(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。Int. Cl.^{7}
A61K 38/18
C12N 15/12
C07K 14/475

(11) 공개번호 10-2004-0094709
(43) 공개일자 2004년11월10일

(21) 출원번호	10-2004-7012977		
(22) 출원일자	2004년08월20일		
번역문 제출일자	2004년08월20일		
(86) 국제출원번호	PCT/US2003/005151	(87) 국제공개번호	WO 2003/072715
(86) 국제출원출원일자	2003년02월21일	(87) 국제공개일자	2003년09월04일

(30) 우선권주장	60/357,845	2002년02월21일	미국(US)
	60/434,644	2002년12월20일	미국(US)

(71) 출원인
와이어쓰
미합중국 뉴저지주 07940-0874 매디슨 파이브 지랄다 팜즈

(72) 발명자
힐제니퍼제이
미국 02143 매사추세츠주 서머빌 서머 스트리트 157알 아파트먼트 16

울프맨닐엠
미국 02030 매사추세츠주 도버 필립스 레인 5

(74) 대리인
특허법인코리아나

심사청구 : 없음

(54) 폴리스타틴 도메인을 포함하는 단백질

요약

본 발명은 성장 및 분화 인자-8 (GDF-8) 의 수준 또는 활성을 조절하기 위한 하나 이상의 폴리스타틴 도메인을 포함하는 단백질 GASP1 의 용도에 관한 것이다. 보다 구체적으로, 본 발명은 GDF-8 의 수준 또는 활성 조절에 관련된 장애를 치료하기 위한 GASP1 의 용도에 관한 것이다. 본 발명은 특히 근육 조직의 증가가 치료적으로 유익한, 근육 질환 및 장애의 치료에 유용하다. 본 발명은 또한 대사, 지방세포 조직 및 골 퇴행에 관련된 질환 및 장애의 치료에 유용하다.

대표도

도 4

명세서

본 출원은 2002 년 2 월 21 일에 출원된 미국 가출원 제 60/357,846 호 및 2002 년 12 월 20 일에 출원된 미국 가출원 제 60/434,645 호를 우선권으로 청구한다.

기술분야

본 발명은 성장 및 분화 인자-8 (GDF-8) 의 수준 또는 활성을 조절하기 위한 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 용도에 관한 것이다. 더 구체적으로는, 본 발명은 GDF-8 의 수준 또는 활성의 조절과 연관된 장애의 치료를 위한, 폴리스타틴 자체를 제외한, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 용도에 관한 것이다. 특히 근육 조직의 증가가 치료적으로 유익한, 근육 질환 및 장애의 치료에 유용하다. 본 발명은 또한 대사, 지방세포 조직 및 골 퇴행과 관련된 질환 및 장애의 치료에 유용하다.

배경기술

미오스타틴 (myostatin) 으로도 공지된 성장 및 분화 인자-8 (GDF-8) 은 구조적으로 관련된 성장 인자의 형질전환 성장 인자-베타 (TGF-β) 수퍼패밀리의 멤버로서, 이들 전부 중요한 생리학적 성장-조절 및 형태형성 특성을 갖는다 (Kingsley 등 (1994) Genes Dev., 8: 133-46; Hoodless 등 (1998) Curr. Topics Microbiol. Immunol., 228: 235-72). GDF-8 은 골격근 질량의 음성적 조절자이며, 그의 생물학적 활성을 조절하는 인자의 동정에 상당한 관심이 집중되고 있다. 예를 들어, GDF-8 은 발생 시 및 성인 골격근에서 고발현된다. 트랜스제닉 마우스에서 GDF-8 이 없는 돌연변이는 골격근의 두드러진 비대 및 증식을 특징으로 한다 (McPherron 등 (1997) Nature, 387: 83-90). 골격근 질량에서의 유사한 증가가 소에서의 자연발생적인 GDF-8 의 돌연변이에서 뚜렷하다 (Ashmore 등 (1974) Growth, 38: 501-507; Swatland and Kieffer (1994) J. Anim. Sci., 38: 752-757; McPherron and Lee (1997) Proc. Natl. Acad. Sci. U.S.A., 94: 12457-12461; 및 Kambadur 등 (1997) Genome Res., 7: 910-915). 최근 연구는 또한 인간에서의 HIV-감염과 연관된 근육 손실에 GDF-8 단백질 발현의 증가가 동반된다는 것을 보였다 (Gonzalez-Cadavid 등 (1998) Proc. Natl. Acad. Sci. U.S.A., 95: 14938-43). 또한, GDF-8 은 근육 특이적인 효소 (예를 들어, 크레아틴 키나아제) 의 생산 및 근육모세포 증식을 조절할 수 있다 (WO 00/43781).

다수의 인간 및 동물 장애가 근육 조직의 손실 또는 기능적 부전과 연관되어 있다. 오늘날까지, 상기 장애에 대해서는 신뢰할 수 있거나 유효한 치료법이 거의 없다. 그러나, 상기 장애와 연관된 심각한 증상들은, 상기 장애를 앓는 환자에서 근육 조직의 양을 증가시키는 치료법을 이용함으로써 실질적으로 감소될 수 있다. 상태를 치유시키지 않으면서도, 상기 치료법들은 상기 환자들의 삶의 질을 현저히 개선시키며, 상기 질환의 일부 영향을 완화시킬 수 있다. 따라서, 상기 장애를 앓는 환자에서 근육 조직의 전반적인 증가에 기여할 수 있는 신규한 치료법을 밝혀내는 것이 당분야에 요구된다.

골격근에서의 그의 성장 조절 및 형태형성 특성에 더하여, GDF-8 은 또한 다수의 기타 생리학적 과정 (예를 들어, 글루코오스 항상성) 뿐만 아니라, 비정상적인 상태, 예컨대 제 2 형 당뇨병 및 지방세포 조직 장애, 예컨대 비만의 발생에 관련될 수 있다. 예를 들어, GDF-8 은 전구지방세포의 지방세포로의 분화를 조절한다 (Kim 등 (2001) B.B.R.C. 281.: 902-906). 따라서, GDF-8 의 조절은 상기 질환의 치료에도 또한 유용할 수 있다.

GDF-8 단백질은 아미노 말단 프로펩티드 및 카르복시 말단 성숙 도메인으로 이루어진 전구체 단백질로서 합성된다 (McPherron and Lee, (1997) Proc. Nat. Acad. Sci. U.S.A., 94: 12457-12461). 절단 전에, 전구체 GDF-8 단백질은 단독이량체를 형성한다. 이어서, 아미노 말단 프로펩티드는 성숙 도메인으로부터 절단된다. 절단된 프로펩티드는 성숙 도메인 이량체에 비공유적으로 결합된 채 남아, 그의 생물학적인 활성을 불활성화시킬 수 있다 (Miyazono 등 (1988) J. Biol. Chem., 263: 6407-6415; Wakefield 등 (1988) J. Biol. Chem., 263: 7646-7654; 및 Brown 등 (1990) Growth Factors, 3: 35-43). 2 개의 GDF-8 프로펩티드가 GDF-8 성숙 이량체에 결합하는 것으로 여겨진다 (Thies 등 (2001) Growth Factors, 18: 251-259). 상기의 불활성화 특성으로 인해, 프로펩티드는 '잠재-연합 펩티드 (latency-associated peptide)' (LAP) 로 공지되어 있으며, 성숙 도메인 및 프로펩티드의 복합체는 일반적으로 '소형 잠재 복합체' 로 불린다 (Gentry 및 Nash (1990) Biochemistry, 29: 6851-6857; Derynck 등 (1995) Nature, 316: 701-705; 및 Massague (1990) Ann. Rev. Cell Biol., 12: 597-641). 기타 단백질들이 또한 GDF-8 또는 구조적으로 연관된 단백질에 결합하며, 이들의 생물학적 활성을 저해하는 것으로 공지되어 있다. 상기 저해 단백질에는 폴리스타틴이 포함된다 (Gamer 등 (1999) Dev. Biol., 208: 222-232). GDF-8 의 성숙 도메인은 프로펩티드가 제거되는 경우 단독이량체로서 활성화되는 것으로 여겨진다.

명백하게, GDF-8 은 다수의 중요한 생물학적 과정의 조절에 관여된다. 상기 과정들에서 그의 중요한 기능으로 인해, GDF-8 은 치료적 개입을 위한 바람직한 표적일 수 있다. 특히, GDF-8 의 활성을 저해하는 치료제는 근육 조직의 증가가 치료적으로 유익한 인간 또는 동물 장애의 치료에 사용될 수 있다.

하나 이상의 폴리스타틴 도메인을 포함하는 공지된 단백질들은 다수의 생물학적 과정들, 특히 TGF-β 수퍼패밀리 신호전달의 조절 및 세포외 매트릭스-매개 과정, 예컨대 세포 유착의 조절에서 중요한 역할을 담당한다. 폴리스타틴, 폴리스타틴 관련 유전자 (FLRG, FSRP), 및 폴리스타틴-관련 단백질 (FRP) 은 모두 TGF-β 에 의한 전사 조절을 통해 (Bartholin 등 (2001), Oncogene, 20: 5409-5419; Shibanuma 등 (1993) Eur. J. Biochem. 217: 13-19) 또는 그의

TGF-β신호전달 경로의 길항능을 통해 (Phillips and de Kretser (1998) Front. Neuroendocrin., 19: 287-322; Tsuchida 등 (2000) J. Biol Chem., 275: 40788-40796; Patel 등 (1996) Dev. Biol., 178: 327-342; Amthor 등 (1996) Dev. Biol., 178: 343-362), TGF-β신호전달에 연관된다. 괄호 안의 단백질 명칭은 대안적인 명칭이다.

인슐린 성장 인자 결합 단백질 7 (IGFBP7, mac25) 은, 하나 이상의 폴리스타 틴 도메인을 포함하며, 인슐린에 결합하고, 인슐린 수용체와의 후속 상호작용을 차단한다. 또한, IGFBP7 은 TGF-β패밀리 멤버인 액티빈에 결합하는 것으로 나타났다 (Kato (2000) Mol. Med., 6: 126-135).

아그린 (agrin) 및 아그린 관련 단백질은 상류방향에 9 개의 폴리스타틴 도메인을 포함하며, 신경 세포로부터 분비되어 아세틸콜린 수용체 및 시냅스 형성에 관여되는 기타 분자들의 응집을 촉진한다. 폴리스타틴 도메인이 성장 인자를 시냅스로 편재시키는데 작용할 수 있다는 것이 제안되었다 (Patthy 등 (1993) Trends Neurosci., 16: 76-81).

오스테오넥틴 (SPARC, BM40) 및 헤빈 (SCI, mast9, QR1) 은 세포외 매트릭스 단백질과 상호작용하며 세포 성장 및 유착을 조절하는 밀접하게 관련된 단백질이다 (Motamed (1999) Int. J. Biochem. Cell. Biol., 31: 1363-1366; Girard 및 Springer (1996) J. Biol. Chem., 271: 4511-4517). 상기 단백질들은 하나 이상의 폴리스타틴 도메인을 포함한다.

기타 폴리스타틴 도메인 단백질이 NCBI 데이타베이스 (National Center for Biotechnology Information, Bethesda, Maryland, USA) 에 등재되어 있거나 밝혀져 있지만, 이들의 기능은 현재 알려져 있지 않다. 상기 단백질에는 U19878 (G01639, 토모레귤린-1 (tomoregulin-1) 과 매우 유사), T46914, 인간 GASP1 (GDF-assocoated serum protein 1; 본원에 기재됨; 도 7), 인간 GASP2 (WFIKKN; Trexler 등 (2001) Proc. Natl. Acad. Sci. U.S.A., 98: 3705-3709; 도 9), 및 테스티칸 (testican; SPOCK) 단백질의 프로테오글리칸 패밀리 (Alliel 등 (1993) Eur. J. Biochem., 214. 347-350) 가 포함된다. 마우스 GASP1 (도 6) 및 마우스 GASP2 (도 8) 에 대한 아미노산 및 뉴클레오티드 서열이 또한 Celera 데이타베이스 (Rockville, MD) 로부터 결정되었다. 본원에 기재된 바와 같이, 클로닝된 마우스 GASP1 의 뉴클레오티드 서열은, 예측된 아미노산 서열을 변화시키지 않는 워블 코돈 (wobble codon) 에서의 일부 염기 치환을 제외하고는 예측된 Celera 서열과 일치했다 (도 13 참조).

발명의 상세한 설명

발명의 개요

따라서, 본 발명은 독특한 구조적 특징, 즉 하나 이상의 폴리스타틴 도메인의 존재를 포함하는, 폴리스타틴 이외의 단백질에 관한 것이다. 폴리스타틴 자체는 본 발명에 포함되지 않는다. 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은, GDF-8 단백질이 단량체 형태이든, 이량체 활성 형태이든, 또는 GDF-8 잠재 복합체 중에 착화된 형태이든 상관없이 성숙 GDF-8 단백질 또는 그의 절편에 특이적으로 반응성이다. 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 성숙 GDF-8 단백질 상의 에피토프에 결합하여 GDF-8 와 관련된 하나 이상의 생물학적 활성을, 동일한 단백질에 의해 결합되지 않은 성숙 GDF-8 단백질에 비해 감소시킬 수 있다.

본 발명은 세포에 대한 GDF-8 의 영향을 조절하는 방법을 제공한다. 상기 방법은 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 유효량을 투여하는 것을 포함한다. 본 발명은 또한 하나 이상의 폴리스타틴 도메인을 포함하는 단백질을 코딩하는 DNA 분자를 투여함으로써 세포 내에서 단백질의 발현시키는 방법이 포함된다.

본 발명에 따르면, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은, 근육 조직의 증가가 치료적으로 유익한 의학적 상태를 치료 또는 예방하기 위해 치료적으로 유효한 용량으로 환자에게 투여될 수 있다. 구현예에는, GDF-8 의 생산, 대사 또는 활성과 연관된 세포 및 조직에 관여되는 질환, 장애 및 손상의 치료가 포함된다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 약학 제제로 제조될 수 있다. 상기 약학 제제는, 단량체 형태이든, 이량체 활성 형태이든, 또는 GDF-8 잠재 복합체 중에 착화된 형태이든 상관없이 성숙 GDF-8 단백질 또는 그의 절편의 결합을 보조하는 기타 성분들을 포함할 수 있다.

또한, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은, 단량체이든, 이량체 활성 형태이든, 또는 GDF-8 잠재 복합체 내에 착화되어 있든 성숙 GDF-8 단백질 또는 그의 절편을 정량적으로 또는 정성적으로 검출하기 위한 진단 도구로서 이용될 수 있다. 예를 들어, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 세포, 체액, 조직 또는 기관 내에서의 GDF-8 단백질의 존재, 부재 또는 양을 검출하기 위해 이용될 수 있다. 검출된 성숙 GDF-8 단백질의 존재 또는 양은 본원에 제시된 하나 이상의 의학적 상태와 연관될 수 있다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은, 단량체 형태이든, 이량체 활성 형태이든, 또는 GDF-8 잠재 복

합체 중에 착화된 형태이든 상관없이 성숙 GDF-8 단백질 또는 그의 절편을 검출하기 위한 진단용 키트 중에 제공되어, 본원에 기재된 하나 이상의 의학적 상태와 상기 결과를 연관짓는 것을 도울 수 있다. 상기 키트는 하나 이상의 폴리스타틴 도메인을 포함하는, 표지되거나 표지되지 않은 하나 이상의 단백질 및 상기 단백질에 결합하는 하나 이상의 제제, 예컨대 표지된 항체를 포함할 수 있다. 키트는 또한 실험의 검출 결과와 비교할 수 있는 적당한 생물학적 표준 및 대조군 표본을 포함할 수 있다. 또한 완충액 또는 세척 용액 및 키트를 이용하기 위한 지시서를 포함할 수 있다. 실험을 수행할 수 있게 하는 구조적인 성분, 예컨대 스틱, 비드, 종이, 칼럼, 바이알 또는 겔이 포함될 수 있다.

도면의 간단한 설명

도 1 은 야생형 마우스 혈청 유래의 GDF-8 복합체의 항체 정제를 나타낸다. 은으로 염색된 환원 겔은 아가로스 비드에 공유적으로 커플링된 JA16 모노클로날 항체를 이용하여 야생형 마우스 혈청으로부터 정제된 단백질을 나타낸다. mock-커플링된 비드를 이용한 대조군 정제 (0) 를 함께 수행하였다. 완충액 (mock 용출), 경합 펩티드 및 SDS 표본 완충액을 사용한 후속 용출은, JA16-콘쥬게이트된 비드로부터 유래한 펩티드로 특이적으로 용출된 2 개의 가시적인 단백질 (화살표로 나타냄) 을 나타내었다.

도 2 는 일반 마우스 혈청 유래의 친화도 정제 표본 중 성숙 및 미가공 GDF-8 의 동정을 나타낸다. 도 2A 는 친화도 정제 표본 내에서 보이는 12 kDa 밴드로부터 동정된 GDF-8 유래 펩티드 (서열 번호: 19) 의 대표적인 MS/MS 스펙트럼을 나타낸다. N-말단 절편 이온 (b 이온) 및 C-말단 절편 이온 (y 이온) 이 모두 보인다. 특히, 가장 진한 y 절편 이온은 프롤린 함유 펩티드의 일반적인 특징인, 프롤린 잔기 이전의 절편화로부터 초래된다. 도 2B 는 GDF-8 의 성숙 영역을 인식하는 폴리클로날 항체를 붙인 웨스턴 블롯을 나타내며, 친화도 정제 표본 내 GDF-8 의 존재를 확인시킨다. GDF-8 의 성숙 및 미가공 형태가 모두 보인다.

도 3 은 GDF-8 프로펩티드 및 폴리스타틴-유사 관련 유전자 (FLRG) 가 일반 마우스 혈청으로부터 단리된 순환 GDF-8 에 결합하는 것을 나타낸다. 36 kDa 밴드에서 동정된 GDF-8 프로펩티드 (서열 번호: 23) (도 3A) 및 FLRG (서열 번호: 30) (도 3C) 유래 펩티드의 대표적인 MS/MS 스펙트럼을 나타낸다. 도 3B 는 GDF-8 의 프로펩티드 영역을 특이적으로 인식하는 폴리클로날 항체를 붙인 친화도 정제 GDF-8 복합체의 웨스턴 블롯을 나타내며, GDF-8 복합체 내에서의 상기 단백질의 질량 스펙트럼 동정을 확인시킨다. 잘려진 프로펩티드 및 미가공 GDF-8 이 모두 보인다 -- 더 긴 노출시, 미가공 GDF-8 을 또한 SDS 용출 표본에서 볼 수 있다. 도 3D 는 FLRG 에 대한 모노클로날 항체를 붙인 친화도 정제 GDF-8 복합체의 웨스턴 블롯을 나타낸다.

도 4 는 혈청 중 GDF-8 프로펩티드, FLRG, 및 주요 GDF-8 결합 단백질로서의 신규 단백질이 동정된 대규모 GDF-8 정제의 분석을 통한 결과를 나타낸다. 은으로 염색된 겔을 음성 대조군 및 JA16 면역침전물의 두 펩티드 용출 표본으로부터 13 개의 조각으로 절단했다. 각 조각 내의 단백질을 트립신으로 소화시키고, 나노플로우-LC-MS/MS 및 데이타베이스 검색을 이용하여 동정했다. JA16 표본에 독특한 단백질에는 미가공 및 성숙 GDF-8, GDF-8 프로펩티드, FLRG 및 신규 멀티도메인 프로테아제 저해제 (*G* DF- *a* ssociated *s* erum *p* rotein 1, GASP1) 만이 포함되었다. 상기 단백질은 겔의 표시된 영역으로부터 동정되었다.

도 5 는 신규 멀티도메인 프로테아제 저해제인 GASP1 이 혈청 중 GDF-8 에 결합하는 것을 보여준다. 도 5A (서열 번호 31 로 정해진 펩티드) 및 5B (서열 번호 33 으로 정해진 펩티드) 는, 도 4 의 은으로 염색된 겔의 밴드 3 에서 동정된 2 개의 GASP1 펩티드로부터의 대표적인 MS/MS 스펙트럼을 보여준다.

도 6A 은 마우스 GASP1 에 대해 예측된 뉴클레오티드 서열을 나타낸다. 도 6B 는 마우스 GASP1 에 대해 예측된 대안적 뉴클레오티드 서열을 나타낸다. 도 6C 는 도 6A 및 6B 에서 나타낸 뉴클레오티드 서열에 의해 코딩되는 예측된 아미노산 서열을 나타낸다. 2 개의 뉴클레오티드 서열에 의해 코딩되는 단백질 서열은, 뉴클레오티드 차이가 모두 워블 코돈 위치에 있기 때문에 동일하다. 폴리스타틴 도메인은 굵은 글자 및 밑줄로 나타낸다.

도 7A 은 인간 GASP1 의 예측된 뉴클레오티드 서열을 나타낸다. 도 7B 은 대응하는 예측된 아미노산 서열을 나타낸다. 폴리스타틴 도메인은 굵은 글자 및 밑줄로 나타낸다. 도 7C 은 대안적 개시 부위를 이용한 인간 GASP1 의 예측된 뉴클레오티드 서열을 나타낸다. 도 7D 는 대응하는 예측된 아미노산 서열을 나타낸다. 폴리스타틴 도메인은 굵은 글자 및 밑줄로 나타낸다. 서열의 말단은 별표로 표시했다.

도 8A 는 마우스 GASP2 에 대해 예측된 뉴클레오티드 서열을 나타내며, 도 8B 은 대응하는 예측된 아미노산 서열을 나타낸다. 폴리스타틴 도메인은 굵은 글자 및 밑줄로 나타낸다.

도 9A 는 인간 GASP2 에 대해 예측된 뉴클레오티드 서열을 나타내며, 도 9B 는 대응하는 예측된 아미노산 서열을 나타낸다. 폴리스타틴 도메인은 굵은 글자 및 밑줄로 나타낸다.

도 10 은 마우스 GASP1 이 다수의 성체 조직 및 발생 도중 발현된다는 것을 나타낸다. 상기 도는 마우스 GASP1 의

조직 발현 프로필을 나타낸다. GASP1 의 551 bp 절편은 Clontech (Palo Alto, CA) 으로부터의 보정된 제 1 쇄 cDNA 패널로부터 증폭되었다. 글리세르알데히드-3-포스페이트 데히드로게나제 (G3PDH) 의 일부를 대조군으로서 증폭했다. G3PDH 발현은 골격근에서는 높고 고환에서는 낮은 것으로 공지되어 있다. cDNA 패널은 G3PDH 에 더하여 β-액틴, 포스포리파아제 A2 및 리보솜 단백질 S29 에 대해서도 보정되었다.

도 11 은 인간 혈청에서 단리된 단백질을 나타낸다. JA16 면역침전물 또는 대조군 표본 (0) 유래의 단백질을 mock PBS 용출, 경합 펩티드 용출 또는 SDS 용출로 용출했다. 겔의 표시된 영역 내 단백질을 트립신으로 소화시키고, LS-MS/MS 및 데이타베이스 검색으로 분석했다. JA16 표본에는 존재하나 대조군 표본에는 존재하지 않는 단백질은 성숙 GDF-8 (밴드 16), GDF-8 프로펩티드 및 FLRG (밴드 11) 및 인간 GASP1 (밴드 4) 이다. 도 11B 는 성숙 GDF-8 을 인식하는 항체를 붙인 동일한 JA16 면역침전물의 웨스턴 블롯을 나타낸다. 인간 혈청으 로부터 단리된 성숙 및 미가공 GDF-8 에 해당하는 밴드가 보인다.

도 12 는 GDF-8 유래 펩티드 및 밴드 4, 11 및 16 (도 11) 으로부터 단리된 연합 단백질의 대표적인 질량 스펙트럼을 나타낸다. 펩티드 서열 및 N-말단 (b 이온) 및 C-말단 (y 이온) 이 나타난다. 동정된 펩티드의 완전한 목록은 표 1 에 제공된다. GASP1 펩티드 (서열 번호: 44) (도 12A), FLRG 펩티드 (서열 번호: 41) (도 12B), GDF-8 프로펩티드 펩티드 (서열 번호: 24) (도 12C), 및 성숙 GDF-8 펩티드 (서열 번호: 13) (도 12D) 로부터의 스펙트럼이 나타난다.

도 13 는 클로닝된 마우스 GASP1 의 뉴클레오티드 (서열 번호: 48) 및 아미노산 (서열 번호: 49) 서열을 나타낸다. JA16 친화도-정제 표본의 질량 스펙트럼측정으로 동정된 펩티드를 밑줄쳤다. 서열의 말단은 별표로 표시된다.

도 14A 는 GASP1 의 도메인 구조를 나타낸다. GASP1 는 아미노산 29 이후에 시그날 서열/절단 부위를 갖는다. 또한, GASP1 는 2 개의 Kunitz/BPTI 세린 프로테아제 저해제 도메인, 폴리스타틴 도메인 (Kazal 세린 프로테아제 저해제 모티브를 포함) 및 메탈로프로테아제를 저해할 수 있는 넥트린 도메인을 포함한다. 도 14B 는 마우스 및 인간 게놈 서열로부터 예측되는 GASP1 및 GASP2 의 계통수를 나타낸다. 마우스 및 인간 GASP1 는 90% 상동성이다. GASP1 및 GASP2 는 54% 상동성이다.

도 15 는 재조합적으로 제조된 GASP1 이 GDF-8 및 GDF-8 프로펩티드 모두에 각각 결합한다는 것을 나타낸다. (A) JA16 을 사용하여, 재조합적으로 정제된 GDF-8 및/또는 프로펩티드가 보충된 mock- 또는 GASP1-V5-His 전달감염된 COS 세포 조건화된 배지로부터 GDF-8 을 면역침전시켰다. 항-V5 (상단 패널), 항-GDF-8 (중간 패널), 또는 항-프로펩티드 폴리클로날 항체를 사용한 웨스턴 블롯을 이용하여, 면역침전물 중 상기 단백질이 존재하는지 여부를 결정하였다. (B) 재조합적으로 제조된 GASP1 단백질은 재조합적으로 정제된 GDF-8 및/또는 프로펩티드가 보충된 mock- 또는 GASP1-V5-His 조건화 배지로부터 항-V5 태그 항체에 의해 면역침전시켰다. 면역침전물을 (A) 에서와 같이 웨스턴 블롯으로 분석했다.

도 16 은 GASP1 이 액티빈 또는 TGF-β이 아닌 GDF-8 및 연관성이 높은 BMP-11 의 생물학적 활성을 저해한다는 것을 나타낸다. mock (빈 동그라미) 또는 GASP1-V5-His (칠한 사각형) 전달감염체로부터의 조건화된 배지의 여러 희석물을 (A) 10 ng/ml GDF-8, (B) 10 ng/ml BMP-11, (C) 10 ng/ml 액티빈, 또는 (D) 0.5 ng/ml TGF-β와 인큐베이션했다. 이어서, 상기 표본들을 A204 (A-C) 또는 RD (D) 세포에서 루시퍼라아제 리포터 활성 분석하여, 첨가된 성장 인자의 활성을 결정했다. 루시퍼라아제 활성은 상대적인 루시퍼라아제 단위로 나타낸다. 각각의 성장 인자로 얻어지는 활성은 칠한 마름모 및 짧은 점선으로 나타낸다. 분석에서 임의의 성장 인자가 첨가되지 않은 백그라운드 활성은 아무 표시없이 긴 점선으로 나타낸 바와 같이 낮다.

도 17 은 GDF-8 의 GASP1 저해 정도를 나타낸다. 정제된 GASP1 을 RD 세포에서의 $(CAGA)_{12}$ (서열 번호: 53) 루시퍼라아제 리포터 분석에서 20 ng/ml 의 미오스타틴을 저해하는 그 능력에 대해 시험했다 (칠한 사각형). GDF-8 단독에 의한 활성은 칠한 마름모 및 짧은 점선으로 나타낸다. 성장 인자를 첨가하지 않은 경우 존재하는 활성은 긴 점선으로 나타낸다.

정의

용어 ' **폴리스타틴 도메인** ' 은, 시스테인 풍부 반복부를 특징으로 하는 아미노산 도메인 또는 아미노산 도메인을 코딩하는 뉴클레오티드 도메인을 의미한다. 폴리스타틴 도메인은 전형적으로 65-90 개 아미노산 길이를 포함하며, 10 개의 보존된 시스테인 잔기 및 Kazal 세린 프로테아제 저해제 도메인과 유사한 영역을 포함한다. 일반적으로, 시스테인 잔기들 사이의 루프 영역은 폴리스타틴 도메인에서의 서열 가변성을 나타내나, 일부 보존성이 명백하다. 4 번째 및 5 번째 시스테인 사이의 루프는 일반적으로, 1 또는 2 개의 아미노산을 포함하여 작다. 7 번째 및 8 번째 시스테인 사이의 루프에서의 아미노산은 일반적으로 (G,A)-(S,N)-(S,N,T)-(D,N)-(G,N) 의 공통 서열에 이어 (T,S)-Y 모티브를 포함하여 가장 보존성이 높다. 9 번째 및 10 번째 시스테인 사이의 영역은 일반적으로, 또다른 아미노산으로 분리되는 2 개의 소수성 잔기 (구체적으로 V, I 또는 L) 을 포함하는 모티브를 포함한다.

용어 ' **하나 이상의 폴리스타틴 도메인을 포함하는 단백질** ' 은 하나 이상, 그러나 가능하다면 하나 초과의 폴리스타틴 도메인을 포함하는 단백질을 의미한다. 상기 용어는 또한 천연 단백질, 특히 GDF-8 결합 활성을 보유한 것과 연합되는 공지된 생물학적 활성을 유지하는, 아미노산 서열에 대한 보존성 또는 비보존성 변화로 개질된 서열을 갖는 서열을 포함하는 상기 단백질의 임의의 변이체 (절편; 치 환, 부가 또는 결실 돌연변이가 있는 단백질; 및 융합 단백질 포함) 도 의미한다. 상기 단백질들은 천연 또는 합성인 임의의 원으로부터 유래될 수 있다. 상기 단백질은 인간 또는 소, 닭, 설치류, 래트, 돼지, 양, 칠면조, 비비원숭이 및 어류를 포함하는 동물 원으로부터 유래될 수 있다. 폴리스타틴 자체는 본 발명에 포함되지 않는다.

용어 ' **GDF-8** ' 또는 ' **GDF-8 단백질** ' 은 특이적 성장 및 분화 인자를 의미한다. 상기 용어에는 단백질의 전장 미가공 전구체 형태 뿐만 아니라, 번역 후 절단으로 얻어지는 성숙 형태 및 프로펩티드 형태가 포함된다. 상기 용어는 또한, 아미노산 서열에 대한 보존성 또는 비보존성 변화로 개질된 서열을 포함하는, 단백질과 연합되는 공지된 생물학적 활성을 유지하는 GDF-8 의 임의 절편을 의미한다. 상기 GDF-8 분자들은 임의의 천연 또는 합성 원으로부터 유도될 수 있다. 상기 단백질은 인간 또는 소, 닭, 설치류, 래트, 돼지, 양, 칠면조, 비비원숭이 및 어류를 포함하는 동물 원으로부터 유래될 수 있다. 각종 GDF-8 분자들은 [McPherron 등 (1997) Proc. Natl. Acad. Sci. USA, 94: 12457-12461] 에 기재되어 있다.

' **성숙 GDF-8** ' 은 GDF-8 전구체 단백질의 카르복시 말단 도메인으로부터 절단된 단백질을 의미한다. 성숙 GDF-8 은 단량체로서, 단독이량체로서, 또는 GDF-8 잠재 복합체 중에 존재할 수 있다. 생체 내 또는 시험관 내 조건에 따라서, 성숙 GDF-8 은 상기 상이한 형태의 일부 또는 전부 사이에서 평형을 구축할 수 있다. 이는 단독이량체일 때 생물학적으로 활성인 것으로 여겨진다. 그의 생 물학적으로 활성인 형태에서, 성숙 GDF-8 은 또한 ' **활성 GDF-8** ' 로 불린다.

' **GDF-8 프로펩티드** ' 는 GDF-8 전구체 단백질의 아미노-말단 도메인으로부터 절단된 폴리펩티드를 의미한다. GDF-8 프로펩티드는 성숙 GDF-8 상의 프로펩티드 결합 도메인에 결합할 수 있다.

' **GDF-8 잠재 복합체** ' 는 성숙 GDF-8 단독이량체 및 GDF-8 프로펩티드 사이에 형성된 단백질의 복합체를 의미한다. 2 개의 GDF-8 프로펩티드가 단독이량체인 2 개 분자의 성숙 GDF-8 와 연합하여 불활성 4 량체 복합체를 형성하는 것으로 여겨진다. 상기 잠재 복합체는 하나 이상의 GDF-8 프로펩티드 대신, 또는 이에 부가하여 기타 GDF 저해제를 포함할 수 있다.

어구 ' **GDF-8 활성** ' 은 활성 GDF-8 단백질과 연합된 하나 이상의 생리학적 성장-조절 또는 형태형성 활성을 의미한다. 예를 들어, 활성 GDF-8 은 골격근의 음성적 조절자이다. 활성 GDF-8 은 또한 근육 특이적 효소 (예를 들어, 크레아틴 키나아제) 의 생산을 조절하고, 근육모세포 증식을 자극하며, 전구지방세포의 지방세포로의 분화를 조절할 수 있다. GDF-8 은 또한 말초 조직, 특히 골격근 또는 지방세포에서 인슐린 및 글루코오스 섭취에 대한 감수성을 증가시키는 것으로 여겨진다. 따라서, GDF-8 생물학적 작용에는, 근육 형성의 저해, 근육 세포 성장의 저해, 근육 발생의 저해, 근육 질량의 감소, 근육 특이적 효소의 조절, 근육모세포 증식의 저해, 전구지방세포의 지방세포로의 분화 조절, 인슐린에 대한 감수성 증가, 글루코오스 섭취 조절, 글루코오스 항상성 및 신경 세포 발생 및 유지의 조절이 포함되지만, 이에 제한되는 것은 아니다.

