



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0064600
(43) 공개일자 2019년06월10일

(51) 국제특허분류(Int. Cl.)
C07K 14/00 (2006.01) A61K 38/00 (2006.01)
A61P 35/00 (2006.01) C07K 14/31 (2006.01)
G01N 33/68 (2006.01)
(52) CPC특허분류
C07K 14/00 (2013.01)
A61K 38/00 (2013.01)
(21) 출원번호 10-2019-7011467
(22) 출원일자(국제) 2017년09월22일
심사청구일자 없음
(85) 번역문제출일자 2019년04월22일
(86) 국제출원번호 PCT/US2017/052887
(87) 국제공개번호 WO 2018/057847
국제공개일자 2018년03월29일
(30) 우선권주장
62/399,123 2016년09월23일 미국(US)

(71) 출원인
듀크 유니버시티
미국, 노스캐롤라이나 27705, 더럼, 슈트 306, 에
르윈 로드 2812
(72) 발명자
칠르코티, 아슈토슈
미국 27701 노스캐롤라이나주 더럼 글로리아 애비
뉴 1001
탱, 니콜라스
미국 27705 노스캐롤라이나주 더럼 스위트 306 에
어빈 로드 2821 듀크 유니버시티 내
켈리, 가렛
미국 27705 노스캐롤라이나주 더럼 스위트 306 에
어빈 로드 2812 듀크 유니버시티 내
(74) 대리인
양영준, 이상남

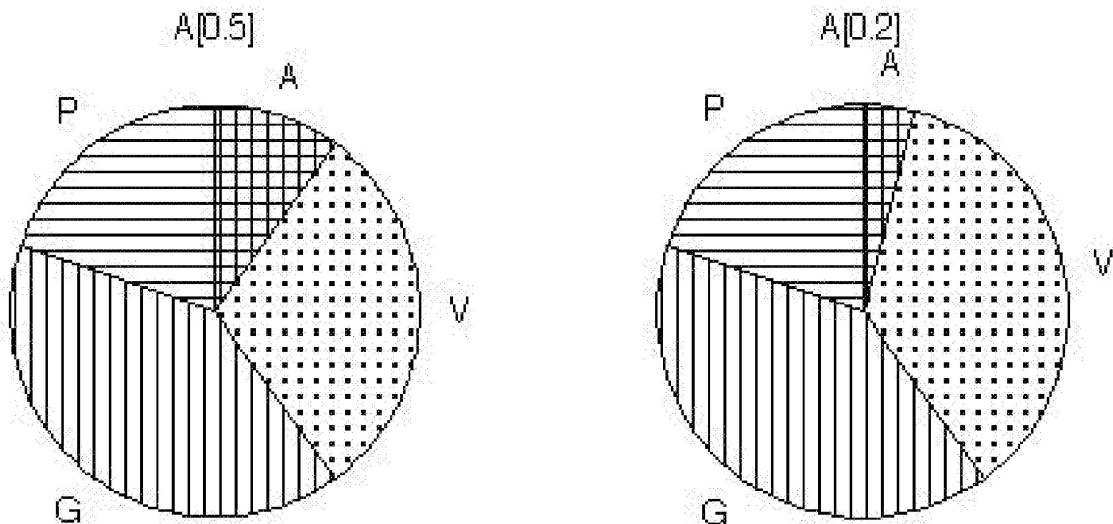
전체 청구항 수 : 총 44 항

(54) 발명의 명칭 LCST 거동을 갖는 비구조화된 비-반복적 폴리펩티드

(57) 요약

본원에서는 임의의 식별가능한 반복 모티프가 결합되어 있는 비구조화된 폴리펩티드를 기술한다. 본원에서는 또 한 비구조화된 폴리펩티드 중 적어도 1개의 것 및 적어도 1개의 결합 폴리펩티드를 포함하는 융합 단백질을 기술한다. 치료를 필요로 하는 대상체에서 질환을 치료하는 방법을 추가로 기술한다. 본 방법은 대상체에게 개시된 융합 단백질의 유효량을 투여하는 단계를 포함할 수 있다.

대표도



(52) CPC특허분류

A61P 35/00 (2018.01)

C07K 14/31 (2013.01)

G01N 33/68 (2013.01)

C07K 2319/00 (2013.01)

C07K 2319/02 (2013.01)

C07K 2319/21 (2013.01)

G01N 2800/52 (2013.01)

명세서

청구범위

청구항 1

식별가능한 반복 모티프를 갖지 않는 비구조화된 폴리펩티드로서, 여기서 폴리펩티드는 하한 임계 용액 온도 (LCST) 미만에서 가용성이거나, 상한 임계 용액 온도 (UCST) 초과에서 가용성이거나, 또는 그의 조합이며, 여기서 LCST 및 UCST는 각각 독립적으로 약 0℃ 내지 약 100℃인 비구조화된 폴리펩티드.

청구항 2

제1항에 있어서, LCST 및 UCST가 각각 독립적으로 약 25℃ 내지 약 37℃인 비구조화된 폴리펩티드.

청구항 3

제1항에 있어서, 비구조화된 폴리펩티드가 적어도 50개 아미노산의 서열을 포함하며,

여기서 아미노산 중 적어도 10%는 프롤린 (P)이고,

여기서 아미노산 중 적어도 20%는 글리신 (G)인

비구조화된 폴리펩티드.

청구항 4

제1항에 있어서, 비구조화된 폴리펩티드가, 아미노산 중 적어도 40%는 발린 (V), 알라닌 (A), 류신 (L), 리신 (K), 트레오닌 (T), 이소류신 (I), 티로신 (Y), 세린 (S), 및 페닐알라닌 (F)으로 이루어진 군으로부터 선택되는 것인 서열을 포함하는 것인 비구조화된 폴리펩티드.

청구항 5

제1항에 있어서, 언어적 복잡성 점수가 15 초과인 비구조화된 폴리펩티드.

청구항 6

제1항에 있어서, 비구조화된 폴리펩티드가 서열식별번호: 7-18 중 임의의 것의 50개 아미노산 하위서열을 포함하는 것인 비구조화된 폴리펩티드.

청구항 7

제1항에 있어서, 비구조화된 폴리펩티드가 3개의 연속적 동일한 아미노산을 함유하지 않는 서열을 포함하며,

여기서 어떠한 5-10개 아미노산 하위서열도 비구조화된 폴리펩티드에서 1회 초과로 발생하지 않고,

여기서 비구조화된 폴리펩티드가 프롤린 (P)으로 개시 및 종결하는 하위서열을 포함할 때, 하위서열은 적어도 1개의 글리신 (G)을 추가로 포함하는 것인

비구조화된 폴리펩티드.

청구항 8

제1항에 있어서, 비구조화된 폴리펩티드가 적어도 50개 아미노산의 서열을 포함하며,

여기서 아미노산 중 적어도 10%는 프롤린 (P)이고,

여기서 아미노산 중 적어도 20%는 글리신 (G)이고,

여기서 아미노산 중 적어도 40%는 발린 (V), 알라닌 (A), 류신 (L), 리신 (K), 트레오닌 (T), 이소류신 (I), 티로신 (Y), 세린 (S), 및 페닐알라닌 (F)으로 이루어진 군으로부터 선택되고,

여기서 서열은 3개의 연속적 동일한 아미노산을 함유하지 않고,

여기서 어떠한 5-10개 아미노산 하위서열도 비구조화된 폴리펩티드에서 1회 초과로 발생하지 않고,
여기서 비구조화된 폴리펩티드가 프롤린 (P)으로 개시 및 종결하는 하위서열을 포함할 때, 하위서열은 적어도 1개의 글리신 (G)을 추가로 포함하는 것인 비구조화된 폴리펩티드.

청구항 9

제1항에 있어서, 서열식별번호: 7-18로부터 선택되는 아미노산 서열을 포함하는 비구조화된 폴리펩티드.

청구항 10

적어도 1개의 결합 폴리펩티드 및 제1항 내지 제9항 중 어느 한 항의 적어도 1개의 비구조화된 폴리펩티드를 포함하는 융합 단백질.

청구항 11

제10항에 있어서, 융합 단백질이 복수개의 비구조화된 폴리펩티드를 포함하는 것인 융합 단백질.

청구항 12

제10항 또는 제11항에 있어서, 융합 단백질이 복수개의 결합 폴리펩티드를 포함하는 것인 융합 단백질.

청구항 13

제12항에 있어서, 적어도 2개의 인접한 결합 폴리펩티드 사이에 위치하는 링커를 추가로 포함하는 융합 단백질.

청구항 14

제12항에 있어서, 적어도 2개의 인접한 비구조화된 폴리펩티드 사이에 위치하는 링커를 추가로 포함하는 융합 단백질.

청구항 15

제13항 또는 제14항에 있어서, 링커가 적어도 1개의 글리신 및 적어도 1개의 세린을 포함하는 것인 융합 단백질.

청구항 16

제15항에 있어서, 링커가 서열식별번호: 21 ((Gly4Ser)₃)로 이루어진 아미노산 서열을 포함하는 것인 융합 단백질.

청구항 17

제13항 또는 제14항에 있어서, 링커가 서열식별번호: 22 (PQPQPKPQPKPEPEPQPG)로 이루어진 아미노산 서열을 포함하는 것인 융합 단백질.

청구항 18

제12항 내지 제17항 중 어느 한 항에 있어서, 복수개의 결합 폴리펩티드가 올리고머를 형성하는 것인 융합 단백질.

청구항 19

제10항 내지 제18항 중 어느 한 항에 있어서, 결합 폴리펩티드가 표적에 결합하고, 융합 단백질이 1개 초과 표적에 결합하는 것인 융합 단백질.

청구항 20

제10항 내지 제19항 중 어느 한 항에 있어서, 결합 폴리펩티드가 단백질 A를 포함하는 것인 융합 단백질.

청구항 21

제10항 내지 제20항 중 어느 한 항에 있어서, 적어도 1개의 비구조화된 폴리펩티드가 열 반응성 폴리펩티드를 포함하는 것인 융합 단백질.

청구항 22

제10항 내지 제21항 중 어느 한 항에 있어서, 적어도 1개의 결합 폴리펩티드와 적어도 1개의 비구조화된 폴리펩티드 사이에 위치하는 적어도 1개의 링커를 추가로 포함하는 융합 단백질.

청구항 23

제22항에 있어서, 융합 단백질이 적어도 1개의 결합 폴리펩티드와 적어도 1개의 비구조화된 폴리펩티드 사이에 복수개의 링커를 포함하는 것인 융합 단백질.

청구항 24

제10항 내지 제23항 중 어느 한 항에 있어서, 적어도 1개의 결합 폴리펩티드가 적어도 1개의 비구조화된 폴리펩티드에 대해 N-말단에 위치하는 것인 융합 단백질.

청구항 25

제10항 내지 제23항 중 어느 한 항에 있어서, 적어도 1개의 결합 폴리펩티드가 적어도 1개의 비구조화된 폴리펩티드에 대해 C-말단에 위치하는 것인 융합 단백질.

청구항 26

치료를 필요로 하는 대상체에서 질환을 치료하는 방법이며, 대상체에게 제10항 내지 제25항 중 어느 한 항에 따른 융합 단백질의 유효량을 투여하는 단계를 포함하는 방법.

청구항 27

제26항에 있어서, 융합 단백질이 제어 방출 제제로 투여되는 것인 방법.

청구항 28

제26항에 있어서, 융합 단백질이 대상체에게 투여 시 세포를 형성하는 것인 방법.

청구항 29

제26항 내지 제28항 중 어느 한 항에 있어서, 융합 단백질이 대상체에게 정맥내로, 동맥내로, 또는 복강내로 투여되는 것인 방법.

청구항 30

제26항 내지 제29항 중 어느 한 항에 있어서, 질환이 암을 포함하는 것인 방법.

청구항 31

제30항에 있어서, 융합 단백질이 종양내로 투여되는 것인 방법.

청구항 32

제26항 내지 제31항 중 어느 한 항에 있어서, 질환이 적어도 1개의 결합 폴리펩티드의 표적과 연관된 질환인 방법.

청구항 33

하기 단계를 포함하는, 대상체에서 질환을 진단하는 방법:

대상체로부터의 샘플을 제10항 내지 제25항 중 어느 한 항에 따른 융합 단백질과 접촉시키는 단계; 및

융합 단백질의 표적에의 결합을 검출하여 샘플 내 표적의 존재를 결정하며, 여기서 샘플 내 표적의 존재는 대상체에서의 질환을 나타내는 것인 단계.

청구항 34

제33항에 있어서, 질환이 암, 대사 질환, 자가면역 질환, 심혈관 질환, 및 정형외과적 장애로부터 선택되는 것인 방법.

청구항 35

하기 단계를 포함하는, 샘플 내 표적의 존재를 결정하는 방법:

샘플을 제10항 내지 제25항 중 어느 한 항의 융합 단백질과, 복합체가 융합 단백질과 샘플 내 표적 사이에 형성되도록 하는 조건 하에 접촉시키는 단계; 및

복합체의 존재를 검출하며, 여기서 복합체의 존재는 샘플 내 표적을 나타내는 것인 단계.

청구항 36

제35항에 있어서, 샘플이 대상체로부터 수득되고, 방법이 질환을 진단하는 것, 예후하는 것, 또는 대상체의 치료의 효능을 사정하는 것을 추가로 포함하는 것인 방법.

청구항 37

제36항에 있어서, 방법이 대상체의 치료의 효능을 사정하는 것을 추가로 포함하는 경우에는, 방법은 효능을 개선시키기 위해 필요에 따라 대상체의 치료를 변형시키는 것을 추가로 포함하는 것인 방법.

청구항 38

하기 단계를 포함하는, 치료를 필요로 하는 대상체에서 질환에 대한 치료의 유효성을 결정하는 방법:

대상체로부터의 샘플을 제10항 내지 제25항 중 어느 한 항의 융합 단백질과, 복합체가 융합 단백질과 샘플 내 표적 사이에 형성되도록 하는 조건 하에 접촉시키는 단계;

샘플 내 복합체의 수준을 결정하며, 여기서 복합체의 수준은 샘플 내 표적의 수준을 나타내는 것인 단계; 및

샘플 내 표적의 수준을 표적의 대조군 수준과 비교하며, 여기서 표적의 수준이 대조군 수준과 상이한 경우에는, 치료가 질환을 치료하는데 있어서 효과적이거나 또는 비효과적인 것으로 결정되는 것인 단계.

청구항 39

하기 단계를 포함하는, 대상체에서 질환을 진단하는 방법:

대상체로부터의 샘플을 제10항 내지 제25항 중 어느 한 항의 융합 단백질과 접촉시키는 단계;

샘플 내 표적의 수준을 결정하는 단계; 및

샘플 내 표적의 수준을 표적의 대조군 수준과 비교하며, 여기서 대조군 수준과 상이한 표적의 수준은 대상체에서의 질환을 나타내는 것인 단계.

청구항 40

제38항 또는 제39항에 있어서, 대조군 수준이 대상체가 치료를 시작한 기간 이전의 또는 그 동안의 시점에서의 대상체 내 수준에 상응하고, 샘플이 이후 시점에 대상체로부터 채취되는 것인 방법.

청구항 41

제38항 또는 제39항에 있어서, 샘플이 대상체가 치료를 받고 있는 기간 동안의 시점에 대상체로부터 채취되고, 대조군 수준이 무질환 수준 또는 대상체가 치료를 시작한 기간 이전의 시점에서의 수준에 상응하는 것인 방법.

청구항 42

제38항 및 제40항 내지 제42항 중 어느 한 항에 있어서, 방법이, 치료가 질환을 치료하는데 있어서 비효과적인 것으로 결정되는 경우에, 치료를 변형시키거나 또는 대상체에게 상이한 치료를 투여하는 것을 추가로 포함하는 것인 방법.

청구항 43

제26항 내지 제42항 중 어느 한 항에 있어서, 융합 단백질이 리포터로 표지되는 것인 방법.

청구항 44

제26항 내지 제43항 중 어느 한 항에 있어서, 질환이 암, 대사 질환, 자가면역 질환, 심혈관 질환, 및 정형외과 적 장애로부터 선택되는 것인 방법.

발명의 설명

기술 분야

[0001]

관련 출원에 대한 상호 참조

[0002]

본 출원은 2016년 9월 23일자 출원된 미국 특허 가출원 번호 62/399,123을 우선권으로 주장하며, 이는 그 전문 이 본원에 참조로 포함된다.

[0003]

연방 정부 지원 연구에 관한 진술

[0004]

본 발명은 미국 국립 보건원(National Institutes of Health)에 의해 수여된 연구 보조금 GM061232 하에 정부 지원으로 이루어졌다. 정부는 본 발명에서 특정 권리를 갖는다.

배경 기술

[0005]

단백질은 특이성에 대해, 및 임상 표적에 대한 선택성에 대해 조작되었을 때 유용한 치료제가 될 수 있다. 단백질은 대개는 그의 복잡성, 다목적성, 내약성 및 다양성에 기인하여 소분자 약물에 대한 우수한 대안이 될 수 있고, 긴 반감기, 특이성, 및 선택성에 기인하여 요법을 위한 관심의 대상이 될 수 있다. 비록 단백질 조작을 통해서 관심 단백질 또는 수용체에 대해 표적화되는 강력한 치료제를 개발할 수 있지만, 신체는 상기 단백질 요법을 소거시키는 많은 메카니즘을 갖고 있다. 따라서, 신뢰할 수 있고, 광범위하게 적용될 수 있는 단백질 전달 해결 방안에 대한 필요가 있다.

발명의 내용

[0006]

한 측면에서, 식별가능한 반복 모티프를 갖지 않는 비구조화된 폴리펩티드로서, 여기서 폴리펩티드는 하한 임계 용액 온도 (LCST) 미만에서 가용성이거나, 상한 임계 용액 온도 (UCST) 초과에서 가용성이거나, 또는 그의 조합 이며, 여기서 LCST 및 UCST는 각각 독립적으로 약 0°C 내지 약 100°C인 비구조화된 폴리펩티드를 개시한다.

[0007]

또 다른 측면에서, 적어도 1개의 결합 폴리펩티드 및 개시된 비구조화된 폴리펩티드 중 적어도 1개를 포함하는 융합 단백질을 개시한다.

