(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。Int. Cl. 7 C07K 14/47	(11) 공개번호 특2002-0089352 (43) 공개일자 2002년11월29일

---

(21) 출원번호	10-2002-7010630		
(22) 출원일자	2002년08월14일		
번역문 제출일자	2002년08월14일		
(86) 국제출원번호	PCT/FR2001/00461	(87) 국제공개번호	WO 2001/60857
(86) 국제출원출원일자	2001년02월15일	(87) 국제공개일자	2001년08월23일

---

(81) 지정국

국내특허 : 아랍에미리트, 안티구아바부다, 알바니아, 아르메니아, 오스트리아, 오스트레일리아, 아제르바이잔, 보스니아-헤르체고비나, 바베이도스, 불가리아, 브라질, 벨라루스, 벨리즈, 캐나다, 스위스, 중국, 코스타리카, 쿠바, 체코, 독일, 덴마크, 도미니카연방, 알제리, 에스토니아, 스페인, 핀랜드, 영국, 그레나다, 그루지야, 가나, 감비아, 크로아티아, 헝가리, 인도네시아, 이스라엘, 인도, 아이슬란드, 일본, 케냐, 키르기즈, 북한, 대한민국, 카자흐스탄, 세인트루시아, 스리랑카, 라이베리아, 레소토, 리투아니아, 룩셈부르크, 라트비아, 모로코, 몰도바, 마다가스카르, 마케도니아, 몽고, 말라위, 멕시코, 모잠비크, 노르웨이, 뉴질랜드, 폴란드, 포르투칼, 루마니아, 러시아, 수단, 스웨덴, 싱가포르, 슬로베니아, 슬로바키아, 시에라리온, 타지키스탄, 투르크메니스탄, 터어키, 트리니다드토바고, 탄자니아, 우크라이나, 우간다, 미국, 우즈베키스탄, 베트남, 유고슬라비아, 남아프리카, 짐바브웨,

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

EA 유라시아특허: 아르메니아, 아제르바이잔, 벨라루스, 키르기즈, 카자흐스탄, 몰도바, 러시아, 타지키스탄, 투르크메니스탄,

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

OA OAPI특허: 부르키나파소, 베넹, 중앙아프리카, 콩고, 코트디브와르, 카메룬, 가봉, 기네, 말리, 모리타니, 니제르, 세네갈, 차드, 토고, 기네비쏘,

(30) 우선권주장	00/01,980	2000년02월17일	프랑스(FR)
	60/198,489	2000년04월18일	미국(US)

(71) 출원인

아방티 파르마 소시에테 아노님
프랑스 앙토니 에프-92160 아브뉴 레이몽 아롱 20
엥스띠뛰 나시오날 드 라 상트 에 드 라 러쉐르쉐 메디칼
프랑스 파리 세덱스 13 에프-75654 뤼 드 톨비악 101

(72) 발명자 코우트니코바하나
프랑스에프-67100스트라스부르뤼드라그룻소35
브리스알렉시스
프랑스에프-75006파리뤼생앙드레데아르스51
푸르니에알랭
프랑스에프-92290샤트네이말라브리아브뉴로제살렝그로28
프라디에로랑
프랑스에프-91370베리에르아브뉴캉바세레23
프라데카트린
프랑스에프-94320티에아브뉴드제네랄드골30
아르눌드-르기녜이자벨
프랑스에프-94430셴베르쉬르마른뤼드브리112
로제-몽퉈마리-프랑수와즈
프랑스에프-92160앙토니뤼데바코네21
코르티올가
프랑스에프-75013파리뤼드셰발레르151

(74) 대리인 차윤근

심사청구 : 없음

---

(54) 파킨 유전자 활성의 조절에 유용한 조성물

요약

본 발명은 신규 화합물 및 특히 약학적, 진단적 또는 약리학적 표적으로서 이의 용도에 관한 것이다. 좀더 구체적으로는, 본 발명은 PAP1으로 불리는 신규 단백질, 및 적어도 부분적으로 파킨 유전자 활성을 조절할 수 있는 신규 펩타이드 또는 조성물에 관한 것이다.

대표도

도 1

색인어

parkin 유전자, PAP1

명세서

기술분야

본 발명은 파킨(parkin)의 활성 조절에 사용될 수 있는 조성물 및 방법에 관한 것이다. 본 발명은 특히 파킨의 파트너인 신규 단백질(PAP1으로 언급)과 이 단백질로부터 유도되거나 이 단백질과 상동한 펩타이드 또는 폴리펩타이드에 관한 것이다. 본 발명은 또한 파킨의 활성을 적어도 부분적으로 조절할 수 있는 화합물, 특히 파킨과 PAP1 간의 상호작용을 방해할 수 있는 화합물에 관한 것이다. 본 발명은 치료 또는 진단 영역에 사용될 수 있거나, 신약 개발을 가능하게 하는 약리학적 표적을 형성하기 위해 사용될 수 있다.

배경기술

파킨 유전자는 파킨슨병의 특정 가족형(상염색체 열성 청소년)으로 돌연변이된다(Kitada 등, 1998). 파킨슨병(Lewy, 1912)은 55세 이상인 집단의 1% 이상에 영향을 미치는 가장 통상적인 신경퇴행성 질환 중 하나이다. 이 질환을 앓고 있는 환자는 경직, 운동완만 및 휴면 중의 불안정으로 특징지워지는 용어 " 파킨슨 증후군" 으로 함께 분류되는 신경학적 장애를 가진다. 이러한 증상은 뇌 흑색질의 도파민성 뉴론의 퇴행 결과이다.

파킨슨병을 앓고 있는 대부분의 사례는 가족성 병력을 가지지 않는다. 그러나, 질환의 단성형에 상응하는 특정 가족성 사례가 존재한다. 현재, 상이한 3개의 유전자만이 희귀한 특정 유전성 유형에서 동정되었다. 첫번째 유형은 원인 유전자가 알파 시누클레인(Synuclein)을 암호화하는 상염색체 우성 유형에 상응한다(Polymeropoulos 등, 1997). 이 단백질은 파킨슨병에 대한 마커로 사용되는 세포질내 봉입체(루이체로 불림)의 풍부한 구성물이다(Lewy, 1912). 두번째 유형(이 또한 상염색체 우성)은 유비퀴틴 카복시-말단 하이드롤라제 L1으로 불리는 하이드롤라제를 암호화하는 유전자에서의 돌연변이와 관련있다(Leroy 등, 1998). 이 효소는 유비퀴틴 중합체 또는 접합체를 유비퀴틴 단량체로 가수분해시키는 것으로 추정된다. 세번째 유형은 상염색체 열성 전달되고 종종 40세 전에 발병하며 루이체가 부재한다는 점에서 상기의 유형과 상이하다. 이들 질환은 파킨슨병의 치료에 사용되는 도파민의 전구체인 레보도파에 더욱 활발하게 반응한다. 이 유형에 관여하는 유전자는 파킨으로 불리는 신규 단백질을 암호화한다(Kitada 등, 1998).

파킨 유전자는 염색체 6의 500,000개 이상의 염기쌍의 게놈 영역(6q25.2-q27)을 커버링하는 12개의 엑손으로 이루어진다. 현재, 질환의 초기에 이 유전자의 두 가지 유형의 주요 돌연변이가 알려져 있다; 엑손 2 내지 9를 커버링하는 영역에서 가변적인 크기의 결실, 또는 정지 코돈의 이른 출현 또는 아미노산의 변화를 제공하는 점 돌연변이(Kitada 등, 1998; Abbas 등, 1999; Lucking 등, 1998; Hattori 등, 1998). 이러한 돌연변이의 특성 및 상염색체 열성 전달 방법이 파킨슨병을 유발하는 파킨의 기능 저하를 설명해준다.

이 유전자는 다수의 조직에서, 특히 흑색질에서 발현된다. 이 유전자에 상응하고, 상이한 대체 스플라이싱으로부터 기원하는 다수의 전사체가 존재한다(Kitada 등, 1998; Sunada 등, 1998). 뇌에서, 두 가지 유형의 메신저 RNA가 발견되는데, 이들 중 하나가 엑손 5에 상응하는 부분이 결여되었다. 백혈구에서, 엑손 3, 4 및 5를 암호화하는 영역을 포함하지 않는 파킨 메신저 RNA가 동정되었다. 뇌에 존재하는 가장 긴 파킨 메신저 RNA는 2960개의 염기를 포함하고 465개 아미노산의 단백질을 암호화한다.

이 단백질은 그것의 N-말단 부분에서 유비퀴틴과 약한 상동성을 가진다. 이의 절반 C-말단은 징크 핑거 도메인(zinc finger domain)과 마찬가지로 시스테인-풍부 영역에 상응하고 금속과 결합할 수 있는, IBR(In Between Ring) 도메인에 의해 분리되는 두 개의 링 핑거 모티프(ring finger motif)를 포함한다(Morett, 1999). 파킨이 세포질 및 멜라민을 함유하는 흑색질 뉴론의 골지체에 위치한다(Shimura 등, 1999)는 것이 면역세포화학에 의해 알려졌다. 또한, 이 단백질은 파킨슨병 환자의 특정 루이체에 존재한다. 파킨의 세포내 기능이 아직 설명되지는 않았지만, 연접소포, 단백질의 성숙 또는 퇴행 및 세포 성장, 분열 및 진화의 조절에서 트랜스포터 역할을 할 수 있다. 상염색체 열성 청소년형에서, 파킨은 부재하고, 이에 따라 이 기능의 저하가 질환의 원인이 된다는 것이 확인된다.

이에 따라 도파민성 뉴론의 퇴행 과정에서 파킨 단백질 역할의 정확한 설명이 파킨슨병, 좀더 일반적으로는 중추신경계 질환의 이해와 이에의 치료적 접근을 위한 주요 이점을 구성한다.

본 발명은 생리학적 조건하에서 이 단백질과 상호작용하는 파킨의 파트너의 동정에 있다. 이 파트너는 파킨의 활성, 특히 도파민성 뉴론의 퇴행 및/또는 신경병의 발병에서 활성을 조절할 수 있는 화합물을 제조하거나 조사하기 위한 신규 약리학적 표적을 나타낸다. 이 단백질, 항체, 상응 핵산 및 특정 탐침 또는 프라이머 또한 생물학적 샘플, 특히 신경 조

직 샘플에서 단백질을 검출하거나 분석하기 위해 사용될 수 있다. 이들 단백질 또는 핵산과, 파킨과 본 발명의 폴리펩타이드 간의 상호작용을 조절할 수 있는 본 발명에 따른 임의 화합물 또한 파킨의 활성을 조절하기 위한 치료적 접근법에서 사용될 수 있다.

좀더 구체적으로 말하면 본 발명은 PAP1(파킨 수반 단백질 1) 또는 LY111으로 언급되고, 파킨과 상호작용하는 신규 인간 단백질의 본 출원인에 의한 설명으로부터 초래된다. PAP1 단백질(서열번호 1 또는 2의 서열)은 시냅토태그민(synaptotagmin)과 어느 정도 상동성을 보이고 좀더 구체적으로 말하면 파킨의 중심 영역(서열번호 3 또는 4의 서열로 표시)과 상호작용할 수 있다. PAP1 단백질은 또한 인간 기원의 다양한 조직, 특히 폐(서열번호 12 및 13) 및 뇌(서열번호 42 및 43)를 사용하여 클로닝되고, 서열분석되며 특징규명되는데, 그 이유는 스플라이싱 변이체(서열번호 14, 15, 44 및 45)에 상응하는 짧은 형태를 가지기 때문이다.

본 발명은 또한 파킨의 기능 조정에 관여하는 PAP1 단백질의 특정 영역의 동정 및 특징규명으로부터 초래된다. 이 단백질의 존재 및 이의 기능에 관여하는 영역의 설명이 특히 약제로 사용될 수 있는 신규 화합물 및/또는 조성물의 제조와 이러한 화합물의 산업적 스크리닝 방법 개발을 가능하게 한다.

따라서 본 발명의 첫번째 주제는 PAP1 단백질(또는 이의 상동체)과 파킨(특히 인간 파킨) 간의 상호작용을 적어도 부분적으로 조절할 수 있거나, 이들 단백질 간의 상호작용을 방해할 수 있는 화합물에 관한 것이다.

본 발명의 또다른 주제는 PAP1 단백질과 이의 단편, 유도체 및 상동체로 이루어진다.

본 발명의 또다른 일면은 PAP1 단백질 또는 이의 단편, 유도체 및 상동체를 암호화하는 핵산과, 이러한 핵산을 포함하는 임의의 벡터, 이러한 핵산 또는 벡터를 함유하는 임의의 재조합 세포 및 세포에 이러한 핵산을 포함하는 임의의 비인간 포유동물로 구성된다.

본 발명은 또한 PAP1 단백질과 이의 단편, 유도체 및 상동체와 결합할 수 있는 항체, 특히 폴리클로날 또는 모노클로날 항체, 더욱 바람직하게는 PAP1 단백질과 결합할 수 있고 파킨과의 상호작용을 적어도 부분적으로 억제할 수 있는 항체에 관한 것이다.

본 발명의 또다른 일면은 PAP1에 대해 특이적이고 pap1 유전자 또는 이 유전자의 영역을 임의의 생물학적 샘플에서 검출하거나 증폭시키는 데 사용될 수 있는 뉴클레오타이드 탐침 또는 프라이머에 관한 것이다.

본 발명은 또한 유전자 이상을 검출하기 위한 약학 조성물 및 방법, 전술한 바와 같은 폴리펩타이드 생산방법 및 활성 화합물을 스크리닝하거나 특징규명하기 위한 방법에 관한 것이다.

전술한 바와 같이, 본 발명의 첫번째 일면은 PAP1 단백질(또는 이의 상동체)과 파킨 간의 상호작용을 적어도 부분적으로 방해할 수 있는 화합물로 이루어진다.

본 발명의 목적상, 용어 PAP1 단백질은 단백질 자체와, 이의 모든 상동형을 언급한다. " 상동형" 는 다양한 세포 기원, 특히 인간 기원의 세포 또는 다른 유기체로부터 유도되는 고려 중인 단백질과 등가이고, 동일한 유형의 활성을 지니는 임의의 단백질을 언급하는 것으로 이해된다. 이러한 상동체는 또한 서열번호 2 서열의 PAP1 단백질의 자연 변이체, 특히 다형성 또는 스플라이싱 변이체도 포함한다. 이러한 상동체는 암호화 핵산(특히 서열번호 1 서열의 핵산) 간의 하이브리드화 실험에 의해 수득될 수 있다. 본 발명의 목적상, 이러한 유형의 서열만이 청구되고 있는 PAP1 단백질의 그것에 필적하는 생리학적 행동을 유도하는 상당한 퍼센티지의 동일성을 가져야만 한다. 표현 " 상당한 퍼센티지의 동일성" 은 적어도 60%, 바람직하게는 80%, 더욱 바람직하게는 90%, 훨씬 더 바람직하게는 95%의 퍼센티지를 의미하는 것으로 이해된다. 이러한 측면에서, 서열번호 2 서열의 변이체 및/또는 상동체가 서열번호 13, 15, 43 및 45의 서열로 기술되고, 인간 기원의 조직을 사용하여 동정된다. 따라서 용어 PAP1은 또한 이러한 폴리펩타이드도 포함한다.

본 발명의 목적상, 두 뉴클레오타이드 또는 아미노산 서열 간의 " 동일성 퍼센티지" 는 비교창을 통해 최적으로 정렬된 두 개의 서열을 비교함으로써 측정될 수 있다.

따라서 두 서열의 최적의 정렬을 달성하기 위해 비교창에서 뉴클레오타이드 또는 폴리펩타이드 서열의 일부에 참조 서열(이러한 첨가 또는 결실이 일어나지 않음)과 비교하여 첨가 또는 결실(예를 들면, gaps)이 일어날 수 있다.

퍼센티지는 비교되는 두 서열(핵산 또는 펩타이드 서열)에 대해 동일한 핵산 염기 또는 아미노산 잔기가 관찰되는 위치의 수를 측정한 다음, 두 염기 또는 아미노산 잔기 사이에 동일한 위치의 수를 비교창의 전체 위치 수로 나누고, 이어서 서열 동일성의 퍼센티지를 수집하기 위해 결과에 100을 곱함으로써 계산된다.

비교를 위한 서열의 최적의 정렬은 컴퓨터에서 WISCONSIN GENETICS SOFTARE PACKAGE, GENETICS COMPUTER GROUP (GCG), 575 Science Doctor, Madison, WISCONSIN에 포함되어 있는 공지 알고리듬을 사용하여 제공될 수 있다.

설명의 방법으로, 서열 동일성의 퍼센티지는 BLAST 프로그램(1996년 3월의 BLAST 1.4.9, 1998년 2월의 BLAST 2.0.4 및 1998년 9월의 BLAST 2.0.6 버전)과 전용 디폴트 파라미터(Altschul 등, J. Mol. Biol., (1990) 215: 403-410; Altschul 등, Nucleic Acids Res. (1997) 25: 3389-3402)를 사용하여 제공될 수 있다. Blast는 Altschul 등(상기)의 알고리듬을 사용하여 " 청구" 참조 서열과 유사한/상동성인 서열에 대해 연구하였다. 사용되는 청구 서열 및 데이터베이스는 기본 펩타이드 또는 핵산일 수 있고, 임의의 조합도 가능하다.

본 발명에 따른 화합물의 방해는 다양한 방법으로 자체가 밝혀질 수 있다. 따라서, 화합물은 PAP1 단백질 또는 이의 상동형과 파킨 간의 상호작용을 적어도 부분적으로 느리게 하거나, 억제하거나 자극할 수 있다. 바람직하게는, 이들은 시험관내에서, 예를 들면 이중-하이브리드 유형 시스템 또는 두 폴리펩타이드 간의 상호작용을 검출하기 위한 무세포 시스템에서 이 상호작용을 조절할 수 있는 화합물이다. 본 발명에 따른 화합물은 바람직하게는 이 상호작용을 적어도 부분적으로 조절할 수 있고, 바람직하게는 이 반응을 화합물 부재시의 대조와 비교하여, 적어도 20%, 더욱 바람직하게는 적어도 50% 증가시키거나 억제할 수 있는 화합물이다.

특정 양태에서, 이들은 서열번호 4의 서열로 표시되는 파킨의 영역과 서열번호 2, 13, 15, 43 또는 45의 서열로 표시되는 PAP1 단백질의 영역 간의 상호작용을 방해할 수 있는 화합물이다.

본 발명의 특정 양태에 따라서, 화합물은 PAP1 단백질, 또는 이의 상동형과 파킨 간의 상호작용의 도메인 수준에서 결합할 수 있다,

본 발명에 따른 화합물은 특성과 기원에 있어서 다양할 수 있다. 특히, 이들은 펩타이드, 핵산(즉, 염기, 특히 DNA 또는 RNA 분자의 스트링을 포함), 지질 또는 사카라이드 유형, 항체 등, 좀더 일반적으로는 임의 유기 또는 무기 분자의 화합물일 수 있다.

첫번째 변형에 따라서, 본 발명의 화합물은 특성에 있어서 펩타이드이다. 용어 " 펩타이드" 는 필요시 다른 화합물 또는 화학 그룹으로 변형되거나 이들과 배합되는 예를 들면, 펩타이드, 폴리펩타이드, 단백질 또는 항체(또는 항체 단편 또는 유도체)와 같은 아미노산의 스트링을 포함하는 임의의 분자를 언급한다. 이러한 측면에서, 용어 " 펩타이드" 는 좀더 구체적으로 말하면 많아야 50개의 아미노산, 더욱 바람직하게는 많아야 40개 아미노산의 스트링을 포함하는 분자를 언급한다. 폴리펩타이드(또는 단백질)는 바람직하게는 50 내지 500개, 또는 그 이상의 아미노산을 포함한다.

첫번째 바람직한 양태에 따라서, 본 발명의 화합물은 서열번호 2 서열의 전체 혹은 일부 펩타이드 또는 이의 유도체, 특히 서열번호 13, 15, 43 또는 45 서열의 전체 혹은 일부 펩타이드 또는 이들 서열의 유도체, 좀더 구체적으로 말하면 서열번호 2, 13, 15, 43 또는 45의 서열을 포함하는 PAP1 단백질을 포함하는 펩타이드 화합물이다.

본 발명의 목적상, 용어 " 유도체" 는 유전 암호의 축퇴성 때문에 고려 중인 서열과 상이하고, 유전적 및/또는 화학적 특성의 한 가지 이상의 변화에 의해 수득되는 임의의 서열과, 서열번호 1 서열의 핵산 또는 이 서열의 단편, 예를 들면 서열번호 12, 14, 42 또는 44 서열의 핵산 또는 이들 서열의 단편과 하이브리드화되는 서열에 의해 암호화되고 PAP1 단백질 또는 이의 상동체와 파킨 간의 상호작용을 방해할 수 있는 임의의 펩타이드를 언급한다. " 유전적 및/또는 화학적 특성의 변화" 는 하나 이상의 잔기의 돌연변이, 치환, 결실, 첨가 및/또는 변형을 의미하는 것으로 이해될 수 있다. 용어 " 유도체" 는 또한 고려 중인 서열과 상동하고, 다른 세포원, 특히 인간 기원의 세포 또는 다른 유기체로부터 유도되며, 동일한 유형의 활성을 지니는 서열을 포함한다. 이러한 상동 서열은 하이브리드화 실험에 의해 수득될 수 있다. 하이브리드화는 핵산 라이브러리로, 탐침으로서 네이티브 서열 또는 이 서열의 단편을 사용하여 다양한 하이브리드화 조건하에서 수행될 수 있다(Maniatis 등, 1989). 또한, 용어 " 단편" 또는 " 부분" 은 적어도 5개의 연속 잔기, 바람직하게는 적어도 9개의 연속 잔기, 훨씬 더 바람직하게는 적어도 15개의 연속 잔기를 포함하는 고려 중인 분자의 일부를 언급한다. 통상의 단편은 적어도 25개의 연속 잔기를 포함할 수 있다.

