



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0139150

(43) 공개일자 2015년12월11일

(51) 국제특허분류(Int. Cl.)

C12P 1/04 (2006.01) C07D 417/02 (2006.01)

C07D 417/06 (2006.01) C12N 1/20 (2006.01)

C12N 15/09 (2006.01) C12P 17/14 (2006.01)

C12R 1/61 (2006.01)

(21) 출원번호 10-2014-0067239

(22) 출원일자 2014년06월02일

심사청구일자 2015년05월21일

(71) 출원인

이화여자대학교 산학협력단

서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)

(72) 발명자

윤여준

서울 강남구 압구정로 151, 126동 202호 (압구정동, 현대아파트)

송재경

충청남도 아산시 신창면 온천대로 725

(뒷면에 계속)

(74) 대리인

손민

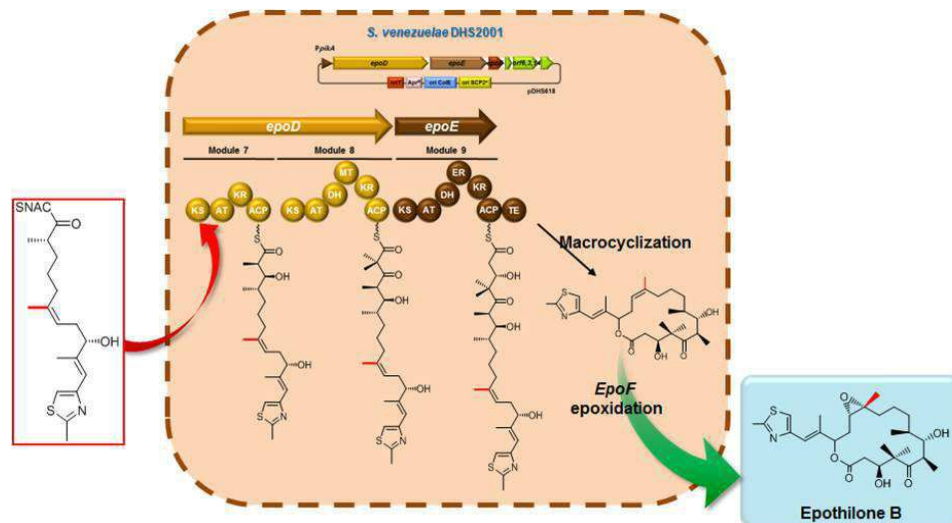
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 방선균에서 에포틸론 생산을 증대시키기 위한 에포틸론의 전구체, 및 이를 이용한 에포틸론의 생산 방법

(57) 요약

본 발명은 방선균에서 에포틸론 생산을 증대시키기 위한 화합물, 및 이를 포함한 배지에 소란지움 셀룰로스의 에포틸론 생합성 유전자 *epoD*, *epoE*, *epoF*, *orf6*, *orf3*, 및 *orf14*가 도입된 방선균을 배양하는 단계를 포함하는, 수율이 향상된 에포틸론의 생산 방법에 관한 것으로, 방선균에서의 에포틸론 생산 수율을 향상시킬 수 있으며, *epoA*, *epoP*, *epoB*, 및 *epoC*가 도입되지 않은 방선균에서도 에포틸론을 생산할 수 있다.

대표도 - 도1b



(72) 발명자

김웅수

서울특별시 강남구 언주로30길 26 삼성타워팰리스
G동 5006호

김원석

서울 동대문구 답십리로66길 46-10

김은지

서울 송파구 송파대로 567, 508동 305호 (잠실동,
잠실5단지아파트)

김은지

서울 서대문구 신촌로 215-2, 803호 (북아현동, 전
진빌딩)

김면수

서울 서대문구 이화여대8길 3, 4층 IP레지던스 42
3호 (대현동, 연화프라자)

앤드류, 래스웰 크리스 제임스

경기도 수원시 영통구 영통로 290번길 26 두산아파
트 806동 2003호

나지혜

서울특별시 강서구 방화대로 49길 6-16 중앙하이빌
501호

이 발명을 지원한 국가연구개발사업

과제고유번호 1711002649

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 원천기술개발사업-글로벌프론티어사업

연구과제명 (지능형바이오시스템)거대 인공 이차대사산물 생합성 유전자집단의 발현시스템 개발

기 여 율 1/2

주관기관 이화여자대학교 산학협력단

연구기간 2011.09.29 ~ 2020.08.31이 발명을 지원한 국가연구개발사업

과제고유번호 1711000101

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 기초연구사업(학술진흥)(2013)-중견연구자지원사업-도약연구지원사업(도전)

연구과제명 미생물 유래 천연물의 생합성경로 규명 및 재설계

기 여 율 1/2

주관기관 이화여자대학교 산학협력단

연구기간 2013.06.01 ~ 2016.05.31

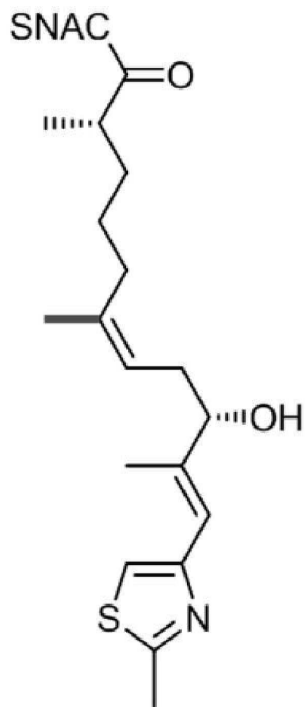
명세서

청구범위

청구항 1

하기 화학식 3으로 표시되는 화합물을 포함한 배지에, 소란지움 셀룰로섬(*Sorangium cellulosum*)의 에포틸론 생합성 유전자 *epoD*, *epoE*, *epoF*, *orf6*, *orf3*, 및 *orf14*가 도입된 방선균을 배양하는 단계를 포함하는, 수율이 향상된 에포틸론의 생산 방법.

[화학식 3]



청구항 2

제1항에 있어서,

상기 방선균은 스트렙토마이시스 베네주엘레 또는 스트렙토마이시스 실라갈라인, 수율이 향상된 에포틸론의 생산 방법.

청구항 3

제1항에 있어서,

상기 에포틸론 생합성 유전자는 *epoA*, *epoP*, *epoB*, 및 *epoC*를 추가로 포함하는 것인, 수율이 향상된 에포틸론의 생산 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 에포틸론 생합성 유전자는 스트렙토마이시스 베네주엘라(*Streptomyces venezuelae*)에 코돈 최적화된 것인, 수율이 향상된 에포틸론의 생산 방법.

청구항 5

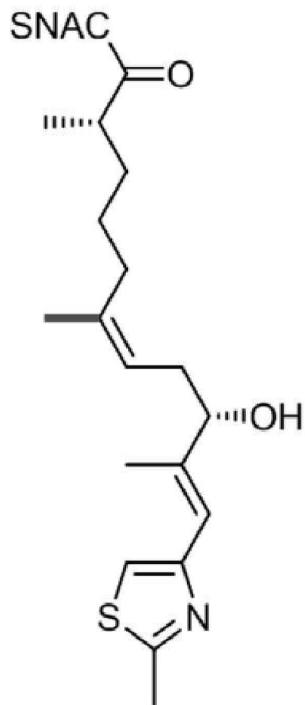
제1항 내지 제3항 중 어느 한 항에 있어서,

상기 에포틸론 생합성 유전자는 pDHS702, pDHS618, pSBAC, 및 pSET152로 이루어지는 군에서 선택되는 어느 하나의 벡터에 합성된 벡터로 도입되는, 수율이 향상된 에포틸론의 생산 방법.

청구항 6

하기 화학식 3으로 표시되는 화합물을 포함하는, 에포틸론 생산용 배지 조성물.

[화학식 3]



청구항 7

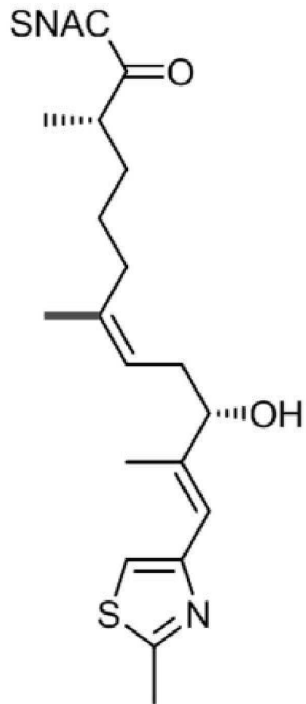
제6항에 있어서,

상기 배지는 스트렙토마이시스 베네주엘라 또는 스트렙토마이시스 실라칼라의 배양용인, 에포틸론 생산용 배지 조성물.

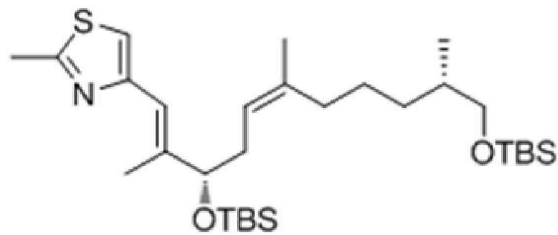
청구항 8

하기 화학식 4로 표시되는 스킨저 중간체(Schinzler's intermediate)에 옥시데이션(oxidation), 디프로텍션(deprotection), 및 티오에스터(thioester) 합성 반응을 가하는 단계를 포함하는, 하기 화학식 3으로 표시되는 에포틸론 전구체 화합물의 제조방법.

[화학식 3]



[화학식 4]

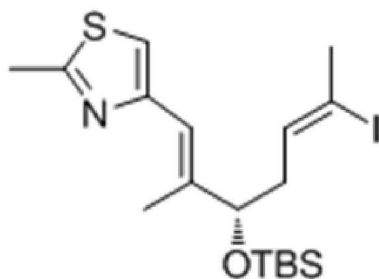


청구항 9

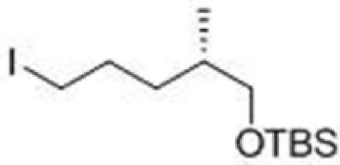
제8항에 있어서,

상기 스킨저 중간체는 하기 화학식 5로 표시되는 요오드화비닐(vinyl iodide) 및 하기 화학식 6으로 표시되는 요오드화물로부터 합성되는 것인, 에포틸론 전구체 화합물의 제조방법.

[화학식 5]



[화학식 6]



청구항 10

제8항에 있어서,

상기 제조방법은 (a) 티오아세트아미드(thioacetamide, 1)로부터 합성된 포스페이트 에스터(phosphate ester, 4)를 합성하는 단계;

(b) L-말산(L-malic acid, 5)으로부터 케톤(ketone, 11)을 합성하는 단계;

(c) 상기 포스페이트 에스터(phosphate ester, 4) 및 상기 케톤(ketone, 11)로부터 요오드화비닐(vinyl iodide, 15)을 합성하는 단계;

(d) (R)-벤질-옥사졸리디논((R)-Benzyl-oxazolidinone, 16)으로부터 제1요오드화물(21)을 합성하는 단계;

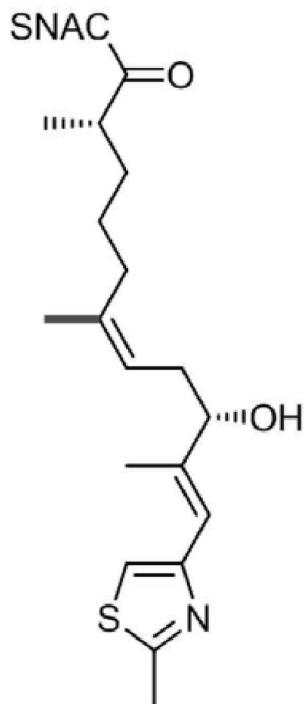
(e) 상기 요오드화비닐(vinyl iodide, 15) 및 제1요오드화물(21)로부터 N-아세틸 시스타민 티오에스터(N-acetyl cysteamine(SNAC) Thioester, 25)를 합성하는 단계를 포함하는 것인, 제1항에 따른 전구체 화합물의 제조방법.

청구항 11

하기 화학식 3으로 표시되는,

방선균의 에포틸론 생산 증대를 위한 에포틸론의 전구체 화합물.

[화학식 3]



발명의 설명

기술 분야

[0001] 본 발명은 방선균에서 에포틸론 생산을 증대시키기 위한 화합물, 이를 포함한 에포틸론 생산용 배지 조성물, 이를 이용한 수율이 향상된 에포틸론의 생산방법, 및 이의 제조방법에 관한 것이다.

배경 기술

[0002] 식물 및 미생물 등에서 생산되는 폴리케타이드(polyketide) 계열의 화합물은 복잡한 구조의 천연 화합물(natural product)로서, 현재까지 약 10,000 가지 이상의 화합물이 알려져 있다. 그 중 에포틸론은 16환계 폴리케타이드 매크로락톤(16-membered polyketide macrolactone)의 일종으로 항암 화학요법제인 탁솔(Taxol)과 유사하게 튜불린(tubulin)에 결합하여 미세소관(microtubule)의 분해를 억제하는 강력한 항암제이다.

[0003] 에포틸론 B로부터 반합성 되는 익사베필론(성분명: Ixabepilone, 상품명: Ixampra, Bristol-Myers Squibb)은 2007년 하반기에 유방암의 치료제로 FDA의 승인을 얻었다. 특히 안트라사이클린계(anthracycline)와 탁산계(taxane) 치료에 내성이 생겼거나, 또는 탁산계에 내성이 있고 더 이상의 안트라사이클린계 치료가 금기인 전이성 유방암 또는 국소 진행성 유방암의 치료에 사용이 가능하다. 또한 익사베필론은 탁솔과 다르게 부작용을 유발하는 가용화제(solubilizing agent) 없이 제형화할 수 있으며, 통상적인 화학 치료법(chemotherapy)에 내성이 있는 암세포에 강한 항암 작용을 보이는 장점이 있어 차세대 항암제로 대두되고 있다.

[0004] 에포틸론의 생산은 미소박테리움(myxobacterium)인 소란지움 셀룰로섬(*Sorangium cellulosum*)에 의해 생산되는 이차대사산물로써 분리되었으나, 16시간에 이르는 긴 성장시간(doubling time)과 유전자 조작의 어려움으로 인하여 그 산업적 생산을 위한 유용성이 떨어지는 문제점이 있다.

[0005] 한편, 경제적인 산업적 가치가 매우 높으나 목적하는 이차대사산물의 대량생산에 기술적 어려움이 있는 경우, 현재의 화학합성을 대체할 수 있는 효율적 방안 중 하나로 대량 생산하고자 하는 이차대사산물의 생합성 유전자집단을 자연생산 생물체로부터 확보하고, 이를 최적화된 이종숙주에서 발현하는 이종숙주 생산시스템(heterologous expression system)이 연구되고 있다. 미국 및 독일의 연구팀에서 이종숙주 생산시스템을 이용하여 스트렙토마이시스 실리칼라, 미소코쿠스 잔서스(*Myxococcus xanthus*), 에스케리치아 콜라이(*Escherichia coli*), 그리고 스트렙토마이시스 베네주엘라에서 에포틸론의 생합성을 시도하여 왔다. 독일 Rolf Miller 연구팀에서는 에포틸론 생합성 유전자를 최적화하여 합성한 후 미소코쿠스 잔서스에서 발현하는 기술(Owald C et al., ACS Synth Biol., 2012. 10)을 보고한 바 있다.

[0006] 또한, 본 발명자들은 방선균 중 스트렙토마이시스 베네주엘라(*Streptomyces venezuelae*)에서의 에포틸론 생산방법에 대하여 출원한 바 있다(한국 특허등록 제 10-0949313호). 그러나 상기 발명들은 에포틸론 B의 경우 약 0.1 $\mu\text{g/L}$, 에포틸론 D의 경우 약 0.4 $\mu\text{g/L}$ 의 수율을 나타내어, 에포틸론을 산업 수요에 맞추어 생산할 만큼 수율이 높지 않다는 문제가 있었는데, 에포틸론의 생산 효율을 높이기 위한 기술이 여전히 요구되고 있다.

[0007] 이러한 배경 하에, 본 발명자들은 유용한 항암물질인 에포틸론의 수율을 높이는 방법을 개발하고자 예의 연구 노력한 결과, 에포틸론 생합성 유전자 전체 클러스터 없이 소란지움 셀룰로섬에서 유래한 에포틸론 생합성 유전자 *epoD*, *epoE*, *epoF*, *orf6*, *orf3*, 및 *orf14*가 도입된 방선균에 신규 합성한 전구체 화합물을 처리함으로써, *epoA*, *epoP*, *epoB*, 및 *epoC* 유전자 없이도 에포틸론을 높은 수율로 생산함을 확인하고 본 발명을 완성하였다.

발명의 내용

해결하려는 과제

[0008] 본 발명의 주된 목적은 방선균에서 에포틸론 생산을 증대시키기 위한, 에포틸론의 전구체 화합물을 제공하는 것이다.

[0009] 본 발명의 다른 목적은 상기 화합물을 포함하는 에포틸론 생산용 배지 조성물을 제공하는 것이다.

[0010] 본 발명의 다른 목적은 상기 화합물을 포함한 배지에, 소란지움 셀룰로섬(*Sorangium cellulosum*)의 에포틸론 생합성 유전자 *epoD*, *epoE*, *epoF*, *orf6*, *orf3*, 및 *orf14*가 도입된 방선균을 배양하는 단계를 포함하는, 수율이 향상된 에포틸론의 생산 방법을 제공하는 것이다.

[0011] 본 발명의 또 다른 목적은 스킨저 중간체(Schinzler's intermediate)에 옥시데이션(oxydation), 디프로텍션(deprotection), 및 티오에스터(thioester) 합성 반응을 가하는 단계를 포함하는, 상기 전구체 화합물의 제조방법을 제공하는 것이다.

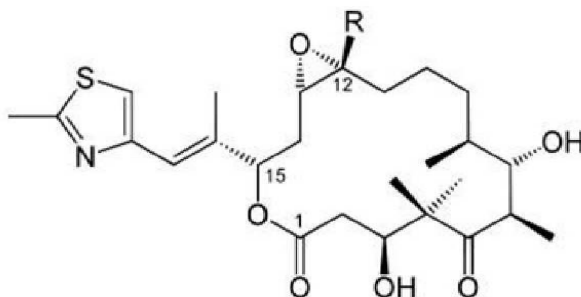
과제의 해결 수단

[0012] 상기의 목적을 달성하기 위한 하나의 양태로서, 본 발명은 방선균에서 에포틸론 생산을 증대시키기 위한 에포틸론의 전구체 화합물을 제공한다.

[0013] 이하, 본 발명에서의 에포틸론 생산을 증대시키기 위한 합성된 전구체 화합물을 구체적으로 설명한다.

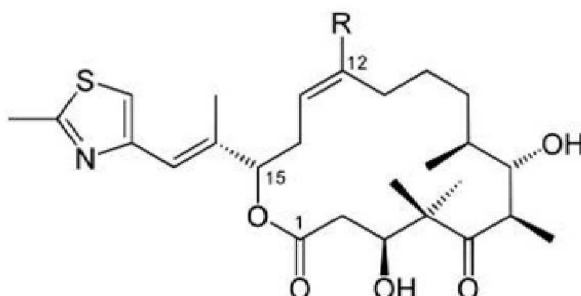
[0014] 본 발명에서 용어, 에포틸론은 메틸티아졸(methylthiazole)기가 올레핀 결합(olefinic bond)에 의하여 결합된 16환계 폴리케타이드 매크로락톤(polyketide macrolactone)의 일종으로, 저성장(low-growing) 미소박테리움(myxobacterium)인 소란지움 셀룰로섬에 의해 생산되는 이차 대사산물로써 최초 분리되었다. 상기 에포틸론은 하기 화학식 1로 표시되며 R이 H인 에포틸론 A, 하기 화학식 1로 표시되며 R이 CH₃인 에포틸론 B, 하기 화학식 2로 표시되며 R이 H인 에포틸론 C, 하기 화학식 2로 표시되며 R이 CH₃인 에포틸론 D를 포함할 수 있다. 에포틸론 C와 D는 에포틸론 A와 B의 생합성 초기단계의 중간물질이다.

화학식 1



[0015]

화학식 2



[0016]

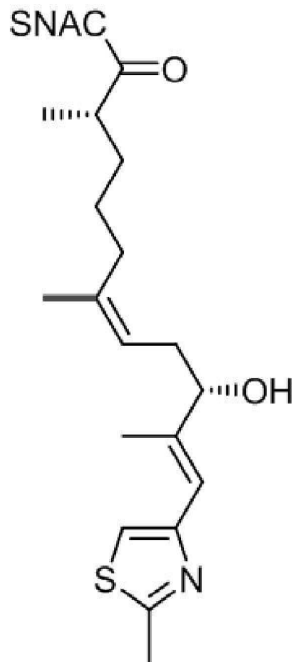
[0017] 본 발명의 에포틸론은, 에포틸론의 유사체, 유도체, 역사베필론을 포함하는 의미로 사용될 수 있다.

[0018] 상기 에포틸론은 튜블린 중합 에세이에서 항종양 활성을 가지는 것으로 보고되었으며(Bollag, D. et al.,

Cancer Res., 55:2325, 1995), 탁솔이나 탁소텔 등의 항암제처럼 미세소관(microtubule)을 안정화시키고 세포 분열을 억제함으로써 암 세포의 증식을 막는 것으로 알려져 있는 유용한 의약품의 원료이다.

본 발명에서 용어, 전구체는 화학반응 등에서 최종적으로 얻을 수 있는 특정 물질이 되기 전 단계의 물질을 의미하며, 에포틸론의 수율을 향상시키기 위한 본 발명의 목적상, 상기 화학식 1 또는 화학식 2로 표시되는 에포틸론의 전구체를 의미하며, 특히 하기 화학식 3으로 표시되는 합성된 화합물을 의미한다.

화학식 3



본 발명의 전구체는 바람직하게는 도 1a의 에포틸론 생합성 모듈 6번에 작용하는 전구체로서 티오에스터 SNAC(thioester SNAC)의 형태인 화합물을 의미한다.

본 발명에 따른 합성된 전구체는 배지에 포함되는 방법 등으로 방선균에 처리될 경우, 에포틸론 생합성 유전자를 도입한 방선균에서 에포틸론의 생산을 증대시킬 수 있다. 현재까지 에포틸론을 생산하기 위하여 방선균에 적용될 수 있는 전구체 화합물은 보고된바 없으며 본 발명자들에 이르러 최초로 합성되어 효과가 확인된 것이다.

에포틸론의 생합성 과정은 도 1a에 도시된 *epoA*, *epoP*, *epoB*, *epoC*, *epoD*, *epoE*, 및 *epoF*의 유전자 클러스트에 의하여 일어나는 것으로 알려져 있다. 그러나, 이러한 대량의 유전자 클러스터를 이중속주세포에 도입 시, 도입방법 또는 도입 전달체인 벡터 내에 포함시킬 수 있는 유전자 크기 등의 제한으로 인하여, 효과적인 도입에 한계가 있었다. 본 발명자들은 에포틸론 생합성을 시작하는 제1효소부터의 일련의 합성이 아닌, 일부 유전자 도입 및 상기 에포틸론의 전구체 도입에 의하여 효과적으로 에포틸론을 제조할 수 있음을 발견하고, 상기 전구체 화합물을 새로운 생합성 도구로 제공하기에 이르렀다. 이를 통하여 대량 유전자 클러스터의 도입에 따른 방선균의 성장 저해도 방지할 수 있다.

또 하나의 양태로서, 본 발명은 상기 전구체 화합물을 포함하는 에포틸론 생산용 배지 조성물을 제공한다.

배지 조성물은 세포가 배양되기 위하여 필요로 하는 물질을 포함하는 조성물이다. 본 발명의 목적 상, 상기 배지조성물은 상기 화학식 3에 따른 에포틸론의 전구체 화합물을 포함하며, 미생물이 상기 전구체 화합물을 에포틸론의 생합성 경로에서 기질로 사용함으로써 에포틸론 생합성 수율을 높이는 효과가 있다. 상기 미생물은 에포틸론 생합성 유전자를 포함하는 미생물이면 제한되지 않으나, 바람직하게는 방선균, 더욱 바람직하게는 스트렙토마이시스 베네주엘레 또는 스트렙토마이시스 실라칼라일 수 있다.

