



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0116102
(43) 공개일자 2016년10월07일

(51) 국제특허분류(Int. Cl.)

C07H 17/02 (2006.01) A61K 31/395 (2006.01)
C07D 225/02 (2006.01) C12N 15/76 (2006.01)
C12N 9/10 (2006.01) C12P 17/10 (2006.01)
C12R 1/465 (2006.01)

(52) CPC특허분류

C07H 17/02 (2013.01)
A61K 31/395 (2013.01)

(21) 출원번호 10-2015-0041636

(22) 출원일자 2015년03월25일
심사청구일자 2015년03월25일

(71) 출원인

이화여자대학교 산학협력단

서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)

(72) 발명자

윤여준

서울특별시 강남구 압구정로 151 126동 202호 (압구정동, 신현대아파트)

강한영

충청북도 청주시 서원구 신율로 13 대우푸르지오아파트 407동301호

(뒷면에 계속)

(74) 대리인

손민

전체 청구항 수 : 총 10 항

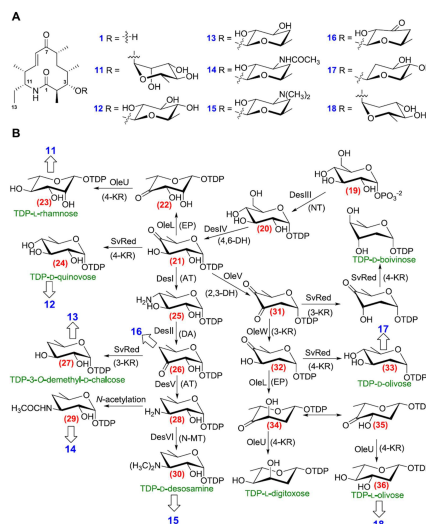
(54) 발명의 명칭 신규 마크로락탐 글리코사이드 유도체, 이의 화학효소적 제조방법 및 이를 포함하는 항균용 조성물

(57) 요약

본 발명은 신규 마크로락탐 글리코사이드 유도체, 이의 화학효소적 제조방법 및 이를 유효성분으로 포함하는 항균용 조성물에 관한 것이다.

본 발명의 신규한 마크로락탐 글리코사이드 유도체를 제조하는 화학효소적 방법은 내재적으로 마크로락톤 글리코사이드 유도체를 생산하는 미생물 및 상기 마크로락톤 글리코사이드 유도체의 생합성에 관여하는 효소를 이용하여 외부로부터 제공되는 화학적으로 합성된 마크로락탐을 글리코실화함으로써 비천연 마크로락탐 글리코사이드 유도체를 제조하는 방법으로서, 본 발명의 방법에 의해 제조된 마크로락탐 글리코사이드 유도체는 천연 마크로락톤 글리코사이드와 비교하여 동등 이상의 항균활성을 가지며 나아가 개선된 대사 안정성을 나타내므로 신규한 구조의 항균 물질 발굴에 유용하게 사용할 수 있다.

대표도 - 도4



(52) CPC특허분류

C07D 225/02 (2013.01)
C12N 15/76 (2013.01)
C12N 9/1048 (2013.01)
C12P 17/10 (2013.01)
C12Y 204/00 (2013.01)
C12Y 204/01248 (2013.01)
C12R 1/465 (2013.01)

신데, 프라모드 비

서울특별시 서대문구 봉원사길 52 201호 (대신동, 해맞이타운)

(72) 발명자

이동건

대구광역시 남구 앞산순환로 657 105동 801호 (봉덕동, 대덕1차아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711015538
부처명	미래창조과학부
연구관리전문기관	한국연구재단
연구사업명	원천기술개발사업-글로벌프론티어사업
연구과제명	거대인공이차대사산물 생합성유전자집단의방선균발현시스템개발
기 여 율	1/2
주관기관	이화여자대학교 산학협력단
연구기간	2011.09.29 ~ 2020.08.31이 발명을 지원한 국가연구개발사업
과제고유번호	1711012398
부처명	미래창조과학부
연구관리전문기관	한국연구재단
연구사업명	기초연구사업(학술진흥)(2013)-중견연구자지원사업-도약연구지원사업(도전)
연구과제명	미생물유래천연물의생합성경로규명및재설계
기 여 율	1/2
주관기관	이화여자대학교 산학협력단
연구기간	2014.06.01 ~ 2015.05.31

명세서

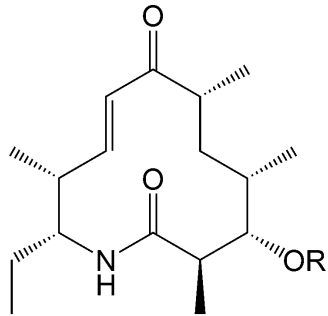
청구범위

청구항 1

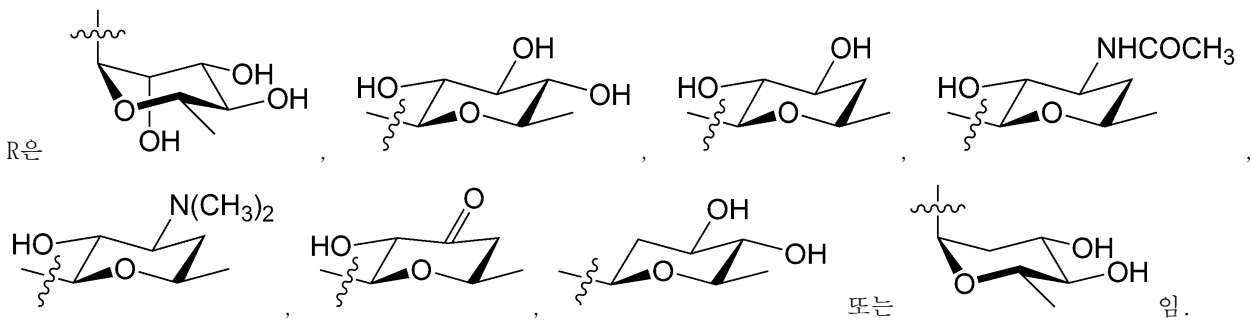
하기 화학식 1로 표시되는 마크로락탐 글리코시드 유도체 또는 이의 약학적으로 허용가능한 염:

상기 식에서,

[화학식 1]



상기 식에서,

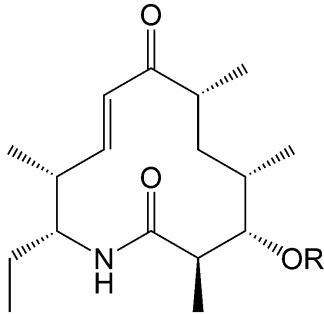


청구항 2

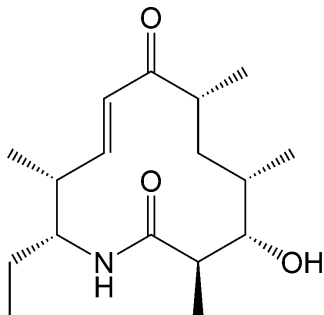
제1항의 마크로락탐 글리코시드 유도체의 화학효소적(chemoenzymatic) 제조방법으로서,

내재적으로 마크로라이드를 생합성하는 스트렙토마이세스 속 미생물에서 피크로마이신 및 마크로라이드의 당 부분인 데조자민(desosamine)을 생합성 및 전이시키는 단백질을 코딩하는 유전자 클러스터를 게놈으로부터 결실시킨 미생물에, 글리코실전달효소 및 데옥시당 생합성효소를 코딩하는 핵산을 포함하는 발현벡터가 도입된 재조합 미생물을 하기 화학식 2로 표시되는 화합물을 포함하는 배지에 배양하는 단계를 포함하는, 하기 화학식 1로 표시되는 마크로락탐 글리코시드 유도체의 제조방법:

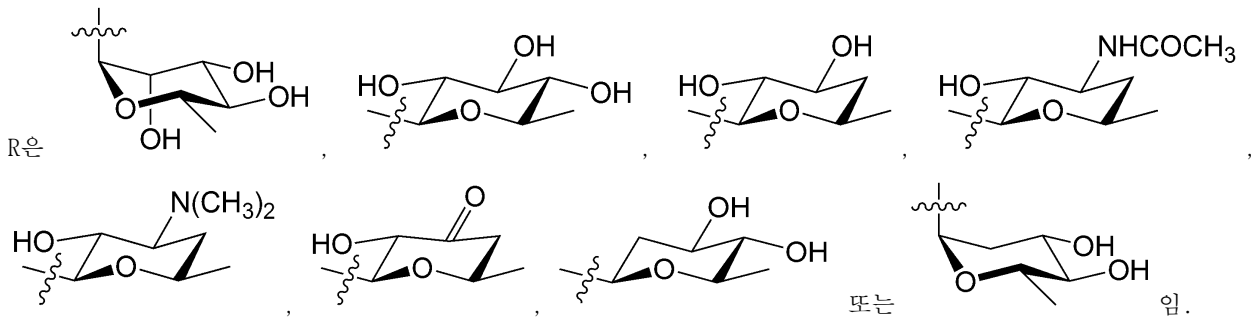
[화학식 1]



[화학식 2]



상기 식에서,



청구항 3

제2항에 있어서,

상기 스트렙토마이세스 속 미생물은 스트렙토마이세스 베네주엘라인 것인 제조방법.

청구항 4

제2항에 있어서,

상기 글리코실전달효소는 DesVII 및 DesVIII의 조합인 것인 제조방법.

청구항 5

제2항에 있어서,

상기 글리코실전달효소 및 테옥시당 생합성시키는 단백질을 코딩하는 핵산은 pLRHM2, pDQNV, pDDSS 또는 pLOLV2의 유전자 조합을 포함하는 것인 제조방법.

청구항 6

제5항에 있어서,

상기 pLRHM2, pDQNV, pDDSS 및 pLOLV2는 각각 서열번호 1 내지 4의 서열을 갖는 것인 제조방법.

청구항 7

제2항에 있어서,

배양된 미생물 또는 배양액으로부터 마크로락탐 글리코시드 유도체를 회수하는 단계를 추가로 포함하는 것인 제조방법.

청구항 8

제1항의 마크로락탐 글리코시드 유도체 또는 이의 약학적으로 허용가능한 염을 유효성분으로 포함하는 항균용 조성물.

청구항 9

제8항에 있어서,

상기 조성물은 에리트로마이신-민감성 균주 또는 에리트로마이신-저항성 균주에 대하여 항균 활성을 갖는 것인 조성물.

청구항 10

제9항에 있어서,

상기 조성물은 E. 패시움 ATCC 19434, S. 아우레우스 ATCC 25923, E. 패시움 P00558 및 S. 아우레우스 P00740 으로 구성된 군으로부터 선택되는 미생물에 대하여 항균 활성을 갖는 것인 조성물.

발명의 설명

기술 분야

[0001] 본 발명은 신규 마크로락탐 글리코시드 유도체, 이의 화학효소적 제조방법 및 이를 유효성분으로 포함하는 항균용 조성물에 관한 것이다.

배경 기술

[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 생산하는 균주를 발굴하였으나, 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 고려하여 설계할 수 있으므로 생물학적 활성을 예상할 수 있어 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, 마크로라이드는 폴리케타이드 마크로락톤 환과 그에 부착된 하나 또는 그 이상의 테옥시당으로 구성된다. 상기 마크로라이드는 스테필로코커스 및 엔테로코커스와 같은 그람-양성 박테리아에 대해 주로 항균 활성을 나타낸다. 마크로라이드 중에서 테옥시당 모이어티가 표적 분자와 특이적으로 접촉함으로써 분자의 생물학적 활성에 크게 영향을 주므로, 테옥시당 모이어티의 구조가 중요하며, 따라서, 상기 언급한 PKS 유전자 변형에 의한 폴리케타이드 구조 변형 외에도 글리코실화와 같은 후-PKS 테일러링 단계(post-PKS tailoring step)가 반드시 필요하다. 그러나, 마크로라이드 계열 항생물질에 대해서도 저항성을 갖는 병원체들이 출현함으로써 인하여 신규 마크로라이드 유도체를 발굴하기 위한 연구가 계속되고 있다. 기존에는 화학적 및 효소적 합성을 이용한 연구가 많이 진행되어 왔으나, 테옥시당 모이어티의 화학적 수식 또는 뉴클레오타이드-활성화되는 테옥시당의 시험관 내 효소적 합성이 용이하지 않다는 문제가 있다.

[0006] YC-17은 나르보마이신과 함께 스트렙토마이세스 베네주엘레 균주로부터 생산되는 글리코실화된 12원-마크로락톤 고리를 포함하는 화합물로서 항균활성을 갖는 것으로 알려져 있다.

[0007] 이에 본 발명자들은 항균활성을 갖는 신규한 마크로라이드 유사 화합물 및 이의 제조방법을 발굴하기 위하여 예의 연구 노력한 결과, 특정 글리코실전달효소 및 테옥시당 생합성효소를 발현하도록 형질전환된 S. 베네주엘레 균주를 화학적으로 합성된 천연 마크로락톤 글리코시드 유도체인 YC-17과 유사한 구조를 갖는 마크로락탐 아글리콘을 제공하여 배양함으로써 신규한 마크로락탐 글리코시드 유도체를 화학효소적으로 제조할 수 있으며, 이와 같이 제조된 화합물들은 항균활성을 가질뿐만 아니라 개선된 대사 안정성을 나타냄을 확인하고 본 발명을 완성하였다.

발명의 내용

해결하려는 과제

[0008] 본 발명의 하나의 목적은 신규 마크로락탐 글리코시드 유도체 또는 이의 약학적으로 허용가능한 염을 제공하는 것이다.

[0009] 본 발명의 다른 목적은 내재적으로 마크로라이드를 생합성하는 스트렙토마이세스 속 미생물에서 피크로마이신 및 마크로라이드의 당 부분인 데조자민(desosamine)을 생합성 및 전이시키는 단백질을 코딩하는 유전자 클러스터를 계놈으로부터 결실시킨 미생물에, 글리코실전달효소 및 테옥시당 생합성효소를 코딩하는 핵산을 포함하는 발현벡터가 도입된 재조합 미생물을 하기 화학식 2로 표시되는 화합물을 포함하는 배지에 배양하는 단계를 포함하는, 상기 마크로락탐 글리코시드 유도체의 화학효소적 제조방법을 제공하는 것이다.

[0010] 본 발명의 또 다른 목적은 상기 마크로락탐 글리코시드 유도체 또는 이의 약학적으로 허용가능한 염을 유효성분으로 포함하는 항균용 조성물을 제공하는 것이다.

[0011] 본 발명의 또 다른 목적은 상기 마크로락탐 글리코시드 유도체 또는 이의 약학적으로 허용가능한 염을 이를 필요로 하는 개체에 투여하는 단계를 포함하는 미생물 감염성 질환의 예방 또는 치료방법을 제공하는 것이다.

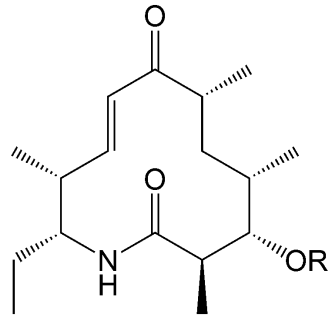
[0012] 본 발명의 또 다른 목적은 상기 마크로락탐 글리코시드 유도체 또는 이의 약학적으로 허용가능한 염을 이용하여 에리트로마이신-민감성 균주, 에리트로마이신-저항성 균주를 포함하는 미생물의 생존 및/또는 증식을 억제하는 방법을 제공하는 것이다.

과제의 해결 수단

[0013] 상기 목적을 달성하기 위한 하나의 양태로서, 본 발명은 하기 화학식 1로 표시되는 마크로락탐 글리코시드 유도체 또는 이의 약학적으로 허용가능한 염을 제공한다:

[0014] 상기 식에서,

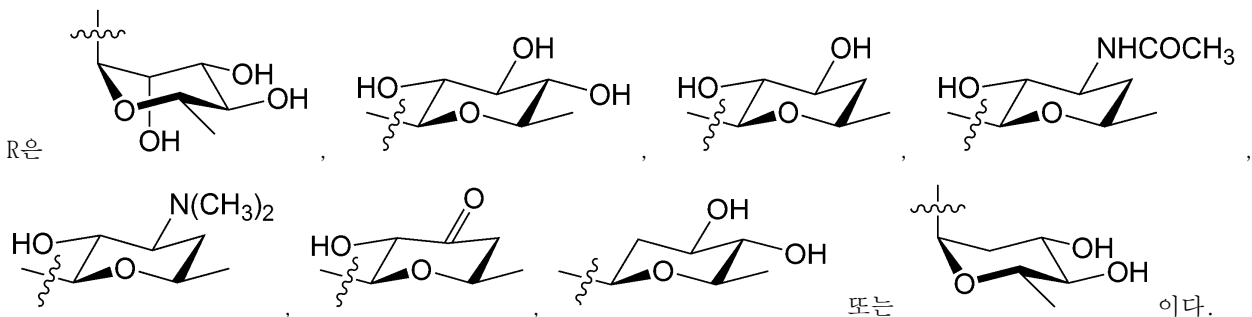
[0015] [화학식 1]



[0016]

[0017] 상기 식에서,

[0018]



[0019] 수소와 탄소에 대한 핵자기공명 스펙트럼을 이용하여 화합물을 동정하여, 상기 화학식 1로 표시되는 일련의 마크로락탐 글리코시드 유도체는 신규한 화합물임을 확인하였다.

[0020] 본 발명의 화합물은 용매화물 또는 전구약물(pro-drug) 형태일 수 있으며, 이들 모두 본 발명의 범위에 포함된다. 용매화물은 바람직하게는 수화물 및 에탄올화물을 포함한다.

[0021] 본 발명의 화합물은 염, 특히 약학적으로 허용가능한 염의 형태로 존재할 수 있다. 염으로는 약학적으로 허용가능한 유리산(free acid)에 의해 형성된 산부가염과 같이, 당 업계에서 통상적으로 사용되는 염을 제한 없이 사용할 수 있다. 본 발명의 용어 "약학적으로 허용 가능한 염"이란 개체에게 비교적 비독성이고 무해한 유효작용을 갖는 농도로서 이 염에 기인한 부작용이 화학식 1로 표시되는 화합물의 이로인 효능을 저하시키지 않는 상기 화합물의 염의 모든 유기 또는 무기부가염을 의미한다.

[0022] 산부가염은 통상의 방법, 예를 들어 화합물을 과량의 산 수용액에 용해시키고, 이 염을 수산화성 유기 용매, 예를 들어 메탄올, 에탄올, 아세톤 또는 아세토니트릴을 사용하여 침전시켜서 제조한다. 동 물량의 화합물 및 물 중의 산 또는 알코올(예, 글리콜 모노메틸에테르)을 가열하고, 이어서 상기 혼합물을 증발시켜 건조시키거나, 또는 석출된 염을 흡인 여과시킬 수 있다.

[0023] 이때, 유리산으로는 유기산과 무기산을 사용할 수 있으며, 무기산으로는 염산, 인산, 황산, 질산, 주석산 등을 사용할 수 있고 유기산으로는 메탄술폰산, p-톨루엔술폰산, 아세트산, 트리플루오로아세트산, 말레인산(maleic acid), 숙신산, 옥살산, 벤조산, 타르타르산, 푸마르산(fumaric acid), 만테르산, 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.

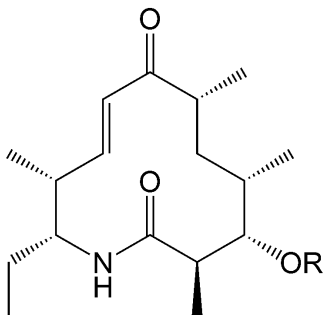
[0024] 또한, 염기를 사용하여 약학적으로 허용 가능한 금속염을 만들 수 있다. 알칼리 금속염 또는 알칼리 토금속 염은, 예를 들어 화합물을 과량의 알칼리 금속 수산화물 또는 알칼리 토금속 수산화물 용액 중에 용해시키고, 비용해 화합물 염을 여과한 후 여액을 증발, 건조시켜 얻는다. 이때, 금속염으로는 특히 나트륨, 칼륨, 또는 칼슘 염을 제조하는 것이 제약상 적합하나 이들에 제한되는 것은 아니다. 또한 이에 대응하는 은염은 알칼리 금속 또는 알칼리 토금속 염을 적당한 은염(예, 질산은)과 반응시켜 얻을 수 있다.

[0025] 본 발명의 화합물의 약학적으로 허용 가능한 염은, 달리 지시되지 않는 한, 상기 화학식 1의 화합물에 존재할 수 있는 산성 또는 염기성 기의 염을 포함한다. 예를 들어, 약학적으로 허용 가능한 염으로는 히드록시기의 나트륨, 칼슘 및 칼륨염 등이 포함될 수 있고, 아미노기의 기타 약학적으로 허용 가능한 염으로는 히드로브롬화물, 황산염, 수소 황산염, 인산염, 수소 인산염, 이수소 인산염, 아세테이트, 숙시네이트, 시트레이트, 타르트레이트, 락테이트, 만델레이트, 메탄술포네이트(메실레이트) 및 p-톨루엔술포네이트(토실레이트) 염 등이 있으며, 당 업계에 알려진 염의 제조방법을 통하여 제조될 수 있다.

[0026] 본 발명의 화합물은 당업계에서 통상적으로 사용되는 방법에 따라 합성할 수 있으며, 바람직하게는 본 발명의 방법을 사용하여 재조합 미생물로부터 생산할 수 있다. 본 발명의 화합물의 구체적인 제조방법은 후술할 것이다.

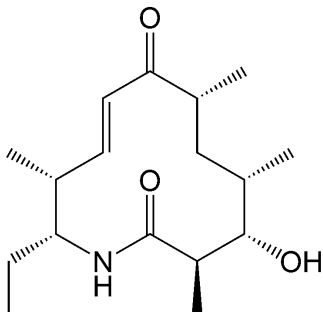
[0027] 다른 하나의 양태로서, 본 발명은 상기 신규 마크로락탐 글리코시드 유도체의 화학효소적(chemoenzymatic) 제조 방법으로서, 내재적으로 마크로라이드를 생합성하는 스트렙토마이세스 속 미생물에서 피크로마이신 및 마크로라이드의 당 부분인 테조자민(desosamine)을 생합성 및 전이시키는 단백질을 코딩하는 유전자 클러스터를 게놈으로부터 결실시킨 미생물에, 글리코실전달효소 및 테옥시당 생합성효소를 코딩하는 핵산을 포함하는 발현벡터가 도입된 재조합 미생물을 하기 화학식 2로 표시되는 화합물을 포함하는 배지에 배양하는 단계를 포함하는, 하기 화학식 1로 표시되는 마크로락탐 글리코시드 유도체의 제조방법을 제공한다:

[0028] [화학식 1]



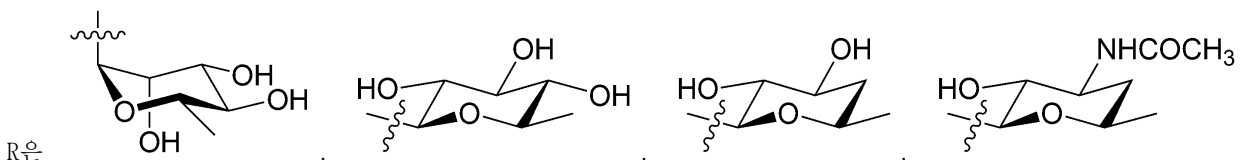
[0029]

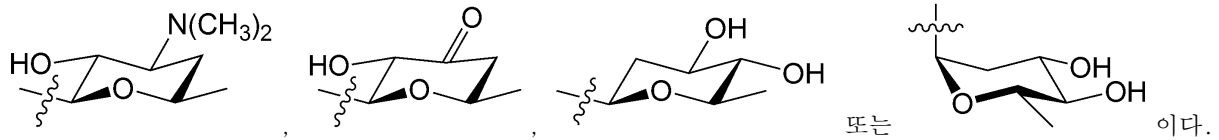
[0030] [화학식 2]



[0031]

[0032] 상기 식에서,





이다.

[0034] 본 발명은 유전자 조작을 통해 내재적으로 마크로락톤 계열의 마크로라이드를 생합성하는 미생물로부터 피크로마이신 및 마크로라이드의 당 부분인 데조자민을 생합성 및 전이시키는 단백질을 코딩하는 유전자 클러스터를 게놈으로부터 결실시키고, 글리코실전달효소 및 데옥시당 생합성효소를 코딩하는 핵산을 포함하는 발현백터를 도입하여 재조합 미생물을 제조한 뒤 이를 화학적으로 합성한 원하는 구조 예컨대, 마크로락탐을 기초로 하는 아글리콘을 공급하면서 배양하여 마크로락톤 계열의 마크로라이드 글리코시드 유도체를 생산하는 미생물로부터 마크로락탐 계열의 마크로라이드 글리코시드 유도체를 화학효소적 방법에 의해 생산할 수 있음을 최초로 발견한 것이 특징이다. 특히, 글리코실전달효소로서 DesVII 및 DesVIII의 조합을 코딩하는 핵산을 발현백터에 도입하여 성공적으로 마크로락탐 계열의 마크로라이드 글리코시드 유도체를 획득할 수 있음을 확인하였다. 이는 글리코시드 전달효소 DesVII/DesVIII가 마크로락톤이 아닌 마크로락탐 골격을 인식하고 이를 글리코실화할 수 있음을 최초로 발견한 것이다.

[0035] 본 발명에서 마크로락탐 글리코시드 유도체 제조에 사용한 미생물은 내재적으로 마크로라이드를 생합성하는 스트렙토마이세스 속 미생물로, 내재적으로 마크로라이드를 생합성하는 미생물이란, 내재적으로 마크로라이드 생합성에 필요한 모든 효소를 발현하는 미생물을 의미한다. 상기 "내재적"이란 야생형 미생물이 천연의 상태로 가지고 있는 활성을 의미한다. 본 발명에서 마크로라이드는 하나 이상의 데옥시당, 예컨대, 클라디노스 및 데조자민 등이 결합된 거대고리형 락톤을 포함하는 마크로락톤 계열의 화합물일 수 있다. 예컨대, D-데조자민이 결합된 12원 고리형 마크로라이드 항생 물질인 YC-17일 수 있다.

[0036] 본 발명의 스트렙토마이세스 속 미생물은 스트렙토마이세스 베네주엘라이일 수 있으나, 이에 제한되지 않고 내재적으로 마크로라이드를 생합성할 수 있는 모든 스트렙토마이세스 속 미생물을 포함한다. 스트렙토마이세스 베네주엘라이는 유전자 조작이 쉽고 상대적으로 빠르게 성장하므로(배가시간, 약 1시간), 대량물질 생산을 위하여 요구되는 배양기간이 짧다. 유전자 조작되지 않은 천연 상태의 상기 미생물은 단일 폴리케타이드 신타아제(PKS) 및 다수의 PKS 후 수식 효소들(데조자민 당의 생합성 및 전이에 연관된 des 클러스터에 의한 글리코실화 및 PikC P450 모노옥시게나아제에 의한 히드록실화)의 작용을 통하여 12원 및 14원 고리형 마크로라이드, 메티마이신 및 피크로마이신을 모두 생산할 수 있다. 본 발명의 구체적 실시예에서는, 스트렙토마이세스 베네주엘라이 ATCC 15439를 사용하였으나, 이에 제한되지 않는다.

[0037] 본 발명의 제조방법은, 내재적으로 생산가능한 다양한 12원 또는 14원 고리형 마크로라이드 등의 글리코시드 유도체가 무분별하게 합성되는 것을 방지하고 원하는 화합물을 효율적으로 제조하기 위하여, 상기 스트렙토마이세스 속 미생물로서 피크로마이신 및 마크로라이드의 당 부분인 데조자민(desosamine)을 생합성 및 전이시키는 단백질을 코딩하는 유전자 클러스터를 게놈으로부터 결실시킨 변이균주에 글리코실전달효소 및 데옥시당 생합성효소를 코딩하는 핵산을 포함하는 발현백터가 도입된 재조합 미생물을 사용하는 것이 특징이다.

[0038] 상기 피크로마이신(GenBank #: AF079138.1) 및 마크로라이드의 당 부분인 데조자민(desosamine; GenBank#: AF079762.1)을 생합성 및 전이시키는 단백질을 코딩하는 유전자 클러스터는 각각 PikAI-PikAII-PikAIII-PikAIV-PikAV 및 DesVIII-DesVII-DesVI-DesR-DesV-DesIV-DesIII-DesII-DesI의 단백질 조합으로 구성될 수 있다. 상기 미생물은 게놈으로부터 상기 각각의 단백질을 코딩하는 유전자인 서열번호 5 내지 18의 염기서열로 구성된 유전자 클러스터가 결실된 것일 수 있다.

[0039] 게놈상에서 목적인 유전자만의 특이적 결실은 당 분야에서 확립된 방법을 통하여 수행할 수 있으며, 방법은 특별히 제한되지 않으며, 상동 재조합(homologous recombination)법을 사용할 수도 있다. 목적인 단백질의 N 말단과 C 말단을 코딩하는 핵산 사이에 선별마커를 포함하는 백터로 스트렙토마이세스 속 균주를 형질전환시켜 게놈과 백터 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.

- [0040] 전술한 바와 같이, 바람직하게 상기 글리코실전달효소로서 DesVII 및 DesVIII의 조합을 코딩하는 핵산을 포함하는 발현백터를 도입한 재조합 미생물을 사용할 수 있다.
- [0041] 상기 DesVII은 글리코실전달효소의 하나로 dTDP-3-디메틸아미노-3,4,6-트리데옥시-알파-D-글루코피라노스:10-데옥시메티노라이드 3-디메틸아미노-4,6-디데옥시-알파-D-글루코스전달효소의 시스템 명을 갖는 효소이며 하기 화학반응을 촉매한다:
- [0042] $\text{dTDP-3-디메틸아미노-3,4,6-트리데옥시-알파-D-글루코피라노스} + 10\text{-데옥시메티노라이드} \rightleftharpoons \text{dTDP} + 10\text{-데옥시메티마이신}.$
- [0043] DesVII은 다양한 크기의 고리형태의 마크로락톤에 TDP-D-데조자민을 부착시키는 글루코실전달효소일 수 있다.
- [0044] 한편, DesVIII은 마크로라이드 항생물질, 네오메티마이신 및 피크로마이신의 생합성에 있어서 DesVII 글루코실전달효소에 의한 TDP-D-데조자민의 전달을 향상시키는 보조 단백질(auxiliary protein)이다.
- [0045] 예컨대, 상기 글리코실전달효소 및 데옥시당 생합성시키는 단백질을 코딩하는 핵산은 pLRHM2, pDQNV, pDDSS 또는 pLOLV2의 유전자 조합을 포함할 수 있으며, 상기 pLRHM2, pDQNV, pDDSS 및 pLOLV2는 각각 서열번호 1 내지 4의 서열을 갖는 것일 수 있다.
- [0046] 바람직하게는 스트렙토마이세스 베네주엘레 YJ028에 상기 서열번호 1 내지 4 중 어느 하나의 서열을 포함하는 백터를 도입하여 제작한 변이균주를 사용할 수 있다. 상기 균주는 본 발명의 실시예에 기재된 방법에 의해 제조할 수 있으나, 이에 제한되지 않는다.
- [0047] 본 발명에서 용어, "배양"은 미생물을 적당히 인공적으로 조절된 환경조건에서 생육시키는 것을 의미한다. 본 발명의 신규한 마크로락탐 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, 바람직하게 종래 스트렙토마이세스 속 균주의 배양에 이용되고 있는 공지의 영양원을 사용할 수 있다. 예컨대, 본 발명의 미생물의 배양은 말론산(malonic acid), 에탄올, 메티오닌(methionine), 탄소원 및 질소원을 포함하는 배지에서 이루어질 수 있다. 이때, 상기 탄소원의 비제한적인 예로는 전분, 포도당, 옥수수 기름, 글리세롤, 말토스, 만노스 및 이노시톨이 있다. 바람직하게, 전분, 포도당 및 옥수수기름을 사용할 수 있으나, 이에 제한되지 않는다. 상기 질소원의 비제한적인 예로는 면실밀, 옥수수 침지액, 옥수수 침지분, 대두분, 펩톤 및 효모 엑기스가 있다. 바람직하게, 면실밀 및 옥수수 침지액을 사용할 수 있으나, 이에 제한되지 않는다. 종균 배양액을 상기 배양 배지를 포함한 발효조에 식균하여 배양할 수 있다. 배양방법으로는 호기적 조건에서는 진탕배양 또는 정지배양이 가능하다. 배양온도는 상기 각 조건들에서 배양할 경우 조건에 따라 약간씩 상이하기는 하나, 보통 20 내지 37℃에서 배양하는 것이 적당하며, 바람직하게는 25 내지 30℃에서 배양할 수 있다. 또한, 배양기간 역시 당업계에서 사용되는 공지의 기간 동안 배양할 수 있으며, 필요에 따라 기간이 조정될 수 있다. 바람직하게 진탕배양, 정지배양의 경우 모두 4 내지 7일간 배양할 수 있으나, 이에 제한되지 않는다.
- [0048] 바람직하게, 본 발명의 제조방법은 배양된 미생물 또는 배양액으로부터 마크로락탐 글리코시드 유도체를 회수하는 단계를 추가로 포함할 수 있다.
- [0049] 상기 배양된 미생물 또는 배양액으로부터 마크로락탐 유도체를 회수하는 방법은 당업계에 널리 알려진 방법을 제한없이 사용하여 수행할 수 있다. 상기 마크로락탐 유도체 회수 방법은 여과, 음이온 교환 크로마토그래피, 결정화, HPLC 또는 이들을 조합하여 사용할 수 있으나, 이들 예에 한정되는 것은 아니다.
- [0050] 또 하나의 양태로서, 본 발명은 상기 마크로락탐 글리코시드 유도체 또는 이의 약학적으로 허용가능한 염을 유효성분으로 포함하는 항균용 조성물을 제공한다.

- [0051] 본 발명에서 "항균용 조성물"은 미생물의 생존 및/또는 증식을 억제하는 작용을 하는 조성물을 의미한다. 상기 조성물은 에리트로마이신-민감성 균주, 에리트로마이신-저항성 균주 또는 이들 균주 모두에 대하여 항균 활성을 가지는 것일 수 있다.
- [0052] 구체적으로, 상기 조성물은 E. 패시움 ATCC 19434, S. 아우레우스 ATCC 25923, E. 패시움 P00558 및 S. 아우레우스 P00740으로 구성된 균으로부터 선택되는 미생물에 대하여 항균 활성을 가질 수 있다. 따라서, 상기 조성물은 상기 미생물에 의해 발생하는 감염성 질환을 예방 또는 치료에 사용될 수 있다.
- [0053] 본 발명의 구체적인 실시예에서는 본 발명의 신규 마크로락탐 글리코사이드 유도체 중 L-람노실-AZDM(11), D-퀴노보실-AZDM(12) 및 D-올리보실-AZDM(17)은 에리트로마이신-민감성 균주인 E. 패시움 ATCC 19434 및 S. 아우레우스 ATCC 25923에 대해서 뿐만 아니라 에리트로마이신-저항성 균주에 대해서도 공지의 항균 물질인 에리트로마이신에 비해 현저히 향상된 항균 활성을 나타내며, 3-O-데메틸-D-칼코실-AZDM(13) 및 3-케토-4,6-디데옥시-D-글루코실-AZDM(16)은 에리트로마이신-민감성 균주에 대해 에리트로마이신과 동등 이상의 항균 활성을 나타냄을 확인하였다(표 4).
- [0054] 상기 조성물은 약학적으로 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, 텍스트로오스, 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.
- [0055] 또한, 본 발명의 조성물은 각각 통상의 방법에 따라 산제, 과립제, 정제, 캡슐제, 현탁액, 에멀전, 시럽, 에어로졸 등의 경구형 제형, 외용제, 좌제 또는 멸균 주사용액의 형태로 제형화하여 사용할 수 있다. 상세하게는, 제형화할 경우 통상 사용하는 충전제, 증량제, 결합제, 습윤제, 봉해제, 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. 비수성용제 및 현탁제로는 프로필렌글리콜, 폴리에틸렌 글리콜, 올리브 오일과 같은 식물성 오일, 에틸올레이트와 같은 주사가능한 에스테르 등이 사용될 수 있다. 좌제의 기제로는 위템솔, 마크로골, 트윈 61, 카카오지, 라우린지, 글리세로젤라틴 등이 사용될 수 있다.
- [0056] 또한, 본 발명은 상기 화학식 1의 화합물 또는 그의 약제학적으로 허용가능한 염을 이를 필요로 하는 개체에 투여하는 단계를 포함하는 미생물 감염성 질환의 예방 또는 치료방법을 제공한다.
- [0057] 또한, 본 발명은 상기 화학식 1의 화합물 또는 그의 약제학적으로 허용가능한 염을 이용하여 에리트로마이신-민감성 균주, 에리트로마이신-저항성 균주를 포함하는 미생물의 생존 및/또는 증식을 억제하는 방법을 제공한다.
- [0058] 본 발명에서 사용된 용어, "개체"는 미생물 감염성 질환이 발병하였거나 발병할 수 있는 인간을 포함한 모든 동물을 의미한다. 본 발명의 조성물을 개체에게 투여하여 상기 미생물 감염성 질환을 효과적으로 예방 또는 치료할 수 있다. 본 발명의 화합물 또는 이의 약제학적으로 허용가능한 염은 기존의 미생물 감염성 질환 치료제와 병행하여 투여될 수 있다.

