



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2008-0085082
(43) 공개일자 2008년09월22일

(51) Int. Cl.

C07K 19/00 (2006.01) C07K 14/605 (2006.01)
C07K 14/76 (2006.01)

(21) 출원번호 10-2008-7019148(분할)
(22) 출원일자 2008년08월04일
심사청구일자 없음
(62) 원출원 특허 10-2003-7007541
원출원일자 2003년06월05일
심사청구일자 2006년11월29일
번역문제출일자 2008년08월04일
(86) 국제출원번호 PCT/US2001/043165
국제출원일자 2001년11월29일
(87) 국제공개번호 WO 2002/46227
국제공개일자 2002년06월13일
(30) 우선권주장
60/251,954 2000년12월07일 미국(US)

(71) 출원인

일라이 릴리 앤드 캄파니

미국 46285 인디애나주 인디애나폴리스 릴리 코포
레이트 센터

(72) 발명자

글래스너, 볼프강

미국 46254 인디애나주 인디애나폴리스 필드스톤
코트 7512

미카노빅, 라드밀라

미국 46236 인디애나주 인디애나폴리스 화이트 오
크 트레일 7126

츠창, 생-형, 레인보우

미국 46033 인디애나주 카멜 릴리 뮤즈 4963

(74) 대리인

김영, 장수길

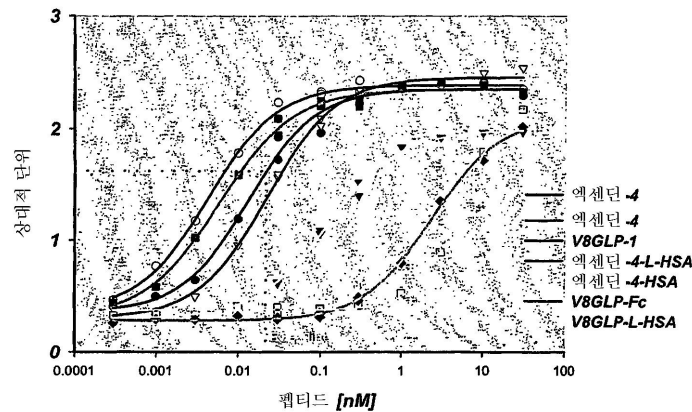
전체 청구항 수 : 총 1 항

(54) G L P-1 융합 단백질

(57) 요약

본 발명은 글루카곤-유사 펩티드의 생체내 반감기를 연장시키는 효과를 지닌 단백질과 융합된, 글루카곤-유사 화
합물에 관한 것이다. 상기 융합 단백질은 인슐린-비의존성 당뇨병뿐만 아니라, 다양한 다른 장애를 치료하기 위
해 이용될 수 있다.

대표도 - 도8



특허청구의 범위

청구항 1

N-말단 및 C-말단을 지닌 제1 폴리펩티드가 N-말단 및 C-말단을 지닌 제2 폴리펩티드에 융합되어 있는 이중성 융합 단백질의 치료적 유효량을 투여하는 것을 포함하고, 이때 제1 폴리펩티드는 GLP-1 화합물이고, 제2 폴리펩티드는

- a) 인간 알부민;
- b) 인간 알부민 유사체; 및
- c) 인간 알부민의 단편

으로 구성된 군에서 선택된 것이며, 제1 폴리펩티드의 C-말단이 제2 폴리펩티드의 N-말단에 융합되어 있는 것인, 인슐린 비의존성 당뇨병을 앓는 환자의 치료 방법.

명세서

발명의 상세한 설명

기술 분야

- <1> 본 발명은 펩티드의 생체내 반감기를 연장시키는 효과를 지닌 단백질과 융합된, 글루카곤-유사 펩티드의 유사체 및 유도체를 비롯한 글루카곤-유사 펩티드에 관한 것이다. 이러한 융합 단백질은 인슐린-비의존성 당뇨병뿐만 아니라, 다양한 다른 장애를 치료하기 위해 이용될 수 있다.

배경 기술

- <2> 글루카곤-유사 펩티드 1 (GLP-1)은 음식 섭취시 장내 L-세포에 의해 분비되는, 아미노산 37개로 구성된 펩티드이다. 이것은 인슐린 분비를 자극 (인슐린 분비 작용 (insulinotropic action))함으로써 세포에 의한 글루코스 흡수를 유발하여 혈청내 글루코스의 양을 감소시키는 것으로 알려져 있다 (예를 들면, 문헌 [Mojsov, S., (1992) *Int. J. Peptide Protein Research*, 40: 333-343] 참조). 그러나, GLP-1은 낮은 활성을 갖고 있다. GLP-1의 제6 위치와 제7 위치 사이가 내부 절단 (endogenous cleavage)되면 더욱 강력한 생물학적 활성을 지닌 GLP-1(7-37)OH 펩티드가 생성된다. 다수의 GLP-1 유사체 및 유도체가 공지되어 있으며, 본원에서는 이들을 "GLP-1 화합물"로 명명한다. 이러한 GLP-1 유사체로는 엑센딘 (Exendin)이 포함되며, 이것은 독도마뱀 (Gila-monster)의 독에서 발견된 펩티드이다. 엑센딘은 천연 GLP-1과 서열 상동성을 갖고 있으며, GLP-1(7-37)OH의 경우와 마찬가지로 GLP-1 수용체와 결합함으로써 다수의 활성을 유발하는 신호 전달 캐스케이드를 개시할 수 있다.
- <3> GLP-1 화합물은 생리학적으로 중요한 다양한 활성을 갖고 있다. 예를 들면, GLP-1은 인슐린 방출을 자극하고, 글루카곤 분비를 낮추고, 위 배출 (gastric emptying)을 저해하며, 글루코스 이용도를 증가시키는 것으로 알려져 있다 (문헌 [Nauck, M. A., et al. (1993) *Diabetologia* 36: 741-744; Gutniak, M., et al. (1992) *New England J. of Med.* 326: 1316-1322; Nauck, M. A., et al., (1993) *J. Clin. Invest.* 91: 301-307] 참조).
- <4> GLP-1은 인슐린-비의존성 당뇨병 (NIDDM)에 대한 치료제로서 매우 유망하다. NIDDM과 관련된 인슐린 내성을 치료하기 위한 많은 경구용 약물이 시판중이다. 그러나 상기 질병이 진행됨에 따라, 환자는 인슐린 방출을 자극하는 치료를 받아야 하고, 결국 인슐린 주사 투여를 수반한 치료를 받아야 한다. 그러나 인슐린 방출을 자극하는 현행 약물은 실제로 인슐린을 투여하는 경우와 같이 저혈당증을 유발할 수 있다. 그러나, GLP-1의 활성은 혈액내 글루코스 농도에 의해 조절된다. 상기 농도가 특정 역치 수준까지 감소될 경우에는 GLP-1이 활성화되지 않는다. 따라서, GLP-1을 이용한 치료와 관련해서는 저혈당증의 위험이 없다.
- <5> 그러나 GLP-1 펩티드의 빠른 소멸 및 짧은 반감기로 인하여, GLP-1 펩티드를 이용한 치료법의 유용성이 제한되었다. 예를 들면, GLP-1(7-37)은 단지 3분 내지 5분의 혈청내 반감기를 갖고 있다. GLP-1(7-36) 아미드는 피하 투여시 약 50분의 작용 시간을 갖고 있다. 내인성 프로테아제 절단에 대해 내성을 지닌 유사체 및 유도체조차도, 24시간에 걸친 반복 투여를 피할 수 있을 정도로 긴 반감기를 갖고 있지는 않다. 장기간에 걸쳐 치료제를 혈액내에서 높은 농도로 유지하는 것이 바람직할 경우에는, 치료제의 빠른 소멸은 반복적인 투여가 필요하기

때문에 불리한 것이다. 또한 지효성 화합물은, 과거에 단지 경구 약물 치료만을 받았던 당뇨병 환자에게 특히 중요하다. 대체로 이러한 환자들은 약물을 다중 주사투여하는 치료법으로 전환하는 것이 매우 어렵다.

발명의 내용

해결 하고자하는 과제

<6> 본 발명은, GLP-1 수용체 활성화와 관련된 신호 전달 활성을 보유하며, 연장된 반감기를 갖는 이중성 GLP-1 융합 단백질을 제공하는 것이다.

과제 해결수단

<7> 본 발명은 혈장내 반감기가 짧은 화합물을 전달하는 것과 관련된 상기 문제점들을 극복하였다. 본 발명의 화합물에는 면역글로블린의 Fc 부분 또는 알부민과 같은 긴 순환형 반감기를 지닌 다른 단백질과 융합된 GLP-1 화합물이 포함된다.

<8> 일반적으로, 크기가 작은 치료용 펩티드는 그의 경미한 구조 변화로 인해 안정성 및(또는) 생물학적 활성이 영향을 받을 수 있기 때문에 취급하기 어렵다. 이것은 현재 개발중인 GLP-1 화합물에서 더욱 그러하다. 예를 들면, GLP-1(7-37)OH는 주로 알파 나선 구조에서 주로 베타 쉬트(sheet) 구조로 형태 변화를 하는 경향이 있다. 상기 베타 쉬트 형태는 불활성인 것으로 판단되는 응집물을 형성한다. 따라서, 반감기가 증가된 생물학적 활성 GLP-1 융합 단백질을 개발할 수 있다는 것은 놀라운 것이었다. GLP-1(7-37)OH만으로 작업하는 난점과, 부착된 작은 GLP-1 펩티드에 비해 크기가 더 큰 융합 파트너를 고려하면, 본 발명은 매우 예상치 못한 것이었다.

<9> 본 발명의 화합물에는, 제1 폴리펩티드가 GLP-1 화합물이고, 제2 폴리펩티드는

<10> a) 인간 알부민;

<11> b) 인간 알부민 유사체; 및

<12> c) 인간 알부민의 단편

<13> 으로 구성된 군에서 선택된 것이며, 제1 폴리펩티드의 C-말단이 제2 폴리펩티드의 N-말단에 융합되어 있는, N-말단 및 C-말단을 지닌 제2 폴리펩티드에 융합된 N-말단 및 C-말단을 지닌 제1 폴리펩티드를 포함하는 이중성 융합 단백질이 포함된다.

<14> 또한 본 발명의 화합물에는, 제1 폴리펩티드가 GLP-1 화합물이고, 제2 폴리펩티드는

<15> a) 인간 알부민;

<16> b) 인간 알부민 유사체; 및

<17> c) 인간 알부민의 단편

<18> 으로 구성된 군에서 선택된 것이며, 제1 폴리펩티드의 C-말단이 펩티드 링커(linker)를 통해 제2 폴리펩티드의 N-말단에 융합되어 있는, N-말단 및 C-말단을 지닌 제2 폴리펩티드에 융합된 N-말단 및 C-말단을 지닌 제1 폴리펩티드를 포함하는 이중성 융합 단백질이 포함된다. 바람직하게는, 상기 펩티드 링커가

<19> a) 글리신이 풍부한 펩티드;

<20> b) 서열 $[Gly-Gly-Gly-Gly-Ser]_n$ 을 갖고, 여기에서 n은 1, 2, 3, 4, 5 또는 6인 펩티드; 및

<21> c) 서열 $[Gly-Gly-Gly-Gly-Ser]_3$ 을 갖는 펩티드

<22> 로 구성된 군에서 선택된 것이다.

<23> 본 발명의 추가적인 화합물에는, 제1 폴리펩티드가 GLP-1 화합물이고, 제2 폴리펩티드는

<24> a) 면역글로블린의 Fc 부분;

<25> b) 면역글로블린 Fc 부분의 유사체; 및

<26> c) 면역글로블린의 Fc 부분의 단편

<27> 으로 구성된 군에서 선택된 것이며, 제1 폴리펩티드의 C-말단이 제2 폴리펩티드의 N-말단에 융합되어 있는, N-

말단 및 C-말단을 지닌 제2 폴리펩티드에 융합된 N-말단 및 C-말단을 지닌 제1 폴리펩티드를 포함하는 이중성 융합 단백질이 포함된다. 상기 GLP-1 화합물은 펩티드 링커를 통해 제2 폴리펩티드에 융합될 수 있다. 바람직하게는, 상기 펩티드 링커가

- <28> a) 글리신이 풍부한 펩티드;
- <29> b) 서열 [Gly-Gly-Gly-Gly-Ser]_n을 갖고, 여기에서 n은 1, 2, 3, 4, 5 또는 6인 펩티드; 및
- <30> c) 서열 [Gly-Gly-Gly-Gly-Ser]₃을 갖는 펩티드
- <31> 로 구성된 군에서 선택된 것이다.
- <32> 일반적으로 상기 이중성 융합 단백질의 일부인 GLP-1 화합물이, GLP-1(7-37)OH, GLP-1(7-36)OH 또는 엑센딘-4에서의 상응하는 아미노산과 상이한 아미노산을 6개 이하로 갖는 것이 바람직하다. 상기 GLP-1 화합물이, GLP-1(7-37)OH, GLP-1(7-36)OH 또는 엑센딘-4에서의 상응하는 아미노산과 상이한 아미노산을 5개 이하로 갖는 것이 더욱 바람직하다. GLP-1 화합물이, GLP-1(7-37)OH, GLP-1(7-36)OH 또는 엑센딘-4에서의 상응하는 아미노산과 상이한 아미노산을 4개, 3개 또는 2개 이하로 갖는 것이 가장 바람직하다. 바람직하게는, 상기 이중성 융합 단백질의 일부인 GLP-1 화합물은 8 위치에 글리신 또는 발린을 갖는다.
- <33> 또한 본 발명에는, 본원에 기재된 이중성 융합 단백질을 코딩하는 폴리뉴클레오티드, 상기 폴리뉴클레오티드를 포함하는 벡터, 및 본원에 기재된 벡터가 형질감염되거나 또는 형질전환된 숙주세포가 포함된다. 또한 이중성 융합 단백질이 검출 가능한 양으로 발현되는 조건하에서, 본원에 기재된 폴리뉴클레오티드를 전사 및 번역하는 단계를 포함하는, 이중성 융합 단백질의 생산 방법이 포함된다.
- <34> 또한 본 발명에는, 혈액내 글루코스 농도의 정상화가 필요한 포유동물에게 본원에 기재된 이중성 융합 단백질을 치료상 유효량으로 투여하는 것을 포함하는, 상기 포유동물에서 혈액내 글루코스 농도를 정상화시키는 방법이 포함된다.
- <35> 본 발명의 이중성 융합 단백질은 인간 알부민, 인간 알부민 유사체, 인간 알부민 단편, 면역글로블린의 Fc 부분, 면역글로블린 Fc 부분의 유사체 또는 면역글로블린 Fc 부분의 단편과 융합된 GLP-1 화합물을 포함한다. 상기 GLP-1 화합물의 C-말단은 알부민 또는 Fc 단백질의 N-말단에 직접적으로 또는 펩티드 링커를 통해 융합될 수 있다. 이러한 이중성 융합 단백질은 생물학적 활성을 갖고 있으며, 천연 GLP-1와 비교하여 증가된 반감기를 갖고 있다.
- <36> 바람직하게는, 이중성 융합 단백질의 일부분을 구성하는 상기 GLP-1 화합물로서, 췌장내 β-세포상의 GLP-1 수용체와 결합함으로써 인슐린 분비 활성을 나타내도록 천연 GLP-1(7-37)OH와 충분한 상동성을 지닌, 자연 발생의 또는 비-자연 발생의 아미노산 약 25개 내지 약 39개로 이루어진 폴리펩티드가 포함된다. 전형적으로 GLP-1 화합물은 GLP-1(7-37)OH, GLP-1(7-37)OH 유사체, GLP-1(7-37)OH 단편 또는 GLP-1(7-37)OH 유사체 단편의 아미노산 서열을 갖는 폴리펩티드를 포함한다. GLP-1(7-37)OH는 서열 1의 아미노산 서열을 갖는다:

```

7   8   9  10  11  12  13  14  15  16  17
His-Ala-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-Ser-
18  19  20  21  22  23  24  25  26  27  28
Ser-Tyr-Leu-Glu-Gly-Gln-Ala-Ala-Lys-Glu-Phe-
29  30  31  32  33  34  35  36  37
Ile-Ala-Trp-Leu-Val-Lys-Gly-Arg-Gly

```

(서열 1)

- <37>
- <38> 당업계와 관례에 따라, GLP-1(7-37)OH의 아미노 말단을 잔기 번호 7로 정하고, 카르복시 말단을 번호 37로 정하였다. 폴리펩티드내 다른 아미노산들은, 서열 1에 나타난 것과 같이 연속적으로 번호를 부여하였다. 예를 들면, 위치 12는 페닐알라닌이고, 위치 22는 글리신이다.
- <39> 또한 GLP-1 화합물에는 "GLP-1 단편"이 포함된다. GLP-1 단편은 GLP-1(7-37)OH, 또는 이것의 유사체 또는 유도체의 N-말단 및(또는) C-말단에서 1개 이상의 아미노산을 절단하여 얻은 폴리펩티드이다. GLP-1(7-37)OH를 나타내기 위해 사용된 명명법은 GLP-1 단편에서도 적용할 수 있다. 예를 들면 GLP-1(9-36)OH는, N-말단에서 2개의 아미노산을 절단하고, C-말단에서 1개의 아미노산을 절단하여 얻은 GLP-1 단편을 나타낸다. 상기 단편내 아미노산은, GLP-1(7-37)OH내 상응하는 아미노산의 번호와 동일한 번호로 나타내었다. 예를 들면 GLP-1(7-37)OH에서와 같이, GLP-1(9-36)OH에서 N-말단의 글루탐산은 위치 9이고, 위치 12는 페닐알라닌이며, 위치 22는

글리신이다. GLP-1(7-36)OH의 경우, GLP-1(7-37)OH의 위치 37에 있는 글리신이 결실되었다.

- <40> 또한 GLP-1 화합물에는, 1개 이상의 아미노산이 GLP-1(7-37)OH, 또는 이것의 단편 또는 유사체의 N-말단 및(또는) C-말단에 부가된 폴리펩티드가 포함된다. 이러한 형태의 GLP-1 화합물은 약 39개 이하의 아미노산을 갖는 것이 바람직하다. "연장된" GLP-1 화합물내의 아미노산은 GLP-1(7-37)OH내 상응하는 아미노산의 번호와 동일한 번호로 나타내었다. 예를 들면, GLP-1(7-37)OH의 N-말단에 2개의 아미노산을 부가하여 얻어진 GLP-1 화합물에서 N-말단의 아미노산은 위치 5에 있고, GLP-1(7-37)OH의 C-말단에 1개의 아미노산을 부가하여 얻어진 GLP-1 화합물에서 C-말단의 아미노산은 위치 38에 있다. 따라서 GLP-1(7-37)OH와 같이, 상기 두 경우의 "연장된" GLP-1 화합물에서 위치 12는 페닐알라닌이고 위치 22는 글리신이다. 연장된 GLP-1 화합물에서의 아미노산 위치 1 내지 위치 6은, 바람직하게는 GLP-1(1-37)OH에서의 상응 위치에 있는 아미노산과 동일하거나 또는 보존적으로 치환된다. 연장된 GLP-1 화합물의 아미노산 38 내지 45는, 바람직하게는 글루카곤 또는 엑센딘-4에서 상응 위치에 있는 아미노산과 동일하거나 또는 보존적으로 치환된다.
- <41> 본 발명의 GLP-1 화합물에는 "GLP-1 유사체"가 포함된다. GLP-1 유사체는 인슐린 분비 활성을 갖도록, GLP-1(7-37)OH 또는 GLP-1(7-37)OH 단편과 충분한 상동성을 갖고 있다. 바람직하게는 GLP-1 유사체가, GLP-1(7-37)OH 또는 GLP-1(7-37)OH 단편내 상응 위치에 있는 아미노산과 1개, 2개, 3개, 4개 또는 5개의 아미노산이 서로 상이하도록 변형된, GLP-1(7-37)OH 또는 GLP-1(7-37)OH 단편의 아미노산 서열을 갖는다. GLP-1 화합물을 나타내기 위해 본원에 사용된 명명법에서, 치환된 새로운 아미노산 및 이것의 위치를 원래 구조의 앞에 나타내었다. 예를 들면, Glu²²-GLP-1(7-37)OH는 GLP-1(7-37)OH의 위치 22에서 정상적으로 발견되는 글리신이 글루탐산으로 치환된 GLP-1 화합물을 나타내고, Val⁸-Glu²²-GLP-1(7-37)OH는 GLP-1(7-37)OH의 위치 8에서 정상적으로 발견되는 알라닌과 위치 22에서 정상적으로 발견되는 글리신이 각각 발린과 글루탐산으로 치환된 GLP-1 화합물을 나타낸다.
- <42> 또한 본 발명의 GLP-1 화합물에는 "GLP-1 유도체"가 포함된다. GLP-1 유도체는, GLP-1 또는 GLP-1 유사체의 아미노산 서열을 갖고 있지만 추가적으로 1개 이상의 아미노산 측쇄기, α-탄소 원자, 말단 아미노기 또는 말단 카르복실산기의 화학적 변형을 갖는 분자로 정의된다. 상기 화학적 변형에는 화합물 잔기 부가, 새로운 결합 형성 및 화합물 잔기 제거가 포함되지만, 이것에만 제한되는 것은 아니다. 아미노산 측쇄기의 변형에는 리신 ε-아미노기의 아실화, 아르기닌, 히스티딘 또는 리신의 N-알킬화, 글루탐산 또는 아스파르트산내 카르복실산기의 알킬화, 및 글루타민 또는 아스파라긴의 탈아미드화가 포함되지만, 이것에만 제한되는 것은 아니다. 말단 아미노기의 변형에는 테스-아미노, N-저급 알킬, N-디-저급 알킬 및 N-아실의 변형이 포함되지만, 이것에만 제한되는 것은 아니다. 말단 카르복시기의 변형에는 아마이드, 저급 알킬 아마이드, 디알킬 아마이드 및 저급 알킬 에스테르의 변형이 포함되지만, 이것에만 제한되는 것은 아니다. 저급 알킬은 C₁-C₄ 알킬이다. 또한 1개 이상의 측쇄기 또는 말단기는, 통상의 지식을 가진 단백질 화학자에게 알려진 보호기에 의해 보호될 수도 있다. 아미노산의 α-탄소는 모노메틸화 또는 디메틸화될 수 있다.
- <43> GLP-1 화합물 자체가 GLP-1 수용체와 결합하여 신호 전달을 유도할 수 있는 한, 모든 GLP-1 화합물은 본 발명에 기재된 이중성 융합 단백질의 일부가 될 수 있다. GLP-1 수용체 결합 및 신호 전달은, 각각 EP 제619,322호 및 미국특허 제5,120,712호에 기재된 것과 같은 시험관내 분석을 이용하여 평가할 수 있다.
- <44> 활성을 지닌 다수의 GLP-1 단편, 유사체 및 유도체가 당업계에 잘 알려져 있으며, 이러한 모든 유사체 및 유도체도 본 발명에 기재된 이중성 융합 단백질의 일부가 될 수도 있다. 신규 GLP-1 유사체뿐만 아니라 당업계에 공지된 GLP-1 유사체 및 유도체들의 몇몇 예가 본원에 제공된다.
- <45> 예를 들면, 당업계에 공지된 몇몇 GLP-1 유사체 및 GLP-1 단편에는, GLP-1(7-34) 및 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) 및 Lys¹⁸-GLP-1(7-37)이 포함된다. GLP-1(7-34) 및 GLP-1(7-35)와 같은 GLP-1 유사체는 미국특허 제5,118,666호에 개시되어 있다. 또한 GLP-1(7-36)과 같이 인슐린 분비 특성을 갖는, GLP-1의 생물학적으로 가공된 형태도 알려져 있다. 생물학적 활성을 갖는 또다른 공지의 GLP-1 화합물들이 미국특허 제5,977,071호 (Hoffmann, et al.), 미국특허 제5,545,618호 (Buckley, et al.) 및 문헌 [Adelhorst, et al., *J. Biol. Chem.* 269: 6275 (1994)]에 개시되어 있다.
- <46> 바람직한 GLP-1 유사체의 군은 화학식 1의 GLP-1 유사체로 구성되어 있다 (서열 2).

화학식 1

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-Phe-
 29 30 31 32 33 34 35 36 37 38 39
 Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-
 40 41 42 43 44 45
 Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

(서열 2)

<47>

<48>

상기 식에서,

<49>

위치 8의 Xaa는 Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp 또는 Lys이고;

<50>

위치 9의 Xaa는 Glu, Asp 또는 Lys이고;

<51>

위치 11의 Xaa는 Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp 또는 Lys이고;

<52>

위치 14의 Xaa는 Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp 또는 Lys이고;

<53>

위치 16의 Xaa는 Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp, Trp 또는 Lys이고;

<54>

위치 17의 Xaa는 Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp 또는 Lys이고;

<55>

위치 18의 Xaa는 Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Trp, Tyr 또는 Lys이고;

<56>

위치 19의 Xaa는 Tyr, Phe, Trp, Glu, Asp, Gln 또는 Lys이고;

<57>

위치 20의 Xaa는 Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Trp, Tyr 또는 Lys이고;

<58>

위치 21의 Xaa는 Glu, Asp 또는 Lys이고;

<59>

위치 22의 Xaa는 Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp 또는 Lys이고;

<60>

위치 23의 Xaa는 Gln, Asn, Arg, Glu, Asp 또는 Lys이고;

<61>

위치 24의 Xaa는 Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp 또는 Lys이고;

<62>

위치 25의 Xaa는 Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp 또는 Lys이고;

<63>

위치 26의 Xaa는 Lys, Arg, Gln, Glu, Asp 또는 His이고;

<64>

위치 27의 Xaa는 Leu, Glu, Asp 또는 Lys이고;

<65>

위치 30의 Xaa는 Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp 또는 Lys이고;

<66>

위치 31의 Xaa는 Trp, Phe, Tyr, Glu, Asp 또는 Lys이고;

<67>

위치 32의 Xaa는 Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp 또는 Lys이고;

<68>

위치 33의 Xaa는 Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp 또는 Lys이고;

<69>

위치 34의 Xaa는 Asn, Lys, Arg, Glu, Asp 또는 His이고;

<70>

위치 35의 Xaa는 Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp 또는 Lys이고;

<71>

위치 36의 Xaa는 Gly, Arg, Lys, Glu, Asp 또는 His이고;

<72>

위치 37의 Xaa는 Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp 또는 Lys이거나, 또는 결실되어 있고;

<73>

위치 38의 Xaa는 Ser, Arg, Lys, Glu, Asp 또는 His이거나, 또는 결실되어 있고;

<74>

위치 39의 Xaa는 Ser, Arg, Lys, Glu, Asp 또는 His이거나, 또는 결실되어 있고;

<75>

위치 40의 Xaa는 Gly, Asp, Glu 또는 Lys이거나, 또는 결실되어 있고;

- <76> 위치 41의 Xaa는 Ala, Phe, Trp, Tyr, Glu, Asp 또는 Lys이거나, 또는 결실되어 있고;
- <77> 위치 42의 Xaa는 Ser, Pro, Lys, Glu 또는 Asp이거나, 또는 결실되어 있고;
- <78> 위치 43의 Xaa는 Ser, Pro, Glu, Asp 또는 Lys이거나, 또는 결실되어 있고;
- <79> 위치 44의 Xaa는 Gly, Pro, Glu, Asp 또는 Lys이거나, 또는 결실되어 있고;
- <80> 위치 45의 Xaa는 Ala, Ser, Val, Glu, Asp 또는 Lys이거나, 또는 결실되어 있으며;
- <81> 단, 위치 37, 38, 39, 40, 41, 42, 43 또는 44에 있는 아미노산이 결실되면, 결실된 아미노산 이후에 존재하는 각각의 아미노산도 또한 결실된다.
- <82> 바람직하게는, 화학식 1의 GLP-1 화합물은 GLP-1(7-37)OH 또는 엑센딘-4에서의 상응하는 아미노산과 6개 미만의 아미노산이 상이하다. 더 바람직하게는, GLP-1(7-37)OH 또는 엑센딘-4에서의 상응하는 아미노산과 5개 미만의 아미노산이 상이하다. 더욱 더 바람직하게는, GLP-1(7-37)OH 또는 엑센딘-4에서의 상응하는 아미노산과 4개 미만의 아미노산이 상이하다.
- <83> 본 발명의 GLP-1 화합물에는, 이것의 C-1-6-에스테르, 아마이드, C-1-6-알킬아미드 또는 C-1-6-디알킬아미드와 같은 화학식 1의 유도체가 포함된다. 이 거명을 통해 그 전문이 본원에 참고문헌으로 포함된 W0 99/43706호에는 화학식 1의 GLP-1 화합물의 유도체가 기재되어 있다. 본 발명에는 W0 99/43706호에 기재된 것과 같이 유도체화된 화학식 1의 화합물 및 유도체화되지 않은 화학식 1의 화합물이 포함된다.
- <84> 또다른 바람직한 GLP-1 화합물의 군은 화학식 2의 GLP-1 유사체로 구성되어 있다 (서열 3).

화학식 2

7 8 9 10 11 12 13 14 15 16 17
Xaa-Xaa-Xaa-Gly-Xaa-Xaa-Thr-Ser-Asp-Xaa-Ser-
18 19 20 21 22 23 24 25 26 27 28
Xaa-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Ala-Xaa-Xaa-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Xaa-Leu-Xaa-Xaa-Xaa-Xaa-R

(서열 3)

- <85>
- <86> 상기 식에서,
- <87> 위치 7의 Xaa는 L-히스티딘, D-히스티딘, 테스아미노-히스티딘, 2-아미노-히스티딘, β -히드록시-히스티딘, 호모히스티딘, α -플루오로메틸-히스티딘 또는 α -메틸-히스티딘이고;
- <88> 위치 8의 Xaa는 Gly, Ala, Val, Leu, Ile, Ser 또는 Thr이고;
- <89> 위치 9의 Xaa는 Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu 또는 His이고;
- <90> 위치 11의 Xaa는 Asp, Glu, Arg, Thr, Ala, Lys 또는 His이고;
- <91> 위치 12의 Xaa는 His, Trp, Phe 또는 Tyr이고;
- <92> 위치 16의 Xaa는 Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu 또는 Ala이고;
- <93> 위치 18의 Xaa는 His, Pro, Asp, Glu, Arg, Ser, Ala 또는 Lys이고;
- <94> 위치 22의 Xaa는 Gly, Asp, Glu, Gln, Asn, Lys, Arg 또는 Cys이고;
- <95> 위치 23의 Xaa는 His, Asp, Lys, Glu, Gln 또는 Arg이고;
- <96> 위치 24의 Xaa는 Glu, Arg, Ala 또는 Lys이고;
- <97> 위치 26의 Xaa는 Trp, Tyr, Phe, Asp, Lys, Glu 또는 His이고;
- <98> 위치 27의 Xaa는 Ala, Glu, His, Phe, Tyr, Trp, Arg 또는 Lys이고;
- <99> 위치 30의 Xaa는 Ala, Glu, Asp, Ser 또는 His이고;

- <100> 위치 31의 Xaa는 Asp, Glu, Ser, Thr, Arg, Trp 또는 Lys이고;
- <101> 위치 33의 Xaa는 Asp, Arg, Val, Lys, Ala, Gly 또는 Glu이고;
- <102> 위치 34의 Xaa는 Glu, Lys 또는 Asp이고;
- <103> 위치 35의 Xaa는 Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His 또는 Glu이고;
- <104> 위치 36의 Xaa는 Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu 또는 His이며;
- <105> 위치 37의 R은 Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly 또는 Gly-Pro이거나, 또는 결실되어 있다.
- <106> 또다른 바람직한 GLP-1 화합물의 군은 화학식 3의 GLP-1 유사체로 구성되어 있다 (서열 4).

