



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0014673
(43) 공개일자 2013년02월08일

(51) 국제특허분류(Int. Cl.)
C12N 1/21 (2006.01) C12P 1/04 (2006.01)
C12N 15/31 (2006.01) C12N 15/70 (2006.01)
(21) 출원번호 10-2012-0083186
(22) 출원일자 2012년07월30일
심사청구일자 2012년07월30일
(30) 우선권주장
1020110075715 2011년07월29일 대한민국(KR)

(71) 출원인
경상대학교산학협력단
경상남도 진주시 진주대로 501 (가좌동)
(72) 발명자
김선원
경상남도 진주시 내동면 독산리 휴먼빌아파트
101-1404
장희정
경상남도 진주시 주악동 111-10 305호
(뒷면에 계속)
(74) 대리인
이준호, 박국진, 장영태, 노준태

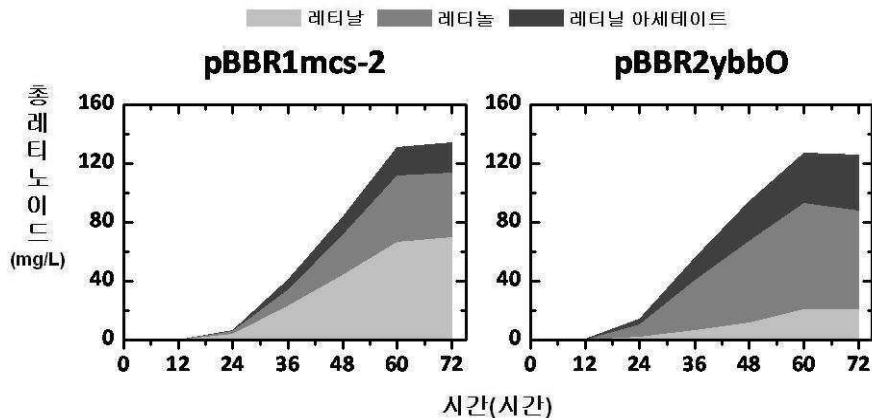
전체 청구항 수 : 총 33 항

(54) 발명의 명칭 Y b b O 유전자가 결실 또는 증폭된 에세리키아 속 미생물 및 그를 이용한 레티노이드의 생산 방법

(57) 요약

본 발명은 YbbO 유전자가 결실 또는 증폭된 에세리키아 속 미생물 및 그를 이용한 레티노이드의 생산방법에 관한 것으로서, 보다 상세하게는 레티노이드 생산능을 갖는 에세리키아(Escherichia) 속 모균주의 YbbO 단백질을 코딩하는 유전자가 감쇄 또는 결실되거나, 증폭된 에세리키아 속 미생물; 이 미생물을 배양시켜 얻어진 배양물로부터 레티노이드를 분리하는 것을 포함하는 레티노이드의 생산 방법에 관한 것이다. 이러한 발명들은 모두 YbbO 단백질의 레티날과 레티놀에 대한 효소 활성을 확인한 것에 기초한다.

대표도 - 도3



(72) 발명자

윤상환

충청남도 부여군 부여읍 동남리 268-2 신안빌라
B-202

하보경

경상남도 진주시 문산읍 삼곡리 고려아파트 504호

특허청구의 범위

청구항 1

레티노이드 생산능을 갖는 에세리키아(Escherichia) 속 모균주의 YbbO 단백질을 코딩하는 유전자가 감쇄 또는 결실, 또는 증폭되거나,

상기 유전자의 뉴클레오티드 서열과 80% 이상의 상동성을 갖는 외래 유전자가 도입되어 증폭된 에세리키아 속 미생물.

청구항 2

청구항 1에 있어서, 상기 YbbO 단백질을 코딩하는 유전자는 서열번호 24의 뉴클레오티드 서열을 갖는 것인, 미생물.

청구항 3

청구항 1에 있어서, 상기 외래 유전자는 서열번호 49 내지 56 중 어느 하나의 뉴클레오티드 서열을 갖는 것인, 미생물.

청구항 4

청구항 1에 있어서, 상기 YbbO 단백질은 서열번호 23의 아미노산 서열을 갖는 것인, 미생물.

청구항 5

청구항 3에 있어서, 상기 외래 유전자는 서열번호 41 내지 48 중 어느 하나의 아미노산 서열을 갖는 단백질을 코딩하는 것인, 미생물.

청구항 6

청구항 1에 있어서, 상기 미생물은 대장균인 미생물.

청구항 7

청구항 6에 있어서, 상기 대장균은 DH5 α, MG1655, BL21(DE), S17-1, XL1-Blue, BW25113 또는 이들의 조합인, 미생물.

청구항 8

청구항 1에 있어서, 상기 YbbO 단백질을 코딩하는 유전자의 감쇄 또는 결실에 의해 상기 모균주에 비해 레티날의 생산능이 향상된, 미생물.

청구항 9

청구항 1에 있어서, 상기 YbbO 단백질을 코딩하는 유전자 또는 상기 유전자의 뉴클레오티드 서열과 80% 이상의

상동성을 갖는 외래 유전자의 증폭에 의해 상기 모균주에 비해 레티놀의 생산능이 향상된, 미생물.

청구항 10

청구항 1에 있어서, 상기 모균주의 에탄올아민 이용 단백질 E (Ethanolamine utilization protein E: EutE)를 코딩하는 유전자; 푸트레신 이용 경로 단백질 C (Putrescine utilization pathway protein C: PuuC)를 코딩하는 유전자; 또는 이들의 조합이 추가로 감쇄 또는 결실된, 미생물.

청구항 11

청구항 10에 있어서, 상기 에탄올아민 이용 단백질 E (EutE)는 서열번호 19의 아미노산 서열을 갖는 것인, 미생물.

청구항 12

청구항 10에 있어서, 상기 푸트레신 이용 경로 단백질 C (PuuC)는 서열번호 20의 아미노산 서열을 갖는 것인, 미생물.

청구항 13

청구항 10에 있어서, 상기 에탄올아민 이용 단백질 E (EutE)를 코딩하는 유전자는 서열번호 21의 뉴클레오티드 서열을 갖는 것인, 미생물.

청구항 14

청구항 10에 있어서, 상기 푸트레신 이용 경로 단백질 C (PuuC)를 코딩하는 유전자는 서열번호 22의 뉴클레오티드 서열을 갖는 것인, 미생물.

청구항 15

청구항 1에 있어서, 상기 레티노이드는 레티날, 레티놀, 레티닐에스테르 및 레티노산으로 이루어진 군에서 선택되는 적어도 하나인, 미생물.

청구항 16

청구항 1에 있어서, 상기 모균주는 서열번호 1의 엔테로코커스 패칼리스 (Enterococcus faecalis) 유래의 아세틸-CoA 아세틸트랜스퍼라제/하이드록시메틸글루타릴 (HMG)-CoA 리덕타제를 코딩하는 유전자; 서열번호 2의 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제를 코딩하는 유전자; 서열번호 3의 스트렙토코커스 뉴모니아 (Streptococcus pneumoniae) 유래의 메발로네이트 키나제를 코딩하는 유전자; 서열번호 4의 스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제를 코딩하는 유전자; 서열번호 5의 스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 유전자; 서열번호 6의 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 유전자; 서열번호 7의 판토에아 아글루메란스 (Pantoea agglomerans) 유래의 제라닐제라닐 피로포스페이트 (GGPP) 신타제를 코딩하는 유전자; 서열번호 8의 판토에아 아글루메란스 유래의 피토엔 신타제를 코딩하는 유전자; 서열번호 9의 판토에아 아글루메란스 유래의 피토엔 데히드로게나제를 코딩하는 유전자; 및 서열번호 10의 판토에아 아나나티스 (Pantoea ananatis) 유래의 라이코펜-β-시클라제를 코딩하는 유전자로 형질전환된 것인, 미생물.

청구항 17

청구항 16에 있어서, 상기 모균주는 서열번호 13의 배양되지 않은 해양 박테리아 (uncultured marine bacterium) 66A03 유래의 β -카로틴 모노옥시게나제를 코딩하는 유전자; 서열번호 14의 생쥐 (*Mus musculus*) 유래의 β -카로틴 15,15'-모노옥시게나제를 코딩하는 유전자; 서열번호 15의 나트로노모나스 파라오니스 (*Natronomonas pharaonis*) ATCC35678 유래의 brp 유사 단백질 2 (brp-like protein 2)를 코딩하는 유전자; 및 서열번호 16 또는 17의 할로박테리움 살리나룸 (*Halobacterium salinarum*) ATCC700922 유래의 β -카로틴 모노옥시게나제를 코딩하는 유전자로 이루어진 군으로 선택되는 하나 이상의 유전자로 더 형질전환된 것인, 미생물.

청구항 18

청구항 16에 있어서, 상기 모균주는 대장균에서 코돈 사용 최적화된 서열번호 18의 염기서열을 갖는 유전자로 더 형질전환된 것인, 미생물.

청구항 19

청구항 1에 있어서, 상기 모균주는 서열번호 11의 대장균 유래 1-데옥시자일롤로즈-5-포스페이트 (DXP) 신타제를 코딩하는 유전자로 형질전환된 것인, 미생물.

청구항 20

청구항 1에 있어서, 상기 모균주는 서열번호 12의 헤마토코커스 플루비알리스 (*Haematococcus pluvialis*) 유래의 IPP 이소머라제를 코딩하는 유전자로 형질전환된 것인, 미생물.

청구항 21

청구항 1에 있어서, 상기 모균주는 기탁번호 KCTC 11254BP의 대장균 DH5 α /pTDHB/pSNA 또는 기탁번호 KCTC 11255BP의 대장균 DH5 α /pTDHBSR/pSNA인, 미생물.

청구항 22

청구항 1 내지 21 중 어느 항의 미생물을 배양하는 단계; 및
배양물로부터 레티노이드를 분리하는 단계;를 포함하는,
미생물로부터 레티노이드를 생산하는 방법.

청구항 23

청구항 22에 있어서, 상기 배양은 친유성 물질을 포함하는 배지 중에서 수행되는 것인, 방법.

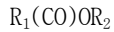
청구항 24

청구항 22에 있어서, 상기 분리는 친유성 물질 상으로부터 되는 것인, 방법.

청구항 25

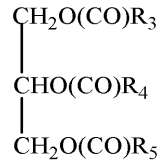
청구항 23 또는 24에 있어서, 상기 친유성 물질은 탄소수 8 내지 50의 알칸 화합물; 하기 화학식 1의 화합물; 하기 화학식 2의 화합물; 또는 이들의 조합인, 방법:

[화학식 1]



(식 중, R_1 및 R_2 는 각각 독립적으로 탄소수 8 내지 50의 알킬을 나타내고, CO는 카르보닐기를 나타냄),

[화학식 2]



(식 중, R_3 , R_4 및 R_5 는 각각 독립적으로 탄소수 8 내지 50의 알킬을 나타내고, CO는 카르보닐기를 나타냄).

청구항 26

청구항 23 또는 24에 있어서, 상기 친유성 물질은 옥탄, 데칸, 도데칸, 테트라데칸, 피토스쿠알란, 미네랄 오일, 이소프로필 미리스테이트, 세틸 에틸헥사노에이트, 디옥타노일 데카노일 글리세롤, 스쿠알란, 또는 이들의 조합인, 방법.

청구항 27

청구항 22에 있어서, 상기 레티노이드는 레티날, 레티놀, 레티닐에스테르 및 레티노산으로 이루어진 군에서 선택되는 적어도 하나인, 방법.

청구항 28

청구항 22에 있어서, 상기 레티노이드는 화장품, 식품 또는 의약품의 소재인, 방법.

청구항 29

Ybb0 단백질; 또는 상기 Ybb0 단백질을 코딩하는 유전자의 아미노산 서열과 80% 이상의 상동성을 갖는 유전자가 코딩하는 단백질을 포함하는 효소 조성물을 첨가함으로써 레티날로부터 레티놀로의 전환을 촉진시키는 방법.

청구항 30

청구항 29에 있어서, 상기 Ybb0 단백질은 서열번호 23의 아미노산 서열을 갖는 것인, 방법.

청구항 31

청구항 29에 있어서, 상기 Ybb0 단백질을 코딩하는 유전자의 아미노산 서열과 80% 이상의 상동성을 갖는 유전자가 코딩하는 단백질은 서열번호 41 내지 48 중 어느 하나의 아미노산 서열을 갖는 것인, 방법.

청구항 32

청구항 29에 있어서, 상기 YbbO 단백질은 서열번호 24의 뉴클레오티드 서열을 갖는 유전자로 코딩되는 것인, 방법.

청구항 33

청구항 29에 있어서, 상기 YbbO 단백질을 코딩하는 유전자의 아미노산 서열과 80% 이상의 상동성을 갖는 유전자가 코딩하는 단백질은 서열번호 49 내지 56 중 어느 하나의 뉴클레오티드 서열을 갖는 유전자로 코딩되는 것인, 방법.

명세서

기술분야

[0001] 레티노이드 생산능을 갖는 에세리키아 속 미생물 및 그 미생물을 이용하여 레티노이드를 생산하는 방법에 관한 것이다.

배경기술

[0002] 레티노이드는 비타민 A와 화학적으로 관련된 친지질성 이소프레노이드 분자의 클래스이다. 레티노이드는 알코올 (예를 들면, 레티놀), 알데히드 (예를 들면, 레티날), 카르복실산 (예를 들면, 레티노산), 또는 에스테르 (예를 들면, 레티닐 아세테이트) 기능기와 함께, β -이온 고리 및 다불포화(polyunsaturated) 결사슬로 구성된다. 이들은 시력보호, 골 발달 및 재생, 항산화 효과와 더불어 피부 노화 방지와 같은 인체 건강에 있어 필수적인 역할을 하고 특정 암의 위험을 감소시킨다고 알려져 있다.

[0003] 레티노이드는 최근 수년간 주름개선 및 피부 질환 치료를 위한 효과적인 화장품 및 의약품 원료로서 큰 관심을 받아 왔다. 레티노이드 시장 규모는 세계적으로 약 160억불 정도로 추정된다. 화학적으로 합성된 레티노이드는 대표적인 상업적 원료이다. 레티놀은 펜타디엔 유도체의 환원에 의해 화학적으로 합성된 레티날의 산성화 또는 가수분해로부터 생산된다. 그러나 이러한 화학적 과정은 복잡한 정제 단계 및 원하지 않는 부산물 형성과 같은 단점을 갖는다. 동물은 과일 및 야채로부터 얻은 카로티노이드로부터 레티노이드를 생산하는 반면, 식물은 레티노이드를 합성할 수 없다. 레티노이드 합성의 전체 경로는 보조기(prosthetic group)로서 레티날을 갖는 박테리 오로돕신 또는 프로테오로돕신을 포함하는 미생물에서만 가능하다. 그러나, 미생물은 레티날의 단백질 결합 형태를 생산하므로 자유 레티노이드의 대량 생산에는 적합하지 않다. 지금까지 생물학적 생산을 위해서 효소를 이용한 일부 제한적인 시도가 있었지만 성공적인 결과는 없었다. 따라서, 대사적으로 형질전환된 미생물을 사용하는, 레티노이드 생산을 위한 생명공학적 방법의 개발이 필요하다.

[0004] 레티노이드는 그의 반응성 있는 콘쥬게이트된 이중 결합으로 인해 화학적으로 매우 불안정하고, 열, 산소 및 빛에 의해 쉽게 산화되고 이성체화된다. 또한 레티노이드는 생물학적으로 레티노산을 통해 쉽게 분해된다. 따라서, 레티노이드를 보다 효율적으로 생산하는 방법이 요구되고 있다.

발명의 내용

해결하려는 과제

[0005] 일 양상은 YbbO 단백질을 코딩하는 유전자가 감쇄 또는 결실, 또는 증폭되거나, 상기 유전자의 뉴클레오티드 서열과 80% 이상의 상동성을 갖는 외래 유전자가 도입되어 증폭된 에세리키아 속 미생물을 제공한다.

[0006] 다른 양상은 이 미생물을 이용하여 레티노이드를 생산하는 방법을 제공한다.

[0007] 또 다른 양상은 YbbO 단백질; 또는 YbbO 단백질을 코딩하는 유전자의 뉴클레오티드 서열과 80% 이상의 상동성을

갖는 유전자에 의해 코딩되는 단백질을 포함하는 효소 조성물로 레티놀 생성을 촉진시키는 방법을 제공한다.

과제의 해결 수단

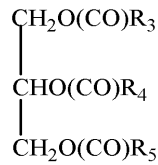
- [0008] 일 양상은 레티노이드 생산능을 갖는 에세리키아(*Escherichia*) 속 모균주의 YbbO 단백질을 코딩하는 유전자가 감쇄 또는 결실, 또는 증폭되거나, 상기 유전자의 뉴클레오타이드 서열과 80% 이상의 상동성을 갖는 외래 유전자가 도입되어 증폭된 에세리키아 속 미생물에 관한 것이다.
- [0009] 모균주는 레티노이드 생산능을 갖는 야생형 에세리키아 속 미생물 또는 형질전환된 에세리키아 속 미생물일 수 있다. 야생형 에세리키아 속 미생물은 내재적 레티노이드 합성의 경로로서 MEP 경로를 가지고 있는 것으로 알려져 있다. 형질전환된 에세리키아 속 미생물은 레티노이드 합성의 내재적 MEP 경로에 연관된 유전자, 외래 MVA 경로에 연관된 유전자, 또는 이들의 조합이 도입된 것일 수 있다. MVA 경로 유전자는 아세틸-CoA로부터 IPP를 생산하는데 관여하는 외래 메발로네이트 경로의 효소를 코딩하는 유전자일 수 있다. 또한, 상기 IPP로부터 β -카로틴을 합성하는데 관여하는 효소를 코딩하는 유전자가 도입된 균주일 수 있다. 두 카피 이상의 IPP 이소머라제가 도입되어 IPP로부터 DMAPP로의 전환이 촉진된 것일 수 있다. 따라서, 상기 미생물은 레티노이드를 고농도로 생산할 수 있다. 도 1은 레티날 생합성의 MEP 경로 및 외래의 MVA 경로를 도식적으로 나타낸 도면이다.
- [0010] 야생형 에세리키아 속 미생물은 예를 들면, 대장균일 수 있다. 상기 대장균은 DH5 α , MG1655, BL21(DE), S17-1, XL1-Blue, BW25113 또는 이들의 조합일 수 있다.
- [0011] 모균주는 예를 들면, 서열번호 1의 엔테로코커스 패칼리스 (*Enterococcus faecalis*) 유래의 아세틸-CoA 아세틸트랜스퍼라제/하이드록시메틸글루타릴 (HMG)-CoA 리덕타제를 코딩하는 유전자, 서열번호 2의 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제를 코딩하는 유전자, 서열번호 3의 스트렙토코커스 뉴모니아 (*Streptococcus pneumoniae*) 유래의 메발로네이트 키나제를 코딩하는 유전자, 서열번호 4의 스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제를 코딩하는 유전자, 서열번호 5의 스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 유전자, 서열번호 6의 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 유전자, 서열번호 7의 판토에아 아글루메란스 (*Pantoea agglomerans*) 유래의 제라닐제라닐 피로포스페이트 (GGPP) 신타제를 코딩하는 유전자, 서열번호 8의 판토에아 아글루메란스 유래의 피토엔 신타제를 코딩하는 유전자, 서열번호 9의 판토에아 아글루메란스 유래의 피토엔 테히드로게나제를 코딩하는 유전자, 및 서열번호 10의 판토에아 아나나티스 (*Pantoea ananatis*) 유래의 라이코펜- β -시클라제를 코딩하는 유전자로 형질전환된 것일 수 있다.
- [0012] 모균주는 서열번호 1 내지 10의 유전자로 형질전환되고, 또한 서열번호 13의 배양되지 않은 해양 박테리아 (uncultured marine bacterium) 66A03 유래의 β -카로틴 모노옥시게나제를 코딩하는 유전자, 서열번호 14의 생쥐 (*Mus musculus*) 유래의 β -카로틴 15,15'-모노옥시게나제를 코딩하는 유전자, 서열번호 15의 나트로노모나스 파라오니스 (*Natronomonas pharaonis*) ATCC35678 유래의 brp 유사 단백질 2 (brp-like protein 2: brp2)을 코딩하는 유전자, 및 서열번호 16 또는 17의 할로박테리움 살리나룸 (*Halobacterium salinarum*) ATCC700922 유래의 β -카로틴 모노옥시게나제를 코딩하는 유전자로 이루어진 군으로 선택되는 하나 이상의 유전자로 더 형질전환된 것일 수 있다.
- [0013] 상기 모균주는 서열번호 12의 헤마토코커스 플루비알리스 (*Haematococcus pluvialis*) 유래의 IPP 이소머라제를 코딩하는 유전자로 더 형질전환된, 레티노이드를 생산하는 것일 수 있다.
- [0014] 상기 모균주는 서열번호 11의 대장균 유래 1-데옥시자일록소스-5-포스페이트 (DXP) 신타제(dxs)를 코딩하는 유전자로 형질전환된 것일 수 있다. DXP는 내재적 MEP 경로에서 속도 결정 단계에 해당하는 효소이므로 DXP 신타제를 코딩하는 유전자가 추가적으로 도입됨으로써 미생물은 β -카로틴을 고농도로 생산할 수 있게 된다.
- [0015] 모균주는 예컨대 기탁번호 KCTC 11254BP의 대장균 DH5 α /pTDHB/pSNA (KOREAN COLLECTION FOR TYPE CULTURE, 2008. 1. 2. 기탁) 또는 기탁번호 KCTC 11255BP의 대장균 DH5 α /pTDHBSR/pSNA (KOREAN COLLECTION FOR TYPE CULTURE, 2008. 1. 2. 기탁)일 수 있다. 특히, 대장균 DH5 α /pTDHBSR/pSNA는 배지 중의 탄소원으로부터 레티노이드를 높은 생산성으로 생산할 수 있다.
- [0016] 일 구체예에서, 상기 미생물은 서열번호 1의 엔테로코커스 패칼리스 (*Enterococcus faecalis*) 유래의 아세틸-CoA 아세틸트랜스퍼라제/하이드록시메틸글루타릴 (HMG)-CoA 리덕타제를 코딩하는 유전자, 서열번호 2의 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제를 코딩하는 유전자, 서열번호 3의 스트렙토코커스 뉴모니아

(*Streptococcus pneumoniae*) 유래의 메발로네이트 키나제를 코딩하는 유전자, 서열번호 4의 스트렙토코커스 뉴모니에 유래의 포스포메발로네이트 키나제를 코딩하는 유전자, 서열번호 5의 스트렙토코커스 뉴모니에 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 유전자, 서열번호 6의 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 유전자, 서열번호 7의 판토에아 아글루메란스 (*Pantoea agglomerans*) 유래의 제라닐제라닐 피로포스페이트 (GGPP) 신타제를 코딩하는 유전자, 서열번호 8의 판토에아 아글루메란스 유래의 피토엔 신타제를 코딩하는 유전자, 서열번호 9의 판토에아 아글루메란스 유래의 피토엔 데히드로게나제를 코딩하는 유전자, 서열번호 10의 판토에아 아나나티스 (*Pantoea ananatis*) 유래의 라이코펜-β-시클라제를 코딩하는 유전자, 서열번호 11의 대장균 유래 1-데옥시자일롤로즈-5-포스페이트 (DXP) 신타제를 코딩하는 유전자 및 서열번호 12의 헤마토코커스 플루비알리스 (*Haematococcus pluvialis*) 유래의 IPP 이소머라제를 코딩하는 유전자로 형질전환된 에세리키아 속 미생물로서, 서열번호 13의 배양되지 않은 해양 박테리아 (uncultured marine bacterium) 66A03 유래의 β-카로틴 모노옥시게나제를 코딩하는 유전자, 서열번호 14의 생쥐 (*Mus musculus*) 유래의 β-카로틴 15,15'-모노옥시게나제를 코딩하는 유전자, 서열번호 15의 나트로노모나스 파라오니스 (*Natronomonas pharaonis*) ATCC35678 유래의 brp 유사 단백질 2 (brp-like protein 2)을 코딩하는 유전자, 및 서열번호 16 또는 17의 할로박테리움 살리나룸 (*Halobacterium salinarum*) ATCC700922 유래의 β-카로틴 모노옥시게나제를 코딩하는 유전자로 이루어진 군에서 선택되는 하나 이상의 유전자로 더 형질전환된 에세리키아 속 미생물일 수 있다. 서열번호 13의 배양되지 않은 해양 박테리아 (uncultured marine bacterium) 66A03 유래의 β-카로틴 모노옥시게나제를 코딩하는 유전자는 대장균에서 코돈 사용 최적화된 서열번호 18의 염기서열을 갖는 것일 수 있다.

- [0017] 본 명세서에 있어서, 용어 "레티노이드 (retinoids)"는 비타민 A에 화학적으로 관련된 화합물의 부류를 나타낸다. 레티노이드의 구조는 시클릭 말단기, 폴리엔 측쇄 및 극성 말단기로 구성된다. 상기 폴리엔 측쇄의 C=C 이중결합이 교대로 되어 형성된 공액 시스템 (conjugated system)은 레티노이드의 색 (보통 노란색, 오렌지 또는 적색)을 띄게 한다. 많은 레티노이드는 발색소(chromophore)이다. 측쇄 및 말단기를 변화시킴으로써 다양한 레티노이드가 생성될 수 있다. 상기 레티노이드는 레티날, 레티놀, 레티노산, 레티닐 아세테이트, 또는 이들의 조합일 수 있다. 또한, 상기 레티노이드는 레티날, 레티놀, 레티노산, 레티닐 아세테이트, 또는 이들의 조합의 생체 내 분해 산물일 수 있다.
- [0018] 레티노이드는 기본 탄소수가 20인 물질로서, 결합하는 지방산 보조기에 따라 최종 탄소수가 달라질 수 있는데, 예컨대 아세테이트 결합시에는 최종 탄소수가 22이고 올레산 결합시에는 탄소수가 38일 수 있다.
- [0019] 본 발명에서 "Ybb0 단백질"은 서열번호 23의 아미노산 서열을 갖는 단백질일 수 있고, 이는 서열번호 24의 뉴클레오티드 서열을 갖는 유전자에 의해 코딩될 수 있다.
- [0020] 본 발명에서 Ybb0 단백질을 코딩하는 유전자의 뉴클레오티드 서열과 80% 이상의 상동성을 갖는 외래 유전자는 서열번호 49 내지 56의 뉴클레오티드 서열을 갖는 것일 수 있다. 외래 유전자는 Ybb0 단백질을 코딩하는 유전자와 80% 이상, 85% 이상, 90% 이상, 95% 이상 또는 99% 이상의 상동성을 갖는다.
- [0021] 서열번호 49는 본원의 Ybb0 단백질을 코딩하는 유전자와 상동성 100%인 *Shigella sonnei* Ss046의 GENE ID: 3666309 ybb0 유전자의 뉴클레오티드 서열이다.
- [0022] 서열번호 50은 본원의 Ybb0 단백질을 코딩하는 유전자와 상동성 92%인 *Citrobacter rodentium* ICC168의 GENE ID: 8710313 ROD_05481인 유전자의 뉴클레오티드 서열이다.
- [0023] 서열번호 51은 본원의 Ybb0 단백질을 코딩하는 유전자와 상동성 91%인 *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2의 GENE ID: 1247022 ybb0 유전자의 뉴클레오티드 서열이다.
- [0024] 서열번호 52는 본원의 Ybb0 단백질을 코딩하는 유전자와 상동성 87%인 *Enterobacter cloacae* EcWSU1의 GENE ID: 11485505 ybb0 유전자의 뉴클레오티드 서열이다.
- [0025] 서열번호 53은 본원의 Ybb0 단백질을 코딩하는 유전자와 상동성 84%인 *Enterobacter aerogenes* KCTC 2190의 GENE ID: 10792827 EAE_13085 유전자의 뉴클레오티드 서열이다.
- [0026] 서열번호 54는 본원의 Ybb0 단백질을 코딩하는 유전자와 상동성 83%인 *Klebsiella pneumoniae* 342의 GENE ID: 6935503 KPK_4209 유전자의 뉴클레오티드 서열이다.
- [0027] 서열번호 55는 본원의 Ybb0 단백질을 코딩하는 유전자와 상동성 83%인 *Klebsiella oxytoca* KCTC 1686의 GENE ID: 11664596 KOX_13160 유전자의 뉴클레오티드 서열이다.