[0059] 본 발명에서의 용어 "예방 또는 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. 이는 또한 미생물 감염성 질환 증상의 감소, 개선, 그 증상의 고통 경감, 미생물 감염성 질환 발생을 감소 또는 치료결과를 증가시키는 환자의 어떠한 다른 변화를 포함한다.

[0060] 본 발명에서 사용된 용어, "투여"는 적절한 방법으로 환자에게 소정의 물질을 도입하는 것을 의미하며, 본 발명의 조성물의 투여 경로는 목적 조직에 도달할 수 있는 한 어떠한 일반적인 경로를 통하여 투여될 수 있다. 복강내 투여, 정맥내 투여, 근육내 투여, 피하 투여, 피내 투여, 경구 투여, 국소 투여, 비내 투여, 폐내 투여, 직장내 투여될 수 있으나, 이에 제한되지는 않는다.

[0061] 본 발명의 약학적 조성물은 약학적으로 유효한 양으로 투여한다. 상기 용어, "약학적으로 유효한 양"은 의학적 치료에 적용 가능한 합리적인 수혜/위험 비율로 질환을 치료하기에 충분하며 부작용을 일으키지 않을 정도의 양을 의미하며, 유효 용량 수준은 환자의 성별, 연령, 체중, 건강상태, 질병의 종류, 중증도, 약물의 활성, 약물에 대한 민감도, 투여 방법, 투여 시간, 투여 경로, 및 배출 비율, 치료기간, 배합 또는 동시에 사용되는 약물을 포함한 요소 및 기타 의학 분야에 잘 알려진 요소에 따라 당업자에 의해 용이하게 결정될 수 있다. 일반적으로, 활성물질을 약 0.01 mg/kg/일 내지 1000 mg/kg/일의 용량으로 투여할 수 있다. 경구 투여하는 경우, 50 내지 500 mg/kg의 범위가 적합할 수 있으며, 1일 1회 이상 투여할 수 있다.

발명의 효과

[0062] 본 발명의 신규한 마크로락탐 글리코시드 유도체를 제조하는 화학효소적 방법은 내재적으로 마크로락톤 글리코시드 유도체를 생산하는 미생물 및 상기 마크로락톤 글리코시드 유도체의 생합성에 관여하는 효소를 이용하여 외부로부터 제공되는 화학적으로 합성된 마크로락탐을 글리코실화함으로써 비천연 마크로락탐 글리코시드 유도체를 제조하는 방법으로서, 본 발명의 방법에 의해 제조된 마크로락탐 글리코시드 유도체는 천연 마크로락톤 글리코시드와 비교하여 동등 이상의 항균활성을 가지며 나아가 개선된 대사 안정성을 나타내므로 신규한 구조의 항균 물질 발굴에 유용하게 사용할 수 있다.

도면의 간단한 설명

[0063] 도 1은 AZDM 1에 대한 역합성 분석 결과를 나타낸 도이다.

도 2는 AZDM 1의 합성 경로를 개략적으로 나타낸 도이다.

도 3은 재조합 S. 베네주엘레 균주의 배양물에 대한 HPLC-ESI-MS 크로마토그램을 나타낸 도이다. (A) 내지 (D)는 각각 변이주 YJ028/pLRHM2의 배양물로부터 화합물 11([M+H]⁺=442) 및 12([M+H]⁺=442), YJ028/pDQNV의 배양물로부터 화합물 12([M+H]⁺=442), YJ028/pDDSS의 배양물로부터 화합물 12([M+H]⁺=442), 13([M+H]⁺=426), 14([M+H]⁺=467), 15([M+H]⁺=453) 및 16([M+H]⁺=424), 및 YJ028/pLOLV2의 배양물로부터 화합물 12([M+H]⁺=442), 17([M+H]⁺=426) 및 18([M+H]⁺=426)이 검출되었음을 나타낸다.

도 4는 (A) 마크로락탐 글리코시드 11-18의 구조 및 (B) 플라스미드에 의해 유도되는 상이한 테옥시당들에 대해 제안된 생합성 경로를 나타낸 도이다. 각 단계를 수행하는 단백질의 기능을 괄호에 표시하였다: NT는 뉴클레오타이드 전달효소, 4,6-DH는 4,6-탈수효소, 2,3-DH는 2,3-탈수효소, AT는 아미노전이효소, DA는 탈아미노효소, EP는 에피머라아제, 4-KR은 4-케토리덕타아제, N-MT는 N-메틸전달효소, 3-KR은 3-케토리덕타아제, SvRed는 S. 베네주엘레 경로 독립적 환원효소를 나타냄.

도 5는 AZDM 글리코시드 11-13, 16 및 17에서의 주요 COSY, HMBC 및 NOESY 상관관계를 나타낸 도이다.

도 6은 에리트로마이신 에스테라아제 어세이 결과를 나타낸 도이다. (A) 내지 (C)는 각각 에스테라아제와 인큐베이션 후 에리트로마이신, L-람노실-10-dm1 및 L-람노실-AZDM(11)의 HPLC-ESI-MS 크로마토그램을 나타낸다: ◇는 선형 에리트로마이신, ●는 에리트로마이신, ▲는 L-람노실-10-dm1 및 ▽는 L-람노실-AZDM에 상응. (D) 및 (E)는 각각 에리트로마이신과 선형 에리트로마이신의 MS 스펙트럼을 나타낸다.

도 7은 인공위액(simulated gastric fluid; SGF) 어세이 결과를 나타낸 도이다. (A) 내지 (C)는 각각 SGF와 인

큐베이션 후 에리트로마이신, L-람노실-10-dm1 및 L-람노실-AZDM(**11**)의 HPLC-ESI-MS 크로마토그램을 나타낸다: ●는 에리트로마이신, ◇는 5-데조자미닐-에리트로노라이드 A, ▲는 L-람노실-10-dm1 및 ▽는 L-람노실-AZDM에 상응. (D) 및 (E)는 각각 에리트로마이신과 5-데조자미닐-에리트로노라이드 A의 MS 스펙트럼을 나타낸다.

도 8은 간 마이크로솜 어세이 결과를 나타낸 도이다. (A) 내지 (C)는 각각 간 마이크로솜과 인큐베이션 후 에리트로마이신, L-람노실-10-dm1 및 L-람노실-AZDM(**11**)의 HPLC-ESI-MS 크로마토그램을 나타낸다: ◇는 N-데스메틸-에리트로마이신, ●는 에리트로마이신, ◎는 히드록실화된 L-람노실-10-dm1, ▲는 L-람노실-10-dm1 및 ▽는 L-람노실-AZDM에 상응. (D) 및 (E)는 각각 N-데스메틸-에리트로마이신과 수산화된 L-람노실-10-dm1의 MS 스펙트럼을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0064] 이하, 실시예를 통하여 본 발명을 보다 상세히 설명하고자 한다. 이들 실시예는 본 발명을 보다 구체적으로 설명하기 위한 것으로, 본 발명의 범위가 이들 실시예에 의해 한정되는 것은 아니다.

[0065] **제조예 1: 아자-(10-데옥시메티노라이드)(aza-(deoxymethinolide); AZDM; 1)의 합성**

[0066] AZDM **1**의 역합성 분석을 도 1에 나타내었다. AZDM **1**을 주요 단계로서 EDC 커플링 및 고리밀폐 상호교환(Ring Closing Metathesis; RCM)을 이용하여 주요 아민 화합물 **2** 및 공지의 카르복시산 **3**으로부터 합성할 수 있었다.

[0067] 본 발명에 따른 AZDM **1**의 합성방법을 도 2에 요약하였다. 프로판알로부터 시작하는 4단계의 입체선택적 방식으로 원하는 아민 화합물 **2**를 준비하였다. 프로판알의 크로틸화(Crotylation)를 성공적으로 수행하여 알콜 화합물 **4**를 수득하였다. 알콜 화합물 **4**의 메실화(mesylation)와 이어지는 소듐 아자이드와의 성공적인 친핵성 치환 및 그 결과로 수득한 아자이드의 LiAlH₄에 의한 환원에 의해 아민 화합물 **2**를 수득하였다. 상기 주요 아민을, EDC 커플링에 의해 카르복시산 **3**과 성공적으로 결합시켜 락탐 유도체 **6**를 수득하였다. 일차 TBS 에테르를 선택적으로 절단하여 알콜 화합물 **7**를 수득하였다. 이어서 DMP로 산화시켜 알데하이드 **8**를 수득하였다. 비닐마그네슘브로마이드와 알데하이드 **8**의 그리나르 반응(Grignard reaction) 후, DMP 산화에 의해 비닐 케톤 **9**를 제조하였다. 2세대 그럽 촉매(second generation Grubbs catalyst)와 함께 RCM에 의해 성공적으로 보호된 마크로락탐 **10**을 제조하였다. TBAF로 TBS 에테르를 절단하여 최종적으로 원하는 AZDM **1**을 수득하였다. 합성된 마크로락탐 **1**의 구조는 MS 데이터 및 1D(¹H- 및 ¹³C-) 및 2D(COSY, HSQC, 및 HMBC) NMR 분석으로 확인하였다.

[0068] **1.1. (3S,4R)-4-메틸-5-헥센-3-올((3S,4R)-4-methyl-5-hexen-3-ol; 4)의 합성**

[0069] THF(15 mL) 용매에 포타슘 tert-부톡사이드(4.6 g, 41.5 mmol)을 가하여 교반한 혼합물에 트랜스-2-부텐(7.2 mL, 80 mmol) 및 n-부틸리튬(26.0 mL의 1.6 M 헥산용액, 41.5 mmol)을 첨가하였다. 용액을 -78℃에서 30분 동안 교반한 후, 상기 용액을 B-메톡시디이소피노캄페일보란(B-methoxydiisopinocampheylborane, 50 mL(1.0 M 에테르 용액), 50 mmol) 용액에 적가하였다. 상기 용액을 -78℃에서 30분 동안 교반하고, BF₃·OEt₂(7.0 mL, 56.0 mmol) 및 프로판알(4.3 mL, 58.5 mmol)을 -78℃에서 연속적으로 적가하였다. 상기 용액을 -78℃에서 3시간 동안 교반한 후 3 N NaOH(30 mL) 및 30% H₂O₂(12 mL)로 처리하였다. 상기 혼합물을 1시간 동안 가열하여 환류하고 에테르(3×40 mL)로 추출하였다. 유기상을 분리하여 물(100 mL)과 염수(100 mL)로 세척하였다. 유기층을 건조(MgSO₄)하여 농축하였다. 플래쉬 크로마토그래피(헥산/EtOAc = 7:1)로 정제하여 무색 오일로서 알콜 화합물, **4**(2.8 g, 61%)를 수득하였다.

[0070] $[\alpha]_D^{30} +12.5$ (c 0.88, CHCl₃);

[0071] IR (film) ν_{max} 3363, 2963, 2932, 2875, 1640, 1460, 1109, 968 cm⁻¹;

[0072] ¹H NMR (400 MHz, CDCl₃) δ 5.76 (ddd, J = 8.3, 11.3, 16.3 Hz, 1H), 5.10 (m, 2H), 3.33 (m, 1H), 2.22 (m,

1H), 1.60 (m, 1H), 1.40 (m, 1H), 1.03 (d, $J = 6.8$ Hz, 3H), 0.97 (t, $J = 7.4$ Hz, 3H);

[0073] ^{13}C NMR (100 MHz, CDCl_3) δ 140.4, 116.2, 76.0, 43.7, 27.0, 16.3, 10.0;

[0074] HR-ESI-MS m/z 137.0936 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_7\text{H}_{14}\text{ONa}$, 137.0937).

[0075] **1.2. (3S,4R)-3-메탄술폰닐옥시-4-메틸-5-헥센((3S,4R)-3-methanesulfonyloxy-4-methyl-5-hexen; 5)의 합성**

[0076] CH_2Cl_2 (25 mL)에 용해시킨 상기 제조예 1.1로부터 수득한 알콜 화합물 **4**(780 mg, 6.8 mmol)의 교반한 용액에 0℃에서 트리에틸아민(3.3 mL, 23.9 mmol) 및 메탄술폰닐클로라이드(1.6 mL, 20.5 mmol)를 첨가하였다. 상기 반응 혼합물을 0℃에서 20분 동안 교반한 후 실온까지 가열하였다. 실온에서 2시간 더 교반한 후, 물(30 mL)을 첨가하고, 혼합물을 CH_2Cl_2 (3×30 mL)로 추출하였다. 유기층을 분리하여 건조(MgSO_4)시켜 농축하였다. 플래쉬 크로마토그래피(헥산/EtOAc = 7:1)로 정제하여 무색 오일로서 표제 화합물 메탄술폰네이트 에스테르, **5**(1.16 g, 89%)를 수득하였다.

[0077] $[\alpha]_D^{29} -2.25$ (c 1.59, CHCl_3);

[0078] IR (film) ν_{max} 2971, 2879, 1641, 1463, 1418, 1341, 1175, 1048, 973 cm^{-1} ;

[0079] ^1H NMR (400 MHz, CDCl_3) δ 5.74 (m, 1H), 5.11 (m, 2H), 4.62 (ddd, $J = 4.9, 6.2, 11.0$ Hz, 1H), 3.02 (s, 3H), 2.60 (m, 1H), 1.71 (m, 2H), 1.10 (d, $J = 6.9$ Hz, 3H), 0.98 (t, $J = 7.4$ Hz, 3H);

[0080] ^{13}C NMR (100 MHz, CDCl_3) δ 138.4, 116.5, 88.1, 41.1, 38.7, 24.6, 15.7, 9.7;

[0081] HR-ESI-MS m/z 215.0711 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_8\text{H}_{16}\text{O}_3\text{SNa}$, 215.0712).

[0082] **1.3. (3R,4R)-3-아민-4-메틸-5-헥센((3R,4R)-3-amine-4-methyl-5-hexene; 2)의 합성**

[0083] HMPA(hexamethylphosphoramide, 10 mL)에 용해시킨 상기 제조예 1.2.로부터 수득한 메탄술폰네이트 에스테르, **5**(1.16 g, 6.03 mmol) 용액에 실온에서 소듐 아자이드(1.57 g, 24.1 mmol)를 첨가하였다. 상기 용액을 40℃까지 가열하고 해당 온도에서 2시간 동안 교반하였다. 용액을 실온까지 냉각시킨 후, 물(20 mL)을 첨가하였다. 혼합물을 Et_2O (3×20 mL)로 추출하고, 유기상을 분리하여 물(40 mL)과 염수(40 mL)로 세척하였다. 유기층을 건조(MgSO_4)시키고, 농축하였다. 잔여물을 THF(30 mL)에 용해시키고 0℃에서 일부를 LiAlH_4 (7.2 mL(1.0 M THF 용액), 7.2 mmol) 및 에탄올(0.42 mL, 7.2 mmol)의 THF(10 mL) 용액에 서서히 가하였다. 상기 용액을 0℃에서 10분 동안 교반한 후 실온까지 가열하였다. 실온에서 2시간 더 교반한 후, 포화 Na_2SO_4 수용액(0.6 mL), 및 무수 Na_2SO_4 (1.0 g)을 상기 용액에 첨가하였다. 혼합물을 농축하고 플래쉬 크로마토그래피(메탄올/EtOAc = 1:10)로 정제하여 무색 액체로서 표제 아민 화합물 **2**(670 mg, 98%)를 수득하였다.

[0084] $[\alpha]_D^{29} +18.8$ (c 1.56, CHCl_3);

[0085] IR (film) ν_{max} 3341, 2936, 1640, 1544, 1460, 1380, 1053 cm^{-1} ;

[0086] ^1H NMR (400 MHz, CDCl_3) δ 5.77 (ddd, $J = 7.3, 9.9, 17.5$ Hz, 1H), 4.99 (m, 2H), 2.55 (ddd, $J = 4.7, 4.7, 8.9$ Hz, 1H), 2.29 (br s, 2H), 2.18 (m, 1H), 1.48 (m, 1H), 1.24 (m, 1H), 0.94 (d, $J = 6.9$ Hz, 3H), 0.89 (t, $J = 7.4$ Hz, 3H);

- [0087] ^{13}C NMR (100 MHz, CDCl_3) δ 141.5, 114.6, 56.7, 42.5, 26.8, 14.1, 10.8;
- [0088] HR-ESI-MS m/z 114.1279 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_7\text{H}_{16}\text{N}$, 114.1277).
- [0089] 1.4. (2R,3S,4S,6R)-3,7-비스(*tert*-부틸디메틸실릴옥시)-2,4,6-트리메틸-N-((3R,4R)-4-메틸헥스-5-엔-3-일)헵탄아미드((2R,3S,4S,6R)-3,7-bis(*tert*-butyldimethylsilyloxy)-2,4,6-trimethyl-N-((3R,4R)-4-methylhex-5-en-3-yl)heptanamide; 6)의 합성
- [0090] CH_2Cl_2 (10 mL)에 용해시킨 카르복시산 3(450 mg, 1.04 mmol) 용액에 DIPEA(233 μL , 1.35 mmol), HOBt(182 mg, 1.35 mmol), 및 EDC·HCl(259 mg, 1.35 mmol)를 첨가하였다. 상기 혼합물을 0℃에서 15분 동안 교반하고 CH_2Cl_2 (1 mL)에 용해시킨 상기 제조예 1.3.으로부터 수득한 아민 화합물 2(235 mg, 2.08 mmol) 용액을 첨가하였다. 상기 혼합물을 0℃에서 2시간 동안 교반한 후 실온까지 가열하였다. 실온에서 16시간 더 교반한 후, 포화 NH_4Cl (10 mL)을 상기 용액에 첨가하고 혼합물을 CH_2Cl_2 (3×10 mL)로 추출하였다. 유기층을 분리하여 건조(MgSO_4)시키고 농축하였다. 플래쉬 크로마토그래피(헥산/EtOAc = 7:1)로 정제하여 무색 오일로서 표제 아미드 화합물 6(445 mg, 81%)을 수득하였다.
- [0091] $[\alpha]_D^{25} +4.6$ (c 1.41, CHCl_3);
- [0092] IR (film) ν_{max} 3306, 2958, 2858, 1639, 1540, 1464, 1383, 1256, 1074, 913 cm^{-1} ;
- [0093] ^1H NMR (400 MHz, CDCl_3) δ 5.73 (ddd, J = 7.9, 9.8, 17.7 Hz, 1H), 5.49 (bd, J = 9.4 Hz, 1H), 5.01 (m, 2H), 3.81 (m, 1H), 3.49 (dd, J = 4.8, 9.7 Hz, 1H), 3.24 (dd, J = 7.2, 9.7 Hz, 1H), 2.40 (m, 1H), 2.31 (m, 1H), 1.77~1.57 (m, 4H), 1.47 (m, 1H), 1.21 (m, 1H), 1.15 (d, J = 7.1 Hz, 3H), 1.02 (d, J = 6.9 Hz, 3H), 0.89 (m, 27H), 0.05 (m, 12H);
- [0094] ^{13}C NMR (100 MHz, CDCl_3) δ 174.9, 140.6, 115.2, 77.6, 67.9, 54.0, 44.8, 42.0, 36.5, 36.2, 33.8, 26.0, 25.9, 24.4, 18.4, 18.3(×2), 16.9, 16.3, 15.7, 10.7, -4.1, -5.4;
- [0095] HR-ESI-MS m/z 550.4087 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{29}\text{H}_{61}\text{NO}_3\text{Si}_2\text{Na}$, 550.4082).
- [0096] 1.5. (2R,3S,4S,6R)-3-(*tert*-부틸디메틸실릴옥시)-7-히드록시-2,4,6-트리메틸-N-((3R,4R)-4-메틸헥스-5-엔-3-일)헵탄아미드((2R,3S,4S,6R)-3-(*tert*-butyldimethylsilyloxy)-7-hydroxy-2,4,6-trimethyl-N-((3R,4R)-4-methylhex-5-en-3-yl)heptanamide; 7)의 합성
- [0097] 상기 제조예 1.4.로부터 수득한 아미드 화합물 6(445 mg, 0.84 mmol)을 0℃에서 메탄올(10 mL)에 용해시켰다. 상기 용액에 DL-10-캠포술포산(58 mg, 0.25 mmol)을 첨가하였다. 상기 용액을 0℃에서 1시간 동안 교반하였다. Et_3N (200 μL , 1.43 mmol)을 첨가하여 상기 반응을 종결시켰다. 용액을 농축시키고 플래쉬 크로마토그래피(헥산/EtOAc = 4:1)로 정제하여 무색 오일로서 표제 일차 알콜 화합물 7(297 mg, 86%)을 수득하였다.
- [0098] $[\alpha]_D^{31} +6.8$ (c 1.05, CHCl_3);
- [0099] IR (film) ν_{max} 3298, 2960, 2931, 1643, 1546, 1461, 1380, 1254, 1069, 911 cm^{-1} ;
- [0100] ^1H NMR (400 MHz, CDCl_3) δ 5.72 (m, 1H), 5.58 (bd, J = 9.3 Hz, 1H), 5.01 (m, 2H), 3.77 (m, 2H), 3.54 (dd, J = 4.3, 10.9 Hz, 1H), 3.33 (dd, J = 5.4, 10.8 Hz, 1H), 2.76 (bs, 1H), 2.43 (m, 1H), 2.32 (m,

1H), 1.81 (m, 1H), 1.63 (m, 3H), 1.24 (m, 1H), 1.16 (d, $J = 7.0$ Hz, 3H), 1.03 (d, $J = 6.9$ Hz, 3H), 0.93 (m, 15H), 0.86 (t, $J = 7.4$ Hz, 3H), 0.08 (s, 6H);

[0101] ^{13}C NMR (100 MHz, CDCl_3) δ 175.4, 140.5, 115.5, 77.7, 66.6, 54.3, 45.3, 41.8, 35.2, 34.8, 32.8, 26.0, 24.2, 18.3, 17.7, 16.6, 16.1, 10.5, -4.0, -4.2;

[0102] HR-ESI-MS m/z 436.3219 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{23}\text{H}_{47}\text{NO}_3\text{SiNa}$, 436.3217).

[0103] 1.6. (2R,3S,4S,6R)-3-((tert-부틸디메틸실릴)옥시)-2,4,6-트리메틸-N-((3R,4R)-4-메틸헥스-5-엔-3-일)-7-옥소헵탄아미드((2R,3S,4S,6R)-3-((tert-butyltrimethylsilyl)oxy)-2,4,6-trimethyl-N-((3R,4R)-4-methylhex-5-en-3-yl)-7-oxoheptanamide; 8)의 합성

[0104] CH_2Cl_2 (20 mL)에 용해시킨 상기 제조예 1.5.으로부터 수득한 알콜 화합물 7(297 mg, 0.72 mmol) 용액에 데스-마틴 퍼아이오딘산(Dess-Martin periodinane; DMP, 611 mg, 1.44 mmol)을 첨가하였다. 상기 용액을 실온에서 2시간 동안 교반한 후 포화 NaHCO_3 (10 mL) 및 Na_2SO_4 (5 mL)를 첨가하였다. 혼합물을 CH_2Cl_2 (3×15 mL)로 추출하고, 유기층을 분리, 건조(MgSO_4)하여 농축시켰다. 플래쉬 크로마토그래피(헥산/EtOAc = 5:1)로 정제하여 무색 액체로서 표제 알데하이드 화합물 8(246 mg, 83%)을 수득하였다.

[0105] $[\alpha]_D^{29} +29.4$ (c 1.70, CHCl_3);

[0106] IR (film) ν_{max} 3313, 2931, 1726, 1641, 1537, 1461, 1254, 1066, 836 cm^{-1} ;

[0107] ^1H NMR (400 MHz, CDCl_3) δ 9.59 (d, $J = 1.9$ Hz, 1H), 5.71 (m, 1H), 5.64 (bd, $J = 9.4$ Hz, 1H), 5.01 (m, 2H), 3.80 (m, 2H), 2.46 (m, 2H), 2.31 (m, 1H), 1.93 (ddd, $J = 3.5, 9.5, 13.8$ Hz, 1H), 1.59 (m, 3H), 1.25 (m, 1H), 1.16 (d, $J = 7.1$ Hz, 3H), 1.11 (d, $J = 7.2$ Hz, 3H), 0.99 (d, $J = 6.9$ Hz, 3H), 0.93 (m, d, $J = 6.9$ Hz, 3H), 0.88 (m, 12H), 0.08 (s, 3H), 0.07 (s, 3H);

[0108] ^{13}C NMR (100 MHz, CDCl_3) δ 205.4, 174.9, 140.8, 115.0, 54.2, 45.4, 44.3, 41.8, 36.2, 32.8, 26.1, 24.5, 18.3, 17.2, 16.2, 16.0, 14.9, 10.7, -3.9, -4.0;

[0109] HR-ESI-MS m/z 434.3062 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{23}\text{H}_{45}\text{NO}_3\text{SiNa}$, 434.3061).

[0110] 1.7. (2R,3S,4S,6R)-3-((tert-부틸디메틸실릴)옥시)-2,4,6-트리메틸-N-((3R,4R)-4-메틸헥스-5-엔-3-일)-7-옥소논-8-엔아미드((2R,3S,4S,6R)-3-((tert-butyltrimethylsilyl)oxy)-2,4,6-trimethyl-N-((3R,4R)-4-methylhex-5-en-3-yl)-7-oxonon-8-enamide; 9)의 합성

[0111] THF(10 mL)에 용해시킨 상기 제조예 1.6.으로부터 수득한 알데하이드 화합물 8(246 mg, 0.60 mmol)의 교반한 용액에 0°C에서 비닐마그네슘브로마이드(1 M THF 용액)(1.20 mL, 1.20 mmol)을 첨가하였다. 1시간 후, Et_2O (10 mL)를 첨가하여 상기 용액을 희석하고 포화 NH_4Cl 수용액(10 mL)을 첨가하였다. 유기층을 분리하고 수층을 Et_2O (3×10 mL)로 추출하였다. 수득한 유기층을 합하여 건조(MgSO_4)시키고 농축하였다. 잔여물을 플래쉬 크로마토그래피(헥산/EtOAc = 5:1)로 정제하여 무색 오일로서 표제 비닐 알콜 화합물(201 mg, 76%)을 수득하였다.

[0112] 상기 수득한 비닐 알콜 화합물(201 mg, 0.46 mmol)을 CH_2Cl_2 (10 mL)에 용해시키고, DMP(390 mg, 0.92 mmol)를 첨가한 후, 혼합 용액을 실온에서 2시간 동안 교반하였다. 반응이 완결된 후, 포화 NaHCO_3 수용액(10 mL) 및 Na_2SO_4 (5 mL)를 첨가하였다. 혼합물을 CH_2Cl_2 (3×10 mL)로 추출하였다. 유기층을 분리, 건조(MgSO_4)하여 농축시

켰다. 플래쉬 크로마토그래피(헥산/EtOAc = 5:1)로 정제하여 무색 액체로서 표제 비닐 케톤 화합물 **9**(145 mg, 72%)를 수득하였다.

[0113] $[\alpha]_D^{28} +32.4$ (c 0.53, CHCl_3);

[0114] IR (film) $\nu_{\max}$ 3302, 2928, 2358, 1726, 1645, 1532, 1461, 1379, 1254, 1064, 910 cm^{-1} ;

[0115] ^1H NMR (400 MHz, CDCl_3) δ 6.44 (dd, J = 10.4, 17.5 Hz, 1H), 6.28 (dd, J = 1.4, 17.5 Hz, 1H), 5.80 (m, 2H), 5.67 (ddd, J = 8.0, 10.3, 17.2 Hz, 1H), 4.98 (m, 2H), 3.74 (m, 2H), 2.93 (m, 1H), 2.61 (m, 1H), 2.21 (m, 1H), 2.02 (m, 1H), 1.27 (m, 2H), 1.15 (m, 6H), 0.91 (m, 18H), 0.08 (s, 3H), 0.07 (s, 3H);

[0116] ^{13}C NMR (100 MHz, CDCl_3) δ 204.7, 175.4, 141.1, 135.6, 128.3, 114.7, 77.6, 54.4, 45.7, 42.1, 41.2, 36.2, 34.1, 26.2, 24.8, 18.9, 18.4, 17.8, 16.6, 16.0, 10.6, -3.7, -3.8;

[0117] HR-ESI-MS m/z 460.3219 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{25}\text{H}_{47}\text{NO}_3\text{SiNa}$, 460.3217).

[0118] 1.8. (E)-(3R,4S,5S,7R,11R,12R)-4-((tert-부틸디메틸실릴)옥시)-12-에틸-3,5,7,11-테트라메틸아자사이클로도덱-9-엔-2,8-디온((E)-(3R,4S,5S,7R,11R,12R)-4-((tert-butyl dimethylsilyl)oxy)-12-ethyl-3,5,7,11-tetramethylazacyclododec-9-ene-2,8-dione; **10**)의 합성

[0119] 상기 실시예 1.7.로부터 수득한 비닐 케톤 화합물 **9**(145 mg, 0.33 mmol)를 실온에서 CH_2Cl_2 (20 mL)에 용해시켰다. 그럽스 촉매(Grubbs' catalyst, 2세대)(56 mg, 20 mol%)를 첨가하고 실온에서 19시간 동안 교반하여 밝은 갈색 액체를 형성하였다. 상기 혼합물을 농축하고 잔여물을 플래쉬 크로마토그래피(헥산/EtOAc = 4:1)로 정제하여 흰색 고체로서 생성물 **10**(83 mg, 61%)을 수득하였다.

[0120] mp: 184-186 $^{\circ}\text{C}$;

[0121] $[\alpha]_D^{29} +35.9$ (c 0.69, CHCl_3);

[0122] IR (film) $\nu_{\max}$ 3300, 2928, 1735, 1691, 1632, 1544, 1462, 1379, 1255, 1092, 1050, 1023, 979 cm^{-1} ;

[0123] ^1H NMR (400 MHz, CDCl_3) δ 6.73 (dd, J = 5.3, 15.9 Hz, 1H), 6.32 (dd, J = 1.3, 15.9 Hz, 1H), 5.23 (bd, J = 10.3 Hz, 1H), 4.09 (m, 1H), 3.64 (d, J = 8.8 Hz, 1H), 2.73 (m, 1H), 2.52 (m, 1H), 2.30 (m, 1H), 1.48 (m, 3H), 1.36 (m, 2H), 1.21 (d, J = 6.9 Hz, 6H), 1.00 (d, J = 7.0 Hz, 3H), 0.95 (m, 3H), 0.92 (m, 3H), 0.90 (s, 9H), 0.08 (s, 3H), 0.07 (s, 3H);

[0124] ^{13}C NMR (100 MHz, CDCl_3) δ 205.1, 174.8, 147.3, 125.6, 79.9, 77.2, 50.5, 46.8, 44.7, 37.7, 34.5, 33.8, 29.7, 26.6, 26.3, 18.5, 17.8, 11.0, 9.6, -3.2, -3.5;

[0125] HR-ESI-MS m/z 432.2903 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{23}\text{H}_{43}\text{NO}_3\text{SiNa}$, 432.2904).

[0126] 1.9. (3R,4S,5S,7R,9E,11R,12R)-12-에틸-4-히드록시-3,5,7,11-테트라메틸아자사이클로도덱-9-엔-2,8-디온((3R,4S,5S,7R,9E,11R,12R)-12-ethyl-4-hydroxy-3,5,7,11-tetramethylazacyclododec-9-ene-2,8-dione; aza-10-deoxymethynolide, AZDM; **1**)의 합성

[0127] 실온에서 건조 THF(10 mL)에 용해시킨 상기 실시예 1.8.로부터 수득한 생성물 **10**(83 mg, 0.20 mmol)의 교반한 용액에 1.0 M TBAF(tetrabutylammonium fluoride, 400 μL , 0.40 mmol)를 첨가하였다. 3시간 후, 반응 혼합물

을 농축하였다. 플래쉬 크로마토그래피(헥산/EtOAc = 1:2)로 정제하여 흰색 고체로서 표제 알콜 화합물 **1**(46 mg, 78%)을 수득하였다.

[0128] mp: 166-168 °C;

[0129] $[\alpha]_D^{28} +42.8$ (c 0.31, CHCl_3);

[0130] IR (film) ν_{max} 3280, 2925, 2854, 1740, 1687, 1631, 1545, 1459, 1382, 1254, 1152, 979 cm^{-1} ;

[0131] ^1H NMR (500 MHz, CD_3OD) δ 6.72 (dd, J = 16, 6Hz, 1H), 6.42 (d, J = 16 Hz, 1H), 4.00 (m, 1H), 3.42 (d, J = 10 Hz, 1H), 2.73 (m, 1H), 2.50 (m, 1H), 2.48 (m, 1H), 1.65 (t, J = 13.5 Hz, 1H), 1.53 (m, 2H), 1.35 (td, J = 12.5, 4 Hz, 1H), 1.29 (m, 1H), 1.25 (d, J = 7.0 Hz, 3H), 1.23 (d, J = 7.0 Hz, 3H), 1.06 (d, J = 6.5 Hz, 3H), 0.94 (d, J = 6.5 Hz, 3H), 0.93 (t, J = 7.5 Hz, 3H);

[0132] ^{13}C NMR (125 MHz, CD_3OD) δ 207.8, 177.9, 149.7, 126.5, 79.2, 52.0, 46.7, 45.8, 39.9, 34.7, 34.3, 27.2, 17.8, 17.8, 17.7, 11.6, 10.1;

[0133] (+)-ESI-MS m/z 318.0 $[\text{M}+\text{Na}]^+$, 313.0 $[\text{M}+\text{NH}_4]^+$, 296.0 $[\text{M}+\text{H}]^+$, 277.9 $[\text{M}+\text{H}-\text{H}_2\text{O}]^+$, 259.9 $[\text{M}+\text{H}-2\text{H}_2\text{O}]^+$,

[0134] HR-ESI-MS m/z 318.2041 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{17}\text{H}_{29}\text{NO}_3\text{Na}$, 318.2040).

[0135] 제조예 2: 박테리아 균주, 배양 조건 및 유전자 조작

[0136] SPA 아가(1 g 효모 추출물, 1 g 비프 추출물, 2 g 트립토스, 10 g 글루코스, 미량의 황산제1철(ferrous sulfate), 및 15 g 아가/L) 상에 *S. 베네주엘레* 변이균주를 번식시켰다. 원형질체 준비(protoplast preparation) 및 *S. 베네주엘레*의 형질전환은 표준 프로토콜에 따라 수행하였다(Xuan, R. *et al.*, *J. Org. Chem.*, 2008, 73: 1456-1461). *S. 베네주엘레*의 형질전환주(transformants)를 선택하기 위하여, 티오스테렙톤(thiostrepton)을 보충한 R2YE 아가 플레이트를 사용하였다. 유전체 DNA(genomic DNA)를 얻기 위하여 *S. 베네주엘레* ATCC 15439 및 스트렙토마이세스 안티바이오티쿠스(antibioticus) ATCC 11891를 사용하였다. 플라스미드 리트머스28(New England Biolabs) 및 에스케리키아 콜라이 DH5 α (Invitrogen)를 이용하여 통상의 서브클로닝을 수행하였다. 테옥시당류 및 이들의 전달에 관여하는 유전자들을 발현하기 위하여, 구성 ermE* 프로모터 및 티오 스트렙톤 저항성 마커를 포함하는 고복제수(high-copy-number) E. 콜라이-스트렙토마이세스 셔틀 벡터, pSE34를 사용하였다.