이러한 유도체 또는 단편은 상이한 목적, 특히 예를 들면, 치료 효능을 증가시키거나 부작용을 감소시키는 목적 또는 신규한 약력학적 및/또는 생물학적 성질을 부여하는 목적으로 생산될 수 있다.

PAP1 단백질 및 이의 상동형으로부터 유도되는 펩타이드로서 특히 파킨과 상호작용할 수 있는 임의의 펩타이드가 언급될 수 있지만, 비기능성으로 만드는 이펙터 영역을 지니는 펩타이드도 언급될 수 있다. 이러한 펩타이드는 PAP1 단백질 및 상동형에서 이 이펙터 영역의 결실, 돌연변이 또는 분열에 의해 수득될 수 있다. 이러한 변형은 예를 들면, 시험관내 돌연변이 유발에 의해, 부가의 요소 또는 합성 서열을 도입함으로써 또는 원래 요소의 결실 또는 치환에 의해 수행될 수 있다. 전술한 바와 같은 이러한 유도체가 제조되면, 파킨의 결합 부위에서 PAP1 단백질 및 상동형 결합의 부분 억제제로서의 이의 활성이 설명될 수 있다. 물론 당업자에게 알려져 있는 임의 기술이 사용될 수 있다.

이들은 또한 전술한 서열의 단편일 수 있다. 이러한 단편은 다양한 방법으로 생산될 수 있다. 좀더 구체적으로 말하면 이들은 당업자에게 알려져 있는 펩타이드 합성기를 사용하여, 본 출원에서 주어진 서열을 기초로 하여 화학적으로 합성될 수 있다. 이들은 또한 목적하는 펩타이드를 암호화하는 뉴클레오타이드 서열의 숙주 세포에서의 발현에 의해 유전적으로 합성될 수도 있다. 이 경우에, 뉴클레오타이드 서열은 본 출원에서 주어진 펩타이드 서열과 유전 암호를 기초로 하여 올리고뉴클레오타이드 합성기를 사용하여 화학적으로 제조될 수 있다. 뉴클레오타이드 서열은 또한 당업자에게 알려져 있는 기술에 따라 효소 절단, 리게이션, 클로닝 등에 의해 또는 이러한 서열로부터 개발된 탐침으로 DNA 라이브러리를 스크리닝함으로써 본 출원에서 주어진 서열로부터 제조될 수 있다.

또한, 본 발명의 펩타이드, 즉 PAP1 단백질 및 이의 상동형과 파킨 간의 상호작용을 적어도 부분적으로 조절할 수 있는 펩타이드는 또한 PAP1 단백질 및 이의 상동형의 파킨에서의 상호작용 부위에 상응하는 서열을 가지는 펩타이드일 수 있다.

본 발명에 따른 다른 펩타이드는 이들의 세포성 표적과 상호작용하기 위해 전술한 펩타이드와 경쟁할 수 있는 펩타이드이다. 이러한 펩타이드는 특히 고려 중인 펩타이드의 서열을 기초로 하여 합성될 수 있고, 전술한 펩타이드와의 경쟁능이 측정될 수 있다.

본 발명의 특정 주제는 PAP1 단백질에 관한 것이다. 좀더 구체적으로 말하면 서열번호 2의 서열 또는 이 서열의 단편 또는 유도체를 포함하는 PAP1 단백질, 예를 들면 서열번호 13, 15, 43 또는 45의 서열 또는 이들 서열의 단편을 가지는 PAP1 단백질이다.

본 발명의 또다른 주제는 전술한 폴리펩타이드에 대한 폴리클로날 또는 모노클로날 항체 또는 항체 단편이나 유도체로 이루어진다. 이러한 항체는 당업자에게 알려져 있는 방법으로 생산될 수 있다. 특히, 이러한 항체는 동물을 본 발명의 펩타이드 화합물(특히 서열번호 2의 전체 혹은 일부 서열을 포함하는 폴리펩타이드 또는 펩타이드)에 대해 면역화시키고, 혈액을 샘플링한 다음, 항체를 분리함으로써 제조될 수 있다. 이러한 항체는 또한 당업자에게 알려져 있는 기술에 따라 하이브리도마를 제조함으로써 생산될 수도 있다.

더욱 바람직하게는, 본 발명의 항체 또는 항체 단편은 청구된 펩타이드의 파킨과의 상호작용을 적어도 부분적으로 조절하는 능력을 가진다.

또한, 이러한 항체는 또한 생물학적 샘플에서 PAP1의 발현을 검출하고/하거나 분석하고, 이에 따라 활성 상태에 대한 정보를 제공하는 데 사용될 수 있다.

항체 단편 또는 유도체는 예를 들면, Fab 또는 Fab'2 단편, 일본쇄 항체(ScFv) 등이다. 이들은 특히 이들이 유도되는 항체의 항원 특이성을 보존하는 임의의 단편 또는 유도체이다.

본 발명에 따른 항체는 더욱 바람직하게는 서열번호 2, 13, 15, 43 또는 45의 서열을 포함하는 PAP1 단백질, 특히 파킨과의 상호작용에 관여하는 이 단백질의 영역과 결합할 수 있다. 이들 항체(또는 단편이나 유도체)는 더욱 바람직하게는 서열번호 2 서열의 잔기 1 내지 344 사이에서 서열에 존재하는 에피토프와 결합할 수 있다.

본 발명은 또한 펩타이드가 아니고 배타적으로 펩타이드가 아니며, 약제로 사용될 수 있는 화합물에 관한 것이다. 실제로, 본 출원에 기술되어 있는 활성 단백질 모티프로부터 PAP1의 활성 조절제이고, 배타적으로 펩타이드가 아니며 약학적 용도와 상용성인 분자를 특히 펩타이드가 아니거나 배타적으로 펩타이드의 특성이 아닌 구조를 가지는 펩타이드의 활성 모티프를 복제함으로써 제조하는 것이 가능하다.

본 발명의 주제는 또한 본 발명에 따른 펩타이드 화합물을 암호화하는 임의의 핵산이다. 좀더 구체적으로 말하면 서열번호 1, 12, 14, 42 또는 44의 전체 혹은 일부 서열을 포함하는 핵산 또는 이의 유도체일 수 있다. " 유도된 서열" 은 본 발명의 목적상, 서열번호 1로 표시되는 서열 또는 이 서열의 단편과 하이브리드화되고 본 발명에 따른 펩타이드 화합물을 암호화하는 임의의 서열 및 유전 암호의 축퇴성으로 인해 후자로부터 초래되는 서열을 의미하는 것으로 이해된다. 본 발명에 따른 핵산은 예를 들면, 서열번호 12, 14, 42 또는 44의 전체 혹은 일부 핵산 서열을 포함한다.

본 발명은 또한 서열번호 1로 표시되는 서열 또는 이 서열의 단편과 상당한 퍼센티지의 동일성을 가지고, PAP1 단백질의 그것과 유사한 생리학적 행동을 보이는 펩타이드 화합물을 암호화하는 서열에 관한 것이다. 표현 " 상당한 퍼센티지의 동일성" 은 적어도 60%, 바람직하게는 80%, 더욱 바람직하게는 90%, 훨씬 더 바람직하게는 95%의 퍼센티지를 의미하는 것으로 이해된다.

본 발명의 다양한 뉴클레오타이드 서열은 인공 기원일 수 있거나 아닐 수도 있다. 이들은 게놈, cDNA 또는 RNA 서열, 하이브리드 서열 또는 합성 또는 반합성 서열일 수 있다. 이들 서열은 DNA 라이브러리(cDNA 라이브러리, 게놈 DNA 라이브러리)를 스크리닝함으로써, 화학적 합성에 의해, 라이브러리의 스크리닝에 의해 수득되는 서열의 화학적 또는 효소적 변화를 포함하는 혼합된 방법에 의해, 또는 핵산 또는 단백질 데이터베이스에서 상동성을 조사함으로써 수득될 수 있다. 전술한 하이브리드화는 바람직하게는 Sambrook 등(1989, pages 9.52 - 9.55)에 의해 기술된 조건하에서 수행된다.

이것은 유리하게는 매우 엄격한 하이브리드화 조건하에서 수행된다. 본 발명의 목적상, 표현 " 매우 엄격한 하이브리드화 조건" 은 하기의 조건을 의미하는 것으로 이해될 것이다:

1 - 막 경쟁 및 예비하이브리드화:

- 혼합: 연어 정액 DNA 40 ㎕(10 mg/㎖)

+ 인간 태반 DNA 40 ㎕(10 mg/㎖)

- 96℃에서 5분간 변성시킨 다음 혼합물을 얼음으로 플런징하고,

- 2X SSC 완충액을 제거한 다음 포름아미드 혼합물 4 ㎖를 막을 함유하는 하이브리드화 관에 붓고,

- 두 변성 DNA의 혼합물을 첨가하며,

- 42℃에서 5 내지 6시간 동안 회전시키면서 배양한다.

2 - 표지된 탐침 경쟁:

- 비특이적 하이브리드화 양에 따라 Cot 1 DNA 10 내지 50 ㎕를 표지되고 정제된 탐침에 첨가하고,

- 95℃에서 7 내지 10분간 변성시킨 다음,

- 65℃에서 2 내지 5시간 동안 배양한다.

3 - 하이브리드화:

- 예비하이브리드화 혼합물을 제거한다.

- 40 ㎕의 연어 정액 DNA + 40 ㎕의 인간 태반 DNA 혼합물; 5분 동안 96℃에서 변성시킨 후 얼음에 찔러둔다.

- 4 ml의 포름아미드 혼합물, 두 DNA 및 표지된 탐침/변성된 Cot 1 DNA의 혼합물을 하이브리드화 관에 첨가한다.

- 15 내지 20시간 동안 42℃에서 회전하며 배양한다.

4 - 세척:

- 주위 온도에서 2X SSC로 1회 세척하여 세정한다.

- 주위 온도에서 2X SSC와 0.1% SDS로 5분 동안 2회 세척한다.

- 0.1X SSC와 0.1% SDS로 65℃에서 15분 동안 2회 세척한다.

Saran으로 막을 싸고 노출시킨다.

상기 기술된 하이브리드화 조건은 고 엄중 조건하에서, 다양한 길이의 20 뉴클레오타이드와 수백개의 뉴클레오타이드로의 핵산 분자의 하이브리드화에 적당하다.

상기 기술된 하이브리드화 조건은 하이브리드화가 조사된 핵산의 길이 또는 당업자에게 공지된 기술에 따라 선택된 표지의 유형의 함수에 따라 조정될 수 있음은 물론이다.

적당한 하이브리드화 조건은, 예를 들면 HAMES 및 HIGGINS(1985)의 연구(Nucleic acid Hybridization: a practical appproach, Hames and Higgins Ed., IRL Press, Oxford) 또는 F. AUSUBEL 등(1999)의 연구(Currents Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.)에 함유된 교시에 따라 조정될 수 있다.

본 발명의 목적을 위하여, 특정 핵산은 서열번호 2 또는 이 서열의 단편 또는 유도체 서열을 포함하는 폴리펩타이드, 특히 인간 PAP1 단백질을 암호화한다. 유리하게도 핵산은 서열번호 1, 12, 14, 42 또는 44의 핵산 서열을 포함한다.

이러한 핵산은 본 발명의 펩타이드 화합물을 생산하는 데 사용될 수 있다. 따라서, 본 출원은 본 발명에 따른 핵산을 함유하는 세포가 핵산을 발현하기 위한 조건하에 배양되는, 또한 생산된 펩타이드 화합물이 회수되는 것에 따라 이러한 펩타이드 화합물을 제조하기 위한 방법에 관한 것이다. 이러한 경우에, 펩타이드 화합물을 암호화하는 부분은 일반적으로는 숙주 세포내 이의 발현을 허용하는 시그널의 조절하에 배치된다. 이러한 시그널(프로모터, 터미네이터, 분비 리더 서열 등)의 선택은 사용되는 숙주 세포의 기능에 따라 다양할 수 있다. 또한, 본 발명의 핵산은 자율적으로 복제할 수 있거나 삽입될 수 있는 벡터의 부분을 형성할 수 있다. 더욱 구체적으로는, 자율-복제되는 벡터는 선택된 숙주에서 자가 복제할 수 있는 서열을 사용하여 제조된다. 삽입 벡터에 관해서는, 예를 들면 숙주 게놈의 특정 지역에 상동성인 서열을 사용하여 제조될 수 있으며, 이는 상동 재조합에 의해 벡터의 삽입을 허용한다. 이는 플라스미드, 에피좀, 염색체, 바이러스 등의 유형의 벡터일 수 있다.

재조합 경로를 통해 본 발명의 펩타이드 화합물을 생산하기 위해 사용될 수 있는 숙주 세포는 진핵 및 원핵 숙주이다. 적당한 진핵 숙주로는 동물 세포 중 효모 또는 진균을 들 수 있다. 특히, 효모로는 사카로마이세스(Saccharomyces), 클루이베로마이세스(Kluyveromyces), 피히아 (Pichia), 쉬반니오마이세스(Schwanniomyces), 또는 한세눌라(Hansenula) 속을 들 수 있다. 동물 세포로는, COS, CHO, C127, PC12 등을 들 수 있다. 진균으로는, 더욱 구체적으로 아스페르길루스(Aspergillus) 아종 또는 트리코데르마(Tricoderma) 아종을 들 수 있다. 원핵 숙주로서, 하기 박테리아의 사용이 바람직하다: 이.콜라이(E.coli), 바실러스(Bacillus) 또는 스트렙토마이세스(Streptomyces).

본 발명의 주제는 또한 이들의 세포에 본 발명에 따른 핵산 또는 벡터를 포함하는 비인간 포유류이다,

이러한 포유류(설치류, 개, 토끼 등)는 구체적으로 PAP1의 특성을 연구하기 위해 또한 치료 목적으로 화합물을 동정하기 위해 사용될 수 있다. 이러한 트랜스제닉 동물의 게놈의 변형은 " 녹-인(knock-1n)" 또는 " 녹-아웃(knock-out)" 에 의해 1 이상의 유전자의 변경 또는 변형의 결과일 수 있다. 이러한 변형은 통상의 변경제 또는 변이원을 사용하여, 또는 부위-지향 돌연변이에 의해 생산될 수 있다. 이러한 게놈의 변형은 이의 야생형 또는 돌연변이 형태의 (1) 유전자(들)의 삽입 또는 (1) 유전자(들)의 치환의 결과일 수 있다. 게놈의 변형은 유리하게도 생식 줄기 세포에 대해, 또한 유리하게도 전핵에 대해 수행될 수 있다. 트랜스제네시스는 변형된 유전자를 포함하는 발현 카셋트를 2개의 수정된 전핵으로 마이크로주입함으로써 수행할 수 있다. 따라서, 본 발명에 따른 동물은 핵산을 포함하는 발현 카셋트의 주입에 의해 수득될 수 있다. 우선적으로, 이러한 핵산은 게놈 DNA(gDNA) 또는 상보적 DNA(cDNA)일 수 있는 DNA이다.

본 발명에 따른 트랜스제닉 동물의 작제는 당업자에게 공지된 통상의 기술에 따라 수행할 수 있다. 당업자는 구체적으로 트랜스제닉 동물의 생산, 특히 트랜스 제닉 마우스의 생산을 미국 특허 제 4,873,191호, 5,464,764호 및 5,789,215호에 기재하였으며, 이러한 문헌의 내용은 본원에 참조로 인용된다.

간단히, 본 발명에 따른 핵산을 포함하는 폴리뉴클레오타이드 작제물은 ES형의 줄기 세포주로 삽입된다. 폴리뉴클레오타이드 작제물은 바람직하게는 Thomas 등(1987, Cell, Vol. 51: 503-512)에 의해 기재된 바와 같이 일렉트로포레이션 (electroporation)에 의해 삽입된다.

일렉트로포레이션 단계를 겪은 세포는, 적당하고 상동 재조합 이벤트를 따르는 게놈으로 삽입된 외생성 폴리뉴클레오타이드 작제물을 가진 양성 세포를 선별하기 위하여, 이어서 폴리뉴클레오타이드 작제물의 존재에 대해 스크리닝(예를 들면 마커를 사용하는 선별, 또는 대안적으로는 PCR 또는 써던 블롯팅과 같은 전기영동 겔상에서의 DNA의 분석에 의해)된다. 이러한 기술은 예를 들면, MANSOUR 등(Nature(1988) 336: 348-352)에 의해 기술되었다.

다음으로, BRADLEY(1987, Production and Analysis of Chimaeric mice. In: E.J. ROBERTSON(Ed., teratocarcinomas and embryonic stem cells: A practical approach. IRL Press, Oxford, page 113)에 의해 기재된 바와 같이, 양성으로 선별된 세포를 분리하고 클로닝한 후, 3.5일된 마우스 배반포에 주입하였다. 이어서 배반포를 암컷 숙주 동물로 도입하고 배를 일정 기간 동안 발생하도록 했다.

대안에 따라서, ES 형 중 양성으로 선별된 세포를 WOOD 등(1993, Proc. Natl. Acad. Sci. USA, Vol. 90: 4582-4585) 또는 NAGY 등(1993, Proc. Natl. Acad. Sci. USA, vol. 90: 8424-8428)에 의해 기술된 바와 같이 8 내지 16 세포기(morulae)에서 2.5일된 배와 접촉시키고, ES 세포는 생식선을 발현하는 세포를 포함하는 배반포를 강력히 콜로니화하기 위해 내부화된다.

이어서 폴리뉴클레오타이드 작제물(트랜스유전자)이 삽입된 것들을 측정하기 위해 시험한다.

본 발명에 따른 핵산은 또한 유전적 안티센스 또는 약제로 사용될 수 있는 안티센스 올리고뉴클레오타이드를 제조하는데 사용될 수 있다. 안티센스 서열은 주어진 서열의 암호화 스트랜드에 상보적인 짧은 길이의 올리고뉴클레오타이드이며, 결과적으로 mRNA 전사체와 특이적으로 하이브리드화할 수 있어, 단백질로의 해독을 억제한다. 따라서 본 발명의 주제는 적어도 부분적으로, 파킨(parkin)에 대한 PAP1 단백질의 상호작용을 억제할 수 있는 안티센스 서열이다. 이러한 서열은 상기 정의된 핵산 서열의 전부 또는 부분으로 이루어질 수 있다. 이들은 일반적으로는 파킨과 상호작용하는 펩타이드를 암호화하는 서열에 대해 상보적인 서열, 또는 서열의 단편이다. 이러한 올리고뉴클레오타이드는 단편화 등에 의해, 또는 화학 합성에 의해 수득될 수 있다.

청구된 서열은 유전자 요법과 관련해서, 파킨과 PAP1 단백질의 상호작용을 조절할 수 있는 안티센스 서열 또는 펩타이드의 생체내 전달 및 발현을 위해서 사용될 수 있다. 이러한 관점에서, 서열은 생체내 이들의 투여를 허용하는 바이러스 또는 비바이러스 벡터에 삽입될 수 있다(Kahn 등, 1991). 본 발명에 따른 바이러스 벡터로서, 가장 바람직하게는 아데노바이러스, 레트로바이러스, 아데노-어소시에이티드 바이러스(AAV) 또는 허프스바이러스형 벡터를 들 수 있다. 본 출원의 주제는 또한 본 발명에 따른 폴리펩타이드, 특히 서열번호 2 또는 이 서열의 유도체의 전부 또는 부분, 예를 들면 서열번호 12, 14, 42 또는 44의 서열 또는 이 서열의 유도체의 전부 또는 부분을 포함하는 폴리펩타이드 또는 펩타이드를 암호화하는 핵산 서열을 포함하는 결손 재조합 바이러스이다.

본 발명은 또한 합성일 수 있거나 또는 합성이지 않을 수 있는, 또한 상기 정의된 뉴클레오타이드 서열 또는 이의 상보적 스트랜드와 하이브리드화할 수 있는 뉴클레오타이드 탐침의 제조를 허용한다. 이러한 탐침은 시험관내 PAP1의 발현 또는 과다발현에 대한 검출용으로, 또는 대안적으로는 유전적 비정상(부정확한 스플라이싱, 다형형, 점 돌연변이 등)에 대한 제시용 진단 도구로서 사용될 수 있다. 이러한 탐침은 또한 상기 정의된 펩타이드를 암호화하는 상동 핵산 서열을 다른 세포원 및 바람직하게는 인간 기원의 세포로부터 검출하고 분리하기 위해 사용될 수 있다. 본 발명의 탐침은 일반적으로는 적어도 10개 염기를 포함하며, 또한 예를 들면, 하나의 상기 언급된 서열 또는 이의 상보적 스트랜드 전부 이하를 포함할 수 있다. 바람직하게는, 이 탐침은 이의 사용 전에 표지된다. 이를 위해서, 당업자에게 공지된 다양한 기술(방사성, 형광, 효소, 화학적 표지 등)이 사용될 수 있다.

본 발명은 또한, PAP1을 암호화하는 핵산의 전부 또는 부분을 증폭시키기 위한 프라이머 또는 프라이머의 쌍, 예를 들면 서열번호 16 내지 41에서 선택된 서열의 프라이머에 관한 것이다.

본 발명의 주제는 또한, 활성 성분으로서, 적어도 상기 정의된 1 화합물, 특히 펩타이드 화합물을 포함하는 임의 약학 조성물이다.

