



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년07월16일
(11) 등록번호 10-2278343
(24) 등록일자 2021년07월12일

- (51) 국제특허분류(Int. Cl.)
C12N 15/63 (2006.01) C12N 15/70 (2006.01)
C12N 15/74 (2006.01) C12N 15/79 (2006.01)
C12N 9/10 (2006.01) C12N 9/18 (2006.01)
C12P 13/02 (2006.01)
- (52) CPC특허분류
C12N 15/63 (2013.01)
C12N 15/70 (2013.01)
- (21) 출원번호 10-2020-7024620(분할)
- (22) 출원일자(국제) 2013년03월12일
심사청구일자 2020년09월23일
- (85) 번역문제출일자 2020년08월26일
- (65) 공개번호 10-2020-0105535
- (43) 공개일자 2020년09월07일
- (62) 원출원 특허 10-2014-7031804
원출원일자(국제) 2013년03월12일
심사청구일자 2018년03월12일
- (86) 국제출원번호 PCT/US2013/030502
- (87) 국제공개번호 WO 2013/154721
국제공개일자 2013년10월17일
- (30) 우선권주장
61/623,711 2012년04월13일 미국(US)
- (56) 선행기술조사문헌
Journal of Natural Products (2004)
67(8):1283-1286
(뒷면에 계속)

- (73) 특허권자
게노마티카 인코포레이티드
미국 92121 캘리포니아주 샌디에고 넥서스 센터
드라이브 4757
- (72) 발명자
루츠, 제이슨, 제이.
미국 캘리포니아 94080 사우스 샌프란시스코 게이
트웨이 블러바드 600
델 카다이에, 스티븐
미국 캘리포니아 94080 사우스 샌프란시스코 게이
트웨이 블러바드 600
- (74) 대리인
특허법인(유)화우

전체 청구항 수 : 총 12 항

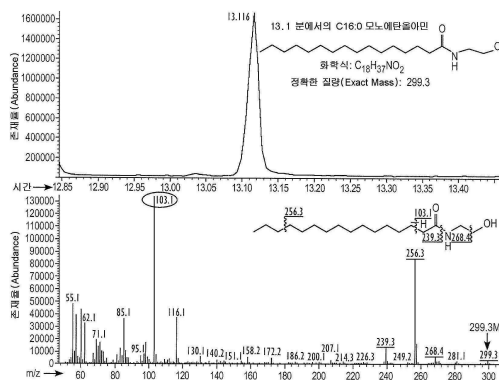
심사관 : 김승범

(54) 발명의 명칭 알카놀아미드들 및 아미도아민들의 미생물에 의한 생산 및 이의 용도

(57) 요약

본 발명은 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 효소를 발현시키도록 조작되는 재조합 미생물에 관한 것이다. 본 발명은 탄소 공급원의 존재 하에서 상기 재조합 미생물을 배양시킴으로써 지방족 아미드를 생산하는 방법을 더 포괄한다.

대표도



(52) CPC특허분류

C12N 15/74 (2013.01)
C12N 15/79 (2013.01)
C12N 9/1029 (2013.01)
C12N 9/18 (2013.01)
C12P 13/02 (2013.01)
C12Y 203/01086 (2013.01)
C12Y 301/01005 (2013.01)
Y02P 20/52 (2015.11)

(56) 선행기술조사문헌

Applied and Environmental Microbiology (2010)
76(5):1633-1641
Archives of Biochemistry and Biophysics (2005)
433(1):212-226
JP2010538602 A
The Journal of Biological Chemistry (1995)
270(9):4216-4219
The Journal of Biological Chemistry (2004)
279(12):11163-11169

명세서

청구범위

청구항 1

생체 내에서(in vivo) 지방족 아미드를 생산하기 위한 재조합 미생물에 있어서,

여기서 상기 재조합 미생물은 에탄올아민 생산을 위한 재조합 경로(recombinant pathway)를 포함하고,

에탄올아민 생산을 위한 경로는 과발현된 (i) 포스포글리세레이트 뮤타아제(phosphoglycerate mutase), (ii) 포스포글리세레이트 탈수소효소(phosphoglycerate dehydrogenase), (iii) 3-포스포세린 아미노기전이효소(3-phosphoserine aminotransferase), 및 (iv) 3-포스포세린 포스페이트(3-phosphoserine phosphate), 및 (v) 이종 세린 탈카르복실화효소(heterologous serine decarboxylase)를 인코딩하는 핵산을 포함하며, 경로는 에탄올아민을 생체 내에서 생산하고,

상기 재조합 미생물은 외인성 티오에스테라아제(exogenous thioesterase) (E.C. 3.1.2.14 또는 E.C. 3.1.1.5) 및 서열번호 1을 포함하는 아미노산 서열을 갖는 외인성 팔미토일푸트레신 합성효소(exogenous palmitoylputrescine synthase, PPS) 폴리펩티드; 또는 외인성 티오에스테라아제(E.C. 3.1.2.14 또는 E.C. 3.1.1.5) 및 서열번호 3을 포함하는 아미노산 서열을 갖는 외인성 N-(4-아미노-2-하이드록실부틸)테트라데칸아미드 합성효소(AhtS) 폴리펩티드;를 추가로 포함하며,

상기 PPS 또는 AhtS 폴리펩티드는 지방족 아미드를 생산하기 위한 에탄올아민 및 아실 티오에스테르의 전환을 촉진(catalyze)시키는, 재조합 미생물.

청구항 2

제1항에 있어서,

상기 포스포글리세레이트 뮤타아제는 *gpm AB*, 상기 포스포글리세레이트 탈수소효소는 *serA*, 상기 3-포스포세린 아미노기전이효소는 *serC*, 및 상기 3-포스포세린 포스페이트는 *serB*인, 재조합 미생물.

청구항 3

제1항에 있어서,

분해효소 에탄올아민 암모니아-분해효소(ethanolamine ammonia-lyase, *eutABC*) 및 세린 탈아미노기효소(serine deaminase, *sdaAB*)를 인코딩하는 유전자가 결실된, 재조합 미생물.

청구항 4

제1항에 있어서,

상기 아실 티오에스테르는 상기 재조합 미생물에 의해 생산된 지방족 아실-ACP 또는 지방족 아실-CoA인, 재조합 미생물.

청구항 5

제1항에 있어서,

상기 티오에스테라아제는 *tesA*인, 재조합 미생물.

청구항 6

제1항에 있어서,

상기 미생물은 박테리아, 남조류(cyanobacteria), 조류(algae) 또는 균류(fungi)인, 재조합 미생물.

청구항 7

제6항에 있어서,

상기 미생물은 *아트로박터*(*Arthrobacter*) AK19, *아시네토박터* 종(*Acinetobacter sp.*) 균주 M-1, *대장균* B, *대장균* C, *대장균* K 또는 *대장균* W 세포인, 재조합 미생물.

청구항 8

지방족 아미드 생산 방법에 있어서,

(a) 제1항 내지 제7항 중 어느 한 항에 따른 재조합 미생물을 제공하는 단계; 및

(b) 폴리펩티드에 대한 적어도 하나의 기질의 존재 하에, 에탄올아민 생산을 위한 경로의 발현과 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드를 인코딩하는 핵산 서열의 발현에 적합한 조건들 하에서, 배양 배지 내에서 상기 재조합 미생물을 배양하는 단계를 포함하는, 지방족 아미드 생산 방법.

를 포함하는, 지방족 아미드 생산 방법.

청구항 9

제8항에 있어서,

일차 아민은 상기 배양 배지 내에 첨가되지 않는, 지방족 아미드 생산 방법.

청구항 10

제8항에 있어서,

지방족 아미드를 상기 배양 배지로부터 분리시키는 단계를 더 포함하는, 지방족 아미드 생산 방법.

청구항 11

제8항에 있어서,

상기 지방족 아미드는 C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19 또는 C20 지방족 알카놀아미드인, 지방족 아미드 생산 방법.

청구항 12

제8항에 있어서,

상기 지방족 아미드는 C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19 또는 C20 지방족 아미도아민인, 지방족 아미드 생산 방법.

발명의 설명

기술 분야

[0001] 관련 출원들에 대한 상호 참조

[0002] 본 출원은 2012년 4월 13일에 출원된 미국 가출원 61/623,711의 이익을 주장하며, 이는 본 명세서에서 전문이 인용참조된다.

[0003] 서열 목록(SEQUENCE LISTING)

[0004] 본 출원은 EFS-Web을 통하여 ASCII 포맷으로 제출된 서열 목록을 포함하며, 이는 본 명세서에서 전문이 인용참조된다. 2013년 3월 11일에 작성된 상기 ASCII 복사본(copy)은 LS00041PCT_SL.txt로 파일명을 지정(named)하였고, 크기는 69,769바이트(bytes)이다.

[0005] 본 발명의 분야

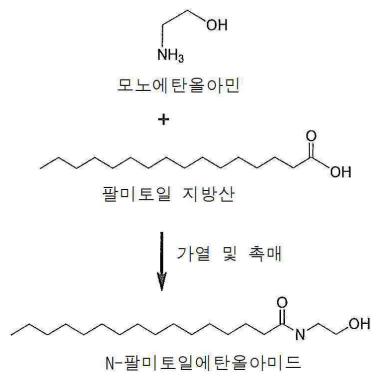
[0006] 본 발명은 탄소 공급원(carbon source)의 존재 하에서 배양되는 경우에 지방족 아미드(fatty amides)를 생산하기 위하여 효소를 발현시키도록 조작되는(engineered) 미생물에 관한 것이다.

배경 기술

[0007] 지방족 아미드는 매우 다양한 생화학적 및 생리학적 기능들을 갖는 동물성 및 식물성 지질의 내인성 성분 (endogenous components)이다[Bachur et al. (1965) *J. Biol. Chem.* 240:1019-1024]. *N*-팔미토일에탄올아민 (PEA), *N*-아라치도노일 에탄올아미드(arachidonoyl ethanolamide)[아난다미드], *N*-올레오일 에탄올아미드(OEA) 및 *N*-아라치도노일 도파민(NADA)과 같은 내인성 지방족 아미드는 중추 및 말초 신경계에서 신호 분자(signaling molecules)로서의 기능을 한다[예를 들어, Tan et al. (2006) *AAPS J.* 8(3): E461-E465; and Lo Verme et al. (2004) *Mol. Pharmacol.* 67(1):15-19]. PEA는 항염증 활성 및 항-통증(anti-nociceptive) 활성을 가하는 것으로 입증되었고, 통증 치료를 위한 PEA의 약제학적 제형들은 상표명 NORMAST로 유럽에서 이용가능하다[Petrosino et al. (2010) *Biochimie* 92(6):724-7; and Bacci et al. (2011) *ISRN Surgery*, Volume 2011, Article ID 917350, 6 pages; doi:10.5402/2011/917350].

[0008] 또한, 지방족 알카놀아미드 및 지방족 아미노아미드와 같은 지방족 아미드는 매우 다양한 비-약제학적인 상업적 용도를 갖는다. 지방족 알카놀아미드 및 지방족 아미노아미드는 기포제(foaming agents), 계면활성제, 또는 개인 위생 용품(예를 들어, 샴푸, 바디 워시 및 세안제), 화장품 제형(예를 들어, 브러쉬, 마스크라 및 립스틱), 및 가정용 청소 제품(예를 들어, 세탁용 세제, 주방용 세제 및 표면 세정용 조성물)의 생산에 있어서 중간체(intermediates)로서 유용하다. 또한, 지방족 알카놀아미드 및 지방족 아미노아미드는 연료 첨가제로서 유용하다. 매년 전세계 시장에서 100,000 톤의 알카놀아미드가 소비되는 것으로 추정된다[Adlercreutz et al. (2010) *Industrial Biotechnology* 6(4):204-211].

[0009] 고전적으로, 상업적인 이용을 위한 지방족 알카놀아미드는 알카놀아민과, 천연 오일 또는 지방과 같은 공급원료(feedstock) 및 미정제 오일로부터 유래되는 지방산 또는 지방산 메틸 에스테르 간의 값비싼 합성 유기 반응들을 통하여 생산된다[Adlercreutz et al., *supra*, and Frost & Sullivan, "Nonionic Surfactants in the Industrial Triad" (2002)]. 예를 들어, PEA는 하기와 같이, 쇼텐-바우만 반응(Schotten-Baumann reaction)에서 모노에탄올아민과 아자유로부터 유래되는 팔미토일 지방산을 반응시킴으로써 생산될 수 있다:



[0010] 또한, 지방족 알카놀아미드는 생합성으로 생성된다. 예를 들어, OEA는 두 개의 효소 과정을 거쳐 포스파티딜에탄올아민(PE) 및 *sn*-1-올레오일-포스파티딜콜린(PC) 전구체들로부터 생산될 수 있고, 여기서 PE 및 *sn*-1-올레오일-PC는 *N*-아실 전이효소(transferase)와 반응하여 *N*-아실 포스파티딜에탄올아민(NAPE)을 형성하고, 이는 이후에 리소(lyso)-PC와 조합되며, NAPE-특이적 포스포리파아제(phospholipase) D와 반응하여 OEA 및 포스파티딘산(phosphatidic acid)을 형성한다[Astarita et al. (2006) *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 290:R1407-R1412를 참조].

[0012] 지방족 아미드를 합성하기 위한 본 기술분야에 알려진 다른 방법들뿐만 아니라 이러한 방법들은 흔히 비효율적인 반응 단계들을 포함하므로, 경제적인 및 환경적인 시각 모두에서 많은 비용이 든다. 따라서, 지방족 아미드의 생산을 위한 방법들 및 반응물(reagent)들을 개선시킬 필요가 있고, 여기서 아미드의 선두 기(head group)의 타입뿐만 아니라 지방족 사슬의 포화도 및 길이가 효율적으로 제어될 수 있다.

발명의 내용

해결하려는 과제

[0013] 본원 발명의 일 측면에서는 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는(catalyze) 폴리펩티드를 인코딩하는 핵산 서열을 포함하는 재조합 미생물을 제공하고, 여기서 상기 미생물은 탄소 공급원

의 존재 하에서 배양된다. 여기서, 미생물이 탄소 공급원의 존재 하에서 배양되는 경우에, 미생물은 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드를 인코딩하는 핵산 서열을 발현시키도록 조작된다. 일 실시예에서, 탄소 공급원은 탄수화물이다. 또 다른 실시예에서, 폴리펩티드는 팔미토일푸트레신 합성효소(palmitoylputrescine synthase: PPS)이다. 또 다른 실시예에서, 폴리펩티드는 N-(4-아미노-2-하이드록실부틸)테트라테칸아미드 합성효소(AhtS) 폴리펩티드이다.

과제의 해결 수단

- [0014] 본 발명의 또 다른 측면에서는 SEQ ID NO: 1의 아미노산 서열을 갖는 팔미토일푸트레신 합성효소(PPS) 폴리펩티드를 제공한다. 일 실시예에서, PPS 폴리펩티드는 SEQ ID NO: 1의 아미노산 서열에 대하여 적어도 약 70 %, 적어도 약 75 %, 적어도 약 80 %, 적어도 약 85 %, 적어도 약 90 %, 적어도 약 91 %, 적어도 약 92 %, 적어도 약 93 %, 적어도 약 94 %, 적어도 약 95 %, 적어도 약 96 %, 적어도 약 97 %, 적어도 약 98 % 또는 적어도 약 99 % 서열 동일성(sequence identity)을 갖는 아미노산 서열을 포함한다. 또 다른 실시예에서, PPS 폴리펩티드는 SEQ ID NO: 2의 핵산 서열을 포함하는 핵산 서열에 의해 인코딩된다.
- [0015] 본 발명의 또 다른 측면에서는 SEQ ID NO: 3의 아미노산 서열을 갖는 N-(4-아미노-2-하이드록실부틸)테트라테칸아미드 합성효소(AhtS) 폴리펩티드를 제공한다. 일 실시예에서, AhtS 폴리펩티드는 SEQ ID NO: 3의 아미노산 서열에 대하여 적어도 약 70 %, 적어도 약 75 %, 적어도 약 80 %, 적어도 약 85 %, 적어도 약 90 %, 적어도 약 91 %, 적어도 약 92 %, 적어도 약 93 %, 적어도 약 94 %, 적어도 약 95 %, 적어도 약 96 %, 적어도 약 97 %, 적어도 약 98 % 또는 적어도 약 99 % 서열 동일성을 갖는 아미노산 서열을 포함한다. 또 다른 실시예에서, AhtS 폴리펩티드는 SEQ ID NO: 22의 핵산 서열을 포함하는 핵산 서열에 의해 인코딩된다.
- [0016] 게다가, 본 발명의 또 다른 측면에서는 재조합 미생물을 제공하고, 여기서 일차 아민은 3-디메틸아미노-1-프로필아민, (±)-1-아미노-2-프로판올, 2-메톡시에틸아민, 3-아미노-1-프로판올, 2-아미노-1,3-프로판디올, 3-메톡시프로필아민, N-(2-하이드록시에틸)에틸렌디아민, 및 부틸아민, 1,4-디아미노부탄 또는 이들의 조합을 포함하나, 이로 제한되지 않는다.
- [0017] 계속해서, 본 발명의 또 다른 측면에서는 재조합 미생물을 제공하고, 여기서 아실 티오에스테르는 지방족 아실-ACP 또는 지방족 아실-CoA이다. 지방족 아실-ACP 또는 지방족 아실-CoA는 미생물에 의해 생산된다.
- [0018] 본 발명은 지방산 생합성 폴리펩티드, 티오에스테라아제(thioesterase) 폴리펩티드(EC 3.1.2.14 또는 EC 3.1.1.5) 및 아실-CoA 합성효소 폴리펩티드(EC 2.3.1.86) 중 하나 이상을 인코딩하는 핵산 서열을 포함하는 재조합 미생물을 더 포괄한다. 일 실시예에서, 티오에스테라아제 폴리펩티드를 인코딩하는 핵산 서열은 'tesA'이다. 또 다른 실시예에서, 아실 CoA 합성효소 폴리펩티드를 인코딩하는 핵산 서열은 fadD이다. 또 다른 실시예에서, 미생물은 accABCD, FabD, FabH, FabG, FabB, FabA, FabZ, FabF, FabI 및/또는 FadR을 포함하나 이로 제한되지 않는 지방산 생합성 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.
- [0019] 본 발명은 박테리아, 남조류(cyanobacteria), 조류(algae) 및 균류(fungi)를 포함하나 이로 제한되지 않는 미생물을 더 고려한다. 일 실시예에서, 박테리아는 대장균(*E. coli.*)이다. 또 다른 실시예에서, 균류는 효모(yeast) 또는 곰팡이(filamentous fungi)이다. 또 다른 실시예에서, 미생물은 맥주효모균(*Saccharomyces cerevisiae*), 캔디다 리폴리티카(*Candida lipolytica*), 대장균(*Escherichia coli*), 아트로박터(*Arthrobacter*), 로도토룰라 글루티닌스(*Rhodotorula glutinins*), 아시네토박터(*Acinetobacter*), 캔디다 리폴리티카(*Candida lipolytica*), 보트리오코커스 브라우니(*Botryococcus braunii*), 비브리오 퍼니씨(*Vibrio furnissii*), 마이크로코커스 로이테우스(*Micrococcus leuteus*), 스테노트로포모나스 말토티리아(*Stenotrophomonas maltophilia*), 바실러스 서브틸러스(*Bacillus subtilis*), 바실러스 리체노포르미스(*Bacillus licheniformis*), 슈도모너스 푸티다(*Pseudomonas putida*), 슈도모나스 플루오레스스(*Pseudomonas fluorescens*), 스트렙토마이세스 코엘리콜라(*Streptomyces coelicolor*), 시네크코커스 종(*Synechococcus sp.*) PCC7002, 써모시네크코커스 엘롱게이트(*Thermosynechococcus elongates*) BP-1, 프로토테카 모리포르미스(*Prototheca moriformis*), 프로토테카 크루가니(*Prototheca krugani*), 프로토테카 스테그노라(*Prototheca stagnora*), 프로토테카 조프피(*Prototheca zopfii*) 또는 콜레라 프로토테코이드(*Chorella protothecoide*) 세포를 포함하나, 이로 제한되지 않는다. 또 다른 실시예에서, 미생물은 아트로박터 AK 19, 아시네토박터 종(*Acinetobacter sp.*) 균주 M-1, 대장균 B, 대장균 C, 대장균 K 또는 대장균 W 세포를 포함하나, 이로 제한되지 않는다.
- [0020] 본 발명의 또 다른 측면에서는 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드를 인코딩하는 핵산 서열을 포함하는 재조합 미생물을 제공하고, 여기서 일차 아민 및 아실 티오에스테르의

지방족 아미드로의 전환을 촉진시키는 폴리펩티드는 미생물에 내부로부터 발생한다(endogenous).

[0021] 본 발명의 또 다른 측면에서는 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드를 인코딩하는 핵산 서열을 포함하는 재조합 미생물을 제공하고, 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드는 미생물의 외부에서 발생한다.

[0022] 계속해서, 본 발명의 또 다른 측면에서는 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 효소를 인코딩하는 핵산 서열을 포함하는 재조합 미생물을 제공한다. 일 실시예에서, 지방족 아미드는 지방족 알카놀아미드 및/또는 지방족 아미도아민이다. 또 다른 실시예에서, 지방족 아미드는 C14, C16 및/또는 C18 지방족 알카놀아미드 및/또는 C14, C16 또는 C18 지방족 아미도아민이다. 또 다른 실시예에서, 지방족 아미드는 C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19 또는 C20 지방족 알카놀아미드 및/또는 C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19 또는 C20 지방족 아미도아민이다.

[0023] 게다가, 본 발명의 또 다른 측면에서는 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드를 인코딩하는 핵산 서열을 포함하는 재조합 미생물을 제공하고, 여기서 미생물은 세린 탈카르복실화효소(serine decarboxylase) 폴리펩티드를 발현시킨다.

[0024] 본 발명은 지방족 아미드의 생산 방법을 더 포괄하고, 상기 방법은: (a) 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드를 인코딩하는 핵산 서열을 포함하는 재조합 미생물을 제공하는 단계; 및 (b) 상기 폴리펩티드에 대한 적어도 하나의 기질의 존재 하에서 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 상기 폴리펩티드를 인코딩하는 상기 핵산 서열의 발현에 적합한 조건들 하에서 배양 배지 내에서 상기 재조합 미생물을 배양하는 단계를 포함한다. 이 방법은 상기 배양 배지에서 상기 지방족 아미드를 분리시키는 단계를 더 포함할 수 있다. 이 방법은 지방족 아미드를 생산하는데 사용될 수 있다. 일 실시예에서, 지방족 아미드는 지방족 알카놀아미드 및/또는 지방족 아미도아민이다. 또 다른 실시예에서, 지방족 아미드는 C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19 또는 C20 지방족 알카놀아미드 및/또는 C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19 또는 C20 지방족 아미도아민이다.

도면의 간단한 설명

[0025] 본원 발명은 바람직한 실시예들을 도시하는데 도움이 되는, 첨부된 도면들과 함께 읽을 때에 가장 잘 이해된다. 그러나, 본 발명이 도면들에 개시된 특정 실시예들로 제한되지 않는 것으로 이해한다.

도 1은 에탄올아민의 존재 하에서 배양되는 팔미토일푸트레신 합성효소(PPS)를 인코딩하는 발현 벡터로 형질전환되는 *대장균* MG1655 균주 DG5에 의해 생산된 지방 종류(fatty species)의 대표적인 가스 크로마토그래피-질량 분석(GC-MS)의 크로마토그램(chromatogram)이다. 위쪽 패널(upper panel)은 13.1 분의 GC 머무름 시간(retention time)을 갖는 피크(peak)를 나타내고, 이는 아래쪽 패널에 나타난 MS 분석에 의해 N-팔미토일에탄올아미드로 식별되었다.

도 2a 내지 도 2d는 N,O-bis(트리메틸실릴)-트리플루오로아세트아미드 (BSTFA)에 의한 유도체화(derivatization) 후에 N-팔미토일에탄올아미드 생성물의 GC-MS 크로마토그램들이다. 도 2a는 13.3 분의 GC 머무름 시간을 갖는 피크를 나타내고, 이는 도 2e에 나타난 MS 분석에 의해 (TMS)-보호된 N-팔미토일에탄올아미드로 식별되었다. 도 2b 내지 도 2d는 각각 BSTFA 단독, BSTFA 유도체화를 하지 않은 N-팔미토일에탄올아미드 및 바탕 반응(blank reaction)의 대조구 크로마토그램들이다.

도 3은 3-디메틸아미노-1-프로필아민의 존재 하에서 배양되는 PPS를 인코딩하는 발현 벡터로 형질전환되는 *대장균* MG1655 균주 DG5에 의해 생산된 지방족 N-(3-디메틸아미노-1-프로필아민)아미드들의 GC-MS 크로마토그램이다.

도 4는 나타난 일차 아민 공급원(feed)들의 존재 하에서 배양되고 PPS를 인코딩하는 발현 벡터로 형질전환되는 *대장균* MG1655 세포들로부터 얻어진 지방족 아미드 생성물들을 나타낸다.

도 5는 3-디메틸아미노-1-프로필아민의 존재 하에서 배양되는 효소 N-(4-아미노-2-하이드록실부틸)테트라데칸아미드 합성효소(AhtS)를 인코딩하는 발현 벡터로 형질전환되는 *대장균* MG1655 균주에 의해 생산된 지방 종류의 대표적인 GC-MS 크로마토그램이다. 위쪽 패널은 11.4 분의 GC 머무름 시간을 갖는 피크를 나타내고, 이는 아래쪽 패널에 나타난 MS 분석에 의해 C14:0 지방족 N-(3-디메틸아미노-1-프로필아미드)로 식별되었다.

도 6은 본 발명의 방법들에 따라서 유전적으로 변형(genetically modified)될 수 있는 대사 경로들의 개략적인

도식이다.

발명을 실시하기 위한 구체적인 내용

- [0026] 천연 항생물질인 팔미토일푸트레신은 GenBank 수탁 번호 AY632377.1(이후에 "AY632377"이라고 함)(SEQ ID NO: 2)의 핵산 서열에 의해 인코딩되는 팔미토일푸트레신 합성효소(PPS)[GenBank 수탁 번호 AAV33349.1(이후에 "AAV33349"라고 함)](SEQ ID NO: 1)를 발현시키는 박테리아에 의해 생산될 수 있다[Brady *et al.* (2004) *J. Nat. Prod.* 67:1283-1286]. *대장균* 내에서 과발현되는 경우, AY632377에 의해 인코딩되는 PPS는 단 하나의 주요한 푸트레신의 N-아실 유도체(1,4-디아미노부탄), 즉 팔미토일푸트레신을 생산하는 것으로 입증되었다(Brady *et al.*, *supra*). 상기 효소를 인코딩하는 동족체(homologue)인, N-(4-아미노-2-하이드록실부틸)테트라테칸아미드 합성효소(AhtS)(GeneBank 수탁 번호 ACX33975.1)(SEQ ID NO: 3)는 PPS의 아미노산 서열에 대하여 38 % 동일한 아미노산 서열을 갖는다. 배양되지 않은 박테리아 RM44(GenBank GQ869386)로부터의 N-(4-아미노-2-하이드록실부틸)테트라테칸아미드 합성효소(AhtS) 유전자는 (SEQ ID NO: 22)로 나타낸다.

- [0027] 최소한 부분적으로, 본 발명은 PPS 또는 AhtS를 발현시키는 미생물(예를 들어, 박테리아)이 탄소 공급원의 존재 하에서 배양되는 경우에 아실 티오에스테르 전구체들로부터 지방족 아미드들을 생산할 수 있다는 발견을 기반으로 한다. 이론에 얽매이고 싶지는 않지만, PPS는 직접적으로 아실 티오에스테르와 일차 아민 간의 아미드화(amidation)를 촉진시킨다고 믿어진다. 처음에, 미생물은 구체적으로 지방족 아미드를 생산하기 위하여 PPS 또는 AhtS와 같은 효소를 발현시키도록 조작된다. 이에 의해서 미생물은 분지형 또는 비분지형을 포함하는, 원하는 사슬 길이(chain length)의 지방족 아미드를 발생시키는 편리한 생물학적 생공장(venient biological factory)의 역할을 하므로 이롭다. 또한, 다양한 상이한 공급원료들[예를 들어, 옥수수, 사탕수수, 글리세롤, 지렁이풀(switchgrass)]은 용통적으로, 미생물을 위해 필요한 탄소 공급원을 공급하도록 통용되어 사용될 수 있다. 이처럼, 미생물은 요구에 따라 발효를 통하여 수확될 수 있는 지방족 아미드를 생산하는데 사용될 수 있고, 이에 의해서 여전히 값비싼 천연 오일 및 복잡한 합성 화학에 의존하는 번거롭고 많은 비용이 드는 이전 기술 시스템들을 우회할 수 있다. 기포제, 양이온 계면활성제, 샴푸 및 목욕 용품으로서 사용하기 위한 중간체, 화장품 및 의약품의 유효제, 연료 첨가제, 및 이와 유사한 것들을 포함하나 이로 제한되지 않는 다수의 제품들의 생산에 지방족 아미드들이 필요하다.

- [0028] 본 발명은 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드를 인코딩하는 핵산 서열을 발현시키도록 조작된 재조합 미생물을 제공하고, 여기서 상기 미생물은 탄소 공급원의 존재 하에서 배양된다. 일 실시예에서, 탄소 공급원은 탄수화물이다. 좀 더 구체적으로, 미생물은 조작되어, PPS 또는 AhtS와 같은 효소가 알카놀아미드 및 아미도아민과 같은 지방족 아미드들을 생산하기 위해 아실 티오에스테르(예를 들어, 아실-CoA 또는 아실-ACP)와 여하한 일차 아민(예를 들어, 에탄올아민, 아민 3-디메틸아미노-1-프로필아민) 사이의 아미드화를 촉진시키기 위하여 발현되도록 한다. 지방족 알카놀아미드들(예를 들어, 코카미도 프로필 베타인의 합성에 사용되는 중간체들) 및 아미도아민들이 천연 오일(또는 지방) 및 미정제 오일과 같은 공급원료들로부터 어느 정도까지(so far) 합성적으로 생성되기 때문에 이는 신규한 과정이며, 이는 원하는 물질들이 얻어질 때까지 원재료를 정제하는 것에 의존하므로 비효율적인 과정이다. 비교하여, 본원 발명은 생산 방법을 제공하고, 여기서 미생물은 효소들을 발현시키도록 조작되어, 예를 들어 알카놀아미드 및 지방족 N-(3-디메틸아미노-1-프로필아민)아미드가 생화학적으로 합성되도록 하며, 이는 지방족 아미드를 생산하는데 훨씬 더 효과적인 과정이다. 아미노산들 또는 탄수화물들이 필요한 탄소 공급원을 공급하기 위하여 미생물의 발효 배지에 첨가될 수 있다(실시예 3 내지 7을 참조). 대안적으로, 미생물은 *생체 내에서(in vivo)* 자체 일차 아민을 발생시키도록 조작될 수 있다. 예를 들어, 에탄올아민의 생합성은 유전적으로 세린 생합성 및 세린 탈카르복실화(decarboxylation) 경로들을 증가시킴으로써 달성될 수 있다(실시예 8을 참조). 효소적으로, AhtS는 PPS처럼 동일한 아미드 화합물들을 생산하나, C14:0 지방족 티오에스테르 기질들을 선호한다. 두 개의 효소들 모두 EC 족(family) 2.3.1.X.X에 속한다.

- [0029] 본 발명은 재조합 미생물 내에서의 지방족 아미드의 생산 방법을 더 제공한다. 이러한 방법에 의해 생산되는 지방족 아미드들은 지방족 알카놀아미드 및 지방족 아미도아민을 포함하나, 이로 제한되지 않는다. 일 실시예에서, 지방족 아미드는 C14, C16 및/또는 C18 지방족 알카놀아미드이다. 또 다른 실시예에서, 지방족 아미드는 C14, C16 및/또는 C18 지방족 아미도아민이다. 상기 방법은 (a) 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 PPS 또는 AhtS와 같은 폴리펩티드를 인코딩하는 핵산 서열을 발현시키도록 조작된 재조합 미생물을 제공하는 단계; 및 (b) 상기 폴리펩티드에 대한 적어도 하나의 기질의 존재 하에서 상기 폴리펩티드의 발현에 적합한 조건들 하에서 상기 재조합 미생물을 배양시켜, 이로 인해서 지방족 아미드를 생산하

는 단계를 포함한다. 미생물은 탄소 공급원의 존재 하에서 배양된다. 탄소 공급원은 아미노산, 탄수화물 및 지질을 포함하나 이로 제한되지 않는 매우 다양한 상이한 공급원들로부터 선택될 수 있다. 일 실시예에서, 탄소 공급원은 탄수화물이다. 미생물에 의해 생산되는 지방족 아미드는 배양액(culture broth)[예를 들어, 발효액(fermentation broth)]으로부터 분리될 수 있다. 일 실시예에서, 지방족 아미드는 미생물의 세포의 환경으로부터 분리된다. 또 다른 실시예에서, 지방족 아미드는 부분적으로 또는 전적으로, 미생물로부터 자발적으로 분리된다. 또 다른 실시예에서, 지방족 아미드는 선택적으로 하나 이상의 적합한 수송 단백질(transport protein)들의 도움으로, 세포의 환경으로 이송된다. 또 다른 실시예에서, 지방족 아미드는 세포의 환경으로 수동적으로 이송된다.

[0030] "지방족 아미드" 및 "알킬 아미드"라는 용어들은 식 R^1CONHR^2 를 갖는 화합물을 칭하고, 여기서 R^1 은 지방산으로부터 유래되는 지방족 기(aliphatic group)를 나타내며, R^2 는 일차 아민으로부터 유래되는 치환 또는 비치환된 헤테로알케닐기, 치환 또는 비치환된 알케닐기, 치환 또는 비치환된 헤테로알킬기, 또는 치환 또는 비치환된 알킬기를 나타낸다.

[0031] "아실 티오에스테르"는 생산 숙주 미생물의 지방산 생합성 경로에 의해 "활성화되는" 지방산을 칭한다. 아실 티오에스테르는 미생물의 내인성 지방산으로부터 발생될 수 있거나, 또는 아실 티오에스테르는 외인적으로(exogenously) 미생물에 제공되는 지방산으로부터 발생될 수 있다. 아실 티오에스테르의 비-제한적인 예시들은 아실-조효소 A(CoA) 및 아실-아실기 운반 단백질(acyl carrier protein: ACP)이다.

[0032] "아실-CoA"라는 용어는 CoA의 4'-포스포판테테인 잔기(moiety)의 술폰드릴기와 알킬 사슬의 카르보닐 탄소 사이에서 형성되는 아실 티오에스테르를 칭하고, 이는 식 $R^1-C(O)S-CoA$ 를 가지며, 여기서 R^1 은 지방족 기(aliphatic group)이다. "아실-ACP"라는 용어는 ACP에 부착된 4'-포스포판테테인 잔기의 술폰드릴기와 알킬 사슬의 카르보닐 탄소 사이에서 형성되는 아실 티오에스테르를 칭하고, 이는 식 $R^1-C(O)S-ACP$ 를 가지며, 여기서 R^1 은 지방족 기(aliphatic group)이다.

[0033] "지방산"이라는 용어는 식 R^1COOH 를 갖는 카르복실산을 의미한다. R^1 은 지방족 기(aliphatic group), 바람직하게는 알킬기를 나타낸다. R^1 은 4 내지 26 개의 탄소 원자들을 포함할 수 있다. 여하한 실시예들에서, R^1 은 길이가 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 적어도 20 또는 적어도 21 개의 탄소들이다. 대안적으로 또는 추가적으로, R^1 은 길이가 22 이하, 21 이하, 20 이하, 19 이하, 18 이하, 17 이하, 16 이하, 15 이하, 14 이하, 13 이하, 12 이하, 11 이하, 10 이하, 9 이하, 8 이하, 7 이하 또는 6 이하의 탄소들이다. 따라서, R^1 기는 상기의 종단점(endpoint)들 중 여하한 두 개로 제한되는(bounded by) R^1 기를 가질 수 있다. 예를 들어, R^1 기는 길이가 6 내지 16 개의 탄소들, 길이가 10 내지 14 개의 탄소들, 또는 길이가 12 내지 18 개의 탄소들일 수 있다. 일부 실시예들에서, 지방산은 C_6 , C_7 , C_8 , C_9 , C_{10} , C_{11} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} , C_{18} , C_{19} , C_{20} , C_{21} , C_{22} , C_{23} , C_{24} , C_{25} 또는 C_{26} 지방산이다. 여하한 실시예들에서, 지방산은 C_6 , C_8 , C_{10} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} 또는 C_{18} 지방산이다. 일 바람직한 실시예에서, 지방족 아미드는 C_{14} , C_{16} 또는 C_{18} 지방족 알카놀아미드이다. 또 다른 바람직한 실시예에서, 지방족 아미드는 C_{14} , C_{16} 또는 C_{18} 지방족 아미도아민이다. 또 다른 바람직한 실시예에서, 지방족 아미드는 C_8 , C_9 , C_{10} , C_{11} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} , C_{18} , C_{19} 또는 C_{20} 알카놀아미드이다. 또 다른 바람직한 실시예에서, 지방족 아미드는 C_8 , C_9 , C_{10} , C_{11} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} , C_{18} , C_{19} 또는 C_{20} 아미도아민이다.

[0034] 지방산의 R^1 기는 직쇄형 또는 분지쇄형일 수 있다. 분지쇄형은 하나 이상의 분지점(point of branching)들을 가질 수 있고, 사이클릭형 분지들을 포함할 수 있다. 일부 실시예들에서, 분지형 지방산은 C_6 , C_7 , C_8 , C_9 , C_{10} , C_{11} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} , C_{18} , C_{19} , C_{20} , C_{21} , C_{22} , C_{23} , C_{24} , C_{25} 또는 C_{26} 분지형 지방산이다. 특정 실시예들에서, 분지형 지방산은 C_6 , C_8 , C_{10} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} 또는 C_{18} 분지형 지방산이다. 여하한 실시예들에서, 분지형 지방산의 하이드록실기는 첫번째(C_1) 위치에 존재한다.

