



**(19) 대한민국특허청(KR)**  
**(12) 공개특허공보(A)**

(11) 공개번호 10-2020-0038528  
(43) 공개일자 2020년04월13일

- (51) 국제특허분류(Int. Cl.)  
*C07K 14/47* (2006.01) *C07K 14/82* (2006.01)  
*C07K 16/18* (2006.01) *C07K 16/32* (2006.01)  
*C07K 19/00* (2006.01) *C12N 15/63* (2006.01)  
*C12N 9/12* (2006.01) *C12Q 1/02* (2017.01)  
*C12Q 1/6806* (2018.01) *C12Q 1/6813* (2018.01)  
*G01N 33/15* (2006.01)
- (52) CPC특허분류  
*C07K 14/47* (2013.01)  
*C07K 14/82* (2013.01)
- (21) 출원번호 10-2020-7007893
- (22) 출원일자(국제) 2018년08월20일  
 심사청구일자 없음
- (85) 번역문제출일자 2020년03월18일
- (86) 국제출원번호 PCT/JP2018/030688
- (87) 국제공개번호 WO 2019/039439  
 국제공개일자 2019년02월28일
- (30) 우선권주장  
 JP-P-2017-158796 2017년08월21일 일본(JP)
- (71) 출원인  
 다이호야쿠힌교교 가부시키키가이샤  
 일본국 도쿄도 지요다쿠 간다니시키키초 1초메 27반  
 치
- (72) 발명자  
 하야시, 고헤이  
 일본 3002611 이바라키켄 츠쿠바시 오쿠보 3 다이  
 호야쿠힌교교 가부시키키가이샤 내
- 이시다, 게이지  
 일본 3002611 이바라키켄 츠쿠바시 오쿠보 3 다이  
 호야쿠힌교교 가부시키키가이샤 내
- (74) 대리인  
 장수길, 이석재

전체 청구항 수 : 총 23 항

(54) 발명의 명칭 **DCTN1 단백질과 RET 단백질과의 융합 단백질**

**(57) 요약**

DCTN1 단백질의 일부와 RET 단백질의 일부가 융합한 신규 폴리펩티드, 상기 폴리펩티드를 코딩하는 폴리뉴클레오티드, 상기 폴리뉴클레오티드 또는 폴리펩티드의 검출 방법, 상기 폴리뉴클레오티드의 발현 또는 폴리펩티드의 발현 및/또는 활성을 억제하는 화합물을 스크리닝하는 방법, 및 RET를 저해하는 화합물을 유효 성분으로 하는 의약 조성물.

(52) CPC특허분류

*C07K 16/18* (2013.01)

*C07K 16/32* (2013.01)

*C07K 19/00* (2013.01)

*C12N 15/63* (2013.01)

*C12N 9/12* (2013.01)

*C12Q 1/02* (2013.01)

*C12Q 1/6806* (2018.05)

*C12Q 1/6813* (2018.05)

*G01N 33/15* (2019.01)

---

## 명세서

### 청구범위

#### 청구항 1

DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드.

#### 청구항 2

제1항에 있어서, 이하의 (a) 내지 (c)에서 선택되는 폴리펩티드.

(a) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열을 포함하는 폴리펩티드.

(b) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열에 있어서, 1개 또는 수개의 아미노산이 치환, 결실, 또는 부가된 아미노산 서열을 포함하는 폴리펩티드.

(c) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열과 90% 이상의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드.

#### 청구항 3

제1항 또는 제2항에 기재된 폴리펩티드를 코딩하는 폴리뉴클레오티드.

#### 청구항 4

제3항에 있어서, 이하의 (d) 내지 (f)에서 선택되는 폴리뉴클레오티드.

(d) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.

(e) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열에 있어서, 1개 또는 수개의 아미노산이 치환, 결실, 또는 부가된 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.

(f) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열과 90% 이상의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.

#### 청구항 5

제3항에 있어서, 이하의 (g) 내지 (i)에서 선택되는 폴리뉴클레오티드.

(g) 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열을 포함하는 폴리뉴클레오티드.

(h) 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열과 상보적인 염기 서열을 포함하는 폴리뉴클레오티드와 엄격한 조건 하에서 하이브리다이징하는 폴리뉴클레오티드.

(i) 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열과 90% 이상의 동일성을 갖는 폴리뉴클레오티드.

#### 청구항 6

제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드를 포함하는 발현 벡터.

**청구항 7**

제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드를 도입한 세포.

**청구항 8**

제1항 또는 제2항에 기재된 폴리펩티드에 특이적으로 결합하는 항체.

**청구항 9**

시료 중 제1항 또는 제2항에 기재된 폴리펩티드의 존재를 검출하는 방법.

**청구항 10**

시료 중 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하기 위한 프라이머 또는 프로브로서, 당해 프라이머 또는 프로브가 이하의 (j) 내지 (l)에서 선택되는 폴리뉴클레오티드.

(j) DCTN1 단백질을 코딩하는 폴리뉴클레오티드에 하이브리다이징하는 프로브 및 RET 단백질을 코딩하는 폴리뉴클레오티드에 하이브리다이징하는 프로브로 이루어지는 군에서 선택되는 적어도 하나의 프로브인 폴리뉴클레오티드.

(k) DCTN1 단백질을 코딩하는 폴리뉴클레오티드와 RET 단백질을 코딩하는 폴리뉴클레오티드와의 융합점에 하이브리다이징하는 프로브인 폴리뉴클레오티드.

(l) DCTN1 단백질을 코딩하는 폴리뉴클레오티드와 RET 단백질을 코딩하는 폴리뉴클레오티드와의 융합점을 끼워 넣도록 설계된 센스 프라이머와 안티센스 프라이머의 세트인 폴리뉴클레오티드.

**청구항 11**

시료 중 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하는 방법.

**청구항 12**

제9항 또는 제11항에 기재된 검출 방법에 있어서, 시료 중 제1항 또는 제2항에 기재된 폴리펩티드 또는 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출한 경우에, 시료의 유래가 되는 환자가 암 이라고 판정하는 방법.

**청구항 13**

RET를 저해하는 화합물을 유효 성분으로 하는, DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 치료용 의약 조성물.

**청구항 14**

이하의 (1) 및 (2)의 공정을 포함하는 제1항 또는 제2항에 기재된 폴리펩티드의 발현 및/또는 활성화, 또는 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 발현을 억제하는 화합물을 스크리닝하는 방법.

(1) 제1항 또는 제2항에 기재된 폴리펩티드, 제1항 또는 제2항에 기재된 폴리펩티드 또는 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드를 발현하고 있는 세포, 또는 제7항의 세포에 시험 화합물을 접촉하는 공정.

(2) 상기 공정 (1)에 있어서, 제1항 또는 제2항에 기재된 폴리펩티드의 발현 및/또는 활성화, 또는 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 발현이 억제되는지 측정하는 공정, 또는 상기 공정 (1)에 기재된 세포의 증식이 억제되는지 측정하는 공정.

**청구항 15**

제1항 또는 제2항에 기재된 폴리펩티드 또는 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드를, RET를 저해하는 화합물을 사용한 화학 요법이 유효한지 여부의 지표로 하는 방법으로서, 제9항에 기재된 검출

방법에 의해 시료 중에서 제1항 또는 제2항에 기재된 폴리펩티드를 검출한 경우, 및/또는 제11항에 기재된 검출 방법에 의해 시료 중에서 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출한 경우에, RET를 저해하는 화합물을 사용한 화학 요법이 유효하다고 판정하는 방법.

**청구항 16**

DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드 및 당해 폴리펩티드를 코딩하는 폴리뉴클레오티드로 이루어지는 균에서 선택되는 적어도 1종을 포함하는 암을 검출하기 위한 바이오마커.

**청구항 17**

DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 환자에게, RET를 저해하는 화합물을 사용한 화학 요법을 행하는 공정을 포함하는, 암의 치료 방법.

**청구항 18**

피험자 유래의 시료 중에서 제1항 또는 제2항에 기재된 폴리펩티드의 존재를 검출하는 것 및/또는 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하는 것을 행하는 공정, 그리고

제1항 또는 제2항에 기재된 폴리펩티드의 존재가 검출되고/되거나 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재가 검출된 경우에, 당해 피험자에 대하여, RET를 저해하는 화합물을 사용한 화학 요법을 행하는 공정을 포함하는, 암의 치료 방법.

**청구항 19**

DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 환자를 치료하기 위한, RET를 저해하는 화합물.

**청구항 20**

DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 환자를 치료하기 위한 암 치료용 의약 조성물을 제조하기 위한, RET를 저해하는 화합물의 사용.

**청구항 21**

시료 중 제1항 또는 제2항에 기재된 폴리펩티드의 존재를 검출하기 위한 수단 및/또는 시료 중 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하기 위한 수단의, RET를 저해하는 화합물을 사용한 화학 요법이 유효한지 여부의 판정약의 제조 방법.

**청구항 22**

제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하기 위한 항DCTN1 항체 및 항RET 항체의 조합.

**청구항 23**

제1항 또는 제2항에 기재된 폴리펩티드 또는 제3항 내지 제5항 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하기 위한 검출약을 제조하기 위한, 제8항에 기재된 항체, 제22항에 기재된 항체의 조합 또는 제10항에 기재된 프라이머 또는 프로브의 사용.

**발명의 설명**

**기술 분야**

본 발명은 DCTN1 단백질과 RET 단백질과의 융합 단백질인 폴리펩티드, 당해 폴리펩티드를 코딩하는 폴리뉴클레오티드, 당해 폴리펩티드 또는 폴리뉴클레오티드의 검출 방법, 당해 폴리펩티드 또는 폴리뉴클레오티드를 표적으로 한 화합물, 당해 화합물을 스크리닝하는 방법에 관한 것이다.

[0001]

**배경 기술**

- [0002] 암은 일본에 있어서의 사인 제1위인 질환이며, 그 치료법의 개선이 요구되고 있다. 갑상선암의 이환자수는 증가하고 있지만, 많은 증례에서 진행이 느려 초기 단계에 적절한 치료를 하는 것으로 높은 생존율을 나타낸다. 반면에 자각 증상이 거의 없기 때문에 적절한 치료에는 조기의 진단이 불가결하다.
- [0003] 갑상선암은 조직형에 따라 유두암, 여포암, 수양암, 미분화암, 악성 림프종으로 나뉜다. 유두암은 갑상선암의 80% 정도를 차지하고 있고, 또한 미분화암은 저빈도이지만 예후가 매우 나쁜 것이 알려져 있다(비특허문헌 1).
- [0004] 유두암에 있어서, 대부분이 암 유전자의 활성화에 의해 발생하는 것이 알려져 있으며 BRAF 변이 유전자(50 내지 60%), RAS 변이 유전자(10 내지 20%), RET 융합 유전자(5 내지 10%) 등의 유전자 이상이 상호 배타적으로 발생하는 것이 명확해지고 있다. 또한 비소세포 폐암에 있어서도, RET 융합 유전자가 1 내지 2%의 빈도로 EGFR 변이 유전자 등의 다른 드라이버 변이 유전자와 상호 배타적으로 존재하는 것이 보고되어 있다(비특허문헌 2 내지 5).
- [0005] 진행한 갑상선암에 있어서는 약물 치료가 주체가 되어, 다양한 멀티키나아제 저해제가 승인되었지만, 드라이버 변이 유전자 특이적으로 효과를 나타내는 약제는 아직 승인되지 않았다. 한편, 폐암에 있어서는 RET 융합 유전자 양성인 환자에서의 RET를 저해함으로써 효과를 나타내는 것이 보고되어 있으며(비특허문헌 6), 갑상선암에 있어서도 변이 유전자나 융합 유전자와 같은 약제 효과의 지표가 될 수 있는 유전자 이상을 동정할 필요가 있다.
- [0006] 암에 있어서, 드라이버가 되는 변이 유전자(변이 단백질)나 융합 유전자(융합 단백질) 등을 동정하는 것은, 암의 성질을 명백히 함과 함께, 이들 변이 유전자나 융합 유전자를 표적으로 한 새로운 암 치료약이나 검사 방법의 개발에 크게 기여하기 때문에, 강하게 요망되고 있다.
- [0007] 그러나, 암 발생의 드라이버가 될 수 있는 변이 유전자나 융합 유전자 등은 충분히 해명되지 않았으며, 약제의 치료 효과에 관련될 수 있는 유전자 이상을 동정하는 것은 매우 의미가 있다.

**선행기술문헌**

**비특허문헌**

- [0008] (비특허문헌 0001) Cancer, 115(16), pp3801-7(2009)
- (비특허문헌 0002) Oncogene, 22(29), pp4578-80(2003)
- (비특허문헌 0003) Cell, 159(3), pp676-90(2014)
- (비특허문헌 0004) Cancer Discov., 3(6), pp630-5(2013)
- (비특허문헌 0005) Nature, 511(7511), pp543-50(2014)
- (비특허문헌 0006) Lancet Respir Med., 5(1), pp42-50(2017)

**발명의 내용**

**해결하려는 과제**

- [0009] 본 발명은, RET 단백질의 적어도 일부를 포함하는 융합 단백질인 신규 폴리펩티드, 또는 당해 폴리펩티드를 코딩하는 폴리뉴클레오티드, 당해 폴리펩티드 또는 폴리뉴클레오티드의 검출 방법, 당해 폴리펩티드 또는 폴리뉴클레오티드를 표적으로 한 화합물, 당해 화합물을 스크리닝하는 방법을 제공하는 것을 과제로 한다.

**과제의 해결 수단**

- [0010] 본 발명자들은, 상기 과제를 해결하기 위해 예의 검토한 결과, 갑상선암 환자 유래의 세포에 있어서, DCTN1 단백질의 일부와 RET 단백질의 일부가 융합한 신규 폴리펩티드 및 당해 폴리펩티드를 코딩하는 폴리뉴클레오티드를 동정했다. 또한, 암 세포에 있어서의 본 발명의 폴리뉴클레오티드 또는 폴리펩티드의 검출 방법, 상기 폴리뉴클레오티드의 발현 또는 폴리펩티드의 발현 및/또는 활성을 억제하는 화합물을 스크리닝하는 방법을

발견했다. 다종다양하게 있는 단백질 중, DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분과의 조합을 갖는 융합 단백질이 세포 내에서 자연스럽게 발생하는 것, 또한 DCTN1과 RET와의 융합 유전자가 암 드라이버로서 작용하기 때문에, 상기 융합 단백질이, 암 진단에 유효한 것은 신규의 지견이며 또한 종래 기술로부터 예상할 수 없는 사항이다. 또한 당해 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 암 환자를 치료하기 위한 RET를 저해하는 화합물을 유효 성분으로 하는 의약 조성물을 발견하고, 본 발명을 완성하기에 이르렀다.

- [0011] 즉, 본 발명은, 이하의 양태를 제공하는 것이다.
- [0012] 항 1. DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드.
- [0013] 항 2. 이하의 (a) 내지 (c)에서 선택되는 항 1에 기재된 폴리펩티드.
- [0014] (a) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열을 포함하는 폴리펩티드.
- [0015] (b) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열에 있어서, 1개 또는 수개의 아미노산이 치환, 결실, 또는 부가된 아미노산 서열을 포함하는 폴리펩티드.
- [0016] (c) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열과 90% 이상의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드.
- [0017] 항 3. 항 1 또는 2에 기재된 폴리펩티드를 코딩하는 폴리뉴클레오티드.
- [0018] 항 4. 이하의 (d) 내지 (f)에서 선택되는 항 3에 기재된 폴리뉴클레오티드.
- [0019] (d) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.
- [0020] (e) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열에 있어서, 1개 또는 수개의 아미노산이 치환, 결실, 또는 부가된 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.
- [0021] (f) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열과 90% 이상의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.
- [0022] 항 5. 이하의 (g) 내지 (i)에서 선택되는 항 3에 기재된 폴리뉴클레오티드.
- [0023] (g) 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열을 포함하는 폴리뉴클레오티드.
- [0024] (h) 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열과 상보적인 염기 서열을 포함하는 폴리뉴클레오티드와 엄격한 조건 하에서 하이브리다이징하는 폴리뉴클레오티드.
- [0025] (i) 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열과 90% 이상의 동일성을 갖는 폴리뉴클레오티드.
- [0026] 항 6. 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드를 포함하는 발현 벡터.
- [0027] 항 7. 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드를 도입한 세포.
- [0028] 항 8. 항 1 또는 2에 기재된 폴리펩티드에 특이적으로 결합하는 항체.
- [0029] 항 9. 시료 중 항 1 또는 2에 기재된 폴리펩티드의 존재를 검출하는 방법.
- [0030] 항 10. 시료 중 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하기 위한 프라이머 또는

프로브로서, 당해 프라이머 또는 프로브가 이하의 (j) 내지 (1)에서 선택되는 폴리뉴클레오티드.

- [0031] (j) DCTN1 단백질을 코딩하는 폴리뉴클레오티드에 하이브리다이징하는 프로브 및 RET 단백질을 코딩하는 폴리뉴클레오티드에 하이브리다이징하는 프로브로 이루어지는 군에서 선택되는 적어도 하나의 프로브인 폴리뉴클레오티드.
- [0032] (k) DCTN1 단백질을 코딩하는 폴리뉴클레오티드와 RET 단백질을 코딩하는 폴리뉴클레오티드와의 융합점에 하이브리다이징하는 프로브인 폴리뉴클레오티드.
- [0033] (1) DCTN1 단백질을 코딩하는 폴리뉴클레오티드와 RET 단백질을 코딩하는 폴리뉴클레오티드와의 융합점을 끼워 넣도록 설계된 센스 프라이머와 안티센스 프라이머의 세트인 폴리뉴클레오티드.
- [0034] 항 11. 시료 중 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하는 방법.
- [0035] 항 12. 항 9 또는 11에 기재된 검출 방법에 있어서, 시료 중 항 1 또는 2에 기재된 폴리펩티드 또는 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출한 경우에, 시료의 유래가 되는 환자가 암이라고 판정하는 방법.
- [0036] 항 13. RET를 저해하는 화합물을 유효 성분으로 하는, DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 치료용 의약 조성물.
- [0037] 항 14. 이하의 (1) 및 (2)의 공정을 포함하는 항 1 또는 2에 기재된 폴리펩티드의 발현 및/또는 활성화, 또는 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 발현을 억제하는 화합물을 스크리닝하는 방법.
- [0038] (1) 항 1 또는 2에 기재된 폴리펩티드, 항 1 또는 2에 기재된 폴리펩티드 또는 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드를 발현하고 있는 세포, 또는 항 7에 기재된 세포에 시험 화합물을 접촉하는 공정.
- [0039] (2) 상기 공정 (1)에 있어서, 항 1 또는 2에 기재된 폴리펩티드의 발현 및/또는 활성화, 또는 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 발현이 억제되는지 측정하는 공정, 또는 상기 공정 (1)에 기재된 세포의 증식이 억제되는지 측정하는 공정.
- [0040] 항 15. 항 1 또는 2에 기재된 폴리펩티드 또는 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드를, RET를 저해하는 화합물을 사용한 화학 요법이 유효한지 여부의 지표로 하는 방법으로서, 항 9에 기재된 검출 방법에 의해 시료 중에서 항 1 또는 2에 기재된 폴리펩티드를 검출한 경우, 및/또는 항 11에 기재된 검출 방법에 의해 시료 중에서 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출한 경우에, RET를 저해하는 화합물을 사용한 화학 요법이 유효하다고 판정하는 방법.
- [0041] 항 16. DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드 및 당해 폴리펩티드를 코딩하는 폴리뉴클레오티드로 이루어지는 군에서 선택되는 적어도 1종을 포함하는 암을 검출하기 위한 바이오마커.
- [0042] 항 17. DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 환자에게, RET를 저해하는 화합물을 사용한 화학 요법을 행하는 공정을 포함하는, 암의 치료 방법.
- [0043] 항 18. 피험자 유래의 시료 중에서 항 1 또는 2에 기재된 폴리펩티드의 존재를 검출하는 것 및/또는 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하는 것을 행하는 공정, 그리고
- [0044] 항 1 또는 2에 기재된 폴리펩티드의 존재가 검출되고/되거나 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재가 검출된 경우에, 당해 피험자에 대하여, RET를 저해하는 화합물을 사용한 화학 요법을 행하는 공정을 포함하는, 암의 치료 방법.
- [0045] 항 19. DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 환자를 치료하기 위한, RET를 저해하는 화합물.
- [0046] 항 20. DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 환자를 치료하기 위한 암 치료용 의약 조성물을 제조하기 위한, RET를 저해하는 화합물의 사용.
- [0047] 항 21. 시료 중 항 1 또는 2에 기재된 폴리펩티드의 존재를 검출하기 위한 수단 및/또는 시료 중 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하기 위한 수단의, RET를 저해하는 화합물을 사용한 화학 요법이 유효한지 여부의 판정약의 제조 방법.

[0048] 항 22. 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하기 위한 항DCTN1 항체 및 항RET 항체의 조합.

[0049] 항 23. 항 1 또는 2에 기재된 폴리펩티드 또는 항 3 내지 5 중 어느 한 항에 기재된 폴리뉴클레오티드의 존재를 검출하기 위한 검출약을 제조하기 위한, 항 8에 기재된 항체, 항 22에 기재된 항체의 조합 또는 항 10에 기재된 프라이머 또는 프로브의 사용.

**발명의 효과**

[0050] 본 발명에 따르면, 본 발명의 폴리뉴클레오티드 및/또는 폴리펩티드는 암 세포에서 특이적으로 발현하고 있는 것이 나타났다. 또한, 본 발명의 폴리뉴클레오티드, 폴리펩티드 및 당해 폴리뉴클레오티드 및/또는 폴리펩티드를 발현하고 있는 세포는, 본 발명의 폴리뉴클레오티드의 발현 또는 폴리펩티드의 발현 및/또는 활성을 억제하는 화합물을 스크리닝하는 방법에 사용할 수 있다. 또한, 본 발명의 폴리뉴클레오티드 및/또는 폴리펩티드의 존재를 지표로 함으로써, DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 대상 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성 대상의 검출을 행하는 것이 가능하다. 또한, RET를 저해하는 화합물은, DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 치료용 의약 조성물로서 유용하다.

**도면의 간단한 설명**

[0051] 도 1은 액적 디지털(Droplet Digital) PCR(ddPCR)을 사용한 갑상선암 조직 중의 DCTN1-RET 융합 유전자 및 GAPDH의 발현 확인.

도 2는 액적 디지털 PCR(ddPCR)을 사용한 정상 갑상선 조직 중의 DCTN1-RET 융합 유전자 및 GAPDH의 발현 확인.

도 3은 정상 갑상선 조직 및 갑상선암 조직 중의 DCTN1-RET 융합 유전자 전장의 발현 확인.

도 4는 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에 있어서의 DCTN1-RET 융합 단백질의 발현 확인. a) 항 인산화 RET 항체를 사용한 DCTN1-RET 융합 단백질의 검출, b) 항RET 항체를 사용한 DCTN1-RET 융합 단백질의 검출, c) 항DCTN1 항체를 사용한 DCTN1-RET 융합 단백질의 검출.

도 5는 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 3차원 배양에 있어서의 증식 확인. N=3, 평균+SD.

도 6은 생체 내(in vivo)에 있어서의 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 종양 형성성 확인.

도 7은 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에 있어서의 RET siRNA에 의한 인산화 RET의 발현 억제의 확인.

도 8은 RET siRNA에 의한 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 증식 억제 효과의 확인.

도 9는 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에 있어서의 RET를 저해하는 화합물에 의한 인산화 RET의 발현 억제의 확인.

**발명을 실시하기 위한 구체적인 내용**

[0052] 본 발명은, 신규 폴리뉴클레오티드 또는 폴리펩티드, 당해 폴리뉴클레오티드 또는 폴리펩티드의 검출 방법, 당해 폴리뉴클레오티드 또는 폴리펩티드를 표적으로 한 화합물, 당해 화합물을 스크리닝하는 방법에 관한 것이다.

[0053] 본 발명은, DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드(이하, 「본 발명의 폴리펩티드」라고도 칭한다)를 제공한다. 또한, 본 발명은, 당해 폴리펩티드를 코딩하는 폴리뉴클레오티드(이하, 「본 발명의 폴리뉴클레오티드」라고도 칭한다)를 제공한다.

[0054] 본 발명에 관한 「DCTN1(다이낙틴 서브유닛(Dynactin Subunit) 1) 단백질」은, 150 kDa 디네인-연관 폴리펩티드(dynein-associated polypeptide) 단백질, 또는 DAP-150 단백질이라고도 칭해지는 단백질이며, 인간 또는 비인간 포유 동물의 DCTN1 단백질을 포함하고, 바람직하게는 인간 DCTN1 단백질이다. 인간에 있어서 2p13.1에 좌승해 있는 유전자에 코딩되어 있는 단백질이다. 본 발명에 있어서, 「DCTN1 단백질」은, 그의 스플라이스 변이체인 이소형을 포함하고, 인간 유래의 것이면, 예를 들어 GenPept 액세스 번호 NP\_004073, NP\_075408, NP\_001128512, NP\_001128513, NP\_00117765, 또는 NP\_00117766으로 나타나는 아미노산 서열을 포함하는 폴리펩티드를 들 수 있다. 또한, 보다 구체적으로는, 예를 들어 서열번호 25, 서열번호 26, 서열번호 27, 서열번호

28, 서열번호 29, 또는 서열번호 30으로 나타나는 아미노산 서열을 포함하는 폴리펩티드를 들 수 있다. 또한, 본 발명에 있어서, 「DCTN1 단백질의 N 말단 부분」이란, 상기 DCTN1 단백질의 N 말단측에 있는 코일드-코일 도메인의 일부 또는 전부를 포함하는 폴리펩티드이고, 바람직하게는 DCTN1 단백질의 N 말단측에 있는 코일드-코일 도메인의 전부를 포함하는 폴리펩티드이다.

[0055] 본 발명에 관한 「RET 단백질」은, Ret 원종양유전자 단백질, RET 수용체 티로신 키나아제(Receptor Tyrosine Kinase) 단백질, 또는 형질감염 중 재배열(Rearranged During Transfection) 단백질이라고도 칭해지는 단백질이며, 인간 또는 비인간 포유 동물의 RET 단백질을 포함하고, 바람직하게는 인간 RET 단백질이다. 인간에 있어서 10q11.2에 좌측해 있는 유전자에 코딩되어 있는 단백질이다. 본 발명에 있어서, 「RET 단백질」은, 그의 스플라이스 변이체인 이소형을 포함하고, 인간 유래의 것이면, 예를 들어 GenPept 액세스 번호 NP\_066124, 또는 NP\_065681로 나타나는 아미노산 서열을 포함하는 폴리펩티드를 들 수 있다. 또한, 보다 구체적으로는, 예를 들어 서열번호 31, 또는 서열번호 32로 나타나는 아미노산 서열을 포함하는 폴리펩티드를 들 수 있다. 또한, 본 발명에 있어서, 「RET 단백질의 C 말단 부분」이란, 상기 RET 단백질의 C 말단측에 있는 키나아제 도메인을 포함하는 폴리펩티드이다.

[0056] 또한, 본 발명의 「DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드」로서는, 상기 DCTN1 단백질의 N 말단측에 있는 코일드-코일 도메인의 일부 또는 전부를 포함하는 폴리펩티드와, 상기 RET 단백질의 C 말단측에 있는 키나아제 도메인을 포함하는 폴리펩티드가 융합되어 있는 폴리펩티드이고, 바람직하게는 상기 DCTN1 단백질의 N 말단측에 있는 코일드-코일 도메인의 전부를 포함하는 폴리펩티드와, 상기 RET 단백질의 C 말단측에 있는 키나아제 도메인을 포함하는 폴리펩티드가 융합되어 있는 폴리펩티드이고, 보다 바람직하게는 이하의 (a) 내지 (c)에서 선택되는 폴리펩티드이다. 이들 폴리펩티드는 키나아제 활성 및/또는 세포 증식 효과를 갖는 것이 바람직하다.

[0057] (a) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열을 포함하는 폴리펩티드.

[0058] (b) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열에 있어서, 1개 또는 수개의 아미노산이 치환, 결실, 또는 부가된 아미노산 서열을 포함하는 폴리펩티드.

[0059] (c) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열과 90% 이상의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드.

[0060] 보다 바람직하게는 이하의 (a) 내지 (c)에서 선택되는 폴리펩티드이다. 이들 폴리펩티드는 키나아제 활성 또는 세포 증식 효과를 갖는 것이 바람직하다.

[0061] (a) 서열번호 18로 나타나는 아미노산 서열을 포함하는 폴리펩티드.

[0062] (b) 서열번호 18로 나타나는 아미노산 서열에 있어서, 1개 또는 수개의 아미노산이 치환, 결실, 또는 부가된 아미노산 서열을 포함하는 폴리펩티드.

[0063] (c) 서열번호 18로 나타나는 아미노산 서열과 90% 이상의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드.

[0064] 본 발명의 「DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드」에는, 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열에 있어서, 1개 또는 수개의 아미노산이 치환, 결실 또는 부가된 아미노산 서열을 포함하는 폴리펩티드(상기 (b))가 포함된다. 이러한 아미노산 서열을 포함하는 DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드로서는, 예를 들어 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열을 갖는 DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드의 이소형을 들 수 있다. 이들 폴리펩티드는 키나아제 활성 또는 세포 증식 효과를 갖는 것이 바람직하다. 여기서, 결실, 치환 또는 부가되는 수개의 아미노산이란, 예를 들어 1 내지 10개가 바람직하고, 보다 바람직하게는 1 내지 5개의 아미노산이다. 또한, 상기 부가에는, N 말단 또는 C 말단으로의 1 내지 수개의 아미노산의 부가, 또는 양 말단으로의 1 내지 수개의 아미노산의 부가가 포함된다.

[0065] 상기 폴리펩티드의 아미노산이 치환된 폴리펩티드로서는, 예를 들어 GenPept 액세스번호: NP\_066124(서열번호 31) 또는 NP\_065681(서열번호 32)로 나타나는 아미노산 서열을 갖는 RET 단백질의 게이트키퍼 부위인 804번째(서열번호 2 및 서열번호 4에서는 1325번째, 서열번호 6 및 서열번호 8에서는 1191번째, 서열번호 10 및 서열번호 12에서는 1300번째, 서열번호 14 및 서열번호 16에서는 1186번째, 서열번호 18 및 서열번호 20에서는 1283번째, 서열번호 22 및 서열번호 24에서는 1318번째)의 발린이 류신, 메티오닌, 또는 글루탐산으로 치환된 폴리펩티드, 또는 806번째(서열번호 2 및 서열번호 4에서는 1327번째, 서열번호 6 및 서열번호 8에서는 1193번째, 서열번호 10 및 서열번호 12에서는 1302번째, 서열번호 14 및 서열번호 16에서는 1188번째, 서열번호 18 및 서열번호 20에서는 1285번째, 서열번호 22 및 서열번호 24에서는 1320번째)의 티로신이 시스테인, 글루탐산, 세린, 히스티딘, 또는 아스파라긴으로 치환된 폴리펩티드를 들 수 있다.

[0066] 또한, 게이트키퍼 부위 이외의 아미노산인 768번째(서열번호 2 및 서열번호 4에서는 1289번째, 서열번호 6 및 서열번호 8에서는 1155번째, 서열번호 10 및 서열번호 12에서는 1264번째, 서열번호 14 및 서열번호 16에서는 1150번째, 서열번호 18 및 서열번호 20에서는 1247번째, 서열번호 22 및 서열번호 24에서는 1282번째)의 글루탐산이 아스파르트산으로 치환된 폴리펩티드, 883번째(서열번호 2 및 서열번호 4에서는 1404번째, 서열번호 6 및 서열번호 8에서는 1270번째, 서열번호 10 및 서열번호 12에서는 1379번째, 서열번호 14 및 서열번호 16에서는 1265번째, 서열번호 18 및 서열번호 20에서는 1362번째, 서열번호 22 및 서열번호 24에서는 1397번째)의 알라닌이 페닐알라닌, 또는 세린으로 치환된 폴리펩티드, 884번째(서열번호 2 및 서열번호 4에서는 1405번째, 서열번호 6 및 서열번호 8에서는 1271번째, 서열번호 10 및 서열번호 12에서는 1380번째, 서열번호 14 및 서열번호 16에서는 1266번째, 서열번호 18 및 서열번호 20에서는 1363번째, 서열번호 22 및 서열번호 24에서는 1398번째)의 글루탐산이 발린으로 치환된 폴리펩티드, 891번째(서열번호 2 및 서열번호 4에서는 1412번째, 서열번호 6 및 서열번호 8에서는 1278번째, 서열번호 10 및 서열번호 12에서는 1387번째, 서열번호 14 및 서열번호 16에서는 1273번째, 서열번호 18 및 서열번호 20에서는 1370번째, 서열번호 22 및 서열번호 24에서는 1405번째)의 세린이 알라닌, 또는 류신으로 치환된 폴리펩티드, 또는 918번째(서열번호 2 및 서열번호 4에서는 1439번째, 서열번호 6 및 서열번호 8에서는 1305번째, 서열번호 10 및 서열번호 12에서는 1414번째, 서열번호 14 및 서열번호 16에서는 1300번째, 서열번호 18 및 서열번호 20에서는 1397번째, 서열번호 22 및 서열번호 24에서는 1432번째)의 메티오닌이 트레오닌으로 치환된 폴리펩티드를 들 수 있지만, 이들에 한정되지 않는다.

[0067] 또한, 본 발명의 DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드에는, 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열에 있어서 상당하는 서열을 적절하게 얼라인먼트했을 때, 상기 서열번호 중 어느 것에서 나타나는 아미노산 서열의 1개와 90% 이상의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드(상기 (c))가 포함된다. 이들 폴리펩티드는 키나아제 활성 또는 세포 증식 효과를 갖는 것이 바람직하다.

[0068] 여기서, 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열과의 동일성은, 바람직하게는 90% 이상이고, 보다 바람직하게는 95% 이상이고, 더욱 바람직하게는 98% 이상이다. 당해 아미노산 서열의 동일성은, 통상 관용의 방법으로 계산할 수 있다.

[0069] 본 발명의 폴리펩티드는, 본 발명에 관한 폴리펩티드를 구성하는 아미노산 서열 이외에, 단백질 태그를 구성하는 아미노산을 갖고 있어도 된다. 태그의 예로서는, 발현 효율을 향상시키는 태그나, 정제 효율을 향상시키는 태그 등, 당업자에게 주지인 태그를 사용할 수 있고, His 태그, Myc 태그, FLAG 태그 등을 들 수 있다.

[0070] 본 발명의 폴리뉴클레오티드는, 상기 DCTN1 단백질의 N 말단 부분과, RET 단백질의 C 말단 부분이 융합되어 있는 폴리펩티드를 코딩하는 폴리뉴클레오티드이고, 바람직하게는 이하의 (d) 내지 (i)에서 선택되는 폴리뉴클레오티드이다. 이들 폴리뉴클레오티드는, 키나아제 활성 또는 세포 증식 효과를 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드인 것이 바람직하다.

[0071] (d) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.

[0072] (e) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열에 있어서, 1개 또는 수개

의 아미노산이 치환, 결실, 또는 부가된 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.

- [0073] (f) 서열번호 2, 서열번호 4, 서열번호 6, 서열번호 8, 서열번호 10, 서열번호 12, 서열번호 14, 서열번호 16, 서열번호 18, 서열번호 20, 서열번호 22, 또는 서열번호 24로 나타나는 아미노산 서열과 90% 이상의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.
- [0074] (g) 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열을 포함하는 폴리뉴클레오티드.
- [0075] (h) 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열과 상보적인 염기 서열을 포함하는 폴리뉴클레오티드와 엄격한 조건 하에서 하이브리다이징하는 폴리뉴클레오티드.
- [0076] (i) 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열과 90% 이상의 동일성을 갖는 폴리뉴클레오티드.
- [0077] 보다 바람직하게는 이하의 (d) 내지 (i)에서 선택되는 폴리뉴클레오티드이다. 이들 폴리뉴클레오티드는, 키나아제 활성 또는 세포 증식 효과를 갖는 폴리펩티드를 코딩하는 폴리뉴클레오티드인 것이 바람직하다.
- [0078] (d) 서열번호 18로 나타나는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.
- [0079] (e) 서열번호 18로 나타나는 아미노산 서열에 있어서, 1개 또는 수개의 아미노산이 치환, 결실, 또는 부가된 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.
- [0080] (f) 서열번호 18로 나타나는 아미노산 서열과 90% 이상의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드.
- [0081] (g) 서열번호 17로 나타나는 염기 서열을 포함하는 폴리뉴클레오티드.
- [0082] (h) 서열번호 17로 나타나는 염기 서열과 상보적인 염기 서열을 포함하는 폴리뉴클레오티드와 엄격한 조건 하에서 하이브리다이징하는 폴리뉴클레오티드.
- [0083] (i) 서열번호 17로 나타나는 염기 서열과 90% 이상의 동일성을 갖는 폴리뉴클레오티드.
- [0084] 본 발명의 폴리뉴클레오티드는, 이중쇄 DNA 뿐만 아니라, 그것을 구성하는 센스쇄 및 안티센스쇄와 같은 각종 단일쇄 DNA나 RNA도 포함한다. 안티센스쇄는, 프로브 등으로서 이용 가능하다. DNA에는, 예를 들어 클로닝이나 화학 합성 기술 또는 그들의 조합으로 얻어지는 cDNA나 게놈 DNA가 포함된다. 또한, 본 발명에 관한 폴리뉴클레오티드에 대하여, 본 발명에 관한 폴리펩티드를 코딩하는 염기 서열 이외에, 비번역 영역(UTR)의 서열 등의 염기 서열이 부가되어 있어도 된다.
- [0085] 여기서, 엄격한 조건이란, 예를 들어 Molecular Cloning: A Laboratory Manual(Second Edition, J.Sambrook et.al, 1989)에 기재된 조건을 들 수 있다. 즉, 6×SSC(1×SSC의 조성: 0.15M 염화나트륨, 0.015M 시트르산나트륨, pH7.0), 0.5% SDS, 5×덴하르트 및 100mg/mL 청어 정자 DNA를 포함하는 용액에 프로브와 함께 65℃에서 8 내지 16시간 항온하고, 하이브리다이징시키는 조건 등을 들 수 있다.
- [0086] 또한, 서열번호 1, 서열번호 3, 서열번호 5, 서열번호 7, 서열번호 9, 서열번호 11, 서열번호 13, 서열번호 15, 서열번호 17, 서열번호 19, 서열번호 21, 또는 서열번호 23으로 나타나는 염기 서열과의 동일성은, 바람직하게는 90% 이상, 보다 바람직하게는 95% 이상, 더욱 바람직하게는 98% 이상이다. 당해 염기 서열의 동일성은, 통상 관용의 방법으로 계산할 수 있다.
- [0087] 본 명세서 중에 있어서, 「키나아제 활성 또는 세포 증식 효과를 갖는다」의 「키나아제 활성을 갖는다」란, 티로신을 인산화하는 효소로서의 활성을 갖는 것을 의미한다. 또한, 「키나아제 활성 또는 세포 증식 효과를 갖는다」의 「세포 증식 효과를 갖는다」란, 본 발명의 폴리뉴클레오티드 및/또는 폴리펩티드를 세포에 도입함으로써, 도입하지 않은 세포에 비하여, 도입한 세포의 증식능이 향상되는 효과이다. 이러한 효과는, 예를 들어 사이토카인 의존적으로 증식하는 세포주에 폴리뉴클레오티드 및/또는 폴리펩티드를 도입했을 때, 사이토카인 비의존적으로 증식했을 때, 세포 증식 효과를 갖는다고 확인할 수 있다.
- [0088] 본 발명의 폴리뉴클레오티드는, 예를 들어 DCTN1 유전자와 RET 유전자와의 융합 유전자를 유지하는 갑상선암 등

으로부터 조정한 cDNA 라이브러리나 게놈 DNA 라이브러리를 사용하여, 본 발명의 폴리뉴클레오티드의 염기 서열의 일부와 특이적으로 하이브리다이징하는 프라이머를 사용하여, 추출할 수 있다. 이러한 프라이머로서는, 본 발명에 관한 폴리뉴클레오티드 또는 그의 안티센스쇄의 적어도 일부에 특이적으로 하이브리다이징하는 프라이머이면, 어떠한 서열 및 길이의 것을 사용해도 된다. 또한 인공적으로 폴리뉴클레오티드를 합성하는 방법을 들 수 있다(Nat. Methods, 11: 499-507, 2014).

- [0089] 본 발명의 발현 벡터는, 본 발명의 폴리뉴클레오티드를 포함하고, 본 발명의 폴리펩티드를 발현시키는 한, 특별히 한정되는 것은 아니다. 예를 들어, 사용하는 숙주에 따라서 적절히 선택한 공지된 발현 벡터에, 본 발명의 폴리뉴클레오티드를 삽입함으로써 얻어지는 발현 벡터를 들 수 있다.
- [0090] 숙주로서는, 형질 전환이 가능한 생세포이면 특별히 한정되지 않고, 예를 들어, 대장균, 고초균 등의 세균, 효모나 사상균 등의 진균류, Sf9 세포 등의 곤충 세포, 누에 등의 곤충, 동물 세포, 식물 또는 식물 유래 세포를 들 수 있다.
- [0091] 본 발명의 폴리뉴클레오티드를 삽입하기 위한 벡터는, 숙주 안에서 복제 가능한 것이면 특별히 한정되지 않고, 도입하는 숙주의 종류, 도입 방법 등에 따라서 적절히 선택할 수 있다. 예를 들어, 플라스미드 DNA, 파지 DNA, 바이러스 벡터를 들 수 있다. 발현 벡터의 구축에 사용되는 벡터 DNA는, 널리 보급되어 입수가 용이한 것이 사용된다. 예를 들어, pUC19(타카라 바이오), pTV118N(타카라 바이오), pMAMneo(클론 테크), pGEX(GE 헬스케어), pET160(Invitrogen), pDEST(Invitrogen), pIEx(머크 밀리포아), pBacPAK(클론 테크)를 들 수 있다. 또한, 바이러스 벡터로서는, 예를 들어 바큘로바이러스 벡터, 레트로바이러스 벡터, 인간 면역 부전증 바이러스(HIV) 등의 렌티바이러스 벡터, 아데노바이러스 벡터, 아데노 수반 바이러스 벡터(AAV 벡터), 헤르페스 바이러스, 백시니아 바이러스, 폭스 바이러스, 폴리오 바이러스, 신드비스 바이러스, 센다이 바이러스, 시미안 바이러스-40(SV-40) 등의 DNA 바이러스나 RNA 바이러스를 들 수 있다.
- [0092] 당해 발현 벡터를 사용해서 숙주를 형질 전환하기 위해서는, 프로토플라스트 법, 적격 세포법, 일렉트로포레이션법 등을 사용해서 행할 수 있다. 얻어진 형질 전환체는, 자화할 수 있는 탄소원, 질소원, 금속염, 비타민 등을 포함하는 배지를 사용해서 적당한 조건 하에서 배양하면 된다.
- [0093] 본 발명의 폴리뉴클레오티드를 도입한 세포는, 예를 들어 상기한 본 발명의 발현 벡터로 형질 전환된 세포, 게놈 편집에 의해 본 발명의 폴리뉴클레오티드를 도입한 세포를 들 수 있다. 여기서 사용할 수 있는 세포로서는, 상기한 숙주 세포를 들 수 있다. 세포가, 발현 벡터로 형질 전환된 세포인지 확인하는 방법으로서, 예를 들어, 본 발명의 폴리펩티드를 검출하기 위한 방법, 또는 본 발명의 폴리뉴클레오티드를 검출하기 위한 방법을 들 수 있다.
- [0094] 「게놈 편집에 의해 본 발명의 폴리뉴클레오티드를 도입한 세포」로서, 바람직하게는 각각 단독으로 존재하는 DCTN1 유전자와 RET 유전자를 게놈 편집에 의해 융합시킨 유전자를 갖는 세포이고, 보다 바람직하게는 각각 단독으로 존재하는 DCTN1 유전자와 RET 유전자의 DCTN1의 엑손 27과 RET의 엑손 12를 게놈 편집에 의해 융합시킨 유전자를 갖는 세포이다. 이러한 세포는 통상 관용의 방법으로 제작 가능하며, 예를 들어 Cell Rep., 9(4), pp1219-1227(2014), Nat. Commun., 5, 3728(2014)에 기재된 방법을 들 수 있다. 세포가, 게놈 편집에 의해 본 발명의 폴리뉴클레오티드를 도입한 세포인지 확인하는 방법으로서, 예를 들어, 본 발명의 폴리펩티드를 검출하기 위한 방법, 또는 본 발명의 폴리뉴클레오티드를 검출하기 위한 방법을 들 수 있다.
- [0095] 본 발명의 폴리펩티드는, 본 발명의 발현 벡터로 형질 전환된 세포를, 세포 배양에 적합한 배지를 사용하여, 적당한 조건 하에서 배양함으로써 얻어진 배양액 및/또는 세포로부터, 일반적인 방법에 의해 단백질의 채취, 정제를 행함으로써, 얻을 수 있다. 또한, 본 발명의 폴리펩티드는, 본 발명의 폴리뉴클레오티드를 갖는 발현 벡터, 또는 본 발명의 폴리뉴클레오티드를 코딩하는 주형 RNA 또는 주형 DNA를 무세포 단백질 합성계(예를 들어, 인간 세포주 유래의 세포 추출액, 토끼 망상적혈구 추출액, 소맥 배아 추출액, 대장균 추출액)에 도입하고, 적당한 조건 하에서 인큐베이션함으로써 얻어진 반응액으로부터, 일반적인 방법에 의해 단백질의 채취, 정제를 행함으로써, 얻을 수 있다.
- [0096] 본 발명에 있어서, 본 발명의 폴리펩티드에 특이적으로 결합하는 항체는, DCTN1 단백질의 N 말단 부분과 RET 단백질의 C 말단 부분의 융합점에 특이적으로 결합하는 항체를 들 수 있다. 당해 항체는, DCTN1 단백질의 N 말단 부분과 RET 단백질의 C 말단 부분의 융합점에 특이적으로 결합하지만, 야생형의 DCTN1, 또는 야생형의 RET 단백질 중 어디에도 결합하지 않는 항체를 의미한다.
- [0097] 본 발명에 있어서, 「DCTN1 단백질의 N 말단 부분과 RET 단백질의 C 말단 부분의 융합점」에 있어서의 「융합점

」이란, DCTN1 단백질의 N 말단 부분 유래의 폴리펩티드와 RET 단백질의 C 말단 부분 유래의 폴리펩티드가 융합한 점을 의미한다. 서열번호 2에 있어서의 융합점은, 서열번호 2에 있어서의 1-1233번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 2에 있어서의 1234-1635번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 4에 있어서의 융합점은, 서열번호 4에 있어서의 1-1233번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 4에 있어서의 1234-1593번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 6에 있어서의 융합점은, 서열번호 6에 있어서의 1-1099번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 6에 있어서의 1100-1501번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 8에 있어서의 융합점은, 서열번호 8에 있어서의 1-1099번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 8에 있어서의 1100-1459번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 10에 있어서의 융합점은, 서열번호 10에 있어서의 1-1208번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 10에 있어서의 1209-1610번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 12에 있어서의 융합점은, 서열번호 12에 있어서의 1-1208번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 12에 있어서의 1209-1568번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 14에 있어서의 융합점은, 서열번호 14에 있어서의 1-1094번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 14에 있어서의 1095-1496번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 16에 있어서의 융합점은, 서열번호 16에 있어서의 1-1094번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 16에 있어서의 1095-1454번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 18에 있어서의 융합점은, 서열번호 18에 있어서의 1-1191번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 18에 있어서의 1192-1593번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 20에 있어서의 융합점은, 서열번호 20에 있어서의 1-1191번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 20에 있어서의 1192-1551번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 22에 있어서의 융합점은, 서열번호 22에 있어서의 1-1226번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 22에 있어서의 1227-1628번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다. 서열번호 24에 있어서의 융합점은, 서열번호 24에 있어서의 1-1226번째의 DCTN1의 N 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드와 서열번호 24에 있어서의 1227-1586번째의 RET의 C 말단 부분 유래의 아미노산 서열을 갖는 폴리펩티드가 융합한 점이다.