본 발명의 주제는 구체적으로, 활성 성분으로서, 상기 정의된 적어도 하나의 항체 및/또는 하나의 항체 단편을 포함하는 임의 약학 조성물, 및 활성 성분으로서 상기 정의된 적어도 하나의 핵산 또는 하나의 벡터를 포함하는 임의 약학 조성물이다.

본 발명의 주제는 또한 활성 성분으로서, PAP1 단백질과 파킨의 상호작용을 증가 또는 감소시킬 수 있는 화학 분자를 포함하는 임의 약학 조성물이다.

또한, 본 발명의 주제는 상기 정의된 펩타이드, 항체, 화학 분자 및 뉴클레오타이드 서열이 서로 또는 다른 활성 성분과 조합되는 임의 약학 조성물이다.

본 발명에 따른 약학 조성물은 파킨 단백질의 활성을 조절하고, 결과적으로 도파민성 뉴런의 생존을 유지하는 데 사용될 수 있다. 더욱 구체적으로는, 이 약학 조성물은 PAP1 단백질과 파킨의 상호작용을 조절하고자 한다. 이들은, 더욱 바람직하게는, 예를 들면, 파킨슨병과 같은 중추신경계의 질병을 치료하고자 하는 약학 조성물이다.

본 발명의 주제는 또한 파킨의 활성을 조절하고, 또는 중추신경계 질병의 유형화를 위한 상기 기술된 분자의 용도이다. 구체적으로, 본 발명은 파킨의 활성을 적어도 부분적으로 조절하기 위한 이들 분자의 용도에 관한 것이다.

본 발명은 또한 서열번호 2 또는 서열번호 4, 또는 이 서열의 단편(또는 유도체)과 결합할 수 있는 분자를 선택하는 것을 포함하는, 파킨의 기능에 대해 활성인 분자를 스크리닝 또는 특징분석하기 위한 방법에 관한 것이다. 본 방법은, 유리하게도 시험관내에서, 서열번호 2 또는 서열번호 4의 서열, 또는 이 서열의 단편(또는 유도체)을 포함하는 폴리펩타이드와 시험될 분자를 접촉시키고, 또한 서열번호 2(특히 잔기 1 내지 잔기 344 지역) 또는 서열번호 4의 서열과 결합할 수 있는 분자를 선택하는 것을 포함한다. 시험된 분자는 자연적으로(펩타이드, 핵산, 지질, 당 등 또는 이 분자의 혼합물, 예를 들면, 조합 라이브러리 등) 다양할 수 있다. 상기 지적하였듯이, 따라서 분자는 파킨 단백질의 활성을 조절하기 위해 사용될 수 있으며, 신경퇴행성 병리를 치료하기 위한 잠재적 치료제를 나타낸다.

본 발명의 다른 이점은 도해로서 및 비제한적으로서 고려되어야 할 하기 실시예 및 도면을 읽음으로써 명백해질 것이다.

도면의 간단한 설명

도 1: 벡터 pLex9-파킨(135-290)의 도해.

도 2: 1차 5'-RACE 실험의 결과. 8개 클론이 수득되었다. 초기 전기 서열은 도면의 하부에 표시하였다.

도 3: 2차 5'-RACE 실험의 결과. 1차 실험에서 수득된 8개 클론 중 단지 2개 클론(클론 A12 및 D5)만이 확인되었다. 초기 전기 서열은 도면의 하부에 표시하였다. DNA 및 단백질의 완전한 서열은 서열 12 내지 15에 주어졌다.

도 4: 2차 5'-RACE 실험의 클론 C5 및 D4의 구조의 세부사항. 생성되는 컨센서스 서열은 도면의 상부에 표시하였다.

도 5: 인간 뇌에서 분리된 전사체의 구조.

도 6: 인간 뇌의 LY111(전길이)의 핵산 및 단백질 서열. (두줄 밑줄: 징크 핑거 도메인의 보존된 시스테인. 굵은 글씨: $C_2$1 도메인, 이탤릭체: C_2C 도메인.)

도 7: 인간 뇌의 LY111(짧은 버전)의 핵산 서열 및 단백질 서열. (두줄 밑줄: 징크 핑거 도메인의 보존된 시스테인. 굵은 글씨: $C_2$1 도메인, 이탤릭체: C_2C 도메인.)

도 8: Cos-7 세포에서의 발현 후 단(8b) 또는 장(8a) LY111 단백질의 위치.

도 9: 인간 폐로부터 LY111(긴 버전)의 핵산 및 단백질 서열.

도 10: 인간 뇌로부터 LY111(짧은 버전)의 핵산 및 단백질 서열.

발명의 상세한 설명

재료 및 사용된 기술

1) 효모 스트레인:

S. 세레비시애(S.cerevisiae) 속의 L40 스트레인(Mata, his3D200, trp1-901, leu2-3, 112, ade2, LYS2:: (lex AOP)$_4$-HIS3, URA3::(lexAOP)$_8$-LacZ, GAL4, GAL80)을 하나의 단백질 파트너가 LexA 단백질에 융합되었을 때 단백질-단백질 상호작용을 증명하는 데 사용하였다. 후자는 리포터 유전자 LacZ 및 His3의 발현을 조절하는 Lex 반응 인자를 인지할 수 있다.

이를 하기 배양 배지에서 배양하였다.

완전한 YPD 배지: -효모 추출물(10 g/l)(Difco)

-박토펩톤(20 g/l)(Difco)

-글루코스(20 g/l)(Merck)

20 g/l의 아가(Difco)를 첨가하여 이 배지를 고체화했다.

최소 YNB 배지: -효모 질소 염기(아미노산 없음)(6.7 g/l)(Difco)

-글루코스(20 g/l)(Merck)

20 g/l의 아가(Difco)를 첨가하여 이 배지를 고체화했다. CSM 배지[CSM-Leu, -Trp, -His(620 mg/l), CSM-Trp(740 mg/l) 또는 CSM-Leu, -Trp(640 mg/l)(Bio101)] 및/또는 2.5 mM 3-아미노-1,2,4-트리아졸의 첨가에 의해 아미노산 및/또는 3-아미노-1,2,3-트리아졸을 이 배지에 보충할 수 있다.

2) 박테리아 스트레인:

유전자형 supE, hsdΔ5, thi, Δ(lac-proAB), F'[tra D36 pro A^{+} B^{+} lacI^{q}laczΔM15]의 이.콜라이의 TG1 스트레인을, 사용된 재조합 플라스미드의 증폭 및 분리의 수단으로서 플라스미드를 작제하기 위하여 사용하였다. 이를 하기 배지에서 배양하였다.

LB 배지: -NaCl (5g/l)(Prolabo)

-박토트립톤(10 g/l)(Difco)

-효모 추출물(5 g/l)(Difco)

15 g/l의 아가(Difco)를 첨가하여 이 배지를 고체화했다.

앰피실린을 100 ug/ml로 사용하였다; 이 항생제를 마커로서, 이 항생제에 내성인 유전자를 포함하는 플라스미드를 받은 박테리아를 선별하는 데 사용한다.

유전자형 supE44, ara14, galK42, lacY1, Δ(gpt-proA)62, rpsL20(Str ^{r}), xyl-5, mtl-1, recA13, Δ(mcrC-mrr), HsdS $^{-}$ (r $^{-}$ m $^{-}$)의 이.콜라이의 HB101 스트레인을, 인간 림프구 cDNA 라이브러리로부터 기원하는 플라스미드를 증폭 및 분리하기 위한 수단으로서 사용하였다.

이를 하기 배지에서 배양하였다.

M9 배지: - Na_2HPO_4 (7 g/l) (prolabo)

- KH_2PO_4 (3 g/l) (prolabo)

- NH_4Cl(1 g/l) (prolabo)

- NaCl(0.5 g/l) (prolabo)

- 글루코스(20 g/l) (sigma)

- $MgSO_4$ (1 mM) (prolabo)

- 티아민(0.001%) (sigma)

15 g/l의 아가(Difco)를 첨가하여 이 배지를 고체화했다.

HB101 스트레인을 생장시키기 위해 M9 배지에 루이신(50 mg/l) (Sigma) 및 프롤린(50 mg/l) (Sigma)이 첨가되어야 한다.

림프구 cDNA 2-하이브리드 라이브러리로부터 기원하는 플라스미드를 선별하는 동안에, 루이신은 첨가하지 않았는데, 왜냐하면 플라스미드가 Leu2 선별 마커를 함유하기 때문이다.

3) 플라스미드:

pGBT10에 상동성이며, 융합 단백질을 형성하기 위해 LexA 박테리아 리프레서, 터미네이터의 상류를 암호화하는 서열의 하류에 위치된 다중 클로닝 부위를 함유하는 5-kb 벡터 pLex9(pBTM116) (Bartel 등, 1993).

pLex-HaRasVa112; WO 98/21327에 기술된 바와 같이, 플라스미드 pLex9은 Va112 위치에서 돌연변이된 HaRas 단백질을 암호화하는 서열을 함유하며, 이는 포유류 Raf 단백질과 상호작용하는 것으로 공지되어 있다(Vojtek 등, 1993). 이 플라스미드를 L40 스트레인에서 PAP1 단백질의 상호작용의 특이성을 시험하기 위하여 사용하였다.

pLex9-cAPP; FE65의 PTB2 도메인과 상호작용하는 것으로 공지된, APP 단백질의 세포질 도메인을 암호화하는 서열을 함유하는 pLex9 플라스미드. 이 플라스미드를 L40 스트레인에서 PAP1 단백질의 상호작용의 특이성을 시험하기 위하여 사용하였다.

4) 합성 올리고뉴클레오타이드:

TTAAGAATTC GGAAGTCCAG CAGGTAG 서열번호 5

ATTAGGATCC CTACACACAA GGCAGGGAG 서열번호 6

파킨의 중심 영역에 상응하는 PCR 단편을 수득할 수 있게 하는 올리고뉴클레오타이드를 EcoRI 및 BamHI 부위에 실었다.

GCGTTTGGAA TCACTACAG 서열번호 7

GGTCTCGGTG TGGCATC 서열번호 8

CCGCTTGCTT GGAGGAAC 서열번호 9

CGTATTTCTC CGCCTTGG 서열번호 10

AATAGCTCGA GTCAGTGCAG GACAAGAG 서열번호 11

서열을 삽입하는 데 사용된 올리고뉴클레오타이드는 PAP1 유전자에 상응한다.?

올리고뉴클레오타이드를 Applied System ABI 394-08기를 사용하여 합성하였다. 이들은 암모니아로 합성 매트릭스로부터 제거되고, 10 용적의 n-부탄올로 2번 침전된 후, 물에서 취해진다. 광학 밀도를 측정함으로써 정량을 수행하였다($1OD_{260}$ 은 30 ug/ml에 상응).

5) 플라스미드 DNA의 제조

플라스미드 DNA의 소량 및 다량의 제조물을 DNA 정제 키트의 제조업자인, Quiagen에 의해 추천되는 프로토콜에 따라 수행하였다:

- Quiaprep Spin Miniprep 키트, 참조: 27106
- Quiaprep plasmid Miniprep 키트, 참조: 12613.

6) PCR(중합효소 연쇄반응)에 의한 DNA의 효소 증폭:

DNA 매트릭스의 존재하에 100 ul의 최종 용적으로, dNTP(0.2 mM), PCR 완충액(10 mM Tris-HCl pH 8.5, 1 mM $MgCl_2$, 5 mM KCl, 0.01% 젤라틴), 10 내지 20 pmol의 각 한가지 올리고뉴클레오타이드 및 2.5 IU의 Ampli Taq DNA 중합효소(Perkin Elmer)로 PCR 반응을 수행하였다. 혼합물을 2 방울의 액체 석유 젤리로 덮어, 샘플의 증발을 제한한다. 사용된 기기는 Appligene의 " Crocodile II" 이다.

94℃의 변성 온도, 52℃의 하이브리드화 온도 및 72℃의 효소에 의한 신장을 위한 온도를 사용하였다.

7) 연결:

모든 연결 반응은 최종 용적 20 ul로, 100 내지 200 ng의 벡터, 0.1 내지 0.5 ug의 삽입체, 40 IU의 T4 DNA 리가제 효소(Biolab) 및 연결 완충액(50 mM Tris-HCl pH 7.8; 10 mM $MgCl_2$; 10 mM DTT; 1 mM ATP)의 존재하에서, 1시간 동안 37℃에서 수행하였다. 음성 대조군은 삽입체의 부재하에 벡터의 연결로 이루어진다.

8) 박테리아의 형질전환:

플라스미드로의 박테리아의 형질전환은 하기 프로토콜에 따라 수행된다: Chung의 방법(Chung 등, 1989)에 따라 10 ul의 연결 용적을 TG1 박테리아를 형질전환하는 데 사용하였다. 형질전환 후에, 박테리아를 LB 배지 + 앰피실린에 플레이팅 하고, 37℃에서 16시간 동안 배양하였다.

9) DNA의 분리 및 추출:

DNA의 분리는 이의 크기의 함수로서, Maniatis(Maniatis 등, 1989)에 따라 아가로스 겔 전기영동에서 수행한다: TBE 완충액(90 mM Tris 베이스; 80 mM 보레이트; 2 mM EDTA)내 1% 아가로스 겔(Gibco BRL).

10) 플라스미드 DNA의 형광 서열분석:

사용된 서열분석 기술은 Sanger의 방법(Sanger 등, 1977)에서 유래하였으며, Applied Biosystem에 의해 발전된 형광에 의한 서열분석을 위해 변용된다. 사용된 프로토콜은 시스템의 디자이너(Perkin Elmer, 1977)에 의해 기술되었다.

11) 효모의 형질전환:

효모를 형질전환하기 위한 통상의 방법을 사용하여 플라스미드를 효모로 도입시키는데, 이 방법은 Gietz(Gietz 등, 1992)에 의해 발전되고 하기 방법으로 변형된다:

림프구 cDNA 라이브러리로 효모를 형질전환하는 특수한 경우에, 사용되는 효모는 LexA 단백질에 융합되는 파킨의 중심 부분을 암호화하는 pLex9-파킨(135-290) 플라스미드를 함유한다. 이는 30℃에서 흔들며 10^7 세포/ml의 밀도가 획득될 때까지 아미노산 CSM-Trp이 보충된 200 ml의 YNB 최소 배지에서 배양된다. 상기 프로토콜에 따라 효모의 형질전환을 수행하기 위하여, 세포 현탁액을 5 ug의 라이브러리가 첨가된 10개의 50 ul 들이 튜브에 분리시킨다. 열쇼크를 20분간 수행하고, 세포를 원심분리시켜 수집한 후, 100 ml의 YPD 배지에서 1시간 동안 30℃에서, 또한 CSM-Leu, -Trp이 보충된 100 ml의 YNB 배지에서 3시간 동안 30℃에서 재현탁시킨다. 형질전환의 효율은 CSM-Trp, -Leu가 보충된 고체 YNB 배지상에, 형질전환된 세포의 다양한 희석액을 둠으로써 측정된다. 30℃에서 배양 3일 후에, 수득된 콜로니를 세고, 림프구 라이브러리 DNA의 ug 당 형질전환의 정도를 측정하였다.

12) 효모로부터 추출된 플라스미드의 분리:

16시간 동안 30℃에서 배양된 5 ml의 효모 배양액을 원심분리하고, 200 ul의 용해 완충액(1M 소르비톨, 0.1 M KH_2PO_4/K_2HPO_4 pH 7.4m, 12.5 mg/ml 자이몰리아제)에서 취한 후, 1시간 동안 37℃에서 배양한다. 이어서 용해물을 DNA 정제 키트, Quiaprep Spin Miniprep 키트, 참조 27106의 제조자인, Quiagen에 의해 추천되는 프로토콜에 따라 처리한다.

13) 베타-갈락토시다제 활성 분석:

니트로셀룰로스 시트를 각각 서로 분리된 효모 콜로니를 함유하는 페트리 디쉬상에 미리 깔아둔다. 이어서, 효모를 파괴하고 따라서 베타-갈락토시다제 활성을 방출하기 위해서, 이 시트를 30초 동안 액체 질소에 침지시킨다. 해동 후, 니트로셀룰로스의 시트를, N,N-디메틸포름아미드 40 mg/ml에서 15 ㎕의 X-Gal(5-브로모-4-클로로-3-인돌릴-β-D-갈락토사이드)를 함유하는 1.5 ml의 PBS 용액(60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, pH 7)에 미리 침지시켰던 와트만 페이퍼가 들어있는 또다른 페트리 디쉬에, 콜로니가 위를 향하게 하여 배치한다. 이어서, 디쉬를 37℃의 배양기에 넣는다. 막 위의 콜로니가 12시간 후 청색으로 변할 때 시험은 양성으로 칭한다.

실시예

실시예 1: 파킨의 중심부와 LexA 박테리아 억제자 간에 융합이 일어나는 융합 단백질의 발현을 허용하는 벡터의 작제

실시예 1: 파킨의 중심부와 LexA 박테리아 억제자 간에 융합이 일어나는 융합 단백질의 발현을 허용하는 벡터의 작제

이중-하이브리드 시스템을 사용하는 라이브러리의 스크리닝은 LexA 박테리아 억제자와 같은 DNA 결합 단백질에 융합될 파킨의 중심 영역을 요한다. 이 융합 단백질의 발현은 서열번호 3 또는 4의 서열에 제시된 서열에 있는 파킨의 중심 영역을 암호화하는 서열이 LexA 단백질에 상응하는 서열과 동일한 판독 프레임에 도입되는 벡터 pLex9(재료와 방법 참조)를 사용하여 수행된다.

아미노산 135에서 개시하는 파킨의 중심 영역의 156 아미노산에 상응하는 DNA의 468 bp-단편은 5′ 말단에 EcoRI 부위 및 3′ 말단에 BamHI 부위를 도입시킬 수 있게 만든, 올리고뉴클레오타이드(서열번호 5 및 6의 서열)를 사용하는 PCR에 의해 수득된다. PCR 단편은 벡터 pLex9-파킨(135-290)을 제공하기 위하여, 단백질 LexA를 암호화하는 서열의 다운스트림에, 플라스미드 pLex9의 다중 클로닝 부위의 EcoRI 및 BamHI 부위 사이에 도입된다(도 1).

작제물은 DNA 서열분석에 의해 증명된다. 이러한 증명은 이 단편이 PCR 반응 동안 돌연변이되지 않으며, LexA에 상응하는 단편의 것과 동일한 개방 판독 프레임에 융합됨을 보여줄 수 있게 한다.

실시예 2: 림프구 융합 라이브러리의 스크리닝실시예 2: 림프구 융합 라이브러리의 스크리닝

본 발명자는 이중 - 하이브리드법(Fields and Song, 1989)을 이용하였다.

융합 단백질의 스크리닝은 GAL4의 전사활성화 도메인과 융합되고, 실시예 1에 기술된 해당 단백질(파킨의 중심 영역)과 상호작용할 수 있는 단백질 생성 클론을 동정할 수 있게 한다. 이러한 상호작용은 스트레인 L40에서 리포터 유전자 His3과 LacZ의 발현을 유도할 수 있게 되는 전사활성자를 재구성할 수 있게 해준다.

이러한 스크리닝을 수행하기 위하여, 본 발명자는 Richard Benarous(Peytavi et al., 1999)에 의해 공급된, 말초 인간 백혈구로부터 기원하는 cDNA로부터 제조되는 융합 라이브러리를 선택하였다. 효모를 림프구 라이브러리로 형질전환시키고 양성 클론을 전술한 바와 같이 선별한다.

스크리닝 동안, 융합 라이브러리로부터의 각각의 분리된 플라스미드가 적어도 하나의 효모에 플라스미드 pLex9 - 파킨(135 - 290)과 동시에 존재하는 확률을 유지시키는 것이 필요하다. 이러한 확률을 유지하기 위해서는, 효모의 형질전환의 양호한 효율을 갖는 것이 중요하다. 이를 위해, 본 출원인은 DNA ㎍당 2.6 X 10^5 형질전환 세포의 효율을 부여하는 효모 형질전환 프로토콜을 선택하였다. 또한, 두 상이한 플라스미드로 효모를 공형질감염시키면 이러한 효율이 감소되므로, 플라스미드 pLex9 - 파킨(135 - 290)으로 예비 형질전환되는 효모의 사용을 선호하였다. 표현형 His - , Lys - , Leu - , Ade - 의 이러한 스트레인 L40 pLex9 - 파킨 (135 - 290)을 융합 라이브러리로부터의 플라스미드 DNA 50 ㎍으로 형질전환시킨다. 이러한 양의 DNA는 발명자로 하여금, 평가 후, 1.3 x 10^7 형질전환 세포를 수득하도록 허용하였으며, 이러한 값은 라이브러리를 구성하는 분리된 플라스미드의 수보다 약간 상회하는 수에 해당한다. 이러한 결과에 따라, 라이브러리의 플라스미드의 거의 전부가 효모의 형질전환에 사용되었음이 고려될 수 있다. 기능성 전사활성자를 재구성할 수 있는 형질전환 세포의 선별은 2.5 mM 3 - 아미노 - 1,2,4 - 트리아졸 및 620 mg/l의 CSM(Bio101)로 보충시키고 히스티딘, 루이신 및 트립토판을 함유하지 않는 YNB 배지에서 수행되었다.

이러한 선별의 말미에, His+ 표현형을 갖는 다수의 클론이 수득되었다. β - 갈락토시다제 활성 분석을 이들 형질전환체에 대하여 수행하여 다른 리포터 유전자 LacZ의 발현에 의해, 수득된 클론의 이러한 수를 증명한다. 115개의 클론이 His+, βGal+ 이중 표현형을 지녔으며, 이는 단백질 - 단백질 상호작용에 상응할 수 있다.

