(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) Int. Cl.[6]
C12N 15/32
A01N 63/00
C07K 14/325
C12N 1/21

(45) 공고일자 2003년10월10일
(11) 등록번호 10-0376096
(24) 등록일자 2003년03월03일

(21) 출원번호 10-1996-0706280
(22) 출원일자 1996년11월06일
번역문제출일자 1996년11월06일
(86) 국제출원번호 PCT/US1995/05431
(86) 국제출원일자 1995년05월05일
(81) 지정국 국내특허 : 오스트레일리아 캐나다 일본 대한민국 뉴질랜드 EP 유럽특허 : 오스트리아 벨기에 스위스 리히텐슈타인 사이프러스 독일 덴마크 스페인 핀랜드 프랑스 영국 그리스 아일랜드 이탈리아 룩셈부르크 모나코 네덜란드 포르투칼 스웨덴

(65) 공개번호 특1997-0702925
(43) 공개일자 1997년06월10일

(30) 우선권주장 08/239476 1994년05월06일 미국(US)

(73) 특허권자 마이코겐 코포레이션
미국 캘리포니아주 샌디에고 오벨린 드라이브 5501
(72) 발명자 톰슨 마크
미합중국 92126 캘리포니아 샌 디에고 캐미니토 엘버리즈 10549
슈웹 죠지 이.
미합중국 92024 캘리포니아 엔씨니테스 월뉴트뷰 1351
(74) 대리인 김학제, 문혜정

심사관 : 장제환

(54) 슈도모나스플루오레스센스내에서의키메라델타-내독소발현

명세서

<1> 발명의 배경

<2> 토양 미생물 바실루스 츄린겐시스(B. t)는 그람-양성, 포자-형성 세균으로 파라스포랄 결정성 단백질 봉입체(inclusions)의 특징이 있다. 이들 봉입체들은 자주 현미경상으로 특수한 형태의 결정들로 나타난다. 상기 단백질들은 해충에 대해 매우 유독하고 그들의 특성 활성면에서 매우 특이적이다. 특수한 비. 티. (B.t.) 독소 유전자들이 분리되고, 서열결정되었으며 재조합 DNA-계 비. 티. 제품들이 생산되고 사용에 대해 인가받았다. 또한, 곤충 저항력에 대한 내독소 유전자에 의해 유전자 조작된 식물의 이용 및 비.티. 내독소 배송 비히클로서 안정화된 완전한 미생물 세포를 이용하는 것을 포함하는, 유전공학기술을 이용하여, 이들 비. 티. 내독소를 농업적 환경에 배송하기 위한 새로운 접근방법에 개발중이다 [Gaertner, F.H., L. Kim [1988] TIBTECH 6:S4-S7). 따라서, 분리된 비.티. 내독소 유전자들은 상업적으로 가치있는 것이 되고 있다.

<3> 지난 10년 까지만 해도, 비.티. 살충제의 상업적 이용은 좁은 범위의 인시목 (모층) 해충으로 크게 제한되어 왔다. 바실루스 츄린겐시스 아종쿠르스타키의 포자들 및 결정들의 조제들은 인시목 해충들에 대한 상업적 살충제로 다년간 이용되어 왔다. 예를 들어, 바실루스 츄린겐시스 변종 쿠르스타킨 HD·1은 다수의 인시목 해충들의 유충에 대해 독성인 결정성 δ-내독소를 생산한다.

<4> 그러나 근년에, 연구원들은 훨씬 넓은 범위의 해충들에 대하여 특이성을 갖는 비.티. 살충제를 발견하였다. 예를 들어, 비.티.의 다른 종, 즉, 이스라엘렌시스 및 테네브리오니그 (비. 티. M-7로도 알려진 및 비.티. 센디에고로도 알려진)는 각각 쌍시목 및 초시목의 곤충들을 방제하기 위해 상업적으로 이용되어 왔다 (Gaertner, F.H., [1989] "Cellular Delivery Systems for Inseeticidal Proteins: Living ana Non-Living Microorganisms," in Controlled Delivery of Crop Protection Agents, R.M. Wilkins, ed., Taylor and Francis, New York and London, 1990, pp.245-255). 또한 Couch, T.L. (1980) "Mesquito Pathogenicity of Bacillus thuringiensis var. israelensis", Developments in Indurstrial Microbiology 22:61-76: Beegle, C.C., (1987) "Use of Entomogenous Bacteria in Agroecosystems," Developments in Industrial Microbiology 20:97-104 참조. Krieg, A., A.M. Huger, G.A. Langenbruch, W. Schnetter (1983) Z. ang. Ent. 96:500-508은 일설에 의하면 쌍시목의 두 종류의 딱정벌레들에 대하여 활성을 갖는 바실루스 츄린겐시스 변종 테네브리오니스 를 기술하고 있다. 이들은 콜로라도 감자 딱정벌레, 렙티노타르사 데셈리네아타 (Leptinotarsa decemlineata) 및 아젤라스티카 알니 (Agelastica alni)이다.

<5> 최근에, 많은 새로운 비. 티.의 아종들이 동정되었고, 활성 δ-내독소

단백질들을 코드화하는 유전자들이 분리되었다.(H()fte, H., H.R. Whiteley [1989] *Microbiological Reviews* 52(2):242-255). H()fte와 Whiteley는 비. 티. 결정 단백질 유전자들을 4개의 주된 군으로 분류하였다. 이러한 군들은 CryI (인시류-특이적), CryII (인시류 및 쌍시류 특이적), CryIII (초시류-특이적), 및 CryIV (쌍시류 특이적)였다. 다른 해충들에 대하여 특이적으로 독성을 나타내는 균주들의 발견이 보고되었다(Feitelson, .J. S., J. Payne, L. Kim [1992] Bio/Technology 10:271-275).

<6> 비. 티. 결정 단백질 유전자의 대장균 (Escherichia coli)내에서의 클로닝과 발현은 발행된 문헌에 기술되어 있다. (Schnepf, H.E., H.R. Whiteley [1981] Proc, Natl. Acad, Sci. USA 78:2893-2897). 미국 특허 제 4,448,885호와 미국특허 제 4,467,036호는 모두 비. 티. 결정 단백질의 이. 콜라이내에서의 발현을 개시하고 있다. 증가된 독성을 시현하고 표적 해충에 대한 확장된 숙주 범위를 나타내는 잡종 비. 티. 결정 단백질들이 제작되어 왔다. 미국특허 제 5,128,130호와 미국특허 제 5,055,294호 참조. 미국특허 제 4,797,276호와 미국특허 제 4,853,331호는 다양한 환경에서 초시류 해충을 방제하는데 이용될수 있는 비. 츄린겐시스 균주 테네브리오니스 (비. 티. 샌 디에고 균주로도 알려지고, M-7로도 알려진)를 개시하고 있다. 미국특허 제 4,918,006호는 쌍시류에 대해 활성을 갖는 비.티. 독소를 개시하고 있다. 미국특허 제 4,849,217흐는 알팔파 바구미에 대해서 활성을 갖는 바실루스 츄린겐시스 분리주들을 개시하고 있다. 미국특허 제 5,208,077호는 초시류-활성 바실루스 츄린겐시스 분리주를 개시한다. 미국특허 제 5,151,363호 및 미국특허 제 4,948,734호는 선충류에 대해서 활성을 갖는 비.티의 특수한 분리주를 개시하고 있다. 활발할 연구 및 재원의 투자의 결과로, 다른 특허들이 신규한 비.티. 분리주 및 비.티. 분리주들의 신규한 용도에 대하여 발행되었다. 그러나, 신규한 비.티. 분리주 및 공지의 비.티. 분리주들의 신규한 용도에 대한 발견은 경험적인, 예측할 수 없는 분야로 남아 있다.

<7> 대다수의 바실루스 츄린겐시스 δ-내독소 결정 단백질 분자들은 두개의 기능적인 분절들로 구성된다. 단백질분해효소-저항력 코어 독소가 첫 번째 분절로서 대략적으로 단백질 분자의 처음 반쪽에 해당한다. cryIIIA 비.티. δ-내독소의 코어 분절의 3차원 구조는 공지되어 있고 모든 관련 독소들이 동일한 전체적인 구조를 갖는 것으러 제안되었다 (Li. J., J. Carroll, D.J. Ellar [1991] Nature 353:815-821). 상기 분자의 두 번째 반쪽이 두 번째 분절이다. 본원의 목적상, 이러한 두 번째 분절을 이하에서 "프로톡신 분절(protoxin segment)"로 칭한다. 프로톡신 분절은 독소 결정 생성에 관여하는 것으로 생각된다 (Arvidson, H., P.E. Dunn, S. Strand, A.I. Aronson [1989] Molecular Microbiology 3:1533-1534; Choma, C.T., W.K. Surewicz P.R. Carey, M. Pozsgay, T. Raynor, H. Kaplan [1990] Eur. J. Biochem. 189:523-527). 전체의 130 kDa 독소 분자는 곤충의 장내에서 단백질분해효소에 의해 저항력 코어 분절로 신속하게 처리된다. 따라서 프로톡신 분절은 독소 분자의 단백질분해효소 처리를 감소시킴으로써 (Haideit M.Z., B.H. Knowles, D.J. Ellar [1986] Eur. J. Biochem. 156:531-540) 또는 독소 용해도를 감소시킴으로써 (Aronsom, A.I., E.S. Han, W, McGaughey, D. Johnson [1991] Appl. Environ. Microbiol. 57:981-986) 곤충에 대한 코어의 접근가능성을 제한하여 독소에 대해 부분적인 곤충 특이성을 전달할 수 있다.

<8> 독소 도메인들 내에서 연결된 키메라 단백질들은 CryIC 및 CryIA(b) 사이에서 보고된 바 있다 (Honee, G., D. Convents, J. Van Rie, S. Jansens, M. Perferoen, B. Visser [1991] Mol. Microbiol. 5:2799-2806); 그러나, 관련 곤충에 대한 이들 키메라 단백질들의 활성은 Cry I C에 비해 너무 약하거나 또는 검출할 수 없을 정도였다.

<9> Honee 등 (Honee, G., W. Vrizen, B. Visser [1990] Appl. Environ. Mocrobiol. 56:823-825)도 Cry I C 및. Cry I A(b)의 탄뎀 독소 도메인들을 연결함으로써 키메라 융합 단백질들 제조하는 것을 보고하였다. 그 결과로서 수득된 단백질은 개개의 독소들의 결합된 활성들과 동등한 향상된 활성의 스펙트럼을 갖는다; 그러나, 키메라의 활성은 표적 해 충들중 어느 하나에 대해서도 향상되지 않았다.

<10> 발명의 간단한 개요

<11> 본 발명은 슈도모나스에서의 바실루스 츄린겐시스(비.티.) δ-내독소의 발현이 비.티. 독소를 코드화하는 유전자를 변형시킴으로써 현저하게 향상될 수 있다는 발견에 관계한다. 구체적으로, 슈도모나스 플루오레스센스에서의 비.티. 내독소 발현은 네이티브 프로톡신-코드화 분절을 상이한 프로톡신 분절로 대체하여 키메라 유전자를 제공하도록 유전자를 재구성함으로써 향상될 수 있다.

<12> 본 발명의 특수한 실시양태에 있어서, 키메라 유전자들은 이종 프로톡신 분절을 네이티브 cry I F 프로톡신 분절 대신에 치환하도록 조립될 수 있다. 특히, cry I A(b) 유전자의 전부 또는 일부의 프로톡신-코드화 부위는 네이티브 cry I F 독소의 프로톡신을 코드화하는 부위의 전부 또는 일부 대신에 이용될 수 있다. 마찬가지로, cry I F 독소의 진부 또는 일부를 코트화하는 부위가 cry I A(c)/cry I A(b) 키메라 유전자의 프로톡신의 전부 또는 일부를 코드화하는 DNA에 의해 대체된 키메라 유전자도 제작될 수 있다. 특정한 실시양태에 있어서, cry I A(c)/cry I A(b) 키메라 유전자는 436DM로 표시되고 미국특허 제 5,128,130호에 기술된 것이다. 이러한 유전자는 슈도모나스 플루오레스센스 MR436내의 플를라스미드로부터 수득될 수 있다.

<13> 본 발명은 특허청구된 독소를 코드화하는 키메라 유전자의 이용도 포함한다. 상기 키메라 유전자는 광범위한 미생물 또는 식물 숙주내에 삽입될 수 있다. 키메라 유전자를 발현하는 형질전환된 숙주는 본 발명의 인시류-활성 독소를 생산하는데 이용될 수 있다. 형질전환된 숙주들은 살충 독소의 생산에 이

용될 수 있거나 또는 독소를 생산하도록 형질전환된 식물 세포의 경우에, 그 식물은 곤충의 공격에 대해 저항력을 갖게 된다. 본 발명은 키메라 독소 또는 키메라 독소를 코드화하는 유전자를 포함하는 숙주를 인시류 해충의 억제방법에 이용하는 방법에도 추가로 관계한다.

<14> 더욱이, 본 발명은 본 발명의 키메라 독소를 생산하는 실질상 온전한 재조합 세포들의 처리를 포함한다. 세포들은 실질상 온전한 세포들이 표적 해충의 환경에 적용될 때 인시류 활성이 연장되도록 처리된다. 이러한 처리는 선택된 수단이 살충제의 특성에 악영향을 미치지 않고 살충제를 보호하는 세포의 능력을 감소시키지 않는한은 화학적 또는 물리적 수단 또는 화학적 및 물리적 수단의 조합일 수 있다. 처리된 세포는 살충제 독소에 대한 보호 코팅체로 기능한다. 독소는 표적 곤충에 의해 섭취되었을 때 활성으로 된다.

<15> [도면의 간단한 설명]

<16> 제 1도 - BamHI 부위는 pMYC1050로 부터 클레노 폴리머라제에 의한 필-인 (fill-in) 반응에 의해 제거되어 플라스미드 pMYC1050△ BamHI을 제공한다. 클로닝를 용이하게 하기 위해, 대부분의 독소 개방해독틀을 포함하는 NsiI DNA 단편이 pGEM5에 클로닝된다. 그 결과로서 수득되는 플라스미드를 pGEMtox라 칭한다. C=ClaI, H=HindIII.

<17> 제 2도 - BamHI 또는 PvuI 클로닝 부위들은 스플라이스 중첩 연장 (SOE: Splice Overlap Extension) 기술에 의해 독소 DNA 내에 도입되었다. 새로운 부위를 갖는 DNA 단편은 pGEMtox에서 동종 DNA 단편들을 대체하는데 이용된다. 그 결과로서 수득되는 플라스미드들은 pGEMtoxBamHI 또는 pGEMtoxPvuI이다. 화살표 밑의 철자 A 내지 G는 텍스트내에 있는 올리고뉴클레오티프 프라이머에 상당한다. 수직선 위의 철자들은 제한효소부위이다. B=BamHI, C=ClaI, H=HindIII, P=PvuI, S=Sac I.

<18> 제 3도 - BamHI 돌연변이를 포함하는 DNA 단편은 pGEMtoxPvuI에 있는 상동 단편을 대체하는데 이용된다. 양자의 클로닝 부위들을 포함하는 그 결과 플라스미드는 pGEMtoxBamHI/PvuI이다. 발현 벡터를 구성하기 위해서, 독소-포함 NsiI 단편이 넓은 숙주 범위를 갖는 벡터인 pTJS260 내에 클로닝되기 위해 절제된다. B=BamHI, C=ClaI, H=HindIII, P=PvuI.

<19> 제 4도 - 새로운 제한부위를 갖는 NsiI 독소-포함 단편은 pMYC1050△BamHI으로 부터의 벡터-포함 DNA에 연결되어 pMYC2224를 제공한다. cryIF 독소를 포함하는 BemH I-Pua I PCR-유래 DNA 단편은 pMYC2224에 있는 동등한 서열과 교환된다. 그 결과로서 수득된 키메라를 pMYC2239라 칭한다. B=BamHI, C=ClaI ,H=HindIII, N=NsiI, P=PvuI.

<20> 제 5도 - pMYC2047의 작은 ApaI DNA 단편은 pMYC2239의 상동 부위 대신에 대체되어 플라스미드 pMYC2244를 제공한다. 이러한 키메라는 독소 부위의 cryIF 및 프로톡신의 cryIA(b)로 구성된다. C=ClaI, H=HindIII, N=NsiI, P=PvuI.

<21> 제 6도 - SOE에 의해 잠재 코돈 변화(silent codon change)가 cryIF 독소내에 도입된다. 이러한 변화들을 갖는 SpeI-KpnI PCR DNA 단편이 pMYC2047의 상동 독소-포함 단편 대신에 대체된다. 그 결과 플라스미드는 pMYC2243이다. 화살표 하단의 철자 H 내지 S는 텍스트의 올리고뉴클레오티드 프라이머에 상당한다.

<22> 제 7도 - 상동 단편을 pMYC2243의 작은 ApaI DNA 단편으로 대체함으로써 pMYC2244내에 잠재 코돈 변화들이 도입된다. 최종 플라스미드는 pMYC2523이다. P=PvuI.

<23> 제 8도 - 436 프로톡신을 포함하는 키메라 독소는 PCR-생성 PvuI-BstEII 프로톡신 DNA를 pMYC2523내의 상동 단편 대신에 대체함으로써 구성된다. 최종 플라스미드는 pMYC2254이다. 화살표하단의 철자 F 내지 M은 텍스트내의 올리고뉴클레오티드 프라이머를 나타낸다.

<24> 제 9도 - CryIF/CryIA(b) 키메라 단백질 서열 및 잔기-대-잔기 치환. "Cons" 라인은 CryIF/CryIA(b) 키메라 서열을 나타낸다. "Alt" 라인은 436 단백질, CryIA(b) 변이체 단백질 및 CryIF 프로톡신에서 발견된 잔기-대-잔기 치환을 나타낸다.

<25> 서열의 간단한 설명

<26> 서열번호 1은 올리고뉴클레오티드 프라이머 "A"이다.

<27> 서열번호 2는 올리고뉴클레오티드 프라이머 "B"이다.

<28> 서열번호 3은 올리고뉴클레오티드 프라이머 "C"이다.

<29> 서열번호 4는 올리고뉴클레오티드 프라이머 "B"이다.

<30> 서열번호 5는 올리고뉴클레오티드 프라이머 "E"이다.

<31> 서열변호 6은 올리고뉴클레오티드 프라이머 "F"이다.

<32> 서열변호 7은 올리고뉴클레오티드 프라이머 "G"이다.

<33> 서열번호 8은 올리고뉴클레오티드 프라이머 "L"이다.

<34> 서열번호 9는 올리고뉴클레오티드 프라이머 "N"이다,

<35> 서열번호 10은 올리고뉴클레오티드 프라이머 "O"이다.

<36> 서열번호 11은 올리고뉴클레오티드 프라이머 "H"이다.

<37> 서열번호 12는 올리고뉴클레오티드 프라이머 "T"이다.

<38> 서열번호 13은 올리고뉴클레오티드 프라이머 "J"이다.

<39> 서열번호 14는 올리고뉴클레오티드 프라이머 "K"이다.

<40> 서열번호 16는 올리고뉴클레오티드 프라이머 "P"이다.

<41> 서열번호 16은 올리고뉴클레오티드 프라이머 "Q"이다.

<42> 서열번호 17은 올리고뉴클레오티드 프라이머 "M"이다.

<43> 서열번호 18은 pMYC2224의 독소-코드화 DNA서열을 도시한 것이다.

<44> 서열번호 19는 pMYC2224에 의해 코드화된 독소의 추론 아미노산 서열을 도시한 것이다.

<45> 서열번호 20은 pMYC2239의 독소-코드화 DNA 서열을 도시한 것이다.

<46> 서열번호 21은 pMYC2239에 의해 코드화된 독소의 추론 아미노산 서열을 도시한 것이다.

<47> 서열번호 SS는 cryIF/criIA(b) 키메라 독소를 코드화하는 pMYC2244의 독소-코드화 DNA 서열을 도시한 것이다.

<48> 서열번호 23은 pMYC2244에 의해 코드화된 cryIF/cryIA(b) 키메라 독소의 추론 아미노산 서열을 도시한 것이다.

<49> 서열번호 24는 pMYC2243의 독소-코드화 DNA 서열을 도시한 것이다.

<50> 서열번호 25는 pMYC2243에 의해 코드화된 독소의 추론 아미노산 서열을 도시한 것이다.

<51> 서열번호 26은 코돈 개정(codon rework)를 갖는 cryIF/cryIA(b) 키메라 독소를 코드화하는 pMYC2523의 독소-코드화 DNA 서열을 도시한 것이다.

<52> 서열번호 27은 pMYC2523에 의해 코드화된 독소의 추론 아미노산 서열을 도시한 것이다.

<53> 서열번호 28은 cryIF/4S6 키케라 독소를 코드화하는 pMYC2254의 독소-코드화 DNA 서열을 도시한 것이다.

<54> 서열번호 29는 pMYC2254에 의해 코드화된 독소의 추론 아미노산 서열을 도시한 것이다.

<55> 서열변호 30은 cryI 독소의 특징 서열이다. 이 서열은 서열번호 30의 601 잔기에서 끝난다.

<56> 서열번호 31은 서열번호 23에서 아미노산 1043의 앞에 있는 8 아미노산들이다.

<57> 서열번호 32는 네이티브 cryIF 독소의 아미노산 서열을 도시한 것이다.

<58> 서열번호 33은 네이티브 cryIA(b) 독소의 아미노산 서열을 도시한 것이다.

<59> 서열번호 34는 cryIA(b)/cryIA(b) 독소의 아미노산 서열을 도시한 것이다.

발명의 상세한 설명

<60> 발명본 발명은 비.티. 독소를 코드화하는 특정한 키메라 유전자들의 발현이 재조합 슈도모나스 플루오레스센스에서 향상된다는 발견에 관계한다. 키메라 유전자들은 네이티브 프로톡신 부분의 전부 또는 일부가 다른 비.티. 독소로 부터의 프로톡신의 전부 또는 일부로 대체된 독소를 코드화한다. 본원에서 구체적으로 예시된 것은 본원에 기술된 바와 같은 cryIA(b) 또는 cryIA(c)/cryIA(b) 독소 중 하나로 부터 유래된 프로톡신 분절에 결합된 cryIF 코어 N-말단 독소 부분을 필수적으로 포함하는 비.티. 독소를 코드화하는 유전자들이다. 본원에서 이용되는 경우에, "코어" 독소 부분은 독소의 살충 활성을 담당하는, 프로톡신 이외의 전체 길이의 비.티. 독소를 의미한다.

<61> 본 발명에 따라 유용한 플라스미드를 포함하는 세균은 다음과 같다:

배 양 균	기 탁 번 호	미 국 특 허
피. 플루오레스센스(pM3, 130-7)	NRRL B-18332	제 5,055,294호
피. 플루오레스센스 MR436 (pM2, 16-11, aka pMYC436)	NRRL B-18292	제 5,128,130호
이. 콜라이 NM522 (pMYC1603)	NRRL B-18517	제 5,188,960호

<63> 기탁물의 이용가능성이 행정부의 활동에 의해 수여된 특허권을 감손시키는 본 발명 실시의 허락을 구성하는 것이 아님을 이해하여야 한다.

<64> 제 1도 내지 제 8도의 플로우 챠트는 본 발명에 따라 수행될 수 있는 벡터 구성의 개관을 제공한다. BamHI 및 PvuI 클로닝 부위는 스플라이스 중첩 연장(SOE)의 PCR 기술(Horton, R.M., H.D. Hunt, S.N. Ho, J.K. Pollen, L,R. Pease [1989] Gene 77:61-681을 이용한 돌연변이유발에 의해 cry I A(c)/cry I A(b) 키메라 독소 유전자내에 도입되어 플라스미드 pMYC2254를 제공할 수 있다. pMYC1260과 같은 cry I F-포함 플라스미드로 부터의 cry I F 유전자의 일부위는 PCR에 의해 만들어질 수 있고 pMYC2224의 BamHI-PvuI cryIA(c)/cryIA(b) 유전자 단편 대신에 대체될 수 있다. 우리가 pMYC2239로 정한 신규의 플라스미드는 독소/프로톡신 분절 접합부까지 cry I F에 의해 이어지는 cry I A(c)의 짧은 분절로 구성된다. 따라서, 프로톡신 분절은 cry I A(b)(pMYC1050)로 부터 유래되었다. cry I F 클른 (pMYC2047)로 부터 유래된 Apa I 단편은 pMYC2239에 있는 Apa I 단편 대신에 대체되었다. 그 결과 클론 (pMYC2244)은 개시 메

치오닌으로 부터 독소/프로톡신 분절 접합부 까지의 cry I F 및 코드화 부위 말단 까지의 cry I A(b)로 구성된다. 클론 pMYC2243은 제한된 부위에 잠재성 코돈 변화(silent codon mutation)를 도입하기 위해 SOE에 의해 제작되었다. 잠재성 변화를 포함한 pMYC2243으로 부터의 Aps I 단편은 pMYC2244에 있는 Apal 단편 대신에 대체되어 pMYC2523을 만들었다. 키메라 pMYC2523은 변화되지 않은 cry I F 단백질 서열을 포함하는 pMYC2243에 비해 향상된 발현을 시현하였다.

<65> cry I F/436 키메라는 pMYC2523의 PvuI-BstEII 단백질 분절-포함 단편을 cry I A(c)/cry I A(b) 유전자를 포함하는 플라스미드로 부터 PCR에 의해 생성된 동등한 단편으로 대체함으로써 조립될 수 있다. 하나의 이러한 유전자는 436 유전자 (예컨대, pMYC467, 미국특허 제 5,128, 130호 및 동 5,169,760호에 기술된 바와 같은)이다. 이러한 구조물도 네이티브 cry I F 단백질 서열에 비해 향상된 발현을 초래한다.

<66> 본 발명의 키메라 독소는 비.티. 독소의 전체 코어 N-말단 독소 부분을 포함하고, 독소 부분의 말단의 및 포인트 지난 위치에서, 상기 단백질은 이종 프로톡신 서열에 대한 전위부(transition)를 갖는다. 이종 프로톡신 분절에 대한 전위부는 대략적으로 독소/프로톡신 접합부에 존재하거나 또는 대안으로 네이티브 프로톡신의 일부분 (독소 부분을 지나 신장하는)은 하류에 존재하는 이종 프로톡신에 대한 전위부에 의해 유지될 수도 있다. 일례로, 본 발명의 하나의 키메라 독소는 cry I F의 전체 독소 부분(아미노산 1~601) 및 이종 프로톡신 (아미노산 602~C-말단)을 갖는다. 바람직한 실시양태에 있어서, 프로톡신의 이종 부분은 cry I A(b) 또는 436 독소로 부터 유래된다.