화학식 3

```

  7   8   9   10  11  12  13  14  15  16  17
Xaa-Xaa-Glu-Gly-Xaa-Xaa-Thr-Ser-Asp-Xaa-Ser-
18  19  20  21  22  23  24  25  26  27  28
Ser-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Xaa-Lys-Xaa-Phe-
29  30  31  32  33  34  35  36  37
Ile-Xaa-Trp-Leu-Xaa-Xaa-Xaa-Xaa-R

```

(서열 4)

- <107>
- <108> 상기 식에서,
- <109> 위치 7의 Xaa는 L-히스티딘, D-히스티딘, 테스아미노-히스티딘, 2-아미노-히스티딘, β -히드록시-히스티딘, 호모히스티딘, α -플루오로메틸-히스티딘 또는 α -메틸-히스티딘이고;
- <110> 위치 8의 Xaa는 Gly, Ala, Val, Leu, Ile, Ser 또는 Thr이고;
- <111> 위치 11의 Xaa는 Asp, Glu, Arg, Thr, Ala, Lys 또는 His이고;
- <112> 위치 12의 Xaa는 His, Trp, Phe 또는 Tyr이고;
- <113> 위치 16의 Xaa는 Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu 또는 Ala이고;
- <114> 위치 22의 Xaa는 Gly, Asp, Glu, Gln, Asn, Lys, Arg 또는 Cys이고;
- <115> 위치 23의 Xaa는 His, Asp, Lys, Glu 또는 Gln이고;
- <116> 위치 24의 Xaa는 Glu, His, Ala 또는 Lys이고;
- <117> 위치 25의 Xaa는 Asp, Lys, Glu 또는 His이고;
- <118> 위치 27의 Xaa는 Ala, Glu, His, Phe, Tyr, Trp, Arg 또는 Lys이고;
- <119> 위치 30의 Xaa는 Ala, Glu, Asp, Ser 또는 His이고;
- <120> 위치 33의 Xaa는 Asp, Arg, Val, Lys, Ala, Gly 또는 Glu이고;
- <121> 위치 34의 Xaa는 Glu, Lys 또는 Asp이고;
- <122> 위치 35의 Xaa는 Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His 또는 Glu이고;
- <123> 위치 36의 Xaa는 Arg, Glu 또는 His이며;
- <124> 위치 37의 R은 Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly 또는 Gly-Pro이거나, 또는 결실되어 있다.
- <125> 또다른 바람직한 GLP-1 화합물의 군은 화학식 4의 GLP-1 유사체로 구성되어 있다 (서열 5).

화학식 4

7 8 9 10 11 12 13 14 15 16 17
Xaa-Xaa-Glu-Gly-Thr-Xaa-Thr-Ser-Asp-Xaa-Ser-
18 19 20 21 22 23 24 25 26 27 28
Ser-Tyr-Leu-Glu-Xaa-Xaa-Ala-Ala-Xaa-Glu-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Trp-Leu-Val-Lys-Xaa-Arg-R

(서열 5)

<126>

<127>

상기 식에서

<128>

위치 7의 Xaa는 L-히스티딘, D-히스티딘, 테스아미노-히스티딘, 2-아미노-히스티딘, β -히드록시-히스티딘, 호모히스티딘, α -플루오로메틸-히스티딘 또는 α -메틸-히스티딘이고;

<129>

위치 8의 Xaa는 Gly, Ala, Val, Leu, Ile, Ser, Met 또는 Thr이고;

<130>

위치 12의 Xaa는 His, Trp, Phe 또는 Tyr이고;

<131>

위치 16의 Xaa는 Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu 또는 Ala이고;

<132>

위치 22의 Xaa는 Gly, Asp, Glu, Gln, Asn, Lys, Arg 또는 Cys이고;

<133>

위치 23의 Xaa는 His, Asp, Lys, Glu 또는 Gln이고;

<134>

위치 26의 Xaa는 Asp, Lys, Glu 또는 His이고;

<135>

위치 30의 Xaa는 Ala, Glu, Asp, Ser 또는 His이고;

<136>

위치 35의 Xaa는 Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His 또는 Glu이고;

<137>

위치 37의 R은 Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly 또는 Gly-Pro이거나, 또는 결실되어 있다.

<138>

또다른 바람직한 GLP-1 화합물의 군은 화학식 5의 GLP-1 유사체로 구성되어 있다 (서열 6).

화학식 5

7 8 9 10 11 12 13 14 15 16 17
Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-Ser-
18 19 20 21 22 23 24 25 26 27 28
Ser-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Ala-Lys-Glu-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Trp-Leu-Val-Lys-Gly-Arg-R

(서열 6)

<139>

<140>

상기 식에서,

<141>

위치 7의 Xaa는 L-히스티딘, D-히스티딘, 테스아미노-히스티딘, 2-아미노-히스티딘, β -히드록시-히스티딘, 호모히스티딘, α -플루오로메틸-히스티딘 또는 α -메틸-히스티딘이고;

<142>

위치 8의 Xaa는 Gly, Ala, Val, Leu, Ile, Ser 또는 Thr이고;

<143>

위치 22의 Xaa는 Gly, Asp, Glu, Gln, Asn, Lys, Arg 또는 Cys이고;

<144>

위치 23의 Xaa는 His, Asp, Lys, Glu, 또는 Gln이고;

<145>

위치 24의 Xaa는 Ala, Glu, His, Phe, Tyr, Trp, Arg 또는 Lys이고;

<146>

위치 30의 Xaa는 Ala, Glu, Asp, Ser 또는 His이며;

<147>

위치 37의 R은 Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly 또는 Gly-Pro이거나, 또는 결실되어 있다.

<148>

바람직한 화학식 1, 2, 3, 4 및 5의 GLP-1 화합물은 GLP-1 유사체 또는 GLP-1 유사체의 단편을 포함하며, 여기

에서 상기 유사체 또는 단편은 위치 8에 알라닌이 아닌 다른 아미노산을 갖는다 (위치 8 유사체). 바람직하게는, 상기 위치 8 유사체는 천연 GLP-1(7-37)OH의 상응하는 아미노산과 비교하여, 위치 9, 11, 12, 16, 18, 22, 23, 24, 26, 27, 30, 31, 33, 34, 35, 36 및 37에서 1개 이상의 추가적으로 치환된 아미노산을 갖는다. 또한 상기 유사체가 천연 GLP-1(7-37)OH 또는 GLP-1(7-36)OH의 상응하는 아미노산과 비교하여 6개 이하의 치환된 아미노산을 갖는 것이 바람직하다. 더 바람직한 유사체는 천연 GLP-1(7-37)OH 또는 GLP-1(7-36)OH의 상응하는 아미노산과 비교하여 5개 이하의 치환된 아미노산을 갖거나, 또는 천연 GLP-1(7-37)OH 또는 GLP-1(7-36)OH의 상응하는 아미노산과 비교하여 4개 이하의 치환된 아미노산을 갖는다. 더욱 더 바람직하게는, 상기 유사체가 본래의 GLP-1(7-37)OH 또는 GLP-1(7-36)OH에서 상응하는 아미노산과 비교하여 3개 이하의 치환된 아미노산을 갖는다. 가장 바람직하게는, 상기 유사체가 천연 GLP-1(7-37)OH의 상응하는 아미노산과 비교하여 2개 이하의 치환된 아미노산을 갖는다.

- <149> 화학식 2, 3, 4 및 5의 화합물은 응집되어 불용성 형태를 생성하는 경향이 낮은 것으로 밝혀졌다. 이러한 경향은, 상대적으로 작은 GLP-1 펩티드가 더 큰 단백질과 융합됨에도 불구하고 활성 구조를 유지해야 하는 융합 단백질에 있어서 또한 중요한 것이다. 본 발명의 융합 단백질에 포함된 바람직한 화학식 2, 3, 4 및 5의 GLP-1 화합물은, 위치 22의 글리신 및 바람직하게는 위치 8의 알라닌이 다른 아미노산으로 치환된 GLP-1 유사체 또는 GLP-1 유사체의 단편을 포함한다.
- <150> 위치 22가 아스파르트산, 글루탐산, 아르기닌 또는 리신일 경우, 바람직하게는 위치 8이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌이며, 더 바람직하게는 발린 또는 글리신이다. 위치 22가 시스테인산과 같은 설폰산일 경우, 바람직하게는 위치 8이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌이며, 더 바람직하게는 발린 또는 글리신이다.
- <151> 다른 바람직한 GLP-1 화합물에는, 바람직하게는 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌이고, 더 바람직하게는 발린 또는 글리신이며, 위치 30이 글루탐산, 아스파르트산, 세린 또는 히스티딘이고, 더 바람직하게는 글루탐산인 것을 제외한 GLP-1(7-37)OH의 서열을 갖는 화학식 4 (서열 5)의 GLP-1 유사체가 포함된다.
- <152> 다른 바람직한 GLP-1 화합물에는, 바람직하게는 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌이고, 더 바람직하게는 발린 또는 글리신이며, 위치 37이 히스티딘, 리신, 아르기닌, 트레오닌, 세린, 글루탐산, 아스파르트산, 트립토판, 티로신 또는 페닐알라닌이고, 더 바람직하게는 히스티딘인 것을 제외한 GLP-1(7-37)OH의 서열을 갖는 화학식 4 (서열 5)의 GLP-1 유사체가 포함된다.
- <153> 다른 바람직한 GLP-1 화합물에는, 바람직하게는 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌이고, 더 바람직하게는 발린 또는 글리신이며, 위치 22가 글루탐산, 리신, 아스파르트산 또는 아르기닌이고, 더 바람직하게는 글루탐산 또는 리신이며, 위치 23이 리신, 아르기닌, 글루탐산, 아스파르트산 및 히스티딘이고, 더 바람직하게는 리신 또는 글루탐산인 것을 제외한 GLP-1(7-37)OH의 서열을 갖는 화학식 4 (서열 5)의 GLP-1 유사체가 포함된다.
- <154> 다른 바람직한 GLP-1 화합물에는, 바람직하게는 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌이고, 더 바람직하게는 발린 또는 글리신이며, 위치 22가 글루탐산, 리신, 아스파르트산 또는 아르기닌이고, 더 바람직하게는 글루탐산 또는 리신이며, 위치 27이 알라닌, 리신, 아르기닌, 트립토판, 티로신, 페닐알라닌 또는 히스티딘이고, 더 바람직하게는 알라닌인 것을 제외한 GLP-1(7-37)OH의 서열을 갖는 화학식 5 (서열 6)의 GLP-1 유사체가 포함된다.
- <155> 다른 바람직한 GLP-1 화합물에는, 위치 8의 아미노산, 및 위치 9, 위치 11, 위치 12, 위치 16, 위치 18, 위치 22, 위치 23, 위치 24, 위치 26, 위치 27, 위치 30, 위치 31, 위치 33, 위치 34, 위치 35, 위치 36 및 위치 37로 구성된 군에서 선택된 1개, 2개 또는 3개 위치의 아미노산이 천연 GLP-1(7-37)OH에서의 상응하는 위치에 있는 아미노산과 상이한 것을 제외한 GLP-1(7-37)OH의 서열을 갖는 화학식 2의 GLP-1 유사체가 포함된다.

<156>

다른 바람직한 화학식 2의 GLP-1 화합물에는

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-

<157>

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, 및 Gly⁸-Lys²²-Glu²³-GLP-1(7-37)OH

<158>

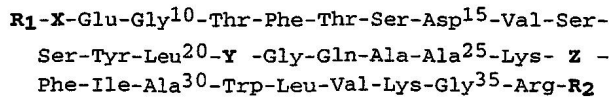
<159>

가 포함된다.

<160>

본 발명에 유용한 또다른 바람직한 GLP-1 유사체 및 유도체의 군은 화학식 6의 분자로 구성되어 있다 (서열 7).

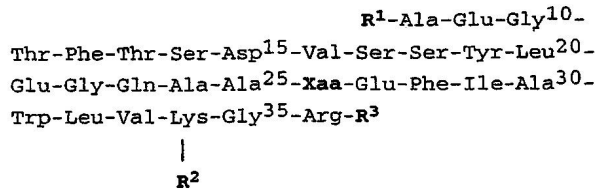
화학식 6



(서열 7)

- <161>
- <162> 상기 식에서,
- <163> R₁은 L-히스티딘, D-히스티딘, 데스아미노-히스티딘, 2-아미노히스티딘, β-히드록시-히스티딘, 호모히스티딘, 알파-플루오로메틸-히스티딘 및 알파-메틸-히스티딘으로 구성된 군에서 선택된 것이고;
- <164> X는 Ala, Gly, Val, Thr, Ile 및 알파-메틸-Ala으로 구성된 군에서 선택된 것이고;
- <165> Y는 Glu, Gln, Ala, Thr, Ser 및 Gly으로 구성된 군에서 선택된 것이고;
- <166> Z는 Glu, Gln, Ala, Thr, Ser 및 Gly으로 구성된 군에서 선택된 것이며;
- <167> R₂는 Gly-OH이다.
- <168> 본 발명에 유용한 또다른 바람직한 GLP-1 화합물의 군은 WO 91/11457호에 개시되어 있으며, 본질적으로 GLP-1(7-34), GLP-1(7-35), GLP-1(7-36) 또는 GLP-1(7-37), 또는 이들의 아마이드, 및
- <169> (a) 위치 26 및(또는) 위치 34의 리신에 대한 글리신, 세린, 시스테인, 트레오닌, 아스파라긴, 글루타민, 티로신, 알라닌, 발린, 이소루신, 루신, 메티오닌, 페닐알라닌, 아르기닌 또는 D-리신으로의 치환; 또는 위치 36의 아르기닌에 대한 글리신, 세린, 시스테인, 트레오닌, 아스파라긴, 글루타민, 티로신, 알라닌, 발린, 이소루신, 루신, 메티오닌, 페닐알라닌, 리신 또는 D-아르기닌으로의 치환;
- <170> (b) 위치 31의 트립토판에 대한 내산화성 아미노산으로의 치환;
- <171> (c) 위치 16의 발린에 대한 티로신으로의 치환 ; 위치 18의 세린에 대한 리신으로의 치환; 위치 21의 글루탐산에 대한 아스파르트산으로의 치환; 위치 22의 글리신에 대한 세린으로의 치환; 위치 23의 글루타민에 대한 아르기닌으로의 치환; 위치 24의 알라닌에 대한 아르기닌으로의 치환; 및 위치 26의 리신에 대한 글루타민으로의 치환 중 1개 이상의 치환; 및
- <172> (d) 위치 8의 알라닌에 대한 글리신, 세린 또는 시스테인으로의 치환; 위치 9의 글루탐산에 대한 아스파르트산, 글리신, 세린, 시스테인, 트레오닌, 아스파라긴, 글루타민, 티로신, 알라닌, 발린, 이소루신, 루신, 메티오닌 또는 페닐알라닌으로의 치환; 위치 10의 글리신에 대한 세린, 시스테인, 트레오닌, 아스파라긴, 글루타민, 티로신, 알라닌, 발린, 이소루신, 루신, 메티오닌 또는 페닐알라닌으로의 치환; 및 위치 15의 아스파르트산에 대한 글루탐산으로의 치환 중 1개 이상의 치환; 및
- <173> (e) 위치 7의 히스티딘에 대한 글리신, 세린, 시스테인, 트레오닌, 아스파라긴, 글루타민, 티로신, 알라닌, 발린, 이소루신, 루신, 메티오닌 또는 페닐알라닌, 또는 D- 또는 N-아실화된 또는 알킬화된 형태의 히스티딘으로의 치환;
- <174> 으로 구성된 군에서 선택된 1개 이상의 변형을 갖는 제약상 허용 가능한 염으로 구성되어 있으며, (a), (b), (d) 및 (e)에서의 치환된 아미노산은 임의로 D-형태일 수도 있으며, 위치 7에서의 치환된 아미노산은 임의로 N-아실화된 또는 N-알킬화된 형태일 수도 있다.
- <175> 투여된 GLP-1의 생체내에서의 빠른 불활성화가 효소 디펩티딜-펩티다제 IV (DPP IV)에 의한 것일 수 있기 때문에 (예를 들면, 문헌 [Mentlein, R., et al., *Eur. J. Biochem.*, 214: 829-835 (1993)] 참조), 융합 단백질내에서 DPP IV로부터 보호된 GLP-1 유사체 및 유도체가 바람직하며, 융합 단백질내 GLP-1 화합물이 Gly⁸-GLP-1(7-37)OH, Val⁸-GLP-1(7-37)OH, α-메틸-Ala⁸-GLP-1(7-37)OH 또는 Gly⁸-Gln²¹-GLP-1(7-37)OH인 것이 더 바람직하다.
- <176> 본 발명에 유용한 또다른 바람직한 GLP-1 화합물의 군은, 이 거명을 통해 본원에 참고문헌으로 포함되는 미국 특허 제5,512,549호에 개시되어 있는 화학식 7 (서열 8)의 화합물로 구성되어 있다.

화학식 7



(서열 8)

<177>

<178>

상기 식에서,

<179>

R^1 은 4-이미다조프로피오닐, 4-이미다조아세틸 및 4-이미다조- α , α -디메틸-아세틸로 구성된 군에서 선택된 것이고;

<180>

R^2 는 C_6 - C_{10} 비분지형 아실로 구성된 군에서 선택된 것이거나, 존재하지 않고;

<181>

R^3 은 Gly-OH 및 NH_2 로 구성된 군에서 선택된 것이며;

<182>

Xaa는 Lys 또는 Arg이다.

<183>

본 발명에 유용한 더 바람직한 화학식 4의 화합물에서는 Xaa가 Arg이고, R^2 가 C_6 - C_{10} 비분지형 아실이다. 본 발명에 유용한 더욱 더 바람직한 화학식 4의 화합물에서는 Xaa가 Arg이고, R^2 가 C_6 - C_{10} 비분지형 아실이며, R^3 이 Gly-OH이다. 본 발명에 유용한 다른 매우 바람직한 화학식 4의 화합물에서는 Xaa가 Arg이고, R^2 가 C_6 - C_{10} 비분지형 아실이고, R^3 이 Gly-OH이며, R^1 이 4-이미다조프로피오닐이다. 본 발명에 유용한 특히 바람직한 화학식 4의 화합물에서는 Xaa가 Arg이고, R^2 가 C_8 비분지형 아실이고, R^3 이 Gly-OH이며, R^1 이 4-이미다조프로피오닐이다.

<184>

바람직하게는 상기 GLP-1 화합물은, 유사체 또는 단편의 주쇄가 위치 8에서 알라닌 이외의 다른 아미노산을 갖는 GLP-1 유사체 (위치 8 유사체)를 포함한다. 또한 상기 주쇄의 위치 7에는 L-히스티딘, D-히스티딘, 또는 테스아미노-히스티딘, 2-아미노-히스티딘, β -히드록시-히스티딘, 호모히스티딘, α -플루오로메틸-히스티딘 또는 α -메틸-히스티딘과 같은 변형된 형태의 히스티딘도 포함될 수 있다. 바람직하게는 상기 위치 8 유사체가 천연 GLP-1(7-37)OH에서의 상응하는 아미노산과 비교하여 위치 12, 16, 18, 19, 20, 22, 25, 27, 30, 33 및 37에서 1개 이상의 추가적으로 치환된 아미노산을 갖는다. 더 바람직하게는 상기 위치 8 유사체가 천연 GLP-1(7-37)OH에서의 상응하는 아미노산과 비교하여 위치 16, 18, 22, 25 및 33에서 1개 이상의 추가적으로 치환된 아미노산을 갖는다.

<185>

바람직한 실시양태에서 상기 GLP-1 유사체는, 위치 12의 아미노산이 트립토판 및 티로신으로 구성된 군에서 선택된 GLP-1(7-37)OH이다. 더 바람직하게는 위치 12에서의 치환에 더하여, 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌으로 치환되고, 더 바람직하게는 발린 또는 글리신으로 치환된다. 더욱 더 바람직하게는 위치 12 및 8에서의 치환에 더하여, 위치 22의 아미노산이 글루탐산으로 치환된다.

<186>

또다른 바람직한 실시양태에서 상기 GLP-1 유사체는, 위치 16의 아미노산이 트립토판, 이소루신, 루신, 페닐알라닌 또는 티로신으로 구성된 군에서 선택된 GLP-1(7-37)OH이다. 더 바람직하게는 위치 16에서의 치환에 더하여, 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌으로 치환되고, 더 바람직하게는 발린 또는 글리신으로 치환된다. 더욱 더 바람직하게는 위치 16 및 8에서의 치환에 더하여, 위치 22의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 16 및 8에서의 치환에 더하여, 위치 30의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 16 및 8에서의 치환에 더하여, 위치 37의 아미노산이 히스티딘으로 치환된다.

<187>

또다른 바람직한 실시양태에서 상기 GLP-1 유사체는, 위치 18의 아미노산이 트립토판, 티로신, 페닐알라닌, 리신, 루신 및 이소루신 (바람직하게는, 트립토판, 티로신 및 이소루신)으로 구성된 군에서 선택된 GLP-1(7-37)OH

이다. 더 바람직하게는 위치 18에서의 치환에 더하여, 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌으로 치환되고, 더 바람직하게는 발린 또는 글리신으로 치환된다. 더욱 더 바람직하게는 위치 18 및 8에서의 치환에 더하여, 위치 22의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 18 및 8에서의 치환에 더하여, 위치 30의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 18 및 8에서의 치환에 더하여, 위치 37의 아미노산이 히스티딘으로 치환된다.

<188> 또다른 바람직한 실시양태에서 상기 GLP-1 유사체는, 위치 19의 아미노산이 트립토판 및 페닐알라닌으로 구성된 군에서 선택된 (바람직하게는, 트립토판인) GLP-1(7-37)OH이다. 더 바람직하게는 위치 19에서의 치환에 더하여, 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌으로 치환되고, 더 바람직하게는 발린 또는 글리신으로 치환된다. 더욱 더 바람직하게는 위치 19 및 8에서의 치환에 더하여, 위치 22의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 19 및 8에서의 치환에 더하여, 위치 30의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 19 및 8에서의 치환에 더하여, 위치 37의 아미노산이 히스티딘으로 치환된다.

<189> 또다른 바람직한 실시양태에서 상기 GLP-1 유사체는, 위치 20의 아미노산이 페닐알라닌, 티로신 또는 트립토판인 GLP-1(7-37)OH이다. 더 바람직하게는 위치 20에서의 치환에 더하여, 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌으로 치환되고, 더 바람직하게는 발린 또는 글리신으로 치환된다. 더욱 더 바람직하게는 위치 20 및 8에서의 치환에 더하여, 위치 22의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 20 및 8에서의 치환에 더하여, 위치 30의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 20 및 8에서의 치환에 더하여, 위치 37의 아미노산이 히스티딘으로 치환된다.

<190> 또다른 바람직한 실시양태에서 상기 GLP-1 유사체는, 위치 25의 아미노산이 발린, 이소루신 및 루신으로 구성된 군에서 선택된 (바람직하게는, 발린인) GLP-1(7-37)OH이다. 더 바람직하게는 위치 25에서의 치환에 더하여, 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌으로 치환되고, 더 바람직하게는 발린 또는 글리신으로 치환된다. 더욱 더 바람직하게는 위치 25 및 8에서의 치환에 더하여, 위치 22의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 25 및 8에서의 치환에 더하여, 위치 30의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 25 및 8에서의 치환에 더하여, 위치 37의 아미노산이 히스티딘으로 치환된다.

<191> 또다른 바람직한 실시양태에서 상기 GLP-1 유사체는, 위치 27의 아미노산이 이소루신 및 알라닌으로 구성된 군에서 선택된 GLP-1(7-37)OH이다. 더 바람직하게는 위치 27에서의 치환에 더하여, 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌으로 치환되고, 더 바람직하게는 발린 또는 글리신으로 치환된다. 더욱 더 바람직하게는 위치 27 및 8에서의 치환에 더하여, 위치 22의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 27 및 8에서의 치환에 더하여, 위치 30의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 27 및 8에서의 치환에 더하여, 위치 37의 아미노산이 히스티딘으로 치환된다.

<192> 또다른 바람직한 실시양태에서 상기 GLP-1 유사체는, 위치 33의 아미노산이 이소루신인 GLP-1(7-37)OH이다. 더 바람직하게는 위치 33에서의 치환에 더하여, 위치 8의 아미노산이 글리신, 발린, 루신, 이소루신, 세린, 트레오닌 또는 메티오닌으로 치환되고, 더 바람직하게는 발린 또는 글리신으로 치환된다. 더욱 더 바람직하게는 위치 33 및 8에서의 치환에 더하여, 위치 22의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 33 및 8에서의 치환에 더하여, 위치 30의 아미노산이 글루탐산으로 치환된다. 또한 바람직하게는 위치 33 및 8에서의 치환에 더하여, 위치 37의 아미노산이 히스티딘으로 치환된다.

<193> GLP-1 화합물은 1개 이상의 하기 위치에서 아미노산이 치환되었다: 8, 12, 16, 18, 19, 20, 22, 25, 27, 30, 33 및 37. 이러한 GLP-1 화합물은 GLP-1(7-37)OH와 비교하여 증가된 효능을 나타내며, 화학식 9 (서열 12)의 아미노산 서열을 포함하고 있다.

화학식 9

**Xaa₇-Xaa₈-Glu-Gly-Thr-Xaa₁₂-Thr-Ser-Asp-Xaa₁₆-Ser-
Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Gln-Ala-Xaa₂₅-Lys-Xaa₂₇-
Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Lys-Gly-Arg-Xaa₃₇**

(서열 12)

<194>

<195> 상기 식에서,

- <196> Xaa₇은 L-히스티딘, D-히스티딘, 데스아미노-히스티딘, 2-아미노-히스티딘, β -히드록시-히스티딘, 호모히스티딘, α -플루오로메틸-히스티딘 또는 α -메틸-히스티딘이고;
- <197> Xaa₈은 Ala, Gly, Val, Leu, Ile, Ser 또는 Thr이고;
- <198> Xaa₁₂는 Phe, Trp 또는 Tyr이고;
- <199> Xaa₁₆은 Val, Trp, Ile, Leu, Phe 또는 Tyr이고;
- <200> Xaa₁₈은 Ser, Trp, Tyr, Phe, Lys, Ile, Leu 또는 Val이고;
- <201> Xaa₁₉는 Tyr, Trp 또는 Phe이고;
- <202> Xaa₂₀은 Leu, Phe, Tyr 또는 Trp이고;
- <203> Xaa₂₂는 Gly, Glu, Asp 또는 Lys이고;
- <204> Xaa₂₅는 Ala, Val, Ile 또는 Leu이고;
- <205> Xaa₂₇은 Glu, Ile 또는 Ala이고;
- <206> Xaa₃₀은 Ala 또는 Glu이고;
- <207> Xaa₃₃은 Val 또는 Ile이며;
- <208> Xaa₃₇은 Gly, His 또는 NH₂이거나, 또는 존재하지 않는다.

<209>

바람직한 화학식 9의 GLP-1 화합물에는

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-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 및 Thr⁸-His³⁷-GLP-1(7-36)NH₂

<210>

가 포함된다.

<211>

<212>

다중 치환된 바람직한 화학식 9의 GLP-1 화합물에는, 위치 8이 발린 또는 글리신이고, 위치 22가 글루탐산이고, 위치 16이 티로신, 루신 또는 트립토판이고, 위치 18이 티로신, 트립토판 또는 이소루신이고, 위치 25가 발린이며, 위치 33이 이소루신인 GLP-1(7-37)OH가 포함된다. 다른 바람직한 GLP-1 화합물에는 하기의 화합물이 포함된다:

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, 및 Val⁸-Ile¹⁸-Glu²²-GLP-1(7-37)OH.

<213>

<214>

또한 본 발명의 GLP-1 화합물에는 엑센딘 화합물이 포함된다. 엑센딘-3 및 엑센딘-4는 독도마뱀과 도마뱀 독소

(Helodermatidae lizard venom)에서 최초로 단리된, 생물학적 활성을 갖는 펩티드이며, 이 화합물은 GLP-1 수용체와 결합하고, 포유동물 벽세포 (parietal cells)내에서 cAMP-의존성 H^+ 생성을 자극하는 것으로 알려져 있다. 엑센딘-3과 엑센딘-4는 각각 39개의 아미노산으로 이루어진 펩티드이며, GLP-1과 약 53 %의 상동성을 갖는다. 이들은 GLP-1 활성화에 대해 강력한 효능제 (agonist)로서 작용한다. 특히 엑센딘(9-39 아미노산)으로 알려진, N-말단이 절단된 엑센딘 유도체는 엑센딘-3, 엑센딘-4 및 GLP-1의 억제제이다.

<215> 전형적으로 엑센딘 화합물은 엑센딘-3, 엑센딘-4, 또는 이들의 유사체 또는 단편의 서열을 갖는 폴리펩티드를 포함한다. 엑센딘-3 및 엑센딘-4는 미국특허 제5,424,286호에 개시되어 있다.

<216> 엑센딘-3은 서열 9의 아미노산 서열을 갖는다.

```

7   8   9  10  11  12  13  14  15  16  17
His-Ser-Asp-Gly-Thr-Phe-Thr-Ser-Asp-Leu-Ser-
18  19  20  21  22  23  24  25  26  27  28
Lys-Gln-Met-Glu-Glu-Glu-Ala-Val-Arg-Leu-Phe-
29  30  31  32  33  34  35  36  37  38  39
Ile-Glu-Trp-Leu-Lys-Asn-Gly-Gly-Pro-Ser-Ser-
40  41  42  43  44  45
Gly-Ala-Pro-Pro-Pro-Ser

```

(서열 9)

<217>

<218> 엑센딘-4는 서열 10의 아미노산 서열을 갖는다.

```

7   8   9  10  11  12  13  14  15  16  17
His-Gly-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Leu-Ser-
18  19  20  21  22  23  24  25  26  27  28
Lys-Gln-Met-Glu-Glu-Glu-Ala-Val-Arg-Leu-Phe-
29  30  31  32  33  34  35  36  37  38  39
Ile-Glu-Trp-Leu-Lys-Asn-Gly-Gly-Pro-Ser-Ser-
40  41  42  43  44  45
Gly-Ala-Pro-Pro-Pro-Ser

```

(서열 10)

<219>

<220> GLP-1 화합물에는, 엑센딘 또는 엑센딘 유사체의 N-말단 및(또는) C-말단에서 1개 이상의 아미노산을 절단한 후 얻어진 폴리펩티드로서의 엑센딘 단편도 포함된다. 또한, GLP-1 화합물에는 엑센딘 또는 엑센딘 단편의 N-말단 및(또는) C-말단에 1개 이상의 아미노산이 부가된 엑센딘 폴리펩티드도 포함된다. 이러한 형태의 엑센딘 화합물은 약 45개 이하의 아미노산을 갖는다.

<221> 또한 GLP-1 화합물에는 "엑센딘 유사체"가 포함된다. 엑센딘 유사체는 엑센딘-4, 엑센딘-3 또는 이것의 단편과 충분한 상동성을 갖고 있어서, 이 유사체는 인슐린 분비 활성을 갖는다. 엑센딘 단편 및(또는) 유사체의 활성은 EP 제619,322호 및 미국특허 제5,120,712호에 기재된 것과 같은 시험관내 분석을 이용하여 평가할 수 있다.

<222> 바람직하게는, 엑센딘 유사체가 엑센딘-4 또는 엑센딘-4 단편에서 상응하는 위치에 있는 아미노산과 1개, 2개, 3개, 4개 또는 5개의 아미노산이 서로 상이하도록 변형된, 엑센딘-4 또는 엑센딘-4 단편의 아미노산 서열을 갖는다. 엑센딘 화합물을 나타내기 위해 본원에 사용된 명명법에서, 치환된 새로운 아미노산 및 이것의 위치를 원래 구조의 앞에 나타내었다. 예를 들면, Val⁸-엑센딘-4는 엑센딘-4의 위치 8에서 정상적으로 발견되는 글리신이 발린으로 치환된 엑센딘 화합물을 나타낸다.

<223> 또다른 바람직한 GLP-1 화합물의 군은 화학식 8 (서열 11)의 GLP-1/엑센딘-4 유사체로 구성되어 있다.

화학식 8

7 8 9 10 11 12 13 14 15 16 17
Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Xaa-Ser-
18 19 20 21 22 23 24 25 26 27 28
Xaa-Xaa-Xaa-Glu-Xaa-Xaa-Ala-Xaa-Xaa-Xaa-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Trp-Leu-Xaa-Xaa-Gly-Xaa-R

(서열 11)

<224>

상기 식에서,

<225>

위치 7의 Xaa는 L-히스티딘, D-히스티딘, 데스아미노-히스티딘, 2-아미노-히스티딘, β -히드록시-히스티딘, 호모히스티딘, α -플루오로메틸-히스티딘 또는 α -메틸-히스티딘이고;

<226>

위치 8의 Xaa는 Gly, Ala 또는 Val이고;

<227>

위치 16의 Xaa는 Leu 또는 Val이고;

<228>

위치 18의 Xaa는 Lys 또는 Ser이고;

<229>

위치 19의 Xaa는 Gln 또는 Tyr이고;

<230>

위치 20의 Xaa는 Met 또는 Leu이고;

<231>

위치 22의 Xaa는 Glu 또는 Gln이고;

<232>

위치 23의 Xaa는 Glu 또는 Gln이고;

<233>

위치 25의 Xaa는 Val 또는 Ala이고;

<234>

위치 26의 Xaa는 Arg 또는 Lys이고;

<235>

위치 27의 Xaa는 Leu 또는 Glu이고;

<236>

위치 30의 Xaa는 Glu 또는 Ala이고;

<237>

위치 33의 Xaa는 Val 또는 Lys이고;

<238>

위치 34의 Xaa는 Asn 또는 Lys이고;

<239>

위치 36의 Xaa는 Gly 또는 Arg이며;

<240>

위치 37의 R은 Gly, Pro 또는 Pro-Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser이거나, 또는 존재하지 않는다. 상기 부류의 범위내에 포함되는 서로 다른 18종의 활성을 표 6에 나타내었다.