실시예 3: 선별된 효모 클론에서 라이브러리 플라스미드의 분리실시예 3: 선별된 효모 클론에서 라이브러리 플라스미드의 분리

파킨의 중심 영역과 상호작용할 수 있는 단백질을 동정하기 위하여, 이중 - 하이브리드 스크리닝 동안 선별되었던 효모 내 함유된 융합 라이브러리 플라스미드를 추출하였다. 이의 다량을 수득할 수 있도록 하기 위해, 이러한 분리는 양성 효모 스트레인으로부터의 DNA 추출물로 이.콜라이를 미리 형질전환시킬 것을 요한다. 이 추출물에 함유되어 있는 라이브러리 플라스미드가 효모/이.콜라이 셔틀 플라스미드이므로, 이는 박테리아에서 용이하게 복제될 수 있다. 라이브러리 플라스미드는 루이신 - 결핍 배지상에서, 루이신에 대해 영영 요구성 HB101 박테리아를 상보화시킴으로써 선별된다.

효모로부터의 DNA 추출물로 형질전환시킨 후 수득된 박테리아 콜로니로부터의 플라스미드 DNA는 제한효소로 분해시킨 다음 DNA 단편을 아가로스 겔상에서 분리시킴으로써 분석된다. 분석된 115개 클론 중에서, 다른 것들과 상이한 프로필을 보인, 라이브러리 플라스미드를 함유하는 하나의 클론을 수득하였다. pGAD - Ly111b로 명명된 이 플라스미드를 좀더 정확하게 검토하였다.

실시예 4: 동정된 플라스미드에 함유된 삽입체의 서열 결정실시예 4: 동정된 플라스미드에 함유된 삽입체의 서열 결정

동정된 플라스미드에 함유된 삽입체의 서열분석은 먼저, 림프구 cDNA 라이브러리의 삽입의 EcoRI 부위에 근접한, 서열 GAL4TA 영역에 상보적인, 올리고뉴클레오타이드 서열번호 7을 사용하여; 이어서, 두 번째로, 서열분석 과정 중에 수득되는 삽입체의 서열에 상응하는, 올리고뉴클레오타이드 서열번호 8 내지 서열번호 11을 사용하여 수행한다. 수득된 서열은 서열번호 1의 서열상에 제시되어 있다. 이렇게 동정된 단백질은 PAP1(Parkin - Associated Protein 1)로 언급된다.

이 삽입체의 서열을 데이터뱅크 GENBank 및 EMBL(European Molecular Biology Lab)에 들어있는 서열과 비교한 결과, 시냅토태그민 계통의 다양한 멤버와단백질 수준에서, 25%의 상동성을 보였다. 시냅토태그민은 뇌와 타 조직에서 발현되는 적어도 11종의 상이한 유전자에 의해 암호화되는 막 단백질 계통의 일부이다. 이들은 C_2 로 명명되는 두 칼슘 - 조절 도메인과 하나의 막횡단 도메인을 함유한다. 이 도메인에서, 시냅토태그민과 PAP1 단백질 간의 상동성이 발견되었다. 다른 유의한 상동성은 관찰되지 않았다.

실시예 5: 파킨의 중심 영역과 PAP1 단백질 간의 상호작용 특이성의 분석 실시예 5: 파킨의 중심 영역과 PAP1 단백질 간의 상호작용 특이성의 분석

PAP1 단백질에 상응하는 단편과 파킨의 중심 영역 간의 상호작용 특이성을 측정하기 위해, 다른 비관련 단백질과의 특이적 상호작용을 위한 2 - 하이브리드 시험을 수행하였다. 이 시험을 수행하기 위해, 본 출원인은 스트레인 L40을, 플라스미드 pLex9 - 파킨(135 - 290) 대신에, LexA의 DNA 결합 도메인과 융합되는, APP 또는 HaRasVal12 단백질의 세포질 도메인을 각각 암호화하는 대조 플라스미드 plex9 - cAPP 또는 pLex9 - HaRasVal12로, 및 2 - 하이브리드 라이브러리의 스크리닝에서 분리되는 플라스미드로 형질전환시켰다. β - Gal 활성 분석은 단백질 - 단백질 상호작용을 측정하기 위하여, 다양한 플라스미드로 형질전환시킨 세포 상에서 수행된다. 분석 결과에 따라, 2 - 하이브리드 라이브러리의 스크리닝 동안에 분리되었던 플라스미드로, 및 플라스미드 pLex9 - 파킨(135 - 290)으로 형질전환시킨 효모만이 β - Gal 활성을 지녔으며, 이는 따라서 파킨의 중심 영역과 PAP1 단백질 간의 상호작용을 보여준다. 이러한 상호작용은 따라서 특이적인 것으로 판명되었는데, 이유는 PAP1의 이러한 단편이 CAPP 또는 HaRasVal112 단백질과 상호작용하는 것으로 보이지는 않기 때문이다.

따라서 이러한 결과는 PAP1으로 언급되는, 파킨과 특이적으로 상호작용할 수 있는 신규 단백질의 존재를 보여준다. 시냅토태그민과 관련있는 이 단백질은 공지 단백질과 유의한 상동성을 보이지 않으며, 항체, 탐침 또는 펩타이드의 생산을 위해, 또는 활성 분자의 스크리닝을 위해 치료 또는 진단 적용에 사용될 수 있다.

실시예 6: 인간 폐 DNA 라이브러리로부터 PAP1 유전자의 클로닝 실시예 6: 인간 폐 DNA 라이브러리로부터 PAP1 유전자의 클로닝

인간 PAP1 유전자의 완전한 서열을 동정하고 변이형의 존재를 특징규명하기 위해, 두 일렉트로닉 연장법을 서열번호 1의 서열을 사용하여 수행하였다. 서열번호 1의 서열에 비해 330 bp 연장을 포함하는, 각각 1644 bp 및 1646 bp의 두 일렉트로닉 서열이 수득되었다. 그러나, 이들 서열의 분석은 해독 후 자명해지는, 컨센서스 영역에서의 차이를 보여주었다. 따라서, 420aa ORF가 한 경우에 수득되었고, 230aa ORF가 다른 서열에서 수득되었다. 수득된 단백질 서열을 공지의 서열과 비교하며, 인간 시냅토개민 I(p65)(p21579)과, 293개의 중첩 아미노산에 걸쳐서 24% 상동성을 드러냈다. 시냅토개민 I의 기능은 시냅스 지역에서 시냅스낭의 트래피킹 동안 막 상호작용에서의 조절 역할일 수 있다. 이는 특정의 특이성으로 산 인지질과 결합한다. 또한, 시냅토개민과 활성화 단백질 키나제 C에 대한 수용체 간의 칼슘 - 의존형 상호작용이 보고되었다. 시냅토개민은 또한 3종의 타 단백질, 즉 뉴렉신, 신택신 및 ap2와 결합할 수 있다. 동정된 서열과 시냅토개민 계열 간의 상동성의 돌연한 조기 소멸을 가정하면, 동정된 서열이 천연 서열에 비해 결실을 지니고 있을 수 있다. 이러한 가설을 증명하고 서열을 확인하기 위해, RT - PCT 서열분석 실험을 1644 bp 서열을 사용하여 수행하였다. 수득할 서열은 시냅토개민과 일한 정도의 상동성을 갖는 420aa ORF를 포함한다.

좀더 큰 서열을 수득하고, 수득된 서열이 스플라이싱 형에 상응하는지 여부를 증명하기 위한 일환으로, 5′ - RACE 연장 실험을 인간 폐 cDNA의 제조시 oligos L1 및 L2를 사용하여, 확인된 서열의 3′ 영역으로부터 개시하였다.

수득된 결과를 도 2에 나타내었고 6종의 상이한 말단 5′ 말단에 상응하는 8 클론의 동정을 보여준다. 이들 중 셋은 ORF(클론 A12, F2, F12)를 방해하는 정지 코돈을 함유하고 클론 A3은 OPR을 함유하지 않는다. 각종 전사체의 존재는 RT - PCR 및 네스팅 RT - PCR에 의해 확인되었다(표 1).

[표 1]

RT - PCR	1차	2차 PCR U3 - L3	2차 PCR U1 - L4	2차 PCR C - B
U3 - L3	170			
A - L4	153		+	
A - L3	스미어	+		
U1 - 14	130			
U1 - L3	스미어	+		
U1 - B	415			+
U2 - B	515			+
기대된 크기		170	130	120

프라이머 U3 - L3와 C - B의 쌍은 서열의 공통 단편에 특이적이며, oligos A 및 U1은 초기 서열과 클론 C11에 대해 특이적이며, oligo L4는 초기 서열에 특이적이며 프라이머 U2는 클론 A3에 특이적이다. 두 번째 5′ - RACE는 다양한 클론에 공통인 영역에 위치된 oligos L3 및 L7으로 수행되었다(도 2). 수득된 결과는 도 3과 4에 주어졌다. 각종 전사체의 존재는 RT - PCR 및 네스팅 RT - PCT에 의해 확인되었다(표 2).

[표 2]

RT - PCR	결과	2차 PCRC - B	2차 PCRU3 - B	2차 PCRU5 - L7
U4 - F	스미어	+	+	+
U5 - F	스미어	+	+	+
U3 - F	1550 bp	+	+	
기대된 크기(bp)		120	385	530

프라이머와 올리고뉴클레오타이드의 서열을 표 3과 4에 나타내었다(서열번호 16 - 37).

[표 3]

서열번호LY111_ U4 CCAGTTCTGCCTGTTCATC 23 내지 41 16 LY111 _ U5 TTCAAAACACAGAGGAGGAG 319 내지 338 17 LY111_ U3 GAATTTGGTCAGTTTAGAGG 759 내지 778 18 LY111 _ L7 TTCTGGGATTTGGAGAGCTTTTTCAC 851 내지 825 19 LY111 _ L6 TCTGTCTGTCCCACACACTGCC 914 내지 892 20 LY111 _ L3 GACTGGCTCCGTCTCTCTG 928 내지 910 21LY111 _ C AAGCAACAGAATCTCCCATCC 1029 내지 1049 22 LY111 _ B GCATTGTCAAAATTGCCCATC 1147 내지 1127 23LY111 _ E AGGCGGAGAAATACGAAGAC 1543 내지 1562 24 LY111 _ D GCAGAGTGAGACAGCCCTTAAC 1767 내지 1746 25Ly111 _ L2 CTTCCTCAGGACTGGCGACTTCAG 1811 내지 1782 26Ly111 _ L1 CAAGCGGTCGTTCATTCCAAAGAG 1934 내지 1913 27 LY111 _ F AAGAGGAGATAACCCACCAGAG 2288 내지 2269 28

[표 4]

LY111 _ A TCGTAGAGCAGCAGGTCCAAG 14 내지 34 46LY111 _ U1 AGGGCTGCTGGCTATTTTTC 36 내지 55 29LY111 _ L4 TAAGAAATGGGTTGTGAAC 148 내지 166 30 LY111 _ C AAGCAACAGAATCTCCCATCC 1029 내지 1049 31LY111 _ B GCATTGTCAAAATTGCCCATC 1147 내지 1127 32LY111 _ E AGGCGGAGAAATACGAAGAC 1543 내지 1562 33LY111 _ D GCAGAGTGAGACAGCCCTTAAC 1767 내지 1746 34Ly111 _ L2 CTTCCTCAGGACTGGCGACTTCAG 1811 내지 1782 35 Ly111 _ L1 CAAGCGGTCGTTCATTCCAAAGAG 1934 내지 1913 36LY111 _ F AAGAGGAGATAACCCACCAGAG 2288 내지 2269 37

결과의 이러한 세트는 인간의 폐로부터 동정된 PAP1 단백질의 길이가 긴 이소형(도 9, 서열번호 12와 13) 및 길이가 짧은 이소형(도 10, 서열번호 14와 15)에 상응하는 컨센서스 서열을 확인할 수 있게 한다. 실시예의 나머지에서, 이 단백질은 또한 LY111이라는 용어를 사용하여 언급되고 있다. 길이가 긴 이소형은 서열번호 12의 잔기 237 - 2069에 위치된 1833 bp ORF에 의해 암호화되고, 610개의 아미노산을 포함한다. 폴리아데닐화 시그널은 뉴클레오타이드 2315로부터 개시하는 부위에 위치되어 있다. 길이가 짧은 이소형은 서열번호 14의 잔기 429 - 1370에 위치된 942 bp ORF에 의해 암호화되고, 313개의 아미노산을 포함한다. 폴리아데닐화 시그널은 뉴클레오타이드 1616으로부터 개시하는 부위에 위치되어 있다.

이어서, 탐침(amplimer CD 및 E - F)를 가지고 다양한 인간 조직에 대하여 노던 블롯팅 실험을 수행하며 근육에서 6 kb 전사체, 심장에서 전사체(3 kb) 및 또한 태아의 간에서 6 kb 전사체를 드러날 수 있게 해준다. 실시예 7은 또한, 인간 태아 뇌에서 전사체의 클로닝을 기술하고 있다.

다양한 상동성 연구가 다양한 단백질 데이터베이스에서 수행되었고, 그 결과를 하기의 표 5에 나타내었다.

[표 5]

데이터뱅크	상동성
Genpept116	G5926736 (AB025258) 그래뉴필린 - a동일성: 31%((215/679), 상동성(POS): 46%(322/679)
	G5926738 (AB025259) 그래뉴필린 - b동일성: 31%(150/479), 상동성 (POS): 47%(230/479)
	G1235722 (D70830) Doc2 베타 (호모 사피엔스)동일성: 25%(74/292), 상동성(POS): 43%(127/292)
	G289718 (L15302) 시냅토개민 - 1동일성: 26%(77/293), 상동성 (POS): 45%(133/293)
Swissprot	SP: SYT1 _ CAEEL 시냅토개민 1동일성: 26%(77/293), 상동성 (POS): 45%(133/293)
	SP:SYT2 _ MOUSE 시냅토개민 II동일성: 24%(72/293), 상동성 (POS): 44% (131/293)

실시예 7: 인간 태아 뇌의 상보성 DNA로부터 PAP1 (LY111B)의 두 완전 길이 전사체(LY111B)의 클로닝

실시예 7: 인간 태아 뇌의 상보성 DNA로부터 PAP1 (LY111B)의 두 완전 길이 전사체(LY111B)의 클로닝

인간의 뇌에서 완전한 길이 Ly111b 전사체의 존재를 확인하기 위하여, 인간 태아의 뇌로부터 유도된 상보 DNA(Marathon Ready cDNA, Clontech)를 사용하고, 프라이머로서 올리고뉴클레오타이드 LyF1(AAT GGA AGG GCG TGA CTC, 도 5, 서열번호 38) 및 HA71 (CCT CAC GCC TGC TGC AAC CTG, 서열번호 39)를 사용하여 PCR을 수행한다. 대략 2 킬로베이스의 약하게 나타난 DNA 단편이 증폭되었다. 이 첫번째 PCR의 산물을 네스팅 PCR을 위한 매트릭스로 사용하고, 올리고뉴클레오타이드 LyEcoF (GCACGAATTC ATG GCC CAA GAA ATA GAT CTG, 서열번호 40) 및 HA72 (CTG TCT TCG TAT TTC TCC GCC TTG, 서열번호 41)으로 수행하였다. 증폭 산물을 제한 효소 EcoRI(올리고뉴클레오타이드 LyEcoF 중으로 통합된) 및 BstEII(도 5)로 분해시켜 발현 벡터 pcDNA3 중으로 삽입시킨 다음, 이들의 서열을 결정한다. 수득된 클론의 서열 분석은 인간 태아의 뇌에 두 개의 잠재적인 완전 길이 Ly111b 전사체의 존재를 밝혀주었다(도 5). 이들 전사체(Ly111b$_{\text{fullA}}$)의 첫 번째는 인간의 폐에서 동정된 mRNA에 상응하고(실시예 6) 609 아미노산 단백질(pLy111b$_{\text{fullA}}$; 도 5와 6, 서열번호 42-43)을 암호화한다. 제 2 (Ly111b$_{\text{fullB}}$)는 아마도 공통 1차 mRNA의 대안적인 스플라이싱의 산물을 나타낸다. Ly111b$_{\text{fullA}}$ 와 동일한 이 전사체에서, 인간의 폐에서 확인된 서열의 뉴클레오타이드 752 내지 956 간의 서열은 부재한다(서열번호 42). 따라서, Ly111b$_{\text{fullB}}$ 는 pLy111$_{\text{fullA}}$ 와 동일한 541 아미노산 단백질(pLy111b$_{\text{fullB}}$)를 암호화하지만, 여기에서 아미노산 172 내지 240 간의 도메인(도 5와 7, 서열번호 44-45)은 소실되었다. 두 단백질 pLy111b$_{\text{fullA}}$/$_{\text{fullB}}$ 는 효모에서 확인된(초기 서열 Ly111b, 도 5), 아미노산 135-290을 포함하는 파킨 단편과의 상호작용 도메인을 통합시키며, 따라서 이론적으로는 이러한 상호작용을 유지시킬 수 있다.

pLy111b$_{\text{fullA}}$/$_{\text{fullB}}$ 단백질은 RIM/랩필린 계열에 속한다.

pLy111b$_{\text{fullA}}$/$_{\text{fullB}}$ 는 RIM/랩필린 계열(Wang Y, Sugita S & Sudhof TG. The RIM/NIM family of neuronal C2 domain proteins. J Biol Chem (2000) 275, 20033-20044)의 단백질과 및 특히 그래뉼로필린(Wang Jie, Takeuchi T, Yokota H & Izumi T. Novel Rabphilin-3-like protein associates with insulin-containing granules in pancreatic beta cells. J Biol Chem (1999) 274, 28542-28548)과 상동성을 보인다. 이들은 N-말단 부분내 징크 핑거 도메인 및 C-말단 부분내 두 C_2 도메인의 존재를 특징으로 한다(도 6과 7). RIM/랩필린 계열 단백질의 징크 핑거 도메인은 Rab 단백질과의 상호작용에 연관되어 있다. GTP-결합 단백질인 후자는 진핵생물 세포에서 막 트래피킹 기구의 필수 성분이다.

또한, RIM/랩필린 계열 단백질의 C_2 도메인이 인지질과의 상호작용을 통해 막에 결합할 수 있음이 기술되어 있다.

cos-7 계통의 세포에서 pLy111b$_{\text{fullA}}$/$_{\text{fullB}}$ 단백질의 발현: 파킨으로 동시 위치선정

Ly111b$_{\text{fullA/B}}$ 전사체의 암호화 서열을 진핵생물 발현 벡터 pcDNA3 중으로 N-말단 myc 에피토프를 암호화하는 서열(pcDNA3-mycLy111b$_{\text{fullA/B}}$)을 갖는 프레임에 삽입시킨다. 이들 벡터를 사용하여 형질감염시킨 cos-7 계통의 세포는 예상된 분자량에 해당하는, 대략 67 kDa (pcDNA3-mycLy111b$_{\text{fullA}}$) 및 60 kDa (pcDNA3-mycLy111b$_{\text{fullB}}$)의 겉보기 분자량을 갖는 단백질을 생성한다. N-말단 myc 에피토프에 대한 항체를 사용하여 면역표지함으로써 드러난 이들 단백질은 작은 반점 방식으로, cos-7 계통 세포의 세포질, 익스텐션 및 종종 핵에 비균질 분포한다(도 8a, b, 칼럼 A). cos-7 계통의 세포에서, 안티-파킨 항체 Asp5를 사용하여 드러나듯, 이들이 파킨으로 과잉발현될 때, 유사 분포는 관찰될 수 있는 이들 단백질의 동시 위치추정이다(도 8a, b, 칼럼 c).

참조문헌

Abbas, N. *et al.* (1999). A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. *Hum Mol Genet.* **8**, 567-574.

Bartel, P.L. *et al.* (1993). D.A. Hartley Ed, Oxford University Press, 153

Chung, C.T. *et al.* (1989). One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution, *Proc. Natl. Acad. Sci. USA.* **86**, 2172-2175.

Fields, S. and **Song,** O. (1989). A novel genetic system to detect protein-protein interactions. *Nature.* **340,** 245-246.

Gietz, R.D. *et al.* (1992). Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.,* **20**, 1425.

Hattori, N. *et al.* (1998). Molecular genetic analysis of a novel Parkin gene in Japanese families with autosomal recessive juvenile parkinsonism: evidence for variable homozygous deletions in the Parkin gene in affected individuals. *Ann Neurol.* **6**, 935-941.

Kahn, A. *et al.,* (1991) Thérapie génique: espoirs et limites, *Médecine et Sciences.* **7**, 705-714.

Kitada, T. *et al.* (1998). Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature.* **392,** 605-608.

Leroy, E. *et al.* (1998). The ubiquitin pathway in Parkinson's disease. *Nature.* **395**, 451-452.

Lewy, F.H. (1912). in *Handbuch der Neurologie* (Lewandowski, M., ed) pp 920-933, Springer, Berlin.

Lucking, C. *et al.* (1998). Homozygous deletions in parkin gene in European and North African families with autosomal recessive juvenile parkinsonism. *Lancet.* **352,** 1355-1356.

Maniatis, T. *et al.* (1989). Molecular cloning, second edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Morett, E. (1999). A novel transactivation domain in parkin. *Trends Biochem Sci.* **24**, 229-231.

Peytavi, R. *et al.* (1999). HEED, the product of the human homolog of the murine eed gene, binds to the matrix protein of HIV-1. *J. Biol. Chem.* **274,** 1635-1645.

Polymeropoulos, M.H. *et al.* (1997). Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science.* **276**, 2045-2047.

Sanger, F. *et al.* (1997). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467.

Shimura, H. *et al.* (1999). Immunohistochemical and subcellular localization of Parkin protein: absence of protein in autosomal recessive juvenile parkinsonism patients. *Ann Neurol.* **45**, 668-672.