<67> 본 발명이 속하는 기술분야의 숙련가들은 비.티. 독소들은, 심지어 cry I F와 같은 하나의 특수한 클래스에 속한다고 해도, 독소 부분으로 부터 프로톡신 부부까지의 전위부의 정확한 위치 및 길이가 어느 정도 달라질 것이라는 사실을 이해할 것이다. 일반적으로, cry I A(b) 및 cry I F 독소들은 길이가 약 1150 내지 약 1200 아미노산들이다. 독소 부분으로 부터 프로톡신 부분까지의 전위부는 일반적으로 전체 길이 독소의 약 50% 내지 약 60% 사이를 차지할 것이다. 본 발명의 키메라 독소는 전체 길이의 이러한 코어 N-말단 독소 부분을 포함한다. 따라서 키메라 독소는 전체 길이 cry I F 비.티. 독소의 적어도 약 50%를 포함할 것이다. 이것은 일반적으로 적어도 약 590 아미노산들일 것이다. 프로톡신 부분과 관련하여, cry I A(b) 프로톡신 부분의 전체 길이는 독소 부분의 말단에서 부터 분자의 C-말단 까지 신장한다. 본 발명의 키메라 독소에서 포함해야할 가장 중요한 부분은 이러한 부분의 마지막 약 100내지 150 아미노산들이다. 본원에서 구체적으로 예시된 키메라 독소에서, cry I A(b) 분자의 적어도 아미노산 1043 (서열번호 23) 내지 C-말단이 이용된다. 서열번호 23에서 아미노산 1043은 서열 Tyr Pro Asn Asn Thr Val Thr Cys (서열번호 31) 다음에 놓인다. 이러한 아미노산 서열은 키메라 독소에서 그 너머에는 항상 이종 아미노산들이 존재하는 분자의 프로톡신 분절 내에서의 위치를 나타낸다. 다른 예에서, 서열번호 31로 도시된 펩타이드는 아미노산 1061 내지 1068에 존재한다. 이 경우에, 아미노산들 1069 내지 C-말단은 바람직하게 이종이다(서열번호 29). 서열번호 31에 도시된 펩타이드는 제 9 도에서 위치 1061 내지 1068에서 발견될 수 있다. 따라서, 본 발명의 키메라 독소에서 이종 DNA (cry I F 코어 N-말단 독소 부분에 비해)를 포함해야만 하는 것은 적어도 전체 비.티. 단백질의 마지막 약 5 내지 10%이다. 본원에 포함된 구체적인 실시예들에 있어서, 이종 프로톡신 서열들은 아미노산 640부터 C-말단 까지 존재한다.

<68> 따라서, 본 발명의 바람직한 실시양태는 약 1150 내지 약1200 아미노산 길이의 키메라 비.티. 독소로서, 여기서 키메라 독소는 전체 cry I F 분자의 적어도 약 50 내지 60%이지만, 전체 분자의 약 90 내지 95%를 초과하지 않는 cry I F 코어 N-말단 독소 부분을 포함한다. 키메라 독소는 cry I A(b) 또는 436 분자의 적어도 약 5 내지 10%를 포함하는 cry I A(b) 또는 436 프로톡신 C-말단 부분을 추가로 포함한다. 따라서 cry I F로 부터 cry I A(b)까지의 전위부분 또는 436 서열로의 전위는 전체 분자의 약 50% 와 약 95% 사이의 프로톡신 분절(또는 독소와 프로톡신 분절들의 접합부)내에서 일어난다. 본원에 제시된 구체적인 실시예에 있어서, cry I F 서열로 부터 이종 프로톡신 서열로의 전위는 서열번호 31에 도시된 펩타이드의 말단 이전에서 일어난다.

<69> 본 발명의 하나의 구체적인 실시예는 제 9도에 도시된 키메라 독소이다. 다른 구조물들은 본원에 제공된 교시의 도움을 이용함으로써 본 발명이 속하는 기술분야의 숙련가들에 의해 만들어질 수 있고 이용될 수 있다. cry I 단백질의 코어 독소 분절은 특징적으로 다음 서열로 끝난다

: Val/Leu Tyr/Ile Ile Asp Arg/Lys Ile/Phe Glu Ile/Phe/Leu

Ile/Leu/Val Pro/Leu/Ala/Val Glu/Thr/Asp (서열번호 30). 이러한

<71> 서열은 서열번호 23의 잔기 601에서 끝난다. 또한 cry I 독소들 (잔기 601 다음에 이어지는)의 프로톡신 분절들은 독소 분절들 보다 더 많은 서열 유사성을 갖는다. 이러한 서열 유사성 때문에, cry I F 서열과 cry I A(b) 또는 436 서열사이의 키메라 단백질을 만드는 프로톡신 분절에서의 전위점은 당업자들에 의해 쉽게 결정될 수 있다. cry I 유전자들의 부분 단백질분해에 관한 데이타의 연구로 부터, 이성분 및 최소-보존 아미노산 부위들이 제 9 도의 위치 1061 · 1068의, 보존된 cry I 프로톡신 서열 다음에서 발견된다. 따라서, 본 발명의 키메라 독소는 전체 cry I F 독소 및 독소 분절(전술한 바와 같은)의 말단과 서열번호 31에 도시된 펩타이드 서열의 말단 사이의 임의의 위치에서 이와 대응하는 cry I A(b) 또는 436 서열로 전위되는 일부분의 cry I F 프로톡신을 포함할 수 있다. 바람직하게, 키메라 독소의 C-말단의 아미노산 서열은 cry I A(b) 서열 또는 436 유전자로 부터의 서열 또는 이들 서열들 중 하나의 동등물을 포함한다.

<72> Cry I F 독소들 및 이들 독소들을 코드화하는 유전자들은 당업계에 주지되어 있다. Cry I F 유전자들 및 독소들은, 예를들어, Chambers et al., (1991) J. Bacteriol. 173 : 3966에 기술되어 있다. cry I A(b) 유전자들 및 독소들은 예컨대 H@fte et al., (1986) Eur.J.Bioehem. 161 : 273); Geiser et al. (1986) Gene 48 : 109 : 및 Haider et al. (1988) Nucleic Acids Res. 16 : 10927에 기술되어 있다. 본원에 포함된 교시의 도움을 받은 숙련된 당업자들은 cry I F 분자의 독소 N-말단 부분 및 cry I A(b) 독소

들의 C-말단 프로톡신 부분을 코드화하는 DNA를 쉽게 동정 및 이용할 수 있을 것이다,

<73> 제 9 도는 본 발명의 독소들에 이용될 수 있는 아미노산 치환들의 예들을 제공한다. 독소의 활성을 변화시키지 않으면서 독소 서열에 대해 여러가지 돌연변이들이 만들어질 수 있다는 사실도 당업계에 잘 알려져 있다. 더욱이, 유전코드의 퇴화성으로 인해, 여러가지 DNA 서열들이 특정 독소를 코드화하는데 이용될 수 있다. 이들 대안의 DNA 및 아미노산 서열들은 본 발명이 속하는 기술분야의 당업자들에 의해 본 발명에 따라 이용될 수 있다.

<74> 본 발명의 프로톡신 치환 기술은 독소의 발현을 증가시키기 위해 다른 클래스의 비.티. 내독소들에 대해서도 이용될 수 있다. 이 기술은 서열번호 30에 도시된 특징적인 서열을 갖는 다른 비.티. 독소들에 가장 잘 적용될 수 있을 것이다.

<75> 본 발명은 신규한 키메라 독소들 및 이들 독소들을 코드화하는 유전자들 뿐만 아니라 이들 신규한 독소들 및 유전자들의 이용방법들도 포함한다. 예를들어, 본 발명의 유전자는 숙주 세포를 형질전환시키는데 이용될 수 있다. 유전자를 발현하고 키메라 독소를 생산하는 숙주 세포들은 살충 조성물에 이용되거나 또는 형질전환된 식물 세포의 경우에는, 형질전환된 세포 자체에 곤충 저항성들 부여하는데 이용될 수 있다.

<76> 유전자들 및 독소들. 본 발명에 따라 이용가능한 유전자들 및 독소들은 개시된 전체 길이의 서열들 뿐만 아니라 본원에 구체적으로 예시된 특징적인 살충 활성을 보유하는 이들 서열들의 단편, 변이체, 및 돌연변이 체를 포함한다. 본원에서 사용되는 경우에, 유전자의 "변이체" 또는 "변형물" 이라 함은 동일한 독소를 코드화하거나 살충 활성을 갖는 동등한 독소를 코드화하는 뉴클레오티드 서열들을 의미한다. 본원에서 사용될 때, "동등한 독소(equivalent toxins)" 라는 용어는 청구된 독소들과 동일하거나 본질적으로 동일한 표적해충에 대한 생물학적 활성을 갖는 독소들을 말한다.

<77> 활성 독소를 코드화하는 유전자들이 몇가지 방법들에 의해 동정되고 수득될 수 있다는 사실은 당업자들에게 자명한 사항이다. 본 발명에 따라 이용가능한 특수한 유전자들(또는 독소 또는 프로톡신 도메인을 코드화하는 그의 일부분)은 전술한 배양균 기탁기관에 기탁된 재조합 분리주로부터 수득될 수 있다. 이들 유전자들 또는 이들의 일부 또는 변이체들도 예컨대 유전자 합성기를 이용함으로써 합성에 의해 인공적으로 만들어질 수 있다. 유전자들의 변형물들은 점돌연변이를 만드는 표준기술을 이용함으로써 용이하게 만들어질 수 있다. 또한, 이들 유전자들의 단편들은 표준방법에 따라 시판되는 엑소뉴클레아제 또는 엔도뉴클레아제를 이용함으로써 만들어질 수 있다. 예를들어, Ba131과 같은 효소들은 이들 유전자들의 말단들로 부터 뉴클레오티드들을 체계적으로 절단하는데 이용될 수 있다. 대안으로, 부위-직접 돌연변이 유발(site-directea mutagensis)가 이용될 수 있다. 또한 활성 단편을 코드화하는 유전자들은 여러가지 제한효소들을 이용함으로써 수득될 수 있다. 단백질분해효소들은 이들 독소들의 활성 단편들을 직접적으로 수득하는데 이용될 수 있다.

<78> 예시된 독소의 살충 활성을 보유하는 단편들 및 동등물들은 본 발명의 범위내에 포함될 것이다. 또한, 유전암호의 중복성 때문에, 여러가지 상이한 DNA 서열들이 본원에 개시된 아미노산 서열을 코드화할 수 있다. 동일하거나 본질적으로 같은 독소를 코드화하는 이들 대용의 DNA 서열들을 만드는 것은 본 발명이 속하는 기술분야의 숙련가들의 기술의 범위내에 포함되는 것이다. 이들 변이체 DNA 서열들은 본 발명의 범위내에 포함된다. 본원에서 이용되는 경우에, "본질적으로 동일한" 서열이라 함은 살충 활성 또는 발현 수준에 실질적으로 영향을 미치지 않는 아미노산 치환, 결실, 부가 또는 삽입을 갖는 서열들을 말한다. 살충활성을 보유하는 단편들도 이러한 정의에 포함된다.

<79> 본 발명의 독소들 및 유전자들을 동정하는 추가의 방법은 올리고뉴클레오티드 프로브의 이용을 통하는 것이다. 이들 프로브들은 검출가능한 튜클레오티드 서열들이다. 이들 서열들은 적당한 표지에 의해 검출가능하거나 또는 국제출원 제 WO93/16094호에 기술된 바와 같이 본래부터 형광물질로 만들어질 수 있다. 본 발명이 속하는 기술분야에 주지되어 있는 바와 같이 프로브 분자와 핵산 시료가 두 분자들 사이에 강한 결합을 형성함으로써 혼성화한다면, 프로브와 시료가 실질적인 상동성을 갖는 것으로 추정할 수 있다. 바람직하게, 혼성화는 예컨대, Keller, G.H., M.M. Manak(1987) DNA Probes, Stoektos Press, New York, NY, pp. 169-170에 기술되어 있는 바와 같은, 당업계에 주지되어 있는 기술에 의해 엄격한 조건하에서 수행된다. 프로브의 검출은 공지의 방법으로 혼성화가 일어났는지 여부를 확인할 수 있는 수단을 제공한다. 이러한 프로브 분석은 본 발명의 독소-코드화 유전자들을 확인하는 신속한 수단을 제공한다. 바람직하게, 이러한 유전자들은 그의 코어 독소-코드화 부분들이 cry I A(b) 또는 436 프로톡신-코드화 부분과 함께 이용되어 본 발명의 키메라 유전자를 만들 수 있는 cry I F 유전자들일 것이다. 본 발명에 따라 프로프로 이용되는 뉴클레오티드 분절들은 DNA 합성기 및 표준방법을 이용함으로써 합성될 수 있다. 이들 뉴클레오티드 서열들은 본 발명의 유전자들을 증폭시키기 위해 PCR 프라이머로도 이용될 수 있다.

<80> 본 발명의 특정한 키메라 독소들은 본원에 구체적으로 예시되어 있다. 본 발명이 예시된 독소와 동일하거나 유사한 살충 활성을 갖는 변이체 또는 동등한 독소들(및 동등한 독소들을 코드화하는 뉴클레오티드 서열들)을 포함한다는 것은 자명한 사실이다. 동등한 독소들은 예시된 독소와 아미노산 상동성을 갖을 것이다. 이러한 아미노산 상동성은 일반적으로 75% 초과이고, 바람직하게 90% 초과이며, 가장 바람직하게 95%를 초과할 것이다. 아미노산 상동성은 생물학적 활성을 담당하거나 궁극적으로 생물학적 활성을 담당하는 3차원 배열의 결정에 관련되는 독소의 주요 부분에서 가장 높을 것이다. 이와 관련하여, 아미노산 치환들이 활성에 결정적이 아닌 부위에서 일어나거나 분자의 3차원 배열에 영향을 미치지 않는 보존적 아미노산 치환이라면 특정한 아미노산 치환이 허용가능하고 예상될 수 있다. 예를들어, 아미노산들은 다음 클래스들로 분류될 수 위치될 수있다: 비극성, 비하전극성, 염기성 및 산성. 하나의 클래스의 아미노산이 동일한 클래스내의 다른 아미노산으로 대체되는 보존적 치환은 그러한 치환이 그 화합물의 생물학적 활성을 실질적으로 변화시키지 않는 한 본 발명의 범위내에 포함된다. 표 1은 각 클래스에 속하는 아미노산들의 예들을 열거한 것이다.

[표 1]

아미노산의 클래스	아미노산들의 예
비극성	**Ala, Val, Leu, Ile, Pro, Met, Phe, Trp**
비하전 극성	**Gly, Ser, Thr, Cys, Tyr, Asn, Gln**
산성	**Asp, Glu**
염기성	**Lys, Arg, His**

<82> 및및 경우에, 비-보존적 치환도 일어날 수 있다. 결정적인 요소는 이러한 치환들이 독소의 생물학적 활성을 현저하게 저하시켜서는 안된다는 것이다.

<83> 재조합 숙주들. 본 발명의 키메라 독소를 코드화하는 유전자는 넓은 범위의 미생물 또는 식물 숙주내에 삽입될 수 있다. 독소 유전자의 발현은 직접적으로 또는 간접적으로 살충 키메라 독소의 세포내 생산 및 보유를 결과시킨다. 예컨대, 슈도모나스와 같은 적당한 미생물 숙주에 의해, 미생물들은 해충의 서식지에 적용되고, 그곳에서 증식 및 섭취될 것이다. 그 결과는 해충의 방제이다. 그 대신에, 독소 유전자를 수용하고있는 미생물들은 독소의 활성을 연장시키고 세포를 안정화하는 조건하에서 처리될 수 있다. 독성 활성을 보유하는 처리된 세포는 이어서 표적 해충의 환경에 적용될 수 있다.

<84> 키메라 독소를 코드화하는 유전자가 적당한 벡터를 통해서 미생물 숙주내에 도입되고, 상기 숙주가 살아있는 상태로 환경에 적용될 때, 특정한 숙주 미생물이 이용되는 것이 필수적이다. 미생물 숙주들은 하나 이상의 목적 농작물의 "식물권" (엽면, 엽계, 근계, 및/또는 근면)을 점유하는 것으로 알려진 것들이 선택된다. 이들 미생물들은 특정한 환경(농작물 및 기타 곤충 서식지)에서 야생형 미생물들과 성공적으로 경쟁하여 폴리펩타이드 살충제를 발현하는 유전자의 안정한 유지 및 발현을 제공하고, 바람직하게 환경적 분해 및 불활성화로 부터 살충제를 효과적으로 보호할 수 있도록 선택된다.

<85> 다수의 미생물들이 광범위한 주요 농작물의 엽면(식물잎의 표면) 및/또는 근계 (식물뿌리 주위의 토양)에 거주하는 것으로 알려져 있다. 이들 미생물들은 세균, 조류 및 진균류를 포함한다. 특히 중요한 것은, 예컨데 슈도모나스속 (Pseudomonas), 에르위나아속(Erwinia), 세라티아속(Serratia), 클렙시엘라속(Klebsiella), 크산토모나스속(Xanthomonas), 스트렙토마이세스속(Streptomyces), 리조비움속(Rhizobium), 로도슈도모나스속(Rhodopseudomonas), 메틸로필리어스속(Methylophiius), 아그로박테리움속(Agrobacterium) , 아세토박터속(Acetobacter) , 락토바실러스속(lactobacillus) , 아르스로박터속(Arthrobacter), 아조토박터속(Azotobacter), 류코노스톡속(Leuconostoc) 및 알칼리게네스속(Alcaligenes)속들과 같은 일례의 세균; 진균류, 특히 사카로마이세스속(Saccharomyces), 크립토코커스속(Crytococcus) , 클뤼베로마이세스속(Kluyveromyces) , 스포로볼로마이세스속(Sporobolmyces) , 로도토룰라속(Rodotorula) 및 아우레오바시디움속(Aureobasidium)과 같은 일례의 효모를 포함한다. 특히 중요한 것은 슈도모나스 시린가에(Pseudomonas syringae), 슈도모나스 플루오레스센스(Pseudomonas fluorescens), 세라티아 마르세스센스(Serratia marcescens), 아세토박터 크실리눔(Acetobacter xylinum), 아그로박테리움 투메파시엔스(Agrobacterium tumefaciens), 로도슈도모나스 스페로이데스(Rhodopseudomonas spheroides), 크산토모나스 캄페스트리스(Xanthomonas campestris), 리조비움 멜리오티(Rhiobium melioti) 알칼리게네스 엔트로퍼스(Alcaligenes entrophus), 및 아조토박터 빈란디(Azotobacter vinlandii)와 같은 식물권 세균 ; 및 로도토룰라 루브라(Rhodotorula rubra), 알. 글루티니스(R. goutinis), 알. 마리나(R. marina), 알. 아우란티아카(R. aurantiaca), 크립토코커스 알비두스(Cryptococcus albidus), 시. 디플루엔스(C. diffluens), 시. 라우렌티 (c. laurentii), 시카로마이세스 로세이(Saccharomyces rosei), 에스. 프레토리엔시스(S. pretoriensis), 에스. 세레비시아에(S. cerevisiac), 스포로볼로마이세스 로세우스(Sporobolmyces roseus), 에스. 오도루스(S. odorus), 클뤼베로마이세스 베로나에(Kluyveromyces veronae), 및 아우레오바시디움 폴루란스(Aureobasidium pollulans)와 같은 식물권 효모 종들을 포함한다. 특히 중요한 것은 색소화된 미생물이다.

<86> 키메라 독소를 코드화하는 유전자들을 유전자의 안정한 유지 및 발현을 허용하는 조건하에서 미생물 숙주내로 도입하는데는 여러가지 방법을 이용할 수 있다. 이들 방법들은 당업자들에게 잘 알려져 있고, 예컨대 미국 특허 제 5,135,867(본원에 참고자료로 첨부)에 기술되어 있다.

<87> 세포의 처리. 상술한 바와 같이, 본 발명의 키메라 독소를 생산하는 재조합 세포들은 독성 활성을 연장하고 세포를 안정화하도록 처리될 수 있다. 생성된 마이크로캡슐은 안정화되고 마이크로캡슐이 표적해충의 환경에 적용될 때 독소를 보호할 세포 구조내에 비.티. 독소를 포함한다. 적당한 숙주 세포들은 진핵 세포 또는 원핵세포 모두를 포함하는데, 보통은 포유동물과 같은 고등생물에 대해서 특성인 물질을 생산하지 않는 세포들로 제한된다. 그러나, 그러한 특성 물질들이 불안정하거나 적용 레벨이 포유동물 숙주에 대해 독성을 끼칠 가능성을 배제할 수 있을 정도로 충분히 낮은 경우에는, 고등 생물에 대해 독성인 물질들을 생산하는 유기체도 이용될 수 있다. 숙주로, 긴요한 것은 원핵생물과 진균류와 같은 하등 진핵생물이다.

<88> 세포는 처리될 때 대개 완전하고 몇몇 경우에는 포자들이 이용될 수 있다고 하더라도 포자형태이기 보다는 증식 형태일 것이다.

<89> 미생물 세포, 예컨대, 본 발명의 키메라 독소를 코드화하는 유전자를 포함하는 미생물의 처리는, 그러한 기술이 독소의 특성에 해로운 영향을 미치지 않고, 독소를 보호하는 세포의 능력을 저하시키지 않는한 화학적 또는 물리적 수단, 또는 화학적 및/또는 물리적 수단의 조합에 의한 것일 수 있다. 화학적시약의 예로는 할로겐화제, 특히 원자번호 17-80의 할로겐이다. 보다 구체적으로, 요오드는 온화한 조건하에서 목적으로 하는 결과를 달성하기 위한 충분한 시간 동안 이용될 수 있다. 다른 적합한 기술은 글루타르알데하이드와 같은 알데하이드; 제피란 클로라이즈 및 세틸피리디니움 클로라이드와 같은 항-감염제; 이소프로필 및 에탄올과 같은 알콜: 루골(Lugd) 요오드, 보우인의 픽서티브, 여러가지 산들 및 헬리의 픽서티브와 같은 조직학적 고정액(Humanson, Gretchen, L., Animal Tissue Techniques, W.H. Freeman and Company, 1967); 또는 세포가 숙주 환경에 투여될 때 세포내에서 생산된 독소의 활성을 보존하고 연장시키는 물리적(열) 및 화학적 제제의 조합에 의한 처리를 포함한다. 물리적 수단 예들은 감마-방사 및 X-방사와 같은 단파장 방사, 동결, UV 조사, 동결전조등을 포함한다. 미생물 세포들의 처리 방법은 미국특허 제 4,695,455 및 동 4,695,462호 (본원에 참고자료로 첨부)에 기술되어 있다.

<90> 세포들은 일반적으로 환경 조건에 대한 저항력을 증가시키는 향상된 구조적 안정성을 구비할 것이다. 살충제가 프로폼이기 때문에, 세포 처리의 방법은 표적 해충 병원체에 의한 프로폼의 성숙된 형태로의 성숙과정을 억제하지 않도록 선택되어야만 한다. 예를들어, 포름알데하에드는 단백질들을 가교시켜 폴리펩타이드 살충제의 프로폼의 성숙을 억제할 수 있다. 처리방법은 적어도 독소의 생체이용을 또는 생물활성의 상당한 부분을 유지해야만 한다.

<91> 생산의 목적상 숙주 세포의 선정에 있어 특히 중요한 특성들은 숙주 세포내로의 유전자 삽입의 용이성, 발현 시스템의 유효성, 발현의 효율성, 숙주내에서의 살충제의 안정성, 및 보조적인 유전적 능력의 존재를 포함한다. 살충제 마이크로캡슐로 이용하기 위한 중요한 특성들은 두꺼운 세포벽, 색소 침착 및 봉입체의 세포내 포장 또는 생성과 같은 살충제에 대한 보호 특성; 수성 환경에서의 생존; 포유동물 독성의 부재; 해충들 섭취의 유인; 독소에 대한 손상 없는 치사 및 고정의 용이성등을 포함한다. 다른 고려해야할 사항들은 제형화 및 취급의 용이성, 경제성, 저장 안정성 등을 포함한다.

<92> 세포들의 생장. 본 발명의 키메라 독소를 코드화하는 유전자를 포함하는 세포 숙주는 임의의 편리한 영양 배지내에서 배양될 수 있는데, 이러한 배지서 DNA 구조물은 선택적 이점을 제공하여, 실질적으로 모든 세포들이 재조합 유전자를 보유하도록 선택배지를 제공한다. 이어서 이들 세포들은 종래의 방법에 따라 회수될 수 있다. 대안으로, 세포들은 회수되기 이전에 처리될 수 있다.

<93> 제형화. 본원에 기술된 키메라 독소를 코드화하는 유전자를 포함하는 재조합 미생물들은 미끼 과립으로 제형화되어 토양에 적용될 수 있다. 제형화된 제품들은 농작물 사이클의 후기 단계에서 종자- 코팅 또는 뿌리 처리 또는 전체 식물 처리로 적용될 수도 있다. 식물 및 토양처리는 무기광물(업황산염, 탄산염, 황산염 임산염 등) 또는 식물성 재료(분말화된 옥수수 속대, 쌀겨, 밤껍질 등)과 같은 불활성 재료들과 혼합함으로써 습윤가능한 분말, 과립 또는 가루로 적용될 수 있다. 제형들은 분무-부착 보조약, 안정화제, 기타 살충 첨가물 또는 계면활성제들을 포함할 수 있다. 액체 제형들은 수성계 이거나 비수성계일 수 있고 거품, 겔, 현탁액, 유화 농축액 등으로 적용될 수 있다. 성분들은 유동화제, 계면활성제, 유화제, 분산제 또는 중합체들을 포함할 수 있다.

<94> 당업자들에 의해 이해되는 바와 같이, 살충제 농도는 특정 제형의 특성, 특히 그것이 농축물인지 바로 쓸수 있는 것인지에 크게 의존할 것이다. 살충제는 적어도 1중량%로 존재하고 100중량%일 수도 있다. 건조 제형을 약 1-95중량% 살충제를 포함하는 반면에, 액상 제형은 일반적으로 액체상 내에 용해되어 있는 약 1-60중량%의 고체일 것이다. 제형들은 일반적으로 약 102 부터 약 104 세포들/mg 까지를 구비할 것이다. 이를 제형들은 헥타당 약 50mg (액체 또는 건조) 내지 1kg 또는 그 이상의 양으로 투여될 것이다.

<95> 제형들은 해충의 환경, 예컨대, 토양 및 서식지 에 분무, 산포, 스프링클립 등의 방법에 의해 적용될 수 있다.