[0008]

또 다른 측면에서, 치료를 필요로 하는 대상체에서 질환을 치료하는 방법이며, 대상체에게 개시된 융합 단백질의 유효량을 투여하는 단계를 포함하는 방법을 개시한다.

[0009]

또 다른 측면에서, 대상체에서 질환을 진단하는 방법이며, 대상체로부터의 샘플을 개시된 융합 단백질과 접촉시키는 단계; 및 융합 단백질의 표적에의 결합을 검출하여 샘플 내 표적의 존재를 결정하며, 여기서 샘플 내 표적의 존재는 대상체에서의 질환을 나타내는 것인 단계를 포함하는 방법을 개시한다.

[0010]

또 다른 측면에서, 샘플 내 표적의 존재를 결정하는 방법이며, 샘플을 개시된 융합 단백질과, 복합체가 융합 단백질과 샘플 내 표적 사이에 형성되도록 하는 조건 하에 접촉시키는 단계; 및 복합체의 존재를 검출하며, 여기서 복합체의 존재는 샘플 내 표적을 나타내는 것인 단계를 포함하는 방법을 개시한다.

[0011]

또 다른 측면에서, 치료를 필요로 하는 대상체에서 질환에 대한 치료의 유효성을 결정하는 방법이며, 대상체로부터의 샘플을 개시된 융합 단백질과, 복합체가 융합 단백질과 샘플 내 표적 사이에 형성되도록 하는 조건 하에 접촉시키는 단계; 샘플 내 복합체의 수준을 결정하며, 여기서 복합체의 수준은 샘플 내 표적의 수준을 나타내는 것인 단계; 및 샘플 내 표적의 수준을 표적의 대조군 수준과 비교하며, 여기서 표적의 수준이 대조군 수준과 상이한 경우에는, 치료가 질환을 치료하는데 있어서 효과적이거나 또는 비효과적인 것으로 결정되는 것인 단계를 포함하는 방법을 개시한다.

[0012]

또 다른 측면에서, 대상체에서 질환을 진단하는 방법이며, 대상체로부터의 샘플을 개시된 융합 단백질과 접촉시

키는 단계; 샘플 내 표적의 수준을 결정하는 단계; 및 샘플 내 표적의 수준을 표적의 대조군 수준과 비교하며, 여기서 대조군 수준과 상이한 표적의 수준은 대상체에서의 질환을 나타내는 것인 단계를 포함하는 방법을 개시한다.

도면의 간단한 설명

[0013]

도 1은 A[0.5] 및 A[0.2] 서열 (서열식별번호(SEQ ID NO): 1 및 서열식별번호: 2)의 아미노산 조성을 제시하는 다이어그램이다. 상응하여, 다이어그램은 또한 각각 서열식별번호: 4, 6, 8, 및 10, 및 서열식별번호: 3, 5, 7, 및 9의 폴리펩티드의 아미노산 조성을 제시한다.

도 2는 예시적인 비-반복적 비구조화된 폴리펩티드 (서열식별번호: 11-18)의 아미노산 조성을 제시하는 다이어그램이다. 각각의 비-반복적 비구조화된 폴리펩티드는 240개 아미노산을 포함한다. 각각의 아미노산 서열은 대략 1/6 프롤린 (P) 잔기, 대략 1/3 글리신 (G) 잔기, 및 대략 1/2 X 잔기를 포함하며, 여기서 X는 발린 (V), 알라닌 (A), 류신 (L), 리신 (K), 트레오닌 (T), 이소류신 (I), 티로신 (Y), 세린 (S), 및 페닐알라닌 (F)으로 이루어진 군으로부터 선택되는 1개 이상의 아미노산이다. X에 대해 선택된 각각의 아미노산은 서로 동일한 빈도로 발생할 수 있다 (서열식별번호: 11-18).

도 3a-b는 반복적 (서열식별번호: 3-6) 및 예시적인 비-반복적 폴리펩티드 (서열식별번호: 7-10)의 전이 온도의 특징화를 제시하는 그래프이다. 도 3a는 200개 아미노산을 포함하는 반복적 및 예시적인 비-반복적 폴리펩티드를 보여주는 그래프이다. 도 3b는 400개 아미노산을 포함하는 반복적 및 예시적인 비-반복적 폴리펩티드를 제시하는 그래프이다.

도 4a-b는 예시적인 비-반복적 비구조화된 폴리펩티드 (서열식별번호: 11-18)의 전이 온도의 특징화를 제시하는 그래프이다. 도 4a는 PBS 중 25 μ M의 예시적인 비-반복적 비구조화된 폴리펩티드의 전이 온도의 특징화를 제시하는 그래프이다. 도 4b는 PBS 중 다양한 농도의 우레아에서의 예시적인 비-반복적 비구조화된 폴리펩티드의 전이 온도의 특징화를 제시하는 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0014]

엘라스틴-유사 폴리펩티드 (ELP)는 반복적 폴리펩티드이다. "ELP"는 펜타펩티드 반복 서열 (VPGXG)_n을 포함하는 폴리펩티드를 지칭하며, 여기서 X는 프롤린을 제외한 임의의 아미노산이고, n은 1 이상의 정수이다 (서열식별번호: 23). ELP는 연구되어 왔고, 하한 임계 용액 온도 (LCST) 거동을 갖는 것으로 특징화되어 있다. ELP는 예를 들어, 본원에서 각각 A[0.5] 및 A[0.2]로 지칭되는, GAGVPGVGVP (서열식별번호: 1) 또는 GVGVPGVPGAGVPGVGVP (서열식별번호: 2)의 반복 하위서열을 포함할 수 있다 (문헌 [McDaniel, J. R. et al. (2013) *Biomacromolecules*] 참조, 상기 문헌은 그 전문이 본원에 참조로 포함됨). 예를 들어, A[0.2] rep-200 (서열식별번호: 3, (GVGVPGVPGAGVPGVGVP)₈)은 총 200개 아미노산을 위해 8회 반복되는 A[0.2] 하위서열을 포함한다. A[0.5] rep-200 (서열식별번호: 4, (GAGVPGVGVP)₂₀)은 총 200개 아미노산을 위해 20회 반복되는 A[0.5] 하위서열을 포함한다. A[0.2] rep-400 (서열식별번호: 5, (GVGVPGVPGAGVPGVGVP)₁₆)은 총 400개 아미노산을 위해 16회 반복되는 A[0.2] 하위서열을 포함한다. A[0.5] rep-400 (서열식별번호: 6, (GAGVPGVGVP)₄₀)은 총 400개 아미노산을 위해 40회 반복되는 A[0.5] 하위서열을 포함한다. 상기 서열의 아미노산 조성은 도 1에 도시되어 있다.

[0015]

본원에서는 ELP의 필수 펜타펩티드 서열은 결여되어 있지만, 예상 밖으로 LCST 거동은 여전히 갖는, 비구조화된, 비-반복적 폴리펩티드를 개시한다. 따라서, 개시된 비구조화된 폴리펩티드는 모두 폴리펩티드 내에 식별가능한 반복적 서열을 갖지 않으며, 이는 (CD에 따르면) 2차 구조가 결여되어 있고, 열 반응성이다.

[0016]

1. 정의

[0017]

달리 정의되지 않는 한, 본원에서 사용된 모든 기술 및 과학 용어는 관련 기술분야의 통상의 기술자가 보편적으로 이해하고 있는 것과 동일한 의미를 가진다. 상충되는 경우에, 정의를 포함한 본 명세서가 우선할 것이다. 비록 본원에 기술된 것과 유사하거나 또는 등가인 방법 및 물질이 본 발명의 실시 또는 시험에서 사용될 수 있지만, 바람직한 방법 및 물질은 하기에 기술된다. 본원에서 언급된 모든 공개, 특허 출원, 특허 및 다른 참고 문헌은 그의 전문이 참조로 포함된다. 본원에 개시된 물질, 방법, 및 예는 단지 예시적인 것이며, 제한하려는 것으로 의도되지 않는다.

- [0018] 본원에 사용된 바와 같이, "포함하다(한다)", "수반하다(한다)", "갖는", "갖는다", 할 수 있다", "함유하다(한다)"라는 용어, 및 그의 파생어는 추가 행위 또는 구조의 가능성을 배제하지 않는, 개방형 연결구, 용어, 또는 단어인 것으로 의도된다. 단수 형태는 문맥상 명확하게 달리 명시되지 않는 한, 복수 지시대상을 포함한다. 본 개시내용은 또한 명확하게 기술되어 있는지 여부와 상관없이, 본원에 제시된 실시양태 또는 요소를 "포함하는", "로 이루어진" 및 "로 본질적으로 이루어진" 다른 실시양태를 고려한다.
- [0019] 본원에 달리 나타내지 않는 한, 본원에서 값을 범위로 열거한 것은 단지 상기 범위 내에 포함된 각각의 개별 값을 개별적으로 지칭하는 것에 대한 속기법으로서의 역할을 하는 것으로 의도되며, 각각의 개별 값은 마치 그 값이 본원에서 개별적으로 열거된 것처럼 본 명세서에 포함된다. 예를 들어, 농도 범위가 1% 내지 50%인 것으로 언급되는 경우에, 이는 예컨대 2% 내지 40%, 10% 내지 30%, 또는 1% 내지 3% 등과 같은 값이 본 명세서에서 명백하게 열거된 것으로 의도된다. 이는 단지 구체적으로 의도되는 것의 한 예일 뿐이며, 열거된 최저 값과 최고 값 사이 및 그 값을 포함하는 수치 값의 모든 가능한 조합도 본 명세서에서 명백하게 언급된 것으로 간주되어야 한다.
- [0020] 수량과 관련하여 사용되는 "약"이라는 수식어는 언급된 값을 포함하며, 문맥상 명시되는 의미를 가진다 (예를 들어, 이는 적어도 특정한 수량의 측정과 연관된 오차 정도를 포함함). "약"이라는 수식어는 또한 두 종점의 절대 값으로 정의되는 범위를 개시하는 것으로 간주되어야 한다. 예를 들어, "약 2 내지 약 4"라는 표현은 또한 "2부터 4까지"의 범위를 개시한다. "약"이라는 용어는 나타낸 수치의 $\pm 10\%$ 를 지칭할 수 있다. 예를 들어, "약 10%"는 9% 내지 11%의 범위를 나타낼 수 있고, "약 1"은 0.9-1.1을 의미할 수 있다. "약"의 다른 의미는 예컨대 반올림과 같이 문맥으로부터 자명해질 수 있으며, 이에, 예를 들어, "약 1"은 또한 0.5부터 1.4까지를 의미할 수 있다.
- [0021] 본원에 사용된 바와 같이, "친화도"라는 용어는 결합 폴리펩티드의 그의 표적 (즉, 결합 파트너)에의 결합 강도를 지칭한다.
- [0022] 본원에 사용된 바와 같이, "효능제"라는 용어는 수용체에 결합하고, 수용체를 활성화시켜 생물학적 반응을 일으키는 엔티티를 지칭한다. "길항제"는 효능제의 작용 또는 신호전달을 차단하거나 또는 억제한다. "역 효능제"는 효능제 작용에 반대되는 작용을 일으킨다. 효능제, 길항제, 및 역 효능제의 활성화는 시험관내, 체내, 생체내, 또는 그의 조합에서 결정될 수 있다.
- [0023] 본원에 사용된 바와 같이, "바이오마커"라는 용어는 질환 또는 병태의 확인 및/또는 분류에 유용한 다양한 농도로 대상체에 존재하는 자연 발생 생물학적 분자를 지칭한다. 바이오마커는 질환에 대한 지표 또는 마커로서 사용되는 유전자, 단백질, 폴리뉴클레오티드, 핵산, 리보핵산, 폴리펩티드, 또는 다른 생물학적 분자를 포함할 수 있다. 일부 실시양태에서, 바이오마커는 질환 마커를 포함한다. 예를 들어, 바이오마커는 질환을 앓는 대상체에서 상향조절되어 있거나 또는 하향조절되어 있는 유전자일 수 있다. 또 다른 예로서, 바이오마커는 질환을 앓거나 또는 질환이 발생될 위험이 있는 대상체에서 그 수준이 증가되어 있거나 또는 감소되어 있는 폴리펩티드일 수 있다. 일부 실시양태에서, 바이오마커는 소분자를 포함한다. 일부 실시양태에서, 바이오마커는 폴리펩티드를 포함한다.
- [0024] 본원에 사용된 바와 같이, "대조군", "참조 수준" 및 "참조"라는 용어는 본원에서 상호교환적으로 사용된다. 참조 수준은 미리 결정된 값 또는 범위일 수 있으며, 이는 측정된 결과를 사정하기 위한 벤치마크로서 사용된다. 본원에 사용된 바와 같이, "대조군 그룹"이란, 대조군 대상체 그룹을 지칭한다. 미리 결정된 수준은 대조군 그룹으로부터의 컷오프 값일 수 있다. 미리 결정된 수준은 대조군 그룹으로부터의 평균일 수 있다. 컷오프 값 (또는 미리 결정된 컷오프 값)은 적응 지수 모델 (AIM) 방법에 의해 결정될 수 있다. 컷오프 값 (또는 미리 결정된 컷오프 값)은 환자 그룹의 생물학적 샘플로부터의 수신자 조작 곡선 (ROC) 분석에 의해 결정될 수 있다. 생물학 분야에서 일반적으로 공지된 바와 같은 ROC 분석은 하나의 상태를 또 다른 상태와 구별할 수 있는, 예컨대 CRC를 앓는 환자의 확인에서 각 마커의 성능을 결정할 수 있는 시험 능력의 결정이다. ROC 분석의 설명은 문헌 [P.J. Heagerty et al. (*Biometrics* 2000, 56, 337-44)]에 제공되어 있으며, 상기 문헌의 개시내용은 그 전문이 본원에 참조로 포함된다. 대안적으로, 컷오프 값은 환자 그룹의 생물학적 샘플의 사분위수 분석에 의해 결정될 수 있다. 예를 들어, 컷오프 값은 제25-제75 백분위수 범위의 임의의 값에 상응하는 값, 바람직하게, 제25 백분위수, 제50 백분위수 또는 제75 백분위수, 및 더욱 바람직하게, 제75 백분위수에 상응하는 값을 선택함으로써 결정될 수 있다. 이러한 통계 분석은 관련 기술분야에 공지된 임의의 방법을 사용하여 수행될 수 있고, 임의 개수의 상업적으로 이용가능한 소프트웨어 패키지 (예컨대 애널리즈-잇 소프트웨어 리미티드(Analyse-it Software Ltd.: 영국 리즈); 스타타코프 엘피(StataCorp LP: 미국 텍사스주 칼리지 스테이

선); 에스에이에스 인스티튜드 인크.(SAS Institute Inc.: 미국 노스캐롤라이나주 캐리))를 통해 구현될 수 있다. 표적에 대한 또는 단백질 활성화에 대한 건강한 또는 정상 수준 또는 범위는 표준 관행에 따라 정의될 수 있다.

[0025] 본원에 사용된 바와 같이, "발현 벡터"라는 용어는 원하는 단백질을 코딩하는 핵산 서열이 삽입되거나 또는 도입될 수 있는, 관련 기술분야에 공지된 플라스미드, 바이러스 또는 또 다른 매체를 가리킨다.

[0026] 본원에 사용된 바와 같이, "숙주 세포"라는 용어는 핵산 구축물 또는 발현 벡터를 이용한 형질전환, 형질감염, 형질도입, 접합 등에 감수성인 세포이다. 숙주 세포는 식물, 박테리아, 효모, 진균, 곤충, 동물 등으로부터 유래될 수 있다. 일부 실시양태에서, 숙주 세포는 에스케리키아 콜라이(*Escherichia coli*)를 포함한다.

[0027] 본원에 사용된 바와 같이, "리포터", "리포터 기", "표지" 및 "검출가능한 표지"라는 용어는 본원에서 상호교환적으로 사용된다. 리포터는 검출가능한 신호를 생성할 수 있다. 표지는 시각적 또는 기구적 수단에 의해 검출가능한 신호를 생성할 수 있다. 리포터 기의 신호 전달의 물리적 성질 (예컨대 형광, 전기화학, 핵 자기 공명 (NMR), 및 전자 상자성 공명 (EPR)) 및 화학적 성질이 상이한 다양한 리포터 기가 사용될 수 있다. 다양한 리포터는 신호 생성 물질, 예컨대 색원체, 형광 화합물, 화학발광 화합물, 방사성 화합물 등을 포함한다. 일부 실시양태에서, 리포터는 방사선 표지를 포함한다. 리포터는 빛을 생성하는 모이어티, 예컨대 아크리디늄 화합물, 및 형광을 생성하는 모이어티, 예컨대 플루오레세인을 포함할 수 있다. 일부 실시양태에서, 리포터로부터의 신호는 형광 신호이다. 리포터는 형광단을 포함할 수 있다. 형광단의 예는 아크릴로단 (6-아크릴로일-2-디메틸아미노나프탈렌), 바단 (6-브로모-아세틸-2-디메틸아미노-나프탈렌), 로다민, 나프탈렌, 단질 아지리딘, 4-[N-[(2-아이오도아세톡시)에틸]-N-메틸아미노]-7-니트로벤즈-2-옥사-1,3-디아졸 에스테르 (IANBDE), 4-[N-[(2-아이오도아세톡시)에틸]-N-메틸아미노-7-니트로벤즈-2-옥사-1,3-디아졸 (IANBDA), 플루오레신, 디피로메텐보론 디플루오라이드 (BODIPY), 4-니트로벤조[c][1,2,5]옥사디아졸(NBD), 알렉사(Alexa) 형광 염료, 및 그의 유도체를 포함하나 이에 제한되지는 않는다. 플루오레세인 유도체는, 예를 들어, 5-플루오레세인, 6-카르복시플루오레세인, 3'(6)-카르복시플루오레세인, 5(6)-카르복시플루오레세인, 6-헥사클로로플루오레세인, 6-테트라클로로플루오레세인, 플루오레세인, 및 이소티오시아네이트를 포함할 수 있다.

