



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(51) Int. Cl.

*C12N 15/31* (2006.01)*C12N 15/63* (2006.01)*A61K 48/00* (2006.01)

(45) 공고일자

2006년12월01일

(11) 등록번호

10-0652513

(24) 등록일자

2006년11월24일

(21) 출원번호	10-1999-7010775	(65) 공개번호	10-2001-0012812
(22) 출원일자	1999년11월20일	(43) 공개일자	2001년02월26일
심사청구일자	2003년05월19일		
번역문 제출일자	1999년11월20일		
(86) 국제출원번호	PCT/US1998/010407	(87) 국제공개번호	WO 1998/53075
국제출원일자	1998년05월20일	국제공개일자	1998년11월26일

(81) 지정국

국내특허 : 알바니아, 아르메니아, 오스트리아, 오스트레일리아, 보스니아 헤르체고비나, 바베이도스, 불가리아, 브라질, 벨라루스, 캐나다, 스위스, 중국, 쿠바, 체코, 독일, 덴마크, 에스토니아, 스페인, 핀란드, 영국, 그루지야, 헝가리, 이스라엘, 아이슬란드, 일본, 케냐, 키르키즈스탄, 북한, 대한민국, 카자흐스탄, 세인트루시아, 스리랑카, 리베이라, 레소토, 리투아니아, 룩셈부르크, 라트비아, 몰도바, 마다가스카르, 마케도니아공화국, 몽고, 말라위, 멕시코, 노르웨이, 뉴질랜드, 슬로베니아, 슬로바키아, 타지키스탄, 투르크멘, 터아키, 트리니아드토바고, 우크라이나, 우간다, 우즈베키스탄, 베트남, 폴란드, 포르투칼, 루마니아, 러시아, 수단, 스웨덴, 싱가포르, 가나, 감비아, 기니 비사우, 인도네시아, 시에라리온, 세르비아 앤 몬테네그로, 짐바브웨,

AP ARIPO특허 : 케냐, 레소토, 말라위, 수단, 스와질랜드, 우간다, 가나, 짐바브웨, 감비아,

EA 유라시아특허 : 아르메니아, 아제르바이잔, 벨라루스, 키르키즈스탄, 카자흐스탄, 몰도바, 러시아, 타지키스탄, 투르크멘,

EP 유럽특허 : 오스트리아, 벨기에, 스위스, 독일, 덴마크, 스페인, 프랑스, 영국, 그리스, 아일랜드, 이탈리아, 룩셈부르크, 모나코, 네덜란드, 포르투칼, 스웨덴, 사이프러스, 핀란드,

OA OAPI특허 : 부르키나파소, 베넌, 중앙아프리카, 콩고, 코트디브와르, 카메룬, 가봉, 기니, 말리, 모리타니, 니제르, 세네갈, 차드, 토고,

(30) 우선권주장

08/859,381

1997년05월20일

미국(US)

09/073,010

1998년05월05일

미국(US)

(73) 특허권자

코리사 코퍼레이션

미국, 워싱톤 98104, 시애틀 콜롬비아 스트리트 1124 스윗 200

(72) 발명자

앨더슨마크알

미국워싱턴주98110베인브리지아일랜드그로우애비뉴노스웨스트1116

딜론다빈씨

미국워싱تون주98053레드몬드노스이스트24스트리트21607

스케이키야시에이더블유

미국워싱تون주98117시애틀25애비뉴노스웨스트8327

캠프스-네토안토니오  
미국워싱턴주98110베인브리지아일랜드노스이스트미드쉽코트9308

(74) 대리인

김성기  
신정건  
나영환

심사관 : 정의준

전체 청구항 수 : 총 30 항

---

#### (54) 결핵의 면역치료 및 진단용 화합물과 이의 사용 방법

---

(57) 요약

본 발명은 결핵에 대한 예방 면역성을 유도하기 위한 화합물 및 방법에 관한 것이다. 본 발명의 화합물은 1종 이상의 엠. 투베르콜로시스(*M. tuberculosis*) 단백질의 하나 이상의 면역원성 부분을 포함하는 폴리펩티드 및 이 폴리펩티드를 코딩하는 DNA 분자를 포함한다. 이러한 화합물은 엠. 투베르콜로시스 감염에 대한 면역화용 백신 및/또는 약학 조성물로 제형화될 수 있으며, 또한 결핵을 진단하는 데 사용될 수 있다.

대표도

도 1a

#### 특허청구의 범위

청구항 1.

마이코박테리움 투베르콜로시스(*M. tuberculosis*) 항원의 면역원성 부분을 포함하는 분리된 폴리펩티드로서, 상기 항원은 서열 번호 109에 개시된 아미노산 서열로 구성되는 것인 분리된 폴리펩티드.

청구항 2.

제1항에 있어서, 마이코박테리움 투베르콜로시스 항원을 포함하며, 상기 항원은 서열 번호 109에 개시된 아미노산 서열로 구성되는 것인 분리된 폴리펩티드.

청구항 3.

제1항에 있어서, 서열 번호 110~124 중 어느 하나에 개시된 아미노산 서열을 포함하는 것인 분리된 폴리펩티드.

청구항 4.

제3항에 있어서, 서열 번호 119에 개시된 아미노산 서열을 포함하는 것인 분리된 폴리펩티드.

**청구항 5.**

제4항에 있어서, 서열 번호 119에 개시된 아미노산 서열로 구성되는 것인 분리된 폴리펩티드.

**청구항 6.**

제3항에 있어서, 서열 번호 120에 개시된 아미노산 서열을 포함하는 것인 분리된 폴리펩티드.

**청구항 7.**

제4항에 있어서, 서열 번호 120에 개시된 아미노산 서열로 구성되는 것인 분리된 폴리펩티드.

**청구항 8.**

제2항에 있어서, 마이코박테리움 투베르콜로시스 항원으로 구성되며, 상기 항원은 서열 번호 109에 개시된 아미노산 서열로 구성되는 것인 분리된 폴리펩티드.

**청구항 9.**

마이코박테리움 투베르콜로시스 항원의 면역원성 부분을 포함하는 폴리펩티드를 코딩하는 뉴클레오티드 서열을 포함하는 DNA 분자로서, 상기 항원은 서열 번호 109에 개시된 아미노산 서열로 구성되는 것인 DNA 분자.

**청구항 10.**

제9항에 있어서, 서열 번호 108에 개시된 뉴클레오티드 서열을 포함하는 것인 DNA 분자.

**청구항 11.**

제10항에 있어서, 서열 번호 108에 개시된 뉴클레오티드 서열로 구성되는 것인 DNA 분자.

**청구항 12.**

제10항에 기재된 DNA 분자를 포함하는 발현 벡터.

**청구항 13.**

제12항에 기재된 발현 벡터로 형질전환시킨 분리된 숙주 세포.

**청구항 14.**

제13항에 있어서, 숙주 세포는 이. 콜라이(*E. coli*), 효모 및 포유류 세포로 구성된 군에서 선택되는 것인 숙주 세포.

### 청구항 15.

제1항 내지 제8항 중 어느 한 항에 기재된 1종 이상의 폴리펩티드 및 생리적 허용 담체를 포함하는, 마이코박테리움 투베르콜로시스에 대한 예방 면역성을 유도하기 위한 약학 조성물.

### 청구항 16.

제9항 내지 제11항 중 어느 한 항에 기재된 1종 이상의 DNA 분자 및 생리적 허용 담체를 포함하는, 마이코박테리움 투베르콜로시스에 대한 예방 면역성을 유도하기 위한 약학 조성물.

### 청구항 17.

제1항 내지 제8항 중 어느 한 항에 기재된 1종 이상의 폴리펩티드 및 비특이적 면역 반응 인핸서를 포함하는, 예방 면역성을 유도하기 위한 백신.

### 청구항 18.

제17항에 있어서, 비특이적 면역 반응 인핸서는 면역 증강제(adjuvant)인 백신.

### 청구항 19.

제9항 내지 제11항 중 어느 한 항에 기재된 1종 이상의 DNA 분자 및 비특이적 면역 반응 인핸서를 포함하는, 예방 면역성을 유도하기 위한 백신.

### 청구항 20.

제19항에 있어서, 비특이적 면역 반응 인핸서는 면역 증강제인 백신.

### 청구항 21.

제1항 내지 제8항 중 어느 한 항에 기재된 폴리펩티드를 2 이상 포함하는 융합 단백질.

### 청구항 22.

제1항 내지 제8항 중 어느 한 항에 기재된 폴리펩티드와 공지된 마이코박테리움 투베르콜로시스 항원을 포함하는 융합 단백질.

### 청구항 23.

제21항에 기재된 융합 단백질 및 생리적 허용 담체를 포함하는, 마이코박테리움 투베르콜로시스에 대한 예방 면역성을 유도하기 위한 약학 조성물.

### 청구항 24.

제22항에 기재된 융합 단백질 및 생리적 허용 담체를 포함하는, 마이코박테리움 투베르콜로시스에 대한 예방 면역성을 유도하기 위한 약학 조성물.

### 청구항 25.

제21항에 기재된 융합 단백질 및 비특이적 면역 반응 인핸서를 포함하는, 예방 면역성을 유도하기 위한 백신.

### 청구항 26.

제22항에 기재된 융합 단백질 및 비특이적 면역 반응 인핸서를 포함하는, 예방 면역성을 유도하기 위한 백신.

### 청구항 27.

제25항에 있어서, 비특이적 면역 반응 인핸서는 면역 증강제인 백신.

### 청구항 28.

제26항에 있어서, 비특이적 면역 반응 인핸서는 면역 증강제인 백신.

### 청구항 29.

(a) 제1항 내지 제8항 중 어느 한 항에 기재된 폴리펩티드; 및

(b) 환자의 피부 표면 상에 면역 반응을 유도하기 위하여 상기 폴리펩티드를 환자의 피부 세포와 접촉시키기에 충분한 장치

를 포함하는, 결핵 검출용 진단 키트.

### 청구항 30.

제29항에 있어서, 면역 반응이 경결인 진단 키트.

## 명세서

### 기술분야

본 발명은 일반적으로 마이코박테리움 투베르콜로시스(*Mycobacterium tuberculosis*) 감염을 검출, 치료 및 예방하는 방법에 관한 것이다. 더욱 구체적으로 본 발명은 마이코박테리움 투베르콜로시스 항원, 또는 이의 일부 또는 기타 변형체를 포함하는 폴리펩티드와, 마이코박테리움 투베르콜로시스 감염에 대한 진단 및 백신 접종에 사용하기 위한 폴리펩티드의 용도에 관한 것이다.

### 배경기술

결핵은 일반적으로 마이코박테리움 투베르콜로시스에 감염되어 유발되는 만성 감염성 질병이다. 결핵은 개발도상국에서 주요 질병일 뿐 아니라 전세계 선진국에서도 그 문제성이 증가하고 있으며, 매년 약 8백만건의 신규 사례가 발생하여 3백

만명이 사망한다. 감염이 되어도 상당한 기간 동안 증상이 없을 수 있지만, 가장 일반적으로는 이 질병은 폐의 급성 염증을 보이며 열과 비생산적인 기침을 유발시킨다. 치료하지 않고 그대로 방치하면, 심각한 합병증을 초래하고 일반적으로는 사망하게 된다.

결핵은 일반적으로 장기간에 걸친 항생제 치료법을 이용하여 제어할 수 있지만, 이러한 치료는 질병이 유포되는 것을 막기에는 충분하지 않다. 감염된 개체는 무증상일 수 있지만, 때로는 전염성이 있다. 또한, 치료법에 순응하는 것이 중요하지만, 환자의 행동 방식을 감시하기는 어렵다. 일부 환자는 전 치료 과정을 종결짓지 못하는데, 이렇게 되면 치료 효과가 없으며, 내약제성이 유발될 수 있다.

결핵이 만연되는 것을 억제하는 데에는 효과적인 백신 접종과 질병의 정확한 조기 진단이 필요하다. 현재, 생박테리아로 백신 접종하는 것이 예방 면역을 유도하는 가장 효과적인 방법이다. 이러한 용도로 사용되는 가장 일반적인 마이코박테리움은 마이코박테리움 보비스(*Mycobacterium bovis*)의 무독성 균주인 바실러스 칼메트-구에린(BCG: *Bacillus Calmette-Guerin*)이다. 그러나, BCG의 안전성 및 효능은 논쟁의 대상이 되고 있어서, 미국 등의 일부 국가에서는 일반 대중에게 이 백신을 접종하지 않는다. 진단은 보통, 투베르쿨린 PPD(단백질 정제된 유도체)를 피내에 노출시키는 피부 테스트를 이용하여 행한다. 항원 특이적 T 세포 반응으로 인해 주사한 지 48~72 시간 후 주사 부위에 측정가능한 경결이 생기는지, 이 경결은 마이코박테리아의 항원에 노출되었음을 의미한다. 그러나, 이 테스트 방법의 문제점은 민감성 및 특이성이 있으며, BCG를 접종한 개체를 감염된 개체와 구별할 수가 없다.

대식세포는 엠. 투베르쿨로시스 면역성의 기본 효과기로서 작용하는 것으로 확인되었으며, T 세포는 이러한 면역성의 주요 유도자이다. 엠. 투베르쿨로시스 감염에 대한 보호에 있어서 T 세포가 필수적인 역할을 한다는 것은, 인간 면역결핍 바이러스(HIV) 감염과 관련이 있는 CD4 T 세포의 고갈로 인해 AIDS 환자에서 엠. 투베르쿨로시스가 자주 발생한다는 것으로 알 수 있다. 마이코박테리움 반응성 CD4 T 세포는 감마 인터페론(IFN- $\gamma$ )의 유효한 생산자인 것으로 확인되었으며, 또한 IFN- $\gamma$ 는 마우스에서 대식세포의 마이코박테리아 억제 효과를 촉발하는 것으로 확인되었다. 인간에서 IFN- $\gamma$ 의 역할은 마우스에서보다는 덜 명확하지만, 여러 연구로부터 1,25-디히드록시-비타민 D3가, 단독으로 또는 IFN- $\gamma$  또는 중양 괴사인자-알파와 함께, 인간 대식세포를 활성화시켜 엠. 투베르쿨로시스 감염을 억제한다는 것이 확인되었다. 또한, IFN- $\gamma$ 는 인간 대식세포가 1,25-디히드록시-비타민 D3을 생성하도록 자극하는 것으로 알려져 있다. 유사하게, IL-12는 엠. 투베르쿨로시스 감염에 대한 내성을 자극하는 역할을 하는 것으로 확인되었다. 엠. 투베르쿨로시스 감염의 면역학에 대해서는 Chan 및 Kaufmann 등의 문헌 [Tuberculosis: Pathogenesis, Protection and Control, Bloom (ed.), ASM Press, 1994년 워싱턴 DC]을 참조할 수 있다.

따라서, 결핵을 예방, 치료 및 검출하기 위한 개선된 백신 및 방법이 당업계에 필요하다. 본 발명은 이러한 요구를 만족시키고 나아가 관련된 다른 장점을 제공한다.

## 발명의 상세한 설명

### 발명의 개요

간단히 설명하면, 본 발명은 결핵의 예방 및 진단을 위한 화합물 및 방법을 제공한다. 제1 양태에서, 본 발명은 엠. 투베르쿨로시스 항원의 면역원성 부분, 또는 단지 보존적 치환 및/또는 변형에 있어서만 상이한 상기 항원의 변형체를 포함하는 폴리펩티드를 제공하는데, 상기 항원은 서열 번호 1, 11, 12, 83, 103-108, 125, 127, 129-137, 139 및 140에 개시된 서열, 이 서열의 상보 서열, 및 서열 번호 1, 11, 12, 83, 103-108, 125, 127, 129-137, 139 및 140에 개시된 서열 또는 이의 상보 서열에 보통의 염중 조건하에서 하이브리드화하는 DNA 서열로 구성된 군에서 선택되는 DNA 서열에 의해 코딩되는 아미노산 서열을 포함한다. 제2 양태에서, 본 발명은 서열 번호 16-33, 109, 126, 138, 141, 142에 제공된 서열 및 이의 변형체로 구성된 군에서 선택되는 아미노산 서열을 갖는 엠. 투베르쿨로시스 항원의 면역원성 부분을 포함하는 폴리펩티드를 제공한다.

관련 제3 양태에서, 전술한 폴리펩티드를 코딩하는 DNA 서열, 이를 DNA 서열을 포함하는 발현 벡터 및 이러한 발현 벡터로 형질전환되거나 또는 형질감염된 숙주 세포를 제공한다.

제4 양태에서, 본 발명은 본 발명의 제1 폴리펩티드 및 제2 폴리펩티드를 포함하는 융합 단백질, 또는 대안적으로 본 발명의 폴리펩티드 및 공지된 엠. 투베르쿨로시스 항원을 포함하는 융합 단백질을 제공한다.

제5 양태에서, 본 발명은 1종 이상의 상기 폴리펩티드, 또는 이러한 폴리펩티드를 코딩하는 DNA 분자, 및 생리적 허용 담체를 포함하는 약학 조성물을 제공한다. 또한, 본 발명은 전술한 1종 이상의 폴리펩티드 및 비특이적인 면역 반응 인핸서를 함유하는 백신을, 이러한 폴리펩티드를 코딩하는 1종 이상의 DNA 서열 및 비특이적 면역 반응 인핸서를 함유하는 백신과 함께 제공한다.

제6 양태에서, 본 발명은 1종 이상의 상기 폴리펩티드의 유효량을 환자에게 투여하는 단계를 포함하여 환자의 예방 면역성을 유도하는 방법을 제공한다.

본 발명의 제7 양태에서, 환자의 결핵을 검출하기 위한 방법 및 진단 키트를 제공한다. 이 방법은 상기 1종 이상의 폴리펩티드와 환자의 피부 세포를 접촉시키는 단계 및 환자의 피부상에서의 면역 반응을 검출하는 단계를 포함한다. 진단 키트는 1종 이상의 상기 폴리펩티드와 함께 환자의 피부 세포와 폴리펩티드를 접촉시키기에 충분한 장치를 포함한다.

제8 양태에서, 본 발명은 환자의 결핵을 검출하기 위한 방법을 제공하는데, 이 방법은 서열 번호 2-10, 102, 128에 개시된 서열, 이 서열의 상보 서열 및 서열 번호 2-10, 102, 128에 개시된 서열에 하이브리드화하는 DNA 서열로 구성된 군에서 선택되는 DNA 서열에 의해 코딩되는 1종 이상의 폴리펩티드와 환자의 피부 세포를 접촉시키는 단계; 및 환자의 피부에서의 면역 반응을 검출하는 단계를 포함한다. 이러한 방법에 사용하기 위한 진단 키트도 제공된다.

본 발명의 이들 양태 및 기타 양태는 후술하는 발명의 상세한 설명 및 첨부 도면을 참조하면 분명해질 것이다. 본원에서 언급한 모든 문헌은 그 전문을 본원에서 참고로 인용하였다.

## 실시예

### 실시예 1

#### 인간 PBMC에서 얻은 CD4+ T 세포주를 이용한 엠. 투베르콜로시스 폴리펩티드의 정제 및 특성 규명

본 발명의 엠. 투베르콜로시스 항원은 실질적으로 문헌 [Sanderson et al. J.Exp.Med., 1995, 182:1751-1757]에 기재된 바와 같이 엠. 투베르콜로시스 균주 H37Rv 및 Erdman의 cDNA 라이브러리를 발현 클로닝하여 분리하였고, 이것은 면역 반응성 T 세포주내에서 IFN- $\gamma$  및 PBMC 증식을 유도하는 것으로 확인되었다.

수상돌기 세포를 엠. 투베르콜로시스로 감염시켜 DC-4 및 DC-5라고 칭하는 2가지 CD4+ T 세포주를 만들었다. 구체적으로, 단일 공여체에서 얻은 유착성 PBMC로부터 수상돌기 세포를 얻은 뒤 투베르콜로시스로 감염시켰다. 동일 공여체 유래의 림프구를 한계 희석 조건하에 감염된 수상돌기 세포와 함께 배양하여 CD4+ T 세포주 DC-4 및 DC-5를 만들었다. 이 세포주는 Tb38-1이 아닌 엠. 투베르콜로시스 유래의 미정제 가용성 단백질과 반응하는 것으로 나타났다. 다시, 한계 희석 조건을 사용하여 DC-6이라 칭한 제3의 CD4+ T 세포주를 얻었으며, 이 세포주는 미정제 가용성 단백질 및 Tb38-1과 모두 반응하는 것으로 나타났다.

엠. 투베르콜로시스 균주 H37Rv 및 Erdman으로부터 계놈 DNA를 분리하고, 이것과 람다 ZAP 발현체(미국 캘리포니아주 리졸라에 소재하는 스트라타진 제품)를 사용하여 벡터 pBSK(-)로 발현 라이브러리를 작제하였다. 이 라이브러리를 이. 콜라이로 형질전환시키고, 유도된 이. 콜라이 배양물의 푸울을 수상돌기 세포와 항온배양한 뒤, 그 결과 얻어지는 항온배양된 수상돌기 세포가 CD4+ T 세포주 DC-6의 세포 증식과 IFN- $\gamma$  생성을 자극하는 능력을 하기 실시예 2에 기재된 바와 같이 관찰하였다. 양성의 푸울을 분획하여, 순수 엠. 투베르콜로시스 클론이 얻어질 때까지 재테스트하였다. 19개의 클론을 분리하였고, 이중 9개는 미국 특허 출원 08/533,634에 개시된 종래 동정된 엠. 투베르콜로시스 항원 TbH-9 및 Tb38-1을 포함하는 것으로 확인되었다. 나머지 10개의 클론(이하, Tb224, Tb636, Tb424, Tb436, Tb398, Tb508, Tb441, Tb475, Tb488 및 Tb465)의 cDNA 서열을 분석하여 각각 서열 1~10에 기재하였다. 이에 대응하는 Tb224 및 Tb636의 추론 아미노산 서열은 각각 서열 13과 14에 기재하였다. 이들 두 항원의 오픈 리딩 프레임은 TbH-9와 상동성이 약간 있는 것으로 확인되었다. 또한, Tb224와 Tb636은 중첩 클론인 것으로 확인되었다.

또, Tb424, Tb436, Tb398, Tb508, Tb441, Tb475, Tb488 및 Tb465는 각각 2개의 작은 오픈 리딩 프레임(ORF-1 및 ORF-2라 칭함) 또는 이것의 절두형을 포함하는 것으로 확인되었으며, 이와 함께 각 클론에 대해 ORF-1 및 ORF-2에 약간의 변형이 있는 것도 확인되었다. Tb424, Tb436, Tb398, Tb508, Tb441, Tb475, Tb488 및 Tb465의 ORF-1 및 ORF-2에 해당하는 추론 아미노산 서열은 서열 16과 17, 18과 19, 20과 21, 22와 23, 24와 25, 26과 27, 28과 29 및 30과 31에 각각 제시하였다. 또한, 클론 Tb424와 Tb436은 ORF-U라고 칭하는 분명한 제3의 오픈 리딩 프레임을 포함하는 것

으로 확인되었다. Tb424 및 Tb436의 ORF-U에 대한 추론 아미노산 서열은 각각 서열 32 및 33에 기재하였다. Tb424 및 Tb436은 중첩 클론 또는 유사 중복/트란스포즈 카피인 것으로 확인되었다. 이와 유사하게, Tb398, Tb508 및 Tb465도 Tb475 및 Tb488과 마찬가지로 중첩 클론 또는 유사 중복/트란스포즈 카피인 것으로 확인되었다.

이 서열들을 BLASTN 시스템을 사용하여 젠뱅크의 공지 서열과 비교하였다. 항원 Tb224 및 Tb431과 상동성이 있는 서열은 찾을 수 없었다. Tb636은 종래 엠. 투베르쿨로시스에서 동정된 코스미드와 100% 동일한 것으로 확인되었다. 이와 유사하게 Tb508, Tb488, Tb398, Tb424, Tb436, Tb441, Tb465 및 Tb475도 공지의 엠. 투베르쿨로시스 코스미드와 상동성이 있는 것으로 확인되었다. 또한, Tb488은 엠. 투베르쿨로시스 토포이소머라제 I과 100% 상동성이 있는 것으로 확인되었다.

오픈 리딩 프레임 ORF-1에 대한 17가지 중첩 웨პ티드(각각 1-1~1-17이라 칭함, 서열 34~50)와 오픈 리딩 프레임 ORF-2에 대한 30가지 중첩 웨პ티드(이하, 2-1~2-30이라 칭함, 서열 51~80)를 이하 실시에 3에 기재된 절차에 따라 합성하였다.

합성 웨პ티드, 재조합 ORF-1 및 ORF-2가 PPD 양성 공여체 유래의 PBMC에서 IFN- $\gamma$ 생성과 T 세포 증식을 유도하는 능력은 하기 실시에 2에 기재된 바와 같이 분석하였다. 도 1a와 1b 및 2a와 2b는 각각 D7 및 D160이라 칭한 2가지 공여체의 재조합 ORF-2 및 합성 웨პ티드 2-1~2-16에 의한 T 세포 증식 및 IFN- $\gamma$ 의 자극에 대하여 예시한 것이다. 재조합 ORF-2(이하 MTI라고 칭함)는 2가지 공여체 모두의 PBMC에서 T 세포 증식과 IFN- $\gamma$ 생성을 자극하였다. 각 합성 웨პ티드에 의해 관찰되는 PBMC 자극의 양은 각 공여체마다 달랐는데, 이것은 각 공여체가 ORF-2상의 여러 에피토프를 인식한다는 것을 시사하는 것이다. ORF-1, ORF-2 및 ORF-U가 암호하는 단백질은 각각 MTS, MTI 및 MSF라고 명명하였다.

MSF의 서열에 대한 18가지 중첩 웨პ티드(각각 MSF-1~MSF-18, 서열 번호 84~101)를 합성하고, 엠. 투베르쿨로시스 배양 여과물에 대하여 생성된 CD4+ T 세포주내 T 세포 증식 및 IFN- $\gamma$  생성을 자극하는 각각의 능력을 이하 기재된 바와 같이 조사하였다. MSF-12 및 MSF-13이라 칭한 웨პ티드(각각, 서열 95 및 96)가 가장 높은 반응도를 나타내는 것으로 확인되었다. 또한, Tb224의 오픈 리딩 프레임에 대한 2가지 중첩 웨პ티드(서열 81 및 82)를 합성하였으며, 이는 PPD 양성 공여체 유래의 PBMC에서 T 세포 증식과 IFN- $\gamma$ 생성을 유도하는 것으로 확인되었다.

전술한 방법에 따라 엠. 투베르쿨로시스에 감염된 수상돌기 세포에 대하여 다른 공여체 유래의 2가지 CD4+ T 세포주를 생성하였다. 이 세포주를 사용하여 전술한 엠. 투베르쿨로시스 cDNA 발현 라이브러리를 스크리닝한 결과, Tb867 및 Tb391이라 칭한 2개의 클론을 분리하였다. 서열 분석된 Tb867의 cDNA 서열은 종래 분리된 엠. 투베르쿨로시스 코스미드 SCY22G10과 동일한 것으로 확인되었고, 후보 반응성 오픈 리딩 프레임은 750 아미노산 엠. 투베르쿨로시스 단백질 키나제를 암호하였다. Tb391의 서열 분석된 cDNA 서열을 젠뱅크의 서열과 비교한 결과 공지 서열과 유의적인 상동성을 나타내지 않았다.

또 다른 연구로서, 실질적으로 전술한 바와 같이 엠. 투베르쿨로시스 배양 여과물에 대하여 CD4+ T 세포주를 생성하고, 이것을 사용하여 전술한 엠. 투베르쿨로시스 Erdman cDNA 발현 라이브러리를 스크리닝하였다. Tb431, Tb472, Tb470, Tb838 및 Tb962라 칭한 5개의 반응성 클론을 분리하였다. Tb431, Tb472, Tb470 및 Tb838의 서열 분석된 cDNA 서열은 각각 서열 11, 12, 104 및 105에 기재하였고, Tb962의 서열 분석된 cDNA 서열은 서열 106 및 107에 기재하였다. Tb431에 대응하는 추론 아미노산 서열은 서열 15에 제시하였다.

이어서, Tb472의 전길이 cDNA 서열(서열 108)을 분리하는 실험을 실시하였다. 중첩 웨პ티드를 합성하고, 이를 사용하여 반응성 오픈 리딩 프레임을 확인하였다. Tb472에 의해 암호된 단백질(MSL이라 칭함)의 추론 아미노산 서열은 서열 109에 기재하였다. Tb472 및 MSL의 서열을 전술한 바와 같이 젠뱅크의 서열과 비교한 결과 공지 서열과는 상동성을 나타내지 않는 것으로 확인되었다. MSL 서열에 대한 15가지 중첩 웨პ티드(이하, 각각 MSL-1~MSL-15라고 칭함; 서열 110~124)를 합성하고, 엠. 투베르쿨로시스 배양 여과물에 대하여 생성된 CD4+ T 세포주에서 T 세포 증식과 IFN- $\gamma$  생성을 자극하는 능력을 이하 기재되는 바와 같이 조사하였다. MSL-10(서열 119) 및 MSL-11(서열 120)이라 칭한 웨პ티드가 가장 높은 반응성을 나타내는 것으로 확인되었다.