- [0098] 상기 항체는, 예를 들어 면역 글로불린(IgA, IgD, IgE, IgG, IgM, IgY 등), Fab 프래그먼트, F(ab')<sub>2</sub> 프래그먼트, 단일쇄 항체 프래그먼트(scFv), 싱글 도메인 항체, 디아바디(Diabody) 등(Nat.Rev.Immunol., 6:343-357, 2006)을 들 수 있고, 이들은 인간 항체, 인간화 항체, 키메라 항체, 마우스 항체, 라마 항체, 닭 항체 등의 모노클로날 항체 또는 폴리클로날 항체를 들 수 있지만 이들에 한정되는 것은 아니다.
- [0099] 상기 항체는, 다양한 공지된 방법을 사용해서 제작할 수 있고, 제작 방법은 특별히 한정되는 것이 아니다. 이러한 공지된 방법으로서, 당해 발명의 폴리펩티드, DCTN1 단백질의 N 말단 부분과 RET 단백질의 C 말단 부분의 융합점을 포함하는 폴리펩티드 단편 등을 면역 동물에 접종하고, 당해 동물의 면역계를 활성화시킨 후, 당해 동물의 혈청을 회수하고, 폴리클로날 항체로서 얻는 방법, 또는 하이브리도마법, 파지 디스플레이법 등에 의해 모노클로날 항체를 얻는 방법 등을 들 수 있다.
- [0100] 본 발명의 폴리펩티드의 발현 및/또는 활성화, 또는 본 발명의 폴리뉴클레오티드의 발현을 억제하는 화합물을 스크리닝하는 방법은, 이하의 (1) 및 (2)의 공정을 포함하는 방법에 의해 행할 수 있다.
- [0101] 즉, 본 발명의 스크리닝 방법은,
- [0102] (1) 본 발명의 폴리펩티드, 또는 본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 세포에 시험 화합물을 접촉하는 공정.
- [0103] (2) 상기 공정 (1)에 있어서, 본 발명의 폴리펩티드의 발현 및/또는 활성화, 또는 본 발명의 폴리뉴클레오티드의 발현이 억제되는지 측정하는 공정, 또는 상기 공정 (1)에 기재된 세포의 증식이 억제되는지 측정하는 공정.
- [0104] 을 포함하는 방법에 의해 행할 수 있다.

- [0105] 보다 바람직하게는, 이하의 (1) 및 (2)의 공정을 포함하는 방법이다.
- [0106] (1) 본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 세포에 시험 화합물을 접촉하는 공정.
- [0107] (2) 상기 공정 (1)에 기재된 세포의 증식이 억제되는지 측정하는 공정.
- [0108] 또한, 본 발명의 폴리펩티드의 발현 및/또는 활성화, 또는 본 발명의 폴리뉴클레오티드의 발현을 억제하는 화합물을 스크리닝하는 방법은, 이하의 (1) 내지 (3)의 공정을 포함하는 방법에 의해 행할 수 있다.
- [0109] 즉, 본 발명의 스크리닝 방법은,
- [0110] (1) 본 발명의 폴리펩티드, 또는 본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 세포에 시험 화합물을 접촉하는 공정.
- [0111] (2) 상기 공정 (1)에 있어서, 본 발명의 폴리펩티드의 발현 및/또는 활성화, 또는 본 발명의 폴리뉴클레오티드의 발현이 억제되는지 측정하는 공정, 또는 상기 공정 (1)에 기재된 세포의 증식이 억제되는지 측정하는 공정.
- [0112] (3) 상기 공정 (2)에 있어서, 본 발명의 폴리펩티드의 발현 및/또는 활성화, 또는 본 발명의 폴리뉴클레오티드의 발현이 억제된 경우, 또는 상기 공정 (1)에 기재된 세포의 증식이 억제된 경우, 시험 화합물이 본 발명의 폴리펩티드의 발현 및/또는 활성화, 또는 본 발명의 폴리뉴클레오티드의 발현을 억제한다고 판정하는 공정.
- [0113] 을 포함하는 방법에 의해 행할 수 있다.
- [0114] 보다 바람직하게는, 이하의 (1) 내지 (3)의 공정을 포함하는 방법이다.
- [0115] (1) 본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 세포에 시험 화합물을 접촉하는 공정.
- [0116] (2) 상기 공정 (1)에 기재된 세포의 증식이 억제되는지 측정하는 공정.
- [0117] (3) 상기 공정 (2)에 있어서, 상기 공정 (1)에 기재된 세포의 증식이 억제된 경우, 시험 화합물이 본 발명의 폴리펩티드의 발현 및/또는 활성화, 또는 본 발명의 폴리뉴클레오티드의 발현을 억제한다고 판정하는 공정.
- [0118] 「본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 세포」는, 본 발명의 발현 벡터로 형질 전환된 세포, 계능 편집에 의해 본 발명의 폴리뉴클레오티드를 도입한 세포, 본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 초대 배양 세포, 본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 주화 세포, 본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 암 환자 유래의 세포 등을 들 수 있다. 세포가, 본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 세포인지 확인하는 방법으로서, 예를 들어, 본 발명의 폴리펩티드를 검출하기 위한 방법, 또는 본 발명의 폴리뉴클레오티드를 검출하기 위한 방법을 들 수 있다.
- [0119] 본 발명에 있어서, 「본 발명의 폴리펩티드의 발현 및/또는 활성화, 또는 본 발명의 폴리뉴클레오티드의 발현이 억제된다」 중에서, 「본 발명의 폴리펩티드의 발현, 또는 본 발명의 폴리뉴클레오티드의 발현이 억제된다」란, 예를 들어 본 발명의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 세포에 시험 화합물을 접촉시킨 후, 당해 세포에 있어서의 본 발명의 폴리펩티드 또는 폴리뉴클레오티드의 발현량을, 본 발명의 폴리펩티드 또는 폴리뉴클레오티드의 존재를 검출하는 방법을 사용해서 평가했을 때, 시험 화합물을 접촉시키지 않은 세포와 비교하여, 시험 화합물을 접촉시킨 세포에 있어서, 본 발명의 폴리펩티드 또는 폴리뉴클레오티드의 발현량이 통계학상 유의미하게 저하되었을 때, 본 발명의 폴리펩티드 또는 폴리뉴클레오티드의 발현이 억제되었다고 판단할 수 있다.
- [0120] 또한, 「본 발명의 폴리펩티드의 발현 및/또는 활성화, 또는 본 발명의 폴리뉴클레오티드의 발현이 억제된다」 중에서, 「본 발명의 폴리펩티드의 활성이 억제된다」란, 예를 들어 본 발명의 폴리펩티드, 또는 본 발명의 폴리펩티드를 발현하고 있는 세포를 사용하여, 시험 화합물을 접촉시키지 않은 경우와 비교하여, 시험 화합물을 접촉시킨 경우에, 티로신의 인산화의 비율이 통계학상 유의미하게 저하되었을 때, 본 발명의 폴리펩티드 활성이 억제되었다고 판단할 수 있다.
- [0121] 또한, 본 발명의 폴리펩티드를 발현하고 있는 세포를 사용하여, 시험 화합물을 접촉시키지 않은 경우와 비교하여, 시험 화합물을 접촉시킨 경우에, 세포 증식이 통계학상 유의미하게 억제되었을 때, 본 발명의 폴리펩티드의 활성이 억제되었다고 판단할 수 있다.
- [0122] 본 발명에 있어서, 「티로신의 인산화」란, RET 단백질(다른 단백질과 융합되어 있는 RET 단백질도 포함한다)의

티로신의 인산화뿐만 아니라, RET 하류 시그널 상의 단백질의 티로신의 인산화도 포함된다. RET 하류 시그널 상의 단백질로서는, 예를 들어 STAT, AKT, ERK를 들 수 있다. 바람직하게는, RET 단백질(다른 단백질과 융합되어 있는 RET 단백질도 포함한다)의 티로신의 인산화이다.

- [0123] 또한, 「티로신의 인산화의 비율」은, 예를 들어 인산화 RET 특이적 항체를 사용하여, 웨스턴 블로팅, 면역 침강, 면역 조직 화학, ELISA, 플로우 사이토메트리에 의해 측정 가능하다.
- [0124] 본 발명에 있어서 「시료」란, 생체 시료(예를 들어, 세포, 조직, 장기, 체액(혈액, 림프액 등), 소화액, 오줌)뿐만 아니라, 이들 생체 시료로부터 얻어지는 핵산 추출물(게놈 DNA 추출물, mRNA 추출물, mRNA 추출물로부터 제조된 cDNA 제조물이나 cRNA 제조물 등)이나 단백질 추출물도 포함한다. 또한, 상기 시료는, 포르말린 고정 처리, 알코올 고정 처리, 동결 처리 또는 파라핀 포매 처리가 실시되어 있는 것이어도 된다. 상기 생체 시료로서는, 생체로부터 채취한 것을 사용할 수 있다. 바람직하게는 암 환자 유래의 시료이고, 보다 바람직하게는 종양 세포를 포함하는 시료이다. 또한, 생체 시료의 채취 방법은, 생체 시료의 종류에 따라 적절히 선택할 수 있다.
- [0125] 본 발명에는, 시료 중 본 발명의 폴리펩티드의 존재를 검출하는 방법이 포함된다.
- [0126] 본 발명에 있어서, 시료 중 본 발명의 폴리펩티드의 존재를 검출하는 방법은, 본 발명의 폴리펩티드에 특이적으로 결합하는 항체를 사용한 ELISA법, 웨스턴 블로팅법, 또는 면역 조직 화학 염색법, 또는 DCTN1 단백질에 특이적으로 결합하는 항체 및 RET 단백질에 특이적으로 결합하는 항체를 사용한 FRET(Fluorescence Resonance Energy Transfer; 형광 공명 에너지 전달)법 등, 통상 관용의 검출법에 의해 검출하는 방법을 들 수 있다. 바람직하게는, 본 발명의 폴리펩티드에 특이적으로 결합하는 항체를 사용한 ELISA법, 웨스턴 블로팅법, 또는 면역 조직 화학 염색법이다.
- [0127] DCTN1 단백질에 특이적으로 결합하는 항체 및 RET 단백질에 특이적으로 결합하는 항체로서는, DCTN1 단백질의 상기 융합점으로부터 N 말단 부분에 결합하는 항체 및 RET 단백질의 상기 융합점으로부터 C 말단 부분에 결합하는 항체가 바람직하고, 이들 항체는, 시판품을 사용하는 것, 또는 통상 공지된 방법으로 제작하는 것이 가능하다.
- [0128] 본 발명에 있어서, 시료 중 본 발명의 폴리펩티드의 존재를 검출하는 방법으로서, 바람직하게는 본 발명의 폴리펩티드에 특이적으로 결합하는 항체 또는 DCTN1 단백질에 특이적으로 결합하는 항체 및 RET 단백질에 특이적으로 결합하는 항체를 사용해서 본 발명의 폴리펩티드를 검출하는 공정을 포함하는 것을 특징으로 하는, 본 발명의 폴리펩티드의 존재를 검출하는 방법이고, 보다 바람직하게는 본 발명의 폴리펩티드에 특이적으로 결합하는 항체를 사용해서 본 발명의 폴리펩티드를 검출하는 공정을 포함하는 것을 특징으로 하는, 본 발명의 폴리펩티드의 존재를 검출하는 방법이다. 따라서, 본 발명의 폴리펩티드의 존재를 검출하기 위한 수단으로서, 특별히 한정되지 않지만, 예를 들어 DCTN1 단백질에 특이적으로 결합하는 항체 및 RET 단백질에 특이적으로 결합하는 항체의 조합; 본 발명의 폴리펩티드에 특이적으로 결합하는 항체 등을 들 수 있다.
- [0129] 본 발명에는, 시료 중 본 발명의 폴리뉴클레오티드의 존재를 검출하기 위한 프라이머 또는 프로브가 포함된다. 본 발명에 있어서, 본 발명의 폴리펩티드의 존재를 검출하기 위한 수단으로서, 특별히 한정되지 않지만, 예를 들어 본 발명의 폴리뉴클레오티드의 존재를 검출하기 위한 프라이머 또는 프로브 등을 들 수 있다.
- [0130] 당해 프라이머 또는 프로브로서는, (j) 내지 (l)에서 선택되는 폴리뉴클레오티드;
- [0131] (j) DCTN1 단백질을 코딩하는 폴리뉴클레오티드에 하이브리다이징하는 프로브 및 RET 단백질을 코딩하는 폴리뉴클레오티드에 하이브리다이징하는 프로브로 이루어지는 군에서 선택되는 적어도 하나의 프로브인 폴리뉴클레오티드.
- [0132] (k) DCTN1 단백질을 코딩하는 폴리뉴클레오티드와 RET 단백질을 코딩하는 폴리뉴클레오티드와의 융합점에 하이브리다이징하는 프로브인 폴리뉴클레오티드.
- [0133] (l) DCTN1 단백질을 코딩하는 폴리뉴클레오티드와 RET 단백질을 코딩하는 폴리뉴클레오티드와의 융합점을 끼워 넣도록 설계된 센스 프라이머와 안티센스 프라이머의 세트인 폴리뉴클레오티드.
- [0134] 를 들 수 있다.
- [0135] 본 발명에 있어서, 「DCTN1 단백질을 코딩하는 폴리뉴클레오티드와 RET 단백질을 코딩하는 폴리뉴클레오티드와의 융합점」에 있어서의 「융합점」이란, DCTN1 단백질을 코딩하는 폴리뉴클레오티드와 RET 단백질을 코딩하는

폴리뉴클레오티드가 융합한 점을 의미한다. 서열번호 1에 있어서의 융합점은, 서열번호 1에 있어서의 1-3699번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 1에 있어서의 3700-4905번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 3에 있어서의 융합점은, 서열번호 3에 있어서의 1-3699번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 3에 있어서의 3700-4779번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 5에 있어서의 융합점은, 서열번호 5에 있어서의 1-3297번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 5에 있어서의 3298-4503번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 7에 있어서의 융합점은, 서열번호 7에 있어서의 1-3297번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 7에 있어서의 3298-4377번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 9에 있어서의 융합점은, 서열번호 9에 있어서의 1-3624번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 9에 있어서의 3625-4830번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 11에 있어서의 융합점은, 서열번호 11에 있어서의 1-3624번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 11에 있어서의 3625-4704번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 13에 있어서의 융합점은, 서열번호 13에 있어서의 1-3282번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 13에 있어서의 3283-4488번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 15에 있어서의 융합점은, 서열번호 15에 있어서의 1-3282번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 15에 있어서의 3283-4362번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 17에 있어서의 융합점은, 서열번호 17에 있어서의 1-3573번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 17에 있어서의 3574-4779번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 19에 있어서의 융합점은, 서열번호 19에 있어서의 1-3573번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 19에 있어서의 3574-4653번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 21에 있어서의 융합점은, 서열번호 21에 있어서의 1-3678번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 21에 있어서의 3679-4884번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다. 서열번호 23에 있어서의 융합점은, 서열번호 23에 있어서의 1-3678번째의 DCTN1을 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드와 서열번호 23에 있어서의 3679-4758번째의 RET를 코딩하는 폴리뉴클레오티드 유래의 염기 서열을 갖는 폴리뉴클레오티드가 융합한 점이다.

- [0136] 본 발명에 있어서, 프라이머 또는 프로브는, 본 발명의 폴리뉴클레오티드 서열 정보에 기초하여, 본 발명의 폴리뉴클레오티드와 특이적으로 하이브리다이징하는 폴리뉴클레오티드로서, 통상 공지된 방법에 의해 제작된다. 당해 프라이머 또는 프로브의 염기수는, 10 내지 50 염기, 바람직하게는 15 내지 50 염기, 보다 바람직하게는 18 내지 35 염기이다.
- [0137] 당해 프라이머 또는 프로브는, 본 발명의 폴리뉴클레오티드와 특이적으로 하이브리다이징하는 것이면, 완전히 상보적일 필요는 없다. 이러한 프라이머 또는 프로브는, 대응하는 염기 서열과 비교하여, 70% 이상, 바람직하게는 80% 이상, 보다 바람직하게는 90% 이상, 보다 바람직하게는 95% 이상, 보다 바람직하게는 98% 이상의 동일성을 갖는 폴리뉴클레오티드이다.
- [0138] 본 발명의 프라이머 또는 프로브로서는, 바람직하게는 (i) 서열번호 69, (ii) 서열번호 70, 또는 (iii) 서열번호 71로 표현되는 폴리뉴클레오티드이고, 보다 바람직하게는 (iv) 서열번호 69와 서열번호 70으로 표현되는 센스 프라이머와 안티센스 프라이머의 세트인 폴리뉴클레오티드이고, 보다 바람직하게는 (v) 서열번호 69, 서열번호 70 및 서열번호 71로 표현되는 센스 프라이머, 안티센스 프라이머 및 프로브의 세트인 폴리뉴클레오티드이다.
- [0139] 본 발명에는, 시료 중 본 발명의 폴리뉴클레오티드의 존재를 검출하는 방법이 포함된다.
- [0140] 본 발명에 있어서, 시료 중 본 발명의 폴리뉴클레오티드의 존재를 검출하는 방법은, 노던 블로팅법, 서던 블로팅법, RT-PCR법, 리얼타임 PCR법, 디지털 PCR법, DNA 마이크로어레이법, 계내(in situ) 하이브리다이제이션법,

시퀀스 해석법 등, 통상 관용의 검출법에 의해 검출하는 방법이다.

- [0141] 본 발명에 있어서, 시료 중 본 발명의 폴리뉴클레오티드의 존재를 검출하는 방법에는, 본 발명의 폴리뉴클레오티드를 포함하는 RET 융합 유전자의 폴리뉴클레오티드의 존재를 검출하기 위한 방법도 포함된다. 당해 방법으로서, RET 단백질을 코딩하는 폴리뉴클레오티드에 하이브리다이징하는 프라이머(예를 들어, RET 키나아제 도메인 이후의 3'측의 서열에 하이브리다이징하는 프라이머)를 사용하여, 5'RACE법에 의해 증폭시킨 PCR 산물을 시퀀스 해석하는 방법 등을 들 수 있다.
- [0142] 본 발명에 있어서, 시료 중 본 발명의 폴리뉴클레오티드의 존재를 검출하는 방법으로서, 바람직하게는 본 발명의 프라이머 또는 프로브를 사용해서 본 발명의 폴리뉴클레오티드를 검출하는 공정을 포함하는 것을 특징으로 하는, 본 발명의 폴리뉴클레오티드의 존재를 검출하는 방법이다.
- [0143] 본 발명에는, RET를 저해하는 화합물을 유효 성분으로 하는, DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 치료용 의약 조성물이 포함된다.
- [0144] 보다 바람직하게는, 본 발명의 폴리펩티드의 발현 및/또는 활성, 또는 본 발명의 폴리뉴클레오티드의 발현을 억제하는 화합물을 유효 성분으로 하는, DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 치료용 의약 조성물이 포함된다.
- [0145] 본 발명에 있어서, 「DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암」에 있어서의 「DCTN1 유전자와 RET 유전자와의 융합 유전자 양성인 암」이란, 본 발명의 폴리뉴클레오티드가 발현하고 있는 암이고, 바람직하게는 본 발명의 폴리뉴클레오티드의 존재를 검출하는 방법을 사용하여, 본 발명의 폴리뉴클레오티드가 검출된 암이다.
- [0146] 또한, 본 발명에 있어서, 「DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암」에 있어서의 「DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암」이란, 본 발명의 폴리펩티드가 발현하고 있는 암이고, 바람직하게는 본 발명의 폴리펩티드의 존재를 검출하는 방법을 사용하여, 본 발명의 폴리펩티드가 검출된 암이다.
- [0147] 본 발명의 암 치료용 의약 조성물에 있어서의 유효 성분으로서, RET를 저해하는 화합물이고, 보다 바람직하게는 본 발명의 폴리펩티드의 발현 및/또는 활성, 또는 본 발명의 폴리뉴클레오티드의 발현을 억제하는 화합물이다. 또한, 본 발명의 스크리닝 방법에 의해 선택한 화합물을 유효 성분으로서 사용할 수 있다. 예를 들어, 공지된 RET를 저해하는 화합물을 본 발명의 의약 조성물에 있어서의 유효 성분으로서 사용할 수 있다. RET를 저해하는 화합물로서, RET의 발현 및/또는 활성을 저해할 수 있는 화합물이면, 다른 티로신 키나아제의 발현 및/또는 활성을 저해하는 화합물이어도 되고, 보다 바람직하게는 RET의 활성을 저해할 수 있는 화합물이며, 다른 티로신 키나아제의 발현 및/또는 활성을 저해하는 화합물이어도 된다. 이러한 화합물로서, 예를 들어, 반데타닙, 소라페닙, 수니티닙, 모테사닙, 카보잔티닙, 렌바티닙, 국제공개 제2016/127074호 팸플릿, 국제공개 제2017/043550호 팸플릿, 국제공개 제2017/011776호 팸플릿, 국제공개 제2017/146116호 팸플릿에 기재된 화합물을 들 수 있다.
- [0148] DCTN1 유전자와 RET 유전자와의 융합 유전자 양성 및/또는 DCTN1 단백질과 RET 단백질과의 융합 단백질 양성인 암 치료용 의약 조성물의 유효 성분으로서, RET를 저해하는 화합물이고, 보다 바람직하게는 반데타닙, 카보잔티닙, 렌바티닙, 국제공개 제2017/043550호 팸플릿에 기재된 일반식 (1)로 나타나는 축합 피리미딘 화합물 또는 국제공개 제2017/146116호 팸플릿에 기재된 일반식 (1)로 나타나는 축합 피리미딘 화합물이고, 보다 바람직하게는, 반데타닙, 카보잔티닙, 렌바티닙, 국제공개 제2017/043550호 팸플릿에 기재된 실시예 화합물 1 내지 90 또는 국제공개 제2017/146116호 팸플릿에 기재된 실시예 화합물 1 내지 207이고, 더욱 바람직하게는 반데타닙, 카보잔티닙, 렌바티닙, 4-아미노-1-(tert-부틸)-N-(5-메틸-1H-피라졸-3-일)-1H-피라졸로[3,4-d]피리미딘-3-카르복사미드, 4-아미노-7-(tert-부틸)-N-(5-메틸-1H-피라졸-3-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-7-(1-플루오로-2-메틸프로판-2-일)-N-(5-메틸-1H-피라졸-3-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-N-(5-메틸-1H-피라졸-3-일)-7-(1-메틸시클로프로필)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-7-(2-시클로프로필프로판-2-일)-N-(5-메틸-1H-피라졸-3-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-N-[4-(메톡시메틸)페닐]-7-(1-메틸시클로프로필)-6-(3-모르폴리노프로-1-핀-1-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-N-[4-(메톡시메틸)페닐]-7-(1-메틸시클로프로필)-6-((테트라히드로-2H-피란-4-일)에티닐)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, (R)-4-아미노-N-[4-(메톡시메틸)페닐]-7-(1-메틸시클로프로필)-6-((테트라히드로푸란-2-일)메톡시)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드 또는 4-아미노-N-

[4-(메톡시메틸)페닐]-6-((1-메틸-1H-피라졸-4-일)에티닐)-7-(1-메틸시클로프로필)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드이며, 특히 바람직하게는 4-아미노-1-(tert-부틸)-N-(5-메틸-1H-피라졸-3-일)-1H-피라졸로[3,4-d]피리미딘-3-카르복사미드, 4-아미노-7-(tert-부틸)-N-(5-메틸-1H-피라졸-3-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-7-(1-플루오로-2-메틸프로판-2-일)-N-(5-메틸-1H-피라졸-3-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-N-(5-메틸-1H-피라졸-3-일)-7-(1-메틸시클로프로필)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-7-(2-시클로프로필프로판-2-일)-N-(5-메틸-1H-피라졸-3-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-N-[4-(메톡시메틸)페닐]-7-(1-메틸시클로프로필)-6-(3-모르폴리노프로-1-핀-1-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, 4-아미노-N-[4-(메톡시메틸)페닐]-7-(1-메틸시클로프로필)-6-((테트라히드로-2H-피란-4-일)에티닐)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드, (R)-4-아미노-N-[4-(메톡시메틸)페닐]-7-(1-메틸시클로프로필)-6-((테트라히드로프로판-2-일)메톡시)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드 또는 4-아미노-N-[4-(메톡시메틸)페닐]-6-((1-메틸-1H-피라졸-4-일)에티닐)-7-(1-메틸시클로프로필)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드이다.

[0149] 본 발명에 있어서, 「RET의 발현 및/또는 활성을 저해할 수 있는 화합물」 중에서, 「RET의 발현을 저해할 수 있다」란, 예를 들어 RET의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 세포에 시험 화합물을 접촉시킨 후, 당해 세포에 있어서의 RET의 폴리펩티드 또는 폴리뉴클레오티드의 발현량을 검출했을 때, 시험 화합물을 접촉시키지 않은 세포와 비교하여, 시험 화합물을 접촉시킨 세포에 있어서, RET의 폴리펩티드 또는 폴리뉴클레오티드의 발현량이 저하되었을 때, RET의 발현이 억제되었다고 판단할 수 있다. 이러한 화합물로서는, 상기와 같은 화합물, siRNA, miRNA, 핵산(DNA, RNA) 앵타머 등을 들 수 있다. siRNA로서는, 예를 들어 CACAUGUCAAAAUGUATT(서열번호 74), GGAUUGAAAACAAACUCUATT(서열번호 75), GCUUGUCCCGAGAUGUUUATT(서열번호 76)을 들 수 있으며, 바람직하게는 CACAUGUCAAAAUGUATT(서열번호 74) 또는 GGAUUGAAAACAAACUCUATT(서열번호 75)이다.

[0150] 또한, 「RET의 발현 및/또는 활성을 저해할 수 있는 화합물」 중에서, 「RET의 활성을 저해할 수 있다」란, 예를 들어 티로신의 인산화를 지표로 판단할 수 있다. 티로신의 인산화를 측정하는 방법으로서, 예를 들어, 국제공개 제2017/043550호 팜플릿의 시험예 1에 기재된 방법을 들 수 있다.

[0151] 또한, RET의 폴리펩티드 및/또는 폴리뉴클레오티드를 발현하고 있는 세포를 사용하여, 세포 증식 억제 효과를 지표로 하여, 「RET의 활성을 저해할 수 있다」라고 판단할 수 있다. 세포 증식 억제 효과는, 예를 들어 국제공개 제2017/043550호 팜플릿의 시험예 3 및 시험예 4에 기재된 방법을 들 수 있다.

[0152] 본 발명의 의약 조성물의 대상이 되는 암은, 본 발명의 폴리뉴클레오티드 및/또는 폴리펩티드를 발현하고 있는 한 특별히 제한되지 않지만, 예를 들어 두경부암, 갑상선암, 소화기암(식도암, 위암, 십이지장암, 간암, 담도암(담낭·담관암 등), 췌장암, 소장암, 대장암(결장직장암, 결장암, 직장암 등), 소화관 간질 종양 등), 폐암(비소세포 폐암, 소세포 폐암), 유방암, 난소암, 자궁암(자궁경부암, 자궁체암 등), 신장암, 방광암, 전립선암, 피부암 등을 들 수 있고, 바람직하게는 갑상선암, 또는 폐암(비소세포 폐암, 소세포 폐암)이다. 부연하면, 여기에서 암에는, 원발소 뿐만 아니라, 다른 장기(간장 등)로 전이한 암도 포함한다.

[0153] 본 발명의 폴리펩티드의 발현 및/또는 활성, 또는 본 발명의 폴리뉴클레오티드의 발현을 억제하는 화합물을 유효 성분으로 하는 제제는, 약학적 담체를 배합하여, 다양한 투여 형태에 따른 의약 조성물로서 제조할 수 있다. 당해 형태로서는, 예를 들어 경구제, 주사제, 좌제, 연고제, 첩부제를 들 수 있다. 이들 투여 형태는, 각각 당업자에게 공지 관용인 제제 방법에 의해 제조할 수 있다.

[0154] 약학적 담체로서는, 제제 소재로서 관용인 각종 유기 또는 무기 담체 물질이 사용되고, 고형 제제에 있어서의 부형제, 결합제, 붕괴제, 활택제, 코팅제 등, 액상 제제에 있어서의 용제, 용해 보조제, 현탁화제, 등장화제, pH 조절제·완충제, 무통화제 등으로서 배합된다. 또한, 필요에 따라 방부제, 항산화제, 착색제, 교미·교취제, 안정화제 등의 제제 첨가물을 사용할 수도 있다.

[0155] 경구용 고형 제제를 제조하는 경우에는, 본 발명 화합물에 부형제, 필요에 따라 부형제, 결합제, 붕괴제, 활택제, 착색제, 교미·교취제 등을 첨가한 후, 통상의 방법에 의해 정제, 피복 정제, 과립제, 산제, 캡슐제 등을 제조할 수 있다.

[0156] 경구용 액체 제제를 제조하는 경우에는, 본 발명 화합물에 pH 조절제·완충제, 안정화제, 교미·교취제 등을 첨가해서 통상의 방법에 의해 내복 액제, 시럽제, 엘릭시르제 등을 제조할 수 있다.

[0157] 주사제를 제조하는 경우에는, 본 발명 화합물에 pH 조절제·완충제, 안정화제, 등장화제, 국소 마취제 등을 첨

가하고, 통상의 방법에 의해 피하, 근육내 및 정맥내용 주사제를 제조할 수 있다.

- [0158] 본 발명에는, 본 발명의 폴리펩티드를 검출하기 위한 방법 또는 본 발명의 폴리뉴클레오티드를 검출하기 위한 방법에 있어서, 시료 중 본 발명의 폴리펩티드 또는 본 발명의 폴리뉴클레오티드의 존재를 검출한 경우에, 암이라고 판정하는 방법이 포함된다. 본 발명에 의해 판정할 수 있는 암으로서, 본 발명의 의약 조성물의 대상이 되는 암으로서 열거한 것 등을 들 수 있다. 상기한 바와 같이 본 발명의 폴리펩티드 또는 폴리뉴클레오티드를 사용함으로써, 암의 판정을 할 수 있다. 따라서, 본 발명의 폴리펩티드, 폴리뉴클레오티드는, 암을 검출하기 위한 바이오마커로서 사용할 수도 있다.
- [0159] 본 발명의 폴리펩티드 또는 본 발명의 폴리뉴클레오티드를, RET를 저해하는 화합물을 사용한 화학 요법이 유효한지 여부의 지표로 하는 방법으로서, 본 발명의 검출 방법에 의해 시료 중에서 본 발명의 폴리펩티드를 검출한 경우, 및/또는 본 발명의 검출 방법에 의해 시료 중에서 본 발명의 폴리뉴클레오티드의 존재를 검출한 경우에, RET를 저해하는 화합물을 사용한 화학 요법이 유효하다고 판정하는 방법이 포함된다.
- [0160] 보다 바람직하게는, 본 발명의 폴리펩티드 또는 본 발명의 폴리뉴클레오티드를, 본 발명의 폴리펩티드의 발현 및/또는 활성, 또는 본 발명의 폴리뉴클레오티드의 발현을 억제하는 화합물을 사용한 화학 요법이 유효한지 여부의 지표로 하는 방법으로서, 본 발명의 검출 방법에 의해 시료 중에서 본 발명의 폴리펩티드를 검출한 경우, 및/또는 본 발명의 검출 방법에 의해 시료 중에서 본 발명의 폴리뉴클레오티드의 존재를 검출한 경우에, 본 발명의 폴리펩티드의 발현 및/또는 활성, 또는 본 발명의 폴리뉴클레오티드의 발현을 억제하는 화합물을 사용한 화학 요법이 유효하다고 판정하는 방법이 포함된다.
- [0161] 보다 바람직하게는, 본 발명의 폴리펩티드 또는 본 발명의 폴리뉴클레오티드를, 본 발명의 스크리닝 방법에 의해 얻어진 화합물을 사용한 화학 요법이 유효한지 여부의 지표로 하는 방법으로서, 본 발명의 검출 방법에 의해 시료 중에서 본 발명의 폴리펩티드를 검출한 경우, 및/또는 본 발명의 검출 방법에 의해 시료 중에서 본 발명의 폴리뉴클레오티드의 존재를 검출한 경우에, 본 발명의 스크리닝 방법에 의해 얻어진 화합물을 사용한 화학 요법이 유효하다고 판정하는 방법이 포함된다.
- [0162] 이하, 실시예에 의해 본 발명을 구체적으로 설명하지만, 본 발명은 이들 실시예에 의해 한정되는 것은 아니다.
- [0163] 실시예
- [0164] 실시예 1 DCTN1 유전자와 RET 유전자와의 융합 유전자(DCTN1-RET 융합 유전자)의 취득
- [0165] <1-1 임상 검체 유래 RNA의 추출>
- [0166] Asterand Bioscience에서 구입한 인간 갑상선암 조직으로부터, RNeasy 미니 키트(Qiagen)를 사용해서 이하의 방법에 의해 RNA를 추출했다. 갑상선암 조직에 완충제 RLT를 600  $\mu$ L 첨가하고, QIAshredder 스핀 칼럼에 어플라이한 후 원심하고 (16,000rpm, 2분, 실온), 여액을 회수했다. 회수한 여액에 동량의 70% 에탄올 수용액을 첨가하고 혼화한 후에, RNeasy 미니 칼럼에 어플라이하고 원심하였다(10,000rpm, 15초, 실온). RNeasy 미니 칼럼에 700  $\mu$ L의 완충제 RW1을 첨가하고 원심하였다(10,000rpm, 15초, 실온). 추가로 500  $\mu$ L의 완충제 RPE를 첨가하고 원심하였다(10,000rpm, 15초, 실온). 마찬가지로 다시 500  $\mu$ L의 완충제 RPE를 첨가하고 원심하였다(10,000rpm, 2분, 실온). 다시 RNeasy 미니 칼럼을 원심(16,000rpm, 1분, 실온)하고, 남은 완충제를 제거했다. RNeasy 미니 칼럼에 RNase 무함유 물을 40  $\mu$ L 어플라이한 후, 원심(10,000rpm, 1분, 실온)하고, 여액을 총 RNA(total RNA)로서 회수했다.
- [0167] <1-2 임상 검체 유래 cDNA의 제작>
- [0168] 상기 1-1에서 얻어진 총 RNA로부터, 슈퍼스크립트(SuperScript) VILO cDNA 합성 키트(invitorgen)를 사용하여, 이하의 방법에 의해 cDNA를 합성했다. 500ng의 총 RNA를 용량이 14  $\mu$ L가 되도록 RNase 무함유 물로 제조하고, 4  $\mu$ L의 5 $\times$ VILO 반응 믹스(Reaction Mix) 및 2  $\mu$ L의 10 $\times$ 슈퍼스크립트 효소 믹스(Enzyme Mix)를 첨가하고 혼화했다. 25 $^{\circ}$ C에서 10분간 보온하고 그 후 42 $^{\circ}$ C에서 60분 보온했다. 반응을 정지시키기 위해서 마지막으로 85 $^{\circ}$ C에서 5분간 인큐베이트하여, cDNA를 얻었다.
- [0169] <1-3 클로닝 벡터의 제작 및 정제>
- [0170] DCTN1-RET 융합 유전자를 증폭시키기 위해서, 표 1에 나타낸 바와 같이, 센스 프라이머로서 프라이머 1(서열번호 33)과 안티센스 프라이머로서 프라이머 2(서열번호 34), 추가로 네스티드(nested) PCR에 사용하는 센스 프라이머로서 프라이머 3(서열번호 35)과 안티센스 프라이머로서 프라이머 4(서열번호 36)를 설계했다.

표 1

프라이머1	5'-TGTCAGCTTTGTGCCTGATTGATGT-3'	서열번호 33
프라이머2	5'-GCTGGGCACTGAAGAGAAAGGAATGC-3'	서열번호 34
프라이머3	5'-AGCAGGATGAGTGC GGAGGCAAGC-3'	서열번호 35
프라이머4	5'-TTAACTATCAAACGTGCCATTAATTTTGCCGC-3'	서열번호 36

[0171]

[0172]

이들 프라이머를 사용해서 상기 1-2에서 합성한 cDNA를 주형으로 하여, KOD-Plus-Neo(TOYOBO)를 사용하고, 이하의 방법으로 DCTN1-RET 융합 유전자를 증폭시켰다. 2 μL의 cDNA, 5 μL의 10×KOD-Plus-Neo용 PCR 완충제, 5 μL의 2mM dNTPs, 3 μL의 25mM MgSO<sub>4</sub>, 1 μL의 KOD-Plus-Neo, 1.5 μL의 프라이머 1(10 μM), 1.5 μL의 프라이머 2(10 μM) 및 31 μL의 재증류수(DDW)를 혼합하여, PCR을 행하였다. 이어서 얻어진 PCR 산물을 100배로 희석하고, 2 μL의 희석한 PCR 산물, 5 μL의 10×KOD-Plus-Neo용 PCR 완충제, 5 μL의 2mM dNTPs, 3 μL의 25mM MgSO<sub>4</sub>, 1 μL의 KOD-Plus-Neo, 1.5 μL의 프라이머 3(10 μM), 1.5 μL의 프라이머 4(10 μM) 및 31 μL의 DDW를 혼합하여, 네스티드 PCR을 행하였다.

[0173]

네스티드 PCR 산물을 1% 아가로오스 겔(나카라이)을 사용한 전기 영동에 의해 분리하고, QIAquick 겔 추출 키트(Qiagen)를 사용하여, 겔로부터 PCR 산물을 정제했다.

[0174]

제한 효소 SmaI(NEB)로 절단한 pUC18 DNA(타카라 바이오)와 정제한 PCR 산물과 T4 DNA 리가아제(ligase)(NEB)와 T4 DNA 리가아제 반응 완충제(NEB)를 혼합하고, 하룻밤 16°C에서 인큐베이트했다. 라이게이션 산물을 SmaI(NEB)로 처리하고, 이하의 방법으로 적격 세포로의 형질전환을 행하였다. 50 μL의 대장균 DH5α 적격 세포(타카라 바이오)에 SmaI 처리를 한 라이게이션 산물을 첨가하여 30분 빙상에서 정치했다. 그 후 42°C에서 30초간 히트 쇼크를 가하고, 2분간 빙상에서 정치했다. SOC 배지(타카라 바이오)를 첨가하여, 37°C에서 1시간 진탕 배양한 후, 암피실린 함유 LB 한천 배지 플레이트(UNITECH)에 배양액을 도포하고, 37°C에서 하룻밤 정치했다. 대장균 콜로니를 암피실린 함유 LB 배지(InvivoGen)에 현탁하고 37°C에서 하룻밤 진탕 배양했다. 증식한 대장균으로부터 QIAquick 스핀 미니프랩 키트(Qiagen)를 사용하여, 침부의 프로토콜에 준해서 DCTN1-RET 융합 유전자가 삽입된 플라스미드 DNA를 정제했다.

[0175]

<1-4 서열의 결정>

[0176]

상기 1-3에서 얻어진 플라스미드 DNA를 주형으로 해서, 표 2에 나타낸 시퀀스용 프라이머의 프라이머 5 내지 프라이머 36을 사용하여, BigDye 터미네이터 v3.1 사이클 시퀀싱 키트를 사용하여 시퀀스 반응을 행하고, Applied Biosystems 3730x1 DNA 애널라이저를 사용해서 시퀀스 해석을 실시했다. 시퀀스 해석의 결과, DCTN1-RET 융합 유전자는, DCTN1 변이체 5(GenBank 액세스 번호: NM\_001190836)의 엑손 1 내지 엑손 27의 3'측의 하류에 RET 변이체 2(GenBank 액세스 번호: NM\_020975)의 엑손 12 내지 엑손 20이 융합한 유전자(서열번호 17)였다.

표 2

프라이머 5	5'-AGTACTGGGGTGGCTGGG-3'	서열번호 37
프라이머 6	5'-CACTTTGGACAAGGAGATG-3'	서열번호 38
프라이머 7	5'-ACAGAAGCTGGAGCTGCGG-3'	서열번호 39
프라이머 8	5'-GGACTGGTGTACTCGCTG-3'	서열번호 40
프라이머 9	5'-TCCTAGACTGCAGGAAACAC-3'	서열번호 41
프라이머 10	5'-CATCGAGAAAGTCCAGAC-3'	서열번호 42
프라이머 11	5'-GCTGCTGGAGACATTGAA-3'	서열번호 43
프라이머 12	5'-TCACTGCTGCTCAGCTCA-3'	서열번호 44
프라이머 13	5'-GAGGATCAAAGTGGGAATT-3'	서열번호 45
프라이머 14	5'-AGTATCTGGCCGAGATGAAG-3'	서열번호 46
프라이머 15	5'-GCAAAGACCTGGAGAAGATG-3'	서열번호 47
프라이머 16	5'-AGGACGTTGAAGTCTGACAG-3'	서열번호 48
프라이머 17	5'-CCTTTGCTTCATCCAGAATC-3'	서열번호 49
프라이머 18	5'-GATTTTGTGTTTCTCCAGCTCT-3'	서열번호 50
프라이머 19	5'-CCTGCTTCTCTGAGGAAGAA-3'	서열번호 51
프라이머 20	5'-GGGCTTAGTCTCAGCAAAC-3'	서열번호 52
프라이머 21	5'-GAGCACTCTGCGTGAACCTTA-3'	서열번호 53
프라이머 22	5'-CAGCTTGTTTCATGGTACTGAT-3'	서열번호 54
프라이머 23	5'-TGGTGAGTCCTCACCAG-3'	서열번호 55
프라이머 24	5'-CCTAGAGTTTTTCCAAGAACCA-3'	서열번호 56
프라이머 25	5'-CATTTAACTGGAATCCGACC-3'	서열번호 57
프라이머 26	5'-GACTCTCTCCAGGCCAGTTC-3'	서열번호 58
프라이머 27	5'-GGCTATCAGAAGTAAAACCACC-3'	서열번호 59
프라이머 28	5'-CGAGAGCTGATGGCACTA-3'	서열번호 60
프라이머 29	5'-CTTCATCACAAAGTGAAGTACTTCC-3'	서열번호 61
프라이머 30	5'-CGTACTCCACGATGAGGAG-3'	서열번호 62
프라이머 31	5'-GATTCTGGATGAAGCAAAGG-3'	서열번호 63
프라이머 32	5'-GGAAGTACTTCACTTGTGATGAAG-3'	서열번호 64
프라이머 33	5'-CCCAGCCACCCCAGTACT-3'	서열번호 65
프라이머 34	5'-GTAAAACGACGGCCAGT-3'	서열번호 66
프라이머 35	5'-GTTTTCCAGTCACGAC-3'	서열번호 67
프라이머 36	5'-CAGGAAACAGCTATGAC-3'	서열번호 68

[0177]

[0178]

실시예 2 DCTN1-RET 융합 유전자의 검출

[0179]

Asterand Bioscience에서 구입한 인간 정상 갑상선 조직 유래 RNA와 상기 1-1에서 얻어진 인간 갑상선암 조직 유래 RNA로부터 슈퍼스크립트 VILO cDNA 합성 키트(invitorgen)를 사용하여, 이하의 방법에 의해 cDNA를 합성했다. 280ng의 총 RNA를 용량이 14 μL가 되도록 RNase 무함유 물로 제조하고, 5×VILO 반응 믹스를 4 μL, 10×슈퍼스크립트 효소 믹스를 2 μL 각각 첨가하고 혼합했다. 25℃에서 10분간 보온하고 그 후 42℃에서 60분 보온했다. 반응을 정지시키기 위해서 마지막으로 85℃에서 5분간 인큐베이트했다.

[0180]

DCTN1-RET 융합 유전자의 검출을 위해서, 표 3에 나타난 바와 같이, DCTN-RET 융합 유전자 검출용 센스 프라이머로서 프라이머 37(서열번호 69)과 DCTN-RET 융합 유전자 검출용 안티센스 프라이머로서 프라이머 38(서열번호 70), DCTN1-RET 융합 유전자 검출용 프로브(프로브는 TaqMan MGB 프로브, 형광 색소는 FAM(Thermo Fisher Scientific))으로서 프라이머 39(서열번호 71)를 설계했다.

표 3

프라이머 37	5'-CTGGAGCCACAGTACCCACT-3'	서열번호 69
프라이머 38	5'-TCCAAATTCGCCTTCTCCTA-3'	서열번호 70
프라이머 39	5'-TTCATCAGCCTTCTCAGGGAGGAT-3'	서열번호 71

[0181]

[0182]

얻어진 cDNA를 10배 희석하고 1.1 μL을 주형으로서 사용하고, 11 μL의 프로브용 ddPCR 슈퍼믹스(Bio-Rad), 2 μL의 프라이머 37(10 μM), 2 μL의 프라이머 38(10 μM), 0.6 μL의 프라이머 39(10 μM), 1.1 μL의 GAPDH를 검출하

는 20×HEX 검정(PrimePCR ddPCR 발현 프로브 검정(Expression Probe Assay): GAPDH, 인간, Bio-Rad)를 혼화시켜서, 자동화 액적 제너레이터(Bio-Rad)로 액적을 제작했다. 제작한 액적을 사용해서 PCR을 행하고, 액적 리더(Bio-Rad)로 DCTN1-RET 및 GAPDH 양성인 액적을 카운트했다. 결과를 도 1 및 도 2에 도시한다.

[0183] 또한, 상기에서 합성한 cDNA를 주형으로 하여, KOD-Plus-Neo(TOYOBO)를 사용하여 이하의 방법으로 DCTN1-RET 융합 유전자를 증폭시켰다. 2 μL의 cDNA, 5 μL의 10×KOD-Plus-Neo용 PCR 완충제, 5 μL의 2mM dNTPs, 3 μL의 25mM MgSO<sub>4</sub>, 1 μL의 KOD-Plus-Neo, 1.5 μL의 프라이머 1(10 μM), 1.5 μL의 프라이머 2(10 μM) 및 31 μL의 DDW를 혼화하여, PCR을 행하였다. 이어서 얻어진 PCR 산물을 100배로 희석하고, 2 μL의 희석한 PCR 산물, 5 μL의 10×KOD-Plus-Neo용 PCR 완충제, 5 μL의 2mM dNTPs, 3 μL의 25mM MgSO<sub>4</sub>, 1 μL의 KOD-Plus-Neo, 1.5 μL의 프라이머 3(10 μM), 1.5 μL의 프라이머 4(10 μM) 및 31 μL의 DDW를 혼화하여, 네스티드 PCR을 행하였다. 네스티드 PCR 산물을 1% 아가로오스 겔(나카라이)을 사용한 전기 영동에 의해 분리하고 촬영을 행하였다. 결과를 도 3에 도시한다.

[0184] 도 1, 도 2 및 도 3에 도시한 바와 같이, 인간 갑상선암 조직 유래의 RNA로부터 합성한 cDNA에 있어서는, DCTN1-RET 융합 유전자가 검출된 데 반해, 인간 정상 갑상선 조직 유래 RNA로부터 합성한 cDNA에 있어서는, DCTN1-RET 융합 유전자가 검출되지 않았다. 이상의 결과로부터, DCTN1-RET 융합 유전자는 암 바이오 마커로서 유용한 것이 나타났다.

[0185] 실시예 3 DCTN1-RET 융합 유전자의 발현 벡터의 구축

[0186] 발현 벡터 구축을 위해, 표 4에 나타낸 바와 같이, 센스 프라이머로서 프라이머 40(서열번호 72)과 안티센스 프라이머로서 프라이머 41(서열번호 73)을 설계했다.

표 4

프라이머40	5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGCCACCAGCAGGATGAGTGCGGAGGCAAGCGCCCCG-3'	서열번호72
프라이머41	5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTTTAACTATCAAACTGTCCATTAATTTTGC-3'	서열번호73

[0187]

[0188] 이들 프라이머를 사용해서 상기 1-2에서 합성한 cDNA를 주형으로 하여, Prime STAR Max DNA 폴리머라아제(Polymerase)(TaKaRa)를 사용하여 이하의 방법으로 DCTN1-RET 융합 유전자를 증폭시켰다. 1 μL의 cDNA, 25 μL의 2×Prime STAR Max DNA 폴리머라아제, 1 μL의 프라이머 40(10 μM), 1 μL의 프라이머 41(10 μM) 및 22 μL의 재증류수(DDW)를 혼화하여, PCR을 행하였다. 얻어진 PCR 산물을 1% 아가로오스 겔(나카라이)을 사용한 전기 영동에 의해 분리하고, GFX PCR DNA 및 겔 밴드 정제 키트(GE Healthcare)를 사용하여, 겔로부터 PCR 산물을 정제했다. 이어서 얻어진 정제 PCR 산물을 Gateway BP Clonase II 효소 믹스(ThermoFisher)를 사용해서 이하의 방법으로 Gateway pDONR221 벡터에 삽입하고, 엔트리 벡터를 제작했다. 구체적으로는, 5.0 μL의 정제 PCR 산물, 3.5 μL의 pDONR221(85ng/μL), 4.0 μL의 BP Clonase II 효소 믹스, 7.5 μL의 TE를 혼화하고, 25℃에서 90분간 인큐베이트했다. 인큐베이트 후에 프로테이나아제(Proteinase) K(2mg/mL)를 1 μL 첨가하고, 37℃에서 10분간 인큐베이트함으로써 엔트리 벡터를 얻었다.