- [0028] 서열번호 56은 본원의 YbbO 단백질을 코딩하는 유전자와 상동성 83%인 *Cronobacter sakazakii* ATCC BAA-894의 GENE ID: 5548810 ESA_02773 유전자의 뉴클레오티드 서열이다.
- [0029] 서열번호 49 내지 56의 뉴클레오티드 서열을 갖는 유전자는 기계 순으로 각각 서열번호 41 내지 48의 아미노산 서열을 갖는 단백질을 코딩할 수 있다.
- [0030] YbbO 유전자의 감쇄 또는 결실에 의하여, 레티날의 생산능 또는 레티노이드 중 레티날의 비율이 증가된 것일 수 있다. YbbO 유전자 또는 이와 80% 이상의 상동성을 갖는 외래 유전자의 증폭에 의하여 레티놀의 생산능 또는 레티노이드 중 레티놀의 비율이 증가된 것일 수 있다.
- [0031] 상기 증폭은 외래 레티놀 탈수소효소를 코딩하는 유전자의 도입에 의하여 이루어진 것일 수 있다. 상기 증폭은 일 양상으로 외래 YbbO 유전자의 도입에 의하여 이루어진 것일 수 있다.
- [0032] 본 명세서에 있어서, 용어 "감쇄(attenuation)"는 대상 유전자의 발현이 모균주에 비하여 감소한 것을 나타낸다. 용어 "결실(deletion)"은 대상 유전자의 발현이 상실된 것을 나타낸다.
- [0033] 상기 감쇄 또는 결실은 유전자 서열의 변이, 예를 들면, 치환, 결실, 삽입 또는 그들의 조합에 의하여 발생할 수 있다. 상기 감쇄 또는 결실은 유전자 조절 부의 서열의 변이, 예를 들면, 치환, 결실, 삽입 또는 그들의 조합에 의하여 발생할 수 있다.
- [0034] 용어 "증폭(amplification)"은 유전자 카피의 증가 및/또는 유전자의 발현량 증가를 나타낸다. 상기 증폭은 일 양상으로 외래 유전자의 도입 또는 내재적 유전자의 카피 수의 증가에 의하여 이루어질 수 있다. 상기 증폭을 위해 유전자의 발현량을 늘릴 수 있도록 프로모터 또는 RBS(ribosomal binding site)를 대체할 수 있다.
- [0035] 에세리키아(*Escherichia*) 속 모균주의 YbbO 단백질을 코딩하는 유전자가 감쇄 또는 결실, 또는 증폭되거나, 상기 유전자의 뉴클레오티드 서열과 80% 이상의 상동성을 갖는 외래 유전자가 도입되어 증폭된 에세리키아 속 미생물은 모균주에 비하여 레티노이드, 즉, 레티날, 레티놀, 레티닐에스테르, 레티노산, 또는 이들의 조합 중 적어도 어느 하나의 생산능이 증가된 것일 수 있다. 또한, 상기 미생물은 모균주에 비하여 레티노이드, 즉, 레티날, 레티놀, 레티닐에스테르, 레티노산, 또는 이들의 조합 중 적어도 어느 하나의 생산능이 감소된 것일 수 있다.
- [0036] 모균주는 에탄올아민 이용 단백질 E (Ethanolamine utilization protein E: EutE)를 코딩하는 유전자; 푸트레신 이용 경로 단백질 C (Putrescine utilization pathway protein C: PuuC)를 코딩하는 유전자; 또는 이들의 조합이 추가로 감쇄 또는 결실된 에세리키아 (*Escherichia*) 속 미생물일 수 있다.
- [0037] 에탄올아민 이용 단백질 E (EutE)는 서열번호 19의 아미노산 서열을 갖는 것일 수 있고, 이를 코딩하는 코딩하는 유전자는 서열번호 21의 뉴클레오티드 서열을 갖는 것일 수 있다.
- [0038] 푸트레신 이용 경로 단백질 C (PuuC)는 서열번호 20의 아미노산 서열을 갖는 것일 수 있고, 이를 코딩하는 유전자는 서열번호 22의 뉴클레오티드 서열을 갖는 것일 수 있다.
- [0039] 다른 양상은 상기 미생물을 배양하는 단계; 및 배양물로부터 레티노이드를 분리하는 단계;를 포함하는 상기 미생물로부터 레티노이드를 생산하는 방법에 관한 것이다.
- [0040] 상기 방법은 상기한 미생물을 배양하는 단계를 포함한다. 여기서 미생물은 위에서 설명한 YbbO 단백질을 코딩하는 유전자가 감쇄 또는 결실되거나, 증폭된 에세리키아 속 미생물일 수 있고, 위에서 설명된 에세리키아 속 미생물에 관한 사항이 본 방법의 미생물에도 모두 적용될 수 있다.
- [0041] 상기 배양은 친유성 물질을 포함하는 배지 중에서 배양하는 단계를 포함할 수 있다.
- [0042] 상기 친유성 물질은 탄소수 8 내지 50의 유기 화합물로서 친유성을 갖는 것일 수 있다.
- [0043] 상기 친유성 물질은 탄소수 8 내지 50의 알칸 화합물, 하기 화학식 1의 화합물; 하기 화학식 2의 화합물; 또는 이들의 조합일 수 있다:
- [0044] [화학식 1]
- [0045] $R_1(CO)OR_2$
- [0046] (식 중, R_1 및 R_2 는 각각 독립적으로 탄소수 8 내지 50의 알킬을 나타내고, CO는 카르보닐기를 나타냄),

[0047] [화학식 2]



[0048]

[0049] (식 중, R_3 , R_4 및 R_5 는 각각 독립적으로 탄소수 8 내지 50의 알킬을 나타내고, CO는 카르보닐기를 나타냄).

[0050] 탄소수 8 내지 50의 알칸 화합물은 직쇄 알칸, 분지 알칸, 시클릭 알칸, 또는 이들의 조합일 수 있다. 알칸 화합물은 예를 들면, 탄소수 8 내지 46, 8 내지 40, 8 내지 36, 8 내지 30, 8 내지 26, 8 내지 20, 8 내지 16, 8 내지 12, 8 내지 10, 10 내지 50, 10 내지 46, 10 내지 40, 10 내지 36, 10 내지 30, 10 내지 26, 10 내지 20, 10 내지 16, 10 내지 12, 10 내지 50, 10 내지 46, 12 내지 50, 12 내지 46, 12 내지 36, 12 내지 30, 12 내지 26, 12 내지 20, 또는 12 내지 16의 알칸 화합물일 수 있다.

[0051] 직쇄 알칸은 탄소수 8 (옥탄), 10 (데칸), 12 (도데칸), 14 (테트라데칸), 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50의 알칸, 또는 이들의 조합일 수 있다.

[0052] 분지 알칸은 탄소수 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50의 알칸, 또는 이들의 조합일 수 있다. 분지 알칸은 테르펜 화합물의 포화된 유사체(analogue)일 수 있다. 예를 들면, 피토스쿠알란일 수 있다.

[0053] 직쇄 알칸, 분지 알칸, 및 시클릭 알칸의 조합은 미네랄 오일일 수 있다. 미네랄 오일은 비식물성 원료(미네랄) 유래의 탄소 수 15 내지 40의 알칸의 혼합물일 수 있다. 탄소수 15 내지 40의 알칸은 탄소수 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40의 알칸의 2 이상의 혼합물일 수 있다.

[0054] 미네랄 오일은 경량 미네랄 오일 또는 중량 미네랄 오일일 수 있다. 경량 미네랄 오일(light mineral oil)은 일반적으로 밀도가 $880\sim 920\text{kg/m}^3$ 이며 20°C 에서 비중이 $820\sim 860\text{kg/m}^3$, 40°C 에서 유동성 점도가 14~18cst를 가지는 물질이다. 중량 미네랄 오일(heavy mineral oil)은 일반적으로 밀도가 920kg/m^3 이며 20°C 에서 비중이 $860\sim 900\text{kg/m}^3$, 40°C 에서 유동성 점도가 65~85cst인 물질이다.

[0055] 화학식 1의 화합물에서 R_1 및 R_2 는 각각 독립적으로 직쇄, 분지쇄 또는 고리형의 탄소수 8 내지 50의 알킬이다. R_1 과 R_2 는 각각 독립적으로 탄소수 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 또는 50의 알킬일 수 있다.

[0056] R_1 과 R_2 는 각각 탄소수 8 내지 50, 예를 들면, 탄소수 8 내지 46, 8 내지 40, 8 내지 36, 8 내지 30, 8 내지 26, 8 내지 20, 8 내지 16, 8 내지 12, 8 내지 10, 10 내지 50, 10 내지 46, 10 내지 40, 10 내지 36, 10 내지 30, 10 내지 26, 10 내지 20, 10 내지 16, 10 내지 12, 10 내지 50, 10 내지 46, 12 내지 50, 12 내지 46, 12 내지 36, 12 내지 30, 12 내지 26, 12 내지 20, 또는 12 내지 16의 알킬일 수 있다. R_1 이 탄소수 13의 직쇄 알킬이고 R_2 가 이소프로필일 수 있다. 또한, R_1 이 에틸펜틸기이고 R_2 가 세틸일 수 있다.

[0057] 화학식 2의 화합물에서, R_3 , R_4 및 R_5 는 각각 독립적으로 직쇄, 분지쇄 또는 고리형의 탄소수 8 내지 50의 알킬이다.

[0058] R_3 , R_4 , 및 R_5 는 각각 탄소수 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 또는 50의 알킬일 수 있다. 상기 화합물은 예를 들면, R_3 , R_4 , 및 R_5 가 각각 탄소수 8 내지 50, 예를 들면, 탄소수 8 내지 46, 8 내지 40, 8 내지 36, 8 내지 30, 8 내지 26, 8 내지 20, 8 내지 16, 8 내지 12, 8 내지 10, 10 내지 50, 10 내지 46, 10 내지 40, 10 내지 36, 10 내지 30, 10 내지 26, 10 내지 20, 10 내지 16, 10 내지 12, 10 내지 50, 10 내지 46, 12 내지 50, 12 내지 46, 12 내지 36, 12 내지 30, 12 내지 26, 12 내지 20, 또는 12 내지 16의 알킬일 수 있다.

[0059] 친유성 물질은 옥탄, 데칸, 도데칸, 테트라데칸, 피토스쿠알란, 미네랄 오일, 이소프로필 미리스테이트, 세틸

에틸헥사노에이트, 디옥타노일 데카노일 글리세롤, 스쿠알란, 또는 이들의 조합일 수 있다.

- [0060] 친유성 물질은 생산되는 레티노이드를 안정화시키는 것뿐만 아니라, 미생물에 의한 레티노이드의 생산성을 증가시킬 수 있다. 친유성 물질은 미생물의 생장에 영향을 미치지 않거나 적게 영향을 미치는 것일 수 있다.
- [0061] 배양은 합성, 반합성, 또는 복합 배양 배지에서 이루어질 수 있다. 배양 배지로는 탄소원, 질소원, 비타민 및 미네랄로 구성된 배지를 사용할 수 있다. 예를 들어, MRS (Man-Rogosa-Sharp) 액체 배지 또는 우유가 첨가된 액체 배지를 사용할 수 있다.
- [0062] 배지의 탄소원으로는 전분, 포도당, 자당, 갈락토스, 과당, 글리세롤, 글루코스 또는 이들의 혼합물이 사용될 수 있다. 예를 들면, 글리세롤이 탄소원으로 사용될 수 있다. 질소원으로는 황산암모늄, 질산암모늄, 질산나트륨, 글루탐산, 카사미노산, 효모추출물, 펩톤, 트립톤, 대두박 또는 이들의 혼합물이 사용될 수 있다. 미네랄은 염화나트륨, 인산제이칼륨, 황산마그네슘 또는 이들의 혼합물이 사용될 수 있다.
- [0063] 배양이 발효조에서 이루어지는 경우 글루코스를 배지의 탄소원으로 사용하는 것이 좋다. 시험관 배양의 경우에는 글리세롤을 배지의 탄소원으로 사용하는 것이 좋다.
- [0064] 미생물 배양 배지 내 상기 탄소원, 질소원 및 미네랄 각각은 예를 들면, 리터당 10 내지 100 g, 5 내지 40 g 및 0.5 내지 4 g 을 이용할 수 있다.
- [0065] 상기의 통상의 배양 배지에 첨가되는 비타민은 비타민 A, 비타민 B, 비타민 C, 비타민 D, 비타민 E 또는 이들의 혼합물일 수 있다. 비타민은 통상의 배양 배지에 상기에서 언급된 탄소원, 질소원, 미네랄 등과 함께 첨가되거나, 멸균하여 준비된 배지에 별도로 첨가될 수 있다.
- [0066] 배양은 통상의 대장균 배양 조건으로 수행될 수 있다. 배양은 예를 들어 약 15-45℃, 예를 들면, 15-44℃, 15-43℃, 15-42℃, 15-41℃, 15-40℃, 15-39℃, 15-38℃, 15-37℃, 15-36℃, 15-35℃, 15-34℃, 15-33℃, 15-32℃, 15-31℃, 15-30℃, 20-45℃, 20-44℃, 20-43℃, 20-42℃, 20-41℃, 20-40℃, 20-39℃, 20-38℃, 20-37℃, 20-36℃, 20-35℃, 20-34℃, 20-33℃, 20-32℃, 20-31℃, 20-30℃, 25-45℃, 25-44℃, 25-43℃, 25-42℃, 25-41℃, 25-40℃, 25-39℃, 25-38℃, 25-37℃, 25-36℃, 25-35℃, 25-34℃, 25-33℃, 25-32℃, 25-31℃, 25-30℃, 27-45℃, 27-44℃, 27-43℃, 27-42℃, 27-41℃, 27-40℃, 27-39℃, 27-38℃, 27-37℃, 27-36℃, 27-35℃, 27-34℃, 27-33℃, 27-32℃, 27-31℃ 또는 27-30℃에서 수행될 수 있다.
- [0067] 배양액 중의 배양 배지를 제거하고 농축된 균체만을 회수하거나 제거하기 위해 원심분리 또는 여과과정을 거칠 수 있으며 이러한 단계는 당업자의 필요에 따라 수행할 수 있다. 농축된 균체는 통상적인 방법에 따라 냉동하거나 냉동건조하여 그 활성을 잃지 않도록 보존할 수 있다.
- [0068] 배양의 일 예에 있어서, 배양은 탄소원으로서 글리세롤을 포함하는 배지에서 이루어지는 것일 수 있다. 글리세롤은 배지 중의 유일한 탄소원일 수 있다. 0.5-5.0%(w/v), 예를 들면, 0.5-4.5%(w/v), 0.5-4.0%(w/v), 0.5-3.5%(w/v), 0.5-3.0%(w/v), 0.5-2.5%(w/v), 0.5-2.0%(w/v), 1-5.0%(w/v), 1-4.5%(w/v), 1-4.0%(w/v), 1-3.5%(w/v), 1-3.0%(w/v) 또는 1-2.5%(w/v)의 글리세롤을 포함하는 배지에서 이루어지는 것일 수 있다. 상기 배지는 글리세롤 및 아라비노스가 첨가된 YT 배지일 수 있다. YT 배지는 1.6중량% 트립톤, 1중량% 효모 추출물 및 0.5 중량% NaCl을 포함할 수 있다.
- [0069] 배양은 친유성 물질 존재 하의 배양 배지에서, 예를 들면 배지 표면에 친유성 물질인 도데칸 상(dodecane phase)을 위치시킨 상태로 수행될 수 있다. 배양은 교반되는 상태에서 수행될 수 있다.
- [0070] 교반되는 경우, 100 내지 300rpm, 예를 들면, 100 내지 280rpm, 100 내지 260rpm, 100 내지 240rpm, 100 내지 220rpm, 100 내지 200rpm, 100 내지 180rpm, 100 내지 160rpm, 100 내지 140rpm, 100 내지 120rpm, 120 내지 300rpm, 120 내지 280rpm, 120 내지 260rpm, 120 내지 240rpm, 120 내지 220rpm, 120 내지 200rpm, 120 내지 180rpm, 120 내지 160rpm, 120 내지 140rpm, 150 내지 300rpm, 150 내지 280rpm, 150 내지 260rpm, 150 내지 240rpm, 150 내지 220rpm, 150 내지 200rpm, 150 내지 180rpm, 140 내지 160rpm, 200 내지 300rpm, 200 내지 280rpm, 200 내지 260rpm, 200 내지 240rpm, 200 내지 220rpm, 또는 150 rpm으로 교반될 수 있다.
- [0071] 교반되는 경우, 상기 친유성 물질, 예컨대 도데칸은 배지 중에서 분산되어 세포와 접촉된다. 친유성 물질은 배지 중에 분산됨으로써 미생물과 접촉하는 면적이 넓어져 배양 중 레티노이드를 효율적으로 세포로부터 분리되게 하여 안정화 및/또는 용해시킬 수 있다.
- [0072] 친유성 물질, 예컨대 도데칸 상 없이 상기 기술한 레티노이드를 생산하는 미생물을 배양시켰을 때, 레티노이드

생산은 일정 시점에서 최고치를 나타내고 그 이후에 감소할 수 있다. 이는 미생물 성장의 정체 상태 동안 추가적인 레티노이드 합성이 중단되는 반면, 세포 내에서 그의 산화적 분해가 일어나기 때문일 수 있다.

- [0073] 친유성 물질, 예컨대 도데칸 상의 존재 하에 배양 배지에서 상기 미생물을 배양시키게 되면 생산된 레티노이드가 세포 내에서 분해되기 전에 친유성 물질, 예컨대 도데칸 상에 흡수되게 되어 레티노이드 생산량을 향상시킬 수 있다.
- [0074] 상기 친유성 물질, 예컨대 도데칸 상은 에세리키아 속 미생물의 세포 성장에 영향을 미치지 않고, 소수성 레티노이드의 추출을 위해 소수성이고, 낮은 휘발성을 갖는 것일 수 있다. 도 2는 β -카로틴의 레티날, 레티놀, 레티노산, 및 레티닐 에스테르를 포함하는 레티노이드로의 변환을 나타낸다.
- [0075] 배지 대 친유성 물질의 부피비는 특정 범위의 비로 한정되지 않고, 예컨대, 배지 대 친유성 물질의 부피비가 1:0.1-3.0, 1:0.2-3.0, 1:0.5-3.0, 1:1.0-3.0, 1:1.5-3.0, 1:2.0-3.0, 1:2.5-3.0, 1:0.2-2.5, 1:0.2-2.0, 1:0.2-1.5, 1:0.2-1.0, 1:0.2-0.5, 1:0.5-2.5, 1:0.5-2.0, 1:0.5-1.5, 1:0.5-1.0, 1:0.8-2.5, 1:0.8-2.0, 1:0.8-1.5, 1:0.8-1.2, 1:0.8-1.0 등이 가능하다.
- [0076] 일 구체예에 따르면 상기 배양하는 단계에서 상기 배지는 약 2.0% 농도의 글리세롤을 포함하고, 상기 에세리키아 속 미생물은 대장균 DH5 α 또는 MG1655이고, 상기 배양하는 단계는 배양액 약 7 ml, 약 29°C에서 배양시키는 것일 수 있다.
- [0077] 상기 방법은 또한, 친유성 물질 상으로부터 레티노이드를 분리하는 단계를 포함한다. 상기 레티노이드, 예를 들면, 레티날, 레티놀, 레티닐 에스테르, 또는 이들의 조합을 분리하는 방법은 당업계에 잘 알려져 있다. 예를 들면, 이온교환 크로마토그래피, HPLC 등의 방법에 의하여 분리될 수 있다. 구체적으로, 균체를 회수한 후에 아세톤 등의 용매를 이용한 추출 후에 고순도의 제품을 얻기 위해서는 HPLC 또는 결정화 작업등을 통한 분리정제가 진행될 수 있다.
- [0078] 레티노이드는 화장품, 식품 또는 의약품의 소재로 널리 활용된다.
- [0079] 다른 양상으로, YbbO 단백질; 또는 상기 YbbO 단백질을 코딩하는 유전자의 아미노산 서열과 80% 이상의 상동성을 갖는 유전자가 코딩하는 단백질을 포함하는 효소 조성물을 첨가함으로써 레티날로부터 레티놀로의 전환을 촉진시킬 수 있다.
- [0080] 상기 YbbO 단백질은 앞서 설명한 바와 마찬가지로 서열번호 23의 아미노산 서열을 갖는 것일 수 있고, 상기 YbbO 단백질을 코딩하는 유전자의 아미노산 서열과 80% 이상의 상동성을 갖는 유전자가 코딩하는 단백질은 서열번호 41 내지 48 중 어느 하나의 아미노산 서열을 갖는 것일 수 있다.
- [0081] YbbO 단백질을 코딩하는 유전자의 아미노산 서열과 80% 이상의 상동성을 갖는 유전자는 YbbO 단백질을 코딩하는 유전자의 아미노산 서열과 80% 이상, 85% 이상, 90% 이상, 95% 이상 또는 99% 이상의 상동성을 가질 수 있다.
- [0082] 서열번호 23의 아미노산 서열을 갖는 YbbO 단백질은 서열번호 24의 뉴클레오티드 서열을 갖는 유전자에 의해 코딩될 수 있고, 서열번호 41 내지 48 중 어느 하나의 아미노산 서열을 갖는 단백질은 서열번호 49 내지 56 중 어느 하나의 뉴클레오티드 서열을 갖는 유전자에 의해 코딩될 수 있다.
- [0083] YbbO 단백질을 포함하는 효소 조성물은 용매, 안정제, 각종 첨가제 등 통상의 효소 조성물에 포함되는 성분들을 추가 포함할 수 있다. 또한, YbbO 단백질 이외의 레티날로부터 레티놀로의 전환을 촉진시키는 성분을 추가 포함할 수 있다.

발명의 효과

- [0084] 일 양상에 따른 에세리키아 속 미생물은 증가된 레티노이드 생산능을 갖는다. 또한, 레티노이드 중 특정 산물을 증가된 비율로 생산할 수 있다.
- [0085] 다른 양상에 따른 레티노이드를 생산하는 방법에 의하면, 레티노이드를 효율적으로 생산할 수 있다. 또한, 레티노이드 중 특정 산물을 증가된 비율로 생산할 수 있다.
- [0086] 본 발명의 YbbO 단백질을 코딩하는 유전자가 감쇄 또는 결실, 또는 증폭되거나, 상기 YbbO 단백질을 코딩하는 유전자의 아미노산 서열과 80% 이상의 상동성을 갖는 유전자가 도입되어 증폭된 에세리키아 속 미생물을 친유성 물질을 포함하는 배지 중에서 배양시켜 친유성 물질 상으로부터 레티노이드를 분리시키면 보다 효율적으로 레티

노이드를 생산할 수 있다.

[0087] 본 발명의 미생물을 이용한 레티노이드 생산 방법을 통해 얻어진 레티노이드는 화장품, 식품, 의약품 등의 소재로 널리 활용될 수 있고, 화장품, 식품, 의약품 등의 제조를 위해 특정 레티노이드의 효과적인 생산이 필요할 경우 본 발명의 레티노이드 생산 방법을 활용하기 적합하다.

도면의 간단한 설명

[0088] 도 1은 레티날 생합성의 MEP 경로 및 외래의 MVA 경로를 도식적으로 나타낸 도면이다.
 도 2는 β -카로틴의 레티날, 레티놀, 레티노산, 및 레티닐 에스테르를 포함하는 레티노이드로의 변환을 나타낸 도면이다.
 도 3은 YbbO 유전자 도입이 레티노이드 생산에 미치는 영향을 나타낸 도면이다.
 도 4는 YbbO 유전자 결실이 레티노이드 생산에 미치는 영향을 나타낸 도면이다.
 도 5는 eutE 유전자 또는 puuC 유전자 결실이 레티노이드 생산에 미치는 영향을 나타낸 도면이다.
 도 6은 도데칸이 포함된 배지를 사용한 배양에서의 대장균(pT-DHBSR/pS-NA)의 도데칸 부피에 따른 레티노이드 생산 및 세포 성장을 나타낸 도면이다.
 도 7은 도데칸이 포함된 배지를 사용한 배양에서의 대장균(pT-DHBSR/pS-NA)의 배양 시간 및 도데칸 부피에 따른 레티노이드의 분포가 총 레티노이드에 대한 각 구성성분의 백분율로 나타낸 도면이다.
 도 8 및 도 9는 다양한 알칸 존재하에서 대장균(pT-DHBSR/pS-NA)의 레티노이드 생산 및 세포성장 결과를 나타낸 도면이다.
 도 10, 도 11 및 도 12는 다양한 부피의 경량 미네랄 오일 존재 하에서 대장균(pT-DHBSR/pS-NA)의 레티노이드 생산, 세포성장, 균체당 레티노이드 생산량(cell specific retinoids productivity) 결과를 나타낸 도면이다.
 도 13 및 도 14는 중량 미네랄 오일 존재 하에서 대장균(pT-DHBSR/pS-NA)의 레티노이드 생산 및 세포성장 결과를 나타낸 도면이다.
 도 15 및 도 16은 중량 미네랄 오일 존재 하에서 시험관을 기울여 배양한 경우 대장균(pT-DHBSR/pS-NA)의 레티노이드 생산 및 세포성장 결과를 나타낸 도면이다.
 도 17은 피부 친화적 친유성 물질 존재 하에서 대장균(pT-DHBSR/pS-NA)의 세포 성장 및 pH를 나타낸 도면이다.
 도 18 및 도 19는 피부 친화적 친유성 물질의 다양한 종류와 양에 따른 대장균(pT-DHBSR/pS-NA)의 레티노이드 생산결과를 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

[0089] 이하 본 발명을 실시예를 통하여 보다 상세하게 설명한다. 그러나, 이들 실시예는 본 발명을 예시적으로 설명하기 위한 것으로 본 발명의 범위가 이들 실시예에 한정되는 것은 아니다. 이하의 실시예에서 다음의 실험 재료 및 방법을 사용하였다.

[0090] 실시예 1: YbbO 유전자가 증폭 또는 결실된 대장균 및 그를 이용한 레티노이드의 생산

[0091] 레티노이드 생산능을 갖는 에세리키아 속 미생물인 대장균 DH5 α (pTDHBSR/pSNA) (기탁번호 KCTC 11255BP)에서 YbbO 유전자를 증폭 또는 결실시키고, 그 증폭 또는 결실이 레티노이드 생산에 미치는 영향을 확인하였다.

[0092] (1) YbbO 유전자의 증폭된 대장균의 제조 및 레티노이드 생산

[0093] 대장균 YbbO 유전자 (서열번호 24)를 pBBR1MCS2 플라스미드의 KpnI과 XhoI 제한효소 부위에 삽입하여 pBBR1MCS2(YbbO) 플라스미드를 제조하였다. 다음으로, pBBR1MCS2(YbbO) 플라스미드를 대장균 DH5 α (pTDHBSR/pSNA)에 형질전환에 의하여 도입하여 대장균 DH5 α (pTDHBSR/pSNA/pBBR1MCS2-YbbO)를 얻었다. 이 과정

을 상세히 설명하면 다음과 같다.

표 1

[0094]

효소명	유전자	유전자 서열 (Genbank 허가번호)
야생형 대장균(<i>Escherichia coli</i> MG1655; taxid 511145) 유래의 예상 산화환원효소	<i>ybbO</i>	서열번호 24 (1786701)

표 2

[0095]

유전자	방향	프라이머 서열	제한효소
<i>ybbO</i>	F	서열번호 25	<i>KpnI</i>
<i>ybbO</i>	R	서열번호 26	<i>XhoI</i>

[0096]

게놈 DNA 제조, 제한효소 절단, 형질 전환 및 표준 분자 생물학 기술은 문헌에 기재된 대로 수행하였다 (Sambrook and Russell 2001). PCR은 그의 표준 프로토콜에 의해 *pfu* DNA 폴리머라제(Solgent Co., Korea)를 사용하여 수행하였다.

[0097]

표 1의 유전자 클로닝에 사용된 프라이머 서열과 제한효소는 표 2에 나타내었다. 표 1의 유전자는 해당 유전자를 포함하고 있는 균주의 염색체 DNA를 주형으로 PCR을 통해 증폭되었다. 증폭된 산물을 표 2에 열거된 제한 효소를 이용하여 벡터 pBBR1MCS2 (GenBank 허가번호 U02374)(서열번호 27)에 도입하여, 벡터 pBBR2-ybbO(서열번호 40)를 제조하였다. 그 후, pBBR2-ybbO 플라스미드를 대장균 DH5 α (pTDHBSR/pSNA)에 형질전환에 의하여 도입하여 대장균 DH5 α (pTDHBSR/pSNA/pBBR2-YbbO)를 얻었다.

[0098]

얻은 균주를 2% (w/v) 글리세롤 및 0.2% (w/v) 아라비노오스를 함유한 2YT 배지(리터 당 16g 트립톤, 10g 이스트 추출물, 및 5g NaCl) 5mL와 도데칸 5mL의 혼합 배지 중에서 72 시간 동안 배양하였다.

