(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0117568
(43) 공개일자 2012년10월24일

(51) 국제특허분류(Int. Cl.)
C12N 1/21 (2006.01) *C12N 15/74* (2006.01)
C12P 17/08 (2006.01) *C07D 313/00* (2006.01)
(21) 출원번호 10-2011-0035384
(22) 출원일자 2011년04월15일
심사청구일자 2011년04월15일

(71) 출원인
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
윤여준
서울특별시 강남구 압구정로 201, 126동 202호 (압구정동, 현대아파트)
한아름
서울특별시 관악구 남부순환로218길 32, 우남하우징 2차 506호 (봉천동)
이동건
대구광역시 남구 봉덕동 1329-2번지 대덕맨션 105동 801호
(74) 대리인
손민

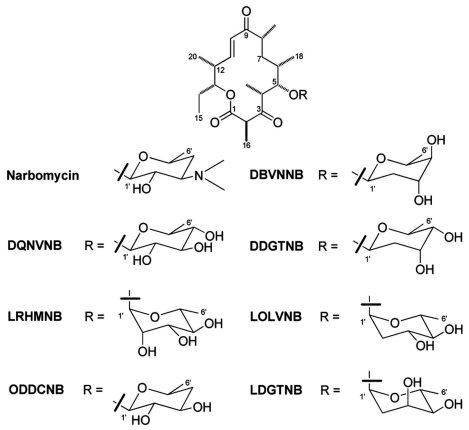
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 **신규 나르보마이신 유도체 화합물, 상기 화합물을 포함하는 항균용 조성물 및 상기 화합물의 생산 방법**

(57) 요 약

본 발명은 신규 나르보마이신 유도체 화합물, 상기 화합물을 포함하는 항균용 조성물 및 상기 화합물의 생산 방법에 관한 것이다. 본 발명의 생산 방법을 통하여 *생체 내* 조합 생합성의 방법으로 나르보마이신에 있어서 다양한 당 부분을 포함하는 나르보마이신 유도체를 합성할 수 있고, 특히 신규하고 항균 활성이 우수한 다양한 나르보마이신 유도체를 합성할 수 있다.

대 표 도 - 도3



이 발명을 지원한 국가연구개발사업

과제고유번호 20100018430

부처명 교육과학기술부

연구사업명 국가지정연구실사업

연구과제명 생합성경로 제설계를 통한 당포함 항생제 및 함암제(glycosylated antibiotic and anticancer)의 유도체 개발

주관기관 이화여자대학교 산학협력단

연구기간 2010.06.01 ~ 2011.05.31이 발명을 지원한 국가연구개발사업

과제고유번호 2010K000890

부처명 교육과학기술부

연구사업명 미생물유전체활용기술개발사업

연구과제명 방선균 생전환 시스템을 통한 폴리케타이드(polyketide) 화합물의 대사다양성 구축

주관기관 이화여자대학교 산학협력단

연구기간 2010.04.01 ~ 2011.03.31

특허청구의 범위

청구항 1

스트렙토마이세스 속 미생물로서, desVIII, desVII, desIII, desIV, desI, desII, desV, desVI, oleL, oleU, oleV, eryBII, urdR 및 oleW로 구성된 군으로부터 선택되는 단백질의 조합 중 desVIII, desVII, desIII 및 desIV 단백질을 포함하는 단백질의 조합을 코딩하는 핵산이 도입된, 나르보마이신 유도체 생산 재조합 미생물.

청구항 2

제1항에 있어서, 상기 단백질의 조합은 desVIII-desVII-desIII-desIV, desVIII-desVII-desIII-desIV-oleL-oleU, desVIII-desVII-desIII-desIV-desI-desII, desVIII-desVII-desIII-desIV-oleV-eryBII-urdR 또는 desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU 의 단백질의 조합인, 재조합 미생물.

청구항 3

제1항에 있어서, 상기 스트렙토마이세스 속 미생물은 스트렙토마이세스 베네주엘래인, 재조합 미생물.

청구항 4

제1항에 있어서, 상기 스트렙토마이세스 속 미생물은, 내재적으로 나르보마이신을 생합성하는 스트렙토마이세스 속 미생물에서 나르보마이신의 당 부분인 D-데소사민을 생합성 및 전이시키는 단백질을 코딩하는 유전자 클러스터가 게놈으로부터 결실된 것인, 재조합 미생물.

청구항 5

제4항에 있어서, 상기 유전자 클러스터는 desVIII-desVII-desIII-desIV-desI-desII-desV-desVI인, 재조합 미생물.

청구항 6

제5항에 있어서, 상기 desVIII-desVII-desIII-desIV-desI-desII-desV-desVI는 서열번호 1의 서열을 갖는 것인, 재조합 미생물.

청구항 7

제1항에 있어서, 상기 재조합 미생물은 서열번호 16, 서열번호 17, 서열번호 18, 서열번호 19 또는 서열번호 20의 핵산서열을 포함하는 발현 벡터가 도입된 것인, 재조합 미생물.

청구항 8

제1항에 있어서, 상기 재조합 미생물은 기탁번호 KFCC11507P 또는 KFCC11508P로 기탁된 것인, 재조합 미생물.

청구항 9

desVIII-desVII-desIII-desIV, desVIII-desVII-desIII-desIV-oleL-oleU, desVIII-desVII-desIII-desIV-desI-desII, desVIII-desVII-desIII-desIV-oleV-eryBII-urdR 또는 desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU 의 단백질의 조합을 코딩하는 핵산을 포함하는 발현 벡터.

청구항 10

제9항에 있어서, 상기 desVIII-desVII-desIII-desIV의 단백질 조합을 코딩하는 핵산은 서열번호 16; desVIII-desVII-desIII-desIV-oleL-oleU의 단백질 조합을 코딩하는 핵산은 서열번호 17; desVIII-desVII-desIII-desIV-desI-desII의 단백질 조합을 코딩하는 핵산은 서열번호 18; desVIII-desVII-desIII-desIV-oleV-eryBII-urdR의 단백질 조합을 코딩하는 핵산은 서열번호 19; 또는 desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU 의 단백질 조합을 코딩하는 핵산은 서열번호 20인, 발현 벡터.

청구항 11

제9항에 있어서, 상기 desVIII-desVII-desIII-desIV의 단백질의 조합을 코딩하는 핵산을 포함하는 발현 벡터는 도 7의 개열지도로 표시되고; desVIII-desVII-desIII-desIV-oleL-oleU의 단백질의 조합을 코딩하는 핵산을 포함하는 발현 벡터는 도 8의 개열지도로 표시되고; desVIII-desVII-desIII-desIV-desI-desII의 단백질의 조합을 코딩하는 핵산을 포함하는 발현 벡터는 도 9의 개열지도로 표시되고; desVIII-desVII-desIII-desIV-oleV-eryBII-urdR의 단백질의 조합을 코딩하는 핵산을 포함하는 발현 벡터는 도 10의 개열지도로 표시되고; 또는 desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU 의 단백질의 조합을 코딩하는 핵산을 포함하는 발현 벡터는 도 11의 개열지도로 표시되는 것인, 발현 벡터.

청구항 12

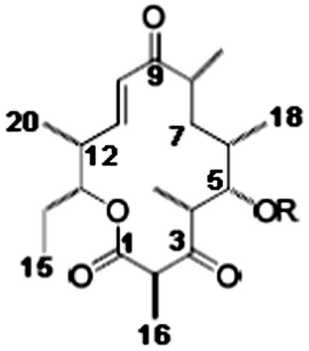
제1항 내지 제8항 중 어느 한 항의 재조합 미생물을 배양하는 단계; 및

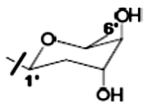
배양된 세포 또는 배양액으로부터 나르보마이신 유도체를 회수하는 단계를 포함하는, 나르보마이신 유도체의 제조방법.

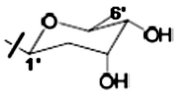
청구항 13

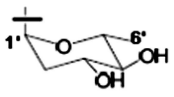
하기 화학식 1로 표시되는 나르보마이신 유도체 화합물, 그의 약제학적으로 허용가능한 염.

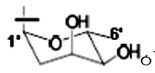
[화학식 1]



상기 식에서 R은 

, 

, 

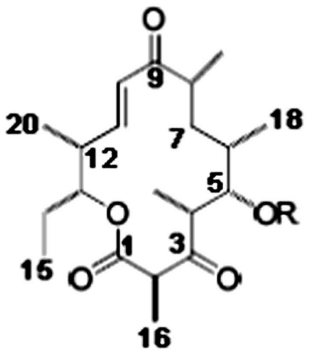
또는 

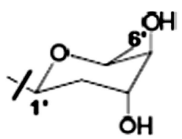
이다.

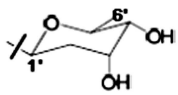
청구항 14

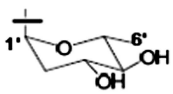
화학식 1의 화합물 또는 약제학적으로 허용가능한 염을 포함하는 항균용 조성물.

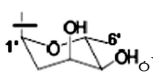
[화학식 1]



상기 식에서 R은 

, 

, 

또는 

이다.

청구항 15

제14항에 있어서, 상기 조성물은 에리트로마이신-감수성 균주 및 에리트로마이신-저항성 균주에 대하여 항균 활성을 가지는 것인, 조성물.

청구항 16

제14항에 있어서, 상기 조성물은, *E. 파에시움* ATCC 19434, *E. 파에시움* P00558, *S. 아우레우스* ATCC 25923 및 *S. 아우레우스* P00740로 구성된 군으로부터 선택된 미생물에 대하여 항균 활성을 가지는 것인, 조성물.

명 세 서

기 술 분 야

[0001] 본 발명은 신규 나르보마이신 유도체 화합물, 상기 화합물을 포함하는 항균용 조성물 및 상기 화합물의 생산 방법에 관한 것이다.

배 경 기 술

[0002] 감염성 세균들이 대부분의 항생제에 대한 내성을 획득하게 되면서, 기존 항생제에 대하여 내성을 갖는 세균이나 곰팡이, 바이러스, 암세포 등을 치료하기 위해, 새로운 종류의 항생제를 얻으려는 노력이 계속되고 있다. 이를 위하여 이제까지는 주로 자연계에 존재하는 미생물을 탐색하여 새로운 항생물질을 만들어내는 균주를 찾았으나, 이러한 신규 항생물질 탐색과정은 그에 들어가는 시간, 노력 및 비용이 엄청나고, 약리작용 및 독성시험까지 완료되어야만 하나의 사용 가능한 항생물질로 인정받을 수 있기 때문에 무척 까다롭고 힘든 작업이다. 따라서 많은 제약회사, 대학 및 연구소에서 기존의 방법과는 다른 돌파구를 기대하고 있다.

[0003] 그에 따라, 항생물질을 만들어내는 균주의 유전자를 조작하여 새로운 구조의 다양한 생리활성물질을 얻어내려는 방법이 시도되고 있다. 이러한 방법은 주로 폴리케타이드(polyketide) 계열 항생물질의 구조변형에 도입되고 있으며, 다양한 생합성 유전자의 조합을 통하여 다양한 구조를 생성할 수 있으므로 이를 조합 생합성(combinatorial biosynthesis)이라 부른다. 이러한 방법은 기존의 화학 합성법 또는 효소 합성법에 비해 탐색과정에서의 시간과 비용이 절약됨은 물론, 기존 항생제의 구조와 활성과의 관계 등을 응용하여 설계하기 때문에 생물학적 활성을 예상할 수 있어, 많은 시간과 비용을 절감할 수 있다.

[0004] 폴리케타이드 계열의 화합물은 미생물로부터 생산되는 복잡한 구조의 화합물로서, 현재까지 약 10,000가지 이상의 화합물이 알려져 있다. 폴리케타이드 계열의 화합물은 폴리케타이드 신타아제(polyketide synthase, PKS)라고 하는 다기능 단백질 복합체(multifunctional protein complex)에 의해 생합성된다. 이러한 PKS 복합체는 연속적인 축합반응에 관여하는 여러 개의 모듈(한 번의 축합반응을 책임지는 촉매 도메인들의 묶음)로 구성되어 있다. 모듈(module)을 구성하는 촉매 도메인(catalytic domain)들의 구성에 따라 폴리케타이드의 구조가 일대일로 대응하여 생합성되므로 PKS 유전자의 변형에 의해 폴리케타이드의 구조를 변형시킬 수 있다.

[0005] 한편, 마크롤라이드는 폴리케타이드 마크로락톤 환과 그에 부착된 하나 또는 그 이상의 데옥시당으로 구성된다. 상기 마크롤라이드는 스태필로코커스 및 엔테로코커스와 같은 그람-양성 박테리아에 대해 주로 항균 활성을 나타낸다. 마크롤라이드 중에서 데옥시당 부분(moiety)이 표적 분자와 특이적으로 접촉함으로써 분자의 생물학적 활성에 크게 영향을 미치기 때문에, 데옥시 당 부분의 구조가 중요하고, 따라서, 상기 언급한 PKS 유전자 변형에 의한 폴리케타이드 구조 변형 외에도 글리코실화와 같은 후-PKS 테일러링 단계(post-PKS tailoring step)가 반드시 필요하다. 마크롤라이드 계열 항생물질 저항성 병원체의 출현으로 인하여 신규 마크롤라이드 유도체를 발견하기 위한 연구가 계속되고 있다. 기존에는 화학적 및 효소적 합성을 이용한 연구가 많이 진행되어왔으나, 데옥시 당 부분의 화학적 수식 또는 뉴클레오티드-활성화되는 데옥시 당의 *시험관 내* 효소적 합성이 쉽지 않은 문제가 있다.

[0006] 나르보마이신은 *스트렙토마이세스 베네주엘래*에서 형성되는 마크롤라이드 계열 물질로서, PikC에 의하여 네오-, 노바-, 및 피크로마이신으로 전환된다(도 1). 나르보마이신은 *스태필로코커스 아우레우스*, *스트렙토코커스 피오제네스*, *코리네박테리움 디프테리애*, 그람-양성 코커스의 에리트로마이신-저항성 균주 등에 대한 우수한 활성을 가지는 것으로 보고되어 왔다. 천연 나르보마이신은 당 부분으로서 데옥시당 D-데소사민을 포함한다. *스트렙토마이세스 베네주엘래*의 글리코실트랜스퍼라아제인 DesVII 및 그의 보조 상대인 DevVIII가 나르보마이신의 아글리콘(비당(非糖) 성분)인 나르보놀라이드로 TDP-데옥시 당을 전이시키고, 여섯 개의 오픈 리딩 프레임(open

reading frame) desI 내지 desVI은 데소사민 생합성에 관여하는 것으로 추정된다. 상기 단백질을 이용하여 *시험관 내* 및 *생체 내*에서 나르보마이신 유도체의 글리코실다양화를 유발할 수 있는 것으로 생각되었으나, *생체 내* 조합 생합성의 방법으로 나르보마이신에 있어서 다양한 당 부분을 포함하는 나르보마이신 유도체를 합성하여 그의 항균 활성을 보고한 예는 없다.

발명의 내용

해결하려는 과제

[0007] 이러한 배경에서, 본 발명자들은 다양한 당 생합성 관련 단백질의 조합을 코딩하는 핵산을 포함하는 카세트를 제작하여 나르보놀라이드를 축적하는 *스트렙토마이세스 베네주엘래*에서 이종 발현시킴으로써 다양하고 신규한 나르보마이신 유도체를 생산할 수 있음을 발견하여 본 발명을 완성하였다.

과제의 해결 수단

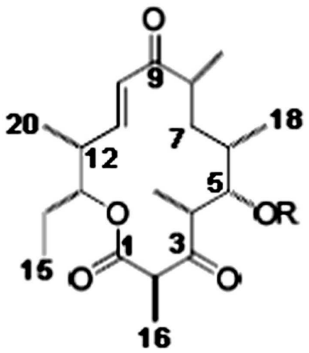
[0008] 본 발명은 스트렙토마이세스 속 미생물로서, desVIII, desVII, desIII, desIV, desI, desII, desV, desVI, oleL, oleU, oleV, eryBII, urdR 및 oleW로 구성된 군으로부터 선택되는 단백질의 조합 중 desVIII, desVII, desIII 및 desIV 단백질을 포함하는 단백질의 조합을 코딩하는 핵산이 도입된, 나르보마이신 유도체 생산 재조합 미생물을 제공하고자 한다.

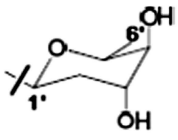
[0009] 또한, 본 발명은 desVIII-desVII-desIII-desIV, desVIII-desVII-desIII-desIV-oleL-oleU, desVIII-desVII-desIII-desIV-desI-desII, desVIII-desVII-desIII-desIV-oleV-eryBII-urdR 또는 desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU 의 단백질의 조합을 코딩하는 핵산을 포함하는 발현 벡터를 제공하고자 한다.

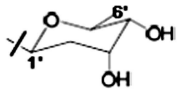
[0010] 또한, 본 발명은 상기 재조합 미생물을 배양하는 단계; 및 배양된 세포 또는 배양액으로부터 나르보마이신 유도체를 회수하는 단계를 포함하는 나르보마이신 유도체의 제조방법을 제공하고자 한다.

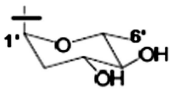
[0011] 또한, 본 발명은 하기 화학식 1로 표시되는 나르보마이신 유도체 화합물, 그의 약제학적으로 허용가능한 염을 제공하고자 한다.

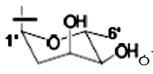
화학식 1

[0012] 

[0013] 상기 식에서 R은 

, 

, 

또는 

이다.

[0014] 또한, 본 발명은 상기 화학식 1의 화합물 또는 약제학적으로 허용가능한 염을 포함하는 항균용 조성물을 제공하고자 한다.

[0015] 하나의 양태로서, 본 발명은 스트렙토마이세스 속 미생물로서, desVIII, desVII, desIII, desIV, desI, desII, desV, desVI, oleL, oleU, oleV, eryBII, urdR 및 oleW로 구성된 군으로부터 선택되는 단백질의 조합 중

desVIII, desVII, desIII 및 desIV 단백질을 포함하는 단백질의 조합을 코딩하는 핵산이 도입된, 나르보마이신 유도체 생산 재조합 미생물을 제공한다. 상기 단백질의 조합은 desVIII-desVII-desIII-desIV, desVIII-desVII-desIII-desIV-oleL-oleU, desVIII-desVII-desIII-desIV-desI-desII, desVIII-desVII-desIII-desIV-oleV-eryBII-urdR 또는 desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU 의 단백질의 조합일 수 있다.

[0016] 본 발명에서 스트렙토마이세스 속 미생물은, 내재적으로 나르보마이신을 생합성할 수 있는 스트렙토마이세스 속 미생물을 의미한다. 상기 내재적으로 나르보마이신을 생합성할 수 있는 미생물이란, 내재적으로 나르보마이신 생합성에 필요한 모든 효소를 발현하는 미생물을 의미한다. 상기 "내재적"이란 야생형 미생물이 천연의 상태로 가지고 있는 활성을 의미한다. 본 발명에서 나르보마이신은 마크롤라이드 계열 물질로서, 도 1에 도시된 바와 같은 구조를 가진다.

[0017] 본 발명의 스트렙토마이세스 속 미생물은 *스트렙토마이세스 베네주엘래*일 수 있으나, 이에 한정되지 않고 나르보마이신을 생산할 수 있는 모든 스트렙토마이세스 속 미생물을 포함한다. *스트렙토마이세스 베네주엘래*는 유전자 조작이 쉽고 상대적으로 빠른 성장 속도 (doubling time, *ca.* 1 h)를 가지므로, 대사물질 생산을 위하여 필요한 배양 기간이 짧다. 상기 미생물은 단일 폴리케타이드 신타아제 (PKS) 및 다수의 PKS 후 수식 효소들 (데소사민 당의 생합성 및 전이에 연관된 des 클러스터에 의한 글리코실화 및 PikC P450 모노옥시게나아제에 의한 히드록실화)의 작용을 통하여 12- 및 14-원환 마크롤라이드, 메티마이신 및 피크로마이신을 모두 생산할 수 있다 (도 1). 본 발명의 구체적 실시예에서, *스트렙토마이세스 베네주엘래* ATCC 15439를 사용할 수 있다.

[0018] 바람직하게는, 상기 스트렙토마이세스 속 미생물은, 내재적으로 나르보마이신을 생합성할 수 있는 스트렙토마이세스 속 미생물에서 나르보마이신의 당 부분인 D-데소사민을 생합성 및 전이시키는 단백질을 코딩하는 유전자 클러스터가 게놈으로부터 결실된 것일 수 있다. 나르보마이신의 당 부분을 생합성 및 전이시키는 단백질을 코딩하는 유전자 클러스터가 게놈으로부터 결실된 스트렙토마이세스 속 미생물은 나르보마이신의 아글리콘(비당(非糖) 성분)인 나르보놀라이드를 축적하게 된다. 상기 유전자 클러스터는 D-데소사민을 생합성 및 전이시키는데 필요한 단백질군을 코딩하는 유전자 클러스터(GenBank: AF079762.1)로서 desVIII-desVII-desIII-desIV-desI-desII-desV-desVI(서열번호 1)로 구성될 수 있다. 상기 미생물은 게놈으로부터 서열번호 1의 염기서열로 구성된 유전자 클러스터가 결실된 것일 수 있다.

[0019] 게놈상에서 목적한 유전자만의 특이적 결실은 당 분야에서 확립된 방법을 통하여 수행할 수 있으며, 방법은 특별히 제한되지 않으며, 상동 재조합(homologous recombination)법을 사용할 수도 있다. 목적한 단백질의 N 말단과 C 말단을 코딩하는 핵산 사이에 선별마커를 포함하는 벡터로 스트렙토마이세스 속 균주를 형질전환시켜 게놈과 벡터 사이에 재조합을 유도할 수 있다. 사용되는 선별마커(selection marker)는 특별히 제한되지 않고, 약물 내성, 영양 요구성, 세포독성제에 대한 내성 또는 표면 단백질의 발현과 같은 선택가능 표현형을 부여하는 마커들이 사용될 수 있다.

[0020] 바람직하게는 *스트렙토마이세스 베네주엘래* YJ003을 사용할 수 있다. 상기 균주는 본 발명의 실시예에 기재된 방법에 의하여 제조할 수 있다.

[0021] 본 발명은 나르보마이신 유도체 생산 경로에서 상기 단백질들이 수행하는 기능을 도 2에서 확인할 수 있는 바와 같이 밝혀내었다. 또한, 상기 단백질의 다양하고 신규한 조합을 발현하는 재조합 미생물을 제조함으로써 변형된 당 부분을 가지는 신규 나르보마이신 유도체를 생합성하였고, 그들의 항균 활성도 평가하였다.

[0022] 본 발명의 나르보마이신 유도체 생산에 필요한 단백질 서열정보는 GenBank: AF079762.1, AF055579.2, U77454.1, AF269227.1에서 확인할 수 있다. 구체적으로, 본 발명에서 DesIII는 글루코오스-1-포스페이트를 기질로 하는 Nucleotidylyl transferase(NT)기능을 가지는 단백질로서, 서열번호 2의 아미노산 서열을 가진다. 본 발명에서 DesIV는 4,6-데히드라타아제(4,6-dehydratase, 4,6-DH) 기능을 하는 단백질로서, 서열번호 3의 아미노산 서열을 가진다. 본 발명에서 DesI은 아미노트랜스퍼라아제(aminotrasferase, AT) 기능을 하는 단백질로서, 서열번호 4의 아미노산 서열을 가진다. 본 발명에서 DesII는 데아미나아제(deaminase, DA) 기능을 하는 단백질로서, 서열번호 5의 아미노산 서열을 가진다. 본 발명에서 DesV는 아미노트랜스퍼라아제(aminotransferase, AT) 기능을 하는 단백질로서, 서열번호 6의 아미노산 서열을 가진다. 본 발명에서 DesVI는 *N*-메틸트랜스퍼라아제(*N*-methyltransferase, N-MT) 기능을 하는 단백질로서, 서열번호 7의 아미노산 서열을 가진다. 본 발명에서 OleL은 에피머라아제(epimerase, EP) 기능을 하는 단백질로서, 서열번호 8의 아미노산 서열을 가진다. 본 발명에서 OleU는 4-케토리덕타아제(4-ketoreductase, 4-KR) 기능을 하는 단백질로서, 서열번호 9의 아미노산 서열을 가진

다. 본 발명에서 OleV는 2,3-데히드라타아제(2,3-dehydratase, 2,3-DH) 기능을 하는 단백질로서, 서열번호 10의 아미노산 서열을 가진다. 본 발명에서 EryBII는 3-케토리덕타아제(3-ketoreductase, 3-KR) 기능을 하는 단백질로서, 서열번호 11의 아미노산 서열을 가진다. 본 발명에서 UrdR은 4-케토리덕타아제(4-ketoreductase, 4-KR) 기능을 하는 단백질로서, 서열번호 12의 아미노산 서열을 가진다. 본 발명에서 OleW는 3-케토리덕타아제(3-ketoreductase, 3-KR) 기능을 하는 단백질로서, 서열번호 13의 아미노산 서열을 가진다. 본 발명에서 DesVIII는 DesVII의 보조 상대 기능을 하는 단백질로서, 서열번호 14의 아미노산 서열을 가진다. 본 발명에서 DesVII은 글리코실트랜스퍼라아제 기능을 하는 단백질로서, 서열번호 15의 아미노산 서열을 가진다.

[0023] 상기 단백질의 조합을 코딩하는 핵산을 도입하는 것은, 해당 핵산을 염색체상에 추가 도입할 수도 있고, 해당 핵산을 벡터 상에 자가 프로모터 또는 강화된 별개의 프로모터와 함께 도입할 수도 있다. 바람직하게는, 발현벡터를 사용해서 스트렙토마이세스 속 미생물을 형질전환시킬 수 있다. 발현 벡터를 이용하여 형질전환된 재조합 미생물을 제조하기 위한 방법으로는 문헌(Sambrook, J., et al., Molecular Cloning, A Laboratory Manual(2판), Cold Spring Harbor Laboratory, 1. 74, 1989)에 기재된 인산칼슘법 또는 염화칼슘/염화루비듐법, 일렉트로포레이션법(electroporation), 전기주입법(electroinjection), PEG 등의 화학적 처리방법, 유전자 총(gene gun) 등을 이용하는 방법 등이 있다.

[0024] 본 발명의 구체적 실시예에서, 재조합 미생물 YJ003/pDDSS(글리코실트랜스퍼라아제 유전자 *desVIII/desVII*와 함께 D-데소사민 [DDSS] 생합성 유전자 조합(*desIII-desIV-desI-desII-desV-desVI*)을 발현하는 플라스미드), YJ003/pDQNV(*desVIII/desVII*와 함께 D-퀴노보오스 [DQNV] 생합성 유전자 조합(*desIII-desIV*)을 발현하는 플라스미드), 및 YJ003/pODDC(*desVIII/desVII*와 함께 3'-*O*-데메틸-D-찰코오스 [ODDC] 생합성 유전자 조합(*desIII-desIV-desI-desII*)을 발현하는 플라스미드)는 나르보마이신, D-퀴노보오실-나르보놀라이드(DQNVNB), 및 3'-*O*-데메틸-D-찰코오실-나르보놀라이드(ODDCNB) 을 각각 생산하였다 (도 2 내지 도 4). 그러나, 재조합 YJ003/pLRHM2(*desVIII/desVII*와 함께 L-람노오스 [LRHM] 생합성 유전자 조합(*desIII-desIV-oleL-oleU*)을 발현하는 플라스미드)의 발효로부터 2개의 유도체인 L-람노오실-나르보놀라이드(LRHMNB) 및 DQNVNB가 동시에 생산되었다(도 3 및 도 4). 상기 균주에서의 DQNVNB의 동시 생산은 *스트렙토마이세스 베네주엘래* 경로와 독립적인 4-케토리덕타아제(SvRed)에 의해서 LRHM의 생합성 중간체인 4-케토-6-데옥시-D-글루코오스가 DQNV로 환원되었기 때문이다(도 2). 재조합 YJ003/pDDGT2(*desVIII/desVII*와 함께 D-디지톡소오스 [DDGT] 생합성 유전자 조합(*desIII-desIV-oleV-eryBII-urdR*)을 발현하는 플라스미드)는 D-디지톡소오실-나르보놀라이드(DDGTNB) 및 D-보이비노오실-나르보놀라이드(DBVNNB)를 생산하였고, 재조합 YJ003/pLOLV2(*desVIII/desVII*와 함께 L-올리보오스 [LOLV] 및 L-디지톡소오스 [LDGT] 생합성 유전자 조합(*desIII-desIV-oleV-oleW-oleL-oleU*)을 발현하는 플라스미드)는 L-올리보오실-나르보놀라이드(LOLVNB), L-디지톡소오실-나르보놀라이드(LDGTNB), 및 DBVNNB를 생산하였다 (도 3 및 도 4). 두 균주에서 DBVNNB를 생산한 것은 천연 *스트렙토마이세스 베네주엘래* 3-케토리덕타아제가 D-디지톡소오스, L-올리보오스 및 L-디지톡소오스의 중간체로서 2,6-다이데옥시-3,4-다이케토-D-글루코오스를 인식하고, 다른 동정되지 않은 *스트렙토마이세스 베네주엘래* 4-케토리덕타아제에 의한 4-케토환원과 조합하여, C3에서 환원시켜 D-보이비노오스를 생산한 것으로 생각된다. YJ003/pLOLV2 에서 LOLVNB 및 LDGTNB가 동시에 생산된 것은 OleU 4-케토리덕타아제가 C-3 에서 C-4로의 토오토메리화(tautomerization) 때문에 형성된 TDP-4-케토-LOLV 및 TDP-4-케토-LDGT상에 작용하여 LOLV 및 LDGT를 각각 합성한 것으로 생각된다(도 2).

[0025] 본 발명의 구체적 실시예에서 제조한 상기 재조합 미생물 중에서, pDDGT2(*desVIII/desVII*와 함께 D-디지톡소오스 [DDGT] 생합성 유전자 조합(*desIII-desIV-oleV-eryBII-urdR*)을 발현하는 플라스미드)가 도입된 재조합 미생물 및 pLOLV2(*desVIII/desVII*와 함께 L-올리보오스 [LOLV] 및 L-디지톡소오스 [LDGT] 생합성 유전자 조합(*desIII-desIV-oleV-oleW-oleL-oleU*)을 발현하는 플라스미드)가 도입된 재조합 미생물을 한국미생물보존센터(Korean Culture Center of Microorganisms, KCCM)에 2011년 3월 10일자로 각각 기탁번호 KFCC11507P, KFCC11508P로 기탁하였다.

[0026] 또 하나의 양태로서, 본 발명은 desVIII-desVII-desIII-desIV, desVIII-desVII-desIII-desIV-oleL-oleU, desVIII-desVII-desIII-desIV-desI-desII, desVIII-desVII-desIII-desIV-oleV-eryBII-urdR 또는 desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU 의 단백질의 조합을 코딩하는 핵산을 포함하는 발현 벡터를 제공한다.

[0027] 본 발명에서 "발현 벡터"란 적당한 숙주세포에서 목적 단백질을 발현할 수 있는 재조합 벡터로서, 유전자 삽입물이 발현되도록 작동가능하게 연결된 필수적인 조절 요소를 포함하는 유전자 작제물로서, 이러한 유전자 작제

물을 제조하기 위해 표준 재조합 DNA 기술을 이용할 수 있다. 본 발명에서 용어, "작동가능하게 연결된(operably linked)"은 일반적 기능을 수행하도록 핵산 발현 조절 서열과 목적하는 단백질을 코딩하는 핵산 서열이 기능적으로 연결(functional linkage)되어 있는 것을 말한다. 예를 들어 프로모터와 단백질을 코딩하는 핵산 서열이 작동가능하게 연결되어 코딩하는 핵산 서열의 발현에 영향을 미칠 수 있다. 발현 벡터의 종류는 원핵세포의 각종 숙주세포에서 목적하는 유전자를 발현하고, 목적하는 단백질을 생산하는 기능을 하면 특별히 한정되지 않지만, 강력한 활성을 나타내는 프로모터와 강한 발현력을 보유하면서 자연 상태와 유사한 형태의 외래 단백질을 대량으로 생산할 수 있는 벡터가 바람직하다. 발현 벡터는 적어도, 프로모터, 개시코돈, 목적하는 단백질을 암호화하는 유전자, 종결코돈 터미네이터를 포함하고 있는 것이 바람직하다. 그 외에 시그널 펩티드를 코딩하는 DNA, 오퍼레이터 서열, 목적하는 유전자의 5측 및 3측의 비해독영역, 선별 마커 영역, 또는 복제가능단위 등을 적절하게 포함할 수도 있다. 개시 코돈 및 종결코돈은 일반적으로 목적 단백질을 코딩하는 핵산 서열의 일부로 간주될 수 있으며, 유전자 작제물이 도입되었을 때 반드시 작용을 나타내야 하며 코딩 서열과 인프레임(in frame)에 있어야 한다. 벡터의 프로모터는 구성적 또는 유도성일 수 있다.

[0028] 본 발명의 구체적 실시예에서, desVIII-desVII-desIII-desIV의 단백질 조합을 코딩하는 핵산은 서열번호 16, desVIII-desVII-desIII-desIV-oleL-oleU의 단백질 조합을 코딩하는 핵산은 서열번호 17, desVIII-desVII-desIII-desIV-desI-desII의 단백질 조합을 코딩하는 핵산은 서열번호 18, desVIII-desVII-desIII-desIV-oleV-eryBII-urdR의 단백질 조합을 코딩하는 핵산은 서열번호 19, desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU 의 단백질 조합을 코딩하는 핵산은 서열번호 20일 수 있다.