<241>

본 발명에 유용한 추가의 엑센딘-유사체가 PCT 특허 공개 WO 99/25728호 (Beeley et al.), WO 99/25727호 (Beeley et al.), WO 98/05351호 (Young et al.), WO 99/40788호 (Young et al.), WO 99/07404호 (Beeley et al.) 및 WO 99/43708호 (Knudsen et al.)에 기재되어 있다.

<242>

본 발명의 GLP-1 융합 단백질은 글리코실화 반응 위치를 포함할 수 있다. 글리코실화 반응은 당 잔기가 단백질의 특정 위치에 부가되는 화학적 변형이다. 단백질의 글리코실화 반응은 성숙 단백질의 올바른 전하, 구조 및 안정성을 보장해주는 역할을 하고, 단백질을 세포 표면으로 표적화하여 결국 이 단백질을 분비하게 할 수 있다. 가장 중요한 것은, 글리코실화 반응이 많은 단백질들의 생체내 소멸 속도에 영향을 준다는 것이다. 당은 O-연결된 또는 N-연결된 형태일 수 있다. 일반적으로, O-연결된 당은 세린 및 트레오닌의 히드록실기 산소에 부가되는 반면에, N-연결된 당은 아스파라긴의 아미드 질소에 부가된다. N-글리코실화 반응이 일어나는 컨센서스 부위 (consensus site)는 Asn-X1-X2이며, 여기에서 X1은 Pro를 제외한 임의의 아미노산이고, X2는 Ser 또는 Thr이다.

<243>

일반적으로 GLP-1 화합물은 생체내에서 글리코실화되지 않는다. 그러나 흥미있게도, Fc 서열에 융합된 C-말단 연장 부분을 갖는 GLP-1 화합물을 포함하는 본 발명의 GLP-1 융합 단백질은 C-말단 연장 부분의 마지막 세린 (SSGAPPPS*)과, Fc의 N-말단 영역내 11 위치의 트레오닌 (AEPKSCDKTHT*CPPC...)에서 글리코실화된다.

<244>

- <245> 이종성 Fc 융합 단백질
- <246> 상기 GLP-1 화합물은 면역글로블린의 Fc 부분에 직접적으로 또는 펩티드 링커를 통해 융합될 수 있다.
- <247> 면역글로블린은 이황결합에 의해 서로 결합된 폴리펩티드 쇄를 지니고 있으며, 전형적으로 2개의 경쇄와 2개의 중쇄를 갖는 분자이다. 각각의 쇄에 있어서, 한 도메인 (V)은 이 분자의 항체 특이성에 따라 가변적인 아미노산 서열을 갖고 있다. 다른 도메인 (C)은 오히려, 동일 군내의 여러 분자들간에 서로 공통적인 불변 서열을 갖고 있다.
- <248> 본원에 사용된, 면역글로블린의 Fc 부분은 면역학 분야에서 일반적으로 사용되는 용어의 의미를 갖는다. 구체적으로, 이 용어는 항체에서 2개의 항원 결합 영역 (Fab 단편)를 제거하여 얻어진 항체 단편을 말한다. 상기 Fab 단편을 제거하는 한 방법으로는 파파인 (papain) 단백질 분해효소를 이용하여 면역글로블린을 분해하는 것이 있다. 따라서 상기 Fc 부분은 비공유결합성 상호작용 및 이황결합을 통해 서로 결합되어 있는 양쪽 중쇄로부터, 거의 동일한 크기를 갖는 불변 영역의 단편으로 형성된다. 상기 Fc 부분에는 힌지 (hinge) 영역이 포함될 수 있고, CH2 및 CH3 도메인을 통해 항체의 C-말단까지 연장될 수 있다. 인간 및 마우스 면역글로블린에 대한 대표적인 힌지 영역은, 이 거명을 통해 본원에 참고문헌으로 포함되는 문헌 [Antibody Engineering, A Practical Guide, Borrebaeck, C. A. K., ed., W. H. Freeman and Co., 1992]에서 찾아볼 수 있다. Fc 부분에는 1개 이상의 글리코실화 반응 위치가 더 포함될 수 있다. 힌지 영역, CH2 및 CH3 도메인, 및 1개의 N-글리코실화 반응 위치 (위치 82)를 지닌 대표적인 Fc 단백질의 아미노산 서열을 도 1에 나타내었다.
- <249> 인간 면역글로블린 Fc 영역에는 서로 다른 이펙터 및 약물동력학적 특성을 갖는 5종류 (IgG, IgA, IgM, IgD 및 IgE)가 있다. IgG는 혈청내 가장 풍부한 면역글로블린이다. 또한 IgG는 혈청내 모든 면역글로블린 중 가장 긴 반감기 (23일)를 갖고 있다. 다른 면역글로블린과는 달리, IgG는 Fc 수용체에 결합한 후 효율적으로 재순환된다. IgG의 하위 군으로서 G1, G2, G3 및 G4의 4가지가 있으며, 이들 각각은 서로 다른 이펙터 기능을 갖고 있다. G1, G2 및 G3는 C1q와 결합하여 보체를 고정시킬 수 있으나, G4는 그렇지 못하다. 비록 G3가 G1보다 더 효율적으로 C1q와 결합할 수는 있지만, G1이 보체-유도된 세포의 용해를 더 효과적으로 매개한다. G2는 보체를 매우 비효율적으로 고정시킨다. IgG내 C1q 결합 위치는 CH2 도메인의 카르복시 말단 영역에 존재한다.
- <250> 모든 IgG 하위 군은 Fc 수용체 (CD16, CD32, CD64)와 결합할 수 있으며, G1 및 G3이 G2 및 G4보다 더 효과적으로 결합한다. IgG의 Fc 수용체 결합 영역은 힌지 영역 및 CH2 도메인의 카르복시 말단 영역에 존재하는 잔기들에 의해 형성된다.
- <251> IgA는 모노머 형태 및 J-쇄에 의해 서로 결합된 다이머 형태로 존재할 수 있다. IgA는 혈청내 두번째로 가장 풍부한 Ig이지만, 단지 6일의 반감기를 갖고 있다. IgA는 3종의 이펙터 기능을 갖고 있다. 이것은 대식세포 및 호산구 (eosinophil) 상의 IgA 특이적 수용체와 결합하여, 각각 식균작용 및 탈과립 (degranulation)을 유발한다. 이것은 또한 알려지지 않은 다른 경로를 통해 보체를 고정시킬 수도 있다.
- <252> IgM은 펜타머이거나 또는 헥사머로서 발현되며, 두 경우 모두 J-쇄에 의해 결합되어 있다. IgM의 혈청내 반감기는 5일이다. 이것은 CH3 도메인내에 존재하는 결합 부위를 통해 C1q와 약하게 결합한다. IgD의 혈청내 반감기는 3일이다. IgD의 이펙터 기능은 명확하지 않다. IgE는 모노머성 Ig이고, 혈청내 반감기가 2.5일이다. IgE는 2개의 Fc 수용체와 결합하여 탈과립을 유발하고, 그 결과 전염증성 물질 (proinflammatory agent)의 방출을 초래한다.
- <253> 원하는 생체내 효과에 따라, 본 발명의 이종성 융합 단백질은 상기한 모든 이소형 (isotype)을 지닐 수 있거나, 또는 보체 및(또는) Fc 수용체 결합 기능이 바뀐 돌연변이된 Fc 영역을 지닐 수 있다. 따라서, 본 발명의 이종성 융합 단백질은 면역글로블린의 Fc 전체 부분, 면역글로블린 Fc 부분의 단편, 또는 GLP-1 화합물과 융합된 이들의 유사체를 지닐 수 있다.
- <254> 본 발명의 융합 단백질은 단일쇄 단백질로 구성되거나, 또는 다중쇄 폴리펩티드로서 존재할 수 있다. 2개 이상의 Fc 융합 단백질들이, Fc 영역간에 자연적으로 형성되는 이황결합을 통해 상호작용을 하도록 생성될 수 있다. 이러한 멀티머는 동종의 GLP-1 화합물들을 갖거나, 또는 상기 융합 단백질 Fc 부분의 N-말단에 융합된 서로 다른 GLP-1 화합물들을 가질 수 있다.
- <255> 상기 융합 단백질의 최종 구조와 상관없이, 상기 Fc 또는 Fc-유사 영역은 N-말단에 융합된 GLP-1 화합물의 생체내 혈장 반감기를 연장시키는 역할을 해야 한다. 또한, 상기 융합된 GLP-1 화합물은 몇몇 생물학적 활성을 보아야 한다. 반감기의 증가는, 융합 단백질의 반감기를 GLP-1 화합물만의 반감기와 비교한 실시예 7에 기재된

방법을 이용하여 증명할 수 있다. 생물학적 활성은 당업계에 공지된 시험관내 및 생체내 방법에 의해 측정될 수 있다. 대표적인 생물학적 분석을 실시예 6, 8 및 9에 기재하였다.

<256> 단백질 분해로 생성된 IgG의 Fc 영역은 완전한 IgG 분자와 동일한 생체내 반감기를 갖고 Fab 단편은 빠르게 분해되기 때문에, 반감기의 연장과 관련된 서열이 CH2 및(또는) CH3 도메인내에 존재하는 것으로 생각된다. 또한 고친화도 Fc 수용체 또는 C1q와 결합하지 않는 IgG 변이체들의 이화 속도는 야생형 모 항체의 소멸 속도와 구별할 수 없다는 것이 문헌상에 알려져 있으며, 이것은 상기 이화 부위가 Fc 수용체 또는 C1q 결합에 관여하는 부위와 구별됨을 시사한다 (Wawrzynczak et al., (1992) *Molecular Immunology* 29: 221). 쥐과 (murine) IgG1 Fc 영역을 이용한 위치-지정 돌연변이 (site-directed mutagenesis) 연구를 통해, IgG1 Fc 영역내 이화 속도를 조절하는 부위가 CH2-CH3 도메인의 경계면에 존재함을 밝혀냈다.

<257> 이러한 연구에 근거하여, 융합 단백질의 반감기를 최적화하기 위해 Fc 영역의 이화 부위를 변형할 수 있다. 본 발명의 이중성 융합 단백질에 사용된 Fc 영역은 IgG1 또는 IgG4 Fc 영역에서 유래된 것이 바람직하다. 상기 Fc 영역은 IgG4이거나 또는 IgG4에서 유래된 것이 더욱 바람직하다. 바람직하게는 IgG Fc 영역이 힌지 영역이 포함된, CH2 영역과 CH3 영역을 모두 지니고 있다.

<258> 이중성 알부민 융합 단백질

<259> 상기 GLP-1 화합물은 알부민, 또는 이들의 유사체, 단편 또는 유도체와 직접적으로 또는 펩티드 링커를 통해 융합될 수 있다.

<260> 일반적으로, 본 발명의 융합 단백질의 일부를 구성하고 있는 알부민 단백질은 임의의 종 (species)에서 클로닝된 알부민으로부터 유래할 수 있다. 그러나, 상기 융합 단백질이 인간에서 면역원성으로 되는 위험을 줄이기 위해 인간 알부민, 및 이것의 단편 및 유사체가 바람직하다. 인간 혈청 알부민 (HSA)은 분자량이 66,500인, 585개 아미노산의 비-글리코실화된 폴리펩티드 단일 쇄로 구성되어 있다. 인간 HSA의 아미노산 서열을 도 2에 나타내었다 (문헌 [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] 참조). 알부민의 다형성 변이체 (polymorphic variant)뿐만 아니라 이것의 유도체 및 단편도 기재되었다 (문헌 [Weitkamp, et al., (1973) *Ann. Hum. Genet.* 37: 219] 참조). 예를 들면, EP 제322,094호에는 HSA의 다양한 짧은 형태가 개시되어 있다. 이러한 단편에는 HSA(1-373), HSA(1-388), HSA(1-389), HSA(1-369) 및 HSA(1-419), 및 1-369와 1-419 사이의 단편이 포함된다. EP 제399,666호에는 HSA(1-177) 및 HSA (1-200), 및 HSA(1-177)과 HSA(1-200)사이의 단편을 비롯한 알부민 단편이 개시되어 있다.

<261> 본 발명의 이중성 융합 단백질에는, 단편, 유사체 및 유도체를 비롯한 임의의 알부민 단백질과 결합된 GLP-1 화합물이 포함되며, 여기에서 상기 융합 단백질은 생물학적으로 활성을 갖고 있고, 단일 GLP-1 화합물보다 더 긴 혈장내 반감기를 갖고 있는 것으로 이해된다. 따라서 융합 단백질의 알부민 부분은, 천연 인간 알부민의 혈장내 반감기와 반드시 동일한 혈장내 반감기를 지닐 필요가 없다. 단편, 유사체 및 유도체는 공지되어 있거나, 또는 천연 인간 알부민 및 목적 GLP-1 화합물의 반감기보다 더 긴 반감기 또는 중간정도의 반감기를 갖도록 제조될 수 있다.

<262> 본 발명의 이중성 융합 단백질에는, GLP-1 화합물 및(또는) 상기 융합 단백질의 Fc 또는 알부민 부분에서 보존적 아미노산 치환을 갖는 단백질이 포함된다. "보존적 치환"은 임의의 아미노산을, 동일한 전기적 순전하를 갖고 거의 동일한 크기 및 형태를 갖는 다른 아미노산으로 치환하는 것이다. 지방족 또는 치환된 지방족 아미노산 측쇄를 갖는 아미노산들은, 측쇄내 탄소 및 헤테로원자의 총 개수가 약 4개 이하로 상이할 경우, 거의 동일한 크기를 갖는다. 이들은 측쇄내의 분지쇄 개수가 1개 이하로 상이할 경우, 거의 동일한 형태를 갖는다. 측쇄에 페닐 또는 치환된 페닐기를 갖는 아미노산들은 거의 동일한 크기 및 형태를 갖는 것으로 간주된다. 본원에서 특별히 언급한 경우를 제외하고는, 보존적 치환은 바람직하게는 자연적으로 발생하는 아미노산으로 이루어진다.

<263> 그러나, 용어 "아미노산"은 가장 넓은 의미로 본원에 사용되었으며, 자연적으로 발생하는 아미노산뿐만 아니라, 아미노산 유사체 및 유도체를 비롯한 비-천연 발생의 아미노산도 포함된다. 비-천연 발생의 아미노산에는 아미노산 부분을 지닌 분자가 포함된다. 상기 광범위한 정의의 관점에서 당업계의 숙련자는, 본원에서 언급하는 아미노산에, 예를 들면 자연적으로 발생하는 프로테오게닉 (proteogenic) L-아미노산; D-아미노산; 아미노산 유사체 및 유도체와 같은, 화학적으로 변형된 아미노산; 노르루신, β -알라닌, 오르니틴, GABA 등과 같은 자연적으로 발생하는 비-프로테오게닉 아미노산; 및 아미노산의 특징적 성질을 갖는 것으로 당업계에 알려진, 화학적으

로 합성된 화합물이 포함되는 것을 알 것이다. 본원에 사용된 용어 "프로테오게닉"은 아미노산이 대사 경로를 통해 세포내에서 펩티드, 폴리펩티드 또는 단백질로 혼입될 수 있음을 나타내는 것이다.

<264> 합성된 비-천연 아미노산, 치환된 아미노산 또는 1개 이상의 D-아미노산을 비롯한 비-천연 아미노산을 본 발명의 이중성 융합 단백질내로 혼입하는 것은 다방면으로 유리할 수 있다. D-아미노산 함유 펩티드 등은 L-아미노산 함유 상대물에 비해, 시험관내 또는 생체내에서 증가된 안정성을 나타낸다. 따라서, D-아미노산이 혼입된 펩티드를 제조하는 것은, 세포내 높은 안정성이 요구되거나 필요할 경우에 특히 유용하다. 더 구체적으로는, D-펩티드 등은 내생성 펩티드 분해효소 및 단백질 분해효소에 대해 저항성을 갖고 있어서, 상기 분자의 향상된 생물학적 이용도를 제공하고, 필요한 경우 생체내에서 연장된 수명을 제공한다. 부가적으로 D-펩티드 등은, T 헬퍼 세포에 대한 주요 조직적합성 복합체 (major histocompatibility complex) 클래스 II-제한된 제시 (presentation)에 의해 효과적으로 프로세싱될 수 없으며, 따라서 모든 생물체에서 체액성 면역 반응을 유도할 가능성이 낮아진다.

<265> 본 발명에 포함된 다양한 폴리펩티드의 구조/기능 분석 외에도, 치환에 사용되는 아미노산을 선택할 경우 많은 인자들이 고려될 수 있다. 이러한 치환을 수행할 때 고려될 수 있는 한 인자로는 아미노산의 하이드로파틱 인덱스 (hydropathic index)이다. 상호작용의 생물학적 기능을 단백질에 부여하는 것에 대한 아미노산의 하이드로파틱 인덱스의 중요성은 키트 (Kyte) 및 돌리틀 (Doolittle)에 의해 논의되었다 (1982, *J. Mol. Biol.*, 157: 105-132). 아미노산의 상대적인 하이드로파틱 특성이 최종 단백질의 2차 구조에 기여하는 것으로 알려졌다. 즉, 상기 인덱스는 단백질의 효소, 기질, 수용체, 리간드, DNA, 항체, 항원 등과 같은 분자와의 상호작용에 영향을 준다. 아미노산의 소수성 및 전하 특성에 기초하여, 각각의 아미노산의 하이드로파틱 인덱스를 하기와 같이 지정하였다: 이소루신 (+4.5); 발린 (+4.2); 루신 (+3.8); 페닐알라닌 (+2.8); 시스테인/시스틴 (+2.5); 메티오닌 (+1.9); 알라닌 (+1.8); 글리신 (-0.4); 트레오닌 (-0.7); 세린 (-0.8); 트립토판 (-0.9); 티로신 (-1.3); 프롤린 (-1.6); 히스티딘 (-3.2); 글루타메이트/글루타민/아스파르테이트/아스파라긴 (-3.5); 리신 (-3.9); 및 아르기닌 (-4.5).

<266> 당업계에 공지된 바와 같이, 펩티드, 폴리펩티드 또는 단백질내의 특정 아미노산은 유사한 하이드로파틱 인덱스 또는 스코어를 지닌 다른 아미노산으로 치환되어, 유사한 또는 향상된 생물학적 활성을 지닌 펩티드 등을 생성할 수 있다. 이러한 치환을 수행하는 데 있어서, ± 2 이내의 하이드로파틱 인덱스를 갖는 아미노산들이 서로 치환되는 것이 바람직하다. 더 바람직한 치환은, 상기 아미노산들이 ± 1 이내의 하이드로파틱 인덱스를 갖는 것이다. 가장 바람직한 치환은, 상기 아미노산들이 ± 0.5 이내의 하이드로파틱 인덱스를 갖는 것이다.

<267> 유사 아미노산들도 또한 친수성에 기초하여 치환될 수 있다. 미국특허 제4,554,101호에는 단백질 인접 아미노산들의 친수성 정도에 의해 결정되는, 단백질의 최대 국소 평균 친수성이 그 단백질의 생물학적 성질과 연관되어 있음이 개시되어 있다. 하기의 친수성 값이 아미노산에 대해 지정되었다: 아르기닌/리신 (+3.0); 아스파르테이트/글루타메이트 (+3.0 $\pm$ 1); 세린 (+0.3); 아스파라긴/글루타민 (+0.2); 글리신 (0); 트레오닌 (-0.4); 프롤린 (-0.5 $\pm$ 1); 알라닌/히스티딘 (-0.5); 시스테인 (-1.0); 메티오닌 (-1.3); 발린 (-1.5); 루신/이소루신 (-1.8); 티로신 (-2.3); 페닐알라닌 (-2.5); 및 트립토판 (-3.4). 따라서, 펩티드, 폴리펩티드 또는 단백질내의 특정 아미노산은 유사한 친수성 스코어를 지닌 다른 아미노산에 의해 치환되어, 유사한 생물학적 활성, 즉 올바른 생물학적 기능을 여전히 보유한 펩티드 등을 생성할 수 있다.

<268> 이러한 치환을 수행하는 데 있어서, ± 2 이내의 하이드로파틱 인덱스를 지닌 아미노산들이 서로 치환되는 것이 바람직하고, ± 1 이내의 경우가 더 바람직하며, ± 0.5 이내의 경우가 가장 바람직하다.

<269> 상기한 바와 같이, 본 발명의 융합 단백질에서의 아미노산 치환은 예를 들면, 아미노산 측쇄 치환체들의 소수성, 친수성, 전하, 크기 등과 같은 아미노산 측쇄 치환체들의 상대적 유사성에 기초할 수 있다. 또한, 치환은 2차 구조 성향에 기초하여 수행될 수도 있다. 예를 들면, 나선형 아미노산은 나선형 구조를 유지하는 아미노산과 교체될 수 있다.

<270> 본 발명의 펩티드내에서 사일런트 (silent) 치환 등을 유발하는 보존적 아미노산 치환을 위한, 앞서 언급한 다양한 특성들이 고려된 예시적인 치환기는, 자연적으로 발생하는 아미노산들이 포함된 군의 다른 구성원들로부터 선택될 수 있다. 아미노산은 하기 4개 군으로 분류될 수 있다: (1) 산성 아미노산; (2) 염기성 아미노산; (3) 중성 극성 아미노산; 및 (4) 중성 비극성 아미노산.

<271> 본 발명의 이중성 융합 단백질을 제조하기 위한 일반적 방법

<272> 비록 본 발명의 이중성 융합 단백질은 다양한 다른 방법에 의해 제조될 수 있지만, 재조합에 의한 방법이 바람

직하다. 본 발명을 위해, 본원에 개시되고 청구된 바와 같이, 하기의 일반적인 분자생물학 용어 및 약어를 다음과 같이 정의하였다. 본원에 사용된 용어 및 약어는 별도로 지정하지 않는 한, 그의 일반적 의미를 갖는다. 예를 들면, "°C"는 섭씨 온도를 나타내고; "mmol"은 밀리몰을 나타내고; "mg"은 밀리그램을 나타내고; "μg"은 마이크로그램을 나타내고; "ml 또는 mL"은 밀리리터를 나타내며; "μl 또는 μL"은 마이크로리터를 나타낸다. 아미노산의 약어는 37 C. F. R. § 1.822 (b) (2) (1994)에 기재된 바와 같다.

- <273> 본원에 사용된 "염기쌍" 또는 "bp"는 DNA 또는 RNA를 나타낸다. 약어 A, C, G 및 T는 DNA 분자에서, 데옥시리보뉴클레오시드인 (데옥시)아데노신, (데옥시)시티딘, (데옥시)구아노신 및 티미딘의 5'-모노포스페이트 형태에 각각 상응하는 것이다. 약어 U, C, G 및 A는 RNA 분자에서, 리보뉴클레오시드인 우리딘, 시티딘, 구아노신 및 아데노신의 5'-모노포스페이트 형태에 각각 상응하는 것이다. 이중 가닥 DNA에 있어서, 염기쌍은 A와 T, 또는 C와 G의 결합관계를 나타낼 수 있다. DNA/RNA에 있어서, 이종이중 (heteroduplex) 염기쌍은 A와 U, 또는 C와 G의 결합관계를 나타낼 수 있다 (용어 "상보적인"의 하기 정의 참조).
- <274> DNA의 "분해" 또는 "제한"은, DNA내 특정 서열에서만 작용하는 제한 효소 ("서열-특이적 엔도뉴클레아제")를 이용한 DNA의 촉매적 절단을 나타낸다. 본원에 사용된 다양한 제한 효소들은 상업적으로 구입 가능하며, 이들의 반응 조건, 보조인자 및 그 외 필요한 사항들은 당업계의 숙련자들에게 공지된 바와 같이 사용하였다. 특정 제한 효소에 대한 적절한 완충액 및 기질의 양은 제조업자에 의해 설명되었거나, 또는 문헌상에서 쉽게 찾아볼 수 있다.
- <275> "라이게이션"은 2개의 이중가닥 핵산 단편들 사이에 포스포디에스테르 결합을 형성시키는 과정을 나타낸다. 별도로 명시하지 않는 한, 라이게이션은 T4 DNA 리가제 (ligase)와 같은 DNA 리가제를 사용하여 공지의 완충액 및 조건으로 수행될 수 있다.
- <276> "플라스미드"는 염색체 이외의 (일반적으로) 자기-복제성 유전 요소를 나타낸다. 플라스미드는 일반적으로 소문자 "p" 뒤에 문자 및(또는) 숫자로 나타낸다. 본원에서의 출발 플라스미드는 제한되지 않은 상태로 상업적 및 공개적으로 구입 가능하거나, 또는 공지의 방법을 이용하여 구입 가능한 플라스미드로부터 제조될 수 있다. 또한, 기재된 플라스미드와 동등한 플라스미드가 당업계에 공지되어 있으며, 이것은 보통의 숙련자들에게 명확할 것이다.
- <277> 본원에 사용된 "재조합 DNA 클로닝 벡터"는 1개 이상의 추가적 DNA 단편이 추가될 수 있거나 또는 추가된 DNA분자를 포함하는 플라스미드 및 파지 등을 비롯한, 임의의 자발적인 복제 물질을 나타낸다.
- <278> 본원에 사용된 "재조합 DNA 발현 벡터"는 삽입된 DNA의 전사를 조절하는 프로모터가 혼입된, 임의의 재조합 DNA 클로닝 벡터를 나타낸다.
- <279> "전사"는 DNA 뉴클레오타이드 서열에 내재하는 정보가 상보적 RNA 서열로 전달되는 과정을 나타낸다.
- <280> "형질감염 (transfection)"은 임의 코딩 서열의 사실상 발현 여부와 관계없이, 발현 벡터를 숙주세포내로 도입하는 것을 나타낸다. 예를 들면 인산칼슘 공침전법(co-precipitation), 리포좀 형질감염법 및 전기천공법(electroporation)과 같은, 수많은 형질감염 방법이 보통의 숙련자에게 공지되어 있다. 성공적인 형질감염은 일반적으로, 상기 벡터의 작동에 대한 임의의 표시가 나타날 때 인지된다.
- <281> "형질전환"은 생물체내로 도입된 DNA가 염색체 이외의 요소로서 복제되거나 또는 염색체내로 통합되어 복제되도록, DNA를 생물체내로 도입하는 것을 나타낸다. 박테리아성 및 진핵성 숙주를 형질전환시키는 방법은 당업계에 잘 알려져 있으며, 핵 주입법, 원형질체 융합법, 또는 염화칼슘을 사용한 칼슘 처리에 의한 방법과 같은 많은 방법들이 문헌 [J. Sambrook, et al., Molecular Cloning: A Laboratory Manual, (1989)]에 요약되어 있다. 일반적으로 DNA를 효모내로 도입할 경우에는 용어 "형질감염"에 대립하는 것으로 용어 "형질전환"을 사용한다.
- <282> 본원에 사용된 "번역"은 전령 (messenger) RNA (mRNA)의 유전 정보를 이용하여 폴리펩티드 사의 합성을 특정짓고 이를 유도하는 과정을 나타낸다.
- <283> "벡터"는 적절한 조절 서열과 결합되었을 경우, 형질감염될 및(또는) 형질전환될 숙주세포에 특정한 성질을 부여하는 적절한 단백질 분자에 상응하는 폴리뉴클레오타이드 서열을 지닌, 유전자 조작시 세포의 형질감염 및(또는) 형질전환에 사용되는 핵산 화합물을 나타낸다. 플라스미드, 바이러스 및 박테리오파지가 적합한 벡터이다. 제한 효소 및 리가제를 사용하여, 근원이 서로 다른 DNA 분자들을 절단하고 결합함으로써 인공 벡터를 제조할 수 있다. 본원에 사용된 용어 "벡터"에는 재조합 DNA 클로닝 벡터 및 재조합 DNA 발현 벡터가 포함된다.

- <284> 본원에 사용된 "상보적인" 또는 "상보성"은 이중 가닥 핵산내에서 수소 결합을 통해 결합되어 있는 염기 (퓨린 및 피리미딘)들의 쌍을 나타낸다. 하기의 염기쌍이 상보적이다: 구아닌과 시토신; 아데닌과 티민; 및 아데닌과 우라실.
- <285> 본원에 사용된 "혼성화"는 핵산의 한 가닥이 염기 짝지음을 통해 상보적인 가닥과 결합하는 과정을 나타낸다. 동일하지는 않지만 매우 유사한 2개의 상보적인 핵산들의 혼성화에 사용되는 조건은, 두 가닥들간의 상보성 정도 및 가닥의 길이에 따라 달라진다. 이러한 기술 및 조건은 당업계의 숙련자들에게 잘 알려져 있다.
- <286> "단리된 아미노산 서열"은 자연적으로 발생하는 서열로부터 위치상 구별되는, 제조되거나 또는 합성된 것과 무관한 임의의 아미노산 서열을 나타낸다.
- <287> "단리된 DNA 화합물"은 위치상 게놈 DNA내 천연 위치로부터 구별되는, 제조되거나 또는 합성된 것과 무관한 임의의 DNA 서열을 나타낸다.
- <288> "단리된 핵산 화합물"은 위치상 천연 위치로부터 구별되는, 제조되거나 또는 합성된 것과 무관한 임의의 RNA 또는 DNA 서열을 나타낸다.
- <289> "프라이머"는 효소적 연장 또는 합성적 연장의 개시 기질로서 작용하는 핵산 단편을 나타낸다.
- <290> "프로모터"는 DNA의 RNA로의 전사를 유도하는 DNA 서열을 나타낸다.
- <291> "프로브"는 다른 핵산 화합물과 혼성화되는, 핵산 화합물 또는 이것의 단편을 나타낸다.
- <292> 혼성화 반응의 "엄격도"는 당업계 보통의 숙련자에 의해 쉽게 결정될 수 있고, 일반적으로 프로브 길이, 세척 온도 및 염 농도에 의존한 경험상의 계산 결과이다. 일반적으로, 긴 프로브는 올바른 어닐링을 위해 더 높은 온도가 필요한 반면, 짧은 프로브는 낮은 온도를 필요로 한다. 일반적으로 혼성화는, 융점 미만의 온도에서 상보적 가닥이 반응계 안에 존재할 경우, 변성된 DNA가 재-어닐링되는 능력에 의존한다. 프로브와 혼성화 가능한 서열 사이에 요구되는 상동성 정도가 높을수록, 이용될 수 있는 상대 온도는 더 높다. 따라서, 높은 상대 온도는 반응을 더 엄격하게 만드는 경향이 있는 반면에, 낮은 온도는 덜 그렇다. 혼성화 반응의 엄격도에 대한 상세한 내용 및 설명은 문헌 [Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, 1995]에서 참조할 수 있다.
- <293> 본원에 사용된 "엄격한 조건" 또는 "높은 엄격도 조건"은, (1) 세척시 예를 들면 50 °C에서 15 mM 염화나트륨/1.5 mM 시트르산나트륨/0.1 % 나트륨 도데실 설페이트와 같이, 낮은 이온 강도 및 높은 온도를 이용하거나; (2) 예를 들면 42 °C에서 0.1 % 소 혈청 알부민/0.1 % 피콜 (ficoll)/0.1 % 폴리비닐피롤리돈/50 mM 인산나트륨 완충액 (pH 6.5)이 포함된 50 % (부피/부피) 포름아미드와 750 mM 염화나트륨/75 mM 시트르산나트륨을 사용하는 것과 같이, 혼성화 과정 동안 포름아미드와 같은 변성화제를 사용하거나; 또는 (3) 42 °C에서 50 % 포름아미드, 5x SSC (750 mM 염화나트륨, 75 mM 시트르산나트륨), 50 mM 인산나트륨 (pH 6.8), 0.1 % 나트륨 피로포스페이트, 5X 덴하르츠 용액 (Denhardt's solution), 초음파로 파쇄된 연어 정자 DNA (50 µg/ml), 0.1 % SDS 및 10 % 텍스트란 설페이트를 사용하고, 42 °C에서 0.2x SSC (30 mM 염화나트륨/3 mM 시트르산나트륨)를 사용하여 세척하고, 55 °C에서 50 % 포름아미드를 사용하여 세척한 후, 55 °C에서 EDTA가 포함된 0.1x SSC를 사용한 높은-엄격도 세척을 이용하는 것일 수 있다.
- <294> "적당히 엄격한 조건"은 문헌 [Sambrook et al. Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, (1989)]에 기재된 바와 같은 것일 수 있고, 상기한 것보다 덜 엄격한 세척 용액 및 혼성화 조건 (예, 온도, 이온 강도 및 %SDS)을 이용하는 것이 포함된다. 적당히 엄격한 조건의 예로는, 20 % 포름아미드, 5x SSC (750 mM 염화나트륨, 75 mM 시트르산나트륨), 50 mM 인산나트륨 (pH 7.6), 5X 덴하르츠 용액, 10 % 텍스트란 설페이트 및 변성되고 전단된 연어 정자 DNA 20 mg/ml이 포함된 용액내에서 밤새 인큐베이션 (37 °C)시킨 후, 약 37 °C 내지 50 °C에서 1x SSC 중에서 필터를 세척하는 것이 있다. 숙련된 기술자는 프로브 길이 등과 같은 인자들을 고려하여 필요에 따라 온도, 이온 강도 등을 조정하는 방법을 알 것이다.
- <295> "PCR"은 열에 안정한 DNA 중합효소를 이용하는, 공지의 중합효소 연쇄 반응을 나타낸다.
- <296> "리더 서열"은 원하는 목적 폴리펩티드를 생성하기 위해, 효소적으로 또는 화학적으로 제거될 수 있는 아미노산 서열을 나타낸다.
- <297> "분비 신호 서열"은 폴리펩티드가 소포체와 같은 세포막 구획과 결합하고, 원형질 막을 통해 이 폴리펩티드가 분비되는 것을 일으키는 역할을 하는, 일반적으로 더 큰 폴리펩티드의 N-말단 영역에 존재하는 아미노산 서열을

나타낸다.