Sunada, Y. *et al.* (1998), Differential expression of the parkin gene in the human brain and peripheral leukocytes. *Neurosci Lett.* **254**, 180-182.

Vojtek, AB. *et al.*, (1993). Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell.* **74**, 205-214.

(57) 청구의 범위

청구항 1.

적어도 부분적으로, PAP1 단백질, 또는 이 단백질의 상동체와 파킨 간의 상호작용을 조절할 수 있는 화합물.

청구항 2.

제 1 항에 있어서, 적어도 부분적으로 상호작용을 늦추거나, 억제하거나 자극함을 특징으로 하는 화합물.

청구항 3.

제 1 항 또는 2 항에 있어서, PAP1 단백질, 또는 이 단백질의 상동체와 파킨 간의 상호작용 도메인과 결합할 수 있음을 특징으로 하는 화합물.

청구항 4.

제 1 항 내지 3 항 중 어느 한 항에 있어서, 펩타이드, 핵산, 지질 또는 사카라이드형 화합물, 또는 항체임을 특징으로 하는 화합물.

청구항 5.

제 4 항에 있어서, 서열번호 2의 펩타이드 서열 또는 이의 유도체의 전부 또는 일부를 포함하는 펩타이드 화합물임을 특징으로 하는 화합물.

청구항 6.

제 4 항에 있어서, 서열이 PAP1 단백질과 파킨의 상호작용 부위의 전부 또는 기능성 부분에 상응하는 영역을 포함하는 펩타이드 화합물임을 특징으로 하는 화합물.

청구항 7.

제 4 항에 있어서, PAP1 단백질로부터 (및/또는 상동형으로부터) 유래하고, 비기능성으로 만든 이펙터 영역을 보유하는 펩타이드 화합물임을 특징으로 하는 화합물.

청구항 8.

서열번호 2의 서열 또는 이 서열의 유도체나 단편을 포함하는 폴리펩타이드.

청구항 9.

제 8 항에 있어서, 서열번호 2의 서열의 적어도 5개 연속 잔기, 바람직하게는 적어도 9개, 더욱 바람직하게는 적어도 15개를 포함하는 폴리펩타이드.

청구항 10.

제 8 항에 있어서, 서열번호 13, 15, 43 또는 45의 서열 또는 이들 서열의 변이체의 전부 또는 일부, 특히 서열번호 13, 15, 43 또는 45의 서열의 적어도 5개의 연속 잔기, 바람직하게는 적어도 9개, 더욱 바람직하게는 적어도 15개의 연속 잔기를 포함하는 폴리펩타이드.

청구항 11.

제 4 항 내지 10 항 중 어느 한 항에 따른 펩타이드 화합물을 암호화하는 핵산.

청구항 12.

제 11 항에 있어서, 서열번호 1, 12, 14, 42 또는 44의 서열 또는 이 서열로부터 유도되는 서열의 전부 또는 일부를 포함함을 특징으로 하는 핵산.

청구항 13.

제 8 항 또는 11 항에 따른 폴리펩타이드를 암호화하는 핵산.

청구항 14.

제 11 항 내지 13 항 중 어느 한 항에 따른 핵산, 또는 이의 상보 스트랜드와 하이브리드화할 수 있는, 핵산, 특히 뉴클레오타이드 탐침.

청구항 15.

제 11 항 내지 14 항 중 어느 한 항에 따른 핵산을 포함하는 벡터.

청구항 16.

제 11 항 내지 14 항 중 어느 한 항에 따른 핵산을 포함하는 결손형 재조합 바이러스.

청구항 17.

서열번호 16 내지 41 및 46의 서열의 핵산으로부터 선택된 핵산.

청구항 18.

제 4 항 내지 10 항 중 어느 한 항에 따른 펩타이드 화합물에 대한 것임을 특징으로 하는 항체 또는 항체 단편 또는 유도체.

청구항 19.

제 18 항에 있어서, 제 9 항 또는 10 항에 따른 폴리펩타이드를 인식함을 특징으로 하는 항체.

청구항 20.

제 1 항 내지 10 항 중 어느 한 항에 따른 적어도 하나의 화합물, 또는 제 18 항 또는 19 항에 따른 항체를 포함하는 약학 조성물.

청구항 21.

적어도 부분적으로, PAP1 단백질, 또는 이 단백질의 상동체와 파킨의 상호작용을 조절할 수 있는, 비 - 펩타이드 화합물 또는 배타적으로 펩타이드 종이 아닌 화합물.

청구항 22.

제 21 항에 있어서, 제 5 항 내지 7 항 중 어느 한 항에 따른 펩타이드의 활성 모티프가 펩타이드가 아니거나 배타적으로 펩타이드 종이 아닌 구조로 복제됨을 특징으로 하는 화합물.

청구항 23.

제 11 항 내지 14 항 중 어느 한 항에 따른 적어도 하나의 핵산, 또는 제 15 항 또는 16 항에 따른 하나의 벡터를 포함하는 약학 조성물.

청구항 24.

제 4 항 내지 10 항 중 어느 한 항에 따른 펩타이드 화합물을 포함하는 약학 조성물.

청구항 25.

제 22 항, 23 항 또는 24 항에 있어서, 신경퇴행성 병리의 치료를 위한 조성물.

청구항 26.

서열번호 2의 서열 또는 서열번호 4의 서열, 또는 이들 서열의 단편과 결합할 수 있는 분자의 선별단계를 포함하는, 활성 분자의 스크리닝 또는 특징규명 방법.

청구항 27.

서열번호 13, 15, 43 및 45 중에서 선택된 서열, 또는 이들 서열의 단편과 결합할 수 있는 분자를 선별하는 단계를 포함하는, 활성 분자의 스크리닝 또는 특징규명 방법.

청구항 28.

제 11 항 내지 14 항 중 어느 한 항에 따른 핵산 또는 제 15 항 또는 16 항에 따른 벡터를 함유하는 세포를 핵산의 발현 조건하에서 배양하고, 생성된 펩타이드 화합물을 회수하는 단계를 포함하는, 제 4 항 내지 10 항 중 어느 한 항에 따른 펩타이드 화합물의 제조방법.

청구항 29.

분리된 형태의 인간 PAP1 단백질.

청구항 30.

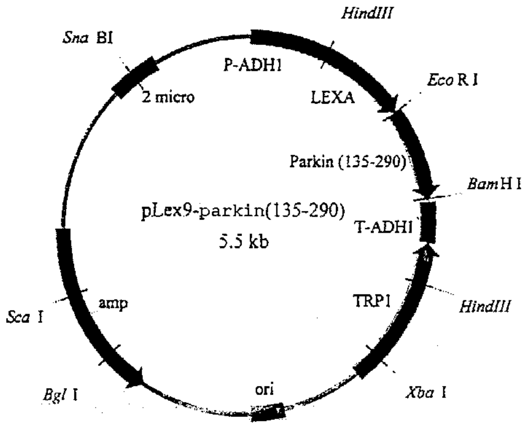
제 11 항 내지 14 항 중 어느 한 항에 따른 핵산 또는 제 15 항 또는 16 항에 따른 벡터를 함유하는 세포.

청구항 31.

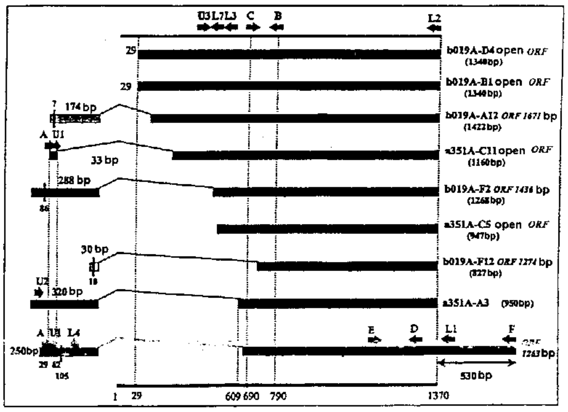
제 11 항 내지 14 항 중 어느 한 항에 따른 핵산을 세포에 포함하는 비 - 인간 포유동물.

도면

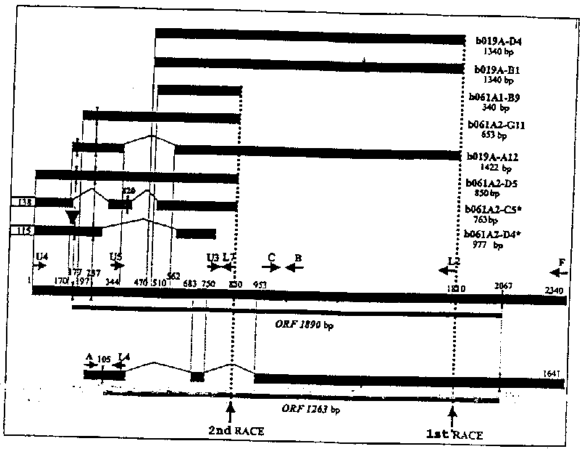
도면 1


HindIII
Sna BI
P-ADH1
Eco R I
2 micro
LEXA
Parkin (135-290)
Bam H I
pLex9-parkin(135-290)
5.5 kb
T-ADH1
amp
TRP1
HindIII
Sca I
Bgl I
ori
Xba I

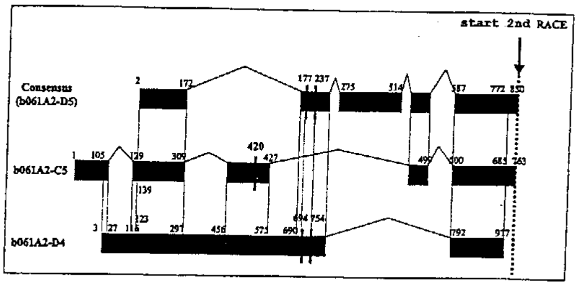
도면 2


U3 L7 L3
C
B
L2
29
b019A-D4 open ORF
(1340bp)
29
b019A-B1 open ORF
(1340bp)
7
174 bp
b019A-A12 ORF 1671 bp
(1422bp)
A U1
33 bp
a351A-C11 open ORF
(1160bp)
288 bp
86
b019A-F2 ORF 1436 bp
(1268bp)
a351A-C5 open ORF
(947bp)
30 bp
18
b019A-F12 ORF 1274 bp
(827bp)
U2
320 bp
a351A-A3 (950bp)
A U1 L4
250bp
29 62
105
E
D
L1
F
ORF
1263 bp
530 bp
1
29
609
690
790
1370

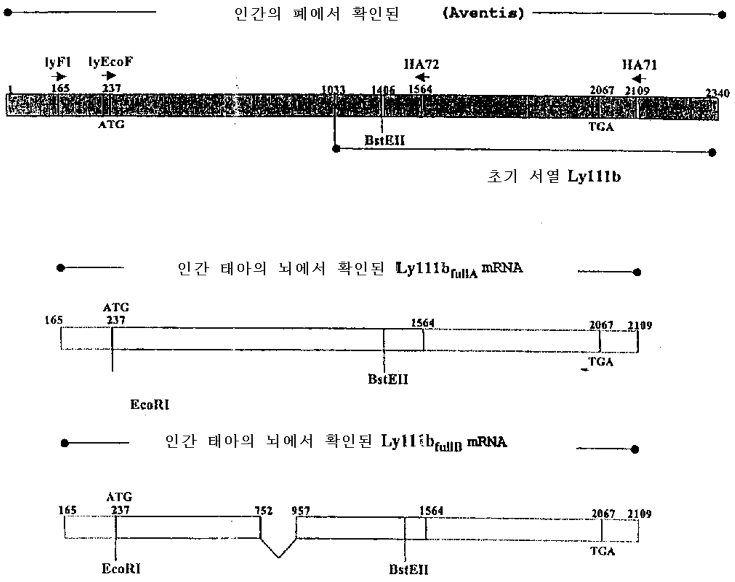
도면 3


b019A-D4
1340 bp
b019A-B1
1340 bp
b061A1-B9
340 bp
b061A2-G11
653 bp
b019A-A12
1422 bp
b061A2-D5
850 bp
b061A2-C5*
763 bp
b061A2-D4*
977 bp
138
115
420
U4
U5
U3
L7
C
B
L2
F
1
170
177
197
237
344
470
510
562
683
750
850
953
1810
2067
2340
ORF 1890 bp
A
105
L4
1641
ORF 1263 bp
2nd RACE
1st RACE

도면 4


start 2nd RACE
Consensus
(b061A2-D5)
2
172
177
237
275
514
587
772
850
b061A2-C5
1
105
129
309
420
427
499
500
685
763
139
b061A2-D4
3
27
116
123
297
456
575
690
694
754
792
977

도면 5



도면 6

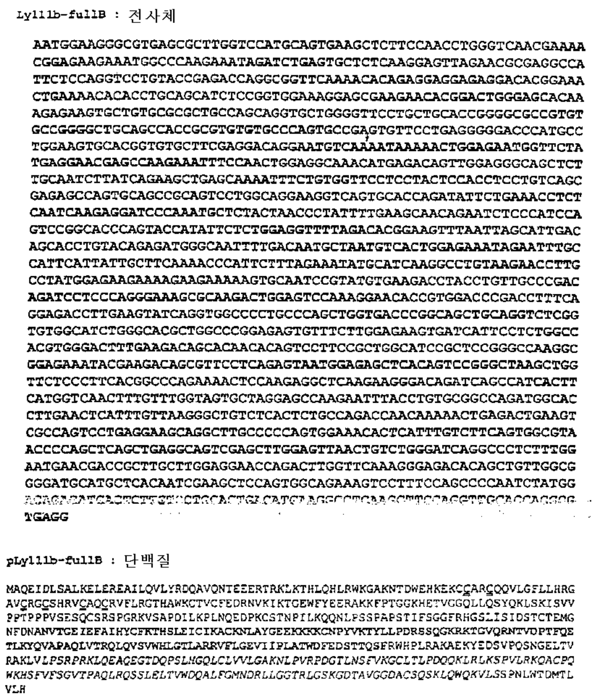
Ly111b-fullA : 전사체

AATGGAAGGGCGTGAGCGCTTGGTCCATGCAGTGAAGCTCTTCCAACCTGGGTCAACGAAAACG
GAGAAGAAATGGCCCAAGAAATAGATCTGAGTGCTCTCAAGGAGTTAGAACGCGAGGCCATTCT
CCAGGTCCTGTACCGAGACCAGGCGGTTCAAAACACAGAGGAGGAGAGGACACGGAAACTGAAA
ACACACCTGCAGCATCTCCGGTGGAAAGGAGCGAAGAACACGGACTGGGAGCACAAAGAGAAGT
GCTGTGCGCGCTGCCAGCAGGTGCTGGGGTTCCTGCTGCACCGGGGCGCCGTGTGCCGGGGCTG
CAGCCACCGCGTGTGTGCCCAGTGCCGAGTGTTCCTGAGGGGGACCCATGCCTGGAAGTGCACG
GTGTGCTTCGAGGACAGGAATGTCAAAATAAAAACTGGAGAATGGTTCTATGAGGAACGAGCCA
AGAAATTTCCAACTGGAGGCAAACATGAGACAGTTGGAGGGCAGCTCTTGCAATCTTATCAGAA
GCTGAGCAAAATTTCTGTGGTTCCTCCTACTCCACCTCCTGTCAGCGAGAGCCAGTGCAGCCGC
AGTCCTGGCAGGTTACAGGAATTTGGTCAGTTTAGAGGATTTAATAAGTCCGTGGAAAATTTGT
TTCTGTCTCTTGCTACCCACGTGAAAAAGCTCTCCAAATCCCAGAATGATATGACTTCTGAGAA
GCATCTTCTCGCCACGGGCCCCAGGCAGTGTGTGGGACAGACAGAGAGACGGAGCCAGTCTGAC
ACTGCGGTCAACGTCACCACCAGGAAGGTCAGTGCACCAGATATTCTGAAACCTCTCAATCAAG
AGGATCCCAAATGCTCTACTAACCCTATTTTGAAGCAACAGAATCTCCCATCCAGTCCGGCACC
CAGTACCATATTCTCTGGAGGTTTTAGACACGGAAGTTTAATTAGCATTGACAGCACCTGTACA
GAGATGGGCAATTTTGACAATGCTAATGTCACTGGAGAAATAGAATTTGCCATTCATTATTGCT
TCAAAACCCATTCTTTAGAAATATGCATCAAGGCCTGTAAGAACCTTGCCTATGGAGAAGAAAA
GAAGAAAAAGTGCAATCCGTATGTGAAGACCTACCTGTTGCCCGACAGATCCTCCCAGGGAAAG
CGCAAGACTGGAGTCCAAAGGAACACCGTGGACCCGACCTTTCAGGAGACCTTGAAGTATCAGG
TGGCCCCTGCCCAGCTGGTGACCCGGCAGCTGCAGGTCTCGGTGTGGCATCTGGGCACGCTGGC
CCGGAGAGTGTTTCTTGGAGAAGTGATCATTCCTCTGGCCACGTGGGACTTTGAAGACAGCACA
ACACAGTCCTTCCGCTGGCATCCGCTCCGGGCCAAGGCGGAGAAATACGAAGACAGCGTTCCTC
AGAGTAATGGAGAGCTCACAGTCCGGGCTAAGCTGGTTCTCCCTTCACGGCCCAGAAAACTCCA
AGAGGCTCAAGAAGGGACAGATCAGCCATCACTTCATGGTCAACTTTGTTTGGTAGTGCTAGGA
GCCAAGAATTTACCTGTGCGGCCAGATGGCACCTTGAACTCATTTGTTAAGGGCTGTCTCACTC
TGCCAGACCAACAAAAACTGAGACTGAAGTCGCCAGTCCTGAGGAAGCAGGCTTGCCCCCAGTG
GAAACACTCATTTGTCTTCAGTGGCGTAACCCCAGCTCAGCTGAGGCAGTCGAGCTTGGAGTTA
ACTGTCTGGGATCAGGCCCTCTTTGGAATGAACGACCGCTTGCTTGGAGGAACCAGACTTGGTT
CAAAGGGAGACACAGCTGTTGGCGGGGATGCATGCTCACAATCGAAGCTCCAGTGGCAGAAAGT
CCTTTCCAGCCCCAATCTATGGACAGACATGACTCTTGTCCTGCACTGACATGAAGGCCTCAAG
GTTCCAGGTTGCAGCAGGCGTGAGG

pLy111b-fullA : 단백질

MAQEIDLSALKELEREAILQVLYRDQAVQNTEEERTRKLKTHLQHLRWKGAKNTDWEHKEKCCARCQQVLGFLLHRG
AVCRGCSHRVCAQCRVFLRGTHAWKCTVCFEDRNVKIKTGEWFYEERAKKFPTGGKHETVGGQLLQSYQKLSKISVV
PPTPPPVSESQCSRSPGRLQEFGQFRGFNKSVENLFLSLATHVKKLSKSQNDMTSEKHLLATGPRQCVGQTERRSQS
DTAVNVTTRKVSAPDILKPLNQEDPKCSTNPILKQQNLPSSPAPSTIFSGGFRHGSLISIDSTCTEMGNFDNANVTG
EIEFAIHYCFKTHSLEICIKACKNLAYGEEKKKKCNPYVKTYLLPDRSSQGKRKTGVQRNTVDPTFQETLKYQVAPA
QLVTRQLQVSVWHLGTLARRVFLGEVIIPLATWDFEDSTTQSFRWHPLRAKAEKYEDSVPQSNGELTVRAKLV*LPSR*
PRKLQEAQEGTDQPSLHGQLCLVVLGAKNLPVRPDGTLNSFVKGCLTLPDQQKLRLKSPVLRKQACPQWKHSFVFSG
*VTPAQLRQSSLELTVWDQALFGMNDRLLGGTRLGSKGDTAVGGDACSQSKLQWQKVLSS*PNLWTDMTLVLH