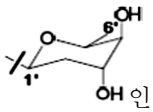
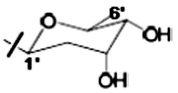
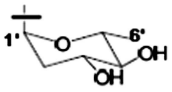
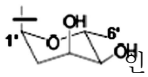
[0029] 또한, 본 발명의 구체적 실시예에서, pDQNV(desVIII-desVII-desIII-desIV의 단백질 조합을 코딩하는 서열번호 16의 핵산을 포함하는 발현벡터)는 도 7의 개열지도, pLRHM2(desVIII-desVII-desIII-desIV-oleL-oleU의 단백질 조합을 코딩하는 서열번호 17의 핵산을 포함하는 발현벡터)는 도 8의 개열지도, pODDC(desVIII-desVII-desIII-desIV-desI-desII의 단백질 조합을 코딩하는 서열번호 18의 핵산을 포함하는 발현벡터)는 도 9의 개열지도, pDDGT2(desVIII-desVII-desIII-desIV-oleV-eryBII-urdR의 단백질 조합을 코딩하는 서열번호 19의 핵산을 포함하는 발현벡터)는 도 10의 개열지도, pLOLV2(desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU의 단백질 조합을 코딩하는 서열번호 20의 핵산을 포함하는 발현벡터)는 도 11의 개열지도, pDDSS(desVIII-desVII-desIII-desIV-desI-desII-desV-desVI의 단백질 조합을 코딩하는 서열번호 1의 핵산을 포함하는 발현벡터)는 도 12의 개열지도를 가질 수 있다.

[0030] 또 다른 양태로서, 본 발명은 상기 재조합 미생물을 배양하는 단계; 및 배양된 세포 또는 배양액으로부터 나르보마이신 유도체를 회수하는 단계를 포함하는, 나르보마이신 유도체의 제조방법을 제공한다.

[0031] 본 발명에서 용어 "배양"은 미생물을 적당히 인공적으로 조절한 환경조건에서 생육시키는 것을 의미한다. 본 발명의 신규 나르보마이신 유도체를 생산하기 위한 미생물 배양은 통상의 미생물이 사용할 수 있는 영양원을 함유하는 배지에서 배양할 수 있다. 미생물의 영양원으로는 당업계에서 통상적으로 사용되는 영양원을 제한 없이 사용할 수 있으며, 바람직하게는 종래 스트렙토마이세스 속 박테리아의 배양에 이용되고 있는 공지의 영양원을 사용할 수 있다. 예를 들어, 말로닉산(malonic acid), 에탄올, 메치오닌(methionine), 탄소원 및 질소원을 포함하는 배지에서 이루어지는 것이 바람직하다. 이때, 상기 탄소원은 전분, 포도당, 옥수수기름, 글리세롤, 말토스, 만노스 및 이노시톨로 구성된 군으로부터 선택된 하나 이상인 것이 바람직하며, 전분, 포도당 및 옥수수기름인 것이 가장 바람직하다. 또한, 상기 질소원은 면실밀, 옥수수 침지액, 옥수수 침지분, 대두분, 펩톤 및 효모엑기스로 구성된 군으로부터 선택된 하나 이상인 것이 바람직하며, 면실밀 및 옥수수침지액인 것이 가장 바람직하다. 종균배양액을 상기 배양 배지를 포함한 발효조에 식균하여 배양하였다. 배양방법으로는 호기적 조건에서는 진탕배양 혹은 정치배양이 가능하다. 배양온도는 상기의 각 조건들에서 배양할 경우 조건에 따라 약간씩 상이하기는 하나, 보통 20-37℃에서 배양하는 것이 적당하며, 바람직하게는 26-30℃에서 배양할 수 있다. 또한, 배양기간 역시 당업계에서 사용되는 공지의 기간 동안 배양할 수 있으며, 필요에 따라 기간이 조정될 수 있다. 바람직하게 진탕배양, 정치배양의 경우 모두 4일 내지 7일간 배양할 수 있다.

[0032] 배양된 세포 또는 배양액으로부터 나르보마이신 유도체를 회수하는 방법은 당업계에 널리 알려진 방법에 따라 수행할 수 있다. 상기 나르보마이신 유도체 회수 방법에는, 여과, 음이온 교환 크로마토그래피, 결정화 및 HPLC 등이 사용될 수 있으나, 이들 예에 한정되는 것은 아니다.

[0033] 또 다른 양태로서, 본 발명은 상기 화학식 1로 표시되는 나르보마이신 유도체 화합물, 그의 약제학적으로 허용

가능한 염을 제공한다. 상기 화학식 1로 표시되는 나르보마이신 유도체 중 R이 인 경우는 D-보이비노오실-나르보놀라이드(DBVNNB), R이  인 경우는 D-디지톡소오실-나르보놀라이드 (DDGTNB), R이  인 경우는 L-올리보오실-나르보놀라이드 (LOLVNB), R이 인 경우는 L-디지톡소오실-나르보놀라이드 (LDGTNB)이다.

[0034] 핵자기공명 스펙트럼을 이용하여 동정한 결과, 상기 화학식 1의 화합물은 신규 나르보마이신 유도체임을 확인하였다.

[0035] 본 발명의 화합물은 용매화물 또는 전구약물(pro-drug) 형태일 수 있는데, 이는 본 발명의 범위 내에 포함된다. 용매화물은 바람직하게는 수화물 및 에탄올화물을 포함한다.

[0036] 본 발명에서의 용어 "약제학적으로 허용가능한 염"은 화합물의 상대적으로 무독성인 무기 및 유기 산 부가염일 수 있다. 상기 염으로는 약학적으로 허용가능한 유리산(free acid)에 의해 형성된 산 부가염이 유용하다. 환자에게 비교적 비독성이고 무해한 유효작용을 갖는 농도로서 이 염에 기인한 부작용이 화학식 1의 화합물의 이로운 효능을 저하시키지 않는 임의의 모든 유기 또는 무기 부가염을 포함한다.

[0037] 산 부가염은 통상의 방법, 예를 들어 화합물을 과량의 산 수용액에 용해시키고, 이 염을 수혼화성 유기 용매, 예를 들어 메탄올, 에탄올, 아세톤 또는 아세토니트릴을 사용하여 침전시켜서 제조한다. 동 몰량의 화합물 및 물 중의 산 또는 알코올(예, 글리콜 모노메틸에테르)을 가열하고, 이어서 상기 혼합물을 증발시켜 건조시키거나, 또는 석출된 염을 흡인 여과시킬 수 있다.

[0038] 이때, 유리산으로는 유기산과 무기산을 사용할 수 있으며, 무기산으로는 염산, 인산, 황산, 질산, 주석산 등을 사용할 수 있고 유기산으로는 메탄설폰산, p-톨루엔설폰산, 아세트산, 트라이플루오로아세트산, 말레인산(maleic acid), 석신산, 옥살산, 벤조산, 타르타르산, 푸마르산(fumaric acid), 만데르산, 프로피온산(propionic acid), 구연산(citric acid), 젖산(lactic acid), 글리콜산(glycollic acid), 글루콘산(gluconic acid), 갈락투론산, 글루탐산, 글루타르산(glutaric acid), 글루쿠론산(glucuronic acid), 아스파르트산, 아스코르브산, 카본산, 바닐릭산, 요오드화수소산(hydroiodic acid) 등을 사용할 수 있으며, 이들에 제한되지 않는다.

[0039] 또한, 염기를 사용하여 약제학적으로 허용가능한 금속염을 만들 수 있다. 알칼리 금속염 또는 알칼리 토금속염은, 예를 들어 화합물을 과량의 알칼리 금속 수산화물 또는 알칼리 토금속 수산화물 용액 중에 용해시키고, 비용해 화합물 염을 여과한 후 여액을 증발, 건조시켜 얻는다. 이때, 금속염으로는 특히 나트륨, 칼륨 또는 칼슘염을 제조하는 것이 제약상 적합하나 이들에 제한되는 것은 아니다. 또한, 이에 대응하는 은염은 알칼리 금속 또는 알칼리 토금속 염을 적당한 은염(예, 질산은)과 반응시켜 얻을 수 있다.

[0040] 화학식 1로 표시되는 화합물의 약제학적으로 허용가능한 염은, 달리 지시되지 않는 한, 상기 화합물에 존재할 수 있는 산성 또는 염기성 기의 염을 포함한다. 예를 들어, 약제학적으로 허용가능한 염으로는 하이드록시기의 나트륨, 칼슘 및 칼륨 염 등이 포함될 수 있고, 아미노기의 기타 약학적으로 허용가능한 염으로는 하이드로브롬화물, 황산염, 수소 황산염, 인산염, 수소 인산염, 이수소 인산염, 아세테이트, 석시네이트, 시트레이트, 타르트레이트, 락테이트, 만델레이트, 메탄설포네이트(메실레이트) 및 p-톨루엔설포네이트(토실레이트) 염 등이 있으며, 당업계에 알려진 염의 제조방법을 통하여 제조될 수 있다.

[0041] 본 발명의 화합물은 당업계에서 통상적으로 사용되는 방법에 따라 합성할 수 있으며, 바람직하게는 본 발명의 방법을 사용하여 재조합 미생물로부터 생산할 수 있다.

[0042] 또 다른 양태로서, 본 발명은 상기 화학식 1의 화합물 또는 약제학적으로 허용가능한 염을 포함하는 항균용 조성물을 제공한다. 본 발명에서 "항균용 조성물"이란, 미생물의 생존 및/또는 증식을 억제하는 작용을 하는 조성물을 의미한다. 상기 조성물은 에리트로마이신-감수성 균주 및 에리트로마이신-저항성 균주에 대하여 항균 활성을 가지는 것일 수 있고, 구체적으로는, *E. 파에시움* ATCC 19434, *E. 파에시움* P00558, *S. 아우레우스* ATCC 25923 및 *S. 아우레우스* P00740로 구성된 군으로부터 선택된 미생물에 대하여 항균 활성을 가질 수 있다. 또한,

상기 조성물은 상기 미생물에 의하여 발생하는 감염성 질환을 치료할 수 있다.

[0043] 상기 조성물은 약제학적으로 허용가능한 담체를 추가로 포함할 수 있다. 본 발명에서의 용어 "약제학적으로 허용가능한 담체"는 임의의 대상 조성물 또는 성분을 하나의 기관, 또는 신체의 부분으로부터 다른 기관, 또는 신체의 부분으로의 운반 또는 수송하는 것에 관여하는 액체 또는 고체 충전제, 희석제, 부형제, 용매 또는 캡슐화 물질과 같은 제약상 허용가능한 물질, 조성물 또는 비히클을 지칭하며, 본 발명의 조성물은, 투여를 위해서 상기 기재한 유효성분 이외에 약제학적으로 허용가능한 담체, 부형제 또는 희석제를 더 포함할 수 있다. 상기 담체, 부형제 및 희석제로는 락토오스, 덱스트로오스, 수크로오스, 소르비톨, 만니톨, 자일리톨, 에리스리톨, 말티톨, 전분, 아카시아 고무, 알지네이트, 젤라틴, 칼슘 포스페이트, 칼슘 실리케이트, 셀룰로오스, 메틸 셀룰로오스, 미정질 셀룰로오스, 폴리비닐 피롤리돈, 물, 메틸히드록시벤조에이트, 프로필히드록시벤조에이트, 탈크, 스테아린산 마그네슘 및 광물유를 들 수 있다.

[0044] 또한, 본 발명의 조성물은 각각 통상의 방법에 따라 산제, 과립제, 정제, 캡슐제, 현탁액, 에멀젼, 시럽, 에어로졸 등의 경구형 제형, 외용제, 좌제 또는 멸균 주사용액의 형태로 제형화하여 사용할 수 있다. 상세하게는, 제형화할 경우 통상 사용하는 충진제, 증량제, 결합제, 습윤제, 붕해제, 계면활성제 등의 희석제 또는 부형제를 사용하여 조제될 수 있다. 경구투여를 위한 고형제제로는 정제, 환제, 산제, 과립제, 캡슐제 등을 포함하나, 이에 한정되는 것은 아니다. 이러한 고형제제는 상기 화학식 1의 화합물 또는 그의 약제학적으로 허용가능한 염에 적어도 하나 이상의 부형제, 예를 들면, 전분, 칼슘 카보네이트, 수크로오스, 락토오스, 젤라틴 등을 섞어 조제될 수 있다. 또한, 단순한 부형제 이외에 스테아린산 마그네슘, 탈크 같은 윤활제들도 사용될 수 있다. 경구를 위한 액상 제제로는 현탁제, 내용 액제, 유제, 시럽제 등을 포함하나, 이에 한정되지 않으며, 흔히 사용되는 단순 희석제인 물, 리퀴드 파라핀 이외에 여러 가지 부형제, 예를 들면 습윤제, 감미제, 방향제, 보존제 등을 첨가하여 조제될 수 있다. 비경구 투여를 위한 제제는 멸균된 수용액, 비수성 용제, 현탁제, 유제, 동결건조 제제 및 좌제를 포함한다. 비수성 용제 및 현탁제로는 프로필렌글리콜, 폴리에틸렌 글리콜, 올리브 오일과 같은 식물성 오일, 에틸올레이트와 같은 주사가능한 에스테르 등이 사용될 수 있다. 좌제의 기제로는 위텝솔, 마크로골, 트윈 61, 카카오지, 라우린지, 글리세로젤라틴 등이 사용될 수 있다.

[0045] 또한, 본 발명은 상기 화학식 1의 화합물 또는 그의 약제학적으로 허용가능한 염을 대상에게 투여하는 단계를 포함하는 미생물 감염성 질환의 치료방법을 제공한다. 또한, 본 발명은 상기 화학식 1의 화합물 또는 그의 약제학적으로 허용가능한 염을 이용하여 에리트로마이신-감수성 균주, 에리트로마이신-저항성 균주를 포함하는 미생물의 생존 및/또는 증식을 억제하는 방법을 제공한다.

[0046] 본 발명에서의 용어 "치료"란, 질환의 발병을 억제하거나 발병을 지연하는 모든 행위를 말하며, 특히 약리학적 활성 물질에 의해 야기되는 동물, 특히 포유류, 및 더욱 특별히 인간에서의 국소적 또는 전신적인 효과를 가리킨다.

[0047] 본 발명의 구체적 실시예에서, DQNVNB는 *S. 아우레우스* P00740의 성장을 저해하는 데에 있어서는 나르보마이신보다 효과적이었다. LOLVNB는 *S. 아우레우스* ATCC 25923에 대하여 선택적 활성을 나타내었는 바, LOLVNB의 3'-가로방향(equatorial)-히드록실기가 항균 활성에 긍정적 영향을 미친다는 것을 의미한다. LRHMNB는 강력한 활성(1.25 - 2.50 μM)을 나타내었다. ODDCNB의 항균 활성(2.5 - 5.0 μM)은 나르보마이신 (10 - 20 μM) 및 DQNVNB (≤ 10 μM) 보다 우수하였다. DBVNNB (≤ 10 μM) 는 *E. 파에시움* ATCC 19434, P00558, 및 *S. 아우레우스* ATCC 25923에 대하여 높은 항균 활성을 나타내었는 바, 이는 DBVNNB의 4'-세로방향(axial)-히드록실기가 MIC를 감소시킨다는 것을 의미한다.

[0048] 본 발명의 조성물은 목적 조직에 도달할 수 있는 한 어떠한 일반적인 경로를 통하여 투여될 수 있다. 따라서, 본 발명의 조성물은 국부, 경구, 비경구, 비내, 정맥내, 근육내, 피하, 안내, 경피 등으로 투여될 수 있고, 용액, 현탁액, 정제, 환약, 캡슐, 서방형 제제 등으로 제형할 수 있다.

[0049] 본 발명의 조성물은 치료학적 유효량으로 투여될 수 있다. 투여량은 환자의 질병 종류 및 중증도, 연령, 성별, 투여방법, 표적 세포, 발현 정도 등 다양한 요인에 따라 달라질 수 있으며, 당 분야의 전문가들에 의해 용이하게 결정될 수 있다.

발명의 효과

[0050] 본 발명의 생산 방법을 통하여 *생체 내* 조합 생합성의 방법으로 나르보마이신에 있어서 다양한 당 부분을 포함하는 나르보마이신 유도체를 합성할 수 있고, 특히 신규하고 항균 활성이 우수한 다양한 나르보마이신 유도체를

합성할 수 있다.

도면의 간단한 설명

[0051] 도 1은 *스트렙토마이세스 베네주엘래*로부터 나르보마이신이 생합성되는 경로 및 나르보마이신의 구조를 나타내는 도면이다.

도 2는 본 발명의 방법에 의하여 변형된 당 부분을 포함하는 나르보마이신 유도체가 생합성되는 경로를 나타낸다. 상기 생합성에 관여하는 단백질의 기능을 괄호안에 표시하였다. NT, 뉴클레오티딜릴 트랜스퍼라아제; 4,6-DH, 4,6-데하이드라타아제; 2,3-DH, 2,3-데하이드라타아제; AT, 아미노트랜스퍼라아제; DA, 데아미나아제; EP, 에피머라아제; 4-KR, 4-케토리덕타아제; N-MT, *N*-메틸트랜스퍼라아제; 3-KR, 3-케토리덕타아제; SvRed, *스트렙토마이세스 베네주엘래* 경로 독립적인 리덕타아제.

도 3은 나르보마이신 및 그의 유도체의 구조를 나타낸다.

도 4는 *스트렙토마이세스 베네주엘래* 균주의 배양액으로부터의 HPLC-ESI-MS 크로마토그램을 나타낸다. (A) YJ003/pDDSS로부터 검출된 나르보마이신 ($[M + H]^+$ = 510). (B) YJ003/pDQNV로부터 검출된 DQNVNB ($[M + NH_4]^+$ = 516). (C) YJ003/pLRHM2로부터 검출된 LRHMNB ($[M + NH_4]^+$ = 516) 및 DQNVNB ($[M + NH_4]^+$ = 516). (D) YJ003/pODDC로부터 검출된 ODDCNB ($[M + NH_4]^+$ = 500). (E) YJ003/pDDGT2로부터 검출된 DBVNNB ($[M + NH_4]^+$ = 500) 및 DDGTNB ($[M + NH_4]^+$ = 500). (F) YJ003/pLOLV2로부터 검출된 DBVNNB ($[M + NH_4]^+$ = 500), LOLVNB ($[M + NH_4]^+$ = 500), 및 LDGTNB ($[M + NH_4]^+$ = 500).

도 5는 나르보마이신의 당 부분의 NOE(Nuclear Overhauser effect) 상관관계를 나타낸다.

도 6은 des 클러스터가 결실된 *스트렙토마이세스 베네주엘래* 결실 돌연변이체를 제조하기 위한 유전자 구조체의 구조를 나타내는 모식도이다.

도 7은 pDQNV(desVIII-desVII-desIII-desIV의 단백질 조합을 코딩하는 핵산을 포함하는 발현벡터)의 개열지도이다.

도 8은 pLRHM2(desVIII-desVII-desIII-desIV-oleL-oleU의 단백질 조합을 코딩하는 핵산을 포함하는 발현벡터)의 개열지도이다.

도 9는 pODDC(desVIII-desVII-desIII-desIV-desI-desII의 단백질 조합을 코딩하는 핵산을 포함하는 발현벡터)의 개열지도이다.

도 10은 pDDGT2(desVIII-desVII-desIII-desIV-oleV-eryBII-urdR의 단백질 조합을 코딩하는 핵산을 포함하는 발현벡터)의 개열지도이다.

도 11은 pLOLV2(desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU의 단백질 조합을 코딩하는 핵산을 포함하는 발현벡터)의 개열지도이다.

도 12는 pDDSS(desVIII-desVII-desIII-desIV-desI-desII-desV-desVI의 단백질 조합을 코딩하는 핵산을 포함하는 발현벡터)의 개열지도이다.

발명을 실시하기 위한 구체적인 내용

[0052] 이하 본 발명을 실시예를 통하여 보다 상세하게 설명한다. 그러나, 이들 실시예는 본 발명을 예시적으로 설명하기 위한 것으로 본 발명의 범위가 이들 실시예에 의해 제한되는 것은 아니다.

[0053] **박테리아 균주, 배양 조건, 및 유전자 조작**

[0054] *스트렙토마이세스 베네주엘래*의 원형질체 제조 및 형질전환은 Kieser, T. et al.의 방법에 따라 수행하였다 (Kieser, T. et al., 2000, Practical *Streptomyces* genetics. John Innes Centre, Norwich, United Kingdom). *스트렙토마이세스 베네주엘래*의 형질전환체는 티오스트렙톤으로 보충된 R2YE 아가 플레이트에서 선택

하였다. *스트렙토마이세스 베네주엘래* 재조합 균주는 SPA 아가(효모 추출물 1 g, 소고기 추출물 1 g, 트립토오스 2 g, 글루코오스 10 g, 미량의 황산철, 및 아가 15 g/리터) 상에서 증식시켰다. *스트렙토마이세스 베네주엘래* ATCC 15439, *스트렙토마이세스 안티비오티커스* ATCC 11891, 및 *사카로폴리스포라 에리트래아* NRRL 2338 를 이용하여 게놈 DNA를 획득하였다.

[0055] 대장균(*E. coli*) DH5α 및 플라스미드 Litmus28 (New England Biolabs)을 이용하여 일반적인 서브클로닝을 하였다. 구성 *ermE** 프로모터 (Schmitt-John, T. et al., 1992. Appl. Microbiol. Biotechnol. 36:493-498) 및 티오스트렙톤 저항성 마커를 포함하는 고효율(high-copy-number) 대장균-*스트렙토마이세스* 셔틀 벡터 pSE34 (Yoon, Y. J. et al., 2002. Chem. Biol. 9:203-214)를 이용하여 데옥시당 및 그들의 부착물의 생합성에 연관된 유전자를 발현시켰다.

[0056] **des 클러스터-결실 돌연변이체의 제조**

[0057] 복제 플라스미드-매개 상동 재조합(replicative plasmid-mediated homologous recombination) 방법(Xue et al., PNAS USA, 95:12111, 1998)을 통하여, 데소사민 생합성이 결실된 *스트렙토마이세스 베네주엘래* 결실 돌연변이체(YJ003)를 제조하였다.

[0058] 플라스미드는 카나마이신 내성 유전자, aphⅡ(Ward et al., Mol. Gen. Genet., 203:468, 1986)를 가지고 있으며, 전체 des 클러스터 대신에 상기 des 클러스터의 업스트림(upstream)과 다운스트림(downstream) 양측면에 위치한 DNA 단편 (flanking DNA fragment)을 가지도록 제조하였다(도 6). 유전자 교체를 위한 구조체(construct)는 원형질체 형질전환(protoplast transformation)을 통해 *스트렙토마이세스 베네주엘래*로 삽입된 pKC1139(Bierman, M. et al., Gene, 116:43-9, 1992)를 사용하여 제조하였다. *스트렙토마이세스 베네주엘래* ATCC 15439를 상기 유전자를 포함하는 플라스미드로 형질전환하였다. 각 형질전환체로부터 얻어진 포자는 카나마이신 선별 아가 플레이트에서 배양하고, 사이클(cycle)은 재조합 가능성을 증대시키기 위하여 3번 반복하였다. 카나마이신-저항성 및 아프라마이신-민감성 표현형을 얻기 위해 목표 유전자를 aphⅡ로 교체한 이중교차(crossover)를 선별 및 스크린하고, 돌연변이 유전자형을 게놈 DNA 서던 블럿 혼성화로 확인하였다.

[0059] des 결실 플라스미드(pYJ003)를 제조하기 위하여, SphⅠ으로 자른 *스트렙토마이세스 베네주엘래* 게놈 DNA를 주형(template)으로 하고, 사용한 5'-des 클러스터 측면에 위치한 부분을 가진 1kb HindⅢ-PstⅠ 단편은 서열번호 21 및 22를 프라이머로 사용하고, 3'-des 클러스터 측면에 위치한 부분을 가진 1kb KpnⅠ-EcoRⅠ 단편은 서열번호 23 및 24를 프라이머로 사용하여, PCR을 수행하였다. PCR에 사용된 반응성분은 주형 DNA 100ng, 각 프라이머 30pmol, pfu DNA polymerase 2.5unit, 각 dNTP 2mM, Tris-HCl 20mM, $(NH_4)_2SO_4$ 10mM, KCl 10mM, Triton X-100 1%, BSA 1mg/ml, $MgSO_4$ 20mM 및 DMSO 10%로, 이들을 섞어서 총 부피가 50㎕가 되도록 하였다.

[0060] 서열번호 21 (forward): 5'-GGCAAGCTTAGCGGGGCGACTGGCGTGCCCACT-3'

[0061] 서열번호 22 (reverse): 5'-GGTCTGCAGTCACCGTGGGTTCTGCCATCTCTT-3'

[0062] 서열번호 23 (forward): 5'-GCTGGTACCGGATGTTCCCTCCGGGCCACCGTC-3'

[0063] 서열번호 24 (reverse): 5'-TGAGAATTCCCTCGCCGTCCTGCCCGCGCTTGG-3'

[0064] **발현 플라스미드 및 *스트렙토마이세스 베네주엘래* 재조합 균주의 제조**

[0065] 복수의 데옥시당 생합성 유전자를 포함하는 DNA 절편을 특이적인 데옥시올리고뉴클레오티드 프라이머 및 주형 DNA를 이용하여 PCR로 증폭하였다 (표 1). PCR은 제조자의 추천 조건 하에서 *Pfu* 폴리머라아제 (Fermentas)를 이용하여 수행하였다.

표 1

[0066]

프라이머	서열(5'-3')	제한효소	DNA 기원 (GenBank accession No.)
desVIII-VII F	TTAATTAAGATATCACCGGCAAGGAAGGACACGACCGCC (서열번호 25)	PacI-EcoRV	*Streptomyces venezuelae* ATCC 15439 (GenBank AF079762.1)
desVIII-VII R	TCTAGACATATGGCGCAGATACAGGGGTGAGGCCTG (서열번호 26)	NdeI-XbaI	

desI-II F	TTAATTAAACTAGTATCGATGACGGTGGCCCGGAGGG (서열번호 27)	PacI-SpeI-ClaI	*Streptomyces venezuelae* ATCC 15439 (GenBank AF079762.1)
desI-II R	TCTAGATGCGGGTCAGCGCAGGAAGCCGCG (서열번호 28)	XbaI	
desIII-IV F	TTAATTAAACTAGTTAACTCGCCACGCCGACCGTT (서열번호 29)	PacI-SpeI	*Streptomyces venezuelae* ATCC 15439 (GenBank AF079762.1)
desIII-IV R	TCTAGAGAGCTCCTCGTAGGCGGCCTT (서열번호 30)	XbaI	
desV F	TTAATTAAACTAGTCAGGTCTCCTTCGCGGACGGCCTC (서열번호 31)	PacI-SpeI	*Streptomyces venezuelae* ATCC 15439 (GenBank AF079762.1)
desV R	TCTAGAACCTGACTAGGCCTGGTCGACCCG (서열번호 32)	XbaI	
desVI F	TTAATTAAACTAGTCCCCAGGCCTCACCCCTGTATCTG (서열번호 33)	PacI-SpeI	*Streptomyces venezuelae* ATCC 15439 (GenBank AF079762.1)
desVI R	TCTAGAGAATTCGCTCAGGCGGGGACGCCGACGAAG (서열번호 34)	XbaI-EcoRI	
oleL F	TTAATTAAACTAGTATCGATATCGCTCCGAGCCCGAAGGGA (서열번호 35)	PacI/ SpeI/ClaI	*Streptomyces antibioticus* ATCC 11891 (GenBank AF055579.2)
oleL R	TCTAGAGCCGGCCAGTACGAGGGCCTT (서열번호 36)	XbaI	
oleU F	TTAATTAAACTAGTGTACCGCGACAACCGC (서열번호 37)	PacI/ SpeI	*S. antibioticus* ATCC 11891 (GenBank AF055579.2)
oleU R	TCTAGAGAAGAGGGCCAGGTCGTGCACGC (서열번호 38)	XbaI	
oleV F	TTAATTAAACTAGTATCGATGGGAATCGCGGAAGCG (서열번호 39)	PacI/SpeI/ClaI	*S. antibioticus* ATCC 11891 (GenBank AF055579.2)
oleV R	TCTAGAATCAGCTCAGGGCCTGGATGC (서열번호 40)	XbaI	
eryBII F	TTAATTAAACTAGTGGTCGTCGGCATCTGCG (서열번호 41)	PacI/SpeI	*Saccharopolyspora erythraea* NRRL 2338 (GenBank U77454.1)
eryBII R	TCTAGACTCGGTTCCTCTTGTGCTCACTGC (서열번호 42)	XbaI	
urdR F	TTAATTAAACTAGTACAGAGATCCAGGACGACGCA (서열번호 43)	PacI/ SpeI	pYJ146
urdR R	TCTAGATCAGATACGGACGGCGGAGGT (서열번호 44)	XbaI	
oleW F	TTAATTAAACTAGTAAGGGAACCCCATGCCCTCCC (서열번호 45)	PacI/SpeI	*S. antibioticus* ATCC 11891 (GenBank AF055579.2)
oleW R	TCTAGAATCAGCACCAGCGCACCCGCGCCA (서열번호 46)	XbaI	

[0067] pYJ146는 Hong, J. S. et al., 2004. FEMS Microbiol. Lett. 238:291-299에 기재되어 있음.

[0068] 상기 다양한 데옥시당을 가지는 나르보놀라이드의 글리코실화에 연관되는 유전자들을 복제형 플라스미드 pSE34에 클로닝하였다. 다양한 유전자를 포함하는 플라스미드를 효율적으로 제작하기 위하여 하나의 유전자를 포함하는 PacI/XbaI 절편을 타 유전자를 포함하는 PacI/SpeI-절단되는 Litmus28에 연결하였다. 상기 절차는 나르보놀라이드가 글리코실화 나르보놀라이드로 전환되는데 필요한 모든 유전자들이 Litmus28 내에서 조합되고 PacI/XbaI로 절단되는 pSE34내로 이동될 때까지 회귀적으로 반복되었다. 조합된 유전자 배열을 표 2에 요약하였다.

표 2

[0069]

플라스미드	유전자 조합	생산물
pDDSS	*desIII-desIV-desI-desII-desV-desVI*	나르보마이신
pDQNV	*desIII-desIV*	DQNVNB
pLRHM2	*desIII-desIV-oleL-oleU*	DQNVNB LRHMNB
pODDC	*desIII-desIV-desI-desII*	ODDCNB

pDDGT2	*desIII-desIV-oleV-eryBII-urdR*	DDGTNB DBVNNB
pLOLV2	*desIII-desIV-oleV-oleW-oleL-oleU*	LOLVNB LDGTNB DBVNNB

[0070] pDDSS를 제작하기 위하여, DesVII/DesVIII 글리코실트랜스퍼라아제/보조 단백질 쌍 및 TDP-D-데소사민 생합성 효소를 코딩하는 유전자를 Litmus28에 순차적으로 클로닝하여 pDDSS_lit28를 제조하였다. pDDSS_lit28의 삽입 DNA를 PacI/XbaI 로 절단하고 pSE34에 연결하여 pDDSS를 제조하였다.

[0071] pDQNV를 제작하기 위하여, ClaI/EcoRI-절단되는 pDDSS_lit28를 Klenow 절편을 이용하여 둔단으로 만들고 연결하여 *desVIII-desVII-desIII-desIV*를 포함하는 PacI/XbaI 절편을 pSE34로 클로닝하였다.

[0072] 다양한 데옥시당의 생합성 및 그들의 전이를 명령하는 다수의 플라스미드(pLRHM2, pODDC, pDDGT2, 및 pLOLV2)를 제작하기 위하여, 다양한 당 유전자 카세트를 포함하는 ClaI/XbaI 절편을 독립적으로 ClaI/XbaI-절단하는 pDDSS_lit28로 클로닝하였다. 제조된 플라스미드의 PacI/XbaI 절편을 pSE34로 각각 클로닝하고, 해당 플라스미드를 *스트렙토마이세스 베네주엘래* 결실 돌연변이체(YJ003)에 도입하였다. 구체적으로, pDDSS(글리코실트랜스퍼라아제 유전자 *desVIII*/*desVII*와 함께 D-데소사민 [DDSS] 생합성 유전자 조합(*desIII-desIV-desI-desII-desV-desVI*)을 발현하는 플라스미드)를 도입하여 재조합 미생물 YJ003/pDDSS를 제조(나르보마이신 생산); pDQNV(*desVIII*/*desVII*와 함께 D-퀴노보오스 [DQNV] 생합성 유전자 조합(*desIII-desIV*)을 발현하는 플라스미드)를 도입하여 재조합 미생물 YJ003/pDQNV를 제조(D-퀴노보오실-나르보놀라이드(DQNVNB) 생산); pODDC(*desVIII*/*desVII*와 함께 3'-*O*-데메틸-D-찰코오스 [ODDC] 생합성 유전자 조합(*desIII-desIV-desI-desII*)을 발현하는 플라스미드)를 도입하여 재조합 미생물 YJ003/pODDC를 제조(3'-*O*-데메틸-D-찰코오실-나르보놀라이드(ODDCNB) 생산); pLRHM2(*desVIII*/*desVII*와 함께 L-람노오스 [LRHM] 생합성 유전자 조합(*desIII-desIV-oleL-oleU*)을 발현하는 플라스미드를 도입하여 재조합 미생물 YJ003/pLRHM2를 제조(L-람노오실-나르보놀라이드(LRHMNB) 및 DQNVNB 생산); pDDGT2(*desVIII*/*desVII*와 함께 D-디지톡소오스 [DDGT] 생합성 유전자 조합(*desIII-desIV-oleV-eryBII-urdR*)을 발현하는 플라스미드를 도입하여 재조합 미생물 YJ003/pDDGT2를 제조(D-디지톡소오실-나르보놀라이드(DDGTNB) 및 D-보이비노오실-나르보놀라이드(DBVNNB) 생산); pLOLV2(*desVIII*/*desVII*와 함께 L-올리보오스 [LOLV] 및 L-디지톡소오스 [LDGT] 생합성 유전자 조합(*desIII-desIV-oleV-oleW-oleL-oleU*)을 발현하는 플라스미드를 도입하여 재조합 미생물 YJ003/pLOLV2를 제조(L-올리보오실-나르보놀라이드(LOLVNB), L-디지톡소오실-나르보놀라이드(LDGTNB), 및 DBVNNB 생산)하였다.