- <298> 본 발명의 이중성 융합 단백질을 코딩하는 DNA 제조
- <299> 야생형 알부민 및 면역글로블린 단백질은 다양한 공급원에서 얻을 수 있다. 예를 들면 상기 단백질들은, 관심 있는 mRNA를 검출 가능한 수준으로 발현하는 조직 또는 세포로부터 제조된 cDNA 라이브러리로부터 얻을 수 있다. 라이브러리는 관심있는 특정 단백질에 대한 공지의 DNA 또는 단백질 서열을 이용하여 디자인된 프로브로 스크리닝될 수 있다. 예를 들면, 면역글로블린 경쇄 또는 중쇄의 불변 영역 서열이 문헌 [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; 및 Morrison, et al. (1984) *Ann. Rev. Immunol.* 2: 239-256]에 기재되어 있다. 알부민의 단백질 및 DNA 서열이 개시된 일부 참고문헌으로는 문헌 [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; 및 Minghetti, et al. (1986) *J. Biol. Chem.* 261: 6747]이 포함된다.
- <300> 선택된 프로브를 사용하여 cDNA 또는 게놈 라이브러리를 스크리닝하는 것은 문헌 [Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY (1989)]에 기재된 것과 같은 표준 방법을 이용하여 수행될 수 있다. 알부민 또는 면역글로블린 단백질을 코딩하는 유전자를 분리하기 위한 다른 방법으로는 PCR 방법을 이용하는 것이 있다 (문헌 [Sambrook et al., 상기 문헌; Dieffenbach et al., *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY (1995)] 참조). PCR 프라이머는 공지의 서열에 기초하여 디자인될 수 있다.
- <301> 일반적으로, 특정한 종 (species)으로부터 클로닝된 야생형 전장 서열은, 융합 단백질의 일부인 GLP-1 화합물의 더 긴 혈장내 반감기를 부여하는 능력을 지닌 유사체, 단편 및 유도체를 생성하기 위한 주형으로서 이용될 수 있다. 본 발명의 이중성 융합 단백질의 Fc 및 알부민 부분은, 인간에 대한 상기 융합 단백질의 잠재적인 면역원성 위험을 줄이기 위해 천연 인간 서열로부터 유래되는 것이 바람직하다.
- <302> 특히, 본 발명에 포함된 융합 단백질의 면역글로블린 부분은 단지 면역글로블린의 Fc 단편만을 갖는 것이 바람직하다. 특정 이펙터 기능의 필요 여부와 융합 단백질의 구조적 특징에 따라서, Fc 단편은 CH2 도메인과 CH3 도메인, 또는 이들의 몇몇 다른 조합과 함께 힌지 영역을 포함할 수 있다. 이러한 Fc 단편은, 상기 단편의 원하는 말단에 해당하는 서열과 혼성화되도록 디자인된 프라이머를 사용한 PCR 기술을 이용하여 제조될 수 있다. 이와 유사하게, 알부민 단편이 필요할 경우에 알부민 내부 서열과 상보적인 PCR 프라이머를 디자인할 수 있다. 또한 PCR 프라이머는 발현 벡터내로의 클로닝을 용이하게 하는 제한 효소 부위가 생성되도록 디자인될 수도 있다.
- <303> 본 발명의 GLP-1 화합물을 코딩하는 DNA는 상기한 바와 같은 클로닝 방법을 비롯한 다양한 다른 방법뿐만 아니라 화학적으로 합성된 DNA에 의해서도 제조될 수 있다. 화학적 합성은 짧은 길이의 코딩된 펩티드를 생성하기에 유용할 수 있다. GLP-1의 아미노산 서열뿐만 아니라 프리프로글루카곤 유전자의 서열도 밝혀졌다 (문헌 [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; Ghiglion, M., et al. (1984) *Diabetologia* 27: 599-600] 참조). 따라서, 천연 GLP-1 화합물 및 이것의 단편을 PCR하기 위한 프라이머가 디자인될 수 있다.
- <304> 그 후, 융합 단백질을 코딩하는 유전자는, GLP-1 화합물을 코딩하는 DNA를 알부민 또는 Fc 단백질을 코딩하는 DNA와 프레임에 맞게 라이게이션시켜 제조될 수 있다. 또한 GLP-1 화합물을 코딩하는 유전자와, 알부민 또는 Fc 단백질을 코딩하는 유전자는 링커 펩티드를 코딩하는 DNA를 통해 프레임에 맞게 결합될 수도 있다.
- <305> 본 발명의 이중성 융합 단백질의 생체내 기능 및 안정성은, 잠재적으로 원치 않는 도메인간의 상호작용을 방지하기 위해 작은 펩티드 링커를 첨가함으로써 최적화될 수 있다. 비록 이러한 링커가 잠재적으로 임의의 길이일 수 있고, 아미노산들의 임의 조합으로 구성될 수 있지만, 링커의 길이는 원치 않는 도메인간의 상호작용 방지, 및(또는) 생물학적 활성 및(또는) 안정성의 최적화를 위해 필요한 길이보다 길지 않는 것이 바람직하다. 일반적으로, 상기 링커는 매우 큰 부피의 측쇄를 지닌 아미노산 또는 중요한 이차 구조를 유발할 수 있는 아미노산을 함유해서는 안된다. 상기 링커는 세린-글리신이 풍부하고, 아미노산 30개 미만의 길이를 갖는 것이 바람직하다. 상기 링커가 아미노산 20개 이하의 길이를 갖는 것이 더 바람직하다. 상기 링커가 아미노산 15개 이하의 길이를 갖는 것이 더욱 더 바람직하다. 바람직한 링커는 반복된 Gly-Gly-Gly-Gly-Ser 서열을 갖는다. 상기

서열이 2회 내지 6회 반복된 서열이 바람직하다. 상기 서열이 3회 내지 4회 반복된 서열이 더욱 더 바람직하다.

<306> 야생형 GLP-1, 알부민, 및 Fc 폴리펩티드 및 이것의 단편을 코딩하는 DNA는 라이게이션 이전에 또는 융합 단백질 전체를 코딩하는 cDNA내에서 돌연변이될 수 있다. 다양한 돌연변이 유발 기술이 당업계에 잘 알려져 있다. 예를 들면, 돌연변이 유발 PCR 방법은 해당하는 단백질내 특정 아미노산 서열을 변화시킬 목적으로 특정 염기 돌연변이를 생성하는 가닥 중복 연장 (strand overlap extension)을 이용한다. 이러한 PCR 돌연변이법은, 2종의 정방향 프라이머 (프라이머 A 및 C) 및 2종의 역방향 프라이머 (프라이머 B 및 D)의, 총 4종 프라이머를 이용한다. 돌연변이된 유전자는 서로 다른 두 단계에서 야생형 주형으로부터 증폭된다. 첫번째 반응에서는 A와 B 사이의 반응 및 별도의 C와 D 사이의 반응을 수행함으로써 유전자를 2등분하여 증폭시며, 여기에서 B 및 C 프라이머는 돌연변이된 유전자의 영역을 향하고 있다. 표적 영역을 향해 상기 프라이머들이 정렬될 때, 이 프라이머들은 치환될 표적 염기들에 대한 불일치 염기들을 갖고 있다. A와 B 사이의 반응 및 C와 D 사이의 반응이 완결되면, 반응 생성물을 분리하고 혼합하여 A와 D 사이의 반응을 위한 주형으로서 이용한다. 이 반응 후, 전장의 돌연변이가 생성물이 얻어진다.

<307> 융합 단백질 전체를 코딩하는 유전자가 생성되면, 이 유전자는 적절한 발현 벡터내로 클로닝될 수 있다. 본 발명의 GLP-1 융합 단백질을 제조하기 위해 이용될 수 있는 특정 전략을 실시예 1에 기재하였다.

<308> 본 발명의 이중성 융합 단백질을 제조할 형태로 발현시키기 위한 일반적 방법

<309> 이중성 융합 단백질 생산을 위해 본원에 기재된 발현 벡터 또는 클로닝 벡터를 사용하여 숙주세포를 형질감염 또는 형질전환시키고, 프로모터 작동 유도, 형질전환체 선별 또는 필요한 서열을 코딩하는 유전자의 증폭에 적합하도록 변형된 통상의 영양 배지내에서 상기 세포를 배양한다. 배양 조건 (예를 들면, 배지, 온도, pH 등)은 과도한 실험을 하지 않더라도 숙련된 기술자에 의해 선택될 수 있다. 일반적으로, 세포 배양의 생산성을 최대화하기 위한 원리, 방법 및 실제 기술은 문헌 [*Mammalian Cell Biotechnology: A Practical Approach*, M. Butler, ed. (IRL Press, 1991)] 및 문헌 [Sambrook, et al., 상기 참조]에서 찾아볼 수 있다. 형질감염 방법은 보통의 숙련된 기술자에게 공지되어 있으며, 예를 들면, CaPO₄법 및 전기천공법이 있다. 포유동물 세포 숙주 시스템에서의 형질전환에 대한 전반적인 측면이 미국특허 제4,399,216호에 개시되어 있다. 효모에서의 형질전환은 전형적으로, 문헌 [van Solingen et al., *J. Bact.* 130 (2): 946-7 (1977)] 및 문헌 [Hsiao et al., *Proc. Natl. Acad. Sci. USA* 76 (8): 3829-33 (1979)]의 방법에 따라 수행된다. 그러나, 핵 미세주입, 전기천공법, 완전한 세포와의 박테리아 원형질체 융합, 또는 다중양이온 (polycation; 예, 폴리브렌 (polybrene) 또는 폴리오미틴 (polyomithine)) 이용법과 같은, 세포내로 DNA를 도입하는 다른 방법도 이용될 수 있다. 포유동물 세포를 형질전환시키는 다양한 기술은 문헌 [Keown, et al., *Methods in Enzymology* 185: 527-37 (1990)] 및 문헌 [Mansour, et al., *Nature* 336 (6197): 348-52 (1988)]에서 참조할 수 있다.

<310> 본원에서 벡터내 핵산 (예를 들면, DNA)을 클로닝하고 발현시키는데 적합한 숙주세포에는 원핵생물, 효모 또는 고등 진핵생물 세포가 포함된다. 적합한 원핵생물에는 그람-음성 또는 그람-양성 생물체 (예를 들면, 이. 콜라이 (*E. coli*)와 같은 엔테로박테리아과 (Enterobacteriaceae)와 같은 진정박테리아 (eubacteria)가 포함되지만, 이것에만 제한되는 것은 아니다. 다양한 이. 콜라이 균주가 공개적으로 구입 가능하며, 예를 들면 이. 콜라이 K12 균주 MM294 (ATCC 31446); 이. 콜라이 X-1776 (ATCC 31537); 이. 콜라이 균주 W3110 (ATCC 27325); 및 K5772 (ATCC 53635)가 있다. 다른 적합한 원핵성 숙주세포에는 에스케리키아 (*Escherichia*; 예, 이. 콜라이), 엔테로박터 (*Enterobacter*), 에르위니아 (*Erwinia*), 클렙시엘라 (*Klebsiella*), 프로테우스 (*Proteus*), 살모넬라 (*Salmonella*; 예, 살모넬라 티피뮤리움 (*Salmonella typhimurium*)), 세르라티아 (*Serratia*; 예, 세르라티아 마르세스칸스 (*Serratia marcescans*)) 및 쉬게일라 (*Shigella*)뿐만 아니라, 바실러스 속 (*Bacilli*; 예, 바실러스 서브틸리스 (*B. subtilis*) 및 바실러스 리케니포르미스 (*B. licheniformis*; 예, 1989.4.12에 공개된 DD266,7 10에 개시되어 있는 바실러스 리케니포르미스 4 1 P)), 슈도모나스 (*Pseudomonas*; 예, 슈도모나스 에루지노사 (*P. aeruginosa*)) 및 스트렙토마이세스 (*Streptomyces*)와 같은 엔테로박테리아과가 포함된다. 상기 예들은 설명을 위한 것이며, 이들로만 제한되는 것은 아니다. 균주 W3110은 재조합 DNA 산물 발효에 이용되는 일반적인 숙주 균주이기 때문에 특히 바람직한 숙주 또는 모 숙주이다. 바람직하게도, 상기 숙주세포는 단백질 분해성 효소를 가장 적은 양으로 분비한다. 예를 들면, 균주 W3110은 숙주 내부에 존재하는 단백질-코딩 유전자내에서의 유전자 돌연변이를 일으키도록 변형될 수 있으며, 이러한 숙주의 예로는 완전 유전자형 (complete genotype) ronA를 갖는 이. 콜라이 W3110 균주 1A2; 완전 유전자형 ton4 ptr3을 갖는 이. 콜라이 W3110 균주 9E4; 완전 유전자형 tonA, ptr3 phoA E15 (argF-lac) I69 degP ompT/can'을 갖는 이. 콜라이

W3110 균주 27C7 (ATCC 55244); 균주 37D6의 가나마이신 (kanamycin)에 대한 내성이 없는 degP 결실 돌연변이인 이. 콜라이 W3110 균주 40B4; 및 미국특허 제4,946,783호 (1990.8.7 허여)에 개시되어 있는, 돌연변이된 세포질 주변 단백질 분해효소를 지닌 이. 콜라이 균주가 포함된다. 별법으로, 예를 들면 PCR 또는 다른 핵산 중합효소 반응과 같은 생체내 클로닝 방법이 적합하다.

<311>

원핵생물뿐만 아니라, 섬유상 진균류 또는 효모와 같은 진핵성 미생물도 융합 단백질 벡터를 클로닝하거나 또는 발현하기에 적합한 숙주이다. 사카로마이세스 세레비시예 (*Saccharomyces cerevisiae*)는 흔히 이용되는 하등 진핵성 숙주 미생물이다. 그 외에, 스킴조사카로마이세스 폼베 (*Schizosaccharomyces pombe*) (문헌 [Beach and Nurse, *Nature* 290: 140-3 (1981)] 및 EP 제139,383호 (1995.5.2 공개)); 예를 들면, 클루이베로마이세스 락티스 (*K. lactis*; MW98-8C, CBS 683, CBS 4574; 문헌 [de Louvencourt et al., *J. Bacteriol.* 154 (2): 737-42 (1983)]), 클루이베로마이세스 프라질리스 (*K. fragilis*; ATCC 12424), 클루이베로마이세스 불가리쿠스 (*K. bulgaricus*; ATCC 16045), 클루이베로마이세스 위케라미 (*K. wickerhamii*; ATCC 24,178), 클루이베로마이세스 왈티 (*K. waltii*; ATCC 56500), 클루이베로마이세스 드로소필라룸 (*K. drosophilorum*; ATCC 36,906; 문헌 [Van den Berg et al., *Bio/Technology* 8(2): 135-9 (1990)]), 클루이베로마이세스 써모톨러란스 (*K. thermotolerans*) 및 클루이베로마이세스 마르시아누스 (*K. marxianus*)와 같은 클루이베로마이세스 (*Kluyveromyces*) 숙주 (미국특허 제4,943,529호 및 문헌 [Fleer, et al., *BiolTechnology* 9(10): 968-75 (1991)]); 야르로위아 (*Yarrowia*; EP 402,226); 피키아 파스토리스 (*Pichia pastoris*; EP 제183,070호; 문헌 [Sreekrishna et al., *J. Basic Microbiol.* 28(4): 265-78 (1988)]); 칸디다 (*Candida*); 트리코더마 레에시아 (*Trichoderma reesia*; EP 제244,234호); 뉴로스포라 크라사사 (*Neurospora crassa*; 문헌 [Case, et al., *Proc. Natl. Acad. Sci. USA* 76(10): 5259-63 (1979)]); 스완니오마이세스 옥시덴틀리스 (*Schwanniomyces occidentalis*; EP 제394,538호 (1990.10.31 공개))와 같은 스완니오마이세스; 및 예를 들면, 뉴로스포라, 페니실리움 (*Penicillium*), 톨리포클라디움 (*Tolypocladium*; WO 91/00357호 (1991.1.10 공개)) 및 아스퍼질러스 (*Aspergillus*) 숙주 (예, 아스퍼질러스 니둘란스 (*A. nidulans*; 문헌 [Ballance et al., *Biochem. Biophys. Res. Comm.* 112(1): 284-9 (1983)], [Tilburn, et al., *Gene* 26(2-3): 205-21 (1983)] 및 [Yelton, et al., *Proc. Natl. Acad. Sci. USA* 81(5): 1470-4 (1984)] 및 아스퍼질러스 나이저 (*A. niger*; 문헌 [Kelly and Hynes, *EMBO J.* 4(2): 475-9 (1985)]))와 같은 섬유상 진균류가 포함된다. 메틸영양요구성 효모 (methylotrophic yeast)는 한세놀라 (*Hansenula*), 칸디다 (*Candida*), 클로엑케라 (*Kloeckera*), 피키아 (*Pichia*), 사카로마이세스 (*Saccharomyces*), 토룰롭시스 (*Torulopsis*) 및 로도토루이아 (*Rhodotoruia*)로 구성된 속 (genus)들로부터 선택된다. 이러한 효모류의 예로서 특정 종 (species)의 목록은 문헌 [C. Antony, *The Biochemistry of Methylotrophs* 269 (1982)]에서 찾아볼 수 있다.

<312>

본 발명의 융합 단백질을 발현하기에 적합한 숙주세포는 다세포 생물체로부터 유래되었다. 무척추동물 세포의 예로는, 곤충세포 (예, 초파리 (*Drosophila*) S2 및 스포돗테라 종 (*Spodoptera* Sp.; 예, 스포돗테라 high5))뿐만 아니라 식물세포도 포함된다. 유용한 포유동물 숙주세포주의 예로는 중국 햄스터 난소 (CHO) 세포 및 COS 세포가 포함된다. 더 구체적인 예로는, SV-40에 의해 형질전환된 원숭이 신장 CV1 세포주 (COS-7, ATCC CRL 1651); 인간 배아 신장 세포주 (293 세포, 또는 현탁 배양에서 성장하도록 서브클로닝된 293 세포, 문헌 [Graham, et al., *J. Gen. Tirol.*, 36(1): 59-74 (1977)]); 중국 햄스터 난소 세포/-DHFR (CHO, 문헌 [Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77(7): 4216-20 (1980)]); 마우스 세르톨리 (sertoli) 세포 (TM4, 문헌 [Mather, *Biol. Reprod.* 23(1): 243-52 (1980)]); 인간 폐 세포 (W138, ATCC CCL 75); 인간 간 세포 (HepG2, HB8065); 및 마우스 유선 종양 (MMT 060562, ATCC CCL 51)이 포함된다. 적절한 숙주세포의 선택은 당업계의 기술범위내에 있는 것으로 생각된다.

<313>

본 발명의 융합 단백질은 제조법에 의해 직접 생산될 수도, 또는 신호 서열, 또는 성숙 융합 단백질의 N-말단에 특정한 절단 위치를 생성하는 다른 부가 서열을 지닌 단백질로서 생산될 수도 있다. 일반적으로, 상기 신호 서열은 벡터내 성분이거나, 또는 벡터내에 삽입된 융합 단백질-코딩 DNA의 일부일 수 있다. 상기 신호 서열은, 예를 들면 알칼라인 포스파타제, 페니실리나제 (penicillinase), lpp 또는 열-안정 엔테로톡신 (enterotoxin) II 리더 (leader)로 구성된 군에서 선택된 원핵성 신호 서열일 수 있다. 효모에서의 분비를 위해 상기 신호 서열은, 예를 들면 효모 인버타제 리더, 알파 인자 리더 (사카로마이세스 α-인자 리더, 및 미국특허 제5,010,182호에 기재된 클루이베로마이세스 cc-인자 리더 포함), 산성 포스파타제 리더, 칸디다 알비칸스 (*C. albicans*) 글루코아미라제 리더 (EP 제362,179호) 또는 WO 90/13646호에 기재된 신호 서열일 수 있다. 포유동물 세포 발현에 있어서, 포유동물 신호 서열은 단백질의 분비를 유도하기 위해 사용될 수 있으며, 그 예로는 동종 또는 관련 종 (species)의 분비 폴리펩티드로부터의 신호 서열, 및 바이러스의 분비성 리더가 있다.

- <314> 발현 벡터와 클로닝 벡터는, 이들 벡터가 1종 이상의 선택된 숙주세포내에서 복제될 수 있도록 하는 핵산 서열을 지니고 있다. 다양한 박테리아, 효모 및 바이러스에 대한 상기 서열은 잘 알려져 있다. 플라스미드 pBR322에서의 복제 개시점 (origin)이 대부분의 그람-음성 박테리아에 적합하고, 2 μ 플라스미드 복제 개시점이 효모에 적합하며, 다양한 바이러스성 복제 개시점 (SV40, 폴리오마 (polyoma), 아데노바이러스 (adenovirus), VSV 또는 BPV)은 포유동물 세포에서의 클로닝 벡터에 유용하다.
- <315> 전형적으로 발현 및 클로닝 벡터는 선별 유전자를 지니고 있을 것이며, 이 유전자는 선별 가능 마커로도 명명된다. 전형적인 선별 유전자는, (a) 항생제 또는 다른 독소 (예를 들면, 엠포실린, 네오마이신, 메토크세이트 (methotrexate) 또는 테트라사이클린)에 대한 내성을 부여하는 단백질을 코딩하거나; (b) 자가영양공급성 결핍 (autotrophic deficiency)을 보충해주는 단백질을 코딩하거나; 또는 (c) 예를 들면, 바실러스 속 (*bacilli*)에서 D-알라닌 라세미화 효소를 코딩하는 유전자와 같이, 복합영양 배지에서 얻을 수 없는 주요 영양물을 공급해주는 단백질을 코딩한다.
- <316> 포유동물 세포에서의 적합한 선별 가능 마커의 예로는, 융합 단백질-코딩 핵산을 획득한 세포를 동정할 수 있도록 하는, DHFR 또는 티미딘 키나제와 같은 것들이 있다. 야생형 DHFR을 이용할 경우에 적절한 숙주세포로는, 문헌 [Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77(7): 4216-20 (1980)]에 기재된 방법으로 제조 및 증식된, DHFR 활성이 결핍된 CHO 세포주가 있다. 효모에서 이용할 수 있는 적합한 선별 유전자로는 효모 플라스미드 YRp7내에 존재하는 *trp1* 유전자가 있다 (문헌 [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* 유전자는, 트립토판을 이용하여 성장할 수 있는 능력이 결여된 효모의 돌연변이 균주 (예를 들면, ATCC 44076, 또는 PEPC1 (문헌 [Jones, *Genetics* 85: 23-33 (1977)])에 대한 선별 마커를 제공한다.
- <317> 일반적으로 발현 및 클로닝 벡터는 mRNA 합성을 유도하기 위한, 융합 단백질-코딩 핵산 서열과 작동 가능하게 연결된 프로모터를 갖고 있다. 다양한 잠재적 숙주세포에 의해 인식되는 프로모터들이 잘 알려져 있다. 원핵성 숙주에서 이용하기에 적합한 프로모터에는 β -락타마제 (lactamase)와 락토스 프로모터 시스템 (문헌 [Chang, et al., *Nature* 275(5681): 617-24 (1978); Goeddel, et al., *Nature* 281(5732): 544-8 (1979)]), 알카라인 포스파타제, 트립토판 (*up*) 프로모터 시스템 (문헌 [Goeddel, *Nucleic Acid Res.* 8(18): 4057-74 (1980)]; EP 제36,776호 (1981.9.30 공개)) 및 하이브리드 프로모터 (예, *tat* 프로모터; 문헌 [deBoer, et al., *Proc. Natl. Acad. Sci. USA* 80(1): 21-5 (1983)])가 포함된다. 또한 박테리아 시스템에서 이용될 수 있는 프로모터들도 융합 단백질을 코딩하는 DNA와 작동 가능하게 연결된 샤인-달가노 (Shine-Dalgarno; S.D.) 서열을 갖고 있을 것이다.
- <318> 효모 숙주에서 이용될 수 있는 적합한 프로모터의 예로는, 3-포스포글리세레이트 키나제에 대한 프로모터 (문헌 [Hitzeman, et al., *J. Biol. Chem.* 255(24): 12073-80 (1980)]), 또는 에놀라제, 글리세르알데히드-3-포스페이트 데히드로게나제, 헥소키나제, 피루베이트 데카르복실라제, 포스포프럭토키나제, 글루코스-6-포스페이트 이소머라제, 3-포스포글리세레이트 뮤타제, 피루베이트 키나제, 트리오스포스페이트 이소머라제, 포스포글루코스 이소머라제 및 글루코키나제와 같은, 다른 당분해성 효소에 대한 프로모터 (문헌 [Hess et al., *J. Adv. Enzyme Reg.* 7:149 (1968); Holland, *Biochemistry* 17(23): 49007 (1978)])가 포함된다.
- <319> 성장 조건에 의해 조절되는 전사에서 추가 이점을 제공하는 유도성 프로모터로서의 다른 효모 프로모터로는, 알콜 데히드로게나제 2, 이소시토크롬 C, 산 포스파타제, 질소 대사와 연관된 분해성 효소, 메탈로티오네인, 글리세르알데히드-3-포스페이트 데히드로게나제, 및 말토스 및 갈락토스 이용을 담당하는 효소에 대한 프로모터 영역이 있다. 효모 발현에서 이용될 수 있는 적합한 벡터 및 프로모터들은 EP 제73,657호에 더 기재되어 있다. 포유동물 숙주세포내 벡터로부터 융합 단백질-코딩 mRNA의 전사는 폴리오마 바이러스, 계두 (fowlpox) 바이러스, 아데노바이러스 (예, 아데노바이러스 2), 소 파필로마 바이러스, 조류 육종 바이러스, 사이토메갈로 바이러스, 레트로바이러스, B형 간염 바이러스 및 시미안 (simian) 바이러스 40 (SV40)와 같은 바이러스의 게놈으로부터 얻은 프로모터, 액틴 프로모터 또는 면역글로블린 프로모터와 같은 이중성 포유동물 프로모터로부터 얻은 프로모터, 및 열-충격 프로모터로부터 얻은 프로모터에 의해 조절될 수 있으며, 단 상기 프로모터들은 숙주세포 시스템에 적합해야 한다.
- <320> 고등 진핵생물에 의한 융합 단백질-코딩 폴리뉴클레오타이드의 전사는 벡터내에 인핸서 (enhancer) 서열을 삽입함으로써 증가될 수 있다. 인핸서는 일반적으로 약 10 내지 300 bp의 DNA내 시스-작용 요소이며, 프로모터에 작용하여 전사를 증가시킨다. 현재 포유동물 유전자에서의 많은 인핸서 서열이 알려져 있다 (글로빈, 엘라스타제, 알부민, α -케토단백질 및 인슐린). 그러나 전형적으로, 진핵성 세포 바이러스에서 유래한 인핸서

가 이용될 것이다. 그 예로는, 복제 개시점의 뒤쪽에 있는 SV40 인핸서 (bp 100-270), 사이토메갈로바이러스 초기 프로모터 인핸서, 복제 개시점의 뒤쪽에 있는 폴리오마 인핸서 및 아데노바이러스 인핸서가 포함된다. 상기 인핸서는 벡터내의 융합 단백질 코딩 서열의 5' 또는 3' 위치에 스플라이싱 (splicing)될 수도 있지만, 바람직하게는 프로모터의 5' 위치에 존재한다.

<321> 또한 진핵성 숙주세포 (효모, 진균류, 곤충, 식물, 동물, 인간 또는 다른 다세포 생물체로부터의 핵 생성 세포)에 사용된 발현 벡터는 전사 종결에 필요한 서열 및 mRNA의 안정화에 필요한 서열도 갖고 있을 것이다. 이러한 서열은 일반적으로 진핵성 또는 바이러스성의 DNA 또는 cDNA의 5' 비번역 영역 또는 때때로 3' 비번역 영역에서 얻을 수 있다. 상기 영역은 융합 단백질을 코딩하는 mRNA의 비번역 부분내에 폴리아데닐화된 단편으로서 전사되는 뉴클레오타이드 단편을 갖고 있다.

<322> 다양한 형태의 융합 단백질은 배양 배지 또는 숙주세포 용해물 (lysate)로부터 회수될 수 있다. 만약 상기 단백질이 막-결합 단백질일 경우, 적합한 세제 용액 (예, 트리톤 (Triton)-X 100)을 이용하거나 또는 효소적 절단에 의해 막으로부터 방출될 수 있다. 융합 단백질의 발현에 이용된 세포는, 예를 들면 동결-해동 순환, 초음파 파쇄, 기계적 파쇄 또는 세포 용해제와 같은, 다양한 물리적 또는 화학적 수단에 의해 파쇄될 수 있다.

<323> 본 발명의 이중성 융합 단백질의 정제

<324> 적절한 숙주세포내에서 본 발명의 이중성 융합 단백질이 발현되면, 유사체는 분리 및 정제될 수 있다. 하기 절차는 적합한 정제 절차의 예시이다: 카르복시메틸 셀룰로스 상에서의 분획화; 세파덱스 (Sephadex) G-75와 같은 겔 여과; DEAE 또는 모노 (Mono)-Q와 같은 음이온 교환 수지; CM 또는 모노 (Mono)-S와 같은 양이온 교환 수지; IgG와 같은 불순물을 제거하기 위한 단백질 A 세파로스 (Protein A sepharose); 에피토프가 결합된 형태의 폴리펩티드를 결합시키기 위한 금속 킬레이팅 칼럼; 역상 HPLC; 크로마토포커싱 (chromatofocusing); 실리카겔; 에탄올 침전; 및 황산암모늄 침전.

<325> 다양한 단백질 정제 방법들이 이용될 수 있고, 이러한 방법들은 당업계에 공지되어 있으며, 예를 들면 문헌 [Deutscher, *Methods in Enzymology* 182: 83-9 (1990)] 및 [Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag, NY (1982)]에 기재되어 있다. 선택된 정제 단계(들)은 이용된 생산 방법의 특성 및 생산된 특정 융합 단백질에 따라 달라질 것이다. 예를 들면, Fc 단편을 포함하는 융합 단백질은 단백질 A (Protein A) 또는 단백질 G (Protein G) 친화 매트릭스를 이용하여 효과적으로 정제될 수 있다. 낮은 pH 또는 높은 pH의 완충액을 사용하여 상기 친화 매트릭스로부터 융합 단백질을 용출시킬 수 있다. 온화한 용출 조건은 융합 단백질의 비가역적 변성을 방지하는 데 도움이 될 것이다. 또한 이미다졸-함유 완충액을 사용할 수도 있다. 실시예 3에는 본 발명의 융합 단백질에 대한 몇몇 성공적인 정제 프로토콜들을 기재하였다.

<326> 본 발명의 이중성 융합 단백질의 특성 분석

<327> 본 발명의 융합 단백질의 특성 분석을 위한 많은 방법들이 존재한다. 이러한 방법들 중 몇몇 방법에는, 단백질 염색법과 병용하는 SDS-PAGE 또는 항-IgG 또는 항-HSA 항체를 이용한 면역블랏팅 (immunoblotting)이 포함된다. 다른 방법에는 매트릭스 보조 레이저 탈착/이온화 질량 분석법 (Matrix Assisted Laser Desorption/Ionization Mass Spectrometry, MALDI-MS), 액체 크로마토그래피/질량 분석법, 등전 포커싱 (isoelectric focusing), 분석용 음이온 교환, 크로마토포커싱 및 원형 이색성 (circular dichroism)이 포함되지만, 이들은 단지 몇가지 예에 불과하다. 면역블랏팅과 병용하는 SDS-PAGE뿐만 아니라 질량 분석법을 이용하여, 몇몇 대표적인 이중성 융합 단백질들의 특성을 분석하였다 (실시예 4 및 5, 도 3 및 4 참조).

