



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0142681
(43) 공개일자 2021년11월25일

- | | |
|--|---|
| <p>(51) 국제특허분류(Int. Cl.)
C07K 14/475 (2006.01) A61K 38/00 (2006.01)
A61P 35/00 (2006.01) C12N 15/62 (2006.01)</p> <p>(52) CPC특허분류
C07K 14/475 (2013.01)
A61K 38/00 (2013.01)</p> <p>(21) 출원번호 10-2021-7033515</p> <p>(22) 출원일자(국제) 2020년03월20일
심사청구일자 없음</p> <p>(85) 번역문제출일자 2021년10월18일</p> <p>(86) 국제출원번호 PCT/EP2020/057773</p> <p>(87) 국제공개번호 WO 2020/188081
국제공개일자 2020년09월24일</p> <p>(30) 우선권주장
19163970.7 2019년03월20일
유럽특허청(EPO)(EP)</p> | <p>(71) 출원인
오슬로 유니버시티세케투스 에이치에프
노르웨이 오슬로 니달렌 피오 박스 4956 (우:
엔-0424)</p> <p>(72) 발명자
아트라마달, 호바르드
노르웨이, 1827 호벨, 리게셈바이엔 156
코스벨, 올레 외르겐
노르웨이, 0852 오슬로, 스티프탐트만 카스' 바이
1</p> <p>(74) 대리인
안소영</p> |
|--|---|

전체 청구항 수 : 총 24 항

(54) 발명의 명칭 **재조합 CCN 도메인 단백질 및 융합 단백질**

(57) 요약

본 발명은 CCN 패밀리 단백질의 구성원의 트롬보스폰딘 유형 1 반복 상동성 도메인에 상응하거나 이와 관련된 아미노산 서열을 갖는 재조합 단백질 및 이의 용도에 관한 것이다. 또한, 본 발명은 융합 파트너와 조합된 CCN 패밀리 단백질의 구성원의 트롬보스폰딘 유형 1 반복 상동성 도메인에 상응하거나 이와 관련된 아미노산 서열 및 선택적으로 링커 영역을 포함하는 융합 단백질에 관한 것이다. 또한, 신규한 단백질분해효소 저항성 Fc 단편이 본원에 개시되어 있다.

(52) CPC특허분류

A61P 35/00 (2018.01)

C12N 15/62 (2013.01)

C07K 2319/30 (2013.01)

명세서

청구범위

청구항 1

단량체 융합 단백질로서,

(i) CCN 패밀리 단백질의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인의 적어도 일부에 상응하는 폴리펩티드;

(ii) (i)의 아미노산 서열에 N 또는 C 말단에 융합된 단량체 융합 파트너; 및

(iii) 선택적으로 (i)의 폴리펩티드와 (ii)의 단량체 융합 파트너 사이의 펩티드 링커를 포함하고,

(i)의 폴리펩티드는 길이가 40 내지 60개의 아미노산이고, 서열번호 37 또는 2 내지 6으로부터 선택되는 아미노산 서열, 또는 서열번호 37 또는 2 내지 6으로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 서열을 포함하고, 서열번호 37 또는 2 내지 6으로부터 선택되는 상기 서열의 시스테인 잔기 모두가 보존되고,

그리고 (ii)의 단량체 융합 파트너 및 (iii)의 펩티드 링커는 IGF 결합 단백질 상동성 도메인, 폰 빌레브란트 인자 유형 C 반복 상동성 도메인, 또는 CCN 패밀리 단백질의 시스테인 노트(knot) 도메인이 아니거나 이를 포함하지 않는 것인 단량체 융합 단백질.

청구항 2

제1항에 있어서, (i)의 폴리펩티드는 길이가 44 내지 57개 아미노산인 융합 단백질.

청구항 3

제1항 또는 제2항에 있어서, (i)의 폴리펩티드는

(a) 서열번호 1 또는 8 내지 12로부터 선택되는 아미노산 서열; 또는

(b) 서열번호 1 또는 8 내지 12로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 아미노산 서열; 또는

(c) (a) 또는 (b)의 아미노산 서열의 일부를 포함하거나 이로 구성되며(consists of), 상기 일부는 각각 서열번호 37, 6, 2, 3, 4 또는 5의 적어도 44개의 아미노산 서열 또는 각각 서열번호 37, 6, 2, 3, 4 또는 5로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 서열을 포함하는 것인 융합 단백질.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 폴리펩티드는 서열번호 37 또는 2 내지 6으로부터 선택되는 아미노산 서열, 또는 서열번호 37 또는 2 내지 6으로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 서열로 구성되는 것인 융합 단백질.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, (iii)의 폴리펩티드 링커는 50개 이하의 아미노산을 포함하는 것인 융합 단백질.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, (i)의 폴리펩티드는 서열번호 37 또는 2 내지 6, 또는 서열번호 1 또는 8 내지 12로부터 선택되는 상기 서열의 2번 위치에 상응하는 위치에 알라닌 잔기를 포함하는 것인 융합 단백질.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, (i)의 아미노산 서열은 서열번호 7, 38, 42 내지 46 또는 47 내지 51로부터 선택되는 아미노산 서열, 또는 이와 적어도 80%의 서열 동일성을 갖는 서열을 포함하고, 단백질은 서열번호 7, 38, 42 내지 46 또는 47 내지 51의 상기 서열의 2번 위치에 상응하는 위치에 알라닌 잔기를 포함하는

것인 융합 단백질.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 단량체 융합 파트너는 혈청 알부민, 트랜스페린, 및 인간 IgG의 단량체 Fc-단편으로 구성된 군으로부터 선택되는 것인 융합 단백질.

청구항 9

제8항에 있어서, 인간 IgG의 상기 단량체 Fc 단편은 IgG1, IgG2 또는 IgG4의 단량체 Fc 단편인 융합 단백질.

청구항 10

제8항 또는 제9항에 있어서, 단량체 Fc 단편은 비 글리코실화된(aglycosylated) 것인 융합 단백질.

청구항 11

제8항 내지 제10항 중 어느 한 항에 있어서, 단량체 Fc 단편은 안정화 이황화 가교(stabilizing disulphide bridge) 및/또는 단백질 분해효소 안정화 돌연변이를 포함하는 것인 융합 단백질.

청구항 12

제8항 내지 제11항 중 어느 한 항에 있어서, 단량체 Fc 단편은 면역 이펙터(effector) 기능을 갖지 않는 것인 융합 단백질.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서, (i)의 아미노산 서열과 단량체 융합 파트너 사이의 펩티드 링커는 서열번호 20 내지 25, 39, 57, 63, 65 또는 67로 구성된 군으로부터 선택되는 아미노산 서열, 또는 이와 80%의 서열 동일성을 갖는 아미노산 서열을 갖는 것인 융합 단백질.

청구항 14

제1항 내지 제8항 중 어느 한 항에 있어서, 융합 단백질은 서열번호 84, 85, 88, 89, 97, 98, 102, 103, 106, 107, 110, 및 111로 구성된 군으로부터 선택되는 아미노산 서열, 또는 이와 80%의 서열 동일성을 갖는 아미노산 서열을 갖는 것인 융합 단백질.

청구항 15

제1항 내지 제14항 중 어느 한 항에 정의된 바와 같은 단량체 융합 단백질을 암호화하는 DNA 분자.

청구항 16

제15항에 있어서, 상기 분자는 신호 서열을 암호화하는 뉴클레오타이드 서열을 추가로 포함하는 것인 DNA 분자.

청구항 17

제15항 또는 제16항에 있어서, 상기 분자는 서열번호 34, 35, 36, 86, 87, 90, 91, 99, 100, 104, 105, 108, 109, 112 또는 113에 기재된 바와 같은 뉴클레오타이드 서열 또는 임의의 전술한 서열과 적어도 80%의 서열 동일성을 갖는 뉴클레오타이드 서열을 포함하는 것인 DNA 분자.

청구항 18

제15항 내지 제17항 중 어느 한 항에 정의된 바와 같은 DNA 분자를 포함하는 발현 벡터.

청구항 19

제18항에 정의된 바와 같은 벡터를 포함하는 숙주 세포.

청구항 20

치료법에 사용하기 위한 제1항 내지 제14항 중 어느 한 항에 따른 융합 단백질.

청구항 21

4개 도메인의 CCN 패밀리 단백질에 기인하는 세포 신호전달 및 세포 생리 기능을 억제하거나 중화함으로써 장애를 치료 또는 예방하는 데 사용하기 위한 제1항 내지 제14항 중 어느 한 항에 따른 융합 단백질.

청구항 22

섬유증, 또는 섬유증을 나타내는 임의의 병태의 치료 또는 예방에 사용하기 위한 제1항 내지 제14항, 제20항 또는 제21항 중 어느 한 항에 따른 융합 단백질.

청구항 23

암의 치료에 사용하기 위한 제1항 내지 제14항 또는 제20항 내지 제22항 중 어느 한 항에 따른 융합 단백질.

청구항 24

서열번호 7, 38, 42 내지 46, 47 내지 51에 제시된 아미노산 서열 또는 이와 적어도 80%의 서열 동일성을 갖는 서열을 포함하거나 이로 구성된 40 내지 60개 아미노산 길이의 단백질로서, 서열번호 7, 38, 42 내지 46, 47 내지 51의 상기 서열의 2번 위치에 상응하는 위치에 알라닌 잔기를 포함하고, 상기 서열의 모든 시스테인 잔기가 보존되는 것인 단백질.

발명의 설명

기술 분야

[0001] 본 발명은 특히 절단되고/되거나 특정 아미노산 변형을 포함하는 단백질을 포함한, CCN 패밀리 단백질의 구성원의 트롬보스폰딘 유형 1 반복 상동성 도메인(도메인 III)에 상응하거나 이와 관련된 아미노산 서열을 갖는 재조합 단백질 및 이의 용도에 관한 것이다. 또한, 본 발명은 선택적으로 링커 펩티드를 통해 융합 파트너와 조합된 CCN 패밀리 단백질의 구성원의 트롬보스폰딘 유형 1 반복 상동성 도메인에 상응하거나 이와 관련된 아미노산 서열을 포함하는 융합 단백질에 관한 것이다. 특히, 융합 파트너는 단량체 단백질이고, 융합 단백질은 단량체성이다.

[0002] 또한, 신규한 단백질분해효소 저항성 Fc 단편이 본원에 개시되어 있다.

배경 기술

[0003] CCN 단백질은 분비된 당단백질의 패밀리아이다. CCN은 본래 유전자 패밀리 중 처음으로 확인된 세 구성원 Cyr61, CTGF 및 NOV에서 파생된 약어로서 만들어졌다. 그러나 이 약어는 최근 세포 연락망 인자(Cellular Communication Network factor)의 약어로 수정되었고, HUGO 유전자 명명위원회(Gene Nomenclature Committee)에서 승인하였다(Perbal B, Tweedie S and Bruford E, J Cell Commun Signal. 2019 Sep;13(3):435). 이 단백질은 종종 세포의 기질(ECM)과 관련된 모세포 단백질(matricellular protein)로 분류된다. CCN 단백질은 세포를 조직으로 조직화하는 스캐폴드 기능의 일부가 아니라 신호전달 단백질로 간주되며, 독립적인 자가분비 또는 주변분비 신호전달 인자로서, 또는 다른 세포의 신호전달 단백질의 변형인자로 기능할 수 있다. 3개의 Wnt 유도성 신호전달 경로 단백질 세트(WISP1/CCN4, WISP2/CCN5 및 WISP3/CCN6)와 함께, 이들은 CCN1~6으로 이름이 변경되어, 포유동물의 6개의 상동 시스테인이 풍부한 단백질의 패밀리를 구성한다.

[0004] CCN 패밀리의 초기 구성원은 분비를 위한 N 말단 펩티드 신호와 4개의 보존된 도메인이 있는 모듈식 구조를 공유한다. 첫 번째 도메인은 인슐린 유사 성장 인자 결합 단백질(IGFBP)에 대한 서열 상동성을 나타내므로 IGF에 대한 친화도는 미미하지만 IGF 결합 단백질 상동성 도메인으로 알려져 있다. 두 번째 도메인은 세포의 기질(ECM) 단백질에서 흔히 볼 수 있는 폰 빌레브란트 인자(von Willebrand factor) 유형 C 반복(VWC) 상동성 도메인으로 알려져 있다. 세 번째 도메인은 인테그린에 대한 CCN 단백질의 부착에 관여할 수 있는 트롬보스폰딘 유형 I 반복 상동성 도메인(TSP-1)으로 알려져 있다. 네 번째 도메인은 헤파린에 결합하는 것으로 보고된 도메인인 시스테인이 풍부한 C 말단 반복부 또는 시스틴 노트(cystine knot) 상동성 도메인이다. CCN 단백질 패밀리의 다섯 번째 구성원인 WISP2(Wnt1 유도성 신호전달 경로 단백질 2)는 CCN5로도 알려져 있고, 카르복실 말단 시스틴 노트 도메인(도메인 4)이 결여되어 있어 비정형적이다. CCN 단백질 패밀리의 TSP-1 상동성 도메인은 34%의 아미노산 서열 동일성과 25%의 서열 유사성을 공유한다(ClustalOmega에 의해 분석 시, 참고로 아래 참고). 4개 도메인의 CCN 단백질은 VWC 상동성 도메인의 4개 시스테인이 다른 패밀리 구성원 사이에서 보존되지 않는 CCN6

을 제외하고 1차 서열 전체에 걸쳐 38개의 보존된 시스테인을 함유한다. 또한, 카르복실 말단 시스틴 노트 상동 도메인이 결합된 CCN5의 경우, IGFBP, VWC 및 TSP-1 상동 도메인의 모든 시스테인이 다른 CCN 패밀리 구성원과 비교하여 보존된다.

- [0005] 종종 힌지 영역이라고 하는 보존되지 않은 단백질 분해효소 민감성 중심 영역은 단백질을 두 개의 절반으로 나눈다. CCN 단백질의 발현은 환경 자극의 변화에 대한 반응으로, 전사, 전사 후 및 번역 수준에서 조절된다.
- [0006] CCN 단백질 패밀리의 도메인 구성에 대한 정보는 예를 들어 문헌[Liu et al., 2017, Journal of Diabetes, 9, pp. 462-474]에서 찾을 수 있다.
- [0007] 세포 수준에서 CCN 단백질은 세포의 기질과 세포 표면의 계면에서 다양한 조절 역할을 할 수 있다. CCN 단백질은 세포 부착, 이동, 증식, 분화, 세포자멸사, 생존, 노화 및 유전자 발현을 조절할 수 있다. 이러한 세포 기능의 하나 이상의 측면을 세포 유형 특이적 방식으로 조절함으로써, CCN은 배발생 중의 심혈관 및 골격 발달뿐만 아니라, 염증, 상처 치유, 및 성인의 조직 손상 및 회복을 비롯한 복잡한 생물학적 과정을 조정한다. 일반적으로, 4개 도메인의 CCN 1~4 및 CCN6(특히 CCN1, CCN2 및 CCN4)은 전섬유증 활성을 나타낼 수 있는 반면, 3개의 도메인 I~III만을 포함하는 CCN5는 항-섬유증 활성을 갖는다.
- [0008] CCN 단백질은 또한 진행성 섬유증, 예를 들어 간 섬유증 및 특발성 폐 섬유증으로 인한 기관 부전, 및 암 침습 및 전이와 같은 광범위한 병리학적 상태에 관여한다. 이와 관련하여 문헌[Jun and Lau, 2011, Nat. Rev. Drug. Discovery, 10(12), pp. 945-963]을 참조하였다. 다시, 일반적으로 말해, 4개 도메인의 CCN 단백질, 특히 CCN2는 다양한 섬유증 질환의 기전에 연루된 것으로 밝혀졌으나, 이러한 질환의 전임상 질환 모델에서는 반대로 증가된 CCN5 수준이 유익할 수 있는 것으로 나타났다.
- [0009] 문헌[Kaasbøll et al., J. Biol. Chem, 293:46, pp. 17953 - 17970]에서, CCN2로도 알려진 결합조직 성장 인자(CTGF)는 생물학적으로 활성인 CCN2를 방출하기 위해 단백질 분해 절단을 필요로 하는 비활성 프레프로단백질(preproprotein)로서 합성 및 분비되며, 도메인 III~IV로 구성된 C 말단 단편의 동종이량체가 생물학적으로 완전히 활성인 형태의 CCN2를 나타내며, 마지막으로, CCN2의 주요한 보고된 모든 활성은 C 말단 도메인 III~IV 단편의 동종이량체에 의해 재현될 수 있다고 보고되었다. Kaasbøll 등에 의해 보고된 활성 분석 결과, 처리되지 않은 전장 CCN2과 도메인 I~II로 구성된 N 말단 단편 둘 다 생물학적으로 활성이 아닌 것으로 나타났다. 또한, 기질 금속 단백질분해효소(MMP) 활성에 의한 전장 CCN2의 단백질 분해 처리가 이의 잠재적인 활성을 방출하는 것으로 밝혀졌다. 종합적으로, Kaasbøll 등에 의해 보고된 결과는 프레프로-CCN2가 N 말단 도메인 I 및 II에 의해 자가 억제됨을 의미한다. 또한, CCN1 및 CCN3의 C-말단 도메인 III 및 IV 단편이 신속한 세포 신호전달의 활성화 및 세포 생리학적 반응을 유도하기에 충분하다는 것이 발견되었다. 그러나 생물학적 활성의 방출을 위해 CCN1 및/또는 CCN3, 또는 임의의 다른 CCN 단백질 패밀리 구성원의 힌지 영역의 엔도펩티다제 절단이 어느 정도 필요한지는 알려져 있지 않다.
- [0010] CCN2는 발달 중에, 증진된 섬유화 및 조직 섬유증을 포함하는 다양한 병리학적 상태, 및 몇몇 암에서 고도로 발현되는 것으로 알려져 있다(Jun and Lau, 2011, 위의 문헌). CCN 단백질이 광범위한 병리학적 상태에 관여하고, 섬유증의 발달에 기계적으로 관여하는 세포의 단백질이며, 건강한 유기체에서 제한된 발현을 나타낸다는 사실은 이들을 매력적인 치료 표적으로 만든다.
- [0011] 문헌[Jeong et al., 2016, J. American College of Cardiology, 67: 13, pp. 1557 - 1568]은 실험적으로 유도된 심장의 압력 과부하 후 쥐의 심장으로의 CCN5의 아데노 관련 바이러스 매개 유전자 전달의 역할을 조사하는 연구에 대해 보고한다. 이 연구는 CCN5가 심근에서 근섬유모세포의 생성을 억제하고 세포자멸사를 증진시킴으로써 확립된 심장 섬유증을 역전시킬 수 있다는 결론을 내리며, 이는 CCN5가 항-심장 섬유증 치료법을 개발하기 위한 플랫폼을 제공할 수 있음을 시사한다.
- [0012] US2008/0207489에서는 CCN5의 발현 또는 평활근 세포에 대한 CCN5 단백질 투여를 포함하는 평활근 증식 기반 장애를 치료하는 방법이 개시되어 있다.
- [0013] EP 2 556 839에서는, 전장 CCN5 또는 CCN2 ΔCT를 암호화하는 뉴클레오티드 서열을 포함하는 유전자 전달체를 포함하는 조성물이 제공되고, 심부전 치료에서의 이의 역할이 제안된다. EP 2 556 839에서의 CCN2 ΔCT는 K251(uniprot 넘버링) 이후에 절단된 CCN2의 아미노산 서열로 정의된다.
- [0014] CCN5의 과발현이 일부 실험 시스템에서 CCN2 과발현의 표현형과 반대 표현형을 초래하는 것으로 보고되었지만(Jeong et al. 위의 문헌, Yoon et al., J Mol Cell Cardiol, 49(2), 294-303 Aug 2010), CCN5에 의한 4개 도메인의 CCN 단백질의 직접적인 길항작용은 저자가 알기로는 보고된 바 없다. 특히, 다른 CCN 패밀리 구성원의

CCN5/WISP2 매개 길항작용에 대한 구조적 기초는 본 발명에 제시된 작업 이전에 알려지지 않았다.

[0015] 본 발명자들은 전장 CCN2(FL-CCN2)가 비활성 전구체인 프레프로단백질이고 도메인 III 및 IV를 포함하는 단편인 CCN2의 모든 생물학적 관련 활성을 전달하는 것으로 보인다는 점을 초기 단계에서 밝혀냈다. 일반적으로 CCN 단백질이 단백질 분해 활성화를 필요로 하는 비활성 프레프로단백질로서 분비되는 정도는 아직 알려지지 않았다. 그러나 본 발명자들(Kaasbøll et al., J. Biol. Chem. (2018) 293(46) 17953-17970) 및 다른 사람들(Butler, G.S. et al. *Matrix Biol* **59**, 23-38 (2017) and Guillon-Munos, A. et al. *J Biol Chem* **286**, 25505- 25518 (2011))에 의해 증명된 바와 같이, 다수의 단백질 분해효소에 대한 전장 CCN 단백질의 감수성은 변형되지 않은 전장 CCN 단백질이 재조합 단백질 제조 중 및 생체 내 투여 후의 안정성의 이유로 약물로서 매우 부적합할 것임을 의미한다. 치료 단백질로서 변형되지 않은 전장 CCN 단백질을 이용하는 것의 이러한 부적합성은 전장 CCN1(Schutze, N. et al.(2005) *Protein Expr Purif* 42, 219-225) 및 전장 CCN6(Schutze, N et al. (2007) *BMC Cell Biol* 8, 45)에 대해 기술된 바와 같이, 전장 CCN 단백질의 융합 단백질에도 적용된다. CCN 단백질 분야에서는 단백질 분해에 대한 이들 단백질의 감수성이 재조합 CCN 단백질 생산이 매우 어려운 이유 중 하나로 잘 알려져 있다. 또한, Kaasbøll 등(J Biol Chem 2018; 293(46):17953-17970)의 신규한 결과에 기초하여, 재조합 전장 CCN 단백질의 활성이 이전의 단백질 분해 처리에 의존할 수 있어, 약동학 및 약력학을 예측할 수 없게 만들기 때문에, 이들은 이상적인 생물학적 약물과는 거리가 멀 수 있다. 또한, Fc-융합 단백질의 경우, 구성요소, 예를 들어, 펩티드 링커, CCN 단편 및 Fc 단편의 단백질 분해 감수성 외에도, 구성요소의 배열이 재조합 융합 단백질의 효능 및 힘에 있어 중요한 것으로 나타났다. 이의 한 가지 예는 본 발명자들이 발표한 논문(Kaasbøll et al. (2018))에 있는데, 여기서 CCN2의 도메인 III-IV를 함유하는 Fc 융합 단백질의 변이체가 미리 쉽게 예측할 수 없는 방식으로 광범위하게 다양한 활성을 갖는 것으로 밝혀졌다.

[0016] CCN 단백질의 작용은 CCN 단백질의 1차 서열에서 유래된 고농도의 합성 펩티드에 의한 길항 효과에 민감한 것으로 보고되었다. 한 가지 예는 Rat2 섬유모세포에서 재조합 CCN2에 의해 자극된 AKT 인산화를 CCN2의 도메인 I, IGFBP 상동성 도메인에서 유래한 펩티드 및 더 적은 정도이지만 CCN2의 도메인 III, TSP-1 반복 상동성 도메인에서 유래한 펩티드에 의해 억제하는 것이다(Moe et al. al., J. Cell Commun. Signal.(2013) 7:31-47). 또 다른 예는 CCN2의 시스틴 노트 상동성 도메인인 도메인 IV로부터 유래된 펩티드에 의한 간 성상 세포의 CCN2(도메인 IV)에 의해 자극된 부착의 억제이다(Gao R and Brigstock DR., J Biol Chem. 2004 Mar 5; 279(10):8848-55). 또한, CCN1의 도메인 III(Leu et al. J. Biol. Chem, 2003, Vol. 278, No. 36, Issue of September 5, pp. 33801-33808, 2003) 및 CCN1, CCN2, CCN3, CCN5 및 CCN6의 도메인 III(Karagiannis EG and Popel The International Journal of Biochemistry & Cell Biology 39 (2007) 2314-2323)으로부터의 펩티드는 HUVEC 세포를 사용한 시험관 내 분석에서 약간의 항혈관신생 효과(Leu et al. J. Biol. Chem, 2003 and Karagiannis EG and Popel, Int J Biochem Cell Biol 39 (2007))와 1064SK 인간 포피 섬유모세포에 대한 항부착효과(Leu et al. J. Biol. Chem, 2003)를 갖는 것으로 보고되었으며, 이들 펩티드는 본 문헌에 기술된 본 발명의 중심인 보존된 시스테인 중 오직 하나(Leu et al. J. Biol. Chem, 2003) 또는 두 개(Karagiannis EG and Popel, Int J Biochem Cell Biol 39, 2007)만 함유한다. HUVEC 세포(Lu, S et al. (2015) *Nat methods* 12, 329-331) 및 정제된 재조합 CCN2(Kaasbøll et al., J. Biol. Chem. 2018)로부터 내인적으로 발현된 CCN2에서 증명된 바와 같이, CCN 단백질의 도메인 III에 있는 시스테인은 이황화 가교를 생성하는 것으로 알려져 있다. C199-C228(uniprot 넘버링)에 걸쳐 있는 CCN2에서 증명된 이황화 가교는 아미노산 쇄가 스스로 다시 접히는 복잡한 3D 구조를 부여한다. 이는 CCN 단백질의 완전한 도메인 III이 진핵생물 시스템에서 생산된 CCN 단백질의 완전한 도메인 III에서와 같이 시스테인 사이의 이황화 가교에 의해 구조적으로 제한되지 않는 짧은 펩티드에 의해 복제될 것으로 기대할 수 없음을 의미한다. 더욱이, 도메인 I, III, 및 IV의 1차 서열로부터 유래된 펩티드에 의한 CCN2 활성의 억제는, CCN2의 특정 도메인으로부터 유래된 펩티드가 4개 도메인의 CCN 단백질의 억제를 부여할 수 있는지에 관한 지식이 당해 분야에서 부족함을 설명한다.

[0017] 본 발명자들은 이제 CCN 패밀리 단백질의 구조-활성 분석 및 CCN2가 생물학적으로 활성이 되기 위해 단백질 분해 처리를 거쳐야 한다는 관찰에 기초하여, CCN5 단백질의 생물학적 활성 부분이 트롬보스폰딘 유형 I 반복 상동 도메인인 도메인 III임을 발견하였다. 이 새로운 지식은 CCN5의 도메인 III뿐만 아니라 CCN 단백질 패밀리의 다른 구성원의 도메인 III에 기초한 생리활성 구조를 제공하는 결과를 가져왔다.

발명의 내용

[0018] 본 발명자들은 CCN5 및 기타 CCN 단백질의 구조-활성 관계에 대한 통찰을 통해 CCN5 신호전달에 기인하는 세포 신호전달 및 세포 생리학적 기능을 재현하고, 또한 다른 4개 도메인의 CCN 단백질(Cyr61(CCN1이라고도 함),

CTGF(CCN2라고도 함), NOV(CCN3이라고도 함), WISP1(CCN4라고도 함) 및 WISP3(CCN6이라고도 함)에 대응할 수 있는 신규한 생물학적 활성 재조합 단백질을 제공하게 되었다. 다시 말해, CCN 단백질의 TSP-1 상동성 도메인인 도메인 III을 기초로 하는 융합 단백질의 형태를 포함하여, CCN5의 생물학적 활성을 재현하거나 가질 수 있고, 4개 도메인의 CCN 단백질인 CCN1~4 또는 CCN6의 효과를 길항하거나 억제할 수 있는 단백질이 제공된다. 특히, 본원의 단백질은 항-섬유증 활성을 갖고 또한 직접적인 항암 활성을 가질 수 있다.

- [0019] 위에서 언급한 바와 같이, 다른 CCN 단백질, 즉 4개 도메인의 CCN 단백질의 도메인 III(TSP-1 상동성 도메인)은 다른 CCN 도메인의 부재 하에 개별적인 도메인으로 제공되는 경우, 놀랍게도 CCN5의 보고된 활성을 재현하기에 충분한 것으로 밝혀졌다. 따라서, 다시 말해서, 4개 도메인의 CCN 단백질의 도메인 III은 다른 CCN 도메인의 부재 하에 개별적인 도메인으로서(즉, 단리된 도메인으로서) 제공되는 경우, CCN5와 동일한 활성을 갖거나, 또는, 대안적으로 CCN5(즉, 전장의 4개 도메인의 CCN 단백질의 활성과 반대 활성의)의 도메인 III/TSP-1 상동성 도메인으로서 발현된다. 따라서, 본 문헌에 개시된 실험에서, 임의의 CCN 단백질의 단리된 TSP-1 상동성 도메인은 CCN5의 TSP-1 상동성 도메인과 동일한 활성을 나타낼 수 있음은 분명하다. CCN5의 경우를 제외하고, 이는 전장의 CCN 단백질이 발휘하는 활성과 동일하지 않을 수 있다.
- [0020] CCN 단백질의 도메인 III이 단량체 융합 파트너에 융합된 단량체 융합 단백질이 본 발명 및 본원의 개시 내용에 따라 특히 유익하고 유용하다는 점이 밝혀졌다.
- [0021] 제1 양태에 따르면, 본 발명은 단량체 융합 단백질로서,
- [0022] (i) CCN 패밀리 단백질의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인의 적어도 일부에 상응하는 폴리펩티드;
- [0023] (ii) (i)의 아미노산 서열에 N 또는 C 말단에 융합된 단량체 융합 파트너; 및
- [0024] (iii) 선택적으로 (i)의 폴리펩티드와 (ii)의 단량체 융합 파트너 사이의 펩티드 링커를 포함하고,
- [0025] (i)의 폴리펩티드는 길이가 40 내지 60개의 아미노산이고, 서열번호 37 또는 2 내지 6으로부터 선택되는 아미노산 서열, 또는 서열번호 37 또는 2 내지 6으로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 서열을 포함하고, 서열번호 37 또는 2 내지 6으로부터 선택되는 상기 서열의 시스테인 잔기 모두가 보존되고,
- [0026] (ii)의 단량체 융합 파트너 및 (iii)의 펩티드 링커는 IGF 결합 단백질 상동성 도메인, 폰 빌레브란트 인자 유형 C 반복 상동성 도메인, 또는 CCN 패밀리 단백질의 시스테인 노트 도메인이 아니거나 이를 포함하지 않는다.
- [0027] 이하에서 더욱 상세히 기술되는 바와 같이, 서열번호 37 및 2 내지 6은 각각 CCN5, CCN3, CCN2, CCN1, CCN4 및 CCN6의 도메인 III의 44개의 아미노산이 절단된 단편을 나타내며, 이들은 이 도메인의 6개의 보존된 시스테인 잔기를 포함한다. 특히, 이 단편은 도메인의 첫 번째 및 마지막 시스테인 잔기가 측면에 배치되어 있다. 이러한 단편은 특히 효과적이며 단백질 가수분해에 대해 저항적인 것으로 밝혀졌다.
- [0028] 구현예에서, (i)의 폴리펩티드는
- [0029] (a) 서열번호 1 또는 8 내지 12로부터 선택되는 아미노산 서열; 또는
- [0030] (b) 서열번호 1 또는 8 내지 12로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 아미노산 서열; 또는
- [0031] (c) (a) 또는 (b)의 아미노산 서열의 일부를 포함하거나 이로 구성되며(consist of), 상기 일부는 각각 서열번호 37, 6, 2, 3, 4 또는 5의 적어도 44개의 아미노산 서열 또는 각각 서열번호 37, 6, 2, 3, 4 또는 5로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 서열을 포함한다.
- [0032] 서열번호 1 및 8 내지 12는 각각 CCN5, CCN6, CCN3, CCN2, CCN1, 및 CCN4의 도메인 III의 약간 더 긴 N 말단이 절단된 단편을 나타낸다. 이들 단편은 각각의 도메인 III으로부터의 일부 추가적인 C 말단 서열과 함께, 각각 서열번호 37, 6, 2, 3, 4, 및 5의 상응하는 서열을 포함한다.
- [0033] 추가의 구현예에서, (i)의 폴리펩티드는 서열번호 37 또는 2 내지 6, 또는 서열번호 1 또는 8 내지 12로부터 선택되는 상기 서열의 2번 위치에 상응하는 위치에 알려진 잔기를 포함한다. 일부 구현예에서, (i)의 아미노산 서열은 서열번호 38 또는 42 내지 46으로부터 선택되는 아미노산 서열, 또는 이와 적어도 80%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 또 다른 구현예에서, (i)의 아미노산 서열은 서열번호 7 또는 47 내지 51로부터 선택되는 아미노산 서열, 또는 이와 적어도 80%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 이와 관련하여 2번 위치에서 이 잔기의 치환은 단백질의 안정성을 촉진하는 데 유익한 것으로 밝혀졌다.

- [0034] 본 발명의 추가 양태에 따르면, 단량체 융합 단백질은 서열번호 84, 85, 88, 89, 97, 98, 102, 103, 106, 107, 110, 111로 구성된 군으로부터 선택되는 아미노산 서열, 또는 이와 80%의 서열 동일성을 갖는 아미노산 서열을 갖는다.
- [0035] 이들 양태의 구현예에서, 단량체 융합 파트너는 혈청 알부민, 트랜스페린 및 단량체 Fc 단편, 특히 IgG, 더욱 특별하게는 인간 IgG의 단량체 Fc 단편으로부터 선택된다.
- [0036] 위에서 언급된 바와 같이, 본원의 도메인 III 단편의 2번 위치의 치환은 단편의 안정성, 특히 단백질 분해효소 분해에 대한 저항성을 향상시킨다. 본원에서는 도메인 III 단편의 이러한 서열 변형된 변이체는 융합 파트너에 연결되지 않고 그 자체로 유용한 단백질을 나타낸다고 믿어지고 제안된다.
- [0037] 따라서, 본 발명의 또 다른 양태는 또한 서열번호 7, 38, 42 내지 46, 또는 47 내지 51에 제시된 아미노산 서열 또는 이와 적어도 80%의 서열 동일성을 갖는 서열을 포함하거나 이로 구성된 40 내지 60개 아미노산 길이의 단백질을 제공하며, 이때, 단백질은 서열번호 7, 38, 42 내지 46, 47 내지 51의 상기 서열의 2번 위치에 상응하는 위치에 알려진 잔기를 포함하고, 상기 서열의 모든 시스템 잔기가 보존된다.
- [0038] 다른 단백질 및 융합 단백질은 또한 하기에 상세히 기술된 바와 같이 본 발명의 추가 양태로서 제공된다.
- [0039] 본 발명의 추가 양태에 따르면, 화학식 I을 포함하는 재조합 단백질로서,
- [0040] Cys - A - Cys - B - Cys - C - Cys - D - Cys - E - Cys - F (식 I)
- [0041] 여기서,
- [0042] A는 화학식 II의 펩티드
- [0043] A1-A2-A3-A5-A6-A7-A8-A9
- [0044] 이고, 여기서 A1은 P, A, V, I 및 L로 구성된 군으로부터 선택되고; A2는 E, D, A, I, L 및 V로 구성된 군으로부터 선택되고; A3은 G, Q, Y, S, N, W, F로 구성된 군으로부터 선택되고; A4는 A, I, L, V, S, T로 구성된 군으로부터 선택되고; A5는 T, Y, N, G, Q 및 S로 구성된 군으로부터 선택되는 아미노산이고; A6은 A, V, I, L, P, S, E, D, K, R, 및 H로 구성된 군으로부터 선택되는 아미노산이고; A7은 W이고; A8은 G, T, S, Q, Y, N, P, A, V, I 및 L로 구성된 군으로부터 선택되고; A9는 A, P, L, I, V, Q로 구성된 군으로부터 선택되는 아미노산이고;
- [0045] B는 화학식 III의 펩티드
- [0046] B1-B2-B3
- [0047] 이고, 여기서 B1은 G, Q, N, S, Y 및 T로 구성된 군으로부터 선택되는 아미노산이고, B2는 T, S, N, F, Q, H, R 및 K로 구성된 군으로부터 선택되는 아미노산이고; B3은 G, Q, N, S, Y, T로 구성된 군으로부터 선택되는 아미노산이고; 여기서 B1 내지 B3 중 하나는 부재하고;
- [0048] C는 화학식 IV의 펩티드
- [0049] C1-C2-C3-C4-C5-C6-C7-C8-C9-C10-C11-C12-C13-C14
- [0050] 이고, 여기서 C1은 G, Q, N, S, Y 및 T로 구성된 군으로부터 선택되는 아미노산이고; C2는 K, R, H, M, T, S, A, L, I 및 V로 구성된 군으로부터 선택되는 아미노산이고; C3은 G, Q, N, S, Y 및 T로 구성된 군으로부터 선택되는 아미노산이고; C4는 M, F, A, I, L, V 및 W로 구성된 군으로부터 선택되는 아미노산이고; C5는 G, Q, N, S, T, Y, A, I, L 및 V로 구성된 군으로부터 선택되는 아미노산이고; C6은 G, Q, N, S 및 T로 구성된 군으로부터 선택되는 아미노산이고; C7은 H, R 및 L로 구성된 군으로부터 선택되는 아미노산이고; C8은 A, L, I 및 V로 구성된 군으로부터 선택되는 아미노산이고; C9는 G, Q, N, S, T 및 Y로 구성된 군으로부터 선택되는 아미노산이고; C10은 G, Q, N, S, T, Y(바람직하게는 N)로 구성된 군으로부터 선택되는 아미노산이고; C11은 V, P, A, I, L, G, Q, N, S, T, Y, R, K, D 및 E로 구성된 군으로부터 선택되는 아미노산이고; C12는 G, Q, N, S, Y 및 T로 구성된 군으로부터 선택되는 아미노산이고; C13은 H, K, R, A, L, I, V, P, G, Q, N, S, Y 및 T로 구성된 군으로부터 선택되는 아미노산이고; C14는 F, P, W, G, Q, N, S, Y, T, E, 및 D로 구성된 군으로부터 선택되는 아미노산이고;
- [0051] D는 화학식 V의 펩티드

- [0052] D1-D2-D3-D4-D5-D6-D7-D8
- [0053] 이고, 여기서 D1은 R, K, H, D, E, W, P로 구성된 군으로부터 선택되는 아미노산이고; D2는 P, A, L, I, V, M, W, D 및 E로 구성된 군으로부터 선택되는 아미노산이고; D3은 D, E, A, L, I, V, R, K 및 H로 구성된 군으로부터 선택되는 아미노산이고; D4는 G, Q, S, Y, T, R, L, K 및 H로 구성된 군으로부터 선택되는 아미노산이고; D5는 G, Q, N, S, Y, T, D 및 E로 구성된 군으로부터 선택되는 아미노산이고; D6은 H, R, K, G, Q, N, S, Y, 및 T로 구성된 군으로부터 선택되는 아미노산이고; D7은 L, H 및 R로 구성된 군으로부터 선택되는 아미노산이고; D8은 A, L, I 및 V로 구성된 군으로부터 선택되는 아미노산이고;
- [0054] E는 화학식 VI의 펩티드
- [0055] E1-E2-E3-E4
- [0056] 이고, 여기서 E1은 P, A, L, I, V, M, W, G, Q, N, S, T, Y, D, 및 E로 구성된 군으로부터 선택되는 아미노산이고; E2는 P, A, L, I, V, M, W, G, Q, N, S, T, Y로 구성된 군으로부터 선택되는 아미노산이고; E3은 R, K, H, G, Q, N, S, T 및 Y로 구성된 군으로부터 선택되는 아미노산이고; E4는 P, A, L, I 및 V로 구성된 군으로부터 선택되는 아미노산이고; F는 부재하거나 또는 약 13개 이하의 아미노산 서열이고,
- [0057] 여기서 제조합 단백질은 40 내지 60개의 아미노산을 포함하는 것인, 제조합 단백질이 제공된다.
- [0058] 위의 양태의 일 구현예에 따르면, 화학식 I의 제조합 단백질이 제공되며, 여기서
- [0059] A1은 P, I, 및 L로 구성된 군으로부터 선택되고; A2는 E, V, 및 A로 구성된 군으로부터 선택되고; A3은 W, Q, 및 Y로 구성된 군으로부터 선택되고; A4는 S, T, 및 A로 구성된 군으로부터 선택되고; A5는 T 및 S로 구성된 군으로부터 선택되는 아미노산이고; A6은 A, E, P, S 및 K로 구성된 군으로부터 선택되는 아미노산이고; A7은 W이고; A8은 G, S 및 T로 구성된 군으로부터 선택되고; A9는 P, Q 및 A로 구성된 군으로부터 선택되는 아미노산이고;
- [0060] B1은 세린(S)이고; B2는 T, K 및 R로 구성된 군으로부터 선택되는 아미노산이고; B3은 T 및 S로 구성된 군으로부터 선택되는 아미노산이고;
- [0061] C1은 아미노산 G이고; C2는 T, L 및 M으로 구성된 군으로부터 선택되는 아미노산이고; C3은 G이고; C4는 M, F, I 및 V로 구성된 군으로부터 선택되는 아미노산이고; C5는 S 및 A로 구성된 군으로부터 선택되는 아미노산이고; C6은 T 및 N으로 구성된 군으로부터 선택되는 아미노산이고; C7은 R이고; C8은 V 및 I로 구성된 군으로부터 선택되는 아미노산이고; C9는 S 및 T로 구성된 군으로부터 선택되는 아미노산이고; C10은 아스파라긴 N이고; C11은 Q, R, D, V 및 E로 구성된 군으로부터 선택되는 아미노산이고; C12는 아스파라긴 N이고; C13은 R, A, P 및 S로 구성된 군으로부터 선택되는 아미노산이고; C14는 F, Q, S, E 및 N으로 구성된 군으로부터 선택되는 아미노산이고;
- [0062] D1은 R, E, 및 W로 구성된 군으로부터 선택되는 아미노산이고; D2는 L, M, 및 P로 구성된 군으로부터 선택되는 아미노산이고; D3은 E, L, V 및 R로 구성된 군으로부터 선택되는 아미노산이고; D4는 T, K, 및 Q로 구성된 군으로부터 선택되는 아미노산이고; D5는 Q 및 E로 구성된 군으로부터 선택되는 아미노산이고; D6은 R, T, S, 및 K로 구성된 군으로부터 선택되는 아미노산이고; D7은 아르기닌(R)이고; D8은 L, 및 I로 구성된 군으로부터 선택되는 아미노산이고;
- [0063] E1은 L, M, E, N 및 Y로 구성된 군으로부터 선택되는 아미노산이고; E2는 S, V, L 및 I로 구성된 군으로부터 선택되는 아미노산이고; E3은 Q 및 R로 구성된 군으로부터 선택되는 아미노산이고; E4는 P이고; F는 부재하거나 또는 PPSRGRSPQNSAF, GQPVYSSL, EADLEEN, EQEPEQPTD, DVDIHTLI, 및 DSNILKTIKIP로 구성된 군으로부터 선택되는 아미노산 서열을 포함하는 13개 이하의 아미노산의 펩티드이고,
- [0064] 여기서 제조합 단백질은 총 44 내지 57개의 아미노산을 포함한다.
- [0065] 위의 양태의 또 다른 구현예에 따르면, F가 완전히 부재하거나, 부분적으로 부재하거나, PPSRGRSPQNSAF의 아미노산 서열을 포함하는 약 13개 아미노산의 펩티드인 화학식 I의 제조합 단백질이 제공된다.
- [0066] 더욱 구체적으로, 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 서열번호 38로 구성된 군으로부터 선택되는 아미노산 서열; 및 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 서열번호 38의 아미

노산 서열과 50% 초과 서열 동일성을 갖는 이의 단편 또는 변이체를 포함하는 제조합 단백질이 제공된다.

- [0067] 본 발명의 또 다른 양태에 따르면, 상기 정의된 바와 같은 제조합 단백질이 제공되며, 상기 단백질은 폐결핵된다.
- [0068] 또 다른 양태에 따르면, 본 발명은 융합 단백질로서,
- [0069] (i) CCN 패밀리 단백질의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인;
- [0070] (ii) (i)의 TSP-1 반복 상동성 도메인에 N 또는 C 말단에 융합된 융합 파트너로서, 상기 융합 파트너는 혈청 알부민, 트랜스페린 및 인간 IgG의 Fc 단편으로 구성된 군으로부터 선택되는 것인 융합 파트너;
- [0071] (iii) 선택적으로, TSP-1 반복 상동성 도메인 및 Fc 단편((i)의 TSP-1 반복 상동성 도메인에 N 또는 C 말단에 융합됨) 사이의 펩티드 링커;
- [0072] 를 포함하는 융합 단백질을 제공한다.
- [0073] 위의 양태의 구현예에 따르면, 본 발명의 추가 양태로서 전술한 바와 같은 본 발명에 따른 제조합 단백질을 포함하는 융합 단백질이 제공된다.
- [0074] 본 발명에 따른 융합 단백질의 융합 파트너는 IgG1, IgG2 또는 IgG4의 Fc 단편, 혈청 알부민 및 트랜스페린으로 구성된 군으로부터 선택되는 일 구현예에 따른다.
- [0075] 추가 구현예에 따르면, 융합 파트너 (ii)가 안정화 이황화 가교를 포함하는 IgG1, IgG2 또는 IgG4의 Fc 단편인 융합 단백질이 제공된다. 이러한 돌연변이는 단백질의 열 안정성을 증가시킬 수 있다. 안정화 돌연변이는 공지되어 있고 당해 분야에 보고되어 있다.
- [0076] 또 다른 추가 구현예에 따르면, 융합 파트너 (ii)가 S228P(IgG4 참조), E233P(IgG1 및 IgG4 참조), F234A(IgG4 참조), L234A(IgG1 참조), L234V(IgG1 참조), F234V(IgG4 참조), L235A(IgG1 및 IgG4 참조), ΔG236(IgG1 및 IgG4 참조) 및 ΔK447(IgG1, IgG2 및 IgG4 참조)로 구성된 군으로부터 선택되는 하나 이상의 돌연변이를 포함하는 IgG1, IgG2 또는 IgG4의 Fc 단편인 융합 단백질이 제공된다.
- [0077] 또 다른 구현예에 따르면, 융합 단백질은 서열번호 15, 서열번호 16, 서열번호 17, 서열번호 18, 및 서열번호 19로 구성된 군으로부터 선택되는 Fc 단편을 포함할 수 있다.
- [0078] 또 다른 구현예에 따르면, 융합 단백질은 서열번호 20; 서열번호 21; 서열번호 22, 서열번호 23, 서열번호 24, 서열번호 25, 및 서열번호 39로 구성된 군으로부터 선택되는 링커를 포함할 수 있다.
- [0079] 일 구현예에 따르면, 링커는 아미노산 서열 (EAAAK)_n을 포함하며, 여기서 n은 적어도 4이고, 바람직하게는 n은 8이다.
- [0080] 또 다른 구현예에 따르면, 융합 단백질은 서열번호 26, 서열번호 27; 서열번호 28, 서열번호 29, 서열번호 30, 서열번호 31, 서열번호 40 및 서열번호 41로 구성된 군으로부터 선택되는 아미노산 서열을 포함한다.
- [0081] 또 다른 구현예에 따르면, (ii)의 융합 파트너는 혈청 알부민이다.
- [0082] 본 발명의 제2 양태의 또 다른 구현예에 따르면, (ii)의 융합 파트너는 트랜스페린이다.
- [0083] 또한, 본 발명은 또 다른 양태에 따르면, 본 발명에 따른 제조합 단백질, 단백질 또는 융합 단백질을 암호화하는 핵산 분자(예를 들어, DNA)를 제공한다.
- [0084] 이 양태의 일 구현예에 따르면, 서열번호 34, 서열번호 35, 서열번호 36, 또는 서열번호 86, 87, 90, 91, 99, 100, 104, 105, 108, 109, 112 또는 113에 제시된 바와 같은 핵산 서열 및 서열번호 34, 서열번호 35, 서열번호 36, 또는 서열번호 86, 87, 90, 91, 99, 100, 104, 105, 108, 109, 112 또는 113과 약 80%의 서열 동일성을 갖는 핵산 서열을 포함하는 DNA 서열이 제공된다.
- [0085] 또한, 본 발명의 또 다른 양태에 따르면, 본 발명에 따른 DNA 서열을 포함하는 발현 벡터가 제공된다. 또한, 본 발명에 따른 발현 벡터를 포함하는 숙주 세포가 제공된다.
- [0086] 마지막으로, CCN 패밀리 단백질, 위에서 정의된 바와 같은 단백질 및 융합 단백질의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인은 4개 도메인의 CCN 패밀리 단백질에 기인하는 세포 신호전달 및 세포 생리 기능을 억제하거나 중화함으로써 장애의 치료 또는 예방을 위한 의약으로서 사용하기 위하여 제공된다.