<96> 재료 및 방법

<97> NACS (Bethesda Research Labs, Gaithersburg, MD) 컬럼 크로마토그래피가 전기용출된 DNA의 정제에 이용되었다. 그것은 버퍼를 결합용으로 0.5 × TBS/0.2M NaCl 로 변경 사용하고, 용출용으로 0.5 × TBS/2.0M NaCl로 변경 사용한 것을 제외하고는 제조자의 지시에 따라 수행되었다.

<98> α- [^{32}P] dATP에 의한 DNA의 무작위 프라이밍 라벨링은 제조자의 지시에 따라 키트 (Boehringer-Manmhein Biochemicals, Indianapolis, IN)를 이용하여 수행하였다.

<99> 겔 정제는 선택된 DNA 단편들의 정제를 위해 아가로스-TBS 겔 전기영동, 전기용출, 및 NACS 컬럼 크로마토그래피를 연속적으로 적용하는 것으로, 방법은 당업계에 잘 알려져 있다.

<100> DNA의 폴리메라제 연쇄 반응(PCR) 증폭은 다음과 같은 사이클 매개 변수에 따라 퍼킨 엘머 (Norwalk, CT) 써멀 싸이클러 상에서 25 싸이클 동안 행해졌다: 1분간 94℃, 2분간 37℃, 3분간 72℃ (각각의 72℃ 싸이클은 5초 연장 시간을 갖는다). PCR DNA 산물들은 클로닝 효율성을 증가시키기 위해 단백질 분해효소 K 처리되었다(Crowe, J.S., Cooper H.J., Smith, M.A,, Sims, M.J,, Parker, B,, Gewert, D, [1991] Nucl, Acids. Res. 19:184).

<101> 올리고데옥시리보뉴클레오티드들 (올리고뉴클레오티드들)은 Applied Biosystems(Fostei City, CA) 모델 381A DNA 합성기에 의해 합성되었다. 필요한 경우 정제는 Nensorb 컬럼 (New England Nuclear-Dupont, Wilmimgton, DE)을 이용하여 제조자의 지시에 따라 수행되었다.

<102> 슈도모나스 플루오레스센스의 일렉트로포레이션은 회전 진탕기상 30℃에서 L-육즙 (LB)에서 배양된 지수기(log-phase) 세포들에 대해 행해졌다. 세포들은 얼음-냉각된 멸균 증류수로 2회~3회 세정되고

증류수에 0.03 × 출발 용적으로 농축되었다. 1-20㎕ DNA가 50-300㎕의 세포들과 혼합되었다. Biorad Gene Pulser (Bio-Rad, Richmond, CA)에 대해 선택된 매개변수들은 0.2cm 전극간격을 갖는 큐벳(cuvette)에서 200옴, 25 마이크로패러드 및 2.25 킬로볼트였다. 일렉트로포레이션에 이어서, 1 미리리터의 LB를 첨가하여 세포들을 얼음에서 적어도 2분간 유wl하였다. 이어서 세포들을 진탕하지 않고 30℃에서 2시간 내지 밤새 인큐베이션하였다.

<103> 피. 플루오레스센스에서의 비.티. 독소의 발현은 Manual of Methods for General Bacteriology (P. Gerhardt et al., 1981, American Society for Microbiology, Washington, D.C.)에서 발견된 권장된 배지내에서 행해졌다. 글리세롤은 글루코오스 대신에 대체되었다. 처방은 수돗물로 만들어졌고 pH는 7.2로 조절되었다. 시드 플라스크들은 L-육즙으로 만들어졌다. 다음의 처방에 따랐다.

기본 배지 (for 1 liter)

글리세롤	65 g
$(NH_4)_2SO_4$	1.0 g
Na_2HPO_4	5.24 g
KH_2PO_4	2.77 g
효모 추출물	5.0 g
카사미노산	1.0 g

Metals 44 (for 100 ml)

EDTA	250 mg
$ZnSO_4 \cdot 7H_2O$	1095 mg
$FeSO_4 \cdot 7H_2O$	500 mg
$MnSO_4 \cdot H_2O$	154 mg
$CuSO_4 \cdot 5H_2O$	39.2 mg
$Co(NO_3)_2 \cdot 6H_2O$	24.8 mg
$Na_2B_4O_7 \cdot 10H_2O$	17.7 mg

침전을 지연시키기 위해 6N H_2SO_4를 몇 방울 첨가

헌터(Hunter's)의 미네랄 혼합물

니트랄로아세틱애시드	10g
(KOH에 의해 용해되고 중화된)	
$MgSO_4 \cdot 7H_2O$	14.45g
$CaCl_2 \cdot 2H_2O$	3.33g
$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$	9.25g
$FeSO_4 \cdot 7H_2O$	99mg
Metals 44	50ml
pH는 6.6~6.8로 조정	

<105> 비.티. 독소 발현의 분석을 위한 접종에서, 200㎖의 육즙 당 4ml의 헌터의 미네랄 혼합물이 첨가되었다. 이어서 플라스크에 밤새 배양한 2

용적%의 접종물을 넣었다. 배양균들을 32℃, ≥200 rpm에서 24동안

생
장시켰다. 이 시점에서, 0.75mM IPTG를 도입하고 2g 효모 추출물을 보충하였다. 48시간 및 72시간 담지된

시료에 대해 단백질 겔들을 진행하였다. 약 130KDa 단백질이 레이저 사진 농도계에 의해 정량되었다.

<106> 이하의 것들은 베스트 모드를 포함하는 본 발명을 실시하기 위한 방법을 설명하는 실시예들이다. 이러한 실시예들은 제한적으로 해석되어서는 안된다. 특별한 언급이 없는한, 모든 백분율은 중량%이고 모든 용매 혼합비는 체적비이다.

<107> 실시예 1~스플라이스 중첩 연장(SOE)에 의한 발현 백터의 변형

<108> 클로닝 벡터는 RSF1010으로 부터 유래되고, 광범위한 숙주-범위를 갖는 플라스미드인, PTJS2620 (PTJS260은 Dr. Donald Helinski, U.C. San Diego)에 기초하여 만들어질 수 있다. 벡터 제작에 이용된 시스템의 일례는 유럽특허출원 제 0 471 564호에서 찾을 수 있다. 본원에서 436 유전자로 불리우는 cry I A(c)/cry I A(b) 유전자 및 독소는 미국특허 제 5,055,294호에 기술되어 있다. pMYC1050으로 명명된 플라스미드는 420 유전자로 알려진 cry I A(c)/cry I A(b) 키메라 유전자를 포함한다. pMYC1050은 당업계에 잘 알려져 있는 방법에 의해 독소 유전자와 pM3, 130-7의 프로모터 (미국특허 제 5,055,294호에 개시된)를 pMYC467과 같은 pTJS260-계 벡터 (미국특허 제 5,169,760호에 개시된)내에 재클로닝함으로써 만들어질 수 있다. 특히, pM3,130-7 프로모터와 독소 유전자는 BamHI 내지 NdeI 단편으로 수득될 수 있고 동일한 부위(BamHI 부근 염기 12100 및 NdeI 부근 염기 8000)에 의해 결합된 단편을 대체함으로써 pMYC467 플라스미드내에 위치될 수 있다.

<109> 개량된 벡터는 이상적으로 유일한 BamHI 클로닝 부위를 포함한다. tac 프로모터(Ptac)로 부터 상류에 위치하는 플라스미드 BamHI 부위는 클레노에 의한 블런팅 및 재연결에 의해 제거될 수 있다 (제 1 도). 상기 부위의 부재는 제한효소 분해(restriction digestion)에 의해 확인될 수 있다. 이러한 과정에 따라 생산된 플라스미드를 pMYC1050∧BamHI이라 한다. 상기 구조물은 이제 SOE 돌연변이유발에 의해 플라스미드에 부가된 BamHI 부위를 갖을 수 있다. SOE 돌연변이 유발은 NsiI 독소-포함 DNA 단편을 선택가능한 마커로서 암피실린 내성(bla) 유전자를 이용하는 작은 pGEM5 (Promega Corp., Madison, WI) 벡터내로 2차 클로닝함으로써 촉진될 수 있다 (제 1 도). 단편은 제한 효소절단에 의해 배향될 수 있다. 이러한 방법에 의해 생산된 플라스미드를 pGEMtox라 부른다.

<110> 독소 코드화 부위에 있는 DNA는 제 2 도에 도시된 바와 같이 제한 효소 클로닝 부위들을 도입하기 위해 SOE의 PCR-매개 기술에 의해 돌연변이될 부 있다. 프라이머로 유용한 올리고뉴클레오티드들은 다음과 같다:

"A" (서열번호 1)

5' GCATACTAGTAGGAGATTTCCATGGATAACAATCCGAAC 3'

"B" (서열번호 2)

5' GGATCCGCTTCCCAGTCT 3'

"C" (서열번호 3)

5' AGAGAGTGGGAAGCGGATCCTACTAATCC 3'

"D" (서열번호 4)

5' TGGATACTCGATCGATATGATAATCCGT3'

"E" (서열번호 5)

5' TAATAAGAGCTCCTATGT 3'

"F" (서열번호 6)

5' TATCATATCGATCGAGTATCCAATTTAG3'

"G" (서열번호 7)

5' GTCACATAGCCAGCTGGT3'

<113> pMYC1050 DNA는 프라이머 세트들 A/B, C/D, E/D, 및 F/G를 이음함으로써 PCR 증폭을 위한 주형으로 이용되었다. 증폭된 DNA 단편들을 AB, CD, ED, 및 FG로 명명하였다. 증폭된 DNA들은 당업계에 잘 알려져 있는 아가로스-TBE 겔 전기영동, 전기용출, 및 NACS 컬럼 크로마토그래피에 의해 정제되었다. 정제된 주형 DNA들은 두 번째 세트의 PCR 반응에 이용되었다. 단편 AB 및 CD는 프라이머 A 및 B에 의해 혼합 및 증폭되었다. 별개의 반응에서, 단편 ED 및 FG는 프라이머 E 및 G와 혼함 및 증폭되었다. 증폭된 DNA는 아가로스-TBE 겔 전기영동에 의해 분석되고 상응하게 크기가 증가한 단편들은 절제, 전기용출되고 당업계에 잘 알려진 방법에 의해 NACS 컬럼상에서 절제되었다. 증폭된 DNA 단편들은 참고를 위해 AD 또는 EG라 한다.

<114> 새로운 제한효소 절단부위를 갖는 DNA 단편 AD 또는 EG는 몇가지 2차 클로닝 방법에 의해 독소-포함 DNA내에 도입되었다 (제 2 도 및 제 3 도). pGEMtox는 ClaI 또는 Hind III에 의해 절단되었다. 벡터-포함 DNA는 겔-정제되었다. 단편 AD는 ClaI에 의해 절단되고 ClaI-절단 pGSMtox 벡터 DNA에 연결되었다.

단편 EG는 HindIII에 의해 절단되고 HindIII-절단 pGEMtox 벡터 DNA에 연결되었다. 이.콜라이 균주 NM522는 연결 혼합물들에 의해 형질전환되었다. 정확하게 조립된 구조물들은 분리된 콜로니들로 부터의 플라스미드 DNA의 제한효소 절단에 의해 동정되었다. 새로운 BamHI 부위를 갖는 플라스미드를 pGEMtox BamHI 이라 칭한다. 새로운 PvuI 부위를 갖는 플라스미드를 pGEMtoxPvuI라 칭하였다. 플라스미드 pGEMtox BamHI으로 부터의 BamHI 부위를 포함하는 ClaI 단편은 pGEMtoxPvuI으로 부터의 포스파타제 ClaI 벡터-포함 단편에 연결되었다. 이.콜라이 균주 NM522는 연결 혼합물들에 의해 형질전환되었다. 정확하게 조립된 구조물들은 프라이머 세트 C/D에 의해 PCR 분석 및 제한절단에 의해 확인되었다. 양자의 새로운 제한 효소 부위들을 갖는 플라스미드는 pGEMtox bamHI/PvuI이라 칭한다.

<115> 완전한 발현 벡터는 pGEMtox bamHI/PvuI으로 부터의 삽입물과 pMYC1050∧BamHI으로 부터의 벡터에 의해 조립되었다 (제 3 도 ALC 제 4 도). 겔-정제된 삽일물은 NsiI 절단 및 ScaI 절단(오염 벡터를 제거하기 위해)에 의해 pGEMtox bamHI/PvuI으로 부터 마련되었다. 그것은 겔-정제된 NsiI-절단 벡터-포함 pMYC1050∧Bam HI DNA에 연결되었다. 이. 콜라이 균주 NM622는 연결 혼합물에 의해 형질전환되고 형질전환 혼합물들은 12㎍/m1의 테트라사이클린을 포함하는 LB 아가상에 플레이팅되었다. NdiI 삽입물을 포함하는 콜리니들은 콜로니 혼성화 및 자동방사선사진법에 의해 확인되었다. 삽입물들은 NsiI 클로닝 부위를 연결하는 프라이머 세트 A/D를 이용한 PCR 및 아가로스-TBE 겔 전기영동을 이용함으로써 배향되었다. 정확하게 조립된 플라스미드는 pMYC2224로 불리운다. 독소의 DNA 및 단백질 서열은 각각 서열번호 18 및 19에 나타내었다. 락토오스-유도성 피. 플루오레스센스 균주는 정확하게 조립된 플라스미드 DNA에 의해 일렉트로포레이션되었다. 형질전환 혼합물들은 20-㎍/m1의 테트라사이클린을 포함하는 LB 아가 상에 플레이팅하였다. 플라스미드 DNA는 그 다음의 클ㄹ로닝 실험들에 이용하기 위해 피. 플루오레스센스로 부터 준비되었다.

<116> 실시예 2 ~ crylF 초가변 부위의 pMYC2224 내로의 2차 클로닝

<117> crylF (pMYC1260)으로 부터의 초가변 부위를 포함하는 DNA 단편은 pMYC2224로 부터의 BamHI-PvuI 독소-포함 DNA 단편 대신에 교환되었다 (제 4 도). 암호화 서열이 기존의 BamHI 부위를 포함하기 때문에 BglII가 클로닝을 위해 선택되었다. BamHI 및 BglII의 4-염기 돌출부분들이 서로 상보적이기 때문에 두 부위들을 접합부로 부터 제거함과 동시에 연결을 제공한다. PCR을 위해 새로운 프라이머를 합성할 필요가 있다:

"L" (서열번호 8)

5' GAGTGGGAAGCAGATCTTAATAATGCACAATTAAGG 3'

<119> 독소-포함 DNA 단편은 주형 pMYC1260에 대해 프라이머 L/D를 이용하여 PCR에 의해 생성되었다. DNA는 2차 클로닝을 위해 BglII및 PvuI에 의해 절단되었다. tetAR자리가 많은 PvuI 부위를 포함하기 때문에, 두 개의 서로 다른 단편들에 대해 벡터-포함 DNA를 분리할 필요가 있었다. 첫번째 단편을 수득하기 위해, pMYC2224는 Bam HI × BstEII에 의해 절단되었고, Ptac-tetAR-자리-rep 기능들을 포함하는 큰 DNA 단편은 겔-정제되었다. 두 번째 단편을 수득하기 위해 pMYC2224는 BstEII ×PvuI으로 절단되었고, 벡터-프로톡신 모듈을 포함하는 DNA 단편은 겔-정제되었다. 3-피스 연결체가 구성되어, 이. 콜라이 균주 NM522 형질인전환에 이용되었다. 정확한 플라스미드는 BamHI/BglII 융합 접합부를 연결하는 프라이머 세트 N/O를 이용함으로써 PCR 분석 및 아가로스-TBE 겔 전기영동에 의해 확인되었다.

"N" (*tac* 프로모터) (서열번호 9)

5' TTAATCATCGGCTCGTA 3'

"O" (서열번호 10)

5' ACTCGATCGATATGATA(GA)TCCGT 3'

<121> 정확한 플라스미드는 pMYC 2239로 명명되었다. 그것은 아미노 말단의 crylA(c), 독소/프로톡신 접합부 까지의 crylA(b)로 구성된다. 독소 DNA 및 단백질 서열들은 각각 서열번호 20 및 21이다.

<122> 실시예 3 ~ 피. 플루오레스센스 발현 플라스미드 pMYC1260 및 pMYC2047의 구성

<123> 클로닝된 독소 유전자 crylF는 다음과 같은 방법에 의해 피. 플루오레스센스에서의 발현을 위해 변형될 수 있다:

<124> 1. pTJS1260-유래 벡터(Dr. Donald Helinski, U.C. San Diego)에 있는 pKK223-rrnB 말단 서열들을 포함하는 플라스미드는 pKK223-3(Pharmacia E. col: vector)로 부터의 Ptac 프로모터와 rrnB 터미네이터를 포함하는 BamHI-ScaI 단편을 pMYC1197 (유럽특허 0 417 564호에 기술)의 BamHI 내지 블런트 KpnI 단편내에 연결함으로써 만들어질 수 있다. 조립된 플라스미드는 이.콜라이의 형질전환 및 테트라 사이클린 선택하의 생장에 이어 재생된다.

<125> 2. Ptac-프로포티드 crylF 독소 유전자를 포함하는 플라스미드는 pMYC1603(NRRL B-18517로 부터의)으로 부터의 약 3800bp의 독소 유전자-포함 Nde I-Nde I-단편(DNA 폴리메라제 및 dNTPs를 이용하여 말단이 평활하게된)을 PKK223-3의 평활말단 EcoRI 및 HindII내에 연결함으로써 만들어질 수 있다. Ptac-프로모티드 crylF 독소 플라스미드는 이. 콜라이의 형질전환, 암피실린 선택하에서의 생장 및 당업계에 잘 알려져 있는 기술에 의한 ptac0 프로모터로 부터의 발현에 적당한 방향으로 되어있는 삽입물을 갖는 플라

스미드들을 선별검사한 다음에 재생될 수 있다.

<126> 3. Ptac-프로모티드 cry I F 독소는 플라스미드 pTJS260으로 부터의 BamHI 및 Apa I 말단들을 갖는 2.4kb DNA 단편, 상기 단계 1로 부터의 플라스미드의 복제 부위를 포함하는 8.5kb의 Apa I내지 HindIII 단편, 및 상기 단계 2로부터의 Ptae 프로모터 및 cry I F 독소 유전자를 포함하는 HindIII 내지 일부 BamHI 단편을 이용하여 3-피스 연결에 의해 pTJS260-유래 벡터내에 조립될 수 있다.

<127> 그 결과로서 수득되는 pTJS260-유래 cry I F 독소 발현 벡터 (pMYC1260)는 피. 플루오레스센스내에 일렉트로포레이션에 의해 도입될 수 있다.

<128> 4. pMYC2047은 프리이머 H 및 K를 이용하여 적당한 cry I F 주형의 PCR에 이어 SpeI 및 KpnI 절단하고 겔 정제에 의해 수득된 SpeI 내지 KpnI 단편, 단계 3의 플라스미드로 부터의 약 10kb의 Apa I 내지 Kpn I 단편 및 Ptac 프로모터를 포함하는 pMYC1197로 부터의 약 2600bp의 ApaI 내지 Spe I 단편을 연결함으로써 구성될 수 있다. 정확한 cry I F 독소 발현 플라스미드들은 슈도모나그 플루오레스센스내로의 일렉트로포레이션에 이어 플라스미드들을 제한효소 절단 함으로써 결정된다.

<129> 실시예 4 ~ cryIF/cryIA(b) 키메라의 구성

<130> 아미노말단의 cry I A(c) 분절은 간단하고 직접적인 교환에 의해 cry I F 코드화 서열에 의해 대체될 수 있다(제 5도). tetAR 자리 및 cry I F 코드화 서열 양자 모두 Apa I 부위를 포함한다. tetAR 유전자들의 일부 및 cry I F의 아미노-말단을 포함하는 작은 ApaI 단편은 pMYC2047로 부터 분리되어 pMYC2239로 부터의 큰 ApaI 벡터-포함 단편에 연결될 수 있다. 피. 플루오레센스 락토오스-유도성 균주는 연결 혼합물에 의해 일렉트로포레이션될 수 있고 20㎍/ml로 테트라사이클린을 함유하는 LB 아가 상 플레이팅될 수 있다. 락토오스-유도성 균주들은 당업자들에게 알려져 있고, 예컨대, 미국특허 제 5,169,760호에 기술되어 있다. ApaI 단편의 정확한 방향은 테트라 사이들린 저항력을 재구성한다. 이러한 방법으로 만들어진 클론은 제한 효소 절단에 의해 정확한 것으로 확인되었고 pMYC2244로 명명되었다. 독소 DNA서열은 서열 번호 22에 도시되고, 추론 단백질 서열은 서열 번호 23에 나타내었다.

<131> 실시예 5 ~ cryIF의 제한된 코돈 교정의 구성

<132> 슈도모나스종에서 코돈 이용법은 GenBank/EMBL서열 라이브러리내의 유전자들의 분석에 의해 평가하는 경우, 트리플렛 코돈의 워블 위치에서 G 또는 C를 선호한다. cry I F 유전자의 제한된 부위는 잠재성인 선호되는 워블 위치 변화를 도입하기 위해 SOE에 의해 개정된다(제 6 도). 이용된 올리고들은 다음과 같다:

"H" (서열번호 11)

5' GGACTAGTAAAAAGGAGATAACCATGGAAAATAATATTCAAAATC 3'

"I" (서열번호 12)

5' TCCAGCGGCAGGCGGCCGGTGCTGCGTTCTTCGTTCAGTATTTCTA

CTTCAGGATTATTTAAAC 3'

"J" (서열번호 13)

5' AACGCAGCACCGGCCGCCTGCCGCTGGACATCAGCCTGAGCCTTAC

ACGTTTCCTTTTGAGTGAA 3'

"K" (서열번호 14)

5' CATCAAAGGTACCTGGT3'

<134> pMYC2047 주형에 대해서 프라이머 세트 H/I 또는 J/K를 이용하여 2 가지의 서로 다른 PCB 반응이 행해졌다. 증폭된 DNA 서열들은 HI 또는 JK 명명되었다. 두번째 PCR 반응은 단편 HI 및 JK를 혼합하고 프라이머 세트 H/K에 의해 PCR 증폭시킴으로써 개시되었다. 보다 큰 SOE DNA를 겔-정제하고 SpeI ×KpnI로 절단하였다. 3-피스 연결은 SpeI-ApaI Ptac-tetAR 자리 DNA, ApaI-KpnI 벡터-프로톡신 모듈 DNA및 SpeI-KpnI PCR DNA에 의해 구성되었다. 피. 플루오레센스 락토오스-유도성 균주는 연결 혼합물에 의해 일렉트로포레이션될 수 있다. 정확한 클론들은 프라이머 세트 P/Q를 이용한 PCR 분석 및 아가로스-TBE 겔 전기영동에 의해 동정될 수 있다. 올리고 P (서열번호 15)는 야생형과 코돈-개정 유전자 사이의 구별을 위해 설계되었다.

"P" (서열번호. 15)

5' TGCCGCTGGACATCAGCCTGAG 3'

"Q" (서열번호. 16)

5' TCTAGAGCGGCCGCTTATAC(CT)CGATCGATATGATA(GA)TCCGT 3'

<136> 완전한 플라스미드는 pMYC2243으로 명명되었다. 독소 DNA 서열은 서열번호 24에 도시되어 있다. 독소 단백질 서열은 변하지 않는 것으로 추정되고, 서열번호 25에 도시되어 있다.

<137> 실시예 6 ~ 제한된 코돈 개정을 포함하는 cry I F/cry I A(b) 키메라의 제작

<138> 구조물은 상기 pMYC2244 (cry I F/cry I A(b))에 대한 것과 동일한 ApaI 단편 교환 전략을 이용함으로써 조립되었다(제 7도). 작은 독소 -tetAR 자리 ApaI DNA 단편은 pMYC2243으로 부터 겔-정제되었다. 큰 벡터 - 프로톡신 모듈 ApaI DNA 단편은 pMYC2244로부터 겔 - 정제되었다. 완전한 플라스미드는pMYC2523으로 명명되었다. 추론 DNA 및 단백질 서열은 각각 서열변호 26 및 27이다.

<139> 실시예 7 ~ pMYC2243및 pMYC2523으로 부터의 독소의 비교 발현

<140> 피. 플루오레스센스에서의 독소 발현은 전술한 바와 같이 분석되었다. 유도후 24시간 및 48시간에, pMYC2523-포함 균주는 pMYC2243-포함 균주에 비해 많은 독소를 생산하였다. 스포돕테라 엑시구아에 대한 독소 특이성 활성은 통계적으로 변하지 않았다.

<141> 실시예 8 ~ 제한된 코돈 개정을 포함하는 cry I F/436 키메라의 구성

<142> 두번째 유형의 키메라 독소는 pMYC2523에서 cry I A(b) 프로톡신 대신에 436 프로톡신 모듈로 대체함으로써 조립되었다 (제 8 도). 436 프로톡신 서열은 C- 말단자체를 제외하고는 cry I A(c) 서열로 구성된다(미국특허 5,128,130 및 5,169,760호 참조, 본원에 전문이 참고로 첨부). 클로닝을 위한 프로톡신 DNA는 주형으로 pMYC4679(미국특허 5,169,760호)와 같은 플라스미드를 이용하여 프라이머 세트 F/M에 의해 PCR에 의해 생성되었다.

"M" (서열번호 17)

5' AGGCTTCCATAGATACCTTGTGCG 3'

<144> PCR DNA는 PvuI × BstEII에 의해 절단되었다. 3-피스 연결은 pMYC2523으로 부터의 SpeI-PvuI 독소 DNA, pMYC2523으로 부터의 SpeI-BstEII 벡터 DNA및 PvuI-BstEII PCR 프로톡신 모듈 DNA에 의해 시작되었다. 락토오스 - 유도 피. 플루오레그센스 균주는 연결 혼합물에 의해 일렉트로포레이션 되었다. 정확한 플라스미드는 프라이머 세트 F/G를 이용하여 PCR하고, 아가로스 - TBE 겔 전기영동에 의해 미미한 크기 증가를 스크리닝함으로써 동정되었다. 구조물은 pMYC25S4로 명명되었다. 추론 DNA 및 단백질 서열은 각각 서열번호 28 및 29에 나타내었다.

<145> 실시예 9 ~ pMYC2243 및 pMYC2254로 부터의 독소들의 비교 발현

<146> 피. 플루오레스센스에서의 독소 발현은 상술한 바와 같이 분석되었다. pMYC2254로 부터의 독소 발현은 pMYCS243의 발현을 능가하여 증가되었다.

<147> 실시예 10 ~ 키메라 독소를 코드화하는 유전자의 식물내 삽입

<148> 본 발명의 하나의 양상은 살충 독소를 코드화하는 유전자에 의한 식물의 형질전환이다. 형질전환된 식물들은 표적 해충의 공격에 대해 저항력을 갖는다.