[0189] 얻어진 엔트리 벡터를 50 μL의 대장균 DH5α 적격 세포(타카라 바이오)에 첨가하여 30분 빙상에서 정치했다. 그 후 37℃에서 20초간 히트 쇼크를 가하여, 2분간 빙상에서 정치했다. SOC 배지(타카라 바이오)를 첨가하여, 37℃에서 1시간 진탕 배양한 후, 카나마이신 함유 LB 한천 배지 플레이트에 배양액을 도포하고, 37℃에서 하룻밤 정치했다. 대장균 콜로니를 카나마이신 함유 LB 배지에 현탁하고, 37℃에서 하룻밤 진탕 배양했다. 증식한 대장균으로부터 DCTN1-RET 융합 유전자가 삽입된 플라스미드 DNA(엔트리벡터 크론)를 DNA 자동 분리 장치 GENE PREP STAR PI-480(구라보)에서 정제했다.

[0190] 얻어진 플라스미드와 Gateway LR Clonase II 효소 믹스(ThermoFisher)를 사용하여, 이하의 방법으로, pJTI Fast DEST 벡터에 DCTN1-RET 융합 유전자를 삽입하여, 발현 벡터로 했다. 150ng의 엔트리 벡터 크론, 1 μL의 pJTI Fast DEST 벡터(150ng/μL), 2 μL의 LR Clonase II 효소 믹스 및 TE 완충제를 혼화하고, 전량을 10 μL로 하고, 25℃에서 90분간 인큐베이트했다. 인큐베이트 후에 프로테이나아제 K(2mg/mL)를 1 μL 첨가하고, 37℃에서 10분간 인큐베이트함으로써, DCTN1-RET 융합 유전자가 삽입된 pJTI Fast DEST 벡터(DCTN1-RET 융합 유전자

발현 벡터)를 얻었다. 얻어진 DCTN1-RET 융합 유전자 발현 벡터를 50 µL의 대장균 DH5α 적격 세포(타카라 바이오)에 첨가하여 30분 빙상에서 정치했다. 그 후 37°C에서 20초간 히트 쇼크를 가하여, 2분간 빙상에서 정치했다. SOC 배지(타카라 바이오)를 첨가하여, 37°C에서 1시간 진탕 배양한 후, 압피실린 함유 LB 한천 배지 플레이트에 배양액을 도포하고, 37°C에서 하룻밤 정치했다. 대장균 콜로니를 압피실린 함유 LB 배지에 현탁하고 37°C에서 하룻밤 진탕 배양했다. 증식한 대장균으로부터 DCTN1-RET 융합 유전자가 삽입된 플라스미드 DNA(DCTN1-RET 융합 유전자 발현 벡터)를 플라스미드 플러스 맥시 키트(QIAGEN)를 사용하여 정제했다.

[0191] 실시예 4 DCTN1-RET 융합 유전자 발현 세포의 수립

[0192] <4-1 세포의 수립>

[0193] DCTN1-RET 융합 유전자 발현 세포의 수립을 위한 숙주 세포로서, 마우스 태아 섬유아 세포 NIH/3T3 세포(American Type Culture Collection)를 선택하고, 상기에서 제작한 DCTN1-RET 융합 유전자가 삽입된 발현 벡터를 형질감염함으로써, DCTN1-RET 융합 유전자 발현 세포를 수립했다. 상세한 방법은 이하에 나타내는 바와 같이 실시했다. NIH/3T3 세포는, 통상의 배양(2차원 배양)을 위해서, D-MEM(고글루코오스)(L-글루타민, 페놀레드, 피루브산나트륨, 1500mg/L 탄산수소나트륨 함유)(WAKO)에 10%가 되도록 초생 송아지 혈청(Newborn Calf Serum; NBCS)(GIBCO)을 첨가한 것을 2차원 배양용 배지로서 사용하고, 37°C, 5% CO<sub>2</sub>로 배양한 것을 사용했다. 형질감염을 행하기 전날에, NIH/3T3 세포를 1.5×10<sup>5</sup> cells/2mL에서, 6웰 플레이트(IWAKI)에 파종하고, 37°C, 5% CO<sub>2</sub>로, 하룻밤 인큐베이트했다. 1.5 µg의 DCTN1-RET 융합 유전자 발현 벡터, 1.5 µg의 pJTI phiC31 인테그라아제 벡터를 혼합한 혼합액에, 그 혼합액의 6배량의 ViaFect 형질감염 시약을 첨가한 후, 전량이 300 µL가 되도록 Opti-MEM을 첨가하고, 실온에서 5분간 인큐베이트함으로써 형질감염 용액을 제작했다. NIH/3T3 세포가 파종되어 있는 웰로부터 배지를 300 µL 제거하고, 상기에서 제작한 형질감염 용액을 300 µL 첨가하고, 37°C, 5% CO<sub>2</sub>로, 하룻밤 인큐베이트했다. 다음날, 형질감염 용액을 제거하기 위해서, 배지를 교환했다. 배지를 교환할 때, 새로운 배지에는, 하이그로마이신 B(나카라이테스크)를 500 µg/mL가 되도록 첨가했다. 하이그로마이신 B에 의해, DCTN1-RET 융합 유전자 도입 발현 벡터가 도입되어 있지 않은 세포를 제거했다.

[0194] 형질감염 후, 1주일에 2회 정도의 배지 교환을 하면서, 세포가 증식할 때까지, 배양을 행하였다. 형질감염 22일 후에, 세포를 트립신으로 회수하고, 이하의 방법으로 단세포 클로닝을 행하였다. 회수한 세포의 세포수를 측정하여, 1cell/200 µL가 되도록 배지를 첨가했다. 96웰 플레이트(ThermoFisher)에 200 µL/1웰이 되도록 세포를 파종했다. 파종 후, 매일 관찰을 행하여, 단세포로부터 증식해 온 세포를 취득하고, DCTN1-RET 융합 유전자 발현 세포(DCTN1-RET 융합 유전자 발현 NIH/3T3 세포)로 했다.

[0195] <4-2 목적 단백질의 발현 확인>

[0196] 얻어진 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에 있어서의 DCTN1-RET 융합 단백질의 발현을, 웨스턴 블로팅법으로 확인했다. 즉, 배양 플라스크로부터 배지를 제거하고, PBS로 한번 세정했다. 배양 플라스크 내에 포스파타아제 저해제(phosphatase inhibitor)(로슈)와 프로테아제 저해제(protease inhibitor)(로슈)를 함유한 샘플 희석 농축물(Sample Diluent Concentrate) 2(R&D SYSTEMS)를 첨가하고, 스크레이퍼로 세포 용해액을 회수했다. 회수한 세포 용해액으로부터 원심 분리에 의해, 단백질 샘플을 얻었다. 단백질 샘플은, 단백질 정량을 행하여, 단백질 농도를 일정하게 했다. 일정 농도의 단백질 샘플에 대하여, SDS-PAGE용 환원 시약 함유 샘플 완충 용액(6x)(나카라이테스크)을 첨가하여, 95°C에서 5분간 인큐베이트함으로써, 단백질을 변성시켜서, 웨스턴 블로팅용 샘플을 얻었다. 또한, 네거티브 컨트롤용 샘플로서, 친주인 NIH/3T3 세포를 사용하여, 마찬가지로 방법으로 웨스턴 블로팅용 샘플을 얻었다. 상기 샘플을 사용하여, 이하에 나타내는 방법으로, 단백질의 발현을 확인했다. 4 내지 15퍼센트 아크릴아미드 겔(BIO-RAD) 및 1×Tris/글리신/SDS 완충제를 사용하여, SDS-PAGE 전기영동(200V에서 30분)에서, 단백질을 분리했다. Trans-Blot Turbo RTA 미디 PVDF 트랜스퍼 키트(BIO-RAD)와 Transblot Turbo 전사 시스템(BIO-RAD)을 사용해서 PVDF막에 단백질을 전사하고, PVDF막을 블로킹 원-피(Blocking One-P)에 1시간 침지했다. 블로킹 원-피가 10%가 되도록 TBS-T로 희석한 용액으로 1차 항체(포스포-Ret(Tyr905) 항체(CST), Ret(C31B4) 토끼 mAb(CST) 및 항-Dctn1 항체(ATLAS ANTIBODIES))를 1/1000 농도가 되도록 희석하고, PVDF막을 침지시켜서, 4°C에서 하룻밤 인큐베이트했다. TBS-T로 세정 후, 항-토끼 IgG, HRP-연결 항체(CST)를 1/2000 농도가 되도록 TBS-T로 희석한 2차 항체 희석액으로, PVDF막을 침지시켜서, 실온에서 1시간 인큐베이트했다. TBS-T로 세정 후, SuperSignal 웨스트 듀라 연장 기간 기질(West Dura Extended Duration Substrate)(ThermoFisher) 및 루미노·이미지 애널라이저 Amersham Imager 600(GE 헬스케어)을 사용해서 단백질의 검출을 행하였다. 또한, 검출 단백질의 분자량은, 프리시전 플러스 프로테인 칼레이도스코프 스

탠다드(BIORAD)에 의해 확인했다.

- [0197] 그 결과, 도 4의 a) 및 b)에 나타난 바와 같이, 항pRET 항체와 항RET 항체를 사용한바 내재성의 RET(150 및 175kDa)는 검출되지 않았다. 한편, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에 있어서만, 175kDa 부근에 DCTN1-RET 융합 단백질이라고 추측되는 밴드를 확인할 수 있었다.
- [0198] 또한, 도 4의 c)에 나타난 바와 같이 항DCTN1 항체를 사용한바, 150kDa 부근에 내재성의 DCTN1이 검출되었다. 한편, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에 있어서만, 내재성의 150kDa의 DCTN1의 밴드 상의 175kDa 부근에 밴드가 검출되었다. 즉, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에 있어서만, RET에 대한 항체 및 DCTN1에 대한 항체의 양쪽에 175kDa 부근의 밴드가 검출된 점에서, 제작한 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에서는, DCTN1과 RET가 융합한 단백질이 발현하고 있는 것이 명확해졌다.
- [0199] 실시에 5 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 3차원 배양에 의한 증식 확인
- [0200] NIH/3T3 세포는, 2차원 배양 조건에서는 양호한 증식을 나타내지만, 3차원 배양 조건에서는, 대부분 증식하지 않는다. 그 반면, NIH/3T3 세포에 암 유전자를 발현시킴으로써, 3차원 배양 조건에서도 증식하는 것이 알려져 있다. 그 성질을 이용하여, DCTN1-RET 융합 유전자가 암 유전자인지의 확인을 행하였다. 37℃, 5% CO<sub>2</sub>로, 2차원 배양으로 배양한 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포 및 NIH/3T3 세포를 트립신으로 회수하고, 세포수를 측정했다. 3차원 배양을 행하기 위해서, FcEM 시리즈 제조 키트(닛산 가가쿠 고교 가부시기가이샤)와 D-MEM(고글루코오스)(L-글루타민, 페놀레드, 피루브산나트륨, 1500mg/L 탄산수소나트륨 함유)(WAKO)과 초생 송아지 혈청(NBCS)(GIBCO)을 사용하여, 3차원 배양용 배지를 제작했다. 제작한 3차원 배양용 배지에, 세포를 1000cells/90μL가 되도록 현탁하고, 96웰 투명 검정 둥근 바닥의 타원 마이크로플레이트(Corning)에 90μL/1well이 되도록 파종하고, 37℃, 5% CO<sub>2</sub>로 인큐베이트했다. 파종 다음날(Day1) 및 파종 8일 후(Day8)에 세포 내 ATP 발광 검출 시약인 Celltiter-Glo 2.0 시약(Progema) 및 루미노미터(EnSpire, PerkinElmer)를 사용해서 발광량(counts per second: cps)을 측정하여, 생세포수의 지표로 했다. Day1에서의 측정 결과와 Day8에서의 측정 결과로부터 각 세포의 증식율을 산출했다(N=3).
- [0201] 그 결과, 도 5에 도시한 바와 같이, NIH/3T3 세포로는, Day8의 세포수는 Day1의 세포수와 비교해서 2.4배였던 데 비해, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에서는, 20.9배였다. 또한, NIH/3T3 세포는, 3차원 배양에 있어서 세포의 응집 덩어리는 형성되지 않지만, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에서는, 3차원 배양에 의해 세포의 응집 덩어리가 형성되어 있는 것이 확인되었다.
- [0202] 즉, DCTN1-RET 융합 유전자를 도입함으로써, 세포의 증식이 항진한 것이 명확해지고, DCTN1-RET 융합 유전자가 암 유전자인 것이 시사되었다.
- [0203] 실시에 6 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 생체 내에 있어서의 종양 형성성 확인
- [0204] DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 생체 내에 있어서의 종양 형성성을 확인하기 위해서, 누드마우스를 사용한 이식 실험을 행하였다. 부연하면, 친주인 NIH/3T3 세포는, 누드마우스의 피하에서는 증식하지 않는 것이, 일반적으로 알려져 있고, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포를 누드마우스의 피하에 이식함으로써, DCTN1-RET 융합 유전자가 종양 형성성에 기여하는지, 즉 암 유전자인지의 확인을 할 수 있다. 피이식 동물로서는, 누드마우스(BALB/cAJcl-nu/nu, 니혼 클레아)를 사용했다. DCTN1-RET 융합 유전자 발현 NIH/3T3 세포를 트립신으로 회수하고, 최종적으로 1×10<sup>8</sup> cells/mL가 되도록 PBS에 현탁하고, 동량의 매트릭셀 기저막 매트릭스(Corning)를 첨가하고, 5×10<sup>7</sup> cells/mL로 한 것을 이식용 세포액으로 했다. 25G 주사 바늘과 1mL 시린지를 사용하여, 이식용 세포액을 누드마우스(N=10)의 우측 흉부의 피하에 0.1mL씩 이식했다. 전자 노기스(미츠토요)를 사용하여, 이식 후, 10, 13, 17일째에, 1마리씩 종양의 긴 직경 및 짧은 직경을 측정하고, 이하의 식을 사용해서 종양 체적을 산출했다.
- [0205] 종양 체적(mm<sup>3</sup>)=(긴 직경, mm)×(짧은 직경, mm)×(짧은 직경, mm)/2
- [0206] 종양 체적의 측정 결과를 도 6에 나타낸다. 그 결과, 누드마우스의 피하에 이식된 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포는, 종양을 형성하고, 양호하게 증식하는 것이 확인되고, 생체 내 실험에 있어서도 DCTN1-RET 융합 유전자가 암 유전자인 것이 시사되었다.
- [0207] 실시에 7 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포를 사용한 siRNA에 의한 DCTN1-RET 융합 단백질의 억제와 세포 증식 억제 효과의 확인

[0208] DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에 대한 siRNA 처리에 의한 영향을 확인했다. 사용한 siRNA는, 하기의 표 5에 나타낸 3종류의 RET siRNA와 네거티브 컨트롤로서 Silencer 셀렉트 네거티브 컨트롤 #1 siRNA(Ambion)를 사용했다. 또한, 3종류의 RET siRNA는, 모두 인간 RET를 표적으로 하는 siRNA이지만, RET siRNA1 및 RET siRNA2는 DCTN1-RET 융합 유전자 내의 RET 부분에 결합하는 서열을 포함하고, RET siRNA3은 DCTN1-RET 융합 유전자 내에 결합하는 서열을 포함하지 않는다. 즉, RET siRNA1 및 RET siRNA2는 DCTN1-RET 융합 유전자의 발현을 억제하지만, RET siRNA3은 DCTN1-RET 융합 유전자의 발현을 억제하지 않는 것이 상정되었다. 이하에 siRNA를 사용한 실험의 방법에 대해서 기재했다.

표 5

RET siRNA1	CACAUGUCAUCAAAUUGUATT	서열번호 74
RET siRNA2	GGAUUGAAAACAAACUCUATT	서열번호 75
RET siRNA3	CCACUGCUACCACAAGUUUTT	서열번호 77

[0209]

[0210] DCTN1-RET 융합 유전자 발현 NIH/3T3 세포는, 2차원 배양용 배지를 사용하고, 37℃, 5% CO<sub>2</sub>에서 배양한 것을 사용했다. siRNA 처리를 행하기 전날에, 각 세포를 3×10<sup>5</sup> cells/2mL로, 6웰 플레이트(IWAKI)에 파종하고, 37℃, 5% CO<sub>2</sub>로, 하룻밤 인큐베이트했다. 사전에 물을 사용해서 20 μM으로 제조한 각 siRNA를 12 μL, 4 μL의 리포펙타민(Lipofectamin) RNAiMAX 형질감염 시약(ThermoFisher) 및 384 μL의 Opti-MEM을 혼합하고, 실온에서 15분 인큐베이트함으로써 siRNA 용액을 제작했다. 각 세포가 파종되어 있는 웰에 siRNA 용액을 400 μL 첨가하고, 37℃, 5% CO<sub>2</sub>로, 하룻밤 인큐베이트했다.

[0211] 다음날, 일부는, 단백질 발현 해석용으로 샘플링을 행하고, 일부에 대해서는, 세포 증식 확인용으로 재파종을 행하였다. 단백질 발현 해석용 샘플링 및 단백질 발현 해석은, 1차 항체로서, 포스포-Ret(Tyr905) 항체(CST), Ret(C31B4) 토끼 mAb(CST) 및 GAPDH(D16H11) XP 토끼 mAb(CST)를 사용한 것 이외에는, 상기 <4-2 목적 단백질의 발현 확인>과 마찬가지로 실시했다. 그 결과, 도 7에 나타낸 바와 같이, siRNA 처리하지 않은 세포(무처리)에 비해, 네거티브 컨트롤 siRNA(NC)를 처리한 세포에 있어서, DCTN1-RET 융합 단백질의 발현은 억제되어 있지 않은 것을 확인할 수 있었다. 한편, RET siRNA1 및 RET siRNA2를 처리한 경우, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 DCTN1-RET 융합 단백질의 발현이 억제되는 것이 확인되고, RET siRNA3에서는 억제되지 않는 것이 명확해졌다.

[0212] 다음에 세포 증식 억제 효과의 확인을 위해, 무처리 또는 siRNA 처리된 웰로부터 트립신에 의해 세포를 회수하고, 세포수를 측정했다. 상기 실시예 5와 마찬가지로 방법으로 3차원 배양을 행하고, 파종 당일(Day0) 및 파종 4일 후(Day4)에 실시예 5와 마찬가지로 방법으로 생세포수를 측정했다. Day0에서의 측정 결과와 Day4에서의 측정 결과로부터 각 세포에 있어서의 증식율을 산출했다.

[0213] 그 결과, 도 8에 나타낸 바와 같이, siRNA를 처리하지 않은 세포(무처리) 및 네거티브 컨트롤 siRNA(NC)를 처리한 세포에서는, Day4의 세포수는, Day0의 세포수와 비교하여, 4.9배 및 3.6배였던 데 비해, RET siRNA1 및 RET siRNA2로 처리한 세포에서는, 2.0배 및 2.4배 정도의 증식으로, 현저하게 증식율이 저하되었다. 한편, RET siRNA3로 처리한 세포에서는, 3.7배의 증식으로, 네거티브 컨트롤 siRNA와 동일 정도의 증식율이며, 증식율의 저하는 보이지 않았다. 이들 결과로부터, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 증식은, siRNA에 의해 RET의 발현을 저해한 경우에도 억제되는 것이 명확해졌다.

[0214] 실시예 8 DCTN1-RET 융합 유전자 발현 NIH/3T3 세포를 사용한 세포 증식 억제 효과

[0215] DCTN1-RET 융합 유전자 발현 NIH/3T3 세포에 대한 시험관 내(in vitro) 세포 증식 시험을 행하였다. 상기 실시예 5와 마찬가지로 방법으로, 3차원 배양 및 파종을 행하였다. 파종 후 37℃, 5% CO<sub>2</sub>로 하룻밤 인큐베이트하였다(Day0). RET를 저해한다고 보고되어 있는 카보잔티닙, 반데타닙, 알렉티닙, 렌바티닙, 촉합 피리미딘 화합물(화합물 1 내지 9; 표 6에 나타낸다)을 디메틸술폰으로 10mmol/L의 농도로 용해하고, 추가로 3차원 배양용 배지를 사용하여, 이들 화합물의 최종 농도가 각각 1000, 333, 111, 37.0, 12.3, 4.12, 1.37, 0.457nmol/L가 되도록 희석을 행하였다. 이것을 먼저 설명한 세포가 파종된 플레이트의 각 웰에 0.01mL씩 첨가하여(Day1), 37℃, 5% CO<sub>2</sub>로 7일간 인큐베이트했다. 배양 후(Day8), 모든 웰에 세포 내 ATP 발광 검출 시약인 Celltiter-

Glo 2.0 시약(Progema)를 첨가하고, 루미노미터(EnSpire, PerkinElmer)를 사용해서 발광량(counts per second: cps)을 측정했다.  $T_{day8}$ 과  $C_{day1}$ 의 값의 크기에 따라, 이하의 식으로부터 화합물의 각 농도에 있어서의 Day1로부터의 증식율을 산출하고, 세포 증식을 50% 억제하는 피검 화합물의 농도( $GI_{50}(nM)$ )를 구했다.

[0216]

1)  $T_{day8} \geq C_{day1}$ 인 경우

[0217]

$$\text{증식율}(\%) = (T_{day8} - C_{day1}) / (C_{day8} - C_{day1}) \times 100$$

[0218]

T: 피검 화합물을 첨가한 웰의 cps

[0219]

C: 피검 화합물을 첨가하지 않은 웰의 cps

[0220]

Day1: 피검 화합물을 첨가한 날

[0221]

Day8: 평가일

[0222]

2)  $T_{day8} < C_{day1}$ 인 경우

[0223]

$$\text{증식율}(\%) = (T_{day8} - C_{day1}) / (C_{day1}) \times 100$$

[0224]

T: 피검 화합물을 첨가한 웰의 cps

[0225]

C: 피검 화합물을 첨가하지 않은 웰의 cps

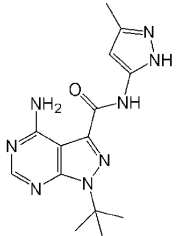
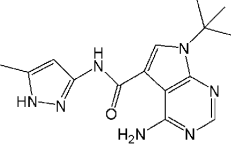
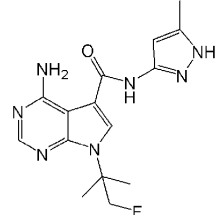
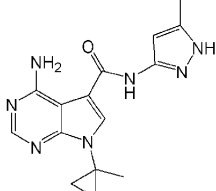
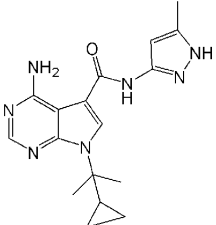
[0226]

Day1: 피검 화합물을 첨가한 날

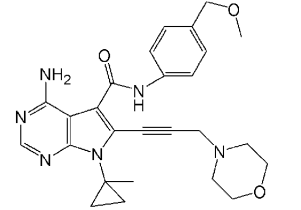
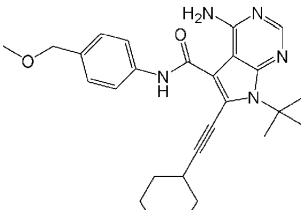
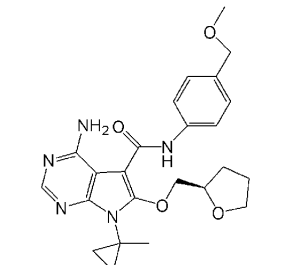
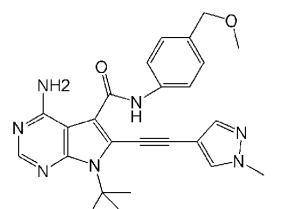
[0227]

Day8: 평가일

표 6

화합물 번호	구조식	화학명	국제공개번호 및 실시예 번호
화합물 1		4-아미노-1-(tert-부틸)-N-(5-메틸-1H-피라졸-3-일)-1H-피라졸로[3,4-d]피리미딘-3-카르복사미드	국제공개 제2017/043550호 실시예 화합물 34
화합물 2		4-아미노-7-(tert-부틸)-N-(5-메틸-1H-피라졸-3-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드	국제공개 제2017/043550호 실시예 화합물 48
화합물 3		4-아미노-7-(1-플루오로-2-메틸프로판-2-일)-N-(5-메틸-1H-피라졸-3-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드	국제공개 제2017/043550호 실시예 화합물 50
화합물 4		4-아미노-N-(5-메틸-1H-피라졸-3-일)-7-(1-메틸시클로프로필)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드	국제공개 제2017/043550호 실시예 화합물 51
화합물 5		4-아미노-7-(2-시클로프로필프로판-2-일)-N-(5-메틸-1H-피라졸-3-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드	국제공개 제2017/043550호 실시예 화합물 52

[0228]

<p>화합물 6</p>		<p>4-아미노-N-[4-(메톡시메틸)페닐]-7-(1-메틸시클로프로필)-6-(3-모르폴리노프로-1-핀-1-일)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드</p>	<p>국제공개 제2017/146116호 실시예 화합물 85</p>
<p>화합물 7</p>		<p>4-아미노-N-[4-(메톡시메틸)페닐]-7-(1-메틸시클로프로필)-6-((테트라히드로-2H-피란-4-일)에틸)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드</p>	<p>국제공개 제2017/146116호 실시예 화합물 87</p>
<p>화합물 8</p>		<p>(R)-4-아미노-N-[4-(메톡시메틸)페닐]-7-(1-메틸시클로프로필)-6-((테트라히드로푸란-2-일)메톡시)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드</p>	<p>국제공개 제2017/146116호 실시예 화합물 89</p>
<p>화합물 9</p>		<p>4-아미노-N-[4-(메톡시메틸)페닐]-6-((1-메틸-1H-피라졸-4-일)에틸)-7-(1-메틸시클로프로필)-7H-피롤로[2,3-d]피리미딘-5-카르복사미드</p>	<p>국제공개 제2017/146116호 실시예 화합물 90</p>

[0229]

[0230]

그 결과, 표 7에 나타낸 바와 같이, 카보잔티닙, 반데타닙, 렌바티닙 및 축합 피리미딘 화합물(화합물 1 내지 9)에 있어서, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 증식이 억제되었다.

표 7

화합물	GI <sub>50</sub> (nmol/L)
화합물 1	8.50
화합물 2	5.80
화합물 3	5.40
화합물 4	5.70
화합물 5	4.20
화합물 6	3.60
화합물 7	11.40
화합물 8	17.90
화합물 9	4.80
카보잔티닙	91.10
반데타닙	68.70
렌바티닙	50.30
알렉티닙	>1000

[0231]

- [0232] 이상의 결과로부터, 상기 RET 저해제는, DCTN1-RET 융합 유전자가 검출된 암에 대한 치료약으로서 유용할 가능성이 시사되었다. 또한, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포를 사용함으로써, DCTN1-RET를 억제하는 화합물의 스크리닝이 가능한 것이 시사되었다.
- [0233] 실시예 9 DCTN1-RET 융합 유전자 발현 세포를 사용한 RET의 인산화 저해
- [0234] DCTN1-RET 융합 유전자 발현 세포에 있어서의 RET의 인산화가, RET를 저해한다고 보고되어 있는 기존의 약제로 저해되는지에 대해서 이하의 방법으로 검토했다.
- [0235] DCTN1-RET 융합 유전자 발현 NIH/3T3 세포는, 2차원 배양용 배지를 사용하고, 37℃, 5% CO<sub>2</sub>로 배양한 것을 사용했다. 약제 처리를 행하기 전날에, 각 세포를 3×10<sup>5</sup> cells/2mL로, 6웰 플레이트(IWAKI)에 파종하고, 37℃, 5% CO<sub>2</sub>로, 하룻밤 인큐베이트했다. 카보잔티닙, 반테타닙, 알렉티닙, 렌바티닙을 디메틸설폭시드로 10mmol/L의 농도로 용해하고, 추가로 PBS를 사용하여, 이들 화합물의 최종 농도가 각각 1000, 100, 10nmol/L가 되도록 희석을 행하였다. 이것을 먼저 설명한 세포가 파종된 플레이트의 각 웰에 20 μL씩 첨가하여(Day1), 37℃, 5% CO<sub>2</sub>로 1시간 인큐베이트했다. 인큐베이트 후, 상기 실시예 7에 기재되어 있는 방법과 마찬가지로 단백질 발현 해석용 샘플링을 행하여, 단백질 발현 해석을 실시했다.
- [0236] 그 결과, 도 9에 나타난 바와 같이, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 인산화 RET 레벨이 카보잔티닙 및 렌바티닙에 의해 현저하게 감소하는 것이 확인되었다. 또한, 상기와 마찬가지로 방법으로, 축합 피리미딘 화합물을 사용하여, RET의 인산화 저해를 평가한 결과, 축합 피리미딘 화합물에 있어서도 RET의 인산화가 현저하게 감소하는 것을 확인할 수 있었다.
- [0237] 이상의 결과로부터, 인산화 RET 레벨을 현저하게 감소시키는 약제는, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 증식을 억제할 수 있는 화합물이며, DCTN1-RET 융합 유전자가 검출된 암에 대한 치료약으로서 유용할 가능성이 시사되었다. 또한, DCTN1-RET 융합 유전자 발현 NIH/3T3 세포의 인산화 RET 레벨을 사용함으로써 RET 저해제의 스크리닝이 가능한 것이 시사되었다.

**서열목록 자유텍스트**

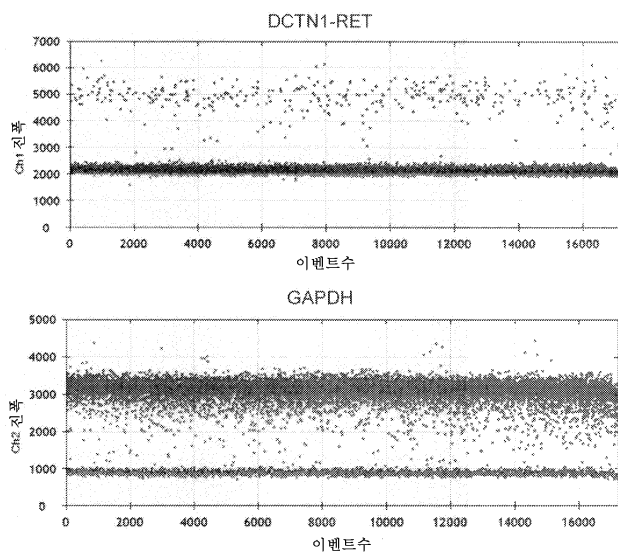
- [0239] 서열번호 1은 DCTN1 변이체 1(v1)[서열번호 25의 일부]과 RET 변이체 2(v2)[서열번호 31의 일부]와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.  
서열번호 2는 DCTN1 v1과 RET v2와의 융합 펩티드의 아미노산 서열을 나타낸다.  
서열번호 3은 DCTN1 v1과 RET 변이체 4(v4)[서열번호 32의 일부]와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.  
서열번호 4는 DCTN1 v1과 RET v4와의 융합 펩티드의 아미노산 서열을 나타낸다.  
서열번호 5는 DCTN1 변이체 2(v2)[서열번호 26의 일부]와 RET v2와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.  
서열번호 6은 DCTN1 v2와 RET v2와의 융합 펩티드의 아미노산 서열을 나타낸다.  
서열번호 7은 DCTN1 v2와 RET v4와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.  
서열번호 8은 DCTN1 v2와 RET v4와의 융합 펩티드의 아미노산 서열을 나타낸다.  
서열번호 9는 DCTN1 변이체 3(v3)[서열번호 27의 일부]과 RET v2와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.  
서열번호 10은 DCTN1 v3과 RET v2와의 융합 펩티드의 아미노산 서열을 나타낸다.  
서열번호 11은 DCTN1 v3과 RET v4와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.  
서열번호 12는 DCTN1 v3과 RET v4와의 융합 펩티드의 아미노산 서열을 나타낸다.  
서열번호 13은 DCTN1 변이체 4(v4)[서열번호 28의 일부]와 RET v2와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.

- 서열번호 14는 DCTN1 v4와 RET v2와의 융합 펩티드의 아미노산 서열을 나타낸다.
- 서열번호 15는 DCTN1 v4와 RET v4와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.
- 서열번호 16은 DCTN1 v4와 RET v4와의 융합 펩티드의 아미노산 서열을 나타낸다.
- 서열번호 17은 DCTN1 변이체 5(v5)[서열번호 29의 일부]와 RET v2와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.
- 서열번호 18은 DCTN1 v5와 RET v2와의 융합 펩티드의 아미노산 서열을 나타낸다.
- 서열번호 19는 DCTN1 v5와 RET v4와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.
- 서열번호 20은 DCTN1 v5와 RET v4와의 융합 펩티드의 아미노산 서열을 나타낸다.
- 서열번호 21은 DCTN1 v6과 RET v2와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.
- 서열번호 22는 DCTN1 v6과 RET v2와의 융합 펩티드의 아미노산 서열을 나타낸다.
- 서열번호 23은 DCTN1 v6과 RET v4와의 융합 펩티드를 코딩하는 폴리뉴클레오티드의 염기 서열을 나타낸다.
- 서열번호 24는 DCTN1 v6과 RET v4와의 융합 펩티드의 아미노산 서열을 나타낸다.
- 서열번호 33은 프라이머의 염기 서열을 나타낸다.
- 서열번호 34는 프라이머의 염기 서열을 나타낸다.
- 서열번호 35는 프라이머의 염기 서열을 나타낸다.
- 서열번호 36은 프라이머의 염기 서열을 나타낸다.
- 서열번호 37은 프라이머의 염기 서열을 나타낸다.
- 서열번호 38은 프라이머의 염기 서열을 나타낸다.
- 서열번호 39는 프라이머의 염기 서열을 나타낸다.
- 서열번호 40은 프라이머의 염기 서열을 나타낸다.
- 서열번호 41은 프라이머의 염기 서열을 나타낸다.
- 서열번호 42는 프라이머의 염기 서열을 나타낸다.
- 서열번호 43은 프라이머의 염기 서열을 나타낸다.
- 서열번호 44는 프라이머의 염기 서열을 나타낸다.
- 서열번호 45는 프라이머의 염기 서열을 나타낸다.
- 서열번호 46은 프라이머의 염기 서열을 나타낸다.
- 서열번호 47은 프라이머의 염기 서열을 나타낸다.
- 서열번호 48은 프라이머의 염기 서열을 나타낸다.
- 서열번호 49는 프라이머의 염기 서열을 나타낸다.
- 서열번호 50은 프라이머의 염기 서열을 나타낸다.
- 서열번호 51은 프라이머의 염기 서열을 나타낸다.
- 서열번호 52는 프라이머의 염기 서열을 나타낸다.
- 서열번호 53은 프라이머의 염기 서열을 나타낸다.
- 서열번호 54는 프라이머의 염기 서열을 나타낸다.
- 서열번호 55는 프라이머의 염기 서열을 나타낸다.
- 서열번호 56은 프라이머의 염기 서열을 나타낸다.

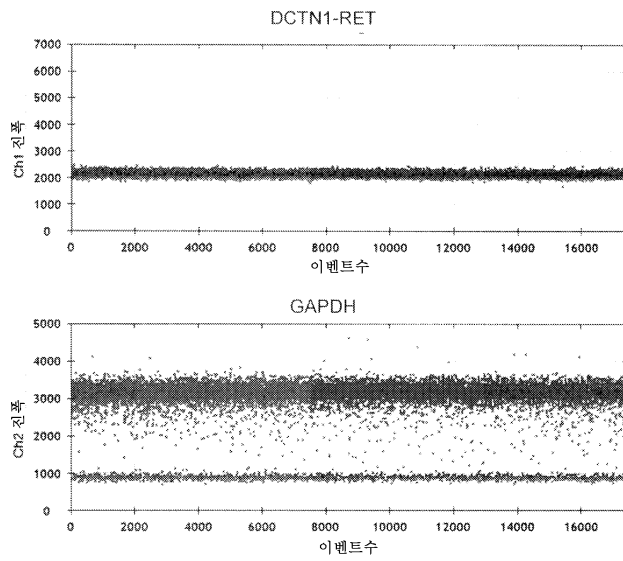
서열번호 57은 프라이머의 염기 서열을 나타낸다.  
 서열번호 58은 프라이머의 염기 서열을 나타낸다.  
 서열번호 59는 프라이머의 염기 서열을 나타낸다.  
 서열번호 60은 프라이머의 염기 서열을 나타낸다.  
 서열번호 61은 프라이머의 염기 서열을 나타낸다.  
 서열번호 62는 프라이머의 염기 서열을 나타낸다.  
 서열번호 63은 프라이머의 염기 서열을 나타낸다.  
 서열번호 64는 프라이머의 염기 서열을 나타낸다.  
 서열번호 65는 프라이머의 염기 서열을 나타낸다.  
 서열번호 66은 프라이머의 염기 서열을 나타낸다.  
 서열번호 67은 프라이머의 염기 서열을 나타낸다.  
 서열번호 68은 프라이머의 염기 서열을 나타낸다.  
 서열번호 69는 프라이머의 염기 서열을 나타낸다.  
 서열번호 70은 프라이머의 염기 서열을 나타낸다.  
 서열번호 71은 프라이머의 염기 서열을 나타낸다.  
 서열번호 72는 프라이머의 염기 서열을 나타낸다.  
 서열번호 73은 프라이머의 염기 서열을 나타낸다.  
 서열번호 74는 RET siRNA의 염기 서열을 나타낸다.  
 서열번호 75는 RET siRNA의 염기 서열을 나타낸다.  
 서열번호 76은 RET siRNA의 염기 서열을 나타낸다.  
 서열번호 77은 RET siRNA의 염기 서열을 나타낸다.

**도면**

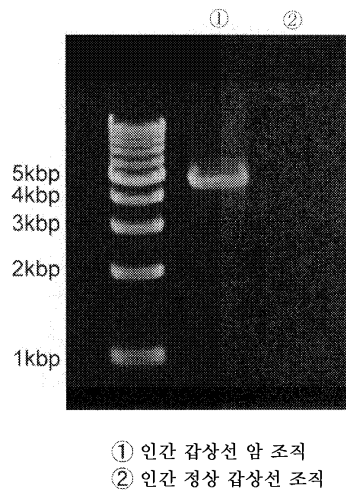
**도면1**



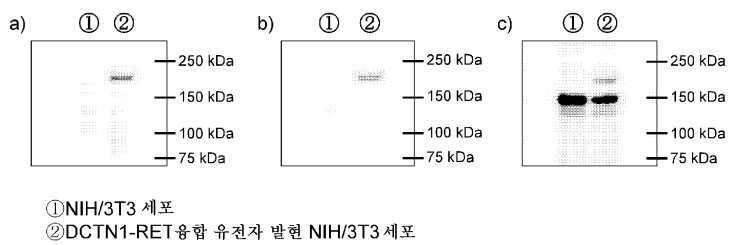
도면2



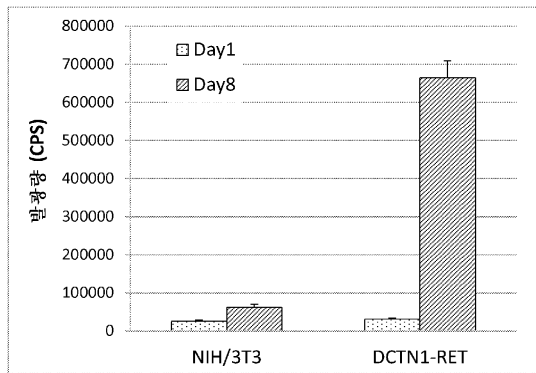
도면3



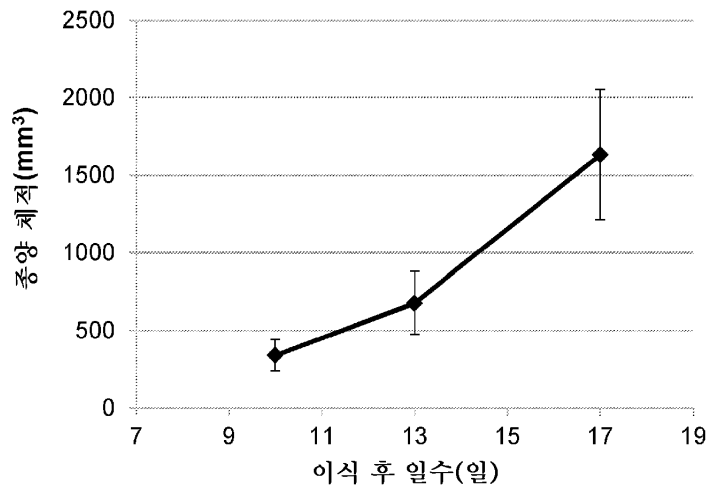
도면4



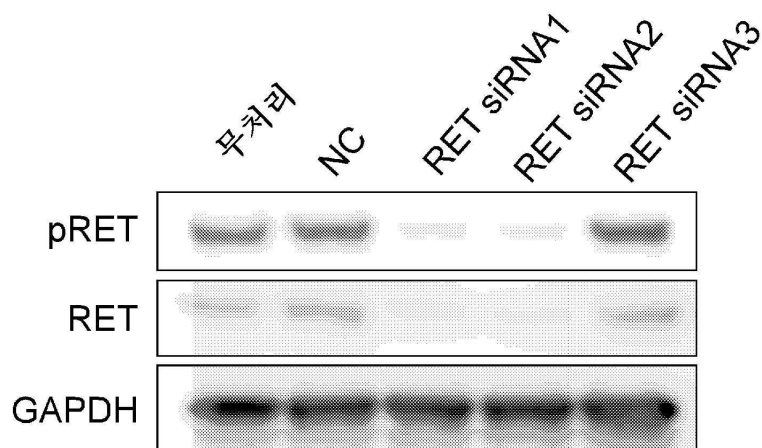
도면5



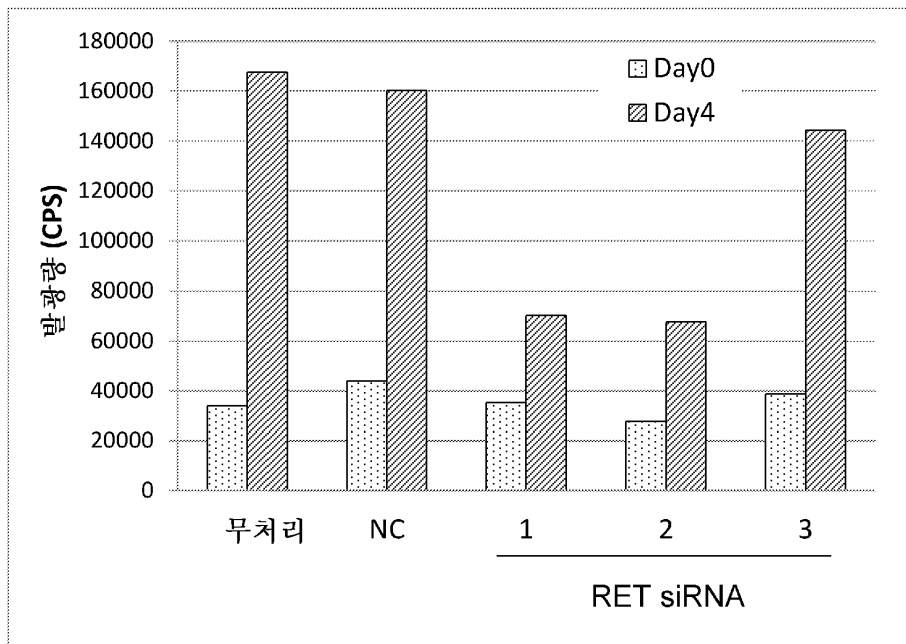
도면6



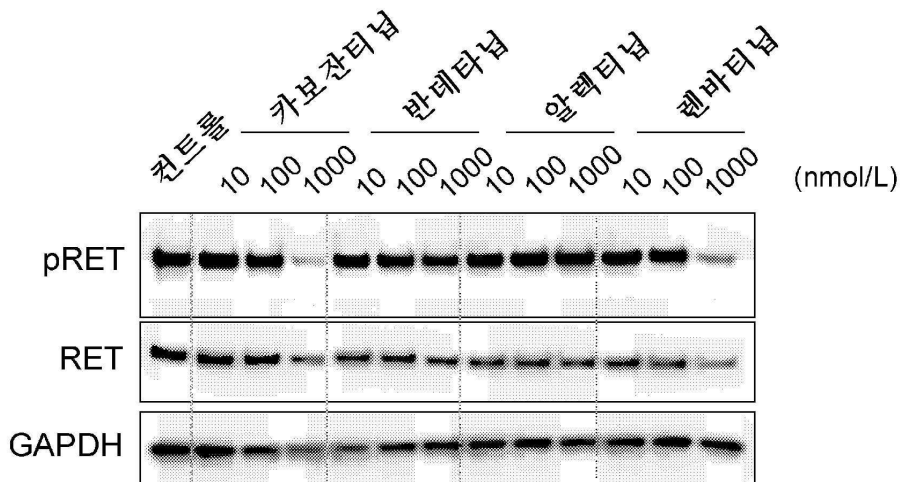
도면7



도면8



도면9



서열 목록

SEQUENCE LISTING

- <110> TAIHO PHARMACEUTICAL CO., LTD.
- <120> Fusion peptide of DCTN1 protein and RET protein
- <130> P17-158W0
- <150> JP 2017-158796
- <151> 2017-08-21
- <160> 77

<170> PatentIn version 3.5

<210> 1

<211> 4908

<212> DNA

<213> Artificial Sequence

<220><223> coding sequence of fusion peptide

<400> 1

```

atggcacaga gcaagaggca cgtgtacagc cggacgcca gcggcagcag gatgagtgcg      60
gaggcaagcg cccggcctct gcgggtgggc tcccgttag aggtgattgg aaaaggccac      120

cgaggcactg tggcctatgt tggagccaca ctgtttgcca ctggcaaatg gtaggcctg      180
attctggatg aagcaaaggg caaaaatgat ggaactgttc aaggcaggaa gtacttcaact      240
tgtgatgaag ggcatggcat ctttgtgctc cagtcccaga tccaggtatt tgaagatgga      300
gcagatacta ctccccaga gacacctgat tcttctgctt caaaagtcct caaaagagag      360
ggaactgata caactgcaaa gactagcaaa ctgcggggac tgaagcctaa gaaggcacccg      420
acagcccgaag agaccacaac tcggcgacc aagcccacgc gccagccag tactggggtg      480
gctggggcca gtagctcctt gggcccctct ggctcagcgt cagcaggtga gctgagcagc      540

agtgagccca gcaccccgcc tcagactccg ctggcagcac ccatcatccc cacgccggtc      600
ctcacctctc ctggagcagt cccccgctt ctttccccat ccaaggagga ggagggacta      660
agggtcagg tcggggacct ggaggagaaa ctagagacc tgagactgaa acgggcagaa      720
gacaaagcaa agctaaaaga gctggagaaa cacaaaatcc agctggagca ggtgcaggaa      780
tggaagagca aaatgcagga gcagcaggcc gacctgcagc ggcgctcaa ggaggcgaga      840
aaggaagcca aggagcgct ggaggcaaaag gaacgtata tggaggagat ggctgatact      900
gctgatgcca ttgagatggc cactttggac aaggagatgg ctgaagagcg ggctgagtcc      960

ctgcagcagg aggtggaggc actgaaggag cgggtggacg agctcactac tgacttagag      1020
atcctcaagg ctgagattga agagaagggc tcagatggcg ctgcatccag ttatcagctc      1080
aagcagcttg aggagcagaa tgcccgcctg aaggatgccc tggtaggat gcgggatctt      1140
tcttctcag agaagcagga gcatgtgaag ctccagaagc tcatggaaaa gaagaaccaa      1200
gagctggaag ttgtgaggca acagcgggag cgtctgcagg aggagctaag ccaggcagag      1260
agcaccattg atagactcaa ggagcaggtg gatgctgctc tgggtgctga ggagatggtg      1320
gagatgctga cagatcggaa cctgaatctg gaagagaaag tgcgcgagtt gagggagact      1380

gtgggagact tggaagcgat gaatgagatg aacgatgagc tgcaggagaa tgcacgtgag      1440

