



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0033906
(43) 공개일자 2009년04월06일

(51) Int. Cl.

C07K 19/00 (2006.01) C12N 15/62 (2006.01)
C12N 15/63 (2006.01) A61K 38/22 (2006.01)

- (21) 출원번호 10-2009-7003679
- (22) 출원일자 2009년02월23일
심사청구일자 2009년02월23일
번역문제출일자 2009년02월23일
- (86) 국제출원번호 PCT/IB2007/002047
국제출원일자 2007년07월13일
- (87) 국제공개번호 WO 2008/012629
국제공개일자 2008년01월31일
- (30) 우선권주장
60/832,582 2006년07월24일 미국(US)
(뒷면에 계속)

(71) 출원인

바이오렉시스 파마슈티칼 코포레이션

미국 뉴욕주 10017 뉴욕 이스트 42번 스트리트 235

(72) 발명자

사텍히 호마윤

미국 뉴욕주 10017 뉴욕 이스트 42번 스트리트 235

프라이어 크리스토퍼 필립

미국 뉴욕주 10017 뉴욕 이스트 42번 스트리트 235

(뒷면에 계속)

(74) 대리인

김창세, 장성구

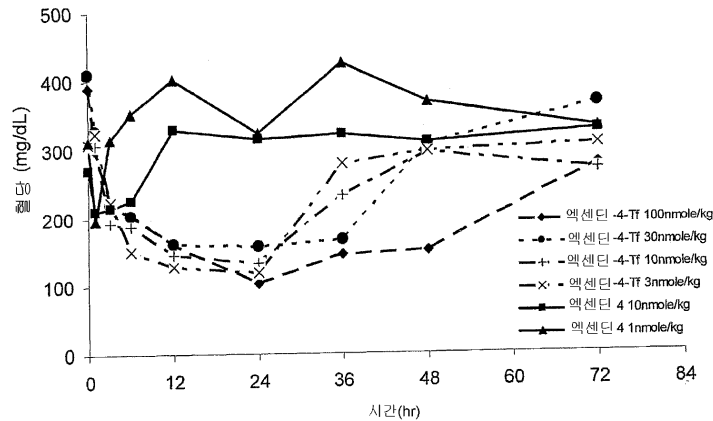
전체 청구항 수 : 총 31 항

(54) 엑센딘 융합 단백질

(57) 요약

본 발명은 폴리펩티드 링커를 통해 트랜스페린(Tf)에 융합된 엑센딘-4(exendin-4)를 포함하는 융합 단백질, 상응하는 핵산 분자, 벡터, 숙주 세포 및 약학적 조성물을 제공한다. 또한, 본 발명은 II형 당뇨병 및 비만의 치료 및 체중 감소를 위한 엑센딘-4/Tf 융합 단백질의 용도를 제공한다.

대표도



(72) 발명자

발랑스 데이비드 제임스

미국 뉴욕주 10017 뉴욕 이스트 42번 스트리트 235

터너 앤드류 존

미국 뉴욕주 10017 뉴욕 이스트 42번 스트리트 235

(30) 우선권주장

60/857,474 2006년11월08일 미국(US)

60/874,965 2006년12월15일 미국(US)

특허청구의 범위

청구항 1

폴리펩티드 링커를 통해 트랜스페린(Tf)에 융합된 엑센딘-4(exendin-4)를 포함하는 융합 단백질.

청구항 2

제1항에 있어서,
링커가 비나선형 폴리펩티드인, 융합 단백질.

청구항 3

제1항에 있어서,
링커가 PEAPTD(서열번호 6), (PEAPTD)₂(서열번호 5), IgG 힌지 링커와 조합된 PEAPTD(서열번호 6) 및 IgG 힌지 링커와 조합된 (PEAPTD)₂(서열번호 5)으로 구성된 군으로부터 선택된, 융합 단백질.

청구항 4

제1항에 있어서,
엑센딘-4가 서열번호 4에 나타낸 아미노산 서열을 갖는, 융합 단백질.

청구항 5

제1항에 있어서,
Tf가 천연 Tf 분자에 비해 감소된 글리코실화를 나타내도록 변경된, 융합 단백질.

청구항 6

제5항에 있어서,
Tf가 서열번호 17에 나타낸 아미노산 서열을 갖는, 융합 단백질.

청구항 7

제1항에 있어서,
엑센딘-4 분자가 융합 단백질의 N-말단, C-말단 또는 N-말단 및 C-말단 둘다에서 융합되어 있는, 융합 단백질.

청구항 8

제1항에 있어서,
융합 단백질의 N-말단이 분비 신호 서열을 추가로 포함하는, 융합 단백질.

청구항 9

제8항에 있어서,
신호 서열이 혈청 Tf, 락토페린, 멜라노트랜스페린 또는 이들의 변이체로부터의 신호 서열인, 융합 단백질.

청구항 10

제8항에 있어서,
신호 서열이 인간 혈청 알부민(HSA)/MF α-1 하이브리드 리더 서열, 변경된 HSA/MF α-1 하이브리드 리더 서열 또는 Tf 신호 서열인, 융합 단백질.

청구항 11

제8항에 있어서,

신호 서열이 서열번호 18에 나타난 인간 Tf 신호 서열인, 융합 단백질.

청구항 12

제1항 내지 제11항 중 어느 한 항의 융합 단백질을 코딩하는 핵산.

청구항 13

제12항의 핵산을 포함하는 벡터.

청구항 14

제13항의 벡터를 포함하는 숙주 세포.

청구항 15

변경된 Tf(mTf)에 융합된 엑센딘-4를 포함하는 융합 단백질로서, 서열번호 23에 나타난 아미노산 서열을 포함하는 융합 단백질.

청구항 16

mTf에 융합된 엑센딘-4를 포함하는 융합 단백질로서, 서열번호 25에 나타난 아미노산 서열을 포함하는 융합 단백질.

청구항 17

서열번호 23 또는 25의 융합 단백질을 코딩하는 핵산.

청구항 18

제17항에 있어서,

서열번호 24 또는 26에 나타난 서열을 포함하는 핵산.

청구항 19

제17항 또는 제18항의 핵산을 포함하는 벡터.

청구항 20

제19항의 벡터를 포함하는 숙주 세포.

청구항 21

제1항 내지 제11항 중 어느 한 항의 융합 단백질 및 약학적으로 허용가능한 담체를 포함하는 약학 조성물.

청구항 22

제15항의 융합 단백질 및 약학적으로 허용가능한 담체를 포함하는 약학 조성물.

청구항 23

제22항에 있어서,

0.5 내지 50 mg의 투여량으로 투여되도록 맞춰진 조성물.

청구항 24

제22항에 있어서,

1 내지 100 mg의 투여량으로 투여되도록 맞춰진 조성물.

청구항 25

제21항 또는 제22항에 있어서,
흡입을 통해 투여되도록 맞춰진 조성물.

청구항 26

폴리펩티드 링커를 통해 Tf에 융합된 엑센딘-4를 포함하는 융합 단백질의 치료 유효량을 II형 당뇨병의 치료 또는 혈당 감소가 필요한 인간 환자에게 투여하는 단계를 포함하는, 상기 인간 환자에서 II형 당뇨병을 치료하거나 혈당을 감소시키는 방법.

청구항 27

제26항에 있어서,
융합 단백질이 서열번호 23에 나타낸 아미노산 서열을 포함하는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 포함하는, 방법.

청구항 28

제27항에 있어서,
서열번호 23에 나타낸 아미노산 서열을 포함하는 융합 단백질을 주당 약 1회, 2주당 약 1회 또는 월당 약 1회의 빈도로 약 0.5 내지 약 50 mg의 투여량으로 투여하는, 방법.

청구항 29

폴리펩티드 링커를 통해 Tf에 융합된 엑센딘-4를 포함하는 융합 단백질의 치료 유효량을 비만의 치료 또는 체중 감소가 필요한 인간 환자에게 투여하는 단계를 포함하는, 상기 인간 환자에서 비만을 치료하거나 체중을 감소시키는 방법.

청구항 30

제29항에 있어서,
융합 단백질이 서열번호 23에 나타낸 아미노산 서열을 포함하는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 포함하는, 방법.

청구항 31

제30항에 있어서,
서열번호 23에 나타낸 아미노산 서열을 포함하는 융합 단백질을 주당 약 1회, 2주당 약 1회 또는 월당 약 1회의 빈도로 약 1 내지 약 100 mg의 투여량으로 투여하는, 방법.

명세서

기술분야

<1> 본 발명은 엑세딘-4 및 트랜스페린을 포함하는 융합 단백질, 및 상승된 당 혈청 농도와 관련된 질환, 예컨대, II형 당뇨병의 치료 및 체중 감소를 위한 상기 융합 단백질의 용도에 관한 것이다. 본 발명의 융합 단백질은 엑세딘-4 및 다른 GLP-1 수용체 작용제를 사용한 치료로부터 효과를 얻을 수 있는 것으로 공지된 다른 질환, 예컨대, I형 당뇨병, 울혈성 심부전, 심근경색, 자극성 장 증후군, 신경 질환, 예컨대, 알츠하이머병 및 헌팅톤병, 및 비-알코올성 비-지방 간 질환의 치료에도 사용될 수 있다.

배경기술

<2> 당뇨병은 다수의 병인 인자로부터 유래된 질환으로서 금식 상태에서 또는 경구 당 내성 시험 동안 당의 투여 후 상승된 혈장 당 농도 또는 고혈당증을 특징으로 하는 질환을 지칭한다. 일반적으로 2가지 유형의 당뇨병이 인

식되어 있다. I형 당뇨병 또는 인슐린 의존성 진성 당뇨병(IDDM)에서, 환자는 당 이용을 조절하는 호르몬인 인슐린을 거의 또는 전혀 생산하지 못한다. II형 당뇨병 또는 비-인슐린 의존성 진성 당뇨병(NIDDM)에서, 환자는 종종 비당뇨병 환자와 동일한 또는 심지어 비당뇨병 환자에 비해 상승된 혈장 인슐린 농도를 보인다. 그러나, 상기 환자는 주요 인슐린-감수성 조직, 즉 근육, 간 및 지방 조직에서 당 및 지질 대사에 대한 인슐린 자극 효과에 대한 내성을 발달시킨다. 혈장 인슐린 농도는 상승되지만 현저한 인슐린 내성을 극복하기에 불충분하여 고혈당증을 야기한다.

- <3> 지속되는 또는 조절되지 않은 고혈당증은 증가한 조기 이환률 및 사망률과 관련되어 있다. 종종 비정상적인 당 항상성은 지질, 지단백질 및 아포지단백질 대사 및 다른 대사적 혈액학적 질환의 변화와 직접적으로 및 간접적으로 관련되어 있다. 예를 들어, II형 진성 당뇨병을 앓는 환자는 특히 관상 심장 질환, 졸중, 말초 혈관 질환, 고혈압, 신장병증 및 신경병증을 비롯한 거대혈관 및 미세혈관 합병증을 앓을 위험이 높다.
- <4> 일반적으로, 비만 및 과체중은 체지방과 상관관계를 갖고 있으며 여러 질환의 위험 척도로서 기능하는 신체 질량 지수(BMI)에 의해 정의된다. BMI는 kg 단위의 체중을 m² 단위의 신장으로 나눔으로써 계산한다(kg/m²). 과체중은 전형적으로 25 내지 29.9 kg/m²의 BMI로서 정의되고, 비만은 30 kg/m² 이상의 BMI로서 정의된다. 예를 들어, 문헌[National Heart, Lung, and Blood Institute, Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults, The Evidence Report, Washington, DC: U.S. Department of Health and Human Services, NIH publication no. 98-4083 (1998)]을 참조한다.
- <5> 과체중이거나 비만인 개체는 고혈압, 이상지혈증, II형(비-인슐린 의존성) 당뇨병, 인슐린 내성, 당 과민성, 고인슐린혈증, 관상 심장 질환, 협심증, 울혈성 심부전, 졸중, 담석, 담낭염, 담석증, 통풍, 골관절염, 폐쇄성 수면 무호흡증 및 호흡 곤란, 쓸개 질환, 일부 형태의 암(예를 들어, 자궁내막암, 유방암, 전립선암 및 결장암) 및 정신적 장애(예컨대, 우울증, 섭식 장애, 왜곡된 신체 이미지 및 약한 자존심)와 같은 질병을 앓을 위험이 높다. 비만이 건강에 미치는 부정적인 결과는 비만을 미국에서 막을 수 있는 사망의 원인 중 두 번째로 높은 원인이 되게 하고 사회적으로 상당한 경제적 및 심리사회적 영향을 미친다. 예를 들어, 문헌[McGinnis M, Foege WH., "Actual Causes of Death in the United States," JAMA 270:2207-12, 1993]을 참조한다.
- <6> 현재, 비만은 그와 관련된 건강 위험을 감소시키기 위한 치료를 요구하는 만성 질환으로서 인식되어 있다. 체중 감소가 중요한 치료 목적이지만, 비만 관리의 주 목적 중 하나는 비만-관련 이환률 및 사망률을 감소시키기 위해 심혈관 수치 및 대사 수치를 개선시키는 것이다. 체중의 5 내지 10% 감소는 대사 수치, 예컨대, 혈당, 혈압 및 지질 농도를 실질적으로 개선시킬 수 있음이 밝혀져 있다. 따라서, 체중의 5 내지 10% 감소는 이환률 및 사망률을 감소시킬 수 있는 것으로 생각된다. 비만을 관리하기 위한 현재 이용가능한 처방 약물은 일반적으로 식이 지방 흡수를 감소시키거나(예컨대, 오를리스타트(orlistat) 음식 섭취를 감소시키고/시키거나 에너지 소비를 증가시킴으로써(예컨대, 시부트라민(sibutramine)) 체중을 감소시킨다.
- <7> II형 당뇨병에 대한 현행 치료법은 외래 인슐린의 투여, 약물의 경구 투여, 식이 요법 및 운동 요법을 포함한다. 2005년, 엑세나티드(exenatide)(엑센딘-4: 바이에타(Byetta[®]))가 메트포민 및/또는 설포닐우레아를 섭취하나 적절한 혈당 조절을 달성하지 못한 II형 당뇨병 환자에 대한 보조 요법으로서 FDA에 의해 승인되었다. 엑세나티드는 엑센딘-4, 즉 힐러 몬스터의 침샘의 내생 생성물인 강력한 GLP-1 수용체 작용제이다. GLP-1처럼, 엑센딘-4는 인크레틴이다. 엑센딘-4는 인슐린친화성을 나타내고, 음식 섭취 및 위 배출을 억제하며 설치류에서 β-세포에 대한 친화성을 나타낸다(Parks et al., Metabolism. 50: 583-589, 2001; Aziz and Anderson, J. Nutr. 132: 990-995, 2002; and Egan et al., J. Clin. Endocrinol. Metab. 87: 1282-1290, 2002). 뿐만 아니라, 그의 N-말단의 2 위치에 글리신이 존재하기 때문에 GLP-1처럼 DPPIV에 대한 기질이 아니다. 엑세나티드의 사용에 있어서 단점은 엑세나티드의 반감기, 즉 t_{1/2}이 2 내지 4시간뿐이기 때문에 매일 2회씩 주사되어야 한다는 점이다(Kolterman et al., J. Clin. Endocrinol. Metab. 88: 3082-3089, 2003 and Fineman et al., Diabetes Care. 26: 2370-2377, 2003).
- <8> 따라서, 혈당 조절을 제공하고 체중을 감소시키는 치료제로서 사용될 수 있는 분해 저항성 GLP-1 수용체 작용제 분자가 오랫동안 요구되어 왔다. 장기간 작용하는 인크레틴 유사체의 개발은 보다 낮은 투약 빈도라는 장점과 함께 당-의존성 인슐린 분비의 연속적인 상승을 통해 혈당 조절을 증가시키는 능력을 제공한다. 본 발명은 생체활성을 유지하면서 엑센딘-4의 생체내 순환 반감기를 연장시키는, 변경된 트랜스페린(Tf)에 융합된 엑센딘-4 분자를 제공함으로써 상기 요구를 충족시킨다. 또한, 본 발명의 융합 단백질의 사용은 현재 인크레틴의 사용과 관련된 구역 및 구토의 높은 발생률을 감소시킬 수 있다.

발명의 상세한 설명

- <13> 본 발명은 펩티드 링커, 바람직하게는 비나선형 펩티드 링커를 통해 Tf 분자에 융합된 엑센딘-4를 포함하는 융합 단백질을 제공한다.
- <14> 바람직하게는, 상기 링커는 PEAPTD(서열번호 6), (PEAPTD)₂(서열번호 5), IgG 힌지 링커와 조합된 PEAPTD(서열번호 6) 및 IgG 힌지 링커와 조합된 (PEAPTD)₂(서열번호 5)으로 구성된 군으로부터 선택된다. 보다 바람직하게는, 링커는 (PEAPTD)₂(서열번호 5)이다.
- <15> 본 발명의 융합 단백질의 Tf 부분(moiety)임의의 포유동물 Tf, 바람직하게는 인간 Tf로부터 유래할 수 있다. 보다 바람직하게는, Tf는 천연 Tf 분자에 비해 감소된 글리코실화를 나타내는 변경된 Tf(mTf)이고, 훨씬 더 바람직하게는 Tf는 서열번호 17에 나타낸 아미노산 서열을 가진다. 다른 바람직한 실시양태에서, Tf는 철 결합 및/또는 Tf 수용체에의 결합을 감소시키도록 변경된다.
- <16> 또 다른 바람직한 실시양태에서, 상기 융합 단백질의 N-말단은 분비 신호 서열, 바람직하게는 혈청 Tf, 락토페린(Lf), 멜라노트랜스페린 또는 이들의 변이체로부터의 신호 서열, 보다 바람직하게는, 인간 혈청 알부민(HSA)/MF α-1 하이브리드 리더 서열, 변경된 HSA/MF α-1 하이브리드 리더 서열 또는 Tf 신호 서열을 추가로 포함하고, 훨씬 더 바람직하게는 상기 신호 서열은 서열번호 18에 나타낸 인간 Tf 신호 서열(nL)이다.
- <17> 바람직한 실시양태에서, 본 발명은 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 포함하는 융합 단백질을 제공하는데, 이때 상기 융합 단백질은 서열번호 23 또는 25에 나타낸 아미노산 서열을 포함하고, 서열번호 25는 N-말단에서 nL 리더 서열을 추가로 포함한다. 다른 바람직한 실시양태에서, 엑센딘-4는 엑센딘-4(1-39)이고 서열번호 4에 나타낸 아미노산 서열을 가지고/가지거나, 엑센딘-4 분자는 융합 단백질의 N-말단, C-말단 또는 N-말단 및 C-말단 둘다에서 융합되어 있다.
- <18> 또한, 본 발명은 상기 융합 단백질을 코딩하는 핵산 분자, 이 핵산 분자를 포함하는 상응하는 벡터, 및 상기 핵산 분자 및 벡터를 포함하는 숙주 세포를 제공한다.
- <19> 또한, 본 발명은 상기 융합 단백질 중 임의의 융합 단백질 및 약학적으로 허용가능한 담체를 포함하는 약학 조성물을 제공한다.
- <20> 바람직한 실시양태에서, 상기 약학 조성물은 서열번호 23의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 포함하고, 일부 실시양태에서 서열번호 23의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 포함하는 상기 약학 조성물은 약 0.5 내지 약 50 mg 또는 약 1 내지 약 100 mg의 투여량으로 투여된다.
- <21> 또 다른 바람직한 실시양태에서, 상기 약학 조성물은 흡입 투여된다.
- <22> 본 발명은 폴리펩티드 링커, 바람직하게는 비나선형 링커를 통해 Tf에 융합된 엑센딘-4를 포함하는 융합 단백질의 치료 유효량을 II형 당뇨병의 치료 또는 혈당 감소가 필요한 인간 환자에게 투여하는 단계를 포함하는, 상기 환자에서 II형 당뇨병을 치료하거나 혈당을 감소시키는 방법도 제공한다.
- <23> 바람직하게는, 상기 방법은 서열번호 23에 나타낸 아미노산 서열을 포함하는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 투여하는 단계를 포함하고, 일부 실시양태에서, 서열번호 23에 나타낸 융합 단백질은 주당 약 1회, 2주당 약 1회, 월당 약 1회의 빈도로 약 0.5 내지 약 50 mg의 투여량으로 투여된다. 또 다른 실시양태에서, 폴리펩티드 링커, 바람직하게는 비나선형 링커를 통해 Tf에 융합된 엑센딘-4, 보다 바람직하게는 서열번호 23에 나타낸 융합 단백질은 동일한 투여량으로 치료 효과를 얻는 데 있어서 엑세나티드보다 덜 빈번하게 투여된다.
- <24> 또한, 본 발명은 폴리펩티드 링커, 바람직하게는 비나선형 링커를 통해 Tf에 융합된 엑센딘-4를 포함하는 융합 단백질의 치료 유효량을 비만의 치료 또는 체중 감소가 필요한 인간 환자에게 투여하는 단계를 포함하는, 상기 환자에서 비만을 치료하거나 체중을 감소시키는 방법을 제공한다. 바람직하게는, 상기 융합 단백질은 서열번호 23에 나타낸 아미노산 서열을 포함하는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 포함하고, 일부 실시양태에서, 서열번호 23에 나타낸 융합 단백질은 주당 약 1회, 2주당 약 1회, 월당 약 1회의 빈도로 약 1 내지 약 100 mg의 투여량으로 투여된다. 또 다른 실시양태에서, 폴리펩티드를 통해 Tf에 융합된 엑센딘-4, 바람직하게는 서열번호 23에 나타낸 융합 단백질은 치료 효과를 얻는 데 있어서 엑세나티드보다 덜 빈번하게 투여

된다.

- <25> 또한, 본 발명은 II형 당뇨병의 치료 또는 혈당 감소가 필요한 환자에서 약 0.5 내지 약 50 mg의 투여량으로 투여되는 II형 당뇨병 치료 또는 혈당 감소용 약제의 제조에 있어서, 또는 비만의 치료 또는 체중의 감소가 필요한 환자에서 약 1 내지 약 100 mg의 투여량으로 투여되는 비만 치료 또는 체중 감소용 약제의 제조에 있어서 엑센딘-4/Tf 융합 단백질의 용도, 또는 이 융합 단백질, 바람직하게는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질, 보다 바람직하게는 서열번호 23에 나타난 융합 단백질을 포함하는 약학 조성물의 용도를 제공한다.
- <26> "엑센딘-4"는 서열번호 4에 나타난 엑센딘-4(1-39); 예를 들어, 엑센딘-4(1-31) 또는 엑센딘-4(1-30)을 생성하기 위해 서열번호 4에 나타난 서열의 C-말단으로부터 제거된 8 내지 9개 이하의 아미노산 잔기를 갖는 엑센딘-4 단편; 및 엑센딘-4(1-39) 또는 상기 다른 엑센딘-4 단편 중 하나와 90% 이상, 바람직하게는 95% 이상 동일한 펩티드를 의미한다.
- <27> 본원에서 사용된 바와 같이, 2개 이상의 DNA 코딩 서열은 DNA 코딩 서열 사이의 인-프레임(in-frame) 융합으로 인해 이 DNA 코딩 서열이 융합 단백질로 번역되는 경우 "연결되어 있다고" 또는 "융합되어 있다"고 기술된다. "연결되어 있는" 또는 "융합되어 있는"은 다른 방법, 예를 들어, 화학적 방법에 의해 융합된 펩티드를 지칭하는데 사용될 수도 있다. Tf 융합과 관련하여 용어 "융합"은 하나 이상의 치료 단백질, 폴리펩티드 또는 펩티드가 Tf의 N-말단 또는 C-말단에 부착되고/되거나 Tf 내의 임의의 2개 아미노산 사이에 삽입되는 것을 포함하나 이들로 한정되지 않는다.
- <28> "약학적으로 허용가능한"은 물질 또는 조성물이 다른 성분 함유 제제 및/또는 치료될 포유동물과 화학적으로 및/또는 독성학적으로 상용가능해야 함을 의미한다.
- <29> "치료 유효량"은 치료 전에 측정된 적당한 대조군 값 또는 비히클-처리 군에서 측정된 적당한 대조군 값에 비해 혈당, 칼로리 섭취, 체중 및/또는 체지방을 감소시키는 본 발명의 엑센딘-4/Tf 융합 단백질의 양을 의미한다.
- <30> 용어 "치료하는", "치료한다" 또는 "치료"는 방지적, 즉 예방적 및 완화적 치료 둘다를 포괄한다.
- <31> 본 발명의 엑센딘-4/Tf 융합 단백질에 대한 상세한 설명
- <32> 본 발명의 엑센딘-4/Tf 융합 단백질은 폴리펩티드 링커를 통해 Tf 펩티드에 융합된 엑센딘-4를 포함한다. 바람직하게는, 전장 엑센딘-4(1-39)(서열번호 4)를 사용하거나, 예를 들어, 엑센딘-4(1-31) 또는 엑센딘-4(1-30)을 생성하기 위해 서열번호 4에 나타난 서열의 C-말단으로부터 제거된 8 또는 9개 이하의 아미노산을 가진 엑센딘-4 단편을 사용한다.
- <33> 바람직하게는, 비나선형 폴리펩티드 링커 부분을 사용하여 엑센딘-4를 Tf에 연결한다.
- <34> 바람직한 링커는 PEAPTDPEAPTD(서열번호 5)이다. 다른 링커는 PEAPTD(서열번호 6), IgG 힌지 링커(서열번호 7 내지 16)와 조합된 PEAPTD(서열번호 6), 및 IgG 힌지 링커(서열번호 7 내지 16)와 조합된 PEAPTDPEAPTD(서열번호 5)로 구성된 군으로부터 선택될 수 있다. 실질적으로 비나선형인 링커 부분을 포함하는 본 발명의 융합 단백질은 실질적으로 비-나선형인 링커를 갖지 않는 유사한 융합 단백질에 비해 증가한 발현 생산성을 나타낼 수 있다. 뿐만 아니라, 실질적으로 비나선형인 링커를 포함하는 엑센딘-4/Tf 융합 단백질은 나선형 폴리펩티드 링커를 가진 유사한 융합 단백질에 비해 증가한 발현 생산성을 나타낼 수 있다.
- <35> 바람직한 엑센딘-4/Tf 융합 단백질은 링커 (PEAPTD)₂(서열번호 5)를 통해 서열번호 17에 나타난 mTf에 연결된 엑센딘-4(1-39)(서열번호 4)를 포함한다. 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질이 생성되는 경우, 이 융합 단백질은 인간 Tf 분비 신호 또는 리더 서열(nL)(서열번호 18)도 포함한다. 상기 바람직한 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 성분들 각각을 코딩하는 핵산 서열은 다음과 같다: nL 리더 서열(서열번호 19), 엑센딘-4(1-39)(서열번호 20), (PEAPTD)₂(서열번호 21) 및 mTf(서열번호 22). nL 리더 서열이 없는 전체 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질에 대한 아미노산 서열은 서열번호 23이고, 이의 상응하는 핵산 서열은 서열번호 24이다. N-말단에서 nL 리더 서열을 가진 전체 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 아미노산 서열은 서열번호 25이고, 이의 상응하는 핵산 서열은 서열번호 26이다.
- <36> 바람직한 mTf가 전술되어 있지만, 임의의 Tf를 사용하여 본 발명의 엑센딘-4/Tf 융합 단백질을 제조할 수 있다.

