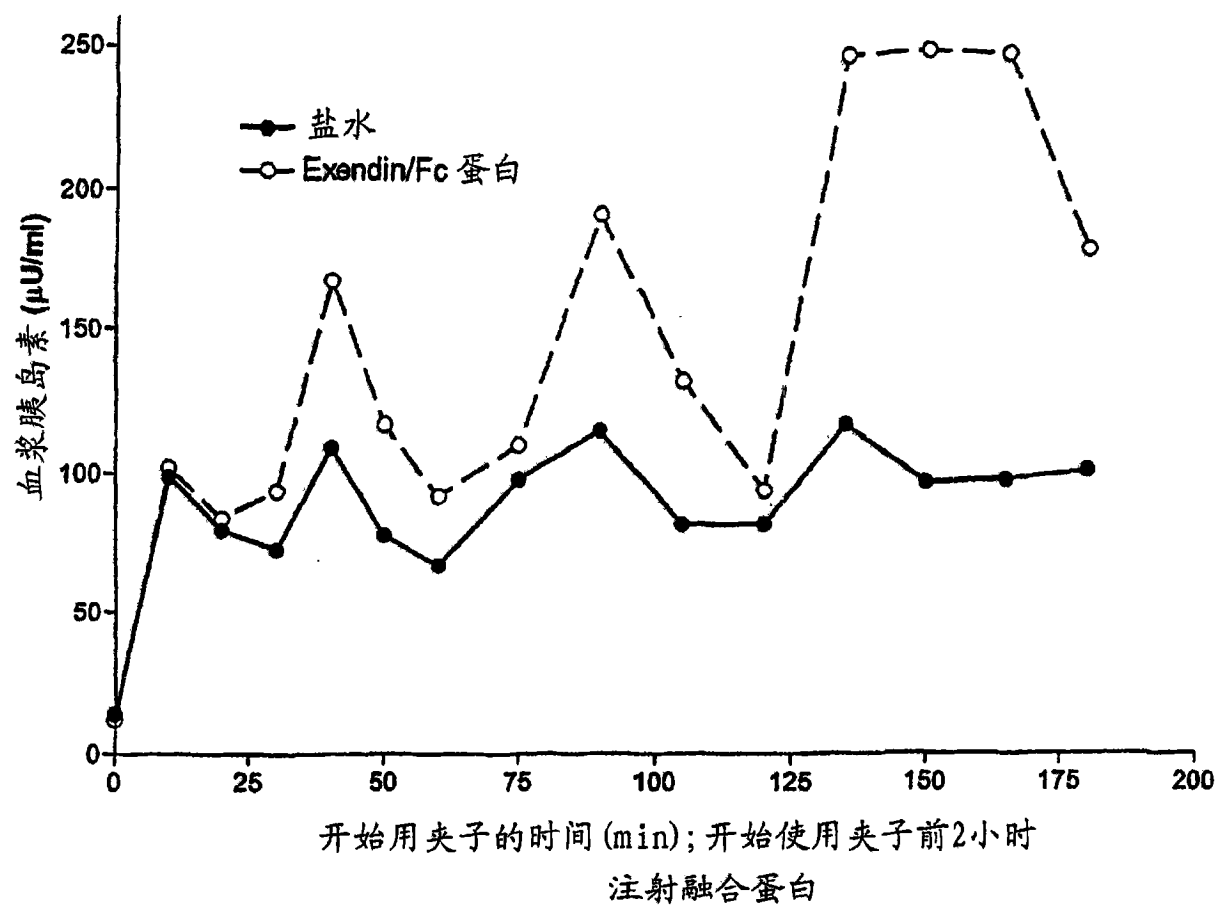


图 11

1
GAGCCCAAATCTTGTGACAAACTCACACATGCCCACCGTGCCCAGCACC
50
TGAATCCTGGGGGACCGTCAGTCTTCCTCTTCCCCCAAACCCAAGG
100
ACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGAC
150
GTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGT
200
GGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCA
250
CGTACCGTGTGGTCAGCGTCCTCACCCTCCTGCACCAGGACTGGCTGAAT
300
GGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCAT
350
CGAGAAAACCATCTCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGT
400
ACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCTG
450
ACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGA
500
GAGCAATGGGCAGCCGGAGAACAACCTACAAGACCACGCCTCCCGTGCTGG
550
ACTCCGACGGCTCCTTCTTCTCTATAGCAAGCTCACCGTGGACAAGAGC
600
AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCT
650
GCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAT
700
AGT [SEQ ID NO: 33]