[0035] 여하한 실시예들에서, 분지형 지방산은 이소- $C_{7:0}$, 이소- $C_{8:0}$, 이소- $C_{9:0}$, 이소- $C_{10:0}$, 이소- $C_{11:0}$, 이소- $C_{12:0}$, 이소-

C_{13:0}, 이소-C_{14:0}, 이소-C_{15:0}, 이소-C_{16:0}, 이소-C_{17:0}, 이소-C_{18:0}, 이소-C_{19:0}, 안테이소-C_{7:0}, 안테이소-C_{8:0}, 안테이소-C_{9:0}, 안테이소-C_{10:0}, 안테이소-C_{11:0}, 안테이소-C_{12:0}, 안테이소-C_{13:0}, 안테이소-C_{14:0}, 안테이소-C_{15:0}, 안테이소-C_{16:0}, 안테이소-C_{17:0}, 안테이소-C_{18:0} 및 안테이소-C_{19:0} 분지형 지방산으로부터 선택된다.

[0036] 분지형 또는 비분지형 지방산의 R¹기는 포화되거나 불포화될 수 있다. 만약 불포화된다면, R¹기는 하나 이상의 불포화 지점들을 가질 수 있다. 일부 실시예들에서, 불포화 지방산은 단일불포화 지방산이다. 여하한 실시예들에서, 불포화 지방산은 C6:1, C7:1, C8:1, C9:1, C10:1, C11:1, C12:1, C13:1, C14:1, C15:1, C16:1, C17:1, C18:1, C19:1, C20:1, C21:1, C22:1, C23:1, C24:1, C25:1 또는 C26:1 불포화 지방산이다. 다른 실시예들에서, 불포화 지방산은 C10:1, C12:1, C14:1, C16:1 또는 C18:1 불포화 지방산이다. 또 다른 실시예들에서, 불포화 지방산은 오메가-7 위치에서 불포화된다. 여하한 실시예들에서, 불포화 지방산은 cis 이중 결합을 포함한다.

[0037] 일차 아민은 식 R²NH₂를 갖는 PPS에 대한 기질로서의 역할을 할 수 있는 여하한 일차 아민일 수 있고, 여기서 R²는 치환 또는 비치환된 알킬, 치환 또는 비치환된 헤테로알킬, 치환 또는 비치환된 알케닐, 또는 치환 또는 비치환된 헤테로알케닐 사슬을 나타낸다. 여하한 실시예들에서, R²는 하이드록실기로 치환되고, 지방족 아미드는 "알카놀아미드" 또는 "지방족 알카놀아미드"로 칭할 수 있다. 다른 실시예들에서, R²는 아미노기를 포함하고, 지방족 아미드는 "아미도아민" 또는 "지방족 아미도아민"으로 칭할 수 있다.

[0038] R²는 1 내지 12 개의 탄소 원자들을 포함할 수 있다. 여하한 실시예들에서, R²는 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 6 또는 적어도 7 개의 탄소들을 포함한다. 대안적으로 또는 추가적으로, R²는 12 이하, 11 이하, 10 이하, 9 이하, 8 이하, 7 이하, 6 이하, 5 이하 또는 4 이하의 탄소들을 포함한다. 따라서, R²기는 상기의 종단점(endpoint)들 중 여하한 두 개로 제한되는 탄소들을 포함할 수 있다. 예를 들어, R²기는 2 내지 8 개의 탄소들, 4 내지 10 개의 탄소들, 또는 3 내지 6 개의 탄소들을 포함할 수 있다. 일부 실시예들에서, R²기는 3, 4, 5 또는 6 개의 탄소 원자들을 포함한다.

[0039] 일차 아민의 R²기는 직쇄형 또는 분지쇄형일 수 있다. 분지쇄형은 하나 이상의 분지점(point of branching)들을 가질 수 있고, 사이클릭형 분지들을 포함할 수 있다.

[0040] 달리 명시되지 않는 한, "알킬"이라는 용어는 그 자체로 또는 또 다른 치환기의 일부로서, 직쇄 또는 분지쇄, 또는 사이클릭 탄화수소 라디칼, 또는 이들의 조합을 의미한다. 또한, 예를 들어 하이드록시알킬, 할로알킬, 아미노알킬, 알킬아미노, 디알킬아미노 등과 같은 기(group)의 일부로서 "알킬"이 존재하는 모든 경우에, 이러한 정의가 적용된다.

[0041] 달리 명시되지 않는 한, "알케닐"이라는 용어는 그 자체로 또는 또 다른 치환기의 일부로서, 예를 들어 약 2 내지 약 12 개의 탄소 원자들을 포함하고 적어도 하나의 탄소-탄소 이중 결합을 포함하는, 직쇄 또는 분지쇄, 또는 사이클릭 탄화수소 라디칼, 또는 이들의 조합을 의미한다.

[0042] 달리 명시되지 않는 한, "헤테로알킬" 및 "헤테로알케닐"이라는 용어들은 직쇄 또는 분지쇄, 또는 사이클릭 탄화수소 라디칼, 또는 이들의 조합을 칭하고, 이들은 명시된 개수의 탄소 원자들 그리고 O, N, Si 및 S로 이루어진 그룹으로부터 선택되는 적어도 하나의 헤테로원자로 이루어지며, 여기서 상기 질소 원자 및 황 원자는 선택적으로 산화될 수 있고, 질소 헤테로원자는 선택적으로 4차화(quaternized)될 수 있다. 헤테로원자(들) O, N, Si 및 S는 헤테로알킬기 또는 헤테로알케닐기의 임의의 내부 위치에 위치될 수 있다. R²기에 대한 예시적인 헤테로알킬기는 -CH₂-CH₂-OH, -CH₂-CH₂-O-CH₃, -CH₂-CHOH-CH₃, -CH₂-CH₂-CH₂-O-CH₃, -CH₂-CH₂-CH₂-N(CH₃)₂, -CH₂-CH₂-CH₂-NH-CH₂-CH₃, -CH₂-CH₂-CH₂-NH-CH₂-CH₂-OH, -CH-(CH₂-OH)₂ 및 -Si(CH₃)₃를 포함하나, 이로 제한되지 않는다. 2 개의 헤테로원자들까지는 예를 들어 -CH₂-CH₂-O-Si(CH₃)₃와 같이, 연속해서 존재할 수 있다.

[0043] "알킬", "헤테로알킬", "알케닐" 및 "헤테로알케닐"이라는 용어들은 나타낸 라디칼의 치환된 형태 및 비치환된 형태 둘 모두를 포함하는 것으로 여겨진다. R²기의 알킬, 헤테로알킬, 알케닐 및 헤테로알케닐 라디칼들에 대한 예시적인 치환기들은 0에서부터 (2m'+1)[여기서, m'는 각각의 라디칼에 대한 총 탄소 원자수임]까지의 범위에

이르는 개수로 -OR', =O, =NR', =N-OR', -NR'R'', -SR', -할로젠, -SiR'R''R''', -OC(O)R', -C(O)R', -CO₂R', -CN 및 -NO₂로부터 선택되는 다양한 기들 중 하나 이상일 수 있으나, 이로 제한되지 않는다. 각각의 R', R'' 및 R'''은 독립적으로, 수소, 비치환 또는 치환된 알킬기, 비치환 또는 치환된 헤테로알킬기, 비치환 또는 치환된 알케닐기, 비치환 또는 치환된 헤테로알케닐기, 알콕시기 또는 티오알콕시기 또는 아릴알킬기를 칭한다.

[0044] 일부 실시예들에서, R²기에 대한 치환기는 -OH이다. 여하한 실시예들에서, R²기는 2, 3 또는 4 개의 하이드록실 기들을 포함하는 폴리하이드록시 알킬 또는 폴리하이드록시 헤테로알킬 잔기이다.

[0045] 또한, R²기의 알킬, 헤테로알킬, 알케닐 또는 헤테로알케닐 사슬은 폴리에틸렌 옥사이드 잔기에 의해 방해받을 수 있다. 여하한 실시예들에서, R²기는 2 이상, 3 이상, 4 이상, 5 이상, 또는 6 이상, 또는 7 이상의 에틸렌 옥사이드 잔기들을 포함하는 폴리에틸렌 옥사이드 잔기를 포함한다. 다른 실시예들에서, R²기는 12 이하, 11 이하, 10 이하, 9 이하, 8 이하, 7 이하, 6 이하 또는 5 이하의 에틸렌 옥사이드 잔기들을 포함하는 폴리에틸렌 옥사이드 잔기를 포함한다. 따라서, R²기는 상기의 종단점(endpoint)들 중 여하한 두 개로 제한되는 에틸렌 옥사이드 잔기들의 개수를 갖는 폴리에틸렌 옥사이드 잔기를 포함할 수 있다. 예를 들어, R²기는 2 내지 10 개, 4 내지 8 개, 또는 3 내지 5 개의 에틸렌 옥사이드 잔기들을 갖는 폴리에틸렌 옥사이드 잔기를 포함할 수 있다.

[0046] 일차 아민은 발효성 탄소 공급원으로부터 미생물 내에서 생산될 수 있다. 예를 들어, 모노에탄올아민은 세린 탈카르복실화효소(SDC)의 작용에 의해 세린으로부터 생체 내에서(*in vivo*) 발생될 수 있다[Rontein *et al.* (2001) *J. Biol. Chem.* 276(38):35523-35529]. 일부 실시예들에서, 미생물은 내인성 SDC 폴리펩티드를 발현시킨다. 다른 실시예들에서, 미생물은 SDC 폴리펩티드를 과발현시키도록 조작된다.

[0047] 푸트레신(1,4-디아미노부탄)은, 각각 아르기닌을 아그마틴으로 전환시키고 아그마틴을 푸트레신으로 전환시키는, 아르기닌 탈카르복실화효소(arginine decarboxylase: ADC) 및 아그마틴 우레오하이드로라아제(agsmatine ureohydrolase: AUH)의 작용에 의해 아르기닌으로부터 생체 내에서(*in vivo*) 발생될 수 있다[Moore *et al.* (1990) *J. Bacteriol* 172(8):4631-4640]. 일부 실시예들에서, 미생물은 내인성 ADC 및 AUH 폴리펩티드들을 발현시킨다. 다른 실시예들에서, 미생물은 ADC 폴리펩티드, AUH 폴리펩티드, 또는 ADC 및 AUH 폴리펩티드들을 발현시키도록 조작된다. 여하한 실시예들에서, ADC는 대장균 MG1655로부터 *speA*(GenBank 수탁 번호 NC_000913) 유전자에 의해 인코딩된다.

[0048] 또한, 일차 아민은 외인적으로 미생물에 제공될 수 있다.

[0049] 본 발명에 사용하기에 적합한 예시적인 일차 아민들은 에탄올아민(모노에탄올아민), 3-디메틸아미노-1-프로필아민, (±)-1-아미노-2-프로판올, 2-메톡시에틸아민, 3-아미노-1-프로판올, 2-아미노-1-3-프로판디올, 3-메톡시프로필아민, N-(2-하이드록시에틸)에틸렌디아민, 부틸아민, 1,4-디아미노부탄 및 이들의 조합들을 포함하나, 이로 제한되지 않는다. 여하한 실시예들에서, 일차 아민은 3-디메틸아미노-1-프로필아민이다.

[0050] 핵산이 발현되고 미생물이 탄소 공급원의 존재 하에서 배양되는 경우에, 본 발명의 방법들 및 제조할 미생물들에 사용하기에 적합한 핵산은 일차 아민 및 아실 티오에스테르를 지방족 아마이드로 전환시킬 수 있는 폴리펩티드를 인코딩하는 서열을 갖는 여하한 핵산일 수 있다.

[0051] 일 실시예에서, 폴리펩티드는 PPS 폴리펩티드이다. 여하한 실시예들에서, PPS 폴리펩티드는 SEQ ID NO: 1의 아미노산 서열, 즉 AAV33349의 아미노산 서열을 포함하거나, 필수적으로 상기 아미노산 서열로 이루어지거나, 또는 상기 아미노산 서열로 이루어진다. 일부 실시예들에서, PPS 폴리펩티드는 SEQ ID NO: 2의 핵산 서열을 포함하는 핵산 서열에 의해 인코딩된다. 다른 실시예들에서, PPS는 폴리펩티드는 SEQ ID NO: 1의 아미노산 서열을 갖는 PPS 폴리펩티드의 동족체이다. 바람직하게, PPS 폴리펩티드는 SEQ ID NO: 1의 아미노산 서열에 대하여 적어도 약 70 %, 적어도 약 75 %, 적어도 약 80 %, 적어도 약 85 %, 적어도 약 90 %, 적어도 약 91 %, 적어도 약 92 %, 적어도 약 93 %, 적어도 약 94 %, 적어도 약 95 %, 적어도 약 96 %, 적어도 약 97 %, 적어도 약 98 % 또는 적어도 약 99 % 동일한 아미노산 서열을 포함하거나, 필수적으로 상기 아미노산 서열로 이루어지거나, 또는 상기 아미노산 서열로 이루어진다.

[0052] 다른 실시예들에서, 폴리펩티드는 N-(4-아미노-2-하이드록실부틸)테트라데칸아미드 합성효소(AhtS) 폴리펩티드이다. 여하한 실시예들에서, AhtS 폴리펩티드는 SEQ ID NO: 3의 아미노산 서열, 즉 GenBank 수탁 번호

ACX33975의 아미노산 서열을 포함하거나, 필수적으로 상기 아미노산 서열로 이루어지거나, 또는 상기 아미노산 서열로 이루어진다. 일부 실시예들에서, AhtS 폴리펩티드는 SEQ ID NO: 22의 핵산 서열을 포함하는 핵산 서열에 의해 인코딩된다. 다른 실시예들에서, AhtS 폴리펩티드는 SEQ ID NO: 3의 아미노산 서열을 갖는 AhtS 폴리펩티드의 동족체이다. 바람직하게, AhtS 폴리펩티드는 SEQ ID NO: 3의 아미노산 서열에 대하여 적어도 약 70 %, 적어도 약 75 %, 적어도 약 80 %, 적어도 약 85 %, 적어도 약 90 %, 적어도 약 91 %, 적어도 약 92 %, 적어도 약 93 %, 적어도 약 94 %, 적어도 약 95 %, 적어도 약 96 %, 적어도 약 97 %, 적어도 약 98 % 또는 적어도 약 99 % 동일한 아미노산 서열을 포함하거나, 필수적으로 상기 아미노산 서열로 이루어지거나, 또는 상기 아미노산 서열로 이루어진다.

[0053] 여하한 실시예들에서, 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드는 미생물에 내부로부터 발생한다. 이러한 실시예들에서, 재조합 미생물은 일차 아민 및 아실 티오에스테르의 지방족 아미드의 전환을 촉진시키는 내인성 폴리펩티드를 과발현시키도록 조작된다.

[0054] 다른 실시예들에서, 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키는 폴리펩티드는 미생물에 외부에서 발생한다(exogenous). 이러한 실시예들에서, 재조합 미생물은 외인성 폴리펩티드를 발현시키도록 조작되어, 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시키도록 한다. 예를 들어, 외인성 폴리펩티드를 인코딩하는 외인성 핵산은 표준 분자 생물학 절차를 통하여 미생물 내로 통합(integrated)될 수 있다. 탄소 공급원과 함께 미생물을 제공하는 것은 미생물이 지방족 아미드 생산을 증가시킬 수 있게 할 것이다.

[0055] 본 명세서에서 사용되는 바와 같이, "상동의(homolog)", "동족체(homologue)" 및 "상동(homologous)"이라는 용어들은 대응하는 폴리뉴클레오티드 또는 폴리펩티드 서열에 대하여 적어도 약 70 % 상동하는 서열을 포함하는 폴리뉴클레오티드 또는 폴리펩티드를 칭한다. 본 기술분야의 당업자는 두 개 이상의 서열들 간의 상동성(homology)을 결정하기 위한 방법들을 익히 알고있다. 예를 들어, 두 개의 서열들 간의 퍼센트 상동성의(percent homology) 결정 및 서열들의 비교는 Basic Local Alignment Search Tool(BLAST)과 같은 수학적 알고리즘을 사용하여 이뤄질 수 있다[Altschul *et al.* (1990) *J. Mol. Biol.* 215(3):403-410].

[0056] "폴리뉴클레오티드"라는 용어는 DNA 또는 RNA의 중합체를 칭하고, 이는 외가닥 또는 이중 가닥일 수 있으며, 비-천연 또는 변화된(altered) 뉴클레오티드들을 포함할 수 있다. "폴리뉴클레오티드", "핵산" 및 "핵산 분자"이라는 용어들은 본 명세서에서 여하한 길이의 뉴클레오티드의 중합 형태, 리보뉴클레오티드(RNA) 또는 디옥시리보뉴클레오티드(DNA) 중 하나를 칭하는 것으로 통용되어 사용된다.

[0057] "폴리펩티드" 및 "단백질"이라는 용어들은 아미노산 잔기들의 중합체를 칭한다. "재조합 폴리펩티드"라는 용어는 재조합 DNA 기술들에 의해 생산된 폴리펩티드를 칭하고, 여기서 일반적으로 발현되는 단백질을 인코딩하는 DNA 또는 RNA는 결국 폴리펩티드 또는 RNA를 생산하기 위해 숙주 세포를 형질전환시키는데 사용되는 적합한 발현 벡터 내로 삽입된다.

[0058] 본 발명의 방법들 및 조성물들에 있어서, 원하는 지방산 또는 이의 아실 티오에스테르 유도체의 생산은 재조합 미생물 내에서 지방산 생산, 분해(degradation) 및/또는 분비의 조절(regulation)에 관련된 하나 이상의 유전자들의 발현을 변화(altering)시킴으로써 향상될 수 있다.

[0059] 일부 실시예들에서, 재조합 미생물은 지방산 생합성 폴리펩티드를 인코딩하는 핵산 서열을 포함한다. 본 명세서에서 사용되는 바와 같이, "지방산 생합성 폴리펩티드"라는 용어는 지방산 생합성에 관련된 여하한 폴리펩티드를 칭한다. 숙주 세포들 내에서의 지방산 생합성 경로는 전구체들 아세틸-CoA 및 말로닐-CoA를 사용한다. 이러한 경로에 대한 단계들은 지방산 생합성의 효소들(*fab*) 및 아세틸-CoA 카르복실화효소(*acc*) 유전자 족(family)들에 의해 촉진된다[예를 들어, Heath *et al.* (2001) *Prog. Lipid Res.* 40(6):467-497 참조]. 아세틸-CoA는 아세틸-CoA 카르복실화효소(EC 6.4.1.2)에 의해 카르복실화되어 말로닐-CoA를 형성한다. 아세틸-CoA 카르복실화효소(EC 6.4.1.2)는 대부분의 원핵생물 내에서 네 개의 개별적인 유전자들(*accA*, *accB*, *accC* 및 *accD*)에 의해 인코딩되는 다중-서브유닛(multi-subunit) 효소이다. *코리네박테리움 글루타미쿠스(Corynebacterium glutamicus)*와 같은 일부 박테리아에서, 아세틸-CoA 카르복실화효소는 각각 *accDA* 및 *accBC*에 의해 인코딩되는 두 개의 서브유닛들 *AccDA*[YP_225123.1] 및 *AccBC*[YP_224991]을 포함한다. 원하는 지방산 또는 지방산 유도체 생성물에 따라서, 특정 *fab* 및/또는 *acc* 유전자들 (또는 이들의 조합들)은 조작된 숙주 세포 내에서 과발현, 변형(modified), 감쇠(attenuated) 또는 결손(deleted)될 수 있다.

[0060] 일부 실시예들에서, 지방산 생합성 폴리펩티드를 인코딩하는 핵산 서열은 *accABCD*를 인코딩한다. 다른 실시예

들에서, 지방산 생합성 폴리펩티드를 인코딩하는 핵산 서열은 FabD, FabH, FabG, FabB, FabA, FabZ, FabF, FabI, 또는 FabV와 같은 다른 유기체로부터 Fab의 기능적 동족체(functional homologue)를 인코딩한다. 본 발명의 방법들 및 조성물들에 사용하기에 적합한 지방산 생합성 폴리펩티드들에 대한 예시적인 GenBank 수탁 번호들은 FabD(AAC74176), FabH(AAC74175), FabG(AAC74177), FabB(POA953), FabA(ACY27485.1), FabZ(ACY27493.1), FabF(AAC74179) 및 FabI(NP_415804)를 포함한다.

[0061] 일부 실시예들에서, 재조합 미생물은 두 개 이상(예를 들어, 3 개 이상, 4 개 이상)의 생합성 폴리펩티드들을 인코딩하는 핵산 서열들(예를 들어, accABCD 및 FabD; FabD, FabH 및 FabG; 또는 FabI, FabG,H,D, FabA,B 및 FabZ)을 포함한다.

[0062] FadR은 지방산 분해 및 지방산 생합성 경로들에 관련된 전사 인자이다[Cronan *et al.* (1998) *Mol. Microbiol.* 29(4):937-943]. FadR은 *fabA*, *fabB*, *ic1R*, *fadA*, *fadB*, *fadD*, *fadE*, *fadI*, *fadJ*, *fadL*, *fadM*, *uspA*, *aceA*, *aceB* 및 *aceK*를 포함하는, 수많은 유전자들의 활성 및/또는 발현을 조절(modulate)하는 것으로 알려져 있다. FadR 표적 유전자들에 의해 인코딩되는 폴리펩티드들에 대한 예시적인 GenBank 수탁 번호들은 *fabA*(NP_415474), *fabB*(BAA16180), (NP_418442), *fadA*(YP_026272.1), *fadB*(NP_418288.1), *fadD*(AP_002424), *fadE*(NP_414756.2), *fadI*(NP_416844.1), *fadJ*(NP_416843.1), *fadL*(AAC75404), *fadM*(NP_414977.1), *uspA*(AAC76520), *aceA*(AAC76985.1), *aceB*(AAC76984.1) 및 *aceK*(AAC76986.1)를 포함한다.

[0063] 일부 실시예들에서, 재조합 미생물은 지방산 생합성 폴리펩티드를 인코딩하는 핵산 서열, 및 FadR을 인코딩하는 핵산 서열을 포함한다. 여하한 실시예들에서, 핵산 서열은 *대장균* MG1655로부터 FadR(NP_415705)을 인코딩한다.

[0064] 티오에스테라아제들(EC 3.1.2.14 또는 EC 3.1.1.5)은 아실-ACP 티오에스테르들로부터 지방산들을 가수분해한다. 아실 티오에스테르 기질의 사슬 길이는 선택된 티오에스테라아제들의 발현을 변형시킴으로써 선택될 수 있다. 여하한 실시예들에서, 숙주 세포는 바람직한 지방산 유도체 기질의 생산을 증가시키기 위하여 하나 이상의 선택된 티오에스테라아제들을 발현시키거나, 과발현시키거나, 감소된 발현을 갖거나, 또는 발현시키지 않도록 조작된다. 예를 들어, C₁₀ 지방산은 C₁₀ 지방산 생산에 선호되는 티오에스테라아제를 발현시키고 C₁₀ 지방산보다 다른 지방산 생산에 선호되는 티오에스테라아제(예를 들어, C₁₄ 지방산 생산에 선호되는 티오에스테라아제)를 감소시킴으로써 생산될 수 있다. 이는 10 개의 탄소들을 포함하는 지방산의 비교적 동종 개체군(homogeneous population)을 야기한다. 다른 실시예들에서, C₁₄ 지방산은 비-C₁₄ 지방산을 생산하는 내인성 티오에스테라아제를 감소시키고 C₁₄-ACP에 선호되는 티오에스테라아제를 발현시킴으로써 생산될 수 있다. 일부 경우들에서, C₁₂ 지방산은 C₁₂-ACP에 선호되는 티오에스테라아제를 발현시키고 우선적으로 비-C₁₂ 지방산을 생산하는 티오에스테라아제를 감소시킴으로써 생산될 수 있다. 아세틸-CoA, 말로닐-CoA 및 지방산의 과생산(overproduction)은, 예를 들어 GC-MS, HPLC 또는 방사성 전구체들을 사용함으로써 본 기술분야에 알려진 방법들을 이용하여 입증될 수 있다.

[0065] 본 발명의 방법들 및 조성물들에 있어서 티오에스테라아제 유전자들의 발현이 변화될 수 있는 티오에스테라아제 유전자들 [및 대응하는 GenBank 수탁 번호(들)]의 비-제한적인 예시들은, *대장균*으로부터의 선도 서열(leader sequence)이 없는 *tesA*(*tesA*)(AAC73596), *대장균*으로부터의 *tesB*(AAC73555), *캘리포니아만월계수* (*Umbellularia californica*)로부터의 *fatB*(Q41635, AAA34215), *쿠페아 후커리아나*(*Cuphea hookeriana*)로부터의 *fatB2*(AAC49269), *쿠페아 후커리아나*로부터의 *fatB3*(Q39513; AAC72881), *시나모눔 캄포럼*(*Cinnamomum camphorum*)으로부터의 *fatB*(Q39473, AAC49151), *애기장대*(*Arabidopsis thaliana*)로부터의 *fatB*[M141T](CAA85388)[Mayer *et al.* (2007) *BMC Plant Biology* 7:1-11], *애기장대*로부터의 *fatA*(NP 189147; NP 193041), *브래디리조븀 자포니쿰*(*Bradyrhizobium japonicum*)으로부터의 *fatA*(CAC39106), *쿠페아 후커리아나*로부터의 *fatA*(AAC72883) 및 *해바라기*(*Helianthus annuus*)로부터의 *fatA1*(AAL79361)을 포함한다.

[0066] 여하한 실시예들에서, 재조합 미생물은 티오에스테라아제를 인코딩하는 핵산 서열을 포함하고, 상기 핵산 서열은 *대장균* MG1655로부터의 '*tesA*(AAC73596)이다.

[0067] 아실-CoA 합성효소(EC 2.3.1.86)는 아실-CoA 티오에스테르의 형성을 촉진시킴으로써 지방산을 활성화시킨다. 본 발명의 방법들 및 조성물들에 있어서 아실-CoA 합성효소 유전자들의 발현이 변화될 수 있는 아실-CoA 합성효소 유전자들의 비-제한적인 예시들은, *fadD*, *fadK*, *BH3103*, *yhfL*, *Pfl-4354*, *EAV15023*, *fadD1*, *fadD2*, *RPC_4074*, *fadDD35*, *fadDD22*, *faa3p* 또는 단백질 ZP_01644857를 인코딩하는 유전자를 포함한다. 아실-CoA 합성효소 유전자들의 구체적인 예시들은 *결핵균*(*M. tuberculosis*) H37Rv로부터의 *fadDD35*[NP_217021], *결핵균*

H37Rv로부터의 *fadDD22*[NP_217464], 대장균으로부터의 *fadD*[NP_416319], 대장균으로부터의 *fadK*[YP_416216], 아시네토박터 중(*Acinetobacter sp.*) ADP1로부터의 *fadD*[YP_045024], 헤모필러스 인플루엔자(*Haemophilus influenza*) RdkW20으로부터의 *fadD*[NP_438551], 로도슈도모나스 파루스트리(*Rhodopseudomonas palustris*) Bis B18로부터의 *fadD*[YP_533919], 바실러스 할로두란스(*Bacillus halodurans*) C-125로부터의 *BH3101*[NP_243969], 슈도모나스 플루오르스스(*Pseudomonas fluorescens*) Pfo-1로부터의 *Pfl-4354*[YP_350082], 코마모나스 테스트스 테론(*Comamonas testosterone*) KF-1로부터의 *EAV15023*[ZP_01520072], *B. 서브틸리스(subtilis)*로부터의 *yhfL*[NP_388908], 녹농균(*P. aeruginosa*) PA01로부터의 *fadDI*[NP_251989], 랄스토니아 솔라나세아룸(*Ralstonia solanacearum*) GM1 1000으로부터의 *fadDI*[NP_520978], 녹농균 PA01로부터의 *fadD2*[NP_251990], 스테노트로포모나스 말토티리아(*Stenotrophomonas maltophilia*) R551-3으로부터의 단백질 ZP_01644857을 인코딩하는 유전자, 맥주효모균으로부터의 *faa3p*[NP_012257], 맥주효모균으로부터의 *faa1p*[NP_014962], 바실러스 서브틸리스로부터의 *Icfa*[CAA99571], 그리고 Shockey *et al.* (2002) *Plant. Physiol.* 129:1710-1722); Caviglia *et al.* (2004) *J. Biol. Chem.* 279:1163-1169; Knoll *et al.* (1994) *J. Biol. Chem.* 269(23):16348-56); Johnson *et al.* (1994) *J. Biol. Chem.* 269:18037-18046; 및 Black *et al.* (1992) *J. Biol. Chem.* 267:25513-25520에 기재된 것들을 포함한다.

- [0068] 일부 실시예들에서, 재조합 미생물은 아실-CoA 합성효소 폴리펩티드를 인코딩하는 핵산 서열을 포함하고, 상기 아실-CoA 합성효소 폴리펩티드는 대장균 MG1655로부터의 *FadD*[NP_416319]이다.
- [0069] 재조합 미생물은 지방산 생합성 폴리펩티드들, 티오에스테라아제 폴리펩티드들 및 아실-CoA 합성효소 폴리펩티드들의 여하한 조합을 인코딩하는 핵산들을 포함할 수 있다. 여하한 실시예들에서, 미생물은 티오에스테라아제 폴리펩티드를 인코딩하는 핵산 서열, 및 아실-CoA 합성효소 폴리펩티드를 인코딩하는 핵산 서열을 포함한다.
- [0070] 본 기술분야의 당업자들은 목적에 따라서(예를 들어, 원하는 지방산 또는 이의 아실 티오에스테르 유도체), 지방산 물질대사에 관련된 구체적인 유전자들 (또는 유전자들의 조합들)이 일차 아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시킬 수 있는 폴리펩티드를 인코딩하는 핵산 서열을 포함하도록 조작된 재조합 미생물 내에서 과발현, 변형, 감쇠 또는 결실될 수 있다는 것을 이해한다. 본 발명에서 사용하기에 적합한 지방산 물질대사에 관련된 유전자들의 추가적인 예시들은 예를 들어, 미국 특허 출원 공개 2011/0162259에 기재되어 있고, 이는 본 명세서에서 전문이 인용참조된다.
- [0071] 일부 실시예들에서, 폴리펩티드는 본 명세서에서 설명되는 폴리펩티드들의 여하한 분절(fragment)이다. "분절"이라는 용어는 네 개의 아미노산 잔기 내지 전체 아미노산 서열에서 하나의 아미노산 잔기를 제외한 크기에 이르는 전체 길이의(full-length) 폴리펩티드 또는 단백질의 짧은 부분을 칭한다. 본 발명의 여하한 실시예들에서, 분절은 폴리펩티드 또는 단백질의 도메인(domain)[예를 들어, 기질 결합 도메인 또는 촉매 도메인]의 전체 아미노산 서열을 칭한다.
- [0072] 일부 실시예들에서, 폴리펩티드는 본 명세서에서 설명되는 여하한 폴리펩티드들의 돌연변이체 또는 변이체이다. 본 명세서에서 사용되는 바와 같은 "돌연변이체" 및 "변이체"라는 용어들은 적어도 하나의 아미노산에 의해 야생형 폴리펩티드와 상이한 아미노산 서열을 갖는 폴리펩티드를 칭한다. 예를 들어, 돌연변이체는 하기의 보존성 아미노산 치환체들 중에서 하나 이상을 포함할 수 있다: 알라닌, 발린, 류신 및 이소류신과 같은 지방족(aliphatic) 아미노산의 또 다른 지방족(aliphatic) 아미노산으로의 대체물(replacement); 세린의 트레오닌으로의 대체물; 트레오닌의 세린으로의 대체물; 아스파르트산 및 글루탐산과 같은 산성 잔기의 또 다른 산성 잔기로의 대체물; 아스파라긴 및 글루타민과 같은 아미드기를 보유하는 잔기의 아미드기를 보유하는 또 다른 잔기로의 대체물; 리신 및 아르기닌과 같은 염기성 잔기의 또 다른 염기성 잔기로의 교환물(exchange); 그리고 페닐알라닌 및 티로신과 같은 방향족 잔기의 또 다른 방향족 잔기로의 대체물. 일부 실시예들에서, 돌연변이형 폴리펩티드는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100 또는 그 이상의 아미노산 치환들, 부가(addition)들, 삽입들 또는 결실들을 갖는다.
- [0073] 폴리펩티드의 바람직한 분절들 또는 돌연변이체들은 대응하는 야생형 폴리펩티드의 생물학적 기능(예를 들어, 효소 활성) 중 일부 또는 전부를 유지한다. 일부 실시예들에서, 분절 또는 돌연변이체는 대응하는 야생형 폴리펩티드의 생물학적 기능 중 적어도 약 75 %, 적어도 약 80 %, 적어도 약 90 %, 적어도 약 95 % 또는 적어도 약 98 % 또는 그 이상을 유지한다. 다른 실시예들에서, 분절 또는 돌연변이체는 대응하는 야생형 폴리펩티드의 생물학적 기능의 약 100 %를 유지한다. 아미노산 잔기들이 생물학적 활성에 영향을 미치지 않고 치환, 삽입 또는 결실될 수 있는지를 결정하기 위한 지침(Guidance)은, 본 기술분야에 잘 알려져 있는 컴퓨터 프로그램, 예를 들어 LASERGENE™ 소프트웨어(DNASTAR, Inc., Madison, WI)를 사용하여 알아낼 수 있다.

- [0074] 또 다른 실시예들에서, 분절 또는 돌연변이체는 "증가된 활성 수준"을 갖는다. "증가된 활성 수준"이란 폴리펩티드가 동일한 조건들 하에서 대응하는 야생형 숙주 세포의 생화학적 및/또는 생물학적 기능의 수준과 비교하여 조작된 세포 내에서 더 높은 수준의 생화학적 또는 생물학적 기능(예를 들어, DNA 결합 또는 효소 활성)을 갖는 것을 의미한다. 활성의 향상 정도는 약 10 % 이상, 약 20 % 이상, 약 50 % 이상, 약 75 % 이상, 약 100 % 이상, 약 200 % 이상, 약 500 % 이상, 약 1000 % 이상, 또는 이 안에서의 여하한 범위일 수 있다.
- [0075] 일부 실시예들에서, 변화된 또는 변형된 발현 수준을 갖는 폴리펩티드 또는 폴리뉴클레오티드는 "과발현"되거나 또는 "증가된 발현 수준"을 갖는다. 본 명세서에서 사용되는 바와 같이, "과발현하다" 및 "증가된 발현 수준"은 동일한 조건들 하에서 대응하는 야생형 세포 내에서 정상적으로 발현되는 것보다 더 큰 농도로 조작된 세포 내에서 폴리펩티드 또는 폴리뉴클레오티드가 발현되도록 야기시키거나 또는 발현하는 것을 의미한다. 예를 들어, 폴리펩티드가 동일한 조건들 하에서 동일한 종류의 비-조작된 숙주 세포 내에서의 농도와 비교하여 조작된 세포 내에서 더 큰 농도로 존재하는 경우에, 폴리펩티드는 조작된 세포 내에서 "과발현"될 수 있다.
- [0076] 다른 실시예들에서, 변화된 발현 수준을 갖는 폴리펩티드 또는 폴리뉴클레오티드는 "감쇠"되거나 "감소된 발현 수준"을 갖는다. 본 명세서에서 사용되는 바와 같이, "감쇠하다" 및 "감소된 발현 수준"은 동일한 조건들 하에서 대응하는 야생형 세포 내에서 정상적으로 발현되는 것보다 더 적은 농도로 조작된 세포 내에서 폴리펩티드 또는 폴리뉴클레오티드가 발현되도록 야기시키거나 또는 발현하는 것을 의미한다.
- [0077] 과발현 또는 감쇠 정도는 1.5 배 이상, 예를 들어 2 배 이상, 3 배 이상, 5 배 이상, 10 배 이상 또는 15 배 이상일 수 있다. 대안적으로 또는 추가적으로, 과발현 또는 감쇠 정도는 500 배 이하, 예를 들어 100 배 이하, 50 배 이하, 25 배 이하 또는 20 배 이하일 수 있다. 따라서, 과발현 또는 감쇠 정도는 상기의 종단점(endpoint)들 중 여하한 두 개로 제한될 수 있다. 예를 들어, 과발현 또는 감쇠 정도는 1.5 내지 500 배, 2 내지 50 배, 10 내지 25 배 또는 15 내지 20 배일 수 있다.
- [0078] 폴리뉴클레오티드 또는 폴리펩티드는 본 기술분야에 알려진 방법들을 사용하여 감쇠될 수 있다. 일부 실시예들에서, 유전자에 의해 인코딩되는 폴리펩티드 또는 유전자의 발현은 유전자의 발현을 제어하는 조절 폴리뉴클레오티드 서열들을 돌연변이시킴으로써 감쇠된다. 다른 실시예들에서, 유전자에 의해 인코딩되는 폴리펩티드 또는 유전자의 발현은 억제 단백질(repressor protein)을 과발현시킴으로써, 또는 억제 단백질을 활성화시키는 외인성 조절 요소(regulatory element)를 제공함으로써 감쇠된다. 또 다른 실시예들에서, DNA- 또는 RNA-기반 유전자 사일런싱 방법(silencing method)들은 폴리뉴클레오티드 또는 유전자의 발현을 감쇠시키는데 사용된다. 일부 실시예들에서, 폴리펩티드 또는 유전자의 발현은, 예를 들어 유전자의 폴리뉴클레오티드 서열의 전체 또는 일부를 결실시킴으로써 완전히 감쇠된다.
- [0079] 폴리뉴클레오티드 또는 폴리펩티드는 본 기술분야에 알려진 방법들을 사용하여 과발현될 수 있다. 일부 실시예들에서, 폴리펩티드의 과발현은 외인성 조절 요소의 사용으로 달성된다. 일반적으로, "외인성 조절 요소"라는 용어는 숙주 세포의 밖에서 유래하는 조절 요소를 칭한다. 그러나, 여하한 실시예들에서, "외인성 조절 요소"라는 용어는 내인성 폴리펩티드의 발현을 제어할 목적으로 기능이 복제되거나 빼앗긴(usurped) 숙주 세포로부터 유래되는 조절 요소를 칭할 수 있다. 예를 들어, 만약 재조합 미생물이 지방산 생합성 폴리펩티드를 인코딩하는 핵산 서열을 포함하는 *대장균* 세포이고, 상기 지방산 생합성 폴리펩티드가 내인성 *fadR* 유전자에 의해 인코딩되는 FadR이라면, 내인성 *fadR*의 발현은 또 다른 *대장균* 유전자로부터 유래되는 프로모터에 의해 제어될 수 있다.
- [0080] 일부 실시예들에서, 외인성 조절 요소는 소분자(small molecule)와 같은 화합물이다. 본 명세서에서 사용되는 바와 같이, "소분자"라는 용어는 약 1,000 g/mol 미만의 분자량을 갖는 화합물 또는 물질을 칭한다.
- [0081] 일부 실시예들에서, 핵산 서열의 발현을 제어하는 외인성 조절 요소는 핵산 서열에 작동가능하게 연결된(operably linked) 발현 제어 서열(expression control sequence)이다. 발현 제어 서열들은 본 기술분야에 알려져 있고, 예를 들어 숙주 세포 내에서 핵산 서열의 발현을 위하여 제공되는 프로모터, 증폭자(enhancers), 아데닐산중합반응 신호(polyadenylation signals), 전사 종결자(transcription terminators), 내부 리보솜 유입점(internal ribosome entry sites: IRES), 리보솜 결합 부위(ribosome binding sites: RBS) 및 이와 유사한 것들을 포함한다. 발현 제어 서열들은 특히 전사에 관련된 세포성 단백질들과 상호작용을 한다[Maniatis *et al.* (1987) *Science* 236:1237-1245]. 예시적인 발현 제어 서열들은, 예를 들어 Goeddel, *Gene Expression Technology: Methods in Enzymology*, Vol. 185, Academic Press, San Diego, Calif.(1990)에 기재된다.
- [0082] "작동가능하게 연결된"이란 적절한 분자들(예를 들어, 전사 활성화인자 단백질들)이 발현 제어 서열(들)에 결합

되는 경우에, 핵산 서열 및 발현 제어 서열(들)이 유전자 발현을 가능하게 하는 방식으로 연결되는 것을 의미한다. 작동가능하게 연결된 프로모터들은 전사 및 번역의 방향에 관하여 선택된 핵산 서열의 상류(upstream)에 위치된다. 작동가능하게 연결된 증폭자들은 선택된 핵산 서열의 상류, 내부(within) 또는 하류(downstream)에 위치될 수 있다.

