



(19) 대한민국특허청(KR)  
 (12) 등록특허공보(B1)

(45) 공고일자 2017년11월16일  
 (11) 등록번호 10-1797713  
 (24) 등록일자 2017년11월08일

- (51) 국제특허분류(Int. Cl.)  
*C12N 15/63* (2006.01) *C12N 15/70* (2006.01)  
*C12N 9/00* (2006.01) *C12N 9/02* (2006.01)  
*C12N 9/88* (2006.01) *C12P 7/24* (2006.01)

- (52) CPC특허분류  
*C12N 15/63* (2013.01)  
*C12N 15/70* (2013.01)  
(21) 출원번호 10-2015-0067557  
(22) 출원일자 2015년05월14일  
 심사청구일자 2015년05월14일  
(65) 공개번호 10-2016-0135011  
(43) 공개일자 2016년11월24일

- (56) 선행기술조사문헌  
CN104498540 A\*  
Hsu 등. Holzforschung. Vol 66, No. 7, 페이지 897-904 (2012.04.19.)\*

\*는 심사관에 의하여 인용된 문헌

- (73) 특허권자  
**한국과학기술원**  
 대전광역시 유성구 대학로 291(구성동)  
 재단법인 지능형 바이오 시스템 설계 및 합성 연구단  
 대전광역시 유성구 대학로 291, 401호(구성동, 한국과학기술원케이아이빌딩)

- (72) 발명자  
**정기준**  
 대전광역시 유성구 문지로 14 1동 304호 (도룡동, 과기원교수아파트)  
**김선창**  
 대전광역시 유성구 대학로 291 (구성동 23, 한국과학기술원)  
 (뒷면에 계속)

- (74) 대리인  
**손민**

전체 청구항 수 : 총 10 항

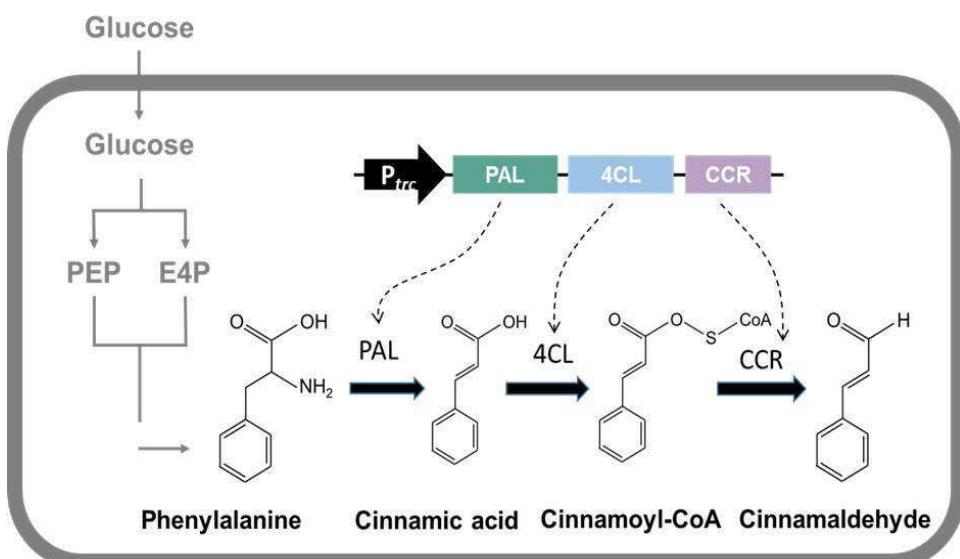
심사관 : 김현태

(54) 발명의 명칭 신남알데하이드의 제조 방법

(57) 요약

본 발명은 재조합 미생물을 이용한 신남알데하이드의 제조방법에 관한 것이다.

**대 표 도** - 도1



## (52) CPC특허분류

*C12N 9/0008* (2013.01)  
*C12N 9/88* (2013.01)  
*C12N 9/93* (2013.01)  
*C12P 7/24* (2013.01)  
*C12Y 102/01044* (2013.01)  
*C12Y 403/01024* (2013.01)  
*C12Y 602/01012* (2013.01)

## (72) 발명자

## 방현배

대전광역시 유성구 대학로 291 (구성동 23, 한국과학기술원)

## 이윤혁

대전광역시 유성구 대학로 291 (구성동 23, 한국과학기술원)

## 정석채

대전광역시 유성구 대학로 291 (구성동 23, 한국과학기술원)

## 이) 발명을 지원한 국가연구개발사업

과제고유번호 2014M3A6A8066443

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 글로벌프론티어사업연구개발사업(지능형 바이오 시스템 설계 및 합성 연구)

연구과제명 친환경 선충 퇴치제 고효율 생산 세포공장 개발

기여율 1/1

주관기관 (재)지능형 바이오 시스템 설계 및 합성 연구단

연구기간 2014.09.01 ~ 2015.08.31

---

## 명세서

### 청구범위

#### 청구항 1

서열번호 1로 표시되는 염기서열로 이루어진 스트렙토마이세스 마리티무스(*Streptomyces maritimus*) 유래의 *pal* (phenylalanine ammonia lyase) 유전자, 서열번호 3으로 표시되는 염기서열로 이루어진 스트렙토마이세스 실리칼라(*Streptomyces coelicolor*) 유래의 *4cl* (4-coumarate:CoA ligase) 유전자, 및 서열번호 5로 표시되는 염기서열로 이루어진 애기장대(*Arabidopsis thaliana*) 유래의 *ccr* (cinnamoyl Co-A reductase) 유전자를 포함하는, 신남알데하이드 생산용 발현 카세트.

#### 청구항 2

삭제

#### 청구항 3

제1항에 있어서,

상기 *pal* 유전자는 서열번호 2로 표시되는 아미노산 서열로 이루어진 단백질을 코딩하는 것인, 발현 카세트.

#### 청구항 4

삭제

#### 청구항 5

제1항에 있어서,

상기 *4cl* 유전자는 서열번호 4로 표시되는 아미노산 서열로 이루어진 단백질을 코딩하는 것인, 발현 카세트.

#### 청구항 6

삭제

#### 청구항 7

제1항에 있어서,

상기 *ccr* 유전자는 서열번호 6로 표시되는 아미노산 서열로 이루어진 단백질을 코딩하는 것인, 발현 카세트.

#### 청구항 8

제1항, 제3항, 제5항 및 제7항 중 어느 한 항의 발현 카세트를 포함하는, 신남알데하이드 생산용 벡터.

#### 청구항 9

제8항의 벡터를 포함하는 신남알데하이드 생산용 형질전환체.

**청구항 10**

제9항에 있어서,

상기 형질전환체는 에스케리치아(*Escherichia*) 속 미생물인 것인, 형질전환체.

**청구항 11**

제9항의 형질전환체를 배양하는 단계를 포함하는, 신남알데하이드의 제조방법.

**청구항 12**

제11항에 있어서,

상기 형질전환체는 폐닐알라닌을 포함하는 배지에서 배양되는 것인, 신남알데하이드의 제조방법.

**청구항 13**

제11항에 있어서,

상기 형질전환체는 카사미노산(casamino acid)을 포함하는 배지에서 배양되는 것인, 신남알데하이드의 제조방법.

**발명의 설명****기술 분야**

[0001]

본 발명은 재조합 미생물을 이용한 신남알데하이드의 제조방법에 관한 것이다.

**배경 기술**

[0003]

신남알데하이드는 계피의 향과 맛을 나타내는 주요 성분으로, 계피 정유의 90%를 차지한다. 주로 계피유 또는 카시아유에 진한 아황산수소나트륨 용액을 첨가한 후, 생성되는 첨가물을 분리하여 에탄올로 세척하고, 묽은 황산 또는 탄산나트륨 수용액에서 분해하여 수증기 증류 후 진공 증류하여 제조한다. 또는 벤즈알데히드, 물, 수산화나트륨의 혼합물을 휘저으면서 아세트알데히드를 방울방울로 떨어뜨려 반응시켜 벤젠으로 추출하고, 감압 하에 분별 증류하여 제조한다.

[0005]

신남알데하이드는 다양한 용도에 사용되고 있다. 예를 들어, 한국공개특허 2003-0033282 는 신남알데하이드의 항산화활성에 대하여 기재하고 있고, 한국등록특허 10-0683113는 신남알데하이드의 비만 치료 효과에 대하여 기재하고 있다. 한국공개특허 2013-0038000는 계피로부터 분리한 트랜스-신남알데하이드의 B형 간염 치료 효과에 대하여 기재하고 있다.

[0007]

이와 같이 다양한 유용한 효과가 알려진 신남알데하이드를 효과적으로 생산하기 위한 방법이 연구될 필요가 있다. 한국등록특허 10-0624236은 벤즈알데하이드 유도체와 바이닐아세테이트를 탄산칼륨과 물 존재 하에서, 아세토니트릴 용매 중에서 가열, 환류시켜 신남알데하이드 유도체를 제조하는 방법에 대하여 기재하고 있다.

[0009]

이러한 배경 하에, 본 발명자들은 신남알데하이드를 효율적으로 생산하는 방법을 개발하고자 예의 연구 노력한 결과, 스트렙토마이세스 마리티무스 유래의 pal 유전자, 스트렙토마이세스 실리칼라 유래의 4cl 유전자, 및 애기장대 유래의 ccr 유전자를 포함하는 재조합 균주를 이용하여 신남알데하이드를 높은 수율로 생산할 수 있음을 확인하고 본 발명을 완성하였다.

**발명의 내용**

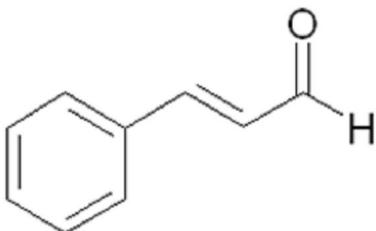
## 해결하려는 과제

- [0011] 본 발명의 주된 목적은 스트렙토마이세스 마리티무스(*Streptomyces maritimus*) 유래의 pal (phenylalanine ammonia lyase) 유전자, 스트렙토마이세스 실리칼라(*Streptomyces coelicolor*) 유래의 4cl (4-coumarate:CoA ligase) 유전자, 및 애기장대(*Arabidopsis thaliana*) 유래의 ccr (cinnamoyl Co-A reductase) 유전자를 포함하는, 신남알데하이드 생산용 발현 카세트를 제공하는 것이다.
- [0012] 본 발명의 다른 목적은 상기 발현 카세트를 포함하는, 신남알데하이드 생산용 벡터를 제공하는 것이다.
- [0013] 본 발명의 다른 목적은 상기 벡터를 포함하는 신남알데하이드 생산용 형질전환체를 제공하는 것이다.
- [0014] 본 발명의 다른 목적은 상기 균주를 배양하는 단계를 포함하는, 신남알데하이드의 제조방법을 제공하는 것이다.

## 과제의 해결 수단

- [0016] 상기의 목적을 달성하기 위한 하나의 양태로서, 본 발명은 신남알데하이드 생합성 유전자를 포함하는, 신남알데하이드 생산용 발현 카세트, 이를 포함하는 벡터, 이를 포함하는 형질전환체, 또는 이를 이용한 신남알데하이드의 제조방법을 제공한다.
- [0017] 본 발명자들은 신남알데하이드 생합성 유전자인 pal (phenylalanine ammonia lyase) 유전자, 4cl (4-coumarate:CoA ligase) 유전자, 및 ccr (cinnamoyl Co-A reductase) 유전자를 도입한 미생물을 제조한 결과, 상기 미생물로부터 1.98 mg/L의 수율로 신남알데하이드를 생산할 수 있음을 확인하였다(도 5 내지 도 8).
- [0018] 또한, 스트렙토마이세스 마리티무스 유래의 PAL, 스트렙토마이세스 실리칼라 유래의 4CL, 애기장대 유래의 CCR의 조합이 다른 미생물 유래의 유전자에 비하여 우수한 활성을 나타냄을 확인하였다(도 4a 및 도 4b).
- [0020] 이하, 본 발명에서의 신남알데하이드 생산용 발현 카세트를 구체적으로 설명한다.
- [0022] 본 명세서에서, "신남알데하이드"는 하기의 화학식 1로 표시되는 화합물을 의미한다.