[0073] 제조한 상기 재조합 미생물 중에서, pDDGT2(*desVIII*/*desVII*와 함께 D-디지톡소오스 [DDGT] 생합성 유전자 조합(*desIII-desIV-oleV-eryBII-urdR*)을 발현하는 플라스미드)가 도입된 재조합 미생물 및 pLOLV2(*desVIII*/*desVII*와 함께 L-올리보오스 [LOLV] 및 L-디지톡소오스 [LDGT] 생합성 유전자 조합(*desIII-desIV-oleV-oleW-oleL-oleU*)을 발현하는 플라스미드)가 도입된 재조합 미생물을 한국미생물보존센터(Korean Culture Center of Microorganisms, KCCM)에 2011년 3월 10일자로 각각 기탁번호 KFCC11507P, KFCC11508P로 기탁하였다.

[0074] **HPLC-ESI-MS 분석 of 나르보마이신 및 그의 글리코실화 유도체**

[0075] *스트렙토마이세스 베네주엘래* 배양물에서 생산된 마크롤라이드를 SPE를 이용하여 추출하여 대사물질을 고성능액체크로마토그래피-전자분무이온화-질량분석기 (high-performance liquid chromatography-electrospray ionization-mass spectrometry, HPLC-ESI-MS)로 분석하였다. 각 배양물을 원심분리하였다(10분동안, 5,000 x g). 5 ml 메탄올 및 5 ml 물로 미리 처리된 OASIS HLB (Waters, Milford, MA) SPE 칼럼으로 상등액을 통과시켰다. 칼럼을 5 ml 20% (vol/vol) 메탄올로 세척하고 *ca.* 30 초동안 공기건조시켰다. 칼럼을 1 ml 0.5% (vol/vol) 메탄올 아세트산으로 3회 용출하였다. 분석 HPLC-ESI-MS를 양이온 모드에서 Waters Nova-Pak C18 칼럼 (150 x 3.9 mm, 5㎛)을 사용하여 Waters/Micromass Quattro micro/MS 인터페이스에서 수행하였다. 분석대상물질을 250 ㎕/분 유속에서 5 mM (wt/vol) 암모늄 아세테이트의 농도 구배를 이용하여 -0.05% (vol/vol) 아세트산 수용액(용액 A) 및 동일한 추가 농도를 가지는 80% (vol/vol) 아세토니트릴 (ACN) (용액 B) 으로서 25분 동안 20 내지 70% 용액 B, 15분 동안 용액 B 90%까지, 9분 동안 90% 용액 B 유지, 및 11분 동안 칼럼 재-평형을 위하여 용액 B 20%까지 - 용출하였다. 다양한 발현 플라스미드를 함유하는 6개의 *스트렙토마이세스 베네주엘래* 재조합 미생물의 SPE 추출물의 HPLC-ESI-MS 분석 결과, 나르보마이신 및 나르보마이신의 7개의 글리코실화 유도

체를 검출하였다(도 4).

[0076] **나르보마이신 및 그의 글리코실화 유도체의 추출, 분리 및 확인**

[0077] 크로마토그래피 분리를 제조용 Spherisorb S5 ODS2 (Waters, 250 x 20 mm, 5 μm) 또는 반-제조용 Watchers 120 ODS-BP (250 x 10 mm, 5 μm, Daiso, Osaka, Japan) 칼럼 상에서 역상 HPLC를 이용하여 수행하였다. ^{1}H, 상관 분광법 (COSY), 이종핵간 다중 결합 상관 (HMBC), 이종핵 단일 양자 상관 (HSQC), 및 핵 오버하우저 효과 분광법 (NOESY) NMR 스펙트럼을 Varian INOVA 500 분광계 (Varian, Inc., Palo Alto, CA)를 이용하여 500 MHz에서 기록하였다. ^{13}C NMR 스펙트럼을 동일한 기계 상에서 125 MHz 에서 기록하였다. 각각의 용매 피크와 관련하여 화학적 시프트가 보고되었다 ($CDCl_3$에 대하여 δ_H 7.27 및 δc 77.0; CD_3OD에 대하여 δ_H 3.30 및 δc 49.0).

[0078] 나르보마이신-생산 균주 *스트렙토마이세스 베네주엘래* YJ003/pDDSS의 전체 배양액 (0.5 L)을 원심 분리하고, 상등액 층을 에틸 아세테이트(Et_2OAc)를 이용하여 용매-용매 분배되도록 하였다. 획득된 추출물을 증발시킨 결과 생성된 갈색 잔여물을 구배 용출을 이용하는 제조용 HPLC를 이용하여 분별하였다. 분석대상물질을 0.1% 수성 아세트산 (용액 C) 및 아세토니트릴-아세트산 (99.9:0.1, v/v) (용액 D)의 농도 구배를 이용하여 0부터 15 분까지 용액 D 15%까지, 53 분에서 용액 D 35%까지, 58 분에서 용액 D 60%까지, 63 - 73 분 동안 다시 15% 용액 D로 5 ml/분의 유속으로 용출하였다. 나르보마이신 포함 분획을 이동상으로서 3 ml/분 30% 수성 CAN을 이용하는 반-제조용 HPLC를 사용하여 더욱 정제하여 비정질의 흰색 고체인 순수한 화합물을 생산하였다(0.67 mg; 머무름 시간 (Rt), 37.6 분; 화학식; ESI-MS $[M + H]^+$ 510; NMR 데이타는 하기 표 3).

표 3

Position	δ_C	δ_H (m, *J* in Hz)
1	169.9	
2	50.9	3.85 (q, 7)
3	207.9	
4	48.9	2.85 (m)
5	79.0	4.18 (brs)
6	36.6	1.74 (m)
7a	37.0	1.45 (m)
7b		1.18 (m)
8	42.2	2.76 (m)
9	203.8	
10	128.5	6.12 (d, 16)
11	147.2	6.67 (dd, 15.5, 6)
12	38.6	2.68 (m)
13	79.0	4.92 (m)
14	22.9	1.60 (m)
		1.53 (m)
15	10.7	0.90 (t, 7)
16	14.5	1.38 (d, 7)
17	13.9	1.36 (d, 7)
18	17.2	1.02 (d, 6.5)
19	15.7	1.12 (d, 7)
20	12.8	1.09 (d, 7)
1′	104.6	4.33 (d, 7)
2′	70.3	3.31 (dd, 7, 2.5)
3′	65.8	2.94 (m)
4′a	31.2	1.60 (m)
4′b		1.53 9m)
5′	69.4	3.58 (m)
6′	21.3	1.25 (d, 6)
7′	40.2	2.38 (s)

[0079]

[0080] D-퀴노보오실-나르보놀라이드(DQNVNB)-생산 균주 YJ003/pDQNV의 전체 배양액 (1.0 L)을 유사하게 처리하여 갈색의 물질을 획득하였다. 상기 잔여물을 구배 용출 (30 분까지 30% 수성 ACN, 40 - 80 분 동안 50% 수성 ACN, 및 81 - 90 분 동안 다시 30% 수성 ACN)을 이용하는 반-제조용 HPLC를 이용하여 2 ml/분의 유속으로 분별하였다. DQNVNB 포함 분획을 이동상으로서 2 ml/분 30% 수성 ACN을 이용하는 반-제조용 HPLC를 사용하여 더욱 정제하여 비정질의 흰색 고체인 순수한 화합물을 생산하였다(4.25 mg; 머무름 시간(Rt), 25.8 분; 화학식; ESI-MS $[M + NH_4]^+$ 516; NMR 데이타는 하기 표 4).

표 4

[0081]

Position	δ_C	δ_H (m, *J* in Hz)	COSY	HMBC
1	169.7			
2	50.7	3.82 (q, 7.0)	H-16	C-1, C-3, C-16
3	209.0			
4	48.7	2.89 (m, 7.5, 5)	H-5, H-17	C-3, C-5, C-6, C-17
5	78.8	4.16 (brs)	H-4	
6	36.5	1.80 (m)		
7a	37.3	1.48 (m)		C-9, C-5, C-8, C-6
7b		1.19 (m)		C-9, C-5, C-8, C-6
8	41.5	2.81 (m, 7)	H-19	C-9, C-7, C-19
9	204.0			
10	128.8	6.05 (d, 16)	H-11	C-9, C-11, C-13, C-12
11	147.5	6.65 (dd, 16, 6)	H-10, H-12	C-9, C-10, C-13, C-12, C-20
12	38.6	2.72 (m)	H-13, H-20	C-11, C-10, C-13, C-20
13	79.0	4.94 (dt, 9.5, 4)	H-12, H-14a, H-14b	C-1, C-11
14a	23.0	1.61 (m)	H-13	C-15
14b		1.52 (m)	H-13	C-15, C-13
15	10.7	0.89 (t, 7.5)		C-13, C-14
16	14.4	1.35 (d, 7.5)	H-2	C-1, C-2, C-3
17	13.5	1.30 (d, 6.5)	H-4	C-3, C-4, C-5
18	17.2	0.98 (d, 6)		C-5, C-6, C-7
19	15.6	1.08 (d, 6.5)	H-8	C-8, C-7, C-9
20	12.6	1.10 (d, 6.5)	H-12	C-13, C-12, C-11
1′	103.1	4.34 (d, 7)	H-2′	C-5
2′	74.7	3.33 (m, 9, 6)	H-1′, H-3′	C-1′, C-3′
3′	76.9	3.46 (bt, 8.1)	H-2′, H-4′	
4′	75.5	3.16 (bt, 8.5)	H-3′, H-5′	C-3′
5′	72.0	3.34 (m, 9, 6)	H-4′, H-6′	C-3′
6′	17.8	1.30 (d, 6.5)	H-5′	C-5′, C-4′

[0082] L-람노오실-나르보놀라이드(LRHMNB)-생산 균주 YJ003/pLRHM2의 전체 배양액 (0.5 L)으로부터 획득한 잔여물을 구배 용출 (55 분까지 35% 수성 ACN, 55 - 65 분 동안 60% 수성 ACN, 및 70 - 80 분 동안 다시 35% 수성 ACN)을 이용하는 제조용 HPLC 를 이용하여 6 ml/분의 유속으로 분별하였다. LRHMNB 포함 분획을 이동상으로서 30% 수성 ACN을 이용하는 반-제조용 HPLC를 사용하여 정제하여 순수한 화합물을 생산하였다(0.54 mg; 머무름 시간 (Rt), 25.2 분; 화학식; ESI-MS $[M + NH_4]^+$ 516; NMR 데이타는 하기 표 5).

표 5

[0083]

Position	δ_C	δ_H (m, *J* in Hz)	COSY	HMBC
1	169.6			
2	50.3	3.81 (q, 7)	H-16	C-1, C-3, C-16
3	208.4[a]			
4	48.4	2.86 (m)	H-5, H-17	
5	78.5[a]	4.33 (brs)	H-4	
6	36.5	1.80 (m)	H-7a, H-18	
7a	37.3	1.46 (m)	H-6, H-8, H-7b	
7b		1.19 (m)	H-7a	
8	41.7	2.83 (m)	H-7a, H-19	
9	204.0[a]			
10	129.1[a]	6.04 (d, 16)	H-11	C-9, C-12
11	147.4	6.62 (dd, 15.5, 5.5)	H-10, H-12	
12	38.6	2.75 (m)	H-11, H-13, H-20	
13	79.2	4.90 (brs)	H-12, H-14a, H-14b	
14a	22.7	1.60 (m)	H-13, H-15	
14b		1.50 (m)	H-13, H-15	
15	10.7	0.91 (t, 7)	H-14a, H-14b	C-13, C-14
16	14.3	1.35 (d, 6.5)	H-2	C-1, C-2, C-3
17	13.5	1.25 (d, 7)	H-4	C-3, C-4
18	17.1	1.00 (d,7)	H-6	C-5, C-6, C-7
19	15.3	1.13 (d, 7)	H-8	C-8, C-7, C-9
20	13.0	1.10 (d, 6.5)	H-12	C-13, C-12, C-11
1′	101.5	4.90 (brs)	H-2′	C-3′, C-5′, C-5
2′	71.1	3.94 (brs)	H-1′, H-3′	
3′	71.9	3.73 (m, 9)	H-2′, H-4′	
4′	73.5	3.50 (m)	H-5′, H-3′	
5′	68.8	3.73 (m, 9, 7)	H-4′, H-6′	
6′	17.6	1.30 (d, 6.5)	H-5′	C-5′, C-4′

[0084] a: HSQC 및 HMBC 스펙트럼에 기초한 수치

[0085] 3'-*O*-데메틸-D-찰코오실-나르보놀라이드(ODDCNB)-생산 균주 YJ003/pODDC의 전체 배양액 (1.0 L)을 동일한 방식으로 처리하였다. 잔여물을 이동상으로서 45% 수성 ACN을 5 ml/분의 유속으로 이용하는 제조용 HPLC 를 이용하여 분별하였다. ODDCNB 포함 분획을 이동상으로서 30% 수성 ACN을 2 ml/분의 유속으로 이용하는 반-제조용 HPLC를 사용하여 정제하여 순수한 화합물을 생산하였다(2.44 mg; 머무름 시간(Rt), 29.6 분; 화학식; ESI-MS $[M + NH_4]^+$ 500; NMR 데이타는 하기 표 6).

표 6

Position	δ_C	δ_H (m, *J* in Hz)	COSY	HMBC
1	169.6			
2	50.5	3.83 (q, 7)	H-16	C-1, C-3, C-16
3	209.1			
4	48.3	2.93 (m)	H-5, H-17	C-3, C-17
5	78.4	4.23 (brs)	H-4	C-18
6	36.4	1.90 (overlapped)	H-7a, H-18	
7a	37.6	1.47 (m)	H-6, H-8	C-9
7b		1.19 (m)		C-9
8	41.2	2.84 (m)	H-7a, H-19	C-7, C-9, C-19
9	203.8			
10	129.2	6.04 (d, 16)	H-11	C-9, C-12
11	147.2	6.64 (dd, 16, 6)	H-10, H-12	C-9, C-12, C-13, C-20
12	38.5	2.74 (m)	H-11, H-13, H-20	C-10, C-11, C-13, C-20
13	79.1	4.94 (dt, 9.5, 3.5)	H-12, H-14a, H-14b	
14a	22.8	1.59 (m)	H-13, H-15	
14b		1.50 (m)	H-13, H-15	C-13
15	10.7	0.91 (t, 7.5)	H-14a, H-14b	C-13
16	14.3	1.37 (d, 7)	H-2	C-1, C-2, C-3
17	13.8	1.32 (d, 7.5)	H-4	C-3, C-4, C-5
18	17.1	1.01 (d, 7)	H-6	C-5, C-6, C-7
19	15.4	1.10 (d, 7.5)	H-8	C-8, C-7, C-9
20	12.9	1.11 (d, 7)	H-12	C-13, C-12, C-11
1′	103.3	4.31 (d, 7.5)	H-2′	C-5, C-3′, C-5′
2′	76.9	3.25 (t, 8.5)	H-1′, H-3′	C-1′, C-3′
3′	71.4	3.70 (ddd, 14.5, 9, 5)	H-2′, H-4′a, H-4′b	C-2′
4′a	39.9	1.95 (ddd, 13, 5, 1.5)	H-5′, H-3′	C-2′, C-3′
4′b		1.42 (m)	H-3′	C-3′, C-5′
5′	68.4	3.62 (ddd, 11.5, 6.5, 1.5)	H-4′, H-6′	C-6′, C-1′, C-3′
6′	21.1	1.25 (d, 6)	H-5′	C-5′, C-4′

[0086]

[0087] D-디지톡소오실-나르보놀라이드(DDGTNB) 및 D-보이비노오실-나르보놀라이드 (DBVNNB)-생산 균주 YJ003/pDDGT2의 전체 배양액 (14.6 L)을 유사하게 처리하고 최종 잔여물을 구배 용출 (45 분까지 40% 수성 ACN, 50 - 60 분 동안 45% 수성 ACN, 및 65 - 75 분 동안 다시 40% 수성 ACN)을 이용하는 제조용 HPLC 를 이용하여 6 ml/분의 유속으로 분별하였다. DBVNNB 및 DDGTNB 포함 분획을 35% 수성 ACN을 2 ml/분의 유속으로 이용하는 반-제조용 HPLC 칼럼에 각각 통과시켜 순수한 DBVNNB (9.16 mg; 머무름 시간(Rt), 29.3 분; 화학식; ESI-MS $[M + NH_4]^+$ 500; NMR 데이타는 하기 표 7) 및 DDGTNB (4.18 mg; 머무름 시간(Rt), 31.2 분; 화학식; ESI-MS $[M + NH_4]^+$ 500; NMR 데이타는 하기 표 8)를 각각 생산하였다.

표 7

Position	δ_C	δ_H (m, *J* in Hz)	COSY	HMBC
1	169.6			
2	50.9	3.80 (q, 7.0)	H-16	C-1, C-3, C-16
3	208.4			
4	48.9	2.84 (quin, 7)	H-5, H-17	C-3, C-5, C-6, C-17
5	79.0	4.17 (d like, 4)	H-4, H-6	C-18
6	36.6	1.80 (m)	H-5, H-7b, H-18	
7a	37.0	1.44 (m)	H-8	C-9, C-8, C-6, C-18, C-19
7b		1.19 (m)	H-6	C-9, C-5
8	42.2	2.78 (m, 6.5, 7)	H-19	C-9, C-7, C-19
9	203.8			
10	128.5[a]	6.11 (d, 16)	H-11	C-9, C-12
11	147.2	6.68 (dd, 15.5, 6)	H-10, H-12	C-9, C-12, C-20
12	38.6	2.73 (m, 6.5, 2.5)	H-13, H-20	C-11
13	79.0	4.92 (dt, 9.5, 4)	H-12, H-14a, H-14b	
14a	22.9	1.59 (m)	H-13, H-15	C-15, C-13
14b		1.52 (m)	H-13, H-15	C-15, C-13
15	10.7	0.90 (t, 7)	H-14a, H-14b	C-13, C-14
16	14.5	1.36 (d, 7)	H-2	C-1, C-2, C-3
17	13.9	1.22 (d, 7.5)	H-4	C-3, C-4, C-5
18	17.2	1.00 (d, 7)	H-6	C-5, C-6, C-7
19	15.7	1.12 (d, 7)	H-8	C-8, C-7, C-9
20	12.8	1.10 (d, 7)	H-12	C-13, C-12, C-11
1′	99.7	4.81 (dd, 8, 4.5)	H-2′a, H-2′b	C-5, C-2′
2′	33.8	1.78 (dd, 8, 3)	H-1′, H-3′	C-1′, C-3′
		1.80 (m)	H-1′, H-3′	C-1′, C-3′
3′	69.6	4.11 (m, 3)	H-2′a, H-2′b, H-4′	C-1′
4′	70.7	3.26 (brs)	H-3′, H-5′	C-5′, C-3′
5′	68.8	4.05 (m, 6.5)	H-4′, H-6′	C-4′, C-1′, C-6′
6′	16.6	1.24 (d, 6.5)	H-5′	C-5′, C-4′

[0088]

표 8

Position	δ_C	δ_H (m, *J* in Hz)	COSY	HMBC
1	169.5			
2	51.0	3.80 (q, 7.0)	H-16	C-1, C-3, C-16
3	208.5[a]			
4	49.0	2.84 (quin, 7.5)	H-5, H-17	C-3, C-17, C-5, C-6
5	79.5	4.15 (brs)	H-4	C-18
6	36.5	1.75 (m)	H-18	
7a	37.0[a]	1.44 (m)	H-8	C-6, C-8, C-9, C-18, C-19
7b		1.19 (m)	H-8	C-6, C-8, C-18, C-19
8	42.5	2.75 (m)	H-7a, H-7b, H-19	C-7, C-9, C-19
9	203.0			
10	129.0[a]	6.12 (d, 16)	H-11	C-9, C-12
11	147.0[a]	6.68 (dd, 16, 6.5)	H-10, H-12	C-9, C-12, C-20
12	38.6	2.73 (m)	H-11, H-20	C-11, C-13
13	78.9	4.92 (dt, 9.5, 5.5)	H-14a	
14a	23.0[a]	1.60 (m)	H-15, H-13	C-15
14b		1.54 (m)	H-15	C-13, C-15
15	10.7	0.90 (t, 7.5)	H-14a, H-14b	C-13, C-14
16	14.6	1.37 (d, 7)	H-2	C-1, C-2, C-3
17	14.0	1.23 (d, 7.5)	H-4	C-3, C-4
18	17.2	1.01 (d, 7)	H-6	C-5, C-6, C-7
19	15.8	1.13 (d, 7)	H-8	C-8, C-7, C-9
20	13.0	1.10 (d, 7)	H-12	C-13, C-12, C-11
1′	98.9	4.84 (d, 9.5)	H-2′b	C-5, C-2′
2′a	37.9	2.10 (dt, 13.5, 2)	H-2′b, H-3′	C-3′, C-4′
2′b		1.70 (ddd, 13.7, 9.9, 2.2)	H-1′, 2′a, H-3′	C-1′
3′	68.5	4.11 (m, 3)	H-2′a, H-4′, H-2′b	C-1′, C-5′
4′	73.2	3.32 (dd, 9.5, 3)	H-3′, H-5′	C-6′
5′	69.5	3.72 (m, 9.5, 6.5)	H-4′, H-6′	C-1′, C-3′, C-4′, C-6′
6′	18.2	1.29 (d, 6.5)	H-5′	C-5′, C-4′

[0089]

[0090] a: HSQC 및 HMBC 스펙트럼에 기초한 수치

[0091] L-올리보오실-나르보놀라이드(LOLVNB) 및 L-디지톡소오실-나르보놀라이드 (LDGTNB)-생산 균주 YJ003/pLOLV2의 전체 배양액 (2.5 L)으로부터 획득한 잔여물을 구배 용출 (30 분까지 30% 수성 ACN, 40 - 80 분 동안 40% 수성 ACN, 및 81 - 90 분 동안 다시 30% 수성 ACN)을 이용하는 반-제조용 HPLC 를 이용하여 2 ml/분의 유속으로 분별하였다. LOLVNB 포함 분획을 이동상으로서 35% 수성 ACN을 2 ml/분의 유속으로 이용하는 반-제조용 HPLC 칼럼에 통과시켜 순수한 화합물(0.55 mg; 머무름 시간(Rt), 28.7 분; 화학식; ESI-MS $[M + NH_4]^+$ 500; NMR 데이타는 하기 표 9)을 생산하였다.

표 9

Position	δ_C	δ_H (m, *J* in Hz)	COSY	HMBC
1	169.6			
2	50.4	3.82 (q, 7)	H-16	C-1, C-3, C-16
3	208.5[a]			
4	48.2[a]	2.85 (m)	H-5, H-17	C-3, C-6, C-17
5	77.5 [a]	4.27 (brs)	H-4	
6	36.5	1.80 (m)	H-18	
7a	37.5[a]	1.43 (m)	H-8	C-6, C-9
7b		1.12 (m)		
8	42.0[a]	2.82 (m)	H-19, H-7a	C-9
9	203.5			
10	129.0 [a]	6.06 (d, 16)	H-11	C-9, C-12
11	147.0	6.63 (m)	H-10, H-12	C-9
12	38.6	2.74 (m)	H-11, H-20	
13	79.2	4.90 (m)	H-14a, H-14b	
14a	22.5	1.60 (m)	H-13, H-15	C-15
14b		1.50 (m)	H-13, H-15	
15	10.7	0.90 (t, 7)	H-14a, H-14b	C-13, C-14
16	14.4	1.37 (d, 7)	H-2	C-1, C-2, C-3
17	13.9	1.24 (d, 7)	H-4	C-3, C-4
18	17.2	1.00 (d, 7)	H-6	C-5, C-6
19	15.2[a]	1.13 (d, 6.5)	H-8	C-8, C-7, C-9
20	13.0[a]	1.10 (d, 7)	H-12	C-12, C-11, C-13
1′	98.9	4.97 (d, 3)	H-2′a	C-3′, C-5′
2′a	38.1	1.70 (td, 12.5, 4)	H-3′, H-2b′, H-1′	C-3′, C-4′, C-1′
2′b		2.12 (dd, 13, 5)	H-3′, H-2a′	C-3′, C-4′, C-1′
3′	69.4	3.90 (ddd, 12, 9, 5)	H-2′a, H-2′b, H-4′	C-4′, C-5′
4′	78.5	3.13 (t, 9)	H-3′, H-5′	C-3′, C-5′
5′	6 8.5	3.73 (m, 9, 6)	H-4′, H-6′	C-4′
6′	17.8	1.29 (d, 6)	H-5′	C-5′, C-4′

[0092]

[0093] a: HSQC 및 HMBC 스펙트럼에 기초한 수치

[0094] LOLVNB 및 LDGTNB-생산 균주 YJ003/pLOLV2의 전체 배양액 (21.6 L)을 이전의균주와 동일하게 처리하여 생산된 잔여물을 구배 용출 (50 분까지 38% 수성 ACN, 55 - 75 분 동안 42% 수성 ACN, 및 76 - 85 분 동안 다시 38% 수성 ACN)을 이용하는 제조용 HPLC 를 이용하여 6 ml/분의 유속으로 분별하였다. LDGTNB 포함 분획을 이동상으로서 35% 수성 ACN을 2 ml/분의 유속으로 이용하는 반-제조용 HPLC 칼럼에 통과시켜 순수한 화합물(1.81 mg; 머무름 시간(Rt), 31.0 분; 화학식; ESI-MS $[M + NH_4]^+$ 500; NMR 데이타는 하기 표 10)을 생산하였다.

표 10

[0095]

Position	δ_C	δ_H (m, *J* in Hz)	COSY	HMBC
1	169.5			
2	50.1	3.80 (q, 7.0)	H-16	C-1, C-3, C-16
3	208.1			
4	48.5[a]	2.85 (m)	H-5, H-17	C-3, H-6, C-17
5	78.5	4.36 (brs)	H-4	
6	36.6	1.82 (m)	H-7a, H-18	
7a	37.2[a]	1.48 (m)	H-6	C-9, C-6, C-18, C-19
7b		1.10 (m)		C-8, C-6, C-18, C-19
8	42.0[a]	2.85 (m)	H-7b, H-7a, H-19	C-9, C-7
9	203.2			
10	129.0	6.02 (d, 16.5)	H-11	C-9, C-12
11	147.4	6.63 (dd, 15.5, 5.5)	H-10, H-12	C-9
12	38.6	2.74 (m)	H-11, H-13, H-20	C-13, C-20
13	79.5	4.91 (dt, 10.5, 3.5)	H-12, H-14b, H-14a	
14a	22.5[a]	1.60 (m)	H-13, H-15, H-14b	
14b		1.48 (m)	H-14a, H-13, H-15	C-13
15	10.7	0.90 (t, 7.5)	H-14a, H-14b	C-13, C-14
16	14.2	1.36 (d, 7.0)	H-2	C-1, C-2, C-3
17	13.7	1.26 (d, 7.0)	H-4	C-3, C-4
18	17.3	1.02 (d, 7.0)	H-6	C-5, C-6, C-7
19	15.5	1.13 (d, 6.5)	H-8	C-8, C-7, C-9
20	13.0[a]	1.09 (d, 7.5)	H-12	C-13, C-12, C-11
1′	99.1	5.03 (brs, 3)	H-2′a, 2′b	C-3′, C-5′
2′a	35.4	2.17 (dd, 14.5, 3.0)	H-1′, H-2′b, H-3′	C-1′, C-3′, C-4′
2′b		1.91 (dt, 15.0, 3.5)	H-1′, H-2′a, H-3′	
3′	67.5	3.96 (brs)	H-2′a, H-2′b, H-4′	C-5′
4′	72.9	3.15 (dd, 10, 3)	H-3′, H-5′	C-5′, C-6′
5′	65.5	3.83 (overlapped)	H-4′, H-6′	C-3′, C-4′, C-6′
6′	17.9	1.33 (d, 6.5)	H-5′	C-4′, C-5′

[0096] a: HSQC 및 HMBC 스펙트럼에 기초한 수치

[0097] **나르보마이신 및 그의 유도체 구조 규명**

[0098] 분리된 대사물질들의 구조를 NMR 분광법 ($_1$H, $_{13}$C, COSY, HSQC, 및 HMBC)을 이용하여 추론하였다. DQNVNB의 ^{13}C NMR 데이타를 통하여 나르보마이신과 같은 14-원 마크롤라이드 글리코시드에서 확인할 수 있는 26 탄소의 존재를 확인하였다. COSY, HSQC, 및 HMBC 스펙트럼을 포함하는 2D NMR 데이타의 해석 결과, DQNVNB의 전체 구조를 구상할 수 있었다. ^{1}H NMR 스펙트럼을 통하여 7개의 메틸기 (δ_H 1.35/ H_3-16, 1.30/ H_3-6', 1.30/H_3-17, 1.10/H_3-19, 1.08/H_3-20, 0.98/H_3-18, 및 0.89/H_3-15), 2개의 메틸렌기 (δ_H 1.48, 1.19/H_2-7 및 1.61, 1.52/H_2-14), 5개의 메타인(methine) 양성자 (δ_H 3.82/H-2, 2.89/H-4, 2.81/H-8, 2.72/H-12, 및 1.80/H-6), 7개의 산소 공급된 메타인 양성자 (δ_H 4.94/H-13, 4.34/H-1', 4.16/H-5, 3.46/H-3', 3.34/H-5', 3.33/H-2', 및 3.16/H-4'), 및 2개의 올레핀 양성자 (δ_H 6.65/H-11 및 6.05/H-10)에 대한 신호를 확인하였다. COSY 스펙트럼을 통하여 아노머 양성자 (δ_H 4.34/H-1') 내지 H-6' (δ_H 1.30)로부터의 6-데옥시헥소오스에 대한 것을 포함하는 5개의 스핀 시스템을 확인하였다. 6-데옥시헥소오스 당의 존재는 ^{13}C NMR 및 HSQC 스펙트럼을 통하여 δc 103.1/C-1'에서의 아노머 탄소와 함께 4개의 산소 공급된 탄소 (δc 76.9/C-3', 75.5/C-4', 74.7/C-2', 72.0/C-5') 및 δc 17.8/C-6' 에서의 메틸 신호를 검출함으로써 재차 확인하였다. HMBC 스펙트럼에서, 아노머 양성자 (δ_H 4.34/H-1')는 탄소 C-5 (δc 78.8)와 연관되어 있었고, 이는 당이 C-5에 부착된 것을 의미한다. 더욱이, 각각의 탄소와 메틸 양성자의 주요 HMBC 상관관계는 그들의 위치를 확인시켜 주었다.

[0099] 글리코실화의 메커니즘이 알려진 것이라면(입체화학이 반전 또는 보존), 당의 입체배치는 아노머 양성자의 커플링 상수로부터 결정될 수 있다. 글리코실트랜스퍼라아제 DesVII는 활성화된 당의 아노머 탄소에서의 입체화학을

반전시킴으로써 당을 전이시킨다. 따라서, 아노머 양성자에 대한 큰 커플링 상수(*ca.* 7 Hz)를 가지는 글리코시드는 $^{4}C_{1}$ 입체구조에서 가로방향의(equatorial) β-글리코시드 연결된 D-당을 포함하는 것을 의미하고, 작은 커플링 상수 (< 5 Hz)는 세로방향의(axial) α-글리코시드 결합을 가지는 $^{1}C_{4}$ 입체구조를 가지는 L-당을 의미한다. 올바른 입체구조 ($^{4}C_{1}$)는 H-1' 및 H-5' 사이의 NOE 크로스 피크를 통하여 재차 확인될 수 있다. 그 결과, 당이 $^{4}C_{1}$ 입체구조를 가지는 D-퀴노보오스로 확인되었다. 아노머 양성자의 큰 커플링 상수 ($J_{1',2'}$ = 7.0 Hz)를 통하여 세로 방향 위치를 확인하였고, 따라서, Klyne의 법칙에 따라 β-글리코시드 연결된 D-당으로 확인되었다. 큰 커플링 상수($J_{2',3'}$ = 9.0 Hz, $J_{3',4'}$ = 9.0 Hz, 및 $J_{4',5'}$ = 9.0 Hz)는 H-2', H-3', H-4', 및 H-5' 양성자가 세로 방향으로 위치한다는 것을 의미한다. 더욱이, H-1' 및 H-5' 사이의 NOE 상관관계(도 5)가 $^{4}C_{1}$ 입체구조를 뒷받침한다.