[0137] 제조예 3: 발현 플라스미드 및 *S. 베네주엘레* 변이균주의 구축

[0138] 특정한 테옥시올리고뉴클레오티드 프라이머와 주형 DNA를 이용한 PCR로 서너가지 테옥시당 생합성 유전자를 포함하는 DNA 단편을 증폭시켰다. *Pfu* 중합효소(Fermentas GmbH, St. Leon-Rot, Germany)를 이용하여 제조업자에 의해 제시된 조건 하에서 PCR을 수행하였다. 다양한 테옥시당류와 아자-10-테옥시메티노라이드(AZDM; **1**)의 글리코실화에 관여하는 유전자들을 복제 플라스미드 pSE34 내에 클로닝하였다. 상이한 유전자들을 함유하는 플라스미드를 효과적으로 구축하기 위하여, 하나의 유전자를 포함하는 PacI/XbaI 단편을 다른 유전자를 함유한 PacI/SpeI-절단된(digested) 리트머스28에 결합시켰다. **1**의 당화된 유도체들로의 전환에 관여하는 모든 유전자들이 리트머스28에 결합될 때까지 상기 과정을 회귀적으로 반복한 후, PacI/XbaI로 절단된 pSE34 내에 옮겼다. 표 1에 조합된 유전자 조직을 나타내었다. pDDSS, pDQNV, pLRAM2, 및 pLOLV2의 구축은 이전에 공지된 방법을 이용하였다(Shinde, P. B. *et al.*, *J. Biotech.*, 2013, 168: 142-148; Han, A. R. *et al.*, *Appl. Microbiol. Biotechnol.*, 2012, 93: 1147-1156). 이들 플라스미드를 *S. 베네주엘레* YJ028 내에 각각 도입하고 결과로 얻은 재조합체들(recombinants)을 합성된 마크로락탐 **1**의 생체 내 글리코실화 반응에 사용하였다.

표 1

[0139]

플라스미드	유전자 조합	생성물
pLRHAM2	<i>desVIII-desVII-desIII-desIV-oleL-oleU</i>	11, 12
pDQNV	<i>desVIII-desVII-desIII-desIV</i>	12
pDDSS	<i>desVIII-desVII-desIII-desIV-desI-desII-desV-desVI</i>	12, 13, 14, 15, 16
pLOLV2	<i>desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU</i>	12, 17, 18

[0140]

실시예 1: AZDM 글리코시드의 생산 및 분석

[0141]

재조합주 YJ028/pDDSS, YJ028/pDQNV, YJ028/pLRAM2, 및 YJ028/pLOLV2를 25 μ g/mL 농도로 항생제 티오스트렙톤(antibiotic thiostrepton)을 함유하는 SGGP 액체 배지(트립톤 4 g, 효모추출물 4 g, MgSO₄ 0.5 g, 글루코스 10 g, 글라이신 2 g, 및 인산칼륨 완충액 0.01 M, 리터 당)에 30°C에서 24시간 동안 배양하였다. 항생제 티오스트렙톤(25 μ g/mL)이 첨가된 50 mL SCM 배지(1.5% 가용 녹말, 2.0% 소이톤(soytone), 0.01% CaCl₂, 0.15% 효모추출물, 및 1.05% MOPS, pH 7.2)가 담긴 250 mL 흠이 있는 삼각 플라스크(baffled Erlenmeyer flask)에 상기 종배양(5 mL)을 접종하였다. 초기에 30°C에서 2일 동안 인큐베이션하고 합성한 AZDM 1을 5 μ g/mL 농도로 공급한 후, 2일 더 인큐베이션하였다. 각각의 성장된 배양물을 원심분리(5,000×g, 10분)하여 상층액을 5 mL 메탄올과 5 mL 물로 미리 처리된(conditioned) OASIS HLB (Waters, Milford, MA) SPE 컬럼에 통과시켰다. 컬럼은 5 mL 20% (vol/vol) 메탄올로 세척하고 약 30초 동안 자연건조하였다(air-dried). 컬럼을 1 mL 0.5% (vol/vol) 메탄올 아세트산으로 3회 용출하고 유기 추출물을 감압 농축하였다. 250 μ L/min 유속으로 HPLC-ESI-MS 분석을 수행하였다. 구배를 사용하여 5 mM 암모늄 아세테이트의 농도 구배를 이용하여-0.05% (vol/vol) 아세트산의 수용액(용액 A) 및 동일한 추가 농도를 갖는 80% (vol/vol) 아세토니트릴 수용액(aqueous acetonitrile)(용액 B)으로, 25분 동안 20 내지 70% 용액 B까지, 15분 동안 90% 용액 B까지, 90% 용액 B에서 9분 동안 유지하고, 컬럼 재평형화를 위하여 추가로 11분 동안 20% 용액 B까지-용출하였다.

[0142]

구체적으로, AZDM 1를 당화하기 위하여, 비천연 TDP-데옥시당의 생합성 및 이에 이어지는 외부로부터 제공된 AZDM(1)과 이들 당의 당화를 수행하는 4종의 미리 구축한 발현 플라스미드를 독립적으로 변이균주 S. 베네주엘레 YJ028에 도입하였다. 이들 플라스미드는 pDDSS(글리코실전달효소 *desVII-desVIII*와 함께 TDP-D-데조자민 생합성 유전자를 발현), pLRHM2(*desVII-desVIII*와 함께 TDP-L-람노스 생합성 유전자를 발현), pDQNV(*desVII-desVIII*와 함께 TDP-D-퀴노보스 생합성 유전자를 발현), 및 pLOLV2(*desVII-desVIII*와 함께 TDP-L-올리보스 및 TDP-L-디지톡소스(digitoxose) 생합성 유전자를 발현)(표 1)를 포함한다. 변이균주들을 적절한 항생제를 선택하여 50 mL SCM 액체 배지에 30°C에서 48시간 동안 배양하였다. AZDM 1(5 mg/L)을 각 성장배지에 첨가하고 추가로 48시간 동안 배양하였다. 배양액의 유기추출물을 준비하여 HPLC-ESI-MS 분석을 수행하였다. m/z 296, 278, 260에서 AZDM 1로부터 발생하는 공통적인 특성 단편에 대한 HPLC 크로마토그램을 모니터링하고, 이로부터 당화된 AZDMs에 대한 피크들을 동정하였다(도 3).

[0143]

L-람노스 생합성 유전자를 발현하는 변이균주 YJ028/pLRHM2는 상대적으로 낮은 수준으로 L-람노실-AZDM(**11**)(도 3A) 및 D-퀴노보실-AZDM(**12**)(도 3A)를 생산하였다. HPLC-ESI-MS 분석에서, 글리코시드 **11** 및 **12**는 각각 15.2분 및 16.5분에서 용출되었고(도 3A), 이들 화합물 모두는 m/z 442에서 나타나는 이들의 양성자 부가 이온과 유사한 MS 스펙트럼을 나타내었다. 이들 화합물의 MS/MS 분석은 모피크와 중성 당 손실에 따른 딸피크 간의 차이를 갖는 특징적인 단편 패턴을 형성하였다. 마크로락탐 글리코시드 **12**는 이미 관찰된 S. 베네주엘레 경로 독립적 4-케토리덕타아제(Sv 4-KR)의 결과로서 발생할 수 있다. Sv 4-KR는 TDP-L-람노스의 생합성 중간체인 TDP-4-케토-6-데옥시-D-글루코스(**21**)를, 이후 *DesVII/DesVIII*에 의해 AZDM(1)에 부착되어 **12**를 제공하는, TDP-D-퀴노보스(**24**)로 환원시킬 수 있다(도 4). 예상되는 바와 같이, 마크로락탐 글리코시드 **12**는 또한 D-퀴노보스 생합성 유전자를 포함하는 변이균주 YJ028/pDQNV에 의해 생산될 수 있다(도 3B). D-데조자민 생합성 유전자를 가진 변이균주 YJ028/pDDSS의 유기추출물의 HPLC-ESI-MS 분석은 양성자 부가 이온 m/z 442, 426, 467, 453, 및 424에 의해 각각 머무름시간 16.5, 18.0, 19.7, 23.2, 및 23.7분에서 5개의 피크를 나타내었으며(도 3C), 이들은 각각 마크로락탐 글리코시드 **12**, 3-O-데메틸-D-칼코실-AZDM(**13**), N-데디메틸-N-아세틸-D-데조자미닐-AZDM(**14**), D-데조자미닐-AZDM(**15**), 및 3-케토-4,6-디데옥시-D-글루코실-AZDM(**16**)으로 동정되었다(도 4A). AZDM으로부터 발생한

공통적인 단편 이온들과 더불어, 마크로락탐 글리코시드 **14** 및 **15**의 양이온 부가 이온들 m/z 467 및 m/z 453은 각각 m/z 172 및 m/z 158에서 단편 이온을 나타내었으며, 이로부터 TDP-당류 **29** 및 **30**을 그들 각각의 AZDM 글리코시드에 전달함을 확인하였다. 모피크와 중성 당의 손실에 상응하는 딸피크 간의 차이에 따른 단편화 패턴에 기초하여 글리코시드 **13** 및 **16**의 구조를 예측하였다. TDP-3-케토-4,6-디데옥시 헥소스(**26**)는 S. 베네주엘레 경로 독립적 3-케토리덕타아제(Sv 3-KR)의 작용에 의해 TDP-3-O-데메틸 D-갈코스(**27**)로 전환될 수 있다(도 4B). 이어지는 AZDM **1**의 당류 **26**, **29**, 및 **27**과의 당화반응은 각각 글리코시드 **16**, **14**, 및 **13**을 생성하였다. 상기 S. 베네주엘레 YJ028/pDDSS 균주에서 미량의 N-아세틸화된-아미노데옥시 헥소스(**29**)를 포함하는 글리코시드 **14**의 검출은 자기보호 반응으로서 N-아세틸화를 제시하였던 이전의 연구결과를 뒷받침하는 것이다. 변이균주 YJ028/pLOLV2는 16.5, 16.7, 및 17.7분의 머무름시간에서 AZDM 글리코시드로부터 발생한 3개의 피크를 나타내었다(도 3D). 이들은 각각 m/z 442, 426, 및 426에서 양성자 부가 이온을 나타내었으며, 이들 각각의 마크로락탐 글리코시드로 동정되었다: 마크로락탐 글리코시드 **12**, D-올리보실-AZDM(**17**), 및 L-올리보실-AZDM(**18**). 당 **33**을 함유하는 글리코시드 **17**의 검출은 중간체 TDP-4-케토-당 **32**에 대한 Sv 4-KR의 작용을 제안하였다(도 4B). S. 베네주엘레 변이균주의 PikC P450에 의해 촉진되는 AZDM 글리코시드의 추가적인 히드록실화는 검출되지 않았다. 이는 활성부위고정(active site anchoring)을 위해 요구되는 N,N-디메틸아미노기가 결여된 변화된 당류를 갖는 마크로락탐 고리를 수용하는 PikC P450 히드록실라라제의 불능에 기인할 수 있다.

[0144] 실시예 2: AZDM 글리코시드의 추출 및 분리

[0145] Brucker Avance 900 MHz, Varian INOVA 500 MHz, 및 Brucker Avance III 400 MHz 분광기로 NMR 스펙트럼을 기록하였다. 내부 참조(internal reference)로서 테트라메틸실란(tramethylsilane; TMS)을 사용하여 ppm 단위로 화학적 이동을 기록하였다. NMR 용매로서 CD_3OD 및 $CDCl_3$ 를 사용하였으며, 화학적 이동 값들은 $^1H/^{13}C$ 핵들에 대해 각각 3.31/49.1 ppm 및 7.27/77.0 ppm로 보정하였다. NMR 데이터 처리에는 Mestrelab Research S.L.의 Mnova 소프트웨어를 이용하였다. 분석용 역상 Nova-Pak C18 컬럼(Waters, Milford, MA, USA; 150×3.9 mm, $4.0 \mu m$)을 이용하여 Micromass Quattro micro MS에 직접 연결된 Waters 2695 분리모듈로 구성된 Waters/Micromass Quattro micro MS interface 상에 HPLC-ESI-MS/MS 스펙트럼을 기록하였다. HR-ESI-MS 데이터는 UPLC와 결합된 Waters SYNAPT G2-S 질량 분광기를 이용하여 획득하였다. 220 nm에 설정된 UV730D UV 검출기와 결합된 SP930D 구매 펌프로 구성된 Acme 9000 HPLC 시스템(YL Instrument Co. Ltd., Korea) 상에서 분취용 Spherisorb S5 ODS2 (Waters, 250×20 mm, $5 \mu m$) 및 준분취용 Watchers 120 ODS-BP (250×10 mm, $5 \mu m$) 컬럼을 이용하여 HPLC 정제를 수행하였다.

[0146] 균주 S. 베네주엘레 YJ028/pLRHM2의 전체 배양액(2.4 L)을 원심분리하고, 상층액을 에틸아세테이트로 용매-용매 분획하였다. 수득한 추출물을 증발시키고, 결과로 얻은 갈색 잔여물을 2 mL/min 유속의 구매용리(60분까지 40% 메탄올 수용액, 61 내지 72분 동안 메탄올 100%까지, 72 내지 82분 동안 40% 메탄올 수용액으로)를 이용한 역상 준분취용 HPLC로 분획하였다. AZDM 글리코시드 **11**을 포함하는 분획을 이동상으로 2 mL/min 유속의 18% 아세트오니트릴 수용액을 이용한 역상 준분취용 HPLC로 정제하여 무정형 흰색 고체로서 순수한 화합물(1.4 mg)을 획득하였다. 균주 S. 베네주엘레 YJ028/pDQNV의 전체 배양액(1.8 L)을 2 mL/min 유속의 구매용리(37분까지 44% 메탄올 수용액, 38 내지 48분 동안 메탄올 100%까지, 49 내지 59분 동안 44% 메탄올 수용액으로)를 이용한 역상 준분취용 HPLC로 분획하였다. AZDM 글리코시드 **12**를 포함하는 분획을 이동상으로 2 mL/min 유속의 20% 아세트오니트릴 수용액을 이용한 역상 준분취용 HPLC로 추가로 정제하여 순수한 화합물(0.3 mg)을 획득하였다. 균주 S. 베네주엘레 YJ028/pDDSS의 전체 배양액(2.4 L)을 위와 유사하게 처리하여 갈색 추출물을 획득하였다. 상기 갈색 잔여물을 2 mL/min 유속의 구매용리(25분까지 50% 메탄올 수용액, 26 내지 45분 동안 메탄올 수용액 60%까지, 46 내지 55분 동안 50% 메탄올 수용액으로)를 이용한 역상 준분취용 HPLC로 분획하였다. AZDM 글리코시드 **16**을 포함하는 분획을 이동상으로 2 mL/min 유속의 28% 아세트오니트릴 수용액을 이용한 역상 준분취용 HPLC로 추가로 정제하여 무정형 흰색 고체로서 순수한 화합물(1.8 mg)을 획득하였다. 유사하게 AZDM 글리코시드 **13**을 표제 화합물을 포함하는 분획을 이동상으로 2 mL/min 유속의 28% 아세트오니트릴 수용액을 이용한 역상 준분취용 HPLC로 정제하여 순수한 화합물(1.2 mg)을 획득하였다. 균주 YJ028/pLOLV2의 전체 배양액(3.0 L)으로부터 수득한 잔여물을 50% 메탄올 수용액으로부터 시작하여 100% 메탄올에 이르는 구매용리를 이용한 역상 열린 컬럼 크로마토그래피로 철저히 분리하였다. 마크로락탐 글리코시드의 존재를 확인하기 위하여 수득한 분획을 HPLC-ESI-MS/MS로 분석하고, AZDM 글리코시드 **17**을 포함하는 분획을 이동상으로 2 mL/min 유속의 23% 아세트오니트릴 수용액을 이용한

역상 준분취용 HPLC로 추가로 분리하여 순수한 화합물(0.8 mg)을 수득하였다.

- [0147] L-람노실-AZDM (11): 무정형 흰색 분말; ^1H - 및 ^{13}C -NMR, 하기 표 2 참조; (+)-ESI-MS m/z 459 $[\text{M}+\text{NH}_4]^+$, 442 $[\text{M}+\text{H}]^+$, 406 $[\text{M}+\text{H}-2\text{H}_2\text{O}]^+$, 296 $[\text{M}+\text{H}-\text{sugar}]^+$, 278 $[\text{M}+\text{H}-\text{sugar}-\text{H}_2\text{O}]^+$, 260 $[\text{M}+\text{H}-\text{sugar}-2\text{H}_2\text{O}]^+$; HR-ESI-MS m/z 464.2614 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{23}\text{H}_{39}\text{NO}_7\text{Na}$, 464.2624).
- [0148] D-퀴노보실-AZDM (12): 무정형 흰색 분말; ^1H - 및 ^{13}C -NMR, 하기 표 2 참조; (+)-ESI-MS m/z 459 $[\text{M}+\text{NH}_4]^+$, 442 $[\text{M}+\text{H}]^+$, 296 $[\text{M}+\text{H}-\text{sugar}]^+$, 278 $[\text{M}+\text{H}-\text{sugar}-\text{H}_2\text{O}]^+$, 260 $[\text{M}+\text{H}-\text{sugar}-2\text{H}_2\text{O}]^+$; HR-ESI-MS m/z 464.2627 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{23}\text{H}_{39}\text{NO}_7\text{Na}$, 464.2624).
- [0149] 3-O-데메틸-D-칼코실-AZDM (13): 무정형 흰색 분말; ^1H - 및 ^{13}C -NMR, 하기 표 2 참조; (+)-ESI-MS m/z 448 $[\text{M}+\text{Na}]^+$, 443.0 $[\text{M}+\text{NH}_4]^+$, 426 $[\text{M}+\text{H}]^+$, 408 $[\text{M}+\text{H}-\text{H}_2\text{O}]^+$, 296 $[\text{M}+\text{H}-\text{sugar}]^+$, 278 $[\text{M}+\text{H}-\text{sugar}-\text{H}_2\text{O}]^+$, 260 $[\text{M}+\text{H}-\text{sugar}-2\text{H}_2\text{O}]^+$; (+)-HR-ESI-MS m/z 448.2673 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{23}\text{H}_{39}\text{NO}_6\text{Na}$, 448.2675).
- [0150] N-데디메틸-N-아세틸-D-데조자미닐-AZDM (14): (+)-ESI-MS m/z 489 $[\text{M}+\text{Na}]^+$, 467 $[\text{M}+\text{H}]^+$, 296 $[\text{M}+\text{H}-\text{sugar}]^+$, 278 $[\text{M}+\text{H}-\text{sugar}-\text{H}_2\text{O}]^+$, 172 $[\text{sugar}]^+$, 113 $[\text{sugar}-28]^+$.
- [0151] D-데조자미닐-AZDM (15): (+)-ESI-MS m/z 475 $[\text{M}+\text{Na}]^+$, 453 $[\text{M}+\text{H}]^+$, 296 $[\text{M}+\text{H}-\text{sugar}]^+$, 278 $[\text{M}+\text{H}-\text{sugar}-\text{H}_2\text{O}]^+$, 158 $[\text{sugar}]^+$.
- [0152] 3-케토-4,6-디테옥시-D-글루코실-AZDM (16): 무정형 흰색 분말; ^1H - 및 ^{13}C -NMR, 하기 표 2 참조; (+)-ESI-MS m/z 446 $[\text{M}+\text{Na}]^+$, 441.0 $[\text{M}+\text{NH}_4]^+$, 424 $[\text{M}+\text{H}]^+$, 296 $[\text{M}+\text{H}-\text{sugar}]^+$, 278 $[\text{M}+\text{H}-\text{sugar}-\text{H}_2\text{O}]^+$, 260 $[\text{M}+\text{H}-\text{sugar}-2\text{H}_2\text{O}]^+$; (+)-HR-ESI-MS m/z 446.2510 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{23}\text{H}_{37}\text{NO}_6\text{Na}$, 446.2519).
- [0153] D-올리보실-AZDM (17): 무정형 흰색 분말; ^1H - 및 ^{13}C -NMR, 하기 표 2 참조; (+)-ESI-MS m/z 448 $[\text{M}+\text{Na}]^+$, 443.0 $[\text{M}+\text{NH}_4]^+$, 426 $[\text{M}+\text{H}]^+$, 296 $[\text{M}+\text{H}-\text{sugar}]^+$, 278 $[\text{M}+\text{H}-\text{sugar}-\text{H}_2\text{O}]^+$, 260 $[\text{M}+\text{H}-\text{sugar}-2\text{H}_2\text{O}]^+$; HR-ESI-MS m/z 448.2672 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{23}\text{H}_{39}\text{NO}_6\text{Na}$, 448.2675).
- [0154] L-올리보실-AZDM (18): 무정형 흰색 분말; (+)-ESI-MS m/z 448 $[\text{M}+\text{Na}]^+$, 426 $[\text{M}+\text{H}]^+$, 296 $[\text{M}+\text{H}-\text{sugar}]^+$, 278 $[\text{M}+\text{H}-\text{sugar}-\text{H}_2\text{O}]^+$, 260 $[\text{M}+\text{H}-\text{sugar}-2\text{H}_2\text{O}]^+$.

표 2

No	11		12 ^a		13		16		17	
	δ_c	δ_H (m, J in Hz)	δ_c	δ_H (m, J in Hz)	δ_c	δ_H (m, J in Hz)	δ_c	δ_H (m, J in Hz)	δ_c	δ_H (m, J in Hz)
1	177.4		177.8		177.8		177.7		177.6	
2	46.1	2.69 (dq, 10, 7)	46.1	2.76 (dq, 9.9, 6.3)	46.2	2.76 (dq, 9.5, 6.5)	46.1	2.76 (m, 9.5, 6.5)	45.9	2.68 (dq, 10, 7)
3	91.2	3.48 (d, 9.5)	87.7	3.54 (d, 9.9)	87.5	3.54 (m, 8.5) ^b	88.4	3.59 (d, 10)	88.5	3.53 (d, 10)
4	34.3	1.33 (m)	34.2	1.32 (over-lapped)	34.3	1.29 (m) ^b	34.3	1.30 (m)	34.2	1.25 (over-lapped)
5	35.3	1.71 (t, 13) 1.41 (td, 13, 3)	35.5	1.71 (br t, 12.6) 1.41 (td, 12.6, 3.6)	35.5	1.72 (t, 12.5) 1.42 (td, 12.5, 3)	35.4	1.76 (t, 14) 1.45 (m)	35.2	1.66 (t, 13.5) 1.40 (m, 13.5, 4)
6	46.4	2.48 (m)	46.4	2.47 (m)	46.4	2.47 (m)	46.5	2.48 (m)	46.4	2.48 (m)
7	207.6		207.9		208.0		207.9		207.4	
8	126.5	6.44 (d, 16)	126.6	6.44 (d, 15.3)	126.6	6.44 (d, 15.5)	126.6	6.44 (d, 15.5)	126.6	6.44 (d, 16)
9	149.7	6.72 (dd, 15.5, 5.5)	149.5	6.72 (dd, 16.2, 6.3)	149.5	6.72 (dd, 15.5, 5.5)	149.6	6.72 (dd, 16, 5.5)	149.5	6.72 (dd, 16, 6)
10	39.6	2.73 (m)	39.7	2.73 (m)	39.7	2.73 (m)	39.7	2.73 (m)	39.6	2.71 (m)
11	51.9	3.97 (m)	51.9	3.97 (ddd, 9.0, 5.3, 3.7)	51.9	3.97 (m)	51.9	3.97 (m) ^b	51.9	3.96 (m)
12	27.2	1.52 (m)	27.2	1.52 (m)	27.2	1.52 (m)	27.2	1.52 (m)	27.2	1.52 (m)
13	11.5	0.93 (t, 7.5)	11.5	0.93 (t, 7.2)	11.5	0.93 (t, 7)	11.5	0.93 (t, 7)	11.5	0.93 (t, 7)
14	17.7 ^b	1.24 (over-lapped) ^b	17.5	1.34 (d, 7.2)	17.6 ^b	1.34 (d, 7)	17.4	1.33 (d)	17.6 ^b	1.22 (d, 6.5)
15	17.9	1.00 (d, 6.5)	17.9 ^b	0.97 (d, 7.2)	17.9	0.97 (d, 7)	18.0	1.04 (d, 7)	17.7 ^b	0.98 (d, 6)
16	17.7 ^b	1.23 (over-lapped) ^b	17.7	1.20 (d, 7.2)	17.7 ^b	1.20 (d, 6.5)	17.7	1.23 (d, 6.5)	17.7 ^b	1.21 (d, 7)
17	10.0	1.07 (d, 6.5)	10.0	1.07 (d, 6.3)	10.0	1.07 (d, 6.5)	10.0	1.07 (d, 7.5)	10.0	1.07 (d, 7)
1'	105.1	4.67 (br s)	105.1	4.27 (d, 8.1)	105.4	4.20 (d, 7.5)	106.4	4.39 (d, 8)	102.7	4.47 (dd, 9.5, 1.5)
2'	72.5 ^b	3.92 (br s)	76.0	3.16 (dd, 9, 9)	77.5	3.05 (t, 8.5)	79.6	3.96 (m) ^b	40.7	2.17 (ddd, 11.9, 5.1, 1.4) 1.43 (dd, 12.7, 2.9)
3'	72.4 ^b	3.59 (dd, 9.5, 3)	78.1	3.28 (t, 9)	72.5	3.54 (m, 8.5) ^b	206.9		72.3	3.48 (ddd, 13.5, 8.5, 4.5)
4'	73.7	3.40 (t, 9.5)	77.2	2.97 (t, 9)	42.2	1.90 (dd, 12.5, 5) 1.27 (m)	48.3	2.44 (d, 6)	78.5	2.87 (t, 9)
5'	70.4	3.71 (dq, 9.5, 7)	72.9	3.23 (dq, 9, 6)	68.8	3.54 (m, 8.5) ^b	68.9	3.70 (m, 12.5, 6.5)	73.1	3.17 (m, 9.2, 6.3)
6'	17.6 ^b	1.24 (over-lapped) ^b	18.0 ^b	1.24 (d, 6.3)	21.2	1.19 (d, 6.5)	21.5	1.32 (d)	18.2	1.25 (d, 6)

^a ¹H- and ¹³C-NMR spectra were measured at 900 MHz and 225 MHz, respectively.

^b Similar values from the same column may be interchanged.

[0155]

[0156] 실시예 3: AZDM 글리코시드(11-13, 16 및 17)의 구조적 특성분석

[0157] 각각의 배양액으로부터 글리코시드 11 내지 18을 수득하였다. 이들 배양액을 가공하고 반복적인 크로마토그래피 분리를 적용하여 순수한 화합물 11(YJ028/pLRHM2 배양액 2.4 L로부터 1.4 mg), 12(YJ028/pDQNV 배양액 1.8 L로부터 0.3 mg), 13(YJ028/pDDSS 배양액 2.4 L로부터 1.2 mg), 16(YJ028/pDDSS 배양액 2.4 L로부터 1.8 mg), 및 17(YJ028/pL0LV2 배양액 3.0 L로부터 0.8 mg)을 수득하였다. 마크로락탐 글리코시드 14, 15, 및 18은 NMR 데이터를 얻기에는 부족한 미량 검출되었다. 1D(¹H- 및 ¹³C-) 및 2D(COSY, HSQC, HMBC, 및 NOESY) NMR, MS/MS, 및 HR-ESI-MS 데이터를 이용하여, 분리된 글리코시드 11-13, 16, 및 17의 구조를 결정하였다. 결합상수값(coupling constant value)(표 2)과 NOE 상관관계(도 5)를 기초로 상기 분리된 마크로락탐 글리코시드에서 부착된 당류의 입체화학 및 형태(conformation)를 확인하였다.

[0158] AZDM 글리코시드 11의 HR-ESI-MS 데이터에 있어서 m/z 464.2614(calcd. 464.2624)에서 관찰된 나트륨 부가 이온으로부터, C₂₃H₃₉NO₇의 분자식이 제안되었다. ¹³C NMR 및 HSQC 스펙트럼으로 마크로락탐 글리코시드 내의 23개 탄소의 존재를 확인하였다: 하나의 케톤(δ_c 207.6/C-7), 하나의 에스테르(δ_c 177.4/C-1), 두 개의 올레핀 탄소(δ_c 149.7/C-9 및 δ_c 126.5/C-8), 하나의 아노머 탄소(anomeric carbon)(δ_c 105.1/C-1'), 5개의 산소화된 메틴 탄소(oxygenated methine carbons)(δ_c 91.2/C-3, 73.7/C-4', 72.5/C-2', 72.4/C-3', 및 70.4/C-5'), 4개의 메틴 탄소(δ_c 51.9/C-11, 46.4/C-6, 46.1/C-2, 및 39.6/C-10), 3개의 메틸렌 탄소(δ_c 35.3/C-5, 34.3/C-4, 및 27.2/C-12), 및 6개의 메틸 탄소(δ_c 17.9/C-15, 17.6/C-6', 17.7/C-14, 17.7/C-16, 11.5/C-13, 및 10.0/C-17). ¹H NMR 스펙트럼은 δ_H 6.72 (H-9) 및 6.44 (H-8)에서 두 개의 올레핀 양성자, δ_H 4.67 (H-1')에서 하나의 아노머 양성자, δ_H 3.97 (H-11)에서 하나의 메틴 양성자, δ_H 3.92 (H-2'), 3.71 (H-5'), 3.59 (H-

3'), 3.48 (H-3), 및 3.40 (H-4')에서 5개의 산소화된 메틴 양성자, 및 δ_H 1.24 (H-6'), 1.24 (H-14), 1.23 (H-16), 1.07 (H-17), 1.00 (H-15), 및 0.93 (H-13)에서 6개 메틸기 신호와 같은 이들 마크로라이드 글리코시드의 특성 신호를 나타내었다. COSY 스펙트럼은 테옥시당에 대한 하나를 포함하는 3개 스핀 시스템의 존재를 나타내었다(도 5). 이들 스핀 시스템 모두는 HMBC 상관관계의 도움으로 연결되었다(도 5). AZDM의 C-3 위치에 대한 당 모이어티의 부착을 HMBC 스펙트럼을 이용하여 확인하였다. HMBC 스펙트럼에서 산소화된 메틴 양성자(δ_H 3.48/H-3)와 아노머 탄소(δ_C 105.1/C-1) 간의 상관관계 뿐만 아니라 아노머 양성자(δ_H 4.67/H-1')와 탄소(δ_C 91.2/C-3) 간의 상관관계가 관찰되었으며, 이로부터 결합점을 확인할 수 있었다. 유사하게, 모든 메틸기들의 위치를 각각의 이웃한 탄소중심에 대한 이들의 HMBC 상관관계를 기초로 결정하였다(도 5). AZDM 글리코시드 **11**과 이의 상대 락톤(lactone counterpart) 간의 주된 차이는 C-11 (δ_C 51.9/11, 74.6/L-람노실-10-dml) 및 H-11 (δ_H 3.97/11, 4.96/L-람노실-10-dml) 피크에 대한 높은 장 이동(shift upfield)이었다. 이들의 적도배향(equatorial orientation)을 나타내는 H-1' (br s), H-2' (br s)에 대한 작은 결합상수값 및 축배향(axial orientation)을 나타내는 H-3' (d, $J = 9.5$ Hz), H-4' (t, $J = 9.5$ Hz), H-5' (dq, $J = 9.5, 7$ Hz)의 큰 결합상수값을 기초로 **11**에서 당은 L-람노스로 확인되었다.

[0159] HR-ESI-MS 및 ^{13}C NMR 데이터로부터 마크로락탐 글리코시드 **12**에 대한 분자식은 $C_{23}H_{39}NO_7$ 로 추론되었다. HR-ESI-MS 스펙트럼은 m/z 464.2627 $[M+Na]^+$ (calcd 464.2624)에서 피크를 나타내었으며 ^{13}C NMR 스펙트럼은 AZDM 글리코시드에 대해 요구되는 23개 탄소의 존재를 나타내었다. **12**에 대한 NMR 데이터(표 2)는 당 및 C-3 위치에 대한 화학적 이동 값에서 기대되는 차이를 제외하고는 **11**에 대한 것과 매우 유사하였다. 아노머 양성자는 δ_H 4.27 (H-1')에서 이중피크(doublet)($J = 8.1$ Hz)으로 나타나는 한편 **11**에서 아노머 양성자는 δ_H 4.67 (H-1')에서 넓은 단일피크(singlet)이었다; 모든 다른 당 양성자들은 큰 결합상수값으로부터 유추되는 바와 같이 축배향을 나타내었다. H-1', H-3', 및 H-5' 간의 NOE 상관관계 또한 축배향을 뒷받침하였으며(도 5), 이는 **12**의 당이 D-퀴노보스임을 나타내었다. 잘 어사인된 COSY, HSQC, 및 HMBC 데이터는 **12**의 구조를 뒷받침하였다.

[0160] AZDM 글리코시드 **13**은 HR-ESI-MS 스펙트럼의 m/z 448.2673에서 관찰된 $[M+Na]^+$ 이온으로부터 어사인된 $C_{23}H_{39}NO_6$ 의 분자식을 가졌다. ^1H - 및 ^{13}C -NMR 데이터(표 2)는 이것이 23개 탄소 중심을 갖는 마크로락탐 **12**의 유사체임을 나타내었다. **13**에 대한 1D- 및 2D-NMR 데이터를 **12**의 데이터와 비교한 결과, 당 모이어티(아노머 탄소를 포함) 중 δ_C 105.4, 77.5, 72.5, 및 68.8에서 단지 4개의 산소화된 탄소 및 하나의 추가적인 메틸렌기(δ_C 42.2; 1.90, 1.27/H-4')가 존재하는 것으로 나타났고, 이는 **13**에 부착된 당 모이어티가 **11** 및 **12**에서의 6-테옥시헥소스를 대신하여 디테옥시헥소스임을 시사하였다. **13**의 COSY 및 HMBC 데이터의 추가적인 분석은 추가적인 메틸렌기가 C-4' 위치에 존재하며, 당 모이어티는 모든 당 양성자에 대한 NOE 상관관계 및 큰 결합상수값에 의해 3-O-테메틸-D-갈코스임을 나타내었다(도 5).

[0161] 마크로락탐 글리코시드 **16**의 분자식은 HR-ESI-MS 스펙트럼의 m/z 446.2510에서 관찰된 나트륨 부가 이온에 기초하여 $C_{23}H_{37}NO_6$ 로 추론하였다. **16**의 NMR 데이터(표 2)는 δ_C 206.9에서 신호를 발생하는 추가적인 케톤 모이어티의 존재를 제외하고는 **13**에 대한 것과 유사하였다. HMBC 스펙트럼은 양성자들 H-2' (δ_H 3.96)와 H-4' (δ_H 2.44) 간의 상관관계와 C-3'로 확인되는 δ_C 206.9에서 카보닐 피크를 나타내었다(도 5). δ_H 4.39 (d, $J = 8.0$ Hz, H-1') 및 δ_H 3.70 (m, H-5')에서의 양성자 신호들 간의 NOE 상관관계(도 5)에 의해 뒷받침되는 부착된 당 모이어티에서 모든 양성자들의 큰 결합상수는 부착된 당이 D-형태의 3-케토-4,6-디테옥시-글루코스임을 제안하였다.

[0162] AZDM 글리코시드 **17**의 HR-ESI-MS 스펙트럼은 m/z 448.2672 (calcd 448.2675)에서 나트륨 부가 이온 피크를 나타내었으며, 이는 이의 분자식이 $C_{23}H_{39}NO_6$ 임을 나타낸다. ^1H - 및 ^{13}C -NMR 데이터의 분석(표 2)은 **17** 역시 디테옥시헥소스 모이어티를 갖는 AZDM 글리코시드임을 제안하였다. COSY 스펙트럼에서, δ_H 4.47/H-1' (d, $J = 9.5, 1.5$ Hz)에서의 아노머 양성자 신호는 δ_H 3.48 (ddd, $J = 13.5, 8.5, 4.5$ Hz)에서 산소화된 메틴 양성자 신호와 추가적인 결합을 나타내는 δ_H 2.17 (ddd, $J = 11.9, 5.1, 1.4$ Hz) 및 δ_H 1.43 (dd, $J = 12.7, 2.9$ Hz)에서의 메틸렌 양성자와 결합되었다. 이는 당이 2,4-디테옥시헥소스임을 나타내었다(도 5). 모든 당 양성자의 큰 결합

상수는 이를 D-올리보스로서 동정하였으며, H-1' (δ_H 4.47), H-3' (δ_H 3.48), 및 H-5' (δ_H 3.17)(도 5) 간의 NOE 상관관계는 추가적으로 이를 뒷받침하였다.