## 화학식 1



- [0023]
- [0025] 본 명세서에서, "신남알데하이드 생합성 유전자"는 pal (phenylalanine ammonia lyase) 유전자, 4cl (4-coumarate:CoA ligase) 유전자, 및 ccr (cinnamoyl Co-A reductase) 유전자를 의미한다.
- [0027] 상기 pal 유전자는 스트렙토마이세스 마리티무스(*Streptomyces maritimus*) 유래일 수 있다. 본 발명의 일 실시 예에서, 스트렙토마이세스 마리티무스 유래의 PAL 효소는 애기장대 PAL 효소에 비하여 효소 활성이 우수함을 확인하였다(도 4a).
- [0028] 비제한적인 일 예로 상기 pal 유전자는 서열번호 1로 표시되는 염기서열을 포함하는 것일 수 있다. 상기 pal 유전자는 서열번호 1로 표시되는 염기서열과 적어도 70% 이상, 바람직하게는 80% 이상, 더욱 바람직하게는 90% 이상, 더 더욱 바람직하게는 95% 이상의 서열 상동성을 갖는 서열로 표시되는 것일 수 있다. 상기 서열 중 하나 이상의 염기가 치환, 결실, 삽입 또는 이들의 조합에 의해 변이된 서열 또한 이에 포함될 수 있음을 자명하다.
- [0029] 비제한적인 일 예로 상기 pal 유전자는 서열번호 2로 표시되는 아미노산 서열을 갖는 PAL 단백질을 코딩하는 것일 수 있다.
- [0030] 상기 4cl 유전자는 스트렙토마이세스 실리칼라(*Streptomyces coelicolor*) 유래일 수 있다.
- [0031] 비제한적인 일 예로 상기 4cl 유전자는 서열번호 3로 표시되는 염기서열을 포함하는 것일 수 있다. 상기 4cl 유

전자는 서열번호 3로 표시되는 염기서열과 적어도 70% 이상, 바람직하게는 80% 이상, 더욱 바람직하게는 90% 이상, 더 더욱 바람직하게는 95% 이상의 서열 상동성을 갖는 서열로 표시되는 것일 수 있다. 상기 서열 중 하나 이상의 염기가 치환, 결실, 삽입 또는 이들의 조합에 의해 변이된 서열 또한 이에 포함될 수 있음을 자명하다.

[0032] 비제한적인 일 예로 상기 4cl 유전자는 서열번호 4로 표시되는 아미노산 서열을 갖는 4CL 단백질을 코딩하는 것일 수 있다.

상기 ccr 유전자는 애기장대(*Arabidopsis thaliana*) 유래일 수 있다.

[0034] 비제한적인 일 예로 상기 ccr 유전자는 서열번호 5로 표시되는 염기서열을 포함하는 것일 수 있다. 상기 ccr 유전자는 서열번호 5로 표시되는 염기서열과 적어도 70% 이상, 바람직하게는 80% 이상, 더욱 바람직하게는 90% 이상, 더 더욱 바람직하게는 95% 이상의 서열 상동성을 갖는 서열로 표시되는 것일 수 있다. 상기 서열 중 하나 이상의 염기가 치환, 결실, 삽입 또는 이들의 조합에 의해 변이된 서열 또한 이에 포함될 수 있음을 자명하다.

[0035] 비제한적인 일 예로 상기 ccr 유전자는 서열번호 6로 표시되는 아미노산 서열을 갖는 CCR 단백질을 코딩하는 것일 수 있다.

[0036] 상기 단백질을 코딩하는 핵산을 화학적으로 합성하여 제조하는 경우, 당업계에 널리 공지된 합성법, 예를 들어 문헌(Engels and Uhlmann, Angew Chem Int Ed Engl., 37:73-127, 1988)에 기술된 방법을 이용할 수 있으며, 트리에스테르, 포스파이트, 포스포르아미다이트 및 H-포스페이트 방법, PCR 및 기타 오토프라이머 방법, 고체 지지체상의 올리고뉴클레오파이어드 합성법 등에 의해서 제조할 수도 있다.

[0037] 본 발명의 일 실시예에서, 스트렙토마이세스 실리칼라 유래의 4CL과 애기장대 유래의 CCR의 조합을 사용할 경우, 애기장대 유래의 4CL과 애기장대 유래의 CCR의 조합에 비하여 효소 활성이 우수함을 확인하였다(도 4b).

[0039] 본 명세서에서 서열과 관련하여 사용된 용어, "상동성"은 주어진 아미노산 서열 또는 염기 서열과 일치하는 정도를 의미하며 백분율로 표시될 수 있다. 본 명세서에서, 주어진 아미노산 서열 또는 염기 서열과 동일하거나 유사한 활성을 갖는 그의 상동성 서열이 "% 상동성"으로 표시된다. 예를 들면, 점수(score), 동일성(identity) 및 유사도(similarity) 등의 매개 변수(parameter)들을 계산하는 표준 소프트웨어, 구체적으로 BLAST 2.0를 이용하거나, 정의된 엄격한 조건하에서 써던 혼성화 실험에 의해 서열을 비교함으로써 확인할 수 있으며, 정의되는 적절한 혼성화 조건은 해당 기술 범위 내이고(Sambrook et al., 1989, Infra), 당업자에게 잘 알려진 방법으로 결정될 수 있다.

[0041] 상기 유전자는 형질전환체에 코돈 최적화된 것일 수 있다. 상기 코돈 최적화는, 숙주에서의 상기 유전자의 발현을 보다 효율적으로 하기 위하여 유전자의 코돈을 상기 숙주 유전자에서 높은 빈도로 사용되는 코돈으로 치환하는 것을 의미한다. 최적화를 위한 방법은 당업계에서 형질전환체에서의 단백질 발현을 증가시키기 위하여 사용되는 방법이라면 제한되지 않고 사용할 수 있다.

[0043] 본 명세서에서, "발현 카세트"는 신남알데하이드 생합성 유전자를 포함하고 있어서 신남알데하이드를 발현시킬 수 있는 단위 카세트를 의미한다. 또한, 본 발명에서 발현 카세트는 발현구조체와 혼용될 수 있다. 본 발명에 따른 발현 카세트는 본 발명의 목적 상, 균주에 도입되어 신남알데하이드를 생산할 수 있다.

[0045] 다른 양태로서, 본 발명은 상기 발현 카세트를 포함하는, 신남알데하이드 생산용 벡터를 제공한다.

[0047] 본 명세서에서, "벡터" 또는 "발현 벡터"는 적합한 숙주 내에서 목적 단백질을 발현시킬 수 있도록 적합한 조절 서열에 작동 가능하게 연결된 상기 목적 단백질을 코딩하는 폴리뉴클레오티드의 염기서열을 함유하는 DNA 제조물을 의미한다. 상기 조절 서열은 전사를 개시할 수 있는 프로모터, 그러한 전사를 조절하기 위한 임의의 오퍼레이터 서열, 적합한 mRNA 리보솜 결합부위를 코딩하는 서열, 및 전사 및 해독의 종결을 조절하는 서열을 포함한다. 벡터는 적당한 숙주세포 내로 형질전환된 후, 숙주 계놈과 무관하게 복제되거나 기능할 수 있으며, 계놈 그 자체에 통합될 수 있다.

[0048] 상기 벡터는 상기 신남알데하이드 생합성 유전자가 작동 가능하게 연결되어 포함된 것일 수 있다. 본 발명에서 "작동 가능하게 연결된(operably linked)"는 일반적 기능을 수행하도록 핵산 발현조절 서열과 목적하는 단백질 또는 RNA를 코딩하는 핵산 서열이 기능적으로 연결(functional linkage)되어 있는 것을 말한다. 예를 들어 프로모터와 단백질 또는 RNA를 코딩하는 핵산 서열이 작동 가능하게 연결되어 코딩서열의 발현에 영향을 미칠 수 있다. 재조합 벡터와의 작동적 연결은 당해 기술분야에서 잘 알려진 유전자 재조합 기술을 이용하여 제조할 수 있으며, 부위-특이적 DNA 절단 및 연결은 당해 기술 분야에서 일반적으로 알려진 효소 등을 사용할 수 있다.

- [0049] 본 발명에서 사용되는 발현 벡터는 숙주세포 내에서 복제 가능한 것이면 특별히 한정되지 않으며, 당업계에 알려진 임의의 벡터를 이용할 수 있다. 통상 사용되는 벡터의 예로는 천연 상태이거나 재조합된 상태의 플라스미드, 코스미드, 바이러스 및 박테리오파지를 들 수 있다. 예를 들어, 파지 벡터 또는 코스미드 벡터로서 pWE15, M13, λMBL3, λMBL4, λIXII, λASHII, λAPII, λt10, λt11, Charon4A, 및 Charon21A 등을 사용할 수 있으며, 플라스미드 벡터로서 pET계, pTrc계, pBR계, pUC계, pBluescript II 계, pGEM계, pTZ계, pCL계, pMAL계 또는 pHT계 등을 사용할 수 있다. 본 발명에서 사용 가능한 벡터는 특별히 제한되는 것이 아니며 공지된 발현 벡터를 사용할 수 있다. 구체적으로는 pET22b, pTrc99a, pDZ, pACYC177, pACYC184, pCL, pECCG117, pUC19, pBR322, pMW118, pCC1BAC, pMAL-p2x 또는 pHT43 벡터 등을 사용할 수 있다.
- [0050] 본 발명의 벡터는 상동재조합을 일으켜서 염색체 내로 삽입될 수 있으므로 상기 염색체 삽입 여부를 확인하기 위한 선별 마커(selection marker)를 추가로 포함할 수 있다.
- [0051] 적합한 발현벡터는 프로모터, 오퍼레이터, 개시코돈, 종결코돈, 폴리아데닐화 시그널, 인핸서 같은 발현 조절 요소 외에도 막 표적화 또는 분비를 위한 신호 서열 또는 리더 서열을 포함하며 목적에 따라 다양하게 제조될 수 있다. 개시 코돈 및 종결 코돈은 일반적으로 표적 단백질을 코딩하는 뉴클레오타이드 서열의 일부로 간주되며, 유전자 작제물이 투여되었을 때 개체에서 반드시 작용을 나타내야 하며 코딩 서열과 인프레임(in frame)에 포함될 수 있다. 벡터의 프로모터는 구성적 또는 유도성일 수 있다. 또한 발현벡터는 벡터를 함유하는 숙주 세포를 선택하기 위한 선택성 마커를 포함하고, 복제가능한 발현벡터인 경우 복제 기원을 포함할 수 있다. 벡터는 자가 복제하거나 숙주 DNA에 통합될 수 있다.
- [0052] 벡터를 제조하는 방법은 특별히 제한되지 아니하며 당해 기술 분야에서 통상적으로 사용되는 방법이라면 어느 것이든 사용할 수 있다.
- [0054] 다른 양태로서, 본 발명은 상기 벡터를 포함하는 신남알데하이드 생산용 형질전환체를 제공한다.
- [0055] 본 발명에서, "생산"이란 신남알데하이드를 균주 내에서 만들어내는 것뿐만 아니라, 신남알데하이드를 세포 밖, 예컨대 배양액으로 배출하는 것 역시 포함하는 개념이다.
- [0056] 본 명세서에서, "형질전환"은 표적 단백질을 코딩하는 폴리뉴클레오티드를 포함하는 벡터를 숙주세포 내에 도입하여 숙주세포 내에서 상기 폴리뉴클레오티드가 코딩하는 단백질이 발현할 수 있도록 하는 것을 의미할 수 있다. 형질전환된 폴리뉴클레오티드는 숙주세포 내에 발현될 수 있기만 한다면, 숙주세포의 염색체 내에 삽입되어 위치하거나 염색체 외에 위치하거나 상관없이 이를 모두를 포함한다. 또한, 상기 폴리뉴클레오티드는 표적 단백질을 코딩하는 DNA 및 RNA를 포함한다. 상기 폴리뉴클레오티드는 숙주세포 내로 도입되어 발현될 수 있는 것이면, 어떠한 형태로 도입되는 것이든 무관하다. 예를 들면, 상기 폴리뉴클레오티드는, 자체적으로 발현되는데 필요한 모든 요소를 포함하는 유전자 구조체인 발현 카세트(expression cassette)의 형태로 숙주세포에 도입될 수 있다. 상기 발현 카세트는 통상 상기 폴리뉴클레오티드에 작동 가능하게 연결되어 있는 프로모터(promoter), 전사 종결신호, 리보솜 결합부위 및 번역 종결신호를 포함한다. 상기 발현 카세트는 자체 복제가 가능한 발현벡터 형태일 수 있다. 또한, 상기 폴리뉴클레오티드는 그 자체의 형태로 숙주세포에 도입되어, 숙주세포에서 발현에 필요한 서열과 작동 가능하게 연결되어 있는 것일 수도 있다.
- [0057] 상기 형질전환체로 사용할 수 있는 미생물은 에스케리치아 속(Escherichia sp.), 시겔라 속(Shigella sp.), 시트로박터 속(Citrobacter sp.), 살모넬라 속(Salmonella sp.), 엔테로박터 속(Enterobacter sp.) 여시니아 속(Yersinia sp.), 크렙시엘라 속(Klebsiella sp.), 어위니아 속(Erwinia sp.), 코리네박테리움 속(Corynebacterium sp.), 브레비박테리움 속(Brevibacterium sp.), 락토바실러스 속(Lactobacillus sp.), 셀레노모나스 속(Selenomonas sp.), 비브리오 속(Vibrio sp.), 슈도모나스 속 (Pseudomonas sp.), 스트렙토마이시스 속 (Streptomyces sp.), 아카노박테리아 속 (Arcanobacterium sp.), 알칼리젠 속(Alcaligenes sp) 등에 속하는 미생물일 수 있으나, 이에 제한되지는 않는다. 구체적으로 에스케리치아 속 미생물일 수 있다.
- [0059] 다른 양태로서, 본 발명은 상기 형질전환체를 배양하는 단계를 포함하는, 신남알데하이드의 제조방법을 제공한다.
- [0060] 본 발명의 구체적인 일 실시양태에서는, 상기 신남알데하이드 생산능을 보유한 형질전환체를 배양한 다음, 그 배양물에 포함된 신남알데하이드 생산량을 측정한 결과, 상기 미생물로부터 1.98 mg/L의 수율로 신남알데하이드를 생산할 수 있음을 확인하였다(도 5 내지 도 8).
- [0062] 상기 형질전환체를 배양하는 단계는, 특별히 이에 제한되지 않으나, 공지된 회분식 배양방법, 연속식 배양방법,