도면 7



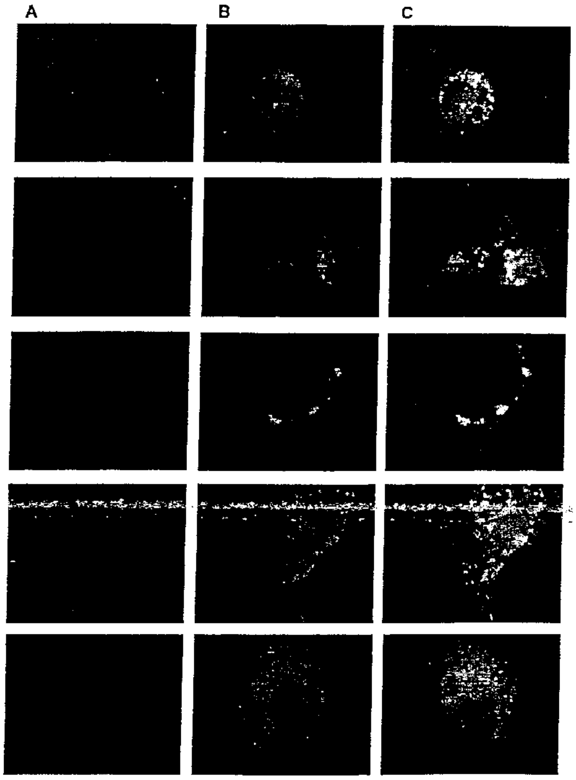
Ly111b-full8 : 전사체

```
AATGGAAGGGCGTGAGCGCTTGGTCCATGCAGTGAAGCTCTTCCAACCTGGGTCAACGAAAA
CGGAGAAGAAATGGCCCAAGAAATAGATCTGAGTGCTCTCAAGGAGTTAGAACGCGAGGCCA
TTCTCCAGGTCCTGTACCGAGACCAGGCGGTTCAAAACACAGAGGAGGAGAGGACACGGAAA
CTGAAAACACACCTGCAGCATCTCCGGTGGAAAGGAGCGAAGAACACGGACTGGGAGCACAA
AGAGAAGTGCTGTGCGCGCTGCCAGCAGGTGCTGGGGTTCCTGCTGCACCGGGGCGCCGTGT
GCCGGGGCTGCAGCCACCGCGTGTGTGCCCAGTGCCGAGTGTTCCTGAGGGGGACCCATGCC
TGGAAGTGCACGGTGTGCTTCGAGGACAGGAATGTCAAAATAAAAACTGGAGAATGGTTCTA
TGAGGAACGAGCCAAGAAATTTCCAACTGGAGGCAAACATGAGACAGTTGGAGGGCAGCTCT
TGCAATCTTATCAGAAGCTGAGCAAAATTTCTGTGGTTCCTCCTACTCCACCTCCTGTCAGC
GAGAGCCAGTGCAGCCGCAGTCCTGGCAGGAAGGTCAGTGCACCAGATATTCTGAAACCTCT
CAATCAAGAGGATCCCAAATGCTCTACTAACCCTATTTTGAAGCAACAGAATCTCCCATCCA
GTCCGGCACCCAGTACCATATTCTCTGGAGGTTTTAGACACGGAAGTTTAATTAGCATTGAC
AGCACCTGTACAGAGATGGGCAATTTTGACAATGCTAATGTCACTGGAGAAATAGAATTTGC
CATTCATTATTGCTTCAAAACCCATTCTTTAGAAATATGCATCAAGGCCTGTAAGAACCTTG
CCTATGGAGAAGAAAAGAAGAAAAAGTGCAATCCGTATGTGAAGACCTACCTGTTGCCCGAC
AGATCCTCCCAGGGAAAGCGCAAGACTGGAGTCCAAAGGAACACCGTGGACCCGACCTTTCA
GGAGACCTTGAAGTATCAGGTGGCCCCTGCCCAGCTGGTGACCCGGCAGCTGCAGGTCTCGG
TGTGGCATCTGGGCACGCTGGCCCGGAGAGTGTTTCTTGGAGAAGTGATCATTCCTCTGGCC
ACGTGGGACTTTGAAGACAGCACAACACAGTCCTTCCGCTGGCATCCGCTCCGGGCCAAGGC
GGAGAAATACGAAGACAGCGTTCCTCAGAGTAATGGAGAGCTCACAGTCCGGGCTAAGCTGG
TTCTCCCTTCACGGCCCAGAAAACTCCAAGAGGCTCAAGAAGGGACAGATCAGCCATCACTT
CATGGTCAACTTTGTTTGGTAGTGCTAGGAGCCAAGAATTTACCTGTGCGGCCAGATGGCAC
CTTGAACTCATTTGTTAAGGGCTGTCTCACTCTGCCAGACCAACAAAAACTGAGACTGAAGT
CGCCAGTCCTGAGGAAGCAGGCTTGCCCCCAGTGGAAACACTCATTTGTCTTCAGTGGCGTA
ACCCCAGCTCAGCTGAGGCAGTCGAGCTTGGAGTTAACTGTCTGGGATCAGGCCCTCTTTGG
AATGAACGACCGCTTGCTTGGAGGAACCAGACTTGGTTCAAAGGGAGACACAGCTGTTGGCG
GGGATGCATGCTCACAATCGAAGCTCCAGTGGCAGAAAGTCCTTTCCAGCCCCAATCTATGG
AGAGAGATGACTCTTGTTCTGCACTGACATGTAGGCCTCAAGCTTCCAGGTTGCACCAGGCG
TGAGG
```

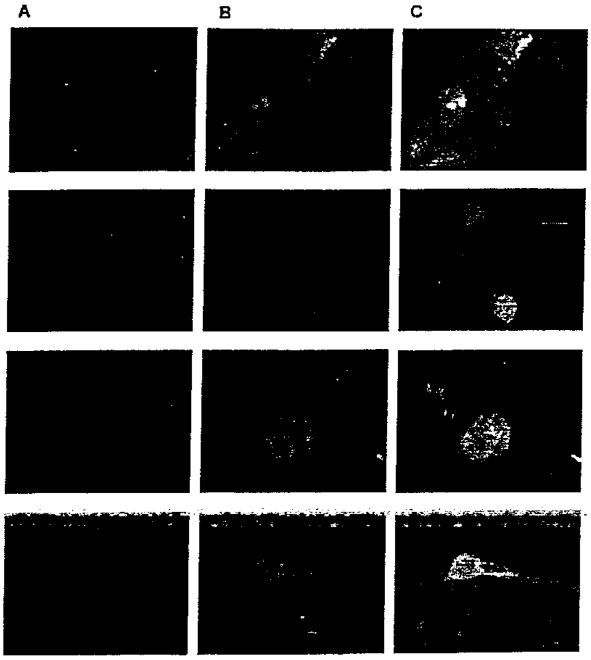
pLy111b-full8 : 단백질

```
MAQEIDLSALKELEREAILQVLYRDQAVQNTEEERTRKLKTHLQHLRWKGAKNTDWEHKEKCCARCQQVLGFLLHRG
AVCRGCSHRVCAQCRVFLRGTHAWKCTVCFEDRNVKIKTGEWFYEERAKKFPTGGKHETVGGQLLQSYQKLSKISVV
PPTPPPVSESQCSRSPGRKVSAPDILKPLNQEDPKCSTNPILKQQNLPSSPAPSTIFSGGFRHGSLISIDSTCTEMG
NFDNANVTGEIEFAIHYCFKTHSLEICIKACKNLAYGEEKKKKCNPYVKTYLLPDRSSQGKRKTGVQRNTVDPTFQE
TLKYQVAPAQLVTRQLQVSVWHLGTLARRVFLGEVIIPLATWDFEDSTTQSFRWHPLRAKAEKYEDSVPQSNGELTV
RAKLVLPSRPRKLQEACEGTDQPSLHGQLCLVVLGAKNLPVRPDGTLNSFVKGCLTLPDQQKLRLKSPVLRKQACPQ
WKHSFVFSGVTPAQLRQSSLELTVWDQALFGMNDRLLGGTRLGSKGDTAVGGDACSQSKLQWQKVLSSPNLWTDMTL
VLH
```

도면 8a



도면 8b



도면 9

```
GGCCTTGGGGCACTGAGGGATGCCAGTTCTGCCTGTTCATCTGGAACCTGGATCTAAGGAGGGAAGAG
GCGTTGCCCCTGCTGGCATAGTCAGGTACCAGCCCAGCCAGGTATTGAACGGGCTGAGCTTTTCATGA
TGGTTCCTGCTGACCTGGAAACATCTTAAATGGAAGGGCGTGAGCGCTTGGTCCATGCAGTGAAGCTC
TTCCAACCTGGGTCAACGAAAACGGAGAAGAA[redacted]GCCCAAGAAATAGATCTGAGTGCTCTCAAGGAG
TTAGAACGCGAGGCCATTCTCCAGGTCCTGTACCGAGACCAGGCGGTTCAAAACACAGAGGAGGAGAG
GACACGGAAACTGAAAACACACCTGCAGCATCTCCGGTGGAAAGGAGCGAAGAACACGGACTGGGAGC
ACAAAGAGAAGTGCTGTGCGCGCTGCCAGCAGGTGCTGGGGTTCCTGCTGCACCGGGGCGCCGTGTGC
CGGGGCTGCAGCCACCGCGTGTGTGCCCAGTGCCGAGTGTTCCTGAGGGGGACCCATGCCTGGAAGTG
CACGGTGTGCTTCGAGGACAGGAATGTCAAAATAAAAACTGGAGAATGGTTCTATGAGGAACGAGCCA
AGAAATTTCCAACTGGAGGCAAACATGAGACAGTTGGAGGGCAGCTCTTGCAATCTTATCAGAAGCTG
AGCAAAATTTCTGTGGTTCCTCCTACTCCACCTCCTGTCAGCGAGAGCCAGTGCAGCCGCAGTCCTGG
CAGGTTACAGGAATTTGGTCAGTTTAGAGGATTTAATAAGTCCGTGGAAAATTTGTTTCTGTCTCTTG
CTACCCACGTGAAAAAGCTCTCCAAATCCCAGAATGATATGACTTCTGAGAAGCATCTTCTCGCCACG
GGCCCCAGGCAGTGTGTGGGACAGACAGAGAGACGGAGCCAGTCTGACACTGCGGTCAACGTCACCAC
CAGGAAGGTCAGTGCACCAGATATTCTGAAACCTCTCAATCAAGAGGATCCCAAATGCTCTACTAACC
CTATTTTGAAGCAACAGAATCTCCCATCCAGTCCGGCACCCAGTACCATATTCTCTGGAGGTTTTAGA
CACGGAAGTTTAATTAGCATTGACAGCACCTGTACAGAGATGGGCAATTTTGACAATGCTAATGTCAC
TGGAGAAATAGAATTTGCCATTCATTATTGCTTCAAAACCCATTCTTTAGAAATATGCATCAAGGCCT
GTAAGAACCTTGCCTATGGAGAAGAAAAGAAGAAAAAGTGCAATCCGTATGTGAAGACCTACCTGTTG
CCCGACAGATCCTCCCAGGGAAAGCGCAAGACTGGAGTCCAAAGGAACACCGTGGACCCGACCTTTCA
GGAGACCTTGAAGTATCAGGTGGCCCCTGCCCAGCTGGTGACCCGGCAGCTGCAGGTCTCGGTGTGGC
ATCTGGGCACGCTGGCCCGGAGAGTGTTTCTTGGAGAAGTGATCATTCCTCTGGCCACGTGGGACTTT
GAAGACAGCACAACACAGTCCTTCCGCTGGCATCCGCTCCGGGCCAAGGCGGAGAAATACGAAGACAG
CGTTCCTCAGAGTAATGGAGAGCTCACAGTCCGGGCTAAGCTGGTTCTCCCTTCACGGCCCAGAAAAC
TCCAAGAGGCTCAAGAAGGGACAGATCAGCCATCACTTCATGGTCAACTTTGTTTGGTAGTGCTAGGA
GCCAAGAATTTACCTGTGCGGCCAGATGGCACCTTGAACTCATTTGTTAAGGGCTGTCTCACTCTGCC
AGACCAACAAAAACTGAGACTGAAGTCGCCAGTCCTGAGGAAGCAGGCTTGCCCCCAGTGGAAACACT
CATTTGTCTTCAGTGGCGTAACCCCAGCTCAGCTGAGGCAGTCGAGCTTGGAGTTAACTGTCTGGGAT
CAGGCCCTCTTTGGAATGAACGACCGCTTGCTTGGAGGAACCAGACTTGGTTCAAAGGGAGACACAGC
TGTTGGCGGGGATGCATGCTCACAATCGAAGCTCCAGTGGCAGAAAGTCCTTTCCAGCCCCAATCTAT
GGACAGACATGACTCTTGTCCTGCACTGACATGAAGGCCTCAAGGTTCCAGGTTGCAGCAGGCGTGAG
GCACTGTGCGTCTGCAGAGGGGCTACGAACCAGGTGCAGGGTCCCAGCTGGAGACCCCTTTGACCTTG
AGCAGTCTCCATCTGCGGCCCTGTCCCATGGCTTAACCGCCTATTGGTATCTGTGTATATTTACGTTA
AACACAATTATGTTACCTAAGCCTCTGGTGGGTAATCTGGACTTATGAG[illegible]TGCCCAGATT
TT[redacted]CGTTGTTACCCATGAAAAAAAAAAAAA
```

단백질 서열

```
MAQEIDLSALKELEREAILQVLYRDQAVQNTEEERTRKLKTHLQHLRWKGAKNTDWEHKEKCCARCQQ
VLGFLLHRGAVCRGCSHRVCAQCRVFLRGTHAWKCTVCFEDRNVKIKTGEWFYEERAKKFPTGGKHET
VGGQLLQSYQKLSKISVVPPTPPPVSESQCSRSPGRLQEFGQFRGFNKSVENLFLSLATHVKKLSKSQ
NDMTSEKHLLATGPRQCVGQTERRSQSDTAVNVTTRKVSAPDILKPLNQEDPKCSTNPILKQQNLPSS
PAPSTIFSGGFRHGSLISIDSTCTEMGNFDNANVTGEIEFAIHYCFKTHSLEICIKACKNLAYGEEKK
KKCNPYVKTYLLPDRSSQGKRKTGVQRNTVDPTFQETLKYQVAPAQLVTRQLQVSVWHLGTLARRVFL
GEVIIPLATWDFEDSTTQSFRWHPLRAKAEKYEDSVPQSNGELTVRAKLVLPSRPRKLQEAQEGTDQP
SLHGQLCLVVLGAKNLPVRPDGTLNSFVKGCLTLPDQQKLRLKSPVLRKQACPQWKHSFVFSGVTPAQ
LRQSSLELTVWDQALFGMNDRLLGGTRLGSKGDTAVGGDACSQSKLQWQKVLSSPNLWTDMTLVLH
```

도면 10

GAAATCATGCCCCTCGTAGAGCAGCAGGTCCAAGCAGGGCTGCTGGCTATTTTTCCAAAAAG
TGAGGCAGTTTAAAAAAAAAGGCGGAGAACTAGAATTATAGAATAATGGCACATTTTGTGTAT
TTGTAAAACTAACGGCTTGCATGGTTCACAACCCATTTCTTATGCCTGTGTTTTCCTTGGCA
GCAAAATTTCTGTGGTTCCTCCTACTCCACCTCCTGTCAGCGAGAGCCAGTGCAGCCGCAGT
CCTGGCAGGAAGGTCAGTGCACCAGATATTCTGAAACCTCTCAATCAAGAGGATCCCAAATG
CTCTACTAACCCTATTTTGAAGCAACAGAATCTCCCATCCAGTCCGGCACCCAGTACCATAT
TCTCTGGAGGTTTTAGACACGGAAGTTTAATTAGCATTGACAGCACCTGTACAGAG[redacted]GGC
AATTTTGACAATGCTAATGTCACTGGAGAAATAGAATTTGCCATTCATTATTGCTTCAAAAC
CCATTCTTTAGAAATATGCATCAAGGCCTGTAAGAACCTTGCCTATGGAGAAGAAAAGAAGA
AAAAGTGCAATCCGTATGTGAAGACCTACCTGTTGCCCGACAGATCCTCCCAGGGAAAGCGC
AAGACTGGAGTCCAAAGGAACACCGTGGACCCGACCTTTCAGGAGACCTTGAAGTATCAGGT
GGCCCCTGCCCAGCTGGTGACCCGGCAGCTGCAGGTCTCGGTGTGGCATCTGGGCACGCTGG
CCCGGAGAGTGTTTCTTGGAGAAGTGATCATTCCTCTGGCCACGTGGGACTTTGAAGACAGC
ACAACACAGTCCTTCCGCTGGCATCCGCTCCGGGCCAAGGCGGAGAAATACGAAGACAGCGT
TCCTCAGAGTAATGGAGAGCTCACAGTCCGGGCTAAGCTGGTTCTCCCTTCACGGCCCAGAA
AACTCCAAGAGGCTCAAGAAGGGACAGATCAGCCATCACTTCATGGTCAACTTTGTTTGGTA
GTGCTAGGAGCCAAGAATTTACCTGTGCGGCCAGATGGCACCTTGAACTCATTTGTTAAGGG
CTGTCTCACTCTGCCAGACCAACAAAAACTGAGACTGAAGTCGCCAGTCCTGAGGAAGCAGG
CTTGCCCCCAGTGGAAACACTCATTTGTCTTCAGTGGCGTAACCCCAGCTCAGCTGAGGCAG
TCGAGCTTGGAGTTAACTGTCTGGGATCAGGCCCTCTTTGGAATGAACGACCGCTTGCTTGG
AGGAACCAGACTTGGTTCAAAGGGAGACACAGCTGTTGGCGGGGATGCATGCTCACAATCGA
AGCTCCAGTGGCAGAAAGTCCTTTCCAGCCCCAATCTATGGACAGACATGACTCTTGTCCTG
CACTGACATGAAGGCCTCAAGGTTCCAGGTTGCAGCAGGCGTGAGGCACTGTGCGTCTGCAG
AGGGGCTACGAACCAGGTGCAGGGTCCCAGCTGGAGACCCCTTTGACCTTGAGCAGTCTCCA
TCTGCGGCCCTGTCCCATGGCTTAACCGCCTATTGGTATCTGTGTATATTTACGTTAAACAC
AATTATGTTACCTAAGCCTCTGGTGGGTTATCTCCTCTTTGAGATGTAGAAAATGGCCAGAT
TTT[redacted]CGTTGTTACCCATGAAAAAAAAAAAAAA

단백질 서열

MGNFDNANVTGEIEFAIHYCFKTHSLEICIKACKNLAYGEEKKKKCNPYVKTYLLPDRSSQG
KRKTGVQRNTVDPTFQETLKYQVAPAQLVTRQLQVSVWHLGTLARRVFLGEVIIPLATWDFE
DSTTQSFRWHPLRAKAEKYEDSVPQSNGELTVRAKLVLPSRPRKLQEAQEGTDQPSLHGQLC
LVVLGAKNLPVRPDGTLNSFVKGCLTLPDQQKLRLKSPVLRKQACPQWKHSFVFSGVTPAQL
RQSSLELTVWDQALFGMNDRLLGGTRLGSKGDTAVGGDACSQSKLQWQKVLSSPNLWTDMTL
VLH