[0163] 본 발명에서는 마크로락탐 글리코시드의 생산 및 DesVII/DesVIII-촉매된 당화반응의 분석을 통해, 몇가지 신규한 사항을 발견하였다. 먼저, 본 발명자들은 선행연구를 통해 DesVII/DesVIII이 비천원 마크로락톤 및 나아가 비고리형 아글리콘에 다양한 당류의 전달을 촉진시킴을 확인하였음에도 불구하고, 이러한 결과들은 마크로라이드 글리코실전달효소 DesVII/DesVIII가 일련의 신규한 당화된 마크로락탐을 제공하는 마크로락탐 아글리콘을 인식하고 이후 과정을 진행할 수 있음을 최초로 나타내는 것이다. 둘째로, L-디지톡소실-나르보노라이드, D-보이비노실-나르보노라이드, 및 D-보이비노실-10-dml를 플라스미드 pL0LV2를 포함하는 변이균주로부터 이전에 분리하였음에도 불구하고, 본 발명에서 L-디지톡소스 및 D-보이비노스를 포함하는 AZDM 글리코시드는 검출되지 않았다. 셋째, 현재까지 알려진 바, 본 발명은 케토-당 **26**을 포함하는 글리코시드의 분리에 대한 최초의 보고이며, 이로부터 TDP-D-데조자민 (**30**)의 생합성 중간체로서 뉴클레오티드 당 **26**의 존재를 확인하였다. 또한, 선행연구에서, *desV* 결손 S. 베네주엘라 변이주가 당 **26**을 갖는 예상되는 메티마이신 유사체 대신에 3-O-데메틸 D-칼코실 메티마이신을 생산함을 확인하였다. 이러한 마지막 두 가지 결과는 다른 당 공여체에 대한 DesVII/DesVIII의 글리코실화 효율이 아글리콘의 구조에 따라 변할 수 있음을 나타내는 것이다(표 3). 그러나, 각각의 당 공여체의 세포 내 농도 및 다른 당 공여체 및 수용체에 대한 정확한 글리코실화 효율은 생체 내 시스템에서 측정할 수 없다. 그러나, 각각의 당 유전자 카세트 발현 수준을 동일화 하기 위하여 동일한 발현 플라스미드와 클로닝 전략을 사용하였으므로, 각각의 유전자 카세트 발현 수준은 유사할 것으로 기대된다. 따라서, 마크로락탐 글리코시드에 대한 생산 속도의 차이는 변이균주 성장거동에서의 차이 또는 당 유전자 카세트 발현 수준에서의 차이 보다는 각각의 기질에 따라 변화하는 글리코실전달효소의 촉매 효율에 주로 기인한다.

표 3

sugar	macrolactam glycosides (mg/L)	YC-17 analogues ^a (mg/L)
L-rhamnosyl	0.58	0.50
D-quinovosyl	0.16	3.40
D-desosaminyl	trace	1.10
3-O-demethyl-D-chalcosyl	0.50	not detected
D-olivovosyl	0.26	0.10
L-olivovosyl	trace	0.30
3-keto-4,6-dideoxy-D-glucosyl	0.75	not detected
N-dedimethyl-N-acetyl-D-desosaminyl	trace	not detected
L-digitoxosyl	not detected	not detected
D-boivinovosyl	not detected	0.02

^aPreviously reported.^{9a}

[0164]

실시예 4: AZDM 글리코시드의 항균활성분석

[0165]

[0166] 에리트로마이신-저항성 균주 엔테로코쿠스 패시움(*Enterococcus faecium*) P00558 및 스탕필로코쿠스 아우레우스(*Staphylococcus aureus*) P00740을 공지의 방법으로 분리하였다. 에리트로마이신-민감성 균주 E. 패시움 ATCC 19434 및 S. 아우레우스 ATCC 25923은 American Type Culture Collection(ATCC, Manassas, VA, USA)로부터 입수하였다. 공지의 방법으로 3회 반복하여 항균 어세이를 수행하였다.

[0167] 분리된 화합물들의 에리트로마이신 저항성 임상적 분리체 엔테로코쿠스 패시움 P00558 및 스탕필로코쿠스 아우레우스 P00740에 대해서 뿐만 아니라 에리트로마이신-민감성 E. 패시움 ATCC 19434 및 S. 아우레우스 ATCC 25923에 대한 시험관 내 항균활성을 평가하였다(표 4). 시험한 균주들에 대한 대부분의 AZDM 글리코시드의 활성은 그들의 상대 마크로락톤의 것과 유사하였음에도 불구하고(표 5), AZDM 글리코시드 **11** 및 **12**는 에리트로마이신-민감성 병원균에 대해 에리트로마이신과 비교하여 4배 더 높은 활성을 나타내었으며, 에리트로마이신-저항성 병원균주에 대해서도 활성을 나타내었다. 각각 6-데옥시헥소스 L-람노스 및 D-퀴노보스를 포함하는 AZDM 글리코시드 **11**(MIC, 7.5-30 μ M) 및 **12**(7.5-30 μ M)는 디데옥시헥소스 모이어티를 포함하는 글리코시드(**13**, **16**, 및 **17**)에 비해 시험한 모든 균주에 대해 더 나은 활성을 나타내었다(30-120 μ M). 이는 **11** 및 **12**의 추가적인 히드록실기가 그들의 항균활성에 바람직한 영향을 끼칠 수 있다는 이전의 관찰을 뒷받침하는 것이다.

표 4

No.	MIC (μ M) ^a			
	<i>E. faecium</i> ATCC19434	<i>E. faecium</i> P00558	<i>S. aureus</i> ATCC25923	<i>S. aureus</i> P00740
11	7.5	30.0	15.0	30.0
12	7.5	30.0	15.0	15.0
13	30.0	120.0	60.0	120.0
16	30.0	120.0	60.0	120.0
17	30.0	>120.0	60.0	60.0
Erythromycin	30.0	>120.0	60.0	120.0

^a*E. faecium* ATCC 19434 and *S. aureus* ATCC 25923 are erythromycin-susceptible pathogens and *E. faecium* P00558 and *S. aureus* P00740 are clinically isolated erythromycin-resistant pathogens.

[0168]

표 5

Sugar	MIC (μ M) ^a							
	macrolactam glycosides				macrolactone glycosides ^b			
	<i>E. faecium</i> ATCC19434	<i>E. faecium</i> P00558	<i>S. aureus</i> ATCC25923	<i>S. aureus</i> P00740	<i>E. faecium</i> ATCC19434	<i>E. faecium</i> P00558	<i>S. aureus</i> ATCC25923	<i>S. aureus</i> P00740
L-rhamnosyl	7.5	30.0	15.0	30.0	11.3	5.6	5.6	5.6
D-quinovosyl	7.5	30.0	15.0	15.0	11.3	11.3	22.6	22.6
D-olivovosyl	30.0	>120.0	60.0	60.0	23.2	46.5	46.5	23.2
Erythromycin	—	—	—	—	30.0	>120.0	60.0	120.0

^a*E. faecium* ATCC 19434 and *S. aureus* ATCC 25923 are erythromycin-susceptible pathogens and *E. faecium* P00558 and *S. aureus* P00740 are clinically isolated erythromycin-resistant pathogens.

^bPreviously reported.²⁴

[0169]

[0170]

실시예 5: 에리트로마이신 에스테라아제와의 안정성 연구

[0171]

Bioneer(Republic of Korea)에 의해 *E. coli*로부터 *ereB*를 인코딩하는 DNA를 합성하였다. 상기 DNA 단편(fragment)을 NdeI와 BamHI로 절단(digest)하여 N-말단 His6-태크를 포함하는 pET15b(Novagen, Germany) 내로 복제하였다. T7 프로모터 하에 과발현을 위하여 *E. coli* BL21(DE3)pLysS(Novagen)에 발현 플라스미드를 도입하였다. 50 μ g/mL 암피실린 및 25 μ g/mL 클로람페니콜(chloramphenicol)을 첨가한 LB 배지에서 세포를 배양하였다. 배양액 1 리터 당 10 mL 밤샘 종균 배양(overnight starter culture)을 접종하였다. 배양물을 37°C에서 광학밀도(OD600)가 0.6에 달할 때까지 성장시킨 후, 0.1 mM 이소프로필- β -D-티오갈락토피라노사이드(isopropyl- β -D-thiogalactopyranoside; IPTG)로 3시간 동안 발현을 유도하였다. 원심분리(6,000 $\times$ g, 10분)하여 세포를 회수하여, 용해완충액(300 mM NaCl 및 10 mM 이미다졸을 포함하는 50 mM 인산나트륨 완충액, pH 8.0)에 재현탁한 후, 초음파처리하여 용해시켰다. 정제된 세포 용해물을 Ni-NTA 초유동 컬럼(Qiagen)에 통과시켰다. 완충액(300 mM NaCl 및 20 mM 이미다졸을 포함하는 50 mM 인산나트륨 완충액, pH 8.0)으로 컬럼을 세척한 후, N-His6-태그된 제조합 단백질을 용출 완충액(300 mM NaCl 및 500 mM 이미다졸을 포함하는 50 mM 인산나트륨 완충액, pH 8.0)으로 용출하였다. 정제된 단백질을 13% SDS-PAGE에 적용하여 쿠마시 블루 염색으로 가시화하였다. 결과로 얻은 단백질을 150 mM NaCl 및 2% (v/v) 글리세롤을 포함하는 50 mM HEPES (pH 7.5)에 대해 투석하고, 시험관 내 반응에 사용하기 전까지 -80°C에 보관하였다. 8 μ g 에리트로마이신, L-람노실-10-dm1, 및 AZDM 글리코시드 11을 별도의 2 mL 유리 바이알에 옮겼다. 각 바이알에 100 μ L 인산완충액을 첨가하고, 2 μ M *EreB*를 첨가하여 반응을 개시하였다. 이들 반응 혼합물을 37°C의 진탕수조(150 rpm)에서 0, 1, 2, 및 24시간 동안 인큐베이션하였다. 150 μ L 차가운 아세트니트릴을 첨가하여 반응을 중지하였다. 반응 혼합물을 4°C에서 원심분리(3,000 $\times$ g, 10분)하고, 상층액을 분리하여 HPLC-ESI-MS/MS 분석을 수행하였다. 본 실험은 3회 반복하여 수행하였다.

[0172]

E. 패슈 P00558 및 S. 아우레우스 P00740의 저항성 기전은 알려지지 않았으므로, 마크로라이드-저항성 병원균의 치료에 대한 마크로락탐 약물분자구조(pharmacophore)의 잠재적 장점을 연구하기 위하여 에리트로마이신, L-람노실-10-dm1, 및 AZDM 글리코시드 11의 에리트로마이신-저항성 에스케리키아 콜라이로부터 분리된 잘 알려진 마크로라이드 에스테라아제 *EreB*에 대한 민감성을 시험하였다. *ereB* 유전자(GenBank Accession No. P05789)를 합성하고 과발현시켜 *E. coli* BL21(DE3)pLysS로부터 정제하였다. 에리트로마이신 및 *EreB*를 포함하는 반응 혼합

물의 HPLC-ESI-MS 분석(도 6A)으로부터, 에리트로마이신(머무름시간 29.5분)은 초기에 부분적으로 이의 선형적 형태(머무름시간 22분)로 전환되는 것이 관찰되었다. 또한, 1시간 인큐베이션 시, 에리트로마이신은 불활성 산물로 완전히 전환되었다(도 6D 및 6E). 기질 및 생성물 모두의 MS 스펙트럼은 이전에 보고된 바와 유사하게 나타났다. L-람노실-10-dm1 및 AZDM 글리코시드 11을 EreB 와 별도로 인큐베이션하고 이어서 HPLC-ESI-MS 분석을 수행하였을 때(도 6B 및 6C), 두 개 화합물은 24시간 인큐베이션 후에도 EreB에 의해 가수분해되지 않음을 확인하였다. 이는 EreB가 14-원 마크로라이드에 특이적인 에스테라아제일 수 있으며 12-원 마크로라이드(L-람노실-10-dm1 및 AZDM 글리코시드 11)를 기질로서 인식하지 못할 수 있음을 나타낸다. 그럼에도 불구하고, 이와 같은 결과는 AZDM 글리코시드 11과 같은 마크로락탐 글리코시드가 에리트로마이신을 분해할 수 있는 미생물성 에스테라아제에 민감하지 않음을 나타내며, E. 패숨 P00558 및 S. 아우레우스 P00740의 저항성 기전이 알려지지 않았음에도 불구하고, 이러한 마크로락탐 글리코시드들은 그들의 항균 활성에 의해 나타난 바와 같이 여전히 에리트로마이신-저항성 병원균에 대해 활성임을 나타낸다(표 4).

[0173] 실시예 6: 인공위액(simulated gastric fluid; SGF)에서의 안정성 연구

[0174] SGF에서 안정성을 결정하기 위하여 AZDM 글리코시드 11 및 L-람노실-10-dm1을 선택하였다. 또한 에리트로마이신을 대조군으로서 안정성 어세이에 적용하였다. 미국약전의 기재에 따라 SGF(리터 당 NaCl 2 g, 펩신 3.2 g, 및 HCl 7 mL)를 준비하였다. 1 mL 메탄올에 각 화합물 0.25 mg을 용해시켜 3가지 모액을 준비하였다. 각 화합물의 200 μ L 분액을 2 mL 바이알에 옮기고, 메탄올을 증발시킨 후 1 mL SGF를 각 바이알에 첨가하였다. 이들 용액으로부터 분액(100 μ L)을 2 mL 용량의 6개 유리 바이알에 37°C의 진탕수조(120 rpm)에서 0, 30, 60, 및 120분 동안 인큐베이션하였다. 전체 실험을 3회 반복하여 수행하였고 분해 프로파일을 확인하기 위하여 HPLC-ESI-MS/MS를 이용하여 상기 조건에서 반응 혼합물을 분석하였다.

[0175] AZDM 글리코시드 11 및 이의 마크로락톤 유도체(L-람노실-10-dm1)는 에리트로마이신에 비해 강한 항균활성을 나타내므로, 마크로락탐 글리코시드, 마크로락톤 글리코시드, 및 에리트로마이신 사이의 위 안정성에서의 차이를 규명하기 위하여 인공위액 안정성 어세이를 실시하였다. 에리트로마이신을 SGF와 함께 인큐베이션했을 때, 반응 혼합물의 HPLC-ESI-MS 분석결과는 인큐베이션 후 즉시 에리트로마이신의 일부가 분해되고 2시간 인큐베이션 후에는 거의 완전히 약물이 사라졌으며, 클래디노스(cladinose) 당이 결합된 이의 분해 생성물 5-데조자미널-에리트로노리드 A를 생성하였다(도 7A). 에리트로마이신 피크는 29.5분에 용출되었고 m/z 734에서 $[M+H]^+$ 이온을 나타내었으며, 5-데조자미널-에리트로노리드 A에 기인하는 피크는 33분에서 용출되었고 m/z 562 및 540에서 각각 $[M+Na]^+$ 및 $[M+H]^+$ 이온을 나타내었다(도 7D 및 7E). 에리트로마이신과 5-데조자미널-에리트로노리드 A의 양성자 부가물들 간의 차이는 예상되는 바와 같이 클래디노스에 기인하는 질량 손실로 설명되었다. 그러나, 마크로락탐 11 및 L-람노실-10-dm1은 SGF에서 2시간 인큐베이션 후에도 손상되지 않았다(도 7B 및 7C). Cladinose는 산성 조건하에서 에리트로마이신으로부터 직접 절단되어 분해 생성물을 발생시키는 것으로 알려져 있다. 대조적으로, AZDM 글리코시드 11 및 L-람노실-10-dm1의 글리코시드 결합(glycosidic linkage)은 SGF에서 보다 안정한 것으로 나타났으며, AZDM 글리코시드 11 및 L-람노실-10-dm1이 위장의 산성 환경에서 살아남도록 하였다.

[0176] 실시예 7: 랫트 간 마이크로솜을 이용한 대사 안정성(metabolic stability) 연구

[0177] 유사하게 AZDM 글리코시드 11, L-람노실-10-dm1, 및 에리트로마이신을 DMSO에 용해시켜 2 μ g/ μ L의 모액을 준비하였다. 각 화합물의 20 μ L DMSO 용액을 2 mL 유리 바이알에 옮기고 655 μ L 인산완충액을 첨가한 후 500 μ g 랫트 간 마이크로솜 단백질(Sigma-Aldrich, USA)을 첨가하였다. 각 반응 혼합물의 100 μ L 분액을 2 mL 용량의 6개 유리 바이알에 옮기고 37°C의 진탕수조(150 rpm)에서 5분 동안 전인큐베이션하였다. 500 μ g NADPH를 각 바이알에 첨가하고 반응 혼합물을 추가로 0, 30, 60, 및 120분 동안 인큐베이션하였다. 150 μ L 차가운 아세트 니트릴을 첨가하여 반응을 종료하였다. 반응 혼합물을 4°C에서 10분 동안 원심분리(3,000 $\times$ g, 10분)하여 상층액을 HPLC-ESI-MS/MS로 분석하였다. 상기 실험은 3회 반복하여 실시하였다.

[0178] AZDM 글리코시드 11 및 이의 마크로락톤 유도체는 에리트로마이신에 비해 강한 항균활성을 나타내므로, 마크로락탐 글리코시드, 마크로락톤 글리코시드, 및 에리트로마이신 사이의 대사 안정성에서의 차이를 규명하기 위해

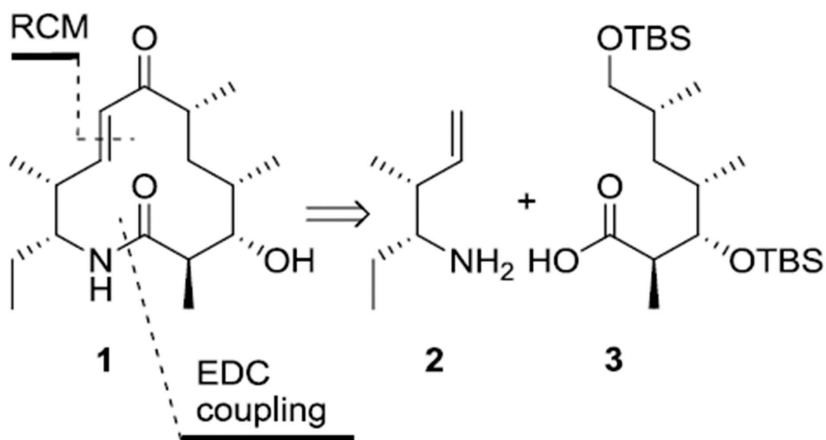
여 간 마이크로솜 대사 안정성 어세이를 실시하였다. 에리트로마이신은 담즙을 통해 즉시 제거되는 비활성 N-데스메틸-에리트로마이신으로 주로 간에서 빠른 대사를 경험하는 것과 같이 짧은 제거 반감기를 갖는다. 다른 마크로라이드 항생제들 역시 데스클래디노실-, N-데스메틸-, 및 히드록시-유도체로 다양한 간 가수분해 및 산화 효소에 의해 대사된다. 대사 패턴을 확인하기 위하여, 에리트로마이신, L-람노실-10-dm1, 및 AZDM 글리코시드 **11**을 랫트 간 마이크로솜 및 NADPH와 함께 인큐베이션하였다. 그 결과, 에리트로마이신 및 L-람노실-10-dm1은 2 시간 동안 인큐베이션 후 각각 N-데스메틸-에리트로마이신 및 히드록실화된 L-람노실-10-dm1로 대사되었다(도 8A, 8B, 8D 및 8E). 에리트로마이신을 포함하는 반응 혼합물의 HPLC-ESI-MS 분석 결과는 2개의 주요 피크를 나타내었다(도 8A): 29.5분에서 용출된 하나의 피크는 m/z 734에서 $[M+H]^+$ 이온에 의해 에리트로마이신으로서, 27 분에서 용출된 다른 하나의 피크는 m/z 742에서 $[M+Na]^+$ 이온 및 m/z 720에서 $[M+H]^+$ 이온을 나타내었다(도 8D). 상기 신규한 대사체는 에리트로마이신에 대해 관찰된 14 Da의 차이에 기초하여 N-데스메틸-에리트로마이신으로 확인되었다. [데조자민] $^+$ 이온 피크가 14 Da 차이인 m/z 158에서 나타났으므로, 상기 결과는 m/z 144에서 $[N-데스메틸-데조자민]^+$ 이온 피크의 존재에 추가로 뒷받침된다. 유사하게, L-람노실-10-dm1과 어세이의 HPLC-ESI-MS 분석 또한 랫트 간 마이크로솜과 인큐베이션 후 2개의 피크를 나타내었다(도 8B). 22.7분에서 용출된 피크는 각각 m/z 465, 460, 및 279에서 특징적인 $[M+Na]^+$, $[M+NH_4]^+$, 및 $[M+H-rhamnose-H_2O]^+$ 이온들을 나타내는 것으로, 이로부터 L-람노실-10-dm1 구조임을 확인하였다. 반면, 15분에서 용출된 피크는 m/z 481, 295, 및 277에서 $[M+Na]^+$, $[M+H-람노스-H_2O]^+$ 및 $[M+H-람노스-2H_2O]^+$ 이온을 나타내었다(도 8E). 상응하는 이온과 L-람노실-10-dm1의 것 사이의 18 Da의 차이는 15분에서의 대사체가 이의 히드록실화된 유사체임을 나타내는 것이다. L-람노실-10-dm1과 이의 히드록실화된 유사체로부터의 스펙트럼을 비교한 결과는 각각 m/z 279에서 m/z 295로, 및 m/z 261에서 m/z 277로의 단편 이온 $[M+H-람노스-H_2O]^+$ 및 $[M+H-람노스-2H_2O]^+$ 에서의 이동을 나타내었으며, 이는 람노스 당 상에서가 아닌 아글리콘 코어 상에서 히드록실화가 일어났음을 나타내는 것이다. 클래이트로마이신은 16-히드록시-클래이트로마이신으로 대사되는 것으로 알려져 있으므로, 수득한 L-람노실-10-dm1의 히드록실화된 대사체는 C-12 위치에서 기능화될 수 있다. 그러나, AZDM 글리코시드 **11**은 랫트 간 마이크로솜과의 2시간 동안의 인큐베이션에도 안정한 것으로 확인되었다(도 8C).

[0179] 이러한 결과는 화합물 **11**과 같은 AZDM 글리코시드가 일차 통과 대사(first-pass metabolism)를 이상적으로 생존하여 에리트로마이신 및 관련 마크로락톤 글리코시드 보다 더 오래 체내에 잔류할 수 있음을 나타내는 것이다. 따라서, AZDM 글리코시드는 아지트로마이신과 유사하게 지속적인 항균활성을 나타낼 수 있음을 나타내는 것이다. 종합적으로 본 발명의 AZDM 글리코시드는 에리트로마이신보다 위산에서 더 안정하고, 에리트로마이신 및 마크로락톤 글리코시드보다 더 대사적으로 안정하며, 에리트로마이신 에스테라아제에 의해 가수분해되지 않으므로, 이와 같은 안정성을 바탕으로 항균제로 유용하게 사용될 수 있다.

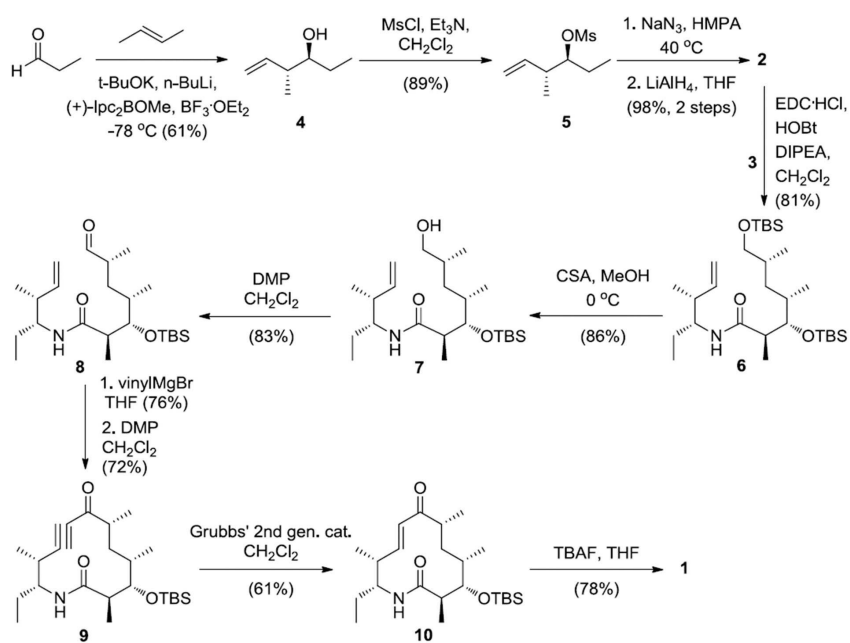
[0180] 종합적으로, 비천연 테옥시헥소오스 당의 세포 기반 생성과 아글리콘 생물전환을 결합시킨 합성 화학에 의한 YC-17에 기초한 마이크로락탐 아글리콘의 개발은 다양한 세트의 생물학적으로 활성인 마크로락탐 글리코시드의 생산을 가능하게 하였다. 본 발명은 마크로라이드 항생제의 다양하게 당화된 마크로락탐 유사체의 최초의 예시이며, 또한, 신규한 마크로락탐 글리코시드의 어레이를 제조하기 위한 마크로락탐을 수용하고 당화하는 DesVII/DesVIII 데조자미닐 글리코실전달효소의 최초의 예이다. 몇몇 신규한 AZDM 글리코시드들은 에리트로마이신-민감성 및 -저항성 병원균 모두에 대해 에리트로마이신 보다 향상된 항균활성을 나타냄을 확인하였다. 또한, 상기 본 발명의 신규한 항생제는 에리트로마이신을 분해하는 세균성 마크로라이드 에스테라아제에 의한 분해에 저항성을 나타내었다. 본 발명의 AZDM 글리코시드들은 또한 에리트로마이신에 비해 SGF에서 향상된 시험관 내 안정성을 나타내며, 랫트 간 마이크로솜 어세이에서 에리트로마이신 및 상응하는 마크로락톤 글리코시드 모두 보다 더 높은 안정성을 나타낸다. 이러한 결과는 YC-17의 다양한 신규한 당화된 마크로락탐 유사체를 제조하기 위하여 내재적으로 유연한 생합성 효소(inherently flexible biosynthetic enzymes)와 화학합성을 결합한 성공적인 적용을 나타내는 것이다. 일련의 어세이는 이들 신규한 분자 중 몇몇이 개선된 약리학적 성질을 가지며 약물 발굴 및 개발 프로그램을 위한 잠재적 항생제 활성을 갖는 신규한 마크로락탐 글리코시드를 생성하기 위한 구조적 다양화(diversification)를 위한 강력한 전략으로서 이와 같은 접근법을 지지함을 나타낸다.

도면

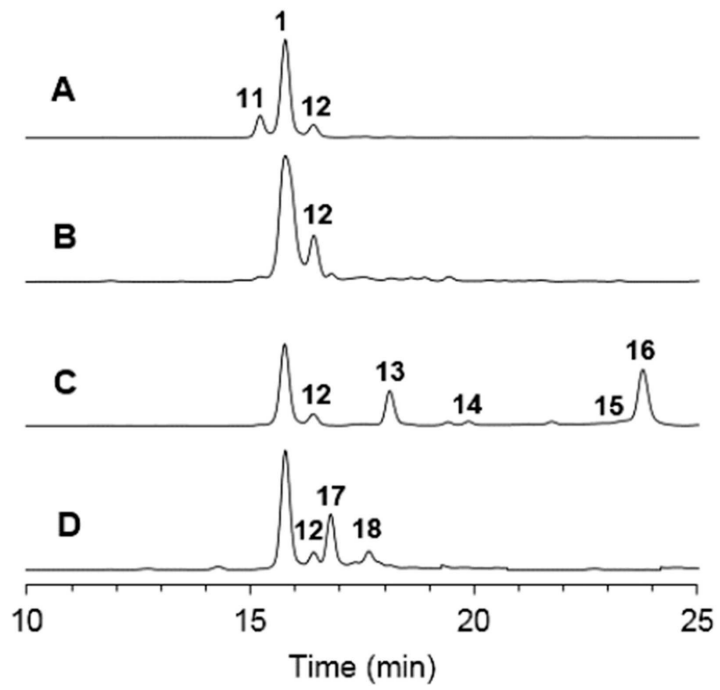
도면1



도면2

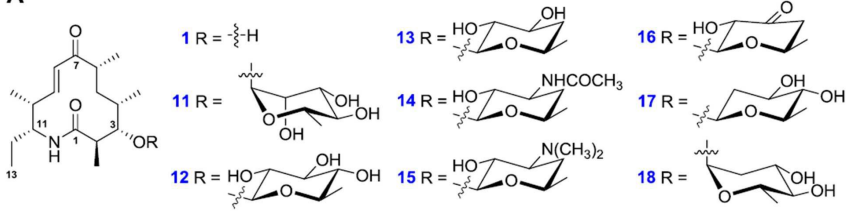


도면3

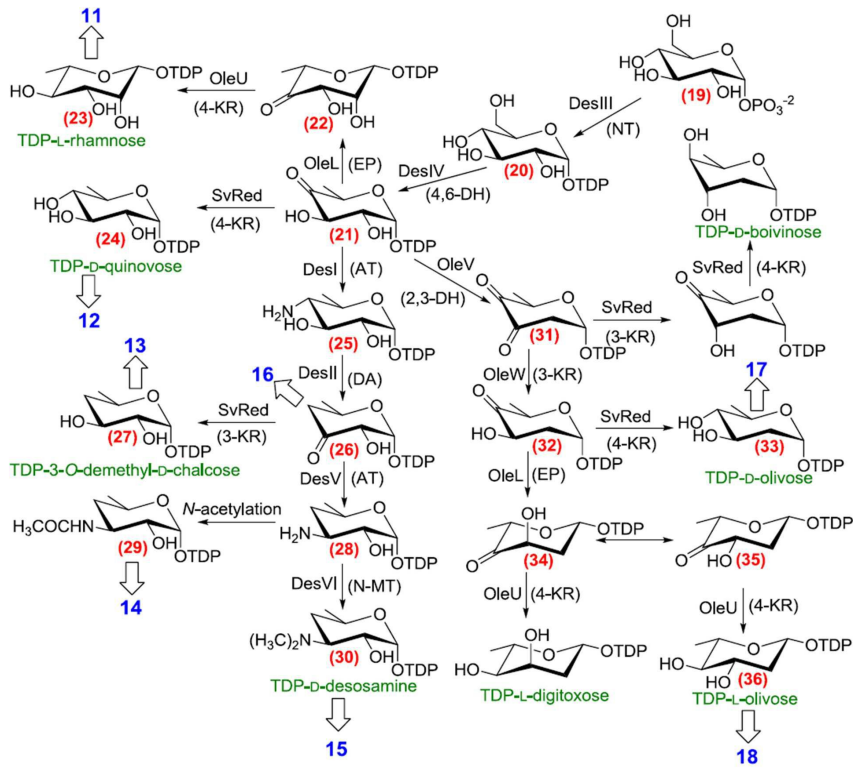


도면4

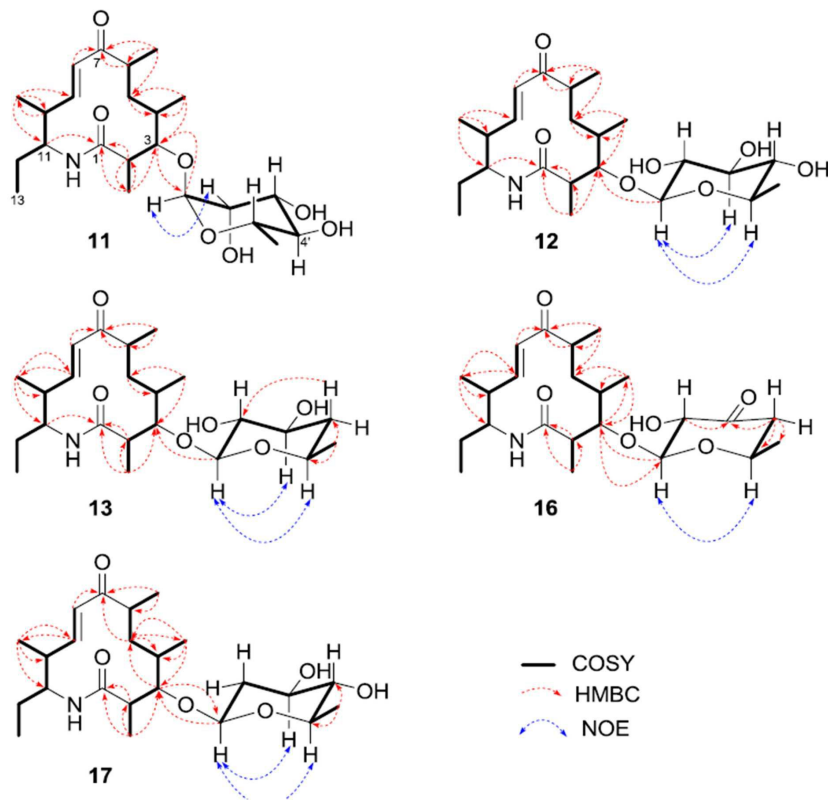
A



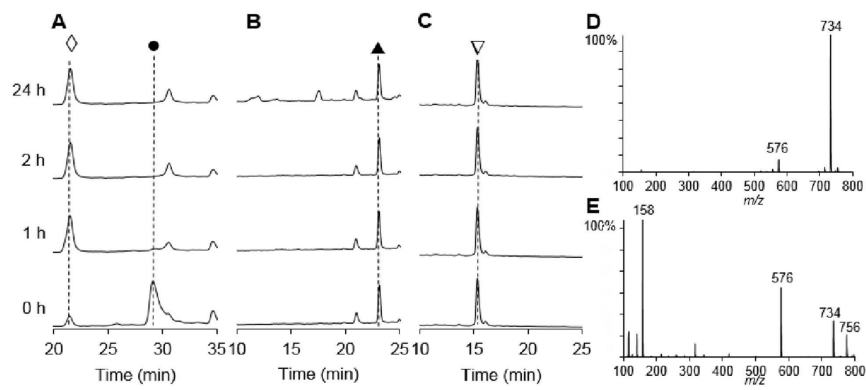
B



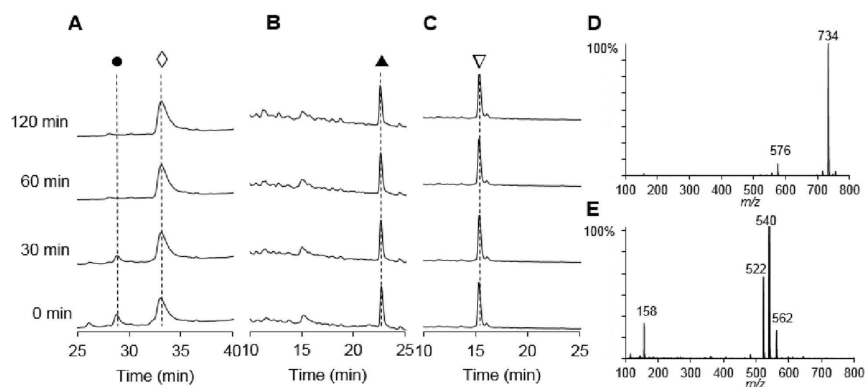
도면5



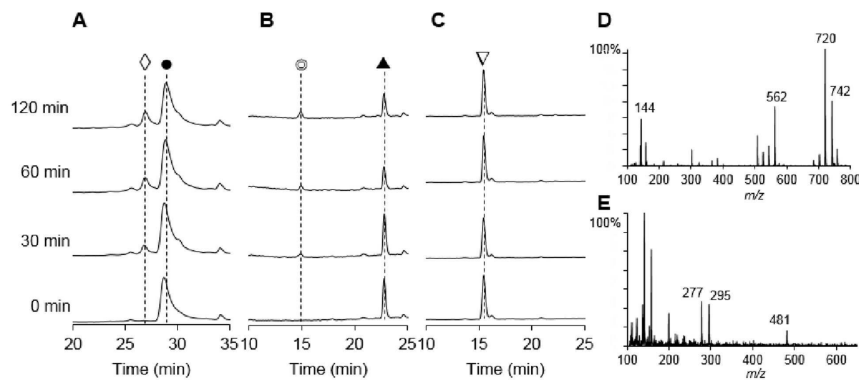
도면6



도면7



도면8



서열 목록

- <110> Ewha University - Industry Collaboration Foundation
- <120> Novel macrolactam glycoside derivatives, chemoenzymatic method
for preparation thereof and antibacterial composition comprising
the same
- <130> KPA141378-KR
- <160> 18
- <170> KopatentIn 2.0
- <210> 1
- <211> 6603
- <212> DNA
- <213> Artificial Sequence
- <220><223> pLRHM2_desVIII-desVII-desIII-desIV-oleL-oleU
- <400> 1