[0083] 일부 실시예들에서, 핵산 서열은 재조합 벡터를 거쳐 숙주 세포로 제공되고, 이는 폴리뉴클레오티드 서열에 작동가능하게 연결된 프로모터를 포함한다. 여하한 실시예들에서, 프로모터는 유도적(inducible), 구조적(constitutive) 또는 세포소기관(organelle)-특이적 프로모터이다. 여하한 실시예들에서, 발현 제어 서열은 본 기술분야에 알려진 방법들[예를 들어, Datsenko et al.(2000) *Proc. Natl. Acad. Sci. U.S.A.* 97(12):6640-6645]을 사용하여 상동 재조합(homologous recombination)에 의해 숙주 세포의 게놈 내로의 발현 제어 서열의 통합에 의하여 내인성 핵산 서열에 작동가능하게 연결된다.

[0084] 본 명세서에서 사용되는 바와 같이, "벡터"라는 용어는 핵산 분자에 연결되는 또 다른 핵산 서열을 이송할 수 있는 핵산 분자를 칭한다. 유용한 벡터 중 한 가지 타입은 에피솜(episome)[즉, 염색체의 복제(extra-chromosomal replication)가 가능한 핵산]이다. 유용한 벡터는 벡터와 연결되는 핵산들의 자동 복제 및/또는 발현이 가능한 벡터이다. 작동가능하게 연결되는 유전자들의 발현을 지시할 수 있는 벡터는 본 명세서에서 "발현 벡터"로 칭한다. 일반적으로, 재조합 DNA 기술들에서 유용한 발현 벡터는 종종 "플라스미드"의 형태이며, 이는 일반적으로 이의 벡터 형태에 있어서 염색체에 결합되지 않는 원형 이중 가닥 DNA 루프(loops)를 칭한다. 그러나, 등가적인 기능들을 수행하고 이후에 본 기술분야에 알려지는 다른 형태들의 발현 벡터도 포함된다.

[0085] 일부 실시예들에서, 재조합 벡터는 (a) 핵산 서열에 작동가능하게(operatively) 연결된 발현 제어 서열; (b) 핵산 서열에 작동가능하게 연결된 선택 마커; 및 (c) 핵산 서열에 작동가능하게 연결된 표적화 서열(targeting sequence)로 이루어진 그룹에서 선택되는 적어도 하나의 서열을 포함한다.

[0086] 발현 벡터의 설계는 형질전환되는 숙주 세포의 선택, 원하는 폴리펩티드의 발현 수준 등과 같은 요인들에 의존할 수 있다는 것을 본 기술분야의 당업자들은 인지할 것이다. 본 명세서에 설명된 발현 벡터들은 본 명세서에 설명되는 바와 같이 폴리뉴클레오티드 서열들에 의해 인코딩되는, 융합 폴리펩티드들을 포함하는 폴리펩티드들을 생산하기 위하여 숙주 세포들에 도입될 수 있다.

[0087] 원핵 및 진핵 세포들 둘 모두에 대한 적합한 발현 시스템들은 본 기술분야에 잘 알려져 있다; 예를 들어, Sambrook et al., "Molecular Cloning: A Laboratory Manual", second edition, Cold Spring Harbor Laboratory(1989)를 참조한다. 유도성(inducible), 비-융합 대장균 발현 벡터의 예시들은 pTrc[Amann et al., (1988) *Gene* 69:301-315] 및 pET 11d[Studier et al., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA, pp.60-89(1990)]를 포함한다. 여하한 실시예들에서, 본 발명의 폴리뉴클레오티드 서열은 박테리오파지 T5로부터 유래된 프로모터에 작동가능하게 연결된다. 효모 내에서의 발현을 위한 벡터의 예시들은 pYepSec1[Baldari et al. (1987) *EMBO J.* 6:229-234], pMFa[Kurjan et al. (1982) *Cell* 30:933-943], pJRY88[Schultz et al. (1987) *Gene* 54:113-123], pYES2(Invitrogen Corp., San Diego, CA) 및 picZ(Invitrogen Corp., San Diego, CA)를 포함한다.

[0088] 벡터들은 종래의 형질전환 또는 형질감염(transfection) 기술들을 통하여 원핵 및 진핵 세포들 내로 도입될 수 있다. 본 명세서에서 사용되는 바와 같이, "형질전환" 및 "형질감염"이라는 용어들은 외래(foreign) 핵산(예를 들어, DNA)을 숙주 세포 내로 도입하기 위하여 본 기술분야에서 인정되는 다양한 기술들을 칭하며, 여기에는 인산칼슘 또는 염화칼슘 공침(co-precipitation), DEAE-텍스트란 매개 형질감염(DEAE-dextran-mediated transfection), 리포펙션(lipofection) 또는 전기천공(electroporation)을 포함한다. 벡터를 흡수(taken up)한 세포들을 선택하기 위한 방법들뿐만 아니라 숙주 세포를 형질전환 또는 형질감염시키는데 적합한 방법들은, 예를 들어 Sambrook et al.(*supra*)에서 찾아볼 수 있다.

[0089] "재조합 미생물"은 본 명세서에서 설명되는 생성물(예를 들어, 지방족 아미드)을 생산하는데 사용되는 숙주 세포이다. "재조합 숙주 세포", "조작된 미생물" 또는 "조작된 숙주 세포"라고도 칭하는 재조합 미생물은, 하나 이상의 핵산들 또는 폴리펩티드들의 발현이 동일한 조건들 하에서 대응하는 야생형 숙주 세포 내에서의 발현과 비교하여 변화 또는 변형된 숙주 세포이다. 본 명세서에서 설명되는 본 발명의 여하한 측면들에서, 숙주 세포는 박테리아 세포, 남조류 세포, 조류(algae) 세포 및 균류 세포(예를 들어, 곰팡이 세포 또는 효모 세포)를 포함할 수 있으나, 이로 제한되지 않는다.

[0090] 일부 실시예들에서, 숙주 세포는 그람-양성(Gram-positive) 박테리아 세포이다. 여타 실시예들에서, 숙주 세포

는 그람-음성(Gram-negative) 박테리아 세포이다.

- [0091] 일부 실시예들에서, 숙주 세포는 *에스체리키아(Escherichia)* 속, *바실러스(Bacillus)* 속, *락토바실러스(Lactobacillus)* 속, *로도코커스(Rhodococcus)* 속, *슈도모나스(Pseudomonas)* 속, *아스페르길러스(Aspergillus)* 속, *트리코더마(Trichoderma)* 속, *뉴로스포라(Neurospora)* 속, *푸사리움(Fusarium)* 속, *후미콜라(Humicola)* 속, *리조무코(Rhizomucor)* 속, *클루이베로마이세스(Kluyveromyces)* 속, *피치아(Pichia)* 속, *무코(Mucor)* 속, *미셀리오프토라(Myceliophthora)* 속, *페니실리움(Penicillium)* 속, *파네로차에테(Phanerochaete)* 속, *느타리(Pleurotus)* 속, *트라메테스(Trametes)* 속, *크리소스포리움(Chrysosporium)* 속, *사카로마이세스(Saccharomyces)* 속, *스테노트로파모나스(Stenotrophomonas)* 속, *스키조사카로마이세스(Schizosaccharomyces)* 속, *시네키크코커스(Synechococcus)* 속, *야로위아(Yarrowia)* 속 또는 *스트렙토마이세스(Streptomyces)* 속으로부터 선택된다.
- [0092] 여하한 실시예들에서, 숙주 세포는 *맥주효모균(Saccharomyces cerevisiae)*, *캔디다 리폴리티카(Candida lipolytica)*, *대장균(Escherichia coli)*, *아트로박터(Arthrobacter)*, *로도토룰라 글루티닌스(Rhodotorula glutinins)*, *아시네토박터(Acinetobacter)*, *캔디다 리폴리티카(Candida lipolytica)*, *보트리오코커스 브라우니(Botryococcus braunii)*, *비브리오 퍼니썬(Vibrio furnissii)*, *마이크로코커스 로이테우스(Micrococcus leuteus)*, *스테노트로포모나스 말토펠리아(Stenotrophomonas maltophilia)*, *바실러스 서브틸리스(Bacillus subtilis)*, *바실러스 리첸노포르미스(Bacillus licheniformis)*, *슈도모나스 푸티다(Pseudomonas putida)*, *슈도모나스 플로레스센스(Pseudomonas floescens)*, *스트렙토마이세스 코엘리콜라(Streptomyces coelicolor)*, *프로토테카 모리포르미스(Prototheca moriformis)*, *프로토테카 크루가니(Prototheca krugani)*, *프로토테카 스테그노라(Prototheca stagnora)*, *프로토테카 조프피(Prototheca zopfii)* 또는 *콜레라 프로토썬코이드(Chorella protothecoide)* 세포이다.
- [0093] 일부 실시예들에서, 숙주 세포는 *아트로박터 AK 19*, *아시네토박터 종(Acinetobacter sp.)* 균주 M-1, *대장균 B*, *대장균 C*, *대장균 K* 또는 *대장균 W* 세포이다.
- [0094] 다른 실시예들에서, 숙주 세포는 *바실러스 렌투스(Bacillus lentus)* 세포, *바실러스 브레비스(Bacillus brevis)* 세포, *바실러스 스테아로썬오피러스(Bacillus stearothermophilus)* 세포, *바실러스 리첸노포르미스(Bacillus licheniformis)* 세포, *바실러스 알칼로필루스(Bacillus alkalophilus)* 세포, *바실러스 코아굴란스(Bacillus coagulans)* 세포, *바실러스 키르쿨란스(Bacillus circulans)* 세포, *바실러스 푸밀리스(Bacillus pumilis)* 세포, *바실러스 튜링기엔시스(Bacillus thuringiensis)* 세포, *바실러스 클라우시(Bacillus clausii)* 세포, *바실러스 메가테리움(Bacillus megaterium)* 세포 또는 *바실러스 아밀로리퀴파시엔스(Bacillus amyloliquefaciens)* 세포이다.
- [0095] 다른 실시예들에서, 숙주 세포는 *트리코더마 코닝기(Trichoderma koningii)* 세포, *트리코더마 비리데(Trichoderma viride)* 세포, *트리코더마 르에세이(Trichoderma reesei)* 세포, *트리코더마 롱기브라키아툼(Trichoderma longibrachiatum)* 세포, *아스페르길러스 아와모리(Aspergillus awamori)* 세포, *아스페르길러스 푸미가테스(Aspergillus fumigates)* 세포, *아스페르길러스 포에티두스(Aspergillus foetidus)* 세포, *아스페르길러스 니둘란스(Aspergillus nidulans)* 세포, *아스페르길러스 니게르(Aspergillus niger)* 세포, *아스페르길러스 오리제(Aspergillus oryzae)* 세포, *후미콜라 인솔렌스(Humicola insolens)* 세포, *후미콜라 라누기노세(Humicola lanuginosa)* 세포, *로도코커스 오파쿠스(Rhodococcus opacus)* 세포, *리조무코 미에헤이(Rhizomucor miehei)* 세포 또는 *무코 미에헤이(Mucor michei)* 세포이다.
- [0096] 또 다른 실시예들에서, 숙주 세포는 숙주 세포는 *스트렙토마이세스 리비단스(Streptomyces lividans)* 세포 또는 *스트렙토마이세스 무리누스(Streptomyces murinus)* 세포이다.
- [0097] 또 다른 실시예들에서, 숙주 세포는 *방선균(Actinomycetes)* 세포이다.
- [0098] 다른 실시예들에서, 숙주 세포는 진핵 식물(eukaryotic plant), 조류, 남세균(cyanobacterium), 녹색-황 세균, 녹색 비-황(green non-sulfur) 세균, 자색 황세균, 자색 비-황 세균, 극한성 생물(extremophile), 효모, 균류, 이들의 조작된 유기체 또는 합성 유기체로부터의 세포이다. 일부 실시예들에서, 숙주 세포는 광 의존성이거나 탄소를 고정시킨다. 일부 실시예들에서, 숙주 세포는 광 의존성이거나 탄소를 고정시킨다. 일부 실시예들에서, 숙주 세포는 독립영양적 활성을 갖는다. 일부 실시예들에서, 숙주 세포는 빛의 존재 하에서와 같이, 광독립영양적 활성(photoautotrophic activity)을 갖는다. 일부 실시예들에서, 숙주 세포는 빛의 부재 하에서 종속영양적 또는 혼합영양적이다. 여하한 실시예들에서, 숙주 세포는 *애기장대(Avabidopsis thaliana)*,

파니쿰 비르가툼(*Panicum virgatum*), 미스칸투스 기간테우스(*Miscanthus giganteus*), 제아 메이즈(*Zea mays*), 클라미도모나스 레인하르티(*Chlamydomonas reinhardtii*), 두나리엘라 살리나(*Dunaliella salina*), 시네코코커스 종(*Synechococcus Sp.*) PCC 7002, 시네코코커스 종 PCC 7942, 시네코시스티스 종(*Synechocystis Sp.*) PCC 6803, 씨모시네코코커스 엘롱게이트(*Thermosynechococcus elongates*) BP-1, 클로로비움 테피둠(*Chlorobium tepidum*), 클로로프렉서스 아우란티쿠스(*Chloroflexus auranticus*), 크로마티움 비노썸(*Chromatium vinosum*), 로도스피리움 루브룸(*Rhodospirillum rubrum*), 로도박터 캡슐라투스(*Rhodobacter capsulatus*), 로도슈도모나스 파루스리스(*Rhodopseudomonas palustris*), 클로스트리디움 융달리(*Clostridium ljungdahlii*), 클로스트리디우씨 모셀룸(*Clostridiethermocellum*), 페니실리움 크리소게눔(*Penicillium chrysogenum*), 피치아 파스토리스(*Pichia pastoris*), 스킴조사카로마이세스 포롬브(*Schizosaccharomyces pombe*), 슈도모나스 플루오르센스(*Pseudomonas fluorescens*) 또는 자이모모나스 모빌리스(*Zymomonas mobilis*)로부터의 세포이다.

[0099] 본 명세서에서 사용되는 바와 같이, "생산을 위하여 허용되는 조건들"이라는 용어는 숙주 세포가 원하는 생성물, 이를 테면 지방족 아미드를 생산하도록 하는 여하한 조건들을 의미한다. 유사하게, "발현에 적합한 조건들"이라는 용어는 숙주 세포가 폴리펩티드를 합성하도록 하는 여하한 조건들을 의미한다. 예를 들어, 적합한 조건들은 발효 조건들을 포함한다. 발효 조건들은 온도 범위들, 통기 수준들(levels of aeration) 및 배지 조성과 같은 많은 파라미터들을 포함할 수 있다. 이러한 조건들의 각각은 개별적으로 그리고 조합적으로, 숙주 세포가 성장할 수 있도록 한다. 예시적인 배양 배지는 액체 배지(broths) 또는 겔(gels)을 포함한다. 일반적으로, 배지는 숙주 세포에 의해 직접적으로 대사작용될 수 있는 탄소 공급원을 포함한다. 또한, 효소들은 탄소 공급원의 기동(mobilization)[예를 들어, 전분 또는 셀룰로오스의 발효성 당들로의 해중합반응(depolymerization)] 및 차후의 물질대사를 촉진시키기 위하여 배지에 사용될 수 있다.

[0100] 본 명세서에서 사용되는 바와 같이, "탄소 공급원"이라는 구는 원핵 또는 단순한 진핵 세포 성장을 위한 탄소 공급원으로서 사용되는데 적합한 화합물 또는 기질을 칭한다. 탄소 공급원은 중합체, 탄수화물, 산(acid), 알코올, 알데히드, 케톤, 아미노산, 펩티드 및 기체(예를 들어, CO 및 CO₂)를 포함하나 이로 제한되지 않는 다양한 형태로 존재할 수 있다. 예시적인 탄소 공급원은 글루코오스, 프럭토오스, 만노오스, 갈락토오스, 자일로오스 및 아라비노오스와 같은 단당류; 프럭토-올리고당 및 갈락토-올리고당과 같은 올리고당류; 녹말, 셀룰로오스, 펙틴 및 자일란과 같은 다당류; 수크로오스, 말토오스 및 투라노오스(turanose)와 같은 이당류; 메틸 셀룰로오스 및 소듐 카복시메틸 셀룰로오스와 같은 셀룰로오스 물질 및 변이체들; 숙시네이트, 락테이트 및 아세테이트와 같은 포화 또는 불포화 지방산류; 에탄올, 메탄올 및 글리세롤과 같은 알코올류, 또는 이들의 혼합물들을 포함하나, 이로 제한되지 않는다. 또한, 탄소 공급원은 글루코오스와 같은, 광합성의 산물일 수 있다. 여하한의 바람직한 실시예들에서, 탄소 공급원은 바이오매스이다. 다른 바람직한 실시예들에서, 탄소 공급원은 수크로오스이다.

[0101] "바이오매스"라는 용어는 탄소 공급원으로부터 유래되는 여하한 생물학적 물질을 칭한다. 일부 실시예들에서, 바이오매스는 탄소 공급원으로 처리(process)되고, 이는 생물전환에 적합하다. 다른 실시예들에서, 바이오매스는 탄소 공급원으로서의 추가적인 처리를 요구하지 않는다. 탄소 공급원은 바이오연료로 전환될 수 있다. 예시적인 바이오매스의 공급원은 옥수수, 사탕수수 또는 지팡이풀(switchgrass)과 같은 식물성 물질 또는 식물(vegetation)이다. 또 다른 예시적인 바이오매스의 공급원은 동물성 물질[예를 들어, 우분(cow manure)]과 같은 신진대사 노폐물(metabolic waste products)이다. 추가 예시적인 바이오매스의 공급원은 조류 및 여타 수산 식물을 포함한다. 또한, 바이오매스는 발효 찌꺼기, 목초(ensilage), 짚, 재목, 오수(sewage), 쓰레기, 셀룰로오스성 대도시 폐기물(cellulosic urban waste) 및 남은 음식들(food leftovers)을 포함하나 이로 제한되지 않는 공업, 농업, 임업 및 가정으로부터의 폐기물(waste products)을 포함한다. 또한, "바이오매스"라는 용어는 탄수화물(예를 들어, 단당류, 이당류 또는 다당류)들과 같은, 탄소 공급원들을 칭할 수 있다.

[0102] 폴리펩티드의 발현 또는 생성물의 생산을 가능하게 할 만큼 조건들이 충분한 지를 결정하기 위하여, 재조합 숙주 세포는, 예를 들어 약 4, 8, 12, 24, 36, 48, 72 또는 그 이상의 시간 동안 배양될 수 있다. 배양 동안 및/또는 배양 이후, 샘플이 얻어지고, 상기 조건들이 생산 또는 발현을 가능하게 하는지를 결정하기 위하여 분석될 수 있다. 예를 들어, 샘플 내의 숙주 세포들 또는 숙주 세포들이 성장된 배지는 원하는 생성물의 존재에 대하여 테스트될 수 있다. 지방족 아미드의 존재에 대하여 테스트하는 경우, MS, 박층 크로마토그래피(TLC), 고성능 액체 크로마토그래피(HPLC), 액체 크로마토그래피(LC), 불꽃 이온화 검출기(FID)에 연결된 GC, GC/MS 및 LC/MS와 같은 분석법(assays)이 사용될 수 있으나, 이로 제한되지 않는다. 폴리펩티드의 발현에 대하여 테스트하는 경우에, 웨스턴 블로팅(Western blotting) 및 점 블로팅(dot blotting)과 같은 기술들이 사용될 수 있으나, 이로 제한되지 않는다.

- [0103] 본 발명의 방법들에 있어서, 지방족 아미드의 생산 및 분리는 발효 조건들을 최적화시킴으로써 향상될 수 있다. 일부 실시예들에서, 발효 조건들은 탄화수소 생성물로 전환되는 탄소 공급원의 백분율(percentage)을 증가시켜 최적화된다. 정상적인 세포 생명주기(lifecycles) 동안, 탄소는 지질, 당류, 단백질, 유기산(organic acids) 및 핵산을 생산하는 것과 같은 세포 기능들에 사용된다. 성장에 관련된 활성들을 위해 필요한 탄소의 양을 감소시키는 것은 생성물로의 탄소 공급원의 전환 효율성을 증가시킬 수 있다. 예를 들어, 이는 맨 먼저 숙주 세포들을 원하는 밀도[예를 들어, 성장의 대수 기(log phase)의 피크(peak)에서 달성되는 밀도]로 성장시킴으로써 달성될 수 있다. 이러한 시점에서, 복제 검사점 유전자(replication checkpoint gene)들은 세포들의 성장을 중지시키는데 이용(harnessed)될 수 있다. 특히, [Camilli *et al.* (2006) *Science* 311:1113; Venturi (2006) *FEMS Microbiol. Rev.* 30:274-291; 및 Reading *et al.* (2006) *FEMS Microbiol. Lett.* 254:1-11에 논평된] 퀴럼 센싱 메카니즘(quorum sensing mechanisms)은 *p53*, *p21*과 같은 검사점 유전자(checkpoint gene)들 또는 다른 검사점 유전자들을 활성화시키는데 사용될 수 있다.
- [0104] 대장균 내에서의 성장 및 세포 복제를 중지시키기 위하여 활성화될 수 있는 유전자들은 *umuDC* 유전자들을 포함한다. *umuDC* 유전자들의 과발현은 정상기(stationary phase)에서 지수함수적 증식(exponential growth)으로의 진행을 중지시킨다[Murli *et al.* (2000) *J. Bacteriol.* 182:1127-1135]. UmuC는 보통 자외선(UV) 및 화학적 돌연변이 유발(chemical mutagenesis)로부터 기인하는 비-코딩 손상(non-coding lesions)을 넘어서 트랜스리전 합성(translesion synthesis)을 실행할 수 있는 DNA 중합효소(polymerase)이다. *umuDC* 유전자 생성물들은 트랜스리전 합성(translesion synthesis)의 과정에 관련되고, 또한 DNA 서열 손상 검사점(damage checkpoint)의 역할을 한다. *umuDC* 유전자 생성물들은 UmuC, UmuD, *umuD'*, *UmuD'₂C*, *UmuD'₂* 및 *UmuD₂*를 포함한다. 동시에, 생성물-생산 유전자들은 활성화될 수 있고, 이로 인해서 지방족 아미드 또는 이의 중간체가 만들어지는 동안에 사용되는 복제 및 유지(maintenance) 경로들의 필요성(need)을 최소화시킨다. 또한, 숙주 세포들은 적절한 최종 생성물 생산 유전자들과 함께 이러한 유전자의 *데노보* 합성(*de novo* synthesis)을 통하여 *prpBCDE* 프로모터 시스템 하에서 pBAD24으로 대장균으로부터 *umuC* 및 *umuD*를 발현시키도록 조작될 수 있다.
- [0105] 숙주 세포는 재조합 효소 복합체(recombinant cellulosome)를 발현시키도록 추가적으로 조작될 수 있고, 이는 숙주 세포가 탄소 공급원으로서 셀룰로오스성 물질(cellulosic material)을 사용할 수 있게 한다. 본 발명의 방법들에서 사용하기에 적합한 예시적인 효소 복합체들은, 예를 들어 국제 특허 출원 공개 WO 2008/100251에 기재된 효소 복합체들을 포함한다. 또한, 숙주 세포는 미국 특허 번호 5,000,000; 5,028,539; 5,424,202; 5,482,846 및 5,602,030에 개시된 방법들에 따라서 탄소 공급원으로서 셀룰로오스성 물질을 사용하고, 효율적으로 탄소를 동화(assimilate)시키도록 조작될 수 있다. 또한, 숙주 세포는 수크로오스가 탄소 공급원으로서 사용될 수 있도록 인버타아제(invertase)를 발현시키도록 조작될 수 있다.
- [0106] 본 발명의 발효 방법들에 대한 일부 실시예들에서, 발효실(fermentation chamber)은 연속적인 반응이 진행(undergoing)되고, 이로 인해서 안정한 환원성 환경(reductive environment)을 만드는 발효를 내부에 포함한다(enclose). 전자 균형은 (기체 형태의) 이산화탄소의 방출에 의해 유지될 수 있다. 또한, NAD/H 및 NADP/H 균형을 증대시키기 위한 노력들은 전자 균형을 안정화시키는 것을 용이하게 할 수 있다. 또한, 세포 내의 NADPH의 이용률(availability)은 NADH:NADPH 수소전달효소(transhydrogenase)를 발현시키도록 숙주 세포를 조작함으로써 향상될 수 있다. 하나 이상의 NADH:NADPH 수소전달효소들의 발현은 해당과정(glycolysis)에서 생산된 NADH를 NADPH로 전환시키며, 이는 지방족 아미드들과 이의 중간체들의 생산을 향상시킬 수 있다.
- [0107] 소규모 생산의 경우에, 조작된 숙주 세포들은 예를 들어, 약 100 mL, 500 mL, 1 L, 2 L, 5 L 또는 10 L의 배치(batches)에서 성장될 수 있고; 발효될 수 있으며; 원하는 핵산 서열, 이를 테면 PPS를 인코딩하는 핵산 서열을 발현시키도록 유도될 수 있다. 대규모 생산의 경우에, 조작된 숙주 세포들은 약 10 L, 100 L, 1000 L, 10,000 L, 100,000 L 및 1,000,000 L 또는 그 이상의 배치(batches)에서 성장될 수 있고; 발효될 수 있으며; 원하는 핵산 서열을 발현시키도록 유도될 수 있다.
- [0108] 일반적으로, 본 발명의 방법들에 의해 생산되는 지방족 아미드는 숙주 세포로부터 분리된다. 지방족 아미드와 같은 생성물들 대하여 본 명세서에서 사용되는 바와 같이 "분리된"이라는 용어는 세포 성분(cellular components), 세포 배양 배지, 또는 화학적 또는 합성 전구체로부터 분리되는 생성물들을 칭한다. 본 명세서에 설명된 방법들에 의해 생산되는 지방족 아미드는 세포질뿐만 아니라 발효액(fermentation broth) 내에서 상대적으로 혼합되지 않을 수 있다. 따라서, 지방족 아미드 및 이의 유도체는 세포 내 또는 세포 외의 유기상(organic phase)에 수집될 수 있다. 유기상 내의 생성물들의 수집은 세포 기능에 지방족 아미드의 영향을 줄일 수 있고, 숙주 세포가 더 많은 생성물을 생산할 수 있게 한다.

- [0109] 일부 실시예들에서, 본 발명의 방법들에 의해 생산되는 지방족 아마이드는 정제된다. 본 명세서에서 사용되는 바와 같이, "정제한다", "정제된" 또는 "정제"라는 용어들은, 예를 들어 분리(isolation) 또는 구분(separation)에 의해 주변 환경으로부터 분자의 이동(removal) 또는 분리를 의미한다. "실질적으로 정제된" 분자는 결합되는 다른 성분들로부터 적어도 약 60 % (예를 들어, 적어도 약 70 %, 적어도 약 75 %, 적어도 약 85 %, 적어도 약 90 %, 적어도 약 95 %, 적어도 약 97 %, 적어도 약 99 %) 유리된다(free). 본 명세서에서 사용되는 바와 같이, 또한 이러한 용어들은 샘플로부터 오염물질들의 제거를 칭한다. 예를 들어, 오염물질들의 제거는 샘플 내에서 지방족 아마이드의 백분율(percentage)의 증가를 유발할 수 있다. 예를 들어, 지방족 아마이드가 재조합 숙주 세포에서 생산되는 경우에, 상기 지방족 아마이드는 숙주 세포 단백질의 제거에 의해 정제될 수 있다. 정제 후, 샘플 내 지방족 아마이드의 백분율이 증가된다.
- [0110] 본 명세서에서 사용되는 바와 같이, "정제한다", "정제된" 및 "정제"라는 용어들은 절대 순도(absolute purity)를 요구하지 않는 상대적인 용어들이다. 따라서, 예를 들어 지방족 아마이드가 숙주 세포에서 생산되는 경우에, 정제된 지방족 아마이드는 다른 세포 성분(예를 들어, 핵산, 폴리펩티드, 지질, 탄수화물 또는 기타 탄화수소)로부터 실질적으로 분리되는 지방족 아마이드이다.
- [0111] 추가적으로, 정제된 지방족 아마이드 제조물(preparation)은 발효 후에 존재할지도 모르는 것들과 같은 오염물질들로부터 실질적으로 유리된 지방족 아마이드인 지방족 아마이드 제조물이다. 일부 실시예들에서, 샘플의 중량 당 적어도 약 50 %가 지방족 아마이드로 구성되어 있는 경우에, 지방족 아마이드는 정제된다. 다른 실시예들에서, 샘플의 중량 당 적어도 약 60 %, 예를 들어 적어도 약 70 %, 적어도 약 80 %, 적어도 약 85 %, 적어도 약 90 %, 적어도 약 92 %가 지방족 아마이드로 구성되어 있는 경우에, 지방족 아마이드는 정제된다. 대안적으로 또는 추가적으로, 샘플의 중량 당 약 100 % 미만, 예를 들어 약 99 % 미만, 약 98 % 미만, 약 95 % 미만, 약 90 % 미만 또는 약 80 % 미만이 지방족 아마이드로 구성되어 있는 경우에, 지방족 아마이드는 정제된다. 따라서, 정제된 지방족 아마이드는 상기의 중단점(endpoint)들 중 여하한 두 개로 제한되는 순도 수준(purity level)을 가질 수 있다. 예를 들어, 샘플의 적어도 약 80 % 내지 95 %, 적어도 약 85 % 내지 99 %, 또는 적어도 약 90 % 내지 98 %가 지방족 아마이드로 구성되어 있는 경우에, 지방족 아마이드는 정제될 수 있다.
- [0112] 지방족 아마이드는 세포 외 환경에 존재할 수 있거나 또는 숙주 세포의 세포 외 환경으로부터 분리될 수 있다. 여하한 실시예들에서, 지방족 아마이드는 숙주 세포에서 분리될 수 있다. 다른 실시예들에서, 지방족 아마이드는 세포 외 환경으로 이송된다. 또 다른 실시예들에서, 지방족 아마이드는 수동적으로 세포 외 환경으로 이송된다. 지방족 아마이드는 국제 특허 출원 공개 WO 2010/042664 및 WO 2010/062480에 기재된 것들과 같은 본 기술분야에 알려진 방법들을 사용하여 숙주 세포로부터 분리될 수 있다.
- [0113] 본 명세서에서 설명되는 방법들은 동종(homogeneous) 화합물들의 생산을 야기할 수 있고, 여기서 생산된 지방족 아마이드의 적어도 약 60 %, 적어도 약 70 %, 적어도 약 80 %, 적어도 약 90 % 또는 적어도 약 95 %는 6 개 미만의 탄소들, 5 개 미만의 탄소들, 4 개 미만의 탄소들, 3 개 미만의 탄소들 또는 약 2 개 미만의 탄소들로 다양한 지방족 사슬들을 가질 것이다. 대안적으로 또는 추가적으로, 본 명세서에서 설명되는 방법들은 동종 화합물들의 생산을 야기할 수 있고, 여기서 생산된 지방족 아마이드의 약 98 % 미만, 약 95 % 미만, 약 90 % 미만, 약 80 % 미만 또는 약 70 % 미만은 6 개 미만의 탄소들, 5 개 미만의 탄소들, 4 개 미만의 탄소들, 3 개 미만의 탄소들 또는 약 2 개 미만의 탄소들로 다양한 지방족 사슬들을 가질 것이다. 따라서, 지방족 아마이드는 상기의 중단점(endpoint)들 중 여하한 두 개로 제한되는 동질도(degree of homogeneity)를 가질 수 있다. 예를 들어, 지방족 아마이드는 동질도를 가질 수 있으며, 여기서 생산된 지방족 아마이드의 약 70 % 내지 95 %, 약 80 % 내지 98 % 또는 약 90 % 내지 95 %는 6 개 미만의 탄소들, 5 개 미만의 탄소들, 4 개 미만의 탄소들, 3 개 미만의 탄소들 또는 약 2 개 미만의 탄소들로 다양한 지방족 사슬들을 가질 것이다. 또한, 이러한 화합물들은 상대적으로 일정한 포화도로 생산될 수 있다.
- [0114] 본 발명의 방법들의 결과로서, 일차 아민 및 아실 티오에스테르의 지방족 아마이드로의 전환을 촉진시키는 폴리펩티드를 인코딩하는 핵산 서열을 포함하도록 조작된 재조합 미생물에 의해 생산되는 지방족 아마이드의 역가(titer), 수율(yield) 또는 생산률(productivity) 중 하나 이상은, 대응하는 야생형 미생물의 역가, 수율 또는 생산률에 대하여 증가된다.
- [0115] "역가"라는 용어는 숙주 세포 배양의 단위 부피 당 생산된 지방족 아마이드의 질량(quantity)을 칭한다. 본 명세서에 설명된 조성물들 및 방법들의 여하한 측면에서, 지방족 아마이드는 25 mg/L 이상, 50 mg/L 이상, 75 mg/L 이상, 100 mg/L 이상, 125 mg/L 이상, 150 mg/L 이상, 175 mg/L 이상, 200 mg/L 이상, 250 mg/L 이상, 300 mg/L 이상, 350 mg/L 이상, 400 mg/L 이상, 450 mg/L 이상, 500 mg/L 이상, 600 mg/L 이상, 700 mg/L 이상, 800

mg/L 이상, 900 mg/L 이상 또는 1000 mg/L 이상의 역가로 생산된다. 대안적으로 또는 추가적으로, 지방족 아미드는 2000 mg/L 이하, 1900 mg/L 이하, 1800 mg/L 이하, 1700 mg/L 이하, 1600 mg/L 이하, 1500 mg/L 이하, 1400 mg/L 이하, 1300 mg/L 이하, 1200 mg/L 이하, 1100 mg/L 이하, 1000 mg/L 이하, 900 mg/L 이하, 800 mg/L 이하, 700 mg/L 이하, 600 mg/L 이하, 500 mg/L 이하, 400 mg/L 이하, 300 mg/L 이하 또는 200 mg/L 이하의 역가로 생산된다. 따라서, 지방족 아미드는 상기의 중단점(endpoint)들 중 여하한 두 개로 제한되는 역가로 생산된다. 예를 들어, 지방족 아미드는 150 내지 1000 mg/L, 200 내지 500 mg/L, 500 내지 1500 mg/L 또는 300 내지 1300 mg/L의 역가로 생산될 수 있다. 다른 실시예들에서, 지방족 아미드는 2000 mg/L 초과, 5000 mg/L 초과, 10,000 mg/L 초과 또는 그 이상, 이를 테면 50 g/L, 70 g/L, 100 g/L, 120 g/L, 150 g/L 또는 200 g/L의 역가로 생산된다.