- [0087] 일 양태에서, 치료법에 사용하기 위한, 본원에 정의된 바와 같은 융합 단백질 또는 단백질이 제공된다.
- [0088] 융합 단백질 또는 단백질은 4개 도메인의 CCN 단백질의 활성, 특히 4개 도메인의 CCN 단백질의 원하지 않거나 비정상적인 활성과 관련된 병태의 치료에 사용하기 위한 것일 수 있다. 활성은 섬유증 효과와 관련될 수 있다. 활성은 전섬유증(pro-fibrotic) 활성일 수 있다.
- [0089] 또 다른 양태에서, 섬유증, 또는 섬유증을 나타내는 임의의 병태(즉, 섬유증 병태 또는 질환)의 치료 또는 예방에 사용하기 위한, 본원에 정의된 바와 같은 융합 단백질 또는 단백질이 제공된다. 추가 양태에서, 암의 치료에 사용하기 위한, 본원에 정의된 바와 같은 융합 단백질 또는 단백질이 제공된다. 또한, 염증성 또는 자가면역 질환, 또는 대사 질환의 치료에 사용하기 위한, 본원에 정의된 바와 같은 융합 단백질 또는 단백질이 제공된다.
- [0090] 또한, 본 발명의 이러한 양태에 따라 본원에 정의되거나 기술된 바와 같은 병태 또는 질환을 치료 또는 예방하기 위한 의약의 제조를 위한, 본원에 정의된 바와 같은 단백질 또는 융합 단백질의 용도가 제공된다.
- [0091] 이러한 양태는 또한 본원에 정의되거나 기술된 바와 같은 병태 또는 질환 치료 또는 예방하는 데 사용하기 위한, 본원에 정의된 바와 같은 단백질 또는 융합 단백질을 포함하는 조성물(예를 들어, 제약 조성물)을 포함한다.
- [0092] 이러한 양태는 또한 본원에 정의되거나 기술된 바와 같은 병태 또는 질병을 치료 또는 예방하는 방법을 포함하고, 상기 방법은 본원에 정의된 바와 같은 단백질 또는 융합 단백질, 특히 유효량의 상기 단백질 또는 융합 단백질을 이를 필요로 하는 대상체에게 투여하는 단계를 포함한다.
- [0093] 발명의 상세한 설명
- [0094] CCN5의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인이 CCN5/WISP2의 세포 신호전달 기능을 부여하는 완전 활성 구조라는 놀라운 발견에 기초한 본 발명은 언급된 바와 같다. CCN5의 TSP-1 반복 상동성 도메인의 활성에 대한 이러한 새로운 통찰에 기초하여, 본 발명자들은 4개 도메인의 CCN 단백질, 즉, CCN1, CCN2, CCN3, CCN4 및 CCN6으로 인한 세포 신호전달 및 세포 생리 기능을 억제하거나 중화하는 데 사용될 수 있는 본 발명에 따른 단백질, 재조합 단백질 및 융합 단백질을 제공한다. 이전에는 CCN5의 활성을 재현하는 데 필요한 전장의 CCN5의 부분이 밝혀지지 않았기 때문에, 그리고 전장의 CCN 분자는 단백질 분해에 매우 민감하고 활성의 균질한 형태로 생산하기가 어렵기 때문에 이 새로운 통찰은 CCN5를 기초로 하는 안정적인, 균질한, 약물 유사 분자의 형성을 가능하게 하는 데 결정적인 것으로 여겨진다. 또한, CCN5의 특정 부분이 예를 들어 전장 단백질의 일시적인 과발현 시 관찰된 활성을 재현하기에 충분할 수 있다는 의미는 CCN 단백질이 모세포 단백질로서 작용한다는 당해 분야의 일반적인 의견과는 대조된다. CCN 단백질 및 일반적으로 모세포 단백질의 작용 기전에 대한 일반적인 견해는 CCN 단백질의 다른 부분이 다양한 다른 ECM 단백질 및 세포 표면 수용체와 상호작용하여 세포 신호전달의 직접적인 조절자로서 작용하기 보다는 이들의 활성을 조절한다는 것이다. CCN5의 TSP-1 반복 상동성 도메인의 활성에 대한 새로운 지식과 CCN 단백질 패밀리의 다른 구성원과의 구조적으로 밀접한 관계에 대한 지식은 다른 CCN 패밀리 단백질의 TSP-1 반복 상동성 도메인도 4개 도메인의 CCN 패밀리 단백질의 세포 신호전달 기능을 억제하는 데 활용될 수 있음을 시사한다. 일 양태에 따르면, 본 발명의 재조합 단백질 및 융합 단백질은 A549 세포에서 AKT(Ser473)의 인산화를 억제한다.
- [0095] 상기 세포 신호전달의 억제는 다양한 장애의 치료와 관련이 있다. 예를 들어, CCN2는 여러 질환, 특히 증진된 섬유 발생 및 조직 섬유화가 특징적인 병태생리학적 특징인 질환에 연루되어 있다.
- [0096] 예를 들어, CCN2 단독의 과발현은 폐에서 섬유증을 유발하기에 충분함이 밝혀졌다(문헌[Sonnlyal et al., *Arthritis Rheum* **62**, 1523-1532 (2010)] 참조). CCN2는 또한 블레오마이신 유발 폐 섬유증(Bonniaud, P. et al. *Am J Respir Cell Mol Biol* **31**, 510-516 (2004)), 방사선 유발 폐 섬유증(Bickelhaupt, S. et al. *J Natl Cancer Inst* **109** (2017)), 및 PTEN(포스파타제 및 텐신 상동체(Phosphatase And Tensin Homolog)) 발현 상실로 인한 폐 섬유증(Parapuram, S.K. et al. *Matrix Biol* **43**, 35-41 (2015))에 필수적인 것으로 밝혀졌다. 또한, 다른 유발 물질의 부재 하에, CCN2가 폐 Clara 세포(Wu, S. et al. *Am J Respir Cell Mol Biol* **42**, 552-563 (2010)), 폐포 유형 II 상피 세포(Chen, S. et al. *Am J Physiol Lung Cell Mol Physiol* **300**, L330-340)에서 발현 및 분비될 때, 섬유모세포 특이적 프로모터로부터 발현될 때(Sonnlyal et al (2010), *supra*, Sonnylal, S. et al., *J Cell Sci* **126**, 2164-2175 (2013)) 또는 아데노바이러스에 의해 전달될 때(Bonniaud, P. et al., *Am J Respir Crit Care Med* **168**, 770-778 (2003)), 폐 섬유증을 유도하는 것으로 밝혀졌다. 따라서, 여러 보고는 모두 CCN2가 피부 또는 폐에서 섬유증을 유발하기에 충분할 뿐만 아니라, 여러 질환 모델에서 완전히 발달된 섬유증 표현형에 필요하다는 결론을 뒷받침한다. 폐 섬유증은 인간 질환 특발성 폐 섬유증(IPF)의 특징이지만,

만성 폐쇄성 폐 질환(COPD)(Jang, J.H. *et al.*, *COPD* **14**, 228-237 (2017)) 및 전신 경화증의 환경에서도 발생한다. 실제로, 폐 섬유증은 전신 경화증 환자의 최대 40%에서 주요 사망 원인으로 보고된 바 있다(Tyndall AJ *et al.*, *Ann Rheum Dis.* 2010 Oct;69(10):1809-15). CCN2 및 WISP1과 같은 다른 CCN 단백질도 인간 환자에서 두 IPF(Konigshoff, M. *et al.*, *J Clin Invest* **119**, 772-787 (2009) 및 COPD(Jang *et al.*, 위의 문헌)의 병태생리에 연루되어 있다.

[0097] 또 다른 예는 신생물 장애이다. 예를 들어, 유방암의 환경에서 CCN2는 삼중 음성 유방암 모델(MDA-MB-231)에서 골 전이에 기여하는 것으로 나타났다(Kang, Y. *et al.*, *Cancer Cell* **3**, 537-549 (2003)). 또한, 높은 수준의 CCN2를 발현하는 세포주인 삼중 음성 유방암 세포(MDA-MB-231)에서 CCN2의 녹다운(Chen, P.S. *et al.*, *J Cell Sci* **120**, 2053-2065 (2007))은 이들 세포에서 이동 능력을 감소시킨 반면, 낮은 내인성 CCN2 발현과 함께, 호르몬 수용체 양성 MCF-7 유방암 세포주에서 CCN2의 과발현(Chen *et al.*, 위의 문헌)은 MCF-7 유방암 세포의 이동 능력을 증가시켰다(Chen *et al.*, 위의 문헌, Chien, W. *et al.*, *Int J Oncol* **38**, 1741-1747 (2011)). 이후의 보고에서는 또한 MCF-7 세포에서 CCN2의 과발현이 화학내성을 증가시킨 반면, MDA-MB-231 세포에서 CCN2의 녹다운은 화학내성을 감소시킴을 발견하였다(Wang, M.Y. *et al.*, *Cancer Res* **69**, 3482-3491 (2009)). CCN2에 의해 부여된 화학내성의 증가는 다른 유방암 세포에 대해서도 보고되었다(Lai, D *et al.*, *Cancer Res* **71**, 2728-2738 (2011)). 또한, 과발현 또는 녹다운 연구를 통해, CCN2가 유방암 세포의 상피에서 중간엽으로의 전이(EMT) 및 부착 비의존성 성장(맘모스피어(mammosphere) 형성)의 능력 증가에 기여하는 것으로 나타났다(Chen *et al.*, 위의 문헌, Zhu, X. *et al.*, *Oncotarget* **6**, 25320-25338 (2015)). CCN2에 의해 유도된 증가된 화학내성 및 증진된 EMT의 발견은 다른 암 유형에서도 EMT와 화학내성 사이의 연관성과 일치한다(Fischer, K.R. *et al.*, *Nature* **527**, 472-476 (2015), Zheng, X. *et al.*, *Nature* **527**, 525-530 (2015)).

[0098] 특정 양태에서, 본 발명은 CCN 패밀리 단백질의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인의 적어도 일부에 상응하는 폴리펩티드를 포함하는 위에서 정의된 바와 같은 단량체 융합 단백질을 제공하며, 여기서 TSP-1 상동성 도메인 서열은 절단 및/또는 변형될 수 있지만, 여기서 도메인의 시스테인 잔기는 보존된다. 이 폴리펩티드는 편의상 본원에서 "TSP-1 폴리펩티드"로 지칭될 수 있으며, 따라서 이 용어는 특정 천연 TSP-1 상동성 도메인 서열에만 제한을 전달하거나 암시하지 않는 것으로 이해되어야 한다. 용어 "TSP-1 폴리펩티드"는 "TSP-1 도메인 단백질" 또는 TSP-1 도메인 서열"과 동의어로 또는 상호교환적으로 사용될 수 있다.

[0099] 아래의 실시예에서 증명되는 바와 같이, 놀랍게도 단량체 융합 파트너는 이량체 융합 단백질을 생산하는 IgG 단백질로부터 유래된 Fc 단편과 같은 이량체 융합 파트너와 비교하여 활성 및 안정적인 단백질을 생산하는 데 유리하다는 점이 발견되었다. 단량체성 융합체는 이들이 포함하는 TSP-1 도메인 폴리펩티드의 활성을 유지한다. 또한, 이 단백질은 단백질 분해와 관련하여 이를 포함하여 안정적이다. 아래에서 추가로 기술되는 바와 같이, 특히 위에서 언급한 Ala 치환을 포함하여, TSP-1 폴리펩티드의 아미노산 서열을 변형함으로써 단백질 분해에 대한 저항성을 개선할 수 있다.

[0100] 따라서, 융합 단백질의 성분 (i)의 폴리펩티드는 서열번호 37 또는 2 내지 6으로부터 선택되는 상기 서열, 또는 서열번호 1 또는 8 내지 12로부터 선택되는 서열과 비교하여, 삽입, 결실, 치환, 돌연변이 또는 이의 임의의 조합을 포함할 수 있으나, 이 폴리펩티드는 상기 서열과 적어도 80%의 서열 동일성을 보유하고, 상기 서열의 모든 시스테인 잔기는 보존된다.

[0101] 또 다른 양태에서, 본 발명은 CCN 패밀리 단백질의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인의 적어도 일부에 상응하나 융합 단백질의 맥락에서는 그렇지 않은 폴리펩티드로 구성되거나 이를 포함하는 단백질(예를 들어, 재조합 단백질)을 제공하며, 여기서 TSP-1 도메인 서열은 절단 및/또는 변형될 수 있고, 서열번호 37 또는 2 내지 6, 또는 서열번호 1 또는 8 내지 12의 2번 위치에 상응하는 위치에서 Ala 치환을 포함하지만, 도메인의 시스테인 잔기는 보존된다. 환원하면, TSP-1-도메인 단백질은 융합 파트너와 같은 또 다른 성분 없이, 또는 이와 독립적으로 제공될 수 있다. 따라서, TSP-1 도메인 단백질은 또 다른 단백질 도메인 또는 성분 또는 다른 기능적 또는 구조적 단백질 서열에 융합되거나 연결되지 않는다. 편의상 이러한 단백질은 "Ala 치환 단백질"로 지칭될 수 있다.

[0102] 본원에서 사용되는 용어 "보존된"은 주어진 서열의 잔기가 결실되거나 치환되지 않음을 의미한다. 환원하면, "보존된"이라는 용어는 "유지된"이라는 용어와 동의어로(그리고 상호교환적으로) 사용된다. 이는 단순히 시스테인 잔기가 서열로부터 제거되지 않음을 의미한다. 따라서, 위의 맥락에서 이는 서열번호 37 또는 2 내지 6 또는 1 또는 8 내지 12로부터 선택되는 서열의 시스테인 잔기가 결실되거나 치환되지 않음을 의미한다. 보존된 잔기 사이(예를 들어, 보존된 시스테인 사이)의 추가 잔기의 삽입 또는 보존되지 않은(예를 들어, 비 시스테인) 잔기

의 결실은 본래 기준 서열(예를 들어, 서열번호 37 또는 2 내지 6으로부터 선택되는 서열)에서 이들의 위치와 비교하여 폴리펩티드 서열 내의 보존된 잔기의 위치를 변경할 수 있다. 그러나 이러한 잔기는 본원에 정의된 바와 같이 여전히 "보존된" 것으로 간주된다. 따라서, "보존된"이라는 용어는 시스테인 잔기의 위치(또는 더욱 구체적으로는, 위치 번호)에 대한 제한 또는 한정을 의미하지 않는다.

- [0103] 일부 구현예에서, (i)의 폴리펩티드는
- [0104] (a) 서열번호 1 또는 8 내지 12로부터 선택되는 아미노산 서열; 또는
- [0105] (b) 서열번호 1 또는 8 내지 12로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 아미노산 서열; 또는
- [0106] (c) (a) 또는 (b)의 아미노산 서열의 일부를 포함하거나 이로 구성되며, 상기 일부는 각각 서열번호 37, 6, 2, 3, 4 또는 5의 적어도 44개의 아미노산 서열 또는 각각 서열번호 37, 6, 2, 3, 4 또는 5로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 서열을 포함한다.
- [0107] 위에서 언급된 바와 같이, (ii)의 단량체 융합 파트너 및 (iii)의 펩티드 링커는 IGF 결합 단백질 상동성 도메인, 폰 빌레브란트 인자 유형 C 반복 상동성 도메인, 또는 CCN 패밀리의 단백질의 시스테인 노트 도메인이 아니거나 이를 포함하지 않는다. 대안적으로, 본 발명의 융합 단백질에 존재할 수 있는 CCN 패밀리의 단백질의 유일한 도메인은 TSP-1 상동성 도메인이다.
- [0108] 유사하게 그리고 비슷하게, 융합 단백질이 아닌 Ala 치환 단백질의 맥락에서, 단백질은 (TSP-1 도메인 단백질 이외의) 임의의 다른 CCN 도메인을 포함하지 않는다.
- [0109] 일부 구현예에서, (i)의 폴리펩티드, 또는 Ala 치환 단백질은 위에서 정의된 바와 같은 TSP-1 상동성 도메인의 일부만을 포함할 수 있다. 본 발명자들은 요구되는 TSP-1 도메인의 최소 단편이 서열번호 37, 6, 2, 3, 4 또는 5의 44개 아미노산 서열임을 발견하였다. 따라서, 일부 구현예에서, (i)의 폴리펩티드는 길이가 적어도 44개 아미노산이다. 일부 구현예에서, (i)의 폴리펩티드는 길이가 44 내지 57개 아미노산이다. 그러나 위에서 언급한 바와 같이 시스테인 잔기 사이에 있는 44개 아미노산 최소 단편에서 아미노산의 하나 이상의 결실이 있을 수 있다. 따라서, 일부 구현예에서 TSP-1 폴리펩티드의 길이는 44개 미만의 잔기, 즉 40 내지 43개의 잔기일 수 있다.
- [0110] 일부 구현예에서, (i)의 폴리펩티드는 서열 37 또는 2 내지 6으로부터 선택되는 아미노산 서열, 또는 서열 37 또는 2 내지 6으로부터 선택되는 서열과 적어도 80%의 서열 동일성을 갖는 서열로 구성된다.
- [0111] 전술한 바와 같이, 융합 단백질을 포함하는 본 발명의 단백질은 CCN5의 활성, 보다 특히 생물학적 활성을 나타낸다(즉, 발휘하거나 갖는다). 구현예에서, 단백질은 CCN5의 TSP-1 상동성 도메인의 활성을 보유하거나 나타내거나 가질 수 있다. 대안적으로, 단백질은 CCN 단백질의 단리된 TSP-1 상동성 도메인의 활성, 특히 생물학적 활성을 나타내는(또는 발휘하거나 갖는) 것으로 정의될 수 있다. 전술한 내용은 도메인의 임의의 활성, 및 TSP-1 상동성 도메인의 항섬유증 효과를 반영하는 특정 활성화에 적용될 수 있다. 이러한 활성화는 도메인의 임의의 특정 생물학적 효과를 기초로 하여, 임의의 편리한 분석 또는 방법을 사용하여 분석(또는 테스트 또는 검출)될 수 있다.
- [0112] 주어진 단백질의 활성화는 AKT의 인산화에 미치는 단백질의 영향을 분석함으로써 편리하게 평가될 수 있음을 주목한다. 특히, 주어진 단백질은 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 AKT(Ser-473)의 인산화를 억제하는 능력에 대해 분석될 수 있다. 당업자는 동일한 활성을 평가하거나 또는 다른 관련된 항섬유증 활성을 평가하기 위해 다른 유사한 분석이 고안될 수 있음을 이해할 것이다.
- [0113] 위에서 언급한 바와 같이, 본 발명의 다른 양태에서, 위의 화학식 I에 따른 아미노산 서열을 포함하는, 4개 도메인의 CCN 단백질에 기인하는 세포 신호전달 및 세포 생리 기능을 억제하거나 중화하는 재조합 단백질이 제공된다.
- [0114] Cys - A - Cys - B - Cys - C - Cys - D - Cys - E - Cys - F
- [0115] 여기서 A, B, C, D, E 및 F는 위에서, 그리고 첨부된 청구항에서 정의된 바와 같다.
- [0116] 화학식 I은 모두 6개의 보존된 시스테인을 포함하며, 아미노산이 단백질의 활성화에 영향을 미치지 않고 치환될 수 있음(아래에서 추가로 논의되는 보존적 치환)을 고려한, 구조적으로 관련된 CCN 패밀리의 단백질(CCN 1 내지 CCN6)의 TSP-1 반복 상동성 도메인의 정렬 결과이다. 상이한 CCN 단백질의 TSP1 반복 상동성 도메인의 첫 번째 보존된 시스테인의 위치는 화학식 I의 재조합 단백질의 1번 위치로 정의된다.

- [0117] 보존된 시스테인 사이의 5개 세그먼트는 각각 A, B, C, D 및 E이다.
- [0118] 첫 번째 세그먼트 A는 식 A1-A2-A3-A5-A6-A7-A8-A9에 의해 정의되며, 여기서 A1 내지 A9는 위에서 정의한 바와 같다. 세그먼트 A의 7번 위치(A7)에 있는 아미노산은 CCN 패밀리 단백질의 모든 구성원에서 트립토판(W)이며 보존되는 것으로 여겨진다.
- [0119] 제2 세그먼트 B는 화학식 B1-B2-B3에 의해 정의되며, 여기서 B1 내지 B3은 위에서 정의한 바와 같다. 일 구현예에 따르면, B1 및 B3은 세린 또는 트레오닌이다.
- [0120] 세 번째 세그먼트 C는 식 C1-C2-C3-C4-C5-C6-C7-C8-C9-C10-C11-C12-C13-C14에 의해 정의되며, 여기서 아미노산 C1 내지 C14는 위에서 정의한 바와 같다. 일 구현예에 따르면, 아미노산 C1 및 C3은 글리신(G)이다. 또 다른 구현예에 따르면, C7은 아르기닌(R)이고, C10 및 C12는 둘 다 아스파라긴(N)이다.
- [0121] 네 번째 세그먼트 D는 식 D1-D2-D3-D4-D5-D6-D7-D8에 의해 정의되며, 여기서 아미노산 D1 내지 D8은 위에서 정의한 바와 같다. 일 구현예에 따르면, D7은 아르기닌(R)이다.
- [0122] 다섯 번째 세그먼트 E는 식 E1-E2-E3-E4에 의해 정의되며, 여기서 아미노산 E1 내지 E4는 위에서 정의한 바와 같다. 본 발명의 일 구현예에 따르면, E4는 프롤린이다.
- [0123] 마지막 시스테인 다음에는 0 내지 13개의 아미노산을 포함하는 가변 길이(F)의 카르복실 말단 펩티드 단편이 있다.
- [0124] F는 CCN 패밀리 단백질의 TSP-1 반복 상동성 도메인의 아미노산 서열과 비교하여 결실되거나 단축될 수 있다. 일 구현예에 따르면, F는 부재한다. 또 다른 구현예에 따르면, F는 PPSRGRSPQNSAF, GQPVYSSL, EADLEEN, EQEPEQPTD, DVDIHTLI, 및 DSNILKTIKIP로 구성된 군으로부터 선택되는 펩티드로 구성된다. 이 구현예의 일 양태에 따르면, 재조합 단백질은 서열번호 8 내지 12에 도시된 바와 같은 아미노산 서열의 형태를 취할 수 있다.
- [0125] 또 다른 양태에 따르면, 본 발명은 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 및 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 서열번호 38로 구성된 군으로부터 선택되는 아미노산 서열; 및 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 및 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 서열번호 38의 아미노산 서열과 적어도 50%의 서열 동일성을 갖는 이의 단편 또는 변이체를 포함하는 재조합 단백질을 제공한다.
- [0126] 일 양태에 따르면, 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 및 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 서열번호 38로 구성된 군으로부터 선택되는 아미노산 서열; 및 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 및 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 서열번호 38의 아미노산 서열과 50% 초과와 서열 동일성을 갖는 이의 단편 또는 변이체로 구성된 재조합 단백질이 제공된다.
- [0127] 본원에서 사용되는 "재조합 단백질"은 재조합 핵산에 의해 암호화된 단백질이다. 이들은 아래에서 추가로 개시되는 바와 같이 숙주 세포의 재조합 핵산으로부터 발현된다.
- [0128] "재조합 핵산"은 본원에서 기원 또는 조작으로 인해, 자연에서 회합되는 폴리뉴클레오타이드의 전부 또는 일부와 회합되지 않고/않거나 아래에서 추가로 개시되는 바와 같이 자연에서 연결되는 것이 아닌 폴리뉴클레오타이드에 연결되는 핵산 분자를 기술하는 데 사용된다.
- [0129] 당업자는 본 발명에 따른 재조합 단백질 및 융합 단백질의 아미노산 서열의 변형이 상기 단백질의 활성을 변경하지 않고 도입될 수 있음을 인지할 것이다. 아미노산은 일반적으로 소수성 또는 친수성 및/또는 극성 또는 비극성 측쇄를 갖는 것으로 분류된다. 동일한 생화학적 특성을 갖는 또 다른 아미노산에 대한 한 아미노산의 치환은 일반적으로 보존적 치환으로 알려져 있다.
- [0130] 아미노산의 보존적 치환에는 다음의 군 내의 아미노산 간에 이루어진 치환이 포함된다:
- [0131] · MILV
- [0132] · FYW
- [0133] · KRH

- [0134] · AG
- [0135] · ST
- [0136] · QN
- [0137] · ED
- [0138] 일반적으로, 보존적 아미노산 치환은 아미노산 치환이 이루어진 단백질의 상대적인 전하나 크기 특성을 변화시키지 않는 아미노산 치환을 지칭하며, 따라서 단백질의 구조가 거의 변하지 않기 때문에 생물학적 활성은 크게 변경되지 않는다.
- [0139] 당업자는 구조적 및 화학적-물리적 특성이 보존되는 한 하나 또는 몇 개의 아미노산이 결실, 아미노산 서열에 삽입 또는 추가되는 경우에 단백질의 생물학적 활성이 또한 유지될 수 있음을 인지할 것이다.
- [0140] 본원에서 아미노산 앞에 사용될 때 기호 "Δ"는 표시된 아미노산의 결실을 지칭하며, 예를 들어 ΔK447은 K447이 존재하지 않는 단백질로 이해되어야 한다. 특정 아미노산의 결실은 또한 본원에서 대안적으로 "-"로 표시되며, 예를 들어 K447-는 또한 K447이 존재하지 않는 단백질로서 이해되어야 한다.
- [0141] 따라서, 본 발명은 첨부된 청구범위에 개시된 바와 같은 재조합 단백질 및 융합 단백질을 포함하며, 여기서 전술한 바와 같은 이러한 변형(아미노산의 치환, 결실, 삽입 및 추가)은 생물학적 활성, 즉, 4개 도메인의 CCN-패밀리 단백질 CCN1, CCN2, CCN3, CCN4 및 CCN6에 기인하는 세포 신호전달 및 세포 생리 기능을 억제하거나 중화하는 능력을 본질적으로 변경하지 않고 도입될 수 있음을 이해해야 한다.
- [0142] 본 명세서 전반에 걸쳐 아미노산 서열이 언급된다. 본원에서 아미노산 서열을 지칭할 때, 때때로 본 명세서에서 "uniprot 넘버링" 또는 Eu 넘버링을 참조하여 해당 아미노산 서열 또는 단백질의 변형을 언급한다. uniprot 넘버링은 uniprot 데이터베이스(UniProt Consortium, *Nucleic Acids Res.* 2019 Jan 8;47(D1):D506-D515)에서 사용되는 넘버링을 지칭한다. uniprot 넘버링은 CCN 단백질의 아미노산 넘버링을 지칭할 때 사용된다. Eu 넘버링은 Eu 항체의 넘버링을 나타내며(Edelman et al., 1969, *Proc Natl Acad Sci USA* 63:78-85), 야생형과 상이한 돌연변이 또는 키메라가 있거나 없는 인간 IgG 클래스의 Fc 단편의 아미노산을 지칭할 때 사용된다. Eu 넘버링 시스템은 예를 들어 IMGT(international ImmunoGeneTics information system) 과학 차트의 IMGT로부터 접근할 수 있다. IMGT는 문헌[Lefranc M-P, *Biomolecules.* 2014 Dec; 4(4): 1102-1139]에 기술되어 있다.
- [0143] 본원에서 사용되는 바와 같이, "서열 동일성"을 지칭할 때, 제2 서열과 적어도 x%의 동일성을 갖는 서열은, 두 서열이 전체 정렬을 통해 최적으로 정렬된 경우 제2 아미노산 서열의 총 길이에 비해 x%가 제2 서열 중 매치된 아미노산과 동일한 제1 서열의 아미노산의 수를 나타낸다는 것을 의미한다. 두 서열은 x가 최대일 때 최적으로 정렬된다. 동일성 백분율의 정렬 및 결정은 수동 또는 자동으로 수행될 수 있다.
- [0144] 아미노산 서열 동일성 백분율을 결정하기 위한 정렬은 예를 들어, 공개적으로 이용 가능한 컴퓨터 소프트웨어, 예컨대, ClustalOmega(Sievers F, Higgins DG (2018) *Protein Sci* 27:135-145), Protein BLAST(National Center for Biotechnology Information(NCBI), 미국) 또는 Megalign(DNASTAR) 소프트웨어와 같은 상업적으로 이용 가능한 소프트웨어를 이용하여, 당해 분야의 기술 내의 다양한 방식으로 달성될 수 있다. 당업자는 비교되는 서열의 전체 길이에 걸쳐 최대 정렬을 달성하는 데 필요한 임의의 알고리즘을 포함하여, 정렬을 측정하기 위한 적절한 매개변수를 결정할 수 있다. NCBI BLAST는 아미노산 서열 동일성을 결정하는 데 사용되는 소프트웨어의 또 다른 예이다(MacWilliam et al., *Nucleic Acids Res.* 2013 Jul; 41(Web Server issue): W597-W600).
- [0145] 본 발명의 일 양태에 따르면, 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 및 서열번호 38로부터 선택되는 아미노산 서열; 및 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 및 서열번호 38의 아미노산 서열과 적어도 50%의 서열 동일성을 갖는 이의 단편 또는 변이체를 포함하는 재조합 단백질이 제공된다.
- [0146] 또 다른 양태에 따르면, 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 및 서열번호 38로 구성된 군으로부터 선택되는 아미노산 서열과 적어도 60%, 70%, 80%, 90%, 또는 95%의 서열 동일성을 갖는 아미노산 서열을 포함하는 재조합 단백질이 제공된다.
- [0147] 생물학적으로 활성인 단백질 및 펩티드는 인간 질환의 임상 관리에서 중요한 위치를 차지한다. 그러나 많은 단

백질 및 펩티드는 작은 크기 및/또는 단백질 분해 대사로 인해 신장 여과에 의해 제거되기에 이상적인 약동학적 특성이 부족하여 문제가 된다. 이러한 인자는 단백질 또는 펩티드의 순환 농도를 효과적인 치료 수준으로 유지하기 위해 일정한 주입 또는 빈번한 피하 투여를 투여해야 하는 것과 같이 치료를 필요로 하는 대상체에게 약물을 투여할 때 제한 또는 과제를 부과할 수 있다. 약물의 일정하거나 매우 빈번한 투여의 필요성은 환자와 의사 모두에게 명백한 도전과 불편으로 인해 임상적으로 바람직하지 않다.

- [0148] 생물학적으로 활성인 펩티드 또는 단백질의 반감기를 연장하기 위한 한 가지 전략은 폐길화(PEGylation)라는 과정에 의해 관심 펩티드 또는 단백질에 폴리에틸렌 글리콜(PEG) 기를 연결하는 것이다(예를 들어, 문헌[Dozie et al. (2015), *Int. J. Mol Sci*, 16(10) 25831 - 25864] 참조). 단백질 폐길화의 일반적인 전략은 단백질의 작용기와 PEG 분자의 상보적인 기를 반응시켜 단백질-PEG 접합체를 형성하는 것이다. PEG 모이어티는 유연성, 친수성, 다양한 크기 및 낮은 독성으로 인해 단백질의 안정성 및 순환 반감기를 증가시키는 몇 가지 장점을 제공한다.
- [0149] 따라서, 일 구현예에서 본 발명은 재조합 단백질이 폐길화되어 있는 전술한 바와 같은 재조합 단백질을 제공한다. 본 발명에 따른 융합 단백질은 또한 폐길화될 수 있다.
- [0150] 융합 단백질
- [0151] 펩티드 및 단백질의 의학적 용도와 관련된 문제를 제거하는 또 다른 방법은 융합 단백질을 제조하여 생체 활성 단백질 또는 펩티드의 반감기를 연장하는 것이다(예를 들어, 문헌[Valeria et al. (2017), "A New Approach to Drug Therapy: Fc-Fusion Technology", *Prim Health Care*, 7:255, doi:10.4172/2167- 1079.1000255] 참고). 단백질 또는 펩티드를 유전자 재조합을 통해 전달체 단백질에 공유결합으로 융합함으로써 관심 단백질의 분자량을 신장 여과의 역치인 약 60 내지 70kDa으로 증가시킬 수 있다.
- [0152] 본 발명은 융합 단백질로서,
- [0153] (i) CCN 패밀리 단백질의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인;
- [0154] (ii) (i)의 TSP-1 반복 상동성 도메인에 N 또는 C 말단에 융합된 융합 파트너로서, 상기 융합 파트너는 혈청 알부민, 트랜스페린 및 면역글로불린 Fc 단편으로 구성된 군으로부터 선택되는 것인 융합 파트너;
- [0155] (iii) 선택적으로, (i)의 TSP-1 반복 상동성 도메인 및 Fc 단편(TSP-1 반복 상동성 도메인에 N 또는 C 말단에 융합됨) 사이의 펩티드 링커;
- [0156] 를 포함하는 융합 단백질을 제공한다.
- [0157] 본 명세서 전반에 걸쳐, TSP-1 반복 상동성 도메인은 또한 CCN 패밀리 단백질의 제3 도메인을 지칭하는 도메인 III으로 언급될 수 있고, 지칭할 수 있다.
- [0158] 바람직한 일 양태에서, 융합 파트너는 단량체 융합 파트너이고, 결과적으로 단량체성인 융합 단백질을 생성한다. 이러한 융합 단백질, 특히 이의 TSP-1 도메인은 위에서 정의되고 아래에서 추가로 설명된다.
- [0159] 그러나 본 개시 내용은 또한 TSP-1 도메인 단백질 성분 및 융합 파트너 성분 모두와 관련하여, 다른 구현예를 포함한다.
- [0160] 이러한 일 구현예에 따르면, TSP-1 반복 상동성 도메인은 위에서 정의된 바와 같은 화학식 I의 재조합 단백질이다.
- [0161] TSP-1-반복-상동성-도메인은 또 다른 구현예에 따라 서열번호 1 내지 12, 37 및 38에 기재된 서열 중 어느 하나에 정의된 바와 같은 아미노산 서열을 갖는 재조합 단백질, 또는 위에서 정의된 바와 같은 화학식 I의 재조합 단백질이다.
- [0162] 일 구현예에 따르면, TSP-1 반복 상동성 도메인은 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 및 서열번호 38로 구성된 군으로부터 선택되는 아미노산 서열과 적어도 50%, 60%, 70%, 80%, 90%, 또는 95%의 서열 동일성을 갖는 아미노산 서열을 포함하는 재조합 단백질이다.
- [0163] 단백질은 본질적으로 단백질 분해효소 분해에 민감하다. 본 발명에 따른 재조합 단백질 및 융합 단백질의 단백질 분해효소 분해를 방지하기 위하여, 아미노산 서열에 대한 변형이 단백질 분해효소 저항성 재조합 단백질 및 융합 단백질을 제공하기 위하여, 예를 들어, 부위 지정 돌연변이 유발에 의해 도입될 수 있다. 예를 들어, 점

돌연변이는 융합 단백질 또는 제조합 단백질의, 서열번호 1 내지 12, 37 또는 38에 정의된 바와 같은 CCN 패밀리 단백질, 또는 더욱 특별하게는 서열번호 1 내지 6, 8 내지 12 또는 37 중 어느 하나에 정의된 바와 같은 단백질의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인에 도입될 수 있다. 일 구현예에 따르면, 점 돌연변이가 도입되어 단백질 분해의 감수성을 감소시킨다. 본 발명의 제조합 단백질 및 융합 단백질의 더 적은 단백질 분해를 초래하는 점 돌연변이의 비 제한적인 예는 서열번호 7에 제시된 바와 같이, CCN5의 도메인 III의 195번 위치에서 프롤린의 알라닌으로의 대체(P195A)에 상응하는 점 돌연변이를 도입하는 것이다. 다른 CCN 패밀리 구성원의 도메인 III에서 유래한 아미노산 서열에도 유사한 돌연변이가 도입될 수 있다. 서열번호 38은 Ala 치환을 포함하는 CCN5의 TSP-1 도메인의 절단된 44개 아미노산의 서열을 나타낸다. 서열번호 42 내지 46은 각각 Ala 치환을 포함하는 CCN1, 2, 3, 4 및 6의 TSP-1 상동성 도메인의 절단된 44개 아미노산의 서열을 나타낸다. 서열번호 47 내지 51은 각각 Ala 치환을 포함하는 CCN1, 2, 3, 4 및 6의 더 긴 TSP-1 상동성 서열을 나타낸다. 임의의 이러한 서열, 또는 이와 적어도 80%의 서열 동일성을 갖는 서열이 본 발명에 따라 사용될 수 있다.

[0164] 위에서 언급된 바와 같이, 바람직한 구현예에서, 본 발명에 따른 융합 단백질의 융합 파트너 (ii)는 단량체성이다. 임의의 단량체 융합 파트너가 사용될 수 있다. 따라서 융합 파트너는 TSP-1 상동성 도메인 단백질 성분에 융합될 때 단량체 형태로 발생하고 남아 있는 임의의 단백질 또는 이의 일부(예를 들어, 단백질 도메인)일 수 있다. 따라서, 단량체 융합 파트너 및 TSP-1 상동 도메인 단백질을 포함하는 융합 단백질은 단량체로 남게 된다. 즉, 자체적으로 이량체화하거나 더 높은 다량체를 형성하지 않는다.

[0165] 다양한 단백질이 가능한 융합 파트너로 알려져 있으며, 천연 단백질, 또는 이의 단편 또는 아미노산 서열 변형 변이체뿐만 아니라, 합성 단백질 또는 아미노산 호모폴리머를 포함할 수 있다. 이러한 단백질은 특히 IgG, 혈청 알부민 또는 트랜스페린의 Fc 단편을 포함한다.

[0166] 융합 파트너는 자연에서 제1 CCN TSP-1 상동성 폴리펩타이드와 조합하여(예를 들어, 직접적으로 또는 간접적으로 인접하여 또는 연결되어) 존재하지 않고, 합성 또는 인공 조합에서 제1 CCN TSP-1 상동성 폴리펩타이드에 연결되는 제2 폴리펩타이드(또는 제2 아미노산 서열)로서 본원에서 광범위하게 정의된다. 따라서, 융합 단백질은 함께 연결되거나 융합된 적어도 2개의 아미노산 서열 또는 폴리펩타이드의 비 천연 조합을 포함한다.

[0167] 융합 파트너는 적어도 6, 8, 9, 10, 15, 20, 25, 30, 40 또는 50개 이상의 아미노산 길이인 아미노산 서열일 수 있다. 일반적으로, 융합 파트너는 기능적 폴리펩타이드, 또는 환언하면, 예를 들어, 융합 단백질을 안정화하는(제1 폴리펩타이드를 더욱 안정적으로 만드는), 또는 이의 혈청 반감기를 증가시키는, 기능 또는 특성을 융합 단백질에 부여하는 폴리펩타이드이다. 따라서, 융합 파트너는 구조적 단백질이거나 구조적 기능을 가질 수 있거나, 또는 융합 단백질에 활성 또는 특성, 예를 들어 결합 활성을 부여할 수 있다(예를 들어, 융합 파트너는 결합 쌍의 구성원일 수 있거나, 또는 친화도 결합 파트너 동일 수 있다). 대표적인 예에서 융합 파트너는 알부민(특히 혈청 알부민), 피브리노겐, 글루타티온 S-트랜스퍼라제, 트랜스페린, 스트렙타비딘 또는 스트렙타비딘 유사 단백질, 또는 면역글로불린, 또는 이의 일부, 특히 면역글로불린의 Fc 부분(예를 들어, IgG1, IgG2, IgG3 또는 IgG4), 또는 이의 일부 또는 변형일 수 있다. 적합한 혈청 알부민은 소 혈청 알부민(BSA), 마우스 혈청 알부민(MSA) 및 특히 인간 혈청 알부민(HSA)을 포함한다. 다른 가능한 융합 파트너는 융합 단백질의 약동학적 특성을 개선하는 작용을 할 수 있는 폴리펩타이드, 예를 들어 합성 폴리펩타이드, 예컨대, 호모 아미노산 중합체, 프롤린-알라닌-세린 중합체, 또는 예를 들어, 문헌[Strohl, 2015, BioDrugs 29, 215-239]에 기술된 바와 같은 엘라스틴 유사 펩티드를 포함한다. 치료 단백질과 함께 사용하기 위해 당해 분야에 공지된 임의의 융합 파트너가 사용될 수 있다.

[0168] 구현예에서, 본 발명에 따른 융합 단백질의 융합 파트너 (ii)는 IgG의 Fc 단편(임의의 서브클래스 또는 임의의 서브클래스의 키메라), 혈청 알부민 또는 트랜스페린일 수 있다.

[0169] 융합 파트너는 융합 단백질의 TSP-1 상동성 도메인 단백질 성분에, 예를 들어, 본원에 정의된 바와 같은 CCN5 또는 임의의 다른 CCN 단백질의 TSP-1 반복 상동성 도메인에, N 또는 C 말단에 연결될 수 있다. 이는 아래에서 추가로 설명되는 바와 같이 직접적으로 또는 링커를 통해 간접적으로 연결될 수 있다.

[0170] Fc 단편은 이량체를 형성하는 경향이 있으며, 융합 단백질에 사용되는 경우, 융합 단백질 구축물은 융합 단백질의 2개 카피를 포함하는 경향이 있다. 그러나, 단량체 Fc 단편 및 이를 포함하는 단량체 융합 단백질이 제조될 수 있음은 당해 분야에 공지되어 있다.

[0171] 따라서, 융합 파트너가 Fc 단편인 경우, 이는 바람직하게는 단량체 Fc 단편, 예컨대, 임의의 클래스의 인간 IgG의 단량체 Fc 단편이다. 상이한 클래스의 Fc 영역의 일부를 포함하는 키메라 Fc 단편이 포함되며, 변형된 서열

을 갖는 Fc 단편도 포함된다.

- [0172] Fc 융합 단백질은 IgG 이소형의 Fc 단편에 연결된 이펙터 도메인으로 구성된 키메라 단백질을 기초로 하는 성장하는 부류의 단백질 치료제이다. 바이오의약품의 대표적인 예는 예를 들어 류마티스 관절염의 치료에 사용되는 에타네르셉트(Fc 단편에 연결된 TNF- α 수용체)이다. Fc 융합 생물약제학적 단백질의 또 다른 예는 애플리버셉트이다. 습성 황반변성 및 전이성 결장직장암 치료에 사용되는 VEGF 수용체-Fc-융합 단백질인 애플리버셉트. Fc 단편 융합 단백질을 만드는 주요 근거는 신장 배설을 배제하고 신생아 Fc 수용체를 통한 신장 근위 세뇨관 재흡수를 향상시키기에 충분한 분자 질량의 증가로 인한 반감기의 연장을 얻는 것이다. 또한, 내피 세포 상의 신생아 Fc 수용체(FcRn)에 대한 Fc 융합 단백질의 pH 의존적 결합은, 다르게는 세포내 이입 및 후속적인 리소좀 분해로 예정되어 있는, Fc 기반 융합 단백질이 재순환되어 순환으로 다시 방출되도록 한다.
- [0173] 일 구현예에 따르면, 융합 파트너 (ii)가 인간 IgG의 모든 서브클래스를 포함하는 인간 IgG(면역글로불린 G, 면역글로불린 γ 로도 알려짐)로부터의 Fc-단편인 융합 단백질이 제공된다. 본 발명의 또 다른 구현예에 따르면, 융합 파트너가 IgG1, IgG2 또는 IgG4의 Fc 단편인 융합 단백질이 제공된다. 바람직하게는, 인간 IgG의 Fc 단편은 서브클래스 IgG4(서열번호 13) 또는 IgG2(서열번호 14)의 Fc 단편이다.
- [0174] IgG1, IgG2 및 IgG4는 약 3주의 더 긴 반감기로 인해 종종 IgG3보다 선호된다. 당업자는 Fc 융합 파트너로서 특정 서브클래스의 IgG 이소형의 선택이 최종 화합물에 대한 원하는 반감기 연장 및 세포독성 활성 수준에 따라 다를 수 있음을 인지할 것이다. 암 또는 자가면역 질환의 치료를 위해 표시되는 치료 항체는 Fc 수용체에 대한 높은 친화도 및 면역 이펙터 기능을 발휘하는 강력한 능력 때문에 대부분의 경우 IgG1 서브클래스에 속한다. 반면에, IgG2 및 IgG4는 면역 이펙터 기능이 역효과를 일으킬 수 있기 때문에 면역 이펙터 기능의 결여가 요구될 경우 치료 후보의 백본으로 사용하기 위한 바람직한 IgG 서브클래스이다. 면역 이펙터 기능을 활성화하는 Fc 단편의 경향은 Ig 이소형 및 서브클래스에 따라 다르고, 상이한 면역 이펙터 기능에 대해 변한다. 적합한 IgG 서브클래스의 Fc 단편을 선택하는 것 외에도, IgG 서브클래스의 아미노산 서열은 면역 이펙터 기능을 활성화하는 Fc 단편의 능력을 감소시키기 위하여, 예를 들어 부위 지정 돌연변이 유발에 의해 변형될 수 있다.
- [0175] Fc 단편은 단량체를 형성하는 것, 또는 보다 정확하게는 단량체 형태를 유지하거나 갖는 것이 선택될 수 있거나, 또는 단량체 구조를 허용하거나 촉진하는 돌연변이를 도입하도록 변형될 수 있다. 이러한 돌연변이는 본원에서 "단량체 생성 돌연변이"로 지칭된다. 단량체 생성 돌연변이를 포함하는 Fc 단편의 예는 서열번호 54 및 55이다. 당업자는 이러한 돌연변이를 도입하고 단량체 Fc 돌연변이체를 선택하는 방법을 알고 있다.
- [0176] Fc 단편의 면역 이펙터 활성화 기능의 회피
- [0177] 예를 들어, 제2형 당뇨병의 주 1회 치료에 사용되는 GLP-1 효능제-Fc-단편 융합 단백질인 생물약제 융합 단백질 둘라글루타이드(Trulicity™)에서 잘 특성화된 돌연변이 F234A 및 L235A는 IgG4 Fc 단편의 힌지 영역에 도입되어 면역 이펙터 기능을 활성화하는 능력을 감소시킨다.
- [0178] 본 발명의 일 구현예에 따르면, Fc 융합 파트너는 IgG4의 Fc 단편으로, 여기서 IgG4 Fc 단편은 예를 들어, 위의 F234A 및 L235A 돌연변이를 포함하는, 면역 이펙터 기능을 피하도록 변형된다.
- [0179] 단백질 분해효소 저항성 Fc 단편
- [0180] 제조 수율과 생물학적 반감기를 모두 감소시킬 수 있는 또 다른 인자는 융합 단백질의 엔도펩티다제 절단이다. 단백질 분해의 위험을 감소시키거나 제거하기 위하여, 특히 단백질 분해 절단에 민감한 부위에 돌연변이를 도입함으로써 Fc 단편의 아미노산 서열에 변형을 도입할 수 있다. EP 특허 출원 EP2654780B1에서, IgG1 불변 영역의 Fc 도메인은 생성된 변형된 Fc 함유 단백질이 단백질 분해에 저항성을 갖도록 하기 위하여 E233-L234-L235-G236을 P233-V234-A235(G236 결실)(EU 넘버링)로 대체함으로써 변형되었다.
- [0181] CCN5의 도메인 III에 연결된 IgG4의 Fc 단편에서 EP2654780B1에 개시된 아미노산 변형의 통합은 엔도펩티다제에 대한 충분한 저항성을 제공하지 않는 것으로 여겨졌다. 그러나 개선했던 단백질 분해효소 저항성이 본 발명에 따른 융합 파트너로서 사용된 IgG 서브타입의 추가 변형에 의해 달성되었다.
- [0182] 더욱 구체적으로, IgG2의 전체 힌지 영역과 IgG4의 불변 중쇄 2 및 3을 포함하는 융합 단백질이 우수한 단백질 분해 저항성을 나타내는 것으로 밝혀졌다.
- [0183] 문헌[Mueller JP et al., *Mol. Immunol.* (1997), 34(6), pp. 441-452]에서, IgG 항체에서 IgG2/IgG4 키메라의 사용이 개시되어 있다. 발작성 야간 혈색소뇨증 및 비정형 용혈성 요독 증후군의 치료에 사용되는 에쿨리주맙이라는 또 다른 생물약제 단백질 항체는 예를 들어, Fc γ R 의존성 면역 이펙터를 활성화하는 IgG4의 능력을 회피

하는 데 유용한 것으로 나타났다. 나아가, 이러한 키메라 Fc 단편의 IgG4 불변 도메인 2 및 3은 또한 보체 의존성 면역 이펙터 기능을 활성화하는 IgG2의 능력을 회피하는 것으로 나타났다. 이와 관련하여 Rother 등의 보고서를 참조한다. 문헌[Nat. Biotechnol., 25(11), pp. 1256-1264] 및 [Mueller JP 등, 위의 문헌] 참조.

- [0184] 또한, 문헌[Borrok et al. (2017), *J. Pharm. Sci.* 106; 1008-1017]은 항체의 면역 이펙터 기능에 미치는 영향을 연구하기 위하여 Fc 단편에 변형을 도입하는 것(FQY-YTE 돌연변이)을 개시하고 있다. W02017158426A1에서는 항체의 반감기를 개선하기 위하여 Fc 단편에 돌연변이를 도입함으로써 항체를 변형시키는 것이 개시되어 있다. 특히, 면역글로불린의 Fc 영역에서 311, 434, 428, 438, 및 435번 위치 중 하나 이상에서의 변형이 개시되어 있다.
- [0185] 나아가, 문헌[Kinder et al. *J Biol Chem.* 2013 Oct 25;288(43):30843-54]은 단백질 분해효소 저항성 IgG1 항체를 초래한 IgG1의 하부 힌지의 돌연변이(즉, E233P, L234V, L235A, G236-, Eu 넘버링)에 대해 보고하고 있다.
- [0186] 일 구현예에 따르면, 본 발명에 따른 융합 단백질의 Fc 단편은 다음의 돌연변이: S228P, F234A, L235A, K447-(Eu 넘버링)을 포함하는 IgG4 서브클래스의 Fc 단편으로 구성된다. 서열번호 15 참조.
- [0187] 문헌[Jacobsen et al. *J Biol Chem.* 2017 Feb 3;292(5):1865-1875]은 Asn297의 돌연변이로 인해 Fc 단편이 비글리코실화되며, 이에 따라 IgG 이펙터 기능이 결여되었다고 보고하였다. Jacobsen은 또한 일부 변이체(N297G)가 다른 변이체(N297Q 또는 N297A)에 비해 더 나은 안정성과 개발성을 갖는 항체를 생성함을 발견하였다. 모체 IgG1보다 더 나은 안정성을 가져오는 추가 변형(이황화 가교)도 도입되었다.
- [0188] 본 발명에 따르면, 융합 파트너가 Fc 단편인 경우, 이는 예를 들어 서열번호 16에서와 같이 안정화 이황화 가교 없이 또는 이와 함께 비글리코실화될 수 있다.
- [0189] 본 발명자들의 지식으로는, IgG2의 전체 힌지 영역 및 IgG4의 불변 중쇄 2 및 3으로 구성된 Fc 단편은 이전에 상기 Fc 단편을 이펙터 단백질에 연결함으로써 융합 단백질을 제조하는 데 사용된 바 없었다.
- [0190] 일 구현예에 따르면, 본 융합 단백질의 융합 파트너는 비 글리코실화되고 이황화 가교에 의해 안정화된 IgG1의 Fc 단편이고, 여기에서 다음의 돌연변이를 갖는 하부 힌지가 도입되었다: E233P, L234V, L235A, G236-(Eu 넘버링)(서열번호 17).
- [0191] 일 구현예에 따르면, CCN 패밀리의 단백질의 TSP-1 반복 상동성 도메인을 포함하는 융합 단백질의 융합 파트너는 IgG4의 Fc 단편이고, 여기서 다음의 돌연변이가 하부 힌지에 도입되었다: S228P 및 K477- 이외에, E233P, L234V, L235A, G236- 돌연변이(Eu 넘버링)(서열번호 18).
- [0192] 바람직한 일 구현예에서 Fc-단편은 IgG2의 힌지 영역(216 ERKCCVECPAPPPVA-GP 238, Eu 넘버링) 및 임의의 다른 IgG 서브클래스의 키메라이다. 가장 바람직하게는 Fc 단편은 서열번호 19에 제시된 바와 같이, 카르복실 말단의 K477-(Eu 넘버링)이 결실된, IgG2의 힌지 영역 및 IgG4의 불변 중쇄 도메인 2 및 3의 키메라이다. 본 발명의 이 구현예는 개선된 단백질 분해효소 저항성 특성을 갖는 것으로 나타났다(실시예 6 참조).
- [0193] 일 구현예에서, 본 발명의 단량체 융합 단백질의 융합 파트너는 서열번호 54에 제공된 바와 같이, 이황화 가교 안정화되고(R292C, V302C), 비 글리코실화되고(N297G), 단량체 생성 돌연변이(C220Q, C226Q, C229Q, T366R, L368H, P395K, K409T, M428L), Eu 넘버링)가 있는 IgG1의 Fc 단편이다.
- [0194] 추가 구현예에서, 본 발명의 단량체 융합 단백질의 융합 파트너는 서열번호 55에 제공된 바와 같이, IgG2의 힌지 영역과 IgG4의 불변 중쇄 도메인 2 및 3의 키메라이고, 카르복실 말단 K477-이 결실되고, 단량체 생성 돌연변이(C219Q, C220Q, C226Q, C229Q, L351F, T366R, P395K, F405R, Y407E)가 있고, 반감기 연장 돌연변이(M252Y, S254T, T256E)(Eu 넘버링)가 있는 Fc 단편이다.
- [0195] 본 발명에 따른 융합 단백질은 IgG2의 전체 힌지 영역, IgG4의 불변 중쇄 2 및 3, 및 CCN 단백질 패밀리의 구성원의 도메인 III으로 구성된 Fc 단편을 사용하여 예시되지만, 유리한 단백질 분해효소 저항성은 또한 이러한 Fc 단편 키메라가 다른 이펙터 분자, 예를 들어 VEGFR, FGF-21 또는 GLP1에 연결되는 경우 달성된다고 믿어진다. 이펙터 분자는 원하는 약리학 특성을 제공하는 Fc 융합 단백질의 일부이며, Fc 단편은 약동학 특성에 기여한다.
- [0196] 융합 파트너로서의 혈청 알부민
- [0197] 펩티드와 단백질의 반감기를 연장하기 위한 대안적인 전략은 융합 파트너로서 혈청 알부민(SA)을 사용하는 것이다. IgG 및 SA는 대부분의 다른 순환 단백질의 경우 며칠 이하인 것과 비교하여 약 19일의 연장된 반감기를 공

유한다. SA는 또한 신생아 Fc 수용체(FcRn)에 대해 친화성을 가지며, 세포내 분해로부터 구조된다(문헌 [Andersen et al. (2014), *J Biol Chem*, 289(19); pp 13492- 13502] 참조).