[0099]

배양은 길이 15 cm, 지름 25 mm의 시험관에 상기 혼합 배지를 넣고 각 균주를 접종한 후, 진탕 배양기 (shaking incubator)에 약 250 rpm으로 교반하면서 약 29℃에서 배양하였다. 대조군으로는 YbbO 유전자가 도입되지 않은 벡터가 도입된 DH5 α (pTDHBSR/pSNA/pBBR1MCS2)를 사용하였다.

[0100]

도 3은 YbbO 유전자 도입이 레티노이드 생산에 미치는 영향을 나타낸 도면이다. 도 3에 나타낸 바와 같이, 대조군 균주(pBBR1MCS-2)에 비하여, YbbO 유전자 도입 균주(pBBR2-ybbO)에서 전체 레티노이드 중 레티놀과 레티닐 아세테이트의 비율이 증가하였다. 전체 레티노이드 생산량은 큰 차이가 없었다. 대조군에서 레티날: 레티놀: 레티닐 아세테이트는 70mg/L: 45 mg/L: 20 mg/L이었으나, YbbO 유전자 균주에서는 21mg/L: 67 mg/L: 39 mg/L이었다.

[0101]

YbbO 유전자 도입 균주에 의한 레티노이드 생산시 레티닐 아세테이트 비율 증가는 레티놀 비율의 증가에 의해 야기된 것으로 여겨진다. 반면, YbbO 유전자 도입 균주에 의한 레티노이드 생산시 레티날은 소량 생산되었다.

[0102]

(2) YbbO 유전자의 결실된 대장균의 제조 및 레티노이드 생산

[0103]

대장균 DH5 α Δ ybbO 균주는 Datsenko와 Wanner의 방법(Datsenko KA and Wanner BL., 2000 Proc Natl Acad Sci U.S.A., 97(12):6640-6645)에 따라서 다음과 같이 제조되었다.

[0104]

ybbO 유전자를 상동성 재조합에 의하여 선발 마커 유전자인 카나마이신 유전자로 대체하기 위하여, pKD13 플라이스미드(Datsenko KA and Wanner BL., 2000 Proc Natl Acad Sci U.S.A., 97(12):6640-6645)의 카나마이신 유전자에 결합하면서 ybbO 유전자의 상류 말단과 하류 말단의 염기 서열을 갖는 PCR 프라이머를 제조하였다.

[0105]

정방향 프라이머는 ybbO 유전자의 상류 말단의 50bp 염기서열 다음에, 카나마이신 유전자 결합 염기 서열 20bp로 구성되어 있으며, 역방향 프라이머는 ybbO 유전자의 하류 말단의 50bp 염기서열 다음에, 카나마이신 유전자 결합 염기 서열 20bp로 구성되어 있다. 사용한 프라이머는 표 3에 표시하였다. 서열번호 28과 29의 올리고뉴클레오타이드를 프라이머로 하고, pKD13을 주형으로 PCR 반응을 수행하였다.

[0106]

상기의 정제된 PCR 반응산물은 pKD46 (Datsenko KA and Wanner BL., 2000 Proc Natl Acad Sci U.S.A.,

97(12):6640-6645)을 포함한 DH5 α 콤펙턴트 세포 (competent cell)로 형질전환하였다. 카나마이신 저항성을 가진 평판 배지의 콜로니를 콜로니 PCR을 통하여 카나마이신 유전자가 상동성 재조합에 의해서 ybbO 유전자를 대체했는지를 확인하였다. 콜로니 PCR에 사용된 프라이머는 대장균 DH5 α 염색체 상에서 ybbO 유전자 바로 인접한 영역에 결합하도록 제조하였다.

[0107] 그 결과, ybbO 유전자가 결실된 것을 콜로니 PCR 방법에 의하여 표 3의 서열번호 30과 서열번호 31의 프라이머를 이용해 확인하였다.

표 3

[0108]

명칭	서열번호
KO YbbO-F	28
KO YbbO-R	29
KO YbbOCF-F	30
KO YbbOCF-R	31

[0109] 얻은 균주를 2% (w/v) 글리세롤 및 0.2% (w/v) 아라비노오스를 함유한 2YT 배지(리터 당 16g 트립톤, 10g 이스트 추출물, 및 5g NaCl) 5mL와 도데칸 5mL의 혼합 배지 중에서 60 시간 동안 배양하였다. 배양은 길이 15 cm, 지름 25 mm의 시험관에 상기 혼합 배지를 넣고 각 균주를 접종한 후, 진탕 배양기 (shaking incubator)에 약 250 rpm으로 교반하면서 약 29°C에서 배양하였다. 대조군으로는 YbbO 유전자가 결실되지 않은 DH5 α (pTDHBSR/pSNA)를 사용하였다.

[0110] 도 4는 YbbO 유전자 결실이 레티노이드 생산에 미치는 영향을 나타낸 도면이다. 도 4에 나타낸 바와 같이, 대조군 균주(DH5 α)에 비하여, YbbO 유전자 결실 균주(DH5 α Δ ybbO)에서 전체 레티노이드 생산량이 높았다.

[0111] 또한, 전체 레티노이드 중 레티날이 레티놀보다 더 많이 생산되었다. 즉, 60시간 대조군에서, 레티날: 레티놀: 레티닐 아세테이트는 43mg/L: 22 mg/L: 9.7 mg/L이었으나, YbbO 유전자 결실 균주에서, 65mg/L: 25 mg/L: 9.5 mg/L이었다. 따라서, 얻어진 균주를 사용하여 레티날 비율이 높은 레티노이드를 생산할 수 있다.

[0112] 또한, 실시예 2에서 확인된 바와 같이, EutE, PuuC, 또는 그 조합을 결실시킴으로써, 레티날 비율이 높은 레티노이드를 생산할 수 있다.

[0113] 실시예 2: EutE 유전자와 PuuC 유전자가 결실된 대장균 및 그를 이용한 레티노이드 생산

[0114] 레티노이드 생산능을 갖는 에세리키아 (Escherichia) 속 미생물인 DH5 α (pTDHBSR/pSNA) (기탁번호 KCTC 11255BP)에서 EutE 유전자 및 PuuC 유전자를 결실시키고, 그 결실이 레티노이드 생산에 미치는 영향을 확인하였다.

[0115] (1) DH5 α (pTDHBSR/pSNA)의 eutE 유전자 또는 puuC 유전자 결실

[0116] 결실은 Datsenko와 Wanner의 방법 (Datsenko KA and Wanner BL., 2000 Proc Natl Acad Sci U.S.A., 97(12):6640-6645)에 따라서 다음과 같이 제조하였다.

[0117] EutE 유전자를 상동성 재조합에 의하여 선발 마커 유전자인 카나마이신 유전자로 대체하기 위하여, pKD13 플라스미드(Datsenko KA and Wanner BL., 2000 Proc Natl Acad Sci U.S.A., 97(12):6640-6645)의 카나마이신 유전자에 결합하면서 EutE 유전자의 상류 말단과 하류 말단의 염기 서열을 갖는 PCR 프라이머를 제조하였다.

[0118] 정방향 프라이머는 EutE 유전자의 상류 말단의 50bp 염기서열 다음에, 카나마이신 유전자 결합 염기 서열 20bp로 구성되어 있으며, 역방향 프라이머는 EutE 유전자의 하류 말단의 50bp 염기서열 다음에, 카나마이신 유전자 결합 염기 서열 20bp로 구성되어 있다. 사용한 프라이머는 하기 표 4에 표시하였다. 서열번호 32와 서열번호 33의 올리고뉴클레오타이드를 프라이머로 하고, pKD13을 주형으로 PCR 반응을 수행하였다.

[0119] 상기의 정제된 PCR 반응산물은 pKD46 (Datsenko KA and Wanner BL., 2000 Proc Natl Acad Sci U.S.A., 97(12):6640-6645)을 포함한 DH5 α 콤펙턴트 세포 (competent cell)로 형질전환하였다. 카나마이신 저항성을 가진 평판 배지의 콜로니를 콜로니 PCR을 통하여 카나마이신 유전자가 상동성 재조합에 의해서 EutE 유전자를 대

체했는지를 확인하였다. 콜로니 PCR에 사용된 프라이머는 대장균 DH5 α 염색체 상에서 EutE 유전자 바로 인접한 영역에 결합하도록 제조하였다.

[0120] PuuC 유전자를 상동성 재조합에 의하여 선발마커 유전자인 클로람페니콜 유전자로 대체하기 위하여, pKD3 플라이스미드 (Datsenko KA and Wanner BL., 2000 Proc Natl Acad Sci U.S.A., 97(12):6640-6645)의 클로람페니콜 유전자에 결합하면서 PuuC 유전자의 상류 말단과 하류 말단의 염기 서열을 갖는 PCR 프라이머를 제조하였다.

[0121] 정방향 프라이머는 PuuC 유전자의 상류 말단의 50bp 염기서열 다음에, 클로람페니콜 유전자 결합 염기 서열 20bp로 구성되어 있으며, 역방향 프라이머는 PuuC 유전자의 하류 말단의 50bp 염기서열 다음에, 클로람페니콜 유전자 결합 염기 서열 20bp로 구성되어 있다. 사용한 프라이머는 하기 표 4에 표시하였다. 서열번호 36과 서열번호 37의 올리고뉴클레오타이드를 프라이머로 하고, pKD3을 주형으로 PCR 반응을 수행하였다.

[0122] 상기의 정제된 PCR 반응산물은 pKD46 (Datsenko KA and Wanner BL., 2000 Proc Natl Acad Sci U.S.A., 97(12):6640-6645)을 포함한 DH5 α 콤펙트 세포 (competent cell)로 형질전환하였다. 클로람페니콜 저항성을 가진 평판 배지의 콜로니를 콜로니 PCR을 통하여 클로람페니콜 유전자가 상동성 재조합에 의해서 PuuC 유전자를 대체했는지를 확인하였다. 콜로니 PCR에 사용된 프라이머는 대장균 DH5 α 염색체 상에서 PuuC 유전자 바로 인접한 영역에 결합하도록 제조하였다.

[0123] 그 결과, EutE 유전자 및 PuuC 유전자가 결실된 것을 콜로니 PCR 방법에 의하여 EutE는 하기 표 4의 서열번호 34와 서열번호 35의 프라이머로, PuuC는 하기 표 4의 서열번호 38과 서열번호 39의 프라이머로 확인하였다.

표 4

명칭	서열번호
KO eutE-F	32
KO eutE-R	33
KO eutECF-F	34
KO eutECF-R	35
KO puuC-F	36
KO puuC-R	37
KO puuCCF-F	38
KO puuCCF-R	39

[0125] (2) 레티노이드 생산의 비교

[0126] (1)에서 얻은 균주를 2% (w/v) 글리세롤 및 0.2% (w/v) 아라비노오스를 함유한 2YT 배지(리터 당 16g 트립톤, 10g 이스트 추출물, 및 5g NaCl) 5mL와 도데칸 5mL의 혼합 배지 중에서 60 시간 동안 배양하였다.

[0127] 배양은 길이 15 cm, 지름 25 mm의 시험관에 상기 혼합 배지를 넣고 각 균주를 접종한 후, 진탕 배양기 (shaking incubator)에 약 250 rpm으로 교반하면서 약 29℃에서 배양하였다. 대조군으로는 상기 유전자가 결실되지 않은 DH5 α (pTDHBSR/pSNA)를 사용하였다.

[0128] 도 5는 EutE 유전자, 또는 PuuC 유전자 결실이 레티노이드 생산에 미치는 영향을 나타낸 도면이다. 도 5에 나타낸 바와 같이, 대조군 균주에 비하여, 결실 균주에서 전체 레티노이드의 양 (레티날, 레티놀 및 레티닐 아세테이트의 총량) 및 레티날의 비율이 배양 60시간에서 크게 증가하였다. 대조군 균주의 경우 배양 60시간 총 레티노이드 생산량이 75 mg/L, EutE 유전자가 결실된 경우 110.3 mg/L, PuuC 유전자가 결실된 경우 102.2 mg/L를 나타내었다.

[0129] 이는 EutE 유전자 및 PuuC 유전자가 레티날을 레티노산으로 산화시키는 옥시다제 활성을 코딩하기 때문인 것으로 여겨진다. 그러나, 특정한 기작에 한정되는 것은 아니다.

[0130] 실시예 3: 도데칸 존재하 대장균 배양과 레티노이드 생산

[0131] 도데칸 함유 배지 중에서 대장균(pT-DHBSR/pS-NA)를 배양하여, 도데칸이 레티노이드 생산 및 세포 성장에 미치는 영향을 확인하였다.

[0132] 2% (w/v) 글리세롤 및 0.2% (w/v) 아라비노오스를 함유한 2YT 배지(리터 당 16g 트립톤, 10g 이스트 추출물, 및

5g NaCl) 5mL와 도데칸 (0, 1, 2, 3, 4, 5, 6mL)의 혼합 배지 중에서 72 시간 동안 배양하였다. 배양은 길이 15 cm, 지름 25 mm의 시험관에 상기 혼합 배지를 넣고 각 균주를 접종한 후, 진탕 배양기 (shaking incubator)에 약 250 rpm으로 교반하면서 약 29℃에서 배양하였다.

[0133] 도 6은 도데칸 함유 배지에서 대장균(pT-DHBSR/pS-NA)의 레티노이드 생산 및 세포 성장을 나타낸 도면이다. 레티노이드 생산에서 레티날, 레티놀 및 레티닐 아세테이트를 각각 밝은 회색, 어두운 회색 및 검은색으로 나타내었다. 세포 배양시에 오버레이한 도데칸 부피 0 mL, 1 mL, 2 mL, 3 mL, 4 mL, 5 mL 및 6 mL를 각각 ■, ●, ▲, □, ○, △ 및 ☆로 나타내었다.

[0134] 도 7은 배양 시간 및 도데칸 부피에 따른 레티노이드의 분포가 총 레티노이드에 대한 각 구성성분의 백분율로 나타낸 도면이다. 레티날, 레티놀 및 레티닐 아세테이트를 각각 밝은 회색, 어두운 회색 및 검은색으로 나타내었다.

[0135] 도면의 레티노이드 함량은 배양액중의 도데칸층만을 원심분리로 분리, 회수하여 HPLC를 통해 측정된 값이다.

[0136] 측정 결과, 레티노이드는 세포 또는 세포 파쇄물 중에는 존재하지 않거나 무시할 수 있는 정도로만 존재하였으며, 세포가 제거된 배양액 즉, 도데칸 층에 존재하였다.

[0137] 도 6 및 도 7에 나타난 바와 같이, 도데칸이 첨가되지 않은 것에 비하여, 도데칸 첨가에 의하여 레티노이드 생산이 증가하였다.

[0138] 도 6에 나타난 바와 같이, 레티노이드 생산은 도데칸 부피에 비례하였다. 5 mL 도데칸을 갖는 배양의 72 시간에서 136 mg/L의 가장 높은 레티노이드 생산을 얻었고 이는 1 mL 도데칸으로 할 경우(65 mg/L)보다 약 2 배 높은 수치이다. 5 mL 도데칸으로 72 시간 이상 연장된 배양을 시켰을 때 추가적인 레티노이드 생산의 증가는 없었고 그의 최고치를 분해 없이 유지하였다(데이터는 나타나지 않음). 0, 24, 및 48 시간에서 2 mL의 도데칸을 배양에 추가하는 것에 의해 도데칸 첨가 부피를 6 mL까지 증가시켰다. 6 mL 도데칸 첨가 배양에서 5 mL 도데칸을 갖는 배양에 비교하였을 때 전체 레티노이드 생산의 증가는 없었다. 6 mL 도데칸을 초기에 첨가 배양에서도 레티노이드 생산의 증가는 없었다(데이터는 나타나지 않음). 도데칸을 갖는 모든 배양액에서의 세포 성장은 도데칸이 없을 때보다 약간 높았다(도 6).

[0139] 도 7은 도데칸 첨가 부피에 따라 얻은 레티노이드들의 분포를 나타낸 도면이다. 도데칸 첨가가 있는 경우와 없는 경우 얻은 레티날 및 레티놀 비율에서 현저한 레티노이드 분포 차이가 나타난다. 48 시간에서 레티노이드 중 레티날의 비율은 도데칸 첨가 배양에서 약 51% (w/w)이고 도데칸 첨가가 없는 배양에서 23%인 반면, 레티놀 비율은 도데칸 첨가 배양에서 30% 내지 39%이고 도데칸 첨가 없는 배양에서 59%이다. 따라서, 도데칸 첨가는 레티날의 비율을 높이지만 레티놀의 비율을 감소시킨다. 세포에서 레티날로부터의 레티놀 형성의 반응 순서를 고려하였을 때, 레티날은 도데칸에 의해 그의 레티놀로의 전환 전에 세포로부터 추출되는 것으로 생각된다. 48 시간에서의 레티닐 아세테이트 비율은 도데칸 첨가 있는 배양 및 없는 배양 모두에서 20% 미만이고, 이는 레티날 및 레티놀에 비해 상대적으로 낮은 수치이다. 도데칸 첨가 배양에서 레티닐 아세테이트 비율은 배양 시간이 길어짐에 따라 감소하는데 이는 레티닐 아세테이트 형성을 위한 세포 활동이 배양 동안 감소함을 나타낸다. 결론적으로, 도데칸 첨가는 세포 성장의 정체 상태에서 레티노이드 생산의 감소를 방지하고 레티노이드 생산을 증가시켰다.

[0140] 본 발명의 레티노이드의 인-시추 추출에는 세포벽 분해를 위한 리소자임이 필요하지 않다. 레티노이드(C20, 이소프레노이드 분자)는 세포벽의 손실 없이 세포로부터 효과적으로 방출될 수 있다. 레티노이드 생산의 2-상 배양에서, β-카로틴은 레티노이드의 직접적인 전구체이기 때문에 세포 내에 계속 유지되어야 한다. β-카로틴이 도데칸 상에서 추출되는 경우, 이는 시토솔에 위치한 BCD(M)O에 의해 절단될 수 있다.

[0141] β-카로틴은 분자 크기 때문에 세포로부터 방출될 수 없어 도데칸에 의해 추출되지 않으므로, β-카로틴의 2-상 배양에서 세포 내에서 계속 유지될 수 있다.

[0142] 실시예 4: 친유성 물질을 포함하는 배지에서 레티노이드의 생산

[0143] 본 실시예에서는 다양한 친유성 물질 함유 배지 중에서 대장균(pT-DHBSR/pS-NA)을 배양하여, 친유성 물질이 레티노이드 생산 및 세포 성장에 미치는 영향을 확인하였다.

[0144] (1) 알칸을 포함하는 배지에서 레티노이드의 생산

[0145] 위 대장균(pT-DHBSR/pS-NA)을 2% (w/v) 글리세롤 및 0.2% (w/v) 아라비노오스를 함유한 2YT 배지(리터 당 16g 트립톤, 10g 이스트 추출물, 및 5g NaCl) 5mL와 옥탄, 데칸, 도데칸 또는 테트라데칸 5mL의 혼합 배지 중에서 72 시간 동안 배양하였다. 배양은 길이 15 cm, 지름 25 mm의 시험관에 상기 혼합 배지를 넣고 각 균주를 접종한 후, 진탕 배양기 (shaking incubator)에 약 250 rpm으로 교반하면서 약 29℃에서 배양하였다.

[0146] 도 8은 알칸 존재하에서 레티노이드 생산 결과를 나타낸 도면이다. 도 9는 알칸 존재하에서 레티노이드 생산 균주의 성장 결과를 나타낸 도면이다.

[0147] 도 8에 나타난 바와 같이, 데칸을 사용한 경우에 총 108 mg/L의 레티노이드가 생산되었다. 그 외 균체 증식, pH 및 균체 내 β -카로틴의 양은 알칸에 따라 큰 차이를 보이지 않았다. 그 결과 도데칸에 비해 데칸이 레티노이드 생산에 더 유리한 것으로 여겨진다. 옥탄을 사용한 경우, 레티날 및 레티놀의 생산은 다른 알칸과 유사하였으나, 레티닐 아세테이트는 거의 생산하지 않았다. 테트라데칸은 전체 레티노이드 생산량이 다른 알칸에 비하여 낮았다.

[0148] (2) 미네랄 오일을 포함 배지에서 레티노이드의 생산

[0149] (2.1) 경량 미네랄 오일

[0150] 경량 미네랄 오일은 알칸에 비하여 가격이 싸다는 장점이 있다. 2% (w/v) 글리세롤 및 0.2% (w/v) 아라비노오스를 함유한 2YT 배지(리터 당 16g 트립톤, 10g 이스트 추출물, 및 5g NaCl) 5mL와 다른 부피의 경량 미네랄 오일 (Sigma, Catalog No.M8410)의 혼합 배지 중에서 72 시간 동안 배양하였다.

[0151] 배양은 길이 15 cm, 지름 25 mm의 시험관에 상기 혼합 배지를 넣고 각 균주를 접종한 후, 진탕 배양기 (shaking incubator)에 약 250 rpm으로 교반하면서 약 29℃에서 배양하였다.

[0152] 도 10은 경량 미네랄 오일 존재하에서 레티노이드 생산 결과를 나타낸 도면이다. 도 11은 경량 미네랄 오일 존재하에서 균주의 성장 결과를 나타낸 도면이다.

[0153] 도 10에 나타난 바와 같이, 도데칸 5mL에서 136.1mg/L이 생산될 때, 2mL 경량 미네랄 오일에서 158mg/L이 생산되었다. 도 11에 나타난 바와 같이, pH는 도데칸 이외의 경우 큰 차이를 보이지 않았다. 반면, 균체 성장은 경량 미네랄 오일의 양이 증가할수록 감소하였다. 이는 경량 미네랄 오일의 높은 점도와 비중으로 인하여 배지와 미네랄 오일이 충분히 혼합되지 않았기 때문인 것으로 예상된다. 균체 성장의 감소로 인해 레티노이드 생산도 감소하였다.

[0154] 도 12는 균체당 레티노이드 생산량 (cell specific retinoids productivity)을 나타낸 도면이다. 도 12에 나타난 바와 같이, 미네랄 오일의 양에 상관없이 약 5mg/L/OD_{600nm}의 비생산성을 보였다.

[0155] (2.2) 중량 미네랄 오일

[0156] 중량 미네랄 오일은 경량 미네랄 오일에 비하여 싸다. 2% (w/v) 글리세롤 및 0.2% (w/v) 아라비노오스를 함유한 2YT 배지(리터 당 16g 트립톤, 10g 이스트 추출물, 및 5g NaCl) 5mL와 중량 미네랄 오일(Daejung, Catalog No.5658-4400) 2mL의 혼합 배지 중에서 96 시간 동안 배양하였다. 배양은 길이 15 cm, 지름 25 mm의 시험관에 상기 혼합 배지를 넣고 각 균주를 접종한 후, 진탕 배양기 (shaking incubator)에 약 250 rpm으로 교반하면서 약 29℃에서 배양하였다.

[0157] 도 13은 중량 미네랄 오일 존재하에서 레티노이드 생산 결과를 나타낸 도면이다. 도 14는 중량 미네랄 오일 존재하에서 균주의 성장 결과를 나타낸 도면이다. 도 13 및 도 14에 나타난 바와 같이, 경량 미네랄 오일, 및 도데칸에 비하여 중량 미네랄 오일의 경우 세포 성장이 적었다. 또한, 레티노이드 96.5mg/L을 생산하였다. 이는 중량 미네랄 오일의 점도에 의하여 배지와 미네랄 오일이 잘 혼합되지 않았기 때문인 것으로 예측된다.

[0158] 사용된 시험관 (test tube)를 기울여서 배양기에 배치시키는 것을 제외하고는 위와 동일하게 세포를 배양하였다. 시험관을 기울임으로써 교반의 효과가 더 증대되어 배지와 미네랄 오일이 더 잘 혼합될 수 있었다.

[0159] 도 15는 시험관을 기울여 배양한 경우 레티노이드 생산 결과를 나타낸 도면이다. 도 16은 시험관을 기울여 배양

한 경우 균주 성장 결과를 나타낸 도면이다. 도 15 및 도 16에 나타난 바와 같이, 기울여 배양한 경우 세포 성장 및 레티노이드 생산이 증가하였다. 구체적으로, 시험관을 수직으로 한 경우, 96시간에서 88.2mg/L의 레티노이드가 생산되었으나, 시험관을 기울인 경우, 173.9mg/L이 생산되었다.

[0160] 이는 경량 및 중량 미네랄 오일은 높은 점도로 인하여 배지와와의 혼합이 레티노이드 생산에 중요한 인자라는 것을 나타낸다. 따라서, 배양 중 적절한 교반을 제공함으로써 레티노이드에 사용될 수 있다.

[0161] (3) 피부 친화적 친유성 물질을 포함 배지에서 레티노이드의 생산

[0162] 피부 친화적 친유성 물질을 포함하는 배지에서 레티노이드를 생산하였다. 피부 친화적 친유성 물질은 이소프로필 미리스테이트 (IPM), 디옥타노일-데카노일 글리세롤(ODO), 세틸 에틸헥사노에이트 (CEH) 및 피토스쿠알란을 사용하였다.

[0163] DH5 α에 pT-DHBSR/pSNA가 형질전환된 균주 DH5 α(pT-DHBSR/pSNA)를 사용하고, 5ml 배지에 2ml의 중량 미네랄 오일을 각각 첨가하고, 2% (w/v) 글리세롤 및 0.2% (w/v) 아라비노오스를 함유한 2YT 배지(리터 당 16g 트립톤, 10g 이스트 추출물, 및 5g NaCl) 5ml와 이소프로필 미리스테이트 (IPM, Sigma, Catalog No.172472), 디옥타노일-데카노일 글리세롤(ODO), 세틸 에틸헥사노에이트 (CEH), 또는 피토스쿠알란((PHYTOSQUALAN[®], Sophim; 분자식 C₃₀H₆₂; hydrogenated form of squalane; extracted from Olive) 2mL 또는 5mL의 혼합 배지 중에서 72 시간 동안 배양하였다.

[0164] 배양은 길이 15 cm, 지름 25 mm의 시험관에 상기 혼합 배지를 넣고 각 균주를 접종한 후, 진탕 배양기 (shaking incubator)에 약 250 rpm으로 교반하면서 약 29℃에서 배양하였다.

[0165] 도 17은 피부 친화적 친유성 물질 존재하에서 세포 성장 및 pH를 나타낸 도면이다. 도 18 및 도 19는 피부 친화적 친유성 물질의 양에 따른 레티노이드 생산결과를 나타낸 도면이다.

[0166] 도 18 및 도 19에 나타난 바와 같이, 도데칸을 제외한 친유성 물질에서 5mL에 비하여 2mL에서 레티노이드 생산량이 많았다. 즉, 경량 미네랄 오일, IPM, ODO, CEH, 및 피토스쿠알란의 배지 5mL에 대하여 약 2mL를 사용한 경우, 레티노이드 생산량이 많았다. IPM, ODO, CEH 및 피토스쿠알란 중 IPM에서 가장 많은 레티노이드가 생산되었다. 특히, IPM 2mL를 첨가한 경우, 180mg/L의 레티노이드가 생산되었다. IPM의 경우, 균체 성장이 비슷한 것을 고려하면, 균체당 비생산성이 높은 것으로 예측된다.