서열 분석된 Tb838의 cDNA 서열을 젠뱅크 중의 서열과 비교한 결과 종래 분리된 엠. 투베르쿨로시스 코스미드 SCY07H7과 동일한 것으로 확인되었다. 클론 Tb962의 서열 분석된 cDNA 서열과 젠뱅크의 서열을 비교한 결과, 1개가 박토페리틴의 일부를 암호하는 것인 2개의 종래 동정된 엠. 투베르쿨로시스 코스미드와 약간의 상동성을 나타내는 것으로 확인되었다. 하지만, 재조합 박토페리틴은 Tb962를 분리하는 데 사용된 T 세포주와 반응성이 없는 것으로 확인되었다.

전술한 클론 Tb470을 사용하여 TbH9과 상동성을 나타내고 Mtb40이라 칭한 40 kDa 항원을 암호하는 것으로 밝혀진 전길이 오픈 리딩 프레임(서열 125)을 회수하였다. 서열 분석된 Mtb40의 아미노산 서열은 서열 126에 기재하였다. 이와 유사하게, 이어서 서열 83에 기재된 Tb431의 전길이 cDNA 서열을 분리하였으며, 이는 Mtb40을 암호하는 오픈 리딩 프레임을 포함하는 것으로 확인되었다. 또한, Tb470과 Tb431은 U-ORF 유사 항원을 암호하는 잠재적 오픈 리딩 프레임을 포함하는 것으로 확인되었다.

엠. 투베르콜로시스 배양 여과물에 대하여 생성된 복수의 CD4+ T 세포주로 엠. 투베르콜로시스 Erdman cDNA 발현 라이브러리를 스크리닝하여 Tb366, Tb433 및 Tb439라고 칭한 3개의 클론을 분리하였다. 서열 분석된 Tb366, Tb433 및 Tb439의 cDNA 서열을 각각 서열 127, 128 및 129에 기재하였다. 이 서열을 젠뱅크의 서열과 비교한 결과 Tb366과 유의적 상동성이 없는 것으로 나타났다. Tb433은 종래 동정된 엠. 투베르콜로시스 항원 MPT83과 약간의 상동성이 있는 것으로 확인되었다. Tb439는 종래 동정된 엠. 투베르콜로시스 코스미드 SCY02B10과 100% 동일성이 있는 것으로 확인되었다.

실질적으로 전술한 바와 같이 엠. 투베르콜로시스 PPD에 대하여 CD4+ T 세포주를 생성하고, 이것을 사용하여 상기 엠. 투베르콜로시스 Erdman cDNA 발현 라이브러리를 스크리닝하였다. 1개의 반응성 클론(Tb372라고 칭함)을 분리하고, 서열 분석된 cDNA 서열을 서열 번호 130 및 131에 기재하였다. 이 서열을 젠뱅크의 서열과 비교한 결과 유의적 상동성이 있는 서열은 발견되지 않았다.

이어서, 전술한 바와 같이 투베르콜로시스로 8일간 감염시킨 수상돌기 세포에 대하여 생성된 CD4+ T 세포주로 엠. 투베르콜로시스 cDNA 발현 라이브러리를 스크리닝하여 Tb390R5C6 및 Tb390R2C11로 칭한 2개의 클론을 분리하였다. 서열 분석된 Tb390R5C6의 cDNA 서열은 서열 132에 기재하고, Tb390R2C11의 cDNA 서열은 서열 133 및 134에 기재하였다. Tb390R5C6은 종래 동정된 엠. 투베르콜로시스 코스미드와 100% 동일성이 있는 것으로 확인되었다.

그 다음, 전술한 방법을 사용하여 제조된 엠. 투베르콜로시스 게놈 DNA 라이브러리를 다음과 같이 스크리닝하였다. 엠. 투베르콜로시스 Erdman 균주의 게놈 DNA를 평균 2 kb의 크기로 무작위 절단하고 클리나우 폴리머라제로 평활 말단화한 뒤, EcoRI 어댑터를 부가하였다. 이 삽입체를 스크린 파지 벡터(미국 위스콘신주 매디슨에 소재하는 노바젠 제품)에 결찰시키고 파지메이커 추출물(노바젠 제품)을 사용하여 시험관내에서 팩키징하였다. 이 파지 라이브러리(Erd λ스크린 라이브러리라 칭함)를 증폭시키고 그 일부를 자가서브클로닝 기작으로 이. 콜라이 균주 BM25.8(노바젠)을 사용하여 플라스미드 발현 라이브러리로 전환시켰다. pSCREEN 재조합체를 함유하는 BM25.8 배양물로부터 플라스미드 DNA를 정제하고 이것을 사용하여 발현 숙주 균주 BL21(DE3)pLysS의 컴파턴트 세포를 형질전환시켰다. 형질전환된 세포를 96웰 미량역가 평판에 각 웨л당 약 50개의 콜로니를 보유하도록 분액하였다. 96웰 플라스미드 라이브러리 포맷의 레플리카 평판을 재조합 단백질을 발현하도록 IPTG로 유도하였다. 유도 후, 평판을 원심분리하여 이. 콜라이를 침전시키고, 이것을 전술한 바와 같이 PPD 양성 공여체(공여체 160)로 만든 CD4+ T 세포주의 T 세포 발현 클로닝에 직접 사용하였다. 엠. 투베르콜로시스 T 세포 항원을 발현하는 이. 콜라이를 포함하는 푸울을 각각의 콜로니로 분리시킨 뒤, 유사 방식으로 재분석하여 양성군을 동정하였다.

Erdλ스크린 라이브러리의 하나의 96웰 평판으로 공여체 160 유래의 T 세포주를 스크리닝하여 총 9개의 양성군을 얻었다. 공여체 160 유래의 T 세포로 전술한 pBSK 라이브러리를 스크리닝하는 것에 관한 종래의 실험은 대부분 또는 모든 양성 클론이 TbH-9, Tb38-1 또는 MTI(미국 특허 출원 08/533,634호에 개시)이거나 이것의 변형체일 것이라는 것을 암시하였다. 하지만, 서던 분석 결과 단지 3개의 웰만이 TbH-9, Tb38-1 및 MTI의 혼합 프로브와 하이브리드하는 것으로 나타났다. 나머지 6개 양성 웰 중 2개는 동일한 것으로 확인되었다. 분리된 2개의 클론(Y1-26C1 및 Y1-86C11이라 칭함)에 대하여 서열 분석된 5' cDNA 서열을 각각 서열 번호 135 및 136에 제시하였다. hTcc#1이라 칭한 분리된 클론의 전길이 cDNA 서열은 서열 137에 제시하고, 대응하는 추론 아미노산 서열은 서열 138에 제시하였다. 전술한 바와 같이 hTcc#1의 서열과 젠뱅크의 서열을 비교한 결과, 종래 동정된 엠. 투베르콜로시스 MTCY07H7B.06과 약간의 상동성이 있는 것으로 확인되었다.

## 실시예 2

### 엠. 투베르콜로시스 항원에 의한 T 세포 증식과 인터페론-γ생성 유도

재조합 엠. 투베르콜로시스 항원이 T 세포 증식과 인터페론-γ생성을 유도하는 활성이 있는지 다음과 같이 측정하였다.

단백질은 문헌[Skeiky et al. J.Exp.Med., 1995, 181:1527-1537]에 기재된 바와 같이 IPTG로 유도하고 젤 용출로 정제하였다. 그 다음, 정제한 폴리펩티드를 PBMC 제조물에서 T 세포 증식을 유도할 수 있는지에 대하여 스크리닝하였다. PPD 피부 테스트에서 양성 반응을 보인다고 알려지고 그 T 세포가 PPD에 대해 반응하여 증식하는 것으로 알려진 PBMC를 10%의 수집된 인간 혈청과 젠타마이신 50 µg/ml이 보강된 RPMI 1640을 포함하는 배지에서 배양하였다. 정제된 폴리펩티드를 0.5~10 µg/ml의 농도로 첨가하는 것을 2중으로 실시하였다. 96웰 등근비탁 평판에서 200 µl의 부피로 배양한지 6일 후, 각 웰로부터 배지 50 µl를 취하여 하기하는 바와 같이 IFN-γ농도를 측정하였다. 그 다음, 이 평판을 삼중수소화된 티미딘 1 µCi/웰로 18 시간 동안 펄스시키고, 수거한 뒤, 삼중수소의 흡수율을 가스 신틸레이션 계수기로 측정하였다. 2중의 실시 모두에서의 증식이 배지 단독물에서 배양된 세포에서 관찰되는 증식의 3배 이상인 분획을 양성으로 간주하였다.

IFN-γ는 효소 결합 면역흡착 분석법(ELISA)을 이용하여 측정하였다. PBS 중에 함유된 인간 IFN-γ에 대해 지향성인 마우스 모노클로날 항체(미국 캘리포니아주 산디에고에 소재하는 PharMingen 제품)로 ELISA 평판을 실온에서 4 시간 동안 코팅시켰다. 그 다음, 웰을 5%(W/V) 탈지 분유 함유의 PBS로 실온에서 1 시간 동안 차단시켰다. 이 평판을 PBS/0.2% TWEEN-20으로 6회 세척하고, ELISA 평판 내에서 배양 배지로 1:2로 희석한 시료를 실온에서 하룻밤 동안 항온처리하였다. 이 평판을 다시 세척하고, PBS/10% 정상 염소 혈청으로 1:3000배 희석된 폴리클로날 토키 항인간 IFN-γ혈청을 각 웰에 첨가하였다. 그 다음 이 평판을 실온에서 2 시간 동안 항온처리하고 세척한 뒤, 양고추냉이 퍼옥시다제에 결합된 항토끼 IgG(미국 미주리주 세인트 루이스에 소재하는 시그마 케미컬 컴퍼니 제품)를 PBS/5% 탈지분유 중의 1:2000 희석물로 첨가하였다. 실온에서 2시간 동안 항온처리한 후, 평판을 세척하고 TMB 기질을 첨가하였다. 이 반응을 20분 후 1 N 황산을 첨가하여 중단시켰다. 기준 파장을 570 nm로 하고, 450 nm에서 흡광도를 측정하였다. 2중의 테스트 배양물에서 배양된 세포의 평균 OD가 배지 단독물에서만 배양된 세포의 평균 OD에 비하여 2배 이상이고 표준 편차가 3으로 나타난 분획은 양성인 것으로 간주하였다.

### 실시예 3

#### 마우스 엠. 투베르쿨로시스 모델로부터 생성된 CD4+ T 세포주를 이용한 엠. 투베르쿨로시스 폴리펩티드의 정제 및 특성 규명

C57BL/6 마우스를 엠. 투베르쿨로시스로 감염시키면 약 2~3주 동안 진행성 질병이 유발된다. 그 후, 강한 예방적 T 세포 매개의 면역 반응을 유도시키면 이 질병의 진행이 중단된다. 이 감염 모델을 이용하여 예방적 엠. 투베르쿨로시스 항원을 인식할 수 있는 T 세포주를 만들었다.

구체적으로, 28일 동안 엠. 투베르쿨로시스로 감염시킨 C57BL/6 마우스로부터 비장 세포를 분리하여, 이를 전술한 바와 같이 특이적인 항-엠. 투베르쿨로시스 T 세포주를 생성하는 데 이용하였다. 이렇게 하여 얻은 CD4+ T 세포주와, C57BL/6 마우스 유래의 정상 항원 제시(비장) 세포를 함께 사용하여 전술한 엠. 투베르쿨로시스 Erdλ스크린 라이브러리를 스크리닝하였다. T 세포에 대하여 자극 효과가 큰 것으로 확인된 반응성 라이브러리 푸울 중 1가지를 선택하고, 이에 대응하는 활성 클론(Y288C10이라 칭함)을 분리하였다.

클론 Y288C10의 서열을 분석한 결과, 2개의 잠재적 유전자를 나란히 포함하는 것으로 확인되었다. 서열 분석된 상기 2가지 유전자의 cDNA 서열(mTCC#1 및 mTCC#2라 칭함)을 각각 서열 139 및 140으로 제시하고, 이에 대응하는 추론 아미노산 서열을 각각 서열 141 및 142에 제시하였다. 이 서열을 젠뱅크의 서열과 비교한 결과 엠. 투베르쿨로시스 코스미드 MTY21C12에서 종래 밝혀진 미지 서열과 동일한 것으로 나타났다. mTCC#1 및 mTCC#2의 추론 아미노산 서열은 전술한 바와 같이 종래 동정된 TbH9 단백질군의 구성원과 약간의 상동성이 있는 것으로 확인되었다.

### 실시예 4

#### 합성 폴리펩티드의 합성

HPTU(O-벤조트리아졸-N,N,N',N'-테트라메틸우로늄 혼합스페이트) 활성화와 함께 Fmoc 화학을 사용하여 밀리포어 9050 웹티드 합성기로 폴리펩티드를 합성하였다. 이 웹티드의 아미노 말단에 Gly-Cys-Gly 서열을 부착하여 웹티드의 접합 또는 표지화 방법을 제공하였다. 고체 지지체로부터 웹티드의 절단은 다음과 같은 절단 혼합물, 트리플루오로아세트산:에탄디티올:티오아니솔:물:페놀 (40:1:2:2:3)을 사용하여 실시할 수 있다. 2 시간 동안 절단한 후, 웹티드를 저온의 메틸-t-부틸-에테르에서 침전시킬 수 있다. 그 다음, 웹티드 펠릿을 0.1% 트리플루오로아세트산(TFA)을 함유하는 물

에 용해하고, 동결건조시킨 뒤 C18 역상 HPLC로 정제하였다. 물(0.1% TFA 함유)에 0~60% 구배된 아세토니트릴(0.1% TFA 함유)을 사용하여 펩티드를 용출시켰다. 순수 분획을 동결건조한 후, 펩티드를 전기분무 질량 분광분석법 및 아미노산 분석법으로 특징을 규명하였다.

이상, 예시를 목적으로 본 발명의 구체적인 양태를 설명하였으나, 본 발명의 취지와 영역을 벗어나지 않는 한 다양한 변형이 이루어질 수 있음을 자명한 것이다.

삭제

## 도면의 간단한 설명

도 1a 및 도 1b는 재조합 ORF-2 및 ORF-2에 대한 합성 펩티드에 의한, 제1 PPD 양성 공여체(이하 D7로 칭함)로부터 유도된 T 세포에서의 증식 및 인터페론-γ 생성의 자극을 각각 보여준다.

도 2a 및 도 2b는 재조합 ORF-2 및 ORF-2에 대한 합성 펩티드에 의한, 제2 PPD 양성 공여체(이하 D160으로 칭함)로부터 유도된 T 세포에서의 증식 및 인터페론-γ 생성의 자극을 각각 보여준다.

### 발명의 상세한 설명

전술한 바와 같이, 본 발명은 일반적으로 결핵을 예방, 치료 및 진단하기 위한 조성물 및 방법에 관한 것이다. 본 발명의 조성물은 엠. 투베르쿨로시스 항원의 하나 이상의 면역원성 부분 또는 단지 보존적 치환 및/또는 변형에 있어서만 차이가 있는 상기 항원의 변형체를 포함하는 폴리펩티드를 함유한다. 본원에서 사용된 "폴리펩티드"란 용어는 아미노산 잔기가 펩티드 공유결합에 의해 결합되어 있는 전길이의 단백질을 비롯한 임의 길이의 아미노산쇄(즉, 항원)를 포함하는 개념이다. 따라서, 상기 항원중 하나의 면역원성 부분을 포함하는 폴리펩티드는 면역원성 부분으로만 구성되거나 또는 추가 서열을 포함할 수 있다. 추가 서열은 천연의 엠. 투베르쿨로시스 항원에서 유래한 것이거나 또는 이종성일 수 있고, 이러한 서열은 면역원성일 수 있다(그러나, 필수 요건은 아님).

본원에서 사용된 "면역원성"이란 환자, 예컨대 인간 및/또는 생물학적 시료에서 면역 반응(예컨대, 세포성 면역 반응)을 유도하는 능력을 말한다. 구체적으로, 면역원성인 항원(및 이러한 항원의 면역원성 부분 또는 기타 변형체)은 T 세포, NK 세포, B 세포 및 대식세포로 이루어진 군에서 선택되는 하나 이상의 세포(이들 세포는 엠. 투베르쿨로시스 면역 개체로부터 유래함)를 포함하는 생물학적 시료중에서 세포 증식, 인터루킨-12 생성 및/또는 인터페론-γ생성을 자극할 수 있다. 1종 이상의 엠. 투베르쿨로시스 항원의 하나 이상의 면역원성 부분을 포함하는 폴리펩티드는 일반적으로 환자에서 결핵을 검출하거나 또는 결핵에 대한 예방 면역을 유도하는 데 사용할 수 있다.

또한 본 발명의 조성물 및 방법은 상기 폴리펩티드의 변형체를 포함한다. 본원에서 사용된 폴리펩티드 "변형체"란 보존적 치환 및/또는 변형에 있어서만 언급되는 폴리펩티드와 상이한 것으로서, 그 폴리펩티드의 치료성, 항원성 및/또는 면역원성을 보유한 폴리펩티드이다. 폴리펩티드 변형체는 동정된 폴리펩티드와 약 70% 이상, 더욱 바람직하게는 약 90% 이상, 가장 바람직하게는 95% 이상의 동일성을 나타낸다. 면역 반응성을 가진 폴리펩티드의 경우, 상기 폴리펩티드중 하나의 아미노산 서열을 변형시키고, 변형된 폴리펩티드의 면역 반응성을 평가함으로써 변형체를 동정할 수도 있다. 진단용 결합제의 생성에 유용한 폴리펩티드의 경우에, 변형된 폴리펩티드가 결핵의 유무를 검출하는 항체를 생성시키는 능력을 평가함으로써 변형체를 동정할 수 있다. 한편, 본 발명의 진단 방법에 유용하게 이용할 수 있는 본원에서 청구하는 항원의 변형체는 변형된 폴리펩티드가 결핵에 감염된 환자의 혈청에 존재하는 항체를 검출하는 능력을 평가하여 동정할 수 있다. 이러한 변형 서열은, 예컨대, 본원에 기재하는 대표적인 절차를 사용하여 제조하고 테스트할 수 있다.

"보존적 치환"이란 펩티드 화학 분야의 당업자가 폴리펩티드의 2차 구조 및 수치료성이 실질적으로 변화되지 않는 것으로 예측할 수 있도록, 아미노산을 유사한 성질을 가진 또 다른 아미노산으로 치환하는 것을 말한다. 일반적으로, 다음 아미노산 군이 보존적 변화를 나타낸다: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; 및 (5) phe, tyr, trp, his.

변형체는 또한 폴리펩티드의 면역원성, 2차 구조 및 수치료성(hydropathic)에 미치는 영향이 최소인 아미노산의 결실 또는 부가와 같은 방법으로 변형시킬 수 있다. 예컨대, 폴리펩티드는 해독과 동시에 또는 해독 후에 단백질의 전달을 지령하는 단백질의 N-말단의 시그널(또는 리더) 서열에 접합시킬 수 있다. 폴리펩티드는 또한 폴리펩티드의 합성, 정제 또는 동정을 용이하게 하기 위한 링커 또는 기타 서열(예컨대, 폴리-His)에, 또는 폴리펩티드의 고체 지지체에의 결합을 강화시키기 위한 링커 또는 기타 서열에 접합시킬 수 있다. 예컨대, 폴리펩티드는 면역글로불린 Fc 영역에 접합시킬 수 있다.

일반적으로, 엠. 투베르쿨로시스 항원 및 이 항원을 코딩하는 DNA 서열은 다양한 방법 중 임의의 것을 이용하여 제조할 수 있다. 예컨대, 엠. 투베르쿨로시스에서 유래한 게놈 또는 cDNA 라이브러리를 하나 이상의 엠. 투베르쿨로시스 면역 개체로부터 얻은 말초혈 단핵 세포(PBMC)나 T 세포주 또는 T 세포 클론을 사용하여 직접 스크리닝할 수 있다. 직접 라이브러리 스크리닝은 일반적으로 엠. 투베르쿨로시스 면역 개체에서 유래한 T 세포에서 증식 및/또는 인터페론-γ생성을 유도하는 능력에 대해, 발현된 재조합 단백질의 풀풀(pool)을 분석하여 수행할 수 있다. 상기한 바와 같이 항체 반응성에 기초하여 잠재 T 세포 항원을 먼저 스크리닝할 수 있다.

한편, 항원들을 코딩하는 DNA 서열을 엠. 투베르콜로시스에 감염된 환자로부터 얻은 혈청을 사용하여 적절한 엠. 투베르콜로시스 계놈 또는 cDNA 발현 라이브러리를 스크리닝하여 동정할 수 있다. 그러한 스크리닝은 일반적으로 당업자에게 널리 알려진 기법, 예컨대, 문헌[Sambrook 등, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, 뉴욕, 1989]에 기재된 기법을 이용하여 수행할 수 있다.

그 후 정제된 항원을, 예컨대 본원에 개시된 대표적인 방법을 이용하여 적절한 면역 반응(예컨대, 세포성 면역 반응)을 유도하는 능력에 대해 평가한다. 그 다음, 면역원성 항원을 전통적인 에드먼(Edman) 화학과 같은 기법을 사용하여 부분적으로 서열 분석할 수 있다. Edman 및 Berg의 문헌[Eur. J. Biochem. 80:116-132, 1967] 참조. 또한, 면역원성 항원은 이 항원을 코딩하는 DNA 서열을 발현 벡터내로 삽입하고 이를 적절한 숙주에서 발현시킴으로써 재조합에 의해 생산할 수 있다.

본 발명의 항원을 코딩하는 DNA 서열은 또한 분리된 항원의 부분 아미노산 서열에서 유래한 축퇴성 올리고뉴클레오티드에 하이브리드화하는 DNA 서열에 대해 적절한 엠. 투베르콜로시스 cDNA 또는 계놈 DNA 라이브러리를 스크리닝하여 얻을 수도 있다. 그러한 스크리닝에 사용하는 축퇴성 올리고뉴클레오티드 서열은 문헌[Sambrook 등, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, 뉴욕, 1989] 및 여기에 인용된 참조 문헌에 기재된 바와 같이 디자인하고, 합성하여, 스크리닝을 수행할 수 있다. 당해 분야에 널리 알려진 방법으로 상기 올리고뉴클레오티드를 사용하는 폴리머라제 연쇄 반응(PCR)을 이용하여 cDNA 또는 계놈 라이브러리로부터 핵산 프로브를 분리할 수도 있다. 그 후, 분리된 프로브를 사용하여 라이브러리 스크리닝을 수행할 수 있다.

제조 방법과는 무관하게, 본원에 개시된 항원(및 그 면역원성 부분)은 면역원 반응을 유도하는 능력을 가진다. 보다 구체적으로, 항원은 엠. 투베르콜로시스 면역 개체에서 유래한 T 세포, NK 세포, B 세포 및/또는 대식세포에서 증식 및/또는 시토킨 생성(즉, 인터페론- $\gamma$  및/또는 인터루킨-12 생성)을 유도하는 능력을 갖는다. 항원에 대한 면역원성 반응을 평가하는데 사용하기 위한 세포 유형의 선택은 목적하는 반응에 따라 좌우된다. 예컨대, 인터루킨-12 생성은 B 세포 및/또는 대식세포를 함유하는 제제를 사용하여 평가하는 것이 가장 쉽다. 엠. 투베르콜로시스 면역 개체는 엠. 투베르콜로시스에 대한 효과적인 T 세포 반응을 일으킴으로써 결핵의 발병에 저항하는 것으로 간주되는 개체(실질적으로 질병의 증상이 없는 개체)를 말한다. 이러한 개체는 결핵 단백질(PPD)에 대한 피내 피부 테스트 반응이 강한 양성을 보이고(즉, 직경이 약 10 mm 보다 큰 경계를 나타냄), 결핵의 어떠한 정후나 증상도 없음에 근거하여 식별할 수 있다. 엠. 투베르콜로시스-면역 개체에서 유래한 T 세포, NK 세포, B 세포 및 대식세포는 당업자에게 알려진 방법을 사용하여 준비할 수 있다. 예컨대, PBMC(즉, 말초혈 단핵 세포)의 제제는 성분 세포의 추가 분리 없이 이용될 수 있다. 일반적으로 PBMC는, 예컨대 피콜(Ficoll<sup>TM</sup>)(뉴욕주 윈스롭 래보리토리즈)을 통한 밀도 원심분리를 이용하여 준비할 수 있다.

또한, 본원에 기재된 분석에 사용하기 위한 T 세포는 PBMC로부터 직접 정제할 수도 있다. 한편, 마이코박테리아 단백질에 대해 반응성이 있는 농축된 T 세포주 또는 개체의 마이코박테리아 단백질에 대해 반응성이 있는 T 세포 클론을 이용할 수도 있다. 이러한 T 세포 클론은 예컨대, 엠. 투베르콜로시스 면역 개체에서 유래한 PBMC를 마이코박테리아 단백질과 함께 2~4주 동안 배양하여 생성시킬 수 있다. 이 과정은 마이코박테리아 단백질에 특이적인 T 세포만을 증산하여 그러한 세포만으로 구성된 세포주를 생성시킨다. 당업자에게 알려진 방법에 의해 이들 세포를 클로닝하고 개개의 단백질을 사용하여 테스트하여, 개체의 T 세포 특이성을 보다 정확히 확인할 수 있다. 일반적으로, 엠. 투베르콜로시스 면역 개체에서 유래한 T 세포, NK 세포, B 세포 및/또는 대식세포를 사용하여 수행한 증식 및/또는 시토킨 생성(즉, 인터페론- $\gamma$  및/또는 인터루킨-12 생성)에 대한 분석에서 테스트에 양성 반응을 보이는 항원은 면역원성이 있는 것으로 간주된다. 이 분석은, 예컨대 후술하는 대표적인 절차를 사용하여 수행할 수 있다. 이러한 항원의 면역원성 부분은 유사한 분석을 이용하여 확인할 수 있고, 본원에 개시된 폴리펩티드내에 존재할 수 있다.

폴리펩티드(예, 면역원성 항원 또는 그 부분 또는 기타 변형체)가 세포 증식을 유도하는 능력은 세포들(예컨대, T 세포 및/또는 NK 세포)을 폴리펩티드와 접촉시키고, 이들 세포의 증식을 측정함으로써 평가한다. 일반적으로, 약 10<sup>5</sup>개 세포를 평가하기에 충분한 폴리펩티드의 양은 약 10 ng/ml ~ 약 100  $\mu$ g/ml, 바람직하게는 약 10  $\mu$ g/ml이다. 세포와 폴리펩티드의 항온처리는 통상적으로 37°C에서 약 6일 동안수행한다. 폴리펩티드와의 항온처리 후에, 증식 반응에 대해 세포를 분석하는데, 증식 반응은 당업계에 공지된 방법, 예컨대 방사능 표지된 티미딘 필스에 세포를 노출시키는 단계 및 세포내 DNA로 병입된 표지를 측정하는 단계를 포함하는 방법으로 평가할 수 있다. 일반적으로, 증식을 백그라운드(즉, 폴리펩티드 없이 배양한 세포에 대해 관찰된 증식)보다 3배 이상 증가시킨 폴리펩티드는 증식을 유도할 수 있는 것으로 본다.

세포내에서 인터페론- $\gamma$  및/또는 인터루킨-12의 생성을 자극하는 폴리펩티드의 능력은, 폴리펩티드와 세포를 접촉시키는 단계, 및 세포가 생성하는 인터페론- $\gamma$  또는 인터루킨-12의 양을 측정하는 단계를 수행하여 평가할 수 있다. 일반적으로,

약  $10^5$ 개 세포를 평가하기에 충분한 폴리펩티드의 양은 약 10 ng/ml ~ 약 100 µg/ml, 바람직하게는 약 10 µg/ml이다. 폴리펩티드는, 고체 지지체, 예컨대 비아이드 또는 미국 특허 제4,897,268호 및 제5,075,109호에 개시된 것과 같은 생분해성 미소구상에 고정시킬 수 있으나, 반드시 그럴 필요는 없다. 세포와 폴리펩티드의 항온 처리는 통상 37°C에서 약 6일 동안 수행한다. 폴리펩티드와의 항온처리 후에, 인터페론-γ 및/또는 인터루킨-12(또는 하나 이상의 이의 서브유닛)에 대해 세포를 분석하는데, 이때 당업계에 공지된 방법, 예컨대 효소 결합 면역흡착법(ELISA) 또는 IL-12 P70 헤테로이량체인 경우 T 세포의 증식을 측정하는 방법과 같은 생물학적 분석 방법으로 평가할 수 있다. 일반적으로, 배양된 상청액(1 ml당  $10^4$  ~  $10^5$ 개 T 세포 함유) 1 ml당 인터페론-γ 50 pg 이상을 생성하는 폴리펩티드는 인터페론-γ의 생성을 자극하는 것으로 볼 수 있다.  $10^5$ 개 대식세포 또는 B 세포당(또는  $3 \times 10^5$ 개 PBMC당) IL-12 P70 서브유닛 10 pg/ml 이상, 및/또는 IL-12 P40 서브유닛 100 pg/ml 이상의 생성을 자극하는 폴리펩티드는 IL-12의 생성을 자극하는 것으로 볼 수 있다.

일반적으로, 면역원성 항원이란 엠. 투베르쿨로시스 면역 개체 약 25% 이상으로부터 유래된 T 세포, NK 세포, B 세포 및/또는 대식 세포에서 증식 및/또는 시토킨 생성(즉, 인터페론-γ 및/또는 인터루킨-12 생성)을 자극하는 항원이다. 이들 면역원성 항원 중에서, 치료 특성이 우수한 폴리펩티드는 상기 분석에서 반응 정도와 반응이 관찰된 개체의 비율을 기초로 구별해 낼 수 있다. 또한, 치료 특성이 우수한 항원은 엠. 투베르쿨로시스 면역이 없는 개체의 약 25% 이상에서 유래된 세포에서 시험관내 증식 및/또는 시토킨 생성을 자극하지 않으며, 따라서 특별히 엠. 투베르쿨로시스 반응성 세포에 기인하지 않는 반응을 제거한다. 엠. 투베르쿨로시스 면역 개체에서 유래한 T 세포, NK 세포, B 세포 및/또는 대식 세포 제제의 고비율에서 반응을 유도하는 항원(다른 개체로부터의 세포 제조물 중에서의 반응 빈도는 낮음)은 치료 특성이 우수하다.

