



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0112580
(43) 공개일자 2014년09월23일

(51) 국제특허분류(Int. Cl.)
C12N 15/53 (2006.01) C12N 15/63 (2006.01)
C12P 7/06 (2006.01)
(21) 출원번호 10-2014-7024702 (분할)
(22) 출원일자(국제) 2010년07월30일
심사청구일자 없음
(62) 원출원 특허 10-2013-7000550
원출원일자(국제) 2010년07월30일
심사청구일자 2013년01월08일
(85) 번역문제출일자 2014년09월02일
(86) 국제출원번호 PCT/US2010/043962
(87) 국제공개번호 WO 2011/155954
국제공개일자 2011년12월15일
(30) 우선권주장
12/802,560 2010년06월09일 미국(US)

(71) 출원인
코스카타, 인코포레이티드
미국 일리노이주 60555 워렌빌 스위트 100 위버
파크웨이 4575
(72) 발명자
리브즈 앤드류
미국 일리노이주 60661 시카고 유니트 316 노쓰
클린턴 스트리트 226
다타 라틴
미국 일리노이주 60657 시카고 유니트 #3 웨스
트 멜로즈 422
(74) 대리인
장훈

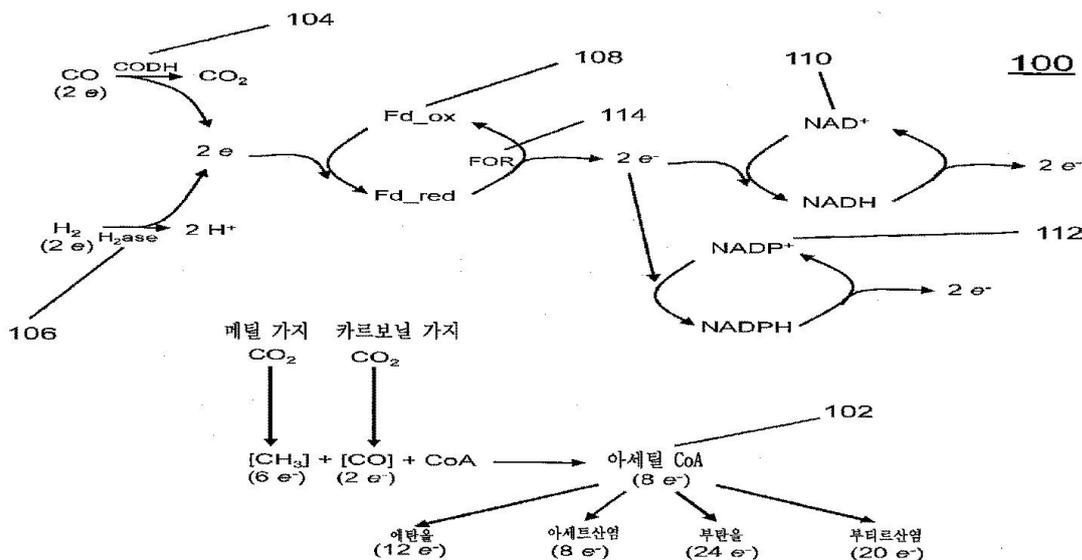
전체 청구항 수 : 총 7 항

(54) 발명의 명칭 합성가스로 에탄올을 생성하기 위한 주요 클로스트리듐 촉매 메커니즘을 암호화하는 유전자의 클로닝 및 발현 및 그에 관한 기능 특성분석

(57) 요약

주요 초산형성 클로스트리듐 종의 유전자 서열을 서열 분석하고 분리하였다. 목적하는 유전자를 동정하고, 기능성을 확인하였다. 클로스트리듐 종에서의 대사 촉매 활성을 위한 목적하는 주요 유전자는 CODH 활성을 암호화하는 세 가지-유전자 오페론, PTA-ACK를 암호화하는 두 가지-유전자 오페론, 및 신규 아세틸 조효소 A 환원 효소를 포함한다. 두 가지 오페론 및 아세틸 coA 환원효소의 프로모터 영역을 조작하여 에탄올 생성을 증가시킨다.

대표도



특허청구의 범위

청구항 1

아세틸 조효소 A 환원효소 및 프로모터를 인코딩하는 폴리펩티드를 암호화하는 뉴클레오티드 서열을 포함하는, 서열번호 3과 98% 이상 동일한, 분리된 폴리뉴클레오티드.

청구항 2

제1항의 폴리뉴클레오티드를 포함하는 벡터.

청구항 3

제1항의 폴리뉴클레오티드를 함유하는 분리된 형질전환체.

청구항 4

제2항의 벡터를 함유하는 분리된 형질전환체.

청구항 5

제1항의 뉴클레오티드 서열에 대한 안티센스 핵산으로서, 효소의 발현을 억제하는, 안티센스 핵산.