<110> AVENTIS PHARMA
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE M
<120> COMPOUNDS WHICH ARE CAPABLE OF MODULATING THE ACTIVITY OF PARKIN,
NUCLEOTIDE SEQUENCES AND USES
<130> PRJ00004
<160> 46
<170> KopatentIn 1.71
<210> 1
<211> 1313
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)..(1032)
<400> 1

```
cag aat ctc cca tcc agt ccg gca ccc agt acc ata ttc tct gga ggt           48
Gln Asn Leu Pro Ser Ser Pro Ala Pro Ser Thr Ile Phe Ser Gly Gly
  1               5                  10                  15
ttt aga cac gga agt tta att agc att gac agc acc tgt aca gag atg           96
Phe Arg His Gly Ser Leu Ile Ser Ile Asp Ser Thr Cys Thr Glu Met
             20                  25                  30
```

```
ggc aat ttt gac aat gct aat gtc act gga gaa ata gaa ttt gcc att          144
Gly Asn Phe Asp Asn Ala Asn Val Thr Gly Glu Ile Glu Phe Ala Ile
         35                  40                  45
cat tat tgc ttc aaa acc cat tct tta gaa ata tgc atc aag gcc tgt          192
His Tyr Cys Phe Lys Thr His Ser Leu Glu Ile Cys Ile Lys Ala Cys
     50                  55                  60
aag aac ctt gcc tat gga gaa gaa aag aag aaa aag tgc aat ccg tat          240
Lys Asn Leu Ala Tyr Gly Glu Glu Lys Lys Lys Lys Cys Asn Pro Tyr
 65                  70                  75                  80
gtg aag acc tac ctg ttg ccc gac aga tcc tcc cag gga aag cgc aag          288
Val Lys Thr Tyr Leu Leu Pro Asp Arg Ser Ser Gln Gly Lys Arg Lys
                 85                  90                  95
act gga gtc caa agg aac acc gtg gac ccg acc ttt cag gag acc ttg          336
Thr Gly Val Gln Arg Asn Thr Val Asp Pro Thr Phe Gln Glu Thr Leu
            100                 105                 110
aag tat cag gtg gcc cct gcc cag ctg gtg acc cgg cag ctg cag gtc          384
Lys Tyr Gln Val Ala Pro Ala Gln Leu Val Thr Arg Gln Leu Gln Val
        115                 120                 125
tcg gtg tgg cat ctg ggc acg ctg gcc cgg aga gtg ttt ctt gga gaa          432
Ser Val Trp His Leu Gly Thr Leu Ala Arg Arg Val Phe Leu Gly Glu
    130                 135                 140
gtg atc att tct ctg gcc acg tgg gac ttt gaa gac agc aca aca cag          480
Val Ile Ile Ser Leu Ala Thr Trp Asp Phe Glu Asp Ser Thr Thr Gln
145                 150                 155                 160
tcc ttc cgc tgg cat ccg ctc cgg gcc aag gcg gag aaa tac gaa gac          528
Ser Phe Arg Trp His Pro Leu Arg Ala Lys Ala Glu Lys Tyr Glu Asp
                165                 170                 175
agc gtt cct cag agt aat gga gag ctc aca gtc cgg gct aag ctg gtt          576
Ser Val Pro Gln Ser Asn Gly Glu Leu Thr Val Arg Ala Lys Leu Val
            180                 185                 190
ctc cct tca cgg ccc aga aaa ctc caa gag gct caa gaa ggg aca gat          624
Leu Pro Ser Arg Pro Arg Lys Leu Gln Glu Ala Gln Glu Gly Thr Asp
        195                 200                 205
cag cca tca ctt cat ggt caa ctt tgt ttg gta gtg cta gga gcc aag          672
Gln Pro Ser Leu His Gly Gln Leu Cys Leu Val Val Leu Gly Ala Lys
    210                 215                 220
aat tta cct gtg cgg cca gat ggc acc ttg aac tca ttt gtt aag ggc          720
Asn Leu Pro Val Arg Pro Asp Gly Thr Leu Asn Ser Phe Val Lys Gly
225                 230                 235                 240
tgt ctc act ctg cca gac caa caa aaa ctg aga ctg aag tcg cca gtc          768
Cys Leu Thr Leu Pro Asp Gln Gln Lys Leu Arg Leu Lys Ser Pro Val
                245                 250                 255
ctg agg aag cag gct tgc ccc cag tgg aaa cac tca ttt gtc ttc agt          816
Leu Arg Lys Gln Ala Cys Pro Gln Trp Lys His Ser Phe Val Phe Ser
            260                 265                 270
ggc gta acc cca gct cag ctg agg cag tcg agc ttg gag tta act gtc          864
Gly Val Thr Pro Ala Gln Leu Arg Gln Ser Ser Leu Glu Leu Thr Val
        275                 280                 285
tgg gat cag gcc ctc ttt gga atg aat gac cgc ttg ctt gga gga acc          912
Trp Asp Gln Ala Leu Phe Gly Met Asn Asp Arg Leu Leu Gly Gly Thr
```

```
    290                 295                 300
aga ctt ggt tca aag gga gac aca gct gtt ggc ggg gat gca tgc tca          960
Arg Leu Gly Ser Lys Gly Asp Thr Ala Val Gly Gly Asp Ala Cys Ser
305                 310                 315                 320
cta tcg aag ctc cag tgg cag aaa gtc ctt tcc agc ccc aat cta tgg         1008
Leu Ser Lys Leu Gln Trp Gln Lys Val Leu Ser Ser Pro Asn Leu Trp
                325                 330                 335
aca gac atg act ctt gtc ctg cac    tgacatga aggcctcaag gttccaggtt       1060
Thr Asp Met Thr Leu Val Leu His
            340
gcagcaggcg tgaggcactg tgcgtctgca gaggggctac gaaccaggtg cagggtccca       1120
gctggagacc cctttgacct tgagcagtct ccatctgcgg ccctgtccca tggcttaacc       1180
gcctattggt atctgtgtat atttacgtta aacacaatta tgttacctaa gcctctggtg       1240
ggttatctcc tctttgagat gtagaaaatg gccagatttt aataaacgtt gttacccatg       1300
aaaaaaaaaa aaa                                                          1313
<210>  2
<211>  344
<212>  PRT
<213>  Homo sapiens
<400>  2
Gln Asn Leu Pro Ser Ser Pro Ala Pro Ser Thr Ile Phe Ser Gly Gly
  1               5                  10                  15
Phe Arg His Gly Ser Leu Ile Ser Ile Asp Ser Thr Cys Thr Glu Met
             20                  25                  30
Gly Asn Phe Asp Asn Ala Asn Val Thr Gly Glu Ile Glu Phe Ala Ile
         35                  40                  45
His Tyr Cys Phe Lys Thr His Ser Leu Glu Ile Cys Ile Lys Ala Cys
     50                  55                  60
Lys Asn Leu Ala Tyr Gly Glu Glu Lys Lys Lys Lys Cys Asn Pro Tyr
 65                  70                  75                  80
Val Lys Thr Tyr Leu Leu Pro Asp Arg Ser Ser Gln Gly Lys Arg Lys
                 85                  90                  95
Thr Gly Val Gln Arg Asn Thr Val Asp Pro Thr Phe Gln Glu Thr Leu
            100                 105                 110
Lys Tyr Gln Val Ala Pro Ala Gln Leu Val Thr Arg Gln Leu Gln Val
        115                 120                 125
Ser Val Trp His Leu Gly Thr Leu Ala Arg Arg Val Phe Leu Gly Glu
    130                 135                 140
Val Ile Ile Ser Leu Ala Thr Trp Asp Phe Glu Asp Ser Thr Thr Gln
145                 150                 155                 160
Ser Phe Arg Trp His Pro Leu Arg Ala Lys Ala Glu Lys Tyr Glu Asp
                165                 170                 175
Ser Val Pro Gln Ser Asn Gly Glu Leu Thr Val Arg Ala Lys Leu Val
            180                 185                 190
Leu Pro Ser Arg Pro Arg Lys Leu Gln Glu Ala Gln Glu Gly Thr Asp
        195                 200                 205
Gln Pro Ser Leu His Gly Gln Leu Cys Leu Val Val Leu Gly Ala Lys
    210                 215                 220
Asn Leu Pro Val Arg Pro Asp Gly Thr Leu Asn Ser Phe Val Lys Gly
225                 230                 235                 240
```

```
Cys Leu Thr Leu Pro Asp Gln Gln Lys Leu Arg Leu Lys Ser Pro Val
                245                 250                 255
Leu Arg Lys Gln Ala Cys Pro Gln Trp Lys His Ser Phe Val Phe Ser
            260                 265                 270
Gly Val Thr Pro Ala Gln Leu Arg Gln Ser Ser Leu Glu Leu Thr Val
        275                 280                 285
Trp Asp Gln Ala Leu Phe Gly Met Asn Asp Arg Leu Leu Gly Gly Thr
    290                 295                 300
Arg Leu Gly Ser Lys Gly Asp Thr Ala Val Gly Gly Asp Ala Cys Ser
305                 310                 315                 320
Leu Ser Lys Leu Gln Trp Gln Lys Val Leu Ser Ser Pro Asn Leu Trp
                325                 330                 335
Thr Asp Met Thr Leu Val Leu His
            340
<210>    3
<211>    471
<212>    DNA
<213>    Homo sapiens
<220>
<221>    CDS
<222>    (1)..(468)
<400>    3
gga agt cca gca ggt aga tca atc tac aac agc ttt tat gtg tat tgc           48
Gly Ser Pro Ala Gly Arg Ser Ile Tyr Asn Ser Phe Tyr Val Tyr Cys
  1               5                  10                  15
aaa ggc ccc tgt caa aga gtg cag ccg gga aaa ctc agg gta cag tgc           96
Lys Gly Pro Cys Gln Arg Val Gln Pro Gly Lys Leu Arg Val Gln Cys
             20                  25                  30
agc acc tgc agg cag gca acg ctc acc ttg acc cag ggt cca tct tgc          144
Ser Thr Cys Arg Gln Ala Thr Leu Thr Leu Thr Gln Gly Pro Ser Cys
         35                  40                  45
tgg gat gat gtt tta att cca aac cgg atg agt ggt gaa tgc caa tcc          192
Trp Asp Asp Val Leu Ile Pro Asn Arg Met Ser Gly Glu Cys Gln Ser
     50                  55                  60
cca cac tgc cct ggg act agt gca gaa ttt ttc ttt aaa tgt gga gca          240
Pro His Cys Pro Gly Thr Ser Ala Glu Phe Phe Phe Lys Cys Gly Ala
 65                  70                  75                  80
cac ccc acc tct gac aag gaa aca tca gta gct ttg cac ctg atc gca          288
His Pro Thr Ser Asp Lys Glu Thr Ser Val Ala Leu His Leu Ile Ala
                 85                  90                  95
aca aat agt cgg aac atc act tgc att acg tgc aca gac gtc agg agc          336
Thr Asn Ser Arg Asn Ile Thr Cys Ile Thr Cys Thr Asp Val Arg Ser
            100                 105                 110
ccc gtc ctg gtt ttc cag tgc aac tcc cgc cac gtg att tgc tta gac          384
Pro Val Leu Val Phe Gln Cys Asn Ser Arg His Val Ile Cys Leu Asp
        115                 120                 125
tgt ttc cac tta tac tgt gtg aca aga ctc aat gat cgg cag ttt gtt          432
Cys Phe His Leu Tyr Cys Val Thr Arg Leu Asn Asp Arg Gln Phe Val
    130                 135                 140
cac gac cct caa ctt ggc tac tcc ctg cct tgt gtg     ta g                 471
```

```
His Asp Pro Gln Leu Gly Tyr Ser Leu Pro Cys Val
145                 150                 155
<210>  4
<211>  156
<212>  PRT
<213>  Homo sapiens
<400>  4
Gly Ser Pro Ala Gly Arg Ser Ile Tyr Asn Ser Phe Tyr Val Tyr Cys
  1               5                  10                  15
Lys Gly Pro Cys Gln Arg Val Gln Pro Gly Lys Leu Arg Val Gln Cys
             20                  25                  30
Ser Thr Cys Arg Gln Ala Thr Leu Thr Leu Thr Gln Gly Pro Ser Cys
         35                  40                  45
Trp Asp Asp Val Leu Ile Pro Asn Arg Met Ser Gly Glu Cys Gln Ser
     50                  55                  60
Pro His Cys Pro Gly Thr Ser Ala Glu Phe Phe Phe Lys Cys Gly Ala
 65                  70                  75                  80
His Pro Thr Ser Asp Lys Glu Thr Ser Val Ala Leu His Leu Ile Ala
                 85                  90                  95
Thr Asn Ser Arg Asn Ile Thr Cys Ile Thr Cys Thr Asp Val Arg Ser
            100                 105                 110
Pro Val Leu Val Phe Gln Cys Asn Ser Arg His Val Ile Cys Leu Asp
        115                 120                 125
Cys Phe His Leu Tyr Cys Val Thr Arg Leu Asn Asp Arg Gln Phe Val
    130                 135                 140
His Asp Pro Gln Leu Gly Tyr Ser Leu Pro Cys Val
145                 150                 155
<210>    5
<211>    27
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    5
ttaagaattc ggaagtccag caggtag                                           27
<210>    6
<211>    29
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    6
attaggatcc ctacacacaa ggcagggag                                         29
<210>    7
<211>    19
<212>    DNA
<213>    Artificial Sequence
<220>
```

```
<223>    Artificial sequence
         description: oligonucleotide
<400>    7
gcgtttggaa tcactacag                                                    19
<210>    8
<211>    17
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    8
ggtctcggtg tggcatc                                                      17
<210>    9
<211>    18
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    9
ccgcttgctt ggaggaac                                                     18
<210>    10
<211>    18
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    10
cgtatttctc cgccttgg                                                     18
<210>    11
<211>    28
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    11
aatagctcga gtcagtgcag gacaagag                                          28
<210>    12
<211>    2347
<212>    DNA
<213>    Homo sapiens
<400>    12
ggccttgggg cactgaggga tgccagttct gcctgttcat ctggaacctg gatctaagga       60
gggaagaggc gttgcccctg ctggcatagt caggtaccag cccagccagg tattgaacgg      120
gctgagcttt tcatgatggt tcctgctgac ctggaaacat cttaaatgga agggcgtgag      180
cgcttggtcc atgcagtgaa gctcttccaa cctgggtcaa cgaaaacgga gaagaaatgg      240
cccaagaaat agatctgagt gctctcaagg agttagaacg cgaggccatt ctccaggtcc      300
```

```
tgtaccgaga ccaggcggtt caaaacacag aggaggagag gacacggaaa ctgaaaacac        360
acctgcagca tctccggtgg aaaggagcga agaacacgga ctgggagcac aaagagaagt        420
gctgtgcgcg ctgccagcag gtgctggggt tcctgctgca ccggggcgcc gtgtgccggg        480
gctgcagcca ccgcgtgtgt gcccagtgcc gagtgttcct gagggggacc catgcctgga        540
agtgcacggt gtgcttcgag gacaggaatg tcaaaataaa aactggagaa tggttctatg        600
aggaacgagc caagaaattt ccaactggag gcaaacatga gacagttgga gggcagctct        660
tgcaatctta tcagaagctg agcaaaattt ctgtggttcc tcctactcca cctcctgtca        720
gcgagagcca gtgcagccgc agtcctggca ggttacagga atttggtcag tttagaggat        780
ttaataagtc cgtggaaaat ttgtttctgt ctcttgctac ccacgtgaaa aagctctcca        840
aatcccagaa tgatatgact tctgagaagc atcttctcgc cacgggcccc aggcagtgtg        900
tgggacagac agagagacgg agccagtctg acactgcggt caacgtcacc accaggaagg        960
tcagtgcacc agatattctg aaacctctca atcaagagga tcccaaatgc tctactaacc       1020
ctattttgaa gcaacagaat ctcccatcca gtccggcacc cagtaccata ttctctggag       1080
gttttagaca cggaagttta attagcattg acagcacctg tacagagatg ggcaattttg       1140
acaatgctaa tgtcactgga gaaatagaat ttgccattca ttattgcttc aaaacccatt       1200
ctttagaaat atgcatcaag gcctgtaaga accttgccta tggagaagaa aagaagaaaa       1260
agtgcaatcc gtatgtgaag acctacctgt tgcccgacag atcctcccag ggaaagcgca       1320
agactggagt ccaaaggaac accgtggacc cgacctttca ggagaccttg aagtatcagg       1380
tggcccctgc ccagctggtg acccggcagc tgcaggtctc ggtgtggcat ctgggcacgc       1440
tggcccggag agtgtttctt ggagaagtga tcattcctct ggccacgtgg gactttgaag       1500
acagcacaac acagtccttc cgctggcatc cgctccgggc caaggcggag aaatacgaag       1560
acagcgttcc tcagagtaat ggagagctca cagtccgggc taagctggtt ctcccttcac       1620
ggcccagaaa actccaagag gctcaagaag ggacagatca gccatcactt catggtcaac       1680
tttgtttggt agtgctagga gccaagaatt tacctgtgcg gccagatggc accttgaact       1740
catttgttaa gggctgtctc actctgccag accaacaaaa actgagactg aagtcgccag       1800
tcctgaggaa gcaggcttgc cccccagtgga aacactcatt tgtcttcagt ggcgtaaccc       1860
cagctcagct gaggcagtcg agcttggagt taactgtctg ggatcaggcc ctctttggaa       1920
tgaacgaccg cttgcttgga ggaaccagac ttggttcaaa gggagacaca gctgttggcg       1980
gggatgcatg ctcacaatcg aagctccagt ggcagaaagt cctttccagc cccaatctat       2040
ggacagacat gactcttgtc ctgcactgac atgaaggcct caaggttcca ggttgcagca       2100
ggcgtgaggc actgtgcgtc tgcagagggg ctacgaacca ggtgcagggt cccagctgga       2160
gacccctttg accttgagca gtctccatct gcggccctgt cccatggctt aaccgcctat       2220
tggtatctgt gtatatttac gttaaacaca attatgttac ctaagcctct ggtgggttat       2280
ctcctctttg agatgtagaa aatggccaga ttttaataaa cgttgttacc catgaaaaaa       2340
aaaaaaa                                                                 2347
<210>    13
<211>    610
<212>    PRT
<213>    Homo sapiens
<400>    13
Met Ala Gln Glu Ile Asp Leu Ser Ala Leu Lys Glu Leu Glu Arg Glu
  1               5                  10                  15
Ala Ile Leu Gln Val Leu Tyr Arg Asp Gln Ala Val Gln Asn Thr Glu
             20                  25                  30
Glu Glu Arg Thr Arg Lys Leu Lys Thr His Leu Gln His Leu Arg Trp
         35                  40                  45
Lys Gly Ala Lys Asn Thr Asp Trp Glu His Lys Glu Lys Cys Cys Ala
     50                  55                  60
Arg Cys Gln Gln Val Leu Gly Phe Leu Leu His Arg Gly Ala Val Cys
 65                  70                  75                  80
```

```
Arg Gly Cys Ser His Arg Val Cys Ala Gln Cys Arg Val Phe Leu Arg
                 85                  90                  95
Gly Thr His Ala Trp Lys Cys Thr Val Cys Phe Glu Asp Arg Asn Val
            100                 105                 110
Lys Ile Lys Thr Gly Glu Trp Phe Tyr Glu Glu Arg Ala Lys Lys Phe
        115                 120                 125
Pro Thr Gly Gly Lys His Glu Thr Val Gly Gly Gln Leu Leu Gln Ser
    130                 135                 140
Tyr Gln Lys Leu Ser Lys Ile Ser Val Val Pro Pro Thr Pro Pro Pro
145                 150                 155                 160
Val Ser Glu Ser Gln Cys Ser Arg Ser Pro Gly Arg Leu Gln Glu Phe
                165                 170                 175
Gly Gln Phe Arg Gly Phe Asn Lys Ser Val Glu Asn Leu Phe Leu Ser
            180                 185                 190
Leu Ala Thr His Val Lys Lys Leu Ser Lys Ser Gln Asn Asp Met Thr
        195                 200                 205
Ser Glu Lys His Leu Leu Ala Thr Gly Pro Arg Gln Cys Val Gly Gln
    210                 215                 220
Thr Glu Arg Arg Ser Gln Ser Asp Thr Ala Val Asn Val Thr Thr Arg
225                 230                 235                 240
Lys Val Ser Ala Pro Asp Ile Leu Lys Pro Leu Asn Gln Glu Asp Pro
                245                 250                 255
Lys Cys Ser Thr Asn Pro Ile Leu Lys Gln Gln Asn Leu Pro Ser Ser
            260                 265                 270
Pro Ala Pro Ser Thr Ile Phe Ser Gly Gly Phe Arg His Gly Ser Leu
        275                 280                 285
Ile Ser Ile Asp Ser Thr Cys Thr Glu Met Gly Asn Phe Asp Asn Ala
    290                 295                 300
Asn Val Thr Gly Glu Ile Glu Phe Ala Ile His Tyr Cys Phe Lys Thr
305                 310                 315                 320
His Ser Leu Glu Ile Cys Ile Lys Ala Cys Lys Asn Leu Ala Tyr Gly
                325                 330                 335
Glu Glu Lys Lys Lys Lys Cys Asn Pro Tyr Val Lys Thr Tyr Leu Leu
            340                 345                 350
Pro Asp Arg Ser Ser Gln Gly Lys Arg Lys Thr Gly Val Gln Arg Asn
        355                 360                 365
Thr Val Asp Pro Thr Phe Gln Glu Thr Leu Lys Tyr Gln Val Ala Pro
    370                 375                 380
Ala Gln Leu Val Thr Arg Gln Leu Gln Val Ser Val Trp His Leu Gly
385                 390                 395                 400
Thr Leu Ala Arg Arg Val Phe Leu Gly Glu Val Ile Ile Pro Leu Ala
                405                 410                 415
Thr Trp Asp Phe Glu Asp Ser Thr Thr Gln Ser Phe Arg Trp His Pro
            420                 425                 430
Leu Arg Ala Lys Ala Glu Lys Tyr Glu Asp Ser Val Pro Gln Ser Asn
        435                 40                 445
Gly Glu Leu Thr Val Arg Ala Lys Leu Val Leu Pro Ser Arg Pro Arg
    450                 455                 460
Lys Leu Gln Glu Ala Gln Glu Gly Thr Asp Gln Pro Ser Leu His Gly
465                 470                 475                 480
```

```
Gln Leu Cys Leu Val Val Leu Gly Ala Lys Asn Leu Pro Val Arg Pro
                485                 490                 495
Asp Gly Thr Leu Asn Ser Phe Val Lys Gly Cys Leu Thr Leu Pro Asp
            500                 505                 510
Gln Gln Lys Leu Arg Leu Lys Ser Pro Val Leu Arg Lys Gln Ala Cys
        515                 520                 525
Pro Gln Trp Lys His Ser Phe Val Phe Ser Gly Val Thr Pro Ala Gln
    530                 535                 540
Leu Arg Gln Ser Ser Leu Glu Leu Thr Val Trp Asp Gln Ala Leu Phe
545                 550                 555                 560
Gly Met Asn Asp Arg Leu Leu Gly Gly Thr Arg Leu Gly Ser Lys Gly
                565                 570                 575
Asp Thr Ala Val Gly Gly Asp Ala Cys Ser Gln Ser Lys Leu Gln Trp
            580                 585                 590
Gln Lys Val Leu Ser Ser Pro Asn Leu Trp Thr Asp Met Thr Leu Val
        595                 600                 605
Leu His
    610
<210>    14
<211>    1648
<212>    DNA
<213>    Homo sapiens
<400>    14
gaaatcatgc ccctcgtaga gcagcaggtc caagcagggc tgctggctat ttttccaaaa         60
agtgaggcag tttaaaaaaa aggcggagaa ctagaattat agaataatgg cacattttgt        120
gtatttgtaa aactaacggc ttgcatggtt cacaacccat ttcttatgcc tgtgttttcc        180
ttggcagcaa aatttctgtg gttcctccta ctccacctcc tgtcagcgag agccagtgca        240
gccgcagtcc tggcaggaag gtcagtgcac cagatattct gaaacctctc aatcaagagg        300
atcccaaatg ctctactaac cctattttga agcaacagaa tctcccatcc agtccggcac        360
ccagtaccat attctctgga ggttttagac acggaagttt aattagcatt gacagcacct        420
gtacagagat gggcaatttt gacaatgcta atgtcactgg agaaatagaa tttgccattc        480
attattgctt caaaacccat tctttagaaa tatgcatcaa ggcctgtaag aaccttgcct        540
atggagaaga aaagaagaaa aagtgcaatc cgtatgtgaa gacctacctg ttgcccgaca        600
gatcctccca gggaaagcgc aagactggag tccaaaggaa caccgtggac ccgacctttc        660
aggagacctt gaagtatcag gtggcccctg cccagctggt gacccggcag ctgcaggtct        720
cggtgtggca tctgggcacg ctggcccgga gagtgtttct tggagaagtg atcattcctc        780
tggccacgtg ggactttgaa gacagcacaa cacagtcctt ccgctggcat ccgctccggg        840
ccaaggcgga gaaatacgaa gacagcgttc ctcagagtaa tggagagctc acagtccggg        900
ctaagctggt tctcccttca cggcccagaa aactccaaga ggctcaagaa gggacagatc        960
agccatcact tcatggtcaa ctttgtttgg tagtgctagg agccaagaat ttacctgtgc       1020
ggccagatgg caccttgaac tcatttgtta agggctgtct cactctgcca gaccaacaaa       1080
aactgagact gaagtcgcca gtcctgagga agcaggcttg cccccagtgg aaacactcat       1140
ttgtcttcag tggcgtaacc ccagctcagc tgaggcagtc gagcttggag ttaactgtct       1200
gggatcaggc cctctttgga atgaacgacc gcttgcttgg aggaaccaga cttggttcaa       1260
agggagacac agctgttggc ggggatgcat gctcacaatc gaagctccag tggcagaaag       1320
tcctttccag ccccaatcta tggacagaca tgactcttgt cctgcactga catgaaggcc       1380
tcaaggttcc aggttgcagc aggcgtgagg cactgtgcgt ctgcagaggg gctacgaacc       1440
aggtgcaggg tcccagctgg agacccettt gaccttgagc agtctccatc tgcggccctg       1500
tcccatggct taaccgccta ttggtatctg tgtatattta cgttaaacac aattatgtta       1560
cctaagcctc tggtgggtta tctcctcttt gagatgtaga aaatggccag attttaataa       1620
```

```
acgttgttac ccatgaaaaa aaaaaaaa                                          1648
<210>    15
<211>    313
<212>    PRT
<213>    Homo sapiens
<400>    15
Met Gly Asn Phe Asp Asn Ala Asn Val Thr Gly Glu Ile Glu Phe Ala
  1               5                  10                  15
Ile His Tyr Cys Phe Lys Thr His Ser Leu Glu Ile Cys Ile Lys Ala
             20                  25                  30
Cys Lys Asn Leu Ala Tyr Gly Glu Glu Lys Lys Lys Lys Cys Asn Pro
         35                  40                  45
Tyr Val Lys Thr Tyr Leu Leu Pro Asp Arg Ser Ser Gln Gly Lys Arg
     50                  55                  60
Lys Thr Gly Val Gln Arg Asn Thr Val Asp Pro Thr Phe Gln Glu Thr
 65                  70                  75                  80
Leu Lys Tyr Gln Val Ala Pro Ala Gln Leu Val Thr Arg Gln Leu Gln
                 85                  90                  95
Val Ser Val Trp His Leu Gly Thr Leu Ala Arg Arg Val Phe Leu Gly
            100                 105                 110
Glu Val Ile Ile Pro Leu Ala Thr Trp Asp Phe Glu Asp Ser Thr Thr
        115                 120                 125
Gln Ser Phe Arg Trp His Pro Leu Arg Ala Lys Ala Glu Lys Tyr Glu
    130                 135                 140
Asp Ser Val Pro Gln Ser Asn Gly Glu Leu Thr Val Arg Ala Lys Leu
145                 150                 155                 160
Val Leu Pro Ser Arg Pro Arg Lys Leu Gln Glu Ala Gln Glu Gly Thr
                165                 170                 175
Asp Gln Pro Ser Leu His Gly Gln Leu Cys Leu Val Val Leu Gly Ala
            180                 185                 190
Lys Asn Leu Pro Val Arg Pro Asp Gly Thr Leu Asn Ser Phe Val Lys
        195                 200                 205
Gly Cys Leu Thr Leu Pro Asp Gln Gln Lys Leu Arg Leu Lys Ser Pro
    210                 215                 220
Val Leu Arg Lys Gln Ala Cys Pro Gln Trp Lys His Ser Phe Val Phe
225                 230                 235                 240
Ser Gly Val Thr Pro Ala Gln Leu Arg Gln Ser Ser Leu Glu Leu Thr
                245                 250                 255
Val Trp Asp Gln Ala Leu Phe Gly Met Asn Asp Arg Leu Leu Gly Gly
            260                 265                 270
Thr Arg Leu Gly Ser Lys Gly Asp Thr Ala Val Gly Gly Asp Ala Cys
        275                 280                 285
Ser Gln Ser Lys Leu Gln Trp Gln Lys Val Leu Ser Ser Pro Asn Leu
    290                 295                 300
Trp Thr Asp Met Thr Leu Val Leu His
305                 310
<210>    16
<211>    19
<212>    DNA
<213>    Artificial Sequence
```