용어 ' **단리된** ' 또는 ' **정제된** ' 은 그의 천연 환경으로부터 실질적으로 자유로운 분자를 의미한다. 예를 들어, 단리된 단백질에는 그가 유래된 세포 또는 조직 원으로부터의 세포성 물질 또는 기타 오염 단백질이 실질적으로 없다. 어구 ' **세포성 물질이 실질적으로 없는** ' 은 단리된 단백질이 적어도 70% 내지 80% (w/w) 순수하거나, 적어도 80%-89% (w/w) 순수하거나, 적어도 90-95% 순수하거나, 또는 적어도 96%, 97%, 98%, 99% 또는 100% (w/w) 순수한 제제를 의미한다.

용어 ' **LC-MS/MS** ' 는 특정 질량/전하 비율의 분자 이온을 단리하고, 상기 이온을 분절화하여, 분절화된 이온의 질량/전하 비율을 기록하도록 프로그램된 질량 분광광도계가 있는 액체 크로마토그래피를 의미한다. 펩티드 표본을 분석하는 경우, 상기 기법은 액체 크로마토그래피를 통한 복합 표본의 상류 분리에 이어, 분절 이온 질량의 기록 및 펩티드 서열의 후속 결정을 가능하게 한다.

용어 ' **MS/MS** ' 는 특정 질량/전하 비율을 가진 분자 이온을 단리하고, 상기 이온을 분절화하고, 생성 분절 이온의 질량/전하 비율을 기록하기 위해 질량 분광광도계를 이용하는 과정을 의미한다. 상기 분절 이온은 펩티드의 서열에 대한 정보를 제공한다.

용어 ' **처치함** ' 또는 ' **처치** ' 는 치료적인 처치 및 예방 또는 방지적 측정을 모두 의미한다. 처치가 필요한 해당자에는, 특정 의학적 장애를 이미 가진 개인 뿐만 아니라 궁극적으로 상기 장애를 입을 수 있는 이들 (예를 들어, 방지적 측정이 필요한 이들) 이 포함될 수 있다. 용어 처치에는 장애의 잠재적인 원인을 알리는 측정 및 그의 원인에 반드시 영향을 주지 않고도 의학적 장애의 증상을 감 소시키는 측정이 모두 포함된다. 따라서, 삶의 질 개선 및 증상의 완화도, 장

애의 원인을 역행시키는 측정이므로, 처치로서 간주된다.

용어 ' **의학적 장애** ' 는 근육, 골 또는 글루코오스 항상성의 장애를 의미하며, GDF-8 및/또는 TGF-β수퍼패밀리의 기타 멤버 (예를 들어, BMP 11) 와 연관된 장애를 포함한다. 상기 장애의 예에는, 대사성 질환 및 장애, 예컨대 인슐린 의존성 (제 1 형) 당뇨병, 인슐린 비의존성 (제 2 형) 당뇨병, 고혈당증, 손상된 글루코오스 관용, 대사성 증후군 (예를 들어, 증후군 X), 및 외상 (예를 들어, 화상 또는 질소 불균형) 에 의해 유도되는 인슐린 저항성, 및 지방세포 조직 장애 (예컨대, 비만); 근육 및 신경근육 장애, 예컨대 근이영양증 (이에 제한되지는 않으나, 중증 또는 양성 X-연관 근이영양증, 사지연관 이영양증, 안견갑상완근 이영양증, 강직성 이영양증, 원위 근이영양증, 진행성 위축성 안근마비, 안인두 이영양증, 듀센씨 (Duchenne's) 근이영양증 및 파쿠야마 유형 (Fakuyama type) 선천성 근이영양증이 포함된다); 근위축성 측삭 경화증 (ALS); 근위축; 기관 위축; 쇠약; 손목 굴증후근; 선천성 폐색성 폐질환; 선천성 근육병증; 선천성 강직증; 가족성 주기 마비; 발작성 미오글로빈뇨증; 중증 근무력증; 이튼 램버트 (Eaton-Lambert) 증후군; 속발성 근무력증; 탈신경위축; 격발성 근위축; 및 사르코페니아, 악액질 및 기타 근육 소모 증후군이 포함되지만, 이에 제한되는 것은 아니다. 기타 예에는, 특히 노년 및/또는 폐경후 여성에서의, 골다공증; 글루코코르티코이드 유도 골다공증; 골감소증; 골관절염; 골다공증 관련 골절; 및 근육 조직의 외상성 또는 만성 손상이 포함된다. 또다른 추가적 예에는, 만성 글루코코르티코이드 치료법 으로 인한 낮은 골 질량, 조발성 생식기 부전, 안드로겐 억제, 비타민 D 결핍, 속발성 부갑상선항진증, 영양 결핍증 및 거식증이 포함된다.

용어 ' **질량 증가** ' 는 하나 이상의 폴리스타틴 도메인을 포함하는 단백질을 사용한 처치 후에, 처치 전 근육 질량의 양에 비해 더 많은 양의 근육의 존재를 의미한다.

용어 ' **치료적으로 유익한** ' 은 장애의 증상 개선, 장애의 진행 서행화, 또는 장애의 진행 정지를 의미한다. 치료적으로 유리함은 장애의 일면, 예컨대 하나 이상의 폴리스타틴 도메인을 포함하는 하나 이상의 단백질을 투여하기 전후의 근육 질량의 양을 비교함으로써 결정된다.

용어 ' **조절** ' 은 단백질의 활성, 거동 또는 양을 증가, 감소 또는 저해함으로써 단백질의 특성을 변화시키는 것을 의미한다. 예를 들어, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 그의 활성을 저해함으로써 GDF-8 을 조절할 수 있다.

용어 ' **안정화 개질** ' 은 단백질을 안정화시키고, 단백질의 시험관 내 반감기를 증강시키고, 단백질의 순환 반감기의 증강시키고/시키거나 단백질의 단백분해를 감소시킬 수 있는 것으로 당분야에서 또는 본원에서 공지된 임의의 개질을 의미한다. 상기 안정화 개질에는, 융합 단백질 (예를 들어, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질 및 제 2 단백질을 포함하는 융합 단백질을 포함), 글리코실화 부위의 개질 (예를 들어, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질에 대한 글리코실화 부위의 부가를 포함), 및 탄수화물 부분의 개질 (예를 들어, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질로부터의 탄수화물 부분의 제거를 포함) 이 포함되지만, 이에 제한되는 것은 아니다. 융합 단백질을 포함하는 안정화 개질의 경우 (예를 들어, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질에 제 2 단백질을 융합시키는 경우), 제 2 단백질은 ' **안정화 부분** ' 또는 ' **안정화 단백질** ' 로 명명될 수 있다. 예를 들어, 하나 이상의 폴리스타틴 도메인을 포함하는 인간 단백질이 IgG 분자에 융합되는 경우, IgG 는 안정화 단백질 또는 안정화 부분으로서 작용한다. 본원에서 사용되는 바와 같이, 융합 단백질의 제 2 단백질에 대한 언급에 부가하여, '안정화 부분' 은 또한 비단백질성 개질, 예컨대 탄수화물 부분 또는 비단백질성 중합체를 포함한다.

용어 ' **IgG 분자의 Fc 영역** ' 은 당업자에게 널리 공지된 바와 같이 이소형 (isotype) IgG 의 면역글로불린의 Fc 도메인을 의미한다. IgG 분자의 Fc 영역은, IgG 분자의 생체 내 혈청 반감기를 증가시키는데 관여하는 IgG 분자 (IgG1, IgG2, IgG3 및 IgG4) 의 부분이다.

' **시험관 내 반감기** ' 는 생명체의 몸체 외부에서 측정된 단백질의 안정성을 의미한다. 시험관 내 반감기를 측정하기 위한 분석은 당분야에 널리 공지되어 있으며, SDS-PAGE, ELISA, 세포-기재 분석, 펄스-체이스 (pulse-chase), 웨스턴 블롯, 노던 블롯 등이 포함되지만, 이에 제한되는 것은 아니다. 상기 및 기타 유용한 분석은 당분야에 널리 공지되어 있다.

' **생체 내 반감기** ' 는 유기체 내에서의 단백질의 안정성을 의미한다. 생체 내 반감기는, 생체 내 혈청 반감기, 순환 반감기 및 본원의 실시예에 나타낸 분 석을 포함하지만 이에 제한되지는 않는 당분야에 공지된 다수의 방법으로 측정될 수 있다.

' **생체 내 혈청 반감기** ' 는 유기체의 혈액 중 순환하는 단백질의 반감기를 의미한다. 당분야에 공지된 방법을 이용하여 생체 내 혈청 반감기를 측정할 수 있다. 예를 들어, 방사활성 단백질을 동물에 투여하여 혈청 중 표지된 단백질의 양을 경시적으로 모니터링할 수 있다.

명세서 및 도면에 열거된 서열의 확인을 돕기 위해 하기 표를 제공하며, 이는 서열 번호, 도면 위치 및 서열의 설명을

제시한다.

서열 번호	참고	설명
1	도 6A	예측된 마우스 GASP1 뉴클레오티드 서열
2	도 6B	예측된 마우스 GASP1 대안적 뉴클레오티드 서열
3	도 6C	서열 번호 1 및 2 에 의해 코딩되는 예측된 마우스 GASP1 아미노산 서열
4	도 7A	예측된 인간 GASP1 뉴클레오티드 서열
5	도 7B	서열 번호 4 에 의해 코딩되는 예측된 인간 GASP1 아미노산 서열
6	도 7C	예측된 인간 GASP1 뉴클레오티드 서열, 대안적 개시 부위
7	도 7D	예측된 인간 GASP1 아미노산 서열, 서열 번호 6 에 의해 코딩되는 대안적 개시 부위
8	도 8A	예측된 마우스 GASP2 뉴클레오티드 서열
9	도 8B	서열 번호 8 에 의해 코딩되는 예측된 마우스 GASP2 아미노산 서열
10	도 9A	예측된 인간 GASP2 뉴클레오티드 서열
11	도 9B	서열 번호 10 에 의해 코딩되는 예측된 인간 GASP2 아미노산 서열
12	실시예 2	경합 펩티드
13-20	표 1, 실시예 5, 6	마우스 GDF-8 펩티드
21-27	표 1, 실시예 5, 6	마우스 GDF-8 프로펩티드 펩티드
28-30	표 1, 실시예 5	마우스 FLRG 펩티드
31-35	표 1, 실시예 5, 7	마우스 GASP1 펩티드
36-37	표 1, 실시예 8	인간 GDF-8 펩티드
38-39	표 1, 실시예 8	인간 GDF-8 프로펩티드 펩티드
40-42	표 1, 실시예 8	인간 FLRG 펩티드
43-45	표 1, 실시예 8,	인간 GASP1 펩티드
46	실시예 7	전방 프라이머
47	실시예 7	후방 프라이머
48	도 13	클로닝된 마우스 GASP1 뉴클레오티드 서열
49	도 13	서열 번호 48 에 의해 코딩되는 클로닝된 마우스 GASP1 아미노산 서열
50	실시예 9	전방 프라이머
51	실시예 9	후방 프라이머
52	실시예 9	예시적 N-말단 펩티드 서열
53	실시예 11	합성 올리고뉴클레오티드

하나 이상의 폴리스타틴 도메인을 포함하는 단백질

본 발명은 폴리스타틴 이외의, 독특한 구조적 특징, 즉 하나 이상의 폴리스타틴 도메인을 포함하는 단백질에 관한 것이다. 폴리스타틴 자체는 본 발명에 포함되지 않는다. 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 GDF-8 에 결합하여 이를 저해할 것으로 여겨진다. 하나 이상의 폴리스타틴 도메인을 갖는 단백질의 예에는, 폴리스타틴-유사 관련 유전자 (FLRG), FRP (flik, tsc36), 아그린, 오스테오넥틴 (SPARC, BM40), 헤빈 (SC1, mast9, QR1), IGFBP7 (mac25), 및 U19878 이 포함되지만, 이에 제한되지는 않는다. 도 6 및 7 에 제공되는 뉴클레오티드 및 아미노산 서열을 포함하는 GASP1 및 도 8 및 9 에 제공되는 뉴클레오티드 및 아미노산 서열을 포함하는 GASP2 는 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 또다른 예이다.

상기 언급된 바와 같이, 폴리스타틴 도메인은 시스테인 풍부 반복부를 특징으로 하는 아미노산 도메인 또는 아미노산 도메인을 코딩하는 뉴클레오티드 도메인으로 정의된다. 폴리스타틴 도메인은 전형적으로는 65-90 개 아미노산 길이를 포함하며, 10 개의 보존성 시스테인 잔기 및 Kazal 세린 프로테아제 저해제 도메인과 유사한 영역을 포함한다. 일반적으로, 시스테인 잔기들 사이의 루프 영역은 폴리스타틴 도메인에서의 서열 가변성을 나타내나, 일부 보존성은 명백하다. 4 번째 및 5 번째 시스테인들 사이의 루프는 일반적으로 1 또는 2 개의 아미노산만을 포함하여 작다. 7 번째 및 8 번째 시스테인 사이의 루프에서의 아미노산은 일반적으로 (G,A)-(S,N)-(S,N,T)-(D,N)-(G,N) 의 공통 서열에 이어 (T,S)-Y 모티브를 포함하여 가장 보존성이 높다. 9 번째 및 10 번째 시스테인 사이의 영역은 일반적으로 또다른 아미노산으로 분리되는 2 개의 소수성 잔기 (구체적으로, V, I 또는 L) 를 포함하는 모티브를 포함한다.

GDF-8 에 결합할 수 있는, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 다양한 방법을 이용하여 단리될 수 있다. 예를 들어, 본 발명에서 예시된 바와 같이, GDF-8 을 이용하는 친화도 정제를 이용할 수 있다. 또한, cDNA 라이브러리의 낮은 엄격성 스크리닝 (low stringency screening) 을 이용하거나, 또는 폴리스타틴 도메인에 의해 만들어진 탐침을 이용하는 축퇴성 (degenerate) PCR 기법을 이용할 수 있다. 더 많은 게놈 데이타를 이용가능해질수록, 다수의 서열 프로파일링 및 분석 프로그램, 예컨대 MotifSearch (Genetics Computer Group, Madison, WI), Profile Search (GCG), 및 BLAST (NCBI) 을 이용한 유사성 검색이 이용되어 공지된 폴리스타틴 도메인과 상당한 상동성을 갖는 신규 단백질을 찾을 수 있다.

당업자는 GDF-8 또는 하나 이상의 폴리스타틴 도메인을 포함하는 단백질이 모두 이들의 생물학적 특성이 변형되지 않으면서 이들의 각각의 아미노산 서열에 임의 수의 보존성 변화를 포함할 수 있음을 인지할 것이다. 상기 보존성 아미노산 개질은 아미노산 측쇄 치환기의 상대적인 유사성, 예를 들어 이들의 소수성, 친수성, 전하, 크기 등에 근거한다. 각종 예상되는 특징을 고려한 보존성 치환의 예는 당업자에게 널리 공지되어 있으며, 하기가 포함된다: 아르기닌 및 라이신; 글루타메이트 및 아스파르테이트; 세린 및 트레오닌; 글루타민 및 아스파라긴; 및 발린, 류신 및 이소류신. 더욱이, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질을 이용하여 하나 이상의 폴리스타틴 도메인을 포함하는 기능적 절편을 생성시킬 수 있다. 상기 절편들이 GDF-8 에 결합하여 이를 저해할 것으로 예측된 다. 본 발명의 하나의 구현예에 있어서, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은, 단량체 형태이든, 이량체 활성 형태이든, 또는 GDF-8 잠재 복합체 중에 착화된 형태이든 상관없이 성숙 GDF-8 또는 이들의 절편에 0.001 내지 100 nM, 또는 0.01 내지 10 nM, 또는 0.1 내지 1 nM 의 친화도로 특이적으로 결합한다.

뉴클레오티드 및 단백질 서열

항상 필요한 것은 아니지만, 원하는 경우, 당업자는 하나 이상의 폴리스타틴 도메인을 포함하는 신규 단백질의 아미노산 또는 핵산 서열을 결정할 수 있다. 예를 들어, 본 발명은 도 6 - 9 에 나타낸 바와 같이 GASP1 및 GASP2 에 대한 아미노산 및 뉴클레오티드 서열을 제공한다.

본 발명은 또한 상기 핵산 및 아미노산 서열의 변이체, 상동체 및 절편을 포함한다. 예를 들어, 상기 핵산 및 아미노산 서열은 천연 단백질의 핵산 또는 아미노산 서열과 적어도 70% 내지 79%, 또는 적어도 80% 내지 89%, 또는 적어도 90% 내지 95%, 또는 적어도 96% 내지 100% 의 상동성을 가진 서열을 포함할 수 있다. 당업자는 GDF-8 에 결합하는 영역이 결합에 관여되지 않는 단백질의 다른 부분에 비해 더 적은 서열 변이를 관용할 수 있음을 인지할 것이다. 따라서, 단백질의 상기 비결합 영역은 단백질의 결합 특성을 현저히 변형시키지 않으면서 실질적인 변이를 포함할 수 있다. 그러나, 당업자는 또한 다수의 변화가 그의 표적에 대한 단백질의 친화도를 특이적으로 증가시킬 수 있다는 것을 인지할 것이다. 상기 친화도를 증가시키는 변화는 전형적으로, 단백질을 상기 결합 영역에서 변형시키고 GDF-8 에의 결합능 또는 결합의 강도에 대해 시험함으로써 결정된 다. 결합 영역의 내부이든 또는 외부이든 상관없이 상기의 모든 변형은 본 발명의 범위에 포함된다.

상대적인 서열 유사성 또는 상동성은 Sequence Analysis Software Package ™ (Version 10; Genetics Computer Group, Inc., University of Wisconsin Biotechnology Center, Madison, WI) 의 'Best Fit' 또는 'Gap' 프로그램을 이용하여 결정될 수 있다. 'Gap' 은 Needleman 및 Wunsch 의 알고리즘 (Needleman and Wunsch, 1970) 을 이용하여 대응부 (match) 의 수를 최대화하고 갭 (gap) 의 수를 최소화하는 2 개의 서열의 배치를 얻는다. 'BestFit' 은 2 개의 서열 사이의 유사성의 최적 절편의 최적 배치를 수행한다. 최적 배치는, Smith 및 Waterman 의 국소적 상동성 알고리즘 (Smith and Waterman, 1981; Smith 등, 1983) 을 이용하여 대응부의 수를 최대화하기 위해 갭을 삽입함으로써 얻어진다.

상기 기재된 Sequence Analysis Software Package 는 본문에 개시된 뉴클레오티드 및 아미노산 서열의 상동체를 동정하기 위한 다수의 기타 유용한 서열 분석 도구를 포함한다. 예를 들어, 'BLAST' 프로그램 (Altschul 등, 1990) 은 특화된 데이타베이스 (예를 들어, NCBI 에서 유지하는 서열 데이타베이스) 내에서의 탐색 서열 (펩티드 또는 핵산) 과 유사한 서열을 검색하며; 'FastA' (Lipman and Pearson, 1985; Pearson 및 Lipman, 1988; Pearson 등, 1990) 은 탐색 서열 및 동일한 유형 (핵산 또는 단백질) 의 서열군 사이의 유사성에 대한 Pearson 및 Lipman 검색을 수행하며; 'TfastA' 는 단백질 탐색 서열 및 뉴클레오티드 서열의 임의의 군 사이의 유사성에 대한 Pearson 및 Lipman 검색을 수행하며 (이는 비교 수행 전 6 개의 모든 리딩 프레임 내에서 뉴클레오티드 서열을 번역한다); 'FastX' 는 프레임변환을 고려하여, 뉴클레오티드 탐색 서열 및 일군의 단백질 서열 사이의 유사성에 대한 Pearson 및 Lipman 검색을 수행한다. 'TfastV' 는 프레임변환을 고려하여, 단백질 탐색 서열 및 뉴클레오티드 서열의 임의의 군 사이의 유사성에 대한 Pearson 및 Lipman 검색을 수행한다 (이는 비교 수행 전 핵산 서열의 두 가닥을 번역한다).

개질된 단백질

본 발명은 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 절편을 포함한다. 상기 절편들은 폴리스타틴 도메인의 전부 또는 일부를 포함할 것이다. 절편들은 폴리스타틴 도메인 및 N-말단 사이 및/또는 폴리스타틴 도메인 및 C-말단 사이 서열의 전부, 일부를 포함하거나 또는 포함하지 않을 수 있다.

당업자는 특정 아미노산이, 단백질의 활성, 예를 들어 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 결합 특징에 악영향을 미치지 않고도 단백질 구조 내의 다른 아미노산에 대해 치환될 수 있다는 것을 이해하고 있다. 따라서, 본 발명의 발명자들은 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 아미노산 서열, 또는 상기 단백질을 코딩하는 DNA 서열에, 이들의 생물학적 유용성 또는 활성에 상당한 손실없이 각종 변화가 가해질 수 있다는 것을 고려했다. 상기 변화에는 모티브 서열의 결실, 삽입, 절단, 치환, 융합, 셔플링 등이 포함될 수 있다.

상기 변화의 수행에서, 아미노산의 친소수성 지수 (hydropathic index) 가 고려될 수 있다. 단백질 상에서의 상호 생물학적 기능을 부여하는데 있어서 친소수성 아미노산 지수의 중요성은 당 분야에서 일반적으로 이해되어 있다 (Kyte 및 Doolittle (1982) J. Mol. Biol, 157: 105-132). 아미노산의 상대적 친소수성 특징은 생성 단백질의 2 차 구조에 기여하며, 이어서 단백질과 기타 분자들, 예를 들어 효소, 기질, 수용체, DNA, 항체, 항원 등의 상호작용을 정의하는 것으로 인정된다.

각각의 아미노산은 그의 소수성 및 전하 특징을 근거로 친소수성 지수가 지정된다 (Kyte 및 Doolittle, 1982); 이들은 이소류신 (+4.5), 발린 (+4.2), 류신 (+3.8), 페닐알라닌 (+2.8), 시스테인/시스틴 (+2.5), 메티오닌 (+1.9), 알라닌 (+1.8), 글리신 (-0.4), 트레오닌 (-0.7), 세린 (-0.8), 트립토판 (-0.9), 타이로신 (-1.3), 프롤린 (-1.6), 히스티딘 (-3.2), 글루타메이트 (-3.5), 글루타민 (-3.5), 아스파르테이트 (-3.5), 아스파라긴 (-3.5), 라이신 (-3.9) 및 아르기닌 (-4.5) 이다. 상기 변화의 수행에 있어서, 그의 친소수성 지수가 ±2 이내, ±1 이내, 및 ±0.5 이내에 있는 아미노산의 치환이 가능하다.

또한, 유사 아미노산의 치환이 친수성에 근거하여 효과적으로 수행될 수 있음이 이해된다. U. S. 특허 제 4,554,101 호에는 그의 인접 아미노산의 친수성에 의해 조절되는 단백질의 최대 국소 평균 친수성이 단백질의 생물학적 특성과 연관된다는 것이 기재되어 있다.

U. S. 특허 제 4,554,101 호에 상세히 기재된 바와 같이, 하기의 친수성 값 이 아미노산 잔기에 대해 지정되었다: 아르기닌 (+3.0), 라이신 (+3.0), 아스파르테이트 (+3.0±1), 글루타메이트 (+3.0±1), 세린 (+0.3), 아스파라긴 (+0.2), 글루타민 (+0.2), 글리신 (0), 트레오닌 (-0.4), 프롤린 (-0.5±1), 알라닌 (-0.5), 히스티딘 (-0.5), 시스테인 (-1.0), 메티오닌 (-1.3), 발린 (-1.5), 류신 (-1.8), 이소류신 (-1.8), 타이로신 (-2.3), 페닐알라닌 (-2.5) 및 트립토판 (-3.4). 상기 변화의 수행에 있어서, 친수성 값이 +2 이내, ±1 이내, 및 ±0.5 이내에 있는 아미노산의 치환이 가능하다 .

개질은 상기 단백질의 구조 및 생물학적 기능이 상기 변화에 의해 영향받지 않도록 보존성일 수 있다. 상기 보존성 아미노산 개질은 아미노산 측쇄 치환기의 상대적인 유사성, 예를 들어, 이들의 소수성, 친수성, 전하, 크기 등에 근거한다. 각종 예견되는 특징을 고려한 예시적 보존성 치환이 당업자에게 널리 공지되어 있으며, 하기를 포함한다: 아르기닌 및 라이신; 글루타메이트 및 아스파르테이트; 세린 및 트레오닌; 글루타민 및 아스파라긴; 및 발린, 류신 및 이소류신. 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 아미노산 서열은, 그의 표적 항원에 대한 상기 단백질의 결합에 악영향을 미치지 않는 한 임의 수의 보존성 변화를 갖도록 개질될 수 있다. 상기 변화들은 하나 이상의 폴리스타틴 도메일을 포함하는 단백질 결합부의 내부 또는 외부에 도입될 수 있다. 예를 들어, 단백질의 항원 결합부의 내부에 도입되는 변화들은 그의 표적에 대한 단백질의 친화도를 증가시키도록 고안될 수 있다.

안정화 개질

안정화 개질은 단백질을 안정화시킬 수 있고, 단백질의 시험관 내 및/또는 생체 내 반감기를 증강시킬 수 있고, 단백질의 순환 반감기를 증강시키고/시키거나 단백질의 단백분해를 감소시킬 수 있다. 상기 안정화 개질에는 융합 단백질, 글리코실화 부위의 개질, 및 탄수화물 부분의 개질이 포함되지만, 이에 제한되는 것은 아니다. 안정화제 단백질은 개질된 GDF 프로펩티드의 전체 안정성을 증강시키는 임의의 단백질일 수 있다. 당업자에게 인지될 바와 같이, 상기 융합 단백질은 임의로 프로펩티드 부분 및 안정화 부분 사이에 링커 펩티드를 포함할 수 있다. 당분야에 널리 공지된 바와 같이, 융합 단백질은 제 2 단백질이 제 1 단백질의 프레임 내에 융합되어, 생성되는 번역된 단백질에 제 1 및 제 2 단백질이 둘 다 포함되도록 제조된다. 예를 들어, 본 발명에 있어서, 융합 단백질은 하나 이상의 폴리스타틴 도메인을 포함하는 단백질이 제 2 단백질 (예컨대, 안정화제 단백질 부분) 에 융합되도록 제조될 수 있다. 상기 융합 단백질은 생성되는 번역된 단백질이 프로펩티드 부분 및 안정화제 부분을 둘 다 포함하도록 제조된다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 글리코실화되거나, 알부민 또는 비단백질성 중합체에 연결될 수 있다. 예를 들어, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 U. S. 특허 제 4,640,835 호; 제 4,496,689 호; 제 4,301,144 호; 제 4,670,417 호; 제 4,791,192 호 또는 제 4,179,337 호에 나타낸 방식으로, 여러 비단백질성 중합체 중 하나, 예컨대 폴리에틸렌 글리콜, 폴리프로필렌 글리콜, 또는 폴리옥시알킬렌에 연결될 수 있다. 단백질은 예를 들어 중합체로의 공유 콘쥬게이션에 의해 화학적으로 개질되어, 이들의 순환 반감기가 증가된다. 중합체, 및 이들을 펩티드에 부착시키는 방법이 또한 U. S. 특허 제 4,766,106 호; 제 4,179,337 호; 제 4,495,285 호; 및 제 4,609,546 호에 나타나 있다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 PEG 화될 수 있다. PEG 화는 체내에서 단백질의 반감기를 연장시키기 위해 폴리에틸렌 글리콜 (PEG) 을 단백질에 부착시키는 공정이다. 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 PEG 화는 GDF-8 의 최적 저해를 위해 필요한 단백질의 투여 용량 또는 빈도를 감소시킬 수 있다. 상기 기법의 개론은 [Bhadra 등 (2002) Pharmazie, 57: 5-29; 및 Harris 등 (2001) Clin. Pharmacokinet., 40: 539-551] 에 제공된다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 IgG 분자의 Fc 영역에 연결될 수 있다. 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 IgG 분자의 Fc 영역 근처에 융합되거나, 링커 펩티드를 통해 IgG 분자의 Fc 영역에 부착될 수 있다. 상기 링커 펩티드의 이용은 단백질 생화학 분야에 널리 공지되어 있다. Fc 영역은, 예를 들어 IgG1 또는 IgG4 유래일 수 있다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 변형된 글리코실화 패턴을 갖도록 (즉, 본래 또는 천연 글리코실화 패턴이 변형되도록) 개질될 수 있다. 본원에서 사용되는 '변형된' 이란, 하나 이상의 탄수화물 부분이 결실되고/되거나 하나 이상의 글리코실화 부위가 본래 단백질에 첨가된 것을 의미한다.

단백질의 글리코실화은 전형적으로 N-연결 또는 0-연결된다. N-연결이란 아스파라긴 잔기 측쇄로의 탄수화물 부분의 부착을 나타낸다. 트리펩티드 서열 아스파라긴-X-세린 및 아스파라긴-X-트레오닌 (여기서, X 는 프롤린을 제외한 임의의 아미노산이다) 이 아스파라긴 측쇄로의 탄수화물 부분의 효소적 부착을 위한 인식 서열이다. 따라서, 폴리펩티드 내 어느 하나의 상기 트리펩티드 서열의 존재는 잠재적 글리코실화 부위를 생성한다. 0-연결 글리코실화란 당 N-아세틸갈락토사민, 갈락토오스, 또는 자일로오스 중 하나의 히드록시아미노산, 가장 일반적으로는 세린 또는 트레오닌으로의 부착을 나타내지만, 5-히드록시프롤린 또는 5-히드록시라이신도 또한 사용될 수 있다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질로의 글리코실화 부위의 첨가는 아미노산 서열이 하나 이상의 상술된 트리펩티드 서열을 포함하도록 변형시켜 편리하게 수행된다 (N-연결 글리코실화 부위를 위해). 변형은 또한 본래 단백질의 서열에 하나 이상의 세린 또는 트레오닌 잔기를 부가하거나 치환시켜 수행될 수 있다 (0-연결 글리코실화 부위를 위해). 용이하게는, 단백질 아미노산 서열을 DNA 수준에서의 변화를 통해 변형시킬 수 있다.

단백질 상에서 탄수화물 부분의 수를 증가시키는 또다른 방법은 단백질의 아미노산 잔기에 글리코시드를 화학적 또는 효소적으로 커플링하는 것이다. 상기 절차는 N- 또는 0-연결 글리코실화를 위한 글리코실화 가능성을 갖는 숙주 세포에서 GDF 펩티드 저해제의 제조를 필요로 하지 않는다는 점에서 유리하다. 이용되는 커플링 방식에 따라, 당은 (a) 아르기닌 및 히스티딘, (b) 자유 카르복실기, (c) 예컨대 시스테인의, 자유 술프히드릴기, (d) 예컨대 세린, 트레오닌, 또는 히드록시프롤린의, 자유 히드록실기, (e) 예컨대 페닐알라닌, 티로신 또는 트립토판 의, 방향족 잔기, 또는 (f) 글루타민의 아미드기에 부착될 수 있다. 상기 방법은 WO 87/05330, 및 [Aplin and Wriston (1981) CRC Crit. Rev. Biochem., 22: 259-306] 에 기재되어 있다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질에 존재하는 임의 탄수화물 부분의 제거는 화학적 또는 효소적으로 수행될 수 있다. 화학적 탈글리코실화에는 트리플루오로메탄술폰산, 또는 균등한 화합물에 단백질을 노출시키는 것이 요구된다. 상기 처리는 아미노산 서열은 온전하게 남겨두면서, 연결 당 (N-아세틸글루코사민 또는 N-아세틸갈락토사민) 을 제외한 대부분의 또는 전부의 당을 절단시킨다.

화학적 탈글리코실화는 [Hakimuddin 등 (1987) Arch. Biochem. Biophys., 259: 52; 및 Edge 등 (1981) Anal. Biochem., 118: 131] 에 기재되어 있다. GDF 펩티드 저해제 상의 탄수화물 부분의 효소적 절단은 [Thotakura 등 (1987) Meth. Enzymol., 138: 350] 에 기재된 바와 같은 다양한 엔도- 및 엑소-글리코시다아제를 사용하여 수행될 수 있다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 단백질 알부민 또는 알부민 유도체에 연결될 수 있다. 단백질 및 폴리펩티드를 알부민 또는 알부민 유도체로 연결시키는 방법은 당분야에 널리 공지되어 있다. 예를 들어, U. S. 특허 제 5,116,944 호를 참조한다.

약학 조성물

본 발명은 하나 이상의 폴리스타틴 도메인을 포함하는 단백질을 함유하는 조 성물을 제공한다. 상기 조성물은 환자로의 투여 및 약학적 사용에 적합할 수 있다. 조성물은 전형적으로 하나 이상의 폴리스타틴 도메인을 포함하는 하나 이상의 단백질 및 약학적으로 허용가능한 부형제를 함유한다. 본원에서 사용되는 어구 '약학적으로 허용가능한 부형제'에는 약학 투여에 상용성인 임의의 모든 용매, 분산 매질, 코팅제, 항균 및 항진균제, 등장화제 및 흡수 지연제 등이 포함된다. 약학 활성 물질에 대한 상기 매질 및 제제의 사용은 당분야에 널리 공지되어 있다. 조성물은 또한 보충의, 부가적, 또는 증강된 치료적 기능을 제공하는 다른 활성 화합물을 함유할 수 있다. 약학 조성물은 또한 투여용 지시서와 함께 용기, 팩 또는 분배기 내에 포함될 수 있다.

본 발명의 약학 조성물은 그 목적하는 투여 경로와 상용성이 되도록 제형화된다. 투여를 위한 방법은 당업자에게 공지되어 있다. 예를 들어 정맥내, 근육내, 또는 피하 투여될 수 있다.

피하 적용을 위해 사용되는 용액 또는 현탁액에는 전형적으로 하나 이상의 하기 성분이 포함된다: 멸균 희석제, 예컨대 주사용수, 식염수 용액, 고정유, 폴리에틸렌 글리콜, 글리세린, 프로필렌 글리콜 또는 다른 합성 용매; 항균제, 예컨대 벤질 알콜 또는 메틸 파라벤; 항산화제, 예컨대 아스코르브산 또는 나트륨 비술파이트; 킬레이트화제, 예컨대 에틸렌디아민테트라 아세트산; 완충액, 예컨대 아세테이트, 시트레이트 또는 포스페이트; 및 삼투성 조절용 제제, 예컨대 염화나트륨 또는 덱스트로오스. pH 는 산 또는 염기, 예컨대 염산 또는 수산화나트륨으로 조절될 수 있다. 상기 제제는 유리 또는 플라스틱으로 제조되는 앰플, 일회용 주 사기 또는 다회 용량 바이알에 밀봉될 수 있다.