상기 배지 조성물에 포함되는 물질은 특별히 제한되지 아니하며 당해 기술 분야에서 통상적으로 사용되는 방법

이라면 어느 것이든 사용할 수 있다.

- [0027] 또 하나의 양태로서, 본 발명은 상기 전구체 화합물을 포함한 배지에, 에포틸론 생합성 유전자 *epoD*, *epoE*, *epoF*, *orf6*, *orf3*, 및 *orf14*가 도입된 방선균을 배양하는 단계를 포함하는, 수율이 향상된 에포틸론의 생산 방법을 제공한다.
- [0028] 하나의 구체예로서, 상기 방선균은 스트렙토마이시스 베네주엘레 또는 스트렙토마이시스 실라칼라일 수 있다.
- [0029] 또 하나의 구체예로서, 상기 에포틸론 생합성 유전자는 바람직하게는 소란지움 셀룰로섬(*Sorangium cellulosum*)에서 유래한 것일 수 있다. 소란지움 셀룰로섬은 기존에 에포틸론을 생산하는 점액 세균으로 이용되어 왔다. 그러나 소란지움 셀룰로섬은 에포틸론의 수율이 낮을 뿐만 아니라 성장시간(더블링 타임)이 길며 유전자 조작의 곤란으로 인하여 유용성이 떨어진다는 단점이 있었다. 또한 에포틸론의 유기합성방법이 보고된바 있지만, 합성방법이 복잡하다는 단점이 있다. 최근, 에포틸론 생합성 유전자 클러스터의 규명으로 인하여 대장균(*E. coli*)에서의 에포틸론의 이중생산이 연구되고 있으나, 생산되는 에포틸론의 양이 채 1 μ g/L도 되지 않으며, 낮은 온도에서(15 $^{\circ}$ C) 장기간 배양(6일)이 요구되는 단점이 있었다.
- [0030] 이에 본 발명자들은 에포틸론의 생산수율을 향상시키기 위하여 이중숙주인 방선균에 본 발명에 따른 합성된 전구체 화합물을 처리하여 배양한 결과, 에포틸론의 생산수율이 향상되었음을 확인하였다.
- [0031] 또 하나의 구체예로서, 상기 에포틸론 생합성 유전자는 바람직하게는 소란지움 셀룰로섬의 유전자를 스트렙토마이시스 베네주엘레 또는 스트렙토마이시스 실라칼라의 코돈에 최적화하여 합성한 것일 수 있다. 상기 최적화된 에포틸론 생합성 유전자는 *epoD*(서열번호 5), *epoE*(서열번호 6), *epoF*(서열번호 7), *orf6*(서열번호 8), *orf3*(서열번호 9), 및 *orf14*(서열번호 10)일 수 있다.
- [0032] 에포틸론을 소란지움 셀룰로섬 또는 이중숙주인 방선균에서 생산하기 위해서는 도 1a에 도시된 것과 같은 일련의 에포틸론 생합성 유전자를 모두 필요로 하나, 본 발명에 따른 합성된 전구체 화합물을 균주에 처리할 경우, *epoD*, *epoE*, *epoF*, *orf6*, *orf3*, 및 *orf14*만 도입하였더라도 에포틸론을 생산할 수 있다.
- [0033] 또 하나의 구체예로서, 상기 에포틸론 생합성 유전자에는 *epoA*, *epoP*, *epoB*, 및 *epoC*가 추가로 포함될 수 있다. 상기 에포틸론 생합성 유전자는 바람직하게는 소란지움 셀룰로섬(*Sorangium cellulosum*)에서 유래한 것일 수 있다. 더욱 바람직하게는 소란지움 셀룰로섬의 유전자를 스트렙토마이시스 베네주엘레 또는 스트렙토마이시스 실라칼라의 코돈에 최적화하여 합성한 것일 수 있다. 상기 최적화된 에포틸론 생합성 유전자는 *epoA*(서열번호 1), *epoP*(서열번호 2), *epoB*(서열번호 3), 및 *epoC*(서열번호 4)일 수 있다.
- [0034] 또 하나의 구체예로서, 상기 에포틸론 생합성 유전자는 pDHS702, pDHS618, pSBAC, 및 pSET152로 이루어지는 군에서 선택되는 어느 하나의 벡터에 합성된 벡터의 형태로 방선균에 도입되어 에포틸론을 생산할 수 있다.
- [0035] 또 하나의 구체예로서, 본 발명에 따른 에포틸론의 생산방법이 스트렙토마이시스 베네주엘레를 이용할 경우, 상기 합성된 벡터는 pDHS702-optSoC_APBC_Sv, pDHS618-optSoC_DEF_Sv, pDHS702-optPD_APBC_Sv, pDHS618-optPD_DEF_Sv, pSBAC-optSoC_APBC_Sv, pSBAC-optSoC_DEF_Sv, pSBAC-optPD_APBCDEF_Sv, pSET152-optSoC_DEF_Sv, 및 pSET152-optPD_DEF_Sv으로 이루어지는 군에서 선택되는 어느 하나 이상의 벡터일 수 있으나, 이에 제한되는 것은 아니다. 상기 합성된 벡터의 개열지도를 도 3a 내지 도 3i에 도시하였다.
- [0036] 또 하나의 구체예로서, 본 발명의 에포틸론 생산방법은 스트렙토마이시스 베네주엘레에 (a) pDHS702-optSoC_APBC_Sv와 pDHS618-optSoC_DEF_Sv를 도입하는 단계, (b) pDHS702-optPD_APBC_Sv와 pDHS618-optPD_DEF_Sv를 도입하는 단계, (c) pSBAC-optSoC_APBC_Sv와 pSBAC-optSoC_DEF_Sv를 도입하는 단계, (d) pSBAC-optPD_APBCDEF_Sv를 도입하는 단계, (e) pDHS618-optSoC_DEF_Sv를 도입하는 단계, (f) pDHS618-optPD_DEF_Sv를 도입하는 단계, (g) pSBAC-optSoC_DEF_Sv를 도입하는 단계, (h) pSET152-optSoC_DEF_Sv를 도입하는 단계, (i) pSET152-optPD_DEF_Sv를 도입하는 단계를 포함하는 단계 중 어느 하나의 단계를 포함할 수 있다.
- [0037] 또 하나의 구체예로서, 본 발명에 따른 에포틸론의 생산방법이 스트렙토마이시스 실리칼라를 이용할 경우, 상기 합성된 벡터는 pDHS702-optSoC_APBC_Sc 또는 pDHS618-optSoC_DEF_Sc일 수 있으나, 이에 제한되는 것은 아니다. 상기 합성된 벡터의 개열지도를 도 4a 및 도 4b에 도시하였다.
- [0038] 또 하나의 구체예로서, 본 발명의 에포틸론 생산방법은 스트렙토마이시스 실리칼라에 pDHS702-optSoC_APBC_Sc

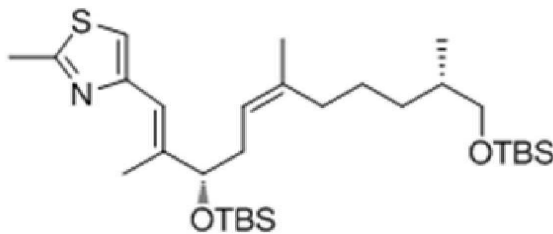
또는 pDHS618-optSoC_DEF_Sc를 도입하는 단계를 포함할 수 있다.

[0039] 상기 합성된 벡터는 소란지움 셀룰로섬의 유전자간 영역, 또는 이중속주로 사용되는 스트렙토마이시스 베네주엘라의 피크로마이신 또는 데소사민 유전자 집단의 유전자간 영역을 포함할 수 있다. 상기 유전자간 영역(intergenic region)은 유전자와 유전자 사이에 위치하는 DNA서열로, 비해독성(noncoding) 부분을 의미한다 (Tropp, Burton E. (2008). Molecular Biology: Genes to Proteins). 이는 인트라제닉 영역(intragenic region)인 인트론(intron)과는 차이가 있다. 기능이 없는 것으로 알려져 있으나, 발현 유전자의 특이적인 발현 증대를 위한 유전자간 영역이 어느 부분인지는 전혀 보고된 바 없다.

[0040] 또 하나의 양태로서, 본 발명은 상기 전구체 화합물의 제조방법을 제공한다.

[0041] 상기 제조방법은 하기 화학식 4로 표시되는 스킨저 중간체(Schinker's intermediate, *Chem. Eur. J.* 1999, 5, 2492-2500)에 옥시데이션(oxidation), 디프로텍션(deprotection), 및 티오에스터(thioester) 합성 반응을 가하는 단계를 포함할 수 있다.

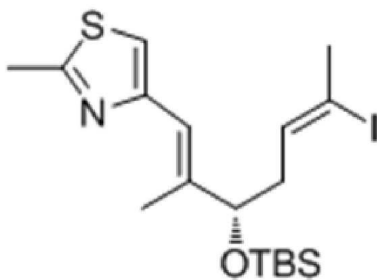
화학식 4



[0042]

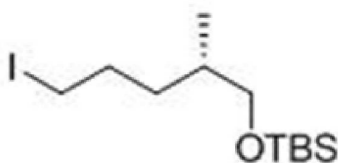
[0043] 하나의 구체예로서, 상기 스킨저 중간체는 하기 화학식 5로 표시되는 요오드화비닐(vinyl iodide) 및 하기 화학식 6으로 표시되는 요오드화물로부터 합성될 수 있다.

화학식 5



[0044]

화학식 6



[0045]

[0046] 또 하나의 구체예로서, 상기 제조방법은 (a) 티오아세트아미드(thioacetamide, 1)로부터 합성된 포스페이트 에스터(phosphate ester, 4)를 합성하는 단계; (b) L-말산(L-malic acid, 5)으로부터 케톤(ketone, 11)을 합성하는 단계; (c) 상기 포스페이트 에스터(phosphate ester, 4) 및 상기 케톤(ketone, 11)로부터 요오드화비닐(vinyl iodide, 15)을 합성하는 단계; (d) (R)-벤질-옥사졸리디논((R)-Benzyl-oxazolidinone, 16)으로부터 제1요오드화물(21)을 합성하는 단계; (e) 상기 요오드화비닐(vinyl iodide, 15) 및 제1요오드화물(21)로부터 N-아세틸 시스타민 티오에스터(N-acetyl cysteamine(SNAC) Thioester, 25)를 합성하는 단계를 포함한다. 상기 제조방법은 도 2a 및 도 2b에 도시하였다.

발명의 효과

[0047] 본 발명에 따른 합성된 전구체, 및 이를 이용한 에포틸론의 생산 방법은 방선균에서의 에포틸론 생산 수율을 향상시킬 수 있다.

[0048] 또한, 본 발명에 따른 합성된 전구체, 및 이를 이용한 에포틸론의 생산 방법은 *epoA*, *epoP*, *epoB*, 및 *epoC*가 도입되지 않은 방선균에서도 에포틸론을 생산할 수 있다.

도면의 간단한 설명

[0049] 도 1a은 에포틸론 폴리케타이드 생합성 유전자(PKS)의 모듈을 나타낸 것이며, 도 1b는 에포틸론 생산을 위한 전구체 화합물 공급의 개념도를 나타낸 것이다.

도 2a는 본 발명에 따른 전구체 화합물의 합성반응을 나타낸 것이며, 도 2b는 상기 도 2a를 요약하여 나타낸 것이다.

도 3a 내지 도 3i는 본 발명에 따른 에포틸론 생산방법에 사용되는, 스트렙토마이시스 베네주엘라에 도입하기 위한 벡터의 개열지도를 나타낸 것이다.

도 4a 및 도 4b는 본 발명에 따른 에포틸론 생산방법에 사용되는, 스트렙토마이시스 실리칼라에 도입하기 위한 벡터의 개열지도를 나타낸 것이다.

도 5는 본 발명에 따른 에포틸론 생산방법에 따라 생산된 에포틸론 D를 나타낸 것이다(ESI-MS/MS).

발명을 실시하기 위한 구체적인 내용

[0050] 이하 본 발명을 하기 예에 의해 상세히 설명한다. 다만, 하기 예는 본 발명을 예시하기 위한 것일 뿐, 하기 예에 의해 본 발명의 범위가 제한되는 것은 아니다.

[0051] 균주, 플라스미드, 배양조건

[0052] 본 발명에서 서브-클로닝(subcloning)을 위하여 대장균(*E. coli*) DH5(Invitrogen), 플라스미드 pLitmus28(New England Biolabs)을 사용하였다. 본 발명에 사용된 대장균 DH5는 서브클로닝에 널리 이용되는 가장 잘 알려진 균주 중에 하나이다. 이 균주는 *redAI* 과 *endAI* 유전자의 변이로 인하여 클로닝된 내부 유전자(insert)의 안정성과 플라스미드의 획득 수율을 높여주는 장점이 있다. 한편, 플라스미드 pLitmus28은 3kb 이하의 크기를 가지며, 암피실린(ampicillin)에 저항성을 가지고 다양한 개열지도(cleavage map)를 가지고 있어서 서브클로닝에 널리 이용되는 플라스미드이다.

[0053] 에포틸론을 스트렙토마이세스 베네주엘라에서 발현시키기 위한 대장균-스트렙토마이시스 셔틀(shuttle) 벡터로 pDHS702, pDHS618 또는 pSBAC를 사용하였다. 상기 벡터들은 *pikA* I 프로모터와 각각 티오스트렙톤(thiostrepton) 및 아프라마이신(apramycin) 내성 유전자를 포함하고 있다. 상기 셔틀벡터 pDHS702와 pDHS618은 미국 미시간 대학교 미생물학과 데이비드 셔먼(David H. Sherman) 교수로부터 확보하였다.

[0054] 또한 스트렙토마이시스 실리칼라에서 에포틸론 유전자를 발현하기 위해서 pDHS702와 pDHS618에서 *pikA* I 프로모터를 *act* 프로모터로 교체하여 발현하였다.

[0055] 조작된 돌연변이 균주인 스트렙토마이시스 베네주엘라 DHS2001은 한국 등록특허 제10-0949313호에 언급된 균주

를 사용하였다.

[0056] 대장균은 플라스미드의 선별을 위하여 암피실린 (ampicillin, 50 $\mu\text{g/mL}$) 또는 아프라마이신 (apramycin, 25 $\mu\text{g/mL}$)이 첨가된 LB배지에서 배양하였다.

[0057] 실시예 1: 에포틸론 생합성 유전자 합성 및 발현벡터 제조

[0058] 에포틸론 생합성 유전자는 소란지움 셀룰로섬의 유전자 정보를 바탕으로 이중숙주 스트렙토마이시스 베네주엘라의 코돈 선호도에 맞추어 바이오니아에서 ~2kb의 사이즈로 합성하였다. 합성된 에포틸론 생합성 유전자는 *epoA* (서열번호 1), *epoP*(서열번호 2), *epoB*(서열번호 3), *epoC*(서열번호 4), *epoD*(서열번호 5), *epoE*(서열번호 6), *epoF*(서열번호 7), *orf6*(서열번호 8), *orf3*(서열번호 9), 및 *orf14*(서열번호 10)이다.

[0059] 상기 합성된 유전자 단편들은 각각의 제한효소를 이용하여 서브클로닝 단계를 거쳐 결합시켜 pDHS702, pDHS618 및 pSBAC에 도입하였다. 서브클로닝 및 도입 방법은 한국 등록특허 제10-0949313호에 기재된 방법을 일부 변형하여 사용하였다.

[0060] 이를 통하여 도 3a 내지 도 3i에 도시된 개열지도와 같이 벡터 pDHS702-optSoC_APB_C_Sv, pDHS618-optSoC_DEF_Sv, pDHS702-optPD_APB_C_Sv, pDHS618-optPD_DEF_Sv, pSBAC-optSoC_APB_C_Sv, pSBAC-optSoC_DEF_Sv, pSBAC-optPD_APB_CDEF_Sv, pSET152-optSoC_DEF_Sv, 및 pSET152-optPD_DEF_Sv를 완성하였다. 또한, 도 4a 및 도 4b에 도시된 개열지도와 같이 벡터 pDHS702-optSoC_APB_C_Sc, 및 pDHS618-optSoC_DEF_Sc를 완성하였다.

[0061] 실시예 2: 에포틸론 전구체 thioesterSNAC의 유기합성

[0062] 도 2a 및 도 2b에 따라 본 발명의 전구체 화합물, 즉 에포틸론 N-아세틸 시스타민 티오에스터(Epothilone N-acetyl cysteamine (SNAC) Thioester, 25)를 합성하였다.

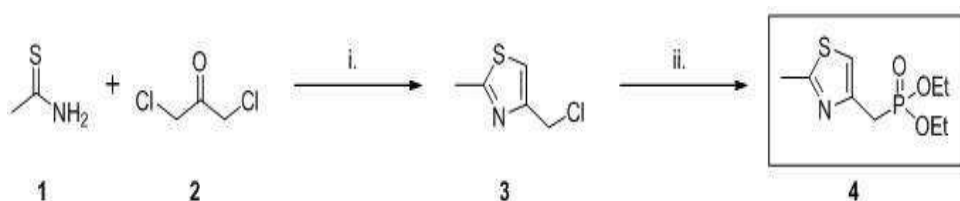
[0063] (1) 스킨저 중간체에 의한 합성

[0064] 먼저, 탄소-탄소 커플링 반응 중 하나인 네기시 (Negishi) 커플링 반응을 진행하여 스킨저 중간체(Schinzer's intermediate, 22; Chem. Eur. J. 1999, 5, 2492-2500)를 준비하였다. 그 후, 상기 스킨저 중간체(22)로부터 옥시데이션(oxydation), 디프로텍션(deprotection), 및 티오에스터(thioester) 반응을 가하여, 상기 방법을 이용하여 원하는 물질인 본 발명의 전구체 화합물(25)을 합성하였다.

[0065] (2) 티오아세트아미드 및 L-말산으로부터의 합성

[0066] 새로운 탄소-탄소 결합을 형성하기 위해서 필요한 반응 커플링 파트너인 요오드화비닐(vinyl iodide, 15)과 제1 요오드화물(primary iodide, 21)을 준비하기 위해서, 포스페이트 에스터(phosphate ester, 4; JOC 2004, 69, 9269)와 케톤(ketone, 11; Chem. Eur. J. 1999, 5, 2492-2500)을 각각 티오아세트아미드(thioacetamide, 1), L-말산(L-malic acid, 5), R)-벤질-옥사졸리디논((R)-Benzyl-oxazolidinone, 16)을 이용하여 본 발명의 전구체 화합물을 합성하였다(도 2a). 구체적인 합성단계는 아래와 같다.

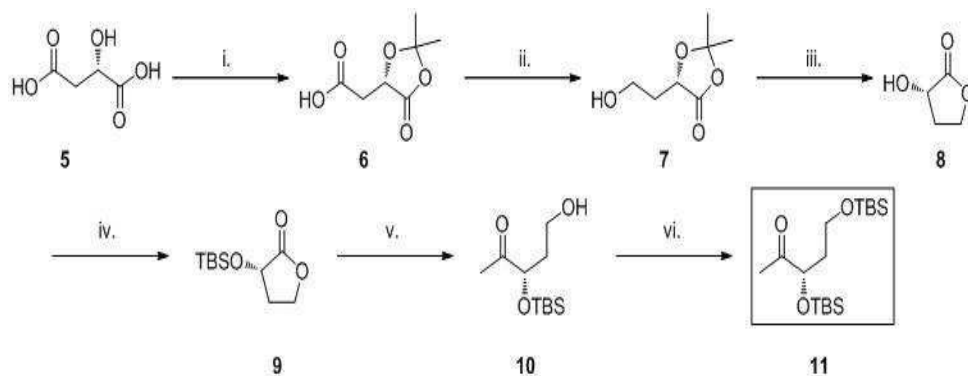
[0067] 합성단계 (a)



i. 에탄올, 55℃, 7 h, 80% (lit 81%); 및

ii. 트리에틸포스파이트(triethylphosphite), Δ , 6 h, 86% (lit 86%))

합성단계 (b)



i. 디메톡시프로판(dimethoxypropane), pTSA (1 mol%), RT, 3 h, 75% (lit 84%);

ii. BH₃-THF, THF, 0℃, 1.75 h, 이후 RT, overnight;

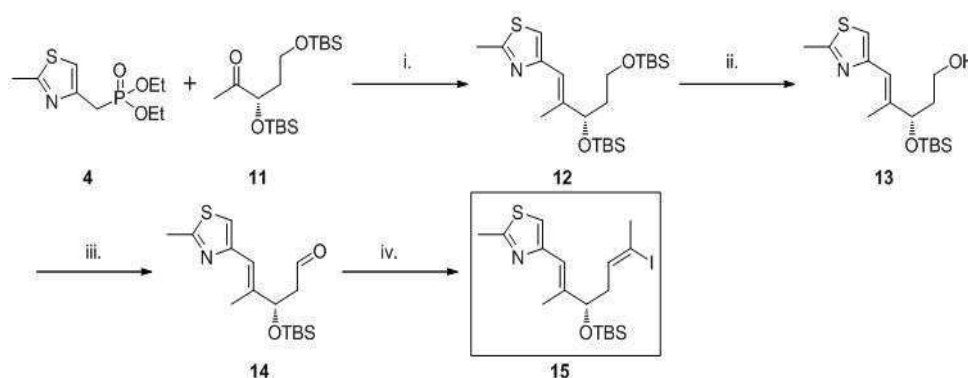
iii. pTSA, CH₂Cl₂, RT, 밤새(overnight), 71%, 2회 이상 (lit 85% over two steps);

iv. TBSCl, 이미다졸(imidazole), DMF, RT, 밤새, 88% (lit 93%);

v. MeLi, THF, -78℃, 3.5 h; 및

vi. TBSCl, imidazole, DMF, RT, 밤새, 88% over two steps (lit 80% over two steps)

합성단계 (c)



i. nBuLi, THF, -78℃, 1 h, 이후 케톤(ketone), -78 내지 실온(RT)에서 밤새도록, 73%. with 18% recovery of starting material (lit 79% with 22% recovery);

ii. THF:TFA:H₂O (4:1:1), 0℃, 35 min, 94% (lit 92% HF-Pyr);

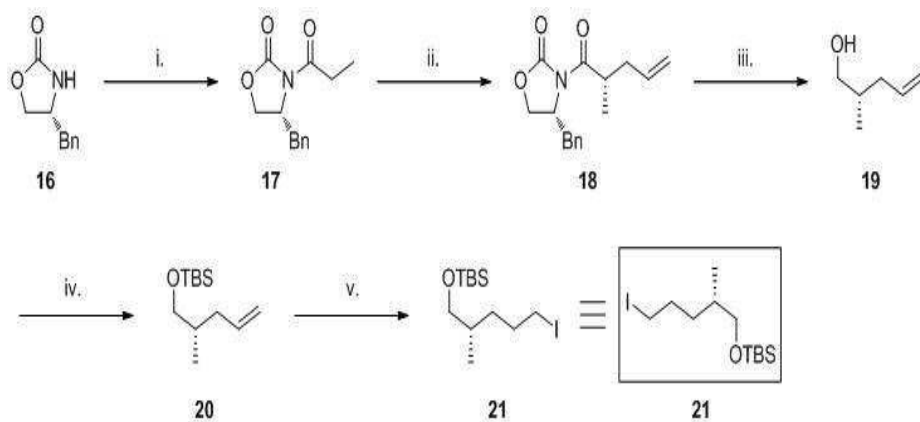
iii. DMSO, 옥살릴 클로라이드(oxalyl chloride), CH₂Cl₂, -78℃, 10 min, 이후 알코올(alcohol), 30 min, 이후 NEt₃, -78℃ 내지 RT, 30 min, 92% (lit 82%); 및

iv. 요오드화 에틸 트리페닐포스포늄(Ethyltriphenylphosphonium iodide), nBuLi, THF, 0℃ to RT, 20 min, 이후 I₂, THF에 첨가, -78℃, 15 min, 이후 -30℃, 30 min, 이후 NaHMDS, -30℃, 15 min, 이후 알데히드

(aldehyde), -30°C , 10 min, 53% (lit 54%)

[0085]

합성단계 (d)



[0086]

[0087]

i. nBuLi, THF, -78°C , 5 min, 이후 프로피오닐 클로라이드(propionyl chloride), -78°C , 30 min, 이후 -78°C 내지 RT, 97% (lit 91 96%);