[0100] ODDCNB의 1D 및 2D NMR 데이타는 DQNVNB와 유사하였다. 이는 상기 물질 역시 나르보마이신 유도체라는 것을 의미한다. 유일한 차이점은 당에서 추가적인 메틸렌기의 존재이고, 이는 다이데옥시헥소오스임을 의미한다. COSY 스펙트럼에서, 양성자 H-1' (δ_H 4.31)는 H-3' (δ_H 3.70)에 커플링된 H-2' (δ_H 3.25)에 커플링되어 있음을 확인하였다. 양성자 H-3'는 메틸렌 H-4' 양성자 (δ_H 1.95 및 1.42)에 커플링되었고, 차례로 H3-6' 양성자(δ_H 1.25)에 커플링된 H-5' 양성자 (δ_H 3.62)에 커플링되어 있음을 확인하였다. 이는 4,6-다이데옥시헥소오스 부분의 존재를 의미한다. 커플링 상수에 의하면, 양성자 H-2', H-3', 및 H-5'는 세로 방향인 것으로 확인되었다. 이는 당이 3-*O* 데메틸-찰코오스임을 의미한다. 큰 커플링 상수 ($J_{1',2'}$ = 7.5 Hz)는 일반적으로 $^{4}C_{1}$ 입체구조에서 D-당에 대하여 β-글리코시드 결합을 의미한다. 이것은 H-1' 및 H-5' 사이의 NOE 크로스 피크(도 5)에 의하여 뒷받침된다. 이러한 방법으로, 상기 획득된 1D 및 2D NMR 데이타는, 당 부분이 C-5에 부착된 3-*O*-데메틸-D-찰코오스임을 나타낸다.

[0101] DBVNNB의 NMR 데이타는 ODDCNB와 유사하였다. 이는 다이데옥시헥소오스 당을 포함하는 나르보마이신 유도체의 존재를 의미한다. COSY 스펙트럼에서, 아노머 양성자 H-1' (δ_H 4.81, dd, J = 8, 4.5 Hz)는 H-3' (δ_H 4.11)에 커플링된 메틸렌 양성자 (H_2-2': δ_H 1.78/1.80)에 커플링되어 있음을 확인하였다. 양성자 H-3' H-4' (δ_H 3.26)에 커플링되었고, 또한 차례로 H3-6' 양성자(δ_H 1.24)에 커플링된 H-5' (δ_H 4.05)에 커플링되어 있음을 확인하였다. 이는 당 부분이 2,6-다이데옥시헥소오스임을 의미한다. 아노머 양성자의 큰 커플링 상수(δ_H 4.81, dd, $J_{1',2'}$ = 8.0, 4.5 Hz)는 가로방향의 β-글리코시드 결합을 의미하고 따라서, $^{4}C_{1}$ 입체구조에 있어서 입체화학이 D-당으로 결정되었다. H-3' (J = 3 Hz) 및 H-4' (brs)의 작은 커플링 상수는 상기 두 양성자가 모두 가로방향으로 위치하고 각각의 히드록실기가 세로방향이라는 것을 의미한다. 따라서, 당은 D-보이비노오스로 확인되었고 상기 데이터는 공지된 보고에 잘 부합하였다. 또한, $^{4}C_{1}$ 입체구조는 H-1' 및 H-5' 사이의 NOE 크로스 피크(도 5)를 통하여 확인되었다.

[0102] DDGTNB의 NMR 데이타는 당 양성자의 커플링 상수에서의 차이를 제외하고는 DBVNNB와 유사하였다. NMR 데이터에 기초하여 다이데옥시당은 D-디지톡소오스로 확인되었고, 아노머 양성자의 큰 커플링 상수($J_{1',2'}$ = 9.5 Hz)는 세로방향의 양성자를 가지는 β-글리코시드 연결된 D 당을 나타낸다. ^{1}H NMR 스펙트럼은 *e,a* 및 *e,e* 커플링에 일반적인 H-3' 의 작은 커플링 상수 (δ_H 4.11, m, J = 3 Hz)를 통하여 3'-OH기의 세로방향을 확증해주었고, 따라서, 디지톡소오스 및 올리보오스 당을 구별하기 위하여 유용하였다. 이축 커플링 ($J_{4',5'}$ = 9.5 Hz)이 H-4' 및 H-5'사이에서 관찰되었는 바, 이는 가로방향의 5'-CH_3 및 4'-OH를 의미한다. 디지톡소오스 당의 $^{4}C_{1}$ 입체구조 (D-당에 일반적인)는 H-1' 및 H-5' 사이의 NOE 상관관계를 이용하여 확인하였다(도 5).

[0103] LOLVNB의 NMR 데이타는 H-3' (δ_H 3.90, ddd, J =12.0, 9.0, 5.0 Hz)에 대한 커플링 상수 및 스플릿 패턴에서의 주요 변화를 제외하고는 DDGTNB와 유사하였다. 커플링 상수 $J_{3',2'ax}$ = 12.0 Hz 및 $J_{3',4'}$ = 9.0 Hz는 H-3' 양성자가 세로방향으로 위치하는 반면 3'-OH가 가로방향으로 위치한다는 것을 의미한다. 따라서, DDGTNB에서 디지톡소오스인 것과 달리, 당 부분이 올리보오스로 확인되었다. 당의 입체배치는 아노머 양성자의 커플링 상수(J = 3 Hz)

및 NOE 크로스 피크(도 5)를 통하여 L-올리보오스로 결정되었다.

[0104] LDGTNB의 ^{1}H NMR 스펙트럼은 당 부분의 화학적 시프트에 있어서의 약간의 차이를 제외하고는 DDGTNB와 유사하였다. 가장 현저한 차이는 아노머 양성자 신호 (δ_H 5.03, brs, $J_{1',2'}$ = 3 Hz)이고, 이는 DDGTNB (δ_H 4.84, $J_{1',2'}$ = 9.5 Hz)에서와 달리, 세로방향의 글리코시드 연결의 존재를 의미한다. 1D 및 2D NMR 데이타는 LDGTNB의 구조를 뒷받침하였다. 또한, ^{1}H NMR 스펙트럼은 커플링 상수 ($J_{3',4'}$ = 3.0 Hz 및 $J_{4',5'}$ = 10.0 Hz)를 통하여, 가로방향의 H-3' 및 세로방향의 H-4' 및 H-5'를 확인하여, 디지톡소오스 당으로 확인하였다. NOESY 스펙트럼에서, DDGTNB에서 관찰된 H-1' 및 H-5' 사이의 NOE 상관관계 대신에, H-2'ax (δ_H 1.91, dt, J = 15.0, 3.5 Hz) 및 H-4' (δ_H 3.15, dd, J = 10.0, 3.0 Hz) 양성자 사이의 상관관계가 관찰되었다(도 5). 따라서, H-1'부터 C-5까지의 HMBC 크로스 피크에 의하여 측정된 바와 같이 당을 C-5 (δ_C 78.5)에 부착된 L-디지톡소오스로 확인하였다. 상기 데이타는 L-디지톡소오스에 대하여 보고된 데이터에 잘 부합되었다.

[0105] HPLC-ESI-MS 및 1D (^{1}H 및 ^{13}C) 및 2D (COSY, HMBC, HSQC, 및 NOESY) NMR 데이터를 통하여 나르보마이신 및 LRHMNB를 확인하였다. 상기 화합물에 있어서 당의 입체배치(configuration) 및 입체구조(conformation)를 커플링 상수 및 NOE 크로스 피크를 통하여 결정하였다.

[0106] **임상적 균주의 분리 및 항균제 감수성 측정**

[0107] 에리트로마이신-저항성 균주 *E. 파에시움* P00558 및 *S. 아우레우스* P00740를 대한민국 대구에 소재한 3차 대학병원에서 분리하였다. 상기 분리물을 속 및 종 동정 및 항균제 감수성 측정을 위하여 MicroScan WalkAway 96 시스템 (Dade Behring, West Sacramento, CA)를 이용하여 처리하였다. MicroScan 그람 양성 MIC/콤보 (PC1A) 패널 (Dade Behring), 그람 음성 MIC/ 콤보 (NC44) 패널 (Dade Behring), 및 그람 음성 구분점(breakpoint) 콤보 (NBC39) 패널 (Dade Behring)을 이용하여 항균제 감수성 및 *E. 파에시움 및 S. 아우레우스*의 종 레벨을 각각 동정하였다. 시험을 위한 접종물 현탁액을 동일한 깨끗한 18- 내지 24-시간-계대배양 플레이트를 이용하여 같은 날에 처리하였다. 미량희석 항균제 감수성 시험을 위하여 Prompt Inoculation 시스템-D (Dade Behring)를 이용하여 접종물을 표준화하였다. 모든 항균제 감수성 시험은 최신 Clinical and Laboratory Standards Institute 방법에 따라 수행하였다. 품질 관리는 *S. 아우레우스* ATCC 25923를 이용하여 관찰하였다.

[0108] **항균 활성 어세이**

[0109] 에리트로마이신-감수성 균주 *E. 파에시움* ATCC 19434 및 *S. 아우레우스* ATCC 25923 및 에리트로마이신-저항성 균주 *E. 파에시움* P00558 및 *S. 아우레우스* P00740를 이용하여 중성의 데옥시당을 포함하는 나르보마이신 및 그의 유도체의 항균 활성을 측정하였다. 에리트로마이신을 대조군으로서 사용하였다. 박테리아 세포 (2×10^{7} /ml)를 LB 또는 Mueller-Hinton 배지에 접종하고 96-웰 미세적정 플레이트로 0.1 ml/웰씩 나누었다. MIC를 National Committee for Clinical Laboratory Standards에 따른 항생제의 연속 2배 희석을 이용하여 측정하였다. 37 ℃에서 접종 24시간 후, 시험 개체의 성장을 방해하는 최소 항생제 농도를 MIC로서 결정하였다. 성장은 미세적정 ELISA Reader (Molecular Devices Emax, CA) 를 이용하여 620 nm에서의 흡광도를 관찰함으로써 어세이하였다. 그 결과를 하기 표 11에 나타내었다.

표 11

[0110]

물질	MIC (μM)			
	ATCC 19434	P00558	ATCC 25923	P00740
나르보마이신	10	10	10	20
DQNVNB	10	10	10	<10
LRHMNB	2.5	2.5	<2.5	1.25
ODDCNB	5	5	2.5	5
DBVNNB	<10	10	10	20
DDGTNB	10	<20	<20	20

LOLVNB	10	20	5	10
LDGTNB	10	20	10	20
에리트로마이신	10	>20	10	>40

[0111] D-데소사민 포함 나르보마이신은 에리트로마이신-감수성 균주에 대하여 에리트로마이신 (10 μM) 과 유사한 항균 활성을 나타내었다. 반면 에리트로마이신-저항성 균주에 대하여는 에리트로마이신 (> 20 μM)보다 더 높은 활성 (10 - 20 μM)이 관찰되었다. DQNVNB는 *E. 파에시움* ATCC 19434, *E. 파에시움* P00558, 및 *S. 아우레우스* ATCC 25923에 대하여 개선된 활성을 나타내지 않았지만, *S. 아우레우스* P00740의 성장을 저해하는 데에 있어서는 나르보마이신보다 효과적이었다. LOLVNB는 *S. 아우레우스* ATCC 25923에 대하여 선택적 활성을 나타내었다. 본 실험예에서, LRHMNB는 강력한 활성 (1.25 - 2.50 μM)을 나타내었다. ODDCNB 역시 모든 균주에 대하여 유의적인 활성 (2.5 - 5.0 μM)을 나타내었다. LRHMNB 의 *시험관 내 항균* 활성(1.25 - 2.50 μM) 을 LOLVNB 의 그것(5 - 20 μM)과 비교한 결과, LRHMNB의 2'-히드록실기의 구조-활성 정보를 확인하였고, 이는 항균 활성의 가장 중요한 결정 요인일 수 있다. ODDCNB의 항균 활성(2.5 - 5.0 μM)은 나르보마이신 (10 - 20 μM) 및 DQNVNB (≤ 10 μM) 보다 우수하였다. 이는 나르보마이신의 3'-다이메틸아미노 기가 3'-히드록실기로 교체된 것 및 DQNVNB에 있는 4'-히드록실기가 없는 것이 ODDCNB의 항균 활성을 개선시킨 것임을 확인하였다. *S. 아우레우스* ATCC 25923 및 P00740에 대한 LOLVNB (5 - 10 μM) 및 LDGTNB (10 - 20 μM)의 MIC를 비교한 결과, LOLVNB의 3'-가로방향(equatorial)-히드록실기가 항균 활성에 긍정적 영향을 미친다는 것을 의미한다. DBVNNB (≤ 10 μM) 는 *E. 파에시움* ATCC 19434, P00558, 및 *S. 아우레우스* ATCC 25923에 대하여 DDGTNB (< 20 μM) 보다 더 높은 항균 활성을 나타내었는 바, 이는 DBVNNB의 4'-세로방향(axial)-히드록실기가 MIC를 감소시킨다는 것을 의미한다.

수탁번호

[0112]

기탁기관명 : 한국미생물보존센터(국내)

수탁번호 : KFCC11507P

수탁일자 : 20110310

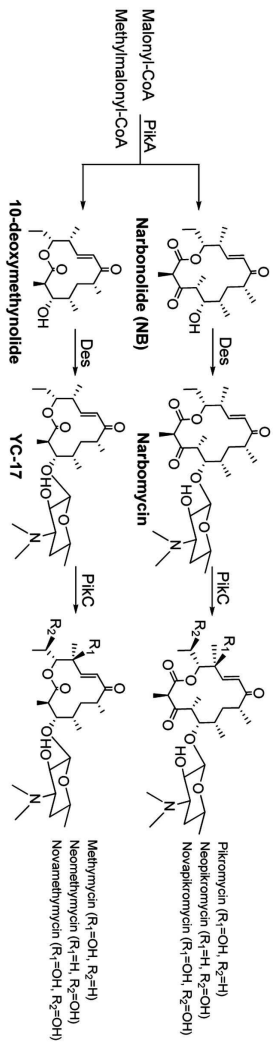
기탁기관명 : 한국미생물보존센터(국내)

수탁번호 : KFCC11508P

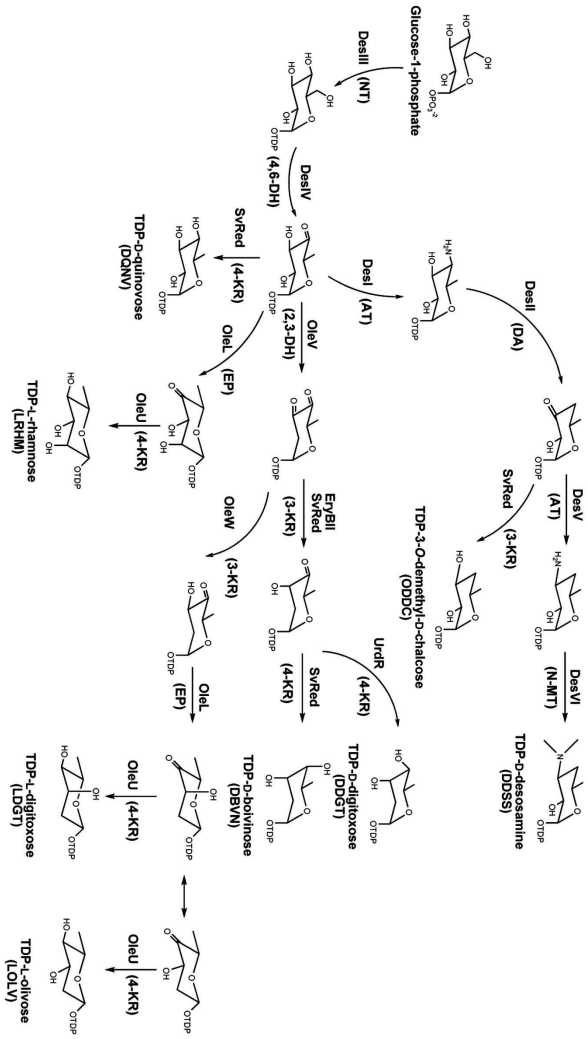
수탁일자 : 20110310

도면

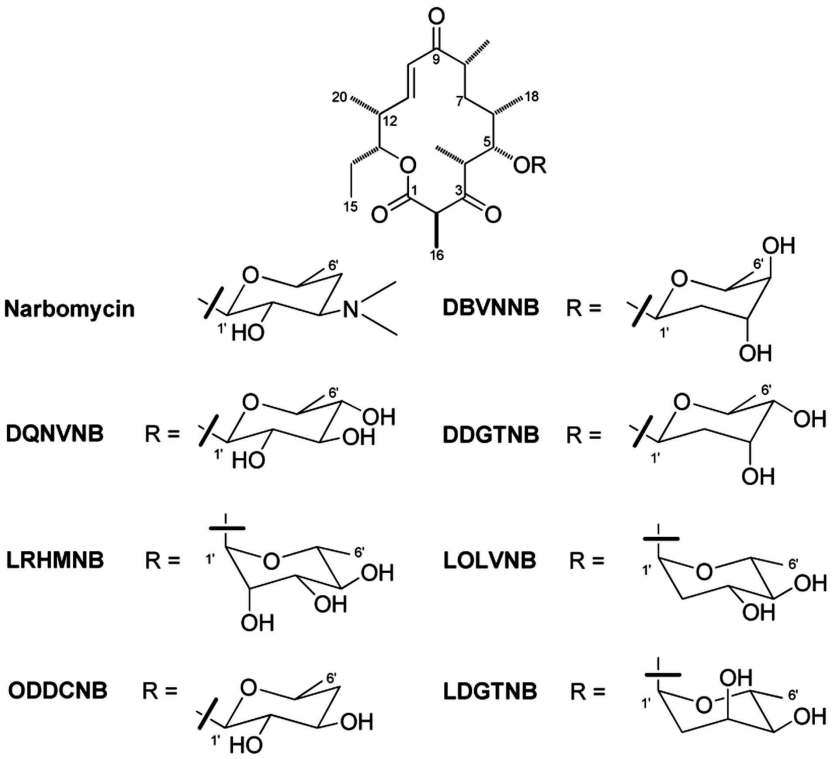
도면1


Malonyl-CoA
Methylmalonyl-CoA
PikA
10-deoxymethynolide
Narbonolide (NB)
Des
YC-17
Narbomycin
PikC
Methymycin (R1=OH, R2=H)
Neomethymycin (R1=H, R2=OH)
Novamethymycin (R1=OH, R2=OH)
Pikromycin (R1=OH, R2=H)
Neopikromycin (R1=H, R2=OH)
Novapikromycin (R1=OH, R2=OH)

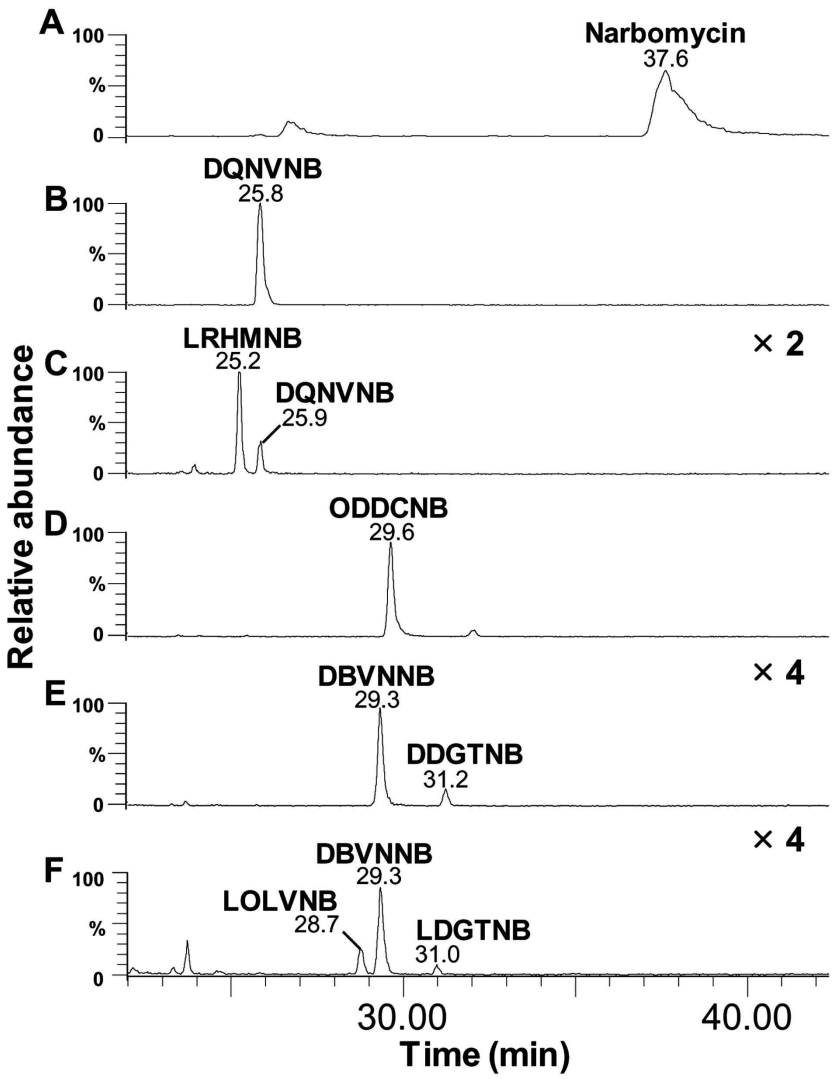
도면2


Glucose-1-phosphate
DesIII (NT)
DesIV (4,6-DH)
DesI (AT)
DesII (DA)
DesV (AT)
DesVI (N-MT)
TDP-D-desosamine (DDSS)
SvRed (3-KR)
TDP-3-O-demethyl-D-chalcose (ODDC)
UrdR (4-KR)
TDP-D-digitoxose (DDGT)
EryBII SvRed (3-KR)
OleV (2,3-DH)
SvRed (4-KR)
TDP-D-quinovose (DQNV)
OleL (EP)
OleU (4-KR)
TDP-L-rhamnose (LRHM)
OleW (3-KR)
OleL (EP)
TDP-D-boivinose (DBVN)
OleU (4-KR)
TDP-L-digitoxose (LDGT)
OleU (4-KR)
TDP-L-olivose (LOLV)

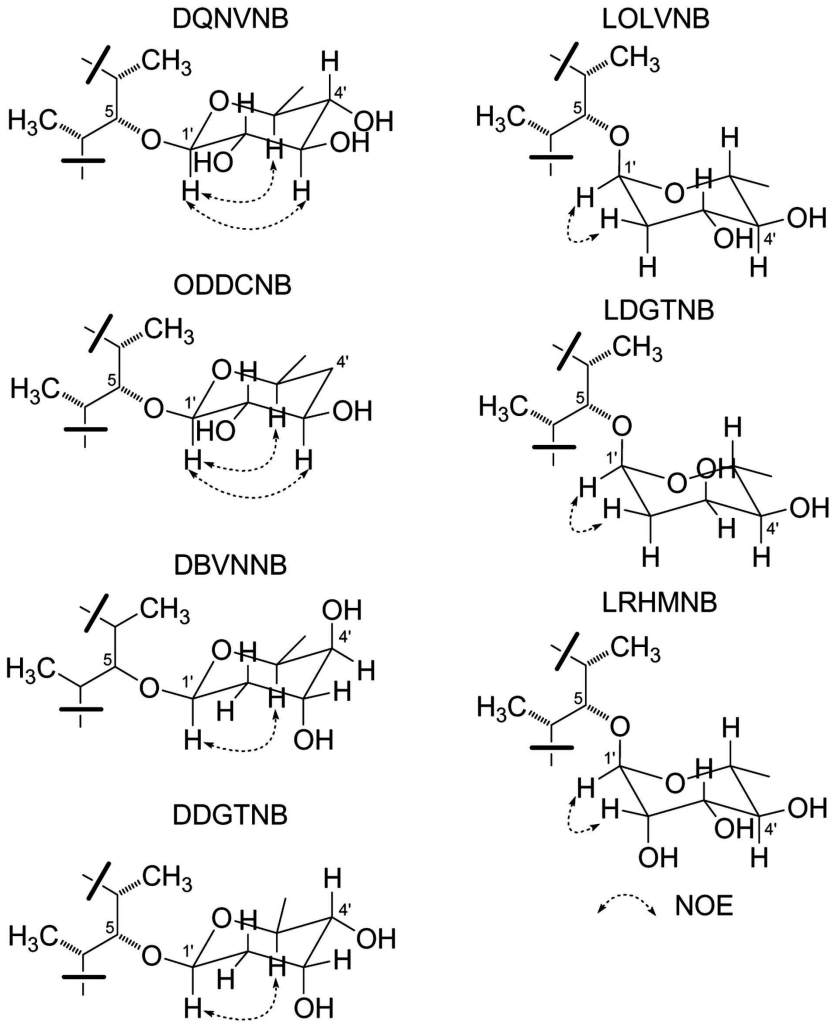
도면3


Narbomycin
DBVNNB R =
DQNVNB R =
DDGTNB R =
LRHMNB R =
LOLVNB R =
ODDCNB R =
LDGTNB R =

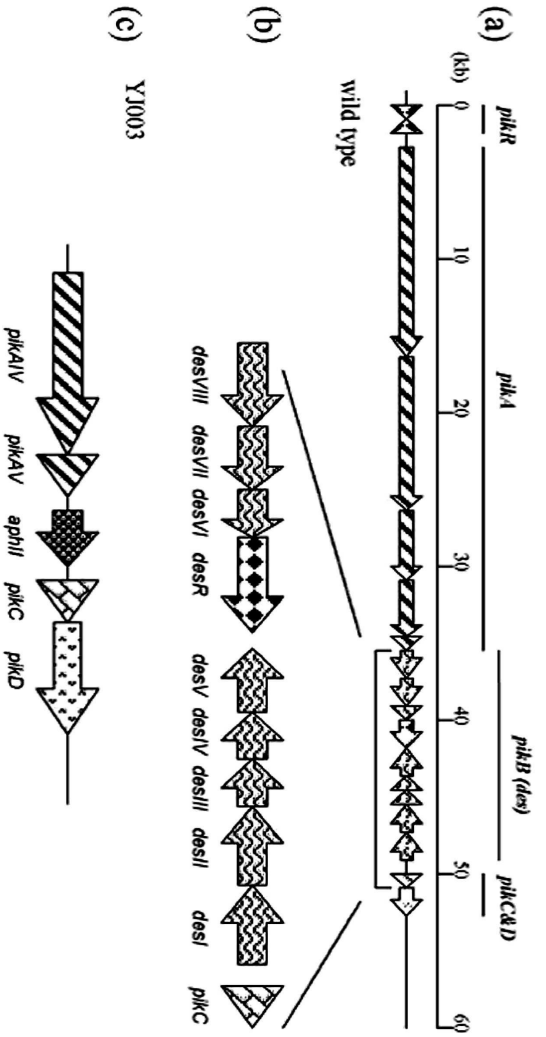
도면4


A
100
%
0
Narbomycin
37.6
B
DQNVNB
25.8
C
LRHMNB
25.2
DQNVNB
25.9
× 2
D
ODDCNB
29.6
E
DBVNNB
29.3
DDGTNB
31.2
× 4
F
LOLVNB
28.7
DBVNNB
29.3
LDGTNB
31.0
× 4
Relative abundance
30.00
40.00
Time (min)

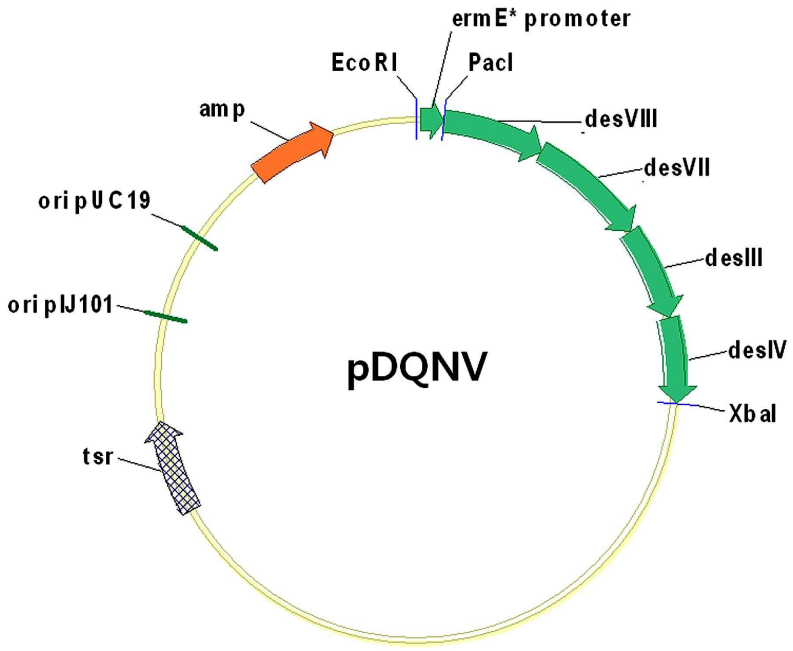
도면5



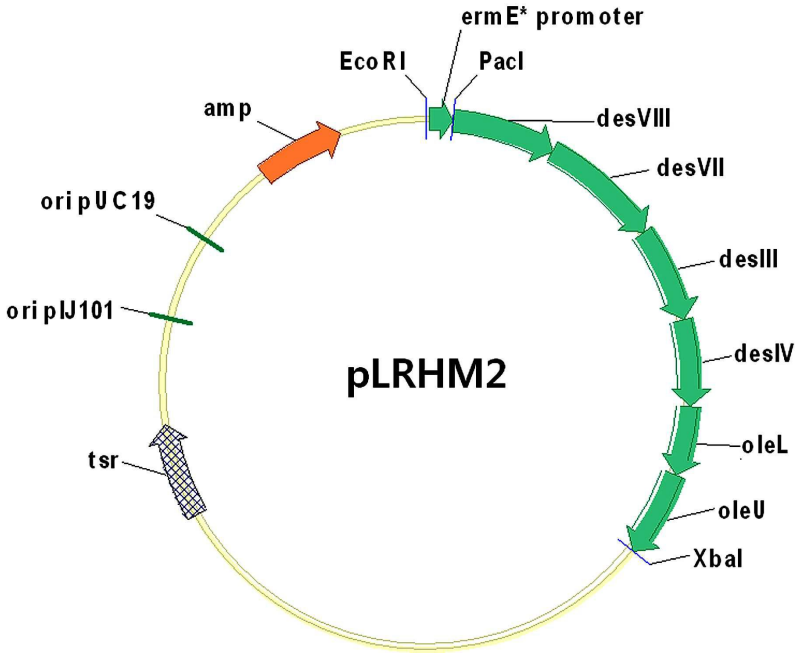
도면6


(a)
(kb)
0
10
20
30
40
50
60
pikR
pikA
pikB (des)
pikC&D
wild type
(b)
desVIII desVII desVI desR desV desIV desIII desII desI pikC
(c)
YJ003
pikAIV pikAV aphII pikC pikD

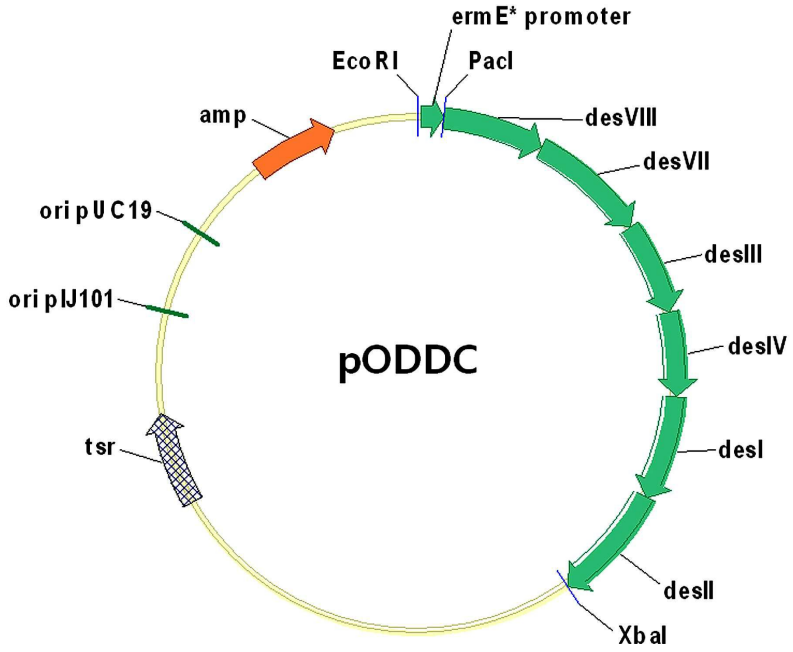
도면7


ermE* promoter
EcoRI
PacI
amp
desVIII
desVII
desIII
desIV
XbaI
oripUC19
oripIJ101
pDQNV
tsr

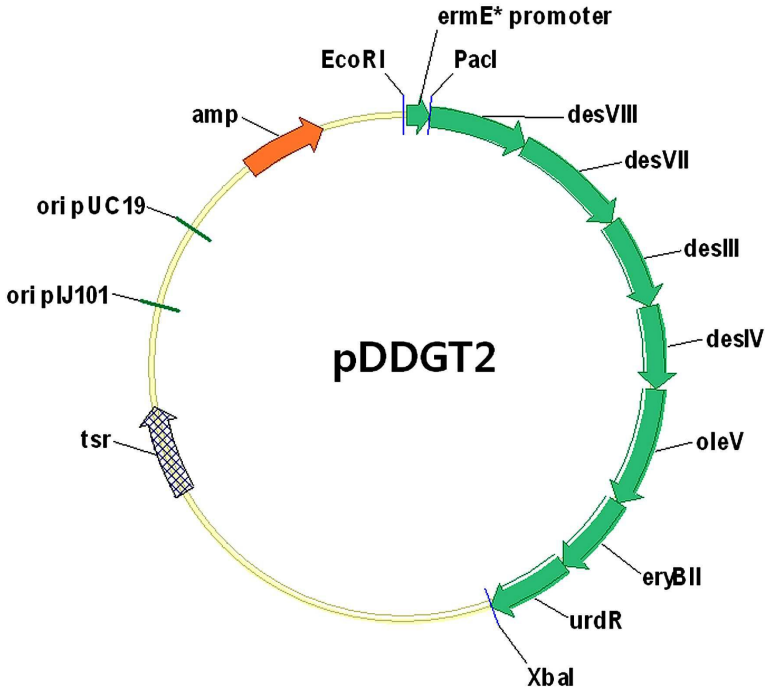
도면8


ermE* promoter
EcoRI
PacI
amp
desVIII
desVII
oripUC19
desIII
oriplJ101
pLRHM2
desIV
oleL
tsr
oleU
XbaI