주사용으로 적합한 약학 조성물에는 멸균 수용액 또는 분산액, 및 멸균 주사가능 용액 또는 분산액의 즉석 제조용 멸균 분말이 함유된다. 정맥내 투여를 위해 적합한 담체에는 생리 식염수, 정균수, Cremophor EL ™ (BASF, Parsippany, NJ) 또는 인산염 완충 식염수 (PBS) 가 포함된다. 모든 경우에 있어서, 조성물은 멸균되어야 하며, 주사가 용이하도록 하는 정도로 유체여야 한다. 제조 및 보관 조건 하에서 안정해야 하며, 미생물, 예컨대 세균 및 진균의 오염 작용에 대해 보호되어야 한다. 담체는, 예를 들어 물, 에탄올, 폴리올 (예를 들어 글리세롤, 프로필렌 글리콜, 및 액체 폴리에틸렌 글리콜 등), 및 이들의 적합한 혼합물을 포함하는 용매 또는 분산 매질일 수 있다. 적절한 유동성은, 예를 들어 코팅제, 예컨대 레시틴의 사용에 의해, 분산의 경우 목적하는 입자 크기의 유지에 의해, 및 계면활성제의 사용에 의해 유지될 수 있다. 미생물의 작용 예방은 다양한 항균 및 항진균제, 예를 들어 파라벤, 클로로부탄올, 페놀, 아스코르브산, 티메로살 등으로 얻어질 수 있다. 많은 경우, 조성물 중에 등장화제, 예를 들어 당, 폴리알콜, 예컨대 만니톨, 소르비톨, 염화나트륨이 함유될 수 있다. 주사용 조성물의 흡수 연장은 조성물 중에 흡수를 지연시키는 제제, 예를 들어 알루미늄 모노스테아레이트 및 젤라틴을 함유시켜 얻어질 수 있다.

하나의 구현예에 있어서, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 화합물을 신체로부터의 신속한 제거에 대해 보호할 담체와 함께, 예컨대 임플란 트 및 마이크로캡슐화 전달계를 포함하는 조절 방출 제형으로 제조된다. 생분해성, 생체적합성 중합체, 예컨대 에틸렌 비닐 아세테이트, 폴리무수물, 폴리글리콜산, 콜라겐, 폴리오르토에스테르 및 폴리락트산이 사용될 수 있다. 상기 제형의 제조 방법은 당업자에게 자명할 것이다. 상기 재료는 또한 Alza Corporation 및 Nova Pharmaceuticals, Ind. 에서 시판되는 것으로 입수할 수 있다. 하나 이상의 폴리스타틴 도메인을 포함하는 단백질을 함유하는 리포좀 현탁액이 또한 약학적으로 허용가능한 담체로 사용될 수 있다. 이들은 당업자에게 공지된, 예를 들어 U. S. 특허 제 4,522,811 호에 기재된 방법에 따라 제조될 수 있다.

치료적으로 유용한 제제, 예컨대 성장 호르몬 (예컨대, BMP, TGF-β, FGF, IGF), 시토카인 (예컨대, 인터류킨 및 CDF), 항생제 및 치료받는 상태에 대해 유익한 임의의 다른 치료제가 임의로, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질 중에 함유되거나, 이와 동시에 또는 순차적으로 투여될 수 있다.

투여량의 균일성 및 투여의 용이성을 위해, 투여 단위형으로 조성물을 제형화하는 것이 특히 유리하다. 본원에서 사용되는 투여 단위형이란 치료될 대상체를 위한 단위 투여량으로 맞춰진 물리적 개별 단위를 나타내며; 각각의 단위에는 요구되는 약학 담체와 연합되어 목적하는 치료 효과를 나타내도록 계산된 소정량의 활성 화합물이 포함된다. 본 발명의 투여 단위형에 대한 명시는 활성 화합물의 독특한 특징 및 얻어질 특정 치료 효과, 및 개인들의 치료를 위한 활성 화합물 합성의 당분야에 존재하는 한계에 의해 결정되며, 이에 직접 의존한다.

치료 적응증

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 인간 또는 동물에서 다양한 의학적 장애를 예방, 진단 또는 치료하는데 유용하다. 따라서, 본 발명은, 질환의 증상을 완화시키는데 충분한 양의, 하나 이상의 폴리스타틴 도메인을 포함하는 하나 이상의 단백질을 함유하는 조성물을 대상체에 투여하여, 근육 세포 및 조직에 관련된 질환 및 장애를 치료하는 방법을 제공한다. 상기 장애에는 중증 또는 양성 X-연관 근이영양증, 사지연관 이영양증, 안견갑상완골 이영양증, 강직성 이영양증, 원위 근이영양증, 진행성 위축성 안근마비, 안인두 이영양증, 듀센씨 근이영양증, 및 파쿠야마 유형 선천성 근이영양증을 포함하지만, 이에 제한되지는 않는 근이영양증; 근위축성 측삭 경화증 (ALS); 근위축; 기관 위축; 쇠약; 손목굴 증후군; 선천성 폐색성 폐질환; 선천성 근육병증; 선천성 강직증; 가족성 주기마비; 발작성 미오글로빈뇨; 중증근무력증; 이튼-램버트 증후군; 속발성 근무력증; 탈신경 위축; 격발성 근위축; 및 사르코페니아, 악액질 및 다른 근육 소모 증후군이 포함된다. 본 발명은 또한 근조직의 외상성 또는 만성 손상에 관한 것이다.

근육 질환 및 장애를 위한 치료법을 제공하는데 덧붙여, 본 발명은 또한 비정상적 글루코오스 항상성에서 야기되는 대사성 질환 또는 장애를 예방 또는 치료하는 방법을 제공한다. 상기 질환 또는 장애에는 대사성 질환 및 장애 (예컨대, 인슐린 의존성 (제 1 형) 당뇨병, 인슐린 비의존성 (제 2 형) 당뇨병), 고혈당증, 손상된 글루코오스 관용, 대사성 증후군 (예컨대, 증후군 X), 외상 (예컨대, 화상 또는 질소 불균형) 에 의해 유도되는 인슐린 저항성 및 비만, 지방세포 조직 장애 (예컨대 비만), 또는 골 퇴행성 질환 (예컨대 골다공증, 특히 노년성 및/또는 폐경후 여성에서; 글루코코르티코이드 유도 골다공증; 골감소증; 골관절염; 및 골다공증 관련 골절) 이 포함된다. 또다른 예에는 만성 글루코코르티코이드 치료법으로 인한 낮은 골 질량, 조발성 생식기 부전, 안드로겐 억제, 비타민 D 결핍, 속발성 부갑상선항진증, 영양 결핍 및 신경성 식욕부진이 포함된다.

정상적인 글루코오스 항상성에는 혈중 글루코오스 수준의 미묘한 변화에 반응하는 췌장 베타 세포에 의한 인슐린 분비의 미세하게 튜닝된 조정이 필요하다. 인슐린의 기본 작용 중 하나는 혈액으로부터 조직, 특히 근육 및 지방으로의 글루코오스 섭취를 자극하는 것이다.

따라서, 본 발명은 질환 증상을 완화시키기 충분한 양의, 하나 이상의 폴리스타틴 도메인을 포함하는 하나 이상의 단백질을 함유하는 조성물을 대상체에게 투여하여, 당뇨병 및 관련 장애, 예컨대 비만 또는 고혈당증을 치료하는 방법을 제공한다. 특히 제 2 형 또는 인슐린 비의존성 당뇨병 (NIDDM) 은 하기 3 가지를 특징으로 한다: (1) 말초 조직, 특히 골격근 및 지방세포에서 글루코오스 섭취에 대한 인슐린 작용에의 저항성, (2) 간의 글루코오스 생산을 저해하는 손상된 인슐린 작용, 및 (3) 탈조절된 인슐린 분비 (DeFronzo (1997) Diabetes Rev. 5: 177-269). 따라서, 제 2 형 당뇨병을 앓는 대상체는 인슐린에 대한 감수성 및 세포에 의한 글루코오스 섭취를 증가시키는, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질을 투여하여 본 발명에 따라 치료될 수 있다.

유사하게, 인슐린 기능장애 (예컨대 저항성, 불활성 또는 결핍) 및/또는 세포 내로의 불충분한 글루코오스 수송을 특징으로 하는 다른 질환 및 대사성 장애가 또한 인슐린에 대한 감수성 및 세포에 의한 글루코오스 섭취를 증가시키는, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질을 투여하여 본 발명에 따라 치료될 수 있다.

단백질을 사용하는 치료 방법

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 상기 단백질에 결합되지 않은 GDF-8 단백질에 비해, GDF-8 단백질 (단량체형, 이량체형 또는 GDF-8 잠재 복합체에 착화된 형태와 무관) 과 연합된 하나 이상의 활성을 저해하거나 감소시키는데 사용될 수 있다. 하나의 구현예에 있어서, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질에 결합된 성숙 GDF-8 단백질의 활성은 폴리스타틴 도메인을 갖는 단백질에 결합되지 않은 성숙 GDF-8 단백질에 비해 50% 이상, 또는 60, 62, 64, 66, 68, 70, 72, 72, 76, 78, 80, 82, 84, 86, 또는 88% 이상, 또는 90, 91, 92, 93, 또는 94% 이상, 또는 95% 이상 내지 100% 저해된다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질을 함유하는 약학 제제는 치료적 유효량으로 투여된다. 본원에서 사용되는 단백질의 '유효량' 은 목적하는 생물학적 결과를 얻기 위해 GDF-8 의 활성을 감소시키는데 충분한 투여량이다. 목적하는 생물학적 결과는 근육 질량의 증가, 근육 강도의 증가, 개선된 대사, 감소된 비만증, 또는 개선된 글루코오스 항상성을 포함하는 임의의 치료적 이득일 수 있다. 상기 개선은 제지방 (lean) 및 지방 체중 (예컨대 이중 x-선 주사 분석), 근육 강도, 혈청 지질, 혈청 렙틴, 혈청 글루코오스, 글리케이션된 헤모글로빈, 글루코오스 관용, 및 당뇨병의 이차 합병증에서의 개선을 측정하는 방법을 포함하는 다양한 방법으로 측정될 수 있다.

일반적으로, 치료적 유효량은 대상체에서 의학적 상태의 중증도 뿐만 아니라 대상체의 연령, 상태 및 성별에 따라 변할 수 있다. 투여량은 의사에 의해 결정될 수 있고, 필요하다면 치료의 목적 효과에 따라 조절될 수 있다. 하나 이상의 폴리스타틴 도메인을 포함하는 하나 이상의 단백질을 투여하기 위해 적절한 투여량은 5 mg 내지 100 mg, 15 mg 내지 85 mg, 30 mg 내지 70 mg, 또는 40 mg 내지 60 mg 범위일 수 있다. 단백질은 단회 용량으로, 또는 간격마다, 예컨대 1 일 1 회, 1 주 1 회 및 1 개월 1 회 투여될 수 있다. 투여 스케줄은 GDF-8 에 대한 단백질의 친화도, 단백질의

반감기 또는 환자 상태의 중증도에 따라 조절될 수 있다. 일반적으로, 조성물은 투여 후 최대한의 시간 길이 동안 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 순환 수준을 최대화시키기 위해, 볼루스 용량으로 투여된다. 연속 점적이 또한 볼루스 용량 후에 이용될 수 있다.

상기 화합물의 독성 및 치료 유효성은 세포 배양 또는 실험 동물에서의 표준 약학 절차, 예컨대 LD_{50} (집단의 50% 에 치사적인 용량) 및 ED_{50} (집단의 50% 에 치료적으로 유효한 용량) 을 결정하기 위한 절차로 결정될 수 있다. 독성 및 치료 효과 사이의 용량비가 치료 지수이며, 비 LD_{50}/ED_{50} 으로 나타낼 수 있다. 큰 치료 지수를 나타내는, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질이 사용 될 수 있다.

세포 배양 분석 및 동물 연구에서 수득한 데이타를 인간에서 사용하기 위한 투여량 범위 평가에 이용할 수 있다. 상기 화합물의 투여량은 독성이 거의 또는 전혀 없이 ED_{50} 을 포함하는 순환 농도의 범위 내에 속할 수 있다. 투여량은 사용되는 투여형 및 이용되는 투여 경로에 따라 상기 범위 내에서 변할 수 있다. 본 발명에서 사용되는 하나 이상의 폴리스타틴 도메인을 포함하는 임의의 단백질에 있어서, 치료적 유효 용량은 세포 배양 분석에서 먼저 추정될 수 있다. 용량은 동물 모델에서, 세포 배양에서 결정된 바와 같은 IC_{50} (즉, 증상의 최대 저해의 절반을 일으키는 시험 단백질의 농도) 을 포함하는 순환 혈장 농도 범위를 달성하도록 제형화될 수 있다. 혈장 내 수준은, 예를 들어 고성능 액체 크로마토그래피로 측정될 수 있다. 임의의 특정 투여량의 효과는 적합한 생분석으로 모니터링될 수 있다. 적합한 생분석의 예에는 GDF-8 단백질/수용체 결합 분석, 크레아틴 키나아제 분석, 지방세포에서의 글루코오스 섭취에 근거한 분석 및 면역학적 분석이 포함된다.

DNA 투여 방법

본 발명은 또한 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 생체 내 제조를 위한 유전자 치료법을 제공한다. 상기 치료법은 폴리뉴클레오티드 서열을 본원에 열거된 바와 같은 장애를 갖는 세포 또는 조직 내로 도입하여 그 치료 효과를 얻을 것이다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 폴리뉴클레오티드 서열의 전달은 재조합 발현 벡터, 예컨대 키메릭 바이러스 또는 콜로이드성 분산계를 이용하여 얻어질 수 있다. 표적 리포좀이 폴리뉴클레오티드 서열의 치료적 전달을 위해 이용될 수 있다. 유전자 치료법에 이용될 수 있는 다양한 바이러스성 벡터에는 아데노바이러스, 헤르페스 바이러스, 박시니아, 또는 RNA 바이러스, 예컨대 레트로바이러스가 포함된다. 레트로바이러스성 벡터는 설치류 또는 조류 레트로바이러스의 유도체일 수 있다. 단일 외부 유전자가 삽입될 수 있는 레트로바이러스성 벡터의 예에는 하기가 포함되지만, 이에 제한되는 것은 아니다: 몰로니 설치류 백혈병 바이러스 (MoMuLV), 하베이 설치류 육종 바이러스 (HaMuSV), 설치류 유선암 바이러스 (MuMTV), 및 라우스 육종 바이러스 (RSV). 여러 부가적 레트로바이러스성 벡터에는 여러 유전자가 도입될 수 있다. 모든 상기 벡터에는 형질도입된 세포가 동정 및 생성될 수 있도록 하는 선택 마커에 대한 유전자가 전달되거나 도입될 수 있다. 예를 들어, 특정 표적 세포 상의 수용체에 대한 리간드를 코딩하는 또다른 유전자와 함께, 관심 GDF 프로펩티드 폴리뉴클레오티드 서열을 바이러스성 벡터 내로 삽입하므로써, 벡터가 이제 표적 특이적이 된다.

레트로바이러스성 벡터는, 예를 들어 당, 당지질, 또는 단백질을 부착하여 표적 특이적으로 만들 수 있다. 표적화는 항체를 이용하여 수행될 수 있다. 당업자는 특정 폴리뉴클레오티드 서열이 레트로바이러스성 게놈 내로 삽입되거나 바이러스 외피에 부착되어, 하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 폴리뉴클레오티드를 포함하는 레트로바이러스성 벡터가 표적 특이적으로 전달될 수 있음을 인지할 것이다. 하나의 구현예에 있어서, 벡터는 근육 세포 또는 근육 조직으로 표적화된다.

재조합 레트로바이러스는 결손되어 있으므로, 이들은 LTR 내의 조절 서열의 조절 하에서 레트로바이러스의 전체 구조 유전자를 코딩하는 플라스미드를 포함하는 헬퍼 세포주를 필요로 한다. 상기 플라스미드에는 캡슐화를 위한 RNA 전사물을 인식하는 패키징 기전을 가능케하는 뉴클레오티드 서열이 없다. 패키징 신호가 결실된 헬퍼 세포주에는, 예를 들어 PSI.2, PA317 및 PA12 가 포함되지만, 이에 제한되는 것은 아니다. 게놈이 팩키징되지 못하므로, 상기 세포주는 빈 바이러스 입자 (virion) 를 만든다. 패키징 신호는 손상되지 않았지만 구조 유전자는 다른 관심 유전자로 대체된 상기 세포 내로 레트로바이러스성 벡터가 도입되는 경우, 벡터가 패키징될 수 있고, 벡터 바이러스 입자가 제조된다.

이와는 달리, 다른 조직 배양 세포는 종래 인산칼슘 전달감염에 의해, 레트로바이러스성 구조 유전자인 gag, pol 및 env 을 코딩하는 플라스미드로 직접 전달감염될 수 있다. 이어서, 상기 세포는 관심 유전자를 포함하는 벡터 플라스미드로 전달감염된다. 생성 세포는 레트로바이러스성 벡터를 배양 배지 내로 방출한다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질의 폴리뉴클레오티드를 위한 또다른 표적화 전달계는 콜로이드성 분산계이다. 콜로이드성 분산계에는 거대분자 복합체, 나노캡슐, 마이크로스피어, 비드 및 수중유 에멀젼, 마이셀, 혼합 마이셀, 및 리포좀을 포함하는 지질 기재 시스템이 포함된다. 리포좀은 시험관 내 및 생체 내 전달 비히클로서 유

용한 인공 막 소포이다. RNA, DNA 및 온전한 바이러스 입자는 수성 내부 내에서 캡슐화되어 생물학적 활성형으로 세포에 전달될 수 있다 (예를 들어, Fraley 등 (1981) Trends Biochem. Sci., 6: 77 참조). 리포좀 비히클을 이용한 효율적 유전자 전달 방법은 당분야에 공지되어 있다 (예를 들어, Mannino 등 (1988) Biotechniques, 6: 682 참조). 보통, 리포좀 조성물은 통상 스테로이드, 특히 콜레스테롤과 조합된 인지질의 조합물이다. 다른 인지질 또는 다른 지질도 사용될 수 있다. 리포좀의 물리적 특징은 pH, 이온 강도 및 2 가 양이온의 존재에 의존한다.

리포좀 제조에 유용한 지질의 예에는 포스파티딜 화합물, 예컨대 포스파티딜글리세롤, 포스파티딜콜린 포스파티딜세린, 포스파티딜에탄올아민, 스핑고지질, 세레브로시드 및 갱글리오시드가 포함된다. 인지질의 예에는 난 (egg) 포스파티딜콜린 디팔미토일포스파티딜콜린 및 디스테아로일포스파티딜콜린이 포함된다. 리포좀 표적화는 또한, 예를 들어 기관 특이성, 세포 특이성 및 소기관 특이성에 근거할 수 있으며, 당분야에 공지되어 있다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질을 발현하는 세포를 GDF-8 반응의 조절에 사용하기 위한 부위로 전달하는데 사용될 수 있는 광범위한 방법이 존재한다. 본 발명의 하나의 구현예에 있어서, 폴리스타틴 단백질을 발현하는 세포는 직접 적용에 의해, 예를 들어 상기 세포 표본을 조직 손상 부위 내로 직접 주사하여 전달될 수 있다. 상기 세포는 정제될 수 있다. 상기 세포는 세포를 손상 부위에 국소화시키기 위해, 이들의 운동성을 부분적으로 저해시키는 배지 또는 매트릭스 중에 전달될 수 있다. 상기 배지 또는 매트릭스는 반고체, 예컨대 겔 유사 중합체를 포함하는 겔 또는 페이스트일 수 있다. 이와는 달리, 배지 또는 매트릭스는 세포의 고체 매트릭스 내로의 이동을 허용하고, 이들을 거기에서 유지시키면서 세포 증식을 허용할 고형의 다공성 고체 형태일 수 있다.

GDF-8 의 검출 및 단리 방법

하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 생체 내 또는 시험관 내 GDF-8 의 수준 또는 존재를 검출하기 위해 이용될 수 있다. 의학적 상태와 상기 단백질의 수준 또는 존재를 연관시켜, 당업자는 관련 의학 상태를 진단할 수 있다. 하나 이상의 폴리스타틴 도메인을 포함하는 단백질로 진단될 수 있는 의학 상태를 본원에 나타낸다.

상기 검출 방법은 당분야에 널리 공지되어 있으며, ELISA, 방사면역분석, 면역블롯, 웨스턴 블롯, 면역형광, 면역침전 및 기타 대등 기법이 포함된다. 하나 이상의 폴리스타틴 도메인을 포함하는 단백질은 또한 GDF-8 을 검출하기 위한 하나 이상의 상기 기법을 도입시킨 진단 키트에서 제공될 수 있다. 상기 키트에는 단백질의 검출 및 키트의 이용을 보조하는 다른 성분, 팩키징, 설명서 또는 다른 물질이 포함될 수 있다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질이 진단 목적을 위한 것인 경우, 이들을, 예를 들어 리간드기 (예컨대 비오틴) 또는 검출가능한 마커기 (예컨대 형광기, 방사선동위원소 또는 효소) 로 개질시키는 것이 바람직할 수 있다. 필요하다면, 단백질을 종래 기법을 이용하여 표지할 수 있다. 적합한 표지에는 형광단, 발색단, 방사활성 원자, 다전자 시약, 효소 및 특정 결합 파트너를 갖는 리간드가 포함된다. 효소는 전형적으로 이들의 활성에 의해 검출된다. 예를 들어, 홀스래디쉬 페록시다아제는 통상 분광광도계로 정량가능한, 3,3',5,5'-테트라메틸벤지딘 (TMB) 을 청색 안료로 전환시키는 그 능력에 의해 검출된다. 다른 적합한 결합 파트너에는 비오틴 및 아비딘 또는 스트렙타비민, IgG 및 단백질 A, 및 당분야에 공지된 여러 수용체-리간드 커플이 포함된다. 다른 변경 및 가능성이 당업자에게 쉽게 자명할 것이며, 본 발명의 범위 내에서 대등하게 간주된다.

하나 이상의 폴리스타틴 도메인을 포함하는 단백질 또는 이들의 절편은 또한 정제 공정에서 GDF-8 의 단리를 위해 유용할 수 있다. 한 유형의 공정에 있어서, 단백질은 예를 들어 칼럼 또는 수지 내로의 도입을 통해 고정화될 수 있다. 단백질은 GDF-8 에 결합한 후, 결합 GDF-8 을 방출시키는 조건에 놓여 사용된다. 상기 공정이 GDF-8 의 상업적 제조를 위해 이용될 수 있다.

하기 실시예는 본 발명의 구현예를 제공한다. 당업자는 본 발명의 요지 또는 범위를 변경시키지 않고 수행될 수 있는 여러 개질 및 변화를 인지할 것이다. 상기 개질 및 변화는 본 발명의 범위 내에 포함되는 것으로 여겨진다. 실시예는 어떠한 방식으로든 본 발명을 제한하지 않는다. 명세서 및 청구범위에서의 모든 수는, 예를 들어 본 발명의 범위 내인 것으로 간주될 투여량의 작은 변화에서와 같이 '약' 이라는 용어로 변형되는 것으로 이해된다.

실시예

실시예 1: JA16 이 콘쥬게이션된 비드의 정제

N-히드록시숙신이미딜-활성화 비드 (4% 비드화 아가로오스, Sigma H-8635, St Louis MO) 를 MilliQ-$H_2$0 중에 세척하고, 최종 농도 10 mg JA16/수지 ml 이 될 수 있게 하는 비로 항-GDF-8 JA16 모노클로날 항체 (100 mM MOPS, pH 7.5 중 3-4 ㎍/㎕) 와 함께 4℃ 에서 4 시간 동안 인큐베이션하였다. 비드를 100 mM MOPS pH 7.5 및 인산염 완충 식염수 (PBS) (Ausubel 등 (1999) Current Protocols in Molecular Biology, John Wiley amp; Sons) 로

광범위하게 세척하고, 사용 전까지 PBS 중 4℃ 에서 보관하였다. 대조군 비드를 JA16 항체 없이 동일하게 제조하였다.

실시예 2: 친화도 정제

총 40 ㎕ 의 패킹된, JA16 이 콘쥬게이션된 비드 또는 대조군 비드를 15 ml 의 일반 Balb/C 마우스 혈청 (Golden West Biologicals, Temecula CA) 또는 30 ml 의 풀링된 일반 인간 혈청 (ICN Biomedical, Aurora OH) 과 4℃ 에서 3 시간 동안 인큐베이션하였다. 비드를 ~10 ml 의 냉 1 % Triton X-100/PBS 중에 2 회, ~10 ml 의 냉 0.1 % Triton X-100/PBS 중에 2 회, 및 ~1 ml 의 냉 PBS 중에 2 회 세척하였다. 단백질을 3 개의 후속 단계에서 비드로부터 용출하였다. 첫번째로, 비드를 'mock 용출' 로 처리하였다 (여기서, 100 ㎕ 의 PBS 를 비드에 첨가하고, 30 분 동안 4℃ 에서 인큐베이션하였다. 상청액을 수집하여, 30 ㎕ 의 4 ×LDS 표본 완충액 (Invitrogen, Carlsbad CA) 과 합하였다). 두번째로, 비드를 '펩티드 용출' 처리하였다 (여기서, 100 ㎕ 의 PBS 중 1 ㎍/㎕ 의 경합 펩티드 (서열: DFGLDSDEHSTESRSSRYPLTVDFEAFGWD-COOH (서열 번호: 12)) 를 비드에 첨가하고, 30 분 동안 4℃ 에서 다시 인큐베이션하였다. 상청액을 앞에서와 같이 수집하였다). 세번째로, 비드를 'SDS 용출' 기법으로 처리하였다 (여기서, 30 ㎕ 의 4 ×LDS 완충액 (Invitrogen) 및 100 ㎕ 의 PBS 를 비드에 첨가하고, 10 분 동안 80℃ 로 가열한 후, 상청액을 깨끗한 튜브로 옮겼다).

각각의 용출 단계에서 방출된 단백질의 은으로 염색된 겔을 도 1 에 나타내었다. 도 1 에 나타낸 은으로 염색된 겔 중 대략 12 및 36 kDa 의 2 개 단백질 밴드는 콘쥬게이션되지 않은 대조군 비드가 아니라 JA16 이 콘쥬게이션된 비드로부터 특이적으로 용출되었다.

실시예 3: 질량 분광측정

표본을 80℃ 에서 10 분 동안 NuPage 10 ×환원제 (Invitrogen) 로 환원시키고, 암실에서 22℃ 에서 30 분 동안 110 μM 요오도아세트아미드로 알킬화시켰다. 표본으로 제조자의 지시에 따라 즉시 MES 완충계 중의 10% NuPage 비스-Tris 겔 (Invitrogen) 을 걸고, 글루타르알데히드가 없는 시스템을 이용하여 은으로 염색하였다 (Shevchenko 등 (1996) Anal. Chem., 68: 850-858). 밴드를 절단하고 Abimed Digest Pro (Langenfeld, Germany) 또는 ProGest Investigator (Genomics Solutions, Ann Arbor MI) 중에서 서열분석 등급의 개질된 트립신 (Promega, Madison WI) 으로 소화시켰다. 소화된 표본의 부피를 증발로 감소시키고, 1% 아세트산을 보충하여 최종 부피 ~20 ㎕ 로 만들었다. 표본 (5-10 ㎕) 을 Picofrit 바늘 (New Objectives, Woburn MA) 에 패킹된 10 cm ×75 ㎛ 내부 직경 의 C_{18} 역상 칼럼 상에 로딩하였다. MS/MS 데이타를 LCQ Deca 또는 LCQ Deca XP (Finnigan, San Jose CA) 질량 분광광도계를 이용하여 수집하고, Sequest 프로그램 (Finnigan) 을 이용하여 NCBI 비중복 데이타베이스에 대해 검색하였다. 달리 언급되지 않는 한, 본 문헌에 열거된 모든 펩티드 서열은 수작업에 의해 고품질로 간주되고, Sequest 스코어링 시스템에서 X_{corr} 스코어 > 2.5 를 얻은 MS/MS 스펙트럼에 해당된다.

실시예 4: 웨스턴 블롯

단백질을 0.45 ㎛ 니트로셀룰로오스막 (Invitrogen) 에 옮기고, 하룻밤동안 4℃ 에서 블로킹 완충액 (Tris 완충 식염수 (TBS: 10 mM Tris-Cl, pH 7.5, 150 mM NaCl) 중 5% 탈지 분유) 으로 블로킹하였다. 이어서, 블롯을 실온에서 1-3 시간 동안 블로킹 완충액 중 1:1000 으로 희석된 일차 항체를 붙이고 (probed), 5 ×TBS 로 세척하고, 실온에서 1-3 시간 동안 블로킹 완충액 중 홀스래디쉬 페록시다아제가 콘쥬게이션된 이차 항체를 붙이고, 앞에서와 같이 세척하였다. West Pico Substrate (Pierce) 를 이용하여 방사활성측정에 의해 신호를 검출하였다.

실시예 5: GDF-8 의 단리

선행 실시예에 기재된 방법을 이용한 실험으로 GDF-8 을 단리하였다. 그 환원형인 GDF-8 이 12 kDa 이므로, 도 1 에 나타낸 은으로 염색된 겔에서 더 낮은 위치 밴드에서의 단백질을 성숙 GDF-8 로 예측하였다. 상기 가설을 확인하기 위해, 상기 밴드를 절단하고, 이를 트립신으로 소화시켜, LC-MS/MS 에 의해 생성 펩티드의 MS/MS 스펙트럼을 수득하였다. 트립신으로 분해된 6 개의 펩티드에 해당하는 MS/MS 스펙트럼은, 도 2A 및 표 1 에 나타낸 바와 같이, 겔의 상기 영역으로부터 성숙 GDF-8 이 단리되었음을 확인시켰다.

표 1 에는 마우스 및 인간 혈청으로부터의 JA16 면역침전물에서 확인된 GDF-8 (서열 번호: 13-20), GDF-8 프로펩티드 (서열 번호: 21-27), FLRG (서열 번호: 28-30), 및 GASP1 (서열 번호: 31-35) 에서 유래된 펩티드가 열거되었다. 단백질 서열 중 바로 직전의 아미노산은 각각의 펩티드에 대해 괄호 안에 나타내었고, 펩티드의 하전 상태 (z) 및 Sequest 프로그램 연관 계수 (X_{corr}, 신뢰성 측정치) 를 열거하였다. 표에서 서열 기재 번호는 단지 단리된 펩티드 및 이들의 서열 만을 나타낸다. 괄호 내 선행 아미노산은 펩티드에 포함되지 않으며, 단지 참고를 위해 제공된다. 모든 스펙트럼을 수작업으로 확인하였다.

흥미롭게도, 웨스턴 블롯에 또한 비가공 전장 GDF-8 에 해당하는 밴드 (43 kDa) 가 포함되었으며, 이는 상기 분자의 일부가 단백분해 가공을 거치지 않고 혈청 내로 분비됨을 시사하였다 (도 2B). 미가공 GDF-8 의 존재는 우리측 질량 분광측정에서 확인되었다 (데이타는 나타내지 않음). 따라서, 친화도 정제 방법으로 일반 마우스 혈청으로부터 GDF-8 이 효과적으로 단리되었다.

JA16 항체가 GDF-8 및 상당히 관련된 단백질 BMP/GDF-11 을 둘 다 인지하지만, 질량 분광측정에 의해 우리측 친화도 정제 표본에서 BMP-11 펩티드의 증거는 보지 못했다.

[표 1]

JA16 면역침전물에서 동정된 펩티드

	마우스 혈청	z	X_{corr}
GDF-8 (성숙)	(K) ANYCSGECEFVFLQK (서열 번호 :13)	3+	4.63
	(K) MSPINMLYFNGK (서열 번호 :14)	2+	3.81
	(R) DFGLDCDEHSTESR (SEQ ID NO:15)	2+	3.47
	(K) ANYCSGECEFVFLQK (서열 번호 :16)	2+	3.31
	(K) M*SPINMLYFNGK (서열 번호 :17)	3+	2.95
	(R) YPLTVDFEAFGWDWIIAPK (서열 번호 :18)	2+	2.86
	(K) M*SPINM*LYFNGK (서열 번호 :19)	2+	2.51
	(R) GSAGPCCTPTK (서열 번호:20)	2+	2.43
GDF-8 (프로펩티드)	(K) LDM*SPGTGIWQSIDVK (서열 번호:21)	2+	3.82
	(K) ALDENGHDLAVTFPGPGEDGLNPFLEVK (서열 번호 : 22)	3+	3.17
	(K) LDMSPGTGIWQSIDVK (서열 번호 :23)	2+	2.98
	(R) ELIDQYDVQR (서열 번호 :24)	2+	2.97
	(K) TPTTVFVQILR (서열 번호 :25)	2+	2.91
	(K) AQLWIYLRPVK (서열 번호 :26)	2+	2.77
	(K) EGLCNACAWR (서열 번호 :27)	2+	2.75
폴리스타틴 -유사 관련 유전자 (FLRG)	(R) PQSCLVDQTGSAHCVVCR (서열 번호 :28)	3+	3.34
	(K) DSCDGVECGPGK (서열 번호 :29)	2+	2.99
	(K) SCAQVVCPR (서열 번호 :30)	2+	2.59
신규 멀티도메인	(R) ECETDQECETYEK (서열 번호 :31)	2+	2.98
	(R) ADFPLSVVR (서열 번호 :32)	2+	2.56
프로테아제 저해제 (GASP1)	(R) EACEESCPFPR (서열 번호 :33)	2+	2.95
	(R) SDFVILGR (서열 번호 :34)	2+	2.73
	(R) VSELTEEQDSGR (서열 번호 :35)	2+	3.88

M* = 산화된 메티오닌

	인간 혈청	z	X_{corr}
GDF-8 (성숙)	(K) ANYCSGECEFVFLQK (서열 번호 :36)	2+	4.21
	(R) DFGLDCDEHSTESR (서열 번호 :37)	3+	2.08
GDF-8 (프로펩티드)	(K) ALDENGHDLAVTFPGPGEDGLNPFLEVK (서열 번호 :38)	3+	3.71
	(R) ELIDQYDVQR (서열 번호 :39)	2+	3.01
폴리스타틴-유사 관련 유전자 (FLRG)	(R) PQSCVVDQTGSAHCVVCR (서열 번호 :40)	3+	3.37
	(R) CECAPDCSGLPAR (서열 번호 :41)	2+	3.21
	(R) LQVCGSDGATYR (서열 번호 :42)	2+	3.06
멀티도메인 프로테아제 저해제 (GASP1)	(R) VSELTEEPDSGR (서열 번호 :43)	2+	2.44
	(R) CYMDAEACSK (서열 번호 :44)	2+	2.69
	(K) GITLAVVTCR (서열 번호 :45)	2+	2.42

실시예 6: GDF-8 에 결합된 단백질의 단리

일단 친화도 정제 기법이 일반 마우스 혈청으로부터 GDF-8 을 성공적으로 단리할 수 있다는 것을 확인한 후, 자연 조건 하에 GDF-8 과 결합하는 단백질의 동정을 진행하였다. 도 1 에 나타낸 은으로 염색된 겔 상의 36 kDa 밴드를 상술된 바와 같이 분석하였다. 질량 분광측정으로, JA16 으로 면역침전된 표본에 특이적인 겔의 상기 영역에서 두 단백질을 동정하였다. 이들을 GDF-8 프로펩티드 및 폴리스타틴-유사 관련 유전자 (follistatin-like related gene, FLRG) 로 결정하였다. 상기 단백질 각각에서 동정된 펩티드를 표 1 에 나타내었다 (서열 번호: 13-27). 고품질 MS/MS

스펙트럼에서, GDF-8 프로펩티드로부터 6 개의 독특한 펩티드, 및 FLRG 로부터 3 개의 독특한 펩티드가 발견되었고; 대표적 펩티드를 도 3A 및 3C 에 나타내었다. 또한, 상기 단백질 모두의 존재가 각각 GDF-8 프로펩티드 및 FLRG 에 특이적인 폴리클로날 항체로의 웨스턴 블롯으로 확인되었다 (도 3B 및 3D). 즉, 순환 GDF-8 은 생체 내에서 GDF-8 프로펩티드 및 FLRG 에 결합하는 것으로 나타났다.