[0088]

ii. NaHMDS, THF, -78°C , 1 h, 이후 알릴이오다이드(allyl iodide), 4 h, 이후 -78°C to RT, 95% 24:1 d.r. (lit 61%, one isomer);

[0089]

iii. LAH, Et₂O, 0°C to RT, 6.5 h;

[0090]

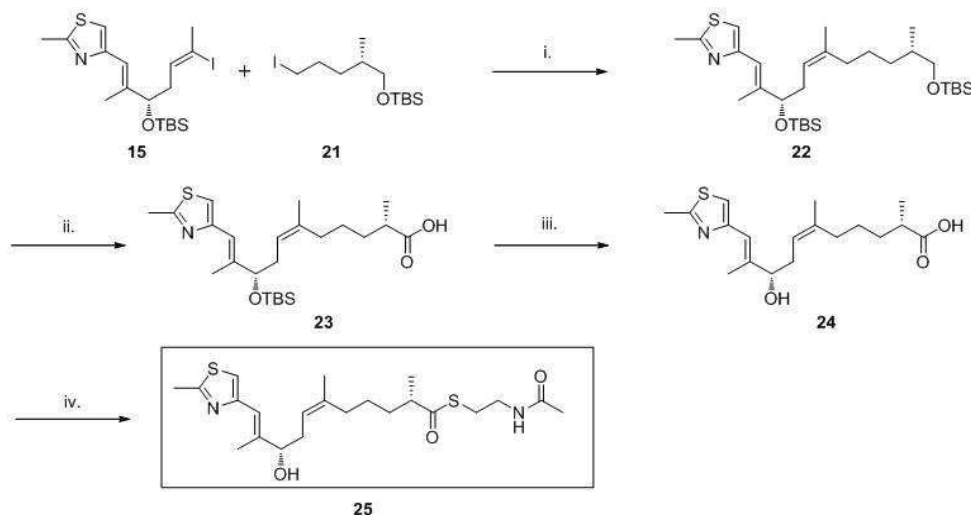
iv. TBSCl, 이미다졸(imidazole), DMF, RT, overnight, 75% over two steps (lit 80% over two steps); 및

[0091]

v. BH₃-THF, THF, 0°C , 1 h, 이후 MeOH, NaOAc (1M MeOH), ICl (1M CH₂Cl₂), 0°C , 30 min, 48% (lit 60%)

[0092]

합성단계 (e)



[0093]

[0094]

i. Zn/Cu couple, 1,2-디브로모에탄(dibromoethane), PhH, Δ , 10 sec, TMSCl, RT, 5 min, 이후 DMAc, 제1요오드화물(primary iodide), PhH, 60°C , 2.5 h, 이후 TESOTf, DMAc, Δ , 1 h, 이후 Pd(PPh₃)₄, RT, 5 min, 이후 요오드화비닐(vinyl iodide), PhH, 60°C , 30 min, 97% (lit 84%);

[0095]

ii. PDC, DMF, RT, 3 days, 66%;

[0096]

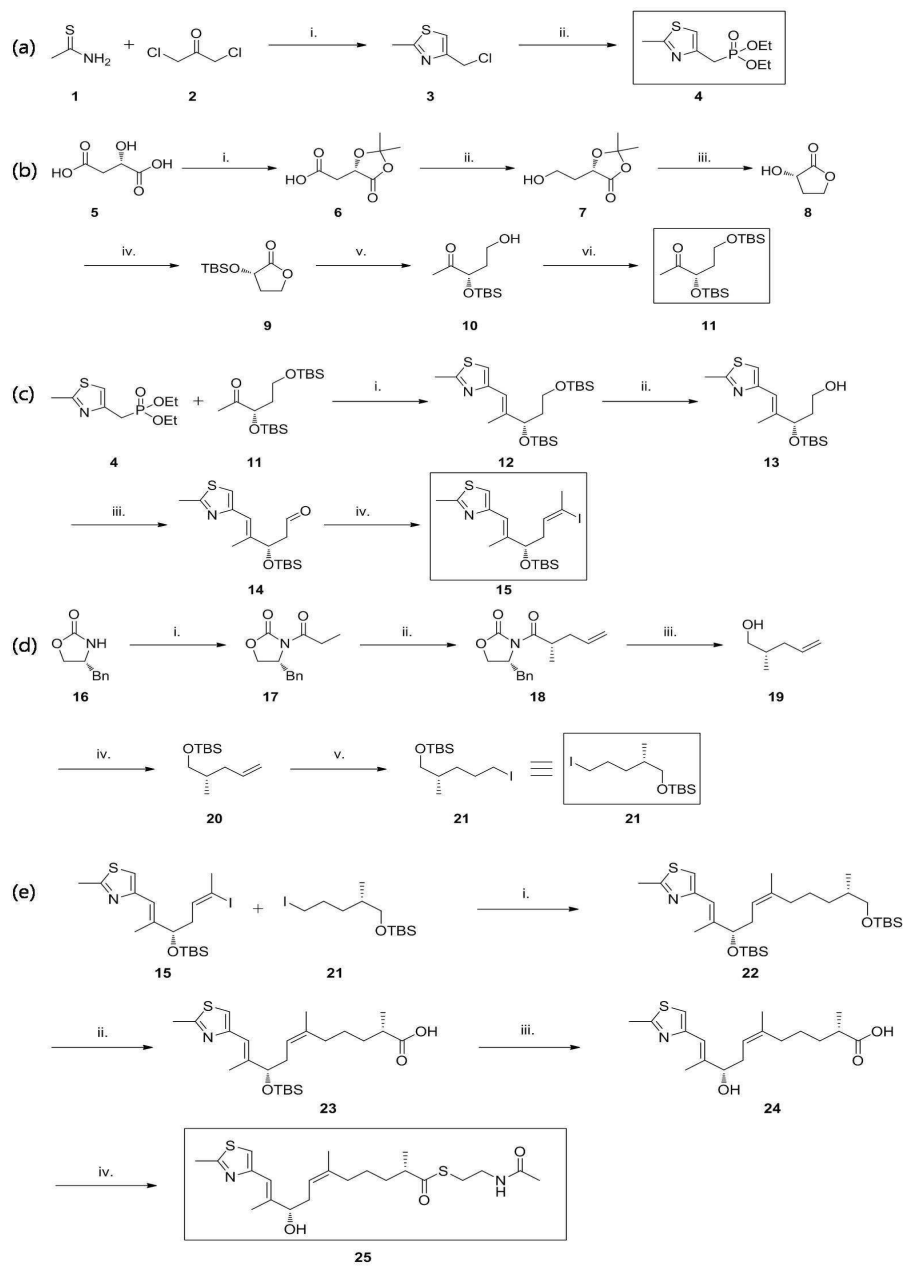
iii. TBAF, THF, RT, 36 h; 및

- [0097] iv. DMAP, HSNAC, HOBt, EDC-HCl, CH₂Cl₂, 63% over two steps
- [0098] 실시예 3: 이중숙주에서 에포틸론의 발현
- [0099] (1) 스트렙토마이시스 베네주엘레에의 형질전환
- [0100] 실시예 1에서 제조한 pDHS702-optSoC_APB_C_Sv와 pDHS618-optSoC_DEF_Sv를 스트렙토마이시스 베네주엘레 DHS2001로 도입하고, 재조합된 균주를 아프라마이신 (25μg/mL), 티오스트렙톤 (50μg/mL) 및 카나마이신 (kanamycin) (50μg/mL)를 함유한 R2YE 배지를 이용하여 30℃에서 4일간 배양하였다. R2YE 조성은 103g sucrose (sigma), 0.25g K₂SO₄ (sigma), 10.12g MgCl₂·6H₂O (sigma), 10g glucose (sigma), 0.1g casamino acids (Difco), 5g yeast extract (Difco), 5.73g TES buffer (sigma), 2mL trace element solution, 22g agar (1L 기준)이며, 고온 멸균 (autoclave) 후에 10mL 0.5% KH₂PO₄ (sigma), 4mL 5M CaCl₂·2H₂O (sigma), 15mL L-proline(sigma), 7mL 1N NaOH(sigma) (1L 기준)을 첨가하여 제조하였다.
- [0101] 또한 벡터 pDHS702-optPD_APB_C_Sv와 벡터 pDHS618-optPD_DEF_Sv의 세트, 또는 벡터 pSBAC-optSoC_APB_C_Sv와 벡터 pSBAC-optSoC_DEF_Sv의 세트를 각각 스트렙토마이시스 베네주엘레 DHS2001 균주로 도입하고 상기와 동일한 방법으로 배양하였다.
- [0102] 또한, pSBAC-optPD_APB_C_DEF_Sv, pDHS618-optSoC_DEF_Sv, pDHS618-optPD_DEF_Sv, pSBAC-optSoC_DEF_Sv, pSET152-optSoC_DEF_Sv, 또는 pSET152-optPD_DEF_Sv를 각각 스트렙토마이시스 베네주엘레 DHS2001 균주로 도입하고 상기와 동일한 방법으로 배양하였다.
- [0103] 또한, pDHS702-optSoC_APB_C_Sc와 pDHS618-optSoC_DEF_Sc를 스트렙토마이시스 실리칼라 M1154에 도입하고 상기와 동일한 방법으로 배양하였다.
- [0104] 또한, pDHS618-optSoC_DEF_Sc를 스트렙토마이시스 실리칼라 M1154에 도입하고 상기와 동일한 방법으로 배양하였다.
- [0105] (2) 스트렙토마이시스 베네주엘레 DHS2001에 전구체 화합물 처리
- [0106] 상기 합성된 벡터가 도입된 스트렙토마이시스 베네주엘레 DHS2001 균주들은 항생제를 섞은 R2YE의 액체배지 또는 고체배지에서 30℃의 온도에서 3일간 배양하였다. 실시예 2에 따라 합성된 전구체 화합물의 공급은 1) 1일째 한번만 공급하는 방법 및 2) 1일째와 2일째 두 번 공급하는 방법을 사용하였으며, 1 내지 4 mM를 공급하였다.
- [0107] (3) 스트렙토마이시스 실리칼라 M1154에 전구체 화합물 처리
- [0108] 상기 합성된 벡터가 도입된 스트렙토마이시스 실리칼라 M1154 균주들은 항생제를 섞은 R2YE의 액체배지 또는 고체배지에서 30℃의 온도에서 6일간 배양하였다. 그리고 실시예 2에 따라 합성된 전구체 화합물의 공급은 1) 2일째 한번만 공급하는 방법 및 2) 2일째와 4일째 두 번 공급하는 방법을 사용하였으며, 1 내지 5 mM를 공급하였다.
- [0109] (4) 에포틸론의 생산여부 확인
- [0110] 상기 전구체 화합물이 공급되어 배양된 균주가 포함된 배양 용액을 동량의 에틸아세테이트(ethyl acetate)와 혼합하고, 분별깔때기로 층 분리를 2회 반복하고, 유기층을 진공회전농축기 (rotary evaporator)로 농축하였다. 얻어진 용출물을 100μL의 메탄올 (methanol)로 다시 녹이고 상기 용액을 Phenomenex Synergi Polar-RP 컬럼 (1504.6mm, 4μm)을 장착한 LC-ESI-MS/MS (액체크로마토그래피/전자스프레이 이온화-질량분석기; Waters/Micromass QuattroMicro/MS)로 분석하였다.
- [0111] 분석조건은 250μL/min의 유속으로 용출시켰고, 0.1%(v/v) 포름산 (formic acid) 이 첨가된 증류수 용액(A)과 0.1%(v/v)의 포름산이 첨가된 80%(v/v) 아세토니트릴 (acetonitrile) 용액(B)를 이용하여 수행하였다. B를 30%

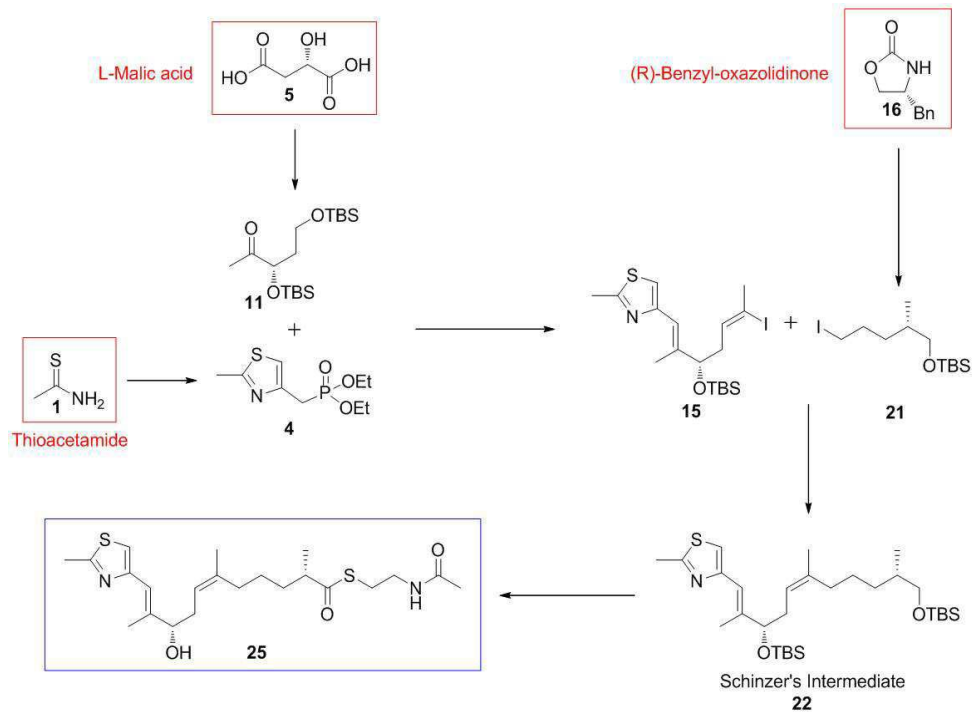
에서 50%로 5분간, 50%에서 90%로 3분간 적용하고, 90%에서 4분간 유지시킨 후 30%로 15분간의 구배로 컬럼의 재평형을 위하여 용출시켰다.

- [0112] 분석한 결과, 35분의 용출물이 에포틸론 D이며, 에포틸론 표준물질로 동일하게 실험한 결과 용출시간 및 분획 형태가 동일함을 확인하였다(도 5).
- [0113] 특히, 에포틸론 생합성 유전자 클러스터 전체를 도입하지 않고, 벡터 pDHS618-optPD_DEF_Sv만 도입되어 형질전환된 스트렙토마이시스 베네주엘레의 경우, 에포틸론 B, D가 생산되는 것을 확인하였다. 특히 에포틸론 D는 약 10 μ g 생산되는 것을 확인하였다.
- [0114] 이는 에포틸론 생합성 유전자 클러스터 전체가 도입되지 않고 epoD 이후의 유전자 클러스터만 도입된 방선균에, 본 발명에 따른 에포틸론의 전구체 화합물, 또는 이를 포함하는 배지 조성물을 처리할 경우 에포틸론 B, D를 생산할 수 있음을 확인한 것이다.
- [0115] 또한, pDHS702-optPD_APB_C_Sv와 pDHS618-optPD_DEF_Sv가 도입된 스트렙토마이시스 베네주엘레에 본 발명에 따른 에포틸론의 전구체를 공급해준 경우 에포틸론 B, D가 생산되는 것을 확인하였다. 특히 에포틸론 D는 약 10 μ g 생산되는 것을 확인하였다.
- [0116] 또한, pDHS702-optSoC_APB_C_Sc와 pDHS618-optSoC_DEF_Sc가 도입된 스트렙토마이시스 실라칼라에 본 발명에 따른 에포틸론의 전구체를 공급해준 경우 에포틸론 B, D가 생산되는 것을 확인하였다. 또한, 상기와 같은 액체 배양이 아닌 고체배양으로 고체배지를 제조하여 pDHS702-optPD_APB_C_Sv와 pDHS618-optPD_DEF_Sv가 도입된 스트렙토마이시스 베네주엘레에서도 에포틸론 B, D가 생산되는 것을 확인하였다.
- [0117] 이와 같이, 본 발명에 따른 합성된 전구체, 및 이를 이용한 에포틸론의 생산 방법은 방선균에서의 에포틸론 생산 수율을 향상시킬 수 있으며, *epoA*, *epoP*, *epoB*, 및 *epoC*가 도입되지 않은 방선균에서도 에포틸론을 생산할 수 있다는 장점이 있다.
- [0118] 이상의 설명으로부터, 본 발명이 속하는 기술분야의 당업자는 본 발명이 그 기술적 사상이나 필수적 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 이와 관련하여, 이상에서 기술한 실시 예들은 모든 면에서 예시적인 것이며 한정적인 것이 아닌 것으로서 이해해야만 한다. 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허 청구범위의 의미 및 범위 그리고 그 등가 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