수탁번호

[0167]

기탁기관명 : 한국생명공학연구원

수탁번호 : KCTC11254BP

수탁일자 : 20080107

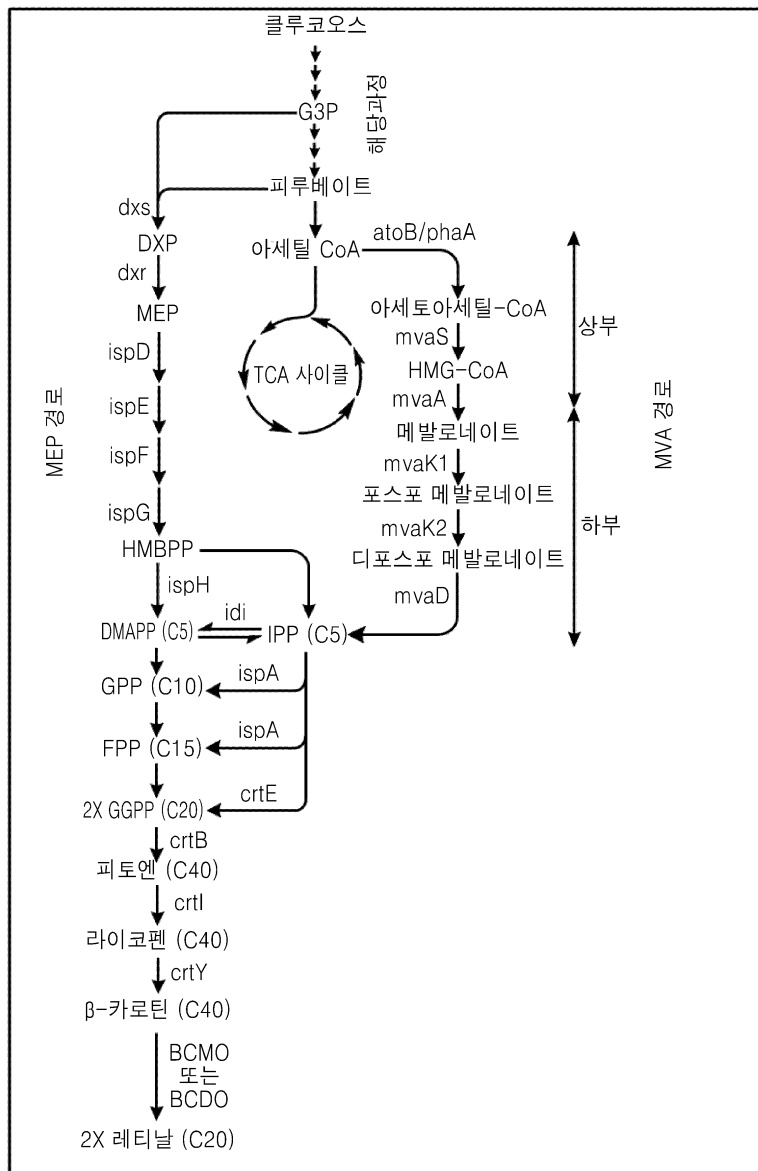
기탁기관명 : 한국생명공학연구원

수탁번호 : KCTC11255BP

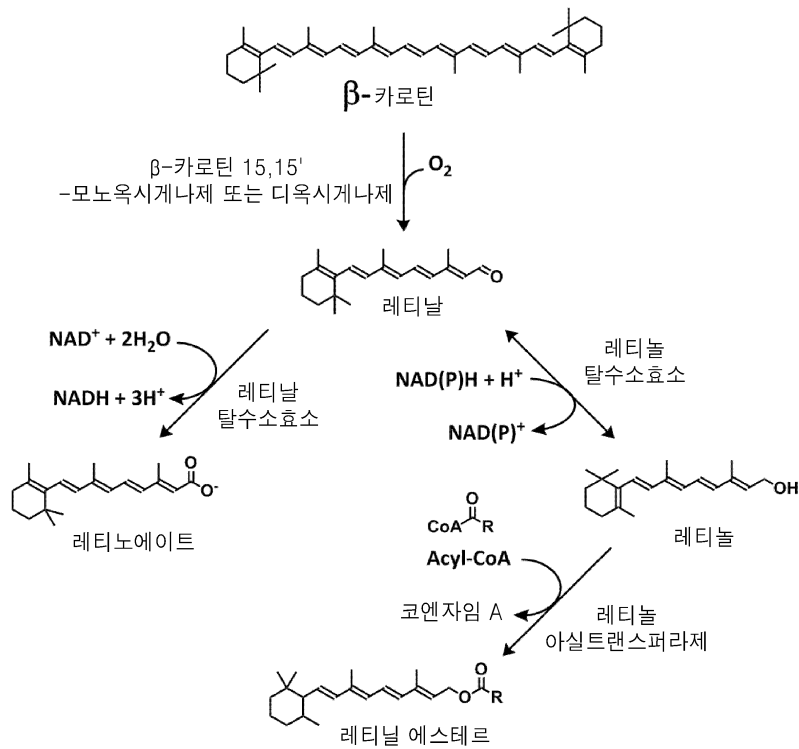
수탁일자 : 20080107

도면

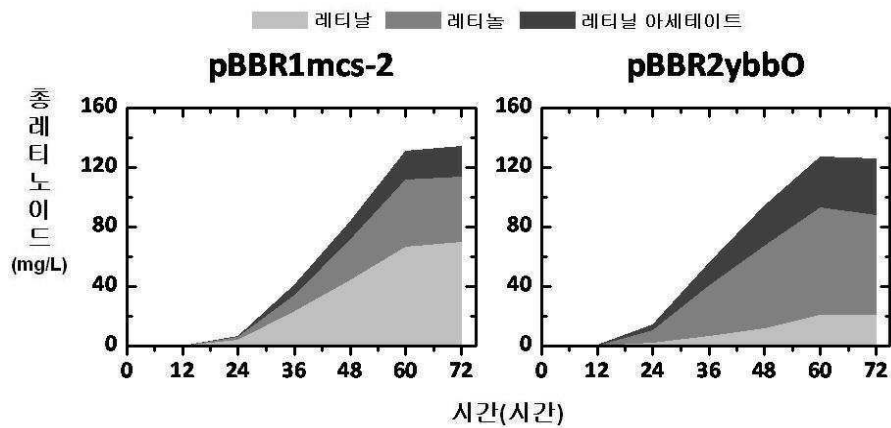
도면1



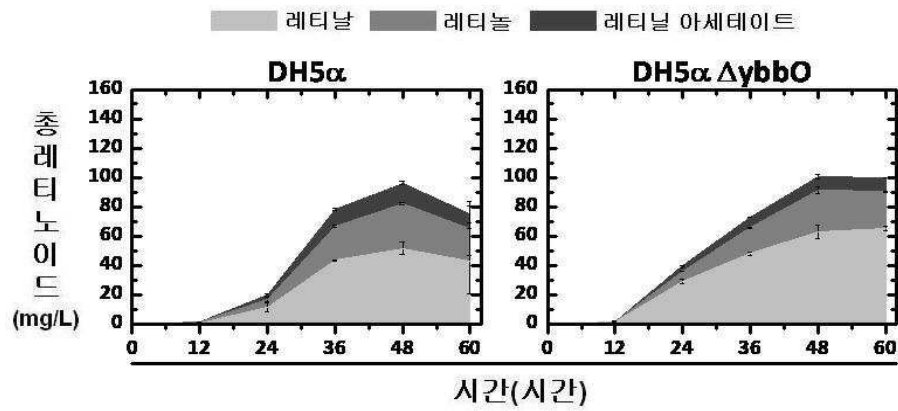
도면2



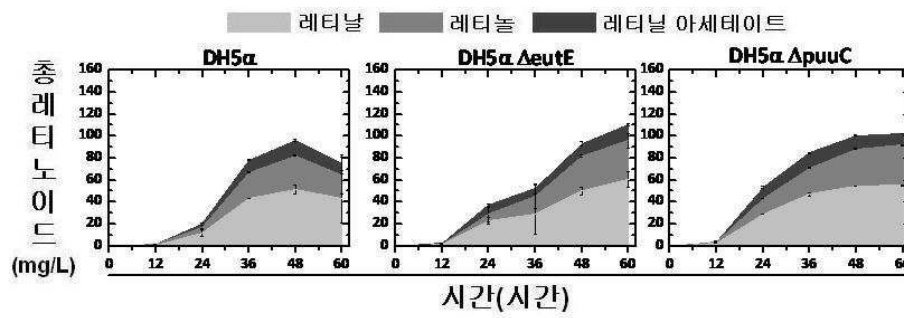
도면3



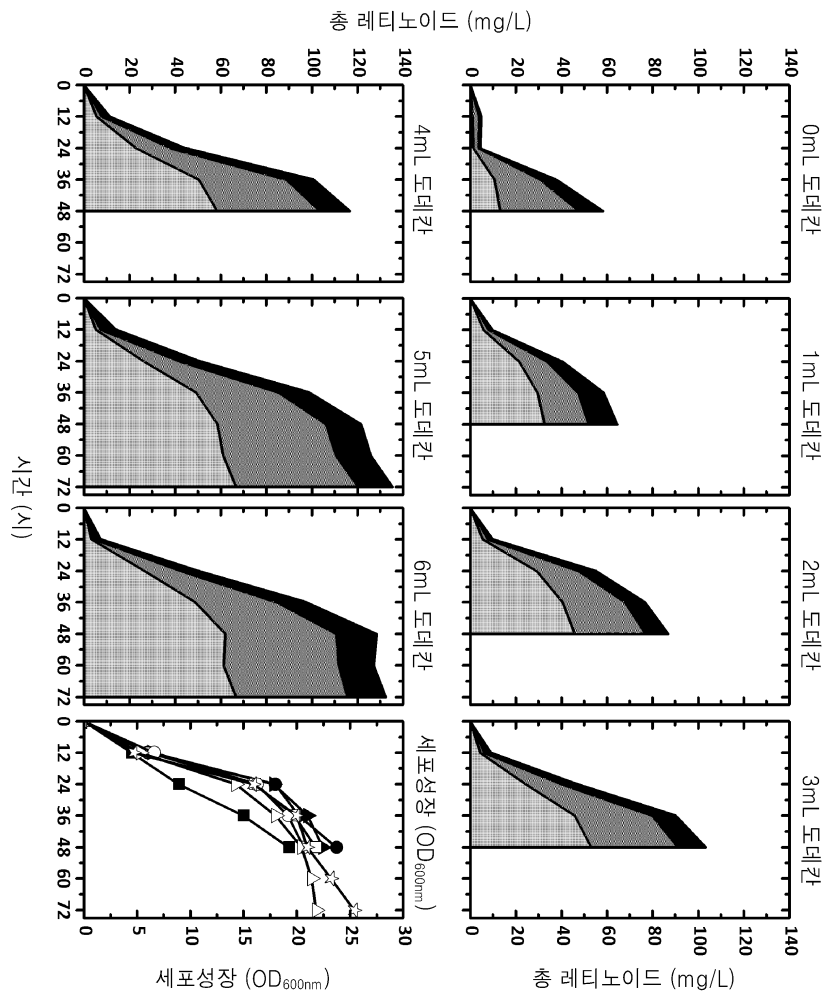
도면4



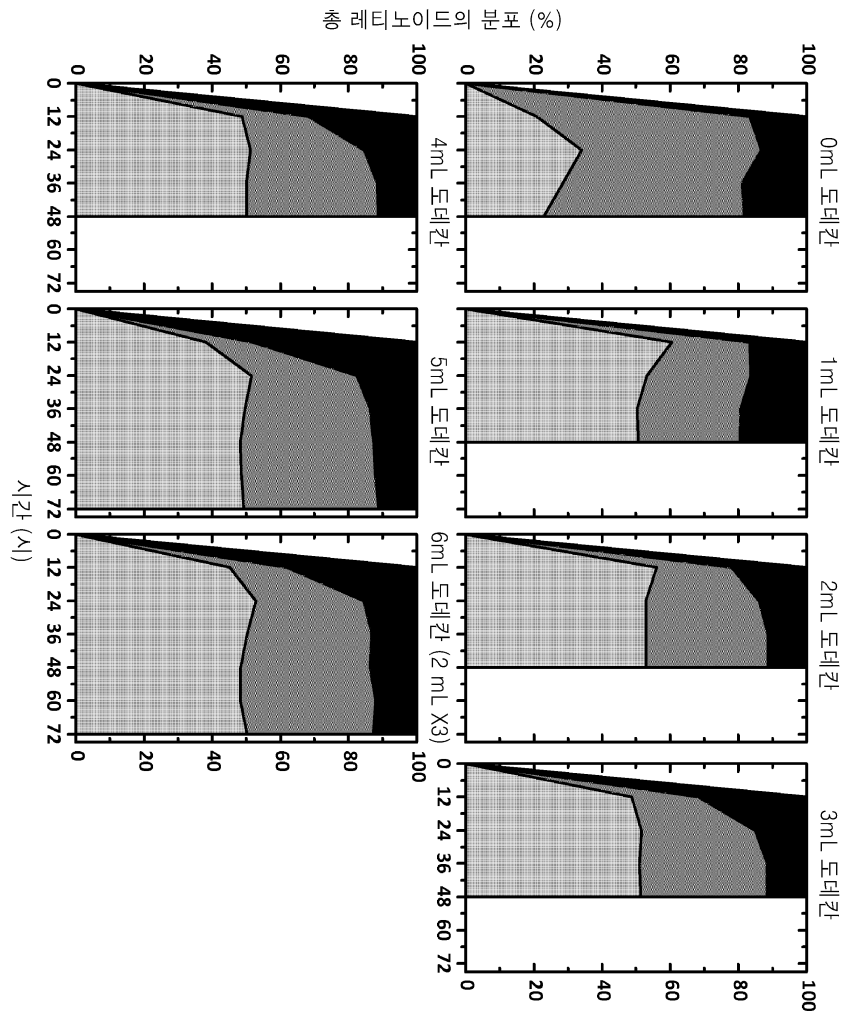
도면5



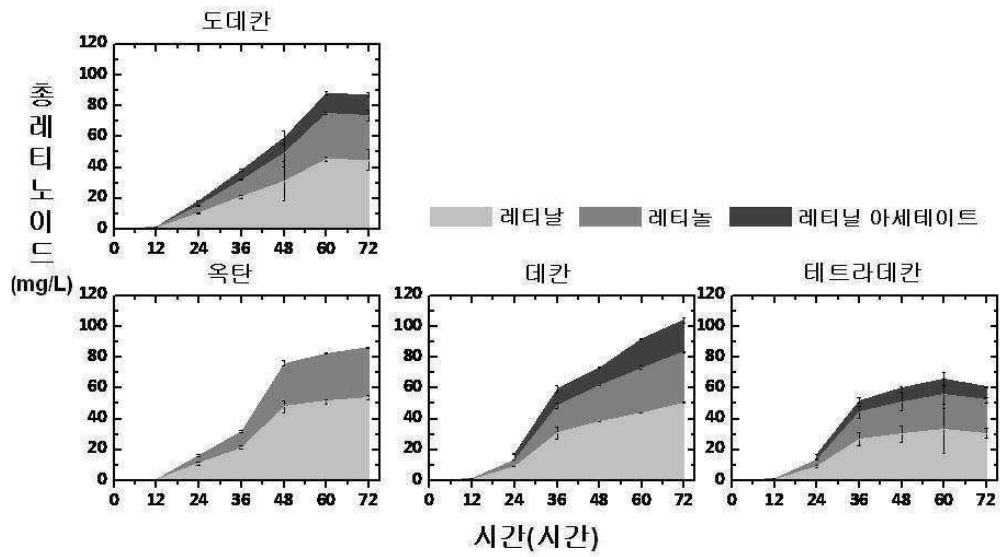
도면6



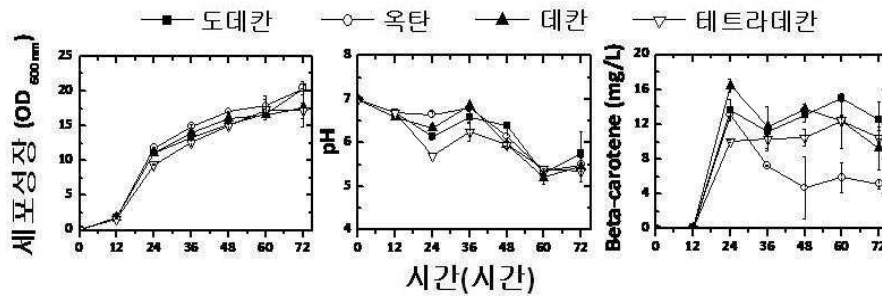
도면7



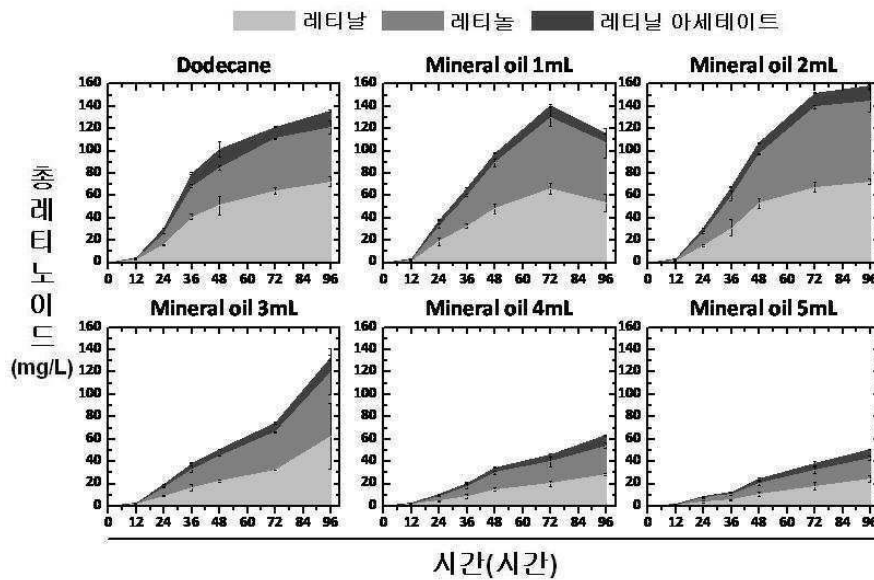
도면8



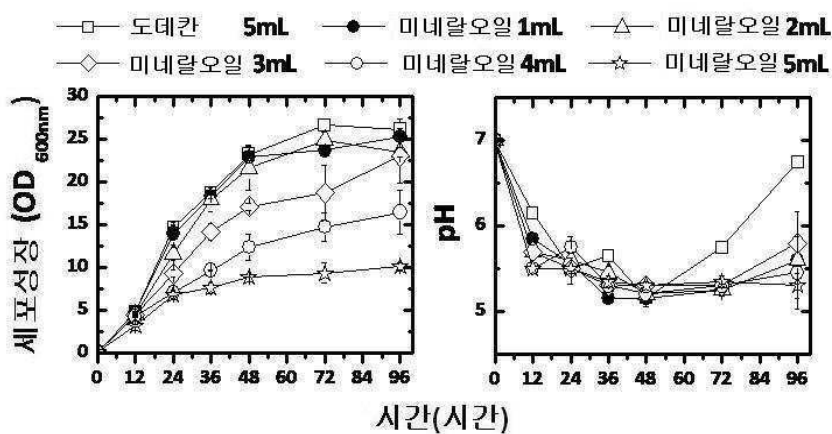
도면9



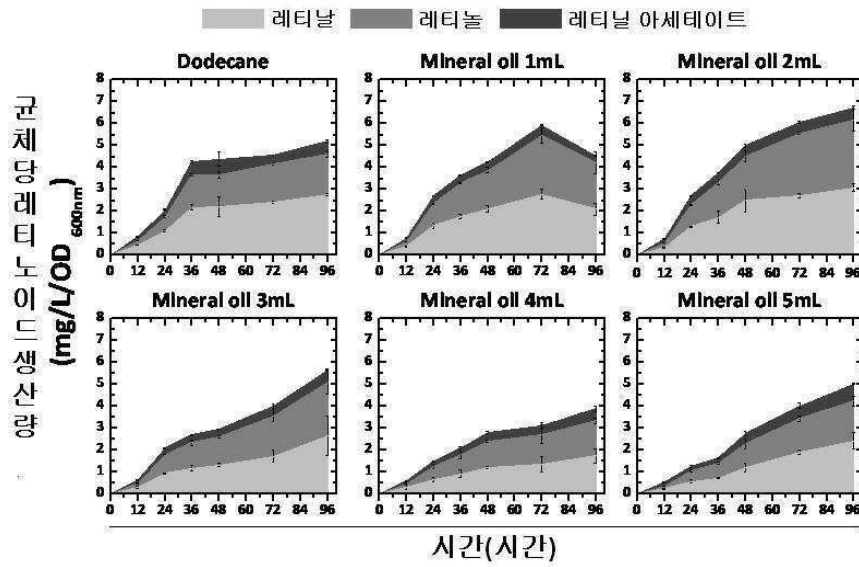
도면10



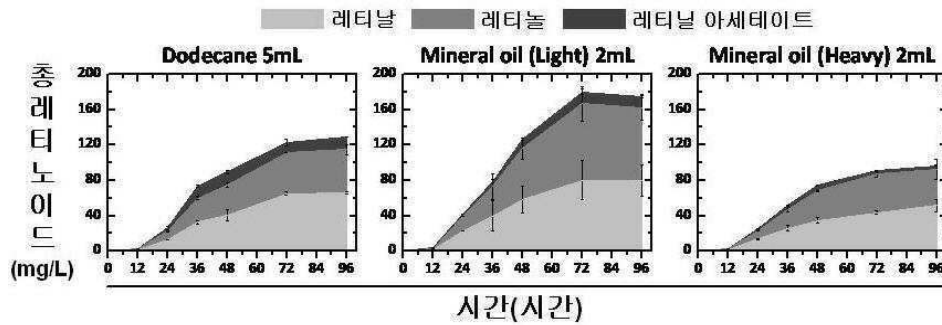
도면11



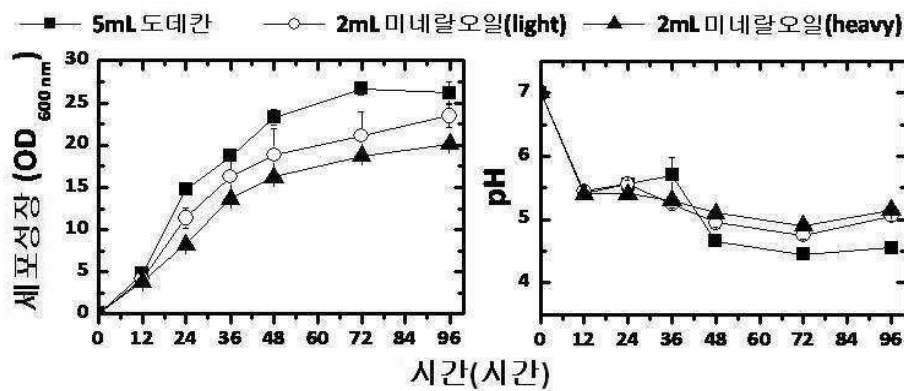
도면12



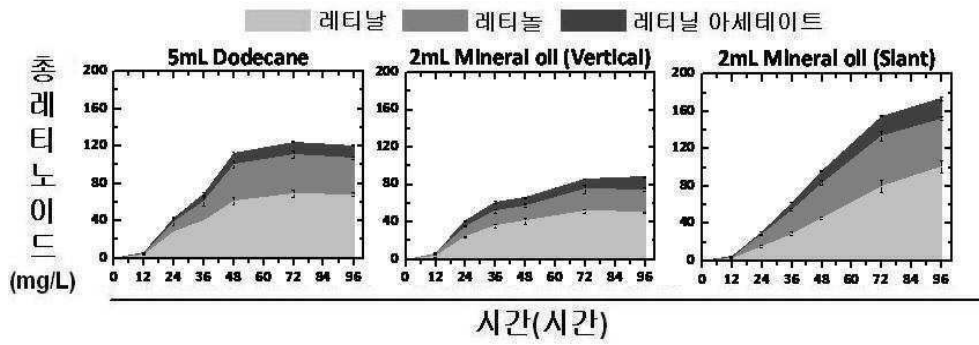
도면13



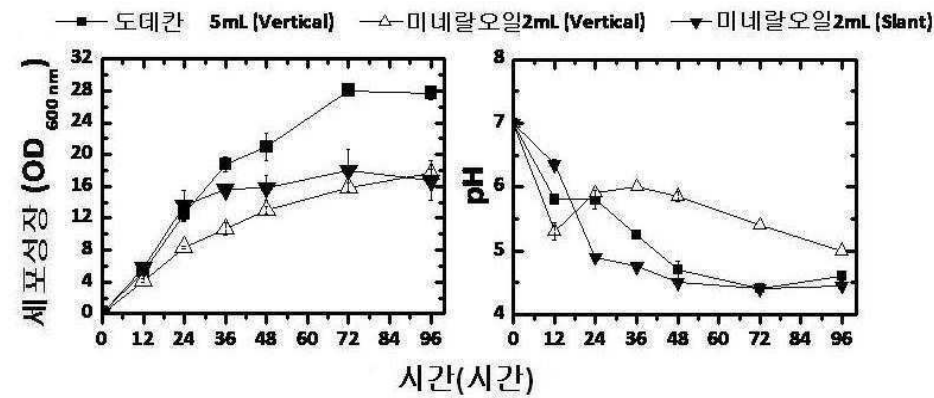
도면14



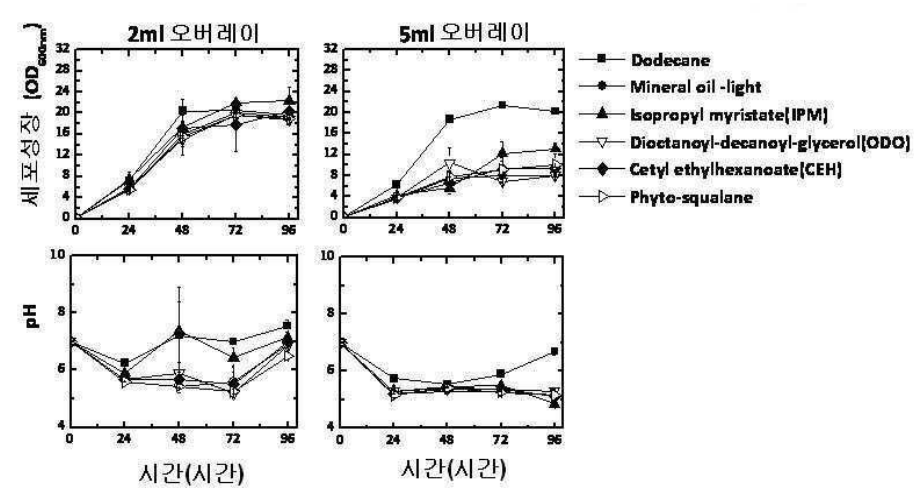
도면15



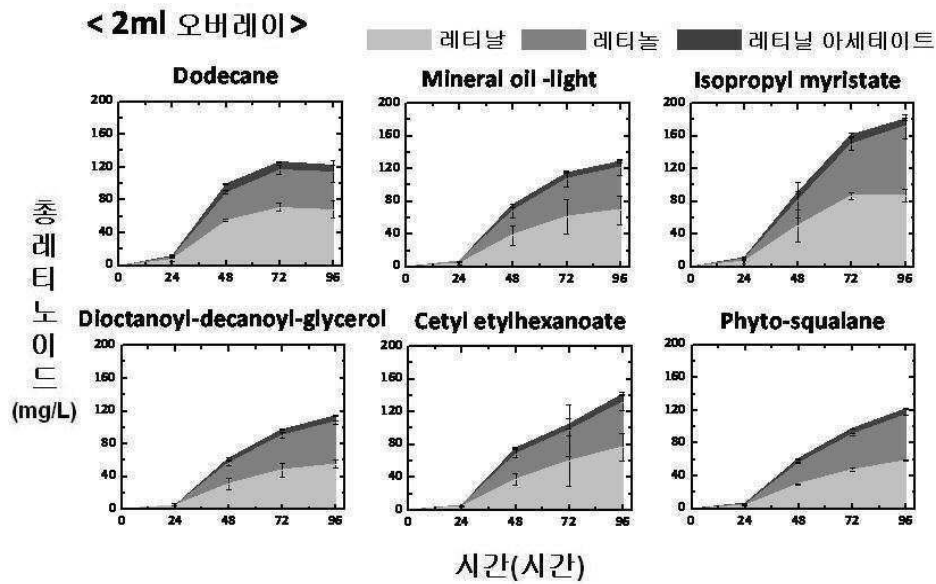
도면16



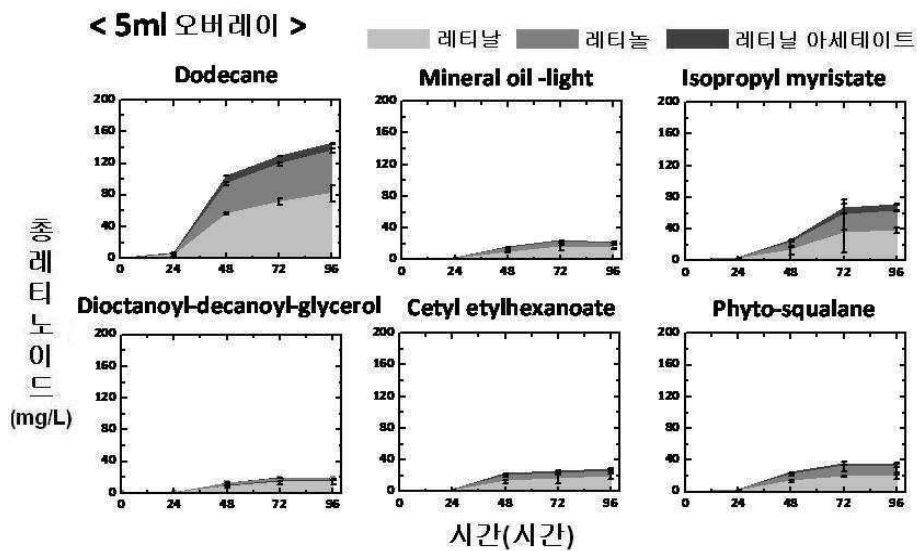
도면17



도면18



도면19



서열 목록

- <110> INDUSTRY-ACADEMIC COOPERATION FOUNDATION GYEONGSANG NATIONAL UNIVERSITY
- <120> A method of producing retinoids using a microorganism of
Escherichia genus having enhanced retinoids productivity
- <130> PN093621
- <160> 56
- <170> Kopatent In 1.71
- <210> 1
- <211> 803
- <212> PRT