청구항 6

제1항의 폴리뉴클레오티드를 함유하는 혐기성, 에탄올 생성 미생물을 분리 및 정제하고, 발효 생물 반응기 내에서 상기 미생물을 사용하여 합성가스를 발효시킨 후, 제1항의 환원효소를 통해 아세틸 CoA로부터 에탄올 생산을 촉진하기 위한 충분한 성장 조건을 제공함을 포함하는, 에탄올 생성 방법.

청구항 7

제1항의 아세틸 조효소 A 환원효소의 활성을 증가시키거나 뉴클레오티드 서열의 과다발현을 일으키도록 제1항의 뉴클레오티드 서열의 프로모터 영역을 변형 또는 복제함을 포함하여, 제1항의 뉴클레오티드 서열을 포함한 미생물에서 에탄올 생성(ethanogenesis)을 증가시키는 방법.

명세서

기술분야

- [0001] 관련 미국 출원 데이터
- [0002] 이 출원은 일부 계속 출원으로서 2008년 12월 16일에 출원된 미국 특허 출원 12/336,238에 대한 우선권 주장 출원이다. 그 출원 전체가 이 안에 참조로서 통합되었다.
- [0003] 이 발명은 CO, H₂ 및 CO 및/또는 H₂를 함유한 혼합물을 바이오 연료로 생물학적으로 전환하는 데 사용되는 미생물의 새로운 유전자 서열에 대한 클로닝 및 발현에 관한 것이다.

배경기술

- [0004] 합성가스(syngas)는 일산화탄소(CO) 기체, 이산화탄소(CO₂) 기체, 수소(H₂) 기체 및 CH₄, N₂, NH₃, H₂S 및 기타 미량 기체 등 다른 휘발성 기체의 혼합물이다. 합성가스는 바이오매스, 유기 폐기물, 석탄, 석유, 플라스틱 또는 기타 탄소 함유 재료나 개질된 천연가스 등 다양한 유기 물질의 가스화를 통해 생성된다.
- [0005] 합성가스를 함유한 대기 중에서 성장한 초산형성 클로스트리듐 (acetogenic Clostridia) 미생물은 합성가스 성분인 CO, CO₂ 및 H₂을 흡수하여 지방족 C₂-C₆ 알코올과 지방족 C₂-C₆ 유기산을 생성할 수 있다. 이 합성가스 성분은 도 1에 표시된 Wood-Ljungdahl 대사 경로 100을 활성화하며 이 경로에서 경로 내 주요 중간 생성물인 아세틸 조효소 A 102가 형성된다.

- [0006] Wood-Ljungdahl 경로 100을 활성화하는 효소는 일산화탄소 탈수소효소(CODH) 104와 수소화효소(H_2ase) 106이다. 이 효소는 합성가스 내 CO 및 H_2 로부터 전자를 포획하여 황화철(FeS) 전자 운반체 단백질인 페리독신 108로 전달한다. 합성가스 발효 과정 동안 산화환원전위가 매우 낮기 때문에(보통 400mV와 500mV 사이) 초산형성 클로스트리듐의 Wood-Ljungdahl 경로 100에서 주요한 전자 운반체는 페리독신 108이다. 전자 전달 시 페리독신 108은 전자 상태를 Fe^{3+} 에서 Fe^{2+} 로 변화시킨다. 페리독신과 결합된 전자는 페리독신 산화환원효소 114(FORs)의 작용을 통하여 보조인자 NAD^+ 110 및 $NADP^+$ 112로 전달된다. 환원된 뉴클레오티드 보조 인자(NAD^+ 및 $NADP^+$)는 Wood-Ljungdahl 경로 100에서 중간 화합물 생성에 사용되어 아세틸-CoA 102를 형성한다.
- [0007] Wood-Ljungdahl 경로 100을 통한 아세틸-CoA 102 형성은 도 2에 더 상세하게 표시되어 있다. CO_2 202나 CO 208은 경로에 대한 기질을 제공한다. CO_2 202의 탄소가 환원되어 포름산 탈수소효소(FDH) 204에 의해 먼저 포름산염으로 환원되고 다음으로는 메틸 테트라히드로엽산 중간 생성물 206으로 추가 환원되는 연속적인 환원을 통해 메틸 그룹으로 환원된다. CO 208의 탄소는 경로의 2차 가지를 통하여 일산화탄소 탈수소효소(CODH) 104에 의해 카르보닐 그룹 210으로 환원된다. 그런 다음 두 개의 탄소 일부분은 일산화탄소 탈수소효소(CODH/ACS) 복합체의 일부인 아세틸-CoA 합성효소(ACS) 212의 작용을 통하여 Acetyl-CoA 102로 축합된다. 아세틸-CoA 102는 초산형성 클로스트리듐 내 C_2-C_6 알코올과 산을 생성하는 중간 대사물이다.
- [0008] 아세틸 CoA 102로부터 에탄올의 생성은 두 가지 가능한 경로 중 하나를 통해 이루어진다. 알데히드 탈수소효소는 아세트 알데히드의 생성을 촉진하며 아세트 알데히드는 1차 알코올 탈수소효소의 작용에 의해 에탄올로 산화된다. 다른 방법으로, 균질 초산생성 미생물에서는 NADPH-의존 아세틸 CoA 환원효소("AR")가 아세틸 CoA로부터 직접 에탄올 생성을 촉진한다.
- [0009] Wood-Ljungdahl 경로 100은 아세트산염 214가 생성될 때(도 2) ATP 생성과 관련하여 전기적으로 중성이다. 에탄올 216이 생성될 때 메틸렌 테트라히드로엽산이 메틸 테트라히드로엽산 206으로 환원되는 단계에서 1 ATP가 소모되며 따라서 이 프로세스는 1 ATP 때문에 전기적으로는 순 (-)를 띄게 된다. 아세틸- PO_4 218이 아세트산염 214로 전환될 때 경로가 균형을 이루게 된다.
- [0010] 초산형성 클로스트리듐 유기체는 이온 구배에 의한 인산화를 통해 세포 에너지를 생성한다. CO 환경에서 성장 시 막횡단 전위가 발생하며 ADP로부터 ATP를 합성하는 데 사용된다. 프로세스를 중재하는 효소는 수소화효소, NADH 탈수소효소, 일산화탄소 탈수소효소 및 메틸렌 테트라히드로엽산 환원효소를 포함한다. ATP 생성 단계에 관여하는 것처럼 보이는 막 운반체는 퀴논, 메나퀴논 및 시토크롬을 포함한다.
- [0011] 초산형성 클로스트리듐은 Wood-Ljungdahl 경로 100을 통해 상업적 관심물질인 에탄올, n-부탄올, 헥산올, 아세트산 및 부티르산 등의 C_2-C_6 알코올 및 산의 혼합물을 생성한다. 예를 들어 아세트산염 및 에탄올은 C. ragsdalei에 의해 일부 발효 상태에 의존하여 다양한 비율로 생성된다. 그러나, 원하는 산물, 예를 들어 에탄올 같은 알코올 생산 비용은, 해당하는 산, 예컨대 아세트산염의 생산을 줄이거나 아예 없애므로써 생산을 극대화할 경우, 크게 낮출 수 있다. 따라서 Wood-Ljungdahl 경로 100을 통한 선택된 C_2-C_6 알코올 또는 산의 생산성을 향상하기 위해서는 경로 내 주요 효소의 효소 활성을 조절함으로써 초산형성 클로스트리듐을 대사적으로 처리하는 것이 바람직하다.