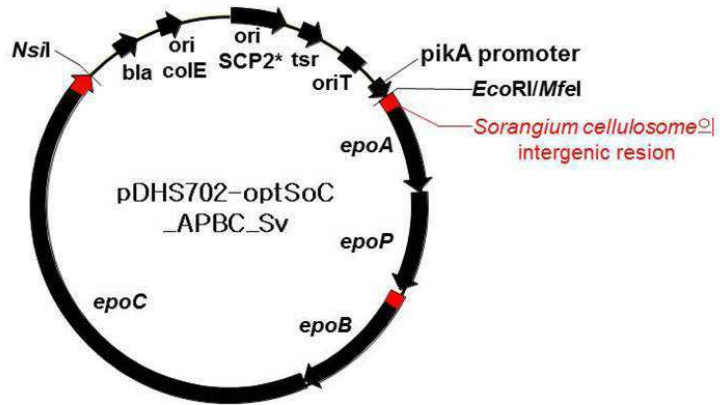
도면2a



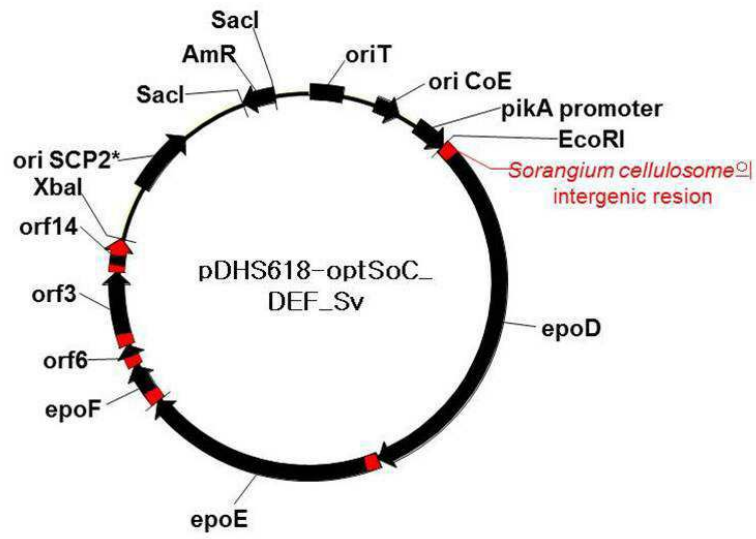
도면2b



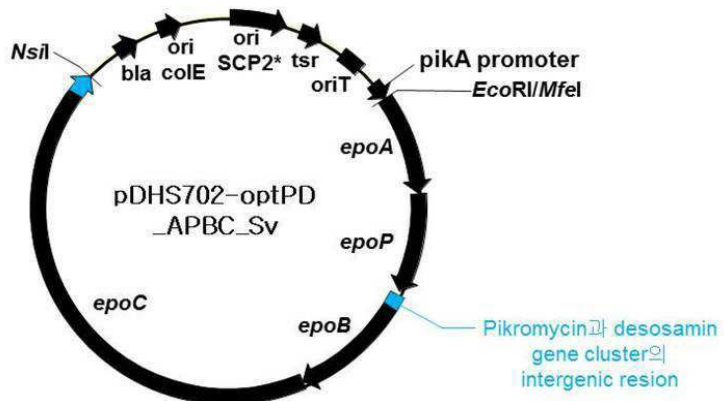
도면3a



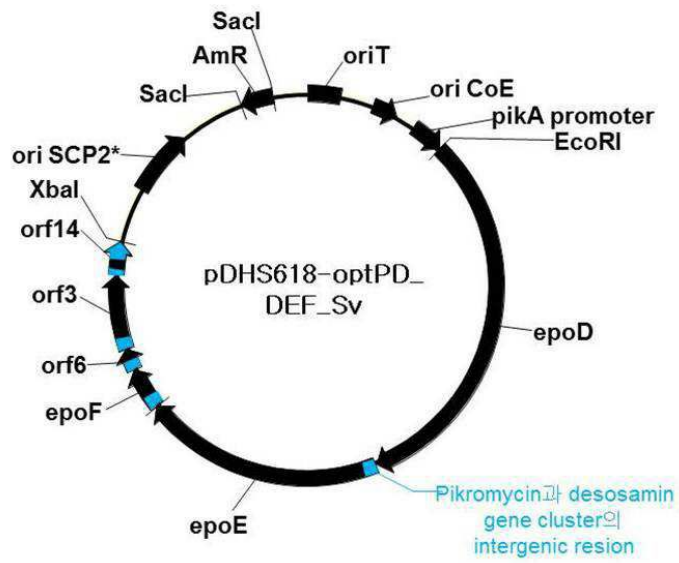
도면3b



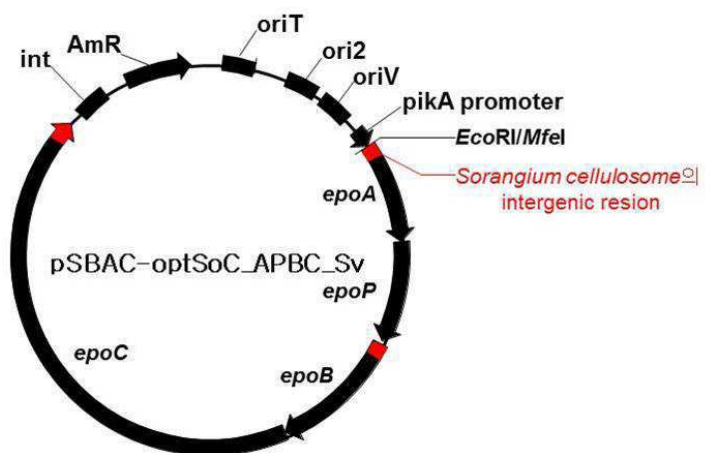
도면3c



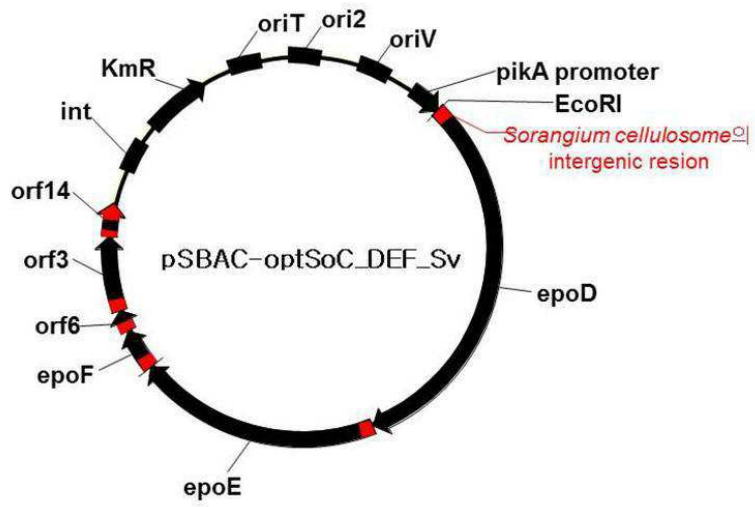
도면3d



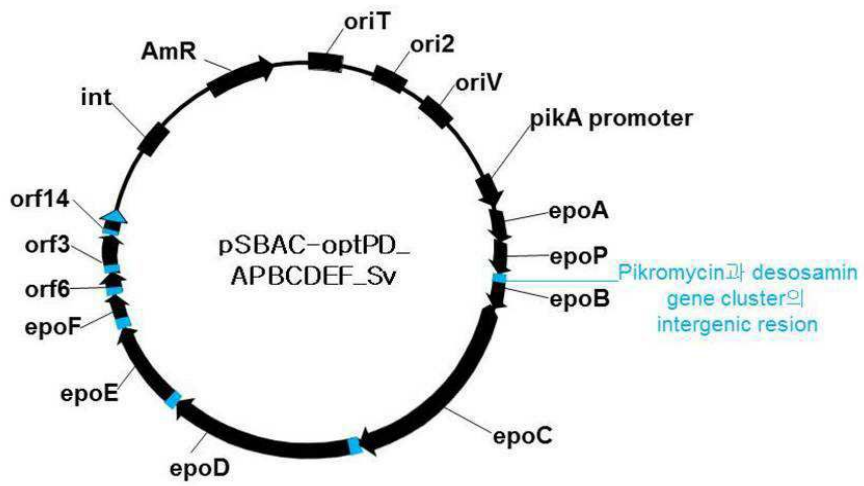
도면3e



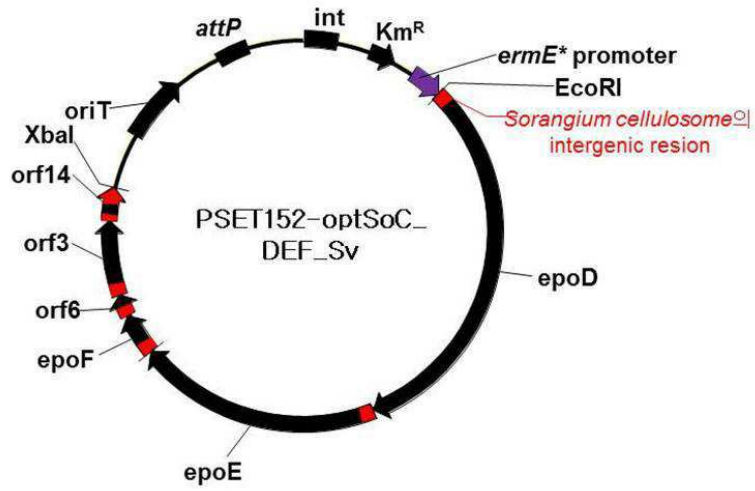
도면3f



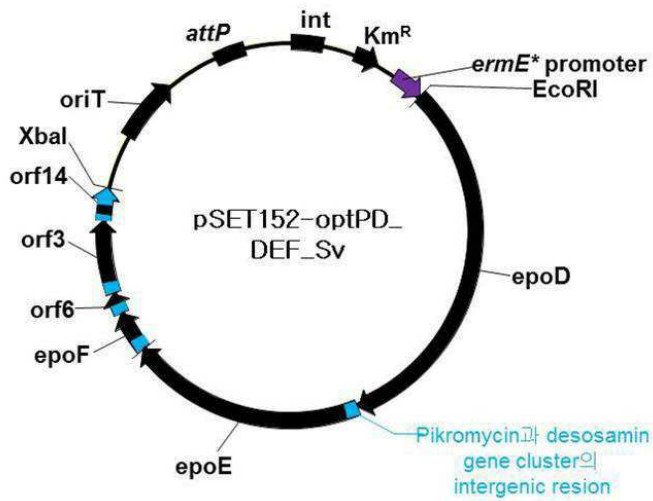
도면3g



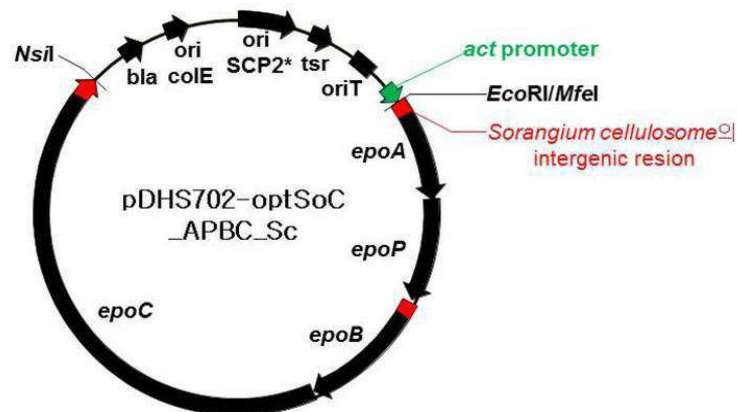
도면3h



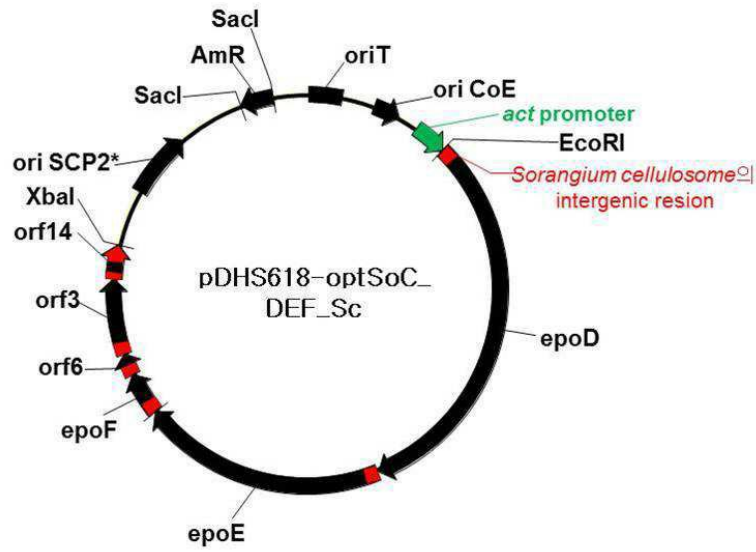
도면3i



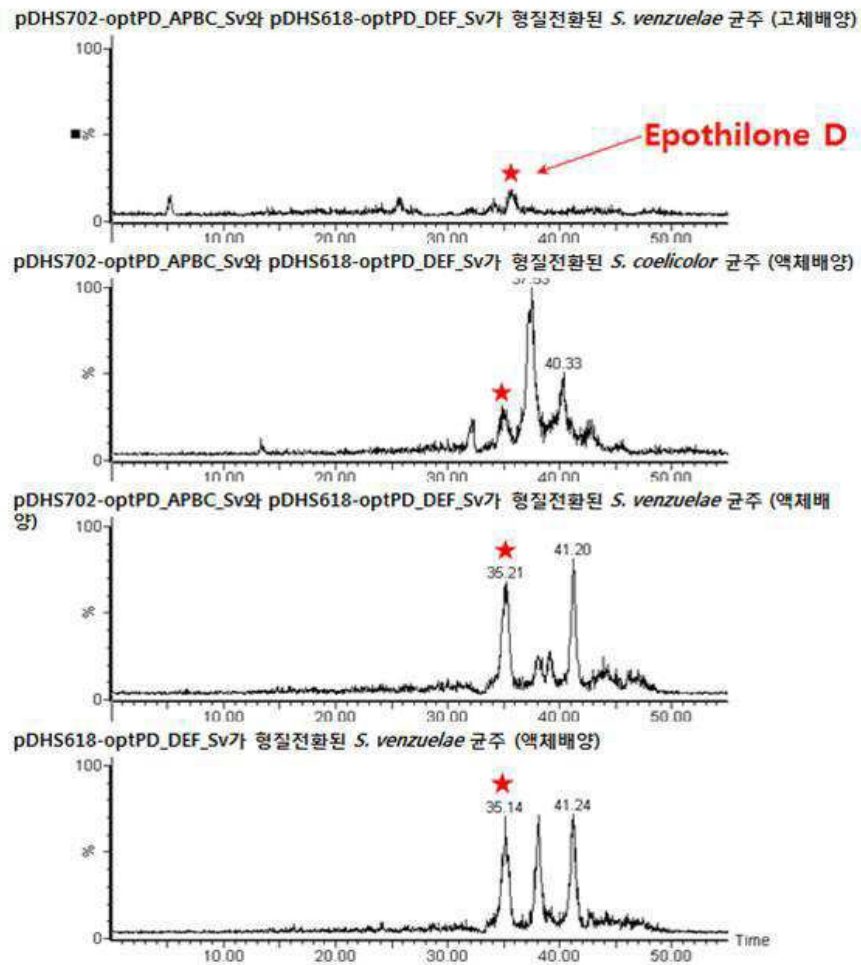
도면4a



도면4b



도면5



서열목록

<110> EWHA UNIVERSITY - INDUSTRY COLLABORATION FOUNDATION

<120> SYNTHETIC PRECURSOR OF EPOTHILONE FOR IMPROVING PRODUCTION OF EPOTHILONE AND METHOD FOR PREPARING EPOTHILONE USING THE SAME

<130> KPA140463-KR

<160> 10

<170> KopatentIn 2.0

<210> 1

<211> 4266

<212> DNA

<213> Artificial Sequence

<220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in
Streptomyces venezuelae

<400> 1

atggccgacc gcccgatcga gcgcgccgcc gaggaccgga tcgccatcgt cggagcctcc 60

tgccgcctcc cgggcggcgt catcgacctc tccggcttct ggaccctcct cgaaggttcc 120

cgcgacaccg tcggccgcgt cccggccgaa cgctgggacg ccgctgcctg gttcgaccg 180

gaccccgacg ccccgggcaa gacccgggtc acccgcgctt ctttctcag tgacgtcgcc 240

tgtctcgacg catccttctt cggatatctc ccgcgcgagg cgctccgcat ggaccgggcc 300

caccgtctcc tctctgaggt ctgctgggag gccctcgaga acgccgcat cgcctcgtcc 360

gccctcgtcg gcacggagac cggcgtcttc atcggaatcg gcccgagcga gtacgaggcc 420

gccctcccgc agggccaccg ctccgccgag atcgacgccc acggcgggtct cggcaccatg 480

ccctccgtcg gcgcgggccc catcagctac gccctcgccc tccgcggccc gtgcgtcgcc 540

gtcgataccg cctactctc ctcctcgtc gccgtgcacc tcgctgcca gtcctcctgt 600

tccggcgagt gctccaccgc cctcgccggc ggcgtctccc tcatgtctc cccgtccacc 660

ctcgtctggc tctccaagac ccgcgcctc gcccgcgacg gccgtgcaa ggccttctcc 720

gccgaagccg acggcttcgg tcgcggtgag ggctgcgccg tcgtcgtcct gaagcgctc 780

tccggcgccc gggccgacgg cgaccgcatc ctgcctgca tccgcggctc cgccatcaac 840

cacgacggag ccagctccgg gctcaccgtc ccgaacggct cctcccagga gatcgtctc 900

aagcgcgccc tcgccgacgc cggtcgcgcc gcctcctcgg tcggctacgt cgaggcgcac 960

ggcaccggga cgacctcgg cgaccgcatc gagatccagg ccctcaacgc cgtctacggc 1020

ctcggtcgcg acgtcgccac cccgtctc ctcggctcgg tcaagaccaa cctcgccac 1080

ccggaatacg cctccggcat caccggctc ctcaaggctg tactctcct ccagcacggc 1140

cagatcccgg cccacctcca cgcacaggcc ctcaaccccc gcatctctctg gggcgacctc	1200
cgcctcaccg tcaccgcgcg ccgcaccccg tggccggact ggaacacgcc gcgccgcgcc	1260
ggtgtctcct ccttcggcat gtcggaacg aacgcccacg tcgtcctcga ggaggccccg	1320
gcccaccct gcaccccgcc ggcccctgag cgtccggccg agtcctctgt cctctccgcg	1380
cggaccgcct ccgcctcga tgcacaggcc gcccgccctc gcgaccacct cgagacctac	1440
ccgtcccagt gcctcggcga cgtcgctttc tccctcgcca ctacgcgctc cgccatggag	1500
caccgcctcg ccgtcgcgc caccatccgt gagggcctcc gcgtgcctc cgacgtgcc	1560
gcccagggtc agacctcccc gggcgccgtc aggtccatcg ccgactctc caggggcaag	1620
cttgcccttc tcttcaccgg ccagggtgcc cagacgctcg gcatgggccc cggcctctac	1680
gacgtctggt ccgccttcg cgaggccttc gacctctcg tccgcctctt caaccaggag	1740
ctcgaccgcc cgtccgcga ggtcatgtgg gccgagccgg cctccgtcga cgcgcacctc	1800
ctcgaccaga ccgccttcac ccagccggcc ctgttcacct tcgagtacgc gctcgcggcc	1860
ctctggcgct cctggggggt cgagccggag ctcgtcgccg ggcaactcat cggcgagctc	1920
gtcgccgct gcgtcgcgg cgtctttctc ctcgaggacg ccgtcttct cgtcgccgcc	1980
cgcggtcgtc tcattcaggc cctcccgcc ggccggcgca tggcttccat cgaggccccg	2040
gaggccgacg tcgcgcgcgc cgtcgccgg caccgcccgt ccgtctccat cgcgcgcgtc	2100
aacgccccgg accaggctgt catcgccggc gccggccagc cggtcacgc aatcgccgcc	2160
gccatggcgg ccccgggggc ccgcaccaag gccctccacg tcagccacgc cttccactcc	2220
ccgtcatgg ccccgatgct cgaggccttc ggccgcgtcg ccgagtcgt ctctaccgc	2280
cgcgcgacga tcgtcctgt ctccaacctc tccgggaagg cctgtaccga cgaggctctc	2340
tccccgggat actgggtcgg ccacgcccgc gaggtcgtac gcttcgcca tggagtcaag	2400
gccctccacg ccgcaggcgc cggcaccttc gtcgaggtcg gccgaagtc caccctctc	2460
ggtctcgtcc cggcctgcat gcccgacgcc cgtccggccc tctcgcctc ctcccgccc	2520
ggccgcgacg aaccggccac cgtctctgag gcccttggcg gcctctgggc cgtcgccggc	2580
ctcgtctcct gggccggtct cttcccgctc ggccggccgc gcgtcccgt cccgacctac	2640
ccgtggcagc gcgagcgcta ctggatgac accaaggcgg acgacgccgc ccgcccgcac	2700
cgcgcgccc cgggcgcgg ccacgacgag gtcgaggagg gcggcgccgt ccgtggcggc	2760
gaccgccgt ccgccgcct cgaccacccg ccgccggaat ccggccgccc cgagaaggtc	2820
gaggccgccc gcgacaggcc gttccgactc gagatcgacg agccgggct cctcgaccac	2880
ctcgtcctcc gcgtaccga acgcccgcgc ccgggcctcg gcgaggttga gatcgccgtc	2940

gacgccgcgg gcctctcctt caacgacgtg cagctcgccc tcggcatggt cccggacgac	3000
ctccccggca agccgaaccc gccgctctc ctcggcgggg agtgcgccgg ccgcatcgtc	3060
gccgtcggcg agggcgtcaa cggcctcgtc gtcggccagc cggtcacgc actctccgcc	3120
ggcgcgttcg ccaccacgt caccaccagc gccgcactcg tcctcccgcg accgcaggcc	3180
ctctccgcca tcgaggccgc cgccatgccg gtcgcctacc tcaccgcgtg gtacgccctc	3240
gaccgcatcg cccgcctcca gccggggcaa cgcgtcctca tccacgccgc caccggcggc	3300
gtcggcctcg ccgccgtgca gtggggccag cacgtcggcg ccgaggtaca cgccaccgcc	3360
ggcacgccgg agaagcgagc ctacctgag tcctcggcg tgcgtacgt ctccgactcc	3420
cgctccgacc gcttcgtggc cgacgtccgc gcctggaccg gcggggaggg cgtcgacgtc	3480
gtctcaact cctctccgg cgagctcacc gacaagtct tcaacctct ccgtccac	3540
ggcgcgttcg tcgagctcgg caagcgcgac tgctacgccg acaaccagct cggcctccgc	3600
ccgttctcc gcaacctctc cttctccctc gtcgacctcc gcggcatgat gctcgagcgc	3660
ccggccccgtg tccgcgcct cctcgaggag ctctcggcc tcacgccgc cggtgtcttc	3720
accccgccgc cgatgccac gctcccatc gcccgctgg ccgacgcctt ccgtccatg	3780
gcccaggccc agcacctcgg caagctcgtc ttgacctcg gagaccgga ggtccagatc	3840
cgcatccga cccatgccgg cgccggcccg tccaccggcg accgcgacct cctcgaccgc	3900
ctgcctccg ccgccccgc cgcccgccc gccgcctcg aggccttct cgcacccag	3960
gtctcccagg tcctccgac cccggaaac aaggtcggag ccgaggcct cttaccgc	4020
ctcggcatgg actcctcat ggccgtcgag ctccgcaacc gcatgaagc ctccctcaag	4080
ctcaagctct ccaccactt cctctccacc tccccgaaca tcgacctct cgcacagac	4140
ctctcgacg cctggccac cgcctctcc ctcgagcgc tggtgcca gaacctccgc	4200
gccggcgtcc agaagactt cgtctctcc ggagccgacc aggactggga gatcatcgcg	4260
ctgtga	4266
<210> 2	
<211> 4233	
<212> DNA	
<213> Artificial Sequence	
<220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in	
Streptomyces venezuelae	
<400> 2	
gtgaccatca accagctcct caacgagctc gagcaccagg gtatcaagct cgcggccgac	60

ggcgagcgcc tccagatcca ggccccgaag aacgccctca acccgaacct cctcgccccgc	120
atctccgagc acaagagcac catcctcacg atgtctcgcc agcgctccc ggccggagtcc	180
atcgteccgg ccccgccga gaggcacgcc ccgtteccgc tcaccgacat ccaggagtcc	240
tactggctcg gccgcaccgg ggctttcacg gtccctccg gcatccacgc ctaccgcgag	300
tacgactgca cggacctga cgtcccgcgc ctctcccgcg ctttcgcaa ggtcgtcgcc	360
cgccacgaca tgctccgcgc ccacacctc ccggacatga tgcaggtcat cgagccgaag	420
gtcgacgccg acatcgagat catcgacctg cggggcctcg accgctccac ccgcgaggcc	480
cgctctgtct cgctccgcga cgcgatgtcc caccgcatct acgacaccga gcgcccccg	540
ctctaccacg tagtcgctt ccgcctcgac gagcgccaga cccgcctcgt cctcagcatc	600
gacctcatca acgtcgacct cggcagcctc agcatcatct tcaaggactg gctctcgttc	660
tacgaggacc cggagacgtc cctcccgttc ctcgagctgt cctaccgca ctacgtctc	720
gccctcgaat cccgcaagaa gtccgaggcg caccagaggt ccatggacta ctggaagcgc	780
cgcacgcgcg agtccccgcc gccgccgacc ctcccgatga aggccgaccc gtccacctc	840
aaggagatcc gcttcgccca caccgagcag ttggtcccggt cggactcctg gggccggctc	900
aagcgccgcg ttggcgagcg cggccttacc ccgacgggcg tcatcctcgc cgctttctcc	960
gaggtcatcg gccgtggtc cgcgtccccg cgcttcaccc tcaacatcac cctcttcaac	1020
cgctccccg tccacccccg cgtcaacgac ataaccggcg acttcacctc catggtcctg	1080
ctggacatcg acacaaccg cgacaagtcc ttcgagcagc gcgccaagcg catccaggag	1140
cagctctggg aggccatgga ccaactgcgac gtctccggca tcgaggtgca gcgcgaagcc	1200
gccccgctcc tcggcatcca gcgcggcgcc ctcttcccgg tegtctcac ctccgccctc	1260
aaccagcagg tcgtggcggt caccagctc cagcgctcgc gcaccccggt ctacacctcc	1320
accagacgc cgagctcct cctcgaccac cagctctacg aacacgacgg cgacctcgtg	1380
ctgcctggg acatcgctga cggcgtcttc ccgccggacc tcctcgacga catgctcgag	1440
gcctacgtcg tcttctccg ccgcctcacc gaggagccct ggggggagca ggtccgctgc	1500
tcctccccgc ccgccagct cgaggccgc gcctccgcca acgccacgaa cgccctctc	1560
tcgagcaca cgtccacgg cctcttcgcc gcccgctcg agcagctccc gatgcagctc	1620
gcggtcgtct ccgccgcaa gacctcacc tacgaggagc tcagccgccc ctcccgcgg	1680
ctcggggccc gcctccgca gcaggcgcc cgccgaaca cctcgtcgc ggtcgtcatg	1740
gagaagggtt gggagcaggt cgtcgccgtc ctcgcggtcc tcgagtcgg cgccgcctac	1800

gtcccgatcg acgcggacct gccggccgag cgcattccact acctcctcga ccacggcgag	1860
gtcaagctcg tcttcaccca gccgtggctc gacggcaagc tctcctggcc gccgggcac	1920
cagcgctcc tegtctccga ggccggcgtc gagggcgacg gcgaccagcc gcccatgatg	1980
ccgatccaga ccccgctcga cctcgcctac gtcatttata cctcgggctc caccggcctc	2040
ccgaaggcg tcatgatcga ccaccggcg gccgtcaaca ccatcctcga catcaacgag	2100
cgttcgaga tcggcccggg cgaccgctc ctgcacctct cctccctcag cttcgacctc	2160
tccgtgtacg acgtcttcgg gatcctgcc gccggcgga ccatcgtcgt cccggatgcc	2220
agcaagctcc gcgaccggcg ccattgggccc gagctcatcg agcgcgagaa ggtaaccgtc	2280
tggaactccg tgccggccct catgcgcatg ctctcgagc atttcgaggg ccgcccggac	2340
tccttcgccc gctccctcgg gctctccctc ctctccggcg actggatccc ggtgggcctc	2400
ccggcgagc tccaggccat ccgcccggcg gtctccgtca tctccctcgg cggcgccacc	2460
gaggcctcca tctggtccat cgggtaccgg gtccgcaacg tcgacctcag ctgggcctcc	2520
atcccgtacg ggcgcccgt ccgcaaccag acgttccag tcttcgacga agccctcgaa	2580
ccgcgcccgg tctgggtccc gggccagctc tacatcgcg gcgtggggct cgccctcggc	2640
tactggcgcg acgaggagaa gacgcgaag tcttccctc tgcacccgga gaccggcgag	2700
cgctctaca agaccggcga cctcggccgc tacctccgg acggcaacat cgaattcatg	2760
ggcccgagg acaaccagat caagctccgc gggtaccggt tcgagctcgg cgagattgag	2820
gagacctca agtccaccc gaacgtccgc gacccgtga tctcccggt cggcaacgac	2880
gctcgaaca agtctctct cgcctacgtc gtgccggagg ggacccgccc ccgcgccgcc	2940
gagcaggatg cctccctcaa gaccgagcgc atcgacgccc gcgcccacgc cgccgaagcc	3000
gacggcctct ccgacggcga gcgcgtccag ttcaagctgg cccggcacgg cctccgccc	3060
gacctcgacg gcaagccggt cgtcgatctc accggccagg acccgcgca agccggcctc	3120
gacgtctacg cccgccgccc ctccgtccgc acgttcttg aggccccgat cccgttcgtc	3180
gagttcgggc gcttctctc ctgcctctcc tccgtggaac cggacggcg caccctcccc	3240
aagttccgt acccgctcc gccgtccacc tacccggtcc agacctacgc ctacgtgaag	3300
tccggccga tcgagggcgt cgacgaagcg ttctactact accaccggt cgagcaccgt	3360
ctgtgaagc tctccgacca cggcatcgag cggcgcccc acgtcgacca aaacttcgac	3420
gtcttcgacg aggcggcggt caacctctc ttctcgccc gcatcgacgc catcgagtcc	3480
ttgtacggct ccagctccc cgagttctgc ctgctcgagg cgggtacat ggcccagctg	3540

ctcatggagc aggcccttc ctgaacatc ggcgtctgcc cggtcggcca gttcaacttc 3600
gagcaggatc gcccggtcct cgacctccga cactccgacg tctacgtcca cggcatgctc 3660
ggcgcccgcg tcgaccccg gtagttccag gtctgcaccc tcggccagga ctctccccg 3720
cgccgcgcca ccaccccgcg cgcccccgcg ggccgcgagc agcacttcgc cgacatgctc 3780
cgcgacttcc tccgaccaa gctcccgag tacatggatc cgaccgtctt cgtcgagctc 3840
gacgccctcc cgtcacgtc caacggcaag gtcgaccga aggccctccg cgagcgcaag 3900
gacacctcct ccccgcgcca ctccggccac accgccccgc gcgacgcgt cgaggagatc 3960

ctcgtcgccg tcgtccgca ggtactcggc ctcgaggatc tcgggctgca gcagtcttc 4020
gtcgacctc ggcgacctc catccacatc gtcgcatgc gtcctcct ccagaaacgc 4080
ctcgaccgag agatcgccat caccgagctc ttccagtacc cgaacctggg ctgctcgcc 4140
tccggcctcc gcccgactc ggcgcatc gaccagcgc cgaacatgca ggaccgctc 4200
gaggtccgcc gcaaggcgcg ccgcccgtcc tga 4233