<213> Enterococcus faecalis

<400> 1

Met Lys Thr Val Val Ile Ile Asp Ala Leu Arg Thr Pro Ile Gly Lys

1 5 10 15

Tyr Lys Gly Ser Leu Ser Gln Val Ser Ala Val Asp Leu Gly Thr His

20 25 30

Val Thr Thr Gln Leu Leu Lys Arg His Ser Thr Ile Ser Glu Glu Ile

35 40 45

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

50 55 60

Pro Ala Arg Gln Ile Ala Ile Asn Ser Gly Leu Ser His Glu Ile Pro

65 70 75 80

Ala Met Thr Val Asn Glu Val Cys Gly Ser Gly Met Lys Ala Val Ile

85 90 95

Leu Ala Lys Gln Leu Ile Gln Leu Gly Glu Ala Glu Val Leu Ile Ala

100 105 110

Gly Gly Ile Glu Asn Met Ser Gln Ala Pro Lys Leu Gln Arg Phe Asn

115 120 125

Tyr Glu Thr Glu Ser Tyr Asp Ala Pro Phe Ser Ser Met Met Tyr Asp

130 135 140

Gly Leu Thr Asp Ala Phe Ser Gly Gln Ala Met Gly Leu Thr Ala Glu

145 150 155 160

Asn Val Ala Glu Lys Tyr His Val Thr Arg Glu Glu Gln Asp Gln Phe

165 170 175

Ser Val His Ser Gln Leu Lys Ala Ala Gln Ala Gln Ala Glu Gly Ile

180 185 190

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

195 200 205

Lys Asp Glu Gly Ile Arg Pro Asn Ser Ser Val Glu Lys Leu Gly Thr

210 215 220

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

[illegible]

Gln Gly Phe Lys Thr Val Asn Gln Gln Arg Leu Met Arg Gly Gln Ile
485 490 495

Val Phe Tyr Asp Val Ala Asp Ala Glu Ser Leu Ile Asp Glu Leu Gln
500 505 510

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

Ile Val Lys Arg Gly Gly Gly Leu Arg Asp Leu Gln Tyr Arg Ala Phe
530 535 540

Asp Glu Ser Phe Val Ser Val Asp Phe Leu Val Asp Val Lys Asp Ala
545 550 555 560

Met Gly Ala Asn Ile Val Asn Ala Met Leu Glu Gly Val Ala Glu Leu
565 570 575

Phe Arg Glu Trp Phe Ala Glu Gln Lys Ile Leu Phe Ser Ile Leu Ser
580 585 590

Asn Tyr Ala Thr Glu Ser Val Val Thr Met Lys Thr Ala Ile Pro Val
595 600 605

Ser Arg Leu Ser Lys Gly Ser Asn Gly Arg Glu Ile Ala Glu Lys Ile
610 615 620

Val Leu Ala Ser Arg Tyr Ala Ser Leu Asp Pro Tyr Arg Ala Val Thr
625 630 635 640

His Asn Lys Gly Ile Met Asn Gly Ile Glu Ala Val Val Leu Ala Thr
645 650 655

Gly Asn Asp Thr Arg Ala Val Ser Ala Ser Cys His Ala Phe Ala Val
660 665 670

Lys Glu Gly Arg Tyr Gln Gly Leu Thr Ser Trp Thr Leu Asp Gly Glu
675 680 685

Gln Leu Ile Gly Glu Ile Ser Val Pro Leu Ala Leu Ala Thr Val Gly
690 695 700

Gly Ala Thr Lys Val Leu Pro Lys Ser Gln Ala Ala Ala Asp Leu Leu
705 710 715 720

Ala Val Thr Asp Ala Lys Glu Leu Ser Arg Val Val Ala Ala Val Gly
725 730 735

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

740 745 750
Gln Lys Gly His Met Ala Leu Gln Ala Arg Ser Leu Ala Met Thr Val
755 760 765
Gly Ala Thr Gly Lys Glu Val Glu Ala Val Ala Gln Gln Leu Lys Arg
770 775 780
Gln Lys Thr Met Asn Gln Asp Arg Ala Leu Ala Ile Leu Asn Asp Leu
785 790 795 800
Arg Lys Gln

<210> 2
<211> 383
<212> PRT
<213> Enterococcus faecalis

<400> 2
Met Thr Ile Gly Ile Asp Lys Ile Ser Phe Phe Val Pro Pro Tyr Tyr
1 5 10 15
Ile Asp Met Thr Ala Leu Ala Glu Ala Arg Asn Val Asp Pro Gly Lys
20 25 30
Phe His Ile Gly Ile Gly Gln Asp Gln Met Ala Val Asn Pro Ile Ser
35 40 45
Gln Asp Ile Val Thr Phe Ala Ala Asn Ala Ala Glu Ala Ile Leu Thr
50 55 60
Lys Glu Asp Lys Glu Ala Ile Asp Met Val Ile Val Gly Thr Glu Ser
65 70 75 80
Ser Ile Asp Glu Ser Lys Ala Ala Ala Val Val Leu His Arg Leu Met
85 90 95
Gly Ile Gln Pro Phe Ala Arg Ser Phe Glu Ile Lys Glu Ala Cys Tyr
100 105 110
Gly Ala Thr Ala Gly Leu Gln Leu Ala Lys Asn His Val Ala Leu His
115 120 125
Pro Asp Lys Lys Val Leu Val Val Ala Ala Asp Ile Ala Lys Tyr Gly

130 135 140
 Leu Asn Ser Gly Gly Glu Pro Thr Gln Gly Ala Gly Ala Val Ala Met
 145 150 155 160
 Leu Val Ala Ser Glu Pro Arg Ile Leu Ala Leu Lys Glu Asp Asn Val
 165 170 175
 Met Leu Thr Gln Asp Ile Tyr Asp Phe Trp Arg Pro Thr Gly His Pro
 180 185 190
 Tyr Pro Met Val Asp Gly Pro Leu Ser Asn Glu Thr Tyr Ile Gln Ser
 195 200 205
 Phe Ala Gln Val Trp Asp Glu His Lys Lys Arg Thr Gly Leu Asp Phe

 210 215 220
 Ala Asp Tyr Asp Ala Leu Ala Phe His Ile Pro Tyr Thr Lys Met Gly
 225 230 235 240
 Lys Lys Ala Leu Leu Ala Lys Ile Ser Asp Gln Thr Glu Ala Glu Gln
 245 250 255
 Glu Arg Ile Leu Ala Arg Tyr Glu Glu Ser Ile Ile Tyr Ser Arg Arg
 260 265 270
 Val Gly Asn Leu Tyr Thr Gly Ser Leu Tyr Leu Gly Leu Ile Ser Leu
 275 280 285

 Leu Glu Asn Ala Thr Thr Leu Thr Ala Gly Asn Gln Ile Gly Leu Phe
 290 295 300
 Ser Tyr Gly Ser Gly Ala Val Ala Glu Phe Phe Thr Gly Glu Leu Val
 305 310 315 320
 Ala Gly Tyr Gln Asn His Leu Gln Lys Glu Thr His Leu Ala Leu Leu
 325 330 335
 Asp Asn Arg Thr Glu Leu Ser Ile Ala Glu Tyr Glu Ala Met Phe Ala
 340 345 350
 Glu Thr Leu Asp Thr Asp Ile Asp Gln Thr Leu Glu Asp Glu Leu Lys

 355 360 365
 Tyr Ser Ile Ser Ala Ile Asn Asn Thr Val Arg Ser Tyr Arg Asn
 370 375 380

<210> 3
 <211> 292
 <212> PRT
 <213> Streptococcus pneumoniae
 <400> 3
 Met Thr Lys Lys Val Gly Val Gly Gln Ala His Ser Lys Ile Ile Leu
 1 5 10 15
 Ile Gly Glu His Ala Val Val Tyr Gly Tyr Pro Ala Ile Ser Leu Pro
 20 25 30
 Leu Leu Glu Val Glu Val Thr Cys Lys Val Val Ser Ala Glu Ser Pro
 35 40 45
 Trp Arg Leu Tyr Glu Glu Asp Thr Leu Ser Met Ala Val Tyr Ala Ser
 50 55 60
 Leu Glu Tyr Leu Asp Ile Thr Glu Ala Cys Val Arg Cys Glu Ile Asp
 65 70 75 80
 Ser Ala Ile Pro Glu Lys Arg Gly Met Gly Ser Ser Ala Ala Ile Ser
 85 90 95
 Ile Ala Ala Ile Arg Ala Val Phe Asp Tyr Tyr Gln Ala Asp Leu Pro
 100 105 110
 His Asp Val Leu Glu Ile Leu Val Asn Arg Ala Glu Met Ile Ala His
 115 120 125
 Met Asn Pro Ser Gly Leu Asp Ala Lys Thr Cys Leu Ser Asp Gln Pro
 130 135 140
 Ile Arg Phe Ile Lys Asn Val Gly Phe Thr Glu Leu Glu Met Asp Leu
 145 150 155 160
 Ser Ala Tyr Leu Val Ile Ala Asp Thr Gly Val Tyr Gly His Thr Arg
 165 170 175
 Glu Ala Ile Gln Val Val Gln Asn Lys Gly Lys Asp Ala Leu Pro Phe
 180 185 190
 Leu His Ala Leu Gly Glu Leu Thr Gln Gln Ala Glu Val Ala Ile Ser
 195 200 205
 Gln Lys Tyr Ala Glu Gly Leu Gly Leu Ile Phe Ser Gln Ala His Leu

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

 Gly Leu Gly Gly Cys Ile Ile Ala Leu Val Thr Asn Leu Thr His Ala
 260 265 270
 Gln Glu Leu Ala Glu Arg Leu Glu Glu Lys Gly Ala Val Gln Thr Trp
 275 280 285
 Ile Glu Ser Leu
 290
 <210> 4
 <211> 336
 <212> PRT
 <213> Streptococcus pneumoniae
 <400> 4
 Met Ile Ala Val Lys Thr Cys Gly Lys Leu Tyr Trp Ala Gly Glu Tyr
 1 5 10 15
 Ala Ile Leu Glu Pro Gly Gln Leu Ala Leu Ile Lys Asp Ile Pro Ile

 20 25 30
 Tyr Met Arg Ala Glu Ile Ala Phe Ser Asp Ser Tyr Arg Ile Tyr Ser
 35 40 45
 Asp Met Phe Asp Phe Ala Val Asp Leu Arg Pro Asn Pro Asp Tyr Ser
 50 55 60
 Leu Ile Gln Glu Thr Ile Ala Leu Met Gly Asp Phe Leu Ala Val Arg
 65 70 75 80
 Gly Gln Asn Leu Arg Pro Phe Ser Leu Lys Ile Cys Gly Lys Met Glu
 85 90 95

 Arg Glu Gly Lys Lys Phe Gly Leu Gly Ser Ser Gly Ser Val Val Val
 100 105 110
 Leu Val Val Lys Ala Leu Leu Ala Leu Tyr Asn Leu Ser Val Asp Gln
 115 120 125

Asn Leu Leu Phe Lys Leu Thr Ser Ala Val Leu Leu Lys Arg Gly Asp
 130 135 140
 Asn Gly Ser Met Gly Asp Leu Ala Cys Ile Val Ala Glu Asp Leu Val
 145 150 155 160
 Leu Tyr Gln Ser Phe Asp Arg Gln Lys Ala Ala Ala Trp Leu Glu Glu
 165 170 175
 Glu Asn Leu Ala Thr Val Leu Glu Arg Asp Trp Gly Phe Phe Ile Ser
 180 185 190
 Gln Val Lys Pro Thr Leu Glu Cys Asp Phe Leu Val Gly Trp Thr Lys
 195 200 205
 Glu Val Ala Val Ser Ser His Met Val Gln Gln Ile Lys Gln Asn Ile
 210 215 220
 Asn Gln Asn Phe Leu Ser Ser Ser Lys Glu Thr Val Val Ser Leu Val
 225 230 235 240
 Glu Ala Leu Glu Gln Gly Lys Ala Glu Lys Val Ile Glu Gln Val Glu
 245 250 255
 Val Ala Ser Lys Leu Leu Glu Gly Leu Ser Thr Asp Ile Tyr Thr Pro
 260 265 270
 Leu Leu Arg Gln Leu Lys Glu Ala Ser Gln Asp Leu Gln Ala Val Ala
 275 280 285
 Lys Ser Ser Gly Ala Gly Gly Gly Asp Cys Gly Ile Ala Leu Ser Phe
 290 295 300
 Asp Ala Gln Ser Ser Arg Asn Thr Leu Lys Asn Arg Trp Ala Asp Leu
 305 310 315 320
 Gly Ile Glu Leu Leu Tyr Gln Glu Arg Ile Gly His Asp Asp Lys Ser
 325 330 335
 <210> 5
 <211> 317
 <212> PRT
 <213> Streptococcus pneumoniae
 <400> 5
 Met Asp Arg Glu Pro Val Thr Val Arg Ser Tyr Ala Asn Ile Ala Ile

1 5 10 15
 Ile Lys Tyr Trp Gly Lys Lys Lys Glu Lys Glu Met Val Pro Ala Thr
 20 25 30

 Ser Ser Ile Ser Leu Thr Leu Glu Asn Met Tyr Thr Glu Thr Thr Leu
 35 40 45
 Ser Pro Leu Pro Ala Asn Val Thr Ala Asp Glu Phe Tyr Ile Asn Gly
 50 55 60
 Gln Leu Gln Asn Glu Val Glu His Ala Lys Met Ser Lys Ile Ile Asp
 65 70 75 80
 Arg Tyr Arg Pro Ala Gly Glu Gly Phe Val Arg Ile Asp Thr Gln Asn
 85 90 95
 Asn Met Pro Thr Ala Ala Gly Leu Ser Ser Ser Ser Ser Gly Leu Ser

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

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

 245 250 255

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

<210> 6
 <211> 182
 <212> PRT
 <213> E. coli K12 MG1655
 <400> 6

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

50 55 60
 Ser Val Cys Gly His Pro Gln Leu Gly Glu Ser Asn Glu Asp Ala Val
 65 70 75 80
 Ile Arg Arg Cys Arg Tyr Glu Leu Gly Val Glu Ile Thr Pro Pro Glu
 85 90 95
 Ser Ile Tyr Pro Asp Phe Arg Tyr Arg Ala Thr Asp Pro Ser Gly Ile
 100 105 110
 Val Glu Asn Glu Val Cys Pro Val Phe Ala Ala Arg Thr Thr Ser Ala
 115 120 125

Leu Gln Ile Asn Asp Asp Glu Val Met Asp Tyr Gln Trp Cys Asp Leu
 130 135 140
 Ala Asp Val Leu His Gly Ile Asp Ala Thr Pro Trp Ala Phe Ser Pro

145 150 155 160
Trp Met Val Met Gln Ala Thr Asn Arg Glu Ala Arg Lys Arg Leu Ser
165 170 175

Ala Phe Thr Gln Leu Lys
180

<210> 7

<211> 307

<212> PRT

<213> Pantoea agglomerans

<400> 7

Met Val Ser Gly Ser Lys Ala Gly Val Ser Pro His Arg Glu Ile Glu

1 5 10 15
Val Met Arg Gln Ser Ile Asp Asp His Leu Ala Gly Leu Leu Pro Glu
20 25 30

Thr Asp Ser Gln Asp Ile Val Ser Leu Ala Met Arg Glu Gly Val Met
35 40 45

Ala Pro Gly Lys Arg Ile Arg Pro Leu Leu Met Leu Leu Ala Ala Arg
50 55 60

Asp Leu Arg Tyr Gln Gly Ser Met Pro Thr Leu Leu Asp Leu Ala Cys

65 70 75 80
Ala Val Glu Leu Thr His Thr Ala Ser Leu Met Leu Asp Asp Met Pro
85 90 95

Cys Met Asp Asn Ala Glu Leu Arg Arg Gly Gln Pro Thr Thr His Lys
100 105 110

Lys Phe Gly Glu Ser Val Ala Ile Leu Ala Ser Val Gly Leu Leu Ser
115 120 125

Lys Ala Phe Gly Leu Ile Ala Ala Thr Gly Asp Leu Pro Gly Glu Arg
130 135 140

Arg Ala Gln Ala Val Asn Glu Leu Ser Thr Ala Val Gly Val Gln Gly
145 150 155 160
Leu Val Leu Gly Gln Phe Arg Asp Leu Asn Asp Ala Ala Leu Asp Arg
165 170 175

Thr Pro Asp Ala Ile Leu Ser Thr Asn His Leu Lys Thr Gly Ile Leu
180 185 190

Phe Ser Ala Met Leu Gln Ile Val Ala Ile Ala Ser Ala Ser Ser Pro
195 200 205

Ser Thr Arg Glu Thr Leu His Ala Phe Ala Leu Asp Phe Gly Gln Ala
210 215 220

Phe Gln Leu Leu Asp Asp Leu Arg Asp Asp His Pro Glu Thr Gly Lys
225 230 235 240

Asp Arg Asn Lys Asp Ala Gly Lys Ser Thr Leu Val Asn Arg Leu Gly
245 250 255

Ala Asp Ala Ala Arg Gln Lys Leu Arg Glu His Ile Asp Ser Ala Asp
260 265 270

Lys His Leu Thr Phe Ala Cys Pro Gln Gly Gly Ala Ile Arg Gln Phe
275 280 285

Met His Leu Trp Phe Gly His His Leu Ala Asp Trp Ser Pro Val Met
290 295 300

Lys Ile Ala
305

<210> 8
<211> 309
<212> PRT
<213> Pantoea agglomerans
<400> 8

Met Ser Gln Pro Pro Leu Leu Asp His Ala Thr Gln Thr Met Ala Asn
1 5 10 15

Gly Ser Lys Ser Phe Ala Thr Ala Ala Lys Leu Phe Asp Pro Ala Thr
20 25 30

Arg Arg Ser Val Leu Met Leu Tyr Thr Trp Cys Arg His Cys Asp Asp
35 40 45

Val Ile Asp Asp Gln Thr His Gly Phe Ala Ser Glu Ala Ala Ala Glu
50 55 60

Glu Glu Ala Thr Gln Arg Leu Ala Arg Leu Arg Thr Leu Thr Leu Ala

65					70				75				80			
Ala	Phe	Glu	Gly	Ala	Glu	Met	Gln	Asp	Pro	Ala	Phe	Ala	Ala	Phe	Gln	
				85				90				95				
Glu	Val	Ala	Leu	Thr	His	Gly	Ile	Thr	Pro	Arg	Met	Ala	Leu	Asp	His	
				100				105				110				
Leu	Asp	Gly	Phe	Ala	Met	Asp	Val	Ala	Gln	Thr	Arg	Tyr	Val	Thr	Phe	
				115				120				125				
Glu	Asp	Thr	Leu	Arg	Tyr	Cys	Tyr	His	Val	Ala	Gly	Val	Val	Gly	Leu	
				130				135				140				
Met	Met	Ala	Arg	Val	Met	Gly	Val	Arg	Asp	Glu	Arg	Val	Leu	Asp	Arg	
145				150				155				160				
Ala	Cys	Asp	Leu	Gly	Leu	Ala	Phe	Gln	Leu	Thr	Asn	Ile	Ala	Arg	Asp	
				165				170				175				
Ile	Ile	Asp	Asp	Ala	Ala	Ile	Asp	Arg	Cys	Tyr	Leu	Pro	Ala	Glu	Trp	
				180				185				190				
Leu	Gln	Asp	Ala	Gly	Leu	Thr	Pro	Glu	Asn	Tyr	Ala	Ala	Arg	Glu	Asn	
				195				200				205				
Arg	Ala	Ala	Leu	Ala	Arg	Val	Ala	Glu	Arg	Leu	Ile	Asp	Ala	Ala	Glu	
210				215				220								
Pro	Tyr	Tyr	Ile	Ser	Ser	Gln	Ala	Gly	Leu	His	Asp	Leu	Pro	Pro	Arg	
225				230				235				240				
Cys	Ala	Trp	Ala	Ile	Ala	Thr	Ala	Arg	Ser	Val	Tyr	Arg	Glu	Ile	Gly	
				245				250				255				
Ile	Lys	Val	Lys	Ala	Ala	Gly	Gly	Ser	Ala	Trp	Asp	Arg	Arg	Gln	His	
				260				265				270				
Thr	Ser	Lys	Gly	Glu	Lys	Ile	Ala	Met	Leu	Met	Ala	Ala	Pro	Gly	Gln	
				275				280				285				
Val	Ile	Arg	Ala	Lys	Thr	Thr	Arg	Val	Thr	Pro	Arg	Pro	Ala	Gly	Leu	
290				295				300								
Trp Gln Arg Pro Val																
305																
<210> 9																

<211> 486

<212> PRT

<213> Pantoea agglomerans

<400> 9

Met Lys Lys Thr Val Val Ile Gly Ala Gly Phe Gly Gly Leu Ala Leu

1 5 10 15
Ala Ile Arg Leu Gln Ala Ala Gly Ile Pro Thr Val Leu Leu Glu Gln
20 25 30
Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Trp His Asp Gln Gly Phe
35 40 45
Thr Phe Asp Ala Gly Pro Thr Val Ile Thr Asp Pro Thr Ala Leu Glu
50 55 60
Ala Leu Phe Thr Leu Ala Gly Arg Arg Met Glu Asp Tyr Val Arg Leu
65 70 75 80

Leu Pro Val Lys Pro Phe Tyr Arg Leu Cys Trp Glu Ser Gly Lys Thr
85 90 95
Leu Asp Tyr Ala Asn Asp Ser Ala Glu Leu Glu Ala Gln Ile Thr Gln
100 105 110
Phe Asn Pro Arg Asp Val Glu Gly Tyr Arg Arg Phe Leu Ala Tyr Ser
115 120 125
Gln Ala Val Phe Gln Glu Gly Tyr Leu Arg Leu Gly Ser Val Pro Phe
130 135 140
Leu Ser Phe Arg Asp Met Leu Arg Ala Gly Pro Gln Leu Leu Lys Leu

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

210 215 220
 Asn Gly Met Val Lys Leu Phe Thr Asp Leu Gly Gly Glu Ile Glu Leu
 225 230 235 240
 Asn Ala Arg Val Glu Glu Leu Val Val Ala Asp Asn Arg Val Ser Gln
 245 250 255
 Val Arg Leu Ala Asp Gly Arg Ile Phe Asp Thr Asp Ala Val Ala Ser
 260 265 270
 Asn Ala Asp Val Val Asn Thr Tyr Lys Lys Leu Leu Gly His His Pro
 275 280 285
 Val Gly Gln Lys Arg Ala Ala Ala Leu Glu Arg Lys Ser Met Ser Asn

 290 295 300
 Ser Leu Phe Val Leu Tyr Phe Gly Leu Asn Gln Pro His Ser Gln Leu
 305 310 315 320
 Ala His His Thr Ile Cys Phe Gly Pro Arg Tyr Arg Glu Leu Ile Asp
 325 330 335
 Glu Ile Phe Thr Gly Ser Ala Leu Ala Asp Asp Phe Ser Leu Tyr Leu
 340 345 350
 His Ser Pro Cys Val Thr Asp Pro Ser Leu Ala Pro Pro Gly Cys Ala
 355 360 365

 Ser Phe Tyr Val Leu Ala Pro Val Pro His Leu Gly Asn Ala Pro Leu
 370 375 380
 Asp Trp Ala Gln Glu Gly Pro Lys Leu Arg Asp Arg Ile Phe Asp Tyr
 385 390 395 400
 Leu Glu Glu Arg Tyr Met Pro Gly Leu Arg Ser Gln Leu Val Thr Gln
 405 410 415
 Arg Ile Phe Thr Pro Ala Asp Phe His Asp Thr Leu Asp Ala His Leu
 420 425 430
 Gly Ser Ala Phe Ser Ile Glu Pro Leu Leu Thr Gln Ser Ala Trp Phe

 435 440 445
 Arg Pro His Asn Arg Asp Ser Asp Ile Ala Asn Leu Tyr Leu Val Gly
 450 455 460

Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Val Ala Ser Ala
465 470 475 480
Lys Ala Thr Ala Ser Leu
485
<210> 10
<211> 382
<212> PRT
<213> *Pantoea ananatis*
<400> 10
Met Gln Pro His Tyr Asp Leu Ile Leu Val Gly Ala Gly Leu Ala Asn
1 5 10 15

Gly Leu Ile Ala Leu Arg Leu Gln Gln Gln Gln Pro Asp Met Arg Ile
20 25 30
Leu Leu Ile Asp Ala Ala Pro Gln Ala Gly Gly Asn His Thr Trp Ser
35 40 45
Phe His His Asp Asp Leu Thr Glu Ser Gln His Arg Trp Ile Ala Pro
50 55 60
Leu Val Val His His Trp Pro Asp Tyr Gln Val Arg Phe Pro Thr Arg
65 70 75 80
Arg Arg Lys Leu Asn Ser Gly Tyr Phe Cys Ile Thr Ser Gln Arg Phe
85 90 95
Ala Glu Val Leu Gln Arg Gln Phe Gly Pro His Leu Trp Met Asp Thr
100 105 110
Ala Val Ala Glu Val Asn Ala Glu Ser Val Arg Leu Lys Lys Gly Gln
115 120 125
Val Ile Gly Ala Arg Ala Val Ile Asp Gly Arg Gly Tyr Ala Ala Asn
130 135 140
Ser Ala Leu Ser Val Gly Phe Gln Ala Phe Ile Gly Gln Glu Trp Arg
145 150 155 160

Leu Ser His Pro His Gly Leu Ser Ser Pro Ile Ile Met Asp Ala Thr
165 170 175
Val Asp Gln Gln Asn Gly Tyr Arg Phe Val Tyr Ser Leu Pro Leu Ser

180	185	190	
Pro Thr Arg Leu Leu Ile Glu Asp Thr His Tyr Ile Asp Asn Ala Thr			
195	200	205	
Leu Asp Pro Glu Cys Ala Arg Gln Asn Ile Cys Asp Tyr Ala Ala Gln			
210	215	220	
Gln Gly Trp Gln Leu Gln Thr Leu Leu Arg Glu Glu Gln Gly Ala Leu			
225	230	235	240
Pro Ile Thr Leu Ser Gly Asn Ala Asp Ala Phe Trp Gln Gln Arg Pro			
245	250	255	
Leu Ala Cys Ser Gly Leu Arg Ala Gly Leu Phe His Pro Thr Thr Gly			
260	265	270	
Tyr Ser Leu Pro Leu Ala Val Ala Val Ala Asp Arg Leu Ser Ala Leu			
275	280	285	
Asp Val Phe Thr Ser Ala Ser Ile His His Ala Ile Thr His Phe Ala			
290	295	300	
Arg Glu Arg Trp Gln Gln Gln Gly Phe Phe Arg Met Leu Asn Arg Met			
305	310	315	320
Leu Phe Leu Ala Gly Pro Ala Asp Ser Arg Trp Arg Val Met Gln Arg			
325	330	335	
Phe Tyr Gly Leu Pro Glu Asp Leu Ile Ala Arg Phe Tyr Ala Gly Lys			
340	345	350	
Leu Thr Leu Thr Asp Arg Leu Arg Ile Leu Ser Gly Lys Pro Pro Val			
355	360	365	
Pro Val Leu Ala Ala Leu Gln Ala Ile Met Thr Thr His Arg			

370	375	380	
<210> 11			
<211> 620			
<212> PRT			
<213> E. coli K12 MG1655			
<400> 11			
Met Ser Phe Asp Ile Ala Lys Tyr Pro Thr Leu Ala Leu Val Asp Ser			
1	5	10	15

Thr Gln Glu Leu Arg Leu Leu Pro Lys Glu Ser Leu Pro Lys Leu Cys
 20 25 30
 Asp Glu Leu Arg Arg Tyr Leu Leu Asp Ser Val Ser Arg Ser Ser Gly
 35 40 45
 His Phe Ala Ser Gly Leu Gly Thr Val Glu Leu Thr Val Ala Leu His
 50 55 60
 Tyr Val Tyr Asn Thr Pro Phe Asp Gln Leu Ile Trp Asp Val Gly His
 65 70 75 80
 Gln Ala Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Lys Ile Gly
 85 90 95
 Thr Ile Arg Gln Lys Gly Gly Leu His Pro Phe Pro Trp Arg Gly Glu
 100 105 110
 Ser Glu Tyr Asp Val Leu Ser Val Gly His Ser Ser Thr Ser Ile Ser
 115 120 125
 Ala Gly Ile Gly Ile Ala Val Ala Ala Glu Lys Glu Gly Lys Asn Arg
 130 135 140
 Arg Thr Val Cys Val Ile Gly Asp Gly Ala Ile Thr Ala Gly Met Ala
 145 150 155 160
 Phe Glu Ala Met Asn His Ala Gly Asp Ile Arg Pro Asp Met Leu Val
 165 170 175
 Ile Leu Asn Asp Asn Glu Met Ser Ile Ser Glu Asn Val Gly Ala Leu
 180 185 190
 Asn Asn His Leu Ala Gln Leu Leu Ser Gly Lys Leu Tyr Ser Ser Leu
 195 200 205
 Arg Glu Gly Gly Lys Lys Val Phe Ser Gly Val Pro Pro Ile Lys Glu
 210 215 220
 Leu Leu Lys Arg Thr Glu Glu His Ile Lys Gly Met Val Val Pro Gly
 225 230 235 240
 Thr Leu Phe Glu Glu Leu Gly Phe Asn Tyr Ile Gly Pro Val Asp Gly
 245 250 255
 His Asp Val Leu Gly Leu Ile Thr Thr Leu Lys Asn Met Arg Asp Leu