예를 들면, 야생형 인간 Tf는 유전자 복제로부터 유래한 것으로 보이는 2개의 주 도메인 또는 로브(lobe), 즉 N 로브(약 330개 아미노산) 및 C 로브(약 340개 아미노산)를 가진 약 75 kDa의 679개 아미노산 단백질(글리코실화에 대해 설명되어 있지 않음)이다. 진뱅크(GenBank) 검색 번호 NMJD01063, XM_002793, M12530, XM_039845, XM_039847 및 S95936을 참조한다(상기 번호 및 서열번호 2 및 3(서열번호 2는 nL 인간 Tf 리더 서열의 추가 19개 아미노산 서열을 포함함)은 참고로 본원에 도입되어 있음). 상기 2개의 도메인은 시간이 지나면서 달라지지만 상당한 정도의 동일성/유사성을 보유한다.

- <37> N 및 C 로브 각각은 2개의 서브도메인, N1과 N2 및 C1과 C2로 더 분류된다. Tf의 기능은 철을 체내 세포로 수송하는 것이다. 이 과정은 모든 세포, 특히 활발하게 성장하는 세포 상에서 발현되는 Tf 수용체(TfR)에 의해 매개된다. TfR은 Tf의 철 결합 형태(수용체당 2개의 분자가 결합됨)를 인식하여 세포내이입을 일으키고, 이로 인해 TfR/Tf 결합체가 엔도솜으로 수송된다. 엔도솜 내의 pH가 극소적으로 낮아짐으로써 결합된 철이 방출, TfR/Tf 결합체가 세포 표면으로 재순환되고 Tf(그의 철-비결합 형태인 apoTf로서 공지됨)가 방출된다. 수용체 결합은 Tf의 C 도메인을 통해 일어난다. C 도메인 내의 2개의 글리코실화 부위는 글리코실화되지 않은 철 결합 Tf가 수용체에 결합되기 때문에 수용체 결합에 관여하는 것으로 보이지 않는다.
- <38> Tf 분자 각각은 2개의 철 이온(Fe^{3+})을 운반할 수 있다. 상기 분자들은 N1 서브도메인과 N2 서브도메인 및 C1 서브도메인 과 C2 서브도메인 사이의 공간에서 결합체를 형성하여 분자의 구조적 변화를 일으킨다.
- <39> 서열번호 3의 인간 Tf의 경우, 철 결합 부위는 적어도 아미노산 Asp 63(천연 Tf 신호 서열을 포함하는 서열번호 2의 Asp 82), Asp 392(서열번호 2의 Asp 411), Tyr 95(서열번호 2의 Tyr 114), Tyr 426(서열번호 2의 Tyr 445), Tyr 188(서열번호 2의 Tyr 207), Tyr 514 또는 517(서열번호 2의 Tyr 533 또는 Tyr 536), His 249(서열번호 2의 His 268) 및 His 585(서열번호 2의 His 604)를 포함한다. 힌지 영역은 적어도 서열번호 3의 N 도메인 아미노산 잔기 94-96, 245-247 및/또는 316-318뿐만 아니라 C 도메인 아미노산 잔기 425-427, 581-582 및/또는 652-658를 포함한다. 서열번호 3의 인간 Tf의 카보네이트 결합 부위는 적어도 아미노산 Thr 120(서열번호 2의 Thr 139), Thr 452(서열번호 2의 Thr 471), Arg 124(서열번호 2의 Arg 143), Arg 456(서열번호 2의 Arg 475), Ala 126(서열번호 2의 Ala 145), Ala 458(서열번호 2의 Ala 477), Gly 127(서열번호 2의 Gly 146), 및 Gly 459(서열번호 2의 Gly 478)를 포함한다.
- <40> 인간 Tf 변이체, 소, 돼지, 양, 개, 토끼, 래트, 마우스, 햄스터, 바늘두더지, 오리너구리, 닭, 개구리, 빨벌레, 원숭이, 다른 소과 동물, 고양이 및 양서류 종을 포함하는 임의의 동물 Tf 분자를 사용하여 본 발명의 융합 단백질을 생성할 수 있지만, 변경된 엑센딘-4/Tf 융합 단백질은 인간 유래의 단백질인 것이 바람직하다. 이 Tf 서열들 모두가 진뱅크 및 다른 공개 데이터베이스에서 용이하게 입수할 수 있다. 인간 Tf 핵산 서열이 이용가능하고(서열번호 1 및 상기 검색 번호 참조) 인간 Tf 핵산 서열을 사용하여 Tf 또는 Tf의 도메인과 선택된 치료 분자 사이의 유전적 융합체를 제조할 수 있다. 융합체는 관련 분자 예컨대, 락토 Tf(Lf)(진뱅크 검색 번호: NM 002343) 또는 뮤린 멜라노트랜스페린(진뱅크 검색 번호: NM_013900)으로부터 제조될 수도 있다.
- <41> 멜라노트랜스페린은 악성 흑색종 세포에서 높은 농도로 발견되는 글리코실화된 단백질이고 처음에는 인간 흑색종 항원 p97(Brown et al., 1982, Nature, 296: 171-173)로 지칭되었다. 멜라노트랜스페린은 인간 혈청 Tf, 인간 Lf 및 닭 Tf와 높은 서열 유사성을 갖는다(Brown et al., Nature, 296: 171-173, 1982; Rose et al., Proc. Natl. Acad. Sci. USA, 83: 1261-1265, 1986). 그러나, 이 수용체들과 달리, 멜라노트랜스페린에 대한 세포내 수용체가 동정되지 않았다. 멜라노트랜스페린은 철에 가역적으로 결합하고 2가지 형태로 존재하는데, 이 두 가지 형태 중 하나는 글리코실 포스파티딜이노시톨 앵커에 의해 세포막에 결합되고, 다른 형태는 가용성을 나타내고 활발하게 분비된다(Baker et al., FEBS Lett, 298, 1992: 215-218; Alemany et al., J. Cell Sci., 104: 1155-1162, 1993; Food et al., J. Biol. Chem. 274: 701 1-7017, 1994).
- <42> Lf, 즉 천연 방어 철-결합 단백질은 항균 활성, 항진균 활성, 항바이러스 활성, 항신생물 활성 및 소염 활성을 갖는 것으로 밝혀져 있다. 상기 단백질은 일반적인 균무리(flora)에 흔히 노출되는 외부 분비물, 즉 모유, 눈물, 코 배출물, 침, 기관지 점액, 위장액, 자궁-질 점액 및 정액에 존재한다. 또한, Lf는 순환하는 다형핵 호중구(PMN)의 2차 특이적 과립의 주 성분이다. 아포단백질은 부패 영역 중의 PMN이 탈과립화될 때 방출된다. Lf의 주요 기능은 체액 및 염증 영역에서 자유 철을 제거하여 자유 라디칼-매개 손상을 억제하고 미생물 및 신생물 세포를 침윤시키는 금속의 이용률을 감소시킨다는 것이다. 성체에서 ^{125}I Lf의 전환률을 조사한 연구에서, Lf가 간 및 비장에 의해 급속히 흡수되고 방사성이 간 및 비장에서 수주 동안 지속됨을 보여주었다(Bennett et al., Clin. Sci. (Lond.) 57: 453-460, 1979).

- <43> 본 발명의 엑센딘-4/Tf 융합 단백질의 Tf 부분은 Tf 스플라이스 변이체를 포함한다. 한 예에서, Tf 스플라이스 변이체는 인간 Tf의 스플라이스 변이체일 수 있다. 구체적으로, 인간 Tf 스플라이스 변이체는 진뱅크 검색 번호 AAA61140의 Tf 스플라이스 변이체일 수 있다.
- <44> 본 발명의 엑센딘-4/Tf 융합 단백질의 Tf 부분은 Lf 스플라이스 변이체를 포함한다. 한 예에서, 인간 혈청 Lf 스플라이스 변이체는 호중구 Lf의 신규 스플라이스 변이체일 수 있다. 구체적으로, 호중구 Lf 스플라이스 변이체는 진뱅크 검색 번호 AAA59479에 나타난 서열일 수 있다. 또한, 호중구 Lf 스플라이스 변이체는 신규 스플라이스 변이 영역을 포함하는 아미노산 서열 EDCIALKGEADA(서열번호 27)를 포함할 수 있다.
- <45> 별법으로, 본 발명의 엑센딘-4/Tf 융합 단백질의 Tf 부분은 펠라노트랜스페린 변이체를 포함한다.
- <46> 변경된 Tf 융합체는 임의의 Tf 단백질, 단편, 도메인 또는 개조된 도메인을 사용하여 제조할 수 있다. 예를 들어, 융합 단백질은 천연 Tf 신호 서열을 갖거나 갖지 않은 전장 Tf 서열을 사용하여 제조할 수 있다. Tf 융합 단백질은 단일 Tf 도메인, 예컨대, 개별 N 또는 C 도메인, 또는 2N 또는 2C 도메인을 포함하는 변경된 형태의 Tf를 사용하여 제조할 수도 있다(미국 특허 공보 제2006/0130158호 참조). 단일 C 도메인과 치료 단백질의 융합체가 생성될 수 있고, 이때 상기 C 도메인은 글리코실화를 감소시키거나, 억제하거나 예방하도록 변경된다. 별법으로, 단일 N 도메인의 사용은 Tf 글리코실화 부위가 C 도메인 및 N 도메인 내에 존재하기 때문에 유리하다. 바람직하게는, Tf 융합 단백질은 고 농도로 발현되는 단일 N 도메인을 가진다.
- <47> 본원에서 사용된 바와 같이, N-유사 도메인처럼 작용하도록 변경된 C 말단 도메인 또는 로브는 천연 또는 야생형 N 도메인 또는 로브와 실질적으로 유사한 글리코실화 패턴 또는 철 결합 성질을 나타내도록 변경된다. 바람직하게는, C 도메인 또는 로브는 글리코실화되지 않고 관련 C 도메인 영역 또는 아미노산을 천연 또는 야생형 N 도메인의 상응하는 영역 또는 부위에 존재하는 C 도메인 영역 또는 아미노산으로 치환시킴으로써 철에 결합하지 않도록 변경된다.
- <48> 본원에서 사용된 바와 같이, "2개의 N 도메인 또는 로브"를 포함하는 Tf 부분은 천연 C 도메인 또는 로브를 천연 또는 야생형 N 도메인 또는 로브 또는 변경된 N 도메인 또는 로브로 치환하도록 변경된 Tf 분자, 또는 야생형 또는 변경된 N 도메인과 실질적으로 유사한 작용을 하도록 변경된 C 도메인을 함유하는 Tf 분자를 포함한다.
- <49> 2개 도메인의 중첩(wiss PDB Viewer 3.7b2, Iterative Magic Fit) 및 직접적 아미노산 정렬(ClustalW multiple alignment)에 의한 상기 2개 도메인의 분석은 상기 2개 도메인이 시간이 지남에 따라 달라짐을 보여준다. 아미노산 정렬은 상기 2개의 도메인 사이의 42% 동일성 및 59% 유사성을 보여준다. 그러나, N 도메인의 약 80%가 C 도메인과 구조적으로 동일하다. C 도메인은 N 도메인에 비해 여러 과외의 이황화 결합도 가진다.
- <50> 한 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 Tf의 N 말단 로브를 2개 이상 포함한다. 추가 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 인간 혈청 Tf로부터 유래된 Tf의 N 말단 로브를 2개 이상 포함한다.
- <51> 엑센딘-4/Tf 융합 단백질의 Tf 부분은 서열번호 3의 Asp63, Gly65, Tyr95, Tyr188 및 His249로 구성된 군으로부터 선택된 1개 이상의 아미노산 잔기에서 돌연변이를 가진 Tf의 N 말단 로브 2개 이상; 서열번호 3의 Lys206 또는 His207에서 돌연변이를 가진 재조합 인간 혈청 Tf N-말단 로브 돌연변이체; 또는 Tf의 C 말단 로브 2개 이상을 포함할 수도 있다. 추가 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 인간 혈청 Tf로부터 유래된 Tf의 C 말단 로브를 2개 이상 포함한다.
- <52> 추가 실시양태에서, C 말단 로브는 글리코실화를 허용하지 않는 서열번호 3의 Asn413 및 Asn611 중 1개 이상의 아미노산의 돌연변이를 추가로 포함한다.
- <53> 또 다른 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 서열번호 3의 Asp392, Tyr426, Tyr514, Tyr517 및 His585로 구성된 군으로부터 선택된 1개 이상의 아미노산 잔기에서 돌연변이를 가진 Tf의 C 말단 로브를 2개 이상 포함하는데, 이때 상기 돌연변이체는 금속에 대한 결합 능력을 보유한다. 대체 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 서열번호 3의 Tyr426, Tyr514, Tyr517 및 His585로 구성된 군으로부터 선택된 1개 이상의 아미노산 잔기에서 돌연변이를 가진 Tf의 C 말단 로브를 2개 이상 포함하는데, 이때 상기 돌연변이체는 금속에 대한 감소된 결합 능력을 보유한다. 또 다른 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 서열번호 3의 Asp392, Tyr426, Tyr517 및 His585로 구성된 군으로부터 선택된 1개 이상의 아미노산 잔기에서 돌연변이를 가진 Tf의 C 말단 로브를 2개 이상 포함하는데, 이때 돌연변이체는 금속에 대한 결합 능력을 보유하지 않으며 실질적으로 N 도메인처럼 작용한다.

- <54> Tf의 C 도메인이 융합 단백질의 일부인 경우, 2개의 N-결합 글리코실화 부위, 즉 서열번호 3의 N413 및 N611에 상응하는 아미노산 잔기는 글리코실화 또는 과다만노실화(hypermannosylation)를 방지하고 융합 단백질 및/또는 치료 단백질의 혈청 반감기를 연장하도록(아시알로-Tf를 생성하도록, 또는 일부 경우 모노시알로-Tf 또는 다이시알로-Tf를 생성하도록) 효모 시스템에서의 발현을 위해 돌연변이될 수 있다. N413 및 N611에 상응하는 Tf 아미노산 외에, 글리코실화를 방지하거나 실질적으로 감소하도록 N-X-S/T 글리코실화 부위 내의 인접한 잔기를 돌연변이시킬 수 있다. 미국 특허 제5,986,067호를 참조한다. 또한, 피치아 파스토리스(*Pichia pastoris*)에서 발현된 Tf의 N 도메인이 글리코실화를 방지하도록 돌연변이될 수 있거나 변경될 수 있는 S32에서 단일 헥소스로 O-결합 글리코실화된다고 보고되어 있다.
- <55> 따라서, 엑센딘-4/Tf 융합 단백질은 감소된 글리코실화를 나타내는 변경된 Tf 분자(아시알로-Tf, 모노시알로-Tf 및 다이시알로-Tf를 포함하나 이들로 한정되지 않음)를 포함할 수도 있다. 또 다른 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 글리코실화를 방지하도록 돌연변이된 재조합 Tf 돌연변이체를 포함한다. 엑센딘-4/Tf 융합 단백질의 Tf 부분은 완전히 글리코실화된 재조합 트랜스페린 돌연변이체도 포함할 수 있다. 추가 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 N-결합 글리코실화를 방지하도록 돌연변이된 재조합 인간 혈청 Tf 돌연변이체를 포함하는데, 이때 서열번호 3의 Asn413 및 Asn611 중 하나 이상이 글리코실화를 허용하지 않는 아미노산으로 돌연변이된다. 또 다른 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 글리코실화를 방지하거나 글리코실화를 실질적으로 감소시키도록 돌연변이된 재조합 인간 혈청 Tf 돌연변이체를 포함하는데, 이때 예를 들어, N-X-S/T 글리코실화 부위 내의 인접 잔기, 예컨대, S/T 잔기를 돌연변이시킨다. 나아가, 글리코실화는 세린 또는 스레오닌 잔기를 돌연변이시킴에 의해 감소할 수 있거나 방지될 수 있다. 또한, X를 프롤린으로 치환하면 글리코실화가 억제되는 것으로 공지되어 있다.
- <56> 하기에 더 상세히 논의된 바와 같이, 본 발명의 변경된 Tf 융합 단백질은 철 및/또는 Tf 수용체에 결합하지 않도록 개조될 수도 있다. 본 발명의 다른 실시양태에서, 철 결합은 보유되고, Tf의 철 결합 능력은 치료 단백질 또는 펩티드를 상피 또는 내피 세포막을 가로질러 세포 내부로 전달하는 데 이용될 수 있다. 철 및/또는 Tf 수용체에 결합하는 이 실시양태는 치료 단백질의 혈청 반감기를 연장하도록 글리코실화를 감소시키거나 방지하게끔 개조될 것이다. N 도메인은 철이 적재되어 있는 경우 단독으로 TfR에 결합하지 않을 것이고, 철 결합 C 도메인은 TfR에 결합할 것이지만 전체 분자와 동일한 친화도로 TfR에 결합하지는 않을 것이다.
- <57> 별법으로, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 돌연변이를 가진 재조합 Tf 돌연변이체를 포함할 수 있고, 이때 상기 돌연변이체는 금속 이온과의 결합 능력을 보유하지 않는다. 대체 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 돌연변이를 가진 재조합 Tf 돌연변이체를 포함할 수 있고, 이때 상기 돌연변이체는 야생형 혈청 Tf보다 더 약한 금속 이온에 대한 결합 친화성을 나타낸다. 대체 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 돌연변이를 가진 재조합 Tf 돌연변이체를 포함할 수 있고, 이때 상기 돌연변이체는 야생형 혈청 Tf보다 더 강한 금속 이온에 대한 결합 친화성을 나타낸다.
- <58> 또 다른 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 돌연변이를 가진 재조합 Tf 돌연변이체를 포함할 수 있고, 이때 상기 돌연변이체는 TfR에 대한 결합 능력을 보유하지 않는다. 예를 들어, 엑센딘-4 및 본 발명의 Tf 융합 단백질은 세포 표면 GLP-1 수용체에 결합할 수 있으나 Tf 수용체에는 결합할 수 없다. 이러한 융합 단백질은 세포 표면에서, 즉 세포로 들어가지 않고 치료 활성을 나타낼 수 있다.
- <59> 별법으로, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 야생형 혈청 Tf보다 더 약한 TfR에 대한 결합 친화성을 나타내는 돌연변이를 가진 재조합 Tf 돌연변이체; 야생형 혈청 Tf보다 더 강한 TfR에 대한 결합 친화성을 나타내는 돌연변이를 가진 재조합 Tf 돌연변이체; 카보네이트 이온에 대한 결합 능력을 보유하지 않는 돌연변이를 가진 재조합 Tf 돌연변이체; 야생형 혈청 Tf보다 더 약한 카보네이트 이온에 대한 결합 친화성을 가진 돌연변이를 가진 재조합 Tf 돌연변이체; 야생형 혈청 Tf보다 더 강한 카보네이트 이온에 대한 결합 친화성을 가진 돌연변이를 가진 재조합 Tf 돌연변이체를 포함할 수 있다.
- <60> 또 다른 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 서열번호 3의 Asp63, Gly65, Tyr95, Tyr188, His249, Asp392, Tyr426, Tyr514, Tyr517 및 His585로 구성된 군으로부터 선택된 1개 이상의 아미노산 잔기에서 돌연변이를 가진 재조합 인간 혈청 Tf 돌연변이체를 포함하는데, 이때 상기 돌연변이체는 금속 이온에 결합하는 능력을 보유한다. 대체 실시양태에서, 재조합 인간 혈청 Tf 돌연변이체는 서열번호 3의 Asp63, Gly65, Tyr95, Tyr188, His249, Asp392, Tyr426, Tyr514, Tyr517 및 His585로 구성된 군으로부터 선택된 1개 이상의 아미노산 잔기에서 돌연변이를 갖는데, 이때 상기 돌연변이체는 금속 이온에 대한 감소된 결합 능력을 보유한다. 또 다른 실시양태에서, 재조합 인간 혈청 Tf 돌연변이체는 서열번호 3의 Asp63, Gly65, Tyr95, TyM

88, His249, Asp392, Tyr426, Tyr517 및 His585로 구성된 군으로부터 선택된 1개 이상의 아미노산 잔기에서 돌연변이를 갖는데, 이때 상기 돌연변이체는 금속 이온에 대한 결합 능력을 보유하지 않는다.

- <61> 또 다른 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 서열번호 3의 Lys206 또는 His207에서 돌연변이를 가진 재조합 인간 혈청 Tf 돌연변이체를 포함하는데, 이때 상기 돌연변이체는 야생형 인간 혈청 Tf(미국 특허 제5,986,067호 참조)보다 더 강한 금속 이온에 대한 결합 친화성을 나타낸다. 대체 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 서열번호 3의 Lys206 또는 His207에서 돌연변이를 가진 재조합 인간 혈청 Tf 돌연변이체를 포함하는데, 이때 상기 돌연변이체는 야생형 인간 혈청 Tf보다 더 약한 금속 이온에 대한 결합 친화성을 나타낸다. 추가 실시양태에서, 엑센딘-4/Tf 융합 단백질의 Tf 부분은 서열번호 3의 Lys206 또는 His207에서 돌연변이를 가진 재조합 인간 혈청 Tf 돌연변이체를 포함하는데, 이때 상기 돌연변이체는 금속 이온에 결합하지 않는다.
- <62> 통상적으로 이용되는 분자적 기법, 예를 들어, 문헌[Sambrook et al. Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, 1989]에 개시된 기법을 포함하나 이로 한정되지 않은 임의의 이용가능한 기법을 이용하여 본 발명의 엑센딘-4/Tf 융합 단백질을 생성할 수 있다. 당업계에 잘 공지되어 있는 부위-특이적 돌연변이유발을 달성하기 위한 기법을 이용하여 뉴클레오타이드 치환을 수행하는 경우, 코딩되는 아미노산 변경은 바람직하게는 성질 면에서 미미한 변경, 즉 보존적 아미노산 치환이지만, 특히 Tf 융합 단백질, 예컨대, 감소한 글리코실화, 감소한 철 결합 등을 나타내는 변경된 Tf 단백질의 변경된 Tf 부분을 생성하는 경우 다른 비-보존적 치환도 고려할 수 있다. 구체적으로, 아미노산의 치환, 적은 결실 또는 삽입, 전형적으로 1 내지 약 30개의 아미노산의 결실 또는 삽입; Tf 도메인 사이의 삽입; 및 작은 아미노- 또는 카복실-말단 연장부, 예컨대, 아미노-말단 메티오닌 잔기, 또는 Tf 도메인 또는 결합 Tf 단백질과 엑센딘-4 사이의 50, 40, 30, 20 또는 10개의 잔기로 구성된 작은 링커 펩티드, 또는 정제를 용이하게 하는 작은 연장부, 예컨대, 폴리-히스티딘 트랙(tract), 항원성 에피토프 또는 결합 도메인이 고려된다.
- <63> 보존적 아미노산 치환의 예는 동일 군, 예컨대, 염기성 아미노산 군(예를 들어, 아르기닌, 라이신, 히스티딘), 산성 아미노산 군(예컨대, 글루탐산 및 아스파르트산), 극성 아미노산 군(예를 들어, 글루타민 및 아스파라긴), 소수성 아미노산 군(예를 들어, 류신, 이소류신, 발린), 방향족 아미노산 군(예를 들어, 페닐알라닌, 트립토판, 티로신) 및 작은 아미노산 군(예를 들어, 글리신, 알라닌, 세린, 스레오닌, 메티오닌) 내에서 수행된 치환이다.
- <64> 비-보존적 치환은 한 군 내의 아미노산을 또 다른 군 내의 아미노산으로 치환시키는 것을 포함한다. 예를 들어, 비-보존적 치환은 극성 아미노산을 소수성 아미노산으로 치환시키는 것을 포함할 것이다. 뉴클레오타이드 치환의 일반적인 설명을 위해서는 예를 들어, 문헌[Ford et al., Prot. Exp. Pur. 2: 95-107, 1991]을 참조한다. 비-보존적 치환, 결실 및 삽입은 철과의 결합을 나타내지 않거나 철과의 감소된 결합을 나타내고/내거나, Tf 융합 단백질과 TfR의 결합을 나타내지 않거나 융합 단백질과 TfR의 감소된 결합을 나타내고/내거나, 글리코실화를 나타내지 않거나 감소된 글리코실화를 나타내는 본 발명의 Tf 융합 단백질을 생성하는데 특히 유용하다.
- <65> 철 결합 및/또는 수용체 결합은 결실, 치환 또는 삽입을 포함하는 돌연변이를 서열번호 3의 Tf N 도메인 잔기 Asp63, Tyr95, Tyr188, His249 및/또는 C 도메인 잔기 Asp392, Tyr426, Tyr514 및/또는 His585 중 하나 이상에 상응하는 아미노산 잔기 내로 도입함으로써 감소시킬 수 있거나 파괴할 수 있다. 철 결합은 서열번호 3의 아미노산 Lys206, His207 또는 Arg632를 돌연변이시킴에 의해 영향을 받을 수도 있다. 카보네이트 결합은 결실, 치환 또는 삽입을 포함하는 돌연변이를 서열번호 3의 Tf N 도메인 잔기 Thr120, Arg124, Ala126, Gly 127 및/또는 C 도메인 잔기 Thr452, Arg456, Ala458 및/또는 Gly 459 중 하나 이상에 상응하는 아미노산 잔기 내로 도입함으로써 감소시킬 수 있거나 파괴할 수 있다.
- <66> TfR과의 결합은 결실, 치환 또는 삽입을 포함하는 돌연변이를 철 결합에 대해 전술한 Tf N 도메인 잔기들 중 하나 이상에 상응하는 아미노산 잔기 내로 도입함으로써 감소시킬 수 있거나 파괴할 수 있다.
- <67> 전술한 바와 같이, 글리코실화는 결실, 치환 또는 삽입을 포함하는 돌연변이를 C 도메인 잔기 N413 및/또는 N611에 상응하는 N-X-S/T 부위 주변의 Tf C 도메인 잔기들 중 하나 이상에 상응하는 아미노산 잔기 내로 도입함으로써 감소시킬 수 있거나 방지할 수 있다(미국 특허 제5,986,067호 참조). 예를 들어, N413 및/또는 N611은 Glu 잔기로 돌연변이될 수 있다.
- <68> 본 발명의 Tf 융합 단백질이 글리코실화, 철 결합, 카보네이트 결합 및/또는 수용체 결합을 방지하도록 변경되지 않는 경우, 철 및/또는 카보네이트 이온은 상기 융합 단백질로부터 제거될 수 있거나 절단될 수 있다. 예를

들어, 이용가능한 데글리코실레이즈(deglycosylase)를 사용하여 상기 융합 단백질로부터 글리코실화 잔기, 특히 Tf 부분에 부착된 당 잔기를 절단할 수 있고/있거나, 글리코실화 효소가 없는 효모를 사용하여 글리코실화를 방지할 수 있고/있거나, 재조합 세포를 글리코실화를 방지하는 물질, 예를 들어, 투니카마이신(tunicamycin)의 존재 하에 성장시킬 수 있다.

<69> 융합 단백질 상의 탄수화물도 상기 융합 단백질을 데글리코실레이즈로 처리하여 감소시킬 수 있거나 완전히 효소적으로 제거할 수 있다. 데글리코실레이즈는 당업계에 잘 공지되어 있다. 데글리코실레이즈의 예는 갈락토시다이스, PNGase A, PNGase F, 글루코시다이스, 만노시다이스, 푸코시다이스 및 엔도 H 데글리코실레이즈를 포함하나 이들로 한정되지 않는다.