[0116] "수율"이라는 용어는 숙주 세포 내에서 투입된(input) 탄소 공급원이 생성물(즉, 지방족 아미드)로 전환되는 효율을 칭한다. 산소-함유 탄소 공급원들(예를 들어, 글루코오스 및 다른 탄수화물 기반 공급원들)의 경우, 산소는 이산화탄소의 형태로 방출되어야 한다. 따라서, 두 개의 산소 원자들이 방출될 때마다, 탄소 원자도 방출되어, (지방산 유래 생성물들에 대하여) 대략 34 % (w/w)의 최대 이론적 물질대사의 효율성에 이른다. 그러나, 이러한 수치(figure)는 다른 유기 화합물들 및 탄소 공급원들의 경우에 변한다. 본 명세서에 보고된 전형적인 수율은 대략 5 % 미만이다. 본 명세서의 방법들에 따라서 지방족 아미드를 생산하도록 조작된 숙주 세포는 약 3 % 이상, 약 5 % 이상, 약 10 % 이상, 약 15 % 이상, 약 18 % 이상 또는 약 20 % 이상의 수율을 가질 수 있다. 대안적으로 또는 추가적으로, 수율은 약 30 % 이하, 약 27 % 이하, 약 25 % 이하, 약 22 % 이하, 약 20 % 이하, 약 17 % 이하, 약 13 % 이하 또는 약 10 % 이하이다. 따라서, 이러한 수율은 상기의 중단점(endpoint)들 중 여하한 두 개로 제한될 수 있다. 예를 들어, 본 발명의 제조합 미생물에 의해 생산되는 지방족 아미드의 수율은 약 5 % 내지 약 25 %, 약 10 % 내지 약 25 %, 약 10 % 내지 약 22 %, 약 15 % 내지 약 27 % 또는 약 18 % 내지 약 22 %일 수 있다. 다른 실시예들에서, 수율은 30 % 초과이다.

[0117] "생산률"이라는 용어는 숙주 세포 배양의 단위 밀도 당 숙주 세포 배양의 단위 부피 당 생산된 지방족 아미드의 질량을 칭한다. 본 발명에 설명된 조성물들 및 방법들의 여하한 측면에서, 제조합 미생물에 의해 생산된 지방족 아미드의 생산률은 약 3 mg/L/OD₆₀₀ 이상, 약 6 mg/L/OD₆₀₀ 이상, 약 9 mg/L/OD₆₀₀ 이상, 약 12 mg/L/OD₆₀₀ 이상, 약 15 mg/L/OD₆₀₀ 이상, 약 18 mg/L/OD₆₀₀ 이상 또는 약 20 mg/L/OD₆₀₀ 이상이다. 대안적으로 또는 추가적으로, 생산률은 약 50 mg/L/OD₆₀₀ 이하, 약 40 mg/L/OD₆₀₀ 이하, 약 30 mg/L/OD₆₀₀ 이하, 약 25 mg/L/OD₆₀₀ 이하, 약 20 mg/L/OD₆₀₀ 이하, 약 17 mg/L/OD₆₀₀ 이하 또는 약 10 mg/L/OD₆₀₀ 이하이다. 따라서, 생산률은 상기의 중단점(endpoint)들 중 여하한 두 개로 제한될 수 있다. 예를 들어, 생산률은 약 3 내지 약 30 mg/L/OD₆₀₀, 약 6 내지 약 20 mg/L/OD₆₀₀ 또는 약 15 내지 약 30 mg/L/OD₆₀₀일 수 있다.

[0118] 또한, 본 발명은 본 명세서에서 설명된 방법들 및 제조합 미생물들에 의해 생산되는 지방족 아미드를 제공한다. 본 발명의 방법들 및 제조합 미생물들에 의해 생산되는 바이오생산물(bioprodut)[예를 들어, 지방족 아미드]은 이중 탄소-동위원소 지문기술(dual carbon-isotopic fingerprinting) 또는 ¹⁴C 연대측정(dating)을 기초로 석유 화학의 탄소로부터 유래되는 유기 화합물들과 구별될 수 있다. 추가적으로, 생물자원 탄소(biosourced carbon)의 특이적인 공급원(예를 들어, 글루코오스 대 글리세롤)은 이중 탄소-동위원소 지문기술에 의해 결정될 수 있다(예를 들어, 미국 특허 7,169,588 참조).

[0119] 석유 기반 유기 화합물들로부터 바이오생산물을 구별하는 능력은 거래 중에(in commerce) 이러한 물질들을 추적하는데 유익하다. 예를 들어, 생물학적 기반의 탄소 동위원소 프로파일들 및 석유 기반의 탄소 동위원소 프로파일들 모두를 포함하는 유기 화합물들 또는 화학제품들은 석유 기반 물질들로부터 만들어진 유기 화합물들 및 화학제품들과 구별될 수 있다. 따라서, 본 발명의 방법들에 따라서 제조되는 지방족 아미드는 이의 독특한 탄소 동위원소 프로파일을 기반으로 하여 거래 중에 추적(followed)될 수 있다.

[0120] 바이오생산물들은 각각의 연료에서 안정한 탄소 동위원소 비율(¹³C/¹²C)을 비교함으로써 석유 기반 유기 화합물들과 구별될 수 있다. 주어진 바이오생산물의 ¹³C/¹²C 비율은 이산화탄소가 고정된 시간에서 대기 중의 이산화탄소에 대한 ¹³C/¹²C 비율의 결과이다. 또한, 이는 정확한 물질대사의 경로를 반영한다. 국부적인 변이들도 발생한다. 석유, C₃ 식물들[활엽(the broadleaf)], C₄ 식물들[목초(the grasses)] 및 해양의 탄산염(marine carbonate)들 모두는 ¹³C/¹²C 및 대응하는 δ¹³C 수치들에서 상당한 차이들을 나타낸다. 더욱이, C₃ 및 C₄ 식물들

의 지질 물질은 물질대사의 경로의 결과로서 동일한 식물들의 탄수화물 성분들로부터 유래된 재료들과 상이하게 분석한다.

[0121] ^{13}C 측정 척도(measurement scale)는 본래 Pee Dee Belemnite(PDB) 석회암에 의해 제로 세트(zero set)로 정의되며, 여기서 수치들은 이러한 재료로부터 천 분의 일 편차로 주어진다. " $\delta^{13}\text{C}$ " 수치는 천 분의 일[퍼밀(per mil)], 약어로는 ‰로 표현되고, 하기와 같이 계산된다:

$$[0122] \delta^{13}\text{C}(\text{‰}) = \left[\left(\frac{^{13}\text{C}/^{12}\text{C}}{\text{샘플}} - \frac{^{13}\text{C}/^{12}\text{C}}{\text{표준물질}} \right) / \frac{^{13}\text{C}/^{12}\text{C}}{\text{표준물질}} \times 1000 \right]$$

[0123] 일부 실시예들에서, 본 발명의 방법들에 따라서 생성된 지방족 아미드는 약 -30 이상, 약 -28 이상, 약 -27 이상, 약 -20 이상, 약 -18 이상, 약 -15 이상, 약 -13 이상 또는 약 -10 이상의 $\delta^{13}\text{C}$ 를 갖는다. 대안적으로 또는 추가적으로, 지방족 아미드는 약 -4 이하, 약 -5 이하, 약 -8 이하, 약 -10 이하, 약 -13 이하, 약 -15 이하, 약 -18 이하 또는 약 -20 이하의 $\delta^{13}\text{C}$ 를 갖는다. 따라서, 지방족 아미드는 상기의 종단점(endpoint)들 중 여하한 두 개로 제한되는 $\delta^{13}\text{C}$ 를 가질 수 있다. 예를 들어, 지방족 아미드는 약 -30 내지 약 -15, 약 -27 내지 약 -19, 약 -25 내지 약 -21, 약 -15 내지 약 -5, 약 -13 내지 약 -7 또는 약 -13 내지 약 -10의 $\delta^{13}\text{C}$ 를 가질 수 있다. 일부 실시예들에서, 지방족 아미드는 약 -10, -11, -12 또는 -12.3의 $\delta^{13}\text{C}$ 를 가질 수 있다. 다른 실시예들에서, 지방족 아미드는 약 -15.4 이상의 $\delta^{13}\text{C}$ 를 갖는다. 또 다른 실시예들에서, 지방족 아미드는 약 -15.4 내지 약 -10.9의 $\delta^{13}\text{C}$, 또는 약 -13.92 내지 약 -13.84의 $\delta^{13}\text{C}$ 를 갖는다.

[0124] 또한, 바이오생성물들은 각각의 화합물에서 ^{14}C 의 양을 비교함으로써 석유 기반 유기 화합물들과 구별될 수 있다. ^{14}C 는 핵의 반감기가 5730년이므로, "더 오래된(older)" 탄소를 함유하는 석유 기반 연료는 "최근의(newer)" 탄소를 함유하는 바이오생성물들과 구별될 수 있다[예를 들어, Currie, "Source Apportionment of Atmospheric Particles", *Characterization of Environmental Particles*, J. Buffle and H. P. van Leeuwen, Eds., Vol. I of the IUPAC Environmental Analytical Chemistry Series, Lewis Publishers, Inc., pp. 3-74(1992) 참조].

[0125] ^{14}C 는 "현대 탄소의 분율"의 단위(f_M)로 주어지는 결과들을 갖는, 가속기 질량 분광 분석(accelerator mass spectrometry: AMS)에 의해 측정될 수 있다. f_M 은 미국표준기술연구소(NIST)의 표준 기준 물질들(SRMs) 4990B 및 4990C에 의해 정의된다. 본 명세서에서 사용되는 바와 같이, "현대 탄소의 분율" 또는 f_M 은 각각 옥살산 기준 HOxI 및 HOxII로 알려진, 미국표준기술연구소(NIST)의 표준 기준 물질들(SRMs) 4990B 및 4990C에 의해 정의되는 바와 같은 동일한 의미를 갖는다. 기초적인 정의는 (AD 1950을 기준으로) 0.95 배의 $^{14}\text{C}/^{12}\text{C}$ 동위원소 비율 HOxI에 관한 것이다. 이는 감소-수정된(decay-corrected) 산업 혁명 이전의 목재와 거의 동등하다. 현재의 생물권(식물 재료)에 대한, f_M 은 대략 1.1이다.

[0126] 일부 실시예들에서, 본 발명의 방법들에 따라서 생산되는 지방족 아미드는 적어도 약 1, 예를 들어 적어도 약 1.003, 적어도 약 1.01, 적어도 약 1.04, 적어도 약 1.111, 적어도 약 1.18 또는 적어도 약 1.124의 $f_M^{14}\text{C}$ 를 갖는다. 대안적으로 또는 추가적으로, 지방족 아미드는 약 1.130 이하, 예를 들어 약 1.124 이하, 약 1.18 이하, 약 1.111 이하 또는 약 1.04 이하의 $f_M^{14}\text{C}$ 를 갖는다. 따라서, 지방족 아미드는 상기의 종단점(endpoint)들 중 여하한 두 개로 제한되는 $f_M^{14}\text{C}$ 를 가질 수 있다. 예를 들어, 지방족 아미드는 약 1.003 내지 약 1.124의 $f_M^{14}\text{C}$, 약 1.04 내지 약 1.18의 $f_M^{14}\text{C}$ 또는 약 1.111 내지 약 1.124의 $f_M^{14}\text{C}$ 를 가질 수 있다.

[0127] ^{14}C 의 또 다른 측정은 현대 탄소의 퍼센트, 즉 pMC로 알려져 있다. ^{14}C 연대를 사용하는 고고학자 또는 지질학자에 경우, AD 1950년은 "측정의 기점 연도(zero years old)"와 같다. 또한, 이는 100 pMC로 나타낸다. 대기 중의 "핵무기 탄소(Bomb carbon)"는 열-핵무기 시험의 정점인 1963년에 정상 수준에 거의 두 배에 달하였다. 대기권 내에서 이의 분포는, 핵무기 탄소의 등장 이후에, AD 1950년 이후로 살아 있는 식물들 및 동물들에 대하여 100 pMC보다 더 큰 수치들을 나타내는 것으로 추정되고 있다. 107.5 pMC에 가까운 오늘날의 수치는 시간이

지나면서 점차 감소하고 있다. 이는 옥수수수와 같은 신선한 바이오매스 재료는 107.5 pMC에 가까운 ¹⁴C 특징(signature)을 부여하는 것을 의미한다. 석유 기반 화합물들은 0의 pMC 수치를 가질 것이다. 현대 탄소와 화석 탄소의 결합은 현대 pMC 함유량(content)의 희석을 야기할 것이다. 107.5 pMC가 현대 바이오매스 재료들의 ¹⁴C 함유량을 나타내며 0 pMC가 석유 기반 생성물들의 ¹⁴C 함유량을 나타낸다고 가정함으로써, 재료에 대한 측정된 pMC 수치는 두 개의 성분 타입에 대한 비율을 반영할 것이다. 예를 들어, 오늘날의 콩으로부터 100 % 유래된 재료는 107.5 pMC에 가까운 방사성 탄소 특징을 가질 것이다. 만약 재료가 석유 기반 생성물로 50 % 희석되었다면, 결과적으로 얻어진 혼합물은 대략 54 pMC의 방사성 탄소 특징을 가질 것이다.

[0128] 생물학적 기반의 탄소 함유량은 107.5 pMC와 같으면 "100 %"이고, 0 pMC와 같으면 "0 %"라고 지정함으로써 유래된다. 예를 들어, 99 pMC로 측정된 샘플은 93 %의 동등한 생물학적 기반의 탄소 함유량을 제공할 것이다. 이 수치는 생물학적 기반의 탄소 결과를 의미하며, 현대의 생물학적 재료 또는 석유 기반 재료로부터 유래되는 분석된 재료 내에서 모든 성분들을 추측하는 것으로 해석된다.

[0129] 일부 실시예들에서, 본 발명의 방법들에 따라서 생산된 지방족 아미드는 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 75, 적어도 약 80, 적어도 약 85, 적어도 약 90, 적어도 약 95, 적어도 약 96, 적어도 약 97 또는 적어도 약 98의 pMC를 갖는다. 대안적으로 또는 추가적으로, 지방족 아미드는 약 108 이하, 약 105 이하, 약 102 이하, 약 99 이하, 약 96 이하, 약 93 이하, 약 90 이하, 약 85 이하 또는 약 80 이하의 pMC를 갖는다. 따라서, 지방족 아미드는 상기의 종단점(endpoint)들 중 여하한 두 개로 제한되는 pMC를 가질 수 있다. 예를 들어, 지방족 아미드는 약 50 내지 약 100; 약 60 내지 약 105; 약 70 내지 약 100; 약 80 내지 약 105; 약 85 내지 약 100; 약 87 내지 약 98; 또는 약 90 내지 약 95의 pMC를 가질 수 있다. 다른 실시예들에서, 본 명세서에서 설명된 지방족 아미드는 약 90, 약 91, 약 92, 약 93, 약 94 또는 약 94.2의 pMC를 갖는다.

[0130] 본 명세서에서 설명되는 여하한 방법들 및 제조합 미생물들에 의해 생산된 지방족 아미드는 그 자체가(per se) 계면활성제 또는 세제로서 직접적으로 사용될 수 있거나, 또는 지방족 아미드는 개인용, 애완동물용 또는 가정용 세정 조성물로 제제화(formulated)될 수 있다. 계면활성제, 세제 및 세정 조성물들 및 이들의 생산 방법들은 본 기술분야의 당업자들에게 잘 알려져 있고, 더 자세한 세부사항은 예를 들어 미국 특허 출원 공개 2011/0206630에 기재되며, 이는 본 명세서에서 전문이 인용참조된다.

[0131] 따라서, 본 발명은 본 명세서에 설명되는 여하한 방법들에 의해 생산되는 지방족 아미드를 포함하는 계면활성제, 세제 및 세정 조성물들을 제공한다. 본 기술분야의 당업자는 계면활성제, 세제 또는 세정 조성물의 의도하는 목적에 따라서, 상이한 지방족 아미드들이 생산되고 사용될 수 있다는 점을 인식할 것이다. 예를 들어, 본 명세서에서 설명되는 지방족 아미드가 계면활성제 또는 세제 생산에 공급원료로서 사용되는 경우에, 본 기술분야의 당업자는 지방족 아미드 공급원료의 특성들이 생산되는 계면활성제 또는 세제 조성물의 특성들에 영향을 미칠 것이라는 점을 인식할 것이다. 따라서, 계면활성제 또는 세제 조성물의 특성들은 공급원료로서 사용하기 위한 특정한 지방족 아미드를 생산함으로써 선택될 수 있다.

[0132] 본 발명의 지방족 아미드 기반의 계면활성제 또는 세제는 본 기술분야에 잘 알려진 다른 계면활성제들 및/또는 세제들과 혼합될 수 있다. 혼합물 내에 존재할 수 있는 지방족 아미드는 혼합물의 총 중량에 기초하여, 10 중량 퍼센트(weight percent: wt.%) 이상, 15 wt.% 이상, 20 wt.% 이상, 30 wt.% 이상, 40 wt.% 이상, 50 wt.% 이상, 60 wt.% 이상 또는 70 wt.% 이상의 양으로 혼합물 내에 존재할 수 있다. 대안적으로 또는 추가적으로, 지방족 아미드는 혼합물의 총 중량에 기초하여, 95 wt.% 이하, 90 wt.% 이하, 80 wt.% 이하, 70 wt.% 이하, 60 wt.% 이하, 50 wt.% 이하 또는 40 wt.% 이하의 양으로 혼합물 내에 존재할 수 있다. 따라서, 지방족 아미드는 상기의 종단점(endpoint)들 중 여하한 두 개로 제한되는 양으로 혼합물 내에 존재할 수 있다. 예를 들어, 지방족 아미드는 15 내지 40 %, 30 내지 90 %, 50 내지 95 % 또는 40 내지 50 %의 양으로 혼합물 내에 존재할 수 있다.

[0133] 지방족 아미드 기반의 계면활성제는 세정 조성물에 세정력과 세척력을 부여하기 위하여 세정 조성물로 제제화될 수 있다. 지방족 아미드는 세정 조성물의 총 중량에 기초하여, 0.001 wt.% 이상, 0.1 wt.% 이상, 1 wt.% 이상, 10 wt.% 이상, 20 wt.% 이상 또는 40 wt.% 이상의 양으로 세정 조성물 내에 존재할 수 있다. 대안적으로 또는 추가적으로, 지방족 아미드는 세정 조성물의 총 중량에 기초하여, 60 wt.% 이하, 50 wt.% 이하, 40 wt.% 이하, 30 wt.% 이하, 15 wt.% 이하 또는 5 wt.% 이하의 양으로 세정 조성물 내에 존재할 수 있다. 따라서, 지방족 아미드는 상기의 종단점(endpoint)들 중 여하한 두 개로 제한되는 양으로 세정 조성물 내에 존재할 수 있다. 예를 들어, 지방족 아미드는 0.1 내지 10 wt.%, 10 내지 15 wt.%, 20 내지 40 wt.% 또는 0.001 내지 5 wt.%의 양

으로 세정 조성물로 제제화될 수 있다.

- [0134] 본 발명의 세정 조성물은 정제(tablet), 과립, 분말 또는 콤팩트(compact)와 같은 고체 형태일 수 있다. 또한, 세정 조성물은 유체(fluid), 겔(gel), 페이스트(paste), 에멀전(emulsion) 또는 농축물(concentrate)과 같은 액체 형태일 수 있다.
- [0135] 이하한 실시예들에서, 본 발명의 세정 조성물은 액체 또는 고체 세탁용 세제 조성물이다. 일부 실시예들에서, 세정 조성물은 경질 표면(hard surface) 세정 조성물이고, 여기서 경질 표면 세정 조성물은 바람직하게 부직포 기재(nonwoven substrate)를 함침(impregnate)시킨다. 본 명세서에서 사용되는 바와 같이, "함침시키다"는 경질 표면 세정 조성물이 부직포 기재와 접촉하게 위치되어, 부직포 기재의 적어도 일부분이 경질 표면 세정 조성물에 의해 침투(penetrated)되도록 한다는 것을 의미한다. 예를 들어, 경질 표면 세정 조성물은 바람직하게 부직포 기재를 포화시킨다. 다른 실시예들에서, 본 발명의 세정 조성물은 차량 관리용(car care) 조성물이고, 이는 견목재(hard wood), 타일(tile), 세라믹, 플라스틱, 가죽, 금속 및/또는 유리와 같은 다양한 표면들을 세정 하는데 유용하다. 일부 실시예들에서, 세정 조성물은 예를 들어, 액체의 식기세척용(liquid hand dishwashing) 조성물, 고체의 자동 식기세척기용(automatic dishwashing) 조성물, 액체의 자동 식기세척기용 조성물 및 탭/단위 투여 형태(tab/unit dose form)의 자동 식기세척기용 조성물과 같은 식기세척용 조성물이다.
- [0136] 다른 실시예들에서, 세정 조성물은 다양한 장비 및 기계를 세정하고 석유 굴착 작업(oil drilling operations)에 사용하기 위하여 산업 환경에서 사용될 수 있다. 예를 들어, 본 발명의 세정 조성물은 석유 굴착을 돕는데 사용되는 경우와 같이, 경도 허용 계면활성제 시스템(hardness tolerant surfactant systems)을 필요로 하는 조성물들에 그리고 경도와 관계없는(free hardness) 접촉 상태가 되는 환경들에 특히 적합할 수 있다.
- [0137] 일부 실시예들에서, 본 발명의 이하한 제조합 미생물들 및 방법들에 의해 생산되는 지방족 아미드는 샴푸, 바디 워시(body wash), 세안제, 또는 액체 또는 고체 비누와 같은 개인 또는 애완동물 관리용(personal or pet care) 조성물로 제제화된다.
- [0138] 본 발명의 이하한 제조합 미생물들 및 방법들에 의해 생산되는 지방족 아미드를 함유하는 세정 조성물은 본 기술분야의 당업자들에게 잘 알려진 다른 세정 부가물(adjunct)들을 포함할 수 있다. 가정용 세정 조성물, 개인 관리용 조성물 및 이와 유사한 것들을 포함하는 대부분의 세정 조성물들에 적용가능한 일반적인 세정 부가물들은, 용제, 가용화제(solubilizing agents), 담체(carriers), 빌더(builders), 효소, 중합체, 거품 촉진제(suds boosters), 거품 억제제[소포제(antifoam)], 염료, 충전제(fillers), 살균제(germicide), 향수성 물질(hydrotropes), 산화 방지제, 향수, 향수-대응물(pro-perfumes), 효소 안정화제(enzyme stabilizing agents), 색소 및 이와 유사한 것들을 포함한다. 일부 실시예들에서, 세정 조성물은 액체 세정 조성물이고, 여기서 상기 조성물은 용제, 킬레이트제, 분산제(dispersants) 및 물로부터 선택되는 하나 이상의 것들을 포함한다. 다른 실시예들에서, 세정 조성물은 고체이고, 여기서 상기 조성물은 예를 들어, 무기 충전제 염(inorganic filler salt)을 더 포함한다. 무기 충전제 염은, 예를 들어 약 10 wt.%에서 약 35 wt.%까지 다양하게 상당한 양으로 존재하는, 고체 세정 조성물들의 통상적인 성분이다. 예를 들어, 적합한 충전제 염은 염화물들 및 황산염들의 알칼리 및 알칼리 토금속 염들을 포함한다. 예시적인 충전제 염은 황산나트륨이다.
- [0139] 가정용 세정 조성물들(예를 들어, 세탁용 세제 및 가정용 표면 세정제)은 표백제, 표백 활성화제, 촉매 물질, 분산제 중합체(dispersant polymers), 은 관리제(silvercare), 변색 방지제 및/또는 부식 방지제(anti-tarnish and/or anti-corrosion agents), 알칼리 공급원(alkalinity source), 가공 보조제(processing aids), 이염 억제제(dye transfer inhibiting agents), 광택제, 구조 탄성화제(structure elasticizing agents), 섬유 유연제, 마모 방지제(anti-abrasion agents) 및 다른 섬유 관리제(fabric care agents)로부터 선택되는 하나 이상의 추가 성분들을 포함할 수 있다. 세정 부가물은 특히 가정용 세정 조성물에 유용하고, 사용 수준은 예를 들어, 미국 특허 번호 5,576,282, 6,306,812 및 6,326,348에 기재된다. 적합한 세탁용 또는 다른 가정용 세정 부가물의 목록은, 예를 들어 국제 특허 출원 공개 WO 99/05245에 기재되어 있다.
- [0140] 개인 관리용, 애완동물 관리용 또는 화장품용 조성물들[예를 들어, 샴푸, 세안제, 손 세정제, 블러쉬(blushes), 브론저(bronzers) 및 이와 유사한 것들]은 컨디셔닝제(conditioning agents)[예를 들어, 비타민, 실리콘, 실리콘 에멀전 안정화 성분], 양이온 셀룰로오스, 또는 중합체[예를 들어, 구아 중합체(guar polymers)], 비듬 방지제, 향균제, 겔 형성제(gel-forming agents), 현탁제(suspending agents), 점도 조절제(viscosity modifiers), 염료, 비휘발성 용제 또는 희석제(수용성 또는 불수용성), 거품 촉진제, 이살충제(pediculicides), pH 조절제(pH adjusting agents), 향수, 보존제, 킬레이트제(chelators), 단백질, 피부 활성제(skin active agents), 자외선 차단제, UV 흡수제, 미네랄, 허브/과일/식품 추출물, 스펅고지질 유도체(sphingolipid derivatives) 및 점

토(clay)로부터 선택되는 하나 이상의 추가 성분들을 포함할 수 있다.

- [0141] 또한, 본 발명은 본 명세서에서 설명되는 여하한 제조합 미생물들 및 방법들에 의해 생산되는 지방족 아미드를 포함하는 연료 첨가제(fuel additive)를 제공한다. 여하한 실시예들에서, 연료 첨가제는 엔진 성능 첨가제(engine performance additive), 세제, 분산제, 마모 방지제(anti-wear agent), 점도 지수 조절제(viscosity index modifier), 마찰 조절제(friction modifier), 산화 방지제, 부식 억제제, 소포제, 이음새 마감제(seal fix), 윤활 첨가제(lubricity additive), 유동점 강하제(pour point depressant), 혼탁점 감소제(cloud point reducer), 매연 억제제(smoke suppressant), 항력 감소 첨가제(drag reducing additive), 금속 불활성화제(metal deactivator), 살생물제(biocide) 및 향 유화제(demulsifier)로부터 선택된다. 연료 첨가제는 미국 특허 출원 공개 2010/0257777에 더욱 상세하게 기재되고, 이는 본 명세서에서 인용참조된다.
- [0142] 여하한 실시예들에서, 여하한 제조합 미생물들 및 방법들에 의해 생산되는 지방족 아미드를 포함하는 연료 첨가제는 지방족 아미드를 위한 용제로서 사용되는 하나 이상의 베이스 오일(base oils) 및 지방족 아미드를 포함하는 패키지(package)에 뒤섞인다. 등급 및/또는 타입에 따라서, 베이스 오일은 예를 들어, 극한 온도 이점(extreme temperature benefits), 항산화 이점 또는 적합한 유동점을 포함하는 첨가제 패키지에 다양한 수준의 성능 이점을 제공할 수 있다.
- [0143] 또한, 본 발명은 본 명세서에서 설명되는 여하한 제조합 미생물들 및 방법들에 의해 생산되는 지방족 아미드, 및 약제학적으로 허용가능한 담체를 포함하는 약제학적 조성물을 제공한다. 약제학적 조성물은, 예를 들어 특정한 징후(particular indication)[예를 들어, 통증 또는 염증]의 치료에 유용한 치료제와 같은 추가 치료제들 뿐만 아니라, 예를 들어 희석제, 보조제(adjuvants), 부형제(excipients), 방부제, pH 조절제 및 이와 유사한 것들과 같은 추가 성분들을 함유할 수 있다.
- [0144] 약제학적 조성물은 고형[예를 들어, 정제, 캡슐, 설하정(sublingual tablet), 분말, 포(sachet)] 조성물일 수 있다. 또한, 약제학적 조성물은 액체[예를 들어, 수성 액체(aqueous liquid), 젤, 로션, 크림] 조성물일 수 있다. 약제학적 조성물은, 예를 들어 경구, 국부(topical), 정맥 내(intravenous), 근육 내(intramuscular), 복강 내(intraperitoneal), 척추 강내(intrathecal), 경막외(epidural), 경피(percutaneous), 피하(subcutaneous), 점막(transmucosal) 및 비강 내(intranasal) 경로들로 이루어진 그룹으로부터 선택되는 투여 경로와 같은, 여하한 적합한 경로로 투여하기 위하여 제제화될 수 있다.
- [0145] 또한, 본 발명은 본 명세서에서 설명되는 여하한 제조합 미생물들 및 방법들에 의해 생산되는 지방족 아미드의 유효량(effective amount)을 피시험체(subject)에 투여하고, 이에 의해서 상기 피시험체의 질병 또는 질환(condition)을 치료하거나 예방하는 단계를 포함하는, 상기 지방족 아미드의 유효량이 필요한 피시험체의 질병 또는 질환을 치료하거나 예방하는 방법을 제공한다.
- [0146] "유효량" 또는 "치료적으로 유효량(therapeutically effective amount)"이란, 인간 또는 동물 피시험체의 질병 또는 질환의 하나 이상의 증상들을 (숙련된 의료 전문가에 의해 판단되는 바와 같이, 어느 정도까지) 완화시키는 양을 의미한다. 추가적으로, "유효량" 또는 "치료적으로 유효량"이란, 질병 또는 질환의 원인이 되거나 원인과 관련된 생리학적 또는 생화학적 파라미터들을 부분적으로 또는 완전히 정상적으로 회복시키는 양을 의미한다. 본 기술분야의 숙련된 임상의는 조성물을 투여했을 경우에 특정한 질병 상태 또는 장애를 치료하거나 또는 예방하기 위한, 조성물의 치료적으로 유효량을 결정할 수 있다. 치료적으로 유효하도록 요구되는 조성물의 정확한 양은, 예를 들어 많은 환자의 특정 고려사항에 더하여, 활성 물질의 특정한 활성, 이용되는 전달 장치(delivery device), 상기 물질의 물리학적 특성, 투여 목적과 같은 수많은 요인들에 달려 있을 것이다. 유효량 또는 치료적으로 유효량으로 투여되어야 할 조성물의 양의 결정은, 보통의 숙련된 임상의의 기술 및 본 기술분야에서 정례적인 것이다.
- [0147] 일부 실시예들에서, 유효량은 투여 단위(dosage unit) 당 1 ng 이상, 예를 들어 10 ng 이상, 100 ng 이상, 1 μg 이상, 10 μg 이상, 100 μg 이상, 1 mg 이상, 10 mg 이상, 50 mg 이상 또는 100 mg 이상의 본 발명의 지방족 아미드일 수 있다. 대안적으로 또는 추가적으로, 유효량은 투여 단위 당 5 g 이하, 1 g 이하, 500 mg 이하, 250 mg 이하, 100 mg 이하, 75 mg 이하, 25 mg 이하, 10 mg 이하 또는 1 mg 이하의 본 발명의 지방족 아미드일 수 있다. 따라서, 지방족 아미드는 상기의 중단점(endpoint)들 중 여하한 두 개로 제한되는 양으로 투여 단위에 존재할 수 있다. 예를 들어, 지방족 아미드는 100 ng 내지 10 mg, 50 mg 내지 250 mg, 1 μg 내지 1 mg 또는 100 mg 내지 500 mg의 양으로 투여 단위에 존재할 수 있다. 본 발명의 지방족 아미드의 유효량을 포함하는 투여 단위는 단일의 일일 투여량(single daily dose)으로 투여될 수 있거나, 또는 총 일일 투여량은 필요에 따라서 일일 두 번, 세 번, 네 번 또는 그 이상 분할된 투여량으로 투여될 수 있다.

- [0148] 질병 또는 질환은 본 발명의 지방족 아마이드로의 치료에 빠른 반응을 보이는(responsive) 생리학적 또는 생화학적 파라미터들 및/또는 하나 이상의 증상들을 갖는 여하한 질병 또는 질환일 수 있다. 일부 실시예들에서, 질병은 염증 질병이고, 본 발명의 지방족 아마이드는 염증을 감소시키는데 충분한 양으로 투여된다. 여하한 실시예들에서, 염증성 질환은 자가면역 질병, 류마티스성 관절염, 다발성 경화증 또는 크론병(Crohn's disease)이다.
- [0149] 여하한 실시예들에서, 질환은 통증이고, 본 발명의 지방족 아마이드는 무통증을 제공하는데 충분한 양으로 투여된다.
- [0150] 다른 실시예들에서, 질환은 고혈압이고, 본 발명의 지방족 아마이드는 혈압을 감소시키는데 충분한 양으로 투여된다.
- [0151] 아난다미드(Anandamide)는 칸나비노이드(cannabinoid: CB) 1 수용체, 그리고 정도는 덜 하지만 CB2 수용체 및 바닐로이드 1 수용체의 내인성 작용제(agonist)이다(Tan *et al.*, *supra*). 인간 및 동물 피시험체로의 아난다미드 투여는 음식 섭취 및 체중 조절; 혈압 감소; 심박동수 감소; 심근 재관류손상(myocardial reperfusion injury)에 대한 보호; 화학적, 기계적 또는 열적 자극에 의해 유발되는 급성 통증(acute pain) 감소; 신경병증성 또는 염증성 유래(neuropathic or inflammatory origin)의 만성적인 통증 감소; 염증 감소; 급성 신경 손상(예를 들어, 외상성 뇌손상, 뇌졸중 및 뇌전증) 및 만성 신경퇴행성 질환[예를 들어, 다발성 경화증, 파킨슨병, 헌팅턴병(Huntington's disease), 루게릭병(amyotrophic lateral sclerosis) 및 알츠하이머병]에 있어서 신경보호(neuroprotection)를 제공; 기관지 확장(bronchodilation)을 촉진; (예를 들어, 녹내장 환자들에 있어서) 안압 감소; 및 종양 세포의 세포사멸(tumor cell apoptosis)을 촉진을 포함하는, 무수한 생리학적 효과들을 갖는 것으로 입증되었다(예를 들어, Pacher *et al.* (2006) *Pharmacol. Rev.* 58(3):389-462 참조).
- [0152] 일부 실시예들에서, 지방족 아마이드는 본 명세서에서 설명되는 제조합 미생물들 및 방법들에 의해 생산되는 아난다미드이고, 질병 또는 질환은 아난다미드의 유효량을 투여함으로써 치료되거나 또는 예방될 수 있는 상기에 언급된 질병들 또는 질환들 중 하나이다.
- [0153] PEA 및 OEA는 과산화소체 증식제-활성화 수용체(peroxisome proliferator-activated receptor)- α (PPAR- α)의 내인성 작용제(agonist)이다[Fu *et al.* (2003) *Nature* 425(6953):90-93; and Lo Verme *et al.*, *supra*]. 인간 및 동물 피시험체로의 PEA 또는 OEA의 투여는 음식 섭취 및 체중 조절; 통증 및 염증 감소; 신경보호(neuroprotection)를 제공; 및 안압 감소를 포함하는, 아난다미드에 의해 유발되는 많은 동일한 반응들을 재현(evoke)하여 입증되었다(예를 들어, Fu *et al.*, *supra*, Tan *et al.*, *supra*, Lo Verme *et al.*, *supra*, 및 미국 특허 번호 6,348,498 및 6,656,972 참조).
- [0154] 일부 실시예들에서, 지방족 아마이드는 본 명세서에서 설명되는 제조합 미생물들 및 방법들에 의해 생산되는 PEA이고, 질병 또는 질환은 PEA의 유효량을 투여함으로써 치료되거나 또는 예방될 수 있는 상기에 언급된 질병들 또는 질환들 중 하나이다. 다른 실시예들에서, 지방족 아마이드는 본 명세서에서 설명되는 제조합 미생물들 및 방법들에 의해 생산되는 OEA이고, 질병 또는 질환은 OEA의 유효량을 투여함으로써 치료되거나 또는 예방될 수 있는 상기에 언급된 질병들 또는 질환들 중 하나이다.
- [0156] **실시예**
- [0157] 하기의 특정한 실시예들은 본 발명을 설명하기 위하여 의도된 것이며, 청구항들의 범위를 제한하는 것으로 해석되어서는 안 된다.
- [0158] **실시예 1**
- [0159] 이 실시예는 유전적으로 조작된(genetically engineered) 미생물의 구성(construction)을 설명하며, 여기서 아실-CoA 탈수소효소, 외막 단백질 수용체(outer membrane protein receptor), 피루브산 포름산 분해효소(pyruvate formate lyase), 젖산 탈수소효소(lactate dehydrogenase) 및 전사 억제인자(transcriptional repressor)가 감쇠되었다.
- [0160] 대장균 MG1655 DV4는 *fadE*(아실-CoA 탈수소효소), *fhuA*(외막 단백질 수용체), *pf1B*(피루브산 포름산 분해효소) 및 *ldhA*(젖산 탈수소효소) 유전자 결실들을 포함하는 유전적으로 조작된 대장균 K 균주이다(미국 특허 출원 공개 2011/0072714 및 2011/0162259를 참조하고, 이는 본 명세서에서 인용참조된다). 전사 억제인자를 인코딩하는 대장균 MG1655의 *fabR* 유전자(GenBank 수탁 번호 AAC76945)는, 본 명세서에서 설명되는 하기의 변형(modification)들과 함께, Datsenko *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97:6640-6645에 따라서

Lambda Red 시스템을 사용하여 *대장균* MG1655 DV4로부터 결실되었다.

[0161] 하기의 두 개의 프라이머들은 결실 균주를 생성하는데 사용되었다:

fabR_del_F: 5'-
ATGTTTATTGCGTTACCGTTCATTCACAATACTGGAGCAATCCAGTATGATTCCG
GGGATCCGTCGACC-3' (SEQ ID NO: 4); 및

fabR_del_R: 5'-
CGTACCTCTATCTTGATTTGCTTGTTCATTACTCGTCCTTCACATTTCTGTAGGC
TGGAGCTGCTTCG-3' (SEQ ID NO: 5).