```

acagaactgg agctgcggga gcagctggac atggcaggcg cgcgggttcg tgaggcccag 1500  
 aagcgtgtgg aggcagccca ggagacggtt gcagactacc agcagaccat caagaagtac 1560  
 cgccagctga ccgccatct acaggatgtg aatcgggaac tgacaaacca gcaggaagca 1620  
 tctgtggaga ggcaacagca gccacctcca gagaccttg acttcaaaat caagtttgct 1680  
 gagactaagg cccatgcca ggcaattgag atggaattga ggcagatgga ggtggcccag 1740  
 gccaatcgac acatgtccct gctgacagcc ttcattgctg acagcttctc tcggccaggt 1800  
  
 ggggaccatg actgcgttct ggtgctgttg ctcatgcctc gtctcatttg caaggcagag 1860  
 ctgatccgga agcaggccca ggagaagttt gaactaagtg agaactgttc agagcggcct 1920  
 gggctcgcgag gagctgctgg ggagcaactc agctttgctg ctggactggt gtactcgctg 1980  
 agcctgctgc aggccacgct acaccgctat gagcatgcc tctctcagtg cagtgtggat 2040  
 gtgtataaga aagtggcgag cctgtaccct gagatgagtg cccatgagcg ctcttggat 2100  
 ttctcattg aactgctgca caaggatcag ctggatgaga ctgtcaatgt ggagcctctc 2160  
 accaaggcca tcaagtacta tcagcatctg tacagcatcc accttgccga acagcctgag 2220  
  
 gactgtacta tgcagctggc tgaccacatt aagttcacgc agagtgtctt ggactgcatg 2280  
 agtgtggagg taggacggct gcgtgccttc ttgcagggtg ggcaggaggc tacagatatt 2340  
 gccctcctgc tccgggatct ggaaacttca tgcagtgaca tccgccagtt ctgcaagaag 2400  
 atccgaagge gaatgccagg gacagatgct cctgggatcc cagctgcact ggcctttgga 2460  
 ccacaggtat ctgacacgct cctagactgc aggaaacact tgactgggt cgtggctgtg 2520  
 ctgcaggagg tggcagctgc tctgcccag ctattgccc cactggcaga gaatgagggg 2580  
 ctacttgtgg ctgctctgga ggaactggct ttcaaagcaa gcgagcagat ctatgggacc 2640  
  
 ccctccagca gccctatga gtgtctgcgc cagtcatgca acatcctcat cagtaccatg 2700  
 aacaagctgg ccacagccat gcaggagggg gattatgatg cagagcggcc ccccagcaag 2760  
 cctccaccgg ttgaactgcg ggctgctgcc ctctctgcag agatcacaga tgctgaaggc 2820  
 ctgggtttga agctcgaaga tgcagagaca gttattaagg agttgaagaa gtcaactcaag 2880  
 attaaggag aggagctaag tgaggccaat gtgcggtga gcctcctgga gaagaagttg 2940  
 gacagtgtg ccaaggatgc agatgagcgc atcgagaaag tccagactcg gctggaggag 3000  
 accaggcac tgctgcgaaa gaaggagaaa gattttgagg agacaatgga tgcaactccag 3060  
  
 gctgacatcg accagctgga ggcagagaag gcagaactaa agcagcgtct gaacagccag 3120  
 tccaaacgca cgattgaggg actccggggc cctcctcctt caggcattgc tactctggtc 3180  
 tctggcattg ctggtgaaga acagcagcga ggagccatcc ctgggcaggc tccagggtct 3240  
 gtgccaggcc cagggtggt gaaggactca ccactgctgc ttcagcagat ctctgcatg 3300

aggctgcaca tctcccagct ccagcatgag aacagcatcc tcaagggagc ccagatgaag 3360  
 gcatccttgg catccctgcc ccctctgcat gttgcaaagc tatcccatga gggccctggc 3420  
 agtgagt tac cagctggagc gctgtatcgt aagaccagcc agctgctgga gacattgaat 3480

caattgagca cacacacgca cgtagtagac atcactcgca ccagccctgc tgccaagagc 3540  
 ccgtcggccc aacttatgga gcaagtggct cagcttaagt ccctgagtga caccgtcgag 3600  
 aagctcaagg atgaggtcct caaggagaca gtatctcagc gccctggagc cacagtaccc 3660  
 actgactttg ccaccttccc ttcacagcc ttctcaggg aggatccaaa gtgggaattc 3720  
 cctcgggaaga acttggttct tggaaaaact ctaggagaag gcgaatttgg aaaagtggtc 3780  
 aaggcaacgg ccttccatct gaaaggcaga gcagggtaca ccacggtggc cgtgaagatg 3840  
 ctgaaagaga acgcctcccc gagtgagctt cgagacctgc tgtcagagtt caacgtcctg 3900

aagcaggtca accaccaca tgtcatcaaa ttgtatgggg cctgcagcca ggatggcccg 3960  
 ctctctctca tcgtggagta cgccaatac ggctccctgc ggggcttctc cgcgagagc 4020  
 cgcaaagtgg ggcctggcta cctgggcagt ggaggcagcc gcaactccag ctccctggac 4080  
 caccgggatg agcgggccct caccatgggc gacctcatct catttgcctg gcagatctca 4140  
 caggggatgc agtacttggc cgagatgaag ctcgttcatc gggacttggc agccagaaac 4200  
 atcctggtag ctgaggggcg gaagatgaag atttcgatt tcggcttgtc ccgagatgtt 4260  
 tatgaagagg attcctacgt gaagaggagc cagggtcggg ttccagttaa atggatggca 4320

attgaatccc tttttgatca tatctacacc acgcaaagtg atgtatggtc ttttgggtgc 4380  
 ctgctgtggg agatcgtgac cctaggggga aaccctatc ctgggattcc tectgagcgg 4440  
 ctcttcaacc ttctgaagac cggccaccgg atggagaggc cagacaactg cagcgaggag 4500  
 atgtaccgcc tgatgctgca atgctggaag caggagccgg acaaaaggcc ggtgtttgcg 4560  
 gacatcagca aagacctgga gaagatgatg gttaagagga gagactactt ggaccttgcg 4620  
 gcgtccactc catctgactc cctgatttat gacgacggcc tctcagagga ggagacaccg 4680  
 ctggtggact gtaataatgc cccctcct cgagccctcc cttccacatg gattgaaaac 4740

aaactctatg gcatgtcaga cccgaactgg cctggagaga gtcctgtacc actcacgaga 4800  
 gctgatggca ctaacactgg gtttccaaga tatccaatg atagtgtata tgctaactgg 4860  
 atgctttcac cctcagcggc aaaattaatg gacacgtttg atagttaa 4908

<210> 2

<211> 1635

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 2

Met Ala Gln Ser Lys Arg His Val Tyr Ser Arg Thr Pro Ser Gly Ser

1                    5                    10                    15

Arg Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg

20                    25                    30

Val Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly

35                    40                    45

Ala Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu

50                    55                    60

Ala Lys Gly Lys Asn Asp Gly Thr Val Gln Gly Arg Lys Tyr Phe Thr

65                    70                    75                    80

Cys Asp Glu Gly His Gly Ile Phe Val Arg Gln Ser Gln Ile Gln Val

85                    90                    95

Phe Glu Asp Gly Ala Asp Thr Thr Ser Pro Glu Thr Pro Asp Ser Ser

100                    105                    110

Ala Ser Lys Val Leu Lys Arg Glu Gly Thr Asp Thr Thr Ala Lys Thr

115                    120                    125

Ser Lys Leu Arg Gly Leu Lys Pro Lys Lys Ala Pro Thr Ala Arg Lys

130                    135                    140

Thr Thr Thr Arg Arg Pro Lys Pro Thr Arg Pro Ala Ser Thr Gly Val

145                    150                    155                    160

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

165                    170                    175

Glu Leu Ser Ser Ser Glu Pro Ser Thr Pro Ala Gln Thr Pro Leu Ala

180                    185                    190

Ala Pro Ile Ile Pro Thr Pro Val Leu Thr Ser Pro Gly Ala Val Pro

195                    200                    205

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

210                    215                    220

Arg Asp Leu Glu Glu Lys Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu

225                    230                    235                    240  
 Asp Lys Ala Lys Leu Lys Glu Leu Glu Lys His Lys Ile Gln Leu Glu  
                                  245                    250                    255  
 Gln Val Gln Glu Trp Lys Ser Lys Met Gln Glu Gln Gln Ala Asp Leu  
                                  260                    265                    270  
 Gln Arg Arg Leu Lys Glu Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu  
  
                                  275                    280                    285  
 Ala Lys Glu Arg Tyr Met Glu Glu Met Ala Asp Thr Ala Asp Ala Ile  
                                  290                    295                    300  
 Glu Met Ala Thr Leu Asp Lys Glu Met Ala Glu Glu Arg Ala Glu Ser  
 305                    310                    315                    320  
 Leu Gln Gln Glu Val Glu Ala Leu Lys Glu Arg Val Asp Glu Leu Thr  
                                  325                    330                    335  
 Thr Asp Leu Glu Ile Leu Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp  
  
                                  340                    345                    350  
 Gly Ala Ala Ser Ser Tyr Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala  
                                  355                    360                    365  
 Arg Leu Lys Asp Ala Leu Val Arg Met Arg Asp Leu Ser Ser Ser Glu  
                                  370                    375                    380  
 Lys Gln Glu His Val Lys Leu Gln Lys Leu Met Glu Lys Lys Asn Gln  
 385                    390                    395                    400  
 Glu Leu Glu Val Val Arg Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu  
  
                                  405                    410                    415  
 Ser Gln Ala Glu Ser Thr Ile Asp Glu Leu Lys Glu Gln Val Asp Ala  
                                  420                    425                    430  
 Ala Leu Gly Ala Glu Glu Met Val Glu Met Leu Thr Asp Arg Asn Leu  
                                  435                    440                    445  
 Asn Leu Glu Glu Lys Val Arg Glu Leu Arg Glu Thr Val Gly Asp Leu  
                                  450                    455                    460  
 Glu Ala Met Asn Glu Met Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu  
  
 465                    470                    475                    480

Thr Glu Leu Glu Leu Arg Glu Gln Leu Asp Met Ala Gly Ala Arg Val  
 485 490 495  
 Arg Glu Ala Gln Lys Arg Val Glu Ala Ala Gln Glu Thr Val Ala Asp  
 500 505 510  
 Tyr Gln Gln Thr Ile Lys Lys Tyr Arg Gln Leu Thr Ala His Leu Gln  
 515 520 525  
 Asp Val Asn Arg Glu Leu Thr Asn Gln Gln Glu Ala Ser Val Glu Arg  
 530 535 540  
 Gln Gln Gln Pro Pro Pro Glu Thr Phe Asp Phe Lys Ile Lys Phe Ala  
 545 550 555 560  
 Glu Thr Lys Ala His Ala Lys Ala Ile Glu Met Glu Leu Arg Gln Met  
 565 570 575  
 Glu Val Ala Gln Ala Asn Arg His Met Ser Leu Leu Thr Ala Phe Met  
 580 585 590  
 Pro Asp Ser Phe Leu Arg Pro Gly Gly Asp His Asp Cys Val Leu Val  
 595 600 605  
 Leu Leu Leu Met Pro Arg Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys  
 610 615 620  
 Gln Ala Gln Glu Lys Phe Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro  
 625 630 635 640  
 Gly Leu Arg Gly Ala Ala Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu  
 645 650 655  
 Val Tyr Ser Leu Ser Leu Leu Gln Ala Thr Leu His Arg Tyr Glu His  
 660 665 670  
 Ala Leu Ser Gln Cys Ser Val Asp Val Tyr Lys Lys Val Gly Ser Leu  
 675 680 685  
 Tyr Pro Glu Met Ser Ala His Glu Arg Ser Leu Asp Phe Leu Ile Glu  
 690 695 700  
 Leu Leu His Lys Asp Gln Leu Asp Glu Thr Val Asn Val Glu Pro Leu  
 705 710 715 720  
 Thr Lys Ala Ile Lys Tyr Tyr Gln His Leu Tyr Ser Ile His Leu Ala

725                                 730                                 735  
 Glu Gln Pro Glu Asp Cys Thr Met Gln Leu Ala Asp His Ile Lys Phe  
 740                                 745                                 750  
 Thr Gln Ser Ala Leu Asp Cys Met Ser Val Glu Val Gly Arg Leu Arg  
 755                                 760                                 765  
 Ala Phe Leu Gln Gly Gly Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu  
 770                                 775                                 780  
 Arg Asp Leu Glu Thr Ser Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys  
  
 785                                 790                                 795                                 800  
 Ile Arg Arg Arg Met Pro Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala  
 805                                 810                                 815  
 Leu Ala Phe Gly Pro Gln Val Ser Asp Thr Leu Leu Asp Cys Arg Lys  
 820                                 825                                 830  
 His Leu Thr Trp Val Val Ala Val Leu Gln Glu Val Ala Ala Ala Ala  
 835                                 840                                 845  
 Ala Gln Leu Ile Ala Pro Leu Ala Glu Asn Glu Gly Leu Leu Val Ala  
  
 850                                 855                                 860  
 Ala Leu Glu Glu Leu Ala Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr  
 865                                 870                                 875                                 880  
 Pro Ser Ser Ser Pro Tyr Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu  
 885                                 890                                 895  
 Ile Ser Thr Met Asn Lys Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr  
 900                                 905                                 910  
 Asp Ala Glu Arg Pro Pro Ser Lys Pro Pro Pro Val Glu Leu Arg Ala  
  
 915                                 920                                 925  
 Ala Ala Leu Arg Ala Glu Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys  
 930                                 935                                 940  
 Leu Glu Asp Arg Glu Thr Val Ile Lys Glu Leu Lys Lys Ser Leu Lys  
 945                                 950                                 955                                 960  
 Ile Lys Gly Glu Glu Leu Ser Glu Ala Asn Val Arg Leu Ser Leu Leu  
 965                                 970                                 975

Glu Lys Lys Leu Asp Ser Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu

	980		985		990
Lys Val Gln Thr Arg Leu Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys					
	995		1000		1005
Glu Lys Glu Phe Glu Glu Thr Met Asp Ala Leu Gln Ala Asp Ile					
	1010		1015		1020
Asp Gln Leu Glu Ala Glu Lys Ala Glu Leu Lys Gln Arg Leu Asn					
	1025		1030		1035
Ser Gln Ser Lys Arg Thr Ile Glu Gly Leu Arg Gly Pro Pro Pro					
	1040		1045		1050
Ser Gly Ile Ala Thr Leu Val Ser Gly Ile Ala Gly Glu Glu Gln					
	1055		1060		1065
Gln Arg Gly Ala Ile Pro Gly Gln Ala Pro Gly Ser Val Pro Gly					
	1070		1075		1080
Pro Gly Leu Val Lys Asp Ser Pro Leu Leu Leu Gln Gln Ile Ser					
	1085		1090		1095
Ala Met Arg Leu His Ile Ser Gln Leu Gln His Glu Asn Ser Ile					
	1100		1105		1110
Leu Lys Gly Ala Gln Met Lys Ala Ser Leu Ala Ser Leu Pro Pro					
	1115		1120		1125
Leu His Val Ala Lys Leu Ser His Glu Gly Pro Gly Ser Glu Leu					
	1130		1135		1140
Pro Ala Gly Ala Leu Tyr Arg Lys Thr Ser Gln Leu Leu Glu Thr					
	1145		1150		1155
Leu Asn Gln Leu Ser Thr His Thr His Val Val Asp Ile Thr Arg					
	1160		1165		1170
Thr Ser Pro Ala Ala Lys Ser Pro Ser Ala Gln Leu Met Glu Gln					
	1175		1180		1185
Val Ala Gln Leu Lys Ser Leu Ser Asp Thr Val Glu Lys Leu Lys					
	1190		1195		1200
Asp Glu Val Leu Lys Glu Thr Val Ser Gln Arg Pro Gly Ala Thr					

1205                      1210                      1215  
 Val Pro Thr Asp Phe Ala Thr Phe Pro Ser Ser Ala Phe Leu Arg

1220                      1225                      1230  
 Glu Asp Pro Lys Trp Glu Phe Pro Arg Lys Asn Leu Val Leu Gly

1235                      1240                      1245  
 Lys Thr Leu Gly Glu Gly Glu Phe Gly Lys Val Val Lys Ala Thr

1250                      1255                      1260  
 Ala Phe His Leu Lys Gly Arg Ala Gly Tyr Thr Thr Val Ala Val

1265                      1270                      1275  
 Lys Met Leu Lys Glu Asn Ala Ser Pro Ser Glu Leu Arg Asp Leu

1280                      1285                      1290  
 Leu Ser Glu Phe Asn Val Leu Lys Gln Val Asn His Pro His Val

1295                      1300                      1305  
 Ile Lys Leu Tyr Gly Ala Cys Ser Gln Asp Gly Pro Leu Leu Leu

1310                      1315                      1320  
 Ile Val Glu Tyr Ala Lys Tyr Gly Ser Leu Arg Gly Phe Leu Arg

1325                      1330                      1335  
 Glu Ser Arg Lys Val Gly Pro Gly Tyr Leu Gly Ser Gly Gly Ser

1340                      1345                      1350  
 Arg Asn Ser Ser Ser Leu Asp His Pro Asp Glu Arg Ala Leu Thr

1355                      1360                      1365  
 Met Gly Asp Leu Ile Ser Phe Ala Trp Gln Ile Ser Gln Gly Met

1370                      1375                      1380  
 Gln Tyr Leu Ala Glu Met Lys Leu Val His Arg Asp Leu Ala Ala

1385                      1390                      1395  
 Arg Asn Ile Leu Val Ala Glu Gly Arg Lys Met Lys Ile Ser Asp

1400                      1405                      1410  
 Phe Gly Leu Ser Arg Asp Val Tyr Glu Glu Asp Ser Tyr Val Lys

1415                      1420                      1425  
 Arg Ser Gln Gly Arg Ile Pro Val Lys Trp Met Ala Ile Glu Ser

1430                      1435                      1440

Leu Phe Asp His Ile Tyr Thr Thr Gln Ser Asp Val Trp Ser Phe  
 1445 1450 1455  
 Gly Val Leu Leu Trp Glu Ile Val Thr Leu Gly Gly Asn Pro Tyr  
 1460 1465 1470  
 Pro Gly Ile Pro Pro Glu Arg Leu Phe Asn Leu Leu Lys Thr Gly  
 1475 1480 1485  
 His Arg Met Glu Arg Pro Asp Asn Cys Ser Glu Glu Met Tyr Arg  
 1490 1495 1500  
 Leu Met Leu Gln Cys Trp Lys Gln Glu Pro Asp Lys Arg Pro Val  
 1505 1510 1515  
 Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys Met Met Val Lys Arg  
 1520 1525 1530  
 Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr Pro Ser Asp Ser Leu  
 1535 1540 1545  
 Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu Thr Pro Leu Val Asp  
 1550 1555 1560  
 Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu Pro Ser Thr Trp Ile  
 1565 1570 1575  
 Glu Asn Lys Leu Tyr Gly Met Ser Asp Pro Asn Trp Pro Gly Glu  
 1580 1585 1590  
 Ser Pro Val Pro Leu Thr Arg Ala Asp Gly Thr Asn Thr Gly Phe  
 1595 1600 1605  
 Pro Arg Tyr Pro Asn Asp Ser Val Tyr Ala Asn Trp Met Leu Ser  
 1610 1615 1620  
 Pro Ser Ala Ala Lys Leu Met Asp Thr Phe Asp Ser  
 1625 1630 1635  
 <210> 3  
 <211> 4782  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> coding sequence of fusion peptide

<400> 3

atggcacaga gcaagaggca cgtgtacagc cggacgcca gcggcagcag gatgagtgcg 60  
gaggcaagcg cccggcctct gcgggtgggc tcccgtgtag aggtgattgg aaaagccac 120  
cgaggcactg tggcctaigt tggagccaca ctgtttgcca ctggcaaatg ggtaggcgtg 180  
attctggatg aagcaaaggc caaaaatgat ggaactgttc aaggcaggaa gtacttact 240  
tgtgatgaag ggcattggcat ctttgtgcgc cagtcccaga tccaggtatt tgaagatgga 300  
gcagatacta ctccccaga gacacctgat tcttctgctt caaaagtctt caaaagagag 360  
ggaactgata caactgcaaa gactagcaaa ctgcggggac tgaagcctaa gaaggcaccg 420  
  
acagcccgaag agaccacaac tcggcgaccc aagcccacgc gcccagccag tactggggtg 480  
gctggggcca gtagctccct gggcccctct ggctcagcgt cagcaggtga gctgagcagc 540  
agtgagccca gcaccccgcc tcagactccg ctggcagcac ccatcatccc cacgccggtc 600  
ctcacctctc ctggagcagt cccccgctt ccttccccat ccaaggagga ggagggacta 660  
agggctcagg tgcgggacct ggaggagaaa ctagagacc tgagactgaa acgggcagaa 720  
gacaaagcaa agctaaaaga gctggagaaa cacaaaatcc agctggagca ggtgcaggaa 780  
tgaagagca aatgcagga gcagcaggcc gacctgcagc ggcgctcaa ggaggcgaga 840  
  
aaggaagcca aggaggcgtt ggaggcaaag gaacgctata tggaggagat ggctgatact 900  
gctgatgcca ttgataggc cactttggac aaggagatgg ctgaagagcg ggctgagtc 960  
ctgcagcagg aggtggaggc actgaaggag cgggtggacg agctcactac tgacttagag 1020  
atcctcaagg ctgagattga agagaagggc tcagatggcg ctgcatccag ttatcagctc 1080  
aagcagcttg aggagcagaa tgccccctg aaggatgccc tggtaggat gcgggatctt 1140  
tcttctcag agaagcagga gcatgtgaag ctccagaagc tcatggaaa gaagaaccaa 1200  
gagctggaag ttgtgaggca acagcgggag cgtctgcagg aggagctaag ccaggcagag 1260  
  
agcaccattg atgagctcaa ggagcaggtg gatgctgctc tgggtgctga ggagatggtg 1320  
gagatgctga cagatcgga cctgaatctg gaagagaaag tgcgcgagtt gagggagact 1380  
gtgggagact tgaagcgat gaatgagatg aacgatgagc tgcaggagaa tgcacgtgag 1440  
acagaactgg agctgcggga gcagctggac atggcaggcg cgcgggttcg tgaggcccag 1500  
aagcgtgtgg aggcagccca ggagacggtt gcagactacc agcagacat caagaagtac 1560  
cgccagctga ccgcccctct acaggatgtg aatcggaac tgacaaacca gcaggaagca 1620  
tctgtggaga ggcaacagca gccacctca gagaccttg acttcaaat caagtttgct 1680  
  
gagactaagg cccatgcca ggcaattgag atggaattga ggcagatgga ggtggcccag 1740  
gccaatcgac acatgtccct gctgacagcc tcatgctctg acagcttctt tcggccaggt 1800

ggggaccatg actgcgttct ggtgctgttg ctcatgcctc gtctcatttg caaggcagag 1860  
 ctgatccgga agcaggccca ggagaagtth gaactaagtg agaactgttc agagcggcct 1920  
 gggtcgcgag gagctgctgg ggagcaactc agctttgctg ctggactggt gfactcgtg 1980  
 agcctgctgc aggccacgct acaccgctat gagcatgccc tctctcagtg cagtgtggat 2040  
 gtgtataaga aagtgggcag cctgtaccct gagatgagtg cccatgagcg ctcttggat 2100  
  
 ttctcattg aactgctgca caaggatcag ctggatgaga ctgtcaatgt ggagcctctc 2160  
 accaaggcca tcaagtacta tcagcatctg tacagcatcc accttgccga acagcctgag 2220  
 gactgtacta tgcagctggc tgaccacatt aagtccacgc agagtgtctt ggactgcatg 2280  
 agtgtggagg taggacggct gcgtgccttc ttgcagggtg ggacaggaggc tacagatatt 2340  
 gccctcctgc tccgggatct ggaaacttca tgcaagtaca tccgccagtt ctgcaagaag 2400  
 atccgaaggc gaatgccagg gacagatgct cctgggatcc cagctgcaact ggcccttggg 2460  
 ccacaggtat ctgacacgct cctagactgc aggaaacact tgacgtgggt cgtggctgtg 2520  
  
 ctgcaggagg tggcagctgc tgctgcccag ctcttggccc cactggcaga gaatgagggg 2580  
 ctacttgtgg ctgctctgga ggaactggct ttcaaagcaa gcgagcagat ctatgggacc 2640  
 cctccagca gccctatga gtgtctgctc cagtcatgca acatcctcat cagtaccatg 2700  
 aacaagctgg ccacagccat gcaggagggg gagtatgatg cagagcggcc ccccagcaag 2760  
 cctccaccgg ttgaactgcg ggctgctgcc ctctctgctg agatcacaga tgctgaaggc 2820  
 ctgggttga agctcgaaga tcgagagaca gttattaagg agttgaagaa gtcactcaag 2880  
 attaaggagg aggagctaag tgaggccaat gtgcggctga gcctcctgga gaagaagttg 2940  
  
 gacagtgtg ccaaggatgc agatgagcgc atcgagaaag tccagactcg gctggaggag 3000  
 acccaggcac tgcctgcaaa gaaggagaaa gattttgagg agacaatgga tgcaactcag 3060  
 gctgacatcg accagctgga ggcagagaag gcagaactaa agcagcgtct gaacagccag 3120  
 tccaaacgca cgattgaggg actccggggc cctcctcctt caggcattgc tactctggtc 3180  
 tctggcattg ctggtgaaga acagcagcga ggagccatcc ctgggcagcg tccagggtct 3240  
 gtgccaggcc cagggtctgt gaaggactca ccaactgctgc ttcagcagat ctctgcatg 3300  
 aggctgcaca tctcccagct ccagcatgag aacagcatcc tcaaggagc ccagatgaag 3360  
  
 gcatccttg catccttgc ccctctgcat gttgcaaagc tatcccatga gggccttggc 3420  
 agtgagttac cagctggagc gctgtatcgt aagaccagcc agctgctgga gacattgaat 3480  
 caattgagca cacacacgca cgtagtagac atcactcga ccagccctgc tgccaagagc 3540  
 ccgtcggccc aacttatgga gcaagtggct cagcttaagt cctgagtgga caccgtcgag 3600  
 aagctcaagg atgaggtcct caaggagaca gtatctcagc gccctggagc cacagtacc 3660

actgactttg ccaccttccc ttcacagacc ttctcaggg aggatccaaa gtgggaattc 3720  
 cctcggaga acttggttct tggaaaaact ctaggagaag gcgaatttgg aaaagtggc 3780

aaggcaacgg ctttccatct gaaaggcaga gcagggtaca ccacggtggc cgtgaagatg 3840  
 ctgaaagaga acgcctcccc gaggtagctt cgagacctgc tgtcagagtt caacgtcctg 3900  
 aagcaggtca accaccaca tgtcatcaaa ttgtatgggg cctgcagcca ggatggcccg 3960  
 ctctcctca tegtggagta cgccaatac ggctccctgc ggggcttct cecgagagc 4020  
 cgcaaagtgg ggcttgcta cctgggcagt ggaggcagcc gcaactccag ctccctggac 4080  
 caccggatg agcgggccct caccatgggc gacctcatct catttgcctg gcagatctca 4140  
 caggggatgc agtatctggc cgagatgaag ctcgttcatc gggacttggc agccagaaac 4200

atcctggtag ctgaggggcg gaagatgaag atttcgatt tcggcttgtc ccgagatgtt 4260  
 tatgaagagg attcctactg gaagaggagc cagggtcggg ttccagttaa atggatggca 4320  
 attgaatccc tttttgatca tatctacacc acgcaaagtg atgtatggtc ttttgggtgc 4380  
 ctgctgtggg agatcgtgac cctaggggga aaccctatc ctgggattcc tctgagcgg 4440  
 ctcttcaacc ttctgaagac cggccaccgg atggagaggc cagacaactg cagcgaggag 4500  
 atgtaccgcc tgatcgtgca atgctggaag caggagccgg acaaaaggcc ggtgtttgcg 4560  
 gacatcagca aagacctgga gaagatgatg gttaagagga gagactactt ggaccttgcg 4620

gcgtccactc catctgactc cctgatttat gacgacggcc tctcagagga ggagacaccg 4680  
 ctggtggact gtaataatgc cccctcct cgagccctcc cttccacatg gattgaaaac 4740  
 aaactctatg gtagaatttc ccatgcattt actagattct ag 4782

<210> 4

<211> 1593

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 4

Met Ala Gln Ser Lys Arg His Val Tyr Ser Arg Thr Pro Ser Gly Ser  
 1                    5                    10                    15

Arg Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg

                  20                    25                    30

Val Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly  
                   35                    40                    45

Ala Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu  
 50 55 60  
 Ala Lys Gly Lys Asn Asp Gly Thr Val Gln Gly Arg Lys Tyr Phe Thr  
 65 70 75 80  
 Cys Asp Glu Gly His Gly Ile Phe Val Arg Gln Ser Gln Ile Gln Val  
 85 90 95  
 Phe Glu Asp Gly Ala Asp Thr Thr Ser Pro Glu Thr Pro Asp Ser Ser  
 100 105 110  
 Ala Ser Lys Val Leu Lys Arg Glu Gly Thr Asp Thr Thr Ala Lys Thr  
 115 120 125  
 Ser Lys Leu Arg Gly Leu Lys Pro Lys Lys Ala Pro Thr Ala Arg Lys  
 130 135 140  
 Thr Thr Thr Arg Arg Pro Lys Pro Thr Arg Pro Ala Ser Thr Gly Val  
 145 150 155 160  
 Ala Gly Ala Ser Ser Ser Leu Gly Pro Ser Gly Ser Ala Ser Ala Gly  
 165 170 175  
 Glu Leu Ser Ser Ser Glu Pro Ser Thr Pro Ala Gln Thr Pro Leu Ala  
 180 185 190  
 Ala Pro Ile Ile Pro Thr Pro Val Leu Thr Ser Pro Gly Ala Val Pro  
 195 200 205  
 Pro Leu Pro Ser Pro Ser Lys Glu Glu Glu Gly Leu Arg Ala Gln Val  
 210 215 220  
 Arg Asp Leu Glu Glu Lys Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu  
 225 230 235 240  
 Asp Lys Ala Lys Leu Lys Glu Leu Glu Lys His Lys Ile Gln Leu Glu  
 245 250 255  
 Gln Val Gln Glu Trp Lys Ser Lys Met Gln Glu Gln Gln Ala Asp Leu  
 260 265 270  
 Gln Arg Arg Leu Lys Glu Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu  
 275 280 285  
 Ala Lys Glu Arg Tyr Met Glu Glu Met Ala Asp Thr Ala Asp Ala Ile

290                      295                      300  
 Glu Met Ala Thr Leu Asp Lys Glu Met Ala Glu Glu Arg Ala Glu Ser  
 305                      310                      315                      320  
 Leu Gln Gln Glu Val Glu Ala Leu Lys Glu Arg Val Asp Glu Leu Thr  
                          325                      330                      335  
 Thr Asp Leu Glu Ile Leu Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp  
  
                          340                      345                      350  
 Gly Ala Ala Ser Ser Tyr Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala  
                          355                      360                      365  
 Arg Leu Lys Asp Ala Leu Val Arg Met Arg Asp Leu Ser Ser Ser Glu  
                          370                      375                      380  
 Lys Gln Glu His Val Lys Leu Gln Lys Leu Met Glu Lys Lys Asn Gln  
 385                      390                      395                      400  
 Glu Leu Glu Val Val Arg Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu  
  
                          405                      410                      415  
 Ser Gln Ala Glu Ser Thr Ile Asp Glu Leu Lys Glu Gln Val Asp Ala  
                          420                      425                      430  
 Ala Leu Gly Ala Glu Glu Met Val Glu Met Leu Thr Asp Arg Asn Leu  
                          435                      440                      445  
 Asn Leu Glu Glu Lys Val Arg Glu Leu Arg Glu Thr Val Gly Asp Leu  
                          450                      455                      460  
 Glu Ala Met Asn Glu Met Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu  
  
 465                      470                      475                      480  
 Thr Glu Leu Glu Leu Arg Glu Gln Leu Asp Met Ala Gly Ala Arg Val  
                          485                      490                      495  
 Arg Glu Ala Gln Lys Arg Val Glu Ala Ala Gln Glu Thr Val Ala Asp  
                          500                      505                      510  
 Tyr Gln Gln Thr Ile Lys Lys Tyr Arg Gln Leu Thr Ala His Leu Gln  
                          515                      520                      525  
 Asp Val Asn Arg Glu Leu Thr Asn Gln Gln Glu Ala Ser Val Glu Arg  
  
                          530                      535                      540

Gln Gln Gln Pro Pro Pro Glu Thr Phe Asp Phe Lys Ile Lys Phe Ala  
 545                      550                      555                      560

Glu Thr Lys Ala His Ala Lys Ala Ile Glu Met Glu Leu Arg Gln Met  
                                  565                      570                      575

Glu Val Ala Gln Ala Asn Arg His Met Ser Leu Leu Thr Ala Phe Met  
                                  580                      585                      590

Pro Asp Ser Phe Leu Arg Pro Gly Gly Asp His Asp Cys Val Leu Val  
                                  595                      600                      605

Leu Leu Leu Met Pro Arg Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys  
                                  610                      615                      620

Gln Ala Gln Glu Lys Phe Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro  
 625                      630                      635                      640

Gly Leu Arg Gly Ala Ala Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu  
                                  645                      650                      655

Val Tyr Ser Leu Ser Leu Leu Gln Ala Thr Leu His Arg Tyr Glu His  
                                  660                      665                      670

Ala Leu Ser Gln Cys Ser Val Asp Val Tyr Lys Lys Val Gly Ser Leu  
                                  675                      680                      685

Tyr Pro Glu Met Ser Ala His Glu Arg Ser Leu Asp Phe Leu Ile Glu  
                                  690                      695                      700

Leu Leu His Lys Asp Gln Leu Asp Glu Thr Val Asn Val Glu Pro Leu  
 705                      710                      715                      720

Thr Lys Ala Ile Lys Tyr Tyr Gln His Leu Tyr Ser Ile His Leu Ala  
                                  725                      730                      735

Glu Gln Pro Glu Asp Cys Thr Met Gln Leu Ala Asp His Ile Lys Phe  
                                  740                      745                      750

Thr Gln Ser Ala Leu Asp Cys Met Ser Val Glu Val Gly Arg Leu Arg  
                                  755                      760                      765

Ala Phe Leu Gln Gly Gly Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu  
                                  770                      775                      780

Arg Asp Leu Glu Thr Ser Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys



Ser Gln Ser Lys Arg Thr Ile Glu Gly Leu Arg Gly Pro Pro Pro  
 1040 1045 1050  
 Ser Gly Ile Ala Thr Leu Val Ser Gly Ile Ala Gly Glu Glu Gln  
 1055 1060 1065  
 Gln Arg Gly Ala Ile Pro Gly Gln Ala Pro Gly Ser Val Pro Gly  
 1070 1075 1080  
 Pro Gly Leu Val Lys Asp Ser Pro Leu Leu Leu Gln Gln Ile Ser  
 1085 1090 1095  
 Ala Met Arg Leu His Ile Ser Gln Leu Gln His Glu Asn Ser Ile  
 1100 1105 1110  
 Leu Lys Gly Ala Gln Met Lys Ala Ser Leu Ala Ser Leu Pro Pro  
 1115 1120 1125  
 Leu His Val Ala Lys Leu Ser His Glu Gly Pro Gly Ser Glu Leu  
 1130 1135 1140  
 Pro Ala Gly Ala Leu Tyr Arg Lys Thr Ser Gln Leu Leu Glu Thr  
 1145 1150 1155  
 Leu Asn Gln Leu Ser Thr His Thr His Val Val Asp Ile Thr Arg  
 1160 1165 1170  
 Thr Ser Pro Ala Ala Lys Ser Pro Ser Ala Gln Leu Met Glu Gln  
 1175 1180 1185  
 Val Ala Gln Leu Lys Ser Leu Ser Asp Thr Val Glu Lys Leu Lys  
 1190 1195 1200  
 Asp Glu Val Leu Lys Glu Thr Val Ser Gln Arg Pro Gly Ala Thr  
 1205 1210 1215  
 Val Pro Thr Asp Phe Ala Thr Phe Pro Ser Ser Ala Phe Leu Arg  
 1220 1225 1230  
 Glu Asp Pro Lys Trp Glu Phe Pro Arg Lys Asn Leu Val Leu Gly  
 1235 1240 1245  
 Lys Thr Leu Gly Glu Gly Glu Phe Gly Lys Val Val Lys Ala Thr  
 1250 1255 1260  
 Ala Phe His Leu Lys Gly Arg Ala Gly Tyr Thr Thr Val Ala Val

1265                      1270                      1275  
 Lys Met Leu Lys Glu Asn Ala Ser Pro Ser Glu Leu Arg Asp Leu

1280                      1285                      1290  
 Leu Ser Glu Phe Asn Val Leu Lys Gln Val Asn His Pro His Val

1295                      1300                      1305  
 Ile Lys Leu Tyr Gly Ala Cys Ser Gln Asp Gly Pro Leu Leu Leu

1310                      1315                      1320  
 Ile Val Glu Tyr Ala Lys Tyr Gly Ser Leu Arg Gly Phe Leu Arg

1325                      1330                      1335  
 Glu Ser Arg Lys Val Gly Pro Gly Tyr Leu Gly Ser Gly Gly Ser

1340                      1345                      1350  
 Arg Asn Ser Ser Ser Leu Asp His Pro Asp Glu Arg Ala Leu Thr

1355                      1360                      1365  
 Met Gly Asp Leu Ile Ser Phe Ala Trp Gln Ile Ser Gln Gly Met

1370                      1375                      1380  
 Gln Tyr Leu Ala Glu Met Lys Leu Val His Arg Asp Leu Ala Ala

1385                      1390                      1395  
 Arg Asn Ile Leu Val Ala Glu Gly Arg Lys Met Lys Ile Ser Asp

1400                      1405                      1410  
 Phe Gly Leu Ser Arg Asp Val Tyr Glu Glu Asp Ser Tyr Val Lys

1415                      1420                      1425  
 Arg Ser Gln Gly Arg Ile Pro Val Lys Trp Met Ala Ile Glu Ser

1430                      1435                      1440  
 Leu Phe Asp His Ile Tyr Thr Thr Gln Ser Asp Val Trp Ser Phe

1445                      1450                      1455  
 Gly Val Leu Leu Trp Glu Ile Val Thr Leu Gly Gly Asn Pro Tyr

1460                      1465                      1470  
 Pro Gly Ile Pro Pro Glu Arg Leu Phe Asn Leu Leu Lys Thr Gly

1475                      1480                      1485  
 His Arg Met Glu Arg Pro Asp Asn Cys Ser Glu Glu Met Tyr Arg

1490                      1495                      1500

Leu Met Leu Gln Cys Trp Lys Gln Glu Pro Asp Lys Arg Pro Val  
 1505 1510 1515

Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys Met Met Val Lys Arg  
 1520 1525 1530

Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr Pro Ser Asp Ser Leu  
 1535 1540 1545

Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu Thr Pro Leu Val Asp  
 1550 1555 1560

Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu Pro Ser Thr Trp Ile  
 1565 1570 1575

Glu Asn Lys Leu Tyr Gly Arg Ile Ser His Ala Phe Thr Arg Phe  
 1580 1585 1590

<210> 5

<211> 4506

<212> DNA

<213> Artificial Sequence

<220><223> coding sequence of fusion peptide

<400> 5

atgatgagac aggcaccgac agcccgaaag accacaactc ggcgacccaa gcccacgcgc 60  
 ccagccagta ctgggggtggc tggggccagt agtcacctgg gccctcttgg ctcagcgtca 120  
 gcaggtgagc tgagcagcag tgagcccagc accccggctc agactccgct ggcagcaccc 180  
 atcatcccca cgcgggtcct cacctctcct ggagcagtec ccccgcttcc ttecccatcc 240  
 aaggaggagg agggactaag ggctcaggtg cgggacctgg aggagaaact agagacctg 300  
  
 agactgaaac gggcagaaga caaagcaaag ctaaaagagc tggagaaaca caaatccag 360  
 ctggagcagg tgcaggaatg gaagagcaaa atgcaggagc agcaggccga cctgcagcgg 420  
 cgctcaagg aggcgagaaa ggaagccaag gaggcgctgg aggcaaagga acgctatatg 480  
 gaggagatgg ctgatactgc tgatgccatt gagatggcca ctttggacia ggagatggct 540  
 gaagagcggg ctgagtcctt gcagcaggag gtggaggcac tgaaggagcg ggtggacgag 600  
 ctactactg acttagagat cctcaaggct gagattgaag agaagggctc agatggcgct 660  
 gcatccagtt atcagctcaa gcagcttgag gacgagaatg cccgcctgaa ggatgcctg 720  
  
 gtgaggatgc gggatctttc ttctcagag aagcaggagc atgtgaagct ccagaagctc 780

atggaaaaga agaaccaaga gctggaagt gtgaggcaac agcgggagcg tctgcaggag 840  
 gagctaagcc aggcagagag caccattgat gagctcaagg agcaggtgga tgctgctctg 900  
 ggtgctgagg agatggtgga gatgctgaca gatcggaacc tgaatctgga agagaaagtg 960  
 cgcgagtga gggagactgt gggagacttg gaagcgatga atgagatgaa cgatgagctg 1020  
 caggagaatg cactgagac agaactggag ctgctgggagc agctggacat ggcaggcgcg 1080  
 cgggttcgtg aggcccagaa gcgtgtggag gcagcccagg agacggttgc agactaccag 1140  
  
 cagaccatca agaagtaccg ccagctgacc gcccatctac aggatgtgaa tcgggaactg 1200  
 acaaaccagc aggaagcatc tgtggagagg caacagcagc cacctccaga gaccttgac 1260  
 ttcaaaatca agtttgctga gactaaggcc catgccaagg caattgagat ggaattgagg 1320  
 cagatggagg tggcccaggc caatcgacac atgtccctgc tgacagcctt catgcctgac 1380  
 agcttcttc ggccaggtgg ggaccatgac tgcgttctgg tgctgttctt catgcctcgt 1440  
 ctcatctgca aggcagagct gatccggaag caggcccagg agaagtttga actaagtgag 1500  
 aactgttcag agcggcctgg gctgcgagga gctgctgggg agcaactcag ctttgctgct 1560  
  
 ggactggtgt actcgtgag cctgetgcag gccacgetac accgctatga gcatgcctc 1620  
 tctcagtga gtgtggatgt gtataagaaa gtgggcagcc tgtaccctga gatgagtgcc 1680  
 catgagcgt ccttggattt cctcattgaa ctgctgcaca aggatcagct ggatgagact 1740  
 gtcaatgtgg agcctctcac caagccatc aagtactatc agcatctgta cagcatccac 1800  
 cttgccgaac agcctgagga ctgtactatg cagctggctg accacattaa gttcacgcag 1860  
 agtgctctgg actgcatgag tgtggaggta ggacggctgc gtgccttctt gcagggtggg 1920  
 caggaggcta cagatattgc cctcctgctc cgggatctgg aaacttcag cagtgacatc 1980  
  
 cgccagtct gcaagaagat ccgaaggcga atgccaggga cagatgctcc tgggatccca 2040  
 gctgcaactgg ccttggacc acaggtatct gacacgetcc tagactgcag gaaacacttg 2100  
 acgtgggtcg tggctgtgct gcaggagggtg gcagctgctg ctgccagct cattgcccc 2160  
 ctggcagaga atgaggggct acttgtggct gctctggagg aactggcttt caaagcaagc 2220  
 gagcagatct atgggacccc ctccagcagc ccctatgagt gtctgcgcca gtcacgcaac 2280  
 atcctcatca gtacatgaa caagctggcc acagccatgc aggaggggga gtatgatgca 2340  
 gagcggcccc ccagcaagc tccaccggtt gaactgcggg ctgctgcctt tcgtgcagag 2400  
  
 atcacagatg ctgaaggcct gggtttgaag ctgcaagatc gagagacagt tattaaggag 2460  
 ttgaagaagt cactcaagat taaggagag gagctaagtg aggccaatgt gcggctgagc 2520  
 ctctggaga agaagtggga cagtgtgcc aaggatgcag atgagcgcac cgagaaagt 2580  
 cagactcggc tggaggagac ccaggcactg ctgcgaaaga aggagaaaga gtttgaggag 2640

acaatggatg cactccaggc tgacatcgac cagctggagg cagagaaggc agaactaaag 2700  
cagcgtctga acagccagtc caaacgcacg attgaggac tccggggccc tcctccttca 2760  
ggcattgcta ctctggtctc tggcattgct ggtgaagaac agcagcgagg agccatccct 2820

gggcaggctc cagggtctgt gccaggccca gggtctgtga aggactcacc actgctgctt 2880  
cagcagatct ctgccatgag gctgcacatc tcccagctcc agcatgagaa cagcatcctc 2940  
aagggagccc agatgaaggc atccttggca tcctgcccc ctctgcatgt tgcaaagcta 3000  
tcccatgagg gccttgagc tgagttacca gctggagcgc tgtatcgtaa gaccagccag 3060  
ctgctggaga cattgaatca attgagcaca cacacgcacg tagtagacat cactcgcacc 3120  
agccctgctg ccaagagccc gtcggcccaa cttatggagc aagtggctca gcttaagtcc 3180  
ctgagtgaca ccgtcgagaa gctcaaggat gaggtcctca aggagacagt atctcagcgc 3240

cctggagcca cagtaccac tgactttgcc accttccctt catcagcctt cctcaggag 3300  
gatccaaagt ggaattccc tcggaagaac ttggttcttg gaaaaactct aggagaaggc 3360  
gaatttggaa aagtggtaaa ggcaacggcc ttccatctga aaggcagagc aggttacacc 3420  
acggtggccg tgaagatgct gaaagagaac gcctccccga gtgagcttcg agacctgctg 3480  
tcagagtcca acgtcctgaa gcaggtcaac caccacatg tcatcaaatt gtatggggcc 3540  
tgcagccagg atggcccgt cctcctcacc gtggagtacg ccaaatacgg ctccctgcgg 3600  
ggcttctcc gcgagagccc caaagtgggg cctggctacc tgggcagtgg aggcagccgc 3660

aactccagct ccttggacca cccgatgag cgggccctca ccatggcga cctcatctca 3720  
tttgcttggc agatctcaca ggggatgcag tatctggccg agatgaagct cgttcatcgg 3780  
gacttggcag ccagaaacat cctggtagct gaggggcgga agatgaagat ttcgatttc 3840  
ggcttgtccc gagatgttta tgaagaggat tcctacgtga agaggagcca gggtcggatt 3900  
ccagttaaat ggatggcaat tgaatccctt tttgatcata tctacaccac gcaaagtgat 3960  
glatggtctt ttggtgtcct gctgtgggag atcgtgacc tagggggaaa ccctatcct 4020  
gggattctc ctgagcggt cttcaacctt ctgaagacc gccaccgat ggagaggcca 4080

gacaactgca gcgaggagat gtaccgcctg atgctgcaat gctggaagca ggagccggac 4140  
aaaaggccgg tgtttgcgga catcagcaaa gacctggaga agatgatggt taagaggaga 4200  
gactacttgg accttgcggc gtccactcca tctgactccc tgatttatga cgacggcctc 4260  
tcagaggagg agacaccgct ggtggactgt aataatgccc ccctccctcg agccctccct 4320  
tccacatgga ttgaaaacaa actctatggc atgtcagacc cgaactggcc tggagagagt 4380  
cctgtaccac tcacgagagc tgatggcact aacactgggt ttccaagata tccaaatgat 4440

agtgatatg ctaactggat gctttcaccc tcagcgcaa aattaatgga cacgtttgat 4500

agttaa 4506

<210> 6

<211> 1501

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 6

Met Met Arg Gln Ala Pro Thr Ala Arg Lys Thr Thr Thr Arg Arg Pro

1 5 10 15

Lys Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser Ser Ser

20 25 30

Leu Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser Ser Glu

35 40 45

Pro Ser Thr Pro Ala Gln Thr Pro Leu Ala Ala Pro Ile Ile Pro Thr

50 55 60

Pro Val Leu Thr Ser Pro Gly Ala Val Pro Pro Leu Pro Ser Pro Ser

65 70 75 80

Lys Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu Glu Lys

85 90 95

Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys Leu Lys

100 105 110

Glu Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu Trp Lys

115 120 125

Ser Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu Lys Glu

130 135 140

Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg Tyr Met

145 150 155 160

Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr Leu Asp

165 170 175

Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu Val Glu



Lys Ala Ile Glu Met Glu Leu Arg Gln Met Glu Val Ala Gln Ala Asn  
 435 440 445  
 Arg His Met Ser Leu Leu Thr Ala Phe Met Pro Asp Ser Phe Leu Arg  
 450 455 460  
 Pro Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met Pro Arg  
 465 470 475 480  
 Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu Lys Phe  
 485 490 495  
 Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly Ala Ala  
 500 505 510  
 Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu Ser Leu  
 515 520 525  
 Leu Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln Cys Ser  
 530 535 540  
 Val Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met Ser Ala  
 545 550 555 560  
 His Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys Asp Gln  
 565 570 575  
 Leu Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile Lys Tyr  
 580 585 590  
 Tyr Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu Asp Cys  
 595 600 605  
 Thr Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala Leu Asp  
 610 615 620  
 Cys Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln Gly Gly  
 625 630 635 640  
 Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr Ser  
 645 650 655  
 Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met Pro  
 660 665 670  
 Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro Gln