<328> 예를 들면, 표 3 (실시예 5 참조)은 몇몇 대표적인 융합 단백질들에 대한 계산된 분자량뿐만 아니라 질량 분석법에 의해 결정된 질량을 나타내고 있다. 또한, 도 3 및 도 4는 SDS-PAGE에 의해 결정된, 몇몇 대표적인 융합 단백질의 분자량을 나타내고 있다. 시험된 모든 이중성 융합 단백질은 일시적으로 발현되고 분리되었다. 또한, Igκ 신호 서열을 절단하여 올바른 N-말단을 지닌 단백질을 얻었다.

<329> 또한, 표 3은 질량 분석법에 의해 결정된 질량이 예측 질량보다 더 큰 몇몇 경우를 나타내고 있다. 이것은 Fc 부분의 글리코실화 및 C-말단 연장에 의한 것이다. 융합 단백질의 효소적 분해 이후, 역상 HPLC 및 질량 분석법을 이용하여 당 부분을 지닌 펩티드 분획을 동정할 수 있다. 그 후, 상기 분획에 대한 N-말단 아미노산 서열 분석을 수행하여 잠재적 글리코실화 부위를 동정할 수 있다. 예를 들면, 엑센딘-4-Fc (서열 29)의 특성 분석 결과는 위치 39의 세린과 위치 50의 트레오닌이 O-연결된 글리코실화가 되어 있고, 위치 122의 아스파라긴은 N-연결된 글리코실화가 되어 있음을 나타내고 있다.

<330> 활성에 대해서도 몇몇 대표적인 GLP-1 융합 단백질을 시험하였다. 시험관내 및 생체내에서의 GLP-1 활성을 검

출하기 위한 많은 방법들이 존재한다 (실시예 6, 7, 8 및 9 참조). 표 4 (실시예 6)는 몇몇 GLP-1 융합체와 관련된 GLP-1 수용체의 활성을 나타내고 있다. 측정값은 Val⁸-GLP-1(7-37)OH와 관련된 활성에 대한 값이다. 시험된 모든 융합 단백질은 GLP-1 수용체 활성을 갖고 있었다. 낮은 수준의 시험관내 활성은 반드시 낮은 생체내 효과를 암시하는 것은 아니다. 상기 융합 단백질들의 실질적인 반감기 증가로 인해, 약한 시험관내 활성은 일반적으로 낮은 생체내 활성에 대한 예측 수단이 되지 못한다. 도 7 및 실시예 7은 본 발명의 융합 단백질과 관련하여 연장된 반감기를 나타내고 있다. 예를 들면, Val⁸-GLP-1-Fc은 원숭이에서 약 45시간의 반감기를 갖고, Val⁸-GLP-1-HSA은 원숭이에서 약 87시간의 반감기를 갖고, Gly⁸-Glu²²-GLP-1-CEX-링커-IgG1은 개에서 IV 투여 후 약 55시간의 반감기를 가지며, Gly⁸-Glu²²-GLP-1-CEX-링커-IgG1은 개에서 SC 투여 후 약 38시간의 반감기를 갖는다.

<331> 본 발명의 조성물

<332> 물리적 안정성은 치료용 단백질 제제에서 또한 중요한 특성이다. GLP-1 화합물은 가공시 발생하는 구조적 변화로 인해, 제조 및 제제화하기에 특히 어려웠다. 예를 들면, 몇몇 GLP-1 화합물은 일반적으로 응집되는 경향을 갖고 있다. 또한 몇몇 GLP-1 화합물은, 가용성의 활성 α-나선 형태에서, 불용성의 잠재적 불활성 β-쉬트 형태로 전환되는 것으로 나타났다. GLP-1 화합물과, IgG의 Fc 영역 또는 알부민과 같은 큰 단백질과의 융합은, 상기 GLP-1 화합물의 반감기를 증가시킬뿐만 아니라 GLP-1 화합물의 물리적 및 구조적 안정성에도 기여한다. 예를 들면, PBS 중의 Val⁸-GLP-1-링커-HSA는 37 °C에서 약 30일까지 안정하다.

<333> 본 발명의 이중성 융합 단백질은 1종 이상의 부형제와 함께 제제화될 수 있다. 본 발명의 활성 융합 단백질을 제약상 허용 가능한 완충액과 혼합할 수도 있으며, pH는 허용 가능한 안정성을 제공하고 투여 (예를 들면, 비경구 투여)에 적합하도록 조정될 수 있다.

<334> 임의로는, 1종 이상의 제약상 허용 가능한 항미생물제가 첨가될 수도 있다. 메타-크레졸 및 페놀은 바람직한 제약상 허용 가능한 미생물 제제이다. 이온 강도 또는 장성 (tonicity)을 조정하기 위해, 1종 이상의 제약상 허용 가능한 염을 첨가할 수 있다. 상기 제제의 등장성을 더 조정하기 위해 1개 이상의 부형제가 첨가될 수 있다. 글리세린은 등장성-조정 부형제의 한 예이다. 제약상 허용 가능하다는 것은 인간 또는 다른 동물에 투여하기에 적합하다는 의미이며, 따라서 독성 요소 또는 바람직하지 않은 불순물을 함유하고 있지 않고, 본원의 활성 화합물의 활성을 방해하지 않는다.

<335> 본 발명의 이중성 융합 단백질의 제약상 허용 가능한 염 형태가 본 발명에 사용될 수 있다. 산 부가 염을 형성하기 위해 흔히 사용되는 산으로는 염산, 브롬화수소산, 요오드화수소산, 황산, 인산 등과 같은 무기산, 및 p-톨루엔설폰산, 메탄설폰산, 옥살산, p-브로모페닐-설폰산, 탄산, 숙신산, 시트르산, 벤조산, 아세트산 등과 같은 유기산이 있다. 바람직한 산 부가 염은 염산 및 브롬화수소산과 같은 무기산을 사용하여 형성된 것이다.

<336> 염기 부가 염으로는 암모늄, 알칼리 또는 알카라인 토금속의 수산화물; 탄산염; 중탄산염 등과 같은 무기 염기에서 유래한 것들이 포함된다. 따라서 본 발명의 상기 염을 제조하는 데 유용한 염기로는 수산화나트륨, 수산화칼륨, 수산화암모늄, 탄산칼륨 등이 포함된다.

<337> 조성물의 투여

<338> 투여는 보통의 숙련된 의사에 의해 효과적이라고 알려진 임의의 경로를 통해 수행될 수 있다. 말초 비경구 투여는 이러한 방법 중 하나이다. 비경구 투여는 일반적으로 의학 문헌내에서, 살균된 주사기 또는 다른 기계적 도구 (예를 들면, 주입 펌프)에 의해 투약물을 체내로 주사하는 것으로서 이해된다. 말초 비경구 경로에는 정맥내, 근육내, 피하 및 복막내 투여 경로가 포함된다.

<339> 본 발명의 이중성 융합 단백질은 비경구 경로 이외의 경로인, 직장, 직장, 비강 또는 하기도 경로에 의해 투여될 수도 있다. 이러한 비경구 경로 이외의 경로 중, 하기도 경로 및 경구 경로가 바람직하다.

<340> 본 발명의 융합 단백질은 매우 다양한 질병 및 장애를 치료하는 데 이용될 수 있다. 본 발명의 융합 단백질은 주로, "GLP-1 수용체"로서 언급되는 수용체에 작용함으로써 생물학적 효과를 발휘한다. 따라서 GLP-1 수용체 자극 또는 GLP-1 화합물의 투여에 의해 바람직한 반응을 나타내는 질병 및(또는) 장애를 지닌 대상체는 본 발명의 GLP-1 융합 단백질을 사용하여 치료될 수 있다. 이러한 대상체를 "GLP-1 화합물을 사용한 치료가 필요한" 또는 "GLP-1 수용체 자극이 필요한" 대상체라고 한다. 이러한 대상체에는 인슐린 비의존성 당뇨병, 인슐린 의존성 당뇨병, 뇌졸중 (WO 00/16797호 참조), 심근경색 (WO 98/08531호 참조), 비만 (WO 98/19698호 참조), 수

술 후 이화작용 변화 (미국특허 제6,006,753호 참조), 기능적 소화불량 및 과민성 대장증후군 (WO 99/64060호 참조) 대상체가 포함된다. 또한 GLP-1 화합물을 사용한 예방 치료가 필요한 대상체, 예를 들면 인슐린 비의존성 당뇨병으로 발전할 수 있는 위험을 지닌 대상체도 포함된다 (WO 00/07617호 참조). 글루코스 내성이 손상된 또는 글루코스 차단이 손상된 대상체, 대상체의 신장 및 신체에 대한 정상 체중보다 약 25 %를 초과하는 체중을 갖는 대상체, 췌장이 일부 절제된 대상체, 1명 이상의 조상이 인슐린 비의존성 당뇨병을 지닌 대상체, 임신성 당뇨병에 걸렸던 대상체, 및 급성 또는 만성 췌장염에 걸렸던 대상체는 인슐린 비의존성 당뇨병으로 발전할 수 있는 위험이 있다.

<341> GLP-1 화합물의 "유효량"은 GLP-1 수용체 자극이 필요한 환자에게 투여했을 때, 허용 불가능한 부작용을 유발하지 않으면서 원하는 치료 및(또는) 예방 효과를 나타내는 양이다. "원하는 치료 효과"에는 1개 이상의 하기 효과가 포함된다: 1) 질병 또는 장애와 관련된 징후(들)의 완화 효과; 2) 질병 또는 장애와 관련된 징후 개시의 지연 효과; 3) 치료를 하지 않은 경우와 비교시 수명 증가 효과; 및 4) 치료를 하지 않은 경우보다 삶의 질 향상 효과. 예를 들면, 당뇨병 치료를 위한 GLP-1 화합물의 "유효량"은 치료를 하지 않은 경우보다 혈중 글루코스 농도를 더 잘 조절할 수 있고, 따라서 망막증, 신경병증 또는 신장 질병과 같은 당뇨병 합병증의 개시를 지연시킬 수 있는 양이다. 당뇨병 방지를 위한 GLP-1 화합물의 "유효량"은 치료를 하지 않은 경우와 비교시, 설포닐 우레아, 티아졸리딘디온, 인슐린 및(또는) 비스구아니딘과 같은 항-저혈당증 약제를 사용한 치료가 필요한, 혈중 글루코스 수준의 증가 개시를 지연시킬 수 있는 양이다.

<342> 환자의 혈중 글루코스를 정상화하기에 효과적인 융합 단백질의 투여량은, 예를 들면 환자의 성별, 체중 및 나이, 혈중 글루코스 조절 불능의 심각성, 투여경로 및 생체 이용도, 융합 단백질의 약물 동력학적 측면, 효능 및 제형 등을 비롯한 다양한 인자들에 의존한다.

<343> 본 발명은 알부민 단백질, 알부민 단편, 알부민 유사체, Fc 단백질, Fc 단편 또는 Fc 유사체와의 융합으로 인해 향상된 생화학적 및 생물물리학적 성질을 갖는 GLP-1 화합물을 포함한다. 이러한 이중성 단백질은 숙주세포 내에서 성공적으로 발현될 수 있고, GLP-1 수용체 활성화와 관련된 신호 전달 활성을 보유하며, 연장된 반감기를 갖는다.

효 과

<344> 본 발명에 의해, GLP-1 수용체 활성화와 관련된 신호 전달 활성을 보유하며, 연장된 반감기를 갖는 이중성 GLP-1 융합 단백질을 얻을 수 있다.

발명의 실시를 위한 구체적인 내용

<345> 하기 실시예는 본 발명을 더 설명하기 위해 제시되었다. 본 발명의 범위는 단지 하기 실시예로만 구성된 것으로서 해석되어서는 안된다. 당업계의 숙련자들은 기재된 특정 시약, 장비 및 절차가 단순히 설명을 위한 것이고, 어떠한 경우라도 본 발명을 제한하고자 하는 것이 아님을 알 것이다.

<346> 실시예 1: 이중성 융합 단백질을 코딩하는 DNA 제조

<347> 실시예 1a: Val⁸-GLP-1(7-37)-Fc를 코딩하는 DNA 제조

<348> 인간 IgG1의 Fc 부분을 cDNA 라이브러리로부터 분리하였는데, 이는 힌지 영역 전체와 CH2 및 CH3 도메인을 포함하는 것이었다. 염기쌍 696개를 지닌 상기 인간 IgG1 Fc 부분의 단편을 포유동물 발현 벡터 pJB02내의 *NheI* 및 *Eco47III* 위치에 서브클로닝하여 pJB02/Fc (도 5 참조)를 제조하였다. Val⁸-GLP-1(7-37)에 융합된, Igκ 분비 신호 서열을 코딩하는 DNA는, 4종의 중복 및 상보적 올리고뉴클레오타이드들의 시험관내 혼성화에 의해 제조되었다:

5'- CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
CCAGGTTCCACTGGTGACCACTG - 3' (서열 12)

5'- GAGGGCACCTTCACCTCCGACGTGTCTCTCTATCTGGAGGGCCAGGCCGCAAGGA
GTTTCATCGCCTGGCTGGTGAAGGGAAGAGGC - 3' (서열 13)

5'- TGAAGGTGCCCTCCACGTGGTCACCACTGGAACCTGGAACCCAGAGCAGCAGTA
CCCATAGCAGGAGTGTGTCTGTCTCCATGGTGG - 3' (서열 14)

5'- GCCTCTTCCCTTCACCAAGCCAGGCGATGAACTCCTTGGCGGCCTGGCCCTCCAGA
TAGGAGGACACGTCGGAGG - 3' (서열 15)

<349>

<350> 각각의 올리고뉴클레오타이드 등가량 (올리고 각각의 최종 농도: 1 pm/ μ l)을 사용하여 혼성화 반응을 수행하였다. 올리고뉴클레오타이드 혼합물을 라이게이션 완충액 (50 mM 트리스-HCl, pH 7.5, 10 mM MgCl₂, 10 mM DTT, 1mM ATP, 25 μ g/ml 소 혈청 알부민)내에서 5분 동안 100 °C로 가열한 후, 2시간 이상에 걸쳐 30 °C까지 냉각시켰다.

<351> 결과의 혼성화 산물을 *NheI* 및 *Eco47III*로 절단시킨 pJB02/Fc 벡터 골격에, 실온에서 2시간 동안 또는 16 °C에서 밤새 라이게이션시켰다. XL-1 블루 컴피턴트 (competent) 세포 (스트라타진 (Stratagene)社)를 상기 라이게이션 산물로 형질전환시켰다. 상기 클론을 *NcoI*으로 절단하여, 펩티드 코딩 삽입체 (코작 (Kozak) 서열 및 신호 펩티드의 첫번째 Met을 코딩하는 삽입체)의 존재에 대해 상기 재조합 플라스미드를 스크리닝하고, 서열을 분석하였다. 형질감염 분석에 사용된 결과의 발현 플라스미드를 pJB02-V8-GLP-1-Fc (도 5)로 명명하였다.

<352> 실시예 1b: Val⁸-GLP-1(7-37)-HSA를 코딩하는 DNA 제조

<353> 플라스미드 HSA/pcDNA3.1GS를 인비트로젠 (Invitrogen) 社 (카다로그 번호: H-M12523M-pcDNA.3.1/GS)에서 구입하여, 인간 혈청 알부민을 코딩하는 cDNA를 단리하기 위한 주형으로서 사용하였다. PCR을 이용하여, 5' 말단에서 DNA 코딩 리더 서열뿐만 아니라 아미노산 6개의 프로-펩티드가 제거된 HSA cDNA를 제조하였다. 또한, 정지 코돈 (stop codon)을 상기 HSA 코딩 서열의 3' 말단에 직접 부가하였다. 마지막으로, 클로닝을 용이하게 하기 위해 5' 및 3' 말단에 제한 효소 부위를 부가하였다. 인비트로젠 社에서 구입한 본래의 벡터내에 존재하는 HSA DNA 서열은, 천연 인간 서열과 비교하여 유전자 3' 영역 (위치 667)에 1개의 염기가 치환되어 있다. 상기 치환으로 Asp 대신 Asn에 대한 코돈이 생성된다. 따라서, 상기한 가닥 중복 PCR 돌연변이 방법을 이용하여, 상기 코돈이 그 위치에서 Asp를 코딩하도록 치환시켰다. 결과의 HSA 코딩 DNA를 pJB02내의 *NheI* 및 *HindIII* 위치에 클로닝하여 pJB02-HSA (도 6)를 제조하였다.

<354> 실시예 1a에 기재한 방법으로, Val⁸-GLP-1(7-37) 서열에 융합된 Ig κ 리더 서열을 제작하였다. 상기 DNA를 pJB02-HSA의 *NheI* 및 *FspI* 위치에 라이게이션시켜 pJB02-Val⁸-GLP-1-HSA를 제조하였다.

<355> 실시예 1c: Val⁸-GLP-1(7-37)-링커-HSA를 코딩하는 DNA 제조

<356> 실시예 1b에 기재한 방법으로 벡터 pJB02-HSA를 제조하였다. 링커 서열 [GGGS]₃을 코딩하는 DNA를, HSA 코딩 DNA의 5' 말단에 프레임에 맞게 라이게이션시켜 pJB02-링커-HSA (도 7)를 제조하였다. 실시예 1a에 기재한 방법으로, Ig κ 리더 서열을 코딩하고, Val⁸-GLP-1(7-37) 서열 및 링커 서열의 5' 부분에 융합된 DNA를 제작하였다. 이 DNA를 pJB02의 *NheI* 및 *BspEI* 위치에 라이게이션시켜 pJB02-Val⁸-GLP-1-링커-HSA를 제조하였다.

<357> 실시예 1d: 엑센딘-4-Fc를 코딩하는 DNA 제조

<358> 실시예 1a에 기재한 방법으로 플라스미드 pJB02/Fc를 제조하였다. 엑센딘-4에 융합된, Ig κ 신호 서열 코딩 DNA는 하기의 중복 및 상보적 올리고뉴클레오타이드들의 시험관내 혼성화에 의해 제조되었다:

5' - CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCACCGGTCAC - 3' (서열 16)

5' - GGAGAGGGAACCTTACCAGCGACCTGAGCAAGCAGATGGAGGAGGAGGCCGT
GAGACTG - 3' (서열 17)

5' - TTCATCGAGTGGCTGAAGAACGGAGGACCAAGCAGCGGAGCCCCCTCCTCCT
AGC - 3' (서열 18)

5' - GAACCTGGAACCCAGAGCAGCAGTACCCATAGCAGGAGTGTGTCTGTCTCCA
TGGTGG - 3' (서열 19)

5' - CTCCTCCTCCATCTGCTTGCTCAGGTCGCTGGTGAAGGTTCCCTCTCCGTGA
CCGGTG - 3' (서열 20)

5' - GCTAGGAGGAGGGGCTCCGCTGCTTGGTCCTCCGTTCTTCAGCCACTCGAT
GAACAGTCTCACGGC - 3' (서열 21)

<359>

<360> 실시예 1a에 기재한 방법으로 혼성화 반응을 수행하였다. 혼성화 산물을, 실시예 1a에 기재한 것과 같이 *NheI* 및 *Eco47III*로 절단된 pJB02 벡터에 라이게이션시켜 pJB02-엑센딘-4-Fc를 제조하였다.

<361>

실시예 1e: 엑센딘-4-HSA를 코딩하는 DNA 제조

<362>

실시예 1b에 기재한 방법으로 플라스미드 pJB02-HSA를 제조하였다. 엑센딘-4에 융합된 Ig κ 신호 서열 코딩 DNA는, 실시예 1d에 기재된 것과 동일하게, 중복 및 상보적 올리고뉴클레오타이드들의 시험관내 혼성화에 의해 제조되었다. 또한 상기 혼성화 방법으로 혼성화 반응을 수행하였다. DNA를 pJB02-HSA의 단일 *NheI* 및 *FspI* 위치에 클로닝하여 pJB02-엑센딘-4-HSA를 제조하였다.

<363>

실시예 1f: 엑센딘-4-링커-HSA를 코딩하는 DNA 제조

<364>

실시예 1c에 기재한 방법으로 플라스미드 pJB02-링커-HSA를 제조하였다. 실시예 1d에 기재한 방법으로, Ig κ 신호 서열을 코딩하고 엑센딘-4와 링커 서열의 5' 부분에 융합된 DNA를 제조하였다. 이 DNA를 pJB02-링커-HSA의 단일 *NheI* 및 *BspEI* 위치에 클로닝하여 pJB02-엑센딘-4-링커-HSA를 제조하였다.

<365>

실시예 1g: Val⁸-GLP-1/C-Ex-Fc를 코딩하는 DNA 제조

<366>

실시예 1d에 기재한 방법으로 플라스미드 pJB02-엑센딘-4-Fc를 제조하였다. *AgeI* 및 *Eco47III*를 사용하여, 벡터에서 엑센딘-4 코딩 DNA를 잘라내었다. Val⁸-GLP-1/C-Ex 코딩 DNA는, 하기의 중복 및 상보적 올리고뉴클레오타이드들의 시험관내 혼성화에 의해 제조되었다:

5' -CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCTCTATCTGGA
GGGCCAGGCCGCCA - 3' (서열 22)

5' - AGGAATTCATCGCCTGGCTGGTGAAGGGCCGGGCGAGCAGCGG
AGCCCCCTCTCCTAGC - 3' (서열 23)

5' - CTCCAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCTCCAC
GTGA - 3' (서열 24)

5' - GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGCCCTTACCAGCCAGGCGA
TGAATTCCTTGGCGGCCTGGCC - 3' (서열 25)

<367>

<368> 실시예 1a에 기재한 방법으로 혼성화 반응을 수행하였다. 혼성화 산물을 상기 pJB02-엑센딘-4-Fc 발현 벡터의 엑센딘-4 위치에 라이게이션시켜 pJB02-Val⁸-GLP-1/C-Ex-Fc를 제조하였다.

<369>

실시예 1h: Val⁸-Glu²²-GLP-1-Fc를 코딩하는 DNA 제조

<370>

실시예 1d에 기재한 방법으로 플라스미드 pJB02-엑센딘-4-Fc를 제조하였다. *AgeI* 및 *Eco47III*를 사용하여, 벡

터에서 엑센딘-4 코딩 DNA를 잘라내었다. Val⁸-Glu²²-GLP-1 코딩 DNA는, 하기의 중복 및 상보적 올리고뉴클레오타이드들의 시험관내 혼성화에 의해 제조되었다:

5' -CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGCTCCTATCTCGA
GGAGCAGGCCGCCA - 3' (서열 26)

5' - AGGAGTTCATCGCCTGGCTGGTGAAGGGCCGGGGC - 3' (서열 27)

5' - GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGCGGGC
TGCTC - 3' (서열 28)

5' - CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCT
CCACGTGA - 3' (서열 29)

<371>

<372>

실시예 1a에 기재한 방법으로 혼성화 반응을 수행하였다. 혼성화 산물을 상기 pJB02-엑센딘-4-Fc 발현 벡터의 엑센딘-4 위치에 라이게이션시켜 pJB02-Val⁸-Glu²²-GLP-1-Fc를 제조하였다.

<373>

실시예 1i: Val⁸-Glu²²-GLP-1/C-Ex-Fc를 코딩하는 DNA 제조

<374>

실시예 1d에 기재한 방법으로 플라스미드 pJB02-엑센딘-4-Fc를 제조하였다. AgeI 및 Eco47III를 사용하여, 벡터에서 엑센딘-4 코딩 DNA를 잘라내었다. Val⁸-Glu²²-GLP-1/C-Ex 코딩 DNA는, 하기의 중복 및 상보적 올리고뉴클레오타이드들의 시험관내 혼성화에 의해 제조되었다:

5' - CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGCTCCTATCTCGA
GGAGCAGGCCGCCA - 3' (서열 30)

5' - AGGAATTCATCGCCTGGCTGGTGAAGGGCCGGGGCAGCAGCGGA
GCCCCCTCCTCCTAGC - 3' (서열 31)

5' - CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCC
TCCACGTGA - 3' (서열 32)

5' - GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGCCCTTCACCAGCCAGGCGA
TGAATTCCTTGCGGCCTGCTC - 3' (서열 33)

<375>

<376>

실시예 1a에 기재한 방법으로 혼성화 반응을 수행하였다. 혼성화 산물을 상기 pJB02-엑센딘-4-Fc 발현 벡터의 엑센딘-4 위치에 라이게이션시켜 pJB02-Val⁸-Glu²²-GLP-1/C-Ex-Fc를 제조하였다.

<377>

실시예 1j: Gly⁸-GLP-1-Fc를 코딩하는 DNA 제조

<378>

실시예 1d에 기재한 방법으로 플라스미드 pJB02-엑센딘-4-Fc를 제조하였다. AgeI 및 Eco47III를 사용하여, 벡터에서 엑센딘-4 코딩 DNA를 잘라내었다. Gly⁸-GLP-1 코딩 DNA는, 하기의 중복 및 상보적 올리고뉴클레오타이드들의 시험관내 혼성화에 의해 제조되었다:

5' - CCGGTCACGGCGAGGGCACCTTCACTAGTGACGTGCTCCTATCTGGA
GGGCCAGGCCGCCA - 3' (서열 34)

5' - AGGAGTTCATCGCCTGGCTGGTGAAGGGCCGGGGC - 3' (서열 35)

5' - CTCCAGATAGGAGGACACGTCACCTAGTGAAGGTGCCCTC
GCCGTGA - 3' (서열 36)

5' - GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGCGGGC
CTGGCC - 3' (서열 37)

<379>

<380>

실시예 1a에 기재한 방법으로 혼성화 반응을 수행하였다. 혼성화 산물을 상기 pJB02-엑센딘-4-Fc 발현 벡터의

엑센딘-4 위치에 라이게이션시켜 pJB02-Gly⁸-GLP-1-Fc를 제조하였다.

<381> 실시예 2: 이중성 융합 단백질의 발현

<382> HEK 293 EBNA 세포 (부착 및 부유 세포)를 일시적 형질감염시켜, 실시예 1에서 제조된 DNA에 의해 코딩되는 융합 단백질을 발현시켰다. 세포 수를 계수하고, 시딩 (seeding)하여 24시간 후에 형질감염시켰다. 퓨진 (FuGene, 상표명) 6 형질감염 시약 (로쉬 몰레큘라 바이오케미칼스 (Roche Molecular Biochemicals) 社, 카다로그 번호 1814443)을 옵티엠이엠 (OptiMEM; 기브코/비알엘 (Gibco/BRL) 社)과 혼합하고, 실온에서 5분 동안 인큐베이션한 후 DNA를 첨가하여 형질감염 혼합물을 제조하고, 이 혼합물을 추가로 15분 동안 더 인큐베이션시켰다. 형질감염 직전에 새로운 배양 배지를 플레이트에 첨가하였다. 형질감염에 대해 표 1 및 표 2에 더 자세히 나타내었다.

표 1

<383>

293 EBNA 세포의 일시적 형질감염에 사용된 시약					
조직배양 접시	시딩한 세포수	DNA (μ g)	퓨진 (Fugene) (μ l)	옵티엠이엠 (OptiMEM) 배지 (ml)	배양배지의 부피 (ml)
35 mm	5×10^5	1.5	9	0.102	2
100 mm	2×10^6	12	73	0.820	10
700 cm ² (RB)	2×10^7	65	400	4.0	100

표 2

<384>

배지의 조성	
배양 및 형질감염 배지	수거용 배지
DMEM:F12 (3:1)	하이프리트랙 (Hybritech) 염기
5 % FBS	1 mM Ca ²⁺
20 mM HEPES	20 mM HEPES
2 mM L-글루타민	1 μ g/ml 누셀린 (Nuselin; 인간 인슐린)
50 μ g/ml 제네티신 (geneticin; G418 NEO)	1 μ g/ml 인간 트랜스페린 (transferrin)
50 μ g/ml 토브로마이신 (tobromycin)	50 μ g/ml 토브로마이신

<385> 소규모 형질감염 (35 mm 내지 100 mm 접시)의 경우, 형질감염 후 24시간 뒤에 세포를 PBS로 세척하고 수거용 배지를 첨가하고, 수일 동안 매 24시간 마다 배지를 수거하고 교체하였다. 대규모 형질감염 (700 cm² 롤러 병)의 경우, 형질감염 후 48시간 뒤에 상기 롤러 병 (Roller Bottel, RB)을 PBS로 세척하고, 수거용 배지로 교체하였다. 연속적으로 10일 이상 동안 매 24시간 마다 배지를 수거하고 교체하였다. 관례적으로, 이후의 단백질 정제에서 단지 10개의 배지 수거액만을 사용하였다.

<386> 실시예 3: 이중성 융합 단백질의 정제

<387> 실시예 3a: Val⁸-GLP-1-Fc의 정제

<388> CUNO 여과 시스템을 이용하여, 대규모 형질감염에서 얻은 약 4.5 리터의 조건 배지 (conditioned medium; 융합 단백질 발현 수준이 약 20 μ g/ml임)를 여과하고, 10 K 여과 막을 사용한 프로플럭스 접선 흐름 여과 시스템 (ProFlux tangential flow filtration system)을 이용하여 부피 250 ml로 농축시켰다. 1x PBS (pH 7.4)로 포화된 하이트랩 단백질 A 칼럼 (HiTrap protein A column)을 이용하여 분당 2 ml의 유속으로 Val⁸-GLP-1-Fc를 상기 칼럼에 결합시키고, pH 3.3의 50 mM 시트르산을 사용하여 용리하였다. 4 ml의 1x PBS 및 100 μ l의 1 M 트리 스 (pH 8)가 들어있는 튜브내에 분획 (1 ml)을 수집하였다.