<210> 3

<211> 5499

<212> DNA

<213> Artificial Sequence

<220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in

Streptomyces venezuelae

<400> 3

atgaagagc aggagtctc cgccatgcc gtcacggga tgcggcgcg cttccgggc 60
gcccgcgacc tcgacgagtt ctggcgcaac ctccgacg gcaccgagc cgtccagcgc 120
ttctccgagc aggagctcgc cgctccggt gtcgaccg ccctcgtcct cgaccgtcc 180
tacgtccgcg ccggcagcgt ccttgaggac gtcgacgct tcgacgcgc cttcttcggc 240
atctccccgc gcgaggccga gcttatggac ccgcagcacc gcatttcat ggagtgcgcc 300
tgggaggccc tcgagaacgc cggatacgac ccgaccgct acgagggtc catcggcgtc 360

tacccggcg ccaacatgct ctctacctg accagcaacc tccagagca cccggccatg 420
atgcgtggc cgggtggtt ccagacctc atcggaacg acaaggacta cctcgccacc 480
cacgtctcct accgctgaa cctccgcggc ccgtccatct ccgtccagac ggctgttcc 540
acctccctc tcgccgtcca cctcgcatgc atgtccctc tcgaccgtga gtgcgacatg 600
gctctcgccg gcggcatcac cgccgcacg ccggtacgt ctacccgag 660
ggcgcatct tctccccgga cggccactgc cgcgcttcg acgccaaggc caacggcacc 720

atcatgggca acggatgcgg cgtcgtgctc ctcaagccgc tgcaccgcgc cctctccgac	780
ggtgaccggt tccgcgccgt catcctcggc tccgccacca acaacgacgg cggccgcaag	840
atcggttca cggccccgtc tgaggtcggc caggcccagg ccatcatgga ggccctcgcc	900
ctcggggcgc tgaagcccg ctccatccag tacatcgaga cccacggcac cggcacctc	960
ctcggcgacg ccatcgaaac cgccgccctc cgccgcgtct tcggcaggga cgctccgcc	1020
cgcaggtcct gcgccatcgg ctccgtcaag accggaatcg gccacctga gtccgcccgc	1080
ggcatcgccg gcctcatcaa gaccgtctc gccctcgaa accgccagct cccccgtcc	1140
ctcaacttcg agtccccgaa cccgtccatc gacttcgcct cctccccgtt ctactcaac	1200
acctccctca aggactgaa caccggctcc accccgcgcc gcgcggcgt ctctctctc	1260
ggcatcggcg gcaccaacgc ccacgtggtc ctcgaggagg cccccgcgc caagctccg	1320
gccgccgccc cgccccgtc cgccgagctc ttctgtgtct ccgccaagtc cgccgccgcc	1380
ctcgacgtcg ccgccgccg cctccgcgac cacctgcagg cccaccaggg gatctccctc	1440
ggcgacgtcg ccttctccct cgcgaccacc cgctccccga tggagcacg cctcgccatg	1500
gccccccga gcccgaggc cctccgcgaa ggctcgacg ccgccgccg cggccagacc	1560
ccgccgggtg ccgtccgcgg ccgtgtctcc ccgggcaacg tcccgaaggt cgtcttcgtc	1620
ttccccgggc agggttccca gtgggtcggc atgggtcgcc agtccttcgc cgaggagccg	1680
gtgttccacg ccgccctctc cgctgcgac cgcccatcc aggcgaggc cggctggtcc	1740
ctctcgccg aactcgccgc cgacgaaggc tcctccagc tcgagcgcat cgacgtcgtc	1800
cagccggtcc tcttcgcct cgccgtgcc ttccgccccc tctggcgag ctggggcgtc	1860
gccccggacg tcgtcattgg aactccatg ggcgaggtcg ccgccccca cgtcgccggc	1920
gccctctcct tggaggacgc agtcgccatc atctgccgcc gctccgcct cctccgcgc	1980
atctccggcc agggcgagat ggccgtcacc gagctgtccc tggccgaggc cgaggccgcc	2040
ctccgggtc acgaggaccg cgtctccgtc gccgtcagca actccccgc ctccaccgtc	2100
ctctccggcg agccggccgc catcggcgag gtctctcct ccctcaacgc caagggggtc	2160
ttctgccgcc gcgtaaaggc cgacgtgcc tccactccc cgcaggtcga cccgtccgc	2220
gaggacctcc tcgccccct cggcgccctc cgccggggg ccgccgccgt cccgatgcgc	2280
tccaccgtca ccggcgctat ggtcgccggc ccggaactcg gcgccaacta ctggatgaac	2340
aacctccgcc agccgggtccg ctccgccgag gtctgtccagg cccaactcca gggcgccac	2400
ggcctcttcg tcgagatgtc gccgcacccc atctcacca cctccgtcga ggagatgcgc	2460

cgcgccgccc agcgagccgg tgccgcagta ggctccctcc gccggggcca ggacgaacgc	2520
ccggccatgc tcgaggccct cggcaccctc tgggcccagg gttaccgggt cccgtggggc	2580
cgctctttcc cggccggcgg ccgcccgcga ccgtcccca cctaccgtg gcagcgcgag	2640
cgctactgga tcgaggcgcc ggccaagtcc gcccgggcg accgccgtgg cgtccgcgcc	2700
ggcgggcacc cgctcctcgg cgagatgcag accctctcca ccagacctc caccgcctc	2760
tgggaaacca ccctgcacct caagcgcctc ccgtggctcg gagatcaccg cgtccagggc	2820
gccgtcgtct tcccgggcgc cgcctacctc gagatggcca tctcctcggg cgccgaggcc	2880
ctcggggacg gcccgtctca gatcaccgac gtctcctcg ccgaggccct cgccttcgcc	2940
ggtgacgcgg ccgtcctcgt ccaggctcgc actaccgagc agccgtcagg ccgcctccag	3000
ttccagatcg ccagccgggc cccgggcgcc ggccatgcct cttccgcgt ccacgcccgc	3060
ggagctctcc tcagggtcga gcgcaccgag gtcccgccg gtctcacct ctccgccgtc	3120
cgccccgcc tccaggcctc catcccggcc gccgccacct acgccgagct caccgagatg	3180
ggcctccagt acggcccggc cttccagggt atcgccgagc tctggcgagg cgaggagagag	3240
gcccctcgcc gcgtccgct cccggacgcc gccggatccg ccgccgagta ccgcctccac	3300
cccgcctcc tcgacgctg cttccagatc gtcggctccc tcttcgccc ctcgggcgag	3360
gccacgccgt ggggtccggg cgagctcggc tccctccgcc tgctccagcg cccttcgggc	3420
gaactctggt gccacgcccg cgtcgtcaac cacggccacc agaccccga ccgccagggc	3480
gccgacttct gggctcgtga ctctccggc gccgtcgtcg ccgaggctcg cggcctcgtg	3540
gcccagcgcc tcccggggcg cgtccgccgc cgcgaggagg acgactggt cctcgagctc	3600
gagtgggagc cggccgccc cggcaccgcc aaggtcaacg ccggacgctg gtcctcctc	3660
ggcgggtggg gcggcctcgg agccgcctc cgcgccatgc tcgaagccgg cgcccacgca	3720
gtcgtccacg ccgccgagaa caacacctc gcccccggcg tccgcgact cctcgccaag	3780
gccttcgacg ggcaggcccc gaccgccgtg gtcacctcg gctccctga cggcggcggc	3840
gagctcgacc cgggactcgg cccccaggga gccctcgacg ccccgcgctc cgccgacgtc	3900
tccccggacg ccctcgacc ggccctcgtc cgcggctcg actcggtcct ctggaccgtc	3960
caggcgctcg ccggcatggg cttccgcgac gccccgcgcc tctggctctt gaccgcggc	4020
gcccaggccg tcggtgccgg cgacgtctcc gtcaccagg ccccgctcct cggcctcggc	4080
cgcgtcatcg ccatggagca tgccgacctc cgttgcgccc gcgtcgacct cgaccggcc	4140
cgcccgaag gcgagctcg cgcctcctc gcagagctcc tcgtgacga cgccgaggcc	4200
gaggctgccc tcaggggcgg cgagcgctgc gtcgcgcga tcgtccgccg ccagccggag	4260
accgcccgc gaggtcgeat cgagtctgc gtcccgaccg acgtcaccat ccgcgccgac	4320

tccacctacc tcgtcaccgg cggcctcggc ggccctcggc tctccgtcgc aggttggtc	4380
gccgagcgcg gcgccgtca cctcgtcctc gtcggccgat ccggcgccgc ttccgtcgag	4440
cagcgcgccg ccgtcgccgc cctcgaggcc cgcggcgccc gcgtcacgt cgccaaggcc	4500
gacgtcgccg accgagccca gctcgagcgc atcctccgcg aggtcaccac cagcggcatg	4560
ccgtccgcg gcgtcgtcca cgcggccggc atcctcgacg acggcctcct catgcagcag	4620
accccgccgc gcttccgcaa ggtcatggcc ccgaaggctc agggcgccct ccacctccac	4680
gccctcaccg gcgaggcccc gctctccttc ttctcctct acgcctccgg cgtgggcctc	4740
ctcggtctcc cgggacaggg caactacgcc gccccaaca cattctcga cgccctcgcc	4800
caccaccgcc gcgcccaggg tctcccggcc ctctccgtcg actggggcct cttcgccgag	4860
gtcggcatgg ccgcccgcga ggaggaccgt ggcgctcgcc tcgtctcccg cggtatgcgc	4920
tccctcaccg cggacgaggg cctctccgcc ctgcccgc tccttgagtc cggccgcgcc	4980
caggtcggcg tcatgccgt caacccccgc ctctgggtcg aactctacc gcccgccgc	5040
agttcccga tgtctcccg cctcgtcacc gccaccgcg cctcggccgg cggcccgcc	5100
ggcgacggcg acctctccg ccgctcgcc gccccgagc cgtccgccg ctccgcctc	5160
ctcgagccgc tctccgcgc ccagatctcc caggtctcc gcctcccga gggcaagatc	5220
gaggtggacg ccccgctcac ctccctcggc atgaactccc tcatgggtct cgagctcgc	5280
aaccgcatcg aggccatgct cggcatcacc gtcccgcca ccctctctg gacctaccg	5340
accgtcgccg ccctctccg ccacctcgcc cgcgaggcct gcgaggccgc cccggtcgag	5400
tccccgcaca ccaccgcga ttcggccgta gagatcgagg aaatgtcca ggacgacctc	5460
accagctca tcgccgcaa gttcaaggcc ctacgtga	5499
<210> 4	
<211> 21774	
<212> DNA	
<213> Artificial Sequence	
<220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in	
Streptomyces venezuelae	
<400> 4	
gtgaccaccc gtggcccgc cgcacgagc aaccgctca agcaggccgc catcatcatc	60
cagcgccctg aggagcgct cgcggcctc gccaggccg agctcgagcg caccgagccg	120
atgcccatcg tcggcatcgg ctgccgttc ccggcgggcg ccgacgcccc ggaggccttc	180
tgggagctcc tcgacgccga gcgcgacgcc gtccagccgc tcgacatcg ctgggccttc	240

gtcggcgctcg ccccggtcga ggccgtcccg cactgggccc gcctcctcac cgagccgatac	300
gactgcttcg acgcgcctt cttcggcatac tccccgcgcg agggccgctc cctcgacccg	360
cagcaccgcc tctcctcga agtcgcctgg gagggcctcg aggacgcggg catccgcgcg	420
cgctccatcg acggctcccg caccggcgtc ttctgctggcg ccttcaccgc cgactacgcc	480
cgcaccgtcg cccgcctccc gcgcgaggag cgcgacgcct actccgccac cggcaacatg	540
ctctccatcg ccgcggcgcg cctcagctac accctcgccc tccagggccc gtgcctcacc	600
gtcgacaccg cctgctctc ctcctctgc gccatccacc tcgcctgccg ctcctccgc	660
gccggtgagt ccgacctgc cctcgccggc ggctgtctcg cctcctctc cccggacatg	720
atggaggccg ccgcccgcac ccaggccctc tccccggacg gccgctgccg caccttcgac	780
gccctcgcca acggcttcgt ccgcggcgag ggctgcggcc tcgtcgtcct caagcgctc	840
tccgatgccc agcgcgacgg cgaccggatc tgggccttga tccgcggctc cgccatcaac	900
cacgacggtc gctccaccgg cctcaccgcc ccgaacgtcc tcgcccagga gaccgtctc	960
cgcgaggccc tccgtccgc ccacgtcgag gccggcgccg tcgactacgt cgagaccac	1020
ggcaccggca cctccctcgg cgacctatc gaggtcgagg ccctccgcgc caccgtcggc	1080
ccgccccgt ccgacggcac ccgtgcgtc ctcggcgccg tcaagacaa catcgccac	1140
ctcgaggccg ccgcggcggt cgccggcctc atcaaggccg cctctcctt caccacgag	1200
cgcacccgc gcaacctcaa cttccgact ctcaaccgc gcacccgct cgagggctcc	1260
gccctcgccc tcgccaccga gccggtccg tggccgcga ccgaccgcc gcgcttcgcc	1320
ggcgtctcct ccttcggcat gtccggcacc aacccccag tcgtcctcga ggaggccccg	1380
gccgtcgaac tctggccggc cgcgccggag cgctccgccc agctcctcgt cctctccgga	1440
aagtccgagg gcgcctcga cgcaccggcc gcccgccctc gcgagcact cgacatgcac	1500
ccggaactcg gcctcggcga cgtcgcttc tcccttgcca caaccgctc cgccatgtcc	1560
caccgcctcg ccgtcgccgt cactccgcg gagggcctcc tcgccgcct ctcgcccgtc	1620
gccagggcc agaccccgcc cggcgccgcc cgtgcacatg cctcctctc ccgcggcaag	1680
ctcgcttcc tcttcaccgg ccaggcgcc cagaccccg gcattggcg cggcctgtgc	1740
gccgctggc cggccttcg cgaggcttc gaccgtgcg tcgccctctt cgaccgcgag	1800
ctcgaccgcc cgctccgga ggtcatgtgg gccgaggccg gctccgcca gtcctctc	1860
ctcgaccaga ccgcttcac ccaaccggcc ctcttcgccc tcgagtacg cctcaccgcc	1920
ctctggcgct cctgggggggt cgagccggag ctctctctc gccattccat cggcgagctc	1980
gtcgtgcct gcgtcgccgg cgtcttctc cttgaggacg gcgtccgct cgtcgccgcc	2040
cgcggccgcc tcatgcagg cctcagtgc ggccggcgcca tggatccct cggcgccccg	2100

gaggcggaag tcgccgccgc cgtcgccccg cagccgcct ccgtttctat cgccgccgtc	2160
aacggccccg agcaggtcgt catcgccggc gtcgagcagg ccgtccaggc catcgccgca	2220
ggcttcgccg ccccgggcgc cgcaccaag cgcctccacg tctccacgc cttccactcc	2280
ccgtcatgg agccgatgct cgaggagtgc ggccgcgtcg ccgcctccgt cacctatcgc	2340
cgcctcgtcg tctccctcgt ctccaacctc tccggcaagg tcgtcacga cgagctctcc	2400
gccccgggct actgggtccg ccacgtccgc gaggccgtcc gcttcgccga cggcgtcaag	2460
gcccctccacg aggcgggcgc cggcaccttc gttgaggtcg gccgaagcc gacctctctc	2520
ggcctctctc ccgcctgcct cccggaggcc gagccgaccc tctcgcctc cctccgcgcc	2580
ggccgcgagg aggcggccgg cgtcctcgag gccctcgcc gcctctgggc cgccggcggc	2640
tccgtctcct ggccgggctt cttcccgcgc gccggccgcc gcgtcccgct cccgacctac	2700
ccgtggcagc gccagcgcta ctggatcgag gccccggccg agggcctcgg cgccaccgcc	2760
gccgacgccc tcgccagtg gttctaccgc gtcgactggc cggagatgcc gcgtctctcc	2820
gtcgactccc gcccgcccg ctccggcggc tggctcgttc tcgccgaccg cggcggcgtc	2880
ggtgaggccg ccgccgccgc cctctctctc cagggtcgt cctgcgccgt cctccacgcc	2940
ccggccgagg cctccgccgt cgccgagcag gtcacccagg ccctcggcgg ccgcaacgac	3000
tggcaggcg tctctacct ctggggctc gacgccgtcg tcgaggcagg cgcctccgcc	3060
gaggaggtcg ccaaggtcac ccacctgcc gccgccccgg tctcgcctt catccaggcc	3120
ctcggcacgc gcccggcctc ccccgccctc tggatcgta ccccgggcgc ctgcaccgtc	3180
ggcgccgagc cggacgccgc cccgtgccag gccgccctct ggggcatggg ccgcgtcgcc	3240
gccctcgagc acccgggctc ctggggcggc ctggtcgacc tcgacccgga ggagtcccg	3300
accgaggtcg aggcctcgt cgccgagctc ctctccccgg acccgagga ccagctcgcc	3360
ttccgccagg gcccgcccg cgccgccgc ctctcgccg ccccgccgga gggcaacgcc	3420
gccccggtct cctctccgc cgagggtcc tacctcgta ccggcgccct cggcgccctc	3480
ggcctctctg tcgcccgctg gctcgtcgag cgcggcgccg gccacctcgt cctcatctcc	3540
cgccacggcc tcccggaccg cgaggagtgg ggccgcgacc agccgccgga ggtccgcgcc	3600
cgcacgcgcg ccatcgaggc cctcgaggcc caggcgccc gcgtaccgt cgccgcgtc	3660
gacgtcgccg acccgaggg catggccgcc ctctcgccg ccgtcgagcc gccgtccga	3720
ggcgtcgtcc acgcgcggg cctctcgac gacggcctcc tcgccacca ggacgcggc	3780
agactcgccc gggctctccg cccgaaggtc gaggcgccct gggctctcca cacctcacc	3840

cgcgagcagc cgctcgacct cttcgctctt ttctcctccg cctccggcgt cttcggtccc	3900
atcgccagg gctcctacgc cgccggcaac gccttctctg acgccctcgc cgacctcagg	3960
cgaccccagg gcctcgccgc cctctccatc gcctggggcc tctgggcca gggcgcatg	4020
ggctcccagg ccagcgagc ggagcacgag gcctccggca tctgggcat gccgacctcc	4080
cgcgccctgg ccgcatgga gtggctcttc ggaccccgcc ccaccagcg cgtcgtcatc	4140
cagatggact gggccacgc cggcgccgcc ccgcgagcgc cctcccgcg ccgcttctgg	4200
gaccgcctcg tcaccgccac caaggaggcc tcctcctccg ccgtcccggc cgtcgagcgc	4260
tggcgcaacg cctccgtcgt cgagaccgc tcgccctct acgagctcgt ccgcggcgtc	4320
gtcgccggcg tcatgggctt caccgaccag ggcaactctg acgtccggcg cggttcgcc	4380
gagcagggcc tcgactccct catggccgtc gagatccga agcgctcca gggcgagctc	4440
ggcatgccgc tctccgccac cctcgcttc gaccaccga ccgtcgagcg cctcgtcag	4500
tacctctct cccaggccct cgagctccag gaccgcaccg acgtccgtc cgtccgcctc	4560
ccggccaccg aagaccgat cgccatcgtc ggccggcct gccgttccc gggcggcgtc	4620
gaggacctcg agtctactg gcagctctc accgaggcg tcgtcgtctc caccgaggtc	4680
ccggccgacc gctggaacgg cgccgacggc cgcgtcccg gctccggcga ggccagcgc	4740
cagacctacg tcccgcgcgg cggttcttc cgcgaggctg agaccttga cgccgccttc	4800
ttccacatct ccccgcgca ggccatgtc ctcgaccgc agcagcgct cctctcag	4860
gtctcctggg aggccatga gcgcgccggc caggaccgt ccgcctccg cgagtcccg	4920
accggcgtct tcgtcggcg cgccccgaac gactacgccc agcgctcca ggagctgcc	4980
gacgaggcgg caggcctcta ctccggcacc ggcaacatgc tctcgtcgc cgccggccgc	5040
ctctctctct tcctgggct ccatggccc accctcgccg tcgataccgc ctgtctctc	5100
tcctcgtcg cctccacct cggtgtcag tcctccgcc gcggcgagt cgaccaggcc	5160
ctcgtcggcg gcgtcaacat gctcctctcc ccgaagacct tcgccctct ctcgccatg	5220
cacgccctct ccccgggcgg ccgtgcaag accttctccg ccgacgccga cggctacgcc	5280
cgcgccgagg gctgcgccgt cgtcgtctc aagcgctct ccgacgccga gcgcgaccgc	5340
gaccgatcc tcgccgtcat ccgggcacc gccatcaacc acgacggccc gtctcggc	5400
ctcaccgtcc cgtcgggccc ggcccaggag gccctctcc gccaggcct cgcccacgcc	5460
ggcgtcgtcc cgccgacgt cgacttcgtc gactgccacg gcaccggcac cgccctcgcc	5520
gaccgatcg aggtccgcgc cctctccgac gtctacggcc agggccgccc ggccgaccgc	5580

ccgtcatcc tcggcgccgc caaagccaac ctcgccaca tggagcccgc cgccggcctc	5640
gcccggcctcc tcaagcccggt cctcgccctc ggccaggagc agatcccggc ccagccggag	5700
ctcgccgagc tcaacccgct cctcccggtg gaggccctcc cggtcgccgt cgcccgccgc	5760
gccgtcccggt ggccgcgcac cgaccgcccg cgcttcgccc gcgtctctc cttcgcatg	5820
tccggcacca acgcccacgt cgtcctcgag gaggccccag ccgtcgagct ctggccggcc	5880
gccccggagc gatccgccga gctcctcgtc ctctccgca agtccgaggg cgccctcgac	5940
gcccaggccg cccgcctccg cgagcacctc gacatgcacc cggagctcgg cctcggcgac	6000
gtcgcttctt cctcgccac cacacgtcc gccatgaacc accgcctcgc cgtcgccgtc	6060
acctcccgcg agggcctcct cgccgcccctc tccgccgtcg cccaggggcca gaccccgccg	6120
ggcgccgccc gctcatcgc ctctctctcc cgcggaagc tcgcttctt cttaccggc	6180
caggcgcccc agacccggg catgggcgc ggccctctcg ccgcctggcc cgcttccgc	6240
gaggccttcg accgctgctg cgccctcttc gatcgcgagt tggaccgcc gtcgcgag	6300
gtcatgtggg ccgagccggg ctctgccgag tccctctctc tcgaccagac cgcttcacc	6360
cagccggccc tcttcacggt cgagtacgc ctaccgccc tctggcgctc ctggggcgtc	6420
gagccggagc tcgtcgccgg ccattccgcc ggcgaaactg tcgccgctg cgtcgccggc	6480
gtcttctccc tcgaggacgg cgttcgcctc gtcgccgcc gcggccgact catgcagggc	6540
ctctccgccc gcggcgccat ggtctccctc ggcgccccgg agggcgaggt cgccgccgcc	6600
gtcgccccgc acgcccctc cgtctccatc gccgccgtca acggcccga gcaggtcgtc	6660
atcgccggcg tcgagcaggc cgtccaggcc atcgccgcc gcttcgccgc ccgcccgcc	6720
cgcacaaaac gcctccacgt ctccacgcc tccactccc cgtcatgga gccgatgtc	6780
gaggagttcg gcccgctgc cgctccgtc acctaccgcc gcccgctcgt ctccctcgtc	6840
tccaacctct ccggcaaggt cgtcgccgac gagctctccg ccccgggcta ctgggtccgc	6900
cacgtccgcg agggcgctcg cttcgccgac ggcgtaagg ccctccacga ggccggcgcc	6960
ggcaccttcg tcgaggtcgg cccgaagccg acctcctcg gcctctccc ggctgcctc	7020
ccggaggccg aaccgacct cctcgccctc ctccgcgcc gccgcagga ggccgccggc	7080
gtctctgagg ccctcgccgc cctctgggcc gccggcggt ccgtctctg gcctggcgtc	7140
ttcccgaccg ccggccgccg cgtcccgctc ccgacctacc cgtggcagcg tcagcgctac	7200
tggccggaca tcgagccgga ttcccgccgc cagcccgcc cgcacccgac ccagggtgg	7260
ttctaccgcg tcgactggcc ggagatcccg cgcagtctcc agaagtcca ggaggcctcc	7320
cgcggctcct ggctcgtcct cgtgacaag ggccggcgctg gcgaggccgt cgccgccgcc	7380
ctctccaccc gcggttgcc gtgcgtcgtc ctccacgcc cggcagagac ctccgccacc	7440