- [0198] 본 발명의 일 구현예에서, 융합 단백질이 전술한 바와 같이 제공되며, 여기에서 융합 파트너는 혈청 알부민, 바람직하게는 인간 혈청 알부민이다.
- [0199] 일 구현예에서, 단량체 융합 단백질이 전술한 바와 같이 제공되며, 여기에서 융합 단백질은 서열번호 101에 제공된 바와 같이 인간 혈청 알부민의 아미노산 25 내지 606을 포함한다.
- [0200] 본 발명의 추가 구현예에서, 알부민, 예를 들어 인간 혈청 알부민은 증가된 또는 감소된 반감기를 초래하는 pH 의존성이 있거나 없이, FcRn 친화도를 변경함으로써 예를 들어, 반감기를 증가시키거나 감소시키도록 변형된다.
- [0201] 융합 파트너로서의 트랜스페린
- [0202] 그러나 펩티드와 단백질의 반감기를 연장하기 위한 대안적인 전략은 트랜스페린의 자연적으로 긴 반감기를 활용하여 융합 파트너로서 트랜스페린을 사용하는 것이다. (Strohl W. *BioDrugs*. 2015; 29(4): 215-239). 트랜스페린은 글리코실화 또는 비 글리코실화 형태로 사용될 수 있다.
- [0203] 본 발명의 일 구현예에서, 융합 단백질이 전술한 바와 같이 제공되며, 여기에서 융합 파트너는 트랜스페린, 바람직하게는 인간 트랜스페린이다.
- [0204] 일 구현예에서, 단량체 융합 단백질이 전술한 바와 같이 제공되며, 여기에서 융합 단백질은 서열번호 53에 제공된 바와 같이 인간 트랜스페린의 아미노산 20 내지 698을 포함한다.
- [0205] 링커
- [0206] 또 다른 구현예에 따르면, 본 발명에 따른 융합 단백질은 융합 파트너와 이펙터 분자 사이에 펩티드 링커를 선택적으로 포함할 수 있다. 즉, 링커는 CCN 단백질의 TSP-1 반복 상동성 도메인에 N 또는 C 말단에 융합된다(TSP-1 도메인 단백질/폴리펩티드).
- [0207] 임의의 펩티드 링커가 사용될 수 있으며(CCN 단백질 서열이 아닌 한), 이들 중 다수가 당해 분야에 공지되고 기술되어 있다. 링커는 가요성 링커 서열일 수 있다(이는 가요성 링커 서열 모티프의 반복을 포함할 수 있다). 당해 분야에 공지된 전형적인 링커는 작은 비 극성(예를 들어, 글리신) 또는 극성(예를 들어, 세린 또는 트레오닌) 잔기가 풍부하고, 일반적으로 글리신 및 세린 잔기(GS)의 스트레치 또는 기타 아미노산 잔기, 예컨대, 알라닌, 리신 및/또는 글루탐산(A, K, 및/또는 E), 또는 실제로 임의의 아미노산으로 구성된다. 일반적으로 사용되는 링커는 (GGGGS) 링커(서열번호 121)이며, 이는 링커의 반복 단위((GGGGS)_n, 여기서 n의 카피 수는 조정될 수 있음, 예를 들어, 1 내지 10, 1 내지 6, 1 내지 4 등)로서 제공될 수 있다. 링커는 1 내지 50, 1 내지 45, 1 내지 40, 1 내지 30, 1 내지 25, 1 내지 20, 1 내지 15, 1 내지 12, 1 내지 10, 예를 들어, 1 내지 8, 1 내지 6, 1 내지 5, 또는 1 내지 4개 아미노산 길이일 수 있다. 다양한 상이한 링커가 하기 실시예에서 기술되고 사용되며, 이들 중 임의의 것이 본 발명의 임의의 융합 단백질에서 사용될 수 있다.
- [0208] 일부 구현예에서, 링커는 50개 이하의 아미노산을 포함한다.
- [0209] 펩티드 링커의 특성은 이펙터 기능의 유지를 추가로 개선할 수 있다. 그러나 펩티드 링커는 엔도펩티다제 절단 및 융합 단백질의 제거에 민감할 수 있다. 산재된 세린 잔기가 있거나 없는 글리신이 있는 펩티드 링커가 일반적으로 사용되지만, 이 디자인이 항상 원하는 활성 및 엔도펩티다제에 대한 저항성을 갖는 융합 단백질을 생성하는 것은 아니다. 뉴로트로핀 결합 단백질-Fc-융합 단백질을 개시하는 US20180273603에서는, 서열 A(EAAAAK)A(그 안의 서열번호 14)의 반복을 포함하는 α-나선형 링커가 제안된다. 또한, US2018/0127478은 Fc 융합 단백질에서 제안된 서열 EAAAAK의 1 내지 3개 반복으로 구성된 아미노산 링커의 사용을 개시하고 있다.
- [0210] 본 발명에 따르면, 펩티드 서열 EAAAAK(본원의 서열번호 21)로 구성된 링커가 또한 TSP-1 상동 도메인과 융합 파트너(Fc 단편) 사이에 포함될 수 있다. 더욱 바람직하게는, 링커는 아미노산 서열 EAAAAK의 반복으로 구성된다.
- [0211] 링커가 본 발명의 융합 단백질에 포함되는 경우, 링커는 융합 파트너와 이펙터 분자, 즉 CCN 단백질의 도메인 III 사이에 위치한다. 링커는 CCN 단백질의 도메인 III의 C 말단 또는 N 말단에 도입될 수 있다.
- [0212] 또한, 나선형 링커는 현탁액 CHO 세포에서 재조합 단백질의 발현 후 엔도펩티다제 절단에 대해 저항성이 있었다. 이는 제조 목적과 생체 내 효능 모두에 중요하다. 또한, Fc 융합 단백질에서 Fc 단편과 이펙터 도메인 사이의 α-나선형 링커의 도입은 Fc 융합 단백질의 응집 경향을 감소시키는 것으로 나타났다.

- [0213] 이러한 결과는 CCN 계열 단백질의 도메인 III을 이펙터 단백질로서 포함하는 융합 단백질에서 나타나지만, 다른 이펙터 분자를 Fc 단편 본 발명에 따른 α -나선형 링커에 결합하는 경우에도 유리한 응집 감소 경향 및 단백질 분해효소 저항성 효과가 얻어지는 것으로 여겨진다.
- [0214] 따라서, 본 발명은 Fc 단편과 이펙터 분자 사이에 화학식 aa1-aa2-(EAAAK)_n-aa3-aa4-aa5(n 은 4 이상임)의 펩티드 링커 서열을 갖는 Fc 단편을 포함하는 Fc 융합 단백질을 제공하며, 여기서 aa1, aa2, aa3, aa4, aa5는 독립적으로 부재하거나 아미노산이다. 링커는 Fc 단편의 N 말단 또는 C 말단에 위치할 수 있다. 일 구현예에 따르면, n 은 8이다. 또 다른 구현예에 따르면, aa1은 트레오닌(T)이고, aa1, aa2, aa3, aa4 및 aa5는 Ala(A)이다. 일 구현예에 따르면, 위의 Fc 융합 단백질의 링커는 서열번호 22, 서열번호 23, 서열번호 24, 및 서열번호 25로 구성된 군으로부터 선택된다. 특히, (EAAAK)-반복으로 구성된 링커, 즉 (EAAAK)_n(여기서 n 은 8)로 구성된 링커를 포함하는 본 발명에 따른 융합 단백질의 사용은 유리하게도 더 적은 응집을 초래하는 것으로 나타났다.
- [0215] 본 발명에 따라 사용될 수 있는 대안적인 링커는 서열번호 20(TEGRMD)에 제시된 바와 같은 아미노산 서열을 갖는 링커이다.
- [0216] 따라서, 일 구현예에서, 본 발명은 융합 파트너와 CCN5 상의 도메인 III(예를 들어, 서열번호 1 내지 서열번호 12, 37 또는 38) 사이에 링커 펩티드의 도입을 포함할 수 있다. 서열번호 20의 링커를 포함하는 융합 단백질의 비 제한적인 예는 각각 서열번호 28, 서열번호 29 및 서열번호 30에 제시되어 있다.
- [0217] 본 발명이 (서열번호 20에서와 같이) 펩티드 링커의 N 말단에 유전적으로 융합된 CCN5의 도메인 III 및 (서열번호 15에서와 같이) 다음의 돌연변이(S228P, F234A, L235A, K447-, Eu-넘버링)를 포함하는 IgG 서브타입 IgG4의 Fc 단편을 취하는 경우, 완전한 서열은 CCN5(dIII)-Fc_{v2}로도 명명되는 서열번호 28에서와 같을 것이다.
- [0218] 본 발명이 (서열번호 20에서와 같이) 펩티드 링커의 N 말단에 유전적으로 융합된 CCN5의 도메인 III 및 서열번호 18에 제시된 바와 같이 다음의 돌연변이 (S228P, E233P, F234V, L235A, G236-, K447-, Eu-넘버링)를 포함하는 IgG 서브타입 IgG4의 Fc 단편을 취하는 경우, 생성된 서열은 CCN5(dIII)-Fc_{v2.1}로도 명명되는 서열번호 29에 제시된 바와 같을 것이다.
- [0219] 본 발명이 (서열번호 20에서와 같이) 펩티드 링커의 N 말단에 유전적으로 융합된 CCN5의 도메인 III 및 서열번호 19에 제시된 바와 같이, IgG 서브타입 IgG2/4 서브타입의 키메라 Fc 단편을 취하는 경우, 생성된 서열은 CCN5(dIII)-Fc_{v2.3}으로도 명명되는 서열번호 30에 제시된 바와 같을 것이다.
- [0220] 본 발명이 (서열번호 25에서와 같이) 펩티드 링커의 N 말단에 유전적으로 융합된 (서열번호 1에서와 같이) CCN5의 도메인 III 및 서열번호 19에 제시된 바와 같이 IgG 서브타입 IgG2/4 서브타입의 키메라 Fc 단편을 취하는 경우, 생성된 서열은 CCN5(dIII)-HLn8-Fc_{v2.3}으로도 명명되는 서열번호 31에 제시된 바와 같을 것이다.
- [0221] 본 발명의 바람직한 구현예에 따르면, 본 발명에 따른 융합 단백질로서,
- [0222] 1) CCN 패밀리 단백질, 특히 CCN5의 도메인 III으로의 점 돌연변이(서열번호 7 참조, 상기 도메인 III의 감소된 단백질 분해 감수성을 초래함);
- [0223] 2) 인간 IgG4 및 인간 IgG2의 Fc 단편의 조작된 키메라(서열번호 19, 이는 Fc 융합 단백질에 사용되는 이전에 기술된 Fc 단편 백본에 비해 단백질 분해 감수성을 감소시킴);를 포함하고,
- [0224] 3) 펩티드 링커의 최적화된 구성(서열번호 21 내지 25 참조)(이는 단백질 분해 감수성을 감소시키고, 융합 단백질의 생물학적 활성을 증진시키며, 융합 단백질의 응집 경향을 감소시킴)을 포함하는 융합 단백질이 제공된다.
- [0225] 일부 구현예에서, (i)의 아미노산 서열과 단량체성 융합 파트너 사이의 펩티드 링커는 서열번호 20 내지 25 또는 39로 구성된 군으로부터 선택되는 아미노산 서열, 또는 이와 80%의 서열 동일성을 갖는 아미노산 서열을 갖는다.
- [0226] 본 발명에 따라 사용될 수 있는 대안적인 링커 서열은 서열번호 57, 63, 65, 67 및 121에 제공된다.
- [0227] 재조합 발현
- [0228] 본 발명에 따른 재조합 단백질 및 융합 단백질은 상기 단백질을 암호화하는 뉴클레오티드 서열의 발현을 가능하게 하는 숙주 세포를 배양함으로써 제조될 수 있다. 당업자는 일반적으로 이용 가능한 유전 공학 기술 및 재조합 DNA 발현 시스템을 사용하여 다양한 숙주 세포 시스템에서 이중 발현에 의한 재조합 단백질의 제조를 위한 단리된 핵산 서열의 발현을 제공하는 다양한 이용 가능한 생명공학 기술에 대해 잘 알고 있다. 예를 들어, 문헌

[“Recombinant Gene Expression Protocols, in Methods in Molecular Biology, 1997, Ed. Rocky S Tuan, Human Press (ISSN 1064-3745)] 또는 문헌[Sambrook et al., Molecular Cloning: A laboratory Manual (third edition), 2001, CSHL Press, (ISBN 978-087969577-4)] 참조. 예를 들어, 본 발명에 따른 재조합 단백질을 암호화하는 핵산 서열은 적합한 숙주 세포에서 원하는 단백질 암호화 핵산 서열의 발현을 유도하도록 특별히 개조한 모든 필요한 전사 및 번역 조절 서열을 포함하는 적합한 발현 벡터에 삽입될 수 있다. 적합한 발현 벡터는 예를 들어, 플라스미드, 코스미드, 바이러스 또는 인공 효모 염색체(YAC)이다.

- [0229] 본 발명의 재조합 단백질을 암호화하는 DNA 서열은 당업자에게 잘 알려진 방법 또는 당업자에게 잘 알려진 상업적 공급업체, 예를 들어, Genscript, Thermo Fisher Scientific 등에 의해 합성될 수 있다.
- [0230] 이 양태의 일 구현예에 따르면, 서열번호 86, 87, 90, 91, 99, 100, 104, 105, 108, 109, 112, 또는 113에 제시된 바와 같은 핵산 서열, 또는 임의의 진술한 서열과 적어도 80%의 서열 동일성을 갖는 서열을 포함하는 DNA 분자가 제공된다. 이러한 DNA 분자를 포함하는 발현 벡터가 또한 제공된다. 이 양태의 또 다른 구현예에 따르면, 이러한 벡터를 포함하는 숙주 세포가 또한 제공된다.
- [0231] 발현되고 재조합 단백질을 제조하는 데 사용되는 DNA 서열은 게이트웨이 클로닝 시스템을 사용하여 일반적으로 엔트리 벡터(entry vector)로 알려진 벡터에 삽입될 수 있다(Esposito et al, 2009, “Gateway Cloning for Protein Expression”, in Methods in Molecular Biology, 498, pp. 31-54). 엔트리 벡터에 클로닝된 유전자는 재조합에 의해 다양한 발현 벡터에 쉽게 도입될 수 있다. 예를 들어, 본 발명에 따른 재조합 단백질 또는 융합 단백질을 암호화하는 합성된 서열은 BP 게이트웨이 재조합효소 클로닝에 의해 재조합되어 적합한 숙주 세포, 예컨대, 대장균 세포에서 플라스미드를 증식시키는 데 사용될 수 있는 엔트리 벡터를 생성할 수 있다. 바람직한 구현예에서, 플라스미드의 효율적인 증식을 허용하도록 돌연변이된 대장균 세포, 예를 들어 One Shot Top10™ 세포가 사용된다.
- [0232] 본 발명의 일 구현예에 따르면, 프로모터에 작동 가능하게 연결된 본 발명에 따른 재조합 단백질 또는 융합 단백질을 암호화하는 DNA 서열을 포함하는 발현 벡터가 제조된다. 당업자는 본원에서 사용되는 “프로모터”가 특정 유전자의 전사를 제어하고 개시하는 DNA 암호화 서열의 DNA 상류(5'-프라임) 영역을 지칭함을 인지할 것이다. 프로모터는 전사를 개시하기 위하여 RNA 중합효소 및 기타 단백질의 인식 및 결합을 제어한다. “작동 가능하게 연결된”은 프로모터와 제2 서열 사이의 기능적 연결을 지칭하며, 여기에서 프로모터 서열은 제2 서열에 상응하는 DNA 서열의 전사를 개시하고 매개한다. 일반적으로, 작동 가능하게 연결된 것은 연결되는 핵산 서열이 연속적임을 의미한다.
- [0233] 하기에 언급된 목적지 벡터로부터 생성된 것과 같은 엔트리 벡터 및 발현 벡터는 예를 들어 Qiagen™의 QIAprep™ Spin Miniprep 키트 또는 QIAGEN™ 플라스미드 플러스 맥시 키트를 사용하는 것과 같이, 당업자에게 잘 알려진 표준 플라스미드 분리 기술을 이용하여 분리될 수 있다.
- [0234] 본 발명에 따른 재조합 단백질 또는 융합 단백질을 암호화하는 DNA 서열을 함유하는 엔트리 벡터를 사용하는 경우, 상기 엔트리 벡터는 LR 게이트웨이 재조합효소를 사용하여 목적지 벡터와 추가로 재조합되어 발현 벡터를 생성할 수 있다. 그런 다음, 발현 벡터를 사용하여 적절한 숙주 세포에서 단백질 암호화 DNA 서열을 발현할 수 있다. 적용 가능한 목적지 벡터의 비 제한적인 예는 예를 들어, 문헌[Kaasbøl et al., *J. Biol. Chem.*, 293:46, pp. 17953 - 17970]에 기술된 바와 같은 pUCOE-DHFR-DEST이다.
- [0235] 또한, 생성된 발현 벡터는 표준 제한 효소 소화 및 DNA 겔 전기영동에 의해 검증될 수 있다.
- [0236] 본 발명의 일 양태에 따르면, 화학식 I의 재조합 단백질을 암호화하는 핵산 서열을 포함하는 발현 벡터가 제공된다. 본 발명의 또 다른 양태에 따르면, 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 및 서열번호 8 내지 12, 37, 38, 84, 85, 88, 89, 97, 98, 102, 103, 106, 107, 110 및 111로 구성된 군으로부터 선택되는 아미노산 서열; 및 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 및 서열번호 38, 및 서열번호 84, 85, 88, 89, 97, 98, 102, 103, 106, 107, 110 및 111의 아미노산 서열과 적어도 50%의 서열 동일성을 갖는 이의 단편 또는 변이체를 포함하는 발현 벡터가 제공된다.
- [0237] 또 다른 양태에 따르면, 서열번호 1, 서열번호 2, 서열번호 3, 서열번호 4, 서열번호 5, 서열번호 6, 서열번호 7, 서열번호 8, 서열번호 9, 서열번호 10, 서열번호 11, 서열번호 12, 서열번호 37, 및 서열번호 38, 및 서열번호 84, 85, 88, 89, 97, 98, 102, 103, 106, 107, 110 및 111로 구성된 군으로부터 선택되는 아미노산 서열과 적어도 60%, 70%, 80%, 90%, 또는 95%의 서열 동일성을 갖는 아미노산 서열을 포함하는 재조합 단백질을 암호화

하는 발현 벡터가 제공된다.

[0238] 본 발명의 또 다른 구현예에 따르면, 본 발명에 따른 융합 단백질을 암호화하는 발현 벡터가 제공된다.

[0239] 당업자는 유전자 부호의 변성 및 다양한 유기체의 특정 코돈에 대한 선호도를 잘 알고 있다. 따라서, 숙주 세포의 선택에 따라, 본 발명의 재조합 단백질 및 융합 단백질을 암호화하는 핵산 서열은 숙주 세포의 바람직한 코돈에 맞추어 조정될 수 있다. 따라서, 본 발명의 단백질의 아미노산은 하기 표에 제시된 바와 같이 코돈의 임의의 조합에 의해 암호화될 수 있다:

아미노산	단일 문자 아미노산 부호	DNA 코돈
------	------------------	--------

[0240]

이소류신	I	ATT, ATC, ATA
류신	L	CTT, CTC, CTA, CTG, TTA, TTG
발린	V	GTT, GTC, GTA, GTG
페닐알라닌	F	TTT, TTC
메티오닌	M	ATG
시스테인	C	TGT, TGC
알라닌	A	GCT, GCC, GCA, GCG
글리신	G	GGT, GGC, GGA, GGG
프롤린	P	CCT, CCC, CCA, CCG
트레오닌	T	ACT, ACC, ACA, ACG
세린	S	TCT, TCC, TCA, TCG, AGT, AGC
티로신	Y	TAT, TAC
트립토판	W	TGG
글루타민	Q	CAA, CAG
아스파라긴	N	AAT, AAC
히스티딘	H	CAT, CAC
글루탐산	E	GAA, GAG
아스파르트산	D	GAT, GAC
리신	K	AAA, AAG
아르기닌	R	CGT, CGC, CGA, CGG, AGA, AGG
정지 코돈	정지	TAA, TAG, TGA

[0241]

[0242] 바람직하게는 코돈은 선택된 숙주 세포에 따라 고발현을 위해 더욱 최적화된다.

[0243] 재조합 DNA 기술에 의한 단백질의 발현을 위하여, 본 발명의 특정 구현예에 더하여, 신호 펩티드를 암호화하는 DNA 서열은 바람직하게는 단백질 서열의 N 말단 단부에 부가된다. 신호 펩티드는 숙주 세포에서 합성 중에 및/또는 합성 후에 융합 단백질의 국소화를 유도하는 역할을 할 수 있다. 따라서, 이는 융합 단백질의 분비를 유도

하는 서열일 수 있다. 이러한 신호 펩티드 서열의 사용은 당해 분야에 잘 알려져 있다. 신호 펩티드는 임의의 형태를 취할 수 있으며, 예를 들어, 이는 IgGk-쇄 신호 펩티드를 구성할 수 있거나, 또는 인간 혈청 알부민(서열번호 32)으로부터의 신호 펩티드를 구성할 수 있다.

[0244] 인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 서열번호 28의 N 말단에 부가되는 경우, 발현되는 단백질 서열은 서열번호 33, 또는 서열번호 85, 89, 98, 103, 107, 또는 111에 제시된 바와 같을 수 있다.

[0245] 또한, 재조합 DNA 기술에 의한 단백질의 발현을 위하여, 본 발명의 일 특정 구현예에 따르면, 서열번호 33에 제시된 바와 같은 아미노산 서열, 서열번호 34 또는 서열번호 86, 90, 99, 104, 108, 또는 112에 제시된 바와 같은 뉴클레오티드 서열을 갖는 단백질이 사용될 수 있고, 여기에서 번역 정지 코돈은 암호화 서열의 3' 말단에 부가된다.

[0246] 본 발명이 서열번호 34의 뉴클레오티드 서열에 의해 구현되는 경우, 뉴클레오티드 서열은 바람직하게는 서열번호 35 또는 서열번호 86, 90, 99, 104, 108, 또는 112에서와 같이, Kozak 서열, 예를 들어, GCCACC가 암호화 서열의 바로 5' 말단에 부가된다. DNA 서열은 예를 들어 게이트웨이 재조합효소 attB 부위와 같은 서브클로닝을 가능하게 하기 위하여 추가로 DNA 요소가 양 측면에 위치할 수 있다. 그러나 임의의 클로닝 또는 합성 전략을 사용하여 DNA 서열을 생성하고 발현 벡터로의 서브클로닝을 촉진할 수 있다. DNA 서열이 게이트웨이 재조합효소 부위를 포함하여 서브클로닝이 가능한 경우, 뉴클레오티드 서열은 서열번호 36 또는 서열번호 87, 91, 100, 105, 109, 또는 113에 제시된 바와 같을 수 있다.

[0247] 본 발명의 융합 단백질의 재조합 단백질을 암호화하는 핵산 서열을 포함하는 수득된 발현 벡터는 원하는 단백질의 생산에 적합한 숙주 세포에 도입될 수 있다. 다양한 상업적으로 입수 가능한 또는 독점적인 숙주 세포가 사용될 수 있다. 예를 들어, 발현 벡터는 진핵생물 숙주 세포, 예컨대, CHO 세포, 예를 들어 CHO DG44 DHFR(디하이드로폴레이트 환원효소^{-/-}) 현탁 배양 맞춤형 CHO 세포로 전달될 수 있다. 발현 벡터를 사용한 숙주 세포의 형질 감염은 당업자에게 잘 알려진 방법, 예를 들어 전기천공법에 의해 수행될 수 있다.

[0248] 적합한 배양 배지에서 숙주 세포를 배양하면, 숙주 세포에서 발현 벡터에 의해 암호화되는 본 발명에 따른 재조합 단백질 또는 융합 단백질이 생성되고, 생성된 단백질은 당업자에게 잘 알려진 방법에 의해 수집 및 정제될 수 있다.

[0249] 발현 벡터는 배양 배지 내로 발현된 단백질 또는 융합 단백질의 분비를 위하여, 일반적으로 "신호 펩티드"로 알려진 신호 서열을 포함할 수 있다.

[0250] 세포 배양 배지로부터 분리된 재조합 단백질의 단리 및 정제를 위하여, 큰 입자 및 바이오매스를 제거하기 위해 하나 이상의 전처리 또는 정화 단계가 일반적으로 먼저 사용된다. 적용 가능한 전처리 단계의 비 제한적인 예는 예를 들어 역삼투, 원심분리, 여과 방법 및 정용여과, 또는 이들의 조합이다. 그런 다음, 수득된 단백질은 일반적으로 당업자에게 잘 알려진 다양한 크로마토그래피 방법 중 하나 이상, 예를 들어 친화도 크로마토그래피, 이온 교환 크로마토그래피, 혼합 모드 크로마토그래피, 소수성 상호작용 크로마토그래피, 크기 배제 크로마토그래피 또는 기타 크로마토그래피 기술, 또는 이들의 조합에 의해 정제된다.

[0251] 예를 들어, 적합한 숙주 세포에 의해 발현된 재조합 단백질 또는 융합 단백질은 친화도 크로마토그래피 방법을 이용하여, 예컨대, MabSelect™ SuRe™ 배지, 예를 들어, FPLC 크로마토그래피 시스템, 예를 들어, 5 mm의 UV 플로우 셀이 장착된 BioRad NGC Discover™ 10 Pro 시스템에 장착된 5 ml의 HiTrap MabSelect™ SuRe™ 컬럼을 이용하여 정제될 수 있다. 정제할 단백질을 포함하는 샘플을 로딩한 후, 컬럼은 일반적으로 하나 이상의 적용 가능한 세척 완충액으로 1회 이상 세척되고, 그 후 단백질은 적용 가능한 용리 완충액을 사용하여 용리된다. 얻어진 단백질은 위에 열거된 크로마토그래피 방법 중 하나 이상을 사용하여 추가로 정제할 수 있다.

[0252] 다양한 변형이 예를 들어 발현을 용이하게 하기 위하여 당업자에게 잘 알려진 기술을 이용하여 본 발명의 재조합 단백질을 암호화하는 핵산 서열에 도입될 수 있음을 이해해야 한다. 부위 지정 돌연변이가 유발의 사용에 의해, 서열을 발현하는 데 사용되는 원하는 숙주에 대해 암호화 서열을 맞춤 개조하고, 따라서 재조합 단백질을 생산하도록 변형이 도입될 수 있다. 당업자는 숙주 특이적 코돈의 존재 및 숙주 특이적 코돈에 맞춘 이중 핵산 서열의 조정이 위에서 언급된 바와 같이 발현 효율을 증가시킨다는 사실을 잘 알고 있다. 예를 들어, 단리 및 정제를 용이하게 하기 위하여, 즉 이러한 목적에 유용한 펩티드 또는 단백질을 암호화하는 서열을 추가함으로써, 다른 변형이 또한 도입될 수 있다. 또한, 숙주 세포로부터 원하는 재조합 단백질의 분비를 제공하는 신호 펩티드를 암호화하는 핵산 서열은 또한 본 발명의 핵산 서열에 연결될 수 있다.

- [0253] 또한, 본 발명은 본 발명에 따른 재조합 단백질 또는 융합 단백질의 생산에 적합한 숙주 세포를 제공한다. 재조합 단백질의 생산을 위해 특수하게 맞춤 개조된 다양한, 상업적으로 입수 가능한 숙주 세포, 원핵 숙주 세포 및 진핵 숙주 세포가 사용될 수 있다. 적합한 숙주 세포의 비 제한적 예는 예를 들어, CHO 세포, HEK293 세포, 피치아 파스토리스(Pichia pastoris) 세포, NS0 세포 또는 대장균 세포이다.
- [0254] 마지막으로, 본 발명은 또한 CCN 패밀리 단백질에 기인하는 세포 신호전달 및 세포 생리 기능을 억제하거나 중화함으로써 장애의 치료 또는 예방을 위한 의약으로서 사용하기 위한, CCN 패밀리 단백질의 트롬보스폰딘 유형 1 반복(TSP-1) 상동성 도메인 및 상기 TSP-1 반복 상동성 도메인을 포함하는 융합 단백질에 관한 것이다.
- [0255] 일 양태에서, 본 발명은 치료법에 사용하기 위한, 본원에 정의된 바와 같은 단백질, 예를 들어, 융합 단백질을 제공한다.
- [0256] 일부 양태에서, 단백질, 예를 들어, 융합 단백질은 섬유증, 또는 섬유증을 나타내는 임의의 병태(즉, 임의의 섬유증 병태 또는 장애)의 치료 또는 예방에 사용하기 위한 것일 수 있다. 섬유증은 예를 들어, 폐, 눈, 심장, 골격근, 복막, 신장, 간, 췌장, 담관, 피부, 혈관 또는 기타 전신계를 비롯한 임의의 조직 또는 기관에 영향을 미칠 수 있다. 특히, 섬유증을 나타내는 병태는 특발성 폐 섬유증을 포함한, 임의의 병인에 의한 것일 수 있는 폐 섬유증, 기관지폐 이형성증, 망막 섬유증, 당뇨병성 망막병증, 연령 관련 황반 변성, 망막 박리, 산소 유발 망막병증, 녹내장, 심장 섬유증, 이식 후 섬유증, 심근병증 관련 섬유증, 근육 섬유증, 뒤센 근이영양증, 복막 섬유증, 당뇨병성 신병증, 만성 신장 질환(신장 섬유증), 급성 신장 손상, 세뇨관간질 섬유증, 만성 동종이식편 신병증, 간 섬유증, 비 알코올성 지방간염, 지방간 질환, 만성 췌장염, 담도 섬유증, 켈로이드, 흉터, 전신 경화증, 죽상동맥경화증, 경막의 섬유증으로부터 선택될 수 있다.
- [0257] 심장 섬유증과 관련하여, 치료 또는 예방되는 병태는 보존된 박출률이 있거나 없는 심장 비대 및 심부전을 포함할 수 있다.
- [0258] 추가 양태에서, 본 발명은 염증성 또는 자가면역 질환의 치료에 사용하기 위한, 본원에 정의된 바와 같은 단백질, 예를 들어, 융합 단백질을 제공한다. 일부 구현예에서, 염증성 질환은 류마티스 관절염, 근위축성 측삭 경화증(ALS), 염증성 장 질환, 궤양성 대장염, 크론병으로부터 선택된다.
- [0259] 추가 양태에서, 본 발명은 암의 치료에 사용하기 위한, 본원에 정의된 바와 같은 단백질, 예를 들어, 융합 단백질을 제공한다. 이와 관련하여 4개 도메인의 CCN 단백질은 단리된 암세포에서 발암성 반응을 유도할 수 있을 뿐만 아니라 암세포 또는 종양 기질에 직접 작용하여 전이, 화학내성 및 면역요법 내성에 기여할 수 있는 것으로 알려져 있다. 따라서, 4개 도메인 CCN 단백질의 효과 또는 활성을 억제하는 본원의 단백질의 활성은 암 치료에 대한 근거를 제공한다. 암은 임의의 악성 또는 전암성 신 생물 병태일 수 있다. 이는 임의의 조직 또는 기관의 것일 수 있다. 일 구현예에서 암은 고형 종양으로 나타날 수 있다. 또 다른 구현예에서 암은 조혈계의 것일 수 있거나 조혈계에 있을 수 있다. 이는 원발성 암 또는 이차 암 또는 전이일 수 있다. 따라서, 암은 췌장암, 유방암, 전립선암, 자궁경부암, 난소암, 간암, 방광암, 뇌암, 혈액암, 뼈암, 피부암, 폐암 또는 위암일 수 있다. 일부 구현예에서, 암은 췌장암, 췌관 선암종, 유방암, 전립선암, 자궁경부암, 난소 암종, 간암, 간세포 암종, 요로상피 방광암, 뇌암, 교모세포종, 급성 림프모구성 백혈병, 골육종, 흑색종, 중피종, 위암종, 구강 편평세포암종, 식도암, 결장직장암, 폐암으로부터 선택된다.
- [0260] 추가 양태에서, 본 발명은 대사 질환의 치료에 사용하기 위한, 본원에 정의된 바와 같은 단백질, 예를 들어, 융합 단백질을 제공한다. 대사 질환은 인슐린 저항성 또는 포도당 불내증일 수 있거나 이와 관련될 수 있다. 일부 구현예에서, 대사 질환은 제2형 당뇨병 및 대사 증후군으로부터 선택된다.
- [0261] 본 발명의 융합 단백질은 또한 전술한 병태의 치료 방법에 사용될 수 있다. 유사하게, 본 발명의 융합 단백질은 전술한 병태의 치료에 사용하기 위한 의약의 제조 방법에 사용될 수 있다.

도면의 간단한 설명

- [0262] **도 1**은 CCN5(dIII)-Fcv2(서열번호 28의 서열에 정의된 바와 같은 본 발명의 구현예)의 세포 생리학적 및 세포 신호전달을 나타낸다.
- a) 는 서열번호 28의 CCN5(dIII)-Fcv2 융합 단백질이 A549 폐암 세포에서 Akt(세린-473)의 인산화를 농도 의존적으로 억제함을 보여준다.
- b) 는 서열번호 28의 CCN5(dIII)-Fcv2 융합 단백질이 인간 폐 섬유모세포 세포주인 IMR90에서 증식을 억제함을

보여준다.

c) 는 서열번호 28의 CCN5(dIII)-Fc_{v2} 융합 단백질이 에스트로겐 수용체 양성 유방암 세포주 MCF-7 및 삼중 음성 유방암 세포주 MDA-MB-231의 구체 형성 능력(부착 비의존성 성장)을 억제함을 보여준다.

d) 는 서열번호 28의 CCN5(dIII)-Fc_{v2} 융합 단백질이 TGF- β 유도성 SMAD 리포터 활성을 용량 의존적으로 억제함을 보여준다(SMAD 단백질은 표준 TGF- β 조절성 전사 인자임).

모든 오차 막대는 SD를 보여준다. 통계적 유의성은 Dunnetts 사후 검정을 사용하여 일원 분산분석으로 계산하였다($p < 0.05$ 는 *로 표시함).

도 2는 CCN5(dIII)가 IgG의 Fc 단편에 융합된 본 발명의 구현예의 단백질 분해효소 감수성에 미치는 Fc 단편의 인지 영역의 상이한 변이체의 영향을 보여주며, 여기서 테스트한 융합 단백질은 각각 서열번호 28; 서열번호 29; 및 서열번호 30에 제시된 바와 같은 서열을 포함한다. 아래 실시예 6 참조.

도 3은 CCN5(dIII)를 IgG의 Fc 단편과 연결하는 펩티드 링커의 구조에 따른 본 발명의 구현예의 응집 경향을 보여주며, 여기서 테스트한 융합 단백질은 서열번호 30 및 서열번호 31에 제시된 바와 같은 서열을 포함한다.

도 4는 펩티드 링커에 C 말단으로 연결되고 Fc 힌지를 통해 Fc 단편에 연결되는 TSP-1 반복 상동성 도메인을 포함하는 본 발명에 따른 융합 단백질을 도시한다.

도 5는 서열번호 7(Fc-HLn8-CCN5(dIII))-P195A, 서열번호 41)에 도시된 바와 같이, 본 발명의 구현예가 CCN5 TSP-1 반복 상동성 도메인의 프롤린195의 돌연변이를 포함하는 경우, CCN5 TSP-1 반복 상동성 도메인의 야생형 P195 변이체(Fc-HLn8-CCN5(dIII), 서열번호 40)와 비교하여, 엔도펩티다제 절단에 대한 감소된 감수성을 보여준다.

도 6은 단백질 A 포획 크로마토그래피에 의해 정제된 서열번호 58에 상응하는 단백질의 생산을 보여준다. 환원제 베타-메르캅토에탄올의 부재(- 라인) 하에 이량체가 존재함을 알 수 있다. 그러나 베타-메르캅토에탄올의 존재(+ 라인) 하에, 1차 생성물은 서열번호 58에 의해 암호화된 부분 모두(TSP-1 상동성 도메인 단편, 펩티드 링커 및 Fc 단편)를 함유하는 온전한 융합 단백질이 아니라, Fc 단편만으로 구성된 절단 단편임을 알 수 있다.

도 7은 단백질 A 포획 크로마토그래피에 의해 정제된, 절단된 C 말단 꼬리를 갖는, 서열번호 27에 상응하는 단백질의 생산을 보여준다. 이 단백질은 C 말단 꼬리가 포함된 서열번호 58에 상응하는 단백질보다 단백질 분해효소 분해에 대해 유의미하게 더 저항적임을 알 수 있다.

도 8은 단백질 A 포획 크로마토그래피에 의해 정제된, 서열번호 27에 상응하는 단백질과 유사한, 서열번호 73에 상응하는 단백질의 생산을 보여준다. 다시, 베타-메르캅토에탄올의 존재(+ 라인) 하에, 단백질이 서열번호 58에 상응하는 단백질보다 단백질 분해효소 분해에 더 저항적임을 알 수 있다.

도 9는 안정적으로 형질감염된 세포에서 생산된 서열번호 41에 상응하는 단백질의 다양한 농도를 투여한 후 A549 폐암 세포에서 포스포-AKT(Ser-473) 수준을 측정된 분석 결과를 보여준다. 이 단백질은 AKT의 인산화를 억제하지 않음을 알 수 있다.

도 10은 안정적으로 형질감염된 세포에서 생산된 서열번호 80에 상응하는 단백질의 다양한 농도를 투여한 후 A549 폐암 세포에서 포스포-AKT(Ser-473) 수준을 측정된 분석 결과를 보여준다. 이 단백질은 AKT의 인산화를 유의하게 억제하지 않으며 실제로는 더 높은 농도에서 심지어 포스포-AKT를 증가시킬 수 있음을 알 수 있다.

도 11은 일시적으로 형질감염된 세포에서 생산된 서열번호 80에 상응하는 단백질의 다양한 농도를 투여한 후 A549 폐암 세포에서 포스포-AKT(Ser-473) 수준을 측정된 분석 결과를 보여준다. 일시적으로 형질감염된 세포에서 생산될 경우, 이 단백질은 AKT의 인산화에 대해 농도 의존적 억제 활성을 가짐을 알 수 있다.

도 12는 서열번호 84, 94 및 106에 상응하는 단백질의 다양한 농도를 투여한 후 A549 폐암 세포에서 포스포-AKT(Ser-473) 수준을 측정된 분석 결과를 보여준다. 이들 각각의 단백질은 AKT의 인산화에 대한 농도 의존적 억제 활성을 가지고 있음을 알 수 있다.

도 13은 서열번호 88에 상응하는 단백질의 다양한 농도를 투여한 후 A549 폐암 세포에서 포스포-AKT(Ser-473) 수준을 측정된 분석 결과를 보여준다. 이 단백질은 10 μ g/ml 초과 농도에서 AKT의 인산화를 억제할 수 있음을 알 수 있다.

도 14는 서열번호 102 및 97에 상응하는 단백질의 다양한 농도를 투여한 후 A549 폐암 세포에서 포스포-

AKT(Ser-473) 수준을 측정한 분석 결과를 보여준다. 두 단백질은 모두 AKT의 인산화에 대한 농도 의존적 억제 활성을 가지고 있음을 알 수 있다.

도 15는 서열번호 110에 상응하는 단백질의 다양한 농도를 투여한 후 A549 폐암 세포에서 포스포-AKT 수준을 측정한 분석 결과를 보여준다. 이 단백질은 AKT의 인산화에 대한 농도 의존적 억제 활성을 가지고 있음을 알 수 있다.

도 16은 서열번호 106에 상응하는 단백질을 포함하는 여러 실험의 결과를 보여준다.

a)는 이 단백질이 TGF-베타 및 CCN2 둘 다에 의해 유도된 인간 폐 섬유모세포의 이동을 억제함을 보여준다.

b)는 이 단백질이 TGF-베타 및 CCN2 둘 다에 의해 유도된 스크래치 상처의 봉합을 억제함을 보여준다.

c)는 이 단백질이 전섬유증인 것으로 알려진 유전자 COL1A1의 발현에 대한 TGF-베타 유도를 부분적으로 억제함을 보여준다.

d)는 이 단백질이 전섬유증인 것으로 알려진 유전자 FN1의 발현에 대한 TGF-베타 유도를 부분적으로 억제함을 보여준다.

e)는 이 단백질이 전섬유증인 것으로 알려진 유전자 ACTA2의 발현에 대한 TGF-베타 유도를 부분적으로 억제함을 보여준다.

f)는 이 단백질이 전섬유증인 것으로 알려진 유전자 CCN2의 발현에 대한 TGF-베타 유도를 부분적으로 억제함을 보여준다.

발명을 실시하기 위한 구체적인 내용

실시예

실시예 1

본 발명에 따른 융합 단백질의 발현

이 실시예에서, CCN5의 아미노산 194-246(서열번호 1), 펩티드 링커(서열번호 20)의 융합된 N 말단 및 서열번호 15의 IgG, IgG4 서브클래스의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질(CCN5(dIII)-FcV2)(즉, 서열번호 28에 따른 융합 단백질)의 제조가 설명된다. 융합 단백질은 서열번호 32의 알부민으로부터 유래하는 N 말단 신호 서열이 추가로 첨부되었고, 하기에 개시된 바와 같이 포유동물 세포에서 발현되었다.

서열번호 36의 서열에 제시된 DNA 서열은 상업적 공급업체에 의해 합성되고 서열 검증되었다. 합성된 서열은 엔트리 벡터를 생성하기 위하여 BP 게이트웨이 재조합효소 클로닝에 의해 pDonrZeo와 재조합되었다. 플라스미드의 효율적인 증식을 허용하도록 돌연변이된 적격(competent) 대장균(One Shot Top10™ 세포)의 형질감염 후, Qiagen™의 QIAprep™ Spin Miniprep 키트를 사용하여 표준 플라스미드 단리 기술로 엔트리 벡터를 단리하였다. 플라스미드 단리 후, 당업자에게 잘 알려진 표준 기술에 따라 제한 효소 소화에 이은 DNA 겔 전기영동에 의해 엔트리 벡터를 검증하였다.

서열번호 35의 서열을 함유하는 엔트리 벡터는 LR 게이트웨이 재조합효소를 사용하여 목적지 벡터와 추가로 재조합되었다. 사용된 목적지 벡터는 문헌[Kaasbøl et al., 2018, 위 참조]에 의해 기술된 바와 같이 pUCOE-DHFR-DEST였다.

플라스미드의 효율적인 증식을 허용하도록 돌연변이된 적격 대장균(One Shot Top10™ 세포)의 형질감염 후, QIAGEN™ 플라스미드 플러스 맥시 키트를 사용하여 표준 플라스미드 단리 기술로 발현 벡터를 단리하였다. 당업자에게 잘 알려진 표준 기술에 따라 표준 제한 효소 소화 및 DNA 겔 전기영동에 의해 생성된 발현 벡터를 검증하였다. 그런 다음, Expifectamine™ CHO 형질감염 키트(Gibco Cat.#: A29129)의 제조업체가 제공한 "Max Titer" 프로토콜에 따라, 그리고 문헌[Kaasbøl et al., 2018, 위 참조]에서 간략하게 설명된 바와 같이, 생성된 발현 벡터를 ExpiCHO 현탁액 배양물에 적응된 CHO 세포로 옮겼다. 세포를 4℃에서 20분 동안 4750 g에서 원심분리하여 형질감염 6일 후에 침강시키고, 상청액 세포 배양 배지를 수확하였다. 100% 이소프로판올 중 0.1 M의 PMSF를 1 mM의 농도로 첨가하고, 0.5 M의 EDTA를 2 mM의 농도로 첨가하였다. 그런 다음, 96%의 에탄올을 최종 농도가 약 3%가 되도록 첨가하였다. 크로마토그래피 정제 전에 1 M의 TrisHCl pH 7.4를 25 mM의 최종 농도가

되도록 첨가하였다.

- [0270] 정제의 포획 단계는 단백질 A 크로마토그래피 매질을 사용한 친화도 크로마토그래피에 의해 수행하였다. 이 실험에 사용된 매질은 rProtein A FF(GE Healthcare)였다. 위의 문헌에서 기술된 바와 같이 수확되고 보충된 세포 배양 배지 60 mL로부터 발현된 재조합 단백질의 정제를 위하여 5 mL HiTrap™ rProtein A FF 컬럼(GE Healthcare)을 사용하였다. HiTrap™ rProtein A FF 컬럼을 5 mm UV 플로우 셀이 장착된 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)에 장착하고, 25 mM의 TrisHCl pH 7.4, 25 mM의 NaCl 및 3%의 에탄올을 함유하는 완충액으로 평형화하였다. 재조합 단백질을 함유하는 수확된 세포 배양 배지를 샘플 펌프로 2.5 ml/분의 속도로 로딩한 후, 6의 컬럼 부피의 세척 완충액(25 mM의 TrisHCl pH 7.4, 25 mM의 NaCl 및 3%의 에탄올)으로 세척한 후, 3% 에탄올 중 0.1 M의 시트르산나트륨, pH 3.0으로 용리시켰다. 100 mAU를 초과하는 UV 280 nm의 흡광도를 갖는 용리액을 1 M의 TrisHCl pH 9.0 1 mL로 미리 채워진 저단백질 결합 튜브에서 3 mL의 분획으로 수집하였다. UV 흡광도 피크를 함유하는 분획을 Vivaspin® 20 mL, 30 kDa MWCO 농축 장치를 사용하여 500 µL까지 농축시켰다. 농축 후, 샘플을 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)의 샘플 로딩 루프에 로딩하였다. FPLC 크로마토그래피 시스템에는 50 mM의 NaCl, 20 mM의 HEPES pH 7.0으로 평형화된 Superdex® 200 Increase 10/300 GL 컬럼(GE Healthcare)이 장착되었다. 샘플을 주입하고, 0.25 mL/분의 유속으로 사전 평형 완충액(50 mM NaCl, 20 mM HEPES pH 7.0)으로 컬럼을 관류시켰다. 주요 UV 280 nm 흡광도 피크는 정제된 재조합 단백질(CCN5(dIII)-Fcγ2, 서열번호 28)을 함유하는 것으로 밝혀졌다. 수집된 분획의 10 µL 샘플을 Mini-PROTEAN® TGX Stain-Free™ 프리캐스트 겔을 사용하여 SDS-PAGE를 수행하고, 단리된 재조합 단백질을 ChemiDoc™ 영상화 시스템(BioRad)을 사용하여 시각화하였다.
- [0271] 재조합 단백질이 다양한 발현 시스템에서 생산될 수 있고 유사한 결과를 갖는 다양한 크로마토그래피 방법에 의해 정제될 수 있음은 당업자에게 널리 알려져 있다.
- [0272] 실시예 2
- [0273] CCN5의 아미노산 194~246(서열번호 1), 펩티드 링커(서열번호 20)의 융합된 N 말단 및 서열번호 15의 IgG, 서브클래스 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질(CCN5(dIII)-Fcγ2)을 암호화하는 DNA 서열을 서열번호 28에 따른 재조합 단백질을 생산하기 위하여 발현시켰다.
- [0274] 수득된 단백질을 A549 인간 폐암 세포에서 생존 촉진(prosurvival) 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 테스트하였다(도 1a). 조직 배양 처리된 Corning Incorporated Costar® 96 웰 멸균 폴리스티렌 플레이트를 피브로넥틴(BioWhittaker® 돌베코(Dulbecco) 인산염 완충 식염수(Lonza Cat. No. 17-512F, 이하 PBS로 지칭함)에서 10 µg/mL까지 희석된 Sigma Cat# F1141)으로 코팅하였다. 피브로넥틴을 함유하는 코팅 용액을 웰에 100 µL/웰의 부피로 분배하고 실온에서 1시간 동안 인큐베이션한 후, 코팅 용액을 경사분리하고, 100 µL의 PBS 또한 피브로넥틴이 코팅된 웰에 분배하였고, 이 또한 경사분리하였다. 최대 80%의 합류도(confluency)의 밀도를 유지하기 위해 계대 배양된 A549 세포를 효소 처리(Accutase®, biowest®의 Cat No. L0950-100)에 의해 분리하고, 10%의 열 불활성화 소 태아 혈청(FBS)(실온까지 평형화된 FBS(Cat. No. 16000-044 from gibco)가 있는 500 mL의 플라스크를 30분 동안 진탕하면서 60℃ 온도의 수조에서 인큐베이션하였다) 및 50 µg/ml의 겐수마이신(Sanofi)이 보충된 고 글루코스의 돌베코 변형 이글 배지(Gibco Cat. No.: 41965-039)에 110000개 세포/mL의 농도까지 희석하고, 100 µL의 세포 용액을 피브로넥틴이 코팅된 웰에 분배하였다. 모든 세포 인큐베이션은 37℃의 온도, 실내 공기 및 5% CO2의 습한 분위기를 유지하는 세포 배양 인큐베이터에서 수행되었다. 밤새 인큐베이션한 후, A549 세포를 PBS에서 2회 세척하고, FBS는 없이 50 µg/ml의 겐수마이신(Sanofi)이 있는 고 글루코스 함유 돌베코 변형 이글 배지(DMEM, Gibco Cat. No.: 41965-039) 90 µL를 웰에 분배하였다. FBS가 없는 배지에서 18시간 인큐베이션한 후, 세포를 해당 재조합 단백질의 10 µL 용액으로 자극하였다. 60분 동안 자극한 후, 배지를 경사분리하고, Cisbio 포스포-AKT(Ser473) 키트(Cisbio Inc, Cat. No.: 64AKSPEG)에서 제공하는 대로 차단 시약이 포함된 용해 완충액 50 µL를 첨가하여 세포를 수확하였다. 차단 시약이 있는 용해 완충액을 첨가한 후, 96웰 플레이트를 500 rpm에서 PST-60HL 플러스(ThermoFisher) 플레이트 진탕기에서 60분 동안 인큐베이션한다. 진탕 후, 용해된 샘플을 각 웰에서 백색 벽의 HTRF 96 웰 저부피 플레이트(Cisbio Inc., Cat.#: 66PL96025)로 16 µL를 옮기기 전에 트리튜레이션(trituration)하였다. 인산화된 AKT(Ser473)의 양을 분석하기 위하여, 4 µL의 표지된 항체 혼합물(Cisbio Inc의 포스포-AKT d2 및 포스포-AKT 크립테이트의 50/50 부피/부피 혼합물, Cat. No.: 64AKSPEG)을 각 웰에 첨가하고(음성 대조군 웰에는 크립테이트 항체만을 사용함), 플레이트를 접착성 플라스틱 필름으로 밀봉하고 TR-FRET 기록 헤드 및 337 nm 방출 및 615 nm 및 665 nm 여기 필터가 장착된 PolarStar Omega 플레이트 판독기(BMG Labtech, 독일)에서 판독하기 전에 밤새 4℃에서 인큐베이션하였다. 665 nm와 615 nm 여기 기록 사이의 비율을 블랭크로 보정하였고, 재조합 단

백질 자극된 켈의 값은 비히클 자극된 켈의 백분율로서 표시하였다.