<70> 그럼에도 불구하고, 일부 경우, 융합 단백질의 Tf 부분이 완전히 글리코실화되어 있는 상태로 경구 전달하는 것이 바람직할 수 있다.

<71> Tf의 3차원적 구조를 변경하도록 Tf를 추가로 돌연변이시킬 수 있는데, 예를 들어, 철 결합 및 Tf 수용체 인식에 필요한 구조적 변화를 방지하도록 힌지 영역을 변경시킬 수 있다. 예를 들어, 돌연변이를 N 도메인 아미노산 잔기 94-96, 245-247 및/또는 316-318뿐만 아니라 C 도메인 아미노산 잔기 425-427, 581-582 및/또는 652-658 내에 또는 주변에 도입할 수 있다. 또한, 돌연변이를 Tf 구조 및 기능을 변경하도록 이 부위들의 측부 영역 내에 또는 주변에 도입할 수 있다.

<72> 엑센딘-4/Tf 융합 단백질은 치료 단백질의 반감기 또는 생체이용률을 연장시키기 위한 담체 단백질로서 작용할 수 있을 뿐만 아니라, 일부 경우 치료 단백질을 세포 내부로 및/또는 혈액-뇌 장벽(BBB)를 가로질러 전달하는 데 있어서 담체 단백질로서 작용할 수 있다. 대체 실시양태에서, 융합 단백질은 변경된 Tf 분자를 포함하고, 이때 Tf는 BBB를 가로지르는 능력을 보유하지 않는다.

<73> 또 다른 실시양태에서, 엑센딘-4/Tf 융합 단백질은 변경된 Tf 분자를 포함하고, 이때 Tf 분자는 TfR에 결합하여 치료 펩티드를 세포 내부로 전달하는 능력을 보유한다. 대체 실시양태에서, 엑센딘-4/Tf 융합 단백질은 변경된 Tf 분자를 포함하고, 이때 Tf 분자는 TfR에 결합하여 치료 펩티드를 세포 내부로 수송하는 능력을 보유하지 않는다.

<74> 추가 실시양태에서, 엑센딘-4/Tf 융합 단백질은 변경된 Tf 분자를 포함하고, 이때 Tf 분자는 TfR에 결합하여 치료 펩티드를 세포 내부로 수송하는 능력 및 BBB를 가로지르는 능력을 보유한다. 대체 실시양태에서, 엑센딘-4/Tf 융합 단백질은 변경된 Tf 분자를 포함하고, 이때 Tf 분자는 BBB를 가로지르는 능력을 보유하지만 TfR에 결합하여 치료 펩티드를 세포 내부로 수송하는 능력을 보유하지는 않는다.

<75> 본 발명의 변경된 융합 단백질은 펩티드 결합 또는 변경된 펩티드 결합에 의해 서로 연결된 아미노산으로 구성될 수 있고 20개의 유전자-코딩 아미노산 이외의 아미노산을 함유할 수 있다. 폴리펩티드는 번역 후 프로세싱과 같은 천연 과정에 의해 또는 당업계에 잘 공지되어 있는 화학적 변경 기법에 의해 변경될 수 있다. 이러한 변경은 기본 교체에 잘 기재되어 있고 논문 및 연구 문헌에 보다 상세히 기재되어 있다.

<76> 돌연변이는 펩티드 골격, 아미노산 측쇄 및 아미노산 또는 카복시 말단을 포함하는 폴리펩티드의 임의의 위치에서 일어날 수 있다. 동일한 유형의 변경이 소정의 한 폴리펩티드 내의 여러 부위에서 동일한 정도 또는 다양한 정도로 존재할 수 있다. 또한, 소정의 한 폴리펩티드가 많은 유형의 변경을 포함할 수 있다. 폴리펩티드는 유비퀴틴화(ubiquitination)로 인해 분지될 수 있고 분지를 갖거나 갖지 않는 고리 구조를 가질 수 있다. 고리형, 분지형 및 분지된 고리형 폴리펩티드는 천연 번역 후 프로세스로부터 발생할 수 있거나 합성 방법에 의해 제조될 수 있다. 변경은 아세틸화, 아실화, ADP-리보실화, 아미드화, 플라빈의 공유결합, 헴(heme) 잔기의 공유결합, 뉴클레오타이드 또는 뉴클레오타이드 유도체의 공유결합, 지질 또는 지질 유도체의 공유결합, 포스포티딜이노시톨의 공유결합, 가교결합, 고리화, 이황화 결합 형성, 탈메틸화, 공유 가교결합의 형성, 시스테인의 형성, 글리코실화, GPI 앵커 형성, 하이드록실화, 요오딘화(iodination), 메틸화, 미리스틸화, 산화, 페그화, 단백질분해 프로세싱, 인산화, 프레닐화, 라세미체화, 설파화(sulfation), 단백질에의 아미노산의 전달-RNA 매개 부가, 예컨대, 아르기닐화, 및 유비퀴틴화를 포함한다(예를 들어, 문헌[Proteins - Structure and Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Post-translational Covalent Modification of Proteins, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12, 1983; and Seifter et al. Meth. Enzymol. 182:626-646, 1990] 참조).

<77> 엑센딘-4/Tf를 코딩하는 핵산 분자

- <78> 본 발명은 엑센딘-4/Tf 융합 단백질을 코딩하는 핵산 분자도 제공한다. 바람직한 핵산 분자는 (PEAPTD)₂(서열번호 5)을 통해 mTf에 연결된 엑센딘-4(1-39)의 아미노산 서열인 서열번호 23을 코딩한다. 예시적 핵산 서열은 서열번호 24에 나타나 있다. 가장 바람직하게는, 본 발명의 핵산 서열은 인간 Tf 분비 신호 또는 리더 서열을 나타내는 추가 N-말단 19개 아미노산을 가진 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 아미노산 서열인 서열번호 25를 코딩한다. 서열번호 25를 코딩하는 예시적 핵산 서열은 서열번호 26으로서 표시된다.
- <79> 엑센딘-4/Tf 융합 단백질을 코딩하는 서열은 C-말단에서 정지 코돈(예를 들어, tga, taa, tag)도 포함할 수 있고 화학적 합성법; cDNA 또는 게놈 라이브러리 스크리닝, 발현 라이브러리 스크리닝 및/또는 cDNA의 중합효소 연쇄 반응(PCR) 증폭으로부터 얻은 Tf 폴리뉴클레오타이드 서열 및 야생형 엑센딘-4의 유전적 돌연변이를 포함하나 이들로 한정되지 않는 다양한 방법으로 용이하게 얻을 수 있다. 엑센딘-4/Tf 융합 단백질을 코딩하는 핵산 분자는 부위 지정 돌연변이유발, PCR 증폭 또는 다른 적절한 방법을 이용하여 제조할 수 있고, 이때 상기 프라이머는 원하는 점 돌연변이를 가진다. 본원에 기재된 재조합 DNA 방법 및 돌연변이유발 방법은 일반적으로 문헌[Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989, and Current Protocols in Molecular Biology, Ausubel et al., Green Publishers Inc. and Wiley and Sons, 1994]에 기재되어 있다.
- <80> 엑센딘-4/Tf 융합 단백질의 아미노산 서열을 코딩하는 핵산 폴리뉴클레오타이드는 발현된 단백질의 성질에 기초하여 양성 클론을 검출하는 발현 클로닝에 의해 확인될 수 있다. 전형적으로, 핵산 라이브러리는 항체 또는 다른 결합 파트너(예를 들어, 수용체 또는 리간드)와 발현되어 숙주 세포 표면에 표시되는 클로닝된 단백질의 결합에 의해 스크리닝된다. 항체 또는 결합 파트너는 원하는 클론을 발현하는 세포를 확인하기 위한 검출 표지로 변경된다.
- <81> 하기 설명에 따라 수행되는 재조합 발현 기법은 엑센딘-4/Tf 융합 단백질 코딩 폴리뉴클레오타이드를 제조하고 코딩된 폴리펩티드를 발현시키기 위해 수행될 수 있다. 예를 들어, 당업자는 엑센딘-4/Tf 융합 단백질의 아미노산 서열을 코딩하는 핵산 서열을 적절한 벡터 내로 삽입함으로써 다량의 원하는 뉴클레오타이드 서열을 용이하게 제조할 수 있다. 그 다음, 상기 서열을 사용하여 검출 프로브 또는 증폭 프라이머를 생성할 수 있다. 별 방법으로, 엑센딘-4/Tf 융합 단백질의 아미노산 서열을 코딩하는 폴리뉴클레오타이드를 발현 벡터 내로 삽입할 수 있다. 상기 발현 벡터를 적절한 숙주 내로 도입함으로써 코딩된 엑센딘-4/Tf 융합 단백질을 다량으로 제조할 수 있다.
- <82> 적합한 핵산 서열을 수득하는 또 다른 방법은 PCR이다. 이 방법에서, cDNA는 역전사 효소를 사용하여 폴리(A)+RNA 또는 총 RNA로부터 제조한다. 그 다음, 전형적으로 엑센딘-4/Tf 융합 단백질의 아미노산 서열을 코딩하는 cDNA의 두 분리된 영역에 상보적인 2개의 프라이머를 Tag 중합효소와 같은 중합효소와 함께 상기 cDNA에 첨가하고, 상기 중합효소는 상기 2개의 프라이머 사이의 cDNA 영역을 증폭한다.
- <83> 폴리펩티드의 아미노 말단을 코딩하는 DNA 단편은 메티오닌 잔기를 코딩하는 ATG를 가질 수 있다. 이 메티오닌은 상기 숙주 세포 내에서 생성된 폴리펩티드가 그 세포로부터 분비되도록 디자인되는 지에 따라 엑센딘-4/Tf 융합 단백질의 성숙 형태에 존재할 수 있거나 존재할 수 없다. 이소류신을 코딩하는 코돈도 출발 부위로서 사용될 수 있다. 당업자에게 공지되어 있는 다른 방법도 이용할 수 있다. 일부 실시양태에서, 핵산 변이체는 소정의 숙주 세포 내에서 엑센딘-4/Tf 융합 단백질이 최적으로 발현되도록 변경된 코돈을 함유한다. 구체적인 코돈 변경은 엑센딘-4/Tf 융합 단백질 및 발현용으로 선택된 숙주 세포에 달려 있을 것이다. 이러한 코돈 최적화는 다양한 방법으로, 예를 들어, 소정의 숙주 세포에서 고도로 발현된 유전자에 사용하기에 바람직한 코돈을 선별함으로써 수행할 수 있다. 고도로 발현된 유전자의 코돈 선호도에 대한 코돈 빈도 표, 예컨대, "Eco_high.Cod"를 도입하는 컴퓨터 알고리즘이 이용될 수 있고 위스콘신 대학 팩키지 버전 9.0(Genetics Computer Group, 미국 위스콘신 매디슨 소재)에 의해 제공된다. 다른 유용한 코돈 빈도 표는 "Celegans_high.cod," "Celegans_low.cod," "Drosophila_high.cod," "Human_high.cod," "Maize_high.cod," 및 "Yeastjiigh.cod"를 포함한다.
- <84> 벡터
- <85> 엑센딘-4/Tf 융합 단백질의 아미노산 서열을 코딩하는 핵산 분자는 표준 라이게이션(ligation) 기법을 이용하여 적절한 발현 벡터 내로 삽입시킨다. 상기 벡터는 전형적으로 사용된 구체적 숙주 세포에서 기능할 수 있도록 선택된다(즉, 상기 벡터는 유전자의 증폭 및/또는 유전자의 발현이 일어날 수 있도록 숙주 세포 기구와 상용가능해야 함). 엑센딘-4/Tf 융합 단백질의 아미노산 서열을 코딩하는 핵산 분자는 원핵 숙주 세포, 효모, 곤충

(바칼로바이러스 시스템) 및/또는 진핵 숙주 세포에서 증폭될 수 있고/있거나 발현될 수 있다. 발현 벡터에 대해서는 문헌[Meth. Enz., vol. 185, D. V. Goeddel, Academic Press, 1990]을 참조한다.

- <86> 전형적으로, 임의의 숙주 세포에서 사용되는 발현 벡터는 플라스미드 유지 및 외래 뉴클레오타이드 서열의 클로닝 및 발현을 위한 서열을 포함할 것이다. 이러한 서열은 일부 실시양태에서 "플래깅 서열"로서 총칭되고, 전형적으로 하기 뉴클레오타이드 서열들 중 하나 이상을 포함할 것이다: 프로모터, 1개 이상의 인핸서 서열, 복제 기점, 전사 종결 서열, 공여 스플라이스 부위 및 수용 스플라이스 부위를 포함하는 완전한 인트론 서열, 폴리펩티드 분비를 위한 리더 서열을 코딩하는 서열, 리보솜 결합 부위, 폴리아데닐화 서열, 발현될 폴리펩티드를 코딩하는 핵산을 삽입시키기 위한 폴리링커 영역, 및 선별가능한 마커 요소. 이 서열들 각각은 하기에 논의되어 있다.
- <87> 경우에 따라, 벡터는 "태그"-코딩 서열, 즉 엑센던-4/Tf 융합 단백질 코딩 서열의 5' 또는 3' 말단에 위치한 올리고뉴클레오타이드를 포함할 수 있고, 상기 올리고뉴클레오타이드 서열은 폴리His(예컨대, 헥사His) 또는 또 다른 "태그", 예컨대, FLAG, HA(헤마글루티닌 인플루엔자 바이러스) 또는 myc(이들에 대한 시판되는 항체가 존재함)를 코딩한다. 이 태그는 전형적으로 폴리펩티드의 발현 시 폴리펩티드에 융합되어 있고 숙주 세포로부터의 엑센던-4/Tf 융합 단백질의 친화 정제를 위한 수단으로서 작용할 수 있다. 친화 정제는 예를 들어, 친화 매트릭스로서 태그에 대한 항체를 사용하는 컬럼 크로마토그래피에 의해 달성될 수 있다. 경우에 따라, 태그는 추후에 다양한 수단, 예컨대, 절단을 위한 특정 펩티데이즈를 사용함으로써, 예를 들어, 아미노산 서열들 중 하나의 업스트림에 존재하는 FLAG 태그 서열의 3'을 엔테로카이네이즈로 절단함으로써 정제된 엑센던-4/Tf 융합 단백질로부터 제거할 수 있다.
- <88> 플래깅 서열은 상동성(즉, 숙주 세포와 동일한 종 및/또는 균주로부터 유래함), 이종성(즉, 숙주 세포 종 또는 균주와 다른 종으로부터 유래함), 혼성(즉, 하나보다 많은 공급원으로부터 유래한 플래깅 서열의 조합물) 또는 합성 플래깅 서열일 수 있거나, 정상적으로 엑센던-4 발현을 조절하는 기능을 하는 천연 서열일 수 있다. 플래깅 서열의 공급원은 플래깅 서열이 숙주 세포 기구 내에서 기능할 수 있고 숙주 세포 기구에 의해 활성화될 수 있는 한 임의의 원핵 또는 진핵 유기체, 임의의 척추 또는 무척추 유기체 또는 임의의 식물일 수 있다.
- <89> 유용한 플래깅 서열은 당업자에게 잘 공지되어 있는 여러 방법들 중 임의의 방법에 의해 획득될 수 있다. 전형적으로, 본원에서 유용한 플래깅 서열은 맵핑(mapping) 및/또는 제한 엔도뉴클레이즈 절단에 의해 이미 확인되어 있을 것이므로 적절한 제한 엔도뉴클레이즈를 사용하여 적절한 조직 공급원으로부터 단리할 수 있다. 일부 경우, 플래깅 서열의 전장 뉴클레오타이드 서열은 공지되어 있다. 이때, 플래깅 서열은 핵산 합성 또는 클로닝에 대해 본원에 기재된 방법을 이용하여 합성할 수 있다.
- <90> 플래깅 서열의 전부 또는 단지 일부가 공지되어 있는 경우, 동일한 또는 또 다른 종으로부터의 적절한 올리고뉴클레오타이드 및/또는 플래깅 단편을 사용한 PCR 및/또는 게놈 라이브러리 스크리닝을 이용하여 플래깅 서열의 전장 뉴클레오타이드 서열을 얻을 수 있다. 플래깅 서열이 공지되어 있지 않은 경우, 플래깅 서열을 포함하는 DNA 단편은 예를 들어, 코딩 서열 또는 심지어 또 다른 유전자 또는 유전자들을 포함할 수 있는 큰 DNA 단편으로부터 단리할 수 있다. 단리는 제한 엔도뉴클레이즈 절단을 수행하여 적절한 DNA 단편을 생성한 후 아가로스 겔 정제, 퀴아젠(Qiagen[®]) 컬럼 크로마토그래피((Qiagen, 미국 캘리포니아주 채츠워스 소재) 또는 당업자에게 공지된 다른 방법을 이용하여 정제함으로써 달성할 수 있다. 이 목적을 달성하기에 적합한 효소의 선택은 당업자에게 용이할 것이다.
- <91> 복제 기점은 전형적으로 시판되는 원핵 발현 벡터의 일부이고 숙주 세포에서 벡터의 증폭을 돕는다. 일부 경우 벡터를 일정한 카피수까지 증폭하는 것은 엑센던-4/Tf 융합 단백질의 최적 발현에 중요할 수 있다. 선택된 벡터가 복제 기점 부위를 포함하지 않는 경우, 복제 기점은 공지된 서열에 기초하여 화학적으로 합성되어 벡터 내로 라이게이션될 수 있다. 예를 들어, 플라스미드 pBR322(New England Biolabs, 미국 매사추세츠주 베벌리 소재)로부터의 복제 기점은 대부분의 그람 음성 박테리아에 적합하고, 다양한 복제 기점(예를 들어, SV40, 폴리오마, 아데노바이러스, 수포성 구내염 바이러스(VSV) 또는 파필로마바이러스, 예컨대, HPV 또는 BPV)은 포유동물 세포에서 벡터를 클로닝하는 데 유용하다. 일반적으로, 복제 기점은 포유동물 발현 벡터의 경우에는 필요하지 않다(예를 들어, SV40 복제 기점만이 초기 프로모터를 포함하고 있기 때문에 종종 사용됨).
- <92> 전사 종결 서열은 폴리펩티드 코딩 영역의 3' 말단에 위치되어 있고 전사를 종결하는 기능을 한다. 통상적으로, 진핵 세포 내의 전사 종결 서열은 G-C 풍부 단편에 이은 폴리-T 서열이다. 상기 서열은 라이브러리로부터 용이하게 클로닝될 수 있거나 심지어 벡터의 일부로서 구입할 수 있으나, 핵산 합성 방법, 예컨대, 본

원에 기재된 방법을 이용하여 용이하게 합성할 수도 있다.

- <93> 선별가능한 마커 유전자 요소는 선별 배양 배지에서 성장된 숙주 세포의 생존 및 성장에 필요한 단백질을 코딩한다. 전형적인 선별 마커 유전자는 (a) 항생제 또는 다른 독소, 예를 들어, 앰피실린, 테트라사이클린 또는 원핵 숙주 세포의 경우 카나마이신에 대한 내성을 부여하는 단백질; (b) 세포의 영양요구성 결핍을 보충하는 단백질; 또는 (c) 복합 배지로부터 이용될 수 없는 중요한 영양분을 공급하는 단백질을 코딩한다. 바람직한 선별 마커는 카나마이신 내성 유전자, 앰피실린 내성 유전자 및 테트라사이클린 내성 유전자이다. 네오마이신 내성 유전자도 원핵 숙주 세포 및 진핵 숙주 세포에서 선별하는 데 사용될 수 있다.
- <94> 다른 선별 유전자는 사용하여 발현될 유전자를 증폭할 수 있다. 증폭은 성장에 중요한 단백질의 생성을 위해 필요한 유전자가 재조합 세포의 연속 세대의 염색체 내에서 직렬식으로 반복되는 과정이다. 포유동물 세포에 적합한 선별 마커의 예로는 다이하이드로폴레이트 리덕테이즈(DHFR) 및 티미딘 카이네이즈를 들 수 있다. 포유동물 세포 형질전환체는 형질전환체만이 벡터에 존재하는 선별 유전자에 의해 생존할 수 있는 선별 압력 하에 둔다. 선별 압력은 배지 중의 선별제의 농도가 연속적으로 변경되어 엑세던-4/Tf 융합 단백질을 코딩하는 DNA 및 선별 유전자 둘다의 증폭을 이끌어내는 조건 하에 형질전환된 세포를 배양함으로써 부가한다. 그 결과, 증가한 양의 엑세던-4/Tf 융합 단백질이 증폭된 DNA로부터 합성된다.
- <95> 리보솜 결합 부위는 통상적으로 mRNA의 번역 개시에 필요하고 샤인-달가노 서열(원핵생물) 또는 코작 서열(진핵생물)을 특징으로 한다. 리보솜 결합 부위는 전형적으로 프로모터에 대해 3' 및 발현될 엑세던-4/Tf 융합 단백질의 코딩 서열에 대해 5'에 위치한다. 샤인-달가노 서열은 다양하지만 전형적으로 폴리퓨린(즉, A-G 함량이 높음)이다. 많은 샤인-달가노 서열이 확인되었고, 이들 각각은 본원에 기재된 방법을 이용하여 용이하게 합성할 수 있고 진핵 벡터에서 사용할 수 있다.
- <96> 용어 "분비 신호 서열", "신호 서열" 또는 "분비 리더 서열"은 상호교환적으로 사용되고 예를 들어, 미국 특허 제6,291,212호 및 제5,547,871호에 기재되어 있다. 분비 신호 서열, 신호 서열 또는 분비 리더 서열은 분비 펩티드를 코딩한다. 분비 펩티드는 세포로부터 성숙 폴리펩티드 또는 단백질의 분비를 유도하는 작용을 하는 아미노산 서열이다. 분비 펩티드는 일반적으로 소수성 아미노산으로 이루어진 코어를 특징으로 하고 전형적으로 (그러나, 예외가 없는 것은 아님) 새로 합성된 단백질의 아미노 말단에서 발견된다. 분비 펩티드는 분비 과정 동안 매우 자주 성숙 단백질로부터 절단된다. 분비 펩티드는 분비 경로를 통과할 때 성숙 단백질로부터 신호 펩티드의 절단을 가능하게 하는 프로세싱 부위를 포함할 수 있다. 프로세싱 부위는 신호 펩티드 내에 코딩될 수 있거나, 예를 들어, 시험관내 돌연변이유발에 의해 신호 펩티드에 부가될 수 있다.
- <97> 분비 펩티드는 본 발명의 융합 단백질의 분비를 유도하는 데 사용될 수 있다. 다른 분비 펩티드와 함께 사용될 수 있는 이러한 분비 펩티드 중 하나는 알파 교배 인자 리더 서열이다. 분비 신호 서열, 신호 서열 또는 분비 리더 서열은 단백질을 분비시키는 일련의 복합적 번역 후 프로세싱 단계들에 필요하다. 온전한 신호 서열이 존재하는 경우, 발현된 단백질은 조면 소포체(rough endoplasmic reticulum)의 루멘(lumen)으로 들어간 후 골기체를 통해 분비 소포체로 수송되고 최종적으로 세포 외부로 수송된다. 일반적으로, 분비 서열은 개시 코돈 바로 뒤에 위치하고 분비될 단백질의 아미노 말단에 위치하는 신호 펩티드를 코딩한다. 대부분의 경우, 신호 서열은 신호 펩티데이즈로 지칭되는 특정한 프로테이즈에 의해 절단된다. 바람직한 신호 서열은 바이러스, 포유동물 또는 효모 발현 벡터를 사용한 재조합 단백질 발현의 프로세싱 및 유출을 개선시킨다.
- <98> 한 실시양태에서, 천연 Tf 신호 서열을 사용하여 본 발명의 융합 단백질을 발현하고 분비시킬 수 있다. Tf 분자는 다양한 종류의 분비물, 예컨대, 혈액, 눈물 및 모유에 존재하기 때문에, 많은 다양한 Tf 신호 펩티드가 존재한다. 예를 들어, Tf 신호 펩티드는 혈청 Tf, Lf 또는 멜라노트랜스페린으로부터 유래할 수 있다. 천연 Tf 신호 펩티드도 다양한 종, 예컨대, 곤충, 포유동물, 어류, 개구리, 오리, 닭 또는 다른 종으로부터 유래할 수 있다. 바람직하게는, 신호 펩티드는 포유동물 Tf 분자로부터 유래한다. 보다 바람직하게는, 신호 펩티드는 인간 혈청 Tf로부터 유래한다. 다양한 포유동물 Tf 분자로부터의 신호 펩티드 서열은 미국 특허 공보 제 2006/0205037호에 기재되어 있다.
- <99> Tf 유래의 신호 서열을 사용하여 이종 단백질을 분비할 수 있는데, 예를 들어, Tf 신호 서열에 대한 이종성을 나타내는 원하는 임의의 단백질을 Tf 신호를 사용하여 발현하고 분비할 수 있다. 구체적으로, Tf 신호 서열을 사용하여 재조합 효모로부터 단백질을 분비할 수 있다. 바람직하게는, 신호 펩티드는 인간 혈청 Tf(서열번호 19에 의해 코딩된 서열번호 18)로부터 유래한다. 다른 바람직한 신호 펩티드는 HSA/MF α-1(서열번호 41에 의해 코딩된 서열번호 40) 및 변경된 HSA/MF α-1(서열번호 43에 의해 코딩된 서열번호 42).