gccgagctcg tcaccgaggc cgccggcggc cgctccgact ggcaggtcgt cctctacctc	7500
tggggcctcg acgccgtcgt cggcgccgag gcctccatcg acgagatcgg cgacgccacc	7560
cgccgcgcca cgccccggt cctcggcctc gcccgcttcc tctccaccgt ctctgctcc	7620
ccgcgcctct gggtcgtcac ccgcggcgcc tgcatactcg gcgacgagcc ggcaatcgcc	7680
ccctgccagg ccgcctctcg gggcatgggc cgctcgccg ccctcgagca cccggcgca	7740
tggggcgccc tcgtcacct cgacccgcgc gcctccccgc cgaggcctc cccgattgac	7800
ggcgagatgc tcgtaccga gctcctctcc caggagaccg aggaccagct cgccttccgc	7860
cacggcgccc gccacgccgc ccgcctcgtc gccccccgc cgaggcgcca ggccgccccg	7920
gtctccctct ccgccgaggc ctctacctc gtcaccggcg gcctcgccgg cctcggcctc	7980
atcgtcgccc agtggctcgt cgagctcggc gcccgccacc tcgtcctcac ctcccgccgc	8040
ggcctcccgg accgccaggc ctggtgcgag cagcagccgc cggagatcgg cgcccccatc	8100
gccgcccgtc aggcctcga ggcccgcggc gcccgcgta ccgtcgccgc cgtcgacgtc	8160
gccgacgtc agccgatgac cgccctcgtc tcctccgtgg agccgccgt ccgcggcgtc	8220
gtccacgcc cggcgctc cgtcatgcgc ccgtcgccg agaccgacga gacctctc	8280
gagtcgtcc tcccccga ggtcgccgc tcctggctcc tccaccgct cctccacggc	8340
cgcccgctc acctttcgt ccttttcc tccggcgccg ccgtctgggg ctccactcc	8400
cagggcgct acgccgcgc caacgcctc ctcgacggcc tcgccacct ccgccgtcc	8460
cagtcctcc cgccctc cgtcgctgg ggcctctggg ccgaggcgcc catggccgac	8520
gccgaggccc acgccgctt gtccgacatc ggcgtcctcc cgatgtccac ctccgccgc	8580
ctctccgcc tccagcgct cgtcgagacc ggcgtgccc agcgaccgt caccgcatg	8640
gactgggccc gtttcgccc ggtctacacc gcccgcgcc gcccaacct cctctccgc	8700
ctcgtcgcc gcccgacat catcgccccg tcccccgccg ccgccgccac ccgcaactgg	8760
cgccgctct ccgtcgcca ggcccgctc gccctccacg agatcgcca cggtgccgtc	8820
gcccgctcc tcggcttct cgacccgtcc gccctgacc cgggcatggg cttcaacgag	8880
cagggcctcg actccttgat ggccgtcgag atccgcaacc tcctccaggc cgagctcgat	8940
gtccgctct ccaccacct cgcttcgac caccgaccg tccagcgct cgtcgagcac	9000
ctctcgtcg acgtcctca gctcgaggac cgctccgaca cccagcacgt ccgtccctc	9060
gcctccgacg agccgatcgc catcgtcgcc gccgcctgcc gcttcccgcc cggcgctcag	9120
gacctcagat cctactggca gctcctgcc gagggcgctc tcgtatccgc cgaggtccc	9180
gccgaccgt gggacgccgc cgactgtac gaccggacc cggagatccc gggccgcacc	9240
tacgtacca agggcgctt cctccgcgac ctccagcgc tcgacgccac cttcttccgc	9300

atctccccgc gcgaggccat gtccctcgac ccgcagcagc gcctcctcct cgaggctctcc	9360
tgggaggccc tcgagtcgcg cggcatcgcc ccggacacac tccgcgactc cccgaccggc	9420
gtcttcgtcg gcgccggccc gaacgagtac tacaccacgc gcctcgtgg cttcaccgac	9480
ggcgccggcg gcctctacgg cggcaccggc aacatgctct ccgtcacgcg cggacgcctc	9540
tccttcttcc tcggcctcca cggcccgcacc ctgcctatgg acaccgcctg ctctcctcc	9600
ctcgtcgac tccacctgcg ctgccagagt ctccgcctcg gcgagtgcga ccaggccctc	9660
gtcggcgcg tcaacgtcct cctgcgcccg gagaccttcg tcctcctctc ccgcatgcgc	9720
gccctctccc cggacggcgg ctgcaagacc ttctccggcg acgccgacgg ctacgcccg	9780
ggcgaggcgt gcgccgtcgt cgtcctcaag cgcctccgcg acccccagcg cgccggcgac	9840
tccatcctcg cctcatccg cggctccgcc gtcaaccacg accgcccgtc ctccggtctc	9900
accgtcccga accgcccggc ccagcaggcc ctctccgcc aggcctctc ccaggccggc	9960
gtctccccgg tcgacgtcga cttcgtcgag tgccacggca ccggcacgcg cctcggcgac	10020
ccgatcgagg tccaggcct ctccgaggtc tacggcccgg gccgtccgg cgaccgcccg	10080
ctcgtcctcg gcgccgcaa ggccaacgtc gccacctcg aggcgcctc cggcctcgcc	10140
tcctcctca aggcgtcct cgcctccgc cagagcaga tcccggccca gccggagctc	10200
ggcgagctca acccgcacct cccgtggaac acctcccgg tcgccgtccc gcgcaaggcc	10260
gtcccgtggg gcccgggcgc ccgccgcgc cgcgcggcg tctccgctt cggcctctcc	10320
ggcaccaacg tccacgtcgt cctcgaggag gcccggagg tcgagccggc cccggccgcc	10380
ccggccccc cggtcgagct cgtcgtctc tcgccaaagt ccgccgcgc cctcgacgcc	10440
gccgcgccc gcctctccg ccacctctc gccaccgg agctctcct cggcgacgtc	10500
gccttctccc tcgccaccac ccgtccccg atggagcacc gcctcgccat cgccaccacc	10560
tcccgcgagg cctccgcgg cgcctcgac gccgccgcc agcaaaagac cccgcagggc	10620
gccgtccgcg gcaaggccgt ctctcccgc ggcaagctcg ccttctctt caccggccag	10680
ggcgcacaga tgccgggcat gggccgggc ctctacaga cctggccggc cttccgcgag	10740
gccttcgacc gctgcgtcgc cctcttcgac cgcgaaatcg accagccgt ccgcgaggtc	10800
atgtgggccc caccgggcct cgcacaggct gccgcctcg accagaccgc ctacgccag	10860
ccggccctct tcgccctega gtacgcctc gccccctct ggcgctctg gggcgtcgag	10920
ccgcacgtcc tcctcgcca ttcaatcggc gagctcgtcg ccgctgtgt cgccggcgtc	10980
ttctccctcg aggacgccgt ccgcctcgtc gcccccgcg gccgcctcat gcaggccctc	11040

ccggccggcg gcgccatggt cgccatcgcc gcctccgagg ccgaggtcgc cgcctccgtc	11100
gccccgcacg ccgccaccgt ctccatcgcc gccgtcaacg gcccggacgc cgtcgtcatc	11160
gccggcgccg aggtccaggt cttggccctc ggccgccacct tcgccgcccg cggcatccgc	11220
accaagagac ttgccgtctc ccacgccttc cactccccgc tcatggaccg gatgctcgag	11280
gacttccagc gcgtcgccgc caccatcgcc taccgcgccc cggaccgccc ggtcgtctcc	11340
aacgtcaccg gccacgtcgc tggtcgggag atcgccaccc cggagtactg ggtccgccac	11400
gtccgctccg ccgtccgctt cggcgacggc gccaaaggccc tccacgccgc cggcgccgcc	11460
acattcgtcg aggtcggccc gaagccggtc ctctccggcc tctccccgc ctgcctcggc	11520
gaggccgacg ccgtcctcgt cccgtccctc cgcgccgacc gctccgagtg tgaagtcgtc	11580
ctcgccgccg tcggcgccgt gtacgcctgg ggccggcgccc tcgactggaa gggcgtcttc	11640
cctgacggcg cccgccgctg cgcctcccg atgtaccctt ggacgcgcga gcgccactgg	11700
atggacctca ccccgcgctc cgcgccccc gccggcatcg ccggccgctg gccgtcgcg	11760
ggagtcggcc tctgcatgcc gggcgccgtc ctccaccacg tctctccat cggcccgcg	11820
caccagccgt tctcggcga tcacctgctc ttcggcaagg tcgtcgtccc gggcgccctc	11880
cacgtcgccg tcctcctctc catcgccgcc gagcgctggc cggagcgcg catcgagctc	11940
accggcgctg agttcctcaa ggccatagcc atggagccgg accaggaggt cgagctccac	12000
gccgtcctca ccccgagggc cgcggcgac ggctacctct tcgagctcgc caccctcgcc	12060
gccccggaga ccgagcgccg ctggaccacc cagccccgc gcccgctcca gccgaccgac	12120
ggcgccccgg gcgcctccc gcgcctcgag gtctcagagg accgcgccat ccagccgctt	12180
gacttcgccc gcttctctga ccgctctcc gccgtccga tcggctgggg cccgctctgg	12240
cgctggctcc aggacggccg cgtcggcgac gaggcctccc tcgccacctt agtcccgacc	12300
taccgaacg cccacgacgt cgcgccgtc caccgatcc tctcgacaa cggtctcgcc	12360
gtctccctcc tctccaccg ctccgagccg gaggacgacg gcaccccgcc gctcccgttc	12420
gccgtcgagc gcgtccgctg gtggcgcgcc ccggtcggcc gcgtccgctg cggcgcgctc	12480
ccgctctccc aggccttcgg cgtctctcc ttcgtcctcg tcgacgagac cggcgaggtc	12540
gtcggcagg tcgagggtt cgtctgccgc cgcgtccgc gcgaggtctt cctccgccag	12600
gagtcggcg cctccaccg cgcctctat cgcctcgact ggccggaggc cccgctcccg	12660
gacgccccgg ccgagcgcat cgaggagtcc tgggtcgtcg tcgccgccc gggctccgag	12720
atggccggcg ccttcgccac ccgcctcaac cgctgcgtcc tcgccgagcc gaagggcctc	12780
gaggccgccc tcgccggcgt ctccccggcc ggctctatct gcctctggga ggccggcgcc	12840
cacgaggagg ccccgccgc cgcacgacg gtcgcaaccg agggcctctc cgtcgtccag	12900

gcccctccgcg accgcgccgt ccgcctctgg tgggtcacca tgggcgccgt cgcgctcgag	12960
gcccggcgagc gcgtccaggt cgcaccgcc cctgtctggg gcctcggccg caccgtcatg	13020
caggagcgcc cggagctctc ctgcacctc gtcgacctc agccggaggc cgacgccgcc	13080
cgctccgccg acgtcctcct ccgcgagctc ggccgcgccg acgacgaaac ccaggtagcc	13140
ttcgcgtccg gcaagcgccg cgtcgcccgc ctctcaagg ccaccactcc ggagggcctc	13200
ctcgtcccgg acgccgagtc ctaccgcctc gaggccggcc agaagggcac cctcgaccag	13260
ctccgcctcg ccccgccca gcgcgcgcc cccggcccgg gcgaggtcga gatcaaggtc	13320
accgcctccg gcctcaactt ccgaaccgtc ctgccttcc tcggcatgta cccgggcgac	13380
gcccggccga tggggggcga ctgcgccggc gtcgccaccg ccgtcggcca gggcgctccg	13440
catgttgcg tcggcgacgc cgtcatgact ctggcaccc tccaccgctt cgtcaccgtc	13500
gacgcccgcc tcgtcgtccg ccagccggcc ggcctcacc cggcccaggc cgccaccgtc	13560
ccggtcgctt tcctaccgc ctggctcgcc ctccaccacc tcggaaacct ccgcgcggc	13620
gagcgcgtcc tcctaccgc cgcgccggc ggctcggca tggcagccgt ccagatcgcc	13680
cgttggtcg gcgcggaggt cttgccacc gcttccccgt ccaagtgggc agccgtccag	13740
gccatggcg tcgccgcac ccacatcgcc tcctccgca ccctcgagtt cgccgagacc	13800
ttccccaag tcaccggcg ccgcggcgtc gacgtcgtc tcaacgcct cgccggcgag	13860
ttcgtcgacg cctccctct cctcctctc accggcgac gcttctcga gatgggcaag	13920
accgacatcc gcgaccgcg ccgctcgcc gccgccacc cgggcgtccg ctaccgcgtc	13980
ttcgacatcc tcgagctcg cccggaccgc accgcgaga tcctcgagcg cgtggtcgag	14040
ggcttcgcg ccggccatct ccgtgcctc ccggtccag ccttcgcat caccaaggcc	14100
gaggccgct tccgttcat ggcccaggcc cgccaccagg gcaaggtcgt cctcctccg	14160
gccccgtccg ccgccccgt cgccccgacc ggcaccgtc tcctaccgg cggcctcggc	14220
gccctcggcc tcacgtcg ccgtggctc gccagcagg gcgtccgca catggtctc	14280
acaggccgcc gcggcctcga cactccgggc gccccaagg ccgtcgccga gatcgaggca	14340
ctcggcgccc gcgtcacat cgccgcctc gacgtcgcc accgcaacgc cctcgaggcc	14400
gtctccagg ccatcccgc cgaatggcg ctccaggcg tcctccacc gcgaggcgcc	14460
ctcgacgag gcgtcctcga cgagcagacc accgaccgt tctcccggt cctcgccccg	14520
aaggtcaccg gcgcctgga cctccacgag ctaccgcg gcaacgaact cgcttcttc	14580
gtactcttct cctcatgtc cggcctctc ggctccgccc gccagtcaa ctacggccc	14640
gcaaacacct tcctcgacgc cctcgccgc caccgccgc ccgaggccct cgccgcacag	14700
tcctcgcct ggggcccgtg gtccgacggc ggcatggcg ccggcctctc tgccgcctc	14760

caggcccccc tcgcccccca cggcatgggc gccctctccc cagcccaggg caccgccctc	14820
cttgccagg ccctcgcccg cccggagacc cagctcgcg ccatgtccct cgacgtccgc	14880
gccgcctccc aggcctccgg cgcgcgctc ccgccgtct ggccgcacct cgtccgcgc	14940
gagggccccc acaccgccgc cggcgcccag ggcccccctg ccgccgcct cggcgccctc	15000
ccggaggccc gacgcgcga cgaggtccgc aaggtcgtcc aggccagat cgcgcgctc	15060
ctctcctggt ccgcgcctc cgcgtcccg gtcgaccgc cgtctcga tctcggccta	15120
gactccctca ccgcgtcga gtcgcgaac gtctcgcc agcgcgtcgg cgccaccctc	15180
ccggccacc tcgccttcga ccaccgacc gtcgacgcc tcaccgctg gtctctgac	15240
aaggtcctcg ccgtcgcca gccgtccgc tctccgcca agtctcccc gcaggtcgcc	15300
ctcgacgagc cgatcgccat catcggcac ggctgtcgt tcccggcgg cgtcgccgac	15360
ccggagtctt tctggcgct cctcgaggag ggctccgac ccgtcgtcga ggtccgcac	15420
gagcgctggg acatcgatgc cttctacgac ccggaccgg acgtccgagg caagatgacc	15480
accgcttcg gcggttctt ctcgacatc gaccgttcg acccggtctt cttcggcac	15540
tccccgcgc aggccaccac catggaccgc cagcagcgc tctctcga gacctctgg	15600
gaggccttcg agcgcgcgg catctcccg gagcgtctca tgggtcga caccggcgtc	15660
ttctcgac tctctacca ggagtacgc gccctcgcc ggccatcga ggcttcgac	15720
ggctacctcg gcaccggcac caccgcctc gtcgcctcc gccgatctc ctacgtctc	15780
ggcctcaagg gccgtccct caccgtcgc accgcctgct cctctccct cgtcgccgc	15840
cacctgcct gccaggccct ccgcgcggc gagtgtccg tcgccctgc cggcggcgta	15900
gccctcatgc tcaccccggc cacttcgtc gatttctcc gccctcgccg cctcgccccg	15960
gacggccgct gcaagtctt ctcgcccgc gccgacggc tcggctggc cgagggtgc	16020
gcatgtctc tctcaagcc gtcgcgcac gccagcgc acggcgacc gatctcgcc	16080
gtcatccgc gcaccgccgt caaccaggac ggccgctcca atggcctcac cgcgccaac	16140
ggctcctccc agcaggaggt catccgccgc gccctcagc aggcggcct cgcgccgcc	16200
gacgtctct acgtcgagt ccacggcacc ggaccaccc tcggcgacc gatcgaggc	16260
caggccctcg gcgcgtctt cgcaccggc cgcgcgtcc accgccgt cgtcataggc	16320
tctgtcaagt ccaacatcgg ccataccag gccgcgcgg gcgtcgccg cgtcatcaag	16380
gtagccctcg cctcgagcg cgccctcat ccgcgtccc tccacttga cgcgccaac	16440
ccgcacatt cgtggtcga gtcgcgcgc caggtcgcc ccaagccgt cgagtggacc	16500

cgcaacggcg tcccgcgcgc gcgcggcgtc tctctcttcg gcgtctccgg caccaacgcc	16560
cacgtcgtcc tcgaggaggc cccggccgcc gccttcgccc cggccgcgcgc ccgtctcgcc	16620
gagctcttcg tctcttcgcg caagtccgcc gccgccctcg acgcacaggc ggcccgcctc	16680
tccgcccatt tcgtcgccca cccggagctc ggcttcggcg acctcgctt ctcctcgcc	16740
accacccgct ccccgatgac ctaccgcctc gccgttgccg ccacctccg cgaggccctc	16800
tccgcccgc tcgacaccgc cgcgggggg caggccccgc cggccgcggc ccgcgccac	16860
gcctccaccg gctccgcccc gaaggtcgtc ttcgtcttc cgggccaggg ctcccagtgg	16920
ctcgcatgg gccagaagct cctctccgag gagccggtct tccgcgacgc cctctcgcc	16980
tgcgaccgcg ccatccaggc cgaggccggc tggctcttgc tcgccgagct cgccgccgac	17040
gagaccacat cccaactcgg cgcctcgcg gtctgtccgc cggccctctt cgccatcgag	17100
gtcgccctct ccgccctctg gcgtctctgg ggcttcgagc cggacgccgt cgtcgccac	17160
tccatggcg aggtcgccgc cgcggcgctc gccggcgccc tctccctcga agacgccgc	17220
gccatcatct gccgcgcgc cctctctctt cgcgcctct cggccaggg tgagatggc	17280
gtctgcgagc tctccctcgc cgaggctgaa gccgccctcc tcggctacga ggaccgcctc	17340
tccgtagccg tctccaactc cccgcgtcc accgtctctg ccggcgagcc ggccgcctc	17400
gccgaggctc tcgcatctct cgcgcccaag ggctctctct gccgccggt caaggctgac	17460
gtcgccctcc actccccga gatcgaccgc ctccgcgacg agctctcgc cgcctcggc	17520
gagctcgaac cgcgccaggc caccgtctcc atgcgtcca ccgtcacctc caccatcatg	17580
gccggcccg agctcgtcgc ctctactgg gccgacaacg tccgccagcc ggtccgcttc	17640
gccgaggccg tccagtcct catggaggac ggccacggcc tcttcgtcga gatgtccca	17700
caccgatcc tcaccacctc cgtcgaggag atccgccgcg ccaccaagcg cgaggcgctc	17760
gccgtcgggt cctccgcgc gcgccaagac gagcgctct ccatgctcga ggccctcggc	17820
gccctctggg tccacggcca ggccgtcggc tgggagcgcc tcttctcgc cggcggcgc	17880
ggcctccgc gcgtccgct cccgacctac ccgtggcagc gcgagcgcta ctgggtcgac	17940
gccccgaccg gcggcgccgc cggcggtcc cgttcgccc acgccggctc ccaccgcctc	18000
ctcgcgaaa tgcagacct ctcaccacag cgtccacc gcgtctggga gaccaccctc	18060
gacctcaagc gcctcccggt gctcggcgac caccgcgtcc agggcgccgt cgtcttccg	18120
ggcgccgct acctcgagat ggccctctcc tccggcgccg aagccctcgg cgtggcccg	18180
ctccaggct cgcagctgt cctcgccgag gccctcgct tcgccgacga cccccggcc	18240
gctgtccagg tcatggccac cgaggagcga ccggcgccgc tccagttcca cgtcgctcc	18300
agagtcccg gccacggcg gcgcgcctc cgtccacgc cccgcggcgt cctccgcag	18360

atcgagcgcg ccgaggtccc ggcccgcctc gacctcgag ccctccgcgc ccgcctccag	18420
gcctccgccc cgcccgccgc cacctacgcc gccctcgccg agatgggcct cgagtacggc	18480
ccggccttcc agggcctcgt cgagctctgg cgcggcgagg gggaggcgct tggccgcgtg	18540
cgctccccgg aagccgcccgg ctccccggcc gcctgccgcc tccacccggc cctcctcgac	18600
gcctgcttcc acgtctctc cgccttcgcc gaccgcggcg aggccacccc gtgggtcccg	18660
gtcgagatcg gctccctccg ctggttccag cgcccgccg gcgagctctg gtgtcacgcc	18720
cgtccgctct cccacggcaa gccgaccccg gaccgccgt ccaccgactt ctgggtcgtc	18780
gactccaccg gcgccatcgt cgccgagatc tccggcctcg tgcgccagcg cctcgccggc	18840
ggcgtccgcc gcccgagga ggacgactgg ttcatggagc cggcctggga gccgaccgcc	18900
gtcccgggct ccgaggtcat ggccggccgc tggctcctca tcggctccgg cggcggcctc	18960
ggcggccccc tccactccgc cctcaccgag gccggccact ccgtagtcca cgcgaccggc	19020
cgcggcacca gcgcgcggc cctccaggcc ctctcaccg cctccttga cggccaggcc	19080
ccgacgtccg tcgtccacct cggtctcctc gacgagcgcg gcgtcctga cgcgacgcc	19140
ccgttcgatg ccgatgcct cgaggagtcc ctctccgcg gctcgactc cgtcctctgg	19200
accgtccagg ccgtcgccgg cgccggcttc cgcgaccgc cggcctctg gctcgtcacc	19260
cgcggcgccc aggccatcgg cgccggcgac gtctccgtcg cccaggcccc cctcctcggc	19320
ctcgcccgcg tcctgcctc cgagcacgcc gagctccgt gcgcccgcac cgacctcgac	19380
ccggccccgc gcgacggcga ggtcgacgag ctctcgccg agctcctgc cgacgacgcc	19440
gaggaggagg tcgccttccg cggcggcgag cgccgcgtcg cccgcctcgt ccgccgcctc	19500
ccagagaccg actgccgga gaagatcgag ccggccgagg gccgccgtt ccgcctcgag	19560
atcgacggct ccggcgctct cgacgacctc gtctccgcg ccaccgagcg ccgccgccg	19620
ggccccggcg aggtcgagat tgccgtcgag gccgccggcc tcaacttct cgacgtcatg	19680
cgcgcatg gcatctaccc tgggccgggc gacggcccgg tcgcctcgg cgccgagtgc	19740
tccggccgca tcgtcgccat gggcgagggc gtcgagtccc tccgcatcgg ccaggacgtc	19800
gtcgccgtcg ccccgttctc ctccggcacc cacgtacca tcgatgccg catgctgcc	19860
ccgcgcccgg ccgcactcac cgccgccag gccgccccc tccgggtcgc cttcatgacc	19920
gcctggtacg gcctcgtcca cctcggcgc ctccgcgccg gcgagcgct cctcatccac	19980
tcagccaccg gcggcacccg cctcgccgcc gtccagatcg cccgccact cggcgccgag	20040
atcttcgcca ccgcgggtac cccggagaag cgcgcctggc tccgcgagca gggcatcgcc	20100
cacgtcatgg actccgctc cctcgacttc gccgagcagg tcctcgccgc caccaagggc	20160
gagggcgctg acgtcgtcct caactccctc tccggcgccg ccatcgacgc ctccctctcc	20220

accctcgtcc cggacggcgc cttcatcgaa ctcggcaaga ccgacatcta cgccgatcgc	20280
tccttgggccc ttgcacactt ccgcaagtcc ctctcctact ccgccgtcga cctcgccggc	20340
ctcgccgtcc gccgcccga gcgcgtcgcc gccctcctcg ccgaggtcgt cgatctcctc	20400
gcccgccggcgc cctccagcc gctcccgtc gagatcttcc cgctctcccg cgccgccgac	20460
gccttccgca agatggccca ggcccagcac ctcggcaagc tcgtcctcgc cctcgaggac	20520
ccggacgtcc gcatccgct cccggggcgc tccggcgtcg ccataccgcgc cgacggcgcc	20580
tacctcgtca ccggcggcct cggcggcctc ggctctcag tcgccggctg gctcgccgag	20640
cagggcgccg gccacctcgt cctcgtcggc cgctccggcg ccgtctccgc cgagcagcag	20700
accgccgtcg ccgccctcga ggcccacggc gcccgcgta ccgtcgcccg cgccgacgtc	20760
gccgaccgcg ccagatgga gcgcatctc cgcgaggta ccgcctccgg catgccctc	20820
cgccggcgtcg tccacgcgc cggcacctc gacgacggac tcctcatgca gcagaccccg	20880
gcccgcttcc gcgcgctcat ggccccgaag gtcgcggcg cctccacct ccacgccctc	20940
accgcgaag ccccgctctc cttcttcgtc ctctacgct ccggcgccgg cctcctcggc	21000
tccccgggccc agggcaacta cgccgccgc aacaccttcc tcgacgcct cgcccaccac	21060
cgccgcgccc agggcctccc ggccctctcc atcgactggg gcctcttcgc cgacgtcggc	21120
ctcgccgcgc gccagcagaa ccgcggcgcc cgctcgtca ccccgggcac ccgtccctc	21180
accccgagc agggcctctg ggccctcgag cgctcctcg acggcgaccg caccagggc	21240
ggcgtcatgc cgttcgacgt ccgccaatgg gtcgagttct acccgccgc cgctcctcc	21300
cgccgcctct cccgcctcat gaccgcgcgc cgctcgcct ccggccgcct cgccggcgac	21360
cgcgacttgc tcgagcgct cgccaccgc gaggcggcg ccgagccgg catgctccag	21420
gaggtcgtcc gcgcccaggt ctcccagtc ctccgcctct ccgagggcaa gctcgacgtc	21480
gacgccccgc tcactcct cggcattggac tccctcatgg gcctcgagct ccgcaaccgc	21540
atcgaggccg tcctcggcat caccatgcc gccaccctcc tctggaccta cccgaccgtc	21600
gccgccctct ccgccacct cgctccac gtcgtctcca ccggcgacgg cgagtccgcc	21660
cgcccgccgg acaccggctc cgtcgcccg accaccacg aggtcgctc cctcgacgag	21720
gacggcctct tcgccctcat cgacgagtc ctcgcccgcg ccggcaagcg ctga	21774
<210> 5	
<211> 11394	
<212> DNA	
<213> Artificial Sequence	
<220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in	