실시예 7: GDF-8 에 결합하는 신규 단백질의 단리

생체 내 순환 GDF-8 복합체의 주성분을 분석하기 위해, 야생형 마우스 혈청으로부터 천연 GDF-8 및 그 연합 단백질을 아가로오스가 콘쥬게이션된 항 GDF-8 모노클로날 항체 JA16 으로의 친화도 정제로 단리하였다. JA16 에 결합된 단백질에 PBS 완충액 단독 (mock 용출), JA16 결합을 위해 GDF-8 과 경합할 수 있는 펩티드, 및 SDS 세제의 후속 용출 단계를 수행하였다. 상기 표본을 농축하고, 1 차 SDS-PAGE 겔을 걸어, 은 염색으로 시각화하였다 (도 4). JA16 정제 표본에 독 특한 2 개의 밴드가 보였다 - GDF-8 로 동정된 12 kDa 밴드, 및 GDF-8 프로펩티드 및 FLRG 를 둘 다 포함하는 36 kDa 밴드.

생체 내에서 GDF-8 에 결합된 다른 단백질을 동정할 수 있는지를 결정하기 위해, 정제를 대략 5 배 스케일링 업하고, JA16 면역복합체에는 존재하지만 음성 대조군에는 존재하지 않는 단백질을 검색하기 위해 질량 분광측정을 이용하였다. 상기 목적을 달성하기 위해, 도 4 에 나타낸 바와 같이 10 내지 200 kDa 사이의 분자량에 해당하는 은으로 염색된 겔 영역을 13 개 겔 조각으로 절단하였다. 상기 조각 각각을 겔 내 트립신 소화 및 LC-MS/MS 로 처리하였따. 공지된 단백질의 비중복 NCBI 데이타베이스에 대한 생성 MS/MS 스펙트럼의 비교는 JA16 면역침전물에 특이적인 임의의 부가적 단백질은 나타내지 않았으나, 앞에서 기술된 단백질 (성숙 GDF-8, GDF-8 프로펩티드, 미가공 GDF-8, 및 FLRG) 은 모두 상기 표본에서 동정되었다 (도 4). JA16 면역복합체 및 음성 대조군 표본 둘 다에서 발견된 백그라운드 단백질에는 풍부한 혈청 단백질, 예컨대 알부민, 면역글로불린 및 보체 단백질이 포함되었다. JA16 표본에는 상당히 관련된 단백질 BMP-11/GDF-11 을 포함하는 다른 TGF-β수퍼패밀리 멤버의 증거는 없었다. 즉, JA16 항체는 본 실험에서 GDF-8 을 특이적으로 정제하였다.

흥미롭게도, JA16 이 시험관 내에서 GDF-8/폴리스타틴 복합체를 면역침전시킬 수 있음에도 불구하고, 우리의 GDF-8 면역복합체에서 폴리스타틴의 증거는 발견하지 못하였다 (데이타는 나타내지 않음). 폴리스타틴은 ActRIIB 수용체와 GDF-8 의 연합을 길항하여 GDF-8 활성을 억제하는 것으로 나타나 있다 (Lee and McPherron (2001) Proc. Natl. Acad. Sci. U.S.A., 98: 9306-9311). 우리의 결과는, 정상 조건 하에서 폴리스타틴이 순환 GDF-8 복합체의 활성 조절에 주요 역할을 담당하지 않음을 시사한다.

상기 MS/MS 절차에 의한 단백질의 동정이 검색된 데이타베이스의 컨텐츠에 의존하므로, Celera 마우스 게놈 서열에서 예측된 단백질의 데이타베이스에 대해 13 개 표본에서 수집된 MS/MS 스펙트럼을 비교하여, 도 4 로부터의 데이타를 추가 분석하였다. 상기 분석으로 JA16 으로 정제된 표본에 특이적인 부가적 단백질이 동정되었고, 이를 GDF-연합 혈청 단백질 1 (GASP1) 로 명명하였다. 상기 단백질이 최초 동정되었으므로, 상기 서열을 접근 번호 gi|20914039 으로 공개 게놈 서열분석 시도에 의한 NCBI nr 데이타베이스에 부가하였다.

GASP1 서열에 해당하는 5 개 펩티드를 고품질 MS/MS 스펙트럼에 기준하여 동정하였다 (표 1 (서열 번호: 31-35); 도 5A 및 B). GASP1 펩티드에 해당하는 스펙트럼이 밴드 3 에서 발견되었고, 70-80 kDa 단백질이 포함되었다. 그러나, 상기 단백질에 해당하는 특이적 밴드는, 아마도 상기 영역에서 백그라운드 면역글로불린 및 알부민의 풍부함으로 인해 보이지 않았다 (도 4 참조). 2.3 을 초과하는 Sequest X_{corr} 스코어는 일반적으로 2^{+} 이온에 대해 유의미한 것으로 간주된다. 우연하게도, 실험에서 동정된 펩티드 중 하나 (서열 = ECETDQECETYEK (서열 번호: 31)) 가 상기 단백질을 코딩하는 2 개 엑손 사이 연결부 (junction) 에 위치하여, 상기 경우에 대한 Celera 사의 유전자 예측 알고리즘의 정확성을 확인시켰다.

GASP1 전사물 및 단백질의 서열을 GASP1 의 실제 클로닝 전에 예측하였다 (도 6). GASP1 은 예측 분자량 63 kDa 을 갖는 571 개 아미노산 단백질로 예측되었다. 이는 그 N-말단에 추정 시그날 서열/절단 부위를 가지며, 아미노산 314 및 514 에서 2 개의 N-글리코실화 가능 부위를 갖는다. Pfam 및 BLAST ([Altschul 등 (1990) J. Mol. Biol., 215: 403-410; Bateman 등 (2002) Nucleic Acids Res., 30: 276-280] 의 기법에 따름) 에 의한 GASP1 단백질 서열의 분석은 GASP1 이 WAP 도메인, 폴리스타틴/Kazal 도메인, 면역글로불린 도메인, 2 개의 연계 (tandem) Kunitz 도메인, 및 netrin 도메인을 포함하는 여러 보존 도메인을 포함한다는 것을 나타내었다 (도 14A). 처음에 유장 산성 단백질에서 동정된 WAP 도메인은 4 개의 디술피드 코어를 형성하는 8 개 시스테인을 포함하며, 종종 항프로테아제 활성을 갖는 단백질에서 발견된다 (Hennighausen and Sippel (1982) Nucleic Acids Res., 10: 2677-2684; Seemuller 등 (1986) FEBS Lett., 199: 43-48). 폴리스타틴 도메인이 GDF-8 및 GASP1 사이의 상호작용을 매개하는 것으로 여겨진다. 폴리스타틴 도메인의 C-말단 영역에는 Kazal 세린 프로테아제 저해제 도메인에 대한 유사성이 포함되어 있다. GASP1 의 경우에 있어서, 상기 영역은 폴리스타틴 또는 FLRG 에서보다 Kazal 도메인에 더욱 더 가까이 연관되어 있어서, 상기 영역이 부가적 프로테아제 저해제 기능을 가질 수 있음을 시사하였다. 처음에 소 췌장 트립신 저해제에서 동정된 Kunitz 도메인은 또한 세린 프로테아제를 저해하므로, 상기 단백질 패밀리의 조절에서 GASP1 의 역할이 성립될 수 있다. 또한, netrin 도메인은 메탈로프로테아제의 저해에 연관되어 있다 (Banyai and Patthy, 1999

; Mott 등 2000). 즉, 상기 보존 영역의 존재에 근거하여, GASP1 은 아마도 프로테아제의 활성을 저해하여, 잠재 GDF-8 복합체의 GDF-8 가공 또는 활성화를 조절할 수 있을 것이다.

마우스 Celera 전사물 데이타베이스에 대한 BLAST 검색에서, GASP1 와 > 50% 상동성을 갖는 단백질을 발견하여 GASP2 로 명명하였다. GASP2 에는 GASP1 에서와 동일한 도메인 구조가 포함되어, 상기 단백질이 다중 (multivalent) 프로테아제 저해제 패밀리의 2 개 멤버를 정의함을 시사하였다 (도 14B). 흥미롭게도, GASP2 가 아닌 GASP1 에 해당하는 펩티드만이 우리의 JA16 으로 정제된 표본에서 발견되었다. 상기 결과는 GASP1 및 GASP2 가 상이한 생물학적 특이성을 가질 수 있음을 시사한다. GASP1 및 GASP2 가 둘 다 인간에서 보존되어 있다 (마우스와 > 90% 상동성). 인간 GASP1 에 대한 서열은 이제 접근 번호 gi┃18652308 하에 NCBI nr 데이타베이스에서 이용가능하다. 인간 혈청 중 GDF-8 의 농도가 마우스 혈청에서 발견된 것보다 상당히 더 낮았지만 (Hill 등 (2002) J. Biol. Chem., 277: 40735-40741), 단백질의 질량 분광측정 분석의 감도로 인해 인간 혈청의 JA16 면역침전물로부터 GASP1 의 인간 상동체에 해당하는 3 펩티드를 동정할 수 있었다 (표 1). 상기 펩티드 중 어느 것도 대응 음성 대조군에서는 발견되지 않았다. 또한, 상기 실험에서 인간 GASP2 의 증거는 존재하지 않았다. 즉, GASP1 및 GDF-8 사이의 상호작용은 마우스 및 인간 사이에서 보존되어 있다.

GDF-8 은 골격근에서 거의 배타적으로 생성된다. GASP1 mRNA 의 조직 분포를 결정하기 위해, GASP1 의 551 bp 절편을 여러 마우스 조직 및 단계별 배아에 서 제조된 제 1 쇄 cDNA 로부터 증폭하였다 (도 10). 마우스 GASP1 절편은 제조자의 지시에 따라 Advantage cDNA PCR 키트 (Clontech) 를 이용하여 보정된 마우스 제 1 쇄 cDNA 패널 (Clontech, Palo Alto CA) 로부터 증폭하였다 (전방 프라이머: 5' TTGGCCACTGCCACCACAATCTCAACCACTT 3' (서열 번호: 46); 후방 프라이머: 5' TCTCAGCATGGCCATGCCGCCGTCGA 3' (서열 번호: 47)). GASP1 은 거의 광범위하게 발현되며, 특히 골격근 및 심장에서 고발현되는 것으로 나타났다. 유의미한 발현이 뇌, 폐 및 고환에서 또한 나타났다. 대조적으로, 간 및 신장은 상대적으로 저수준의 GASP1 mRNA 를 발현하였다. 발생학적으로, GASP1 mRNA 수준은 거의 일정하게 유지되었고, 아마도 마우스 배아발생 7 일 내지 11 일 사이에 약한 증가하였다.

실시예 8: 인간 및 마우스 혈청 중 GDF-8

인간 혈청 중 GDF-8 의 농도는 마우스 혈청 중에서 발견되는 것보다 상당히 더 낮다. GDF-8 이 치료적 표적으로서 잠재성을 가지므로, 인간에서 순환 GDF-8 복합체의 조성을 결정하는 것이 목표였다. 이에 관한 지식으로 마우스 모델의 타당성을 결정하고, 대안적 치료 표적을 동정할 수 있을 것이다. 즉, GDF-8 의 JA16-기재 친화도 정제를 인간 혈청을 사용하여 반복하였다. 마우스에 비해 인간 혈청 중에서의 더 낮은 수준의 GDF-8 로 인해, 성숙 GDF-8 및 GDF-8 프로펩티드/FLRG 에 해당하는 밴드는 보이지 않았다 (도 11A). 그러나, GDF-8 의 성숙 영역을 인지하는 폴리클로날 항체로의 웨스턴 블롯은 JA16 으로 정제된 표본 중 성숙 및 미가공 GDF-8 의 존재를 나타내었다 (도 11B).

성숙 GDF8 과 공동 정제되는 단백질을 동정하기 위해, 질량 분광측정의 높은 감도를 이용하였다. 음성 대조군 및 JA16 이 콘쥬게이션된 비드 둘 다로부터 펩티드 용출 표본에 해당하는 레인을 16 조각으로 절단하였다. 상기 겔 조각으로, 앞에서와 같이 겔 내 트립신 소화, 나노플로우 LC-MS/MS, 및 Sequest 를 이용한 분석을 수행하였다.

흥미롭게도, 음성 대조군이 아닌 JA16 표본에서만 특이적으로 동정된 단백질은 성숙 GDF-8, GDF-8 프로펩티드, 인간 FLRG, 및 GASP1 의 인간 상동체였다. 상기 단백질 각각에서 발견된 펩티드를 표 1 에 나타내었고 (서열 번호: 36-45), 대표적 MS/MS 스펙트럼을 도 12 에 나타내었다. 즉, 생체 내 GDF-8 복합체는 마우스 및 인간 사이에서 보존된 것으로 나타났다.

실시예 9: 마우스 GASP1 의 클로닝 및 특징분석

예측 GASP1 서열을 동정한 후, 마우스 GASP1 의 실제 서열을 결정하는 것이 목표였다. Celera 예측 서열에 근거하여, GASP1 코딩 서열을 하기 프라이머 (fp: 5' CACCATGTGTGCCCCAGGGTATCATCGGTTCTGG 3' (서열 번호: 50); rp: 5' TTGCAAGCCCAGGAAGTCCTTGAGGAC 3' (서열 번호: 51)) 를 이용하여 PfuTurbo 폴리머라아제 (Stratagene) 로의 PCR 에 의해 마우스 심장 QUICKCLONE cDNA (Clontech) 로부터 증폭하였다. 상기 반응으로부터의 PCR 산물을 1% 아가로오스 겔 상에서 대략 1700 염기쌍의 단일 주요 밴드로 나타났다. 이어서 증폭된 DNA 를 pcDNA3.1D/V5-His-TOPO 벡터 (Invitrogen) 의 TOPO 부위 내로 클로닝하여, 제조자의 지시에 따라 프레임 내에 C-말단 V5-His 태그를 포함시켰다. 전장 cDNA 삽 입물을 이중쇄 모두 서열분석하였다. 마우스 GASP1 클론의 뉴클레오티드 서열을 도 13 에 나타내었다. 상기 클론은 예측된 아미노산 서열을 변화시키지 않는 워블 코돈에서의 일부 염기 치환 (즉, 288C:G; 294G:A; 615G:A; 738A:G; 768C:T; 1407A:G; 1419A:G; 및 1584C:G (여기서, 나타낸 위치의 첫번째 염기는 Celera 에 의해 보고된 것이고, 두번째 염기는 클론의 서열분석에서 수득된 것이다); 도 6A 및 B 참조) 을 제외하고, 예측된 Celera 서열과 일치하였다.

GASP1 단백질의 N-말단 가공을 결정하기 위해, C-말단 V5-His 태그를 갖도록 클로닝된 마우스 GASP1 (GASP1-V5-His) 을 코딩하는 포유류 발현 벡터로 COS1 세포를 전달감염시켰다. 무혈청 조건화 배지를 48 시간 후에 얻고,

항 V5 폴리클로날 항체 (Sigma) 로 웨스턴 블롯 분석에 의해 분석하였다. 보다 구체적으로, 무혈청 Dulbecco's 개질 Eagle's 배지 중에서 FuGENE 6 시약 (Roche) 을 이용하여 GASP1-V5-His/pcDNA3.1D-V5-His-TOPO 또는 빈 벡터로 COS1 세포를 전달감염시킨 48 시간 후에 조건화 배지를 수집하였다.

대략 80 kDa 의 단일 밴드가 나타났으며, GASP1 이 조건화 배지 내로 분비됨이 확인되었다 (데이타는 나타내지 않음). 상기 조건화 배지의 대략 10 ml 을 His-친화도 칼럼에 흘리고, 역상 크로마토그래피로 추가 정제하였다. 상기 정제 방식으로, 쿠마시 (Coomassie) 염색된 SDS-PAGE 겔 상에서 추정 크기의 전장 GASP1 밴드를 얻었다. 상기 밴드의 Edman 서열분석으로 N-말단 서열 L-P-P-I-R-Y-S-H-A-G-I (서열 번호: 52) 이 결정되었다. 즉, GASP1 의 아미노산 1-29 는 가공 및 분비 도중 제거되는 시그날 서열을 형성하였다.

실시예 10: GDF-8 프로펩티드 및 성숙 GDF-8 에 대한 재조합적으로 생산된 GASP1 의 결합

다음에, 재조합적으로 생산된 GASP1 이 GDF-8 에 대해 마우스 혈청에서 단리된 GASP1 과 동일한 결합 패턴을 갖는지를 결정하였다. 재조합 단백질로의 면역침전을 위해, 벡터- 또는 GASP1-전달감염된 세포로부터의 400 μℓ 의 조건화 배지를 1.2 μg 의 재조합 정제 GDF-8 및/또는 GDF-8 프로펩티드 단백질 (Thies 등 2001) 과 합하였다. JA16 (10 μℓ 패킹 부피) 또는 항-V5 (30 μℓ) 가 콘쥬게이션된 아가로오스 비드를 4℃ 에서 2 시간 동안 보충된 조건화 배지와 인큐베이션하고, 인산염 완충 식염수 (PBS) 중 냉 1% Triton 중에 2 회 및 PBS 중에 2 회 세척하였다. 비드를 DTT 가 들어있는 50 μℓ 의 1 ×LDS 완충액 중에 재현탁하였다. 상술한 바와 같이 웨스턴 블롯을 수행하였다 (Hill 등 2002).

GDF-8 및 GASP1 사이의 상호작용을 확인하고 추가로 분석하기 위해, 정제된 재조합 GDF-8 및 정제된 재조합 GDF-8 프로펩티드와 벡터 대조군 또는 GASP1-V5-His 로 전달감염된 COS1 세포의 조건화 배지를 인큐베이션하였다. 이어서 GDF-8 을 JA16 이 콘쥬게이션된 아가로오스 비드로 면역침전시키고, 웨스턴 블롯을 이용하여 GASP1 및 GDF-8 프로펩티드의 공동 정제를 살펴보았다 (도 15A). GASP1 (레인 3) 및 GDF-8 프로펩티드 (레인 1) 둘 다 GDF-8 와 함께 면역침전되어, GDF-8 이 상기 단백질 모두와 상호작용할 수 있음을 증명하였다. GDF-8 의 부재 하에서는 GASP1 또는 프로펩티드 어느 것도 JA16 면역침전물에서 검출되지 않아 (레인 4), 상기 실험에서의 비특이적 결합 가능성이 배제되었다. 3 개 단백질이 모두 존재하는 경우, GASP1 및 GDF-8 프로펩티드 둘 다가 GDF-8 와 함께 풀링 다운되어, 상기 단백질이 3 원 복합체를 형성할 수 있는 가능성을 제시하였다 (레인 5). 그러나, 상기 실험은 GASP1 및 프로펩티드가 개별 GDF-8 분자 상의 동일 에피토프에 결합하는 가능성을 배제하지 않는다.

GASP1 및 GDF-8 사이의 상호작용을 추가 확인하기 위해, GDF-8 및/또는 GDF-8 프로펩티드 재조합 단백질이 보충된 조건화 배지로부터 GASP1 을 풀링 다운시켜 역 (reverse) 면역침전을 수행하였다. 이를 수행하기 위해, GASP1 상에서 C-말단 V5-His 태그의 V5 에피토프에 대해 제조된, 아가로오스가 콘쥬게이션된 모노클로날 항체를 사용하였다. 예상된 바와 같이, GDF-8 은 GASP1 과 함께 면역침전되어 (도 15B, 레인 3 및 5), 상기 단백질 사이의 직접적 상호작용이 더욱 확인되었다. 놀랍게도, GDF-8 프로펩티드 또한 GDF-8 의 부재 하에서도 GASP1 과 공동 정제되어 (레인 4), GDF-8 프로펩티드가 GASP1 에 직접 결합할 수 있음을 시사하였다. 즉, GASP1 은 GDF-8 및 GDF-8 프로펩티드 모두에 독립적으로 결합하였다. 이는 성숙 GDF-8 에 배타적으로 결합하는 또다른 폴리스타틴-도메인 단백질, FLRG 와는 대조적이다 (Hill 등 (2002) J. Biol. Chem., 277: 40735-40741). GDF-8 및 프로펩티드 둘 다의 첨가는 프로펩티드만이 첨가된 경우에 비해 GASP1 에 대해 더 적은 프로펩티드 결합을 일관되게 나타내었다. 상기 관찰은 GASP1 이 GDF-8 의 작은 잠재 복합체에 결합하지 않을 수 있음을 시사한다.

실시예 11: 액티빈 (Activin) 또는 TGF-β1 이 아닌 GDF-8 및 BMP-11 활성의 GASP1 매개 저해

루시퍼라아제 리포터 구축물 pGL3-(CAGA)$_{12}$ (서열 번호: 53) (Dennler 등 (1998) EMBO J., 17: 3091-3100) 을 A204 또는 RD 횡문근육종 세포 내로 일시적으로 전달감염시켰다. 벡터 또는 GASP1 이 전달감염된 세포로부터의 조건화 배지의 희석물을 10 ng/ml GDF-8, 10 ng/ml BMP-11, 10 ng/ml rh 액티빈 A (R amp; D Systems), 또는 0.5 ng/ml rh TGF-β1 (R amp; D Systems) 과 함께 37℃ 에서 30 분 동안 인큐베이션하였다. 루시퍼라아제 활성을 [Thies 등 (2001) Growth Factors, 18: 251-259 및 Zimmers 등 (2002) Science, 296: 1486-1488] 에 따라 측정하였다. 상기 분석에서, A204 세포는 GDF-8, BMP-11, 및 액티빈에 반응하였지만, TGF-β1 에는 잘 반응하지 않았다. RD 세포는 GDF-8 및 TGF-β1 둘 다에 반응하였다. 따라서, GDF-8, BMP-11, 및 액티빈에 대한 GASP1 의 저해능을 시험하기 위해서는 A204 세포를, TGF-β및 GDF-8 에 대한 활성을 모니터링하기 위해서는 RD 세포를 이용하였다. GDF-8 에 대한 결과는 A204 세포로 나타내었지만, RD 세포에서도 유사하였다. 각각의 상기 성장 인자에 의해 유도되는 루시퍼라아제 활성의 농도 의존성을 측정하는 표준 곡선을 각 실험에 대해 만들었다 (데이타는 나타내지 않음). 이용된 성장 인자 농도는, 농도에서의 작은 변화가 루시퍼라아제 활성에서의 측정가능한 변화를 일으키는 상기 곡선의 직선 영역 내에 속하였다.

두 개의 폴리스타틴-도메인 단백질인 폴리스타틴 및 FLRG 는 (CAGA)$_{12}$ (서열 번호: 53) 루시퍼라아제 전사 리포터 분석에서 GDF-8 활성을 저해하지만, 또한 관 련 단백질인 액티빈 및 BMP-11 의 생물학적 활성도 저해하였다. (

CAGA)$_{12}$ (서열 번호: 53) 리포터 분석에서 GDF-8, BMP-11, 액티빈, 및 TGF-β1 의 활성에 대한 GASP1 의 저해능을 또한 시험하였다.

V5-His 태그를 갖는 GASP1 또는 벡터 대조군으로 전달감염된 COS 세포로부터의 조건화 배지의 다양한 희석물을 정제된 재조합 GDF-8 (10 ng/ml), BMP-11 (10 ng/ml), 액티빈 (10 ng/ml), 또는 TGF-β1 (0.5 ng/ml) 과 인큐베이션하고, (CAGA)$_{12}$ (서열 번호: 53) 리포터 구축물을 발현하는 횡문근육종 세포에서의 성장 인자 활성에 대해 분석하였다. GASP1 은 농도 의존적 방식으로 GDF-8 활성을 강력히 저해하였다 (도 16A). 성숙 GDF-8 및 BMP-11 이 매우 보존되어 있고 단지 11 개 아미노산 만이 다르므로, 예측될 수 있는 바와 같이, GASP1 은 상기 분석에서 BMP-11 을 유사하게 저해하였다 (도 16B). 놀랍게도, GASP1 은 액티빈 또는 TGF-β1 의 활성은 저해하지 않아서 (도 16C 및 D), 폴리스타틴 자체에 의해 나타나지 않는 매우 높은 수준의 특이성을 시사하였다. 즉, GASP1 은 GDF-8 및 BMP-11 의 그 저해에 있어서 특이성을 나타내었다.

GDF-8 에 대한 GASP1 의 친화도를 리포터 유전자 분석에서 GDF-8 의 저해에 대한 IC50 을 결정하여 평가하였다. GASP1-V5-His 단백질을 코발트 친화도 칼럼 상에서 조건화 배지로부터 정제하고, 상술된 바와 같이 용출하였다. GASP1 을 포함하는 분획을 BioSepS3000 칼럼 (Phenomenex) 을 이용하여 PBS 중 크기 배제 크로마토그래피로 추가 정제하였다. 도 17 에 나타낸 바와 같이, GASP1 은 대 략 3 nM 의 IC50 을 가지며 GDF-8 을 저해하였다.

실시예 12: 근육 장애의 치료

GASP1 을 표 2 에 따라 GDF-8 의 기능에 관련된 질환 또는 장애를 앓는 환자에게 투여할 수 있다. 환자는 단회로 또는 간격마다, 예컨대 1 일 1 회 조성물을 섭취하였으며, 이들의 질환 또는 장애의 증상이 개선되었다. 예를 들어, 근육 질량, 근육 활성 및 또는 근육 긴장성으로 측정되는 바와 같은 근육 장애에 관련된 증상이 개선되었다. 이는 본 발명의 조성물이 GDF-8 기능과 관련된 질환 또는 장애, 예컨대 근육 장애의 치료를 위해 유용함을 나타낸다.

[표 2]

환자	질환	투여 경로	투여량	투여 빈도	예측 결과
1	근이영양증	피하	25 mg	1 일 1 회	근육 질량의 증가 및 근육 활성의 개선
2	〃	〃	50 mg	〃	〃
3	〃	〃	50 mg	1 주 1 회	〃
4	〃	〃	50 mg	1 개월 1 회	〃
5	〃	근육내	25 mg	1 일 1 회	〃
6		〃	50 mg	〃	
7	〃	〃	50 mg	1 주 1 회	〃
8	〃	〃	50 mg	1 개월 1 회	〃
9	〃	정맥내	25 mg	1 일 1 회	〃
10		〃	50 mg	〃	〃
11	〃	〃	50 mg	1 주 1 회	〃
12	〃		50 mg	1 개월 1 회	〃
13	당뇨병	피하	50 mg	1 일 1 회	혈당치 조절의 개선
14	〃	〃	50 mg	1 주 1 회	〃
15	〃	근육내	50 mg	〃	〃
16	〃	정맥내	50 mg	〃	〃
17	비만	피하	50 mg	1 일 1 회	체중 감소 및 근육 질량의 증가
18	〃	근육내	50 mg	1 주 1 회	〃
19		정맥내	50 mg	〃	

본 출원을 통해 언급된 모든 참고문헌, 특허 및 공개 특허 출원의 전체 내용 이 본원에 참고로 도입된다. 상기 상세한 설명은 단지 예시 목적으로 주어진다. 광범위한 변화 및 개질이 상술된 구현예에 수행될 수 있다. 따라서, 모든 대등물을 포함하는 하기 청구범위로 본 발명의 범위가 한정되는 것으로 이해되어야 한다.

(57) 청구의 범위

청구항 1.

하기를 함유하는 약학 조성물:

i) GASP1, 및

ii) 하나 이상의 약학적으로 허용가능한 담체.

청구항 2.

제 1 항에 있어서, GASP1 이 안정화 개질을 갖는 조성물.

청구항 3.

제 2 항에 있어서, 개질이 IgG 분자의 Fc 영역으로의 융합인 조성물.

청구항 4.

제 3 항에 있어서, IgG 분자가 IgG1 또는 IgG4, 또는 이들의 유도체인 조성물.

청구항 5.

제 4 항에 있어서, IgG 분자가 IgG1 또는 이들의 유도체인 조성물.

청구항 6.

제 3 항에 있어서, IgG 분자가 링커 펩티드에 의해 GASP1 에 융합되는 조성물.

청구항 7.

제 2 항에 있어서, 개질이 변형된 글리코실화 부위를 포함하는 조성물.

청구항 8.

제 2 항에 있어서, 개질이 하나 이상의 탄수화물 부분을 포함하는 조성물.

청구항 9.

제 2 항에 있어서, 개질이 알부민 또는 알부민 유도체를 포함하는 조성물.

청구항 10.

제 2 항에 있어서, 개질이 비단백질성 중합체를 포함하는 조성물.

청구항 11.

제 2 항에 있어서, 개질이 PEG 화를 포함하는 조성물.

청구항 12.

GASP1 및 하기로부터 선택되는 하나 이상의 기타 키트 성분을 포함하는 진단 키트:

i) GASP1 에 결합하는 하나 이상의 제제;

ii) 하나 이상의 완충액 및/또는 용액; 및

iii) 하나 이상의 구조적 성분.

청구항 13.

GASP1 을 코딩하는 핵산을 포함하는 재조합 세포.

청구항 14.
제 13 항에 있어서, GASP1 이 안정화 개질을 갖는 재조합 세포.

청구항 15.
GASP1 을 투여하고, GASP1 이 GDF-8 과 상호작용하도록 하는 것을 포함하는 GDF-8 의 조절 방법.

청구항 16.
치료적으로 유효한 용량의 GASP1 을 투여하고, GASP1 이 GDF-8 과 상호작용하도록 하는 것을 포함하는, 의학적 장애를 앓는 환자의 치료 방법.

청구항 17.
GASP1 을 코딩하는 핵산을 투여하고, 상기 핵산이 GASP1 로 번역되도록 하고, 번역된 GASP1 이 GDF-8 과 상호작용하도록 하는 것을 포함하는, 의학적 장애를 앓는 환자의 치료 방법.

청구항 18.
GASP1 을 코딩하는 핵산을 세포에 투여하고, 상기 핵산이 세포에 들어가도록 하고, 상기 세포가 GASP1 을 발현하도록 하는 것을 포함하는, GASP1 을 코딩하는 핵산의 발현 방법.

청구항 19.
제 16 항 내지 제 18 항 중 어느 한 항에 있어서, GASP1 이 안정화 개질을 갖는 방법.

청구항 20.
제 19 항에 있어서, 개질이 IgG 분자의 Fc 영역으로의 융합인 방법.

청구항 21.
제 20 항에 있어서, IgG 분자가 IgG1 또는 IgG4, 또는 이들의 유도체인 방법.

청구항 22.
제 21 항에 있어서, IgG 분자가 IgG1 또는 이들의 유도체인 방법.

청구항 23.
제 20 항에 있어서, IgG 분자가 링커 펩티드에 의해 GASP1 에 융합되는 방법.

청구항 24.
제 19 항에 있어서, 개질이 변형된 글리코실화 부위를 포함하는 방법.

청구항 25.
제 19 항에 있어서, 개질이 하나 이상의 탄수화물 부분을 포함하는 방법.

청구항 26.
제 19 항에 있어서, 개질이 알부민 또는 알부민 유도체를 포함하는 방법.

청구항 27.
제 19 항에 있어서, 개질이 비단백질성 중합체를 포함하는 방법.

청구항 28.
제 19 항에 있어서, 개질이 PEG 화를 포함하는 방법.

청구항 29.
제 16 항에 있어서, 근육 조직의 질량 또는 양의 증가가 환자에게 치료적으로 유익한 방법.

청구항 30.
제 16 항에 있어서, 장애가 근육 장애인 방법.

청구항 31.
제 30 항에 있어서, 근육 장애가 근이영양증인 방법.

청구항 32.

제 31 항에 있어서, 근이영양증이 중증 또는 양성 X-연관 근이영양증, 사지연관 이영양증, 안견갑상완골 이영양증, 강직성 이영양증, 원위 근이영양증, 진행성 위축성 안근마비, 안인두 이영양증, 듀센씨 (Duchenne's) 근이영양증 및 파쿠야마 (Fakuyama)-유형 선천성 근이영양증으로부터 선택되는 방법.

청구항 33.

제 30 항에 있어서, 장애가 근위축성 측삭 경화증, 선천성 폐색성 폐질환, 선천성 근육병증, 선천성 강직증, 가족성 주기마비, 발작성 미오글로빈뇨, 중증근무력증, 이튼-램버트 (Eaton-Lambert) 증후군, 속발성 근무력증, 탈신경 위축, 기관 위축, 쇠약, 손목굴 증후군, 근위축, 원위 근위축, 사르코페니아, 악액질 및 기타 근육 소모 증후군으로부터 선택되는 방법.