- [0275] 실시예 3
- [0276] CCN5의 아미노산 194~246(서열번호 1), 펩티드 링커(서열번호 20)의 융합된 N 말단 및 서열번호 15의 IgG, 서브클래스 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질(CCN5(dIII)-Fc_v2)을 암호화하는 DNA 서열을 서열번호 28에 따른 재조합 단백질을 생산하기 위하여 발현시켰다.
- [0277] 수득된 단백질을 IMR90 인간 폐 섬유모세포에서 전섬유증 TGF- β 자극에 의한 전사(SMAD2/3 결합 시스템-요소로부터)를 억제하는 능력에 대해 테스트하였다. (도 1d). Rat2 세포 대신 2500개의 IMR90 폐 섬유모세포/켈을 사용하는 것을 제외하고는, 문헌[Kaasbøll et al. (2018), 위 참조]에 기술된 바와 같이 기술적으로 분석을 수행하였다. 자극에 사용된 단백질은 도 1d에 표시된 바와 같았다. IMR90 세포는 사용 전에 A549 세포에 대해 설명한 바와 같이(위 참조) 계대배양하였다. IMR90 세포는 계대 20 이전, 즉 복제 노쇠에 도달하기 전에 사용되었다.
- [0278] 실시예 4
- [0279] CCN5의 아미노산 194~246(서열번호 1), 펩티드 링커(서열번호 20)의 융합된 N 말단 및 서열번호 15의 IgG, 서브클래스 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질(CCN5(dIII)-Fc_v2)을 암호화하는 DNA 서열을 서열번호 28에 따른 재조합 단백질을 생산하기 위하여 발현시켰다.
- [0280] 수득된 단백질을 인간 폐 섬유모세포 세포주 IMR90의 증식을 억제하는 능력에 대해 테스트하였다(도 1b). IMR90 세포는 사용 전에 A549 세포에 대해 설명한 바와 같이(위 참조) 계대배양하였다. IMR90 세포는 계대 20 이전, 즉 복제 노쇠에 도달하기 전에 사용되었다. 실험을 위하여, A549 세포에 대해 설명한 바와 같이(위 참조), IMR90 세포를 수확하고, PBS에 세척하고, 실험 2에 대해 설명한 바와 같이(위 참조), 겐수마이신을 함유하고, 1% FBS를 함유하는 DMEM에서 희석하고, 12000/켈의 밀도로 xCELLigence 임피던스 플레이트에 시딩하였다. 2시간 후, 세포를 해당 재조합 단백질의 10 μ L 용액 또는 FBS로 자극하였고, 문헌[Kaasbøll et al., (2018), 위 참조]에서 설명한 바와 같이 CellTiter-Glo®(Promega Inc.)로 수확하기 전에 추가로 72시간 동안 인큐베이션하였다.
- [0281] 실시예 5
- [0282] CCN5의 아미노산 194~246(서열번호 1), 펩티드 링커(서열번호 20)의 융합된 N 말단 및 서열번호 15의 IgG, 서브클래스 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질(CCN5(dIII)-Fc_v2)을 암호화하는 DNA 서열을 서열번호 28에 따른 재조합 단백질을 생산하기 위하여 발현시켰다.
- [0283] 수득된 단백질을 문헌[Kaasbøll et al., 위 참조]에서 설명한 바와 같이, 에스트로겐 수용체 양성 유방암 세포주인 MCF-7과 삼중 음성 유방암 세포주인 MDA-MB-231의 구체 형성 능력(부착 비의존성 성장)을 억제하는 능력에 대해 테스트하였다(도 1c). MDA-MB-231 세포는 문헌[Kaasbøll et al., 위 참조]에서 MCF-7 세포주에 대해 설명한 바와 같이 처리하였다. MCF-7 및 MDA-MB-231 세포주를 A549 세포주에 대해 설명한 바와 같이(위 참조) 계대배양하였다.
- [0284] 실시예 6
- [0285] CCN5의 아미노산 194~246(서열번호 1), 펩티드 링커(서열번호 20)의 융합된 N 말단 및, 서열번호 15의 IgG, 서브클래스 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질(CCN5(dIII)-Fc_v2); 서열번호 18의 IgG, 서브클래스 IgG4의 Fc 단편(S228P, E233P, F234V, L235A, G236-, K447-, Eu-넘버링)을 포함하는 융합 단백질(CCN5(dIII)-Fc_v2.1); 또는 IgG2/4 서브클래스의 키메라 Fc 단편(서열번호 19)을 포함하는 융합 단백질(CCN5(dIII)-Fc_v2.3)을 암호화하는 DNA 서열을 서열번호 28(CCN5(dIII)-Fc_v2), 서열번호 29(CCN5(dIII)-Fc_v2.1) 및 서열번호 30(CCN5(dIII)-Fc_v2.3)에 따른 재조합 단백질을 생산하기 위하여 발현시켰다.
- [0286] 구체적으로, 서열번호 28(CCN5(dIII)-Fc_v2), 서열번호 29(CCN5(dIII)-Fc_v2.1) 및 서열번호 30(CCN5(dIII)-Fc_v2.3)의 발현을 암호화하는 발현 벡터를 Expifectamine™ CHO 형질감염 키트(Gibco Cat.#: A29129)의 제조업체가 제공한 "Max Titer" 프로토콜에 따라, 그리고 문헌[Kaasbøll et al., 위 참조]에서 간략하게 설명된 바와 같이, ExpicHO 현탁액 배양물에 적응된 CHO 세포로 형질감염시켰다. 세포를 Heraeus biofuge pico 벤치탑 원심분리기에서 13000 rpm으로 5분 동안원심분리하여 형질감염 6일 후에 침강시키고, 상청액 세포 배양 배지를 수확하였다. 수확된 세포 배양 상청액의 샘플을 Mini-PROTEAN® TGX Stain-Free™ 프리캐스트 겔을 사용하여 SDS-PAGE에 의해 분리하고, 재조합 단백질을 ChemiDoc™ 영상화 시스템(BioRad)을 사용하여 시각화하였다. 그런 다

음, 분리된 단백질을 웨스턴 블롯 분석을 위하여 반건식 블롯팅 시스템인 트랜스-블롯 터보(Trans-Blot Turbo)(Bio-Rad)를 사용하여 PVDF 막으로 옮겼다. 블롯은 시각화를 위하여 SuperSignal™ West Femto Maximum Sensitivity Substrate(ThermoFisherScientific) 및 ChemiDoc™ 영상화 시스템(BioRad)과 함께 사용된 홀스래디쉬 퍼옥시다제(Invitrogen Cat.#: A10654)에 접합된 항인간 IgG4 항체로 프로브되었다.

[0287] 도 2에는, (서열번호 19로 제시된) IgG2/4 키메라로 구성된 Fc 단편 골격의 개선된 단백질 분해효소 저항성을 나타내는 데이터가 제시되어 있다.

[0288] 면역 이펙터 침묵화 IgG4 힌지를 갖는 IgG4 Fc 단편에 융합된 CCN5/WISP2(도메인 III)(서열번호 28에 정의된 바와 같음); IgG2에 기초한 돌연변이를 포함하는 동일한 IgG4 백본인 CCN5(도메인 III)-Fcγ2(서열번호 29에 정의된 바와 같음); CCN5(도메인 III)-Fcγ2.1, 또는 IgG2로부터의 완전한 힌지 영역을 갖는 동일한 IgG4 백본(서열번호 30에 정의된 바와 같음); CCN5(도메인 III)-Fcγ2.3을 ExpiCHO 시스템에서 발현시켰고, 조정 배지(CM)는 6일 후에 수확되었다. SDS-PAGE 겔의 웨스턴 블로팅 및 총 단백질 염색은 CCN5(도메인 III)-Fcγ2.3 변이체가 배양 중에 존재하는 단백질 분해효소에 가장 덜 민감함을 보여준다. Fc 단편에 대한 항-IgG4 항체의 면역 반응성은 IgG2로부터의 서열 치환으로 부분적으로 손실되고, 따라서 일반적인 단백질 염색에 비해 단백질 수준을 과소평가한다는 점에 주목한다.

[0289] 실시예 7

[0290] CCN5의 아미노산 194~246(서열번호 1), 서열번호 20에 기술된 펩티드 링커의 융합된 N 말단 및 IgG2/4 서브클래스의 키메라 Fc 단편(서열번호 19)을 포함하는 융합 단백질을 암호화하는 DNA 서열. (CCN5(dIII)-Fcγ2.3) 또는 서열번호 25에 기술된 펩티드 링커 및 IgG2/4 서브클래스의 키메라 Fc 단편(서열번호 19)(CCN5(dIII)-HLn8-Fcγ2.3)을 서열번호 30(CCN5(dIII)-Fcγ2.3) 및 서열번호 31(CCN5(dIII)-HLn8-Fcγ2.3)에 따른 제조합 단백질을 생산하기 위하여 발현시켰다.

[0291] 구체적으로, 서열번호 30(CCN5(dIII)-Fcγ2.3) 및 서열번호 31(CCN5(dIII)-HLn8-Fcγ2.3)의 발현을 암호화하는 발현 벡터를 Expifectamine™ CHO 형질감염 키트(Gibco Cat.#: A29129)의 제조업체가 제공한 "Max Titer" 프로토콜에 따라, 그리고 문헌[Kaasbøll et al., 위 참조]에서 간략하게 설명된 바와 같이, ExpiCHO 현탁액 배양물에 적응된 CHO 세포로 형질감염시켰다. 세포를 Heraeus biofuge pico 벤치탑 원심분리기에서 13000 rpm으로 5분 동안 원심분리하여 형질감염 4일 후에 침강시키고, 상청액 세포 배양 배지를 수확하였다. 수확된 세포 배양 상청액의 샘플을 Mini-PROTEAN® TGX Stain-Free™ 프리캐스트 겔을 사용하는 SDS-PAGE에 의해 분리하였다. 분리된 단백질을 웨스턴 블롯 분석을 위하여 반건식 블롯팅 시스템인 트랜스-블롯 터보(Trans-Blot Turbo)(Bio-Rad)를 사용하여 PVDF 막으로 옮겼다. 블롯은 시각화를 위하여 SuperSignal™ West Femto Maximum Sensitivity Substrate(ThermoFisherScientific) 및 ChemiDoc™ 영상화 시스템(BioRad)과 함께 사용된 홀스래디쉬 퍼옥시다제(Invitrogen Cat.#: A10654)에 접합된 항인간 IgG4 항체로 프로브되었다.

[0292] 도 3에서, 본 발명의 구현예가 서열번호 25에 도시된 바와 같은 펩티드 링커를 포함할 때 응집에 대한 감소된 경향을 보여주는 데이터가 제공된다.

[0293] 다양한 펩티드 링커를 통해 키메라 IgG2/4 Fc 단편의 아미노 말단에 융합된 CCN5(도메인 III)를 발현하는 일시적으로 형질감염된 CHO 현탁액 세포로부터의 CM의 비환원 SDS-PAGE. 웨스턴 블롯은 서열번호 31에 도시된 바와 같은 아미노 서열; (dIII)-HLn8-Fcγ2.3을 갖는 융합 단백질이 서열번호 30에 도시된 바와 같은 아미노산 서열; CCN5(도메인 III)-Fcγ2.3을 갖는 본 발명의 융합 단백질보다 더 낮은 응집 경향을 가짐을 보여준다. 이 결과는 서열번호 25의 서열에 정의된 펩티드 링커가 서열번호 20의 서열에 정의된 펩티드 링커를 함유하는 융합 단백질과 비교하여 융합 단백질의 더 낮은 응집 경향을 제공함을 증명한다.

[0294] 실시예 8

[0295] CCN5의 아미노산 194~246(서열번호 1) 또는 195번 위치의 아미노산(프롤린)이 알라닌으로 치환된 CCN5의 아미노산 194~246(서열번호 7), 펩티드 링커(서열번호 39)의 융합된 C 말단 및 서열번호 15의 IgG 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질을 암호화하는 DNA 서열을 서열번호 40(Fc-HLn8-CCN5(dIII)) 또는 서열번호 41(Fc-HLn8-CCN5(dIII)-P195A)에 따른 제조합 단백질을 생산하기 위하여 발현시켰다.

[0296] 구체적으로, 서열번호 40(Fc-HLn8-CCN5(dIII)) 및 서열번호 41(Fc-HLn8-CCN5(dIII)-P195A)의 발현을 암호화하는 발현 벡터를 Expifectamine™ CHO 형질감염 키트(Gibco Cat.#: A29129)의 제조업체가 제공한 "Max Titer" 프로토콜에 따라, 그리고 문헌[Kaasbøll et al., 위 참조]에서 간략하게 설명된 바와 같이, ExpiCHO 현탁액 배

양물에 적응된 CHO 세포로 형질감염시켰다. 세포를 Heraeus biofuge pico 벤치탑 원심분리기에서 13000 rpm으로 5분 동안 원심분리하여 형질감염 3일 후에 침강시키고, 상청액 세포 배양 배지를 수확하였다. 수확된 세포 배양 상청액의 샘플을 Mini-PROTEAN® TGX Stain-Free™ 프리캐스트 겔을 사용하는 SDS-PAGE에 의해 분리하였다. 분리된 단백질을 웨스턴 블롯 분석을 위하여 반건식 블롯팅 시스템인 트랜스-블롯 터보(Trans-Blot Turbo)(Bio-Rad)를 사용하여 PVDF 막으로 옮겼다. 블롯은 시각화를 위하여 SuperSignal™ West Fento Maximum Sensitivity Substrate(ThermoFisherScientific) 및 ChemiDoc™ 영상화 시스템(BioRad)과 함께 사용된 홀스래디쉬 퍼옥시다제(Invitrogen Cat.#: A10654)에 접합된 항인간 IgG4 항체로 프로브되었다.

[0297] 도 5에서, 본 발명의 구현예가 서열번호 7에 도시된 바와 같이 CCN5 TSP-1 반복 상동성 도메인의 프롤린195의 돌연변이를 포함할 때 엔도펩티다제 절단에 대한 감소된 감수성을 나타내는 데이터가 제시된다.

[0298] P195A 돌연변이를 포함하거나(Fc-HLn8-CCN5(dIII)-P195A) CCN5 TSP-1 반복 상동성 도메인의 야생형 P195 변이체를 발현하는(Fc-HLn8-CCN5(dIII)) 서열번호 15에 기술된 바와 같이, IgG4 Fc 단편의 카르복실 말단에 융합된 CCN5(도메인 III)를 발현하는 일시적으로 형질감염된 CHO 현탁액 세포로부터 CM의 환원 SDS-PAGE. 이 블롯은 P195A 돌연변이가 CCN5의 TSP-1 반복 상동성 도메인에 대한 단백질 분해 저항성을 제공함을 보여준다.

[0299] 실시예 9

[0300] 서열번호 58에 상응하는 단백질 서열(CCN5(dIII)-SL-Fcv0)을 초래하는, 인간 CCN5의 아미노산 194-250(서열번호 56), 펩티드 링커(서열번호 57)의 융합된 N 말단 및 서열번호 15의 인간 IgG, IgG4 서브클래스의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질이 개시된다. 융합 단백질은 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열이 추가로 첨부되어 서열번호 59에 상응하는 융합 단백질을 생성하였고, 하기에 개시된 바와 같이 포유동물 세포에서 발현되었다.

[0301] 서열번호 59의 융합 단백질을 암호화하는 DNA 서열을 랩스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 60의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 61의 DNA 서열이 생성되었다. 서열번호 61의 서열은 상업적 공급업체에 의해 합성되고 검증되었다. 합성된 서열은 엔트리 벡터를 생성하기 위하여 BP 게이트웨이 재조합효소 클로닝에 의해 pDonrZeo와 재조합되었다. 플라스미드의 효율적인 증식을 허용하도록 돌연변이된 적격(competent) 대장균(One Shot Top10™ 세포)의 형질감염 후, Qiagen™의 QIAprep™ Spin Miniprep 키트를 사용하여 표준 플라스미드 분리 기술로 엔트리 벡터를 분리하였다. 플라스미드 분리 후, 당업자에게 잘 알려진 표준 기술에 따라 제한 효소 소화에 이은 DNA 겔 전기영동에 의해 엔트리 벡터를 검증하였다.

[0302] 서열번호 60의 서열을 함유하는 엔트리 벡터는 LR 게이트웨이 재조합효소를 사용하여 목적지 벡터와 추가로 재조합되었다. 사용된 목적지 벡터는 문헌[Kaasbøll et al., 2018, J. Biol. Chem, 293:46, pp. 17953 - 17970]에 의해 기술된 바와 같이 pUCOE-DHFR-DEST였다.

[0303] 플라스미드의 효율적인 증식을 허용하도록 돌연변이된 적격 대장균(One Shot Top10™ 세포)의 형질감염 후, QIAGEN™ 플라스미드 플러스 맥시 키트를 사용하여 표준 플라스미드 분리 기술로 발현 벡터를 분리하였다. 당업자에게 잘 알려진 표준 기술에 따라 표준 제한 효소 소화 및 DNA 겔 전기영동에 의해 생성된 발현 벡터를 검증하였다. 그런 다음, Expifectamine™ CHO 형질감염 키트(Gibco Cat.#: A29129)의 제조업체가 제공한 "Max Titer" 프로토콜에 따라, 그리고 문헌[Kaasbøll et al., 2018, 위 참조]에서 간략하게 설명된 바와 같이, 생성된 발현 벡터를 현탁액 배양물에 적응된 ExpiCHO 세포로 형질감염시켰다. 세포를 4℃에서 20분 동안 4750xg에서 원심분리하여 형질감염 4일 후에 침강시키고, 상청액 세포 배양 배지를 수확하였다. 100% 이소프로판올 중 0.1 M의 PMSF를 1 mM의 농도로 첨가하고, 0.5 M의 EDTA를 2 mM의 농도로 첨가하였다. 그런 다음, 96%의 에탄올을 최종 농도가 약 3%가 되도록 첨가하였다. 크로마토그래피 정제 전에 1 M의 TrisHCl pH 7.4를 25 mM의 최종 농도가 되도록 첨가하였다.

[0304] 단백질 A 크로마토그래피 매질을 사용한 친화도 크로마토그래피에 의해 단백질을 정제하였다. 이 실험에 사용된 크로마토그래피 매질은 rProtein A FF(GE Healthcare)였다. 위의 문헌에서 기술된 바와 같이 수확되고 보충된 세포 배양 배지 120 mL로부터 발현된 재조합 단백질의 정제를 위하여 5 mL HiTrap™ rProtein A FF 컬럼(GE Healthcare)을 사용하였다. HiTrap™ rProtein A FF 컬럼을 5 mm UV 플로우 셀이 장착된 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)에 장착하고, 25 mM의 TrisHCl pH 7.4, 25 mM의 NaCl 및 3%의 에탄올을 함유하는 완충액으로 평형화하였다. 재조합 단백질을 함유하는 수확된 세포 배양 배지를 샘플 펌프로

2.5 ml/분의 속도로 로딩한 후, 10의 컬럼 부피의 세척 완충액(25 mM의 TrisHCl pH 7.4, 25 mM의 NaCl 및 3%의 에탄올)으로 세척한 후, 3% 에탄올 중 0.1 M의 시트르산나트륨, pH 3.0으로 용리시켰다. 3 mL의 용리된 분획을 1 M의 TrisHCl pH 9.0 1 mL로 미리 채워진 저단백질 결합 튜브에 수집하였다. 단백질을 280 nm UV 흡광도로 모니터링되었으며, UV 280 nm 흡광도 피크를 함유하는 합한 분획의 10 µL 샘플을 환원제 β-메르캅토에탄올의 존재 또는 부재 하에 Mini-PROTEAN® TGX Stain-Free™ 프리캐스트 겔을 사용하여 SDS-PAGE를 수행하고, 분리된 재조합 단백질을 ChemiDoc™ 영상화 시스템(BioRad)을 사용하여 시각화하였다.

[0305] 재조합 단백질이 다양한 발현 시스템에서 생산될 수 있고 유사한 결과를 갖는 다양한 크로마토그래피 방법에 의해 정제될 수 있음은 당업자에게 널리 알려져 있다.

[0306] 도 6에서, 서열번호 58에 상응하는 단백질의 발현 및 정제는 환원제 β-메르캅토에탄올의 부재 하에 예상보다 높게 이동하는 단백질을 초래하여 이량체 형성을 나타내는 것으로 나타났다. 그러나 환원제 β-메르캅토에탄올의 존재 하에 정제된 단백질을 함유하는 레인에서 알 수 있는 바와 같이, 서열번호 58에 상응하는 단백질의 발현 및 정제는 온전한 단백질이 아닌 절단 단편을 주로 초래한다.

[0307] 실시예 10

[0308] 서열번호 58에 상응하는 단백질의 단백질 분해 저항성을 증가시키기 위하여 서열번호 58의 서열의 여러 변이체를 생성하였다. DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드 및 실시예 9에 기술된 바와 같이 발현된 단백질을 생성하기 위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었다. 변이체는 아래에 열거된 바와 같은 변형이 있는 단백질을 포함한다:

[0309] 1) 서열번호 32의 알부민 유래 N 말단 신호서열 아미노 말단에서 시작하고, 서열번호 63에 상응하는 펩티드 링커의 절단 및 서열번호 15의 Fc 단편과 조합되어, 서열번호 64에 상응하는 서열을 초래하는, 서열번호 62에 상응하는 돌연변이(P245L)를 포함하는 아미노산 194 내지 249로 구성된 CCN5의 단편,

[0310] 2) 서열번호 32의 알부민 유래 N 말단 신호서열 아미노 말단에서 시작하고, 서열번호 65에 상응하는 펩티드 링커의 변형 및 서열번호 15의 Fc 단편과 조합되어, 서열번호 66에 상응하는 서열을 초래하는, 서열번호 1에 상응하는 아미노산 194 내지 246으로 구성된 CCN5의 단편,

[0311] 3) 서열번호 32의 알부민 유래 N 말단 신호서열 아미노 말단에서 시작하고, 서열번호 67에 상응하는 펩티드 링커의 변형 및 서열번호 15의 Fc 단편과 조합되어, 서열번호 68에 상응하는 서열을 초래하는, 서열번호 1에 상응하는 아미노산 194 내지 246으로 구성된 CCN5의 단편,

[0312] 4) 서열번호 32의 알부민 유래 N 말단 신호서열 아미노 말단에서 시작하고, 서열번호 65에 상응하는 펩티드 링커의 변형 및 서열번호 19의 Fc 단편과 조합되어, 서열번호 69에 상응하는 서열을 초래하는, 서열번호 1에 상응하는 아미노산 194 내지 246으로 구성된 CCN5의 단편.

[0313] 실시예 12에 개시된 단백질의 이러한 반복(1 내지 4, 위 참조)은 실시예 9에 기술된 바와 같이 수행된 ExpiCHO 시스템에서 발현시키는 동안 단백질 분해 절단에 대한 저항성에서 약간의 개선을 나타내었다. 그러나 서열번호 64, 서열번호 66, 서열번호 68 및 서열번호 69에 상응하는 단백질의 발현은 단백질 분해 저항성의 정도가 정제된 온전한 단백질의 생산을 허용하기에 여전히 불충분함을 보여주었다.

[0314] 실시예 11

[0315] CCN5의 아미노산 194 내지 237(여기서, 195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨)(서열번호 38), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 카르복실 말단 K477-(Eu 넘버링)이 결실된 IgG 서브타입 IgG2/4의 키메라 Fc 단편(서열번호 19)을 포함하여 서열번호 27을 초래하는 융합 단백질을 생성하였다. 융합 단백질은 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열이 추가로 첨부되어 서열번호 70에 상응하는 융합 단백질을 생성하였다. 서열번호 70의 융합 단백질을 암호화하는 DNA 서열을 햄스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 71의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 72의 DNA 서열이 생성되었다.

[0316] DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드 및 실시예 9에 기술된 바와 같이 ExpiCHO 세포의 일시적인 형질감염에 의해 발현된 서열번호 70에 상응하는 단백질을 생성하기 위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었다.

- [0317] 세포를 4℃에서 20분 동안 4750xg에서 원심분리하여 형질감염 6일 후에 침강시키고, 상청액 세포 배양 배지를 수확하였다. 100% 이소프로판올 중 0.1 M의 PMSF를 0.1 mM의 농도가 되도록 첨가하였다. 크로마토그래피 정제 전에 1 M의 시트르산나트륨 pH 5.5를 30 mM의 최종 농도가 되도록 첨가하였다.
- [0318] 단백질은 1 mL의 HiTrap™ MabSelectSuRe™ 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 BioScale™ Mini Bio-Gel® P-6 10 mL 컬럼(BioRad)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다. 컬럼을 5 mm UV 플로우 셀이 장착된 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)에 장착하였다. MabSelectSuRe™ 컬럼을 제1 컬럼 전환 밸브에 장착하고 30 mM의 시트르산나트륨 pH 5.5로 구성된 완충액으로 평형화시켰고, Bio-Gel® 컬럼은 제2 컬럼 전환 밸브에 장착하고 완충액 A2(100 mM NaH2PO4/Na2HPO4 pH 6.5)로 평형화시켰다. Bio-Gel® 컬럼을 함유하는 제2 컬럼 전환 밸브를 바이패스하도록 설정한 상태에서, 재조합 단백질을 함유하는 수확된 세포 배양 배지 140 mL를 샘플 펌프를 사용하여 2.0 ml/분의 속도로 MabSelectSuRe™ 컬럼에 로딩한 후, 5 컬럼 부피의 세척 완충액 A1(30 mM 시트르산나트륨 pH 5.5)로 세척한 다음, 5 컬럼 부피의 세척 완충액 A3(30 mM 시트르산나트륨, 0.5 M NaCl, pH 5.5)에 이어, 3 컬럼 부피의 세척 완충액 A1로 세척하였다. 용리 완충액(30 mM의 시트르산 pH 3.4)으로 용리시키기 전에, 제2 컬럼 전환 밸브에 장착된 Bio-Gel® 컬럼을 유동 경로로 들어가도록 전환시켰다. 2 mL의 용리 완충액으로 용리시킨 후, MabSelectSuRe™ 컬럼을 유동 경로의 밖으로 전환시키고, 정제된 단백질을 완충액 A2를 사용하여 Bio-Gel® 컬럼으로부터 용리시켰다. 단백질 용리를 280 nm UV 흡광도로 모니터링하였으며, 흡광도가 100 mAU를 초과하면 수집이 촉발되었다. 수집된 분획을 합하고, 10 µL의 샘플을 환원제 β-메르캅토에탄올의 존재 하에 Mini-PROTEAN® TGX Stain-Free™ 프리캐스트 겔을 사용하여 SDS-PAGE를 수행하고, 단리된 재조합 단백질을 ChemiDoc™ 영상화 시스템(BioRad)을 사용하여 시각화하였다.
- [0319] 도 7에서, CCN5의 카르복실 말단 꼬리가 절단된, 서열번호 27에 상응하는 단백질의 발현 및 정제는, 세포 배양 배지가 실시예 9(도 6)에 비해 계대 배양 후 추가로 2일 후에 수확되었음에도 불구하고, (서열번호 58, 64, 66, 68 및 69에서와 같이) CCN5의 모든 카르복실 말단 아미노산이 포함된 변이체보다 실질적으로 더욱 단백질 분해 저항성이 있음을 보여주었다.
- [0320] 실시예 12
- [0321] CCN3의 아미노산 206 내지 249(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 카르복실 말단 K477-(Eu 넘버링)이 결실된 IgG 서브타입 IgG2/4의 키메라 Fc 단편(서열번호 19)을 포함하여 서열번호 73의 융합 단백질을 초래하는 융합 단백질을 생성하였다. 융합 단백질은 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열이 추가로 첨부되어 서열번호 74에 상응하는 융합 단백질을 생성하였다. 서열번호 74의 융합 단백질을 암호화하는 DNA 서열을 램스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 75의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 76의 DNA 서열이 생성되었다.
- [0322] DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드 및 실시예 9에 기술된 바와 같이 Exp CHO 세포의 일시적인 형질감염에 의해 발현된 서열번호 74에 상응하는 단백질을 생성하기 위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었다.
- [0323] 세포를 4℃에서 20분 동안 4750xg에서 원심분리하여 형질감염 5일 후에 침강시키고, 상청액 세포 배양 배지를 수확하였다. 100% 이소프로판올 중 0.1 M의 PMSF를 0.1 mM의 농도가 되도록 첨가하였다. 크로마토그래피 정제 전에 1 M의 시트르산나트륨 pH 5.5를 30 mM의 최종 농도가 되도록 첨가하였다.
- [0324] 단백질은 5 mL의 HiTrap™ MabSelectSuRe™ 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다. 컬럼을 5 mm UV 플로우 셀이 장착된 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)에 장착하였다. MabSelectSuRe™ 컬럼을 제1 컬럼 전환 밸브에 장착하고 30 mM의 시트르산나트륨 pH 5.5로 구성된 완충액으로 평형화시켰고, HiPrep™ 컬럼은 제2 컬럼 전환 밸브에 장착하고 완충액 A2(100 mM NaH2PO4/Na2HPO4 pH 6.5)로 평형화시켰다. HiPrep™ 컬럼을 함유하는 제2 컬럼 전환 밸브를 바이패스하도록 설정한 상태에서, 재조합 단백질을 함유하는 수확된 세포 배양 배지 260 mL를 샘플 펌프를 사용하여 3.5 ml/분의 속도로 MabSelectSuRe™ 컬럼에 로딩한 후, 5 컬럼 부피의 세척 완충액 A1(30 mM 시트르산나트륨 pH 5.5)로 세척한 다음, 5 컬럼 부피의 세척 완충액 A3(30 mM 시트르산나트륨, 0.5 M NaCl, pH 5.5)에 이어, 2 컬럼 부피의 세척 완충액 A1로 세척하였다. 용리 완충액(30 mM의 시트르산 pH 3.4)으로 용리시키기 전에, 제2 컬럼 전환 밸브에

장착된 HiPrep™ 컬럼을 유동 경로로 들어가도록 설정하였다. 10 mL의 용리 완충액으로 용리시킨 후, MabSelectSuRe™ 컬럼을 유동 경로의 밖으로 전환시키고, 정제된 단백질을 완충액 A2를 사용하여 HiPrep™ 컬럼 으로부터 용리시켰다. 단백질 용리를 280 nm에서의 UV 흡광도로 모니터링하였으며, 흡광도가 100 mAU를 초과하 면 수집이 촉발되었다. 수집된 분획을 합하고, 10 µL의 샘플을 환원제 β-메르캅토에탄올의 존재 또는 부재 하에 Mini-PROTEAN® TGX Stain-Free™ 프리캐스트 겔을 사용하여 SDS-PAGE를 수행하고, 단리된 재조합 단백질을 ChemiDoc™ 영상화 시스템(BioRad)을 사용하여 시각화하였다.

[0325] 도 8에서, 서열번호 27에 개시된 바와 같은 상동성 CCN5로부터 유래된 아미노산(도메인 III/TSP-1 상동성 도메인)을 함유하는 융합 단백질과 유사한, 서열번호 73에 개시된 바와 같은, CCN3/Nov로부터 유래된 아미노산(도메인 III/TSP-1 상동 도메인)을 함유하는 융합 단백질이, 실시예 11에 기술되고 도 7에 제시된 바와 같이, CCN5로부터 유래된 아미노산을 함유하는 융합 단백질보다 단백질 분해에 대해 비슷하거나 더 우수한 저항성을 가짐을 알 수 있다.

[0326] 실시예 13

[0327] 서열번호 41에 상응하는 단백질 서열을 초래하는, CCN5의 아미노산 194 내지 246(여기서, 195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨)(서열번호 7), 펩티드 링커(서열번호 39)의 융합된 C 말단 및 서열번호 15의 IgG, IgG4 서브클래스의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질에 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열이 첨부되어 서열번호 77에 상응하는 융합 단백질을 생성하였다. 서열번호 77의 융합 단백질을 암호화하는 DNA 서열을 햄스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 78의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 79의 DNA 서열이 생성되었다.

[0328] 서열번호 79의 서열은 상업적 공급업체에 의해 합성되고 서열 검증되었다. 합성된 서열은 엔트리 벡터를 생성하기 위하여 BP 게이트웨이 y재조합효소 클로닝에 의해 pDonrZeo와 재조합되었다. 플라스미드의 효율적인 증식을 허용하도록 돌연변이된 적격(competent) 대장균(One Shot Top10™ 세포)의 형질감염 후, Qiagen™의 QIAprep™ Spin Miniprep 키트를 사용하여 표준 플라스미드 단리 기술로 엔트리 벡터를 단리하였다. 플라스미드 단리 후, 당업자에게 잘 알려진 표준 기술에 따라 제한 효소 소화에 이은 DNA 겔 전기영동에 의해 엔트리 벡터를 검증하였다.

[0329] 서열번호 78의 서열을 함유하는 엔트리 벡터는 LR 게이트웨이 재조합효소를 사용하여 목적지 벡터와 추가로 재조합되었다. 사용된 목적지 벡터는 문헌[Kaasbøll et al., 2018, J. Biol. Chem., 293:46, pp. 17953 - 17970]에 의해 기술된 바와 같이 pUCOE-DHFR-DEST였다.

[0330] 플라스미드의 효율적인 증식을 허용하도록 돌연변이된 적격 대장균(One Shot Top10™ 세포)의 형질감염 후, QIAGEN™ 플라스미드 플러스 맥시 키트를 사용하여 표준 플라스미드 단리 기술로 발현 벡터를 단리하였다. 당업자에게 잘 알려진 표준 기술에 따라 표준 제한 효소 소화 및 DNA 겔 전기영동에 의해 생성된 발현 벡터를 검증하였다. 그런 다음, 생성된 발현 벡터를 Neon 형질감염 시스템(ThermoFisherScientific)을 사용하는 전기천공법에 의해 현탁액 배양 적응된 DG44 CHO 세포로 옮겼다.

[0331] 세포를 (문헌[Kaasbøll et al., 위 참조]에 기술된 바와 같이) 진탕기 플랫폼 상에서 8% CO2와 함께 37°C에서 유지되는 세포 배양 인큐베이터의 통기형 Erlenmeyer 플라스크에서 유지시켰다. 형질감염된 세포를 CD DG44 세포 배지(Gibco, Cat.# 12610-010)에서 밤새 유지한 후, HyClone™ ActiPro™ 배지(하이포잔틴 및 티미딘 불포함, GE Healthcare)로 옮기고, 생존율이 80%에 도달할 때까지 계대 배양하였고, 이 시점에서 배지에 0.1 µM의 메토평렉세이트를 보충하였다. 0.1 µM의 메토평렉세이트를 첨가한 후, 생존율이 다시 80%에 도달할 때까지 세포를 계대배양하였고, 이 시점에서 배지에 1 µM의 메토평렉세이트를 보충하였다. 생존율이 98%를 초과하고 배가 시간이 26시간 미만으로 감소할 때까지 세포를 다시 계대배양하였고, 이 시점에서 세포 풀은 안정적으로 형질감염된 것으로 간주되었다. 일단 안정적인 세포 풀이 확립되면, 세포 배양 부피를 확장하여 1*10^6개 세포/mL의 밀도의 생산을 위해 안정적으로 형질감염된 세포의 시딩을 할 수 있게 하였다. 세포 배양물을 계대 배양 후 3일째부터 매일 4/0.4% v/v HyClone™ Cell Boost™ 7a/7b로 보충하였다. 10일 후 세포를 4°C에서 20분 동안 4750xg에서 원심분리하여 침강시키고, 상청액 세포 배양 배지를 수확하였다. 100% 이소프로판올 중 0.1 M의 PMSF를 0.1 mM의 농도가 되도록 첨가하였다. 크로마토그래피 정제 전에 1 M의 시트르산나트륨 pH 5.5를 30 mM의 최종 농도가 되도록 첨가하였다.

- [0332] 실시예 15에 기술된 바와 같이, 단백질은 5 mL의 HiTrap™ MabSelectSuRe™ 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다. 정제된 단백질 제조물(단백질 분해 처리의 징후를 나타내지 않음)을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 9에서 알 수 있는 바와 같이, CHO 현탁액 세포의 안정적으로 형질감염된 플로부터 생산된 서열번호 41에 상응하는 정제된 단백질은 놀랍게도 AKT(세린 473)의 인산화를 억제하는 능력의 징후를 나타내지 않았다.
- [0333] 실시예 14
- [0334] 서열번호 80에 상응하는 단백질 서열을 초래하는, CCN3의 아미노산 206 내지 249(서열번호 44)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨), 펩티드 링커(서열번호 22)의 융합된 N 말단 및 IgG 서브타입 IgG2/4의 키메라 Fc 단편(서열번호 19)을 포함하는 융합 단백질을 생성하고, 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열을 첨부하여, 서열번호 81에 상응하는 융합 단백질을 생성하였다. 서열번호 81의 융합 단백질을 암호화하는 DNA 서열을 햄스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 82의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 83의 DNA 서열이 생성되었다.
- [0335] 서열번호 83의 서열은 상업적 공급업체에 의해 합성되고 서열 검증되었다. 합성된 서열은 엔트리 벡터를 생성하기 위하여 BP 게이트웨이 재조합효소 클로닝에 의해 pDonrZeo와 재조합되었다. 플라스미드의 효율적인 증식을 허용하도록 돌연변이된 적격(competent) 대장균(One Shot Top10™ 세포)의 형질감염 후, Qiagen™의 QIAprep™ Spin Miniprep 키트를 사용하여 표준 플라스미드 단리 기술로 엔트리 벡터를 단리하였다. 플라스미드 단리 후, 당업자에게 잘 알려진 표준 기술에 따라 제한 효소 소화에 이은 DNA 겔 전기영동에 의해 엔트리 벡터를 검증하였다.
- [0336] 서열번호 82의 서열을 함유하는 엔트리 벡터는 LR 게이트웨이 재조합효소를 사용하여 목적지 벡터와 추가로 재조합되었다. 사용된 목적지 벡터는 문헌[Kaasbøll et al., 2018, J. Biol. Chem., 293:46, pp. 17953 - 17970]에 의해 기술된 바와 같이 pUCOE-DHFR-DEST였다.
- [0337] 플라스미드의 효율적인 증식을 허용하도록 돌연변이된 적격 대장균(One Shot Top10™ 세포)의 형질감염 후, QIAGEN™ 플라스미드 플러스 맥시 키트를 사용하여 표준 플라스미드 단리 기술로 발현 벡터를 단리하였다. 당업자에게 잘 알려진 표준 기술에 따라 표준 제한 효소 소화 및 DNA 겔 전기영동에 의해 생성된 발현 벡터를 검증하였다. 그런 다음, 생성된 발현 벡터를 ExpiCHO™ 안정적인 생산 배지(Gibco Cat.#: A3711001)의 제조업체가 제공한 “Creation and Scale up of a Stable Cell Line Using ExpiCHO™ Products” 프로토콜에 따라 ExpiCHO 현탁액 배양 적응된 CHO 세포로 옮겼다. 세포를 (문헌[Kaasbøll et al., 위 참조]에 기술된 바와 같이) 진탕기 플랫폼 상에서 8% CO₂와 함께 37°C에서 유지되는 세포 배양 인큐베이터의 통기형 Erlenmeyer 플라스크에서 유지시켰다. 형질감염된 세포를 0.1 μM의 메토타렉세이트가 보충된 ExpiCHO™ 발현 배지로 옮기기 전에 ExpiCHO™ 발현 배지에서 밤새 유지시켰다. 그런 다음, 생존율이 다시 80%에 도달할 때까지 세포를 계대배양하였고, 이 시점에서 배지에 1 μM의 메토타렉세이트를 보충하였다. 생존율이 95%를 초과하고 배가 시간이 20시간 미만으로 감소할 때까지 세포를 다시 계대배양하였고, 이 시점에서 세포 풀은 안정적으로 형질감염된 것으로 간주되었다. 일단 안정적인 세포 풀이 확립되면, 세포 배양 부피를 확장하여 1*10⁶개 세포/mL의 밀도의 생산을 위해 안정적으로 형질감염된 세포의 시딩을 할 수 있게 하였다. 5일 후 세포를 4°C에서 20분 동안 4750xg에서 원심분리하여 침강시키고, 상청액 세포 배양 배지를 수확하였다. 100% 이소프로판올 중 0.1 M의 PMSF를 0.1 mM의 농도가 되도록 첨가하였다. 크로마토그래피 정제 전에 1 M의 시트르산나트륨 pH 5.5를 30 mM의 최종 농도가 되도록 첨가하고 2 M의 L-아르기닌 pH 4.0을 100 mM의 최종 농도가 되도록 첨가하였다.
- [0338] 완충액 A1, A2, A3 및 B1에 100 mM의 L-아르기닌을 첨가하는 것을 제외하고는 실시예 12에 기술된 것과 동일한 프로토콜을 이용하여, 단백질은 5 mL의 HiTrap™ MabSelect Prisma™ 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다. 정제된 단백질 제조물(단백질 분해 처리의 징후를 나타내지 않음)을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 10에서 알 수 있는 바와 같이, CHO 현탁액 세포의 안정적으로 형질감염된 플로부터 생산된 서열번호 80에 상응하는 정제된 단백질은 놀랍게도 AKT(세린 473)의 인산화를 억제하는 능력의 징후

를 나타내지 않았으며, 이는 CCN5로부터 유래된 아미노산(서열번호 41, 실시예 13, 도 9) 또는 CCN3로부터 유래된 아미노산(서열번호 80)을 함유하는 온전한 이량체 Fc-융합 단백질 제조물 중 어느 것도 생물학적으로 활성이 아님을 보여준다.

[0339] 실시예 15

[0340] 서열번호 80에 상응하는 단백질 서열을 초래하는 CCN3의 아미노산 206 내지 249(서열번호 44)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨), 펩티드 링커(서열번호 22)의 융합된 N 말단 및 IgG 서브타입 IgG2/4의 키메라 Fc 단편(서열번호 19)을 포함하며, 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열이 추가로 첨부되어 서열번호 81에 상응하는 융합 단백질을 암호화하는, 서열번호 82를 함유하는 실시예 14에 기술된 발현 플라스미드를 실시예 9에 기술된 바와 같이 ExpiCHO™ 세포의 일시적인 형질감염에 의해 발현시켰다. 형질감염 6일 후, 실시예 14에 기술된 바와 같이 세포를 침강시키고, 실시예 14에 기술된 바와 같이 배지를 보충하였다. 실시예 14에 기술된 바와 같이, 단백질은 5 mL의 HiTrap™ MabSelect Prisma™ 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다.

[0341] 정제된 단백질 제조물(부분적으로 단백질 분해 처리됨)을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 11에서 알 수 있는 바와 같이, 일시적으로 형질감염된 ExpiCHO™ 세포로부터 생산된 서열번호 80에 상응하는 정제된 단백질은 AKT(세린 473)의 인산화를 억제하는 농도 의존적 능력을 나타내어, 서열번호 80에 상응하는 융합 단백질을 생산하는 데 사용된 발현 시스템 및 결과적으로 관찰된 단백질 분해 처리의 정도가 생성된 단백질 제조물의 활성, 또는 이의 결여에 크게 영향을 미침을 보여준다.

[0342] 실시예 16

[0343] 서열번호 84에 상응하는 단백질 서열을 초래하는, CCN3의 아미노산 206 내지 249(서열번호 44)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 단량체 유도 및 반감기 연장 돌연변이를 갖는 Fc 단편(서열번호 55)을 포함하는 융합 단백질에 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열을 첨부하여, 서열번호 85에 상응하는 융합 단백질을 생성하였다. 서열번호 85의 융합 단백질을 암호화하는 DNA 서열을 햄스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 86의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 86의 DNA 서열이 생성되었다. DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드를 생성하기 위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었고, 실시예 9에 기술된 바와 같이 ExpiCHO™ 세포의 일시적인 형질감염에 의해 발현되었다. 형질감염 5일 후, 실시예 14에 기술된 바와 같이 세포를 침강시키고, 실시예 14에 기술된 바와 같이 배지를 보충하였다. 실시예 14에 기술된 바와 같이, 단백질은 5 mL의 HiTrap™ MabSelect Prisma™ 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다.

[0344] 예상한 단량체 형태를 나타낸 정제된 단백질 제조물을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 12에서 알 수 있는 바와 같이, 서열번호 84에 상응하는 정제된 단백질은 AKT(세린 473)의 인산화를 억제하는 농도 의존적 능력을 나타내어, CCN 단백질의 도메인 III/TSP-1 상동성 도메인으로부터의 아미노산을 포함하는 또 다른 단량체 융합 단백질이 A549 인간 폐암 세포에서 AKT(세린-473)의 인산화를 억제하는 능력을 가졌음을 보여주었다.

[0345] 실시예 17

[0346] 서열번호 88에 상응하는 단백질 서열을 초래하는, CCN3의 아미노산 206 내지 249(서열번호 44)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 단량체 유도 및 안정성 유도 돌연변이를 갖는 Fc 단편(서열번호 54)을 포함하는 융합 단백질에 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열을 첨부하여, 서열번호 89에 상응하는 융합 단백질을 생성하였다. 서열번호 89의 융합 단백질을 암호화하는 DNA 서열을 햄스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 90의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 91의 DNA 서열이 생성되었다. DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드를 생성하기

위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었고, 실시예 9에 기술된 바와 같이 ExpiCHO™ 세포의 일시적인 형질감염에 의해 발현되었다. 형질감염 6일 후, 실시예 14에 기술된 바와 같이 세포를 침강시키고, 실시예 14에 기술된 바와 같이 배지를 보충하였다. 실시예 14에 기술된 바와 같이, 단백질은 5 mL의 HiTrap™ MabSelect Prisma™ 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다.

[0347] 예상한 단량체 형태를 우세하게 나타낸 정제된 단백질 제조물을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 13에서 알 수 있는 바와 같이, 서열번호 88에 상응하는 정제된 단백질은 AKT(세린 473)의 인산화를 억제하는 농도 의존적 능력을 나타내어, CCN 단백질의 도메인 III/TSP-1 상동성 도메인으로부터의 아미노산을 포함하는 또 다른 단량체 융합 단백질이 A549 인간 폐암 세포에서 AKT(세린-473)의 인산화를 억제하는 능력을 가졌음을 보여주었다.

[0348] 실시예 18

[0349] 서열번호 94에 상응하는 단백질 서열을 초래하는, CCN3의 아미노산 206 내지 249(서열번호 44)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨), 펩티드 링커(서열번호 93)의 융합된 N 말단 및, 6xHis 태그, HaloTag 및 Sumo* 요소를 포함하는 다기능 태그(서열번호 92)를 포함하는 융합 단백질에 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열을 첨부하여, 서열번호 114에 상응하는 융합 단백질을 생성하였다. 서열번호 114의 융합 단백질을 암호화하는 DNA 서열을 햄스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 95의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 96의 DNA 서열이 생성되었다. DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드를 생성하기 위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었고, 실시예 15에 기술된 바와 같이 ExpiCHO™ 세포의 일시적인 형질감염에 의해 발현되었다. 형질감염 5일 후, 실시예 14에 기술된 바와 같이 세포를 침강시켰다. 크로마토그래피 정제 전에 100% 이소프로판올 중 0.1 M의 PMSF를 0.1 mM의 농도가 되도록 첨가하고, 1 M의 시트르산나트륨 pH 5.5를 30 mM의 최종 농도가 되도록 첨가하고, 2 M의 L-아르기닌 pH 4.0을 0.1 M의 최종 농도가 되도록 첨가하고, 이미다졸을 5 mM의 최종 농도가 되도록 첨가하였다.

[0350] 단백질은 5 mL의 HiTrap™ HisTrap™ excel 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다. 컬럼을 5 mm UV 플로우 셀이 장착된 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)에 장착하였다. HisTrap™ 컬럼을 제1 컬럼 전환 밸브에 장착하고 5 mM의 이미다졸, 50 mM의 NaCl, 100 mM의 L-아르기닌으로 구성된 A1 완충액으로 평형화시켰고, HiPrep™ 컬럼은 제2 컬럼 전환 밸브에 장착하고 완충액 A2(100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌 pH 6.5)로 평형화시켰다. HiPrep™ 컬럼을 함유하는 제2 컬럼 전환 밸브를 바이패스하도록 설정한 상태에서, 재조합 단백질을 함유하는 수확된 세포 배양 배지 250 mL를 샘플 펌프를 사용하여 3.5 ml/분의 속도로 HisTrap™ 컬럼에 로딩한 후, 5 컬럼 부피의 세척 완충액 A1로 세척한 다음, 5 컬럼 부피의 세척 완충액 A3(5 mM의 이미다졸, 0.5 M의 NaCl, 100 mM의 L-아르기닌)에 이어, 2 컬럼 부피의 세척 완충액 A1로 세척하였다. 용리 완충액(250 mM의 이미다졸, 50 mM의 NaCl, 100 mM의 L-아르기닌)으로 용리시키기 전에, 제2 컬럼 전환 밸브에 장착된 HiPrep™ 컬럼을 유동 경로로 들어가도록 설정하였다. 10 mL의 용리 완충액으로 용리시킨 후, HisTrap™ 컬럼을 유동 경로의 밖으로 전환시키고, 정제된 단백질을 완충액 A2를 사용하여 HiPrep™ 컬럼으로부터 용리시켰다. 단백질 용리를 280 nm에서의 UV 흡광도로 모니터링하였으며, 흡광도가 60 mAU를 초과하면 수집이 촉발되었다.

[0351] 예상한 단량체 형태를 나타낸 정제된 단백질 제조물을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 12에서 알 수 있는 바와 같이, 서열번호 94에 상응하는 정제된 단백질은 AKT(세린 473)의 인산화를 억제하는 농도 의존적 능력을 보여주어, CCN 단백질의 도메인 III/TSP-1 상동성 도메인으로부터의 아미노산을 포함하는 또 다른 단량체 융합 단백질이 A549 인간 폐암 세포에서 AKT(세린-473)의 인산화를 억제하는 능력을 가졌음을 보여주었다.

[0352] 실시예 19

[0353] 서열번호 97을 초래하는, CCN5의 아미노산 194 내지 237(여기서, 195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨)(서열번호 38), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 인간 혈청 알부민의 아미노산 25 내지 609(서열번호 52)를 포함하는 융합 단백질에 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호

서열을 첨부하여 서열번호 98에 상응하는 융합 단백질을 생성하였다. 서열번호 98의 융합 단백질을 암호화하는 DNA 서열을 햄스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 99의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 100의 DNA 서열이 생성되었다.

[0354] DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드를 생성하기 위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었고, AKT의 구조적 활성 형태를 발현하도록 조작된 DG44 CHO 현탁액 세포를 사용하여 실시예 13에 기술된 바와 같은 서열번호 98의 단백질을 발현하는 CHO 현탁액 세포의 안정적인 풀(pool)을 생성하였다. 일단 안정적인 세포 풀이 확립되면, 세포 배양 부피를 확장하여 1×10^6 개 세포/mL의 밀도의 생산을 위해 안정적으로 형질감염된 세포의 시딩을 할 수 있게 하였다. 6일 후 세포를 4°C에서 20분 동안 4750xg에서 원심분리하여 침강시키고, 상층액 세포 배양 배지를 수확하였다. 크로마토그래피 정제 전에 100% 이소프로판올 중 0.1 M의 PMSF를 0.1 mM의 농도가 되도록 첨가하고, 0.5 M의 EDTA를 2 mM의 최종 농도가 되도록 첨가하고, 1 M의 시트르산나트륨 pH 5.5를 30 mM의 최종 농도가 되도록 첨가하고, 2 M의 L-아르기닌 pH 4.0을 0.1 M의 최종 농도가 되도록 첨가하였다.