ttaattaaga tatccccgac cagccttatac gaaggagcgg cgaagagatg gcagaaccca	60
cggtgaccga cgacctgacg ggggccctca cgcagccccc gctgggccgc accgtccgcg	120
cggtggccga ccgtgaactc ggcacccacc tcttgagac ccgcggcatc cactggatcc	180
acgcccgaa cggcgacccg tacgccaccg tgctgcgcgg ccaggcggac gaccctatc	240
ccgcgtacga gcggtgctg gcccgcggcg cgtctctctt cagcccgacg ggcagctggg	300
tcaccgccga tcacgcctg gcggcgagca tcctctgtc gacggacttc ggggtctccg	360
gcgccgacgg cgtcccgtg ccgcagcagg tcctctcgta cggggagggc tgtccgtgg	420
agcgcgagca ggtgctgccg gcggccggtg acgtgccgga gggcgggcag cgtgccgtgg	480
tcgaggggat ccaccgggag acgctggagg gtctcgccg ggacccgtcg gcgtcgtacg	540
ccttcgagct gctgggcgtt ttcttccgcc cggcgggtgac ggccgctgcc gccgccgtgc	600

tgggtgttcc cgcgaccgg cgcgcgact tcgcggatct gctggagcgg ctccggccgc	660
tgtccgacag cctgctggcc ccgcagtcct tgcggacggt acgggcggcg gacggcgcg	720
tggccgagct caccggcgtg ctgcccatt cggacgactc ccccggggcc ctgctgtcgg	780
cgctcggggt caccgcagcc gtccagctca ccgggaacgc ggtgctcgcg ctctcgcgc	840
atcccagca gtggcgggag ctgtgcgacc ggcccgggct cgcgcgccg gcggtggagg	900
agacctccg ctacgacccg ccggtgcagc tcgacgcccg ggtggtccgc ggggagacgg	960
agctggcggg ccggcggctg ccggccgggg cgcatgtcgt cgtcctgacc gccgcgaccg	1020
gccgggaccc ggaggtcttc accgacccgg agcgttcga cctcgcgcgc cccgacgccc	1080
ccgcgcacct cgcgtgcac cccgccgtc cgtacggccc ggtggcgtcc ctggtccggc	1140
ttcaggcgga ggtcgcgtg cggaccctgg ccggcgcttt ccccgggctg cggcaggcgg	1200
gggacgtgct cccccccg cgcgcgcctg tcggcccggg gccgtgagc gtcccgtca	1260
gcagctcctg agacaccggg gcccgggtcc gcccggcccc cttcggagc gaccggacgg	1320
ctcggaccac ggggacggct cagaccgtcc cgtgtgtccc cgtccggctc ccgtccgcc	1380
catccgccc ctccaccggc aaggaaggac acgacgcat gcgcgtcctg ctgacctgt	1440
tcgcacatca cagcactac tacggcctgg tgcccctggc ctgggcgtg ctgcgcgcc	1500
ggcacgaggt gcggtcgc agccagcccg cgtcacgga caccatcac gggtccgggc	1560
tcgcccggt gccggtcgc accgaccacc tcattcacga gtaccgggtg cggatggcgg	1620
gcgagccgcg ccgaacct ccggcgatcg cttcgacga ggcccgtccc gagccgtgg	1680
actgggacca cgcctcggc atcgaggcga tcctcgcgcc gtacttccat ctgctcgcca	1740
acaacgactc gatggtcgac gacctcgtc acttcgccg gtcttgagc cggaacctgg	1800
tgtgtggga gccgacgacc tacgcgggcg ccgtcgcgc ccaggtcac ggtgccgcgc	1860
acgccgggt cctgtggggg cccgacgtga tgggcagcgc ccgcgcaag ttcgtcgcgc	1920
tgcgggaccg gcagccgcc gagcaccgc aggacccac cgcgagtggt ctgacgtgga	1980
cgtcgaccg gtacggcgc tccttgaag aggagctgct caccggccag ttcacgatcg	2040
accgacccc gccgagcctg cgctcgaca cgggcctgcc gaccgtcggg atgcgttatg	2100
ttccgtacaa cggcacgtc gtctgcccgg actggctgag tgagccgcc gcgcggcccc	2160
gggtctgct gacctcggc gtctccgcg gtgaggtcct cggcggcgac ggcgtctgc	2220
agggcgacat cctggaggcg ctgcgcgacc tcgacatcga gctcgtgcc acgctcgacg	2280
cgagtcagcg gccgagatc cgcaactacc cgaagcacac ccggttcacg gacttcgtgc	2340
cgatgcacg gctcctgcc agctgctcgg cgatcatcca ccacggcggg gcgggcacct	2400
acgcgaccgc cgtgatcaac gcggtgccg aggtcatgct ccgagagctg tgggacgcgc	2460

cggtcaaggc gcgggccgtc gccgagcagg gggcggggtt cttcctgccg ccggccgagc	2520
tcacgccgca ggccgtgcgg gacgccgtcg tccgcatcct cgacgacccc tcggtcgcca	2580
ccgccgca ccggctgcgc gaggagacct tcggcgaccc caccgccgcc gggatcgtec	2640
ccgagctgga gcggctcgcc gcgcagcacc gccgccgcc ggccgacgcc cggcactgag	2700
ccgcacccct cgcgccaggc ctcacccctg tatctcatat gtctagttaa ctgccacgc	2760
cgaccgttat caccggcgcc ctgctgctag ttccgagaa tgaagggaat agtcctggcc	2820
ggcgggagcg gaactcggtt gcatccggcg acctcggtca ttctgaagca gattcttccg	2880
gtctacaaca aaccgatgat ctactatccg ctgtcggttc tcatgctcgg cggtattcgc	2940
gagattcaaa tcatctcgac ccccagcac atcgaactct tccagtcgt tctcggaac	3000
ggcaggcacc tgggaataga actcgactat gcggtccaga aagagccgc aggaatcgcg	3060
gacgcacttc tcgtcgagc cgagcacatc ggcgacgaca cctgcgccct gatcctgggc	3120
gacaacatct tcacgggcc cggcctctac acgctcctgc gggacagcat cgcgcgcctc	3180
gacggctcgc tgctcttcgg ctaccggctc aaggaccccg agcggtagcg cgtcgccgag	3240
gtggacgca cgggcgggtt gaccgacctc gtcgagaagc ccgtcaagcc gcgtccaac	3300
ctcgccgtca cggcctcta cctctacgac aacgacgtcg tcgacatcg caagaacatc	3360
cggccctcgc cgcgcggcga gctggagatc accgacgtca accgcgtcta cctggagcgg	3420
ggccgggccg aactcgtcaa cctgggccgc ggcttcgcct ggctggacac cggcacccac	3480
gactcgtcc tcggggccgc ccagtacgtc caggtcctgg aggagcgga gggcgtctgg	3540
atcggggcc ttgaggagat cgccttcgc atgggcttca tcgacgccga ggcctgtcac	3600
ggcctgggag aaggcctctc ccgaccgag tacggcagct atctgatgga gatcgccggc	3660
cgcgagggag cccgtgagg gcacctcgc gccgacgct tcccagacc gacagcgcca	3720
ccgacagtgc gaccacacc gcgaccgca ccgccaccga cagtgcgacc cacaccgga	3780
cctacagcgc gaccgaaagg aagacggcag tgcggcttct ggtgaccgga ggtgcgggct	3840
tcatcggtc gcaattcgtg cggcagctcc tcgccgggac gtaccccgac gtgcccgccg	3900
atgaggtgat cgtcctggac agcctcacct acgcgggcaa ccgcgccaac ctgcgccgg	3960
tggacgcca cccgcactg cgcttcgtcc acggcgacat ccgcgacgcc ggcctcctcg	4020
cccgggaact gcgcggcgtg gacgccatcg tccactcgc ggccgagagc cacgtggacc	4080
gtccatcgc gggcgctgcc gtgttcaccg agaccaacgt gcagggcacg cagacgtgc	4140
tccagtgcgc cgtcgacgcc ggcgtcgcc gggtcgtgca cgtctccacc gacgaggtgt	4200

acgggtcgat cgactccggc tcctggaccg agagcagccc gctggagccc aactcgccct	4260
acgcggcgtc caaggccggc tccgacctcg ttgcccgcgc ctaccaccgg acgtacggcc	4320
tcgacgtacg gatcacccgc tgctgcaaca actacgggccc gtaccagcac cccgagaagc	4380
tcacccccct cticgtgacg aacctcctcg acggcgggac gctcccgtg tacggcgacg	4440
gcgcgaacgt ccgcgagtgg gtgcacaccg acgaccactg ccggggcatc gcgctcgtec	4500
tcgcgggcgg ccgggccggc gagatctacc acatcggcgg cggcctggag ctgaccaacc	4560
gcgaactcac cggcatcctc ctggactcgc tcggcgccga ctggctcctg gtccggaagg	4620
tcgccgaccg caagggccac gacctgcgct actccctcga cggcggcgag atcgagcgcg	4680
agctcggcta ccgcccgcag gtctccttcg cggacggcct cgcgcgacc gtccgctggt	4740
accgggagaa ccgcggctgg tgggagccgc tcaaggcgac cggcccgcag ctgcccacca	4800
ccgcgtgga ggtgtccggt tgagcagccg ccgcgagacc ccccggtcc ccttcctcga	4860
cctcaaggcc gcctacgagg agctctctag tatcgctccg agcccgaagg gaaaatcgag	4920
tgcccaattc ggcagaatcg ggatcgatgg agttactcga cgtcgacggg gcctggttat	4980
acaccccgga aatcatcgcg gacgaacggg gagaattcct cgaatggttc cggggtcgga	5040
cattccagga gaagatcggc caccctctct cgctggccca ggccaactgc tcggtgtccc	5100
gcaaggcgtt ctgcgcggca tccacttcgc cgacgcccc gcccgccag gccaaagtacg	5160
tcacctgcgc ctccggcacc gtgctcgacg tggctgtcga cgtacgccgg ggctcgccca	5220
ccttcggccg atgggccgcc gtccgactcg acgcggcccg ccaccagggg ctctacctgg	5280
ccgaaggact cggccacgcg ttcattggccc tcaccgacga cgccacggtc gtctacctct	5340
gtcacagcc ctacgtggcc gaggccgagc ggcccgtaga ccctctcgac ccggcgatcg	5400
gcatcgaatg gccgacggac atcgacatcg tcctgtcgg cgaagggaca cccacgcacc	5460
gtccctggcg caggccgcgg agaccggcga tcctgccgga ctacgaggga gtgcccgcg	5520
ccttacatcg cggaggcggc cggcgtggga ccggcccgtg aaggccctcg tactggccgg	5580
ctctagttag gtagtacgc gacaaccgcg cctggtggga acccctgaag aagcgcccg	5640
ccggtcccgc cccccccc agaggcagcg gcccatgaga tggctgatca ccggcgccgc	5700
cggaatgctg ggccgggaac tcgtccggcg cctcgccgag aacgaggagg acgtcgcggc	5760
cctcgccac gaccacctcg acgtaccgc accctcgcc gtgcgggcgg cactcgccga	5820
gcaccgtccc gggatcgtcg tcaactgcgc gcctacacg gccgtcgacg acgccgagac	5880
ggacgaggcc gccgctgccc tcctcaacgc cgaggcgccc cggctgctgg ccgaggccct	5940

gcgccccac cggcggcacg gcctcgtcca cctgtccacc gactacgtct tccccggcga 6000
cgccccaccc ccctacgccg aggaccaccc cagggtccc cgcagcgctt acggacgcac 6060
caaacgggac ggcgagcaag cgggtctgac ggactgccc accgccaccg tgctgcgcac 6120
cgcttggtg tacgggcgca ccggccgcag ctctgtccgc acgatgatcg aacgggaggc 6180
gcgcggcgga gccatcgacg tcgtcgccga ccagcgcggc cagcccacct ggaccggcga 6240
cctcgccgac cgcacatcg ccgtcggccg gcaccccggc gtccacggca tcctgcacgc 6300
caccaacgcc ggctccgcca cctggtacga cctggcaca gaggtcttcc ggctcctcga 6360

cgccgacccc gggcgggtcc ggcccaccac cggcgccgcc ttccgcagac ccgccccccg 6420
ccccgcctac agcgtcctcg gccacgaccg ctggcgcggg accggcctgg caccctgcg 6480
tgactggcgc tcggccctgc gcgaggcgtt ccccgacatc ctgcccgcg aacaccacc 6540
gaccggcgga ggagcagcat gaaacgaggc gtgcacgacc tggccctctt ccctaggtct 6600
aga 6603

<210> 2

<211> 4891

<212> DNA

<213> Artificial Sequence

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

<400> 2

ttaattaaga tatccccgac cagccttatt gaaggagcgg cgaagagatg gcagaacca 60
cggtagccga cgacctgacg ggggcccctca cgcagccccc gctgggcccgc accgtccgcg 120
cggtagccga ccgtgaactc ggcacccacc tcctggagac ccgcggcatc cactggatcc 180
acgccgcgaa cggcgacccg tacgccaccg tgctgcgcgg ccaggcggac gaccctatc 240
ccgcgtacga gcgggtgcgt gcccgcggcg cgtctctctt cagcccagc ggagctggg 300
tcaccgccga tcacgccctg gcggcgagca tcctctgctc gacggacttc ggggtctccg 360
gcgccgacgg cgtcccgtg ccgcagcagg tcctctcgta cggggagggc tgtccgtgg 420

agcgcgagca ggtgctgccg gcggccggtg acgtgccgga gggcgggcag cgtgccgtgg 480
tcgaggggat ccaccgggag acgtggagg gtctcgccgc ggaccctcg gcgtcgtacg 540
ccttcgagct gctgggcggt ttcttccgcc cggcggtagc ggccgtgcc gccgccgtgc 600
tgggtgttcc cgcggaccgg cgcgcggact tcgcggatct gctggagcgg ctccggccgc 660
tgtccgacag cctgctggcc ccgcagtcct tcgggacggt acgggcggcg gacggcgcgc 720
tgccgagct caccgcgtg ctgccgatt cggacgactc cccgggggcc ctgctgtcgg 780

cgctcggggt caccgcagcc gtccagctca ccgggaacgc ggtgctcgcg ctcttcgcmc	840
atcccagca gtggcgggag ctgtgcgacc ggcccgggct cgcggcggcc gcggtggagg	900
agacctccg ctacgacctg ccggtgcagc tcgacgcccg ggtggtccgc ggggagacgg	960
agctggcggg ccggcggctg ccggccgggg cgcattgctg cgtcctgacc gccgcgaccg	1020
gccgggagcc ggaggtcttc accgacctgg agcgtctcga cctcgcgcgc cccgacgccg	1080
ccgcgcacct cgcgctgcac cccgccggtc cgtacggccc ggtggcgtcc ctggtccggc	1140
ttcaggcgga ggtcgcgctg cggacctgg ccggcggttt ccccgggctg cggcaggcgg	1200
gggacgtgct ccgccccgc cgcgcgcctg tcggccgagg gccgctgagc gtcccgtca	1260
gcagctcctg agacaccggg gccccggtcc gcccgcccc ccttcggacg gaccggacgg	1320
ctcggaccac ggggacggct cagaccgtcc cgtgtgtccc cgtccggctc ccgtccgcc	1380
catccgccc ctccaccggc aaggaaggac acgacgcat gcgcgtcctg ctgacctgt	1440
tcgcacatca cagcactac tacggcctgg tgcccctggc ctgggcgctg ctgcgcgccg	1500
ggcacgaggt gcgggtcgcc agccagcccg cgtcacgga caccatcac gggtccgggc	1560
tcgcccggt gccggtcgcc accgaccacc tcattcacga gtaccgggtg cggatggcgg	1620
gcgagccgcg ccgaacct cggcgatcg ccttcgacga ggcccgccc gagccgtgg	1680
actgggacca cgcctcgcc atcgaggcga tcctcgcccc gtacttccat ctgctcgcca	1740
acaacgactc gatggtcgac gacctgtcg acttcgccc gtcctggcag ccggacctgg	1800
tgctgtggga gccgacgacc tacgcgggcg ccgtcgccgc ccaggtcac ggtgccgcgc	1860
acgcccggt cctgtggggg cccgacgtga tgggcagcgc ccgccgaag ttcgtcggc	1920
tcggggaccg gcagccgcc gagcaccgcg aggacccac cgcggagtgg ctgacgtgga	1980
cgctcgaccg gtacggcgcc tccttcgaag aggagctgct caccggccag ttcacgatcg	2040
accgacccc gccgagcctg cgcctcgaca cgggcctgcc gaccgtcggg atgcgttatg	2100
ttccgtacaa cggcacgtcg gtcgtgccgg actggctgag tgagccccc gcgcggcccc	2160
gggtctgctt gacctcgcc gtctccgcgc gtgaggtcct cggcggcgac ggcgtctgc	2220
agggcgacat cctggaggcg ctgcgcgacc tcgacatcga gtcgtcgcc acgctcgacg	2280
cgagtacgcg gccgagatc cgcaactacc cgaagcacac ccggttcacg gacttcgtgc	2340
cgatgcacgc gtcctcgcc agctgctcgg cgatcatcca ccacggcggg gcgggcacct	2400
acgcgaccgc cgtgatcaac gcggtgccgc aggtcatgct cgcgagctg tgggacgcgc	2460
cggtcaaggc gcggccgct gccgagcagg gggcggggtt cttcctgcc ccggccgagc	2520

tcacgccgca ggccgtgcgg gacgccgtcg tccgcatcct cgacgacccc tcggtcgcca	2580
ccgccgcgca ccggctgcgc gaggagacct tcggcgaccc caccgccgcc gggatcgccc	2640
ccgagctgga gcggctcgcc gcgcagcacc gccgcccgcc ggccgacgcc cggcactgag	2700
ccgcacccct cgcgccaggc ctaccccttg tatctcatat gtctagttaa ctcgccacgc	2760
cgaccgttat caccggcgcc ctgctgctag ttccgagaa tgaagggaat agtcctggcc	2820
ggcgggagcg gaactcggtt gcatccggcg acctcggtca ttccgaagca gattcttccg	2880
gtctacaaca aaccgatgat ctactatccg ctgtcggttc tcatgctcgg cggtatcgc	2940
gagattcaaa tcattctgac cccccagcac atcgaactct tccagtcgct tctcggaac	3000
ggcaggcacc tgggaataga actcgactat gcggtccaga aagagccgc aggaatcgcg	3060
gacgcacttc tcgtcggagc cgagcacatc ggcgacgaca cctgcgcctt gatcctgggc	3120
gacaacatct tccacgggccc ggccctctac acgctcctgc gggacagcat cgcgcgcctc	3180
gacggctgcg tgctcttcgg ctaccgggtc aaggaccccg agcggtagcg cgtcgccgag	3240
gtggacgcga cgggccggct gaccgacctc gtcgagaagc ccgtcaagcc gcgtccaac	3300
ctcgccgtea cggccctcta cctctacgac aacgacgtcg tcgacatcgc caagaacatc	3360
cggccctcgc cgcgcggcga gctggagatc accgacgtca accgcgtcta cctggagcgg	3420
ggccggggccg aactcgtcaa cctggggcgc ggcttcgcct ggctggacac cggcaccac	3480
gactcgctcc tgcggggcgc ccagtacgtc caggctcctgg aggagcggca gggcgtctgg	3540
atcgccggcc ttgaggagat cgccttcgc atgggcttca tcgacgccga ggccgtcac	3600
ggcctgggag aaggcctctc ccgcaccgag tacggcagct atctgatgga gatcgccggc	3660
cgcgagggag cccgtgagg gcacctcgc gccgacgct tcccacgacc gacagcgcca	3720
ccgacagtgc gaccacacc gcgacccgca ccgccaccga cagtgcgacc cacaccgca	3780
cctacagcgc gaccgaaagg aagacggcag tgcggcttct ggtgaccgga ggtgcgggt	3840
tcattcggtc gcattctgtg cggcagctcc tcgccggggc gtaccccgac gtgccgccg	3900
atgagtgat cgtcctggac agcctcacct acgcgggcaa ccgcgccaac ctgcgcccg	3960
tggacgcgga cccgcgactg cgcttcgtcc acggcgacat ccgcgacgcc ggcctcctcg	4020
cccgggaact gcgcggcgtg gacgccatcg tccattcgc ggccgagagc cacgtggacc	4080
gctccatcgc gggcgcttc gtgttcaccg agaccaacgt gcagggcacg cagacgtgc	4140
tccagtgcgc cgtcgacgcc ggcgtcgcc gggtcgtgca cgtctccacc gacgaggtgt	4200
acgggtcgat cgactccggc tccctggaccg agagcagccc gctggagccc aactcgcct	4260
acgcggcgtc caaggccgga tccgacctcg ttgcccgcc ctaccaccgg acgtacggcc	4320
tcgacgtacg gatcaccgc tgctgcaaca actacgggcc gtaccagcac cccgagaagc	4380

tcacccccct cticgtgacg aacctcctcg acggcgggac gctcccgtg tacggcgacg	4440
gcgcgaacgt ccgcgagtggt gtgcacaccg acgaccactg ccggggcatc gcgctcgtcc	4500
tcgcgggcgg ccgggccggc gagatctacc acatcggcgg cggcctggag ctgaccaacc	4560
gcgaactcac cggcatcctc ctggactcgc tcggcgccga ctggctcctg gtcggaagg	4620
tcgccgaccg caagggccac gacctgcgt actccctcga cggcggcgag atcgagcgcg	4680
agctcggcta ccgccgcgag gtctccttcg cggacggcct cgcgcggacc gtccgtggt	4740
accgggagaa ccgcggctgg tgggagccgc tcaaggcgac cggcccgag ctgcccgcca	4800
ccgcgtgga ggtgtccgcg tgagcagccg cgcgcgagacc ccccgctcc ctttcctcga	4860
cctcaaggcc gcctacgagg agctctctag a	4891
<210> 3	
<211> 9715	
<212> DNA	
<213> Artificial Sequence	
<220><223> pDDSS_desVIII-desVII-desIII-desIV-desI-desII-desV-desVI	
<400> 3	
ttaattaaga tatccccgac cagccttatt gaaggagcgg cgaagagatg gcagaacca	60
cgggtgaccga cgacctgacg ggggccctca cgcagccccc gctgggccgc accgtccgcg	120
cgggtggccga ccgtgaactc ggcaccacc tcctggagac ccgcggcatc cactggatcc	180
acgccgcgaa cggcgaccgc tacgccaccg tgcctgcgcg ccaggcggac gaccctatc	240
ccgcgtacga gcgggtgcgt gcccgcggcg cgtctcctt cagcccgacg ggcagctggg	300
tcaccgccga tcacgccctg cgcgcgagca tcctctctc gacggacttc ggggtctccg	360
gcgccgacgg cgtcccgggt ccgcagcagg tcctctcgta cggggagggc tgtccgtgg	420
agcgcgagca ggtgctgccg gcggccgggt acgtgccgga gggcgggcag cgtgccgtgg	480
tcgaggggat ccaccgggag acgtggagg gtctcgccg ggaccctcg gcgtcgtacg	540
ccttcgagct gctgggcggt ttcttccgc cggcgggtgac ggccgctgcc gccgccgtgc	600
tgggtgttcc cgcggaccgg cgcgcggact tcgcggatct gctggagcgg ctccggccgc	660
tgtccgacag cctgctggcc ccgcagtcg tcgcggacgt acgggcggcg gacggcgcg	720
tggccgagct caccgcgtg ctgcgcgatt cggacgactc ccccggggcc ctgctgtcgg	780
cgtcgggggt caccgcagcc gtccagctca ccgggaacgc ggtgctcgcg ctctcgcgc	840
atcccgagca gtggcgggag ctgtgcgacc ggcccgggct cgcggcggcc gcggtggagg	900
agaccctccg ctacgaccgc ccggtgcagc tcgacgccg ggtggtccgc ggggagacgg	960

agctggcggg ccggcggctg ccggccgggg cgcattgtct cgtcctgacc gccgcgaccg	1020
gccgggaccc ggaggtcttc acggacccgg agcgtcttca cctcgcgcgc cccgacgccg	1080
ccgcgcacct cgcgtgcac cccgccgtc cgtacggccc ggtggcgtcc ctggtccggc	1140
ttcaggcgga ggtcgcgtc cggaccctgg ccgggcgttt ccccggtcgt cggcaggcgg	1200
gggacgtgct ccgccccgc cgcgcgcctg tcggccgcgg gccgtgagc gtcccgtca	1260
gcagctcctg agacaccggg gccccggtcc gcccgcccc cttcggacg gaccggacgg	1320
ctcggaccac ggggacggct cagaccgtcc cgtgtgtccc cgtccggctc ccgtccgcc	1380
catccgccc ctccaccggc aaggaaggac acgacgcat gcgcgtctc ctgacctcgt	1440
tcgcacatca cagcactac tacggcctgg tgccccggc ctgggcgtc ctcgccgcc	1500
ggcacgaggt gcgggtgcc agccagccc cgtcacgga caccatcacc ggtccgggc	1560
tcgccgggt gccggtcggc accgaccacc tcattcacga gtaccgggtg cggatggcgg	1620
gcgagccgcg cccgaacct cggcgatcg ctttcgacga ggcccgctcc gagccgtgg	1680
actgggacca cgcctcggc atcgaggcga tcctcgcgcc gtacttccat ctgctcgcca	1740
acaacgactc gatggtcgac gacctcgtc atttcgccg gtcttgccag ccggacctgg	1800
tgctgtggga gccgacgacc tacgcgggcg ccgtcgcgc ccaggtcacc ggtgccgcgc	1860
acgccgggt cctgtgggg cccgacgtga tgggcagcgc ccgccgaag ttcgtcgcgc	1920
tcggggaccg gcagccgcc gagcaccgc aggacccac cgcggagtgg ctgacgtgga	1980
cgtcgcacc gtacggccc tccttcgaag aggagctgt caccggccag ttcacgatcg	2040
acccgacccc gccgagcctg gcctcgaca cgggcctgcc gaccgtcggg atgcgttatg	2100
ttccgtacaa cggcacgtcg gtcgtgccgg actggctgag tgagccgcc gccgcggccc	2160
gggtctgcct gacctcggc gtctccgcgc gtgaggtcct cggcgccgac ggcgtctcgc	2220
agggcgacat cctggaggcg ctgccgacc tcgacatcga gctcgtgcc acgctcgacg	2280
cgagtacgc gcccgagatc cgcaactacc cgaagcacac ccggttcacg gacttcgtgc	2340
cgatgcacgc gtcctgccg agctgctcgg cgatcatcca ccacggcggg gcgggcacct	2400
acgcgaccgc cgtgatcaac gcggtgccgc aggtcatgct cgcgagctg tgggacgcgc	2460
cggtaaggc gcgggcccgc gccgagcagg gggcggggtt cttcctgcc cggccgagc	2520
tcacgccga gcccggtcgg gacgccgtcg tcgcatacct cgacgacccc tcggtcgcca	2580
ccgccgcga ccggtcgcgc gaggagacct tcggcgaccc caccgccgc gggatcgtcc	2640
ccgagctgga gcggctgcc gcgcagcacc gcccccgc gcccgacgc cggcactgag	2700
ccgacccct cgcgccagg ctcacccctg tatctcatat gtctagttaa ctcgccacgc	2760
cgaccgttat caccggcgcc ctgctgctag ttccgagaa tgaagggaat agtcctggcc	2820

ggcgggagcg gaactcggct gcatccggcg acctcggta tttcgaagca gattcttccg	2880
gtctacaaca aaccgatgat ctactatccg ctgtcggttc tcatgctcgg cggatttcgc	2940
gagattcaaa tcatctcgac cccccagcac atcgaactct tccagtcgct tctcggaaac	3000
ggcaggcacc tgggaataga actcgactat gcggtccaga aagagcccg aggaatcgcg	3060
gacgcacttc tctcggagc cgagcacatc ggcgacgaca cctgcgcct gatcctgggc	3120
gacaacatct tccacgggcc cggcctctac acgctcctgc gggacagcat cgcgcgcctc	3180
gacggctcgc tgctcttcgg ctaccgggtc aaggaccccg agcggtagcg cgtcgccgag	3240
gtggacgca cgggccggct gaccgacctc gtcgagaagc ccgtcaagcc gcgtccaac	3300
ctcgccgtca cggcctcta cctctacgac aacgacgtcg tcgacatcgc caagaacatc	3360
cggcctcgc cgcgcggcga gctggagatc accgacgtca accgcgtcta cctggagcgg	3420
ggccggggcg aactcgtcaa cctggggcgc ggcttcgcct ggctggacac cggcacccac	3480
gactcgtcc tgcggggcgc ccagtcgtc caggtcctgg aggagcggca gggcgtctgg	3540
atcggggcc ttgaggagat cgccttcgc atgggcttca tcgacgccga ggcctgtcac	3600
ggcctgggag aaggcctctc ccgcaccgag tacggcagct atctgatgga gatcgccggc	3660
cgcgagggag ccccgtagg gcacctcgc gccgacgct tcccacgacc gacagcgcca	3720
ccgacagtgc gaccacacc gcgaccgca ccgccaccga cagtgcgacc cacaccgca	3780
cctacagcgc gaccgaaagg aagacggcag tgcggcttct ggtgaccgga ggtgcgggct	3840
tcatcggtc gacttcgtg cggcagctcc tcgccggggc gtaccccgac gtgcccgccg	3900
atgagtgat cgtcctggac agcctcacct acggggcaa ccgcgccaac ctgccccgg	3960
tggacgcga cccgcgactg cgcttcgtcc acggcgacat ccgcgacgc ggcctcctcg	4020
cccgggaact gcggcgctg gacgccatc tccacttcgc ggccgagagc cacgtggacc	4080
gtccatcgc gggcgctcc gtgttcaccg agaccaacgt gcagggcag cagacgtgc	4140
tccagtgcgc cgtcgacgc ggcgtcggc gggtcgtgca cgtctccacc gacgaggtgt	4200
acgggtcgat cgactccgc tctggaccg agagcagccc gctggagccc aactcgcct	4260
acgcggcgtc caaggccgc tccgacctg ttccccgc ctaccaccg acgtacggcc	4320
tcgacgtac gatcaccgc tctgcaaca actacgggc gtaccagcac cccgagaagc	4380
tcatccccct ctctgtgacg aaactcctc acggcgggac gctcccgctg tacggcgacg	4440
gcgcgaacgt ccgcgagtg gtgcacaccg acgacctg ccggggcatc gcgtcgtcc	4500
tcggggcgcg cggggccgc gagatctacc acatcgcgcg cggcctggag ctgaccaacc	4560

gcgaactcac cggcatcctc ctggactcgc tcggcgccga ctggtcctcg gtccggaagg	4620
tcgccgaccg caagggccac gacctgcgct actccctcga cggcggcgag atcgagcgcg	4680
agctcggcta ccgccgcgag gtctccttcg cggacggcct cgcgcggacc gtccgctggt	4740
accgggagaa ccgcggctgg tgggagccgc tcaaggcgac cggcccgag ctgcccgcga	4800
ccgccgtgga ggtgtccgag tgagcagccg cgcgcgagacc ccccgctcc ccttcctcga	4860
cctcaaggcc gctacgagg agctctctag tgctagcgac ggtggcccgg agggaaacatc	4920
cgtgaaaagc gccttatccg acctcgcat tctcggcggc cccgccgctt tcgaccagcc	4980
gctcctcgtg gggcgggcca accgcatcga ccgcgccagg ctgtacgagc ggctcgaccg	5040
ggccctcgac agccagtggc tgtccaacgg cggcccgtc gtccgcgagt tcgaggagcg	5100
cgtcgccggg ctgcggggg tccggcatgc cgtggccacc tgcaacgcca cggccgggct	5160
ccagctcctc gcgcacgcc cggcctcacc cggcgaagt atcatgccgt cgatgacgtt	5220
cgcgccacc ccgcacgcac tgcgtggat cggcctcacc cggctcttcg ccgacatcga	5280
cccggacacc ggcaacctcg acccggacca ggtggccgcc gcggtcacac cccgcacctc	5340
ggcgtcgtc ggcgccacc tctggggccg ccctgcgcc gccgaccagc tgcggaaggt	5400
cgcgcagcag caggcctgc ggctgtactt cgacgcccg cagccctcg gctgcgcggt	5460
cgacggccgg cccgccgga gcctcggcga cgcgaggtc ttcagcttc acgccacaa	5520
ggccgtcaac gccttcgagg gcggcgccgt cgtaccgac gacgccgacc tcgccgccc	5580
gatccgcgcc ctccacaact tcggcttcga cctgccggc ggagccccg ccggcgggac	5640
caacgccaag atgagcagg ccgcgcgcgc catgggcctc acctccctcg acgcgtttcc	5700
cgaggtcatc gaccggaacc ggcgcaacca cgcgcctac cgcgagcacc tcgcggacct	5760
ccccggcgtc ctctgcgcg accacgacc ccacggcctc aacaaccacc agtacgtgat	5820
cgtcgagatc gacgaggcca ccaccggcat ccaccgcgac ctctcatgg aggtcctgaa	5880
ggccgaaggc gtgcacacc gcgcctactt ctgcggggc tgccacgagc tggagccgta	5940
ccgcggcgag ccgcacgcc cgctgccgca caccgaacgc ctgcgcgcg cggtgctgtc	6000
cctgccgacc ggcaccgcca tcggcgacga cgacatccgc cgggtcgccg acctgctgcg	6060
tctctgcgcg acccgcggcc gcgaactgac cgcgcgccac cgcgacacgg ccccccccc	6120
gctcgcggcc cccagacat ccacggccac gatggagcgc tccgatgac cgcggcgcc	6180
ctttccgcca ccggccggc cgaacgctgc gcgcacccg gagccgatct gggggcggcg	6240
gtccacgccg tcggccagac cctcggccc ggcgccctcg tgccggccga cgaggccgga	6300

acgacgcgcc gccacctcgt ccggctcgcc gtgcgtacg gcaacagccc cttcaccccg	6360
ctggaggagg cccgccacga cctgggcgtc gaccgggacg ccttccggcg cctcctcgcc	6420
ctgttcgggc aggtcccga gctccgcacc gcggtcgaga ccggccccgc cggggcgtag	6480
tggaagaaca cctgctccc gctcgaacag cgcggcgtct tcgacgcggc gctcgccagg	6540
aagccctctt tcccgtagc cgctggcctc taccccgcc cgacctgcat gttccgctgc	6600
catttctgcg tccgtgtgac cggcggccgc tacgacctg ccgcctcga cggcggaac	6660
gccatgttcc ggctggctcat cgacgagata ccccgggca acccctcggc gatgtacttc	6720
tccggcgcc tggagccgt caccaacccc ggctcgga gcctggccgc gcacgccacc	6780
gaccacggcc tgcggccac cgtctacacg aactccttcg cgctcaccga gcgcacctg	6840
gagcgccagc ccggcctctg gggcctgcac gccatccga cctcgtcta cggcctcaac	6900
gacgaggagt acgagcagc caccggcaag aaggccgct tccgcccgt ccgcgagaac	6960
ctgcgccgt tccagcagct gcgcggcag cgcgagtcgc cgatcaacct cggcttcgcc	7020
tacatcgtgc tccggggccg tgcctccgc ctgctcgacc tggctgactt catcgccgac	7080
ctcaacgacg ccgggcaggg caggacgatc gacttcgtca acattcgga ggactacagc	7140
ggccgtgacg acggcaagct gccgcaggag gagcgggccc agctccagga ggccctcaac	7200
gccttcgagg agcgggtccg cgagcgcacc cccggactcc acatcgacta cggctacgcc	7260
ctgaacagcc tgcgcaccgg ggccgacgcc gaactgtgc ggatcaagcc cgccaccatg	7320
cggcccaccg cgcacccga ggtcgcggtg caggctgatc tcctcggcga cgtgtacctg	7380
taccgcgagg ccggcttccc cgacctggac ggcgcgacc gctacatgc gggccgctg	7440
accccgaca cctccctcac cgaggtcgtc agggacttcg tcgagcgcg cggcgaggtg	7500
gcggccgtcg acggcgacga gtacttcatg gacggcttcg atcaggctgt caccgcccgc	7560
ctgaaccagc tggagcgga cggcgggac ggtgggagg aggccgcgg cttcctgcgc	7620
tgaccgcgat ctagtagggt ctcttcgcg gacggcctcg cgcggaccgt ccgctggtag	7680
cgggagaacc gcggctggtg ggagccgtc aaggcgaccg ccccgagct gcccgcacc	7740
gccgtggagg tgtccgctg agcagccgcg ccgagacccc ccgctcccc ttcctcgacc	7800
tcaaggccgc ctacgaggag ctccgcgcgg agaccgacgc cgcgatgcc cgcgtcctcg	7860
actcggggcg ctacctctc ggaccgaac tcgaaggatt cgaggcgag ttcgcccgt	7920
actgcgagac ggaccagcc gtcggcgtga acagcgggat ggacgcctc cagctcgccc	7980
tccgcggcct cggcatcgga cccggggacg aggtgatcgt cccctcgac acgtacatcg	8040
ccagctggct cgcggtgtcc gccaccggcg cgacccccgt gcccgtag cgcacgagg	8100
accacccac cctggaccgc ctgctcgtcg agaaggcgat cccccccgc acccgggcg	8160

tcctccccgt ccacctctac gggcaccccg ccgacatgga cgcctccgc gagctcgcg	8220
accggcacgg cctgcacatc gtcgaggacg ccgcgcaggc ccacggcgcc cgctaccggg	8280
gccggcggat cggcgccggg tcgtcgggtgg ccgcgttcag cttctaccg ggcaagaacc	8340
tcggctgctt cggcgacggc ggcgccgtcg tcaccggcga ccccgagctc gccgaacggc	8400
tccggatgct ccgcaactac ggctcgcggc agaagtacag ccacgagacg aagggcacca	8460
actccgcct ggacgagatg caggccgccg tctgctggat ccggctcgcc cacctggaca	8520
gctggaacgg ccgcaggctg gcgctggccg cggagtacct ctccgggctc gccggactgc	8580
ccggcatcgg cctgccgtg accgcgccg acaccgacc ggtctggcac ctcttcaccg	8640
tgcgcaccga gcgcgcgac gagctgcgca gccacctcga cgcgcggcg atcgacacc	8700
tcacgacta cccggtacc gtgcacctct cgcgcgcta cgcgggcgag gcaccgccgg	8760
aaggctcgt cccgcgggc gagagcttcg cgcggcagg cctcagcctg ccgatcgcc	8820
cgcacctgga gcgcgcgag gcgctcggg tgatcgacgc cgtgcgcgaa tgggccgagc	8880
gggtcgacca ggcctagtc ggttctagtc aattgcccc ggctcacc ctgtatctgc	8940
gccgggggac gccccggcc caccctcga aagaccgaaa gcaggagcac cgtgtacgaa	9000
gtcgaccacg ccgacgtcta cgacctctt tacctgggtc gcggcaagga ctacgcgcc	9060
gaggcctccg acatcgccga cctggtgcgc tccgtacc cagaggctc ctgctcctg	9120
gacgtggcct gcggtacggg cagcatctg gagcattca ccaaggagt cggcgacacc	9180
gccggcctgg agctgtccga ggacatgct acccacgcc gcaagcggct gcccgacgc	9240
acgtccacc agggcgacat gcgggacttc cggctcgcc ggaagtctc cgccgtggtc	9300
agcatgttca gtcctgctg ctacctgaag acgaccgagg aactcggcgc gcccgctgcc	9360
tcgttcgagg agcacctgga gcccgtggc gtcgtcgtc tcgagccgtg gtggttccc	9420
gagaccttcg ccgacggctg ggtcagcgcc gacgtcgtc gccgtgacgg gcgcaccgtg	9480
gcccgtgtct cgactcgtt gcgggagggg aacgcgacgc gcatggaggt ccacttcacc	9540
gtggccgacc cgggcaagg cgtgcggcac ttctccgacg tccatctcat caccctgtc	9600
caccaggccg agtacgagg cgcggtcacg gccgccggg tcgcgctcga gtacctggag	9660
ggcggcccg cgggcccgtg cctcttcgtc ggcgtcccc cctgagtcta gacct	9715
<210> 4	
<211> 9085	
<212> DNA	
<213> Artificial Sequence	
<220><223> pL0LV2_desVIII-desVII-desIII-desIV-oleV-oleW-oleL-oleU	