치료 특성이 우수한 항원은 백신으로 투여할 때 실험 동물에서 엠. 투베르쿨로시스 감염의 심각도를 감소시키는 능력을 기초로 하여 동정할 수 있다. 실험 동물에 사용하기에 적합한 백신 제제는 하기에 상세히 설명되어 있다. 효능은 실험 감염 후에 박테리아 수를 약 50% 이상 감소시키고/시키거나, 사망율을 약 40% 이상 감소시키는 항원의 능력을 근거로 하여 결정할 수 있다. 적절한 실험 동물로는 마우스, 기니아 피그 및 영장류가 있다.

일반적으로 진단 특성이 우수한 항원은, 활성 투베르쿨로시스가 있는 개체에서 수행되는 피내 피부 테스트에서 반응을 유발할 수 있지만, 엠. 투베르쿨로시스에 감염되지 않은 개체에서 수행되는 테스트에서는 반응을 유발할 수 없는 능력을 기초로 하여 동정할 수 있다. 피부 테스트는 일반적으로 후술하는 바와 같이 수행할 수 있으며, 5 mm 이상의 경계 반응을 나타내면 양성인 것으로 간주한다.

본원에 개시된 항원의 면역원성 부분은 공지된 기법, 예컨대 Paul의 문헌 [Fundamental Immunology, 3d ed., Raven Press, 1993, p243-247] 및 본원의 참고 인용문에 요약된 바와 같은 기법을 사용하여 제조 및 동정할 수 있다. 이러한 기법은 면역원성이 있는 천연 항원의 폴리펩티드 부분을 스크리닝하는 단계를 포함한다. 본원에 개시된 대표적인 증식 및 시토킨 생성 분석법은 대체로 상기 스크리닝에 사용될 수 있다. 폴리펩티드의 면역원성 부분은, 대표적인 분석법내에서 전길이 항원이 생성하는 것과 거의 유사한 면역 반응(예, 증식, 인터페론-γ 생성 및/또는 인터루킨-12 생성)을 유발하는 부분이다. 즉, 항원의 면역원성 부분은 본원에 개시된 모델 증식 분석법에서 전길이 항원이 유도하는 증식의 약 20% 이상, 바람직하게는 약 100%를 유도할 수 있다. 또한 면역원성 부분은 본원에 개시된 모델 분석에서 전길이 항원이 유도하는 인터페론-γ 및/또는 인터루킨-12의 약 20% 이상, 바람직하게는 약 100%의 생성을 자극할 수 있다.

엠. 투베르쿨로시스 항원의 부분 및 기타 변형체는 합성 또는 재조합 수단으로 제조할 수 있다. 약 100개 아미노산 미만, 일반적으로 약 50개 아미노산 미만의 합성 폴리펩티드는 당업계에 공지된 기법을 이용하여 제조할 수 있다. 예를 들어, 이러한 폴리펩티드는 상업적으로 이용되는 고상 기법, 예컨대 Merrifield 고상 합성법을 이용하여 합성할 수 있으며, 이 방법에서는 아미노산이 성장하는 아미노산 쇄에 순차적으로 부가된다. Merrifield의 문헌 [J.Am.Chem.Soc. 85:2149-2146, 1963] 참조. 폴리펩티드의 자동 합성을 위한 장치는 미국 캘리포니아주 포스터 시티의 퍼킨 엘머/아플라이드 바이오시스템즈 디비전과 같은 공급업체로부터 시판되며, 제조업자의 지시에 따라 작동시킬 수 있다. 천연 항원의 변형체는, 표준 돌연변이 유발 기법, 예컨대 올리고뉴클레오티드에서 유도된 부위 특이적 돌연변이 유발법을 이용하여 제조할 수 있다. 표준 기법을 사용하여 DNA 서열의 일부를 제거하여 절두된 폴리펩티드를 제조할 수도 있다.

천연 항원의 부분 및/또는 변형체를 포함하는 재조합 폴리펩티드는 당업계에 공지된 각종 기법을 이용하여 폴리펩티드를 코딩하는 DNA 서열로부터 쉽게 제조할 수 있다. 예를 들어, 재조합 단백질을 배양 배지내로 분비하는 적절한 숙주/벡터계에서 얻은 상청액을 먼저 시판용 필터를 사용하여 농축시킬 수 있다. 농축 후에, 농축물을 친화도 매트릭스 또는 이온 교환 수지와 같은 적절한 정제 매트릭스에 적용할 수 있다. 마지막으로, 하나 이상의 역상 HPLC 단계를 사용하여 재조합 단백질을 추가로 정제할 수 있다.

당업계에 공지된 임의의 각종 발현 벡터를 사용하여 본 발명의 재조합 폴리펩티드를 발현시킬 수 있다. 재조합 폴리펩티드를 코딩하는 DNA 문자를 포함하는 발현 벡터로 형질전환 또는 형질감염시킨 임의의 적절한 숙주 세포에서 발현시킬 수 있다. 적절한 숙주 세포로는 원핵세포, 효모 및 고등 진핵세포가 있다. 사용된 숙주 세포는 이. 콜라이, 효모 또는 포유류 세포, 예컨대 COS 또는 CHO가 바람직하다. 이러한 방식으로 발현된 DNA 서열은 자연 발생적인 항원, 자연 발생적인 항원의 부분 또는 이의 변형체를 코딩할 수 있다.

일반적으로, 본원에 개시된 폴리펩티드는 제조 방법과 관계없이 실질적으로 순수한 형태로 제조된다. 폴리펩티드는 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 가장 바람직하게는 약 99% 이상의 순도를 갖는다. 특정 바람직한 구체예에서, 하기에 상세히 설명된 바와 같이, 본원에 개시된 하나 이상의 방법에 사용하기 위한 약학 조성물 또는 백신내로 실질적으로 순수한 폴리펩티드를 병입시킨다.

한 구체예에서, 본 발명은 (a) 서열 번호 1-12, 83, 102-108, 125, 127-137, 139 및 140의 DNA 서열; (b) 이 DNA 서열의 상보 서열, 또는 (c) (a) 또는 (b)의 서열에 실질적으로 상동성인 DNA 서열에 의해 코딩되는 1종 이상의 아미노산 서열을 포함하는 엠. 투베르콜로시스 항원(또는 이러한 항원의 변형체)의 면역원성 부분을 하나 이상 포함하는 폴리펩티드를 개시하고 있다. 관련 구체예에서, 본 발명은 서열 번호 16-33, 109, 126, 138, 141, 142에 제공된 서열 및 이의 변형체로 구성된 군에서 선택되는 아미노산 서열을 가진 엠. 투베르콜로시스 항원의 면역원성 부분을 하나 이상 포함하는 폴리펩티드를 제공한다.

본원에서 제공된 엠. 투베르콜로시스 항원은 본원에 구체적으로 기술된 1종 이상의 DNA 서열에 실질적으로 상동성인 DNA 서열에 의해 코딩되는 변형체를 포함한다. 본원에서 사용한 "실질적인 상동성"은 보통의 엔중 조건하에 하이브리드화 할 수 있는 DNA 서열을 의미한다. 적절한 보통의 엔중 조건은  $5 \times SSC$ , 0.5% SDS, 1.0 mM EDTA(pH 8.0) 용액에서의 예비 세척; 50~65°C,  $5 \times SSC$ 에서 밤새, 또는 중간 상동성의 경우 45°C, 0.5  $\times SSC$ 에서 하이브리드화; 및 0.1% SDS를 함유하는  $2 \times SSC$ , 0.5  $\times SSC$  및 0.2  $\times SSC$ 를 각각 사용한 65°C에서의 20분간의 2회 세척을 포함한다. 이러한 하이브리드화 DNA 서열은 본 발명의 범위에 속하며, 암호 축퇴성으로 인해 하이브리드화 DNA 서열에 의해 코딩되는 면역원성 폴리펩티드를 코딩하는 뉴클레오티드 역시 본 발명에 속한다.

관련 양태에서, 본 발명은 본 발명의 제1 폴리펩티드 및 제2 폴리펩티드를 포함하는 융합 단백질, 또는 대안적으로 본 발명의 폴리펩티드 및 공지된 엠. 투베르콜로시스 항원, 예컨대 Andersen 및 Hansen 등의 문헌 [*Infect. Immun.* 57:2481-2488, 1989 (젠햄크 수탁 번호 M30046)]에 개시된 38 kD 항원, 또는 엠. 보비스(수탁 번호 U34848) 및 엠. 투베르콜로시스(Sorensen 등, *Infec. Immun.* 63:1710-1717, 1995)에서 이전에 동정된 ESAT-6을 포함하는 융합 단백질을 제공한다. 이러한 융합 단백질의 변형체도 제공된다. 본 발명의 융합 단백질은 제1 폴리펩티드 및 제2 폴리펩티드 사이에 링커 웨프티드를 포함할 수 있다.

본 발명의 융합 단백질을 코딩하는 DNA 서열은 공지된 재조합 DNA 기법을 사용하여 작제하여 제1 폴리펩티드 및 제2 폴리펩티드를 코딩하는 별개의 DNA 서열을 적절한 발현 벡터내로 회합시킨다. 제1 폴리펩티드를 코딩하는 DNA 서열의 3' 말단을, 웨프티드 링커를 사용하거나, 또는 사용하지 않고, 제2 폴리펩티드를 코딩하는 DNA 서열의 5' 말단에 결찰시켜, 서열의 리딩 프레임이 2개의 DNA 서열의 mRNA가 제1 폴리펩티드 및 제2 폴리펩티드 둘다의 생물학적 활성을 보유하는 단일 융합 단백질로 해독되도록 하는 상태가 되게 한다.

웨프티드 링커 서열을 사용하여 각 폴리펩티드를 2차 구조 및 3차 구조로 풀딩시키기에 충분한 거리로 제1 폴리펩티드 및 제2 폴리펩티드를 분리시킬 수 있다. 이러한 웨프티드 링커 서열은 당업계에 공지된 표준 기법을 사용하여 융합 단백질내로 병입시킨다. 적절한 웨프티드 링커 서열은 다음과 같은 인자를 기초로 선택할 수 있다. (1) 가요성이 있는 연장된 구조를 형성할 수 있는 능력; (2) 제1 폴리펩티드 및 제2 폴리펩티드상의 기능성 애피토프와 상호작용할 수 있는 2차 구조를 형성할 수 있는 능력의 부재; (3) 폴리펩티드 기능성 애피토프와 반응할 수 있는 소수성 잔기 또는 하전된 잔기의 결핍. 바람직한 웨프티드 링커 서열의 예로는 Gly, Asn 및 Ser 잔기가 있다. 기타 중성에 가까운 천연 아미노산, 예컨대 Thr 및 Ala 역시 링커 서열에 사용할 수 있다. 링커로서 유용하게 사용될 수 있는 아미노산 서열은 Maratea 등의 문헌 [*Gene* 40:39-46, 1985; Murphy 등, *Proc.Natl.Acad.Sci. USA* 83:8258-8262, 1986], 미국 특허 제4,935,233호 및 미국 특허 제4,751,180호에 개시된 아미노산 서열을 포함한다. 링커 서열은 길이가 1개~약 50개인 아미노산일 수 있다. 제1 폴리펩티드 및 제2 폴리펩티드가 기능성 도메인을 분리하고 입체 방해를 방지하는 데 사용될 수 있는 비필수 N 말단 아미노산 영역을 보유하는 경우, 웨프티드 서열은 필요하지 않다.

결찰된 DNA 서열은 적절한 전사 또는 해독 조절 성분에 작동가능하게 결합된다. DNA 발현을 담당하는 조절 성분은 제1 폴리펩티드를 코딩하는 DNA 서열의 5'에만 위치한다. 유사하게, 종지 코돈은 제2 폴리펩티드를 코딩하는 DNA 서열의 3'에만 존재한다.

또 다른 양태에서, 본 발명은 1종 이상의 상기 폴리펩티드 또는 융합 단백질(또는 이러한 폴리펩티드를 코딩하는 DNA 분자)을 사용하여 환자에서 결핵에 대한 예방 면역성을 유도하는 방법을 제공한다. 본원에서 "환자"는 임의의 온혈 동물, 바람직하게는 인간을 의미한다. 환자는 질병에 걸린 사람일 수도 있고, 또는 검출가능한 질병 및/또는 감염이 없는 사람일 수도 있다. 즉, 예방 면역성은 결핵을 예방하거나 치료하기 위해 유도할 수 있다.

이러한 양태에서, 폴리펩티드, 융합 단백질 또는 DNA 분자는 일반적으로 약학 조성물 및/또는 백신내에 존재한다. 약학 조성물은 1종 이상의 폴리펩티드를 포함할 수 있으며, 각 폴리펩티드는 1종 이상의 상기 서열(또는 이의 변형체)과 생리적 허용 담체를 포함할 수 있다. 백신은 1종 이상의 상기 폴리펩티드 및 비특이적 면역 반응 인핸서, 예컨대 면역 증강제 (adjuvant) 또는 리포좀(그 내부에 폴리펩티드가 병입되어 있음)을 포함할 수 있다. 이러한 약학 조성물 및 백신은 기타 엠. 투베르콜로시스 항원을, 조합 폴리펩티드내로 병입시키거나 또는 별도의 폴리펩티드내에 존재하도록 하여 포함할 수도 있다.

대안적으로, 백신은 폴리펩티드가 동일계에서 생성될 수 있도록 전술한 1종 이상의 폴리펩티드를 코딩하는 DNA를 포함할 수 있다. 이러한 백신에서, DNA는 핵산 발현계, 박테리아 발현계 및 바이러스 발현계를 비롯하여 당업계에 공지된 임의의 각종 전달계내에 존재할 수 있다. 적절한 핵산 발현계는 환자내에서의 발현을 위해 필수적인 DNA 서열(예, 적절한 프로모터 및 종결 시그널)을 포함한다. 박테리아 전달계는 세포 표면상에 폴리펩티드의 면역원성 부분을 발현하는 박테리아(예, 바실러스-칼메트-구에린)을 투여하는 것을 포함한다. 바람직한 구체예에서, DNA는 바이러스 발현계(예, 백시니아 또는 기타 폭스 바이러스, 레트로바이러스 또는 아데노바이러스)를 사용하여 도입할 수 있으며, 바이러스 발현계는 비병원성(불완전), 복제 캠피던트 바이러스의 사용을 포함할 수 있다. 이러한 발현계로 DNA를 병입시키기 위한 기술은 당업계에 공지되어 있다. DNA는, 예컨대 Ulmer 등의 문헌[Science 259:1745-1749, 1993]에 개시되어 있고, Cohen의 문헌[Science 259:1691-1692, 1993]에 검토된 바와 같이 "나출(裸出; naked)상태"일 수 있다. 나출 상태의 DNA 흡수율은 세포 내로 효과적으로 이송되는 생분해성 비이드상에 DNA를 코팅하여 증가시킬 수 있다.

관련 양태에서, 전술한 DNA 백신은 본 발명의 폴리펩티드 또는 공지된 엠. 투베르콜로시스 항원, 예컨대 전술한 38 kD 항원과 동시에 또는 순차적으로 투여할 수 있다. 예를 들어, 전술한 바와 같이 "나출 상태로" 또는 전달계를 이용하여 본 발명의 폴리펩티드를 코딩하는 DNA를 투여한 다음 항원을 투여하여 백신의 예방 면역 효과를 증강시킬 수 있다.

투여 경로 및 빈도뿐 아니라 용량을 개체별로 다양하게 하여, 현재 BCG를 사용한 면역화에 이용되고 있는 것과 동일하게 사용할 수 있다. 일반적으로 약학 조성물 및 백신은 주사(예, 피내, 근육내, 정맥내 또는 피하), 비강내(예, 흡출) 또는 경구로 투여할 수 있다. 1~36 주간 동안 1~3회 용량을 투여할 수 있다. 3~4 개월의 간격으로 3회 용량을 투여하고, 그 후 추가 백신 접종을 주기적으로 한다. 환자 개개인마다 다른 프로토콜이 적절할 수 있다. 적절한 용량은 전술한 바와 같이 투여했을 때 1~2년 동안 면역화된 환자내에서 엠. 투베르콜로시스 감염으로부터 환자를 보호하기에 충분한 면역 반응을 유발시킬 수 있는 폴리펩티드 또는 DNA의 양이다. 일반적으로, 1회 용량내에 존재하는 (또는 1회 용량내의 DNA에 의해 숙주내에서 생성되는) 폴리펩티드의 양은 숙주 1 kg당 약 1 pg~약 100 mg, 통상적으로 약 10 pg~약 1 mg, 바람직하게는 약 100 pg~약 1  $\mu$ g일 수 있다. 적절한 용량 크기는 환자의 체격에 따라 다양할 수 있지만, 통상 약 0.1 ml~약 5 ml이다.

당업자에게 공지된 임의의 적절한 담체가 본 발명의 약학 조성물에 사용될 수 있지만, 담체의 유형은 투여 방식에 따라 달라진다. 피하 주사와 같은 비경구 투여의 경우, 담체는 물, 염수, 알코올, 지질, 왁스 또는 완충액을 포함하는 것이 바람직하다. 경구 투여의 경우, 임의의 전술한 담체 또는 고형 담체, 예컨대 만니톨, 락토스, 전분, 스테아르산마그네슘, 나트륨 사카린, 활석, 셀룰로스, 글루코스, 슈크로스 및 탄산마그네슘을 사용할 수 있다. 생분해성 미소구(예, 폴리락틱 갈락티드) 역시 본 발명의 약학 조성물에 대한 담체로서 사용할 수 있다. 적절한 생분해성 미소구는, 미국 특허 제4,897,268호 및 제5,075,109호에 개시되어 있다.

임의의 각종 면역 증강제를 본 발명의 백신에 사용하여 면역 반응을 비특이적으로 증강시킬 수 있다. 대부분의 면역 증강제는 급속한 이화작용으로부터 항원을 보호하도록 고안된 물질, 예컨대 수산화나트륨 또는 광유, 및 면역 반응의 비특이적 자극제, 예컨대 지질 A, 보르타델라 퍼투시스(*Bordetella pertussis*) 또는 마이코박테리움 투베르콜로시스(*Mycobacterium tuberculosis*)를 포함한다. 적절한 면역 증강제는, 예컨대 프로인트 불완전 보조제 및 프로인트 완전 보조제(디프코 래보러티리즈)와 머크 보조제 65(미국 뉴저지주 라웨이 머크 앤드 캄파니 인코포레이티드) 같이 시판되고 있다. 기타 적절한 면역 증강제로는 명반, 생분해성 미소구, 모노포스포릴 지질 A 및 퀄(quill) A가 있다.

또 다른 양태에서, 본 발명은 피부 테스트를 이용하여 결핵을 진단하는 데 전술한 1종 이상의 폴리펩티드를 사용하는 방법을 제공한다. 본원에서 사용한 "피부 테스트"는, 전술한 1종 이상의 폴리펩티드의 피내 주사후에 지연형 과민(DTH) 반응(예, 팽윤, 적화(赤化), 피부염)을 측정하여 환자에서 직접 수행되는 임의의 분석법이다. 이러한 주사는 환자의 피부 세포와 폴리펩티드(들)를 접촉시키기에 충분한 임의의 장치, 예컨대 투베르쿨린 주사기 또는 1 ml 주사기를 사용하여 수행할 수 있다. 반응은 바람직하게는 주사 후 48 시간 이상, 더욱 바람직하게는 주사 후 48~72 시간 후에 측정한다.

DTH 반응은 세포 매개 면역 반응으로서, 이 반응은 미리 테스트 항원(즉, 사용된 폴리펩티드의 면역원성 부분 또는 이의 변형체)에 노출된 환자에게서 보다 크게 나타난다. 반응은 룰러(ruler)를 사용하여 시각적으로 측정할 수 있다. 일반적으로 직경이 약 0.5 cm 이상, 바람직하게는 약 1.0 cm 이상인 반응은 결핵 감염임을 나타내는 양성 반응이며, 이 경우 활성 질병으로 나타날 수도 나타나지 않을 수도 있다.

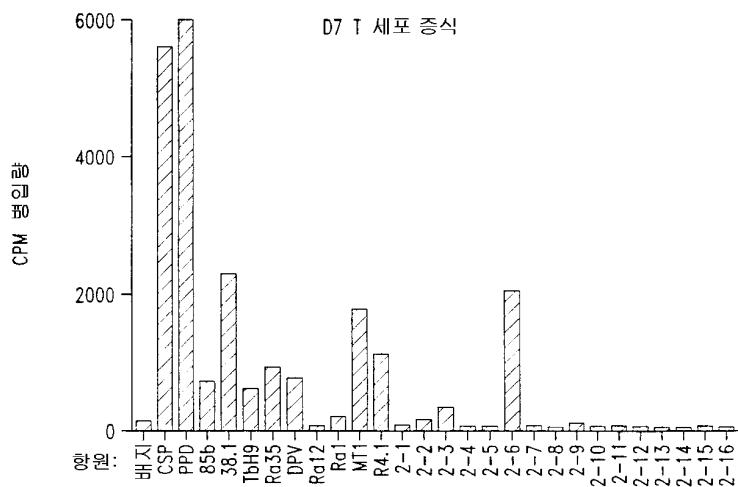
본 발명의 폴리펩티드는 전술한 바와 같은 폴리펩티드 및 생리적 허용 담체를 포함하는 약학 조성물로서 피부 테스트에 사용하기 위해 제형화하는 것이 바람직하다. 이러한 조성물은 통상적으로 1종 이상의 상기 폴리펩티드를 0.1 ml의 부피내에 약 1 µg~약 100 µg, 바람직하게는 약 10~50 µg 양으로 포함할 수 있다. 이러한 약학 조성물내에 사용된 담체는 폐놀 및/또는 트윈 80<sup>TM</sup>과 같은 적절한 보존제가 함유된 염수 용액이다.

바람직한 구체예에서, 피부 테스트에 사용되는 폴리펩티드는 반응 기간 동안 주사 부위에 남아 있기에 충분한 크기이다. 일반적으로, 아미노산 길이가 9개 이상인 폴리펩티드면 충분하다. 또한, 폴리펩티드는 주사한지 수시간내에 대식세포에 의해 분해되어 T 세포에 제시되는 것이 바람직하다. 이러한 폴리펩티드는 상기 1종 이상의 서열 및/또는 기타 면역원성 또는 비면역원성 서열의 반복체를 포함할 수 있다.

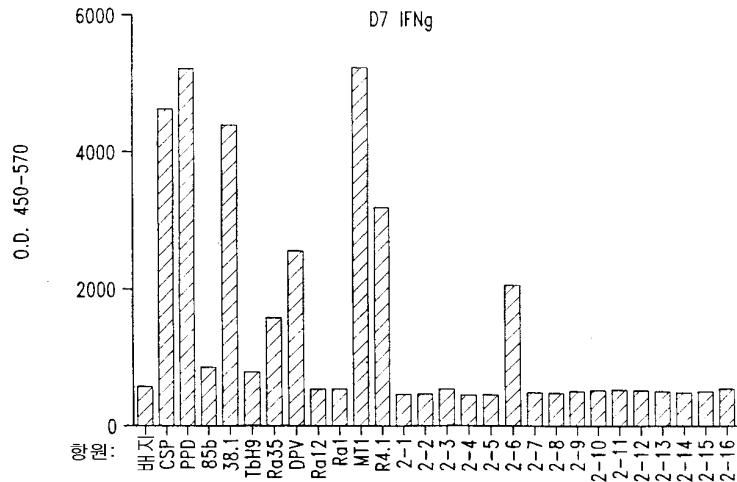
다음 실시예는 예시를 위해 제공되는 것이며, 본 발명을 제한하려는 것은 아니다.

## 도면

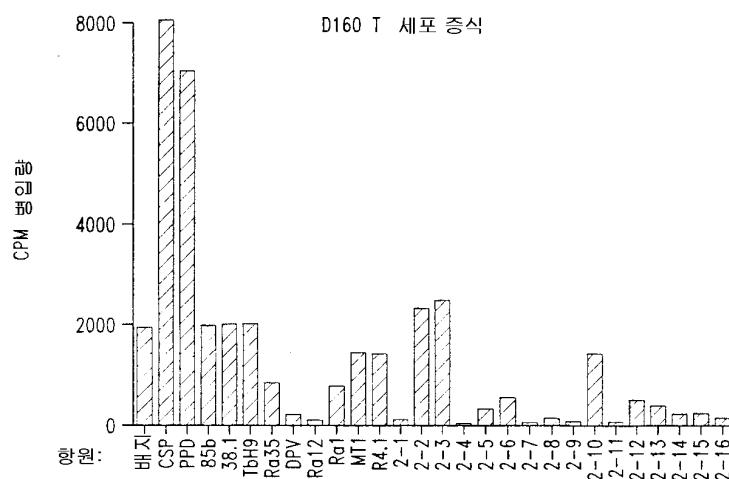
도면1a



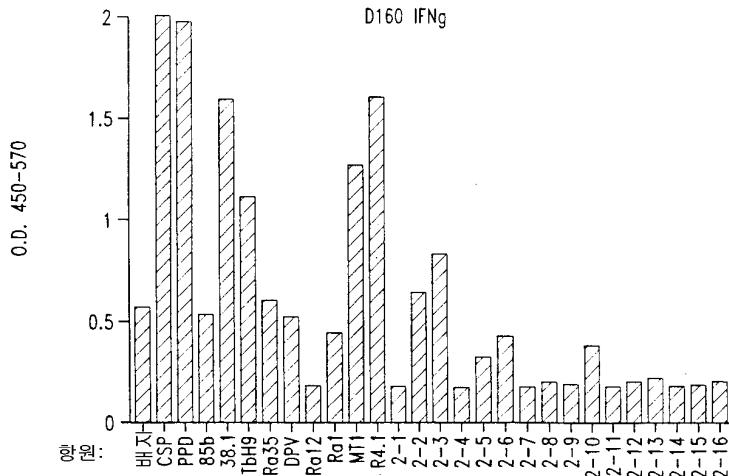
도면1b



도면2a



## 도면2b



## 서열목록

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: Corixa Corporation
- (B) STREET: Suite 200, 1124 Columbia Street
- (C) CITY: Seattle
- (D) STATE: Washington
- (E) COUNTRY: USA
- (F) POSTAL CODE (ZIP): 98104
- (G) TELEPHONE: (206) 754-5830
- (H) TELEFAX: (206) 754-5994
- (I) TELEX:

(ii) TITLE OF INVENTION: Compounds for Immunotherapy and Diagnosis  
of Tuberculosis and Methods of Their Use

(iii) NUMBER OF SEQUENCES: 144

## (iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Townsend and Townsend and Crew LLP
- (B) STREET: Two Embarcadero Center, Eighth Floor
- (C) CITY: San Francisco
- (D) STATE: California
- (E) COUNTRY: USA
- (F) ZIP: 94111-3834

## (v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Diskette
- (B) COMPUTER: IBM compatible
- (C) OPERATING SYSTEM: Windows
- (D) SOFTWARE: FastSEQ for Windows Version 2.0b

## (vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: PCT/US98/10407
- (B) FILING DATE: 20-MAY-1998
- (C) CLASSIFICATION:

- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/859,381
  - (B) FILING DATE: 20-MAY-1997
  - (A) APPLICATION NUMBER: US 09/073,010
  - (B) FILING DATE: 05-MAY-1998
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Bastian, Kevin L.
  - (B) REGISTRATION NUMBER: 34,774
  - (C) REFERENCE/DOCKET NUMBER: 14058-87-1PC
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (415) 576-0200
  - (B) TELEFAX: (415) 576-0300
  - (C) TELEX:

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1886 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGCTCTGGTG ACCACCAACT	TCTTCGGTGT	CAACACCATC	CCGATCGCCC	TCAACGAGGC	60
CGACTACCTG CGCATGTGGA	TCCAGGCCGC	CACCGTCATG	AGCCACTATC	AAGCCGTCGC	120
GCACGAAATC TGGTGTCTCC	ATGAATANGC	CAGTCGGGA	AAGCCGTGGG	CCAGTATCAC	180
CACGGGTGCG CCGGGCTCAC	CGGCCTCGAC	CACTCGCAGT	CGCACGCCGT	TGGTATCAAC	240
TAACCGTNCN GTANGTGC	GC	CCATCGTCTC	ACCAAATCAC	ACCGGGCACC	300
GGGCTTGGGG AGCANCCAGA	GGCGATTGTC	GCGGGTGCTG	CCGCGCATCA	TTGATCGGCC	360
GGCCGGACCA NTCGGGC	CTCTGACGTC	CGGATCNAC	TTCCTGTGCA	GCTGGCATGG	420
CTACAGCTCA CAGTGA	CTGC	CCCACGATTG	CCGGCCAGGT	CCAGTTCAA	480
TTCGCGGACA AAAGCAGCAG	GTCAACCAAC	CGCAGTCAGT	CGAGGGTCCC	AAACGTGAGC	540
CAATCGGTGA AATGGCTTGC	TGCA	CCGGTCACAG	GCTTAGCCGA	CAGCACCGGA	600
ATAGCTCAGG CGGGCTATAG	AGTCCTATAG	AAACATTTC	TGATAGAATT	AACCGCTGTC	660
TTGGCGTGAT CTTGATA	CG	CTCGCCGTGC	GACGGTTGG	CTCAGTAGCT	720
TAACCCATCC TCGGCAGGTG	TCTACTAAGG	CGAGACACCG	CATTGGTGGG	GCTGCATCGC	780
AAATCGGTCC GAGCATGTAG	CACTGCCGTT	ATCCCCGGAT	AGCAAACCCAC	CCGGAACCA	840
GGCTATCCCA GTCGCTCTCC	GACGGAGGCC	GTTCGCTTT	CCGTTGCCG	ATAACTCCCG	900
AGTGGATATC GGCGTTATCA	NATTCA	AGGCTTACGGT	TTTCTTCGCA	AGGTACCGGT	960
TTCGGATATC TCGGACGGAT	AATTACTAAA	ACTTCAGTGG	TTTAGATAAG	GCCGCCGCAA	1020
TACTTCGCCG ATCTTGCCGA	GCGCAACGGA	TTTCCATCGT	CGGTTTTCGT	CGCCTTATCA	1080
AACATGATCG GAGATAATGA	CAGATCGGCC	TAGCTAGGTG	TTTAGCGGAC	GCGATTAGG	1140
ACAACCGAGA TTTGCTTGC	CTCGCAACCA	TGAGAGCGCC	CCGCTTCGAC	GCCGAATCGG	1200
GTGAGTGATG GTGGGTTAGC	ACAGCCCTGA	TTGCGCCACC	GGCGAGGTGA	TTGTGCCCGC	1260
CACGAGGCCG CCGCCGGCTA	GCCCCATGAG	CACGNTATAT	AGACTCTCCT	GCAACAGATC	1320
TCATACCGAT CGAAGGCGAA	GCGCAGGCAT	CGACGTCGGA	GACACTGCCT	TGGGATCGCG	1380
CCGCCTACAC GGCAGGTTGGC	GCATTGTCGC	AGCGCAGTTG	CAGGAGGGCA	AATGTGCGCA	1440
GACGATGTAG TCGACAACAA	GTGNACATGC	CGTCTTCACG	AACTCAAAAC	TGACGATCTG	1500
CTTAGCATGA AAAAAGTGT	TGACATCGGC	CAAGCATGAC	AGCCAGACTG	TAGGCCTACG	1560
CGTGCAATGC AGAACCAAGG	NTATGCATGG	AATCGACGAC	CGTTGAGATA	GGCGGCAGGC	1620
ATGAGCAGAG CGTTCATCAT	CGATCCAACG	ATCAGTGC	TTGACGGCTT	GTACGACCTT	1680
CTGGGGATTG GAATACCAA	CCAAGGGGGT	ATCCTTTACT	CCTCACTAGA	GTACTTCGAA	1740
AAAGCCCTGG AGGAGCTGGC	AGCAGCGTTT	CCGGGTGATG	GCTGGTTAGG	TTCGGCCGCG	1800
GACAAATACG CCGGAAAAAA	CCGCAACCAC	GTGAATT	TCCAGGA	GGCAGACCTC	1860

GATCGTCAGC TCATCAGCCT GATCCA

1886

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2305 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGCACCGCGCT GGCCGCGCAA TACACCGAAA TTGCAACGGA ACTCGCAAGC GTGCTCGCTG	60
CGGTGCAGGC AAGCTCGTGG CAGGGGCCA GCGCCGACCG GTTCGTCGTC GCCCATCAAC	120
CGTTCCGGTA TTGGCTAACCC CACGCTGCCA CGGTGGCCAC CGCAGCAGCC GCCGCGCACN	180
AAACGGCCGC CGCCGGGTAT ACGTCCGCAT TGGGGGGCAT GCCTACGCTA GCCGAGTTGG	240
CGGCCAACCA TGCCATGCAC GGCGCTCTGG TGACCACCAA CTTCTTCGGT GTCAACACCCA	300
TCCCGATCGC CCTCAACGAG GCCGACTACC TGCGCATGTG GATCCAGGCC GCCACCGTCA	360
TGAGCCACTA TCAAGCCGTC GCGCACGAAA GCGTGGCGGC GACCCCCAGC ACGCCGCCGG	420
CGCCGCGAGAT AGTGACCAGT GCGGCCAGCT CGGGGGCTAG CAGCAGCTTC CCCGACCCGA	480
CCAAATTGAT CCTGCAGCTA CTCAAGGATT TCCTGGAGCT GCTGCGCTAT CTGGCTGTTG	540
AGCTGCTGCC GGGGCCGCTC GGCGACCTCA TCGCCCAGGT GTTGGACTGG TTCATCTCGT	600
TCGTGTCCGG TCCAGTCTTC ACGTTCTCG CCTACCTGGT GCTGGACCCA CTGATCTATT	660
TCGGACCGTT CGCCCCGCTG ACGAGTCCGG TCCTGTTGCC TGCTGTGGAG TTACGCAACC	720
GCCTCAAAAC CGCCACCGGA CTGACGCTGC CACCTACCGT GATTTTCGAT CATCCCACTC	780
CCACTGCGGT CGCCGAGTAT GTCGCCAGC AAATGTCTGG CAGCCGCCA ACGGAATCCG	840
GTGATCCGAC GTCGAGGTT GTCGAACCCG CTCGTGCCGA ATTCCGGACG AGTGTGTT	900
ATCAAATCCC CCCGAGACCT GCGGACACCC GGCGCGCTTG CCGACATCGA GATGATGTCC	960
CGCGAGATAG CAGAATTGCC CAACATCGT ATGGTGCAGG GCTTGACCCG ACCGAACGGG	1020
GAACCTCTGA AGGAGACCAA GGTCTCGTT CAGGCTGGTG AAGTGGCGG CAAGCTGAC	1080
GAAGCGACCA CCCTGCTCGA AGAGCACCGGA GGCAGCTGG ACCAGCTGAC CGGGGGTGC	1140
CACCAAGTGG CCGACGCCCT CGCCCAAATA CGCAACGAAA TCAATGGGGC CGTGGCCAGC	1200
TCGAGCGGGA TAGTCAACAC CCTGCAGGCC ATGATGGACC TGATGGCGG TGACAAGACC	1260
ATCCGACAAC TGGAAAATGC GTCCCAATAT GTCGGCGCA TGCGGGCTCT GGGGGACAAT	1320
CTGAGCGGGA CCGTCACCGA TGCCGAACAA ATCGCCACTT GGGCCAGCCC TATGGTCAAC	1380
GCCCTCAACT CCAGCCCGGT GTGTAACAGC GATCCCGCCT GTCGGACGTC GCGCGCACAG	1440
TTGGCGGCGA TTGTCAGGC GCAGGACGAC GGCGCTGCTCA GGTCCATCAG AGCGCTAGCC	1500
GTCACCCCTGC AACAGACGCA GGAATACCAAG ACACTCGCCC GGACGGTGAG CACACTGGAC	1560
GGGCAACTGA AGCAAGTCGT CAGCACCCCT AAAGCGGTG ACGGCCTACC CACCAAATTG	1620
GCTCAAATGC AGCAAGGAGC CAACGCTCTC GCCGACGGCA GCGCAGCGCT GCGGGCAGGC	1680
GTGCAGGAAT TGGTCGATCA GGTCAAAAAG ATGGGCTCAG GGCTCAACGA GGCGCCGAC	1740
TTCCCTGTTGG GGATCAAGCG GGATGCGGAC AAGCCGTCAA TGGCGGGCTT CAACATTCCA	1800
CCCGCAGATT TTTCGAGGGC CGAGTTCAAG AAGGGCGCCC AGATTTCTTG TCGGGCCGAT	1860
GGTCATGCGG CGCGGTACTT CGTGCAGAGC GCGCTGAATC CGGCCACAC CGAGGCGATG	1920
GATCAGGTCA ACGATATCCT CCGTGTGCG GATTCCGCGC GACCGAATAC CGAACCTCGAG	1980
GATGCCACGA TAGGTCTGGC GGGGGTTCCG ACTGCGCTGC GGGATATCCG CGACTACTAC	2040
AACAGCGATA TGAAATTCA CGTCATTGCG ACGATCGTTA TCGTATTCTT GATTCTCGTC	2100
ATTCTGNTGC GCGCACTTGT GGNNTCCGATA TATCTGATAG GCTCGGTGCT GATTCTTAC	2160
TTGTCGGGCC TAGGCATAGG AACTTTCGTT TTCCAATTGA TACTGGGCCA GGAAATGCAT	2220
TGGAGCCTGC CGGGACTGTC CTTCATATTA TTGGTTGCCA TCGGCGCTGA CTACAACATG	2280
CTGCTCATTT CACGCATCCG CGACG	2305

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1742 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGCTCTCTT	TCAACGTCAT	AAGTTGGTG	GGCCAGTCGG	CCGCCTGTC	ATATGGCACC	60
ATAAACGCGT	GTCCCATGGA	TACCCGGACC	GCACGACGGT	AGAGCGGATC	AGGCCAGCG	120
GTGCCGAACA	CTACCGCGTC	CACGCTCAGC	CCTGCCGCGT	TGCGGAAGAT	CGAGCCCAGG	180
TTCTCATGGT	CGTTAACGCC	TTCCAACACT	GCGACGGTGC	GCGCCCCGGC	GACCACCTGA	240
GCAACGCTCG	GCTCCGGCAC	CCGGCGCGCG	GCTGCCAACA	CCCCACGATT	GAGATGGAAG	300
CCGATCACCC	GTGCCATGAC	ATCAGCCGAC	GCTCGATAGT	ACGGCGGCC	GACACCGGCC	360
AGATCATCCT	TGAGCTCGGC	CAGCCGGCGG	TCGGTGCCGA	ACAGCGCCAG	CGCGTGAAC	420
CGTGAGGCCA	GCATGCGCTG	CACCACCAGC	ACACCCCTCGG	CGATCACCAA	CGCCTTGCCG	480
GTCGGCAGAT	CGGGACNACN	GTCGATGCTG	TTCAGGTCAC	GGAAATCGTC	GAGCCGTGGG	540
TCGTCGGGAT	CGCAGACGTC	CTGAACATCG	AGGCCGTCGG	GGTGCTGGC	ACAACGGCCT	600
TCGGTCACCG	GCTTCGTG	ACCAGAGCCA	GCATCAGATC	GGCGCGCTG	CGCAGGATGT	660
CACGCTCGCT	GCGGTTCAAGC	GTCGCGAGCC	GCTCAGCCAG	CCACTCTTGC	AGAGAGCCGT	720
TGCTGGGATT	AATTGGGAGA	GGAAGACAGC	ATGTCGTTCG	TGACCACACA	GCCGGAAGCC	780
CTGGCAGCTG	CGGCGGCGAA	CCTACAGGGT	ATTGGCACGA	CAATGAACGC	CCAGAACGCG	840
GCCGCGGCTG	CTCCAACCAC	CGGAGTAGTG	CCCGCAGCCG	CCGATGAAGT	ATCAGCGCTG	900
ACCGCGGCTC	AGTTTGCTGC	GCACGCGCAG	ATGTACCAAA	CGGTCAGCGC	CCAGGCCGCG	960
GCCATTACAG	AAATGTTCGT	GAACACGCTG	GTGGCCAGTT	CTGGCTCATA	CGCGGCCACC	1020
GAGGGCGGCCA	ACGCAGCCGC	TGCGGCTGA	ACGGGCTCGC	ACGAACCTGC	TGAAGGAGAG	1080
GGGGAACATC	CGGAGTTCTC	GGGTCAGGGG	TTGCGCCAGC	GCCCAGCCGA	TTCAGNTATC	1140
GGCGTCCATA	ACAGCAGACG	ATCTAGGCAT	TCAGTACTAA	GGAGACAGGC	AACATGGCCT	1200
CACGTTTTAT	GACGGATCCG	CATGCGATGC	GGGACATGGC	GGGCGTTTT	GAGGTGCACG	1260
CCCAGACGGT	GGAGGACGAG	GTCGCGCGA	TGTGGCGTC	CGCGCAAAC	ATTCCGGTG	1320
CGGGCTGGAG	TGGCATGGCC	GAGGCGACCT	CGCTAGACAC	CATGACCTAG	ATGAATCAGG	1380
CGTTTCGCAA	CATCGTGAAC	ATGCTGCACG	GGGTGCGTGA	CGGGCTGGTT	CGCGACGCCA	1440
ACAANTACGA	ACAGCAAGAG	CAGGCCTCCC	AGCAGATCCT	GAGCAGNTAG	CGCCGAAAGC	1500
CACAGCTGNG	TACGNNTTCT	CACATTAGGA	GAACACCAAT	ATGACGATTA	ATTACCAAGTT	1560
CGGGGACGTC	GACGCTCATG	GCGCCATGAT	CCCGCCTCAG	GCGGCGTCGC	TTGAGGCAGGA	1620
GCATCAGGCC	ATCGTTCTG	ATGTGTTGGC	CGCGGGTGC	TTTGGGGCG	GCGCCGGTTC	1680
GGTGGCTTGC	CAGGAGTTCA	TTACCCAGTT	GGGCCGTAAC	TTCCAGGTGA	TCTACGAGCA	1740
GG						1742

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2836 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTTGATTCCG	TTCGCGCGC	CGCCGAAGAC	CACCAACTCC	GCTGGGGTGG	TCGCACAGGC	60
GGTTGCGTCG	GTCAGCTGGC	CGAATCCAA	TGATTGGTGG	CTCNGTCCGG	TTGCTGGGCT	120
CGATTACCCC	CACGGAAAGG	ACGACGATCG	TTCGTTGCT	CGGTCACTCG	TACTTGGCGA	180
CGGGCATGGC	CGGGTTCTT	ACCTCGATCG	CACAGCAGCT	GACCTTCGGC	CCAGGGGGCA	240
CAACGGCTGG	CTCCGGCGGA	GCCTGGTACC	CAACGCCACA	ATTGCGCCGC	CTGGGTGCAG	300
GCCCGGCGGT	GTCGGCGAGT	TTGGCGCGGG	CGGAGCCGGT	CGGGAGGTTG	TCGGTGCAGC	360
CAAGTTGGC	CGTCGCGGCT	CCGGCCTTCG	CGGAGAAGCC	TGAGGCAGGC	ACGCCGATGT	420
CCGTCATCGG	CGAAGCGTCC	AGCTGCGGTC	AGGGAGGCCT	GCTTCGAGGC	ATACCGCTGG	480
CGAGAGCGGG	CGGGCGTACA	GGCGCCTTCG	CTCACCGATA	CGGGTTCCGC	CACAGCGTGA	540
TTACCCGGTC	TCCGTCGGCG	GGATAGCTT	CGATCCGGTC	TGCGCGCCCG	CCGGAAATGC	600

TGCAGATAGC	GATCGACCGC	GCCGGTCGGT	AAACGCCGCA	CACGGCACTA	TCAATGCGCA	660
CGGCGGGCGT	TGATGCCAAA	TTGACCGTCC	CGACGGGGCT	TTATCTGCGG	CAAGATTCA	720
TCCCCAGCCC	GGTCGGTGGG	CCGATAAAATA	CGCTGGTCAG	CGCGACTCTT	CCGGCTGAAT	780
TCGATGCTCT	GGGCGCCCGC	TCGACGCCGA	GTATCTCGAG	TGGGCCGCAA	ACCCGGTCAA	840
ACGCTGTTAC	TGTGGCGTTA	CCACAGGTGA	ATTTGCAGTG	CCAACCTGGTG	AACACTTGCG	900
AACGGGTGGC	ATCGAAATCA	ACTTGTGCG	TTGCAGTGAT	CTACTCTCTT	GCAGAGAGCC	960
GTTGCTGGGA	TTAATTGGGA	GAGGAAGACA	GCATGTCGTT	CGTGACCACA	CAGCCGGAAG	1020
CCCTGGCAGC	TGCGGCGGCG	AACCTACAGG	GTATTGGCAC	GACAATGAAC	GCCCAGAACG	1080
CGGCCCGCGC	TGCTCCAACC	ACCGGAGTAG	TGCCCGCAGC	CGCCGATGAA	GTATCAGCGC	1140
TGACCGCGGC	TCAGTTGCT	GCGCACGCGC	AGATGTACCA	AACGGTCAGC	GCCCAGGCCG	1200
CGGCCATTCA	CGAAATGTTC	GTGAACACGC	TGGTGGCCAG	TTCTGGCTCA	TACCGGGCCA	1260
CCGAGGCGGC	CAACGCAGCC	GCTGCCGGCT	GAACGGGCTC	GCACGAACCT	GCTGAAGGAG	1320
AGGGGGAACA	TCCGGAGTTC	TCGGGTCAGG	GGTTGCGCCA	GCGCCCAGCC	GATTCACTA	1380
TCGGCGTCCA	TAACACGAGA	CGATCTAGGC	ATTCAGTACT	AAGGAGACAG	GCAACATGGC	1440
CTCACGTTT	ATGACGGATC	CGCATGCGAT	GCAGGACATG	GCAGGGCCGTT	TTGAGGTGCA	1500
CGCCCAGACG	GTGGAGGACG	AGGCTCGCCG	GATGTGGCG	TCCGCGCAA	ACATTCCGG	1560
TGCGGGCTGG	AGTGGCATGG	CCGAGGCGAC	CTCGCTAGAC	ACCATGACCT	AGATGAATCA	1620
GGCGTTTCGC	AACATCGTGA	ACATGCTGCA	CGGGGTGCGT	GACGGGCTGG	TTCGCGACGC	1680
CAACAACTAC	GAACAGCAAG	AGCAGGCCCTC	CCAGCAGATC	CTGAGCAGCT	AGGCCGAAA	1740
GCCACAGCTG	CGTACGCTT	CTCACATTAG	GAGAACACCA	ATATGACGAT	TAATTACCA	1800
TTCGGGGACG	TCGACGCTCA	TGGGCCATG	ATCCGCGCTC	AGGCAGGCGTC	GCTTGAGGCG	1860
GAGCATCAGG	CCATCGTTCG	TGATGTGTTG	GCCGCGGGTG	ACTTTTGGGG	CGGCGCCGGT	1920
TCGGTGGCTT	GCCAGGAGTT	CATTACCCAG	TTGGGCCGTA	ACTTCCAGGT	GATCTACGAG	1980
CAGGCCAACG	CCCACGGGCA	GAAGGGCGAG	GCTGCCGGCA	ACAACATGGC	GCAAACCGAC	2040
AGGCCCGTCG	GCTCCAGCTG	GGCCTAAAAC	TGAACCTTCAG	TCGCGGCAGC	ACACCAACCA	2100
GCCGGTGTG	TGCTGTGTCC	TGCAGTTAAC	TAGCACTCGA	CCGCTGAGGT	AGCGATGGAT	2160
CAACAGAGTA	CCCACCGCGA	CATCACCGTC	AACGTCGACG	GCTTCTGGAT	GCTTCAGGCG	2220
CTACTGGATA	TCCGCCACGT	TGCGCCTGAG	TTACGTTGCC	GGCGTACGT	CTCCACCGAT	2280
TCCAATGACT	GGCTAAACGA	GCACCCGGGG	ATGGCGGTCA	TGCGCGAGCA	GGGCATTGTC	2340
GTCAACGACG	CGGTCAACGA	ACAGGTCGCT	GCCCGGATGA	AGGTGCTTGC	CGCACCTGAT	2400
CTTGAAGTCG	TCGCCCTGCT	GTCACGCGGC	AAGTGCTGT	ACGGGGTCAT	AGACGACGAG	2460
AACCAGCCGC	CGGGTTCGCG	TGACATCCCT	GACAATGAGT	TCCGGGTGGT	GTGGGCCCGG	2520
CGAGGCCAGC	ACTGGGTGTC	GGCGGTACGG	GTTGGCAATG	ACATCACCGT	CGATGACGTG	2580
ACGGTCTCGG	ATAGCGCCTC	GATGCCGCA	CTGGTAATGG	ACGGTCTGGA	GTGATTACAC	2640
CACGCCGACC	CAGCCCGAT	CAACGCCGTC	AACGTGCCAA	TGGAGGAGAT	CTCGTGCCGA	2700
ATTCCGGACG	AGGCACGAGG	CGGTGTCGGT	GACGACGGGA	TCGATCACGA	TCATCGACCG	2760
GCCGGGATCC	TTGGCGATCT	CGTTGAGCAC	GACCCGGGCC	CGCGGGAAGC	TCTGCGACAT	2820
CCATGGGTT	TTCCCG					2836

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 900 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AACATGCTGC	ACGGGGTGC	TGACGGGCTG	GTTCGCGACG	CCAACAACTA	CGAGCAGCAA	60
GAGCAGGCCT	CCCAGCAGAT	CCTCAGCAGC	TAACGTCAGC	CGCTGCAGCA	CAATACTTT	120
ACAAGCGAAG	GAGAACAGGT	TCGATGACCA	TCAACTATCA	GTTCGGTGAT	GTGACGCTC	180
ACGGCGCCAT	GATCCCGCCT	CAGGCCGGGT	TGCTGGAGGC	CGAACATCAG	GCCATCATTC	240
GTGATGTGTT	GACCGCGAGT	GACTTTGGG	GCGGCGCCGG	TTCGGCGGCC	TGCCAGGGGT	300
TCATTACCCA	ATTGGGCCGT	AACTTCCAGG	TGATCTACGA	ACAGGCAAC	GCCCACGGGC	360

AGAAGGTGCA	GGCTGCCGGC	AACAACATGG	CGCAAACCGA	CAGCGCCGTC	GGCTCCAGCT	420
GGGCCTGACA	CCAGGCCAAG	GCCAGGGACG	TGGTGTACGA	GTGAAGGTTTC	CTCGCGTGAT	480
CCTTCGGGTG	GCAGTCTAGG	TGGTCAGTGC	TGGGGTGTG	GTGGTTGCT	GCTTGGCGGG	540
TTCTTCGGTG	CTGGTCAGTG	CTGCTCGGGC	TCGGGTGAGG	ACCTCGAGGC	CCAGGTAGCG	600
CCGTCCCTCG	ATCCATTCTG	CGTGTGTTTC	GGCGAGGACG	GCTCCGACGA	GGCGGATGAT	660
CGAGGCGCGG	TCGGGAAAGA	TGCCCACGAC	GTCGTTCGG	CGTCGTACCT	CTCGGTTGAG	720
GCGTTCCTGG	GGGTTGTTGG	ACCAGATTG	GCGCCAGATC	TTCTTGGGA	AGGGGGTGAA	780
CGCCAGCAGG	TCGGTGCAGG	CGGTGTCGAN	GTGCTCGGCC	ACCGCAGGGA	GTGGTCGGT	840
CAGAGCGTCG	AGTACCCGAT	CATATTGGGC	AACAACGTAT	TCGGCGTTGG	GCTGGTCGTA	900

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1905 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCTCGCCGGA	TGTGGGCGTC	CGCGCAAAAC	ATTTCCGGTG	CGGGCTGGAG	TGGCATGGCC	60
GAGGCGACCT	CGCTAGACAC	CATGGCCCAG	ATGAATCAGG	CGTTTCGCAA	CATCGTGAAC	120
ATGCTGCACG	GGGTGCGTGA	CGGGCTGGTT	CGCGACGCCA	ACAACACTACGA	GCAGCAAGAG	180
CAGGCCTCCC	AGCAGATCCT	CAGCAGCTAA	CGTCAGCCGC	TGCAGCACAA	TACTTTACA	240
AGCGAAGGAG	AACAGGTTCG	ATGACCATCA	ACTATCAGTT	CGGTGATGTC	GACGCTCACG	300
GCGCCATGAT	CCGCGCTCAG	GCCGGGTTGC	TGGAGGCCGA	GCATCAGGCC	ATCATTGTCG	360
ATGTGTTGAC	CGCGAGTGAC	TTTGGGGCG	GCGCCGGTTC	GGCAGGCTGC	CAGGGGTTCA	420
TTACCCAGTT	GGGCCGTAAC	TTCCAGGTGA	TCTACGAACA	AGCCAACACC	CACGGGCAGA	480
AGGTGCAAGC	TGCCGGCAAC	AACATGGGCC	AAACCGACAG	CGCCGTCNGC	TCCAGCTGGG	540
CCTGACACCA	GGCCAAGGCC	AGGGACGTGG	TGTACNAGTG	AAGGTTCTC	GCCTGATCCT	600
TCGGGTGGCA	GTCTAGGTGG	TCAGTGCTGG	GGTGGTTGGT	GTGTTGCTGCT	TGGCGGGTTC	660
TTCGGTGCTG	GTCAGTGCTG	CTCAGGCTCG	GGTGAGGACC	TCGAGGCCA	GGTAGCGCCG	720
TCCTTCGATC	CATTGTCGT	GTTGTTCGGC	GAGGACNGCT	CCGACGANGC	GGATGATCGA	780
GGCGCGGTGCG	GGGAAGATGC	CCACGACGTC	GGTTCGGCGT	CGTACCTCTC	GGTTGAAGCG	840
TTCCTGGGGG	CCACCGCTTG	GCGCCNANGC	ACTCCACGCC	AATTGTCNC	ACCTAACAGC	900
GGTGGCCAAC	GACTATGACT	ACGACACCGT	TTTGCCAGG	GCCCTCNAAA	GGATCTGCGC	960
GTCCCGGCCA	CACGCTTTT	GCGATAAGTA	CCTCCGGCAA	TTCTATGAGT	GTACTGCGGN	1020
CCCGGAAAAC	CGCAAGGGAG	TTGGGTGTGA	CGGTTNTTGC	AAATGACGGG	CBAATCCGGC	1080
GGCCAGCTGG	CAGAATTGCG	AGATTCTTG	ATCAACGTCC	CGTCACGCGA	CACCGGGCGA	1140
ATCCAGGAAT	CTCACATCGT	TTTTATTATC	GCGATCTCCG	AAATGTCGA	ACACCGCGTT	1200
TTCGCGCCTC	GCCAATAGGA	AAGCCGATCC	TTACCGGGCC	ATTGAAAGA	TGGTCGCGGA	1260
ACGTGCGGGA	CACCAATGGT	GTCTCTTCC	CGATAGAGAC	GGGGTCATCA	ATCGACAAGT	1320
GGTCGGCGAC	TACGTACCGA	ACTGGCGGA	GTGGAATGG	TTGCCCCGGG	CGGGCGGGC	1380
GTTGAAGAAG	CTACGGGCAT	GGGCTCCGTA	CATCGTTGTC	GTGACAAACC	AGCAGGGCGT	1440
GGGTGCCGGA	TTGATGAGCG	CCGTCGACGT	GATGGTGATA	CATCGGCACC	TCCAAATGCA	1500
GCTTGCATCC	GATGGCGTGC	TGATAGATGG	ATTTCAAGTT	TGCCCCGAC	ACCGTTCGCA	1560
GCGGTGTGGC	TGCCGTAAGC	CGAGACCGGG	TCTGGTCCTC	GACTGGCTCG	GACGACACCC	1620
CGACAGTGTAG	CCATTGCTGA	GCATCGTGGT	TGGGGACAGC	CTCAGCGATC	TTGACATTGG	1680
CACACAACGT	CGCCGCTGCT	GCCGGTGCAT	GTGCCAGTGT	CCAGATAAGG	GGGCCAGTT	1740
CTGGCGGTGT	CGCTGACGCG	TCATTGACT	CGCTCTGGGA	GTTCGCTGTC	GCAGTCGGAC	1800
ATGCGCGGGG	GGAGCGGGGC	TAATGGCGAT	CTTGGCGGGG	CGAGCGCCGT	NGCGGNTCGG	1860
ACTNNNGCGGT	GGCGGGACAG	ACGTGGAACC	GTACTCGAGC	CAGTT		1905