- <389> SDS-PAGE, 및 조르박스 (Zorbax) C8 칼럼을 이용한 역상-HPLC에 의해 측정된 융합 단백질 함유 분획을 합하고, 이것을 1x PBS (pH 7.4)로 포화된 수퍼덱스 (Superdex) 75 60/60 칼럼에 분당 10 ml의 유속으로 로딩하였다. 양성 분획 (튜브당 20 ml)들을 수집하고 합하였다. 그 후, 합해진 분획을 0.1 % TFA/물로 포화된 C4 역상 크로마토그래피에 분당 3 ml의 유속으로 로딩하였다. 5 % B (아세트니트릴 중 0.1 % TFA)에서 100 % B의 농도구배를 사용하여 70분 동안 Val⁸-GLP-1-Fc를 용리하였다. 용리 분획 (튜브당 3 ml)들을 수집하였다. 진공 건조에 의해 아세트니트릴을 제거하고, H₂O 1 ml을 첨가하였다. 정제된 샘플 (약 32 ml)을 4 리터의 1x PBS (pH 7.4)로 2회 투석하였다.
- <390> 그 후, 투석된 샘플을 밀렉스 (MILLEX)-GV 0.22 µm 필터 유닛을 이용하여 여과하고, 280 nm에서의 흡광도를 이용하여 농도를 결정하였다.
- <391> 실시예 3b: Val⁸-GLP-1-HSA 또는 Val⁸-GLP-1-링커-HSA의 정제
- <392> CUNO 여과 시스템을 이용하여 조건 배지 약 6.5 리터 (융합 단백질 발현 수준: 약 10 µg/ml)를 여과하고, 10 K 여과 막을 사용한 프로플릭스 접선 흐름 여과 시스템을 이용하여 부피 380 ml로 농축시켰다.
- <393> 20 mM 트리스 (pH 7.4)로 포화된 패스트 플로우 Q 칼럼 (Fast Flow Q column; Pharmacia 社) 50 ml을 이용하여 분당 5 ml의 유속으로 융합 단백질을 상기 칼럼에 결합시켰다. 20 mM 트리스 (pH 7.4)와 1 M NaCl이 포함된 용액 (10 CV)의 0 %에서 50 %의 농도구배를 이용하여 단백질을 용리한 후, 100 % B (2 CV)로 용리하였다.
- <394> 융합 단백질이 포함된 분획을 합하고, 이것을 0.1 % TFA/물로 포화된 C4 역상 크로마토그래피에 분당 5 ml의 유속으로 로딩하였다. 20 % B (아세트니트릴 중 0.1 % TFA)에서 90 % B의 농도구배를 이용하여 120분 동안 융합 단백질을 용리하였다. 분획 (튜브당 3.5 ml)을 수집하였다. 아세트니트릴을 진공 건조에 의해 제거하였다.
- <395> 1x PBS (pH 7.4)를 사용하여 합해진 샘플 약 9 ml을 40 ml로 희석하고, 4 리터의 1x PBS (pH 7.4)로 밤새 투석하였다. 샘플을 여과하고, 280 nm에서의 흡광도를 이용하여 농도를 결정하였다.
- <396> 실시예 3c: 엑센딘-4-Fc의 정제
- <397> CUNO 여과 시스템을 이용하여 약 4 리터의 조건 배지 (융합 단백질 발현 수준: 약 8 µg/ml)를 여과하고, 30 K 여과 막을 사용한 프로플릭스 접선 흐름 여과 시스템을 이용하여 부피 250 ml로 농축시켰다.
- <398> 1x PBS (pH 7.4)로 포화된 하이트랩 단백질 A 칼럼 5 ml을 이용하여 분당 2 ml의 유속으로 엑센딘-4-Fc를 상기 칼럼에 결합시키고, 50 mM 시트르산 (pH 3.3)을 사용하여 용리하였다. 융합 단백질이 포함된 분획을 합하고, 여과하고, 4 리터의 1x PBS로 밤새 투석하였다. 그 후 투석된 샘플을, 1x PBS (pH 7.4)와 0.5 M NaCl이 포함된 용액으로 포화된 수퍼덱스 75 60/60 칼럼에 분당 10 ml의 유속으로 로딩하였다. 융합 단백질이 포함된 분획 (튜브당 20 ml)을 수집하고, 합하고, 약 1 mg/ml의 농도로 농축하였다. 그 후, 농축한 샘플을 밀렉스-GV 0.22 µm 필터 유닛을 이용하여 여과하였다.
- <399> 실시예 3d: 엑센딘-4-HSA 및 엑센딘-4-링커-HSA의 정제
- <400> CUNO 여과 시스템을 이용하여 약 1.1 리터의 조건 배지 (융합 단백질 발현 수준: 약 6 µg/ml)를 여과하고, 30 K 여과 막을 사용한 프로플릭스 접선 흐름 여과 시스템을 이용하여 부피 175 ml로 농축시켰다.
- <401> 20 mM 트리스 (pH 7.4)로 포화된 하이트랩 Q-세파로스 칼럼 (Pharmacia 社) 5 ml을 이용하여 분당 2 ml의 유속으로 융합 단백질을 상기 칼럼에 결합시켰다. 20 mM 트리스 (pH 7.4)와 1 M NaCl이 포함된 용액 (12 CV)의 0 %에서 50 %의 농도구배를 이용하여 단백질을 용리한 후, 100 % B (4 CV)로 용리하였다.
- <402> 융합 단백질이 포함된 분획을 합하고, 0.1 % TFA/물로 포화된 C4 역상 크로마토그래피에 분당 5 ml의 유속으로 로딩하였다. 10 % B (아세트니트릴 중 0.1 % TFA)에서 100 % B의 농도구배를 이용하여 70분 동안 융합 단백질을 용리하였다. 융합 단백질이 포함된 분획 (튜브당 10 ml)을 수집하였다. 아세트니트릴을 진공 건조에 의해 제거하였다.
- <403> 합해진 샘플 약 8 ml을 4 리터의 1x PBS (pH 7.4)로 밤새 투석하였다. 이 샘플을 여과하고, 280 nm에서의 흡광도를 이용하여 농도를 결정하였다. 그 후, 투석된 샘플을 1x PBS (pH 7.4)와 0.5 M NaCl이 포함된 용액으로 포화된 수퍼덱스 200 26/60 칼럼에 분당 2 ml의 유속으로 로딩하였다. 융합 단백질이 포함된 분획 (튜브당 3

mL)을 수집하고, 합하고, 농축하고, 여과하였다.

<404> 실시예 4: SDS-PAGE를 이용한 융합 단백질의 특성 분석

<405> 정제된 융합 단백질과, 다양한 융합 단백질 발현 벡터로 형질감염된 세포들의 조건 배지를 분석하기 위해, SDS-PAGE 이후 면역블랏팅 (immunoblotting)을 이용하였다. 노벡스 (Novex) 16 % 트리스-글리신 프리캐스트 (Precast) 겔 (EC6498), 러닝 (running) 완충액 (10x, LC2675) 및 샘플 완충액 (L2676)을 사용한 노벡스 파워 이즈 (PowerEase) 500 시스템을 이용하여 SDS-PAGE를 수행하였다. 로딩하기 전에, 50 mM DTT를 사용하여 샘플을 환원시키고 95 °C에서 3분 내지 5분 동안 가열하였다.

<406> SDS-PAGE 겔을 러닝한 후, 물과 트랜스퍼 완충액 (20 % 메탄올이 포함된 1x 트리스-글리신 세프라버프 (Seprabuff); Owl Scientific 社; 카다로그 번호 ER26-S)을 사용하여 겔에서 SDS를 씻어내었다. 노벡스 트랜스퍼 장치를 이용하여 PVDF (BioRad 社, 카다로그 번호 162-0174) 및 니트로셀룰로스 막 (BioRad 社, 카다로그 번호 1703965 또는 1703932)에 트랜스퍼하였다. 실온에서 90분 동안 30 V 내지 35 V로 트랜스퍼를 수행하였다. 0.1 % 트윈 (Tween)-20 (Sigma 社, 카다로그 번호 P-7949)이 포함된 1x PBS 및 5 % 탈지유 (BioRad 社, 카다로그 번호 170-6404)를 사용하여 4 °C에서 1시간 내지 12시간 동안 막을 블로킹 (blocking)시켰다. 1x PBS+5 % 탈지유를 사용하여 항체를 희석하고, 이 용액에 상기 블랏을 4 °C에서 1시간 내지 2시간 동안 인큐베이션시켰다. 인큐베이션 사이에, 1x PBS 및 0.2 % 트윈-20을 사용하여 상기 블랏을 실온에서 5분씩 4회 세척하였다. PBS는 기브코 (GIBCO) 10x PBS (카다로그 번호 70011)를 사용하여 최종 조성이 1mM 1가 인산칼륨, 3 mM 2가 인산나트륨, 153 mM의 염화나트륨 (pH 7.4)인 것으로 만들거나, 또는 시그마 (Sigma) 社의 PBS 파워치 (카다로그 번호 1000-3)를 사용하여 최종 조성이 120 mM NaCl, 2.7 mM KCl 및 10 mM 포스페이트 (25 °C에서 pH 7.4)인 것으로 만들었다.

<407> 1차 항체로는 폴리클로날 (polyclonal) 염소 항-IgG1 항체 또는 토끼 항-HSA 항체를 사용하였다. 2차 항체로는 항-염소 IgG HRP 항체 또는 항-토끼 IgG HRP 항체를 사용하였다. 상기 2차 항체는 1:5000으로 희석되었다. 블랏을 현상하기 위해 ECL 시스템 (Amersham Pharmacia Biotech 社, 카다로그 번호 RN2108 및 RPN1674H)을 이용하였다.

<408> 도 3A는 정제된 Fc 단백질을, pJB02-Val⁸-GLP-1-Fc 및 pJB02-엑센딘-4-Fc로 형질감염된 세포의 조건 배지에서 정제된 단백질과 비교한 것이다. 이동성의 감소는 융합 단백질의 GLP-1 부분으로 인한 크기 증가와 일치한다. 이와 유사하게, 도 3B는 정제된 HSA를, pJB02-Val⁸-GLP-1-HSA, pJB02-Val⁸-GLP-1-링커-HSA, pJB02-엑센딘-4-HSA 또는 pJB02-엑센딘-4-링커-HSA로 형질감염된 세포의 조건 배지에서 정제된 단백질과 비교한 것이다. 도 4는 정제된 융합 단백질 제제를 나타낸 것이다.

<409> 실시예 5: 질량 분석기를 이용한 융합 단백질의 특성 분석

<410> 모든 실험은 시간 지연 초점 (Time Lag Focusing) 전자장치, 리플렉트론 (Reflectron; 범위 0 Da 내지 8000 Da의 펩티드 분석에 사용), 선형 검출기 (고 질량/우수 신호 분석시 사용) 및 후-가속 검출기 (Post Acceleration Detector, 또는 P.A.D.; 고 질량/매우 낮은 신호 분석에 사용)를 갖춘 마이크로매스 (Micromass) 토프스펙 (TofSpec)-2E 질량 분석기 상에서 수행하였다. 선형 모드에서 상기 기기의 유효 경로 길이는 1.2 미터이고, 리플렉트론 모드에서는 2.3 미터이다. 선형 및 리플렉트론 모드 검출용으로 2개의 이중 마이크로-채널 플레이트 검출기가 설치되어 있다. 레이저는 337 nm에서 초당 5회의 레이저 샷 (Laser shot)으로 작동하는 레이저 사이언스 社 (Laser Science Inc.)의 VSL-337i 질소 레이저를 사용하였다. 모든 데이터는 2 GHz, 8 비트 내부 계수화 장치 (digitizer)를 이용하여 획득하였고, 스펙트럼당 50회 이하의 레이저 샷들의 평균을 산출하였다.

<411> 상기 기기는 본원의 GLP-1 융합 단백질의 분석을 위해 선형 모드에서 작동되었다. 선형 검출기는 MALDI-TOF-MS 기기의 플라이트 (flight) 튜브 하단으로 이동하는 이온들을 검출하는 장치이다. 이 장치는 시간에 따른 과잉 이온의 양을 측정하여, 신호 변환을 위해 계수화 장치로 신호를 전달한다. 상기 계수화 장치는 아날로그 신호를 디지털화하는 전환장치이며, 질량 분석기의 신호가 컴퓨터에 전달되도록 하고, 여기에서 상기 신호는 사용 가능한 m/z 스펙트럼으로 재구성된다.

<412> 재결정화된 포화 시나핀산 (sinapinic acid) 용액 (50/50 Acn/H₂O 및 0.1 % TFA에 희석됨)을 이온화 매트릭스로서 이용하였다. 시나핀산은 10 kDa 초과 단백질에 대해 적합한 매트릭스이다. 분석된 샘플의 정확한 질량 측정을 위해, 고유 질량을 갖는 대조 단백질을 내부 및 외부 보정 화일에 사용하였다. 샘플을 매트릭스에 대해 1:2로 희석하여 모두 분석하였다. 상기 기기는 최초로 하기 선형 검출기 조건으로 설정하였다.

- <413> 공급원 전압 : 20.0 keV 펄스 전압 : 3.0 keV
- <414> 추출 전압 : 20.0 kev 조립성 레이저 (Laser Coarse) : 50
- <415> 초점 전압 : 16.0 kev 미립성 레이저 (Laser Fine) : 50
- <416> 선형 검출기 : 3.7 kev
- <417> P.A.D. : (오프 라인)
- <418> 최상의 신호/잡음 비율 및 최고 해상도를 얻기 위해, 상기 설정을 (필요에 따라) 변형하였다. 표 3은 서로 다른 GLP-1 융합 단백질의 특성 분석을 나타낸 것이다.

표 3

융합 단백질	예측 질량 (KDa)	질량 분석기로 측정된 질량 (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
Val ⁸ -Glu ²² -GLP-1-CEX-IgG1	60.69	65.86
엑센딘-4-IgG1	60.69	65.86
Val ⁸ -GLP-1-링커-HSA	70.70	69.89, 70.74
엑센딘-4-HSA	70.56	70.62
엑센딘-4-링커-HSA	71.56	71.62

CEX는 C-말단 연장을 말하고, 서열 Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser을 포함한다.

링커는 Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser-Gly- Gly-Gly-Gly-Ser 이다.

- <419>
- <420> 실시예 6: 이중성 융합 단백질의 활성
- <421> EP 제619,322호 (Gelfand, et al.) 및 미국특허 제5,120,712호에 기재되어 있는 것과 같은 시험관내 분석법을 이용하여, 본 발명의 융합 단백질의 GLP-1 수용체 활성화 능력을 평가하였다. Val⁸-GLP-1(7-37)OH의 활성화에 대한 상기 화합물들의 활성을 표 4에 나타내었다. 도 8에는 Val⁸-GLP-1 및 엑센딘-4 융합 단백질에 대한 시험관내 투여량 반응 곡선을 나타내었다. 또한, 표 5a 및 표 5b에는 생물학적으로 활성을 지닌 융합 단백질을 제조하기 위해 Fc 또는 알부민 단백질에 융합될 수 있는 많은 GLP-1 유사체들의 시험관내 활성을 나타내었다. 상기 활성을 GLP-1(7-37)OH와 비교하였다.

표 4

GLP-1 융합 단백질의 시험관내 활성

융합 단백질	시험관내 활성 (Val ⁸ -GLP-1의 %)
Val ⁸ -GLP-1-IgG1	1
엑센딘-4-IgG1	240
Val ⁸ -GLP-1-링커-HSA	0.2
엑센딘-4-HSA	20
엑센딘-4-링커-HSA	90
엑센딘-4	500
Val ⁸ -Glu ²² -GLP-1-IgG1	3.7
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-링커-IgG1	150
엑센딘-4-C2-IgG1	250
엑센딘-4-링커-IgG1	330
Gly ⁸ -Glu ²² -GLP-1-CEX-링커-HSA	4
Gly ⁸ -Glu ²² -GLP-1-CEX-링커-IgG4	80

CEX는 C-말단 연장을 말하고, 서열 Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser 을 포함한다.

링커는 Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser-Gly- Gly-Gly-Gly-Ser 이다.

C2는 Ser-Ser-Gly-Ala-Ser-Ser-Gly-Ala이다.

표 3 및 표 4에 기재된 융합 단백질의 아미노산 서열을 서열 13 내지 서열 31에 나타내었다.

Val⁸-GLP-1-인간 혈청 알부민 아미노산 서열을 서열 13에 나타내었다.

```

1 HVEGTFSTDV SSYLEGQAAK EFWIWLKGR GDAHKSEVAH RFDKLGEEFN KALVLIAPAQ
61 YLQQCPFEDH VKLVNEVTEF AKTCVADESA ENCDKSLHTL FGDKLCTVAT LRETYGEMAD
121 CCAKQEPERN ECFLQHKDDN PNLPRLVNPE VDMCTAFHD NEETFLKKYL YEIARRHPYF
181 YAPELLFFAK RYKAAPTECC QAADKAACLL PKLDELDEG KASSAKQRLK CASLQKFGER
241 AFKAWAVARL SQRFPAEFA EVSKLVTDLT KVHTECHGD LLECADDRAD LAKYICENQD
301 SISKLEKCC EKPLLEKSHC IAEVENDEMP ADLPSLAADF VESKDVCKNY AEAKDVFLGM
361 FLYEYARRHP DYSVVLLRL AKTYETLEK CCAAADPHC YAKVFDEFKP LVEEPQNLIK
421 QNCELFEQLG EYKFNALLV RYTKVPQVS TPTLVEVSRN LGKVGSKCKC HPEAKRMPCA
481 EDYLSVVLNQ LCVLHEKTPV SDRVTKCCTE SLVNRPRCFS ALEVDETVVP KEPNAETFTF
541 HADICTLSEK ERQIKQTAL VELVKHKPKA TKEQLKAVMD DPAFVEKCC KADKETCFA
601 EEGKLVAAAS QAALGL (서열 13)

```

Val⁸-GLP-1-링커-인간 혈청 알부민 아미노산 서열을 서열 14에 나타내었다.

```

1 HVEGTFSTDV SSYLEGQAAK EFWIWLKGR GGGGSGGGG SGGGSDAHK SEVAHRFKDL
61 GEENFKALVL IAFAYLQQC PFEDHVKLVN EVTEFAKTCV ADESAENCDK SLHTLFGDKL
121 CTVATLRETY GEMADCCAKQ EPERNECFLO HKDDNPNLPR LVRFEVDVMC TAFHDNEETF
181 LKKYLYEIAR RHPYFYAPEL LFFAKRYKAA FTECCQAADK AACLLPKLDE LRDEGKASSA
241 KQRLKASLQ KGERAFKAW AVARLSQRF KAFAEVSKL VTDLTKVHTE CCHGDLLECA
301 DDRADLAKYI CENQDSISK LKCECEKPLL EKSHCIAEVE NDMPADLPS LAADFVESKD
361 VCKNYAEAKD VFLGMFLYFY ARRHDPYSV LLLRLAKTYE TTLEKCCAAA DPHECYAKVF
421 DEFKPLVEEP QNLKQNCLE FEQLGEYKFO NALLVRYTKK VPQVSTPTLV EVSRNLGKVG
481 SKCKHPEAK RMPCAEDYLS VVLNQLCVLH EKTPVSDRVT KCCTESLVNR RPCFSALEVD
541 ETVPKEFNA ETFTPHADIC TLSEKERQIK QTALVELVK HKPKATKEQL KAVMDDFAAF
601 VEKCKKADK ETCFAEEGKK LVAASQAALG L (서열 14)

```

<428> Gly⁸-Glu²²-GLP-1-CEx-링커-인간 혈청 알부민 아미노산 서열을 서열 15에 나타내었다.

```

1 HEGTFTSDV SSYLEQAAK EFLAWLVKGR GSSGAPPPSG GGGGSGGGGS GGGGSDAHKS
61 EVAHRFKDLG EENFKALVLI AFAQYLQQCP FEDHVKLVNE VTEFAKTCVA DESAENCDSK
121 LHTLFGDKLC TVATLRETYG EMADCCAKQE PERNECFLQH KDDNPNLPR LVRPEVDVMCT
181 AFHDNEETFL KKYLYEIARR HPYFYAPELL FFAKRYKAAF TECCQAADKA ACLLPKLDEL
241 RDEGKASSAK QRLKASLQK FGERAFKAWA VARLSQRFPK AEFAEVSKLV TDLTKVHTEC
301 CHGDLLECAD DRADLAKYIC ENQDSISSKL KECCEKPLLE KSHCIAEVEN DEMPADLPSL
361 AADFVESKDV CKNYAEAKDV FLGMFLYEYA RRHPDYSVVL LLRLAKTYET TLEKCCAAAD
421 PHECYAKVFD EFKPLVEEPQ NLIKQNCLEF EQLGEYKFQN ALLVRYTKKV PQVSTPTLVE
481 VSRNLGKVG S KCKKHPEAKR MPCAEDYLSV VLNQLCVLHE KTPVSDRVTK CCTESLVNRR
541 PCFSALEVDE TYVPKEFNAE TPTFHADICT LSEKERQIKK QTALVELVKH KPKATKEQLK
601 AVMDFFAAVF EKCKKADDE TCFABEGKKL VAASQAALGL (서열 15)

```

<429>

<430> 엑센딘-4-인간 혈청 알부민 아미노산 서열을 서열 16에 나타내었다.

```

1 HEGTFTSDL SKQMEEAVR LFIEWLKNNG PSSGAPPPSD AHKSEVAHRF KDLGEENFKA
61 LVLIIFAQYL QQCPFEDHVK LVNEVTEFAK TCVADESAEN CDKSLHTLFG DKLC TVATLR
121 ETYGEADCC AKQEPERNEC FLQHKDDNPN LPRLVRPEVD VMCTAFHDNE ETFLKKYLYE
181 IARRHPYFYA PELLFFAKRY KAAFTECCQA ADKAACLLPK LDEL RDEGKA SSAQRLKCA
241 SLQKFGERAF KAWAVARLSQ RFPKAFAEV SKLVTDLT KV HTECCHGDL ECAADDRADLA
301 KYICENQDSI SSKLKECCEK PLLEKSHCIA EVENDEMPAD LPSLAADFVE SKDVCKNYAE
361 AKDVFLGMPL YEYARRHPDY SVLLLLRLAK TYETTLEKCC AADPHECYA KVFDEFKPLV
421 EEPQNLIKQN CELFEQLGEY KFNALLVRY TKKVPQVSTP TLVEVSRNLG KVGSKCKKHP
481 EAKRMPCAED YLSVVLNQLC VLHEKTPVSD RVTCKCTESL VNRRPCFSAL EVDETYVPKE
541 FNAETFTFHA DICTLSEKER QIKKQTALVE LVKHKPKATK EQLKAVMDDF AAFVEKCKKA
601 DDKETCFABE GKKLVAASQA ALGL (서열 16)

```

<431>

<432> 엑센딘-4-링커-인간 혈청 알부민 아미노산 서열을 서열 17에 나타내었다.

```

1 HEGTFTSDL SKQMEEAVR LFIEWLKNNG PSSGAPPPSG GGGGSGGGGS GGGGSDAHKS
61 EVAHRFKDLG EENFKALVLI AFAQYLQQCP FEDHVKLVNE VTEFAKTCVA DESAENCDSK
121 LHTLFGDKLC TVATLRETYG EMADCCAKQE PERNECFLQH KDDNPNLPR LVRPEVDVMCT
181 AFHDNEETFL KKYLYEIARR HPYFYAPELL FFAKRYKAAF TECCQAADKA ACLLPKLDEL
241 RDEGKASSAK QRLKASLQK FGERAFKAWA VARLSQRFPK AEFAEVSKLV TDLTKVHTEC
301 CHGDLLECAD DRADLAKYIC ENQDSISSKL KECCEKPLLE KSHCIAEVEN DEMPADLPSL
361 AADFVESKDV CKNYAEAKDV FLGMFLYEYA RRHPDYSVVL LLRLAKTYET TLEKCCAAAD
421 PHECYAKVFD EFKPLVEEPQ NLIKQNCLEF EQLGEYKFQN ALLVRYTKKV PQVSTPTLVE
481 VSRNLGKVG S KCKKHPEAKR MPCAEDYLSV VLNQLCVLHE KTPVSDRVTK CCTESLVNRR
541 PCFSALEVDE TYVPKEFNAE TPTFHADICT LSEKERQIKK QTALVELVKH KPKATKEQLK
601 AVMDFFAAVF EKCKKADDE TCFABEGKKL VAASQAALGL (서열 17)

```

<433>

<434> Val⁸-GLP-1-IgG1 아미노산 서열을 서열 18에 나타내었다.

```

1 HVEGTFTSDV SSYLEGQAAK EFLAWLVKGR GAEPKSCDKT HTCPCPAPE LLGGPSVFLF
61 PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VNAKTKPRE EQYNSTYRVV
121 SVLTVLHQDW LNGKEYCKKV SNKALPAPIE KTISKAGQFP REPQVYTLPP SREEMTKNQV
181 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TTPVLDSDGS FFLYSKLTVD KSRWQQGNVF
241 SCSVMHEALH NHYTQKSLSL SPGK (서열 18)

```

<435>

<436> Val⁸-GLP-1-CEx-IgG1 아미노산 서열을 서열 19에 나타내었다.

```

1 HVEGTFTSDV SSYLEGQAAK EFLAWLVKGR GSSGAPPPSA EPKSCDKTHT CPFCPAPELL
61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NMYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYCKVSN KALPAPIEKT ISKAGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTT PVLDSDGSFF LYSKLTVDKS
241 RWQQGNVFSV SVMHEALHNN YTQKSLSLSP GK (서열 19)

```

<437>

<438> Val⁸-Glu²²-GLP-1-IgG1 아미노산 서열을 서열 20에 나타내었다.

```

1 HVEGTFTSDV SSYLEEQAAK EFLAWLVKGR GAEPKSCDKT HTCPPCPAPE LLGGPSVFLF
61 PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNATKPRE EQYNSTYRVV
121 SVLTVLHQDW LNGKEYCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV
181 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSGDS FFLYSLKLTVD KSRWQQGNVF
241 SCSVMHEALH NHYTQKSLSL SPGK (서열 20)

```

<439>

<440> Val⁸-Glu²²-GLP-1-CEX-IgG1 아미노산 서열을 서열 21에 나타내었다.

```

1 HVEGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGAPPPSA EPKSCDKTHT CPPCPAPELL
61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTP PVLDSGDSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHNNH YTQKSLSLSP GK (서열 21)

```

<441>

<442> Gly⁸-Glu²²-GLP-1-C2-IgG1 아미노산 서열을 서열 22에 나타내었다.

```

1 HVEGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGASSGAA EPKSCDKTHT CPPCPAPELL
61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTP PVLDSGDSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHNNH YTQKSLSLSP GK (서열 22)

```

<443>

<444> Gly⁸-Glu²²-GLP-1-CEX-링커-IgG1 아미노산 서열을 서열 23에 나타내었다.

```

1 HVEGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGAPPPSG GGGSGGGGSG GGSAEPKSC
61 DKTHTCPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
121 GVEVHNATK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTPPVLDS
241 DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNNHYTQKS LSLSPGK (서열 23)

```

<445>

<446> Gly⁸-Glu²²-GLP-1-CEX-링커-IgG4 아미노산 서열을 서열 24에 나타내었다.

```

1 HVEGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGAPPPSG GGGSGGGGSG GGSAESKYG
61 PPCPCPAPE FLGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSEDPEV QFNWYVDGVE
121 VHNATKPRE EQFNSTYRVV SVLTVLHQDW LNGKEYCKV SNKGLPSSIE KTISKAKGQP
181 REPQVYTLPP SQEEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSGDS
241 FFLYSRLTVD KSRWQEGNVF SCSVMHEALH NHYTQKSLSL SLGK (서열 24)

```

<447>

<448> Gly⁸-Glu²²-GLP-1-CEX-2링커-IgG1 아미노산 서열을 서열 25에 나타내었다.

```

1 HVEGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGAPPPSG GGGSGGGGSG GGGSGGGGSG
61 GGGSGGGGSA EPKSCDKTHT CPPCPAPELL GGPSVFLFPP KPKDTLMISR TPEVTCVVVD
121 VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYCKVSN
181 KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG
241 QPENNYKTP PVLDSGDSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNNH YTQKSLSLSP
301 GK (서열 25)

```

<449>

<450> Gly⁸-Glu²²-GLP-1-2링커-IgG1 아미노산 서열을 서열 26에 나타내었다.

```

1 HVEGTFTSDV SSYLEEQAAK EFLAWLVKGR GGGGGSGGGG SGGGGSGGGG SGGGGSGGGG
61 SAEPKSCDKT HTCPPCPAPE LLGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV
121 KFNWYVDGVE VHNATKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYCKV SNKALPAPIE
181 KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT
241 TPPVLDSGDS FFLYSLKLTVD KSRWQQGNV SCSVMHEALH NHYTQKSLSL SPGK

```

<451> (서열 26)

<452> Gly⁸-Glu²²-GLP-1-2CEX-IgG1 아미노산 서열을 서열 27에 나타내었다.

```

1 HEGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGAPPPSS SGAPPPSAEP KSCDKTHTCP
61 PCPAPELLGG PSVFLFPPKP KDTLMISRTPEVTCTVVDVS HEDPEVKFNW YVDGVEVHNA
121 KTKPREEQYN STYRVSVLT VLHQDWLNGK EYCKKVSNA LPAPIEKTIS KAKGQPREPQ
181 VYTLPPSREE MTKNQVSLTCLVKGFYPSDI AVEWESNGQP ENNYKTTPPV LDSGGSFFLY
241 SKLTVDKSRW QQGNVFSCSV MHEALHNHYT QKSLSLSPGK (서열 27)

```

<453>

<454> Gly⁸-Glu²²-Val²⁵-Ile³³-GLP-1-CEX-링커-IgG1 아미노산 서열을 서열 28에 나타내었다.

```

1 HEGTFTSDV SSYLEEQAVK EFLAWLIKGR GSSGAPPPSG GGGSGGGGSG GGSAAEPKSC
61 DKTHTCPPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
121 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTTPVLDS
241 DGSFFLYSKL TVDKSRWQQG NVFSCSVME ALHNHYTQKS LSLSPGK (서열 28)

```

<455>

<456> 엑센딘-4-IgG1 아미노산 서열을 서열 29에 나타내었다.

```

1 HEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSA EPKSCDKTHT CPPCPAPELL
61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK (서열 29)

```

<457>

<458> 엑센딘-4-C2-IgG1 아미노산 서열을 서열 30에 나타내었다.

```

1 HEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGASSGAA EPKSCDKTHT CPPCPAPELL
61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK (서열 30)

```

<459>

<460> 엑센딘-4-링커-IgG1 아미노산 서열을 서열 31에 나타내었다.

```

1 HEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGSAAEPKSC
61 DKTHTCPPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
121 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTTPVLDS
241 DGSFFLYSKL TVDKSRWQQG NVFSCSVME ALHNHYTQKS LSLSPGK (서열 31)

```

<461>

표 5aa

시험관내 GLP-1 유사체 활성

<u>GLP-1 화합물</u>	<u>GLP-1 수용체 활성화도</u>
GLP-1 (7-37) OH	1.0
Val ⁸ -GLP-1 (7-37) OH	0.47 (n = 6)
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

<462>

표 5ab

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

<463>

표 5ac

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

<464>

표 5ad

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

<465>

표 5ba

시험관내 GLP-1 유사체 활성

GLP-1 화합물	GLP-1 수용체 활성화도
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
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
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

<466>

표 5bb

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
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

표 6

GLP/엑셀던 유사체의 시험관내 활성

펩티드 서열	시험관내 활성 (Val ⁸ -GLP-1(7-37)OH의 %)
HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGP-NH2	6.21
HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPP PS-NH2	6.75, 3.25
HVEGTFTSDLSKQMEEEAVRLFIAWLVKGRG	2.86
HVEGTFTSDVSSYLEEEAVRLFIAWLVKGRG	1.47
HVEGTFTSDLSKQMEGQAAKEFIAWLVKGRG	0.11
HVEGTFTSDVSKQMEGQAAKEFIAWLVKGRG	0.04
HGEGTFTSDLSKQMEGQAAKEFIEWLKNGGP-NH2	1.44
HGEGTFTSDLSKQMEEEAAKEFIEWLKNGGP-NH2	2.80
HGEGTFTSDVSSYLEEEAVRLFIEWLKNGGP-NH2	5.40
HGEGTFTSDLSVSSYLEEEAVRLFIEWLKNGGP-NH2	5.07
HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRPSSGAPPPS -NH2	3.30
HAEGTFTSDVSKQLEEEAAKEFIAWLVKGRG	2.15
HVEGTFTSDVSSYLEGQAAKEFIEWLKNGGP-NH2	2.36
HGEGTFTSDLSKQMEEEAVRLFIAWLVKGRG	3.25
HVEGTFTSDVSSYLEEEAAKEFIAWLVKGRG	1.00
HVEGTFTSDVSSYLEGQAAKEFIAWLKNGRG	0.20
HVEGTFTSDVSSYLEGQAAKEFIAWLVKGRG	1.00
HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG	2.12

실시예 7: Val⁸-GLP-1-IgG1 및 Val⁸-GLP-1-HSA의 생체내 약물동력학

키노몰구스 원숭이 (cynomolgus monkey)를 이용하여 Val⁸-GLP-1-IgG1 및 Val⁸-GLP-1-HSA에 대한 약물동력학적

연구를 수행하였다. 정제된 Val⁸-GLP-1-IgG1 또는 Val⁸-GLP-1-HSA를 5.6 nmol/kg의 양으로 원숭이에게 투여하였다. 상기 화합물은 정맥내 볼루스 투여의 방법으로 투여되었다. 화합물 투여 전, 및 투여 후 0.083, 0.25, 0.5, 1, 4, 8, 12, 24, 48, 72, 96, 120, 144, 168 및 216시간 후에 혈액을 채취하여 EDTA가 들어있는 튜브에 넣었다. Val⁸-GLP-1(7-37)의 N-말단 (7-16) 영역에 대해 1차적인 특이성을 갖는 폴리클로날 항혈청을 사용한 방사면역분석법 (radioimmunoassay)을 이용하여 면역반응성 Val⁸-GLP-1의 혈장내 농도를 측정하였다. 도 9는 단일 정맥 투여량을 2마리의 키노몰구스 원숭이에게 투여한 후, Val⁸-GLP-1-Fc 및 Val⁸-GLP-1-링커-HSA의 혈장내 농도를 나타낸 것이다. Fc 융합 단백질은 약 45시간의 반감기를 갖고 있었으며, 알부민 융합 단백질은 약 87시간의 반감기를 갖고 있었다.

<471> 실시예 8: 엑센딘-4-IgG1의 생체내 약물동력학

<472> 장기간에 걸쳐 캐놀러 (cannular)가 삽입된 보통의 수컷 비글 (beagle) 개 2마리를 밤새 금식시킨 후 시험하였다. 동맥 및 정맥 혈관 접근 포트가 이용되었고, 피부를 관통하여 카테테르 (catheter)를 두부 정맥에 삽입하고 고정시켰다. 동물을 우리에게 가두고, 그의 카테테르를 스윙벨 (swivel)/테터 (tether) 시스템에 부착하였다. 융합 단백질 엑센딘-4-IgG1 (11.8 μM)이 포함된 용액을 두부 정맥 카테테르를 통해 정맥내에 주입 (1.0 mmol/kg)하였다. 그 후, 염수 10 mL을 사용하여 상기 카테테르를 세정하였다. 2시간 후, 고혈당성 (150 mg/dl) 클램프 (hyperglycemic clamp)를 시작하여, 3시간 동안 계속하였다. 융합 단백질, 글루코스 및 인슐린의 혈장내 농도를 측정하기 위해, 동맥혈 샘플을 상기의 5시간에 걸쳐 계속 채취하였다.