Ile Ala Gly Glu Glu Gln Gln Arg Gly Ala Ile Pro Gly Gln Ala Pro

930 935 940

Gly Ser Val Pro Gly Pro Gly Leu Val Lys Asp Ser Pro Leu Leu Leu

945 950 955 960

Gln Gln Ile Ser Ala Met Arg Leu His Ile Ser Gln Leu Gln His Glu

965 970 975

Asn Ser Ile Leu Lys Gly Ala Gln Met Lys Ala Ser Leu Ala Ser Leu

980 985 990

Pro Pro Leu His Val Ala Lys Leu Ser His Glu Gly Pro Gly Ser Glu

995 1000 1005

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

1010 1015 1020

Thr Leu Asn Gln Leu Ser Thr His Thr His Val Val Asp Ile Thr

1025 1030 1035

Arg Thr Ser Pro Ala Ala Lys Ser Pro Ser Ala Gln Leu Met Glu

1040 1045 1050

Gln Val Ala Gln Leu Lys Ser Leu Ser Asp Thr Val Glu Lys Leu

1055 1060 1065

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

1070 1075 1080

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

1085 1090 1095

Arg Glu Asp Pro Lys Trp Glu Phe Pro Arg Lys Asn Leu Val Leu

1100 1105 1110

Gly Lys Thr Leu Gly Glu Gly Glu Phe Gly Lys Val Val Lys Ala

1115 1120 1125

Thr Ala Phe His Leu Lys Gly Arg Ala Gly Tyr Thr Thr Val Ala

1130 1135 1140

Val Lys Met Leu Lys Glu Asn Ala Ser Pro Ser Glu Leu Arg Asp

1145 1150 1155

Leu Leu Ser Glu Phe Asn Val Leu Lys Gln Val Asn His Pro His

1160	1165	1170
Val Ile Lys Leu Tyr Gly Ala	Cys Ser Gln Asp Gly	Pro Leu Leu
1175	1180	1185
Leu Ile Val Glu Tyr Ala Lys	Tyr Gly Ser Leu Arg	Gly Phe Leu
1190	1195	1200
Arg Glu Ser Arg Lys Val Gly	Pro Gly Tyr Leu Gly	Ser Gly Gly
1205	1210	1215
Ser Arg Asn Ser Ser Ser Leu	Asp His Pro Asp Glu	Arg Ala Leu
1220	1225	1230
Thr Met Gly Asp Leu Ile Ser	Phe Ala Trp Gln Ile	Ser Gln Gly
1235	1240	1245
Met Gln Tyr Leu Ala Glu Met	Lys Leu Val His Arg	Asp Leu Ala
1250	1255	1260
Ala Arg Asn Ile Leu Val Ala	Glu Gly Arg Lys Met	Lys Ile Ser
1265	1270	1275
Asp Phe Gly Leu Ser Arg Asp	Val Tyr Glu Glu Asp	Ser Tyr Val
1280	1285	1290
Lys Arg Ser Gln Gly Arg Ile	Pro Val Lys Trp Met	Ala Ile Glu
1295	1300	1305
Ser Leu Phe Asp His Ile Tyr	Thr Thr Gln Ser Asp	Val Trp Ser
1310	1315	1320
Phe Gly Val Leu Leu Trp Glu	Ile Val Thr Leu Gly	Gly Asn Pro
1325	1330	1335
Tyr Pro Gly Ile Pro Pro Glu	Arg Leu Phe Asn Leu	Leu Lys Thr
1340	1345	1350
Gly His Arg Met Glu Arg Pro	Asp Asn Cys Ser Glu	Glu Met Tyr
1355	1360	1365
Arg Leu Met Leu Gln Cys Trp	Lys Gln Glu Pro Asp	Lys Arg Pro
1370	1375	1380
Val Phe Ala Asp Ile Ser Lys	Asp Leu Glu Lys Met	Met Val Lys
1385	1390	1395

Arg Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr Pro Ser Asp Ser  
 1400 1405 1410

Leu Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu Thr Pro Leu Val  
 1415 1420 1425

Asp Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu Pro Ser Thr Trp  
 1430 1435 1440

Ile Glu Asn Lys Leu Tyr Gly Met Ser Asp Pro Asn Trp Pro Gly  
 1445 1450 1455

Glu Ser Pro Val Pro Leu Thr Arg Ala Asp Gly Thr Asn Thr Gly  
 1460 1465 1470

Phe Pro Arg Tyr Pro Asn Asp Ser Val Tyr Ala Asn Trp Met Leu  
 1475 1480 1485

Ser Pro Ser Ala Ala Lys Leu Met Asp Thr Phe Asp Ser  
 1490 1495 1500

<210> 7

<211> 4380

<212> DNA

<213> Artificial Sequence

<220><223> coding sequence of fusion peptide

<400> 7

atgatgagac aggcaccgac agcccgaaag accacaactc ggcgacccaa gccacgcgc 60  
 ccagccagta ctggggtggc tggggccagt agctccctgg gccctctgg ctcagcgtca 120  
 gcaggtgagc tgagcagcag tgagcccagc accccggctc agactccgct ggcagcacc 180  
  
 atcatccca cgccgtctct cacctctct ggagcagtcc ccccgcttcc ttccccatcc 240  
 aaggaggagg agggactaag ggctcagggtg cgggacctgg aggagaaact agagaccctg 300  
 agactgaaac gggcagaaga caaagcaaag ctaaaagagc tggagaaaca caaatccag 360  
 ctggagcagg tgcaggaatg gaagagcaaa atgcaggagc agcaggccga cctgcagcgg 420  
 cgctcaagg aggcgagaaa ggaagccaag gaggcgctgg aggcaaagga acgctatatg 480  
 gaggagatgg ctgatactgc tgatgccatt gagatggcca ctttggacaa ggagatggct 540  
 gaagagcggg ctgagtcctt gcagcaggag gtggaggcac tgaaggagcg ggtggacgag 600  
  
 ctcaactg acttagagat cctcaaggct gagattgaag agaagggctc agatggcgct 660

gcatccagtt atcagctcaa gcagcttgag gagcagaatg cccgcctgaa ggatgcctg 720  
 gtgaggatgc gggatctttc ttcctcagag aagcaggagc atgtgaagct ccagaagctc 780  
 atggaaaaga agaaccaaga gctggaagtt gtgaggcaac agcgggagcg tctgcaggag 840  
 gagctaagcc aggcagagag caccattgat gagctcaagg agcagggtgga tgctgctctg 900  
 ggtgctgagg agatggtgga gatgctgaca gatcggaacc tgaatctgga agagaaagtg 960  
 cgcgagtga gggagactgt gggagacttg gaagcgatga atgagatgaa cgatgagctg 1020  
  
 caggagaatg cacgtgagac agaactggag ctgcgggagc agctggacat ggcaggcgcg 1080  
 cgggttcgtg aggccagaa gcgtgtggag gcagcccagg agacggttgc agactaccag 1140  
 cagaccatca agaagtaccg ccagctgacc gccatctac aggatgtgaa tcgggaactg 1200  
 acaaaccagc aggaagcatc tgtggagagg caacagcagc cacctccaga gaccttgac 1260  
 ttcaaaaatca agtttgctga gactaaggcc catgccaagg caattgagat ggaattgagg 1320  
 cagatggagg tggcccaggc caatcgacac atgtccctgc tgacagcctt catgcctgac 1380  
 agcttccttc gggcagggtg ggaccatgac tgcgttctgg tgctgttgct catgcctcgt 1440  
  
 ctcatittga aggcagagct gatccggaag caggcccagg agaagtttga actaagtgag 1500  
 aactgttcag agcggcctgg gctgcgagga gctgctgggg agcaactcag ctttgctgct 1560  
 ggactgggtg actcgtgag cctgctgcag gccacgctac accgctatga gcatgcctc 1620  
 tctcagtga gtgtggatgt gtataagaaa gtgggcagcc tgtaccctga gatgagtgcc 1680  
 catgagcgt ccttgattt cctcattgaa ctgctgcaca aggatcagct ggatgagact 1740  
 gtcaatgtgg agcctctcac caagccatc aagtactatc agcatctga cagcatccac 1800  
 cttgccgaac agcctgagga ctgtactatg cagctggctg accacattaa gttcacgag 1860  
  
 agtgctctgg actgcatgag tgtggaggta ggacggctgc gtgccttctt gcagggtggg 1920  
 caggaggcta cagatattgc cctcctgctc cgggatctgg aaacttcag cagtgacatc 1980  
 cgccagttct gcaagaagat ccgaaggcga atgccaggga cagatgctcc tgggatccca 2040  
 gctgcaactg cctttggacc acaggtatct gacacgctcc tagactgcag gaaacacttg 2100  
 acgtgggtcg tgctgtgct gcaggagggtg gcagctgctg ctgcccagct cattgcccga 2160  
 ctggcagaga atgaggggct acttgtggct gctctggagg aactggcttt caaagcaagc 2220  
 gagcagatct atgggacccc ctccagcagc ccctatgagt gtctgcgcca gcatgcaac 2280  
  
 atcctcatca gtacatgaa caagctggcc acagccatgc aggaggggga gtatgatgca 2340  
 gagcgcccc ccagcaagcc tccaccggtt gaactgcggg ctgctgcctt tcgtgcagag 2400  
 atcacagatg ctgaaggcct gggtttgaag ctgcaagatc gagagacagt tattaaggag 2460  
 ttgaagaagt cactcaagat taagggagag gagctaagtg aggccaatgt gcggctgagc 2520

ctcctggaga agaagttgga cagtgtgcc aaggatgcag atgagcgcat cgagaaagt 2580  
cagactcggc tggaggagac ccaggcactg ctgcgaaaga aggagaaaga gtttgaggag 2640  
acaatggatg cactccaggc tgacatcgac cagctggagg cagagaaggc agaactaaag 2700

cagcgtctga acagccagtc caaacgcacg attgaggac tccggggccc tctccttca 2760  
ggcattgcta ctctggtctc tggcattgct ggtgaagaac agcagcgagg agccatccct 2820  
gggcaggctc cagggtctgt gccaggccca gggctggtga aggactcacc actgctgctt 2880  
cagcagatct ctgcatgag gctgcacatc tcccagctcc agcatgagaa cagcatcctc 2940  
aaggagcccc agatgaagc atccttggca tcctgcccc ctctgcatgt tgcaaagcta 3000  
tcccatgagg gcctggcag tgagttacca gctggagcgc tgtatcgtaa gaccagccag 3060  
ctgctggaga cattgaatca attgagcaca cacacgcacg tagtagacat cactcgcacc 3120

agccctgctg ccaagagccc gtcggcccaa ctatggagc aagtggctca gettaagtec 3180  
ctgagtgaca ccgtcgagaa gctcaaggat gaggtcctca aggagacagt atctcagcgc 3240  
cctggagcca cagtaccac tgactttgcc accttccctt catcagcctt cctcaggag 3300  
gatccaaagt gggaattccc tcggaagaac ttggttcttg gaaaaactct aggagaagc 3360  
gaatttggaa aagtggtaaa ggcaacggcc ttccatctga aaggcagagc agggtaacc 3420  
acggtggccc tgaagatgct gaaagagaac gcctccccga gtgagcttcg agacctgctg 3480  
tcagagtta acgtcctgaa gcaggtaac cacccatg tcatcaaatt gtatggggcc 3540

tgcagccagg atggcccgt cctctcatc gtggagtacg ccaaatacgg ctccctgcgg 3600  
ggcttctcc gcgagagccg caaagtgggg cctggctacc tgggcagtgg aggcagccgc 3660  
aactccagct ccctggacca cccgatgag cgggccctca ccatggcga cctcatctca 3720  
tttgctggc agatctcaca ggggatgcag tatctggccg agatgaagct cgttcatcgg 3780  
gacttggcag ccagaaacat cctggtagct gagggcgga agatgaagat ttcgatttc 3840  
ggcttgtccc gagatgttta tgaagaggat tcctacgtga agaggagcca gggtcggatt 3900  
ccagttaaat ggatggcaat tgaatccctt tttgatcata tctacaccac gcaaagtgat 3960

gtatggtctt ttggtgtcct gctgtgggag atcgtgacc taggggaaa ccctatcct 4020  
gggattctc ctgagcggt cttcaacctt ctgaagacc gccaccgat ggagaggcca 4080  
gacaactgca gcgaggagat gtaccgctg atgctgcaat gctggaagca ggagccggac 4140  
aaaagccgg tgtttcgga catcagcaa gacctggaga agatgatggt taagaggaga 4200  
gactacttg accctgcggc gtccactcca tctgactccc tgatttatga cgacgcctc 4260  
tcagaggagg agacaccgt ggtggactgt aataatgccc cctccctcg agccctcct 4320

tccacatgga ttgaaaacaa actctatggt agaatttccc atgcatttac tagattctag 4380

<210> 8

<211> 1459

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 8

Met Met Arg Gln Ala Pro Thr Ala Arg Lys Thr Thr Thr Arg Arg Pro

1 5 10 15

Lys Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser Ser Ser

20 25 30

Leu Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser Ser Glu

35 40 45

Pro Ser Thr Pro Ala Gln Thr Pro Leu Ala Ala Pro Ile Ile Pro Thr

50 55 60

Pro Val Leu Thr Ser Pro Gly Ala Val Pro Pro Leu Pro Ser Pro Ser

65 70 75 80

Lys Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu Glu Lys

85 90 95

Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys Leu Lys

100 105 110

Glu Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu Trp Lys

115 120 125

Ser Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu Lys Glu

130 135 140

Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg Tyr Met

145 150 155 160

Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr Leu Asp

165 170 175

Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu Val Glu

180 185 190

Ala Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu Ile Leu  
 195 200 205

Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser Ser Tyr  
 210 215 220

Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp Ala Leu  
 225 230 235 240

Val Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His Val Lys  
 245 250 255

Leu Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val Val Arg  
 260 265 270

Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu Ser Thr  
 275 280 285

Ile Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala Glu Glu  
 290 295 300

Met Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu Lys Val  
 305 310 315 320

Arg Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn Glu Met  
 325 330 335

Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu Leu Arg  
 340 345 350

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

Val Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr Ile Lys  
 370 375 380

Lys Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg Glu Leu  
 385 390 395 400

Thr Asn Gln Gln Glu Ala Ser Val Glu Arg Gln Gln Gln Pro Pro Pro  
 405 410 415

Glu Thr Phe Asp Phe Lys Ile Lys Phe Ala Glu Thr Lys Ala His Ala  
 420 425 430

Lys Ala Ile Glu Met Glu Leu Arg Gln Met Glu Val Ala Gln Ala Asn

435                                      440                                      445  
 Arg His Met Ser Leu Leu Thr Ala Phe Met Pro Asp Ser Phe Leu Arg  
 450                                      455                                      460  
 Pro Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met Pro Arg  
 465                                      470                                      475                                      480  
 Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu Lys Phe  
                                     485                                      490                                      495  
 Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly Ala Ala  
  
                                     500                                      505                                      510  
 Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu Ser Leu  
                                     515                                      520                                      525  
 Leu Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln Cys Ser  
                                     530                                      535                                      540  
 Val Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met Ser Ala  
 545                                      550                                      555                                      560  
 His Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys Asp Gln  
  
                                     565                                      570                                      575  
 Leu Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile Lys Tyr  
                                     580                                      585                                      590  
 Tyr Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu Asp Cys  
                                     595                                      600                                      605  
 Thr Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala Leu Asp  
                                     610                                      615                                      620  
 Cys Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln Gly Gly  
  
 625                                      630                                      635                                      640  
 Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr Ser  
                                     645                                      650                                      655  
 Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met Pro  
                                     660                                      665                                      670  
 Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro Gln  
                                     675                                      680                                      685

Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp Val Val

690 695 700

Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile Ala Pro

705 710 715 720

Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu Leu Ala

725 730 735

Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser Pro Tyr

740 745 750

Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met Asn Lys

755 760 765

Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg Pro Pro

770 775 780

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

785 790 795 800

Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg Glu Thr

805 810 815

Val Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu Glu Leu

820 825 830

Ser Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu Asp Ser

835 840 845

Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr Arg Leu

850 855 860

Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe Glu Glu

865 870 875 880

Thr Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala Glu Lys

885 890 895

Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg Thr Ile Glu

900 905 910

Gly Leu Arg Gly Pro Pro Pro Ser Gly Ile Ala Thr Leu Val Ser Gly

915 920 925

Ile Ala Gly Glu Glu Gln Gln Arg Gly Ala Ile Pro Gly Gln Ala Pro



Val Ile Lys Leu Tyr Gly Ala Cys Ser Gln Asp Gly Pro Leu Leu  
 1175 1180 1185

Leu Ile Val Glu Tyr Ala Lys Tyr Gly Ser Leu Arg Gly Phe Leu  
 1190 1195 1200

Arg Glu Ser Arg Lys Val Gly Pro Gly Tyr Leu Gly Ser Gly Gly  
 1205 1210 1215

Ser Arg Asn Ser Ser Ser Leu Asp His Pro Asp Glu Arg Ala Leu  
 1220 1225 1230

Thr Met Gly Asp Leu Ile Ser Phe Ala Trp Gln Ile Ser Gln Gly  
 1235 1240 1245

Met Gln Tyr Leu Ala Glu Met Lys Leu Val His Arg Asp Leu Ala  
 1250 1255 1260

Ala Arg Asn Ile Leu Val Ala Glu Gly Arg Lys Met Lys Ile Ser  
 1265 1270 1275

Asp Phe Gly Leu Ser Arg Asp Val Tyr Glu Glu Asp Ser Tyr Val  
 1280 1285 1290

Lys Arg Ser Gln Gly Arg Ile Pro Val Lys Trp Met Ala Ile Glu  
 1295 1300 1305

Ser Leu Phe Asp His Ile Tyr Thr Thr Gln Ser Asp Val Trp Ser  
 1310 1315 1320

Phe Gly Val Leu Leu Trp Glu Ile Val Thr Leu Gly Gly Asn Pro  
 1325 1330 1335

Tyr Pro Gly Ile Pro Pro Glu Arg Leu Phe Asn Leu Leu Lys Thr  
 1340 1345 1350

Gly His Arg Met Glu Arg Pro Asp Asn Cys Ser Glu Glu Met Tyr  
 1355 1360 1365

Arg Leu Met Leu Gln Cys Trp Lys Gln Glu Pro Asp Lys Arg Pro  
 1370 1375 1380

Val Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys Met Met Val Lys  
 1385 1390 1395

Arg Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr Pro Ser Asp Ser

1400                      1405                      1410  
 Leu Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu Thr Pro Leu Val

1415                      1420                      1425  
 Asp Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu Pro Ser Thr Trp

1430                      1435                      1440  
 Ile Glu Asn Lys Leu Tyr Gly Arg Ile Ser His Ala Phe Thr Arg

1445                      1450                      1455  
 Phe

<210> 9

<211> 4833

<212> DNA

<213> Artificial Sequence

<220><223> coding sequence of fusion peptide

<400> 9

atggcacaga gcaagaggca cgtgtacagc cggacgcca gcggcagcag gatgagtgcg          60  
 gaggcaagcg cccggcctct gcgggtgggc tcccgtgtag aggtgattgg aaaaggccac          120  
 cgaggcactg tggcctatgt tggagccaca ctgtttgcca ctggcaaatg gtaggcctg          180  
  
 attctggatg aagcaaaggg caaaaatgat ggaactgttc aaggcaggaa gtacttcact          240  
 tgtgatgaag ggcatggcat ctttgtgcmc cagtcccaga tccaggtatt tgaagatgga          300  
 gcagatacta cttcccaga gacacctgat tcttctgctt caaaagtcct caaaagagag          360  
 ggaactgata caactgcaaa gactagcaaa ctgcccacgc gccagccag tactgggggtg          420  
 gctggggcca gtagctcctt gggcccctct ggctcagcgt cagcaggtga gctgagcagc          480  
 agtgagccca gcaccccggc tcagactccg ctggcagcac ccatcatccc cacgccggtc          540  
 ctcacctctc ctggagcagt cccccgctt ccttcccct ccaaggagga ggagggacta          600  
  
 agggctcagg tgcgggacct ggaggagaaa ctagagacc tgagactgaa acgggcagaa          660  
 gacaaagcaa agctaaaaga gctggagaaa cacaaaatcc agctggagca ggtgcaggaa          720  
 tggaagagca aaatgcagga gcagcaggcc gacctgcagc ggcgctcaa ggagcgagaga          780  
 aaggaagcca aggaggcgtt ggaggcaaag gaacgctata tggaggagat ggctgatact          840  
 gctgatgcca ttgatagggc cactttggac aaggagatgg ctgaagagcg ggctgagtcc          900  
 ctgcagcagg aggtggaggc actgaaggag cgggtggacg agctcactac tgacttagag          960

atcctcaagg ctgagattga agagaagggc tcagatggcg ctgcatccag ttatcagctc 1020  
  
 aagcagcttg aggagcagaa tgcccgcctg aaggatgccc tggtaggat gcgggatctt 1080  
 tcttctcag agaagcagga gcatgtgaag ctccagaagc tcatggaaaa gaagaaccaa 1140  
 gagctggaag ttgtgagca acagcgggag cgtctgcagg aggagctaag ccaggcagag 1200  
 agcaccattg atgagctcaa ggagcaggtg gatgctgctc tgggtgctga ggagatggtg 1260  
 gagatgctga cagatcggaa cctgaatctg gaagagaaag tgcgcgagtt gagggagact 1320  
 gtgggagact tggaaagcat gaatgagatg aacgatgagc tgcaggagaa tgcacgtgag 1380  
 acagaactgg agctgcggga gcagctggac atggcaggcg cgcgggttcg tgaggcccag 1440  
  
 aagcgtgtgg aggcagccca ggagacggtt gcagactacc agcagacat caagaagtac 1500  
 cgccagctga ccgccatct acaggatgtg aatcggaac tgacaaacca gcaggaagca 1560  
 tctgtggaga ggcaacagca gccacctcca gagaccttg acttcaaat caagtttget 1620  
 gagactaagg cccatgcaa ggcaattgag atggaattga ggcagatgga ggtggcccag 1680  
 gccaatcgac acatgtcctt gctgacagcc ttcattgctg acagcttctt tcggccaggt 1740  
 ggggaccatg actgcgttct ggtgctgttg ctcatgctc gtctcattg caaggcagag 1800  
 ctgatccgga agcaggccca ggagaagttt gaactaagtg agaactgttc agagcgcct 1860  
  
 gggctgcgag gagctgctgg ggagcaactc agctttgctg ctggactggt gfactcgctg 1920  
 agcctgctgc aggccacgct acaccgctat gagcatgccc tctctcagtg cagtgtggat 1980  
 gtgtataaga aagtgggagc cctgtacct gagatgagtg cccatgagcg ctcttggat 2040  
 ttctcattg aactgctgca caaggatcag ctggatgaga ctgtcaatgt ggagcctctc 2100  
 accaaggcca tcaagtacta tcagcatctg tacagcatcc acctgcccga acagcctgag 2160  
 gactgacta tgcagctggc tgaccacatt aagttcacgc agagtgtctt ggactgcatg 2220  
 agtgtggagg taggacggct gcgtgccttc ttgcagggtg ggcaggaggc tacagatatt 2280  
  
 gccctcctgc tccgggatct ggaaacttca tgcagtgaca tccgccagtt ctgcaagaag 2340  
 atccgaagge gaatgccagg gacagatgct cctgggatcc cagctgcaact ggcccttggga 2400  
 ccacaggtat ctgacacgct cctagactgc aggaaacact tgactgggtt cgtggctgtg 2460  
 ctgcaggagg tggcagctgc tgctgcccag ctattgccc cactggcaga gaatgagggg 2520  
 ctacttgtgg ctgctctgga ggaactggct ttcaaagcaa gcgagcagat ctatgggacc 2580  
 ccctccagca gccctatga gtgtctgctc cagtcatgca acatcctcat cagtaccatg 2640  
 aacaagctgg ccacagccat gcaggagggg gattatgatg cagagcggcc ccccagcaag 2700

cctccaccgg ttgaactgcg ggctgctgcc cttcgtgcag agatcacaga tgctgaaggc 2760  
 ctgggtttga agctcgaaga tcgagagaca gttattaagg agttgaagaa gtcactcaag 2820  
 attaagggag aggagctaag tgaggccaat gtgcggctga gcctcctgga gaagaagttg 2880  
 gacagtgtg ccaaggatgc agatgagcgc atcgagaaag tccagactcg gctggaggag 2940  
 acccaggcac tgctgcgaaa gaaggagaaa gagtttgagg agacaatgga tgcactccag 3000  
 gctgacatcg accagctgga ggcagagaag gcagaactaa agcagcgtct gaacagccag 3060  
 tccaaacgca cgattgaggg actccggggc cctcctcctt caggcattgc tactctggtc 3120  
  
 tctggcattg ctggtggagc catccctggg caggctccag ggtctgtgcc aggcccaggg 3180  
 ctggtgaagg acicaccact gctgctttag cagatctctg ccatgaggct gcacatctcc 3240  
 cagctccagc atgagaacag catcctcaag ggagcccaga tgaaggcatc cttggcatcc 3300  
 ctgccccctc tgcatgttg aaagetatcc catgagggcc ctggcagtga gttaccagct 3360  
 ggagcgtgt atcgtaagac cagccagctg ctggagacat tgaatcaatt gagcacacac 3420  
 acgcacgtag tagacatcac tcgcaccagc cctgctgcca agagcccgtc ggcccaactt 3480  
 atggagcaag tggtcagct taagtccctg agtgacaccg tcgagaagct caaggatgag 3540  
  
 gtccctcaagg agacagtatc tcagcgcct ggagccacag taccactga ctttgccacc 3600  
 ttcccttcat cagccttct cagggaggat ccaaagtgga aattccctcg gaagaacttg 3660  
 gtctttgaa aaactctagg agaaggcga tttggaaaag tggccaaggc aacggccttc 3720  
 catctgaaag gcagagcagg gtacaccacg gtggccgtga agatgctgaa agagaacgcc 3780  
 tccccgagtg agcttcgaga cctgctgtca gatttcaac tctgaagca ggtcaaccac 3840  
 ccacatgtca tcaaattgta tggggcctgc agccaggatg gcccgtcct cctcatctg 3900  
 gagtacgcca aatcggctc cctgcggggc ttctccgag agagccgcaa agtggggcct 3960  
  
 ggctacctgg gcagtggagg cagccgcaac tccagctccc tggaccacce ggatgagcgg 4020  
 gccctacca tgggcgacct catctcattt gcctggcaga tctcacaggg gatgcagtat 4080  
 ctggccgaga tgaagctctg tcatcgggac ttggcagcca gaaacatcct ggtagctgag 4140  
 gggcgaaga tgaagatttc ggatttcggc ttgtcccag atgtttatga agaggattec 4200  
 tacgtgaaga ggagccaggg tcggattcca gttaaatgga tggcaattga atccctttt 4260  
 gatcatatct acaccacgca aagtgatgta tggctctttg gtgtcctgct gttggagatc 4320  
 gtgaccctag ggggaaacce ctatcctggg attcctcctg agcggctctt caaccttctg 4380  
  
 aagaccggcc accggatgga gaggccagac aactgcagcg aggagatgta ccgcctgatg 4440  
 ctgcaatgct ggaagcagga gccggacaaa aggccgtgt ttgcggacat cagcaaagac 4500  
 ctggagaaga tgatggttaa gaggagagac tacttggacc ttgcggcgtc cactccatct 4560

gactccctga tttatgacga cggcctctca gaggaggaga caccgctggt ggactgtaat 4620  
aatgcccccc tcctcgcagc cctcccttcc acatggattg aaaacaaact ctatggcatg 4680  
tcagaccgca actggcctgg agagagtccct gtaccactca cgagagctga tggcactaac 4740  
actgggtttc caagatatcc aaatgatagt gtatatgcta actggatgct ttcacctca 4800

gcggcaaaat taatggacac gtttgatagt taa 4833

<210> 10

<211> 1610

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 10

Met Ala Gln Ser Lys Arg His Val Tyr Ser Arg Thr Pro Ser Gly Ser

1 5 10 15

Arg Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg

20 25 30

Val Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly

35 40 45

Ala Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu

50 55 60

Ala Lys Gly Lys Asn Asp Gly Thr Val Gln Gly Arg Lys Tyr Phe Thr

65 70 75 80

Cys Asp Glu Gly His Gly Ile Phe Val Arg Gln Ser Gln Ile Gln Val

85 90 95

Phe Glu Asp Gly Ala Asp Thr Thr Ser Pro Glu Thr Pro Asp Ser Ser

100 105 110

Ala Ser Lys Val Leu Lys Arg Glu Gly Thr Asp Thr Thr Ala Lys Thr

115 120 125

Ser Lys Leu Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser

130 135 140

Ser Ser Leu Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser

145 150 155 160

Ser Glu Pro Ser Thr Pro Ala Gln Thr Pro Leu Ala Ala Pro Ile Ile  
 165 170 175  
 Pro Thr Pro Val Leu Thr Ser Pro Gly Ala Val Pro Pro Leu Pro Ser  
 180 185 190  
 Pro Ser Lys Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu  
 195 200 205  
 Glu Lys Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys  
 210 215 220  
 Leu Lys Glu Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu  
 225 230 235 240  
 Trp Lys Ser Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu  
 245 250 255  
 Lys Glu Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg  
 260 265 270  
 Tyr Met Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr  
 275 280 285  
 Leu Asp Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu  
 290 295 300  
 Val Glu Ala Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu  
 305 310 315 320  
 Ile Leu Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser  
 325 330 335  
 Ser Tyr Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp  
 340 345 350  
 Ala Leu Val Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His  
 355 360 365  
 Val Lys Leu Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val  
 370 375 380  
 Val Arg Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu  
 385 390 395 400  
 Ser Thr Ile Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala

405 410 415  
 Glu Glu Met Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu

420 425 430  
 Lys Val Arg Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn

435 440 445  
 Glu Met Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu

450 455 460  
 Leu Arg Glu Gln Leu Asp Met Ala Gly Ala Arg Val Arg Glu Ala Gln

465 470 475 480  
 Lys Arg Val Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr

485 490 495  
 Ile Lys Lys Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg

500 505 510  
 Glu Leu Thr Asn Gln Gln Glu Ala Ser Val Glu Arg Gln Gln Gln Pro

515 520 525  
 Pro Pro Glu Thr Phe Asp Phe Lys Ile Lys Phe Ala Glu Thr Lys Ala

530 535 540  
 His Ala Lys Ala Ile Glu Met Glu Leu Arg Gln Met Glu Val Ala Gln

545 550 555 560  
 Ala Asn Arg His Met Ser Leu Leu Thr Ala Phe Met Pro Asp Ser Phe

565 570 575  
 Leu Arg Pro Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met

580 585 590  
 Pro Arg Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu

595 600 605  
 Lys Phe Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly

610 615 620  
 Ala Ala Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu

625 630 635 640  
 Ser Leu Leu Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln

645 650 655

Cys Ser Val Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met  
 660 665 670  
 Ser Ala His Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys  
 675 680 685  
 Asp Gln Leu Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile  
 690 695 700  
 Lys Tyr Tyr Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu  
 705 710 715 720  
 Asp Cys Thr Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala  
 725 730 735  
 Leu Asp Cys Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln  
 740 745 750  
 Gly Gly Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu  
 755 760 765  
 Thr Ser Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg  
 770 775 780  
 Met Pro Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly  
 785 790 795 800  
 Pro Gln Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp  
 805 810 815  
 Val Val Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile  
 820 825 830  
 Ala Pro Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu  
 835 840 845  
 Leu Ala Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser  
 850 855 860  
 Pro Tyr Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met  
 865 870 875 880  
 Asn Lys Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg  
 885 890 895  
 Pro Pro Ser Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg



Val Asp Ile Thr Arg Thr Ser Pro Ala Ala Lys Ser Pro Ser Ala  
 1145 1150 1155

Gln Leu Met Glu Gln Val Ala Gln Leu Lys Ser Leu Ser Asp Thr  
 1160 1165 1170

Val Glu Lys Leu Lys Asp Glu Val Leu Lys Glu Thr Val Ser Gln  
 1175 1180 1185

Arg Pro Gly Ala Thr Val Pro Thr Asp Phe Ala Thr Phe Pro Ser  
 1190 1195 1200

Ser Ala Phe Leu Arg Glu Asp Pro Lys Trp Glu Phe Pro Arg Lys  
 1205 1210 1215

Asn Leu Val Leu Gly Lys Thr Leu Gly Glu Gly Glu Phe Gly Lys  
 1220 1225 1230

Val Val Lys Ala Thr Ala Phe His Leu Lys Gly Arg Ala Gly Tyr  
 1235 1240 1245

Thr Thr Val Ala Val Lys Met Leu Lys Glu Asn Ala Ser Pro Ser  
 1250 1255 1260

Glu Leu Arg Asp Leu Leu Ser Glu Phe Asn Val Leu Lys Gln Val  
 1265 1270 1275

Asn His Pro His Val Ile Lys Leu Tyr Gly Ala Cys Ser Gln Asp  
 1280 1285 1290

Gly Pro Leu Leu Leu Ile Val Glu Tyr Ala Lys Tyr Gly Ser Leu  
 1295 1300 1305

Arg Gly Phe Leu Arg Glu Ser Arg Lys Val Gly Pro Gly Tyr Leu  
 1310 1315 1320

Gly Ser Gly Gly Ser Arg Asn Ser Ser Ser Leu Asp His Pro Asp  
 1325 1330 1335

Glu Arg Ala Leu Thr Met Gly Asp Leu Ile Ser Phe Ala Trp Gln  
 1340 1345 1350

Ile Ser Gln Gly Met Gln Tyr Leu Ala Glu Met Lys Leu Val His  
 1355 1360 1365

Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Ala Glu Gly Arg Lys

1370	1375	1380
Met Lys Ile Ser Asp Phe Gly Leu Ser Arg Asp Val Tyr Glu Glu		
1385	1390	1395
Asp Ser Tyr Val Lys Arg Ser Gln Gly Arg Ile Pro Val Lys Trp		
1400	1405	1410
Met Ala Ile Glu Ser Leu Phe Asp His Ile Tyr Thr Thr Gln Ser		
1415	1420	1425
Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Val Thr Leu		
1430	1435	1440
Gly Gly Asn Pro Tyr Pro Gly Ile Pro Pro Glu Arg Leu Phe Asn		
1445	1450	1455
Leu Leu Lys Thr Gly His Arg Met Glu Arg Pro Asp Asn Cys Ser		
1460	1465	1470
Glu Glu Met Tyr Arg Leu Met Leu Gln Cys Trp Lys Gln Glu Pro		
1475	1480	1485
Asp Lys Arg Pro Val Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys		
1490	1495	1500
Met Met Val Lys Arg Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr		
1505	1510	1515
Pro Ser Asp Ser Leu Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu		
1520	1525	1530
Thr Pro Leu Val Asp Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu		
1535	1540	1545
Pro Ser Thr Trp Ile Glu Asn Lys Leu Tyr Gly Met Ser Asp Pro		
1550	1555	1560
Asn Trp Pro Gly Glu Ser Pro Val Pro Leu Thr Arg Ala Asp Gly		
1565	1570	1575
Thr Asn Thr Gly Phe Pro Arg Tyr Pro Asn Asp Ser Val Tyr Ala		
1580	1585	1590
Asn Trp Met Leu Ser Pro Ser Ala Ala Lys Leu Met Asp Thr Phe		
1595	1600	1605

Asp Ser

1610

<210> 11

<211> 4707

<212> DNA

<213> Artificial Sequence

<220><223> coding sequence of fusion peptide

<400> 11

```

atggcacaga gcaagaggca cgtgtacagc cggacgcca gcggcagcag gatgagtgcg      60
gaggcaagcg cccggcctct gcgggtgggc tcccgtag aggtgattgg aaaaggccac      120
cgaggcactg tggcctatgt tggagccaca ctgtttgcc actggcaaatg gtaggcctg      180
attctggatg aagcaaaggg caaaaatgat ggaactgttc aaggcaggaa gtacttcaact      240

tgtgatgaag ggcattggcat ctttgtgccc cagtcccaga tccaggtatt tgaagatgga      300
gcagatacta ctccccaga gacacctgat tcttctgctt caaaagtctt caaaagagag      360
ggaactgata caactgcaaa gactagcaaa ctgccacgc gccagccag tactggggtg      420
gctggggcca gtagctcctt gggcccctct ggctcagcgt cagcaggtga gctgagcagc      480
agtgagccca gcaccccgcc tcagactccg ctggcagcac ccatcatccc cacgccggtc      540
ctcacctctc ctggagcagt cccccgctt ccttcccac ccaaggagga ggagggacta      600
agggctcagg tgcgggacct ggaggagaaa ctagagacc tgagactgaa acgggcagaa      660

gacaaagcaa agctaaaaga gctggagaaa cacaaaatcc agctggagca ggtgcaggaa      720
tggaagagca aaatgcagga gcagcaggcc gacctgcagc ggcgcctcaa ggagcgaga      780
aaggaagcca aggagcgctt ggaggcaaag gaacgctata tggaggagat ggctgatact      840
gctgatgcca ttgagatggc cactttggac aaggagatgg ctgaagagcg ggctgagtcc      900
ctgcagcagg aggtggaggc actgaaggag cgggtggacg agctcactac tgacttagag      960
atcctcaagg ctgagattga agagaagggc tcagatggcg ctgcatccag ttatcagctc     1020
aagcagcttg aggagcagaa tgcccgcctg aaggatgccc tggtaggat gcgggatctt     1080

tcttcctcag agaagcagga gcatgtgaag ctccagaagc tcatggaaaa gaagaaccaa     1140
gagctggaag ttgtgaggca acagcgggag cgtctgcagg aggagctaag ccaggcagag     1200
agcaccattg atgagctcaa ggagcagggt gatgctgctc tgggtgctga ggagatggtg     1260
gagatgctga cacatcgga cctgaatctg gaagagaaag tgcgcgagtt gagggagact     1320
gtgggagact tgaagcgat gaatgagatg aacgatgagc tgcaggagaa tgcacgtgag     1380
acagaactgg agctgcggga gcagctggac atggcaggcg cgcgggttcg tgaggcccag     1440

```

aagcgtgtgg aggcagccca ggagacggtt gcagactacc agcagacat caagaagtac 1500  
  
cgccagctga ccgccatct acaggatgtg aatcggaac tgacaaacca gcaggaagca 1560  
tctgtggaga ggcaacagca gccacctcca gagaccttg acttcaaat caagtttget 1620  
gagactaagg cccatgcca ggcaattgag atggaattga ggcagatgga ggtggcccag 1680  
gccaatgcac acatgtccct gctgacagcc ttcattgctg acagcttctc tcggccaggt 1740  
ggggaccatg actgcgttct ggtgctgttg ctcatgcctc gtctcattg caaggcagag 1800  
ctgatccgga agcaggccca ggagaagttt gaactaagtg agaactgttc agagcggcct 1860  
gggctgcgag gagctgctgg ggagcaactc agctttgctg ctggactggt gtactcgtg 1920  
  
agcctgctgc aggccacgct acaccgctat gagcatgccc tctctcagt cagtgtggat 1980  
gigtataaga aagtgggcag cctgtacct gagatgagt cccatgagcg ctcttggat 2040  
ttctcattg aactgctgca caaggatcag ctggatgaga ctgtcaatg ggagcctctc 2100  
accaaggcca tcaagtacta tcagcatctg tacagcatcc acctgcca acagcctgag 2160  
gactgtacta tgcagctggc tgaccacatt aagtccacgc agagtctct ggactgcatg 2220  
agtgtggagg taggacggct gcgtgccttc ttgcagggtg ggcaggaggc tacagatatt 2280  
gccctctgc tccgggatct ggaaacttca tgcagtgaca tccgccagtt ctgcaagaag 2340  
  
atccgaaggc gaatgccagg gacagatgct cctgggatcc cagctgact ggcctttgga 2400  
ccacaggtat ctgacacgct cctagactgc aggaaacact tgactgggt cgtggctgtg 2460  
ctgcaggagg tggcagctgc tgctgcccag ctattgccc cactggcaga gaatgagggg 2520  
ctacttgtgg ctgctctgga ggaactggtt ttcaaagcaa gcgagcagat ctatgggacc 2580  
ccctccagca gccctatga gtgtctgctc cagtcatgca acatcctcat cagtaccatg 2640  
aacaagtgg ccacagccat gcaggagggg gattatgat cagagcggcc cccagcaag 2700  
cctccaccgg ttgaactgcg ggctgctgcc ctctgtgag agatcacaga tgctgaaggc 2760  
  
ctgggtttga agctcgaaga tcgagagaca gttattaagg agttgaagaa gtcactcaag 2820  
attaaggag aggagctaag tgaggccaat gtgcggtga gcctcctgga gaagaagttg 2880  
gacagtgtg ccaaggatgc agatgagcgc atcgagaaag tccagactcg gctggaggag 2940  
accagggcac tgctgcgaaa gaaggagaaa gattttgagg agacaatgga tgcactccag 3000  
gctgacatcg accagctgga ggcagagaag gcagaactaa agcagcgtct gaacagccag 3060  
tccaaacgca cgattgaggg actccgggct cctcctcctt caggcattgc tactctggtc 3120  
tctggcattg ctggtggagc catcctggg caggctccag ggtctgtgcc aggccaggg 3180

ctggtgaagg acicaccact gctgcttcag cagatctctg ccatgaggct gcacatctcc 3240  
 cagctccagc atgagaacac catcctcaag ggagcccaga tgaaggcatc cttggcatcc 3300  
 ctgccccctc tgcattgttc aaagetatcc catgagggcc ctggcagtga gttaccagct 3360  
 ggagcgtgt atcgtaagac cagccagctg ctggagacat tgaatcaatt gagcacacac 3420  
 acgcacgtag tagacatcac tcgcaccagc cctgctgcca agagcccctc ggcccaactt 3480  
 atggagcaag tggctcagct taagtccctg agtgacaccg tcgagaagct caaggatgag 3540  
 gtctcaagg agacagiatc tcagcgcctt ggagccacag taccactga ctttgccacc 3600

ttcccttcat cagccttctt cagggaggat ccaaagtggg aattcctcgc gaagaacttg 3660  
 gtctttggaa aaactctagg agaaggcga tttggaaaag tggccaagc aacggccttc 3720  
 catctgaaag gcagagcagg gtacaccacg gtggccgtga agatgctgaa agagaacgcc 3780  
 tccccgagtg agcttcgaga cctgctgtca gatttcaacg tctgaaagca ggtcaaccac 3840  
 ccacatgtca tcaaattgta tggggcctgc agccaggatg gcccgtcctt cctcatctg 3900  
 gactaccca aatcggctc cctgccccgc ttctccgcg agagccgcaa agtggggcct 3960  
 ggctacctgg gcagtggagg cagccgcaac tccagctccc tggaccaccc ggatgagcgg 4020

gccctacca tgggcgacct catctcattt gcctggcaga tctcacagg gatgcagtat 4080  
 ctggccgaga tgaagctcgt tcatcgggac ttggcagcca gaaacatcct gtagctgag 4140  
 gggcgaaga tgaagatttc ggatttcggc ttgtcccag atgtttatga agaggattcc 4200  
 tacgtgaaga ggagccagg tcgattcca gttaaatgga tggcaattga atccctttt 4260  
 gatcatatct acaccacgca aagtgatgta tggctctttg gtgtcctgct gtgggagatc 4320  
 gtgaccttag ggggaaaccc ctatcctggg attcctcctg agcggctctt caaccttctg 4380  
 aagaccggcc accggatgga gaggccagac aactgcagcg aggagatgta ccgctgatg 4440

ctgcaatgct ggaagcagga gccggacaaa aggccgtgt ttgcggacat cagcaaagac 4500  
 ctggagaaga tgatggttaa gaggagagac tacttggacc ttgcggcgtc cactccatct 4560  
 gactccctga tttatgacga cggcctctca gaggaggaga caccgtggt ggactgtaat 4620  
 aatgcccccc tcctegagc cctccttcc acatggattg aaaacaaact ctatggtaga 4680  
 atttccatg catttactag attctag 4707

<210> 12

<211> 1568

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 12

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

245 250 255  
 Lys Glu Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg

260 265 270  
 Tyr Met Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr

275 280 285  
 Leu Asp Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu

290 295 300  
 Val Glu Ala Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu

305 310 315 320  
 Ile Leu Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser

325 330 335  
 Ser Tyr Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp

340 345 350  
 Ala Leu Val Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His

355 360 365  
 Val Lys Leu Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val

370 375 380  
 Val Arg Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu

385 390 395 400  
 Ser Thr Ile Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala

405 410 415  
 Glu Glu Met Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu

420 425 430  
 Lys Val Arg Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn

435 440 445  
 Glu Met Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu

450 455 460  
 Leu Arg Glu Gln Leu Asp Met Ala Gly Ala Arg Val Arg Glu Ala Gln

465 470 475 480  
 Lys Arg Val Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr

485 490 495

Ile Lys Lys Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg  
 500 505 510

Glu Leu Thr Asn Gln Gln Glu Ala Ser Val Glu Arg Gln Gln Gln Pro  
 515 520 525

Pro Pro Glu Thr Phe Asp Phe Lys Ile Lys Phe Ala Glu Thr Lys Ala  
 530 535 540

His Ala Lys Ala Ile Glu Met Glu Leu Arg Gln Met Glu Val Ala Gln  
 545 550 555 560

Ala Asn Arg His Met Ser Leu Leu Thr Ala Phe Met Pro Asp Ser Phe  
 565 570 575

Leu Arg Pro Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met  
 580 585 590

Pro Arg Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu  
 595 600 605

Lys Phe Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly  
 610 615 620

Ala Ala Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu  
 625 630 635 640

Ser Leu Leu Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln  
 645 650 655

Cys Ser Val Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met  
 660 665 670

Ser Ala His Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys  
 675 680 685

Asp Gln Leu Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile  
 690 695 700

Lys Tyr Tyr Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu  
 705 710 715 720

Asp Cys Thr Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala  
 725 730 735

Leu Asp Cys Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln

740                      745                      750  
 Gly Gly Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu  
 755                      760                      765  
 Thr Ser Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg  
  
 770                      775                      780  
 Met Pro Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly  
 785                      790                      795                      800  
 Pro Gln Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp  
 805                      810                      815  
 Val Val Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile  
 820                      825                      830  
 Ala Pro Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu  
  
 835                      840                      845  
 Leu Ala Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser  
 850                      855                      860  
 Pro Tyr Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met  
 865                      870                      875                      880  
 Asn Lys Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg  
 885                      890                      895  
 Pro Pro Ser Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg  
  
 900                      905                      910  
 Ala Glu Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg  
 915                      920                      925  
 Glu Thr Val Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu  
 930                      935                      940  
 Glu Leu Ser Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu  
 945                      950                      955                      960  
 Asp Ser Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr  
  
 965                      970                      975  
 Arg Leu Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe  
 980                      985                      990

Glu Glu Thr Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala  
 995 1000 1005  
 Glu Lys Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg  
 1010 1015 1020  
 Thr Ile Glu Gly Leu Arg Gly Pro Pro Pro Ser Gly Ile Ala Thr  
 1025 1030 1035  
 Leu Val Ser Gly Ile Ala Gly Gly Ala Ile Pro Gly Gln Ala Pro  
 1040 1045 1050  
 Gly Ser Val Pro Gly Pro Gly Leu Val Lys Asp Ser Pro Leu Leu  
 1055 1060 1065  
 Leu Gln Gln Ile Ser Ala Met Arg Leu His Ile Ser Gln Leu Gln  
 1070 1075 1080  
 His Glu Asn Ser Ile Leu Lys Gly Ala Gln Met Lys Ala Ser Leu  
 1085 1090 1095  
 Ala Ser Leu Pro Pro Leu His Val Ala Lys Leu Ser His Glu Gly  
 1100 1105 1110  
 Pro Gly Ser Glu Leu Pro Ala Gly Ala Leu Tyr Arg Lys Thr Ser  
 1115 1120 1125  
 Gln Leu Leu Glu Thr Leu Asn Gln Leu Ser Thr His Thr His Val  
 1130 1135 1140  
 Val Asp Ile Thr Arg Thr Ser Pro Ala Ala Lys Ser Pro Ser Ala  
 1145 1150 1155  
 Gln Leu Met Glu Gln Val Ala Gln Leu Lys Ser Leu Ser Asp Thr  
 1160 1165 1170  
 Val Glu Lys Leu Lys Asp Glu Val Leu Lys Glu Thr Val Ser Gln  
 1175 1180 1185  
 Arg Pro Gly Ala Thr Val Pro Thr Asp Phe Ala Thr Phe Pro Ser  
 1190 1195 1200  
 Ser Ala Phe Leu Arg Glu Asp Pro Lys Trp Glu Phe Pro Arg Lys  
 1205 1210 1215  
 Asn Leu Val Leu Gly Lys Thr Leu Gly Glu Gly Glu Phe Gly Lys

1220                      1225                      1230  
 Val Val Lys Ala Thr Ala Phe His Leu Lys Gly Arg Ala Gly Tyr  
 1235                      1240                      1245  
 Thr Thr Val Ala Val Lys Met Leu Lys Glu Asn Ala Ser Pro Ser  
 1250                      1255                      1260  
 Glu Leu Arg Asp Leu Leu Ser Glu Phe Asn Val Leu Lys Gln Val  
  