청구항 34.

제 30 항에 있어서, 장애가 근육 조직에 대한 외상성 손상 및 근육 조직에 대한 만성 손상으로부터 선택되는 근육 장애인 방법.

청구항 35.

제 16 항에 있어서, 장애가 대사 질환 또는 장애인 방법.

청구항 36.

제 35 항에 있어서, 장애가 인슐린 의존성 (제 1 형) 당뇨병, 인슐린 비의존성 (제 2 형) 당뇨병, 고혈당증, 손상된 글루코오스 관용, 대사성 증후군 (예를 들 어, 증후군 X), 외상 (화상 또는 질소 불균형) 으로 유도되는 인슐린 저항성 또는 비만인 방법.

청구항 37.

제 16 항에 있어서, 장애가 비만과 같은 지방세포 조직 장애인 방법.

청구항 38.

제 16 항에 있어서, 장애가 골다공증, 글루코코르티코이드 유도 골다공증, 골감소증, 골관절염, 또는 골다공증 관련 골절과 같은 골 퇴행성 질환, 또는 만성 글루코코르티코이드 치료법으로 인한 낮은 골 질량, 조발성 생식기 부전, 안드로겐 억제, 비타민 D 결핍, 속발성 부갑상선항진증, 영양 결핍증 및 거식증을 포함하는 기타 장애인 방법.

청구항 39.

제 16 항에 있어서, GASP1 이 1 회, 또는 매일, 매주, 또는 매개월 간격으로 투여되는 방법.

청구항 40.

제 16 항에 있어서, GASP1 이 5 mg 내지 100 mg 의 용량으로 투여되는 방법.

청구항 41.

제 16 항에 있어서, GASP1 이 15 mg 내지 85 mg 의 용량으로 투여되는 방법.

청구항 42.

제 16 항에 있어서, GASP1 이 30 mg 내지 70 mg 의 용량으로 투여되는 방법.

청구항 43.

제 16 항에 있어서, GASP1 이 40 mg 내지 60 mg 의 용량으로 투여되는 방법.

청구항 44.

제 1 항에 있어서, GASP1 이 하기로부터 선택되는 조성물:

i) 서열 번호 5;

ii) 서열 번호 7; 및

iii) i) 또는 ii) 의 치환, 부가 및/또는 결실 돌연변이.

청구항 45.

제 1 항에 있어서, GASP1 이 하기로부터 선택되는 뉴클레오티드 서열로 코딩되는 조성물:

i) 서열 번호 4;

ii) 서열 번호 6;

iii) i) 또는 ii) 에 의해 코딩되는 서열의 치환, 부가 및/또는 결실 돌연변이를 코딩하는 뉴클레오티드 서열; 및

iv) i) 또는 ii) 에 의해 코딩되는 것과 동일한 아미노산 서열을 코딩하는 뉴클레오티드 서열.

청구항 46.
하기를 함유하는 약학 조성물:

i) 하기로부터 선택되는 GASP-1 의 절편:

a) 폴리스타틴 도메인을 나타내는 서열 번호 5 의 아미노산 174-239;

b) 폴리스타틴 도메인을 나타내는 서열 번호 7 의 아미노산 110-175; 및

c) a) 또는 b) 의 치환, 부가 및/또는 결실 돌연변이; 및

ii) 하나 이상의 약학적으로 허용가능한 담체.

청구항 47.
제 46 항에 있어서, GASP1 절편이 하기로부터 선택되는 뉴클레오티드 서열로 코딩되는 조성물:

i) 폴리스타틴 도메인을 코딩하는 서열 번호 4 의 뉴클레오티드 520-717;

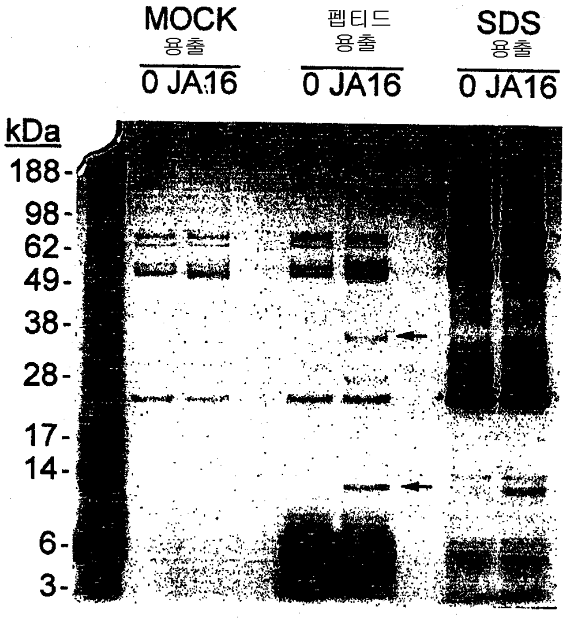
ii) 폴리스타틴 도메인을 코딩하는 서열 번호 6 의 뉴클레오티드 328-525;

iii) i) 또는 ii) 로 코딩되는 서열의 치환, 부가 및/또는 결실 돌연변이를 코딩하는 뉴클레오티드 서열; 및

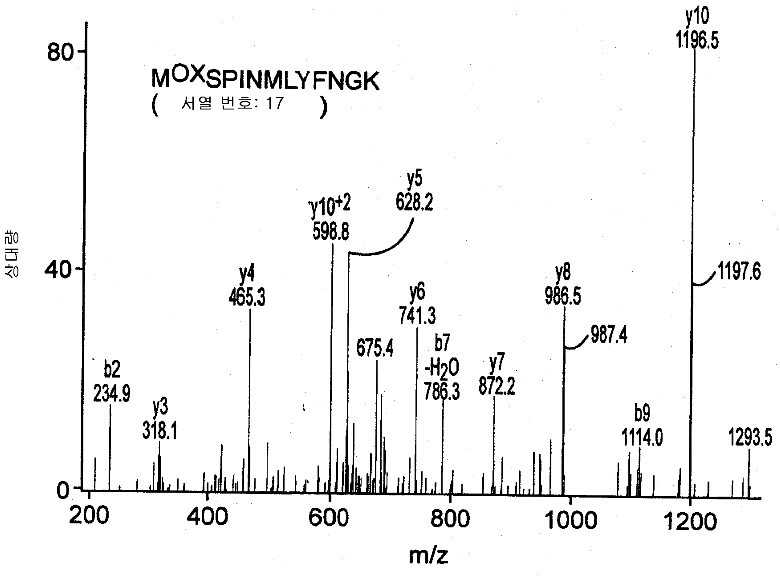
iv) i) 또는 ii) 에 의해 코딩되는 것과 동일한 아미노산 서열을 코딩하는 뉴클레오티드 서열.

도면

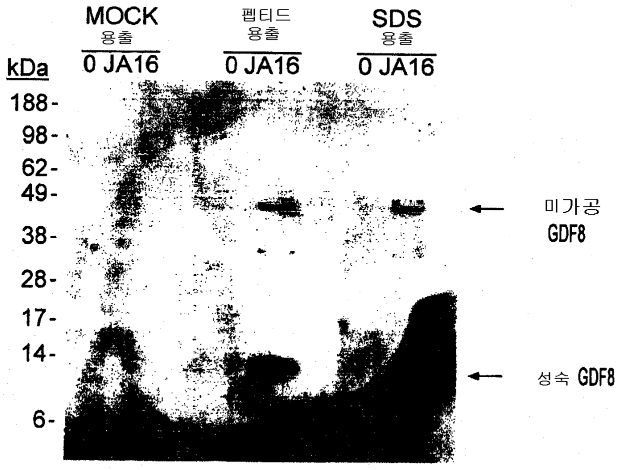
도면1



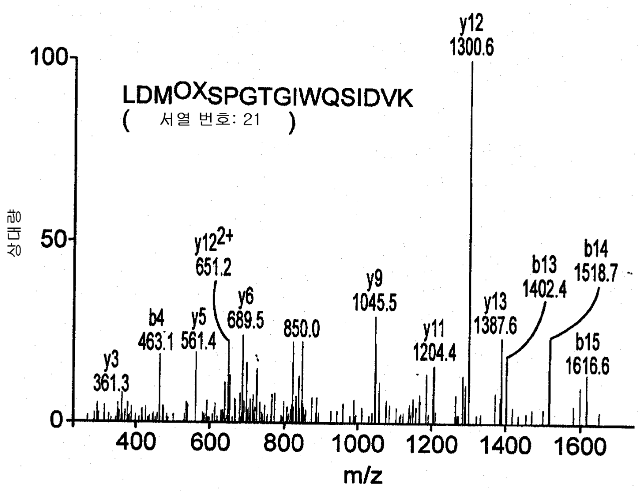
도면2A



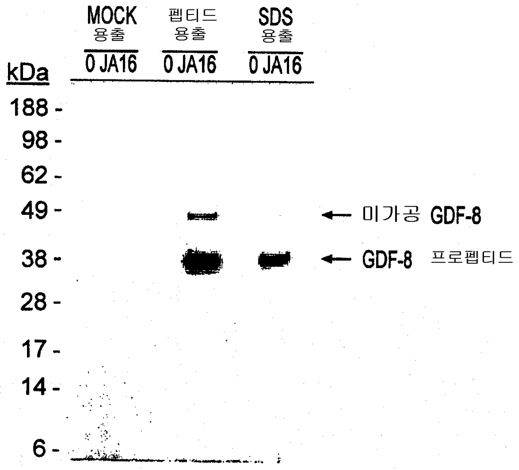
도면2B



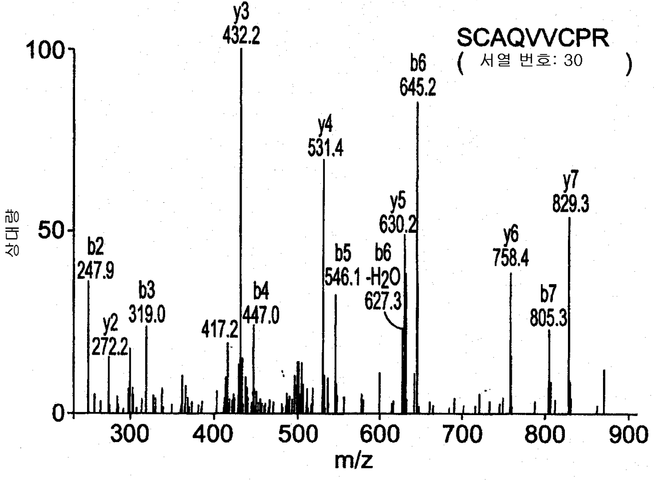
도면3A



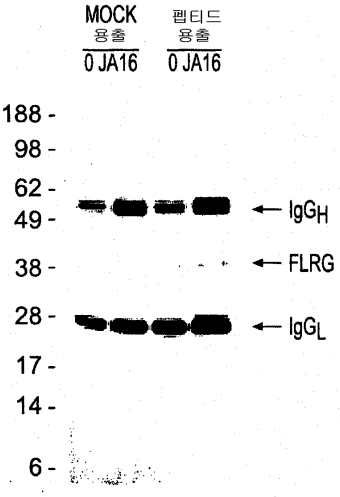
도면3B



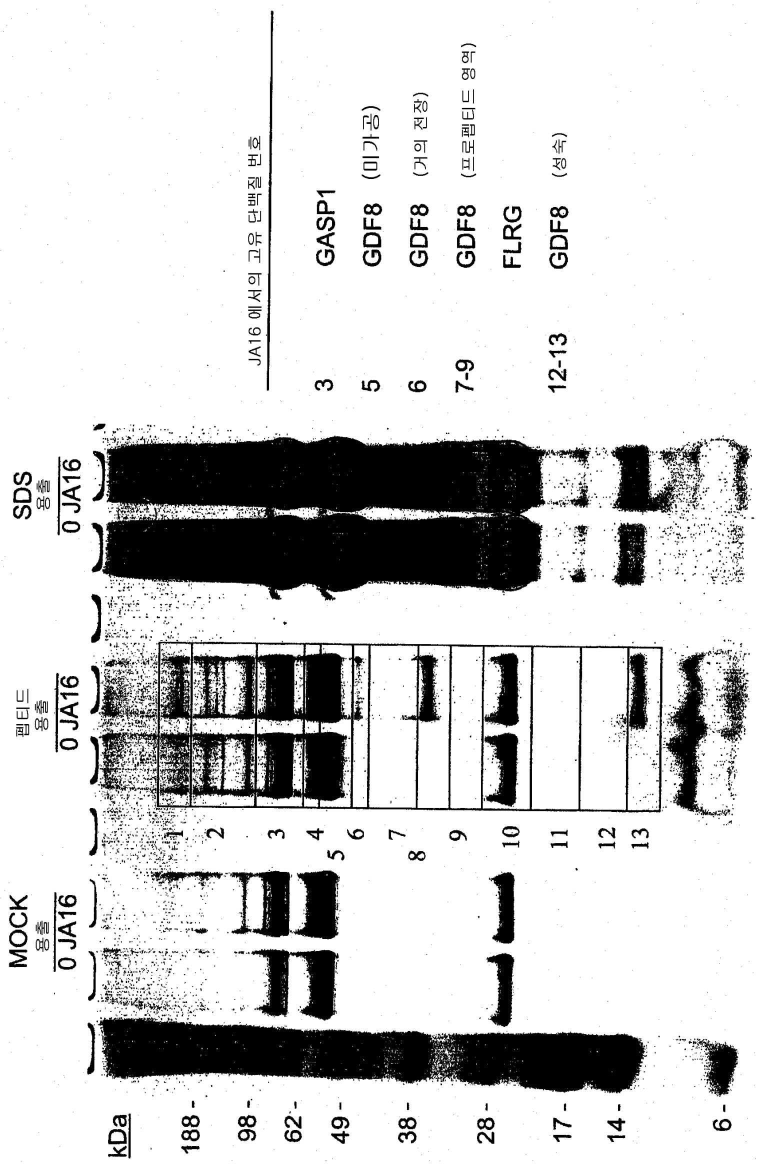
도면3C



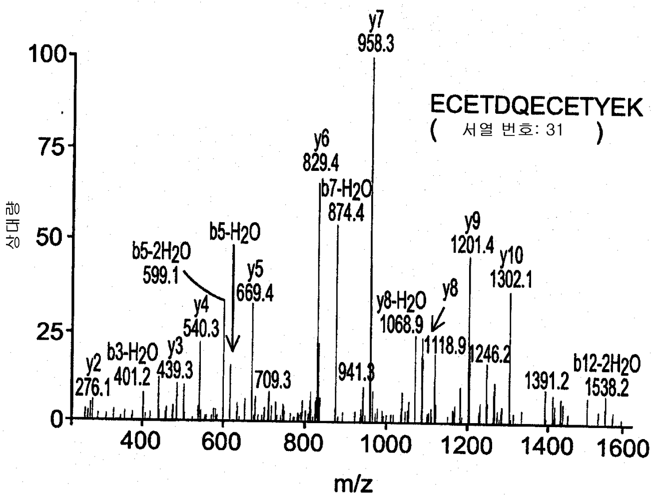
도면3D



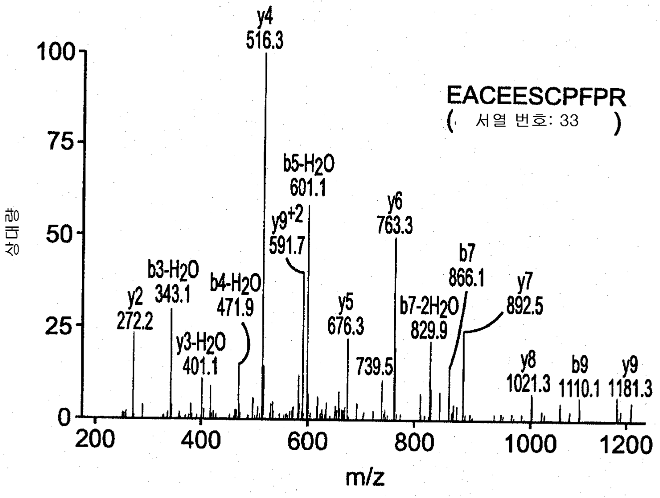
도면4


JA16 에서의 고유 단백질 번호
3 GASP1
5 GDF8 (미가공)
6 GDF8 (거의 전장)
7-9 GDF8 (프로펩티드 영역)
FLRG
12-13 GDF8 (성숙)
SDS 용출
0 JA16
펩티드 용출
0 JA16
MOCK 용출
0 JA16
1 2 3 4 5 6 7 8 9 10 11 12 13
kDa
188
98
62
49
38
28
17
14
6

도면5A


ECETDQECETYEK
(서열 번호: 31)
100
75
50
25
0
상대량
400
600
800
1000
1200
1400
1600
m/z
y7 958.3
y6 829.4
b7-H2O 874.4
b5-H2O
b5-2H2O 599.1
y5 669.4
y4 540.3
y3 439.3
b3-H2O 401.2
y2 276.1
709.3
941.3
y8-H2O 1068.9
y8
1118.9
y9 1201.4
y10 1302.1
1246.2
1391.2
b12-2H2O 1538.2

도면5B



도면6A

예측된 마우스 GASP1 뉴클레오티드 서열

서열 번호: 1

```
   1 atgtgtgccc cagggtatca tcggttctgg tttcactggg ggctgctgtt gctgctgctc
  61 ctcgaggctc cccttcgagg cctagcactg ccacccatcc gatactccca tgcgggcatc
 121 tgccccaacg acatgaaccc caacctctgg gtggatgccc agagcacctg caagcgagag
 181 tgtgaaacag accaggaatg tgagacctat gagaaatgct gccccaatgt gtgtgggacc
 241 aagagctgtg tggcagcccg ctacatggat gtgaaaggga agaagggccc tgtgggcatg
 301 cccaaggagg ccacatgtga ccatttcatg tgcctgcagc agggctctga gtgtgacatc
 361 tgggacggcc agcccgtgtg taagtgcaaa gatcgctgtg agaaggagcc cagcttcacc
 421 tgtgcctctg atggccttac ctactacaac cgttgcttca tggacgccga agcctgctcc
 481 aagggcatca cactgtctgt ggtcacctgt cgttatcact tcacctggcc taacaccagc
 541 cctccaccgc ctgagaccac ggtgcatccc accaccgcct ctccggagac tctcgggctg
 601 gacatggcag ccccggccct gctcaaccac cctgtccatc agtcagtcac cgtgggtgag
 661 actgtgagtt tcctctgtga cgtggtaggc cggcctcggc cagagctcac ttgggagaaa
 721 cagctggagg accgagaaaa tgttgtcatg aggcccaacc acgtgcgcgg taatgtggtg
 781 gtcactaaca ttgcccagct ggtcatctac aacgtccagc cccaggatgc tggcatatac
 841 acctgtacag ctcgaaatgt cgctggtgtc ctgagggctg acttcccgtt gtcggtggtc
 901 aggggtggtc aggccagggc cacttcagag agcagtctca atggcacagc ttttccagca
 961 acagagtgcc tgaagccccc agacagtgag gactgtggag aggagcagac acgctggcac
1021 ttcgacgccc aggctaacaa ctgcctcact ttcacctttg gccactgcca ccacaatctc
1081 aaccactttg agacctacga ggcctgtatg ctggcttgta tgagtgggcc attggccacc
1141 tgcagcctgc ctgccctgca agggccttgc aaagcttatg tcccacgctg ggcctacaac
1201 agccagacag gcctatgcca gtccttcgtc tatggcggct gtgagggcaa cggtaacaac
1261 tttgaaagcc gtgaggcttg tgaggagtcg tgtcccttcc cgaggggtaa ccagcactgc
1321 cgggcctgca agccccggca aaaacttgtt accagcttct gtcggagtga ctttgtcatc
1381 ctgggcaggg tctctgagct gaccgaagag caagactcag gccgtgccct ggtgaccgtg
1441 gatgaggtct taaaagatga gaagatgggc ctcaagtttc tgggccggga gcctctggaa
1501 gtcaccctgc ttcatgtaga ctggacctgt ccttgcccca acgtgacagt gggtgagaca
1561 ccactcatca tcatgggga ggtcgacggc ggcatggcca tgctgagacc cgatagcttt
1621 gtgggggcat cgagcacacg gcgggtcagg aagctccgtg aggtcatgta caagaaaacc
1681 tgtgacgtcc tcaaggactt cctgggcttg caatga
```

도면6B

예측된 마우스 GASP1 대안 뉴클레오티드 서열

서열 번호: 2

```
   1 atgtgtgccc cagggtatca tcggttctgg tttcactggg ggctgctgtt
  51 gctgctgctc ctcgaggctc cccttcgagg cctagcactg ccacccatcc
 101 gatactccca tgcgggcatc tgccccaacg acatgaaccc caacctctgg
 151 gtggatgccc agagcacctg caagcgagag tgtgaaacag accaggaatg
 201 tgagacctat gagaaatgct gccccaatgt gtgtgggacc aagagctgtg
 251 tggcagcccg ctacatggat gtgaaaggga agaaggggcc tgtaggcatg
 301 cccaaggagg ccacatgtga ccatttcatg tgcctgcagc agggctctga
 351 gtgtgacatc tgggacggcc agcccgtgtg taagtgcaaa gatcgctgtg
 401 agaaggagcc cagcttcacc tgtgcctctg atggccttac ctactacaac
 451 cgttgcttca tggacgccga agcctgctcc aagggcatca cactgtctgt
 501 ggtcacctgt cgttatcact tcacctggcc taacaccagc cctccaccgc
 551 ctgagaccac ggtgcatccc accaccgcct ctccggagac tctcgggctg
 601 gacatggcag ccccagccct gctcaaccac cctgtccatc agtcagtcac
 651 cgtgggtgag actgtgagtt tcctctgtga cgtggtaggc cggcctcggc
 701 cagagctcac ttgggagaaa cagctggagg accgagagaa tgttgtcatg
 751 aggcccaacc acgtgcgtgg taatgtggtg gtcactaaca ttgcccagct
 801 ggtcatctac aacgtccagc cccaggatgc tggcatatac acctgtacag
 851 ctcgaaatgt cgctggtgtc ctgagggctg acttcccgtt gtcggtggtc
 901 aggggtggtc aggccagggc cacttcagag agcagtctca atggcacagc
 951 ttttccagca acagagtgcc tgaagccccc agacagtgag gactgtggag
1001 aggagcagac acgctggcac ttcgacgccc aggctaacaa ctgcctcact
1051 ttcacctttg gccactgcca ccacaatctc aaccactttg agacctacga
1101 ggcctgtatg ctggcttgta tgagtgggcc attggccacc tgcagcctgc
1151 ctgccctgca agggccttgc aaagcttatg tcccacgctg ggcctacaac
1201 agccagacag gcctatgcca gtccttcgtc tatggcggct gtgagggcaa
1251 cggtaacaac tttgaaagcc gtgaggcttg tgaggagtcg tgtcccttcc
1301 cgaggggtaa ccagcactgc cgggcctgca agccccggca aaaacttgtt
1351 accagcttct gtcggagtga ctttgtcatc ctgggcaggg tctctgagct
1401 gaccgaggag caagactcgg gccgtgccct ggtgaccgtg gatgaggtct
1451 taaaagatga gaagatgggc ctcaagtttc tgggccggga gcctctggaa
1501 gtcaccctgc ttcatgtaga ctggacctgt ccttgcccca acgtgacagt
1551 gggtgagaca ccactcatca tcatgggggga ggtggacggc ggcatggcca
1601 tgctgagacc cgatagcttt gtggggggcat cgagcacacg gcgggtcagg
1651 aagctccgtg aggtcatgta caagaaaacc tgtgacgtcc tcaaggactt
1701 cctgggcttg caatga
```

도면6C

예측된 마우스 GASP1 아미노산 서열

서열 번호: 3

1 MCAPGYHRFW FHWGLLLLLL LEAPLRGLAL PPIRYSHAGI CPNDMNPNLW VDAQSTCKRE

61 CETDQECETY EKCCPNVCGT KSCVAARYMD VKGKKGPVGM PKEA**TCDHFM** **CLQQGSECDI**

121 **WDGQPVCKCK** **DRCEKEPSFT** **CASDGLTYYN** **RCFMDAEACS** **KGITLSVVTC** RYHFTWPNTS

181 PPPPETTVHP TTASPETLGL DMAAPALLNH PVHQSVTVGE TVSFLCDVVG RPRPELTWEK

241 QLEDRENVVM RPNHVRGNVV VTNIAQLVIY NVQPQDAGIY TCTARNVAGV LRADFPLSVV

301 RGGQARATSE SSLNGTAFPA TECLKPPDSE DCGEEQTRWH FDAQANNCLT FTFGHCHHNL

361 NHFETYEACM LACMSGPLAT CSLPALQGPC KAYVPRWAYN SQTGLCQSFV YGGCEGNGNN

421 FESREACEES CPFPRGNQHC RACKPRQKLV TSFCRSDFVI LGRVSELTEE QDSGRALVTV

481 DEVLKDEKMG LKFLGREPLE VTLLHVDWTC PCPNVTVGET PLIIMGEVDG GMAMLRPDSF

541 VGASSTRRVR KLREVMYKKT CDVLKDFLGL Q

도면7A

예측된 인간 GASP1 뉴클레오티드 서열

서열 번호: 4

1 atgaatccca acctctgggt ggacgcacag agcacctgca ggcgggagtg tgagacggac
61 caggagtgtg agatggacca ggtgagtggg atccagaagc cacagtgtga ggcagaccag
121 gtgaatgggg tccagaagcc gcaatgtgag atggaccaga agtgggagtg tgaggttgac
181 caggtgagtg gggtccagaa gccggtgtgt gaggcggacc aggtgagtgg ggtccagaag
241 ccacagtgtg agatggacca ggtgagtggg atccagaagc tggagtgtga ggcggaccag
301 aagtgggagt atgaggtgga ccaggtgagt ggggtccaga agccacagtg tgagatggac
361 caggtgagtg ggatccagaa gctggagtgt gaggcggacc aggagtgtga gacctatgag
421 aagtgctgcc ccaacgtatg tgggaccaag agctgcgtgg cggcccgcta catggacgtg
481 aaagggaaga agggcccagt gggcatgccc aaggaggcca catgtgacca cttcatgtgt
541 ctgcagcagg gctctgagtg tgacatctgg gatggccagc ccgtgtgtaa gtgcaaagac
601 cgctgtgaga aggagcccag ctttacctgc gcctcggacg gcctcaccta ctataaccgc
661 tgctacatgg atgccgaggc ctgctccaaa ggcatcacac tggccgttgt aacctgccgc
721 tatcacttca cntggcccaa caccagcccc ccaccacctg agaccaccat gcaccccacc
781 acagcctccc cagagacccc tgagctggac atggcggccc ctgcgctgct caacaaccct
841 gtgcaccagt cggtcaccat gggtgagaca gtgagcttcc tctgtgatgt ggtgggccgg
901 ccccggcctg agatcacctg ggagaagcag ttggaggatc gggagaatgt ggtcatgcgg
961 cccaaccatg tgcgtggcaa cgtggtggtc accaacattg cccagctggt catctataac
1021 gcccagctgc aggatgctgg gatctacacc tgcacggccc ggaacgtggc tggggtcctg
1081 agggctgatt tcccgctgtc ggtggtcagg ggtcatcagg ctgcagccac ctcagagagc
1141 agccccaatg gcacggcttt cccggcggcc gagtgcctga agcccccaga cagtgaggac
1201 tgtggcgaag agcagacccg ctggcacttc gatgcccagg ccaacaactg cctgaccttc
1261 accttcggcc actgccaccg taacctcaac cactttgaga cctatgaggc ctgcatgctg
1321 gcctgcatga gcgggccgct ggccgcgtgc agcctgcccg ccctgcaggg gccctgcaaa
1381 gcctacgcgc ctcgctgggc ttacaacagc cagacgggcc agtgccagtc ctttgtctat
1441 ggtggctgcg agggcaatgg caacaacttt gagagccgtg aggcctgtga ggagtcgtgc
1501 cccttcccca gggggaacca gcgctgtcgg gcctgcaagc ctcggcagaa gctcgttacc
1561 agcttctgtc gcagcgactt tgtcatcctg ggccgagtct ctgagctgac cgaggagcct
1621 gactcgggcc gcgccctggt gactgtggat gaggtcctaa aggatgagaa aatgggcctc
1681 aagttcctgg gccaggagcc attggaggtc actctgcttc acgtggactg ggcatgcccc
1741 tgccccaacg tgaccgtgag cgagatgccg ctcatcatca tgggggaggt ggacggcggc
1801 atggccatgc tgcgccccga tagctttgtg ggcgcatcga gtgcccgccg ggtcaggaag
1861 cttcgtgagg tcatgcacaa gaagacctgt gacgtcctca aggagtttct tggcttgcac
1921 tga

도면7B

인간 GASP1 의 예측된 아미노산 서열

서열 번호: 5

1 MNPNLWVDAQ STCRRECETD QECEMDQVSG IQKPQCEADQ VNGVQKPQCE MDQKWECEVD
61 QVSGVQKPVC EADQVSGVQK PQCEMDQVSG IQKLECEADQ KWEYEVDQVS GVQKPQCEMD
121 QVSGIQKLEC EADQECETYE KCCPNVCGTK SCVAARYMDV KGKKGPVGMP KEA**TCDHFMC**
181 **LQQGSECDIW** **DGQPVCKCKD** **RCEKEPSFTC** **ASDGLTYYNR** **CYMDAEACSK** **GITLAVVTC**R
241 YHFTWPNTSP PPPETTMHPT TASPETPELD MAAPALLNNP VHQSVTMGET VSFLCDVVGR
301 PRPEITWEKQ LEDRENVVMR PNHVRGNVVV TNIAQLVIYN AQLQDAGIYT CTARNVAGVL
361 RADFPLSVVR GHQAAATSES SPNGTAFPAA ECLKPPDSED CGEEQTRWHF DAQANNCLTF
421 TFGHCHRNLN HFETYEACML ACMSGPLAAC SLPALQGPCK AYAPRWAYNS QTGQCQSFVY
481 GGCEGNGNNF ESREACEESC PFPRGNQRCR ACKPRQKLVT SFCRSDFVIL GRVSELTEEP
541 DSGRALVTVD EVLKDEKMGL KFLGQEPLEV TLLHVDWACP CPNVTVSEMP LIIMGEVDGG
601 MAMLRPDSFV GASSARRVRK LREVMHKKTC DVLKEFLGLH

도면7C

대안적인 개시 부위를 이용한 인간 GASP1 의 예측된 뉴클레오티드 서열

서열 번호: 6

1 atgtgggccc caaggtgtcg ccggttctgg tctcgctggg agcaggtggc agcgctgctg
61 ctgctgctgc tactgctcgg ggtgccccg cgaagcctgg cgctgccgcc catccgctat
121 tcccacgccg gcatctgccc caacgacatg aatcccaacc tctgggtgga cgcacagagc
181 acctgcaggc gggagtgtga gacggaccag gagtgtgaga cctatgagaa gtgctgcccc
241 aacgtatgtg ggaccaagag ctgcgtggcg gcccgctaca tggacgtgaa agggaagaag
301 ggcccagtgg gcatgcccaa ggaggccaca tgtgaccact tcatgtgtct gcagcagggc
361 tctgagtgtg acatctggga tggccagccc gtgtgtaagt gcaaagaccg ctgtgagaag
421 gagcccagct ttacctgcgc ctcggacggc ctcacctact ataaccgctg ctacatggat
481 gccgaggcct gctccaaagg catcacactg gccgttgtaa cctgccgcta tcacttcacc
541 tggcccaaca ccagcccccc accacctgag accaccatgc accccaccac agcctcccca
601 gagacccctg agctggacat ggcggcccct gcgctgctca acaaccctgt gcaccagtcg
661 gtcaccatgg gtgagacagt gagtttcctc tgtgatgtgg tgggccggcc ccggcctgag
721 atcacctggg agaagcagtt ggaggatcgg gagaatgtgg tcatgcggcc caaccatgtg
781 cgtggcaacg tggtggtcac caacattgcc cagctggtca tctataacgc ccagctgcag
841 gatgctggga tctacacctg cacggcccgg aacgtggctg gggtcctgag ggctgatttc
901 ccgctgtcgg tggtcagggg tcatcaggct gcagccacct cagagagcag ccccaatggc
961 acggctttcc cggcggccga gtgcctgaag cccccagaca gtgaggactg tggcgaagag
1021 cagacccgct ggcacttcga tgcccaggcc aacaactgcc tgaccttcac cttcggccac
1081 tgccaccgta acctcaacca ctttgagacc tatgaggcct gcatgctggc ctgcatgagc
1141 gggccgctgg ccgcgtgcag cctgcccgcc ctgcaggggc cctgcaaagc ctacgcgcct
1201 cgctgggctt acaacagcca gacgggccag tgccagtcct ttgtctatgg tggctgcgag
1261 ggcaatggca acaactttga gagccgtgag gcctgtgagg agtcgtgccc cttccccagg
1321 gggaaccagc gctgtcgggc ctgcaagcct cggcagaagc tcgttaccag cttctgtcgc
1381 agcgactttg tcatcctggg ccgagtctct gagctgaccg aggagcctga ctcgggccgc
1441 gccctggtga ctgtggatga ggtcctaaag gatgagaaaa tgggcctcaa gttcctgggc
1501 caggagccat tggaggtcac tctgcttcac gtggactggg catgcccctg ccccaacgtg
1561 accgtgagcg agatgccgct catcatcatg gggaggtgg acggcggcat ggccatgctg
1621 cgccccgata gctttgtggg cgcatcgagt gcccgccggg tcaggaagct tcgtgaggtc
1681 atgcacaaga agacctgtga cgtcctcaag gagtttcttg gcttgcactg a