유가식 배양방법 등에 의해 수행됨이 바람직하다. 이때, 배양조건은 특별히 이에 제한되지 않으나, 염기성 화합물(예: 수산화나트륨, 수산화칼륨 또는 암모니아) 또는 산성 화합물(예: 인산 또는 황산)을 사용하여 적정 pH (예컨대, pH 5 내지 9, 바람직하게는 pH 6 내지 8, 가장 바람직하게는 pH 6.8)를 조절할 수 있고, 산소 또는 산소-함유 가스 혼합물을 배양물에 도입시켜 호기성 조건을 유지할 수 있으며, 배양온도는 20 내지 45, 바람직하게는 25 내지 40를 유지할 수 있고, 약 10 내지 160 시간 동안 배양함이 바람직하다.

[0063] 상기 배양을 위하여 사용되는 배지는 탄소 공급원으로는 당 및 탄수화물(예: 글루코오스, 슈크로오스, 락토오스, 프리토오스, 말토오스, 몰라세, 전분 및 셀룰로오스), 유지 및 지방(예: 대두유, 해바라기씨유, 땅콩유 및 코코넛유), 지방산(예: 팔미트산, 스테아르산 및 리놀레산), 알코올(예: 글리세롤 및 에탄올) 및 유기화합물(예: 아세트산) 등을 개별적으로 사용하거나 또는 혼합하여 사용할 수 있고; 질소 공급원으로는 질소-함유 유기화합물(예: 펩톤, 효모 추출액, 육즙, 맥아 추출액, 옥수수 침지액, 대두 박분 및 우레아), 또는 무기 화합물(예: 황산암모늄, 염화암모늄, 인산암모늄, 탄산암모늄 및 질산암모늄) 등을 개별적으로 사용하거나 또는 혼합하여 사용할 수 있으며; 인 공급원으로 인산 이수소칼륨, 인산수소이칼륨, 이에 상응하는 나트륨 함유 염 등을 개별적으로 사용하거나 또는 혼합하여 사용할 수 있고; 기타 금속염(예: 황산마그네슘 또는 황산철), 아미노산 및 비타민과 같은 필수성장-촉진 물질을 포함할 수 있다. 본 명세서에서 상기 배지는 배양액과 동일한 의미로 사용될 수 있다.

[0065] 일 예로 상기 제조방법은 배양된 형질전환체 또는 그의 배양물로부터 신남알데하이드를 회수하는 단계를 추가로 포함할 수 있다.

[0066] 본 명세서에서, "배양물"은 미생물 배양 결과 얻어지는 물질로, 상기 배지, 배양되는 미생물, 및 배양되는 미생물로부터 분비되는 물질을 모두 포함하는 것일 수 있다. 예를 들어 균체를 배양하기 위해 필요한 영양 공급원, 예를 들어 탄소원, 질소원 등 이외에 무기염 성분, 아미노산, 비타민, 핵산 및/또는 기타 일반적으로 배양 배지 (또는 배양액)에 함유될 수 있는 성분들이 포함되어 있을 수 있다. 또한 예를 들어 균체가 생산·분비한 효소 등이 포함되어 있을 수 있다.

[0067] 배양에 의하여 생산된 신남알데하이드는 배지 중으로 분비되거나 세포 내에 잔류할 수 있으므로, 상기 배양물은 미생물 배양에 의하여 생산된 신남알데하이드를 포함할 수 있다.

[0068] 본 발명의 상기 배양 단계에서 생산된 신남알데하이드를 회수하는 방법은 배양방법, 예를 들어 회분식, 연속식 또는 유가식 배양 방법 등에 따라 당해 분야에 공지된 적합한 방법을 이용하여 배양액으로부터 신남알데하이드를 수집할 수 있다.

[0070] 일 예로, 상기 형질전환체는 폐닐알라닌을 포함하는 배지에서 배양되는 것일 수 있다. 상기 형질전환체에 포함된 pal 유전자는 폐닐알라닌을 기질로 사용하므로, 폐닐알라닌을 포함하는 배지에서 배양할 경우 우수한 효율로 신남알데하이드를 생산할 수 있다.

[0072] 일 예로, 상기 형질전환체는 카사미노산(casamino acid)을 포함하는 배지에서 배양되는 것일 수 있다. 본 발명의 일 실시예에서, 신남알데하이드 생합성 유전자를 포함하는 형질전환체를 카사미노산을 포함하는 배지에서 배양한 결과 최고 생산 수율이 1.98 mg/L로 높게 나타난 반면, 카사미노산을 포함하지 않은 배지에서 배양한 결과 최고 생산 수율은 1.23 mg/L로 나타났다.

### 발명의 효과

[0074] 본 발명에 따른 신남알데하이드 생산용 발현 카세트, 이를 포함하는 벡터, 이를 포함하는 형질전환체, 및 이를 이용한 신남알데하이드의 제조방법을 이용할 경우 우수한 효율로 신남알데하이드를 생산할 수 있다.

### 도면의 간단한 설명

[0076] 도 1은 본 발명의 일 실시예에 따른 신남알데하이드 생산 시스템에 대한 모식도를 나타낸 것이다.

도 2는 본 발명의 일 실시예에서 사용된 PAL, 4CL, CCR의 발현을 위하여 구축한 플라스미드를 나타낸 것이다.

도 3은 정제된 PAL, 4CL, CCR 단백질을 SDS-PAGE로 정량하여 나타낸 것이다.

도 4a는 최종적으로 생성된 신남산(cinnamate, CA)의 농도를 도 4b는 신남알데하이드(cinnamaldehyde, CAD)의 농도를 정량하여 나타낸 것이다.

도 5은 복합 배지 1에서 형질전환체를 배양한 후, 시간대 별로 측정한 광학 밀도 (OD600; open circle)와 이에

따른 신남알데하이드의 농도 (CAD con.; closed circle)를 나타낸 것이다.

도 6은 복합 배지 2에서 형질전환체를 배양한 후, 시간대 별로 측정한 광학 밀도 (OD600; open circle)와 이에 따른 신남알데하이드의 농도 (CAD con.; closed circle)를 나타낸 것이다.

도 7은 제한 배지 1에서 형질전환체를 배양한 후, 시간대 별로 측정한 광학 밀도 (OD600; open circle)와 이에 따른 신남알데하이드의 농도 (CAD con.; closed circle)를 나타낸 것이다.

도 8은 제한 배지 2에서 형질전환체를 배양한 후, 시간대 별로 측정한 광학 밀도 (OD600; open circle)와 이에 따른 신남알데하이드의 농도 (CAD con.; closed circle)를 나타낸 것이다.

### 발명을 실시하기 위한 구체적인 내용

[0077]

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

[0079]

#### 실시예 1. 재조합 유전자 발현을 위한 균주 및 플라스미드 제작

[0081]

먼저, 애기장대 (*Arabidopsis thaliana*) cDNA로부터 하기 표 1에 기재된 서열의 BHB 22-F 및 BHB 22-R, BHB 20-F 및 BHB 20-R, BHB 19-F 및 BHB 19-R를 이용하여 중합효소 연쇄반응을 통해 pal(phenylalanine ammonia-lyase, 서열번호 17), 4cl(4-coumarate:CoA ligase, 서열번호 18), ccr(cinnamoyl-CoA reductase, 서열번호 5) 유전자를 얻었다.

[0082]

또한 코돈 최적화되어 합성한 스트렙토마이세스 마리티무스 (*Streptomyces maritimus*) pal 유전자로부터 하기 표 1에 기재된 서열의 BHB 31-F 및 BHB 31-R를 이용하여 중합효소 연쇄반응을 통해 pal 유전자 (서열번호 19)를 얻었다. 또한, 스트렙토마이세스 실리칼라 (*Streptomyces coelicolor*) 유전체 DNA (genomic DNA)로부터 하기 표 1에 기재된 서열의 BHB 21-F 및 BHB 21-R을 이용하여 중합효소 연쇄반응을 통해 4cl 유전자(서열번호 3)를 얻었다.

### 표 1

BHB 19-F	GCATCTAGAAACACAAACAAGGAAGATAATGCACCACCACCACCAAC ACATGCCAGTCGACGTAGCC	서열번호 7
BHB 19-R	ATGCGCGGCCGCTTATCAAGACCCGATCTTAATGCCATTTC	서열번호 8
BHB 20-F	GCATCTAGACCGAAATCAAAGGAACACCAACGTATGCACCACCACCAC CACATGGGCCACAAGAACAAAG	서열번호 9
BHB 20-R	ATGCGCGCCGCTTATCACAATCCATTGCTAGTTTGCCC	서열번호 10
BHB 21-F	GCATCTAGACGAATACTGGAGGACCTAACAGTATGCACCACCACCAC CACATGTTCCGCAGCGAGTAC	서열번호 11
BHB 21-R	ATGCGCGCCGCTTATCATCGGGCTCCCTGAGCT	서열번호 12
BHB 22-F	GCATCTAGACCTTAAGGAGGTCTATCTTCATATGCACCACCACCAC CACATGGAGATTAACGGGGCAC	서열번호 13
BHB 22-R	ATGCGCGCCGCTTATCAACATATTGGAATGGGAGCTCC	서열번호 14
BHB 31-F	GCATTCTAGACCCAACGAAGGGGAACCACACAATATGCACCACCACCACCA CCACACCTTCGTTATTGAACTGGATATGAATGTTACCC	서열번호 15
BHB 31-R	ATGCGCGCCGCTTATCAGTGTGCTGCCACGGCTG	서열번호 16

[0086]

각각의 중합효소 연쇄반응 산물들을 제한효소 XbaI과 HindIII로 절단하고, 하기 표 2 및 도 2에 기재된 각 벡터에 라이게이션 (ligation) 하였다.