<400> 4

ttaattaaga tatccccgac cagccttatc gaaggagcgg cgaagagatg gcagaaccca	60
cgggtgaccga cgacctgacg ggggccccca cgcagccccc gctgggccgc accgtccgcg	120
cgggtggccga ccgtgaactc ggcacccacc tcctggagac ccgcggcatc cactggatcc	180
acgccgcgaa cggcgaccgg tacgccaccg tgctgcgcgg ccaggcggac gacccgtatc	240
ccgcgtacga gggggtgcgt gcccgcggcg cgtctctctt cagcccagc ggcagctggg	300
tcaccgccga tcacgccctg gcggcgagca tcctctgctc gacggacttc ggggtctccg	360
gcgccgacgg cgtcccggcg ccgcagcagg tcctctcgta cggggagggc tgtccgctgg	420
agcgcgagca ggtgctgccg gcggccggcg acgtgccgga gggcgggcag cgtgccgtgg	480
tcgaggggat ccaccgggag acgctggagg gtctcgccg ggacccgtcg gcgtcgtacg	540
ccttcgagct gctgggcggt ttctgctgcc cggcgggtgac ggccgctgcc gccgcccgtc	600
tgggtgttcc cgcggaccgg cgcgcggact tcgcggatct gctggagcgg ctccggccgc	660
tgctcgacag cctgctggcc ccgcagtcctc tcgggacggg acgggcggcg gacggcgcg	720
tggccgagct caccgcgctg ctgcgcgatt cggacgactc ccccggggcc ctgctgtcgg	780
cgctcggggg caccgcagcc gtccagctca ccgggaacgc ggtgctcgcg ctctcgcgc	840
atcccagca gtggcgggag ctgtcgacc ggcccgggct cgcggcgcc gccgtggagg	900
agacctccg ctacgacccg ccgggtgcagc tcgacgccc ggtggtccgc ggggagacgg	960
agctggcggg ccggcggctg ccggccgggg cgcattgctg cgtcctgacc gccgcgaccg	1020
gccgggaccc ggaggtcttc accgacccgg agcgtctcga cctcgcgcg cccgacgccg	1080
ccgcgcacct cgcgctgcac cccgcggctc cgtacggccc ggtggcgtcc ctggtccggc	1140
ttcaggcgga ggtcgcgctg cggaccctgg ccgggcgttt ccccgggctg cggcaggcgg	1200
gggacgtgct cgcgcccgcc cgcgcgctg tcggccgagg gccgtgagc gtcccggtea	1260
gcagctcctg agacaccggg gcccgggtcc gcccgcccc ctttcggacg gaccggacgg	1320
ctcggaccac ggggacggct cagaccgtcc cgtgtgtccc cgtccggctc ccgtccgcc	1380
catccgccc ctccaccggc aaggaaggac acgacgcat gcgcgtcctg ctgacctcgt	1440
tcgcacatca cagcaactac tacggcctgg tgcccctggc ctgggcgctg ctgcgcccg	1500
ggcacgaggt gcgggtcgcc agccagcccc cgtcacgga caccatcacc gggtcgggc	1560
tcgcccggt gccggctggc accgaccacc tcattccaga gtaccgggtg cggatggcgg	1620
gcgagccgcg ccgaacat ccggcgatcg ctttcgacga ggcccgtccc gagccgtgg	1680
actgggacca cgcctcggc atcgaggcga tcctcgcccc gtacttccat ctgctcgcca	1740
acaacgactc gatggctgac gacctcgtcg acttcgcccc gtcctggcag ccggacctgg	1800

tgctgtggga gccgacgacc tacgcgggcg ccgtcgccgc ccaggtcacc ggtgccgcgc	1860
acgcccgggt cctgtggggg cccgacgtga tgggcagcgc ccgccgcaag ttcgtcgcg	1920
tgccggaccg gcagccgccc gagcaccgcg aggacccac cgccgagtgg ctgacgtgga	1980
cgctcgaccg gtacggcgcc tccttcgaag aggagctgct caccggccag ttcacgatcg	2040
acccgacccc gccgagcctg cgcctcgaca cgggcctgcc gaccgtcggg atgcgttatg	2100
ttccgtacaa cggcacgtcg gtctgtccgg actggctgag tgagccgccc gcgcggcccc	2160
gggtctgcct gaccctcggc gtctccgcgc gtgaggtcct cggcggcgac ggcgtctcgc	2220
agggcgacat cctggaggcg ctgcccgacc tcgacatcga gctcgtcgcc acgtcgacg	2280
cgagtacgcg cgccgagatc cgcaactacc cgaagcacac ccggttcacg gacttcgtgc	2340
cgatgcacgc gtcctgccc agctgctcgg cgatcatcca ccacggcggg gcgggcacct	2400
acgcgaccgc cgtgatcaac gcggtgccgc aggtcatgct cgccgagctg tgggacgcgc	2460
cggtaaggc gcgggcccgc gccgagcagg gggcggggtt ctctctgccg ccggccgagc	2520
tcacgccga gccggtcggg gacgccgtcg tccgcatcct cgacgacccc tcggtcgcca	2580
ccgccgcga ccggctgcgc gaggagacct tcggcgaccc caccgccgcc gggatcgtec	2640
ccgagctgga gcggctcgcc gcgcagcacc gcccccgcc ggccgacgcc cggcactgag	2700
ccgcacccct cgccccaggc ctcacccctg tatctcatat gtctagttaa ctgccacgc	2760
cgaccgttat caccggcgcc ctgctgctag ttccgagaa tgaagggaat agtcctggcc	2820
ggcgggagcg gaactcggtt gcatccggcg acctcggtca ttctgaagca gattcttcg	2880
gtctacaaca aaccgatgat ctactatccg ctgtcggttc tcatgctcgg cggtatctgc	2940
gagattcaaa tcatctcgac ccccagcac atcgaactct tccagtctgt tctcgaaac	3000
ggcaggcacc tgggaataga actcgactat gcggtccaga aagagcccgc aggaatcgcg	3060
gacgcacttc tcgtcggagc cgagcacatc ggcgacgaca cctgcgcctt gatcctgggc	3120
gacaacatct tccacgggcc cggcctctac acgtctctgc gggacagcat cgcggcctc	3180
gacggctgcg tgctcttcgg ctaccgggtc aaggaccccg agcggtacgg cgtcgccgag	3240
gtggacgcga cgggccggct gaccgacctc gtcgagaagc ccgtcaagcc gcgtccaac	3300
ctcgccgtca ccggcctcta cctctacgac aacgacgtcg tcgacatcgc caagaacatc	3360
cggccctcgc cgcgcggcga gctggagatc accgacgtca accgcgtcta cctggagcgg	3420
ggccgggccc aactcgtcaa cctgggccgc ggcttcgcct ggctggacac cggcaccac	3480
gactcgtcc tcggggccgc ccagtacgtc caggtcctgg aggagcggca gggcgtctgg	3540

atcgcgggcc ttgaggagat cgccttccgc atgggcttca tcgacgccga ggcctgtcac	3600
ggcctgggag aaggcctctc ccgcaccgag tacggcagct atctgatgga gatcgccggc	3660
cgcgagggag ccccgtaggg gcacctcgcg gccgacgcgt tcccacgacc gacagcgcca	3720
ccgacagtgc gacccacacc gcgacccgca ccgccaccga cagtgcgacc cacaccgca	3780
cctacagcgc gaccgaaagg aagacggcag tgcggcttct ggtgaccgga ggtgcgggct	3840
tcatcggtc gcacttcgtg cggcagctcc tcgccggggc gtaccccgac gtgcccgcg	3900
atgaggtgat cgtcctggac agcctcacct acgcgggcaa ccgcgccaac ctgccccgg	3960
tggacgcgga cccgcgactg cgcttcgtcc acggcgacat ccgcgacgcc ggcctcctcg	4020
cccgggaact gcgcggcgtg gacgccatcg tccacttcgc ggccgagagc cacgtggacc	4080
gtccatcgc gggcgctcc gtgttcaccg agaccaacgt gcagggcacg cagacgtgc	4140
tccagtgcgc cgtcgacgcc ggctcgcc gccgtcgtca cgtctccacc gacgaggtgt	4200
acgggtcgat cgactccggc tectggaccg agagcagccc gctggagccc aactcgccct	4260
acgcgcgctc caaggccggc tccgacctcg ttccccgcgc ctaccaccgg acgtacggcc	4320
tcgacgtacg gatcacccgc tgctgcaaca actacgggcc gtaccagcac cccgagaagc	4380
tcatccccct ctctgtgacg aacctcctcg acggcgggac gctcccgtg tacggcgacg	4440
gcgcgaacgt ccgcgagtgg gtgcacaccg acgaccactg ccggggcatc gcgtcgtcc	4500
tcgcggcgcg ccgggccggc gagatctacc acatcggcgg ccggcctggag ctgaccaacc	4560
gcgaactcac cggcatcctc ctggactcgc tcggcgccga ctggctcctg gtccggaagg	4620
tcgccgaccg caaggccac gacctgcgt actccctcga cggcggcgag atcgagcgcg	4680
agctcggtc cgcgccgag gtctccttcg cggacggcct cgcgcggacc gtccgtggt	4740
accgggagaa ccgcggctgg tgggagccgc tcaaggcgac cgcgccgag ctgcccgca	4800
ccgccgtgga ggtgtccgc tgagcagccg cgcgagacc ccccggtcc cttcctcga	4860
cctcaaggcc gcctacgagg agctctctag tgggaatcgc ggaagcggcc gggttcgcg	4920
ccgtcctgga tcacaatgat atggggaatt cccgcgatga gcgaagcaat gggatcggt	4980
ccgacggcgg gcagtgaagt ctctcgcacc tgcgcgttcc tgtcctggtt ggacgcgcgc	5040
cgccgggcca atcgctgac ggtggaacac gtaccgttca gggagtatac ggggtggcaa	5100
ttcgacgaga acacggggaa cctccgacat accagcggtc gtttcttctc catcgaagga	5160
ctccgggtac gcacggacca ctgctgggtc ggaagctgga cccagcccat tatcgtgcaa	5220
ccggagatag gcattctcgg cctcctggtc aagaggttcg acggcatcct gcacgtcctg	5280

gtgcaggcaa agaacgaacc gggtaataatc ggcggccttc agctctcccc caccgtccag	5340
gccactcgca gcaattacac ccgcgtccac cgcgccggcg gtgtcagata cctggagtac	5400
ttcgcgctcc cccgcgggcg cggtcgggtc ctccgggacg tctccaatc cgagcagggc	5460
tcctggttcc tgcacaagcg gaaccggaac atgggtggtcg aggcctgga cgacgtgccc	5520
ctcgacgacg acttccactg gatcagcctc ggcgggctgc ggaagctgct gctgaggccg	5580
catctggtga acatggacac ccgcacggtg ctgtcctgcc ttccccgga tccggcaccg	5640
gacggccggc agccgccggc gcccgcggct cccttcgccg ccgcggtcac gcggtccctc	5700
acccggggtg ccaccgcctt gcacaccatg ggcgagatcc tcggctggct gaccgacgag	5760
cgggtccggc gggaactggt gcagcagcgg gtgccgtgg aggagaccgc gttcagcggc	5820
tggcggcgtg acgaccacgc catcgcgcac aaggacggcg actacttccg ggtcatcggg	5880
gtgagcgtcc ggccagcag ccgggaggtg tcctcgtgga gccagccgtt gctggcccc	5940
gtcggccccg gtctggccgc ctctctcacc agcgcgatcc gcggcgtcct ccacgtcctg	6000
ctgcacgccc gcaccgaagc cggcctgtc aacggcccgg agatggcacc gaccgtgcag	6060
tgcgccccg tcaactaccg tgcggtgcc gccgagtacc ggcccgccta cctcgactac	6120
gtgctgtcgg ccgatcccg acgcatccgc tacgacacc tccagtcgga ggaggcgggc	6180
cgttccacc acgcggagaa ccgtacgtc gtggtggagg ccgaggacga ctccccgtc	6240
gaggtgcccc gcgacttccg ctggctcacc ctgcaccaga tctcgcct gctgcaccac	6300
agcaactatg tcaacgtgga ggcgcgacg ctcgctgcct gcattccagc tgagctgac	6360
tagtgctagc aagggaaccc catgccctcc ccccgctgc gcttcggtgt gctgggtgcc	6420
gccgacatcg ccttcgccg caccgtgcc gcgctgctc cccaccgga cgtcaccgtg	6480
gtcgcggtct ccagccggga caggccccg gccgccggtt tcgcccggc gttcgggtgc	6540
gaggcggtec cgggccacca ggcgtgctg gaccgcgacg acatcgacg gctgtacgtc	6600
ccgtgccgg tgatggtgca cagccctgg gtggaggcgg cgctgctgc gggcaggcac	6660
gtcctggtag agaagccgt gaccgcgacc cgctccggag ccgaggacct gatgccctg	6720
gcacgtccc gcggcctggt cctgatggag aacttcacgt cgctgcacca cgcacagcac	6780
ggcacgtca cggacctct cggggacggg acgatcggc aactgcgtc gctgtccgcg	6840
gcgttcagca tcccgcgaa gccggagggc gacatccgt accagcccga cgtgggcgga	6900
ggcgccctgc tcgacatcg gatctaccc ctgcggggcg cactgcactt cctgggcccc	6960
gacctgcac cgcccgccg ggtcctgcgt cgggagcggc gccggaacgt ggtggtctcc	7020
ggccacgtgc tctcaccac accgcacggc gtctggccg agctggcctt cggcatggag	7080
cacgcctacc gctcggagta caccctctc ggccacggcg gccgcctgc cctggaccgt	7140

gccttcaccc cgcccagagc ccaccggcca cgtgtggaga tccaccggca ggacgcctg	7200
gacatcgctg acctgccgc ggacgcgcaa ttcccaacc tctccggga ctctgtactg	7260
gctgtccgcg aggggtcccg ccggctcaca cagcaccacg ccgacgccgt acgccaagcc	7320
gatctcgctg agcgcgctat ggcggtggcg cgggtgcgct ggtgctgtct agtatcgctc	7380
cgagcccgaa gggaaaatcg agtgcccaat tcggcagaat cgggatcgat ggagttactc	7440
gacgtcgacg gggcctggtt atacaccccg gaaatcatgc gggacgaacg gggagaattc	7500
ctcgaatggt tccggggtcg gacattccag gagaagatcg gccacccct ctcgtggcc	7560
caggccaact gctcggtgtc ccgaaggcg ttctgcgagg catccacttc gccgacgccc	7620
ccgccggcc aggccaagta cgtcacctgc gcctccggca ccgtgctcga cgtggctgctc	7680
gacgtacgcc ggggctcgcc caccitcgcc cgtgggccc ccgtccgact cgacgcggcc	7740
cgccaccagg ggctctacct ggccgaagga ctggccacg cgttcatggc cctcaccgac	7800
gacgccacgg tctctacct ctgctcacag ccctacgtgg ccgaggccga gcgggcccga	7860
gacctctcg acccgcgat cggcatcgaa tggccgacgg acatcgacat cgtccctgtc	7920
ggcgaaggga caccacgca ccgtccctgg cgcaggccgc ggagaccgg catcctgccc	7980
gactacgagg gatlccccg cgccttacat cgcggaggcg gccggcgtgg gaccggcccc	8040
tgaaggccct cgtactggcc ggctctagtt acgtagtacc gcgacaaccg cgcctggtgg	8100
gaacccctga agaagcgcc ggccggctcc gcccccccc cgagaggcag cggcccatga	8160
gatggctgat caccggcgc gccggaatgc tggccggga actcgtccgg cgcctcgccg	8220
agaacgagga ggacgtcgcg gccctcgcc acgaccacct cgacgtcacc cgacctccg	8280
ccgtgcgggc ggactcgcc gagcacctc ccggatcgt cgtcaactgc gccgcctaca	8340
cggcgtcga cgacccgag acggacgagg ccgccgtgc cctcctcaac gccgaggcgc	8400
cccggctgct ggccgagggc ctgcgcccc accggcggca cggcctcgtc cacctgtcca	8460
ccgactacgt ctttccggc gacgcccga cccctacgc cgaggaccac cccacggctc	8520
ccgcagcgc ctacggacgc accaaacggg acggcgagca agcggtgctg acggcactgc	8580
ccaccgccac cgtgctgcgc accgcctggc tgtacggcg caccggccgc agcttcgtcc	8640
gcacgatgat cgaacgggag gcgcgcggcg gagccatcga cgtcgtcgcc gaccagcgcg	8700
gccagccac ctggaccgc gacctcgcc accgcatcat cgccgtcgcc cggcaccccc	8760
gcgtccacgg catcctgcac gccaccaacg ccggctccgc cacctggtac gacctggcac	8820
aagaggtctt ccggctcctc gacgccgacc ccggcggggt ccggcccacc accggcgccg	8880
ccttcgcgag acccgcccc gcggcgccct acagctcct cggccacgac cgctggcgcg	8940
ggaccggcct ggaccccctg cgtgactggc gctcgccct gcgcgaggcg ttccccgaca	9000

tcctcgccgc ggaacaccca ccgacccggc gaggagcagc atgaaacgag gcgtgcacga 9060

cctggccctc ttccctaggt ctaga 9085

<210> 5

<211> 13842

<212> DNA

<213> Artificial Sequence

<220><223> PikAI

<400> 5

atgtcttcag ccggaattac caggaccggt gcgagaacac cggtagacagg gcgtggggcg 60

gcagcgtggg acacggggga agtgcgggtc cgacgggggt tgccccctgc cggccccgat 120

catgcggagc actccttctc tcgtgctcct accggtgatg tgcgcgccga attgattcgt 180

ggagagatgt cgacagtgtc caagagttag tccgaggaat tcgtgtccgt gtcgaacgac 240

gccggttccg cgcacggcac agcggaaacc gtcgccgtcg tcggcatctc ctgccgggtg 300

cccggcgccc gggaccgag agagtctctg gaactcctgg cggcaggcgg ccaggccgtc 360

accgacgtcc ccgcggaccg ctggaacgcc ggcgacttct acgaccgga ccgtccgcc 420

cccggccgct cgaacagccg gtggggcggg ttcacgagg acgtcgaccg gttcgacgcc 480

gccttcttcg gcatctcgcc ccgcgaggcc gcggagatgg acccgagca gcggtcggcc 540

ctggagctgg gctgggagge cctggagcgc gccgggatcg acccgtctc gctcaccggc 600

accgcaccg gcgtcttcgc cggcgccatc tgggacgact acgccaccct gaagcaccgc 660

cagggcggcg ccgcgatcac ccgcacacc gtcaccggcc tccaccgagg catcatcgcg 720

aaccgactct cgtacacgt cgggtccgc ggccccagca tggctcgtga ctccggccag 780

tcctcgtcgc tcgtcgccgt ccacctcgcg tgcgagagcc tgcggcgagg cgagtcgag 840

ctcgccctcg ccggcggcgt ctcgctcaac ctggtgccgg acagcatcat cggggcgagc 900

aagttcggcg gcctctcccc cgacggccgc gcctacacct tcgacgcgcg cgccaacggc 960

tacgtacgcg gcgaggcggg cggtttcgtc gtcctgaage gcctctcccg ggccgtcgcc 1020

gacggcgacc cggtgctcgc cgtgatccgg ggcagcgccg tcaacaacgg cggcgccgcc 1080

cagggatga cgacccccga cgcgcaggcg caggaggcgg tgctccgga ggcccacgag 1140

cgggcgggga ccgcgccggc cgactgagg tacgtcagc tgcacggcac cggcaccccc 1200

gtgggcgacc cgatcgaggc cgctgcgtc ggcccgccc tcggcaccgg ccgcccggcc 1260

ggacacggc tcctggtcgg ctgggtcaag acgaacatcg gccacctgga gggcgcggcc 1320

ggcatcgccg gcctcatcaa ggccgtcctg gcggtccgcg gtcgcgcgt gcccgccagc 1380

ctgaactacg agaccccgaa cccggcgatc ccgttcgagg aactgaacct cgggtgaac	1440
acggagtacc tgccgtggga gccggagcac gacgggcagc ggatggctgt cggcgtgtcc	1500
tcgttcggca tgggcggcac gaacgcgat gtcgtgctcg aagaggcccc cgggggttgt	1560
cgaggtgctt cggtcgtgga gtcgacggtc ggccgggtcgg cggtcggcgg cggtgtggtg	1620
ccgtgggttg tgtcggcgaa gtccgctgcc gcgctggacg cgcagatcga gcggcttgcc	1680
gcgttcgcct cgcgggatcg tacggatggt gtcgacgcgg gcgctgtcga tgcgggtgct	1740
gtcgatgcgg gtgctgtcgc tcgctactg gccggcgggc gtgctcagtt cgagcaccgg	1800
gccgtcgtcg tcggcagcgg gccggacgat ctggcggcag cgctggccgc gcctgagggt	1860
ctggtccggg gcgtggcttc cgggtgcggg cgagtggcgt tcgtgttccc cgggcagggc	1920
acgcagtggg ccggcatggg tgccgaactg ctggactctt ccgcggtgtt cgcggcggcc	1980
atggccgaat gcgaggccgc actctccccg tacgtcgact ggtcgttga gcccgctgta	2040
cggcagggcc cgggtgcgcc cacgtggag cgggtcgatg tcgtgcagcc tgtgacgttc	2100
gccgtcatgg tctcgtggc tcgctgttg cagcaccacg gggtgacgcc ccaggcggtc	2160
gtcggccact cgcaggcgca gatcgccgc gcgtacgtcg ccggtgccct gaggctggac	2220
gacgccgtc gtgtcgtgac cctgcgcagc aagtccatcg cgcgccacct cgcggcaag	2280
ggcggcatgc tgtccctcgc gctgagcgag gacgccgtcc tggagcgact ggccgggttc	2340
gacgggctgt ccgtcgccgc tgtgaacggg ccaccgcca ccgtggtctc cggtgacccc	2400
gtacagatcg aagagcttgc tcgggcgtgt gaggccgatg gggtccgtgc gcgggtcatt	2460
cccgtcgact acggtccca cagccggcag gtcgagatca tcgagagcga gctcgccgag	2520
gtcctcgccg ggctcagccc gcaggctccg cgcgtgccgt tcttctcgac actcgaaggc	2580
gcctggatca ccgagccgt gctcgacggc ggctactggt accgcaacct gcgccatcgt	2640
gtgggttcg ccccgccgt cgagacctg gccaccgacg agggcttcac ccacttcgtc	2700
gaggtcagcg cccacccgt cctcaccatg gccctcccc ggaccgtcac cggcttgccg	2760
acctgcgtc gcgacaacgg cggtcaggac cgcctagtcg cctccctcgc cgaagcatgg	2820
gccaacggac tcgcggtcga ctggagcccc ctctccct ccgcgaccgg ccaccactcc	2880
gacctccca cctacgcgtt ccagaccgag cgcactggc tgggcgagat cgaggcgctc	2940
gccccggcgg gcgagccggc ggtgcagccc gccgtctcc gcacggaggc ggccgagccg	3000
gcggagctcg accgggacga gcagctgcgc gtgatcctgg acaagggtccg ggcgagacg	3060
gcccaggtgc tggggtacgc gacaggcggg cagatcgagg tcgaccggac cttccgtgag	3120
gccggttgca cctccctgac cggcgtggac ctgcgcaacc ggatcaacgc cgccttcggc	3180

gtacggatgg cgccgtccat gatcttcgac ttccccaccc ccgaggctct cgccggagcag	3240
ctgtctctcg tcgtgcacgg ggaggcggcg gcgaaccggg ccggtgcgga gccggctccg	3300
gtggcggcgg ccggtgccgt cgacgagccg gtggcgatcg tcggcatggc ctgccgcctg	3360
cccgttgggg tcgcctcgcc ggaggacctg tggcggctgg tggccggcgg cggggacgcg	3420
atctcgaggt tcccgcagga ccgcggctgg gacgtggagg ggctgtacca cccggatccg	3480
gagcaccgcc gcacgtcgta cgtccgccag ggcggtttca tcgagaacgt cgccggcttc	3540
gacgcggcct tcttcgggat ctgccgcgc gaggccctcg ccatggacc gcagcagcgg	3600
ctctctctcg aaacctctg ggaggccgtc gaggacccg ggatcgacc gacctcctg	3660
cggggacggc aggtcggcgt cttcactggg gcgatgacc acgagtacgg gccgagcctg	3720
cgggacggcg gggaaggcct cgacggctac ctgctgacc gcaacacggc cagcgtgatg	3780
tcgggccgcg tctcgtaac actcggcctt gaggggcccc ccctgacgtt ggacacggcc	3840
tgctctctgt cgctggctgc cctgcacct gccgtgcagg ccctgcgca gggcgaggtc	3900
gacatggcgc tcgccggcgg cgtggccgtg atgccacgc ccgggatgtt cgtcgagttc	3960
agccggcagc ggggcttggc cggggacggc cggtcgaagg cgttcgccgc gtcggcggac	4020
ggcaccagct ggtccgaggc cgtcggcgtc ctctctctcg agcgctgtc ggacgccgc	4080
cgcaacggac accaggtcct cgcggtcgtc cgcggcagcg ccttgaacca ggacggcgcg	4140
agcaacggcc tcacggctcc gaacggggccc tcgcagcagc gcgtcatccg gcgcgcgtg	4200
gcggacggcc ggctgacgac ctccgacgtg gacgtcgtcg aggcacacgg cacgggcacg	4260
cgactcggcg acccgatcga ggcgaggcc ctgatcgcca cctacggcca gggccgtgac	4320
gacgaacagc cgctgcgcct cgggtcgttg aagtccaaca tcgggcacac ccaggccgcg	4380
gccggcgtct ccggtgtcat caagatggtc caggcgatgc gccacggact gctgccgaag	4440
acgtgcacg tcgacgagcc ctccgaccag atcgactggt cggtcggcgc cgtggaactc	4500
ctcaccgagg ccgtcgactg gccggagaag caggacggcg ggctgcgccg gccgccgtc	4560
tcctccttcg ggatcagcgg caccaatgcg catgtggtgc tcgaagagc cccggtggtt	4620
gtcgagggtg cttcggctgt cgagccgtcg gttggcgggt cggcggctcg cggcgggtgtg	4680
acgccttggg tgggtgcggc gaagtccgt gccgcgtcg acgcgcagat cgagcggctt	4740
gccgcatcgc ctcgcggga tcgtacggat gacgccagc ccggtgctgt cgacgcgggc	4800
gctgtcgtc acgtactggc tgacgggcgt gctcagttcg agcaccgggc cgtcgcgtc	4860
ggcgccgggg cggacgacct cgtacaggcg ctggccgatc cggacgggct gatacgcgga	4920

acggcttccg gtgtcgggcg agtggcggtt gtgttccccg gtcagggcac gcagtgggct	4980
ggcatgggtg ccgaactgct ggactcttcc gcggtgttcg cggcggccat ggccagtggt	5040
gaggccgcgc tgtccccgta cgtcgactgg tcgctggagg ccgtcgtacg gcaggccccc	5100
ggtgcgcccc cgctggagcg ggtcgatgtc gtgcagcctg tgacgttcgc cgtcatggtc	5160
tcgctggctc gcgtgtggca gcaccacggt gtgacgcccc aggcggctgt cggccactcg	5220
caggcgagga tcgcccgccg gtacgtcgcc ggagccctgc ccctggacga cggcccccgc	5280
gtcgtcaccg tgcgcagcaa gtccatcgcc gccacactcg ccggcaaggg cggcatgctg	5340
tccttcgcgc tgaacgagga cgccgtcttg gagcgactga gtgacttcga cgggctgtcc	5400
gtcgccgcgc tcaacgggcc caccgccact gtctgttcgg gtgacccctg acagatcgaa	5460
gagcttgctc aggcgtgcaa ggcggacgga ttccgcgcgc ggatcattcc cgtcgactac	5520
gcgtcccaca gccggcaggt cgagatcatc gagagcgagc tcgccaggt cctcgccggt	5580
ctcagcccgcc aggccccgcg cgtgccgttc ttctcgacgc tcgaaggcac ctggatcacc	5640
gagcccgctc tcgacggcac ctactggtac cgcaacctcc gtcaccgct cggcttcgcc	5700
cccgccatcg agaccctggc cgtcgacgag ggcttcacgc acttcgtcga ggtcagcgcc	5760
caccccgctc tcacatgac cctccccgag accgtcaccg gcctcggcac cctccgtcgc	5820
gaacagggag gccaaagagc tctggtcacc tcgctcgccg aggcgtgggt caacgggctt	5880
cccgtggcat ggacttcgtc cctgccccgc acggcctccc gcccggtct gccacctac	5940
gccttcagg ccgagcgcta ctggctcgag aacactccc cgcacctggc caccggcgac	6000
gactggcgct accgcatcga ctggaagcgc ctcccggccg ccgaggggtc cgagcgaccc	6060
ggcctgtccg gccgtggct cgccgtcacg ccggaggacc actccgcgca ggccgcccgc	6120
gtgctcaccg cgtggtcga cgccggggcg aaggtcgagg tgctgacggc cggggcggac	6180
gacgaccgtg aggccctcgc cgccggctc accgactga cgaccggtga cggcttcacc	6240
ggcgtggctc cgtcctcga cggactcgtc ccgcaggtcg cctgggtcca ggcgtcggc	6300
gacgccgga tcaaggcgcc cctgtggtcc gtcaccagg gcgcggtctc cgtcgacgt	6360
ctcgacaccc ccgccaccc cgaccgggccc atgctctggg gcctcggccg cgtcgtcgcc	6420
cttgagcacc ccgaacgtg ggccggcctc gtcgacctcc ccgccagcc cgatgccgcc	6480
gccctcgccc acctcgtcac cgcactctcc ggccgccacc gcgaggacca gatcgccatc	6540
cgcaccaccg gactccacgc ccgccgctc gcccgcgac ccctccacgg acgtcggccc	6600
accccgact ggcagcccca cggcaccgtc ctcatcaccg gcggcaccgg agccctcggc	6660
agccacgcc cagctggat ggcccaccac ggagccgaac acctcctct cgtcagccgc	6720
agcggcgaac aagccccgg agccacccaa ctaccgccg aactcaccgc atcgggcgcc	6780

cgcgtcacca tcgccgcctg cgacgtcgcc gacccccacg ccatgcgcac cctcctcgac	6840
gccatccccg ccgagacgcc cctcaccgcc gtcgtccaca ccgccggcgc gctcgacgac	6900
ggcatcgtgg acacgtgac cgccgagcag gtcggcgggg cccaccgtgc gaaggccgtc	6960
ggcgccctcg tgctcgacga gctgaccggg gacctcgacc tcgacgcgtt cgtgctcttc	7020
tcgtccgtgt cgagcactct gggcatcccc ggtcagggca actacgcccc gcacaacgcc	7080
tacctcgacg cctcgcggc tcgccgcggg gccaccggcc ggtccgccgt ctcggtggcc	7140
tggggaccgt gggacggtgg cggcatggcc gccggtgacg gcgtggccga gcggctgcgc	7200
aaccacggcg tgcccgcat ggaccgggaa ctgcacctgg ccgactgga gtcgcgcctc	7260
ggccgggacg agaccgcgat caccgtcgcg gacatcgact gggaccgctt ctacctcgcg	7320
tactcctccg gtcgcccga gccctcgtc gaggagctgc ccgaggtgcg gcgcatcatc	7380
gacgcacggg acagcgccac gtccggacag ggccgggagct ccgccaggg cgccaacccc	7440
ctggccgagc ggctggccgc cgcggtccc ggcgagcgta cggagatcct cctcggtctc	7500
gtacgggcgc aggccgccgc cgtgctccgg atgcgttcgc cggaggacgt cgccgccgac	7560
cgcgcttca aggacatcgg ctctgactcg ctgccggtg tcgagctgcg caacaggctg	7620
acccgggca ccgggtcca gctgccgcg acgtctgtct tcgaccaccc gacgccgtg	7680
gccctcgtgt cgtgctccg cagcgagttc ctcggtgacg aggagacggc ggacgcccg	7740
cggctccggg cgtgccgcg gactgtcgtt gccggtgccg gcgccggcgc cggcacgat	7800
gccgacgacg atccgatcg gatcgtcgcg atgagctgcc gctaccccg tgacatccgc	7860
agcccgagg acctgtggcg gatgtgtcc gaggcgggcg agggcatcac gccgttcccc	7920
accgaccgcg gctgggacct cgacggcctg tacgacccg acccgacgc gctcggcagg	7980
gcgtacgtcc gcgaggcg gttcttcac gacgcggccg agttcgacgc ggagtcttc	8040
ggcgtctgc cgcgcgaggc gctggccatg gaccgcagc agcggtgct cctgacgacg	8100
tcctgggagg ccttcgagcg ggccggcatc gagccggcat cgctgcgcgg cagcagcacc	8160
ggtgtcttca tcggcctctc ctaccaggac tacgcggccc gcgtccgaa cgcgccggt	8220
ggcgtggagg gttacctgt gaccggcagc acgccgagcg tcgcgtcggg ccgtatcgcg	8280
tacaccttcg gtctgaagg gcccgcgacg accgtcgaca ccgctgctc gtcgtcgtg	8340
accgacctgc acctggcggt gcgggcgctg cgacggcg agtgacgat ggcgtcgc	8400
ggtggcgtgg cgatgatggc gacccgcac atgttcgtgg agttcagccg tcagcggcg	8460
ctgccccgg acggccgacg caaggccttc tcggcgacg ccgacgggtt cggcgccgcg	8520
gaggcgctgc gctgctgct cgtggagcgg ctctcgacg cgccggcgaa cggtcacccg	8580
gtgctcgccg tggtcggcg tactccgtc aaccaggacg gcgccagaa cgggctgacc	8640