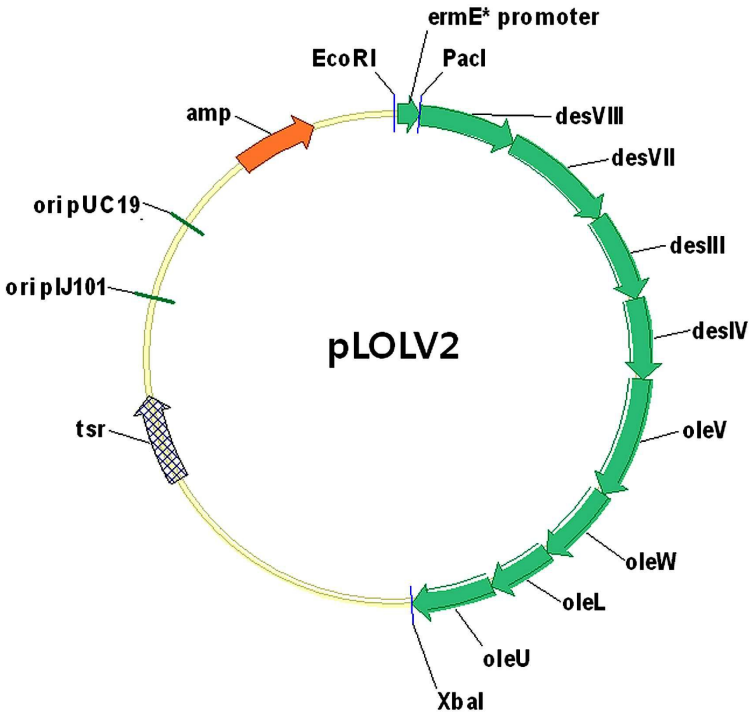
도면9


ermE* promoter
EcoRI
PacI
amp
desVIII
desVII
oripUC19
desIII
oriplJ101
pODDC
desIV
desI
tsr
desII
XbaI

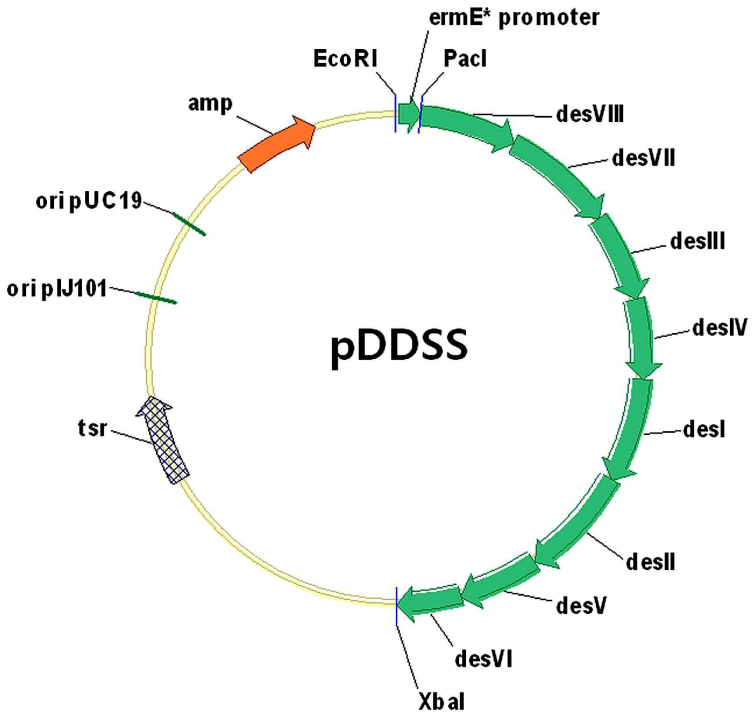
도면10


ermE* promoter
EcoRI
PacI
amp
desVIII
desVII
ori pUC19
desIII
ori pIJ101
pDDGT2
desIV
tsr
oleV
eryBII
urdR
XbaI

도면11


ermE* promoter
EcoRI
PacI
amp
desVIII
desVII
ori pUC19
desIII
ori pIJ101
pLOLV2
desIV
tsr
oleV
oleW
oleL
oleU
XbaI

도면12



서 열 목 록

<110> Ewha University - Industry Collaboration Foundation

<120> Novel narbomycin derivatives compounds, antibacterial composition comprising the same and manufacturing method of the same

<130> PA110076/KR

<160> 46

<170> KopatentIn 1.71

<210> 1

<211> 9715

<212> DNA

<213> Artificial Sequence

<220><223> desVIII-desVII-desIII-desIV-desI-desII-desV-desVI

<400> 1

ttaattaaga tatccccgac cagccttatc gaaggagcgg cgaagagatg gcagaaccca 60

cggtgaccga cgacctgacg ggggccctca cgcagccccc gctgggccgc accgtccgcg 120

cggtggccga ccgtgaactc ggcacccacc tcctggagac ccgcggcatc cactggatcc 180

acgccgcgaa cggcgacccg tacgccaccg tgctgcgcgg ccaggcggac gacccgtatc 240

ccgcgtacga gcgggtgcgt gcccgcggcg cgctctcctt cagcccgacg ggcagctggg 300

tcaccgccga tcacgccctg gcggcgagca tcctctgctc gacggacttc ggggtctccg 360

gcgccgacgg cgtcccggtg ccgcagcagg tcctctcgta cggggagggc tgtccgctgg 420

agcgcgagca ggtgctgccg gcggccggtg acgtgccgga gggcgggcag cgtgccgtgg 480

tcgaggggat ccaccgggag acgctggagg gtctcgcgcc ggacccgtcg gcgtcgtacg 540

ccttcgagct gctgggcggt ttcgtccgcc cggcggtgac ggccgctgcc gccgccgtgc 600

tgggtgttcc cgcggaccgg cgcgcggact tcgcggatct gctggagcgg ctccggccgc 660

tgtccgacag cctgctggcc ccgcagtccc tgcggacggt acgggcggcg gacggcgcgc 720

tggccgagct cacggcgctg ctcgccgatt cggacgactc ccccggggcc ctgctgtcgg 780

cgctcggggt caccgcagcc gtccagctca ccgggaacgc ggtgctcgcg ctcctcgcgc 840

atcccgagca gtggcgggag ctgtgcgacc ggcccgggct cgcggcggcc gcggtggagg 900

agaccctccg ctacgacccg ccggtgcagc tcgacgcccg ggtggtccgc ggggagacgg 960

agctggcggg ccggcggctg ccggccgggg cgcatgtcgt cgtcctgacc gccgcgaccg 1020

gccgggaccc ggaggtcttc acggacccgg agcgcttcga cctcgcgcgc cccgacgccg 1080

ccgcgcacct cgcgctgcac cccgccggtc cgtacggccc ggtggcgtcc ctggtccggc 1140

ttcaggcgga ggtcgcgctg cggaccctgg ccgggcgttt ccccgggctg cggcaggcgg 1200

gggacgtgct ccgcccccgc cgcgcgcctg tcggccgcgg gccgctgagc gtcccggtca 1260

gcagctcctg agacaccggg gccccggtcc gcccggcccc ccttcggacg gaccggacgg 1320

ctcggaccac ggggacggct cagaccgtcc cgtgtgtccc cgtccggctc ccgtccgccc 1380

catcccgccc ctccaccggc aaggaaggac acgacgccat gcgcgtcctg ctgacctcgt 1440

tcgcacatca cacgcactac tacggcctgg tgcccctggc ctgggcgctg ctcgccgccg 1500

ggcacgaggt gcgggtcgcc agccagcccg cgctcacgga caccatcacc gggtccgggc 1560

tcgccgcggt gccggtcggc accgaccacc tcatccacga gtaccgggtg cggatggcgg 1620

gcgagccgcg cccgaaccat ccggcgatcg ccttcgacga ggcccgtccc gagccgctgg 1680

actgggacca cgccctcggc atcgaggcga tcctcgcccc gtacttccat ctgctcgcca 1740

acaacgactc gatggtcgac gacctcgtcg acttcgcccg gtcctggcag ccggacctgg 1800

tgctgtggga gccgacgacc tacgcgggcg ccgtcgccgc ccaggtcacc ggtgccgcgc 1860

acgcccgggt cctgtggggg cccgacgtga tgggcagcgc ccgccgcaag ttcgtcgcgc 1920

tgcgggaccg gcagccgccc gagcaccgcg aggaccccac cgcggagtgg ctgacgtgga 1980

cgctcgaccg gtacggcgcc tccttcgaag aggagctgct caccggccag ttcacgatcg 2040

acccgacccc gccgagcctg cgcctcgaca cgggcctgcc gaccgtcggg atgcgttatg 2100

ttccgtacaa cggcacgtcg gtcgtgccgg actggctgag tgagccgccc gcgcggcccc 2160

gggtctgcct gaccctcggc gtctccgcgc gtgaggtcct cggcggcgac ggcgtctcgc 2220

agggcgacat cctggaggcg ctcgccgacc tcgacatcga gctcgtcgcc acgctcgacg 2280
cgagtcagcg cgccgagatc cgcaactacc cgaagcacac ccggttcacg gacttcgtgc 2340
cgatgcacgc gctcctgccg agctgctcgg cgatcatcca ccacggcggg gcgggcacct 2400
acgcgaccgc cgtgatcaac gcggtgccgc aggtcatgct cgccgagctg tgggacgcgc 2460
cggtcaaggc gcgggccgtc gccgagcagg gggcggggtt cttcctgccg ccggccgagc 2520
tcacgccgca ggccgtgcgg gacgccgtcg tccgcatcct cgacgacccc tcggtcgcca 2580
ccgccgcgca ccggctgcgc gaggagacct tcggcgaccc caccccggcc gggatcgtcc 2640

ccgagctgga gcggctcgcc gcgcagcacc gccgcccgcc ggccgacgcc cggcactgag 2700
ccgcacccct cgccccaggc ctcacccctg tatctcatat gtctagttaa ctcgccacgc 2760
cgaccgttat caccggcgcc ctgctgctag tttccgagaa tgaagggaat agtcctggcc 2820
ggcgggagcg gaactcggct gcatccggcg acctcggtca tttcgaagca gattcttccg 2880
gtctacaaca aaccgatgat ctactatccg ctgtcggttc tcatgctcgg cggtattcgc 2940
gagattcaaa tcatctcgac cccccagcac atcgaactct tccagtcgct tctcggaaac 3000
ggcaggcacc tgggaataga actcgactat gcggtccaga aagagcccgc aggaatcgcg 3060

gacgcacttc tcgtcggagc cgagcacatc ggcgacgaca cctgcgccct gatcctgggc 3120
gacaacatct tccacgggcc cggcctctac acgctcctgc gggacagcat cgcgcgcctc 3180
gacggctgcg tgctcttcgg ctacccggtc aaggaccccg agcggtacgg cgtcgccgag 3240
gtggacgcga cgggccggct gaccgacctc gtcgagaagc ccgtcaagcc gcgctccaac 3300
ctcgccgtca ccggcctcta cctctacgac aacgacgtcg tcgacatcgc caagaacatc 3360
cggccctcgc cgcgcggcga gctggagatc accgacgtca accgcgtcta cctggagcgg 3420
ggccgggccg aactcgtcaa cctgggccgc ggcttcgcct ggctggacac cggcacccac 3480

gactcgctcc tgcgggccgc ccagtacgtc caggtcctgg aggagcggca gggcgtctgg 3540
atcgcgggcc ttgaggagat cgccttccgc atgggcttca tcgacgccga ggcctgtcac 3600
ggcctgggag aaggcctctc ccgcaccgag tacggcagct atctgatgga gatcgccggc 3660
cgcgagggag ccccgtgagg gcacctcgcg gccgacgcgt tcccacgacc gacagcgcca 3720
ccgacagtgc gacccacacc gcgacccgca ccgccaccga cagtgcgacc cacaccgcga 3780
cctacagcgc gaccgaaagg aagacggcag tgcggcttct ggtgaccgga ggtgcgggct 3840
tcatcggctc gcacttcgtg cggcagctcc tcgccggggc gtaccccgac gtgcccgccg 3900

atgaggtgat cgtcctggac agcctcacct acgcgggcaa ccgcgccaac ctcgccccgg 3960

tggacgcgga cccgcgactg cgcttcgtcc acggcgacat ccgcgacgcc ggcctcctcg 4020

cccgggaact gcgcggcgtg gacgccatcg tccacttcgc ggccgagagc cacgtggacc 4080

gctccatcgc gggcgcgtcc gtgttcaccg agaccaacgt gcagggcacg cagacgctgc 4140

tccagtgcgc cgtcgacgcc ggcgtcggcc gggtcgtgca cgtctccacc gacgaggtgt 4200

acgggtcgat cgactccggc tcctggaccg agagcagccc gctggagccc aactcgccct 4260

acgcggcgtc caaggccggc tccgacctcg ttgcccgcgc ctaccaccgg acgtacggcc 4320

tcgacgtacg gatcacccgc tgctgcaaca actacgggcc gtaccagcac cccgagaagc 4380

tcatccccct cttcgtgacg aacctcctcg acggcgggac gctcccgctg tacggcgacg 4440

gcgcgaacgt ccgcgagtgg gtgcacaccg acgaccactg ccggggcatc gcgctcgtcc 4500

tcgcgggcgg ccgggccggc gagatctacc acatcggcgg cggcctggag ctgaccaacc 4560

gcgaactcac cggcatcctc ctggactcgc tcggcgccga ctggtcctcg gtccggaagg 4620

tcgccgaccg caagggccac gacctgcgct actccctcga cggcggcgag atcgagcgcg 4680

agctcggcta ccgcccgcag gtctccttcg cggacggcct cgcgcggacc gtccgctggt 4740

accgggagaa ccgcggctgg tgggagccgc tcaaggcgac cgccccgcag ctgcccgcca 4800

ccgccgtgga ggtgtccgcg tgagcagccg cgccgagacc ccccgcgtcc ccttcctcga 4860

cctcaaggcc gcctacgagg agctctctag tgctagcgac ggtggcccgg agggaacatc 4920

cgtgaaaagc gccttatccg acctcgcatt cttcggcggc cccgccgctt tcgaccagcc 4980

gctcctcgtg gggcggccca accgcatcga ccgcgccagg ctgtacgagc ggctcgaccg 5040

ggccctcgac agccagtggc tgtccaacgg cggcccgctc gtccgcgagt tcgaggagcg 5100

cgtcgccggg ctcgccgggg tccggcatgc cgtggccacc tgcaacgcca cggccgggct 5160

ccagctcctc gcgcacgccg ccggcctcac cggcgaagtg atcatgccgt cgatgacgtt 5220

cgccgccacc ccgcacgcac tgcgctggat cggcctcacc ccggtcttcg ccgacatcga 5280

cccggacacc ggcaacctcg acccggacca ggtggccgcc gcggtcacac cccgcacctc 5340

ggccgtcgtc ggcgtccacc tctggggccg cccctgcgcc gccgaccagc tgcggaaggt 5400

cgccgacgag cacggcctgc ggctgtactt cgacgccgcg cacgccctcg gctgcgcggt 5460

cgacggccgg cccgccggca gcctcggcga cgccgaggtc ttcagcttcc acgccaccaa 5520

ggccgtcaac gccttcgagg gcggcgccgt cgtcaccgac gacgccgacc tcgccgcccg 5580

gatccgcgcc ctccacaact tcggcttcga cctgcccggc ggcagccccg ccggcgggac 5640

caacgccaag atgagcgagg ccgccgccgc catgggcctc acctccctcg acgcgtttcc 5700

cgaggtcatc gaccggaacc ggcgcaacca cgccgcctac cgcgagcacc tcgcggacct 5760

ccccggcgtc ctcgtcgccg accacgaccg ccacggcctc aacaaccacc agtacgtgat 5820

cgtcgagatc gacgaggcca ccaccggcat ccaccgcgac ctcgtcatgg aggtcctgaa 5880

ggccgaaggc gtgcacaccc gcgcctactt ctcgccgggc tgccacgagc tggagccgta 5940

ccgcgggcag ccgcacgccc cgctgccgca caccgaacgc ctcgccgcgc gcgtgctgtc 6000

cctgccgacc ggcaccgcca tcggcgacga cgacatccgc cgggtcgccg acctgctgcg 6060

tctctgcgcg acccgcggcc gcgaactgac cgcgcgccac cgcgacacgg cccccgcccc 6120

gctcgcggcc ccccagacat ccacgcccac gattggacgc tcccgatgac cgccccgcc 6180

ctttccgcca ccgccccggc cgaacgctgc gcgcaccccg gagccgatct gggggcggcg 6240

gtccacgccg tcggccagac cctcgccgcc ggcggcctcg tgccgcccga cgaggccgga 6300

acgaccgccc gccacctcgt ccggctcgcc gtgcgctacg gcaacagccc cttcaccccg 6360

ctggaggagg cccgccacga cctgggcgtc gaccgggacg ccttccggcg cctcctcgcc 6420

ctgttcgggc aggtcccgga gctccgcacc gcggtcgaga ccggccccgc cggggcgtac 6480

tggaagaaca ccctgctccc gctcgaacag cgcggcgtct tcgacgcggc gctcgccagg 6540

aagcccgtct tcccgtacag cgtcggcctc taccccggcc cgacctgcat gttccgctgc 6600

cacttctgcg tccgtgtgac cggcgcccgc tacgacccgt ccgccctcga cgccggcaac 6660

gccatgttcc ggtcggtcat cgacgagata cccgcgggca acccctcggc gatgtacttc 6720

tccggcggcc tggagccgct caccaacccc ggcctcggga gcctggccgc gcacgccacc 6780

gaccacggcc tgcggcccac cgtctacacg aactccttcg cgctcaccga gcgcaccctg 6840

gagcgccagc ccggcctctg gggcctgcac gccatccgca cctcgctcta cggcctcaac 6900

gacgaggagt acgagcagac caccggcaag aaggccgcct tccgccgcgt ccgcgagaac 6960

ctgcgccgct tccagcagct gcgcgccgag cgcgagtcgc cgatcaacct cggcttcgcc 7020

tacatcgtgc tcccgggccg tgcctcccgc ctgctcgacc tggtcgactt catcgccgac 7080

ctcaacgacg ccgggcaggg caggacgatc gacttcgtca acattcgcga ggactacagc 7140

ggccgtgacg acggcaagct gccgcaggag gagcgggccg agctccagga ggccctcaac 7200

gccttcgagg agcgggtccg cgagcgcacc cccggactcc acatcgacta cggctacgcc 7260

ctgaacagcc tgcgcaccgg ggccgacgcc gaactgctgc ggatcaagcc cgccaccatg 7320

cggcccaccg cgcacccgca ggtcgcggtg caggtcgatc tcctcggcga cgtgtacctg 7380

taccgcgagg ccggcttccc cgacctggac ggcgcgaccc gctacatcgc gggccgcgtg 7440

acccccgaca cctccctcac cgaggtcgtc agggacttcg tcgagcgcgg cggcgaggtg 7500

gcggccgtcg acggcgacga gtacttcatg gacggcttcg atcaggtcgt caccgcccgc 7560

ctgaaccagc tggagcgcga cgccgcggac ggctgggagg aggcccgcgg cttcctgcgc 7620

tgacccgcat ctagtcaggt ctccttcgcg gacggcctcg cgcggaccgt ccgctggtac 7680

cgggagaacc gcggctggtg ggagccgctc aaggcgaccg ccccgcagct gcccgccacc 7740
gccgtggagg tgtccgcgtg agcagccgcg ccgagacccc ccgcgtcccc ttcctcgacc 7800
tcaaggccgc ctacgaggag ctccgcgcgg agaccgacgc cgcgatcgcc cgcgtcctcg 7860
actcggggcg ctacctcctc ggacccgaac tcgaaggatt cgaggcggag ttcgccgcgt 7920
actgcgagac ggaccacgcc gtcggcgtga acagcgggat ggacgccctc cagctcgccc 7980
tccgcggcct cggcatcgga cccggggacg aggtgatcgt cccctcgcac acgtacatcg 8040
ccagctggct cgcggtgtcc gccaccggcg cgacccccgt gcccgtcgag ccgcacgagg 8100

accaccccac cctggacccg ctgctcgtcg agaaggcgat caccccccgc acccgggcgc 8160
tcctccccgt ccacctctac gggcaccccg ccgacatgga cgccctccgc gagctcgcgg 8220
accggcacgg cctgcacatc gtcgaggacg ccgcgcaggc ccacggcgcc cgctaccggg 8280
gccggcggat cggcgccggg tcgtcggtgg ccgcgttcag cttctacccg ggcaagaacc 8340
tcggctgctt cggcgacggc ggcgccgtcg tcaccggcga ccccgagctc gccgaacggc 8400
tccggatgct ccgcaactac ggctcgcggc agaagtacag ccacgagacg aagggcacca 8460
actcccgcct ggacgagatg caggccgccg tgctgcggat ccggctcgcc cacctggaca 8520

gctggaacgg ccgcaggtcg gcgctggccg cggagtacct ctccgggctc gccggactgc 8580
ccggcatcgg cctgccggtg accgcgcccg acaccgaccc ggtctggcac ctcttcaccg 8640
tgcgcaccga gcgccgcgac gagctgcgca gccacctcga cgcccgcggc atcgacaccc 8700
tcacgcacta cccggtaccc gtgcacctct cgcccgccta cgcgggcgag gcaccgccgg 8760
aaggctcgct cccgcgggcc gagagcttcg cgcggcaggt cctcagcctg ccgatcggcc 8820
cgcacctgga gcgcccgcag gcgctgcggg tgatcgacgc cgtgcgcgaa tgggccgagc 8880
gggtcgacca ggcctagtca ggttctagtc aattgcccca ggcctcaccc ctgtatctgc 8940

gccgggggac gcccccggcc caccctccga aagaccgaaa gcaggagcac cgtgtacgaa 9000
gtcgaccacg ccgacgtcta cgacctcttc tacctgggtc gcggcaagga ctacgccgcc 9060
gaggcctccg acatcgccga cctggtgcgc tcccgtaccc ccgaggcctc ctcgctcctg 9120
gacgtggcct gcggtacggg cacgcatctg gagcacttca ccaaggagtt cggcgacacc 9180
gccggcctgg agctgtccga ggacatgctc acccacgccc gcaagcggct gcccgacgcc 9240
acgctccacc agggcgacat gcgggacttc cggctcggcc ggaagttctc cgccgtggtc 9300
agcatgttca gctccgtcgg ctacctgaag acgaccgagg aactcggcgc ggccgtcgcc 9360

tcgttcgcgg agcacctgga gcccggtggc gtcgtcgtcg tcgagccgtg gtggttcccg 9420

gagaccttcg ccgacggctg ggtcagcgcc gacgtcgtcc gccgtgacgg gcgcaccgtg 9480

gcccgtgtct cgcactcggt gcgggagggg aacgcgacgc gcatggaggt ccacttcacc 9540

gtggccgacc cgggcaaggg cgtgcggcac ttctccgacg tccatctcat caccctgttc 9600

caccaggccg agtacgaggc cgcgttcacg gccgccgggc tgcgcgtcga gtacctggag 9660

ggcggcccgt cgggccgtgg cctcttcgtc ggcgtccccg cctgagtcta gaccc 9715

<210> 2

<211> 292

<212> PRT

<213> Artificial Sequence

<220><223> DesIII protein

<400> 2

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

Ala Thr Ser Val Ile Ser Lys Gln Ile Leu Pro Val Tyr Asn Lys Pro
20 25 30

Met Ile Tyr Tyr Pro Leu Ser Val Leu Met Leu Gly Gly Ile Arg Glu
35 40 45

Ile Gln Ile Ile Ser Thr Pro Gln His Ile Glu Leu Phe Gln Ser Leu

50 55 60

Leu Gly Asn Gly Arg His Leu Gly Ile Glu Leu Asp Tyr Ala Val Gln
65 70 75 80

Lys Glu Pro Ala Gly Ile Ala Asp Ala Leu Leu Val Gly Ala Glu His
85 90 95

Ile Gly Asp Asp Thr Cys Ala Leu Ile Leu Gly Asp Asn Ile Phe His
100 105 110

Gly Pro Gly Leu Tyr Thr Leu Leu Arg Asp Ser Ile Ala Arg Leu Asp
115 120 125

Gly Cys Val Leu Phe Gly Tyr Pro Val Lys Asp Pro Glu Arg Tyr Gly
130 135 140

Val Ala Glu Val Asp Ala Thr Gly Arg Leu Thr Asp Leu Val Glu Lys
145 150 155 160

Pro Val Lys Pro Arg Ser Asn Leu Ala Val Thr Gly Leu Tyr Leu Tyr
165 170 175
Asp Asn Asp Val Val Asp Ile Ala Lys Asn Ile Arg Pro Ser Pro Arg
180 185 190
Gly Glu Leu Glu Ile Thr Asp Val Asn Arg Val Tyr Leu Glu Arg Gly

195 200 205
Arg Ala Glu Leu Val Asn Leu Gly Arg Gly Phe Ala Trp Leu Asp Thr
210 215 220
Gly Thr His Asp Ser Leu Leu Arg Ala Ala Gln Tyr Val Gln Val Leu
225 230 235 240
Glu Glu Arg Gln Gly Val Trp Ile Ala Gly Leu Glu Glu Ile Ala Phe
245 250 255
Arg Met Gly Phe Ile Asp Ala Glu Ala Cys His Gly Leu Gly Glu Gly
260 265 270

Leu Ser Arg Thr Glu Tyr Gly Ser Tyr Leu Met Glu Ile Ala Gly Arg
275 280 285
Glu Gly Ala Pro
290

<210> 3
<211> 337
<212> PRT
<213> Artificial Sequence
<220><223> DesIV protein
<400> 3
Met Arg Leu Leu Val Thr Gly Gly Ala Gly Phe Ile Gly Ser His Phe
1 5 10 15
Val Arg Gln Leu Leu Ala Gly Ala Tyr Pro Asp Val Pro Ala Asp Glu
20 25 30

Val Ile Val Leu Asp Ser Leu Thr Tyr Ala Gly Asn Arg Ala Asn Leu
35 40 45
Ala Pro Val Asp Ala Asp Pro Arg Leu Arg Phe Val His Gly Asp Ile
50 55 60

Arg Asp Ala Gly Leu Leu Ala Arg Glu Leu Arg Gly Val Asp Ala Ile
65 70 75 80
Val His Phe Ala Ala Glu Ser His Val Asp Arg Ser Ile Ala Gly Ala
85 90 95
Ser Val Phe Thr Glu Thr Asn Val Gln Gly Thr Gln Thr Leu Leu Gln

100 105 110
Cys Ala Val Asp Ala Gly Val Gly Arg Val Val His Val Ser Thr Asp
115 120 125
Glu Val Tyr Gly Ser Ile Asp Ser Gly Ser Trp Thr Glu Ser Ser Pro
130 135 140
Leu Glu Pro Asn Ser Pro Tyr Ala Ala Ser Lys Ala Gly Ser Asp Leu
145 150 155 160
Val Ala Arg Ala Tyr His Arg Thr Tyr Gly Leu Asp Val Arg Ile Thr
165 170 175

Arg Cys Cys Asn Asn Tyr Gly Pro Tyr Gln His Pro Glu Lys Leu Ile
180 185 190
Pro Leu Phe Val Thr Asn Leu Leu Asp Gly Gly Thr Leu Pro Leu Tyr
195 200 205
Gly Asp Gly Ala Asn Val Arg Glu Trp Val His Thr Asp Asp His Cys
210 215 220
Arg Gly Ile Ala Leu Val Leu Ala Gly Gly Arg Ala Gly Glu Ile Tyr
225 230 235 240
His Ile Gly Gly Gly Leu Glu Leu Thr Asn Arg Glu Leu Thr Gly Ile

245 250 255
Leu Leu Asp Ser Leu Gly Ala Asp Trp Ser Ser Val Arg Lys Val Ala
260 265 270
Asp Arg Lys Gly His Asp Leu Arg Tyr Ser Leu Asp Gly Gly Glu Ile
275 280 285
Glu Arg Glu Leu Gly Tyr Arg Pro Gln Val Ser Phe Ala Asp Gly Leu
290 295 300
Ala Arg Thr Val Arg Trp Tyr Arg Glu Asn Arg Gly Trp Trp Glu Pro

305 310 315 320

Leu Lys Ala Thr Ala Pro Gln Leu Pro Ala Thr Ala Val Glu Val Ser
325 330 335

Ala

<210> 4

<211> 415

<212> PRT

<213> Artificial Sequence

<220><223> DesI protein

<400> 4

Met Lys Ser Ala Leu Ser Asp Leu Ala Phe Phe Gly Gly Pro Ala Ala
1 5 10 15

Phe Asp Gln Pro Leu Leu Val Gly Arg Pro Asn Arg Ile Asp Arg Ala
20 25 30

Arg Leu Tyr Glu Arg Leu Asp Arg Ala Leu Asp Ser Gln Trp Leu Ser

35 40 45

Asn Gly Gly Pro Leu Val Arg Glu Phe Glu Glu Arg Val Ala Gly Leu
50 55 60

Ala Gly Val Arg His Ala Val Ala Thr Cys Asn Ala Thr Ala Gly Leu
65 70 75 80

Gln Leu Leu Ala His Ala Ala Gly Leu Thr Gly Glu Val Ile Met Pro
85 90 95

Ser Met Thr Phe Ala Ala Thr Pro His Ala Leu Arg Trp Ile Gly Leu
100 105 110

Thr Pro Val Phe Ala Asp Ile Asp Pro Asp Thr Gly Asn Leu Asp Pro

115 120 125

Asp Gln Val Ala Ala Ala Val Thr Pro Arg Thr Ser Ala Val Val Gly
130 135 140

Val His Leu Trp Gly Arg Pro Cys Ala Ala Asp Gln Leu Arg Lys Val
145 150 155 160

Ala Asp Glu His Gly Leu Arg Leu Tyr Phe Asp Ala Ala His Ala Leu

```
                165                 170                 175
Gly Cys Ala Val Asp Gly Arg Pro Ala Gly Ser Leu Gly Asp Ala Glu

            180                 185                 190
Val Phe Ser Phe His Ala Thr Lys Ala Val Asn Ala Phe Glu Gly Gly
        195                 200                 205
Ala Val Val Thr Asp Asp Ala Asp Leu Ala Ala Arg Ile Arg Ala Leu
    210                 215                 220
His Asn Phe Gly Phe Asp Leu Pro Gly Gly Ser Pro Ala Gly Gly Thr
225                 230                 235                 240
Asn Ala Lys Met Ser Glu Ala Ala Ala Ala Met Gly Leu Thr Ser Leu
                245                 250                 255

Asp Ala Phe Pro Glu Val Ile Asp Arg Asn Arg Arg Asn His Ala Ala
            260                 265                 270
Tyr Arg Glu His Leu Ala Asp Leu Pro Gly Val Leu Val Ala Asp His
        275                 280                 285
Asp Arg His Gly Leu Asn Asn His Gln Tyr Val Ile Val Glu Ile Asp
    290                 295                 300
Glu Ala Thr Thr Gly Ile His Arg Asp Leu Val Met Glu Val Leu Lys
305                 310                 315                 320
Ala Glu Gly Val His Thr Arg Ala Tyr Phe Ser Pro Gly Cys His Glu

                325                 330                 335
Leu Glu Pro Tyr Arg Gly Gln Pro His Ala Pro Leu Pro His Thr Glu
            340                 345                 350
Arg Leu Ala Ala Arg Val Leu Ser Leu Pro Thr Gly Thr Ala Ile Gly
        355                 360                 365
Asp Asp Asp Ile Arg Arg Val Ala Asp Leu Leu Arg Leu Cys Ala Thr
    370                 375                 380
Arg Gly Arg Glu Leu Thr Ala Arg His Arg Asp Thr Ala Pro Ala Pro
385                 390                 395                 400

Leu Ala Ala Pro Gln Thr Ser Thr Pro Thr Ile Gly Arg Ser Arg
                405                 410                 415
```

<210> 5

<211> 485

<212> PRT

<213> Artificial Sequence

<220><223> DesII protein

<400> 5

Met Thr Ala Pro Ala Leu Ser Ala Thr Ala Pro Ala Glu Arg Cys Ala
1 5 10 15

His Pro Gly Ala Asp Leu Gly Ala Ala Val His Ala Val Gly Gln Thr
20 25 30

Leu Ala Ala Gly Gly Leu Val Pro Pro Asp Glu Ala Gly Thr Thr Ala
35 40 45

Arg His Leu Val Arg Leu Ala Val Arg Tyr Gly Asn Ser Pro Phe Thr
50 55 60

Pro Leu Glu Glu Ala Arg His Asp Leu Gly Val Asp Arg Asp Ala Phe
65 70 75 80

Arg Arg Leu Leu Ala Leu Phe Gly Gln Val Pro Glu Leu Arg Thr Ala
85 90 95

Val Glu Thr Gly Pro Ala Gly Ala Tyr Trp Lys Asn Thr Leu Leu Pro
100 105 110

Leu Glu Gln Arg Gly Val Phe Asp Ala Ala Leu Ala Arg Lys Pro Val
115 120 125

Phe Pro Tyr Ser Val Gly Leu Tyr Pro Gly Pro Thr Cys Met Phe Arg
130 135 140

Cys His Phe Cys Val Arg Val Thr Gly Ala Arg Tyr Asp Pro Ser Ala
145 150 155 160

Leu Asp Ala Gly Asn Ala Met Phe Arg Ser Val Ile Asp Glu Ile Pro
165 170 175

Ala Gly Asn Pro Ser Ala Met Tyr Phe Ser Gly Gly Leu Glu Pro Leu
180 185 190

Thr Asn Pro Gly Leu Gly Ser Leu Ala Ala His Ala Thr Asp His Gly
195 200 205

Leu Arg Pro Thr Val Tyr Thr Asn Ser Phe Ala Leu Thr Glu Arg Thr
210 215 220
Leu Glu Arg Gln Pro Gly Leu Trp Gly Leu His Ala Ile Arg Thr Ser
225 230 235 240
Leu Tyr Gly Leu Asn Asp Glu Glu Tyr Glu Gln Thr Thr Gly Lys Lys
245 250 255