260 265 270
 Lys Gly Pro Gln Phe Leu His Ile Met Thr Lys Lys Gly Arg Gly Tyr
 275 280 285
 Glu Pro Ala Glu Lys Asp Pro Ile Thr Phe His Ala Val Pro Lys Phe
 290 295 300
 Asp Pro Ser Ser Gly Cys Leu Pro Lys Ser Ser Gly Gly Leu Pro Ser
 305 310 315 320
 Tyr Ser Lys Ile Phe Gly Asp Trp Leu Cys Glu Thr Ala Ala Lys Asp
 325 330 335
 Asn Lys Leu Met Ala Ile Thr Pro Ala Met Arg Glu Gly Ser Gly Met
 340 345 350
 Val Glu Phe Ser Arg Lys Phe Pro Asp Arg Tyr Phe Asp Val Ala Ile
 355 360 365
 Ala Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Ile Gly Gly
 370 375 380
 Tyr Lys Pro Ile Val Ala Ile Tyr Ser Thr Phe Leu Gln Arg Ala Tyr
 385 390 395 400
 Asp Gln Val Leu His Asp Val Ala Ile Gln Lys Leu Pro Val Leu Phe
 405 410 415
 Ala Ile Asp Arg Ala Gly Ile Val Gly Ala Asp Gly Gln Thr His Gln
 420 425 430
 Gly Ala Phe Asp Leu Ser Tyr Leu Arg Cys Ile Pro Glu Met Val Ile
 435 440 445
 Met Thr Pro Ser Asp Glu Asn Glu Cys Arg Gln Met Leu Tyr Thr Gly
 450 455 460
 Tyr His Tyr Asn Asp Gly Pro Ser Ala Val Arg Tyr Pro Arg Gly Asn
 465 470 475 480
 Ala Val Gly Val Glu Leu Thr Pro Leu Glu Lys Leu Pro Ile Gly Lys
 485 490 495
 Gly Ile Val Lys Arg Arg Gly Glu Lys Leu Ala Ile Leu Asn Phe Gly
 500 505 510

Thr Leu Met Pro Glu Ala Ala Lys Val Ala Glu Ser Leu Asn Ala Thr
515 520 525

Leu Val Asp Met Arg Phe Val Lys Pro Leu Asp Glu Ala Leu Ile Leu
530 535 540

Glu Met Ala Ala Ser His Glu Ala Leu Val Thr Val Glu Glu Asn Ala
545 550 555 560

Ile Met Gly Gly Ala Gly Ser Gly Val Asn Glu Val Leu Met Ala His
565 570 575

Arg Lys Pro Val Pro Val Leu Asn Ile Gly Leu Pro Asp Phe Phe Ile
580 585 590

Pro Gln Gly Thr Gln Glu Glu Met Arg Ala Glu Leu Gly Leu Asp Ala
595 600 605

Ala Gly Met Glu Ala Lys Ile Lys Ala Trp Leu Ala
610 615 620

<210> 12

<211> 305

<212> PRT

<213> Haematococcus pluvialis

<400> 12

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

Ser Ala Gln Gln Pro Ser Cys Ala His Ala Arg Leu Gln Phe Lys Leu
20 25 30

Arg Ser Met Gln Met Thr Leu Met Gln Pro Ser Ile Ser Ala Asn Leu
35 40 45

Ser Arg Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser Thr Trp
50 55 60

Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu Cys Ile
65 70 75 80

Leu Val Asp Val Glu Asp Asn Ile Thr Gly His Ala Ser Lys Leu Glu
85 90 95

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

100 105 110
 Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Leu Gln Gln
 115 120 125
 Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn Thr Cys

 130 135 140
 Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp Gln Leu
 145 150 155 160
 Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala Ala Ile
 165 170 175
 Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu Pro Ala
 180 185 190
 Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala Asp Val
 195 200 205

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

 275 280 285
 Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile Asn Glu
 290 295 300
 Ala
 305
 <210> 13
 <211> 275
 <212> PRT
 <213> Uncultured marine bacterium 66A03
 <400> 13

Met Gly Leu Met Leu Ile Asp Trp Cys Ala Leu Ala Leu Val Val Phe
1 5 10 15
Ile Gly Leu Pro His Gly Ala Leu Asp Ala Ala Ile Ser Phe Ser Met
20 25 30
Ile Ser Ser Ala Lys Arg Ile Ala Arg Leu Ala Gly Ile Leu Leu Ile
35 40 45
Tyr Leu Leu Leu Ala Thr Ala Phe Phe Leu Ile Trp Tyr Gln Leu Pro
50 55 60
Ala Phe Ser Leu Leu Ile Phe Leu Leu Ile Ser Ile Ile His Phe Gly
65 70 75 80
Met Ala Asp Phe Asn Ala Ser Pro Ser Lys Leu Lys Trp Pro His Ile
85 90 95
Ile Ala His Gly Gly Val Val Thr Val Trp Leu Pro Leu Ile Gln Lys
100 105 110
Asn Glu Val Thr Lys Leu Phe Ser Ile Leu Thr Asn Gly Pro Thr Pro
115 120 125
Ile Leu Trp Asp Ile Leu Leu Ile Phe Phe Leu Cys Trp Ser Ile Gly
130 135 140
Val Cys Leu His Thr Tyr Glu Thr Leu Arg Ser Lys His Tyr Asn Ile
145 150 155 160
Ala Phe Glu Leu Ile Gly Leu Ile Phe Leu Ala Trp Tyr Ala Pro Pro
165 170 175
Leu Val Thr Phe Ala Thr Tyr Phe Cys Phe Ile His Ser Arg Arg His
180 185 190
Phe Ser Phe Val Trp Lys Gln Leu Gln His Met Ser Ser Lys Lys Met
195 200 205
Met Ile Gly Ser Ala Ile Ile Leu Ser Cys Thr Ser Trp Leu Ile Gly
210 215 220
Gly Gly Ile Tyr Phe Phe Leu Asn Ser Lys Met Ile Ala Ser Glu Ala
225 230 235 240
Ala Leu Gln Thr Val Phe Ile Gly Leu Ala Ala Leu Thr Val Pro His

245 250 255
 Met Ile Leu Ile Asp Phe Ile Phe Arg Pro His Ser Ser Arg Ile Lys
 260 265 270
 Ile Lys Asn
 275
 <210> 14
 <211> 566
 <212> PRT
 <213> Mus musculus
 <400> 14
 Met Glu Ile Ile Phe Gly Gln Asn Lys Lys Glu Gln Leu Glu Pro Val
 1 5 10 15
 Gln Ala Lys Val Thr Gly Ser Ile Pro Ala Trp Leu Gln Gly Thr Leu
 20 25 30

 Leu Arg Asn Gly Pro Gly Met His Thr Val Gly Glu Ser Lys Tyr Asn
 35 40 45
 His Trp Phe Asp Gly Leu Ala Leu Leu His Ser Phe Ser Ile Arg Asp
 50 55 60
 Gly Glu Val Phe Tyr Arg Ser Lys Tyr Leu Gln Ser Asp Thr Tyr Ile
 65 70 75 80
 Ala Asn Ile Glu Ala Asn Arg Ile Val Val Ser Glu Phe Gly Thr Met
 85 90 95
 Ala Tyr Pro Asp Pro Cys Lys Asn Ile Phe Ser Lys Ala Phe Ser Tyr

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

Glu Ala Gly Asn Val Leu Asn Met Gly Thr Ser Val Val Asp Lys Gly
 180 185 190
 Arg Thr Lys Tyr Val Ile Phe Lys Ile Pro Ala Thr Val Pro Asp Ser
 195 200 205
 Lys Lys Lys Gly Lys Ser Pro Val Lys His Ala Glu Val Phe Cys Ser
 210 215 220
 Ile Ser Ser Arg Ser Leu Leu Ser Pro Ser Tyr Tyr His Ser Phe Gly
 225 230 235 240
 Val Thr Glu Asn Tyr Val Val Phe Leu Glu Gln Pro Phe Lys Leu Asp
 245 250 255
 Ile Leu Lys Met Ala Thr Ala Tyr Met Arg Gly Val Ser Trp Ala Ser
 260 265 270
 Cys Met Ser Phe Asp Arg Glu Asp Lys Thr Tyr Ile His Ile Ile Asp
 275 280 285
 Gln Arg Thr Arg Lys Pro Val Pro Thr Lys Phe Tyr Thr Asp Pro Met
 290 295 300
 Val Val Phe His His Val Asn Ala Tyr Glu Glu Asp Gly Cys Val Leu
 305 310 315 320
 Phe Asp Val Ile Ala Tyr Glu Asp Ser Ser Leu Tyr Gln Leu Phe Tyr
 325 330 335
 Leu Ala Asn Leu Asn Lys Asp Phe Glu Glu Lys Ser Arg Leu Thr Ser
 340 345 350
 Val Pro Thr Leu Arg Arg Phe Ala Val Pro Leu His Val Asp Lys Asp
 355 360 365
 Ala Glu Val Gly Ser Asn Leu Val Lys Val Ser Ser Thr Thr Ala Thr
 370 375 380
 Ala Leu Lys Glu Lys Asp Gly His Val Tyr Cys Gln Pro Glu Val Leu
 385 390 395 400
 Tyr Glu Gly Leu Glu Leu Pro Arg Ile Asn Tyr Ala Tyr Asn Gly Lys
 405 410 415
 Pro Tyr Arg Tyr Ile Phe Ala Ala Glu Val Gln Trp Ser Pro Val Pro
 420 425 430

Thr Lys Ile Leu Lys Tyr Asp Ile Leu Thr Lys Ser Ser Leu Lys Trp
435 440 445
Ser Glu Glu Ser Cys Trp Pro Ala Glu Pro Leu Phe Val Pro Thr Pro
450 455 460
Gly Ala Lys Asp Glu Asp Asp Gly Val Ile Leu Ser Ala Ile Val Ser
465 470 475 480
Thr Asp Pro Gln Lys Leu Pro Phe Leu Leu Ile Leu Asp Ala Lys Ser
485 490 495
Phe Thr Glu Leu Ala Arg Ala Ser Val Asp Ala Asp Met His Leu Asp
500 505 510
Leu His Gly Leu Phe Ile Pro Asp Ala Asp Trp Asn Ala Val Lys Gln
515 520 525
Thr Pro Ala Glu Thr Gln Glu Val Glu Asn Ser Asp His Pro Thr Asp
530 535 540
Pro Thr Ala Pro Glu Leu Ser His Ser Glu Asn Asp Phe Thr Ala Gly
545 550 555 560
His Gly Gly Ser Ser Leu
565
<210> 15
<211> 350
<212> PRT
<213> Natronomonas pharaonis ATCC35678
<400> 15
Met Ser Asn Ala Ser Leu Arg Pro Ser Gly Thr Ala Ser Ala Thr Leu
1 5 10 15
Phe Arg Leu Ala Phe Leu Pro Gly Trp Ala Val Ile Ala Ala Thr Thr
20 25 30
Gly Ala Phe Leu Val Gly Ala Ser Leu Pro Leu Thr Tyr Gln Leu Ile
35 40 45
Pro Leu Ala Ala Ser Val Val Leu Leu Gly Leu Pro His Gly Ala Val
50 55 60
Asp His Leu Ala Leu Pro Arg Thr Arg Asn Glu Arg Val Thr Val Arg

65					70				75				80			
Trp	Leu	Ala	Ala	Ile	Gly	Val	Leu	Tyr	Ala	Val	Val	Gly	Gly	Leu	Tyr	
				85				90				95				
Ala	Ala	Val	Trp	Phe	Leu	Ala	Pro	Val	Gly	Ala	Val	Ala	Ala	Phe	Ile	
				100				105				110				
Phe	Met	Thr	Trp	Val	His	Trp	Gly	Gln	Gly	Glu	Ile	Tyr	Pro	Leu	Val	
				115				120				125				
Ala	Leu	Ala	Asp	Ala	Asp	His	Leu	Asp	Gly	Arg	Leu	Glu	Arg	Gly	Leu	
				130				135				140				
Thr	Ala	Ala	Ile	Arg	Gly	Ala	Leu	Pro	Met	Leu	Val	Pro	Phe	Val	Ala	
				145				150				155				
												160				
Phe	Pro	Asp	Gln	Tyr	Glu	Leu	Val	Val	Thr	Thr	Leu	Val	Gly	Leu	Phe	
				165				170				175				
Asp	Ala	Asp	Ala	Ala	Ala	Thr	Ala	Ala	Ala	Ala	Phe	Thr	Pro	Thr	Ala	
				180				185				190				
Arg	Leu	Ala	Val	Ala	Val	Thr	Val	Gly	Gly	Leu	Val	Ala	Val	Thr	Leu	
				195				200				205				
Gly	Ile	Gly	Ala	Val	Ala	Ala	Ser	Glu	Thr	Gly	Trp	Gly	Pro	Trp	Leu	
				210				215				220				
Leu	Asp	Ala	Gly	Glu	Thr	Gly	Leu	Leu	Ile	Leu	Phe	Phe	Ala	Ala	Val	
				225				230				235				
												240				
Pro	Pro	Ile	Phe	Ala	Ile	Gly	Leu	Tyr	Phe	Cys	Phe	Trp	His	Ser	Leu	
				245				250				255				
Arg	His	Ile	Val	Arg	Leu	Leu	Ala	Val	Asp	Asn	Arg	Ala	Ala	Pro	Ala	
				260				265				270				
Leu	Asp	Gly	Arg	Arg	Tyr	Gly	Ala	Ala	Leu	Ala	Arg	Phe	Ala	Arg	Asp	
				275				280				285				
Ala	Ala	Pro	Leu	Ser	Ala	Ala	Ser	Leu	Val	Leu	Leu	Gly	Leu	Leu	Tyr	
				290				295				300				
Leu	Ala	Val	Pro	Gly	Ser	Val	Asp	Ser	Pro	Leu	Ser	Leu	Val	Gly	Thr	
				305				310				315				
												320				

Tyr Leu Val Leu Ile Ala Val Leu Thr Leu Pro His Val Val Val Val

325 330 335

Ala Trp Met Asp His Glu Gln Arg Leu Trp Arg Pro Gly Ala

340 345 350

<210> 16

<211> 284

<212> PRT

<213> Halobacterium salinarum ATCC700922

<400> 16

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

1 5 10 15

Gly Thr Val Thr Val Arg Trp Leu Ala Val Val Gly Val Leu Tyr Leu

20 25 30

Val Leu Gly Gly Gly Tyr Ala Ala Ala Trp Phe Phe Ala Pro Val Pro

35 40 45

Ala Ala Phe Ala Phe Val Ala Ile Thr Trp Leu His Trp Gly Gln Gly

50 55 60

Asp Leu Tyr Pro Leu Leu Asp Phe Leu Asp Val Asp Tyr Leu Asp Thr

65 70 75 80

Arg Pro Arg Arg Ala Ala Thr Val Leu Ile Arg Gly Gly Leu Pro Met

85 90 95

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

100 105 110

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

115 120 125

Asp Pro Arg Val Arg Leu Trp Leu Gly Val Ala Phe Ala Ala Ala Thr

130 135 140

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

145 150 155 160

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

165 170 175

Val Val Pro Pro Val Phe Ala Val Gly Val Tyr Phe Cys Val Trp His

180 185 190
 Ser Val Arg His Val Ala Arg Ala Ile Ala Val Asp Gly Ser Val His
 195 200 205
 Pro Ser Leu Arg Ala Gly Asp Ile Leu Gly Pro Leu Ala Arg Phe Gly
 210 215 220
 Val Glu Ala Ala Pro Met Thr Ala Ala Ala Leu Ala Leu Gly Gly Val

 225 230 235 240
 Leu Trp Trp Ala Val Pro Asn Pro Pro Thr Thr Leu Glu Ser Gly Ala
 245 250 255
 Ala Leu Tyr Leu Val Leu Ile Ala Val Leu Thr Leu Pro His Val Ala
 260 265 270
 Val Val Thr Trp Met Asp Arg Val Gln Gly Val Leu
 275 280
 <210> 17
 <211> 284
 <212> PRT
 <213> Halobacterium salinarum ATCC700922
 <400>
 17
 Met Pro His Gly Ala Ile Asp Tyr Leu Ala Leu Pro Arg Ala Val Thr
 1 5 10 15
 Gly Thr Val Thr Val Arg Trp Leu Ala Val Val Gly Val Leu Tyr Leu
 20 25 30
 Val Leu Gly Gly Gly Tyr Ala Ala Ala Trp Phe Phe Ala Pro Val Pro
 35 40 45
 Ala Ala Phe Ala Phe Val Ala Ile Thr Trp Leu His Trp Gly Gln Gly
 50 55 60
 Asp Leu Tyr Pro Leu Leu Asp Phe Leu Asp Val Asp Tyr Leu Asp Thr

 65 70 75 80
 Arg Pro Arg Arg Ala Ala Thr Val Leu Ile Arg Gly Gly Leu Pro Met
 85 90 95
 Leu Val Pro Leu Leu Gly Phe Pro Glu Arg Tyr Arg Ser Val Val Asp
 100 105 110

Ala Phe Ala Ala Pro Phe Gly Gly Ser Val Gly Asp Leu Ala Val Phe
115 120 125

Asp Pro Arg Val Arg Leu Trp Leu Gly Val Ala Phe Ala Ala Ala Thr
130 135 140

Val Ala Val Leu Ala Ala Gly Arg Arg Arg Thr His Ser Pro Gly Ala
145 150 155 160

Trp Arg Val Asp Ala Ala Glu Thr Leu Leu Leu Trp Val Phe Phe Phe
165 170 175

Val Val Pro Pro Val Phe Ala Val Gly Val Tyr Phe Cys Val Trp His
180 185 190

Ser Val Arg His Val Ala Arg Ala Ile Ala Val Asp Gly Ser Val His
195 200 205

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

210 215 220

Val Glu Ala Ala Pro Met Thr Ala Ala Ala Leu Ala Leu Gly Gly Val
225 230 235 240

Leu Trp Trp Ala Val Pro Asn Pro Pro Thr Thr Leu Glu Ser Gly Ala
245 250 255

Ala Leu Tyr Leu Val Leu Ile Ala Val Leu Thr Leu Pro His Val Ala
260 265 270

Val Val Thr Trp Met Asp Arg Val Gln Gly Val Leu
275 280

<210> 18

<211> 2412

<212> DNA

<213> Enterococcus faecalis

<400> 18

ttgaaaacag tagttattat tgatgcatta cgaacaccaa ttggaaaata taaaggcagc 60

ttaagtcaag taagtgccgt agacttagga acacatgtta caacacaact tttaaaaaga 120

cattccacta tttctgaaga aattgatcaa gtaatctttg gaaatgtttt acaagctgga 180

aatggccaaa atccgcgacg acaaatagca ataaacagcg gtttatctca tgaaattccc 240

gcaatgacag ttaatgaggt ctgcggatca ggaatgaagg ccgttatattt ggcgaacaaa 300

ttgattcaat taggagaagc ggaagtttta attgctggcg ggattgagaa tatgtcccaa	360
gcacctaaat tacaacgatt taattacgaa acagaaagct atgatgcgcc tttttctagt	420
atgatgtacg atgggttaac ggatgccttt agtgggtcaag caatgggctt aactgctgaa	480
aatgtggcgcg aaaagtatca tgtaactaga gaagagcaag atcaattttc tgtacattca	540
caattaaaag cagctcaagc acaagcagaa gggatatctg ctgacgaaat agccccatta	600
gaagtatcag gaacgcttgt ggagaaagat gaagggattc gccctaattc gagcgttgag	660
aagctaggaa cgcttaaac agtttttaaa gaagacggta ctgtaacagc agggaatgca	720
tcaaccatta atgatggggc ttctgctttg attattgctt cacaagaata tgccgaagca	780
cacggctctt cttattttagc tattattcga gacagtgtgg aagtcggtat tgatccagcc	840
tatatgggaa tttcgccgat taaagccatt caaaaactgt tagcgcgcaa tcaacttact	900
acggaagaaa ttgatctgta tgaaatcaac gaagcatttg cagcaacttc aatcgtggtc	960
caaagagaac tggctttacc agaggaaaag gtcaacatit atgggtggcgg tatttcatta	1020
ggtcatgcga ttgggtccac aggtgctcgt ttattaacga gtttaagtta tcaattaaat	1080
caaaaagaaa agaaatatgg agtggcttct ttatgtatcg gcggtggctt aggactcgt	1140
atgctactag agagacctca gcaaaaaaaaa aacagccgat tttatcaaat gagtcttgag	1200
gaacgcctgg cttctcttct taatgaaggc cagatttctg ctgatacaaa aaaagaatit	1260
gaaaatacgg ctttatcttc gcagattgcc aatcatatga ttgaaaatca aatcagtga	1320
acagaagtgc cgatgggcgt tggcttacat ttaacagtgg acgaaactga ttatttggt	1380
ccaatggcga cagaagagcc ctcatgtatt gcggctttga gtaatggtgc aaaaatagca	1440
caaggattta aaacagtga tcaacaacgt ttaatgcgtg gacaaatcgt tttttacgat	1500
gttgacagc cagatcatt gattgatgaa ctacaagtaa gagaacgga aatttttcaa	1560
caagcagagt taagtatcc atctatcgtt aaacgcggcg gcggttaag agatttgcaa	1620
tatcgtgctt ttgatgaatc atttgtatct gtcgactttt tagtagatgt taaggatgca	1680
atggggggcaa atatcgttaa cgctatgttg gaaggtgtgg ccgagttgtt ccgtgaatgg	1740
tttgcggagc aaaagatttt attcagtatt ttaagtaatt atgccacgga gtcggttgtt	1800
acgatgaaaa cgctatttcc agtttcacgt ttaagtaagg ggagcaatgg ccgggaaatt	1860
gctgaaaaaa ttgtttttagc ttacgctat gcttcattag atccttatcg ggcagtcacg	1920
cataacaaag ggatcatgaa tggcattgaa gctgtcgttt tagctacagg aatgataca	1980
cgcgctgtta gcgcttcttg tcatgctttt gcggtgaagg aaggctgcta ccaaggtttg	2040

actagtgtga cgctggatgg cgaacaacta attggtgaaa tttcagttcc gcttgcgtta 2100
gccacggttg gcggtgccac aaaagtctta cctaaatctc aagcagctgc tgatttgta 2160
gcagtgcagg atgcaaaaga actaagtcga gtagtagcgg ctgttggttt ggcacaaaat 2220
ttagcggcgt tacgggcctt agtctctgaa ggaattcaaa aaggacacat ggctctacaa 2280
gcacgttctt tagcgatgac ggtcggagct actggtaaag aagttgaggc agtcgctcaa 2340
caattaaaac gtcaaaaaac gatgaaccaa gaccgagcct tggtatattt aaatgattta 2400
agaaaacaat aa 2412

<210> 19

<211> 467

<212> PRT

<213> E coli DH5a(pTDHBSR/pSNA)

<400> 19

Met Asn Gln Gln Asp Ile Glu Gln Val Val Lys Ala Val Leu Leu Lys

1 5 10 15

Met Gln Ser Ser Asp Thr Pro Ser Ala Ala Val His Glu Met Gly Val

20 25 30

Phe Ala Ser Leu Asp Asp Ala Val Ala Ala Ala Lys Val Ala Gln Gln

35 40 45

Gly Leu Lys Ser Val Ala Met Arg Gln Leu Ala Ile Ala Ala Ile Arg

50 55 60

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

65 70 75 80

Glu Thr Gly Met Gly Arg Val Glu Asp Lys Phe Ala Lys Asn Val Ala

85 90 95

Gln Ala Arg Gly Thr Pro Gly Val Glu Cys Leu Ser Pro Gln Val Leu

100 105 110

Thr Gly Asp Asn Gly Leu Thr Leu Ile Glu Asn Ala Pro Trp Gly Val

115 120 125

Val Ala Ser Val Thr Pro Ser Thr Asn Pro Ala Ala Thr Val Ile Asn

130 135 140

Asn Ala Ile Ser Leu Ile Ala Ala Gly Asn Ser Val Ile Phe Ala Pro

145 150 155 160

His Pro Ala Ala Lys Lys Val Ser Gln Arg Ala Ile Thr Leu Leu Asn
 165 170 175
 Gln Ala Ile Val Ala Ala Gly Gly Pro Glu Asn Leu Leu Val Thr Val
 180 185 190
 Ala Asn Pro Asp Ile Glu Thr Ala Gln Arg Leu Phe Lys Phe Pro Gly
 195 200 205
 Ile Gly Leu Leu Val Val Thr Gly Gly Glu Ala Val Val Glu Ala Ala
 210 215 220
 Arg Lys His Thr Asn Lys Arg Leu Ile Ala Ala Gly Ala Gly Asn Pro
 225 230 235 240
 Pro Val Val Val Asp Glu Thr Ala Asp Leu Ala Arg Ala Ala Gln Ser
 245 250 255
 Ile Val Lys Gly Ala Ser Phe Asp Asn Asn Ile Ile Cys Ala Asp Glu
 260 265 270
 Lys Val Leu Ile Val Val Asp Ser Val Ala Asp Glu Leu Met Arg Leu
 275 280 285
 Met Glu Gly Gln His Ala Val Lys Leu Thr Ala Glu Gln Ala Gln Gln
 290 295 300
 Leu Gln Pro Val Leu Leu Lys Asn Ile Asp Glu Arg Gly Lys Gly Thr
 305 310 315 320
 Val Ser Arg Asp Trp Val Gly Arg Asp Ala Gly Lys Ile Ala Ala Ala
 325 330 335
 Ile Gly Leu Lys Val Pro Gln Glu Thr Arg Leu Leu Phe Val Glu Thr
 340 345 350
 Thr Ala Glu His Pro Phe Ala Val Thr Glu Leu Met Met Pro Val Leu
 355 360 365
 Pro Val Val Arg Val Ala Asn Val Ala Asp Ala Ile Ala Leu Ala Val
 370 375 380
 Lys Leu Glu Gly Gly Cys His His Thr Ala Ala Met His Ser Arg Asn
 385 390 395 400
 Ile Glu Asn Met Asn Gln Met Ala Asn Ala Ile Asp Thr Ser Ile Phe

405 410 415

Val Lys Asn Gly Pro Cys Ile Ala Gly Leu Gly Leu Gly Gly Glu Gly

420 425 430

Trp Thr Thr Met Thr Ile Thr Thr Pro Thr Gly Glu Gly Val Thr Ser

435 440 445

Ala Arg Thr Phe Val Arg Leu Arg Arg Cys Val Leu Val Asp Ala Phe

450 455 460

Arg Ile Val

465

<210> 20

<211> 495

<212> PRT

<213> E coli DH5a(pTDHBSR/pSNA)