Streptomyces venezuelae

<400> 5

atgaccgacc gcgagggtca gtcctcagag cgtctccgag aggtcacctt cgccctccgc	60
aagaccttca acgagcgaga caccctcagag ctcgagaaga ccgagccgat cgccatcgtg	120
ggcatcgggt gccgtttccc gggcgggcgcc ggcaccccgaggccttctg ggagctcctc	180
gacgacggcc gcgaccccat ccgcccgtc gaagagcgct gggccctcgt aggcgtcgac	240
ccggcgagac acgtcccgcg ctggggcgcc ctcctcaccg aggccatcga tggcttcgac	300
gccgcttctt tcggcatcgc ccgcgcgag gcccgctccc tcgaccgca gcaccgctc	360
ctcctcaggg tcgcctggga gggcttcgaa gacccggca tcccgccgag ctcctcgtc	420
ggctcccgca ccggcgcttt cgtcgggcgtc tgcgccaccg agtacctcca cgccgcccgc	480
gcccaccagc cgcgcgagga acgagcgcc tactccacca ccggcaacat gctttccatc	540
gcccgggac gcctctccta caccctcgga ctcagggcc cgtgccttac cgtcgatacc	600
gcctgctcct cctcctcgt cgccatccac ctcgctgcc gttcctccg cgcccgcgaa	660
tccgacctcg cctcgcgg cggcgtcaac atgtcctct ccccggaac catgcgcgcc	720
ctcgcccgca ccagggcct ctcgccgaac ggccgctgcc agaccttga cgctccgcc	780
aacggcttcg tcccgggcga gggctgcggc ctcacgtcc tcaagcgct ctcgacgcc	840
cgcccgagc gcgaccgat ctggggcctc atcccgggct ccgccatcaa ccaggatggc	900
cgtccaccg gcctaccgc accgaacgtc ctcgcccagg gcgcctcct ccgagggcc	960
ctccgaacg ccggcgctga ggccgaggcc atcggtaca tcgagacga ccgcccgcg	1020
acctcctcg gcgaccgat cgagatcgag gccctccgag ccgtcgtcgg ccggggccgc	1080
gcccagggcg cccgctgct cctcgggcgt gtaagacca acctcgcca tctcgagggc	1140
gcccggggcg tcgcccgcct catcaaggcc acctctccc tccaccaga gcgtatccg	1200
cgcaacctca acttcgcac cctcaaccg cgcacccga tcgagggcac cgccctcgc	1260
ctcgccaccg agccggctcc gtggccgcg accggccgca cccgcttcg cgcgctctcc	1320
tccttcggca tgtccggcac caaccccaac gtcgtcctcg aggaagcccc ggccgtcgaa	1380
ccggaggcgg ccgaccgga gcgcccgcg gagctcttcg tcctctccg caagtccgc	1440
gcccctcgc acgtcaggc cgcccgctc cgcgaccacc tcgagaagca cgtcgagctc	1500
ggcctcggcg acgtgcctt ctcctcgc accaccgct ccgcatgga gcaccgctc	1560
gccgtcggcg cctcctccg cgaggccctc cgggcgccc tctccgccc cgcccagggc	1620

cacacccgcg cgggcgcgct ccgcggccgc gcctccggcg gctccgcccc gaaggtcgtc	1680
ttcgtcttcc cgggccaggg ctcccagtggt gtcggaatgg gccgcaagct catggccgag	1740
gaaccggtct tccgcgcgcg cctcgaaggc tgcgaccgcg ccatcgaggc agaggccggt	1800
tggctccctcc tcggcgagct ctccgccgac gaggccgcct cccagctcgg ccgcatcgac	1860
gtcgtccagc cggctctctt cgccatggag gtcgccctct ccgccctctg gcgctcctgg	1920
ggcgtcgagc cggaggccgt cgtcggccac tccatgggcg aggtagccgc cggccacgtc	1980
gccggcgcg cgtccctcga ggacgccgtc gccatcatct gccgccgtc ccgctctctc	2040
cgccgcatct ccggccaggg cgagatggcc ctgctcgagc tctccctcga ggaggccgag	2100
gcccctca gaggtcacga gggccgctc tcggtcgccg tctccaactc ccccgctcc	2160
actgtctctg ccggcgagcc ggccgctc tcagaggtcc tcgccgctt caccgccaag	2220
ggcgtcttct ggccaggt caaggtcgac gtcgcatccc actccccga ggtcgaccg	2280
ctccgcgagg agtcatcg cgccctcggc gccatccgc cgccgcccgc cgcgtcccg	2340
atgcgtcca ctgtaccgg cggcgtcatc gccggcccgg agtcggcg ctcctactgg	2400
gccgacaacc tccccagcc ggtccgttc gccccgccc cccaggccct cctcgagggc	2460
ggccggccc tcttcacga aatgtcccg caccgatcc tcgtccgccc gtcgacgag	2520
attcagaccg ccgccgagca gggaggcgcc gcagtcggct cctccgccc cgccaggac	2580
gaacgcgcca cctcctcga ggccctcggc accctctggg cctccggcta cccggtctcc	2640
tgggcccgc tcttccggc cggcggtcgc cgcgtcccgc tcccagcta cccgtggcag	2700
cacgagcgt gctggatcga ggtcgagccg gacccccgc gcctcgccgc cgcgaccg	2760
accaaggact ggttctacc caccgactgg ccggaggtcc cgccgcccgc ccctaagtcc	2820
gagaccgccc atggctctg gctcctctc gccgaccgc gcggcgtcgg agaggccgtc	2880
gccgcccgc tctccaccg cggcctctcc tgcaccgtcc tccacgctc cgtgacgcc	2940
tccaccgtcg ccgagcaggt ctccgaggcc gcctcccgc gcaacgactg gcaggcgctc	3000
ctctacctct ggggcctcga cgccgtcgtc gacccggcg cctccgcca cgaggtctcc	3060
gaggccacc gccgcgccac cgccccgtc ctccgctcgt tacgttct ctcgcccgc	3120
ccgaccgcgc cgcgttctg ggtcgtcacc cgcggcgct gcaccgtcgg cggcgagccg	3180
gaggcctccc tctccaggc cgccctctgg ggctcgcgc gattcgccgc tctcgagcac	3240
ccggccgct gggcgccgt cgtcgacct gacctcaga agtccccgac cgagatcgaa	3300
ccgctcgtcg ccgagctct cagccctgac gccgaggacc agtcgcctt ccgctccggc	3360

cgccgccacg ccgcccgcct cgtcgccgcc ccgccggagg gcgacgtcgc cccgatctcc	3420
ctctccgcgc agggctccta cctcgtcacc ggccgcctcg gcggcctcgg cctcctcgtc	3480
gcccgtggc tgcgcgagcg cggcgcccgc cacctcgtcc tcacctcccg ccacggactc	3540
ccggagcgcc aggcctccgg cggcgagcag ccgccggagg cccgcgccg catcgccgcc	3600
gtcgagggcc tcgaggccca gggcgcccgc gtcaccgtcg ccgccgtcga cgtcgcgag	3660
gccgatccga tgaccgcct cctcgccgcc atcgagccgc cgtccgagg cgtcgtccac	3720
gcccgccgcg tcttcccgg cgcacacctc gccgagaccg acgaggccct cctcgaatct	3780
gtctctcgac cgaaggtcgc cggcagttgg ctctccacc gtctctccg cgaccgaccg	3840
ctcgacctct tgcctctctt ctctccggc gcccccgtct gggcgggcaa gggccagggc	3900
gcctacgcgc ccgccaacgc ctctctcgac ggctcgcgc accaccgcgc cgccactcg	3960
ctcccgcccc tctccctcgc ctggggcctc tgggcccagg gcggcatggt cgacgccaag	4020
gcccacgccc gactctcga catcggcgtc ctccgatgg ccaccgccc ggccctctcc	4080
ggcctcgagc gcctcgtcaa cacctccgc gtccagcgt cgtcaccgc catggactgg	4140
gcccgtctcg ccccgtctc cgcgcgccgc ggccgcccga acctctctc cgccctcgtc	4200
gctgaggacg agcgcgccgc ctccccgcgc gtcccgaccg ccaaccgcat ctggcgcggc	4260
ctcagcgtcg ccgagtccg ctccgccttg tacgagctcg tcccgggcgt cgcgccgtc	4320
ctcggtctct ccgaccggg cgactcgcac gtcggccgag gcttcgccga gcagggcctc	4380
gactccctca tggccctcga gatccgaac cgcctccagc gcgagctcgg cgagcgcctc	4440
tccgccacc tcgccttcga ccaccgacc gtcgagcgcc tcgtcgcca cctgctcacc	4500
gacgtcttga agctcgagga ccgctccgac acccgccaca tccgtccgt cgcgcggac	4560
gacgacatcg ccatcgctcg cgcgcgatgc cgttcccgg gcggagacga aggcctcgag	4620
acctactggc gccacctcgc cgagggcgat gtcgtctcca ccgaggtccc ggccgaccgc	4680
tggcgcgccg ccgactggta cgaccggac ccggaggctc cgggcccgcac ctacgtcgcc	4740
aaggcgccct tcctccgca cgtccgtcc ctcgacgcgc ctttcttcgc catctcccgc	4800
cgcgaggcca tgtccctcga cccgcagcag cgcctcctcc tcgaagtctc ctgggaggcc	4860
atcgagcgcg ccggccagga cccgatggcc cttcgcgagt ccgccaccgg cgtcttcgtc	4920
ggcatgatcg gctccgagca cgcggaacgc gtccagggcc tcgacgacga cgcgcctc	4980
ctctacggca ccaccggcaa cctcctctca gtcgccgcgc gccgcctctc cttcttctc	5040
ggcctccacg gcccaccat gaccgtcgat accgcctgtt cctctccct cgtcgccctc	5100
catctcgct gtcagtcct ccgctcggga gactgcgacc aggcctcgc cggcggtcc	5160
tccgtctcc tctccccgc ctcttcgtc gccgatccc gcatgcgct cctctcccgc	5220

gacggccgct gcaagacctt ctccgccgcc gccgacggat tcgcccgcgc cgagggtctgc	5280
gccgtcgtcg tctcaagcg tctccgcgac gccccagcgc accgcgaccc gatcctcgcc	5340
gtcgtccgct ccaccgcat caaccacgac ggcccgctct ccggcctcac cgtcccgctcc	5400
ggcccgggccc agcaggccct cctccgccag gccctcgccc aggcgggctc cgcccgggcc	5460
gaggtcgact tcgtcagtg ccacggcacc ggccaccgcc tcggcgaccc gatcgaggtc	5520
caggccctcg gcgcgtcta cggccgcggc cgcccgccgc agcggccgct ctggctcggc	5580
gctgtaaagg ccaacctcgg ccacctcgag gccgccgccg gcctcgccgg cgtcctcaag	5640
gtcctcctcg cctcgagca cgacgagatc ccggctcaac cggagctcga cgagctcaac	5700
ccgcacatcc cgtgggcccga gctcccggtc gccgtcgtcc gcccgccgct gccgtggccg	5760
cgcgggcggc gcccgcgccg cgccggcgtc tccgcttcg gcctcagtgg caccaacgcc	5820
cacgtcgtcc tcgaggaggc ccggcgccgc gagccggctc ccggcgctcc ggagcgcgcc	5880
gccgagctct tcgtcctctc cgccaagtcc gcccgccgcc tcgacgcca ggccggccgc	5940
ctccgcgacc acctcgagaa gcacgtcgag ctccggcctc gcgacgtcgc cttctcctc	6000
gccaccaccc gctccgcat ggagcaccga ctcccgctcg ccgctcctc ccgcgaggcc	6060
ttgcgcggcg cctctccgc cgccggccag ggccacaccc cgccggggcg cgtccgcggc	6120
cgccctccg gcggctccgc ccgaaggctc gtcttcgtct tcccgggcca gggctccag	6180
tgggtcggca tgggcccga gctcatggcc gaggagccgg tcttcgcgc cgccctcgag	6240
ggctcgacc gcgccatcga ggccgaagcg ggctggctcc tctcggcga gctctccgc	6300
gacgaggccg cctcccagct cggacgcatc gacgtcgtcc agccggctct cttcgccatg	6360
gaagtgcgcc tctccgctt ctggcgctcc tggggcgctc aaccagaggc cgtcgtcggc	6420
cactccatgg gcgaggctgc cgccggccac gtcgccggcg cctctcctt cgaggacgcc	6480
gtcgccatca tctgccgcg ctcccgctc ctccggcgca tctccggaca gggagagatg	6540
gccctcgtcg agctctcct cgaggaggcc gaggccgctc tcccgggcca cgagggccgc	6600
ctctccgtcg ccgtctcaa ctccccgcgc tccaccgtcc tcgcccgcga gccggccgcc	6660
ctctccgagg tctcggcg cctcaccgcc aaggcgctct tctggcgcca ggtcaaggctc	6720
gacgtcgctt cccactcccc gcaggtcgac ccgtccgcg aggagctcat cgccgcctt	6780
ggcgccatcc gcccgcgcg cgccggcgtc ccgatgcgt ccaccgtcac cggcgggctc	6840
atcgctggcc cggagctcgg cgctcctac tggcccgaca acctccgcca gccggtccgc	6900
ttcgccggcg ccgcccaggc cctcctcgag ggccggcccg cctcttcat cgagatgtcc	6960
ccgacccga tctcgtccc gccgtcgac gagatccaga cggccggcga gcaggcgga	7020
gcccgctcg gctcctccg gcgcgccag gacgaacgc ccacctcct cgaagccctc	7080

ggcacctgt gggcctccgg ctaccgggtc tctgggccc gcctcttccc ggccggcggc	7140
cgcccgctcc cgtccccac ctaccgtgg cagcacgagc gctactggat cgaggactcc	7200
gtccacggct ctaagccgtc cctccgcctc cgccagctcc gcaacggcgc caccgaccac	7260
ccgtctctcg gcgccccgt cctcgtctcc gcccgcccgg gcgcccacct ctgggaacag	7320
gccctcagcg acgagcgtct ctctacctc tccgagcacc gcgtccacgg cgaggccgtc	7380
ctcccgctcg ccgctacgt cgagatggcc ctcccgccg gcgtcgacct ctacggcacc	7440
gccacctcg tgctcagca gctcgccctc gagcgcgcct tggccgtccc gtccgagggc	7500
ggtcgcatcg tccaggtcgc cctctccgag gagggcccgg gccgcgcctc cttccaggtc	7560
tctcccgcg aagaggcagg ccgtctctgg gttcgccacg ccaccggcca cgtctgctcc	7620
ggccagtctt ccgccgtcgg cgcctcaag gagccccgt gggagatcca gcgccgtgc	7680
ccgtccgtcc tctctccga ggccctctac ccgtctctca acgagcacgc cctcgactac	7740
ggcccggtgt tccagggcgt cgaacaggtc tggctcgca ccggcgaggt cctcgccgc	7800
gtccggctcc caggcgacat ggctctctcc tccggcgcgt accgcatcca cccggccctc	7860
ctcgagcct gcttccaggt cctaccgcc ctctcacca cccccgaac catcgagatc	7920
cgccgccgcc tcaccgacct ccacgagccg gacctccgc gctccgcgc cccggtcaac	7980
caggccgtct ccgacacctg gctctgggac gcccccctgg acggcggccg ccgccaatcc	8040
gcctccgtcc cggctgacct cgtctcggc tccttccacg ccaagtggga ggttatggag	8100
cgcctcgccc aggctacat catcggcacc ctccgcatct ggaacgtctt ctgcgccgcc	8160
ggcgagcgcc acaccatcga cgagctctc gtccggctcc agatctccgt cgtctaccgc	8220
aaggtcatca agcgtggat ggagcacctc gtgccattg gcatctctgt cggcgacggc	8280
gagcacttcg tctctccca gccgtcccg gagccggacc tcgccgcgt cctcgaggag	8340
gccggccgcg tcttcgccga cctccggtc ctctcgagt ggtgcaagtt cgccggcgag	8400
cgctcgccg acgtctcac cggcaagacc ctgcccctcg agatctctt cccggcgggc	8460
tccttcgata tggccgagcg catctaccgc gattccccga tcgccgcta ctccaacggc	8520
atcgtacgcg gcgtcgtcga gtccgcagcc cgcgtcgtcg ccccgagcgg catgttctcc	8580
atctcgaga tcggcgccgg caccggcgcc accaccgccg ccgtctctcc ggtctctctc	8640
ccggaccgca ccgagtacca cttaccgac gtctccccgc tcttctcgc ccgcgcagag	8700
cagcgcttcc gcgactacce gtctctcaa tacggcatcc tcgacgtcga ccaggaaccg	8760
gccggacagg gctacgcaca ccagcgcttc gacgtcatcg tagccgcaa cgtcatccac	8820

gccacccgcg acatccgcgc caccgccaag cgctctctct cctcctcgc cccgggcggc	8880
ctcctcgtcc tcgtcgaggg caccggccac ccgatctggt tcgacatcac caccggcctc	8940
atcgagggct ggcagaagta cgaggacgac cttcgcatcg accacccgct cctcccggcc	9000
cgcacctggt gcgacgtcct ccgccgcgtc ggcttcggcg acgccgtctc cctcccgggc	9060
gacggctccc cggccggcat cctcggccag cagtcatcc tctcccgcgc cccgggcatc	9120
gccggcgccg cctgcgactc ctccggcgag tccgccaccg agagcccggc cgcgcgcgcc	9180
gtccgacagg agtgggcca cggtccgcc gacgtcgtgc accgcatggc cctcgagcgc	9240
atgtacttcc atcgccgtcc gggccgccag gtctgggtcc acggccgcct ccgaccgggt	9300
ggcggcgccct tcaccaaggc cctcgccggc gacctcctcc tcttcgagga caccggacag	9360
gtcgtcgccg aggtccaggg cctccgcctc ccgcagctcg aggcctccgc cttcgccccg	9420
cgcgaccgcg gcgaggagt gctctacgca ctcgagtggc agcgcaagga cccgatcccg	9480
gaggccccgg ccgccgcctc gtctcctcc gccggcgccct ggctcgtcct catggaccag	9540
ggcggcaccc gcgccgccct cgtatccctc ctcgagggcc gcggcgaggc ctgcgtccgc	9600
gtcatcgccg gcaccgccta cgctgcctc gccccgggcc tctaccaggt cgacccggcc	9660
cagccggacg gcttccacac cctcctcaga gacgccttcg gcgaagaccg catctgccgc	9720
gccgtcgtcc acatgtggtc cctcgacgcc accgccgccg gcgagcgcg caccgccgag	9780
tccctccagg ccgaccagct cctcggctcc ctctccgcc tctccctcgt ccaggccctc	9840
gtccgccgcc gctggcgcaa catgccgcgc ctctggctcc tcacccgcgc cgtccacgcc	9900
gtcggcgccg aagacgccgc cgctccgtc gcccaggccc cggcttgggg cctcggccgc	9960
acctcgcgc tcgagcacc ggagctccgc tgcacactcg tcgacgtcaa cccggccccg	10020
tccccggagg acgccgcgc cctcgccgtc gagctcggcg cctccgaccg cgaagatcag	10080
gtcgcctccc gctccgacgg ccgtacgtc gccgcctcg tccgtctctc cttctccggc	10140
aagccggcta cagactgcgg catccgcgcc gacggctcct acgtcatcac cgacggcatg	10200
ggccgtgtcg gcctctcgt cgcacagtgg atggtcatgc agggcgcccc ccacgtcgtc	10260
ctcgtcgacc gcggcggcgc ctccgaggcc tcccgcgacg ccctccgcag tatggccgag	10320
gccggagccg aagtccagat agtcgaagcc gacgtcgcgc gccgcgacga cgtggccccg	10380
ctcctctcca agatcgaacc gtccatgccg ccgtccgcg gcatcgtcta cgtcgacggc	10440
accttccagg gcgactctc catgctcgag ctcgacgccc gccgcttcaa ggagtggatg	10500
tatccgaagg tcctcggcgc ctggaacctc cagccctca cccgcgaccg ttcctcgcac	10560
ttcttcgtcc tctactctc cggcacctcc ctctcggcc tcccgggcca gggctcccgc	10620
gccgccggcg acgtttcct cgacgccatc gccaccacc gctgcaaggt cggcctcacc	10680

gccaatgtcca tcaactgggg cctcctctcc gaggcctcct ccccgccac cccgaacgac 10740

ggcggcgccc gcctcgagta ccgcgcatg gagggcctca ccctcgagca gggcgccgcc 10800

gcctcggac gcctcctcgc ccgcccgcgc gcacaggctg gcgtcatgcg cctcaatttg 10860

cgtcagtggc tcgagttcta cccgaacgct gcccgccctg cctcctgggc cgagctcctc 10920

aaggaacgcg accgcgcga ccgcgcgcc tccaacgcct ccaacctccg cgaggccctc 10980

cagtccgccc gcccgagga ccgccaactc atcctcgaga agcacctctc cgagctcctc 11040

ggccgcggcc tccgctccc gccggagcgc atcgagcgtc acgtcccgtt ctccaacctc 11100

ggcatggact cctcatcgg cctcgagctc cgcaaccgca tcgaggccgc tctcggcac 11160

accgtcccgg ccacctcct ttggacctac cctaacgtg ccgccctctc cggtcctcctc 11220

ctcgacatcc tcttcccgaa cgccggcgcc acccacgcc cggccaccga gcgcgagaag 11280

tccttcgaaa acgacgcgc cgacctgaa gccctccgcg gcatgaccga cgagcagaag 11340

gacgccctcc tcgccgagaa gctcgcccag ctcgcccaga tcgtcgggga gtga 11394

<210> 6

<211> 7320

<212> DNA

<213> Artificial Sequence

<220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in Streptomyces venezuelae