Ala Ala Phe Arg Arg Val Arg Glu Asn Leu Arg Arg Phe Gln Gln Leu
260 265 270
Arg Ala Glu Arg Glu Ser Pro Ile Asn Leu Gly Phe Ala Tyr Ile Val
275 280 285
Leu Pro Gly Arg Ala Ser Arg Leu Leu Asp Leu Val Asp Phe Ile Ala
290 295 300
Asp Leu Asn Asp Ala Gly Gln Gly Arg Thr Ile Asp Phe Val Asn Ile
305 310 315 320
Arg Glu Asp Tyr Ser Gly Arg Asp Asp Gly Lys Leu Pro Gln Glu Glu

325 330 335
Arg Ala Glu Leu Gln Glu Ala Leu Asn Ala Phe Glu Glu Arg Val Arg
340 345 350
Glu Arg Thr Pro Gly Leu His Ile Asp Tyr Gly Tyr Ala Leu Asn Ser
355 360 365
Leu Arg Thr Gly Ala Asp Ala Glu Leu Leu Arg Ile Lys Pro Ala Thr
370 375 380
Met Arg Pro Thr Ala His Pro Gln Val Ala Val Gln Val Asp Leu Leu
385 390 395 400

Gly Asp Val Tyr Leu Tyr Arg Glu Ala Gly Phe Pro Asp Leu Asp Gly
405 410 415
Ala Thr Arg Tyr Ile Ala Gly Arg Val Thr Pro Asp Thr Ser Leu Thr
420 425 430
Glu Val Val Arg Asp Phe Val Glu Arg Gly Gly Glu Val Ala Ala Val
435 440 445
Asp Gly Asp Glu Tyr Phe Met Asp Gly Phe Asp Gln Val Val Thr Ala
450 455 460

Arg Leu Asn Gln Leu Glu Arg Asp Ala Ala Asp Gly Trp Glu Glu Ala

465 470 475 480

Arg Gly Phe Leu Arg

485

<210> 6

<211> 379

<212> PRT

<213> Artificial Sequence

<220><223> DesV protein

<400> 6

Met Ser Ser Arg Ala Glu Thr Pro Arg Val Pro Phe Leu Asp Leu Lys

1 5 10 15

Ala Ala Tyr Glu Glu Leu Arg Ala Glu Thr Asp Ala Ala Ile Ala Arg

20 25 30

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

35 40 45

Glu Ala Glu Phe Ala Ala Tyr Cys Glu Thr Asp His Ala Val Gly Val

50 55 60

Asn Ser Gly Met Asp Ala Leu Gln Leu Ala Leu Arg Gly Leu Gly Ile

65 70 75 80

Gly Pro Gly Asp Glu Val Ile Val Pro Ser His Thr Tyr Ile Ala Ser

85 90 95

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

100 105 110

His Glu Asp His Pro Thr Leu Asp Pro Leu Leu Val Glu Lys Ala Ile

115 120 125

Thr Pro Arg Thr Arg Ala Leu Leu Pro Val His Leu Tyr Gly His Pro

130 135 140

Ala Asp Met Asp Ala Leu Arg Glu Leu Ala Asp Arg His Gly Leu His

145 150 155 160

Ile Val Glu Asp Ala Ala Gln Ala His Gly Ala Arg Tyr Arg Gly Arg

165 170 175

Arg Ile Gly Ala Gly Ser Ser Val Ala Ala Phe Ser Phe Tyr Pro Gly

180 185 190

Lys Asn Leu Gly Cys Phe Gly Asp Gly Gly Ala Val Val Thr Gly Asp

195 200 205

Pro Glu Leu Ala Glu Arg Leu Arg Met Leu Arg Asn Tyr Gly Ser Arg

210 215 220

Gln Lys Tyr Ser His Glu Thr Lys Gly Thr Asn Ser Arg Leu Asp Glu

225 230 235 240

Met Gln Ala Ala Val Leu Arg Ile Arg Leu Ala His Leu Asp Ser Trp

245 250 255

Asn Gly Arg Arg Ser Ala Leu Ala Ala Glu Tyr Leu Ser Gly Leu Ala

260 265 270

Gly Leu Pro Gly Ile Gly Leu Pro Val Thr Ala Pro Asp Thr Asp Pro

275 280 285

Val Trp His Leu Phe Thr Val Arg Thr Glu Arg Arg Asp Glu Leu Arg

290 295 300

Ser His Leu Asp Ala Arg Gly Ile Asp Thr Leu Thr His Tyr Pro Val

305 310 315 320

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

325 330 335

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

340 345 350

Ile Gly Pro His Leu Glu Arg Pro Gln Ala Leu Arg Val Ile Asp Ala

355 360 365

Val Arg Glu Trp Ala Glu Arg Val Asp Gln Ala

370 375

<210> 7

<211> 237

<212> PRT

<213> Artificial Sequence

<220><223> DesVI protein

<400> 7

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

65 70 75 80
Pro Asp Ala Thr Leu His Gln Gly Asp Met Arg Asp Phe Arg Leu Gly
85 90 95
Arg Lys Phe Ser Ala Val Val Ser Met Phe Ser Ser Val Gly Tyr Leu
100 105 110
Lys Thr Thr Glu Glu Leu Gly Ala Ala Val Ala Ser Phe Ala Glu His
115 120 125
Leu Glu Pro Gly Gly Val Val Val Val Glu Pro Trp Trp Phe Pro Glu
130 135 140

Thr Phe Ala Asp Gly Trp Val Ser Ala Asp Val Val Arg Arg Asp Gly
145 150 155 160
Arg Thr Val Ala Arg Val Ser His Ser Val Arg Glu Gly Asn Ala Thr
165 170 175
Arg Met Glu Val His Phe Thr Val Ala Asp Pro Gly Lys Gly Val Arg
180 185 190
His Phe Ser Asp Val His Leu Ile Thr Leu Phe His Gln Ala Glu Tyr
195 200 205
Glu Ala Ala Phe Thr Ala Ala Gly Leu Arg Val Glu Tyr Leu Glu Gly

210 215 220
Gly Pro Ser Gly Arg Gly Leu Phe Val Gly Val Pro Ala
225 230 235

<210> 8

<211> 204

<212> PRT

<213> Artificial Sequence

<220><223> OleL

<400> 8

Met Glu Leu Leu Asp Val Asp Gly Ala Trp Leu Tyr Thr Pro Glu Ile
1 5 10 15

Met Arg Asp Glu Arg Gly Glu Phe Leu Glu Trp Phe Arg Gly Arg Thr
20 25 30

Phe Gln Glu Lys Ile Gly His Pro Leu Ser Leu Ala Gln Ala Asn Cys
35 40 45

Ser Val Ser Arg Lys Ala Phe Cys Ala Ala Ser Thr Ser Pro Thr Pro
50 55 60

Pro Pro Gly Gln Ala Lys Tyr Val Thr Cys Ala Ser Gly Thr Val Leu
65 70 75 80

Asp Val Val Val Asp Val Arg Arg Gly Ser Pro Thr Phe Gly Arg Trp
85 90 95

Ala Ala Val Arg Leu Asp Ala Ala Arg His Gln Gly Leu Tyr Leu Ala

100 105 110

Glu Gly Leu Gly His Ala Phe Met Ala Leu Thr Asp Asp Ala Thr Val
115 120 125

Val Tyr Leu Cys Ser Gln Pro Tyr Val Ala Glu Ala Glu Arg Ala Val
130 135 140

Asp Pro Leu Asp Pro Ala Ile Gly Ile Glu Trp Pro Thr Asp Ile Asp
145 150 155 160

Ile Val Pro Val Gly Glu Gly Thr Pro Thr His Arg Pro Trp Arg Arg
165 170 175

Pro Arg Arg Pro Gly Ile Leu Pro Asp Tyr Glu Gly Val Pro Gly Ala
180 185 190

Leu His Arg Gly Gly Gly Arg Arg Gly Thr Gly Pro
195 200

<210> 9

<211> 295

<212> PRT

<213> Artificial Sequence

<220><223> OleU protein

<400> 9

Met Arg Trp Leu Ile Thr Gly Ala Ala Gly Met Leu Gly Arg Glu Leu
1 5 10 15

Val Arg Arg Leu Ala Glu Asn Glu Glu Asp Val Ala Ala Leu Gly His
20 25 30

Asp His Leu Asp Val Thr Arg Pro Ser Ala Val Arg Ala Ala Leu Ala
35 40 45

Glu His Arg Pro Gly Ile Val Val Asn Cys Ala Ala Tyr Thr Ala Val
50 55 60

Asp Asp Ala Glu Thr Asp Glu Ala Ala Ala Ala Leu Leu Asn Ala Glu
65 70 75 80

Ala Pro Arg Leu Leu Ala Glu Gly Leu Arg Pro His Arg Arg His Gly
85 90 95

Leu Val His Leu Ser Thr Asp Tyr Val Phe Pro Gly Asp Ala Arg Thr
100 105 110

Pro Tyr Ala Glu Asp His Pro Thr Ala Pro Arg Ser Ala Tyr Gly Arg
115 120 125

Thr Lys Arg Asp Gly Glu Gln Ala Val Leu Thr Ala Leu Pro Thr Ala
130 135 140

Thr Val Leu Arg Thr Ala Trp Leu Tyr Gly Arg Thr Gly Arg Ser Phe
145 150 155 160

Val Arg Thr Met Ile Glu Arg Glu Ala Arg Gly Gly Ala Ile Asp Val
165 170 175

Val Ala Asp Gln Arg Gly Gln Pro Thr Trp Thr Gly Asp Leu Ala Asp
180 185 190

Arg Ile Ile Ala Val Gly Arg His Pro Gly Val His Gly Ile Leu His
195 200 205

Ala Thr Asn Ala Gly Ser Ala Thr Trp Tyr Asp Leu Ala Gln Glu Val
210 215 220
Phe Arg Leu Leu Asp Ala Asp Pro Gly Arg Val Arg Pro Thr Thr Gly
225 230 235 240

Ala Ala Phe Arg Arg Pro Ala Pro Arg Pro Ala Tyr Ser Val Leu Gly
245 250 255
His Asp Arg Trp Arg Gly Thr Gly Leu Ala Pro Leu Arg Asp Trp Arg
260 265 270
Ser Ala Leu Arg Glu Ala Phe Pro Asp Ile Leu Ala Ala Glu His Pro
275 280 285
Pro Thr Arg Arg Gly Ala Ala
290 295

<210> 10
<211> 474
<212> PRT
<213> Artificial Sequence
<220><
223> OleV protein
<400> 10
Met Ile Trp Gly Ile Pro Ala Met Ser Glu Ala Met Gly Ser Val Pro
1 5 10 15
Thr Ala Gly Ser Glu Val Ser Ser Thr Cys Ala Phe Leu Ser Trp Leu
20 25 30
Asp Ala Arg Arg Arg Ala Asn Arg Leu Thr Val Glu His Val Pro Phe
35 40 45
Arg Glu Leu Ser Gly Trp Gln Phe Asp Glu Asn Thr Gly Asn Leu Arg
50 55 60

His Thr Ser Gly Arg Phe Phe Ser Ile Glu Gly Leu Arg Val Arg Thr
65 70 75 80
Asp His Cys Trp Phe Gly Ser Trp Thr Gln Pro Ile Ile Val Gln Pro
85 90 95
Glu Ile Gly Ile Leu Gly Leu Leu Val Lys Arg Phe Asp Gly Ile Leu
100 105 110

His Val Leu Val Gln Ala Lys Asn Glu Pro Gly Asn Ile Gly Gly Leu
115 120 125

Gln Leu Ser Pro Thr Val Gln Ala Thr Arg Ser Asn Tyr Thr Arg Val

130 135 140

His Arg Gly Gly Gly Val Arg Tyr Leu Glu Tyr Phe Ala Ser Pro Arg
145 150 155 160

Gly Arg Gly Arg Val Leu Ala Asp Val Leu Gln Ser Glu Gln Gly Ser
165 170 175

Trp Phe Leu His Lys Arg Asn Arg Asn Met Val Val Glu Ala Leu Asp
180 185 190

Asp Val Pro Leu Asp Asp Asp Phe His Trp Ile Ser Leu Gly Gly Leu
195 200 205

Arg Lys Leu Leu Leu Arg Pro His Leu Val Asn Met Asp Thr Arg Thr
210 215 220

Val Leu Ser Cys Leu Pro Pro Asp Pro Ala Pro Asp Gly Arg Gln Pro
225 230 235 240

Pro Ala Pro Ala Ala Pro Phe Ala Ala Ala Val Thr Arg Ser Leu Thr
245 250 255

Arg Gly Ala Thr Ala Leu His Thr Met Gly Glu Ile Leu Gly Trp Leu
260 265 270

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

275 280 285

Glu Glu Thr Ala Phe Ser Gly Trp Arg Arg Asp Asp His Ala Ile Ala
290 295 300

His Lys Asp Gly Asp Tyr Phe Arg Val Ile Gly Val Ser Val Arg Ala
305 310 315 320

Ser Ser Arg Glu Val Ser Ser Trp Ser Gln Pro Leu Leu Ala Pro Val
325 330 335

Gly Pro Gly Leu Ala Ala Phe Val Thr Arg Arg Ile Arg Gly Val Leu
340 345 350

His Val Leu Leu His Ala Arg Thr Glu Ala Gly Leu Leu Asn Gly Pro

355 360 365

Glu Met Ala Pro Thr Val Gln Cys Arg Pro Leu Asn Tyr Arg Ala Val

370 375 380

Pro Ala Glu Tyr Arg Pro Ala Tyr Leu Asp Tyr Val Leu Ser Ala Asp

385 390 395 400

Pro Gly Arg Ile Arg Tyr Asp Thr Leu Gln Ser Glu Glu Gly Gly Arg

405 410 415

Phe His His Ala Glu Asn Arg Tyr Val Val Val Glu Ala Glu Asp Asp

420 425 430

Phe Pro Val Glu Val Pro Arg Asp Phe Arg Trp Leu Thr Leu His Gln

435 440 445

Ile Leu Ala Leu Leu His His Ser Asn Tyr Val Asn Val Glu Ala Arg

450 455 460

Ser Leu Val Ala Cys Ile Gln Ala Leu Ser

465 470

<210> 11

<211> 333

<212> PRT

<213> Artificial Sequence

<220><223> EryBII protein

<400> 11

Met Thr Thr Asp Ala Ala Thr His Val Arg Leu Gly Arg Ser Ala Leu

1 5 10 15

Leu Thr Ser Arg Leu Trp Leu Gly Thr Val Asn Phe Ser Gly Arg Val

20 25 30

Glu Asp Asp Asp Ala Leu Arg Leu Met Asp His Ala Arg Asp Arg Gly

35 40 45

Ile Asn Cys Leu Asp Thr Ala Asp Met Tyr Gly Trp Arg Leu Tyr Lys

50 55 60

Gly His Thr Glu Glu Leu Val Gly Arg Trp Leu Ala Gln Gly Gly Gly

65 70 75 80

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

85 90 95

Glu Arg Val Asn Asp Ser Gly Leu Ser Ala Arg His Ile Ile Ala Ser

100 105 110

Cys Glu Gly Ser Leu Arg Arg Leu Gly Val Asp His Ile Asp Val Tyr

115 120 125

Gln Met His His Ile Asp Arg Ser Ala Pro Trp Asp Glu Val Trp Gln

130 135 140

Ala Met Asp Ser Leu Val Ala Ser Gly Lys Val Ser Tyr Val Gly Ser

145 150 155 160

Ser Asn Phe Ala Gly Trp His Ile Ala Ala Ala Gln Glu Asn Ala Ala

165 170 175

Arg Arg His Ser Leu Gly Met Val Ser His Gln Cys Leu Tyr Asn Leu

180 185 190

Ala Val Arg His Ala Glu Leu Glu Val Leu Pro Ala Ala Gln Ala Tyr

195 200 205

Gly Leu Gly Val Phe Ala Trp Ser Pro Leu His Gly Gly Leu Leu Ser

210 215 220

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

225 230 235 240

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

245 250 255

Tyr Glu Lys Phe Cys Arg Asn Leu Gly Glu Asp Pro Ala Glu Val Gly

260 265 270

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

275 280 285

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

290 295 300

Thr Leu Asp Glu Gln Ala Leu Ser Glu Leu Asp Glu Ile Phe Pro Ala

305 310 315 320

Val Ala Ser Gly Gly Ala Ala Pro Glu Ala Trp Leu Gln

325 330

<210> 12

<211> 247

<212> PRT

<213> Artificial Sequence

<220><223> UrdR protein

<400> 12

Met Asp Ile Val Gly Asn Gly Phe Leu Ala Arg Asn Leu Arg Pro Leu
1 5 10 15

Ala Glu Arg His Pro Asp Thr Val Ala Leu Ala Ala Gly Val Ser Trp
20 25 30

Ala Ser Gly Thr Ser Asp Ala Asp Phe Ala Arg Glu Ala Ala Leu Leu
35 40 45

Arg Asp Thr Ala Lys Gln Cys Ala Ala Thr Gly Arg Arg Leu Leu Phe
50 55 60

Phe Ser Thr Ala Ala Thr Gly Met Tyr Gly Leu Ala Glu Gly Pro Gly
65 70 75 80

Arg Glu Asp Thr Pro Val Thr Pro Cys Thr Pro Tyr Gly Ala His Lys
85 90 95

Leu Ala Leu Glu Glu Leu Leu Arg Asp Ser Gly Ala Asp His Val Ile
100 105 110

Leu Arg Leu Gly His Leu Val Gly Pro Asp Gln Pro Glu His Gln Leu
115 120 125

Leu Pro Thr Leu Val Arg His Leu Arg Glu Gly Ala Val Arg Ile His
130 135 140

Arg Gly Ala Ala Arg Asp Leu Ile Asp Val Ser Asp Val Val Thr Ile
145 150 155 160

Val Asp Cys Leu Leu Gly Leu Asp Leu Lys Ala Glu Thr Val Asn Val
165 170 175

Ala Ser Gly Tyr Ala Val Pro Val Lys Asp Ile Val Glu Leu Leu Arg
180 185 190

Arg Arg Leu Gly Val Glu Ala Arg Tyr Glu Phe Gln Asp Ala Gly Gly
195 200 205

Gln His Val Ile Ser Ile Glu Lys Leu Arg Ala Leu Val Pro Gln Val

210 215 220

Gln Asp Met Gly Phe Gly Pro Asp Tyr Tyr Arg Arg Ile Leu Ser Asp

225 230 235 240

Phe Thr Ser Ala Val Arg Ile

245

<210> 13

<211> 328

<212> PRT

<213> Artificial Sequence

<220><223> OleW protein

<400> 13

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

1 5 10 15

Ala Leu Arg Arg Thr Val Pro Ala Leu Leu Ala His Pro Asp Val Thr

20 25 30

Val Val Ala Val Ser Ser Arg Asp Thr Ala Arg Ala Ala Arg Phe Ala

35 40 45

Ala Ala Phe Gly Cys Glu Ala Val Pro Gly His Gln Ala Leu Leu Asp

50 55 60

Arg Asp Asp Ile Asp Ala Leu Tyr Val Pro Leu Pro Val Met Val His

65 70 75 80

Thr Pro Trp Val Glu Ala Ala Leu Leu Arg Gly Arg His Val Leu Val

85 90 95

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

100 105 110

Leu Ala Arg Ser Arg Gly Leu Val Leu Met Glu Asn Phe Thr Ser Leu

115 120 125

His His Ala Gln His Gly Thr Val Thr Asp Leu Leu Arg Asp Gly Thr

130 135 140

Ile Gly Glu Leu Arg Ser Leu Ser Ala Ala Phe Thr Ile Pro Pro Lys

145 150 155 160

Pro Glu Gly Asp Ile Arg Tyr Gln Pro Asp Val Gly Gly Gly Ala Leu

165 170 175

Leu Asp Ile Gly Ile Tyr Pro Leu Arg Ala Ala Leu His Phe Leu Gly

180 185 190

Pro Asp Leu His Ala Ala Gly Ala Val Leu Arg Arg Glu Arg Arg Arg

195 200 205

Asn Val Val Val Ser Gly His Val Leu Leu Thr Thr Pro His Gly Val

210 215 220

Val Ala Glu Leu Ala Phe Gly Met Glu His Ala Tyr Arg Ser Glu Tyr

225 230 235 240

Thr Leu Phe Gly Thr Ala Gly Arg Leu Arg Leu Asp Arg Ala Phe Thr

245 250 255

Pro Pro Glu Thr His Arg Pro Arg Val Glu Ile His Arg Gln Asp Ala

260 265 270

Leu Asp Ile Val Asp Leu Pro Pro Asp Ala Gln Phe Ala Asn Leu Val

275 280 285

Arg Asp Phe Val Leu Ala Val Arg Glu Gly Pro Gly Arg Leu Thr Gln

290 295 300

His His Ala Asp Ala Val Arg Gln Ala Asp Leu Val Glu Arg Val Met

305 310 315 320

Ala Val Ala Arg Val Arg Trp Cys

325

<210> 14

<211> 402

<212> PRT

<213> Artificial Sequence

<220><223> DesVIII protein

<400> 14

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

1 5 10 15

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

20 25 30

Thr Arg Gly Ile His Trp Ile His Ala Ala Asn Gly Asp Pro Tyr Ala

35 40 45

Thr Val Leu Arg Gly Gln Ala Asp Asp Pro Tyr Pro Ala Tyr Glu Arg
50 55 60
Val Arg Ala Arg Gly Ala Leu Ser Phe Ser Pro Thr Gly Ser Trp Val
65 70 75 80
Thr Ala Asp His Ala Leu Ala Ala Ser Ile Leu Cys Ser Thr Asp Phe
85 90 95
Gly Val Ser Gly Ala Asp Gly Val Pro Val Pro Gln Gln Val Leu Ser
100 105 110
Tyr Gly Glu Gly Cys Pro Leu Glu Arg Glu Gln Val Leu Pro Ala Ala

115 120 125
Gly Asp Val Pro Glu Gly Gly Gln Arg Ala Val Val Glu Gly Ile His
130 135 140
Arg Glu Thr Leu Glu Gly Leu Ala Pro Asp Pro Ser Ala Ser Tyr Ala
145 150 155 160
Phe Glu Leu Leu Gly Gly Phe Val Arg Pro Ala Val Thr Ala Ala Ala
165 170 175
Ala Ala Val Leu Gly Val Pro Ala Asp Arg Arg Ala Asp Phe Ala Asp
180 185 190

Leu Leu Glu Arg Leu Arg Pro Leu Ser Asp Ser Leu Leu Ala Pro Gln
195 200 205
Ser Leu Arg Thr Val Arg Ala Ala Asp Gly Ala Leu Ala Glu Leu Thr
210 215 220
Ala Leu Leu Ala Asp Ser Asp Asp Ser Pro Gly Ala Leu Leu Ser Ala
225 230 235 240
Leu Gly Val Thr Ala Ala Val Gln Leu Thr Gly Asn Ala Val Leu Ala
245 250 255
Leu Leu Ala His Pro Glu Gln Trp Arg Glu Leu Cys Asp Arg Pro Gly

260 265 270
Leu Ala Ala Ala Ala Val Glu Glu Thr Leu Arg Tyr Asp Pro Pro Val
275 280 285

Gln Leu Asp Ala Arg Val Val Arg Gly Glu Thr Glu Leu Ala Gly Arg
290 295 300
Arg Leu Pro Ala Gly Ala His Val Val Val Leu Thr Ala Ala Thr Gly
305 310 315 320
Arg Asp Pro Glu Val Phe Thr Asp Pro Glu Arg Phe Asp Leu Ala Arg
325 330 335

Pro Asp Ala Ala Ala His Leu Ala Leu His Pro Ala Gly Pro Tyr Gly
340 345 350
Pro Val Ala Ser Leu Val Arg Leu Gln Ala Glu Val Ala Leu Arg Thr
355 360 365
Leu Ala Gly Arg Phe Pro Gly Leu Arg Gln Ala Gly Asp Val Leu Arg
370 375 380
Pro Arg Arg Ala Pro Val Gly Arg Gly Pro Leu Ser Val Pro Val Ser
385 390 395 400
Ser Ser

<

210> 15

<211> 426

<212> PRT

<213> Artificial Sequence

<220><223> DesVII protein

<400> 15

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

50 55 60
Arg Met Ala Gly Glu Pro Arg Pro Asn His Pro Ala Ile Ala Phe Asp
65 70 75 80

Glu Ala Arg Pro Glu Pro Leu Asp Trp Asp His Ala Leu Gly Ile Glu
85 90 95
Ala Ile Leu Ala Pro Tyr Phe His Leu Leu Ala Asn Asn Asp Ser Met
100 105 110
Val Asp Asp Leu Val Asp Phe Ala Arg Ser Trp Gln Pro Asp Leu Val
115 120 125

Leu Trp Glu Pro Thr Thr Tyr Ala Gly Ala Val Ala Ala Gln Val Thr
130 135 140
Gly Ala Ala His Ala Arg Val Leu Trp Gly Pro Asp Val Met Gly Ser
145 150 155 160
Ala Arg Arg Lys Phe Val Ala Leu Arg Asp Arg Gln Pro Pro Glu His
165 170 175
Arg Glu Asp Pro Thr Ala Glu Trp Leu Thr Trp Thr Leu Asp Arg Tyr
180 185 190
Gly Ala Ser Phe Glu Glu Glu Leu Leu Thr Gly Gln Phe Thr Ile Asp

195 200 205
Pro Thr Pro Pro Ser Leu Arg Leu Asp Thr Gly Leu Pro Thr Val Gly
210 215 220
Met Arg Tyr Val Pro Tyr Asn Gly Thr Ser Val Val Pro Asp Trp Leu
225 230 235 240
Ser Glu Pro Pro Ala Arg Pro Arg Val Cys Leu Thr Leu Gly Val Ser
245 250 255
Ala Arg Glu Val Leu Gly Gly Asp Gly Val Ser Gln Gly Asp Ile Leu
260 265 270

Glu Ala Leu Ala Asp Leu Asp Ile Glu Leu Val Ala Thr Leu Asp Ala
275 280 285
Ser Gln Arg Ala Glu Ile Arg Asn Tyr Pro Lys His Thr Arg Phe Thr
290 295 300
Asp Phe Val Pro Met His Ala Leu Leu Pro Ser Cys Ser Ala Ile Ile
305 310 315 320
His His Gly Gly Ala Gly Thr Tyr Ala Thr Ala Val Ile Asn Ala Val
325 330 335

Pro Gln Val Met Leu Ala Glu Leu Trp Asp Ala Pro Val Lys Ala Arg

340 345 350

Ala Val Ala Glu Gln Gly Ala Gly Phe Phe Leu Pro Pro Ala Glu Leu

355 360 365

Thr Pro Gln Ala Val Arg Asp Ala Val Val Arg Ile Leu Asp Asp Pro

370 375 380

Ser Val Ala Thr Ala Ala His Arg Leu Arg Glu Glu Thr Phe Gly Asp

385 390 395 400

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

405 410 415

His Arg Arg Pro Pro Ala Asp Ala Arg His

420 425

<210> 16

<211> 4891

<212> DNA

<213> Artificial Sequence

<220><223> desVIII-desVII-desIII-desIV

<400> 16

ttaattaaga tatccccgac cagccttatc gaaggagcgg cgaagagatg gcagaaccca 60

cggtgaccga cgacctgacg ggggccctca cgcagccccc gctgggccgc accgtccgcg 120

cggtggccga ccgtgaactc ggcacccacc tcctggagac ccgcggcatc cactggatcc 180

acgccgcgaa cggcgacccg tacgccaccg tgctgcgcgg ccaggcggac gacccgtatc 240

ccgcgtacga gcgggtgcgt gcccgcggcg cgctctcctt cagcccgacg ggcagctggg 300

tcaccgccga tcacgccctg gcggcgagca tcctctgctc gacggacttc ggggtctccg 360

gcgccgacgg cgtcccggtg ccgcagcagg tcctctcgta cggggagggc tgtccgctgg 420

agcgcgagca ggtgctgccg gcggccggtg acgtgccgga gggcgggcag cgtgccgtgg 480

tcgaggggat ccaccgggag acgctggagg gtctcgcgcc ggacccgtcg gcgtcgtacg 540

ccttcgagct gctgggcggt ttcgtccgcc cggcggtgac ggccgctgcc gccgccgtgc 600

tgggtgttcc cgcggaccgg cgcgcggact tcgcggatct gctggagcgg ctccggccgc 660

tgtccgacag cctgctggcc ccgcagtccc tgcggacggt acgggcggcg gacggcgcgc 720

tggccgagct cacggcgctg ctcgccgatt cggacgactc ccccggggcc ctgctgtcgg 780

cgctcggggt caccgcagcc gtccagctca ccgggaacgc ggtgctcgcg ctcctcgcgc 840

atcccgagca gtggcgggag ctgtgcgacc ggcccgggct cgcggcggcc gcggtggagg 900

agaccctccg ctacgacccg ccggtgcagc tcgacgcccg ggtggtccgc ggggagacgg 960

agctggcggg ccggcggctg ccggccgggg cgcatgtcgt cgtcctgacc gccgcgaccg 1020

gccgggaccc ggaggtcttc acggacccgg agcgcttcga cctcgcgcgc cccgacgccg 1080

ccgcgcacct cgcgctgcac cccgccggtc cgtacggccc ggtggcgtcc ctggtccggc 1140

ttcaggcgga ggtcgcgctg cggaccctgg ccgggcgttt ccccgggctg cggcaggcgg 1200

gggacgtgct ccgcccccgc cgcgcgcctg tcggccgcgg gccgctgagc gtcccggtca 1260

gcagctcctg agacaccggg gccccggtcc gcccggcccc ccttcggacg gaccggacgg 1320

ctcggaccac ggggacggct cagaccgtcc cgtgtgtccc cgtccggctc ccgtccgccc 1380

catcccgccc ctccaccggc aaggaaggac acgacgccat gcgcgtcctg ctgacctcgt 1440

tcgcacatca cacgcactac tacggcctgg tgcccctggc ctgggcgctg ctcgccgccg 1500

ggcacgaggt gcgggtcgcc agccagcccg cgctcacgga caccatcacc gggtccgggc 1560

tcgccgcggt gccggtcggc accgaccacc tcatccacga gtaccgggtg cggatggcgg 1620

gcgagccgcg cccgaaccat ccggcgatcg ccttcgacga ggcccgtccc gagccgctgg 1680

actgggacca cgccctcggc atcgaggcga tcctcgcccc gtacttccat ctgctcgcca 1740

acaacgactc gatggtcgac gacctcgtcg acttcgcccg gtcctggcag ccggacctgg 1800

tgctgtggga gccgacgacc tacgcgggcg ccgtcgccgc ccaggtcacc ggtgccgcgc 1860

acgcccgggt cctgtggggg cccgacgtga tgggcagcgc ccgccgcaag ttcgtcgcgc 1920

tgcgggaccg gcagccgccc gagcaccgcg aggaccccac cgcggagtgg ctgacgtgga 1980

cgctcgaccg gtacggcgcc tccttcgaag aggagctgct caccggccag ttcacgatcg 2040

acccgacccc gccgagcctg cgcctcgaca cgggcctgcc gaccgtcggg atgcgttatg 2100

ttccgtacaa cggcacgtcg gtcgtgccgg actggctgag tgagccgccc gcgcggcccc 2160

gggtctgcct gaccctcggc gtctccgcgc gtgaggtcct cggcggcgac ggcgtctcgc 2220

agggcgacat cctggaggcg ctcgccgacc tcgacatcga gctcgtcgcc acgctcgacg 2280

cgagtcagcg cgccgagatc cgcaactacc cgaagcacac ccggttcacg gacttcgtgc 2340

cgatgcacgc gctcctgccg agctgctcgg cgatcatcca ccacggcggg gcgggcacct 2400

acgcgaccgc cgtgatcaac gcggtgccgc aggtcatgct cgccgagctg tgggacgcgc 2460

cggtcaaggc gcgggccgtc gccgagcagg gggcggggtt cttcctgccg ccggccgagc 2520

tcacgccgca ggccgtgcgg gacgccgtcg tccgcatcct cgacgacccc tcggtcgcca 2580

ccgccgcgca ccggctgcgc gaggagacct tcggcgaccc caccccggcc gggatcgtcc 2640

ccgagctgga gcggctcgcc gcgcagcacc gccgcccgcc ggccgacgcc cggcactgag 2700

ccgcacccct cgccccaggc ctcacccctg tatctcatat gtctagttaa ctcgccacgc 2760

cgaccgttat caccggcgcc ctgctgctag tttccgagaa tgaagggaat agtcctggcc 2820

ggcgggagcg gaactcggct gcatccggcg acctcggtca tttcgaagca gattcttccg 2880

gtctacaaca aaccgatgat ctactatccg ctgtcggttc tcatgctcgg cggtattcgc 2940

gagattcaaa tcatctcgac cccccagcac atcgaactct tccagtcgct tctcggaaac 3000

ggcaggcacc tgggaataga actcgactat gcggtccaga aagagcccgc aggaatcgcg 3060

gacgcacttc tcgtcggagc cgagcacatc ggcgacgaca cctgcgccct gatcctgggc 3120

gacaacatct tccacgggcc cggcctctac acgctcctgc gggacagcat cgcgcgcctc 3180

gacggctgcg tgctcttcgg ctacccggtc aaggaccccg agcggtacgg cgtcgccgag 3240

gtggacgcga cgggccggct gaccgacctc gtcgagaagc ccgtcaagcc gcgctccaac 3300

ctcgccgtca ccggcctcta cctctacgac aacgacgtcg tcgacatcgc caagaacatc 3360

cggccctcgc cgcgcggcga gctggagatc accgacgtca accgcgtcta cctggagcgg 3420

ggccgggccg aactcgtcaa cctgggccgc ggcttcgcct ggctggacac cggcacccac 3480

gactcgctcc tgcgggccgc ccagtacgtc caggtcctgg aggagcggca gggcgtctgg 3540

atcgcgggcc ttgaggagat cgccttccgc atgggcttca tcgacgccga ggcctgtcac 3600

ggcctgggag aaggcctctc ccgcaccgag tacggcagct atctgatgga gatcgccggc 3660

cgcgagggag ccccgtgagg gcacctcgcg gccgacgcgt tcccacgacc gacagcgcca 3720

ccgacagtgc gacccacacc gcgacccgca ccgccaccga cagtgcgacc cacaccgcga 3780

cctacagcgc gaccgaaagg aagacggcag tgcggcttct ggtgaccgga ggtgcgggct 3840

tcatcggctc gcacttcgtg cggcagctcc tcgccggggc gtaccccgac gtgcccgccg 3900

atgaggtgat cgtcctggac agcctcacct acgcgggcaa ccgcgccaac ctcgccccgg 3960

tggacgcgga cccgcgactg cgcttcgtcc acggcgacat ccgcgacgcc ggcctcctcg 4020

cccgggaact gcgcggcgtg gacgccatcg tccacttcgc ggccgagagc cacgtggacc 4080

gctccatcgc gggcgcgtcc gtgttcaccg agaccaacgt gcagggcacg cagacgctgc 4140

tccagtgcgc cgtcgacgcc ggcgtcggcc gggtcgtgca cgtctccacc gacgaggtgt 4200

acgggtcgat cgactccggc tcctggaccg agagcagccc gctggagccc aactcgccct 4260

acgcggcgtc caaggccggc tccgacctcg ttgcccgcgc ctaccaccgg acgtacggcc 4320

tcgacgtacg gatcacccgc tgctgcaaca actacgggcc gtaccagcac cccgagaagc 4380

tcatccccct cttcgtgacg aacctcctcg acggcgggac gctcccgctg tacggcgacg 4440

gcgcgaacgt ccgcgagtgg gtgcacaccg acgaccactg ccggggcatc gcgctcgtcc 4500
tcgcgggcgg ccgggccggc gagatctacc acatcggcgg cggcctggag ctgaccaacc 4560
gcgaactcac cggcatcctc ctggactcgc tcggcgccga ctggtcctcg gtccggaagg 4620
tcgccgaccg caagggccac gacctgcgct actccctcga cggcggcgag atcgagcgcg 4680
agctcggcta ccgcccgcag gtctccttcg cggacggcct cgcgcggacc gtccgctggt 4740
accgggagaa ccgcggctgg tgggagccgc tcaaggcgac cgccccgcag ctgcccgcca 4800
ccgccgtgga ggtgtccgcg tgagcagccg cgccgagacc ccccgcgtcc ccttcctcga 4860