gcgccaacg gaccctcgca gcagcgggtg atccggcagg cgctcgccga cccccggtg	8700
gcacccggcg acatcgacgc cgtcgagacg cacggcacgg gaacctcgct gggcgacccc	8760
atcgaggccc agggcctcca ggccacgtac ggcaaggagc ggcccgcgga acggccgctc	8820
gccatcggct ccgtgaagtc caacatcgga cacaccagg ccgcggccgg tgcggcgggc	8880
atcatcaaga tggctctcgc gatgcgccac ggacacctgc cgaagacct ccacgccac	8940
gagccgagcc cgcacgtcga ctgggcgaac agcggcctgg ccctcgtcac cgagccgatc	9000
gactggccgg ccggcaccgg tccgcgccgc gccgccgtct cctccttcgg catcagcggg	9060
acgaacgcgc acgtcgtgct ggagcaggcg ccggatgctg ctggtgaggt gcttggggcc	9120
gatgagggtc ctgagggtc tgagacggta gcgatggctg ggacggctgg gacctccgag	9180
gtcgtgagg gctctgagg ctcgaggcc ccgcggccc ccggcagccg tgaggcgtcc	9240
ctccccgggc acctgcctg ggtgctgtcc gccaaaggac agcagtcgt gcgcggccag	9300
gccgcgccc tgcacgcgtg gctgtccgag ccgcgcccg acctgtcgga cgcggacgga	9360
ccggccccgc tgcgggacgt cgggtacacg ctgccacga gccgtaccgc cttcgcgcac	9420
cgcgccggcg tgaccgcgc gcaccgggac gggttcctgg acgggctggc cacgtggcc	9480
cagggcggca cctcgccca cgtccacctg gacaccgcc gggacggcac caccgcgttc	9540
ctcttcaccg gccagggcag tcagcgcccc ggcgccggcc gtgagctgta cgaccggcac	9600
cccgtcttcg cccgggcgct cgacgagatc tgcgccacc tcgacggtca cctcgaactg	9660
cccctgctcg acgtgatgtt cgcggccgag ggacgcgcgg aggccgcgct gctcgacgag	9720
acgcggtaca cgcagtgcgc gctgttcgcc ctggaggtcg cgctcttcg gctcgtcgag	9780
agctggggca tgcggccggc cgactgtct ggtaactcgg tcggcgagat cgccgccgcg	9840
cacgtcgccg gtgtgttct ctcgccgac gccgccgcc tggtcgccgc gcgcggccgg	9900
ctcatgcagg agctgccgc cggtaggcgc atgctgccg tccaggccgc ggaggacgag	9960
atccgcgtgt ggctggagac ggaggagcgg tacgcgggac gtctggacgt cgccgccgtc	10020
aacggcccc aggcccgct cctgtccgc gacgcggacg cggcgcgga ggcggaggcg	10080
tactggtccg ggctcgccg caggacccgc gcgtgcggg tcagccacgc cttccactcc	10140
gcgcacatgg acggcatgct cgacgggttc cgcgccgtcc tggagacggt ggagtcccg	10200
cgcctctccc tgacctggt ctcgaacgtc accggcctgg ccgccggccc ggacgacctg	10260
tgcgacccc agtactgggt ccggcacgtc cgcggcaccg tccgttctc cgacggcgtc	10320
cgtgtcctgc gcgacctcg cgtgcggacc tgcttgagc tgggccccga cggggtcctc	10380

accgcatgg cgcccgacgg cctcgcggac acccccgcgg attccgctgc cggctcccc	10440
gtcggctctc ccgccggctc tcccgcggac tccgccgccc gcgcgctccg gccccggccg	10500
ctgctcgtgg cgtgctgctg ccgcaagcgg tcggagaccg agaccgtcgc ggacgccctc	10560
ggcagggcgc acgcccacgg caccggaccc gactggcacg cctggttcgc cggctccggg	10620
gcgcaccgcg tggacctgcc cacgtactcc ttccggcgcg accgctactg gctggacgcc	10680
ccggcggccg acaccgcggt ggacaccgcc ggccctcggc tcggcaccgc cgaccaccg	10740
ctgctcggcg ccgtggtcag ccttcggac cgggacggcc tgcctgtcac cggccgcctc	10800
tccttcgca cccaccgtg gctcgcggac cacgccgtcc tggggagcgt cctgtcccc	10860
ggcgccgca tggctgaact cgccgcgcac gctgcggagt ccgccggtct gcgtgacgtg	10920
cgggagctga cctccttga accgctggta ctccccgagc acggtggcgt cgagctgcgc	10980
gtgacggtcg ggccgccggc cggagagccc ggtggcgagt cggccgggga cggcgcacgg	11040
cccgtctccc tccactcgcg gctcggcgc gcgcccgcg gtaccgcctg gtcctgccac	11100
gcgaccggtc tgcctggccac cgaccggccc gagcttccc tcgcgccga ccgtgcggcc	11160
atgtggccgc cgcagggcgc cgaggaggtg ccgctcgacg gtctctacga gcggctcgac	11220
gggaacggcc tcgccttcgg tccgtgttc caggggctga acgcggtgtg gcggtacgag	11280
ggtgaggtct tcgccacat cgcgtcccc gccaccacga atgcgaccgc gcccgcgacc	11340
gcgaacggcg ggggagtg gcggcgggcc cctacggca tccaccgcg cctgtctgac	11400
gcttcgtcgc acccatcgc ggtcggcggc ctgctcgacg agcccagct cgtccgcgtc	11460
cccttccact ggagcgggtg caccgtgcac gcggccggtg ccgcggcggc ccgggtccgt	11520
ctgcctccg cggggacgga cgccgtctcg ctgtccctga cggacggcga gggacgccg	11580
ctggtctccg tggaacggct cacgtcgcg ccggtcaccg ccgatcaggc ggcggcgagc	11640
cgcgtcggcg ggetgatgca ccgggtggcc tggcgtccgt acgccctcgc ctgctccgc	11700
gaacaggacc cgcacgccac ttctgacggg ccgaccgccg tcctcggcaa ggacgagctg	11760
aaggtcgcg ccgccctgga gtccgcgggc gtggaagtgc ggtctaccc cgacctggcc	11820
gcgtgtccc aggacgtggc ggccggcgcc ccggcgcccc gtaccgtcct tgcgcgctg	11880
cccgcgggtc ccgccacgg cggcgcgag ggtgtacggg gcacggtggc ccggacgctg	11940
gagctgtcc aggcctggct ggccgacgag cacctcgcgg gcaccgcct gtcctggtc	12000
acccgcggtg cggctcggga ccccagggg tcggcgccc acgatggcg cgaggacctg	12060
tcgcacgcgg ccgcctgggg tctcgtacgg accgcgcaga ccgagaacct cggccgttc	12120
ggccttctcg acctggccga cgacgcctcg tcgtaccgga cctgcgctc ggtgctctcc	12180
gacgcgggcc tgcgcgacga accgcagctc gcctgcacg acggcaccat caggctggcc	12240

cgcttggcct ccgtccggcc cgagaccggc accgccgcac cggcgctcgc cccggagggc 12300
acggtcctgc tgaccggcgg caccggcggc ctgggcggac tggtcgccc gcacgtggtg 12360
ggcgagtggg gcgtacgacg cctgtctgtg gtgagccggc ggggcacgga cccccgggc 12420
gccgacgagc tcgtgcacga gctggaggcc ctgggagccg acgtctcggt ggccgcgtgc 12480

gacgtcgccg accgcaagc ctcaccgcc gtactcgacg ccatccccgc cgaacaccg 12540
ctcaccggg tcgtccacac ggcaggcgtc ctctccgacg gcaccctccc gtccatgacg 12600
acggaggacg tggaaacgt actgcggccc aaggctgacg ccgcgttctt cctcgacgaa 12660
ctcacctega cggccgata cgacctggca gcgttcgtca tgttctctc cggcccgcc 12720
gtcttcggtg gcgcggggca ggggccttac gccgcccca accccacct cgacgcctc 12780
gcctggcgcc gccgggcagc cggactcccc gccctctccc tcggctgggg cctctgggcc 12840
gagaccagcg gcatgaccgg cgagctcggc caggcggacc tgcgccgat gagccgcgcg 12900

ggcatcggcg ggatcagcga cgccgagggc atcgcgctcc tcgacgccgc cctccgcgac 12960
gaccgccacc cggctctgtt gcccctgcgg ctcgacgccg ccgggctgcg ggacgcggcc 13020
gggaacgacc cggccggaat cccggcgtc ttccgggacg tcgtcggcgc caggaccgtc 13080
cgggccccgc cgtccgcggc ctccgcctcg acgacagccg ggacggccgg cacgccgggg 13140
acggcggacg gcgcggcgga aacggcggcg gtcacgtcgc ccgaccgggc cgccaccgtg 13200
gacgggcccc caccgcagcg cctgtctgtc gattctgtc tcggcgaggt cgccgaagta 13260
ctcgccacg ccccggttca ccggtatgac gccgaacggg gcttctctga cctcggttc 13320

gactccctga ccgccgtga actccgcaac cggctcaact ccgccgtgg cctcgccctc 13380
ccggcgacc tggtcttctga ccaccaagc ccggcgccac tcgcctccca cctggacgcc 13440
gagctgccg gcggcgctc ggaccaggac ggagccggga accggaacgg gaacgagaac 13500
gggacgacgg cgtcccgag caccgccgag acggacgcgc tgcggcaca actgaccgc 13560
ctggaaggcg ccttggtgtg gacggcctc tcggacgcc ccgggagcga agaagtcctg 13620
gagcacctgc ggtccctgcg ctcgatggtc acggcgaga ccgggaccgg gaccgcgtc 13680
ggagccccgg acggcgccgg gtccggcgcc gaggaccggc cctgggcggc gggggacgga 13740

gccgggggcg ggagttagga cggcgcgga gtcccgact tcatgaacgc ctcgccgag 13800
gaactcttcg gctctctga ccaggacccc agcacggact ga 13842

<210> 6

<211> 11220

<212> DNA

<213> Artificial Sequence

<220><223> PikAII

<400> 6

gtgtccacgg tgaacgaaga gaagtacctc gactacctgc gtcgtgccac ggcggacctc	60
cacgaggccc gtggccgcct ccgcgagctg gaggcgaagg cgggcgagcc ggtggcgatc	120
gtcggcatgg cctgccgcct gcccggcggc gtcgcctcgc ccgaggacct gtggcggctg	180
gtggccggcg gcgaggacgc gatctcggag ttccccagg accgcggctg ggacgtggag	240
ggcctgtacg acccgaaccc ggaggccacg ggcaagagtt acgcccgcga ggccggattc	300
ctgtacgagg cgggcgagtt cgacgccgac ttcttcggga tctcgcgcg cgaggccctc	360
gccatggacc cgcagcagcg tctcctcctg gaggcctcct gggaggcggt cgagcacgcc	420
gggatcccgg cggccaccgc gcgcggcacc tcggtcggcg tcttcaccgg cgtgatgtac	480
cacgactacg ccacccgtct caccgatgtc ccggagggca tcgagggcta cctgggcacc	540
ggcaactccg gcagtgtcgc ctccggccgc gtcgcgtaca cgcttggcct ggaggggccg	600
gccgtcacgg tcgacaccgc ctgctcgtcc tcgctggtcg ccctgcacct cgccgtgcag	660
gccctgcga agggcgaggt cgacatggcg ctccggcg gcgtgacgt catgtcgacg	720
cccagacct tcgtcgagtt cagccgtcag cgcgggctgg cgcggacgg ccggtcgaag	780
tccttctcgt cgacggccga cggcaccagc tggtcgagg gcgtcggcgt cctcctcgtc	840
gagcgctgt ccgacgcgcg tcgcaagggc catcggtacc tcgccgtggt ccggggcacc	900
gccgtcaacc aggacggcgc cagcagcggc ctacaggctc cgaacgggcc gtcgcagcag	960
cgcgtcatcc gacgtgccct ggcgagcgc cggctcacga cctccgacgt ggacgtcgtc	1020
gaggcccacg gcacgggtac gcgactcggc gacccgatcg aggcgcaggc cgtcatcgcc	1080
acgtacgggc agggccgtga cggcgaacag ccgctgcgcc tcgggtcgtt gaagtccaac	1140
atcgacaca ccaggccgc cgcgggtgtc tcggcgta tcaagatggt ccaggcgatg	1200
cgccacggcg tcctgccgaa gacgtccac gtggagaagc cgacggacca ggtggactgg	1260
tcgcggggcg cggtcgagct gctcaccgag gccatggact ggccggacaa gggcgacggc	1320
ggactgcga gggccgcggt ctctccttc ggcgtcagcg ggacgaacgc gcacgtcgtg	1380
ctcgaagagg ccccgccggc cgaggagacc cctgcctccg aggcgacccc ggccgtcgag	1440
ccgtcggtcg gcgccgcct ggtgccgtgg ctggtgtcgg cgaagactcc ggccgcgctg	1500
gacgcccaga tcggacgcct cgcccgcttc gcctcgcagg gccgtacgga cgcgccgat	1560
ccggcgccgg tcgtcgcgt actggccggc gggcgcccg agttcgagca ccgggccgtc	1620
gtgtcggca ccggacagga cgatttcgcg caggcgctga ccgtccgga aggactgata	1680
cgcggcacgc cctcggacgt gggccgggtg gcgttcgtgt tccccggtca gggcacgcag	1740

tgggcccggga tgggcgccga actcctcgac gtgtcgaagg agttcgcggc ggccatggcc	1800
gagtgcgaga gcgcgtcttc cegctatgtc gactggctgc tggaggccgt cgtccggcag	1860
gcgccgggcg cggccacgct ggagcgggtc gacgtcgtcc agcccgtgac cttcgtgttc	1920
atggtttcgc tggcgaaggt ctggcagcac cacggcgtga cggcgaggc cgtcgtcggc	1980
cactcgcagg gcgagatcgc cgccgcgtac gtgcgggtg ccctcacct cgacgacgcc	2040
gcccgcgtcg tcacctgcg cagcaagtcc atcgccccc acctcgccgg caaggcggc	2100
atgatctccc tcgccctcag cgaggaagcc acccggcagc gcatcgagaa cctccacgga	2160
ctgtcgatcg ccgcgtcaa cggcccccacc gccaccgtgg ttctgggcga cccacccag	2220
atccaagagc tcgtcaggc gtgtgaggcc gacggggtcc gcgcacggat catccccgtc	2280
gactacgctt cccacagcg ccacgtcgag accatcgaga gcgaactcgc cgaggctctc	2340
gccgggctca gcccgcggac acctgagggt cgtttcttct cgacactcga aggcgcctgg	2400
atcacggagc cgggtgctga cggcacctac tggtaaccga acctccgcca ccgcgtcggc	2460
ttcgcccccg ccgtcgagac cctcgccacc gacgaaggct tcacccactt catcgaggtc	2520
agcgcaccac ccgtctcac catgacctc ccgagaccg tcaccggcct cggcacctc	2580
cgcccggaac agggaggcca ggagcgtctg gtacactcac tcgccgaagc ctggaccaac	2640
ggcctacca tcgactgggc gccgtcttc cccaccgcaa ccggccacca ccccgagtc	2700
cccactacg ctttcagcg ccgtcactac tggctccag actccccgc cgtccagggc	2760
tccgtgcagg actcctggcg ctaccgcac gactggaagc gcctcgcgtt cggcgacgcg	2820
tccgagcgcg ccgggtgttc cgggcgttgg ctgtctgtcg tccccgagga ccgttccgcc	2880
gaggccgccc cgggtgtcgc cgcgctgtcc ggcgccggcg ccgacccgt acagctggac	2940
gtgtccccgc tgggcgaccg gcagcggctc gccgcgacgc tgggcgaggc cctggcggcg	3000
gccggtggag ccgtcgacgg cgtctctctg ctgtctcgtt gggacgagag cgcgcacccc	3060
ggccaccccg ccccttcac ccggggcacc ggccgaccc tcacctggt gcaggcgtg	3120
gaggacgccc gcgtcgccg cccgtgttgg tgcgtgaccc acggcgcggt gtccgtcggc	3180
cgggcgcacc acgtcacctc cccgcccag gccatggtgt ggggcatggg ccgggtcgcc	3240
gccctggagc acccgagcg gtggggcggc ctgatcgacc tgcctcgga cggcgaccg	3300
gcggcccttg accgcatgac caggctctc gccggcggtt cgggtgagga ccaggctcgc	3360
gtacgcctt ccgggtgtgt cggcccgcc ctctcccgcg cctccctccc ggcgcacggc	3420
acggcttcgc cgtggtggca ggccgacggc acggtgtctg tcaccggtgc cgaggagcct	3480
gcggccgccc aggcgcacg ccggctggcc cgcgacggcg ccggacacct cctcctccac	3540

accaccccct ccggcagcga aggcgccgaa ggcacctccg gtgccgccga ggactccggc	3600
ctcgccgggc tctcgccga actcgccgac ctgggcgcga cggccaccgt cgtgacctgc	3660
gacctcacgg acgcggaggc ggccgcccg ctgctcgccg gcgtctccga cgcgcacccg	3720
ctcagcgccg tctccacct gccgccacc gtcgactccg agccgctcgc cgcgaccgac	3780
gcggacgcgc tgcctgtgt cgtgaccgcg aaggccaccg ccgcgtcca cctggaccgc	3840
ctcctcgggg aggcgcggc tgccggaggc cgtccgccg tcttggtcct cttctcctcg	3900
gtcgccgcga tctggggcgg cgccggtcag ggcgcgtacg ccgccgttac ggccttctc	3960
gacgccctcg ccggtcagca ccgggccgac ggccccaccg tgacctcggg ggcctggagc	4020
ccctgggagg gcagccgcgt caccgagggt gcgaccgggg agcggctgcg ccgcctcggc	4080
ctgcgcccc tgcccccgc gacggcgctc accgccctgg acaccgcgt cggccacggc	4140
gacaccgccg tcacgatcgc cgacgtcgac tggtcgagct tgcccccgg cttcaccacg	4200
gccccggcgg gcacctcct cgccgatctg cccgaggcgc gccgcgcgt cgacgagcag	4260
cagtcgacga cggccgccga cgacaccgtc ctgagccgcg agctcggtgc gtcaccggc	4320
gccgaacagc agcgccgtat gcaggagtgt gtccgcgagc acctcgccgt ggtcctcaac	4380
cacctctcc ccgaggcgt cgacacgggg cgggccttcc gtgacctcgg attcgactcg	4440
ctgacggcgg tcgagctccg caaccgcctc aagaaccca ccggcctggc cctcccggcc	4500
actctggtct tcgactacc gacccccgg acgtggcgg agttcctcct cgcggagatc	4560
ctgggcgagc aggcgcgtgc ggccgagcag cttccggtgg acggcggggt cgacgacgag	4620
cccgtcgca tctcgcat ggcgtgccgc ctgccggcg gtgtcgctc gccggaggac	4680
ctgtggcggc tgggtggccg cgccgaggac gcgatctccg gcttcccgca ggaccgcggc	4740
tgggacgtgg aggggtgta gcaccggac ccggacgcgt ccgggcggac gtactgccgt	4800
gccggtggct tctcgacga ggccggcgag ttcgacgccg acttcttcgg gatctcgccg	4860
cgcgaggccc tcgcatgga ccgcagcag cggtcctcc tggagacctc ctgggaggcc	4920
gtcgaggacg ccgggatcga ccgacctcc cttcaggggc agcaggtcgg cgtgttcgcg	4980
ggcaccaacg gccccacta cgagccgtg ctccgcaaca ccgccagga tcttgagggt	5040
tacgtcggga cgggcaacgc cgccagcatc atgtcgggc gtgtctcgta caccctcggc	5100
ctggagggcc cggccgtcac ggtcgacacc gcctgtcct cctcgctggt cgccctgcac	5160
ctcgccgtgc agccctcgc caaggcgaa tgcggactgg cgctcgccgg cgggtgtgacg	5220
gtcatgtcga cccccagac gtctgtggag ttcagccggc agcgcgggct cgcggaggac	5280

ggccggtcga aggcgttcgc cgcgtcggcg gacggcttcg gcccgggcga gggcgtcggc	5340
atgtcctcgc tcgagcgctt gtcggacgcc cgcgcgaacg gacaccgtgt gctggcggtc	5400
gtgcgcggca gcgcggtcaa ccaggacggc gcgagcaacg gcctgaccgc cccgaacggg	5460
ccctcgcagc agcgcgtcat ccggcgcgcg ctgcgggacg cccgactgac gaccgcccac	5520
gtggacgtcg tcgaggccca cggcacgggc acgcgactcg gcgaccgat cgaggcacag	5580
gccctcatcg ccacctacgg ccaggggcgc gacaccgaac agccgttcgc cctggggtcg	5640
ttgaagtcca acatcgga caaccaggcc gccgcgggtg tctccggcat catcaagatg	5700
gtccaggcga tgcgccacgg cgtcttcgag aagacgtcc acgtggaccg gccgtcggac	5760
cagatcgact ggtcggcggg cacggtcgag ctgctcaccg aggccatgga ctggccgagg	5820
aagcaggagg gcgggctgag ccgcgcggcc gtctctcct tcggcatcag cggcacgaac	5880
gcgcacatcg tgcctgaaga agccccggtc gacgaggacg ccccgggcga cgagccgtcg	5940
gtcggcgggtg tgggtccgtg gctcgtgtcc gcgaagactc cggccgcgct ggacgcccag	6000
atcggacgcc tcgccgctt cgcctcgcag ggccgtacgg acgcccgcga tccgggcgcg	6060
gtcgtcgcg tactggccgg cgggcgtgag cagttcgagc accgggcccgt cgcgtcggc	6120
accggacagg acgacctggc ggccgcactg gccgcgcctg agggctctgt cgggggtgtg	6180
gcctccggtg tgggtcgagt ggcgttcgtg ttcccgggac agggcacgca gtgggcccgg	6240
atgggtgccg aactcctcga cgtgtcgaag gatttcgagg cggccatggc cgagtgcgag	6300
gccgcgctcg ctccgtacct ggactggctg ctggaggccg tcgtccgaca ggccccggc	6360
gcgcccacgc tggagcgggt cgatgtctc cagcccgtga cgttcgccgt catggtctcg	6420
ctggcgaagg tctggcagca ccacggggtg accccgcaag ccgtcgtcgg cactcgcag	6480
ggcgagatcg ccgcccgta cgtcggcgtt gccctgagcc tggacgacgc cgtcgtgtc	6540
gtgacctgc gcagcaagtc catcggcgcc cactcgcgg gccaggcgg catgctgtcc	6600
ctcgcgtga gcgaggcggc cgttgtggag cgactggccg ggttcgacgg gctgtccgtc	6660
gccgcgctca acgggcctac cgccaccgtg gtttcgggag accgaccca gatcaagag	6720
ctcgtcagg cgtgtgaggc cgacggggtc cgcgcacgga tcatccccgt cgactacgcc	6780
tcccacagcg ccacgtcga gaccatcgag agcgaactcg ccgacgtcct ggccgggttg	6840
tccccccaga caccacgtt cccctcttc tccacctcg aaggcgcctg gatcaccgaa	6900
cccgcctcg acggcggtta ctggtaccgc aacctcgcc atcgtgtggg ctccgccccg	6960
gccgtcga aa cctggccac cgacgaaggc ttaccact tcgtcgaggt cagcggccac	7020
cccgtctca ccatggcgtt ccccgagacc gtcaccggac tcggcaccct ccgccgtgac	7080
aacggcggac agcaccgctt caccacctc ctgcgcagg cctgggcaa cggcctcacc	7140

gtcgactggg cctctctcct ccccaccacg accaccacc ccatctgcc cacctacgcc	7200
ttccagaccg agcgctactg gccgcagccc gacctctccg ccgccggtga catcacctcc	7260
gccggtctcg gggcgccga gcaccgctg ctccggcgg ccgtggcgct cgcggactcc	7320
gacggctgcc tgcacacggg gagcctctcc ctccgtacgc acccctggct ggccgaccac	7380
gcggtggccg gcaccgtgct gctgccggga acggcgctcg tggagctggc gttccgagcc	7440
ggggaccagg tcggttgca tctggtcgag gagctaccc tcgacgcgcc gctcgtgctg	7500
ccccgtcgtg gcgcgggtccg tgtgcagctg tccgtcggcg cgagcgacga gtccggcgct	7560
cgtaccttcg ggctctacgc gcaccggag gacgcgcgg gcgaggcgga tgggacgcgg	7620
cacgccaccg gtgtgctggc cggcgtgctg gaccgcaccg ccccgctcgc gcaccggag	7680
gcctggccgc cggcgccgc cgagccggtg gacgtggacg gtctgtacga gcgtctcgcg	7740
gcgaacggct acggctacgg cccctcttc cagggcgtcc gtggtgtctg gcggcgtggc	7800
gacgaggtgt tcgccagct ggccctgccg gccgaggtcg ccggtgccga gggcgcgcg	7860
ttcgcccttc acccgcgct gctcgacgcc gccgtgcagg ccggcggtgc gggccggggc	7920
gttcggcgcg ggcacgcgcg tgcgttcgc ctggagcggg atctcctgta cgcggtcggc	7980
gccaccgcc tcccgctgcg gctggcccc gccggcccgg acacggtgtc cgtgagcgcc	8040
gccgactcct ccgggcagcc ggtgttcgcc gcggactccc tcacggtgct gccgctcgac	8100
cccgcgacg tggcgccct cagcgaccgc actctggacg cgtgcacct gctggagtgg	8160
accgcctggg acggtgccgc gcagccctg cccggcgcg tctgtctggg cggcgacgcc	8220
gacggtctcg ccggcgctg gcgcgccgt ggaccgagg tctgtcctt cccggacctt	8280
acggacctgg tggaggcgt cgaccggggc gagaccccg ccccggcgac cgtcctggtg	8340
gcctgcccc ccggcgccc cgatgggccc gagcatgtcc gcgaggccct gcacgggtcg	8400
ctcgcgctga tgcaggcctg gctggccgac gagcggttca ccgatggcg cctggtgctc	8460
gtgaccgcg acgcggtcgc cggcggttc ggcgacggc tgcggtccac gggacaggcc	8520
gccgtctggg gcctcgccg gtccgcgcag acggagagcc cggcgcggtt cgtcctgctc	8580
gacctcgccg gggaagccc gacggccggg gacgccacc cggggacgg cctgacgacc	8640
ggggacgcca ccgtcgccg cacctctgga gacccgccc tcggcagcgc cctcgcgacc	8700
gcccctggct cggcgagcc gcagctcgcc ctccgggacg gggcgctcct cgtacccgc	8760
ctggcgcggg ccggcgccc cggcgccg gacggcctcg ccggcgccga cggcctcgcc	8820
gctctgccgc tggcgccgc tccggccctc tggcgtctgg agcccggtac ggacggcagc	8880
ctggagagcc tcacggcgcc gccggcgac gccgagacc tcgccccga gccgctcgcc	8940
ccgggacagg tccgcatcg gatccgggc accggtctca acttcgcga cgtcctgac	9000

gccctcggca tgiacccga tccggcgctg atgggcaccg agggagccgg cgtggtcacc	9060
gcgaccggcc ccggcgctac gcacctcgcc cccggcgacc gggcatggg cctgctctcc	9120
ggcgcgtagc ccccggtcgt cgtggcggac ggcgggaccg tcgcgcggat gcccgagggg	9180
tggacgttcg ccagggcgcc ctccgtgccg gtggtgttcc tgacggccgt ctacgcctg	9240
cgcgacctgg cggacgtcaa gcccggcgag cgcctcctgg tccactccgc cgcgggtggc	9300
gtgggcatgg ccgccgtgca gctcggccgg cactggggcg tggaggtcca cggcacggcg	9360
agtcacggga agtgggacgc cctgcgcgcg ctggcctgg acgacgcga catgcctcc	9420
tccgcaccc tggacttga gtccgcgttc cgtgccgtt ccggcggggc gggcatggac	9480
gtcgtactga actcgctgc ccgcgagttc gtcgacgcct cgctgcgcct gctcgggccg	9540
ggcgcccggt tcgtggagat ggggaagacc gacgtccgcg acgcggagcg ggtcgccgcc	9600
gaccaccccg gtgtcgcta ccgcgccttc gacctgggcg aggcggggcc ggagcggatc	9660
ggcgagatgc tcgccaggt catgccctc ttcgaggacg ggggtctccg gcacctgccc	9720
gtcacacct gggacgtgcg ccgggcccgc gacgccttc ggcacgtcag ccaggcccgc	9780
cacacgggca aggtcgtcct cacgatgccg tcgggcctcg acccgagggg tacggtcctg	9840
ctgaccggcg gcaccggtcg gctggggggc atcgtggccc ggcacgtggt gggcgagtgg	9900
ggcgtagcgc gccgtctgt cgtgagccgg cggggcacgg acgccccggg cgccggcgag	9960
ctcgtgcacg agctggaggg cctgggagcc gacgtctcgg tggccgcgtg cgacgtcgcc	10020
gaccgcgaag cctcaccgc cgtactcgac tcgatccccg ccgaacaccc gtcaccgcg	10080
gtcgtccaca cggcaggcgt cctctccgac ggcacctcc cctcgatgac agcggaggat	10140
gtggaacacg tactgcgtcc caaggtcgac gccgcgttcc tcctcgacga actcacctcg	10200
acgcccggct acgacctggc agcgttcgtc atgttctcct ccgccccgc cgtcttcggt	10260
ggcgcggggc agggcgcta cgccgccgc aacgccaccc tcgacgcct cgcctggcg	10320
cgccggacag ccggactccc cgccctctcc ctcggtggg gcctctgggc cgagaccagc	10380
ggcatgaccg gcggactcag gcacaccgac cgtcgcggc tggcccgttc cggggcgacg	10440
cccatggaca gcgagctgac cctgtccctc ctggacggg ccatgcgccg cgacgacctg	10500
gcgtcgtcc cgatgcctt ggacgtcgcc gcgtccgcg ccagcagcg cgacggcatg	10560
ctggcgccgc tgcctagcgg gtcacccgc ggatcgggg tcggcgggcg gccggtcaac	10620
cagcgaggg cagccgccg aggcgcgggc gagcgggaca cggacctcg cggcgggctc	10680
gccgcgatga caccggacga ccgggtcgcg cacctgcggg acctcgtccg tacgcacgtg	10740

gcgaccgtcc tgggacacgg caccgccgagc cgggtggacc tggagcgggc cttccgcgac 10800
accggttttc actcgctcac cgccgtcgaa ctccgcaacc gtctcaacgc cgcgaccggg 10860
ctgcggtgc cggccacgt ggtcttcgac caccaccacc cgggggagct cgcggggcac 10920
ctgctcgacg aactcgccac ggcccggggc gggtcctggg cggaaggcac cgggtccgga 10980
gacacggcct cggcgaccga tcggcagacc acggcgggcc tcgccgaact cgaccggctg 11040
gaaggcgtgc tcgcctccct cgcgcccgc gccggcggcc gtccggagct cgcgcccgg 11100
ctcagggcgc tggccgcggc cctgggggac gacggcgacg acgccaccga cctggacgag 11160

gcgtccgacg acgacctctt ctcttcac gacaaggagc tgggcgactc cgacttctga 11220
11220

<210> 7

<211> 4689

<212> DNA

<213> Artificial Sequence

<220><223> PikAIII

<400> 7

atggcgaaca acgaagacaa gtcgccgac tacctcaagc gcgtcacgc cgagctgcag 60
cagaacacca ggctctcg cgagatcgag ggacgcacgc acgagccggt ggcatcggtg 120
ggcatggcct gccgcctgcc gggcggtgtc gcctcgcccg aggacctgtg gcagctgggtg 180
gccggggacg gggacgcgat ctcgaggttc ccgcaggacc gcggtggga cgtggagggg 240

ctgtacgacc ccgaccgga cgcgtccggc aggacgtact gccggtccgg cggattcctg 300
cacgacgccg gcgagttcga cgccgacttc ttcgggatct cgccgcgcga ggccctcgcc 360
atggaccgc agcagcgact gtccctcacc accgcgtggg aggcgatcga gagcgcgggc 420
atcgaccga cggccctgaa gggcagcggc ctcgcgctct tcgtcgccgg ctggcacacc 480
ggctacacct cggggcagac caccgccgtg cagtgcgccg agctggaggg ccacctggc 540
agcggcgcg cgtgggctt cctgtccggc cgtatcgct acgtcctcg tacggacgga 600
ccggccctga ccgtggacac ggcctgctc tctcgttg tcgcctgca cctcgccgtg 660

caggccctcc gcaagggcga gtgcgacatg gccctcgccg gtggtgtcac ggtcatgcc 720
aacgcggacc tgctcgtgca gttcagccgg cagcgcgggc tggccgcgga cggccggtcg 780
aaggcgttcg ccacctcgcc ggacggcttc gggcccgcg agggcgccgg agtcctgctg 840
gtggagcgcc tctcgacgc ccgcccaac ggacaccgga tctcgcggt cgtcccgcc 900
agcgcggtca accaggacgg cgccagcaac ggcctcacgg ctccgcacgg gccctcccag 960
cagcgctca tccgacgggc cctggcggac gcccggtcg cgccgggtga cgtggacgtc 1020

gtcagggcgc acggcacggg cacgcggctc ggcgacccga tcgaggcgca ggccctcatc	1080
gccacctacg gccaggagaa gagcagcgaa cagccgctga ggctgggcgc gttgaagtcg	1140
aacatcgggc acacgcaggc cgcgcccggt gtcgcaggtg tcatcaagat ggtccaggcg	1200
atgcgccacg gactgctgcc gaagacgctg cacgtcgacg agccctcgga ccagatcgac	1260
tggtcggcgg gcacggtgga actcctcacc gaggccgtcg actggccgga gaagcaggac	1320
ggcgggctgc gccgcgcggc tgtctcctcc ttcggcatca gcgggacgaa cgcgcacgtc	1380
gtcctggagg agggccccggc ggtcgaggac tccccggccg tcgagccgcc ggccggtggc	1440
ggtgtggtgc cgtggccggt gtccgcgaag actccggccg cgctggacgc ccagatcggg	1500
cagctcgccg cgtacgcgga cggtcgtacg gacgtggatc cggcgggtggc cggccgcgcc	1560
ctggtcgaca gccgtacggc gatggagcac cgcgcggtcg cggtcggcga cagccgggag	1620
gcactcgggg acgccctcg gatgccggaa ggactggtac gcggcacgtc ctgggacgtg	1680
ggccgggtgg cgttcgtctt ccccgccag ggacgcagt gggccggcat gggcgccgaa	1740
ctccttgaca gtcaccgga gttcgctgcc tcgatggccg aatgcgagac cgcgtctcc	1800
cgctacgtcg actggtctt tgaagccgtc gtccgacagg aaccggcgcc acccacgtc	1860
gaccgcgtcg acgtcgtcca gcccgtagc ttcgctgtca tggctctcgt ggccaaggtc	1920
tggcagcacc acggcatcac cccccaggcc gtctcggcc actcgcaggg cgagatcgcc	1980
gccgcgtacg tcgccgtgac actcacctc gacgacgccg cccgcgtcgt caccctgcgc	2040
agcaagtcca tcgccccca cctcggcggc aaggcgccga tgatctccct cgcctcgac	2100
gaggcggcgg tctgaagcg actgagcgac ttcgacggac tctccgtcgc cgccgtcaac	2160
ggccccaccg ccaccgtcgt ctccggcgac ccgaccaga tcgaggaaact cggccgcacc	2220
tgcgaggccg acggcgctcg tgcgcggatc atcccggtcg actacgcctc ccacagccgg	2280
caggtcgaga tcatcgagaa ggagctggcc gaggtcctcg ccggactcgc cccgcaggt	2340
ccgcacgtgc cgttcttctc caccctcgaa ggacactgga tcaccgagcc ggtgctcgac	2400
ggcacctact ggtaccgcaa cctgcgcat cgcgtgggct tcgccccgc cgtggagacc	2460
ttggcggttg acggcttcac ccacttcac gaggtcagcg cccacccgt cctcaccatg	2520
accctccccg agaccgtcac cggcctcggc accctccgcc gcgaacaggg aggccaggag	2580
cgtctggtca cctcactcgc cgaagcctgg gccaacggcc tcaccatcga ctgggcgccc	2640
atcctcccca ccgcaaccgg ccaccacccc gagctcccca cctacgcctt ccagaccgag	2700
cgttcttgge tgcagagctc cgcgcccacc agcgcgcggc acgactggcg ttaccgcgtc	2760