### 표 2

[0088]

Strain	Relevant Characteristics
<i>E. coli</i> MG1655	$F^- \lambda^- i\lambda^- rfb^- 50rph^- 1$
<i>E. coli</i> BL21(DE3)	$F^+ ompT gal dcm lon hsdS_B(r_B^- m_B^-) \lambda (DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])$
<i>E. coli</i> W3110	$F^- \lambda^- rph^- 1 INV(rrnD, rrnE)$

<i>E. coli</i> NST74	<i>E. coli</i> W3110 derivative ( <i>aroF aroG tyrR pheA pheAo</i> )
Plasmid	Relevant Characteristics
pET22b	Amp <sup>R</sup> , <i>T7</i> promoter
pTrc99a	Amp <sup>R</sup> , <i>trc</i> promoter
pHB-I01	pET22b derivative, His <sub>8</sub> -tag <i>A. thaliana</i> PAL
pHB-I02	pET22b derivative, His <sub>8</sub> -tag <i>A. thaliana</i> 4CL
pHB-I03	pET22b derivative, His <sub>8</sub> -tag <i>A. thaliana</i> CCR
pHB-I04	pET22b derivative, His <sub>8</sub> -tag <i>S. maritimus</i> PAL
pHB-I05	pET22b derivative, His <sub>8</sub> -tag <i>S. coelicolor</i> 4CL
pHB-P01	pTrc99a derivative, FLAG-tag <i>S. maritimus</i> PAL, His <sub>8</sub> -tag <i>S. coelicolor</i> 4CL, and His <sub>8</sub> -tag <i>A. thaliana</i> CCR

- [0090] 라이제이션에 완료된 pHB-I01, pHB-I02, pHB-I03, pHB-I04, 또는 pHB-I05 백터는 대장균 균주 MG1655을 거쳐 최종적으로 BL21(DE3)에 형질전환 시켰다. pHB-P01 백터는 대장균 균주 NST74에 형질전환 시켰다.
- [0092] **실시예 2. PAL, 4CL, CCR 단백질의 분리 정제**
- [0094] 대장균 균주 BL21(DE3)과 NST74 균주를 2% 포도당과 100 μg/mL 암피실린(ampicillin)을 포함하는 LB (Luria-Bertani) 배지에 접종하였다. 이를 37°C, 200 rpm 조건으로 12시간 동안 배양한 후, 신선한 LB 배지에 1/100 부피만큼 옮겼다. 그 후 동일한 조건으로 OD<sub>600</sub>이 0.6에 도달할 때까지 배양하였다.
- [0095] 그 후 BL21(DE3) 균주는 단백질의 생산을 위해 25°C, 200 rpm 조건으로 30분 간 적응시킨 후 1 mM의 IPTG (isopropyl-β-D-thiogalactopyranoside)을 첨가하고 6시간 동안 추가로 배양하였다. 한편 NST74 균주는 37°C, 200 rpm 조건에서 1 mM IPTG를 첨가한 후 배양하였다.
- [0096] 그 후, 상기 배양액을 4°C, 6000 rpm 조건으로 10분 동안 원심분리하고 상등액을 제거하였다. 이를 버퍼(50 mM potassium phosphate, 300 mM sodium chloride, pH 7.0)로 재현탁 (resuspension)하고, 초음파 세포 파쇄기를 이용하여 파쇄하여 세포 혼탁액을 얻었다. 이후 상기 세포 혼탁액을 4°C, 10000 rpm 조건에서 10분 동안 원심분리하고, 수용성의 상층액 부분을 얻었다.
- [0098] 그 후, 상기 세포 혼탁액으로부터 IMAC (immobilized metal affinity chromatography)를 이용하여 8X 히스티딘 태그(Histidine tag)를 지니고 있는 PAL, 4CL, CCR 단백질들을 정제하였다.
- [0100] 또한, 상기 수용성의 상층액 부분을 0.45 μm으로 필터링한 후, 바인딩 버퍼 (binding buffer: 50 mM potassium phosphate, 300 mM sodium chloride, pH 7.0)로 전처리된 Talon®-금속 친화 수지에 가하였다. 이후, 상기 수지를 10 mL의 세척 버퍼 (washing buffer: 50 mM potassium phosphate, 300 mM sodium chloride, 15 mM imidazole, pH 7.0)로 세척한 후, 1 mM의 일루션 버퍼 (elution buffer: 50 mM potassium phosphate, 300 mM sodium chloride, 150 mM imidazole, pH 7.0)를 이용하여 최종적으로 단백질들을 정제하였다.
- [0102] **실험예 1. PAL, 4CL, CCR 단백질의 정량**
- [0104] 상기 정제된 단백질을 12% (w/v) SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis)를 이용한 텐시토메트리 (densitometry) 방법과 BSA (bovine serum albumin)를 이용한 총 단백질 정량법을 통해 정량하여 도 3에 나타내었다.
- [0106] 도 3에서 레인 M은 단백질 크기 표시이고, 레인 1부터 3까지는 애기장대 PAL 효소의 전체(total) 부분, 수용성 (soluble) 부분, 정제 후 용리(elution) 부분이고, 레인 4부터 6까지는 애기장대 4CL 효소의 total, soluble, 정제 후 elution, 레인 7부터 9까지는 애기장대 CCR 효소의 total, soluble, 정제 후 elution, 레인 10부터 12까지는 스트렙토마이세스 마리티무스 PAL 효소의 total, soluble, 정제 후 elution 부분, 그리고 레인 13부터 15까지는 스트렙토마이세스 실리칼라 4CL 효소의 total, soluble, 정제 후 elution 부분을 나타낸 것이다.
- [0107] 각 단백질의 크기는 애기장대 PAL은 78 kDa (solid arrow), 애기장대 4CL은 61kDa (dashed arrow), 애기장대 CCR은 37kDa (closed triangle), 스트렙토마이세스 마리티무스 PAL은 56kDa (open triangle), 스트렙토마이세스 실리칼라 4CL은 55kDa (open triangle)으로 나타나 PAL, 4CL, CCR 단백질이 각각 잘 생성되었음을 확인하였

다.

## [0109] 실험예 2. 효소의 활성 분석

[0111] 상기 실시예 2에서 정제한 PAL, 4CL, CCR 단백질들의 활성은 세포 외 반응을 통해 분석하였다.

[0112] 우선, PAL의 활성 분석을 위하여 100 mM Tris-HCl, 0.2 mM 페닐알라닌, 정제한 PAL 효소 200  $\mu\text{g}/\text{mL}$ 를 반응시켰다. 또한, 4CL 및 CCR의 활성 분석을 위하여 400 mM Tris-HCl, 5 mM ATP, 5 mM 염화마그네슘 (magnesium chloride), 0.3 mM 코엔자임 A (Coenzyme A), 0.5 mM 트랜스-신남산(*trans*-cinnamate), 및 정제한 4CL 및 CCR 효소를 각각 50  $\mu\text{g}/\text{mL}$ 씩 반응시켰다.

[0113] 상기 반응을 30 °C에서 1시간 동안 진행한 후, PAL 효소에 의해 생성된 신남산, 그리고 4CL과 CCR 효소에 의해 생성된 신남알데하이드(cinnamaldehyde)를 역상 고성능 액체 크로마토그래피 (reverse-phase high-performance liquid chromatography, reverse-phase HPLC)를 통해 분석하였다. 구체적으로, ZORBAX Eclipse AAA 컬럼 (150 X 4.6 mm; 3.5  $\mu\text{m}$ ; Agilent, CA, USA)에서, 이동상 A (mobile phase A)는 0.1% 트리플루오로아세트산 (trifluoroacetic acid), 이동상 B (mobile phase B)는 아세토나이트릴 (acetonitrile)을 사용하였다. 아세토나이트릴 (acetonitrile)의 비율은 처음 1분 간은 10%, 이후 9분 간 70%로 점차적으로 변경하였다. 컬럼의 온도는 40°C, 유속은 1 mL/min으로 고정하였다.

[0114] 표준용액 (standard solution; 1, 10, 50, 100, 200 mg/L cinnamate와 0.41, 4.1, 41, 82, 136.7 mg/L cinnamaldehyde)을 사용하여 표준곡선 (standard curve)을 작성하였으며, 이를 이용하여 최종적으로 생성된 신남산(cinnamate, 도 4a에 CA로 표기)과 신남알데하이드(cinnamaldehyde, 도 4b에 CAD로 표기)를 정량하였다.

[0116] 그 결과, 도 4a에 나타난 것과 같이, 애기장대 유래의 PAL에 비하여 스트렙토마이세스 마리티무스 (*Streptomyces maritimus*) 유래의 PAL의 활성이 우수한 것으로 나타났다. 또한, 도 4b에 나타난 것과 같이, 스트렙토마이세스 마리티무스 유래의 4CL과 애기장대 유래의 CCR의 조합을 사용할 경우, 애기장대 유래의 4CL과 애기장대 유래의 CCR의 조합에 비하여 효소 활성이 우수한 것으로 나타났다.

## [0118] 실험예 3. 세포 고농도 배양 (High-cell density cultivation)

[0120] 먼저, R/2 semi-defined 배지 (6.75 g/L potassium dihydrogen phosphate, 2 g/L ammonium phosphate dibasic, 0.85 g/L citric acid, 0.7 g/L magnesium sulfate heptahydrate, 5 mL/L trace metal solution (TMS; 10 g/L iron sulfate heptahydrate, 2.2 g/L zinc sulfate heptahydrate, 2 g/L calcium chloride dehydrate, 1 g/L copper sulfate pentahydrate, 0.58 g/L manganese sulfate pentahydrate, 0.1 g/L ammonium heptamolybdate tetrahydrate, 0.02 g/L sodium tetraborate decahydrate, pH 6.8)에 적응된, pHB-P01을 포함하는 대장균 균주 NST74를 200 mL 양으로 37°C, 200 rpm에서 12시간 동안 배양하였다.

[0121] 그 후, 동일한 배지 1.8 L에 접종하여 총 2 L 양으로 5L 크기의 생물 반응기에서 고농도 배양하였다. pH가 6.77 보다 낮아지는 경우에는 50% (v/v) 암모니아 (ammonia)를, 6.86보다 높아지는 경우에는 먹이 배지 (feeding solution)를 공급하였다. 먹이 배지로는 복합 배지 1 (complex feeding solution 1; 500 g/L glucose, 75 g/L yeast extract, 20 g/L magnesium sulfate heptahydrate), 복합 배지 2 (complex feeding solution 2; 500 g/L glucose, 100 g/L 카사미노산(casamino acid), 20 g/L magnesium sulfate heptahydrate), 제한 배지 1 (defined feeding solution; 700 g/L glucose, 20 g/L magnesium sulfate heptahydrate), 또는 제한 배지 2 (defined feeding solution; 500 g/L glucose, 20 g/L magnesium sulfate heptahydrate, 0.81 g/L phenylalanine)를 사용하였다.

[0122] 온도는 37°C로 유지시켰고, 용존 산소 (DO)는 휘저음 속도 (agitation rate)를 1000 rpm까지 상승시킨 후, 산소를 공급하여 40%로 유지시켰다. 광학 밀도(optical density, OD<sub>600</sub>)가 60에 도달하게 되면, 단백질 (효소)을 생성시키기 위하여 IPTG 1 mM를 공급하였다. 거품 억제제는 멸균 후 수동으로 공급하였다.

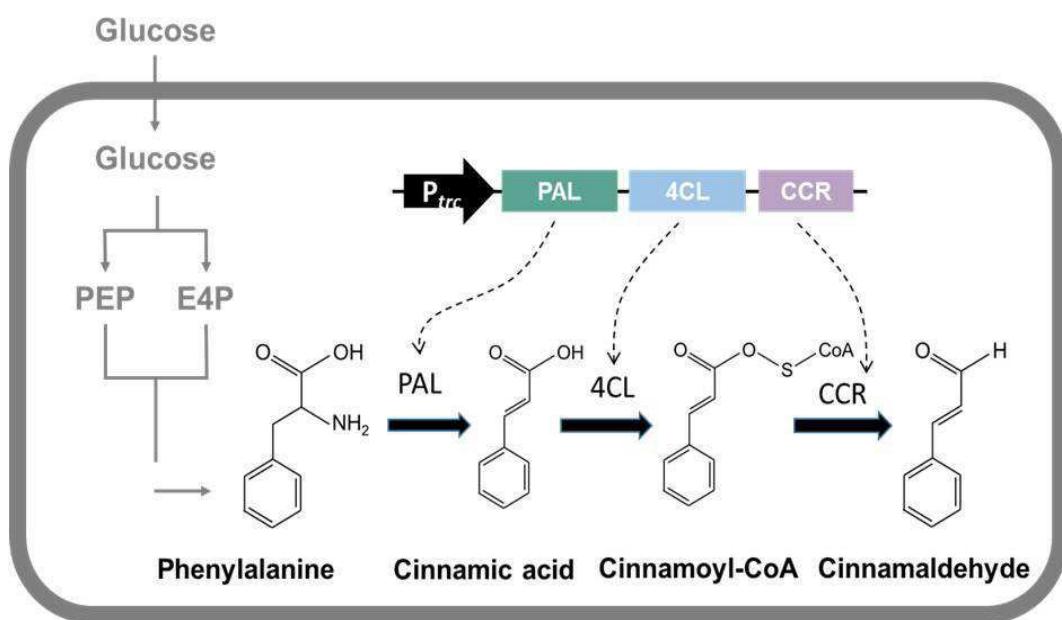
[0123] 배양한 대장균 균주는 시간대 별로 광학 밀도 (OD<sub>600</sub>)가 4가 되도록 따로 모아두어 단백질 분석에 이용하였다. 배양액의 상등액은 0.22  $\mu\text{m}$  필터링 (filter)하여 역상 고성능 액체 크로마토그래피 (reverse-phase HPLC) 분석을 통해 신남알데하이드 정량에 이용하였다.