cctcaaggcc gcctacgagg agctctctag a 4891

<210> 17
<211> 6603
<212> DNA
<213> Artificial Sequence
<220><223> desVIII-desVII-desIII-desIV-oleL-oleU
<400> 17

ttaattaaga tatccccgac cagccttatc gaaggagcgg cgaagagatg gcagaaccca 60
cggtgaccga cgacctgacg ggggccctca cgcagccccc gctgggccgc accgtccgcg 120
cggtggccga ccgtgaactc ggcacccacc tcctggagac ccgcggcatc cactggatcc 180
acgccgcgaa cggcgacccg tacgccaccg tgctgcgcgg ccaggcggac gacccgtatc 240

ccgcgtacga gcgggtgcgt gcccgcggcg cgctctcctt cagcccgacg ggcagctggg 300
tcaccgccga tcacgccctg gcggcgagca tcctctgctc gacggacttc ggggtctccg 360
gcgccgacgg cgtcccggtg ccgcagcagg tcctctcgta cggggagggc tgtccgctgg 420
agcgcgagca ggtgctgccg gcggccggtg acgtgccgga gggcgggcag cgtgccgtgg 480
tcgaggggat ccaccgggag acgctggagg gtctcgcgcc ggacccgtcg gcgtcgtacg 540
ccttcgagct gctgggcggt ttcgtccgcc cggcggtgac ggccgctgcc gccgccgtgc 600
tgggtgttcc cgcggaccgg cgcgcggact tcgcggatct gctggagcgg ctccggccgc 660

tgtccgacag cctgctggcc ccgcagtccc tgcggacggt acgggcggcg gacggcgcgc 720
tggccgagct cacggcgctg ctcgccgatt cggacgactc ccccggggcc ctgctgtcgg 780
cgctcggggt caccgcagcc gtccagctca ccgggaacgc ggtgctcgcg ctcctcgcgc 840
atcccgagca gtggcgggag ctgtgcgacc ggcccgggct cgcggcggcc gcggtggagg 900
agaccctccg ctacgacccg ccggtgcagc tcgacgcccg ggtggtccgc ggggagacgg 960

agctggcggg ccggcggctg ccggccgggg cgcatgtcgt cgtcctgacc gccgcgaccg 1020

gccgggaccc ggaggtcttc acggacccgg agcgcttcga cctcgcgcgc cccgacgccg 1080

ccgcgcacct cgcgctgcac cccgccggtc cgtacggccc ggtggcgtcc ctggtccggc 1140

ttcaggcgga ggtcgcgctg cggaccctgg ccgggcgttt ccccgggctg cggcaggcgg 1200

gggacgtgct ccgcccccgc cgcgcgcctg tcggccgcgg gccgctgagc gtcccggtca 1260

gcagctcctg agacaccggg gccccggtcc gcccggcccc ccttcggacg gaccggacgg 1320

ctcggaccac ggggacggct cagaccgtcc cgtgtgtccc cgtccggctc ccgtccgccc 1380

catcccgccc ctccaccggc aaggaaggac acgacgccat gcgcgtcctg ctgacctcgt 1440

tcgcacatca cacgcactac tacggcctgg tgcccctggc ctgggcgctg ctcgccgccg 1500

ggcacgaggt gcgggtcgcc agccagcccg cgctcacgga caccatcacc gggtccgggc 1560

tcgccgcggt gccggtcggc accgaccacc tcatccacga gtaccgggtg cggatggcgg 1620

gcgagccgcg cccgaaccat ccggcgatcg ccttcgacga ggcccgtccc gagccgctgg 1680

actgggacca cgccctcggc atcgaggcga tcctcgcccc gtacttccat ctgctcgcca 1740

acaacgactc gatggtcgac gacctcgtcg acttcgcccg gtcctggcag ccggacctgg 1800

tgctgtggga gccgacgacc tacgcgggcg ccgtcgccgc ccaggtcacc ggtgccgcgc 1860

acgcccgggt cctgtggggg cccgacgtga tgggcagcgc ccgccgcaag ttcgtcgcgc 1920

tgcgggaccg gcagccgccc gagcaccgcg aggaccccac cgcggagtgg ctgacgtgga 1980

cgctcgaccg gtacggcgcc tccttcgaag aggagctgct caccggccag ttcacgatcg 2040

acccgacccc gccgagcctg cgcctcgaca cgggcctgcc gaccgtcggg atgcgttatg 2100

ttccgtacaa cggcacgtcg gtcgtgccgg actggctgag tgagccgccc gcgcggcccc 2160

gggtctgcct gaccctcggc gtctccgcgc gtgaggtcct cggcggcgac ggcgtctcgc 2220

agggcgacat cctggaggcg ctcgccgacc tcgacatcga gctcgtcgcc acgctcgacg 2280

cgagtcagcg cgccgagatc cgcaactacc cgaagcacac ccggttcacg gacttcgtgc 2340

cgatgcacgc gctcctgccg agctgctcgg cgatcatcca ccacggcggg gcgggcacct 2400

acgcgaccgc cgtgatcaac gcggtgccgc aggtcatgct cgccgagctg tgggacgcgc 2460

cggtcaaggc gcgggccgtc gccgagcagg gggcggggtt cttcctgccg ccggccgagc 2520

tcacgccgca ggccgtgcgg gacgccgtcg tccgcatcct cgacgacccc tcggtcgcca 2580

ccgccgcgca ccggctgcgc gaggagacct tcggcgaccc caccccggcc gggatcgtcc 2640

ccgagctgga gcggctcgcc gcgcagcacc gccgcccgcc ggccgacgcc cggcactgag 2700

ccgcacccct cgccccaggc ctcacccctg tatctcatat gtctagttaa ctcgccacgc 2760

cgaccgttat caccggcgcc ctgctgctag tttccgagaa tgaagggaat agtcctggcc 2820
ggcgggagcg gaactcggct gcatccggcg acctcggtca tttcgaagca gattcttccg 2880
gtctacaaca aaccgatgat ctactatccg ctgtcggttc tcatgctcgg cggtattcgc 2940
gagattcaaa tcatctcgac cccccagcac atcgaactct tccagtcgct tctcggaaac 3000
ggcaggcacc tgggaataga actcgactat gcggtccaga aagagcccgc aggaatcgcg 3060
gacgcacttc tcgtcggagc cgagcacatc ggcgacgaca cctgcgccct gatcctgggc 3120
gacaacatct tccacgggcc cggcctctac acgctcctgc gggacagcat cgcgcgcctc 3180

gacggctgcg tgctcttcgg ctacccggtc aaggaccccg agcggtacgg cgtcgccgag 3240
gtggacgcga cgggccggct gaccgacctc gtcgagaagc ccgtcaagcc gcgctccaac 3300
ctcgccgtca ccggcctcta cctctacgac aacgacgtcg tcgacatcgc caagaacatc 3360
cggccctcgc cgcgcggcga gctggagatc accgacgtca accgcgtcta cctggagcgg 3420
ggccgggccg aactcgtcaa cctgggccgc ggcttcgcct ggctggacac cggcacccac 3480
gactcgctcc tgcgggccgc ccagtacgtc caggtcctgg aggagcggca gggcgtctgg 3540
atcgcgggcc ttgaggagat cgccttccgc atgggcttca tcgacgccga ggcctgtcac 3600

ggcctgggag aaggcctctc ccgcaccgag tacggcagct atctgatgga gatcgccggc 3660
cgcgagggag ccccgtgagg gcacctcgcg gccgacgcgt tcccacgacc gacagcgcca 3720
ccgacagtgc gacccacacc gcgacccgca ccgccaccga cagtgcgacc cacaccgcga 3780
cctacagcgc gaccgaaagg aagacggcag tgcggcttct ggtgaccgga ggtgcgggct 3840
tcatcggctc gcacttcgtg cggcagctcc tcgccggggc gtaccccgac gtgcccgccg 3900
atgaggtgat cgtcctggac agcctcacct acgcgggcaa ccgcgccaac ctcgccccgg 3960
tggacgcgga cccgcgactg cgcttcgtcc acggcgacat ccgcgacgcc ggcctcctcg 4020

cccgggaact gcgcggcgtg gacgccatcg tccacttcgc ggccgagagc cacgtggacc 4080
gctccatcgc gggcgcgtcc gtgttcaccg agaccaacgt gcagggcacg cagacgctgc 4140
tccagtgcgc cgtcgacgcc ggcgtcggcc gggtcgtgca cgtctccacc gacgaggtgt 4200
acgggtcgat cgactccggc tcctggaccg agagcagccc gctggagccc aactcgccct 4260
acgcggcgtc caaggccggc tccgacctcg ttgcccgcgc ctaccaccgg acgtacggcc 4320
tcgacgtacg gatcacccgc tgctgcaaca actacgggcc gtaccagcac cccgagaagc 4380
tcatcccccc cttcgtgacg aacctcctcg acggcgggac gctcccgctg tacggcgacg 4440

gcgcgaacgt ccgcgagtgg gtgcacaccg acgaccactg ccggggcatc gcgctcgtcc 4500

tcgcgggcgg ccgggccggc gagatctacc acatcggcgg cggcctggag ctgaccaacc 4560

gcgaactcac cggcatcctc ctggactcgc tcggcgccga ctggtcctcg gtccggaagg 4620

tcgccgaccg caagggccac gacctgcgct actccctcga cggcggcgag atcgagcgcg 4680

agctcggcta ccgcccgcag gtctccttcg cggacggcct cgcgcggacc gtccgctggt 4740

accgggagaa ccgcggctgg tgggagccgc tcaaggcgac cgccccgcag ctgcccgcca 4800

ccgccgtgga ggtgtccgcg tgagcagccg cgccgagacc ccccgcgtcc ccttcctcga 4860

cctcaaggcc gcctacgagg agctctctag tatcgctccg agcccgaagg gaaaatcgag 4920

tgcccaattc ggcagaatcg ggatcgatgg agttactcga cgtcgacggg gcctggttat 4980

acaccccgga aatcatgcgg gacgaacggg gagaattcct cgaatggttc cggggtcgga 5040

cattccagga gaagatcggc caccccctct cgctggccca ggccaactgc tcggtgtccc 5100

gcaaggcgtt ctgcgcggca tccacttcgc cgacgccccc gcccggccag gccaagtacg 5160

tcacctgcgc ctccggcacc gtgctcgacg tggtcgtcga cgtacgccgg ggctcgccca 5220

ccttcggccg atgggccgcc gtccgactcg acgcggcccg ccaccagggg ctctacctgg 5280

ccgaaggact cggccacgcg ttcatggccc tcaccgacga cgccacggtc gtctacctct 5340

gctcacagcc ctacgtggcc gaggccgagc gggccgtaga ccctctcgac ccggcgatcg 5400

gcatcgaatg gccgacggac atcgacatcg tccctgtcgg cgaagggaca cccacgcacc 5460

gtccctggcg caggccgcgg agacccggca tcctgccgga ctacgaggga gtgcccggcg 5520

ccttacatcg cggaggcggc cggcgtggga ccggcccgtg aaggccctcg tactggccgg 5580

ctctagttac gtagtaccgc gacaaccgcg cctggtggga acccctgaag aagcggccgg 5640

ccggtcccgc cgccccccg agaggcagcg gcccatgaga tggctgatca ccggcgccgc 5700

cggaatgctg ggccgggaac tcgtccggcg cctcgccgag aacgaggagg acgtcgcggc 5760

cctcggccac gaccacctcg acgtcacccg accctccgcc gtgcgggcgg cactcgccga 5820

gcaccgtccc gggatcgtcg tcaactgcgc cgcctacacg gccgtcgacg acgccgagac 5880

ggacgaggcc gccgctgccc tcctcaacgc cgaggcgccc cggctgctgg ccgagggcct 5940

gcgccccccac cggcggcacg gcctcgtcca cctgtccacc gactacgtct ttcccggcga 6000

cgcccgcacc ccctacgccg aggaccaccc cacggctccc cgcagcgcct acggacgcac 6060

caaacgggac ggcgagcaag cggtgctgac ggcactgccc accgccaccg tgctgcgcac 6120

cgcctggctg tacgggcgca ccggccgcag cttcgtccgc acgatgatcg aacgggaggc 6180

gcgcggcgga gccatcgacg tcgtcgccga ccagcgcggc cagcccacct ggaccggcga 6240

cctcgccgac cgcatcatcg ccgtcggccg gcaccccggc gtccacggca tcctgcacgc 6300

caccaacgcc ggctccgcca cctggtacga cctggcacaa gaggtcttcc ggctcctcga 6360

cgccgacccc gggcgggtcc ggcccaccac cggcgccgcc ttccgcagac ccgccccccg 6420

ccccgcctac agcgtcctcg gccacgaccg ctggcgcggg accggcctgg cacccctgcg 6480

tgactggcgc tcggccctgc gcgaggcgtt ccccgacatc ctcgccgcgg aacacccacc 6540

gacccggcga ggagcagcat gaaacgaggc gtgcacgacc tggccctctt ccctaggtct 6600

aga 6603

<210> 18

<211> 7635

<212> DNA

<213> Artificial Sequence

<220><223> desVIII-desVII-desIII-desIV-desI-desII

<400> 18

ttaattaaga tatccccgac cagccttatc gaaggagcgg cgaagagatg gcagaaccca 60

cggtgaccga cgacctgacg ggggccctca cgcagccccc gctgggccgc accgtccgcg 120

cggtggccga ccgtgaactc ggcacccacc tcctggagac ccgcggcatc cactggatcc 180

acgccgcgaa cggcgacccg tacgccaccg tgctgcgcgg ccaggcggac gacccgtatc 240

ccgcgtacga gcgggtgcgt gcccgcggcg cgctctcctt cagcccgacg ggcagctggg 300

tcaccgccga tcacgccctg gcggcgagca tcctctgctc gacggacttc ggggtctccg 360

gcgccgacgg cgtcccggtg ccgcagcagg tcctctcgta cggggagggc tgtccgctgg 420

agcgcgagca ggtgctgccg gcggccggtg acgtgccgga gggcgggcag cgtgccgtgg 480

tcgaggggat ccaccgggag acgctggagg gtctcgcgcc ggacccgtcg gcgtcgtacg 540

ccttcgagct gctgggcggt ttcgtccgcc cggcggtgac ggccgctgcc gccgccgtgc 600

tgggtgttcc cgcggaccgg cgcgcggact tcgcggatct gctggagcgg ctccggccgc 660

tgtccgacag cctgctggcc ccgcagtccc tgcggacggt acgggcggcg gacggcgcgc 720

tggccgagct cacggcgctg ctcgccgatt cggacgactc ccccggggcc ctgctgtcgg 780

cgctcggggt caccgcagcc gtccagctca ccgggaacgc ggtgctcgcg ctcctcgcgc 840

atcccgagca gtggcgggag ctgtgcgacc ggcccgggct cgcggcggcc gcggtggagg 900

agaccctccg ctacgacccg ccggtgcagc tcgacgcccg ggtggtccgc ggggagacgg 960

agctggcggg ccggcggctg ccggccgggg cgcatgtcgt cgtcctgacc gccgcgaccg 1020

gccgggaccc ggaggtcttc acggacccgg agcgcttcga cctcgcgcgc cccgacgccg 1080

ccgcgcacct cgcgctgcac cccgccggtc cgtacggccc ggtggcgtcc ctggtccggc 1140

ttcaggcgga ggtcgcgctg cggaccctgg ccgggcgttt ccccgggctg cggcaggcgg 1200

gggacgtgct ccgcccccgc cgcgcgcctg tcggccgcgg gccgctgagc gtcccggtca 1260

gcagctcctg agacaccggg gccccggtcc gcccggcccc ccttcggacg gaccggacgg 1320

ctcggaccac ggggacggct cagaccgtcc cgtgtgtccc cgtccggctc ccgtccgccc 1380

catcccgccc ctccaccggc aaggaaggac acgacgccat gcgcgtcctg ctgacctcgt 1440

tcgcacatca cacgcactac tacggcctgg tgcccctggc ctgggcgctg ctcgccgccg 1500

ggcacgaggt gcgggtcgcc agccagcccg cgctcacgga caccatcacc gggtccgggc 1560

tcgccgcggt gccggtcggc accgaccacc tcatccacga gtaccgggtg cggatggcgg 1620

gcgagccgcg cccgaaccat ccggcgatcg ccttcgacga ggcccgtccc gagccgctgg 1680

actgggacca cgccctcggc atcgaggcga tcctcgcccc gtacttccat ctgctcgcca 1740

acaacgactc gatggtcgac gacctcgtcg acttcgcccg gtcctggcag ccggacctgg 1800

tgctgtggga gccgacgacc tacgcgggcg ccgtcgccgc ccaggtcacc ggtgccgcgc 1860

acgcccgggt cctgtggggg cccgacgtga tgggcagcgc ccgccgcaag ttcgtcgcgc 1920

tgcgggaccg gcagccgccc gagcaccgcg aggaccccac cgcggagtgg ctgacgtgga 1980

cgctcgaccg gtacggcgcc tccttcgaag aggagctgct caccggccag ttcacgatcg 2040

acccgacccc gccgagcctg cgcctcgaca cgggcctgcc gaccgtcggg atgcgttatg 2100

ttccgtacaa cggcacgtcg gtcgtgccgg actggctgag tgagccgccc gcgcggcccc 2160

gggtctgcct gaccctcggc gtctccgcgc gtgaggtcct cggcggcgac ggcgtctcgc 2220

agggcgacat cctggaggcg ctcgccgacc tcgacatcga gctcgtcgcc acgctcgacg 2280

cgagtcagcg cgccgagatc cgcaactacc cgaagcacac ccggttcacg gacttcgtgc 2340

cgatgcacgc gctcctgccg agctgctcgg cgatcatcca ccacggcggg gcgggcacct 2400

acgcgaccgc cgtgatcaac gcggtgccgc aggtcatgct cgccgagctg tgggacgcgc 2460

cggtcaaggc gcgggccgtc gccgagcagg gggcggggtt cttcctgccg ccggccgagc 2520

tcacgccgca ggccgtgcgg gacgccgtcg tccgcatcct cgacgacccc tcggtcgcca 2580

ccgccgcgca ccggctgcgc gaggagacct tcggcgaccc caccccggcc gggatcgtcc 2640

ccgagctgga gcggctcgcc gcgcagcacc gccgcccgcc ggccgacgcc cggcactgag 2700

ccgcacccct cgccccaggc ctcacccctg tatctcatat gtctagttaa ctcgccacgc 2760

cgaccgttat caccggcgcc ctgctgctag tttccgagaa tgaagggaat agtcctggcc 2820

ggcgggagcg gaactcggct gcatccggcg acctcggtca tttcgaagca gattcttccg 2880

gtctacaaca aaccgatgat ctactatccg ctgtcggttc tcatgctcgg cggtattcgc 2940

gagattcaaa tcatctcgac cccccagcac atcgaactct tccagtcgct tctcggaaac 3000

ggcaggcacc tgggaataga actcgactat gcggtccaga aagagcccgc aggaatcgcg 3060

gacgcacttc tcgtcggagc cgagcacatc ggcgacgaca cctgcgccct gatcctgggc 3120

gacaacatct tccacgggcc cggcctctac acgctcctgc gggacagcat cgcgcgcctc 3180

gacggctgcg tgctcttcgg ctacccggtc aaggaccccg agcggtacgg cgtcgccgag 3240

gtggacgcga cgggccggct gaccgacctc gtcgagaagc ccgtcaagcc gcgctccaac 3300

ctcgccgtca ccggcctcta cctctacgac aacgacgtcg tcgacatcgc caagaacatc 3360

cggccctcgc cgcgcggcga gctggagatc accgacgtca accgcgtcta cctggagcgg 3420

ggccgggccg aactcgtcaa cctgggccgc ggcttcgcct ggctggacac cggcacccac 3480

gactcgctcc tgcgggccgc ccagtacgtc caggtcctgg aggagcggca gggcgtctgg 3540

atcgcgggcc ttgaggagat cgccttccgc atgggcttca tcgacgccga ggcctgtcac 3600

ggcctgggag aaggcctctc ccgcaccgag tacggcagct atctgatgga gatcgccggc 3660

cgcgagggag ccccgtgagg gcacctcgcg gccgacgcgt tcccacgacc gacagcgcca 3720

ccgacagtgc gacccacacc gcgacccgca ccgccaccga cagtgcgacc cacaccgcga 3780

cctacagcgc gaccgaaagg aagacggcag tgcggcttct ggtgaccgga ggtgcgggct 3840

tcatcggctc gcacttcgtg cggcagctcc tcgccggggc gtaccccgac gtgcccgccg 3900

atgaggtgat cgtcctggac agcctcacct acgcgggcaa ccgcgccaac ctcgccccgg 3960

tggacgcgga cccgcgactg cgcttcgtcc acggcgacat ccgcgacgcc ggcctcctcg 4020

cccgggaact gcgcggcgtg gacgccatcg tccacttcgc ggccgagagc cacgtggacc 4080

gctccatcgc gggcgcgtcc gtgttcaccg agaccaacgt gcagggcacg cagacgctgc 4140

tccagtgcgc cgtcgacgcc ggcgtcggcc gggtcgtgca cgtctccacc gacgaggtgt 4200

acgggtcgat cgactccggc tcctggaccg agagcagccc gctggagccc aactcgccct 4260

acgcggcgtc caaggccggc tccgacctcg ttgcccgcgc ctaccaccgg acgtacggcc 4320

tcgacgtacg gatcacccgc tgctgcaaca actacgggcc gtaccagcac cccgagaagc 4380

tcatccccct cttcgtgacg aacctcctcg acggcgggac gctcccgctg tacggcgacg 4440

gcgcgaacgt ccgcgagtgg gtgcacaccg acgaccactg ccggggcatc gcgctcgtcc 4500

tcgcgggcgg ccgggccggc gagatctacc acatcggcgg cggcctggag ctgaccaacc 4560

gcgaactcac cggcatcctc ctggactcgc tcggcgccga ctggtcctcg gtccggaagg 4620

tcgccgaccg caagggccac gacctgcgct actccctcga cggcggcgag atcgagcgcg 4680

agctcggcta ccgcccgcag gtctccttcg cggacggcct cgcgcggacc gtccgctggt 4740

accgggagaa ccgcggctgg tgggagccgc tcaaggcgac cgccccgcag ctgcccgcca 4800

ccgccgtgga ggtgtccgcg tgagcagccg cgccgagacc ccccgcgtcc ccttcctcga 4860
cctcaaggcc gcctacgagg agctctctag tgctagcgac ggtggcccgg agggaacatc 4920
cgtgaaaagc gccttatccg acctcgcatt cttcggcggc cccgccgctt tcgaccagcc 4980
gctcctcgtg gggcggccca accgcatcga ccgcgccagg ctgtacgagc ggctcgaccg 5040
ggccctcgac agccagtggc tgtccaacgg cggcccgctc gtccgcgagt tcgaggagcg 5100
cgtcgccggg ctcgccgggg tccggcatgc cgtggccacc tgcaacgcca cggccgggct 5160
ccagctcctc gcgcacgccg ccggcctcac cggcgaagtg atcatgccgt cgatgacgtt 5220

cgccgccacc ccgcacgcac tgcgctggat cggcctcacc ccggtcttcg ccgacatcga 5280
cccggacacc ggcaacctcg acccggacca ggtggccgcc gcggtcacac cccgcacctc 5340
ggccgtcgtc ggcgtccacc tctggggccg cccctgcgcc gccgaccagc tgcggaaggt 5400
cgccgacgag cacggcctgc ggctgtactt cgacgccgcg cacgccctcg gctgcgcggt 5460
cgacggccgg cccgccggca gcctcggcga cgccgaggtc ttcagcttcc acgccaccaa 5520
ggccgtcaac gccttcgagg gcggcgccgt cgtcaccgac gacgccgacc tcgccgcccg 5580
gatccgcgcc ctccacaact tcggcttcga cctgcccggc ggcagccccg ccggcgggac 5640

caacgccaag atgagcgagg ccgccgccgc catgggcctc acctccctcg acgcgtttcc 5700
cgaggtcatc gaccggaacc ggcgcaacca cgccgcctac cgcgagcacc tcgcggacct 5760
ccccggcgtc ctcgtcgccg accacgaccg ccacggcctc aacaaccacc agtacgtgat 5820
cgtcgagatc gacgaggcca ccaccggcat ccaccgcgac ctcgtcatgg aggtcctgaa 5880
ggccgaaggc gtgcacaccc gcgcctactt ctcgccgggc tgccacgagc tggagccgta 5940
ccgcgggcag ccgcacgccc cgctgccgca caccgaacgc ctcgccgcgc gcgtgctgtc 6000
cctgccgacc ggcaccgcca tcggcgacga cgacatccgc cgggtcgccg acctgctgcg 6060

tctctgcgcg acccgcggcc gcgaactgac cgcgcgccac cgcgacacgg cccccgcccc 6120
gctcgcggcc ccccagacat ccacgcccac gattggacgc tcccgatgac cgcccccgcc 6180
ctttccgcca ccgccccggc cgaacgctgc gcgcaccccg gagccgatct gggggcggcg 6240
gtccacgccg tcggccagac cctcgccgcc ggcggcctcg tgccgcccga cgaggccgga 6300
acgaccgccc gccacctcgt ccggctcgcc gtgcgctacg gcaacagccc cttcaccccg 6360
ctggaggagg cccgccacga cctgggcgtc gaccgggacg ccttccggcg cctcctcgcc 6420
ctgttcgggc aggtcccgga gctccgcacc gcggtcgaga ccggccccgc cggggcgtac 6480

tggaagaaca ccctgctccc gctcgaacag cgcggcgtct tcgacgcggc gctcgccagg 6540

aagcccgtct tcccgtacag cgtcggcctc taccccggcc cgacctgcat gttccgctgc 6600

cacttctgcg tccgtgtgac cggcgcccgc tacgacccgt ccgccctcga cgccggcaac 6660

gccatgttcc ggtcggtcat cgacgagata cccgcgggca acccctcggc gatgtacttc 6720

tccggcggcc tggagccgct caccaacccc ggcctcggga gcctggccgc gcacgccacc 6780

gaccacggcc tgcggcccac cgtctacacg aactccttcg cgctcaccga gcgcaccctg 6840

gagcgccagc ccggcctctg gggcctgcac gccatccgca cctcgctcta cggcctcaac 6900

gacgaggagt acgagcagac caccggcaag aaggccgcct tccgccgcgt ccgcgagaac 6960

ctgcgccgct tccagcagct gcgcgccgag cgcgagtcgc cgatcaacct cggcttcgcc 7020

tacatcgtgc tcccgggccg tgcctcccgc ctgctcgacc tggtcgactt catcgccgac 7080

ctcaacgacg ccgggcaggg caggacgatc gacttcgtca acattcgcga ggactacagc 7140

ggccgtgacg acggcaagct gccgcaggag gagcgggccg agctccagga ggccctcaac 7200

gccttcgagg agcgggtccg cgagcgcacc cccggactcc acatcgacta cggctacgcc 7260

ctgaacagcc tgcgcaccgg ggccgacgcc gaactgctgc ggatcaagcc cgccaccatg 7320

cggcccaccg cgcacccgca ggtcgcggtg caggtcgatc tcctcggcga cgtgtacctg 7380

taccgcgagg ccggcttccc cgacctggac ggcgcgaccc gctacatcgc gggccgcgtg 7440

acccccgaca cctccctcac cgaggtcgtc agggacttcg tcgagcgcgg cggcgaggtg 7500

gcggccgtcg acggcgacga gtacttcatg gacggcttcg atcaggtcgt caccgcccgc 7560

ctgaaccagc tggagcgcga cgccgcggac ggctgggagg aggcccgcgg cttcctgcgc 7620

tgacccgcat ctaga 7635

<210> 19

<211> 8375

<212> DNA

<213> Artificial Sequence

<220><223> desVIII-desVII-desIII-desIV-oleV-eryBII-urdR

<400> 19

ttaattaaga tatccccgac cagccttatc gaaggagcgg cgaagagatg gcagaaccca 60

cggtgaccga cgacctgacg ggggccctca cgcagccccc gctgggccgc accgtccgcg 120

cggtggccga ccgtgaactc ggcacccacc tttaattaag atatccccga ccagccttat 180

cgaaggagcg gcgaagagat ggcagaaccc acggtgaccg acgacctgac gggggccctc 240

acgcagcccc cgctgggccg caccgtccgc gcggtggccg accgtgaact cggcacccac 300

ctcctggaga cccgcggcat ccactggatc cacgccgcga acggcgaccc gtacgccacc 360

gtgctgcgcg gccaggcgga cgacccgtat cccgcgtacg agcgggtgcg tgcccgcggc 420
gcgctctcct tcagcccgac gggcagctgg gtcaccgccg atcacgccct ggcggcgagc 480
atcctctgct cgacggactt cggggtctcc ggcgccgacg gcgtcccggt gccgcagcag 540
gtcctctcgt acggggaggg ctgtccgctg gagcgcgagc aggtgctgcc ggcggccggt 600
gacgtgccgg agggcgggca gcgtgccgtg gtcgagggga tccaccggga gacgctggag 660
ggtctcgcgc cggacccgtc ggcgtcgtac gccttcgagc tgctgggcgg tttcgtccgc 720
ccggcggtga cggccgctgc cgccgccgtg ctgggtgttc ccgcggaccg gcgcgcggac 780

ttcgcggatc tgctggagcg gctccggccg ctgtccgaca gcctgctggc cccgcagtcc 840
ctgcggacgg tacgggcggc ggacggcgcg ctggccgagc tcacggcgct gctcgccgat 900
tcggacgact ccccccggggc cctgctgtcg gcgctcgggg tcaccgcagc cgtccagctc 960
accgggaacg cggtgctcgc gctcctcgcg catcccgagc agtggcggga gctgtgcgac 1020
cggcccgggc tcgcggcggc cgcggtggag gagaccctcc gctacgaccc gccggtgcag 1080
ctcgacgccc gggtggtccg cggggagacg gagctggcgg gccggcggct gccggccggg 1140
gcgcatgtcg tcgtcctgac cgccgcgacc ggccgggacc cggaggtctt cacggacccg 1200