[0028] 본원에 사용된 바와 같이, "샘플" 또는 "시험 샘플"이라는 용어는 그 중의 표적의 존재 및/또는 수준이 검출되거나 또는 결정되는 임의의 샘플을 의미할 수 있다. 샘플은 액체, 용액, 에멀전, 또는 현탁액을 포함할 수 있다. 샘플은 의학 샘플을 포함할 수 있다. 샘플은 임의의 생물학적 유체 또는 조직, 예컨대 혈액, 전혈, 혈액의 분획, 예컨대 혈장 및 혈청, 근육, 장액, 땀, 타액, 소변, 눈물, 윤활액, 골수, 뇌척수액, 코 분비물, 객담, 양수, 기관지폐포 세척액, 위 세척액, 구토, 분변물, 폐 조직, 말초 혈액 단핵 세포, 전체 백혈구, 림프절 세포, 비장 세포, 편도 세포, 암 세포, 종양 세포, 담즙, 소화액, 피부, 또는 그의 조합을 포함할 수 있다. 일부 실시양태에서, 샘플은 분취물을 포함한다. 다른 실시양태에서, 샘플은 생물학적 유체를 포함한다. 샘플은 관련 기술분야에 공지된 임의의 수단에 의해 획득될 수 있다. 샘플은 환자로부터 획득된 그대로 직접 사용될 수도 있거나, 또는 본원에서 논의된 바와 같은 일부 방식으로 또는 다르게는 관련 기술분야에 공지된 바와 같이 샘플의 특징을 변형시키기 위해 예컨대 여과, 증류, 추출, 농축, 원심분리, 간접 성분의 불활성화, 시약의 첨가 등에 의해 미리 처리될 수 있다.

[0029] 본원에 사용된 바와 같이, 본원에 사용된 "대상체"라는 용어는 본원에 기술된 융합 단백질을 원하거나 또는 그를 필요로 하는 포유동물을 의미할 수 있다. 대상체는 인간 또는 비-인간 동물일 수 있다. 대상체는 포유동물일 수 있다. 포유동물은 영장류 또는 비-영장류일 수 있다. 포유동물은 영장류, 예컨대 인간; 비-영장류, 예컨대 개, 고양이, 말, 소, 돼지, 마우스, 래트, 낙타, 라마, 염소, 토끼, 양, 햄스터, 및 기니 피그; 또는 비-인간 영장류, 예컨대 예를 들어, 원숭이, 침팬지, 고릴라, 오랑우탄, 및 긴팔원숭이일 수 있다. 대상체는 임의의 연령 또는 발달 단계, 예컨대, 예를 들어, 성인, 청소년, 또는 유아일 수 있다.

[0030] 본원에 사용된 바와 같이, "전이" 또는 "상 전이"라는 용어는 열 반응성 폴리펩티드의 응집을 지칭한다. 상 전이는 LCST 또는 역 전이 온도 T_i 로 불리는 특정 온도에서 급격하게 및 가역적으로 발생한다. 전이 온도 (LCST 또는 T_i) 미만에서, 열 반응성 폴리펩티드는 고도로 가용성이다. 전이 온도 초과로 가열되었을 때, 열 반응성 폴리펩티드는 소수성 방식으로 붕괴되고, 응집하여, 별도의 겔-유사 상을 형성한다. 상 전이 거동은 폴리펩티드의 조절형 및/또는 저속형 방출을 위하여 대상체의 조직 내에서의 약물 배포를 형성하기 위해 사용될 수 있다. "역 전이 사이클링"은 열 반응성 폴리펩티드에 대한 단백질 정제 방법을 지칭한다. 단백질 정제 방법은 가용성 상 및 불용성 상을 통해 용액을 사이클링하기 위해 열 반응성 폴리펩티드의 가역적 상 전이 거동의 사용

을 수반할 수 있고, 이로써 오염물질을 제거할 수 있고, 크로마토그래피에 대한 필요를 제거할 수 있다.

- [0031] 본원에 사용된 바와 같이, "하위서열"이라는 용어는 연속적 아미노산의 또 다른 서열 내에 발생하는 연속적 아미노산의 서열을 지칭한다. 하위서열은 적어도 2개 아미노산을 포함한다. 일부 실시양태에서, 하위서열의 길이는 2 내지 50개, 2 내지 20개, 2 내지 15개, 또는 2 내지 10개의 연속 아미노산 길이이다. 일부 실시양태에서, 하위서열은 3, 4, 5, 6, 7, 8, 9, 또는 10개의 연속 아미노산을 포함한다.
- [0032] 본원에 사용된 바와 같이, "실질적으로 동일한"이라는 용어는 제1 및 제2 아미노산 서열이 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 또는 1100개 아미노산의 영역에 걸쳐 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99%라는 것을 의미할 수 있다.
- [0033] 본원에 사용된 바와 같이, "치료" 또는 "치료하는"이라는 용어는 질환으로부터 대상체를 보호하는 것을 지칭할 경우에, 질환을 예방, 억제, 억압, 호전, 또는 완전히 제거하는 것을 의미한다. 질환을 예방하는 것은 질환의 발병 이전에 대상체에 본 발명의 조성물을 투여하는 것을 수반한다. 질환을 억제하는 것은 질환의 유도 이후에, 그러나, 그의 임상적 출현 이전에 대상체에게 본 발명의 조성물을 투여하는 것을 수반한다. 질환을 억압하거나 또는 호전시키는 것은 질환의 임상적 출현 이후에 대상체에게 본 발명의 조성물을 투여하는 것을 수반한다.
- [0034] 본원에 사용된 바와 같이, "결합가"라는 용어는 잠재적인 결합 단위 또는 결합 부위를 지칭한다. "다가"라는 용어는 다중의 잠재적인 결합 단위를 지칭한다. "다량체" 및 "다가"라는 용어는 본원에서 상호교환적으로 사용된다.
- [0035] 본원에 사용된 바와 같이, 폴리뉴클레오티드와 관련하여 "변이체"라는 용어는 (i) 참조되는 뉴클레오티드 서열의 일부 또는 단편; (ii) 참조되는 뉴클레오티드 서열 또는 그의 일부의 상보체; (iii) 참조되는 폴리뉴클레오티드 또는 그의 상보체와 실질적으로 동일한 폴리뉴클레오티드; 또는 (iv) 엄격한 조건 하에, 참조되는 폴리뉴클레오티드, 그의 상보체, 또는 그와 실질적으로 동일한 서열에 하이브리드화하는 폴리뉴클레오티드를 의미한다.
- [0036] "변이체"는 아미노산의 삽입, 결실 또는 보존적 치환에 의해 아미노산 서열이 상이하지만, 적어도 하나의 생물학적 활성을 보유하는 펩티드 또는 폴리펩티드인 것으로 추가로 정의될 수 있다. "생물학적 활성"의 대표적인 예는 특정 항체 또는 폴리펩티드에 의해 결합되거나 또는 면역 반응을 촉진시킬 수 있는 능력을 포함한다. 변이체는 실질적으로 동일한 서열을 의미할 수 있다. 변이체는 그의 기능적 단편을 의미할 수 있다. 변이체는 또한 폴리펩티드의 다중의 카피를 의미할 수 있다. 다중의 카피는 직렬식으로 있거나, 링커에 의해 이격되어 있을 수 있다. 변이체는 또한 적어도 하나의 생물학적 활성을 보유하는 아미노산 서열을 갖는 참조되는 폴리펩티드와 실질적으로 동일한 아미노산 서열을 갖는 폴리펩티드를 의미할 수 있다. 아미노산의 보존적 치환, 즉, 유사한 특성 (예컨대 친수성, 하전된 영역의 정도 및 분포)을 갖는 상이한 아미노산으로의 아미노산의 대체는 관련 기술분야에서 전형적으로 작은 변화를 수반하는 것으로 인식된다. 이러한 작은 변화는 부분적으로, 아미노산의 소수친수성 지수를 고려함으로써 확인될 수 있다. 문헌 [Kyte et al., *J. Mol. Biol.* 1982, 157, 105-132]를 참조한다. 아미노산의 소수친수성 지수는 소수성 및 전하를 고려하는 것에 기초한다. 소수친수성 지수가 유사한 아미노산이 치환되고, 여전히 그대로 단백질 기능을 보유할 수 있다고 관련 기술분야에 공지되어 있다. 한 측면에서, 소수친수성 지수가 ± 2 인 아미노산이 치환된다. 아미노산의 소수성은 또한 생물학적 기능을 보유하는 폴리펩티드를 생성하는 치환을 보이는데 사용될 수 있다. 미국 특허 번호 4,554,101 (상기 특허는 그 전체가 본원에 참조로 포함됨)에서 논의된 바와 같이, 폴리펩티드와 관련하여 아미노산의 친수성을 고려해 봄으로써 항원성 및 면역원성과 우수한 연관성을 갖는 것으로 보고된 유용한 척도인, 상기 폴리펩티드의 최대 국소 평균 친수성을 계산해 낼 수 있게 된다. 관련 기술분야에서 이해되는 바와 같이, 친수성 값이 유사한 아미노산의 치환을 통해 생물학적 활성, 예를 들어, 면역원성을 보유하는 폴리펩티드를 얻을 수 있다. 치환은 서로 친수성 값이 ± 2 내에 있는 아미노산으로 수행될 수 있다. 아미노산의 소수성 지수 및 친수성 값 둘 다는 상기 아미노산의 특정 측쇄에 의해 영향을 받는다. 상기 관찰결과와 일관되게, 생물학적 기능과 상용성인 아미노산 치환은 소수성, 친수성, 전하, 크기, 및 다른 특성으로 나타나는 바와 같이, 아미노산의 상대적 유사성, 특히, 아미노산의 측쇄에 의존하는 것으로 이해된다.
- [0037] 변이체는 전장의 전체 유전자 서열, 또는 그의 단편에 걸쳐 실질적으로 동일한 폴리뉴클레오티드 서열일 수 있다. 폴리뉴클레오티드 서열은 전장의 유전자 서열 또는 그의 단편에 걸쳐 80%, 81%, 82%, 83%, 84%, 85%,

86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일할 수 있다. 변이체는 전장의 아미노산 서열, 또는 그의 단편에 걸쳐 실질적으로 동일한 아미노산 서열일 수 있다. 아미노산 서열은 전장의 아미노산 서열, 또는 그의 단편에 걸쳐 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일할 수 있다.

2. 비구조화된 폴리펩티드

본원에서는 비구조화된 폴리펩티드를 개시한다. 비구조화된 폴리펩티드는 CD에 의해 관찰된 바, 2차 구조가 결여되어 있고, 임의의 식별가능한 반복적 서열이 결여되어 있는, 임의의 적합한 폴리펩티드를 포함할 수 있다. 추가로, 비구조화된 폴리펩티드는 주어진 농도에서 그의 LCST 미만 및/또는 그의 UCST 초과에서 가용성일 수 있으며, 이로써 폴리펩티드에 상 전이 특징을 부여할 수 있고, 그 결과, "열 반응성 폴리펩티드"로서 지칭될 수 있다. LCST는 그보다 낮은 온도에서 폴리펩티드가 혼화성인 온도이다. UCST는 그보다 높은 온도에서 폴리펩티드가 혼화성인 온도이다. LCST는 UCST와 관련하여 분리된 별개의 온도이다. 일부 실시양태에서, 비구조화된 폴리펩티드는 오직 UCST 거동만을 가질 수 있다. 일부 실시양태에서, 비구조화된 폴리펩티드는 오직 LCST 거동만을 가질 수 있다. 일부 실시양태에서, 비구조화된 폴리펩티드는 UCST 및 LCST 거동 둘 다를 가질 수 있다. 상기 실시양태에서, UCST는 LCST보다 높다. 비구조화된 폴리펩티드의 LCST는 약 0℃ 내지 약 100℃, 예컨대 약 10℃ 내지 약 50℃, 또는 약 20℃ 내지 약 42℃일 수 있다. 비구조화된 폴리펩티드의 UCST는 약 0℃ 내지 약 100℃, 예컨대 약 10℃ 내지 약 50℃, 또는 약 20℃ 내지 약 42℃일 수 있다. 일부 실시양태에서, 비구조화된 폴리펩티드의 전이 온도(들) (LCST 및/또는 UCST)는 실온(약 25℃) 내지 체온(약 37℃)일 수 있다. 비구조화된 폴리펩티드는 비구조화된 폴리펩티드가 대상체에게 투여되는 농도에서 체온 미만 또는 체온 초과인 그의 LCST 및/또는 UCST를 가질 수 있다. 열 반응성 비구조화된 폴리펩티드는 다양한 온도 및 농도에서 상 전이될 수 있다. 열 반응성 비구조화된 폴리펩티드는 제2 폴리펩티드의, 그의 접합 대상에의 결합 또는 효능에는 영향을 주지 않을 수 있다. 추가로, 열 반응성 비구조화된 폴리펩티드는 임의의 개수의 원하는 전이 온도, 분자량, 및 포맷으로 조정될 수 있다.

비구조화된 폴리펩티드는 다양한 양 및 유형의 아미노산을 포함할 수 있다. 예를 들어, 비구조화된 폴리펩티드는, 아미노산 중 적어도 10%가 프롤린 (P)이고, 아미노산 중 적어도 20%가 글리신 (G)인, 적어도 50개 아미노산의 서열을 포함할 수 있다. 일부 실시양태에서, 비구조화된 폴리펩티드는 아미노산 중 적어도 40%가 발린 (V), 알라닌 (A), 류신 (L), 리신 (K), 트레오닌 (T), 이소류신 (I), 티로신 (Y), 세린 (S), 및 페닐알라닌 (F)으로 이루어진 군으로부터 선택되는 것인 서열을 포함할 수 있다. 일부 실시양태에서, 비구조화된 폴리펩티드는, 어떠한 5-10개 아미노산 하위서열도 비구조화된 폴리펩티드에서 1회 초과로 발생하지 않고, 비구조화된 폴리펩티드가 프롤린 (P)으로 개시 및 종결하는 하위서열을 포함할 때, 하위서열은 적어도 1개의 글리신 (G)을 추가로 포함하는 것인, 3개의 연속적 동일한 아미노산을 함유하지 않는 서열을 포함할 수 있다.

일부 실시양태에서, 비구조화된 폴리펩티드는, 아미노산 중 적어도 10%가 프롤린 (P)이고; 아미노산 중 적어도 20%가 글리신 (G)이고; 아미노산 중 적어도 40%가 발린 (V), 알라닌 (A), 류신 (L), 리신 (K), 트레오닌 (T), 이소류신 (I), 티로신 (Y), 세린 (S), 및 페닐알라닌 (F)으로 이루어진 군으로부터 선택되고; 서열은 3개의 연속적 동일한 아미노산을 함유하지 않고; 어떠한 5-10개 아미노산 하위서열도 비구조화된 폴리펩티드에서 1회 초과로 발생하지 않고; 비구조화된 폴리펩티드가 프롤린 (P)으로 개시 및 종결하는 하위서열을 포함할 때, 하위서열은 적어도 1개의 글리신 (G)을 추가로 포함하는 것인, 적어도 50개 아미노산의 서열을 포함할 수 있다.

(하기 실시예에서 논의되는 바와 같은) 컴퓨터 알고리즘에 의해 확인되는 비-반복적 서열 중 임의의 것의 더 짧은 하위서열은 열 반응성 거동을 유지할 수 있다. 상기 언급된 동일한 서열 규칙에 의해 제약을 받기 때문에, 장의 서열과 유사한 조성이 충분히 비-반복적일 경우에, 50개 아미노산 정도로 짧은 하위서열을 포함하는 비구조화된 폴리펩티드가 알고리즘의 요건을 충족시킬 수 있다. 선행 연구에서는 본질적 무질서는 길이가 50개 미만의 아미노산 길이인 비구조화된 펩티드에 의해 코딩될 수 있는 것으로 제시된 바 있다 (문헌 [Radivojac et al. (2007) Biophys J.] 참조, 상기 문헌은 그 전문이 본원에 참조로 포함됨). 추가로, 50개 아미노산 정도로 짧은 ELP는 열 반응성 LCST 거동을 보일 수 있으며, 여기서 측정된 전이 온도는 0 내지 100℃의 범위이다 (문헌 [Aladini et al. (2016) J Pept Sci.] 참조, 상기 문헌은 그 전문이 본원에 참조로 포함됨). 따라서, 일부 실시양태에서, 비구조화된 폴리펩티드는 서열식별번호: 7-18 중 임의의 것의 50개 아미노산 하위서열을 포함할 수 있다.

상기 언급된 바와 같이, 비구조화된 폴리펩티드에는 아미노산의 식별가능한 반복 모티프가 결여되어 있을 수 있다. 비구조화된 폴리펩티드의 반복성은 그의 언어적 복잡성 점수를 특징으로 할 수 있다. 언어적 복잡성 점수

는 주어진 서열 중의 독특한 하위서열의 총 개수를 동일한 알파벳 및 윈도우 길이에 대하여 가능한 독특한 하위서열의 총 개수로 나눈 값으로 정의된다 (문헌 [Troyanskaya et al. (2002) Bioinformatics] 참조, 상기 문헌은 그 전문이 본원에 참조로 포함됨). 언어적 복잡성 점수에 관한 추가의 상세한 설명은 하기 실시예 1에서 살펴볼 수 있다. 비구조화된 폴리펩티드의 언어적 복잡성 점수는 15 초과, 16 초과, 17 초과, 18 초과, 19 초과, 또는 20 초과일 수 있다.

[0044] 일부 실시양태에서, 비구조화된 폴리펩티드는 서열식별번호: 7-18로 이루어진 군으로부터 선택되는 아미노산 서열을 포함할 수 있다.

[0045] 3. 융합 단백질

[0046] 본원에서는 또한 상기 기술된 바와 같은, 비구조화된 폴리펩티드 중 적어도 1개를 포함할 수 있는 융합 단백질을 개시한다. 융합 단백질은 적어도 1개의 결합 폴리펩티드 및 적어도 1개의 링커를 추가로 포함할 수 있다.