- <100> 신호 서열을 확실히 효율적으로 제거하기 위해, 일부 경우 신호 서열과 성숙 단백질 사이에 짧은 전구-펩티드 서열을 포함시키는 것이 바람직할 수 있고, 이때 상기 전구-펩티드의 C-말단 부분은 프로테이즈, 예컨대, 효모 kex2p 프로테이즈에 대한 인식 부위를 포함한다. 바람직하게는, 전구-펩티드 서열은 길이가 약 2 내지 12개 아미노산, 보다 바람직하게는 약 4 내지 8개 아미노산이다. 이러한 전구-펩티드의 예는 Arg-Ser-Leu-Asp-Lys-Arg(서열번호 113), Arg-Ser-Leu-Asp-Arg-Arg(서열번호 114), Arg-Ser-Leu-Glu-Lys-Arg(서열번호 115) 및 Arg-Ser-Leu-Glu-Arg-Arg(서열번호 116)이다.
- <101> 발현 및 클로닝 벡터는 전형적으로 숙주 유기체에 의해 인식되고 엑센딘-4/Tf 융합 단백질을 코딩하는 분자에 작동가능하게 연결된 프로모터를 포함할 것이다. 프로모터는 구조 유전자의 전사를 조절하는, 상기 구조 유전자의 개시 코돈의 업스트림(즉, 5') (일반적으로 약 100 내지 1000 bp 이내)에 위치한 비전사 서열이다. 프로모터는 보편적으로 2가지 부류, 즉 유도성 프로모터 및 기본구성성 프로모터 중 하나로 분류된다. 유도성 프로모터는 배양 조건의 일부 변화, 예컨대, 영양분의 존재 또는 부재 또는 온도의 변화에 반응하여 그 자신의 조절 하에 있는 DNA로부터 증가한 수준의 전사를 개시한다. 이에 반해, 기본구성성 프로모터는 연속적인 유전자 생성물의 생성을 개시한다. 즉, 유전자 발현에 대한 조절이 거의 없거나 전혀 없다. 다양한 잠재적 숙주 세포에 의해 인식되는 많은 수의 프로모터가 잘 공지되어 있다. 적절한 프로모터는 제한 효소 분해에 의해 공급원 DNA로부터 프로모터를 제거하고 원하는 프로모터 서열을 벡터 내로 삽입함으로써 엑센딘-4/Tf 융합 단백질을 코딩하는 DNA에 작동가능하게 연결된다. 천연 엑센딘-4 또는 Tf 프로모터 서열을 사용하여 엑센딘-4/Tf 융합 단백질 핵산 분자의 증폭 및/또는 발현을 유도할 수 있다. 그러나, 이종 프로모터는 이것이 천연 프로모터에 비해 발현될 단백질의 더 많은 전사 및 더 높은 수율을 허용하는 경우 및 사용을 위해 선택된 숙주 세포 시스템과 상용가능한 경우 바람직하다.
- <102> 또한, 효모 숙주와 함께 사용하기에 적합한 프로모터는 당업계에 잘 공지되어 있고 하기에 더 논의되어 있다. 효모 인핸서를 효모 프로모터와 함께 사용하는 것이 유리하다. 포유동물 숙주 세포와 함께 사용하기에 적합한 프로모터는 잘 공지되어 있고 바이러스, 예컨대, 폴리오마 바이러스, 조류폭스 바이러스, 아데노바이러스(예컨대, 아데노바이러스 2), 소류 파필로마 바이러스, 양서류 육종 바이러스, 사이토메갈로바이러스, 레트로바이러스, B형 간염 바이러스 및 가장 바람직하게는 원숭이 바이러스 40(SV40)의 게놈으로부터 수득된 것을 포함하나 이들로 한정되지 않는다. 다른 적합한 포유동물 프로모터는 이종 포유동물 프로모터, 예를 들어, 열-충격 프로모터 및 액틴 프로모터를 포함한다.
- <103> 원핵 숙주와 함께 사용하기에 적합한 프로모터는 베타-락테메이즈 및 락토스 프로모터 시스템; 이. 콜라이 T7 유도성 RNA 중합효소; 알칼리성 포스파테이즈; 트립토판(trp) 프로모터 시스템; 및 tac 프로모터와 같은 하이브리드 프로모터를 포함한다. 다른 공지된 박테리아 프로모터도 적합하다. 이의 서열은 공개되어 있기 때문에 당업자는 임의의 유용한 제한 부위를 공급할 필요가 있을 때 링커 또는 어댑터를 사용하여 상기 서열을 원하는 DNA 서열에 라이게이션시킬 수 있다.
- <104> 엑센딘-4/Tf 융합 단백질의 발현을 조절하는 데 있어서 흥미로울 수 있는 추가 프로모터는 SV40 초기 프로모터 영역(Bemoist and Chambon, Nature 290:304-10, 1981); CMV 프로모터; 라우스 육종 바이러스의 3' 장 말단 반복부(3' long terminal repeat)에 포함된 프로모터(Yamamoto et al, Cell 22:787-97, 1980); 헤페스 티미딘 카이네이즈 프로모터(Wagner et al., Proc. Natl. Acad. Sci. U.S.A. 78:1444-45, 1981); 메탈로티오닌 유전자의 조절 서열(Brinster et al., Nature 296:39-42, 1982); 원핵 발현 벡터, 예컨대, 베타-락테메이즈 프로모터 (Villa-Kamaroff et al., Proc. Natl. Acad. Sci. U.S.A. 75:3727-31, 1978); 또는 tac 프로모터(DeBoer et al., Proc. Natl. Acad. Sci. U.S.A., 80:21-25, 1983)를 포함하나 이들로 한정되지 않는다.
- <105> 인핸서 서열을 벡터 내로 삽입시켜 고등 진핵생물에서 엑센딘-4/Tf 융합 단백질을 코딩하는 DNA의 전사를 증가시킬 수 있다. 인핸서는 프로모터에 작용하여 전사를 증가시키는, 통상적으로 그 길이가 약 10 내지 300 bp인 DNA의 시스-작용 요소이다. 인핸서는 전사 단위체에 대해 5' 및 3'에서 발견된다. 포유동물 유전자(예를 들어, 글로빈, 엘라스테이즈, 알부민, 알파-태아단백질 및 인슐린)로부터 이용가능한 여러 인핸서 서열이 공지되어 있다. 그러나, 전형적으로 바이러스로부터의 인핸서가 사용될 것이다. SV40 인핸서, 사이토메갈로바이러스 초기 프로모터 인핸서, 폴리오마 인핸서 및 아데노바이러스 인핸서는 진핵 프로모터의 활성화를 위한 예시적 인핸서 요소이다. 인핸서는 엑센딘-4/Tf 융합 단백질 코딩 핵산 부위에 대해 5' 또는 3' 위치에서 벡터 내로 스플라이싱될 수 있지만, 인핸서는 전형적으로 프로모터에 대해 5' 위치에 존재한다.
- <106> 발현 벡터는 출발 벡터, 예컨대, 시판되는 벡터로부터 구축될 수 있다. 이러한 벡터는 원하는 플랭킹 서열의 전부를 포함할 수 있거나 포함하지 않을 수 있다. 본원에 기재된 플랭킹 서열 중 하나 이상이 벡터 내에 이미

존재하지 않는 경우, 플랭킹 서열은 개별적으로 얻어 벡터 내로 라이게이션될 수 있다. 플랭킹 서열 각각을 얻는 데 사용되는 방법은 당업자에게 잘 공지되어 있다.

<107> 본 발명에서 사용하기에 적합한 효모 벡터는 예를 들어, 미국 특허 제6,291,212호에 기재되어 있고 YRp7(Struhl et al., Proc. Natl. Acad. Sci. USA 76: 1035-1039, 1978), YEpl3(Broach et al., Gene 8: 121-133, 1979), pJDB249 및 pJDB219(Beggs, Nature 275:104-108, 1978), pPPC000 δ , pSeCHSA, pScNHSa, pC4 및 이들의 유도체를 포함한다. 유용한 플라스미드 벡터는 pRS403-406, pRS413-416, 및 시판되는 피키아 벡터(Stratagene Cloning Systems, 미국 캘리포니아주 라 졸라 소재)를 포함한다. 플라스미드 pRS403, pRS404, pRS405 및 pRS406은 효모 삽입 플라스미드(Yips)이고 효모 선별 마커 HIS3, TRP1, LEU2 및 URA3을 포함한다. 플라스미드 pRS413-41.6은 효모 센트로미어(Centromere) 플라스미드(YCps)이다.

<108> 이러한 벡터는 일반적으로 우성 표현형을 나타내는 임의의 수의 유전자들 중 하나일 수 있는 선별 마커를 포함할 것이고, 상기 우성 표현형에 대한 표현형적 분석이 형질전환체가 선별되게 할 수 있게 하기 위해 존재한다. 바람직한 선별 마커는 숙주 세포 영양요구성을 보충하는 마커, 항생제 내성을 제공하는 마커, 또는 세포가 특정한 탄소 공급원을 이용할 수 있게 하는 마커이고, LEU2(상기 Broach et al.), URA3(Botstein et al., Gene 8: 17, 1979), HIS3(상기 Struhl et al.) 또는 POT1(Kawasaki and Bell, 유럽 특허 제171,142호)을 포함한다. 다른 적절한 선별 마커는 효모 세포에 대해 클로람페니콜 내성을 부여하는 CAT 유전자를 포함한다. 효모에서 사용하기에 바람직한 프로모터는 효모 해당(glycolytic) 유전자(Hitzeman et al., J Biol. Chem. 225: 12073-12080, 1980; Alber and Kawasaki, J. Mol. Appl. Genet. 1 : 419-434, 1982; Kawasaki, 미국 특허 제 4,599,311호) 또는 알코올 데히드로게네이스 유전자(Young et al., in Genetic Engineering of Microorganisms for Chemicals, Hollaender et al., p. 355, Plenum, N.Y., 1982; Ammerer, Meth. Enzymol. 101: 192-201, 1983)로부터의 프로모터를 포함한다. 이와 관련하여, 특히 바람직한 프로모터는 TPH 프로모터(Kawasaki, 미국 특허 제4,599,311호) 및 ADH2-4^C 프로모터(미국 특허 제6,291,212호((Russell et al., Nature 304: 652-654, 1983) 참조)이다. 발현 단위체는 전사 종결요소를 포함할 수 있다. 바람직한 전사 종결요소는 TPH 종결요소(상기 Alber and Kawasaki)이다. 다른 바람직한 벡터 및 바람직한 요소, 예컨대, 효모 발현 시스템의 프로모터 및 종결요소는 유럽 특허 제0258067호, 제0286424호, 제0317254호, 제0387319호, 제0386222호, 제0424117호, 제0431880호, 제1002095호, 제0828759호, 제0764209호, 제0749478호 및 제0889949호; PCT 공보 제WO 00/44772호 및 제WO 94/04687호; 및 미국 특허 제5,739,007호, 제5,637,504호, 제5,302,697호, 제5,260,202호, 제5,667,986호, 제5,728,553호, 제5,783,423호, 제5,965,386호, 제6150,133호, 제6,379,924호 및 제5,714,377호에 개시되어 있다.

<109> 본 발명의 융합 단백질은 효모 외에 섬유상 진균, 예를 들어, 진균 아스퍼길러스(*Aspergillus*) 균주에서 발현될 수 있다. 유용한 프로모터의 예는 아스퍼길러스 니둘란스(*Aspergillus nidulans*) 해당 유전자로부터 유래한 프로모터, 예컨대, adh3 프로모터(McKnight et al., EMBO J. 4: 2093-2099, 1985) 및 tpi 프로모터를 포함한다. 적합한 종결요소의 한 예는 adh3 종결요소이다(상기 McKnight et al.). 이러한 요소를 사용하는 발현 단위체를 예를 들어, 아스퍼길러스의 염색체 DNA 내로 삽입될 수 있는 벡터 내로 클로닝할 수 있다.

<110> 다른 벡터는 박테리아, 곤충 및 포유동물 숙주 세포와 상용가능한 벡터이다. 이러한 벡터는 특히, pCRII, pCR3 및 pcDNA3.1(Invitrogen, 미국 캘리포니아주 칼스배드 소재), pBSII(Stratagene), pET15(Novagen, 미국 위스콘신주 매디슨 소재), pGEX(Pharmacia Biotech, 미국 뉴저지주 피스카타웨이 소재), pEGFP-N2(Clontech, 미국 캘리포니아주 팔로 알토 소재), pETL(BlueBac11, Invitrogen), pDSR-알파(PCT 공보 제WO 90/14363호) 및 pFastBacDual(Gibco-BRL, 미국 뉴욕주 그랜드 아일랜드 소재)을 포함한다.

<111> 추가 적합한 벡터는 코스미드, 플라스미드 변경된 바이러스를 포함하나 이들로 한정되지 않고, 벡터 시스템은 선택된 숙주 세포와 상용가능해야 함을 인식할 것이다. 이러한 벡터는 플라스미드, 예컨대, 블루스크립트(Bluescript[®]) 플라스미드 유도체(고카피수 ColE1-기재 파지미드, Stratagene), Taq-중폭 PCR 생성물을 클로닝하기 위해 디자인된 PCR 클로닝 플라스미드(예를 들어, TOPO[®] TA 클로닝[®] 키트, PCR2.1[®] 플라스미드 유도체, Invitrogen) 및 포유동물, 효모 또는 바이러스 벡터, 예컨대, 마콜로바이러스 발현 시스템(pBacPAK 플라스미드 유도체, Clontech)를 포함하나 이들로 한정되지 않는다.

<112> 발현 벡터는 원하는 코딩 서열의 다운스트림에 위치한 폴리아데닐화 신호를 포함한다. 폴리아데닐화 신호는 SV40으로부터의 초기 또는 후기 폴리아데닐화 신호(상기 Kaufman and Sharp), 아테노바이러스 5 E1B 영역으로부터의 폴리아데닐화 신호 및 인간 성장 호르몬 유전자 종결요소를 포함한다(DeNoto et al., Nucl. Acid Res. 9:

3719-3730, 1981). 특히 바람직한 폴리아데닐화 신호는 V_H 유전자 종결요소(미국 특허 제6,291,212호 참조)이다. 발현 벡터는 프로모터와 RNA 스플라이스 부위 사이에 위치하는 비코딩 바이러스 리더 서열, 예컨대, 아데노바이러스 2 삼원 리더를 포함할 수 있다. 바람직한 벡터는 인헨서 서열, 예컨대, SV40 인헨서(미국 특허 제 6,291,212호 참조) 및 마우스 인헨서(Gillies, Cell 33: 717-728, 1983)를 포함할 수도 있다. 발현 벡터는 아데노바이러스 VA RNA를 코딩하는 서열도 포함할 수 있다.

<113> 벡터를 제작하고 엑센딘-4/Tf 융합 단백질을 코딩하는 핵산 분자를 벡터의 적절한 부위 내로 삽입시킨 후, 완성된 벡터를 증폭 및/또는 폴리펩티드 발현에 적합한 숙주 세포 내로 삽입시킬 수 있다. 엑센딘-4/Tf 융합 단백질에 대한 발현 벡터를 선택된 숙주 세포 내로 형질전환시키는 것은 형질감, 감염, 전기천공, 미세주사, 리포펙션, DEAE-텍스트란 방법 또는 다른 공지된 기법과 같은 방법을 포함하는 잘 공지된 방법으로 달성할 수 있다. 선택되는 방법은 부분적으로 사용되는 숙주 세포의 종류의 함수일 것이다. 이 방법 및 다른 적절한 방법은 당업자에게 잘 공지되어 있고 예를 들어, 문헌[Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, 1989]에 기재되어 있다. 본 발명의 융합 단백질을 포함하는 클로닝된 DNA 서열은 예를 들어, 인산칼슘-매개 형질감염(Wigler et al., Cell 14: 725, 1978; Corsaro and Pearson, Somatic Cell Genetics 7: 603, 1981; Graham and Van der Eb, Virology 52: 456, 1973)에 의해 배양된 포유동물 세포 내로 도입될 수 있다. 클로닝된 DNA 서열을 포유동물 세포 내로 도입하기 위한 다른 기법, 예컨대, 전기천공(Neumann et al., EMBO J. 1: 841-845, 1982) 또는 리포펙션도 이용될 수 있다. 클로닝된 DNA가 도입되어 있는 세포를 확인하기 위해, 선별 마커는 일반적으로 원하는 유전자 또는 cDNA와 함께 세포 내로 도입된다. 배양된 포유동물 세포에서 사용하기에 바람직한 선별 마커는 약물, 예컨대, 네오마이신, 하이그로마이신 및 메토포렉세이트에 대한 내성을 부여하는 유전자를 포함한다. 선별 마커는 증폭가능한 선별 마커일 수 있다. 바람직한 증폭가능한 선별 마커는 DHFR 유전자이다. 특히 바람직한 증폭 마커는 DHFR^r(미국 특허 제 6,291,212호 참조) cDNA(Simonsen and Levinson, Proc. Natl. Acad. Sci. USA 80: 2495-2499, 1983)이다. 선별 마커는 Thilly(Mammalian Cell Technology, Butterworth Publishers, 미국 매사추세츠주 스톤햄 소재)에 의해 평가되어 있고 선별 마커의 선택은 당업계의 통상의 기술 수준 내에 있다.

<114> 숙주 세포

<115> 본 발명은 본 발명의 엑센딘-4/Tf 융합 단백질을 발현하도록 형질전환된 세포, 바람직하게는 효모 세포를 포함한다. 형질전환된 숙주 세포 자체 외에, 본 발명은 영양 배지 중의 상기 세포 배양물, 바람직하게는 단일클론(클론적으로 균질한) 배양물, 또는 단일클론 배양물로부터 유래된 배양물도 포함한다. 폴리펩티드가 분비되는 경우, 배지는 세포와 함께 또는 세포 없이(세포가 여과되거나 원심분리에 의해 제거되는 경우) 폴리펩티드를 함유할 것이다.

<116> 본 발명의 엑센딘-4/Tf 융합 단백질을 생성하는 데 특히 유용한 숙주 세포는 메틸친화성 효모인 피키아 파스토리스(Steinlein et al., Protein Express. Purif. 6:619-624, 1995)이다. 피키아 파스토리스는 그의 알코올 옥시테이즈 프로모터가 단리되어 클로닝되었기 때문에 외래 단백질의 생성에 뛰어난 숙주가 되도록 개발되어 왔고, 그의 형질전환체는 1985년 처음 보고되었다. 파키아 파스토리스는 당의 부재 하에 탄소 공급원으로서 메탄올을 사용할 수 있다. 파키아 파스토리스 발현 시스템은 메탄올의 대사에서 첫 번째 단계를 촉매진하는 촉매 효소인 알코올 옥시테이즈를 발현하는 유전자를 조절하는 메탄올-유도 알코올 옥시테이즈(AOX1) 프로모터를 사용할 수 있다. 이 프로모터는 특징규명되어 일련의 피키아 파스토리스 발현 벡터 내로 도입된다. 피키아 파스토리스에서 생성된 단백질은 전형적으로 정확히 폴딩되고 배지 내로 분비되기 때문에, 유전공학적으로 개조된 피키아 파스토리스의 발효는 이 콜라이 발현 시스템에 대한 뛰어난 대안을 제공한다. 과상풍 독소 단편, 보르다텔라 페르투스 퍼택틴(*Bordatella pertussis pertactin*), 인간 혈청 알부민 및 라이소자임을 비롯한 다수의 단백질을 이 시스템을 사용하여 제조하였다.

<117> 효모 사카로마이세스 세레비지에(*Saccharomyces cerevisiae*) 균주가 또 다른 바람직한 숙주이다. 바람직한 실시양태에서, 효모 세포, 보다 구체적으로 당단백질의 아스파라긴-관련 글리코실화에 필요한 유전자에서 유전적 결함을 가진 사카로마이세스 세레비지에 숙주 세포가 사용된다. 이러한 결함을 가진 사카로마이세스 세레비지에 숙주 세포는 많은 이용가능한 효모 균주가 글리코실화 또는 과다만노실화를 방지하거나 감소시키도록 변경되었을지라도 표준 돌연변이 및 선별 기법을 이용하여 제조할 수 있다. 문헌[Ballou et al., J. Biol. Chem. 255: 5986-5991, 1980]에는 아스파라긴-관련 글리코실화에 영향을 미치는 유전자에서 결함을 가진 만도단백질 생합성 돌연변이체의 단리가 기재되어 있다. 문헌[Gentsch and Tanner, Glycobiology 7:481-486, 1997]에는 효모 내의 단백질의 O-글리코실화에서 첫 번째 단계를 담당하는 효소를 코딩하는 6종 이상의 유전자(PMT1 내지

PMT6)로 이루어진 패밀리가 기재되어 있다. 이 유전자들 중 하나 이상에서 결함을 가진 돌연변이체는 감소한 O-결합 글리코실화 및/또는 N-결합 글리코실화의 변경된 특이성을 보인다.

- <118> 한 실시양태에서, 숙주는 PCT 공보 제WO 05/061718호에 기재된 사카로마이세스 세레비지에 균주이다. 예를 들어, 숙주는 녹아웃된 PDH 또는 다른 샤페론의 숙주 버전을 가진 균주 내에서 한 카피의 PDH 유전자 또는 임의의 다른 샤페론 유전자를 보유하는 pSAC35-기재 플라스미드를 함유할 수 있다. 이러한 제작물은 증가한 안정성을 부여한다.
- <119> 이중 단백질의 생성을 최적화하기 위해, 숙주 균주가 단백질 분해 활성을 감소시키는 돌연변이, 예컨대, 사카로마이세스 세레비지에 pep4 돌연변이(Jones, Genetics 85: 23-33, 1977)를 보유하는 것도 바람직하다. 다른 프로테이즈 코딩 영역에서 돌연변이를 보유하는 숙주 균주는 본 발명의 엑센딘-4/Tf 융합 단백질을 다량으로 생성하는 데 특히 유용하다.
- <120> 숙주 세포는 적절한 조건 하에 배양될 때 엑센딘-4/Tf 융합 단백질을 합성하고, 이 엑센딘-4/Tf 융합 단백질은 추후에 배양 배지로부터 모아질 수 있거나(숙주 세포가 상기 단백질을 배지 내로 분비하는 경우) 상기 단백질을 생산하는 숙주 세포로부터 직접 모아질 수 있다(상기 단백질이 분비되지 않는 경우). 적절한 숙주 세포의 선택은 다양한 인자, 예컨대, 원하는 발현 수준, 활성에 바람직한 또는 필요한 폴리펩티드 변경(예컨대, 글리코실화 또는 인산화) 및 생물학적 활성 분자로의 폴딩(folding)의 용이함에 달려 있을 것이다.
- <121> 다른 숙주 세포는 원핵 숙주 세포(예컨대, 이. 콜라이) 또는 진핵 숙주 세포(예컨대, 곤충 또는 척추 세포)일 수 있다. 다수의 적절한 숙주 세포는 당업계에 공지되어 있고 어메리칸 타입 컬처 콜렉션(ATCC)(미국 버지니아주 마나사스 소재)으로부터 입수할 수 있다. 그 예로는 포유동물 세포, 예컨대, 차이니스 햄스터 난소 세포(CHO), CHO DHFR(-) 세포(Urlaub et al., Proc. Natl. Acad. Sci. U.S.A. 97:4216-20, 1980), 인간 배아 신장(HEK) 293 또는 293T 세포, 또는 3T3 세포를 들 수 있으나 이들로 한정되지 않는다. 형질전환, 배양, 증폭, 스크리닝, 생성물 생산 및 정제에 적합한 포유동물 숙주 세포 및 방법의 선택은 당업계에 공지되어 있다. 다른 적합한 포유동물 세포주는 원숭이 COS-1 및 COS-7 세포주, 및 CV-1 세포주이다. 다른 예시적 포유동물 숙주 세포는 형질전환된 세포주를 비롯한 영장류 세포주 및 설치류 세포주를 포함한다. 정상적인 이배체 세포, 일차 조직의 시험관내 배양물로부터 유래한 세포 균주, 및 일차 체외이식조직도 적합하다. 후보 세포는 선별 유전자에서 유전형적으로 결핍되어 있을 수 있거나 우성적으로 작용하는 선별 유전자를 함유할 수 있다. 다른 적합한 포유동물 세포주는 마우스 신경모세포종 N2A 세포, HeLa, 마우스 L-929 세포, 스위스 Balb-c 또는 NIH 마우스로부터 유래된 3T3 세포주, BHK 또는 HaK 햄스터 세포주를 포함하나 이들로 한정되지 않는다. 이 세포주 각각은 단백질 발현 분야의 숙련된 자에게 공지되어 있고 입수될 수 있다.
- <122> 유사하게, 박테리아 세포도 적합한 숙주 세포로서 유용하다. 예를 들어, 이. 콜라이의 다양한 균주(예를 들어, HB101, DH5 α , DH10, 및 MC1061)는 생물공학 분야에서 숙주 세포로서 잘 공지되어 있다. 사카로마이세스 서브틸리스(*Saccharomyces subtilis*), 슈도모나스(*Pseudomonas*) 종, 다른 바실러스(*Bacillus*) 종 및 스트렙토마이세스(*Streptomyces*) 종도 사용될 수 있다.
- <123> 또한, 원하는 경우, 곤충 세포 시스템을 엑센딘-4/Tf 융합 단백질의 발현을 위해 이용할 수 있다. 이러한 시스템은 예를 들어, 문헌[Kitts et al., Biotechniques 14:810-17, 1993; Lucklow, Curr. Opin. Biotechnol. 4:564-72, 1993; and Lucklow et al., J. Virol., 67:4566-79, 1993]에 기재되어 있다. 바람직한 곤충 세포는 Sf-9 및 Hi5(Invitrogen)이다.
- <124> 엑센딘-4/Tf 융합 단백질 생산
- <125> 본 발명의 DNA 제작물을 함유하는 숙주 세포는 적절한 성장 배지 중에서 성장한다. 본원에서 사용된 바와 같이, 용어 "적절한 성장 배지"는 세포의 성장에 필요한 영양분을 함유하는 배지를 의미한다. 영양분은 탄소 공급원, 질소 공급원, 필수 아미노산, 비타민, 미네랄 및 성장 인자를 포함할 수 있다. 일반적으로, 성장 배지는 예를 들어, DNA 제작물 상의 선별 마커 또는 DNA 제작물과 함께 동시 형질감염된 선별 마커에 의해 보충되는 필수 영양분에서의 결핍 또는 약물 선별에 의해 DNA 제작물을 함유하는 세포를 선별할 것이다. 예를 들어, 효모 세포는 바람직하게는 탄소 공급원, 예컨대, 수크로스, 비-아미노산 질소 공급원, 무기 염, 비타민 및 필수 아미노산 보충물을 포함하는 화학적으로 정의된 배지 중에서 성장한다. 배지의 pH는 2보다 크고 8보다 작은 pH, 바람직하게는 pH 5.5 내지 6.5에서 유지된다. 안정한 pH를 유지하는 방법은 완충 및 일정한 pH 조절을 포함한다. 바람직한 완충제는 석신산 및 비스-트리스(Sigma Chemical Co., 미국 미조리주 세인트 루이스 소재)를 포함한다. 아스파라긴-관련 글리코실화에 필요한 유전자에서 결함을 가진 효모 세포는 바람직하게는 삼투압 안

정화제를 함유하는 배지 중에서 성장한다. 바람직한 삼투압 안정화제는 0.1 내지 1.5 M, 바람직하게는 0.5 또는 1.0 M의 농도로 배지 내로 보충된 소르비톨이다.