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2921 base pairs

- (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGGATGCCG	TGGTGGTTGG	TATTGCCCAA	ACCCTGGCGC	TGGTCCCCGG	GGTATCCAGG	60
TCCGGGTCGA	CCATCAGCGC	TGGACTGTTT	CTCGGACTCG	ACCGTGAAC	GGCCGCCCGA	120
TTCGGATTCC	TGCTGGCCAT	TCCAGCGGTG	TTCGCCTCCG	GGTTGTTCTC	GTTGCCGAC	180
GCATTCCACC	CGGTAACCGA	GGGCATGAGC	GCTACTGGCC	CGCAGTTGCT	GGTGGCCACC	240
CTGATCGCGT	TCGTCCCTCGG	TCTGACCGCG	GTGGCCTGGC	TGCTCGCGTT	TCTGGTGCAG	300
CACAACATGT	ACTGGTTCGT	CGGCTACCGG	GTGCTCGTCG	GGACGGGCAT	GCTCGTGCTG	360
CTGGCTACCG	GGACGGTAGC	CGCGACATGA	CCGTCATCTT	GCTACGCCAT	GCCCGTTCCA	420
CCTCGAACAC	CGCGGGCGTG	CTGGCCGGCC	GGTCCGGCGT	CGACCTCGAC	GAGAAGGGGC	480
GCGAGCAGGC	CACCGGGTTG	ATCGATCGAA	TTGGTGACCT	GCCGATCCGG	GCGGTGCGGT	540
CTTCTCCAAT	GCTGGGGTGT	CAACGCACCG	TCGAACCGCT	GGCCGAGGCG	CTGTGCCTGG	600
AGCCGCTCAT	CGATGACCGG	TTCTCCGAAG	TCGACTACGG	CGAATGGACT	GGCAGAAAAA	660
TCGGTGACCT	GGTCGACGAG	CCGTTGTGGC	GGGTAGTCCA	GGCCCACCCC	AGCGCGGCGG	720
TGTTTCCCGG	CGGTGAGGGT	TTGGCGCAGG	TGCAAGACGTG	GTTGTCCTGA	CGGATTTCCA	780
TGCCGGGGAA	CACCAAGACC	GGATCGGCAC	TGGCGGTCGC	CGGCAGAAC	CCGGCCGCCA	840
ATAGGGCGAC	CGTCGCTGCG	AATGCGCGT	GTACCAGGCG	GACCACCTTG	AACCTCCATC	900
CGTCGGGGCC	AAGCGCATCG	CCCGCCGCGC	GTTACGGCTA	AGGCACGATCA	AAACCCGACG	960
GTAATACTTC	GGCAATGTCG	GGTCNCGACG	TTACCGAGAC	GTGACCAGNG	AGGCNGCGGC	1020
ATTGGATTTA	TCGATGGTGC	GCGGTTCCCA	NCCCGGCGGT	CCGAANACGT	AGCCCAGCCG	1080
ATCCCGCAGA	CGTGGTGCAG	ACCGCCAGTC	ACGCACGATC	GCCACGTACT	CGCGGGTCTG	1140
CAGCTCCAG	ATGTTGAACG	TGTCGACCCG	CTTGGTCAGG	CCATAATGCG	GTCGGAATAG	1200
CTCCGGCTGA	AAGCTACCGA	ACAGGCGGTC	CCAGATGATG	AGGATGCCGC	CATAGTTCTT	1260
GTCCANATAC	ACCGGGTCCA	TTCCGTGGT	GACCCGGTGG	TGCGACGGGG	TATTGAAGAC	1320
GAATTCGAAC	CACCGCGGCA	GCCTGTCGAT	CCGCTCGGTG	TGCACCCAGA	ACTGGTAGAT	1380
CAAGTTCAGC	GACCAATTGC	AGAACACCAT	CCAAGGGGGA	AGCCCCATCA	GTGGCAGCGG	1440
AACCCACATG	AGAATCTCGC	CGCTGTTGTT	CCANTTTCTG	GCGCAGCGCG	GTGGCGAAGT	1500
TGAAGTATTTC	GCTGGAGTGA	TGCGCCTGGT	GGGTAGCCCA	GATCAGCCGA	ACTCGGTGGG	1560
CGATGCGGTG	ATAGGAGTAG	TACAGCAGAT	CGACACCAAC	GATCGCGATC	ACCCAGGTGT	1620
ACCACCGGTG	GGCGGACAGC	TGCCAGGGGG	CAAGGTAGGC	ATAGATTGCG	GCATAACCGA	1680
GCAGGGCAAG	GGACTTCCAG	CCGGCGGTGG	TGGCTATCGA	AACCAGCCCC	ATCGAGATGC	1740
TGGCCACCGA	GTCGCGGGTG	AGGTAAGCGC	CCGAGGCGGG	CCGTGGCTGC	CCGGTAGCAG	1800
CGGTCTCGAT	GCTTTCAGC	TTGCGGGCCG	CCGTCCATTG	GAGAATCAGC	AGCAATAGAA	1860
AACATGGAAT	GGCGAACAGT	ACCGGGTCCC	GCATTTCTC	GGGCAGCGCT	GAGAAGAATC	1920
CGCGCACGGC	ATGGCCGAGG	CGACCTCGNT	AGACACCATG	ACCCAGATGA	ATCAGGCCTT	1980
TCGCAACATC	GTGAACATGC	TGCACGGGGT	GCGTGACCGG	CTGGTTCCGC	ACGCCAACAA	2040
NTACGAACAG	CAAGAGCAGG	CCTCCCAGCA	GATCCTCAGC	AGCTGACCCG	GCCCGACGAC	2100
TCAGGAGGAC	ACATGACCAT	CAACTATCAA	TTCGGGGACG	TCGACGCTCA	CGGCACCATG	2160
ATCCCGCCTC	AGGCCGGTC	GCTGGAGGCC	GAGCATCAGG	CCATCATTTG	TGATGTGTTG	2220
ACCGCGAGTG	ACTTTGGGG	CGGCGCCGGT	TCGGCGGCCT	GCCAGGGGTT	CATTACCCAG	2280
CTGGGCCGTA	ACTTCCAGGT	GATNTACGAG	CAGGCCAACG	CCCACGGGCA	GAAGGTGCAG	2340
GCTGCCGGCA	ACAACATGGC	ACAAACCGAC	AGCGCCGTG	GCTCCAGCTG	GGCATAAAGN	2400
TGGCTTAAGG	CCCGCGCCGT	CAATTACAAC	GTGGCCGCAC	ACCGGTTGGT	GTGTGGCCAC	2460
GTTGTTATCT	GAACGACTAA	CTACTTCGAC	CTGCTAAAGT	CGGCGCGTTG	ATCCCCGGTC	2520
GGATGGTGCT	GAACCTGGAA	GATGGCCTCA	ATGCCCTTGT	TGCGGAAGGG	ATTGAGGCCA	2580
TCGTGTTTCG	TACTTTAGGC	GATCAGTGCT	GGTTGTGGGA	GTCGCTGCTG	CCCGACGAGG	2640
TGCGCCGACT	GCCCGAGGAA	CTGGCCCGGG	TGGACGCATT	GTTGGACGAT	CCGGCGTTCT	2700
TCGCCCCGTT	CGTGCCGTTC	TTCGACCCGC	GCAGGGGCCG	GCCGTCGACG	CCGATGGAGG	2760
TCTATCTGCA	GTTGATGTTT	GTGAAGTTCC	GCTACCGGCT	GGGCTATGAG	TCGCTGTGCC	2820

GGGAGGTGGC TGATTCGATC ACCTGACGGC GGTTTGCCG CATTGCCTG GACGGGTCGG	2880
TGCCGCATCC GACCACATTG ATGAAGCTCA CCACCGTTC C	2921

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1704 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCGATCGTC GTCAACGANG TCGACCGTCA CCACGGACTG ATCAACAAGT TCGCAGGCGA	60
CGCCGCCCTG GCCATCTTCG GAGCCCCGAA CCGCCTCGAC CGTCCCAGAG ACGCCGCGCT	120
GGCCGCCGCC CGGGCCATAN CCGANCAGCT GGCCNACGAG ATGCCCGAGG TCCAAGCCGG	180
CATCGGGGTG GCGGCAGGCC ANATCGTCGC CGGCAATGTC GGCGCCAAGC AAAGATTCTNA	240
ATACACAGTG GTCGGCAAGC CGGTCAACCA NGCGGCCGA TTGTGCGAAC TGGCAAATC	300
ACACCCCGCG CGATTGGGTG TCGCCCGCTC GGCTCATGGT CACCCAATTC AAGGACTACT	360
TTGGCCTGGC GCACGACCTG CCGAAGTGGG CGAGTGAAGG CGCCAAAGCC GCGGGTGAGG	420
CCGCCAAGGC GTTGCCTGGC GCCGTTCCGG CCATCCGAG TGCTGGCCTG AGCGGCGTTG	480
CGGGCGCCGT CGGTCAAGCG GCGTCGGTCG GGGGATTGAA GGTTCCGCC GTTGGACCG	540
CCACGACCCC GGCGGCGAGC CCCGCGGTGC TGGCGGCCGTC CAACGGCCTC GGAGCCGCGG	600
CCGCGCTGA AGGTTCGACA CACCGTTG GCGGGATGCC GCTCATGGGT ANCAGTGCCG	660
GACGTGCGTT TAACAACCTTC GCTGCCCTC GATACGGATT CAAGCCGACC GTGATGCC	720
AACCGCCGGC TGGCGGATGA CCAACTACGT TCGTTGATCG AGGATCGAAT TCNACGATTC	780
AAAGGGAGGA ATTCAATATGA CCTCNCGTTT TATGACGGAT CCGCACGCNA TNCGGGACAT	840
GGCGGGCCGT TTTGAGGTGC ACGCCCAGAC GGTGGAGGAC GAGGCTNGCN GGATGTGGC	900
GTCCCGCGAA AACATTCCG GTGCGGGCTG GAGTGGCATG GCCGAGGCCA CCTCGNTAGA	960
CACCATGGCC CAGATGAATC AGGCCTTTCN CAACATCGTG AACATGCTGC ACGGGGTNG	1020
TGACGGGCTG GTTCGCGACG CCAACAACTA CGAACAGCAA GAGCAGGCC CCCAGCAGAT	1080
CCTCAGCAGC TGACCCGGCC CGACGACTCA GGAGGACACA TGACCATCAA CTATCAATT	1140
GGGGACGTG ACGCTCATGG CGCCATGATC CGCGCTNTGG CCGGGTTGCT GGAGGCCGAG	1200
CATCAGGCCA TCATTCTGA TGTGTTGACC GCGAGTGACT TTTGGGCCG CGCCGGTTC	1260
CGGGCCTGCC AGGGGTTCAT TACCCAGTTG GGCGCTAACT TCCAGGTGAT TTACGAGCAG	1320
GCCAACGCC ACGGGCAGAA GGTGCAGGCT GCCGGCAACA ACATGGCACA AACCGACAGC	1380
GCCGTNGGNT CCAGCTGGC CTAACCCGGG TCNTAAGTTG GGTCCCGCGCA GGGCGGGCCG	1440
ATCAGCGTNG ACTTTGGCGC CCGATACACG GGCATNTNT NGTCGGGAAC ACTCGGCC	1500
CGTCAGNTGC CCGCTTCCCC TTGTTNGCGC ACGTGCTCG TGATGGCTT GACGACCGCT	1560
TCGCCGGCCG GGCCAATCAA TTGGTCGCGC TTGCTNTAG CCCATTGCG CGACGCCCGC	1620
GGCGCCGCGA GTTGTCCCTT GAAATAAGGA ATCACAGCAC GGGCGAACAG CTCATAGGAG	1680
TGAAAGGTTG CGTGGCGGG GCCC	1704

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2286 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCGTCTTGGC GTCTGGCGC ATTGTGATCT GGGCCANTTG CCCCTCCACC CAGACCGCGC	60
CCAGCTTGTG GATCCAGCCC GCGACCCGGA TTGCCACCCTC GCGAACCGGG AACGGATTCT	120
CCGCTGAATT CTGGGTCACT TCGCAGTCGC CGGGGTGATC CTGTTGGCGA NCAGCGTCTG	180
GAACGGGCCT CNAACCGTG CCGTAAGCCC AGCGTGTACG CCGTCAGCCC GACGCCGATG	240
CCGAATGCCCT TGCCGCCAA GCTGAGCCGC CGGGCTCCA CCAAGAGCGT CACGGTGAGC	300

CAGCCAACCA	GATGCAAGGC	GACGATCACC	GCGAAGTGC	GAATTGGCA	CGAGAGGTGC	360
TGGAAATCCA	GCAATACGCC	CGCGAGCCGA	TCTCGTTGGA	CCAGACCATC	GGCGACGANG	420
GCGACAGNCA	GCTTGGCGAT	TTCATCGAAA	ACAGCGAGGC	GGTGGTGNC	GTCGACGCGG	480
TGTCTTCAC	TTTGCTGCAT	GATCAACTGC	ANTCGGTGCT	GGACACGCTC	TCCGAGCGTG	540
AGGCGGGCGT	GGTGC GGCTA	CGCTTCGGCC	TTACCGACGG	CCAGCCGCGC	ACCCTTGACG	600
AGATCGGCCA	GGTCTACGGC	GTGACCCGGG	AACGCATCCG	CCAGATCGAA	TCCAAGACTA	660
TGTGAAAGTT	GCGCCATCCG	AGCCGCTCAC	AGGTCTCGCG	CGACTATCGT	GCCGAATTG	720
GCACGAGCCG	TTTGAGGTG	CACGCCAGA	CGGGTGGAGGA	CGAGGCTCGC	CGGATGTGGG	780
CGTCCGCGCA	AAACATTTCC	GGTGC GGCGT	GGAGTGGCAT	GGCCGANGCG	ACCTCGCTAG	840
ACACCATGGC	CCAGATGAAT	CAGGC GTTTC	GCAACATCGT	GAACATGCTG	CACGGGGTGC	900
GTGACGGGCT	GGTTCGCGAC	GCCAACAACT	ACGAACAGCA	AGAGCAGGCC	TCCCAGCAGA	960
TCCTCAGCAG	CTGACCCGGC	CCGACGACTC	AGGAGGACAC	ATGACCATCA	ACTATCAATT	1020
CGGGGACGTC	GACGCTCATG	GCGCCATGAT	CCCGCCTCTG	GCCGGGTTGC	TGGAGGCCGA	1080
GCATCAGGCC	ATCATTCTG	ATGTGTTGAC	CGCGAGTGC	TTTGGGGCG	GCGCCGGTTC	1140
GGCGGCCTGC	CAGGGGTTCA	TTACCCAGTT	GGGCGTAAC	TTCCAGGTGA	TCTACGAGCA	1200
GGCCAACGCC	CACGGGCAGA	AGGTGCAGGC	TGCCGGCAAC	AACATGGCAC	AAACCGACAG	1260
CGCCGTCGGC	TCCAGCTGGG	CCTAACCCGG	GTCCTAAGTT	GGGTCCGCGC	AGGGCGGGCC	1320
GATCAGCGTC	GACTTGGCG	CCCGATAACAC	GGGCATGTNG	TNGTCGGGAA	CACTGCGCCC	1380
CGTCAGCTG	CCCGCTTCCC	CTTGGTCGGC	GACGTGCTCG	GTGATGGCTT	TGACGACCGC	1440
TTCGCCGGCG	CGGCCAATCA	ATTGGTCGCG	CTTGCCTCTA	GCCTCGTGCC	GAATTGGCA	1500
CGAGGGTGC	GGTGC CGCGC	TATCGGCAGC	ACGTGAGCTC	CACGACGAAC	TCATCCCAGT	1560
GCTGGGTTCC	GC GGAGTTCG	GCATCGCGT	GTCGGCCGGA	AGGGCCATCG	CCGCCACAT	1620
CGCGCTCAA	GCCCCGTTCG	AGTACACCCT	CATCGGCAGC	CCGGTCAACG	AGGCCGCCCCG	1680
GCTCACCGAA	CTGGCCAAAG	TCGAGGATGG	CCACGTTCTG	CGCTCGCGA	TCGCGGTCA	1740
TGGGCCCTG	GACGCCGAAG	CATTGTGTTG	GGATGTTGGC	GAGGTGGTTG	AGCTCCGCGG	1800
ACGTGCTGCA	CCCACCCAAAC	TAGCCAGGCC	AATGAATNTG	GCNGCACCCG	AAGAGGTTTC	1860
CAGCGAAGTA	CGCGGCTAGT	CGCGCTTGGC	TGCNTTCTTC	GCCGGCACCT	TCCGGGCAGC	1920
TTTCCTGGCT	GGCCGTTTG	CCGGACCCCG	GGCTCGCGA	TCGGCCAACA	GCTCGCGGC	1980
GCGCTCGTC	GTTATGGAAG	CCACGTNGTC	GCCCTTACGC	AGGCTGGCAT	TGGTCTCACC	2040
GTCGGTGACG	TACGGCCCGA	ATCGGCCGTC	CTTGATGACC	ATTGGCTTGC	CAGACGCCGG	2100
ATNTGNTCCC	AGCTCGCGA	GC GGCGGAGC	CGAAGCGCTT	TGCCGGCCAC	GACNTTCGG	2160
CTCTGNGTAG	ATNTTCAGGG	CTTCGTCGAG	CGNGATGGTG	AATATATGGT	CTTCGGTGAC	2220
CAGTGATCGA	GAATCGTTGC	CGCGCTTGTAG	ATACGGTCNG	TAGCGCCCGT	TCTGCGCGGT	2280
GATNTC						2286

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1136 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGCATCTTC	CCCGACCGCG	CCTCGATCAT	CCGCCTCGTC	GGAGCCGTC	TCGCCGAACA	60
ACACGACGAA	TGGATCGAAG	GACGGCGCTA	CCTGGGCCTC	GAGGTCTCA	CCCGAGCCCG	120
AGCAGCACTG	ACCAGCACCG	AAGAACGCC	AAGCAGCAA	CCACCAACAC	CCCAGCACTG	180
ACCACCTAGA	CTGCCACCCG	AAGGATCACG	CGAGGAACCT	TCACTCGTAC	ACCACGTCCC	240
TGGCCTTGGC	CTGGTGTCA	GCCCAGCTGG	AGCCGACGCC	GCTGTCGTT	TGCCCCATGT	300
TGTTGCCGGC	AGCCTGCACC	TTCTGCCCGT	GGGCCTTGGC	CTGCTCGTAG	ATCACCTGGA	360
AGTTACGGCC	CAACTGGGTA	ATGAACCCCT	GGCAGGCCGC	CGAACCGCG	CCGCCCCAAA	420
AGTCACTCGC	GGTCAACACA	TCACGAATGA	TGGCCTGATG	CTCGGCCCTCC	AGCAACCCGG	480
CCTGAGCGCG	GATCATGGCG	CCGTGAGCGT	CGACATCACC	GAAC TGATAG	TTGATGGTCA	540
TCGAACCTGT	TCTCCTTCG	TTG TAAAGT	ATTGTGCTGC	AGCGGCTGAC	GTTAGCTGCT	600

GAGGATCTGC	TGGGAGGCCT	GCTCTTGCT	CGTGCCGAAT	TCGGCAGGAG	AGGCCGCCTT	660
CGAAGAAATC	CTTGAGAAT	TCGCCAAGGC	CGTCGACCCA	GCATGGGTC	AGCTGCCAG	720
CCCGGCCGGC	TGGCAACCCT	TCCCGCTCGA	GAAAGACCTG	GAGGAATACC	AGTACAAAC	780
GACCTCCCAG	ACGTCCGAGA	CGGTGACGGC	GGTCACGTC	CCGCTCTCC	TGCTGGCGGG	840
CCACGCTTGT	CAGACGTGTG	GGTTTACAAC	GGGCGGGCGT	ACGACCTGAG	TGAGTGGATT	900
TCCAAGCATC	CCGGCGGC	CTTNTTCATT	GGGCGGACCA	AGAACCGCGA	CATCACCGCA	960
ATCGTCAAGT	CCTACCACATCG	TGATCCGGCG	ATTGTCGAGC	GAATCTGCA	GC GGAGGTAC	1020
CGGTTGGGCC	GCGACGCAAC	CCCTAGGGAC	ATCCACCCCCA	AGCACAATGC	ACCGGCATTT	1080
CTGTTCAAAG	ACGACTTCAA	CAGCTGGCGG	GACACCCCCA	AGTATCGATT	NGACGA	1136

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 967 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGAGCGCCAA	CCCTACCGTC	GGTTCGTCAC	ACGGACCGCA	TGGCCTGCTC	CGCGGACTGC	60
CGCTAGGGTC	GCGGATCACT	CGGCGTAGCG	GCGCCTTGC	CCACCGATAT	GGGTTCCGTC	120
ACAGTGTGGT	TGCCCGCCCG	CCATCGGCCG	GATAACGCCA	TGACCTCAGC	TCGGCAGAAA	180
TGACAATGCT	CCCAAAGGCG	TGAGCACCCG	AAGACAACTA	AGCAGGAGAT	CGCATGCCGT	240
TTGTGACTAC	CCAACCAGAA	GCACTGGCGG	CGGGCGGCCGG	CAGTCTGCAG	GGAATCGGCT	300
CCGCATTGAA	CGCCCAGAAT	GC GGCTGCGG	CGACTCCCAC	GACGGGGGTG	GTCCGGCGGC	360
CGCCGATGAA	NTGTCGGCGC	TGACGGCGGC	TCAGTTCGCG	GCACACGCC	AGATCTATCA	420
GGCCGTCAGC	GCCCAGGCCG	CGGCGATTCA	CGAGATGTT	GTCAACACTC	TACAGATGAG	480
CTCAGGGTCG	TATGCTGCTA	CCGAGGCCGC	CAACGCCGCC	CGGGCCGGNT	AGAGGAGTCA	540
CTGCGATGGA	TTTTGGGGCG	TTGCCGCCGG	AGGTCAATT	GGTGCAGGATG	TATGCCGTTC	600
CTGGCTCGGC	ACCAATGGTC	GCTGCGCGT	CGGCCTGGAA	CGGGTTGCC	CGGGAGCTGA	660
TTTCGGCGGC	CACCGGTTAT	GAGACGGTGA	TCACTCAGCT	CAGCAGTGA	GGGTGGCTAG	720
GTCCGGCGTC	AGCGGCGATG	GCCGAGGCAG	TTGCCGCCGA	TGTGGCGTGG	ATGAGTGCCG	780
CTGCGGCCGA	AGCCGAGCAG	CGGGCCACAC	AGGCCAGGGC	CGCCGCGGCC	GCTTTGAGG	840
CGGCGTTG	CGCGACGGTG	CCTCCGCCGT	TGATCGCGGC	CAACCGGGCT	TCGTTGATGC	900
AGCTGATCTC	GACGAATGTC	TTTGGTCAGA	ACACCTCGGC	GATCGCGGCC	GCCGAAGCTC	960
AGTACGG						967

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 585 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TGGATTCCGA	TAGCGGTTTC	GGCCCCTCGA	CGGGCGACCA	CGGCGCGCAG	GCCCTCGAAC	60
GGGGGGCCGG	GACGCTGGGA	TTCGCCGGGA	CCGCAACCAA	AGAACGCCGG	GTCCGGCGG	120
TCGGGCTGAC	CGCACTGGCC	GGTGATGAGT	TCGGCAACGG	CCCCCGGATG	CCGATGGTGC	180
CGGGGACCTG	GGAGCAGGGC	AGCAACGAGC	CCGAGGCCGCC	CGACGGATCG	GGGAGAGGGG	240
GAGGCGACGG	CTTACCGCAC	GACAGCAAGT	AACCGAATT	CGAATCACGT	GGACCCGTAC	300
GGGTCGAAAG	GAGAGATGTT	ATGAGCCTT	TGGATGCTCA	TATCCCACAG	TTGGTGGCCT	360
CCCAGTCGGC	GTTTGCCGCC	AAGGCCGGGC	TGATGCCGCA	CACGATCGGT	CAGGCCGAGC	420
AGGCGGCCGAT	GTCGGCTCAG	CGCTTTCACC	AGGGGGAGTC	GTCGGCGGCC	TTTCAGGCCG	480
CCCATGCCCG	GTGGTGGCG	CGGGCCGCCA	AAGTCAACAC	CTTGGTGGAT	GTCGCCGAGG	540
CGAATCTGGG	TGAGGCCGCC	GGTACCTATG	TGGCCGCCGA	TGCTG		585

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ala	Leu	Val	Thr	Thr	Asn	Phe	Phe	Gly	Val	Asn	Thr	Ile	Pro	Ile	Ala
1															15
Leu	Asn	Glu	Ala	Asp	Tyr	Leu	Arg	Met	Trp	Ile	Gln	Ala	Ala	Thr	Val
	20														30
Met	Ser	His	Tyr	Gln	Ala	Val	Ala	His	Glu	Ile	Trp	Cys	Leu	His	Glu
		35													45
Xaa	Ala	Ser	Ser	Gly	Lys	Pro	Trp	Ala	Ser	Ile	Thr	Thr	Gly	Ala	Pro
		50													60
Gly	Ser	Pro	Ala	Ser	Thr	Thr	Arg	Ser	Arg	Thr	Pro	Leu	Val	Ser	Thr
	65				70										80
Asn	Arg	Xaa	Val	Xaa	Ala	Pro	Ile	Val	Ser	Pro	Asn	His	Thr	Gly	His
		85													95
Arg	Pro	Glu	Lys	Gly	Leu	Gly	Ser	Xaa	Gln	Arg	Arg	Leu	Ser	Arg	Val
		100													110
Leu	Pro	Arg	Ile	Ile	Asp	Arg	Pro	Ala	Gly	Pro	Xaa	Gly	Pro	Pro	Leu
		115													125
Thr	Ser	Gly	Ser	His	Phe	Leu	Cys	Ser	Trp	His	Gly	Tyr	Ser	Ser	Gln
		130													140

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 352 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

His	Ala	Leu	Ala	Ala	Gln	Tyr	Thr	Glu	Ile	Ala	Thr	Glu	Leu	Ala	Ser
1															15
Val	Leu	Ala	Ala	Val	Gln	Ala	Ser	Ser	Trp	Gln	Gly	Pro	Ser	Ala	Asp
		20													30
Arg	Phe	Val	Val	Ala	His	Gln	Pro	Phe	Arg	Tyr	Trp	Leu	Thr	His	Ala
		35													45
Ala	Thr	Val	Ala	Thr	Ala	Ala	Ala	Ala	His	Xaa	Thr	Ala	Ala	Ala	
		50													60
Gly	Tyr	Thr	Ser	Ala	Leu	Gly	Gly	Met	Pro	Thr	Leu	Ala	Glu	Leu	Ala
	65				70										80
Ala	Asn	His	Ala	Met	His	Gly	Ala	Leu	Val	Thr	Thr	Asn	Phe	Phe	Gly
															95
Val	Asn	Thr	Ile	Pro	Ile	Ala	Leu	Asn	Glu	Ala	Asp	Tyr	Leu	Arg	Met
		100													110
Trp	Ile	Gln	Ala	Ala	Thr	Val	Met	Ser	His	Tyr	Gln	Ala	Val	Ala	His
		115													125
Glu	Ser	Val	Ala	Ala	Thr	Pro	Ser	Thr	Pro	Pro	Ala	Pro	Gln	Ile	Val
		130													140

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

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 141 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Asp Phe Gly Ala Leu Pro Pro Glu Val Asn Ser Val Arg Met Tyr  
 1 5 10 15  
 Ala Val Pro Gly Ser Ala Pro Met Val Ala Ala Ser Ala Trp Asn  
     20 25 30  
 Gly Leu Ala Ala Glu Leu Ser Ser Ala Ala Thr Gly Tyr Glu Thr Val  
     35 40 45  
 Ile Thr Gln Leu Ser Ser Glu Gly Trp Leu Gly Pro Ala Ser Ala Ala  
     50 55 60  
 Met Ala Glu Ala Val Ala Pro Tyr Val Ala Trp Met Ser Ala Ala Ala  
 65 70 75 80  
 Ala Gln Ala Glu Gln Ala Ala Thr Gln Ala Arg Ala Ala Ala Ala  
     85 90 95  
 Phe Glu Ala Ala Phe Ala Ala Thr Val Pro Pro Pro Leu Ile Ala Ala  
     100 105 110  
 Asn Arg Ala Ser Leu Met Gln Leu Ile Ser Thr Asn Val Phe Gly Gln  
     115 120 125  
 Asn Thr Ser Ala Ile Ala Ala Glu Ala Gln Tyr Gly  
     130 135 140

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

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

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 67 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met	Thr	Ile	Asn	Tyr	Gln	Phe	Gly	Asp	Val	Asp	Ala	His	Gly	Ala	Met
1						5			10				15		
Ile	Arg	Ala	Gln	Ala	Ala	Ser	Leu	Glu	Ala	Glu	His	Gln	Ala	Ile	Val
						20		25				30			
Arg	Asp	Val	Leu	Ala	Ala	Gly	Asp	Phe	Trp	Gly	Gly	Ala	Gly	Ser	Val
						35		40			45				
Ala	Cys	Gln	Glu	Phe	Ile	Thr	Gln	Leu	Gly	Arg	Asn	Phe	Gln	Val	Ile
					50		55				60				
Tyr	Glu	Gln													
		65													

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

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

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:  
 Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met  
 1 5 10 15  
 Ile Arg Ala Gln Ala Ala Ser Leu Glu Ala Glu His Gln Ala Ile Val  
 20 25 30  
 Arg Asp Val Leu Ala Ala Gly Asp Phe Trp Gly Gly Ala Gly Ser Val  
 35 40 45  
 Ala Cys Gln Glu Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile  
 50 55 60  
 Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn  
 65 70 75 80  
 Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala  
 85 90

## (2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:  
 Asn Met Leu His Gly Val Arg Asp Gly Leu Val Arg Asp Ala Asn Asn  
 1 5 10 15  
 Tyr Glu Gln Gln Glu Gln Ala Ser Gln Gln Ile Leu Ser Ser  
 20 25 30

## (2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 94 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:  
 Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met  
 1 5 10 15  
 Ile Arg Ala Gln Ala Gly Leu Leu Glu Ala Glu His Gln Ala Ile Ile  
 20 25 30  
 Arg Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala Gly Ser Ala  
 35 40 45  
 Ala Cys Gln Gly Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile  
 50 55 60  
 Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn  
 65 70 75 80  
 Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala  
 85 90

## (2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 69 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Ala	Arg	Arg	Met	Trp	Ala	Ser	Ala	Gln	Asn	Ile	Ser	Gly	Ala	Gly	Trp
1			5				10						15		
Ser	Gly	Met	Ala	Glu	Ala	Thr	Ser	Leu	Asp	Thr	Met	Ala	Gln	Met	Asn
			20			25						30			
Gln	Ala	Phe	Arg	Asn	Ile	Val	Asn	Met	Leu	His	Gly	Val	Arg	Asp	Gly
			35			40				45					
Leu	Val	Arg	Asp	Ala	Asn	Asn	Tyr	Glu	Gln	Gln	Glu	Gln	Ala	Ser	Gln
			50			55				60					
Gln	Ile	Leu	Ser	Ser											
			65												