<149> 키메라 독소를 코드화하는 유전자는, 본원에 기술된 바와 같이 당업계에 잘 알려져 있는 여러가지 기술들을 이용함으로써 식물 세포내에 삽입될 수 있다. 예를들어, 이. 콜라이에서의 복제 시스템 및 형질전환된 세포의 선별을 가능케하는 마커를 포함하는 다수의 클로닝벡터들은 외래 유전자들을 고등식물내에 삽입하기 위한 조제품으로 이용가능하다. 예를들어, 이러한 벡터들은 pBR32, pUC 시리즈, M13mp 시리즈, pACYC184, 등을 포함한다. 따라서, 비. 티. 독소를 코드화하는 서열은 벡터내 적당한 제한부위에 삽입될 수 있다. 그 결과 플라스미드는 이. 콜라이 내에 형질전환시키는데 이용된다. 이. 콜라이 세포들은 적당한 영양 배지에서 배양된 다음 회수되고 용해된다. 플라스미드는 재생된다. 서열분석, 제한분석, 전기영동, 및 기타 생화학-분자생물학적 방법들이 분석방법으로 일반적으로 실시된다. 각각의 조작후에, 이용된 DNA 서열은 절제되어 다음 DNA 서열에 연결될 수 있다. 각각의 플라스미드 서열은 동일하거나 또는 상이한 플라스미드에 클로닝될 수 있다. 목적 유전자를 식물에 삽입하는 방법에 따라, 다른DNA 서열들도 필요할 수 있다. 예를들어,식물 세포의 형질전환을 위해 Ti 또는 Ri 플라스미드가 이용된다면, 적어도 Ti 또는 Ri 플라스미드 T-DNA의 우측 접경, 그러나 자주 좌측 및 우측 접경이 삽입되는 유전자의 측접 부위로 연결되어야만 한다.

<150> 식물세포의 형질전환을 위한 T-DNA의 이용은 활발하게 연구되었고 유럽특허 0 120 516호: Hoekema (1985) In: The Binary Plant Vector System, Offset-durkkerij Kanters B.V., Alblasssrdam, Chapter 5; Fraley et al., Crit Rer. Plant Sci. 4:1-46: 및 An et al., (1985) EMBO J. 4:277-287에 충분히 기술되어 있다.

<151> 일단 삽입된 DNA가 게놈에 통합되면, 그것은 그곳에서 비교적 안정하게 되고, 대개 다시 분리되어 나오지 않는다. 그것은 보통 형질전환된 식물 세포에 살생물제 또는 카나마이신, G 418, 블레오마이신, 히그로마이신, 또는 클로람페니콜과 같은 항생제에 대한 내성을 부여하는 선택 마커를 포함한다. 따라서, 이용된 개개의 마커는 삽입된 DNA를 포함하지 않는 세포들 대신에 형질전환된 세포들이 선택되도록 한다.

<152> 다수의 기술들이 DNA를 식물 숙주 세포내에 삽입하는데 이용될 수 있다. 이들 기술들은 형질전환제로 아그로박테리움 투메파시엔스 또는 아그로박테리움 리조게네스를 이용하여 T-DNA에 의해 형질전환시키는 것, 융합, 주입, 또는 일레트로포레이션 및 기타 가능한 방법들은 포함한다. 만약 아그로박테리아가 형질전환을 위해 이용된다면, 삽입될 DNA는 특수한 플라스미드, 즉, 중간벡터 또는 바이너리 벡터내에 클로닝되어야 한다. 중간벡터는 T-DNA에 있는 서열과 상동성을 갖는 서열들에 의해 상동 재조합에 의해 Ti 또는 Ri 플라스미드내에 통합될 수 있다. Ti 또는 Ri 플라스미드도 T-DNA의 전이에 필요한 vir 부위를 포함한다. 중간벡터들은 스스로 아그로박테리아내에서 복제할 수 없다. 중간벡터는 헬퍼 플라스미드(접합)에 의해 아그로박테리움 투메파시엔스 내로 전이될 수 있다. 바이너리 벡터들은 스스로 이. 콜라이 및 아그로 박테리아 내에서 복제할 수 있다. 이들은 선택 마커 유전자 및 우측 및 좌측 T-DNA 접경 부위에 의해 프레임되는 링커 또는 플리커를 포함한다. 이들은 직접적으로 아그로박테리아내에 형진진활될 수 있다

(Holsters et al. [1978] Mol. Gen. Genent. 163:181~187). 숙주 세포로 이용된 아그로박테리움은 vir 부위를 갖는 플라스미드를 포함한다. vir 부위는 T-DNA를 식물 세포내로 전이하는데 필요하다. 추가의 T-DNA가 포함될 수 있다. 이와같이 형질전환된 세균은 식물세포의 형질전환에 이용된다. 식물 외식편은 DNA를 식물 세포내로 전이하기 위해 아그로박테리움 투메파시엔스 또는 아그로박테리움 리조게네스와 함께 배양될 수 있다. 이어서 전체 식물들은 적당한 배지에서 감염된 식물 재료(예컨대, 잎 조각, 줄기의 분절, 뿌리)뿐만 아니라 프로토플라스트 또는 현탁액-배양 세포)로 부터 재생될 수 있다. 이어서 이렇게 해서 수득된 식물들은 삽입된 DNA의 존재여부에 대해 시험될 수 있다. 주입 또는 일렉트로포레이션의 경우에 플라스미드에 대해서는 어떠한 특별한 요구도 부가되지 않는다. 예를 들어 pUC 유도체들과 같은 통상의 플라스미드들을 이용하는 것이 가능하다.

<153> 형질전환된 세포들은 식물내에서 보통 방법으로 생장한다. 그들은 배아세포를 형성할 수 있고 형질전환된 특성을 자손 식물로 전달할 수 있다. 이러한 식물들은 보통의 방법으로 생장되고 동일한 형질전환된 우전적 요소 또는 다른 유전적 요소들은 갖는 식물들과 교배될 수 있다. 그 결과의 잡종 개체들은 유사한 표현형 특성을 갖는다.

<154> 본 발명의 바람직한 실시양태에 있어서, 식물들은 코돈 용법이 식물에 최적화된 유전자들에 의해 형질전환될 것이다. 또한, 바람직하게, 절단된 독소를 코드화하는 독소가 이용될 것이다. 절단된 독소는 일반적으로 전체 길이 독소의 약 55% 내지 약 80%를 코드화한다. 식물에 이용하기 위한 합성 유전자를 만드는 방범은 당업계에 공지되어 있다.

<155> 실시예 11 ~ 키메라 독소를 코드화하는 유전자의 곤충 바이러스내로의 클로닝

<156> 다수의 바이러스들이 곤충들을 감염시키는 것으로 알려져 있다. 예를들어, 이들 바이러스들은 바큘로바이러스 및 엔토모폭시바이러스들을 포함한다. 본 발명의 하나의 실시양태에 일어서, 살충 독소를 코드화하는 유전자들은, 본원에 기술된 바와 같이, 곤충 바이러스의 계놈내에 위치되어, 바이러스의 병원성을 증가시절 수 있다. 키메라 독소 유전자를 포함하는 곤충 바이러스를 구성하는 방법은 본 발명이 속하는 기술분야의 당업자들에게 잘 알려져 있고 이들에 의해 쉽게 실시될 수 있다. 이러한 방법들은, 예컨대 Merryweather 등 (Merryweather, A.T., U. Weyer M.P.G. Harris, M. Hirst, T. Booth, R.D. Possee (1990) J. Gen. Virol. 71:1535-1544) 및 Martens등 (Martens, J.W.M., G. Honee, D. Zuidema, J.W.M. van tent, B. Visser, J.M. Vlak (1990) APPl. Environmental Microbiol. 56(9): 2764-2770에 기술되어 있다.

<157> 본원에 기술된 실시예 및 실시양태들은 설명만을 위한 것이며, 더욱이 본 발명에 대한 다수의 변형 또는 변화들이 당업자들에게 제안될 것이며 본원의 정신과 범위 및 청부된 특허청구의 범위내에 포함되어야 한다.

서열표

(1) 일반 정보:

(i) 출원인정보:

출 원 인 : 마이코젠 코퍼레이션
거리주소 : 5501 오리린 드라이브
도 시 : 샌 디에고
주 : 캘리포니아
국 가 : 미합중국
우편번호 : 92121
tel. : (619) 453-8030 fax. : (619)453-6991

(ii) 발명의 명칭: 슈도모나스 플루오레스센스에서 델타-엔도톡신 발현의 증가

(iii) 서열의 수: 34

(iv) 통신주소:

(A) 수신인: 데이비드 알. 살리완칙
(B) 거리: 2421 노스웨스트 41st 스트리트, 슈트 A-1
(C) 도시: 게이네스빌
(D) 주: 플로리다
(E) 나라: 미합중국
(F) 우편번호: 32606

(v) 컴퓨터 판독 형태:

(A) 미디엄 타입: 플로피 디스크
(B) 컴퓨터: IBM PC 호환기종
(C) 작동 시스템: PC-DOS / MS-DOS
(D) 소프트웨어: PatentIn 릴리스 #1.0, 버젼 #1.25

(vi) 현출원 데이타

(A) 출원번호: 미합중국
(B) 출원일:
(C) 분류:

(viii) 변호사/대리인 정보:

(A) 성명: 데이비드 알. 살리완칙
(B) 등록번호: 31,794
C): 참고/도킷 번호: MA 83

(ix) 전기통신 정보:

(A) 전화번호: (904) 375-8100
(B) 팩스번호: (904) 372-5800

(2) 서열 번호 1에 대한 정보:

(i) 서열의 특징:

(A) 서열의 길이 : 39 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형

(ii) 분자 유형 : DNA (게놈)

(xi) 서열: 서열 번호 1:

GCATACTAGT AGGAGATTTC CATGGATAAC AATCCGAAC 39

(2) 서열 번호 2에 대한 정보:

(i) 서열의 특징:

(A) 서열의 길이 : 18 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 2:

GGATCCGCTT CCCAGTCT 18

(2) 서열 번호 3에 대한 정보:
(i) 서열의 특징:
(A) 서열의 길이 : 29 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 3:

AGAGAGTGGG AAGCGGATCC TACTAATCC 29

(2) 서열 번호 4에 대한 정보 :
(i) 서열의 특징:
(A) 서열의 길이 : 28 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 4:

TGGATACTCG ATCGATATGA TAATCCGT 28

(2) 서열 번호 5에 대한 정보 :
(i) 서열의 특징:
(A) 서열의 길이 : 18 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 5 :

TAATAAGAGC TCCTATGT 18

(2) 서열 번호 6에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 28 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 6 :

TATCATATCG ATCGAGTATC CAATTTAG 28

(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 11 :

GGACTAGTAA AAAGGAGATA ACCATGGAAA ATAATATTCA AAATC 45

(2) 서열 번호 12에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 64 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 12 :

TCCAGCGGCA GGCGGCCGGT GCTGCGTTCT TCGTTCAGTA TTTCTACTTC AGGATTATTT 60
AAAC 64

(2) 서열 번호 13에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 65 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 13 :

AACGCAGCAC CGGCCGCCTG CCGCTGGACA TCAGCCTGAG CCTTACACGT TTCCTTTTGA 60
GTGAA 65

(2) 서열 번호 14에 대한 정보 :
(i) 서열의 특징:
(A) 서열의 길이 : 17 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)

(xi) 서열 : 서열 번호 14 :

CATCAAAGGT ACCTGGT 17

(2) 서열 번호 15에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 22 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)

CCGGATTCTA GAGATTGGGT AAGGTATAAT CAATTTAGAA GAGAATTAAC ACTAACTGTA 720

TTAGATATCG TTGCTCTGTT CCCGAATTAT GATAGTAGAA GATATCCAAT TCGAACAGTT 780

TCCCAATTAA CAAGAGAAAT TTATACAAAC CCAGTATTAG AAAATTTTGA TGGTAGTTTT 840

CGAGGCTCGG CTCAGGGCAT AGAAAGAAGT ATTAGGAGTC CACATTTGAT GGATATACTT 900

AACAGTATAA CCATCTATAC GGATGCTCAT AGGGGTTATT ATTATTGGTC AGGGCATCAA 960

ATAATGGCTT CTCCTGTAGG GTTTTCGGGG CCAGAATTCA CTTTTCCGCT ATATGGAACT 1020

ATGGGAAATG CAGCTCCACA ACAACGTATT GTTGCTCAAC TAGGTCAGGG CGTGTATAGA 1080

ACATTATCGT CCACTTTATA TAGAAGACCT TTTAATATAG GGATAAATAA TCAACAACTA 1140

TCTGTTCTTG ACGGGACAGA ATTTGCTTAT GGAACCTCCT CAAATTTGCC ATCCGCTGTA 1200

TACAGAAAAA GCGGAACGGT AGATTCGCTG GATGAAATAC CGCCACAGAA TAACAACGTG 1260

CCACCTAGGC AAGGATTTAG TCATCGATTA AGCCATGTTT CAATGTTTCG TTCAGGCTTT 1320

AGTAATAGTA GTGTAAGTAT AATAAGAGCT CCTATGTTCT CTTGGATACA TCGTAGTGCT 1380

GAATTTAATA ATATAATTCC TTCATCACAA ATTACACAAA TACCTTTAAC AAAATCTACT 1440

AATCTTGGCT CTGGAACTTC TGTCGTTAAA GGACCAGGAT TTACAGGAGG AGATATTCTT 1500

CGAAGAACTT CACCTGGCCA GATTTCAACC TTAAGAGTAA ATATTACTGC ACCATTATCA 1560

CAAAGATATC GGGTAAGAAT TCGCTACGCT TCTACCACAA ATTTACAATT CCATACATCA 1620

ATTGACGGAA GACCTATTAA TCAGGGGAAT TTTTCAGCAA CTATGAGTAG TGGGAGTAAT 1680

TTACAGTCCG GAAGCTTTAG GACTGTAGGT TTTACTACTC CGTTTAACTT TTCAAATGGA 1740

TCAAGTGTAT TTACGTTAAG TGCTCATGTC TTCAATTCAG GCAATGAAGT TTATATAGAT 1800

CGAATTGAAT TTGTTCCGGC AGAAGTAACC TTTGAGGCAG AATATGATTT AGAAAGAGCA 1860

CAAAAGGCGG TGAATGAGCT GTTTACTTCT TCCAATCAAA TCGGGTTAAA AACAGATGTG 1920

ACGGATTATC ATATCGATCG AGTATCCAAT TTAGTTGAGT GTTTATCTGA TGAATTTTGT 1980

CTGGATGAAA AAAAAGAATT GTCCGAGAAA GTCAAACATG CGAAGCGACT TAGTGATGAG 2040

CGGAATTTAC TTCAAGATCC AAACTTTAGA GGGATCAATA GACAACTAGA CCGTGGCTGG 2100

AGAGGAAGTA CGGATATTAC CATCCAAGGA GGCGATGACG TATTCAAAGA GAATTACGTT 2160

ACGCTATTGG GTACCTTTGA TGAGTGCTAT CCAACGTATT TATATCAAAA AATAGATGAG 2220

TCGAAATTAA AAGCCTATAC CCGTTACCAA TTAAGAGGGT ATATCGAAGA TAGTCAAGAC 2280

TTAGAAATCT ATTTAATTCG CTACAATGCC AAACACGAAA CAGTAAATGT GCCAGGTACG 2340

GGTTCCTTAT GGCCGCTTTC AGCCCCAAGT CCAATCGGAA AATGTGCCCA TCATTCCCAT 2400

CATTTCTCCT TGGACATTGA TGTTGGATGT ACAGACTTAA ATGAGGACTT AGGTGTATGG 2460

GTGATATTCA AGATTAAGAC GCAAGATGGC CATGCAAGAC TAGGAAATCT AGAATTTCTC 2520

GAAGAGAAAC CATTAGTAGG AGAAGCACTA GCTCGTGTGA AAAGAGCGGA GAAAAAATGG 2580

AGAGACAAAC GTGAAAAATT GGAATGGGAA ACAAATATTG TTTATAAAGA GGCAAAAGAA 2640

TCTGTAGATG CTTTATTTGT AAACTCTCAA TATGATAGAT TACAAGCGGA TACCAACATC 2700

(xi) 서열 : 서열 번호 15 :

TGCCGCTGGA CATCAGCCTG AG 22

(2) 서열 번호 16에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 41 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 16 :

TCTAGAGCGG CCGCTTATAC YCGATCGATA TGATARTCCG T 41

(2) 서열 번호 17에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 24 염기
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 17 :

AGGCTTCCAT AGATACCTTG TGCG 24

(2) 서열 번호 18에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 3465 염기쌍
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 18 :

ATGGATAACA ATCCGAACAT CAATGAATGC ATTCCTTATA ATTGTTTAAG TAACCCTGAA 60

GTAGAAGTAT TAGGTGGAGA AAGAATAGAA ACTGGTTACA CCCCAATCGA TATTTCCTTG 120

TCGCTAACGC AATTTCTTTT GAGTGAATTT GTTCCCGGTG CTGGATTTGT GTTAGGACTA 180

GTTGATATAA TATGGGGAAT TTTTGGTCCC TCTCAATGGG ACGCATTTCT TGTACAAATT 240

GAACAGTTAA TTAACCAAAG AATAGAAGAA TTCGCTAGGA ACCAAGCCAT TTCTAGATTA 300

GAAGGACTAA GCAATCTTTA TCAAATTTAC GCAGAATCTT TTAGAGAGTG GGAAGCGGAT 360

CCTACTAATC CAGCATTAAG AGAAGAGATG CGTATTCAAT TCAATGACAT GAACAGTGCC 420

CTTACAACCG CTATTCCTCT TTTTGCAGTT CAAAATTATC AAGTTCCTCT TTTATCAGTA 480

TATGTTCAAG CTGCAAATTT ACATTTATCA GTTTTGAGAG ATGTTTCAGT GTTTGGACAA 540

AGGTGGGGAT TTGATGCCGC GACTATCAAT AGTCGTTATA ATGATTTAAC TAGGCTTATT 600

GGCAACTATA CAGATTATGC TGTACGCTGG TACAATACGG GATTAGAACG TGTATGGGGA 660

GCGATGATTC ATGCGGCAGA TAAACGCGTT CATAGCATTC GAGAAGCTTA TCTGCCTGAG 2760

CTGTCTGTGA TTCCGGGTGT CAATGCGGCT ATTTTTGAAG AATTAGAAGG GCGTATTTTC 2820

ACTGCATTCT CCCTATATGA TGCGAGAAAT GTCATTAAAA ATGGTGATTT TAATAATGGC 2880

TTATCCTGCT GGAACGTGAA AGGGCATGTA GATGTAGAAG AACAAAACAA CCACCGTTCG 2940

GTCCTTGTTG TTCCGGAATG GGAAGCAGAA GTGTCACAAG AAGTTCGTGT CTGTCCGGGT 3000

CGTGGCTATA TCCTTCGTGT CACAGCGTAC AAGGAGGGAT ATGGAGAAGG TTGCGTAACC 3060

ATTCATGAGA TCGAGAACAA TACAGACGAA CTGAAGTTTA GCAACTGTGT AGAAGAGGAA 3120

GTATATCCAA ACAACACGGT AACGTGTAAT GATTATACTG CGACTCAAGA AGAATATGAG 3180

GGTACGTACA CTTCTCGTAA TCGAGGATAT GACGGAGCCT ATGAAAGCAA TTCTTCTGTA 3240

CCAGCTGATT ATGCATCAGC CTATGAAGAA AAAGCATATA CAGATGGACG AAGAGACAAT 3300

CCTTGTGAAT CTAACAGAGG ATATGGGGAT TACACACCAC TACCAGCTGG CTATGTGACA 3360

AAAGAATTAG AGTACTTCCC AGAAACCGAT AAGGTATGGA TTGAGATCGG AGAAACGGAA 3420

GGAACATTCA TCGTGGACAG CGTGGAATTA CTTCTTATGG AGGAA 3465

(2) 서열 번호 19에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 1155 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 19 :

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
1 5 10 15

Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
20 25 30

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
35 40 45

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
85 90 95

```
Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
            100                 105                 110

Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
        115                 120                 125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
    130                 135                 140

Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
145                 150                 155                 160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
                165                 170                 175

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
            180                 185                 190

Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val
        195                 200                 205

Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg
    210                 215                 220

Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
225                 230                 235                 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro
                245                 250                 255

Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val
            260                 265                 270

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu
        275                 280                 285

Arg Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
    290                 295                 300

Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Tyr Trp Ser Gly His Gln
305                 310                 315                 320

Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro
                325                 330                 335

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp
    370                 375                 380

Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val
385                 390                 395                 400

Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
                405                 410                 415

Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His
            420                 425                 430

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445
```

```
Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
    450                 455                 460

Ile Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr
465                 470                 475                 480

Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly
                485                 490                 495

Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg
            500                 505                 510

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

Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg
    530                 535                 540

Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn
545                 550                 555                 560

Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn
                565                 570                 575

Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn
            580                 585                 590

Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu
        595                 600                 605

Val Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val
    610                 615                 620

Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val
625                 630                 635                 640

Thr Asp Tyr His Ile Asp Arg Val Ser Asn Leu Val Glu Cys Leu Ser
                645                 650                 655

Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys
            660                 665                 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
        675                 680                 685

Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr
    690                 695                 700

Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
705                 710                 715                 720

Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
                725                 730                 735

Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg
            740                 745                 750

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
        755                 760                 765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
    770                 775                 780

Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His
785                 790                 795                 800

His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp
                805                 810                 815
```

```
Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala
            820                 825                 830

Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu
        835                 840                 845

Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg
    850                 855                 860

Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu
865                 870                 875                 880

Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala
                885                 890                 895

Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser
            900                 905                 910

Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn
        915                 920                 925

Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser
    930                 935                 940

Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly
945                 950                 955                 960

Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn
                965                 970                 975

Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser
            980                 985                 990

Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr
        995                 1000                1005

Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile
    1010                1015                1020

Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu
1025                1030                1035                1040

Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Thr Gln
                1045                1050                1055

Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Gly
            1060                1065                1070

Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr Ala Ser Ala Tyr
        1075                1080                1085

Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp Asn Pro Cys Glu Ser
    1090                1095                1100

Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr
1105                1110                1115                1120

Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
                1125                1130                1135

Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu
            1140                1145                1150

Met Glu Glu
        1155
```

(2) 서열 번호 20에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 3450 염기쌍
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)
(xi) 서열 : 서열 번호 20 :

```
ATGGATAACA ATCCGAACAT CAATGAATGC ATTCCTTATA ATTGTTTAAG TAACCCTGAA      60
GTAGAAGTAT TAGGTGGAGA AAGAATAGAA ACTGGTTACA CCCCAATCGA TATTTCCTTG     120
TCGCTAACGC AATTTCTTTT GAGTGAATTT GTTCCCGGTG CTGGATTTGT GTTAGGACTA     180
GTTGATATAA TATGGGGAAT TTTTGGTCCC TCTCAATGGG ACGCATTTCT TGTACAAATT     240
GAACAGTTAA TTAACCAAAG AATAGAAGAA TTCGCTAGGA ACCAAGCCAT TTCTAGATTA     300
GAAGGACTAA GCAATCTTTA TCAAATTTAC GCAGAATCTT TTAGAGAGTG GGAAGCGGAT     360
CTTAATAATG CACAATTAAG GGAAGATGTG CGTATTCGAT TTGCTAATAC AGACGACGCT     420
TTAATAACAG CAATAAATAA TTTTACACTT ACAAGTTTTG AAATCCCTCT TTTATCGGTC     480
TATGTTCAAG CGGCGAATTT ACATTTATCA CTATTAAGAG ACGCTGTATC GTTTGGGCAG     540
GGTTGGGGAC TGGATATAGC TACTGTTAAT AATCATTATA ATAGATTAAT AAATCTTATT     600
CATAGATATA CGAAACATTG TTTGGACACA TACAATCAAG GATTAGAAAA CTTAAGAGGT     660
ACTAATACTC GACAATGGGC AAGATTCAAT CAGTTTAGGA GAGATTTAAC ACTTACTGTA     720
TTAGATATCG TTGCTCTTTT TCCGAACTAC GATGTTAGAA CATATCCAAT TCAAACGTCA     780
TCCCAATTAA CAAGGGAAAT TTATACAAGT TCAGTAATTG AGGATTCTCC AGTTTCTGCT     840
    ATACCTA ATGGTTTTAA TAGGGCGGAA TTTGGAGTTA GACCGCCCCA TCTTATGGAC     900
    ATGAATT CTTTGTTTGT AACTGCAGAG ACTGTTAGAA GTCAAACTGT GTGGGGAGGA     960
CACTTAGTTA GTTCACGAAA TACGGCTGGT AACCGTATAA ATTTCCCTAG TTACGGGGTC    1020
TTCAATCCTG GTGGCGCCAT TTGGATTGCA GATGAGGATC CACGTCCTTT TTATCGGACA    1080
TTATCAGATC CTGTTTTTGT CCGAGGAGGA TTTGGGAATC CTCATTATGT ACTGGGGCTT    1140
AGGGGAGTAG CATTTCAACA AACTGGTACG AACCACACCC GAACATTTAG AAATAGTGGG    1200
ACCATAGATT CTCTAGATGA AATCCCACCT CAGGATAATA GTGGGGCACC TTGGAATGAT    1260
TATAGTCATG TATTAAATCA TGTTACATTT GTACGATGGC CAGGTGAGAT TTCAGGAAGT    1320
GATTCATGGA GAGCTCCAAT GTTTTCTTGG ACGCACCGTA GTGCAACCCC TACAAATACA    1380
ATTGATCCGG AGAGGATTAC TCAAATACCA TTGGTAAAAG CACATACACT TCAGTCAGGT    1440
ACTACTGTTG TAAGAGGGCC CGGGTTTACG GGAGGAGATA TTCTTCGACG AACAAGTGGA    1500
GGACCATTTG CTTATACTAT TGTTAATATA AATGGGCAAT TACCCCAAAG GTATCGTGCA    1560
AGAATACGCT ATGCCTCTAC TACAAATCTA AGAATTTACG TAACGGTTGC AGGTGAACGG    1620
ATTTTTGCTG GTCAATTTAA CAAAACAATG GATACCGGTG ACCCATTAAC ATTCCAATCT    1680
TTTAGTTACG CAACTATTAA TACAGCTTTT ACATTCCCAA TGAGCCAGAG TAGTTTCACA    1740
```