[0355] 단백질은 3 mL의 CaptureSelect™ 인간 알부민 친화도 매트릭스(ThermoFisherScientific)로 패키징된 Tricorn 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다. 컬럼을 5 mm UV 플로우 셀이 장착된 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)에 장착하였다. CaptureSelect™ 함유 컬럼을 제1 컬럼 전환 밸브에 장착하고 100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌 pH 6.5로 구성된 A1 완충액으로 평형화시켰고, HiPrep™ 컬럼은 제2 컬럼 전환 밸브에 장착하고 완충액 A1(100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌 pH 6.5)로 평형화시켰다. HiPrep™ 컬럼을 함유하는 제2 컬럼 전환 밸브를 바이패스하도록 설정한 상태에서, 재조합 단백질을 함유하는 수확된 세포 배양 배지 500 mL를 샘플 펌프를 사용하여 2.0 mL/분의 속도로 CaptureSelect™ 함유 컬럼에 로딩한 후, 5 컬럼 부피의 세척 완충액 A1로 세척한 다음, 5 컬럼 부피의 세척 완충액 A2(100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌, 0.25 M NaCl, pH 6.5)에 이어, 5 컬럼 부피의 세척 완충액 A1로 세척하였다. 용리 완충액(30 mM의 시트르산, pH 3.5 + 0.5 M의 L-아르기닌)으로 용리시키기 전에, 제2 컬럼 전환 밸브에 장착된 HiPrep™ 컬럼을 유동 경로로 들어가도록 설정하였다. 10 mL의 용리 완충액으로 용리시킨 후, CaptureSelect™ 함유 컬럼을 유동 경로의 밖으로 전환시키고, 정제된 단백질을 완충액 A1를 사용하여 HiPrep™ 컬럼으로부터 용리시켰다. 단백질 용리를 280 nm에서의 UV 흡광도로 모니터링하였으며, 흡광도가 100 mAU를 초과하면 수집이 촉발되었다.

[0356] 예상한 단량체 형태를 나타낸 정제된 단백질 제조물을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 14에서 알 수 있는 바와 같이, 서열번호 97에 상응하는 정제된 단백질은 AKT(세린 473)의 인산화를 억제하는 농도 의존적 능력을 나타내어, CCN 단백질의 도메인 III/TSP-1 상동성 도메인으로 부터의 아미노산을 포함하는 또 다른 단량체 융합 단백질이 A549 인간 폐암 세포에서 AKT(세린-473)의 인산화를 억제하는 능력을 가졌음을 보여주었다.

[0357] 실시예 20

[0358] 인간 혈청 알부민의 융합 단백질(아미노산 25 내지 606, 서열번호 101)을 인간 CCN5의 아미노산 194 내지 246 (여기서, 195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨)(서열번호 7)에 연결되는 펩티드 링커(서열번호 22)에 C 말단에 융합하여 서열번호 103을 생성하였다. 서열번호 102에 상응하는 융합 단백질에 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열을 첨부하여 서열번호 103에 상응하는 융합 단백질을 생성하였다. 서열번호 103의 융합 단백질을 암호화하는 DNA 서열을 햄스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 104의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 105의 DNA 서열이 생성되었다.

[0359] DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드를 생성하기 위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었고, AKT의 구조적 활성 형태를 발현하도록 조작된 DG44 CHO 현탁액 세포를 사용하여 실시예 13에 기술된 바와 같은 서열번호 104의 단백질을 발현하는 CHO 현탁액 세포의 안정적인 풀(pool)을 생성하였다. 일단 안정적인 세포 풀이 확립되면, 세포 배양 부피를 확장하여 1×10^6 개 세포/mL의 밀도의 생산을 위해 안정적으로 형질감염된 세포의 시딩을 할 수 있게 하였다. 6일 후 세포를 4°C에서 20분 동안 4750xg에서 원심분리

하여 침강시키고, 상청액 세포 배양 배지를 수확하였다. 크로마토그래피 정제 전에 100% 이소프로판올 중 0.1 M의 PMSF를 0.1 mM의 농도가 되도록 첨가하고, 0.5 M의 EDTA를 2 mM의 최종 농도가 되도록 첨가하고, 1 M의 시트르산나트륨 pH 5.5를 30 mM의 최종 농도가 되도록 첨가하고, 2 M의 L-아르기닌 pH 4.0을 0.1 M의 최종 농도가 되도록 첨가하였다.

[0360] 샘플 로딩 유속이 2.0 mL/분 대신 0.37mL/분인 것을 제외하고는 실시예 19에 기술된 바와 같이, 단백질은 3 mL의 CaptureSelect™ 인간 알부민 친화도 매트릭스(ThermoFisherScientific)로 패키징된 Tricorn 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다.

[0361] 예상한 단량체 형태를 나타낸 정제된 단백질 제조물을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 14에서 알 수 있는 바와 같이, 서열번호 102에 상응하는 정제된 단백질은 AKT(세린 473)의 인산화를 억제하는 농도 의존적 능력을 나타내어, CCN 단백질의 도메인 III/TSP-1 상동성 도메인으로부터의 아미노산을 포함하는 또 다른 단량체 융합 단백질이 A549 인간 폐암 세포에서 AKT(세린-473)의 인산화를 억제하는 능력을 가졌음을 보여주었다.

[0362] 실시예 21

[0363] 서열번호 106을 초래하는, CCN3의 아미노산 206 내지 249(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 인간 혈청 알부민의 아미노산 25 내지 609(서열번호 52)를 포함하는 융합 단백질에 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열을 첨부하여 서열번호 107에 상응하는 융합 단백질을 생성하였다. 서열번호 107의 융합 단백질을 암호화하는 DNA 서열을 랩스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 108의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 109의 DNA 서열이 생성되었다.

[0364] DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드를 생성하기 위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었고, 실시예 9에 기술된 바와 같이 ExpiCHO™ 세포의 일시적인 형질감염에 의해 발현되었다. 형질감염 6일 후, 실시예 14에 기술된 바와 같이 세포를 침강시키고, 실시예 19에 기술된 바와 같이 배지를 보충하였다. 단백질은 10 mL의 CaptureSelect™ 인간 알부민 친화도 매트릭스(ThermoFisherScientific)로 패키징된 Tricorn 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다. 컬럼을 5 mm UV 플로우 셀이 장착된 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)에 장착하였다. CaptureSelect™ 함유 컬럼을 제1 컬럼 전환 밸브에 장착하고 100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌 pH 6.5로 구성된 A1 완충액으로 평형화시켰고, HiPrep™ 컬럼은 제2 컬럼 전환 밸브에 장착하고 완충액 A1(100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌 pH 6.5)로 평형화시켰다. HiPrep™ 컬럼을 함유하는 제2 컬럼 전환 밸브를 바이패스하도록 설정한 상태에서, 재조합 단백질을 함유하는 수확된 세포 배양 배지 500 mL를 샘플 펌프를 사용하여 1.0 mL/분의 속도로 CaptureSelect™ 함유 컬럼에 로딩한 후, 3 컬럼 부피의 세척 완충액 A1로 세척한 다음, 2 컬럼 부피의 세척 완충액 A2(100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌, 0.25 M NaCl, pH 6.5)에 이어, 3 컬럼 부피의 세척 완충액 A1로 세척하였다. 용리 완충액(30 mM의 시트르산, pH 3.5 + 0.1 M의 L-아르기닌)으로 용리시키기 전에, 제2 컬럼 전환 밸브에 장착된 HiPrep™ 컬럼을 유동 경로로 들어가도록 설정하였다. 15 mL의 용리 완충액으로 용리시킨 후, CaptureSelect™ 함유 컬럼을 유동 경로의 밖으로 전환시키고, 정제된 단백질을 완충액 A1를 사용하여 HiPrep™ 컬럼으로부터 용리시켰다. 단백질 용리를 280 nm에서의 UV 흡광도로 모니터링하였으며, 흡광도가 100 mAU를 초과하면 수집이 촉발되었다.

[0365] 예상한 단량체 형태를 함유한 정제된 단백질 제조물을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 12에서 알 수 있는 바와 같이, 서열번호 106에 상응하는 정제된 단백질은 AKT(세린 473)의 인산화를 억제하는 농도 의존적 능력을 나타내어, CCN 단백질의 도메인 III/TSP-1 상동성 도메인으로부터의 아미노산을 포함하는 또 다른 단량체 융합 단백질이 A549 인간 폐암 세포에서 AKT(세린-473)의 인산화를 억제하는 능력을 가졌음을 보여주었다.

[0366] 실시예 22

[0367] 서열번호 110을 초래하는, CCN3의 아미노산 206 내지 249(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로

로 치환됨)(서열번호 44), 펩티드 링커(서열번호 22)의 융합된 N 말단 및 인간 혈청 알부민의 아미노산 25 내지 609(서열번호 52)를 포함하는 융합 단백질에 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열을 첨부하여 서열번호 111에 상응하는 융합 단백질을 생성하였다. 서열번호 111의 융합 단백질을 암호화하는 DNA 서열을 햄스터 세포에서의 단백질 발현을 위하여 코돈 최적화하였고(상업적인 공급업체의 알고리즘에 의해), 번역을 위한 KOZAK 서열을 5' 말단에 첨부하고, 3' 말단에 정지 코돈을 도입하여 서열번호 112의 DNA 서열을 생성하였다. DNA 서열은 게이트웨이 attB 부위가 양 말단에 추가로 첨부되어 서열번호 113의 DNA 서열이 생성되었다.

[0368] DNA 서열은 실시예 9에 기술된 바와 같은 플라스미드를 생성하기 위해 서브클로닝되기 전에 상업적 공급업체에 의해 합성 및 검증되었고, 실시예 9에 기술된 바와 같이 ExpiCHO™ 세포의 일시적인 형질감염에 의해 발현되었다. 형질감염 6일 후, 실시예 14에 기술된 바와 같이 세포를 침강시키고, 실시예 19에 기술된 바와 같이 배지를 보충하였다. 단백질은 10 mL의 CaptureSelect™ 인간 알부민 친화도 매트릭스(ThermoFisherScientific)로 패키징된 Tricorn 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 53 mL의 HiPrep™ 26/10 탈염 컬럼(GE Healthcare)을 이용한 탈염으로 구성된 탠덤 크로마토그래피에 의해 정제되었다. 컬럼을 5 mm UV 플로우 셀이 장착된 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)에 장착하였다. CaptureSelect™ 함유 컬럼을 제1 컬럼 전환 밸브에 장착하고 100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌 pH 6.5로 구성된 A1 완충액으로 평형화시켰고, HiPrep™ 컬럼은 제2 컬럼 전환 밸브에 장착하고 완충액 A1(100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌 pH 6.5)로 평형화시켰다. HiPrep™ 컬럼을 함유하는 제2 컬럼 전환 밸브를 바이패스하도록 설정한 상태에서, 제조된 단백질을 함유하는 수확된 세포 배양 배지 300 mL를 샘플 펌프를 사용하여 1.0 mL/분의 속도로 CaptureSelect™ 함유 컬럼에 로딩한 후, 3 컬럼 부피의 세척 완충액 A1로 세척한 다음, 2 컬럼 부피의 세척 완충액 A2(100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌, 0.25 M NaCl, pH 6.5)에 이어, 3 컬럼 부피의 세척 완충액 A1로 세척하였다. 용리 완충액(30 mM의 시트르산, pH 3.5 + 0.5 M의 L-아르기닌)으로 용리시키기 전에, 제2 컬럼 전환 밸브에 장착된 HiPrep™ 컬럼을 유동 경로로 들어가도록 설정하였다. 15 mL의 용리 완충액으로 용리시킨 후, CaptureSelect™ 함유 컬럼을 유동 경로의 밖으로 전환시키고, 정제된 단백질을 완충액 A1를 사용하여 HiPrep™ 컬럼으로부터 용리시켰다. 단백질 용리를 280 nm에서의 UV 흡광도로 모니터링하였으며, 흡광도가 100 mAU를 초과하면 수집이 촉발되었다.

[0369] 예상한 단량체 형태를 함유한 정제된 단백질 제조물을 실시예 2에 기술된 바와 같이 A549 인간 폐암 세포에서 생존 촉진 신호전달(AKT의 세린-473 인산화)을 억제하는 능력에 대해 후속적으로 테스트하였다. 도 15에서 알 수 있는 바와 같이, 서열번호 110에 상응하는 정제된 단백질은 AKT(세린 473)의 인산화를 억제하는 농도 의존적 능력을 나타내어, CCN 단백질의 도메인 III/TSP-1 상동성 도메인으로부터의 아미노산을 포함하는 또 다른 단량체 융합 단백질이 A549 인간 폐암 세포에서 AKT(세린-473)의 인산화를 억제하는 능력을 가졌음을 보여주었다.

[0370] 실시예 23

[0371] CCN3의 아미노산 206 내지 249(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 인간 혈청 알부민의 아미노산 25 내지 609(서열번호 52)를 포함하며, 서열번호 32의 알부민으로부터 유래하는 분비를 위한 N 말단 신호 서열이 추가로 첨부되어 서열번호 107에 상응하는 융합 단백질을 암호화하는, 서열번호 108을 함유하는 실시예 21에 기술된 발현 플라스미드를 사용하여 실시예 14에 기술된 바와 같은 안정적으로 형질감염된 ExpiCHO™ 세포의 풀을 생성하였다. 서열번호 106에 상응하는 분비된 단백질을 함유하는 조정 배지의 배치를 생산하기 위하여, 안정적으로 형질감염된 세포의 풀을 확장하여 1*10⁶개 세포/mL의 밀도의 안정적으로 형질감염된 세포 250 mL의 부피를 시딩할 수 있도록 하였다. 계대 배양 후 2일째부터 시작하여 이틀마다 세포 배양물에 5% v/v의 2X EfficientFeed™ C+(Gibco™)를 보충하였고, 계대 배양 후 2일째에 3%의 글루코스(10% w/v)를 보충하였고, 계대 배양 후 6일째에 5%의 글루코스(10% w/v)를 보충하였다. 9일 후 세포를 4℃에서 20분 동안 4750xg에서 원심분리하여 침강시키고, 상청액 세포 배양 배지를 수확하였다. 100% 이소프로판올 중 0.1 M의 PMSF를 0.1 mM의 농도가 되도록 첨가하였다. 크로마토그래피 정제 전에 1 M의 시트르산나트륨 pH 5.5를 30 mM의 최종 농도가 되도록 첨가하고 2 M의 L-아르기닌 pH 4.0을 100 mM의 최종 농도가 되도록 첨가하였다.

[0372] 단백질은 10 mL의 CaptureSelect™ 인간 알부민 친화도 매트릭스(ThermoFisherScientific)로 패키징된 Tricorn 컬럼(GE Healthcare)을 사용한 포획 단계에 바로 이어지는 2개의 병렬 연결된 Superdex 200 Increase 10/300 GL(GE Healthcare) 컬럼을 이용한 크기 배제 크로마토그래피로 구성된 2D-크로마토그래피에 의해 정제되었다. 컬럼을 5 mL의 샘플 루프에 연결된 유출 밸브 및 5 mm UV 플로우 셀이 장착된 FPLC 크로마토그래피 시스템(BioRad NGC Discover™ 10 Pro 시스템)에 장착하였다. CaptureSelect™ 함유 컬럼을 제1 컬럼 전환 밸브에 장

착하고 100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌 pH 6.5로 구성된 A1 완충액으로 평형화시켰고, Superdex 200 Increase 컬럼은 제2 컬럼 전환 밸브에 장착하고 완충액 A1(100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌 pH 6.5)로 평형화시켰다. Superdex 200 Increase 컬럼을 함유하는 제2 컬럼 전환 밸브를 바이패스하도록 설정한 상태에서, 재조합 단백질을 함유하는 수확된 세포 배양 배지 120 mL를 샘플 펌프를 사용하여 3.9 mL/분의 속도로 CaptureSelect™ 함유 컬럼에 로딩하였다. CaptureSelect™ 함유 컬럼에 재조합 단백질을 함유하는 수확된 세포 배양 배지를 로딩한 후, 3 컬럼 부피의 완충액 A1로 세척한 다음, 2 컬럼 부피의 완충액 A2(100 mM NaH₂PO₄/Na₂HPO₄, 100 mM L-아르기닌, 0.25 M NaCl, pH 6.5)에 이어, 3 컬럼 부피의 완충액 A1로 세척하였다. CaptureSelect™ 함유 컬럼을 15 mL 완충액 B1(30 mM의 시트르산, 0.5 M의 L-아르기닌, pH 3.5)으로 용리시켰고, 이때 시스템은 흡광도가 1200 mAU를 초과하는 용리액을 샘플 루프로 수집하도록 설정되었다. CaptureSelect™ 함유 컬럼의 용리 후, 제1 컬럼 전환 밸브에 연결된 CaptureSelect™ 함유 컬럼을 유동 경로의 밖으로 전환시키고, 제2 컬럼 전환 밸브를 Superdex 200 Increase 함유 컬럼을 유동 경로로 전환하도록 설정하였다. 그런 다음, 용리된 단백질을 함유하는 CaptureSelect™ 함유 컬럼의 용리액을 0.5 mL/분의 속도로 완충액 A1이 함유된 Superdex 200 Increase를 함유하는 컬럼에 로딩하였다. 단백질 용리를 280 nm에서의 UV 흡광도로 모니터링하였으며, 흡광도가 200 mAU를 초과하면 수집이 촉발되었다.

[0373] 예상한 단량체 형태를 함유한 정제된 단백질 제조물을 후속적으로 정상 인간 폐 섬유모세포(NHFL)(Lonza Bioscience, Catalog #: CC-2512)의 TGF β 유도 및 활성 CCN2 유도 활성을 억제하는 능력에 대해 테스트하였다. NHFL를 상업적 공급업체(Lonza Bioscience)의 설명서에 따라 모든 첨가제(2%의 소 태아 혈청, 인슐린, hFGF-B, 겐타마이신/암포테리신-B)를 함유한 완전 성장 배지(Lonza Bioscience bullet kit (Cat#: CC-3132))에 계대 배양하여 최대 80%의 함유도의 밀도를 유지하였다. 활성-CCN2는 CCN2의 도메인 3 내지 4로 구성되었고, 문헌[Kaasbøll et al., 2018, 위 참조]에 의해 기술된 바와 같이 제조 및 정제되었다.

[0374] 서열번호 106에 상응하는 단백질이 NHFL의 활성-CCN2 및 TGF β -유도된 세포 이동에 미치는 영향을 테스트하기 위하여(트랜스웰 분석/변형된 보이덴 챔버 분석), 세포를 먼저 트립신/EDTA로 분리하고, 트립신 중화 시약(Lonza Bioscience, Cat.# CC-5034)으로 중화시키고, 기초 성장 배지(섬유모세포 기초 배지(LonzaBioscience Cat.#: CC-3131, 겐수마이신(50 μ g/mL) 외의 다른 첨가제 없음))에 재현탁시킨 후, 5 μ m의 기공 크기의 트랜스웰 인서트(24-웰 플레이트, Corning® Transwell®, Cat. # CLS3402-48EA, SigmaAldrich (Merck KGaA)) 상단에 있는 웰당 100 μ L의 부피 중 30000개의 세포를 시딩하였다. 웰의 하부 챔버는 겐수마이신 이외의 다른 첨가제가 없는 500 μ L의 기초 성장 배지에 용해된 테스트 물질 또는 비히클 대조군을 함유하였다. 20시간의 인큐베이션 후, 인서트를 웰에서 제거하고, 인산염 완충 식염수(PBS, Lonza Bioscience, Cat. #: 17-512F)에 담가 2회 세척한 후, 4% 포르말데히드(Solvec, Swe., Cat.# : 621092)에 37°C에서 15분 동안 고정시켰다. PBS로 2회 세척하기 전에 PBS 중 0.1%의 Triton X-100으로 10분 동안 처리하여 세포를 투과화시켰다. 막이 건조되기 전에 면봉으로 긁어서 인서트의 상단에 있는 이동하지 않은 세포를 제거하였다. 인서트의 아래쪽에 이동된 세포의 핵을 암소에서 15분 동안 Hoechst 33342 20 mM(PBS에 1:5000 희석, ThermoFisherScientific, Cat. #: 62249)로 염색한 후, PBS에 담가 2회 세척하였다. 막을 트랜스웰 인서트에서 잘라낸 후, 유리 쪽으로 이동한 세포가 있는 유리 슬라이드 위에 얹고, 1방울의 ProLong™ Gold Antifade(ThermoFisherScientific, Cat. #: P36934)를 바르고, 유리 커버슬립을 얹고, 각 웰의 5~10개의 이미지를 Zeiss Axio Observer Z.1 영상화 시스템에서 캡처하였다. ImageJ 소프트웨어 v1.51k(Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, <https://imagej.nih.gov/ij/>, 1997-2018.)를 사용하여 이미지를 반자동으로 분석하였다. 도 16a에서 볼 수 있는 바와 같이, 서열번호 106에 상응하는 단백질은 TGF β 및 활성-CCN2 둘 다에 의해 유도된 이동을 억제한다.

[0375] 활성-CCN2 및 TGF β -유도된 스크래치 상처 분석에 미치는 서열번호 106에 상응하는 단백질의 영향을 테스트하기 위하여, NHFL를 트립신/EDTA로 분리하고, 트립신 중화 시약(Lonza Bioscience, Cat.# CC-5034)으로 중화한 후, 조직 배양 처리된 12-웰 플레이트(Corning Costar®, Cat. # 3513)에서 1 mL의 부피 중 100000개의 세포를 시딩하였다. 시딩 다음 날, 세포를 0.9%의 NaCl로 2회 세척하고, 완전 성장 배지를 기초 성장 배지로 교체하였다. 기초 성장 배지에서 16 내지 20시간 동안 인큐베이션한 후, 평균 12.5 μ L의 피펫 팁(ThermoFisherScientific, Cat#: 94420053)으로 세포 단층에 스크래치를 만들고, 세포를 PBS로 1회 세척한 후, 세포를 테스트 물질 또는 비히클과 함께 1 mL의 기초 성장 배지에서 인큐베이션하였다. 세포를 추가 24시간 동안 인큐베이션한 후, PBS에서 3회 세척한 후, 4% 포르말데히드에서 37°C에서 15분 동안 고정시켰다. 고정 후, 세포를 부드럽게 진탕하면서 PBS에서 3x3분 동안 다시 세척하고, 부드럽게 진탕하면서 PBS 중 0.1%의 Triton X-100으로 10분 동안 투과화시켰다. 세포의 핵을 암소에서 15분 동안 Hoechst 33342 20 mM(PBS에 1:5000 희석, ThermoFisherScientific, Cat. #: 62249)로 염색한 후, PBS에서 부드럽게 진탕하면서 3x5분 세척하였다. ProLong™ Gold

Antifade(ThermoFisherScientific, Cat. #: P36934) 1방울을 장착 전에 적용하고, 나머지 간극을 중심으로 하는 5개의 이미지를 Zeiss Axio Observer Z.1 영상화 시스템을 사용하여 각 웰에서 캡처하였다. 스크래치 상처의 길이를 따라 3개의 고정된 간격으로 스크래치 후 남은 간극을 측정하여 이미지를 분석하였다. 각 웰의 모든 이미지로부터의 모든 측정값의 평균을 계산하고, 하나의 생물학적 복제물로 계수하였다. 도 16b에서 볼 수 있는 바와 같이, 서열번호 106에 반응하는 단백질은 TGF β 및 활성-CCN2 둘 다에 의해 유도된 스크래치 상처의 봉합을 억제한다.

[0376]

TGF β -유도된 유전자 조절에 미치는 서열번호 106에 반응하는 단백질의 영향을 테스트하기 위하여, NHLF를 트립신/EDTA로 분리하고, 트립신 중화 시약(Lonza Bioscience, Cat.# CC-5034)으로 중화한 후, 조직 배양 처리된 12-웰 플레이트(Corning Costar®, Cat. # 3513)에서 1 mL의 부피 중 100000개의 세포를 시딩하였다. 시딩 다음 날, 세포를 0.9%의 NaCl로 2회 세척하고, 완전 성장 배지를 0.1%의 열-불활성화 소 태아 혈청(Cat. # 16000-044, Gibco™, 열-불활성화는 실시예 2에 기술된 바와 같이 수행됨)이 보충된 기초 성장 배지로 교체하였다. 6 시간 동안 0.1% 소 태아 혈청을 함유하는 기초 성장 배지에서 인큐베이션한 후, 테스트 물질 또는 비히클 대조군을 웰에 첨가하였다. 96시간 후 웰을 PBS에서 2회 세척하고, 제조업체의 프로토콜에 따라 Qiagen RNeasy RNA 추출 키트(Cat. # 74106)를 사용하여 RNA를 추출하였다. RNA 농도는 NanoDrop® ND-1000 분광광도계(NanoDrop Technologies, 미국)로 정량화하고, 뉴클레아제가 없는 물로 희석하여 최종 RNA 농도가 50 ng/ μ L이 되도록 한 후, 제조업체의 프로토콜에 따라 각 복제물의 200 ng의 RNA를 사용하여 TaqMan™ 역전사 키트(Cat. # N8080234)를 활용하여 cDNA를 생성하였다. 각각의 TaqMan™ 분석 및 TaqMan Fast Advanced Master Mix(ThermoFisherScientific Cat. # 4444557)를 사용하여 생성된 cDNA 샘플에서 차등 유전자 발현 분석을 분석하였다. TaqMan™ 실시간 PCR 반응은 제조사의 프로토콜에 따라 Applied Biosystems StepOnePlus 실시간 PCR 시스템을 사용하여 각 샘플에 대해 기술적 3중 반복으로 실행되었다. 각 샘플에서 단일 값을 산출하기 위하여 기술적 3중 반복의 평균을 내기 전에, 상이한 전사체들의 상대적인 양을 표준 곡선에서 계산하였다. 모든 유전자 발현 결과는 GAPDH(ThermoFisherScientific, Cat. # Hs02786624_g1) mRNA 수준과 관련되었으며, 비히클 대조군-자극된 웰의 평균의 배수로 표현되도록 정규화되었다. 도 19c 내지 f에서 알 수 있는 바와 같이, 서열번호 107에 반응하는 단백질은 일반적으로 전섬유증 유전자인 것으로 간주되는 TGF β -유도 유전자 COL1A1("콜라겐 유형 1 α -1", ThermoFisherScientific, Cat. # Hs00164004_m1), FN1("피브로넥틴 1", ThermoFisherScientific, Cat. # Hs01549976_m1), ACTA2("평활근 액틴 α -2", ThermoFisherScientific, Cat. # Hs00426835_g1) 및 CCN2(ThermoFisherScientific, Cat. # Hs00170014_m1)의 부분적 억제를 제공한다.

[0377] 명세서 및 서열목록에 언급된 서열번호의 개요

서열번호	서열 정보
1	인간 WISP2/CCN5의 아미노산 194 ~ 246(도메인 III/TSP-1 상동성 도메인)(긴 단편)
2	인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편)
3	인간 CTGF/CCN2의 아미노산 199 ~ 242(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편)
4	인간 Cyr61/CCN1의 아미노산 229 ~ 272(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편)
5	인간 WISP1/CCN4의 아미노산 216 ~ 259(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편)
6	인간 WISP3/CCN6의 아미노산 209 ~ 252(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편)
7	인간 WISP2/CCN5의 아미노산 194 ~ 246, (긴 단편) 여기서 195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨.
8	인간 WISP3/CCN6의 아미노산 209 ~ 263(도메인 III/TSP-1 상동성 도메인)(긴 단편)
9	인간 NOV/CCN3의 아미노산 206 ~ 258(도메인 III/TSP-1 상동성 도메인)(긴 단편)
10	인간 CTGF/CCN2의 아미노산 199 ~ 250(도메인 III/TSP-1 상동성 도메인)(긴 단편)
11	인간 Cyr61/CCN1의 아미노산 229 ~ 280(도메인 III/TSP-1 상동성 도메인)(긴 단편)
12	인간 WISP1/CCN4의 아미노산 216 ~ 267(도메인 III/TSP-1 상동성 도메인)(긴 단편)
13	IgG4의 Fc 단편
14	IgG2의 Fc 단편

[0378]

15	IgG4의 돌연변이된 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링 참조)
16	IgG1의 비 글리코실화된(N297G), 이황화 가교 안정화된(R292C, V302C) Fc 단편, Eu 넘버링
17	이황화 가교 안정화(R292C, V302C), 비 글리코실화(N297G) 및 하부 힌지 영역에 단백질 분해효소 안정화 돌연변이(E233P, L234V, L235A, G236-, Eu 넘버링)가 있는, IgG1의 Fc 단편
18	IgG4의 Fc 단편으로, 하부 힌지 영역에 단백질 분해효소 안정화 돌연변이(E233P, F234V, L235A, G236-, Eu-넘버링) 및 S228P 돌연변이 및 K447- 결실이 있음.
19	카르복실 말단의 K477-(Eu 넘버링)이 결실된, IgG2의 힌지 영역 및 IgG4의 불변 중쇄 도메인 2 및 3의 키메라인 Fc 단편.
20	링커(TEGRMD)
21	링커(EAAAK)
22	링커(EAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAK) n=8
23	링커(TAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAK)
24	링커(EAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKAAA)
25	링커(TAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKAAA)
26	서열번호 19의 IgG2/4 Fc 단편에 대해 N 말단에 추가로 연결되는 서열번호 21의 링커의 N 말단에 연결된, CCN5의 도메인 III(서열번호 37)을 포함하는 본 발명의 융합 단백질.
27	서열번호 19의 IgG2/4 Fc 단편에 대해 N 말단에 추가로 연결되는 서열번호 21의 링커의 N 말단에 연결된, 195번 위치의 프롤린이 알라닌으로 교체된, CCN5의 도메인 III(서열번호 38)을 포함하는 본 발명의 융합 단백질.
28	CCN5의 아미노산 194-246(서열번호 1), 펩티드 링커(서열번호 20)의 융합된 N 말단 및 서열번호 15의 IgG, 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 본 발명의 융합 단백질(CCN5(dIII)-Fc _v 2)
29	CCN5의 아미노산 194-246(서열번호 1), 펩티드 링커(서열번호 20)의 융합된 N 말단 및 서열번호 18의 IgG 서브타입 IgG4의 Fc 단편(S228P, E233P, F234V, L235A, G236-, K447-, Eu-넘버링)을 포함하는 융합 단백질(CCN5(dIII)-Fc _v 2.1)

[0379]

30	CCN5의 아미노산 194~246(서열번호 1), 펩티드 링커(서열번호 20)의 융합된 N 말단 및 IgG 서브타입 IgG2/4 서브타입의 키메라 Fc 단편(서열번호 19)을 포함하는 융합 단백질(CCN5(dIII)- Fc _{v2.3}).
31	CCN5의 아미노산 194~246(서열번호 1) 링커(서열번호 25)의 융합된 N 말단 및 IgG 서브타입 IgG2/4 서브타입의 키메라 Fc 단편(서열번호 19)을 포함하는 융합 단백질(CCN5(dIII)- HLn8-Fc _{v2.3}).
32	인간 혈청 알부민으로부터의 신호 펩티드(MKWVTFISLLFLFSSAYS)
33	서열번호 28의 융합 단백질로서, 인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 융합 단백질의 N 말단에 첨부된 융합 단백질.
34	서열번호 33의 융합 단백질을 암호화하는 DNA 서열
35	서열번호 33의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.
36	서열번호 35의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 DNA 서열.
37	인간 WISP2/CCN5의 아미노산 194 ~ 237(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편)
38	인간 WISP2/CCN5의 아미노산 194 ~ 237(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편), 여기서 195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨.
39	링커(AEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKAAA)
40	CCN5의 아미노산 194~246(서열번호 1), 펩티드 링커(서열번호 39)의 융합된 C 말단 및 서열번호 15의 IgG 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질(Fc-HLn8-CCN5(dIII)).
41	CCN5의 아미노산 194~246(서열번호 7)(195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨), 펩티드 링커(서열번호 39)의 융합된 C 말단 및 서열번호 15의 IgG 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질(Fc-HLn8-CCN5(dIII)-P195A).

[0380]

42	인간 Cyr61/CCN1의 아미노산 229 ~ 272(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편), 여기서 230번 위치의 아미노산(이소류신)이 알라닌으로 치환됨.
43	인간 CTGF/CCN2의 아미노산 199 ~ 242(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편), 여기서 200번 위치의 아미노산(류신)이 알라닌으로 치환됨.
44	인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편), 여기서 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨.
45	인간 WISP1/CCN4의 아미노산 216 ~ 259(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편), 여기서 217번 위치의 아미노산(이소류신)이 알라닌으로 치환됨.
46	인간 WISP3/CCN6의 아미노산 209 ~ 252(도메인 III/TSP-1 상동성 도메인)(44개 aa 단편), 여기서 210번 위치의 아미노산(류신)이 알라닌으로 치환됨.
47	인간 Cyr61/CCN1의 아미노산 229 ~ 280(도메인 III/TSP-1 상동성 도메인)(긴 단편), 여기서 230번 위치의 아미노산(이소류신)이 알라닌으로 치환됨.
48	인간 CTGF/CCN2의 아미노산 199 ~ 250(도메인 III/TSP-1 상동성 도메인)(긴 단편), 여기서 200번 위치의 아미노산(류신)이 알라닌으로 치환됨.
49	인간 NOV/CCN3의 아미노산 206 ~ 258(도메인 III/TSP-1 상동성 도메인)(긴 단편), 여기서 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨.
50	인간 WISP1/CCN4의 아미노산 216 ~ 267(도메인 III/TSP-1 상동성 도메인)(긴 단편), 여기서 217번 위치의 아미노산(이소류신)이 알라닌으로 치환됨.
51	인간 WISP3/CCN6의 아미노산 209 ~ 263(도메인 III/TSP-1 상동성 도메인)(긴 단편), 여기서 210번 위치의 아미노산(류신)이 알라닌으로 치환됨.
52	인간 혈청 알부민의 아미노산 25-609
53	인간 세로트랜스페린의 아미노산 20-698

[0381]

54	이황화 가교 안정화(R292C, V302C), 비 글리코실화(N297G) 및 단량체 생성 돌연변이(C220Q, C226Q, C229Q, T366R, L368H, P395K, K409T, M428L), Eu 넘버링)가 있는, IgG1의 Fc 단편
55	IgG2의 힌지 영역과 IgG4의 불변 중쇄 도메인 2 및 3의 키메라이고, 카르복실 말단 K477이 결실되고, 단량체 생성 돌연변이(C219Q, C220Q, C226Q, C229Q, L351F, T366R, P395K, F405R, Y407E)가 있고, 반감기 연장 돌연변이(M252Y, S254T, T256E)(Eu 넘버링)가 있는 Fc 단편.
56	인간 WISP2/CCN5의 아미노산 194 ~ 250(도메인 III/TSP-1 상동성 도메인)
57	링커(IEGRMD)
58	CCN5의 아미노산 194~250(서열번호 56), 펩티드 링커(서열번호 57)의 융합된 N 말단 및 서열번호 15의 IgG 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질.
59	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, CCN5의 아미노산 194~250(서열번호 56), 펩티드 링커(서열번호 57)의 융합된 N 말단 및 서열번호 15의 IgG 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질
60	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 59의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.
61	서열번호 59의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 60의 DNA 서열.
62	인간 WISP2/CCN5의 아미노산 194 ~ 249(도메인 III/TSP-1 상동성 도메인)(긴 단편), 여기서 245번 위치의 아미노산(프롤린)이 류신으로 치환됨.
63	링커(GRMD)
64	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 WISP2/CCN5의 아미노산 194 ~ 249(도메인 III/TSP-1 상동성 도메인), 여기서 245번 위치의 아미노산(프롤린)이 류신으로 치환됨(서열번호 62), 펩티드 링커(서열번호 63)의 융합된 N 말단 및 서열번호 15의 IgG 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)

[0382]

65	링커(TEGRMD)
66	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 WISP2/CCN5의 아미노산 194 ~ 246(도메인 III/TSP-1 상동성 도메인)(서열번호 1), 펩티드 링커(서열번호 65)의 융합된 N 말단 및 서열번호 15의 IgG 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)
67	링커(TAEAAAKA)
68	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 WISP2/CCN5의 아미노산 194 ~ 246(도메인 III/TSP-1 상동성 도메인)(서열번호 1), 펩티드 링커(서열번호 67)의 융합된 N 말단 및 서열번호 15의 IgG 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)
69	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 WISP2/CCN5의 아미노산 194 ~ 246(도메인 III/TSP-1 상동성 도메인)(서열번호 1), 펩티드 링커(서열번호 65)의 융합된 N 말단 및 서열번호 19의 카르복실 말단의 K477-(Eu 넘버링)이 결실된, IgG2의 힌지 영역 및 IgG4의 불변 중쇄 도메인 2 및 3으로 구성된 키메라 Fc 단편.
70	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, (서열번호 19의 IgG2/4 Fc 단편에 대해 N 말단에 추가로 연결되는 서열번호 21의 링커의 N 말단에 연결된, 195번 위치의 프롤린이 알라닌으로 치환된, CCN5의 도메인 III의 아미노산 194~237(서열번호 38)을 포함하는) 서열번호 27의 융합 단백질.
71	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 70의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.
72	서열번호 70의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 71의 DNA 서열.
73	인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 서열번호 19의 카르복실 말단의 K477-(Eu 넘버링)이 결실된, IgG2의 힌지 영역 및 IgG4의 불변 중쇄 도메인 2 및 3으로 구성된 키메라 Fc 단편.

[0383]

74	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 서열번호 19의 카르복실 말단의 K477-(Eu 넘버링)이 결실된, IgG2의 힌지 영역 및 IgG4의 불변 중쇄 도메인 2 및 3으로 구성된 키메라 Fc 단편을 포함하는 융합 단백질(서열번호 73)
75	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 74의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.
76	서열번호 74의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 75의 DNA 서열.
77	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, CCN5의 아미노산 194-246(서열번호 7)(195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨), 펩티드 링커(서열번호 30)의 융합된 C 말단 및 서열번호 15의 IgG 서브타입 IgG4의 Fc 단편(S228P, F234A, L235A, K447-, Eu-넘버링)을 포함하는 융합 단백질(Fc-HLn8-CCN5(dIII)-P195A)(서열번호 41)
78	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 77의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.
79	서열번호 77의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 78의 DNA 서열.
80	인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 22)의 융합된 N 말단 및 서열번호 19의 카르복실 말단의 K477-(Eu 넘버링)이 결실된, IgG2의 힌지 영역 및 IgG4의 불변 중쇄 도메인 2 및 3으로 구성된 키메라 Fc 단편.

[0384]

81	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 22)의 융합된 N 말단 및 서열번호 19의 카르복실 말단의 K477-(Eu 넘버링)이 결실된, IgG2의 힌지 영역 및 IgG4의 불변 중쇄 도메인 2 및 3으로 구성된 키메라 Fc 단편을 포함하는 융합 단백질(서열번호 81)
82	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 81의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.
83	서열번호 81의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 82의 DNA 서열.
84	인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 서열번호 55의 IgG2의 힌지 영역과 IgG4의 불변 중쇄 도메인 2 및 3의 키메라이고, 카르복실 말단 K477이 결실되고, 단량체 생성 돌연변이(C219Q, C220Q, C226Q, C229Q, L351F, T366R, P395K, F405R, Y407E) 및 반감기 연장 돌연변이(M252Y, S254T, T256E)(Eu 넘버링)가 있는 Fc 단편을 포함하는 융합 단백질
85	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 서열번호 55의 IgG2의 힌지 영역과 IgG4의 불변 중쇄 도메인 2 및 3의 키메라이고, 카르복실 말단 K477이 결실되고, 단량체 생성 돌연변이(C219Q, C220Q, C226Q, C229Q, L351F, T366R, P395K, F405R, Y407E) 및 반감기 연장 돌연변이(M252Y, S254T, T256E)(Eu 넘버링)가 있는 Fc 단편을 포함하는 융합 단백질(서열번호 84)
86	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 85의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.

[0385]

87	서열번호 85의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 86의 DNA 서열.
88	인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 서열번호 54의 이황화 가교 안정화(R292C, V302C), 비 글리코실화(N297G) 및 단량체 생성 돌연변이(C220Q, C226Q, C229Q, T366R, L368H, P395K, K409T, M428L), Eu 넘버링)가 있는, IgG1의 Fc 단편을 포함하는 융합 단백질
89	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 21)의 융합된 N 말단 및 서열번호 54의 이황화 가교 안정화(R292C, V302C), 비 글리코실화(N297G) 및 단량체 생성 돌연변이(C220Q, C226Q, C229Q, T366R, L368H, P395K, K409T, M428L), Eu 넘버링)가 있는, IgG1의 Fc 단편을 포함하는 융합 단백질(서열번호 88)
90	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 89의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.
91	서열번호 89의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 90의 DNA 서열.
92	6xHis 태그, HaloTag 및 Sumo* 요소로 구성된 다기능 융합 태그
93	GS-링커
94	인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 93)의 융합된 C 말단 및 서열번호 92의 6xHis 태그, HaloTag 및 Sumo* 요소로 구성된 다기능 융합 태그를 포함하는 융합 단백질
95	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 114의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.

[0386]

96	서열번호 114의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 95의 DNA 서열.
97	인간 혈청 알부민의 아미노산 25~609(서열번호 52)에 대해 N 말단에 추가로 연결되는, 서열번호 21의 링커의 N 말단에 연결된, CCN5의 아미노산 194~237(여기서, 195번 위치의 프롤린이 알라닌으로 교체됨)(서열번호 38)을 포함하는 본 발명의 융합 단백질
98	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 혈청 알부민의 아미노산 25~609(서열번호 52)에 대해 N 말단에 추가로 연결되는, 서열번호 21의 링커의 N 말단에 연결된, CCN5의 아미노산 194~237(여기서, 195번 위치의 프롤린이 알라닌으로 교체됨)(서열번호 38)을 포함하는 융합 단백질(서열번호 97)
99	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 98의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.
100	서열번호 98의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 99의 DNA 서열.
101	인간 혈청 알부민의 아미노산 25~606
102	인간 혈청 알부민의 아미노산 25~606(서열번호 101)에 대해 C 말단에 추가로 연결되는, 서열번호 22의 링커의 C 말단에 연결된, 인간 WISP2/CCN5의 아미노산 194~246(여기서, 195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨)(서열번호 7)을 포함하는 본 발명의 융합 단백질
103	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 혈청 알부민의 아미노산 25~606(서열번호 101)에 대해 C 말단에 추가로 연결되는, 서열번호 22의 링커의 C 말단에 연결된, 인간 WISP2/CCN5의 아미노산 194~246(여기서, 195번 위치의 아미노산(프롤린)이 알라닌으로 치환됨)(서열번호 7)을 포함하는 융합 단백질(서열번호 102).
104	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 103의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.

[0387]

105	서열번호 103의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 104의 DNA 서열.
106	인간 혈청 알부민의 아미노산 25~609(서열번호 52)에 대해 N 말단에 추가로 연결되는, 서열번호 21의 링커의 N 말단에 연결된, 인간 NOV/CCN3의 아미노산 206~249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44)를 포함하는 본 발명의 융합 단백질
107	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 혈청 알부민의 아미노산 25~609(서열번호 52)에 대해 N 말단에 추가로 연결되는, 서열번호 21의 링커의 N 말단에 연결된, 인간 NOV/CCN3의 아미노산 206~249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44)를 포함하는 융합 단백질(서열번호 106)
108	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 107의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.
109	서열번호 108의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 108의 DNA 서열.
110	인간 혈청 알부민의 아미노산 25~609(서열번호 52)에 대해 N 말단에 추가로 연결되는, 서열번호 22의 링커의 N 말단에 연결된, 인간 NOV/CCN3의 아미노산 206~249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44)를 포함하는 본 발명의 융합 단백질
111	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 혈청 알부민의 아미노산 25~609(서열번호 52)에 대해 N 말단에 추가로 연결되는, 서열번호 22의 링커의 N 말단에 연결된, 인간 NOV/CCN3의 아미노산 206~249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44)을 포함하는 융합 단백질(서열번호 110)
112	햄스터 세포에서의 발현을 위하여 코돈 최적화된 서열번호 111의 융합 단백질을 암호화하고 5' 말단에 KOZAK 서열인 GCCACC가 추가로 첨부되고 3' 말단에 번역 정지 코돈이 추가로 첨부된 DNA 서열.

[0388]

113	서열번호 111의 융합 단백질을 암호화하고 양 말단에 게이트웨이 AttB 재조합효소 부위를 추가로 포함하는 서열번호 112의 DNA 서열.
114	인간 혈청 알부민으로부터의 신호 펩티드(서열번호 32)가 아미노 말단에 첨부된, 인간 NOV/CCN3의 아미노산 206 ~ 249(도메인 III/TSP-1 상동성 도메인)(여기서, 207번 위치의 아미노산(이소류신)이 알라닌으로 치환됨)(서열번호 44), 펩티드 링커(서열번호 93)의 융합된 C 말단 및 서열번호 92의 6xHis 태그, HaloTag 및 Sumo* 요소로 구성된 다기능 융합 태그를 포함하는 융합 단백질(서열번호 94)
115	화학식 I의 잠재적 섹션 F
116	화학식 I의 잠재적 섹션 F
117	화학식 I의 잠재적 섹션 F
118	화학식 I의 잠재적 섹션 F
119	화학식 I의 잠재적 섹션 F
120	화학식 I의 잠재적 섹션 F
121	GGGGS 링커

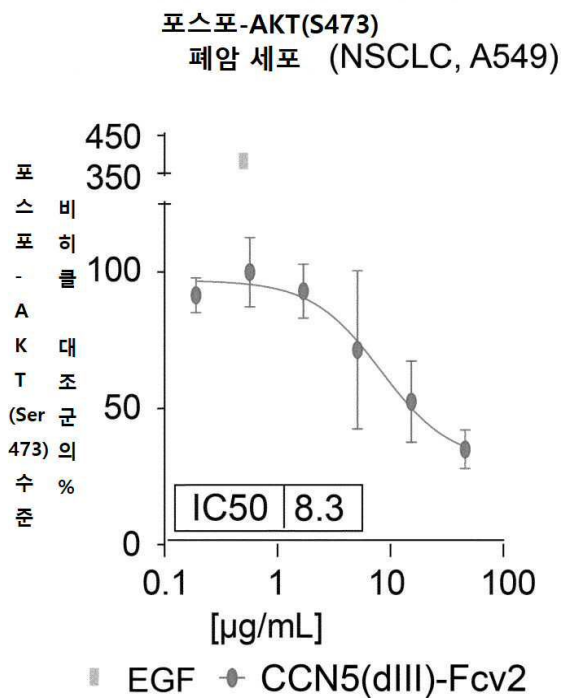
[0389]

[0390]

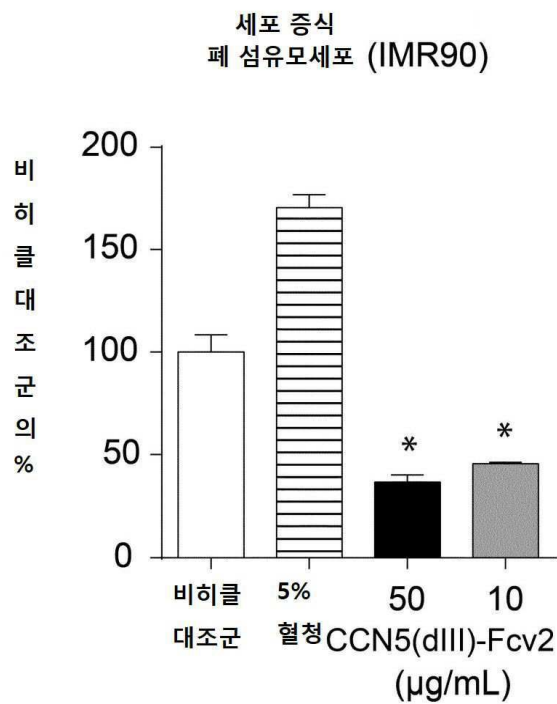
위의 "발명의 상세한 설명"에 기술된 바와 같이, uniprot 데이터베이스에 따른 CCN 단백질의 넘버링. 위의 "발명의 상세한 설명"에 기술된 바와 같은 Eu 넘버링 시스템에 따른 Fc 단편의 넘버링.