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met	Thr	Ile	Asn	Tyr	Gln	Phe	Gly	Asp	Val	Asp	Ala	His	Gly	Ala	Met
1			5				10					15			
Ile	Arg	Ala	Gln	Ala	Gly	Leu	Leu	Glu	Ala	Glu	His	Gln	Ala	Ile	Ile
			20			25				30					
Arg	Asp	Val	Leu	Thr	Ala	Ser	Asp	Phe	Trp	Gly	Gly	Ala	Gly	Ser	Ala
			35			40				45					
Ala	Cys	Gln	Gly	Phe	Ile	Thr	Gln	Leu	Gly	Arg	Asn	Phe	Gln	Val	Ile
			50			55			60						
Tyr	Glu	Gln	Ala	Asn	Thr	His	Gly	Gln	Lys	Val	Gln	Ala	Ala	Gly	Asn
			65			70			75			80			
Asn	Met	Ala	Gln	Thr	Asp	Ser	Ala	Val	Xaa	Ser	Ser	Trp	Ala		
			85			90									

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Gly	Met	Ala	Glu	Ala	Thr	Ser	Xaa	Asp	Thr	Met	Thr	Gln	Met	Asn	Gln
1			5			10						15			
Ala	Phe	Arg	Asn	Ile	Val	Asn	Met	Leu	His	Gly	Val	Arg	Asp	Gly	Leu
			20			25			30						
Val	Arg	Asp	Ala	Asn	Xaa	Tyr	Glu	Gln	Glu	Gln	Ala	Ser	Gln	Gln	
			35			40			45						
Ile	Leu	Ser	Ser												
			50												

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met	Thr	Ile	Asn	Tyr	Gln	Phe	Gly	Asp	Val	Asp	Ala	His	Gly	Ala	Met
1															15
Ile	Arg	Ala	Gln	Ala	Gly	Ser	Leu	Glu	Ala	Glu	His	Gln	Ala	Ile	Ile
	20						25				30				
Ser	Asp	Val	Leu	Thr	Ala	Ser	Asp	Phe	Trp	Gly	Gly	Ala	Gly	Ser	Ala
	35						40				45				
Ala	Cys	Gln	Gly	Phe	Ile	Thr	Gln	Leu	Gly	Arg	Asn	Phe	Gln	Val	Xaa
	50						55				60				
Tyr	Glu	Gln	Ala	Asn	Ala	His	Gly	Gln	Lys	Val	Gln	Ala	Ala	Gly	Asn
	65						70				75				80
Asn	Met	Ala	Gln	Thr	Asp	Ser	Ala	Val	Gly	Ser	Ser	Trp	Ala		
	85						90								

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 98 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met	Thr	Ser	Arg	Phe	Met	Thr	Asp	Pro	His	Ala	Met	Arg	Asp	Met	Ala
1															15
Gly	Arg	Phe	Glu	Val	His	Ala	Gln	Thr	Val	Glu	Asp	Glu	Ala	Arg	Arg
	20						25								30
Met	Trp	Ala	Ser	Ala	Gln	Asn	Ile	Ser	Gly	Ala	Gly	Trp	Ser	Gly	Met
	35						40								45
Ala	Glu	Ala	Thr	Ser	Leu	Asp	Thr	Met	Ala	Gln	Met	Asn	Gln	Ala	Phe
	50						55								60
Arg	Asn	Ile	Val	Asn	Met	Leu	His	Gly	Val	Arg	Asp	Gly	Leu	Val	Arg
	65						70								80
Asp	Ala	Asn	Asn	Tyr	Glu	Gln	Gln	Glu	Gln	Ala	Ser	Gln	Gln	Ile	Leu
	85						90								95
Ser	Ser														

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met	Thr	Ile	Asn	Tyr	Gln	Phe	Gly	Asp	Val	Asp	Ala	His	Gly	Ala	Met
1															15
Ile	Arg	Ala	Xaa	Ala	Gly	Leu	Leu	Glu	Ala	Glu	His	Gln	Ala	Ile	Ile

20	25	30
Ser Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala Gly Ser Ala		
35	40	45
Ala Cys Gln Gly Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile		
50	55	60
Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn		
65	70	75
Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala		
85	90	

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Arg Phe Glu Val His Ala Gln Thr Val Glu Asp Glu Ala Arg Arg Met			
1	5	10	15
Trp Ala Ser Ala Gln Asn Ile Ser Gly Ala Gly Trp Ser Gly Met Ala			
20	25	30	
Xaa Ala Thr Ser Leu Asp Thr Met Ala Gln Met Asn Gln Ala Phe Arg			
35	40	45	
Asn Ile Val Asn Met Leu His Gly Val Arg Asp Gly Leu Val Arg Asp			
50	55	60	
Ala Asn Asn Tyr Glu Gln Gln Glu Gln Ala Ser Gln Gln Ile Leu Ser			
65	70	75	80
Ser			

## (2) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met			
1	5	10	15
Ile Arg Ala Leu Ala Gly Leu Leu Glu Ala Glu His Gln Ala Ile Ile			
20	25	30	
Ser Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala Gly Ser Ala			
35	40	45	
Ala Cys Gln Gly Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile			
50	55	60	
Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn			
65	70	75	80
Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala			
85	90		

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Gln Glu Gln Ala Ser Gln Gln Ile Leu Ser Ser

1 5 10

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met

1 5 10 15

Ile Arg Ala Gln Ala Gly Leu Leu Glu Ala Glu His Gln Ala Ile Ile

20 25 30

Arg Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala Gly Ser Ala

35 40 45

Ala Cys Gln Gly Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile

50 55 60

Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn

65 70 75 80

Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala

85 90

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 99 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met Ser Phe Val Thr Thr Gln Pro Glu Ala Leu Ala Ala Ala Ala Ala

1 5 10 15

Asn Leu Gln Gly Ile Gly Thr Thr Met Asn Ala Gln Asn Ala Ala Ala

20 25 30

Ala Ala Pro Thr Thr Gly Val Val Pro Ala Ala Ala Asp Glu Val Ser

35 40 45

Ala Leu Thr Ala Ala Gln Phe Ala Ala His Ala Gln Met Tyr Gln Thr

50 55 60

Val Ser Ala Gln Ala Ala Ala Ile His Glu Met Phe Val Asn Thr Leu

65 70 75 80

Val Ala Ser Ser Gly Ser Tyr Ala Ala Thr Glu Ala Ala Asn Ala Ala

85 90 95

Ala Ala Gly

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 99 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met	Ser	Phe	Val	Thr	Thr	Gln	Pro	Glu	Ala	Leu	Ala	Ala	Ala	Ala	Ala
1				5				10						15	
Asn	Leu	Gln	Gly	Ile	Gly	Thr	Thr	Met	Asn	Ala	Gln	Asn	Ala	Ala	Ala
				20				25					30		
Ala	Ala	Pro	Thr	Thr	Gly	Val	Val	Pro	Ala	Ala	Asp	Glu	Val	Ser	
		35				40					45				
Ala	Leu	Thr	Ala	Ala	Gln	Phe	Ala	Ala	His	Ala	Gln	Met	Tyr	Gln	Thr
					50			55			60				
Val	Ser	Ala	Gln	Ala	Ala	Ile	His	Glu	Met	Phe	Val	Asn	Thr	Leu	
65					70				75			80			
Val	Ala	Ser	Ser	Gly	Ser	Tyr	Ala	Ala	Thr	Glu	Ala	Ala	Asn	Ala	Ala
				85				90				95			
Ala	Ala	Gly													

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Asp	Pro	His	Ala	Met	Arg	Asp	Met	Ala	Gly	Arg	Phe	Glu	Val	His
1				5				10				15		

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Arg	Asp	Met	Ala	Gly	Arg	Phe	Glu	Val	His	Ala	Gln	Thr	Val	Glu
1				5				10				15		

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Arg	Phe	Glu	Val	His	Ala	Gln	Thr	Val	Glu	Asp	Glu	Ala	Arg	Arg
1				5				10				15		

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Ala	Gln	Thr	Val	Glu	Asp	Glu	Ala	Arg	Arg	Met	Trp	Ala	Ser	Ala
1				5				10						15

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Asp	Glu	Ala	Arg	Arg	Met	Trp	Ala	Ser	Ala	Gln	Asn	Ile	Ser	Gly
1				5				10						15

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Met	Trp	Ala	Ser	Ala	Gln	Asn	Ile	Ser	Gly	Ala	Gly	Trp	Ser	Gly
1				5				10						15

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Gln	Asn	Ile	Ser	Gly	Ala	Gly	Trp	Ser	Gly	Met	Ala	Glu	Ala	Thr
1				5				10						15

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Ala	Gly	Trp	Ser	Gly	Met	Ala	Glu	Ala	Thr	Ser	Leu	Asp	Thr	Met	Thr
1				5					10					15	

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Met	Ala	Glu	Ala
1	5		
Thr	Ser	Leu	Asp
Thr			Met
			Ala
			Gln
			Met
			Asn
			Gln

## (2) INFORMATION FOR SEQ ID NO:43:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Ser	Leu	Asp	Thr
1	5		
Met			Met
Ala			Asn
Gln			Gln
			Ala
			Phe
			Arg
			Asn
			Ile

## (2) INFORMATION FOR SEQ ID NO:44:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Ala	Gln	Met	Asn
1	5		
Gln			Gln
			Ala
			Phe
			Arg
			Asn
			Ile
			Val
			Asn
			Met
			Leu
			His

## (2) INFORMATION FOR SEQ ID NO:45:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Ala	Phe	Arg	Asn
1	5		
			Ile
			Val
			Asn
			Met
			Leu
			His
			Gly
			Val
			Arg
			Asp
			Gly

## (2) INFORMATION FOR SEQ ID NO:46:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Val	Asn	Met	Leu
1	5		
			His
			Gly
			Val
			Arg
			Asp
			Gly
			Leu
			Val
			Arg
			Asp
			Ala

## (2) INFORMATION FOR SEQ ID NO:47:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Gly	Val	Arg	Asp
1	5		
			Gly
			Leu
			Val
			Arg
			Asp
			Ala
			Asn
			Asn
			Tyr
			Glu
			Gln

## (2) INFORMATION FOR SEQ ID NO:48:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Leu	Val	Arg	Asp
Ala	Asn	Asn	Tyr
Glu	Gln	Gln	Glu
Gln	Ala	Ser	

1	5	10	15
---	---	----	----

## (2) INFORMATION FOR SEQ ID NO:49:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Asn	Asn	Tyr	Glu
Gln	Gln	Glu	Gln
Ala	Ser	Gln	Gln
Ile	Leu	Ser	Ser

1	5	10	15
---	---	----	----

## (2) INFORMATION FOR SEQ ID NO:50:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Met	Ala	Ser	Arg
Phe	Met	Thr	Asp
Pro	His	Ala	Met
Arg	Asp	Met	Ala

1	5	10	15
---	---	----	----

Gly

## (2) INFORMATION FOR SEQ ID NO:51:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Met	Thr	Ile	Asn
Tyr	Gln	Phe	Gly
Asp	Asp	Val	Asp
Ala	His	Gly	Ala

1	5	10	15
---	---	----	----

## (2) INFORMATION FOR SEQ ID NO:52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Gln	Phe	Gly	Asp
Asp	Val	Asp	Ala
Ala	His	Gly	Ala
Met	Ile	Arg	Ala

1	5	10	15
---	---	----	----

## (2) INFORMATION FOR SEQ ID NO:53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Asp	Ala	His	Gly	Ala	Met	Ile	Arg	Ala	Gln	Ala	Ala	Ser	Leu	Glu
1														
														15

## (2) INFORMATION FOR SEQ ID NO:54:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Met	Ile	Arg	Ala	Gln	Ala	Ala	Ser	Leu	Glu	Ala	Glu	His	Gln	Ala
1														
														15

## (2) INFORMATION FOR SEQ ID NO:55:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Ala	Ala	Ser	Leu	Glu	Ala	Glu	His	Gln	Ala	Ile	Val	Arg	Asp	Val
1														
														15

## (2) INFORMATION FOR SEQ ID NO:56:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Ala	Glu	His	Gln	Ala	Ile	Val	Arg	Asp	Val	Leu	Ala	Ala	Gly	Asp
1														
														15

## (2) INFORMATION FOR SEQ ID NO:57:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Ile	Val	Arg	Asp	Val	Leu	Ala	Ala	Gly	Asp	Phe	Trp	Gly	Gly	Ala
1														
														15

## (2) INFORMATION FOR SEQ ID NO:58:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Leu	Ala	Ala	Gly	Asp	Phe	Trp	Gly	Gly	Ala	Gly	Ser	Val	Ala	Cys	Gln
1	5														

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Phe	Trp	Gly	Gly	Ala	Gly	Ser	Val	Ala	Cys	Gln	Glu	Phe	Ile	Thr
1	5													

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Gly	Ser	Val	Ala	Cys	Gln	Glu	Phe	Ile	Thr	Gln	Leu	Gly	Arg	Asn
1	5													

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

Gln	Glu	Phe	Ile	Thr	Gln	Leu	Gly	Arg	Asn	Phe	Gln	Val	Ile	Tyr	Glu
1	5														

Gln Ala

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Arg	Asn	Phe	Gln	Val	Ile	Tyr	Glu	Gln	Ala	Asn	Ala	His	Gly	Gln
1	5													

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid

- (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Tyr | Glu | Gln | Ala | Asn | Ala | His | Gly | Gln | Lys | Val | Gln | Ala | Ala |
| 1   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|     | 5   |     |     |     |     |     |     |     |     | 10  |     |     |     | 15  |
- (2) INFORMATION FOR SEQ ID NO:64:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Ala | His | Gly | Gln | Lys | Val | Gln | Ala | Ala | Gly | Asn | Asn | Met | Ala |
| 1   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|     | 5   |     |     |     |     |     |     |     |     | 10  |     |     |     | 15  |
- (2) INFORMATION FOR SEQ ID NO:65:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Val | Gln | Ala | Ala | Gly | Asn | Asn | Met | Ala | Gln | Thr | Asp | Ser | Ala |
| 1   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|     | 5   |     |     |     |     |     |     |     |     | 10  |     |     |     | 15  |
- (2) INFORMATION FOR SEQ ID NO:66:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 16 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Asn | Asn | Met | Ala | Gln | Thr | Asp | Ser | Ala | Val | Gly | Ser | Ser | Trp | Ala |
| 1   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|     | 5   |     |     |     |     |     |     |     |     | 10  |     |     |     |     | 15  |
- (2) INFORMATION FOR SEQ ID NO:67:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Ala | His | Gly | Ala | Met | Ile | Arg | Ala | Leu | Ala | Gly | Leu | Leu | Glu |
| 1   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|     | 5   |     |     |     |     |     |     |     |     | 10  |     |     |     | 15  |
- (2) INFORMATION FOR SEQ ID NO:68:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:  
 Asp Ala His Gly Ala Met Ile Arg Ala Gln Ala Gly Leu Leu Glu  
 1               5                           10                           15
- (2) INFORMATION FOR SEQ ID NO:69:  
 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 15 amino acids  
     (B) TYPE: amino acid  
     (C) STRANDEDNESS: single  
     (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:  
 Met Ile Arg Ala Leu Ala Gly Leu Leu Glu Ala Glu His Gln Ala  
 1               5                           10                           15
- (2) INFORMATION FOR SEQ ID NO:70:  
 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 15 amino acids  
     (B) TYPE: amino acid  
     (C) STRANDEDNESS: single  
     (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:  
 Met Ile Arg Ala Gln Ala Gly Leu Leu Glu Ala Glu His Gln Ala  
 1               5                           10                           15
- (2) INFORMATION FOR SEQ ID NO:71:  
 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 15 amino acids  
     (B) TYPE: amino acid  
     (C) STRANDEDNESS: single  
     (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:  
 Ala Gly Leu Leu Glu Ala Glu His Gln Ala Ile Ile Ser Asp Val  
 1               5                           10                           15
- (2) INFORMATION FOR SEQ ID NO:72:  
 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 15 amino acids  
     (B) TYPE: amino acid  
     (C) STRANDEDNESS: single  
     (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:  
 Ala Gly Leu Leu Glu Ala Glu His Gln Ala Ile Ile Arg Asp Val  
 1               5                           10                           15
- (2) INFORMATION FOR SEQ ID NO:73:  
 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 15 amino acids  
     (B) TYPE: amino acid  
     (C) STRANDEDNESS: single  
     (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Ala Glu His Gln Ala Ile Ile Ser Asp Val Leu Thr Ala Ser Asp  
 1                   5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:74:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

Ala Glu His Gln Ala Ile Ile Arg Asp Val Leu Thr Ala Ser Asp  
 1                   5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:75:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Ile Ile Ser Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala  
 1                   5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:76:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Ile Ile Arg Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala  
 1                   5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:77:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala Gly Ser Ala Ala Cys Gln  
 1                   5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:78:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

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

## (2) INFORMATION FOR SEQ ID NO:79:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Gly	Ser	Ala	Ala	Cys	Gln	Gly	Phe	Ile	Thr	Gln	Leu	Gly	Arg	Asn
1														
														15

## (2) INFORMATION FOR SEQ ID NO:80:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Gln	Gly	Phe	Ile	Thr	Gln	Leu	Gly	Arg	Asn	Phe	Gln	Val	Ile	Tyr
1														
														15

## (2) INFORMATION FOR SEQ ID NO:81:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Val	Thr	Thr	Asn	Phe	Phe	Gly	Val	Asn	Thr	Ile	Pro	Ile	Ala	Leu	Asn
1															
															15
Glu	Ala	Asp	Tyr	Leu	Arg	Met	Trp	Ile							
20															25

## (2) INFORMATION FOR SEQ ID NO:82:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Asn	Glu	Ala	Asp	Tyr	Leu	Arg	Met	Trp	Ile	Gln	Ala	Ala	Thr	Val	Met
1															
															15
Ser	His	Tyr	Gln	Ala	Val	Ala	His	Glu							
20															25

## (2) INFORMATION FOR SEQ ID NO:83:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 967 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

TGAGCGCCAA	CCCTACCGTC	GGTTCGTCAC	ACGGACCGCA	TGGCCTGCTC	CGCGGACTGC	60
CGCTAGGGTC	CGGGATCACT	CGGCGTAGCG	GCGCCTTGCG	CCACCGATAT	GGGTTCCGTC	120
ACAGTGTGGT	TGCCCGCCCG	CCATCGGCCG	GATAACGCCA	TGACCTCAGC	TCGGCAGAAA	180
TGACAATGCT	CCCAAAGGCG	TGAGCACCCG	AAGACAACTA	AGCAGGAGAT	CGCATGCCGT	240
TTGTGACTAC	CCAACCAGAA	GAECTGGCGG	CGGCGGCCGG	CAGTCTGCAG	GGAATCGGCT	300
CCGCATTGAA	CGCCCAGAAT	GCGGCTGCCG	CGACTCCCAC	GACGGGGTG	GTCCGGCGGC	360
CGCCGATGAA	NTGTCGGCGC	TGACGGCGGC	TCAGTTCGCG	GCACACGCC	AGATCTATCA	420
GGCCGTCAGC	GCCCCAGGCCG	CGCGATTCA	CGAGATGTTC	GTCAACACTC	TACAGATGAG	480
CTCAGGGTCG	TATGCTGCTA	CCGAGGCCG	CAACGCCGCC	GCGGCCGGNT	AGAGGAGTCA	540
CTGCGATGGA	TTTGGGGCG	TTGCCGCCG	AGGTCAATT	GGTGCGGATG	TATGCCGTT	600
CTGGCTCGGC	ACCAATGGTC	GCTGCGCGT	CGGCCTGGAA	CGGGTTGGCC	GCGGAGCTGA	660
GTTCGGCGGC	CACCGTTAT	GAGACGGTGA	TCACTCAGCT	CAGCAGTGAG	GGGTGGCTAG	720
GTCCGGCGTC	AGCGGCGATG	GCCGAGGCCAG	TTGCGCCGTA	TGTGGCGTGG	ATGAGTGCCG	780
CTGCGGCCA	AGCCGAGCAG	GCGGCCACAC	AGGCCAGGGC	CGCCGCCGCC	GCTTTGAGG	840
CGGCCTTGC	CGCGACGGTG	CCTCCGCCGT	TGATCGCGGC	CAACCGGGCT	TCGTTGATGC	900
AGCTGATCTC	GACGAATGTC	TTGGTCAGA	ACACCTCGGC	GATCGCGGCC	GCCGAAGCTC	960
AGTACGG						967

## (2) INFORMATION FOR SEQ ID NO:84:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Met	Ser	Phe	Val	Thr	Thr	Gln	Pro	Glu	Ala	Leu	Ala	Ala	Ala	Ala
1														
														15

## (2) INFORMATION FOR SEQ ID NO:85:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Thr	Gln	Pro	Glu	Ala	Leu	Ala	Ala	Ala	Ala	Asn	Leu	Gln	Gly
1													
													15

## (2) INFORMATION FOR SEQ ID NO:86:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Leu	Ala	Ala	Ala	Ala	Ala	Asn	Leu	Gln	Gly	Ile	Gly	Thr	Thr	Met
1														
														15

## (2) INFORMATION FOR SEQ ID NO:87:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

Ala Asn Leu Gln Gly Ile Gly Thr Thr Met Asn Ala Gln Asn Ala  
 1               5               10               15

## (2) INFORMATION FOR SEQ ID NO:88:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Ile Gly Thr Thr Met Asn Ala Gln Asn Ala Ala Ala Ala Ala Pro  
 1               5               10               15

## (2) INFORMATION FOR SEQ ID NO:89:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Asn Ala Gln Asn Ala Ala Ala Ala Pro Thr Thr Gly Val Val  
 1               5               10               15

## (2) INFORMATION FOR SEQ ID NO:90:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

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

## (2) INFORMATION FOR SEQ ID NO:91:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

Thr Thr Gly Val Val Pro Ala Ala Ala Asp Glu Val Ser Ala Leu  
 1               5               10               15

## (2) INFORMATION FOR SEQ ID NO:92:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

```

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

```

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

```

Glu Val Ser Ala Leu Thr Ala Ala Gln Phe Ala Ala His Ala Gln
   1           5           10          15

```

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

Thr Ala Ala Gln Phe Ala Ala His Ala Gln Met Tyr Gln Thr Val  
 1                   5                   10                   15

(2) INFORMATION FOR SEO ID NO: 95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

Ala Ala His Ala Gln Met Tyr Gln Thr Val Ser Ala Gln Ala Ala  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

Met Tyr Gln Thr Val Ser Ala Gln Ala Ala Ala Ile His Glu Met Phe  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

Ser Ala Glu Ala Ala Ala Ile His Glu Met Phe Val Asp Thr Leu

1 5 10 15

## (2) INFORMATION FOR SEQ ID NO:98:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

Ala Ile His Glu Met Phe Val Asn Thr Leu Val Ala Ser Ser Gly  
1 5 10 15

## (2) INFORMATION FOR SEQ ID NO:99:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

Phe Val Asn Thr Leu Val Ala Ser Ser Gly Ser Tyr Ala Ala Thr  
1 5 10 15

## (2) INFORMATION FOR SEQ ID NO:100:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

Val Ala Ser Ser Gly Ser Tyr Ala Ala Thr Glu Ala Ala Asn Ala  
1 5 10 15

## (2) INFORMATION FOR SEQ ID NO:101:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

Ser Tyr Ala Ala Thr Glu Ala Ala Asn Ala Ala Ala Gly  
1 5 10

## (2) INFORMATION FOR SEQ ID NO:102:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1784 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

ATTCGTTCT	GCCGCAGCTA	AATCCCGGGG	ACATCGTCGC	CGGCCAGTAC	GAGGTCAAAG	60
GCTGCATCGC	GCACGGCGGA	CTGGGCTGGA	TCTACCTCGC	TCTCGACCAG	AATGTCAACG	120
GCCGTCCGGT	GGTGCTCAAG	GGCCTGGTGC	ATTCCGGTGA	TGCCGAAGCG	CAGGCAATGG	180

CGATGGCCGA	ACGCCAGTTC	CTGGCCGAGG	TGGTGCACCC	GTCGATCGT	CAGATCTTCA	240
ACTTTGTCGA	GCACACCGAC	AGGCACGGGG	ATCCGGTCGG	CTACATCGT	ATGGAATAACG	300
TCGGCGGGCA	ATCGCTAAA	CGCAGCAAGG	GTCANAAACT	GCCCCTCGCG	GAGGCCATCG	360
CCTACCTGCT	GGAGATCCTG	CCGGCGCTGA	GCTACCTGCA	TTCCCATCGC	TTGGTCTACA	420
ACGACCTGAA	GCCGGAAAAC	ATCATGCTGA	CCGAGGAACA	GCTCAAGCTG	ATCGACCTGG	480
GCGCGGTATC	CGGGATCAAC	TCGTTCGGCT	ACCTCTACGG	GACCCCAGGC	TTCCAGGC	540
CCGAGATCGT	CGGGACCGGT	CCGACGGTGG	CCACCGACAT	CTACACCGTG	GGACGCACGC	600
TCGCGGCGCT	CACGCTGGAC	CTGCCACCCC	GCAATGGCCG	TTATGTGGAT	GGGCTACCCG	660
AAGACGACCC	GGTGTGAAA	ACCTACGACT	CTTACGGCCG	GTTGCTCGCG	AGGGCCATCG	720
ACCCCGATCC	CGGGCAACGG	TTCACCACCG	CCGAAGAGAT	GTCCCGCAA	TTGACGGGCG	780
TGTTGCGGGA	GGTGGTCGCC	CAGACACCGG	GGTGGCGCGG	CCAGGCTATC	AACGATCTTC	840
AGTCCCAGTC	GGTCGACATT	TGGAGTGGAC	TGCTGGTGGC	GCACACCGAC	GTGTATCTGG	900
ACGGGCAGGT	GCACGCGGAG	AAGCTGACCG	CCAACGAGAT	CGTGACCGCG	CTGTCGGTGC	960
CGCTGGTCGA	TCCGACCGAC	GTCGCAGCTT	CGGTCCCTGCA	GGCCACGGTG	CTCTCCCAGC	1020
CGGTGCAGAC	CCTAGACTCG	NTGCGCGCG	CCCGCCACGG	TGCGCTGGAC	GCCGACGGCG	1080
TCGATTNTCC	GAGTCAGTGG	AGCTGCCGCT	AATGGAAGTC	CGCGCGCTGC	TGGATCTCGG	1140
CGATGTGGCC	AAGGCCACCC	GAAAACCTCGA	CGATCTGGCC	GAACCGTTG	GCTGGCGATG	1200
GCGATTGGTC	TGGTACCGGG	CCGTCGCCGA	GCTGCTCACCC	GGCGACTATG	ACTCGGCCAC	1260
CAAACATTC	ACCGAGGTGC	TGGATACCTT	TCCCGGCGAG	CTGGGCCCA	AGCTCGCCCT	1320
GGCCGCCACC	GCCGAACTAG	CCGGCAACAC	CGACGAACAC	AAGTTCTATC	AGACGGTGTG	1380
GAGCACCAAC	GACGGCGTGA	TCTCGCGGC	TTTCGGACTG	GCCAGAGCCC	GGTCGGCCGA	1440
AGGTGATCGG	GTCGGCGCCG	TGCGCACGCT	CGACGAGGTA	CCGCCCACCTT	CTCGGCATTT	1500
CACCACGGCA	CGGCTGACCA	GCGCGGTGAC	TCTGTTGTCC	GGCCGGTCAA	CGAGTGAAGT	1560
CACCGAGGAA	CAGATCCGCG	ACGCCGCCCG	AAGAGTGGAG	GCGCTGCC	CGACCGAAC	1620
ACGCGTGCTG	CAGATCCGCG	CCCTGGTGCT	GGGTGGCGCG	CTGGACTGGC	TGAAGGACAA	1680
CAAGGCCAGC	ACCAACCACA	TCCTCGTTT	CCCCTTCACC	AGTCACGGGC	TGCGGCTGGG	1740
TGTCGAGGCG	TCACTGCGCA	GCCTGGCCCG	GGTAGCTCCC	ACTC		1784

## (2) INFORMATION FOR SEQ ID NO:103:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 766 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

ACAARACACT	CGGYGGCKGC	CGMTCCGGCC	TGATCGTCGG	TGATCAGCYT	CGTGCCAAAY	60	
TCGGCACAAG	GTGCGCGCTR	CCCAANGAGT	TCTTCGCCGC	RGTGCGMGCM	KAAC	120	
ATCN	TGGT	GGTGCCTC	CGCANAAACCC	CGAACCTAA	ACCCATTAA	ACCGGGCAGG	180
AAGTTTCCTA	CATYTACCCN	RGS	MANCCAA	CCGGGCCGCC	NANAAMCCG	TCCTGGANTC	240
CGANC	GGT	CCGGTGTTCG	CCGCACTGCT	GACCGGCACG	GARTATCCGC	AGGC	300
GGCCAACCGCG	TGGGTGCAAC	TGGCCTACGG	TGCGCACCAS	GACGCCATCA	CCGGCTCGGA	360	
GTCCGACCA	GTACTCAATG	CTGGCGACCA	CACCA	CAGACCAAAC	TGGTGCACGC	420	
CGATCTCCAG	GCGCGCCGGC	CCGGTGGCAT	ACGGATTGGT	CGAAACCAAT	CCGAAGGAAT	480	
TCATCACCGA	CGGT	CACCGA	AAACGATCGC	CCC	AAATGGGN	GGACNACCN	540
ATTNACCGTT	NAACAAGTTG	GN	GTAGGTTTC	TTTGATATCG	AKCAACCGAT	ACGGAKCGGM	600
CCGCGGAATG	GTAGACCACC	ACCAGTGC	NCAMGTMGTG	CACCA	GGTTTG	GTCATCGCCC	660
GCAGATCGGT	GACCCGCCA	AGC	GTCCGG	ATGCGGAGAT	GASGGTGACC	AGCCYGGTTG	720
ACCTGTTGAT	CAGGTTNTCC	CAGT	GCCACG	TCGGCAGCTG	GCCGGT		766