GTAGGTGCTG ATACTTTTAG TTCAGGGAAT GAAGTTTATA TAGACAGATT TGAATTGATT 1800

CCAGTTACTG CAACATTTGA AGCAGAATAT GATTTAGAAA GAGCACAAAA GGCGGTGAAT 1860

GCGCTGTTTA CTTCTATAAA CCAAATAGGG ATAAAAACAG ATGTGACGGA TTATCATATC 1920

GATCGAGTAT CCAATTTAGT TGAGTGTTTA TCTGATGAAT TTTGTCTGGA TGAAAAAAAA 1980

GAATTGTCCG AGAAAGTCAA ACATGCGAAG CGACTTAGTG ATGAGCGGAA TTTACTTCAA 2040

GATCCAAACT TTAGAGGGAT CAATAGACAA CTAGACCGTG GCTGGAGAGG AAGTACGGAT 2100

ATTACCATCC AAGGAGGCGA TGACGTATTC AAAGAGAATT ACGTTACGCT ATTGGGTACC 2160

TTTGATGAGT GCTATCCAAC GTATTTATAT CAAAAAATAG ATGAGTCGAA ATTAAAAGCC 2220

TATACCCGTT ACCAATTAAG AGGGTATATC GAAGATAGTC AAGACTTAGA AATCTATTTA 2280

ATTCGCTACA ATGCCAAACA CGAAACAGTA AATGTGCCAG GTACGGGTTC CTTATGGCCG 2340

CTTTCAGCCC CAAGTCCAAT CGGAAAATGT GCCCATCATT CCCATCATTT CTCCTTGGAC 2400

ATTGATGTTG GATGTACAGA CTTAAATGAG GACTTAGGTG TATGGGTGAT ATTCAAGATT 2460

AAGACGCAAG ATGGCCATGC AAGACTAGGA AATCTAGAAT TTCTCGAAGA GAAACCATTA 2520

GTAGGAGAAG CACTAGCTCG TGTGAAAAGA GCGGAGAAAA AATGGAGAGA CAAACGTGAA 2580

AAATTGGAAT GGGAAACAAA TATTGTTTAT AAAGAGGCAA AAGAATCTGT AGATGCTTTA 2640

TTTGTAAACT CTCAATATGA TAGATTACAA GCGGATACCA ACATCGCGAT GATTCATGCG 2700

GCAGATAAAC GCGTTCATAG CATTCGAGAA GCTTATCTGC CTGAGCTGTC TGTGATTCCG 2760

GGTGTCAATG CGGCTATTTT TGAAGAATTA GAAGGGCGTA TTTTCACTGC ATTCTCCCTA 2820

TATGATGCGA GAAATGTCAT TAAAAATGGT GATTTTAATA ATGGCTTATC CTGCTGGAAC 2880

GTGAAAGGGC ATGTAGATGT AGAAGAACAA AACAACCACC GTTCGGTCCT TGTTGTTCCG 2940

GAATGGGAAG CAGAAGTGTC ACAAGAAGTT CGTGTCTGTC CGGGTCGTGG CTATATCCTT 3000

CGTGTCACAG CGTACAAGGA GGGATATGGA GAAGGTTGCG TAACCATTCA TGAGATCGAG 3060

AACAATACAG ACGAACTGAA GTTTAGCAAC TGTGTAGAAG AGGAAGTATA TCCAAACAAC 3120

ACGGTAACGT GTAATGATTA TACTGCGACT CAAGAAGAAT ATGAGGGTAC GTACACTTCT 3180

CGTAATCGAG GATATGACGG AGCCTATGAA AGCAATTCTT CTGTACCAGC TGATTATGCA 3240

TCAGCCTATG AAGAAAAAGC ATATACAGAT GGACGAAGAG ACAATCCTTG TGAATCTAAC 3300

AGAGGATATG GGGATTACAC ACCACTACCA GCTGGCTATG TGACAAAAGA ATTAGAGTAC 3360

TTCCCAGAAA CCGATAAGGT ATGGATTGAG ATCGGAGAAA CGGAAGGAAC ATTCATCGTG 3420

GACAGCGTGG AATTACTTCT TATGGAGGAA 3450

(2) 서열 번호 21에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 1150 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질

(xi) 서열 : 서열 번호 21 :

```
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
1               5                  10                  15

Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
            20                  25                  30

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
        35                  40                  45

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
    50                  55                  60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
65                  70                  75                  80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
                85                  90                  95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
            100                 105                 110

Ser Phe Arg Glu Trp Glu Ala Asp Leu Asn Asn Ala Gln Leu Arg Glu
        115                 120                 125

Asp Val Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala
    130                 135                 140

Ile Asn Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val
145                 150                 155                 160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val
                165                 170                 175

Ser Phe Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His
            180                 185                 190

Tyr Asn Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu
        195                 200                 205

Asp Thr Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg
    210                 215                 220

Gln Trp Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val
225                 230                 235                 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro
                245                 250                 255

Ile Gln Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val
            260                 265                 270

Ile Glu Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg
        275                 280                 285

Ala Glu Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser
    290                 295                 300

Leu Phe Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly
305                 310                 315                 320

His Leu Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro
                325                 330                 335

Ser Tyr Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu
            340                 345                 350
```

```
Asp Pro Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg
        355                 360                 365

Gly Gly Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala
    370                 375                 380

Phe Gln Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly
385                 390                 395                 400

Thr Ile Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala
                405                 410                 415

Pro Trp Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg
            420                 425                 430

Trp Pro Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe
        435                 440                 445

Ser Trp Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu
    450                 455                 460

Arg Ile Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly
465                 470                 475                 480

Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg
                485                 490                 495

Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly
            500                 505                 510

Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr
        515                 520                 525

Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly
    530                 535                 540

Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser
545                 550                 555                 560

Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln
                565                 570                 575

Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val
            580                 585                 590

Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu Ala
        595                 600                 605

Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr
    610                 615                 620

Ser Ile Asn Gln Ile Gly Ile Lys Thr Asp Val Thr Asp Tyr His Ile
625                 630                 635                 640

Asp Arg Val Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys Leu
                645                 650                 655

Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu
            660                 665                 670

Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile Asn
        675                 680                 685

Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln
    690                 695                 700

Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Leu Gly Thr
705                 710                 715                 720
```

```
Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser
                725                 730                 735

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

Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu
        755                 760                 765

Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Pro
    770                 775                 780

Ser Pro Ile Gly Lys Cys Ala His His Ser His His Phe Ser Leu Asp
785                 790                 795                 800

Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val
                805                 810                 815

Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu
            820                 825                 830

Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val
        835                 840                 845

Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp
    850                 855                 860

Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu
865                 870                 875                 880

Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala
                885                 890                 895

Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr
            900                 905                 910

Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu
        915                 920                 925

Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg
    930                 935                 940

Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn
945                 950                 955                 960

Val Lys Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val
                965                 970                 975

Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val
            980                 985                 990

Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly
        995                 1000                1005

Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr Asp
    1010                1015                1020

Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro Asn Asn
1025                1030                1035                1040

Thr Val Thr Cys Asn Asp Tyr Thr Ala Thr Gln Glu Glu Tyr Glu Gly
                1045                1050                1055

Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Gly Ala Tyr Glu Ser Asn
            1060                1065                1070

Ser Ser Val Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr
        1075                1080                1085
```

```
Thr Asp Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly
1090                1095                1100

Asp Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr
1105                1110                1115                1120

Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly
                1125                1130                1135

Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
            1140                1145                1150
```

(2) 서열 번호 22에 대한 정보 :

(i) 서열의 특징:
- (A) 서열의 길이 : 3444 염기쌍
- (B) 서열의 형 : 핵산
- (C) 쇄의 수 : 1 본쇄
- (D) 형태 : 선형

(ii) 분자 유형 : DNA (게놈)

```
ATGGAGAATA ATATTCAAAA TCAATGCGTA CCTTACAATT GTTTAAATAA TCCTGAAGTA      60
GAAATATTAA ATGAAGAAAG AAGTACTGGC AGATTACCGT TAGATATATC CTTATCGCTT     120
ACACGTTTCC TTTTGAGTGA ATTTGTTCCA GGTGTGGGAG TTGCGTTTGG ATTATTTGAT     180
TTAATATGGG GTTTTATAAC TCCTTCTGAT TGGAGCTTAT TTCTTTTACA GATTGAACAA     240
TTGATTGAGC AAAGAATAGA AACATTGGAA AGGAACCGGG CAATTACTAC ATTACGAGGG     300
TTAGCAGATA GCTATGAAAT TTATATTGAA GCACTAAGAG AGTGGGAAGC AAATCCTAAT     360
AATGCACAAT TAAGGGAAGA TGTGCGTATT CGATTTGCTA ATACAGACGA CGCTTTAATA     420
ACAGCAATAA ATAATTTTAC ACTTACAAGT TTTGAAATCC CTCTTTTATC GGTCTATGTT     480
CAAGCGGCGA ATTTACATTT ATCACTATTA AGAGACGCTG TATCGTTTGG GCAGGGTTGG     540
GGACTGGATA TAGCTACTGT TAATAATCAT TATAATAGAT TAATAAATCT TATTCATAGA     600
TATACGAAAC ATTGTTTGGA CACATACAAT CAAGGATTAG AAAACTTAAG AGGTACTAAT     660
ACTCGACAAT GGGCAAGATT CAATCAGTTT AGGAGAGATT TAACACTTAC TGTATTAGAT     720
ATCGTTGCTC TTTTTCCGAA CTACGATGTT AGAACATATC CAATTCAAAC GTCATCCCAA     780
TTAACAAGGG AAATTTATAC AAGTTCAGTA ATTGAGGATT CTCCAGTTTC TGCTAATATA     840
CCTAATGGTT TTAATAGGGC GGAATTTGGA GTTAGACCGC CCCATCTTAT GGACTTTATG     900
AATTCTTTGT TTGTAACTGC AGAGACTGTT AGAAGTCAAA CTGTGTGGGG AGGACACTTA     960
GTTAGTTCAC GAAATACGGC TGGTAACCGT ATAAATTTCC CTAGTTACGG GGTCTTCAAT    1020
CCTGGTGGCG CCATTTGGAT TGCAGATGAG GATCCACGTC CTTTTTATCG GACATTATCA    1080
GATCCTGTTT TTGTCCGAGG AGGATTTGGG AATCCTCATT ATGTACTGGG GCTTAGGGGA    1140
GTAGCATTTC AACAAACTGG TACGAACCAC ACCCGAACAT TTAGAAATAG TGGGACCATA    1200
GATTCTCTAG ATGAAATCCC ACCTCAGGAT AATAGTGGGG CACCTTGGAA TGATTATAGT    1260
CATGTATTAA ATCATGTTAC ATTTGTACGA TGGCCAGGTG AGATTTCAGG AAGTGATTCA    1320
```

TGGAGAGCTC CAATGTTTTC TTGGACGCAC CGTAGTGCAA CCCCTACAAA TACAATTGAT 1380

CCGGAGAGGA TTACTCAAAT ACCATTGGTA AAAGCACATA CACTTCAGTC AGGTACTACT 1440

GTTGTAAGAG GGCCCGGGTT TACGGGAGGA GATATTCTTC GACGAACAAG TGGAGGACCA 1500

TTTGCTTATA CTATTGTTAA TATAAATGGG CAATTACCCC AAAGGTATCG TGCAAGAATA 1560

CGCTATGCCT CTACTACAAA TCTAAGAATT TACGTAACGG TTGCAGGTGA ACGGATTTTT 1620

GCTGGTCAAT TTAACAAAAC AATGGATACC GGTGACCCAT TAACATTCCA ATCTTTTAGT 1680

TACGCAACTA TTAATACAGC TTTTACATTC CCAATGAGCC AGAGTAGTTT CACAGTAGGT 1740

GCTGATACTT TTAGTTCAGG GAATGAAGTT TATATAGACA GATTTGAATT GATTCCAGTT 1800

ACTGCAACAT TTGAAGCAGA ATATGATTTA GAAAGAGCAC AAAAGGCGGT GAATGCGCTG 1860

TTTACTTCTA TAAACCAAAT AGGGATAAAA ACAGATGTGA CGGATTATCA TATCGATCGA 1920

GTATCCAATT TAGTTGAGTG TTTATCTGAT GAATTTTGTC TGGATGAAAA AAAAGAATTG 1980

TCCGAGAAAG TCAAACATGC GAAGCGACTT AGTGATGAGC GGAATTTACT TCAAGATCCA 2040

AACTTTAGAG GGATCAATAG ACAACTAGAC CGTGGCTGGA GAGGAAGTAC GGATATTACC 2100

ATCCAAGGAG GCGATGACGT ATTCAAAGAG AATTACGTTA CGCTATTGGG TACCTTTGAT 2160

GAGTGCTATC CAACGTATTT ATATCAAAAA ATAGATGAGT CGAAATTAAA AGCCTATACC 2220

CGTTACCAAT TAAGAGGGTA TATCGAAGAT AGTCAAGACT TAGAAATCTA TTTAATTCGC 2280

TACAATGCCA AACACGAAAC AGTAAATGTG CCAGGTACGG GTTCCTTATG GCCGCTTTCA 2340

GCCCCAAGTC CAATCGGAAA ATGTGCCCAT CATTCCCATC ATTTCTCCTT GGACATTGAT 2400

GTTGGATGTA CAGACTTAAA TGAGGACTTA GGTGTATGGG TGATATTCAA GATTAAGACG 2460

CAAGATGGCC ATGCAAGACT AGGAAATCTA GAATTTCTCG AAGAGAAACC ATTAGTAGGA 2520

GAAGCACTAG CTCGTGTGAA AAGAGCGGAG AAAAAATGGA GAGACAAACG TGAAAAATTG 2580

GAATGGGAAA CAAATATTGT TTATAAAGAG GCAAAAGAAT CTGTAGATGC TTTATTTGTA 2640

AACTCTCAAT ATGATAGATT ACAAGCGGAT ACCAACATCG CGATGATTCA TGCGGCAGAT 2700

AAACGCGTTC ATAGCATTCG AGAAGCTTAT CTGCCTGAGC TGTCTGTGAT TCCGGGTGTC 2760

AATGCGGCTA TTTTTGAAGA ATTAGAAGGG CGTATTTTCA CTGCATTCTC CCTATATGAT 2820

GCGAGAAATG TCATTAAAAA TGGTGATTTT AATAATGGCT TATCCTGCTG GAACGTGAAA 2880

GGGCATGTAG ATGTAGAAGA ACAAAACAAC CACCGTTCGG TCCTTGTTGT TCCGGAATGG 2940

GAAGCAGAAG TGTCACAAGA AGTTCGTGTC TGTCCGGGTC GTGGCTATAT CCTTCGTGTC 3000

ACAGCGTACA AGGAGGGATA TGGAGAAGGT TGCGTAACCA TTCATGAGAT CGAGAACAAT 3060

ACAGACGAAC TGAAGTTTAG CAACTGTGTA GAAGAGGAAG TATATCCAAA CAACACGGTA 3120

ACGTGTAATG ATTATACTGC GACTCAAGAA GAATATGAGG GTACGTACAC TTCTCGTAAT 3180

CGAGGATATG ACGGAGCCTA TGAAAGCAAT TCTTCTGTAC CAGCTGATTA TGCATCAGCC 3240

TATGAAGAAA AAGCATATAC AGATGGACGA AGAGACAATC CTTGTGAATC TAACAGAGGA 3300

TATGGGGATT ACACACCACT ACCAGCTGGC TATGTGACAA AAGAATTAGA GTACTTCCCA 3360

GAAACCGATA AGGTATGGAT TGAGATCGGA GAAACGGAAG GAACATTCAT CGTGGACAGC 3420

GTGGAATTAC TTCTTATGGA GGAA 3444

(2) 서열 번호 23에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 1148 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 23 :

Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
1 5 10 15

Asn Pro Glu Val Glu Ile Leu Asn Glu Glu Arg Ser Thr Gly Arg Leu
20 25 30

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
35 40 45

Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
50 55 60

Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
65 70 75 80

Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
85 90 95

Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
100 105 110

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
115 120 125

Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
130 135 140

Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
145 150 155 160

Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe
165 170 175

Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
180 185 190

Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr
195 200 205

Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
210 215 220

Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
225 230 235 240

Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln
245 250 255

Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu
260 265 270

```
Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu
        275                 280                 285

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
    290                 295                 300

Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu
305                 310                 315                 320

Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr
                325                 330                 335

Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
            340                 345                 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
        355                 360                 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
    370                 375                 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
385                 390                 395                 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
                405                 410                 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
            420                 425                 430

Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
        435                 440                 445

Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
    450                 455                 460

Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
465                 470                 475                 480

Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
                485                 490                 495

Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu
            500                 505                 510

Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu
        515                 520                 525

Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe
    530                 535                 540

Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser
545                 550                 555                 560

Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser
                565                 570                 575

Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile
            580                 585                 590

Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu Ala Glu Tyr
        595                 600                 605

Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Ile
    610                 615                 620

Asn Gln Ile Gly Ile Lys Thr Asp Val Thr Asp Tyr His Ile Asp Arg
625                 630                 635                 640
```

```
Val Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu
                645                 650                 655

Lys Lys Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp
            660                 665                 670

Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln
        675                 680                 685

Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly
    690                 695                 700

Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Leu Gly Thr Phe Asp
705                 710                 715                 720

Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu
                725                 730                 735

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

Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val
        755                 760                 765

Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Pro Ser Pro
    770                 775                 780

Ile Gly Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp
785                 790                 795                 800

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
                805                 810                 815

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
            820                 825                 830

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg
        835                 840                 845

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr
    850                 855                 860

Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val
865                 870                 875                 880

Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile
                885                 890                 895

His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro
            900                 905                 910

Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu
        915                 920                 925

Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val
    930                 935                 940

Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys
945                 950                 955                 960

Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val
                965                 970                 975

Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys Pro
            980                 985                 990

Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly
        995                 1000                1005
```

```
Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr Asp Glu Leu
    1010                1015                1020

Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro Asn Asn Thr Val
1025                1030                1035                1040

Thr Cys Asn Asp Tyr Thr Ala Thr Gln Glu Glu Tyr Glu Gly Thr Tyr
                1045                1050                1055

Thr Ser Arg Asn Arg Gly Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser
            1060                1065                1070

Val Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp
        1075                1080                1085

Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr
    1090                1095                1100

Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro
1105                1110                1115                1120

Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe
                1125                1130                1135

Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
            1140                1145
```

(2) 서열 번호 24에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 3522 염기쌍
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)

(xi) 서열 : 서열 번호 24 :

```
ATGGAAAATA ATATTCAAAA TCAATGCGTA CCTTACAATT GTTTAAATAA TCCTGAAGTA      60

GAAATACTGA ACGAAGAACG CAGCACCGGC CGCCTGCCGC TGGACATCAG CCTGAGCCTT     120

ACACGTTTCC TTTTGAGTGA ATTTGTTCCA GGTGTGGGAG TTGCGTTTGG ATTATTTGAT     180

TTAATATGGG GTTTTATAAC TCCTTCTGAT TGGAGCTTAT TTCTTTTACA GATTGAACAA     240

TTGATTGAGC AAAGAATAGA AACATTGGAA AGGAACCGGG CAATTACTAC ATTACGAGGG     300

TTAGCAGATA GCTATGAAAT TTATATTGAA GCACTAAGAG AGTGGGAAGC AAATCCTAAT     360

AATGCACAAT TAAGGGAAGA TGTGCGTATT CGATTTGCTA ATACAGACGA CGCTTTAATA     420

ACAGCAATAA ATAATTTTAC ACTTACAAGT TTTGAAATCC CTCTTTTATC GGTCTATGTT     480

CAAGCGGCGA ATTTACATTT ATCACTATTA AGAGACGCTG TATCGTTTGG GCAGGGTTGG     540

GGACTGGATA TAGCTACTGT TAATAATCAT TATAATAGAT TAATAAATCT TATTCATAGA     600

TATACGAAAC ATTGTTTGGA CACATACAAT CAAGGATTAG AAAACTTAAG AGGTACTAAT     660

ACTCGACAAT GGGCAAGATT CAATCAGTTT AGGAGAGATT TAACACTTAC TGTATTAGAT     720

ATCGTTGCTC TTTTTCCGAA CTACGATGTT AGAACATATC CAATTCAAAC GTCATCCCAA     780

TTAACAAGGG AAATTTATAC AAGTTCAGTA ATTGAGGATT CTCCAGTTTC TGCTAATATA     840

CCTAATGGTT TTAATAGGGC GGAATTTGGA GTTAGACCGC CCCATCTTAT GGACTTTATG     900
```

AATTCTTTGT TTGTAACTGC AGAGACTGTT AGAAGTCAAA CTGTGTGGGG AGGACACTTA 960

GTTAGTTCAC GAAATACGGC TGGTAACCGT ATAAATTTCC CTAGTTACGG GGTCTTCAAT 1020

CCTGGTGGCG CCATTTGGAT TGCAGATGAG GATCCACGTC CTTTTTATCG GACATTATCA 1080

GATCCTGTTT TTGTCCGAGG AGGATTTGGG AATCCTCATT ATGTACTGGG GCTTAGGGGA 1140

GTAGCATTTC AACAAACTGG TACGAACCAC ACCCGAACAT TTAGAAATAG TGGGACCATA 1200

GATTCTCTAG ATGAAATCCC ACCTCAGGAT AATAGTGGGG CACCTTGGAA TGATTATAGT 1260

CATGTATTAA ATCATGTTAC ATTTGTACGA TGGCCAGGTG AGATTTCAGG AAGTGATTCA 1320

TGGAGAGCTC CAATGTTTTC TTGGACGCAC CGTAGTGCAA CCCCTACAAA TACAATTGAT 1380

CCGGAGAGGA TTACTCAAAT ACCATTGGTA AAAGCACATA CACTTCAGTC AGGTACTACT 1440

GTTGTAAGAG GGCCCGGGTT TACGGGAGGA GATATTCTTC GACGAACAAG TGGAGGACCA 1500

TTTGCTTATA CTATTGTTAA TATAAATGGG CAATTACCCC AAAGGTATCG TGCAAGAATA 1560

CGCTATGCCT CTACTACAAA TCTAAGAATT TACGTAACGG TTGCAGGTGA ACGGATTTTT 1620

GCTGGTCAAT TTAACAAAAC AATGGATACC GGTGACCCAT TAACATTCCA ATCTTTTAGT 1680

TACGCAACTA TTAATACAGC TTTTACATTC CCAATGAGCC AGAGTAGTTT CACAGTAGGT 1740

GCTGATACTT TTAGTTCAGG GAATGAAGTT TATATAGACA GATTTGAATT GATTCCAGTT 1800

ACTGCAACAT TTGAAGCAGA ATATGATTTA GAAAGAGCAC AAAAGGCGGT GAATGCGCTG 1860

TTTACTTCTA TAAACCAAAT AGGGATAAAA ACAGATGTGA CGGATTATCA TATTGATCAA 1920

GTATCCAATT TAGTGGATTG TTTATCAGAT GAATTTTGTC TGGATGAAAA GCGAGAATTG 1980

TCCGAGAAAG TCAAACATGC GAAGCGACTC AGTGATGAGC GGAATTTACT TCAAGATCCA 2040

AACTTCAAAG GCATCAATAG GCAACTAGAC CGTGGTTGGA GAGGAAGTAC GGATATTACC 2100

ATCCAAAGAG GAGATGACGT ATTCAAAGAA AATTATGTCA CACTACCAGG TACCTTTGAT 2160

GAGTGCTATC CAACGTATTT ATATCAAAAA ATAGATGAGT CGAAATTAAA ACCCTATACT 2220

CGTTATCAAT TAAGAGGGTA TATCGAGGAT AGTCAAGACT TAGAAATCTA TTTGATCCGC 2280

TATAATGCAA AACACGAAAC AGTAAATGTG CTAGGTACGG GTTCTTTATG GCCGCTTTCA 2340

GTCCAAAGTC CAATCAGAAA GTGTGGAGAA CCGAATCGAT GCGCGCCACA CCTTGAATGG 2400

AATCCTGATC TAGATTGTTC CTGCAGAGAC GGGGAAAAAT GTGCACATCA TTCGCATCAT 2460

TTCTCCTTGG ACATTGATGT TGGATGTACA GACTTAAATG AGGACTTAGA TGTATGGGTG 2520

ATATTCAAGA TTAAGACGCA AGATGGCCAT GCAAGACTAG GAAATCTAGA GTTTCTCGAA 2580

GAGAAACCAT TAGTCGGGGA AGCACTAGCT CGTGTGAAAA GAGCAGAGAA AAAATGGAGA 2640

GATAAACGTG AAAAATTGGA ATTGGAAACA AATATTGTTT ATAAAGAGGC AAAAGAATCT 2700

GTAGATGCTT TATTTGTAAA CTCTCAATAT GATCAATTAC AAGCGGATAC GAATATTGCC 2760

ATGATTCATG CGGCAGATAA ACGTGTTCAT AGAATTCGGG AAGCGTATCT TCCAGAGTTA 2820

TCTGTGATTC CGGGTGTAAA TGTAGACATT TTCGAAGAAT TAAAAGGGCG TATTTTCACT 2880

GCATTCTTCC TATATGATGC GAGAAATGTC ATTAAAAACG GTGATTTCAA TAATGGCTTA 2940

```
TCATGCTGGA ACGTGAAAGG GCATGTAGAT GTAGAAGAAC AAAACAACCA CCGTTCGGTC      3000

CTTGTTGTTC CGGAATGGGA AGCAGAAGTG TCACAAGAAG TTCGTGTCTG TCCGGGTCGT      3060

GGCTATATCC TTCGTGTCAC AGCGTACAAG GAGGGATATG GAGAAGGTTG CGTAACCATT      3120

CATGAGATCG AGAACAATAC AGACGAACTG AAGTTTAGCA ACTGCGTAGA AGAGGAAGTC      3180

TATCCAAACA ACACGGTAAC GTGTAATGAT TATACTGCAA ATCAAGAAGA ATACGGGGGT      3240

GCGTACACTT CCCGTAATCG TGGATATGAC GAAACTTATG GAAGCAATTC TTCTGTACCA      3300

GCTGATTATG CGTCAGTCTA TGAAGAAAAA TCGTATACAG ATGGACGAAG AGACAATCCT      3360

TGTGAATCTA ACAGAGGATA TGGGGATTAC ACACCACTAC CAGCTGGCTA TGTGACAAAA      3420

GAATTAGAGT ACTTCCCAGA AACCGATAAG GTATGGATTG AGATCGGAGA AACGGAAGGA      3480

ACATTCATCG TGGACAGCGT GGAATTACTC CTTATGGAGG AA                         3522
```

(2) 서열 번호 25에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 1174 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 25 :

```
Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
1               5                  10                  15

Asn Pro Glu Val Glu Ile Leu Asn Glu Glu Arg Ser Thr Gly Arg Leu
            20                  25                  30

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
        35                  40                  45

Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
    50                  55                  60

Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
65                  70                  75                  80

Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
                85                  90                  95

Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
            100                 105                 110

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
        115                 120                 125

Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
    130                 135                 140

Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
145                 150                 155                 160

Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe
                165                 170                 175

Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
            180                 185                 190
```