도면7D

대안적인 개시 부위를 이용한 인간 GASP1 의 예측된 아미노산 서열

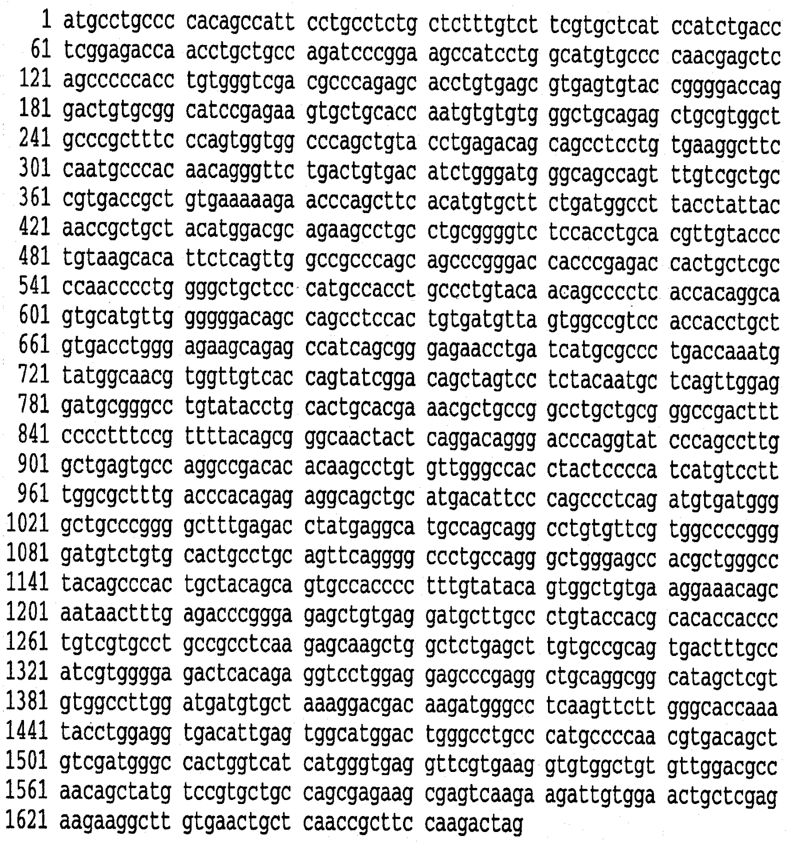
서열 번호: 7

1 MWAPRCRRFW SRWEQVAALL LLLLLLGVPP RSLALPPIRY SHAGICPNDM NPNLWVDAQS
61 TCRRECETDQ ECETYEKCCP NVCGTKSCVA ARYMDVKGKK GPVGMPKEA**T** **CDHFMCLQQG**
121 **SECDIWDGQP** **VCKCKDRCEK** **EPSFTCASDG** **LTYYNRCYMD** **AEACSKGITL** **AVVTC**RYHFT
181 WPNTSPPPPE TTMHPTTASP ETPELDMAAP ALLNNPVHQS VTMGETVSFL CDVVGRPRPE
241 ITWEKQLEDR ENVVMRPNHV RGNVVVTNIA QLVIYNAQLQ DAGIYTCTAR NVAGVLRADF
301 PLSVVRGHQA AATSESSPNG TAFPAAECLK PPDSEDCGEE QTRWHFDAQA NNCLTFTFGH
361 CHRNLNHFET YEACMLACMS GPLAACSLPA LQGPCKAYAP RWAYNSQTGQ CQSFVYGGCE
421 GNGNNFESRE ACEESCPFPR GNQRCRACKP RQKLVTSFCR SDFVILGRVS ELTEEPDSGR
481 ALVTVDEVLK DEKMGLKFLG QEPLEVTLLH VDWACPCPNV TVSEMPLIIM GEVDGGMAML
541 RPDSFVGASS ARRVRKLREV MHKKTCDVLK EFLGLH*

도면8A

예측된 마우스 GASP2 뉴클레오티드 서열

서열 번호: 8

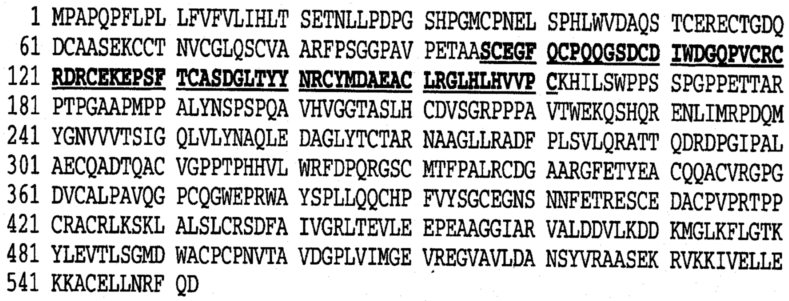


1 atgcctgccc cacagccatt cctgcctctg ctctttgtct tcgtgctcat ccatctgacc
61 tcggagacca acctgctgcc agatcccgga agccatcctg gcatgtgccc caacgagctc
121 agcccccacc tgtgggtcga cgcccagagc acctgtgagc gtgagtgtac cggggaccag
181 gactgtgcgg catccgagaa gtgctgcacc aatgtgtgtg ggctgcagag ctgcgtggct
241 gcccgctttc ccagtggtgg cccagctgta cctgagacag cagcctcctg tgaaggcttc
301 caatgcccac aacagggttc tgactgtgac atctgggatg ggcagccagt ttgtcgctgc
361 cgtgaccgct gtgaaaaaga acccagcttc acatgtgctt ctgatggcct tacctattac
421 aaccgctgct acatggacgc agaagcctgc ctgcggggtc tccacctgca cgttgtaccc
481 tgtaagcaca ttctcagttg gccgcccagc agcccgggac cacccgagac cactgctcgc
541 ccaacccctg gggctgctcc catgccacct gccctgtaca acagcccctc accacaggca
601 gtgcatgttg gggggacagc cagcctccac tgtgatgtta gtggccgtcc accacctgct
661 gtgacctggg agaagcagag ccatcagcgg gagaacctga tcatgcgccc tgaccaaatg
721 tatggcaacg tggttgtcac cagtatcgga cagctagtcc tctacaatgc tcagttggag
781 gatgcgggcc tgtatacctg cactgcacga aacgctgccg gcctgctgcg ggccgacttt
841 ccccttccg ttttacagcg ggcaactact caggacaggg acccaggtat cccagccttg
901 gctgagtgcc aggccgacac acaagcctgt gttgggccac ctactcccca tcatgtcctt
961 tggcgctttg acccacagag aggcagctgc atgacattcc cagccctcag atgtgatggg
1021 gctgcccggg gctttgagac ctatgaggca tgccagcagg cctgtgttcg tggccccggg
1081 gatgtctgtg cactgcctgc agttcagggg ccctgccagg gctgggagcc acgctgggcc
1141 tacagcccac tgctacagca gtgccacccc tttgtataca gtggctgtga aggaaacagc
1201 aataactttg agacccggga gagctgtgag gatgcttgcc ctgtaccacg cacaccaccc
1261 tgtcgtgcct gccgcctcaa gagcaagctg gctctgagct tgtgccgcag tgactttgcc
1321 atcgtgggga gactcacaga ggtcctggag gagcccgagg ctgcaggcgg catagctcgt
1381 gtggccttgg atgatgtgct aaaggacgac aagatgggcc tcaagttctt gggcaccaaa
1441 tacctggagg tgacattgag tggcatggac tgggcctgcc catgccccaa cgtgacagct
1501 gtcgatgggc cactggtcat catgggtgag gttcgtgaag gtgtggctgt gttggacgcc
1561 aacagctatg tccgtgctgc cagcgagaag cgagtcaaga agattgtgga actgctcgag
1621 aagaaggctt gtgaactgct caaccgcttc caagactag

도면8B

예측된 마우스 GASP2 아미노산 서열

서열 번호: 9



1 MPAPQPFLPL LFVFVLIHLT SETNLLPDPG SHPGMCPNEL SPHLWVDAQS TCERECTGDQ
61 DCAASEKCCT NVCGLQSCVA ARFPSGGPAV PETAA**SCEGF** **QCPQQGSDCD** **IWDGQPVCRC**
121 **RDRCEKEPSF** **TCASDGLTYY** **NRCYMDAEAC** **LRGLHLHVVP** CKHILSWPPS SPGPPETTAR
181 PTPGAAPMPP ALYNSPSPQA VHVGGTASLH CDVSGRPPPA VTWEKQSHQR ENLIMRPDQM
241 YGNVVVTSIG QLVLYNAQLE DAGLYTCTAR NAAGLLRADF PLSVLQRATT QDRDPGIPAL
301 AECQADTQAC VGPPTPHHVL WRFDPQRGSC MTFPALRCDG AARGFETYEA CQQACVRGPG
361 DVCALPAVQG PCQGWEPRWA YSPLLQQCHP FVYSGCEGNS NNFETRESCE DACPVPRTPP
421 CRACRLKSKL ALSLCRSDFA IVGRLTEVLE EPEAAGGIAR VALDDVLKDD KMGLKFLGTK
481 YLEVTLSGMD WACPCPNVTA VDGPLVIMGE VREGVAVLDA NSYVRAASEK RVKKIVELLE
541 KKACELLNRF QD

도면9A

인간 GASP2 의 예측된 뉴클레오티드 서열

서열 번호: 10

1 atgcccgccc tacgtccact cctgccgctc ctgctcctcc tccggctgac ctcgggggct
61 ggcttgctgc cagggctggg gagccacccg ggcgtgtgcc ccaaccagct cagccccaac
121 ctgtgggtgg acgcccagag cacctgtgag cgcgagtgta gcagggacca ggactgtgcg
181 gctgctgaga agtgctgcat caacgtgtgt ggactgcaca gctgcgtggc agcacgcttc
241 cccggcagcc cagctgcgcc gacgacagcg gcctcctgcg agggctttgt gtgcccacag
301 cagggctcgg actgcgacat ctgggacggg cagcccgtgt gccgctgccg cgaccgctgt
361 gagaaggagc ccagcttcac ctgcgcctcg gacggcctca cctactacaa ccgctgctat
421 atggacgccg aggcctgcct gcggggcctg cacctccaca tcgtgccctg caagcacgtg
481 ctcagctggc cgcccagcag cccggggccg ccggagacca ctgcccgccc cacacctggg
541 gccgcgcccg tgcctcctgc cctgtacagc agcccctccc cacaggcggt gcaggttggg
601 ggtacggcca gcctccactg cgacgtcagc ggccgcccgc cgcctgctgt gacctgggag
661 aagcagagtc accagcgaga gaacctgatc atgcgccctg atcagatgta tggcaacgtg
721 gtggtcacca gcatcgggca gctggtgctc tacaacgcgc ggcccgaaga cgccggcctg
781 tacacctgca ccgcgcgcaa cgctgctggg ctgctgcggg ctgacttccc actctctgtg
841 gtccagcgag agccggccag ggacgcagcc cccagcatcc cagccccggc cgagtgcctg
901 ccggatgtgc aggcctgcac gggccccact tccccacacc ttgtcctctg gcactacgac
961 ccgcagcggg gcggctgcat gaccttcccg gcccgtggct gtgatggggc ggcccgcggc
1021 tttgagacct acgaggcatg ccagcaggcc tgtgcccgcg gccccggcga cgcctgcgtg
1081 ctgcctgccg tgcagggccc ctgccggggc tgggagccgc gctgggccta cagcccgctg
1141 ctgcagcagt gccatccctt cgtgtacggt ggctgcgagg gcaacggcaa caacttccac
1201 agccgcgaga gctgcgagga tgcctgcccc gtgccgcgca caccgccctg ccgcgcctgc
1261 cgcctccgga gcaagctggc gctgagcctg tgccgcagcg acttcgccat cgtggggcgg
1321 ctcacggagg tgctggagga gcccgaggcc gccggcggca tcgcccgcgt ggcgctcgag
1381 gacgtgctca aggatgacaa gatgggcctc aagttcttgg gcaccaagta cctggaggtg
1441 acgctgagtg gcatggactg ggcctgcccc tgccccaaca tgacggcggg cgacgggccg
1501 ctggtcatca tgggtgaggt gcgcgatggc gtggccgtgc tggacgccgg cagctacgtc
1561 cgcgccgcca gcgagaagcg cgtcaagaag atcttggagc tgctggagaa gcaggcctgc
1621 gagctgctca accgcttcca ggactagccc ccgcaggggc ctgcgccacc ccgtcctggt
1681 gaataaacgc actcc

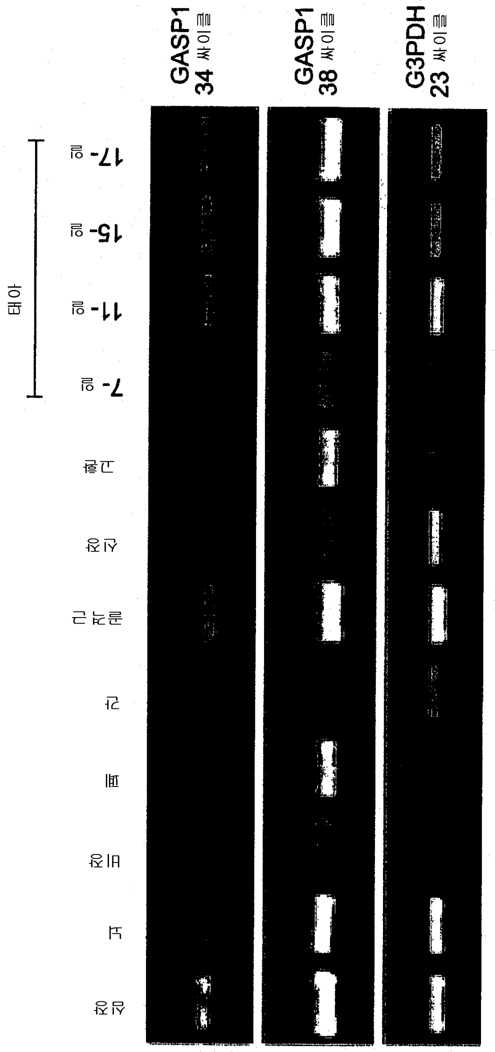
도면9B

인간 GASP-2 의 예측된 아미노산 서열

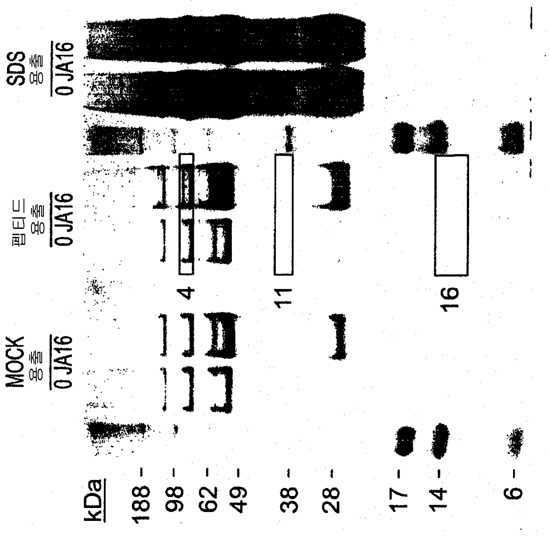
서열 번호: 11

1 MPALRPLLPL LLLLRLTSGA GLLPGLGSHP GVCPNQLSPN LWVDAQSTCE RECSRDQDCA
61 AAEKCCINVC GLHSCVAARF PGSPAAPTTA A**SCEGFVCPQ** **QGSDCDIWDG** **QPVCRCRDRC**
121 **EKEPSFTCAS** **DGLTYYNRCY** **MDAEACLRGL** **HLHIVPC**KHV LSWPPSSPGP PETTARPTPG
181 AAPVPPALYS SPSPQAVQVG GTASLHCDVS GRPPPAVTWE KQSHQRENLI MRPDQMYGNV
241 VVTSIGQLVL YNARPEDAGL YTCTARNAAG LLRADFPLSV VQREPARDAA PSIPAPAECL
301 PDVQACTGPT SPHLVLWHYD PQRGGCMTFP ARGCDGAARG FETYEACQQA CARGPGDACV
361 LPAVQGPCRG WEPRWAYSPL LQQCHPFVYG GCEGNGNNFH SRESCEDACP VPRTPPCRAC
421 RLRSKLALSL CRSDFAIVGR LTEVLEEPEA AGGIARVALE DVLKDDKMGL KFLGTKYLEV
481 TLSGMDWACP CPNMTAGDGP LVIMGEVRDG VAVLDAGSYV RAASEKRVKK ILELLEKQAC
541 ELLNRFQD

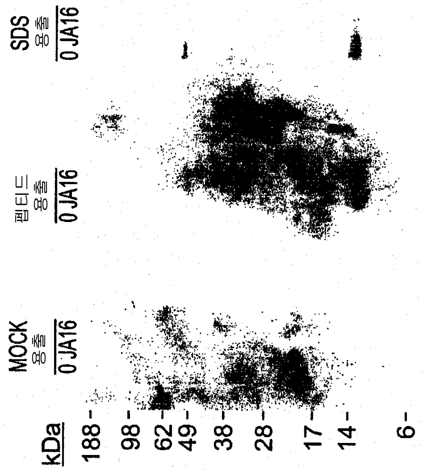
도면10


GASP1
34 싸이클
GASP1
38 싸이클
G3PDH
23 싸이클
심장
뇌
비장
폐
간
골격근
신장
고환
7-일
11-일
15-일
17-일
태아

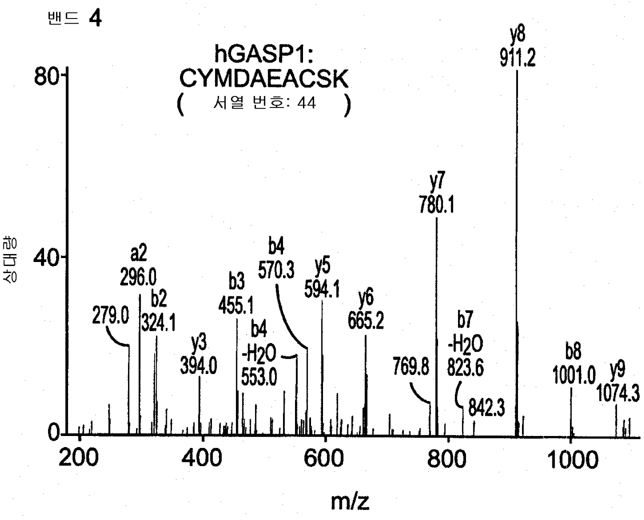
도면11A


kDa
188
98
62
49
38
28
17
14
6
MOCK
용출
0 JA16
펩티드
용출
0 JA16
SDS
용출
0 JA16
4
11
16

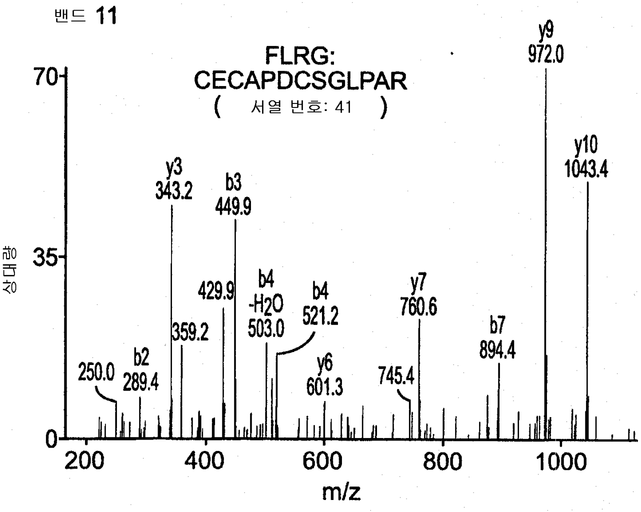
도면11B



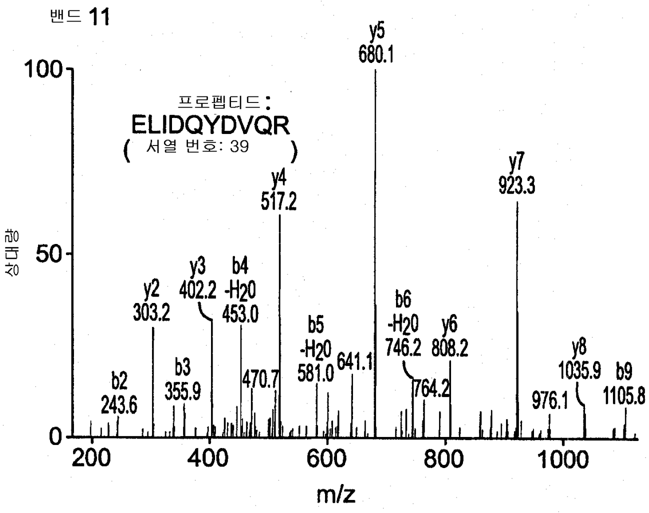
도면12A



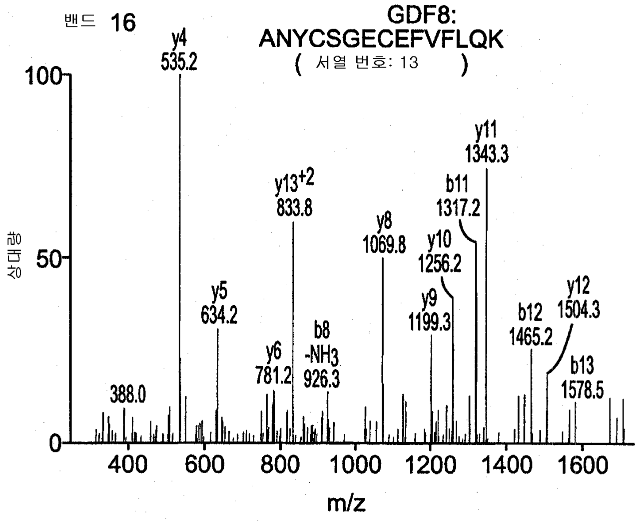
도면12B



도면12C


밴드 11
프로펩티드:
ELIDQYDVQR
(서열 번호: 39)
100
50
0
상대량
y5 680.1
y4 517.2
y7 923.3
y3 402.2
b4 -H20 453.0
y2 303.2
b5 -H20 581.0
641.1
b6 -H20 746.2
y6 808.2
764.2
470.7
b2 243.6
b3 355.9
y8 1035.9
b9 1105.8
976.1
200
400
600
800
1000
m/z

도면12D


밴드 16
GDF8:
ANYCSGECEFVFLQK
(서열 번호: 13)
100
50
0
상대량
y4 535.2
y13+2 833.8
y11 1343.3
b11 1317.2
y8 1069.8
y10 1256.2
y9 1199.3
y5 634.2
y12 1504.3
b12 1465.2
b8 -NH3 926.3
y6 781.2
b13 1578.5
388.0
400
600
800
1000
1200
1400
1600
m/z

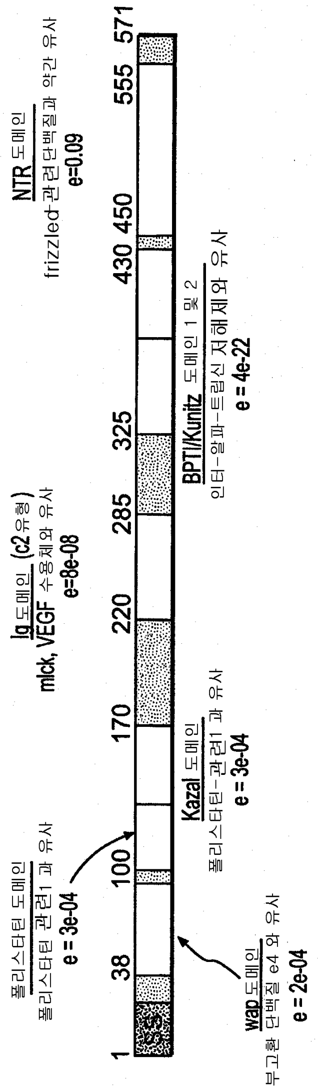
도면13

클로닝된 마우스 GASP1 뉴클레오티드 및 아미노산 서열

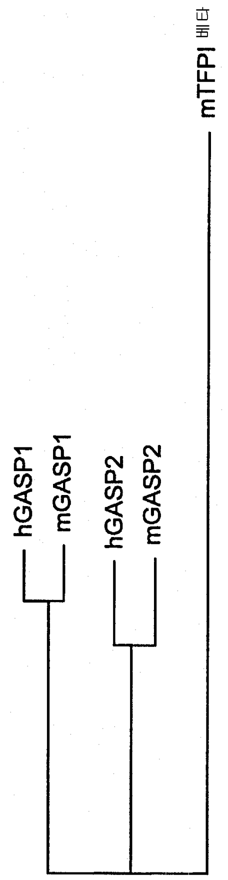
서열 번호: 48

```
      M  C  A  P  G  Y  H  R  F  W  F  H  W  G  L  L  L  L  L  L  L  E  A  P  L
   1  ATGTGTGCCCCAGGG TATCATCGGTTCTGG TTTCACTGGGGGCTG CTGTTGCTGCTGCTC CTCGAGGCTCCCCTT
      R  G  L  A  L  P  P  I  R  Y  S  H  A  G  I  C  P  N  D  M  N  P  N  L  W
  76  CGAGGCCTAGCACTG CCACCCATCCGATAC TCCCATGCGGGCATC TGCCCCAACGACATG AACCCCAACCTCTGG
      V  D  A  Q  S  T  C  K  R  E  C  E  T  D  Q  E  C  E  T  Y  E  K  C  C  P
 151  GTGGATGCCCAGAGC ACCTGCAAGCGAGAG TGTGAAACAGACCAG GAATGTGAGACCTAT GAGAAATGCTGCCCC
      N  V  C  G  T  K  S  C  V  A  A  R  Y  M  D  V  K  G  K  K  G  P  V  G  M
 226  AATGTGTGTGGGACC AAGAGCTGTGTGGCA GCCCGCTACATGGAT GTGAAAGGGAAGAAG GGGCCTGTAGGCATG
      P  K  E  A  T  C  D  H  F  M  C  L  Q  Q  G  S  E  C  D  I  W  D  G  Q  P
 301  CCCAAGGAGGCCACA TGTGACCATTTCATG TGCCTGCAGCAGGGC TCTGAGTGTGACATC TGGGACGGCCAGCCC
      V  C  K  C  K  D  R  C  E  K  E  P  S  F  T  C  A  S  D  G  L  T  Y  Y  N
 376  GTGTGTAAGTGCAAA GATCGCTGTGAGAAG GAGCCCAGCTTCACC TGTGCCTCTGATGGC CTTACCTACTACAAC
      R  C  F  M  D  A  E  A  C  S  K  G  I  T  L  S  V  V  T  C  R  Y  H  F  T
 451  CGTTGCTTCATGGAC GCCGAAGCCTGCTCC AAGGGCATCACACTG TCTGTGGTCACCTGT CGTTATCACTTCACC
      W  P  N  T  S  P  P  P  P  E  T  T  V  H  P  T  T  A  S  P  E  T  L  G  L
 526  TGGCCTAACACCAGC CCTCCACCGCCTGAG ACCACGGTGCATCCC ACCACCGCCTCTCCG GAGACTCTCGGGCTG
      D  M  A  A  P  A  L  L  N  H  P  V  H  Q  S  V  T  V  G  E  T  V  S  F  L
 601  GACATGGCAGCCCCA GCCCTGCTCAACCAC CCTGTCCATCAGTCA GTCACCGTGGGTGAG ACTGTGAGTTTCCTC
      C  D  V  V  G  R  P  R  P  E  L  T  W  E  K  Q  L  E  D  R  E  N  V  V  M
 676  TGTGACGTGGTAGGC CGGCCTCGGCCAGAG CTCACTTGGGAGAAA CAGCTGGAGGACCGA GAGAATGTTGTCATG
      R  P  N  H  V  R  G  N  V  V  V  T  N  I  A  Q  L  V  I  Y  N  V  Q  P  Q
 751  AGGCCCAACCACGTG CGTGGTAATGTGGTG GTCACTAACATTGCC CAGCTGGTCATCTAC AACGTCCAGCCCCAG
      D  A  G  I  Y  T  C  T  A  R  N  V  A  G  V  L  R  A  D  F  P  L  S  V  V
 826  GATGCTGGCATATAC ACCTGTACAGCTCGA AATGTCGCTGGTGTC CTGAGGGCTGACTTC CCGTTGTCGGTGGTC
      R  G  G  Q  A  R  A  T  S  E  S  S  L  N  G  T  A  F  P  A  T  E  C  L  K
 901  AGGGGTGGTCAGGCC AGGGCCACTTCAGAG AGCAGTCTCAATGGC ACAGCTTTTCCAGCA ACAGAGTGCCTGAAG
      P  P  D  S  E  D  C  G  E  E  Q  T  R  W  H  F  D  A  Q  A  N  N  C  L  T
 976  CCCCCAGACAGTGAG GACTGTGGAGAGGAG CAGACACGCTGGCAC TTCGACGCCCAGGCT AACAACTGCCTCACT
      F  T  F  G  H  C  H  H  N  L  N  H  F  E  T  Y  E  A  C  M  L  A  C  M  S
1051  TTCACCTTTGGCCAC TGCCACCACAATCTC AACCACTTTGAGACC TACGAGGCCTGTATG CTGGCTTGTATGAGT
      G  P  L  A  T  C  S  L  P  A  L  Q  G  P  C  K  A  Y  V  P  R  W  A  Y  N
1126  GGGCCATTGGCCACC TGCAGCCTGCCTGCC CTGCAAGGGCCTTGC AAAGCTTATGTCCCA CGCTGGGCCTACAAC
      S  Q  T  G  L  C  Q  S  F  V  Y  G  G  C  E  G  N  G  N  N  F  E  S  R  E
1201  AGCCAGACAGGCCTA TGCCAGTCCTTCGTC TATGGCGGCTGTGAG GGCAACGGTAACAAC TTTGAAAGCCGTGAG
      A  C  E  E  S  C  P  F  P  R  G  N  Q  H  C  R  A  C  K  P  R  Q  K  L  V
1276  GCTTGTGAGGAGTCG TGTCCCTTCCCGAGG GGTAACCAGCACTGC CGGGCCTGCAAGCCC CGGCAAAAACTTGTT
      T  S  F  C  R  S  D  F  V  I  L  G  R  V  S  E  L  T  E  E  Q  D  S  G  R
1351  ACCAGCTTCTGTCGG AGTGACTTTGTCATC CTGGGCAGGGTCTCT GAGCTGACCGAGGAG CAAGACTCGGGCCGT
      A  L  V  T  V  D  E  V  L  K  D  E  K  M  G  L  K  F  L  G  R  E  P  L  E
1426  GCCCTGGTGACCGTG GATGAGGTCTTAAAA GATGAGAAGATGGGC CTCAAGTTTCTGGGC CGGGAGCCTCTGGAA
      V  T  L  L  H  V  D  W  T  C  P  C  P  N  V  T  V  G  E  T  P  L  I  I  M
1501  GTCACCCTGCTTCAT GTAGACTGGACCTGT CCTTGCCCCAACGTG ACAGTGGGTGAGACA CCACTCATCATCATG
      G  E  V  D  G  G  M  A  M  L  R  P  D  S  F  V  G  A  S  S  T  R  R  V  R
1576  GGGGAGGTGGACGGC GGCATGGCCATGCTG AGACCCGATAGCTTT GTGGGGGCATCGAGC ACACGGCGGGTCAGG
      K  L  R  E  V  M  Y  K  K  T  C  D  V  L  K  D  F  L  G  L  Q  *
1651  AAGCTCCGTGAGGTC ATGTACAAGAAAACC TGTGACGTCCTCAAG GACTTCCTGGGCTTG CAATGA
```

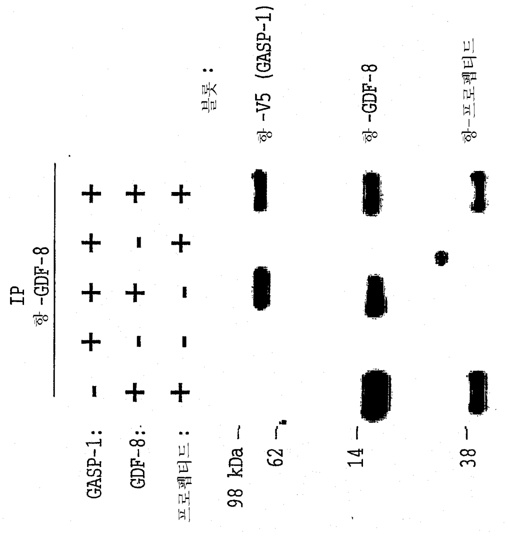
도면14A


1
38
100
170
220
285
325
430
450
555
571
SS
폴리스타틴 도메인
폴리스타틴 관련1 과 유사
e = 3e-04
Ig 도메인 (c2 유형)
mlck, VEGF 수용체와 유사
e=8e-08
NTR 도메인
frizzled-관련단백질과 약간 유사
e=0.09
wap 도메인
부고환 단백질 e4 와 유사
e = 2e-04
Kazal 도메인
폴리스타틴-관련1 과 유사
e = 3e-04
BPTI/Kunitz 도메인 1 및 2
인터-알파-트립신 저해제와 유사
e = 4e-22

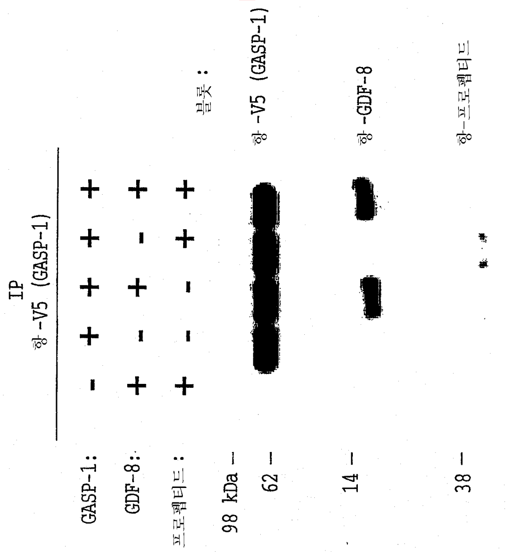
도면14B


hGASP1
mGASP1
hGASP2
mGASP2
mTFPI 베타

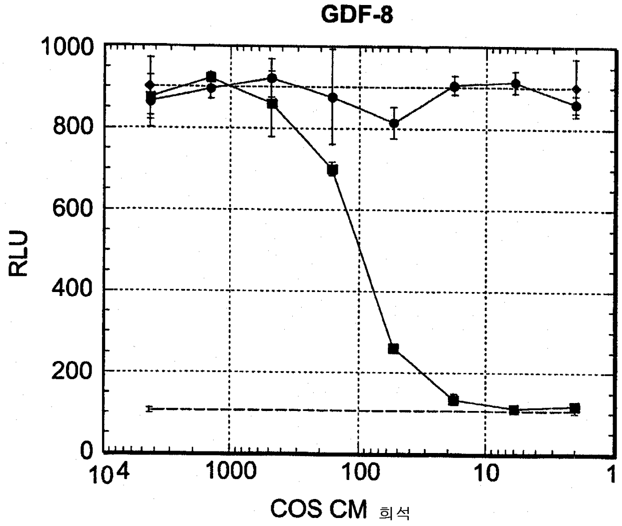
도면15A


IP
항-GDF-8
GASP-1: - + + + +
GDF-8: + - + - +
프로펩티드 : + - - + +
98 kDa
62
14
38
블롯 :
항-V5 (GASP-1)
항-GDF-8
항-프로펩티드