[0125] 도 5는 상기 먹이 배지로 복합 배지 1을 사용한 후, 시간대 별로 측정한 광학 밀도 (OD<sub>600</sub>; open circle)와 이에 따른 신남알데하이드의 농도 (CAD con.; closed circle)를 나타낸다. 최고 생산 수율은 1.23 mg/L로 나타났다.

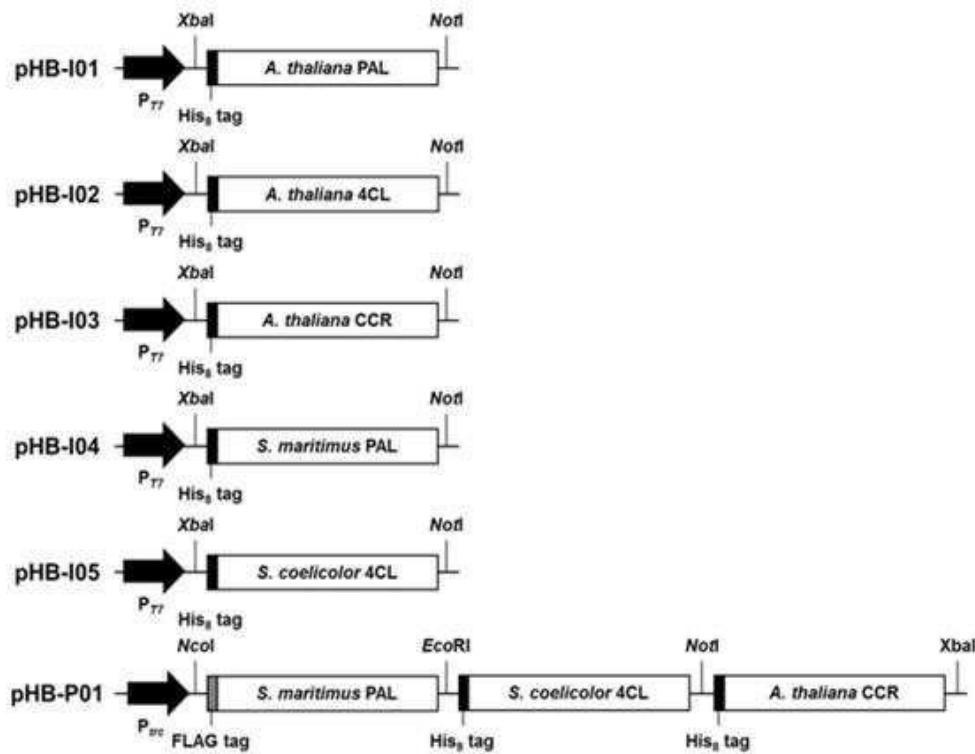
- [0127] 도 6은 상기 벽이 배지로 복합 배지 2를 사용한 후, 시간대 별로 측정한 광학 밀도 (OD600; open circle)와 이에 따른 신남알데하이드의 농도 (CAD con.; closed circle)를 나타낸다. 최고 생산 수율은 1.98 mg/L로 나타났다.
- [0129] 도 7은 상기 벽이 배지로 제한 배지 1을 사용한 후, 시간대 별로 측정한 광학 밀도 (OD600; open circle)와 이에 따른 신남알데하이드의 농도 (CAD con.; closed circle)를 나타낸다. 최고 생산 수율은 0.71 mg/L로 나타났다.
- [0131] 도 8은 상기 벽이 배지로 제한 배지 2를 사용한 후, 시간대 별로 측정한 광학 밀도 (OD600; open circle)와 이에 따른 신남알데하이드의 농도 (CAD con.; closed circle)를 나타낸다. 최고 생산 수율은 1.05 mg/L로 나타났다.
- [0133] 이상의 설명으로부터, 본 발명이 속하는 기술분야의 당업자는 본 발명이 그 기술적 사상이나 필수적 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 이와 관련하여, 이상에서 기술한 실시 예들은 모든 면에서 예시적인 것이며 한정적인 것이 아닌 것으로서 이해해야만 한다. 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특히 청구범위의 의미 및 범위 그리고 그 등가 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