[0162]

[0163] *fabR_del_F* 및 *fabR_del_R* 프라이머들은 상기에(*supra*) Datsenko *et al.*,에 기재된 바와 같이, PCR에 의해 플라스미드 pKD13로부터 카나마이신 저항성(kanamycin resistance: Km^R) 카세트(cassette)를 증폭시키는데 사용되었다. 이후에, 결과적으로 얻어진 PCR 생성물은 상기에(*supra*) Datsenko *et al.*,에 기재된 바와 같이, 사전에 3 내지 4 시간 동안 아라비노오스로 유도되는 세포들인, 플라스미드 pKD46을 함유하는 전기 적격(electro-competent) *대장균* MG1655 DV4 세포들을 형질전환시키는데 사용되었다. 37 °C에서 SOC 배지 내에서 3 시간 동안 성장 이후, 세포들은 50 µg/mL의 카나마이신을 함유하는 Luria agar 플레이트 상에 평판배양(plated)되었다. 카나마이신에 대해 저항성이 있는 콜로니들이 식별되었고, 37 °C에서 하룻밤 배양된 후 분리되었다. *fabR* 유전자의 파괴(Disruption)는 *대장균 fabR* 유전자 측면에 위치하는(flanking) 프라이머들을 사용하여 확인되었다.

[0164]

하기의 프라이머들을 사용하여 *fabR*의 결실 확인이 실시되었다:

fabR_3: 5'-GCGACGCGCGCACCTTGCTTAACCCAGGCC-3' (SEQ ID NO: 6)

[0165]

fabR_4: 5'-CGCATCTTCGCGCCAATCCAGAACC-3' (SEQ ID NO: 7).

[0166]

결실을 확인한 후, 단일 콜로니가 상기에(*supra*) Datsenko *et al.*,에 기재된 방법에 따라서 Km^R 마커를 제거하는데 사용되었다. *fadE*, *fhuA*, *pf1B*, *ldhA* 및 *fabR* 유전자 결실들을 갖는 결과적으로 얻어진 MG1655 *대장균* 균주는 *대장균* MG1655 Δ*fadE* Δ*fhuA* Δ*pf1B* Δ*ldhA* Δ*fabR* 또는 *대장균* MG1655 DG5로 명명되었다.

[0167]

이 실시예는 *대장균* MG1655 DG5의 구성을 나타내며, 이는 아실-CoA 탈수소효소, 외막 단백질 수용체, 피루브산 포름산 분해효소, 젖산 탈수소효소 및 전사 억제자가 감쇠된, 유전적으로 조작된 미생물이다.

[0168]

실시예 2

[0169]

이 실시예는 유전적으로 조작된 미생물의 구성을 설명하며, 여기서 티오에스테라아제('tesA) 및 아실-CoA 합성효소(fadD)를 인코딩하는 뉴클레오티드 서열들은 유도성 프로모터의 제어 하에서 미생물의 염색체로 통합된다.

[0170]

'tesA는 선도 서열이 없는(leaderless) *대장균 tesA* 유전자를 포함하는 뉴클레오티드 서열(GenBank entry AAC73596, Accession U00096.2)이다. 'tesA는 *대장균* 티오에스테라아제(EC 3.1.1.5, 3.1.2.-)를 인코딩하고, 여기서 맨 처음의 25 아미노산들이 결실되고, 26 위치의 아미노산인 알라닌이 메티오닌으로 대체되었다. 이후에, 이 메티오닌은 'tesA의 맨 처음 아미노산이 된다[Cho *et al.* (1995) *J. Biol. Chem.* 270:4216-4219]. *대장균 fadD*(GenBank entry AAC74875; Accession U00096.2)는 아실-CoA 합성효소를 인코딩한다.

[0171]

'tesA 또는 'tesA-fadD를 함유하는 pACYC-Ptrc 플라스미드의 구성

[0172]

'tesA 유전자는 pETDuet-1-'tesA 플라스미드로부터 얻어졌고, 이는 상기에 설명된 바와 같이 'tesA 유전자를 *Nde* I /*Avr* II 분해(digested) pETDuet-1 플라스미드(Novagen, Madison, WI)로 클로닝(cloning)시킴으로써 구성되었다(미국 특허 출원 공개 2010/0242345 및 국제 특허 출원 공개 WO 2007/136762를 참조하고, 이는 본 명세서에서 이들의 전문이 인용참조된다). *fadD* 유전자는 pHZ1.61 플라스미드(SEQ ID NO: 8)로부터 얻어졌고, 이는 상기에 설명된 바와 같이 *fadD* 유전자를 pCDFDuet-1 플라스미드(Novagen, Madison, WI)로 클로닝시킴으로써 구성되었다(또한, 미국 특허 출원 공개 2010/0257777을 참조하고, 이는 본 명세서에서 전문이 인용참조된다). 'tesA 및 *fadD* 유전자들은 하기의 프라이머들 및 고 정확도(high fidelity) PHUSION™ 증합효소(New England Biolabs, Inc., Ipswich, MA)를 사용하여, 각각 pETDuet-1-'tesA 및 pHZ1.61로부터 증폭되었다:

'*tesA*Forward- 5'-CTCTAGAAATAATTTAACTTTAAGTAGGAGAUAGGTAC
CCATGGCGGACACGTTATTGAT-3' (SEQ ID NO: 9)

'*tesA*Reverse- 5'-CTTCGAATTCATTTAAATTATTCTAGAGTCATTATGAGTC
ATGATTTACTAAAGGC-3' (SEQ ID NO: 10)

*fadD*Forward- 5'-CTCTAGAAATAATTTTAGTTAAGTATAAGAAGGAGATATAACC
ATGGTGAAGAAGGTTTGGCTTAA-3' (SEQ ID NO: 11)

*fadD*Reverse- 5'-CTTCGAATTCATTTAAATTATTCTAGAGTTATCAGGCTTTA
TTGTCCAC-3' (SEQ ID NO: 12).

[0173]

[0174]

pACYC-'*tesA* 플라스미드를 구성하기 위하여, '*tesA* PCR 생성물 및 pACYC-*Ptrc* 벡터(SEQ ID NO: 13)는 *NcoI* 및 *EcoRI*로 분해(digested)되었다. 하룻밤 동안 T4 DNA 연결효소(ligase)[New England Biolabs, Ipswich, MA]로 결찰(ligation)한 후, DNA 생성물은 TOP10® ONE SHOT® 세포들(Invitrogen, Carlsbad, CA)로 형질전환되었다. pACYC-*Ptrc* 벡터에 '*tesA*의 삽입은 제한효소 분해(restriction digestion)에 의해 확인되었다. 또한, IN-FUSION™ 클로닝을 위한 중복된 분절들(overlapping fragments) 및 *SwaI* 제한 부위(restriction site)는 '*tesA* 삽입체의 3'-말단에서 생성되었다.

[0175]

pACYC-*Ptrc*-'*tesA-fadD* 플라스미드를 구성하기 위하여, pACYC-*Ptrc*-'*tesA* 플라스미드는 하룻밤 동안 *SwaI*에 의해 제한효소 분해(restriction digestion)되었다. pHZ1.61로부터 증폭되는 *fadD* PCR 생성물은 IN-FUSION™ PCR 클로닝 시스템(Clontech, Mountain View, CA)을 사용하여 '*tesA* 유전자의 하류에(downstream)클로닝되었다. *fadD*의 삽입은 제한효소 분해(restriction digestion)로 확인되었다. *fadD*의 삽입은 '*tesA* 유전자의 뒤에 *SwaI* 부위를 파괴하나, *fadD*의 3'-말단에서 새로운 *SwaI* 부위를 재생성한다.

[0176]

Tn7tes 및 Tn7tesfad 플라스미드들의 구성

[0177]

pACYC-*Ptrc*-'*tesA* 및 pACYC-*Ptrc*-'*tesA-fadD* 플라스미드들은 각각 *Ptrc*-'*tesA* 및 *Ptrc*-'*tesA-fadD* 카세트들을 발생시키기 위하여 템플릿(template)들로서 사용되었다. 하기의 프라이머들은 카세트들을 얻기 위하여 사용되었다:

IFF: 5'-GGGTCAATAGCGGCCGCCAATTCGCGCGGAAGGCG-3' (SEQ ID NO: 14)

[0178]

IFR: 5'-TGGCGCGCCTCTAGGGCATTACGCTGACTTGACGGG-3' (SEQ ID NO: 15).

[0179]

플라스미드 pGRG25(GenBank 수탁 번호 DQ460223)는 *Not I* 및 *AvrII*(New England Biolabs, Inc., Ipswich, MA)에 의해 제한효소 분해(restriction digestion)되었다. *Ptrc*-'*tesA* 카세트는 IN-FUSION™ PCR 클로닝 시스템(Clontech, Mountain View, CA)을 사용하여 pGRG25의 *Not I* 및 *AvrII* 제한 부위들에 클로닝되어, Tn7tes 플라스미드(SEQ ID NO: 16)를 생성하였고, 여기서 *lacI_q* 및 *Ptrc*-'*tesA* 유전자들은 왼쪽 및 오른쪽 Tn7 말단들의 양 측면에 위치된다. *Ptrc*-'*tesA-fadD* 카세트는 유사하게 pGRG25의 *Not I* 및 *AvrII* 제한 부위들에 클로닝되고, 이로 인해서 Tn7tesfad 플라스미드(SEQ ID NO: 17)를 생성하였고, 여기서 *lacI_q* 및 *Ptrc*-'*tesA-fadD* 유전자들은 왼쪽 및 오른쪽 Tn7 말단들의 양 측면에 위치된다.

[0180]

대장균 MG1655 DG5 Tn7-'*tesA* 및 대장균 MG1655 DG5 Tn7-'*tesA-fadD*의 발생(Generation)

[0181]

플라스미드들 Tn7tes 및 Tn7tesfad는 각각 별도로 McKenzie et al., *BMC Microbiology*, 6:39(2006)에 기재된 프로토콜(protocol)을 사용하여 (실시예 1에 설명된) 균주 대장균 MG1655 DG5 내로 전기천공(electroporated)되었다. 전기천공(electroporation) 후에, 암피실린-저항성 세포들은 하룻밤 동안 32 °C에서 0.1 % 글루코오스 및 100 µg/mL 카르베니실린(carbenicilin)을 함유하는 LB 배지 내에서 성장시키고, 뒤이어 하룻밤 동안 32 °C에서 0.1 % 아라비노오스를 함유하는 LB 플레이트에서 성장시켜 Tn7-전위 분획물(transposition fraction)들을 포함하는 세포들에 대하여 선택함으로써 선택된다. 단일 콜로니들은 암피실린이 있는 또는 없는 LB 배지 상에 스트리킹(streaked)되었고, Tn7tes 또는 Tn7tesfad 플라스미드들을 회복시키기 위하여 42 °C에서 하룻밤 동안 성장되었다. 따라서, *lacI_q* 및 *Ptrc*-'*tesA* 또는 *lacI_q* 및 *Ptrc*-'*tesA-fadD* 유전자들은 *pstS* 및 *glsM* 유전자들 사이에 위치되는 대장균 MG1655 염색체 상의 attTn7 부위에 통합되었다. 이러한 유전자들의 통합은 하기의 프라이머들을 사용하여 PCR 및 시퀀싱(sequencing)함으로써 확인되었다:

attTn7.A: 5'-GATGCTGGTGGCGAAGCTGT-3' (SEQ ID NO: 18)

[0182]

attTn7.C: 5'-GTTGCGACGGTGGTACGCATAAC-3' (SEQ ID NO: 19).

- [0183] 따라서, 결과적으로 얻어진 균주들은 *대장균* MG1655 DG5 Tn7-'*tesA* 및 *대장균* MG1655 DG5 Tr7-'*tesA-fadD*라는 명칭을 붙였다.
- [0184] 이 실시예의 결과들은 유전적으로 조작된 미생물들의 발생을 설명하며, 여기서 티오에스테라아제를 인코딩하는 뉴클레오티드 서열(즉, *대장균* MG1655 DG5 Tn7-'*tesA*) 또는 티오에스테라아제 및 아실-CoA 합성효소를 인코딩하는 뉴클레오티드 서열(즉, *대장균* MG1655 DG5 Tn7-'*tesA-fadD*)은 유도성 프로모터의 제어 하에서 숙주 세포의 염색체 내로 통합되었다.
- [0185] **실시예 3**
- [0186] 이 실시예는 유전적으로 조합된 미생물 내에서 팔미토일푸트레신 합성효소를 인코딩하는 유전자를 발현시킴으로써 N-팔미토일에탄올아미드를 생산하는 방법을 설명한다.
- [0187] GenBank 수탁 번호 AY632377[Brady, S.F., *et al.* (2004) *J. Nat. Prod.*, 67: 1283-1286](SEQ ID NO: 2)로서 식별되는, 박테리아성 팔미토일푸트레신 합성효소(PPS)를 인코딩하는 유전자는, DNA2.0(Menlo Park, CA)에 의해 합성되었다. pJ201:30127로 칭하는 DNA2.0 플라스미드는 개시 코돈 측면에 위치하는 *NcoI* 부위 및 종결 코돈 측면에 위치하는 *PmeI* 부위를 포함하도록 설계되었다. pJ201:30127 내의 PPS 인코딩 유전자는 최적화된 코돈이 아니다. PPS 인코딩 유전자는 발현 벡터 OP80로 서브클로닝(subcloned)되었고, 이는 이전에 설명되었다(미국 특허 출원 공개 No. 2010/0154293를 참조하고, 이는 본 명세서에서 전문이 인용참조된다). OP80 벡터는 pCL1920을 기반으로 하며, 이는 IPTG-유동성 *trc* 프로모터의 제어 하에서 작동가능하게 연결된 유전자들을 발현시키는 적은 복제 플라스미드(low copy plasmid)이다. PPS 유전자를 발현시키는 OP80 벡터를 구성하기 위하여, 플라스미드들 OP80 및 pJ201:30127은 정제되었고 *NcoI/PmeI*(New England Biolabs, Inc., Ipswich, MA)로 제한 효소 분해(restriction digestions)되었다. PPS 인코딩 유전자는 T4 DNA 연결효소로 *NcoI/PmeI* 분해 OP80 플라스미드에 결합되었다. 결합 반응물(ligation reaction)은 TOP10® *대장균* 세포들로 형질전환되었고, 상기 세포들은 스펙티노마이신(spectinomycin)을 함유하는 LB 한천(agar) 상에 평판배양(plated)되었다. 열 개의 콜로니들이 선택되었고, 상기 콜로니들을 배양하고, 플라스미드 DNA를 분리시키며, 그리고 *NcoI/PmeI*로 플라스미드 DNA를 분해시킴으로써 PPS 인코딩 삽입체에 대하여 테스트하였다. 하나의 콜로니는 제한효소 분해(restriction digestion)에 의한 PPS 인코딩 삽입체를 함유하는 플라스미드에 양성(positive)이었다. 플라스미드는 서열 분석에 의해서 PPS 인코딩 유전자를 함유하는 것으로 확인되었고, "OP80-PPS"로 칭하였다. 이후에, OP80-PPS 플라스미드는 (실시예 1에 설명된) *대장균* MG1655 균주들 DG5, (실시예 2에 설명된) DG5 Tn7-'*tesA* 및 (실시예 2에 설명된) DG5 Tn7-'*tesA-fadD*로 형질전환되었다. 대조구로서, OP80 빈 벡터(empty vector)는 *대장균* MG1655 균주들 DG5, DG5 Tn7-'*tesA* 및 DG5 Tn7-'*tesA-fadD*로 형질전환되었다. 여섯 개의 모든 균주들은 추가 글루코스 오스가 없는 LB 배지 내에서 배양되었다. 각각의 배양물(culture)이 1.2의 OD₆₀₀에 이르면, IPTG는 에탄올아민 [0.1 %(v/v)]와 함께, 1 mM의 최종 농도에 첨가되었다. 24시간 후, 배양물들은 수확되었고, 에틸 아세테이트 [에틸 아세테이트의 1 용량(volume)에 배양물의 2 용량]로 추출되었으며, 그리고 유기 분획물(organic fraction)이 수집되었고, GC-MS에 의해 지방 종류의 분석을 위하여 이용되었다.
- [0188] 모든 샘플들은 풀 스캔 모드(full scan mode)[50-500 m/z]의 질량 검출기용(mass detector) 전자 이온화(electron ionization: EI)를 포함하는, 분리용 30 m×0.25 mm 0.25 μm 필름 Agilent HP-5-MS 컬럼을 갖춘 GC-MS(5975B VL MSD를 갖는 Agilent 6850 GC)에 의해 분석되었다. 1 μL의 에틸 아세테이트 추출물은 300 °C로 셋팅된 Agilent 비분취 주입구(splitless inlet) 상에 주입되었다. 컬럼의 온도는 다음과 같이 프로그램되었다: 5분 간 100 °C, 20 °C/분으로 320 °C까지 상승, 그리고 5 분간 320 °C로 유지. 운반 기체(carrier gas)인 헬륨은 1.2 mL/분의 흐름 속도로 설정되었다. OP80-PPS로 형질전환된 *대장균* MG1655 균주 DG5의 질량 스펙트럼의 해석을 포함하는 대표적인 크로마토그램은 도 1에 나타난다. 13.1 분의 머무름 시간을 갖는 피크(peak)는 OP80-PPS로 형질전환된 모든 *대장균* MG1655 균주들 DG5, DG Tn7-'*tesA* 및 DG Tn7-'*tesA-fadD*의 GC-MS 기록(traces)에서 식별되었다. 이 피크는 N-팔미토일에탄올아미드로 식별되었다. N-팔미토일에탄올아미드로 식별된 상기 피크는 빈 벡터(empty vector)로 형질전환된 여하한 대조구 균주들에서는 나타나지 않았다.
- [0189] 대략 200 mg/L의 N-팔미토일에탄올아미드는 OP80-PPS로 형질전환된 *대장균* MG1655 균주 DG5에서 생산되었다고 추정되었다. *대장균* MG1655 균주 DG5는 지방족 아실-ACP를 생산하도록 조작된 반면에, DG Tn7-'*tesA*는 유리 지방산을 생산하도록 조작되었고, DG Tn7-'*tesA-fadD*는 지방족 아실-CoA를 생산하도록 조작되었다. 이러한 결과들은 세 개의 모든 균주들이 N-팔미토일에탄올아미드를 생산한다는 것을 입증하므로, 모(parental) DG5 균주에서의 N-팔미토일에탄올아미드의 생산은 유전자 AY632377에 의해 인코딩되는 PPS 효소가 일차 에탄올아민의 존재 하에서 기질로서 C16:0 아실-ACP를 사용한다는 것을 시사한다. 도 1에서 식별된 N-팔미토일에탄올아미드는

N,O-bis(트리메틸실릴)-트리플루오로아세트아미드(BSTFA)에 의한 유도체화(derivatization)에 의해 추가적으로 특징지어졌고, 여기서 하이드록실기는 트리메틸실릴기로 보호(trimethylsilyl-protected)되었다. 도 2a에 나타낸 바와 같이, 13.3 분의 머무름 시간을 갖는 GC-MS 기록(trace) 내의 피크는 트리메틸실릴(TMS)-보호된 생성물로서 식별되었다.

[0190] 이 실시예의 결과들은, 지방족 아실-ACP를 생산하도록 조작된 박테리아 균주 내에서 유전자 AY632377에 의해 인코딩되는 PPS 효소를 발현시킴으로써 *N*-팔미토일에탄올아미드를 생산하는 방법을 설명한다. 또한, 이러한 결과들은 단일 효소(즉, PPS)가 *생체 내에서(in vivo)* 일차 아민의 선두 기(head group) 및 아실-ACP로부터의 직접적인 이차 아미드화를 실시할 수 있다는 것을 시사한다.

[0191] **실시예 4**

[0192] 이 실시예는 유전적으로 조작된 미생물 내에서 팔미토일푸트레신 합성효소를 인코딩하는 유전자를 발현시킴으로써 *N*-팔미토일에탄올아미드를 생산하는 방법을 나타낸다.

[0193] OP80 빈 벡터(empty vector) 또는 OP80-PPS 플라스미드는 실시예 3에 설명된 바와 같이, *대장균* MG1655 균주들 DG5, DG5 Tn7- '*tesA*' 및 DG5 Tn7- '*tesA-fadD*'로 형질전환되었다. 각각의 여섯 개의 균주들은 추가 글루코오스가 없는 LB 액체 배지(broth) 내에서 하룻밤 동안 배양되었다. 이후에, 하룻밤 동안 배양된 각각의 배양물들은 글루코오스를 함유하는 영양분이 풍부한 배지 내에 접종(inoculated)되었다. 좀 더 구체적으로, 1 mL의 하룻밤 동안 배양된 각각의 배양물들은 세 개의(in triplicate) 영양분이 풍부한 2N-BT[2 % 글루코오스, 질소 제한 배지(nitrogen limited medium), 0.2 M Bis-Tris, pH 7.0, 0.1 % Triton]: 1 mM IPTG, 항생제 및 1 % 에탄올아민을 함유하는 LB 배지(1:10)에 접종되었고, pH-제어 인큐베이터에서 배양되었다. 48시간 후, 각각의 배양물의 OD₆₀₀이 기록되었고, 세포들이 수확되었으며, 에틸 아세테이트[에틸 아세테이트의 1 용량(volume)에 배양물의 2 용량]로 추출되었다. 유기 분획물(organic fraction)들이 수집되었고, 실시예 3에 설명된 바와 같이 GC-MS 분석을 위하여 이용되었다. 13.1 분의 머무름 시간을 갖는 피크는 OP80-PPS로 형질전환된 *대장균* MG1655 균주 DG5 Tn7- '*tesA-fadD*'에 대한 GC-MS 기록(trace)에서 가장 풍부하게(most abundant) 발견되었다. 13.1 분의 머무름 시간을 갖는 피크는 OP80-PPS로 형질전환된 균주 DG5에서 두번째로 가장 풍부하였고, OP80-PPS로 형질전환된 균주 DG5 Tn7- '*tesA*'에서 세번째로 가장 풍부하였다. 13.1 분의 피크는 모든 균주들에서 *N*-팔미토일에탄올아미드로 식별되었다. *N*-팔미토일에탄올아미드로 식별된 피크는 빈 벡터(empty vector)로 형질전환된 여하한 *대조구* DG5 균주들에서는 나타나지 않았다.

[0194] 상기에 언급한 바와 같이, *대조구* MG1655 균주 DG5 Tn7- '*tesA-fadD*'는 지방족 아실-CoA를 생산하도록 조작되었다. 그러므로, DG5 Tn7- '*tesA-fadD*' 균주에서의 *N*-팔미토일에탄올아미드의 생산은, 유전자 AY632377에 의해 인코딩되는 PPS 효소가 일차 에탄올아민의 존재 하에서 기질로서 지방족 아실-CoA도 사용할 수 있다는 것을 시사한다.

[0195] 이 실시예의 결과들은, 지방족 아실-ACP를 생산하도록 조작된 박테리아 균주 내에서 유전자 AY632377에 의해 인코딩되는 PPS 효소를 발현시킴으로써 *N*-팔미토일에탄올아미드를 생산하는 방법을 나타낸다. 또한, 이러한 결과들은 단일 효소(즉, PPS)가 *생체 내에서(in vivo)* 일차 아민의 선두 기(head group) 및 지방족 티오에스테르로부터의 직접적인 이차 아미드화를 실시할 수 있다는 것을 시사한다.

[0196] **실시예 5**

[0197] 이 실시예는 유전적으로 조작된 미생물 내에서 PPS를 인코딩하는 유전자를 발현시킴으로써 포화 및 불포화 지방족 아미드들을 생산하는 방법을 나타낸다.

[0198] 지방족 3-디메틸아미노-1-프로필아미드는 양쪽성 세제(amphoteric detergent)인 코카미도프로필 베타인(cocamidopropyl betaine: CAPB)의 합성에 있어서 전구체이다. 지방족 *N*-(3-디메틸아미노-1-프로필아민)아미드가 PPS를 발현시키도록 유전적으로 조작된 미생물 내에서 생산될 수 있는지를 결정하기 위하여, *대장균* MG1655 균주 DG5 세포들은 OP80 빈 벡터(empty vector) 또는 OP80-PPS 플라스미드로 형질전환되었고, 추가 글루코오스가 없는 LB 액체 배지(broth) 내에서 하룻밤 동안 배양되었다. 이후에, 하룻밤 동안 배양된 각각의 배양물들은 글루코오스를 함유하는 영양분이 풍부한 배지 내에 접종되었다. 좀 더 구체적으로, 1 mL의 하룻밤 동안 배양된 각각의 배양물들은 세 개의(in triplicate) 영양분이 풍부한 2N-BT[2 % 글루코오스, 질소 제한 배지(nitrogen limited medium), 0.2 M Bis-Tris, pH 7.0, 0.1 % Triton]: 1 mM IPTG, 항생제 및 1 % 일차 아민 3-디메틸아미노-1-프로필아민을 함유하는 LB 배지(1:10)에 접종되었고, pH-제어 인큐베이터에서 배양되었다.

60시간 후, 각각의 배양물의 OD₆₀₀이 기록되었고, 세포들이 수확되었으며, 에틸 아세테이트[에틸 아세테이트의 1 용량(volume)에 배양물의 2 용량]로 추출되었다. 유기 분획물들이 수집되었고, 실시예 3에 설명된 바와 같이 GC-MS 분석을 위하여 이용되었다. 크로마토그램은 이온 58에 대한 추출 이온 크로마토그램(extraction ion chromatogram)에 의해 분석되었고, 상기 이온 58은 지방족 N-(3-디메틸아미노-1-프로필아민)아미드에 대한 일반적인 이온이다. C12:0(12.8 분), C14:0 (13.2 분), C16:1(13.5 분), C16:0 (13.6 분) 및 C18:1(14.3 분) 지방족 사슬들을 함유하는 지방족 N-(3-디메틸아미노-1-프로필아민)아미드들이 GC-MS 기록(trace)에서 식별되었다(도 3). C16:0 지방족 사슬을 함유하는 지방족 아미드들이 GC-MS 기록(trace)에서 가장 풍부한 것으로 식별되었다(도 3).

[0199] 이 실시예의 결과들은, 3-디메틸아미노-1-프로필아민을 함유하는 배지 내에서 유전자 AY632377에 의해 인코딩되는 PPS 효소를 추가적으로 발현시키고 아실-ACP를 생산하도록 조작된 *대장균* 균주를 배양시킴으로써 0 또는 1 개의 불포화를 갖는 다양한 지방족 사슬 길이를 갖는 지방족 아미드들을 생산하는 방법을 나타낸다.

[0200] **실시예 6**

[0201] 이 실시예는 유전적으로 조작된 미생물에 다양한 일차 아민들을 공급함으로써 다양한 지방족 아미드들을 생산하는 방법을 설명한다.

[0202] OP80 빈 벡터(empty vector) 또는 OP80-PPS 플라스미드는 실시예 3에 설명된 바와 같이, *대장균* MG1655 균주들 DG5, DG5 Tn7-*'tesA* 및 DG5 Tn7-*'tesA-fadD*로 형질전환되었다. 이후에, 하룻밤 동안 배양된 각각의 배양물들은 글루코오스를 함유하는 영양분이 풍부한 배지 내에 접종되었다. 좀 더 구체적으로, 1 mL의 하룻밤 동안 배양된 각각의 배양물들은 영양분이 풍부한 2N-BT[2 % 글루코오스, 질소 제한 배지(nitrogen limited medium), 0.2 M Bis-Tris, pH 7.0, 0.1 % Triton]: 1 mM IPTG, 항생제, 그리고 (±)-1-아미노-2-프로판올; 2-메톡시에틸아민; 3-아미노-1-프로판올; 2-아미노-1,3-프로판디올; 3-메톡시프로필아민; N-(2-하이드록시에틸)에틸렌디아민 또는 부틸아민 중 하나의 1 %를 함유하는 LB 배지(1:10)에 접종되었다. 배양물들은 pH-제어 인큐베이터에서 60 시간 동안 배양되었다. 각각의 배양물의 OD₆₀₀이 기록되었고, 세포들이 수확되었으며, 에틸 아세테이트[에틸 아세테이트의 1 용량(volume)에 배양물의 2 용량]로 추출되었다. 유기 분획물들이 수집되었고, 실시예 3에 설명된 바와 같이 GC-MS 분석을 위하여 이용되었다. 지방족 아미드들은 일차 아민이 제공된 각각의 PPS-발현 *대장균* 균주들로부터 얻어졌다. 각각의 공급된 기질로부터 얻어진 지방족 아미드 생성물은 도 4에 표현되었다. 지방족 아미드 생성물들은 빈 벡터(empty vector)로 형질전환된 *대장균* 균주들에서는 검출되지 않았다.

[0203] 이 실시예의 결과들은, 일차 아민의 공급 타입(primary amine feed type)을 변화시킴으로써 유전자 AY632377에 의해 인코딩되는 PPS 효소를 추가적으로 발현시키고 지방족 티오에스테르P를 생산하도록 유전적으로 조작된 *대장균* 균주에서 다른(distinct) 지방족 아미드 종류를 생산하는 방법을 설명한다.

[0204] **실시예 7**

[0205] 이 실시예는 유전적으로 조작된 미생물 내에서의 PPS의 동족체(homolog)의 발현이 PPS가 발현되는 경우에 생산되는 동일한 지방족 아미드 화합물들을 생산하는 것을 나타낸다.

[0206] AY632377(SEQ ID NO: 2)에 의해 인코딩되는 PPS는 BLAST 분석에 의해 알려진 여하한 다른 서열에 대하여 20 % 초과와 서열 동일성을 갖지 않도록 사전에 결정되었다[Brady *et al.* (2004) *J. Not. Prod.*, 67:1283-1286 참조]. 유전자 AY632377, 즉 GenBank 수탁 번호 AAV33349(SEQ ID NO: 1)에 의해 인코딩되는 PPS 효소의 아미노산 서열은 생명공학 정보 국립 센터(National Center for Biotechnology Information: NCBI) 데이터베이스의 BLAST 검색되었다. 효소, N-(4-아미노-2-하이드록실부틸)테트라메칸아미드 합성효소(AhtS)(GenBank 수탁 번호 ACX33975.1)(SEQ ID NO: 3)를 인코딩하는 동족체는, 유전자 AY632377에 의해 인코딩되는 PPS의 아미노산 서열에 대하여 38 % 동일한 아미노산 서열을 갖는 것으로 확인되었다. AhtS를 인코딩하는 유전자는 GENEART™(Life Technologies, Grand Island, NY)에 의해 합성되었고, 하기와 같이 발현 벡터 OP80으로 클로닝되었다. 플라스미드 OP80은 정제되었고, *NcoI/PmeI*(New England Biolabs, Inc., Ipswich, MA)로 제한효소 분해(restriction digestions)되었다. AhtS 유전자는 하기의 프라이머들에 의해 IN-FUSION™ PCR 클로닝 시스템(Clontech, Mountain View, CA)을 사용하여 OP80으로 클로닝되었다:

3.10.10-2_InfusF: 5'-GAGGAATAAACCATGCCCCATTCTTGAAAGCGTGGG-3' (SEQ ID NO: 20) 및
 3.10.10-2_InfusR: 5'-AGCTGGAGACCGTTTAAACTTATAAACCGCTGTTTGTGCGCAACCG-3' (SEQ ID NO: 21).

[0207]

[0208]

두 콜로니들은 제한효소 분해(restriction digestion)에 의한 AhtS-인코딩 삽입체를 함유하는 플라스미드에 양성(positive) 이었다. 플라스미드는 서열 분석에 의해서 AhtS-인코딩 유전자를 함유하는 것으로 확인되었고, "OP80-AhtS"로 칭하였다. OP80-AhtS 플라스미드는 *대장균* MG1655 균주들 DG5, Tn7-'*tesA* 및 DG5 Tn7-'*tesA-fadD*로 형질전환되었다. 세 개의 균주들은 배양되었고, 실시예 4에 설명된 바와 같이 유도(induced)되었으며, 그리고 각각의 배양물에 1%의 최종 농도로 3-디메틸아미노-1-프로필아민, (±)-1-아미노-2-프로판올 또는 에탄올아민 중 하나가 공급되었다. pH-제어 인큐베이터에서 60 시간 동안 배양한 후, 배양물들은 수확되었고, 에틸아세테이트[에틸 아세테이트의 1 용량(volume)에 배양물의 2 용량]로 추출되었다. 유기 분획물들이 수집되었고, 실시예 3에 설명된 바와 같이 GC-MS 분석을 위하여 이용되었다.

[0209]

지방족 아미드들은 각각의 일차 아민들이 제공된 각각의 AhtS-발현 *대장균* 균주들로부터 얻어졌다. 3-디메틸아미노-1-프로필아민이 공급된 OP80-AhtS로 형질전환된 *대장균* MG1655 DG5 Tn7-'*tesA-fadD*에 의해 생산되는 생성물들에 대한 대표적인 GC-MS 기록(trace)은 도 5에 제공된다. 11.4 분의 GC 머무름 시간을 갖는 피크는 MS 분석에 의해 C14:0 지방족 N-(3-디메틸아미노-1-프로필아미드)로 확인되었다(도 5). 지방족 아미드 생성물은 빈 벡터(empty vector)로 형질전환된 *대장균* 균주들에서는 검출되지 않았다. 가장 많은 양의 아미드들은 *대장균* MG1655 균주 DG5 Tn7-'*tesA-fadD*에서 생산된 반면에, 가장 적은 양의 아미드들은 *대장균* MG1655 균주 DG5 Tn7-'*tesA*에서 생산되었다. DG5 Tn7-'*tesA-fadD* 균주에서의 아미드의 생산 증가는 AhtS 효소가 C14:0 지방족 티오에스테르 기질들에 대한 선호를 갖는다는 것을 시사한다. 또한, 이러한 데이터는 AhtS 및 PPS 둘 모두가 지방족 아미드를 만들기 위하여 *대장균* 내에서 각각의 두 개의 지방족 티오에스테르 기질들(즉, 지방족-ACPs 및 지방족-CoAs)을 사용할 수 있다는 점을 시사한다.

[0210]

이 실시예의 결과들은 AhtS 효소가 PPS 효소처럼 일차 아민과 지방족 티오에스테르 간의 동일한 타입의 반응을 촉진시킨다는 것을 입증한다. 또한, 이 실시예에 나타난 데이터는 AhtS 효소가 C14:0 지방족 티오에스테르 기질들에 대한 선호를 갖는다는 것을 나타낸다.

[0211]

실시예 8

[0212]

이 실시예는 본 발명에 따라서 지방족 아미드의 발생(generation)에 있어서 출발 물질(starting material)로서 유용한 일차 아민을 생체 내에서(*in vivo*) 발생시키는 방법을 제공한다.

[0213]

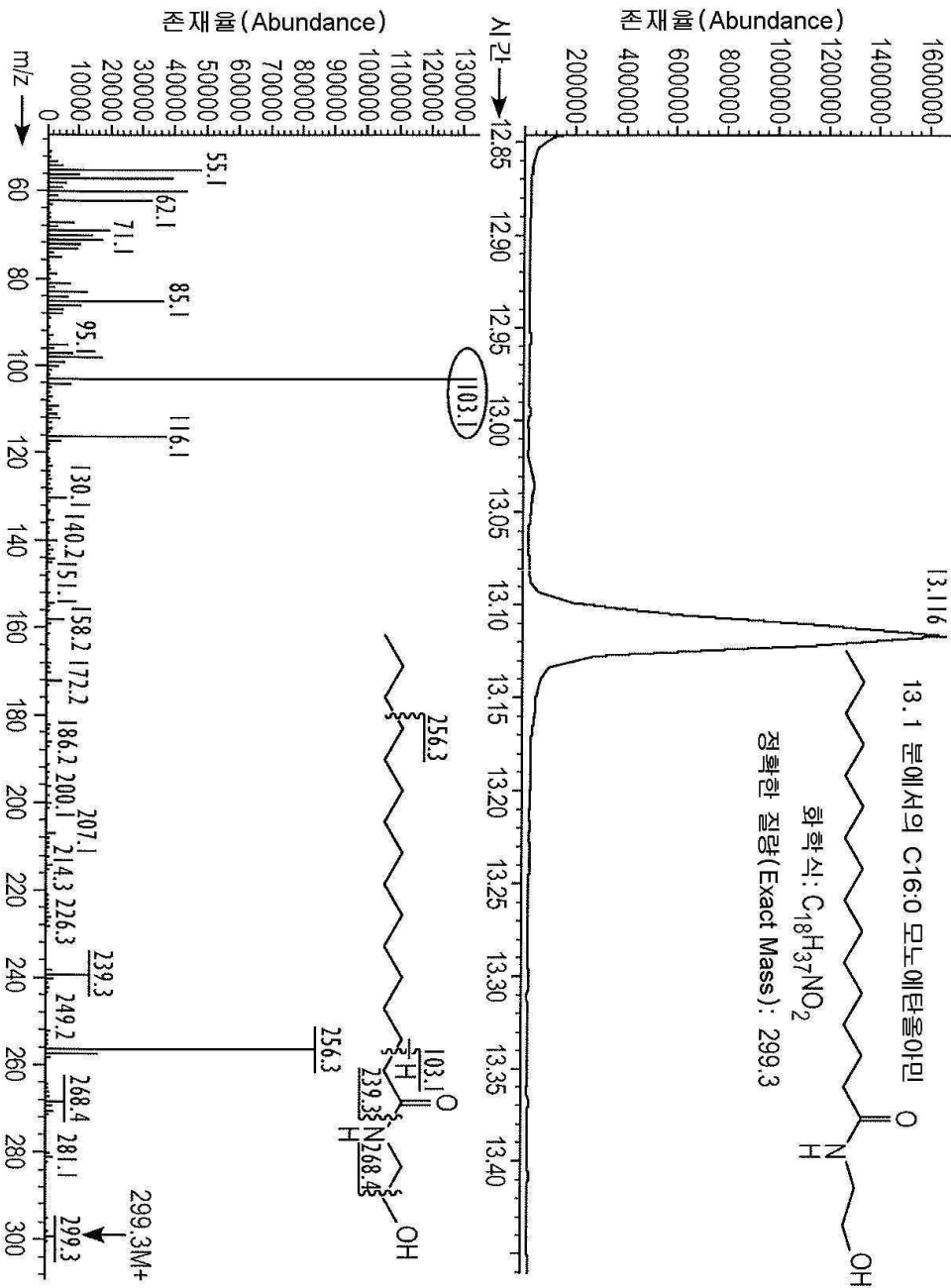
생체 내에서(*in vivo*) 에탄올아민의 생산은 세린 생합성 및 세린 탈카르복실화 경로들을 유전적으로 증가시킴으로써 달성될 수 있다. 이렇게 하기 위해서, 당분해 중간체(glycolytic intermediate)인 3-포스포글리세레이트는 포스포글리세레이트 뮤타아제(phosphoglycerate mutases)(*gpm AB*)를 과발현시키기 위하여 *대장균* 균주 MG1655를 조작함으로써 증가된다. 세린 생산 경로는 포스포글리세레이트 탈수소효소(*serA*), 3-포스포세린 아미노기전이효소(aminotransferase)(*serC*) 및 3-포스포세린 포스페이트(*serB*)를 과발현시킴으로써 조작된다. 세린을 에탄올아민으로 탈카르복실화시키는 이종(heterologous) 세린 탈카르복실화효소(SDC)가 숙주에서 발현된다. 균주가 에탄올아민 및 세린을 신진대사 시키는 것을 방지하기 위하여, 분해 효소(degradation enzyme)들인 에탄올아민 암모니아-분해효소(*eutABC*) 및 세린 탈아미노기효소(deaminase)(*sdaAB*)를 인코딩하는 유전자들은 결실된다(도 6). 이후에, 지방족 아미드는 PPS 또는 AhtS와 같은 폴리펩티드를 과발현시킴으로써 에탄올아민을 생산하는 재조합 미생물 내에서 생산될 수 있으며, 상기 폴리펩티드는 에탄올아민 및 아실 티오에스테르의 지방족 아미드로의 전환을 촉진시킨다.