图 12

1
GATGCGCACAAAGAGTGAGGTTGCTCATCGGTTTAAAGATTGTTGGGAGAAGA
50
AAATTTCAAAGCCTTGGTGTGATTGCCTTTGCTCAGTATCTTCAGCAGT
100
GTCCATTTGAAGATCATGTAAAATTAGTGAATGAAGTAACTGAATTTGCA
150
AAAACATGTGTTGCTGATGAGTCAGCTGAAAATTGTGACAAATCACTTCA
200
TACCCTTTTTTGGAGACAAATTATGCACAGTTGCAACTCTTCGTGAAACCT
250
ATGGTGAAATGGCTGACTGCTGTGCAAAACAAGAACCTGAGAGAAATGAA
300
TGCTTCTTGCAACACAAAGATGACAACCCAAACCTCCCCGATTGGTGAG
350
ACCAGAGGTTGATGTGATGTGCACTGCTTTTCATGACAATGAAGAGACAT
400
TTTTGAAAAATACTTATATGAAATTGCCAGAAGACATCCTTACTTTTAT
450
GCCCCGGAACCTCTTTTCTTTGCTAAAAGGTATAAAGCTGCTTTTACAGA
500
ATGTTGCCAAGCTGCTGATAAAGCTGCCTGCCTGTTGCCAAAGCTCGATG
550
AACTTCGGGATGAAGGGAAGGCTTCGTCTGCCAAACAGAGACTCAAGTGT
600
GCCAGTCTCCAAAAATTTGGAGAAAGAGCTTTCAAAGCATGGGCAGTAGC
650
TCGCCTGAGCCAGAGATTTCCCAAAGCTGAGTTTGCAGAAGTTTCCAAGT
700
TAGTGACAGATCTTACCAAAGTCCACACGGAATGCTGCCATGGAGATCTG
750
CTTGAATGTGCTGATGACAGGGCGGACCTTGCCAAGTATATCTGTGAAAA
800
TCAAGATTGATCTCCAGTAAACTGAAGGAATGCTGTGAAAAACCTCTGT
850
TGGAAAAATCCCACTGCATTGCCGAAGTGGAATATGATGAGATGCCTGCT
900
GACTTGCCCTTCATTAGCTGCTGATTTTGTGAAAGTAAGGATGTTTGCAA
950
AAACTATGCTGAGGCAAAGGATGTCTTCCTGGGCATGTTTTTGTATGAAT
1000
ATGCAAGAAGGCATCCTGATTACTCTGTCTGCTGCTGCTGAGACTTGCC

图 13

1050
AAGACATATGAAACCACTCTAGAGAAGTGCTGTGCCGCTGCAGATCCTCA
1100
TGAATGCTATGCCAAAGTGTTGATGAATTTAAACCTCTTGTGGAAGAGC
1150
CTCAGAATTTAATCAAACAAAATTGTGAGCTTTTTTGAGCAGCTTGGAGAG
1200
TACAAATTCAGAATGCGCTATTAGTTCGTTACACCAAGAAAGTACCCCA
1250
AGTGTCAACTCCAACCTCTTGTAGAGGTCTCAAGAAACCTAGGAAAAGTGG
1300
GCAGCAAATGTTGTAAACATCCTGAAGCAAAAAGAATGCCCTGTGCAGAA
1350
GACTATCTATCCGTGGTCCTGAACCAGTTATGTGTGTTGCATGAGAAAAC
1400
GCCAGTAAGTGACAGAGTCACCAAATGCTGCACAGAATCCTTGGTGAACA
1450
GGCGACCATGCTTTTCAGCTCTGGAAGTCGATGAAACATACGTTCCCAA
1500
GAGTTTAATGCTGAAACATTACCTTCCATGCAGATATATGCACACTTTC
1550
TGAGAAGGAGAGACAAATCAAGAAACAACTGCACTTGTTGAGCTCGTGA
1600
AACACAAGCCCAAGGCAACAAAAGAGCAACTGAAAGCTGTTATGGATGAT
1650
TTCGCAGCTTTTGTAGAGAAGTGCTGCAAGGCTGACGATAAGGAGACCTG
1700
CTTTGCCGAGGAGGGTAAAAAAGCTTGTGCTGCAAGTCAAGCTGCCTTAG
1750
GCTTATAATGAC [SEQ ID NO: 35]

图 13 续