<473> 상기 시험 결과를 이전의 유사 시험, 즉 피하 주사로 상기 2마리 동물에게 식염 볼루스를 투여하고, 3시간 후에 이 3시간 동안의 고혈당성 (150 mg/dl) 클램프를 이용한 시험 결과와 비교하였다.

<474> 상기 두 세트의 시험에서, 혈장내 글루코스 농도는 베크만 (Beckman) 글루코스 분석기를 이용하여 측정하였다. 혈장내 인슐린 농도는 린코 리서치 사 (Linco Research, Inc)의 사원이 소속 회사 연구실에서 개발된 RIA 키트를 이용하여 측정하였다. 결과를 도 10 및 도 11에 나타내었다.

<475> 실시예 9: Gly⁸-Glu²²-GLP-1-CEX-링커-IgG1의 생체내 약물동력학

<476> 보통의 수컷 비글 개 3마리로 구성된 2개의 군에 Gly⁸-Glu²²-GLP-1-CEX-링커-IgG1 0.1 mg/kg을 피하 (SC) 또는 정맥내 (IV)에 투여하였다. 방사면역분석법을 이용하여, 상기 IV군과 SC군에 대해 투여 전 30분에서 투여 후 216시간까지 수집한 샘플에서 면역반응성 Gly⁸-Glu²²-GLP-1-CEX-링커-IgG1의 혈장내 농도를 측정하였다. 그 후, 상기 농도를 이용하여 보고된 약물동력학적 파라미터를 결정하였다. IV 투여된 Gly⁸-Glu²²-GLP-1-CEX-링커-IgG1의 평균 소멸 반감기는 약 55시간이었고, 체내 총 소멸율 (total body clearance)은 1.5 mL/h/kg이었다. SC 투여된 Gly⁸-Glu²²-GLP-1-CEX-링커-IgG1의 평균 소멸 반감기는 약 38시간이었다.

도면의 간단한 설명

<477> 본 발명은 하기 도면을 참고로 하여 더 설명된다:

<478> 도 1: 힌지 (hinge) 영역, CH2 및 CH3 도메인을 포함하는 IgG1 Fc 아미노산 서열.

<479> 도 2: 인간 혈청 알부민 아미노산 서열.

<480> 도 3: A. IgG1-Fc 및 GLP-1-Fc 융합 단백질의 분자량을 나타내는 SDS-PAGE 겔 및 동일 겔의 면역블랏 (immunoblot)(레인 1, 표준 MW; 레인 2, 정제된 Fc; 레인 3, 위장 형질감염 (mock transfection)된 배지; 레인 4, Val⁸-GLP-1-Fc; 레인 5, 엑센딘-4-Fc). B. 인간 HSA 및 GLP-1-HSA 융합 단백질의 분자량을 나타내는 SDS-PAGE 겔 및 동일 겔의 면역블랏 (레인 1, 표준 MW; 레인 2, 정제된 HSA; 레인 3, 위장 형질감염된 배지; 레인 4, Val⁸-GLP-1-HSA; 레인 5, Val⁸-GLP-1-[Gly-Gly-Gly-Gly-Ser]₃-HSA; 레인 6, 엑센딘-4-HSA; 레인 7, 엑센딘-4-[Gly-Gly-Gly-Gly-Ser]₃-HSA).

<481> 도 4: 정제된 Fc, 알부민 및 GLP-1 융합 단백질의 SDS-PAGE 겔 (레인 1, 표준 MW; 레인 2, 정제된 Fc; 레인 3, Val⁸-GLP-1-Fc; 레인 4, 엑센딘-4-Fc; 레인 5, 표준 MW; 레인 6, Val⁸-GLP-1-HSA; 레인 7, 엑센딘-4-HSA; 레인

8, 엑센딘-4-[Gly-Gly-Gly-Gly-Ser]₃-HSA).

- <482> 도 5: 도 1에 나타난 Fc 영역을 포함하는 발현 클로닝 벡터.
- <483> 도 6: 도 2에 나타난 알부민 서열을 포함하는 발현 클로닝 벡터.
- <484> 도 7: 프레임에 맞게 융합된 15개 아미노산의 링커 및 도 2에 나타난 알부민 서열을 코딩하는 DNA를 포함하는 발현 클로닝 벡터.
- <485> 도 8: GLP-1 융합 단백질의 시험관내 투여량 반응 활성.
- <486> 도 9: GLP-1 Fc 및 HSA 융합 단백질의 약물동력학.
- <487> 도 10: 금식시킨 보통의 개 2마리에서의 엑센딘-Fc에 대한 당역학적 (glucodynamic) 반응.
- <488> 도 11: 금식시킨 보통의 개 2마리에서의 엑센딘-Fc에 대한 인슐린 분비 반응.
- <489> 도 12: 인간 IgG1 Fc 영역을 코딩하는 DNA 서열.
- <490> 도 13: 인간 알부민 단백질을 코딩하는 DNA 서열.

도면

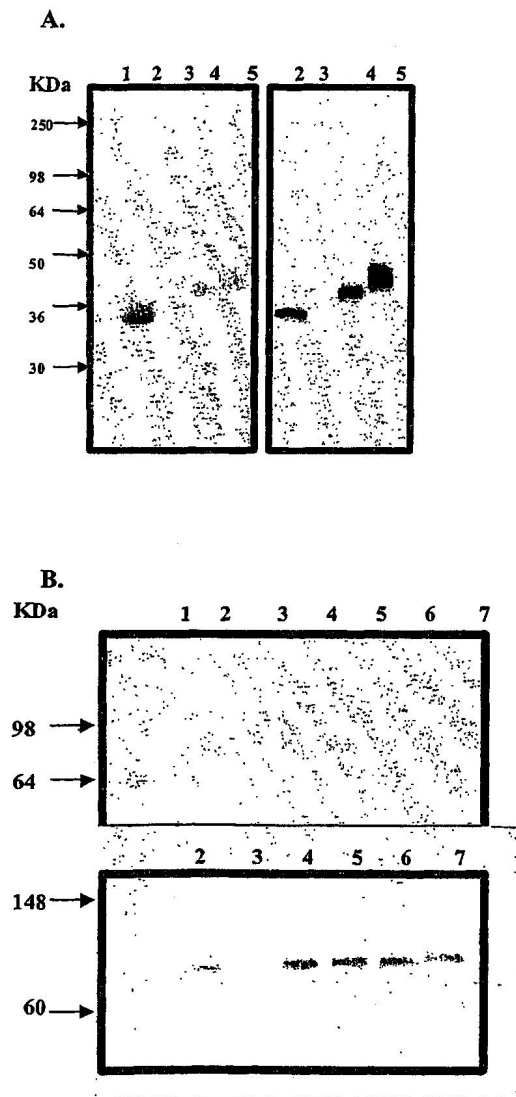
도면1

				5					10					15
Ala	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys
				20					25					30
Pro	Ala	Pro	Glu	Lys	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro
				35					40					45
Lys	Pro	Lys	Asp	Thr	Lys	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr
				50					55					60
Cys	Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe
				65					70					75
Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys
				80					85					90
Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val
				95					100					105
Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys
				110					115					120
Cys	Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr
				125					130					135
Ile	Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr
				140					145					150
Leu	Pro	Pro	Ser	Arg	Glu	Glu	Met	Thr	Lys	Asn	Gln	Val	Ser	Leu
				155					160					165
Thr	Cys	Leu	Val	Lys	Gly	Phe	Tyr	Pro	Ser	Asp	Ile	Ala	Val	Glu
				170					175					180
Trp	Glu	Ser	Asn	Gly	Gln	Pro	Glu	Asn	Asn	Tyr	Lys	Thr	Thr	Pro
				185					190					195
Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	Phe	Phe	Leu	Tyr	Ser	Lys	Leu
				200					205					210
Thr	Val	Asp	Lys	Ser	Arg	Trp	Gln	Gln	Gly	Asn	Val	Phe	Ser	Cys
				215					220					225
Ser	Val	Met	His	Glu	Ala	Leu	His	Asn	His	Tyr	Thr	Gln	Lys	Ser
				230										
Leu	Ser	Leu	Ser	Pro	Gly	Lys	(서열 32)							

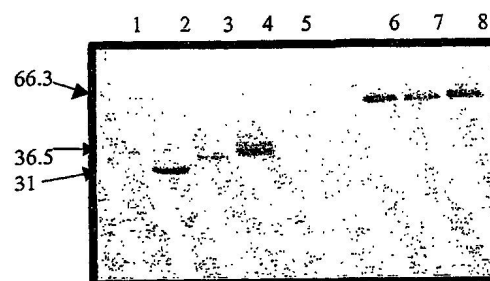
도면2

Asp	20	His	Lys	Ser	Glu	Val	Ala	His	Arg	Phe	Lys	Asp	Leu	Gly	Glu	Glu	Asn
Phe	Lys	Ala	Leu	Val	Leu	Ile	Ala	Phe	Ala	Gln	Tyr	Leu	Gln	Gln	Cys	Pro	Phe
Glu	Asp	His	Val	Lys	Leu	Val	Asn	Glu	Val	Thr	Glu	Phe	Ala	Lys	Thr	Cys	Val
Ala	Asp	Glu	Ser	Ala	Glu	Asn	Cys	Asp	Lys	Ser	Leu	His	Thr	Leu	Phe	Gly	Asp
Lys	Leu	Cys	Thr	Val	Ala	Thr	Leu	Arg	Glu	Thr	Tyr	Gly	Glu	Met	Ala	Asp	Cys
Cys	Ala	Lys	Gln	Glu	Pro	Glu	Arg	Asn	Glu	Cys	Phe	Leu	Gln	His	Lys	Asp	Asp
Asn	Pro	Asn	Leu	Pro	Arg	Leu	Val	Arg	Pro	Glu	Val	Asp	Val	Met	Cys	Thr	Ala
Phe	His	Asp	Asn	Glu	Glu	Thr	Phe	Leu	Lys	Lys	Tyr	Leu	Tyr	Glu	Ile	Ala	Arg
Arg	His	Pro	Tyr	Phe	Tyr	Ala	Pro	Glu	Leu	Leu	Phe	Phe	Ala	Lys	Arg	Tyr	Lys
Ala	Ala	Phe	Thr	Glu	Cys	Cys	Gln	Ala	Ala	Asp	Lys	Ala	Ala	Cys	Leu	Leu	Pro
Lys	Leu	Asp	Glu	Leu	Arg	Asp	Glu	Gly	Lys	Ala	Ser	Ser	Ala	Lys	Gln	Arg	Leu
Lys	Cys	Ala	Ser	Leu	Gln	Lys	Phe	Gly	Glu	Arg	Ala	Phe	Lys	Ala	Trp	Ala	Val
Ala	Arg	Leu	Ser	Gln	Arg	Phe	Pro	Lys	Ala	Glu	Phe	Ala	Glu	Val	Ser	Lys	Leu
Val	Thr	Asp	Leu	Thr	Lys	Val	His	Thr	Glu	Cys	Cys	His	Gly	Asp	Leu	Leu	Glu
Cys	Ala	Asp	Asp	Arg	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	Cys	Glu	Asn	Gln	Asp	Ser
Ile	Ser	Ser	Lys	Leu	Lys	Glu	Cys	Cys	Glu	Lys	Pro	Leu	Leu	Glu	Lys	Ser	His
Cys	Ile	Ala	Glu	Val	Glu	Asn	Asp	Glu	Met	Pro	Ala	Asp	Leu	Pro	Ser	Leu	Ala
Ala	Asp	Phe	Val	Glu	Ser	Lys	Asp	Val	Cys	Lys	Asn	Tyr	Ala	Glu	Ala	Lys	Asp
Val	Phe	Leu	Gly	Met	Phe	Leu	Tyr	Glu	Tyr	Ala	Arg	Arg	His	Pro	Asp	Tyr	Ser
Val	Val	Leu	Leu	Leu	Arg	Leu	Ala	Lys	Thr	Tyr	Glu	Thr	Thr	Leu	Glu	Lys	Cys
Cys	Ala	Ala	Ala	Asp	Pro	His	Glu	Cys	Tyr	Ala	Lys	Val	Phe	Asp	Glu	Phe	Lys
Pro	Leu	Val	Glu	Glu	Pro	Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	Glu	Leu	Phe	Glu
Asn	Leu	Gly	Glu	Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	Val	Arg	Tyr	Thr	Lys	Lys
Val	Pro	Gln	Val	Ser	Thr	Pro	Thr	Leu	Val	Glu	Val	Ser	Arg	Asn	Leu	Gly	Lys
Val	Gly	Ser	Lys	Cys	Cys	Lys	His	Pro	Glu	Ala	Lys	Arg	Met	Pro	Cys	Ala	Glu
Asp	Tyr	Leu	Ser	Val	Val	Leu	Asn	Gln	Leu	Cys	Val	Leu	His	Glu	Lys	Thr	Pro
Val	Ser	Asp	Arg	Val	Thr	Lys	Cys	Cys	Thr	Glu	Ser	Leu	Val	Asn	Arg	Arg	Pro
Cys	Phe	Ser	Ala	Leu	Glu	Val	Asp	Glu	Thr	Tyr	Val	Pro	Lys	Glu	Phe	Asn	Ala
Glu	Thr	Phe	Thr	Phe	His	Ala	Asp	Ile	Cys	Thr	Leu	Ser	Glu	Lys	Glu	Arg	Gln
Ile	Lys	Lys	Gln	Thr	Ala	Leu	Val	Glu	Leu	Val	Lys	His	Lys	Pro	Lys	Ala	Thr
Lys	Glu	Gln	Leu	Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala	Phe	Val	Glu	Lys	Cys
Cys	Lys	Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala	Glu	Glu	Gly	Lys	Lys	Leu	Val
Ala	Ala	Ser	Gln	Ala	Ala	Leu	Gly	Leu	(서열 34)								

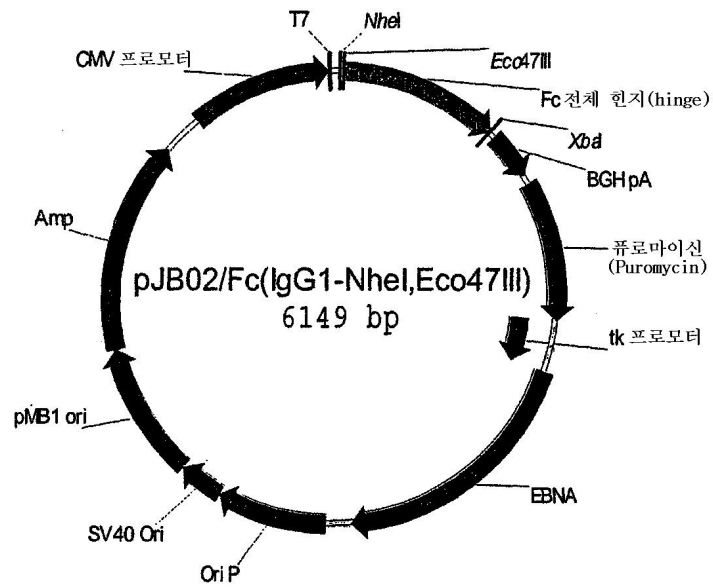
도면3



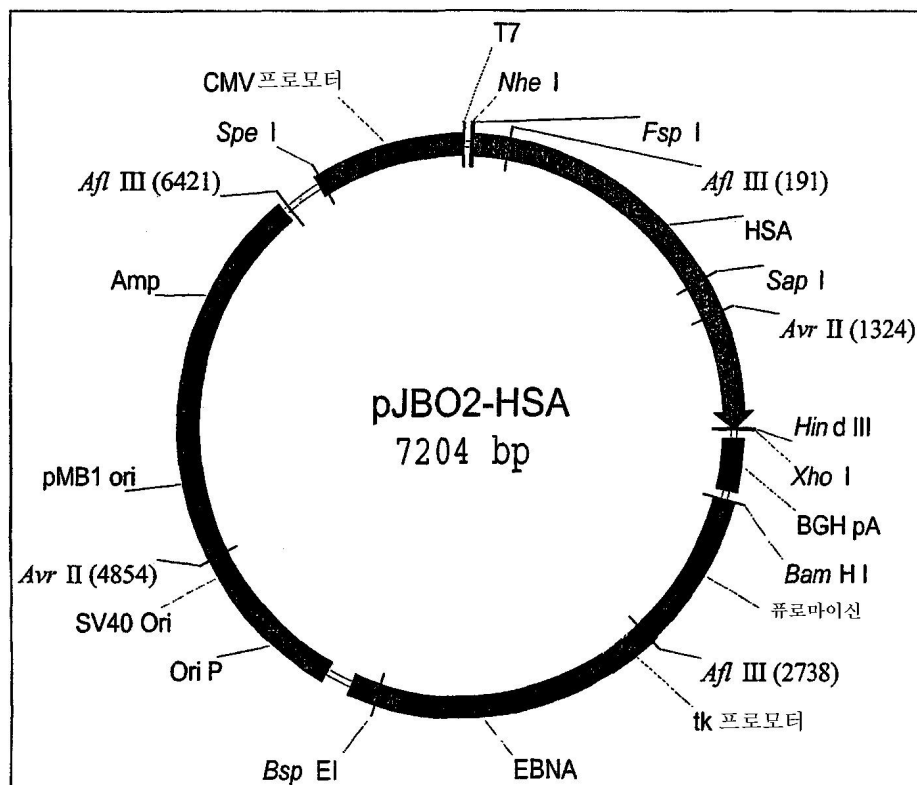
도면4



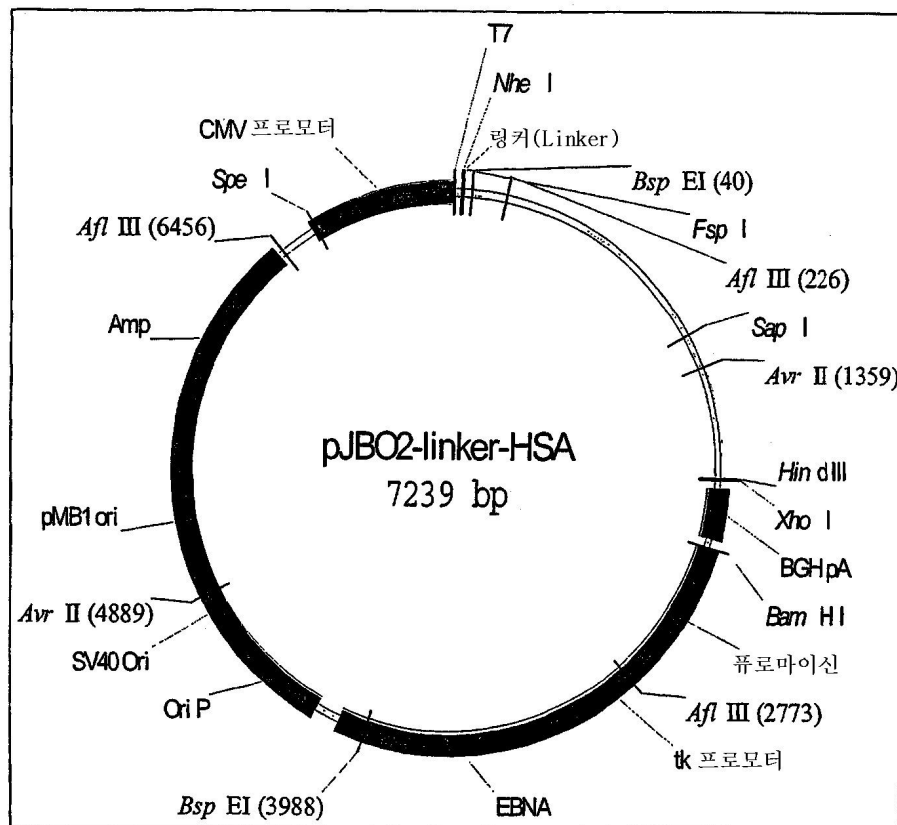
도면5



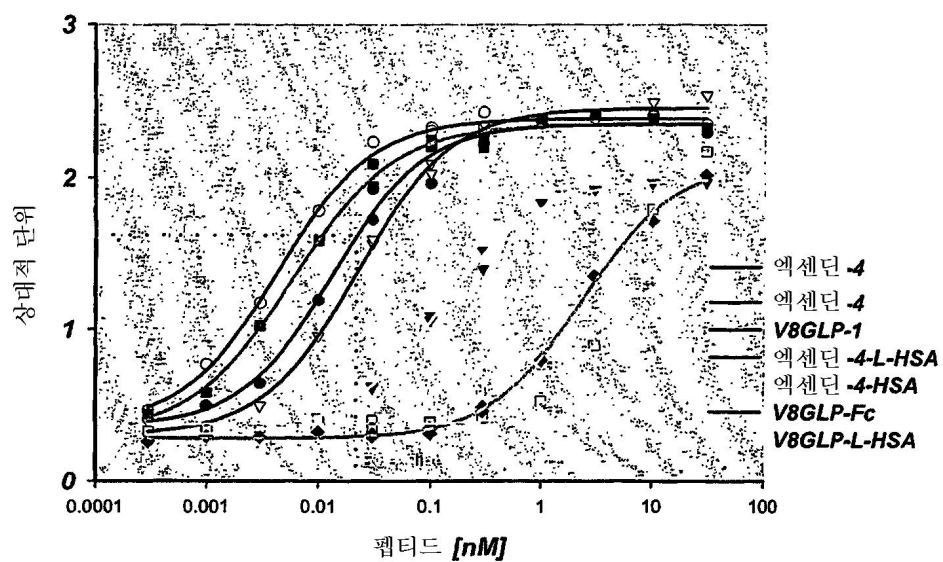
도면6



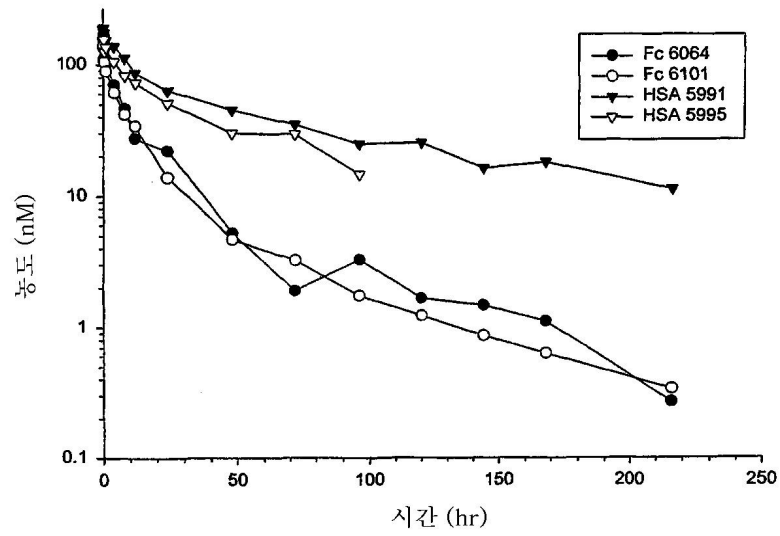
도면7



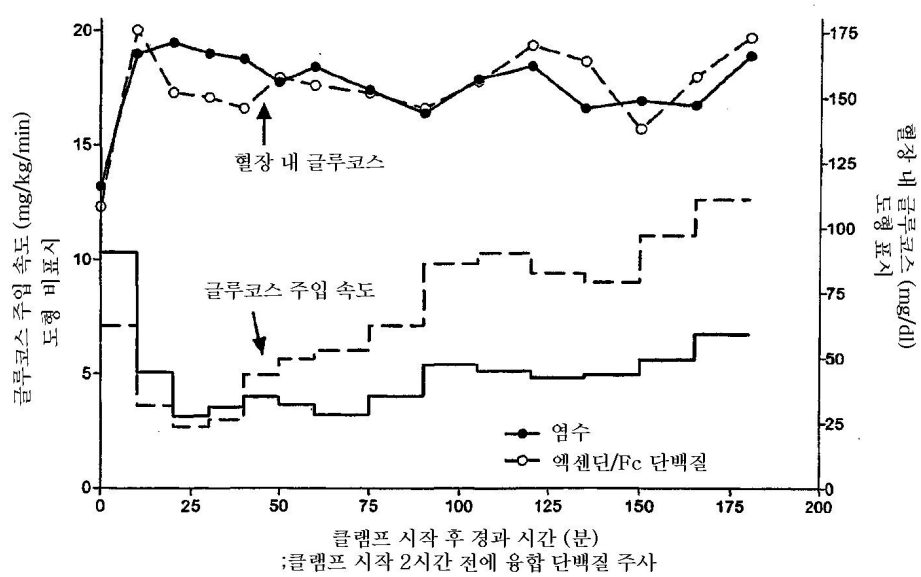
도면8



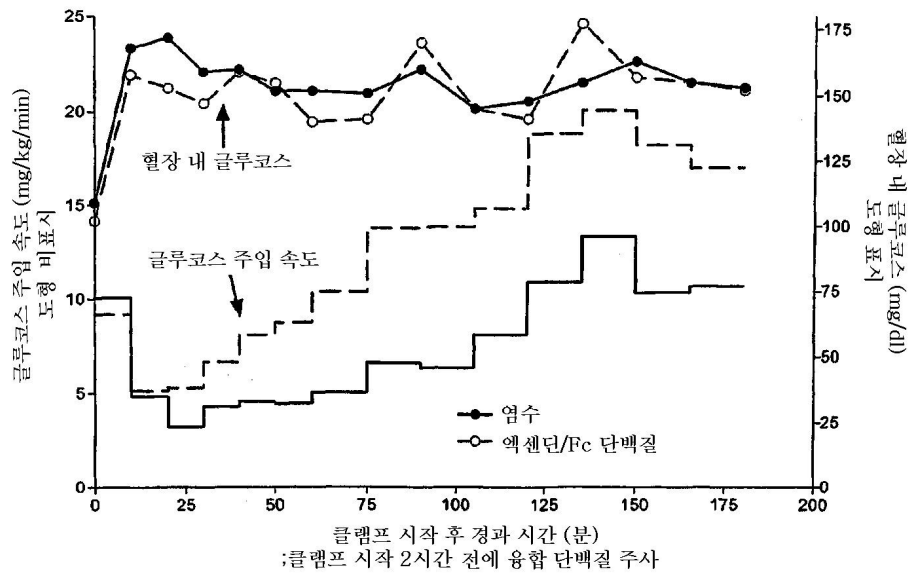
도면9



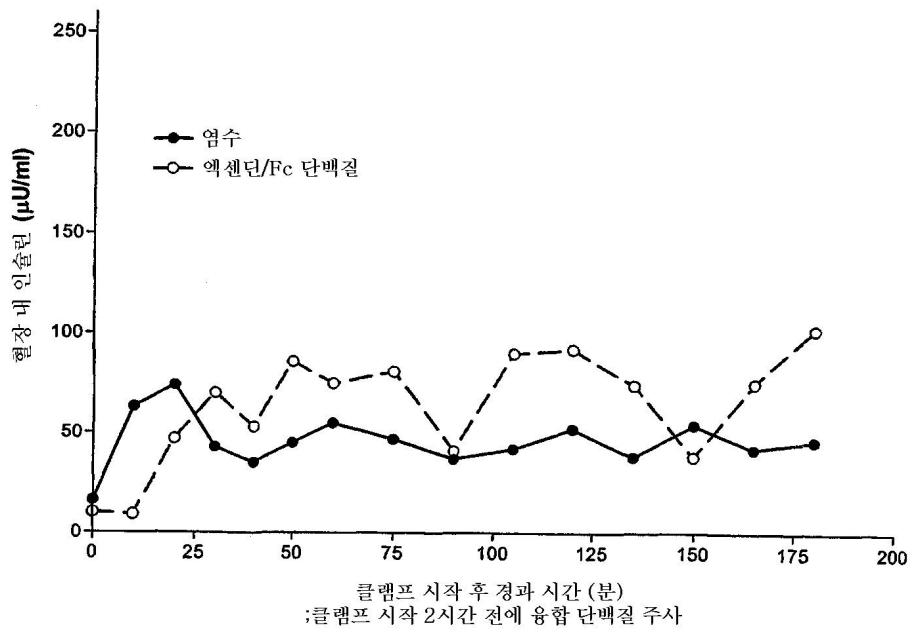
도면10a



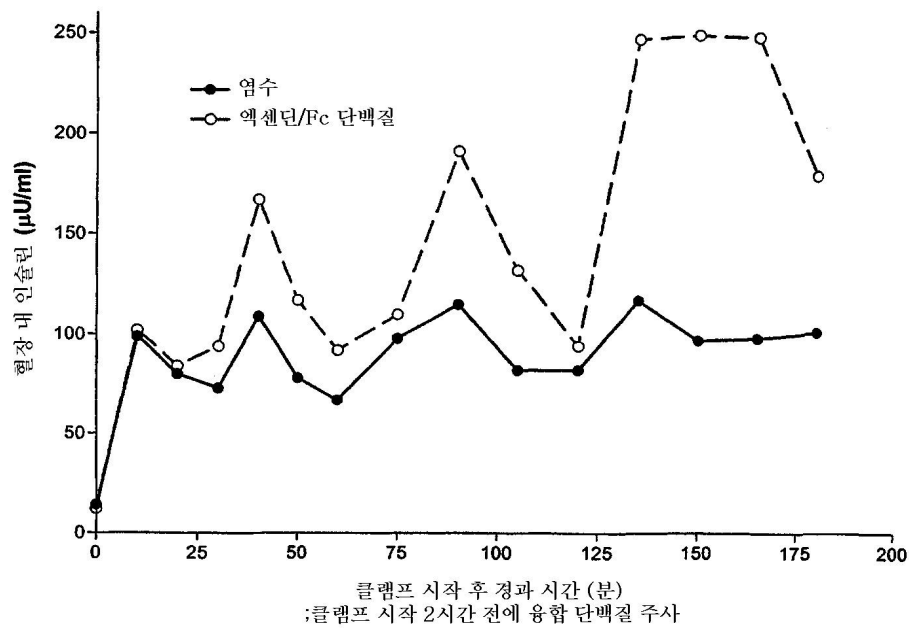
도면10b



도면11a



도면11b



도면12

```

1
GAGCCCAAATCTTGTGACAAACTCACACATGCCACCGTGCCCAGCACC
50
TGAACTCCTGGGGGACCGTCAGTCTTCCTCTTCCCCCAAACCCAAGG
100
ACACCCCTCATGATCTCCCGACCCCTGAGGTACATGCGTGGTGGTGAC
150
GTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGT
200
GGAGGTGCATAATGCCAAGACAAAGCCGGGAGGAGCAGTACAACAGCA
250
CGTACCGTGTGGTCAGCGTCCTCACCCTCTGCACCAGGACTGGCTGAAT
300
GGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGCCCCAT
350
CGAGAAAACCATCTCCAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGT
400
ACACCCTGCCCCCATCCCGGAGGAGATGACCAAGAACCAGGTGAGCCTG
450
ACCTGCCTGGTCAAAGGCTTCTATCCAGCGACATCGCCGTGGAGTGGGA
500
GAGCAATGGGCAGCCGGAGAACAACCTACAAGACCACGCCTCCCGTGTGG
550
ACTCCGACGGCTCCTTCTCTCTATAGCAAGCTCACCGTGGACAAGAGC
600
AGGTGGCAGCAGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCT
650
GCACAACCACTACACGAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAT
700
AGT (서열 33)