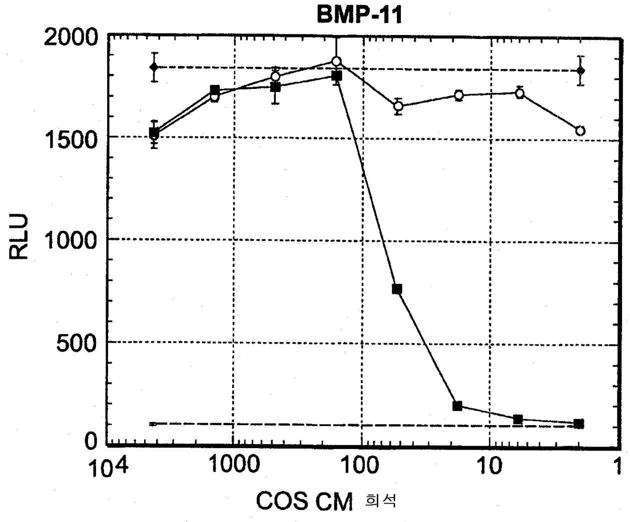
도면15B



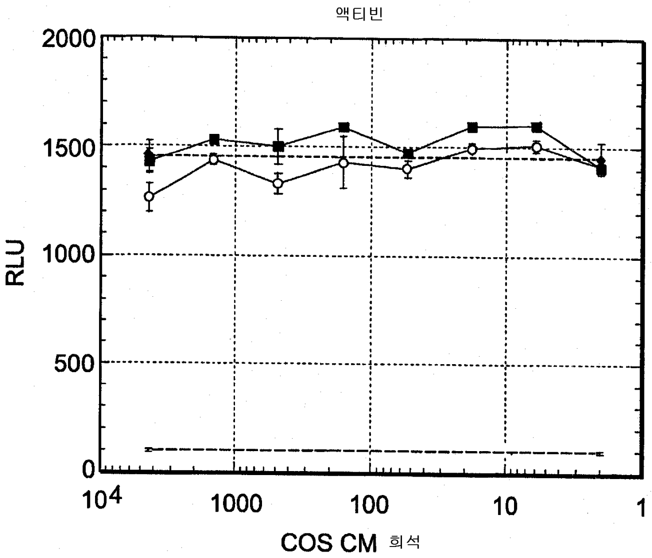
도면16A



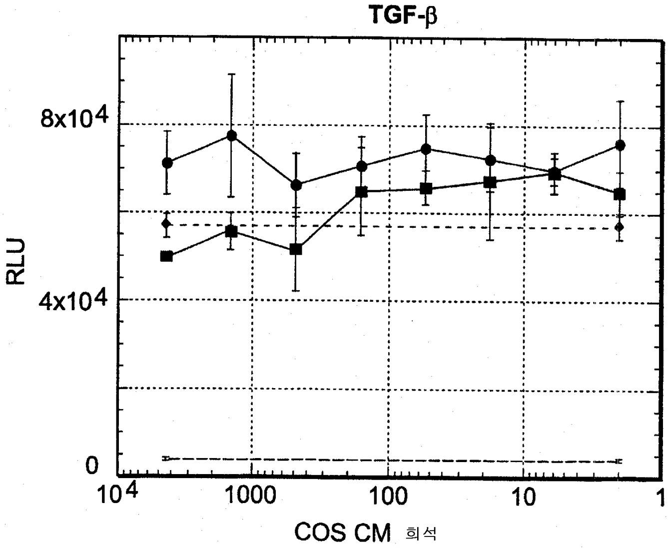
도면16B



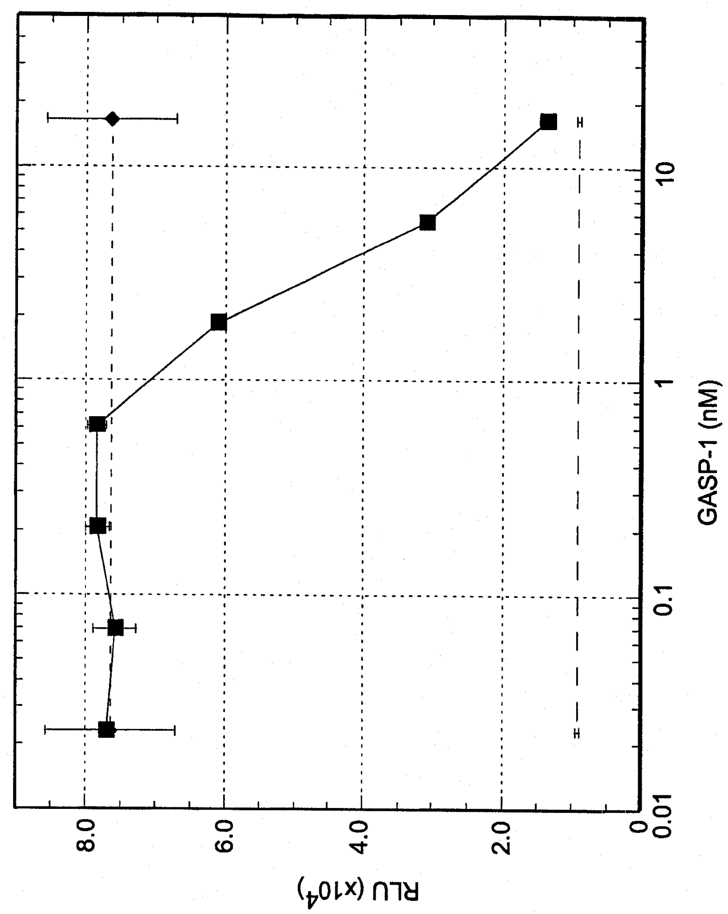
도면16C


액티빈
2000
1500
1000
500
0
RLU
10^4
1000
100
10
1
COS CM 희석

도면16D


TGF-β
8x10^4
4x10^4
0
RLU
10^4
1000
100
10
1
COS CM 희석

도면17



<110> HILL, JENNIFER J.
WOLFMAN, NEIL M.

<120> FOLLISTATIN DOMAIN CONTAINING PROTEINS

<130> 08702.0015-00

<150> 60/357,846
<151> 2002-02-21

<150> 60/434,645
<151> 2002-12-20

<160> 53

<170> PatentIn Ver. 2.1

<210> 1

<211> 1716

<212> DNA

<213> Mus sp.

<400> 1

atgtgtgccc cagggtatca tcggttctgg tttcactggg ggctgctgtt gctgctgctc 60

ctcgaggctc cccttcgagg cctagcactg ccacccatcc gatactccca tgcgggcatc 120

tgccccaacg acatgaaccc caacctctgg gtggatgccc agagcacctg caagcgagag 180

tgtgaaacag accaggaatg tgagacctat gagaaatgct gccccaatgt gtgtgggacc 240

aagagctgtg tggcagcccg ctacatggat gtgaaaggga agaagggccc tgtgggcatg 300

cccaaggagg ccacatgtga ccatttcatg tgcctgcagc agggctctga gtgtgacatc 360

tgggacggcc agcccgtgtg taagtgcaaa gatcgctgtg agaaggagcc cagcttcacc 420

tgtgcctctg atggccttac ctactacaac cgttgcttca tggacgccga agcctgctcc 480

aagggcatca cactgtctgt ggtcacctgt cgttatcact tcacctggcc taacaccagc 540

cctccaccgc ctgagaccac ggtgcatccc accaccgcct ctccggagac tctcgggctg 600

gacatggcag ccccggccct gctcaaccac cctgtccatc agtcagtcac cgtgggtgag 660

actgtgagtt tcctctgtga cgtggtaggc cggcctcggc cagagctcac ttgggagaaa 720

cagctggagg accgagaaaa tgttgtcatg aggcccaacc acgtgcgcgg taatgtggtg 780

gtcactaaca ttgcccagct ggtcatctac aacgtccagc cccaggatgc tggcatatac 840 acctgtacag ctcgaaatgt cgctggtgtc ctgagggctg acttcccgtt gtcggtggtc 900

aggggtggtc aggccagggc cacttcagag agcagtctca atggcacagc ttttccagca 960

acagagtgcc tgaagccccc agacagtgag gactgtggag aggagcagac acgctggcac 1020

ttcgacgccc aggctaacaa ctgcctcact ttcacctttg gccactgcca ccacaatctc 1080

aaccactttg agacctacga ggcctgtatg ctggcttgta tgagtgggcc attggccacc 1140

tgcagcctgc ctgccctgca agggccttgc aaagcttatg tcccacgctg ggcctacaac 1200

agccagacag gcctatgcca gtccttcgtc tatggcggct gtgagggcaa cggtaacaac 1260

tttgaaagcc gtgaggcttg tgaggagtcg tgtcccttcc cgaggggtaa ccagcactgc 1320

cgggcctgca agccccggca aaaacttgtt accagcttct gtcggagtga ctttgtcatc 1380

ctgggcaggg tctctgagct gaccgaagag caagactcag gccgtgccct ggtgaccgtg 1440

gatgaggtct taaaagatga gaagatgggc ctcaagtttc tgggccggga gcctctggaa 1500

gtcaccctgc ttcatgtaga ctggacctgt ccttgcccca acgtgacagt gggtgagaca 1560

ccactcatca tcatggggga ggtcgacggc ggcatggcca tgctgagacc cgatagcttt 1620

gtgggggcat cgagcacacg gcgggtcagg aagctccgtg aggtcatgta caagaaaacc 1680

tgtgacgtcc tcaaggactt cctgggcttg caatga 1716

<210> 2

<211> 1716

<212> DNA

<213>   Mus sp.

<400>   2

atgtgtgccc cagggtatca tcggttctgg tttcactggg ggctgctgtt gctgctgctc   60

ctcgaggctc cccttcgagg cctagcactg ccacccatcc gatactccca tgcgggcatc   120

tgccccaacg acatgaaccc caacctctgg gtggatgccc agagcacctg caagcgagag   180

tgtgaaacag accaggaatg tgagacctat gagaaatgct gccccaatgt gtgtgggacc   240

aagagctgtg tggcagcccg ctacatggat gtgaaaggga agaaggggcc tgtaggcatg   300

cccaaggagg ccacatgtga ccatttcatg tgcctgcagc agggctctga gtgtgacatc   360

tgggacggcc agcccgtgtg taagtgcaaa gatcgctgtg agaaggagcc cagcttcacc   420

tgtgcctctg atggccttac ctactacaac cgttgcttca tggacgccga agcctgctcc   480

aagggcatca cactgtctgt ggtcacctgt cgttatcact tcacctggcc taacaccagc   540

cctccaccgc ctgagaccac ggtgcatccc accaccgcct ctccggagac tctcgggctg   600

gacatggcag ccccagccct gctcaaccac cctgtccatc agtcagtcac cgtgggtgag   660

actgtgagtt tcctctgtga cgtggtaggc cggcctcggc cagagctcac ttgggagaaa   720

cagctggagg accgagagaa tgttgtcatg aggcccaacc acgtgcgtgg taatgtggtg   780

gtcactaaca ttgcccagct ggtcatctac aacgtccagc cccaggatgc tggcatatac   840

acctgtacag ctcgaaatgt cgctggtgtc ctgagggctg acttcccgtt gtcggtggtc   900

aggggtggtc aggccagggc cacttcagag agcagtctca atggcacagc ttttccagca   960 acagagtgcc tgaagccccc agacagtgag gactgtggag aggagcagac acgctggcac 1020

ttcgacgccc aggctaacaa ctgcctcact ttcacctttg gccactgcca ccacaatctc 1080

aaccactttg agacctacga ggcctgtatg ctggcttgta tgagtgggcc attggccacc 1140

tgcagcctgc ctgccctgca agggccttgc aaagcttatg tcccacgctg ggcctacaac 1200

agccagacag gcctatgcca gtccttcgtc tatggcggct gtgagggcaa cggtaacaac 1260

tttgaaagcc gtgaggcttg tgaggagtcg tgtcccttcc cgaggggtaa ccagcactgc 1320

cgggcctgca agccccggca aaaacttgtt accagcttct gtcggagtga ctttgtcatc 1380

ctgggcaggg tctctgagct gaccgaggag caagactcgg gccgtgccct ggtgaccgtg 1440

gatgaggtct taaaagatga gaagatgggc ctcaagtttc tgggccggga gcctctggaa 1500

gtcaccctgc ttcatgtaga ctggacctgt ccttgcccca acgtgacagt gggtgagaca 1560

ccactcatca tcatggggga ggtggacggc ggcatggcca tgctgagacc cgatagcttt 1620

gtgggggcat cgagcacacg gcgggtcagg aagctccgtg aggtcatgta caagaaaacc 1680

tgtgacgtcc tcaaggactt cctgggcttg caatga 1716

<210> 3
<211> 571
<212> PRT
<213> Mus sp.

<400> 3

Met Cys Ala Pro Gly Tyr His Arg Phe Trp Phe His Trp Gly Leu Leu

1 5 10 15

Leu Leu Leu Leu Leu Glu Ala Pro Leu Arg Gly Leu Ala Leu Pro Pro

20 25 30

Ile Arg Tyr Ser His Ala Gly Ile Cys Pro Asn Asp Met Asn Pro Asn

35 40 45

Leu Trp Val Asp Ala Gln Ser Thr Cys Lys Arg Glu Cys Glu Thr Asp

50 55 60

Gln Glu Cys Glu Thr Tyr Glu Lys Cys Cys Pro Asn Val Cys Gly Thr

65 70 75 80

Lys Ser Cys Val Ala Ala Arg Tyr Met Asp Val Lys Gly Lys Lys Gly

85 90 95

Pro Val Gly Met Pro Lys Glu Ala Thr Cys Asp His Phe Met Cys Leu

100 105 110

Gln Gln Gly Ser Glu Cys Asp Ile Trp Asp Gly Gln Pro Val Cys Lys

115 120 125

Cys Lys Asp Arg Cys Glu Lys Glu Pro Ser Phe Thr Cys Ala Ser Asp

130 135 140

Gly Leu Thr Tyr Tyr Asn Arg Cys Phe Met Asp Ala Glu Ala Cys Ser

145 150 155 160

Lys Gly Ile Thr Leu Ser Val Val Thr Cys Arg Tyr His Phe Thr Trp

165 170 175

Pro Asn Thr Ser Pro Pro Pro Pro Glu Thr Thr Val His Pro Thr Thr

180 185 190

Ala Ser Pro Glu Thr Leu Gly Leu Asp Met Ala Ala Pro Ala Leu Leu

195 200 205

Asn His Pro Val His Gln Ser Val Thr Val Gly Glu Thr Val Ser Phe

210 215 220

Leu Cys Asp Val Val Gly Arg Pro Arg Pro Glu Leu Thr Trp Glu Lys

225 230 235 240

Gln Leu Glu Asp Arg Glu Asn Val Val Met Arg Pro Asn His Val Arg

245 250 255

Gly Asn Val Val Val Thr Asn Ile Ala Gln Leu Val Ile Tyr Asn Val

260 265 270

Gln Pro Gln Asp Ala Gly Ile Tyr Thr Cys Thr Ala Arg Asn Val Ala

275 280 285

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

290 295 300

Ala Arg Ala Thr Ser Glu Ser Ser Leu Asn Gly Thr Ala Phe Pro Ala

305 310 315 320

Thr Glu Cys Leu Lys Pro Pro Asp Ser Glu Asp Cys Gly Glu Glu Gln

325 330 335

Thr Arg Trp His Phe Asp Ala Gln Ala Asn Asn Cys Leu Thr Phe Thr

340 345 350

Phe Gly His Cys His His Asn Leu Asn His Phe Glu Thr Tyr Glu Ala

355 360 365

Cys Met Leu Ala Cys Met Ser Gly Pro Leu Ala Thr Cys Ser Leu Pro
370 375 380

Ala Leu Gln Gly Pro Cys Lys Ala Tyr Val Pro Arg Trp Ala Tyr Asn
385 390 395 400

Ser Gln Thr Gly Leu Cys Gln Ser Phe Val Tyr Gly Gly Cys Glu Gly
405 410 415

Asn Gly Asn Asn Phe Glu Ser Arg Glu Ala Cys Glu Glu Ser Cys Pro
420 425 430

Phe Pro Arg Gly Asn Gln His Cys Arg Ala Cys Lys Pro Arg Gln Lys
435 440 445

Leu Val Thr Ser Phe Cys Arg Ser Asp Phe Val Ile Leu Gly Arg Val
450 455 460

Ser Glu Leu Thr Glu Glu Gln Asp Ser Gly Arg Ala Leu Val Thr Val
465 470 475 480

Asp Glu Val Leu Lys Asp Glu Lys Met Gly Leu Lys Phe Leu Gly Arg
485 490 495

Glu Pro Leu Glu Val Thr Leu Leu His Val Asp Trp Thr Cys Pro Cys
500 505 510

Pro Asn Val Thr Val Gly Glu Thr Pro Leu Ile Ile Met Gly Glu Val
515 520 525

Asp Gly Gly Met Ala Met Leu Arg Pro Asp Ser Phe Val Gly Ala Ser
530 535 540

Ser Thr Arg Arg Val Arg Lys Leu Arg Glu Val Met Tyr Lys Lys Thr
545   550   555   560

Cys Asp Val Leu Lys Asp Phe Leu Gly Leu Gln
                565   570

<210> 4
<211> 1923
<212> DNA
<213> Homo sapiens

<220>
<221> modified_base
<222> (732)
<223> a, t, c or g

<400> 4
atgaatccca acctctgggt ggacgcacag agcacctgca ggcgggagtg tgagacggac 60

caggagtgtg agatggacca ggtgagtggg atccagaagc cacagtgtga ggcagaccag 120

gtgaatgggg tccagaagcc gcaatgtgag atggaccaga agtgggagtg tgaggttgac 180

caggtgagtg gggtccagaa gccggtgtgt gaggcggacc aggtgagtgg ggtccagaag 240

ccacagtgtg agatggacca ggtgagtggg atccagaagc tggagtgtga ggcggaccag 300

aagtgggagt atgaggtgga ccaggtgagt ggggtccaga agccacagtg tgagatggac 360

caggtgagtg ggatccagaa gctggagtgt gaggcggacc aggagtgtga gacctatgag 420

aagtgctgcc ccaacgtatg tgggaccaag agctgcgtgg cggcccgcta catggacgtg 480 aaagggaaga agggcccagt gggcatgccc aaggaggcca catgtgacca cttcatgtgt 540

ctgcagcagg gctctgagtg tgacatctgg gatggccagc ccgtgtgtaa gtgcaaagac 600

cgctgtgaga aggagcccag ctttacctgc gcctcggacg gcctcaccta ctataaccgc 660

tgctacatgg atgccgaggc ctgctccaaa ggcatcacac tggccgttgt aacctgccgc 720

tatcacttca cntggcccaa caccagcccc ccaccacctg agaccaccat gcaccccacc 780

acagcctccc cagagacccc tgagctggac atggcggccc ctgcgctgct caacaaccct 840

gtgcaccagt cggtcaccat gggtgagaca gtgagcttcc tctgtgatgt ggtgggccgg 900

ccccggcctg agatcacctg ggagaagcag ttggaggatc gggagaatgt ggtcatgcgg 960

cccaaccatg tgcgtggcaa cgtggtggtc accaacattg cccagctggt catctataac 1020

gcccagctgc aggatgctgg gatctacacc tgcacggccc ggaacgtggc tggggtcctg 1080

agggctgatt tcccgctgtc ggtggtcagg ggtcatcagg ctgcagccac ctcagagagc 1140

agccccaatg gcacggcttt cccggcggcc gagtgcctga agcccccaga cagtgaggac 1200

tgtggcgaag agcagacccg ctggcacttc gatgcccagg ccaacaactg cctgaccttc 1260

accttcggcc actgccaccg taacctcaac cactttgaga cctatgaggc ctgcatgctg 1320

gcctgcatga gcgggccgct ggccgcgtgc agcctgcccg ccctgcaggg gccctgcaaa 1380

gcctacgcgc ctcgctgggc ttacaacagc cagacgggcc agtgccagtc ctttgtctat 1440

ggtggctgcg agggcaatgg caacaacttt gagagccgtg aggcctgtga ggagtcgtgc 1500 cccttcccca gggggaacca gcgctgtcgg gcctgcaagc ctcggcagaa gctcgttacc 1560

agcttctgtc gcagcgactt tgtcatcctg ggccgagtct ctgagctgac cgaggagcct 1620

gactcgggcc gcgccctggt gactgtggat gaggtcctaa aggatgagaa aatgggcctc 1680

aagttcctgg gccaggagcc attggaggtc actctgcttc acgtggactg ggcatgcccc 1740

tgccccaacg tgaccgtgag cgagatgccg ctcatcatca tgggggaggt ggacggcggc 1800

atggccatgc tgcgccccga tagctttgtg ggcgcatcga gtgcccgccg ggtcaggaag 1860

cttcgtgagg tcatgcacaa gaagacctgt gacgtcctca aggagtttct tggcttgcac 1920

tga 1923

<210> 5

<211> 640

<212> PRT

<213> Homo sapiens

<400> 5

Met Asn Pro Asn Leu Trp Val Asp Ala Gln Ser Thr Cys Arg Arg Glu
1 5 10 15

Cys Glu Thr Asp Gln Glu Cys Glu Met Asp Gln Val Ser Gly Ile Gln
20 25 30

Lys Pro Gln Cys Glu Ala Asp Gln Val Asn Gly Val Gln Lys Pro Gln
35 40 45

Cys Glu Met Asp Gln Lys Trp Glu Cys Glu Val Asp Gln Val Ser Gly

50 55 60

Val Gln Lys Pro Val Cys Glu Ala Asp Gln Val Ser Gly Val Gln Lys
65 70 75 80

Pro Gln Cys Glu Met Asp Gln Val Ser Gly Ile Gln Lys Leu Glu Cys
85 90 95

Glu Ala Asp Gln Lys Trp Glu Tyr Glu Val Asp Gln Val Ser Gly Val
100 105 110

Gln Lys Pro Gln Cys Glu Met Asp Gln Val Ser Gly Ile Gln Lys Leu
115 120 125

Glu Cys Glu Ala Asp Gln Glu Cys Glu Thr Tyr Glu Lys Cys Cys Pro
130 135 140

Asn Val Cys Gly Thr Lys Ser Cys Val Ala Ala Arg Tyr Met Asp Val
145 150 155 160

Lys Gly Lys Lys Gly Pro Val Gly Met Pro Lys Glu Ala Thr Cys Asp
165 170 175

His Phe Met Cys Leu Gln Gln Gly Ser Glu Cys Asp Ile Trp Asp Gly
180 185 190

Gln Pro Val Cys Lys Cys Lys Asp Arg Cys Glu Lys Glu Pro Ser Phe
195 200 205

Thr Cys Ala Ser Asp Gly Leu Thr Tyr Tyr Asn Arg Cys Tyr Met Asp
210 215 220

Ala Glu Ala Cys Ser Lys Gly Ile Thr Leu Ala Val Val Thr Cys Arg
225 230 235 240

Tyr His Phe Thr Trp Pro Asn Thr Ser Pro Pro Pro Pro Glu Thr Thr

245 250 255

Met His Pro Thr Thr Ala Ser Pro Glu Thr Pro Glu Leu Asp Met Ala

260 265 270

Ala Pro Ala Leu Leu Asn Asn Pro Val His Gln Ser Val Thr Met Gly

275 280 285

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

290 295 300

Ile Thr Trp Glu Lys Gln Leu Glu Asp Arg Glu Asn Val Val Met Arg

305 310 315 320

Pro Asn His Val Arg Gly Asn Val Val Val Thr Asn Ile Ala Gln Leu

325 330 335

Val Ile Tyr Asn Ala Gln Leu Gln Asp Ala Gly Ile Tyr Thr Cys Thr

340 345 350

Ala Arg Asn Val Ala Gly Val Leu Arg Ala Asp Phe Pro Leu Ser Val

355 360 365

Val Arg Gly His Gln Ala Ala Ala Thr Ser Glu Ser Ser Pro Asn Gly

370 375 380

Thr Ala Phe Pro Ala Ala Glu Cys Leu Lys Pro Pro Asp Ser Glu Asp

385 390 395 400

Cys Gly Glu Glu Gln Thr Arg Trp His Phe Asp Ala Gln Ala Asn Asn

405 410 415

```
Cys Leu Thr Phe Thr Phe Gly His Cys His Arg Asn Leu Asn His Phe
            420                 425                 430

Glu Thr Tyr Glu Ala Cys Met Leu Ala Cys Met Ser Gly Pro Leu Ala
        435                 440                 445

Ala Cys Ser Leu Pro Ala Leu Gln Gly Pro Cys Lys Ala Tyr Ala Pro
    450                 455                 460

Arg Trp Ala Tyr Asn Ser Gln Thr Gly Gln Cys Gln Ser Phe Val Tyr
465                 470                 475                 480

Gly Gly Cys Glu Gly Asn Gly Asn Asn Phe Glu Ser Arg Glu Ala Cys
                485                 490                 495

Glu Glu Ser Cys Pro Phe Pro Arg Gly Asn Gln Arg Cys Arg Ala Cys
            500                 505                 510

Lys Pro Arg Gln Lys Leu Val Thr Ser Phe Cys Arg Ser Asp Phe Val
        515                 520                 525

Ile Leu Gly Arg Val Ser Glu Leu Thr Glu Glu Pro Asp Ser Gly Arg
    530                 535                 540

Ala Leu Val Thr Val Asp Glu Val Leu Lys Asp Glu Lys Met Gly Leu
545                 550                 555                 560

Lys Phe Leu Gly Gln Glu Pro Leu Glu Val Thr Leu Leu His Val Asp
                565                 570                 575

Trp Ala Cys Pro Cys Pro Asn Val Thr Val Ser Glu Met Pro Leu Ile
            580                 585                 590

Ile Met Gly Glu Val Asp Gly Gly Met Ala Met Leu Arg Pro Asp Ser
```

595 600 605

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

610 615 620

Met His Lys Lys Thr Cys Asp Val Leu Lys Glu Phe Leu Gly Leu His

625 630 635 640

<210> 6

<211> 1731

<212> DNA

<213> Homo sapiens

<400> 6

atgtgggccc caaggtgtcg ccggttctgg tctcgctggg agcaggtggc agcgctgctg 60

ctgctgctgc tactgctcgg ggtgccccg cgaagcctgg cgctgccgcc catccgctat 120

tcccacgccg gcatctgccc caacgacatg aatcccaacc tctgggtgga cgcacagagc 180

acctgcaggc gggagtgtga gacggaccag gagtgtgaga cctatgagaa gtgctgcccc 240

aacgtatgtg ggaccaagag ctgcgtggcg gcccgctaca tggacgtgaa agggaagaag 300

ggcccagtgg gcatgcccaa ggaggccaca tgtgaccact tcatgtgtct gcagcagggc 360

tctgagtgtg acatctggga tggccagccc gtgtgtaagt gcaaagaccg ctgtgagaag 420

gagcccagct ttacctgcgc ctcggacggc ctcacctact ataaccgctg ctacatggat 480

gccgaggcct gctccaaagg catcacactg gccgttgtaa cctgccgcta tcacttcacc 540

tggcccaaca ccagcccccc accacctgag accaccatgc accccaccac agcctcccca 600 gagacccctg agctggacat ggcggcccct gcgctgctca acaaccctgt gcaccagtcg 660

gtcaccatgg gtgagacagt gagtttcctc tgtgatgtgg tgggccggcc ccggcctgag 720

atcacctggg agaagcagtt ggaggatcgg gagaatgtgg tcatgcggcc caaccatgtg 780

cgtggcaacg tggtggtcac caacattgcc cagctggtca tctataacgc ccagctgcag 840

gatgctggga tctacacctg cacggcccgg aacgtggctg gggtcctgag ggctgatttc 900

ccgctgtcgg tggtcagggg tcatcaggct gcagccacct cagagagcag ccccaatggc 960

acggctttcc cggcggccga gtgcctgaag cccccagaca gtgaggactg tggcgaagag 1020

cagacccgct ggcacttcga tgcccaggcc aacaactgcc tgaccttcac cttcggccac 1080

tgccaccgta acctcaacca ctttgagacc tatgaggcct gcatgctggc ctgcatgagc 1140

gggccgctgg ccgcgtgcag cctgcccgcc ctgcaggggc cctgcaaagc ctacgcgcct 1200

cgctgggctt acaacagcca gacgggccag tgccagtcct ttgtctatgg tggctgcgag 1260

ggcaatggca acaactttga gagccgtgag gcctgtgagg agtcgtgccc cttccccagg 1320

gggaaccagc gctgtcgggc ctgcaagcct cggcagaagc tcgttaccag cttctgtcgc 1380

agcgactttg tcatcctggg ccgagtctct gagctgaccg aggagcctga ctcgggccgc 1440

gccctggtga ctgtggatga ggtcctaaag gatgagaaaa tgggcctcaa gttcctgggc 1500

caggagccat tggaggtcac tctgcttcac gtggactggg catgcccctg ccccaacgtg 1560

accgtgagcg agatgccgct catcatcatg gggaggtgg acggcggcat ggccatgctg 1620 cgccccgata gctttgtggg cgcatcgagt gcccgccggg tcaggaagct tcgtgaggtc 1680

atgcacaaga agacctgtga cgtcctcaag gagtttcttg gcttgcactg a 1731

<210> 7

<211> 576

<212> PRT

<213> Homo sapiens

<400> 7

Met Trp Ala Pro Arg Cys Arg Arg Phe Trp Ser Arg Trp Glu Gln Val

1 5 10 15

Ala Ala Leu Leu Leu Leu Leu Leu Leu Leu Gly Val Pro Pro Arg Ser

20 25 30

Leu Ala Leu Pro Pro Ile Arg Tyr Ser His Ala Gly Ile Cys Pro Asn

35 40 45

Asp Met Asn Pro Asn Leu Trp Val Asp Ala Gln Ser Thr Cys Arg Arg

50 55 60

Glu Cys Glu Thr Asp Gln Glu Cys Glu Thr Tyr Glu Lys Cys Cys Pro

65 70 75 80

Asn Val Cys Gly Thr Lys Ser Cys Val Ala Ala Arg Tyr Met Asp Val

85 90 95

Lys Gly Lys Lys Gly Pro Val Gly Met Pro Lys Glu Ala Thr Cys Asp

100 105 110

His Phe Met Cys Leu Gln Gln Gly Ser Glu Cys Asp Ile Trp Asp Gly

115 120 125

Gln Pro Val Cys Lys Cys Lys Asp Arg Cys Glu Lys Glu Pro Ser Phe

130 135 140

Thr Cys Ala Ser Asp Gly Leu Thr Tyr Tyr Asn Arg Cys Tyr Met Asp

145 150 155 160

Ala Glu Ala Cys Ser Lys Gly Ile Thr Leu Ala Val Val Thr Cys Arg

165 170 175

Tyr His Phe Thr Trp Pro Asn Thr Ser Pro Pro Pro Pro Glu Thr Thr

180 185 190

Met His Pro Thr Thr Ala Ser Pro Glu Thr Pro Glu Leu Asp Met Ala

195 200 205

Ala Pro Ala Leu Leu Asn Asn Pro Val His Gln Ser Val Thr Met Gly

210 215 220

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

225 230 235 240

Ile Thr Trp Glu Lys Gln Leu Glu Asp Arg Glu Asn Val Val Met Arg

245 250 255

Pro Asn His Val Arg Gly Asn Val Val Val Thr Asn Ile Ala Gln Leu

260 265 270

Val Ile Tyr Asn Ala Gln Leu Gln Asp Ala Gly Ile Tyr Thr Cys Thr

275 280 285

Ala Arg Asn Val Ala Gly Val Leu Arg Ala Asp Phe Pro Leu Ser Val

290 295 300

```
Val Arg Gly His Gln Ala Ala Ala Thr Ser Glu Ser Ser Pro Asn Gly
305                 310                 315                 320

Thr Ala Phe Pro Ala Ala Glu Cys Leu Lys Pro Pro Asp Ser Glu Asp
                325                 330                 335

Cys Gly Glu Glu Gln Thr Arg Trp His Phe Asp Ala Gln Ala Asn Asn
            340                 345                 350

Cys Leu Thr Phe Thr Phe Gly His Cys His Arg Asn Leu Asn His Phe
        355                 360                 365

Glu Thr Tyr Glu Ala Cys Met Leu Ala Cys Met Ser Gly Pro Leu Ala
    370                 375                 380

Ala Cys Ser Leu Pro Ala Leu Gln Gly Pro Cys Lys Ala Tyr Ala Pro
385                 390                 395                 400

Arg Trp Ala Tyr Asn Ser Gln Thr Gly Gln Cys Gln Ser Phe Val Tyr
                405                 410                 415

Gly Gly Cys Glu Gly Asn Gly Asn Asn Phe Glu Ser Arg Glu Ala Cys
            420                 425                 430

Glu Glu Ser Cys Pro Phe Pro Arg Gly Asn Gln Arg Cys Arg Ala Cys
        435                 440                 445

Lys Pro Arg Gln Lys Leu Val Thr Ser Phe Cys Arg Ser Asp Phe Val
    450                 455                 460

Ile Leu Gly Arg Val Ser Glu Leu Thr Glu Glu Pro Asp Ser Gly Arg
465                 470                 475                 480
```

```
Ala Leu Val Thr Val Asp Glu Val Leu Lys Asp Glu Lys Met Gly Leu
                485                 490                 495

Lys Phe Leu Gly Gln Glu Pro Leu Glu Val Thr Leu Leu His Val Asp
            500                 505                 510

Trp Ala Cys Pro Cys Pro Asn Val Thr Val Ser Glu Met Pro Leu Ile
        515                 520                 525

Ile Met Gly Glu Val Asp Gly Gly Met Ala Met Leu Arg Pro Asp Ser
    530                 535                 540

Phe Val Gly Ala Ser Ser Ala Arg Arg Val Arg Lys Leu Arg Glu Val
545                 550                 555                 560

Met His Lys Lys Thr Cys Asp Val Leu Lys Glu Phe Leu Gly Leu His
                565                 570                 575
```

<210> 8

<211> 1659

<212> DNA

<213> Mus sp.