<400> 20

Met Asn Phe His His Leu Ala Tyr Trp Gln Asp Lys Ala Leu Ser Leu

1 5 10 15

Ala Ile Glu Asn Arg Leu Phe Ile Asn Gly Glu Tyr Thr Ala Ala Ala

20 25 30

Glu Asn Glu Thr Phe Glu Thr Val Asp Pro Val Thr Gln Ala Pro Leu

35 40 45

Ala Lys Ile Ala Arg Gly Lys Ser Val Asp Ile Asp Arg Ala Met Ser

50 55 60

Ala Ala Arg Gly Val Phe Glu Arg Gly Asp Trp Ser Leu Ser Ser Pro

65 70 75 80

Ala Lys Arg Lys Ala Val Leu Asn Lys Leu Ala Asp Leu Met Glu Ala

85 90 95

His Ala Glu Glu Leu Ala Leu Leu Glu Thr Leu Asp Thr Gly Lys Pro

100 105 110

Ile Arg His Ser Leu Arg Asp Asp Ile Pro Gly Ala Ala Arg Ala Ile

115 120 125

Arg Trp Tyr Ala Glu Ala Ile Asp Lys Val Tyr Gly Glu Val Ala Thr

130 135 140

Thr Ser Ser His Glu Leu Ala Met Ile Val Arg Glu Pro Val Gly Val

145 150 155 160

Ile Ala Ala Ile Val Pro Trp Asn Phe Pro Leu Leu Leu Thr Cys Trp

165 170 175

Lys Leu Gly Pro Ala Leu Ala Ala Gly Asn Ser Val Ile Leu Lys Pro

180 185 190

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

195 200 205

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

210 215 220

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

225 230 235 240

Ala Phe Thr Gly Ser Thr Arg Thr Gly Lys Gln Leu Leu Lys Asp Ala

245 250 255

Gly Asp Ser Asn Met Lys Arg Val Trp Leu Glu Ala Gly Gly Lys Ser

260 265 270

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

275 280 285

Ala Thr Ala Ala Gly Ile Phe Tyr Asn Gln Gly Gln Val Cys Ile Ala

290 295 300

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

305 310 315 320

Leu Leu Lys Gln Gln Ala Gln Asn Trp Gln Pro Gly His Pro Leu Asp

325 330 335

Pro Ala Thr Thr Met Gly Thr Leu Ile Asp Cys Ala His Ala Asp Ser

340 345 350

Val His Ser Phe Ile Arg Glu Gly Glu Ser Lys Gly Gln Leu Leu Leu

355 360 365

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

370 375 380

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

385					390					395					400
Pro	Val	Leu	Val	Val	Thr	Arg	Phe	Thr	Ser	Glu	Glu	Gln	Ala	Leu	Gln
					405					410					415
Leu	Ala	Asn	Asp	Ser	Gln	Tyr	Gly	Leu	Gly	Ala	Ala	Val	Trp	Thr	Arg
					420					425					430
Asp	Leu	Ser	Arg	Ala	His	Arg	Met	Ser	Arg	Arg	Leu	Lys	Ala	Gly	Ser

435					440					445									
Val	Phe	Val	Asn	Asn	Tyr	Asn	Asp	Gly	Asp	Met	Thr	Val	Pro	Phe	Gly				
450					455					460									
Gly	Tyr	Lys	Gln	Ser	Gly	Asn	Gly	Arg	Asp	Lys	Ser	Leu	His	Ala	Leu				
465					470					475					480				
Glu	Lys	Phe	Thr	Glu	Leu	Lys	Thr	Ile	Trp	Ile	Ser	Leu	Glu	Ala					
485					490					495									

<210>	21
<211>	1404
<212>	DNA
<213>	E coli DH5a(pTDHBSR/pSNA)
<400>	

21	
atgaatcaac aggatattga acaggtggtg aaagcggtag tgctgaaaat gcaaagcagt	60
gacacgccgt ccgccgccgt tcatgagatg ggcgttttcg cgtccctgga tgacgccgtt	120
gcggcagcca aagtcgcccc gcaagggtta aaaagcgtgg caatgcgcca gttagccatt	180
gctgccattc gtgaagcagg cgaaaaacac gccagagatt tagcggaact tgccgtcagt	240
gaaaccggca tggggcgcggt tgaagataaa ttgtcaaaaa acgtcgctca ggcgcgcggc	300
acaccaggcg ttgagtgcct ctctccgcaa gtgctgactg gcgacaacgg cctgacccta	360
attgaaaacg caccctgggg cgtggtggct tcggtgacgc ctccactaa cccggcggca	420
accgtaatta acaacgccat cagcctgatt gcccggggca acagcgtcat ttttgccccg	480
catccggcgg cgaaaaaagt ctcccagcgg gcgattacgc tgctcaacca ggcgattgtt	540
gccgcaggtg ggccggaaaa ct tactggtt actgtggcaa atccggatat cgaaaccgcg	600
caacgcttgt tcaagtttcc gggatatcggc ctgctggtgg taaccggcgg cgaagcggta	660
gtagaagcgg cgcgtaaaca caccaataaa cgtctgattg ccgcaggcgc tggcaaccgc	720
ccggtagttg ttgatgaaac cgccgacctc gcccggtgcc ctcagttccat cgtcaaaggc	780

gcttcttttcg ataacaacat catttgtgcc gacgaaaagg tactgattgt tgttgatagc 840

gtagccgatg aactgatgcg tctgatggaa ggccagcacg cggtgaaact gaccgcagaa 900

caggcgacgc agctgcaacc ggtgttgcgtg aaaaatatcg acgagcgcg aaaggcacc 960

gtcagccgtg actgggttgg tgcgcacgca ggcaaaatcg cggcggcaat cggccttaaa 1020

gttccgcaag aaacgcgcct gctgtttgtg gaaaccaccg cagaacatcc gtttgccgtg 1080

actgaactga tgatgccgtt gttgcccgtc gtgcgcgtcg ccaacgtggc ggatgccatt 1140

gcgctagcgg tgaactgga aggcggttgc caccacacgg cggcaatgca ctgcgcgaac 1200

atcgaaaaca tgaaccagat ggcgaaatgct attgatacca gcattttcgt taagaacgga 1260

ccgtgcattg cggggttggg gctgggcggg gaaggctgga ccacatgac catcaccacg 1320

ccaaccggtg aaggggtaac cagcgcgcgt acgtttgtcc gtctgcgtcg ctgtgtatta 1380

gtcgatgcgt ttcgcattgt ttaa 1404

<210> 22

<211> 1488

<212> DNA

<213> E coli DH5a(pTDHBSR/pSNA)

<400> 22

atgaattttc atcatctggc ttactggcag gataaagcgt taagtctcgc cattgaaaac 60

cgcttattta ttaacggtga atatactgct gcggcggaag atgaacatt tgaaaccgtt 120

gatccggtca cccaggcacc gctggcgaaa attgcccgcg gcaagagcgt cgatatcgac 180

cgctgcgatga gcgcagcacg cggcgatattt gaacgcggcg actggtcact ctcttctccg 240

gctaaacgta aagcgggtact gaataaactc gccgatttaa tggaagccca cgccgaagag 300

ctggcactgc tggaactct cgacaccggc aaaccgattc gtcacagtct gcgtgatgat 360

attcccggcg cggcgcgcgc cattcgctgg tacgccgaag cgatcgaaa agtgatggc 420

gaagtggcga ccaccagtag ccatgagctg gcgatgatcg tgcgtgaacc ggtcggcgtg 480

attgccgcca tcgtgccgtg gaacttcccg ctgttctga cttgctggaa actcggccccg 540

gcgctggcgg cgggaaacag cgtgattcta aaaccgtctg aaaaatcacc gtcagtgcg 600

attcgtctcg cggggttggc gaaagaagca ggcttgccgg atggtgtgtt gaacgtggtg 660

acgggttttg gtcatgaagc cgggcaggcg ctgtcgcgtc ataacgatat cgacgccatt 720

gcctttaccg gttcaaccg tactcgggaaa cagctgtga aagatgcggg cgacagcaac 780

atgaaacgcg tctggctgga agcgggcggc aaaagcgcca acatcgtttt cgctgactgc 840

ccggatttgc aacaggcggc aagcgccacc gcagcaggca tttctacaa ccaggacag 900

gtgtgcatcg ccggaacgcg cctgttgctg gaagagagca tcgccgatga attcttagcc 960
ctgttaaaac agcaggcgca aaactggcag cggggccatc cacttgatcc cgcaaccacc 1020

atgggcacct taatcgactg cgcccacgcc gactcggtec atagctttat tcgggaagge 1080
gaaagcaaag ggcaactgtt gttggatggc cgtaacgccg ggctggctgc cgccatcggc 1140
ccgaccatct ttgttgatgt ggacccgaat gcgtccttaa gtcgcgaaga gattttcggt 1200
ccggtgctgg tggtcacgcg ttccacatca gaagaacagg cgctacagct tgccaacgac 1260
agccagtacg gccttggcgc ggcggtatgg acgcgcgacc tctcccgcgc gcaccgcatg 1320
agccgacgcc tgaagccgg ttccgtcttc gtcaataact acaacgacgg cgatatgacc 1380
gtgccgtttg gcggctataa gcagagcggc aacggtcgcg acaaatccct gcatgccctt 1440

gaaaaattca ctgaactgaa aaccatctgg ataagcctgg aggcctga 1488

<210> 23

<211> 269

<212> PRT

<213> E coli DH5a(pTDHBSR/pSNA)

<400> 23

Met Thr His Lys Ala Thr Glu Ile Leu Thr Gly Lys Val Met Gln Lys

1 5 10 15

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

20 25 30

Leu Glu Leu Lys Arg Gln Gly Phe His Val Leu Ala Gly Cys Arg Lys

35 40 45

Pro Asp Asp Val Glu Arg Met Asn Ser Met Gly Phe Thr Gly Val Leu

50 55 60

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

65 70 75 80

Ile Ala Leu Thr Asp Asn Cys Leu Tyr Gly Ile Phe Asn Asn Ala Gly

85 90 95

Phe Gly Met Tyr Gly Pro Leu Ser Thr Ile Ser Arg Ala Gln Met Glu

100 105 110

Gln Gln Phe Ser Ala Asn Phe Phe Gly Ala His Gln Leu Thr Met Arg

115 120 125

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

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

260 265
 <210> 24
 <211> 810
 <212> DNA
 <213> E coli DH5a(pTDHBSR/pSNA)
 <400> 24
 atgactcata aagcaacgga gatcctgaca ggtaaagtta tgcaaaaatc ggtcttaatt 60
 accggatggt ccagtggaat tggcctggaa agcgcgctcg aattaaaacg ccagggtttt 120
 catgtgctgg caggttgccg gaaaccggat gatgttgagc gcatgaacag catgggattt 180
 accggcgtgt tgatcatct ggattcacca gaaagtgttg atcgcgacg cgacgaggtg 240
 atcgccctga ccgataattg tctgtatggg atctttaaca atgccggatt cggcatgtat 300
 ggcccccttt ccaccatcag ccgtgcgcag atggaacagc agttttccgc caactttttc 360
 ggcgcacacc agtcaccat gcgcctgtta cccgcgatgt taccgcacgg tgaagggcgt 420
 attgtgatga catcatcggg gatgggatta atctccacgc cgggtcgtgg cgcttacgcg 480
 gccagtaaat atgcgctgga ggcgtggtca gatgcactgc gcatggagct gcgccacagc 540

ggaattaaag tcagcctgat cgaacccggt cccattcgta ctgccttcac cgacaacgtc	600
aaccagacgc aaagtataa accagtcgaa aatcccgga tcgccgccg ctttacgttg	660
ggaccggaag cgttggtgga caaagtacgc catgctttta ttagcgagaa gccgaagatg	720
cgctatccgg tgacgctggt gacctgggcg gtaatggtgc ttaagcgct gctgccgggg	780
cgctgatgg acaaatatt gcaggggtga	810
<210> 25	
<211> 47	
<212> DNA	
<213> Artificial Sequence	
<220><223> YbbO-F primer	
<400> 25	
cggtaccga ggagtaata atatgactca taaagcaacg gagatcc	47
<210> 26	
<211> 33	
<212> DNA	
<213> Artificial Sequence	
<220><223> YbbO-R primer	
<400> 26	
cgctcgagtt acccctgcaa tattttgtcc atc	33
<210> 27	
<211> 5144	
<212> DNA	
<213> Artificial Sequence	
<220><223> pBBR1MCS2	
<400> 27	
accttcggga ggcctgaag cccgttctgg acgccctggg gccgttgaat cgggatatgc	60
aggccaagcg cgccgcgac atcaaggccg tggcgcaaaa gctgctgacg gaacagcggg	120
aagtccagcg ccagaaacag gccagcgcc agcaggaacg cggcgcgca catttccccg	180
aaaagtgcc cctgggatga atgtcagcta ctgggctatc tggacaaggg aaaacgcaag	240
cgcaaagaga aagcagtag cttgcagtgg gcttacatgg cgatagctag actgggcggt	300
tttatggaca gcaagcgaac cggaattgcc agctggggcg ccctctggtg aggttgggaa	360
gccctgcaaa gtaaacgga tggttttctt gccccaagg atctgatggc gcaggggatc	420

aagatctgat caagagacag gatgaggatc gtttcgcatg attgaacaag atggattgca	480
cgcaggttct ccggccgctt ggggtggagag gctattcggc tatgactggg cacaacagac	540
aatcgctgct tctgatgccg ccgtgttccg gctgtcagcg caggggccc cggttctttt	600
tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag gacgaggcag cgcggctatc	660
gtggctggcc acgacgggcg ttccttgccg agctgtgctc gacgttgtca ctgaagcggg	720
aagggaactgg ctgctattgg gcgaagtgcc ggggcaggat ctctgtcat ctacattgc	780
tcctgccgag aaagtatcca tcatggctga tgcaatgcgg cggctgcata cgcttgatcc	840
ggctacctgc ccattcgacc accaagcgaa acatcgcatc gagcgagcac gtactcggat	900
ggaagccggt ctgtcgatc aggatgatct ggacgaagag catcaggggc tcgcgccagc	960
cgaactgttc gccaggctca aggcgcgcat gcccgcggc gaggatctcg tcgtgacca	1020
tggcgatgcc tgcttgccga atatcatggt ggaaaatggc cgctttctg gattcatcga	1080
ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttggtta cccgtgatat	1140
tgctgaagag ctggcgccg aatgggctga ccgcttcctc gtgctttacg gtatcgccgc	1200
tcccgatctg cagcgcatcg cttctatcgc cttcttgac gagttctct gagcgggact	1260
ctggggttcg aaatgaccga ccaagcgacg cccaacctgc catcacgaga ttctgattcc	1320
accgccgct tctatgaaag gttgggcttc ggaatcggtt tccgggacgc cggctggatg	1380
atcctccagc gcggggatct catgctggag ttcttcgccc acccccatgg gcaaatatta	1440
tacgcaaggc gacaaggctc tgatgccgct ggcgattcag gttcatcatg ccgtttgtga	1500
tggcttccat gtcggcagaa tgcttaatga attacaacag ttttatgca tgcgccaat	1560
acgcaaaccg cctctccccg cgcgttggc gattcattaa tgcagctggc acgacaggtt	1620
tcccgaactg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc tcactcatta	1680
ggcacccag gctttacact ttatgcttcc ggtcgtatg ttgtgtggaa ttgtgagcgg	1740
ataacaattt cacacaggaa acagctatga ccatgattac gccaagcgcg caattaaccc	1800
tcactaaagg gaacaaaagc tgggtaccgg gccccccctc gaggtcgacg gtatcgataa	1860
gcttgatata gaattcctgc agcccggggg atccactagt tctagagcgg ccgccaccgc	1920
ggtggagctc caattcgccc tatagtgagt cgtattacgc gcgctcactg gccgtcgttt	1980
tacaacgtcg tgactgggaa aacctggcg ttaccaact taatgcctt gcagcacatc	2040
cccccttcgc cagctggcgt aatagcgaag agggccgcac cgatcgccct tcccaacagt	2100
tgccgagcct gaatggcgaa tggaaattgt aagcgtaat attttgttaa aattcgcgtt	2160
aaatttttgt taaatcagct ctttttttaa ccaataggcc gactgcgatg agtggcaggg	2220
cggggcgtaa tttttttaag gcagttattg gtgcccttaa acgcctggtg ctacgcctga	2280

ataagtata ataagcggat gaatggcaga aattcgaaag caaatcgcac ccggtcgtcg	2340
gttcaggga gggtcgttaa atagccgctt atgtctattg ctggtttacc ggtttattga	2400
ctaccggaag cagtgtgacc gtgtgcttct caaatgcctg aggccagttt gctcaggctc	2460
tccccgtgga ggtaataatt gacgatatga tcattttattc tgcctcccag agcctgataa	2520
aaacggtgaa tccgttagcg aggtgccgcc ggcttcatt caggctcagg tggcccggct	2580
ccatgcaccg cgacgcaacg cggggaggca gacaaggtat agggcggcga ggcggctaca	2640
gccgatagtc tggaacagcg cacttacggg ttgctgcga acccaagtgc taccggcgcg	2700
gcagcgtgac ccgtgtcgcc ggctccaacg gctcgccatc gtccagaaaa cacggctcat	2760
cgggcacgga caggcgctgc tgcccgcgc gttcccatc ctcggtttcg gtcaaggctg	2820
gcaggtctgg ttccatgccc ggaatgccgg gctggctggg cggctcctcg ccggggccgg	2880
tcggtagtgt ctgctcgccc ggatacaggg tcgggatgcg gcgcaggctc ccatgcccc	2940
acagcgattc gtcttggtcg tcgtgatcaa ccaccacggc ggactgaac accgacaggc	3000
gcaactggtc gcggggctgg cccacgcca cgcggctatt gaccacgtag gccgacacgg	3060
tgccggggcc gttgagcttc acgacggaga tccagcgctc ggccaccaag tccttgactg	3120
cgtattggac cgtccgcaa gaacgtccga tgagcttga aagtgtcttc tggtgacca	3180
ccacggcgtt ctggtggccc atctgcgcca cgaggtgatg cagcagcatt gccgccgtgg	3240
gtttcctcgc aataagcccc gccacgcct catgcgttt gcgttcggt tgcaaccagt	3300
gaccgggctt gttcttggt tgaatgccga tttctctgga ctgcgtggcc atgcttatct	3360
ccatgcggta gggtgccga cggttgcgc accatgcga atcagctga acttttcggc	3420
agcgcgacaa caattatcg ttgcgtaaaa gtggcagtca attacagatt ttctttaacc	3480
tacgcaatga gctattgcg ggggtgccgc aatgagctgt tgcgtacccc ctttttttaa	3540
gttgttgatt ttaagtctt tcgcatctc cctatatct agttcttttg tgcccaaaga	3600
agggcacccc tgccgggttc cccacgcct tcggcgcggc tccccctcg gcaaaaagtg	3660
gccccctcgg ggttgttga tcgactgcg gcccttcggc cttgccaag gtggcgctgc	3720
cccccttgaa cccccgact cgccgccgtg aggtcgggg ggacggcggg cgggcttcgc	3780
cttcgactgc cccactcgc atagcttgg gtggttcag gcgcgtcaag gccaaagcgc	3840
tgcgcggtcg ctgcgcgagc ctgacccgc ctccacttg gtgtccaacc ggcaagcgaa	3900
gcgcgcaggc cgcaggccgg aggtttttcc ccagagaaaa ttaaaaaaat tgatggggca	3960
aggccgcagg ccgcgcagtt ggagccggtg ggtatgtgt cgaaggctgg gtagccggtg	4020
ggcaatccct gtggtcaagc tcgtgggcag gcgcagcctg tccatcagct tgtccagcag	4080

ggttgtccac gggccgagcg aagcgagcca gccggtggcc gtcgcggcc atcgtccaca	4140
tatccacggg ctggcaaggg agcgcagcga ccgcgcaggg cgaagcccgg agagcaagcc	4200
cgtagggcgc cgcagccgcc gtagcgggtc acgactttgc gaagcaaagt ctagttagta	4260
tactcaagca ttgagtggcc cgccggaggc accgccttgc gctgcccccg tcgagccggt	4320
tggacaccaa aaggaggagg caggcatggc ggcatacgcg atcatgcgat gcaagaagct	4380
ggcgaaaatg ggcaacgtgg cgccagctct caagcacgcc taccgcgagc gcgagacgcc	4440
caacgtgac gccagcagga cgccagagaa cgagcactgg gcggccagca gcaccgatga	4500
agcgatgggc cgactgcgcg agttgctgcc agagaagcgg cgcaaggacg ctgtgttggc	4560
ggtcgagtac gtcatgacgg ccagcccgga atggtggaag tcggccagcc aagaacagca	4620
ggcggcgctt ttcgagaagg cgcacaagtg gctggcggac aagtacgggg cggatcgcat	4680
cgtgacggcc agcateccac gtgacgaaac cagcccgcac atgaccgctg tcgtggtgcc	4740
gctgacgcag gacggcaggc tgtcggccaa ggagttcatc ggcaacaaag cgcagatgac	4800
ccgcgaccag accacgtttg cgcccgctgt ggccgatcta gggctgcaac ggggcatcga	4860
gggcagcaag gcacgtcaca cgcgcttca ggcgtttctac gaggccctgg agcggccacc	4920
agtgggccac gtcacatca gcccgaagc ggtcgagcca cgcgcctatg caccgcaggg	4980
attggccgaa aagctgggaa tctcaaagcg cgttgagacg ccggaagccg tggccgaccg	5040
gctgacaaaa gcggttcggc aggggtatga gcctgcccta caggccgccg caggagcgcg	5100
tgagatgcgc aagaaggccg atcaagccca agagacggcc cgag	5144
<210> 28	
<211> 70	
<212> DNA	
<213> Artificial Sequence	
<220><223> KO Ybb0-F primer	
<400> 28	
gtttattgcc gactggatgg cgaagcagtt gcagccttta gtaaatcatg attccgggga	60
tccgtcgacc	70
<210> 29	
<211> 70	
<212> DNA	
<213> Artificial Sequence	
<220><223> KO Ybb0-R primer	

<400> 29
gacatggggg cttagcgcg cgcttcaact caccctgca atattttgtc ttaggctgg 60
agctgcttcg 70
<210> 30
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> KO YbbOCF-F
<400> 30
gcctgcaaac tatggtcgcc 20
<210> 31
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> KO YbbOCF-R
<400> 31
cagcaccgga gtggtcacg 20
<210> 32
<211> 70
<212> DNA
<213> Artificial Sequence
<220><223> KO EutE-F primer
<400> 32
gtggtgtctg gcggtcaggc aattttccac aaataaggca gaacatcatg attccgggga 60
tccgtcgacc 70
<210> 33
<211> 70
<212> DNA
<213> Artificial Sequence
<220><223>
> KO EutE-R primer
<400> 33
attgttcgtc gtgcgccatc tgttactcct taaacaatgc gaaacgcac ttaggctgg 60
agctgcttcg 70

<210>	34	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	KO EutECF-F	
<400>	34	
	ggtaatcctg acgggcaatg c	21
<210>	35	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	KO EutECF-R	
<400>	35	
	cgacaccaca tcgcaggtgc	20
<210>	36	
<211>	70	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	KO PuuC-F primer	
<400>	36	
	acgttcacaa actgcatata tctgatatagac gtgaaacagg agtcataatg atgggaatta	60
	gccatgggtcc	70
<210>	37	
<211>	70	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	KO PuuC-R primer	
<400>	37	
	cgtagtaact gctggatatgt tcggtcattt caggcctcca ggcttatcca tntaggtctgg	60
	agctgcttcg	70
<210>	38	
<211>	20	
<212>	DNA	

<213>	Artificial Sequence	
<220><223>	KO PuuCCF-F	
<400>	38	
cagaggcggc	ttgaaggatg	20
<210>	39	
<211>	24	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	KO PuuCCF-R	
<400>	39	
gtgattgact	cattcagcgt gtcg	24
<210>	40	
<211>	5959	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	pBBR2-ybb0	
<400>	40	
accttcggga	gcgccctgaag cccgttctgg acgccctggg gccgttgaat cgggatatgc	60
aggccaaggc	cgccgcgac atcaaggccg tgggcgaaaa gctgctgacg gaacagcggg	120
aagtccagcg	ccagaaacag gccagcgcc agcaggaacg cgggcgcgca catttccccg	180
aaaagtgcc	ctgggatga atgtcagcta ctgggctatc tggacaaggg aaaacgcaag	240
cgcaaagaga	aagcaggtag cttgcagtgg gcttacatgg cgatagctag actgggcggt	300
tttatggaca	gcaagcgaac cggaattgcc agctggggcg ccctctggta aggttgggaa	360
gccctgcaaa	gtaaactgga tggctttctt gccccaagg atctgatggc gcaggggatc	420
aagatctgat	caagagacag gatgaggatc gtttcgcatg attgaacaag atggattgca	480
cgcaggttct	ccggccgctt ggggtggagag gctattcggc tatgactggg cacaacagac	540
aatcggtgc	tctgatgccg ccgtgttcg gctgtcagcg caggggcgcc cggttctttt	600
tgtcaagacc	gacctgtccg gtgccctgaa tgaactgcag gacgaggcag cgcggctatc	660
gtggctggcc	acgacgggcg ttccttgccg agctgtgctc gacgttgtca ctgaagcggg	720
aagggaactg	ctgctattgg gcgaagtgcc ggggcaggat ctctgtcat ctacattgc	780
tcctgccgag	aaagtatcca tcatggctga tgcaatcgcg cggtgcata cgcttgatcc	840
ggctacctgc	ccattcgacc accaagcgaa acatcgcatc gagcgagcac gtactcggat	900

ggaagccggt ctgtcgatc aggatgatct ggacgaagag catcaggggc tcgcgccagc	960
cgaactgttc gccaggctca aggcgcgcat gcccgacggc gaggatctcg tcgtgaccca	1020
tggcgatgcc tgcttgccga atatcatggt ggaaaatggc cgcttttctg gattcatcga	1080
ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttggtta cccgtgatat	1140
tgctgaagag cttagcgagg aatgggctga ccgcttcctc gtgctttacg gtatcgccgc	1200
tcccgattcg cagcgcatcg ctttctatcg ctttcttgac gagttcttct gagcgggact	1260
ctggggttcg aaatgaccga ccaagcgacg cccaacctgc catcacgaga ttctgattcc	1320
accgccgcct tctatgaaag gttgggcttc ggaatcgttt tccgggacgc cggttgatg	1380
atcctccagc gcggggatct catgctggag ttcttcgccc acccccatgg gcaaataa	1440
tacgcaaggc gacaaggctc tgatgccgct ggcgattcag gttcatcatg ccgtttgtga	1500
tggcttccat gtcggcagaa tgcttaatga attacaacag tttttatgca tgcgcccaat	1560
acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc acgacaggtt	1620
tcccacttgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc tactcatta	1680
ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg	1740
ataacaattt cacacaggaa acagctatga ccatgattac gccaaagcgc caattaaccc	1800
tactaaagg gaacaaaagc tgggtaccga ggaggttaata atatgactca taaagcaacg	1860
gagatcctga caggtaaagt tatgcaaaaa tcggtcttaa ttaccggatg ttccagtga	1920
attggcctgg aaagcgcgtc cgaattaaaa cgcagggtt ttcatgtgct ggcaggttgc	1980
cggaaaccgg atgatgttga gcgcatgaac agcatgggat ttaccggcgt gttgatcgat	2040
ctggattcac cagaaagtgt tgatcgcgca gccgacgagg tgatcgccct gaccgataat	2100
tgtctgtatg ggatctttaa caatgccgga ttcggcatgt atggccccct ttccaccatc	2160
agccgtgccc agatggaaca gcagttttcc gccaaacttt tcggcgca caagctcacc	2220
atgcgcctgt taccgcgat gttaccgcac ggtgaagggc gtatttgtat gacatcatcg	2280
gtgatgggat taatctccac gccgggtcgt ggcgcttacg cggccagtaa atatgcgtg	2340
gaggcgtggt cagatgcact gcgcatggag ctgcgccaca gcggaattaa agtcagcctg	2400
atcgaaccgg gtccattcgt tactcgcttc accgacaacg tcaaccagac gcaaagtgat	2460
aaaccagtgc aaaaatccgg catcgccgcc cgctttacgt tgggaccgga agcgggtggtg	2520
gacaaagtac gccatgcttt tattagcgag aagccgaaga tgcgtatcc ggtgacgctg	2580
gtgacctggg cggtaatggt gcttaagcgc ctgctgccgg ggcgctgat ggacaaaata	2640
ttgcaggggt aactcgaggt cgacggatc gataagcttg atatcgaatt cctgcagccc	2700
gggggatcca ctagtcttag agcgccgcc acccggttg agtccaatt cgccctatag	2760