## 도면

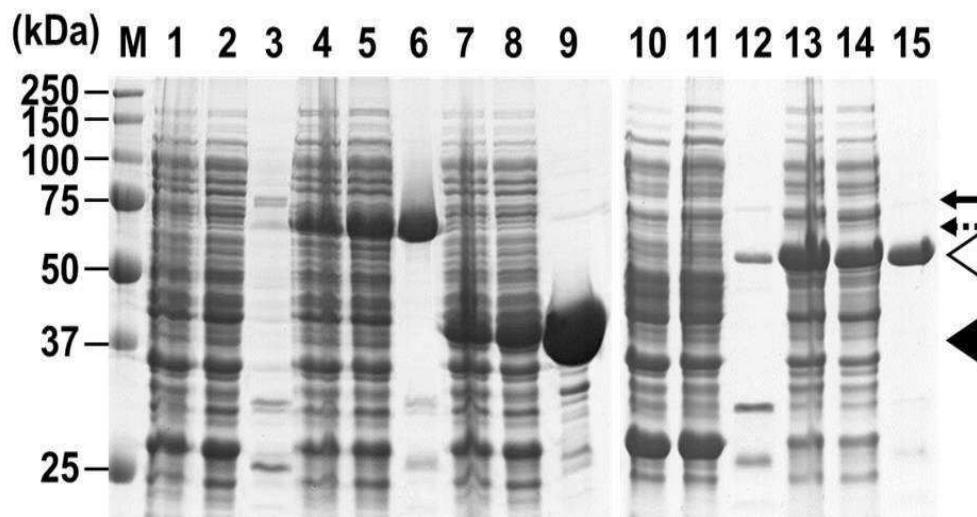
### 도면1



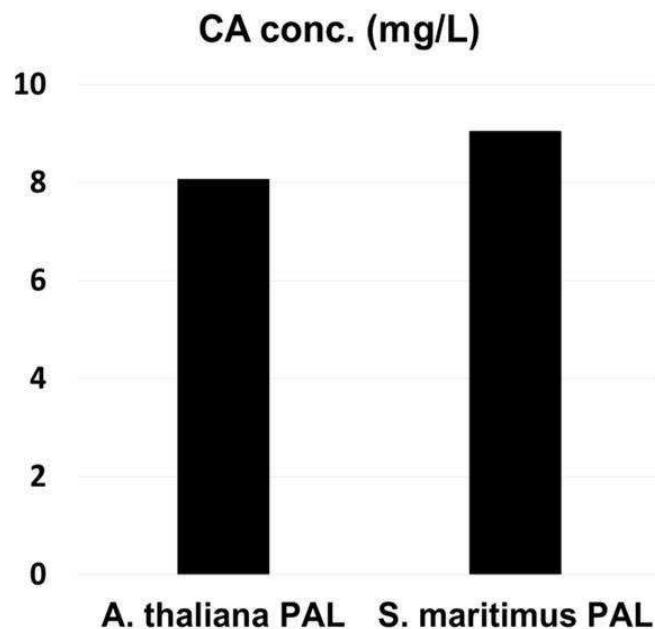
## 도면2



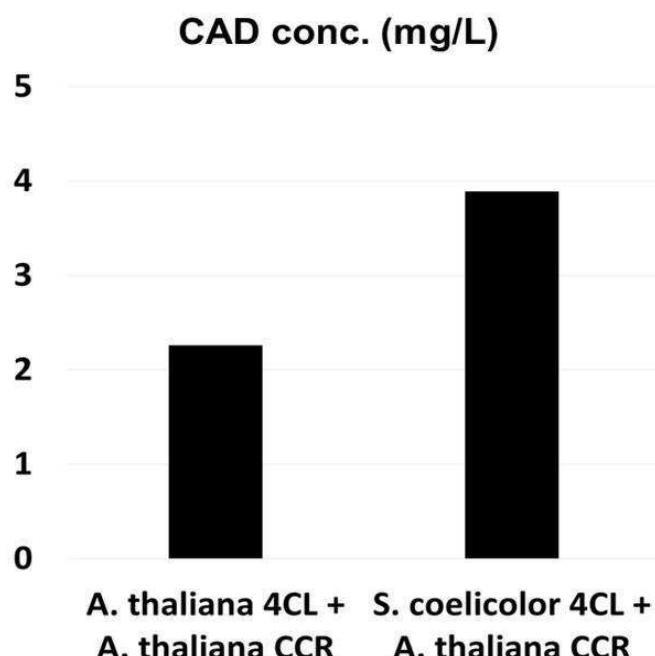
## 도면3



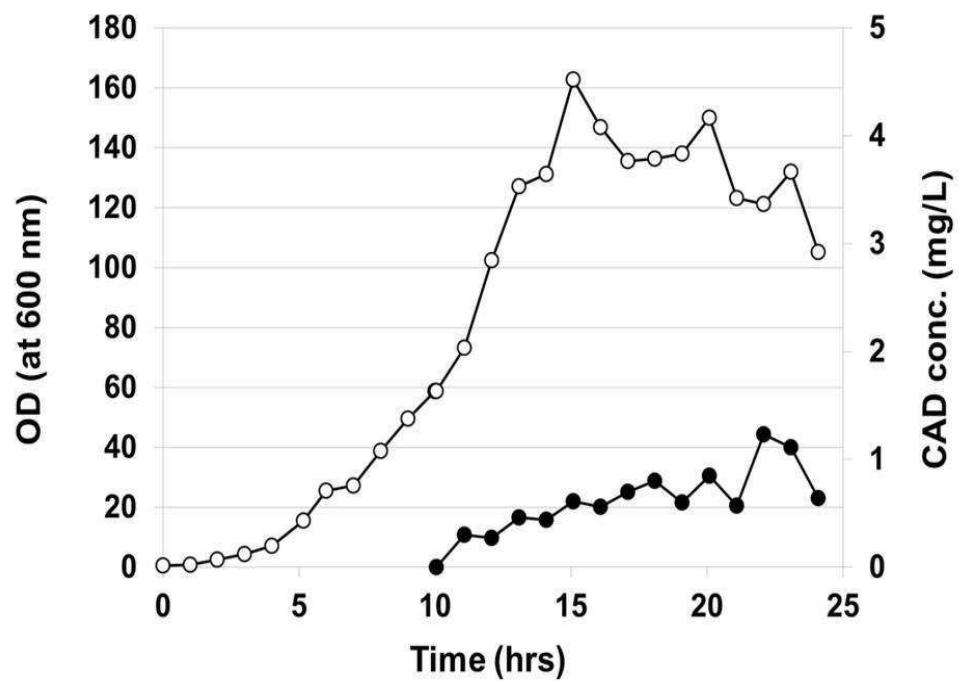
도면4a



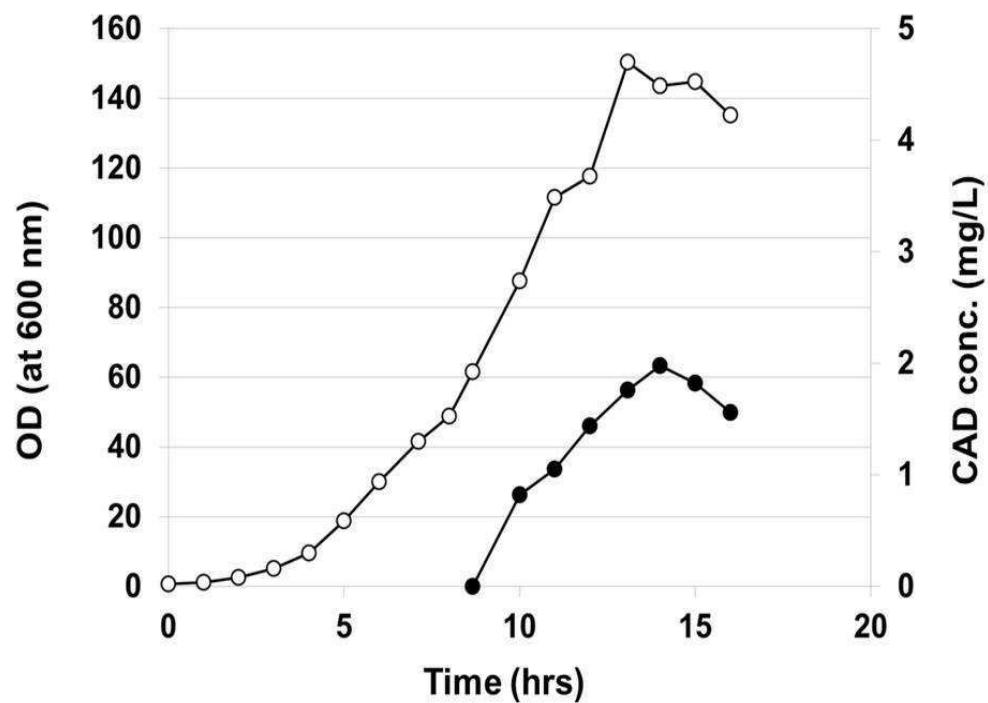
도면4b



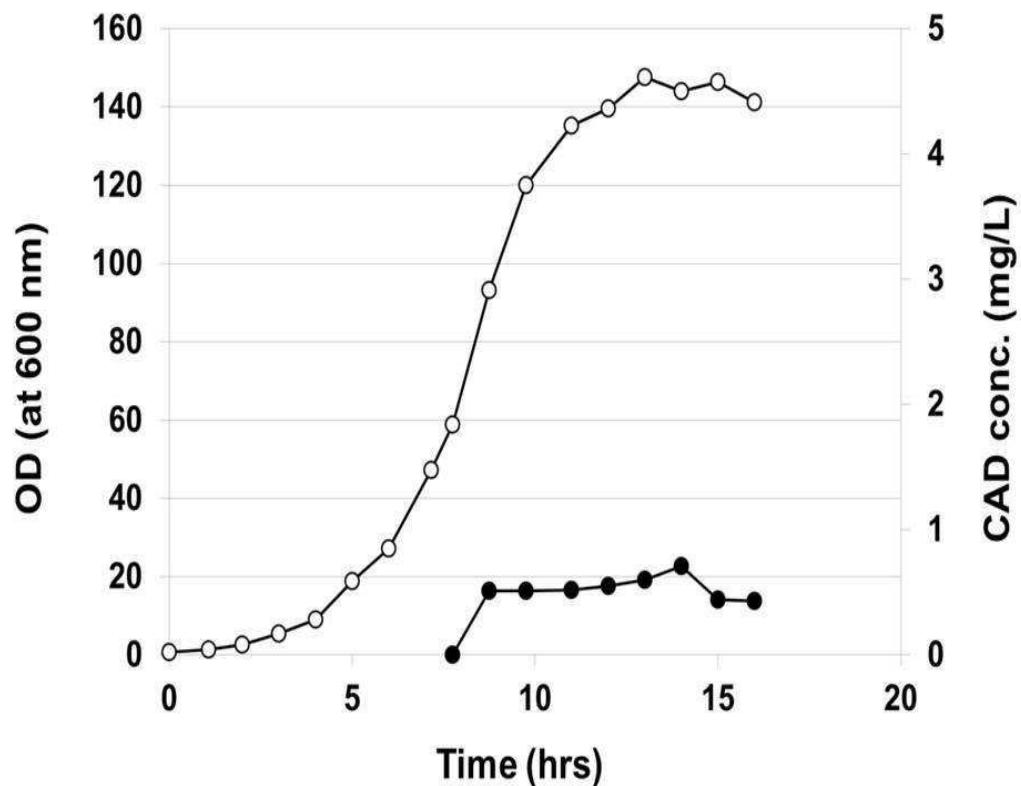
도면5



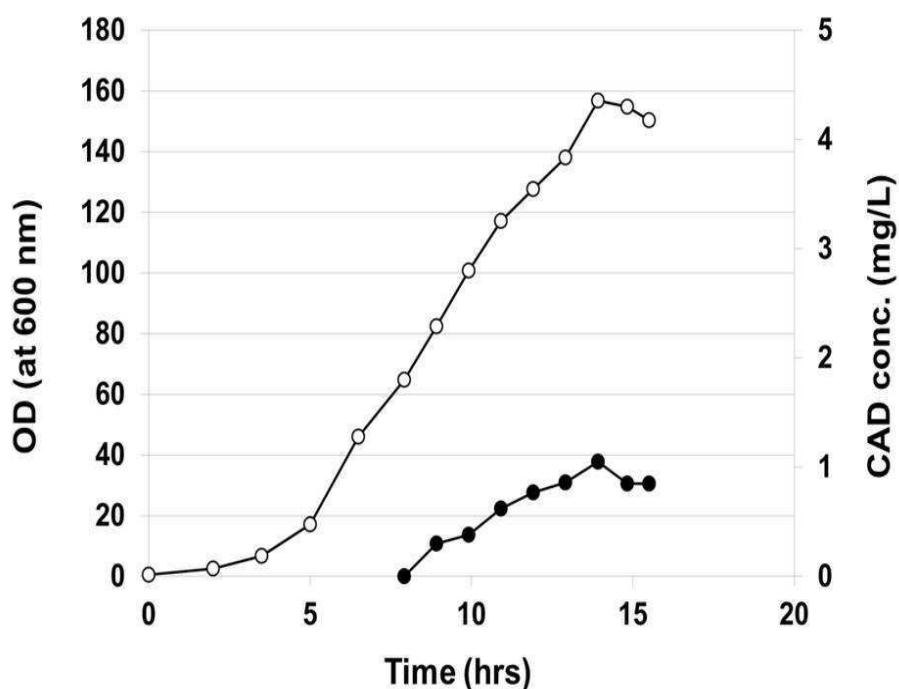
도면6



## 도면7



## 도면8



## 서 열 목 록

&lt;110&gt; Intelligent Synthetic Biology Center

Korea Advanced Institute of Science and Technology

<120> A method for production of cinnamaldehyde

<130> KPA150161-KR

<160> 19

<170> Kopatent In 2.0

<210> 1

<211> 1569

<212> DNA

<213> Streptomyces maritimus

<400> 1

atgaccttcg tcatagagct cgacatgaac gtcacgctcg accaacttga ggacgcggcg	60
cgacagcgca cggccgtgaa gctgtccgca cccgtccgct cccgcgtccg cgcctcgcg	120
gacgtgttgg tgaagttcgt gcaggaccaa cgtgtcatct acggggtaa caccagcatg	180

gggggcttcg tcgaccacct cgtcccggtg tcccaggccc ggcagctcca ggagaacctg	240
atcaacgcgg tcgccaccaa cgtggggcg tatctggacg acacgaccgc ccggaccatc	300
atgctgtccc gcatcgtgtc gctggcgcgc gggaaactccg ccatcaccccc ggcgaatctg	360
gacaagctgg tggccgtact caacgcccggg atcgtgccgt gcatcccgaa gaagggtct	420
ttggccacca gcggtgacct cggcccgctg gccgcgatcg ccctgggtgt cgccgggcag	480
tggaaggccc gctacaacgg tcaagatcatg cccggccggc aggcctgtc cgaggccggc	540
gtcgagccga tggagcttag ctacaaggat ggcctggccc tgcataacgg cacgtcaggc	600

atggtcggcc tgggcaccat ggtcctccag gccgcgcgcc ggctcgtgaa ccgcgtacctg	660
caggtgtccg cgttgtcggt cgagggcctg gcaggcatga cggaaaccgtt cgaccctcg	720
gtgcacggcg tcaagccgca cccggcccgag cgtcagggtgg cctcgccgtt gtgggagggg	780
cttgcgcact cgcacctggc ggtcaacgaa ctggacaccg agcagaccct ggccggagag	840
atgggcacgg tcgccaaggc cggttcgtcg ggcgtcgagg acgcctactc catccgtgc	900
acgcgcaga tcctcggtcc cgtggtcgtat gtgcgtggacc ggatcggggc gaccctgcag	960
gacgagctga actcctccaa cgacaacccg atcgtcctgc cggaggaggc ggagggtttc	1020

cacaacgggc acttccacgg ccagtacgtg gccatggcca tggaccacct gaacatggcc	1080
ctggccaccc tgaccaatct cggcaacccgg cgcgtggacc gcttcctgga caagagcaac	1140
agcaacgggc tgccgcctt cctgtgccgg gaagatccgg gactgcgcct gggcctgatg	1200
ggcggccact tcatgaccgc gtcgtaccgc gcccggacacc gcaccctgac cattccgtat	1260
tccgtgcact ccctcacgag tacggcggac ttccaggaca tcgttcctt cggattcg	1320
gcccggccccc gcccgggaa ggtactcacc aacgctgcct acgtggtggc cttcgagctg	1380

ctgtgcgcct gccaggccgt cgacatccgc ggcgcccaca aactgtcctc cttcacccgc 1440

ccgctctatg agcgcacccg caagatcgtg ccgttcttcg accggacga gaccatcacc 1500

gactacgtcg agaagctggc ggccgacctg atcgccggcg agcccgtcga cgctgccgtg 1560

gcggcgcac 1569

<210> 2

<211> 523

<212> PRT

<213> Streptomyces maritimus

<400> 2

Met Thr Phe Val Ile Glu Leu Asp Met Asn Val Thr Leu Asp Gln Leu

1 5 10 15

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

20 25 30

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

35 40 45

Asp Glu Arg Val Ile Tyr Gly Val Asn Thr Ser Met Gly Gly Phe Val

50 55 60

Asp His Leu Val Pro Val Ser Gln Ala Arg Gln Leu Gln Glu Asn Leu

65 70 75 80

Ile Asn Ala Val Ala Thr Asn Val Gly Ala Tyr Leu Asp Asp Thr Thr

85 90 95

Ala Arg Thr Ile Met Leu Ser Arg Ile Val Ser Leu Ala Arg Gly Asn

100 105 110

Ser Ala Ile Thr Pro Ala Asn Leu Asp Lys Leu Val Ala Val Leu Asn

115 120 125

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

130 135 140

Gly Asp Leu Gly Pro Leu Ala Ala Ile Ala Leu Val Cys Ala Gly Gln

145 150 155 160

Trp Lys Ala Arg Tyr Asn Gly Gln Ile Met Pro Gly Arg Gln Ala Leu

165 170 175

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

Asp Ile Val Ser Phe Gly Phe Val Ala Ala Arg Arg Ala Arg Glu Val

435 440 445

Leu Thr Asn Ala Ala Tyr Val Val Ala Phe Glu Leu Leu Cys Ala Cys

450 455 460

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

465 470 475 480

Pro Leu Tyr Glu Arg Thr Arg Lys Ile Val Pro Phe Phe Asp Arg Asp

485 490 495

Glu Thr Ile Thr Asp Tyr Val Glu Lys Leu Ala Ala Asp Leu Ile Ala

500 505 510

Gly Glu Pro Val Asp Ala Ala Val Ala Ala His

515 520

<210> 3

<211> 1566

<212> DNA

<213> Streptomyces coelicolor

<400> 3

atgttccgca gcgagtagcgc agacgtcccg cccgtcgacc tgcccatcca cgacgccgtg	60
ctcgccgggg ccgcgcctt cgggagcacc cggcgctga tcgacggcac cgacggcacc	120
accctcacct acgaggcagggt ggaccggttc cacccggcgcg tgcggccgc cctggccgag	180
accggcgtgc gcaaggggcga cgtcctcgcc ctgcacagcc ccaacaccgt cgcctcccc	240
ctggccttct acgeccacac ccgcgcgggc gcctccgtca ccacggtgca tccgctcgcg	300
acggcggagg agttcgccaa gcagctgaag gacagcgcgg cccgctggat cgtcaccgtc	360