도면

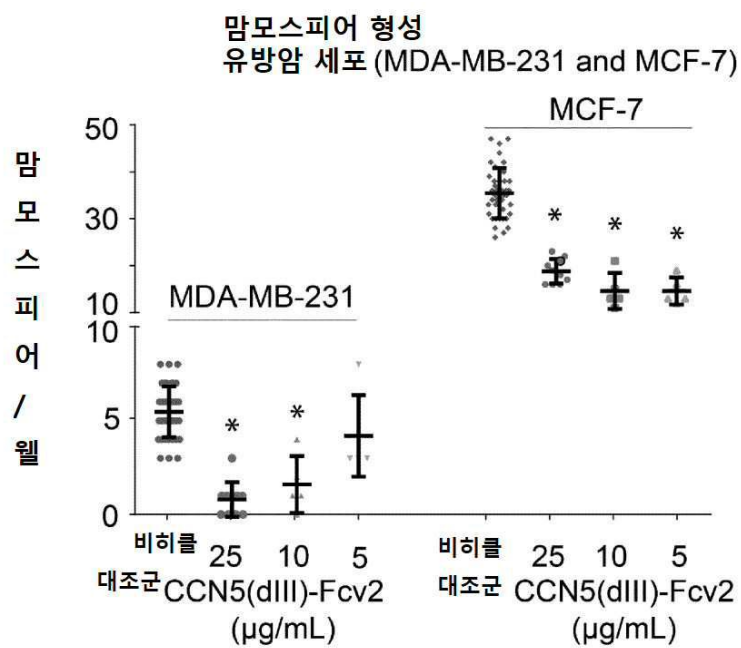
도면1a



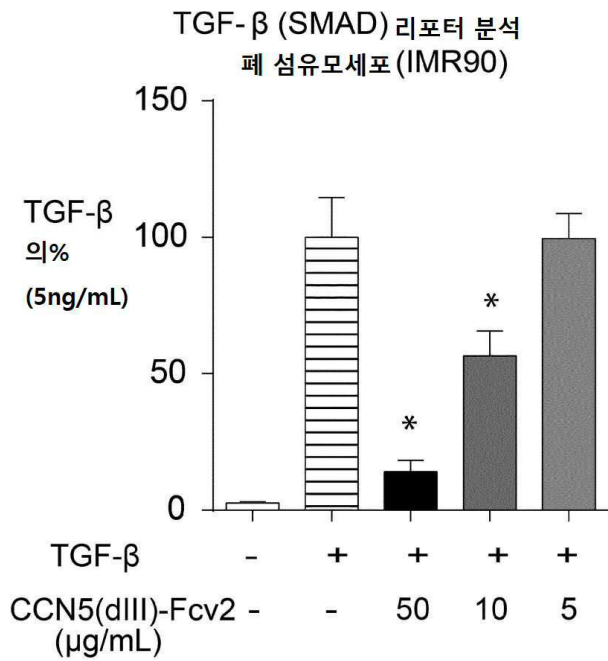
도면1b



도면1c

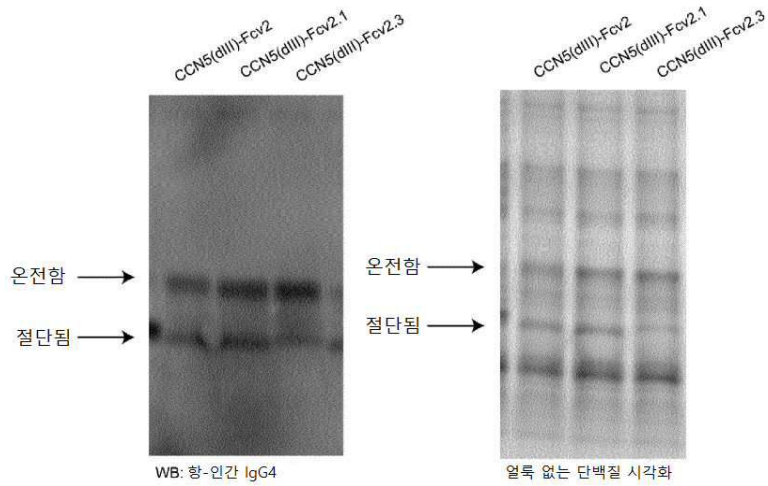


도면1d



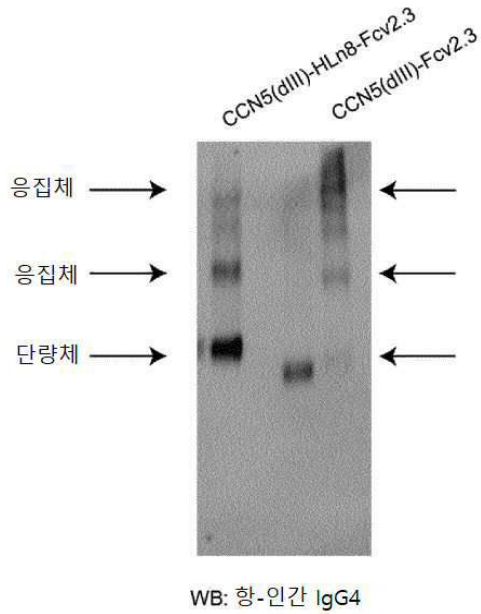
도면2

IgG4로부터의 불변 중쇄 도메인 2 내지 3과 CCN5(도메인 III)-Fc-융합 힌지 변이체로
일시적으로 형질감염된 ExpiCHO 현탁 배양으로부터의 조정 배지

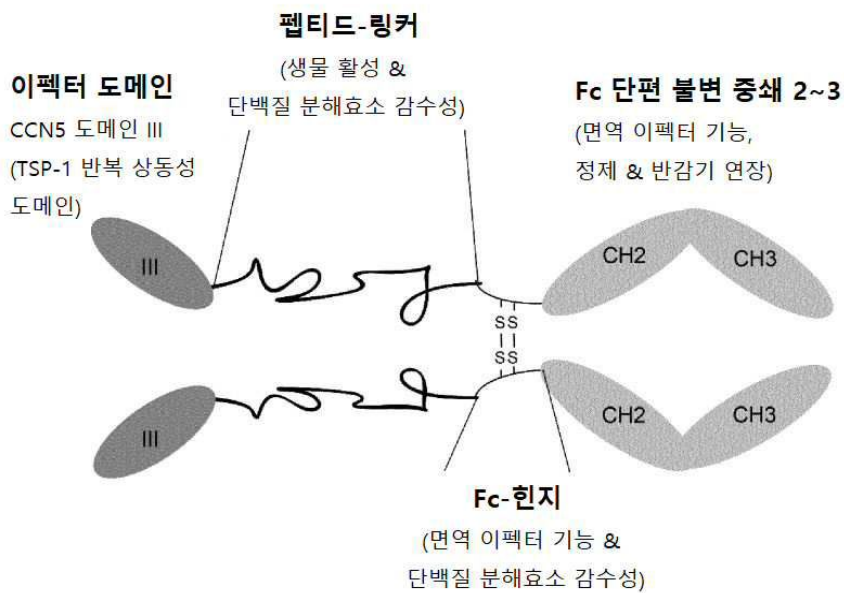


도면3

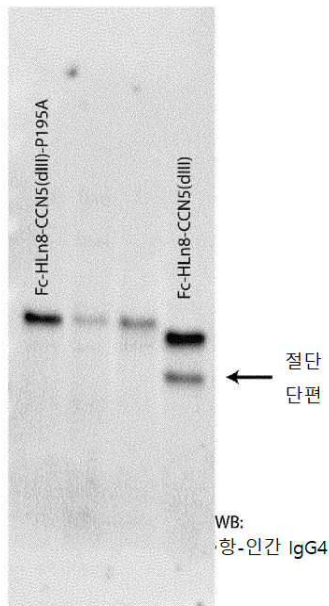
상이한 펩티드 링커와의 CCN(도메인 III)-Fc-융합 변이체로
일시적으로 형질감염된 ExpiCHO 현탁 배양으로부터의 조정 배지



도면4

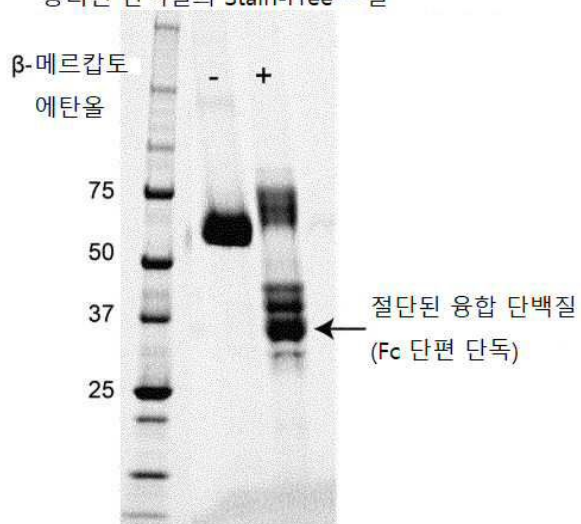


도면5



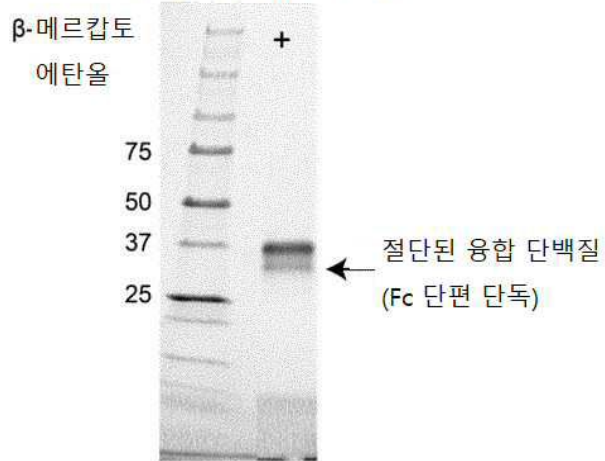
도면6

단백질 A 포획 크로마토그래피에 의해 정제된
서열번호 58(CCN(dIII)-SL-Fc_v0)에 상응하는
용리된 단백질의 Stain-Free™ 겔



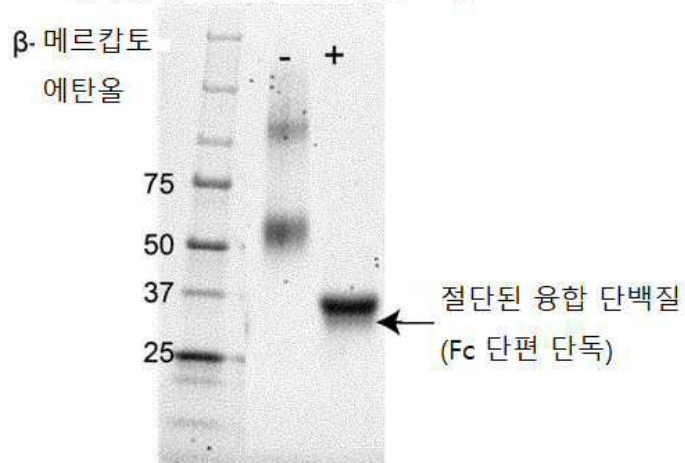
도면7

단백질 A 포획 크로마토그래피에 의해 정제된
서열번호 27에 상응하는 용리된
단백질의 Stain-Free™ 겔

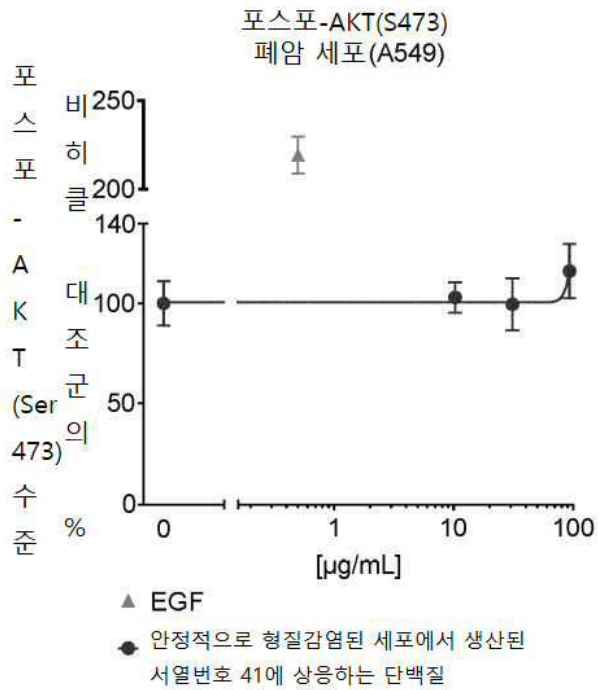


도면8

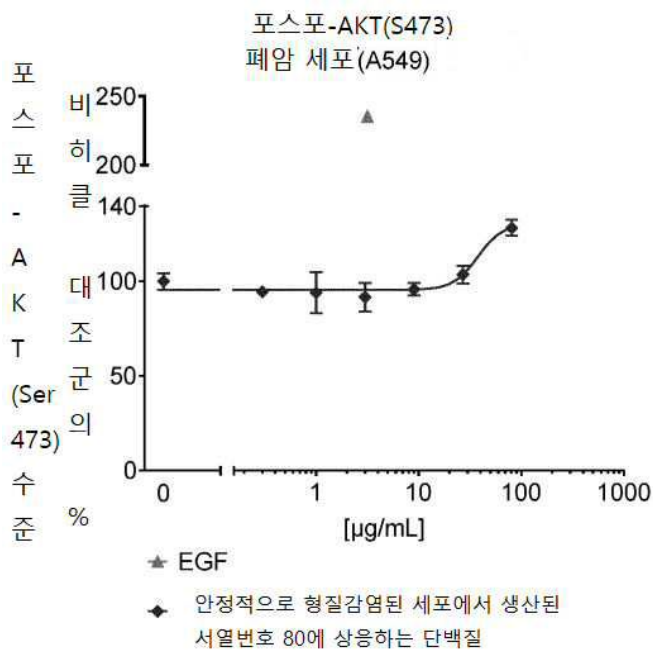
단백질 A 포획 크로마토그래피에 의해 정제된
서열번호 73에 상응하는
용리된 단백질의 Stain-Free™ 겔



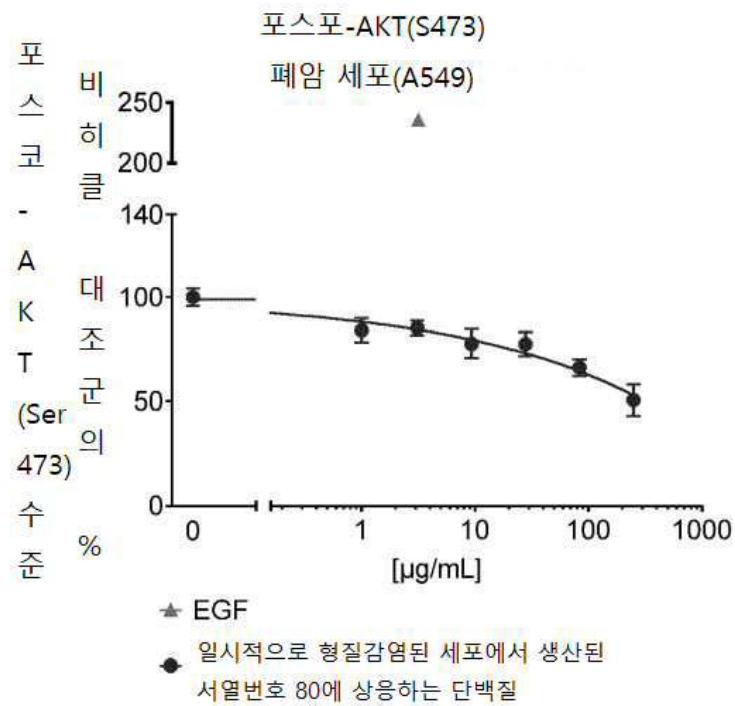
도면9



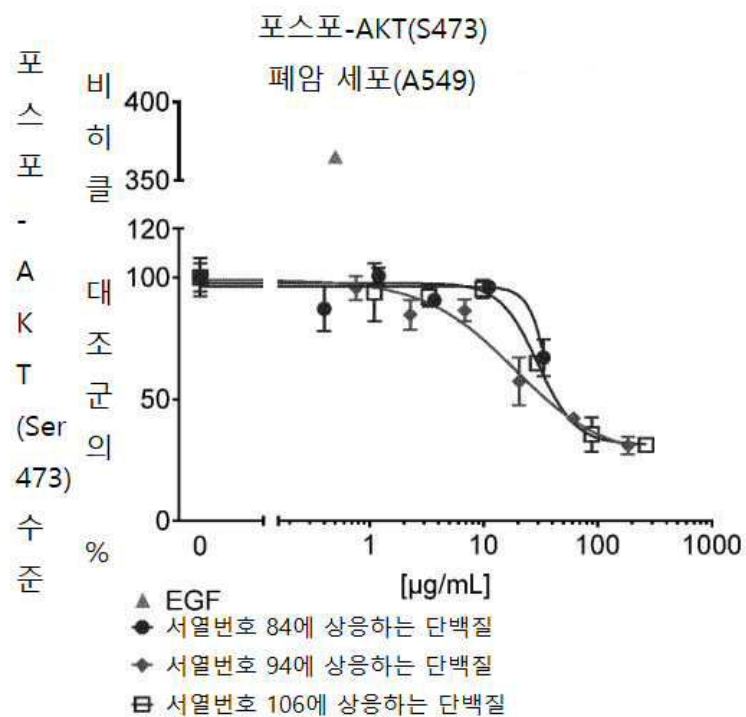
도면10



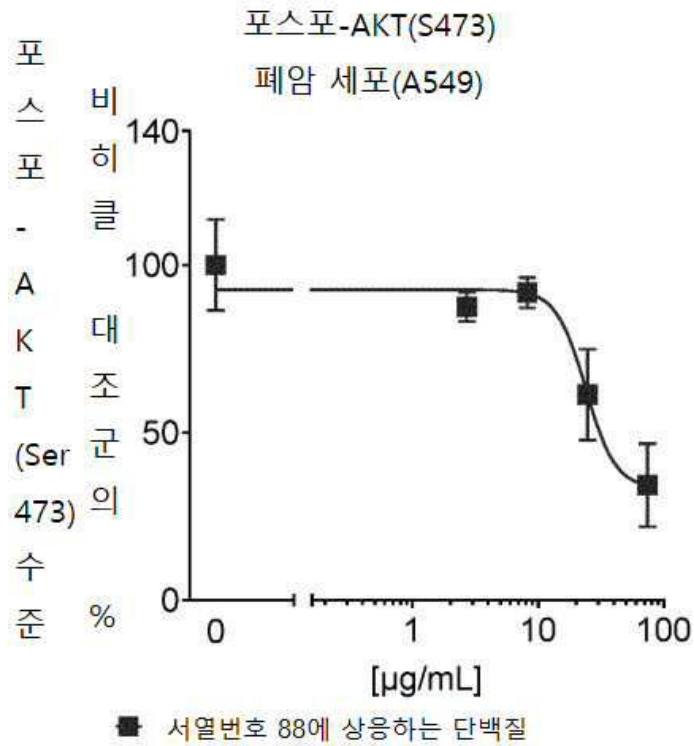
도면11



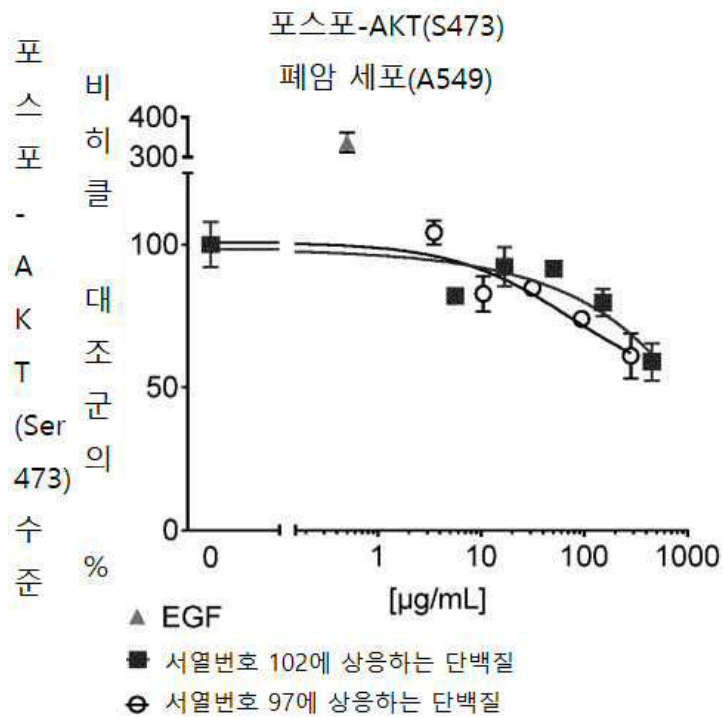
도면12



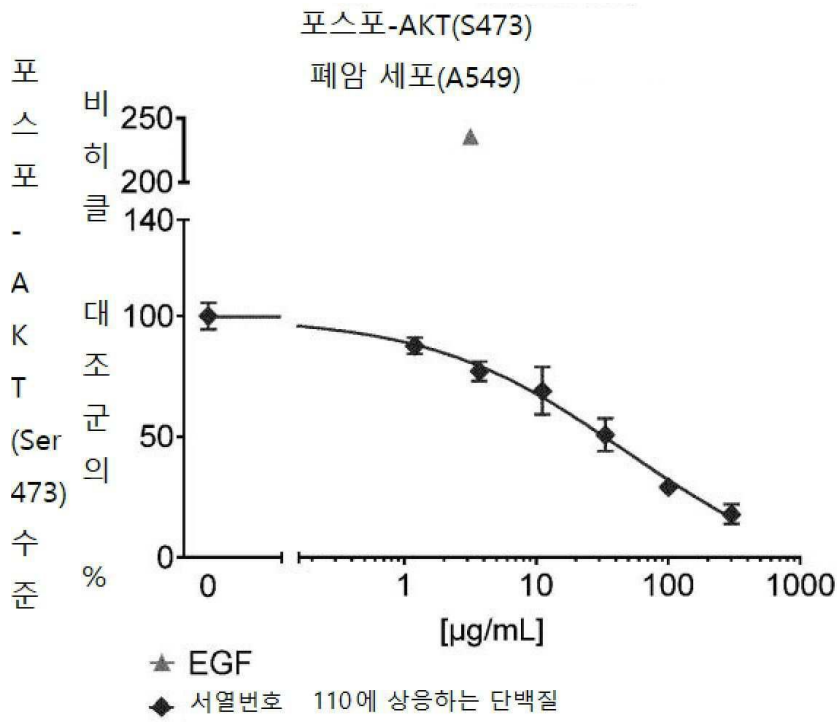
도면13



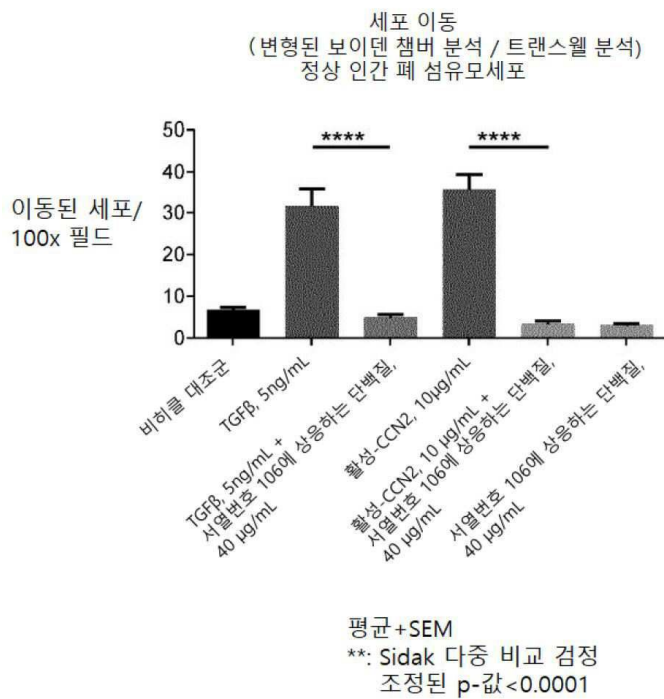
도면14



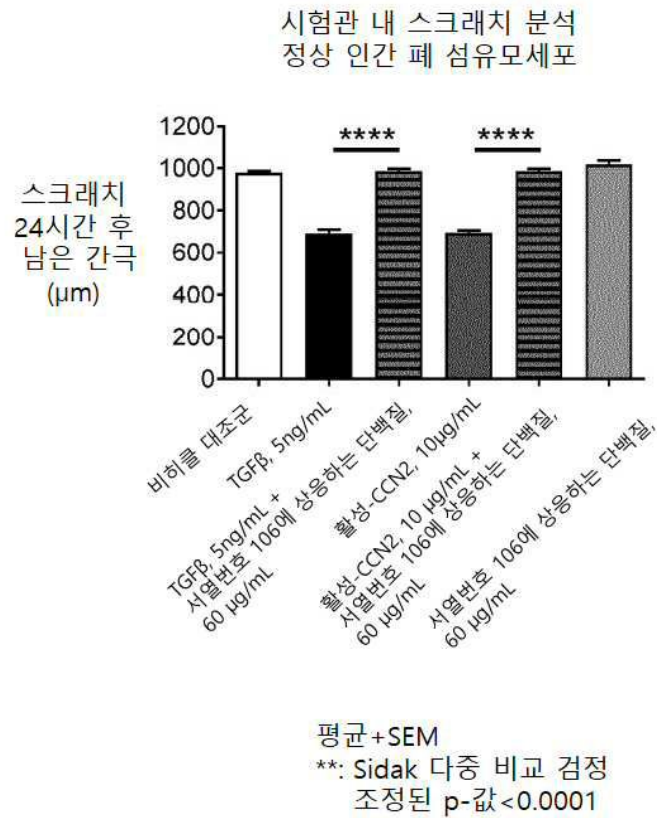
도면15



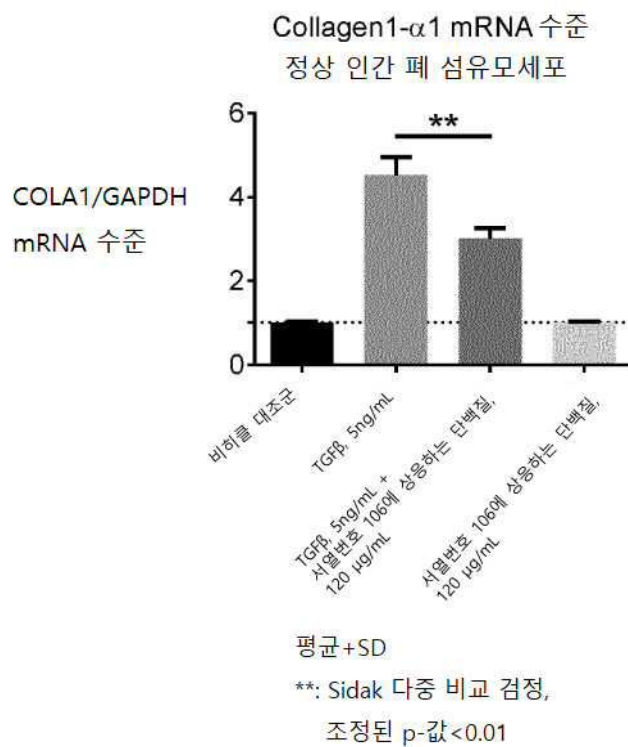
도면16a



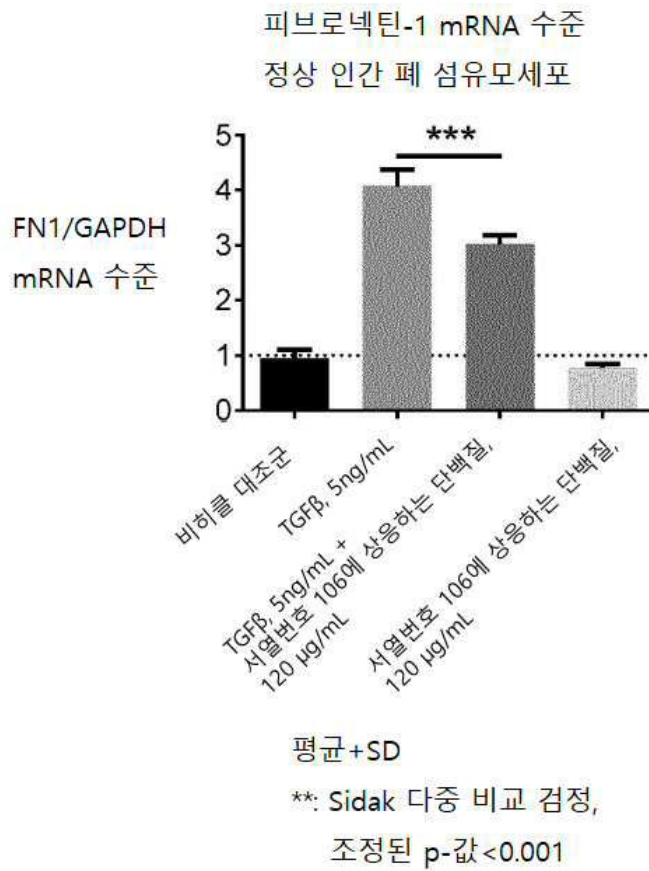
도면16b



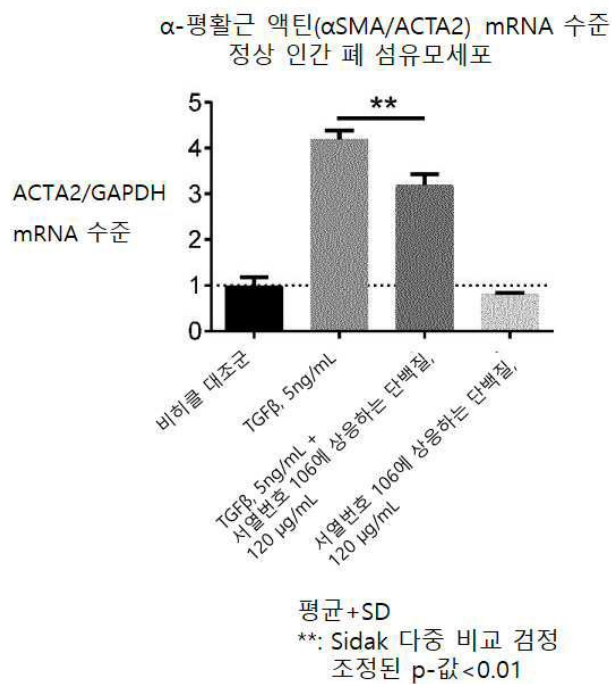
도면16c



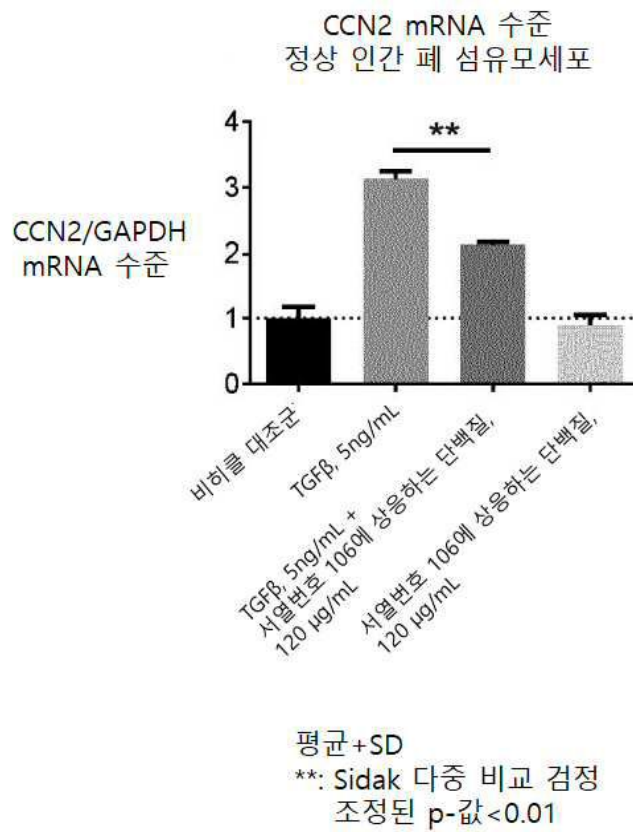
도면16d



도면16e



도면16f



서열 목록

SEQUENCE LISTING

<110> Oslo Universitetssykehus HF

<120> Recombinant CCN domain proteins and fusion proteins

<130> 27.11.145331/01

<150> EP 19163970.7

<151> 2019-03-20

<160> 121

<170> PatentIn version 3.5

<210> 1

<211> 53

<212> PRT

<213> Homo sapiens

<400> 1

Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu
20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg
35 40 45

Gly Arg Ser Pro Gln
50

<210> 2

<211> 44

<212> PRT

<213> Homo sapiens

<400> 2

Cys Ile Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly
1 5 10 15

Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met
20 25 30

Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys
35 40

<210> 3

<211> 44

<212> PRT

<213> Homo sapiens

<400> 3

Cys Leu Val Gln Thr Thr Glu Trp Ser Ala Cys Ser Lys Thr Cys Gly
1 5 10 15

Met Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Ala Ser Cys Arg Leu
20 25 30

Glu Lys Gln Ser Arg Leu Cys Met Val Arg Pro Cys
35 40

<210> 4

<211> 44

<212> PRT

<213> Homo sapiens

<400> 4

Cys Ile Val Gln Thr Thr Ser Trp Ser Gln Cys Ser Lys Thr Cys Gly

1 5 10 15

Thr Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Pro Glu Cys Arg Leu

20 25 30

Val Lys Glu Thr Arg Ile Cys Glu Val Arg Pro Cys

35 40

<210> 5

<211> 44

<212> PRT

<213> Homo sapiens

<400> 5

Cys Ile Ala Tyr Thr Ser Pro Trp Ser Pro Cys Ser Thr Ser Cys Gly

1 5 10 15

Leu Gly Val Ser Thr Arg Ile Ser Asn Val Asn Ala Gln Cys Trp Pro

20 25 30

Glu Gln Glu Ser Arg Leu Cys Asn Leu Arg Pro Cys

35 40

<210> 6

<211> 44

<212> PRT

<213> Homo sapiens

<400> 6

Cys Leu Val Gln Ala Thr Lys Trp Thr Pro Cys Ser Arg Thr Cys Gly

1 5 10 15

Met Gly Ile Ser Asn Arg Val Thr Asn Glu Asn Ser Asn Cys Glu Met

20 25 30

Arg Lys Glu Lys Arg Leu Cys Tyr Ile Gln Pro Cys

35 40

<210> 7

<211> 53

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN5 sequence

<400> 7

Cys Ala Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu

20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg

35 40 45

Gly Arg Ser Pro Gln

50

<210> 8

<211> 55

<212> PRT

<213> Homo sapiens

<400> 8

Cys Leu Val Gln Ala Thr Lys Trp Thr Pro Cys Ser Arg Thr Cys Gly

1 5 10 15

Met Gly Ile Ser Asn Arg Val Thr Asn Glu Asn Ser Asn Cys Glu Met

20 25 30

Arg Lys Glu Lys Arg Leu Cys Tyr Ile Gln Pro Cys Asp Ser Asn Ile

35 40 45

Leu Lys Thr Ile Lys Ile Pro

50 55

<210> 9

<211> 53

<212> PRT

<213> Homo sapiens

<400> 9

Cys Ile Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly

1 5 10 15

Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met

20 25 30

Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Gln Glu Pro

35 40 45

Glu Gln Pro Thr Asp

50

<210> 10

<211> 52

<212> PRT

<213> Homo sapiens

<400> 10

Cys Leu Val Gln Thr Thr Glu Trp Ser Ala Cys Ser Lys Thr Cys Gly

1 5 10 15

Met Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Ala Ser Cys Arg Leu

20 25 30

Glu Lys Gln Ser Arg Leu Cys Met Val Arg Pro Cys Glu Ala Asp Leu

35 40 45

Glu Glu Asn Ile

50

<210> 11

<211> 52

<212> PRT

<213> Homo sapiens

<400> 11

Cys Ile Val Gln Thr Thr Ser Trp Ser Gln Cys Ser Lys Thr Cys Gly

1 5 10 15

Thr Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Pro Glu Cys Arg Leu

20 25 30

Val Lys Glu Thr Arg Ile Cys Glu Val Arg Pro Cys Gly Gln Pro Val

35 40 45

Tyr Ser Ser Leu

50

<210> 12

<211> 52

<212> PRT

<213> Homo sapiens

<400> 12

Cys Ile Ala Tyr Thr Ser Pro Trp Ser Pro Cys Ser Thr Ser Cys Gly
 1 5 10 15
 Leu Gly Val Ser Thr Arg Ile Ser Asn Val Asn Ala Gln Cys Trp Pro
 20 25 30
 Glu Gln Glu Ser Arg Leu Cys Asn Leu Arg Pro Cys Asp Val Asp Ile
 35 40 45

His Thr Leu Ile

50

<210> 13

<211> 229

<212> PRT

<213> Homo sapiens

<400> 13

Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro Ala Pro Glu Phe
 1 5 10 15
 Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr
 20 25 30
 Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val
 35 40 45
 Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val

50

55

60

Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser
 65 70 75 80
 Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu
 85 90 95
 Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser
 100 105 110
 Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro

115

120

125

Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln

130 135 140
Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
145 150 155 160
Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
165 170 175
Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu

180 185 190
Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser
195 200 205
Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
210 215 220

Leu Ser Leu Gly Lys

225

<210> 14

<211> 228

<212> PRT

<213> Homo sapiens

<400> 14

Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro Pro Val

1 5 10 15
Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu
20 25 30
Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser
35 40 45
His Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu
50 55 60
Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr

65 70 75 80
Phe Arg Val Val Ser Val Leu Thr Val Val His Gln Asp Trp Leu Asn
85 90 95
Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ala Pro
100 105 110

Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly Gln Pro Arg Glu Pro Gln
115 120 125
Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val
130 135 140
Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ser Val
145 150 155 160
Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro
165 170 175
Pro Met Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr
180 185 190
Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val
195 200 205
Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
210 215 220
Ser Pro Gly Lys
225
<210> 15
<211> 228
<212> PRT
<213> Artificial Sequence
<220><223> Modified Fc fragment
<400> 15
Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu Ala
1 5 10 15
Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr
20 25 30
Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val
35 40 45
Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val
50 55 60
Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser
65 70 75 80

Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu

85 90 95

Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser

100 105 110

Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro

115 120 125

Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln

130 135 140

Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala

145 150 155 160

Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr

165 170 175

Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu

180 185 190

Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser

195 200 205

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

210 215 220

Leu Ser Leu Gly

225

<210> 16

<211> 232

<212> PRT

<213> Artificial Sequence

<220><223> Modified Fc fragment

<400> 16

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

1 5 10 15

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro

20 25 30

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

35 40 45
Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
50 55 60
Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Cys Glu Glu Gln
65 70 75 80
Tyr Gly Ser Thr Tyr Arg Cys Val Ser Val Leu Thr Val Leu His Gln
85 90 95
Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala

100 105 110
Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
115 120 125
Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr
130 135 140
Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
145 150 155 160
Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr

165 170 175
Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
180 185 190
Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
195 200 205
Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
210 215 220
Ser Leu Ser Leu Ser Pro Gly Lys
225 230

<210> 17

<211> 231

<212> PRT

<213> Artificial Sequence

<220><223> Modified Fc fragment

<400> 17

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

1 5 10 15
 Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
 20 25 30
 Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val
 35 40 45

 Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp
 50 55 60
 Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Cys Glu Glu Gln Tyr
 65 70 75 80
 Gly Ser Thr Tyr Arg Cys Val Ser Val Leu Thr Val Leu His Gln Asp
 85 90 95
 Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu
 100 105 110

 Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
 115 120 125
 Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys
 130 135 140
 Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
 145 150 155 160
 Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
 165 170 175

 Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
 180 185 190
 Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser
 195 200 205
 Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
 210 215 220
 Leu Ser Leu Ser Pro Gly Lys
 225 230
 <210> 18
 <211> 227
 <212> PRT

<213> Artificial Sequence

<220><223> Modified Fc fragment

<400> 18

Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Pro Val

1 5 10 15

Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu

20 25 30

Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser

35 40 45

Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu

50 55 60

Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr

65 70 75 80

Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn

85 90 95

Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser

100 105 110

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln

115 120 125

Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val

130 135 140

Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val

145 150 155 160

Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro

165 170 175

Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr

180 185 190

Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val

195 200 205

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu

210 215 220

Ser Leu Gly

225

<210> 19

<211> 227

<212> PRT

<213> Artificial Sequence

<220><223> Modified Fc fragment

<400> 19

Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro Pro Val

1 5 10 15

Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu

20 25 30

Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser

35 40 45

Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu

50 55 60

Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr

65 70 75 80

Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn

85 90 95

Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser

100 105 110

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln

115 120 125

Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val

130 135 140

Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val

145 150 155 160

Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro

165 170 175

Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr

180 185 190

Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val
 195 200 205

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
 210 215 220

Ser Leu Gly

225

<210> 20

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 20

Thr Glu Gly Arg Met Asp

1 5

<210> 21

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 21

Glu Ala Ala Ala Lys

1 5

<210> 22

<211> 40

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 22

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu

1 5 10 15

Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala

20 25 30

Ala Ala Lys Glu Ala Ala Ala Lys

35 40

<210> 23

<211> 42

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 23

Thr Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala

1 5 10 15

Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

20 25 30

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

35 40

<210> 24

<211> 43

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 24

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu

1 5 10 15

Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala

20 25 30

Ala Ala Lys Glu Ala Ala Ala Lys Ala Ala Ala

35 40

<210> 25

<211> 45

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 25

Thr Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala

1 5 10 15

Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

20 25 30

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala Ala Ala

35 40 45

<210> 26

<211> 276

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 26

Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu

20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Glu Ala Ala Ala

35 40 45

Lys Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro Pro

50 55 60

Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr

65 70 75 80

Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val

85 90 95

Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val

100 105 110

Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser

115 120 125

Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu

130 135 140

Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser

145 150 155 160

Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro

165 170 175
 Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln
 180 185 190
 Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
 195 200 205
 Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
 210 215 220

Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu
 225 230 235 240
 Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser
 245 250 255
 Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
 260 265 270

Leu Ser Leu Gly

275

<210> 27

<211> 276

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 27

Cys Ala Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly
 1 5 10 15
 Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu
 20 25 30
 Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Glu Ala Ala Ala
 35 40 45
 Lys Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro Pro
 50 55 60

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

85 90 95
 Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val
 100 105 110
 Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser
 115 120 125

 Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu
 130 135 140
 Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser
 145 150 155 160
 Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
 165 170 175
 Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln
 180 185 190

 Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
 195 200 205
 Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
 210 215 220
 Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu
 225 230 235 240
 Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser
 245 250 255

 Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
 260 265 270
 Leu Ser Leu Gly
 275
 <210> 28
 <211> 287
 <212> PRT
 <213> Artificial Sequence
 <220><223> Fusion protein
 <400> 28
 Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15
 Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu
 20 25 30

 Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg
 35 40 45
 Gly Arg Ser Pro Gln Thr Glu Gly Arg Met Asp Glu Ser Lys Tyr Gly
 50 55 60
 Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu Ala Ala Gly Gly Pro Ser
 65 70 75 80
 Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
 85 90 95

 Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser Gln Glu Asp Pro
 100 105 110
 Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
 115 120 125
 Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg Val Val
 130 135 140
 Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
 145 150 155 160

 Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr
 165 170 175
 Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
 180 185 190
 Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
 195 200 205
 Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
 210 215 220

 Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
 225 230 235 240
 Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser
 245 250 255

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

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

<210> 29

<211> 286

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 29

Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly
1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu
20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg
35 40 45

Gly Arg Ser Pro Gln Thr Glu Gly Arg Met Asp Glu Ser Lys Tyr Gly

50 55 60

Pro Pro Cys Pro Pro Cys Pro Ala Pro Pro Val Ala Gly Pro Ser Val

65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr

85 90 95

Pro Glu Val Thr Cys Val Val Val Asp Val Ser Gln Glu Asp Pro Glu

100 105 110

Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys

115 120 125

Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg Val Val Ser

130 135 140

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys

145 150 155 160

Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile

165 170 175

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

180 185 190

Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu

195 200 205

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

210 215 220

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

225 230 235 240

Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg

245 250 255

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

260 265 270

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly

275 280 285

<210> 30

<211> 286

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 30

Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu

20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg

35 40 45

Gly Arg Ser Pro Gln Thr Glu Gly Arg Met Asp Glu Arg Lys Cys Cys

50 55 60

Val Glu Cys Pro Pro Cys Pro Ala Pro Pro Val Ala Gly Pro Ser Val

65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr

85 90 95
 Pro Glu Val Thr Cys Val Val Val Asp Val Ser Gln Glu Asp Pro Glu
 100 105 110
 Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
 115 120 125
 Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg Val Val Ser
 130 135 140

 Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
 145 150 155 160
 Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile
 165 170 175
 Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
 180 185 190
 Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
 195 200 205

 Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
 210 215 220
 Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
 225 230 235 240
 Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg
 245 250 255
 Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
 260 265 270

 His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly
 275 280 285
 <210> 31
 <211> 325
 <212> PRT
 <213> Artificial Sequence
 <220><223> Fusion protein
 <400> 31
 Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15
 Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu
 20 25 30
 Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg

 35 40 45
 Gly Arg Ser Pro Gln Thr Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala
 50 55 60
 Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
 65 70 75 80
 Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala
 85 90 95
 Ala Ala Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro

 100 105 110
 Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
 115 120 125
 Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
 130 135 140
 Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly
 145 150 155 160
 Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn

 165 170 175
 Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp
 180 185 190
 Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro
 195 200 205
 Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu
 210 215 220
 Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn

 225 230 235 240
 Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
 245 250 255

Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
260 265 270

Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg
275 280 285

Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys

290 295 300

Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
305 310 315 320

Ser Leu Ser Leu Gly
325

<210> 32

<211> 18

<212> PRT

<213> Homo sapiens

<400> 32

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser

<210> 33

<211> 305

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 33

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr
20 25 30

Cys Gly Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys
35 40 45

Arg Leu Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro

50 55 60
 Ser Arg Gly Arg Ser Pro Gln Thr Glu Gly Arg Met Asp Glu Ser Lys
 65 70 75 80
 Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu Ala Ala Gly Gly
 85 90 95
 Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile
 100 105 110
 Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser Gln Glu
 115 120 125

 Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val His
 130 135 140
 Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg
 145 150 155 160
 Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
 165 170 175
 Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu
 180 185 190

 Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr
 195 200 205
 Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val Ser Leu
 210 215 220
 Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp
 225 230 235 240
 Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val
 245 250 255

 Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp
 260 265 270
 Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val Met His
 275 280 285
 Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu
 290 295 300

Gly

305

<210> 34

<211> 915

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 34

```
atgaaatggg tcacctttat ctccctgctg ttcctgttct cctccgcta ctcttgccct 60
gagtgttcta cagcttgggg cccttgctct accacctgtg gactcggcat ggccaccaga 120
gtgtctaacc agaacagatt ctgccggctg gaaaccagc ggagactgtg cctgtctaga 180
ccctgtcctc ctagcagagg cagatcccct cagaccgagg gcagaatgga cgagtctaag 240
tacggccctc cttgtcctcc atgtcctgct ccagaagctg ctggcggccc ttccgtgttt 300
ctgttcctc caaagcctaa ggacacctg atgatctctc ggacctga agtgacctgc 360
gtggtggtgg atgtgtccca agaggatccc gaggtgcagt tcaattggta cgtggacggc 420
```

```
gtggaagtgc acaacgcaa gaccaagcct agagaggaac agttcaactc cacctacaga 480
gtggtgtccg tgctgaccgt gctgcaccag gattggctga acggcaaaga gtacaagtgc 540
aaggtgtcca acaagggcct gccttcagc atcgaaaaga ccatctcaa ggccaagggc 600
cagcctaggg aaccccaggt ttacacctg cctccaagcc aagaggaaat gaccaagaac 660
caggtgtccc tgacctgctt ggtcaagggc ttctaccctt ccgatatcgc cgtggaatgg 720
gagagcaatg gccagctga gaacaactac aagaccacac ctctgtgct ggactccgac 780
ggctccttct ttctgtactc ccgctgacc gtggacaagt ccagatggca agagggaac 840
```

```
gtgttctcct gctccgtgat gcacaggcc ctgcacaatc actacacca gaagtcctg 900
tctctgtccc tgggc 915
```

<210> 35

<211> 924

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 35

```
gccacatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgcctactct 60
tgccctgagt ggtctacagc ttggggccct tgctctacca cctgtggact cggcatggcc 120
```

accagagtgt ctaaccagaa cagattctgc cggctggaaa cccagcggag actgtgcctg 180
tctagaccct gtctcctag cagaggcaga tccctcaga ccgagggcag aatggacgag 240

tctaagtacg gccctccttg tcttccatgt cctgtccag aagctgctgg cggcccttcc 300
gtgtttctgt tccctccaaa gcctaaggac accctgatga tctctcggac ccctgaagtg 360
acctgcgtgg tgggtgatgt gtcccaagag gatcccagg tgcagttcaa ttggtacgtg 420
gacggcgtgg aagtgcacaa cgccaagacc aagcctagag aggaacagtt caactccacc 480
tacagagtgg tgtccgtgct gaccgtgctg caccaggatt ggctgaacgg caaagagtac 540
aagtgaagg tgtccaacaa gggcctgcct tccagcatcg aaaagaccat ctccaaggcc 600
aagggccagc ctagggaacc ccaggtttac accctgcctc caagccaaga ggaaatgacc 660

aagaaccagg tgtccctgac ctgcctggtc aagggttctt acccttccga tatcgccgtg 720
gaatgggaga gcaatggcca gcctgagaac aactacaaga ccacacctcc tgtgctggac 780
tccgacggct ccttctttct gtactccgc ctgaccgtgg acaagtccag atggcaagag 840
ggcaacgtgt tctcctgctc cgtgatgcac gaggccctgc acaatcacta caccagaag 900
tccctgtctc tgtccctggg ctaa 924

<210> 36
<211> 997
<212> DNA
<213> Artificial Sequence
<220><223> Fusion protein
<400> 36

ggggacaagt ttgtacaaa aagcaggcta tggtagcc accatgaaat gggtcacctt 60

tatctccctg ctgttctgt tctctccgc ctactcttgc cctgagtgg ctacagcttg 120
gggcccttgc tctaccact gtggactcgg catggccacc agagtgtcta accagaacag 180
attctgccgg ctggaaacct agcggagact gtgcctgtct agaccctgtc ctcttagcag 240
aggcagatcc cctcagaccg agggcagaat ggacgagtct aagtacggcc ctccttgtec 300
tccatgtcct gctccagaag ctgctggcgg cccttccgtg tttctgttcc ctccaaagcc 360
taaggacacc ctgatgatct ctcggacccc tgaagtgacc tgcgtggtgg tggatgtgtc 420
ccaagaggat cccgagggtgc agttcaattg gtacgtggac ggctggaag tgcacaacgc 480

caagaccaag cctagagagg aacagttcaa ctccacctac agagtgggtg ccgtgctgac 540
cgtgctgcac caggattggc tgaacggcaa agagtacaag tgcaaggtgt ccaacaaggg 600
cctgccttcc agcatcgaag agaccatctc caaggccaag ggccagccta gggaacccca 660

ggtttacacc ctgcctccaa gccaaagga aatgaccaag aaccaggtgt ccctgacctg 720
 cctgggtcaag ggcttctacc cttccgatat cgccgtggaa tgggagagca atggccagcc 780
 tgagaacaac tacaagacca cacctcctgt gctggactcc gacggctcct tctttctgta 840
 ctcccgctg accgtggaca agtccagatg gcaagagggc aacgtgttct cctgctccgt 900

gatgcacgag gccctgcaca atcactacac ccagaagtcc ctgtctctgt ccctgggcta 960
 atctagaaac ccagctttct tgtacaaagt ggtccccc 997

<210> 37

<211> 44

<212> PRT

<213> Homo sapiens

<400> 37

Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu

20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys

35 40

<210> 38

<211> 44

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN5 sequence

<400> 38

Cys Ala Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu

20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys

35 40

<210> 39

<211> 44

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 39

Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

1 5 10 15

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu

20 25 30

Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala Ala Ala

35 40

<210> 40

<211> 325

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 40

Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu Ala

1 5 10 15

Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr

20 25 30

Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val

35 40 45

Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val

50 55 60

Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser

65 70 75 80

Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu

85 90 95

Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser

100 105 110

Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro

115 120 125

Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln

130 135 140
 Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
 145 150 155 160
 Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
 165 170 175
 Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu
 180 185 190

 Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser
 195 200 205
 Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
 210 215 220
 Leu Ser Leu Gly Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu
 225 230 235 240
 Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala
 245 250 255

 Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala Ala Ala
 260 265 270
 Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly
 275 280 285
 Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu
 290 295 300
 Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg
 305 310 315 320

Gly Arg Ser Pro Gln

325

<210> 41

<211> 325

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 41

Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu Ala

1 5 10 15
 Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr
 20 25 30
 Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val
 35 40 45

 Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val
 50 55 60
 Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser
 65 70 75 80
 Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu
 85 90 95
 Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser
 100 105 110

 Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
 115 120 125
 Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln
 130 135 140
 Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
 145 150 155 160
 Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
 165 170 175

 Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu
 180 185 190
 Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser
 195 200 205
 Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
 210 215 220
 Leu Ser Leu Gly Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu
 225 230 235 240

 Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala
 245 250 255

Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala Ala Ala
260 265 270
Cys Ala Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly
275 280 285
Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu
290 295 300

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg
305 310 315 320
Gly Arg Ser Pro Gln
325

<210> 42

<211> 44

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN1 sequence

<400> 42

Cys Ala Val Gln Thr Thr Ser Trp Ser Gln Cys Ser Lys Thr Cys Gly
1 5 10 15
Thr Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Pro Glu Cys Arg Leu
20 25 30

Val Lys Glu Thr Arg Ile Cys Glu Val Arg Pro Cys
35 40

<210> 43

<211> 44

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN2 sequence

<400> 43

Cys Ala Val Gln Thr Thr Glu Trp Ser Ala Cys Ser Lys Thr Cys Gly
1 5 10 15
Met Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Ala Ser Cys Arg Leu
20 25 30

Glu Lys Gln Ser Arg Leu Cys Met Val Arg Pro Cys

35 40

<210> 44

<211> 44

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN3 sequence

<400> 44

Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly

1 5 10 15

Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met

20 25 30

Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys

35 40

<210> 45

<211> 44

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN4 sequence

<400> 45

Cys Ala Ala Tyr Thr Ser Pro Trp Ser Pro Cys Ser Thr Ser Cys Gly

1 5 10 15

Leu Gly Val Ser Thr Arg Ile Ser Asn Val Asn Ala Gln Cys Trp Pro

20 25 30

Glu Gln Glu Ser Arg Leu Cys Asn Leu Arg Pro Cys

35 40

<210> 46

<211> 44

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN6 sequence

<400> 46

Cys Ala Val Gln Ala Thr Lys Trp Thr Pro Cys Ser Arg Thr Cys Gly
1 5 10 15
Met Gly Ile Ser Asn Arg Val Thr Asn Glu Asn Ser Asn Cys Glu Met
20 25 30
Arg Lys Glu Lys Arg Leu Cys Tyr Ile Gln Pro Cys
35 40
<210> 47
<211> 52
<212> PRT
<213> Artificial Sequence
<220>
><223> Modified CCN1 sequence
<400> 47
Cys Ala Val Gln Thr Thr Ser Trp Ser Gln Cys Ser Lys Thr Cys Gly
1 5 10 15
Thr Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Pro Glu Cys Arg Leu
20 25 30
Val Lys Glu Thr Arg Ile Cys Glu Val Arg Pro Cys Gly Gln Pro Val
35 40 45
Tyr Ser Ser Leu
50
<210> 48
<211> 52
<212> PRT
<213> Artificial Sequence
<220><223> Modified CCN2 sequence
<400> 48
Cys Ala Val Gln Thr Thr Glu Trp Ser Ala Cys Ser Lys Thr Cys Gly
1 5 10 15
Met Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Ala Ser Cys Arg Leu
20 25 30
Glu Lys Gln Ser Arg Leu Cys Met Val Arg Pro Cys Glu Ala Asp Leu
35 40 45
Glu Glu Asn Ile