 1265                      1270                      1275  
 Asn His Pro His Val Ile Lys Leu Tyr Gly Ala Cys Ser Gln Asp  
 1280                      1285                      1290  
 Gly Pro Leu Leu Leu Ile Val Glu Tyr Ala Lys Tyr Gly Ser Leu  
 1295                      1300                      1305  
 Arg Gly Phe Leu Arg Glu Ser Arg Lys Val Gly Pro Gly Tyr Leu  
 1310                      1315                      1320  
 Gly Ser Gly Gly Ser Arg Asn Ser Ser Ser Leu Asp His Pro Asp  
  
 1325                      1330                      1335  
 Glu Arg Ala Leu Thr Met Gly Asp Leu Ile Ser Phe Ala Trp Gln  
 1340                      1345                      1350  
 Ile Ser Gln Gly Met Gln Tyr Leu Ala Glu Met Lys Leu Val His  
 1355                      1360                      1365  
 Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Ala Glu Gly Arg Lys  
 1370                      1375                      1380  
 Met Lys Ile Ser Asp Phe Gly Leu Ser Arg Asp Val Tyr Glu Glu  
  
 1385                      1390                      1395  
 Asp Ser Tyr Val Lys Arg Ser Gln Gly Arg Ile Pro Val Lys Trp  
 1400                      1405                      1410  
 Met Ala Ile Glu Ser Leu Phe Asp His Ile Tyr Thr Thr Gln Ser  
 1415                      1420                      1425  
 Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Val Thr Leu  
 1430                      1435                      1440  
 Gly Gly Asn Pro Tyr Pro Gly Ile Pro Pro Glu Arg Leu Phe Asn  
  
 1445                      1450                      1455

Leu Leu Lys Thr Gly His Arg Met Glu Arg Pro Asp Asn Cys Ser  
 1460 1465 1470  
 Glu Glu Met Tyr Arg Leu Met Leu Gln Cys Trp Lys Gln Glu Pro  
 1475 1480 1485  
 Asp Lys Arg Pro Val Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys  
 1490 1495 1500  
 Met Met Val Lys Arg Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr  
 1505 1510 1515  
 Pro Ser Asp Ser Leu Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu  
 1520 1525 1530  
 Thr Pro Leu Val Asp Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu  
 1535 1540 1545  
 Pro Ser Thr Trp Ile Glu Asn Lys Leu Tyr Gly Arg Ile Ser His  
 1550 1555 1560  
 Ala Phe Thr Arg Phe  
 1565

<210> 13

<211> 4491

<212> DNA

<213> Artificial Sequence

<220><223> coding sequence of fusion peptide

<400> 13

atgatgagac aggcaccgac agcccgaaag accacaactc ggcgacccaa gcccacgcgc 60  
 ccagccagta ctggggtggc tggggccagt agtcctctgg gccctctgg ctcagcgta 120  
 gcaggtgagc tgagcagcag tgagcccagc accccggctc agactccgct ggcagcaccc 180  
 atcatcccca cgccggtcct cacctctcct ggagcagtcc ccccgttcc ttcccatecc 240  
 aaggaggagg agggactaag ggctcaggtg cgggacctgg aggagaaact agagaccctg 300  
 agactgaaac gggcagaaga caaagcaaag ctaaaagagc tggagaaaca caaatccag 360  
 ctggagcagg tgcaggaatg gaagagcaaa atgcaggagc agcaggccga cctgcagcgg 420  
 cgctcaagg aggcgagaaa ggaagccaag gaggcgctgg aggcaaagga acgctatatg 480  
 gaggagatgg ctgatactgc tgatgccatt gagatggcca ctttgacaa ggagatggct 540  
 gaagagcggg ctgagtcctt gcagcaggag gtggaggcac tgaaggagcg ggtggacgag 600

ctcactactg acttagagat cctcaaggct gagattgaag agaagggctc agatggcgct 660  
 gcatccagtt atcagctcaa gcagcttgag gagcagaatg cccgcctgaa ggatgccctg 720  
 gtgaggatgc gggatctttc ttccctcagag aagcaggagc atgtgaagct ccagaagctc 780

atgaaaaaga agaaccaaga gctggaagtt gtgaggcaac agcgggagcg tctgcaggag 840  
 gagctaaacc aggcagagag caccattgat gagctcaagg agcaggtgga tgctgctctg 900  
 ggtgctgagg agatggtgga gatgctgaca gatcggaacc tgaatctgga agagaaagtg 960  
 cgcgagttga gggagactgt gggagacttg gaagcgatga atgagatgaa cgatgagctg 1020  
 caggagaatg cacgtgagac agaactggag ctgctgggagc agctggacat ggcaggcgcg 1080  
 cgggttcgtg aggccagaa gcgtgtggag gcagcccagg agacggttgc agactaccag 1140  
 cagaccatca agaagtaccg ccagctgacc gcccatctac aggatgtgaa tcgggaactg 1200

acaaaccagc aggaagcatc tgtggagagg caacagcagc cacctccaga gacctttgac 1260  
 ttcaaaatca agtttgctga gactaaggcc catgccaagg caattgagat ggaattgagg 1320  
 cagatggagg tggcccaggc caatcgacac atgtccctgc tgacagcctt catgcctgac 1380  
 agcttccctc ggccaggttg ggaccatgac tgcgttctgg tgctgttctt catgcctcgt 1440  
 ctcaattgca aggcagagct gatccggaag caggcccagg agaagtttga actaagtgag 1500  
 aactgttcag agcggcctgg gctgcgagga gctgctgggg agcaactcag ctttctgct 1560  
 ggactggtgt actcgtgag cctgctgcag gccacgctac accgctatga gcatgcctc 1620

tctcagtgca gtgtggaagt gtataagaaa gtgggcagcc tgtaccctga gatgagtgcc 1680  
 catgagcgct ccttgattt cctcattgaa ctgctgcaca aggatcagct ggatgagact 1740  
 gtcaatgtgg agcctctcac caaggccatc aagtactatc agcatctgta cagcatccac 1800  
 cttgccgaac agcctgagga ctgtactatg cagctggctg accacattaa gttcacgcag 1860  
 agtgctctgg actgcatgag tgtggaggta ggacggetgc gtgccttctt gcagggtggg 1920  
 caggaggcta cagatattgc cctcctgctc cgggatctgg aaacttcatg cagtgacatc 1980  
 cgccagttct gcaagaagat ccgaaggcga atgccaggga cagatgctcc tgggatccca 2040

gctgcaactg cctttggacc acaggtatct gacacgctcc tagactgcag gaaacacttg 2100  
 acgtgggtcg tggctgtgct gcaggaggtg gcagctgctg ctgccagct cattgcccc 2160  
 ctggcagaga atgaggggct acttgtggct gctctggagg aactggcttt caaagcaagc 2220  
 gagcagatct atgggacccc ctccagcagc ccctatgagt gtctgcgcca gtcacgcaac 2280  
 atcctcatca gtacatgaa caagctggcc acagccatgc aggaggggga gtatgatgca 2340  
 gagcggcccc ccagcaagc tccaccggtt gaactgcggg ctgctgcctt tcgtgcagag 2400

atcacagatg ctgaaggcct gggtttgaag ctggaagatc gagagacagt tattaaggag 2460  
  
 ttgaagaagt cactcaagat taaggagag gagctaagtg aggccaatgt gcggctgagc 2520  
 ctcttgaga agaagtggc cagtgtgcc aaggatgcag atgagcgcat cgagaaagt 2580  
 cagactcggc tggaggagac ccaggcactg ctgcgaaaga aggagaaaga gtttggagg 2640  
 acaatggatg cactccaggc tgacatcgac cagctggagg cagagaaggc agaactaaag 2700  
 cagcgtctga acagccagtc caaacgcacg attgagggac tccggggccc tectcttca 2760  
 ggcatgtcta ctctggctc tggcattgct ggtggagcca tccctgggca ggctccaggg 2820  
 tctgtgccag gccagggct ggtgaaggac tcaccactgc tgcttcagca gatctctgcc 2880  
  
 atgaggctgc acatctcca gctccagcat gagaacagca tcctcaagg agcccagatg 2940  
 aaggcatcct tggcatcctt gcccctctg catgttgcaa agctatcca tgaggccct 3000  
 ggcatgtagt taccagctgg agcgtgtat cgtaaacca gccagctgct ggagacattg 3060  
 aatcaattga gcacacacac gcacgtagta gacatcactc gcaccagccc tgctccaag 3120  
 agcccgtcgg cccaacttat ggagcaagtg gctcagctta agtcctgag tgacaccgtc 3180  
 gagaagctca aggatgaggt cctcaaggag acagtatctc agcgcctgg agccacagta 3240  
 cccactgact ttgccacctt ccttcatca gccttctca gggaggatcc aaagtgggaa 3300  
  
 ttccctcgga agaacttgg tcttgaaaa actctaggag aaggcgaatt tggaaaagtg 3360  
 gtcaaggcaa cggccttcca tctgaaaggc agagcagggt acaccacggt ggccgtgaag 3420  
 atgctgaaag agaacgcctc cccagtgag ctctgagacc tgctgtcaga gttcaacgtc 3480  
 ctgaagcagg tcaaccacc acatgtcatc aaattgtatg gggcctgcag ccaggatggc 3540  
 ccgctctcc tcatcgtgga gtacgcaaaa tacggctccc tgcggggctt cctccgcgag 3600  
 agccgcaaag tggggctcgg ctacctgggc agtggaggca gccgcaactc cagctcctg 3660  
 gaccaccgg atgagcgggc cctcaccatg ggcgacctca tctcatttgc ctggcagatc 3720  
  
 tcacagggga tgcagtatct ggccgagatg aagctcgttc atcgggactt ggcagccaga 3780  
 aacatcctgg tagctgaggg gcggaagatg aagatttcgg atttcggtt gtccgagat 3840  
 gtttatgaag aggattccta cgtgaagagg agccagggtc ggattccagt taaatgatg 3900  
 gcaattgaat cccttttga tcatatctac accacgcaaa gtgatgtatg gtcttttgg 3960  
 gtctgtctgt gggagatcgt gacctaggg ggaaccctt atcctgggat tectctgag 4020  
 cggctcttca accttctgaa gaccggccac cggatggaga ggccagacaa ctgcagcgag 4080  
 gagatgtacc gcctgatgct gcaatgctgg aagcaggagc cggacaaaag gccggtgttt 4140

gcggacatca gcaaagacct ggagaagatg atggttaaga ggagagacta cttggacctt 4200  
 gcggcgtcca ctccatctga ctccctgatt tatgacgacg gcctctcaga ggaggagaca 4260  
 ccgctggtgg actgtaataa tgccccctc cctcgagccc tcctttccac atggattgaa 4320  
 aacaaactct atggcatgtc agaccggaac tggcctggag agagtcctgt accactcacg 4380  
 agagctgatg gcactaacac tgggtttcca agatatcaa atgatagtgt atatgctaac 4440  
 tggatgcttt caccctcagc ggcaaaatta atggacacgt ttgatagtta a 4491

<210> 14

<211> 1496

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 14

Met Met Arg Gln Ala Pro Thr Ala Arg Lys Thr Thr Thr Arg Arg Pro

1 5 10 15

Lys Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser Ser Ser

20 25 30

Leu Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser Ser Glu

35 40 45

Pro Ser Thr Pro Ala Gln Thr Pro Leu Ala Ala Pro Ile Ile Pro Thr

50 55 60

Pro Val Leu Thr Ser Pro Gly Ala Val Pro Pro Leu Pro Ser Pro Ser

65 70 75 80

Lys Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu Glu Lys

85 90 95

Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys Leu Lys

100 105 110

Glu Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu Trp Lys

115 120 125

Ser Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu Lys Glu

130 135 140

Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg Tyr Met

145 150 155 160

Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr Leu Asp  
 165 170 175  
 Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu Val Glu  
 180 185 190  
  
 Ala Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu Ile Leu  
 195 200 205  
 Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser Ser Tyr  
 210 215 220  
 Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp Ala Leu  
 225 230 235 240  
 Val Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His Val Lys  
 245 250 255  
  
 Leu Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val Val Arg  
 260 265 270  
 Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu Ser Thr  
 275 280 285  
 Ile Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala Glu Glu  
 290 295 300  
 Met Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu Lys Val  
 305 310 315 320  
  
 Arg Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn Glu Met  
 325 330 335  
 Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu Leu Arg  
 340 345 350  
 Glu Gln Leu Asp Met Ala Gly Ala Arg Val Arg Glu Ala Gln Lys Arg  
 355 360 365  
 Val Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr Ile Lys  
 370 375 380  
  
 Lys Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg Glu Leu  
 385 390 395 400  
 Thr Asn Gln Gln Glu Ala Ser Val Glu Arg Gln Gln Gln Pro Pro Pro



Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met Pro  
 660 665 670  
 Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro Gln  
 675 680 685  
 Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp Val Val  
 690 695 700  
  
 Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile Ala Pro  
 705 710 715 720  
 Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu Leu Ala  
 725 730 735  
 Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser Pro Tyr  
 740 745 750  
 Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met Asn Lys  
 755 760 765  
  
 Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg Pro Pro  
 770 775 780  
 Ser Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg Ala Glu  
 785 790 795 800  
 Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg Glu Thr  
 805 810 815  
 Val Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu Glu Leu  
 820 825 830  
  
 Ser Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu Asp Ser  
 835 840 845  
 Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr Arg Leu  
 850 855 860  
 Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe Glu Glu  
 865 870 875 880  
 Thr Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala Glu Lys  
 885 890 895  
  
 Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg Thr Ile Glu



Glu Asn Ala Ser Pro Ser Glu Leu Arg Asp Leu Leu Ser Glu Phe  
 1145 1150 1155  
 Asn Val Leu Lys Gln Val Asn His Pro His Val Ile Lys Leu Tyr  
 1160 1165 1170  
 Gly Ala Cys Ser Gln Asp Gly Pro Leu Leu Leu Ile Val Glu Tyr  
 1175 1180 1185  
 Ala Lys Tyr Gly Ser Leu Arg Gly Phe Leu Arg Glu Ser Arg Lys  
 1190 1195 1200  
  
 Val Gly Pro Gly Tyr Leu Gly Ser Gly Gly Ser Arg Asn Ser Ser  
 1205 1210 1215  
 Ser Leu Asp His Pro Asp Glu Arg Ala Leu Thr Met Gly Asp Leu  
 1220 1225 1230  
 Ile Ser Phe Ala Trp Gln Ile Ser Gln Gly Met Gln Tyr Leu Ala  
 1235 1240 1245  
 Glu Met Lys Leu Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu  
 1250 1255 1260  
  
 Val Ala Glu Gly Arg Lys Met Lys Ile Ser Asp Phe Gly Leu Ser  
 1265 1270 1275  
 Arg Asp Val Tyr Glu Glu Asp Ser Tyr Val Lys Arg Ser Gln Gly  
 1280 1285 1290  
 Arg Ile Pro Val Lys Trp Met Ala Ile Glu Ser Leu Phe Asp His  
 1295 1300 1305  
 Ile Tyr Thr Thr Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu  
 1310 1315 1320  
  
 Trp Glu Ile Val Thr Leu Gly Gly Asn Pro Tyr Pro Gly Ile Pro  
 1325 1330 1335  
 Pro Glu Arg Leu Phe Asn Leu Leu Lys Thr Gly His Arg Met Glu  
 1340 1345 1350  
 Arg Pro Asp Asn Cys Ser Glu Glu Met Tyr Arg Leu Met Leu Gln  
 1355 1360 1365  
 Cys Trp Lys Gln Glu Pro Asp Lys Arg Pro Val Phe Ala Asp Ile

1370	1375	1380	
Ser Lys Asp Leu Glu Lys Met Met Val Lys Arg Arg Asp Tyr Leu			
1385	1390	1395	
Asp Leu Ala Ala Ser Thr Pro Ser Asp Ser Leu Ile Tyr Asp Asp			
1400	1405	1410	
Gly Leu Ser Glu Glu Glu Thr Pro Leu Val Asp Cys Asn Asn Ala			
1415	1420	1425	
Pro Leu Pro Arg Ala Leu Pro Ser Thr Trp Ile Glu Asn Lys Leu			
1430	1435	1440	
Tyr Gly Met Ser Asp Pro Asn Trp Pro Gly Glu Ser Pro Val Pro			
1445	1450	1455	
Leu Thr Arg Ala Asp Gly Thr Asn Thr Gly Phe Pro Arg Tyr Pro			
1460	1465	1470	
Asn Asp Ser Val Tyr Ala Asn Trp Met Leu Ser Pro Ser Ala Ala			
1475	1480	1485	
Lys Leu Met Asp Thr Phe Asp Ser			
1490	1495		
<210> 15			
<211> 4365			
<212> DNA			
<213> Artificial Sequence			
<220><223> coding sequence of fusion peptide			
<400> 15			
atgatgagac aggcaccgac agcccgaaag accacaactc ggcgacccaa gcccacgcgc	60		
ccagccagta ctgggggtggc tggggccagt agtcacctgg gccctctgg ctcagcgtea	120		
gcaggtgagc tgagcagcag tgagcccagc accccggctc agactccgct ggcagcacc	180		
atcatcccca cggcgttct cactctctct ggagcagtcc ccccgcttc tccccatec	240		
aaggaggagg agggactaag ggctcaggtg cgggacctgg aggagaaact agagacctg	300		
agactgaaac gggcagaaga caaagcaaag ctaaaagagc tggagaaaca caaatccag	360		
ctggagcagg tgcaggaatg gaagagcaaa atgcaggagc agcaggccga cctgcagcgg	420		
cgctcaagg aggcgagaaa ggaagccaag gaggcgctgg aggcaaagga acgctatatg	480		

gagagatgg ctgatactgc tgatgccatt gagatggcca ctttggacaa ggagatggct 540  
 gaagagcggg ctgagtcctt gcagcaggag gtggaggcac tgaaggagcg ggtggacgag 600  
 ctcaactactg acttagagat cctcaaggct gagattgaag agaagggctc agatggcgcct 660  
 gcatccagtt atcagctcaa gcagcttgag gaggcagaatg cccgcctgaa ggatgccttg 720  
 gtgaggatgc gggatctttc ttcctcagag aagcaggagc atgtgaagct ccagaagctc 780  
  
 atggaaaaga agaaccaaga gctggaagtt gtgaggcaac agcgggagcg tetgcaggag 840  
 gagctaagcc aggcagagag caccattgat gagctcaagg agcagggtgga tgctgctctg 900  
 ggtgctgagg agatggtgga gatgctgaca gatcggaacc tgaatctgga agagaaagtg 960  
 cgcgagtta gggagactgt gggagacttg gaagcgatga atgagatgaa cgatgagctg 1020  
 caggagaatg cacgtgagac agaactggag ctgcgggagc agctggacat ggcaggcgcg 1080  
 cgggttcgtg aggccagaa gcgtgtggag gcagcccagg agacggttgc agactaccag 1140  
 cagaccatca agaagtaccg ccagctgacc gcccatctac aggatgtgaa tcgggaactg 1200  
  
 acaaaccagc aggaagcatc tgtggagagg caacagcagc cacctccaga gaccttgac 1260  
 ttcaaaaatc agtttctga gactaaggcc catgccaagg caattgagat ggaattgagg 1320  
 cagatggagg tggcccaggc caatcgacac atgtccctgc tgacagcctt catgcctgac 1380  
 agcttccttc ggcaggtgg ggaccatgac tgcgttctgg tgctgttctt catgcctcgt 1440  
 ctcatattga aggcagagct gatccggaag caggcccagg agaagtttga actaagtgag 1500  
 aactgttcag agcggcctgg gctgcgagga gctgctgggg agcaactcag ctttgctgct 1560  
 ggactgggtg actcgtgag cctgctgag gccacgctac accgctatga gcatgcctc 1620  
  
 tctcagtga gtgtggatgt gtataagaaa gtgggcagcc tgtaccctga gatgagtgcc 1680  
 catgagcgt ccttggattt cctcattgaa ctgctgcaca aggatcagct ggatgagact 1740  
 gtcaatgtgg agcctctcac caaggccatc aagtactatc agcatctgta cagcatccac 1800  
 cttgccgaac agcctgagga ctgtactatg cagctggctg accacattaa gttcacgcag 1860  
 agtgctctgg actgcatgag tgtggaggta ggacggctgc gtgccttctt gcagggtggg 1920  
 caggaggcta cagatattgc cctcctgctc cgggatctgg aaacttcag cagtgcacac 1980  
 cgccagttct gcaagaagat ccgaaggcga atgccaggga cagatgctcc tgggatccca 2040  
  
 gctgcaactg cctttggacc acaggtatct gacacgctcc tagactgcag gaaacacttg 2100  
 acgtgggtcg tggctgtgct gcaggaggig gcagctgctg ctgcccagct cattgcccga 2160  
 ctggcagaga atgaggggct acttgtggct gctctggagg aactggcttt caaagcaagc 2220  
 gagcagatct atgggacccc ctccagcagc cctatgagt gtctgcgcca gtcacgcaac 2280  
 atcctcatca gtacatgaa caagctggcc acagccatgc aggaggggga gtatgatgca 2340

gagcgccccc ccagcaagcc tccaccggtt gaactgcggg ctgctgccct tctgtcagag 2400  
 atcacagatg ctgaaggcct gggtttgaag ctggaagatc gagagacagt tattaaggag 2460  
  
 ttgaagaagt cactcaagat taaggagag gagctaagtg aggccaatgt gcggctgagc 2520  
 ctctggaga agaagttaga cagtgtgcc aaggatgcag atgagcgcat cgagaaagt 2580  
 cagactcggc tggaggagac ccaggcactg ctgcgaaaga aggagaaaga gtttgaggag 2640  
 acaatggatg cactccaggc tgacatcgac cagctggagg cagagaaggc agaactaaag 2700  
 cagcgtctga acagccagtc caaacgcacg attgaggac tccggggccc tctccttca 2760  
 ggcatgtcta ctctggtctc tggcattgct ggtggagcca tccctgggca ggctccaggg 2820  
 tctgtgccag gccagggct ggtgaaggac tcaccactgc tgcttcagca gatctctgcc 2880  
  
 atgaggctgc acatctccca gctccagcat gagaacagca tctcaaggg agcccagatg 2940  
 aaggcatcct tggcatcctt gcccccctctg catgttgcaa agctatccca tgaggccct 3000  
 ggcagtgagt taccagctgg agcgtgtat cgtaagacca gccagctgct ggagacattg 3060  
 aatcaattga gcacacacac gcacgtagta gacatcactc gcaccagccc tgctgccaag 3120  
 agcccgtcgg cccaacttat ggagcaagtg gctcagetta agtccctgag tgacaccgtc 3180  
 gagaagctca aggatgaggt cctcaaggag acagtatctc agcgcctgg agccacagta 3240  
 cccactgact ttgccacctt ccttcatca gccttctca gggaggatcc aaagtgggaa 3300  
  
 ttccctcgga agaacttggg tcttgaaaa actctaggag aaggcgaatt tggaaaagtg 3360  
 gtcaaggcaa cggccttcca tctgaaaggc agagcagggt acaccacggt gccctgaag 3420  
 atgctgaaag agaacgcctc cccgagtgag cttcgagacc tgctgtcaga gttcaacgtc 3480  
 ctgaagcagg tcaaccacc acatgtcatc aaattgtatg gggcctgcag ccaggatggc 3540  
 ccgctctcc tcatctgga gtacgcaaaa tacggctccc tgcggggctt cctcccgag 3600  
 agccgcaaag tggggcctgg ctacctgggc agtggaggca gccgcaactc cagctccctg 3660  
 gaccaccggg atgagcgggc ctcacatg ggcgacctca tctcatttgc ctggcagatc 3720  
  
 tcacagggga tgcagtatct ggccgagatg aagctcgttc atcgggactt ggcagccaga 3780  
 aacatcctgg tagctgaggg gcggaagatg aagatttcgg atttcggctt gtcccagat 3840  
 gtttatgaag aggattccta cgtgaagagg agccagggtc ggattccagt taaatggatg 3900  
 gcaattgaat cctttttga tcatatctac accacgcaaa gtgatgtatg gcttttggg 3960  
 gtctgctgt gggagatcgt gaccctaggg ggaaccctc atcctgggat tctcctgag 4020  
 cggctcttca accttctgaa gaccggccac cggatggaga ggccagacaa ctgcagcgag 4080

gagatgtacc gcctgatgct gcaatgctgg aagcaggagc cggacaaaag gccgggtgttt 4140  
 gcggacatca gcaaagacct ggagaagatg atggttaaga ggagagacta cttggacctt 4200  
 gcggcgtcca ctccatctga ctccctgatt tatgacgacg gcctctcaga ggaggagaca 4260  
 ccgctggtgg actgtaataa tgccccctc cctcgagccc tcccttcac atggattgaa 4320  
 aacaaactct atggtagaat ttcccatgca ttactagat tctag 4365

<210> 16

<211> 1454

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 16

Met Met Arg Gln Ala Pro Thr Ala Arg Lys Thr Thr Thr Arg Arg Pro  
 1 5 10 15

Lys Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser Ser Ser  
 20 25 30

Leu Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser Ser Glu  
 35 40 45

Pro Ser Thr Pro Ala Gln Thr Pro Leu Ala Ala Pro Ile Ile Pro Thr  
 50 55 60

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

Lys Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu Glu Lys  
 85 90 95

Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys Leu Lys  
 100 105 110

Glu Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu Trp Lys  
 115 120 125

Ser Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu Lys Glu  
 130 135 140

Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg Tyr Met  
 145 150 155 160

Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr Leu Asp  
                           165                          170                          175  
 Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu Val Glu  
                           180                          185                          190  
 Ala Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu Ile Leu  
                           195                          200                          205  
  
 Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser Ser Tyr  
                           210                          215                          220  
 Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp Ala Leu  
 225                          230                          235                          240  
 Val Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His Val Lys  
                           245                          250                          255  
 Leu Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val Val Arg  
                           260                          265                          270  
  
 Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu Ser Thr  
                           275                          280                          285  
 Ile Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala Glu Glu  
                           290                          295                          300  
 Met Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu Lys Val  
 305                          310                          315                          320  
 Arg Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn Glu Met  
                           325                          330                          335  
  
 Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu Leu Arg  
                           340                          345                          350  
 Glu Gln Leu Asp Met Ala Gly Ala Arg Val Arg Glu Ala Gln Lys Arg  
                           355                          360                          365  
 Val Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr Ile Lys  
                           370                          375                          380  
 Lys Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg Glu Leu  
 385                          390                          395                          400  
  
 Thr Asn Gln Gln Glu Ala Ser Val Glu Arg Gln Gln Gln Pro Pro Pro

	405	410	415
Glu Thr Phe Asp Phe Lys Ile Lys Phe Ala Glu Thr Lys Ala His Ala			
	420	425	430
Lys Ala Ile Glu Met Glu Leu Arg Gln Met Glu Val Ala Gln Ala Asn			
	435	440	445
Arg His Met Ser Leu Leu Thr Ala Phe Met Pro Asp Ser Phe Leu Arg			
	450	455	460
Pro Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met Pro Arg			
465	470	475	480
Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu Lys Phe			
	485	490	495
Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly Ala Ala			
	500	505	510
Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu Ser Leu			
	515	520	525
Leu Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln Cys Ser			
	530	535	540
Val Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met Ser Ala			
545	550	555	560
His Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys Asp Gln			
	565	570	575
Leu Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile Lys Tyr			
	580	585	590
Tyr Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu Asp Cys			
	595	600	605
Thr Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala Leu Asp			
	610	615	620
Cys Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln Gly Gly			
625	630	635	640
Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr Ser			
	645	650	655

Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met Pro  
 660 665 670  
 Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro Gln  
 675 680 685  
 Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp Val Val  
 690 695 700  
 Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile Ala Pro  
 705 710 715 720  
  
 Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu Leu Ala  
 725 730 735  
 Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser Pro Tyr  
 740 745 750  
 Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met Asn Lys  
 755 760 765  
 Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg Pro Pro  
 770 775 780  
  
 Ser Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg Ala Glu  
 785 790 795 800  
 Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg Glu Thr  
 805 810 815  
 Val Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu Glu Leu  
 820 825 830  
 Ser Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu Asp Ser  
 835 840 845  
  
 Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr Arg Leu  
 850 855 860  
 Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe Glu Glu  
 865 870 875 880  
 Thr Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala Glu Lys  
 885 890 895  
 Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg Thr Ile Glu

900	905	910	
Gly Leu Arg Gly Pro Pro Pro Ser Gly Ile Ala Thr Leu Val Ser Gly			
915	920	925	
Ile Ala Gly Gly Ala Ile Pro Gly Gln Ala Pro Gly Ser Val Pro Gly			
930	935	940	
Pro Gly Leu Val Lys Asp Ser Pro Leu Leu Leu Gln Gln Ile Ser Ala			
945	950	955	960
Met Arg Leu His Ile Ser Gln Leu Gln His Glu Asn Ser Ile Leu Lys			
965	970	975	
Gly Ala Gln Met Lys Ala Ser Leu Ala Ser Leu Pro Pro Leu His Val			
980	985	990	
Ala Lys Leu Ser His Glu Gly Pro Gly Ser Glu Leu Pro Ala Gly Ala			
995	1000	1005	
Leu Tyr Arg Lys Thr Ser Gln Leu Leu Glu Thr Leu Asn Gln Leu			
1010	1015	1020	
Ser Thr His Thr His Val Val Asp Ile Thr Arg Thr Ser Pro Ala			
1025	1030	1035	
Ala Lys Ser Pro Ser Ala Gln Leu Met Glu Gln Val Ala Gln Leu			
1040	1045	1050	
Lys Ser Leu Ser Asp Thr Val Glu Lys Leu Lys Asp Glu Val Leu			
1055	1060	1065	
Lys Glu Thr Val Ser Gln Arg Pro Gly Ala Thr Val Pro Thr Asp			
1070	1075	1080	
Phe Ala Thr Phe Pro Ser Ser Ala Phe Leu Arg Glu Asp Pro Lys			
1085	1090	1095	
Trp Glu Phe Pro Arg Lys Asn Leu Val Leu Gly Lys Thr Leu Gly			
1100	1105	1110	
Glu Gly Glu Phe Gly Lys Val Val Lys Ala Thr Ala Phe His Leu			
1115	1120	1125	
Lys Gly Arg Ala Gly Tyr Thr Thr Val Ala Val Lys Met Leu Lys			
1130	1135	1140	

Glu Asn Ala Ser Pro Ser Glu Leu Arg Asp Leu Leu Ser Glu Phe  
 1145 1150 1155  
  
 Asn Val Leu Lys Gln Val Asn His Pro His Val Ile Lys Leu Tyr  
 1160 1165 1170  
 Gly Ala Cys Ser Gln Asp Gly Pro Leu Leu Leu Ile Val Glu Tyr  
 1175 1180 1185  
 Ala Lys Tyr Gly Ser Leu Arg Gly Phe Leu Arg Glu Ser Arg Lys  
 1190 1195 1200  
 Val Gly Pro Gly Tyr Leu Gly Ser Gly Gly Ser Arg Asn Ser Ser  
 1205 1210 1215  
  
 Ser Leu Asp His Pro Asp Glu Arg Ala Leu Thr Met Gly Asp Leu  
 1220 1225 1230  
 Ile Ser Phe Ala Trp Gln Ile Ser Gln Gly Met Gln Tyr Leu Ala  
 1235 1240 1245  
 Glu Met Lys Leu Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu  
 1250 1255 1260  
 Val Ala Glu Gly Arg Lys Met Lys Ile Ser Asp Phe Gly Leu Ser  
 1265 1270 1275  
  
 Arg Asp Val Tyr Glu Glu Asp Ser Tyr Val Lys Arg Ser Gln Gly  
 1280 1285 1290  
 Arg Ile Pro Val Lys Trp Met Ala Ile Glu Ser Leu Phe Asp His  
 1295 1300 1305  
 Ile Tyr Thr Thr Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu  
 1310 1315 1320  
 Trp Glu Ile Val Thr Leu Gly Gly Asn Pro Tyr Pro Gly Ile Pro  
 1325 1330 1335  
  
 Pro Glu Arg Leu Phe Asn Leu Leu Lys Thr Gly His Arg Met Glu  
 1340 1345 1350  
 Arg Pro Asp Asn Cys Ser Glu Glu Met Tyr Arg Leu Met Leu Gln  
 1355 1360 1365  
 Cys Trp Lys Gln Glu Pro Asp Lys Arg Pro Val Phe Ala Asp Ile



gctgagtccc tgcagcagga ggtggaggca ctgaaggagc gggtggacga gctcactact 900  
gacttagaga tcctcaagcg tgagattgaa gagaagggct cagatggcgc tgcattccagt 960  
tatcagctca agcagcttga ggagcagaat gcccgcctga aggatgccct ggtgaggatg 1020  
cgggatcttt cticctcaga gaagcaggag catgtgaagc tccagaagct catggaaaag 1080  
aagaaccaag agctggaagt tgtgaggcaa cagcgggagc gtctgcagga ggagctaagc 1140  
caggcagaga gcaccattga tgagctcaag gagcaggtgg atgctgctct gggtgctgag 1200  
  
gagatggtgg agatgctgac agatcggaac ctgaatctgg aagagaaagt gcgcgagttg 1260  
agggagactg tgggagactt ggaagcgatg aatgagatga acgatgagct gcaggagaat 1320  
gcacgtgaga cagaactgga gctgcgggag cagctggaca tggcaggcgc gcgggttcgt 1380  
gaggcccaga agcgtgtgga ggcagcccag gagacggttg cagactacca gcagaccatc 1440  
aagaagtacc gccagctgac gcacctata caggatgtga atcggaact gacaaaccag 1500  
caggaagcat ctgtggagag gcaacagcag ccacctccag agaccttga cttcaaatc 1560  
aagtttgctg agactaaggc ccatccaag gcaattgaga tggattgag gcagatggag 1620  
  
gtggcccagg ccaatcgaca catgtccctg ctgacagcct tcatgcctga cagcttcctt 1680  
cggccagggt gggaccatga ctgcttctg gtgctgttgc tcatgcctcg tctcatttgc 1740  
aaggcagagc tgatccgaa gcaggcccag gagaagttg aactaagtga gaactgttca 1800  
gagcggcctg ggctgcgagg agctgctggg gagcaactca gctttgctgc tggactggtg 1860  
tactcctga gcctgctgca ggccacgcta caccgctatg agcatgccct ctctcagtgc 1920  
agtgtggatg tgtataagaa agtgggcagc ctgtaccctg agatgagtgc ccatgagcgc 1980  
tccttggatt tcctcattga actgctgcac aaggatcagc tggatgagac tgtcaatgtg 2040  
  
gagcctctca ccaaggccat caagtactat cagcatctgt acagcatcca ccttgccgaa 2100  
cagcctgagg actgtaactat gcagetggct gaccacatta agttcacgca gagtgctctg 2160  
gactgcatga gtgtggaggt aggacggctg cgtgccttct tgcagggtgg gcaggaggct 2220  
acagatattg cctcctgct cgggatctg gaaacttcat gcagtgacat ccgccagttc 2280  
tgcaagaaga tccgaaggcg aatgccaggg acagatgctc ctgggatccc agctgcactg 2340  
gcctttggac cacaggtatc tgacacgctc ctgactgca ggaaacactt gacgtgggtc 2400  
gtggctgtgc tgcaggaggt ggcagctgct gctgccagc tcattgcccc actggcagag 2460  
  
aatgaggggc tacttgtggc tgctctggag gaactggctt tcaaagcaag cgagcagatc 2520  
tatgggacce cctccagcag ccctatgag tgtctgcgc agtcatgcaa catcctcacc 2580  
agtaccatga acaagctggc cacagccatg caggaggggg agtatgatgc agagcggccc 2640  
cccagcaagc ctccaccgt tgaactgcgg gctgctgccc ttcgtgcaga gatcacagat 2700

gctgaaggcc tgggtttgaa gctcgaagat cgagagacag ttattaagga gttgaagaag 2760  
tactcaaga ttaagggaga ggagctaagt gaggccaatg tgcggctgag cctcctggag 2820  
aagaagtgg acagtgctgc caaggatgca gatgagcgca tcgagaaaagt ccagactcgg 2880  
  
ctggaggaga cccaggcact gctcgaag aaggagaaag agtttgagga gacaatggat 2940  
gcactccagg ctgacatcga ccagctggag gcagagaagg cagaactaaa gcagcgtctg 3000  
aacagccagt ccaaacgcac gattgaggga ctccggggcc ctctctctc aggcattgct 3060  
actctggtct ctggcattgc tggtaggagc atccctgggc aggctccagg gtctgtgcca 3120  
ggcccagggc tggtagaaga ctaccactg ctgcttcagc agatctctgc catgaggctg 3180  
cacatctccc agctccagca tgagaacagc atcctcaagg gagcccagat gaaggcatcc 3240  
ttggcatccc tgccccctct gcatgttgca aagctatccc atgagggccc tggcagtgag 3300  
  
ttaccagctg gagcgtgta tcgtaagacc agccagctgc tggagacatt gaatcaattg 3360  
agcacacaca cgcacgtagt agacatcact cgcaccagcc ctgctgcaa gagcccgtcg 3420  
gcccactta tggagcaagt ggctcagctt aagtcctga gtgacaccgt cgagaagctc 3480  
aaggatgagg tctcaagga gacagtatct cagcgcctg gagccacagt acccactgac 3540  
tttgccacct tcccttcac agccttctc agggaggatc caaagtggga attcctcgg 3600  
aagaacttgg ttcttgaaa aactctagga gaaggcgaat ttgaaaaagt ggtcaaggca 3660  
acggccttc atctgaaagg cagagcaggg tacaccacgg tggcctgaa gatgctgaaa 3720  
  
gagaacgcct ccccgagiga gcttcgagac ctgctgtcag agttcaact cctgaagcag 3780  
gtcaaccacc cacatgtcat caaatgtat ggggcctgca gccaggatgg cccgtctctc 3840  
ctcatcgttg agiacgcaa atacggctcc ctgcgggct tctccgca gagccgcaaa 3900  
gtggggcctg gctacctggg cagtggagc agccgcaact ccagctcct ggaccaccg 3960  
gatgagcggg cctcaccat gggcgacct atctcattg cctggcagat ctcacagggg 4020  
atgcagtatc tggccgagat gaagctcgtt catcgggact tggcagccag aaacatcctg 4080  
gtagctgagg ggcggaagat gaagatttcg gatttcggct tgtcccgaga tgtttatgaa 4140  
  
gaggattcct acgtgaagag gagccagggt cggattccag ttaaatggat ggcaattgaa 4200  
tcccttttg atcatactca caccacgcaa agtgatgat ggtcttttg tgcctgctg 4260  
tgggagatcg tgacctagg gggaaacccc tatcctggga ttctctga gggctcttc 4320  
aaccttctga agaccggcca ccgatggag agccagaca actgcagcga ggagatgtac 4380  
gcctgatgc tgcaatgctg gaagcaggag ccggacaaaa ggccggtgtt tgcggacatc 4440  
agcaaagacc tggagaagat gatggttaag aggagagact acttggacct tgcggcgtcc 4500

actccatctg aciccctgat ttatgacgac ggcctctcag aggaggagac accgctggtg 4560  
 gactgtaata atgccccct cctcagagcc ctcccttcca catggattga aaacaaactc 4620  
 tatggcatgt cagaccggaa ctggcctgga gagagtctg taccactcac gagagetgat 4680  
 ggcactaaca ctgggtttcc aagatatcca aatgatagtg tataatgctaa ctggatgctt 4740  
 tcaccctcag cggcaaaatt aatggacacg tttgatagtt aa 4782

<210> 18

<211> 1593

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 18

Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg Val  
 1 5 10 15

Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly Ala  
 20 25 30

Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu Ala  
 35 40 45

Lys Gly Lys Asn Asp Gly Thr Val Gln Gly Arg Lys Tyr Phe Thr Cys  
 50 55 60

Asp Glu Gly His Gly Ile Phe Val Arg Gln Ser Gln Ile Gln Val Phe  
 65 70 75 80

Glu Asp Gly Ala Asp Thr Thr Ser Pro Glu Thr Pro Asp Ser Ser Ala  
 85 90 95

Ser Lys Val Leu Lys Arg Glu Gly Thr Asp Thr Thr Ala Lys Thr Ser  
 100 105 110

Lys Leu Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser Ser  
 115 120 125

Ser Leu Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser Ser  
 130 135 140

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

Thr Pro Val Leu Thr Ser Pro Gly Ala Val Pro Pro Leu Pro Ser Pro  
 165 170 175  
 Ser Lys Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu Glu  
 180 185 190  
 Lys Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys Leu  
 195 200 205  
  
 Lys Glu Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu Trp  
 210 215 220  
 Lys Ser Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu Lys  
 225 230 235 240  
 Glu Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg Tyr  
 245 250 255  
 Met Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr Leu  
 260 265 270  
  
 Asp Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu Val  
 275 280 285  
 Glu Ala Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu Ile  
 290 295 300  
 Leu Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser Ser  
 305 310 315 320  
 Tyr Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp Ala  
 325 330 335  
  
 Leu Val Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His Val  
 340 345 350  
 Lys Leu Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val Val  
 355 360 365  
 Arg Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu Ser  
 370 375 380  
 Thr Ile Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala Glu  
 385 390 395 400  
  
 Glu Met Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu Lys



Ala His Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys Asp  
660 665 670

Gln Leu Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile Lys  
675 680 685

Tyr Tyr Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu Asp  
690 695 700

Cys Thr Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala Leu  
705 710 715 720

Asp Cys Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln Gly  
725 730 735

Gly Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr  
740 745 750

Ser Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met  
755 760 765

Pro Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro  
770 775 780

Gln Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp Val  
785 790 795 800

Val Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile Ala  
805 810 815

Pro Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu Leu  
820 825 830

Ala Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser Pro  
835 840 845

Tyr Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met Asn  
850 855 860

Lys Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg Pro  
865 870 875 880

Pro Ser Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg Ala  
885 890 895

Glu Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg Glu

900	905	910	
Thr Val Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu Glu			
915	920	925	
Leu Ser Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu Asp			
930	935	940	
Ser Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr Arg			
945	950	955	960
Leu Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe Glu			
965	970	975	
Glu Thr Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala Glu			
980	985	990	
Lys Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg Thr Ile			
995	1000	1005	
Glu Gly Leu Arg Gly Pro Pro Pro Ser Gly Ile Ala Thr Leu Val			
1010	1015	1020	
Ser Gly Ile Ala Gly Gly Ala Ile Pro Gly Gln Ala Pro Gly Ser			
1025	1030	1035	
Val Pro Gly Pro Gly Leu Val Lys Asp Ser Pro Leu Leu Leu Gln			
1040	1045	1050	
Gln Ile Ser Ala Met Arg Leu His Ile Ser Gln Leu Gln His Glu			
1055	1060	1065	
Asn Ser Ile Leu Lys Gly Ala Gln Met Lys Ala Ser Leu Ala Ser			
1070	1075	1080	
Leu Pro Pro Leu His Val Ala Lys Leu Ser His Glu Gly Pro Gly			
1085	1090	1095	
Ser Glu Leu Pro Ala Gly Ala Leu Tyr Arg Lys Thr Ser Gln Leu			
1100	1105	1110	
Leu Glu Thr Leu Asn Gln Leu Ser Thr His Thr His Val Val Asp			
1115	1120	1125	
Ile Thr Arg Thr Ser Pro Ala Ala Lys Ser Pro Ser Ala Gln Leu			
1130	1135	1140	

Met Glu Gln Val Ala Gln Leu Lys Ser Leu Ser Asp Thr Val Glu  
 1145 1150 1155

Lys Leu Lys Asp Glu Val Leu Lys Glu Thr Val Ser Gln Arg Pro  
 1160 1165 1170

Gly Ala Thr Val Pro Thr Asp Phe Ala Thr Phe Pro Ser Ser Ala  
 1175 1180 1185

Phe Leu Arg Glu Asp Pro Lys Trp Glu Phe Pro Arg Lys Asn Leu  
 1190 1195 1200

Val Leu Gly Lys Thr Leu Gly Glu Gly Glu Phe Gly Lys Val Val  
 1205 1210 1215

Lys Ala Thr Ala Phe His Leu Lys Gly Arg Ala Gly Tyr Thr Thr  
 1220 1225 1230

Val Ala Val Lys Met Leu Lys Glu Asn Ala Ser Pro Ser Glu Leu  
 1235 1240 1245

Arg Asp Leu Leu Ser Glu Phe Asn Val Leu Lys Gln Val Asn His  
 1250 1255 1260

Pro His Val Ile Lys Leu Tyr Gly Ala Cys Ser Gln Asp Gly Pro  
 1265 1270 1275

Leu Leu Leu Ile Val Glu Tyr Ala Lys Tyr Gly Ser Leu Arg Gly  
 1280 1285 1290

Phe Leu Arg Glu Ser Arg Lys Val Gly Pro Gly Tyr Leu Gly Ser  
 1295 1300 1305

Gly Gly Ser Arg Asn Ser Ser Ser Leu Asp His Pro Asp Glu Arg  
 1310 1315 1320

Ala Leu Thr Met Gly Asp Leu Ile Ser Phe Ala Trp Gln Ile Ser  
 1325 1330 1335

Gln Gly Met Gln Tyr Leu Ala Glu Met Lys Leu Val His Arg Asp  
 1340 1345 1350

Leu Ala Ala Arg Asn Ile Leu Val Ala Glu Gly Arg Lys Met Lys  
 1355 1360 1365

Ile Ser Asp Phe Gly Leu Ser Arg Asp Val Tyr Glu Glu Asp Ser



<212> DNA

<213> Artificial Sequence

<220><223> coding sequence of fusion peptide

<400> 19

```

atgagtgccg aggcaagcgc ccggcctctg cgggtgggct cccgtgtaga ggtgattgga      60
aaaggccacc gaggcaactgt ggcctatgtt ggagccacac tgtttgccac tggcaaatgg      120
gtaggcgtga ttctggatga agcaaagggc aaaaatgatg gaactgttca aggcaggaag      180
tacttcactt gtgatgaagg gcatggcatc tttgtgcgcc agtcccagat ccaggtatTT      240

gaagatggag cagatactac ttccccagag acacctgatt cttctgcttc aaaagtcttc      300
aaaagagagg gaactgatac aactgcaaag actagcaaac tgcccacgcg cccagccagt      360
actgggggtg ctggggccag tagctccctg ggcccctctg gctcagcgtc agcaggtgag      420
ctgagcagca gtgagcccag caccccggct cagactccgc tggcagcacc catcatcccc      480
acgccgttcc tcacctctcc tggagcagtc cccccgttc cttccccatc caaggaggag      540
gagggactaa gggctcaggt gcgggacctg gaggagaaac tagagaccct gagactgaaa      600
cgggcagaag acaaagcaaa gctaaaagag ctggagaaac acaaaatcca gctggagcag      660

gtgcaggaat ggaagagcaa aatgcaggag cagcaggccg acctgcagcg ggcctcaag      720
gaggcgagaa aggaagccaa ggaggcgctg gaggcaaagg aacgctatat ggaggagatg      780
gctgatactg ctgatgccat tgagatggcc actttggaca aggagatggc tgaagagcgg      840
gctgagtccc tgcagcagga ggtggaggca ctgaaggagc ggggtggacga gctcactact      900
gacttagaga tcctcaagcg tgagattgaa gagaagggct cagatggcgc tgcattccagt      960
tatcagctca agcagcttga ggagcagaat gccgcctga aggatgccct ggtgaggatg     1020
cgggatcttt cttcctcaga gaagcaggag catgtgaagc tccagaagct catggaaaag     1080

aagaaccaag agctggaagt tgtgaggcaa cagcgggagc gtctgcagga ggagctaagc     1140
caggcagaga gcaccattga tgagctcaag gagcaggtgg atgctgctct ggggtgctgag     1200
gagatgggtg agatgctgac agatcggaac ctgaatctgg aagagaaagt gcgaggttg     1260
agggagactg tgggagactt ggaagcgatg aatgagatga acgatgagct gcaggagaat     1320
gcacgtgaga cagaactgga gctgcgggag cagctggaca tggcaggcgc ggggttcgt     1380
gaggcccaga agcgtgtgga ggcagcccag gagacggttg cagactacca gcagaccatc     1440
aagaagtacc gccagctgac cgccatcta caggatgtga atcggaact gacaaccag     1500

caggaagcat ctgtggagag gcaacagcag ccacctccag agaccttga cttcaaaatc     1560
aagtttgctg agactaaggc ccatccaag gcaattgaga tggaaattgag gcagatggag     1620