- <126> 이. 콜라이 세포를 배양하기에 적합한 배지는 예를 들어, 루리아 브로쓰(LB) 및/또는 테리픽 브로쓰(TB)를 포함한다. 진핵 세포를 배양하기에 적합한 배지는 로스웰 파크 메모리얼 인스티튜트 배지 1640(RPMI 1640), 미니멀 에센셜 배지(MEM) 및/또는 돌베코스 모디파이드 이글 배지(DMEM)을 포함하고, 이들은 모두 배양될 구체적인 세포주에 필요한 혈청 및/또는 성장 인자로 보충될 수 있다. 곤충 배양에 적합한 배지는 필요에 따라 이스트레이트, 락트알부민 가수분해물 및/또는 태아 소 혈청으로 보충된 그레이스 배지이다.
- <127> 전형적으로, 항생제 또는 형질감염된 또는 형질전환된 세포의 선별적 성장에 유용한 다른 화합물을 배지에 보충물로서 첨가한다. 사용될 화합물은 숙주 세포를 형질전환시키는 데 사용된 플라스미드 상에 존재하는 선별 마커 요소에 의해 결정될 것이다. 예를 들어, 선별 마커 요소가 가나마이신 내성인 경우, 배양 배지에 첨가되는 화합물은 가나마이신일 것이다. 선별적 성장을 위한 다른 화합물은 앰피실린, 테트라사이클린 및 네오마이신을 포함한다.
- <128> 바큘로바이러스/곤충 세포 발현 시스템을 사용하여 본 발명의 변경된 Tf 융합 단백질을 생성할 수도 있다. BacPAK™ 바큘로바이러스 발현 시스템(BD Biosciences (Clontech))은 곤충 숙주 세포에서 고 농도로 재조합 단백질을 발현한다. 표적 유전자는 전달 벡터 내로 삽입되고, 이 전달 벡터는 선형화된 BacPAK6 바이러스 DNA와 함께 숙주 세포 내로 동시에 형질감염된다. BacPAK6 DNA는 바큘로바이러스 게놈의 필수 요소를 갖고 있지 않다. 상기 DNA가 상기 벡터와 재조합되는 경우, 상기 필수 요소는 복원되고 표적 유전자는 바큘로바이러스 게놈으로 전달된다. 재조합 후, 소수의 바이러스 플라크를 픽킹하여 정제하고, 재조합 표현형을 확인한다. 그 다음, 새로이 단리된 재조합 바이러스를 증폭시켜 곤충 세포 배양물을 감염시키는 데 사용함으로써 다량의 원하는 단백질을 생산할 수 있다.
- <129> 본 발명의 엑센딘-4/Tf 융합 단백질은 형질전환 식물 및 동물을 사용하여 생산할 수도 있다. 예를 들어, 양 및 염소는 그들의 모유에서 치료 단백질을 생산할 수 있다. 또한, 담배 식물은 그들의 잎에서 상기 단백질을 함유할 수 있다. 단백질의 형질전환 식물 및 동물 생산은 융합 단백질을 코딩하는 신규 유전자를 유기체 게놈 내로 부가하는 것을 포함한다. 상기 형질전환 유기체는 신규 단백질을 생산할 수 있을 뿐만 아니라 이 능력을 그의 자손에 전달할 수도 있다.
- <130> 숙주 세포에 의해 생산된 엑센딘-4/Tf 융합 단백질의 양은 당업계에 공지된 표준 방법을 이용하여 평가할 수 있다. 이러한 방법은 웨스턴 블롯 분석, SDS-폴리아크릴아마이드 겔 전기영동, 비-변성 겔 전기영동, 고성능 액체 크로마토그래피(HPLC) 분리, 면역침전 및/또는 활성 분석, 예컨대, DNA 결합 겔 이동 분석을 포함하나 이들로 한정되지 않는다.
- <131> 엑센딘-4/Tf 융합 단백질이 숙주 세포주로부터 분비되도록 디자인된 경우, 대부분의 폴리펩티드는 세포 배양 배지 중에서 발견될 수 있다. 그러나, 폴리펩티드가 숙주 세포로부터 분비되지 않는 경우, 상기 폴리펩티드는 세포질 및/또는 핵(진핵 숙주 세포) 또는 세포질(그람 음성 박테리아 숙주 세포)에 존재할 것이다.
- <132> 숙주 세포 세포질 및/또는 핵(진핵 숙주 세포) 또는 세포질(박테리아 숙주 세포)에 위치한 엑센딘-4/Tf 융합 단백질의 경우, (그람 음성 박테리아를 위한 봉입체(inclusion body)를 포함하는) 세포내 물질을 당업자에게 공지된 임의의 표준 기법을 이용하여 숙주 세포로부터 추출할 수 있다. 예를 들어, 프렌치(French) 프레스, 균질화 및/또는 초음파처리 후 원심분리를 이용하여 숙주 세포를 용해시켜 원형질막공간/세포질의 내용물을 방출시킬 수 있다.
- <133> 엑센딘-4/Tf 융합 단백질이 세포질 내에서 봉입체를 형성하는 경우, 봉입체는 종종 세포 내막 및/또는 세포 외막에 결합하여 원심분리 후 펠렛 물질에서 주로 발견될 것이다. 그 다음, 펠렛 물질을 극도의 pH에서 처리하거나 환원제, 예컨대, 알칼리성 pH의 디티오쓰레이톨 또는 산성 pH의 트리스 카복시에틸 포스핀의 존재 하에 무질서유발제, 예컨대, 세제, 구아니딘, 구아니딘 유도체, 우레아 또는 우레아 유도체로 처리하여 봉입체를 파괴시키거나 용해시킬 수 있다. 그 다음, 용해된 엑센딘-4/Tf 융합 단백질은 겔 전기영동, 면역침전 등을 이용하여 분석할 수 있다. 폴리펩티드를 단리하고자 하는 경우, 표준 방법, 예컨대, 본원 및 문헌[Marston et al., Meth. Enz. 182:264-75, 1990]에 기재된 방법을 이용하여 단리할 수 있다.
- <134> 엑센딘-4/Tf 융합 단백질의 발현 시 봉입체가 상당한 정도로 형성되지 않는 경우, 폴리펩티드는 세포 균질화물의 원심분리 후에 상청액에서 주로 발견될 것이다. 상기 폴리펩티드는 본원에 기재된 방법과 같은 방법을 이용하여 상청액으로부터 더 단리할 수 있다.

- <135> 폴리펩티드를 생산하기 위한 다수의 추가 방법이 당업계에 공지되어 있고, 이 방법들을 이용하여 엑센딘-4/Tf 융합 단백질을 생산할 수 있다. 예를 들어, mRNA와 이의 코딩된 펩티드의 융합 단백질의 생산이 기재되어 있는 문헌[Roberts et al., Proc. Natl. Acad. Sci. U.S.A., 94:12297-303, 1997]을 참조한다. 문헌[Roberts, Curr. Opin. Chem. Biol. 3:268-73, 1999]도 참조한다.
- <136> 펩티드 또는 폴리펩티드를 생산하는 방법은 미국 특허 제5,763,192호, 제5,814,476호, 제5,723,323호 및 제5,817,483호에도 기재되어 있다. 상기 방법은 추계적(stochastic) 유전자 또는 이의 단편을 생산한 후, 이 유전자를 상기 추계적 유전자에 의해 코딩된 하나 이상의 단백질을 생산하는 숙주 세포 내로 도입하는 단계를 포함한다. 그 다음, 상기 숙주 세포를 스크리닝하여 원하는 활성을 가진 펩티드 또는 폴리펩티드를 생산하는 클론을 확인한다. 재조합 펩티드 발현을 위한 다른 방법은 미국 특허 제6,103,495호, 제6,210,925호, 제6,627,438호 및 제6,737,250호에 개시되어 있다. 이 방법은 이. 콜라이 및 이. 콜라이의 일반적 분비 경로를 이용한다. 펩티드를 신호 서열에 융합함으로써 펩티드를 분비를 위한 표적물로 만든다.
- <137> 펩티드 또는 폴리펩티드의 또 다른 생산 방법은 PCT 공보 제WO 99/15650호에 기재되어 있다. 유전자 발현을 위한 유전자 발현의 무작위적 활성화로 지칭되는 상기 공개된 방법은 내생 유전자 발현의 활성화 또는 제자리 재조합 방법에 의한 유전자의 과다발현을 포함한다. 예를 들어, 내생 유전자의 발현은 비-상동성 또는 변칙적 재조합에 의해 유전자의 발현을 활성화시킬 수 있는 표적 세포 내로 조절 서열을 도입시킴으로써 활성화시키거나 증가시킨다. 먼저, 표적 DNA를 방사선조사에 노출시키고 유전적 프로모터를 삽입시킨다. 상기 프로모터는 최종적으로 유전자의 전방에 위치하여 유전자의 전사를 개시한다. 이로 인해 원하는 펩티드 또는 폴리펩티드가 발현된다.
- <138> 엑센딘-4/Tf 융합 단백질의 단리/정제
- <139> 분비된 생물학적 활성 엑센딘-4/Tf 융합 단백질은 이 생물학적 활성 융합 단백질의 분비를 허용하는 조건 하에 성장된 숙주 세포의 배지로부터 단리할 수 있다. 세포 물질은 배양 배지로부터 제거되고, 생물학적 활성 융합 단백질은 당업계에 공지된 단리 기법을 이용하여 단리한다. 적합한 단리 기법은 겔 여과, 이온 교환 크로마토그래피 및 친화성 크로마토그래피를 비롯한 다양한 크로마토그래피 방법에 의한 침전 및 분획화를 포함한다.
- <140> 특히 바람직한 정제 방법은 철 결합 또는 금속 킬레이팅 컬럼 상의 친화성 크로마토그래피, 또는 폴리펩티드 융합체의 Tf 또는 치료 단백질에 대해 유도된 항원을 사용한 면역친화성 크로마토그래피이다. 항원은 바람직하게는 고체 지지체 또는 기질에 고정화되거나 부착된다. 한 실시양태에서, 상기 기질은 CNBr에 의해 활성화된 세파로스(Pharmacia LKB Technologies, Inc., 미국 뉴저지주 피스카타웨이 소재)이다. 이 방법으로 결합이 일어나게 하는 조건 하에 배지를 항원/기질과 합한다. 결합체를 세척하여 결합되지 않은 물질을 제거할 수 있고, 결합체 형성에 불리한 조건을 이용하여 엑센딘-4/Tf 융합 단백질을 방출하거나 회수한다. 특히 유용한 회수 방법은 pH의 변화(이때, 고정된 항원은 제1 pH에서 엑센딘-4/Tf 융합 단백질에 대해 높은 친화성을 나타내고 제2 (보다 높거나 보다 낮은) pH에서 감소한 친화성을 나타냄), 일부 무질서유발제의 농도 변화, 또는 세척 사용을 포함한다.
- <141> 용액으로부터의 엑센딘-4/Tf 융합 단백질의 정제는 다양한 기법을 이용하여 달성할 수 있다. 폴리펩티드가 태그, 예컨대, 핵사His 9 또는 다른 작은 펩티드, 예컨대, FLAG(Eastman Kodak Co., 코넥티컷주 뉴 헤븐 소재) 또는 myc(Invitrogen)를 그의 카복실 또는 아미노-말단에서 보유하도록 합성된 경우, 상기 폴리펩티드는 상기 용액을 친화성 컬럼에 통과시킴으로써 1-단계 공정으로 정제할 수 있고, 이때 컬럼 매트릭스는 태그에 대한 높은 친화성을 나타낸다.
- <142> 예를 들어, 폴리히스티딘은 높은 친화성 및 특이성으로 니켈에 결합한다. 따라서, 니켈(예컨대, 퀴아젠[®] 니켈 컬럼)을 정제에 사용할 수 있다. 문헌[Current Protocols in Molecular Biology, § 10.11.8(상기)]을 참조한다.
- <143> 또한, 엑센딘-4/Tf 융합 단백질은 엑센딘-4/Tf 융합 단백질을 특이적으로 인식하여 결합할 수 있는 단일클론 항체의 사용을 통해 정제할 수 있다.
- <144> 엑센딘-4/Tf 융합 단백질이 부분적으로 또는 실질적으로 오염물질을 갖지 않도록 엑센딘-4/Tf 융합 단백질을 부분적으로 또는 완전히 정제하는 것이 바람직한 경우, 당업자에게 공지된 표준 방법을 이용할 수 있다. 이러한 방법은 전기영동 후 전기회수에 의한 분리, 다양한 종류의 크로마토그래피(친화성, 면역친화성, 분자체 및 이온 교환), HPLC 및 분취형 등전 포커싱("Isoprime" 기구/기법, Hoefer Scientific, 미국 캘리포니아주 샌 프란시스코 소재)을 포함하나 이들로 한정되지 않는다. 일부 경우, 2종 이상의 정제 기법을 조합하여 순도의 증가를

달성할 수 있다.

<145> 약학 조성물

<146> 본 발명의 엑센딘-4/Tf 융합 단백질은 일반적으로 약학 조성물의 형태로 투여될 것이다. 약학 조성물은 예를 들어, 경구 투여에 적합한 제형(예컨대, 정제, 캡슐제, 환제, 산제, 용액, 현탁액), 비경구 주사에 적합한 제형(예를 들어, 멸균 용액, 현탁액 또는 에멀전), 비강내 투여에 적합한 제형(예를 들어, 에어로졸 적하 등), 직장 투여에 적합한 제형(예를 들어, 좌제) 또는 경피에 적합한 제형(예를 들어, 패치)일 수 있다. 약학 조성물은 정확한 투여량의 단위 투여에 적합한 단위 투약 제형일 수 있다. 약학 조성물은 본 발명의 엑센딘-4/Tf 융합 단백질을 활성 성분으로서 포함할 것이고 통상적인 약학 담체를 포함할 수 있다. 또한, 약학 조성물은 다른 약학적 물질, 보조제 등을 포함할 수 있다.

<147> 생체활성 펩티드의 다양한 약학 조성물을 제조하는 방법이 제약 분야에 공지되어 있다. 예를 들어, 미국 특허 공보 제2005/0009748호(경구 투여의 경우); 및 미국 특허 공보 제2004/0157777호, 제2005/0002927호 및 제2005/0215475호(경점막 투여, 예를 들어, 비강내 또는 협측 투여)를 참조한다. 문헌[Remington: The Practice of Pharmacy, Lippincott Williams and Wilkins, Baltimore, MD, 20th ed., 2000]도 참조한다.

<148> 종래, 펩티드 및 단백질은 경구 투여된 경우 낮은 생체이용률 때문에 주사에 의해 투여되어 왔다. 이 약물들은 화학적 및 구조적으로 불안정하기 쉽고 위의 산성 조건뿐만 아니라 위 및 위장관 내의 효소에 의해 종종 분해된다. 이 전달 문제를 해결하기 위해, 특정한 경구 전달 기술, 예컨대, 소수성 골격 및 친수성 분자를 가진 중합체로 구성된 약물 운반체로서의 나노입자 내의 캡슐화, 미세입자 내의 캡슐화, 에멀전 형태의 리포솜 내로 삽입, 및 다른 분자에의 접합이 개발되었다. 이들 모든 기술은 본 발명의 융합 분자에서 이용될 수 있다.

<149> 나노입자의 예로는 키토산 및 카보폴로 코팅된 점막부착성 나노입자(Takeuchi et al., Adv. Drug Deliv. Rev. 47: 39-54, 2001), 및 하전된 조합 폴리에스터, 폴리(2-설포부틸-비닐 알코올) 및 폴리(D,L-락트산-코-글리콜산)를 함유하는 나노입자(Jung et al., Eur. J. Pharm. Biopharm. 50: 147-160, 2000)를 들 수 있다. 폴리-N-아이소프로필아크릴아마이드 영역 및 양이온성 폴리-비닐아민 기를 가진 표면 중합체를 함유하는 나노입자는 래트에게 경구 투여된 경우 연어 칼시토닌의 개선된 흡수를 보였다.

<150> 알기네이트 및 펙틴으로 구성되어 있고 폴리라이신으로 강화된 약물 전달 입자는 비교적 산 및 염기에 대한 내성을 나타내며 약물 운반체로서 사용될 수 있다. 이 입자는 생체부착, 증가된 흡수 및 지속된 방출이라는 이점을 겸비하고 있다(Liu et al., J. Pharm. Pharmacol. 51 : 141-149, 1999).

<151> 또한, 합성 소마토스타틴과 같은 펩티드의 N-말단 및 C-말단에 접합되어 양쪽성 계면활성제를 생성하는 리포아미노산 기 및 리포사카라이드 기는 생물학적 활성을 보유하는 조성물을 생성함이 밝혀져 있다(Toth et al., J. Med. Chem. 42(19):4010-4013, 1999).

<152> 다른 펩티드 전달 기법의 예로는 원하는 펩티드 및 니트로소-N-아세틸-D,L-페니실아민 및 카보폴 또는 타우로콜레이트 및 카보폴을 함유하는 카보폴-코팅 점막부착성 에멀전을 들 수 있다. 이들은 래트에게 경구 투여된 경우 혈청 칼슘 농도를 감소시키는 데 있어서 효과적임이 밝혀져 있다(Ogiso et al., Biol. Pharm. Bull. 24: 656-661, 2001). 포스파티딜콜린으로부터 유래한 포스파티딜에탄올을 사용하여 인슐린 담체로서 포스파티딜에탄올을 함유하는 리포솜을 제조하였다. 이 리포솜은 래트에게 경구 투여된 경우 활성을 나타냄이 밝혀져 있다(Kisel et al., Int. J. Pharm. 216: 105-114, 2001).

<153> 인슐린도 인슐린 및 프로테이즈 억제제, 예컨대, 아프로티닌 또는 바시트라신을 함유하는 폴리(비닐 알코올)-겔 구체 내에 제형화되어 왔다. 이 겔 구체의 당-강화 성질은 래트에서 입증되었고, 이때 인슐린은 하부 장에서 주로 방출된다(Kimura et al., Biol. Pharm. Bull. 19: 897-900, 1996).

<154> 또한, 인슐린의 경구 전달은 오일 층에서 계면활성제와 함께 분산되는 폴리(알킬 시아노아크릴레이트)로 만들어진 나노입자(Damge et al., J. Pharm. Sci. 86: 1403-1409, 1997), 및 키토산으로 코팅된 알긴산칼슘 비드(Onal et al., Artif. Cells Blood Substit. Immobil. Biotechnol. 30: 229-237, 2002)를 사용하여 연구하여 왔다.

<155> 다른 방법에서, 펩티드의 N-말단 및 C-말단을 폴리에틸렌 글리콜에 연결한 후 알킬 쇠에 연결하여 효소적 분해에 대한 개선된 내성 및 위장 벽을 통한 개선된 확산을 나타내는 접합체를 형성한다(www.nobexcorp.com).

<156> 바이오포터(BioPORTER[®])는 펩티드와 비공유적으로 상호작용하여 보호 코팅 또는 층을 생성하는 양이온성 지질

혼합물이다. 펩티드-지질 결합체는 세포의 원형질막에 융합할 수 있고 상기 펩티드는 세포 내로 이입된다.

- <157> 출발 물질로서 리포솜을 사용하는 방법에 있어서, 달팽이 모양의 입자가 약학 비히클로서 개발되었다. 펩티드는 주로 음으로 하전된 지질을 함유하는 리포솜 현탁액에 첨가된다. 칼슘의 첨가는 리포솜을 붕괴시키고 지질 이중층으로 구성된 큰 시트에 융합시키는데, 상기 시트는 나중에 자발적으로 둥글게 말리거나 달팽이 형태로 적층된다(미국 특허 제5,840,707호).
- <158> 뿐만 아니라, 본 발명은 액센딘-4/Tf 융합 단백질 제형의 폐 전달을 포함한다. 폐 전달은 다른 투여 경로에 의해 전달하기 어려운 거대분자의 전달에 특히 유망하다. 이러한 폐 전달은 폐에 전달된 약물이 폐포 영역을 통해 혈액 순환계 내로 직접 용이하게 흡수되기 때문에 전신 전달 및 폐 질환을 치료하기 위한 국소 전달 둘다에 효과적일 수 있다.
- <159> 본 발명은 다양한 증상 또는 질환을 치료하기 위한 경구 흡입(폐 전달)용 약물 분산물을 형성하기에 적합한 조성물을 제공한다. 융합 단백질 제형은 지질 분사제, 에어로졸-기체 계측 투약 흡입기(MDI's) 및 건조 산제 분산물 장치와 같은 다양한 수단에 의해 전달될 수 있다. 폐 전달용 조성물을 제제화함에 있어서, 표면 활성제 또는 계면활성제 및 벌크 담체를 포함하는 약학적으로 허용가능한 담체를 통상적으로 첨가하여 안정성, 분산성, 점도 및/또는 벌킹 특성을 제공하여 객체로의 조성물의 균일한 폐 전달을 증가시킨다.
- <160> 표면 활성제 또는 계면활성제는 점막 또는 라이닝(lining)을 통한 폴리펩티드의 흡수를 촉진한다. 유용한 표면 활성제 또는 계면활성제는 지방산 및 이의 염, 담즙산 염, 인지질 또는 알킬 사카라이드를 포함한다. 지방산 및 이의 염의 예로는 카프릴레이트(C₈), 카프레이트(C₁₀), 라우레이트(C₁₂) 및 미리스테이트(C₁₄)의 나트륨, 칼륨 및 라이신 염을 들 수 있다. 담즙산 염의 예로는 콜산, 케노데옥시콜산, 글리코콜산, 타우로콜산, 글리코케노데옥시콜산, 타우로케노데옥시콜산, 데옥시콜산, 글리코데옥시콜산, 타우로데옥시콜산, 리소콜산 및 우르소데옥시콜산을 들 수 있다.
- <161> 인지질의 예로는 단일 쇠 인지질, 예컨대, 라이소포스파티딜콜린, 라이소포스파티딜글리세롤, 라이소포스파티딜 에탄올아민, 라이소포스파티딜이노시톨 및 라이소포스파티딜세린, 또는 이중 쇠 인지질, 예컨대, 디아실포스파티딜콜린, 디아실포스파티딜글리세롤, 디아실포스파티딜에탄올아민, 디아실포스파티딜이노시톨 및 디아실포스파티딜세린을 들 수 있다. 알킬 사카라이드의 예로는 알킬 글루코사이드 또는 알킬 말토사이드, 예컨대, 데실 글루코사이드 및 도데실 말토사이드를 들 수 있다.
- <162> 담체로서 유용한 약학적 부형제는 안정화제, 예컨대, HSA, 벌킹제, 예컨대, 탄수화물, 아미노산 및 폴리펩티드; pH 조절제 또는 완충제; 및 염, 예컨대, 염화나트륨을 포함한다. 이 담체들은 결정 형태 또는 비결정 형태일 수 있거나 결정 형태와 비결정 형태의 혼합물일 수 있다.
- <163> 벌킹제로서 사용되는 탄수화물의 예로는 모노사카라이드, 예컨대, 갈락토스, D-만노스 및 소르보스; 디아사카라이드, 예컨대, 락토스 및 트레할로스; 사이클로덱스트린, 예컨대, 2-하이드록시프로필-베타-사이클로덱스트린; 및 폴리사카라이드, 예컨대, 라피노스, 말토덱스트린 및 에스트란, 알디톨, 예컨대, 만니톨 및 자일리톨을 들 수 있다. 벌킹제로서 사용되는 폴리펩티드의 예로는 아스파르탐을 들 수 있다. 아미노산은 알라닌 및 글리신을 포함하고 글리신이 바람직하다.
- <164> 조성물의 부성분인 첨가제는 분무 건조 과정 동안의 구조적 안정성 및 산제의 분산성 개선을 위해 포함될 수 있다. 이 첨가제는 소수성 아미노산, 예컨대, 트립토판, 티로신, 류신 및 페닐알라닌을 포함한다.
- <165> 적절한 pH 조절제 또는 완충제는 유기 산 및 염기로부터 제조된 유기 염, 예컨대, 시트르산나트륨 및 아스코르브산나트륨을 포함하고, 시트르산나트륨이 바람직하다.
- <166> 폐 전달용 GLP-1 수용체 작용제 융합 조성물은 단위 투여량으로 포장될 수 있고, 이때 치료 유효량의 조성물이 단위 투여량 용기, 예컨대, 블리스터 팩 또는 젤라틴 캡슐 내에 존재한다. 블리스터 팩 또는 젤라틴 캡슐의 제조는 전형적으로 포장 분야에서 일반적으로 잘 공지되어 있는 방법에 의해 수행된다.
- <167> 미국 특허 제6,524,557호는 (a) HFA 추진제; (b) 추진제 중에 분산될 수 있는 약학적 활성 폴리펩티드; 및 (c) C₈ 내지 C₁₆ 지방산 또는 이의 염, 담즙산 염, 인지질 또는 알킬 사카라이드인 계면활성제를 포함하는 약학적 에어로졸 제형을 개시하고 있고, 이때 상기 계면활성제는 하부 호흡관 내에서의 폴리펩티드의 전신 흡수를 증가시킨다. 본 발명은 이러한 제형의 제조 방법 및 환자의 치료에 있어서 상기 제형의 용도도 제공한다.
- <168> 건조 산제 약물의 폐 전달을 위한 한 방법은 가압된 기체 공급원을 제공하기 위한 핸드 펌프를 구비한 핸드-헬

드 조작 장치를 이용하는 것이다. 가압된 기체는 산제 분산 장치, 예컨대, 벤투리 노즐을 통해 급격히 방출되고, 분산된 산제는 환자가 흡입할 수 있게 만들어진다.

- <169> 건조 산제 분산 장치는 여러 특허에 기재되어 있다. 미국 특허 제3,921,637호에는 산제 약물의 단일 캡슐을 통한 관통용 바늘을 구비한 수동 펌프가 기재되어 있다. 약물 치료에 있어서 다수의 용기 디스크 또는 스트립의 사용은 유럽 특허 제0 467 172호; PCT 공보 제WO 91/02558호 및 제WO 93/09832호; 및 미국 특허 제4,627,432호, 제4,811,731호, 제5,035,237호, 제5,048,514호, 제4,446,862호, 제5,048,514호 및 제4,446,862호에 기재되어 있다.
- <170> 단백질 치료제의 에어로졸화는 유럽 특허 제0 289 336호에 개시되어 있다. 치료 에어로졸 제형은 PCT 공보 제WO 90/09781호에 개시되어 있다.
- <171> 치료 방법
- <172> 본 발명의 엑센딘-4/Tf 융합 단백질은 본원에 기재된 질환 상태 또는 증상의 치료를 위한 다른 약학적 물질과 함께 사용될 수 있다. 따라서, 본 발명은 본 발명의 화합물을 다른 약학적 물질과 함께 투여하는 단계를 포함하는 치료 방법도 제공한다.
- <173> 본 발명의 방법 측면에서, 본 발명의 엑센딘-4/Tf 융합 단백질은 단독으로 또는 1종 이상의 다른 약학적 물질과 함께 당업계에 공지된 통상적인 말초 투여 방법 중 임의의 방법에 의해 객체에 개별적으로 또는 함께 말초 투여된다. 따라서, 엑센딘-4/Tf 융합 단백질 또는 조합물은 객체에 비경구(예를 들어, 정맥내, 복강내, 근육내 또는 피하), 비강내, 경구, 설하, 협측, 흡입(예를 들어, 에어로졸), 직장(예를 들어, 좌약제) 또는 경피 투여될 수 있다. 비경구 투여(예를 들어, 주사)가 바람직한 투여 방법이고, 피하 투여가 바람직한 비경구 투여 방법이다. 흡입에 의한 폐 전달도 바람직한 투여 방법이다.
- <174> 비경구 주사에 적합한 조성물은 일반적으로 약학적으로 허용가능한 멸균 수성 또는 비수성 용액, 분산액, 현탁액 또는 에멀전, 및 멸균 주사 용액 또는 분산액으로 재구성하기 위한 멸균 산제를 포함한다. 적합한 수성 및 비수성 담체 또는 희석제(용매 및 비히클을 포함함)의 예로는 물, 에탄올, 폴리올(프로필렌 글리콜, 폴리에틸렌 글리콜, 글리세롤 등), 이들의 적절한 혼합물, 올리브유와 같은 식물성 오일을 포함하는 트라이글리세라이드, 및 에틸 올레에이트와 같은 주사가 가능한 유기 에스터를 들 수 있다.
- <175> 이 비경구 주사용 조성물은 부형제, 예컨대, 보존제, 습윤제, 가용화제, 유화제 및 분산제도 함유할 수 있다. 상기 조성물의 미생물 오염은 다양한 항균제 및 항진균제, 예를 들어, 파라벤, 클로로부탄올, 페놀 및 소르브산을 사용하여 방지할 수 있다. 등장화제, 예를 들어, 당 및 염화나트륨을 포함하는 것도 바람직할 수 있다. 주사가 가능한 약학 조성물의 지연된 흡수는 흡수를 지연할 수 있는 물질, 예를 들어, 알루미늄 모노스테아레이트 및 젤라틴에 의해 달성될 수 있다.
- <176> 본 발명의 엑센딘-4/Tf 융합 단백질은 투여 방식, 객체의 연령 및 체중, 질환의 심각도, 치료할 증상 또는 장애, 및 투여될 엑센딘-4/Tf 융합 단백질의 약리 활성을 포함하는 다수의 인자에 따라 달라지는 투여량으로 객체에 투여될 것이다. 구체적인 환자에 대한 투여량 범위 및 최적 투여량의 결정은 당업계의 통상적인 기술 내에 있다.
- <177> 혈당을 감소시키는 치료를 위한 비경구 주사의 경우, 서열번호 23에 나타낸 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질은 단위 투약 당 약 0.5 내지 50 mg, 보다 바람직하게는 단위 투약 당 0.5 내지 20 mg의 투여량 수준으로 인간에게 투여될 수 있고, 이때 단위 투약은 주당 약 1회, 2주당 약 1회 또는 월당 약 1회 실시한다.
- <178> 체중을 감소시키는 치료를 위한 비경구 주사의 경우, 투여량 범위는 혈당 감소를 위한 경우보다 더 높을 수 있다. 따라서, 체중을 감소시키는 치료를 위한 비경구 투여의 경우, 서열번호 23에 나타낸 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질은 단위 투약 당 약 1 내지 100 mg의 투여량 수준에서 인간에게 투여될 수 있고, 이때 단위 투약은 주당 약 1회, 2주당 약 1회 또는 월당 약 1회 실시한다.
- <179> 본 발명은 인간 환자에서 II형 당뇨병을 치료하거나 혈당을 감소시키는 데 사용하기 위한 본 발명의 엑센딘-4/Tf 융합 단백질을 제공한다. 또한, 인간 환자에서 비만을 치료하거나 음식 섭취를 감소시키는 데 사용하기 위한 본 발명의 엑센딘-4/Tf 융합 단백질을 제공한다. 본 발명의 추가 측면은 인간 환자에서 II형 당뇨병을 치료하거나 혈당을 감소시키기 위한 약제의 제조에 있어서 본 발명의 엑센딘-4/Tf 융합 단백질의 용도를 제공한다. 본 발명의 또 다른 추가 측면은 인간 환자에서 비만을 치료하거나 음식 섭취를 감소시키기 위한 약

제의 제조에 있어서 본 발명의 엑센딘-4/Tf 융합 단백질의 용도를 제공한다. 본 발명의 방법 측면의 특징은 이 측면들 각각에 적용될 수 있다.