tcaccgctcc tgcaccgcgc cggccgggcc gccgaactcg cggcggcgt ccaggagatc	420
ctggctcgac acagcgcgcc cggtcaccgc tccctcgat acatgtcgcc ctgcaccgc	480
cccgaaaccgt ccgtcgccat cgaccggcc gaggacgtcg cggccctgcc gtactctcg	540
ggcaccacccg gcaccccaa gggcgtcatg ctcacacacc ggcagatcgc caccaaccc	600
gcccagctcg aaccgtcgat gccgtccgcg cccggcacc gcttcctcgc cgtgctgccg	660
ttcttccaca tctacggctt gaccggctt atgaacgccc cgtccggct cggccacc	720
gtcggttcc tgcccgctt cgacctggag cagttctcg cggccatcca gaaccaccgc	780

atcaccagcc tgtacgtcgc cccggcgate gtcctggccc tcgccaaaca cccctggtc	840
--	-----

gccgactacg acctctcctc gctgaggtac atcgtagcg ccggccccc gctcgacgcg	900
cgtctcgccg ccgcctgctc gcagcggctc ggcctgccg ccgtcgccca ggcctacggc	960
atgaccgaac tgtccccggg cacccacgtc gtccccctgg acgcgatggc cgacgcgccc	1020
cccgacccg tcggcaggct catcgccgg accgagatgc gcatcgctc cctcaccgac	1080
ccgggcacgg acctccccgc cggagagtcc gggagatcc tcatccggg ccccccagatc	1140
atgaaggcgt acctggccg ccccgacgcc accgcggcca tgatcgacga ggagggtgg	1200

ctgcacaccc gggacgtcg acacgtcgac gccgacggct ggctgttcgt cgtcgaccgc	1260
gtcaaggaac tgaatcaagta caagggcttc caggtggccc ccggcgaact ggaggcccac	1320
ctgctcaccc accccggcgt cgccgacgcg gccgtcgctg ggcctacga cgacgcacggc	1380
aacgaggtac cgacgcctt cgtcgccgc cagccggccg caccggcct cgccggagagc	1440
gagatcatga tgtacgtcg cgaacgcgtc gccccata aacgcgtccg ccgggtcacc	1500
ttcgtcgacg ccgtcccccg cgccgcctcc ggcaagatcc tccggcaca gctcagggag	1560
ccgcga	1566

<210> 4  
<211> 522  
<212> PRT  
<213> Streptomyces coelicolor  
<400> 4

Met Phe Arg Ser Glu Tyr Ala Asp Val Pro Pro Val Asp Leu Pro Ile

1	5	10	15
---	---	----	----

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

20	25	30
----	----	----

Leu Ile Asp Gly Thr Asp Gly Thr Thr Leu Thr Tyr Glu Gln Val Asp

35	40	45
----	----	----

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

50	55	60
----	----	----

Lys Gly Asp Val Leu Ala Leu His Ser Pro Asn Thr Val Ala Phe Pro

65	70	75	80
----	----	----	----

Leu Ala Phe Tyr Ala Ala Thr Arg Ala Gly Ala Ser Val Thr Thr Val

85	90	95
----	----	----

His Pro Leu Ala Thr Ala Glu Glu Phe Ala Lys Gln Leu Lys Asp Ser

100	105	110
-----	-----	-----

Ala Ala Arg Trp Ile Val Thr Val Ser Pro Leu Leu Ser Thr Ala Arg  
 115 120 125  
  
 Arg Ala Ala Glu Leu Ala Gly Gly Val Gln Glu Ile Leu Val Cys Asp  
 130 135 140  
 Ser Ala Pro Gly His Arg Ser Leu Val Asp Met Leu Ala Ser Thr Ala  
 145 150 155 160  
 Pro Glu Pro Ser Val Ala Ile Asp Pro Ala Glu Asp Val Ala Ala Leu  
 165 170 175  
 Pro Tyr Ser Ser Gly Thr Thr Gly Thr Pro Lys Gly Val Met Leu Thr  
 180 185 190  
 His Arg Gln Ile Ala Thr Asn Leu Ala Gln Leu Glu Pro Ser Met Pro  
  
 195 200 205  
 Ser Ala Pro Gly Asp Arg Val Leu Ala Val Leu Pro Phe Phe His Ile  
 210 215 220  
 Tyr Gly Leu Thr Ala Leu Met Asn Ala Pro Leu Arg Leu Gly Ala Thr  
 225 230 235 240  
 Val Val Val Leu Pro Arg Phe Asp Leu Glu Gln Phe Leu Ala Ala Ile  
 245 250 255  
 Gln Asn His Arg Ile Thr Ser Leu Tyr Val Ala Pro Pro Ile Val Leu  
 260 265 270  
  
 Ala Leu Ala Lys His Pro Leu Val Ala Asp Tyr Asp Leu Ser Ser Leu  
 275 280 285  
 Arg Tyr Ile Val Ser Ala Ala Ala Pro Leu Asp Ala Arg Leu Ala Ala  
 290 295 300  
 Ala Cys Ser Gln Arg Leu Gly Leu Pro Pro Val Gly Gln Ala Tyr Gly  
 305 310 315 320  
 Met Thr Glu Leu Ser Pro Gly Thr His Val Val Pro Leu Asp Ala Met  
 325 330 335  
 Ala Asp Ala Pro Pro Gly Thr Val Gly Arg Leu Ile Ala Gly Thr Glu  
  
 340 345 350  
 Met Arg Ile Val Ser Leu Thr Asp Pro Gly Thr Asp Leu Pro Ala Gly

355	360	365
Glu Ser Gly Glu Ile Leu Ile Arg Gly Pro Gln Ile Met Lys Gly Tyr		
370	375	380
Leu Gly Arg Pro Asp Ala Thr Ala Ala Met Ile Asp Glu Glu Gly Trp		
385	390	395
Leu His Thr Gly Asp Val Gly His Val Asp Ala Asp Gly Trp Leu Phe		
405	410	415
Val Val Asp Arg Val Lys Glu Leu Ile Lys Tyr Lys Gly Phe Gln Val		
420	425	430
Ala Pro Ala Glu Leu Glu Ala His Leu Leu Thr His Pro Gly Val Ala		
435	440	445
Asp Ala Ala Val Val Gly Ala Tyr Asp Asp Asp Gly Asn Glu Val Pro		
450	455	460
His Ala Phe Val Val Arg Gln Pro Ala Ala Pro Gly Leu Ala Glu Ser		
465	470	475
Glu Ile Met Met Tyr Val Ala Glu Arg Val Ala Pro Tyr Lys Arg Val		
485	490	495
Arg Arg Val Thr Phe Val Asp Ala Val Pro Arg Ala Ala Ser Gly Lys		
500	505	510
Ile Leu Arg Arg Gln Leu Arg Glu Pro Arg		
515	520	
<210>	5	
<211>	1032	
<212>	DNA	
<213>	Arabidopsis thaliana	
<400>	5	
atgccagtcg acgttagcctc accggccgga aaaaccgtct gcgtcaccgg agctggtgga		60
tacatcgctt cttggattgt taagatactt ctcgagagag gttacacagt caaaggaacc		120
gtacggaatc cagatgatcc gaagaacaca catttggagag aactagaagg agggaaaggag		180
agactgattc tgtgcaaagc agatcttcag gactacgagg ctcttaaggc ggcgattgat		240
ggttgcgacg gcgctttca cacggcttct cctgtcaccc acgatccgga acaaatggtg		300
gagccggccg tgaatggagc caagttgtta attaatgctg cggctgaggc caaggtcaag		360

cgcggtggtca	tcacccctc	cattgggtgcc	gtctacatgg	acccgaaccg	tgaccctgag	420
gctgtcgttg	acgaaagtgtt	ttggaggttat	cttgacttct	gcaaaaacac	caagaatttgg	480
tatttgttacg	gcaagatggtt	ggcggaaacaa	gcggcgtggg	agacagcaaa	ggagaaaggta	540
gttgacttgg	tgggtttgaa	tccgggtgctg	gttcttgac	cgccgttaca	gccgacgatc	600
aacgccagtc	tttaccacgt	cctcaaataat	ctaaccggct	cggttaagac	ttatgctaat	660
ttgactcaag	cttatgtgga	tgttcgcgtat	gtcgcgtgg	ctcatgttct	ggtctatgag	720
gcaccctcg	cctccggacg	ttatctccta	gccgagagtgc	ctcgccaccg	cggggaaagt	780
gttgagattc	tggctaagct	attcccgag	tatcctcttc	cgaccaagtgc	caaggacgag	840
aagaacctta	gagccaagcc	atacaaattt	actaaccaga	agattaagga	cttaggctta	900
gagttcactt	ccaccaagca	aaggctctac	gacacagtca	agagcttaca	agagaaaggc	960
catcttgctc	ctccctctcc	tcctccttca	gcatcgcaag	aatccgtgga	aatggcatt	1020
aagatcggtt	ct					1032
<210>	6					
<211>	344					
<212>	PRT					
<213>	Arabidopsis thaliana					
<400>	6					
Met Pro Val Asp Val Ala Ser Pro Ala Gly Lys Thr Val Cys Val Thr						
1	5	10	15			
Gly Ala Gly Gly Tyr Ile Ala Ser Trp Ile Val Lys Ile Leu Leu Glu						
20	25	30				
Arg Gly Tyr Thr Val Lys Gly Thr Val Arg Asn Pro Asp Asp Pro Lys						
35	40	45				
Asn Thr His Leu Arg Glu Leu Glu Gly Gly Lys Glu Arg Leu Ile Leu						
50	55	60				
Cys Lys Ala Asp Leu Gln Asp Tyr Glu Ala Leu Lys Ala Ala Ile Asp						
65	70	75	80			
Gly Cys Asp Gly Val Phe His Thr Ala Ser Pro Val Thr Asp Asp Pro						
85	90	95				
Glu Gln Met Val Glu Pro Ala Val Asn Gly Ala Lys Phe Val Ile Asn						
100	105	110				

Ala Ala Ala Glu Ala Lys Val Lys Arg Val Val Ile Thr Ser Ser Ile

115 120 125

Gly Ala Val Tyr Met Asp Pro Asn Arg Asp Pro Glu Ala Val Val Asp

130 135 140

Glu Ser Cys Trp Ser Asp Leu Asp Phe Cys Lys Asn Thr Lys Asn Trp

145 150 155 160

Tyr Cys Tyr Gly Lys Met Val Ala Glu Gln Ala Ala Trp Glu Thr Ala

165 170 175

Lys Glu Lys Gly Val Asp Leu Val Val Leu Asn Pro Val Leu Val Leu

180 185 190

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

195 200 205

Lys Tyr Leu Thr Gly Ser Ala Lys Thr Tyr Ala Asn Leu Thr Gln Ala

210 215 220

Tyr Val Asp Val Arg Asp Val Ala Leu Ala His Val Leu Val Tyr Glu

225 230 235 240

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

245 250 255

Arg Gly Glu Val Val Glu Ile Leu Ala Lys Leu Phe Pro Glu Tyr Pro

260 265 270

Leu Pro Thr Lys Cys Lys Asp Glu Lys Asn Pro Arg Ala Lys Pro Tyr

275 280 285

Lys Phe Thr Asn Gln Lys Ile Lys Asp Leu Gly Leu Glu Phe Thr Ser

290 295 300

Thr Lys Gln Ser Leu Tyr Asp Thr Val Lys Ser Leu Gln Glu Lys Gly

305 310 315 320

His Leu Ala Pro Pro Pro Pro Pro Ser Ala Ser Gln Glu Ser Val

325 330 335

Glu Asn Gly Ile Lys Ile Gly Ser

340

<210> 7

<211> 78

<212> DNA

<213> Artificial Sequence

<220><223> BHB 19-F

<400> 7

gcacatctagaa acacaaacaa ggaaggaaga taaatgcacc accaccacca ccaccacac	60
--	----

atgccagtcg acgttagcc	78
----------------------	----

<210> 8

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> BHB 19-R

<400> 8

atgcgcggcc gcttatcaag acccgatctt aatgccattt tc	42
--	----

<210> 9

<211> 80

<212> DNA

<213> Artificial Sequence

<220><223> BHB 20-F

<400> 9

gcacatctagac cgaatcaaa aggaacacca acgtatgcac caccaccacc accaccacca	60
--	----

catggcgcga caagaacaag	80
-----------------------	----

<210> 10

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> BHB 20-R

<400> 10

atgcgcggcc gcttatcaca atccatttgc tagtttgcc c	41
--	----

<210> 11

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> BHB 21-F

<400>	11	
atgcgcggcc gcttatcaca atccattgc tagtttgcc c		41
<210>	12	
<211>	35	
<212>	DNA	
<213>	Artificial Sequence	
<220><223> BHB 21-R		
<400>	12	
atgcgcggcc gcttatcatc gcggctccct gagct		35
<210>	13	
<211>	80	
<212>	DNA	
<213>	Artificial Sequence	
<220><223> BHB 22-F		
<400>	13	
gcatactagac acctaagga ggtctatctt tcatatgcac caccaccacc accaccacca		60
catggagatt aacggggcac		80
<210>	14	
<211>	39	
<212>	DNA	
<213>	Artificial Sequence	
<220><223> BHB 22-R		
<400>	14	
atgcgcggcc gcttatcaac atattgaaat gggagctcc		39
<210>	15	
<211>	96	
<212>	DNA	
<213>	Artificial Sequence	
<220><223> BHB 31-F		
<400>	15	
gcattctaga cccaacgaag gggaaaccac acaaatgca ccaccaccac caccaccacc		60
acaccttcgt tattgaactg gatatgaatg ttaccc		96
<210>	16	
<211>	35	