[0047] 일부 실시양태에서, 융합 단백질은 1개 초과, 2개, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 또는 적어도 20개의 결합 폴리펩티드를 포함할 수 있다. 융합 단백질은 30개 미만, 25개 미만, 또는 20개 미만의 결합 폴리펩티드를 포함할 수 있다. 융합 단백질은 1 내지 30개, 예컨대 1 내지 20개, 또는 1 내지 10개의 결합 폴리펩티드를 포함할 수 있다. 상기 실시양태에서, 결합 폴리펩티드는 서로 동일하거나 또는 상이할 수 있다. 일부 실시양태에서, 융합 단백질은 서로에 대해 직렬식으로 위치하는 1개 초과, 2개, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 또는 적어도 20개의 비구조화된 폴리펩티드를 포함할 수 있다. 융합 단백질은 30개 미만, 25개 미만, 또는 20개 미만의 비구조화된 폴리펩티드를 포함할 수 있다. 융합 단백질은 1 내지 30개, 예컨대 1 내지 20개, 또는 1 내지 10개의 비구조화된 폴리펩티드를 포함할 수 있다. 상기 실시양태에서, 비구조화된 폴리펩티드는 서로 동일하거나 또는 상이할 수 있다. 일부 실시양태에서, 융합 단백질은 서로에 대해 직렬식으로 위치하는 1개 초과, 2개, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 또는 적어도 20개의 비구조화된 폴리펩티드를 포함할 수 있다.

[0048] 일부 실시양태에서, 융합 단백질은 1개 초과, 2개, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 또는 적어도 20개의 비구조화된 폴리펩티드를 포함할 수 있다. 융합 단백질은 30개 미만, 25개 미만, 또는 20개 미만의 비구조화된 폴리펩티드를 포함할 수 있다. 융합 단백질은 1 내지 30개, 예컨대 1 내지 20개, 또는 1 내지 10개의 비구조화된 폴리펩티드를 포함할 수 있다. 상기 실시양태에서, 비구조화된 폴리펩티드는 서로 동일하거나 또는 상이할 수 있다. 일부 실시양태에서, 융합 단백질은 서로에 대해 직렬식으로 위치하는 1개 초과, 2개, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 또는 적어도 20개의 비구조화된 폴리펩티드를 포함할 수 있다.

[0049] 일부 실시양태에서, 융합 단백질은 모듈식 선형 폴리펩티드로서 배열될 수 있다. 예를 들어, 모듈식 선형 폴리펩티드는 하기 구조들: [결합 폴리펩티드]_m-[링커]_k-[비구조화된 폴리펩티드]; [비구조화된 폴리펩티드]-[링커]_k-[결합 폴리펩티드]_m; [결합 폴리펩티드]_m-[링커]_k-[비구조화된 폴리펩티드]-[결합 폴리펩티드]_m-[링커]_k-[비구조화된 폴리펩티드]; 또는 [비구조화된 폴리펩티드]-[결합 폴리펩티드]_m-[링커]_k-[비구조화된 폴리펩티드]-[결합 폴리펩티드]_m 중 하나로 배열될 수 있으며, 여기서 k 및 m은 각각 독립적으로 1 이상의 정수이다. 일부 실시양태에서, m은 20 이하의 정수일 수 있다. 일부 실시양태에서, m은 정수 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20일 수 있다. 일부 실시양태에서, k는 10 이하의 정수일 수 있다. 일부 실시양태에서, k는 정수 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10일 수 있다. 일부 실시양태에서, 적어도 1개의 결합 폴리펩티드는 적어도 1개의 비구조화된 폴리펩티드에 대해 N-말단에 위치할 수 있다. 일부 실시양태에서, 적어도 1개의 결합 폴리펩티드는 적어도 1개의 비구조화된 폴리펩티드에 대해 C-말단에 위치할 수 있다.

[0050] 일부 실시양태에서, 1개 이상의 열 반응성 폴리펩티드를 포함하는 융합 단백질의 전이 온도는 실온 (약 25°C) 내지 체온 (약 37°C)일 수 있다. 열 반응성, 비구조화된 폴리펩티드는 융합 단백질이 대상체에게 투여되는 온도에서 체온 미만 또는 체온 초과인 그의 LCST 및/또는 UCST를 가질 수 있다.

[0051] 융합 단백질은 관련 기술분야에 공지된 방법에 따라 숙주 세포에서 재조합적으로 발현될 수 있다. 융합 단백질은 관련 기술분야에 공지된 임의의 적합한 수단에 의해 정제될 수 있다. 예를 들어, 융합 단백질은 크로마토그래피, 예컨대 액체 크로마토그래피, 크기 배제 크로마토그래피, 또는 친화성 크로마토그래피를 사용하여, 또는 초원심분리 기술을 통해 정제될 수 있다. 일부 실시양태에서, 융합 단백질은 크로마토그래피 없이 정제될 수

있다. 일부 실시양태에서, 융합 단백질은 역 전이 사이클링을 사용하여 정제될 수 있다.

[0052] a. 결합 폴리펩티드

[0053] 결합 폴리펩티드는 적어도 1개의 표적에 결합할 수 있는 임의의 폴리펩티드를 포함할 수 있다. 결합 폴리펩티드는 적어도 1개의 표적에 결합할 수 있다. "표적"은 결합 폴리펩티드가 결합할 수 있는 엔티티일 수 있다. 표적은 예를 들어, 또 다른 폴리펩티드, 세포 표면 수용체, 탄수화물, 항체, 소분자, 또는 그의 조합을 포함할 수 있다. 표적은 바이오마커일 수 있다. 표적은 효능작용을 통해 활성화되거나 또는 길항작용을 통해 차단될 수 있다. 결합 폴리펩티드는 특이적으로 표적에 결합할 수 있다. 결합 폴리펩티드는 표적에 결합함으로써 표적화 모이어티, 효능제, 길항제, 또는 그의 조합으로서 작용할 수 있다.

[0054] 결합 폴리펩티드는 표적에 결합하는 단량체일 수 있다. 단량체는 1개 이상의 표적에 결합할 수 있다. 결합 폴리펩티드는 올리고머를 형성할 수 있다. 결합 폴리펩티드는 동일하거나 또는 상이한 결합 폴리펩티드와 올리고머를 형성할 수 있다. 올리고머는 표적에 결합할 수 있다. 올리고머는 1개 이상의 표적에 결합할 수 있다. 올리고머 내의 1개 이상의 단량체가 1개 이상의 표적에 결합할 수 있다. 일부 실시양태에서, 융합 단백질은 다가일 수 있다. 일부 실시양태에서, 융합 단백질은 다중 표적에 결합할 수 있다. 일부 실시양태에서, 결합 폴리펩티드 단독의 활성화는 융합 단백질의 일부일 때의 결합 단백질의 활성화와 동일한 것일 수 있다.

[0055] 일부 실시양태에서, 결합 폴리펩티드는 1개 이상의 스캐폴드 단백질을 포함할 수 있다. 본원에 사용된 바와 같이, "스캐폴드 단백질"이란, 비교적 안정적이고, 정의된 3차원 구조를 갖는 1개 이상의 폴리펩티드 도메인을 지칭한다. 스캐폴드 단백질은 친화성 조작 능력을 추가로 가질 수 있다. 일부 실시양태에서, 스캐폴드 단백질은 특정 표적에 결합하도록 조작될 수 있다. 결합 폴리펩티드가 1개 초과 스캐폴드 단백질을 포함하는 실시양태에서, 스캐폴드 단백질은 동일하거나 또는 상이할 수 있다.

[0056] 일부 실시양태에서, 결합 폴리펩티드는 단백질 A를 포함할 수 있다. 단백질 A는 원래 스태필로코쿠스 아우레우스(*Staphylococcus aureus*)로부터 유래된 42 kD 단백질이다. 이는 면역글로불린 G 항체의 불변 영역 (Fc)에 높은 결합 친화도를 보일 수 있다. 단백질 A는 모두가 항체 Fc에 결합할 수 있는 5개의 연결 도메인을 포함한다. 고정화된 단백질 A는 다양한 항체의 정제를 위해 사용될 수 있다 (예컨대 문헌 [Hober et al. (2006) J. Chromatogr. B] 참조, 상기 문헌은 그 전문이 본원에 참조로 포함됨).

[0057] b. 링커

[0058] 상기 언급된 바와 같이, 융합 단백질은 적어도 1개의 링커를 추가로 포함할 수 있다. 일부 실시양태에서, 융합 단백질은 1개 초과 링커를 포함할 수 있다. 상기 실시양태에서, 링커는 서로 동일하거나 또는 상이할 수 있다. 융합 단백질은 적어도 1, 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 적어도 20, 적어도 25, 적어도 30, 적어도 35, 적어도 40, 적어도 45, 적어도 50, 적어도 55, 적어도 60, 적어도 65, 적어도 70, 적어도 75, 적어도 80, 적어도 85, 적어도 90, 적어도 95, 또는 적어도 100개의 링커를 포함할 수 있다. 융합 단백질은 500개 미만, 400개 미만, 300개 미만, 또는 200개 미만의 링커를 포함할 수 있다. 융합 단백질은 1 내지 1000개의 링커, 예컨대 10 내지 900, 10 내지 800, 또는 5 내지 500개의 링커를 포함할 수 있다.

[0059] 링커는 결합 폴리펩티드와 비구조화된 폴리펩티드 사이에, 결합 폴리펩티드 사이에, 비구조화된 폴리펩티드 사이에, 또는 그의 조합에 위치할 수 있다. 다중 링커가 서로 인접하게 위치할 수 있다. 다중 링커는 서로 인접하게 및 결합 폴리펩티드와 비구조화된 폴리펩티드 사이에 위치할 수 있다.

[0060] 링커는 임의의 적합한 아미노산 서열 및 길이의 폴리펩티드일 수 있다. 링커는 스페이서 펩티드로서 작용할 수 있다. 링커는 폴리펩티드 도메인 사이에 존재할 수 있다. 링커는 결합 도메인의 활성화는 보존하면서, 결합 폴리펩티드의 결합 도메인을 충분히 이격시켜 놓을 수 있다. 일부 실시양태에서, 링커는 하전된 아미노산을 포함할 수 있다. 일부 실시양태에서, 링커는 가소성일 수 있다. 일부 실시양태에서, 링커는 적어도 1개의 글리신 및 적어도 1개의 세린을 포함할 수 있다. 일부 실시양태에서, 링커는 (Gly₄Ser)₃ (서열식별번호: 21)으로 이루어진 아미노산 서열을 포함할 수 있다. 일부 실시양태에서, 링커는 적어도 1개의 프롤린을 포함할 수 있다. 일부 실시양태에서, 링커는 서열식별번호: 22로 이루어진 아미노산 서열을 포함할 수 있다.

[0061] c. 폴리뉴클레오티드

[0062] 본원에 상술된 융합 단백질을 코딩하는 폴리뉴클레오티드를 추가로 개시한다. 벡터는 본원에 상술된 융합 단백

질을 코딩하는 폴리뉴클레오티드를 포함할 수 있다. 폴리펩티드를 발현시키기 위해, 전사를 지시하는 프로모터, 전사/번역 종결인자, 및 단백질을 코딩하는 핵산의 경우에, 번역 개시를 위한 리보솜 결합 부위를 함유하는 발현 벡터 내로 폴리펩티드를 코딩하는 폴리뉴클레오티드를 서브클로닝할 수 있다. 벡터의 한 예로 pet24가 있다. 박테리아 프로모터는 관련 기술분야에 널리 공지되어 있다. 본원에 상술된 융합 단백질을 코딩하는 폴리뉴클레오티드를 포함하는 발현 벡터로 형질전환되거나 또는 형질감염된 숙주 세포를 추가로 개시한다. 단백질을 발현하기 위한 박테리아 발현 시스템은, 예컨대 *E. 콜라이*(*E. coli*), 바실러스 종(*Bacillus sp.*), 및 살모넬라(*Salmonella*)에서 이용가능하다 ([Paiva et al., *Gene* 1983, 22, 229-235]; [Mosbach et al., *Nature* 1983, 302, 543-545], 상기 문헌 둘 다는 그 전문이 본원에 참조로 포함됨). 상기 발현 시스템을 위한 키트가 상업적으로 이용가능하다. 포유동물 세포, 효모, 및 곤충 세포를 위한 진핵생물 발현 시스템이 관련 기술분야에 널리 공지되어 있고, 또한 상업적으로 이용가능하다. 레트로바이러스 발현 시스템이 또한 사용될 수 있다.

[0063] 4. 투여

[0064] 상기에 상술된 바와 같은 융합 단백질은 제약 분야의 당업자에게 공지된 표준 기술에 따라 제제화될 수 있다. 융합 단백질을 포함하는 상기 조성물은 특정 대상체의 연령, 성별, 체중 및 상태, 및 투여 경로와 같은 인자들을 고려하여 의학 분야의 당업자에게 공지된 투여량으로 및 기술에 의해 투여될 수 있다.

[0065] 융합 단백질은 예방적으로 또는 치료적으로 투여될 수 있다. 예방적 투여에서, 융합 단백질은 반응을 유도하는데 충분한 양으로 투여될 수 있다. 치료적 적용에서, 융합 단백질은 치료 효과를 일으키는데 충분한 양으로 그를 필요로 하는 대상체에게 투여될 수 있다. 이를 달성하는데 적절한 양이 "치료적 유효 용량"인 것으로 정의된다. 이러한 용도에 효과적인 양은 예컨대 투여되는 융합 단백질 요법의 특정 조성물, 투여 방식, 질환의 단계 및 중증도, 환자의 일반적인 건강 상태, 및 처방의의 판단에 의존할 것이다.

[0066] 융합 단백질은 문헌 [Donnelly et al. (*Ann. Rev. Immunol.* 1997, 15, 617-648)]; 펠그너(Felgner) 등 (미국 특허 번호 5,580,859, 1996년 12월 3일 발행); 펠그너 (미국 특허 번호 5,703,055, 1997년 12월 30일 발행); 및 카슨(Carson) 등 (미국 특허 번호 5,679,647, 1997년 10월 21일 발행) (상기 문헌 모두 그의 내용은 그 전문이 본원에 참조로 포함됨)에 기술되어 있는 바와 같이, 관련 기술분야에 공지된 방법에 의해 투여될 수 있다. 융합 단백질은, 예를 들어, 백신 건을 사용하여, 개체에게 투여될 수 있는 입자 또는 비드로 복합체를 형성할 수 있다. 관련 기술분야의 통상의 기술자는 생리적으로 허용되는 화합물을 포함한, 제약상 허용되는 담체의 선택이, 예를 들어, 투여 경로에 의존한다는 것을 인지할 것이다.

[0067] 융합 단백질은 다양한 경로를 통해 전달될 수 있다. 전형적인 전달 경로는 비경구 투여, 예컨대 진피내, 근육내 또는 피하 전달을 포함한다. 다른 경로는 경구 투여, 비내, 질내, 경피, 정맥내, 동맥내, 종양내, 복강내, 및 표피 경로를 포함한다. 일부 실시양태에서, 융합 단백질은 대상체에게 정맥내로, 동맥내로, 또는 복강내로 투여될 수 있다.

[0068] 융합 단백질은 액체 제제, 예컨대 현탁제, 시럽, 또는 엘릭시르 일 수 있다. 융합 단백질은 (예컨대 펠그너 등의 미국 특허 번호 5,703,055; [Gregoriadis, *Liposome Technology*, Vols. I to III (2nd ed. 1993)] (상기 문헌의 내용은 그 전문이 본원에 참조로 포함됨)에 기술된 방법에 의해) 리포솜, 마이크로스피어, 또는 다른 중합체 매트릭스 내에 포함될 수 있다. 리포솜은 인지질 또는 다른 지질로 구성될 수 있으며, 제조 및 투여가 비교적 간단한, 무독성의, 생리적으로 허용되는 대사가 가능한 담체일 수 있다.

[0069] 융합 단백질은 백신으로서 사용될 수 있다. 백신은 전기천공을 통해, 예컨대 미국 특허 번호 7,664,545 (상기 특허는 그의 내용 전문이 본원에 참조로 포함됨)에 기술된 방법에 의해 투여될 수 있다. 전기천공은 미국 특허 번호 6,302,874; 5,676,646; 6,241,701; 6,233,482; 6,216,034; 6,208,893; 6,192,270; 6,181,964; 6,150,148; 6,120,493; 6,096,020; 6,068,650; 및 5,702,359 (상기 특허들은 그의 내용 전문이 본원에 참조로 포함됨)에 기술된 방법 및/또는 장치에 의해 이루어질 수 있다. 전기천공은 최소 침습적 장치를 통해 수행될 수 있다.

[0070] 일부 실시양태에서, 융합 단백질은 제어 방출 제제로 투여될 수 있다. 일부 실시양태에서, 융합 단백질은, 융합 단백질이 투여 이전에 가용성인 상태 그대로 유지되도록 하고, 투여시에는 융합 단백질이 대상체에서 겔-유사 데포로 전이되도록 하는 전이 온도를 갖는 1개 이상의 열 반응성 폴리펩티드를 포함할 수 있다. 일부 실시양태에서, 융합 단백질은, 융합 단백질이 실온에서 가용성인 상태 그대로 유지되도록 하고, 투여시에는 융합 단백질이 대상체에서 겔-유사 데포로 전이되도록 하는 전이 온도를 갖는 1개 이상의 열 반응성 폴리펩티드를 포함할 수 있다. 예를 들어, 일부 실시양태에서, 융합 단백질은 실온 (약 25°C) 내지 체온 (약 37°C)인 전이 온도

를 갖는 1개 이상의 열 반응성 폴리펩티드를 포함할 수 있으며, 이로써 융합 단백질은 투여되어 세포를 형성할 수 있다. 본원에 사용된 바와 같이, "세포"는 시간이 경과함에 따라 융합 단백질을 방출하는, 융합 단백질을 포함하는 겔-유사 조성물을 지칭한다. 일부 실시양태에서, 융합 단백질은 피하로 또는 종양내로 주사되어 세포 (코아세르베이트)를 형성할 수 있다. 세포는 융합 단백질을 조절 방식으로 및/또는 지속적으로 방출할 수 있다. 예를 들어, 세포는 융합 단백질을 순환 또는 종양 내로 지속적으로 방출할 수 있다. 일부 실시양태에서, 융합 단백질은 적어도 1일, 적어도 2일, 적어도 3일, 적어도 4일, 적어도 5일, 적어도 6일, 적어도 7일, 적어도 1주, 적어도 1.5주, 적어도 2주, 적어도 2.5주, 적어도 3.5주, 적어도 4주, 또는 적어도 1개월의 기간에 걸쳐 세포로부터 방출될 수 있다.