gagtgggaagc cgctgacggc ctccggccag gcggacctgt ccgggcgggtg gatcgctgcc	2820
gtcgggagcg agccagaagc cgagctgctg ggcgcgctga aggccgcggg agcggaggtc	2880
gacgtactgg aagccggggc ggacgacgac cgtgaggccc tcgccgcccg gctcaccgca	2940
ctgacgaccg gcgacggctt caccggcgtg gtctcgctcc tcgacgacct cgtgccacag	3000
gtcgctggg tgcaggcact cggcgacgcc ggaatcaagg cggccctgtg gtccgtcacc	3060
cagggcgagg tctccgtcgg acgtctcgac accccgcgg accccgaccg ggccatgctc	3120
tggggcctcg gccgcgtcgt cgcccttgag caccccaac gctgggcccgg cctcgctcac	3180
ctccccgcc agccgatgc cgccgcctc gccacctcg tcaccgact ctccggcgcc	3240
accggcgagg accagatgc catccgacc accggactcc acgcccgcg cctcgccgc	3300
gcacccctcc acggacgtcg gccacccgc gactggcagc cccacggcac cgtcctcatc	3360
accggcgcca ccggagccct cggcagccac gccgcacgt ggatggccca ccacggagcc	3420
gaacacctcc tctcgtcag ccgcagcggc gaacaagccc ccggagccac ccaactcacc	3480
gccgaactca ccgcatcggg cgccgcgtc accatcgccg cctgcgacgt cgccgacccc	3540
cacgcatgc gcacctcct cgacgccatc ccgcgcgaga cgccctcac cgccgtcgtc	3600
cacaccgccc gcgcaccggg cggcgatccg ctggacgtca ccggcccga ggacatgcc	3660
cgcatcctgg gcgcgaagac gagcggcgcc gaggtcctcg acgacctgt ccgcggcact	3720
ccgttgacg cttcgtcct ctactcctcg aacgccgggg tctggggcag cggcagccag	3780
ggcgtctacg cggcgccaa gcgccacctc gacgcgtcg ccggccggcg ccgcgccgg	3840
ggcgagacgg cgacctcgt cgctggggc ctctgggccg gcgacggcat gggccggggc	3900
gccgacgacg cgtactggca gcgtcgcggc atccgtccga tgagccccga ccgcgccctg	3960
gacgaactgg ccaaggccct gagccacgac gagaccttcg tcgccgtggc cgatgtcac	4020
tgggagcgtg tcgcgccgc gtacacgggtg tccgtccca gccttctgt cgacggcgtc	4080
ccggaggccc ggagggcgt cgccgaccc gtccgtgcc cggctcccgg cgacgccgc	4140
gtggcgccga ccgggcagtc gtccgcgtg gccgcgatca ccgcgtccc cgagcccag	4200
cgccggccgg cgtcctcac ctcgtcgt acccacggg cggccgtact cggccattcc	4260
tccccgacc gggtagcccc cggccgtgcc ttaccgagc tcggcttcga ctcgctgacg	4320
gccgtgcagc tccgaacca gctctccag gtggtcgga acaggtccc cgccaccag	4380
gtcttcgacc acccgacgc cgccgacac gccgcgacc tccagaggc gtacctgca	4440
ccggccgagc cgccccgac ggactgggag gggcgggtgc gccgggcct gccgaactg	4500
cccctgacc ggtcgccga cgccggggtc ctgcacaccg tctgcgct caccggcatc	4560
gagcccagc cgggttcgg cggttcggac ggcggcgccg ccgacctgg tcgggagccg	4620

gaggcgtcga tcgacgacct ggacgccgag gccctgatcc ggatggctct cggcccccgt	4680
aacacctga	4689
<210> 8	
<211> 4041	
<212> DNA	
<213> Artificial Sequence	
<220><223> PikAIV	
<400> 8	
atgacgagtt ccaacgaaca gttggtggac gctctgcgcg cctctctcaa ggagaacgaa	60
gaactccgga aagagagccg tcgccgggcc gaccgtcggc aggagcccat ggcgatcgtc	120
ggcatgagct gccggttcgc gggcggaatc cgtcccccgc aggacctctg ggacgccgtc	180
gccgcgggca aggacctggt ctccgaggta ccggaggagc gcggctggga catcgactcc	240
ctctacgacc cggtgcccgc gcgcaagggc acgacgtacg tccgcaacgc cgcgttcctc	300
gacgacgccg ccggttcga cgcgcccttc ttccgggatct cgcgcgcga ggccctcgcc	360
atggaccgc agcagcggca gctcctcgaa gcctcctggg aggtcttcga gcgggcccgc	420
atcgaccccg cgtcgttcgc cggcaccgac gtccgctgt acgtgggctg tggctaccag	480
gactacgcgc cggacatccg ggtcgcgcgc gaaggcacgc gcggttacgt cgtcacccgc	540
aactcctccg ccgtggcctc cgggcgcacg gcgtactccc tcggcctgga gggacccgcc	600
gtgaccgtgg acacggcgtg ctctctctcg ctctgcgcgc tgcacctcgc cctgaagggc	660
ctgcggaacg gcgactgctc gacggcactc gtggcgccgc tggccgtcct cgcgacgccg	720
ggcgcggttc tcgagttcag cagccagcag gccatggccg ccgacggccg gaccaagggc	780
ttcgctcgg cggcgacgg cctcgctgg ggcgaggcg tcgccgtact cctcctcgaa	840
cggctctccg acgcgcggcg caagggccac cgggtcctgg ccgtcgtgcg cggcagcgcc	900
atcaaccagg acggcgcgag caacggcctc acggctccgc acgggccctc ccagcagcac	960
ctgatccgcc aggccctggc cgacgcgcgg ctacacgtcga gcgacgtgga cgtcgtggag	1020
ggccacggca cggggacccg tctcggcgac ccgatcgagg cgcaggcgct gctcgccacg	1080
tacgggcagg ggcgcgcgcc ggggcagccg ctgcggctgg ggacgtgaa gtcgaacatc	1140
gggcacacgc aggccgcttc ggggtgcgc ggtgtcatca agatgggtgca ggcgctgcgc	1200
cacggggtgc tgccgaagac cctgcacgtg gacgagccga cggaccaggt cgactggtcg	1260
gccggttcgc tcgagctgct caccgaggcc gtggactggc cggagcggcc gggccggctc	1320
cgccggcgcg gcgtctccgc gttcggcgtg ggcgggacga acgcgcacgt cgtcctggag	1380

gaggccccgg cggtcgagga gtcccttgcc gtcgagccgc cggccggtgg cggcgtggtg	1440
ccgtggccgg tgtcccgaa gacctcgcc gactggacg ccagatcgg gcagctgcc	1500
gcatacgcg aagaccgcac ggacgtggat ccggcggtgg cggccgcgc cctggtcgac	1560
agccgtacgg cgatggagca ccgcgcggtc gcggtcgcg acagccgga ggcactcgcg	1620
gacgcctgc ggatgccga aggactgta cggggcacgg tcaccgatcc gggccgggtg	1680
gcgttcgtct tccccggca gggcacgcag tgggccggca tgggcgccga actcctcgac	1740
agctcacccg aattcgccgc cgccatggcc gaatgcgaga ccgactctc cccgtacgtc	1800
gactggtctc tcgaagcgt cgtccgacag gctcccagcg caccgacact cgaccgcgtc	1860
gacgtcgtcc agcccgtaac ctccgctc atggtctccc tcgccaaggt ctggcagcac	1920
cacggcatca cccccagggc cgtcatcgcc cactcccagg gcgagatcgc cgcccgctac	1980
gtcgccggtg cctcaccct cgacgacgc gctcgtgctg tgacctcag cagcaagtc	2040
atcgccgcc acctcgccc caaggcggc atgatctccc tcgcctcag cgaggaagcc	2100
acccggcagc gcatcgagaa cctccacgga ctgtcgatcg ccgccgtcaa cgggcctacc	2160
gccacgtgg tttcgggca cccacccag atccaagaac ttgctcaggc gtgtgagcc	2220
gacggcatcc gcgcacgat catccccgtc gactacgct cccacagcg ccacgtcgag	2280
accatcgaga acgaactgc cgacgtctg gcggggtgt cccccagac acccaggtc	2340
cccttctct ccacctcga aggcacctg atcaccgaac ccgcctcga cggcggtac	2400
tggtaccga acctccgca tcgtgtgggc ttgccccgg ccgtcgagac cctcgccacc	2460
gacgaaggct tcaccactt catcgaggtc agcgcacc cgtcctcac catgacctc	2520
cccgacaagg tcaccggcct ggccacctc cgacgcgagg acggcggaca gcaccgcctc	2580
accacctccc ttgccaggc ctgggccaac ggctcgcctc tcgactgggc ctccctcctg	2640
cccgccagg gcgcctcag ccccgccgtc ccgacctc cgacgtacgc ctccagcac	2700
cgctcgtact ggatcagccc cgggggtccc ggcgaggcg ccgcgcacac cgcttcggg	2760
cgcgaggccg tcgccgagc ggggctcgcg tggggcccgg gtgccgagga cctcgacgag	2820
gagggccggc gcagcgccgt actcgcgatg gtgatgcggc aggcggcctc cgtgctcgg	2880
tcgactcgc ccgaagaggt ccccgctgac cggcgctgc gggagatcgg ctccgactc	2940
ctgaccgcc tcgacttcg caaccgcgtc aaccggctga ccggtctcca gctgccgcc	3000
accgtcgtgt tccagaccc gacggcgct gcgctcgcc agcgcatcag cgacgagctg	3060
gccgagcgga actgggcccgt cgccgagccg tcggatcacg agcaggcgga ggaggagaag	3120
gccgcccgtc cggcgggggc ccgctccggg gccgacaccg gcgccggcg cgggatgttc	3180
cgcgccctgt tccggcaggc cgtggaggac gaccggtac gcgagttct cgacgtcctc	3240

gccgaagcct ccgcgttccg cccgcagttc gcctcgcccc aggcctgctc ggagcggctc 3300
gacccgggtgc tgctcgccgg cgggtccgacg gaccgggcgg aaggccgtgc cgttctcgtc 3360
ggctgcaccg gcaccgcggc gaacggcggc ccgcacgagt tctgcggct cagcacctcc 3420

ttccaggagg agcgggactt cctcgccgta cctctccccg gctacggcac gggtacgggc 3480
accggcacgg cctctctccc ggccgatctc gacaccgcgc tcgacgcca gggccgggcg 3540
atcctccggg ccgccgggga cgccccgtc gtctgtctcg ggcactccgg cggcgccctg 3600
ctcgcgacg agctggcctt ccgcctggag cgggcgcacg gcgcgccgc gccccgggatc 3660
gtctgtgtc acccctatcc gccgggcat caggagcca tcgaggtgtg gacgaggcag 3720
ctgggcgagg gcctgttcgc gggcgagctg gagccgatgt ccgatgcgcg gctgctggcc 3780
atgggccggt acgcgcggtt cctcgccggc ccgcggccgg gccgcagcag cgcgccctg 3840

cttctggtcc gtgcctccga accgtgggc gactggcagg aggagcgggg cgaactggct 3900
gcccactggg accttccga caccgtcgcg gacgtgccgg gcgaccactt cacgatgat 3960
cgggaccacg cgccggccgt cgccgaggcc gtctctctct ggctcgacgc catcgagggc 4020
atcgaggggg cgggcaagtg a 4041

<210> 9
<211> 846
<212> DNA
<213> Artificial Sequence
<220><223> PikAV
<400> 9

gtgaccgaca gacctctgaa cgtggacagc ggactgtgga tccggcgctt ccaccccgcg 60
ccgaacagcg cgggtgcggct ggtctgcctg ccgcacgccg gcggtccgc cagctacttc 120

ttccgcttct cggaggagct gcaccctcc gtcgaggccc tgtcgggtgca gtatccgggc 180
cgccaggacc ggcgtgccga gccgtgtctg gagagcgtcg aggagctcgc cgagcatgtg 240
gtcgcgccca ccgaacctg gtggcaggag ggccggctgg ccttcttcgg gcacagcctc 300
ggcgctccg tcgccttcga gacggccgc atcctggaac agcggcacgg ggtacggccc 360
gagggcctgt acgtctccg tcggcgccc ccgtcgttg cgccggaccg gctcgtccac 420
cagctggacg accgggcgtt cctggccgag atccggcggc tcagcggcac cgacgagcgg 480
ttctccagg acgacgagct gctgcggctg gtgctgccc cgctgcgcag cgactacaag 540

gcggcggaga cgtacctgca ccggccgtcc gccaagetca cctgcccgt gatggccctg 600
gccggcgacc gtgaccgaa ggcccgctg aacgaggtgg ccgagtggcg tcggcacacc 660

agcgggccgt tctgcctccg ggcgtactcc ggcggccact tctacctcaa cgaccagtgg	720
cacgagatct gcaacgacat ctccgaccac ctgctcgtca cccgcggcgc gcccgatgcc	780
cgcgtcgtgc agcccccgac cagccttata gaaggagcgg cgaagagatg gcagaaccca	840
cggtga	846
<210> 10	
<211> 1209	
<212> DNA	
<213> Artificial Sequence	
<220><223> DesVIII	
<400> 10	
gtgaccgacg acctgacggg ggccctcacg cagccccgcg tgggccgcac cgtccgcgcg	60
gtggccgacc gtgaactcgg caccacctic ctggagaccg gcggcatcca ctggatccac	120
gccgcgaacg gcgaccgta cgcaccgtg ctgcgcggcc aggcggacga cccgtatccc	180
gcgtacgagc ggggtgcgtg ccgcggcgcg ctctccttca gcccgaacgg cagctgggtc	240
accgccgatc acgccctggc ggcgagcatc ctctgctcga cggacttcgg ggtctccggc	300
gccgacggcg tcccggtgcc gcagcaggtc ctctcgtacg gggagggctg tccgctggag	360
cgcgagcagg tgctgccggc ggccggtgac gtgccggagg gcgggcagcg tgccgtggtc	420
gaggggatcc accgggagac gctggagggt ctgcgcggcg acccgtcggc gtcgtacgcc	480
ttcgagctgc tgggcggttt cgtccgcccg gcggtgacgg ccgctgccgc cgccgtgctg	540
ggtgttcccc cggaaccggc gcgggacttc gcggatctgc tggagcggct ccggccgctg	600
tccgacagcc tgctggcccc gcagtccctg cggacggtac gggcggcgga cggcgcgctg	660
gccgagctca cggcgctgct cgccgattcg gacgactccc ccggggccct gctgtcggcg	720
ctcggggta cgcagccgt ccagctcacc gggaacgcgg tgctcgcgt cctcgcgcat	780
cccgagcagt ggcgggagct gtgcgaccgg cccgggctcg cggcggccgc ggtggaggag	840
acctccgct acgaccgcc ggtgcagctc gacgccggg tggtcgcgg ggagacggag	900
ctggcgggccc ggcggtgcc ggccggggcg catgtcgtcg tcctgaccgc cgcgaccggc	960
cgggaccgg aggtcttcac ggaccggag cgttcgacc tcgcgcgcc cgacgccgc	1020
gcgcacctcg cgtgcaccc cgccggtccg tacggcccgg tggcgtccct ggtccggctt	1080
caggcggagg tcgcgtcgc gacctggcc gggcgtttcc ccgggctgcg gcaggcgggg	1140
gagtgctcc gccccgcg cgcgcctgtc ggccgcgggc cgtgagcgt cccggtcagc	1200
agctcctga	1209

<210> 11

<211> 1281

<212> DNA

<213> Artificial Sequence

<220><223> DesVII

<400> 11

atgcgcgtcc tgctgacctc gttcgacacat cacacgcact actacggcct ggtgccctg 60

gcctgggcgc tgctgcgcgc cgggcacgag gtgcgggtcg ccagccagcc cgcgctcacg 120

gacaccatca ccgggtccgg gctcggcgcg gtgcgggtcg gcaccgacca cctcatccac 180

gagtaccggg tgcggatggc gggcgagccg cggccgaacc atccggcgat cgccttcgac 240

gaggcccgtc ccgagccgct ggactgggac cagccctcg gcatcgaggc gatcctcgcc 300

ccgtacttcc atctgctcgc caacaacgac tcgatggtcg acgacctcgt cgacttcgcc 360

cggctcctggc agccggacct ggtgctgtgg gagccgacga cctacgcggg cgccttcgcc 420

gcccaggtea ccggtgccgc gcacggccgg gtctgtgtgg ggcccagct gatgggcagc 480

gcccccgca agttcgtcgc gctgcgggac cggcagccgc ccgagcacgc cgaggacccc 540

accgcggagt ggctgacgtg gacgctcgac cggtagggcg cctccttcga agaggagctg 600

ctcacggcc agttcacgat gacccgacc ccgcccagcc tgcgcctcga cacgggcctg 660

ccgaccgtcg ggatgcgtta tgttccgtac aacggcacgt cggtagtgcc ggactggctg 720

agtgagccgc ccgcgcggcc ccgggtctgc ctgacctcg gcgtctccgc gcgtgaggtc 780

ctcggcggcg acggcgtctc gcaggcgac atcctggagg cgctcgccga cctcgacatc 840

gagtcgtcg ccacgtcga cgcgagtcag cgcgccgaga tccgcaacta cccgaagcac 900

accggttca cggacttctg gccgatgcac gcgtctctgc cgagctgctc ggcatcatc 960

caccacggcg gggcgggcac ctacgcgacc gccgtgatca acgcggtgcc gcaggatcatg 1020

ctcgcggagc tgtgggacgc gccgtcaag gcgcgggccc tcgccgagca gggggcgggg 1080

tttttctcgc cgcggccga gctcagccg caggccgtgc gggacgccgt cgtccgcatc 1140

ctcgacgacc cctcggtcgc caccgccgcg caccggctgc gcgaggagac cttcggcgac 1200

cccccccg ccgggatcgt ccccgagctg gagcggctcg ccgcgcagca ccgccgccc 1260

ccggccgacg cccggcactg a 1281

<210> 12

<211> 714

<212> DNA

<213> Artificial Sequence

<220><223> DesVI

<400> 12

gtgtacgaag tcgaccacgc cgacgtctac gacctcttct acctgggtcg cggcaaggac 60
tacgccgccg aggcctccga catcgccgac ctgggtcgct cccgtacccc cgaggcctcc 120
tcgtccttg acgtggcctg cggtaggggc acgcatctgg agcacttcac caaggagttc 180

ggcgacaccg cgggcctgga gctgtccgag gacatgtctc cccacgccg caagcggctg 240
cccagacca cgctccacca gggcgacatg cgggacttcc ggctcggccg gaagtcttcc 300
gccgtggtca gcatgttcac ctccgtcggc tacctgaaga cgaccgagga actcggcgcg 360
gccgtcgcct cggtcgcgga gcacctggag cccggtggcg tcgtcgtcgt cgagccgtgg 420
tggttcccgg agaccttcgc cgacggctgg gtcagcgccg acgtcgtccg ccgtgacggg 480
cgcaccgtgg cccgtgtctc gcaactcgtg cgggagggga acgcgacgcg catggaggtc 540
catttcaccg tggccgaccc gggcaagggc gtgcggcact tctccgacgt ccatctcatc 600

acctgttcc accaggccga gtacgaggcc gcgttcacgg ccgccgggct gcgcgtcgag 660
tacctggagg gcggcccgtc gggccgtggc ctcttcgtcg gcgtccccgc ctga 714

<210> 13

<211> 2430

<212> DNA

<213> Artificial Sequence

<220><223> DesR

<400> 13

gtgacaggta agaccgaat accgctgtc cgcccgggcc gcaccacgc cagggccttc 60
acctggccg tcgtcggcac cctgttgccg ggcaccaccg tggcgccgc cgctcccggc 120
gccgccgaca cggccaatgt tcagtacacg agccgggccc cggagctcgt cggccagatg 180
acgtcgacg agaagatcag ctctgtccac tggcgctgg accccgaccg gcagaacgtc 240

ggctaccttc cggcgtgcc gcgtctgggc atcccggagc tgcgtgccgc cgacggcccg 300
aacggcatcc gcctggtggg gcagaccgcc accgcgtgc ccgcgccgt cgccctggcc 360
agcaccttc agacaccat ggccgacagc tacggcaagg tcattggccg cgacggtcgc 420
gcgtcaacc aggacatggt cctgggccc atgatgaaca acatccgggt gccgcacggc 480
ggccggaact acgagacctt cagcaggac ccctggtct cctcgcgcac cgcggtcgcc 540
cagatcaagg gcatccagg tgcgggtctg atgaccacgg ccaagcatt cgcggccaac 600
aaccaggaga acaaccgctt ctccgtgaac gccaatgtcg acgagcagac gctccgcgag 660

atcgagtacc cggcggttcca ggcttcctcc aaggccggcg cggcctcctt catgtgtgcc	720
tacaacggcc tcaacgggaa gccgtcctgc ggcaacgacg agtcctcaa caacgtgctg	780
cgcacgcagt ggggcttcca gggctgggtg atgtccgact ggctcgccac cccgggcacc	840
gacgccatca ccaagggcct cgaccaggag atggcgctcg agtccccg cgacgtccc	900
aaggcgagc ctcgcccgc ggccaagtgc ttggcgagg cgctgaagac ggccgtcctg	960
aacggcacgg tccccaggc ggccgtgacg cggtcggcgg agcggatcgt cggccagatg	1020
gagaagttcg gtctgtcct cgccactccg gcgccggcg ccgagcgga caaggcgggt	1080
gcccaggcgg tgtcccga ggtcgccgag aacggcgcg gtgtcctgcg caacaggggc	1140
caggccctgc cgtcgccgg tgacccggc aagagcatcg cggcatcgg cccgacggcc	1200
gtcgaccca aggtcacgg cctgggcagc gccacgtcg tccggactc ggccggcgcg	1260
ccactcgaca ccatcaagg ccgcgcggt gcgggtcgga cggtagcgt cgagacgggt	1320
gaggagacct tcgggacga gatccggcg gggaacctca gcccgcggt caaccagggc	1380
caccagctcg agccgggcaa ggccggggcg ctgtacgac gcacgtgac cgtgcccgc	1440
gacggcgagt accgcatcg ggtccgtgc accggtggtt acgccacgt gcagctcggc	1500
agccacacca tcgaggccgg tcaggtctac ggcaaggtga gcagcccgt cctcaagctg	1560
accaagggca cgcacaagct cagcatctcg ggcttcgga tgagtccac cccgtctcc	1620
ctggagctgg gctgggtgac gccggcgcg gccgacgga cgtcgcaa ggccgtggag	1680
tcggcgcgga agggccgtac ggccgtcgtc ttgcctacg acgacggc cgaggcgctc	1740
gaccgtcga acctgtcgt gccgggtacg caggacaagc tgatctcggc tgtcgcgac	1800
gccaaccga acacgatcgt ggtcctaac accggttcgt cggctgtgat gccgtggctg	1860
tccaagacc gcgcgtcct ggacatgtgg taccggggc aggcggcgcg cgaggccacc	1920
gccgcgtgc tctacgtga cgtcaaccg agcggcaagc tcacgcagag cttcccggcc	1980
gccgagaacc agcacgggt cgccggcgac ccgacaagct acccggcggt cgacaaccag	2040
cagacgtacc gcgaggcat ccacgtcggg tacgctggt tcgacaagga gaacgtcaag	2100
ccgtgttcc cgttcgggca cggcctgtcg tacacctgt tcacgcagag cggcccgacc	2160
gtcgtgcgt cgtccacggg tggctgaag gtcacgtca cggtcgcaa cagcggaag	2220
cgcgccggc aggaggtcgt ccaggcgtac ctggtgcca gccgaacgt gacggctccg	2280
caggcgaaga agaagctcgt gggctacac aaggtctcg tcgccggcg cgaggcgaag	2340
acggtgacgg tgaacgtga cgcctcag ctgcagacc gttcgtctc gccgacctg	2400
cggggcagcg ccacgtcaa cgtctggtga	2430

<210> 14

<211> 1140

<212> DNA

<213> Artificial Sequence

<220><223> DesV

<400> 14

ctaggcctgg tcgacccgct cggcccattc gcgcacggcg tcgatcaccc gcagcgcttg 60

cgggcgctcc aggtgcgggc cgatcggcag gctgaggacc tgccgcgcga agctctcggc 120

ccgcgggagc gaggcttccg gcggtgcctc gcccgctag gcgggcgaga ggtgcacggg 180

taccgggtag tgcgtgaggg tgctgatgcc gcgggcgtcg aggtggctgc gcagctcgtc 240

gcggcgctcg gtgcgcacgg tgaagaggcg ccagaccggg tcggtgtcgg gcgcggtcac 300

cggcaggccg atcccgggca gtccggcgag ccggagagg tactccgcgg ccagcgccga 360

cctgcggcgg ttccagctgt ccaggtgggc gagccggatc cgcagcacgg cggcctgcat 420

ctcgtccagg cgggagttgg tgcccttcgt ctctgtgctg tacttctgcc gcgagccgta 480

gttcgggagc atccggagcc gtccggcgag ctccgggtcg ccggtgacga cggcgccgcc 540

gtcgccgaag cagccgaggt tcttccccgg gtagaagctg aacgcggcca ccgacgaccc 600

ggcgccgatc cggcgccccc ggtagcgggc gccgtgggcc tgcgcggcgt cctcgacgat 660

gtgcaggccg tgccggtccg cgagctcgcg gagggcgctc atgtcggcgg ggtgcccgta 720

gaggtggacg gggaggagcg cccgggtgcg gggggtgatc gccttctcga cgagcagcgg 780

gtccagggtg ggggtggtcct cgtgcggctc gacgggcacg ggggtcgcgc cgggtggcga 840

caccgcgagc cagctggcga tgtacgtgtg cgaggggacg atcacctcgt ccccggttcc 900

gatgccgagg ccgcggaggg cgagctggag ggcgtccatc ccgtgttca cgcgacggc 960

gtggtccgtc tcgcagtacg gcgcgaactc cgcctcgaat ccttcgagtt cgggtccgag 1020

gaggtagcgc cccgagtcga ggacgcgggc gatcgcgcg tcggtctccg cgcggagctc 1080

ctcgtaggcg gccttgaggt cgaggaaggg gacgcggggg gtctcggcgc ggctgctcac 1140

1140

<210> 15

<211> 1014

<212> DNA

<213> Artificial Sequence

<220><223> DesIV

<400> 15

tcacgcggac acctccacgg cgggtggcggg cagctgcggg gcggtcgcct tgagcggctc 60

ccaccagccg cggttctccc ggtaccagcg gacggtccgc gcgaggccgt ccgcgaagga	120
gacctgcggg cggtagccga gctcgcgctc gatctcgccg ccgtcgaggg agtagcgag	180
gtcgtggccc ttcggtcgg cgaccttccg gaccgaggac cagtcggcgc cgagcgagtc	240
caggaggatg ccggtgagtt cgcggttggf cagctccagg ccgccgccga tgtggtagat	300
ctcgccggcc cgcccgcccg cgaggacgag cgcgatgccc cggcagtggt cgtcggtgtg	360
cacccactcg cggacgttcg cgccgtcgcc gtacagcggg agcgtcccgc cgtcgaggag	420
gttcgtcacg aagaggggga tgagcttctc ggggtgctgg tacggcccg agttgttgca	480
gcagcgggtg atccgtacgt cgaggccgta cgtccggtgg taggcgcggg caacaggtc	540
ggagccggcc ttggacgccg cgtagggcga gttgggctcc agcgggctgc tctcggtcca	600
ggagccggag tcgatcgacc cgtacacctc gtcggtggag acgtgcacga cccggccgac	660
gccggcgctc acggcgcaact ggagcagcgt ctgcgtgccc tgcacgttgg tctcggtgaa	720
cacggacgcg ccccgatgg agcggctcac gtggctctcg gcccggaagt ggacgatggc	780
gtccacgccg cgcagttccc gggcgaggag gccggcgctc cggatgtcgc cgtggacgaa	840
gcgcagtgcg gggtcgcgt ccaccggggc gaggttggcg cggttgccg cgtaggtgag	900
gctgtccagg acgatcacct catcgccggg cagctcgggg tacccccgg cgaggagctg	960
ccgcacgaag tgcgagccga tgaagccgc acctccggtc accagaagcc gcac	1014
<210> 16	
<211> 879	
<212> DNA	
<213> Artificial Sequence	
<220><223> DesIII	
<400> 16	
tcacggggct ccctcgccg cggcgatctc catcagatag ctgccgtact cgggtcggga	60
gaggccttct ccaggccgt gacaggctc ggcgtcgatg aagcccatgc ggaaggcgat	120
ctctcaagg ccccgatcc agacgccctg ccgctctcc aggacctgga cgtactgggc	180
ggcccgcagg agcgagtcgt ggggtccggt gtccagccag gcgaagccgc ggcccaggtt	240
gacgagttcg gcccggcccc gctccaggta gacgcggttg acgtcggtga tctccagtc	300
gccgcgcggc gagggccgga tgttcttggc gatgtcgacg acgtcggtgt cgtagaggta	360
gaggccggtg acggcgaggt tggagcgcg cttgacgggc ttctcgacga ggtcggtcag	420
ccggcccgtc gcgtccacct cggcgacgcc gtaccgctcg gggtccttga ccgggtagcc	480
gaagagcacg cagccgtcga ggcgcgcat gctgtcccgc aggagcgtgt agaggccggg	540

cccgtggaag atgttgtcgc ccaggatcag ggcgcagggtg tcgtcgccga tgtgctcggc 600
 tccgacgaga agtgcgtccg cgattcctgc gggctctttc tggaccgcat agtcgagttc 660
 tattcccagg tgcctgccgt ttccgagaag cgactggaag agttcgatgt gctggggggt 720
 cgagatgatt tgaatctcgc gaataccgcc gagcatgaga accgacagcg gatagtagat 780

catcggtttg ttgtagaccg gaagaatctg cttcgaaatg accgaggtcg ccggatgcag 840
 ccgagttccg ctcccgcggg ccaggactat tcccttcat 879

<210> 17
 <211> 1458
 <212> DNA
 <213> Artificial Sequence
 <220><223> DesII
 <400> 17

tcagcgcagg aagccgcggg cctcctccca gccgtccgcg gcgtcgcgct ccagctggtt 60
 caggcggggc gtgacgacct gatcgaagcc gtccatgaag tactcgtcgc cgtcgacggc 120
 cgccacctcg ccgccgcgct cgacgaagtc cctgacgacc tcggtgaggg aggtgtcggg 180
 ggtcacgcgg ccgcgatgt agcgggtcgc gccgtccagg tcggggaagc cggcctcgcg 240

gtacaggtac acgtcgccga ggagatcgac ctgcaccgcg acctgcgggt gcgcggtggg 300
 ccgatgggtg gcgggcttga tccgagcag ttccggctcg gccccggtgc gcaggctgtt 360
 cagggcgtag ccgtagtcca tgtggagtcc gggggtcgc tcgcggaccc gctcctcgaa 420
 ggcggtgagg gcctcctgga gctcggcccg ctctcctgc ggacagcttg cgtcgtcacg 480
 gccgtgtag tcctcgcaaa tgttgacgaa gtgatcgtc ctgccctgcc cggcgtcgtt 540
 gaggtcggcg atgaagtcga ccaggtcgag caggcgggag gcacggcccg ggagcacgat 600
 gtaggcgaag ccgaggttga tcggcgactc gcgctcggcg cgcagctgct ggaagcggcg 660

caggttctcg cggacgcggc ggaaggcggc ctctctgccg gtggtctgct cgtactctc 720
 gtcgttgagg ccgtagagcg aggtgcggat ggctgcagg cccagaggc cgggctggcg 780
 ctccagggtg cgctcggta gcgcgaagga gtctgttag acggtgggcc gcaggccgtg 840
 gtcggtggcg tgcgcggcca ggctcccag gccggggttg gtgagcggct ccaggccgcc 900
 ggagaagtac atcgccgagg ggttccccgc gggtatctcg tcgatgaccg accggaacat 960
 ggcggtgccg gcgtcgaggg cggacgggtc gtagcgggcg ccggtcacac ggacgcagaa 1020
 gtggcagcgg aacatgcagg tcgggccggg gtagaggccg acgctgtacg ggaagacggg 1080

cttcctggcg agcggccgct cgaagacgcc gcgctgttcg agcgggagca gggtgttctt 1140

ccagtacgcc ccggcggggc cggctctgac cgcggtgcgg agctccggga cctgcccga 1200
cagggcgagg aggcgccgga aggcgtcccg gtcgacgccc aggtcgtggc gggcctcctc 1260
cagcggggtg aaggggctgt tgccgtagcg cacggcgagc cggacgaggt ggcgggcggt 1320
cgttcgggcc tcgtcgggcg gcacgaggcc gccggcggcg agggctctggc cgacggcgtg 1380
gaccgccgcc ccagatcgg ctccgggggtg cgcgcagcgt tcggccgggg cggcggcgga 1440
aaggcggggg gcggtcat 1458

<210> 18

<211> 1248

<212> DNA

<213> Artificial Sequence

<220><223> DesI

<400> 18

tcatcgggag cgtccaatcg tgggcgtgga tgtctggggg gccgcgagcg gggcgggggc 60
cgtgtcgcgg tggcgcgcgg tcagttcgcg gccgcgggtc gcgcagagac gcagcaggtc 120
ggcgacccgg cggatgtcgt cgtcggcgat ggcggtgccg gtcggcaggg acagcacgcg 180
cgcggcgagg cgttcggtgt gcggcagcgg ggcgtgcggc tgcccgcggt acggctccag 240
ctcgtggcag cccggcgaga agtaggcgcg ggtgtgcacg ccttcggcct tcaggacctc 300
catgacgagg tcgcggtgga tgccggtggt ggcctcgtcg atctcgacga tcacgtactg 360

gtggttgttg aggccgtggc ggtcgtggtc ggcgacgagg acgccgggga ggtccgcgag 420
gtgtcgcgg taggcggcgt ggttgcgccg gttccggtcg atgacctcg gaaacgcgtc 480
gagggaggtg aggcccatgg cggcggcggc ctgctcatc ttggcgttgg tcccgcggc 540
ggggctgccg ccgggcaggt cgaagccgaa gttgtggagg gcgcggatcc gggcggcgag 600
gtcggcgtcg tcggtgacga cggcgccgcc ctcaaggcg ttgacggcct tggcggcgtg 660
gaagctgaag acctcggcgt cgccgaggct gccggcgggc cggccgtcga ccgcgcagcc 720
gagggcgtgc gcggcgtcga agtacagccg caggccgtgc tcgtcggcga ccttcgcgag 780

ctggtcggcg gcgcaggggc ggccccagag gtggacgccg acgacggccg aggtgcgggg 840
tgtgaccgcg gcggccacct ggtccgggtc gaggttgccg gtgtccgggt cgatgtcggc 900
gaagaccggg gtgaggccga tccagcgagc tgcgtgcggg gtggcggcga acgtcatcga 960
cggcatgata acctcggcgg tgaggccggc ggcgtgcgcg aggagctgga gcccgccgt 1020
ggcgttcgag gtggccacgg catgccggac cccggcgagc ccggcgacgc gtcctcga 1080
ctcgcggacg agcgggccgc cgttggacag ccactggctg tcgagggccg ggtcgagccg 1140

ctcgtacagc ctggcgcggt cgatgcggtt gggccgcccc acgaggagcg gctggtcgaa 1200

agcggcgggg ccgccgaaga atgcgaggtc ggataaggcg cttttcac 1248