[0214]

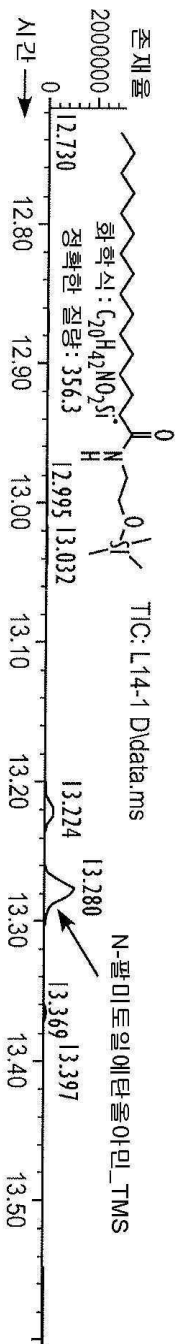
본원 발명에 대한 다양한 변형(modification)들 및 변화(variation)들은 본 기술분야의 당업자들에게 본 발명의 범위 및 정신을 벗어나지 않는다고 인식될 것이다. 본 발명이 특정 바람직한 실시예들과 관련되어 설명될지라도, 청구항들이 이러한 특정 실시예들로 지나치게 제한되어서는 안된다는 것을 이해해야 한다. 실제로, 본 기술분야의 당업자들에 의해 이해되는 본 발명을 실시하기 위한 설명된 모드(mode)들의 다양한 변형들은 청구항들의 범위 내에 포함되는 것으로 의도된다.

도면

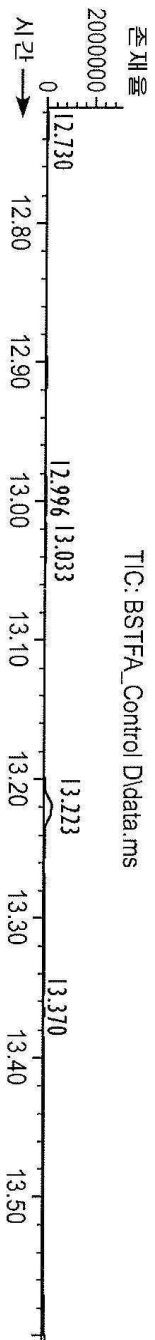
도면1



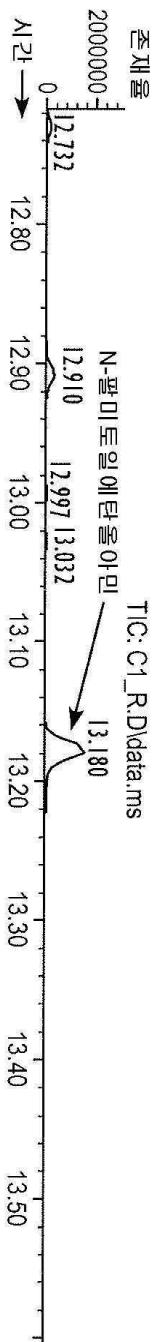
도면2a



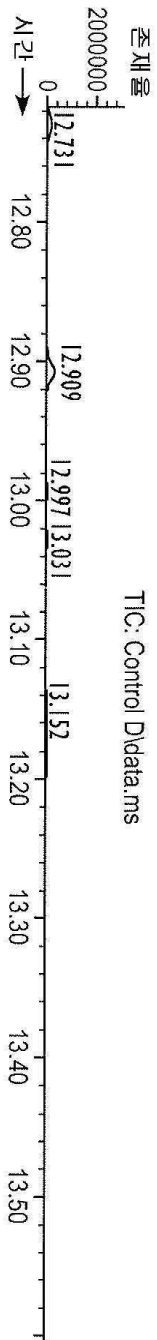
도면2b



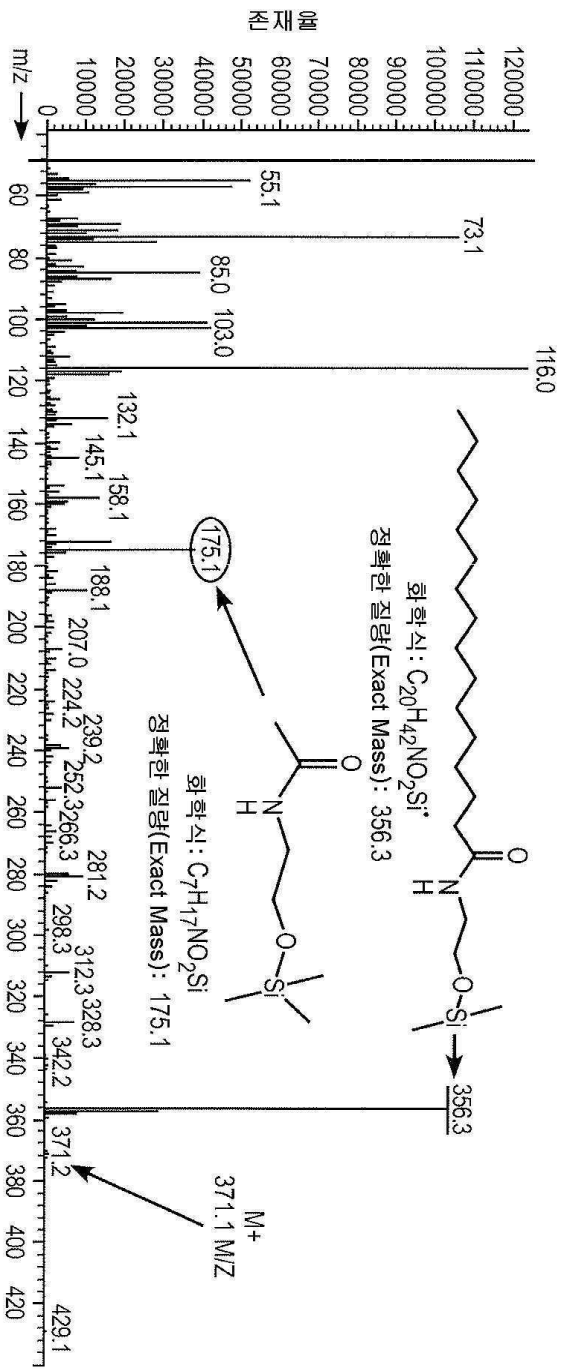
도면2c



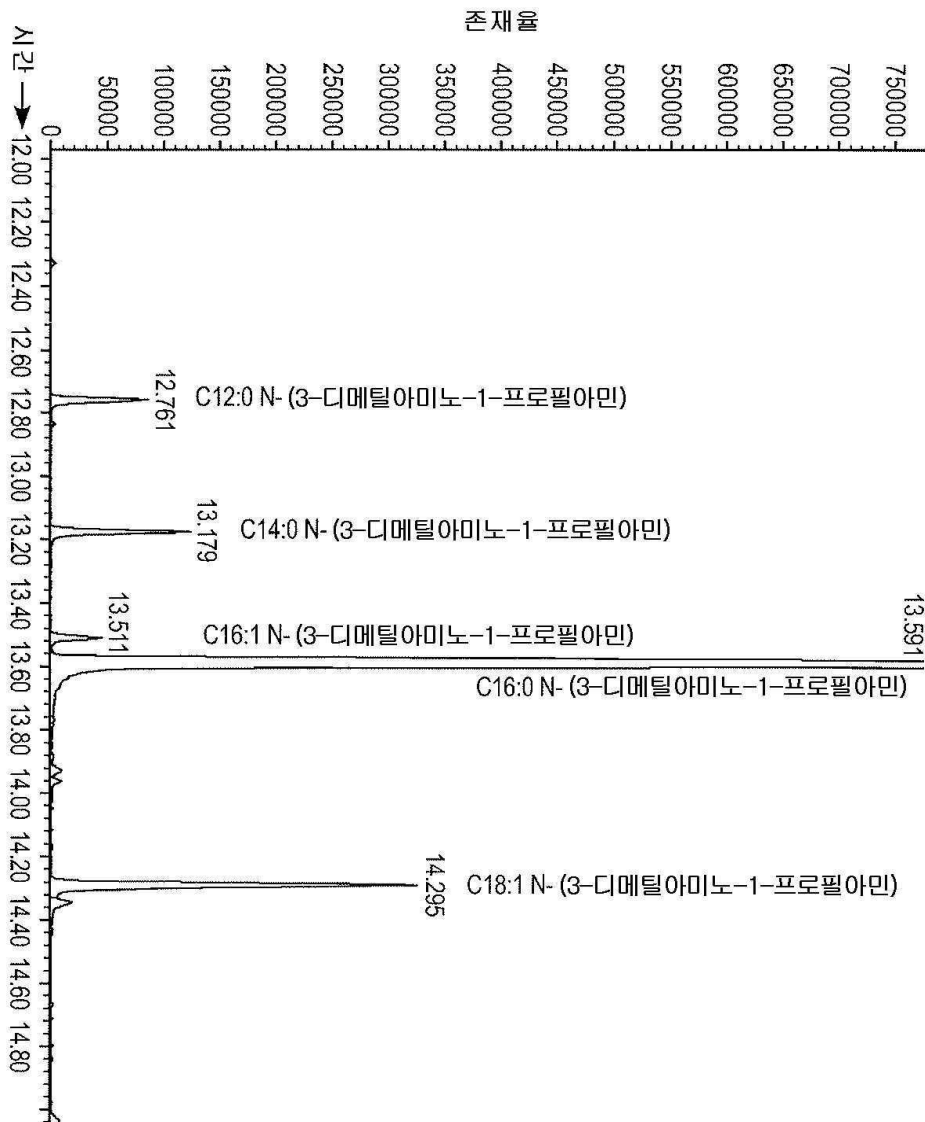
도면2d



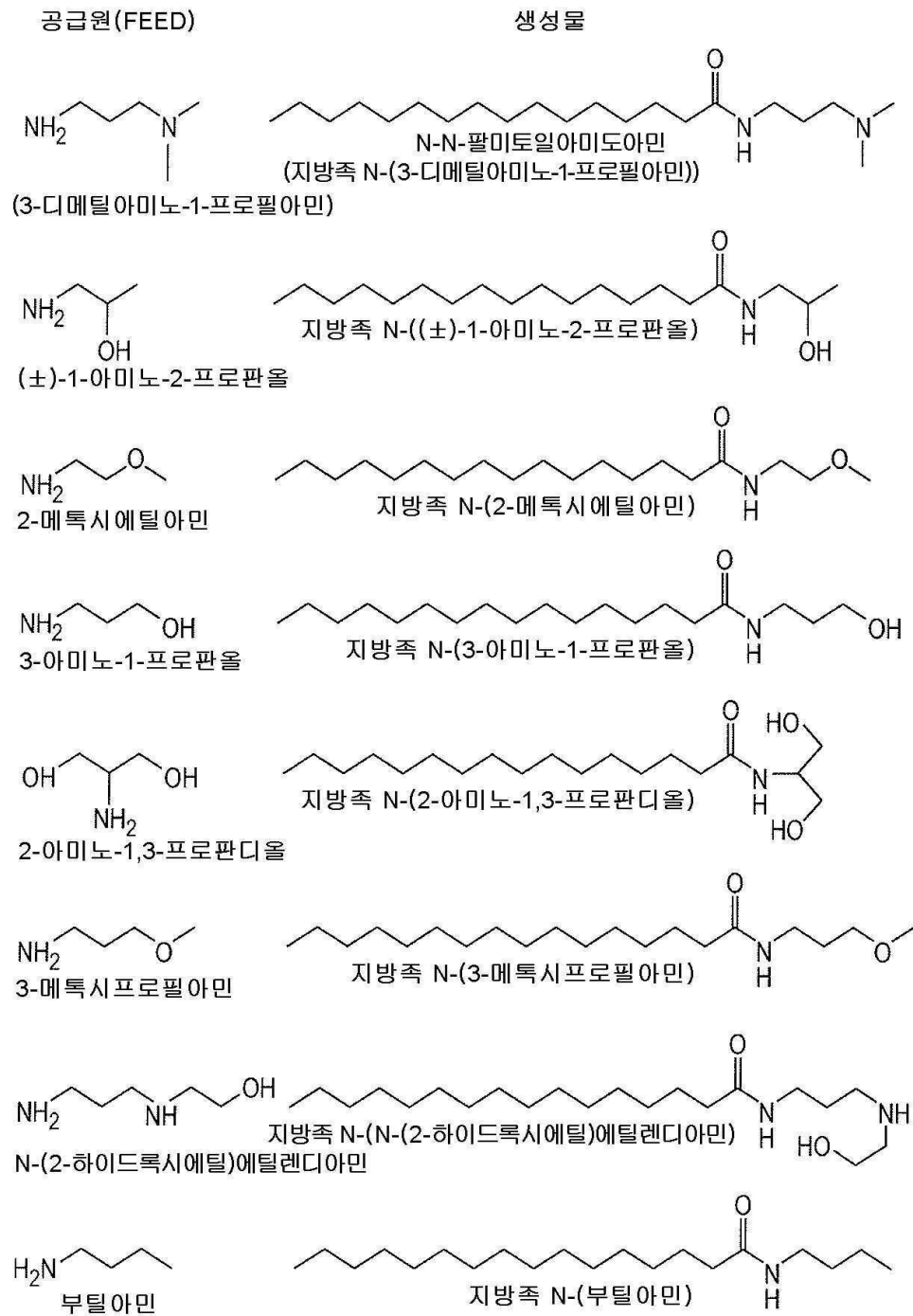
도면2e



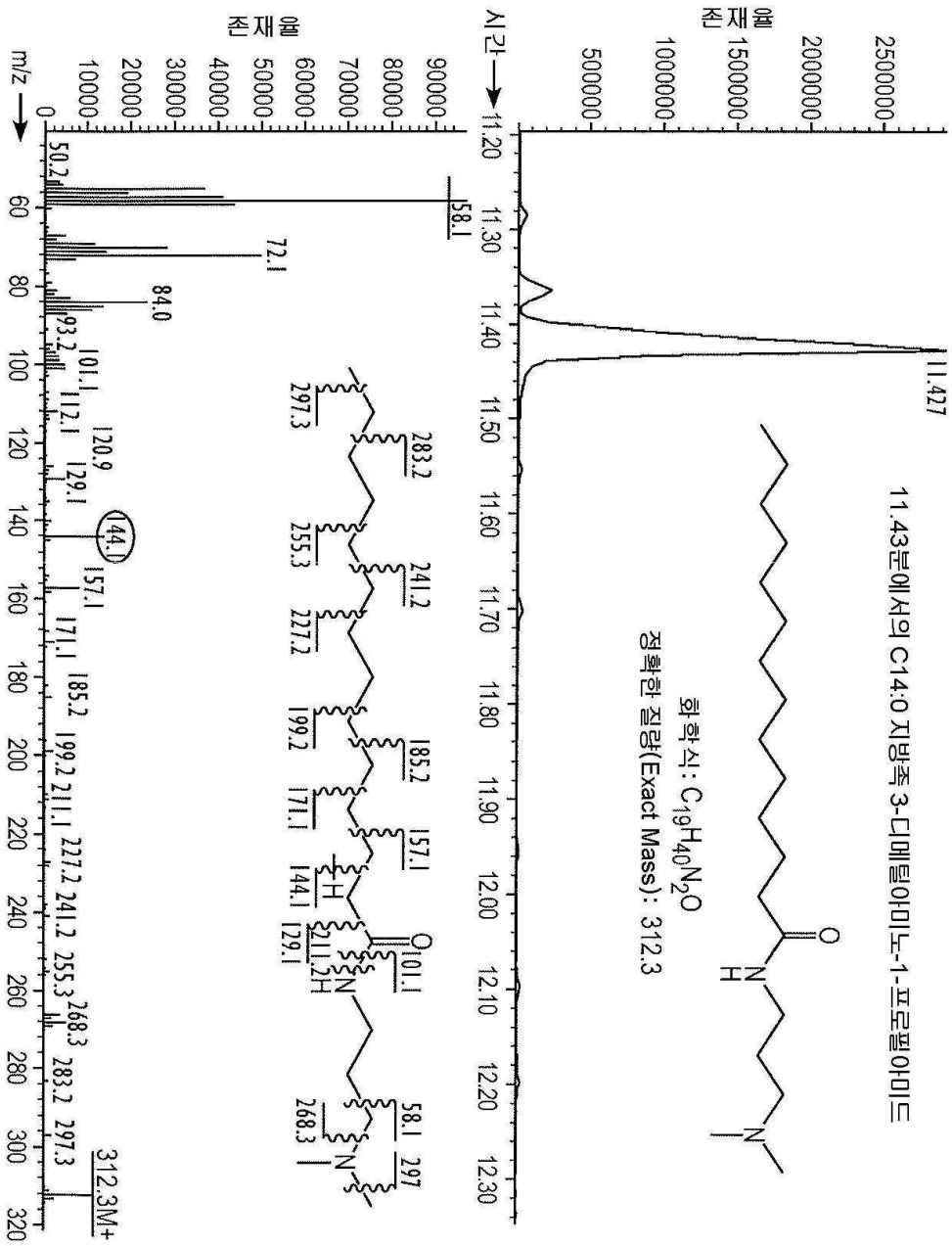
도면3



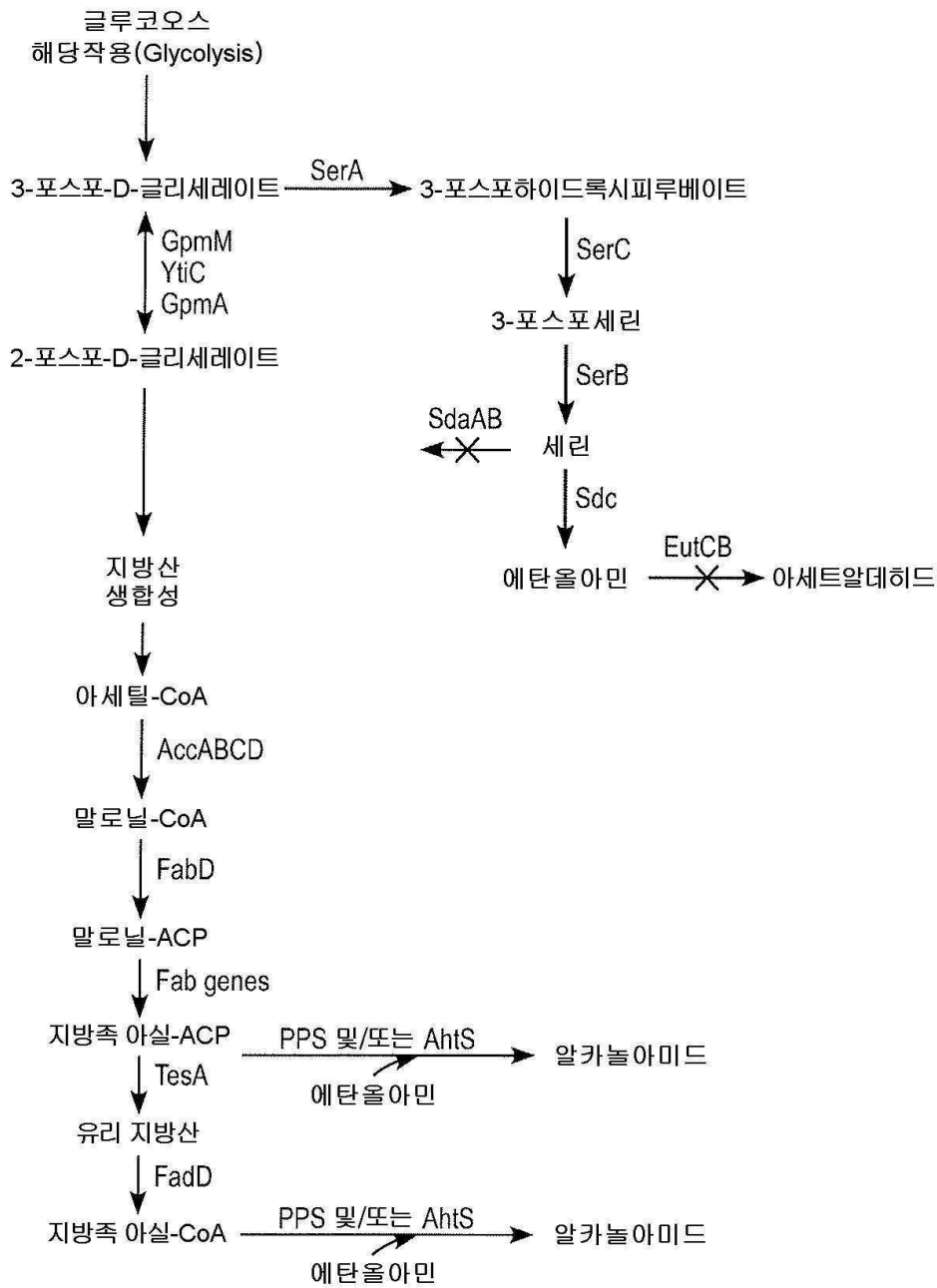
도면4



도면5



도면6



서열 목록

SEQUENCE LISTING

<110> REG Life Sciences, LLC

<120> MICROBIAL PRODUCTION OF ALKANOLAMIDES AND AMIDOAMINES AND USES

THEREOF

<130> FPA11756-D01/C

<140><141><150> 61/623,711

<151> 2012-04-13

<160> 22

<170> PatentIn version 3.5

<210> 1

<211> 207

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown: Uncultured bacterium"

<400> 1

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

1 5 10 15

Ala Leu Glu Glu Ile Glu Glu Ser Val Thr Leu Phe Glu Ala Val Ser

 20 25 30

Ala Glu Met Gly Trp Ser Pro Gly Asp Arg Leu Arg Ala Tyr Ala Ala

 35 40 45

Asn Ala Glu Tyr Leu Met Ala Leu Val Gly Gly Ala Leu Ala Gly Ala

 50 55 60

Val Gln Ile Val Leu Ser Asn Gly His Asp Pro Met Pro Ser Asp His

65 70 75 80

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

 85 90 95

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

 100 105 110

Gly Pro Leu Gly Val Glu Met Trp Arg Leu Cys Arg Asp Ala Gly Lys

 115 120 125

Ser Glu Ile Trp Ile Glu Cys Val Pro Arg Asn Leu Ala Val Tyr Arg

 130 135 140

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

145 150 155 160

Gly Glu Pro Cys Tyr Leu Ala Arg Met Gly Val Glu Ala Val Ala Glu

 165 170 175

Ser Leu Arg Ala Arg Ala Gly Arg Cys Asp Gly Phe Ala Arg Leu Val

180 185 190
 Ala Gln Ala Tyr Arg Asp Thr Gly Trp Arg Ser Met Gly Leu Ala
 195 200 205

<210> 2

<211> 624

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown: Uncultured bacterium"

<400> 2

atggaagt tgtacaatgg cgacggattc gtcacgcaa gggcgacggc cctggaagag 60
 atcaggaggt cggcacact ctctcaggcg gtctccgccg aaatgggatg gagccctggg 120
 gatcggttc gggcatatgc agcgaacgcg gactacctga tggcgctagt cggcggagcg 180
 ctggctgggg cggatgcagat cgtgctctcc aacggtcagc atcctatgcc ttcggaccac 240
 gtctggccgg aacttgcggc gggctcagca tatgacctg cgcacgcgtc agtgatggcg 300

 ctgccccgc gcttccgggg agcggcagga ctattcgggc cgcttggcgt cgagatgtgg 360
 aggctttgcc gggacgccgg gaaatccgag atatggatcg agtgcgtgcc gaggaacctt 420
 gcggtgtacc gaagatgtgg ttggccgctg gaggtggcgg gcgagttgcg gaccatttc 480
 ggcgagcctt gctacctggc gcgaatgggc gtggaggcgg tggcggaatc tctgagggcc 540
 agagccggcc gctgcgatgg gttcggcaga ctggtcgcgc aggcttaccg agacaccgga 600
 tggcggtaaa tggggcttgc gtaa 624

<210> 3

<211> 220

<212> PRT

<213> Unknown

<

<220><221> source

<223> /note="Description of Unknown: Uncultured bacterium RM44"

<400> 3

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

Met Val Gly Ile Leu Glu Gly Ile Ala Ala Glu Leu Ser Trp Arg Pro
 35 40 45
 Gly Thr Gln Leu Arg Asp Tyr Gln Asp Arg Ala Ala His Leu Ala Val
 50 55 60
 Leu Val Gly Ser Glu Ile Val Gly Gly Leu Gln Ile Val Thr Ser Pro
 65 70 75 80
 Ser Ala Asp Cys Leu Pro Tyr Arg Leu Val Trp Pro Glu Val Cys Val
 85 90 95
 Pro Asp Gly Ala Ala Ile Ala Asp Ile Thr Ile Leu Ala Leu Arg Lys
 100 105 110
 Glu Tyr Arg Ala Arg Phe Asn Leu Phe Trp Pro Leu Cys Val Glu Leu
 115 120 125
 Trp Arg His Cys Val Ala Glu Gly Ala Thr Glu Met Arg Leu Glu Ala
 130 135 140
 Thr Pro Asp Thr Leu Arg Leu Tyr Arg Arg Ile Gly Trp Pro Leu Glu
 145 150 155 160
 Val Ile Gly Asp Leu Arg Leu His Trp Asn Glu Pro Cys Phe Leu Cys
 165 170 175
 Arg Met Gly Ile Val Asp Val Ala Gly Ala Met Val Val Arg Ala Leu
 180 185 190
 Gln Ser Ala Thr Tyr Gln Ala Val Leu Ala Gly Met Ser Arg Pro Val
 195 200 205
 Ala Ser Ala Ser Pro Val Ala Thr Asn Ser Gly Leu
 210 215 220
 <210> 4
 <211> 70
 <212> DNA
 <213> Artificial Sequence
 <220><221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 primer"
 <400> 4

atgttttatt gcgttaccgt tcattcaciaa tactggagca atccagtatg attccgggga 60

tccgtcgacc 70

<210> 5

<211> 70

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 5

cgtacctcta tcttgatttg cttgtttcat tactcgtcct tcacatttcc tgtaggctgg 60

agctgcttcg 70

<210> 6

<211> 30

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 6

gcgacgcgcg caccttgctt aaccaggccc 30

<210> 7

<211> 27

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 7

cgcattctcg cgcaatcca gaacacc 27

<210> 8

<211> 6700

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 8

```

ggggaattgt gagcggataa caattcccct gtagaaataa ttttgtttaa cttaataag      60
gagatatacc atggtgaaga aggtttggct taaccgttat cccgcggacg ttccgacgga      120
gatcaaccct gaccgttata aatctctggt agatatgttt gagcagtcgg tcgcgcgcta      180
cgccgatcaa cctgcgtttg tgaatatggg ggaggtaatg accttccgca agctggaaga      240
acgcagtcgc gcgtttgccg cttatttgca acaagggttg gggctgaaga aaggcgatcg      300
cgttgcgctg atgatgccta atttattgca atatccggtg gcgctgtttg gcattttgcg      360
tgccgggatg atcgtcgtaa acgttaacct gttgtatacc ccgctgagc ttgagcatca      420

gcttaacgat agcggcgcat cggcgattgt tatcgtgtct aactttgctc acacactgga      480
aaaagtgggt gataaaaccg ccgttcagca cgtaatctcg acccgatagg gcgatcagct      540
atctacggca aaaggcacgg tagtcaattt cgttgttaaa tacatcaagc gtttggtgcc      600
gaaataccat ctgccagatg ccatttcatt tcgtagcgca ctgcataacg gctaccggat      660
gcagtacgtc aaaccegaac tggtgccgga agatthagct tttctgcaat acaccggcgg      720
caccactggt gtggcgaaag gcgcgatgct gactcaccgc aatatgctgg cgaacctgga      780
acaggttaac gcgacctatg gtccgctgtt gcatccgggc aaagagctgg tggtagcggc      840

gctgccgctg taicacattt ttgccctgac cattaactgc ctgctgttta tcgaactggg      900
tgggcagaac ctgcttatca ctaaccgcg cgatattcca gggttggtaa aagagttagc      960
gaaatatccg ttaccgcta tcacggcgt taacaccttg ttcaatgcgt tgctgaacaa      1020
taaagagttc cagcagctgg atttctccag tctgcatctt tccgcaggcg gagggatgcc      1080
agtgcagcaa gtggtggcag agcgttgggt gaaactgaca ggacagtatc tgctggaagg      1140
ctatggcctt accgagtgtg cgccgctggt cagcgttaac ccatatgata ttgattatca      1200
tagtggtagc atcggtttgc cggtgccgtc gacggaagcc aaactggtgg atgatgatga      1260

taatgaagta ccaccgggtc aaccgggtga gctttgtgtc aaaggaccgc aggtgatgct      1320
gggttactgg cagcgtccgg atgctacaga tgagatcatc aaaaatggct ggttacacac      1380
cggcgacatc gcggtgatgg atgaagaagg gttcctgcgc attgtcgatc gtaaaaaaga      1440
catgattctg gtttccggtt ttaactgcta tccaacgag attgaagatg tcgtcatgca      1500
gcatcctggc gtacaggaag tcgcggctgt tggcgtacct tccggctcca gtggtgaagc      1560

```

ggtgaaaatc ttcgtagtga aaaaagatcc atcgcttacc gaagagtcac tggtagacctt 1620
 ttgccgccgt cagctcacgg gctacaaagt accgaagctg gtggagtffc gtgatgagtt 1680

 accgaaatct aacgtcggaa aaatcttgcg acgagaatta cgtgacgaag cgcgcggcaa 1740
 agtggacaat aaagcctgaa agcttgcggc cgcataatgc ttaagtcgaa cagaaagtaa 1800
 tcgtattgta cacggccgca taatcgaaat taatacgact cactataggg gaattgtgag 1860
 cggataacaa ttccccatct tagtatatta gttaagtata agaaggagat atacatatgc 1920
 gccattaca tccgattgat tttatattcc tgcactaga aaaaagacaa cagcctatgc 1980
 atgtagtggtg tttatcttgg tttcagattc ctgataacgc cccagacacc tttattcaag 2040
 atctggtgaa tgatatccgg atatcaaaat caatccctgt tccaccattc aacaataaac 2100

 tgaatgggct tttttgggat gaagatgaag agtttgattt agatcatcat tttcgtcata 2160
 ttgcaactgc tcatcctggt cgtattcgtg aattgcttat ttatatttca caagagcaca 2220
 gtacgctgct agatcgggca aagcccttgt ggacctgcaa tattattgaa ggaattgaag 2280
 gcaatcgttt tgccatgtac ttcaaaattc accatgcgat ggtcgtatggc gttgctggtg 2340
 tgcggttaat tgaaaaatca ctctccatg atgtaacaga aaaaagtatc gtgccacctt 2400
 ggtgtgttga gggaaaacgt gcaaagcgt taagagaacc taaaacaggt aaaattaaga 2460
 aatcatgtc tggatattaag agtcagcttc aggcgacacc cacagtcatt caagagcttt 2520

 ctcagacagt atttaaagat attggacgta atcctgatca tgtttcaagc tttcaggcgc 2580
 cttgttctat tttgaatcag cgtgtgagct catcgcgacg ttttcagca cagtcttttg 2640
 acctagatcg ttttcgtaat attgccaaat cgttgaatgt gaccattaat gatgtgttac 2700
 tagcggtatg ttctggtgca ttacgtgctg atttgatgag tcataaatagt ttgccttcaa 2760
 aaccattaat tgccatggtt ccagcctcta ttcgcaatga cgattcagat gtcagcaacc 2820
 gtattacgat gattctggca aatttgcaa cccacaaaga tgatccttta caacgtcttg 2880
 aaattatccg ccgtagtggt caaaactcaa agcaacgctt caaacgtatg accagcgatc 2940

 agattctaaa ttatagtgtc gtcgtatatg gccctgcagg actcaacata atttctggca 3000
 tgatgccaaa acgccaagcc ttcaatctgg ttatttccaa tgtgcctggc ccaagagagc 3060
 cactttactg gaatggtgcc aaactgatg cactctacc agcttcaatt gtattagacg 3120
 gtcaagcatt gaatattaca atgaccagtt atttagataa acttgaagtt ggtttgattg 3180
 catgccgtaa tgcatgcca agaatgcaga atttactgac acatttagaa gaagaaattc 3240
 aactatttga aggcgtaatt gcaaagcagg aagatattaa aacagccaat taaaaacaat 3300

aaacttgatt ttttaattta tcagataaaa ctaaagggtt aaattagccc tcttaggctg 3360

 ctgccaccgc tgagcaataa ctagcataac cccttggggc ctctaacgg gtcttgaggg 3420
 gttttttgct gaaacctcag gcatttgaga agcacacggt cacactgctt cggtagtca 3480
 ataaaccggt aaaccagcaa tagacataag cggctattha acgaccctgc cctgaaccga 3540
 cgaccgggtc atcgtggccg gatcttgccg ccctcggct tgaacgaatt gttagacatt 3600
 atttgccgac taccttggtg atctcgcctt tcacgtagtg gacaaattct tccaactgat 3660
 ctgcgcgca ggccaagcga tcttcttctt gtccaagata agcctgtcta gcttcaagta 3720
 tgacgggctg atactgggcc ggcaggcgt ccattgcca gtcggcagcg acatccttcg 3780

 gcgcgatfff gccggtfact gcgctgtacc aaatgcggga caacgtaagc actacatttc 3840
 gctcatgcc agcccagtcg ggcggcgagt tccatagcgt taaggtttca tttagcgcct 3900
 caaatagatc ctgttcagga accggatcaa agagtctctc cgccgctgga cctaccaagg 3960
 caacgctatg ttctcttctt tttgtcagca agatagccag atcaatgtcg atcgtggctg 4020
 gctcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc agttcgcgct 4080
 tagctggata acgccacgga atgatgtcgt cgtgcacaac aatggtgact tetacagcgc 4140
 ggagaatctc gctctctcca ggggaagccg aagtttcaa aaggtcgttg atcaaagctc 4200

 gcccggttgt ttcataagc cttacggtca ccgtaaccag caaatcaata tcaactgtgtg 4260
 gcttcaggcc gccatccact gcggagccgt acaaatgtac ggccagcaac gtcggttcga 4320
 gatggcgtc gatgacgcca actacctctg atagttgagt cgatacttcg gcgatcaccg 4380
 ctccctcat actcttctt tttcaatatt attgaagcat ttatcagggt tattgtctca 4440
 tgagcggata catatttgaa tgtatttaga aaaataaaca aatagctagc tcaactcggtc 4500
 gctacgtcc gggcgtgaga ctgcggcggg cgctcgggac acatacaaag ttaccacag 4560
 attccgtgga taagcagggg actaacatgt gaggcaaac agcagggcgg cgccggtggc 4620

 gtttttccat aggtccgcc ctctgccag agttcacata aacagacgct tttccggtgc 4680
 atctgtggga gccgtgaggc tcaacatga atctgacagt acggcgaaa cccgacagga 4740
 cttaaagatc cccaccgttt ccggcgggtc gctccctctt gcgctctctt gttccgacct 4800
 tgccgtttac cggatacctg ttccgcttt ctcccttacg ggaagtgtgg cgctttctca 4860
 tagctcacac actggtatct cggtcgggtg taggtcgttc gctccaagct gggctgtaag 4920
 caagaactcc ccgttcagcc cgactgctgc gccttatccg gtaactgttc acttgagtcc 4980
 aaccggaaa agcacggtaa aacgccactg gcagcagcca ttggttaactg ggagttcgca 5040

gaggatttgt ttagctaac acgcggttgc tcttgaagtg tgcgccaaag tccggctaca 5100
 ctggaaggac agatttgggt gctgtgctct gcgaaagcca gttaccacgg ttaagcagtt 5160
 ccccaactga cttaaccttc gatcaaacca cctccccagg tggttttttc gtttacaggg 5220
 caaaagatta cgcgcagaaa aaaaggatct caagaagatc cttgatctt ttctactgaa 5280
 ccgctctaga tttcagtca atttatctct tcaaatgtag cacctgaagt cagccccata 5340
 cgataaagt tgtaattctc atgttagtca tgcgccgcgc ccaccggaag gagctgactg 5400
 ggttgaaggc tctcaagggc atcggtcgag atcccgggtgc ctaatgagtg agctaactta 5460

cattaattgc gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc 5520
 attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggccc cagggtgggt 5580
 tttcttttca ccagtgagac gggcaacagc tgattgcctc tcaccgctg gcctgagag 5640
 agttgcagca agcgggtccac gctggtttgc cccagcaggc gaaaatcctg tttgatggtg 5700
 gttaacggcg ggatataaca tgagctgtct tcggtatcgt cgtatcccac taccgagatg 5760
 tccgcaccaa cgcgcagccc ggactcggta atggcgcgca ttgcgccag cgccatctga 5820
 tcgttggcaa ccagcatcgc agtgggaacg atgccctcat tcagcatttg catggtttgt 5880

tgaaaaccgg acatggcact ccagtcgcct tcccgttccg ctatcggctg aatttgattg 5940
 cgagtgagat atttatgcca gccagccaga cgcagacgcg ccgagacaga acttaatggg 6000
 cccgctaaca gcgcgatttg ctggtgacct aatgcgacca gatgctccac gccagtcgc 6060
 gtaccgtctt catgggagaa aataatactg ttgatgggtg tctggtcaga gacatcaaga 6120
 aataacgccg gaacattagt gcaggcagct tccacagcaa tggcatcctg gtcattccagc 6180
 ggatagttaa tgatcagccc actgacgcgt tgcgcgagaa gattgtgcac cgccgcttta 6240
 caggcttcca cgccgcttcg ttctaccatc gacaccacca cgctggcacc cagttgatcg 6300

gcgcgagatt taatgcgcgc gacaatttgc gacggcgcgt gcagggccag actggaggtg 6360
 gcaacccaa tcagcaacga ctgtttgcc gccagttgtt gtgccacgcg gttgggaatg 6420
 taattcagct ccgcatcgc cgcttccact ttttcccgcg ttttcgaga aacgtggctg 6480
 gcctggttca ccacgcggga aacggtctga taagagacac cggcatactc tgcgacatcg 6540
 tataacgtta ctggtttcac attcaccacc ctgaattgac tctcttccgg gcgctatcat 6600
 gccataccgc gaaaggtttt gcgccattcg atggtgtccg ggatctcgac gctctccctt 6660
 atgcgactcc tgcattagga aattaatacg actcactata 6700