tgagtcgtat tacgcgcgct cactggccgt cgttttaca cgtcgtgact gggaaaaccc	2820
tggcggtacc caacttaate gccttcgagc acatccccct ttcgccagct ggcgtaatag	2880
cgaagaggcc cgcaccgac gcccttccca acagttgcgc agcctgaatg gcgaatggaa	2940
attgtaagcg ttaatatattt gttaaaattc gcgttaaatt tttgttaaatt cagctcattt	3000
tttaaccaat aggccgactg cgatgagtgg cagggcgggg cgtaattttt ttaaggcagt	3060
tattggtgcc cttaaacgcc tgggtctacg cctgaataag tgataataag cggatgaatg	3120
gcagaaattc gaaagcaaat tcgaccgggt cgtcgggtca gggcagggtc gttaaatagc	3180
cgcttatgtc tattgtctgt ttaccggttt attgactacc ggaagcagtg tgaccgtgtg	3240
cttctcaaat gcctgaggcc agtttctca ggctctcccc gtggaggtaa taattgacga	3300
tatgatcatt tattctgcct cccagagcct gataaaaacg gtgaatccgt tagcgagggtg	3360
ccgccggctt ccattcaggt cgaggtggcc cggtccatg caccgcgacg caacgcgggg	3420
aggcagacaa ggtatagggc ggcgaggcgg ctacagccga tagtctggaa cagcgcaatt	3480
acgggttgct gcgcaaccca agtctaccg gcgcggcagc gtgaccctg tcggcggtc	3540
caacggctcg ccattctca gaaaacacgg ctcatcggtc atcggcaggc gctgctgccc	3600
gcgcggttcc cattctccg tttcggtcaa ggctggcagg tctggttcca tgcccgaat	3660
gccgggtggt ctggcggtt cctcgccggg gccggtcggt agttgctgct cggccggata	3720
cagggtcggg atgcggcgca ggtcgccatg cccaacagc gattcgtcct ggtcgtcgtg	3780
atcaaccacc acggcggcac tgaacaccga caggcgcaac tggtcgagg gctggcccca	3840
cgccacgagg tcattgacca cgtaggccga cacggtgccg gggccgttga gcttcacgac	3900
ggagatccag cgctcgcca ccaagtctt gactgcgtat tggaccgtcc gcaaagaacg	3960
tccgatgagc ttggaagtgt tcttctggct gaccaccag gcgttctggt ggcccatctg	4020
cgccacgagg tgatgcagca gcattgccgc cgtgggtttc ctgcgaataa gcccggccca	4080
gcctcatgc gctttgcgtt ccgtttgcac ccagtaccg ggcttgttct tggttgaat	4140
gccgatttct ctggactgct tggccatgct tatctccatg cggtagggtg ccgcacggtt	4200
gcggcaccat gcgcaatcag ctgcaacttt tcggcagcgc gacaacaatt atgcgttgcg	4260
taaaagtggc agtcaattac agattttctt taacctacgc aatgagctat tgcggggggt	4320
gccgcaatga gctgttgct accccctttt ttaagtgt tgatttttaa gtctttcgca	4380
tttcgcccta tatctagttc ttgtgtgccc aaagaagggc acccctgcgg ggttcccca	4440
gccttcggc gcggctcccc ctccggcaaa aagtggcccc tccggggctt gttgatcgac	4500
tgcgcggcct tcggccttgc ccaaggtggc gctgccccct tggaaacccc gactcgccg	4560
ccgtgaggct cggggggcag gcggcggggc ttgccttcg actgccccca ctgcatagg	4620

cttgggtcgt tccaggcgcg tcaaggccaa gccgctgcgc ggtcgctgcg cgagccttga 4680

cccgccttcc acttgggtgc caaccggcaa gcgaagcgcg caggccgcag gccggaggct 4740

tttccccaga gaaaattaaa aaaattgatg gggcaaggcc gcaggccgcg cagttggagc 4800

cgggtgggtat gtggtcgaag gctgggtagc cggtagggcaa tccctgtggt caagctcgtg 4860

ggcaggcgca gccgtgccat cagcttgtcc agcagggttg tccacgggcc gagcgaagcg 4920

agccagccgg tggccgctcg cgcccatcgt ccacatatcc acgggctggc aaggagcgcg 4980

agcgaccgcg cagggcgaag cccggagagc aagcccgtag ggccgcgcag ccgccgtagg 5040

cggtcacgac ttgcgaagc aaagtctagt gagtatactc aagcattgag tggcccgccg 5100

gaggcaccgc ctgcgctgc cccgctcgag ccggttggac accaaaaggg aggggcaggc 5160

atggcggcat acgcgatcat gcgatgcaag aagctggcga aaatgggcaa cgtggcggcc 5220

agtctcaagc acgcctaccg cgagcgcgag acgccaacg ctgacgccag caggacgcca 5280

gagaacgagc actgggcggc cagcagcacc gatgaagcga tgggccgact gcgcgagttg 5340

ctgccagaga agcggcgcaa ggacgctgtg ttggcggtcg agtacgtcat gacggccagc 5400

ccggaatggt ggaagtgcgc cagccaagaa cagcaggcgg cgttcttcga gaaggcgac 5460

aagtggctgg cggacaagta cggggcggat cgcacgtga cggccagcat ccaccgtgac 5520

gaaaccagcc cgcacatgac cgcgttcgtg gtgccgtga cgcaggacgg caggctgtcg 5580

gccaaaggagt tcacggcaa caaagcgag atgaccgcg accagaccac gtttgcggcc 5640

gctgtggccg atctagggt gcaacggggc atcagaggca gcaaggcacg tcacacgcgc 5700

attcaggcgt tctacaggc cctggagcgg ccaccagtgg gccacgtcac catcagcccg 5760

caagcggtcg agccacgcgc ctatgcaccg cagggtattg ccgaaaagct gggaatctca 5820

aagcgcgttg agacgccga agccgtggcc gaccggctga caaaagcgt tcggcagggg 5880

tatgagcctg ccctacaggc cgccgcagga gcgcgtgaga tgcgcaagaa ggccgatcaa 5940

gccaagaga cgccccgag 5959

<210> 41

<211> 208

<212> PRT

<213> Shigella sonnei Ss046

<400> 41

Met Met Asn Phe Asn Asn Val Phe Arg Trp His Leu Pro Phe Leu Phe

1 5 10 15

Leu Ala Leu Leu Thr Phe Arg Ala Ala Ala Asp Thr Leu Leu Ile

20 25 30
 Leu Gly Asp Ser Leu Ser Ala Gly Tyr Arg Met Ser Ala Ser Ala Ala
 35 40 45

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

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

 Ala Asp Trp Met Ala Lys Gln Leu Gln Pro Leu Val Asn His Asp Ser
 195 200 205
 <210> 42
 <211> 256
 <212> PRT
 <213> Citrobacter rodentium ICC168
 <400> 42
 Met Gln Lys Ser Val Leu Ile Thr Gly Cys Ser Ser Gly Ile Gly Leu
 1 5 10 15
 Glu Ser Ala Leu Glu Leu Lys Arg Gln Gly Phe Gln Ile Leu Ala Gly
 20 25 30

Cys Arg Lys Pro Glu Asp Val Ala Arg Met Lys Ser Met Gly Phe Thr

35 40 45

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

50 55 60

Asp Glu Val Ile Ala Leu Thr Asp Asn Cys Leu Phe Gly Ile Phe Asn

65 70 75 80

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

85 90 95

Gln Met Glu Gln Gln Phe Ser Thr Asn Phe Phe Gly Ala His Gln Leu

100 105 110

Thr Met Arg Leu Leu Pro Ala Met Leu Pro His Gly Glu Gly Arg Ile

115 120 125

Val Met Thr Ser Ser Val Met Gly Leu Ile Ser Thr Pro Gly Arg Gly

130 135 140

Ala Tyr Ala Ala Ser Lys Tyr Ala Leu Glu Ala Trp Thr Asp Ala Leu

145 150 155 160

Arg Met Glu Leu Arg His Ser Gly Ile Lys Val Ser Leu Ile Glu Pro

165 170 175

Gly Pro Ile Arg Thr Arg Phe Thr Glu Asn Val Asn Gln Thr Gln Ser

180 185 190

Asp Lys Pro Val Glu Asn Pro Gly Ile Ala Ala Arg Phe Thr Leu Asp

195 200 205

Pro Glu Ala Val Val Asp Lys Val Arg His Ala Phe Val Ser Glu Lys

210 215 220

Pro Arg Leu Arg Tyr Pro Val Thr Leu Val Thr Trp Ala Val Met Leu

225 230 235 240

Leu Lys Arg Leu Leu Pro Gly Arg Leu Met Asp Lys Ile Leu His Gly

245 250 255

<210> 43

<211> 256

<212> PRT

<213> Salmonella enterica subsp. enterica serovar Typhimurium str. LT2

<400> 43

Met Gln Lys Ser Val Leu Ile Thr Gly Cys Ser Ser Gly Ile Gly Leu

1 5 10 15

Asp Ser Ala Leu Glu Leu Lys Arg Gln Gly Phe Gln Ile Leu Ala Gly

20 25 30

Cys Arg Lys Pro Asp Asp Val Ala Arg Met Asn Ser Met Gly Phe Thr

35 40 45

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

50 55 60

Asp Glu Val Ile Ala Leu Thr Asp Asn Arg Leu Tyr Gly Ile Phe Asn

65 70 75 80

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

85 90 95

Gln Met Glu Gln Gln Phe Ser Ser Asn Phe Phe Gly Ala His Gln Leu

100 105 110

Thr Met Arg Leu Leu Pro Ala Met Leu Pro His Gly Glu Gly Arg Ile

115 120 125

Val Met Thr Ser Ser Val Met Gly Leu Ile Ser Thr Pro Gly Arg Gly

130 135 140

Ala Tyr Ala Ala Ser Lys Tyr Ala Leu Glu Ala Trp Ser Asp Ala Leu

145 150 155 160

Arg Met Glu Leu Arg His Thr Gly Ile Lys Val Ser Leu Ile Glu Pro

165 170 175

Gly Pro Ile Arg Thr Arg Phe Thr Asp Asn Val Asn Gln Thr Gln Arg

180 185 190

Asp Lys Pro Val Glu Asn Pro Gly Ile Ala Ala Arg Phe Thr Arg Asp

195 200 205

Pro Asp Ala Val Val Ala Lys Val Arg His Ala Phe Ala Ser Asp Lys

210 215 220

Pro Arg Leu Arg Tyr Pro Val Thr Leu Val Thr Trp Ala Val Met Ala

225 230 235 240

Leu Lys Arg Ile Leu Pro Gly Arg Met Leu Asp Lys Ile Leu Gln Ser
245 250 255

<210> 44
<211> 218
<212> PRT
<213> Enterobacter cloacae EcWSU1

<400> 44

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

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

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

195 200 205

Arg Leu Ala Pro Leu Val Lys His Asp Ser

210 215

<210> 45

<211> 255

<212> PRT

<213> Enterobacter aerogenes KCTC 2190

<400> 45

Met Gln Lys Ser Val Leu Val Thr Gly Cys Ser Ser Gly Ile Gly Leu

1 5 10 15

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

20 25 30

Cys Arg Lys Gly Glu Asp Val Ala Arg Met Gln Glu Leu Gly Leu Thr

35 40 45

Gly Ile Leu Leu Asp Leu Asp Asp Pro Gln Ser Val Glu Asn Ala Ala

50 55 60

Ala Glu Val Ile Ala Leu Thr Glu Asn Arg Leu Tyr Gly Leu Phe Asn

65 70 75 80

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

85 90 95

Gln Met Glu Gln Gln Phe Ser Ala Asn Phe Phe Gly Ala His Gln Leu

100 105 110

Thr Met Leu Leu Leu Pro Ala Met Gln Pro His Gly Glu Gly Arg Ile

115 120 125

Val Met Thr Ser Ser Val Met Gly Leu Ile Ser Ser Pro Gly Arg Gly

130 135 140

Ala Tyr Ala Ala Ser Lys Tyr Ala Leu Glu Ala Trp Ser Asp Ala Leu

145 150 155 160

Arg Met Glu Leu Arg His Ser Gly Ile Lys Val Ser Leu Ile Glu Pro

165 170 175

Gly Pro Ile Arg Thr Arg Phe Thr Asn Asn Val Asn Gln Thr Gly Asp

180 185 190
 Lys Thr Val Glu Asn Pro Gly Ile Ala Ala Arg Phe Thr Leu Gly Pro
 195 200 205
 Glu Ala Val Val Glu Lys Val Arg His Ala Phe Val Ser Asp Lys Pro
 210 215 220
 Lys Leu Arg Tyr Pro Val Thr Leu Val Thr His Ala Val Gly Trp Leu
 225 230 235 240
 Lys Arg Leu Leu Pro Ala Arg Ala Met Asp Lys Ile Ile Gln Gly
 245 250 255
 <210> 46
 <211> 269

 <212> PRT
 <213> Klebsiella pneumoniae 342
 <400> 46
 Met Thr Pro Glu Arg Gln Gly Ser Leu Thr Gly Lys Val Met Gln Lys
 1 5 10 15
 Thr Val Leu Val Thr Gly Cys Ser Ser Gly Ile Gly Leu Glu Ser Ala
 20 25 30
 Leu Asp Leu Thr Arg Gln Gly Phe Arg Val Leu Ala Ala Cys Arg Lys
 35 40 45
 Ala Glu Asp Val Ala Arg Met Gln Glu Leu Gly Leu Thr Gly Ile Leu
 50 55 60

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

130 135 140
 Ser Ser Val Met Gly Leu Ile Ala Ser Pro Gly Arg Gly Ala Tyr Ala
 145 150 155 160
 Ala Ser Lys Tyr Ala Leu Glu Ala Trp Ser Asp Ala Leu Arg Met Glu
 165 170 175
 Leu Arg His Ser Gly Ile Gln Val Ser Leu Ile Glu Pro Gly Pro Ile
 180 185 190
 Arg Thr Arg Phe Thr Asp Asn Val Asn Gln Thr Gln Ser Asp Lys Pro
 195 200 205

 Val Glu Asn Pro Gly Ile Ala Ala Arg Phe Thr Leu Gly Pro Glu Ala
 210 215 220
 Val Val Glu Lys Val Arg His Ala Phe Thr Ser Asp Lys Pro Lys Leu
 225 230 235 240
 Arg Tyr Pro Val Thr Leu Val Thr His Ala Val Ala Leu Leu Lys Arg
 245 250 255
 Leu Leu Pro Ala Arg Ala Met Asp Lys Ile Ile His Gly
 260 265
 <210> 47
 <211> 256

 <212> PRT
 <213> Klebsiella oxytoca KCTC 1686
 <400> 47
 Met Gln Lys Ser Val Leu Val Thr Gly Cys Ser Ser Gly Ile Gly Leu
 1 5 10 15
 Glu Ser Ala Leu Asp Leu Lys Arg Gln Gly Phe Asn Val Leu Ala Ala
 20 25 30
 Cys Arg Lys Ala Glu Asp Val Ala Arg Met Gln Glu Leu Gly Leu Thr
 35 40 45
 Gly Val Leu Leu Asp Leu Asp Asp Pro Gln Ser Val Glu Arg Ala Ala
 50 55 60

 Ala Glu Val Ile Ala Leu Thr Asp Asn Arg Leu Tyr Gly Leu Phe Asn
 65 70 75 80

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

 130 135 140
 Ala Tyr Ala Ala Ser Lys Phe Ala Leu Glu Ala Trp Ser Asp Ala Leu
 145 150 155 160
 Arg Met Glu Leu Arg His Ser Gly Val Lys Val Ile Leu Ile Glu Pro
 165 170 175
 Gly Pro Ile His Thr Arg Phe Thr Asn Asn Val Asn Gln Thr Gln Val
 180 185 190
 Asp Lys Pro Val Glu Asn Pro Gly Ile Ala Ala Arg Phe Thr Leu Gly
 195 200 205

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

 <210> 48
 <211> 256
 <212> PRT
 <213> Cronobacter sakazakii ATCC BAA-894
 <400> 48

 Met Gln Lys Ser Leu Leu Ile Thr Gly Cys Ser Ser Gly Ile Gly Leu

 1 5 10 15
 Ala Cys Ala His Glu Leu Arg Arg Gln Gly Phe Gln Ile Leu Ala Ala
 20 25 30
 Cys Arg Lys Pro Asp Asp Val Ala Arg Met Asn Thr Leu Gly Phe Ile

35 40 45
 Gly Ile Glu Leu Asp Leu Asp Ser Pro Gln Ser Val Glu Ala Ala Ala
 50 55 60
 Gln Glu Val Ile Arg Leu Thr Asn Asn Arg Leu Tyr Gly Ile Phe Asn
 65 70 75 80

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

 145 150 155 160
 Arg Met Glu Leu Arg Tyr Ser Gly Val Lys Val Ser Leu Ile Glu Pro
 165 170 175
 Gly Pro Ile Arg Thr Arg Phe Thr Glu Asn Val Asn Gln Thr Gln Ala
 180 185 190
 Asp Lys Pro Val Glu Asn Pro Gly Ile Ala Ala Arg Phe Thr Leu Gly
 195 200 205
 Pro Glu Ala Val Val Ala Lys Val Arg His Ala Phe Glu Ser Pro Asn
 210 215 220

 Pro Arg Ile Arg Tyr Pro Val Thr Leu Val Thr His Ala Val Gly Ile
 225 230 235 240
 Leu Lys Arg Leu Leu Pro Thr Arg Ile Met Asp Lys Ile Leu His Gly
 245 250 255

 <210> 49
 <211> 810
 <212> DNA
 <213> Shigella sonnei Ss046
 <400> 49

tcacccctgc aatattttgt ccatcacgcg ccccggcagc aggcgcttaa gcaccattac	60
tgcccaagtc accagcgcta ccggaatagc catcttcggc ttctcgctaa taaaagcatg	120
gcgtactttg tccaccaccg ctcccggtcc caacgtaaag cgggcggcga tgccgggatt	180
ttcgactggt ttatcacttt gcgtctggtt gacgttgcg gtgaagcgag tacgaatggg	240
accgggttcg atcaggctga ctttaattcc gctgtggcgt agtccatgc gcagcgcgc	300
tgaccatgcc tccagcgcat atttactggc cgcgtaagcg ccacgaccg gtgtggagat	360
taatcccatc accgatgatg tcatcacaat acgcccctca ccgtgcgcta acatcgcggg	420
taacaggcgc atggtgagct ggtgtgcgcc gaaaaagtgt gcggaaaact gctgttccat	480
ctgcgcgagg ctgatgggtg aaagggggcc atacatgccg aatccggcat tgtaaagat	540
cccatacaga caattatcgg tcagggcgat cacttcgtcg gctgcgcgat caacactttc	600
tggcgaatcc agatcgatca acacccgggt aaatcccatg ctgttcatac gctcaacatc	660
atccggtttc cggcaacctg ccagcacatg aaaaccctgg cgttttaatt cgagcgcgct	720
ttccaggcca attccactgg aacatccggt aattaagacc gatttttgca taactttacc	780
tgtcaggatc tccgttgctt tatgagtc	810
<210> 50	
<211> 771	
<212> DNA	
<213> Citrobacter rodentium ICC168	
<400> 50	
tcacccgtgt aaaaattttgt ccatcaggcg tccgggcagc aggcgcttca gcagcataac	60
ggcccaggta accagcgta ccggaatagc aagcctgggt ttctcgctga caaaagcgtg	120
gcgcactttg tcgacgacgg ctcccgatc gagcgtaaac cgcgcggcaa taccggggtt	180
ttccaccggc ttatcgcttt gcgtctgatt gacgttttcg gtaaaacggg tgcaatggg	240
gccgggttcg atcaggctga ctttaatccc gctgtggcgc aatccatac gcagcgcac	300
cgccacgct tccagcgct atttactggc ggcataggcg ccgcgtccgg gcgtggagat	360
cagtcccatc accgatgagg tcatgacaat gcggccttcg ccgtgcggca gcatggcggg	420
caacaggcgc atcgtcagct gatgagcgcc aaaaaagtgt gttgaaaatt gctgtccat	480
ctgtcgcgg ctgatgggtg aaaggggacc gtagacgccg taaccggcgt tgtaaagat	540
cccaaataaa cagttatcgg tcagggcgat cactcatcg gcggcgcgat cgacgtctc	600
gggcgagtcg agatcgagca gtacgccgtt aaatcccata ctctcatgc gggcgacatc	660
ttcaggttta cggcagcctg caaggatttg aaaaccctgg cgcttcagtt caagcgcgt	720

ttccagacca attccgctgg aacaacctgt gattaagacc gatttttgca t 771

<210> 51

<211> 771

<212> DNA

<213> Salmonella enterica subsp. enterica serovar Typhimurium str. LT2

<400> 51

tcagctttgt aaaattttat ccagcatccg gcccggaat atgcgtttca gcgcatcac 60

cgcccagggtg accagcgtaa cgggtagcg cagcctgggt ttatcgcttg caaaagcatg 120

acgcactttc gctactaccg cgtcaggatc gcgggtgaaa cgcgcggcga tgccgggatt 180

ttctaccggt ttatcacgt cgtcttggtt aacgttatcg gtaaacggg tacgaatcgg 240

ccccggttct atcaggctca ctttaatgcc tgtgtgccgt agtccatac gtagcgcgtc 300

agaccaggct tcgagcgcat atttgctggc ggcgtaagcg ccgcgccctg gggtaggaaat 360

cagccccatc accgatgagg tcatacaat gcgtccttcg ccatgcggca acatggcggg 420

aagtaagcgc atggtaagct gatgcgcgcc gaagaaatta ctggaaaact gttgctccat 480

ctgttcacgg ctgatggtag aaagggggcc atacacgccg tagccggcat tgtaaagat 540

cccatacaga cgattatcgg tcaggcgcat cacttcacg gcggcacgat ccacgctttc 600

aggggagtcg agatcgatca acacgccgtt aaatccata ctgttcaccc gggcgacatc 660

atcaggtttg cggcagcccc ccaggatttg aaagccctgg cgtttaagtt caagcgcgt 720

gtccagaccg attccgctgg aacatcctgt tattaagacc gatttttgca t 771

<210> 52

<211> 810

<212> DNA

<213> Enterobacter cloacae EcWSU1

<400> 52

tcaaccgtgt aaaattttgt ccatcatccg gcccggcagc aggcgtttta accagcctac 60

cgcatgggtt accagcgta cgggtagcg catttttggg ttttcgtct caaaagcatg 120

gcgcacttta gcgacgactg cttccggccc gagcgtaaag cgtgcggcga taccgggatt 180

ttcgaccggt ttatccgctt gggcttggtt aacgttttcc gtaaagcggg tgccgatagg 240

gcccggctca atcaggctga ctttaatgcc gctgtgacgc aactccatgc gcagggcac 300

ggaccaggcc tccagcgcat atttgctggc ggcataggcc ccgcggcccc gggttgaaat 360

tagccccatc acggaggagg tcatacaat gcgtccttca ccgtgcggca gcatggcggg 420

gagcaggcgc atggtgagct gatggacgcc aaagaaattg gtggaaaact gttttccag 480

ctgctcgcgg gtgagcgtgt ccaggggacc atacacgccg taaccggcat tgttaaaaag 540

gccgtacaga cggttgtcgg tcagggccag cacctcatcg gctgcgcgtt ccacgctttc 600

cggtgagtct aaatccagta acacaccggi aaagcccagc ccgttcatgc gttcgacgtc 660

gtcgggttta cggcaggccg ccagtacca atatccctga cgcttcaatt cgagtgcact 720

ttcaagaccg attccgctgg aacatcctgt aattaagacc gatttttgca taactttacc 780

tgtcaggatc tccgctgaat taggagtc 810

<210> 53

<211> 768

<212> DNA

<213> Enterobacter aerogenes KCTC 2190

<400> 53

tcagccctgg ataattttgt ccatcgcacg ggccggtagc aggcgtttta accagcctac 60

cgcgtgggtg accagcgtca ccggatagcg cagtttcggc ttgtcgtga cgaaggcgtg 120

acgcactttt tccaccaccg cctccgggcc gagcgtaaag cgcgcggcaa taccgggggtt 180

ttctaccgtc ttgtcgcgg tctggttaac attgttggtg aagcgggtgc ggattggtcc 240

aggctcaatc aggtgactt ttatgccact gtggcggagt tccatgcgca gcgcacgga 300

ccagcctcc agcgcgtatt tactggcggc ataggcgccg cggccggggc tggaaatcag 360

ccccatcacc gaagaggta tgacgatgcg gccctcgccg tgaggctgca tagcgggcag 420

cagcagcatg gtgagctggt gggcgccgaa aaagttagca gaaaactgct gttccatctg 480

ctgacggctg atggtggtga gcgggccgta gacgccgtag ccagcattat tgaacaggcc 540

atacaggcgg ttctcggtca gcgcgatgac ctgcggcgcg gcattttcga cgctttgcgg 600

atcatcgaga tccagcaaaa taccggtcag acccagctcc tgcatacggg cgacatcctc 660

ccccttacgg caggcgcca gtaccggaa gccctggcgc tttaaatcaa gcgcgctttc 720

gaggccaatt ccgtggaac atcctgtaac taaaaccgat ttttgc 768

<210> 54

<211> 810

<212> DNA

<213> Klebsiella pneumoniae 342

<400> 54

atgactcctg aacgccaggg gtcgtgaca ggtaaagtta tgcaaaaaac ggtttttagtg 60

acaggctgtt ccagcggaat tggctctggaa agcgcgctgg atttaaccg ccagggcttt	120
cgcgtgctgg cggcctgccg caaagcggag gacgtcgcgc ggatgcagga gctggggcta	180
accggtattt tgctcgatct ggatgacca cagagcgtgg agcgcgcggc ggcggagggtg	240
attgcgtga ccgataatcg tctctatggc ctgttcaata acgccggtta tggcgtctat	300
ggcccgtga acaccatcag ccgccagcag atggaacagc agttttccgc caactttttc	360
ggcgcgcac cagtgacat gctgctgctg ccggcgatga cgccgcacgg cgaagggcgc	420
atcgtgatga cctcctcggg gatggggctt atcgccagtc ccggtcgcgg ggcctatgcc	480
gccagtaaat acgccctcga ggcgtggtcc gatgcgctac gcatggaatt gcgccacagt	540
ggtattcagg tcagcctgat tgaaccgggc ccgatccgca cccgctttac cgataacgtc	600
aaccagacgc aaagcgataa acccgtagaa aatccgggca tcgccgctcg ctttacctc	660
gggccggagg cgggtgtgga aaaagtgcgc cacgctttta ccagcgataa gccaaagctg	720
cgctaccggg tcactctggt gactcacgcg gtggcgctgc tcaaagcct gctgcctgcc	780
cgcgcgatgg acaaaattat ccacggctga	810
<210> 55	
<211> 770	
<212> DNA	
<213> Klebsiella oxytoca KCTC 1686	
<400> 55	
ttagccctgg ataattttgt ccatcgcgcg ggtcggcagc aggcgtttca gcagggcgac	60
cgcgtcgcc accagagtga ccgggtagcg cagcttcggg ttatcgctga caaaagcatg	120
gcgcactttt tccaccaccg cttccggccc gagagtaaag cgggcggcga tgccgggggtt	180
ttccaccggc ttatcgacct gggctcgtgt gacgttgttg gtgaagcggg tgtggatcgg	240
gcccggtcga atcaggatga ccttaactcc gctgtggcgc agtccatgc gcagcgcatc	300
ggaccaggct tccagcgca atttctggc ggcataggcg ccgcggccgg ggctggagat	360
aattcccatc accgatgagg tcatcacaat acgcccttcg ccgtcgggt ccattgccgg	420
aagcagcagc atcgttagct gatggacgcc gaaaaagtgt gcggaaaact gctgctccat	480
ctgccggcgg ctgatgtcg gcagcgggcc gtagacgcg tagcccgct tgtaaacag	540
gccgtacagg cggttgtcgg tcagggaat cacctcggcg gcggcgctt caacgctctg	600
cggatcgta agatccagca acacaccggg cagccccagc tcctgcatcc gtgcgacatc	660
ttccgctttg cggcaggcgg ccagcacgtt aaagccctgg cgctttaaata caagggcgct	720
ttcaagacca attccgctgg aacatcctgt aactaaaacc gatttttgca	770

<210> 56
 <211> 771
 <212> DNA
 <213> Cronobacter sakazakii ATCC BAA-894
 <400> 56

atgcaaaaat cactcttaat tacaggatgt tccagcggta ttggtcttgc ctgcgcgcac 60
 gagctgcgcc gtcagggctt tcaaatcctt gcggcctgcc gcaagccgga cgatgtggcg 120
 cggatgaata cgctgggggtt catcggcatt gaactcgatc tcgactcccc gcagagcgtg 180
 gaggcgcgcg cgcaggaggt gattcgcttg acgaataacc gcctgtacgg gatittcaat 240
 aacggcggct acggcgtgta cggtcgcgtc tcaacatca gccgcgcgca gatggaacag 300

cagttctcga cgaacttttt cggcgcgcac cagctcacca tgctgctgct gcccgcgatg 360
 ctgccgcacg gcgaagggcg catcgtcaat accagttcgg tcatgggggtt aatttcaacc 420
 cccggacgcg gcgcctatgc cgccagcaaa tatgcgctgg aggcgtggtc tgacgcgctg 480
 cgcattggagc tgcgctacag cggcgtgaag gtgagcctga ttgagcccgg ccctattcgc 540
 acgcgcttta cggaaaacgt caaccagacg caggcggata aaccggtcga aaaccccggc 600
 atcgccgcgc gttttacact cggcccggag gcggtcgtcg ccaaagtgcg ccatgctttt 660
 gagagcccga atccgcgtat tcgctatccg gtcacgctgg tgacgcatgc ggtaggcata 720

ctcaaacggc tactgccgac gcgaatcatg gacaaaattt tacacgggta a 771