<400> 8

```
atgcctgccc cacagccatt cctgcctctg ctctttgtct tcgtgctcat ccatctgacc         60

tcggagacca acctgctgcc agatcccgga agccatcctg gcatgtgccc caacgagctc        120

agccccccacc tgtgggtcga cgcccagagc acctgtgagc gtgagtgtac cggggaccag        180

gactgtgcgg catccgagaa gtgctgcacc aatgtgtgtg ggctgcagag ctgcgtggct        240
```

gcccgctttc ccagtggtgg cccagctgta cctgagacag cagcctcctg tgaaggcttc 300

caatgcccac aacagggttc tgactgtgac atctgggatg ggcagccagt ttgtcgctgc 360

cgtgaccgct gtgaaaaaga acccagcttc acatgtgctt ctgatggcct tacctattac 420

aaccgctgct acatggacgc agaagcctgc ctgcggggtc tccacctgca cgttgtaccc 480

tgtaagcaca ttctcagttg gccgcccagc agcccgggac cacccgagac cactgctcgc 540

ccaaccccctg gggctgctcc catgccacct gccctgtaca acagcccctc accacaggca 600

gtgcatgttg gggggacagc cagcctccac tgtgatgtta gtggccgtcc accacctgct 660

gtgacctggg agaagcagag ccatcagcgg gagaacctga tcatgcgccc tgaccaaatg 720

tatggcaacg tggttgtcac cagtatcgga cagctagtcc tctacaatgc tcagttggag 780

gatgcgggcc tgtatacctg cactgcacga aacgctgccg gcctgctgcg ggccgacttt 840

cccctttccg ttttacagcg ggcaactact caggacaggg acccaggtat cccagccttg 900

gctgagtgcc aggccgacac acaagcctgt gttgggccac ctactcccca tcatgtcctt 960

tggcgctttg acccacagag aggcagctgc atgacattcc cagccctcag atgtgatggg 1020

gctgcccggg gctttgagac ctatgaggca tgccagcagg cctgtgttcg tggccccggg 1080

gatgtctgtg cactgcctgc agttcagggg ccctgccagg gctgggagcc acgctgggcc 1140

tacagcccac tgctacagca gtgccacccc tttgtataca gtggctgtga aggaaacagc 1200

aataactttg agacccggga gagctgtgag gatgcttgcc ctgtaccacg cacaccaccc 1260 tgtcgtgcct gccgcctcaa gagcaagctg gctctgagct tgtgccgcag tgactttgcc 1320

atcgtgggga gactcacaga ggtcctggag gagcccgagg ctgcaggcgg catagctcgt 1380

gtggccttgg atgatgtgct aaaggacgac aagatgggcc tcaagttctt gggcaccaaa 1440

tacctggagg tgacattgag tggcatggac tgggcctgcc catgccccaa cgtgacagct 1500

gtcgatgggc cactggtcat catgggtgag gttcgtgaag gtgtggctgt gttggacgcc 1560

aacagctatg tccgtgctgc cagcgagaag cgagtcaaga agattgtgga actgctcgag 1620

aagaaggctt gtgaactgct caaccgcttc caagactag 1659

<210> 9

<211> 552

<212> PRT

<213> Mus sp.

<400> 9

Met Pro Ala Pro Gln Pro Phe Leu Pro Leu Leu Phe Val Phe Val Leu

1 5 10 15

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

20 25 30

Pro Gly Met Cys Pro Asn Glu Leu Ser Pro His Leu Trp Val Asp Ala

35 40 45

Gln Ser Thr Cys Glu Arg Glu Cys Thr Gly Asp Gln Asp Cys Ala Ala

50 55 60

```
Ser Glu Lys Cys Cys Thr Asn Val Cys Gly Leu Gln Ser Cys Val Ala
 65                  70                  75                  80

Ala Arg Phe Pro Ser Gly Gly Pro Ala Val Pro Glu Thr Ala Ala Ser
                 85                  90                  95

Cys Glu Gly Phe Gln Cys Pro Gln Gln Gly Ser Asp Cys Asp Ile Trp
            100                 105                 110

Asp Gly Gln Pro Val Cys Arg Cys Arg Asp Arg Cys Glu Lys Glu Pro
        115                 120                 125

Ser Phe Thr Cys Ala Ser Asp Gly Leu Thr Tyr Tyr Asn Arg Cys Tyr
    130                 135                 140

Met Asp Ala Glu Ala Cys Leu Arg Gly Leu His Leu His Val Val Pro
145                 150                 155                 160

Cys Lys His Ile Leu Ser Trp Pro Pro Ser Ser Pro Gly Pro Pro Glu
                165                 170                 175

Thr Thr Ala Arg Pro Thr Pro Gly Ala Ala Pro Met Pro Pro Ala Leu
            180                 185                 190

Tyr Asn Ser Pro Ser Pro Gln Ala Val His Val Gly Gly Thr Ala Ser
        195                 200                 205

Leu His Cys Asp Val Ser Gly Arg Pro Pro Pro Ala Val Thr Trp Glu
    210                 215                 220

Lys Gln Ser His Gln Arg Glu Asn Leu Ile Met Arg Pro Asp Gln Met
225                 230                 235                 240

Tyr Gly Asn Val Val Val Thr Ser Ile Gly Gln Leu Val Leu Tyr Asn
```

245 250 255

Ala Gln Leu Glu Asp Ala Gly Leu Tyr Thr Cys Thr Ala Arg Asn Ala

260 265 270

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

275 280 285

Thr Thr Gln Asp Arg Asp Pro Gly Ile Pro Ala Leu Ala Glu Cys Gln

290 295 300

Ala Asp Thr Gln Ala Cys Val Gly Pro Pro Thr Pro His His Val Leu

305 310 315 320

Trp Arg Phe Asp Pro Gln Arg Gly Ser Cys Met Thr Phe Pro Ala Leu

325 330 335

Arg Cys Asp Gly Ala Ala Arg Gly Phe Glu Thr Tyr Glu Ala Cys Gln

340 345 350

Gln Ala Cys Val Arg Gly Pro Gly Asp Val Cys Ala Leu Pro Ala Val

355 360 365

Gln Gly Pro Cys Gln Gly Trp Glu Pro Arg Trp Ala Tyr Ser Pro Leu

370 375 380

Leu Gln Gln Cys His Pro Phe Val Tyr Ser Gly Cys Glu Gly Asn Ser

385 390 395 400

Asn Asn Phe Glu Thr Arg Glu Ser Cys Glu Asp Ala Cys Pro Val Pro

405 410 415

Arg Thr Pro Pro Cys Arg Ala Cys Arg Leu Lys Ser Lys Leu Ala Leu

420 425 430

Ser Leu Cys Arg Ser Asp Phe Ala Ile Val Gly Arg Leu Thr Glu Val
435 440 445

Leu Glu Glu Pro Glu Ala Ala Gly Gly Ile Ala Arg Val Ala Leu Asp
450 455 460

Asp Val Leu Lys Asp Asp Lys Met Gly Leu Lys Phe Leu Gly Thr Lys
465 470 475 480

Tyr Leu Glu Val Thr Leu Ser Gly Met Asp Trp Ala Cys Pro Cys Pro
485 490 495

Asn Val Thr Ala Val Asp Gly Pro Leu Val Ile Met Gly Glu Val Arg
500 505 510

Glu Gly Val Ala Val Leu Asp Ala Asn Ser Tyr Val Arg Ala Ala Ser
515 520 525

Glu Lys Arg Val Lys Lys Ile Val Glu Leu Leu Glu Lys Lys Ala Cys
530 535 540

Glu Leu Leu Asn Arg Phe Gln Asp
545 550

<210> 10
<211> 1695
<212> DNA
<213> Homo sapiens

<400> 10
atgcccgccc tacgtccact cctgccgctc ctgctcctcc tccggctgac ctcgggggct 60

ggcttgctgc cagggctggg gagccacccg ggcgtgtgcc ccaaccagct cagccccaac 120

ctgtgggtgg acgcccagag cacctgtgag cgcgagtgta gcagggacca ggactgtgcg 180

gctgctgaga agtgctgcat caacgtgtgt ggactgcaca gctgcgtggc agcacgcttc 240

cccggcagcc cagctgcgcc gacgacagcg gcctcctgcg agggctttgt gtgcccacag 300

cagggctcgg actgcgacat ctgggacggg cagcccgtgt gccgctgccg cgaccgctgt 360

gagaaggagc ccagcttcac ctgcgcctcg gacggcctca cctactacaa ccgctgctat 420

atggacgccg aggcctgcct gcggggcctg cacctccaca tcgtgccctg caagcacgtg 480

ctcagctggc cgcccagcag cccggggccg ccggagacca ctgcccgccc cacacctggg 540

gccgcgcccg tgcctcctgc cctgtacagc agcccctccc cacaggcggt gcaggttggg 600

ggtacggcca gcctccactg cgacgtcagc ggccgcccgc cgcctgctgt gacctgggag 660

aagcagagtc accagcgaga gaacctgatc atgcgccctg atcagatgta tggcaacgtg 720

gtggtcacca gcatcgggca gctggtgctc tacaacgcgc ggcccgaaga cgccggcctg 780

tacacctgca ccgcgcgcaa cgctgctggg ctgctgcggg ctgacttccc actctctgtg 840

gtccagcgag agccggccag ggacgcagcc cccagcatcc cagccccggc cgagtgcctg 900

ccggatgtgc aggcctgcac gggccccact tccccacacc ttgtcctctg gcactacgac 960

ccgcagcggg gcggctgcat gaccttcccg gcccgtggct gtgatggggc ggcccgcggc 1020

tttgagacct acgaggcatg ccagcaggcc tgtgcccgcg gccccggcga cgcctgcgtg 1080 ctgcctgccg tgcagggccc ctgccggggc tgggagccgc gctgggccta cagcccgctg 1140

ctgcagcagt gccatccctt cgtgtacggt ggctgcgagg gcaacggcaa caacttccac 1200

agccgcgaga gctgcgagga tgcctgcccc gtgccgcgca caccgccctg ccgcgcctgc 1260

cgcctccgga gcaagctggc gctgagcctg tgccgcagcg acttcgccat cgtggggcgg 1320

ctcacggagg tgctggagga gcccgaggcc gccggcggca tcgcccgcgt ggcgctcgag 1380

gacgtgctca aggatgacaa gatgggcctc aagttcttgg gcaccaagta cctggaggtg 1440

acgctgagtg gcatggactg ggcctgcccc tgccccaaca tgacggcggg cgacgggccg 1500

ctggtcatca tgggtgaggt gcgcgatggc gtggccgtgc tggacgccgg cagctacgtc 1560

cgcgccgcca gcgagaagcg cgtcaagaag atcttggagc tgctggagaa gcaggcctgc 1620

gagctgctca accgcttcca ggactagccc ccgcaggggc ctgcgccacc ccgtcctggt 1680

gaataaacgc actcc 1695

<210> 11

<211> 548

<212> PRT

<213> Homo sapiens

<400> 11

Met Pro Ala Leu Arg Pro Leu Leu Pro Leu Leu Leu Leu Leu Arg Leu
1 5 10 15

Thr Ser Gly Ala Gly Leu Leu Pro Gly Leu Gly Ser His Pro Gly Val
20 25 30

Cys Pro Asn Gln Leu Ser Pro Asn Leu Trp Val Asp Ala Gln Ser Thr

35 40 45

Cys Glu Arg Glu Cys Ser Arg Asp Gln Asp Cys Ala Ala Ala Glu Lys

50 55 60

Cys Cys Ile Asn Val Cys Gly Leu His Ser Cys Val Ala Ala Arg Phe

65 70 75 80

Pro Gly Ser Pro Ala Ala Pro Thr Thr Ala Ala Ser Cys Glu Gly Phe

85 90 95

Val Cys Pro Gln Gln Gly Ser Asp Cys Asp Ile Trp Asp Gly Gln Pro

100 105 110

Val Cys Arg Cys Arg Asp Arg Cys Glu Lys Glu Pro Ser Phe Thr Cys

115 120 125

Ala Ser Asp Gly Leu Thr Tyr Tyr Asn Arg Cys Tyr Met Asp Ala Glu

130 135 140

Ala Cys Leu Arg Gly Leu His Leu His Ile Val Pro Cys Lys His Val

145 150 155 160

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

165 170 175

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

180 185 190

Ser Pro Gln Ala Val Gln Val Gly Gly Thr Ala Ser Leu His Cys Asp

195 200 205

Val Ser Gly Arg Pro Pro Pro Ala Val Thr Trp Glu Lys Gln Ser His

210 215 220

Gln Arg Glu Asn Leu Ile Met Arg Pro Asp Gln Met Tyr Gly Asn Val

225 230 235 240

Val Val Thr Ser Ile Gly Gln Leu Val Leu Tyr Asn Ala Arg Pro Glu

245 250 255

Asp Ala Gly Leu Tyr Thr Cys Thr Ala Arg Asn Ala Ala Gly Leu Leu

260 265 270

Arg Ala Asp Phe Pro Leu Ser Val Val Gln Arg Glu Pro Ala Arg Asp

275 280 285

Ala Ala Pro Ser Ile Pro Ala Pro Ala Glu Cys Leu Pro Asp Val Gln

290 295 300

Ala Cys Thr Gly Pro Thr Ser Pro His Leu Val Leu Trp His Tyr Asp

305 310 315 320

Pro Gln Arg Gly Gly Cys Met Thr Phe Pro Ala Arg Gly Cys Asp Gly

325 330 335

Ala Ala Arg Gly Phe Glu Thr Tyr Glu Ala Cys Gln Gln Ala Cys Ala

340 345 350

Arg Gly Pro Gly Asp Ala Cys Val Leu Pro Ala Val Gln Gly Pro Cys

355 360 365

Arg Gly Trp Glu Pro Arg Trp Ala Tyr Ser Pro Leu Leu Gln Gln Cys

370 375 380

His Pro Phe Val Tyr Gly Gly Cys Glu Gly Asn Gly Asn Asn Phe His

385 390 395 400

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

405 410 415

Cys Arg Ala Cys Arg Leu Arg Ser Lys Leu Ala Leu Ser Leu Cys Arg

420 425 430

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

435 440 445

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

450 455 460

Asp Asp Lys Met Gly Leu Lys Phe Leu Gly Thr Lys Tyr Leu Glu Val

465 470 475 480

Thr Leu Ser Gly Met Asp Trp Ala Cys Pro Cys Pro Asn Met Thr Ala

485 490 495

Gly Asp Gly Pro Leu Val Ile Met Gly Glu Val Arg Asp Gly Val Ala

500 505 510

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

515 520 525

Lys Lys Ile Leu Glu Leu Leu Glu Lys Gln Ala Cys Glu Leu Leu Asn

530 535 540

Arg Phe Gln Asp

545

<210> 12

<211> 30

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Illustrative competing
peptide

<400> 12

```
Asp Phe Gly Leu Asp Ser Asp Glu His Ser Thr Glu Ser Arg Ser Ser
  1               5                  10                  15

Arg Tyr Pro Leu Thr Val Asp Phe Glu Ala Phe Gly Trp Asp
             20                  25                  30
```

<210> 13

<211> 15

<212> PRT

<213> Mus sp.

<400> 13

```
Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys
  1               5                  10                  15
```

<210> 14

<211> 12

<212> PRT

<213> Mus sp.

<400> 14

```
Met Ser Pro Ile Asn Met Leu Tyr Phe Asn Gly Lys
```

1 5 10

<210> 15

<211> 14

<212> PRT

<213> Mus sp.

<400> 15

Asp Phe Gly Leu Asp Cys Asp Glu His Ser Thr Glu Ser Arg

1 5 10

<210> 16

<211> 15

<212> PRT

<213> Mus sp.

<400> 16

Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys

1 5 10 15

<210> 17

<211> 12

<212> PRT

<213> Mus sp.

<400> 17

Met Ser Pro Ile Asn Met Leu Tyr Phe Asn Gly Lys

1 5 10

<210> 18

<211> 19

<212> PRT

<213> Mus sp.

<400> 18

Tyr Pro Leu Thr Val Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile

1 5 10 15

Ala Pro Lys

<210> 19

<211> 12

<212> PRT

<213> Mus sp.

<400> 19

Met Ser Pro Ile Asn Met Leu Tyr Phe Asn Gly Lys

1 5 10

<210> 20

<211> 11

<212> PRT

<213> Mus sp.

<400> 20

Gly Ser Ala Gly Pro Cys Cys Thr Pro Thr Lys

1 5 10

<210> 21

<211> 16

<212> PRT

<213> Mus sp.

<400> 21

Leu Asp Met Ser Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val Lys

1 5 10 15

<210> 22

<211> 28

<212> PRT

<213> Mus sp.

<400> 22

Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr Phe Pro Gly Pro

1 5 10 15

Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Lys

20 25

<210> 23

<211> 16

<212> PRT

<213> Mus sp.

<400> 23

Leu Asp Met Ser Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val Lys

1 5 10 15

<210> 24

<211> 10

<212> PRT

<213> Mus sp.

<400> 24

Glu Leu Ile Asp Gln Tyr Asp Val Gln Arg

1 5 10

<210> 25

<211> 11

<212> PRT

<213> Mus sp.

<400> 25

Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg

1 5 10

<210> 26

<211> 11

<212> PRT

<213> Mus sp.

<400> 26

Ala Gln Leu Trp Ile Tyr Leu Arg Pro Val Lys

1 5 10

<210> 27

<211> 10

<212> PRT

<213> Mus sp.

<400> 27

Glu Gly Leu Cys Asn Ala Cys Ala Trp Arg

1 5 10

<210> 28

<211> 18

<212> PRT

<213> Mus sp.

<400> 28

Pro Gln Ser Cys Leu Val Asp Gln Thr Gly Ser Ala His Cys Val Val

1 5 10 15

Cys Arg

<210> 29

<211> 12

<212> PRT

<213> Mus sp.

<400> 29

Asp Ser Cys Asp Gly Val Glu Cys Gly Pro Gly Lys

1 5 10

<210> 30

<211> 9

<212> PRT

<213> Mus sp.

<400> 30

Ser Cys Ala Gln Val Val Cys Pro Arg

1 5

<210> 31

<211> 13

<212> PRT

<213> Mus sp.

<400> 31

Glu Cys Glu Thr Asp Gln Glu Cys Glu Thr Tyr Glu Lys

1 5 10

<210> 32

<211> 9

<212> PRT

<213> Mus sp.

<400> 32

Ala Asp Phe Pro Leu Ser Val Val Arg

1 5

<210> 33

<211> 11

<212> PRT

<213> Mus sp.

<400> 33

Glu Ala Cys Glu Glu Ser Cys Pro Phe Pro Arg

1 5 10

<210> 34

<211> 8

<212> PRT

<213> Mus sp.

<400> 34

Ser Asp Phe Val Ile Leu Gly Arg

1 5

<210> 35

<211> 12

<212> PRT

<213> Mus sp.

<400> 35

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

1 5 10

<210> 36

<211> 15

<212> PRT

<213> Homo sapiens

<400> 36

Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys

1 5 10 15

<210> 37

<211> 14

<212> PRT

<213> Homo sapiens

<400> 37

Asp Phe Gly Leu Asp Cys Asp Glu His Ser Thr Glu Ser Arg

1 5 10

<210> 38

<211> 28

<212> PRT

<213> Homo sapiens

<400> 38

Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr Phe Pro Gly Pro

1 5 10 15

Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Lys

20 25

<210> 39

<211> 10

<212> PRT

<213> Homo sapiens

<400> 39

Glu Leu Ile Asp Gln Tyr Asp Val Gln Arg

1 5 10

<210> 40

<211> 18

<212> PRT

<213> Homo sapiens

<400> 40

Pro Gln Ser Cys Val Val Asp Gln Thr Gly Ser Ala His Cys Val Val

1 5 10 15

Cys Arg

<210> 41

<211> 13

<212> PRT

<213> Homo sapiens

<400> 41

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

1 5 10

<210> 42

<211> 12

<212> PRT

<213> Homo sapiens

<400> 42

Leu Gln Val Cys Gly Ser Asp Gly Ala Thr Tyr Arg

1 5 10

<210> 43

<211> 12

<212> PRT

<213> Homo sapiens

<400> 43

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

1 5 10

<210> 44

<211> 10

<212> PRT

<213> Homo sapiens

<400> 44

Cys Tyr Met Asp Ala Glu Ala Cys Ser Lys

1 5 10

<210> 45

<211> 10

<212> PRT

<213> Homo sapiens

<400> 45

Gly Ile Thr Leu Ala Val Val Thr Cys Arg

1 5 10

<210> 46

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 46

ttggccactg ccaccacaat ctcaaccact t 31

<210> 47

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 47

tctcagcatg gccatgccgc cgtcga 26

<210> 48

<211> 1716

<212> DNA

<213> Mus sp.

<220>

<221> CDS

<222> (1)..(1713)

<400> 48

```
atg tgt gcc cca ggg tat cat cgg ttc tgg ttt cac tgg ggg ctg ctg          48
Met Cys Ala Pro Gly Tyr His Arg Phe Trp Phe His Trp Gly Leu Leu
  1               5                  10                  15

ttg ctg ctg ctc ctc gag gct ccc ctt cga ggc cta gca ctg cca ccc          96
Leu Leu Leu Leu Leu Glu Ala Pro Leu Arg Gly Leu Ala Leu Pro Pro
```

20 25 30

atc cga tac tcc cat gcg ggc atc tgc ccc aac gac atg aac ccc aac 144

Ile Arg Tyr Ser His Ala Gly Ile Cys Pro Asn Asp Met Asn Pro Asn

35 40 45

ctc tgg gtg gat gcc cag agc acc tgc aag cga gag tgt gaa aca gac 192

Leu Trp Val Asp Ala Gln Ser Thr Cys Lys Arg Glu Cys Glu Thr Asp

50 55 60

cag gaa tgt gag acc tat gag aaa tgc tgc ccc aat gtg tgt ggg acc 240

Gln Glu Cys Glu Thr Tyr Glu Lys Cys Cys Pro Asn Val Cys Gly Thr

65 70 75 80

aag agc tgt gtg gca gcc cgc tac atg gat gtg aaa ggg aag aag ggg 288

Lys Ser Cys Val Ala Ala Arg Tyr Met Asp Val Lys Gly Lys Lys Gly

85 90 95

cct gta ggc atg ccc aag gag gcc aca tgt gac cat ttc atg tgc ctg 336

Pro Val Gly Met Pro Lys Glu Ala Thr Cys Asp His Phe Met Cys Leu

100 105 110

cag cag ggc tct gag tgt gac atc tgg gac ggc cag ccc gtg tgt aag 384

Gln Gln Gly Ser Glu Cys Asp Ile Trp Asp Gly Gln Pro Val Cys Lys

115 120 125

tgc aaa gat cgc tgt gag aag gag ccc agc ttc acc tgt gcc tct gat 432

Cys Lys Asp Arg Cys Glu Lys Glu Pro Ser Phe Thr Cys Ala Ser Asp

130 135 140

ggc ctt acc tac tac aac cgt tgc ttc atg gac gcc gaa gcc tgc tcc 480

Gly Leu Thr Tyr Tyr Asn Arg Cys Phe Met Asp Ala Glu Ala Cys Ser

145 150 155 160

aag ggc atc aca ctg tct gtg gtc acc tgt cgt tat cac ttc acc tgg 528
Lys Gly Ile Thr Leu Ser Val Val Thr Cys Arg Tyr His Phe Thr Trp
165 170 175

cct aac acc agc cct cca ccg cct gag acc acg gtg cat ccc acc acc 576
Pro Asn Thr Ser Pro Pro Pro Pro Glu Thr Thr Val His Pro Thr Thr
180 185 190

gcc tct ccg gag act ctc ggg ctg gac atg gca gcc cca gcc ctg ctc 624
Ala Ser Pro Glu Thr Leu Gly Leu Asp Met Ala Ala Pro Ala Leu Leu
195 200 205

aac cac cct gtc cat cag tca gtc acc gtg ggt gag act gtg agt ttc 672
Asn His Pro Val His Gln Ser Val Thr Val Gly Glu Thr Val Ser Phe
210 215 220

ctc tgt gac gtg gta ggc cgg cct cgg cca gag ctc act tgg gag aaa 720
Leu Cys Asp Val Val Gly Arg Pro Arg Pro Glu Leu Thr Trp Glu Lys
225 230 235 240

cag ctg gag gac cga gag aat gtt gtc atg agg ccc aac cac gtg cgt 768
Gln Leu Glu Asp Arg Glu Asn Val Val Met Arg Pro Asn His Val Arg
245 250 255

ggt aat gtg gtg gtc act aac att gcc cag ctg gtc atc tac aac gtc 816
Gly Asn Val Val Val Thr Asn Ile Ala Gln Leu Val Ile Tyr Asn Val
260 265 270

cag ccc cag gat gct ggc ata tac acc tgt aca gct cga aat gtc gct 864
Gln Pro Gln Asp Ala Gly Ile Tyr Thr Cys Thr Ala Arg Asn Val Ala
275 280 285

ggt gtc ctg agg gct gac ttc ccg ttg tcg gtg gtc agg ggt ggt cag 912
Gly Val Leu Arg Ala Asp Phe Pro Leu Ser Val Val Arg Gly Gly Gln

```
    290                 295                 300

gcc agg gcc act tca gag agc agt ctc aat ggc aca gct ttt cca gca          960
Ala Arg Ala Thr Ser Glu Ser Ser Leu Asn Gly Thr Ala Phe Pro Ala
305                 310                 315                 320

aca gag tgc ctg aag ccc cca gac agt gag gac tgt gga gag gag cag         1008
Thr Glu Cys Leu Lys Pro Pro Asp Ser Glu Asp Cys Gly Glu Glu Gln
                325                 330                 335

aca cgc tgg cac ttc gac gcc cag gct aac aac tgc ctc act ttc acc         1056
Thr Arg Trp His Phe Asp Ala Gln Ala Asn Asn Cys Leu Thr Phe Thr
            340                 345                 350

ttt ggc cac tgc cac cac aat ctc aac cac ttt gag acc tac gag gcc         1104
Phe Gly His Cys His His Asn Leu Asn His Phe Glu Thr Tyr Glu Ala
        355                 360                 365

tgt atg ctg gct tgt atg agt ggg cca ttg gcc acc tgc agc ctg cct         1152
Cys Met Leu Ala Cys Met Ser Gly Pro Leu Ala Thr Cys Ser Leu Pro
    370                 375                 380

gcc ctg caa ggg cct tgc aaa gct tat gtc cca cgc tgg gcc tac aac         1200
Ala Leu Gln Gly Pro Cys Lys Ala Tyr Val Pro Arg Trp Ala Tyr Asn
385                 390                 395                 400

agc cag aca ggc cta tgc cag tcc ttc gtc tat ggc ggc tgt gag ggc         1248
Ser Gln Thr Gly Leu Cys Gln Ser Phe Val Tyr Gly Gly Cys Glu Gly
                405                 410                 415

aac ggt aac aac ttt gaa agc cgt gag gct tgt gag gag tcg tgt ccc         1296
Asn Gly Asn Asn Phe Glu Ser Arg Glu Ala Cys Glu Glu Ser Cys Pro
            420                 425                 430
```

ttc ccg agg ggt aac cag cac tgc cgg gcc tgc aag ccc cgg caa aaa 1344

Phe Pro Arg Gly Asn Gln His Cys Arg Ala Cys Lys Pro Arg Gln Lys

435 440 445

ctt gtt acc agc ttc tgt cgg agt gac ttt gtc atc ctg ggc agg gtc 1392

Leu Val Thr Ser Phe Cys Arg Ser Asp Phe Val Ile Leu Gly Arg Val

450 455 460

tct gag ctg acc gag gag caa gac tcg ggc cgt gcc ctg gtg acc gtg 1440

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

465 470 475 480

gat gag gtc tta aaa gat gag aag atg ggc ctc aag ttt ctg ggc cgg 1488

Asp Glu Val Leu Lys Asp Glu Lys Met Gly Leu Lys Phe Leu Gly Arg

485 490 495

gag cct ctg gaa gtc acc ctg ctt cat gta gac tgg acc tgt cct tgc 1536

Glu Pro Leu Glu Val Thr Leu Leu His Val Asp Trp Thr Cys Pro Cys

500 505 510

ccc aac gtg aca gtg ggt gag aca cca ctc atc atc atg ggg gag gtg 1584

Pro Asn Val Thr Val Gly Glu Thr Pro Leu Ile Ile Met Gly Glu Val

515 520 525

gac ggc ggc atg gcc atg ctg aga ccc gat agc ttt gtg ggg gca tcg 1632

Asp Gly Gly Met Ala Met Leu Arg Pro Asp Ser Phe Val Gly Ala Ser

530 535 540

agc aca cgg cgg gtc agg aag ctc cgt gag gtc atg tac aag aaa acc 1680

Ser Thr Arg Arg Val Arg Lys Leu Arg Glu Val Met Tyr Lys Lys Thr

545 550 555 560

tgt gac gtc ctc aag gac ttc ctg ggc ttg caa tga 1716

Cys Asp Val Leu Lys Asp Phe Leu Gly Leu Gln

565 570

<210> 49

<211> 571

<212> PRT

<213> Mus sp.

<400> 49

Met Cys Ala Pro Gly Tyr His Arg Phe Trp Phe His Trp Gly Leu Leu
1 5 10 15

Leu Leu Leu Leu Leu Glu Ala Pro Leu Arg Gly Leu Ala Leu Pro Pro
20 25 30

Ile Arg Tyr Ser His Ala Gly Ile Cys Pro Asn Asp Met Asn Pro Asn
35 40 45

Leu Trp Val Asp Ala Gln Ser Thr Cys Lys Arg Glu Cys Glu Thr Asp
50 55 60

Gln Glu Cys Glu Thr Tyr Glu Lys Cys Cys Pro Asn Val Cys Gly Thr
65 70 75 80

Lys Ser Cys Val Ala Ala Arg Tyr Met Asp Val Lys Gly Lys Lys Gly
85 90 95

Pro Val Gly Met Pro Lys Glu Ala Thr Cys Asp His Phe Met Cys Leu
100 105 110

Gln Gln Gly Ser Glu Cys Asp Ile Trp Asp Gly Gln Pro Val Cys Lys
115 120 125

Cys Lys Asp Arg Cys Glu Lys Glu Pro Ser Phe Thr Cys Ala Ser Asp

130 135 140

Gly Leu Thr Tyr Tyr Asn Arg Cys Phe Met Asp Ala Glu Ala Cys Ser
145 150 155 160

Lys Gly Ile Thr Leu Ser Val Val Thr Cys Arg Tyr His Phe Thr Trp
165 170 175

Pro Asn Thr Ser Pro Pro Pro Pro Glu Thr Thr Val His Pro Thr Thr
180 185 190

Ala Ser Pro Glu Thr Leu Gly Leu Asp Met Ala Ala Pro Ala Leu Leu
195 200 205

Asn His Pro Val His Gln Ser Val Thr Val Gly Glu Thr Val Ser Phe
210 215 220

Leu Cys Asp Val Val Gly Arg Pro Arg Pro Glu Leu Thr Trp Glu Lys
225 230 235 240

Gln Leu Glu Asp Arg Glu Asn Val Val Met Arg Pro Asn His Val Arg
245 250 255

Gly Asn Val Val Val Thr Asn Ile Ala Gln Leu Val Ile Tyr Asn Val
260 265 270

Gln Pro Gln Asp Ala Gly Ile Tyr Thr Cys Thr Ala Arg Asn Val Ala
275 280 285

Gly Val Leu Arg Ala Asp Phe Pro Leu Ser Val Val Arg Gly Gly Gln
290 295 300

Ala Arg Ala Thr Ser Glu Ser Ser Leu Asn Gly Thr Ala Phe Pro Ala
305 310 315 320

Thr Glu Cys Leu Lys Pro Pro Asp Ser Glu Asp Cys Gly Glu Glu Gln

325 330 335

Thr Arg Trp His Phe Asp Ala Gln Ala Asn Asn Cys Leu Thr Phe Thr

340 345 350

Phe Gly His Cys His His Asn Leu Asn His Phe Glu Thr Tyr Glu Ala

355 360 365

Cys Met Leu Ala Cys Met Ser Gly Pro Leu Ala Thr Cys Ser Leu Pro

370 375 380

Ala Leu Gln Gly Pro Cys Lys Ala Tyr Val Pro Arg Trp Ala Tyr Asn

385 390 395 400

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

405 410 415

Asn Gly Asn Asn Phe Glu Ser Arg Glu Ala Cys Glu Glu Ser Cys Pro

420 425 430

Phe Pro Arg Gly Asn Gln His Cys Arg Ala Cys Lys Pro Arg Gln Lys

435 440 445

Leu Val Thr Ser Phe Cys Arg Ser Asp Phe Val Ile Leu Gly Arg Val

450 455 460

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

465 470 475 480

Asp Glu Val Leu Lys Asp Glu Lys Met Gly Leu Lys Phe Leu Gly Arg

485 490 495

Glu Pro Leu Glu Val Thr Leu Leu His Val Asp Trp Thr Cys Pro Cys

500 505 510

Pro Asn Val Thr Val Gly Glu Thr Pro Leu Ile Ile Met Gly Glu Val

515 520 525

Asp Gly Gly Met Ala Met Leu Arg Pro Asp Ser Phe Val Gly Ala Ser

530 535 540

Ser Thr Arg Arg Val Arg Lys Leu Arg Glu Val Met Tyr Lys Lys Thr

545 550 555 560

Cys Asp Val Leu Lys Asp Phe Leu Gly Leu Gln

565 570

<210> 50

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 50

caccatgtgt gccccagggt atcatcggtt ctgg 34

<210> 51

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 51

ttgcaagccc aggaagtcct tgaggac 27

<210> 52

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Illustrative N-terminal peptide sequence

<400> 52

Leu Pro Pro Ile Arg Tyr Ser His Ala Gly Ile

1 5 10

<210> 53

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 53

cagacagaca gacagacaga cagacagaca gacagacaga cagacaga 48