<180> 본 발명의 실시양태는 하기 실시예에 의해 예시될 것이다. 그러나, 본 발명의 실시양태는 이 실시예의 구체적인 세부적 사항으로 한정되지 않고, 이들의 다른 변경이 당업자에게 공지될 것이거나 본 개시내용 및 첨부된 청구범위에 비추어 당업자에게 자명할 것이다. 본원에 인용된 모든 참고문헌은 그들 전체 내용이 참고로 본원에 도입된다.

실시예

<181> 실시예 1: 엑센딘-4/Tf 융합 단백질의 제작

<182> **엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질**

<183> 부위 증첩 연장(SOE) PCR을 이용하여 엑센딘-4(1-39) DNA 서열(서열번호 20)을 pREX0549의 분비 신호 서열(nL)(서열번호 19)과 mTf 서열(서열번호 22) 사이에 삽입하였다. 기질로서 pREX0549를 사용하여 상기 서열을 삽입하기 위해 2종의 프라이머, 즉 P0702(서열번호 28) 및 P0703(서열번호 29)을 디자인하였다.

<184> 상기 DNA 서열은 효모 발현에 최적인 코돈을 사용하는 엑센딘-4 아미노산 서열(서열번호 30)의 역 번역으로 수득하였다. 먼저, *Afl*III 부위의 5' 프라이머, 즉 P0702를 가진 P0177(서열번호 31) 또는 *Bam*HI 부위의 3' 프라이머, 즉 P0703을 가진 P0014(서열번호 32)를 사용하여 2종의 PCR 생성물을 생성하였다.

<185> 이 반응으로부터 얻은 생성물을 겔 정제하고 제2 PCR 라운드에서 외부 프라이머, 즉 P0177 및 P0014만을 사용하여 연결하였다. 이 제2 반응으로부터 얻은 생성물을 겔 정제하고 제한 효소 *Afl*III 및 *Bam*HI으로 분해하였고, pREX0549도 마찬가지로 처리하였다. 이 반응으로부터 얻은 적절한 생성물을 서로 라이게이션시켜 pREX0561을 수득하였고, 이 pREX0561의 *Afl*III 부위와 *Bam*HI 부위 사이에 있는 DNA를 시퀀싱하여 엑센딘-4 서열의 정확한 삽입을 확인하였다. 제한 효소 *Not*I으로 절단하여 pREX0561로부터 발현 카세트를 회수하여, *Not*I으로 절단되고 소 소장 알칼리성 포스파테이즈로 처리된 pSAC35에 라이게이션시켜 pREX0589를 수득하였다.

<186> 전술한 절차와 동일한 절차를 이용하되 pREX0561을 기질로 사용하고 프라이머 P1810(서열번호 33) 및 P1811(서열번호 34)을 사용하여 SOE PCR을 수행함으로써 엑센딘-4 서열의 코딩된 C-말단과 코딩된 mTf 서열의 N-말단 사이에 링커 펩티드(PEAPTD)₂(SEQ ID NO.: 21)를 도입하였다.

<187> 이 PCR로부터 얻은 최종 생성물을 겔 정제하고 제한 효소 *Afl*III 및 *Bam*HI으로 분해하였고, pREX0549도 마찬가지로 처리하였다. 이 반응으로부터 얻은 적절한 생성물을 서로 라이게이션시켜 pREX0935를 수득하였고, 이 pREX0935의 *Afl*III 부위와 *Bam*HI 부위 사이에 있는 DNA를 시퀀싱하여 (PEAPTD)₂(서열번호5)를 코딩하는 서열의 정확한 삽입을 확인하였다. 제한 효소 *Not*I으로 절단하여 pREX0935로부터 발현 카세트를 회수하여, *Not*I으로 절단되고 소 소장 알칼리성 포스파테이즈로 처리된 pSAC35에 라이게이션시켜 pREX0936을 수득하였다. nL 리더 서열을 갖지 않는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 아미노산 서열은 본원에서 서열번호 23으로서 제공되어 있다. 서열번호 23을 코딩하는 핵산 서열은 본원에서 서열번호 24로서 제공되어 있다. nL 리더 서열을 갖는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 아미노산 서열은 본원에서 서열번호 25로서 제공되어 있다. 서열번호 25를 코딩하는 핵산 서열은 본원에서 서열번호 26으로서 제공되어 있다.

<188> **추가 엑센딘-4/Tf 제작물**

<189> 엑센딘-4는 GLP-1에 비해 C-말단에서 추가 9개의 아미노산을 가진다. 자유 펩티드 면에서, 이 추가 잔기들은 GLP-1 수용체에 대한 증가된 친화성 및 보다 높은 프로테이즈 내성을 수여하는 것으로 생각된다. 그러나, 이것은 펩티드의 면역원성에 일정한 정도로 기여할 수 있다. 잔기 32 내지 39를 코딩하는 DNA 서열 및 31 내지 39를 코딩하는 DNA 서열을 각각 결실시킴으로써 GLP-1에 대한 서열 상동성만을 나타내는, 즉 엑센딘-4(1-31) 및 엑센딘-4(1-30)를 제조하기 위해 전술한 바와 실질적으로 동일한 절차를 이용하여 2종의 추가 제작물을 제조하였다. 엑센딘-4(1-31)의 경우, 프라이머 P0904(서열번호 35) 및 P0941(서열번호 36)을 사용하였고 적절한 생성물을 라이게이션시켰다(pREX0629/pREX0658). 엑센딘-4(1-30)의 경우, 프라이머 P0942(서열번호 37) 및 P0943(서열번호 38)을 사용하였고 적절한 생성물을 라이게이션시켰다(pREX0630/pREX0659).

<190> 또한, 제작물은 엑센딘-4(1-39) 서열 및 대안적 링커, 예를 들어, (GGGGS)₃(서열번호 39), PEAPTD(pREX1005)(서

열번호 6) 또는 IgG 힌지(pREX0938) (서열번호 7 내지 16)을 사용하여 제조하였다.

<191> 다른 신호 서열을 가진 추가 엑센딘-4/Tf 제작물 - 상대적 생산성에 대한 영향

<192> 신호 서열 HSA/MF α-1(pREX 1354)(서열번호 40; 서열번호 41에 의해 코딩됨) 및 변경된 HSA/MF α-1(pREX 1345)(서열번호 42; 서열번호 43에 의해 코딩됨)에 연결된 엑센딘-4/mTf(서열번호 23; 서열번호 24에 의해 코딩됨)를 발현시키기 위한 제작물을 제조하였다. 3종의 상이한 신호 서열을 가진 엑센딘-4/mTf를 발현하는 효모 균주들의 생산성의 비교는 Tf 신호 서열(nL)/HSA/MF α-1/변경된 HSA/MF α-1에 대한 상대적 생산성 비가 1/1.75/1.32임을 보여주었다.

<193> 실시예 2: 엑센딘-4/Tf 융합 단백질의 효능 측정

<194> 샘플과의 항온처리 후, 래트 GLP-1 수용체(CHO-GLP-1R)로 형질감염된 CHO 세포에서 GLP-1 수용체-매개 리간드 결합의 결과로서 얻은 cAMP의 측정된 반응으로부터 효능을 계산하였다. 96-웰 조직 배양 플레이트에 CHO-GLP-1R 세포를 시딩하고 밤새 배양하였다. 다음 날, 세포를 크랩스-링거 완충제(KRB)로 린싱하고 포스포다이에스터 레이즈 억제제인 3-아이소부틸-1-메틸잔틴(IBMX, 2 mM)을 함유하는 KRB 중에서 항온처리하여 cAMP를 처리하는 세포내 효소를 억제하였다. 시험 화합물 및 대조군의 연속 희석물을 KRB/IBMX로 준비하고 세포 웰을 샘플 및 대조군과 항온처리하되 1개의 샘플 또는 대조군 당 3개의 세포 웰을 사용하였다. 항온처리 후, 경쟁-기초 형광 면역 분석(캐치포인트(CatchPoint)[®] cAMP 형광 분석 키트, Molecular Devices Corp., 미국 캘리포니아주 서니 베일 소재)을 이용하여 세포내 cAMP 수준의 증가를 측정하기 위해 개별 샘플 용해물을 분석하였다. GLP-1 수용체-매개 리간드 결합 후 세포 내의 cAMP 측정량을 이용하여 생체활성 및 상대적 효능을 측정하였다.

<195> 하기 표 1에 나타난 데이터는 엑센딘-4/Tf 융합체가 GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질보다 GLP-1 수용체를 활성화시키는 데 있어서 더 강력함을 보여준다. 각 융합 단백질의 mTf 부분은 서열번호 17에 나타난 아미노산 서열을 가진다.

표 1

GLP-1/mTf 및 엑센딘-4/Tf 융합 단백질의 효능

플라스미드	제작물			효능 (nM)
pREX0585	GLP-1(7-37;A8G,K34A)	(PEAPTD) ₂	mTf	1.3
pREX0659	엑센딘 -4(1-30)	(PEAPTD) ₂	mTf	0.85
pREX0658	엑센딘 -4(1-31)	(PEAPTD) ₂	mTf	0.33
pREX0936	엑센딘 -4(1-39)	(PEAPTD) ₂	mTf	0.16

<196> 실시예 3: 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 MMP-내성

<198> 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질(서열번호 23)을 시험관 내에서 매트릭스 메탈로프로테아제 I(MMP-1, 콜라게네이즈)에 의한 불활성화에 대한 내성에 대해 시험하였다. 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질 샘플 및 GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질 샘플을 37°C에서 제조합 MMP-1과 48시간 동안 항온처리한 후 활성에 대해 시험하였다. 도 1은 엑센딘-4/Tf 융합 단백질(도 1B)이 GLP-1/mTf 융합 단백질(도 1A)과 대조적으로 MMP-1에 의한 불활성화에 대한 내성을 가짐을 보여준다. 이러한 분해 차이는 분자의 활성 부분에서의 아미노산 서열의 높은 유사성에도 불구하고 일어난다.

<199> 실시예 4: 당뇨병 마우스에서 혈당에 대한 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 효과

<200> 당뇨병(db/db) 마우스에게 다양한 투여량의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질, GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질, 또는 엑센딘-4 펩티드(Bachem, 미국 펜실베이니아주 킹 오브 프루시아 소재)를 주사하고, 혈당측정기를 이용하여 혈액 샘플을 분석함으로써 혈당 농도를 모니터링하였다. 도 2에 보여진 바와 같이, 1.3 nmole/kg만큼 낮은 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 투여량은 피하 주사 후 3시간까지 상기 동물들에서 혈당을 상당히 감소시켰다. 당 농도는 모든 처리군에서 거의 정상화되었고 이 농도는 24시간 동안 지속되었다. 당 농도는 투여된 투여량에 따라 처리 후 48 내지 72시간 사이에 처리 전 농도까지 점차 증가하였다. 엑센딘-4 펩티드는 투여량과 관계없이 엑센딘-4(1-39)(PEAPTD)₂(서열

번호 5) mTf 융합 단백질만큼 많이 혈당을 낮추지 못하였고, 엑센딘-4의 당 농도 감소 효과는 약 12시간까지 완전히 소멸하였다. 이것은 엑센딘-4로 달성할 수 있는 혈당의 최대 감소가 약 37%라는 문헌 보고와 일치하고, 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질에서 관찰된 혈당 감소는 약 70%이었다. 혈당에 대한 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 효과는 동등한 양의 GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질보다 상당히 더 높았고 더 오래 지속되었다.

<201> 실시예 5: 래트 체중에 대한 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 효과

<202> 다양한 투여량의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질(서열번호 23) 및 엑센딘-4 펩티드를 스프라그 다울리 래트에게 피하 주사하였다. mTf 또는 생리식염수를 대조군으로서 사용하였다. (투약 당일 투약하기 전에) 래트의 체중을 매일 측정하였다. 동물들이 음식 및 물에 언제든지 자유롭게 접근할 수 있게 하였다.

<203> 도 3에 보여진 바와 같이, 10 내지 100 nmole/kg 투여량의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질로 처리된 동물은 첫 주사 후 체중 감량을 보였고, 체중 감량은 전체 투여 기간 동안 계속되었다. 5일째 날, 100 nmole/kg 투여량으로 처리된 동물들은 대조군에 비해 평균 75 g(17%)의 체중 감량을 보였고, 체중 감량은 음식 및 물 섭취에 있어서의 강화와 관련되어 있다. 급격한 체중 감량이 관찰되었기 때문에, 약물의 매일 투여를 5일째 날에 중단하였다. 5일이라는 투여 기간 후, 모든 동물들은 유사한 속도로 체중 증가를 보였다. 그러나, 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질 처리 군, 특히 높은 투여량이 처리된 군은 마지막 투여 후 20일째 날에 대조군보다 여전히 체중이 낮았다.

<204> 실시예 6: II형 당뇨병에서의 혈당 조절 및 체중 감소를 위한 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질에 대한 예측 투여량

<205> 공개된 데이터에 의거할 때, 엑세나티드(바이에타®)의 치료적 단위 투여량인 10 µg은 인간에서 200 pg/ml의 Cmax를 나타내었다. 엑세나티드와 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질(서열번호 23) 사이의 분자 크기 차이(4.2 kDa 대 80.5 kDa)는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 혈액 농도 약 3.8 ng/ml가 혈당 강하 효과 면에서 엑세나티드의 치료적 투여량 수준과 동등함을 보여준다. 크기 외에, 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질은 인간 GLP-1 수용체를 발현하는 CHO 세포의 시험관내 시험(도 4)에 의거할 때 엑센딘-4보다 약 5배 덜 강력하다. 따라서, 10 µg의 엑세나티드의 치료 활성화 유사한 치료 활성을 얻기 위해서는 약 20 ng/ml의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 순환 농도가 요구된다.

<206> 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질과 GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질은 크기 및 구조 둘다에서 유사하다. 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 분자량은 80.5 kDa이고 GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 분자량은 79.6 kDa이다. 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질은 GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질보다 약 4 내지 8배 더 강력하다. 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질 1 mg/kg을 시아노물구스 원숭이에게 정맥내 및 피하 투여한 후의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 평균 약동학적 파라미터는 하기 표 2에 기재되어 있다. GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질 0.6 mg/kg을 시아노물구스 원숭이에게 피하 투여한 후의 GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 평균 약동학적 파라미터는 하기 표 3에 기재되어 있다.

표 2

엑센딘-4(1-39) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질 1 mg/kg을 수컷 및 암컷 시아노몰구스 원숭이에게 정맥내 및 피하 투여한 후 엑센딘-4(1-39) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질에 대한 평균 약동학적 파라미터의 요약

파라미터	정맥내	피하
Cmax (ng/mL)	33,981 ± 14,826 (4)	5,236 ± 1,038 (4)
Tmax (h)	0.542 (4)	9.02 (4)
AUC(0-t) (h·ng/mL)	567,364 ± 68,102 (4)	278,067 ± 24,367 (4)
AUC(inf) (h·ng/mL)	572,314 ± 68,660 (4)	280,279 ± 29,261 (3)
λ z (h ⁻¹)	0.0313 ± 0.0137 (4)	0.0252 ± 0.0084 (3)
t½ (h)	25.5 ± 10.3 (4)	29.3 ± 08.2 (3)
CL (mL/min/kg)	1.77 ± 0.24 (4)	-
Vz (mL/kg)	65.9 ± 31.0 (4)	-
F (%)	-	49.0

표 3

GLP-1(7-37, A8G, K34A) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질 0.6 mg/kg을 수컷 및 암컷 시아노몰구스 원숭이에게 피하 투여한 후 GLP-1(7-37, A8G, K34A) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질에 대한 평균 약동학적 파라미터의 요약

파라미터	수컷	암컷
Cmax (ng/mL)	2,922 ± 1,530 (3)	3,173 ± 1,767 (3)
Tmax (h)	12.0 (3)	24.2 (3)
AUC(0-t) (h·ng/mL)	168,087 ± 58,749 (3)	165,474 ± 32,756 (3)
AUC(inf) (h·ng/mL)	171,963 ± 61,998 (3)	186,065 ± 140 (2)
λ z (h ⁻¹)	0.0216 ± 0.0042 (3)	0.0250 ± 0.0002 (2)
t½ (h)	32.9 ± 6.78 (3)	27.7 ± 0.23 (2)
CL (mL/min/kg)	-	-
Vz (mL/kg)	-	-
F (%)	-	-

별도의 실험에서, GLP-1(7-37, A8G, K34A) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질의 생체이용률은 시아노몰구스 원숭이에서 약 50%인 것으로 나타났다.

정맥내 투여 및 피하 투여 둘다에서 엑센딘-4(1-39) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질의 제거 반감기 (t_{1/2}), Tmax 및 생체이용률(F%)은 GLP-1(7-37, A8G, K34A) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질의 원숭이 약동학적 파라미터와 유사하였다.

GLP-1(7-37, A8G, K34A) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질을 사용한 이전 인간 실험은 투여량 30 µg/kg부터 900 µg/kg까지 약동학적 프로필이 선형이고 평균 Tmax가 48시간이며 평균 t_{1/2}이 약 50시간임을 보여주었다. Cmax는 300 µg/kg의 투여량(또는 30 mg/100 kg 환자)에서 758 ± 435 ng/ml이었고 900 µg/kg의 투여량(또는 90 mg/100 kg 환자)에서 1,609 ± 805 ng/ml이었다. 그러나, 상기 용합 단백질은 상기 투여량들 및 1800 µg/kg의 투여량에서 당뇨병 객체에서 혈당 농도에 대한 강력한 효과를 보이지 않았다.

상기 두 화합물 사이의 크기 및 구조의 유사성 및 원숭이에서의 유사한 임상전 약동학적 프로필 때문에, 인간에서의 엑센딘-4(1-39) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질의 약동학적 특성은 GLP-1(7-37, A8G, K34A) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질과 유사할 것으로 예측된다.

GLP-1(7-37, A8G, K34A) (PEAPTD)₂ (서열번호 5) mTf와의 유사성, GLP-1(7-37, A8G, K34A) (PEAPTD)₂ (서열번호 5) mTf에 비해 상대적으로 4 내지 8배 더 높은 엑센딘-4(1-39) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질의 시험관내 효능, 및 엑세나티드에 비해 상대적으로 5배 더 낮은 엑센딘-4(1-39) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질의 시험관내 효능에 의거할 때, 2 mg 투여량의 엑센딘-4(1-39) (PEAPTD)₂ (서열번호 5) mTf 용합 단백질이 당 강하 효과를 나타낼 수 있음은 놀라운 것이다. 또한, 주당 1회 투약 시 객체 당 10 mg의 투여량이 20 ng/ml의 정지-상태 Cmin을 달성하는 데 필요할 것이다. 따라서, 치료적 혈당 강하를 위한 엑센딘-4(1-39) (PEAPTD)₂ (서열번호

5) mTf 융합 단백질(서열번호 23)의 유효 투여량은 주당 1회 투약 시 단위 투약 당 0.5 내지 50 mg이다. 이러한 투여량은 2주당 1회 또는 월당 1회 투여될 수도 있다.

<214> 혈당에 대한 효과 외에, 10 nmole/kg 이상의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질은 마우스 및 래트에서 동물 체중의 감소와 상호관련되어 있었다. 단위 투여량, 즉 10 및 100 nmole/kg의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 마우스에게 투여한 지 24시간 후, 대조군 또는 1 nmole/kg의 엑센딘-4 처리 동물에서의 평균 체중 감소 1%에 비해 상기 마우스의 체중 감소는 평균 6% 및 14%이었다. 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 매일 투여도 100 nmole/kg의 투여량에서 래트에서 16%의 체중 감소를 나타내었다. 체중 감소는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질로 투여된 동물에서 관찰된 감소한 음식 섭취(GLP-1 수용체 활성화의 공지된 약리 효과임)에 기인할 수 있다. 입수가 가능한 데이터는 체중 감소에 필요한 투여량이 혈당 강하에 필요한 투여량보다 2 내지 3배 더 높음을 보여준다. 따라서, 체중 감소를 위한 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질(서열번호 23)의 유효량은 주당 1회 투약 시 단위 투약 당 1.0 내지 100 mg이다. 이러한 투여량은 2주당 1회 또는 월당 1회 투여될 수도 있다.

<215> 실시예 7: 흡입에 의한 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 전달

<216> 에어로테크 II 압축 공기 제트 분사기(CIS-US Inc., 미국 매사추세츠주 베드포드 소재)를 이용하여 에어로졸을 생성하고 1.58 cm 직경의 스테인레스 강철 에어로졸 전달 라인을 통해 24-포트 플로우 패스트 설치류 노출 시스템(IN-TOX, ABQ, 미국 뉴멕시코주 소재) 내로 도입하였다. 챔버의 유출 유동 속도는 약 11.5 ℓ/분이었다. 분사기 압력은 약 30 psi로 유지하였다.

<217> 상기 융합 단백질에 대한 에어로졸화의 효과를 시험하기 위해, 10 mM 히스티딘(pH 7.4) 및 100 mM NaCl 중의 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질 용액 10 mg/ml를 분사한 후, 에어로졸로부터의 농축액 5 ml를 8분간에 걸쳐 바이오샘플러에 모아 완전성 및 활성화에 대해 시험하였다. 에어로졸화 절차는 SDS-PAGE 및 SEC-HPLC에 의한 평가 시 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 구조에 검출가능한 불리한 영향을 미치지 않았고, 외관상의 분해 또는 응집체 형성이 없었다. 회수한 물질은 생물학적 활성도 나타내는 것으로 보였다.

<218> 생체내 시험을 위해, 당뇨병 마우스(db/db)를 흡입 챔버 내에 배치시키고 에어로졸화된 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 다양한 시간 동안 흡입시켰다. 이어서, 흡입 노출 시간의 말기에서 다음 72시간 동안 혈당에 대해 마우스를 모니터링하였다. 챔버 내의 흡입 시간은 동물이 피하 투여된 0.3, 1 및 3 mg/kg의 투여량에 동등한 노출을 받도록 선택되었다. 생체내 활성에서 흡입 투여 경로와 비교하기 위해 대조군으로서 상기 투여량들을 피하 투여하였다. 흡입에 의해 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질을 제공받은 동물에서의 혈당 농도는 상기 단백질에의 노출 후 상당히 낮아졌다. 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질의 순환 농도의 평가는 이 비-최적화된 시스템 및 제형화에 대해 약 10%의 생체이용률을 보여주었다.

도면의 간단한 설명

<9> 도 1은 시험관 내에서 GLP-1(7-37,A8G,K34A)(PEAPTD)₂(서열번호 5) mTf 융합 단백질(GLP-1/Tf)(도 1A)와 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질(엑센딘-4/Tf)(도 1B)의 콜라게네이즈 내성(MMP-1)을 비교한 것을 보여주는 그래프이다.

<10> 도 2는 당뇨병(db/db) 마우스에서 혈당에 대한 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질(엑센딘-4/Tf)의 투여량 효과, 및 엑센딘-4 대조군에 대한 비교 효과를 보여주는 그래프이다. 각 점은 평균 혈당 측정치를 보여준다(n=3).

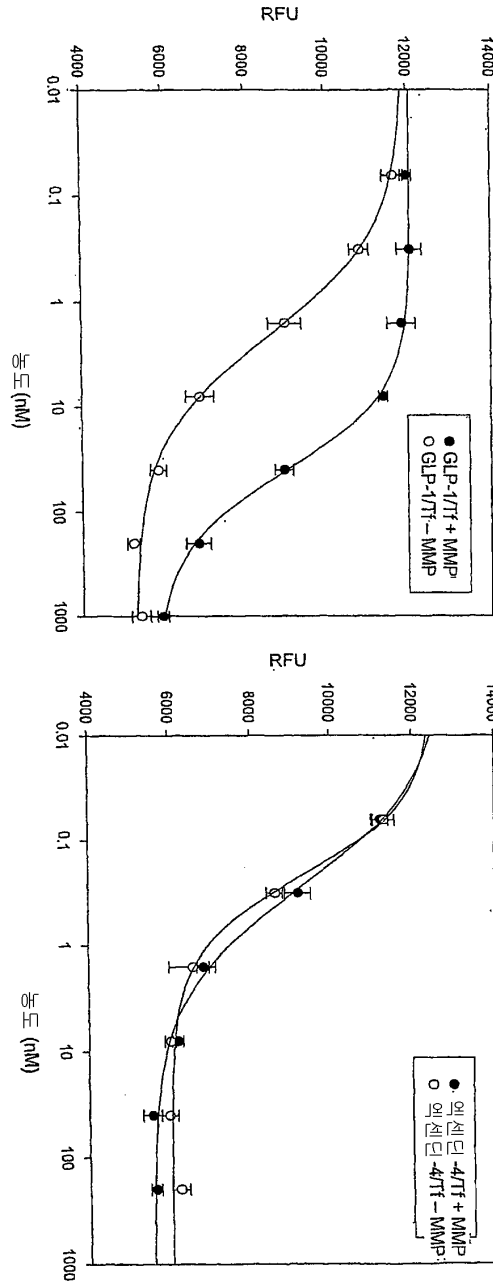
<11> 도 3은 체중에 대한 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질(엑센딘-4/Tf)의 매일 주사의 투여량 효과, 및 엑센딘-4 및 mTf 대조군에 대한 비교 효과를 보여주는 그래프이다.