<174>

```
Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr
        195                 200                 205

Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
    210                 215                 220


Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
225                 230                 235                 240

Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln
                245                 250                 255

Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu
            260                 265                 270

Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu
        275                 280                 285

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
    290                 295                 300

Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu
305                 310                 315                 320

Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr
                325                 330                 335

Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
            340                 345                 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
        355                 360                 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
    370                 375                 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
385                 390                 395                 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
                405                 410                 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
            420                 425                 430

Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
        435                 440                 445

Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
    450                 455                 460

Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
465                 470                 475                 480

Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
                485                 490                 495

Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu
            500                 505                 510

Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu
        515                 520                 525

Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe
    530                 535                 540
```

```
Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser
545                 550                 555                 560

Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser
                565                 570                 575

Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile
            580                 585                 590

Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu Ala Glu Tyr
        595                 600                 605

Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Ile
    610                 615                 620

Asn Gln Ile Gly Ile Lys Thr Asp Val Thr Asp Tyr His Ile Asp Gln
625                 630                 635                 640

Val Ser Asn Leu Val Asp Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu
                645                 650                 655

Lys Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp
            660                 665                 670

Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe Lys Gly Ile Asn Arg Gln
        675                 680                 685

Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Arg Gly
    690                 695                 700

Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Pro Gly Thr Phe Asp
705                 710                 715                 720

Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu
                725                 730                 735

Lys Pro Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln
            740                 745                 750

Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val
        755                 760                 765

Asn Val Leu Gly Thr Gly Ser Leu Trp Pro Leu Ser Val Gln Ser Pro
    770                 775                 780

Ile Arg Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp
785                 790                 795                 800

Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His
                805                 810                 815

His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu
            820                 825                 830

Asn Glu Asp Leu Asp Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp
        835                 840                 845

Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu
    850                 855                 860

Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg
865                 870                 875                 880

Asp Lys Arg Glu Lys Leu Glu Leu Glu Thr Asn Ile Val Tyr Lys Glu
                885                 890                 895

Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln
            900                 905                 910
```

Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg
915 920 925

Val His Arg Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro
930 935 940

Gly Val Asn Val Asp Ile Phe Glu Glu Leu Lys Gly Arg Ile Phe Thr
945 950 955 960

Ala Phe Phe Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe
965 970 975

Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu
980 985 990

Glu Gln Asn Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala
995 1000 1005

Glu Val Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu
1010 1015 1020

Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile
1025 1030 1035 1040

His Glu Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val
1045 1050 1055

Glu Glu Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr
1060 1065 1070

Ala Asn Gln Glu Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly
1075 1080 1085

Tyr Asp Glu Thr Tyr Gly Ser Asn Ser Ser Val Pro Ala Asp Tyr Ala
1090 1095 1100

Ser Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Asp Asn Pro
1105 1110 1115 1120

Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly
1125 1130 1135

Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp
1140 1145 1150

Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu
1155 1160 1165

Leu Leu Leu Met Glu Glu
1170

(2) 서열 번호 26에 대한 정보 :

(i) 서열의 특징:
(A) 서열의 길이 : 3444 염기쌍
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : DNA (게놈)

(xi) 서열 : 서열 번호 26 :

ATGGAAAATA ATATTCAAAA TCAATGCGTA CCTTACAATT GTTTAAATAA TCCTGAAGTA 60

GAAATACTGA ACGAAGAACG CAGCACCGGC CGCCTGCCGC TGGACATCAG CCTGAGCCTT 120

ACACGTTTCC TTTTGAGTGA ATTTGTTCCA GGTGTGGGAG TTGCGTTTGG ATTATTTGAT 180

TTAATATGGG GTTTTATAAC TCCTTCTGAT TGGAGCTTAT TTCTTTTACA GATTGAACAA 240

TTGATTGAGC AAAGAATAGA AACATTGGAA AGGAACCGGG CAATTACTAC ATTACGAGGG 300

TTAGCAGATA GCTATGAAAT TTATATTGAA GCACTAAGAG AGTGGGAAGC AAATCCTAAT 360

AATGCACAAT TAAGGGAAGA TGTGCGTATT CGATTTGCTA ATACAGACGA CGCTTTAATA 420

ACAGCAATAA ATAATTTTAC ACTTACAAGT TTTGAAATCC CTCTTTTATC GGTCTATGTT 480

CAAGCGGCGA ATTTACATTT ATCACTATTA AGAGACGCTG TATCGTTTGG GCAGGGTTGG 540

GGACTGGATA TAGCTACTGT TAATAATCAT TATAATAGAT TAATAAATCT TATTCATAGA 600

TATACGAAAC ATTGTTTGGA CACATACAAT CAAGGATTAG AAAACTTAAG AGGTACTAAT 660

ACTCGACAAT GGGCAAGATT CAATCAGTTT AGGAGAGATT TAACACTTAC TGTATTAGAT 720

ATCGTTGCTC TTTTTCCGAA CTACGATGTT AGAACATATC CAATTCAAAC GTCATCCCAA 780

TTAACAAGGG AAATTTATAC AAGTTCAGTA ATTGAGGATT CTCCAGTTTC TGCTAATATA 840

CCTAATGGTT TTAATAGGGC GGAATTTGGA GTTAGACCGC CCCATCTTAT GGACTTTATG 900

AATTCTTTGT TTGTAACTGC AGAGACTGTT AGAAGTCAAA CTGTGTGGGG AGGACACTTA 960

GTTAGTTCAC GAAATACGGC TGGTAACCGT ATAAATTTCC CTAGTTACGG GGTCTTCAAT 1020

CCTGGTGGCG CCATTTGGAT TGCAGATGAG GATCCACGTC CTTTTTATCG GACATTATCA 1080

GATCCTGTTT TTGTCCGAGG AGGATTTGGG AATCCTCATT ATGTACTGGG GCTTAGGGGA 1140

GTAGCATTTC AACAAACTGG TACGAACCAC ACCCGAACAT TTAGAAATAG TGGGACCATA 1200

GATTCTCTAG ATGAAATCCC ACCTCAGGAT AATAGTGGGG CACCTTGGAA TGATTATAGT 1260

CATGTATTAA ATCATGTTAC ATTTGTACGA TGGCCAGGTG AGATTTCAGG AAGTGATTCA 1320

TGGAGAGCTC CAATGTTTTC TTGGACGCAC CGTAGTGCAA CCCCTACAAA TACAATTGAT 1380

CCGGAGAGGA TTACTCAAAT ACCATTGGTA AAAGCACATA CACTTCAGTC AGGTACTACT 1440

GTTGTAAGAG GGCCCGGGTT TACGGGAGGA GATATTCTTC GACGAACAAG TGGAGGACCA 1500

TTTGCTTATA CTATTGTTAA TATAAATGGG CAATTACCCC AAAGGTATCG TGCAAGAATA 1560

CGCTATGCCT CTACTACAAA TCTAAGAATT TACGTAACGG TTGCAGGTGA ACGGATTTTT 1620

GCTGGTCAAT TTAACAAAAC AATGGATACC GGTGACCCAT TAACATTCCA ATCTTTTAGT 1680

TACGCAACTA TTAATACAGC TTTTACATTC CCAATGAGCC AGAGTAGTTT CACAGTAGGT 1740

GCTGATACTT TTAGTTCAGG GAATGAAGTT TATATAGACA GATTTGAATT GATTCCAGTT 1800

ACTGCAACAT TTGAAGCAGA ATATGATTTA GAAAGAGCAC AAAAGGCGGT GAATGCGCTG 1860

TTTACTTCTA TAAACCAAAT AGGGATAAAA ACAGATGTGA CGGATTATCA TATCGATCGA 1920

GTATCCAATT TAGTTGAGTG TTTATCTGAT GAATTTTGTC TGGATGAAAA AAAAGAATTG 1980

TCCGAGAAAG TCAAACATGC GAAGCGACTT AGTGATGAGC GGAATTTACT TCAAGATCCA 2040

AACTTTAGAG GGATCAATAG ACAACTAGAC CGTGGCTGGA GAGGAAGTAC GGATATTACC 2100

ATCCAAGGAG GCGATGACGT ATTCAAAGAG AATTACGTTA CGCTATTGGG TACCTTTGAT 2160

GAGTGCTATC CAACGTATTT ATATCAAAAA ATAGATGAGT CGAAATTAAA AGCCTATACC 2220

CGTTACCAAT TAAGAGGGTA TATCGAAGAT AGTCAAGACT TAGAAATCTA TTTAATTCGC 2280

TACAATGCCA AACACGAAAC AGTAAATGTG CCAGGTACGG GTTCCTTATG GCCGCTTTCA 2340

GCCCCAAGTC CAATCGGAAA ATGTGCCCAT CATTCCCATC ATTTCTCCTT GGACATTGAT 2400

GTTGGATGTA CAGACTTAAA TGAGGACTTA GGTGTATGGG TGATATTCAA GATTAAGACG 2460

CAAGATGGCC ATGCAAGACT AGGAAATCTA GAATTTCTCG AAGAGAAACC ATTAGTAGGA 2520

GAAGCACTAG CTCGTGTGAA AAGAGCGGAG AAAAAATGGA GAGACAAACG TGAAAAATTG 2580

GAATGGGAAA CAAATATTGT TTATAAAGAG GCAAAAGAAT CTGTAGATGC TTTATTTGTA 2640

AACTCTCAAT ATGATAGATT ACAAGCGGAT ACCAACATCG CGATGATTCA TGCGGCAGAT 2700

AAACGCGTTC ATAGCATTCG AGAAGCTTAT CTGCCTGAGC TGTCTGTGAT TCCGGGTGTC 2760

AATGCGGCTA TTTTTGAAGA ATTAGAAGGG CGTATTTTCA CTGCATTCTC CCTATATGAT 2820

GCGAGAAATG TCATTAAAAA TGGTGATTTT AATAATGGCT TATCCTGCTG GAACGTGAAA 2880

GGGCATGTAG ATGTAGAAGA ACAAAACAAC CACCGTTCGG TCCTTGTTGT TCCGGAATGG 2940

GAAGCAGAAG TGTCACAAGA AGTTCGTGTC TGTCCGGGTC GTGGCTATAT CCTTCGTGTC 3000

ACAGCGTACA AGGAGGGATA TGGAGAAGGT TGCGTAACCA TTCATGAGAT CGAGAACAAT 3060

ACAGACGAAC TGAAGTTTAG CAACTGTGTA GAAGAGGAAG TATATCCAAA CAACACGGTA 3120

ACGTGTAATG ATTATACTGC GACTCAAGAA GAATATGAGG GTACGTACAC TTCTCGTAAT 3180

CGAGGATATG ACGGAGCCTA TGAAAGCAAT TCTTCTGTAC CAGCTGATTA TGCATCAGCC 3240

TATGAAGAAA AAGCATATAC AGATGGACGA AGAGACAATC CTTGTGAATC TAACAGAGGA 3300

TATGGGGATT ACACACCACT ACCAGCTGGC TATGTGACAA AAGAATTAGA GTACTTCCCA 3360

GAAACCGATA AGGTATGGAT TGAGATCGGA GAAACGGAAG GAACATTCAT CGTGGACAGC 3420

GTGGAATTAC TTCTTATGGA GGAA 3444

(2) 서열 번호 27에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 1148 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 27 :

Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
1 5 10 15

Asn Pro Glu Val Glu Ile Leu Asn Glu Glu Arg Ser Thr Gly Arg Leu
20 25 30

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
35 40 45

Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
50 55 60

```
Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
65                  70                  75                  80

Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
                85                  90                  95

Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
            100                 105                 110

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
        115                 120                 125

Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
    130                 135                 140

Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
145                 150                 155                 160

Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe
                165                 170                 175

Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
            180                 185                 190

Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr
        195                 200                 205

Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
    210                 215                 220

Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
225                 230                 235                 240

Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln
                245                 250                 255

Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu
            260                 265                 270

Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu
        275                 280                 285

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
    290                 295                 300

Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu
305                 310                 315                 320

Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr
                325                 330                 335

Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
            340                 345                 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
        355                 360                 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
    370                 375                 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
385                 390                 395                 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
                405                 410                 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
            420                 425                 430
```

```
Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
        435                 440                 445

Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
    450                 455                 460

Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
465                 470                 475                 480

Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
                485                 490                 495

Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu
            500                 505                 510

Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu
        515                 520                 525

Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe
    530                 535                 540

Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser
545                 550                 555                 560

Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser
                565                 570                 575

Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile
            580                 585                 590

Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu Ala Glu Tyr
        595                 600                 605

Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Ile
    610                 615                 620

Asn Gln Ile Gly Ile Lys Thr Asp Val Thr Asp Tyr His Ile Asp Arg
625                 630                 635                 640

Val Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu
                645                 650                 655

Lys Lys Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp
            660                 665                 670

Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln
        675                 680                 685

Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly
    690                 695                 700

Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Leu Gly Thr Phe Asp
705                 710                 715                 720

Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu
                725                 730                 735

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

Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val
        755                 760                 765

Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Pro Ser Pro
    770                 775                 780

Ile Gly Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp
785                 790                 795                 800
```

```
Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
                805                 810                 815

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
            820                 825                 830

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg
        835                 840                 845

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr
    850                 855                 860

Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val
865                 870                 875                 880

Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile
                885                 890                 895

His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro
            900                 905                 910

Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu
        915                 920                 925

Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val
    930                 935                 940

Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys
945                 950                 955                 960

Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val
                965                 970                 975

Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys Pro
            980                 985                 990

Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly
        995                 1000                1005

Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr Asp Glu Leu
    1010                1015                1020

Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro Asn Asn Thr Val
1025                1030                1035                1040

Thr Cys Asn Asp Tyr Thr Ala Thr Gln Glu Glu Tyr Glu Gly Thr Tyr
                1045                1050                1055

Thr Ser Arg Asn Arg Gly Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser
            1060                1065                1070

Val Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp
        1075                1080                1085

Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr
    1090                1095                1100

Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro
1105                1110                1115                1120

Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe
                1125                1130                1135

Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
            1140                1145
```

(2) 서열 번호 28에 대한 정보 :

(ⅰ) 서열의 특징:
(A) 서열의 길이 : 3522 염기쌍
(B) 서열의 형 : 핵산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ⅱ) 분자 유형 : DNA (게놈)

(xi) 서열 : 서열 번호 28 :

ATGGAAAATA ATATTCAAAA TCAATGCGTA CCTTACAATT GTTTAAATAA TCCTGAAGTA 60
GAAATACTGA ACGAAGAACG CAGCACCGGC CGCCTGCCGC TGGACATCAG CCTGAGCCTT 120
ACACGTTTCC TTTTGAGTGA ATTTGTTCGA GGTGTGGGAG TTGCGTTTGG ATTATTTGAT 180
TTAATATGGG GTTTTATAAC TCCTTCTGAT TGGAGCTTAT TTCTTTTACA GATTGAACAA 240
TTGATTGAGC AAAGAATAGA AACATTGGAA AGGAACCGGG CAATTACTAC ATTACGAGGG 300
TTAGCAGATA GCTATGAAAT TTATATTGAA GCACTAAGAG AGTGGGAAGC AAATCCTAAT 360
AATGCACAAT TAAGGGAAGA TGTGCGTATT CGATTTGCTA ATACAGACGA CGCTTTAATA 420
ACAGCAATAA ATAATTTTAC ACTTACAAGT TTTGAAATCC CTCTTTTATC GGTCTATGTT 480
CAAGCGGCGA ATTTACATTT ATCACTATTA AGAGACGCTG TATCGTTTGG GCAGGGTTGG 540
GGACTGGATA TAGCTACTGT TAATAATCAT TATAATAGAT TAATAAATCT TATTCATAGA 600
TATACGAAAC ATTGTTTGGA CACATACAAT CAAGGATTAG AAAACTTAAG AGGTACTAAT 660
ACTCGACAAT GGGCAAGATT CAATCAGTTT AGGAGAGATT TAACACTTAC TGTATTAGAT 720
ATCGTTGCTC TTTTTCCGAA CTACGATGTT AGAACATATC CAATTCAAAC GTCATCCCAA 780
TTAACAAGGG AAATTTATAC AAGTTCAGTA ATTGAGGATT CTCCAGTTTC TGCTAATATA 840
CCTAATGGTT TTAATAGGGC GGAATTTGGA GTTAGACCGC CCCATCTTAT GGACTTTATG 900
AATTCTTTGT TTGTAACTGC AGAGACTGTT AGAAGTCAAA CTGTGTGGGG AGGACACTTA 960
GTTAGTTCAC GAAATACGGC TGGTAACCGT ATAAATTTCC CTAGTTACGG GGTCTTCAAT 1020
CCTGGTGGCG CCATTTGGAT TGCAGATGAG GATCCACGTC CTTTTTATCG GACATTATCA 1080
GATCCTGTTT TTGTCCGAGG AGGATTTGGG AATCCTCATT ATGTACTGGG GCTTAGGGGA 1140
GTAGCATTTC AACAAACTGG TACGAACCAC ACCCGAACAT TTAGAAATAG TGGGACCATA 1200
GATTCTCTAG ATGAAATCCC ACCTCAGGAT AATAGTGGGG CACCTTGGAA TGATTATAGT 1260
CATGTATTAA ATCATGTTAC ATTTGTACGA TGGCCAGGTG AGATTTCAGG AAGTGATTCA 1320
TGGAGAGCTC CAATGTTTTC TTGGACGCAC CGTAGTGCAA CCCCTACAAA TACAATTGAT 1380
CCGGAGAGGA TTACTCAAAT ACCATTGGTA AAAGCACATA CACTTCAGTC AGGTACTACT 1440
GTTGTAAGAG GGCCCGGGTT TACGGGAGGA GATATTCTTC GACGAACAAG TGGAGGACCA 1500
TTTGCTTATA CTATTGTTAA TATAAATGGG CAATTACCCC AAAGGTATCG TGCAAGAATA 1560
CGCTATGCCT CTACTACAAA TCTAAGAATT TACGTAACGG TTGCAGGTGA ACGGATTTTT 1620
GCTGGTCAAT TTAACAAAAC AATGGATACC GGTGACCCAT TAACATTCCA ATCTTTTAGT 1680
TACGCAACTA TTAATACAGC TTTTACATTC CCAATGAGCC AGAGTAGTTT CACAGTAGGT 1740

GCTGATACTT TTAGTTCAGG GAATGAAGTT TATATAGACA GATTTGAATT GATTCCAGTT 1800

ACTGCAACAT TTGAAGCAGA ATATGATTTA GAAAGAGCAC AAAAGGCGGT GAATGCGCTG 1860

TTTACTTCTA TAAACCAAAT AGGGATAAAA ACAGATGTGA CGGATTATCA TATCGATCGA 1920

GTGTCCAATT TAGTTACGTA TTTATCGGAT GAATTTTGTC TGGATGAAAA GCGAGAATTG 1980

TCCGAGAAAG TCAAACATGC GAAGCGACTC AGTGATGAAC GCAATTTACT CCAAGATTCA 2040

AATTTCAAAG ACATTAATAG GCAACCAGAA CGTGGGTGGG GCGGAAGTAC AGGGATTACC 2100

ATCCAAGGAG GGGATGACGT ATTTAAAGAA AATTACGTCA CACTATCAGG TACCTTTGAT 2160

GAGTGCTATC CAACATATTT GTATCAAAAA ATCGATGAAT CAAAATTAAA AGCCTTTACC 2220

CGTTATCAAT TAAGAGGGTA TATCGAAGAT AGTCAAGACT TAGAAATCTA TTTAATTCGC 2280

TACAATGCAA AACATGAAAC AGTAAATGTG CCAGGTACGG GTTCCTTATG GCCGCTTTCA 2340

GCCCAAAGTC CAATCGGAAA GTGTGGAGAG CCGAATCGAT GCGCGCCACA CCTTGAATGG 2400

AATCCTGACT TAGATTGTTC GTGTAGGGAT GGAGAAAAGT GTGCCCATCA TTCGCATCAT 2460

TTCTCCTTAG ACATTGATGT AGGATGTACA GACTTAAATG AGGACCTAGG TGTATGGGTG 2520

ATCTTTAAGA TTAAGACGCA AGATGGGCAC GCAAGACTAG GGAATCTAGA GTTTCTCGAA 2580

GAGAAACCAT TAGTAGGAGA AGCGCTAGCT CGTGTGAAAA GAGCGGAGAA AAAATGGAGA 2640

GACAAACGTG AAAAATTGGA ATGGGAAACA AATATCGTTT ATAAAGAGGC AAAAGAATCT 2700

GTAGATGCTT TATTTGTAAA CTCTCAATAT GATCAATTAC AAGCGGATAC GAATATTGCC 2760

ATGATTCATG CGGCAGATAA ACGTGTTCAT AGCATTCGAG AAGCTTATCT GCCTGAGCTG 2820

TCTGTGATTC CGGGTGTCAA TGCGGCTATT TTTGAAGAAT TAGAAGGGCG TATTTTCACT 2880

GCATTCTCCC TATATGATGC GAGAAATGTC ATTAAAAATG GTGATTTTAA TAATGGCTTA 2940

TCCTGCTGGA ACGTGAAAGG GCATGTAGAT GTAGAAGAAC AAAACAACCA CCGTTCGGTC 3000

CTTGTTGTTC CGGAATGGGA AGCAGAAGTG TCACAAGAAG TTCGTGTCTG TCCGGGTCGT 3060

GGCTATATCC TTCGTGTCAC AGCGTACAAG GAGGGATATG GAGAAGGTTG CGTAACCATT 3120

CATGAGATCG AGAACAATAC AGACGAACTG AAGTTTAGCA ACTGTGTAGA AGAGGAAGTA 3180

TATCCAAACA ACACGGTAAC GTGTAATGAT TATACTGCGA CTCAAGAAGA ATATGAGGGT 3240

ACGTACACTT CTCGTAATCG AGGATATGAC GGAGCCTATG AAAGCAATTC TTCTGTACCA 3300

GCTGATTATG CATCAGCCTA TGAAGAAAAA GCATATACAG ATGGACGAAG AGACAATCCT 3360

TGTGAATCTA ACAGAGGATA TGGGGATTAC ACACCACTAC CAGCTGGCTA TGTGACAAAA 3420

GAATTAGAGT ACTTCCCAGA AACCGATAAG GTATGGATTG AGATCGGAGA AACGGAAGGA 3480

ACATTCATCG TGGACAGCGT GGAATTACTT CTTATGGAGG AA 3522

(2) 서열 번호 29에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 1174 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형

(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 29 :

```
Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
1               5                  10                  15

Asn Pro Glu Val Glu Ile Leu Asn Glu Glu Arg Ser Thr Gly Arg Leu
            20                  25                  30

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
        35                  40                  45

Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
    50                  55                  60

Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
65                  70                  75                  80

Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
                85                  90                  95

Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
            100                 105                 110

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
        115                 120                 125

Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
    130                 135                 140

Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
145                 150                 155                 160

Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe
                165                 170                 175

Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
            180                 185                 190

Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr
        195                 200                 205

Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
    210                 215                 220

Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
225                 230                 235                 240

Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln
                245                 250                 255

Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu
            260                 265                 270

Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu
        275                 280                 285

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
    290                 295                 300

Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu
305                 310                 315                 320

Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr
                325                 330                 335
```

```
Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
            340                 345                 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
        355                 360                 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
    370                 375                 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
385                 390                 395                 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
                405                 410                 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
            420                 425                 430

Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
        435                 440                 445

Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
    450                 455                 460

Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
465                 470                 475                 480

Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
                485                 490                 495

Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu
            500                 505                 510

Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu
        515                 520                 525

Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe
    530                 535                 540

Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser
545                 550                 555                 560

Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser
                565                 570                 575

Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile
            580                 585                 590

Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu Ala Glu Tyr
        595                 600                 605

Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Ile
    610                 615                 620

Asn Gln Ile Gly Ile Lys Thr Asp Val Thr Asp Tyr His Ile Asp Arg
625                 630                 635                 640

Val Ser Asn Leu Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu
                645                 650                 655

Lys Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp
            660                 665                 670

Glu Arg Asn Leu Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln
        675                 680                 685

Pro Glu Arg Gly Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly
    690                 695                 700
```

```
Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp
705                 710                 715                 720

Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu
                725                 730                 735

Lys Ala Phe Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln
            740                 745                 750

Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val
        755                 760                 765

Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro
    770                 775                 780

Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp
785                 790                 795                 800

Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His
                805                 810                 815

His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu
            820                 825                 830

Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp
        835                 840                 845

Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu
    850                 855                 860

Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg
865                 870                 875                 880

Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu
                885                 890                 895

Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln
            900                 905                 910

Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg
        915                 920                 925

Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro
    930                 935                 940

Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr
945                 950                 955                 960

Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe
                965                 970                 975

Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu
            980                 985                 990

Glu Gln Asn Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala
        995                 1000                1005

Glu Val Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu
    1010                1015                1020

Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile
1025                1030                1035                1040

His Glu Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val
                1045                1050                1055

Glu Glu Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr
            1060                1065                1070
```

```
Ala Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly
        1075                1080                1085

Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr Ala
    1090                1095                1100

Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp Asn Pro
1105                1110                1115                1120

Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly
                1125                1130                1135

Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp
            1140                1145                1150

Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu
        1155                1160                1165

Leu Leu Leu Met Glu Glu
    1170
```

(2) 서열 번호 30에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 12 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 30 :

```
Xaa Xaa Ile Asp Xaa Xaa Glu Xaa Xaa Xaa Xaa Xaa
                5                   10
```

(2) 서열 번호 31에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 8 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 31 :

```
Tyr Pro Asn Asn Thr Val Thr Cys
                5
```