```

gtggcccagg ccaatcgaca catgtccctg ctgacagcct tcatgcctga cagcttcctt 1680  
 cggccagggt gggaccatga ctgcttctg gtgctgttgc tcatgcctcg tctcatttgc 1740  
 aaggcagagc tgatccgaa gcaggcccag gagaagtttg aactaagtga gaactgttca 1800  
 gagcggcctg ggctgcgagg agctgctggg gagcaactca gctttgctgc tggactggtg 1860  
 tactcgtga gcctgctgca ggccacgcta caccgctatg agcatgcct ctctcagtgc 1920

agtgtggatg tgtataagaa agtgggcagc ctgtaccctg agatgagtgc ccatgagcgc 1980  
 tccttggatt tcctcattga actgctgcac aaggatcagc tggatgagac tgtcaatgtg 2040  
 gagcctctca ccaaggccat caagtactat cagcatctgt acagcatcca ccttgccgaa 2100  
 cagcctgagg actgtactat gcagctggct gaccacatta agttcacgca gagtgtctctg 2160  
 gactgcatga gtgtggaggt aggacggctg cgtgccttct tgcagggtgg gcaggaggct 2220  
 acagatattg cctcctctgt ccgggatctg gaaacttcat gcagtgcacat ccgccagttc 2280  
 tgcaagaaga tccgaaggcg aatgccaggg acagatgctc ctgggatccc agctgcactg 2340

gcctttggac cacaggtatc tgacacgctc ctagactgca ggaaacactt gacgtgggtc 2400  
 gtggctgtgc tgcaggaggt ggcagctgct gctgcccagc tcattgcccc actggcagag 2460  
 aatgaggggc tacttgtggc tgctctggag gaactggctt tcaaagcaag cgagcagatc 2520  
 tatgggacce cctccagcag ccctatgag tgtctgcgcc agtcatgcaa catcctcacc 2580  
 agtaccatga acaagctggc cacagccatg caggaggggg agtatgatgc agageggccc 2640  
 cccagcaagc ctccaccgt tgaactgcgg gctgctgccc ttcgtgcaga gatcacagat 2700  
 gctgaaggcc tgggtttgaa gctcgaagat cgagagacag ttattaagga gttgaagaag 2760

tcactcaaga ttaagggaga ggagctaagt gaggccaatg tgcggctgag cctcctggag 2820  
 aagaagttag acagtgtctg caaggatgca gatgagcga tcgagaaagt ccagactcgg 2880  
 ctggaggaga cccaggcact gctgcgaaag aaggagaaag agtttgagga gacaatggat 2940  
 gcactccagg ctgacatcga ccagctggag gcagagaagg cagaactaaa gcagcgtctg 3000  
 aacagccagt ccaaacgcac gattgaggga ctccggggcc ctctccttc aggcattgct 3060  
 actctggctc ctggcattgc tggaggagcc atccctgggc aggetccagg gctctgcca 3120  
 ggcccagggc tggatgaagga ctaccactg ctgcttcagc agatctctgc catgaggctg 3180

cacatctccc agctccagca tgagaacagc atcctcaagg gagcccagat gaaggcatcc 3240  
 ttggcatccc tgccccctct gcatgttga aagctatccc atgagggccc tggcagtgag 3300  
 ttaccagctg gagcgtgta tcgtaagacc agccagctgc tggagacatt gaatcaattg 3360  
 agcacacaca cgcacgtagt agacatcact cgcaccagcc ctgctgcaa gagcccgtcg 3420  
 gcccaactta tggagcaagt ggctcagctt aagtcctga gtgacaccgt cgagaagctc 3480

aaggatgagg tcctcaagga gacagtatct cagcgccttg gagccacagt acccaactgac 3540  
 ttgtccacct tccttcacat agccttcctc agggaggatc caaagtggga attccctcgg 3600

aagaacttgg ttcttgaaa aactetagger gaaggcgaat ttgaaaagt ggtcaaggca 3660  
 acggccttcc atctgaaagg cagagcaggg tacaccacgg tggccgtgaa gatgctgaaa 3720  
 gagaacgct ccccgagtga gcttcgagac ctgctgtcag agttcaact cctgaagcag 3780  
 gtcaaccacc cacatgtcat caaattgtat ggggcctgca gccaggatgg cccgctcctc 3840  
 ctcatcgtgg agtacgcaa atacggctcc ctgcggggct tcctccgca gagccgcaaa 3900  
 gtggggcctg gctacctggg cagtggaggc agccgcaact ccagctcct ggaccaccg 3960  
 gatgagcggg ccctacat gggcgacct atctcattg cctggcagat ctcacagggg 4020

atgcagtatc tggccgagat gaagctcgtt catcgggact tggcagccag aaacatcctg 4080  
 gtagctgagg ggcggaagat gaagatttcg gatttcggct tgtcccgaga tgtttatgaa 4140  
 gaggattcct acgtgaagag gagccaggt cggattccag ttaaatggat ggcaattgaa 4200  
 tccctttttg atcatalcta caccacgcaa agtgatgtat ggtcttttgg tgcctgctg 4260  
 tgggagatcg tgacctagg gggaaacecc taccctggga ttcctcctga gcgctcttc 4320  
 aaccttctga agaccggcca ccggatggag aggccagaca actgcagcga ggagatgtac 4380  
 cgctgatgc tgcaatgctg gaagcaggag ccggacaaaa ggccggtgtt tgcggacatc 4440

agcaaagacc tggagaagat gatggttaag aggagagact acttggacct tgcggcgtcc 4500  
 actccatctg actccctgat ttatgacgac ggcctctcag aggaggagac accgctggtg 4560  
 gactgtaata atgccccct ccctcgagcc ctccttcca catggattga aaacaaactc 4620  
 tatggtagaa ttcccatgc atttactaga ttctag 4656

<210> 20

<211> 1551

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 20

Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg Val  
 1 5 10 15

Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly Ala  
 20 25 30

Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu Ala



Glu Ala Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu Ile  
 290 295 300  
 Leu Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser Ser  
 305 310 315 320  
 Tyr Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp Ala  
 325 330 335  
  
 Leu Val Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His Val  
 340 345 350  
 Lys Leu Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val Val  
 355 360 365  
 Arg Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu Ser  
 370 375 380  
 Thr Ile Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala Glu  
 385 390 395 400  
  
 Glu Met Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu Lys  
 405 410 415  
 Val Arg Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn Glu  
 420 425 430  
 Met Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu Leu  
 435 440 445  
 Arg Glu Gln Leu Asp Met Ala Gly Ala Arg Val Arg Glu Ala Gln Lys  
 450 455 460  
  
 Arg Val Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr Ile  
 465 470 475 480  
 Lys Lys Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg Glu  
 485 490 495  
 Leu Thr Asn Gln Gln Glu Ala Ser Val Glu Arg Gln Gln Gln Pro Pro  
 500 505 510  
 Pro Glu Thr Phe Asp Phe Lys Ile Lys Phe Ala Glu Thr Lys Ala His  
 515 520 525  
  
 Ala Lys Ala Ile Glu Met Glu Leu Arg Gln Met Glu Val Ala Gln Ala

530                    535                    540  
 Asn Arg His Met Ser Leu Leu Thr Ala Phe Met Pro Asp Ser Phe Leu  
 545                    550                    555                    560  
 Arg Pro Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met Pro  
                          565                    570                    575  
 Arg Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu Lys  
                          580                    585                    590  
  
 Phe Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly Ala  
                          595                    600                    605  
 Ala Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu Ser  
                          610                    615                    620  
 Leu Leu Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln Cys  
 625                    630                    635                    640  
 Ser Val Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met Ser  
                          645                    650                    655  
  
 Ala His Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys Asp  
                          660                    665                    670  
 Gln Leu Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile Lys  
                          675                    680                    685  
 Tyr Tyr Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu Asp  
                          690                    695                    700  
 Cys Thr Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala Leu  
 705                    710                    715                    720  
  
 Asp Cys Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln Gly  
                          725                    730                    735  
 Gly Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr  
                          740                    745                    750  
 Ser Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met  
                          755                    760                    765  
 Pro Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro  
                          770                    775                    780

Gln Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp Val  
 785                      790                      795                      800

Val Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile Ala  
                                  805                      810                      815

Pro Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu Leu  
                                  820                      825                      830

Ala Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser Pro  
                                  835                      840                      845

Tyr Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met Asn  
                                  850                      855                      860

Lys Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg Pro  
 865                      870                      875                      880

Pro Ser Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg Ala  
                                  885                      890                      895

Glu Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg Glu  
                                  900                      905                      910

Thr Val Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu Glu  
                                  915                      920                      925

Leu Ser Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu Asp  
                                  930                      935                      940

Ser Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr Arg  
 945                      950                      955                      960

Leu Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe Glu  
                                  965                      970                      975

Glu Thr Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala Glu  
                                  980                      985                      990

Lys Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg Thr Ile  
                                  995                      1000                      1005

Glu Gly Leu Arg Gly Pro Pro Pro Ser Gly Ile Ala Thr Leu Val  
                                  1010                      1015                      1020

Ser Gly Ile Ala Gly Gly Ala Ile Pro Gly Gln Ala Pro Gly Ser

1025	1030	1035
Val Pro Gly Pro Gly Leu Val	Lys Asp Ser Pro Leu	Leu Leu Gln
1040	1045	1050
Gln Ile Ser Ala Met Arg Leu	His Ile Ser Gln Leu	Gln His Glu
1055	1060	1065
Asn Ser Ile Leu Lys Gly Ala	Gln Met Lys Ala Ser	Leu Ala Ser
1070	1075	1080
Leu Pro Pro Leu His Val Ala	Lys Leu Ser His Glu	Gly Pro Gly
1085	1090	1095
Ser Glu Leu Pro Ala Gly Ala	Leu Tyr Arg Lys Thr	Ser Gln Leu
1100	1105	1110
Leu Glu Thr Leu Asn Gln Leu	Ser Thr His Thr His	Val Val Asp
1115	1120	1125
Ile Thr Arg Thr Ser Pro Ala	Ala Lys Ser Pro Ser	Ala Gln Leu
1130	1135	1140
Met Glu Gln Val Ala Gln Leu	Lys Ser Leu Ser Asp	Thr Val Glu
1145	1150	1155
Lys Leu Lys Asp Glu Val Leu	Lys Glu Thr Val Ser	Gln Arg Pro
1160	1165	1170
Gly Ala Thr Val Pro Thr Asp	Phe Ala Thr Phe Pro	Ser Ser Ala
1175	1180	1185
Phe Leu Arg Glu Asp Pro Lys	Trp Glu Phe Pro Arg	Lys Asn Leu
1190	1195	1200
Val Leu Gly Lys Thr Leu Gly	Glu Gly Glu Phe Gly	Lys Val Val
1205	1210	1215
Lys Ala Thr Ala Phe His Leu	Lys Gly Arg Ala Gly	Tyr Thr Thr
1220	1225	1230
Val Ala Val Lys Met Leu Lys	Glu Asn Ala Ser Pro	Ser Glu Leu
1235	1240	1245
Arg Asp Leu Leu Ser Glu Phe	Asn Val Leu Lys Gln	Val Asn His
1250	1255	1260

Pro His Val Ile Lys Leu Tyr Gly Ala Cys Ser Gln Asp Gly Pro  
 1265 1270 1275

Leu Leu Leu Ile Val Glu Tyr Ala Lys Tyr Gly Ser Leu Arg Gly  
 1280 1285 1290

Phe Leu Arg Glu Ser Arg Lys Val Gly Pro Gly Tyr Leu Gly Ser  
 1295 1300 1305

Gly Gly Ser Arg Asn Ser Ser Ser Leu Asp His Pro Asp Glu Arg  
 1310 1315 1320

Ala Leu Thr Met Gly Asp Leu Ile Ser Phe Ala Trp Gln Ile Ser  
 1325 1330 1335

Gln Gly Met Gln Tyr Leu Ala Glu Met Lys Leu Val His Arg Asp  
 1340 1345 1350

Leu Ala Ala Arg Asn Ile Leu Val Ala Glu Gly Arg Lys Met Lys  
 1355 1360 1365

Ile Ser Asp Phe Gly Leu Ser Arg Asp Val Tyr Glu Glu Asp Ser  
 1370 1375 1380

Tyr Val Lys Arg Ser Gln Gly Arg Ile Pro Val Lys Trp Met Ala  
 1385 1390 1395

Ile Glu Ser Leu Phe Asp His Ile Tyr Thr Thr Gln Ser Asp Val  
 1400 1405 1410

Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Val Thr Leu Gly Gly  
 1415 1420 1425

Asn Pro Tyr Pro Gly Ile Pro Pro Glu Arg Leu Phe Asn Leu Leu  
 1430 1435 1440

Lys Thr Gly His Arg Met Glu Arg Pro Asp Asn Cys Ser Glu Glu  
 1445 1450 1455

Met Tyr Arg Leu Met Leu Gln Cys Trp Lys Gln Glu Pro Asp Lys  
 1460 1465 1470

Arg Pro Val Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys Met Met  
 1475 1480 1485

Val Lys Arg Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr Pro Ser

1490                      1495                      1500  
 Asp Ser Leu Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu Thr Pro  
 1505                      1510                      1515  
  
 Leu Val Asp Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu Pro Ser  
 1520                      1525                      1530  
 Thr Trp Ile Glu Asn Lys Leu Tyr Gly Arg Ile Ser His Ala Phe  
 1535                      1540                      1545  
 Thr Arg Phe  
 1550  
 <210> 21  
 <211> 4887  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> coding sequence of fusion peptide  
 <400> 21  
 atggcacaga gcaagaggca cgtgtacagc cggacgccca gcggcagcag gatgagtgcg 60  
 gaggcaagcg cccggcctct gcgggtgggc tcccgtgtag aggtgattgg aaaaggccac 120  
  
 cgaggcactg tggcctatgt tggagccaca ctgtttgcca ctggcaaatg gtaggcctg 180  
 attctggatg aagcaaaggg caaaaatgat ggaactgttc aaggcaggaa gtacttcaact 240  
 tgtgatgaag ggcatggcat ctttgtgcmc cagtcccaga tccaggtatt tgaagatgga 300  
 gcagatacta ctccccaga gacacctgat tcttctgctt caaaagtcct caaaagagag 360  
 ggaactgata caactgcaaa gactagcaaa ctggcaccga cagcccgaaa gaccacaact 420  
 cggcgaccca agcccacgcg cccagccagt actggggtgg ctggggccag tagctcctg 480  
 ggcccctctg gctcagcgtc agcaggtgag ctgagcagca gtgagcccag caccggct 540  
  
 cagactccgc tggcagcacc cateatcccc acgccgtcc tcacctctcc tggagcagtc 600  
 cccccgttc ctccccatc caaggaggag gagggactaa gggctcaggt gcgggacctg 660  
 gaggagaaac tagagacctt gagactgaaa cgggcagaag acaaagcaaa gctaaaagag 720  
 ctggagaaac acaaaatcca gctggagcag gtgcaggaat ggaagagcaa aatgcaggag 780  
 cagcaggccg acctgcagcg gcgcctcaag gaggcgagaa aggaagccaa ggaggcgtg 840  
 gaggcaaagg aacgctatat ggaggagatg gctgatactg ctgatgcat tgagatggcc 900  
 actttggaca aggagatggc tgaagagcgg gctgagtccc tgcagcagga ggtggaggca 960

ctgaaggagc ggggtggacga gctcactact gacttagaga tcctcaaggc tgagattgaa 1020  
 gagaagggct catagggcgc tgcattccagt tatcagctca agcagcttga ggagcagaat 1080  
 gcccgcctga aggatgccct ggtgaggatg cgggatcttt cttcctcaga gaagcaggag 1140  
 catgtgaagc tccagaagct catggaaaag aagaaccaag agctggaagt tgtgaggcaa 1200  
 cagcgggagc gtctgcagga ggagctaagc caggcagaga gcaccattga tgagctcaag 1260  
 gagcaggtgg atgctgctct ggggtctgag gagatggtgg agatgctgac agatcggaac 1320  
 ctgaatctgg aagagaaagt gcgcgagttg agggagactg tgggagactt ggaagcgatg 1380  
  
 aatgagatga acgatgagct gcaggagaat gcacgtgaga cagaactgga gctgcgggag 1440  
 cagctggaca tggcaggcgc gcgggttcgt gaggcccaga agcgtgtgga ggcagcccag 1500  
 gagacggttg cagactacca gcagaccatc aagaagtacc gccagctgac cgcccatcta 1560  
 caggatgtga atcgggaact gacaaaccag caggaagcat ctgtggagag gcaacagcag 1620  
 ccacctccag agacctttga cttcaaaatc aagtttctg agactaaggc ccatgccaag 1680  
 gcaattgaga tggaaattgag gcagatggag gtggcccagg ccaatcgaca catgtcctg 1740  
 ctgacagcct tcatgcctga cagcttctct cggccagggtg gggaccatga ctgcgttctg 1800  
  
 gtgctgttgc tcatgcctcg tctcatttgc aaggcagagc tgatccgaa gcaggcccag 1860  
 gagaagtttg aactaagtga gaactgttca gagcggcctg ggctgcgagg agctgctggg 1920  
 gagcaactca gctttgctgc tggactggtg tactcgtga gcctgctgca ggccacgcta 1980  
 caccgctatg agcatgccct ctctcagtgc agtgtggatg tgtataagaa agtgggcagc 2040  
 ctgtaccctg agatgagtgc ccatgagcgc tccttggatt tectcattga actgctgcac 2100  
 aaggatcagc tggatgagac tgtcaatgtg gagcctctca ccaaggccat caagtactat 2160  
 cagcatctgt acagcatcca ccttgccgaa cagcctgagg actgtactat gcagctggct 2220  
  
 gaccacatta agttcacga gagtgcctctg gactgcatga gtgtggaggt aggacggctg 2280  
 cgtgccttct tgcagggtgg gcaggaggct acagatatg ccctcctgct ccgggatctg 2340  
 gaaacttcat gcagtgacat ccgccagttc tgcaagaaga tccgaaggcg aatgccaggg 2400  
 acagatgctc ctgggatccc agctgcactg gcctttggac cacaggtatc tgacacgctc 2460  
 ctagactgca ggaaacactt gacgtgggtc gtggctgtgc tgcaggaggt ggcagctgct 2520  
 gctgcccage tcattgcccc actggcagag aatgaggggc tacttgtggc tgctctggag 2580  
 gaactggctt tcaaagcaag cgagcagatc tatgggaccc cctccagcag ccctatgag 2640  
  
 tgtctgcgcc agtcatgcaa catcctcatc agtaccatga acaagctggc cacagccatg 2700  
 caggaggggg agtatgatgc agagcggccc ccagcaagc ctccaccgtg tgaactcggg 2760  
 gctgctgccc ttcgtgcaga gatcacagat gctgaaggcc tgggtttgaa gctcgaagat 2820

cgagagacag ttattaagga gttgaagaag tcaactcaaga ttaagggaga ggagctaagt 2880  
 gaggccaatg tgcggctgag cctcctggag aagaagtgg acagtgtgc caaggatgca 2940  
 gatgagcgca tcgagaaagt ccagactcgg ctggaggaga cccaggcact gctgcgaaag 3000  
 aaggagaaag agtttgagga gacaatggat gcactccagg ctgacatcga ccagctggag 3060  
  
 gcagagaagg cagaactaaa gcagcgtctg aacagccagt ccaaacgcac gattgagga 3120  
 ctccggggcc ctctccttc aggcattgct actctggtct ctggcattgc tggatgaaga 3180  
 cagcagcgag gagccatccc tgggcaggct ccagggtctg tgccaggccc agggctggtg 3240  
 aaggactcac cactgctgct tcagcagatc tctgcatga ggctgcacat ctcccagctc 3300  
 cagcatgaga acagcatcct caaggagacc cagatgaagg catccttggc atccctgccc 3360  
 cctctgcatg ttgcaaagct atcccatgag ggccttgca gtgagttacc agctggagcg 3420  
 ctgtatcgta agaccagcca gctgctggag acattgaaac aattgagcac acacacgcac 3480  
  
 glagtagaca tcaactgcac cagccctgct gccaagagcc cgtcggccca acttatggag 3540  
 caagtggctc agcttaagtc cctgagtac accgtcgaga agctcaagga tgaggtcctc 3600  
 aaggagacag tatctcagcg ccctggagcc acagtaccca ctgactttgc caccttccct 3660  
 tcatcagcct tcctcagga ggatccaaag tgggaattcc ctcggaagaa cttggttctt 3720  
 ggaaaaactc taggagaagg cgaatttga aaagtgtca aggcaacggc ctccatctg 3780  
 aaaggcagag cagggtacac cacggtggcc gtgaagatgc tgaaagagaa cgcctccccg 3840  
 agtgagcttc gagacctgct gtcagattc aacgtcctga agcaggtcaa ccaccacat 3900  
  
 gtcacaaat tgatggggc ctgcagccag gatggcccgc tcctcctcat cgtggagtac 3960  
 gccaaatagc gctccctcgc gggcttcctc cgcgagagcc gcaaagtggg gcctggctac 4020  
 ctgggcagtg gaggcagcc caactccagc tccttgacc acccggatga gcgggcctc 4080  
 accatggcg accatctc atttgcctgg cagatctcac aggggatgca gtatctgccc 4140  
 gagatgaagc tcgttcatcg ggacttgca gccagaaaca tcctggtagc tgagggcg 4200  
 aagatgaaga tttcgattt cggtctgtcc cgagatgtt atgaagagga ttctactgt 4260  
 aagaggagcc aggtcggat tccagttaa tggatggcaa ttgaatccct tttgatcat 4320  
  
 atctacacca cgcaaagta tgtatggtct tttggtgtcc tgctgtgga gatcgtgacc 4380  
 ctagggggaa acccctatcc tgggattcct cctgagcggc tcttcaacct tctgaagacc 4440  
 ggccaccgga tggagagcc agacaactgc agcagaggaga tgtaccgct gatgctgcaa 4500  
 tgctggaagc aggagccgga caaaaggccg gtgtttcgg acatcagcaa agacctggag 4560  
 aagatgatgg ttaagaggag agactacttg gaccttgcgg cgtccactcc atctgactec 4620  
 ctgatttatg acgacggcct ctgagaggag gagacaccgc tgggtgactg taataatgcc 4680

cccctccctc gageccctccc ttccacatgg attgaaaaca aactctatgg catgtcagac 4740  
  
 ccgaactggc ctggagagag tctgtacca ctcacgagag ctgatggcac taacctggg 4800  
 ttccaagat atccaaatga tagtgtatat gctaaactgga tgctttcacc ctcagcggca 4860  
 aaattaatgg acacgtttga tagttaa 4887  
  
 <210> 22  
 <211> 1628  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> fusion peptide  
 <400> 22  
  
 Met Ala Gln Ser Lys Arg His Val Tyr Ser Arg Thr Pro Ser Gly Ser  
 1 5 10 15  
 Arg Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg  
  
 20 25 30  
 Val Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly  
 35 40 45  
 Ala Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu  
 50 55 60  
 Ala Lys Gly Lys Asn Asp Gly Thr Val Gln Gly Arg Lys Tyr Phe Thr  
 65 70 75 80  
 Cys Asp Glu Gly His Gly Ile Phe Val Arg Gln Ser Gln Ile Gln Val  
  
 85 90 95  
 Phe Glu Asp Gly Ala Asp Thr Thr Ser Pro Glu Thr Pro Asp Ser Ser  
 100 105 110  
 Ala Ser Lys Val Leu Lys Arg Glu Gly Thr Asp Thr Thr Ala Lys Thr  
 115 120 125  
 Ser Lys Leu Ala Pro Thr Ala Arg Lys Thr Thr Thr Arg Arg Pro Lys  
 130 135 140  
 Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser Ser Ser Leu  
  
 145 150 155 160  
 Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser Ser Glu Pro



Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala Glu Glu Met  
 420 425 430  
 Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu Lys Val Arg  
 435 440 445  
 Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn Glu Met Asn  
 450 455 460  
 Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu Leu Arg Glu  
 465 470 475 480  
 Gln Leu Asp Met Ala Gly Ala Arg Val Arg Glu Ala Gln Lys Arg Val  
 485 490 495  
 Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr Ile Lys Lys  
 500 505 510  
 Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg Glu Leu Thr  
 515 520 525  
 Asn Gln Gln Glu Ala Ser Val Glu Arg Gln Gln Gln Pro Pro Pro Glu  
 530 535 540  
 Thr Phe Asp Phe Lys Ile Lys Phe Ala Glu Thr Lys Ala His Ala Lys  
 545 550 555 560  
 Ala Ile Glu Met Glu Leu Arg Gln Met Glu Val Ala Gln Ala Asn Arg  
 565 570 575  
 His Met Ser Leu Leu Thr Ala Phe Met Pro Asp Ser Phe Leu Arg Pro  
 580 585 590  
 Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met Pro Arg Leu  
 595 600 605  
 Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu Lys Phe Glu  
 610 615 620  
 Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly Ala Ala Gly  
 625 630 635 640  
 Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu Ser Leu Leu  
 645 650 655  
 Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln Cys Ser Val



Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg Ala Glu Ile

915 920 925

Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg Glu Thr Val

930 935 940

Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu Glu Leu Ser

945 950 955 960

Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu Asp Ser Ala

965 970 975

Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr Arg Leu Glu

980 985 990

Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe Glu Glu Thr

995 1000 1005

Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala Glu Lys

1010 1015 1020

Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg Thr Ile

1025 1030 1035

Glu Gly Leu Arg Gly Pro Pro Pro Ser Gly Ile Ala Thr Leu Val

1040 1045 1050

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

1055 1060 1065

Gln Ala Pro Gly Ser Val Pro Gly Pro Gly Leu Val Lys Asp Ser

1070 1075 1080

Pro Leu Leu Leu Gln Gln Ile Ser Ala Met Arg Leu His Ile Ser

1085 1090 1095

Gln Leu Gln His Glu Asn Ser Ile Leu Lys Gly Ala Gln Met Lys

1100 1105 1110

Ala Ser Leu Ala Ser Leu Pro Pro Leu His Val Ala Lys Leu Ser

1115 1120 1125

His Glu Gly Pro Gly Ser Glu Leu Pro Ala Gly Ala Leu Tyr Arg

1130 1135 1140

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

1145                      1150                      1155  
 Thr His Val Val Asp Ile Thr Arg Thr Ser Pro Ala Ala Lys Ser

1160                      1165                      1170  
 Pro Ser Ala Gln Leu Met Glu Gln Val Ala Gln Leu Lys Ser Leu

1175                      1180                      1185  
 Ser Asp Thr Val Glu Lys Leu Lys Asp Glu Val Leu Lys Glu Thr

1190                      1195                      1200  
 Val Ser Gln Arg Pro Gly Ala Thr Val Pro Thr Asp Phe Ala Thr

1205                      1210                      1215  
 Phe Pro Ser Ser Ala Phe Leu Arg Glu Asp Pro Lys Trp Glu Phe

1220                      1225                      1230  
 Pro Arg Lys Asn Leu Val Leu Gly Lys Thr Leu Gly Glu Gly Glu

1235                      1240                      1245  
 Phe Gly Lys Val Val Lys Ala Thr Ala Phe His Leu Lys Gly Arg

1250                      1255                      1260  
 Ala Gly Tyr Thr Thr Val Ala Val Lys Met Leu Lys Glu Asn Ala

1265                      1270                      1275  
 Ser Pro Ser Glu Leu Arg Asp Leu Leu Ser Glu Phe Asn Val Leu

1280                      1285                      1290  
 Lys Gln Val Asn His Pro His Val Ile Lys Leu Tyr Gly Ala Cys

1295                      1300                      1305  
 Ser Gln Asp Gly Pro Leu Leu Leu Ile Val Glu Tyr Ala Lys Tyr

1310                      1315                      1320  
 Gly Ser Leu Arg Gly Phe Leu Arg Glu Ser Arg Lys Val Gly Pro

1325                      1330                      1335  
 Gly Tyr Leu Gly Ser Gly Gly Ser Arg Asn Ser Ser Ser Leu Asp

1340                      1345                      1350  
 His Pro Asp Glu Arg Ala Leu Thr Met Gly Asp Leu Ile Ser Phe

1355                      1360                      1365  
 Ala Trp Gln Ile Ser Gln Gly Met Gln Tyr Leu Ala Glu Met Lys

1370                      1375                      1380

Leu Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Ala Glu  
 1385 1390 1395  
 Gly Arg Lys Met Lys Ile Ser Asp Phe Gly Leu Ser Arg Asp Val  
 1400 1405 1410  
 Tyr Glu Glu Asp Ser Tyr Val Lys Arg Ser Gln Gly Arg Ile Pro  
 1415 1420 1425  
 Val Lys Trp Met Ala Ile Glu Ser Leu Phe Asp His Ile Tyr Thr  
 1430 1435 1440  
 Thr Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile  
 1445 1450 1455  
 Val Thr Leu Gly Gly Asn Pro Tyr Pro Gly Ile Pro Pro Glu Arg  
 1460 1465 1470  
 Leu Phe Asn Leu Leu Lys Thr Gly His Arg Met Glu Arg Pro Asp  
 1475 1480 1485  
 Asn Cys Ser Glu Glu Met Tyr Arg Leu Met Leu Gln Cys Trp Lys  
 1490 1495 1500  
 Gln Glu Pro Asp Lys Arg Pro Val Phe Ala Asp Ile Ser Lys Asp  
 1505 1510 1515  
 Leu Glu Lys Met Met Val Lys Arg Arg Asp Tyr Leu Asp Leu Ala  
 1520 1525 1530  
 Ala Ser Thr Pro Ser Asp Ser Leu Ile Tyr Asp Asp Gly Leu Ser  
 1535 1540 1545  
 Glu Glu Glu Thr Pro Leu Val Asp Cys Asn Asn Ala Pro Leu Pro  
 1550 1555 1560  
 Arg Ala Leu Pro Ser Thr Trp Ile Glu Asn Lys Leu Tyr Gly Met  
 1565 1570 1575  
 Ser Asp Pro Asn Trp Pro Gly Glu Ser Pro Val Pro Leu Thr Arg  
 1580 1585 1590  
 Ala Asp Gly Thr Asn Thr Gly Phe Pro Arg Tyr Pro Asn Asp Ser  
 1595 1600 1605  
 Val Tyr Ala Asn Trp Met Leu Ser Pro Ser Ala Ala Lys Leu Met

1610 Asp Thr Phe Asp Ser  
 1615  
 1620  
 <210> 23  
 <211> 4761  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> coding sequence of fusion peptide  
 <400> 23

atggcacaga gcaagaggca cgtgtacagc cggacgccca gcggcagcag gatgagtgcg 60

gaggcaagcg cccggcctct gcgggtgggc tcccgtgtag aggtgattgg aaaaggccac 120

cgaggcactg tggcctatgt tggagccaca ctgtttgcca ctggcaaatg ggtaggcgtg 180

attctggatg aagcaaaggg caaaaatgat ggaactgttc aaggcaggaa gtacttcact 240

tgtgatgaag ggcatggcat ctttgtgcgc cagtcccaga tccaggtatt tgaagatgga 300

gcagatacta ctccccaga gacacctgat tcttctgctt caaaagtctt caaaagagag 360

ggaactgata caactgcaaa gactagcaaa ctggcaccga cagcccgaag gaccacaact 420

cggcgaccca agcccacgcg cccagccagt actggggtgg ctggggccag tagctccctg 480

ggccccctctg gctcagcgtc agcaggtgag ctgagcagca gtgagcccag caccgccgct 540

cagactccgc tggcagcacc catcatcccc acgccgtcc tcacctctcc tggagcagtc 600

cccccgcttc ctccccatc caaggaggag gagggactaa gggctcaggt gcgggacctg 660

gaggagaaac tagagacct gagactgaaa cgggcagaag acaaagcaaa gctaaaagag 720

ctggagaaac acaaaatcca gctggagcag gtgcaggaat ggaagagcaa aatgcaggag 780

cagcaggccg acctgcagcg gcgcctcaag gaggcgagaa aggaagccaa ggaggcgtg 840

gaggcaaagg aacgctatat ggaggagatg gctgatactg ctgatgcat tgagatggcc 900

actttggaca aggagatggc tgaagagcgg gctgagtccc tgcagcagga ggtggaggca 960

ctgaaggagc gggaggacga gctcactact gacttagaga tcctcaaggc tgagattgaa 1020

gagaagggct catatggcgc tgcattcagt tatcagctca agcagcttga ggagcagaat 1080

gccccctga aggatgcctt ggtgaggatg cgggatcttt cttcctcaga gaagcaggag 1140

catgtgaagc tccagaagct catggaaaag aagaaccaag agctggaagt tgtgaggcaa 1200

cagcgggagc gtctgcagga ggagctaagc caggcagaga gcaccattga tgagctcaag 1260

gagcaggtgg atgctgctct ggggtctgag gagatggtgg agatgctgac agatcggaac 1320

ctgaatctgg aagagaaagt gcgcgagttg agggagactg tgggagactt ggaagcgatg 1380  
 aatgagatga acgatgagct gcaggagaat gcacgtgaga cagaactgga gctgcgggag 1440  
 cagctggaca tggcaggcgc gcgggttcgt gaggcccaga agcgtgtgga ggcagcccag 1500  
 gagacggttg cagactacca gcagaccatc aagaagtacc gccagctgac cgcccatcta 1560  
 caggatgtga atcgggaact gacaaaccag caggaagcat ctgtggagag gcaacagcag 1620  
 ccacctccag agacctttga cttcaaaatc aagtttctg agactaagc ccatgccaag 1680  
 gcaattgaga tggaaattgag gcagatggag gtggcccagg ccaatcgaca catgtccctg 1740  
  
 ctgacagcct tcatgcctga cagcttctct cgcccagggtg gggaccatga ctgcgttctg 1800  
 gtgctgttgc tcatgcctcg tctcatttgc aaggcagagc tgatccgga gcaaggcccag 1860  
 gagaagtttg aactaagtga gaactgttca gagcggcctg ggctgcgagg agctgctggg 1920  
 gagcaactca gctttgctgc tggactgggtg tactcgtctga gcctgctgca ggccacgcta 1980  
 caccgctatg agcatgcctt ctctcagtgc agtgtggatg tgtataagaa agtgggcagc 2040  
 ctgtaccctg agatgagtgc ccatgagcgc tccttgatt tcctcattga actgctgcac 2100  
 aaggatcagc tggatgagac tgtcaatgtg gagcctctca ccaaggccat caagtactat 2160  
  
 cagcatctgt acagcatcca ccttgcgga cagcctgagg actgtactat gcagctggct 2220  
 gaccacatta agttcacga gagtgtctctg gactgcatga gtgtggaggt aggacggctg 2280  
 cgtgccttct tgcaggggtg gcaggaggct acagatatg ccctcctgct ccgggatctg 2340  
 gaaacttcat gcagtgcacat ccgccagttc tgcaagaaga tccgaaggcg aatgccaggg 2400  
 acagatgctc ctgggatccc agctgcaactg gcctttggac cacaggtatc tgacacgctc 2460  
 ctagactgca ggaaacactt gacgtgggtc gtggctgtgc tgcaggaggt ggcagctgct 2520  
 gctgcccagc tcattgcccc actggcagag aatgaggggc tacttgtggc tgctctggag 2580  
  
 gaactggctt tcaaagcaag cgagcagatc tatgggaccc cctccagcag cccctatgag 2640  
 tgtctgcgcc agtcatgcaa catcctcatc agtaccatga acaagctggc cacagccatg 2700  
 caggaggggg agtatgatgc agagcggccc ccagcaagc ctccaccgtg tgaactcggg 2760  
 gctgctgccc ttcgtgcaga gatcacagat gctgaaggcc tgggtttgaa getcgaagat 2820  
 cgagagacag ttattaagga gttgaagaag tactcaaga ttaagggaga ggagctaagt 2880  
 gaggccaatg tgcggctgag cctcctggag aagaagtgg acagtgtctc caagatgca 2940  
 gatgagcgca tcgagaaagt ccagactcgg ctggaggaga cccaggcact gctgcgaaag 3000  
  
 aaggagaaag agtttgagga gacaatgat gactccagg ctgacatcga ccagctggag 3060  
 gcagagaagg cagaactaaa gcagcgtctg aacagccagt ccaaacgcac gattgagggg 3120  
 ctccggggcc ctctccttc aggcattgct actctgtct ctggcattgc tggatgaagaa 3180

cagcagcgag gagccatccc tgggcaggct ccagggtctg tgccaggccc agggctggtg 3240  
aaggactcac cactgctgct tcagcagatc tctgccatga ggctgcacat ctcccagctc 3300  
cagcatgaga acagcatcct caaggagacc cagatgaagg catccttggc atccctgccc 3360  
cctctgcatg ttgcaaagct atcccatgag ggccttgcca gtgagttacc agctggagcg 3420

ctgtatcgta agaccagcca gctgctggag acattgaatc aattgagcac acacacgcac 3480  
gtagtagaca teactcgcac cagecctgct gccaagagcc cgctcgccca acttatggag 3540  
caagtggctc agcttaagtc cctgagtac accgtcgaga agctcaagga tgaggctctc 3600  
aaggagacag tatctcagcg ccttgagacc acagtaccca ctgactttgc caccttcctt 3660  
tcatcagcct tcctcagga ggatccaaag tgggaattcc ctcggaagaa cttggttctt 3720  
ggaaaaactc taggagaagg cgaatttga aaagtgtca aggcaacggc cttccatctg 3780  
aaaggcagag cagggtacac cacggtggcc gtgaagatgc tgaaagagaa cgctccccg 3840

agtgagcttc gagacctgct gtcagattc aacgtcctga agcaggtcaa ccaccacat 3900  
gtcatcaaat tgtatggggc ctgcagccag gatggcccgc tcctcctcat cgtggagtac 3960  
gccaataacg gtcctctcgc gggcttctc cgcgagagcc gcaaagtggg gectggctac 4020  
ctgggcagtg gaggcagccg caactccagc tccttgacc acccggatga gcgggcctc 4080  
accatgggcg acctcatctc atttgcctgg cagatctcac aggggatgca gtatctggcc 4140  
gagatgaage tegtcatcg ggacttggca gccagaaaca tcctggtagc tgagggcgcg 4200  
aagatgaaga tttcggattt cggcttgtcc cgagatgttt atgaagagga ttctactgtg 4260

aagaggagcc agggctggat tccagttaa tggatggcaa ttgaatccct ttttgatcat 4320  
atctacacca cgcaaagtga tgtatggtct tttggtgtcc tgctgtggga gatcgtgacc 4380  
ctaggggaa acccctatcc tgggattcct cctgagcggc tcttcaacct tctgaagacc 4440  
ggccaccgga tggagaggcc agacaactgc agcagaggaga tgtaccgct gatgctgcaa 4500  
tgctggaagc aggagccgga caaaaggccg gtgtttgcgg acatcagcaa agacctggag 4560  
aagatgatgg ttaagaggag agactacttg gaccttgcgg cgtccactcc atctgactcc 4620  
ctgatttatg acgacggcct ctgagaggag gagacaccgc tgggtgactg taataatgcc 4680

ccccctcctc gagccctccc ttccatgag attgaaaaca aactctatgg tagaatttcc 4740  
catgcattta ctagattcta g 4761

<210> 24

<211> 1586

<212> PRT

<213> Artificial Sequence

<220><223> fusion peptide

<400> 24

Met Ala Gln Ser Lys Arg His Val Tyr Ser Arg Thr Pro Ser Gly Ser  
 1                    5                    10                    15

Arg Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg  
                   20                    25                    30

Val Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly  
                   35                    40                    45

Ala Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu  
                   50                    55                    60

Ala Lys Gly Lys Asn Asp Gly Thr Val Gln Gly Arg Lys Tyr Phe Thr  
 65                    70                    75                    80

Cys Asp Glu Gly His Gly Ile Phe Val Arg Gln Ser Gln Ile Gln Val  
                   85                    90                    95

Phe Glu Asp Gly Ala Asp Thr Thr Ser Pro Glu Thr Pro Asp Ser Ser  
                   100                    105                    110

Ala Ser Lys Val Leu Lys Arg Glu Gly Thr Asp Thr Thr Ala Lys Thr  
                   115                    120                    125

Ser Lys Leu Ala Pro Thr Ala Arg Lys Thr Thr Thr Arg Arg Pro Lys  
                   130                    135                    140

Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser Ser Ser Leu  
 145                    150                    155                    160

Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser Ser Glu Pro  
                   165                    170                    175

Ser Thr Pro Ala Gln Thr Pro Leu Ala Ala Pro Ile Ile Pro Thr Pro  
                   180                    185                    190

Val Leu Thr Ser Pro Gly Ala Val Pro Pro Leu Pro Ser Pro Ser Lys  
                   195                    200                    205

Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu Glu Lys Leu  
                   210                    215                    220

Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys Leu Lys Glu

225                      230                      235                      240  
 Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu Trp Lys Ser  
                                  245                      250                      255  
 Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu Lys Glu Ala  
                                  260                      265                      270  
 Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg Tyr Met Glu  
                                  275                      280                      285  
  
 Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr Leu Asp Lys  
                                  290                      295                      300  
 Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu Val Glu Ala  
 305                      310                      315                      320  
 Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu Ile Leu Lys  
                                  325                      330                      335  
 Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser Ser Tyr Gln  
                                  340                      345                      350  
  
 Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp Ala Leu Val  
                                  355                      360                      365  
 Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His Val Lys Leu  
                                  370                      375                      380  
 Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val Val Arg Gln  
 385                      390                      395                      400  
 Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu Ser Thr Ile  
                                  405                      410                      415  
  
 Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala Glu Glu Met  
                                  420                      425                      430  
 Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu Lys Val Arg  
                                  435                      440                      445  
 Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn Glu Met Asn  
                                  450                      455                      460  
 Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu Leu Arg Glu  
 465                      470                      475                      480

Gln Leu Asp Met Ala Gly Ala Arg Val Arg Glu Ala Gln Lys Arg Val  
 485 490 495

Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr Ile Lys Lys  
 500 505 510

Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg Glu Leu Thr  
 515 520 525

Asn Gln Gln Glu Ala Ser Val Glu Arg Gln Gln Gln Pro Pro Pro Glu  
 530 535 540

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

Ala Ile Glu Met Glu Leu Arg Gln Met Glu Val Ala Gln Ala Asn Arg  
 565 570 575

His Met Ser Leu Leu Thr Ala Phe Met Pro Asp Ser Phe Leu Arg Pro  
 580 585 590

Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met Pro Arg Leu  
 595 600 605

Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu Lys Phe Glu  
 610 615 620

Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly Ala Ala Gly  
 625 630 635 640

Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu Ser Leu Leu  
 645 650 655

Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln Cys Ser Val  
 660 665 670

Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met Ser Ala His  
 675 680 685

Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys Asp Gln Leu  
 690 695 700

Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile Lys Tyr Tyr  
 705 710 715 720

Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu Asp Cys Thr

	725	730	735
Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala Leu Asp Cys			
	740	745	750
Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln Gly Gly Gln			
	755	760	765
Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr Ser Cys			
	770	775	780
Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met Pro Gly			
785	790	795	800
Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro Gln Val			
	805	810	815
Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp Val Val Ala			
	820	825	830
Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile Ala Pro Leu			
	835	840	845
Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu Leu Ala Phe			
	850	855	860
Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser Pro Tyr Glu			
865	870	875	880
Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met Asn Lys Leu			
	885	890	895
Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg Pro Pro Ser			
	900	905	910
Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg Ala Glu Ile			
	915	920	925
Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg Glu Thr Val			
	930	935	940
Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu Glu Leu Ser			
945	950	955	960
Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu Asp Ser Ala			
	965	970	975

Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr Arg Leu Glu  
 980 985 990

Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe Glu Glu Thr  
 995 1000 1005

Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala Glu Lys  
 1010 1015 1020

Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg Thr Ile  
 1025 1030 1035

Glu Gly Leu Arg Gly Pro Pro Pro Ser Gly Ile Ala Thr Leu Val  
 1040 1045 1050

Ser Gly Ile Ala Gly Glu Glu Gln Gln Arg Gly Ala Ile Pro Gly  
 1055 1060 1065

Gln Ala Pro Gly Ser Val Pro Gly Pro Gly Leu Val Lys Asp Ser  
 1070 1075 1080

Pro Leu Leu Leu Gln Gln Ile Ser Ala Met Arg Leu His Ile Ser  
 1085 1090 1095

Gln Leu Gln His Glu Asn Ser Ile Leu Lys Gly Ala Gln Met Lys  
 1100 1105 1110

Ala Ser Leu Ala Ser Leu Pro Pro Leu His Val Ala Lys Leu Ser  
 1115 1120 1125

His Glu Gly Pro Gly Ser Glu Leu Pro Ala Gly Ala Leu Tyr Arg  
 1130 1135 1140

Lys Thr Ser Gln Leu Leu Glu Thr Leu Asn Gln Leu Ser Thr His  
 1145 1150 1155

Thr His Val Val Asp Ile Thr Arg Thr Ser Pro Ala Ala Lys Ser  
 1160 1165 1170

Pro Ser Ala Gln Leu Met Glu Gln Val Ala Gln Leu Lys Ser Leu  
 1175 1180 1185

Ser Asp Thr Val Glu Lys Leu Lys Asp Glu Val Leu Lys Glu Thr  
 1190 1195 1200

Val Ser Gln Arg Pro Gly Ala Thr Val Pro Thr Asp Phe Ala Thr

1205	1210	1215
Phe Pro Ser Ser Ala Phe Leu Arg Glu Asp Pro Lys Trp Glu Phe		
1220	1225	1230
Pro Arg Lys Asn Leu Val Leu Gly Lys Thr Leu Gly Glu Gly Glu		
1235	1240	1245
Phe Gly Lys Val Val Lys Ala Thr Ala Phe His Leu Lys Gly Arg		
1250	1255	1260
Ala Gly Tyr Thr Thr Val Ala Val Lys Met Leu Lys Glu Asn Ala		
1265	1270	1275
Ser Pro Ser Glu Leu Arg Asp Leu Leu Ser Glu Phe Asn Val Leu		
1280	1285	1290
Lys Gln Val Asn His Pro His Val Ile Lys Leu Tyr Gly Ala Cys		
1295	1300	1305
Ser Gln Asp Gly Pro Leu Leu Leu Ile Val Glu Tyr Ala Lys Tyr		
1310	1315	1320
Gly Ser Leu Arg Gly Phe Leu Arg Glu Ser Arg Lys Val Gly Pro		
1325	1330	1335
Gly Tyr Leu Gly Ser Gly Gly Ser Arg Asn Ser Ser Ser Leu Asp		
1340	1345	1350
His Pro Asp Glu Arg Ala Leu Thr Met Gly Asp Leu Ile Ser Phe		
1355	1360	1365
Ala Trp Gln Ile Ser Gln Gly Met Gln Tyr Leu Ala Glu Met Lys		
1370	1375	1380
Leu Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Ala Glu		
1385	1390	1395
Gly Arg Lys Met Lys Ile Ser Asp Phe Gly Leu Ser Arg Asp Val		
1400	1405	1410
Tyr Glu Glu Asp Ser Tyr Val Lys Arg Ser Gln Gly Arg Ile Pro		
1415	1420	1425
Val Lys Trp Met Ala Ile Glu Ser Leu Phe Asp His Ile Tyr Thr		
1430	1435	1440

Thr Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile  
 1445 1450 1455

Val Thr Leu Gly Gly Asn Pro Tyr Pro Gly Ile Pro Pro Glu Arg  
 1460 1465 1470

Leu Phe Asn Leu Leu Lys Thr Gly His Arg Met Glu Arg Pro Asp  
 1475 1480 1485

Asn Cys Ser Glu Glu Met Tyr Arg Leu Met Leu Gln Cys Trp Lys  
 1490 1495 1500

Gln Glu Pro Asp Lys Arg Pro Val Phe Ala Asp Ile Ser Lys Asp  
 1505 1510 1515

Leu Glu Lys Met Met Val Lys Arg Arg Asp Tyr Leu Asp Leu Ala  
 1520 1525 1530

Ala Ser Thr Pro Ser Asp Ser Leu Ile Tyr Asp Asp Gly Leu Ser  
 1535 1540 1545

Glu Glu Glu Thr Pro Leu Val Asp Cys Asn Asn Ala Pro Leu Pro  
 1550 1555 1560

Arg Ala Leu Pro Ser Thr Trp Ile Glu Asn Lys Leu Tyr Gly Arg  
 1565 1570 1575

Ile Ser His Ala Phe Thr Arg Phe  
 1580 1585

<210> 25

<211> 1278

<212> PRT

<213> Homo sapiens

<400> 25

Met Ala Gln Ser Lys Arg His Val Tyr Ser Arg Thr Pro Ser Gly Ser

1 5 10 15

Arg Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg

20 25 30

Val Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly

35 40 45

Ala Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu



Glu Met Ala Thr Leu Asp Lys Glu Met Ala Glu Glu Arg Ala Glu Ser  
 305 310 315 320

Leu Gln Gln Glu Val Glu Ala Leu Lys Glu Arg Val Asp Glu Leu Thr  
 325 330 335

Thr Asp Leu Glu Ile Leu Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp  
 340 345 350

Gly Ala Ala Ser Ser Tyr Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala  
 355 360 365

Arg Leu Lys Asp Ala Leu Val Arg Met Arg Asp Leu Ser Ser Ser Glu  
 370 375 380

Lys Gln Glu His Val Lys Leu Gln Lys Leu Met Glu Lys Lys Asn Gln  
 385 390 395 400

Glu Leu Glu Val Val Arg Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu  
 405 410 415