[0071] **5. 검출**

[0072] 본원에 사용된 바와 같이, "검출하다" 또는 "~의 존재를 결정하다"라는 용어는 1개 이상의 융합 단백질, 표적, 또는 표적에 결합된 융합 단백질의 검출불가능, 낮은, 보통, 또는 높은 농도의 정성적 측정을 지칭한다. 검출은 시험관내, 생체외, 또는 생체내 검출을 포함할 수 있다. 검출은 1개 이상의 융합 단백질 또는 표적의 부재 대비 1개 이상의 융합 단백질 또는 표적의 존재를 검출하는 것을 포함할 수 있다. 검출은 또한 1개 이상의 융합 단백질 또는 표적의 수준의 정량화를 포함할 수 있다. "정량화하다" 또는 "정량화"라는 용어는 상호교환적으로 사용될 수 있고, 상대적이건 절대적이건, 물질 (예컨대 융합 단백질 또는 표적)의 정량 또는 존재비를 결정하는 프로세스를 지칭할 수 있다. 임의의 적합한 검출 방법이 본 개시내용의 일반 범주 내에 포함된다. 일부 실시양태에서, 융합 단백질은 검출을 위해 그에 부착된 리포터를 포함할 수 있다. 일부 실시양태에서, 융합 단백질은 리포터로 표지될 수 있다. 일부 실시양태에서, 표적에 결합된 융합 단백질의 검출은 웨스턴 블롯 상의 밴드 강도, 유세포 결정법, 방사선 표지 영상화, 세포 결합 결정법, 활성 결정법, 표면 플라즈몬 공명 (SPR), 면역결정법을 포함하나 이에 제한되지는 않는 방법에 의해, 또는 관련 기술분야에 공지된 다양한 다른 방법에 의해 결정될 수 있다.

[0073] 융합 단백질이 표적 결합 및/또는 검출을 위한 항체 모방체인 것을 포함하는 일부 실시양태에서, 임의의 면역결정법이 활용될 수 있다. 면역결정법은 예를 들어, 효소-결합 면역결정법 (ELISA), 방사선면역결정법 (RIA), 경쟁적 억제 결정법, 예컨대 전방 또는 후방 경쟁적 억제 결정법, 형광 분극 결정법, 또는 경쟁적 결합 결정법일 수 있다. ELISA는 샌드위치 ELISA일 수 있다. 표적에의 융합 단백질의 특이적인 면역학적 결합은 융합 단백질에 부착된 직접적 표지를 통해, 또는 간접적 표지, 예컨대 알칼리성 포스파타제 또는 홀스래디쉬 퍼옥시다제를 통해 검출될 수 있다. 고정화된 융합 단백질 사용이 면역결정법 내로 포함될 수 있다. 융합 단백질은 다양한 지지체, 예컨대 자기 또는 크로마토그래피 매트릭스 입자, 검정 플레이트 (예컨대 마이크로타이터 웰)의 표면, 고체 기재 물질의 조각 등 그 위에 고정화될 수 있다. 검정 스트립은 고체 지지체 상에 어레이로 융합 단백질 또는 복수개의 융합 단백질을 코팅함으로써 제조될 수 있다. 이어서, 상기 스트립을 시험 생물학적 샘플 내로 dipping된 후, 세척 및 검출 단계를 통해 신속히 프로세싱하여 측정가능한 신호, 예컨대 유색 스팟을 생성할 수 있다.

[0074] **6. 방법**

[0075] **a. 질환을 치료하는 방법**

[0076] 또 다른 측면에서, 치료를 필요로 하는 대상체에서 질환을 치료하는 방법을 개시한다. 본 방법은 대상체에게 본원에 기술된 바와 같은 융합 단백질의 유효량을 투여하는 단계를 포함할 수 있다. 질환은 암, 대사 질환, 자가면역 질환, 심혈관 질환, 및 정형외과적 장애로부터 선택될 수 있다. 일부 실시양태에서, 질환은 적어도 1개의 결합 폴리펩티드의 표적과 연관된 질환일 수 있다.

[0077] 대사 질환은 체내의 비정상적인 화학 반응이 정상적인 대사 과정을 변경시키는 경우에 발생할 수 있다. 대사 질환은, 예를 들어, 인슐린 저항, 비-알콜성 지방간 질환, 2형 당뇨병, 인슐린 저항 질환, 심혈관 질환, 죽상경화, 지질-관련 대사 장애, 고혈당증, 고인슐린혈증, 고지혈증, 및 글루코스 대사 장애를 포함할 수 있다.

[0078] 자가면역 질환은 체내에 정상적으로 존재하는 물질 및 조직에 대한 신체의 비정상적인 면역 반응으로부터 발생할 수 있다. 자가면역 질환은 루푸스, 류마티스성 관절염, 다발성 경화증, 인슐린 의존성 당뇨병, 중증 근무력증, 그레이브스병, 자가면역 용혈성 빈혈, 자가면역 혈소판감소성 자반증, 굿패스처 증후군, 심상성 천포창, 급성 류마티스열, 연쇄구균 감염 후 사구체신염, 결절성 다발동맥염, 심근염, 건선, 셀리악병, 크론병, 췌양성 장염, 및 섬유근통을 포함할 수 있으나 이에 제한되지는 않는다.

[0079] 심혈관 질환은 심장 또는 혈관을 수반하는 질환 부류이다. 심혈관 질환은, 예를 들어, 관상 동맥 질환 (CAD),

예컨대 협심증 및 심근경색 (심장 마비), 뇌졸중, 고혈압성 심장 질환, 류마티스성 심장 질환, 심근병증, 심장 부정맥, 선천성 심장 질환, 판막성 심장 질환, 심장염, 대동맥류, 말초 동맥 질환, 및 정맥 혈전증을 포함할 수 있다.

[0080] 정형외과적 장애 또는 근골격 장애는 신체 관절, 인대, 근육, 신경, 힘줄, 및 사지, 목, 및 등을 지지하는 구조에서의 손상 또는 통증이다. 정형외과적 장애는 통증을 유발하고, 정상 활동을 손상시키는 변성 질환 및 염증성 병태를 포함할 수 있다. 정형외과적 장애는, 예를 들어, 손목 터널 증후군, 상과염, 및 건염을 포함할 수 있다.

[0081] 암은 유방암, 결장직장암, 결장암, 폐암, 전립선암, 고환암, 뇌암, 피부암, 직장암, 위암, 식도암, 육종, 기관암, 두경부암, 췌장암, 간암, 난소암, 림프구성 암, 자궁경부암, 외음부암, 흑색종, 증피종, 신장암, 방광암, 갑상선암, 골암, 암종, 육종, 및 연부조직암을 포함할 수 있으나 이에 제한되지는 않는다. 일부 실시양태에서, 암은 결장직장암일 수 있다. 일부 실시양태에서, 암은 결장직장 선암종일 수 있다.

[0082] 일부 실시양태에서, 본 개시내용은 관심 종양성 표적에 대한 항체 모방체 개발에서 스캐폴드 단백질을 사용하는 방법을 제공한다. 스캐폴드 단백질 조각이 출현함에 따라, 입체 및 구조적 제한에 의해 제약을 받지 않는 강력한 효능의 단백질 약물을 디자인할 수 있는 가능성이 도래한다. 비록 강력한 효능의 단백질 약물이 진단 또는 치료를 위해 유용할 수는 있지만, 표적 부위로 성공적으로 전달하는 것이 큰 도전과제가 될 수 있다.

[0083] b. 질환을 진단하는 방법

[0084] 또 다른 측면에서, 질환을 진단하는 방법을 개시한다. 본 방법은 대상체에게 본원에 기술된 바와 같은 융합 단백질을 투여하는 단계, 및 융합 단백질의 표적에의 결합을 검출하여 샘플 내 표적의 존재를 결정하는 단계를 포함할 수 있다. 표적의 존재는 대상체에서의 질환을 나타낼 수 있다. 일부 실시양태에서, 본 방법은 대상체로부터의 샘플을 본원에 기술된 바와 같은 융합 단백질과 접촉시키는 단계; 샘플 내 표적의 수준을 결정하는 단계; 및 샘플 내 표적의 수준을 표적의 대조군 수준과 비교하며, 여기서 대조군 수준과 상이한 표적의 수준은 대상체에서의 질환을 나타내는 것인 단계를 포함할 수 있다. 일부 실시양태에서, 질환은 상기 상술된 바와 같이, 암, 대사 질환, 자가면역 질환, 심혈관 질환, 및 정형외과적 장애로부터 선택될 수 있다. 일부 실시양태에서, 표적은 질환 마커 또는 바이오마커를 포함할 수 있다. 일부 실시양태에서, 융합 단백질은 표적에의 결합 및/또는 검출을 위한 항체 모방체로서 작용할 수 있다.

[0085] c. 표적의 존재를 결정하는 방법

[0086] 또 다른 측면에서, 샘플 내 표적의 존재를 결정하는 방법을 개시한다. 본 방법은 샘플을 본원에 기술된 바와 같은 융합 단백질과, 복합체가 융합 단백질과 샘플 내 표적 사이에 형성되도록 하는 조건 하에 접촉시키는 단계, 및 복합체의 존재를 검출하는 단계를 포함할 수 있다. 복합체의 존재는 샘플 내 표적을 나타내는 것일 수 있다. 일부 실시양태에서, 융합 단백질은 검출을 위해 리포터로 표지될 수 있다.

[0087] 일부 실시양태에서, 샘플은 대상체로부터 수득될 수 있고, 본 방법은 진단하는 것, 예측하는 것, 또는 대상체의 치료의 효능을 사정하는 것을 추가로 포함할 수 있다. 본 방법이 대상체의 치료의 효능을 사정하는 것을 포함하는 경우에는, 본 방법은 효능을 개선시키기 위해 필요에 따라 대상체의 치료를 변형시키는 것을 추가로 포함할 수 있다.

[0088] d. 치료의 유효성을 결정하는 방법

[0089] 또 다른 측면에서, 치료를 필요로 하는 대상체에서 질환에 대한 치료의 유효성을 결정하는 방법을 개시한다. 본 방법은 대상체로부터의 샘플을 본원에 상술된 바와 같은 융합 단백질과, 복합체가 융합 단백질과 샘플 내 표적 사이에 형성되도록 하는 조건 하에 접촉시키는 단계, 샘플 내 복합체의 수준을 결정하며, 여기서 복합체의 수준은 샘플 내 표적의 수준을 나타내는 것인 단계, 및 샘플 내 표적의 수준을 표적의 대조군 수준과 비교하며, 여기서 표적의 수준이 대조군 수준과 상이한 경우에는, 치료가 질환을 치료하는데 있어서 효과적이거나 또는 비효과적인 것으로 결정되는 것인 단계를 포함할 수 있다.

[0090] 시점은 질환의 발병 이전, 요법 투여 이전, 요법 투여 동안의 다양한 시점, 요법 종료 이후, 또는 그의 조합을 포함할 수 있다. 대상체에게의 융합 단백질 투여시, 융합 단백질은 표적에 결합할 수 있으며, 여기서 표적의 존재는 다양한 시점에 대상체에서의 질환의 존재를 나타낸다. 일부 실시양태에서, 표적은 질환 마커 또는 바이오마커를 포함할 수 있다. 일부 실시양태에서, 융합 단백질은 표적에의 결합 및/또는 검출을 위한 항체 모방체로서 작용할 수 있다. 다양한 시점에 표적에의 융합 단백질의 결합 비교는 질환이 진행하였는지 여부, 환자가

후기 단계의 질환을 앓고 있는지 여부, 요법이 질환을 치료하거나 또는 예방하는데 작용하고 있는지 여부, 또는 그의 조합을 나타낼 수 있다.

[0091] 일부 실시양태에서, 대조군 수준은 대상체가 치료를 시작한 기간 이전의 또는 그 동안의 시점에서의 대상체 내 수준에 상응할 수 있고, 샘플은 이후 시점에 대상체로부터 채취된다. 일부 실시양태에서, 샘플은 대상체가 치료를 받고 있는 기간 동안의 시점에 대상체로부터 채취될 수 있고, 대조군 수준은 무질환 수준 또는 대상체가 치료를 시작한 기간 이전의 시점에서의 수준에 상응한다. 일부 실시양태에서, 본 방법은, 치료가 질환을 치료하는데 있어서 비효과적인 것으로 결정되는 경우에 치료를 변형시키거나 또는 대상체에게 상이한 치료를 투여하는 것을 추가로 포함할 수 있다.

[0092] 7. 실시예

[0093] 실시예 1: 비-반복적 비구조화된 폴리펩티드의 확인

[0094] 비-반복적 순열화된 버전의 ELP를 제조하고, 조사하였다. 컴퓨터 알고리즘을 사용하여, 각각이 ELP 카운터파트와 아미노산 조성은 동일하지만, 순열화된 또는 재배열된 아미노산 배치를 갖는, 200 또는 400개 아미노산을 포함하는 서열을 확인하였다. 생성된 서열은, 그가 (a) 3개의 연속적 동일한 아미노산을 함유하거나, (b) 서열 내에 1회 초과로 발생하거나, 또는 (c) 0개 이상의, 글리신 (G)이 아닌 아미노산에 의해 이격되어 있는 적어도 2개의 프롤린 (P)을 함유하는, 5 내지 10개 아미노산을 포함하는 적어도 1개의 하위서열을 함유한다면, 거부되었는 바, 이에 충분히 비-반복적이었다. 이러한 제약은, 이들이 함께 구조-과단 잔기인 것으로서 확인되어 있는 프롤린 (P) 및 글리신 (G)이 잘 분포될 수 있도록 촉진시켰다.

[0095] 알고리즘을 사용하여, 각각이 240개 아미노산을 포함하며, 여기서 1/6은 프롤린 (P)이고, 1/3은 글리신 (G)인, 8개의 비-반복적 서열의 패널을 추가로 확인하였다. 나머지 아미노산은 엘라스틴에서 발견되는 9개의 상이한 아미노산의 다양한 조합이었다. 이들 서열의 아미노산 조성은 도 2에 도시되어 있다.

[0096] 언어적 복잡성 점수를 계산하여 각각의 비구조화된 폴리펩티드 서열의 반복성을 정량화하였다. 언어적 복잡성은 주어진 서열 중의 독특한 하위서열의 총 개수를 동일한 알파벳 및 윈도우에 대하여 가능한 독특한 하위서열의 총 개수로 나눈 값으로 정의된다. 단백질 분석 도구, CIDER를 사용하여 점수를 계산하였다 (문헌 [Holehouse et al. (2015) Biophys J.] 참조, 상기 문헌은 그 전문이 본원에 참조로 포함됨). 윈도우 길이는 각 서열에 대한 전체 서열 길이와 같도록 설정해 놓았다. 최종 점수는 서열 길이를 설명하는 언어적 복잡성 점수와 전체 서열 길이의 곱으로서 제시된다. 서열식별번호: 3-6에서 반복적 폴리펩티드 서열, 및 서열식별번호: 7-18에서 비-반복적 폴리펩티드 서열에 대하여 생성된 최종 점수는 하기 표 1에 기록되어 있다. 비-반복적 폴리펩티드 서열에 대한 최종 점수는 모두 15.0을 초과하였다.

[0097] 표 1: 반복적 (서열식별번호: 3-6) 및 비-반복적 (서열식별번호: 7-18) 폴리펩티드 서열의 반복성의 특징화

SEQ ID NO	길이와 언어적 복잡성 점수의 곱
3	7.9
4	7.9
5	8.0
6	8.0
7	28.7
8	35.6
9	32.8
10	42.8
11	50.6
12	51.6
13	50.6
14	83.3
15	85.3
16	88.3
17	86.3
18	147.8

[0098]

[0099] 실시예 2: 비-반복적 비구조화된 폴리펩티드의 발현

[0100] 명시되지 않는다면, 모든 뉴클레오티드 서열은 코돈 스캐폴딩을 사용하여 아미노산 서열로부터 역-번역하였다 (문헌 [Tang, N. C. et al. (2016) Nature Mater.] 참조, 상기 문헌은 그 전문이 본원에 참조로 포함됨). 명

시되지 않는다면, Met-Ser-Lys-Gly-Pro-Gly를 코딩하는 N-말단 리더 서열 (서열식별번호: 19) 및 Gly-Trp-Pro를 코딩하는 C-말단 His-태그 테일 (서열식별번호: 20)을 유전자 내로 포함시켰다. 모든 유전자를 상업적 합성법에 의해 합성하고, 변형된 pET-24a(+) 플라스미드 (진9 인크.(Gen9 Inc.: 미국 매사추세츠주) 및 진스크립트 인크.(Genscript Inc.: 미국 뉴저지주))로 클로닝하였다. 생성된 플라스미드를 BL21 적격 E. 콜라이 세포 내로 형질전환시켰다.

[0101] 콜로니를 2-5 mL의 테리픽 브로쓰(Terrific Broth: TB) 플러스 50 μ g/mL 카나마이신 중에 접종하고, 37°C 및 250 r.p.m.에서 밤새도록 성장시켰다. 1 밀리리터의 스타터 배양물을 1 L의 TB 플러스 50 μ g/mL 카나마이신 중에 접종하고, 37°C 및 250 r.p.m.에서 6-7 h 동안 성장시켰다. 최종 농도 1 mM의 IPTG를 첨가하여 발현을 유도하고, 세포를 추가로 24 h 동안 성장시켰다. 앞서 기술된 바와 같은 역 전이 사이클링 (ITC)에 의해 비구조화된 폴리펩티드를 정제하였다 (문헌 [Christensen et al. (2009) Curr Protoc Protein Sci.] 참조, 상기 문헌은 그 전문이 본원에 참조로 포함됨).