```


도면13a

```

1
GATGCGCACAAAGAGTGAGGTTGCTCATCGGTTTAAAGATTGGGAGAAGA
50
AAATTTCAAAGCCTTGGTGTTGATTGCCTTTGCTCAGTATCTTCAGCAGT
100
GTCCATTTGAAGATCATGTAAAATTAGTGAATGAAGTAACTGAATTTGCA
150
AAAACATGTGTGCTGATGAGTCAGCTGAAAATTGTGACAAATCACTTCA
200
TACCCTTTTTGGAGACAAATTATGCACAGTTGCAACTCTTCGTGAAACCT
250
ATGGTGAAATGGCTGACTGCTGTGCAAAACAAGAACCTGAGAGAAATGAA
300
TGCTTCTTGCAACACAAAGATGACAACCCAAACCTCCCCCGATTGGTGAG
350
ACCAGAGGTTGATGTGATGTGCACTGCTTTTCATGACAATGAAGAGACAT
400
TTTTGAAAAATACTTATATGAAATTGCCAGAAGACATCCTTACTTTTAT
450
GCCCGGAACCTCTTTCTTTGCTAAAAGGTATAAAGCTGCTTTTACAGA
500
ATGTTGCCAAGCTGCTGATAAAGCTGCCTGCCTGTTGCCAAGCTCGATG
550
AACTTCGGGATGAAGGAAGGCTTCGTCTGCCAAACAGAGACTCAAGTGT
600
GCCAGTCTCCAAAAATTTGGAGAAAGAGCTTTCAAAGCATGGGCAGTAGC
650
TCGCCTGAGCCAGAGATTTCCCAAAGCTGAGTTTGAGAAGTTTCCAAGT
700
TAGTGACAGATCTTACCAAAGTCCACACGGAATGCTGCCATGGAGATCTG
750
CTTGAATGTGCTGATGACAGGGCGGACCTTGCCAAGTATATCTGTGAAAA
800
TCAAGATTCGATCTCCAGTAACTGAAGGAATGCTGTGAAAAACCTCTGT
850
TGAAAAATCCCACTGCATTGCCGAAGTGAAAAATGATGAGATGCCTGCT
900
GACTTGCCCTTCATTAGCTGCTGATTTTGTGAAAGTAAGGATGTTTGCAA
950
AAACTATGCTGAGGCAAAGGATGTCTTCCTGGGCATGTTTTTGTATGAAT
1000
ATGCAAGAAGGCATCCTGATTACTCTGTCTGCTGCTGCTGAGACTTGCC

```

도면13b

```

1050
AAGACATATGAAACCACTCTAGAGAAGTGCTGTGCCGCTGCAGATCCTCA
1100
TGAATGCTATGCCAAAGTGTTCGATGAATTTAAACCTCTTGTGGAAGAGC
1150
CTCAGAATTTAATCAAACAAAATTGTGAGCTTTTGTGAGCAGCTTGGAGAG
1200
TACAAATTCCAGAATGCGCTATTAGTTCGTTACACCAAGAAAGTACCCCA
1250
AGTGTCAACTCCAACCTCTGTAGAGGTCTCAAGAAACCTAGGAAAAGTGG
1300
GCAGCAAATGTTGTAAACATCCTGAAGCAAAAAGAATGCCCTGTGCAGAA
1350
GACTATCTATCCGTGGTCCTGAACCAGTTATGTGTGTTGCATGAGAAAAC
1400
GCCAGTAAGTGACAGAGTCACCAAATGCTGCACAGAATCCTTGGTGAACA
1450
GGCGACCATGCTTTTCAGCTCTGGAAGTCGATGAAACATACGTTCCCAA
1500
GAGTTTAAATGCTGAAACATTACCTTCCATGCAGATATATGCACACTTTC
1550
TGAGAAGGAGAGACAAATCAAGAAACAACTGCACTTGTGAGCTCGTGA
1600
AACACAAGCCCAAGGCAACAAAAGAGCAACTGAAAGCTGTTATGGATGAT
1650
TTCGCAGCTTTTGTAGAGAAGTGCTGCAAGGCTGACGATAAGGAGACCTG
1700
CTTTGCCGAGGAGGGTAAAAAACTTGTGCTGCAAGTCAAGCTGCCTTAG
1750
GCTTATAATGAC      (서열 35)

```

서열 목록

<110> Eli Lilly and Company

<120> GLP-1 FUSION PROTEINS

<130> X-13991

<150> US 60/251,954

<151> 2000-06-12

<160> 35

<170> PatentIn version 3.1

<210> 1

<211> 31

<212> PRT

<213> Homo sapiens

<400> 1

His	Ala	Glu	Gly	Thr	Phe	Thr	Ser	Asp	Val	Ser	Ser	Tyr	Leu	Glu	Gly
1				5					10					15	

Gln	Ala	Ala	Lys	Glu	Phe	Ile	Ala	Trp	Leu	Val	Lys	Gly	Arg	Gly
			20					25					30	

<210> 2

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp,

or Lys;

<220>

<221> MISC_FEATURE

<222> (3)..(3)

<223> Xaa at position 3 is Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> Xaa at position 5 is Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp,
or Lys;

<220>

<221> MISC_FEATURE

<222> (8)..(8)

<223> Xaa at position 8 is Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp,
or Lys;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu,
, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (11)..(11)

<223> Xaa at position 11 is Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (12)..(12)

<223> Xaa at position 12 is Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Lys, Trp or Tyr;

<220>

<221> MISC_FEATURE

<222> (13)..(13)

<223> Xaa at position 13 is Tyr, Phe, Trp, Glu, Asp, Gln, or Lys;

<220>

<221> MISC_FEATURE

<222> (14)..(14)

<223> Xaa at position 14 is Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Lys, Trp or Tyr;

<220>

<221> MISC_FEATURE

<222> (15)..(15)

<223> Xaa at position 15 is Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, Trp or Lys;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is Gln, Asn, Arg, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (18)..(18)

<223> Xaa at position 18 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> Xaa at position 20 is Lys, Arg, Gln, Glu, Asp, or His;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Leu, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (25)..(25)

<223> Xaa at position 25 is Trp, Phe, Tyr, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (26)..(26)

<223> Xaa at position 26 is Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (27)..(27)

<223> Xaa at position 27 is Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (28)..(28)

<223> Xaa at position 28 is Asn, Lys, Arg, Glu, Asp, or His;

<220>

<221> MISC_FEATURE

<222> (29)..(29)

<223> Xaa at position 29 is Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> Xaa at position 30 is Gly, Arg, Lys, Glu, Asp, or His;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa at position 32 is Ser, Arg, Lys, Glu, Asp, or His, or is dele

ted;

<220>

<221> MISC_FEATURE

<222> (33)..(33)

<223> Xaa at position 33 is Ser, Arg, Lys, Glu, Asp, or His, or is deleted;

<220>

<221> MISC_FEATURE

<222> (34)..(34)

<223> Xaa at position 34 is Gly, Asp, Glu, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

<222> (35)..(35)

<223> Xaa at position 35 is Ala, Phe, Trp, Tyr, Glu, Asp, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

<222> (36)..(36)

<223> Xaa at position 36 is Ser, Pro, Lys, Glu, or Asp, or is deleted;

<220>

<221> MISC_FEATURE

<222> (37)..(37)

<223> Xaa at position 37 is Ser, Pro, Glu, Asp, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

<222> (38)..(38)

<223> Xaa at position 38 is Gly, Pro, Glu, Asp, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

<222> (39)..(39)

<223> Xaa at position 39 is Ala, Ser, Val, Glu, Asp, or Lys, or is deleted;

<400> 2

His	Xaa	Xaa	Gly	Xaa	Phe	Thr	Xaa	Asp	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1			5					10						15	

Xaa Xaa Xaa Xaa Xaa Phe Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

20

25

30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35

<210> 3

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Gly, Ala, Val, Leu, Ile, Ser, or Thr;

<220>

<221> MISC_FEATURE

<222> (3)..(3)

<223> Xaa at position 3 is Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> Xaa at position 5 is Asp, Glu, Arg, Thr, Ala, Lys, or His;

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> Xaa at position 6 is His, Trp, Phe, or Tyr;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu, or Ala;

<220>

<221> MISC_FEATURE

<222> (12)..(12)

<223> Xaa at position 12 is His, Pro, Asp, Glu, Arg, Ser, Ala, or Lys;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 13 is Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is His, Asp, Lys, Glu, Gln, or Arg;

<220>

<221> MISC_FEATURE

<222> (18)..(18)

<223> Xaa at position 18 is Glu, Arg, Ala, or Lys;

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> Xaa at position 20 is Trp, Tyr, Phe, Asp, Lys, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala, Glu, Asp, Ser, or His;

<220>

<221> MISC_FEATURE

<222> (25)..(25)

<223> Xaa at position 25 is Asp, Glu, Ser, Thr, Arg, Trp, or Lys;

<220>

<221> MISC_FEATURE

<222> (27)..(27)

<223> Xaa at position 27 is Asp, Arg, Val, Lys, Ala, Gly, or Glu;

<220>

<221> MISC_FEATURE

<222> (28)..(28)

<223> Xaa at position 28 is Glu, Lys, or Asp;

<220>

<221> MISC_FEATURE

<222> (29)..(29)

<223> Xaa at position 29 is Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> Xaa at position 30 is Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe

, His, Gly, or is deleted.

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa at position 31 is Pro or is deleted.

<400> 3

Xaa	Xaa	Xaa	Gly	Xaa	Xaa	Thr	Ser	Asp	Xaa	Ser	Xaa	Tyr	Leu	Glu	Xaa
1				5					10					15	

Xaa	Xaa	Ala	Xaa	Xaa	Phe	Ile	Xaa	Xaa	Leu	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			20						25					30	

<210> 4

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Gly, Ala, Val, Leu, Ile, Ser, or Thr;

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> Xaa at position 5 is Asp, Glu, Arg, Thr, Ala, Lys, or His;

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> Xaa at position 6 is His, Trp, Phe, or Tyr;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, or Ala;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is His, Asp, Lys, Glu, or Gln;

<220>

<221> MISC_FEATURE

<222> (18)..(18)

<223> Xaa at position 18 is Glu, His, Ala, or Lys;

<220>

<221> MISC_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is Asp, Lys, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala, Glu, Asp, Ser, or His;

<220>

<221> MISC_FEATURE

<222> (27)..(27)

<223> Xaa at position 27 is Asp, Arg, Val, Lys, Ala, Gly, or Glu;

<220>

<221> MISC_FEATURE

<222> (28)..(28)

<223> Xaa at position 28 is Glu, Lys, or Asp;

<220>

<221> MISC_FEATURE

<222> (29)..(29)

<223> Xaa at position 29 is Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> Xaa at position 30 is Arg, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, or is deleted.

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa at position 32 is Pro, or is deleted.

<400> 4

Xaa Xaa Glu Gly Xaa Xaa Thr Ser Asp Xaa Ser Ser Tyr Leu Glu Xaa

1 5 10 15

Xaa Xaa Xaa Lys Xaa Phe Ile Xaa Trp Leu Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

<210> 5

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Gly, Ala, Val, Leu, Ile, Ser, Met, or Thr;

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> Xaa at position 6 is His, Trp, Phe, or Tyr;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, or Ala;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is His, Asp, Lys, Glu, or Gln;

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> Xaa at position 20 is Asp, Lys, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala, Glu, Asp, Ser, or His;

<220>

<221> MISC_FEATURE

<222> (29)..(29)

<223> Xaa at position 29 is Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, or is deleted.

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa at position 32 is Pro or is deleted.

<400> 5

Xaa	Xaa	Glu	Gly	Thr	Xaa	Thr	Ser	Asp	Xaa	Ser	Ser	Tyr	Leu	Glu	Xaa
1				5				10					15		

Xaa	Ala	Ala	Xaa	Glu	Phe	Ile	Xaa	Trp	Leu	Val	Lys	Xaa	Arg	Xaa	Xaa
			20					25					30		

<210> 6

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Gly, Ala, Val, Leu, Ile, Ser, or Thr;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is His, Asp, Lys, Glu, or Gln;

<220>

<221> MISC_FEATURE

<222> (18)..(18)

<223> Xaa at position 18 is Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala, Glu, Asp, Ser, or His;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa at position 32 is Pro or is deleted.

<400> 6

Xaa	Xaa	Glu	Gly	Thr	Phe	Thr	Ser	Asp	Val	Ser	Ser	Tyr	Leu	Glu	Xaa
1				5					10					15	

Xaa	Xaa	Ala	Lys	Glu	Phe	Ile	Xaa	Trp	Leu	Val	Lys	Gly	Arg	Xaa	Xaa
			20					25					30		

<210> 7

<211> 31

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Ala, Gly, Val, Thr, Ile, and alpha-methyl-Ala;

<220>

<221> MISC_FEATURE

<222> (15)..(15)

<223> Xaa at position 15 is Glu, Gln, Ala, Thr, Ser, and Gly;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Glu, Gln, Ala, Thr, Ser, and Gly;

<400> 7

Xaa Xaa Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Xaa Gly

1	5	10	15
Gln	Ala	Ala	Lys
	Xaa	Phe	Ile
	Ala	Trp	Leu
	Val	Lys	Gly
	Arg	Gly	
	20	25	30

<210> 8

<211> 30

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is Lys or Arg;

<220>

<221> MOD_RES

<222> (27)..(27)

<223> ACETYLTATION

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> Xaa at position 30 is Gly;

<220>

<221> MOD_RES

<222> (30)..(30)

<223> AMIDATION

<400> 8

Ala	Glu	Gly	Thr	Phe	Thr	Ser	Asp	Val	Ser	Ser	Tyr	Leu	Glu	Gly	Gln
1				5				10					15		

Ala	Ala	Xaa	Glu	Phe	Ile	Ala	Trp	Leu	Val	Lys	Gly	Arg	Xaa
		20					25					30	

<210> 9

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 9

His	Ser	Asp	Gly	Thr	Phe	Thr	Ser	Asp	Leu	Ser	Lys	Gln	Met	Glu	Glu
1				5					10				15		

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser
35

<210> 10

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 10

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser
35

<210> 11

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Gly, Ala, or Val;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Leu or Val;

<220>

<221> MISC_FEATURE

<222> (12)..(12)

<223> Xaa at position 12 is Lys or Ser;

<220>

<221> MISC_FEATURE

<222> (13)..(13)

<223> Xaa at position 13 is Gln or Tyr;

<220>

<221> MISC_FEATURE

<222> (14)..(14)

<223> Xaa at position 14 is Met or Leu;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Glu or Gln;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is Glu or Gln;

<220>

<221> MISC_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is Val or Ala;

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> Xaa at position 20 is Arg or Lys;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Leu or Glu;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Glu or Ala;

<220>

<221> MISC_FEATURE

<222> (27)..(27)

<223> Xaa at position 27 is Val or Lys;

<220>

<221> MISC_FEATURE

<222> (28)..(28)

<223> Xaa at position 28 is Asn or Lys;

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> Xaa at position 30 is Gly or Arg; and

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Gly, or Pro;

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa at position 32 is Ser, or is absent.

<220>

<221> MISC_FEATURE

<222> (33)..(33)

<223> Xaa at position 33 is Ser, or is absent.

<220>

<221> MISC_FEATURE

<222> (34)..(34)

<223> Xaa at position 34 is Gly, or is absent.

<220>

<221> MISC_FEATURE

<222> (35)..(35)

<223> Xaa at position 35 is Ala, or is absent.

<220>

<221> MISC_FEATURE

<222> (36)..(36)

<223> Xaa at position 36 is Pro, or is absent.

<220>

<221> MISC_FEATURE

<222> (37)..(37)

<223> Xaa at position 37 is Pro, or is absent.

<220>

<221> MISC_FEATURE

<222> (38)..(38)

<223> Xaa at position 38 is Pro, or is absent.

<220>

<221> MISC_FEATURE

<222> (39)..(39)

<223> Xaa at position 39 is Pro, or is absent.

<220>

<221> MISC_FEATURE

<222> (39)..(39)

<223> Xaa at position 39 is Ser, or is absent.

<400> 11

Xaa	Xaa	Glu	Glu	Thr	Phe	Thr	Ser	Asp	Xaa	Ser	Xaa	Xaa	Xaa	Glu	Xaa
1				5					10					15	

Xaa	Ala	Xaa	Xaa	Xaa	Phe	Ile	Xaa	Trp	Leu	Xaa	Xaa	Gly	Xaa	Xaa	Xaa
		20						25					30		

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
						35

<210> 12

<211> 31

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Ala, Gly, Val, Leu, Ile, Ser, or Thr;

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> Xaa at position 6 is Phe, Trp, or Tyr;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Val, Trp, Ile, Leu, Phe, or Tyr;

<220>

<221> MISC_FEATURE

<222> (12)..(12)

<223> Xaa at position 12 is Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;

<220>

<221> MISC_FEATURE

<222> (13)..(13)

<223> Xaa at position 13 is Tyr, Trp, or Phe;

<220>

<221> MISC_FEATURE

<222> (14)..(14)

<223> Xaa at position 14 is Leu, Phe, Tyr, or Trp;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Gly, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is Ala, Val, Ile, or Leu;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Glu, Ile, or Ala;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala or Glu;

<220>

<221> MISC_FEATURE

<222> (27)..(27)

<223> Xaa at position 27 is Val or Ile; and

<220>

<221> MOD_RES

<222> (30)..(30)

<223> AMIDATION

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Gly, His or is absent.

<400> 12

Xaa	Xaa	Glu	Gly	Thr	Xaa	Thr	Ser	Asp	Xaa	Ser	Xaa	Xaa	Xaa	Glu	Xaa
1				5					10					15	

Gln	Ala	Xaa	Lys	Xaa	Phe	Ile	Xaa	Trp	Leu	Xaa	Lys	Gly	Arg	Xaa
			20					25					30	

<210> 13

<211> 616

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 13

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Asp
20 25 30

Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu
35 40 45

Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln
50 55 60

Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe
65 70 75 80

Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser
85 90 95

Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg
100 105 110

Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu
115 120 125

Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro
130 135 140

Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp
145 150 155 160

Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg
165 170 175

His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr
180 185 190

Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys
195 200 205

Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser
210 215 220

Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg
225 230 235 240

Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys
245 250 255

Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val
260 265 270

His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg
275 280 285

Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser
290 295 300

Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys
305 310 315 320

Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu
325 330 335

Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu
340 345 350

Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg
355 360 365

His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr
370 375 380

Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys
385 390 395 400

Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln
405 410 415

Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr
420 425 430

Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln
435 440 445

Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val
450 455 460

Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala
465 470 475 480

Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu
485 490 495

Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu
500 505 510

Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr
515 520 525

Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile
530 535 540

Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu
545 550 555 560

Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys
565 570 575

Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala
580 585 590

Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala
595 600 605

Ala Ser Gln Ala Ala Leu Gly Leu
610 615

<210> 14

<211> 631

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 14

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Gly
20 25 30

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ala
35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn

50

55

60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Cys
65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala
85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu
100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu
115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg
130 135 140

Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg
145 150 155 160

Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn
165 170 175

Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His
180 185 190

Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys
195 200 205

Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu
210 215 220

Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala
225 230 235 240

Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala

245

250

255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala
260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His
275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala
290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys
305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile
325 330 335

Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala
340 345 350

Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala
355 360 365

Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His
370 375 380

Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu
385 390 395 400

Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr
405 410 415

Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn
420 425 430

Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys

435 440 445

 Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val
 450 455 460

 Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly
 465 470 475 480

 Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu
 485 490 495

 Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys
 500 505 510

 Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val
 515 520 525

 Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val
 530 535 540

 Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys
 545 550 555 560

 Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val
 565 570 575

 Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala
 580 585 590

 Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp
 595 600 605

 Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala
 610 615 620

 Ser Gln Ala Ala Leu Gly Leu

625 630

<210> 15

<211> 640

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 15

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly
35 40 45

Gly Ser Gly Gly Gly Gly Ser Asp Ala His Lys Ser Glu Val Ala His
50 55 60

Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile
65 70 75 80

Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys
85 90 95

Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu
100 105 110

Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys
115 120 125

Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp
130 135 140

Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His
145 150 155 160

Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp
165 170 175

Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys
180 185 190

Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu
195 200 205

Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys
210 215 220

Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu
225 230 235 240

Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala
245 250 255

Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala
260 265 270

Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys
275 280 285

Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp
290 295 300

Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys
305 310 315 320

Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
325 330 335

Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
340 345 350

Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
355 360 365

Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met
370 375 380

Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu
385 390 395 400

Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys
405 410 415

Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe
420 425 430

Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu
435 440 445

Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val
450 455 460

Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu
465 470 475 480

Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro
485 490 495

Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu
500 505 510

Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val
515 520 525

Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser
530 535 540

Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu
545 550 555 560

Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg
565 570 575

Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro
580 585 590

Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala
595 600 605

Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala
610 615 620

Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
625 630 635 640

<210> 16

<211> 624

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 16

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Asp Ala His Lys Ser Glu Val Ala His
35 40 45

Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile
50 55 60

Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys
65 70 75 80

Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu
85 90 95

Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys
100 105 110

Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp
115 120 125

Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His
130 135 140

Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp
145 150 155 160

Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys

165

170

175

Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu
180 185 190

Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys
195 200 205

Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu
210 215 220

Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala
225 230 235 240

Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala
245 250 255

Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys
260 265 270

Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp
275 280 285

Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys
290 295 300

Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
305 310 315 320

Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
325 330 335

Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
340 345 350

Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met

355 Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu
 370 375 380
 Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys
 385 390 395 400
 Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe
 405 410 415
 Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu
 420 425 430
 Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val
 435 440 445
 Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu
 450 455 460
 Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro
 465 470 475 480
 Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu
 485 490 495
 Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val
 500 505 510
 Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser
 515 520 525
 Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu
 530 535 540
 Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg

545 550 555 560

Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro
565 570 575

Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala
580 585 590

Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala
595 600 605

Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
610 615 620

<210> 17

<211> 640

<212> PRT

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> synthetic construct

<400> 17

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly

35 40 45

Gly Ser Gly Gly Gly Gly Ser Asp Ala His Lys Ser Glu Val Ala His
50 55 60

Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile
65 70 75 80

Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys
85 90 95

Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu
100 105 110

Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys
115 120 125

Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp
130 135 140

Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His
145 150 155 160

Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp
165 170 175

Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys
180 185 190

Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu
195 200 205

Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys
210 215 220

Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu
225 230 235 240

Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala
245 250 255

Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala
260 265 270

Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys
275 280 285

Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp
290 295 300

Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys
305 310 315 320

Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
325 330 335

Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
340 345 350

Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
355 360 365

Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met
370 375 380

Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu
385 390 395 400

Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys
405 410 415

Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe
420 425 430

Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu
435 440 445

Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val
450 455 460

Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu
465 470 475 480

Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro
485 490 495

Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu
500 505 510

Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val
515 520 525

Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser
530 535 540

Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu
545 550 555 560

Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg
565 570 575

Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro
580 585 590

Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala
595 600 605

Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala
610 615 620

Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
625 630 635 640

<210> 18

<211> 264

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 18

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ala
20 25 30

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
35 40 45

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
50 55 60

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
65 70 75 80

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
85 90 95

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln

100	105	110
Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln		
115	120	125
Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala		
130	135	140
Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro		
145	150	155
Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr		
165	170	175
Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser		
180	185	190
Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr		
195	200	205
Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr		
210	215	220
Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe		
225	230	235
Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys		
245	250	255
Ser Leu Ser Leu Ser Pro Gly Lys		
260		

<210> 19

<211> 272

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 19

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 20

<211> 264

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 20

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ala
20 25 30

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
35 40 45

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
50 55 60

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
65 70 75 80

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
85 90 95

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
100 105 110

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
115 120 125

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
130 135 140

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
145 150 155 160

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
165 170 175

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
180 185 190

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
195 200 205

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
210 215 220

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
225 230 235 240

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
245 250 255

Ser Leu Ser Leu Ser Pro Gly Lys
260

<210> 21

<211> 272

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 21

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp

210

215

220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 22

<211> 272

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 22

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Ser Ser Gly Ala Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 23

<211> 287

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 23

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
50 55 60

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
85 90 95

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
100 105 110

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
115 120 125

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
130 135 140

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
145 150 155 160

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
165 170 175

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
180 185 190

Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
195 200 205

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
210 215 220

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
225 230 235 240

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
245 250 255

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
260 265 270

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
275 280 285

<210> 24

<211> 284

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 24

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Ser Lys Tyr Gly Pro Pro Cys Pro
50 55 60

Ser Cys Pro Ala Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe
65 70 75 80

Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val
85 90 95

Thr Cys Val Val Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe
100 105 110

Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro

115					120					125						
Arg	Glu	Glu	Gln	Phe	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val	Leu	Thr	
130					135					140						
Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys	Cys	Lys	Val	
145					150					155					160	
Ser	Asn	Lys	Gly	Leu	Pro	Ser	Ser	Ile	Glu	Lys	Thr	Ile	Ser	Lys	Ala	
165					170					175						
Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu	Pro	Pro	Ser	Gln	
180					185					190						
Glu	Glu	Met	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys	Leu	Val	Lys	Gly	
195					200					205						
Phe	Tyr	Pro	Ser	Asp	Ile	Ala	Val	Glu	Trp	Glu	Ser	Asn	Gly	Gln	Pro	
210					215					220						
Glu	Asn	Asn	Tyr	Lys	Thr	Thr	Pro	Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	
225					230					235					240	
Phe	Phe	Leu	Tyr	Ser	Arg	Leu	Thr	Val	Asp	Lys	Ser	Arg	Trp	Gln	Glu	
245					250					255						
Gly	Asn	Val	Phe	Ser	Cys	Ser	Val	Met	His	Glu	Ala	Leu	His	Asn	His	
260					265					270						
Tyr	Thr	Gln	Lys	Ser	Leu	Ser	Leu	Ser	Leu	Gly	Lys					
275					280											

<210> 25

<211> 302

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 25

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
50 55 60

Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr
65 70 75 80

Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
85 90 95

Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro
100 105 110

Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val
115 120 125

Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr
130 135 140

Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val
145 150 155 160

Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys
165 170 175

Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser
180 185 190

Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro
195 200 205

Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val
210 215 220

Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly
225 230 235 240

Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp
245 250 255

Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp
260 265 270

Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His
275 280 285

Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
290 295 300

<210> 26

<211> 294

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 26

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Gly
20 25 30

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
35 40 45

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Glu Pro
50 55 60

Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu
65 70 75 80

Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
85 90 95

Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
100 105 110

Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly
115 120 125

Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn
130 135 140

Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp
145 150 155 160

Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro
165 170 175

Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu
180 185 190

Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn
195 200 205

Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
210 215 220

Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
225 230 235 240

Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys
245 250 255

Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
260 265 270

Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
275 280 285

Ser Leu Ser Pro Gly Lys
290

<210> 27

<211> 280

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 27

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ser Ser Gly Ala Pro Pro Pro Ser Ala
35 40 45

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
50 55 60

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
65 70 75 80

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
85 90 95

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
100 105 110

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
115 120 125

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
130 135 140

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala

145 150 155 160

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
165 170 175

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
180 185 190

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
195 200 205

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
210 215 220

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
225 230 235 240

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
245 250 255

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
260 265 270

Ser Leu Ser Leu Ser Pro Gly Lys
275 280

<210> 28

<211> 287

<212> PRT

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> synthetic construct

<400> 28

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Val Lys Glu Phe Ile Ala Trp Leu Ile Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
50 55 60

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
85 90 95

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
100 105 110

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
115 120 125

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
130 135 140

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
145 150 155 160

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
165 170 175

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
180 185 190

Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
195 200 205

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
210 215 220

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
225 230 235 240

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
245 250 255

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
260 265 270

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
275 280 285

<210> 29

<211> 272

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 29

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser

195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 30

<211> 272

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 30

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Ser Ser Gly Ala Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 31

<211> 287

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 31

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
50 55 60

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
85 90 95

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
100 105 110

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
115 120 125

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
130 135 140

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
145 150 155 160

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
165 170 175

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
180 185 190

Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
195 200 205

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
210 215 220

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
225 230 235 240

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
245 250 255

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
260 265 270

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys

275

280

285

<210> 32

<211> 232

<212> PRT

<213> Homo sapiens

<400> 32

Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10 15

Ala Pro Glu Lys Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
20 25 30

Lys Asp Thr Lys Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
35 40 45

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
50 55 60

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
65 70 75 80

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
85 90 95

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
100 105 110

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
115 120 125

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
130 135 140

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
145 150 155 160

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
165 170 175

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
180 185 190

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
195 200 205

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
210 215 220

Ser Leu Ser Leu Ser Pro Gly Lys
225 230

<210> 33

<211> 703

<212> DNA

<213> Homo sapiens

<400> 33
gagcccaaat cttgtgacaa aactcacaca tgcccaccgt gcccagcacc tgaactcctg 60

gggggaccgt cagtcttcct cttccccca aaaccaagg acaccctcat gatctcccg 120

acccctgagg tcacatgcgt ggtggtggac gtgagccacg aagaccctga ggtcaagttc 180
 aactggtacg tggacggcgt ggaggtgcat aatgccaaga caaagccgcg ggaggagcag 240
 tacaacagca cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat 300
 ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc cagcccccat cgagaaaacc 360
 atctccaaag ccaaagggca gccccgagaa ccacaggtgt acaccctgcc cccatcccgg 420
 gaggagatga ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatcccagc 480
 gacatcgccg tggagtggga gagcaatggg cagccggaga acaactaaa gaccacgcct 540
 cccgtgctgg actccgacgg ctctctcttc ctctatagca agctcacctg ggacaagagc 600
 aggtggcagc aggggaacgt cttctcatgc tccgtgatgc atgaggtctt gcacaaccac 660
 tacacgcaga agagcctctc cctgtctccg ggtaaatgat agt 703

<210> 34

<211> 585

<212> PRT

<213> Homo sapiens

<400> 34

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
 1 5 10 15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
 20 25 30

Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu

35

40

45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
50 55 60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65 70 75 80

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
85 90 95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
100 105 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
115 120 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
130 135 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145 150 155 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
165 170 175

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
180 185 190

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
195 200 205

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
210 215 220

Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys

225 230 235 240
 Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
 245 250 255
 Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
 260 265 270
 Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
 275 280 285
 Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
 290 295 300
 Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
 305 310 315 320
 Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
 325 330 335
 Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
 340 345 350
 Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
 355 360 365
 Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
 370 375 380
 Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Asn Leu Gly Glu
 385 390 395 400
 Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
 405 410 415
 Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
 420 425 430

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
435 440 445

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
450 455 460

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
465 470 475 480

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
485 490 495

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
500 505 510

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
515 520 525

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
530 535 540

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
545 550 555 560

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
565 570 575

Ala Ala Ser Gln Ala Ala Leu Gly Leu
580 585

<210> 35

<211> 1762

<212> DNA

<213> Homo sapiens

<400> 35

gatgcgaca agagtgaggt tgctcatcgg tttaaagatt tgggagaaga aaatttcaaa	60
gccttgggtgt tgattgcctt tgctcagtat cttcagcagt gtccatttga agatcatgta	120
aaattagtga atgaagtaac tgaatttgca aaaacatgtg ttgctgatga gtcagctgaa	180
aatttgaca aatcacttca tacccttttt ggagacaaat tatgcacagt tgcaactctt	240
cgtgaaacct atggtgaaat ggctgactgc tgtgcaaac aagaacctga gagaaatgaa	300
tgcttcttgc aacacaaaga tgacaacca aacctcccc gattgggtgag accagaggtt	360
gatgtgatgt gcactgcttt tcatgacaat gaagagacat ttttgaataa atacttatat	420
gaaattgcca gaagacatcc ttacttttat gccccgaac tccttttctt tgctaaaagg	480
tataaagctg cttttacaga atgttgccaa gctgctgata aagctgcctg cctgttgcca	540
aagctcgatg aacttcggga tgaagggaag gcttcgtctg ccaaacagag actcaagtgt	600
gccagtctcc aaaaatttgg agaaagagct ttcaaagcat gggcagtagc tcgcctgagc	660
cagagatttc ccaaagctga gtttcagaa gtttccaagt tagtgacaga tcttaccaaa	720
gtccacacgg aatgctgcca tggagatctg cttgaatgtg ctgatgacag ggcggacctt	780
gccaaagtata tctgtgaaaa tcaagattcg atctccagta aactgaagga atgctgtgaa	840
aaacctctgt tggaataatc ccactgcatt gccgaagtgg aaaatgatga gatgcctgct	900
gacttgcctt cattagctgc tgattttgtt gaaagtaagg atgtttgcaa aaactatgct	960
gaggcaaagg atgtcttctt gggcatgttt ttgtatgaat atgcaagaag gcatcctgat	1020
tactctgtcg tgctgtgct gagacttgcc aagacatatg aaaccactct agagaagtgc	1080

tgtgccgtg cagatcctca tgaatgctat gccaaagtgt tcgatgaatt taaacctctt	1140
gtggaagagc ctcagaattt aatcaaacaa aattgtgagc tttttgagca gcttggagag	1200
tacaaattcc agaatgcgct attagtctgt tacaccaaga aagiacccca agtgtcaact	1260
ccaactcttg tagaggcttc aagaaacctt ggaaaagtgg gcagcaaatt ttgtaaacat	1320
cctgaagcaa aaagaatgcc ctgtgcagaa gactatctat ccgtggctct gaaccagtta	1380
tgtgtgttg atgagaaaac gccagtaagt gacagagtca ccaaatgctg cacagaatcc	1440
ttggtgaaca ggcgaccatg cttttcagct ctggaagtcg atgaaacata cgttcccaaa	1500
gagtttaatg ctgaaacatt caccttccat gcagatatat gcacactttc tgagaaggag	1560
agacaaatca agaaacaaac tgcacttgtt gagctcgtga aacacaagcc caaggcaaca	1620
aaagagcaac tgaagctgt tatggatgat ttgcagctt ttgtagagaa gtgctgcaag	1680
gctgacgata aggagacctg ctttgccgag gagggtaaaa aacttggtgc tgcaagtcaa	1740
gctgccttag gcttataatg ac	1762