<220>
<223> Artificial sequence
description: oligonucleotide
<400> 16
ccagttctgc ctgttcatc 19
<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 17
ttcaaaacac agaggaggag 20
<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 18
gaatttggtc agtttagagg 20
<210> 19
<211> 26
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 19
ttctgggatt tggagagctt tttcac 26
<210> 20
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 20
tctgtctgtc ccacacactg cc 22
<210> 21
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 21
gactggctcc gtctctctg 19

```
<210>    22
<211>    21
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    22
aagcaacaga atctcccatc c                                           21
<210>    23
<211>    21
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    23
gcattgtcaa aattgcccat c                                           21
<210>    24
<211>    20
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    24
aggcggagaa atacgaagac                                             20
<210>    25
<211>    22
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    25
gcagagtgag acagccctta ac                                          22
<210>    26
<211>    24
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    26
cttcctcagg actggcgact tcag                                        24
<210>    27
<211>    24
<212>    DNA
<213>    Artificial Sequence
<220>
```

<223> Artificial sequence
description: oligonucleotide
<400> 27
caagcggtcg ttcattccaa agag 24
<210> 28
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 28
aagaggagat aacccaccag ag 22
<210> 29
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 29
agggctgctg gctatttttc 20
<210> 30
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 30
taagaaatgg gttgtgaac 19
<210> 31
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 31
aagcaacaga atctcccatc c 21
<210> 32
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 32
gcattgtcaa aattgcccat c 21
<210> 33

<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 33
aggcggagaa atacgaagac 20
<210> 34
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 34
gcagagtgag acagccctta ac 22
<210> 35
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 35
cttcctcagg actggcgact tcag 24
<210> 36
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 36
caagcggtcg ttcattccaa agag 24
<210> 37
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence
description: oligonucleotide
<400> 37
aagaggagat aacccaccag ag 22
<210> 38
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Artificial sequence

```
         description: oligonucleotide
<400>    38
aatggaaggg cgtgacgc                                                   18
<210>    39
<211>    21
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    39
cctcacgcct gctgcaacct g                                               21
<210>    40
<211>    31
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    40
gcacgaattc atggcccaag aaatagatct g                                    31
<210>    41
<211>    24
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    41
ctgtcttcgt atttctccgc cttg                                            24
<210>    42
<211>    2347
<212>    DNA
<213>    Homo sapiens
<400>    42
ggccttgggg cactgaggga tgccagttct gcctgttcat ctggaacctg gatctaagga     60
gggaagaggc gttgcccctg ctggcatagt caggtaccag cccagccagg tattgaacgg    120
gctgagcttt tcatgatggt tcctgctgac ctggaaacat cttaaatgga agggcgtgag    180
cgcttggtcc atgcagtgaa gctcttccaa cctgggtcaa cgaaaacgga gaagaaatgg    240
cccaagaaat agatctgagt gctctcaagg agttagaacg cgaggccatt ctccaggtcc    300
tgtaccgaga ccaggcggtt caaaacacag aggaggagag gacacggaaa ctgaaaacac    360
acctgcagca tctccggtgg aaaggagcga agaacacgga ctgggagcac aaagagaagt    420
gctgtgcgcg ctgccagcag gtgctggggt tcctgctgca ccggggcgcc gtgtgccggg    480
gctgcagcca ccgcgtgtgt gcccagtgcc gagtgttcct gaggggggacc catgcctgga    540
agtgcacggt gtgcttcgag gacaggaatg tcaaaataaa aactggagaa tggttctatg    600
aggaacgagc caagaaattt ccaactggag gcaaacatga gacagttgga gggcagctct    660
tgcaatctta tcagaagctg agcaaaattt ctgtggttcc tcctactcca cctcctgtca    720
gcgagagcca gtgcagccgc agtcctggca ggttacagga atttggtcag tttagaggat    780
ttaataagtc cgtggaaaat ttgtttctgt ctcttgctac ccacgtgaaa aagctctcca    840
aatcccagaa tgatatgact tctgagaagc atcttctcgc cacgggcccc aggcagtgtg    900
```

```
tgggacagac agagagacgg agccagtctg acactgcggt caacgtcacc accaggaagg         960
tcagtgcacc agatattctg aaacctctca atcaagagga tcccaaatgc tctactaacc        1020
ctattttgaa gcaacagaat ctcccatcca gtccggcacc cagtaccata ttctctggag        1080
gttttagaca cggaagttta attagcattg acagcacctg tacagagatg ggcaattttg        1140
acaatgctaa tgtcactgga gaaatagaat ttgccattca ttattgcttc aaaacccatt        1200
ctttagaaat atgcatcaag gcctgtaaga accttgccta tggagaagaa aagaagaaaa        1260
agtgcaatcc gtatgtgaag acctacctgt tgcccgacag atcctcccag ggaaagcgca        1320
agactggagt ccaaaggaac accgtggacc cgacctttca ggagaccttg aagtatcagg        1380
tggcccctgc ccagctggtg acccggcagc tgcaggtctc ggtgtggcat ctgggcacgc        1440
tggcccggag agtgtttctt ggagaagtga tcattcctct ggccacgtgg gactttgaag        1500
acagcacaac acagtccttc cgctggcatc cgctccgggc caaggcggag aaatacgaag        1560
acagcgttcc tcagagtaat ggagagctca cagtccgggc taagctggtt ctcccttcac        1620
ggcccagaaa actccaagag gctcaagaag ggacagatca gccatcactt catggtcaac        1680
tttgtttggt agtgctagga gccaagaatt tacctgtgcg gccagatggc accttgaact        1740
catttgttaa gggctgtctc actctgccag accaacaaaa actgagactg aagtcgccag        1800
tcctgaggaa gcaggcttgc cccagtgga aacactcatt tgtcttcagt ggcgtaaccc        1860
cagctcagct gaggcagtcg agcttggagt taactgtctg ggatcaggcc ctctttggaa        1920
tgaacgaccg cttgcttgga ggaaccagac ttggttcaaa gggagacaca gctgttggcg        1980
gggatgcatg ctcacaatcg aagctccagt ggcagaaagt cctttccagc cccaatctat        2040
ggacagacat gactcttgtc ctgcactgac atgaaggcct caaggttcca ggttgcagca        2100
ggcgtgaggc actgtgcgtc tgcagagggg ctacgaacca ggtgcagggt cccagctgga        2160
gacccctttg accttgagca gtctccatct gcggccctgt cccatggctt aaccgcctat        2220
tggtatctgt gtatatttac gttaaacaca attatgttac ctaagcctct ggtgggttat        2280
ctcctctttg agatgtagaa aatggccaga ttttaataaa cgttgttacc catgaaaaaa        2340
aaaaaaa                                                                  2347
<210>    43
<211>    610
<212>    PRT
<213>    Homo sapiens
<400>    43
Met Ala Gln Glu Ile Asp Leu Ser Ala Leu Lys Glu Leu Glu Arg Glu
  1               5                  10                  15
Ala Ile Leu Gln Val Leu Tyr Arg Asp Gln Ala Val Gln Asn Thr Glu
             20                  25                  30
Glu Glu Arg Thr Arg Lys Leu Lys Thr His Leu Gln His Leu Arg Trp
         35                  40                  45
Lys Gly Ala Lys Asn Thr Asp Trp Glu His Lys Glu Lys Cys Cys Ala
     50                  55                  60
Arg Cys Gln Gln Val Leu Gly Phe Leu Leu His Arg Gly Ala Val Cys
 65                  70                  75                  80
Arg Gly Cys Ser His Arg Val Cys Ala Gln Cys Arg Val Phe Leu Arg
                 85                  90                  95
Gly Thr His Ala Trp Lys Cys Thr Val Cys Phe Glu Asp Arg Asn Val
            100                 105                 110
Lys Ile Lys Thr Gly Glu Trp Phe Tyr Glu Glu Arg Ala Lys Lys Phe
        115                 120                 125
Pro Thr Gly Gly Lys His Glu Thr Val Gly Gly Gln Leu Leu Gln Ser
    130                 135                 140
Tyr Gln Lys Leu Ser Lys Ile Ser Val Val Pro Pro Thr Pro Pro Pro
145                 150                 155                 160
```

```
Val Ser Glu Ser Gln Cys Ser Arg Ser Pro Gly Arg Leu Gln Glu Phe
                165                 170                 175
Gly Gln Phe Arg Gly Phe Asn Lys Ser Val Glu Asn Leu Phe Leu Ser
            180                 185                 190
Leu Ala Thr His Val Lys Lys Leu Ser Lys Ser Gln Asn Asp Met Thr
        195                 200                 205
Ser Glu Lys His Leu Leu Ala Thr Gly Pro Arg Gln Cys Val Gly Gln
    210                 215                 220
Thr Glu Arg Arg Ser Gln Ser Asp Thr Ala Val Asn Val Thr Thr Arg
225                 230                 235                 240
Lys Val Ser Ala Pro Asp Ile Leu Lys Pro Leu Asn Gln Glu Asp Pro
                245                 250                 255
Lys Cys Ser Thr Asn Pro Ile Leu Lys Gln Gln Asn Leu Pro Ser Ser
            260                 265                 270
Pro Ala Pro Ser Thr Ile Phe Ser Gly Gly Phe Arg His Gly Ser Leu
        275                 280                 285
Ile Ser Ile Asp Ser Thr Cys Thr Glu Met Gly Asn Phe Asp Asn Ala
    290                 295                 300
Asn Val Thr Gly Glu Ile Glu Phe Ala Ile His Tyr Cys Phe Lys Thr
305                 310                 315                 320
His Ser Leu Glu Ile Cys Ile Lys Ala Cys Lys Asn Leu Ala Tyr Gly
                325                 330                 335
Glu Glu Lys Lys Lys Lys Cys Asn Pro Tyr Val Lys Thr Tyr Leu Leu
            340                 345                 350
Pro Asp Arg Ser Ser Gln Gly Lys Arg Lys Thr Gly Val Gln Arg Asn
        355                 360                 365
Thr Val Asp Pro Thr Phe Gln Glu Thr Leu Lys Tyr Gln Val Ala Pro
    370                 375                 380
Ala Gln Leu Val Thr Arg Gln Leu Gln Val Ser Val Trp His Leu Gly
385                 390                 395                 400
Thr Leu Ala Arg Arg Val Phe Leu Gly Glu Val Ile Ile Pro Leu Ala
                405                 410                 415
Thr Trp Asp Phe Glu Asp Ser Thr Thr Gln Ser Phe Arg Trp His Pro
            420                 425                 430
Leu Arg Ala Lys Ala Glu Lys Tyr Glu Asp Ser Val Pro Gln Ser Asn
        435                 440                 445
Gly Glu Leu Thr Val Arg Ala Lys Leu Val Leu Pro Ser Arg Pro Arg
    450                 455                 460
Lys Leu Gln Glu Ala Gln Glu Gly Thr Asp Gln Pro Ser Leu His Gly
465                 470                 475                 480
Gln Leu Cys Leu Val Val Leu Gly Ala Lys Asn Leu Pro Val Arg Pro
                485                 490                 495
Asp Gly Thr Leu Asn Ser Phe Val Ly Gly Cys Leu Thr Leu Pro Asp
            500                 505                 510
Gln Gln Lys Leu Arg Leu Lys Ser Pro Val Leu Arg Lys Gln Ala Cys
        515                 520                 525
Pro Gln Trp Lys His Ser Phe Val Phe Ser Gly Val Thr Pro Ala Gln
    530                 535                 540
Leu Arg Gln Ser Ser Leu Glu Leu Thr Val Trp Asp Gln Ala Leu Phe
545                 550                 555                 560
```

```
Gly Met Asn Asp Arg Leu Leu Gly Gly Thr Arg Leu Gly Ser Lys Gly
                565                 570                 575
Asp Thr Ala Val Gly Gly Asp Ala Cys Ser Gln Ser Lys Leu Gln Trp
            580                 585                 590
Gln Lys Val Leu Ser Ser Pro Asn Leu Trp Thr Asp Met Thr Leu Val
        595                 600                 605
Leu His
    610
<210>    44
<211>    1648
<212>    DNA
<213>    Homo sapiens
<400>    44
gaaatcatgc ccctcgtaga gcagcaggtc caagcagggc tgctggctat ttttccaaaa         60
agtgaggcag tttaaaaaaa aggcggagaa ctagaattat agaataatgg cacattttgt        120
gtatttgtaa aactaacggc ttgcatggtt cacaacccat ttcttatgcc tgtgttttcc        180
ttggcagcaa aatttctgtg gttcctccta ctccacctcc tgtcagcgag agccagtgca        240
gccgcagtcc tggcaggaag gtcagtgcac cagatattct gaaacctctc aatcaagagg        300
atcccaaatg ctctactaac cctattttga agcaacagaa tctcccatcc agtccggcac        360
ccagtaccat attctctgga ggttttagac acggaagttt aattagcatt gacagcacct        420
gtacagagat gggcaatttt gacaatgcta atgtcactgg agaaatagaa tttgccattc        480
attattgctt caaaacccat tctttagaaa tatgcatcaa ggcctgtaag aaccttgcct        540
atggagaaga aaagaagaaa aagtgcaatc cgtatgtgaa gacctacctg ttgcccgaca        600
gatcctccca gggaaagcgc aagactggag tccaaaggaa caccgtggac ccgacctttc        660
aggagacctt gaagtatcag gtggccccctg cccagctggt gacccggcag ctgcaggtct        720
cggtgtggca tctgggcacg ctggcccgga gagtgtttct tggagaagtg atcattcctc        780
tggccacgtg ggactttgaa gacagcacaa cacagtcctt ccgctggcat ccgctccggg        840
ccaaggcgga gaaatacgaa gacagcgttc ctcagagtaa tggagagctc acagtccggg        900
ctaagctggt tctcccttca cggcccagaa aactccaaga ggctcaagaa gggacagatc        960
agccatcact tcatggtcaa ctttgtttgg tagtgctagg agccaagaat ttacctgtgc       1020
ggccagatgg caccttgaac tcatttgtta agggctgtct cactctgcca gaccaacaaa       1080
aactgagact gaagtcgcca gtcctgagga agcaggcttg cccccagtgg aaacactcat       1140
ttgtcttcag tggcgtaacc ccagctcagc tgaggcagtc gagcttggag ttaactgtct       1200
gggatcaggc cctctttgga atgaacgacc gcttgcttgg aggaaccaga cttggttcaa       1260
agggagacac agctgttggc ggggatgcat gctcacaatc gaagctccag tggcagaaag       1320
tcctttccag ccccaatcta tggacagaca tgactcttgt cctgcactga catgaaggcc       1380
tcaaggttcc aggttgcagc aggcgtgagg cactgtgcgt ctgcagaggg gctacgaacc       1440
aggtgcaggg tcccagctgg agaccccttt gaccttgagc agtctccatc tgcggccctg       1500
tcccatggct taaccgccta ttggtatctg tgtatattta cgttaaacac aattatgtta       1560
cctaagcctc tggtgggtta tctcctcttt gagatgtaga aaatggccag attttaataa       1620
acgttgttac ccatgaaaaa aaaaaaaa                                           1648
<210>    45
<211>    313
<212>    PRT
<213>    Homo sapiens
<400>    45
Met Gly Asn Phe Asp Asn Ala Asn Val Thr Gly Glu Ile Glu Phe Ala
  1               5                  10                  15
Ile His Tyr Cys Phe Lys Thr His Ser Leu Glu Ile Cys Ile Lys Ala
             20                  25                  30
```

```
Cys Lys Asn Leu Ala Tyr Gly Glu Glu Lys Lys Lys Lys Cys Asn Pro
         35                  40                  45
Tyr Val Lys Thr Tyr Leu Leu Pro Asp Arg Ser Ser Gln Gly Lys Arg
     50                  55                  60
Lys Thr Gly Val Gln Arg Asn Thr Val Asp Pro Thr Phe Gln Glu Thr
 65                  70                  75                  80
Leu Lys Tyr Gln Val Ala Pro Ala Gln Leu Val Thr Arg Gln Leu Gln
                 85                  90                  95
Val Ser Val Trp His Leu Gly Thr Leu Ala Arg Arg Val Phe Leu Gly
            100                 105                 110
Glu Val Ile Ile Pro Leu Ala Thr Trp Asp Phe Glu Asp Ser Thr Thr
        115                 120                 125
Gln Ser Phe Arg Trp His Pro Leu Arg Ala Lys Ala Glu Lys Tyr Glu
    130                 135                 140
Asp Ser Val Pro Gln Ser Asn Gly Glu Leu Thr Val Arg Ala Lys Leu
145                 150                 155                 160
Val Leu Pro Ser Arg Pro Arg Lys Leu Gln Glu Ala Gln Glu Gly Thr
                165                 170                 175
Asp Gln Pro Ser Leu His Gly Gln Leu Cys Leu Val Val Leu Gly Ala
            180                 185                 190
Lys Asn Leu Pro Val Arg Pro Asp Gly Thr Leu Asn Ser Phe Val Lys
        195                 200                 205
Gly Cys Leu Thr Leu Pro Asp Gln Gln Lys Leu Arg Leu Lys Ser Pro
    210                 215                 220
Val Leu Arg Lys Gln Ala Cys Pro Gln Trp Lys His Ser Phe Val Phe
225                 230                 235                 240
Ser Gly Val Thr Pro Ala Gln Leu Arg Gln Ser Ser Leu Glu Leu Thr
                245                 250                 255
Val Trp Asp Gln Ala Leu Phe Gly Met Asn Asp Arg Leu Leu Gly Gly
            260                 265                 270
Thr Arg Leu Gly Ser Lys Gly Asp Thr Ala Val Gly Gly Asp Ala Cys
        275                 280                 285
Ser Gln Ser Lys Leu Gln Trp Gln Lys Val Leu Ser Ser Pro Asn Leu
    290                 295                 300
Trp Thr Asp Met Thr Leu Val Leu His
305                 310
<210>    46
<211>    21
<212>    DNA
<213>    Artificial Sequence
<220>
<223>    Artificial sequence
         description: oligonucleotide
<400>    46
tcgtagagca gcaggtccaa g                                                21
```