[0102] 실시예 3: 비-반복적 비구조화된 폴리펩티드의 특징화

[0103] 비구조화된 폴리펩티드의 역 전이 온도의 특징화를 위해, 캐리(Cary) 300 자외선-가시광선 분광광도계 (애질런트 테크놀로지스(Agilent Technologies))를 이용하여 용액 온도의 함수로서 25 μ M 펩티드 용액의 350 nm 광학 밀도를 모니터링하였다. 형광이 환경의 소수성이 증가함에 따라 증가하는 것인 스피로 오렌지(SYPRO Orange) 염료를 이용하여 역 전이 온도를 또한 모니터링하였다. 온도 반응성 중합체는 수용액 중에서의 소수성 효과에 기인하여 상 전이를 나타낸다. 그러므로, 스텝원플러스((StepOnePlus)TM 실시간 PCR 장치를 사용하여 용액 온도의 함수로서 형광을 모니터링하였다. 역 전이 온도는 혼탁도 또는 형광 구배가 최대일 때의 온도로서 정의되었다.

[0104] PBS 중 25 μ M의 모든 구축물에 대한, 약 0°C 내지 약 100°C의 가열 및 냉각 곡선으로 이루어진 혼탁도 프로파일은 비-반복적 버전의 A[0.2] (서열식별번호: 7 및 9) 및 A[0.5] (서열식별번호: 8 및 10)가 그의 반복적 카운터파트 (서열식별번호: 3 및 5, 및 서열식별번호: 4 및 6)와 유사한 전이 온도를 갖는 LCST 상 거동을 유지한다는 것을 제시한다. 추가로, 비구조화된 폴리펩티드는 0°C 내지 100°C에서 하한 임계 용액 온도 (LCST) 거동을 보였다 (도 3a-b 및 도 4a-b 참조).

서열

[illegible]

[0105]

SEQ ID NO: 10

A[0.5] 비반복, 400개 아미노산:

PAGAVPGGVPVAGGVVPGAGVVPAGAPVGGAPGVVGVPGGGVPGVVPVGVGVGVVPVGGGVGV
VVPAGGGAPGGVGVVVPGVVPGAPAGGVPGVAPGGVGGVPGVVPGGGVPGVAPGVPGVPG
GVVPGVVPGAAPGGGVPAVGVPGVGGGVVPVVGAGAPGGVPGVAVPGAPGGVVPGGVPVGGGV
GGVVPGGVPGAPGVGVVPGVGVPGVPAVGVPGAPGAPGGGVPGGGAPVGGAGGGVPGGGVP
VPGAPGVGPVGVPGVPGVPGVPGVGGAGVPGGVAPGVAVPGGVGPVVGVPVGGVPGAVPGVP
GVGGAPGVPGGGVVGVGVPGGVVPGVPGAGAPGGGAPGAPGVPPGGAPGVVPGVGGVPGAG
GGVVGGGVVVPVGGGVVPGAPGGAPVGGVVGGVP

비-반복적 폴리펩티드:

각각의 아미노산 서열은 1/6 프롤린 (P) 잔기, 1/3 글리신 (G) 잔기, 및 1/2 X 잔기를 포함하며, 여기서 X는 발린 (V), 알라닌 (A), 류신 (L), 리신 (K), 트레오닌 (T), 이소류신 (I), 티로신 (Y), 세린 (S), 및 페닐알라닌 (F)으로 이루어진 군으로부터 선택되는 하나 이상의 아미노산이다. X에 대해 선택된 각각의 아미노산은 서로 동일한 빈도로 발생한다.

SEQ ID NO: 11

PVGVPGVGGAPVVVGGAPVGVGGAGGVVPGAPVGA VAGVGAPGGA VPAGGAPGVGAPGVVP
GAVPGVGVGGVGGGVVVPVVVGVGGVPAGAVPVGGVAPVGVVPV GAGGGAPAGGVPVAVGV
PVVGGGVPGVGVGVAPGAPGAGVGVVAGAGAAAVAPVGVAPGVPGVPGAVVGVPGGVAPA
VVVGAVPGVVGVVPVGAPVVAGGVAVAPVGGAVPGVPGGVPGGGAPGAPVVGVVPVGVVP

SEQ ID NO: 12

PVGSPPGSPGVVPVSGVVVGGGSPVGGSPVVGVPSPGVSPVVGGGVSPGVGVPGVVVPGV
VGGGVGVPGGGVVGGVPVVVGVSVPSPGVSPGGGVVVPVGSVSSGPVVSVVGVSVVPGSPV
GSPSSGVPGVVPSPGGVVGSSVSPSGGGVSVVGVGGSGVPGVPGSGSVVPVGVSPVGGGSGS
GVGSGSPGGSPGGSPGVVPVGVSPGVGGVPGVVGGSGSPGVPGGGVGVGVPG

SEQ ID NO: 13

PVGTPGTVPGGVVPVGVTPVGGGTGTPTTVTVVGGVVVPGPVGGGTGVPVGTPTVVGGVGVGT
GTPGGTVPGGTGPGVVPGTGPGGVVPGTTVPTGTGTPPVGVVPTVGGVTPGVGGVPPVVGTGVP
VPVGGGVGVVGVVPGTVPGVPGVVGGTVPVGTPTTGVPGVVPTGVTGGVGPVGGGVVGG
VVPGVTVPGTVTTGVPVTVGVVVPVGTGGGVVVGGGTTPGGVPGTPGVGG

SEQ ID NO: 14

GAVVVPGGAVGVPAGAVGGVGGVLVPAVGAVPGGGVPAGVGPVAGVVPGGPLPGGAGVGGALG
AAAPGGVPVGAALPGVAGVAPGGGLPGLGAGAGLGLAGALVLPGLGGVPGVPGGGLLPGGVP
LLGLPGVPAGLPGVLPVGLLLPGAPGVPALGVPLGVPGVAPGLPGAGGLPGVGAPGLLPVGGAP
AALGGLGLPGGGALGLLAPGGALVPGVGGVAPVAAGALLPGGAPLGPALLAL

SEQ ID NO: 15

SSGSSGSGFFSGGVPSGGGVVFVPSFFGFFPSGSGGVVPGVPGGGVGVPGFVGVVPGGFPVGVVPG
VFVPSVVGVPGFFVFPGVPGGSSSPGSPFGVFGGGSSGGGFPGGSSPGVVSPGFGFGVPGFVPG
GSVFFFPFGGGFPGGFPSGSFVPGGGFVVPFVSPGSPGFVPSFGSPGGFGGFSPFSSGVGSPGGV
PGVGVVSPGFPGGSPFGFPGSPGGSFSGSPSVFGSGSPGSFP

SEQ ID NO: 16

AAGGTGFFTGGAPGGTPTAFGGFTPTTFAGGGAPGFFAFPGAPGTFTFGAPGAAPFGTTPGTTFG
 FPGGTGTPGAATPGFFGAFGFPGAFAATFGFPGGGAGAAPAGTPGAPTAAGATTPGAGGTTGGA
 PTGGGFAPTFFGFFPTGTGGAAPGAPGGGTTTFPGATPFGFPGTGTPTGTFGAPGGGFFPGAGFGG
 GAPGAPAAAGGGTPTGATPGTFTPGATFPGTFPFGGGFPGGFPT

SEQ ID NO: 17

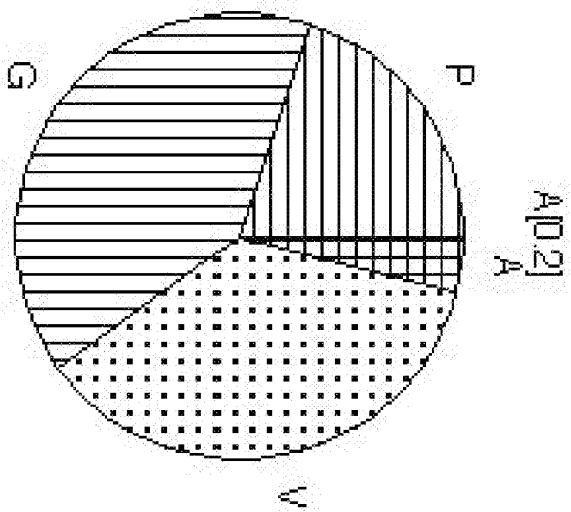
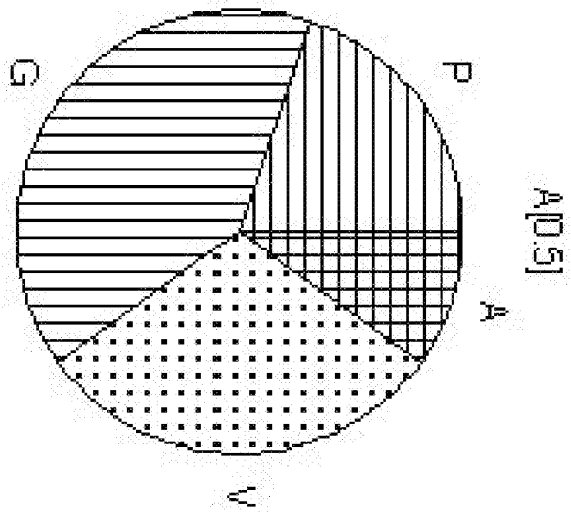
TILPGIGLPGGLPLTIGGILPILLGTTIGGIPGGGIIPGIGGGIPILTLLILPTGLPGTPLTTGTPGITGLG
TPLGIGLIGTGIPGGGTGIPTGIPGTTLPGIPTLGGGIPLIGTPGGIGTGIPGTGTIIGIPGGGLLLIPGIT

PLIGLPGGTPGTPGGGTGTTPTGLPGILLPGGLPGIPTGTPGTGTPIGGLLILLTLPLPLPGLGLPLIP LGTGLTGGTPTGGIPGGIGIPGILTT
SEQ ID NO: 18 PGTYPGYGYVYPTTGGIPGGVVPGGGTKLVPGKKGKGGAKAPGTVPVGAGGGKIVPIYGIAPGKY GYPGGGIVPGITTPGLPTGKKPYGGVPVLYGKLPAGPIPTAGAPGYIAPGVPGGLVKGGTGIAPL GVILVYIGVGGIKGGALPTGGLYPGAGTPGYPVGGGAPAGGIALKPGITPGTAAPGLPGKGGKYT YPGGAPGGTGGVPGYPLALKLGIPTKGGGIGLPYIGLLPGKPGG
SEQ ID NO: 19 Met-Ser-Lys-Gly-Pro-Gly
SEQ ID NO: 20 Gly-Trp-Pro
SEQ ID NO: 21 (Gly ₄ Ser) ₃
SEQ ID NO: 22 PQPQPKPQPKPEPEPQPQG
SEQ ID NO: 23 (VPGXG) _n , 여기서 X는 프롤린을 제외한 임의의 아미노산이고, n은 1 이상의 정수이다.

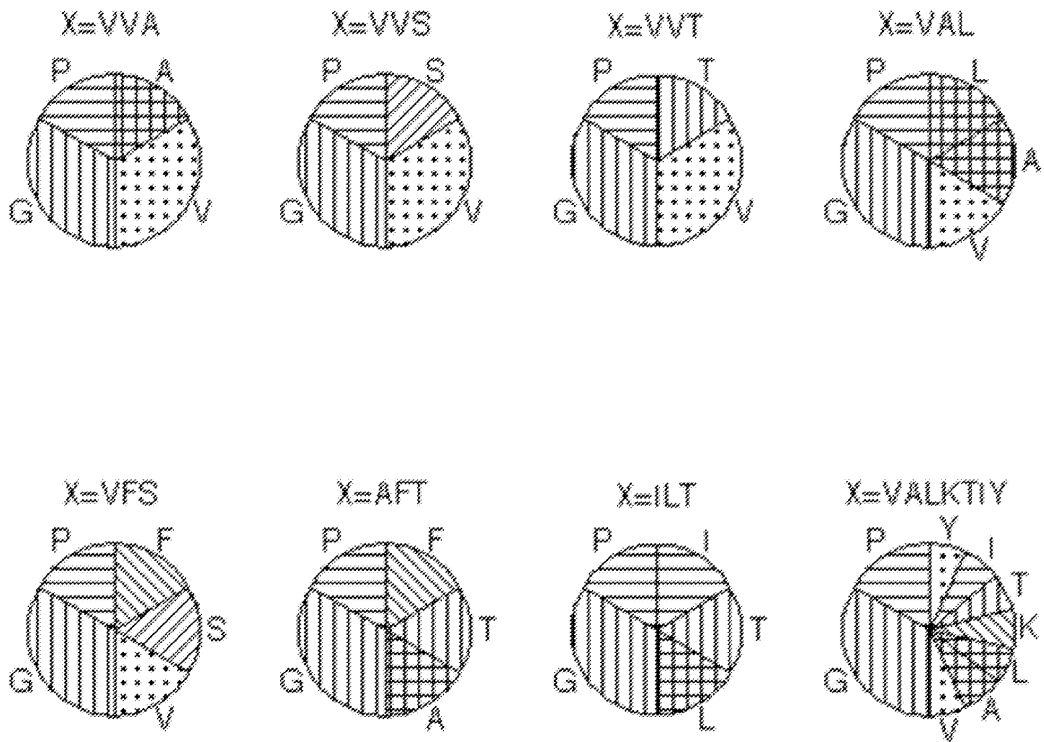
[0107]