50

<210> 49

<211> 53

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN3 sequence

<400> 49

Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly

1 5 10 15

Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met

20 25 30

Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Gln Glu Pro

35 40 45

Glu Gln Pro Thr Asp

50

<210> 50

<211> 52

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN4 sequence

<400> 50

Cys Ala Ala Tyr Thr Ser Pro Trp Ser Pro Cys Ser Thr Ser Cys Gly

1 5 10 15

Leu Gly Val Ser Thr Arg Ile Ser Asn Val Asn Ala Gln Cys Trp Pro

20 25 30

Glu Gln Glu Ser Arg Leu Cys Asn Leu Arg Pro Cys Asp Val Asp Ile

35 40 45

His Thr Leu Ile

50

<210> 51

<211> 55

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN6 sequence

<400> 51

Cys Ala Val Gln Ala Thr Lys Trp Thr Pro Cys Ser Arg Thr Cys Gly

1 5 10 15

Met Gly Ile Ser Asn Arg Val Thr Asn Glu Asn Ser Asn Cys Glu Met

20 25 30

Arg Lys Glu Lys Arg Leu Cys Tyr Ile Gln Pro Cys Asp Ser Asn Ile

35 40 45

Leu Lys Thr Ile Lys Ile Pro

50 55

<210> 52

<211> 585

<212> PRT

<213> Homo sapiens

<400> 52

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

1 5 10 15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln

20 25 30

Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu

35 40 45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys

50 55 60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu

65 70 75 80

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

85 90 95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu

100 105 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His

115 120 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
 130 135 140
 Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
 145 150 155 160
 Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
 165 170 175

 Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
 180 185 190
 Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
 195 200 205
 Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
 210 215 220
 Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
 225 230 235 240

 Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
 245 250 255
 Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
 260 265 270
 Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
 275 280 285
 Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
 290 295 300

 Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
 305 310 315 320
 Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
 325 330 335
 Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
 340 345 350
 Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
 355 360 365

 Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

370 375 380
 Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
 385 390 395 400
 Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
 405 410 415
 Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
 420 425 430

 Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
 435 440 445
 Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
 450 455 460
 Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
 465 470 475 480
 Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
 485 490 495

 Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
 500 505 510
 Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
 515 520 525
 Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
 530 535 540
 Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
 545 550 555 560

 Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
 565 570 575
 Ala Ala Ser Gln Ala Ala Leu Gly Leu
 580 585

 <210> 53
 <211> 679
 <212> PRT
 <213> Homo sapiens
 <400> 53

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
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
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 Ile 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
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> 54
 <211> 232
 <212> PRT
 <213> Artificial Sequence
 <220><223> Modified Fc fragment
 <400> 54
 Glu Pro Lys Ser Gln Asp Lys Thr His Thr Gln Pro Pro Gln Pro Ala
 1 5 10 15

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
20 25 30

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
35 40 45

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
50 55 60

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Cys Glu Glu Gln
65 70 75 80

Tyr Gly Ser Thr Tyr Arg Cys Val Ser Val Leu Thr Val Leu His Gln
85 90 95

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
100 105 110

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
115 120 125

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr
130 135 140

Lys Asn Gln Val Ser Leu Arg Cys His Val Lys Gly Phe Tyr Pro Ser
145 150 155 160

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
165 170 175

Lys Thr Thr Lys Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
180 185 190

Ser Thr Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
195 200 205

Ser Cys Ser Val Leu His Glu Ala Leu His Asn His Tyr Thr Gln Lys
210 215 220

Ser Leu Ser Leu Ser Pro Gly Lys
225 230

<210> 55
<211> 227
<212> PRT

<213> Artificial Sequence

<220><223> Modified Fc fragment

<400> 55

Glu Arg Lys Gln Gln Val Glu Gln Pro Pro Gln Pro Ala Pro Pro Val

1 5 10 15

Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu

20 25 30

Tyr Ile Thr Arg Glu Pro Glu Val Thr Cys Val Val Val Asp Val Ser

35 40 45

Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu

50 55 60

Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr

65 70 75 80

Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn

85 90 95

Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser

100 105 110

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln

115 120 125

Val Tyr Thr Phe Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val

130 135 140

Ser Leu Arg Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val

145 150 155 160

Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Lys

165 170 175

Pro Val Leu Asp Ser Asp Gly Ser Phe Arg Leu Glu Ser Arg Leu Thr

180 185 190

Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val

195 200 205

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu

210 215 220

Ser Leu Gly

225

<210> 56

<211> 57

<212> PRT

<213> Homo sapiens

<400> 56

Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu

20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg

35 40 45

Gly Arg Ser Pro Gln Asn Ser Ala Phe

50 55

<210> 57

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 57

Ile Glu Gly Arg Met Asp

1 5

<210> 58

<211> 291

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 58

Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu

20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg
 35 40 45
 Gly Arg Ser Pro Gln Asn Ser Ala Phe Ile Glu Gly Arg Met Asp Glu
 50 55 60

 Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu Ala Ala
 65 70 75 80
 Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu
 85 90 95
 Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser
 100 105 110
 Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu
 115 120 125

 Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr
 130 135 140
 Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn
 145 150 155 160
 Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser
 165 170 175
 Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln
 180 185 190

 Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val
 195 200 205
 Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val
 210 215 220
 Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro
 225 230 235 240
 Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr
 245 250 255

 Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val
 260 265 270
 Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu

275 280 285
 Ser Leu Gly
 290
 <210> 59
 <211> 309
 <212> PRT
 <213> Artificial Sequence
 <220><223> Fusion protein
 <400> 59
 Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
 1 5 10 15

 Tyr Ser Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr
 20 25 30
 Cys Gly Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys
 35 40 45
 Arg Leu Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro
 50 55 60
 Ser Arg Gly Arg Ser Pro Gln Asn Ser Ala Phe Ile Glu Gly Arg Met
 65 70 75 80

 Asp Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu
 85 90 95
 Ala Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
 100 105 110
 Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
 115 120 125
 Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly
 130 135 140

 Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn
 145 150 155 160
 Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp
 165 170 175
 Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro

180 185 190
Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu
195 200 205

Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn
210 215 220
Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
225 230 235 240
Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
245 250 255
Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg
260 265 270

Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys
275 280 285
Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
290 295 300
Ser Leu Ser Leu Gly

305

<210> 60

<211> 936

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 60

gccaccatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgctactct 60

tgcctgagt ggtctacagc ttggggccct tgcctacca cctgtggact cggcatggcc 120

accagagtgt ctaaccagaa cagattctgc cggctggaaa ccagcggag actgtgtctg 180

tccagacctt gtctcttag cgggggcaga tccctcaga actctgcctt tatcgagggc 240

agaatggacg agtctaagta cggcctcct tgtccacat gtctgtctcc agaagetgct 300

ggcggccctt ccgtgtttct gttccctcca aagcctaagg acaccctgat gatctctcgg 360

acccctgaag tgacctgcgt ggtgtggat gtgtccaag aggatcccga ggtgcagttc 420

aattggtacg tggacggcgt ggaagtgcac aacccaaga ccaagcctag agaggaacag 480

ttcaactcca cctacagagt ggtgtccgtg ctgaccgtgc tgcaccagga ttggctgaac 540

ggcaaagagt acaagtgcaa ggtgtccaac aagggcctgc cttccagcat cgaaaagacc 600

atctccaagg ccaagggccca gcctagggaa ccccaggttt acaccctgcc tccaagccaa 660

gaggaaatga ccaagaacca ggtgtccctg acctgcctgg tcaagggttt ctacccttcc 720

gatatcgccg tggaatggga gagcaatggc cagcctgaga acaactacaa gaccacacct 780

cctgtgctgg actccgacgg ctcttctttt ctgtactccc gcctgaccgt ggacaagtcc 840

agatggcaag agggcaacgt gttctcctgc tccgtgatgc acgaggccct gcacaatcac 900

tacaccaga agtcctgtgc tctgtccctg ggctaa 936

<210> 61

<211> 1009

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 61

ggggacaagt ttgtacaaaa aagcaggcta tggtagcgcc accatgaaat gggtcacctt 60

tatctccctg ctgttctgt tctctccgc ctactcttgc cctgagtggc ctacagcttg 120

gggcctttgc tctaccacct gtggactcgg catggccacc agagtgtcta accagaacag 180

attctgccgg ctggaaaccc agcggagact gtgtctgtcc agaccttgtc ctctagccg 240

gggcagatcc cctcagaact ctgcctttat cgagggcaga atggacgagt ctaagtacgg 300

ccctccttgt ccacatgtc ctgtccaga agctgtggc ggcccttccg tgtttctgtt 360

ccctccaaag cctaaggaca ccctgatgat ctctcggacc cctgaagtga cctgcgtggt 420

ggtggatgtg tccaagagg atcccaggt gcagttcaat tggtagctgg acggcgtgga 480

agtgcacaac gccaaagaca agcctagaga ggaacagttc aactccacct acagagtggc 540

gtccgtgctg accgtgctgc accaggattg gctgaacggc aaagagtaca agtgcaaggt 600

gtccaacaag ggctgcctt ccagcatcga aaagaccatc tccaaggcca agggccagcc 660

tagggaaccc caggtttaca ccctgcctcc aagccaagag gaaatgacca agaaccaggt 720

gtccctgacc tgcttggtca agggcttcta cccttccgat atcgccgtgg aatgggagag 780

caatggccag cctgagaaca actacaagac cacacctcct gtgctggact ccgacggctc 840

cttctttctg tactcccgcc tgaccgtgga caagtccaga tggcaagagg gcaacgtgtt 900

ctctgtctcc gtgatgcacg aggccttgca caatcactac accagaagt ccctgtctct 960

gtccctgggc taatctagaa acccagcttt ctgtacaaa gtggtcccc 1009

<210> 62

<211> 56

<212> PRT

<213> Artificial Sequence

<220><223> Modified CCN5 sequence

<400> 62

Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly
1 5 10 15

Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu
20 25 30

Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg
35 40 45

Gly Arg Ser Leu Gln Asn Ser Ala
50 55

<210> 63

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 63

Gly Arg Met Asp

1

<210> 64

<211> 306

<212> PRT

<213>

Artificial Sequence

<220><223> Fusion protein

<400> 64

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr
20 25 30

Cys Gly Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys

35 40 45
 Arg Leu Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro

 50 55 60
 Ser Arg Gly Arg Ser Leu Gln Asn Ser Ala Gly Arg Met Asp Glu Ser
 65 70 75 80
 Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu Ala Ala Gly
 85 90 95
 Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
 100 105 110
 Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser Gln

 115 120 125
 Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val
 130 135 140
 His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Tyr
 145 150 155 160
 Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
 165 170 175
 Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser Ile

 180 185 190
 Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 195 200 205
 Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val Ser
 210 215 220
 Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 225 230 235 240
 Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro

 245 250 255
 Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr Val
 260 265 270
 Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val Met
 275 280 285

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser

290

295

300

Leu Gly

305

<210> 65

<211> 6

<212

> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 65

Thr Glu Gly Arg Met Asp

1

5

<210> 66

<211> 305

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 66

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala

1

5

10

15

Tyr Ser Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr

20

25

30

Cys Gly Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys

35

40

45

Arg Leu Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro

50

55

60

Ser Arg Gly Arg Ser Pro Gln Thr Glu Gly Arg Met Asp Glu Ser Lys

65

70

75

80

Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu Ala Ala Gly Gly

85

90

95

Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile

100

105

110

Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser Gln Glu
115 120 125
Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val His
130 135 140
Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg
145 150 155 160
Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
165 170 175
Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu
180 185 190
Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr
195 200 205
Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val Ser Leu
210 215 220
Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp
225 230 235 240
Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val
245 250 255
Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp
260 265 270
Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val Met His
275 280 285
Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu
290 295 300
Gly
305
<210> 67
<211> 8
<212> PRT
<213> Artificial Sequence
<220><223> Linker
<400> 67

Thr Ala Glu Ala Ala Ala Lys Ala

1 5

<210> 68

<211> 307

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 68

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala

1 5 10 15

Tyr Ser Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr

20 25 30

Cys Gly Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys

35 40 45

Arg Leu Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro

50 55 60

Ser Arg Gly Arg Ser Pro Gln Thr Ala Glu Ala Ala Ala Lys Ala Glu

65 70 75 80

Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro Glu Ala Ala

85 90 95

Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu

100 105 110

Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser

115 120 125

Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu

130 135 140

Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr

145 150 155 160

Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn

165 170 175

Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser

180 185 190

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln
195 200 205
Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val
210 215 220
Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val
225 230 235 240
Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro
245 250 255
Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr
260 265 270
Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val
275 280 285
Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
290 295 300
Ser Leu Gly
305
<210> 69
<211> 304
<212> PRT
<213> Artificial Sequence
<220><223> Fusion protein
<400> 69
Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15
Tyr Ser Cys Pro Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr
20 25 30
Cys Gly Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys
35 40 45
Arg Leu Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro
50 55 60
Ser Arg Gly Arg Ser Pro Gln Thr Glu Gly Arg Met Asp Glu Arg Lys
65 70 75 80

Cys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro Pro Val Ala Gly Pro

85 90 95

Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser

100 105 110

Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser Gln Glu Asp

115 120 125

Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn

130 135 140

Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg Val

145 150 155 160

Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu

165 170 175

Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys

180 185 190

Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr

195 200 205

Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr

210 215 220

Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu

225 230 235 240

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

245 250 255

Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys

260 265 270

Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val Met His Glu

275 280 285

Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly

290 295 300

<210> 70

<211> 294

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 70

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala

1 5 10 15

Tyr Ser Cys Ala Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr

20 25 30

Cys Gly Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys

35 40 45

Arg Leu Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Glu Ala

50 55 60

Ala Ala Lys Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala

65 70 75 80

Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys

85 90 95

Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val

100 105 110

Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp

115 120 125

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

130 135 140

Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp

145 150 155 160

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

165 170 175

Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg

180 185 190

Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys

195 200 205

Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp

210 215 220

Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
 225 230 235 240
 Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
 245 250 255
 Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser
 260 265 270
 Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
 275 280 285

Leu Ser Leu Ser Leu Gly
 290

<210> 71

<211> 891

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 71

gccaccatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgcctactct 60
 tgcgccgagt ggctacagc ttggggccct tgttctacca cctgtggcct cggcatggcc 120
 accagagtgt ccaaccagaa cagattctgc cgcttgaaa ccagcggag actgtgtttg 180
 tccagacctt gcgaggccgc tgccaaagaa agaaagtgt gcgtggaatg cctccttgt 240
 cctgtctctc ctgtggctgg ccttccgtg tttctgttcc ctcaaagcc taaggacacc 300

ctgatgatct ctcgacccc tgaagtgacc tgcgtggtgg tggatgtgtc ccaagaggat 360
 cccgagggtg agttcaattg gtacgtggac ggctggaag tgcacaacgc caagaccaag 420
 cctagagagg aacagttcaa ctccacctac agagtgtgt cctgtctgac cgtgtgtcac 480
 caggattggc tgaacggcaa agagtacaag tgcaaggtgt ccaacaaggg actgccctcc 540
 agcatcgaaa agaccatctc caaggccaag ggacagccca gagaaccca ggtgtacaca 600
 ctgcctcaa gccaaagga aatgaccaag aaccaggtgt cctgacctg cctggtaag 660
 ggcttctacc cttccgatat cgccgtggaa tgggagtcca atggccagcc tgagaacaac 720

tacaagacca cacctccagt gctggactcc gacggctcct tctttctgta ctccgcctg 780
 accgtggaca agtccagatg gcaagagggc aacgtgttct cctgtccgt gatgcacgag 840
 gccctgcaca atcactacac ccagaagtcc ctgtctctgt ccctgggcta a 891

<210> 72

<211> 964

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 72

```

ggggacaagt ttgtacaaa aagcaggcta tggtagcgcc accatgaaat gggtcacctt      60
tatctccctg ctgttctgt tctctccgc ctactcttgc gccgagtggc ctacagcttg      120
gggcccttgt tctaccacct gtggcctcgg catggccacc agagtgtcca accagaacag      180

attctgccgg ctggaaaccc agcggagact gtgtttgtcc agaccttgcg aggccgctgc      240
caaagaaaga aagtgtctgc tggaaatgcc tccttgtcct gctcctctg tggctggccc      300
ttccgtgttt ctgttcctc caaagcctaa ggacacctg atgatctctc ggacctctga      360
agtgcactgc gtggtggcgg atgtgtccca agaggatccc gaggtgcagt tcaattggta      420
cgtggacggc gtggaagtgc acaacccaa gaccaagcct agagaggaac agttcaactc      480
cacctacaga gtggtgtcgg tgctgacctg gctgcaccag gattggctga acggcaaaga      540
gtacaagtgc aaggtgtcca acaagggact gccctccagc atcgaaaaga ccatctccaa      600

ggccaagggg cagcccagag aaccccaggt gtacacactg cctccaagcc aagaggaaat      660
gaccaagaac caggtgtccc tgacctgcct ggtcaagggc ttctaccctt ccgatatcgc      720
cgtggaatgg gagtccaatg gccagcctga gaacaactac aagaccacac ctccagtgtc      780
ggactccgac ggctccttct ttctgtactc ccgctgacc gtggacaagt ccagatggca      840
agagggcaac gtgttctcct gctccgtgat gcacgaggcc ctgcacaatc actacacca      900
gaagtccttg tctctgtccc tgggctaata tagaaacca gctttcttgt acaaagtgg      960
cccc

```

<210> 73

<211> 276

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 73

```

Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly
1           5           10           15

Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met
          20           25           30

```

Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala Ala Ala
 35 40 45
 Lys Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro Pro
 50 55 60
 Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr
 65 70 75 80
 Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val
 85 90 95
 Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val
 100 105 110
 Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser
 115 120 125
 Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu
 130 135 140
 Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro Ser
 145 150 155 160
 Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
 165 170 175
 Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln
 180 185 190
 Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
 195 200 205
 Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
 210 215 220
 Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu
 225 230 235 240
 Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser Cys Ser
 245 250 255
 Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
 260 265 270
 Leu Ser Leu Gly

275

<210> 74

<211> 294

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 74

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala

1 5 10 15

Tyr Ser Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser

20 25 30

Cys Gly Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys

35 40 45

Glu Met Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala

50 55 60

Ala Ala Lys Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala

65 70 75 80

Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys

85 90 95

Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val

100 105 110

Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp

115 120 125

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

130 135 140

Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp

145 150 155 160

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

165 170 175

Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg

180 185 190

Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys

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

210 215 220
Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
225 230 235 240
Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
245 250 255
Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser
260 265 270
Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser

275 280 285
Leu Ser Leu Ser Leu Gly

290

<210> 75

<211> 891

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 75

gccaccatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgcctactct 60
tgtgccgagc agaccacaga gtggaccgcc tgctctaagt cttgcggcat gggtcttctcc 120
accagagtga ccaaccggaa cagacagtgc gagatgctga agcagaccg gctgtgtatg 180
gttcgacctt gcgaggccgc tgccaaagaa agaaagtgt gcgtggaatg ccctccttgt 240

cctgtctctc ctgtggctgg ccttccgtg tttctgttcc ctccaaagcc taaggacacc 300
ctgatgatct ctcgacccc tgaagtgacc tgcgtggtgg tggatgtgtc ccaagaggat 360
cccaggtgac agttcaattg gtacgtggac ggcgtggaag tgcacaacgc caagaccaag 420
cctagagagg aacagttcaa ctccacctac agagtgtgt cctgtctgac cgtgtctgac 480
caggattggc tgaacggcaa agagtacaag tgcaaggtgt ccaacaaggg cctgccttcc 540
agcatcgaaa agaccatctc caaggccaag ggacagccca gagaaccca ggtgtacaca 600
ctgcctccaa gccaaagagga aatgaccaag aaccaggtgt ccctgacctg cctggtcaag 660

ggcttctacc ctccgatat cgccgtggaa tgggagtcca atggccagcc tgagaacaac 720

tacaagacca cacctccagt gctggactcc gacggctcct tctttctgta ctcccgctg 780

accgtggaca agtccagatg gcaagagggc aacgtgttct cctgctccgt gatgcacgag 840

gccctgcaca atcactacac ccagaagtcc ctgtctctgt ccctgggcta a 891

<210> 76

<211> 964

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 76

ggggacaagt ttgtacaaa aagcaggcta tggtagcgcc accatgaaat gggtcacctt 60

tatctccctg ctgttctgt tctctccgc ctactctgt gccgagcaga ccacagagtg 120

gaccgcctgc tctaagcttt gcggcatggg cttctccacc agagtgacca accggaacag 180

acagtgcgag atgctgaagc agaccggct gtgtatggtt cgaccttgcg aggccgctgc 240

caaagaaaga aagtgtcgc tggaatgcc tcttgtctt gctcctctg tggctggccc 300

ttccgtgttt ctgttccctc caaagcctaa ggacaccctg atgatctctc ggaccctga 360

agtgacctgc gtggtggigg atgtgtcca agaggatccc gaggtgcagt tcaattggta 420

cgtggacggc gtggaagtgc acaacgcaa gaccaagcct agagaggaac agttcaactc 480

cacctacaga gtggtgtccg tgtgaccgt gctgcaccag gattggctga acggcaaaga 540

gtacaagtgc aaggtgtcca acaagggcct gccttccagc atcgaaaaga ccatttccaa 600

ggccaaggga cagcccagag aaccccaggt gtacacactg cctccaagcc aagaggaaat 660

gaccaagaac caggtgtccc tgacctgcct ggtcaagggc ttctaccctt ccgatatcgc 720

cgtggaatgg gagtccaatg gccagcctga gaacaactac aagaccacac ctccagtgt 780

ggactccgac ggctccttct ttctgtactc ccgctgacc gtggacaagt ccagatggca 840

agagggcaac gtgttctct gctccgtgat gcacgaggcc ctgcacaatc actacacca 900

gaagtcctg tctctgtccc tgggctaatc tagaaacca gctttcttgt acaaagtgt 960

cccc 964

<210> 77

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 77

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15
Tyr Ser Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro
20 25 30
Glu Ala Ala Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
35 40 45
Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val
50 55 60
Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp
65 70 75 80
Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe
85 90 95
Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp
100 105 110
Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu
115 120 125
Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
130 135 140
Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys
145 150 155 160
Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
165 170 175
Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
180 185 190
Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
195 200 205
Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser
210 215 220
Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
225 230 235 240
Leu Ser Leu Ser Leu Gly Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala

245 250 255
 Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
 260 265 270
 Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Ala
 275 280 285
 Ala Ala Cys Ala Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr

290 295 300
 Cys Gly Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys
 305 310 315 320
 Arg Leu Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro
 325 330 335
 Ser Arg Gly Arg Ser Pro Gln

340

<210> 78

<211> 1038

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 78

gccacatga aatgggtcac ctttatctcc ctgctgttcc tgttctccag cgcctactcc 60

gagtctaagt acggccctcc ttgtctcca tgtctgctc cagaagctgc tggcggccct 120

tccgtgtttc tgttccctcc aaagcctaag gacacctga tgatctctcg gaccttgaa 180

gtgacctgcg tgggtgtgga tgtgtccaa gaggatcccg aggtgcagtt caattgttac 240

gtggacggcg tggaagtga caacgccaag accaagccta gagaggaaca gttcaactcc 300

acctacagag tgggtgtccgt gctgacctg ctgcaccagg attggctgaa cggcaaagag 360

tacaagtga aggtgtccaa caaggcctg cttccagca tcgaaaagac catctccaag 420

gccaagggcc agcctaggga accccagggt tacacctgc ctccaagcca agaggaaatg 480

accaagaacc aggtgtccct gacctgcctg gtcaagggt tctacccttc cgatatcgcc 540

gtggaatggg agagcaatgg ccagcctgag aacaactaca agaccacacc tctgtgctg 600

gactccgacg gctccttctt tctgtactcc cgcctgaccg tggacaagtc cagatggcaa 660

gagggcaacg tgttctctg ctccgtgatg cagaggccc tgcacaatca ctaccccag 720

aagtcctgt ctctgtccct gggagctgag gccgtgcta aagaagctgc cgctaaagag 780

gccgcagcca aagaggcagc cgccaaagaa gccgctgcaa aagaggctgc tgcaaaagaa 840
gcagcagcta aagaagctgc tgccaaggcc gctgcttgtg ccgaatggtc tacagcttgg 900

ggcccttgcct ctaccacctg tggactcggc atggccacca gagtgtctaa ccagaacaga 960
ttctgccggc tggaaaccca gcggagactg tgctgtctta gacctgtcc tcctagcaga 1020
ggcagatccc ctcaagtga 1038

<210> 79
<211> 1111
<212> DNA
<213> Artificial Sequence
<220><223> Fusion protein
<400> 79

ggggacaagt ttgtacaaaa aagcaggcta tggtagcgcc accatgaaat gggtcacctt 60
tatctccctg ctgttcctgt tctccagcgc ctactccgag tctaagtacg gccctccttg 120
tcctccatgt cctgctccag aagctgctgg cggcccttcc gtgtttctgt tccctccaaa 180

gcctaaggac accctgatga tctctcggac ccctgaagtg acctgcgtgg tggtaggatgt 240
gtccaagag gatccccagg tgcagttcaa ttggtacgtg gacggcgtgg aagtgcacaa 300
cgccaagacc aagcctagag aggaacagtt caactccacc tacagagtgg tgtccgtgct 360
gacctgctg caccaggatt ggctgaacgg caaagagtac aagtgcagg tgtccaacaa 420
gggcctgcct tccagcatcg aaaagacat ctccaaggcc aagggccagc ctagggaacc 480
ccaggtttac accctgcctc caagccaaga ggaaatgacc aagaaccagg tgtccctgac 540
ctgcctggtc aagggttct acccttcga tatcgccgtg gaatgggaga gcaatggcca 600

gcctgagaac aactacaaga ccacacctcc tgtgtggac tccgacggt cttttttct 660
gtactccgc ctgacctgg acaagtccag atggcaagag ggcaacgtgt tctcctgctc 720
cgtgatgcac gaggccctgc acaatcacta caccagaag tccctgtctc tgtccctggg 780
agctgaggcc gctgctaaag aagctgccgc taaagaggcc gcagccaaag aggcagccgc 840
caaagaagcc gctgcaaaag aggtgctgc aaaagaagca gcagctaaag aagctgctgc 900
caaggccgt gcttgtgccg aatgtgtac agcttggggc ccttgtctta ccacctgtg 960
actcgcatg gccaccagag tgtctaacca gaacagattc tgccggtgg aaaccagcg 1020

gagactgtgc ctgtctagac cctgtcctcc tagcagaggc agatcccctc agtgatctag 1080
aaaccagct ttcttgtaca aagtggtecc c 1111

<210> 80

<211> 311

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 80

Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly

1 5 10 15

Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met

20 25 30

Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala Ala Ala

35 40 45

Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

50 55 60

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu

65 70 75 80

Ala Ala Ala Lys Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro

85 90 95

Ala Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro

100 105 110

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

115 120 125

Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val

130 135 140

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

145 150 155 160

Phe Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln

165 170 175

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

180 185 190

Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro

195 200 205

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr

[illegible]

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

290 295 300

Ser Leu Ser Leu Ser Leu Gly

305 310

<210> 81

<211> 329

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 81

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala

1 5 10 15

Tyr Ser Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser
20 25 30

Cys Gly Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys

35 40 45

Glu Met Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala
50 55 60

Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala
65 70 75 80

Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala

85 90 95

Lys Glu Ala Ala Ala Lys Glu Arg Lys Cys Cys Val Glu Cys Pro Pro
 100 105 110
 Cys Pro Ala Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro
 115 120 125
 Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys
 130 135 140
 Val Val Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp
 145 150 155 160

 Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu
 165 170 175
 Glu Gln Phe Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu
 180 185 190
 His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn
 195 200 205
 Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly
 210 215 220

 Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu
 225 230 235 240
 Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr
 245 250 255
 Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn
 260 265 270
 Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe
 275 280 285

 Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn
 290 295 300
 Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr
 305 310 315 320
 Gln Lys Ser Leu Ser Leu Ser Leu Gly
 325

<210> 82

<211> 996

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 82

gccacatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgctactct 60

tgtgccgagc agaccacaga gtggaccgcc tgccttaagt cttgcggcat gggcttctcc 120

accagagtga ccaaccggaa cagacagtgc gagatgctga agcagaccg gctgtgtatg 180

gttcgacctt gcgaggccgc tgccaaagag gctgctgcta aagaagccgc cgcaaaagag 240

gcagcagcaa aagaggctgc cgccaaagag gccgcagcca aagaagcagc agctaaagag 300

gccgctgcaa aagaacggaa gtgctgcgtg gaatgcctc cttgtcctgc tctcctgtg 360

gctggccctt ccgtgtttct gttccctcca aagcctaagg acaccctgat gatctctcg 420

acccctgaag tgacctgcgt ggtggtggat gtgtccaag aggatccga ggtgcagttc 480

aattggtacg tggacggcgt ggaagtgcac aacccaaga ccaagcctag agaggaacag 540

ttcaactcca cctacagagt ggtgtccgtg ctgacctgc tgcaccagga ttggtgaac 600

ggcaaagagt acaagtgcaa ggtgtccaac aaggccctgc cttccagcat cgaaaagacc 660

atctccaagg ccaagggaca gccagagaa cccaggtgt acacactgcc tccaagccaa 720

gaggaaatga ccaagaacca ggtgtccctg acctgcctgg tcaagggtt ctaccttcc 780

gatatcgccg tggaatggga gtccaatggc cagcctgaga acaactaaa gaccacact 840

ccagtgtgg actccgacgg ctcttcttt ctgtactccc gctgaccgt ggacaagtc 900

agatggcaag agggcaactg gttctcctgc tccgtgatgc acgaggccct gcacaatcac 960

tacaccaga agtccctgtc tctgtccctg ggctaa 996

<210> 83

<211> 1069

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 83

ggggacaagt ttgtacaaaa aagcaggcta tggtagccg accatgaaat gggtcacctt 60

tatctccctg ctgttctgt tctcctccgc ctactcttgt gccgagcaga ccacagagtg 120

gaccgcctgc tetaagtctt gcggcatggg cttctccacc agagtgacca accggaacag 180

acagtgcgag atgtgaagc agaccggct gtgtatggtt cgaccttgcg aggccgctgc 240

caaagaggct gctgctaaag aagccgccgc aaaagaggca gcagcaaaag aggctgccgc 300
 caaagaggcc gcagccaaag aagcagcagc taaagaggcc gctgcaaaag aacggaagtg 360
 ctgctggaa tgcctcctt gtectgtcc tctgtggct ggcccttccg tgtttctgtt 420
 ccctccaaag cctaaggaca ccctgatgat ctctcggacc cctgaagtga cctgcgtggt 480
 ggtggatgtg tccaagagg atcccaggt gcagttcaat tggtagctgg acggcgtgga 540
 agtgcacaac gccaagacca agcctagaga ggaacagttc aactccacct acagagtgtt 600
 gtccgtgctg accgtgctgc accaggattg gctgaacggc aaagagtaca agtgcaaggt 660

 gtccaacaag ggctgcctt ccagcatcga aaagaccatc tccaaggcca agggacagcc 720
 cagagaaccc caggtgtaca cactgcctcc aagccaagag gaaatgacca agaaccaggt 780
 gtccctgacc tgcttggtca agggcttcta cccttccgat atgccgtgg aatgggagtc 840
 caatggccag cctgagaaca actacaagac cacacctcca gtgctggact ccgacggctc 900
 ctcttttctg tactccgcc tgaccgtgga caagtccaga tggcaagagg gcaacgtgtt 960
 ctctgtctcc gtgatgcag aggcctgca caatcactac acccagaagt ccctgtctct 1020
 gtccctgggc taatctagaa acccagcttt ctgtacaaa gtggtcccc 1069

<210> 84

<211> 276

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 84

Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly

1 5 10 15

Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met

20 25 30

Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala Ala Ala

35 40 45

Lys Glu Arg Lys Gln Gln Val Glu Gln Pro Pro Gln Pro Ala Pro Pro

50 55 60

Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr

65 70 75 80

Leu Tyr Ile Thr Arg Glu Pro Glu Val Thr Cys Val Val Val Asp Val

85 90 95

Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly Val
100 105 110

Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser
115 120 125

Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu
130 135 140

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

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

Gln Val Tyr Thr Phe Pro Pro Ser Gln Glu Glu Met Thr Lys Asn Gln
180 185 190

Val Ser Leu Arg Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
195 200 205

Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
210 215 220

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

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

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

Leu Ser Leu Gly
275

<210> 85
<211> 294
<212> PRT
<213> Artificial Sequence
<220><223> Fusion protein
<400> 85

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser
 20 25 30
 Cys Gly Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys
 35 40 45
 Glu Met Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala
 50 55 60
 Ala Ala Lys Glu Arg Lys Gln Gln Val Glu Gln Pro Pro Gln Pro Ala
 65 70 75 80
 Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
 85 90 95
 Asp Thr Leu Tyr Ile Thr Arg Glu Pro Glu Val Thr Cys Val Val Val
 100 105 110
 Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp
 115 120 125
 Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe
 130 135 140
 Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp
 145 150 155 160
 Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu
 165 170 175
 Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
 180 185 190
 Glu Pro Gln Val Tyr Thr Phe Pro Pro Ser Gln Glu Glu Met Thr Lys
 195 200 205
 Asn Gln Val Ser Leu Arg Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
 210 215 220
 Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
 225 230 235 240
 Thr Thr Lys Pro Val Leu Asp Ser Asp Gly Ser Phe Arg Leu Glu Ser
 245 250 255
 Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser

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

275 280 285
Leu Ser Leu Ser Leu Gly

290

<210> 86
<211> 891
<212> DNA
<213> Artificial Sequence
<220><223> Fusion protein
<400> 86

gccaccatga aatgggtcac ctttatctcc ctgctgttcc tgttctcctc cgcctactct 60
tgtgccgagc agaccacaga gtggaccgcc tgctctaagt cttgcggcat gggttctctc 120
accagagtga ccaaccggaa cagacagtgc gagatgctga agcagaccgc gctgtgtatg 180
gttcgacctt gcgaggccgc tgccaaagaa agaaagcagc aggtcgagca gcctcctcag 240

cctgctcctc ctgttgctgg ccttccgtg tttctgttcc ctccaaagcc taaggacacc 300
ctgtatcatc cccgcgagcc tgaagtgacc tgcgtggtgg tggatgtgtc ccaagaggat 360
cccagagtgc agttcaattg gtacgtggac ggcgtggaag tgcacaacgc caagaccaag 420
cctagagagg aacagttcaa ctccacctac agagtgggtg ccgtgctgac cgtgctgcac 480
caggattggc tgaacggcaa agagtacaag tgcaaggtgt ccaacaaggg cctgccttcc 540
agcatcgaaa agaccatctc caaggccaag ggacagccca gagaaccca ggtgtacaca 600
ttccctccat ctcaagagga aatgaccaag aaccaggtgt ccctgcggtg cctggtcaag 660

ggcttctacc ctctgatat cgccgtggaa tgggagtcca acggccagcc tgagaacaac 720
tacaagacca ccaagcctgt gctggactcc gacggctcct tccggttga atctagactg 780
accgtggaca agtcccgtg gcaagagggc aacgtgttct cctgctctgt gatgcacgag 840
gccctgcaca accactacac ccagaagtcc ctgtctctgt ccctgggcta a 891

<210> 87
<211> 964
<212> DNA
<213> Artificial Sequence
<220><223> Fusion protein
<400> 87

ggggacaagt ttgtacaaaa aagcaggcta tggtagcgcc accatgaaat gggtcacctt	60
tatctccctg ctgttcctgt tctctccgc ctactcttgt gccgagcaga ccacagagtg	120
gaccgcctgc tctaagtcctt gcggcatggg cttctccacc agagtgacca accggaacag	180
acagtgcgag atgctgaagc agaccggct gtgtatggtt cgaccttgcg aggccgctgc	240
caaagaaaga aagcagcagg tcgagcagcc tctcagcct gctcctctg ttgctggccc	300
ttcctgtttt ctgttcctc caaagcctaa ggacaccctg tacatcacc gcgagcctga	360
agtgcacctgc gtggtggagg atgtgtccca agaggatccc gaggtgcagt tcaattggta	420
cgtggacggc gtggaagtgc acaacgcaa gaccaagcct agagaggaac agttcaactc	480
cacctacaga gtggtgtccg tgctgaccgt gctgcaccag gattggctga acggcaaaga	540
gtacaagtgc aaggtgtcca acaagggcct gccttcacc atcgaaaaga ccatctccaa	600
ggccaaggga cagcccagag aaccccaggt gtacacattc cctccatctc aagaggaaat	660
gaccaagaac caggtgtccc tgcggtgcct ggtcaagggc ttctaccctt ctgatatcgc	720
cgtggaatgg gagtccaacg gccagcctga gaacaactac aagaccacca agcctgtgct	780
ggactccgac ggctccttc ggcttgaatc tagactgacc gtggacaagt cccggtggca	840
agagggcaac gtgttctct gctctgtgat gcacgaggcc ctgcacaacc actacacca	900
gaagtcctg tctctgtccc tgggctaatc tagaaacca gctttcttgt acaaagtgtt	960
cccc	964
<210> 88	
<211> 281	
<212> PRT	
<213> Artificial Sequence	
<220><223> Fusion protein	
<400> 88	
Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly	
1 5 10 15	
Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met	
20 25 30	
Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala Ala Ala	
35 40 45	
Lys Glu Pro Lys Ser Gln Asp Lys Thr His Thr Gln Pro Pro Gln Pro	
50 55 60	

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
65 70 75 80
Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
85 90 95
Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
100 105 110
Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Cys Glu Glu
115 120 125
Gln Tyr Gly Ser Thr Tyr Arg Cys Val Ser Val Leu Thr Val Leu His
130 135 140
Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
145 150 155 160
Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
165 170 175
Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu
180 185 190
Thr Lys Asn Gln Val Ser Leu Arg Cys His Val Lys Gly Phe Tyr Pro
195 200 205
Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
210 215 220
Tyr Lys Thr Thr Lys Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
225 230 235 240
Tyr Ser Thr Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
245 250 255
Phe Ser Cys Ser Val Leu His Glu Ala Leu His Asn His Tyr Thr Gln
260 265 270
Lys Ser Leu Ser Leu Ser Pro Gly Lys
275 280
<210> 89
<211> 299
<212> PRT
<213> Artificial Sequence

<220><223> Fusion protein

<400> 89

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala

1 5 10 15

Tyr Ser Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser

20 25 30

Cys Gly Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys

35 40 45

Glu Met Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala

50 55 60

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

65 70 75 80

Gln Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro

85 90 95

Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr

100 105 110

Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn

115 120 125

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

130 135 140

Glu Glu Gln Tyr Gly Ser Thr Tyr Arg Cys Val Ser Val Leu Thr Val

145 150 155 160

Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser

165 170 175

Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys

180 185 190

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp

195 200 205

Glu Leu Thr Lys Asn Gln Val Ser Leu Arg Cys His Val Lys Gly Phe

210 215 220

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

225	230	235	240
Asn Asn Tyr Lys Thr Thr Lys Pro Val Leu Asp Ser Asp Gly Ser Phe			
	245	250	255
Phe Leu Tyr Ser Thr Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly			
	260	265	270
Asn Val Phe Ser Cys Ser Val Leu His Glu Ala Leu His Asn His Tyr			
	275	280	285
Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys			
	290	295	
<210> 90			
<211> 906			
<212> DNA			
<213> Artificial Sequence			
<220><223> Fusion protein			
<400> 90			
gccacatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgctactct	60		
tgtgccgagc agaccacaga gtggaccgcc tgctctaagt cttgcggcat gggcttctcc	120		
accagagtga ccaaccggaa cagacagtgc gagatgctga agcagaccgc gctgtgtatg	180		
gttcgacctt gcgaggccgc tgccaaagag cctaagagcc aggacaagac ccacacacag	240		
cctccacagc ctgctccaga attgctcgga ggcccttccg tgtttctgtt ccctccaaag	300		
cctaaggaca cctgatgat ctctcggacc cctgaagtga cctgcgtggt ggtggatgtg	360		
tctcacgagg atcccgaagt gaagtccaat tggtagctgg acggcgtgga agtcacaaac	420		
gccaagacaa agccctgcga ggaacagtac ggctccacct acagatgcgt gtccgtgctg	480		
acagtgtgc accaggattg gctgaacggc aaagagtaca agtgcaaggt gtccaacaag	540		
gccctgcctg ctctatcga aaagaccatc tccaaggcca agggccagcc tagagaaccc	600		
caggtgtaca cactgccacc ttctagggac gagctgacca agaaccaggt gtccctgaga	660		
tgccactgta agggcttcta ccctccgat atgccgtgg aatgggagtc taatggacag	720		
cccgagaaca actacaagac caccaagcct gtgctggact ccgacggctc cttcttctctg	780		
tactctaccc tgaccgtgga caagtccaga tggcagcagg gcaacgtgtt ctctgctct	840		
gtgctgcacg aggcctgca caatcactac acccagaagt ccctgtctct gtccctggc	900		
aagtga	906		

<210> 91

<211> 979

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 91

```

ggggacaagt ttgtacaaaa aagcaggcta tggtagcgcc accatgaaat gggtcacctt      60
tatctccctg ctgttctctg tctcctccgc ctactcttgt gccgagcaga ccacagagtg      120
gaccgcctgc tctaagtctt gcggcatggg cttctccacc agagtgacca accggaacag      180
acagtgcgag atgctgaagc agaccgggct gtgtatgggt cgaccttgcg aggccgctgc      240
caaagagcct aagagccagg acaagacca cacacagcct ccacagcctg ctccagaatt      300

gctcggaggc ccttccgtgt ttctgttccc tccaaagcct aaggacaccc tgatgatctc      360
tcggaccctt gaagtgaact gcgtgggtgt gtagtgtgtt cagcaggatc ccgaagtga      420
gttcaattgg tacgtggacg gcgtggaagt gcacaacgcc aagacaaagc cctgcgagga      480
acagtacggc tccacctaca gatgcgtgtc cgtgctgaca gtgctgcacc aggattggct      540
gaacggcaaa gactacaagt gcaagggtgc caacaaggcc ctgcctgctc ctatcgaaaa      600
gacctctcc aaggccaagg gccagcctag agaaccacag gtgtacacac tgccaccttc      660
tagggacgag ctgaccaaga accagggtgc cctgagatgc cacgtgaagg gcttctaccc      720

ctccgatatc gccgtggaat gggagtctaa tggacagccc gagaacaact acaagaccac      780
caagcctgtg ctggactccg acggctcctt cttcctgtac tctaccctga ccgtggacaa      840
gtccagatgg cagcagggca acgtgttctc ctgctctgtg ctgcacgagg ccctgcacaa      900
tactacacc cagaagtccc tgtctctgtc ccctggcaag tgatctagaa acccagcttt      960
cttgtacaaa gtggtcccc                                     979

```

<210> 92

<211> 402

<212> PRT

<213> Artificial Sequence

<220><223> Fusion tag

<400> 92

Gly His His His His His His Gly Ser Glu Ile Gly Thr Gly Phe Pro

1 5 10 15

Phe Asp Pro His Tyr Val Glu Val Leu Gly Glu Arg Met His Tyr Val

20 25 30
 Asp Val Gly Pro Arg Asp Gly Thr Pro Val Leu Phe Leu His Gly Asn
 35 40 45
 Pro Thr Ser Ser Tyr Val Trp Arg Asn Ile Ile Pro His Val Ala Pro
 50 55 60
 Thr His Arg Cys Ile Ala Pro Asp Leu Ile Gly Met Gly Lys Ser Asp

 65 70 75 80
 Lys Pro Asp Leu Gly Tyr Phe Phe Asp Asp His Val Arg Phe Met Asp
 85 90 95
 Ala Phe Ile Glu Ala Leu Gly Leu Glu Glu Val Val Leu Val Ile His
 100 105 110
 Asp Trp Gly Ser Ala Leu Gly Phe His Trp Ala Lys Arg Asn Pro Glu
 115 120 125
 Arg Val Lys Gly Ile Ala Phe Met Glu Phe Ile Arg Pro Ile Pro Thr

 130 135 140
 Trp Asp Glu Trp Pro Glu Phe Ala Arg Glu Thr Phe Gln Ala Phe Arg
 145 150 155 160
 Thr Thr Asp Val Gly Arg Lys Leu Ile Ile Asp Gln Asn Val Phe Ile
 165 170 175
 Glu Gly Thr Leu Pro Met Gly Val Val Arg Pro Leu Thr Glu Val Glu
 180 185 190
 Met Asp His Tyr Arg Glu Pro Phe Leu Asn Pro Val Asp Arg Glu Pro

 195 200 205
 Leu Trp Arg Phe Pro Asn Glu Leu Pro Ile Ala Gly Glu Pro Ala Asn
 210 215 220
 Ile Val Ala Leu Val Glu Glu Tyr Met Asp Trp Leu His Gln Ser Pro
 225 230 235 240
 Val Pro Lys Leu Leu Phe Trp Gly Thr Pro Gly Val Leu Ile Pro Pro
 245 250 255
 Ala Glu Ala Ala Arg Leu Ala Lys Ser Leu Pro Asn Cys Lys Ala Val

 260 265 270

Asp Ile Gly Pro Gly Leu Asn Leu Leu Gln Glu Asp Asn Pro Asp Leu
275 280 285
Ile Gly Ser Glu Ile Ala Arg Trp Leu Ser Thr Leu Glu Ile Ser Gly
290 295 300
Leu Gln Asp Ser Glu Val Asn Gln Glu Ala Lys Pro Glu Val Lys Pro
305 310 315 320
Glu Val Lys Pro Glu Thr His Ile Asn Leu Lys Val Ser Asp Gly Ser
325 330 335
Ser Glu Ile Phe Phe Lys Ile Lys Lys Thr Thr Pro Leu Arg Arg Leu
340 345 350
Met Glu Ala Phe Ala Lys Arg Gln Gly Lys Glu Met Asp Ser Leu Thr
355 360 365
Phe Leu Tyr Asp Gly Ile Glu Ile Gln Ala Asp Gln Thr Pro Glu Asp
370 375 380
Leu Asp Met Glu Asp Asn Asp Ile Ile Glu Ala His Arg Glu Gln Ile
385 390 395 400
Gly Gly

<210> 93

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 93

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Ser Gly Gly
1 5 10 15

Gly Ser Ser Gly

20

<210> 94

<211> 466

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 94

Gly His His His His His His Gly Ser Glu Ile Gly Thr Gly Phe Pro

1 5 10 15
Phe Asp Pro His Tyr Val Glu Val Leu Gly Glu Arg Met His Tyr Val
20 25 30
Asp Val Gly Pro Arg Asp Gly Thr Pro Val Leu Phe Leu His Gly Asn
35 40 45
Pro Thr Ser Ser Tyr Val Trp Arg Asn Ile Ile Pro His Val Ala Pro
50 55 60
Thr His Arg Cys Ile Ala Pro Asp Leu Ile Gly Met Gly Lys Ser Asp

65 70 75 80
Lys Pro Asp Leu Gly Tyr Phe Phe Asp Asp His Val Arg Phe Met Asp
85 90 95
Ala Phe Ile Glu Ala Leu Gly Leu Glu Glu Val Val Leu Val Ile His
100 105 110
Asp Trp Gly Ser Ala Leu Gly Phe His Trp Ala Lys Arg Asn Pro Glu
115 120 125
Arg Val Lys Gly Ile Ala Phe Met Glu Phe Ile Arg Pro Ile Pro Thr

130 135 140
Trp Asp Glu Trp Pro Glu Phe Ala Arg Glu Thr Phe Gln Ala Phe Arg
145 150 155 160
Thr Thr Asp Val Gly Arg Lys Leu Ile Ile Asp Gln Asn Val Phe Ile
165 170 175
Glu Gly Thr Leu Pro Met Gly Val Val Arg Pro Leu Thr Glu Val Glu
180 185 190
Met Asp His Tyr Arg Glu Pro Phe Leu Asn Pro Val Asp Arg Glu Pro

195 200 205
Leu Trp Arg Phe Pro Asn Glu Leu Pro Ile Ala Gly Glu Pro Ala Asn
210 215 220
Ile Val Ala Leu Val Glu Glu Tyr Met Asp Trp Leu His Gln Ser Pro
225 230 235 240

Val Pro Lys Leu Leu Phe Trp Gly Thr Pro Gly Val Leu Ile Pro Pro

245 250 255

Ala Glu Ala Ala Arg Leu Ala Lys Ser Leu Pro Asn Cys Lys Ala Val

260 265 270

Asp Ile Gly Pro Gly Leu Asn Leu Leu Gln Glu Asp Asn Pro Asp Leu

275 280 285

Ile Gly Ser Glu Ile Ala Arg Trp Leu Ser Thr Leu Glu Ile Ser Gly

290 295 300

Leu Gln Asp Ser Glu Val Asn Gln Glu Ala Lys Pro Glu Val Lys Pro

305 310 315 320

Glu Val Lys Pro Glu Thr His Ile Asn Leu Lys Val Ser Asp Gly Ser

325 330 335

Ser Glu Ile Phe Phe Lys Ile Lys Lys Thr Thr Pro Leu Arg Arg Leu

340 345 350

Met Glu Ala Phe Ala Lys Arg Gln Gly Lys Glu Met Asp Ser Leu Thr

355 360 365

Phe Leu Tyr Asp Gly Ile Glu Ile Gln Ala Asp Gln Thr Pro Glu Asp

370 375 380

Leu Asp Met Glu Asp Asn Asp Ile Ile Glu Ala His Arg Glu Gln Ile

385 390 395 400

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

405 410 415

Gly Gly Gly Ser Ser Gly Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala

420 425 430

Cys Ser Lys Ser Cys Gly Met Gly Phe Ser Thr Arg Val Thr Asn Arg

435 440 445

Asn Arg Gln Cys Glu Met Leu Lys Gln Thr Arg Leu Cys Met Val Arg

450 455 460

Pro Cys

465

<210> 95

<211> 1461

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 95

```

gccacatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgcctactct      60
ggccaccacc atcacatca cggctccgag atcggaaccg gctttccttt cgaccctcac      120
tacgtggaag tgctgggcga gagaatgcac tatgtggacg tgggccccag agatggaacc      180
cctgtgctgt ttctgcacgg caaccctacc tccagctacg tgtggcggaa catcatccct      240
cacgtggccc ctacacacag atgtatcgcc cctgacctga tcggcatggg caagtctgac      300

aagcctgacc tgggctactt cttcgacgac cacgtgcggt tcatggacgc ctttatcgag      360
gctctgggcc tcgaagaggt ggtgctggtc atccatgatt ggggctctgc cctgggcttt      420
cactgggcca agagaaaccc cgagagagtg aagggaatcg cttcatgga attcatccgg      480
cctattccta cctgggacga gtggcctgag ttgccagag agacattcca ggccctcaga      540
accaccgacg tgggcagaaa gctgatcatc gaccagaacg tgttcatcga gggcaccctg      600
cctatgggag tcgtcagacc tctgaccgag gtggaaatgg accactacag agagcccttt      660
ctgaaccccg tggaccggga acctctttgg agattcccta acgagctgcc tatcgtggc      720

gagcctgcca atattgtggc cctggtggaa gagtacatgg actggctgca tcagagcccc      780
gtgcctaagc tgctgttttg gggaacaccc ggcgtgctga ttctctctgc tgaagctgct      840
agactggcca agagcctgcc taactgcaag gccgtggata tcggccctgg cctgaatctg      900
ctgcaagagg acaacccgca tctgatcgga tctgagatcg cccggtggct gaggaccctg      960
gaaatcagtg gactgcagga ctccgaagtg aatcaagagg ccaagcctga agtgaagccc     1020
gaagtcaagc ctgagacaca catcaacctg aagggtgtccg acggctccag cgagatcttc     1080
ttcaagatca agaaaaccac acctctgcgg cggtgatgg aagcctttgc caagagacag     1140

ggcaaagaga tggactccct gaccttctg tacgacggca tcgagatcca ggccgatcag      1200
accctgagg acctggacat ggaagataac gacatcattg aggccacag agagcagatc      1260
ggcggctctg gtggtagcgg aggttctggt ggatctggtg gttcttctgg cggcggatct      1320
tctggctgtg ctgagcagac aaccgagtgg accgcctgct ctaagtcttg tggcatgggc      1380
ttctccacca gagtgaccaa ccggaacaga cagtgcgaga tgctgaagca gaccggctg      1440
tgtatggtcc gaccttgcta a