## (2) INFORMATION FOR SEQ ID NO:104:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1231 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

CGGCACGAGA ATGTCGCCTG TGCCTCGATA GCCACTTGC	GCG TGTTGCG CTGCCAGCGG 60
GTCAGCCAGG TCGCCTGGTC CAGGCCATCG GGCCGGCGCA	GGAGCGCGAT GTTGGCCAGA 120
CCCAGGTGTAC GAGAACCGGA CTCGACNAAG TGTCGGCGCT	GACGGCGGCT CAGTCGCGG 180
CACACGCCA GATCTATCAG GCCGTCAGCG CCCAGGCCGC	GGCGATTAC GAGATGTTCG 240
TCAACACTCT ACAGATNANC TCAGGGTCGT ATGCTGCTAC	CGAGGCCGCC AACGCAGGCC 300
CGGCCGGCTA GAGGAGTCAC TGCAGATGGAT TTTGGGGCGT	TGCGCCCGGA GGTCAATTCTG 360
GTGCGGATGT ATGCCGGTCC TGGCTCGGCA CCAATGGTCG	CTGCGGGCGTC GGCCTGGAAC 420
GGGTTGGCCG CGGAGCTGAG TTCGGCGGCC ACCGGTTATG	AGACGGTGAT CACTCAGCTC 480
AGCAGTGAGG GGTGGCTAGG TCCGGCGTCA GCAGGCATGG	CCGAGGCAGT TGCGCCGTAT 540
GTGGCGTGGA TGAGTGCCGC TGCGGCACAA GCGGAGCAGG	CGGCCACACACA GGCCAGGGCC 600
GCCGCGGCCG CTTTGAGGG GGCAGTTGCC GCGACGGTGC	CTCCGCGTGT GATCGCGGCC 660
AACCAGGGCTT CGTTGATGCA GCTGATCTCG ACGAATGTCT	TTGGTCAGAA CACCTCGGCC 720
ATCGCGGCCG CGAAGCTCA GTACGGCGAG ATGTGGGCC	AAGACTCCGC GGCATGTAT 780
GCCTACGCGG GCAGTTCGGC GAGCGCTCG GCGGTACAC	CGTTTAGCAC GCCGCCGCAG 840
ATTGCCAACCG CGACCGCTCA GGGTACCGAG GCCGCGGCCG	TGGCCACCGC CGCCGGTACC 900
GCCCAGTCGA CGCTGACGGA GATGATCACC GGGCTACCCA	ACGCGCTGCA AAGCCTCACC 960
TCACNTCTGT TGCAGTCGTC TAACGGTCCG CTGCGTGGC	TGTGGCAGAT CTTGTTCGGC 1020
ACGCCCAATT TCCCCACCTC AATTTCGGCA CTGCTGACCG	ACCTGCAAGCC CTACGCGAGC 1080
TTNTNTNTATA ACACCGAGGG CCTGCCGTAC TTCAAGCATCG	GCATGGCAA CAACTTCAATT 1140
CAGTCGGCCA AGACCCTGGG ATTGATCGGC TAGGCAGCAC	CGGCTGCGGT CGCGGNTGCT 1200
GGGGATNCCG CCAAGGGCTT GCCTCGTGCC G	1231

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2041 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

CGGCACGAGC TCGTGCCGAT CAGTGCCATT GACGGCTTGT	ACGACCTTCT GGGGATTGGA 60
ATACCCAACC AAGGGGGTAT CCTTTACTCC TCACTAGAGT	ACTTCGAAAA AGCCCTGGAG 120
GAGCTGGCAG CAGCGTTTCC GGGTGATGGC TGGTTAGGTT	CGGCCGCGGA CAAATACGCC 180
GGCAAAAACC GCAACCACGT GAATTTTTTC CAGGAACCTGG	CAGACCTCGA TCGTCAGCTC 240
ATCAGCCTGA TCCACGACCA GGCCAACGCG GTCCAGACGA	CCCAGCACAT CCTGGAGGGC 300
GCCAAGAAAG GTCTCGAGTT CGTGCAGCCCG GTGGCTGTGG	ACCTGACCTA CATCCCGGTC 360
GTCGGGCACG CCCTATCGGC CGCCTTCCAN GCGCCGTTTT	GCGCGGGCGC GATGGCCGTA 420
GTGGCGGGCG CGCTTGCTA CTTGGTCGTG AAAACGCTGA	TCAACGCGAC TCAACTCCTC 480
AAATTGCTTG CCAAATTGGC GGAGTTGGTC GCGGCCGCCA	TTGCGGACAT CATTTCGGAT 540
GTGGCGGACA TCATCAAGGG CATCCTCGGA GAAAGTGTGGG	AGTTCATCAC AAACGCGCTC 600
AACGGCCTGA AAGAGCTTG GGACAAGCTC ACGGGGTGGG	TGACCGGACT GTTCTCTCGA 660
GGGTGGTCGA ACCTGGAGTC CTTCTTGCG GCGTCCCCCG	GCTTGACCGG CGCGACCAGC 720
GGCTTGTGCG AAGTGACTGG CTTGTTCGGT GCGGCCGGTC	TGTCCGCATC GTCGGGCTTG 780
GCTCACGCGG ATAGCCTGGC GAGCTCAGCC AGCTTGC	CCCGCCGG CATTGGGGGC 840
GGGTCCGGTT TTGGGGGCTT GCCGAGCCTG GCTCAGGTCC	ATGCCGCCTC AACTCGGCAG 900
GCGCTACGGC CCCGAGCTGA TGGCCCGGTC GCGGCCGCTG	CCGAGCAGGT CGGGGGCAG 960
TCGCAGCTGG TCTCCGCGCA GGGTTCCCAA GGTATGGCG	GACCCGTAGG CATGGCGGC 1020
ATGCACCCCT CTTCGGGGGC GTCGAAAGGG ACGACGACGA	AGAAGTACTC GGAAGGCGCG 1080
GCGCGGGCA CTGAAGACGC CGAGCGCGC CCAGTCGAAG	CTGACGCGGG CGGTGGCAA 1140

AAGGTGCTGG	TACGAAACGT	CGTCTAACGG	CATGGCGAGC	CAAATCCATT	GCTAGCCAGC	1200
GCCTAACAAAC	GCGCAATGCT	AAACGGAAGG	GACACGATCA	ATGACGGAAA	ACTTGACC GT	1260
CCAGCCCGAG	CGTCTCGGTG	TACTGGCGTC	GCACCATGAC	AACGCGCGG	TCGATGCNTC	1320
CTCGGGCGTC	GAAGCTGCCG	CTGGCCTAGG	CGAATCTGTG	GCGATCACTC	ACGGTCCGTA	1380
CTGCTCACAG	TTCAACGACA	CGTTAAATGT	GTACTTGACT	GCCCACAATG	CCCTGGGCTC	1440
GTCCTTGAT	ACGGCCGGTG	TCGATCTCGC	CAAAAGTCTT	CGAATTGCGG	CGAAGATATA	1500
TAGCGAGGCC	GACGAAGCGT	GGCGCAAGGC	TATCGACGGG	TTGTTTACCT	GACCACGTTT	1560
GCTGCCCGCA	GTGCAGGCCA	CGACGTAGCG	CAGGTCGTGT	CCCTCGTAGG	CGTGGATGCG	1620
ACCGGCCAGC	ACCAGCACCC	GGTGCACCC	GATGGGCACG	GACAGTAGCT	CGCCCGCATG	1680
CCCGGCTGCG	GTTGGCGGC	CAAACCCGGG	CAGTTCGGCC	TGCGGCAGCA	CGGTGGTNNG	1740
GGAGCCCAAC	GCCGCAACGG	CCGGTAACCA	TCCCGACCCG	AGCACGACCG	AGACGTATG	1800
TTCGCCGATC	CCGGTGCCTG	CAGCGATGAC	CTGCCGCCGC	CGCCGGGCCA	GTTCGTGCGG	1860
ATCGGGCGC	GGGTCAAGCA	CACTGGCGA	GCTTAACTGA	GCCGCTCGCC	GGGGAGCGGG	1920
TGCTNGTCGA	TGAGATACTG	CGAGCATGCC	AGCAGCCAGC	GCATCCGACC	GGTCGAGGA	1980
ATTGGTGC GG	CGCCGTGGTG	GCGAGCTGGT	CGAGCTGTCC	CATGCCATCC	ACCTCGTGCC	2040
G						2041

## (2) INFORMATION FOR SEQ ID NO:106:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1202 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GAGCTCACCG	CTATCAACCA	ATACTTCTG	CACTCCAAGA	TGCAGGACAA	CTGGGGTTTT	60
ACCGAGCTGG	CGGCCACAC	CCGCGCGGAG	TCGTTCGACG	AAATGCGGCA	CGCCGAGGAA	120
ATCACCGATC	GCATCTTGT	GCTGGATGGT	TTGCCGAACT	ACCAGCGCAT	CGGTTCGTTG	180
CGTATCGGCC	AGACGCTCCG	CGAGCAATT	GAGGCCGATC	TGGCGATCGA	ATACGACGTG	240
TTGAATCGTC	TCAAGCCAGG	AATCGTCATG	TGCCGGGAGA	AACAGGACAC	CACCAGCGCC	300
GTACTGCTGG	AGAAAATCGT	TGCCGACGAG	GAAGAACACA	TCGACTACTT	GGAAACGCAG	360
CTGGAGCTGA	TGGACAAGCT	AGGAGAGGAG	CTTACTCGG	CGCAGTGC GT	CTCTCGCCCA	420
CCGACCTGAT	GCCCCGTTGA	GGATTCTCCG	ATACCACTCC	GGGCGCCGCT	GACAAGCTCT	480
AGCATCGACT	CGAACAGCGA	TGGGAGGGCG	GATATGGCGG	GCCCCACAGC	ACCGACCACT	540
GCCCCCACCG	CAATCCGAGC	CGGTGGCCCG	CTGCTCAGTC	CGGTGCGACG	CAACATTATT	600
TTCACCGCAC	TTGTGTTCGG	GGTGCTGGTC	GCTGCGACCG	GCCAAACCAT	CGTTGTGCC	660
GCATTGCCGA	CGATCGTCGC	CGAGCTGGC	AGCACCGTTG	ACCAGTCGTG	GGCGGTCA CC	720
AGCTATCTGC	TGGGGGGAAC	ACTSKYKKK	KTGKKGKS KS	KSRMRMKCTC	GGTGATCTGC	780
TCGGCCGCAA	CAGGGTGCTG	CTAGGCTCCG	TCGTGGTCTT	CGTCGTTGGC	TCTGTGCTGT	840
GGGGGTTATC	GCAGACGATG	ACCATGCTGG	CGATCTCTCG	CGCACTGCAG	GGCGTCGGTG	900
CCGGTGGCAT	TTCCGTCACC	GCCTACGCC	TGGCCGCTGA	GGTGGTCCCA	CTGCGGGACC	960
GTGGCCGCTA	CCAGGGCGTC	TTANGTGC GG	TGTTCGGTGT	CAACACGGTC	ACCGGTCCGC	1020
TGCTGGGGGG	CTGGCTCACC	GACTATCTGA	GCTGGCGGTG	GGCGTCCGA	CCACCAGCCC	1080
CATCACCGAC	CCGATCGCGG	TCATCGCGC	GAACACCGCC	CTCGCGCGT	TGCGGGCAGG	1140
TCCCTTGGGG	AACGTGGTCC	CACAGCGCCA	GAACGGTCGG	AAATGCGATG	GCCGACCCAC	1200
AC						1202

## (2) INFORMATION FOR SEQ ID NO:107:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 496 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

GGCGGGCGCA	GTTGGCCAGC	AGTTNGGGCG	GGGGAGCCGG	TTCGGNGACC	AAGAAATCGG	60
CCTGGGCAAG	CAGCCGGGAC	CGCGNACCGT	GATCAGTTNG	GATGCCGGG	ACCGCCGCCG	120
ACCAANGCCA	TTCCGCCGNT	GAGGAAGTCG	GAANTNTGCG	CAGTGATGAC	GCCCTGCTGC	180
AACGCNTCCC	GGATTGCCGA	GC GGATGCC	GCCGAACGGC	GGTGCTCACC	ACCGCGAGC	240
ACCCCTACNG	ACAGGCCCGC	ATAGCTGAAT	GACGCCGGGT	NACCGCCGTC	CCNTCCACCG	300
NGANATCCGC	CCGGANGCAA	AAGATCCGTC	GGCGCTCCGC	CTCGGCGACG	ACAGCCACGT	360
TCACCCGCGC	GTTATCGGTG	GCCGCGATCG	CATACCAGGC	GCCGTCAAGG	TNGCCGTYGC	420
GGTAGTCACG	CACCGACAAG	GTGATYTGTT	CCATCGCCTN	GACGGGGGG	GTGACGCTGG	480
GGGCGATCAM	GTGCAC					496

## (2) INFORMATION FOR SEQ ID NO:108:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 849 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

TGGATTCCGA	TAGCGGTTTC	GGCCCCTCGA	CGGGCGACCA	CGGCGCGCAG	GCCTCCGAAC	60
GGGGGGCCGG	GACGCTGGGA	TTCGCCGGGA	CCGCAACCAA	AGAACGCCGG	GTCCGGCGG	120
TCGGGCTGAC	CGCACTGGCC	GGTGATGAGT	TCGGCAACGG	CCCCCGGATG	CCGATGGTGC	180
CGGGGACCTG	GGAGCAGGGC	AGCAACGAGC	CCGAGGCGCC	CGACGGATCG	GGGAGAGGGG	240
GAGGCGACGG	CTTACCGCAC	GACAGCAAGT	AACCGAATTG	CGAATCACGT	GGACCCGTAC	300
GGGTCGAAAG	GAGAGATGTT	ATGAGCCTTT	TGGATGCTCA	TATCCCACAG	TTGGTGGCCT	360
CCCAGTCGGC	GTTTGCCGCC	AAGGCGGGC	TGATGCGGCA	CACGATCGGT	CAGGCCGAGC	420
AGGCAGCGAT	GTCGGCTCAG	GCGTTTCACC	AGGGGGAGTC	GTCGGCGCG	TTTCAGGCCG	480
CCCATGCCCG	GTTTGTGGCG	GCGGCCGCCA	AAGTCAACAC	CTTGTGGAT	GTCGCGCAGG	540
CGAATCTGGG	TGAGGCCGCC	GGTACCTATG	TGGCCGCCGA	TGCTGCCGCC	GGTCGACCT	600
ATACCGGGTT	CTGATCGAAC	CCTGCTGACC	GAGAGGACTT	GTGATGTCGC	AAATCATGTA	660
CAACTACCCC	GCGATGTTGG	GTCACGCCGG	GGATATGGCC	GGATATGCCG	GCACGCTGCA	720
GAGCTTGGGT	GCCGAGATCG	CCGTGGAGCA	GGCCGCGTTG	CAGAGTGCCT	GGCAGGGCGA	780
TACCGGGATC	ACGTATCAGG	CGTGGCAGGC	ACANTGGTAA	CCANGCCANG	GAAGATTG	840
					TGCGGGCCT	849

## (2) INFORMATION FOR SEQ ID NO:109:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 97 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

Met	Ser	Leu	Leu	Asp	Ala	His	Ile	Pro	Gln	Leu	Val	Ala	Ser	Gln	Ser
1															15
Ala	Phe	Ala	Ala	Lys	Ala	Gly	Leu	Met	Arg	His	Thr	Ile	Gly	Gln	Ala
20															30
Glu	Gln	Ala	Ala	Met	Ser	Ala	Gln	Ala	Phe	His	Gln	Gly	Glu	Ser	Ser
35															45
Ala	Ala	Phe	Gln	Ala	Ala	His	Ala	Arg	Phe	Val	Ala	Ala	Ala	Lys	
50															60
Val	Asn	Thr	Leu	Leu	Asp	Val	Ala	Gln	Ala	Asn	Leu	Gly	Glu	Ala	Ala
65															80
Gly	Thr	Tyr	Val	Ala	Ala	Asp	Ala	Ala	Ala	Ser	Thr	Tyr	Thr	Gly	

85

90

95

Phe

## (2) INFORMATION FOR SEQ ID NO:110:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

Met Ser Leu Leu Asp Ala His Ile Pro Gln Leu Val Ala Ser Gln  
 1               5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:111:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

Ala His Ile Pro Gln Leu Val Ala Ser Gln Ser Ala Phe Ala Ala  
 1               5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:112:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Leu Val Ala Ser Gln Ser Ala Phe Ala Ala Lys Ala Gly Leu Met  
 1               5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:113:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

Ser Ala Phe Ala Ala Lys Ala Gly Leu Met Arg His Thr Ile Gly  
 1               5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:114:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

Lys Ala Gly Leu Met Arg His Thr Ile Gly Gln Ala Glu Gln Ala  
 1               5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:115:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

Arg His Thr Ile Gly Gln Ala Glu Gln Ala Ala Met Ser Ala Gln  
 1               5               10               15

## (2) INFORMATION FOR SEQ ID NO:116:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

Gln Ala Glu Gln Ala Ala Met Ser Ala Gln Ala Phe His Gln Gly  
 1               5               10               15

## (2) INFORMATION FOR SEQ ID NO:117:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

Ala Met Ser Ala Gln Ala Phe His Gln Gly Glu Ser Ser Ala Ala  
 1               5               10               15

## (2) INFORMATION FOR SEQ ID NO:118:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

Ala Phe His Gln Gly Glu Ser Ser Ala Ala Phe Gln Ala Ala His  
 1               5               10               15

## (2) INFORMATION FOR SEQ ID NO:119:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

Glu Ser Ser Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val Ala  
 1               5               10               15

## (2) INFORMATION FOR SEQ ID NO:120:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

Phe Gln Ala Ala His Ala Arg Phe Val Ala Ala Ala Ala Lys Val  
 1                   5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:121:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

Ala Arg Phe Val Ala Ala Ala Lys Val Asn Thr Leu Leu Asp  
 1                   5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:122:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

Ala Ala Ala Lys Val Asn Thr Leu Leu Asp Val Ala Gln Ala Asn  
 1                   5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:123:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

Asn Thr Leu Leu Asp Val Ala Gln Ala Asn Leu Gly Glu Ala Ala  
 1                   5                   10                   15

## (2) INFORMATION FOR SEQ ID NO:124:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 18 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

Val Ala Gln Ala Asn Leu Gly Glu Ala Ala Gly Thr Tyr Val Ala Ala  
 1                   5                   10                   15

Asp Ala

## (2) INFORMATION FOR SEQ ID NO:125:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1752 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

CGGCACGAGA ATGTCGCCTG TGCCTCGATA GCCACTTGCG TGTGGTCGCG CTGCCAGCGG	60
GTCAGCCAGG TCGCCTGGTC CAGGCCATCG GGCCGGCGCA GGAGCGCGAT GTTGGCCAGA	120
CCCAGGTGAC GAGAACCGGA CTCGACNAAG TGTCGGCGCT GACGGGGCT CAGTTCGCGG	180
CACACGCCCA GATCTATCAG GCCGTCAGCG CCCAGGCCGC GGCGATTAC GAGATGTTCG	240
TCAACACTCT ACAGATNANC TCAGGGTCGT ATGCTGCTAC CGAGGCCGCC AACGCGGCCG	300
CGGCCGGCTA GAGGAGTCAC TGCGATGGAT TTTGGGGCGT TGCCGCCGGA GGTCAATTCC	360
GTGCGGATGT ATGCCGGTCC TGGCTCGGCA CCAATGGTCG CTGCGGGTCA GGCCTGGAAC	420
GGGTTGGCCG CGGAGCTGAG TTCGGCGGCC ACCGGTTATG AGACGGTGAT CACTCAGCTC	480
AGCAGTGAGG GGTGGCTAGG TCCGGCGTCA GCGGCGATGG CCGAGGCAGT TGCGCCGTAT	540
GTGGCGTGGA TGAGTGCAG TGCGGCGCAA GCCGAGCAGG CGGCCACACA GCCCAGGGCC	600
GCCGCGGCCG CTTTGAGGC GGCGTTGCC GCGACGGTGC CTCCGCCGTT GATCGCGGCC	660
AACCAGGCTT CGTTGATGCA GCTGATCTCG ACGAATGTCT TTGGTCAGAA CACCTCGGCG	720
ATCGCGGCCG CGGAAGCTCA GTACGGCGAG ATGTGGGCCA AAGACTCCGC GGCATGTAT	780
GCCTACGCGG GCAGTTCGGC GAGCGCTCG GCGGTACAGC CGTTAGCAC GCCGCCGCAG	840
ATTGCCAACCG CGACCGCTCA GGGTACGCAG GCCGCGGCCG TGGCCACCGC CGCCGGTACC	900
GCCCAGTCGA CGCTGACGGA GATGATCACC GGGCTACCCA ACGCGCTGCA AAGCCTCACC	960
TCACNTCTGT TGCAGTCGTC TAACGGTCCG CTGTCGTGGC TGTGGCAGAT CTTGTTCGGC	1020
ACGCCAATT TCCCCACCTC AATTTCGGCA CTGCTGACCG ACCTGCAGCC CTACGCGAGC	1080
TTNTNTATA ACACCGAGGG CCTGCCGTAC TTCAGCATCG GCATGGCAA CAACTTCATT	1140
CAGTCGGCCA AGACCCTGGG ATTGATCGGC TAGGCAGGCAC CGGCTGGGT CGCGGCTGCT	1200
GGGGATGCCG CCAAGGGCTT GCCTGGACTG GGCGGGATGC TCGGTGGCG GCCGGTGGCG	1260
GCAGGTCTGG GCAATGCGGC TTCGGTTGGC AAGCTGTCGG TGCCGCCGGT GTGGANTGGA	1320
CCGTTGCCCG GGTGGTGAC TCCGGGGCT GCTCCGCTAC CGGTGAGTAC GGTAGTGCC	1380
GCCCCGGAGG CGGCAGCCCGG AAGCCTGTT GGCGGCCCTGC CGCTANCTGG TGCGGGCGGG	1440
GCCGGCGCGG GTCCACGCTA CGGATTCCRT CCCACCGTCA TGGCTGCC ACCCTTCGMC	1500
GGGATAGTCG CTGCCGCAAC GTATTAACGC GCCGGCCTCG GCTGGTGTGG TCCGCTGCC	1560
GTGGCAATTG GTCNGCGCCG AAATCTCSGT GGGTTATTTR CGGTGGGATT TTTTCCCGAA	1620
GCCGGGTTCA RCACCGGATT TCCTAACGGT CCCGCKACTC TCGTGCCGAA TTCGCACTA	1680
AGTGACGTCC GGCGGAAACC CGTTGGGTNT GAAAGCTTCA GAAAGGCCCG CTCCCAAGGGG	1740
TTCCGGCAAAC GG	1752

## (2) INFORMATION FOR SEQ ID NO:126:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

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

65	70	75	80
Ala Gln Ala Glu Gln Ala Ala Thr Gln Ala Arg Ala Ala Ala Ala			
85	90	95	
Phe Glu Ala Ala Phe Ala Ala Thr Val Pro Pro Pro Leu Ile Ala Ala			
100	105	110	
Asn Arg Ala Ser Leu Met Gln Leu Ile Ser Thr Asn Val Phe Gly Gln			
115	120	125	
Asn Thr Ser Ala Ile Ala Ala Glu Ala Gln Tyr Gly Glu Met Trp			
130	135	140	
Ala Gln Asp Ser Ala Ala Met Tyr Ala Tyr Ala Gly Ser Ser Ala Ser			
145	150	155	160
Ala Ser Ala Val Thr Pro Phe Ser Thr Pro Pro Gln Ile Ala Asn Pro			
165	170	175	
Thr Ala Gln Gly Thr Gln Ala Ala Ala Val Ala Thr Ala Ala Gly Thr			
180	185	190	
Ala Gln Ser Thr Leu Thr Glu Met Ile Thr Gly Leu Pro Asn Ala Leu			
195	200	205	
Gln Ser Leu Thr Ser Xaa Leu Leu Gln Ser Ser Asn Gly Pro Leu Ser			
210	215	220	
Trp Leu Trp Gln Ile Leu Phe Gly Thr Pro Asn Phe Pro Thr Ser Ile			
225	230	235	240
Ser Ala Leu Leu Thr Asp Leu Gln Pro Tyr Ala Ser Xaa Xaa Tyr Asn			
245	250	255	
Thr Glu Gly Leu Pro Tyr Phe Ser Ile Gly Met Gly Asn Asn Phe Ile			
260	265	270	
Gln Ser Ala Lys Thr Leu Gly Leu Ile Gly Ser Ala Ala Pro Ala Ala			
275	280	285	
Val Ala Ala Ala Gly Asp Ala Ala Lys Gly Leu Pro Gly Leu Gly Gly			
290	295	300	
Met Leu Gly Gly Pro Val Ala Ala Gly Leu Gly Asn Ala Ala Ser			
305	310	315	320
Val Gly Lys Leu Ser Val Pro Pro Val Trp Xaa Gly Pro Leu Pro Gly			
325	330	335	
Ser Val Thr Pro Gly Ala Ala Pro Leu Pro Val Ser Thr Val Ser Ala			
340	345	350	
Ala Pro Glu Ala Ala Pro Gly Ser Leu Leu Gly Gly Leu Pro Leu Xaa			
355	360	365	
Gly Ala Gly Gly Ala Gly Ala Gly Pro Arg Tyr Gly Phe Xaa Pro Thr			
370	375	380	
Val Met Ala Arg Pro Pro Phe Xaa Gly Ile Val Ala Ala Ala Thr Tyr			
385	390	395	400

## (2) INFORMATION FOR SEQ ID NO:127:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 474 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

GGCACGAGCA CCAGTTGACC CGCGAAGAAC CTGACCGCGC CACCCAGCGC CGCCCGCATC	60
ACCGGGCCCCG TCCCACGAAC CTTTCGGTA AACGAGCCAC TCCAGGGAG ATCGGTACCG	120
CCCGACGCAT TTGGTGTAAG GACCACCTCG CGAAGTAGT CCTGGACGGG TGTCCTCGCG	180

CCAACCAGCT	TGTAGACGTG	GCGACGGTCC	TGCTCATACT	CGACGGTCTC	TTCTGCACG	240
AACACCGGCC	ACATGCCTAG	TTTGC GGATG	GCCCCGATGC	CGCCGGGC	GGGATCACCG	300
CGTCGCGCCC	AACTCGATTG	AGCAACGATG	GGCTTGGCCC	AGGTGCC	GTTGCCACCG	360
TCTGTCACGA	GCCGAAACAA	GGTTGCAGCC	GGCGCGCTGC	TGGTCTTGGT	GACCTCGAAC	420
GAAAATTCC	GACCCGACAT	GCGCGACTCC	CGAAACGACA	ACTGAAGCTC	GTGC	474

## (2) INFORMATION FOR SEQ ID NO:128:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1431 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

CTGCGCGCCG	GAAAAAANTA	TTACTGGCAG	GACCGGCAGA	ATGCATGGTG	ATATTCCGGT	60
GATGAGGCCG	CCGAGGAACC	GACTAGTGC	AGGGTCAACA	CATCGTTAT	TCGTTGCCGT	120
TTAGGTCTTG	GATCTGCCGG	GACGGCAACG	AGTTGGCAGG	ACCGCTCACG	CGAGCGCTGT	180
TGACAGAGTC	GGTCACGTC	GAACTCGCCA	CCCGTCAGAT	GCGAATGATA	GCCACATCGG	240
CCACACCATC	GACGGCGTCG	AAGTCGCCGT	CGTGGGTAC	GACCGGCACC	CCTTGCAGCC	300
TGGCAACGGC	AGCGGCCCTC	ACCGGACGGG	ACCGAGATCG	TCGGTGGTGT	CGCCAGTGAG	360
CGTTGCGAGG	TCGCGGGTGC	AATCCCGCAT	CTGCTTGC	ATGCCGAAGC	CGCCGCAGCA	420
GCTCGTCTCG	ACTCAACCAT	CGGCGCCGTG	CGGGCTGCCT	GCGGTCA	GCGCAACGGG	480
TTTGCCTTG	GCAGTGATGG	TGATGTCTTC	GCCGGCCTGC	ACGCGCCGTA	GCAGCCGGC	540
GGTGTGTTG	CGCAGTTCGC	GAGACGCGAC	TTCA	GAGCAGTC	ATGCTGCCGG	600
CGCTGGCGC	GGTGTACCG	TCATGCGCTT	GGGATATCAC	GTGATCTATC	GGCACGAAGC	660
CGCCGGATGA	GCGAGGCAA	CCGCCTACAC	GGGCTGCCTC	GCCTTGACCG	CGCCGAACGT	720
TACTGTGCCG	GGGGCATCAG	CACCGTATCG	ATCATGTACA	CCGTCGCGT	GGCGGTGTGA	780
CTCCGCCACA	TACCAAACGG	GCGTTGTTGA	CCATGAGTCG	TCGCGGGCGC	CTATCACC	840
CAGGTCGGCA	CCTTGCAGGT	CTGATGGGTG	CCGTCGATCC	TGCTCGGACT	CGCCTGGCCG	900
GCTATCACGT	GGTAGGTCAG	GATGCTGCTG	AGCAGCTTGG	CGTCAGTCTT	GAGTTGATCG	960
ATAGTGGCCG	CCGGCAGCTT	GTCGAATGCG	GCGTTGGTGG	GGGCGAAAAC	GGTGTACTCG	1020
CCGCCGTTGA	GGGTGTCGAC	CAGATTACA	TCCGGGTTCA	GCTTGCCGA	CAGAGCCGAG	1080
GTCAGGGTAC	TGAGCATCGG	GTTGTTGGAA	GCCGCGGTAG	CGACCGGGTC	TTGCGCCATT	1140
CCGGCCACCG	ATCCGGGACC	GGTGGGATT	TGCGCCGCGT	ATTGCGCGA	CCCACGACCA	1200
ATCAGGTCCG	CTGCGGTCAG	CCATTGCC	CGTGGTAACG	GGCGCCGCCG	GGCTGGTC	1260
CGGTTTCCGG	CTGGTGTCTT	GCGACACGGG	TTGGTGCTC	GAACAACCCG	CTAAGAACGC	1320
AATCGCGATG	GCTGCGAGGC	TCGCTGCTG	GGCCGGTTTG	GCCTGAACGT	TGATCATCGC	1380
TTCGATTCC	TTGCTTCTGC	GGCGGCC	GGTGGTGG	AACGCCGTCC	TCCTGGTGG	1431