<12> 도 4는 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5) mTf 융합 단백질(엑센딘-4/Tf)과 엑센딘-4에 대한 상대적 효능을 비교하는 그래프이다. 이것은 세포-기재 cAMP 분석에 의해 측정되었다. 엑센딘-4(1-39)(PEAPTD)₂(서열번호 5)

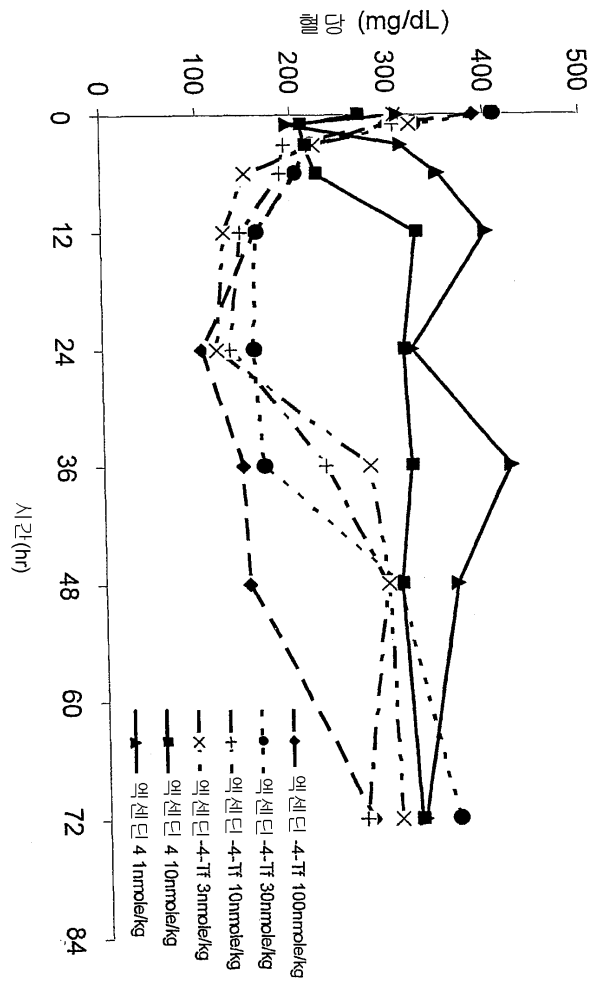
mTf 융합 단백질에 대한 EC50은 31.3 pM이고, 엑센딘-4에 대한 EC50은 6.6 pM이다.

도면

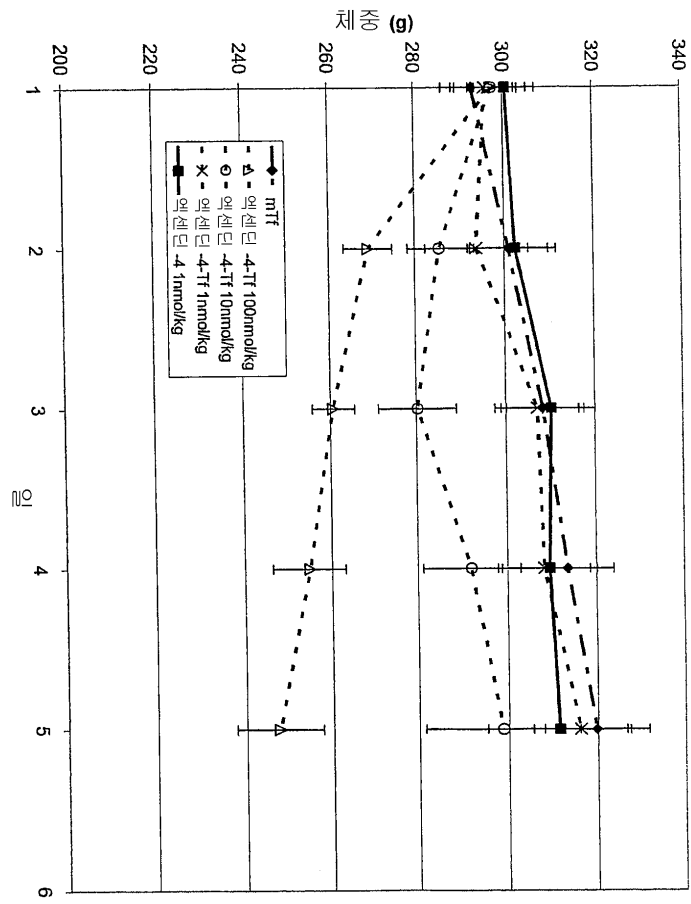
도면1



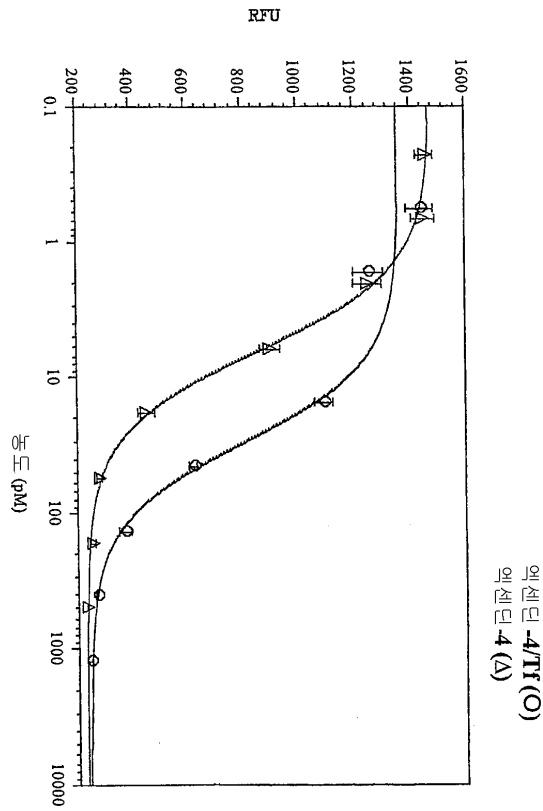
도면2



도면3



도면4



서열목록

SEQUENCE LISTING

<110> Biorexis Pharmaceutical Corporation
Sadeghi, Homayoun
Turner, Andrew J.
Prior, Christopher P.
Ballance, David J.

<120> EXENDIN FUSION PROTEINS

<130> PC19606A

<150> US 60/832,582

<151> 2006-07-24

<150> US/857,474

<151> 2006-11-08

<150> US 60/874,965

<151> 2006-12-15

<160> 38

<170> PatentIn version 3.5

<210> 1

<211> 2318

<212> DNA

<213> Homo sapiens

<400> 1

gcacagaagc gagtccgact gtgctcgctg ctcagcgccg cacccggaag atgaggctcg 60

ccgtgggagc cctgctggtc tgcgccgtcc tggggctgtg tctggctgtc cctgataaaa 120

ctgtgagatg gtgtgcagtg tccgagcatg aggccactaa gtgccagagt ttccgcgacc 180

atatgaaaag cgtcattcca tccgatggtc ccagtgttgc ttgtgtgaag aaagcctcct 240

accttgattg catcagggcc attgcggcaa acgaagcggg tgctgtgaca ctggatgcag 300

gtttggtgta tgatgcttac ctggctccca ataacctgaa gcctgtggtg gcagagtctt 360

atgggtcaaa agaggatcca cagactttct attatgctgt tgctgtggtg aagaaggata 420

gtggcttcca gatgaaccag cttcaggca agaagtctg ccacacgggt ctaggcaggt 480

ccgctgggtg gaacatcccc ataggcttac ttactgtga cttacctgag ccacgtaaac 540

ctcttgagaa agcagtggcc aatttcttct cgggcagctg tgccccttgt gcggatggga 600

cggacttccc ccagctgtgt caactgtgtc cagggtgtgg ctgctccacc cttaaccaat 660

acttcggcta ctcgggagcc ttcaagtgtc tgaaggatgg tgctggggat gtggcctttg 720

tcaagcactc gactatattt gagaacttgg caaacaaggc tgacaggac cagtatgagc 780

tgctttgcct ggacaacacc cggaagccgg tagatgaata caaggactgc cacttggccc 840

aggtcccttc tcataccgtc gtggcccga gtatggcgg caaggaggac ttgatctggg 900

agcttctcaa ccaggccag gaacattttg gcaaagacaa atcaaaagaa ttccaactat 960
 tcagctctcc teatgggaag gacctgctgt ttaaggactc tgcccacggg tttttaaag 1020
 tccccccag gatggatgcc aagatgtacc tgggctatga gtatgtcact gccatccgga 1080
 atctacggga aggcacatgc ccagaagccc caacagatga atgcaagcct gtgaagtgg 1140
 gtgcgtgag ccaccacgag aggctcaagt gtgatgagtg gagtgttaac agttaggga 1200
 aaatagagtg tgtatcagca gagaccaccg aagactgcat cgccaagatc atgaatggag 1260
 aagctgatgc catgagcttg gatggagggt ttgtctacat agcgggcaag tgtggtctgg 1320
 tgctgtctt ggcagaaaac tacaataaga gcgataattg tgaggataca ccagaggcag 1380
 ggtattttg tntagcagtg gtgaagaaat cagcttctga cctcacctgg gacaatctga 1440
 aaggcaagaa gtctgccat acggcagttg gcagaaccgc tggetggaac atccccatgg 1500
 gcctgtctta caataagatc aaccactgca gatttgatga attttcagt gaaggttgtg 1560
 cccctgggtc taagaaagac tccagtctct gtaagctgtg tatgggctca ggcctaaacc 1620
 tgttgaacc caacaacaaa gaggataact acggctacac aggcgcttcc aggtgtctgg 1680
 ttgagaaggg agatgtggcc ttttgaaac accagactgt cccacagaac actgggggaa 1740
 aaaacctga tccatgggct aagaatctga atgaaaaga ctatgagtgt ctgtgccttg 1800
 atgtaccag gaaacctgtg gaggagtatg cgaactgcca cctggccaga gccccgaatc 1860
 acgctgtggt cacacgaaa gataaggaag cttgcgtcca caagatatta cgtcaacagc 1920
 agcacctatt tggaagcaac gtaactgact gctcgggcaa cttttgtttg ttccggtcgg 1980
 aaaccaagga cttctgttc agagatgaca cagtatgttt ggccaaactt catgacagaa 2040
 acacatatga aaaatactta ggagaagaat atgtcaaggc tgttgtaac ctgagaaaat 2100

gtccaccte atcactcctg gaagcctgca cttccgtag accttaaaat ctcagaggta 2160
 gggctgccac caaggtgaag atgggaacgc agatgatcca tgagtttgcc ctggtttcac 2220
 tggccaagt ggitttgct aaccacgtct gtcttcacag ctctgtgtg ccatgtgtgc 2280
 tgaacaaaa ataaaaatta ttattgattt tatatttc 2318

<210> 2
 <211> 698
 <212> PRT
 <213> Homo sapiens

<400> 2

Met Arg Leu Ala Val Gly Ala Leu Leu Val Cys Ala Val Leu Gly Leu
 1 5 10 15

Cys Leu Ala Val Pro Asp Lys Thr Val Arg Trp Cys Ala Val Ser Glu
 20 25 30

His Glu Ala Thr Lys Cys Gln Ser Phe Arg Asp His Met Lys Ser Val
 35 40 45

Ile Pro Ser Asp Gly Pro Ser Val Ala Cys Val Lys Lys Ala Ser Tyr
 50 55 60

Leu Asp Cys Ile Arg Ala Ile Ala Ala Asn Glu Ala Asp Ala Val Thr
 65 70 75 80

Leu Asp Ala Gly Leu Val Tyr Asp Ala Tyr Leu Ala Pro Asn Asn Leu
 85 90 95

Lys Pro Val Val Ala Glu Phe Tyr Gly Ser Lys Glu Asp Pro Gln Thr
 100 105 110

Phe Tyr Tyr Ala Val Ala Val Val Lys Lys Asp Ser Gly Phe Gln Met
 115 120 125

Asn Gln Leu Arg Gly Lys Lys Ser Cys His Thr Gly Leu Gly Arg Ser
 130 135 140

Ala Gly Trp Asn Ile Pro Ile Gly Leu Leu Tyr Cys Asp Leu Pro Glu
 145 150 155 160

Pro Arg Lys Pro Leu Glu Lys Ala Val Ala Asn Phe Phe Ser Gly Ser
 165 170 175

Cys Ala Pro Cys Ala Asp Gly Thr Asp Phe Pro Gln Leu Cys Gln Leu
 180 185 190

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

Gly Ala Phe Lys Cys Leu Lys Asp Gly Ala Gly Asp Val Ala Phe Val
 210 215 220

Lys His Ser Thr Ile Phe Glu Asn Leu Ala Asn Lys Ala Asp Arg Asp
 225 230 235 240

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

Tyr Lys Asp Cys His Leu Ala Gln Val Pro Ser His Thr Val Val Ala
 260 265 270

Arg Ser Met Gly Gly Lys Glu Asp Leu Ile Trp Glu Leu Leu Asn Gln
 275 280 285

Ala Gln Glu His Phe Gly Lys Asp Lys Ser Lys Glu Phe Gln Leu Phe
 290 295 300

Ser Ser Pro His Gly Lys Asp Leu Leu Phe Lys Asp Ser Ala His Gly

Leu Cys Lys Leu Cys Met Gly Ser Gly Leu Asn Leu Cys Glu Pro Asn
 515 520 525

Asn Lys Glu Gly Tyr Tyr Gly Tyr Thr Gly Ala Phe Arg Cys Leu Val
 530 535 540

Glu Lys Gly Asp Val Ala Phe Val Lys His Gln Thr Val Pro Gln Asn
 545 550 555 560

Thr Gly Gly Lys Asn Pro Asp Pro Trp Ala Lys Asn Leu Asn Glu Lys
 565 570 575

Asp Tyr Glu Leu Leu Cys Leu Asp Gly Thr Arg Lys Pro Val Glu Glu
 580 585 590

Tyr Ala Asn Cys His Leu Ala Arg Ala Pro Asn His Ala Val Val Thr
 595 600 605

Arg Lys Asp Lys Glu Ala Cys Val His Lys Ile Leu Arg Gln Gln Gln
 610 615 620

His Leu Phe Gly Ser Asn Val Thr Asp Cys Ser Gly Asn Phe Cys Leu
 625 630 635 640

Phe Arg Ser Glu Thr Lys Asp Leu Leu Phe Arg Asp Asp Thr Val Cys
 645 650 655

Leu Ala Lys Leu His Asp Arg Asn Thr Tyr Glu Lys Tyr Leu Gly Glu
 660 665 670

Glu Tyr Val Lys Ala Val Gly Asn Leu Arg Lys Cys Ser Thr Ser Ser
 675 680 685

Leu Leu Glu Ala Cys Thr Phe Arg Arg Pro
 690 695

<210> 3
 <211> 679
 <212> PRT
 <213> Homo sapiens

<400> 3

Val Pro Asp Lys Thr Val Arg Trp Cys Ala Val Ser Glu His Glu Ala
 1 5 10 15

Thr Lys Cys Gln Ser Phe Arg Asp His Met Lys Ser Val Ile Pro Ser
 20 25 30

Asp Gly Pro Ser Val Ala Cys Val Lys Lys Ala Ser Tyr Leu Asp Cys
 35 40 45

Ile Arg Ala Ile Ala Ala Asn Glu Ala Asp Ala Val Thr Leu Asp Ala
 50 55 60

Gly Leu Val Tyr Asp Ala Tyr Leu Ala Pro Asn Asn Leu Lys Pro Val
 65 70 75 80

Val Ala Glu Phe Tyr Gly Ser Lys Glu Asp Pro Gln Thr Phe Tyr Tyr
 85 90 95

Ala Val Ala Val Val Lys Lys Asp Ser Gly Phe Gln Met Asn Gln Leu
 100 105 110

Arg Gly Lys Lys Ser Cys His Thr Gly Leu Gly Arg Ser Ala Gly Trp
 115 120 125

Asn Ile Pro Ile Gly Leu Leu Tyr Cys Asp Leu Pro Glu Pro Arg Lys
 130 135 140

Pro Leu Glu Lys Ala Val Ala Asn Phe Phe Ser Gly Ser Cys Ala Pro

145 150 155 160
 Cys Ala Asp Gly Thr Asp Phe Pro Gln Leu Cys Gln Leu Cys Pro Gly
 165 170 175
 Cys Gly Cys Ser Thr Leu Asn Gln Tyr Phe Gly Tyr Ser Gly Ala Phe
 180 185 190
 Lys Cys Leu Lys Asp Gly Ala Gly Asp Val Ala Phe Val Lys His Ser
 195 200 205
 Thr Ile Phe Glu Asn Leu Ala Asn Lys Ala Asp Arg Asp Gln Tyr Glu
 210 215 220
 Leu Leu Cys Leu Asp Asn Thr Arg Lys Pro Val Asp Glu Tyr Lys Asp
 225 230 235 240
 Cys His Leu Ala Gln Val Pro Ser His Thr Val Val Ala Arg Ser Met
 245 250 255
 Gly Gly Lys Glu Asp Leu Ile Trp Glu Leu Leu Asn Gln Ala Gln Glu
 260 265 270
 His Phe Gly Lys Asp Lys Ser Lys Glu Phe Gln Leu Phe Ser Ser Pro
 275 280 285
 His Gly Lys Asp Leu Leu Phe Lys Asp Ser Ala His Gly Phe Leu Lys
 290 295 300
 Val Pro Pro Arg Met Asp Ala Lys Met Tyr Leu Gly Tyr Glu Tyr Val
 305 310 315 320
 Thr Ala Ile Arg Asn Leu Arg Glu Gly Thr Cys Pro Glu Ala Pro Thr
 325 330 335
 Asp Glu Cys Lys Pro Val Lys Trp Cys Ala Leu Ser His His Glu Arg
 340 345 350

Leu Lys Cys Asp Glu Trp Ser Val Asn Ser Val Gly Lys Ile Glu Cys
 355 360 365

Val Ser Ala Glu Thr Thr Glu Asp Cys Ile Ala Lys Ile Met Asn Gly
 370 375 380

Glu Ala Asp Ala Met Ser Leu Asp Gly Gly Phe Val Tyr Ile Ala Gly
 385 390 395 400

Lys Cys Gly Leu Val Pro Val Leu Ala Glu Asn Tyr Asn Lys Ser Asp
 405 410 415

Asn Cys Glu Asp Thr Pro Glu Ala Gly Tyr Phe Ala Val Ala Val Val
 420 425 430

Lys Lys Ser Ala Ser Asp Leu Thr Trp Asp Asn Leu Lys Gly Lys Lys
 435 440 445

Ser Cys His Thr Ala Val Gly Arg Thr Ala Gly Trp Asn Ile Pro Met
 450 455 460

Gly Leu Leu Tyr Asn Lys Ile Asn His Cys Arg Phe Asp Glu Phe Phe
 465 470 475 480

Ser Glu Gly Cys Ala Pro Gly Ser Lys Lys Asp Ser Ser Leu Cys Lys
 485 490 495

Leu Cys Met Gly Ser Gly Leu Asn Leu Cys Glu Pro Asn Asn Lys Glu
 500 505 510

Gly Tyr Tyr Gly Tyr Thr Gly Ala Phe Arg Cys Leu Val Glu Lys Gly
 515 520 525

Asp Val Ala Phe Val Lys His Gln Thr Val Pro Gln Asn Thr Gly Gly
 530 535 540

Lys Asn Pro Asp Pro Trp Ala Lys Asn Leu Asn Glu Lys Asp Tyr Glu
 545 550 555 560

Leu Leu Cys Leu Asp Gly Thr Arg Lys Pro Val Glu Glu Tyr Ala Asn
 565 570 575

Cys His Leu Ala Arg Ala Pro Asn His Ala Val Val Thr Arg Lys Asp
 580 585 590

Lys Glu Ala Cys Val His Lys Ile Leu Arg Gln Gln Gln His Leu Phe
 595 600 605

Gly Ser Asn Val Thr Asp Cys Ser Gly Asn Phe Cys Leu Phe Arg Ser
 610 615 620

Glu Thr Lys Asp Leu Leu Phe Arg Asp Asp Thr Val Cys Leu Ala Lys
 625 630 635 640

Leu His Asp Arg Asn Thr Tyr Glu Lys Tyr Leu Gly Glu Glu Tyr Val
 645 650 655

Lys Ala Val Gly Asn Leu Arg Lys Cys Ser Thr Ser Ser Leu Leu Glu
 660 665 670

Ala Cys Thr Phe Arg Arg Pro
 675

- <210> 4
- <211> 39
- <212> PRT
- <213> Heloderma suspectum

<400> 4

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> 5
<211> 12
<212> PRT
<213> Unknown

<220>
<223> linker peptide

<400> 5

Pro Glu Ala Pro Thr Asp Pro Glu Ala Pro Thr Asp
1 5 10

<210> 6
<211> 6
<212> PRT
<213> Artificial

<220>
<223> non-helical linker peptide

<400> 6

Pro Glu Ala Pro Thr Asp
1 5

<210> 7
<211> 16
<212> PRT
<213> Artificial

<220>
<223> peptide linker

<400> 7

Pro Glu Ala Pro Thr Asp Glu Pro Lys Ser Cys Asp Lys Thr His Thr
 1 5 10 15

<210> 8

<211> 31

<212> PRT

<213> Artificial

<220>

<223> peptide linker

<400> 8

Pro Glu Ala Pro Thr Asp Val Glu Pro Lys Ser Ser Asp Lys Thr His
 1 5 10 15

Thr Ser Pro Pro Ser Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
 20 25 30

<210> 9

<211> 31

<212> PRT

<213> Artificial

<220>

<223> peptide linker

<400> 9

Pro Glu Ala Pro Thr Asp Val Glu Pro Lys Ser Ala Asp Lys Thr His
 1 5 10 15

Thr Ala Pro Pro Ala Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
 20 25 30

<210> 10

<211> 21

<212> PRT

<213> Artificial

<220>

<223> peptide linker

<400> 10

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Pro
1 5 10 15

Glu Ala Pro Thr Asp
 20

<210> 11

<211> 31

<212> PRT

<213> Artificial

<220>

<223> peptide linker

<400> 11

Val Glu Pro Lys Ser Ser Asp Lys Thr His Thr Ser Pro Pro Ser Pro
1 5 10 15

Ala Pro Glu Leu Leu Gly Gly Pro Ser Pro Glu Ala Pro Thr Asp
 20 25 30

<210> 12

<211> 31

<212> PRT

<213> Artificial

<220>

<223> peptide linker

<400> 12

Val Glu Pro Lys Ser Ala Asp Lys Thr His Thr Ala Pro Pro Ala Pro

1 5 10 15

Ala Pro Glu Leu Leu Gly Gly Pro Ser Pro Glu Ala Pro Thr Asp
 20 25 30

<210> 13
 <211> 31
 <212> PRT
 <213> Artificial

<220>
 <223> peptide linker

<400> 13

Pro Glu Ala Pro Thr Asp Val Glu Pro Lys Ser Cys Asp Lys Thr His
 1 5 10 15

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
 20 25 30

<210> 14
 <211> 31
 <212> PRT
 <213> Artificial

<220>
 <223> peptide linker

<400> 14

Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
 1 5 10 15

Ala Pro Glu Leu Leu Gly Gly Pro Ser Pro Glu Ala Pro Thr Asp
 20 25 30

<210> 15
 <211> 31
 <212> PRT

<213> Artificial

<220>

<223> peptide linker

<400> 15

Pro Glu Ala Pro Thr Asp Val Glu Pro Lys Ala Ala Asp Lys Thr His
 1 5 10 15

Thr Ala Pro Pro Ala Pro Ala Pro Glu Leu Leu Gly Gly Pro Ala
 20 25 30

<210> 16

<211> 31

<212> PRT

<213> Artificial

<220>

<223> peptide linker

<400> 16

Val Glu Pro Lys Ala Ala Asp Lys Thr His Thr Ala Pro Pro Ala Pro
 1 5 10 15

Ala Pro Glu Leu Leu Gly Gly Pro Ala Pro Glu Ala Pro Thr Asp
 20 25 30

<210> 17

<211> 679

<212> PRT

<213> Homo sapiens

<400> 17

Val Pro Asp Lys Thr Val Arg Trp Cys Ala Val Ser Glu His Glu Ala
 1 5 10 15

Thr Lys Cys Gln Ser Phe Arg Asp His Met Lys Ser Val Ile Pro Ser

20 25 30

Asp Gly Pro Ser Val Ala Cys Val Lys Lys Ala Ser Tyr Leu Asp Cys
35 40 45

Ile Arg Ala Ile Ala Ala Asn Glu Ala Asp Ala Val Thr Leu Asp Ala
50 55 60

Gly Leu Val Tyr Asp Ala Tyr Leu Ala Pro Asn Asn Leu Lys Pro Val
65 70 75 80

Val Ala Glu Phe Tyr Gly Ser Lys Glu Asp Pro Gln Thr Phe Tyr Tyr
85 90 95

Ala Val Ala Val Val Lys Lys Asp Ser Gly Phe Gln Met Asn Gln Leu
100 105 110

Arg Gly Lys Lys Ser Cys His Thr Gly Leu Gly Arg Ser Ala Gly Trp
115 120 125

Asn Ile Pro Ile Gly Leu Leu Tyr Cys Asp Leu Pro Glu Pro Arg Lys
130 135 140

Pro Leu Glu Lys Ala Val Ala Asn Phe Phe Ser Gly Ser Cys Ala Pro
145 150 155 160

Cys Ala Asp Gly Thr Asp Phe Pro Gln Leu Cys Gln Leu Cys Pro Gly
165 170 175

Cys Gly Cys Ser Thr Leu Asn Gln Tyr Phe Gly Tyr Ser Gly Ala Phe
180 185 190

Lys Cys Leu Lys Asp Gly Ala Gly Asp Val Ala Phe Val Lys His Ser
195 200 205

Thr Ile Phe Glu Asn Leu Ala Asn Lys Ala Asp Arg Asp Gln Tyr Glu

210

215

220

Leu Leu Cys Leu Asp Asn Thr Arg Lys Pro Val Asp Glu Tyr Lys Asp
225 230 235 240

Cys His Leu Ala Gln Val Pro Ser His Thr Val Val Ala Arg Ser Met
245 250 255

Gly Gly Lys Glu Asp Leu Ile Trp Glu Leu Leu Asn Gln Ala Gln Glu
260 265 270

His Phe Gly Lys Asp Lys Ser Lys Glu Phe Gln Leu Phe Ser Ser Pro
275 280 285

His Gly Lys Asp Leu Leu Phe Lys Asp Ser Ala His Gly Phe Leu Lys
290 295 300

Val Pro Pro Arg Met Asp Ala Lys Met Tyr Leu Gly Tyr Glu Tyr Val
305 310 315 320

Thr Ala Ile Arg Asn Leu Arg Glu Gly Thr Cys Pro Glu Ala Pro Thr
325 330 335

Asp Glu Cys Lys Pro Val Lys Trp Cys Ala Leu Ser His His Glu Arg
340 345 350

Leu Lys Cys Asp Glu Trp Ser Val Asn Ser Val Gly Lys Ile Glu Cys
355 360 365

Val Ser Ala Glu Thr Thr Glu Asp Cys Ile Ala Lys Ile Met Asn Gly
370 375 380

Glu Ala Asp Ala Met Ser Leu Asp Gly Gly Phe Val Tyr Ile Ala Gly
385 390 395 400

Lys Cys Gly Leu Val Pro Val Leu Ala Glu Asn Tyr Asn Lys Ala Asp
405 410 415

Asn Cys Glu Asp Thr Pro Glu Ala Gly Tyr Phe Ala Val Ala Val Val
 420 425 430

Lys Lys Ser Ala Ser Asp Leu Thr Trp Asp Asn Leu Lys Gly Lys Lys
 435 440 445

Ser Cys His Thr Ala Val Gly Arg Thr Ala Gly Trp Asn Ile Pro Met
 450 455 460

Gly Leu Leu Tyr Asn Lys Ile Asn His Cys Arg Phe Asp Glu Phe Phe
 465 470 475 480

Ser Glu Gly Cys Ala Pro Gly Ser Lys Lys Asp Ser Ser Leu Cys Lys
 485 490 495

Leu Cys Met Gly Ser Gly Leu Asn Leu Cys Glu Pro Asn Asn Lys Glu
 500 505 510

Gly Tyr Tyr Gly Tyr Thr Gly Ala Phe Arg Cys Leu Val Glu Lys Gly
 515 520 525

Asp Val Ala Phe Val Lys His Gln Thr Val Pro Gln Asn Thr Gly Gly
 530 535 540

Lys Asn Pro Asp Pro Trp Ala Lys Asn Leu Asn Glu Lys Asp Tyr Glu
 545 550 555 560

Leu Leu Cys Leu Asp Gly Thr Arg Lys Pro Val Glu Glu Tyr Ala Asn
 565 570 575

Cys His Leu Ala Arg Ala Pro Asn His Ala Val Val Thr Arg Lys Asp
 580 585 590

Lys Glu Ala Cys Val His Lys Ile Leu Arg Gln Gln Gln His Leu Phe
 595 600 605

Gly Ser Asn Val Ala Asp Cys Ser Gly Asn Phe Cys Leu Phe Arg Ser
 610 615 620

Glu Thr Lys Asp Leu Leu Phe Arg Asp Asp Thr Val Cys Leu Ala Lys
 625 630 635 640

Leu His Asp Arg Asn Thr Tyr Glu Lys Tyr Leu Gly Glu Glu Tyr Val
 645 650 655

Lys Ala Val Gly Asn Leu Arg Lys Cys Ser Thr Ser Ser Leu Leu Glu
 660 665 670

Ala Cys Thr Phe Arg Arg Pro
 675

<210> 18
 <211> 19
 <212> PRT
 <213> Homo sapiens

<400> 18

Met Arg Leu Ala Val Gly Ala Leu Leu Val Cys Ala Val Leu Gly Leu
 1 5 10 15

Cys Leu Ala

<210> 19
 <211> 57
 <212> DNA
 <213> Homo sapiens

<400> 19

atgaggctcg ccgtgggagc cctgetggtc tgcgcctcc tggggctgtg tetggcg

57

<210> 20

<211> 117
 <212> DNA
 <213> Heloderma suspectum

<400> 20
 catggtgaag gtacttttac ttctgatttg tctaaacaaa tggaagaaga agctgttaga 60

ttgtttattg aatggttgaa aaatggtggt ccatcttctg gtgctccacc accatct 117

<210> 21
 <211> 12
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> peptide linker