(2) 서열 번호 32에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 1184 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 32 :

```
Cys Arg Tyr Ile Phe Ala Met Pro Glu Pro Met Glu Asn Asn Ile Gln
1               5                   10                  15
```

```
Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn Asn Pro Glu Val Glu Ile
            20                  25                  30

Leu Asn Glu Glu Arg Ser Thr Gly Arg Leu Pro Leu Asp Ile Ser Leu
        35                  40                  45

Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe Val Pro Gly Val Gly Val
    50                  55                  60

Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly Phe Ile Thr Pro Ser Asp
65                  70                  75                  80

Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln Leu Ile Glu Gln Arg Ile
                85                  90                  95

Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr Thr Leu Arg Gly Leu Ala
            100                 105                 110

Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu Arg Glu Trp Glu Ala Asn
        115                 120                 125

Pro Asn Asn Ala Gln Leu Arg Glu Asp Val Arg Ile Arg Phe Ala Asn
    130                 135                 140

Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn Asn Phe Thr Leu Thr Ser
145                 150                 155                 160

Phe Glu Ile Pro Leu Leu Ser Val Tyr Val Gln Ala Ala Asn Leu His
                165                 170                 175

Leu Ser Leu Leu Arg Asp Ala Val Ser Phe Gly Gln Gly Trp Gly Leu
            180                 185                 190

Asp Ile Ala Thr Val Asn Asn His Tyr Asn Arg Leu Ile Asn Leu Ile
        195                 200                 205

His Arg Tyr Thr Lys His Cys Leu Asp Thr Tyr Asn Gln Gly Leu Glu
    210                 215                 220

Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp Ala Arg Phe Asn Gln Phe
225                 230                 235                 240

Arg Arg Asp Leu Thr Leu Thr Val Leu Asp Ile Val Ala Leu Phe Pro
                245                 250                 255

Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln Thr Ser Ser Gln Leu Thr
            260                 265                 270

Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu Asp Ser Pro Val Ser Ala
        275                 280                 285

Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu Phe Gly Val Arg Pro Pro
    290                 295                 300

His Leu Met Asp Phe Met Asn Ser Leu Phe Val Thr Ala Glu Thr Val
305                 310                 315                 320

Arg Ser Gln Thr Val Trp Gly Gly His Leu Val Ser Ser Arg Asn Thr
                325                 330                 335

Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr Gly Val Phe Asn Pro Gly
            340                 345                 350

Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro Arg Pro Phe Tyr Arg Thr
        355                 360                 365

Leu Ser Asp Pro Val Phe Val Arg Gly Gly Phe Gly Asn Pro His Tyr
    370                 375                 380
```

```
Val Leu Gly Leu Arg Gly Val Ala Phe Gln Gln Thr Gly Thr Asn His
385                 390                 395                 400

Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile Asp Ser Leu Asp Glu Ile
                405                 410                 415

Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp Asn Asp Tyr Ser His Val
            420                 425                 430

Leu Asn His Val Thr Phe Val Arg Trp Pro Gly Glu Ile Ser Gly Ser
        435                 440                 445

Asp Ser Trp Arg Ala Pro Met Phe Ser Trp Thr His Arg Ser Ala Thr
    450                 455                 460

Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile Thr Gln Ile Pro Leu Val
465                 470                 475                 480

Lys Ala His Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly
                485                 490                 495

Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala
            500                 505                 510

Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala
        515                 520                 525

Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val
    530                 535                 540

Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr
545                 550                 555                 560

Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr
                565                 570                 575

Ala Phe Thr Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp
            580                 585                 590

Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile
        595                 600                 605

Pro Val Thr Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln
    610                 615                 620

Lys Ala Val Asn Ala Leu Phe Thr Ser Ile Asn Gln Ile Gly Ile Lys
625                 630                 635                 640

Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Asp
                645                 650                 655

Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu
            660                 665                 670

Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln
        675                 680                 685

Asp Pro Asn Phe Lys Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg
    690                 695                 700

Gly Ser Thr Asp Ile Thr Ile Gln Arg Gly Asp Asp Val Phe Lys Glu
705                 710                 715                 720

Asn Tyr Val Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr
                725                 730                 735

Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Pro Tyr Thr Arg Tyr
            740                 745                 750
```

```
Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu
        755                 760                 765

Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Leu Gly Thr Gly
    770                 775                 780

Ser Leu Trp Pro Leu Ser Val Gln Ser Pro Ile Arg Lys Cys Gly Glu
785                 790                 795                 800

Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys
                805                 810                 815

Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser
            820                 825                 830

Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Asp Val
        835                 840                 845

Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly
    850                 855                 860

Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala
865                 870                 875                 880

Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu
                885                 890                 895

Glu Leu Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp
            900                 905                 910

Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn
        915                 920                 925

Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu
    930                 935                 940

Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Val Asp Ile
945                 950                 955                 960

Phe Glu Glu Leu Lys Gly Arg Ile Phe Thr Ala Phe Phe Leu Tyr Asp
                965                 970                 975

Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys
            980                 985                 990

Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg
        995                 1000                1005

Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val
    1010                1015                1020

Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys
1025                1030                1035                1040

Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn
                1045                1050                1055

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro
            1060                1065                1070

Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Asn Gln Glu Glu Tyr
        1075                1080                1085

Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Thr Tyr Gly
    1090                1095                1100

Ser Asn Ser Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys
1105                1110                1115                1120
```

```
Ser Tyr Thr Asp Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly
                1125                1130                1135

Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu
            1140                1145                1150

Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr
        1155                1160                1165

Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
    1170                1175                1180
```

(2) 서열 번호 33에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 1165 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 33 :

```
Cys Arg Tyr Ile Ala Asx Met Pro Glu Pro Met Asp Asn Asn Pro Asn
1               5                   10                  15

Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu Ser Asn Pro Glu Val Glu
            20                  25                  30

Val Leu Gly Gly Glu Arg Ile Glu Thr Gly Tyr Thr Pro Ile Asp Ile
        35                  40                  45

Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser Glu Phe Val Pro Gly Ala
    50                  55                  60

Gly Phe Val Leu Gly Leu Val Asp Ile Ile Trp Gly Ile Phe Gly Pro
65                  70                  75                  80

Ser Gln Trp Asp Ala Phe Leu Val Gln Ile Glu Gln Leu Ile Asn Gln
                85                  90                  95

Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile Ser Arg Leu Glu Gly
            100                 105                 110

Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu Ser Phe Arg Glu Trp Glu
        115                 120                 125

Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu Met Arg Ile Gln Phe
    130                 135                 140

Asn Asp Met Asn Ser Ala Leu Thr Thr Ala Ile Pro Leu Phe Ala Val
145                 150                 155                 160

Gln Asn Tyr Gln Val Pro Leu Leu Ser Val Tyr Val Gln Ala Ala Asn
                165                 170                 175

Leu His Leu Ser Val Leu Arg Asp Val Ser Val Phe Gly Gln Arg Trp
            180                 185                 190

Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg Tyr Asn Asp Leu Thr Arg
        195                 200                 205

Leu Ile Gly Asn Tyr Thr Asp His Ala Val Arg Trp Tyr Asn Thr Gly
    210                 215                 220
```

```
Leu Glu Arg Val Trp Gly Pro Asp Ser Arg Asp Trp Ile Arg Tyr Asn
225                 230                 235                 240

Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu Asp Ile Val Ser Leu
                245                 250                 255

Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro Ile Arg Thr Val Ser Gln
            260                 265                 270

Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val Leu Glu Asn Phe Asp Gly
        275                 280                 285

Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu Gly Ser Ile Arg Ser Pro
    290                 295                 300

His Leu Met Asp Ile Leu Asn Ser Ile Thr Ile Tyr Thr Asp Ala His
305                 310                 315                 320

Arg Gly Glu Tyr Tyr Trp Ser Gly His Gln Ile Met Ala Ser Pro Val
                325                 330                 335

Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro Leu Tyr Gly Thr Met Gly
            340                 345                 350

Asn Ala Ala Pro Gln Gln Arg Ile Val Ala Gln Leu Gly Gln Gly Val
        355                 360                 365

Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg Arg Pro Phe Asn Ile Gly
    370                 375                 380

Ile Asn Asn Gln Gln Leu Ser Val Leu Asp Gly Thr Glu Phe Ala Tyr
385                 390                 395                 400

Gly Thr Ser Ser Asn Leu Pro Ser Ala Val Tyr Arg Lys Ser Gly Thr
                405                 410                 415

Val Asp Ser Leu Asp Glu Ile Pro Pro Gln Asn Asn Asn Val Pro Pro
            420                 425                 430

Arg Gln Gly Phe Ser His Arg Leu Ser His Val Ser Met Phe Arg Ser
        435                 440                 445

Gly Phe Ser Asn Ser Ser Val Ser Ile Ile Arg Ala Pro Met Phe Ser
    450                 455                 460

Trp Ile His Arg Ser Ala Glu Phe Asn Asn Ile Ile Pro Ser Ser Gln
465                 470                 475                 480

Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr Asn Leu Gly Ser Gly Thr
                485                 490                 495

Ser Val Val Lys Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg
            500                 505                 510

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

Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg Tyr Ala Ser Thr Thr Asn
    530                 535                 540

Leu Gln Phe His Thr Ser Ile Asp Gly Arg Pro Ile Asn Gln Gly Asn
545                 550                 555                 560

Phe Ser Ala Thr Met Ser Ser Gly Ser Asn Leu Gln Ser Gly Ser Phe
                565                 570                 575

Arg Thr Val Gly Phe Thr Thr Pro Phe Asn Phe Ser Asn Gly Ser Ser
            580                 585                 590
```

```
Val Phe Thr Leu Ser Ala His Val Phe Asn Ser Gly Asn Glu Val Tyr
        595                 600                 605

Ile Asp Arg Ile Glu Phe Val Pro Ala Glu Val Thr Phe Glu Ala Glu
    610                 615                 620

Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Glu Leu Phe Thr Ser
625                 630                 635                 640

Ser Asn Gln Ile Gly Leu Lys Thr Asp Val Thr Asp Tyr His Ile Asp
                645                 650                 655

Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys Leu Asp
            660                 665                 670

Glu Lys Lys Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser
        675                 680                 685

Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile Asn Arg
    690                 695                 700

Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly
705                 710                 715                 720

Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Leu Gly Thr Phe
                725                 730                 735

Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys
            740                 745                 750

Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser
        755                 760                 765

Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr
    770                 775                 780

Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Pro Ser
785                 790                 795                 800

Pro Ile Gly Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
                805                 810                 815

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
            820                 825                 830

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
        835                 840                 845

Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys
    850                 855                 860

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu
865                 870                 875                 880

Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
                885                 890                 895

Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met
            900                 905                 910

Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
        915                 920                 925

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
    930                 935                 940

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
945                 950                 955                 960
```

```
Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val
                965                 970                 975

Lys Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu
            980                 985                 990

Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys
        995                 1000                1005

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

Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr Asp Glu
1025                1030                1035                1040

Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro Asn Asn Thr
                1045                1050                1055

Val Thr Cys Asn Asp Tyr Thr Ala Thr Gln Glu Glu Tyr Glu Gly Thr
            1060                1065                1070

Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Gly Ala Tyr Glu Ser Asn Ser
        1075                1080                1085

Ser Val Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr
    1090                1095                1100

Asp Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp
1105                1110                1115                1120

Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe
                1125                1130                1135

Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr
            1140                1145                1150

Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
        1155                1160                1165
```

(2) 서열 번호 34에 대한 정보 :

(i) 서열의 특징 :
(A) 서열의 길이 : 1188 아미노산
(B) 서열의 형 : 아미노산
(C) 쇄의 수 : 1 본쇄
(D) 형태 : 선형
(ii) 분자 유형 : 단백질
(xi) 서열 : 서열 번호 34 :

```
Ala Asx Cys Pro Glu Pro Met Asp Asn Asn Pro Asn Ile Asn Glu Cys
1               5                  10                  15

Ile Pro Tyr Asn Cys Leu Ser Asn Pro Glu Val Glu Val Leu Gly Gly
            20                  25                  30

Glu Arg Ile Glu Thr Gly Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu
        35                  40                  45

Thr Gln Phe Leu Leu Ser Glu Phe Val Pro Gly Ala Gly Phe Val Leu
    50                  55                  60

Gly Leu Val Asp Ile Ile Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp
65                  70                  75                  80
```

Ala Phe Leu Val Gln Ile Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu
85 90 95

Phe Ala Arg Asn Gln Ala Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu
100 105 110

Tyr Gln Ile Tyr Ala Glu Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr
115 120 125

Asn Pro Ala Leu Arg Glu Glu Met Arg Ile Gln Phe Asn Asp Met Asn
130 135 140

Ser Ala Leu Thr Thr Ala Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln
145 150 155 160

Val Pro Leu Leu Ser Val Tyr Val Gln Ala Ala Asn Leu His Leu Ser
165 170 175

Val Leu Arg Asp Val Ser Val Phe Gly Gln Arg Trp Gly Phe Asp Ala
180 185 190

Ala Thr Ile Asn Ser Arg Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn
195 200 205

Tyr Thr Asp Tyr Ala Val Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val
210 215 220

Trp Gly Pro Asp Ser Arg Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg
225 230 235 240

Glu Leu Thr Leu Thr Val Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr
245 250 255

Asp Ser Arg Arg Tyr Pro Ile Arg Thr Val Ser Gln Leu Thr Arg Glu
260 265 270

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

Ser Ala Gln Gly Ile Glu Arg Ser Ile Arg Ser Pro His Leu Met Asp
290 295 300

Ile Leu Asn Ser Ile Thr Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr
305 310 315 320

Tyr Trp Ser Gly His Gln Ile Met Ala Ser Pro Val Gly Phe Ser Gly
325 330 335

Pro Glu Phe Thr Phe Pro Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro
340 345 350

Gln Gln Arg Ile Val Ala Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu
355 360 365

Ser Ser Thr Leu Tyr Arg Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln
370 375 380

Gln Leu Ser Val Leu Asp Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser
385 390 395 400

Asn Leu Pro Ser Ala Val Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu
405 410 415

Asp Glu Ile Pro Pro Gln Asn Asn Asn Val Pro Pro Arg Gln Gly Phe
420 425 430

Ser His Arg Leu Ser His Val Ser Met Phe Arg Ser Gly Phe Ser
435 440

```
Ser Ser Val Ser Ile Ile Arg Ala Pro Met Phe Ser Trp Ile His Arg
    450                 455                 460

Ser Ala Glu Phe Asn Asn Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile
465                 470                 475                 480

Pro Ala Val Lys Gly Asn Phe Leu Phe Asn Gly Ser Val Ile Ser Gly
                485                 490                 495

Pro Gly Phe Thr Gly Gly Asp Leu Val Arg Leu Asn Ser Ser Gly Asn
            500                 505                 510

Asn Ile Gln Asn Arg Gly Tyr Ile Glu Val Pro Ile His Phe Pro Ser
        515                 520                 525

Thr Ser Thr Arg Tyr Arg Val Arg Val Arg Tyr Ala Ser Val Thr Pro
    530                 535                 540

Ile His Leu Asn Val Asn Trp Gly Asn Ser Ser Ile Phe Ser Asn Thr
545                 550                 555                 560

Val Pro Ala Thr Ala Thr Ser Leu Asp Asn Leu Gln Ser Ser Asp Phe
                565                 570                 575

Gly Tyr Phe Glu Ser Ala Asn Ala Phe Thr Ser Ser Leu Gly Asn Ile
            580                 585                 590

Val Gly Val Arg Asn Phe Ser Gly Thr Ala Gly Val Ile Ile Asp Arg
        595                 600                 605

Phe Glu Phe Ile Pro Val Thr Ala Thr Leu Glu Ala Glu Tyr Asn Leu
    610                 615                 620

Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Thr Asn Gln
625                 630                 635                 640

Leu Gly Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser
                645                 650                 655

Asn Leu Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg
            660                 665                 670

Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg
        675                 680                 685

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

Arg Gly Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp
705                 710                 715                 720

Val Phe Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys
                725                 730                 735

Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala
            740                 745                 750

Phe Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu
        755                 760                 765

Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val
    770                 775                 780

Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly
785                 790                 795                 800

Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro
                805                 810                 815
```

Asp Leu Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser
820 825 830

His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu
835 840 845

Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His
850 855 860

Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly
865 870 875 880

Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys
885 890 895

Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys
900 905 910

Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln
915 920 925

Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His
930 935 940

Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val
945 950 955 960

Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe
965 970 975

Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn
980 985 990

Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln
995 1000 1005

Asn Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val
1010 1015 1020

Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val
1025 1030 1035 1040

Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu
1045 1050 1055

Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu
1060 1065 1070

Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Thr
1075 1080 1085

Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp
1090 1095 1100

Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr Ala Ser Ala
1105 1110 1115 1120

Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp Asn Pro Cys Glu
1125 1130 1135

Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr Val
1140 1145 1150

Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu
1155 1160 1165

Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
1170 1175

Leu Met Glu Glu
1185

(57) 청구의 범위

청구항 1

바실루스 츄린겐시스 독소를 코드화하는 뉴클레오티드 서열을 포함하는 분리된 폴리뉴클레오티드 분자로서, 여기서 상기 바실루스 츄린겐시스 독소가 cryIF 코어 N-말단 독소 부분 및 cryIA(b)독소 또는 cryIA(c)/cryIA(b) 키메라 독소로 부터의 이종 C-말단 프로톡신 부분을 포함하는 키메라 독소인 분리된 폴리뉴클레오티드 분자.

청구항 2

cryIF 코어 N-말단 독소 부분 및 cryIA(b) 독소 또는 cryIA(b)/cryIA(c) 키메라 독소로 부터의 C-말단 프로톡신 부분을 포함하는 실질상 순수한 키메라 바실루스 츄린겐시스 독소.

청구항 3

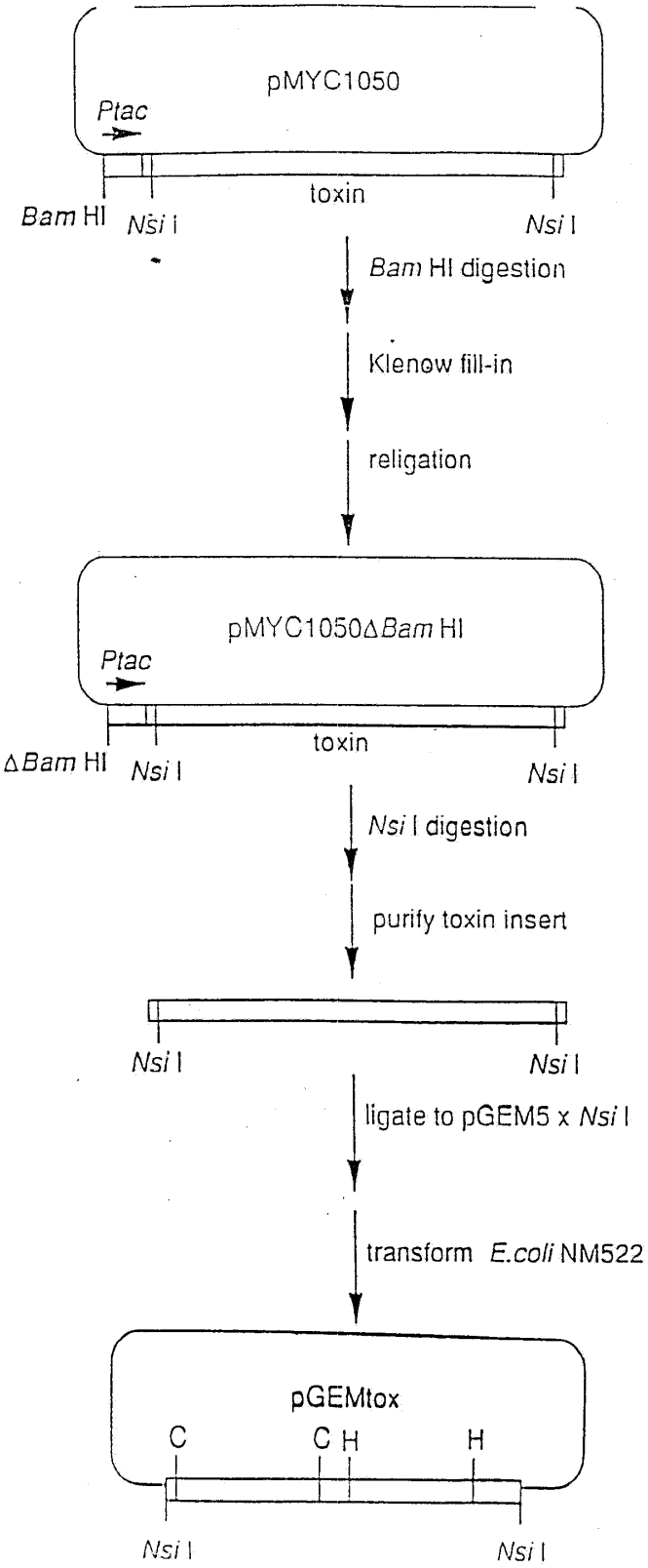
인시류 해충을 인시류-방제 유효량의 제 2항의 독소와 접촉시키는 과정을 포함하는 인시류 해충의 방제방법.

요약

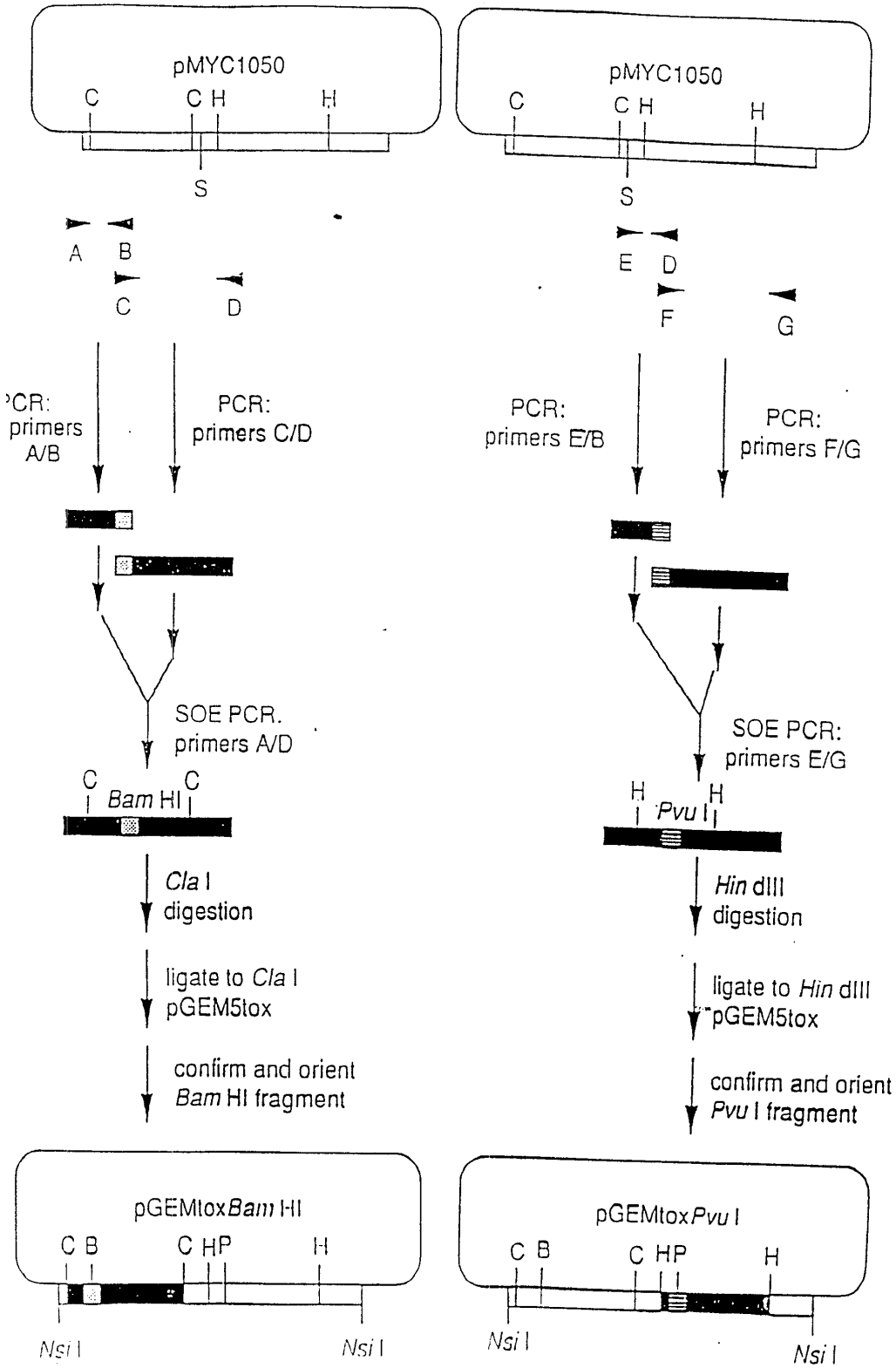
슈도모나스 플루오레스센스에서의 바실루스 츄린겐시스 내독소 발현은 바실루스 츄린겐시스 내독소를 코드화하는 유전자를 변형시킴으로써 향상될 수 있다. 키메라 유전자들은 네이티브 프로톡신을 코드화하는 바실루스 츄린겐시스 유전자의 분절을 상이한 프로톡신 분절을 코드화하는 분절로 대체함으로써 만들어질 수 있다. 본원에 예시될 것은 cryIF/cryI(b) 키메라로서, 여기서 네이티브 cryIF 프로톡신 분절은 cryIA(b) 프로톡신 분절에 의해 치환되어 슈도로나스에서 향상된 cryIF 독소의 발현을 제공한다. 본 발명은 또한 신규한 유전자 및 플라스미드에도 관계한다.