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; BHB 31-R

&lt;400&gt; 16

atgcgcggcc gcttatcagt gtgctgccac ggctg 35

&lt;210&gt; 17

&lt;211&gt; 2175

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 17

atggagatta acggggcaca caagagcaac ggaggaggag tggacgctat gttatgccc 60

ggagacatca agacaaagaa catggtgatc aacgcggagg atcctctcaa ctggggagct 120

gcagcggagc aaatgaaagg tagccattt gatgaagtga agagaatggt tgctgagtt 180

aggaagccag ttgtgaatct tggtggtgag actctgacca ttggacaagt ggctgcgatc 240

tcaactattt gtaacagtgt gaaggtggag ctatcgaga cagctagagc cggtgtgaat 300

gcttagtagt attgggttat ggagagtatg aacaaaggca ctgatagtt tggtgttact 360

actggtttg gtgctacttc tcatcgaga accaaaaacg gtgtcgact tcagaaggaa 420

cttatttagt tccttaacgc cgaaatattc ggaagcacga aagaaacaag ccacacattt 480

ccacactccg ccacaagagc cgccatgtt gtacgaatca acactctctt ccaaggattt 540

tccggtatcc gatttgagat tctcgaaatca attaccattt tcctcaacaa caacatca 600

ccatctctcc ccctccgtgg tacaatcacc gcctccggag atctcggttcc tctctctac 660

atcggccggac ttctcaccgg tcgtccaaat tccaaagcta ctggtccaa cggtgaagct 720

ttaacacgag aggaagcttt caaatttagca ggaatcagct cggatttt tgatctccag 780

cctaaggaag gtctcgctt agtcaatggc acggcggtt gatctggat ggcgtcaatg 840

gtgttattcg aaacgaatgt tctctctt ttggctgaga tttgtcgcc ggtttcgca 900

gaggtgatga gtggtaagcc tggatccacc gatcatctca ctcacagact taaacatcat 960

cccggtcaaa tcaagcgcc ggcgataatg gagcatatcc tcgacggaag ctcgtacatg 1020

aaattagctc agaagcttca cgagatggat ccgttacaga aacctaaaca agatcggtac 1080

gctttcgta ctctcctca atggtaggt cctcaaatcg aagtgtatcc ttacgcaacg 1140

aaatcgatcg agcgtgagat taactccgtc aacgataatc cggtatcgatcgatcg 1200

aacaaggcga ttacgggtgg taacttccaa ggaacaccaa tcggagttt aatggataac 1260

acgagattgg cgatagcagc gattggtaaa ctcatgttg ctcaattctc agagtttg	1320
aatgatttct acaaacaatgg ttaccctcg aatctaaccg cttcgaggaa tccaaatgg	1380
gattatggat tcaagggagc tgagattgca atggcttctt attgttcaga gcttcaatac	1440
ttagctaatc ctgtgactag ccatgttcaa tcagcagagc aacataacca agatgtcaac	1500
tcttgggac taatctcgtc tcgcaaaact tctgaagctg ttgatattct caagcttatg	1560
tcaacaacgt tcctcggtgc gatttgtcaa gctgtggatt tgagacattt ggaggagaat	1620
ttgagacaga ctgtgaagaa cactgtctct caagtggcga agaaagttct tactactgga	1680
gtcaatggtg agttcatcc ttctcgcttc tgcgaaaagg atttactcaa agtttagac	1740
cgtgaacaag tctcacacata cgccgatgt ccttgtacg caacgtaccc gttgattcag	1800
aagctgagac aagttattgt tgaccatgct ttgatcaatg gtgagagtga gaagaatgca	1860
gtgacttcaa tctccataa gattggagct ttcgaggagg agcttaaggc agtgcgtaccg	1920
aaagaagtgg aagcagcaag agcagcctac gataacggaa catcgctat cccgaacagg	1980
atcaaggaat gtaggtcgta tccattgtat agattcgta gggaaagagct tggaacagag	2040
ctttgaccg gagagaaagt gacgtcgct ggagaagagt tcgacaaggt tttcacggcg	2100
atttgtgaag gtaaaatcat tgatccgatg atggaatgtc tcaacgagtg gaacggagct	2160
cccatccaa tatgt	2175
<210> 18	
<211> 1683	
<212> DNA	
<213> Arabidopsis thaliana	
<400> 18	
atggcgccac aagaacaagc agtttcttag gtgatggaga aacagagcaa caacaacaac	60
agtgacgtca tttccgatc aaagttaccc gatattaca tcccgaacca cctatctctc	120
cacgactaca tcttccaaaa catctccgag ttgccacta agccttgctt aatcaacggaa	180
ccaacccggcc acgtgtacac ttactccgac gtccacgtca tctccggcc aatcgccccc	240
aattttcaca aactcggcgt taacccaaac gacgtcgta tgctccctt cccaaactgt	300
cccgagttcg tcctctctt cctcgccgcc tccttcccg ggcacaccgc caccggccca	360
aaccctttct tcactccggc ggagatagct aaacaagccaa aagcctccaa caccaaactc	420
ataatcaccc aagtcgtta cgtcgacaaa atcaaaccac ttcaaaacga cgacggagta	480
gtcatcgct gcacgtacga caacgaatcc gtgcaatcc ctgaaggctg cctccgttcc	540
accgagttga cttagtcgac aaccgaggca tcagaagtcgatc tcgactcggt ggagattca	600

ccggacgacg tggtggact acttactcc tctggcacga cgggattacc aaaaggagtg	660
atgctgactc acaagggtct agtcacgacg gttgctcagc aagtcgacgg cgagaacctg	720
aatcttatt tccacagcga tgacgtata ctctgttt tgcccatgtt tcataatctac	780
gccttaact cgtcatgtt gtgtggctt agagttggc cggcgattct gataatgccg	840
aagtttgaga tcaatctgct attggagctg atccagaggt gtaaagtgac ggtggctcg	900
atggttccgc cgattgtgtt ggccattgcg aagtcttcgg agacggagaa gtatgatttgc	960
agctcgataa gagtgggtgaa atctgggtct gctcccttg gtaaagaact tgaagatgcc	1020
gttaatgcca agtttcccaa tgccaaactc ggtcaggat acggaatgac ggaagcaggt	1080
ccagtgctag caatgtcggtt aggttttgc aaggAACCTT ttccggtaa gtcaggagct	1140
tgtggtaactg ttgttaagaaa tgctgagatg aaaatagtttgc atccagacac cggagattct	1200
cTTTcgagga atcaacccgg tgagatttgtt attcgtggc accagatcat gaaaggttac	1260
ctcaacaatc cggcagctac agcagagacc attgataaag acggttggct tcatactgga	1320
gatattggat tgatcgatga cgatgacgag ctttcatcg ttgatcgatt gaaagaactt	1380
atcaagtata aaggTTTCA ggttagctcg gctgagctag aggcttgct catcggtcat	1440
cctgacatta ctgatgttgc tgggtcgca atgaaagaag aagcagctgg tgaagttcct	1500
gttgcatttgc tggtaatc gaaggattcg gagttatcg aagatgtatgtt gaagcaattc	1560
gtgtcgaaac agttgtgtt ttacaagaga atcaacaaag tggcttcac tgaatccatt	1620
cctaaagctc catcaggaa gatatttggaa aaagatctga gggcaaaact agcaaatttgc	1680
tttgc	1683
<210> 19	
<211> 1569	
<212> DNA	
<213> Artificial Sequence	
<220><223> phenylalanine ammonia lyase (codon optimization)	
<400> 19	
atgaccttcg ttattgaact ggatatgaat gttaccctgg accaacttgg agatgcggcc	60
cgtcagcgta ccccggtgaa actgtctgcc ccgggtcggtt cccgcgtgcg tgcctcacgt	120
gatgttctgg tcaaatttgc tcaaggacgaa cgcgtgatct atggcgtaa cacctcgatg	180
ggcggttgc tggatcatct ggtgccgggtt tcacaaggcgc gtcaatgcgca agaaaacctg	240
attaatgcgg cggcaacgaa tgtgggtgcc tacctggatg acaccacggc acgcaccatt	300
atgctgtcgc gtatcgtag cctggcgccg gcgcaacagcg ctatcagcc ggcgaatctg	360

gataaactgg tcgccgtgct gaacgcaggt attgtgccgt gcatcccgga aaaaggctct	420
ctgggcacca gcggcgcacct gggtccgctg gctgcgatcg ctctggtttg tgcgggcccag	480
tggaaagccc gtataaacgg ccagattatg ccgggtcgcc aagccctgtc cgaagcaggc	540
gtggAACCGA tggaactgtc atacaaagat ggtctggcgc tgattaatgg cacgagcggt	600
atggtgtggtc tgggcacgt ggtgctgcaa gcagcacgtc gcctggttga tcgctatctg	660
caagt cagcg ctctgtctgt ggaaggcctg gcgggtatga ccaaaccgtt tgacccgcgt	720
gttcatggcg tcaaaccgca ccgcggtcag cgtcaagtgc cctctgcct gtgggaaggc	780
ctggctgata gtcacctggc ggtcaacgaa ctggacacgg aacagaccct ggcaggcgaa	840
atgggcaccc tggttaaagc gggttcgctg gctattigaag atgcgtatag catccgttgc	900
acgcgcaga ttctgggtcc ggtgggttatg gttctggacc gcatcggtgc aaccctgcaa	960
gatgaactga atagctctaa cgacaatccg attgtcctgc cggaagaagc ggaagtgttt	1020
cataacggcc atttccacgg tcaatacgtg gcgatggcga tggatcacct gaatatggct	1080
ctggcgaccc ttacgaacct ggctaattcgt cgcgctgatc gttttctgga caaatcaaac	1140
tcgaatggtc tgccggcctt cctgtgtcgt gaagatccgg gtctgcgtct gggctctgatg	1200
ggcggtaat ttatgacggc ctctatcacc gcagaaaccc gtacgctgac cattccgatg	1260
agtgtgcagt ccctgacgtc aaccgcggat ttccaagaca tcgttagttt tggtttcgtc	1320
gctgcacgtc gccccgcga agtcctgacc aatgcccat acgtcggtgc atttgaactg	1380
ctgtgcgcct gtcaggcagt tgatattcgt ggtgcggaca aactgagttc cttcacgcgc	1440
ccgctgtatg aacgcacccg taaaatcgtc ccgttttcg atcgacgacga aacgattacc	1500
gattacgtgg aaaaactggc agccgacctg attgcgggtg aaccggtgga tgacccgcgt	1560
gcagcacac	1569