gagcgcttcg acctcgcgcg ccccgacgcc gccgcgcacc tcgcgctgca ccccgccggt 1260
ccgtacggcc cggtggcgtc cctggtccgg cttcaggcgg aggtcgcgct gcggaccctg 1320
gccgggcgtt tccccgggct gcggcaggcg ggggacgtgc tccgcccccg ccgcgcgcct 1380
gtcggccgcg ggccgctgag cgtcccggtc agcagctcct gagacaccgg ggccccggtc 1440
cgcccggccc cccttcggac ggaccggacg gctcggacca cggggacggc tcagaccgtc 1500
ccgtgtgtcc ccgtccggct cccgtccgcc ccatcccgcc cctccaccgg caaggaagga 1560
cacgacgcca tgcgcgtcct gctgacctcg ttcgcacatc acacgcacta ctacggcctg 1620

gtgcccctgg cctgggcgct gctcgccgcc gggcacgagg tgcgggtcgc cagccagccc 1680
gcgctcacgg acaccatcac cgggtccggg ctcgccgcgg tgccggtcgg caccgaccac 1740
ctcatccacg agtaccgggt gcggatggcg ggcgagccgc gcccgaacca tccggcgatc 1800
gccttcgacg aggcccgtcc cgagccgctg gactgggacc acgccctcgg catcgaggcg 1860
atcctcgccc cgtacttcca tctgctcgcc aacaacgact cgatggtcga cgacctcgtc 1920
gacttcgccc ggtcctggca gccggacctg gtgctgtggg agccgacgac ctacgcgggc 1980
gccgtcgccg cccaggtcac cggtgccgcg cacgcccggg tcctgtgggg gcccgacgtg 2040

atgggcagcg cccgccgcaa gttcgtcgcg ctgcgggacc ggcagccgcc cgagcaccgc 2100

gaggacccca ccgcggagtg gctgacgtgg acgctcgacc ggtacggcgc ctccttcgaa 2160

gaggagctgc tcaccggcca gttcacgatc gacccgaccc cgccgagcct gcgcctcgac 2220

acgggcctgc cgaccgtcgg gatgcgttat gttccgtaca acggcacgtc ggtcgtgccg 2280

gactggctga gtgagccgcc cgcgcggccc cgggtctgcc tgaccctcgg cgtctccgcg 2340

cgtgaggtcc tcggcggcga cggcgtctcg cagggcgaca tcctggaggc gctcgccgac 2400

ctcgacatcg agctcgtcgc cacgctcgac gcgagtcagc gcgccgagat ccgcaactac 2460

ccgaagcaca cccggttcac ggacttcgtg ccgatgcacg cgctcctgcc gagctgctcg 2520

gcgatcatcc accacggcgg ggcgggcacc tacgcgaccg ccgtgatcaa cgcggtgccg 2580

caggtcatgc tcgccgagct gtgggacgcg ccggtcaagg cgcgggccgt cgccgagcag 2640

ggggcggggt tcttcctgcc gccggccgag ctcacgccgc aggccgtgcg ggacgccgtc 2700

gtccgcatcc tcgacgaccc ctcggtcgcc accgccgcgc accggctgcg cgaggagacc 2760

ttcggcgacc ccaccccggc cgggatcgtc cccgagctgg agcggctcgc cgcgcagcac 2820

cgccgcccgc cggccgacgc ccggcactga gccgcacccc tcgccccagg cctcaccgct 2880

gtatctcata tgtctagtta actcgccacg ccgaccgtta tcaccggcgc cctgctgcta 2940

gtttccgaga atgaagggaa tagtcctggc cggcgggagc ggaactcggc tgcatccggc 3000

gacctcggtc atttcgaagc agattcttcc ggtctacaac aaaccgatga tctactatcc 3060

gctgtcggtt ctcatgctcg gcggtattcg cgagattcaa atcatctcga ccccccagca 3120

catcgaactc ttccagtcgc ttctcggaaa cggcaggcac ctgggaatag aactcgacta 3180

tgcggtccag aaagagcccg caggaatcgc ggacgcactt ctcgtcggag ccgagcacat 3240

cggcgacgac acctgcgccc tgatcctggg cgacaacatc ttccacgggc ccggcctcta 3300

cacgctcctg cgggacagca tcgcgcgcct cgacggctgc gtgctcttcg gctacccggt 3360

caaggacccc gagcggtacg gcgtcgccga ggtggacgcg acgggccggc tgaccgacct 3420

cgtcgagaag cccgtcaagc cgcgctccaa cctcgccgtc accggcctct acctctacga 3480

caacgacgtc gtcgacatcg ccaagaacat ccggccctcg ccgcgcggcg agctggagat 3540

caccgacgtc aaccgcgtct acctggagcg gggccgggcc gaactcgtca acctgggccg 3600

cggcttcgcc tggctggaca ccggcaccca cgactcgctc ctgcgggccg cccagtacgt 3660

ccaggtcctg gaggagcggc agggcgtctg gatcgcgggc cttgaggaga tcgccttccg 3720

catgggcttc atcgacgccg aggcctgtca cggcctggga gaaggcctct cccgcaccga 3780

gtacggcagc tatctgatgg agatcgccgg ccgcgaggga gccccgtgag ggcacctcgc 3840

ggccgacgcg ttcccacgac cgacagcgcc accgacagtg cgacccacac cgcgacccgc 3900

accgccaccg acagtgcgac ccacaccgcg acctacagcg cgaccgaaag gaagacggca 3960

gtgcggcttc tggtgaccgg aggtgcgggc ttcatcggct cgcacttcgt gcggcagctc 4020

ctcgccgggg cgtaccccga cgtgcccgcc gatgaggtga tcgtcctgga cagcctcacc 4080

tacgcgggca accgcgccaa cctcgccccg gtggacgcgg acccgcgact gcgcttcgtc 4140

cacggcgaca tccgcgacgc cggcctcctc gcccgggaac tgcgcggcgt ggacgccatc 4200

gtccacttcg cggccgagag ccacgtggac cgctccatcg cgggcgcgtc cgtgttcacc 4260

gagaccaacg tgcagggcac gcagacgctg ctccagtgcg ccgtcgacgc cggcgtcggc 4320

cgggtcgtgc acgtctccac cgacgaggtg tacgggtcga tcgactccgg ctcctggacc 4380

gagagcagcc cgctggagcc caactcgccc tacgcggcgt ccaaggccgg ctccgacctc 4440

gttgcccgcg cctaccaccg gacgtacggc ctcgacgtac ggatcacccg ctgctgcaac 4500

aactacgggc cgtaccagca ccccgagaag ctcatccccc tcttcgtgac gaacctcctc 4560

gacggcggga cgctcccgct gtacggcgac ggcgcgaacg tccgcgagtg ggtgcacacc 4620

gacgaccact gccggggcat cgcgctcgtc ctcgcgggcg gccgggccgg cgagatctac 4680

cacatcggcg gcggcctgga gctgaccaac cgcgaactca ccggcatcct cctggactcg 4740

ctcggcgccg actggtcctc ggtccggaag gtcgccgacc gcaagggcca cgacctgcgc 4800

tactccctcg acggcggcga gatcgagcgc gagctcggct accgcccgca ggtctccttc 4860

gcggacggcc tcgcgcggac cgtccgctgg taccgggaga accgcggctg gtgggagccg 4920

ctcaaggcga ccgccccgca gctgcccgcc accgccgtgg aggtgtccgc gtgagcagcc 4980

gcgccgagac cccccgcgtc cccttcctcg acctcaaggc cgcctacgag gagctctcta 5040

gtgggaatcg cggaagcggc cgggttcggc gccgtcctgg atcacaatga tatggggaat 5100

tcccgcgatg agcgaagcaa tgggatcggt accgacggcc ggcagtgaag tctcctcgac 5160

ctgcgcgttt ctgtcctggt tggacgcgcg ccgccgggcc aatcgcctga cggtggaaca 5220

cgtaccgttc agggagttat cggggtggca attcgacgag aacacgggga acctccgaca 5280

taccagcggt cgtttcttct ccatcgaagg actccgggta cgcacggacc actgctggtt 5340

cggaagctgg acccagccca ttatcgtgca accggagata ggcattctcg gcctcctggt 5400

caagaggttc gacggcatcc tgcacgtcct ggtgcaggca aagaacgaac cgggtaatat 5460

cggcggcctt cagctctccc ccaccgtcca ggccactcgc agcaattaca cccgcgtcca 5520

ccgcggcggc ggtgtcagat acctggagta cttcgcgtcc ccccgcgggc gcggtcgggt 5580

cctcgcggac gtcctccaat ccgagcaggg ctcctggttc ctgcacaagc ggaaccggaa 5640

catggtggtc gaggccctgg acgacgtgcc cctcgacgac gacttccact ggatcagcct 5700

cggcgggctg cggaagctgc tgctgaggcc gcatctggtg aacatggaca cccgcacggt 5760

gctgtcctgc cttcccccgg atccggcacc ggacggccgg cagccgccgg cgcccgcggc 5820

tcccttcgcc gccgcggtca cgcggtccct cacccggggt gccaccgcct tgcacaccat 5880
gggcgagatc ctcggctggc tgaccgacga gcggtcccgg cgggaactgg tgcagcagcg 5940
ggtgccgctg gaggagaccg cgttcagcgg ctggcggcgt gacgaccacg ccatcgcgca 6000
caaggacggc gactacttcc gggtcatcgg ggtgagcgtc cgggccagca gccgggaggt 6060
gtcctcgtgg agccagccgt tgctggcccc cgtcggcccc ggtctggccg ccttcgtcac 6120
caggcggatc cgcggcgtcc tccacgtcct gctgcacgcc cgcaccgaag ccggcctgct 6180
caacggcccg gagatggcac cgaccgtgca gtgccgcccg ctcaactacc gtgcggtgcc 6240

cgccgagtac cggcccgcct acctcgacta cgtgctgtcg gccgatcccg gacgcatccg 6300
ctacgacacc ctccagtcgg aggagggcgg ccgcttccac cacgcggaga accgctacgt 6360
cgtggtggag gccgaggacg acttcccggt cgaggtgccc cgcgacttcc gctggctcac 6420
cctgcaccag atcctcgccc tgctgcacca cagcaactat gtcaacgtgg aggcgcgcag 6480
cctcgtcgcc tgcatccagc ctgagctgat ctagtggtcg tcggcatctg cgaggaactg 6540
gccgcaggaa ggagagaacc acgatgacca ccgacgccgc gacgcacgtg cggctcgggc 6600
gttccgcgct gctcaccagc aggctctggc tcggcacggt gaacttcagc ggacgcgtcg 6660

aggacgacga cgcgctgcgc ctgatggacc acgcccggga ccgcggcatc aactgcctcg 6720
acaccgccga catgtacggc tggcggctct acaagggcca caccgaggag ctggtgggca 6780
ggtggctggc ccagggcggc ggacggcgcg aggacaccgt gctggcgacc aaggtcggcg 6840
gcgagatgag cgagcgcgtc aacgacagcg ggctgtcggc gcggcacatc atcgcctcct 6900
gcgagggatc gctgcgcagg ctgggcgtcg accacatcga cgtctaccag atgcaccaca 6960
tcgaccggtc cgcgccgtgg gacgaggtgt ggcaggccat ggacagcctc gtcgccagcg 7020
gcaaggtctc ctacgtcggc tcgtcgaact tcgcgggctg gcacatcgcc gccgcgcagg 7080

agaacgccgc ccgccgccac tccctgggca tggtctccca ccagtgcctg tacaacctgg 7140
cggtccggca cgccgagctg gaggtgctgc ccgccgcgca ggcctacggg ctcggcgtct 7200
tcgcctggtc gccgctgcac ggcggcctgc tcagcggagc gctggagaag ctggccgcgg 7260
gcaccgcggt gaagtcggcg cagggccgtg cgcaggtgct gttgccgtcc ctgcgcccgg 7320
cgatcgaggc ctacgagaag ttctgccgca acctcggcga agacccggcc gaggtggggc 7380
tcgcatgggt gctgtcccgg cccggcatcg ccggcgccgt catcggcccg cgaacccccg 7440
agcagctcga ctccgcgctg aaggcgtccg cgatgaccct ggacgagcag gcgctgtccg 7500

aactggacga gatcttcccc gcggtggcct ccggcggcgc ggcgccggaa gcctggttgc 7560

agtgagcaca agaggaaccg agtacgtatc tagtgatatc acagagatca ggacgacgca 7620

tggacattgt gggaaatgga ttcctggcgc ggaacctgcg gccgctggcc gaacggcatc 7680

ccgataccgt ggccctcgcc gccggtgtct catgggcgag cggtacctcc gatgccgact 7740

tcgcccggga agccgcactc ctgcgagaca ccgccaagca gtgcgcggcc accggacgac 7800

gactgctgtt cttctcgacc gctgccacgg gcatgtacgg cctggcggag ggacccggcc 7860

gagaggacac gccagtgacg ccgtgcaccc cgtacggcgc gcacaagctg gcgctggagg 7920

aactgctgcg ggactccggt gccgaccatg tgattctcag gctcgggcac ctcgtggggc 7980

ccgaccagcc cgagcaccaa ctgctgccca cgctggtgcg tcacctacgc gaaggcgcgg 8040

tgcgcattca ccgaggcgcg gcccgtgacc tcatcgacgt cagcgacgtc gtcaccattg 8100

tcgactgcct tctcggcctt gacctcaagg ccgagacggt caacgtggcc tccggctacg 8160

ccgttccggt gaaggacatc gtcgaactcc tgcgccgcag gctcggggtc gaggcgcggt 8220

acgagttcca ggatgccggt ggtcagcacg tcatctccat cgagaagctg cgcgcactcg 8280

ttccgcaggt ccaggacatg ggcttcgggc ccgactacta ccggcggatc ctcagtgact 8340

tcacctccgc cgtccgtatc tgacctaggt ctaga 8375

<210> 20

<211> 9085

<212> DNA

<213> Artificial Sequence

<220><223> desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU

<400> 20

ttaattaaga tatccccgac cagccttatc gaaggagcgg cgaagagatg gcagaaccca 60

cggtgaccga cgacctgacg ggggccctca cgcagccccc gctgggccgc accgtccgcg 120

cggtggccga ccgtgaactc ggcacccacc tcctggagac ccgcggcatc cactggatcc 180

acgccgcgaa cggcgacccg tacgccaccg tgctgcgcgg ccaggcggac gacccgtatc 240

ccgcgtacga gcgggtgcgt gcccgcggcg cgctctcctt cagcccgacg ggcagctggg 300

tcaccgccga tcacgccctg gcggcgagca tcctctgctc gacggacttc ggggtctccg 360

gcgccgacgg cgtcccggtg ccgcagcagg tcctctcgta cggggagggc tgtccgctgg 420

agcgcgagca ggtgctgccg gcggccggtg acgtgccgga gggcgggcag cgtgccgtgg 480

tcgaggggat ccaccgggag acgctggagg gtctcgcgcc ggacccgtcg gcgtcgtacg 540

ccttcgagct gctgggcggt ttcgtccgcc cggcggtgac ggccgctgcc gccgccgtgc 600

tgggtgttcc cgcggaccgg cgcgcggact tcgcggatct gctggagcgg ctccggccgc 660

tgtccgacag cctgctggcc ccgcagtccc tgcggacggt acgggcggcg gacggcgcgc 720

tggccgagct cacggcgctg ctcgccgatt cggacgactc ccccggggcc ctgctgtcgg 780

cgctcggggt caccgcagcc gtccagctca ccgggaacgc ggtgctcgcg ctcctcgcgc 840

atcccgagca gtggcgggag ctgtgcgacc ggcccgggct cgcggcggcc gcggtggagg 900

agaccctccg ctacgacccg ccggtgcagc tcgacgcccg ggtggtccgc ggggagacgg 960

agctggcggg ccggcggctg ccggccgggg cgcatgtcgt cgtcctgacc gccgcgaccg 1020

gccgggaccc ggaggtcttc acggacccgg agcgcttcga cctcgcgcgc cccgacgccg 1080

ccgcgcacct cgcgctgcac cccgccggtc cgtacggccc ggtggcgtcc ctggtccggc 1140

ttcaggcgga ggtcgcgctg cggaccctgg ccgggcgttt ccccgggctg cggcaggcgg 1200

gggacgtgct ccgcccccgc cgcgcgcctg tcggccgcgg gccgctgagc gtcccggtca 1260

gcagctcctg agacaccggg gccccggtcc gcccggcccc ccttcggacg gaccggacgg 1320

ctcggaccac ggggacggct cagaccgtcc cgtgtgtccc cgtccggctc ccgtccgccc 1380

catcccgccc ctccaccggc aaggaaggac acgacgccat gcgcgtcctg ctgacctcgt 1440

tcgcacatca cacgcactac tacggcctgg tgcccctggc ctgggcgctg ctcgccgccg 1500

ggcacgaggt gcgggtcgcc agccagcccg cgctcacgga caccatcacc gggtccgggc 1560

tcgccgcggt gccggtcggc accgaccacc tcatccacga gtaccgggtg cggatggcgg 1620

gcgagccgcg cccgaaccat ccggcgatcg ccttcgacga ggcccgtccc gagccgctgg 1680

actgggacca cgccctcggc atcgaggcga tcctcgcccc gtacttccat ctgctcgcca 1740

acaacgactc gatggtcgac gacctcgtcg acttcgcccg gtcctggcag ccggacctgg 1800

tgctgtggga gccgacgacc tacgcgggcg ccgtcgccgc ccaggtcacc ggtgccgcgc 1860

acgcccgggt cctgtggggg cccgacgtga tgggcagcgc ccgccgcaag ttcgtcgcgc 1920

tgcgggaccg gcagccgccc gagcaccgcg aggaccccac cgcggagtgg ctgacgtgga 1980

cgctcgaccg gtacggcgcc tccttcgaag aggagctgct caccggccag ttcacgatcg 2040

acccgacccc gccgagcctg cgcctcgaca cgggcctgcc gaccgtcggg atgcgttatg 2100

ttccgtacaa cggcacgtcg gtcgtgccgg actggctgag tgagccgccc gcgcggcccc 2160

gggtctgcct gaccctcggc gtctccgcgc gtgaggtcct cggcggcgac ggcgtctcgc 2220

agggcgacat cctggaggcg ctcgccgacc tcgacatcga gctcgtcgcc acgctcgacg 2280

cgagtcagcg cgccgagatc cgcaactacc cgaagcacac ccggttcacg gacttcgtgc 2340

cgatgcacgc gctcctgccg agctgctcgg cgatcatcca ccacggcggg gcgggcacct 2400

acgcgaccgc cgtgatcaac gcggtgccgc aggtcatgct cgccgagctg tgggacgcgc 2460

cggtcaaggc gcgggccgtc gccgagcagg gggcggggtt cttcctgccg ccggccgagc 2520

tcacgccgca ggccgtgcgg gacgccgtcg tccgcatcct cgacgacccc tcggtcgcca 2580

ccgccgcgca ccggctgcgc gaggagacct tcggcgaccc caccccggcc gggatcgtcc 2640

ccgagctgga gcggctcgcc gcgcagcacc gccgcccgcc ggccgacgcc cggcactgag 2700

ccgcacccct cgccccaggc ctcacccctg tatctcatat gtctagttaa ctcgccacgc 2760

cgaccgttat caccggcgcc ctgctgctag tttccgagaa tgaagggaat agtcctggcc 2820

ggcgggagcg gaactcggct gcatccggcg acctcggtca tttcgaagca gattcttccg 2880

gtctacaaca aaccgatgat ctactatccg ctgtcggttc tcatgctcgg cggtattcgc 2940

gagattcaaa tcatctcgac cccccagcac atcgaactct tccagtcgct tctcggaaac 3000

ggcaggcacc tgggaataga actcgactat gcggtccaga aagagcccgc aggaatcgcg 3060

gacgcacttc tcgtcggagc cgagcacatc ggcgacgaca cctgcgccct gatcctgggc 3120

gacaacatct tccacgggcc cggcctctac acgctcctgc gggacagcat cgcgcgcctc 3180

gacggctgcg tgctcttcgg ctacccggtc aaggaccccg agcggtacgg cgtcgccgag 3240

gtggacgcga cgggccggct gaccgacctc gtcgagaagc ccgtcaagcc gcgctccaac 3300

ctcgccgtca ccggcctcta cctctacgac aacgacgtcg tcgacatcgc caagaacatc 3360

cggccctcgc cgcgcggcga gctggagatc accgacgtca accgcgtcta cctggagcgg 3420

ggccgggccg aactcgtcaa cctgggccgc ggcttcgcct ggctggacac cggcacccac 3480

gactcgctcc tgcgggccgc ccagtacgtc caggtcctgg aggagcggca gggcgtctgg 3540

atcgcgggcc ttgaggagat cgccttccgc atgggcttca tcgacgccga ggcctgtcac 3600

ggcctgggag aaggcctctc ccgcaccgag tacggcagct atctgatgga gatcgccggc 3660

cgcgagggag ccccgtgagg gcacctcgcg gccgacgcgt tcccacgacc gacagcgcca 3720

ccgacagtgc gacccacacc gcgacccgca ccgccaccga cagtgcgacc cacaccgcga 3780

cctacagcgc gaccgaaagg aagacggcag tgcggcttct ggtgaccgga ggtgcgggct 3840

tcatcggctc gcacttcgtg cggcagctcc tcgccggggc gtaccccgac gtgcccgccg 3900

atgaggtgat cgtcctggac agcctcacct acgcgggcaa ccgcgccaac ctcgccccgg 3960

tggacgcgga cccgcgactg cgcttcgtcc acggcgacat ccgcgacgcc ggcctcctcg 4020

cccgggaact gcgcggcgtg gacgccatcg tccacttcgc ggccgagagc cacgtggacc 4080

gctccatcgc gggcgcgtcc gtgttcaccg agaccaacgt gcagggcacg cagacgctgc 4140

tccagtgcgc cgtcgacgcc ggcgtcggcc gggtcgtgca cgtctccacc gacgaggtgt 4200

acgggtcgat cgactccggc tcctggaccg agagcagccc gctggagccc aactcgccct 4260

acgcggcgtc caaggccggc tccgacctcg ttgcccgcgc ctaccaccgg acgtacggcc 4320

tcgacgtacg gatcacccgc tgctgcaaca actacgggcc gtaccagcac cccgagaagc 4380

tcatccccct cttcgtgacg aacctcctcg acggcgggac gctcccgctg tacggcgacg 4440

gcgcgaacgt ccgcgagtgg gtgcacaccg acgaccactg ccggggcatc gcgctcgtcc 4500

tcgcgggcgg ccgggccggc gagatctacc acatcggcgg cggcctggag ctgaccaacc 4560

gcgaactcac cggcatcctc ctggactcgc tcggcgccga ctggtcctcg gtccggaagg 4620

tcgccgaccg caagggccac gacctgcgct actccctcga cggcggcgag atcgagcgcg 4680

agctcggcta ccgcccgcag gtctccttcg cggacggcct cgcgcggacc gtccgctggt 4740

accgggagaa ccgcggctgg tgggagccgc tcaaggcgac cgccccgcag ctgcccgcca 4800

ccgccgtgga ggtgtccgcg tgagcagccg cgccgagacc ccccgcgtcc ccttcctcga 4860

cctcaaggcc gcctacgagg agctctctag tgggaatcgc ggaagcggcc gggttcggcg 4920

ccgtcctgga tcacaatgat atggggaatt cccgcgatga gcgaagcaat gggatcggta 4980

ccgacggccg gcagtgaagt ctcctcgacc tgcgcgtttc tgtcctggtt ggacgcgcgc 5040

cgccgggcca atcgcctgac ggtggaacac gtaccgttca gggagttatc ggggtggcaa 5100

ttcgacgaga acacggggaa cctccgacat accagcggtc gtttcttctc catcgaagga 5160

ctccgggtac gcacggacca ctgctggttc ggaagctgga cccagcccat tatcgtgcaa 5220

ccggagatag gcattctcgg cctcctggtc aagaggttcg acggcatcct gcacgtcctg 5280

gtgcaggcaa agaacgaacc gggtaatatc ggcggccttc agctctcccc caccgtccag 5340

gccactcgca gcaattacac ccgcgtccac cgcggcggcg gtgtcagata cctggagtac 5400

ttcgcgtccc cccgcgggcg cggtcgggtc ctcgcggacg tcctccaatc cgagcagggc 5460

tcctggttcc tgcacaagcg gaaccggaac atggtggtcg aggccctgga cgacgtgccc 5520

ctcgacgacg acttccactg gatcagcctc ggcgggctgc ggaagctgct gctgaggccg 5580

catctggtga acatggacac ccgcacggtg ctgtcctgcc ttcccccgga tccggcaccg 5640

gacggccggc agccgccggc gcccgcggct cccttcgccg ccgcggtcac gcggtccctc 5700

acccggggtg ccaccgcctt gcacaccatg ggcgagatcc tcggctggct gaccgacgag 5760

cggtcccggc gggaactggt gcagcagcgg gtgccgctgg aggagaccgc gttcagcggc 5820

tggcggcgtg acgaccacgc catcgcgcac aaggacggcg actacttccg ggtcatcggg 5880

gtgagcgtcc gggccagcag ccgggaggtg tcctcgtgga gccagccgtt gctggccccc 5940

gtcggccccg gtctggccgc cttcgtcacc aggcggatcc gcggcgtcct ccacgtcctg 6000

ctgcacgccc gcaccgaagc cggcctgctc aacggcccgg agatggcacc gaccgtgcag 6060

tgccgcccgc tcaactaccg tgcggtgccc gccgagtacc ggcccgccta cctcgactac 6120

gtgctgtcgg ccgatcccgg acgcatccgc tacgacaccc tccagtcgga ggagggcggc 6180
cgcttccacc acgcggagaa ccgctacgtc gtggtggagg ccgaggacga cttcccggtc 6240
gaggtgcccc gcgacttccg ctggctcacc ctgcaccaga tcctcgccct gctgcaccac 6300
agcaactatg tcaacgtgga ggcgcgcagc ctcgtcgcct gcatccagcc tgagctgatc 6360
tagtgctagc aagggaaccc catgccctcc ccccgtctgc gcttcggtgt gctgggtgcc 6420
gccgacatcg ccctgcgccg caccgtgccc gcgctgctcg cccacccgga cgtcaccgtg 6480
gtcgcggtct ccagccggga cacggcccgc gccgcccgtt tcgccgcggc gttcgggtgc 6540

gaggcggtcc cgggccacca ggcgctgctg gaccgcgacg acatcgacgc gctgtacgtc 6600
ccgctgccgg tgatggtgca cacgccctgg gtggaggcgg cgctgctgcg gggcaggcac 6660
gtcctggtgg agaagccgct gaccgcgacc cgctccggag ccgaggacct gatcgccctg 6720
gcacgctccc gcggcctggt cctgatggag aacttcacgt cgctgcacca cgcacagcac 6780
ggcaccgtca cggacctcct gcgggacggg acgatcggcg aactgcgctc gctgtccgcg 6840
gcgttcacga tcccgccgaa gccggagggc gacatccggt accagcccga cgtgggcgga 6900
ggcgccctgc tcgacatcgg gatctacccc ctgcgggcgg cactgcactt cctgggcccc 6960

gacctgcacg cggccggcgc ggtcctgcgt cgggagcggc gccggaacgt ggtggtctcc 7020
ggccacgtgc tgctcaccac accgcacggc gtcgtggccg agctggcctt cggcatggag 7080
cacgcctacc gctcggagta caccctcttc ggcacggccg gccgcctgcg cctggaccgt 7140
gccttcaccc cgcccgagac ccaccggcca cgtgtggaga tccaccggca ggacgccctg 7200
gacatcgtcg acctgccgcc ggacgcgcaa ttcgccaacc tcgtccggga cttcgtactg 7260
gctgtccgcg agggtcccgg ccggctcaca cagcaccacg ccgacgccgt acgccaagcc 7320
gatctcgtcg agcgcgtcat ggcggtggcg cgggtgcgct ggtgctgtct agtatcgctc 7380

cgagcccgaa gggaaaatcg agtgcccaat tcggcagaat cgggatcgat ggagttactc 7440
gacgtcgacg gggcctggtt atacaccccg gaaatcatgc gggacgaacg gggagaattc 7500
ctcgaatggt tccggggtcg gacattccag gagaagatcg gccacccccc ctcgctggcc 7560
caggccaact gctcggtgtc ccgcaaggcg ttctgcgcgg catccacttc gccgacgccc 7620
ccgcccggcc aggccaagta cgtcacctgc gcctccggca ccgtgctcga cgtggtcgtc 7680
gacgtacgcc ggggctcgcc caccttcggc cgatgggccg ccgtccgact cgacgcggcc 7740
cgccaccagg ggctctacct ggccgaagga ctcggccacg cgttcatggc cctcaccgac 7800

gacgccacgg tcgtctacct ctgctcacag ccctacgtgg ccgaggccga gcgggccgta 7860

gaccctctcg acccggcgat cggcatcgaa tggccgacgg acatcgacat cgtccctgtc 7920

ggcgaaggga cacccacgca ccgtccctgg cgcaggccgc ggagacccgg catcctgccg 7980

gactacgagg gagtgcccgg cgccttacat cgcggaggcg gccggcgtgg gaccggcccg 8040

tgaaggccct cgtactggcc ggctctagtt acgtagtacc gcgacaaccg cgcctggtgg 8100

gaacccctga agaagcggcc ggccggtccc gccgcccccc cgagaggcag cggcccatga 8160

gatggctgat caccggcgcc gccggaatgc tgggccggga actcgtccgg cgcctcgccg 8220

agaacgagga ggacgtcgcg gccctcggcc acgaccacct cgacgtcacc cgaccctccg 8280

ccgtgcgggc ggcactcgcc gagcaccgtc ccgggatcgt cgtcaactgc gccgcctaca 8340

cggccgtcga cgacgccgag acggacgagg ccgccgctgc cctcctcaac gccgaggcgc 8400

cccggctgct ggccgagggc ctgcgccccc accggcggca cggcctcgtc cacctgtcca 8460

ccgactacgt ctttcccggc gacgcccgca ccccctacgc cgaggaccac cccacggctc 8520

cccgcagcgc ctacggacgc accaaacggg acggcgagca agcggtgctg acggcactgc 8580

ccaccgccac cgtgctgcgc accgcctggc tgtacgggcg caccggccgc agcttcgtcc 8640

gcacgatgat cgaacgggag gcgcgcggcg gagccatcga cgtcgtcgcc gaccagcgcg 8700

gccagcccac ctggaccggc gacctcgccg accgcatcat cgccgtcggc cggcaccccg 8760

gcgtccacgg catcctgcac gccaccaacg ccggctccgc cacctggtac gacctggcac 8820

aagaggtctt ccggctcctc gacgccgacc ccgggcgggt ccggcccacc accggcgccg 8880

ccttccgcag acccgccccc cgccccgcct acagcgtcct cggccacgac cgctggcgcg 8940

ggaccggcct ggcacccctg cgtgactggc gctcggccct gcgcgaggcg ttccccgaca 9000

tcctcgccgc ggaacaccca ccgacccggc gaggagcagc atgaaacgag gcgtgcacga 9060

cctggccctc ttccctaggt ctaga 9085

<210> 21

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 21

ggcaagctta gcggggcgac tggcgtgccc act 33

<210> 22

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 22

ggtctgcagt caccgtgggt tctgccatct ctt 33

<210> 23

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 23

gctggtaccg gatgttccct ccgggccacc gtc 33

<210> 24

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 24

tgagaattcc ctcgccgtcc tgcccgcgct tgg 33

<210> 25

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 25

ttaattaaga tatcaccggc aaggaaggac acgacgcc 38

<210> 26

<211

> 36

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 26

tctagacata tggcgcagat acaggggtga ggcctg 36

<210> 27

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 27

ttaattaaac tagtatcgat gacggtggcc cggaggg 37

<210> 28

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 28

tctagatgcg ggtcagcgca ggaagccgcg 30

<210> 29

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 29

ttaattaaac tagttaactc gccacgccga ccgtt 35

<210> 30

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 30

tctagagagc tcctcgtagg cggcctt 27

<210> 31

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 31

ttaattaaac tagtcaggtc tccttcgcgg acggcctc 38

<210> 32

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 32

tctagaacct gactaggcct ggtcgacccg 30

<210> 33

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 33

ttaattaaac tagtccccag gcctcacccc tgtatctg 38

<210> 34

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 34

tctagagaat tcgctcaggc ggggacgccg acgaag 36

<210> 35

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 35

ttaattaaac tagtatcgat atcgctccga gcccgaaggg a 41

<210> 36

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 36

tctagagccg gccagtacga gggcctt 27

<210> 37

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 37

ttaattaaac tagtgtaccg cgacaaccgc 30

<210> 38

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 38

tctagagaag agggccaggt cgtgcacgc 29

<210> 39

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 39

ttaattaaac tagtatcgat gggaatcgcg gaagcg 36

<210> 40

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 40

tctagaatca gctcagggcc tggatgc 27

<210> 41

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 41

ttaattaaac tagtggtcgt cggcatctgc g 31

<210> 42

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 42

tctagactcg gttcctcttg tgctcactgc 30

<210> 43

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 43

ttaattaaac tagtacagag atccaggacg acgca 35

<210> 44

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 44

tctagatcag atacggacgg cggaggt 27

<210> 45

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 45

ttaattaaac tagtaaggga accccatgcc ctccc 35

<210> 46

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 46

tctagaatca gcaccagcgc acccgcgcca 30