도면

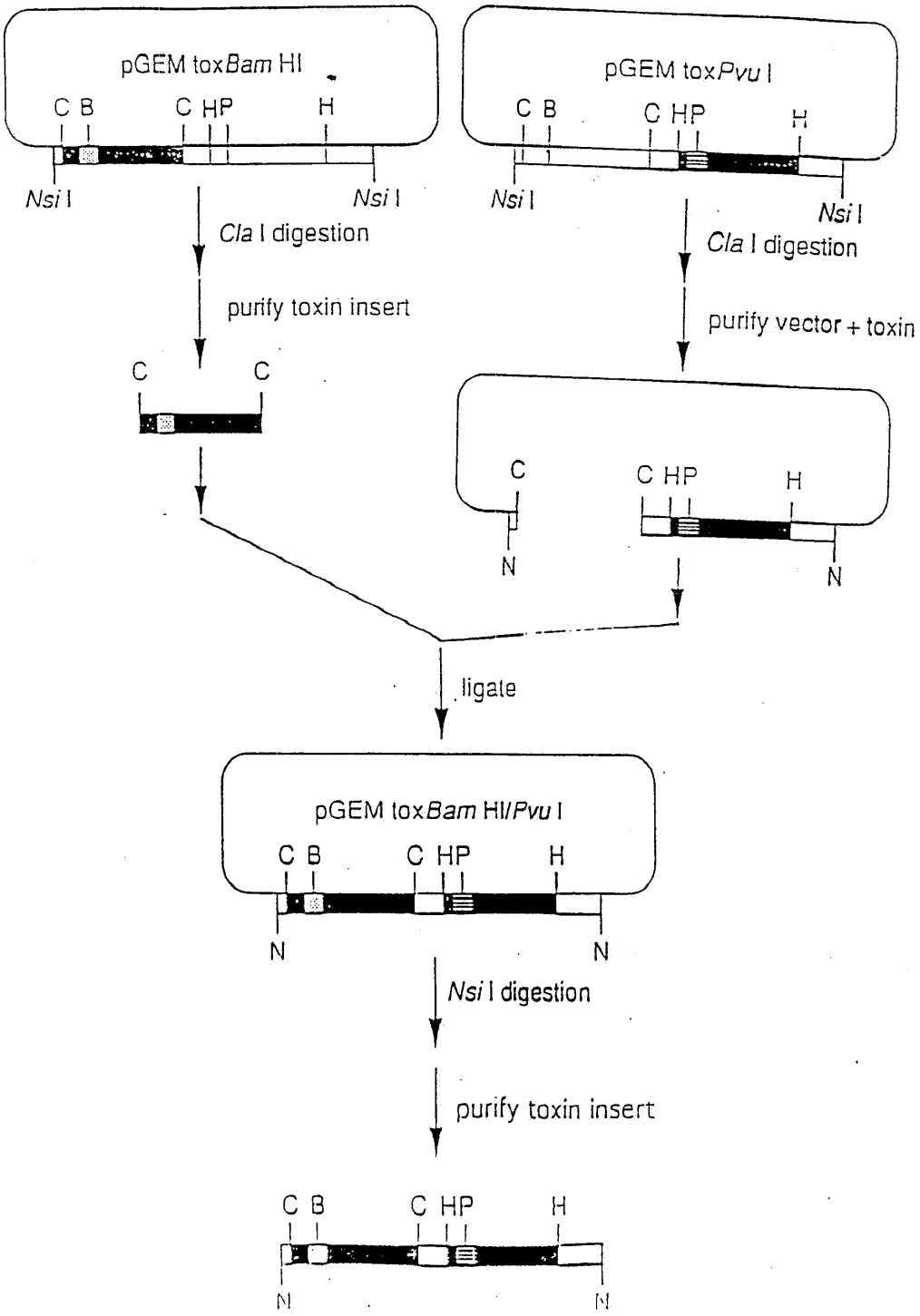
도면1


pMYC1050
Ptac
toxin
Bam HI
Nsi I
Nsi I
Bam HI digestion
Klenow fill-in
religation
pMYC1050ΔBam HI
Ptac
toxin
ΔBam HI
Nsi I
Nsi I
Nsi I digestion
purify toxin insert
Nsi I
Nsi I
ligate to pGEM5 x Nsi I
transform E.coli NM522
pGEMtox
C
C H
H
Nsi I
Nsi I

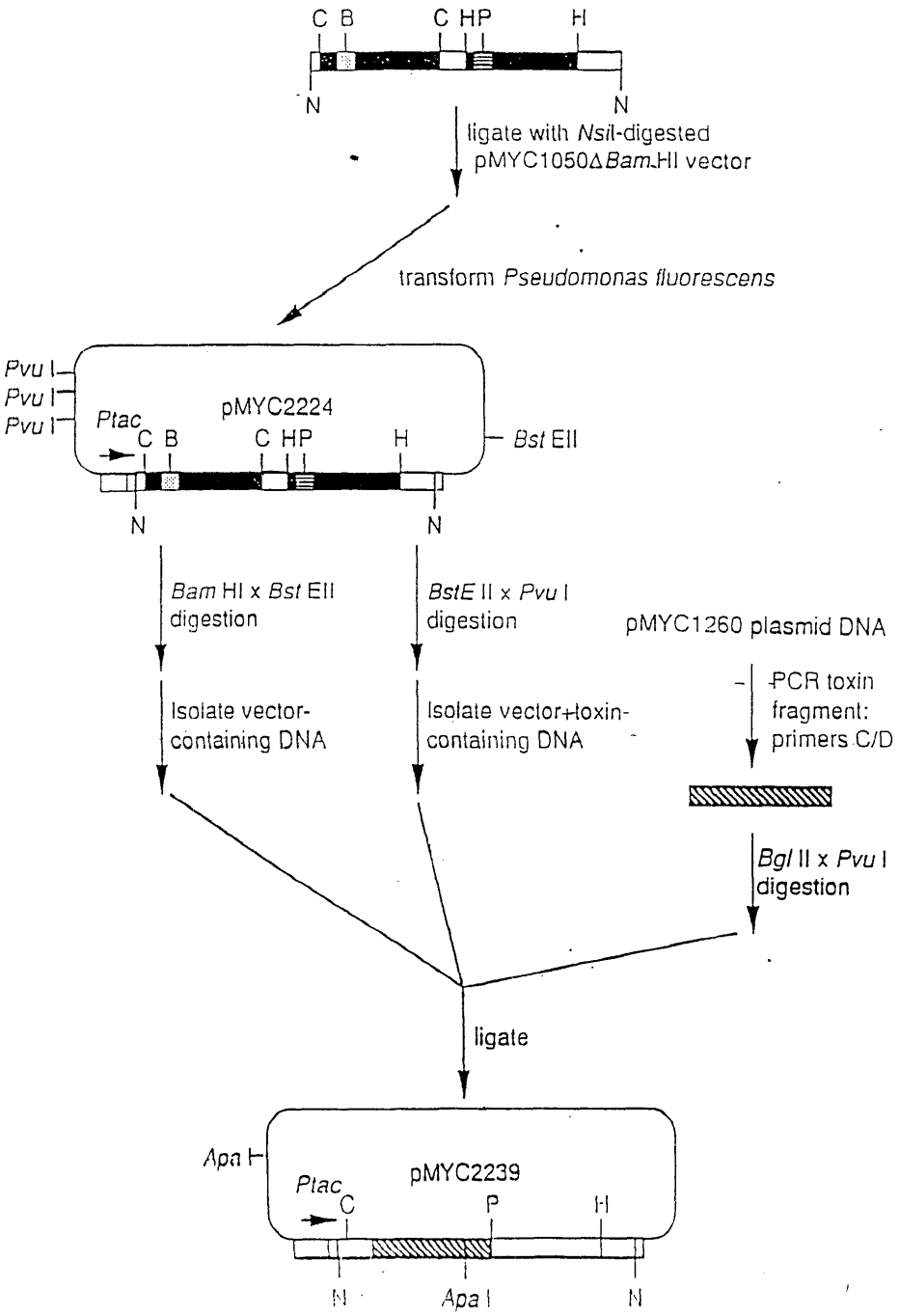
도면2


pMYC1050
C
C H
H
S
A
B
C
D
PCR:
primers
A/B
PCR:
primers C/D
SOE PCR.
primers A/D
C
Bam HI
C
Cla I
digestion
ligate to Cla I
pGEM5tox
confirm and orient
Bam HI fragment
pGEMtoxBam HI
C B
C HP
H
Nsi I
Nsi I
pMYC1050
C
C H
H
S
E
D
F
G
PCR:
primers E/B
PCR:
primers F/G
SOE PCR:
primers E/G
H
Pvu I
H
Hin dIII
digestion
ligate to Hin dIII
pGEM5tox
confirm and orient
Pvu I fragment
pGEMtoxPvu I
C B
C HP
H
Nsi I
Nsi I

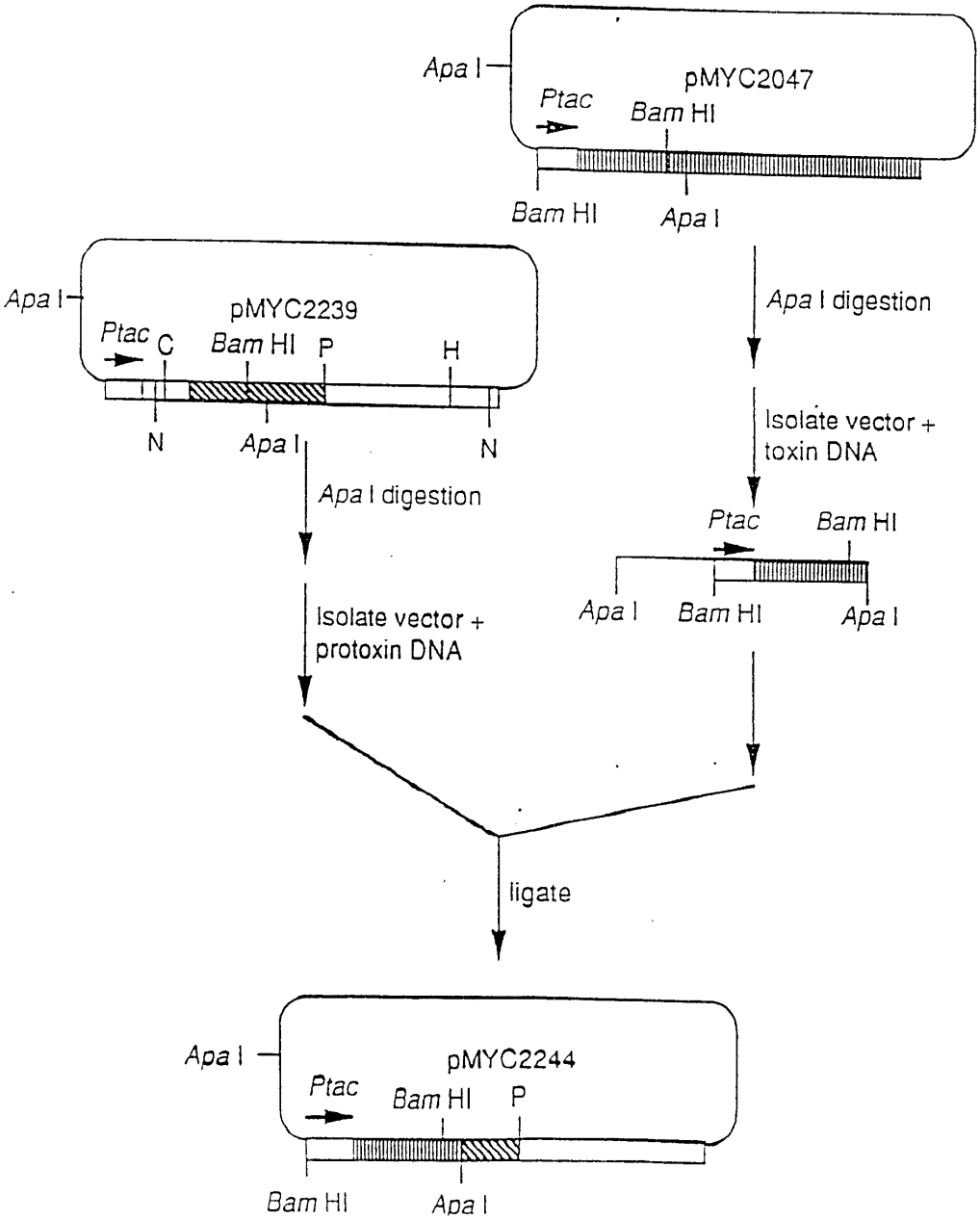
도면3


pGEM tox*Bam* HI
C B
C HP
H
Nsi I
Nsi I
Cla I digestion
purify toxin insert
C
C
pGEM tox*Pvu* I
C B
C HP
H
Nsi I
Nsi I
Cla I digestion
purify vector + toxin
C
C HP
H
N
N
ligate
pGEM tox*Bam* HI/*Pvu* I
C B
C HP
H
N
N
Nsi I digestion
purify toxin insert
C B
C HP
H
N
N

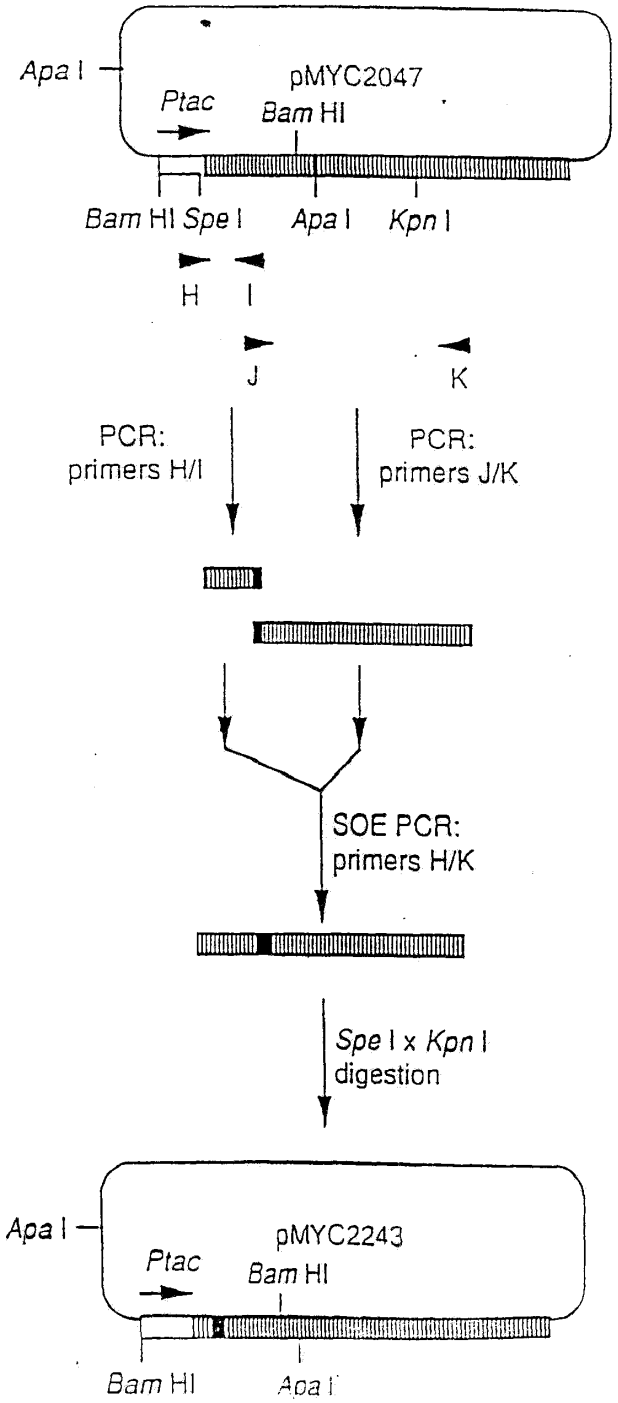
도면4


C B C HP H
N N
ligate with NsiI-digested pMYC1050ΔBamHI vector
transform Pseudomonas fluorescens
Pvu I
Pvu I
Pvu I
Ptac
pMYC2224
C B C HP H
Bst EII
N N
Bam HI x Bst EII digestion
BstE II x Pvu I digestion
pMYC1260 plasmid DNA
PCR toxin fragment: primers C/D
Isolate vector-containing DNA
Isolate vector+toxin-containing DNA
Bgl II x Pvu I digestion
ligate
Apa I
Ptac
pMYC2239
C P H
N Apa I N

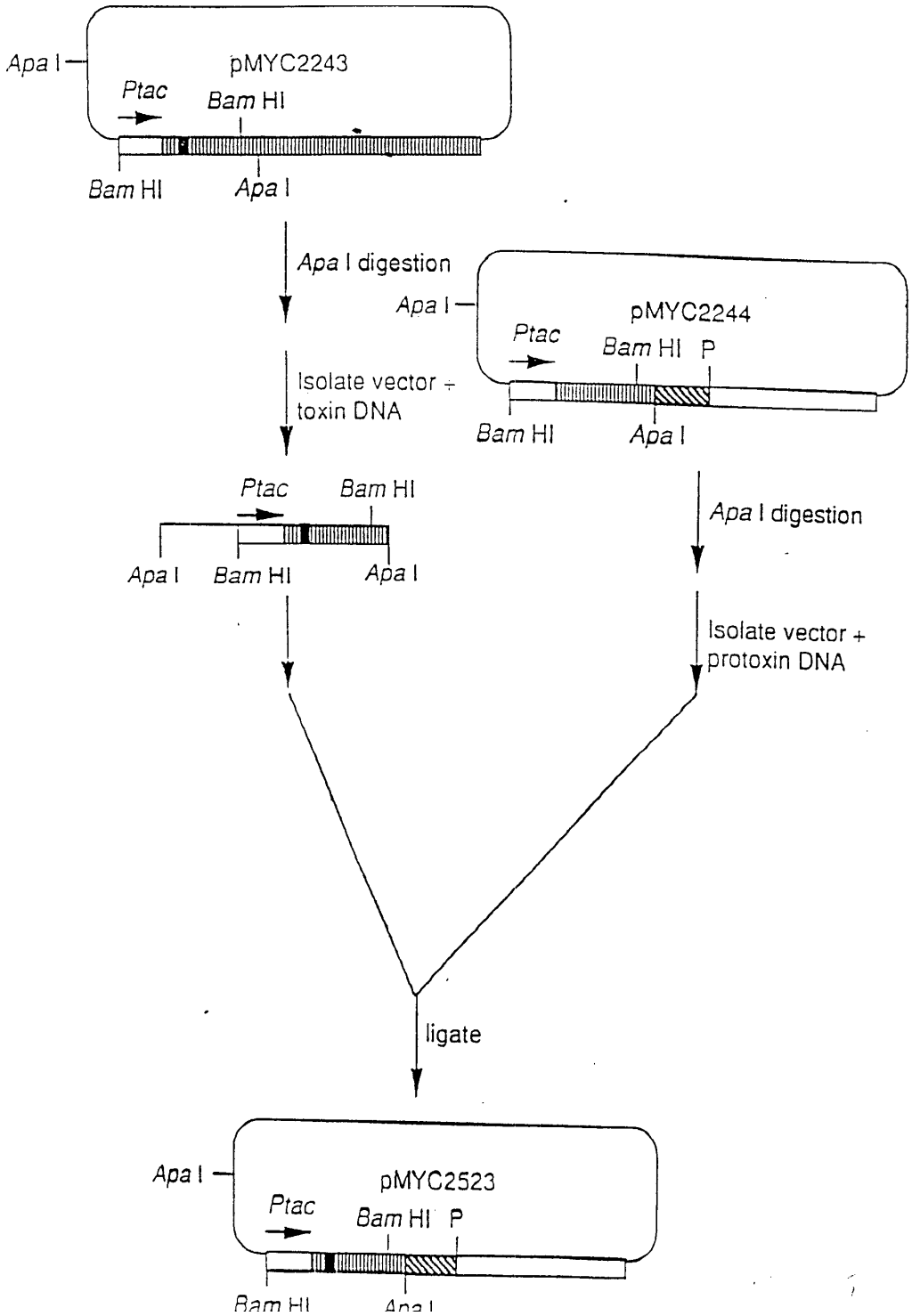
도면5


Apa I
pMYC2047
Ptac
Bam HI
Bam HI
Apa I
Apa I digestion
Apa I
pMYC2239
Ptac
C
Bam HI
P
H
N
Apa I
N
Isolate vector +
toxin DNA
Apa I digestion
Ptac
Bam HI
Apa I
Bam HI
Apa I
Isolate vector +
protoxin DNA
ligate
Apa I
pMYC2244
Ptac
Bam HI
P
Bam HI
Apa I

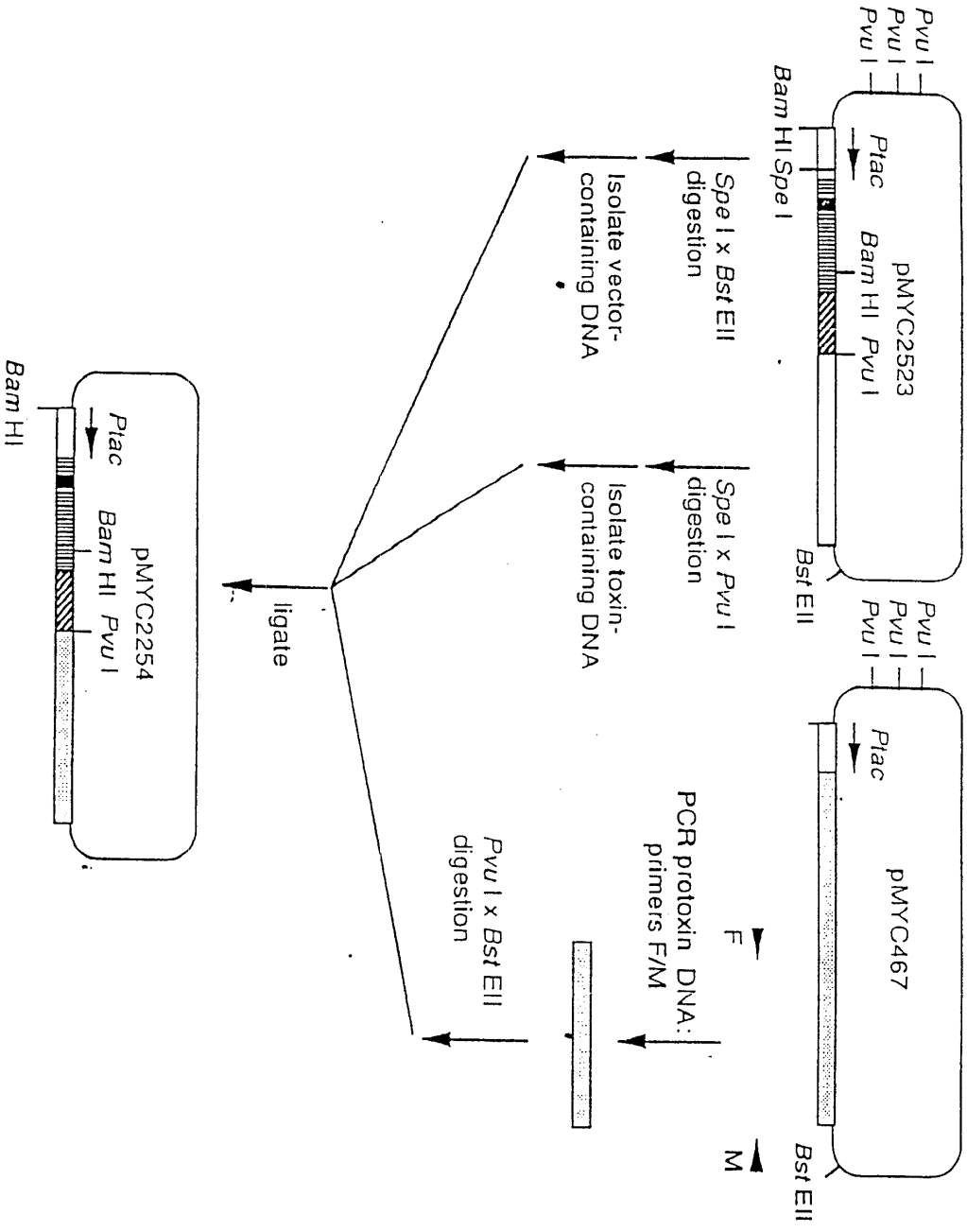
도면6


Apa I
pMYC2047
Ptac
Bam HI
Bam HI Spe I
Apa I
Kpn I
H
I
J
K
PCR:
primers H/I
PCR:
primers J/K
SOE PCR:
primers H/K
Spe I x Kpn I
digestion
Apa I
pMYC2243
Ptac
Bam HI
Bam HI
Apa I

도면7


Apa I
pMYC2243
Ptac
Bam HI
Bam HI
Apa I
Apa I digestion
Apa I
pMYC2244
Ptac
Bam HI P
Isolate vector + toxin DNA
Bam HI
Apa I
Ptac
Bam HI
Apa I
Bam HI
Apa I
Apa I digestion
Isolate vector + protoxin DNA
ligate
Apa I
pMYC2523
Ptac
Bam HI P
Bam HI
Apa I

도면8


Pvu I
Pvu I
Pvu I
Ptac
Bam HI Pvu I
pMYC2523
Bam HI Spe I
Bst EII
Spe I x Bst EII digestion
Isolate vector-containing DNA
Spe I x Pvu I digestion
Isolate toxin-containing DNA
ligate
Pvu I x Bst EII digestion
PCR protoxin DNA: primers F/M
F
M
pMYC467
Bam HI
pMYC2254

```
                                                                                                                  90
     1
Cons MENNIQNQCV PYNCLNNPEV EILNEERSTG RLPLDISLSL TRFLLSEFVP GVGVAFGLFD LIWGFITPSD WSLFLLQIEQ LIEQRIETLE
                                                                                                                 180
     91
Cons RNRAITTLRG LADSYEIYIE ALREWEANPN NAQLREDVRI RFANTDDALI TAINNFTLTS FEIPLLSVYV QAANLHLSLL RDAVSFGQGW
                                                                                                                 270
     181
Cons GLDIATVNNH YNRLINLIHR YTKHCLDTYN QGLENLRGTN TRQWARFNQF RRDLTLTVLD IVALFPNYDV RTYPIQTSSQ LTREIYTSSV
                                                                                                                 360
     271
Cons IEDSPVSANI PNGFNRAEFG VRPPHLMDFM NSLFVTAETV RSQTVWGGHL VSSRNTAGNR INFPSYGVFN PGGAIWIADE DPRPFYRTLS
                                                                                                                 450
     361
Cons DPVFVRGGFG NPHYVLGLRG VAFQQTGTNH TRTFRNSGTI DSLDEIPPQD NSGAPWNDYS HVLNHVTFVR WPGEISGSDS WRAPMFSWTH
                                                                                                                 540
     451
Cons RSATPTNTID PERITQIPLV KAHTLQSGTT VVRGPGFTGG DILRRTSGGP FAYTIVNING QLPQRYRARI RYASTTNLRI YVTVAGERIF
                                                                                                                 630
     541                                                                                                   t
Alt                                                                                                        i
Alt                                                                                  l    n      e   a     p  l i
Alt                                                                                                        p
Cons AGQFNKTMDT GDPLTFQSFS YATINTAFTF PMSQSSFTVG ADTFSSGNEV YIDRFELIPV TATfEAEYdL ERAQKAVNEL FTSSNQIGLK
                                                                                                                 720
     631                                                                                                     s
Alt                                e                                                                       p
Alt  n       Q      t              r          ng            s   kd    p     g   g     r
Cons TDVTDYHIDr VSNLVECLSD EFCLDEKKEL SEKVKHAKRL SDERNLLQDP NFRGINRQLD RGWRGSTDIT IQGGDDVFKE NYVTLLGTFD
                                                                                                                 810
     721                                                                        vq
Alt                                                                  l      r   fe s rKCGE PNRCAPHLEW NPDLDCSCRD
Alt    l              p     e                                                                                 ----------
Cons ECYPTYLYQK IDESKLKAYT RYQLRGYIED SQDLEIYLIR YNAKHETVNV PGTGSLWPLS APSPIG---- ---------- ----------
```

도면9a

```
      811                                                                                                  900
Alt   GE                   i       d                 e   i      gra                                 ql
Cons  --KCAHHSHH FSLDIDVGCT DLNEDLGVWV IFKIKTQDGH ARLGNLEFLE EK-PLVGEAL ARVKRAEKKW RDKREKLEWE TNIVYKEAKE

      901                                                                                                  990
Alt                                                                  d
Alt                 q             t        r q                 vg       k         f
Cons  SVDALFVNSQ YDRLQADTNI AMIHAADKRV HSIREAYLPE LSVIPGVNAA IFEELEGRIF TAFSLYDARN VIKNGDFNNG LSCWNVKGHV

      991                                                                                                 1080
Alt          q                                                t                       f                 n    g
Cons  DVEEQNNHRS VLVVPEWEAE VSQEVRVCPG RGYILRVTAY KEGYGEGCVT IHEIENNTDE LKFSNCVEEE VYPNNTVTCN DYTATQEEYE

      1081                                                                                                1170
Alt    a   c      et g  y         v                                          q
Cons  GTYTSRNRGY DGAYESNSSV PADYASAYEE KAYTDGRRDN PCESNRGYGD YTPLPAGYVT KELEYFPETD KVWIEIGETE GTFIVDSVEL

      1171
Cons  LLMEE
```

도면96