<210> 9

<211> 60

<212> DNA

<213> Artificial Sequence
 <220><221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 primer"
 <220><221> source
 <223> /note="Description of Combined DNA/RNA Molecule: Synthetic
 primer"
 <400> 9
 ctctagaat aatttaactt taagtaggag auaggtaccc atggcggaca cgttattgat 60
 <210> 10
 <211> 58
 <212> DNA
 <213> Artificial Sequence
 <220><221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 primer"
 <400> 10

 cttcgaattc catttaaatt atttctagag tcattatgag tcatgattta ctaaaggc 58
 <210> 11
 <211> 65
 <212> DNA
 <213> Artificial Sequence
 <220><221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 primer"
 <400> 11
 ctctagaat aattttagtt aagtataaga aggagatata ccatggtgaa gaaggtttgg 60
 cttaa 65
 <210> 12
 <211> 51
 <212> DNA
 <213> Artificial Sequence
 <220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 12

cttcgaattc catttaaatt atttctagag ttatcaggct ttattgtcca c 51

<210> 13

<211> 5733

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 13

actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg 60

ctgccataac catgagtgat aacctgagg ccaacttact tctgacaacg atcggaggac 120

cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcggt 180

gggaaccgga gctgaatgaa gccatacca acgacgagcg tgacaccacg atgcctgcag 240

caatggcaac aacgttgcgc aaactattaa ctggcgaact acttactcta gettccggc 300

aacaattaat agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc 360

ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta 420

tcattgcagc actggggcca gatggtaagc cctcccgtat cgtagttatc tacacgacgg 480

ggagttaggc aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga 540

ttaagcattg gtaactgtca gaccaagttt actcatatat actttagatt gattttaaac 600

ttcattttta atttaaaagg atctaggtga agatcctttt tgataatctc atgaccaaaa 660

tccttaacg tgagttttcg ttccactgag cgtcagacc ctttaataaga tgatcttctt 720

gagatcgttt tggctctgcgc gtaatctctt gctctgaaaa cgaaaaaac gccttcagg 780

gcggtttttc gaaggttctc tgagctacca actctttgaa ccgaggtaac tggcttggag 840

gagcgcagtc accaaaactt gtcctttcag tttagcctta accggcgcac gacttcaaga 900

ctaactcctc taaatcaatt accagtggct gctgccagtg gtgcttttgc atgtctttcc 960

gggttggact caagacgata gttaccgat aaggcgcagc ggtcggactg aacggggggt 1020

tcgtgcatac agtccagctt ggagcgaact gcctaccgg aactgagtgt caggcgtgga 1080

atgagacaaa cgcggccata acagcggat gacaccgta aaccgaaagg caggaacagg 1140

agagcgcacg agggagccgc cagggggaaa gccttggtat ctttatagtc ctgtcgggtt 1200
 tcgccaccac tgatttgagc gtcagatttc gtgatgcttg tcaggggggc ggagcctatg 1260
 gaaaaacggc ttgcccggg cccctctact tccctgttaa gtatcttcct ggcaccttcc 1320
 aggaaatctc cgccccgttc gtaagccatt tccgctcgcc gcagtcgaac gaccgagcgt 1380
 agcagtcag tgagcgagga agcggaatat atcctgtatc acatattctg ctgacgcacc 1440

 ggtgcagcct tttttctct gccacatgaa gcacttcaact gacaccctca tcagtgccaa 1500
 catagtaagc cagtatacac tccgctagcg ctgaggctcg cctcgtgaag aagggtgtgc 1560
 tgactcatac caggcctgaa tcgccccatc atccagccag aaagtgaggg agccacggtt 1620
 gatgagagct ttgtttagg tggaccagt ggtgattttg aacttttctg ttgccacgga 1680
 acggtctcgc ttgtcgggaa gatgcgtgat ctgaccttc aactcagcaa aagttcgatt 1740
 tattcaaca agccacgttg tgtctcaaaa tctctgatgt tacattgcac aagataaaaa 1800
 tatatcatca tgaacaataa aactgtctgc ttacataaac agtaatacaa ggggtgttat 1860

 gagccatatt caacgggaaa cgtcttgctc gaggcccgca ttaaattcca acatggatgc 1920
 tgatttatat ggggtataaat gggctcgcg taatgtcggg caatcaggtg cgacaatcta 1980
 tcgattgtat ggggaagccc atgcgccaga gttgtttctg aaacatggca aaggtagcgt 2040
 tgccaatgat gttacagatg agatggtcag actaaactgg ctgacggaat ttatgcctct 2100
 tccgaccate aagcatttta tccgtactcc tgatgatgca tggttactca ccaactcgat 2160
 ccccgggaaa acagcattcc aggtattaga agaatatcct gattcaggtg aaaatattgt 2220
 tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt cctgtttgta attgtccttt 2280

 taacagcgat cgcgtatttc gtctcgctca ggcgcaatca cgaatgaata acggtttggt 2340
 tgatgcgagt gattttgatg acgagcgtaa tggctggcct gttgaacaag tctggaaga 2400
 aatgcataag cttttgcat tctcaccgga ttcagtcgtc actcatggtg atttctcaact 2460
 tgataacctt atttttgacg aggggaaatt aataggttgt attgatgtg gacgagtcgg 2520
 aatcgcagac cgataccagg atcttgccat cctatggaac tgcctcggtg agttttctcc 2580
 ttcattacag aaacggcttt ttcaaaaata tggattgat aatcctgata tgaataaatt 2640
 gcagtttcat ttgatgctcg atgagttttt ctaatcagaa ttggttaatt ggttgaaca 2700

 ctggcagagc attacgctga cttgacggga cggcggtttt gttgaataaa tcgaactttt 2760
 gctgagttga aggatcagat cacgcatctt cccgacaacg cagaccgttc cgtggcaaag 2820
 caaaagtcca aaatcaccaa ctggccacc tacaacaag ctctcatcaa ccgtggtctc 2880
 ctcactttct ggetggatga tggggcgatt caggcctggt atgagtcagc aacaccttct 2940
 tcacgaggca gacctcagc ctcaaatg caggggtaaa agctaaccgc atctttaccg 3000

acaaggcatc cggcagttca acagatcggg aagggtgga tttgctgagg atgaaggtgg 3060
 aggaaggtga tgcattctg gtgaagaagc tcgaccgtct tggcccgac accgccgaca 3120

tgatccaact gataaaagag tttgatgctc aggggtgtagc ggttcggttt attgacgacg 3180
 ggatcagtac cgacggtgat atggggcaaa tggtggtcac catcctgtcg gctgtggcac 3240
 aggctgaacg ccggaggatc ctagagcgca cgaatgaggg ccgacaggaa gcaaagctga 3300
 aaggaatcaa atttggccgc aggcgtaccg tggacaggaa cgtcgtgctg acgcttcac 3360
 agaagggcac tggtgcaacg gaaattgctc atcagctcag tattgcccgc tccacggttt 3420
 ataaaattct tgaagacgaa agggcctcgt gatacgcta tttttatagg ttaatgtcat 3480
 gataataatg gtttcttaga cgtcttaatt aatcaggaga gcgttcaccg acaacaaca 3540

gataaaacga aaggcccagt ctttcgactg agcctttcgt tttatttgat gcctggcagt 3600
 tcctactct cgcatgggga gacccacac taccatcggc gctacggcgt ttcacttctg 3660
 agttcggcat ggggtcaggt gggaccaccg cgctactgcc gccaggcaaa tctgtttta 3720
 tcagaccgct tctgcgttct gatttaatct gtatcaggct gaaaatcttc tctcatccgc 3780
 caaaacagcc aagctggaga ccgtttaaac tcaatgatga tgatgatgat ggtcgacggc 3840
 gctattcaga tcctcttctg agatgagttt ttgttcgggc ccaagcttcg aattcccata 3900
 tggtagcage tgcagatctc gagctcggat ccatggttta ttctctcta ttaatcgat 3960

acattaatat atacctcttt aatttttaat aataaagtta atcgataatt ccggtcgagt 4020
 gcccacagag attgtctgat aaattgttaa agagcagtg cgttcgctt tttctcagcg 4080
 gcgctgttcc ctgtgtgaaa ttgttatecg ctcaaatc cacacattat acgagccgga 4140
 tgattaattg tcaacagctc atttcagaat atttgccaga accgttatga tgcggcgca 4200
 aaaaacatta tccagaacgg gagtgcgctc tgagcgacac gaattatgca gtgatttacg 4260
 acctgcacag ccataccaca gcttcgatg gctgcctgac gccagaagca ttggtgcacc 4320
 gtgcagtcga tgataagctg tcaaaccaga tcaattcgcg ctaactcaca ttaattgcgt 4380

tgcgctcact gcccgtttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg 4440
 gccaacgcgc ggggagaggc ggtttcgctc ttggcgcca ggggggtttt tctttcacc 4500
 agtgagacgg gcaacagctg attgccttc accgcctggc cctgagagag ttgcagcaag 4560
 cggctccagc tggtttgecc cagcaggcga aatcctggt tgatgggtgt tgacggcggg 4620
 atataacatg agctgtcttc ggtatcgtcg tatcccacta ccgagatc cgcaccaacg 4680
 cgagcccgg actcggtaat ggcgcgcat gcgccagcg ccatctgatc gttggcaacc 4740

agcatcgagcag tgggaacgat gccctcattc agcatttgca tggtttgttg aaaaccggac 4800

atggcactcc agtcgccttc ccgttccgct atcggctgaa tttgattgcg agtgagatat 4860

ttatgccage cagccagacg cagacgcgcc gagacagAAC ttaatgggcc cgctaacagc 4920

gcgatttgct ggtgacccaa tgcgaccaga tgcctcacgc ccagtcgcgt accgtcttca 4980

tgggagaaaa taatactggt gatgggtgtc tggtcagaga catcaagaaa taacgccgga 5040

acattagtgc aggcagcttc cacagcaatg gcatcctggt catccagcgg atagttaatg 5100

atcagcccac tgacgcgttg cgcgagaaga ttgtgcaccg ccgctttaca ggcttcgacg 5160

ccgcttcggt ctaccatcga caccaccacg ctggcaccca gttgatcggc gcgagattta 5220

atcgccgcga caatttgca cggcgcgtgc agggccagac tggaggtggc aacgccaatc 5280

agcaacgact gtttgcccgc cagttgttgt gccacgcggt tgggaatgta attcagctcc 5340

gccatcgccg ctccaacttt ttcccgcgtt ttccgagaaa cgtggctggc ctggttcacc 5400

acgcgggaaa cggctctgata agagacaccg gcatactctg cgacatcgta taacgttact 5460

ggtttcacat tcaccacct gaattgactc tcttccgggc gctatcatgc cataccgca 5520

aaggttttgc accattcgat ggtgtcaacg taaatgcatg ccgcttcgcc ttcgcgcgcg 5580

aattgatctg ctgcctcgcg cgtttcgggtg atgacggtga aaacctctga cacatgcagc 5640

tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg 5700

gcgcgtcagc ggggtgttggc ggggccggcc tcg 5733

<210> 14

<211> 36

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 14

gggtcaatag cggccgcaaa ttcgcgcgcg aaggcg 36

<210> 15

<211> 37

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 15

tggcgcgcct cctagggcat tacgctgact tgacggg 37

<210> 16

<211> 15179

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 16

ggccacgatg cgtccggcgt agaggatctg ctcatgtttg acagcttacc atcgatgcat 60

aatgtgcctg tcaaatggac gaagcaggga ttctgcaaac cctatgctac tccgtcaagc 120

cgtaattgt ctgattcgtt accaattatg acaacttgac ggctacatca ttcacttttt 180

cttcacaacc ggcacggaac tcgctcgggc tggccccggt gcatttttta aatacccgcg 240

agaaatagag ttgatgctca aaaccaacat tgcgaccgac ggtggcgata ggcatccggg 300

tgggtgctcaa aagcagcttc gcctggctga tacgttggtc ctgcgccag ctaagacgc 360

taatccctaa ctgctggcgg aaaagatgtg acagacgcga cggcgacaag caaacatgct 420

gtgcgacgct ggcgatatca aaattgctgt ctgccagggt atcgctgatg tactgacaag 480

cctcgcgtac ccgattatcc atcggatgat ggagcgactc gttaatcgct tccatgcgcc 540

gcagtaacaa ttgctcaagc agatttatcg ccagcagctc cgaatagcgc ctttcccctt 600

gccccgcgtt aatgatttgc ccaaacaggc cgtgaaatg cggctggtgc gttcatccg 660

ggcgaaagaa ccccgtattg gcaaatattg acggccagtt aagccattca tgccagtagg 720

cgcgccgacg aaagtaaacc cactggtgat accattcgcg agcctccgga tgaccacgt 780

agtgatgaat ctctcctggc gggaacagca aaatcacc cggcggcaa acaaattctc 840

gtccctgatt tttcaccacc cctgaccgc gaatggtgag attgagaata taacctttca 900

ttcccagcgg tcggtcgata aaaaaatcga gataaccggt ggctcaatc ggcgttaaac 960

ccgccaccag atgggcatta aacgagtatc ccggcagcag gggatcattt tgcgcttcag 1020

ccatactttt catactcccg ccattcagag aagaaaccaa ttgtccatat tgcacagac 1080

attgccgtca ctgcgtcttt tactggctct tctcgtaac caaacggta accccgctta 1140

ttaaaagcat tctgtaacaa agcgggacca aagccatgac aaaaacgcgt aacaaaagtg 1200
 tctataatca cggcagaaaa gtccacattg attatttgca cggcgtcaca ctttgctatg 1260
 ccatagcatt tttatccata agattagcgg atcctacctg acgcttttta tcgcaactct 1320
 ctactgtttc tccatacccg ttttttggg ctagegaatt cgagctcggg acccaagtct 1380
 taaactagac agaatagttg taaactgaaa tcagtccagt tatgctgtga aaaagcatac 1440

 tggacttttg ttatggctaa agcaaaactct tcattttctg aagtgcaaat tgcccgtcgt 1500
 attaaagagg ggcgtggcca agggcatggt aaagactata ttccatggct aacagtacia 1560
 gaagttcctt cttcaggtcg ttcccaccgt atttattctc ataagacggg acgagtcctat 1620
 catttgctat ctgacttaga gcttgctggt tttctcagtc ttgagtggga gagcagcgtg 1680
 ctagatatac gcgagcagtt ccccttatta cctagtata ccaggcagat tgcaatagat 1740
 agtggatta agcatcctgt tattcgtggt gtagatcagg ttatgtctac tgatttttta 1800
 gtggactgca aagatgggcc ttttgagcag tttgctattc aagtcaaacc tgcagcagcc 1860

 ttacaagacg agcgtacctt agaaaaacta gaactagagc gtcgctattg gcagcaaaag 1920
 caaatcctt gggtcatttt tactgataaa gaaataaatc ccgtagtaaa agaaaatatt 1980
 gaatggcttt attcagtgaa aacagaagaa gtttctgcgg agcttttagc acaactatcc 2040
 ccattggccc atatcctgca agaaaaagga gatgaaaaca ttatcaatgt ctgtaagcag 2100
 gttgatattg cttatgattt ggagtttagc aaaacattga gtgagatacg agccttaacc 2160
 gcaaatggtt ttattaagtt caatatttat aagtctttca gggcaataa gtgtgcagat 2220
 ctctgtatta gccaaagtagt gaatatggag gaggttgcgt atgtggcaaa ttaatgaggt 2280

 tgtgctattt gataatgac cgtatcgcat tttggctata gaggatggcc aagttgtctg 2340
 gatgcaata agcgtgata aaggagttcc acaagctagg gctgagttgt tgctaatgca 2400
 gtatttagat gaagccgct tagttagaac tgatgacct tatgtacac ttgatttaga 2460
 agagccgtct gtagattctg tcagcttcca gaagcgcgag gaggattatc gaaaaattct 2520
 tcctattatt aatagtaagg atcgtttcga ccctaaagtc agaagcgaac tcgttgagca 2580
 tgtggccaa gaacataagg ttactaagc tacagtttat aagttgttac gccgttactg 2640
 gcagcgtggt caaacgccta atgcattaat tctgactac aaaaacagcg gtgcaccagg 2700

 ggaaagacgt tcagcgacag gaacagcaaa gattggccga gccagagaat atggttaagg 2760
 tgaaggaacc aagtaacgc ccgagattga acgccttttt aggttgacca tagaaaagca 2820
 cctgttaaat caaaaaggta caaagaccac cgttcctat agacgatttg tggacttggt 2880
 tgctcagtat tttcctcgca ttcccaga ggattacca acactacgtc agtttcgtta 2940
 tttttatgat cgagaatacc ctaaagctca gcgcttaaag tctagagtta aagcaggggt 3000

atataaaaa gacgtacgac ccttaagtag tacagccact tctcaggcgt taggccttgg 3060
 gagtcgttat gagattgatg ccacgattgc tgatatttat ttagtggatc atcatgatcg 3120

 ccaaaaaate ataggaagac caacgcttta cattgtgatt gatgtgttta gtcggatgat 3180
 cacgggcttt tatatcggct ttgaaaatcc gtcttatgtg gtggcgatgc aggcttttgt 3240
 aaatgcttgc tctgacaaaa cggccatttg tgcccagcat gatattgaga ttagtagctc 3300
 agactggccg tigttaggtt tgccagatgt gttgctagcg gaccgtggcg aattaatgag 3360
 tcatcaggtc gaagccttag tttctagttt taatgtgcga gtggaaagtg ctccacctag 3420
 acgtggcgat gctaaaagca tagtggaaag cactttttaga aactacaag ccgagttaa 3480
 gtcctttgca cctggcattg tagagggcag tcggatcaaa agccatggtg aaacagacta 3540

 taggttagat gcatctctgt cggtatattga gttcacacaa attattttgc gtacgatctt 3600
 attcagaaat aaccatctgg tgatggataa atacgatcga gatgctgatt ttctacaga 3660
 tttaccgtct attcctgtcc agctatggca atggggtatg cagcatcgtc caggtagttt 3720
 aagggtctgt gagcaagagc agttgcgagt agcgttactg cctcggcga aggtctctat 3780
 ttcttcattt ggcgttaatt tgtggggttt gtattactcg gggtcagaga ttctgcgtga 3840
 gggttggttg cagcggagca ctgatatagc tagacctcaa cattagaag cgcttatga 3900
 cccagtctg gttgatacga tttatttgtt tccgaagtt ggcagccgtg tattttggcg 3960

 ctgtaatctg acggaacgta gtcggcagtt taaaggcttc tcattttggg aggtttggga 4020
 tatacaagca caagaaaaac acaataaagc caatgcgaag caggatgagt taactaacg 4080
 cagggagctt gaggcgttta ttcagcaaac cattcagaaa gcgaataagt taacgccag 4140
 tactactgag cccaaatcaa cacgattaa gcagattaaa actaataaaa aagaagccgt 4200
 gacctcggag cgtaaaaaac gtgcggagca tttgaagcca agctcttcag gtgatgagge 4260
 taaagtattt cctttcaacg cagtggaagc ggatgatcaa gaagattaca gectaccac 4320
 atacgtgcct gaattatttc aggatccacc agaaaaggat gagtcatgag tgctaccgg 4380

 attcaagcag tttatcgtga tacgggggta gagccttate gtgataatcc ttttatcgag 4440
 gccttaccac cattacaaga gtcagtgaat agtgctgcat cactgaaatc ctctttacag 4500
 ctacttctct ctgacttgca aaagtcccgt gttatcagag ctcatacat ttgtcgtatt 4560
 ccagatgact attttcagcc attaggtacg catttgctac taagtgagcg tatttcggtc 4620
 atgattcgag gtggctacgt aggcagaaat cctaaaacag gagatttaca aaagcattta 4680
 caaatggtt atgagcgtgt tcaaacggga gagttggaga catttcgctt tgaggaggca 4740

cgatctacgg cacaaagctt attgtaatt ggttgttctg gtagtgggaa gacgacctct 4800

cttcacgta ttctagccac gtatcctcag gtgatttacc atcgtgaact caatgtagag 4860

caggtgggtg atttgaaaat agactgctcg cataatggtt cgctaaaaga aatctgcttg 4920

aatTTTTca gagcgttggg tgcagccttg ggctcgaact atgagcgtcg ttatggctta 4980

aaacgcatg gtatagaac catgttggct ttgatgtcgc aaatagccaa tgcacatgct 5040

ttagggttgt tggttattga tgaattcag catttaagcc gctctcgttc ggggtgatct 5100

caagagatgc tgaacttttt tgtgacgatg gtgaatatta ttggcgtacc agtgatgttg 5160

attggtacce ctaaagcacg agagattttt gaggctgatt tgcggtctgc acgtagaggg 5220

gcagggtttg gagctatatt ctgggatcct atacaacaaa cgcaacgtgg aaagcccaat 5280

caagagtgga tcgcttttac ggataatctt tggcaattac agcttttaca acgcaaagat 5340

gcgctgttat cggatgaggt ccgtgatgtg tggatgagc taagccaagg agtgatggac 5400

attgtagtaa aactttttgt actcgtcag ctccgtcgc tagctttagg caatgagcgt 5460

attaccgctg gtttattgcg gcaagtgtat caagatgagt taaagcctgt gcaccccatg 5520

ctagaggcat tacgctcggg tatcccagaa cgcattgctc gttattctga tetagtcgtt 5580

cccgagattg ataacgggtt aatccaactt cagctagata tcgcagcgat acaagaacaa 5640

acaccagaag aaaaagcctt tcaagagtta gataccgaag atcagcgtca tttatatctg 5700

atgctgaaag aggattacga ttcaagcctg ttaattccca ctattaaaa agcgttttagc 5760

cagaatcaa cgatgacaag acaaaagtta ctgcctcttg ttttcagtg gttgatggaa 5820

ggcgaacgg tagtgtcaga actagaaaag cctccaaga gtaaaaaggt ttcggctata 5880

aaggtagtca agcccagcga ctgggatagc ttgcctgata cggatttacg ttatatctat 5940

tcacaacgcc aacctgaaaa aaccatgcat gaacggttaa aagggaagg gtaaatagtg 6000

gatatggcga gcttatttaa acaagcaggt tagccatgag aaactttcct gttccgtact 6060

cgaatgagct gatttatagc actattgcac gggcagcgt ttatcaaggg attgtagtc 6120

ctaagcagct gttggatgag gtgtatggca accgcaagg ggtcgtacc ttaggtctgc 6180

cctcgcattt aggtgtgata gcaagacatc tacatcaaac aggacgttac gctgttcagc 6240

agcttattta ttagcatacc ttattccctt tatatgctcc gttttaggc aaggagcgc 6300

gagacgaagc tattcggtta atggagtacc aagcgaagg tgcggtgcat ttaatgctag 6360

gagtcgctgc ttctagagtt aagagcgata accgcttag atactgcct gattgcgttg 6420

ctcttcagct aaataggtat ggggaagcct ttggcaacg agattggtat ttgcccgctt 6480

tgccatattg tccaaaacac ggtgcttttag tcttctttga tagagctgta gatgatcacc 6540
 gacatcaatt ttgggctttg ggtcatactg agctgctttc agactacccc aaagactccc 6600
 tatctcaatt aacagcacta gctgcttata tagcccctct gttagatgct ccacgagcgc 6660
 aagagctttc cccaagcctt gagcagtgga cgctgtttta tcagcgctta gcgcaggatc 6720
 tagggctaac caaaagcaag cacattcgtc atgacttggg ggcggagaga gtgaggcaga 6780
 ctttttagtga tgaggcacta gagaaactgg atttaaagtt ggcagagAAC aaggacacgt 6840
 gttggctgaa aagtatattc cgtaagcata gaaaagcctt tagttattta cagcatagta 6900

 ttgtgtggca agccttattg ccaaaactaa cggttataga agcgtacag caggcaagtg 6960
 ctcttactga gcactctata acgacaagac ctgttagcca gtctgtgcaa cctaactctg 7020
 aagatttattc tgtaagcat aaagactggc agcaactagt gcataaatac caaggaatta 7080
 aggcggcaag acagctctta gaggggtgggg tgcataacgc ttggctttac cgacatgaca 7140
 gggattggct agttcactgg aatcaacagc atcaacaaga gcgtctggca cccgcccta 7200
 gagttgattg gaaccaaaga gatcgaattg ctgtacgaca actattaaga atcataaagc 7260
 gtctagatag tagccttgat cacccaagag cgacatcgag ctggctgtta aagcaaactc 7320

 ctaacggaac ctctcttga aaaaatctac agaaactgcc tttgtagcg ctttgcttaa 7380
 agcgttactc agagagtgtg gaagattatc aaattagacg gattagccaa gcttttatta 7440
 agcttaaaca ggaagatgtt gagcttaggc gctggcgatt attaagaagt gcaacgttat 7500
 ctaaagagcg gataactgag gaagcacaaa gattcttggg aatggtttat ggggaagagt 7560
 gagtggttag gctagctaca tttaatgaca atgtgcaggt tgtacatatt ggtcatttat 7620
 tccgtaactc gggtcataag gagtggcgta tttttgttg gtttaacca atgcaagaac 7680
 ggaaatggac tcgatttact catttgcctt tattaagtcg agctaaggtg gttaacagta 7740

 caacaaagca aataaataag gcggatcgtg tgattgagtt tgaagcatcg gatcttcaac 7800
 gagccaaaat aatcgatttt cctaactctc cgtccittgc ttccgtacgc aacaaggatg 7860
 gagcgcagag ttcatattat tacgaagctg aaacaccata tagcaagact cgttatcaca 7920
 tcccacagtt agagctagct cggtcattat ttttagggga tcctctagag tcgacctgca 7980
 ggcatgcaag ctiggctgtt ttggcggatg agagaagatt ttcagcctga tacagattaa 8040
 atcagaacgc agaagcggtc tgataaaaca gaatttgctt ggcggcagta gcgcggtggt 8100
 cccacctgac cccatgccga actcagaagt gaaacgccgt agcgcgatg gtagtgtggg 8160

 gtctcccat gcgagagtag ggaactgcca ggcacaaat aaaacgaaag gctcagtcga 8220
 aagactgggc ctttcgtttt atctgttgtt tgcggtgaa cgctctcctg agtaggacaa 8280
 atccgccggg agcggatttg aacgttgca agcaacggcc cggagggtgg cgggcaggac 8340

gcccccata aactgccagg catcaaatta agcagaaggc catcctgacg gatggccttt 8400
 ttgcgtttct acaaactctt ttgtttatit ttctaaatac attcaaatat gtatccgctc 8460
 atgagacaat aaccctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt 8520
 caacatttcc gtgtcgcctt tattcccttt ttgcgccat ttgecttcc tgtttttgct 8580

caccagaaa cgctggtgaa agtaaaagat gctgaagatc agttgggtgc acgagtgggt 8640
 tacatcgaac tggatctcaa cagcggtaag atccttgaga gttttcgccc cgaagaacgt 8700
 ttccaatga tgagcacttt taaagtcttg ctatgtggcg cggattatc ccgtgttgac 8760
 gccggcaag agcaactcgg tcgccgata cactattctc agaatgactt ggttgagtac 8820
 tcaccagtca cagaaaagca tcttacggat ggcatgacag taagagaatt atgcagtgct 8880
 gccataacca tgagtataa cactgcggcc aacttacttc tgacaacgat cggaggaccg 8940
 aaggagctaa ccgctttttt gcacaacatg ggggatcatg taactgcctt tgatcgttgg 9000

gaaccggagc tgaatgaagc cataccaac gacgagcgtg acaccacgat gcctgcagca 9060
 atggcaacaa cgttgcgcaa actattaact ggcgaaactac ttactctage ttcccggcaa 9120
 caattaatag actggatgga ggcggataaa gttgcaggac cacttctgcg ctcgccctt 9180
 ccggctggct ggtttatgct tgataaatct ggagccgggtg agcgtgggtc tcgcggtatc 9240
 attgcagcac tggggccaga tggtaagccc tcccgtatcg tagttatcta cacgacgggg 9300
 agtcaggcaa ctatggatga acgaaataga cagatcgtg agataggtgc ctcaactgatt 9360
 aagcattggt aactgtcaga ccaagtttac tcatatatac tttagattga ttacgcgcc 9420

ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga ccgctacact 9480
 tgccagcgcc ctagcgcgcc ctctttcgc tttcttccct tcctttctcg ccacgttcgc 9540
 cgccggccag cctgcagag caggattccc gttgagcacc gccaggtgcg aataaggac 9600
 agtgaagaag gaacaccgc tcgcggtggt gcctacttca cctatcctgc ccggcggcat 9660
 caccggcgc acaggtcggc ttgctggcgc ctatatgcc gacatcaccg atggggaaga 9720
 tcgggctcgc cacttcgggc tcatgagcgc ttgtttcggc gtgggtatgg tggcaggccc 9780
 cgtggccggg ggactgttgg gcgccatctc cttgcatgca ccattccttg cggcggcgg 9840

gctcaacggc ctcaacctac tactgggctg cttcctaag caggagtcgc ataagggaga 9900
 gcgtcgatec ccgacagtaa gacgggtaag cctgttgatg ataccgctgc ctactgggt 9960
 gcattagcca gtctgaatga cctgtcacgg gataatccga agtggtcaga ctggaatc 10020
 agagggcagg aactgctgaa cagcaaaaag tcagatagca ccacatagca gaccgccat 10080
 aaaacccctt gagaagcccc tgacgggctt ttcttatt atgggtagtt tcttgcgatg 10140
 aatccataaa aggcgcctgt agtgccattt accccattc actgccagag ccgtgagcgc 10200

agcgaactga atgtcacgaa aaagacagcg actcaggtgc ctgatggtcg gagacaaaag 10260

 gaatattcag cgatttgccc gagcttgcga gggtgctact taagccttta gggttttaag 10320
 gtctgttttg tagaggagca aacagcgttt gcgacatcct tttgtaatac tgcggaactg 10380
 actaaagtag tgagttatac acagggctgg gatctattct ttttatcttt ttttattctt 10440
 tctttattct ataaattata accacttgaa tataaacaaa aaaaacacac aaaggtctag 10500
 cggaatttac agagggtcta gcagaattta caagttttcc agcaaaggtc tagcagaatt 10560
 tacagatacc cacaactcaa aggaaaagga ctagtaatta tcattgacta gcccatctca 10620
 attggtatag tgattaaaaat cacctagacc aattgagatg tatgtctgaa ttagttgttt 10680

 tcaaagcaaa tgaactagcg attagtcgct atgacttaac ggagcatgaa accaagctaa 10740
 ttttatgctg tgtggcacta ctcaaccca cgattgaaaa ccctacaagg aaagaacgga 10800
 cggtatcggt cacttataac caatacgttc agatgatgaa catcagtagg gaaaatgctt 10860
 atgggtgatt agctaaagca accagagagc tgatgacgag aactgtgaa atcaggaatc 10920
 ctttggttaa aggctttgag attttccagt ggacaaacta tgccaagttc tcaagcgaaa 10980
 aattagaatt agtttttagt gaagagatat tgccttatct tttccagtta aaaaaattca 11040
 taaaataaa tctggaacat gttaagtctt ttgaaaacaa atactctatg aggatttatg 11100

 agtggttatt aaaagaacta acacaaaaga aaactcaca ggcaaatata gagattagcc 11160
 ttgatgaatt taagttcatg ttaatgcttg aaaataacta ccatgagttt aaaaggctta 11220
 accaatgggt tttgaaacca ataagtaaag atttaaacac ttacagcaat atgaaattgg 11280
 tggttgataa gcgaggccgc ccgactgata cgttgatfff ccaagttgaa ctagatagac 11340
 aatggatct cgtaaccgaa cttgagaaca accagataaa aatgaatggt gacaaaatac 11400
 caacaacat tacatcagat tcttacctac ataacggact aagaaaaaca ctacacgatg 11460
 ctttaactgc aaaaattcag ctcaccagtt ttgaggcaaa atttttgagt gacatgcaaa 11520

 gtaagatga tctcaatggt tcgttctcat ggctcacgca aaaacaacga accacactag 11580
 agaacatact ggctaaatac ggaaggatct gaggttctta tggctcttgt atctatcagt 11640
 gaagcatcaa gactaacaaa caaaagtaga acaactgttc accgttacat atcaaagga 11700
 aaactgtcca tatgcacaga tgaaaacggt gtaaaaaaga tagatacatc agagctttta 11760
 cgagtttttg gtgcatttaa agctgttcac catgaacaga tcgacaatgt aacagatgaa 11820
 cagcatgtaa cacctaagat aacagtgaa accagtataa caaagcaact agaacatgaa 11880
 attgaacacc ttagacaact tgttacagct caacagtcac acatagacag cctgaaacag 11940

gcgatgctgc ttatcgaatc aaagctgccg acaacacggg agccagtgac gcctcccgtg 12000
 gggaaaaaat catggcaatt ctggaagaaa tagcgctttc agcctgtggg cggacaaaat 12060
 agttgggaac tgggaggggt ggaaatggag tttttaagga ttatttaggg aagagtgaca 12120
 aaatagatgg gaactgggtg tagcgtcgta agctaatacg aaaattaaaa atgacaaaat 12180
 agtttggaac tagatttcac ttatctggtt ggtcgacact agtattacc tttatccct 12240
 agatttaaat gatatcggat cctagtaagc cacgttttaa ttaatcagat gggtaaatag 12300
 cggccgcca ttcgcgcgcg aaggcgaagc ggcatgcatt tacgttgaca ccatcgaatg 12360

 gtgcaaaacc tttcgcggtg tggcatgata gcgcccggaa gagagtcaat tcagggtggt 12420
 gaatgtgaaa ccagtaacgt tatacagatg cgagagatg gccggtgtct cttatcagac 12480
 cgtttcccgc gtggtgaacc aggccagcca cgtttctgcg aaaacgcggg aaaaagtgga 12540
 agcggcgatg gcggagctga attacattcc caaccgctg gcacaacaac tggcgggcaa 12600
 acagtcttg ctgattggcg ttgccacctc cagtctggcc ctgcacgcg cgtcgcaaat 12660
 tgtcgcggcg attaaatctc gcgccgatca actgggtgcc agcgtggtgg tgtcgatggt 12720
 agaacgaagc ggcgtcgaag cctgtaaacg ggcggtgcac aatcttctcg cgcaacgct 12780

 cagtgggctg atcattaact atccgctgga tgaccaggat gccattgctg tggaaactgc 12840
 ctgcactaat gttccggcgt tatttcttga tgtctctgac cagacacca tcaacagtat 12900
 tattttctcc catgaagacg gtacgcgact gggcgtggag catctggtcg cattgggtca 12960
 ccagcaaatc gcgctgttag cgggcccatt aagtctctgc tcggcgcgctc tgcgtctggc 13020
 tggctggcat aaatatctca ctcgcaatca aattcagccg atagcggaac gggaaaggcga 13080
 ctggagtgcc atgtccggtt ttcaacaaac catgcaaatg ctgaatgagg gcatcgttcc 13140
 cactgcgatg ctggttgcca acgatcagat ggcgctgggc gcaatgcgcg ccattaccga 13200

 gtcccggctg cgcgttgggt cggatatctc ggtagtggga tacgacgata ccgaagacag 13260
 ctcatgttat atcccgcct caaccacat caaacaggat tttcgcctgc tggggcaaac 13320
 cagcgtggac cgtttgctgc aactctctca gggccaggcg gtgaagggca atcagctggt 13380
 gccctctca ctggtgaaaa gaaaaaccac cctggcgccc aatacgcaaa cgcctctcc 13440
 ccgcgcttg gccgattcat taatgcagct ggcacgacag gttcccgcac tggaaagcgg 13500
 gcagtgagcg caacgaatt aatgtgagtt agcgcgaatt gatctggttt gacagcttat 13560
 catcgactgc acggtgcacc aatgcttctg gcgtcaggca gccatcggaa gctgtggtat 13620

 ggctgtgcag gtcgtaaatc actgcataat tcgtgtcgt caaggcgac tcccgttctg 13680
 gataatgttt tttgcgccga catcataacg gttctggcaa atattctgaa atgagctggt 13740
 gacaattaat catccggctc gtataatgtg tggaaattgt agcggataac aatttcacac 13800

aggaaacagc gccgctgaga aaaagcgaag cggcactgct ctttaacaat ttatcagaca 13860
 atctgtgtgg gactcagacc ggaattatcg attaacttta ttattaaaaa ttaaagaggt 13920
 atatattaat gtatcgatta aataaggagg aataaacat ggccgacacg ttattgattc 13980
 tgggtgatag ccigagcgcc gggatcgaat tgcctgccag cgcggcctgg cctgccttgt 14040

tgaatgataa gtggcagagt aaaacgtcgg tagttaatgc cagcatcagc ggcgacacct 14100
 cgcaacaagg actggcgcgc ctcccgctc tgcctgaaaca gcatcagccg cgttgggtgc 14160
 tggttgaact gggcggcaat gacggtttgc gtggttttca gccacagcaa accgagcaaa 14220
 cgctgcgcca gatcttcagc gatgtcaaag ccgccaacgc tgaaccattg ttaatgcaaa 14280
 tacgtctgcc tgcaaaactat ggtcggcgtt ataataaagc ctttagcgcc atttacccea 14340
 aactcgcaa agagtttgat gttccgctgc tgcctttttt tatggaagag gtctacctca 14400
 agccacaatg gatgcaggat gacggtattc atcccaaccg cgacgcccag ccgtttattg 14460

ccgactggat ggcaagcag ttgcagcctt tagtaaatca tgactcataa tgactctaga 14520
 aataatttaa atggaattcg aagcttgggc ccgaacaaaa actcatctca gaagaggatc 14580
 tgaatagcgc cgtcgacct catcatcatc atcattgagt ttaaaccgctc tccagcttgg 14640
 ctgttttggc ggatgagaga agattttcag cctgatacag attaatcag aacgcagaag 14700
 cgtctgata aaacagaatt tgcctggcgg cagtagcgcg gtggtccac ctgaccccat 14760
 gccgaactca gaagtgaac gccgtagcgc cgatggtagt gtgggtctc cccatgcgag 14820
 agtagggaac tgccaggcat caaataaac gaaaggctca gtcgaaagac tgggcctttc 14880

gttttatctg ttgtttgtcg gtgaacgctc tctgattaa ttaagacgtc ccgtaagtc 14940
 agcgtaatgc cctaggaggc gcgccacggc cgcgtcgacc ccacgccct ctttaatacg 15000
 acgggcaatt tgcacttcag aaaatgaaga gtttgcttta gccataaca aagtccagta 15060
 tgctttttca cagcataact ggactgattt cagtttaca ctattctgtc tagtttaaga 15120
 ctttattgtc atagtttaga tctattttgt tcagtttaag actttattgt ccgccaca 15179