도면

도면1

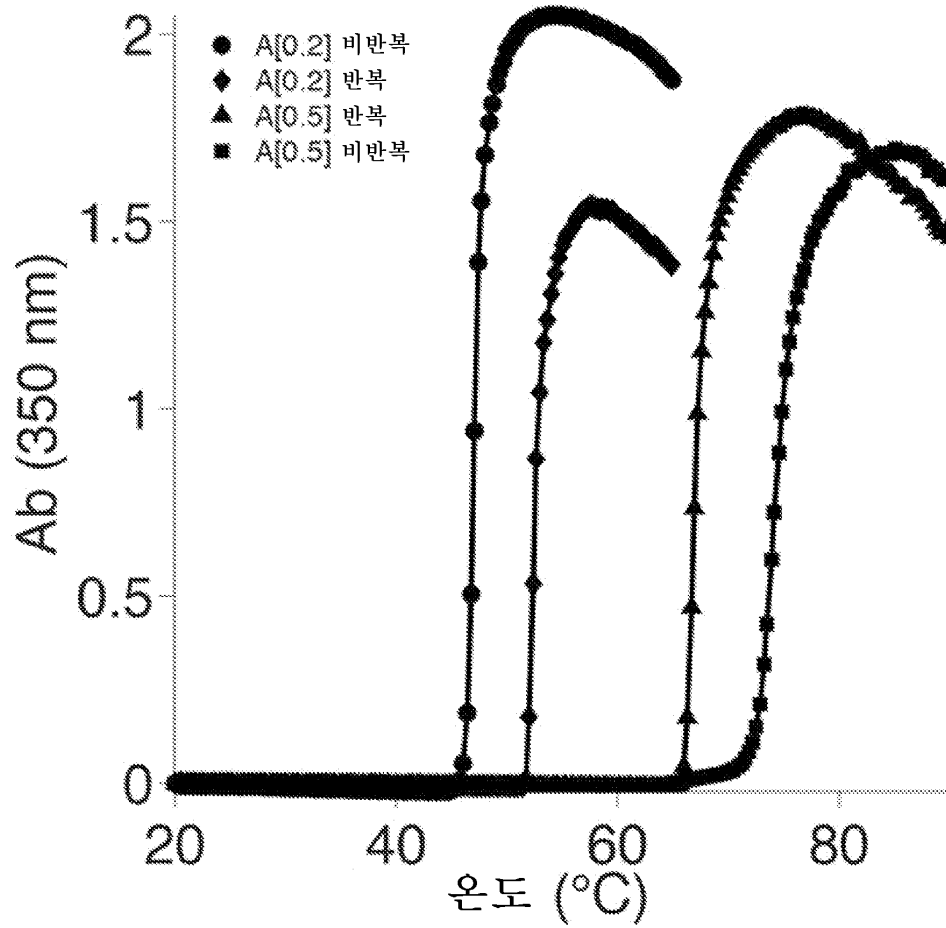


도면2



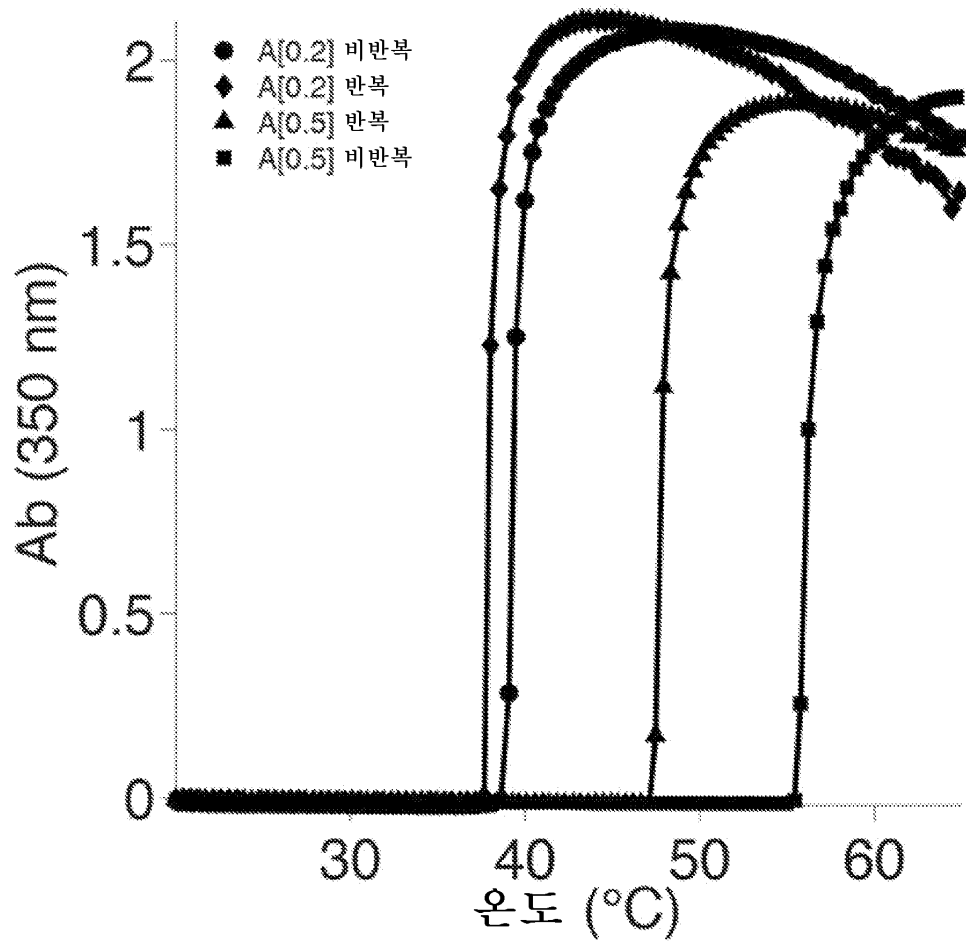
도면3a

200개 아미노산

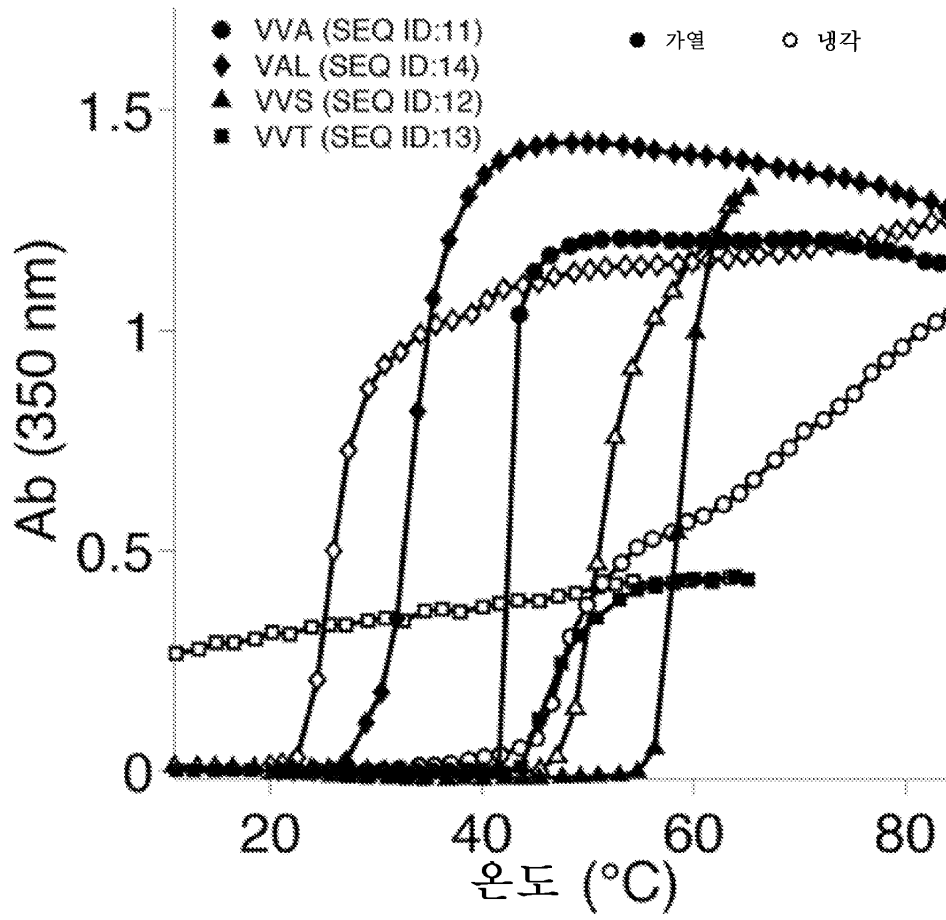


도면3b

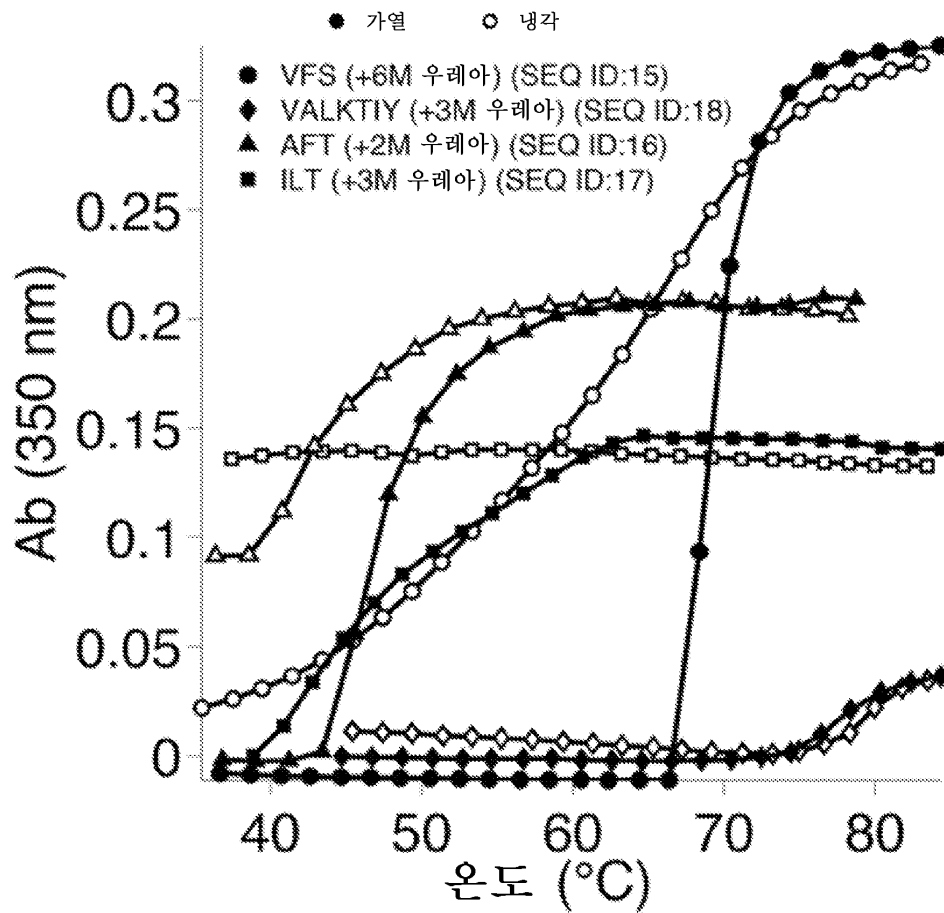
400개 아미노산



도면4a



도면4b



서열 목록

SEQUENCE LISTING

<110> DUKE UNIVERSITY

<120> UNSTRUCTURED NON-REPETITIVE POLYPEPTIDES HAVING LCST BEHAVIOR

<130> 028193-9228-W001

<150> 62/399,123

<151> 2016-09-23

<160> 23

<170> PatentIn version 3.5

<210> 1

<211> 10

<212> PRT

<213> Homo sapien

<400> 1

Gly Ala Gly Val Pro Gly Val Gly Val Pro

1 5 10

<210> 2

<211> 25

<212> PRT

<213> Homo sapien

<400> 2

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly

1 5 10 15

Val Gly Val Pro Gly Val Gly Val Pro

20 25

<210> 3

<211> 200

<212> PRT

<213> Homo sapien

<400> 3

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly

1 5 10 15

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val

20 25 30

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly

35 40 45

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val

50 55 60

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro

65 70 75 80

Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly

85 90 95

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala

100 105 110

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly

115 120 125

Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val

130 135 140
 Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
 145 150 155 160
 Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly

 165 170 175
 Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val
 180 185 190
 Gly Val Pro Gly Val Gly Val Pro
 195 200
 <210> 4
 <211> 200
 <212> PRT
 <213> Homo sapien
 <400> 4
 Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly
 1 5 10 15
 Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala

 20 25 30
 Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
 35 40 45
 Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val
 50 55 60
 Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
 65 70 75 80
 Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly

 85 90 95
 Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala
 100 105 110
 Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
 115 120 125
 Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val
 130 135 140

Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro

145 150 155 160

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly

165 170 175

Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala

180 185 190

Gly Val Pro Gly Val Gly Val Pro

195 200

<210> 5

<211> 400

<212> PRT

<213> Homo sapien

<400> 5

Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly

1 5 10 15

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val

20 25 30

Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly

35 40 45

Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val

50 55 60

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro

65 70 75 80

Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly

85 90 95

Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala

100 105 110

Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly

115 120 125

Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val

130 135 140

Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro
145 150 155 160
Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
165 170 175
Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val
180 185 190
Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly
195 200 205
Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
210 215 220
Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro
225 230 235 240
Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
245 250 255
Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val
260 265 270
Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly
275 280 285
Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
290 295 300
Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
305 310 315 320
Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly
325 330 335
Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val
340 345 350
Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
355 360 365
Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val
370 375 380
Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Val Gly Val Pro

385					390					395					400
<210>	6														
<211>	400														
<212>	PRT														
<213>	Homo sapien														
<400>	6														
Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly
1				5					10					15	
Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala
				20					25					30	
Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly
				35					40					45	
Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val
				50					55					60	
Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro
65					70					75					80
Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly
				85					90					95	
Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala
				100					105					110	
Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly
				115					120					125	
Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val
				130					135					140	
Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro
145					150					155					160
Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly
				165					170					175	
Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala
				180					185					190	
Gly	Val	Pro	Gly	Val	Gly	Val	Pro	Gly	Ala	Gly	Val	Pro	Gly	Val	Gly
				195					200					205	

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val
210 215 220

Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
225 230 235 240

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly
245 250 255

Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala
260 265 270

Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
275 280 285

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val
290 295 300

Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
305 310 315 320

Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly
325 330 335

Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala
340 345 350

Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly
355 360 365

Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro Gly Ala Gly Val
370 375 380

Pro Gly Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val Pro
385 390 395 400

<210> 7

<211> 200

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 7

Pro Gly Gly Gly Val Pro Gly Gly Val Pro Gly Ala Val Pro Gly Val
1 5 10 15

Pro Gly Ala Val Val Pro Val Val Gly Gly Val Pro Gly Val Val Gly

20 25 30

Val Pro Gly Gly Val Val Pro Val Gly Val Val Val Pro Gly Val Pro

35 40 45

Val Gly Val Pro Gly Val Pro Val Val Gly Gly Gly Val Pro Ala Gly

50 55 60

Gly Val Val Pro Gly Gly Gly Val Pro Val Gly Gly Gly Val Val Pro

65 70 75 80

Gly Val Gly Val Val Pro Gly Gly Val Pro Val Val Val Gly Gly Gly

85 90 95

Val Gly Val Pro Gly Val Val Pro Gly Val Pro Gly Val Gly Val Pro

100 105 110

Val Gly Gly Val Pro Gly Gly Gly Val Val Val Pro Gly Gly Val Gly

115 120 125

Val Pro Val Val Gly Val Val Pro Gly Ala Pro Gly Val Pro Gly Val

130 135 140

Pro Gly Gly Val Gly Gly Val Pro Val Gly Val Gly Val Pro Gly Gly

145 150 155 160

Gly Ala Pro Gly Val Val Pro Gly Gly Ala Val Pro Gly Gly Ala Pro

165 170 175

Val Gly Gly Val Gly Gly Gly Val Pro Gly Val Gly Gly Val Pro Gly

180 185 190

Ala Pro Gly Gly Val Pro Gly Gly

195 200

<210> 8

<211> 200

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 8

Pro Gly Gly Val Pro Gly Val Pro Ala Gly Gly Gly Ala Gly Val Pro

1 5 10 15
 Gly Val Pro Gly Ala Pro Ala Gly Gly Gly Ala Pro Gly Val Ala Pro
 20 25 30
 Gly Gly Val Val Pro Gly Val Pro Gly Val Ala Pro Val Val Gly Gly
 35 40 45
 Gly Ala Pro Gly Gly Val Pro Gly Ala Val Pro Gly Val Pro Gly Gly
 50 55 60
 Val Pro Gly Gly Val Val Val Pro Gly Gly Gly Val Gly Ala Pro Gly

65 70 75 80
 Gly Val Ala Pro Val Gly Gly Gly Val Pro Val Val Gly Gly Val Pro
 85 90 95
 Val Gly Gly Val Ala Pro Gly Val Pro Val Gly Val Pro Gly Val Gly
 100 105 110
 Val Gly Val Gly Val Pro Val Gly Gly Gly Val Gly Gly Val Pro Gly
 115 120 125
 Val Val Pro Ala Gly Val Pro Gly Gly Gly Val Ala Gly Val Val Pro

130 135 140
 Val Gly Gly Val Pro Val Val Val Gly Ala Pro Gly Gly Gly Val Gly
 145 150 155 160
 Val Pro Gly Ala Pro Gly Ala Gly Gly Val Pro Gly Gly Ala Pro Gly
 165 170 175
 Ala Pro Gly Val Val Ala Pro Gly Gly Val Pro Val Gly Val Val Val
 180 185 190
 Pro Gly Val Val Pro Gly Gly Gly
 195 200

<210> 9

<211> 400

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 9

Pro Gly Gly Val Pro Gly Val Gly Gly Gly Val Pro Gly Val Pro Gly

1 5 10 15
 Ala Pro Gly Val Val Pro Gly Val Val Val Pro Gly Gly Val Val Pro
 20 25 30
 Val Gly Gly Val Pro Gly Val Pro Val Val Gly Gly Gly Val Pro Val
 35 40 45
 Gly Gly Gly Val Gly Gly Val Pro Gly Ala Pro Gly Gly Gly Val Pro

 50 55 60
 Gly Ala Pro Val Gly Gly Val Pro Val Gly Gly Val Val Pro Gly Val
 65 70 75 80
 Gly Val Pro Gly Ala Gly Gly Gly Val Pro Gly Gly Val Gly Val Val
 85 90 95
 Pro Val Gly Gly Gly Ala Pro Gly Gly Val Pro Gly Val Val Pro Gly
 100 105 110
 Val Pro Val Gly Gly Val Pro Ala Gly Val Val Pro Gly Gly Val Gly

 115 120 125
 Val Gly Gly Gly Val Val Pro Gly Val Val Gly Val Pro Gly Gly Gly
 130 135 140
 Val Gly Gly Gly Val Gly Val Pro Gly Gly Ala Pro Gly Gly Gly Val
 145 150 155 160
 Val Gly Gly Gly Val Pro Gly Gly Gly Val Val Pro Val Val Gly Gly
 165 170 175
 Val Pro Gly Ala Val Pro Gly Val Pro Gly Val Val Val Pro Val Gly

 180 185 190
 Gly Gly Val Val Pro Val Gly Val Pro Gly Gly Val Val Val Pro Gly
 195 200 205
 Val Pro Gly Gly Val Val Gly Val Pro Val Val Gly Val Pro Gly Val
 210 215 220
 Gly Val Val Pro Gly Val Pro Gly Val Pro Val Gly Val Gly Val Val
 225 230 235 240
 Val Pro Gly Val Val Pro Val Gly Val Gly Val Pro Gly Val Pro Gly

 245 250 255

Gly Gly Val Pro Gly Val Gly Val Pro Val Gly Gly Ala Pro Gly Val
260 265 270

Pro Gly Val Pro Gly Val Ala Val Pro Gly Val Val Pro Gly Gly Ala
275 280 285

Pro Gly Val Pro Val Gly Val Val Pro Val Gly Val Val Pro Gly Gly
290 295 300

Gly Val Gly Val Val Pro Gly Gly Val Pro Gly Gly Val Pro Gly Val

305 310 315 320

Pro Ala Gly Val Pro Val Gly Val Pro Val Gly Gly Val Gly Gly Val
325 330 335

Val Pro Gly Gly Gly Ala Pro Gly Val Gly Val Gly Val Pro Ala Gly
340 345 350

Val Pro Gly Gly Val Pro Val Gly Val Gly Gly Val Pro Gly Gly Val
355 360 365

Val Pro Gly Val Pro Ala Gly Gly Gly Val Pro Val Gly Val Pro Gly

370 375 380

Val Gly Gly Val Pro Val Gly Ala Pro Gly Val Pro Gly Gly Val Pro

385 390 395 400

<210> 10

<211> 400

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 10

Pro Ala Gly Ala Val Pro Gly Gly Val Pro Val Ala Gly Gly Val Pro
1 5 10 15

Val Gly Ala Gly Val Val Ala Pro Gly Ala Pro Val Gly Gly Ala Pro
20 25 30

Gly Val Val Gly Val Pro Gly Gly Gly Val Pro Gly Val Pro Val Val
35 40 45

Gly Val Gly Val Gly Val Pro Val Gly Gly Gly Val Gly Val Val Pro
50 55 60

Ala Gly Gly Gly Ala Pro Gly Gly Val Gly Val Val Val Pro Gly Val
65 70 75 80
Val Pro Gly Ala Pro Ala Gly Gly Val Pro Gly Val Ala Pro Gly Gly
85 90 95

Val Gly Gly Val Pro Gly Val Val Pro Gly Gly Gly Val Val Pro Gly
100 105 110
Val Ala Pro Gly Val Pro Gly Val Pro Gly Gly Val Pro Val Gly Val
115 120 125
Pro Val Gly Ala Ala Pro Gly Gly Gly Val Pro Ala Val Gly Val Pro
130 135 140
Gly Val Gly Gly Gly Val Pro Val Val Gly Ala Gly Ala Pro Gly Gly
145 150 155 160

Val Pro Gly Val Ala Val Pro Gly Ala Pro Gly Gly Val Pro Val Gly
165 170 175
Gly Val Pro Gly Gly Gly Val Gly Gly Val Val Pro Gly Gly Val Pro
180 185 190
Gly Ala Pro Val Gly Val Val Pro Gly Val Gly Val Val Pro Gly Val
195 200 205
Pro Ala Gly Val Gly Val Pro Gly Ala Pro Gly Ala Pro Gly Gly Gly
210 215 220

Val Pro Gly Gly Gly Ala Pro Val Gly Gly Ala Gly Gly Val Pro Gly
225 230 235 240
Gly Val Pro Gly Val Pro Gly Ala Pro Gly Val Gly Val Pro Gly Val
245 250 255
Gly Val Pro Val Gly Val Pro Gly Val Pro Gly Val Pro Val Gly Gly
260 265 270
Gly Ala Gly Val Pro Gly Gly Val Ala Pro Gly Val Ala Val Pro Gly
275 280 285