Ser Gln Ala Glu Ser Thr Ile Asp Glu Leu Lys Glu Gln Val Asp Ala  
 420 425 430

Ala Leu Gly Ala Glu Glu Met Val Glu Met Leu Thr Asp Arg Asn Leu  
 435 440 445

Asn Leu Glu Glu Lys Val Arg Glu Leu Arg Glu Thr Val Gly Asp Leu  
 450 455 460

Glu Ala Met Asn Glu Met Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu  
 465 470 475 480

Thr Glu Leu Glu Leu Arg Glu Gln Leu Asp Met Ala Gly Ala Arg Val  
 485 490 495

Arg Glu Ala Gln Lys Arg Val Glu Ala Ala Gln Glu Thr Val Ala Asp  
 500 505 510

Tyr Gln Gln Thr Ile Lys Lys Tyr Arg Gln Leu Thr Ala His Leu Gln  
 515 520 525

Asp Val Asn Arg Glu Leu Thr Asn Gln Gln Glu Ala Ser Val Glu Arg  
 530 535 540

Gln Gln Gln Pro Pro Pro Glu Thr Phe Asp Phe Lys Ile Lys Phe Ala

545                      550                      555                      560  
 Glu Thr Lys Ala His Ala Lys Ala Ile Glu Met Glu Leu Arg Gln Met  
                                  565                      570                      575  
  
 Glu Val Ala Gln Ala Asn Arg His Met Ser Leu Leu Thr Ala Phe Met  
                                  580                      585                      590  
 Pro Asp Ser Phe Leu Arg Pro Gly Gly Asp His Asp Cys Val Leu Val  
                                  595                      600                      605  
 Leu Leu Leu Met Pro Arg Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys  
                                  610                      615                      620  
 Gln Ala Gln Glu Lys Phe Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro  
 625                      630                      635                      640  
  
 Gly Leu Arg Gly Ala Ala Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu  
                                  645                      650                      655  
 Val Tyr Ser Leu Ser Leu Leu Gln Ala Thr Leu His Arg Tyr Glu His  
                                  660                      665                      670  
 Ala Leu Ser Gln Cys Ser Val Asp Val Tyr Lys Lys Val Gly Ser Leu  
                                  675                      680                      685  
 Tyr Pro Glu Met Ser Ala His Glu Arg Ser Leu Asp Phe Leu Ile Glu  
                                  690                      695                      700  
  
 Leu Leu His Lys Asp Gln Leu Asp Glu Thr Val Asn Val Glu Pro Leu  
 705                      710                      715                      720  
 Thr Lys Ala Ile Lys Tyr Tyr Gln His Leu Tyr Ser Ile His Leu Ala  
                                  725                      730                      735  
 Glu Gln Pro Glu Asp Cys Thr Met Gln Leu Ala Asp His Ile Lys Phe  
                                  740                      745                      750  
 Thr Gln Ser Ala Leu Asp Cys Met Ser Val Glu Val Gly Arg Leu Arg  
                                  755                      760                      765  
  
 Ala Phe Leu Gln Gly Gly Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu  
                                  770                      775                      780  
 Arg Asp Leu Glu Thr Ser Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys  
 785                      790                      795                      800

Ile Arg Arg Arg Met Pro Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala  
 805 810 815

Leu Ala Phe Gly Pro Gln Val Ser Asp Thr Leu Leu Asp Cys Arg Lys  
 820 825 830

His Leu Thr Trp Val Val Ala Val Leu Gln Glu Val Ala Ala Ala Ala  
 835 840 845

Ala Gln Leu Ile Ala Pro Leu Ala Glu Asn Glu Gly Leu Leu Val Ala  
 850 855 860

Ala Leu Glu Glu Leu Ala Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr  
 865 870 875 880

Pro Ser Ser Ser Pro Tyr Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu  
 885 890 895

Ile Ser Thr Met Asn Lys Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr  
 900 905 910

Asp Ala Glu Arg Pro Pro Ser Lys Pro Pro Pro Val Glu Leu Arg Ala  
 915 920 925

Ala Ala Leu Arg Ala Glu Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys  
 930 935 940

Leu Glu Asp Arg Glu Thr Val Ile Lys Glu Leu Lys Lys Ser Leu Lys  
 945 950 955 960

Ile Lys Gly Glu Glu Leu Ser Glu Ala Asn Val Arg Leu Ser Leu Leu  
 965 970 975

Glu Lys Lys Leu Asp Ser Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu  
 980 985 990

Lys Val Gln Thr Arg Leu Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys  
 995 1000 1005

Glu Lys Glu Phe Glu Glu Thr Met Asp Ala Leu Gln Ala Asp Ile  
 1010 1015 1020

Asp Gln Leu Glu Ala Glu Lys Ala Glu Leu Lys Gln Arg Leu Asn  
 1025 1030 1035

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

1040	1045	1050
Ser Gly Ile Ala Thr Leu Val	Ser Gly Ile Ala Gly	Glu Glu Gln
1055	1060	1065
Gln Arg Gly Ala Ile Pro Gly	Gln Ala Pro Gly Ser	Val Pro Gly
1070	1075	1080
Pro Gly Leu Val Lys Asp Ser	Pro Leu Leu Leu Gln	Gln Ile Ser
1085	1090	1095
Ala Met Arg Leu His Ile Ser	Gln Leu Gln His Glu	Asn Ser Ile
1100	1105	1110
Leu Lys Gly Ala Gln Met Lys	Ala Ser Leu Ala Ser	Leu Pro Pro
1115	1120	1125
Leu His Val Ala Lys Leu Ser	His Glu Gly Pro Gly	Ser Glu Leu
1130	1135	1140
Pro Ala Gly Ala Leu Tyr Arg	Lys Thr Ser Gln Leu	Leu Glu Thr
1145	1150	1155
Leu Asn Gln Leu Ser Thr His	Thr His Val Val Asp	Ile Thr Arg
1160	1165	1170
Thr Ser Pro Ala Ala Lys Ser	Pro Ser Ala Gln Leu	Met Glu Gln
1175	1180	1185
Val Ala Gln Leu Lys Ser Leu	Ser Asp Thr Val Glu	Lys Leu Lys
1190	1195	1200
Asp Glu Val Leu Lys Glu Thr	Val Ser Gln Arg Pro	Gly Ala Thr
1205	1210	1215
Val Pro Thr Asp Phe Ala Thr	Phe Pro Ser Ser Ala	Phe Leu Arg
1220	1225	1230
Ala Lys Glu Glu Gln Gln Asp	Asp Thr Val Tyr Met	Gly Lys Val
1235	1240	1245
Thr Phe Ser Cys Ala Ala Gly	Phe Gly Gln Arg His	Arg Leu Val
1250	1255	1260
Leu Thr Gln Glu Gln Leu His	Gln Leu His Ser Arg	Leu Ile Ser
1265	1270	1275

<210> 26

<211> 1144

<212> PRT

<213> Homo sapiens

<400> 26

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



Pro Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met Pro Arg  
 465                      470                      475                      480  
 Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu Lys Phe  
                                  485                      490                      495  
 Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly Ala Ala  
                                  500                      505                      510  
 Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu Ser Leu  
                                  515                      520                      525  
 Leu Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln Cys Ser  
                                  530                      535                      540  
 Val Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met Ser Ala  
  
 545                      550                      555                      560  
 His Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys Asp Gln  
                                  565                      570                      575  
 Leu Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile Lys Tyr  
                                  580                      585                      590  
 Tyr Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu Asp Cys  
                                  595                      600                      605  
 Thr Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala Leu Asp  
  
 610                      615                      620  
 Cys Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln Gly Gly  
 625                      630                      635                      640  
 Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr Ser  
                                  645                      650                      655  
 Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met Pro  
                                  660                      665                      670  
 Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro Gln  
  
 675                      680                      685  
 Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp Val Val  
                                  690                      695                      700  
 Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile Ala Pro

705                      710                      715                      720  
 Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu Leu Ala  
                                  725                      730                      735  
 Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser Pro Tyr  
  
                                  740                      745                      750  
 Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met Asn Lys  
                                  755                      760                      765  
 Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg Pro Pro  
                                  770                      775                      780  
 Ser Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg Ala Glu  
 785                      790                      795                      800  
 Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg Glu Thr  
  
                                  805                      810                      815  
 Val Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu Glu Leu  
                                  820                      825                      830  
 Ser Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu Asp Ser  
                                  835                      840                      845  
 Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr Arg Leu  
                                  850                      855                      860  
 Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe Glu Glu  
  
 865                      870                      875                      880  
 Thr Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala Glu Lys  
                                  885                      890                      895  
 Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg Thr Ile Glu  
                                  900                      905                      910  
 Gly Leu Arg Gly Pro Pro Pro Ser Gly Ile Ala Thr Leu Val Ser Gly  
                                  915                      920                      925  
 Ile Ala Gly Glu Glu Gln Gln Arg Gly Ala Ile Pro Gly Gln Ala Pro  
  
                                  930                      935                      940  
 Gly Ser Val Pro Gly Pro Gly Leu Val Lys Asp Ser Pro Leu Leu Leu  
 945                      950                      955                      960

Gln Gln Ile Ser Ala Met Arg Leu His Ile Ser Gln Leu Gln His Glu  
 965 970 975  
 Asn Ser Ile Leu Lys Gly Ala Gln Met Lys Ala Ser Leu Ala Ser Leu  
 980 985 990  
 Pro Pro Leu His Val Ala Lys Leu Ser His Glu Gly Pro Gly Ser Glu  
 995 1000 1005  
 Leu Pro Ala Gly Ala Leu Tyr Arg Lys Thr Ser Gln Leu Leu Glu  
 1010 1015 1020  
 Thr Leu Asn Gln Leu Ser Thr His Thr His Val Val Asp Ile Thr  
 1025 1030 1035  
 Arg Thr Ser Pro Ala Ala Lys Ser Pro Ser Ala Gln Leu Met Glu  
 1040 1045 1050  
 Gln Val Ala Gln Leu Lys Ser Leu Ser Asp Thr Val Glu Lys Leu  
 1055 1060 1065  
 Lys Asp Glu Val Leu Lys Glu Thr Val Ser Gln Arg Pro Gly Ala  
 1070 1075 1080  
 Thr Val Pro Thr Asp Phe Ala Thr Phe Pro Ser Ser Ala Phe Leu  
 1085 1090 1095  
 Arg Ala Lys Glu Glu Gln Gln Asp Asp Thr Val Tyr Met Gly Lys  
 1100 1105 1110  
 Val Thr Phe Ser Cys Ala Ala Gly Phe Gly Gln Arg His Arg Leu  
 1115 1120 1125  
 Val Leu Thr Gln Glu Gln Leu His Gln Leu His Ser Arg Leu Ile  
 1130 1135 1140  
 Ser  
 <210> 27  
 <211> 1253  
 <212> PRT  
 <213> Homo sapiens  
 <400> 27  
 Met Ala Gln Ser Lys Arg His Val Tyr Ser Arg Thr Pro Ser Gly Ser

1                    5                    10                    15  
 Arg Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg  
                          20                    25                    30  
  
 Val Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly  
                          35                    40                    45  
 Ala Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu  
                          50                    55                    60  
 Ala Lys Gly Lys Asn Asp Gly Thr Val Gln Gly Arg Lys Tyr Phe Thr  
 65                    70                    75                    80  
 Cys Asp Glu Gly His Gly Ile Phe Val Arg Gln Ser Gln Ile Gln Val  
                          85                    90                    95  
  
 Phe Glu Asp Gly Ala Asp Thr Thr Ser Pro Glu Thr Pro Asp Ser Ser  
                          100                    105                    110  
 Ala Ser Lys Val Leu Lys Arg Glu Gly Thr Asp Thr Thr Ala Lys Thr  
                          115                    120                    125  
 Ser Lys Leu Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser  
                          130                    135                    140  
 Ser Ser Leu Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser  
 145                    150                    155                    160  
  
 Ser Glu Pro Ser Thr Pro Ala Gln Thr Pro Leu Ala Ala Pro Ile Ile  
                          165                    170                    175  
 Pro Thr Pro Val Leu Thr Ser Pro Gly Ala Val Pro Pro Leu Pro Ser  
                          180                    185                    190  
 Pro Ser Lys Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu  
                          195                    200                    205  
 Glu Lys Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys  
                          210                    215                    220  
  
 Leu Lys Glu Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu  
 225                    230                    235                    240  
 Trp Lys Ser Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu  
                          245                    250                    255

Lys Glu Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg  
 260 265 270  
 Tyr Met Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr  
 275 280 285  
  
 Leu Asp Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu  
 290 295 300  
 Val Glu Ala Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu  
 305 310 315 320  
 Ile Leu Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser  
 325 330 335  
 Ser Tyr Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp  
 340 345 350  
  
 Ala Leu Val Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His  
 355 360 365  
 Val Lys Leu Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val  
 370 375 380  
 Val Arg Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu  
 385 390 395 400  
 Ser Thr Ile Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala  
 405 410 415  
  
 Glu Glu Met Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu  
 420 425 430  
 Lys Val Arg Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn  
 435 440 445  
 Glu Met Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu  
 450 455 460  
 Leu Arg Glu Gln Leu Asp Met Ala Gly Ala Arg Val Arg Glu Ala Gln  
 465 470 475 480  
  
 Lys Arg Val Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr  
 485 490 495  
 Ile Lys Lys Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg



Gly Gly Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu  
 755 760 765  
 Thr Ser Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg  
 770 775 780  
 Met Pro Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly  
 785 790 795 800  
  
 Pro Gln Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp  
 805 810 815  
 Val Val Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile  
 820 825 830  
 Ala Pro Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu  
 835 840 845  
 Leu Ala Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser  
 850 855 860  
  
 Pro Tyr Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met  
 865 870 875 880  
 Asn Lys Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg  
 885 890 895  
 Pro Pro Ser Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg  
 900 905 910  
 Ala Glu Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg  
 915 920 925  
  
 Glu Thr Val Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu  
 930 935 940  
 Glu Leu Ser Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu  
 945 950 955 960  
 Asp Ser Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr  
 965 970 975  
 Arg Leu Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe  
 980 985 990  
  
 Glu Glu Thr Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala

995	1000	1005
Glu Lys Ala Glu Leu Lys Gln Arg Leu Asn Ser Gln Ser Lys Arg		
1010	1015	1020
Thr Ile Glu Gly Leu Arg Gly Pro Pro Pro Ser Gly Ile Ala Thr		
1025	1030	1035
Leu Val Ser Gly Ile Ala Gly Gly Ala Ile Pro Gly Gln Ala Pro		
1040	1045	1050
Gly Ser Val Pro Gly Pro Gly Leu Val Lys Asp Ser Pro Leu Leu		
1055	1060	1065
Leu Gln Gln Ile Ser Ala Met Arg Leu His Ile Ser Gln Leu Gln		
1070	1075	1080
His Glu Asn Ser Ile Leu Lys Gly Ala Gln Met Lys Ala Ser Leu		
1085	1090	1095
Ala Ser Leu Pro Pro Leu His Val Ala Lys Leu Ser His Glu Gly		
1100	1105	1110
Pro Gly Ser Glu Leu Pro Ala Gly Ala Leu Tyr Arg Lys Thr Ser		
1115	1120	1125
Gln Leu Leu Glu Thr Leu Asn Gln Leu Ser Thr His Thr His Val		
1130	1135	1140
Val Asp Ile Thr Arg Thr Ser Pro Ala Ala Lys Ser Pro Ser Ala		
1145	1150	1155
Gln Leu Met Glu Gln Val Ala Gln Leu Lys Ser Leu Ser Asp Thr		
1160	1165	1170
Val Glu Lys Leu Lys Asp Glu Val Leu Lys Glu Thr Val Ser Gln		
1175	1180	1185
Arg Pro Gly Ala Thr Val Pro Thr Asp Phe Ala Thr Phe Pro Ser		
1190	1195	1200
Ser Ala Phe Leu Arg Ala Lys Glu Glu Gln Gln Asp Asp Thr Val		
1205	1210	1215
Tyr Met Gly Lys Val Thr Phe Ser Cys Ala Ala Gly Phe Gly Gln		
1220	1225	1230

Arg His Arg Leu Val Leu Thr Gln Glu Gln Leu His Gln Leu His

1235 1240 1245

Ser Arg Leu Ile Ser

1250

<210> 28

<211> 1139

<212> PRT

<213> Homo sapiens

<400> 28

Met Met Arg Gln Ala Pro Thr Ala Arg Lys Thr Thr Thr Arg Arg Pro

1 5 10 15

Lys Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser Ser Ser

20 25 30

Leu Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser Ser Glu

35 40 45

Pro Ser Thr Pro Ala Gln Thr Pro Leu Ala Ala Pro Ile Ile Pro Thr

50 55 60

Pro Val Leu Thr Ser Pro Gly Ala Val Pro Pro Leu Pro Ser Pro Ser

65 70 75 80

Lys Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu Glu Lys

85 90 95

Leu Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys Leu Lys

100 105 110

Glu Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu Trp Lys

115 120 125

Ser Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu Lys Glu

130 135 140

Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg Tyr Met

145 150 155 160

Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr Leu Asp

165 170 175

Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu Val Glu



Lys Ala Ile Glu Met Glu Leu Arg Gln Met Glu Val Ala Gln Ala Asn  
 435 440 445  
 Arg His Met Ser Leu Leu Thr Ala Phe Met Pro Asp Ser Phe Leu Arg  
 450 455 460  
 Pro Gly Gly Asp His Asp Cys Val Leu Val Leu Leu Leu Met Pro Arg  
 465 470 475 480  
  
 Leu Ile Cys Lys Ala Glu Leu Ile Arg Lys Gln Ala Gln Glu Lys Phe  
 485 490 495  
 Glu Leu Ser Glu Asn Cys Ser Glu Arg Pro Gly Leu Arg Gly Ala Ala  
 500 505 510  
 Gly Glu Gln Leu Ser Phe Ala Ala Gly Leu Val Tyr Ser Leu Ser Leu  
 515 520 525  
 Leu Gln Ala Thr Leu His Arg Tyr Glu His Ala Leu Ser Gln Cys Ser  
 530 535 540  
  
 Val Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met Ser Ala  
 545 550 555 560  
 His Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys Asp Gln  
 565 570 575  
 Leu Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile Lys Tyr  
 580 585 590  
 Tyr Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu Asp Cys  
 595 600 605  
  
 Thr Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala Leu Asp  
 610 615 620  
 Cys Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln Gly Gly  
 625 630 635 640  
 Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr Ser  
 645 650 655  
 Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met Pro  
 660 665 670  
  
 Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro Gln



Ile Ala Gly Gly Ala Ile Pro Gly Gln Ala Pro Gly Ser Val Pro Gly  
 930 935 940  
 Pro Gly Leu Val Lys Asp Ser Pro Leu Leu Leu Gln Gln Ile Ser Ala  
 945 950 955 960  
 Met Arg Leu His Ile Ser Gln Leu Gln His Glu Asn Ser Ile Leu Lys  
 965 970 975  
 Gly Ala Gln Met Lys Ala Ser Leu Ala Ser Leu Pro Pro Leu His Val  
 980 985 990

Ala Lys Leu Ser His Glu Gly Pro Gly Ser Glu Leu Pro Ala Gly Ala  
 995 1000 1005  
 Leu Tyr Arg Lys Thr Ser Gln Leu Leu Glu Thr Leu Asn Gln Leu  
 1010 1015 1020  
 Ser Thr His Thr His Val Val Asp Ile Thr Arg Thr Ser Pro Ala  
 1025 1030 1035  
 Ala Lys Ser Pro Ser Ala Gln Leu Met Glu Gln Val Ala Gln Leu  
 1040 1045 1050

Lys Ser Leu Ser Asp Thr Val Glu Lys Leu Lys Asp Glu Val Leu  
 1055 1060 1065  
 Lys Glu Thr Val Ser Gln Arg Pro Gly Ala Thr Val Pro Thr Asp  
 1070 1075 1080  
 Phe Ala Thr Phe Pro Ser Ser Ala Phe Leu Arg Ala Lys Glu Glu  
 1085 1090 1095  
 Gln Gln Asp Asp Thr Val Tyr Met Gly Lys Val Thr Phe Ser Cys  
 1100 1105 1110

Ala Ala Gly Phe Gly Gln Arg His Arg Leu Val Leu Thr Gln Glu  
 1115 1120 1125  
 Gln Leu His Gln Leu His Ser Arg Leu Ile Ser  
 1130 1135

- <210> 29
- <211> 1236
- <212> PRT
- <213> Homo sapiens

<400> 29

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

Glu Ala Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg Tyr  
 245 250 255  
 Met Glu Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr Leu  
 260 265 270  
 Asp Lys Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu Val  
 275 280 285  
  
 Glu Ala Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu Ile  
 290 295 300  
 Leu Lys Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser Ser  
 305 310 315 320  
 Tyr Gln Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp Ala  
 325 330 335  
 Leu Val Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His Val  
 340 345 350  
  
 Lys Leu Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val Val  
 355 360 365  
 Arg Gln Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu Ser  
 370 375 380  
 Thr Ile Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala Glu  
 385 390 395 400  
 Glu Met Val Glu Met Leu Thr Asp Arg Asn Leu Asn Leu Glu Glu Lys  
 405 410 415  
  
 Val Arg Glu Leu Arg Glu Thr Val Gly Asp Leu Glu Ala Met Asn Glu  
 420 425 430  
 Met Asn Asp Glu Leu Gln Glu Asn Ala Arg Glu Thr Glu Leu Glu Leu  
 435 440 445  
 Arg Glu Gln Leu Asp Met Ala Gly Ala Arg Val Arg Glu Ala Gln Lys  
 450 455 460  
 Arg Val Glu Ala Ala Gln Glu Thr Val Ala Asp Tyr Gln Gln Thr Ile  
 465 470 475 480  
  
 Lys Lys Tyr Arg Gln Leu Thr Ala His Leu Gln Asp Val Asn Arg Glu



Gly Gln Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr  
 740 745 750

Ser Cys Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met  
 755 760 765

Pro Gly Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro  
 770 775 780

Gln Val Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp Val  
 785 790 795 800

Val Ala Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile Ala  
 805 810 815

Pro Leu Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu Leu  
 820 825 830

Ala Phe Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser Pro  
 835 840 845

Tyr Glu Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met Asn  
 850 855 860

Lys Leu Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg Pro  
 865 870 875 880

Pro Ser Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg Ala  
 885 890 895

Glu Ile Thr Asp Ala Glu Gly Leu Gly Leu Lys Leu Glu Asp Arg Glu  
 900 905 910

Thr Val Ile Lys Glu Leu Lys Lys Ser Leu Lys Ile Lys Gly Glu Glu  
 915 920 925

Leu Ser Glu Ala Asn Val Arg Leu Ser Leu Leu Glu Lys Lys Leu Asp  
 930 935 940

Ser Ala Ala Lys Asp Ala Asp Glu Arg Ile Glu Lys Val Gln Thr Arg  
 945 950 955 960

Leu Glu Glu Thr Gln Ala Leu Leu Arg Lys Lys Glu Lys Glu Phe Glu  
 965 970 975

Glu Thr Met Asp Ala Leu Gln Ala Asp Ile Asp Gln Leu Glu Ala Glu

980	985	990
Lys Ala Glu Leu Lys Gln Arg Leu	Asn Ser Gln Ser Lys	Arg Thr Ile
995	1000	1005
Glu Gly Leu Arg Gly Pro Pro	Pro Ser Gly Ile Ala	Thr Leu Val
1010	1015	1020
Ser Gly Ile Ala Gly Gly Ala	Ile Pro Gly Gln Ala	Pro Gly Ser
1025	1030	1035
Val Pro Gly Pro Gly Leu Val	Lys Asp Ser Pro Leu	Leu Leu Gln
1040	1045	1050
Gln Ile Ser Ala Met Arg Leu	His Ile Ser Gln Leu	Gln His Glu
1055	1060	1065
Asn Ser Ile Leu Lys Gly Ala	Gln Met Lys Ala Ser	Leu Ala Ser
1070	1075	1080
Leu Pro Pro Leu His Val Ala	Lys Leu Ser His Glu	Gly Pro Gly
1085	1090	1095
Ser Glu Leu Pro Ala Gly Ala	Leu Tyr Arg Lys Thr	Ser Gln Leu
1100	1105	1110
Leu Glu Thr Leu Asn Gln Leu	Ser Thr His Thr His	Val Val Asp
1115	1120	1125
Ile Thr Arg Thr Ser Pro Ala	Ala Lys Ser Pro Ser	Ala Gln Leu
1130	1135	1140
Met Glu Gln Val Ala Gln Leu	Lys Ser Leu Ser Asp	Thr Val Glu
1145	1150	1155
Lys Leu Lys Asp Glu Val Leu	Lys Glu Thr Val Ser	Gln Arg Pro
1160	1165	1170
Gly Ala Thr Val Pro Thr Asp	Phe Ala Thr Phe Pro	Ser Ser Ala
1175	1180	1185
Phe Leu Arg Ala Lys Glu Glu	Gln Gln Asp Asp Thr	Val Tyr Met
1190	1195	1200
Gly Lys Val Thr Phe Ser Cys	Ala Ala Gly Phe Gly	Gln Arg His
1205	1210	1215

Arg Leu Val Leu Thr Gln Glu Gln Leu His Gln Leu His Ser Arg  
 1220 1225 1230

Leu Ile Ser  
 1235

<210> 30

<211> 1271

<212> PRT

<213> Homo sapiens

<400> 30

Met Ala Gln Ser Lys Arg His Val Tyr Ser Arg Thr Pro Ser Gly Ser  
 1 5 10 15

Arg Met Ser Ala Glu Ala Ser Ala Arg Pro Leu Arg Val Gly Ser Arg  
 20 25 30

Val Glu Val Ile Gly Lys Gly His Arg Gly Thr Val Ala Tyr Val Gly  
 35 40 45

Ala Thr Leu Phe Ala Thr Gly Lys Trp Val Gly Val Ile Leu Asp Glu  
 50 55 60

Ala Lys Gly Lys Asn Asp Gly Thr Val Gln Gly Arg Lys Tyr Phe Thr  
 65 70 75 80

Cys Asp Glu Gly His Gly Ile Phe Val Arg Gln Ser Gln Ile Gln Val  
 85 90 95

Phe Glu Asp Gly Ala Asp Thr Thr Ser Pro Glu Thr Pro Asp Ser Ser  
 100 105 110

Ala Ser Lys Val Leu Lys Arg Glu Gly Thr Asp Thr Thr Ala Lys Thr  
 115 120 125

Ser Lys Leu Ala Pro Thr Ala Arg Lys Thr Thr Thr Arg Arg Pro Lys  
 130 135 140

Pro Thr Arg Pro Ala Ser Thr Gly Val Ala Gly Ala Ser Ser Ser Leu  
 145 150 155 160

Gly Pro Ser Gly Ser Ala Ser Ala Gly Glu Leu Ser Ser Ser Glu Pro  
 165 170 175

Ser Thr Pro Ala Gln Thr Pro Leu Ala Ala Pro Ile Ile Pro Thr Pro  
 180 185 190  
 Val Leu Thr Ser Pro Gly Ala Val Pro Pro Leu Pro Ser Pro Ser Lys  
 195 200 205  
 Glu Glu Glu Gly Leu Arg Ala Gln Val Arg Asp Leu Glu Glu Lys Leu  
 210 215 220  
 Glu Thr Leu Arg Leu Lys Arg Ala Glu Asp Lys Ala Lys Leu Lys Glu  
 225 230 235 240  
 Leu Glu Lys His Lys Ile Gln Leu Glu Gln Val Gln Glu Trp Lys Ser  
 245 250 255  
 Lys Met Gln Glu Gln Gln Ala Asp Leu Gln Arg Arg Leu Lys Glu Ala  
 260 265 270  
 Arg Lys Glu Ala Lys Glu Ala Leu Glu Ala Lys Glu Arg Tyr Met Glu  
 275 280 285  
 Glu Met Ala Asp Thr Ala Asp Ala Ile Glu Met Ala Thr Leu Asp Lys  
 290 295 300  
 Glu Met Ala Glu Glu Arg Ala Glu Ser Leu Gln Gln Glu Val Glu Ala  
 305 310 315 320  
 Leu Lys Glu Arg Val Asp Glu Leu Thr Thr Asp Leu Glu Ile Leu Lys  
 325 330 335  
 Ala Glu Ile Glu Glu Lys Gly Ser Asp Gly Ala Ala Ser Ser Tyr Gln  
 340 345 350  
 Leu Lys Gln Leu Glu Glu Gln Asn Ala Arg Leu Lys Asp Ala Leu Val  
 355 360 365  
 Arg Met Arg Asp Leu Ser Ser Ser Glu Lys Gln Glu His Val Lys Leu  
 370 375 380  
 Gln Lys Leu Met Glu Lys Lys Asn Gln Glu Leu Glu Val Val Arg Gln  
 385 390 395 400  
 Gln Arg Glu Arg Leu Gln Glu Glu Leu Ser Gln Ala Glu Ser Thr Ile  
 405 410 415  
 Asp Glu Leu Lys Glu Gln Val Asp Ala Ala Leu Gly Ala Glu Glu Met



Asp Val Tyr Lys Lys Val Gly Ser Leu Tyr Pro Glu Met Ser Ala His  
 675 680 685  
 Glu Arg Ser Leu Asp Phe Leu Ile Glu Leu Leu His Lys Asp Gln Leu  
 690 695 700  
 Asp Glu Thr Val Asn Val Glu Pro Leu Thr Lys Ala Ile Lys Tyr Tyr  
 705 710 715 720  
 Gln His Leu Tyr Ser Ile His Leu Ala Glu Gln Pro Glu Asp Cys Thr  
 725 730 735  
 Met Gln Leu Ala Asp His Ile Lys Phe Thr Gln Ser Ala Leu Asp Cys  
 740 745 750  
 Met Ser Val Glu Val Gly Arg Leu Arg Ala Phe Leu Gln Gly Gly Gln  
 755 760 765  
 Glu Ala Thr Asp Ile Ala Leu Leu Leu Arg Asp Leu Glu Thr Ser Cys  
 770 775 780  
 Ser Asp Ile Arg Gln Phe Cys Lys Lys Ile Arg Arg Arg Met Pro Gly  
 785 790 795 800  
 Thr Asp Ala Pro Gly Ile Pro Ala Ala Leu Ala Phe Gly Pro Gln Val  
 805 810 815  
 Ser Asp Thr Leu Leu Asp Cys Arg Lys His Leu Thr Trp Val Val Ala  
 820 825 830  
 Val Leu Gln Glu Val Ala Ala Ala Ala Ala Gln Leu Ile Ala Pro Leu  
 835 840 845  
 Ala Glu Asn Glu Gly Leu Leu Val Ala Ala Leu Glu Glu Leu Ala Phe  
 850 855 860  
 Lys Ala Ser Glu Gln Ile Tyr Gly Thr Pro Ser Ser Ser Pro Tyr Glu  
 865 870 875 880  
 Cys Leu Arg Gln Ser Cys Asn Ile Leu Ile Ser Thr Met Asn Lys Leu  
 885 890 895  
 Ala Thr Ala Met Gln Glu Gly Glu Tyr Asp Ala Glu Arg Pro Pro Ser  
 900 905 910  
 Lys Pro Pro Pro Val Glu Leu Arg Ala Ala Ala Leu Arg Ala Glu Ile



Thr His Val Val Asp Ile Thr Arg Thr Ser Pro Ala Ala Lys Ser  
 1160 1165 1170

Pro Ser Ala Gln Leu Met Glu Gln Val Ala Gln Leu Lys Ser Leu  
 1175 1180 1185

Ser Asp Thr Val Glu Lys Leu Lys Asp Glu Val Leu Lys Glu Thr  
 1190 1195 1200

Val Ser Gln Arg Pro Gly Ala Thr Val Pro Thr Asp Phe Ala Thr  
 1205 1210 1215

Phe Pro Ser Ser Ala Phe Leu Arg Ala Lys Glu Glu Gln Gln Asp  
 1220 1225 1230

Asp Thr Val Tyr Met Gly Lys Val Thr Phe Ser Cys Ala Ala Gly  
 1235 1240 1245

Phe Gly Gln Arg His Arg Leu Val Leu Thr Gln Glu Gln Leu His  
 1250 1255 1260

Gln Leu His Ser Arg Leu Ile Ser  
 1265 1270

<210> 31

<211> 1114

<212> PRT

<213> Homo sapiens

<400> 31

Met Ala Lys Ala Thr Ser Gly Ala Ala Gly Leu Arg Leu Leu Leu Leu  
 1 5 10 15

Leu Leu Leu Pro Leu Leu Gly Lys Val Ala Leu Gly Leu Tyr Phe Ser  
 20 25 30

Arg Asp Ala Tyr Trp Glu Lys Leu Tyr Val Asp Gln Ala Ala Gly Thr  
 35 40 45

Pro Leu Leu Tyr Val His Ala Leu Arg Asp Ala Pro Glu Glu Val Pro  
 50 55 60

Ser Phe Arg Leu Gly Gln His Leu Tyr Gly Thr Tyr Arg Thr Arg Leu  
 65 70 75 80

His Glu Asn Asn Trp Ile Cys Ile Gln Glu Asp Thr Gly Leu Leu Tyr



Glu Thr Ser Val Gln Ala Asn Gly Ser Phe Val Arg Ala Thr Val His  
 340 345 350  
 Asp Tyr Arg Leu Val Leu Asn Arg Asn Leu Ser Ile Ser Glu Asn Arg  
 355 360 365  
 Thr Met Gln Leu Ala Val Leu Val Asn Asp Ser Asp Phe Gln Gly Pro  
 370 375 380  
 Gly Ala Gly Val Leu Leu Leu His Phe Asn Val Ser Val Leu Pro Val  
 385 390 395 400  
 Ser Leu His Leu Pro Ser Thr Tyr Ser Leu Ser Val Ser Arg Arg Ala  
 405 410 415  
 Arg Arg Phe Ala Gln Ile Gly Lys Val Cys Val Glu Asn Cys Gln Ala  
 420 425 430  
 Phe Ser Gly Ile Asn Val Gln Tyr Lys Leu His Ser Ser Gly Ala Asn  
 435 440 445  
 Cys Ser Thr Leu Gly Val Val Thr Ser Ala Glu Asp Thr Ser Gly Ile  
 450 455 460  
 Leu Phe Val Asn Asp Thr Lys Ala Leu Arg Arg Pro Lys Cys Ala Glu  
 465 470 475 480  
 Leu His Tyr Met Val Val Ala Thr Asp Gln Gln Thr Ser Arg Gln Ala  
 485 490 495  
 Gln Ala Gln Leu Leu Val Thr Val Glu Gly Ser Tyr Val Ala Glu Glu  
 500 505 510  
 Ala Gly Cys Pro Leu Ser Cys Ala Val Ser Lys Arg Arg Leu Glu Cys  
 515 520 525  
 Glu Glu Cys Gly Gly Leu Gly Ser Pro Thr Gly Arg Cys Glu Trp Arg  
 530 535 540  
 Gln Gly Asp Gly Lys Gly Ile Thr Arg Asn Phe Ser Thr Cys Ser Pro  
 545 550 555 560  
 Ser Thr Lys Thr Cys Pro Asp Gly His Cys Asp Val Val Glu Thr Gln  
 565 570 575  
 Asp Ile Asn Ile Cys Pro Gln Asp Cys Leu Arg Gly Ser Ile Val Gly



Arg Asn Ser Ser Ser Leu Asp His Pro Asp Glu Arg Ala Leu Thr Met  
 835 840 845  
 Gly Asp Leu Ile Ser Phe Ala Trp Gln Ile Ser Gln Gly Met Gln Tyr  
 850 855 860  
 Leu Ala Glu Met Lys Leu Val His Arg Asp Leu Ala Ala Arg Asn Ile  
 865 870 875 880  
 Leu Val Ala Glu Gly Arg Lys Met Lys Ile Ser Asp Phe Gly Leu Ser  
 885 890 895  
 Arg Asp Val Tyr Glu Glu Asp Ser Tyr Val Lys Arg Ser Gln Gly Arg  
 900 905 910  
 Ile Pro Val Lys Trp Met Ala Ile Glu Ser Leu Phe Asp His Ile Tyr  
 915 920 925  
 Thr Thr Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile  
 930 935 940  
 Val Thr Leu Gly Gly Asn Pro Tyr Pro Gly Ile Pro Pro Glu Arg Leu  
 945 950 955 960  
 Phe Asn Leu Leu Lys Thr Gly His Arg Met Glu Arg Pro Asp Asn Cys  
 965 970 975  
 Ser Glu Glu Met Tyr Arg Leu Met Leu Gln Cys Trp Lys Gln Glu Pro  
 980 985 990  
 Asp Lys Arg Pro Val Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys Met  
 995 1000 1005  
 Met Val Lys Arg Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr Pro  
 1010 1015 1020  
 Ser Asp Ser Leu Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu Thr  
 1025 1030 1035  
 Pro Leu Val Asp Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu Pro  
 1040 1045 1050  
 Ser Thr Trp Ile Glu Asn Lys Leu Tyr Gly Met Ser Asp Pro Asn  
 1055 1060 1065  
 Trp Pro Gly Glu Ser Pro Val Pro Leu Thr Arg Ala Asp Gly Thr

1070                      1075                      1080  
 Asn Thr Gly Phe Pro Arg Tyr Pro Asn Asp Ser Val Tyr Ala Asn  
 1085                      1090                      1095  
 Trp Met Leu Ser Pro Ser Ala Ala Lys Leu Met Asp Thr Phe Asp  
 1100                      1105                      1110  
 Ser

<210> 32

<211> 1072

<212> PRT

<213>

> Homo sapiens

<400> 32

Met Ala Lys Ala Thr Ser Gly Ala Ala Gly Leu Arg Leu Leu Leu Leu  
 1                      5                      10                      15  
 Leu Leu Leu Pro Leu Leu Gly Lys Val Ala Leu Gly Leu Tyr Phe Ser  
                     20                      25                      30  
 Arg Asp Ala Tyr Trp Glu Lys Leu Tyr Val Asp Gln Ala Ala Gly Thr  
                     35                      40                      45  
 Pro Leu Leu Tyr Val His Ala Leu Arg Asp Ala Pro Glu Glu Val Pro  
                     50                      55                      60  
  
 Ser Phe Arg Leu Gly Gln His Leu Tyr Gly Thr Tyr Arg Thr Arg Leu  
 65                      70                      75                      80  
 His Glu Asn Asn Trp Ile Cys Ile Gln Glu Asp Thr Gly Leu Leu Tyr  
                     85                      90                      95  
 Leu Asn Arg Ser Leu Asp His Ser Ser Trp Glu Lys Leu Ser Val Arg  
                     100                      105                      110  
 Asn Arg Gly Phe Pro Leu Leu Thr Val Tyr Leu Lys Val Phe Leu Ser  
                     115                      120                      125  
  
 Pro Thr Ser Leu Arg Glu Gly Glu Cys Gln Trp Pro Gly Cys Ala Arg  
                     130                      135                      140  
 Val Tyr Phe Ser Phe Phe Asn Thr Ser Phe Pro Ala Cys Ser Ser Leu  
 145                      150                      155                      160

Lys Pro Arg Glu Leu Cys Phe Pro Glu Thr Arg Pro Ser Phe Arg Ile  
 165 170 175

Arg Glu Asn Arg Pro Pro Gly Thr Phe His Gln Phe Arg Leu Leu Pro  
 180 185 190

Val Gln Phe Leu Cys Pro Asn Ile Ser Val Ala Tyr Arg Leu Leu Glu  
 195 200 205

Gly Glu Gly Leu Pro Phe Arg Cys Ala Pro Asp Ser Leu Glu Val Ser  
 210 215 220

Thr Arg Trp Ala Leu Asp Arg Glu Gln Arg Glu Lys Tyr Glu Leu Val  
 225 230 235 240

Ala Val Cys Thr Val His Ala Gly Ala Arg Glu Glu Val Val Met Val  
 245 250 255

Pro Phe Pro Val Thr Val Tyr Asp Glu Asp Asp Ser Ala Pro Thr Phe  
 260 265 270

Pro Ala Gly Val Asp Thr Ala Ser Ala Val Val Glu Phe Lys Arg Lys  
 275 280 285

Glu Asp Thr Val Val Ala Thr Leu Arg Val Phe Asp Ala Asp Val Val  
 290 295 300

Pro Ala Ser Gly Glu Leu Val Arg Arg Tyr Thr Ser Thr Leu Leu Pro  
 305 310 315 320

Gly Asp Thr Trp Ala Gln Gln Thr Phe Arg Val Glu His Trp Pro Asn  
 325 330 335

Glu Thr Ser Val Gln Ala Asn Gly Ser Phe Val Arg Ala Thr Val His  
 340 345 350

Asp Tyr Arg Leu Val Leu Asn Arg Asn Leu Ser Ile Ser Glu Asn Arg  
 355 360 365

Thr Met Gln Leu Ala Val Leu Val Asn Asp Ser Asp Phe Gln Gly Pro  
 370 375 380

Gly Ala Gly Val Leu Leu Leu His Phe Asn Val Ser Val Leu Pro Val  
 385 390 395 400

Ser Leu His Leu Pro Ser Thr Tyr Ser Leu Ser Val Ser Arg Arg Ala



Ile His Cys Tyr His Lys Phe Ala His Lys Pro Pro Ile Ser Ser Ala  
 660 665 670  
 Glu Met Thr Phe Arg Arg Pro Ala Gln Ala Phe Pro Val Ser Tyr Ser  
 675 680 685  
 Ser Ser Gly Ala Arg Arg Pro Ser Leu Asp Ser Met Glu Asn Gln Val  
 690 695 700  
  
 Ser Val Asp Ala Phe Lys Ile Leu Glu Asp Pro Lys Trp Glu Phe Pro  
 705 710 715 720  
 Arg Lys Asn Leu Val Leu Gly Lys Thr Leu Gly Glu Gly Glu Phe Gly  
 725 730 735  
 Lys Val Val Lys Ala Thr Ala Phe His Leu Lys Gly Arg Ala Gly Tyr  
 740 745 750  
 Thr Thr Val Ala Val Lys Met Leu Lys Glu Asn Ala Ser Pro Ser Glu  
 755 760 765  
  
 Leu Arg Asp Leu Leu Ser Glu Phe Asn Val Leu Lys Gln Val Asn His  
 770 775 780  
 Pro His Val Ile Lys Leu Tyr Gly Ala Cys Ser Gln Asp Gly Pro Leu  
 785 790 795 800  
 Leu Leu Ile Val Glu Tyr Ala Lys Tyr Gly Ser Leu Arg Gly Phe Leu  
 805 810 815  
 Arg Glu Ser Arg Lys Val Gly Pro Gly Tyr Leu Gly Ser Gly Gly Ser  
 820 825 830  
  
 Arg Asn Ser Ser Ser Leu Asp His Pro Asp Glu Arg Ala Leu Thr Met  
 835 840 845  
 Gly Asp Leu Ile Ser Phe Ala Trp Gln Ile Ser Gln Gly Met Gln Tyr  
 850 855 860  
 Leu Ala Glu Met Lys Leu Val His Arg Asp Leu Ala Ala Arg Asn Ile  
 865 870 875 880  
 Leu Val Ala Glu Gly Arg Lys Met Lys Ile Ser Asp Phe Gly Leu Ser  
 885 890 895  
  
 Arg Asp Val Tyr Glu Glu Asp Ser Tyr Val Lys Arg Ser Gln Gly Arg

900 905 910  
 Ile Pro Val Lys Trp Met Ala Ile Glu Ser Leu Phe Asp His Ile Tyr  
 915 920 925  
 Thr Thr Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile  
 930 935 940  
 Val Thr Leu Gly Gly Asn Pro Tyr Pro Gly Ile Pro Pro Glu Arg Leu  
 945 950 955 960

Phe Asn Leu Leu Lys Thr Gly His Arg Met Glu Arg Pro Asp Asn Cys  
 965 970 975  
 Ser Glu Glu Met Tyr Arg Leu Met Leu Gln Cys Trp Lys Gln Glu Pro  
 980 985 990  
 Asp Lys Arg Pro Val Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys Met  
 995 1000 1005  
 Met Val Lys Arg Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr Pro  
 1010 1015 1020

Ser Asp Ser Leu Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu Thr  
 1025 1030 1035  
 Pro Leu Val Asp Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu Pro  
 1040 1045 1050  
 Ser Thr Trp Ile Glu Asn Lys Leu Tyr Gly Arg Ile Ser His Ala  
 1055 1060 1065

Phe Thr Arg Phe  
 1070

<210> 33

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 33

tgtccagctt tgtgctgat tgatgt

26

<210> 34

<211> 26

<212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 34  
 gctgggcact gaagagaaag gaatgc 26  
 <210> 35  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 35  
 agcaggatga gtgcggagc aagc 24  
 <210> 36  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence  
 <220><  
 223> primer  
 <400> 36  
 ttaactatca aacgtgtcca ttaatittgc cgc 33  
 <210> 37  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 37  
 agtactgggg tggtggg 18  
 <210> 38  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 38  
 cactttggac aaggatg 19  
 <210> 39

<211> 18  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> primer  
 <400> 39  
 acagaactgg agctgcgg 18  
 <210> 40  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 40  
 ggactggtgt actcgtg 18  
 <210> 41  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 41  
 tcctagactg caggaaacac 20  
 <210> 42  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> primer  
 <400> 42  
 catcgagaaa gtccagac 18  
 <210> 43  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 43

gctgctggag acattgaa 18  
 <210> 44  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 44  
 tcactgctgc tcagctca 18  
 <210> 45  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 45  
 gaggatccaa agtgggaatt 20  
 <210> 46  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 46  
 agtatctggc cgagatgaag 20  
 <210> 47  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 47  
 gcaaagacct ggagaagatg 20  
 <210> 48  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220><223> primer  
 <400> 48  
 aggacgttga actctgacag 20  
 <210> 49  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 49  
 cctttgcttc atccagaatc 20  
 <210> 50  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 50  
 gattttgtgt ttctccagct ct 22  
 <210> 51  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> primer  
 <400> 51  
 cctgcttctc tgaggaagaa 20  
 <210> 52  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 52  
 gggccttagt ctcagcaaac 20  
 <210> 53  
 <211> 20  
 <212> DNA

<213> Artificial Sequence  
 <220><223> primer  
 <400> 53  
 gagcactctg cgigaactta 20  
 <210> 54  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> primer  
 <400> 54  
 cagcttggtc atgtactga t 21  
 <210> 55  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 55  
 tggtagtcc ttcaccag 18  
 <210> 56  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 56  
 cctagagttt ttccaagaac ca 22  
 <210> 57  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> primer  
 <400> 57  
 catttaactg gaatccgacc 20  
 <210> 58

<211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 58  
 gactctctcc aggccagttc 20  
 <210> 59  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 59  
 ggctatcaga agtaaaacca cc 22  
 <210> 60  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> primer  
 <400> 60  
 cgagagctga tggcacta 18  
 <210> 61  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 61  
 cttcatcaca agtgaagtac ttcc 24  
 <210> 62  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 62  
 cg tactccac gatgaggag 19

<210> 63  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> primer  
 <400> 63  
 gattctggat gaagcaaagg 20  
 <210> 64  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 64  
 ggaagtactt cacttgtgat gaag 24  
 <210> 65  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 65  
 cccagccacc ccagtact 18  
 <210> 66  
 <211> 17  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> primer  
 <400> 66  
 gtaaaacgac ggccagt 17  
 <210> 67  
 <211> 17  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer

<400> 67  
 gttttccag tcacgac 17  
 <210> 68  
 <211> 17  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 68  
 caggaaacag ctatgac 17  
 <210> 69  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> primer  
 <400> 69  
 ctggagccac agtaccact 20  
 <210> 70  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 70  
 tccaaattcg cttctccta 20  
 <210> 71  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 71  
 ttcacagcc ttctcaggg aggat 25  
 <210> 72  
 <211> 41  
 <212> DNA

<213> Artificial Sequence

<220><223> Primer

<400> 72

ggggacaagt ttgtacaaaa aagcaggctt cgccaccagc a 41

<210> 73

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 73

ggggaccact ttgtacaaga aagctgggtt ttaactatca aa 42

<210> 74

<211> 21

<212> RNA

<213> Artificial Sequence

<220><223> RET siRNA1

<220><221> misc\_feature

<223> n stand for thymine

<220><221> misc\_feature

<222> (20)..(21)

<223> n stand for thymine

<400

> 74

cacaugucau caaaauguan n 21

<210> 75

<211> 21

<212> RNA

<213> Artificial Sequence

<220><223> RET siRNA2

<220><221> misc\_feature

<222> (20)..(21)

<223> n stand for thymine

<400> 75

ggauugaaaa caaacucuan n 21

<210> 76  
 <211> 21  
 <212> RNA  
 <213> Artificial Sequence  
 <220><223> RET siRNA  
 <220><221> misc\_feature  
 <222> (20)..(21)  
 <223> n stand for thymine  
 <400> 76

gcuugucccg agauguuuan n 21

<210> 77  
 <211> 21  
 <212> RNA  
 <213> Artificial Sequence  
 <220><223> RET siRNA3  
 <220><221> misc\_feature  
 <222> (20)..(21)  
 <223> n stand for thymine  
 <400> 77

ccacugcuac cacaaguun n 21