<210> 17

<211> 17683

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 17

ggccacgatg cgtccggcgt agaggatctg ctcattgttg acagcttata atcgatgcat 60
 aatgtgcctg tcaaatggac gaagcaggga ttctgcaaac cctatgctac tccgtcaage 120
 cgtcaattgt ctgattcgtt accaattatg acaacttgac ggctacatca ttcacttttt 180
 cttcacaacc ggcacggaac tcgctcgggc tggccccggg gcatttttta aatacccgcg 240
 agaaatagag ttgatcgta aaaccaacat tgcgaccgac ggtggcgata ggcacccggg 300
 tgggtctcaa aagcagcttc gcctggctga tacgttggtc ctgcgccag cttaaagacg 360
 taatccctaa ctgctggcgg aaaagatgtg acagacgca cggcgacaag caaacatgct 420

 gtgcgacgct ggcgatatca aaattgctgt ctgccagggtg atcgtgatg tactgacaag 480
 cctcgcgtac ccgattatcc atcgggtggat ggagcgactc gttaatcgct tccatgcgcc 540
 gcagtaacaa ttgctcaagc agatttatcg ccagcagctc cgaatagcgc ctttccctt 600
 gcccggcgtt aatgatttg ccaaacaggt cgtgaaatg cggctgggtc gcttcatccg 660
 ggcgaaagaa ccccgatttg gcaaatattg acggccagtt aagccattca tgccagtagg 720
 cgcgccgacg aaagtaaac cactgggtgat accattcgcg agcctccgga tgacaccgt 780
 agtgatgaat ctctcctggc gggaacagca aaatatcacc cggtcggcaa acaaattctc 840

 gtccctgatt tttcaccacc ccctgaccgc gaatgggtgag attgagaata taaccttca 900
 ttcccagcgg tcggtcgata aaaaaatcga gataaccgtt ggcctcaatc ggcgttaaac 960
 ccgccaccag atgggcatta aacgagtatc ccggcagcag gggatcattt tgcgcttcag 1020
 ccatactttt catactcccg ccattcagag aagaaaccaa ttgtccatat tgcacagac 1080
 attgccgtea ctgcgtcttt tactggctct tctcgtaac caaacggta accccgctta 1140
 ttaaaagcat tctgtaacaa agcgggacca aagccatgac aaaaacgct aacaaaagtg 1200
 tctataatca cggcagaaaa gtccacattg attatttga cggcgtcaca ctttgctatg 1260

 ccatagcatt tttatecata agattagcgg atcctacctg acgcttttta tcgcaactct 1320
 ctactgtttc tccatacccg ttttttggg ctagcgaatt cgagctcggg acccaagtct 1380
 taaactagac agaatagttg taaactgaaa tcagtccagt tatgctgtga aaaagcatac 1440
 tggacttttg ttatggctaa agcaactct tcattttctg aagtgcaaat tgcccgtcgt 1500
 attaaagagg ggcgtggcca agggcatggt aaagactata ttccatggct aacagtacaa 1560
 gaagtctctt cttcaggtcg ttcccaccgt atttattctc ataagacggg acgagtccat 1620
 catttgctat ctgacttaga gcttctgtt tttctcagtc ttgagtggga gagcagcgtg 1680

 ctagatatac gcgagcagtt ccccttatta cctagtata ccaggcagat tgcaatagat 1740
 agtggtatta agcactcgtt tattcgtggt gtagatcagg ttatgtctac tgatttttta 1800
 gtggactgca aagatggctc ttttgagcag tttgctattc aagtcaaacc tgcagcagcc 1860

ttacaagacg agcgtacctt agaaaaacta gaactagagc gtcgctattg gcagcaaaag 1920
 caaatcctt ggttcatttt tactgataaa gaataaatc ccgtagtaaa agaaaatatt 1980
 gaatggcttt attcagtga aacagaagaa gtttctgcgg agcttttagc acaactatcc 2040
 ccattggccc atatcctgca agaaaaagga gatgaaaaca ttatcaatgt ctgtaagcag 2100

 gttgatattg cttatgattt ggagtttagc aaaacattga gtgagatacg agccttaacc 2160
 gcaaatgggt ttattaagtt caatatttat aagtctttca gggcaataa gtgtgcagat 2220
 ctctgtatta gccaaagtagt gaatatggag gagttgcgct atgtggcaaa ttaatgaggt 2280
 tgtgctatth gataatgac cgtatcgcat tttggctata gaggatggcc aagttgtctg 2340
 gatgcaata agcgtgata aaggagtcc acaagctagg gctgagttgt tgctaagca 2400
 gtatttagat gaaggccgct tagttagaac tgatgacct tatgtacatc ttgatttaga 2460
 agagccgtct gtagattctg tcagcttcca gaagcgcgag gaggattatc gaaaaattct 2520

 tcctattatt aatagtaagg atcgtttcga ccctaaagtc agaagcgaac tcgttgagca 2580
 tgtgtgcca gaacataagg ttactaaggc tacagtttat aagttgttac gccgttactg 2640
 gcagcgtggt caaacgccta atgcattaat tctgactac aaaaacagcg gtgcaccagg 2700
 ggaaagacgt tcagcgacag gaacagcaaa gattggccga gccagagaat atggtaaggg 2760
 tgaaggaacc aagtaacgc ccgagattga acgcttttt aggttgacca tagaaaagca 2820
 cctgttaaat caaaaagta caaagaccac cgttgctat agacgattg tggacttgtt 2880
 tgctcagat tttcctgca ttcccaaga ggattacca aactacgtc agtttcgta 2940

 tttttatgat cgagaatacc ctaaagctca gcgcttaag tctagagta aagcaggggt 3000
 atataaaaa gacgtacgac ccttaagtag tacagccact tctcaggcgt taggccctgg 3060
 gagtcgttat gagattgat ccacgattgc tgatatttat ttagtggatc atcatgatcg 3120
 ccaaaaaate ataggaagac caacgcttta cattgtgatt gatgtgttta gtcggatgat 3180
 cacgggcttt tatatcggt ttgaaaatcc gtcttatgtg gtggcgatgc aggcttttgt 3240
 aatgcttgc tctgacaaaa cggccattg tgcccagcat gatattgaga ttagtagctc 3300
 agactggccg tgtgtaggtt tgccagatgt gttgctagcg gaccgtggcg aattaatgag 3360

 tcatcaggte gaagccttag tttctagttt taatgtcga gtgaaagtg ctccacctag 3420
 acgtggcgat gctaaaggca tagtggaaag cacttttaga aactacaag ccgagtttaa 3480
 gtctttgca cctggcattg tagagggcag tcggatcaaa agccatgggtg aacagacta 3540
 taggttagat gcatctctgt cggtatthga gttcacacaa attattttgc gtacgatctt 3600
 attcagaaat aaccatctgg tgatggataa atacgatcga gatgctgatt ttctacaga 3660
 tttaccgtct attcctgtcc agctatggca atggggtatg cagcatcgta caggtagttt 3720

aagggtctgt gagcaagagc agttgtagt agcgttactg cctcgccgaa aggtctctat 3780

ttcttcattt ggcgttaatt tgtggggttt gtattactcg gggtcagaga ttctgctga 3840
gggttgggtg cagcggagca ctgatatagc tagacctcaa cattagaag cggcttatga 3900
cccagtgctg gttgatacga tttatttgtt tccgaagtt ggcagccgtg tattttggcg 3960
ctgtaatctg acggaacgta gtcggcagtt taaaggtctc tcattttggg aggtttggga 4020
tatacaagca caagaaaaac acaataaagc caatgcgaag caggatgagt taactaaacg 4080
cagggagctt gaggcgttta ttcagcaaac cattcagaaa gcgaataagt taacgccag 4140
tactactgag cccaatcaa cacgattaa gcagattaaa actaataaaa aagaagccgt 4200

gacctcggag cgiaaaaaac gtgcggagca tttgaagcca agctcttcag gtgatgaggc 4260
taaagttatt ctttcaacg cagtggaagc ggatgatcaa gaagattaca gcctaccac 4320
atacgtcct gaattatttc aggatccacc agaaaaggat gaggatgag tgctaccgg 4380
attcaagcag tttatcgtga tacgggggta gaggcttacc gtgataatcc tttatcgag 4440
gccttaccac cattacaaga gtcagtgaat agtgctgcat cactgaaatc ctctttacag 4500
cttacttct ctgacttga aaagtcctgt gttatcagag ctcataccat ttgtctatt 4560
ccagatgact atttcagcc attaggtacg catttgctac taagttagcg tatttcggtc 4620

atgattcgag gtggctacgt aggcagaaat cctaaaacag gagatttaca aaagcattta 4680
caaaatggtt atgagcgtgt tcaaacggga gatttgaga catttcgctt tgaggaggca 4740
cgatctacgg cacaaagctt attgtaatt ggttgttctg gtagtggga gacgacctt 4800
cttcatcgt tctagccac gtatcctcag gtgatttacc atcgtgaact caatgtagag 4860
caggtggtgt atttgaaaat agactgctcg cataatggtt cgctaaaaga aatctgcttg 4920
aatttttca gagcgttga tcgagccttg ggctcgaact atgagcgtcg ttatggetta 4980
aaacgcatg gtatagaaac catgttggct ttgatgtcgc aaatagccaa tgcaatgct 5040

ttagggtgt tggttattga tgaattcag catttaagcc gctctcttc ggttgatct 5100
caagagatgc tgaactttt tgtgacgatg gtgaatatta ttggcgtacc agtgatgtt 5160
attggtacc ctaaagcacg agagattttt gaggctgatt tgcggtctgc acgtagagg 5220
gcagggttg gagctatatt ctgggatcct atacaacaaa cgcaacgtgg aaagcccaat 5280
caagagtga tcgcttttac ggataatctt tggcaattac agcttttaca acgcaaagat 5340
gcgctgttat cggatgaggt ccgtgatgtg tggtatgagc taagccaagg agtgatggac 5400
atttagtaa aactttttgt actcgtcag ctccgtgctc tagctttagg caatgagcgt 5460

attaccgctg gtttattgcg gcaagtgtat caagatgagt taaagcctgt gcaccccatg 5520
 ctagaggcat tacgctcggg tatcccagaa cgcattgctc gttattctga tctagtcgtt 5580
 cccgagattg ataaacgggt aatccaactt cagctagata tcgcagcgat acaagaacaa 5640
 acaccagaag aaaaagccct tcaagagtta gataccgaag atcagcgtca tttatatctg 5700
 atgctgaaag aggattacga ttcaagcctg ttaattccca ctattaaaa agcgtttagc 5760
 cagaatccaa cgatgacaag acaaaaagtta ctgcctcttg ttttgcagtg gttgatggaa 5820
 ggcgaaacgg tagtgtcaga actagaaaag ccctccaaga gtaaaaaggt ttcggctata 5880

 aaggtagtca agcccagcga ctgggatagc ttgcctgata cggatttacg ttatatctat 5940
 tcacaacgcc aacctgaaaa aacctatgcat gaacggttaa aagggaaagg ggtaatagtg 6000
 gatatggcga gcttatttaa acaagcaggt tagccatgag aaactttcct gttccgtact 6060
 cgaatgagct gatttatagc actattgcac gggcagcgtt ttatcaaggg attgttagtc 6120
 ctaagcagct gttggatgag gtgtatggca accgcaaggt ggtcgtacc ttagtctgc 6180
 cctcgcatth aggtgtgata gcaagacatc tacatcaaac aggacgttac gctgttcagc 6240
 agcttattta tgagcatacc ttattccctt tatatgctcc gttttagtagc aaggagcggc 6300

 gagacgaagc tattcggtta atggagtacc aagcgcaagg tgcggtgcat ttaatgctag 6360
 gagtgcgtgc ttctagagtt aagagcgata accgctttag atactgccct gattgcgttg 6420
 ctcttcagct aaataggtat ggggaagcct tttggcaacg agattggat ttgcccgtt 6480
 tgccatattg tccaaaacac ggtgcttttag tcttctttga tagagctgta gatgatcacc 6540
 gacatcaatt ttggccttg ggtcactg agctgctttc agactacccc aaagactccc 6600
 tatctcaatt aacagcacta gctgcttata tagcccctct gttagatgct ccacagcgc 6660
 aagagctttc cccaagcctt gagcagtga cgctgtttta tcagcgctta gcgcaggatc 6720

 tagggetaac caaaagcaag cacattcgtc atgacttggg ggcggagaga gtgaggcaga 6780
 cttttagtga tgaggcacta gagaaactgg atttaaagtt ggcagagaac aaggacacgt 6840
 gttggctgaa aagtatattc cgtaagcata gaaaagcctt tagttattta cagcatagta 6900
 ttgttgggca agccttattg ccaaaactaa cggttataga agcgtacag caggcaagtg 6960
 ctcttactga gcaacttata acgacaagac ctgttagcca gtctgtgcaa cctaactctg 7020
 aagatttate tgtaagcat aaagactggc agcaactagt gcataaatac caaggaatta 7080
 aggcggcaag acagtcttta gaggggtggg tgctatacgc ttggctttac cgacatgaca 7140

 gggattggct agttcactgg aatcaacagc atcaacaaga gcgtctggca cccgcccta 7200
 gaggttgatt gaaccaaaga gatcgaattg ctgtacgaca actattaaga atcataaagc 7260
 gtctagatag tagccttgat cacccaagag cgacatcgag ctggctgtta aagcaactc 7320

ctaacggaac ctctcttgca aaaaatctac agaaactgcc tttgtagcg ctttgcttaa 7380
 agcgttactc agagagtgtg gaagattatc aaattagacg gattagccaa gcttttatta 7440
 agcttaaaca ggaagatgtt gagcttaggc gctggcgatt attaagaagt gcaacgttat 7500
 ctaaagagcg gataactgag gaagcacaaa gattcttgga aatggtttat ggggaagagt 7560

 gagtggttag gctagctaca tttaatgaca atgtgcaggt tgtacatatt ggtcatttat 7620
 tccgtaactc gggtcataag gagtggcgta tttttgtttg gtttaatcca atgcaagaac 7680
 ggaaatggac tcgatttact catttgcctt tattaagtcg agctaagggtg gttaacagta 7740
 caacaaagca aataaataag gcggatcgtg tgattgagtt tgaagcatcg gatcttcaac 7800
 gagccaaaat aatcgatttt cctaactctc cgtcctttgc ttccgtacgc aacaaggatg 7860
 gagcgagag ttcatattat tacgaagctg aaacaccata tagcaagact cgttatcaca 7920
 tcccacagtt agagctagct cggtcattat ttttagggga tcctctagag tcgacctgca 7980

 ggcatgcaag ctitggctgtt ttggcggatg agagaagatt ttcagcctga tacagattaa 8040
 atcagaacgc agaagcggtc tgataaaaca gaatttgcct ggcggcagta gcgcggtggt 8100
 cccacctgac cccatgccga actcagaagt gaaacgccgt agcgcctgat gtagtgtggg 8160
 gtctcccat gcgagagtag ggaactgcca ggcatcaaat aaaacgaaag gctcagtcga 8220
 aagactgggc ctttcgtttt atctgttgtt tgtcggtgaa cgctctctg agtaggacaa 8280
 atccgccggg agcggatttg aacgttgcga agcaacggcc cggagggtgg cgggcaggac 8340
 gcccgcata aactgccagg catcaatta agcagaaggc catcctgacg gatggccttt 8400

 ttgcgtttct acaaactctt ttgtttatit ttctaaatac attcaaatat gtatccgctc 8460
 atgagacaat aaccctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt 8520
 caacatttcc gtgtcgcctt tattcccttt ttgcggeat tttgccttcc tgtttttgct 8580
 cacccagaaa cgctggtgaa agtaaaagat gctgaagatc agttgggtgc acgagtgggt 8640
 tacatcgaac tggatctcaa cagcggtaag atccttgaga gttttcgccc cgaagaacgt 8700
 tttccaatga tgagcacttt taaagtctg ctatgtggcg cggattatc ccgtgttgac 8760
 gccgggcaag agcaactcgg tcgcegcata cactattctc agaatgactt ggttgagtac 8820

 tcaccagtca cagaaaagca tcttacggat ggcatgacag taagagaatt atgcagtgct 8880
 gccataacca tgagtataaa cactcgggcc aacttacttc tgacaacgat cggaggaccg 8940
 aaggagctaa ccgctttttt gcacaacatg ggggatcatg taactcgcct tgatcgttgg 9000
 gaaccggagc tgaatgaagc cataccaaac gacgagcgtg acaccacgat gcctgcagca 9060
 atggcaacaa cgttgcgcaa actattaact ggcgaactac ttaactctagc ttcccggcaa 9120
 caattaatag actggatgga ggcggataaa gttgcaggac cacttctgcg ctcgccctt 9180

ccggctggct ggtttattgc tgataaatct ggagccgggtg agcgtgggtc tcgcggtatc 9240

 attgcagcac tggggccaga tggtaaagccc tcccgtatcg tagttatcta caccagggg 9300
 agtcaggcaa ctatggatga acgaaataga cagatcgctg agataggtgc ctactgatt 9360
 aagcattggt aactgtcaga ccaagtttac tcataatatac tttagattga tttacgcgcc 9420
 ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga ccgtacact 9480
 tgccagcgcc cttagcggccg ctcttttcgc tttcttcct tcctttctcg ccacgttcgc 9540
 cgccggccag cctcgcagag caggattccc gttgagcacc gccaggtgcg aataaggac 9600
 agtgaagaag gaacaccgc tcgcgggtgg gcctacttca cctatctgc ccggcggcat 9660

 caccggcgcc acaggtgcgg ttgctggcgc ctatatgcc gacatcacg atggggaaga 9720
 tcgggctcgc cacttcgggc tcatgagcgc ttgtttcggc gtgggtatgg tggcaggccc 9780
 cgtggccggg ggactgttgg gcgccatctc ctgtcatgca ccattccttg cggcggcgg 9840
 gctcaacggc ctcaacctac tactgggctg cttcctaag caggagtcgc ataagggaga 9900
 gcgtcgatcc ccgacagtaa gacgggtaag cctgttgatg ataccgctgc ctactgggt 9960
 gcattagcca gtctgaatga cctgtcacgg gataatccga agtggtcaga ctggaaaatc 10020
 agagggcagg aactgctgaa cagcaaaaag tcagatagca ccacatagca gaccgccat 10080

 aaaacgccct gagaagcccg tgacgggctt ttcttgatt atggtagtt tccttgcag 10140
 aatccataaa aggcgcctgt agtgccattt accccattc actgccagag ccgtgagcgc 10200
 agcgaactga atgicagaa aaagacagcg actcaggtgc ctgatggtcg gagacaaaag 10260
 gaatattcag cgatttccc gagcttgcga ggggtgctact taagcctta gggttttaag 10320
 gtctgttttg tagaggagca aacagcgttt gcgacatcct tttgtaatac tgcggaactg 10380
 actaaagtag tgagttatac acagggctgg gatctattct tttatcttt tttattctt 10440
 tctttattct ataaattata accactttaa tataaaca aaacacac aaaggtctag 10500

 cggaatttac agagggtcta gcagaattta caagtttcc agcaaaggtc tagcagaatt 10560
 tacagatacc cacaactcaa aggaaaagga ctagtaatta tcattgacta gcccatctca 10620
 attggtatag tgattaaaat cacctagacc aattgagatg tatgtctgaa ttagttgttt 10680
 tcaaagcaaa tgaactagcg attagtcgct atgacttaac ggagcatgaa accaagctaa 10740
 ttttatgctg tttggcacta ctcaaccca cgattgaaaa ccctacaagg aaagaacgga 10800
 cggatcgtt cacttataac caatacgttc agatgatgaa catcagtagg gaaaatgctt 10860
 atggtgtatt agctaaagca accagagagc tgatgacgag aactgtgaa atcaggaatc 10920

ctttggttaa aggccttgag atttccagt ggacaaacta tgccaagttc tcaagcghaa 10980
 aattagaatt agtttttagt gaagagatat tgccttatct tttccagtta aaaaaattca 11040
 taaaatataa tctggaacat gttaagtctt ttgaaaacaa atactctatg aggatttatg 11100
 agtggttatt aaaagaacta acacaaaaga aaactcacia ggcaaatata gagattagcc 11160
 ttgatgaatt taagttcatg ttaatgcttg aaaataacta ccatgagttt aaaaggctta 11220
 accaatgggt tttgaaacca ataagtaaag atttaaacac ttacagcaat atgaaattgg 11280
 tggttgataa gcgaggccgc ccgactgata cgttgatfff ccaagttgaa ctatagatagac 11340

 aatggatct cgtaacghaa cttgagaaca accagataaa aatgaatggg gacaaaatac 11400
 caacaacat tacatcagat tcctacctac ataacggact aagaaaaaca ctacacgatg 11460
 ctttaactgc aaaaattcag ctcaccagt ttagggcaaa atttttgagt gacatgcaaa 11520
 gtaagatga tctcaatggg tctttctcat ggctcacgca aaaacaacga accacactag 11580
 agaacatact ggctaaatac ggaaggatct gaggttctta tggctcttgt atctatcagt 11640
 gaagcatcaa gactaacaaa caaaagtaga acaactgttc accgttcat atcaaaggga 11700
 aaactgtcca tatgcacaga tgaaaacggg gtaaaaaaga tagatacatc agagctttta 11760

 cgagttttg gtgcatttaa agctgttcac catgaacaga tgcacaatgt aacagatgaa 11820
 cagcatgtaa cacctaatag aacaggtgaa accagtataa caaagcaact agaacatgaa 11880
 attgaacacc tgagacaact tgttacagct caacagtcac acatagacag cctgaaacag 11940
 gcgatgctgc ttatcgaatc aaagctgccg acaaacgggg agccagtgc gcctcccgtg 12000
 gggaaaaaat catggcaatt ctggaagaaa tagcgtttc agcctgtggg cggacaaaat 12060
 agttgggaac tgggaggggt ggaatggag ttttaagga ttatttaggg aagagtgaca 12120
 aaatagatgg gaactgggtg tagcgtcgtg agctaatacg aaaattaaa atgacaaaat 12180

 agtttggaa tagatttcac ttatctggtt ggtcgacact agtattacc tgttatccct 12240
 agatttaaat gatatcggat cctagtaagc cacgttttaa ttaatcagat gggccaatag 12300
 cggccgcaa ttcgcgcgag aagcgaagc ggcattgac tacgttgaca ccatcgaatg 12360
 gtgcaaaacc tttcgcggtg tggcatgata gcgcccggaa gagagtcaat tcagggtggg 12420
 gaatgtgaaa ccagtaactg tatacagatg cgcagagtat gccggtgtct ctatcagac 12480
 cgtttcccgc gtgtgaaacc agccagcca cgtttctgag aaaacgcggg aaaaagtgga 12540
 agcggcagat gcggagctga attacattc caaccgctg gcacaacaac tggcgggcaa 12600

 acagtcgtt ctgattggg ttgccacctc cagtctggcc ctgcacgcgc cgtcgcaaat 12660
 tgtcgcggcg attaaatctc gcgccgatca actgggtgcc agcgtgggtg tgtcagatgg 12720
 agaacgaagc ggcgtcgaag cctgtaaagc ggcggtgcac aatcttctc gcacaacgct 12780

cagtgggctg atcattaact atccgctgga tgaccaggat gccattgctg tggaaagctgc 12840
 ctgcactaat gticcggcgt tatttcttga tgtctctgac cagacacca tcaacagtat 12900
 tattttctcc catgaagacg gtacgcgact gggcgtggag catctggtcg cattgggtca 12960
 ccagcaaatc gcgctgttag cgggcccatt aagtctgtc tcggcgcgic tgcgtctggc 13020

 tggctggcat aaatatctca ctgcaatca aattcagccg atagcggaac gggaaggcga 13080
 ctggagtgcc atgtccggtt ttcaacaaac catgcaaatg ctgaatgagg gcatcgttcc 13140
 cactgcatg ctggttgcca acgatcagat ggcgctgggc gcaatgcgcg ccattaccga 13200
 gtccgggctg cgcgttggtg cggatatctc ggtagtggga tacgacgata ccgaagacag 13260
 ctcatgttat atcccgcgt caaccacat caaacaggat tttcgctgc tggggcaaac 13320
 cagcgtggac cgcttgctgc aactctctca gggccaggcg gtgaaggca atcagctgtt 13380
 gcccgtctca ctggtgaaaa gaaaaaccac cctggcgccc aatacgcaaa ccgctctcc 13440

 ccgcgcttg gccattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg 13500
 gcagtgagcg caacgcaatt aatgtgagtt agcgcgaatt gatctggttt gacagcttat 13560
 catcgactgc acggtgcacc aatgettctg gcgtcaggca gccatcggaa getgtggtat 13620
 ggctgtgcag gtcgtaaatc actgcataat tcgtgtcgt caaggcgac tcccgttctg 13680
 gataatgttt ttgcccga catcataacg gttctggcaa atattctgaa atgagctgtt 13740
 gacaattaat catccggctc gtataatgtg tggaaatgtg agcggataac aatttcacac 13800
 aggaaacagc gccgctgaga aaaagcgaag cggcactgct ctttaacaat ttatcagaca 13860

 atctgtgtgg gcactcgacc ggaattatcg attaacttta ttattaaaa ttaaagaggt 13920
 atatattaat gtatcgatta aataaggagg aataaacat ggcggacacg ttattgattc 13980
 tgggtgatag cctgagcgc gggatcgaat tgtctgccag cgcgcctgg cctgccttgt 14040
 tgaatgataa gtggcagagt aaaacgtcgg tagttaatgc cagcatcagc ggcgacacct 14100
 cgcaacaagg actggcgcgc cttccggctc tgctgaaaca gcatcagccg cgttgggtgc 14160
 tggttgaaact gggcggcaat gacggtttgc gtggttttca gccacagca accgagcaaa 14220
 cgctgcgcca gatcttcag gatgtcaaag ccgccaacgc tgaaccattg ttaatgcaaa 14280

 tacgtctgcc tgcaaatat ggtcgcggtt ataatgaagc ctttagcgc atttaccca 14340
 aactcgcaa agagtttgat gttccgctgc tgeccttttt tatggaagag gctacctca 14400
 agccacaatg gatgcaggat gacgtattc atcccaaccg cgacgccag cgtttattg 14460
 ccgactggat ggcgaagcag ttgcagcctt tagtaaatca tgactcataa tgactctaga 14520
 aataatttta gttaaagtata agaaggagat ataccatggt gaagaaggtt tggcttaacc 14580
 gttatcccgc ggacgttccg acggagatca accctgaccg ttatcaatct ctggtagata 14640

tgtttgagca gtcggtcgcg cgctacgccg atcaacctgc gtttgaat atgggggagg 14700

 taatgacctt ccgcaagctg gaagaacgca gtcgcgctt tgccgcttat ttgcaacaag 14760
 ggttggggct gaagaaaggc gatcgcgttg cgttgatgat gcctaattta ttgcaatatic 14820
 cgggtggcgt gtttggcatt ttgctgccc ggatgatcgt cgtaaacgtt aacccttgt 14880
 ataccccgcg tgagcttgag catcagctta acgatagcgg cgcacggcg attgttatcg 14940
 tgtctaactt tgetcacaca ctggaaaaag tggttgataa aaccgctt cagcacgtaa 15000
 ttctgaccg tatggcgat cagctatcta cggcaaaagg cacggtagtc aatttcgtt 15060
 ttaaatacat caagcgttg gtgccgaaat accatctgcc agatgccatt tcatttcgta 15120

 gcgcaactgca taacggctac cggatgcagt acgtcaaacc cgaactggg cgggaagatt 15180
 tagcttttct gcaatacacc ggcgcacca ctggtgtggc gaaaggcgcg atgctgactc 15240
 accgcaatat gctggcgaac ctggaacagg ttaacgcgac ctatggtccg ctgttgcac 15300
 cgggcaaaga gctgggtgtg acggcgtgc cgtgtatca catttttgc ctgaccatta 15360
 actgcctgct gtttatcgaa ctgggtgggc agaacctgct tactactaac ccgcgcgata 15420
 ttccagggtt ggtaaaagag ttagecgaat atccgtttac cgctatcac ggcttaaca 15480
 ccttgttcaa tgcgttgctg aacaataaag agttccagca gctggatttc tccagtctgc 15540

 atctttccg aggcggagg atgccagtgc agcaagtgt ggagagcgt tgggtgaaac 15600
 tgacaggaca gtatctgctg gaaggctatg gccttaccga gtgtgcgcc ctggtcagcg 15660
 ttaaccata tgatattgat tatcatagt gtatcatcgg ttgcccgtg ccgtcgacgg 15720
 aagccaaact ggtggatgat gatgataatg aagtaccacc ggtcaaccg ggtgagcttt 15780
 gtgtcaaagg accgcaggtg atgctgggtt actggcagcg tccgatgct acagatgaga 15840
 tcatcaaaaa tgctgttga cacaccggc acatcgcgt gatgatgaa gaagggttcc 15900
 tgcgcattgt cgatcgtaaa aaagacatga ttctggttc cgttttaac gtctatcca 15960

 acgagattga agatgctgc atgcagcatc ctggcgtaca ggaagtcgc gctgttggcg 16020
 tactttccg ctccagtgt gaagcgggta aatcttctg agtgaaaaa gatccatcgc 16080
 ttaccgaaga gtcactggtg accttttgc gccgtcagct cacgggtac aaagtaccga 16140
 agctggtgga gtttctgat gagttaccga aatctaact cggaataatt ttgcgacgag 16200
 aattactgta cgaagcgcg gcgcaagtgg acaataaagc ctgataactc tagaaataat 16260
 ttaaattgaa ttcgaagctt gggcccgaac aaaaactcat ctcagaagag gatctgaata 16320
 gcgccgtcga ccatcatcat catcatcatt gagtttaaac ggtctccagc ttggctgttt 16380

tggcggatga gagaagattt tcagcctgat acagattaaa tcagaacgca gaagcggctt 16440
 gataaacag aatttgctg gcggcagtag cgcggtggtc ccacctgacc ccattgccaa 16500
 ctgagaagtg aaacccgta gcgccgatgg tagtgtgggg tctcccatg cgagagtagg 16560
 gaactgccag gcatcaaata aaacgaaagg ctcaagtcgaa agactgggcc tttcgtttta 16620
 tctgttgttt gtcggtgaac gctctcctga ttaattaaga cgtctaagaa accattatta 16680
 tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtctt caagaatttt 16740
 ataaaccgtg gagcgggcaa tactgagctg atgagcaatt tccgttgac cagtgcctt 16800

ctgatgaagc gtcagcacga cgttctctg caccgtaacg ctgcccga atttgattcc 16860
 tttcagcttt gcttctctg gccctcatt cgtgcgctct aggatcctcc ggcgttcagc 16920
 ctgtgccaca gccgacagga tggtagccac catttgccc atatcaccgt cggtagctgat 16980
 cccgtctca ataaaccgaa ccgctacacc ctgagcatca aactctttaa tcagttggat 17040
 catgtcggcg gtgtcgcggc caagacggtc gagcttcttc accagaatga catcaccttc 17100
 ctccaccttc atctcagca aatccagccc ttcccgatct gttgaactgc cggatgcctt 17160
 gtcggtaaag atgcggttag cttttacccc tgcattttg agcgtgagg tctgcctctg 17220

gaagaagggtg ttgtgactc ataccaggcc tgaatcgccc catcatccag ccagaaagtg 17280
 agggagccac ggttgatgag agctttgttg taggtggacc agttggtgat tttgaacttt 17340
 tgctttgcca cggaacggtc tgcgtgtcgc ggaagatgcg tgatctgac cttcaactca 17400
 gcaaaagttc gattattca acaaagccgc cgtcccgtca agtcagcgta atgccctagg 17460
 aggcgcgcca cggccgcgtc gacccacgc ccctctttaa tacgacgggc aatttgcaact 17520
 tcagaaaatg aagagtttgc tttagccata acaaaagtcc agtatgcttt ttcacagcat 17580
 aactggactg atttcagttt acaactattc tgtctagttt aagactttat tgtcatagtt 17640

tagatctatt ttgttcagtt taagacttta ttgtccgccc aca 17683

<210> 18

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 18

gatgctggtg gcgaagctgt 20

<210> 19

<211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 primer"
 <400> 19
 gttgcgacgg tggtagcat aac 23

<210> 20
 <211> 35
 <212> DNA
 <213> Artificial Sequence
 <220><221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 primer"
 <400> 20
 gaggaataaa ccatgccat tcttgaaagc gtggg 35

<210> 21
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 primer"
 <400> 21
 agctggagac cgittaaact tataaacgc tgtttgcgc aaccg 45

<210> 22
 <211> 663
 <212> DNA
 <213> Unknown
 <220><221> source
 <223> /note="Description of Unknown: Uncultured bacterium RM44"
 <400> 22

atgccattc ttgaaagcgt gggtttcatg aagacgctat gggagagcgg cggcgctcag	60
gtcgccctga tggaatcgcg ggaagagacc agccacatgg tcggcatcct ggaagggatc	120
gccgccgaac tctcgtggcg tcttggtacg cagcttcgcg attatcagga tagggcgcg	180
cacctggccg ttctggtcgg gtcggagatc gtcggcggct tgcagatcgt cacgtcgccg	240
tcagcggatt gccttccta ccggtcgtc tggccgaag tctgcgtgcc ggacggtgcc	300
gcaatcgcg acatcacgat tctggcgcta cggaaggaat accgcgccg ctcaacctc	360
ttctggccgc tgtgcgtcga gctctggcgg cactgtgtcg cggagggtgc cacggagatg	420
cgctggagg caacgccga tacgtcagg ctctaccgcc gcatcggctg gccgctggag	480
gcatcggcg acctgcgct ccactggaac gagccgtgct tctgtgccc gatggggatc	540
gtagatgtcg cggggcgat ggttgtcgg gccttgcagt ccgccaccta tcaggcgtc	600
ctcgcgggga tgagtcggcc tgtggcgtca gcgtcgccgg ttgcgacaaa cagcggttta	660
tga	663