Gly Val Gly Val Pro Val Val Gly Gly Val Pro Gly Ala Val Pro Gly
290 295 300
Val Pro Gly Val Gly Gly Ala Pro Gly Val Pro Gly Gly Gly Val Val

305 310 315 320
 Gly Val Gly Val Pro Gly Gly Val Val Pro Gly Val Pro Gly Ala Gly
 325 330 335
 Ala Pro Gly Gly Gly Ala Pro Gly Ala Pro Val Gly Val Pro Gly Gly
 340 345 350

 Ala Pro Gly Val Val Pro Gly Val Gly Gly Val Pro Gly Ala Gly Gly
 355 360 365
 Gly Val Val Gly Gly Gly Val Val Pro Val Gly Gly Gly Val Val Pro
 370 375 380
 Gly Ala Pro Gly Gly Ala Pro Val Gly Gly Val Val Gly Gly Val Pro
 385 390 395 400
 <210> 11
 <211> 240
 <212> PRT
 <213> Artificial sequence
 <220><223> Synthetic
 <400> 11
 Pro Val Gly Val Pro Gly Val Gly Gly Ala Pro Val Val Val Gly Gly

 1 5 10 15
 Ala Pro Val Gly Val Gly Gly Ala Gly Gly Val Val Pro Gly Ala Pro
 20 25 30
 Val Gly Ala Val Ala Gly Val Gly Ala Pro Gly Gly Ala Val Pro Ala
 35 40 45
 Gly Gly Ala Pro Gly Val Gly Ala Pro Gly Val Val Pro Gly Ala Val
 50 55 60
 Pro Gly Val Gly Val Gly Gly Val Gly Gly Gly Val Val Val Pro Val

 65 70 75 80
 Val Val Gly Val Gly Gly Val Pro Ala Gly Ala Val Pro Val Gly Gly
 85 90 95
 Val Ala Pro Val Gly Val Pro Val Gly Ala Gly Gly Gly Ala Pro Ala
 100 105 110
 Gly Gly Val Pro Val Ala Val Gly Val Pro Val Val Gly Gly Gly Val

115 120 125
Pro Gly Val Gly Val Val Gly Val Ala Pro Gly Ala Pro Gly Ala Gly

130 135 140
Val Gly Val Val Ala Gly Ala Gly Ala Ala Ala Val Ala Pro Val Gly
145 150 155 160
Val Ala Pro Gly Val Pro Gly Val Pro Gly Ala Val Val Gly Val Pro
165 170 175
Gly Gly Val Ala Pro Ala Val Val Val Gly Ala Val Pro Gly Val Val
180 185 190
Gly Val Val Pro Val Gly Ala Pro Val Val Ala Gly Gly Val Ala Val

195 200 205
Ala Pro Val Gly Gly Ala Val Pro Gly Val Pro Gly Gly Val Pro Gly
210 215 220
Gly Gly Ala Pro Gly Ala Pro Val Val Gly Val Pro Val Gly Val Pro
225 230 235 240

<210> 12

<211> 240

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 12

Pro Val Gly Ser Pro Gly Ser Val Pro Gly Val Val Pro Val Ser Gly
1 5 10 15

Val Val Val Gly Gly Gly Ser Val Pro Gly Ser Ser Pro Val Val Gly
20 25 30
Val Pro Ser Gly Val Pro Val Gly Ser Pro Val Val Gly Gly Gly Val
35 40 45
Ser Pro Gly Val Gly Val Pro Gly Val Val Val Pro Gly Val Pro Val
50 55 60
Gly Gly Gly Val Gly Val Pro Gly Gly Gly Val Val Gly Gly Val Pro
65 70 75 80

Val Val Val Gly Val Ser Val Pro Ser Gly Val Ser Pro Gly Gly Gly
85 90 95

Val Val Val Pro Val Gly Ser Val Ser Ser Gly Val Pro Val Val Ser
100 105 110

Val Val Gly Val Ser Val Val Pro Gly Ser Val Pro Val Gly Ser Pro
115 120 125

Ser Ser Gly Val Pro Gly Val Val Pro Gly Ser Pro Gly Gly Val Val
130 135 140

Gly Ser Ser Val Pro Ser Gly Ser Gly Gly Val Ser Val Pro Val Gly
145 150 155 160

Val Gly Gly Ser Gly Val Pro Gly Val Pro Gly Ser Gly Ser Val Val
165 170 175

Pro Val Gly Val Ser Pro Val Gly Gly Gly Ser Gly Ser Gly Val Gly
180 185 190

Ser Gly Ser Pro Gly Gly Ser Val Pro Gly Gly Ser Pro Gly Val Val
195 200 205

Pro Val Gly Ser Pro Gly Val Gly Gly Val Pro Gly Val Val Gly Gly
210 215 220

Ser Gly Ser Pro Gly Val Pro Gly Gly Val Gly Val Gly Val Pro Gly
225 230 235 240

<210> 13

<211> 240

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 13

Pro Val Gly Thr Pro Gly Thr Val Pro Gly Gly Val Val Pro Val Gly
1 5 10 15

Val Thr Pro Val Gly Gly Gly Thr Gly Thr Pro Thr Thr Val Thr Val
20 25 30

Val Gly Gly Val Val Val Pro Gly Val Pro Gly Gly Gly Thr Pro Gly
35 40 45

Val Pro Val Gly Thr Pro Val Val Gly Gly Val Gly Val Gly Thr Gly
50 55 60
Thr Pro Gly Gly Thr Val Pro Gly Gly Thr Pro Gly Val Val Pro Gly
65 70 75 80
Thr Pro Gly Gly Val Val Pro Gly Thr Thr Val Pro Thr Gly Thr Gly
85 90 95
Thr Thr Pro Val Val Gly Val Pro Thr Val Gly Gly Val Thr Pro Gly
100 105 110
Val Gly Gly Val Pro Val Val Val Gly Thr Gly Val Pro Gly Val Pro
115 120 125
Val Gly Gly Gly Val Gly Val Val Gly Val Val Val Pro Gly Thr Val
130 135 140
Pro Val Gly Val Pro Gly Val Val Gly Gly Thr Val Pro Val Gly Thr
145 150 155 160
Pro Thr Thr Gly Val Pro Gly Val Val Pro Thr Gly Val Thr Pro Gly
165 170 175
Gly Val Gly Val Pro Gly Gly Gly Val Val Gly Gly Val Val Pro Gly
180 185 190
Val Thr Val Pro Val Gly Thr Val Thr Thr Gly Val Pro Val Val Thr
195 200 205
Val Gly Val Val Val Pro Val Gly Thr Val Gly Gly Val Val Val Gly
210 215 220
Gly Gly Thr Thr Pro Gly Gly Val Pro Gly Thr Pro Gly Val Gly Gly
225 230 235 240
<210> 14
<211> 240
<212> PRT
<213> Artificial sequence
<220><223> Synthetic
<400> 14
Gly Ala Val Val Val Pro Gly Gly Ala Val Gly Val Pro Ala Gly Ala
1 5 10 15

Val Gly Gly Val Gly Gly Val Leu Val Pro Ala Val Gly Ala Val Pro
20 25 30

Gly Gly Gly Val Pro Ala Val Gly Val Pro Val Ala Gly Val Val Pro
35 40 45

Gly Gly Leu Pro Gly Gly Ala Gly Val Gly Ala Leu Gly Ala Ala Ala
50 55 60

Pro Gly Gly Val Pro Val Gly Ala Ala Leu Pro Gly Val Ala Gly Val
65 70 75 80

Ala Pro Gly Gly Gly Leu Pro Gly Leu Gly Ala Gly Ala Gly Leu Gly
85 90 95

Leu Ala Gly Ala Leu Val Leu Pro Gly Leu Gly Gly Val Pro Gly Val
100 105 110

Pro Gly Gly Gly Leu Leu Pro Gly Gly Val Pro Leu Leu Gly Leu Pro
115 120 125

Gly Val Pro Ala Gly Leu Pro Gly Val Leu Pro Val Gly Leu Leu Leu
130 135 140

Pro Gly Ala Pro Gly Val Pro Ala Leu Gly Val Pro Leu Gly Val Pro
145 150 155 160

Gly Val Ala Pro Gly Leu Pro Gly Ala Gly Gly Leu Pro Gly Val Gly
165 170 175

Ala Pro Gly Leu Leu Pro Val Gly Gly Ala Pro Ala Ala Leu Gly Gly
180 185 190

Leu Gly Leu Pro Gly Gly Gly Ala Leu Gly Leu Leu Ala Pro Gly Gly
195 200 205

Ala Leu Val Pro Gly Val Gly Gly Val Ala Pro Val Ala Ala Gly Ala
210 215 220

Leu Leu Pro Gly Gly Ala Pro Leu Gly Val Pro Ala Leu Leu Ala Leu
225 230 235 240

<210> 15

<211> 240

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 15

Ser Ser Gly Ser Ser Gly Ser Gly Phe Phe Ser Gly Gly Val Pro Ser

1 5 10 15

Gly Gly Gly Val Val Phe Val Pro Ser Phe Phe Gly Phe Phe Pro Ser

20 25 30

Gly Ser Gly Gly Val Val Pro Gly Val Pro Gly Gly Gly Val Gly Val

35 40 45

Pro Gly Phe Val Gly Val Pro Val Gly Gly Phe Pro Val Gly Val Pro

50 55 60

Gly Val Phe Val Pro Ser Val Val Gly Val Pro Gly Phe Phe Val Phe

65 70 75 80

Pro Gly Val Pro Gly Gly Ser Ser Ser Pro Gly Ser Pro Phe Gly Val

85 90 95

Phe Gly Gly Gly Ser Ser Gly Gly Gly Phe Pro Phe Gly Ser Ser Pro

100 105 110

Gly Val Val Val Ser Pro Gly Phe Gly Phe Gly Val Pro Phe Gly Val

115 120 125

Pro Val Gly Gly Ser Val Phe Phe Phe Pro Phe Gly Gly Gly Phe Pro

130 135 140

Gly Gly Phe Pro Ser Gly Ser Phe Gly Val Pro Gly Gly Gly Phe Val

145 150 155 160

Val Pro Phe Gly Val Ser Pro Gly Ser Pro Gly Phe Val Pro Ser Phe

165 170 175

Gly Ser Pro Gly Gly Phe Gly Gly Phe Ser Pro Phe Ser Ser Gly Val

180 185 190

Gly Ser Pro Gly Gly Val Pro Gly Val Gly Val Val Ser Pro Gly Phe

195 200 205

Pro Gly Gly Ser Pro Phe Gly Phe Pro Gly Ser Pro Gly Gly Ser Phe

210 215 220

Gly Gly Ser Pro Ser Val Phe Gly Ser Gly Ser Pro Gly Ser Phe Pro

225 230 235 240

<210> 16

<211> 240

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 16

Ala Ala Gly Gly Thr Gly Phe Phe Thr Gly Gly Ala Pro Gly Gly Thr

1 5 10 15

Pro Thr Ala Phe Gly Gly Phe Thr Pro Phe Thr Phe Ala Gly Gly Gly

20 25 30

Ala Pro Gly Phe Phe Ala Phe Pro Gly Ala Pro Gly Thr Phe Pro Thr

35 40 45

Phe Gly Ala Pro Gly Ala Ala Pro Phe Gly Thr Thr Pro Gly Thr Phe

50 55 60

Phe Gly Phe Pro Gly Gly Thr Pro Gly Thr Pro Gly Ala Ala Thr Pro

65 70 75 80

Gly Phe Phe Gly Ala Phe Gly Phe Pro Gly Ala Phe Ala Thr Phe Gly

85 90 95

Phe Pro Gly Gly Gly Ala Gly Ala Ala Pro Ala Gly Thr Pro Gly Ala

100 105 110

Pro Thr Ala Ala Gly Ala Thr Thr Pro Gly Ala Gly Gly Thr Thr Gly

115 120 125

Gly Ala Pro Thr Gly Gly Gly Phe Ala Pro Thr Phe Phe Gly Phe Phe

130 135 140

Pro Thr Gly Thr Gly Gly Ala Ala Pro Gly Ala Pro Gly Gly Gly Thr

145 150 155 160

Thr Thr Phe Pro Gly Ala Thr Pro Phe Gly Phe Pro Gly Thr Gly Thr

165 170 175

Pro Gly Thr Phe Gly Ala Pro Gly Gly Gly Phe Phe Pro Gly Ala Gly

180 185 190
Phe Gly Gly Gly Ala Pro Gly Ala Pro Ala Ala Ala Gly Gly Gly Thr
195 200 205
Pro Gly Thr Ala Thr Pro Gly Thr Phe Pro Thr Gly Ala Gly Thr Phe
210 215 220
Pro Gly Thr Phe Pro Phe Gly Gly Gly Phe Pro Gly Gly Phe Pro Thr
225 230 235 240

<210> 17

<211> 240

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 17

Thr Ile Leu Pro Gly Ile Gly Leu Pro Gly Gly Leu Pro Leu Thr Ile
1 5 10 15
Gly Gly Ile Leu Pro Ile Leu Leu Gly Thr Thr Ile Gly Gly Ile Pro
20 25 30
Gly Gly Gly Ile Ile Pro Gly Ile Gly Gly Gly Ile Pro Ile Leu Thr
35 40 45
Leu Leu Ile Leu Pro Thr Gly Leu Pro Gly Thr Pro Leu Thr Thr Gly

50 55 60
Thr Pro Gly Ile Thr Gly Leu Gly Thr Pro Leu Gly Ile Gly Leu Ile
65 70 75 80
Gly Thr Gly Ile Pro Gly Gly Gly Thr Gly Ile Pro Thr Gly Ile Pro
85 90 95
Gly Thr Thr Leu Pro Gly Ile Pro Thr Leu Gly Gly Gly Ile Pro Leu
100 105 110
Ile Gly Thr Pro Gly Gly Ile Gly Thr Ile Gly Ile Pro Gly Thr Gly

115 120 125
Thr Ile Ile Gly Ile Pro Gly Gly Gly Leu Leu Leu Ile Pro Gly Ile
130 135 140
Thr Pro Leu Ile Gly Leu Pro Gly Gly Thr Pro Gly Thr Pro Gly Gly

145 150 155 160
 Gly Thr Gly Thr Thr Pro Thr Gly Leu Pro Gly Ile Leu Leu Pro Gly
 165 170 175
 Gly Leu Pro Gly Ile Pro Thr Gly Thr Pro Gly Thr Gly Thr Pro Ile

 180 185 190
 Gly Gly Leu Leu Ile Leu Leu Thr Leu Pro Gly Leu Pro Gly Leu Gly
 195 200 205
 Leu Pro Gly Leu Ile Pro Leu Gly Thr Gly Leu Thr Gly Gly Thr Pro
 210 215 220
 Thr Gly Gly Ile Pro Gly Gly Ile Gly Ile Pro Gly Ile Leu Thr Thr
 225 230 235 240
 <210> 18
 <211> 240
 <212> PRT
 <213> Artificial sequence

 <220><223> Synthetic
 <400> 18
 Pro Gly Thr Tyr Pro Gly Tyr Gly Tyr Val Tyr Pro Thr Thr Gly Gly
 1 5 10 15
 Ile Pro Gly Gly Val Val Pro Gly Gly Gly Thr Lys Leu Val Pro Gly
 20 25 30
 Lys Gly Lys Gly Gly Ala Lys Ala Pro Gly Thr Val Pro Val Gly Ala
 35 40 45
 Gly Gly Gly Lys Ile Val Pro Ile Tyr Gly Ile Ala Pro Gly Lys Tyr
 50 55 60

 Gly Tyr Pro Gly Gly Gly Ile Val Pro Gly Ile Thr Thr Pro Gly Leu
 65 70 75 80
 Pro Thr Gly Lys Lys Pro Tyr Gly Gly Val Pro Val Leu Tyr Gly Lys
 85 90 95
 Leu Pro Gly Ala Pro Gly Ile Pro Thr Ala Gly Ala Pro Gly Tyr Ile
 100 105 110
 Ala Pro Gly Val Pro Gly Gly Leu Val Lys Gly Gly Thr Gly Ile Ala

115				120				125							
Pro	Leu	Gly	Val	Ile	Leu	Val	Tyr	Ile	Gly	Val	Gly	Gly	Ile	Lys	Gly
130				135				140							
Gly	Ala	Leu	Pro	Thr	Gly	Gly	Leu	Tyr	Pro	Gly	Ala	Gly	Thr	Pro	Gly
145				150				155				160			
Tyr	Pro	Val	Gly	Gly	Gly	Ala	Pro	Ala	Gly	Gly	Ile	Ala	Leu	Lys	Pro
165				170				175							
Gly	Ile	Thr	Pro	Gly	Thr	Ala	Ala	Pro	Gly	Leu	Pro	Gly	Lys	Gly	Gly
180				185				190							

Lys	Tyr	Thr	Tyr	Pro	Gly	Gly	Ala	Pro	Gly	Gly	Thr	Gly	Gly	Val	Pro
		195					200					205			
Gly	Tyr	Pro	Gly	Leu	Ala	Leu	Lys	Leu	Gly	Ile	Pro	Thr	Lys	Gly	Gly
		210				215					220				
Gly	Ile	Gly	Leu	Pro	Tyr	Ile	Gly	Leu	Leu	Pro	Gly	Lys	Pro	Gly	Gly
225					230					235				240	

<210> 19

<211> 6

<212> PRT

<213> Artificial sequence

$\langle 220 \rangle \langle 223 \rangle$ Synthetic

<400> 19

Met Ser Lys Gly Pro Gly

1 5

<210> 20

<211> 3

<212> PRT

<213> Artificial sequence

$\langle 220 \rangle \langle 223 \rangle$ Synthetic

<400> 20

Gly Trp Pro

1

<210> 21

<211> 15

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 21

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 22

<211> 19

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<400> 22

Pro Gln Pro Gln Pro Lys Pro Gln Pro Lys Pro Glu Pro Glu Pro Gln

1 5 10 15

Pro Gln Gly

<210> 23

<211> 5

<212> PRT

<213> Artificial sequence

<220><223> Synthetic

<220><221> Xaa

<222> (4)..(4)

<223> any amino acid except proline

<400> 23

Val Pro Gly Xaa Gly

1 5