```

<210> 96

<211> 1552

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 96

```

ggggacaagt ttgtacaaaa aagcaggcta taagcttgct gccacatga aatgggtcac      60
ctttatctcc ctgctgttcc tgtttctctc cgcctactct ggccaccacc atcaccatca      120
cggctccgag atcggaaccg gctttccttt cgaccctcac tacgtggaag tgctgggcga      180
gagaatgcac tatgtggacg tgggccccag agatggaacc cctgtgctgt tttctgcacgg      240
caacctacc tccagctacg tgtggcggaa catcatccct cacgtggccc ctacacacag      300
atgtatgcc cctgacctga tcggcatggg caagtctgac aagcctgacc tgggctactt      360
cttcgacgac cacgtgcggt tcatggacgc ctttatcgag gctctgggcc tcgaagaggt      420

```

```

ggtgctggtc atccatgatt ggggctctgc cctgggcttt cactgggcca agagaaaccc      480
cgagagagtg aagggaatcg ccttcatgga attcatccgg cctattccta cctgggacga      540
gtggcctgag ttcgccagag agacattcca ggccttcaga accaccgacg tgggcagaaa      600
gctgatcatc gaccagaacg tgttcatcga gggcaccttg cctatgggag tcgtcagacc      660
tctgaccgag gtggaaatgg accactacag agagcccttt ctgaaccccg tggaccggga      720
acctctttgg agattcccta acgagctgcc tatcgctggc gagcctgcca atattgtggc      780
cctggtggaa gatlacatgg actggctgca tcagagcccc gtgcctaagc tgctgttttg      840

```

```

gggaacaccc ggcgtgctga ttctctctgc tgaagctgct agactggcca agagcctgcc      900
taactgcaag gccgtggata tcggccctgg cctgaatctg ctgcaagagg acaaccccga      960
tctgatcgga tctgagatcg cccgtggct gagcaccttg gaaatcagtg gactgcagga      1020
ctccgaagtg aatcaagagg ccaagcctga agtgaagccc gaagtcaagc ctgagacaca      1080
catcaacctg aaggtgtccg acggctccag cgagatcttc ttcaagatca agaaaaccac      1140
acctctcgcg cggtgatgg aagcctttgc caagagacag ggcaaagaga tggactccct      1200
gaccttcttg tacgacggca tcgagateca ggccgatcag acccctgagg acctggacat      1260

```

```

ggaagataac gacatcattg agggccacag agagcagatc ggcggctctg gtggtagcgg      1320
aggtttctgt ggatctggtg gttcttcttg cgcgcatct tctggctgtg ctgagcagac      1380
aaccgagtgg accgcctgct ctaagtcttg tggcatgggc ttctccacca gactgaccaa      1440
ccggaacaga cagtgcgaga tgctgaagca gacccgcttg tgtatggtcc gaccttgcta      1500
aatctagagc ggcccggtta ccaaccagc tttcttgtac aaagtgttcc cc      1552

```

<210> 97

<211> 634

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 97

Cys Ala Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly

1 5 10 15
Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu

20 25 30
Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Glu Ala Ala Ala

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

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

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

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

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

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

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

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

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

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

195 200 205

Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala
 210 215 220
 Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala
 225 230 235 240
 Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly
 245 250 255
 Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe
 260 265 270
 Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr
 275 280 285
 Lys Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp
 290 295 300
 Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile
 305 310 315 320
 Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser
 325 330 335
 His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro
 340 345 350
 Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr
 355 360 365
 Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala
 370 375 380
 Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys
 385 390 395 400
 Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His
 405 410 415
 Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu
 420 425 430
 Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly
 435 440 445
 Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val

450 455 460
 Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly
 465 470 475 480
 Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro
 485 490 495
 Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu
 500 505 510
 His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu

 515 520 525
 Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu
 530 535 540
 Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala
 545 550 555 560
 Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr
 565 570 575
 Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln

 580 585 590
 Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys
 595 600 605
 Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu
 610 615 620
 Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
 625 630
 <210> 98
 <211> 652
 <212> PRT
 <213> Artificial Sequence
 <220><223> Fusion protein

 <400> 98
 Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
 1 5 10 15
 Tyr Ser Cys Ala Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr

20	25	30
Cys Gly Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys		
35	40	45
Arg Leu Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Glu Ala		
50	55	60
Ala Ala Lys Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp		
65	70	75
Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln		
85	90	95
Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu		
100	105	110
Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn		
115	120	125
Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val		
130	135	140
Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys		
145	150	155
Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn		
165	170	175
Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr		
180	185	190
Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu		
195	200	205
Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe		
210	215	220
Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp		
225	230	235
Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly		
245	250	255
Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys		
260	265	270

Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln
 275 280 285
 Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp
 290 295 300
 Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys
 305 310 315 320

 Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp
 325 330 335
 Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu
 340 345 350
 Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp
 355 360 365
 Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys
 370 375 380

 Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu
 385 390 395 400
 Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu
 405 410 415
 Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp
 420 425 430
 Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val
 435 440 445

 Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln
 450 455 460
 Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys
 465 470 475 480
 Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn
 485 490 495
 Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg
 500 505 510

 Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys

515 520 525
 Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys
 530 535 540
 Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val
 545 550 555 560
 Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe
 565 570 575

 His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys
 580 585 590
 Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys
 595 600 605
 Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys
 610 615 620
 Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys
 625 630 635 640

Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
 645 650

<210> 99

<211> 1965

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 99

gccaccatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgctactct	60
tgcgccgagt ggtctacagc ttggggccct tgttctacca cctgtggcct cggcatggcc	120
accagagtgt ccaaccagaa cagattctgc cggctggaaa ccagcggag actgtgcttg	180
tctagacctt gcgaggccgc tgccaaggac gtcataagt ctgaggtggc ccaccggttc	240
aaggacctgg gcgaagagaa ctccaaggcc ctggtgctga tcgccttcgc tcagtacttg	300
cagcagtgcc ccttcgagga ccacgtgaag ctggtcaacg aagtgaccga gttcgccaag	360
acctgcgtgg ccgatgagtc tgccgagaac tgcgacaagt ctctgcacac cctgttcggc	420
gacaagctgt gtaccgtggc taccctgaga gaaacctacg gcgagatggc cgactgctgc	480
gctaagcaag agcccagagag aaacgagtgc ttctgcagc acaaggacga caaccctaac	540

ctgcctagac tcgtgcggcc tgaggtggac gtgatgtgta ccgccttcca cgacaacgag 600
gaaaccttcc tgaagaagta cctgtacgag atcgccagac ggcaccccta cttttacgcc 660

cctgagctgc tgttcttcgc caagcggtag aaggccgcct tcaccgagtg ttgtcaggcc 720
gctgataagg ccgcttgccct gctgcctaaa ctggacgagc tgagagatga aggcaaggcc 780
tccagcgcca agcagagact gaagtgtgcc agcctgcaga agttcggcga gagagccttt 840
aaggcctggg ccgtcgctag actgtcccag agatttccca aggccgagtt tgccgaggtg 900
tccaagctgg ttaccgacct gaccaagggtg cacaccgaat gctgtcacgg cgacctgctg 960
gaatgcgccg atgatatagc cgatctggcc aagtacatct gcgagaacca ggactccatc 1020
tcctccaagc tgaagagtg ctgcgagaag cctctgctgg aaaagtccca ctgtatcgcc 1080

gaggtggaaa acgacgagat gcctgccgat ctgccttctc tggccgccga cttcgtggaa 1140
tctaaggacg tgtgcaagaa ctacgccgag gctaaggatg tgttcttggg catgtttctg 1200
tacgagtacg ctcgccggca ccccgactat tctgttgtgc tgctgctgag actggctaag 1260
acctacgaga caacctcga gaagtgtgtg gccgccgctg atcctcacga gtgttacgcc 1320
aaggtgttcg acgagttcaa gccactggtg gaagaacccc agaacctgat caagcagaat 1380
tgcgagctgt tcgagcagct gggcgagtag aagttccaga acgccctgct cgtgcggtac 1440
accaagaaaag tgccccaggt gtccacacct aactggttg aggtgtcccg gaacctgggc 1500

aaagtgggct ctaagtgtg caagcacccc gaggccaaga gaatgccttg tgccgaggac 1560
tacctgtccg tgggtgtgaa ccagctgtgc gtgctgcacg aaaagacccc tgtgtccgac 1620
agagtgacca agtgtgttac cgagagcctg gtcaacagac ggcttgcctt ctctgccctg 1680
gaagtggacg agacatacgt gccc aaagag ttcaacgccg agacattcac cttccacgcc 1740
gacatctgca cctgtccga gaaagagcgg cagatcaaga aacagaccgc tctggtggaa 1800
ctggtcaagc acaagcccaa ggccacaaaa gaacagctga aggccgtgat ggacgacttc 1860
gccgcctttg tggaaaagtg ttgcaaggcc gacgacaaag agacatgctt cgccgaagag 1920

ggcaagaaac tgggtggccg tttcaggct gctctgggac tttaa 1965

<210> 100
<211> 2038
<212> DNA
<213> Artificial Sequence
<220><223> Fusion protein
<400> 100

ggggacaagt ttgtacaaaa aagcaggcta tggtagcc accatgaaat gggtcacctt 60

tatctccctg ctgttctctgt tctctctccgc ctactcttgc gccgagtggg ctacagcttg	120
gggcccttgt tctaccacct gtggcctcgg catggccacc agagtgtcca accagaacag	180
attctgccgg ctggaaaccc agcggagact gtgcttgtct agaccttgcg aggccgctgc	240
caaggacgct cataagtctg aggtggccca ccggttcaag gacctgggcg aagagaactt	300
caaggccctg gtgctgatcg ccttcgctca gtacttgcag cagtgccct tgcaggacca	360
cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc tgcgtggcgg atgagtctgc	420
cgagaactgc gacaagtctc tgcacacct gttcggcgac aagctgtgta ccgtggctac	480
cctgagagaa acctacggcg agatggccga ctgctgcgct aagcaagagc ccgagagaaa	540
cgagtgcctc ctgcagcaca aggacgacaa ccctaacctg cctagactcg tgcggcctga	600
ggtggacgtg atgtgtaccg ccttcacga caacgaggaa accttctga agaagtacct	660
gtacgagatc gccagacggc acccctactt ttacgccct gagctgctgt tcttcgcaa	720
gcggtacaag gccgccttca ccgagtgttg tcaggccgct gataaggccg ctgacctgct	780
gcctaaactg gacgagctga gagatgaagg caaggcctcc agcgccaagc agagactgaa	840
gtgtgccagc ctgcagaagt tcggcgagag agcctttaag gcctgggccc tcgctagact	900
gtcccagaga ttcccgaagg ccgagtttgc cgaggtgtcc aagctggtta ccgacctgac	960
caaggtgcac accgaatgct gtcacggcga cctgctggaa tgcgccgatg atagagccga	1020
tctggccaag tacatctgcg agaaccagga ctccatctcc tccaagctga aagagtctg	1080
cgagaagcct ctgctggaaa agtcccactg tatcgccgag gtggaaaacg acgagatgcc	1140
tgccgatctg ctttctctgg ccgccgactt cgtggaatct aaggacgtgt gcaagaacta	1200
cgccgaggct aaggatgtgt tcttgggcat gtttctgtac gactacgctc ggccggcacc	1260
cgactattct gttgtgctgc tgctgagact ggctaagacc tacgagacaa cctcgagaa	1320
gtgctgtgcc gccgctgac ctcacgagtg ttacgccaag gtgttcgacg agttcaagcc	1380
actggtggaa gaacccaga acctgatcaa gcagaattgc gagctgttcg agcagctggg	1440
cgagtacaag ttccagaacg cctgctcgt gcggtacacc aagaaagtgc cccaggtgtc	1500
cacacctaca ctggttagg tgctccggaa cctgggcaaa gtgggctcta agtctgcaa	1560
gcaccccgag gccaagagaa tgccttgtgc cgaggactac ctgtccgtgg tgctgaacca	1620
gctgtgcgtg ctgcacgaaa agaccctgt gtccgacaga gtgaccaagt gctgtaccga	1680
gagcctggtc aacagacggc cttgtttctc tgccctggaa gtggacgaga catacgtgcc	1740
caaagagttc aacgccgaga cattcacctt ccacgccgac atctgcaccc tgtccgagaa	1800
agagcggcag atcaagaac agaccgtctt ggtggaactg gtcaagcaca agcccaaggc	1860
caccaagaa cagctgaagg ccgtgatgga cgacttcgcc gcctttgtgg aaaagtgtt	1920

caaggccgac gacaaagaga catgcttcgc cgaagagggc aagaaactgg tggccgcttc 1980

tcaggctgct ctgggacttt aatctagaaa ccagctttc ttgtacaaag tggcccc 2038

<210> 101

<211> 582

<212> PRT

<213> Homo sapiens

<400> 101

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

1 5 10 15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln

20 25 30

Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu

35 40 45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys

50 55 60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu

65 70 75 80

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

85 90 95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu

100 105 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His

115 120 125

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

130 135 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg

145 150 155 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala

165 170 175

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser

180 185 190

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
195 200 205

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
210 215 220

Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
225 230 235 240

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
245 250 255

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
260 265 270

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
275 280 285

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
290 295 300

Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
305 310 315 320

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
325 330 335

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
340 345 350

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
355 360 365

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
370 375 380

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
385 390 395 400

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
405 410 415

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
420 425 430

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys

435 440 445
 Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
 450 455 460
 Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
 465 470 475 480
 Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
 485 490 495

 Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
 500 505 510
 Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
 515 520 525
 Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
 530 535 540
 Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
 545 550 555 560

 Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
 565 570 575
 Ala Ala Ser Gln Ala Ala
 580
 <210> 102
 <211> 675
 <212> PRT
 <213> Artificial Sequence
 <220><223> Fusion protein
 <400> 102
 Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
 1 5 10 15
 Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
 20 25 30

 Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
 35 40 45
 Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys

50 55 60
 Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
 65 70 75 80
 Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
 85 90 95

 Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
 100 105 110
 Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
 115 120 125
 Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
 130 135 140
 Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
 145 150 155 160

 Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
 165 170 175
 Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
 180 185 190
 Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
 195 200 205
 Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
 210 215 220

 Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
 225 230 235 240
 Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
 245 250 255
 Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
 260 265 270
 Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
 275 280 285

 Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
 290 295 300

Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
 305 310 315 320
 Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
 325 330 335
 Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
 340 345 350

 Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
 355 360 365
 Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
 370 375 380
 Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
 385 390 395 400
 Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
 405 410 415

 Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
 420 425 430
 Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
 435 440 445
 Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
 450 455 460
 Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
 465 470 475 480

 Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
 485 490 495
 Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
 500 505 510
 Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
 515 520 525
 Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
 530 535 540

 Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys

545 550 555 560
 Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
 565 570 575
 Ala Ala Ser Gln Ala Ala Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
 580 585 590
 Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu
 595 600 605

Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Cys Ala
 610 615 620
 Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly Leu Gly
 625 630 635 640
 Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu Glu Thr
 645 650 655
 Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg Gly Arg
 660 665 670

Ser Pro Gln
 675

<210> 103

<211> 693

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 103

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
 1 5 10 15
 Tyr Ser Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu
 20 25 30
 Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr
 35 40 45

Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val
 50 55 60
 Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys

65				70				75				80			
Asp	Lys	Ser	Leu	His	Thr	Leu	Phe	Gly	Asp	Lys	Leu	Cys	Thr	Val	Ala
85				90				95							
Thr	Leu	Arg	Glu	Thr	Tyr	Gly	Glu	Met	Ala	Asp	Cys	Cys	Ala	Lys	Gln
100				105				110							
Glu	Pro	Glu	Arg	Asn	Glu	Cys	Phe	Leu	Gln	His	Lys	Asp	Asp	Asn	Pro
115				120				125							
Asn	Leu	Pro	Arg	Leu	Val	Arg	Pro	Glu	Val	Asp	Val	Met	Cys	Thr	Ala
130				135				140							
Phe	His	Asp	Asn	Glu	Glu	Thr	Phe	Leu	Lys	Lys	Tyr	Leu	Tyr	Glu	Ile
145				150				155				160			
Ala	Arg	Arg	His	Pro	Tyr	Phe	Tyr	Ala	Pro	Glu	Leu	Leu	Phe	Phe	Ala
165				170				175							
Lys	Arg	Tyr	Lys	Ala	Ala	Phe	Thr	Glu	Cys	Cys	Gln	Ala	Ala	Asp	Lys
180				185				190							
Ala	Ala	Cys	Leu	Leu	Pro	Lys	Leu	Asp	Glu	Leu	Arg	Asp	Glu	Gly	Lys
195				200				205							
Ala	Ser	Ser	Ala	Lys	Gln	Arg	Leu	Lys	Cys	Ala	Ser	Leu	Gln	Lys	Phe
210				215				220							
Gly	Glu	Arg	Ala	Phe	Lys	Ala	Trp	Ala	Val	Ala	Arg	Leu	Ser	Gln	Arg
225				230				235				240			
Phe	Pro	Lys	Ala	Glu	Phe	Ala	Glu	Val	Ser	Lys	Leu	Val	Thr	Asp	Leu
245				250				255							
Thr	Lys	Val	His	Thr	Glu	Cys	Cys	His	Gly	Asp	Leu	Leu	Glu	Cys	Ala
260				265				270							
Asp	Asp	Arg	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	Cys	Glu	Asn	Gln	Asp	Ser
275				280				285							
Ile	Ser	Ser	Lys	Leu	Lys	Glu	Cys	Cys	Glu	Lys	Pro	Leu	Leu	Glu	Lys
290				295				300							
Ser	His	Cys	Ile	Ala	Glu	Val	Glu	Asn	Asp	Glu	Met	Pro	Ala	Asp	Leu
305				310				315				320			

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

Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr
340 345 350

Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala
355 360 365

Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro
370 375 380

His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu
385 390 395 400

Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu
405 410 415

Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys
420 425 430

Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu
435 440 445

Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met
450 455 460

Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val
465 470 475 480

Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr
485 490 495

Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp
500 505 510

Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His
515 520 525

Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln
530 535 540

Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu
545 550 555 560

Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys

565 570 575
 Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys
 580 585 590
 Leu Val Ala Ala Ser Gln Ala Ala Glu Ala Ala Ala Lys Glu Ala Ala
 595 600 605
 Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala
 610 615 620

Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
 625 630 635 640
 Cys Ala Glu Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys Gly
 645 650 655
 Leu Gly Met Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Arg Leu
 660 665 670
 Glu Thr Gln Arg Arg Leu Cys Leu Ser Arg Pro Cys Pro Pro Ser Arg
 675 680 685

Gly Arg Ser Pro Gln
 690

<210> 104

<211> 2088

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 104

gccaccatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgctactct	60
gacgccca agtctgaggt ggcccacaga ttcaaggacc tgggcgaaga gaacttcaag	120
gccctggtgc tgatcgctt cgctcagtac ttgcagcagt gcccttcga ggaccacgtg	180
aagctggtca acgaagtac cgagttgcc aagacctgcg tggccgatga gtctgccgag	240
aactgcgaca agtctctgca cacctgttc ggcgacaagc tgtgtaccgt ggctaccctg	300
agagaaacct acggcgagat ggccgactgc tgcgctaagc aagagcccga gagaaacgag	360
tgcttctgc agcacaagga cgacaacct aacctgccta gactcgtgcg gcctgaggtg	420
gacgtgatgt gtaccgctt ccacgacaac gaggaacct tcctgaagaa gtacctgtac	480
gagatcgcca gacggcaccc ctacttttac gccctgagc tgctgttctt cgccaagcgg	540

tacaaggccg ccttcaccga gtgttgtag gcccgtgata aggccgcttg cctgctgcct 600

aaactggacg agctgagaga tgaaggcaag gcctccagcg ccaagcagag actgaagtgt 660

gccagcctgc agaagttcgg cgagagagcc ttttaaggcct gggccgctgc tagactgtcc 720

cagagatttc ccaaggccga gtttgccgag gtgtccaagc tggttaccga cctgaccaag 780

gtgcacaccg aatgctgtca cggcgacctg ctggaatgcg ccgatgatag agccgatctg 840

gccaaagtaca tctgcgagaa ccaggactcc atctcctcca agctgaaaga gtgctgcgag 900

aagcctctgc tggaaaagtc ccactgtatc gccgaggtgg aaaacgacga gatgcctgcc 960

gatctgcctt ctctggccgc cgacttcgtg gaatctaagg acgtgtgcaa gaactacgcc 1020

gaggccaagg atgtgttctt gggcatgttt ctgtacgagt acgtcggcg gcaccccgac 1080

tattctgttg tgctgtgct gagactggct aagacctacg agacaaccct cgagaagtgc 1140

tggtccgccc ctgactctca cgagtgttac gccaaagtgt tcgacgagtt caagccactg 1200

gtggaagaac ccagaacct gatcaagcag aattgcgagc tgttcgagca gctgggagag 1260

tacaagttcc agaacgccct gctcgtgcgg tacaccaaga aagtgccca ggtgtccaca 1320

cctacactgg ttgaggtgtc ccggaacctg ggcaaagtgg gctctaagtg ctgcaagcac 1380

cctgaggcca agagaatgcc ttgcgccgag gactacctgt ccgtggtgct gaatcagctg 1440

tgctgtctgc acgaaaagac cctgtgttcc gacagagtga ccaagtgtg taccgagagc 1500

ctggtcaaca gacggccttg cttctctgcc ctggaagtgg acgagacata cgtgcccaaa 1560

gagttcaacg ccgagacatt caccttcac gccgacatct gcaccctgtc cgagaaagag 1620

cggcagatca agaaacagac cgctctgggtg gaactggtca agcacaagcc caaggccacc 1680

aaagaacagc tgaaggccgt gatggacgac ttgcgccctt ttgtggaaaa gtgttgcaag 1740

gccgacgaca aagagacatg cttcgcgcaa gagggcaaga aactggtggc cgcttctcag 1800

gctgtctgagg ccgtgtctaa agaggctgcc gctaaagaag ccgcagccaa agaggcagct 1860

gcaaaagaag ctgtgtcaaa agaggcagcc gccaaagagg ccgtgtctaa agaagcagcc 1920

gccaaagtgtg ctgagtggtc tacagcttgg ggcccctgct ctacaacctg tggactcggc 1980

atggccacca gagtgtctaa ccagaacaga ttctgccggc tggaaacca gcggagactg 2040

tgctgtctga gaccctgtcc tcttagcaga ggcagatccc ctgagtga 2088

<210> 105

<211> 2161

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 105

ggggacaagt ttgtacaaaa aagcaggcta tggtagcc accatgaaat gggtcacctt	60
tatctccctg ctgttctgt tctctccgc ctactctgac gcccacaagt ctgaggtggc	120
ccacagattc aaggacctgg gcgaagagaa cttcaaggcc ctggtgctga tcgccttcgc	180
tcagtacttg cagcagtgcc ctttcgagga ccacgtgaag ctggtcaacg aagtaccga	240
gttcgccaag acctgcgtgg ccgatgagtc tgccgagaac tgcgacaagt ctctgcacac	300
cctgttcggc gacaagctgt gtaccgtggc taccctgaga gaaacctacg gcgagatggc	360
cgactctgc gctaagcaag agccccgagag aaacgagtgc ttcttcgacg acaaggacga	420
caaccctaac ctgcctagac tcgtgcggcc tgaggtggac gtgatgtgta ccgccttcca	480
cgacaacgag gaaaccttcc tgaagaagta cctgtacgag atcgccagac ggcacccta	540
cttttacgcc cctgagctgc tgttcttcgc caagcggtac aaggccgcct tcaccgagt	600
ttgtcaggcc gctgataagg ccgcttcct gctgcctaaa ctggacgagc tgagagatga	660
aggcaaggcc tccagcgcca agcagagact gaagtgtgcc agcctgcaga agttcggcga	720
gagagccttt aaggcctggg ccgtcgctag actgtcccag agatttccca aggccgagtt	780
tgccgaggtg tccaagctgg ttaccgacct gaccaaggtg cacaccgaat gctgtcacgg	840
cgacctgctg gaatgcgccg atgatagagc cgatctggcc aagtacatct gcgagaacca	900
ggactccatc tctccaagc tgaaagagtg ctgcgagaag cctctgctgg aaaagtccca	960
ctgtatcgcc gaggtggaaa acgacgagat gcctgccgat ctgccttctc tggccgccga	1020
cttcgtggaa tctaaggacg tgtgaagaa ctacgccgag gccaaagatg tgttcttggg	1080
catgtttctg tacgagtagc ctgcgaggca ccccgactat tctgttgtgc tgctgctgag	1140
actggctaag acctacgaga caacctcga gaagtgtgt gccgccgtg atcctcacga	1200
gtgttacgcc aagggttctg acgagttcaa gccactgggtg gaagaacccc agaacctgat	1260
caagcagaat tgcgagctgt tcgagcagct gggcgagtag aagttccaga acgcctgct	1320
cgtgcggtac accaagaaag tgccccaggt gtccacacct aacttggttg aggtgtccg	1380
gaacctgggc aaagtgggct ctaagtgtg caagcacct gaggccaaga gaatgccttg	1440
cgccgaggac tacctgtccg tgggtgtgaa tcagctgtgc gtgctgcacg aaaagacccc	1500
tggtgccgac agagtaccca agtgcgttac cgagagcctg gtcaacagac ggccttgctt	1560
ctctgccctg gaagtggacg agacatactg gcccaaagag ttcaacgccg agacattcac	1620
cttcacgcc gacatctgca cctgttccga gaaagagcgg cagatcaaga aacagaccgc	1680
tctggtggaa ctggtcaagc acaagcccaa ggccaccaa gaacagctga aggccgtgat	1740
ggacgacttc gccgcctttg tggaaaagtg ttgcaaggcc gacgacaaag agacatgctt	1800

cgccgaagag ggcaagaaac tggtagccgc ttctcaggct gctgaggccg ctgctaaaga 1860
 ggctgccgct aaagaagccg cagccaaaga ggcagctgca aaagaagctg ctgcaaaaga 1920
 ggcagccgcc aaagaggccg ctgctaaaga agcagccgcc aagtgtgctg agtggcttac 1980
 agcttggggc ccctgctcta caacctgtgg actcggcatg gccaccagag tgtctaacca 2040
 gaacagattc tgccggctgg aaaccagcg gagactgtgc ctgtctagac cctgtcctcc 2100
 tagcagagge agatccctc agtgatctag aaaccagct ttctgtaca aagtgtccc 2160
 c 2161

<210> 106

<211> 634

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 106

Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly

1 5 10 15
 Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met
 20 25 30
 Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala Ala Ala
 35 40 45
 Lys Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly
 50 55 60
 Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu

65 70 75 80
 Gln Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr
 85 90 95
 Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp
 100 105 110
 Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr
 115 120 125
 Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu

130

135

140

Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn
145 150 155 160
Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe
 165 170 175
His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala
 180 185 190
Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys

 195 200 205
Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala
 210 215 220
Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala
225 230 235 240
Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly
 245 250 255
Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe

 260 265 270
Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr
 275 280 285
Lys Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp
 290 295 300
Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile
305 310 315 320
Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser

 325 330 335
His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro
 340 345 350
Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr
 355 360 365
Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala
 370 375 380
Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys

385 390 395 400
 Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His
 405 410 415
 Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu
 420 425 430
 Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly
 435 440 445
 Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val

 450 455 460
 Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly
 465 470 475 480
 Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro
 485 490 495
 Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu
 500 505 510
 His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu

 515 520 525
 Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu
 530 535 540
 Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala
 545 550 555 560
 Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr
 565 570 575
 Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln

 580 585 590
 Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys
 595 600 605
 Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu
 610 615 620
 Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
 625 630
 <210> 107

<211> 652

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 107

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala

1 5 10 15

Tyr Ser Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser

20 25 30

Cys Gly Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys

35 40 45

Glu Met Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala

50 55 60

Ala Ala Lys Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp

65 70 75 80

Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln

85 90 95

Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu

100 105 110

Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn

115 120 125

Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val

130 135 140

Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys

145 150 155 160

Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn

165 170 175

Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr

180 185 190

Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu

195 200 205

Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe
210 215 220

Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp
225 230 235 240

Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly
245 250 255

Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys
260 265 270

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

Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp
290 295 300

Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys
305 310 315 320

Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp
325 330 335

Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu
340 345 350

Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp
355 360 365

Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys
370 375 380

Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu
385 390 395 400

Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu
405 410 415

Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp
420 425 430

Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val
435 440 445

Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln

450 455 460
 Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys
 465 470 475 480
 Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn
 485 490 495
 Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg
 500 505 510

 Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys
 515 520 525
 Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys
 530 535 540
 Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val
 545 550 555 560
 Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe
 565 570 575

 His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys
 580 585 590
 Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys
 595 600 605
 Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys
 610 615 620
 Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys
 625 630 635 640

 Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
 645 650

<210> 108

<211> 1965

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 108

gccacatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgctactct

60

tgtgccgagc agaccacaga gtggaccgcc tgctctaagt cttgcggcat gggcttctcc	120
accagagtga ccaaccggaa cagacagtgc gagatgctga agcagaccgc gctgtgtatg	180
gttcgacctt gcgaggccgc tgccaaggat gctcataagt ctgaggtggc ccaccggttc	240
aaggacctgg gcgaagagaa cttcaaggcc ctggtgctga tcgccttcgc tcagtacttg	300
cagcagtgcc ctttcgagga ccacgtgaag ctggtcaacg aagtaccga gttcgccaag	360
acctgcgtgg ccgatgagtc tgccgagaac tgcgacaagt ctctgcacac cctgttcggc	420
gacaagctgt gtaccgtggc taccctgaga gaaacctacg gcgagatggc cgactgctgc	480
gctaagcaag agcccagag aaacgagtgc ttctgcagc acaaggacga caaccctaac	540
ctgcctagac tcgtgcggcc tgaggtggac gtgatgtgta ccgccttcca cgacaacgag	600
gaaaccttcc tgaagaagta cctgtacgag atcgccagac ggcacccta cttttacgcc	660
cctgagctgc tgttcttcgc caagcggtag aaggccgcct tcaccgagtg ttgtcaggcc	720
gctgataagg ccgcttgctt gctgcctaaa ctggacgagc tgagagatga aggcaaggcc	780
tccagcgcca agcagagact gaagtgtgcc agcctgcaga agttcggcga gagagccttt	840
aaggcctggg ccgtcgttag actgtcccag agatttccca aggcgcgagt ttgccgagtg	900
tccaagctgg ttaccgacct gaccaaggct cacaccgaat gctgtcacgg cgacctgctg	960
gaatgcgccg atgatagagc cgatctggcc aagtacatct gcgagaacca ggactccatc	1020
tcctccaagc tgaagagtg ctgcgagaag cctctgctgg aaaagtcca ctgtatgcc	1080
gaggtggaaa acgacgagat gcctgccgat ctgccttctc tggccgccga cttcgtggaa	1140
tctaaggacg tgtgcaagaa ctacccgag gctaaggatg tgttcctggg catgtttctg	1200
tacgagtacg ctcgccggca ccccgattat agtgtggtgc tgctgctgag actggctaag	1260
acctacgaga caaccctga gaagtgtgt gccgccgctg atcctcacga gtgttacgcc	1320
aaggtgttcg acgagttcaa gccactggtg gaagaacccc agaacctgat caagcagaat	1380
tgcgagctgt tcgagcagct gggcgagtac aagttccaga acgccctgct cgtgcggtac	1440
accaagaaag tgccccaggt gtccacacct aactggttg aggtgtccc gaacctgggc	1500
aaagtgggct ctaagtgtg caagaccccc gagccaaga gaatgccttg cgccgaggat	1560
tacctgtccg tgggtgtgaa ccagctgtgc gtgctgcacg aaaagacccc tgtgtccgac	1620
cgctgacca agtgtgttac agagtccctg gtcaacagac ggccctgctt ctctgccctg	1680
gaagtggacg agacatacgt gcccgaagag ttcaacgccg agacattcac cttccacgcc	1740
gacatctgca cctgtccga gaaagagcgg cagatcaaga aacagaccgc tctggtcgaa	1800
ctggtcaagc acaagcccaa ggccacccaa gaacagctga aggcctgat ggacgacttc	1860

gccgcctttg tggaaaagtg ttgcaaggcc gacgacaaag agacatgctt cgccgaagag 1920

ggcaagaaac tggtagccgc ttctcaggct gctctgggac tttaa 1965

<210> 109

<211> 2038

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 109

ggggacaagt ttgtacaaaa aagcaggcta tggtagccgc accatgaaat gggtcacctt 60

tatctccctg ctgttcctgt tctcctccgc ctactcttgt gccgagcaga ccacagagtg 120

gaccgcctgc tctaagtctt gcggcatggg ctctccacc agagtgacca accggaacag 180

acagtgcgag atgtgaagc agaccggct gtgtatggtt cgaccttgcg aggccgctgc 240

caaggatgct cataagtctg aggtggccca ccggttcaag gacctgggcg aagagaactt 300

caaggccctg gtgtgatcg ctttcgctca gtacttgtag cagtgccctt tcgaggacca 360

cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc tgcgtggccg atgagtctgc 420

cgagaactgc gacaagtctc tgcacacctt gttcggcgac aagctgtgta ccgtggctac 480

cctgagagaa acctacggcg agatggccga ctgctgcgct aagcaagagc ccgagagaaa 540

cgagtgttct ctgcagcaca aggacgacaa ccctaacctg cctagactcg tgcggcctga 600

ggtggacgtg atgtgtaccg ctttcacga caacgaggaa accttctga agaagtacct 660

gtacgagatc gccagacggc accctactt ttacgccctt gagctgctgt tcttcgcaa 720

gcggtacaag gccgccttca ccgagtgttg tcaggccgct gataaggccg cttgcctgct 780

gcctaaactg gacgagctga gagatgaagg caaggcctcc agcgccaagc agagactgaa 840

gtgtgccagc ctgcagaagt tcggcgagag agcctttaag gcctgggccc tcgctagact 900

gtcccagaga tttcccaagg ccgagtttgc cgaggtgtcc aagctggtta ccgacctgac 960

caaggtgcac accgaatgct gtcacggcga cctgctggaa tgcgccgatg atagaccga 1020

tctggccaag tacatctgcg agaaccagga ctccatctcc tccaagtga aagagtgtg 1080

cgagaagcct ctgctgaaa agtccactg tatcgccgag gtgaaaaacg acgagatgcc 1140

tgccgatctg ctttctctgg ccgccgactt cgtggaatct aaggacgtgt gcaagaacta 1200

cgccgaggct aaggatgtgt tcttgggcat gtttctgtac gactacgctc ggcggcaccc 1260

cgattatagt gtggtgtgc tgctgagact ggctaagacc tacgagacaa cctcgagaa 1320

gtgctgtgcc gccgctgac ctcacgagtg ttacgccaag gtgttcgacg agttcaagcc 1380

actgggtggaa gaaccccaga acctgatcaa gcagaattgc gagctgttcg agcagctggg 1440
cgagtacaag ttccagaacg ccctgctcgt gcggtacacc aagaaagtgc cccaggtgtc 1500
cacacctaca ctggttgagg tgtcccggaa cctgggcaaa gtgggctcta agtgctgcaa 1560

gcaccccgag gccaagagaa tgccttgcgc cgaggattac ctgtccgtgg tgctgaacca 1620
gctgtgcgtg ctgcacgaaa agaccctgt gtccgaccgc gtgaccaagt gctgtacaga 1680
gtccctggtc aacagacggc cctgcttctc tgccctggaa gtggacgaga catacgtgcc 1740
caaagagttc aacgccgaga cattcacctt ccacgccgac atctgcaccc tgtccgagaa 1800
agagcggcag atcaagaaac agaccgtctt ggtcgaactg gtcaagcaca agcccaaggc 1860
caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc gcctttgtgg aaaagtgttg 1920
caagcccgac gacaaagaga catgcttcgc cgaagagggc aagaaactgg tggccgcttc 1980

tcaggetgct ctgggacttt aatctagaaa cccagctttc ttgtacaaag tggteccc 2038

<210> 110

<211> 669

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 110

Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser Cys Gly

1 5 10 15

Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys Glu Met

20 25 30

Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala Ala Ala

35 40 45

Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

50 55 60

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu

65 70 75 80

Ala Ala Ala Lys Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys

85 90 95

Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala

100 105 110

Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn
115 120 125
Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu
130 135 140
Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr
145 150 155 160
Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala
165 170 175
Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp
180 185 190
Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys
195 200 205
Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr
210 215 220
Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe
225 230 235 240
Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala
245 250 255
Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu
260 265 270
Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln
275 280 285
Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser
290 295 300
Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr
305 310 315 320
Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu
325 330 335
Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln
340 345 350
Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu

355 360 365
 Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala
 370 375 380
 Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys
 385 390 395 400
 Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr
 405 410 415
 Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg

 420 425 430
 Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala
 435 440 445
 Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu
 450 455 460
 Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu
 465 470 475 480
 Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr

 485 490 495
 Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg
 500 505 510
 Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys
 515 520 525
 Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu
 530 535 540
 Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys

 545 550 555 560
 Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu
 565 570 575
 Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr
 580 585 590
 Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys
 595 600 605

Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr

610

615

620

Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu

625

630

635

640

Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly

645

650

655

Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu

660

665

<210> 111

<211> 687

<212> PRT

<213> Artificial Sequence

<220

><223> Fusion protein

<400> 111

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala

1

5

10

15

Tyr Ser Cys Ala Glu Gln Thr Thr Glu Trp Thr Ala Cys Ser Lys Ser

20

25

30

Cys Gly Met Gly Phe Ser Thr Arg Val Thr Asn Arg Asn Arg Gln Cys

35

40

45

Glu Met Leu Lys Gln Thr Arg Leu Cys Met Val Arg Pro Cys Glu Ala

50

55

60

Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Lys Glu Ala Ala

65

70

75

80

Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala

85

90

95

Lys Glu Ala Ala Ala Lys Asp Ala His Lys Ser Glu Val Ala His Arg

100

105

110

Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala

115

120

125

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

130	135	140	
Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser			
145	150	155	160
Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu			
	165	170	175
Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys			
	180	185	190
Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys			
	195	200	205
Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val			
	210	215	220
Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr			
225	230	235	240
Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu			
	245	250	255
Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln			
	260	265	270
Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg			
	275	280	285
Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser			
	290	295	300
Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg			
305	310	315	320
Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu			
	325	330	335
Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu			
	340	345	350
Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu			
	355	360	365
Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro			
	370	375	380

Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met
 385 390 395 400
 Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp
 405 410 415
 Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe
 420 425 430
 Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu
 435 440 445

 Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala
 450 455 460
 Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys
 465 470 475 480
 Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu
 485 490 495
 Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg
 500 505 510

 Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val
 515 520 525
 Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu
 530 535 540
 Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn
 545 550 555 560
 Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr
 565 570 575

 Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala
 580 585 590
 Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr
 595 600 605
 Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln
 610 615 620
 Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys

625	630	635	640
Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe			
645	650	655	
Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu			
660	665	670	
Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu			
675	680	685	
<210> 112			
<211> 2070			
<212> DNA			
<213> Artificial Sequence			
<220><223> Fusion protein			
<400> 112			
gccacatga aatgggtcac ctttatctcc ctgctgttcc tgttctctc cgctactct			60
tgtgccgagc agaccacaga gtggaccgcc tgctctaagt cttgcggcat gggtttctcc			120
accagagtga ccaaccggaa cagacagtgc gagatgctga agcagaccgc gctgtgtatg			180
gttcgacctt gcgaggccgc tgccaaagag gctgctgcta aagaagccgc cgcaaaagag			240
gcagcagcaa aagaggctgc cgccaaagag gccgcagcca aagaagcagc agctaaagag			300
gccgctgcta aggacgcccc caagtctgaa gtggcccacc ggtttaagga cctgggcgaa			360
gagaacttca aggccttggg gctgatcgcc ttcgctcagt acttgagca gtgccccttc			420
gaggaccacg tgaagctggg caacgaagtg accgagttcg ccaagacctg cgtggccgat			480
gagtctgccg agaactgcga caagtctctg cacacctgt tcggcgacaa gctgtgtacc			540
gtggctaccg tgagagaaac ctacggcgag atggccgact gctgcgctaa gcaagagccc			600
gagagaaacg agtgcttcct gcagcacaag gacgacaacc ctaacctgcc tagactcgtg			660
cggcctgagg tggacgtgat gtgtaccgcc ttccacgaca acgaggaaac cttcctgaag			720
aagtacctgt acgagatcgc cagacggcac ccctactttt acgcccctga gctgctgttt			780
ttcgccaagc ggtacaagc gcgcttcacc gagtgtgtgc aggccgccga taaggccgct			840
tgtctgctgc ctaaactgga cgagctgcgc gacgaaggca aggcctcttc tgctaagcag			900
cggctgaagt gcgccagcct gcagaagttt ggcgagagag ctttcaagc ttgggccgctc			960
gctagactgt cccagagatt tcccaaggcc gattttgccg aggtgtccaa gctgggttacc			1020
gacctgacca aggtgcacac cgaatgctgt cacggcgacc tgctggaatg cgccgatgat			1080

agagccgatac tggccaagta catctgcgag aaccaggact ccatctctc caagctgaaa 1140
gagtgtctgc agaagcctct gctggaaaag tcccactgta tcgccgaggt ggaaaacgac 1200
gagatgcctg ccgactctgc ttctctggcc gccgacttcg tggaaatctaa ggacgtgtgc 1260
aagaactacg ccgaggccaa ggatgtgttc ctgggcatgt ttctgtacga gtacgtcgg 1320

cggcaccg attatagtgt ggtgtctgtg ctgagactgg ctaagaccta cgagacaacc 1380
ctcgagaagt gctgtgccgc cgctgatect cagcagtggt acgccaaggt gttcgacgag 1440
ttcaagccac tgggtgaaga accccagaac ctgatcaagc agaattgcga gctgttcgag 1500
cagctgggag agtacaagtt ccagaacgcc ctgctcgtgc ggtacaccaa gaaagtgtcc 1560
caggtgtcca cacctacact ggttgagggtg tcccgaacc tgggcaaagt gggctctaag 1620
tgctgaagc accctgagc caagagaatg ccttgccgcg aggactacct gtccgtggtg 1680
ctgaatcagc tgtgcgtgct gcacgaaaag acccctgtgt ccgaccgcgt gaccaagtgc 1740

tgtacagagt ccctgggtcaa cagacggccc tgcttctctg ccctggaagt ggacgagaca 1800
tacgtgccc aagagttcaa cgccgagaca ttacattcc acgccgacat ctgcacctg 1860
tccgagaaag agcggcagat caagaaacag accgctctgg tcgagctggt taagcacaag 1920
cccaaggcca ccaagaaca gctgaaggcc gtgatggacg acttcgccgc ctttgtggaa 1980
aagtgttga agcccgacga caaagagaca tgcttcgccg aagagggcaa gaaactggtg 2040
gccgcttctc aggtgtctct gggactttaa 2070

<210> 113

<211> 2143

<212> DNA

<213> Artificial Sequence

<220><223> Fusion protein

<400> 113

ggggacaagt ttgtacaaa aagcaggcta tggtagcgcc accatgaaat gggtcacctt 60
tatctccctg ctgttctgt tctctccgc ctactctgt gccgagcaga ccacagagtg 120
gaccgcctgc tctaagtctt gcggcatggg ctctccacc agagtaccca accggaacag 180
acagtgcgag atgtgaagc agaccggct gtgtatggtt cgaccttgcg agccgctgc 240
caaagaggct gctgctaaag aagccgccgc aaaagaggca gcagcaaaag aggtgccgc 300
caaagaggcc gcagccaaag aagcagcagc taaagaggcc gctgctaagg acgccacaa 360
gtctgaagtg gccaccggt ttaaggacct gggcgaagag aacttcaagg ccctggtgct 420

gatgccttc gctcagtact tgcagcagtg ccccttcgag gaccacgtga agctggtcaa 480

cgaagtgacc gaggttcgcca agacctgcgt ggccgatgag tctgccgaga actgcgacaa 540
gtctctgcac accctgttcg gcgacaagct gtgtaccgtg gctaccctga gagaaaccta 600
cggcgagatg gccgactgct gcgctaagca agagcccagag agaaacgagt gcttctctgca 660
gcacaaggac gacaacccta acctgcctag actcgtgcgg cctgaggtgg acgtgatgtg 720
taccgccttc cagcacaacg aggaaacctt cctgaagaag tacctgtacg agatgccag 780
acggcacccc tacttttacg cccctgagct gctgtttttc gccaaagcgt acaaggccgc 840

cttcaccgag tgttgtcagg ccgccgataa ggccgcttgt ctgctgccta aactggacga 900
gctgcgcgac gaaggcaagg cctcttctgc taagcagcgg ctgaagtgcg ccagcctgca 960
gaagtttggc gagagagcct tcaaggcttg ggccgtcgct agactgtccc agagatttcc 1020
caaggccgag ttgtccgagg tgtccaagct ggttaccgac ctgaccaagg tgcacaccga 1080
atgctgtcac ggcgacctgc tggaaatgcg cgatgataga gccgatctgg ccaagtacat 1140
ctgcgagAAC caggactcca tctctccaa gctgaaagag tgctgcgaga agcctctgct 1200
ggaaaagtcc cactgtatcg ccgaggtgga aaacgacgag atgcctgccg atctgccttc 1260

tctggccgcc gacttcgtgg aatctaagga cgtgtgcaag aactacgccg aggccaagga 1320
tgtgttctg ggcatgttcc tgtacagta cgctcggcgg caccgccatt atagtgtggt 1380
gctgctgctg agactggcta agacctacga gacaaccctc gagaagtgtc gtgccgccgc 1440
tgatctctac gagtgttacg ccaaggtgtt cgacgagttc aagccactgg tggagaacc 1500
ccagaacctg atcaagcaga attgcgagct gttcgagcag ctgggcgagt acaagttcca 1560
gaacgcctg ctctgtcggg acaccaagaa agtgccccag gtgtccacac ctacactggt 1620
tgagggtgcc cggaacctgg gcaaagtggg ctctaagtgc tgcaagcacc ctgaggccaa 1680

gagaatgcct tgcgccgagg actacctgtc cgtggtgctg aatcagctgt gcgtgctgca 1740
cgaaaagacc cctgtgtccg accgcgtgac caagtgtgt acagagtccc tggtaacag 1800
acggccctgc ttctctgccc tggaaagtga cgagacatac gtgcccagg agttcaacgc 1860
cgagacattc accttcacg ccgacatctg caccctgtcc gagaagagc ggagatcaa 1920
gaaacagacc gctctggtcg agctggttaa gcacaagccc aaggccacca aagaacagct 1980
gaaggccgtg atggacgact tcgccgcctt tgtggaaaag tgttgcaagg ccgacgaaa 2040
agagacatgc ttcgccgaag agggcaagaa actggtggcc gcttctcagg ctgctctggg 2100

actttaatct agaaaccag ctttcttgta caaagtggc ccc 2143

<210> 114

<211> 484

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein

<400> 114

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala

1 5 10 15

Tyr Ser Gly His His His His His His Gly Ser Glu Ile Gly Thr Gly

20 25 30

Phe Pro Phe Asp Pro His Tyr Val Glu Val Leu Gly Glu Arg Met His

35 40 45

Tyr Val Asp Val Gly Pro Arg Asp Gly Thr Pro Val Leu Phe Leu His

50 55 60

Gly Asn Pro Thr Ser Ser Tyr Val Trp Arg Asn Ile Ile Pro His Val

65 70 75 80

Ala Pro Thr His Arg Cys Ile Ala Pro Asp Leu Ile Gly Met Gly Lys

85 90 95

Ser Asp Lys Pro Asp Leu Gly Tyr Phe Phe Asp Asp His Val Arg Phe

100 105 110

Met Asp Ala Phe Ile Glu Ala Leu Gly Leu Glu Glu Val Val Leu Val

115 120 125

Ile His Asp Trp Gly Ser Ala Leu Gly Phe His Trp Ala Lys Arg Asn

130 135 140

Pro Glu Arg Val Lys Gly Ile Ala Phe Met Glu Phe Ile Arg Pro Ile

145 150 155 160

Pro Thr Trp Asp Glu Trp Pro Glu Phe Ala Arg Glu Thr Phe Gln Ala

165 170 175

Phe Arg Thr Thr Asp Val Gly Arg Lys Leu Ile Ile Asp Gln Asn Val

180 185 190

Phe Ile Glu Gly Thr Leu Pro Met Gly Val Val Arg Pro Leu Thr Glu

195 200 205

Val Glu Met Asp His Tyr Arg Glu Pro Phe Leu Asn Pro Val Asp Arg

210 215 220

Glu Pro Leu Trp Arg Phe Pro Asn Glu Leu Pro Ile Ala Gly Glu Pro

225 230 235 240

Ala Asn Ile Val Ala Leu Val Glu Glu Tyr Met Asp Trp Leu His Gln

245 250 255

Ser Pro Val Pro Lys Leu Leu Phe Trp Gly Thr Pro Gly Val Leu Ile

260 265 270

Pro Pro Ala Glu Ala Ala Arg Leu Ala Lys Ser Leu Pro Asn Cys Lys

275 280 285

Ala Val Asp Ile Gly Pro Gly Leu Asn Leu Leu Gln Glu Asp Asn Pro

290 295 300

Asp Leu Ile Gly Ser Glu Ile Ala Arg Trp Leu Ser Thr Leu Glu Ile

305 310 315 320

Ser Gly Leu Gln Asp Ser Glu Val Asn Gln Glu Ala Lys Pro Glu Val

325 330 335

Lys Pro Glu Val Lys Pro Glu Thr His Ile Asn Leu Lys Val Ser Asp

340 345 350

Gly Ser Ser Glu Ile Phe Phe Lys Ile Lys Lys Thr Thr Pro Leu Arg

355 360 365

Arg Leu Met Glu Ala Phe Ala Lys Arg Gln Gly Lys Glu Met Asp Ser

370 375 380

Leu Thr Phe Leu Tyr Asp Gly Ile Glu Ile Gln Ala Asp Gln Thr Pro

385 390 395 400

Glu Asp Leu Asp Met Glu Asp Asn Asp Ile Ile Glu Ala His Arg Glu

405 410 415

Gln Ile Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly

420 425 430

Ser Ser Gly Gly Gly Ser Ser Gly Cys Ala Glu Gln Thr Thr Glu Trp

435 440 445

Thr Ala Cys Ser Lys Ser Cys Gly Met Gly Phe Ser Thr Arg Val Thr

450 455 460

Asn Arg Asn Arg Gln Cys Glu Met Leu Lys Gln Thr Arg Leu Cys Met

465 470 475 480

Val Arg Pro Cys

<210> 115

<211> 13

<212> PRT

<213> Artificial Sequence

<220><223> part of recombinant protein

<400> 115

Pro Pro Ser Arg Gly Arg Ser Pro Gln Asn Ser Ala Phe

1 5 10

<210> 116

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> part of recombinant protein

<400> 116

Gly Gln Pro Val Tyr Ser Ser Leu

1 5

<210> 117

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> part of recombinant protein

<400> 117

Glu Ala Asp Leu Glu Glu Asn

1 5

<210> 118

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> part of recombinant protein

<400> 118

Glu Gln Gln Pro Glu Gln Pro Thr Asp

1 5

<210> 119

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> part of recombinant protein

<400> 119

Asp Val Asp Ile His Thr Leu Ile

1 5

<210> 120

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> part of recombinant protein

<400> 120

Asp Ser Asn Ile Leu Lys Thr Ile Lys Ile Pro

1 5 10

<210> 121

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Linker

<400> 121

Gly Gly Gly Gly Ser

1 5