## (2) INFORMATION FOR SEQ ID NO:129:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

GCACGAGAGT	CGTATCTTG	CACCCAGCGC	CCGCTAGGAAA	CCGCTGGCCT	GGCTAACTCA	60
GATGCGGGCG	GCCGTCGATT	CGAGAGGTAA	CCGATCGCCC	GCCGACAATG	GGTTACCCAC	120
CGAGACTGAT	TGCCGCGCAG	CCGCCTTCGA	CGTGTAAAGCG	CCGGTTCGT	CATGCCCGGA	180
ACGGCTGCAC	TCACGGACCT	TCTACGTAGT	ACGTGACGGA	CTTTTACGCA	TTATCGCTGA	240
CGATCTTGC	CTCCCAGGAC	TCCAGAATCT	ACTCGTGCC			279

## (2) INFORMATION FOR SEQ ID NO:130:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1470 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

ACCGCCACCC	GCAGCCCCGGA	ATCACCGTCG	GTAACCTGCG	AATAACAATT	CTTCATCGAC	60
GACTTCGCGA	ACAGCGAAC	CGAGCCCACC	GCCTGATAGC	CTTCTTCCTC	GATGTTCCAA	120
CCGCCGGCGG	CGTCGAACGA	AACGATAACGA	CCCGCGCTCT	GCAGGGTCAGA	CGCATGAATG	180
TCGTAGCCCG	CCAGCAACGG	CAACGCCAGC	AGACCCTGCA	TCGCGGCCGC	CAGATTGCCA	240
CGCACCAATAA	TCGCCAGCCG	GTTGATTTG	CCGGCAAACG	TCAGCGGCAC	ACCCTCGAGC	300
TTCTCGTAGT	GCTCAAGTTC	CACGGCATAAC	AGCCGGCAA	ACTCAACCAGC	GACCGCAGCC	360
GTGCCAGCGA	TGCCGGTAGC	GGTGTAGTCA	TCGGTGATAT	ACACCTTGC	CACATCACGC	420
CCAGAAATCA	TGTTGCCCTG	CGTCGAACGC	CGGTACCCCG	CCATGACAAC	ACCGCCGGGG	480
TATTTCAGCG	CGACAATGGT	GGTCCGTGC	GGCAGTTGCG	CATCGCCGCC	TGCGAGTGGC	540
GCACCGCCGC	TGATGCTTGC	CGGCAGCAAC	TCCGGCGCCT	GGCGGCGCAG	GAAGTCAAGT	600
GAAAGAAGAT	AGGTCTACAG	CGGGTGTCC	AGAGAGTGAA	TTAATGGACA	GGCGATCGGG	660
CAACGGCCAG	GTCACTGTCC	GCCCTTTG	ACGTATGCGC	GGACGAAGTC	CTCGCGTTC	720
TCCTCGAGGA	CGTCGTCGAT	TTCGTCGAGC	AGATCGTCGG	TCTCCTCGGT	CAGCTTTCG	780
CGACGCTCCT	GGCCCGCGGC	GGTGCTGCCG	GCGATGTCGT	CATCATCGCC	GCCGCCACCG	840
CCACGCTTGG	TCTGCTCTTG	CGCCATCGCC	GCCTCCTGCT	TCCTCATGGC	CTTCAAAAG	900
GCCGCGGGTG	CGCGTCACAC	GCCCCGTGTC	TTTCTCTCAC	CTACCGGTCA	ACACCAACGT	960
TTCCCGGCCT	AACCAGGCTT	AGCGAGGCTC	AGCGGTCA	TGCTCTACCA	GCTCCACGGC	1020
ACTGTCCACC	GAATCCAGCA	ACGCACCAAC	ATGCGCCTTA	CTACCCCGCA	ACGGCTCCAG	1080
CGTCGGGATG	CGAACCCAGCG	AGTCGCCGCC	AGGTGCAAGA	TCACCGAGTC	CCAGCTAGCC	1140
CGGGCGATAT	CAGCCCCGAA	CCGGCGCAGG	CATTCGCG	CGGAAATACG	CGCGGGTGT	1200
GGTCGGCGGT	TCTCCACCAGC	ACTCAGCACC	TGGTGTTCG	GTGACTAAC	GCTTTATCGA	1260
GCCGCGCCCG	ACCAGCCGGT	TGTACAGGCC	CTTGTCCAGC	CGGACATCGG	AGTACTGCAG	1320
GTTGACGAGG	TGCAGCCGGG	GCGCCGACCA	GCTCAGGGTTC	TCCCGCTGCC	GGAAACCGTC	1380
GAGCAGCCGC	AGTTTGGCCG	GCCAGTCCAG	CAGCTCCGCG	CAATCCATCG	GGTCACGCTC	1440
GAGCTGATCC	AGCACGTGTG	CCCAAGGTTTC				1470

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1059 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

ATCCCCATCG	CTCCGGCACC	TATCACCAAGG	TAGTCGGTTT	CGATGGTTT	CGCCGGCCCT	60
TGCGTTGGCC	TGGGCCACGG	GTCGTTCATG	GGCCCTCCTG	TGCGGATTGG	AATTTGTGAC	120
AAAGAAATCG	GGCGATCGGT	GAGCAATCGT	CGCCGATGCA	AGACACGCTT	TCGCTGCCGC	180
GGCGTCAGGT	GGAGTTTAGG	CCAGCGTAAC	AAACGTAGACC	GGCCACTGAC	CAAACCCCAA	240
ACCCACAAAC	CCTGGACGCA	TGCGGGTCTC	GGGCGTAAA	TTCCGGGTAG	ATATCGTATA	300
CCGATATCGG	ATGCCGTAGC	CTTATCGAGG	CATGAGACGC	CCGCTAGACC	CACCGATAT	360
TCCAGATGAG	CTGCGGCCAC	GGCTGGGGCT	CTTGGATGCG	GTGGTGATCG	GGCTTGGGTC	420
CATGATCGGT	GCCGGAATCT	TTGCTCGTGC	CGAATTGCGC	ACGAGCTCGT	GCCGAATTG	480
GCACGAGATT	CCAATCCCCA	GAAGGTCGTA	CAAGCCGTCA	ATGGCACTTG	ATCGTTGGAT	540
CGATGATGAA	CGCTCTGCTC	ATGCCTGCCG	CCTATCTCAA	CGGTGTCGA	TTCCATGCAT	600
TAGCCTTGGT	TCTGCATTGC	ACGCGTAGGG	CCTACAGTCT	GGCTGTCATG	CTTGGCCGAT	660
GTCAACAGTT	TTTTCATGC	TAAGCAGATC	GTCAGTTTG	AGTCGTGAA	GACGGCATGT	720
TCACTTGGTG	TCGACTACAT	CGTCTCGCA	CATTGCGCCT	CCTGCAACTG	CGCTGCGACA	780

ATGCGCCAAC CGCCGTGTAG CTCGTGCCGA ATTGGCACG AGGATCCACC GGAGATGGCC	840
GACGACTACG ACGAGGCCTG GATGCTAAC ACCGTGTTCG ACTATCACAA CGAGAACGCA	900
AAAGAAGAGG TCATCCATCT CGTCCCCGAC GTGAACAAGG AGAGGGGCC CATCGAACTC	960
GTAACCAAAGG TAGACAAAGA GGGACATCAG ACTCGTCTAC GATGGGGAGC CACGTTTCA	1020
TACAAGGAAC ATCCTAAGTT TTGATTGGG AACATCCTA	1059

## (2) INFORMATION FOR SEQ ID NO:132:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 153 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

GCACGAGGCA TTGGCGGGCA TCTGCATAAA CGGTGACGTA TCAGCACAAA ACAGCGGAGA	60
GAACAAACATG CGATCAGAAC GTCTCCGGTG GCTGGTAGCC GCAGAAAGTC CGTTCGCCTC	120
GGTGTATTTC GACGACTCGC ACGACTCGT CCG	153

## (2) INFORMATION FOR SEQ ID NO:133:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

CCGCGCGGTC GATCAGCGAG CCAGGCAAAA ACTCCGTCGA GCCCGAGTCG ATGATGGTCA	60
CCCAGCGCAG CATCTGGCGA ACGATCACCT CGATGTGCTT GTCGTGGATC GACACACCTT	120
GGGCGCGGTA GACCTCCTGG ACCTCGCGAA CCAGGTGTAT CTGCACCTCG CGGGGGCCCT	180
GCACCCGAG CACCTCATGC GGGTCGGCCG AGCCTTCCAT CAGCTGCTGG CCCACCTCGA	240
CGTGGTCGCC ATCGGAGAGC ACCCGTTCGG AACCGTCTTC GTGCTTGAAC ACCCGCAGCC	300
GCTGCCGCTT GGAGATCTTG TCGTAGACCA CTTCCCTCACC GCCGTCGTCA GGAACGATGG	360
TGATCTTGTG GAACCGCTCG CGTCCT	387

## (2) INFORMATION FOR SEQ ID NO:134:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 389 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

GTTCAGCACG GCTATCCGAT TGTGCCGTT GCTCGGTGG GTGCTGAACA CGGCATCGAC	60
ATCGTGCTCG ACAACGAATC CCCACTGCTG GCACCGGTCC AGTTCCCTCGC CGAGAACGCTG	120
CTCGGCACCA AAGACGGTCC GGCGCTGGTC CGTGGTGTG GACTGACACC GGTACCGCGC	180
CCCGAACGGC AGTATTACTG GTTCGGCGAG CCAACCGACA CCACAGAGTT TATGGGGCAG	240
CAAGCCGACG ATAACGCCGC ACGCAGGGTG CGCGAGCGTG CCGCCGCCGC TATCGAACAC	300
GGCATCGAGC TGATGCTGGC CGAGCGCGCA GCCGATCCAA ATCGATCCCT GGTCGGACGG	360
CTCTTGCCT CGGACGCCCTA AGGCGCCCC	389

## (2) INFORMATION FOR SEQ ID NO:135:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 480 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

CCCGCGGTCG	GAATGATCCC	CGTCTCGTCG	CGCGCCCATT	TGATGCTGTT	GATGAGCTGT	60
TTGGAGAACG	CCGGTTGGCG	TACCGGTGAG	CCGGAATATC	TGTTGGAAGC	GTCACCGGAT	120
GTNCACATGA	ANTNCNTTGN	CCCNGTNGCG	GTNTTGGNTG	NGGNAACAC	GTGTTGTNTA	180
AGCCTTGNNTG	GNCTCGNAAG	NGCCGTNGAC	GCCTGTGTCG	CCGAAGATAA	TGAGCACCTG	240
ACGGTTGGCG	GGATCGCCGT	TATCCAAGG	AATTCCGAGG	TCGGTCCCAG	AGATGCCGAA	300
GCGTTCCAGG	GTCTTGTGTTG	GGCTGTCCCG	TCCGGTCACC	CACTCGCGA	GGGATGTGGN	360
AGCCCCGGCG	AGCGTGGCAC	CAGGATCCGG	CGCCGCGGCC	GGAGCAGGGT	CGGNNGCTGN	420
NCTGNNTTCC	TNNNGCCNAA	TTNNACTCCN	NCNACAANCT	TGNNNNCCGAC	TCNNACCCGN	480

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 587 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

GCACGAGGCT	ACCGGCGCGT	CGCCCGCCAT	GCCCTGGATG	CACCGTAGC	CACCCGTNCA	60
TNCAGCGGGT	CAGCCGCCGC	GTCCGGGCTT	AACGCTATAG	CAGCTGAAA	CAACCCAGCG	120
CCGGCAATTA	CTTGATGTT	GAACCGATGA	CCATNGCCTN	CGNGTNCAAT	CTCNTCTCTT	180
NGCGCGCCNC	TATTTNNGCC	ATANATTG	TTNNNANNCGN	AACGCTAGAC	GTATCGAGTT	240
CCTTTTCGAC	CACCGGCTCA	ATTGTCAGCA	TCCTATGGGG	AACATGAGCC	CCGCCGCACC	300
GGGCCGTTTC	CAAATGGTGA	CGTCACAACG	GTGTCACAAG	CCAGCGCAAT	GTCCGCGGTA	360
GGGACGCGGC	GGCTGGGATC	GGTGGGTGA	GCGCCCGGCT	TCTCAAAGCG	AGGGGAGCCC	420
CGGGACTCTT	ACCGGCCGAA	GGCGCGGGT	GTCACTGATC	TAGGCTGACG	GCCAGTGGTT	480
GNTNAGCCAA	CAAGGATGAC	NACAAATAAN	CCGAGGANAG	ACANGNGACG	GNCCGANANG	540
CTNANCCGN	NTTGNNCNAA	NNNNACNCAC	TTNTACCGNN	CTTATGN		587

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1200 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

CAGGCATGAG	CAGAGCGTTC	ATCATCGATC	CAACGATCAG	TGCCATTGAC	GGCTTGTACG	60
ACCTTCTGGG	GATTGGAATA	CCCAACCAAG	GGGGTATCCT	TTACTCCTCA	CTAGAGTACT	120
TCGAAAAAACG	CCTGGAGGAG	CTGGCAGCAG	CGTTTCCGGG	TGATGGCTGG	TTAGGTTCGG	180
CCGCGGACAA	ATACGCCGGC	AAAAACCGCA	ACCACGTGAA	TTTTTTCCAG	GAACTGGCAG	240
ACCTCGATCG	TCAGCTCATC	AGCCTGATCC	ACGACCAGGC	CAACGCGGTC	CAGACGACCC	300
GCGACATCCT	GGAGGGCGCC	AAGAAAGGTC	TCGAGTCGT	GCGCCCGGTG	GCTGTGGACC	360
TGACCTACAT	CCCGGTCGTC	GGGCACGCC	TATCGGCCG	CTTCCAGGCG	CCGTTTGCG	420
CGGGCGCGAT	GGCCGTAGTG	GGCGCGCGC	TTGCCTACTT	GGTCGTGAAA	ACGCTGATCA	480
ACGCGACTCA	ACTCCTCAA	TTGCTTGCCA	AATTGGCGGA	GTTGGTCGCG	GCCGCCATTG	540
CGGACATCAT	TTCGGATGTG	CGGGACATCA	TCAAGGGCAC	CCTCGGAGAA	GTGTGGGAGT	600
TCATCACAAA	CGCGCTCAAC	GGCCTGAAAG	AGCTTGGGA	CAAGCTCACG	GGGTGGGTGA	660
CCGGACTGTT	CTCTCGAGGG	TGGTCGAACC	TGGAGTCCTT	CTTTGCGGGC	GTCCCCGGCT	720
TGACCGGCGC	GACCAGCGGC	TTGTCGCAAG	TGACTGGCTT	GTTCGGTGC	GCCGGTCTGT	780
CCGCATCGTC	GGGCTTGGCT	CACGCGGATA	GCCTGGCGAG	CTCAGCCAGC	TTGCCCGCCC	840
TGGCCGGCAT	TGGGGCGGG	TCCGGTTTG	GGGGCTTGCC	GAGCCTGGCT	CAGGTCCATG	900
CCGCCTCAAC	TCGGCAGGCG	CTACGGCCCC	GAGCTGATGG	CCCGGTCGGC	GCCGCTGCCG	960

AGCAGGTCGG CGGGCAGTCG CAGCTGGTCT CCGCGCAGGG TTCCCAAAGGT ATGGGCAGGAC	1020
CCGTAGGCAT GGGCGGCATG CACCCCTCTT CGGGGGCGTC GAAAGGGACG ACGACGAAGA	1080
AGTACTCGGA AGGCAGGGCG GCAGGGCACTG AAGACGCCGA GCGCGGCCA GTGAAAGCTG	1140
ACGCAGGGCGG TGGGCAAAAG GTGCTGGTAC GAAACGTCGT CTAACGGCAT GGCGAGCCAA	1200

## (2) INFORMATION FOR SEQ ID NO:138:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 392 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly	
1 5 10 15	
Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu	
20 25 30	
Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala	
35 40 45	
Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala	
50 55 60	
Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu	
65 70 75 80	
Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln	
85 90 95	
Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val	
100 105 110	
Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala	
115 120 125	
Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val	
130 135 140	
Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala	
145 150 155 160	
Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala	
165 170 175	
Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Thr	
180 185 190	
Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys	
195 200 205	
Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg	
210 215 220	
Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr	
225 230 235 240	
Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala	
245 250 255	
Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser	
260 265 270	
Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Ser Gly Phe	
275 280 285	
Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln	
290 295 300	
Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln	
305 310 315 320	

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met  
                  325                     330                     335  
 Gly Gly Pro Val Gly Met Gly Met His Pro Ser Ser Gly Ala Ser  
                  340                     345                     350  
 Lys Gly Thr Thr Thr Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr  
                  355                     360                     365  
 Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gln  
                  370                     375                     380  
 Lys Val Leu Val Arg Asn Val Val  
                  385                     390

## (2) INFORMATION FOR SEQ ID NO:139:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 439 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

ACGTTTACCC	ATGCCGTCGG	TGCAGAGCAA	CGCCAGACAA	CACAAAGTAG	TCTAATTCCG	60
TTATAAAGCA	GACATTTCCG	TGGTTATGTA	GAAGATGTGCG	ACCGATCAGA	TGAAGCGATC	120
CGCGTCAGGT	GGTATCCGAT	GTCTTTGTG	ACCATCCAGC	CGGTGGTCTT	GGCAGCCGCG	180
ACGGGGGACT	TGCCGACGAT	CGGTACCGCC	GTGAGTGCTC	GGAACACAGC	CGTCTGTGCC	240
CCGACGACGG	GGGTGTTACC	CCCTGCTGCC	AATGACGTGT	CGGTCCGTAC	GGCGGCCCGG	300
TTCACCGCGC	ACACCAAGCA	CTACCGAGTG	GTGAGTAAGC	CGGCCGCGCT	GGTCCATGGC	360
ATGTTCTGTC	CCCTCCCGGC	GGCCACCGCC	GATGCGTATG	CGACCACCGA	GGCCGTCAAT	420
GTGGTCGCGA	CCGGTTAACG					439

## (2) INFORMATION FOR SEQ ID NO:140:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1441 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

GAGGTTGCTG	GCAATGGATT	TCGGGCTTT	ACCTCCGGAA	GTGAATTCAA	GCCGAATGTA	60
TTCCGGTCCG	GGGCCGGAGT	CGATGCTAGC	CGCCGCGGCC	GCCTGGGACG	GTGTGGCCGC	120
GGAGTTGACT	TCCGCCGCGG	TCTCGTATGG	ATCGGTGGTG	TCGACGCTGA	TCGTTGAGCC	180
GTGGATGGGG	CCGGCGGCCG	CCCGCGATGGC	GGCCGCGGCA	ACGCCGTATG	TGGGGTGGCT	240
GGCCGCCACG	CGGGCGCTGG	CGAAGGAGAC	GGCCACACAG	GCGAGGGCAG	CGGGCGGAAGC	300
GTTCGGGACG	GCGTTCGCGA	TGACGGTGCC	ACCATCCCTC	GTCGCGGCCA	ACCGCAGCCG	360
GTTGATGTCG	CTGGTCGCGG	CGAACATTCT	GGGGCAAAAC	AGTGCGGCGA	TCGCGGCTAC	420
CCAGGCGGAG	TATGCCGAAA	TGTGGGCCA	AGACGCTGCC	GTGATGTACA	GCTATGAGGG	480
GGCATCTGCG	GCCGCGTCGG	CGTTGCCGCC	GTTCACTCCA	CCCGTGCAAG	GCACCGGCC	540
GGCCGGGCC	GC GGCGCAG	CCGCGGCCGAC	CCAAGCCGCC	GGTGCAGGCC	CCGTTGCCGA	600
TGCACAGGCG	ACACTGGCCC	AGCTGCC	GGGGATCCTG	AGCGACATTC	TGTCCGCTAC	660
GGCCGCCAAC	GCTGATCCGC	TGACATCGGG	ACTGTTGGGG	ATCGCGTCGA	CCCTCAACCC	720
GCAAGTCGGA	TCCGCTCAGC	CGATAGTGAT	CCCCACCCCG	ATAGGGGAAT	TGGACGTGAT	780
CGCGCTCTAC	ATTGCATCCA	TCGCGACCGG	CAGCATTGCG	CTCGCGATCA	CGAACACGGC	840
CAGACCCCTGG	CACATCGGCC	TATACGGGAA	CGCCGGCGGG	CTGGGACCGA	CGCAGGGCCA	900
TCCACTGAGT	TCGGCGACCG	ACGAGCCGGA	GCCGCACTGG	GGCCCCCTCG	GGGGCGCGGC	960
GCCGGTGTCC	GC GGCGTCG	GCCACGCAGC	ATTAGTCGGA	GCGTTGTCGG	TGCCGCACAG	1020
CTGGACCACG	GCCGCCCCGG	AGATCCAGCT	CGCCGTTCA	GCAACACCCA	CCTTCAGCTC	1080

CAGCGCCGGC	GCCGACCCGA	CGGCCCTAAA	CGGGATGCCG	GCAGGCCCTGC	TCAGCGGGAT	1140
GGCTTTGGCG	AGCCTGGCCG	CACGCGGCAC	GACGGGCGGT	GGCGGCACCC	GTAGCGGCAC	1200
CAGCACTGAC	GGCCAAGAGG	ACGGCCGCAA	ACCCCCGGTA	GTTGTGATTA	GAGAGCAGCC	1260
GCCGCCCGGA	AACCCCCCGC	GGTAAAAGTC	CGGCAACCCT	TCGTCGCCGC	GCGGAAAATG	1320
CCTGGTGAGC	GTGGCTATCC	GACGGGCCGT	TCACACCGCT	TGTAGTAGCG	TACGGCTATG	1380
GACGACGGTG	TCTGGATTCT	CGGCGGCTAT	CAGAGCGATT	TTGCTCGCAA	CCTCAGCAAA	1440
G						1441

## (2) INFORMATION FOR SEQ ID NO:141:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 99 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

Met	Ser	Phe	Val	Thr	Ile	Gln	Pro	Val	Val	Leu	Ala	Ala	Ala	Thr	Gly
1		5				10								15	
Asp	Leu	Pro	Thr	Ile	Gly	Thr	Ala	Val	Ser	Ala	Arg	Asn	Thr	Ala	Val
				20			25							30	
Cys	Ala	Pro	Thr	Thr	Gly	Val	Leu	Pro	Pro	Ala	Ala	Asn	Asp	Val	Ser
					35		40							45	
Val	Leu	Thr	Ala	Ala	Arg	Phe	Thr	Ala	His	Thr	Lys	His	Tyr	Arg	Val
					50		55							60	
Val	Ser	Lys	Pro	Ala	Ala	Leu	Val	His	Gly	Met	Phe	Val	Ala	Leu	Pro
					65		70							80	
Ala	Ala	Thr	Ala	Asp	Ala	Tyr	Ala	Thr	Thr	Glu	Ala	Val	Asn	Val	Val
					85			90						95	
Ala	Thr	Gly													

## (2) INFORMATION FOR SEQ ID NO:142:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 423 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

Met	Asp	Phe	Gly	Leu	Leu	Pro	Pro	Glu	Val	Asn	Ser	Ser	Arg	Met	Tyr
1		5				10								15	
Ser	Gly	Pro	Gly	Pro	Glu	Ser	Met	Leu	Ala	Ala	Ala	Ala	Trp	Asp	
				20			25							30	
Gly	Val	Ala	Ala	Glu	Leu	Thr	Ser	Ala	Ala	Val	Ser	Tyr	Gly	Ser	Val
					35		40							45	
Val	Ser	Thr	Leu	Ile	Val	Glu	Pro	Trp	Met	Gly	Pro	Ala	Ala	Ala	Ala
					50		55							60	
Met	Ala	Ala	Ala	Ala	Thr	Pro	Tyr	Val	Gly	Trp	Leu	Ala	Ala	Thr	Ala
					65		70							80	
Ala	Leu	Ala	Lys	Glu	Thr	Ala	Thr	Gln	Ala	Arg	Ala	Ala	Glu	Ala	
					85			90						95	
Phe	Gly	Thr	Ala	Phe	Ala	Met	Thr	Val	Pro	Pro	Ser	Leu	Val	Ala	Ala
						100		105						110	
Asn	Arg	Ser	Arg	Leu	Met	Ser	Leu	Val	Ala	Ala	Asn	Ile	Leu	Gly	Gln
					115		120							125	

Asn Ser Ala Ala Ile Ala Ala Thr Gln Ala Glu Tyr Ala Glu Met Trp  
     130                       135                       140  
 Ala Gln Asp Ala Ala Val Met Tyr Ser Tyr Glu Gly Ala Ser Ala Ala  
     145                       150                       155                       160  
 Ala Ser Ala Leu Pro Pro Phe Thr Pro Pro Val Gln Gly Thr Gly Pro  
     165                       170                       175  
 Ala Gly Pro Ala Ala Ala Ala Ala Thr Gln Ala Ala Gly Ala Gly  
     180                       185                       190  
 Ala Val Ala Asp Ala Gln Ala Thr Leu Ala Gln Leu Pro Pro Gly Ile  
     195                       200                       205  
 Leu Ser Asp Ile Leu Ser Ala Leu Ala Asn Ala Asp Pro Leu Thr  
     210                       215                       220  
 Ser Gly Leu Leu Gly Ile Ala Ser Thr Leu Asn Pro Gln Val Gly Ser  
     225                       230                       235                       240  
 Ala Gln Pro Ile Val Ile Pro Thr Pro Ile Gly Glu Leu Asp Val Ile  
     245                       250                       255  
 Ala Leu Tyr Ile Ala Ser Ile Ala Thr Gly Ser Ile Ala Leu Ala Ile  
     260                       265                       270  
 Thr Asn Thr Ala Arg Pro Trp His Ile Gly Leu Tyr Gly Asn Ala Gly  
     275                       280                       285  
 Gly Leu Gly Pro Thr Gln Gly His Pro Leu Ser Ser Ala Thr Asp Glu  
     290                       295                       300  
 Pro Glu Pro His Trp Gly Pro Phe Gly Gly Ala Ala Pro Val Ser Ala  
     305                       310                       315                       320  
 Gly Val Gly His Ala Ala Leu Val Gly Ala Leu Ser Val Pro His Ser  
     325                       330                       335  
 Trp Thr Thr Ala Ala Pro Glu Ile Gln Leu Ala Val Gln Ala Thr Pro  
     340                       345                       350  
 Thr Phe Ser Ser Ser Ala Gly Ala Asp Pro Thr Ala Leu Asn Gly Met  
     355                       360                       365  
 Pro Ala Gly Leu Leu Ser Gly Met Ala Leu Ala Ser Leu Ala Ala Arg  
     370                       375                       380  
 Gly Thr Thr Gly Gly Gly Thr Arg Ser Gly Thr Ser Thr Asp Gly  
     385                       390                       395                       400  
 Gln Glu Asp Gly Arg Lys Pro Pro Val Val Val Ile Arg Glu Gln Pro  
     405                       410                       415  
 Pro Pro Gly Asn Pro Pro Arg  
     420

## (2) INFORMATION FOR SEQ ID NO:143:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 97 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

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

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

Phe

(2) INFORMATION FOR SEQ ID NO:144:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 99 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

Cys Arg Leu Cys Leu Asp Ser His Leu Arg Val Val Ala Leu Pro Ala  
   1              5                  10                  15  
 Gly Gln Pro Gly Arg Leu Val Gln Ala Ile Gly Pro Ala Gln Glu Arg  
   20              25                  30  
 Asp Val Gly Gln Thr Arg Cys Thr Arg Thr Gly Leu Asp Xaa Val Ser  
   35              40                  45  
 Ala Leu Thr Ala Ala Gln Phe Ala Ala His Ala Gln Ile Tyr Gln Ala  
   50              55                  60  
 Val Ser Ala Gln Ala Ala Ala Ile His Glu Met Phe Val Asn Thr Leu  
   65              70                  75                  80  
 Gln Xaa Xaa Ser Gly Ser Tyr Ala Ala Thr Glu Ala Ala Asn Ala Ala  
   85              90                  95

Ala Ala Gly