<400> 6

atggccacca ccaacgccgg caagctcgag cacgccctcc tctcatgga caagctcgcc 60

aagaagaacg cgtccctcga gcaggaaagg accgagccga tcgcaatcgt cggcatcggc 120

tgccgcttcc cgggaggcgc cgacaccccg gaggccttct gggagctcct cgactccggc 180

cgcgacgcgg tccagccgct cgaccgcgc tgggccctcg tcggcgcca cccgtccgag 240

gaggtaccgc gctgggccgg cctcctcacc gaggccgtgg acggattcga cgccgccttc 300

ttcggcacct ccccgcgca ggcccgtcc ctcgaccgc agcagcgct cctcctcgag 360

gtcacctggg agggcctcga ggacgccgc atcgccccgc agagtctcga cggatccgc 420

accggcgtct tctcggcgc ctgctcctcc gactactccc acaccgtcgc ccagcagcgc 480

cgcgaggagc aggacgccta cgacatcacc ggcaacaccc tctccgtcgc cgccggccgt 540

ctctcctaca cctcgggtct ccagggcccg tgcctcaccg tcgacaccgc ctgctcctcc 600

tccctcgtcg ccatccacct cgctgccgg tccctccgcg cccgcgagtc cgacctcgcc 660

ctcgccggag gcgtcaacat gctcctctcc tccaagacca tgatcatgct cggcaggatc 720

cagggcctct ccccgacgg ccactgccga acctcgacg cctccgcaa cggttcgtc	780
cgggcgagg gctcgggcat ggtcgctctg aagcgctct cgcacgcca ggcacaggc	840
gaccgcatct gggccctcat ccgcggtcc gccatgaacc aggacggcgg ctcaccggc	900
ctcatggccc cgaacgtcct cgcccaggag gccctcctcc gcgaggcct ccagtcggc	960
cggtcgacg ccggcgccat cggctacgtc gagaccacg gcactggcac ttcgtcggc	1020
gacccgatcg aggtcgaggc cctccgccc gtctcggcc cggcccgcg cgatggctcc	1080
cgctcgctcc tcggcgccgt gaagaccaac ctcgccacc tcgaggcgcg cgccggcgctc	1140
gcgggcctca tcaaggccgc cctcgccctc caccacgaac tcctcccgcg caacctccac	1200
ttccacaccc tcaaccgcg catccgcatc gagggcacgg ccctcgccct cgccaccgag	1260
ccgttccgt ggcccgcgcg cgcccgcccg cgcttcggcg gcgtctcgc attcggcctc	1320
tccggcacca acgtccacgt cgctctcgag gagggcccg ccaccgtcct cgccccggc	1380
acccggggc gctccgcga gctcctcgtc ctctccgca agtcggcgcg cgccctcgac	1440
gcccaggcgg cagcctctc cgcacacatc gccgcctacc cggagcaggg cctcggcgac	1500
gtcgcttct ccttggtctc caccgcctcc ccgatggaac atcgctcgc cgtcgccgc	1560
accagccgcg aggccttcgg ctcgcctcc gaggtcgcg cccagggcca gacccggcc	1620
ggtgccgccc gggcccgcg cgctcctcc ccgggcaagc tcgcttctc cttcgcccga	1680
cagggggccc aagtccggg catgggcgc ggcctctggg aggcctggcc ggccttcgc	1740
gaaaccttcg accgtgcgt caccctctc gaccgcgagc tcaccagcc gctctcgag	1800
gtcatgtggg ccgaaccggg ctctcccgcg tcctccctcc tcgaccagac cgcttcacc	1860
cagccggccc tcttcgctc cgagtacgc ctcgcccgc tcttcgctc ctggggcgctc	1920
gagccggagc tcgtcgccgg ccactccctt ggagaactcg tcgccgctg cgtcgccggc	1980
gtcttctccc tcgaggatgc cgtccgctc gtctcgccc gggccgctc catgcaggcc	2040
ctcccgccg cgggcgccat ggtctccata cgggcgccgg aagctgacgt agccggcgcc	2100
gtcgccccgc acgcccctt ggtctccatc gccgccgta acggcccgga gcaggtcgtc	2160
atcgccggcg ccgagaagt cgtccagcag atcgccgccc ccttcgccc cgcgggcgcc	2220
cgcaccaagc cgtccacgt gtcaccgccc ttccactccc cgctcatgga cccgatgctc	2280
gaggccttca gacggtcac cgaatccgta acctaccgcc gaccgtccat cgcgtcgtc	2340
tccaacctct ccggcaagcc gtgcaccgat gaggtctccg cccccggcta ctgggtcgc	2400
cacggccgcg aggcggtccg cttcgccgac ggcgtcaagg ccctccacgc cgccggcgcc	2460
ggcctcttcg tcgaggtcgg ccgaagccg acctcctcg gcctcgtccc ggctgcctc	2520
ccggacgccc gcccggtcct cctccggccc tcccgcgccc gccgcgacga ggccgctcc	2580

gccctcgaag ccctcggcgg attctgggtc gtcggcggt cggtcacctg gagcggagtc	2640
tccccctccg gcggccgcgc cgtcccgtc ccgacctacc cgtggcaacg cgagcgctac	2700
tgatcgagg ccccggtcga ccgcgaggcc gacggcaccc gccgcgccgc cgccggcggc	2760
caccgcgtcc tcggcgaggt cttctccgtc tccaccacg ccggcctccg cctctgggag	2820
acgacctcg accgcaagcg cctcccgtgg ctccggcgagc accgcgccca gggtagagtt	2880
gtcttcccgg gcgcggccta cctcgagatg gccctcagct ccggcgccga gatcctcggc	2940
gacggcccca tccaggtcac cgacgtcgtc ctcatcgaga ccctcacctt cgccggcgac	3000
acagccgtcc cggtcagggt cgtcaccacc gaggagcgcc cgggcccact ccgtttccag	3060
gtcgctccc gcgagccggg cgagcgccgc gccccgttcc gcatccacgc ccgcggcgtc	3120
ctccgccga tcggccgcgt tgagaccccg gcccgctcca acctcgccgc cctccgcgc	3180
cgctccacg ccgcgcgtccc ggccgccgcc atctacggcg ccctcgccga gatgggcctc	3240
cagtacggcc cggcccttcg ccgcctcgc gagctctggc gcggcgaggg cgaggccctc	3300
ggccgcgtcc gtctcccga ggcagccggc tccgccaccg cctaccagct ccaccggtc	3360
ctctcgacg cctgcgtcca gatgatcgtc ggcccttcg ccgaccgcga cgaggccacc	3420
ccttggggccc cggtaggagt ccgctccgtc cgctcttcc agcgagccc gggcgagctc	3480
tggtgccacg cccgcgtcgt ctccgacggc cagcaggcct cctccagggt gtccgccac	3540
ttcgagctca tggacggaac ccggcccgtc gtcgccgaga tctcccgcct cgtcgtcgag	3600
cgctcgctt caggagtcgc ccgcgcgcac gccgacgact ggttcctcga actcgactgg	3660
gagccggccg ccctcggcgg cccgaagatc accgccgtc gctggctcct cctcggcgag	3720
ggcggcgccc tcggcgctc cctctgtcc gccctcaagg ccgcggcca tgtctcgtc	3780
caccccgccg gcgacgacac ctccaccgcc ggcatgcgc ccctcctcgc caacgccttc	3840
gacggacagg ccccgaccgc agtcgtccac ctctcctccc tcgacggcgg cggacagctc	3900
ggcccgggcc tcggcgccca gggcgccctc gacccccga ggtccccga cgtcgacgcc	3960
gacgtctcgc agtccgcct catgaggggc tgcgactccg tcctctctct cgtccaggcg	4020
ctcgtcggca tggacctccg caacgccccg cgctctggc tcctcaccgc cggcgcccag	4080
gcccccgccg ccggcgacgt ctccgtcgtc caggccccgc tcctcggcct cggccgcacg	4140
atcgccctcg agcagccga actcaggtgt atctccgtc acctcgacce ggccgagccg	4200
gagggcgaag ccgacgcgt cctcgccgag ctctcgcgc acgacgtga agaggaggtc	4260
gccctccgcg gcggcgaccg cctcgtcgcc cgctcgtcc accgcctccc ggacgcccag	4320

cgccgcgaga aggtcgagcc ggccggcgac cgcccgttcc gcctcgagat cgacgagccg	4380
ggcgccctcg accagctcgt cctccgcgcc acggggccgcc gagccccggg cccgggcgag	4440
gtcgagatct ccgtcgaggc cgccggcctc gactccatcg acatccagct cgccctcggc	4500
gtcggcccga acgacctccc gggcgaggag atcgagccgc tcgtcttggg ctccgagtgc	4560
gccggccgca tcgtcgccgt cggcgagggc gtcaacggcc tcgtcgtcgg ccagcctgtc	4620
atcgccctcg ccgccggcgt cttcgctacc cacgtacca cctccgccac cctcgtctc	4680
ccgccccgc tcggcctctc cgccacggag gccgcccca tgccgctcgc ctacctacc	4740
gcctggtacg cctcgacaa ggtcgccac ctccaggccg gcgagcgct cctcatccac	4800
gctgaggccg gcggcgtcgg cctctgcgcc gtccgctggg ccagcgcggt cggcgccgag	4860
gtctacgcca ccgccgacac ccccgaaaac cgcgcctacc tcgagtcctt cggcgtccgc	4920
tacgtctccg actcccgctc cgggcgcttc gtaccgacg tccacgcctg gaccgacggc	4980
gaggcgctcg acgtcgtcct cgactccctc tccggtgagc gcatcgacaa gtcctcatg	5040
gtcctccgcg cctcgggcgc cctcgtcaag ctcgccgcc gcgacgactg cgccgacacc	5100
cagccgggac tcccgcgct cctccgcaac ttctcttct cccaggtcga cctccgcggc	5160
atgatgctcg accagccggc ccgcatccgc gccctcctcg acgagctctt cggcctcgtc	5220
gcccgccgcg ccatctcccc gctcggctcc ggctccgcg tcgggggctc cctcaccccg	5280
ccgccggtcg agaccttccc gatttccgc gccgcagagg ccttccgccg catggcccag	5340
ggccagcacc tcggaagct cgtctcacc ctcgacgacc cggagggtccg catccgcgcc	5400
ccggccgagt cctcggctgc cgtccgcgcc gacggcacgt acctcgtcac cggcgggctc	5460
ggcgccctcg gcctccgggt cgccggctgg ctgcccagc gccgcgccgg ccagctcgtc	5520
ctcgtgggcc gctccggcgc cgctccgcc gagcagcgc ccgccgtcgc cgccctcgag	5580
gcccacggcg ccgcgctcac cgtcgccaag gccgacgtcg ccgaccgctc ccagatcgag	5640
cgctcctcc gcgaggctac cgctccggc atgccgtcc gcggcgtcgt ccacgccgcc	5700
ggcctcgtcg acgacggcct cctcatgcag cagaccccg cccgcttccg caccgtaatg	5760
ggcccgaagg tccagggcgc cctccacctc cacacctca cccgcgaggc cccgctctcc	5820
ttcttctgccc tctacgctc cgccgtggc cttttcggtt ccccgggcca gggaaactac	5880
gcccgccca acgcttctct cgacgcctc agtcaccacc gccgcgcca gggcctcccg	5940
gccctctcca tcgactgggg catgttcacg gaggtcggca tgccgctcgc ccaggaaaac	6000
cgcggtgccc gccagatctc ccgcgcatg cgcgccatca cgccgatga aggcctctcc	6060

gcccctcgccc gcctcctcga gggcgaccgc gtccagaccg gcgtaatccc gatcaccccg 6120
cgccagtggg tcgagttcta cccggccacc gccgccagcc gccgcctctc ccgcctcgtc 6180
acgacgcagc gcgccgtcgc cgaccggacc gccggcgacc gcgacctcct cgagcagctc 6240
gtttccgccc agccgtccgc ccgcgccggc ctctccagg acgtcgtccg cgtccaggtc 6300
tcccacgtcc tccgcctccc ggaggacaag atcgaagtcg acgccccgct ctctccatg 6360
ggcatggact cccttatgtc cctcgagctc cgcaaccgca tcgaggccgc cctcggcgtc 6420
gccgcaccgg cagccctcgg ctggacctac ccgaccgtcg ccgccatcac ccgtggctc 6480

ctcgacgacg ccctcgtcgt ccgtctcggc ggcggtccg acaccgacga gtccaccgcc 6540
tccgccggct ccttcgtcca cgtcctcgc ttcgtcccg tcgtcaagcc gcgcgccgg 6600
ctcttctgct tccatggctc cgccgggtcc ccggagggtc tccgtcctg gtccgagaag 6660
tccgagtggc ccgacctga gatcgtgcc atgtggcacg accgctcct cgctccgag 6720
gacccccgg gcaagaagta cgtccaggaa gccgcctccc tcatccagca ctacccgac 6780
gccccgttcg ccctcgtcgg ctctccctc ggctccgct tcgtcatggg caccgccgtc 6840
gagctcgct cccgcagcgg cgccccggcc ccgtcgcg tcttcacct cggcggttcg 6900

ctcatctcct cctccgagat caccgccgag atggaaaccg atatcatcgc caagctcttc 6960
ttccgcaacg ccgcgggttt cgtccgatcc acccagcagg tccaggccga cgccccgcc 7020
gacaaggcca tcaccgacac catggtcgcc ccggccccgg gcgactcaa ggaaccgccg 7080
gtcaagatcg ccgtcccaat cgtcgccatc gccggctccg acgacgtcat cgtccgccg 7140
tccgacgtcc aggacctcca gtccgcacc accgagcgt tctacatgca cctcctccg 7200
ggcgaccacg agttcctcgt cgaccgtggc cgcgagatca tgcacatcgt cgactccac 7260
ctcaaccgcg tctcgcgcg ccgcaccacc tctccggcc ccgtttcga ggccaagtga 7320

7320

<210> 7

<211> 1260

<212> DNA

<213> Artificial Sequence

<220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in
Streptomyces venezuelae

<400> 7

atgaccagg agcaggcga ccagtcgag acgaagccg ccttcgactt caagcccttc 60
gccccgggt acgccagga ccccttccc gccatcgagc ggctccgca ggccaccccg 120

atcttctact gggacgaggg ccggagctgg gtgtcaccc gctaccacga cgtctccgcc 180

gtgttcaggg acgagcgctt cgccgtctcc cgggaggagt gggaatcttc ggccgagtac 240

agctccgcca tcccggagct ctccgacatg aagaagtacg ggctcttcgg cctgcccccc 300

gaggaccacg cccgctcgcg caagctcgtg aaccgctcct tcacctcgcg ggccatcgac 360

ctctccgcg cggaatcca gcgcaccgtc gaccagctcc tcgacgcccg gtccggccag 420

gaggagtctg acgtcgtcgc cgactacgcg gaaggcatcc cgatcggggc gatctccgcg 480

ctctcaagg tcccgccga gtgcgacgag aagttccgcc gcttcggctc cgccacggcc 540

cgcgcactcg gcgtcgccct cgtcccgag gtgcgacagg agaccaagac cctcgtcgcc 600

tcggtcaccc agggcctcgc cctcctgcac gacgtcctgg acgaacgccg ccgcaaccgc 660

ctggagaacg acgtcctcac gatgtctctc caggccgagg cggacggctc gcgcctctcc 720

accaaggagc tggtcgtctt cgtcggcgcc atcatcgccg ccggcaccga caccaccatc 780

tacctcatcg ccttcgccgt cctcaacctc ctccgctcgc ccgaggccct cgagctcgtc 840

aaggcggagc ccggcctcat gcggaacgcc ctgcgacagg tgcctcgctt cgacaacatc 900

ctccgcatcg gcacggtcgc gttcgccgcg caggatctgg agtactgcgg cgcgtcgatc 960

aagaaggcgc agatggtctt cctcctcatc ccgagcgccc tccgcgacgg aaccgtcttc 1020

tcccccccg acgtcttcga cgtgcgtcga gacacggggg cgtccctcgc gtacggccgg 1080

ggcccgacg tctgccccgg cgtctcctc gcccgccctc aggccgagat cgccgtcggc 1140

accatcttcc gccgttccc ggagatgaag ctgaaggaaa cgccggtgtt cgggtaccac 1200

cccgcttcc ggaacatcga gtccctgaac gtcactctga agccgtcaa ggccgggtga 1260

1260

<210> 8

<211> 555

<212> DNA

<213> Artificial Sequence

<220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in

Streptomyces venezuelae

<400> 8

atgacgtcgc aggaggtccc cggcgcccgc ctggcgcccc agtcgtccct cgtccgggcc 60

cagcacgccg cgcgccacgt ccgcccgtgc acccgcgccg aggagccgcc cgcgctcatg 120

cacggactca ccgagcgcca ggtcctctc tcgctcgtgg ccctggccct cgtgtgtctc 180

accgccgtg cgttcggcga actcgcccgg cgcctccgcc agcccagagt cctcggcgag 240

ctcttcggcg gcgtcgtcct gggtcccagc gtggtcgggg cgctcgcgcc gggcttccac 300
 cgggtcctct tccaggaccc cgcggtcggc gtcgtcctct ccggcatctc ctggatcggg 360

gccctcgtgc tgctcctcat ggcgggcatc gaggtggacg tctccatcct ccgcaaggag 420
 gcccggccgg gcgccctctc ggccctcggc gccatcgccc cgccgctccg caccgccggc 480
 cccctcgtcc agcgcgatga gggggccttc acgtgggacc tggacgtcag cccgcgccgc 540
 agcgcgcagg cctga 555

<210> 9
 <211> 2142
 <212> DNA
 <213> Artificial Sequence
 <220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in
 Streptomyces venezuelae
 <400> 9

atgcacgggc tgacggagcg ccaggctctc ctctcgtcgc tcacgtcgc cctcatcctc 60
 gtcaccgccc gcgcgtccgg cgagctcgcc cgccgcctcc gccagccgga agtctcggc 120
 gagctcttcg gcggcgctgt cctcggcccg agcgtcgtcg gagecctcgc cccgggcttc 180
 caccgcgccc tcttcaggga acccgctgtc ggcgctcgtc tctccggcat ctcttgatc 240
 ggcccccctc tctcctcct catggccggc atcgaggctg acgtcggcat cctccgcaag 300
 gaggcgcgcc cgggcgcctt ctccgccctc ggccccatcg cccccccgt cgccgcgggc 360
 gcagcgttct cggccctcgt cctcgaccgc cccctccgt cggcctctt cctcgggac 420

gtactctcgc tcaccgcgt ctccgtcatc gccaaaggctc tcacgagcg cgagtccatg 480
 cgccgctcct acgccagggt gacctcgcgc gcggcgctcg tctccgaggt cgccgcctgg 540
 gtctcgtcgc ccatgacctc cagctcctac ggggcctccc cgccctcgc cgtcgccgc 600
 tccgccctcc tcgcctccgg ctctctctc ttcattgtcc tcgtcggctg ccgcctcacg 660
 cacctggcca tgcgtgggt cgccgacgcc acccgctct ccaaggcca ggtctccctc 720
 gtctcgtcgc tcacgttctt cgccgccgcc ctacccagc gcctcggctt ccaccgctg 780
 ctggggcct tcgcctcgg ggtctctctc aactccgac cgcgaccaa ccgccgctc 840

ctcgacggcg tccagacct cgtggccgggt ctcttcgccc ccgtcttctt cgtgctcgcc 900
 ggaatgaggg tcgacgtcag ccagctcgc acgccggccg cctggggcac cgtcgccctc 960
 ctctcgcga ccgccaccgc cgccaaggct gtcccgccg ccctcggcgc ccgtctcggc 1020
 ggccctccgc gctcggaggc cgccctcgt gccgtcggtc tcaacatgaa gggcggcacc 1080

gacctcatcg tcgcgatcgt cggagtcgag ctccgctccc tctccaacga ggcctacacc	1140
atgtacgccg tcgtcgccct cgtcaccgtc accgctccc ccgcccctct catctggctc	1200
gagaagcgcg ccccgccac ccaggaagag tccgcccgc tcgagcgga ggaggccgc	1260
cgcgcgctt acatccggg cgtcgagcgc atctctgtcc ccatcgtcgc ccacgcccctc	1320
ccgggggttcg ccacggacat cgtcgagtc atctctgctt ccaagcgga gctcggggag	1380
accgtcgaca tcaccgagct ctccgtcgag cagcaggccc cgggcccgtc ccgcccgcgc	1440
ggcgaggcct ccccgggcct cgccaggctc ggccccggc tccgctcgg catctggcgc	1500
cagcgccgcg agtcctgtg ctccatccag gccatctcc gtgctcccg cgaccagac	1560
ctctctgtca tcggcgcccg ctccccggc cgccccgtg gcatgtctt cggccgctg	1620
caggacgcca tcgtccagcg cgccgaaagc aacgtctctg tcgtctcgg cgaccgccc	1680
gcccgcgagc gcgctccgc ccgcccgcgc ctctctccga tcacggcct cgagtacagc	1740
ttcgcccgcg ccgacctgc cgccacgtc gccctgcct gggacgcgga gctcgtctc	1800
ctctctccg cccagacgga cccgggcgc gtcgtctggc gcgaccgga gccgtcgcgc	1860
gtccgcgcg tggcgcgtc cgtcgtcgc gagccgtct tccgcggtc ccgctcggc	1920
gtccgcgtct ccagccgagt ccacgtcggc gccaccctg ccgacgagat caccgcgag	1980
ctgccccgcg cgccctacga ctgctcgtc ctgggtgct acgaccagc cccctcggc	2040
cgctctacc tcggtaccac cgtcgaatcc gtagtctcc gctcccgcg cccggtggc	2100
ctctctgtgg cccacgtgg caccgcgag caggtccgt ga	2142
<210> 10	
<211> 918	
<212> DNA	
<213> Artificial Sequence	
<220><223> Codon Optimized DNA of Sorangium cellulosum for Expression in	
Streptomyces venezuelae	
<400> 10	
atggagaagg agtcgcgat cgccatctac ggcccatcg ccgccaacgt cgccatcgcc	60
gccgtcaagt tcacgcgcg ggccgtcag ggctcgtcgg cgatgctcag cgagggcgtc	120
cactccctcg tggacaccgc cgacggctc ctctctctcc tcggcaagca ccgacgcgcg	180
cgcccccgcg acgcccagca ccccttcggc caggcaagg agctctactt ctggaccctc	240
atcgtggcca tcatgatctt cgccggccgc ggccgctct ccatctacga gggcatctc	300
cactctctcc acccccgcga gatcgaggac ccgacctgga actacgtcgt cctcggggcc	360

gcagcgggtgt tcgaaggcac ctccctgac atctccatcc acgagttcaa gaagaaggac	420
gggcagggct acctcgcggc catgcgctcc agcaaggacc ccaccacctt cacgacgtc	480
ctcgaggaca gcgccgcgct ggccggcctc acgacgcgct tcctcggcgt ctggctcgga	540
caccgcctcg gcaaccgta cctggacggc gccgcgtcca tcggcatcgg gtcgctctc	600
gccgcggtgg ccgtcttctt cgctcccag tcccgcggcc tgctcgtcgg cgagtcgcc	660
gaccgggagc tcctcgccgc catccgcgcc ctggcctccg cggacccggg cgtgtccgcc	720
gtggggcgcc cgctcacat gcacttcggc ccgcacgagg tcctcgtcgt gtcccgac	780
gagttcgacg cggcgctcac cgctccggc gtggccgagg ccatcgagcg tatcgagacc	840
cgcatccggt ccgagcgccc cgacgtcaag cacatctacg tcgaggcgcg ctccctccac	900
cagcgcgccc gcgcgtga	918