<400> 21

Pro Glu Ala Pro Thr Asp Pro Glu Ala Pro Thr Asp
 1 5 10

<210> 22
 <211> 2043
 <212> DNA
 <213> Homo sapiens

<400> 22
 gtacctgata aaactgtgag atggtgtgca gtgtcggagc atgaggccac taagtccag 60

agtttccgcg accatatgaa aagcgtcatt ccatccgatg gtcccagtgt tgcttgtgtg 120

aagaaagcct cctaccttga ttgcatcagg gccattgcgg caaacgaagc ggatgctgtg 180

acactggatg caggtttggt gtatgatgct tacctggctc ccaataacct gaagcctgtg 240

giggcagagt tctatgggtc aaaagaggat ccacagactt tctattatgc tgttgctgtg 300

gtgaagaagg atagtggctt ccagatgaac cagcttcgag gcaagaagtc ctgccacag 360

ggtctaggca ggtccgctgg gtggaacatc cccataggct tactttactg tgacttacct 420
 gagccacgta aacctcttga gaaagcagtg gccaatcttct tctcgggcag ctgtgccct 480
 tgtgaggatg ggacggactt cccccagctg tgcactgt gtccagggtg tggctgctcc 540
 acccttaacc aatacttcgg ctactcggga gccttcaagt gtctgaagga tggctgctggg 600
 gatgtggcct ttgtcaagca ctgactata ttgagaact tggcaaaca ggctgacagg 660
 gaccgatag agctgctttg cctggacaac acccggaagc cggtagatga atacaaggac 720
 tgccacttgg ccaggtccc ttctcatacc gtcgtggccc gaagtatggg cggcaaggag 780
 gacttgatct gggagcttct caaccaggcc caggaacatt ttggcaaaga caaatcaaaa 840
 gaattccaac tattcagctc tctcatggg aaggacctgc tgtttaagga ctctgccac 900
 gggtttttaa aagtccccc caggatggat gccaatgt acctgggcta tgagtatgtc 960
 actgccatcc ggaatctacg ggaaggcaca tgcccagaag cccaacaga tgaatgcaag 1020
 cctgtgaagt ggtgtgcct gagccaccac gagaggctca agtgtgatga gtggagtgtt 1080
 aacagtgtag ggaaaataga gtgtgtatca gcagagacca ccgaagactg catcgccaag 1140
 atcatgaatg gagaagctga tgccatgagc ttggatggag ggtttgtcta catagcgggc 1200
 aagtgtggtc tggcgcctgt cttggcagaa aactacaata aggctgataa ttgtaggat 1260
 acaccagagg cagggtatct tgctgtagca gtggtgaaga aatcagcttc tgacctcacc 1320
 tgggacaatc tgaaggcaa gaagtcctgc catacggcag ttggcagaac cgctggctgg 1380
 aacatcccca tggcctgct ctacaataag atcaaccact gcagatttga tgaatttttc 1440
 agtgaaggtt gtccccctgg gtctaagaaa gactccagtc tctgtaagct gtgtatgggc 1500
 tcaggcctaa acctctgtga acccaacaac aaagagggat actacggcta cacaggcct 1560

ttcaggtgtc tggttgagaa gggagatgtg gcctttgtga aacaccagac tgtcccacag 1620
 aacctgggg gaaaaaaccc tgatccatgg gctaagaatc tgaatgaaaa agactatgag 1680
 ttgctgtgcc ttgatggtac taggaaacct gtggaggagt atgcgaactg ccacctggcc 1740
 agagccccga atcacgctgt ggtcacacgg aaagataagg aagcatgcgt ccacaagata 1800
 ttacgtcaac agcagcacct atttggaagc aacgtagctg actgctcggg caacttttgt 1860
 ttgttccggt cggaaaccaa ggaccttctg ttcagagatg acacagtatg tttggccaaa 1920
 cttcatgaca gaaacacata tgaaaaatac ttaggagaag aatatgtcaa ggctgttgg 1980
 aacctgagaa aatgctccac ctcatcactc ctggaagcct gcactttccg tcgaccttaa 2040
 taa 2043

<210> 23
 <211> 730
 <212> PRT
 <213> Homo sapiens

<400> 23

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

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

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

Pro Thr Asp Val Pro Asp Lys Thr Val Arg Trp Cys Ala Val Ser Glu
 50 55 60

His Glu Ala Thr Lys Cys Gln Ser Phe Arg Asp His Met Lys Ser Val
65 70 75 80

Ile Pro Ser Asp Gly Pro Ser Val Ala Cys Val Lys Lys Ala Ser Tyr
85 90 95

Leu Asp Cys Ile Arg Ala Ile Ala Ala Asn Glu Ala Asp Ala Val Thr
100 105 110

Leu Asp Ala Gly Leu Val Tyr Asp Ala Tyr Leu Ala Pro Asn Asn Leu
115 120 125

Lys Pro Val Val Ala Glu Phe Tyr Gly Ser Lys Glu Asp Pro Gln Thr
130 135 140

Phe Tyr Tyr Ala Val Ala Val Val Lys Lys Asp Ser Gly Phe Gln Met
145 150 155 160

Asn Gln Leu Arg Gly Lys Lys Ser Cys His Thr Gly Leu Gly Arg Ser
165 170 175

Ala Gly Trp Asn Ile Pro Ile Gly Leu Leu Tyr Cys Asp Leu Pro Glu
180 185 190

Pro Arg Lys Pro Leu Glu Lys Ala Val Ala Asn Phe Phe Ser Gly Ser
195 200 205

Cys Ala Pro Cys Ala Asp Gly Thr Asp Phe Pro Gln Leu Cys Gln Leu
210 215 220

Cys Pro Gly Cys Gly Cys Ser Thr Leu Asn Gln Tyr Phe Gly Tyr Ser
225 230 235 240

Gly Ala Phe Lys Cys Leu Lys Asp Gly Ala Gly Asp Val Ala Phe Val
245 250 255

Lys His Ser Thr Ile Phe Glu Asn Leu Ala Asn Lys Ala Asp Arg Asp

450

455

460

Lys Ala Asp Asn Cys Glu Asp Thr Pro Glu Ala Gly Tyr Phe Ala Val
465 470 475 480

Ala Val Val Lys Lys Ser Ala Ser Asp Leu Thr Trp Asp Asn Leu Lys
485 490 495

Gly Lys Lys Ser Cys His Thr Ala Val Gly Arg Thr Ala Gly Trp Asn
500 505 510

Ile Pro Met Gly Leu Leu Tyr Asn Lys Ile Asn His Cys Arg Phe Asp
515 520 525

Glu Phe Phe Ser Glu Gly Cys Ala Pro Gly Ser Lys Lys Asp Ser Ser
530 535 540

Leu Cys Lys Leu Cys Met Gly Ser Gly Leu Asn Leu Cys Glu Pro Asn
545 550 555 560

Asn Lys Glu Gly Tyr Tyr Gly Tyr Thr Gly Ala Phe Arg Cys Leu Val
565 570 575

Glu Lys Gly Asp Val Ala Phe Val Lys His Gln Thr Val Pro Gln Asn
580 585 590

Thr Gly Gly Lys Asn Pro Asp Pro Trp Ala Lys Asn Leu Asn Glu Lys
595 600 605

Asp Tyr Glu Leu Leu Cys Leu Asp Gly Thr Arg Lys Pro Val Glu Glu
610 615 620

Tyr Ala Asn Cys His Leu Ala Arg Ala Pro Asn His Ala Val Val Thr
625 630 635 640

Arg Lys Asp Lys Glu Ala Cys Val His Lys Ile Leu Arg Gln Gln Gln
645 650 655

His Leu Phe Gly Ser Asn Val Ala Asp Cys Ser Gly Asn Phe Cys Leu
 660 665 670

Phe Arg Ser Glu Thr Lys Asp Leu Leu Phe Arg Asp Asp Thr Val Cys
 675 680 685

Leu Ala Lys Leu His Asp Arg Asn Thr Tyr Glu Lys Tyr Leu Gly Glu
 690 695 700

Glu Tyr Val Lys Ala Val Gly Asn Leu Arg Lys Cys Ser Thr Ser Ser
 705 710 715 720

Leu Leu Glu Ala Cys Thr Phe Arg Arg Pro
 725 730

<210> 24
 <211> 2196
 <212> DNA
 <213> Homo sapiens

<400> 24
 catggtgaag gtacttttac ttctgatttg tctaaacaaa tggaagaaga agctgttaga 60
 ttgtttattg aatggttgaa aaatggtggt ccatcttctg gtgctccacc accatctcca 120
 gaagctccaa ctgatccaga agctccaact gatgtacctg ataaaactgt gagatggtgt 180
 gcagtgtcgg agcatgaggc cactaagtgc cagagtttcc gcgaccatat gaaaagcgtc 240
 attccatccg atggtccag tgttgcttgt gtgaagaaag cctcctacct tgattgcac 300
 agggccattg cggcaaacga agcggatgct gtgacactgg atgcaggttt ggtgtatgat 360
 gettacctgg ctccaataa cctgaagcct gtggtggcag agttctatgg gtcaaaagag 420
 gatccacaga ctttctatta tgctgttgc gtggtgaaga aggatagtgg cttccagatg 480

aaccagcttc gaggaagaa gtctgtccac acgggtctag gcaggtccgc tgggtggaac 540

atcccatag gcttacttta ctgtgactta cctgagccac gtaaacctct tgagaaagca 600

gtggccaatt tcttctcggg cagctgtgcc ccttgtgcgg atgggacgga cttccccag 660

ctgtgtcaac tgtgtccagg gtgtggctgc tccaccctta accaatactt cggctactcg 720

ggagccttca agtgtctgaa ggatgggtgtg ggggatgtgg cctttgtcaa gcactcgact 780

atatttgaga acttggcaaa caagctgac agggaccagt atgagctgct ttgcctggac 840

aacaccgga agccggtaga tgaatacaag gactgccact tggcccaggt cccttctcat 900

accgtcgtgg cccgaagtat gggcggcaag gaggacttga tctgggagct tctcaaccag 960

gcccaggaac attttgcaa agacaaatca aaagaattcc aactattcag ctctctcat 1020

gggaaggacc tgetgtttaa ggactctgcc cacgggtttt taaaagtccc ccccaggatg 1080

gatgccaaga tgiacctggg ctatgagtat gtcactgcca tccggaatct acgggaaggc 1140

acatgccag aagccccaac agatgaatgc aagcctgtga agtgggtgtc gctgagccac 1200

cacgagagcc tcaagtgtga tgagtgaggt gttaacagtg tagggaaaat agagtgtgta 1260

tcagcagaga ccaccgaaga ctgcatgcc aagatcatga atggagaagc tgatgcatg 1320

agcttggatg gaggtttgt ctacatagcg ggcaagtgtg gtctgggtgcc tgtcttggca 1380

gaaaactaca ataaggctga taattgtgag gatacaccag aggcagggtta ttttgctgta 1440

gcagtgggtga agaatcagc ttctgacctc acctgggaca atctgaaagg caagaagtcc 1500

tgccatacgg cagttggcag aaccgtggc tggaacatcc ccatgggcct gctctacaat 1560

aagatcaacc actgcagatt tgatgaattt ttcagtgaag gttgtgcccc tgggtctaag 1620

aaagactcca gtctctgtaa gctgtgtatg ggctcaggcc taaacctctg tgaaccaaac 1680

aacaaagagg gatactacgg ctacacaggc gctttcaggt gtctggttga gaaggagat 1740
 gtggcctttg tgaaacacca gactgtccca cagaacactg ggggaaaaaa cctgatcca 1800
 tgggctaaga atctgaatga aaaagactat gagttgctgt gccttgatgg tactaggaaa 1860
 cctgtggagg agtatgcgaa ctgccacctg gccagagccc cgaatcacgc tgtggtcaca 1920
 cggaaagata aggaagcatg cgtccacaag atattacgtc aacagcagca cctatttga 1980
 agcaacttag ctgactgctc gggcaacttt tgtttgttcc ggtcggaaac caaggacctt 2040
 ctgttcagag atgacacagt atgtttggcc aaacttcatt acagaaacac atatgaaaaa 2100
 tacttaggag aagaatatgt caaggctgtt ggtaacctga gaaaatgctc cacctcatca 2160
 ctcttgaag cctgcacttt ccgtcgacct taataa 2196

<210> 25
 <211> 749
 <212> PRT
 <213> Homo sapiens

<400> 25

Met Arg Leu Ala Val Gly Ala Leu Leu Val Cys Ala Val Leu Gly Leu
 1 5 10 15

Cys Leu Ala His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln
 20 25 30

Met Glu Glu Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly
 35 40 45

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

Pro Glu Ala Pro Thr Asp Val Pro Asp Lys Thr Val Arg Trp Cys Ala
65 70 75 80

Val Ser Glu His Glu Ala Thr Lys Cys Gln Ser Phe Arg Asp His Met
85 90 95

Lys Ser Val Ile Pro Ser Asp Gly Pro Ser Val Ala Cys Val Lys Lys
100 105 110

Ala Ser Tyr Leu Asp Cys Ile Arg Ala Ile Ala Ala Asn Glu Ala Asp
115 120 125

Ala Val Thr Leu Asp Ala Gly Leu Val Tyr Asp Ala Tyr Leu Ala Pro
130 135 140

Asn Asn Leu Lys Pro Val Val Ala Glu Phe Tyr Gly Ser Lys Glu Asp
145 150 155 160

Pro Gln Thr Phe Tyr Tyr Ala Val Ala Val Val Lys Lys Asp Ser Gly
165 170 175

Phe Gln Met Asn Gln Leu Arg Gly Lys Lys Ser Cys His Thr Gly Leu
180 185 190

Gly Arg Ser Ala Gly Trp Asn Ile Pro Ile Gly Leu Leu Tyr Cys Asp
195 200 205

Leu Pro Glu Pro Arg Lys Pro Leu Glu Lys Ala Val Ala Asn Phe Phe
210 215 220

Ser Gly Ser Cys Ala Pro Cys Ala Asp Gly Thr Asp Phe Pro Gln Leu
225 230 235 240

Cys Gln Leu Cys Pro Gly Cys Gly Cys Ser Thr Leu Asn Gln Tyr Phe
245 250 255

Gly Tyr Ser Gly Ala Phe Lys Cys Leu Lys Asp Gly Ala Gly Asp Val

260 265 270

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

Asp Arg Asp Gln Tyr Glu Leu Leu Cys Leu Asp Asn Thr Arg Lys Pro
 290 295 300

Val Asp Glu Tyr Lys Asp Cys His Leu Ala Gln Val Pro Ser His Thr
 305 310 315 320

Val Val Ala Arg Ser Met Gly Gly Lys Glu Asp Leu Ile Trp Glu Leu
 325 330 335

Leu Asn Gln Ala Gln Glu His Phe Gly Lys Asp Lys Ser Lys Glu Phe
 340 345 350

Gln Leu Phe Ser Ser Pro His Gly Lys Asp Leu Leu Phe Lys Asp Ser
 355 360 365

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

Leu Gly Tyr Glu Tyr Val Thr Ala Ile Arg Asn Leu Arg Glu Gly Thr
 385 390 395 400

Cys Pro Glu Ala Pro Thr Asp Glu Cys Lys Pro Val Lys Trp Cys Ala
 405 410 415

Leu Ser His His Glu Arg Leu Lys Cys Asp Glu Trp Ser Val Asn Ser
 420 425 430

Val Gly Lys Ile Glu Cys Val Ser Ala Glu Thr Thr Glu Asp Cys Ile
 435 440 445

Ala Lys Ile Met Asn Gly Glu Ala Asp Ala Met Ser Leu Asp Gly Gly

450

455

460

Phe Val Tyr Ile Ala Gly Lys Cys Gly Leu Val Pro Val Leu Ala Glu
465 470 475 480

Asn Tyr Asn Lys Ala Asp Asn Cys Glu Asp Thr Pro Glu Ala Gly Tyr
485 490 495

Phe Ala Val Ala Val Val Lys Lys Ser Ala Ser Asp Leu Thr Trp Asp
500 505 510

Asn Leu Lys Gly Lys Lys Ser Cys His Thr Ala Val Gly Arg Thr Ala
515 520 525

Gly Trp Asn Ile Pro Met Gly Leu Leu Tyr Asn Lys Ile Asn His Cys
530 535 540

Arg Phe Asp Glu Phe Phe Ser Glu Gly Cys Ala Pro Gly Ser Lys Lys
545 550 555 560

Asp Ser Ser Leu Cys Lys Leu Cys Met Gly Ser Gly Leu Asn Leu Cys
565 570 575

Glu Pro Asn Asn Lys Glu Gly Tyr Tyr Gly Tyr Thr Gly Ala Phe Arg
580 585 590

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

Pro Gln Asn Thr Gly Gly Lys Asn Pro Asp Pro Trp Ala Lys Asn Leu
610 615 620

Asn Glu Lys Asp Tyr Glu Leu Leu Cys Leu Asp Gly Thr Arg Lys Pro
625 630 635 640

Val Glu Glu Tyr Ala Asn Cys His Leu Ala Arg Ala Pro Asn His Ala
645 650 655

Val Val Thr Arg Lys Asp Lys Glu Ala Cys Val His Lys Ile Leu Arg
 660 665 670

Gln Gln Gln His Leu Phe Gly Ser Asn Val Ala Asp Cys Ser Gly Asn
 675 680 685

Phe Cys Leu Phe Arg Ser Glu Thr Lys Asp Leu Leu Phe Arg Asp Asp
 690 695 700

Thr Val Cys Leu Ala Lys Leu His Asp Arg Asn Thr Tyr Glu Lys Tyr
 705 710 715 720

Leu Gly Glu Glu Tyr Val Lys Ala Val Gly Asn Leu Arg Lys Cys Ser
 725 730 735

Thr Ser Ser Leu Leu Glu Ala Cys Thr Phe Arg Arg Pro
 740 745

<210> 26
 <211> 2253
 <212> DNA
 <213> Homo sapiens

<400> 26
 atgaggctcg ccgtgggagc cctgctggtc tgcgccgtcc tggggctgtg tctggcgcac 60
 ggtgaaggta cttttacttc tgatttgtct aaacaaatgg aagaagaagc tgtagattg 120
 tttattgaat ggttgaaaaa tgggtgtcca tcttctgggtg ctccaccacc atctccagaa 180
 gctccaactg atccagaagc tccaactgat gtacctgata aaactgtgag atgggtgtgca 240
 gtgtcggagc atgaggccac taagtgccag agtttccgag accatatgaa aagcgtcatt 300
 ccatccgatg gtcccagtgt tgcttgtgtg aagaaagcct cctaccttga ttgcatcagg 360

gccattgctg caaacgaagc ggatgctgtg aactggatg caggtttgtg gtatgatgct 420
 tacttgctc ccaataacct gaagcctgtg gtggcagagt tctatgggtc aaaagaggat 480
 ccacagactt tctattatgc tgttgctgtg gtgaagaagg atagtggctt ccagatgaac 540
 cagcttcgag gcaagaagtc ctgccacacg ggtctaggca ggtccgctgg gtggaacatc 600
 cccataggct tactttactg tgacttacct gagccacgta aacctcttga gaaagcagtg 660
 gccaatctt tctcgggcag ctgtgccctt tgtgctgatg ggacggactt cccccagctg 720
 tgtcaactgt gtccagggtg tggctgctcc acccttaacc aatacttcgg ctactcggga 780
 gccttcaagt gtctgaagga tggctctggg gatgtggcct ttgtcaagca ctgcactata 840
 ttgagaact tggcaaacaa ggctgacagg gaccagtatg agctgctttg cctggacaac 900
 acccgaagc cggtagatga atacaaggac tgccacttgg cccaggtccc ttctcatacc 960
 gtcgtggccc gaagtatggg cggcaaggag gacttgatct gggagcttct caaccaggcc 1020
 caggaacatt ttggcaaaga caaatcaaaa gaattccaac tattcagctc tctcatggg 1080
 aaggacctgc tgtttaagga ctctgccac gggtttttaa aagtcccccc caggatggat 1140
 gccaatgatg acctgggcta tgagtatgtc actgccatcc ggaatctacg ggaaggcaca 1200
 tgcccagaag cccaacaga tgaatgcaag cctgtgaagt ggtgtgctg gagccaccac 1260
 gagagctca agtgtgatga gtggagtgtt aacagtgtag ggaaaataga gtgtgatca 1320
 gcagagacca ccgaagactg catgccaag atcatgaatg gagaagctga tgccatgagc 1380
 ttggatggag ggtttgtcta catagcgggc aagtgtggtc tggctcctgt cttggcagaa 1440
 aactacaata agctgataa ttgtgaggat acaccagagg cagggtatct tgctgtagca 1500
 gtggtgaaga aatcagcttc tgacctcacc tgggacaatc tgaaaggcaa gaagtctctg 1560

catacggcag ttggcagaac cgctggctgg aacatcecca tgggcctgct ctacaataag 1620
 atcaaccact gcagatttga tgaatttttc agtgaaggtt gtgccctgg gtctaagaaa 1680
 gactccagtc tcigtgaagct gtgtatgggc tcaggcctaa acctctgtga acccaacaac 1740
 aaagagggat actiacggcta cacaggcgtc ttcaggtgtc tggttgagaa gggagatgtg 1800
 gcctttgtga aacaccagac tgtcccacag aacctgggg gaaaaaaccc tgatccatgg 1860
 gctaagaatc tgaatgaaaa agactatgag ttgctgtgcc ttgatgttac taggaaacct 1920
 gtggaggagt atgcgaactg ccacctggcc agagccccga atcacgctgt ggtcacacgg 1980
 aaagataagg aagcatgcgt ccacaagata ttacttcaac agcagcacct atttggaagc 2040
 aacgtagctg actgctcggg caacttttgt ttgttccggt cggaaaccaa ggaccttctg 2100
 ttcagagatg acacagtatg tttggccaaa cttcatgaca gaaacacata tgaaaaatac 2160
 ttaggagaag aatatgtcaa ggctgttggc aacctgagaa aatgctccac ctcactactc 2220
 ctggaagcct gcactttccg tcgaccttaa taa 2253

<210> 27
 <211> 12
 <212> PRT
 <213> Artificial

<220>
 <223> neutrophil lactoferrin splice variant

<400> 27

Glu Asp Cys Ile Ala Leu Lys Gly Glu Ala Asp Ala
 1 5 10

<210> 28
 <211> 95

<212> DNA
 <213> Artificial

<220>
 <223> PCR primer P0702

<400> 28
 caataaaca tctaacagct tcttcttcca ttgtttaga caaatcagaa gtaaaagtac 60

cttcacatg cgccagacac agccccagga cggcg 95

<210> 29
 <211> 107
 <212> DNA
 <213> Artificial

<220>
 <223> PCR primer P0702

<400> 29
 ggaagaagaa gctgtagat tgtttattga atggttgaaa aatggtggtc catcttctgg 60

tgctccacca ccatctgtac ctgataaaac tgtgagatgg tgtgcag 107

<210> 30
 <211> 116
 <212> DNA
 <213> Heloderma suspectum

<400> 30
 atggtgaagg tacttttact tctgatttgt ctaaacaat ggaagaagaa gctgtagat 60

tgtttattga atggttgaaa aatggtggtc catcttctgg tgctccacca ccatct 116

<210> 31
 <211> 18
 <212> DNA
 <213> Artificial

<220>
 <223> PCR primer P0177

<400> 31
 gcgat aaaga gcgcgatg 18

<210> 32
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> PCR primer P0014

<400> 32
 ggctcaggta agtcacagta 20

<210> 33
 <211> 47
 <212> DNA
 <213> Artificial

<220>
 <223> PCR primer P1810

<400> 33
 ccaactgatc cagaagctcc aactgatgta cctgataaaa ctgtgag 47

<210> 34
 <211> 46
 <212> DNA
 <213> Artificial

<220>
 <223> PCR primer 1811

<400> 34
 gcttctggat cagttggagc ttctggagat ggtggtggag caccag 46

<210> 35
 <211> 67
 <212> DNA
 <213> Artificial

<220>

<223> PCR primer P0940

<400> 35

gtggtccacc agaagctcca actgatccag aagctccaac tgatgtacct gataaaactg 60

tgagatg

67

<210> 36

<211> 43

<212> DNA

<213> Artificial

<220>

<223> PCR primer P0941

<400> 36

tggagcttct ggtggaccac catttttcaa ccattcaata aac 43

<210> 37

<211> 65

<212> DNA

<213> Artificial

<220>

<223> PCR primer 0942

<400> 37

ggtggtccag aagctccaac tgatccagaa gctccaactg atgtacctga taaaactgtg 60

agatg

65

<210> 38

<211> 35

<212> DNA

<213> Artificial

<220>

<223> PCR primer 0943

<400> 38

gttggagctt ctggaccacc atttttcaac catte

35