발명의 내용

해결하려는 과제

- [0012] 본 발명의 한 가지 측면은 에탄올 대사 프로세스를 합성가스를 촉매작용하는 효소를 암호화하는 세 가지 주요 오페론에 대한 새로운 서열을 제공하는 것이다: 일산화탄소 탈수소효소, 막관련 전자 전달 단백질, 페리독신 산화환원효소 및 프로모터를 암호화하는 첫 번째 오페론; 아세테이트키나제, 인산아세틸기전이효소와 프로모터를 암호화하는 두 번째 오페론; 아세틸 CoA 환원효소와 프로모터를 암호화하는 세 번째 오페론.
- [0013] 발명의 또 다른 측면은 위에서 설명한 오페론에 대한 폴리뉴클레오티드 서열 유전암호를 포함하는 분리된 벡터 또는 형질전환체를 제공하는 것이다.

- [0014] 발명의 또 다른 측면은 다음과 같은 에탄올 생성 방법을 제공하는 것이다: 일산화탄소 탈수소효소, 막관련 전자 전달 단백질, 페리독신 산화환원효소 및 프로모터를 포함한 오페론, 아세테이트 키나제, 인산아세틸기전이효소 및 프로모터를 암호화하는 오페론, 또는 아세틸 CoA 환원효소와 프로모터를 암호화하는 오페론을 암호화하는 폴리뉴클레오티드를 함유하는 혐기성, 에탄올 생성 미생물을 분리하고 정제하는 단계; 발효 생물 반응기 내에서 상기 미생물을 이용하여 합성가스를 발효시키는 단계; 아세틸 CoA로부터 에탄올 생산을 촉진하기 위해, 충분한 아연(이에 제한되지는 않음)을 포함하여 NADPH의 세포성 생산을 위한 충분한 성장 조건을 제공하는 단계.
- [0015] 발명의 또 다른 측면은, 아세틸 조효소 A 환원효소를 암호화하는 폴리뉴클레오티드를 함유하는 혐기성 에탄올 생성 미생물을 분리하고 정제하고, 발효 생물 반응기에서 상기 미생물을 사용하여 합성가스를 발효시키고, 아세틸 CoA로부터 에탄올 생산을 촉진하기 위해 충분한 아연(이에 제한되지는 않음)을 포함하여 NADPH의 세포성 생산을 위한 충분한 성장 조건을 제공함으로써 에탄올을 생성하는 방법을 제공하는 것이다.
- [0016] 본 발명의 또 다른 측면은 위에서 설명한 오페론 중 하나를 암호화하는 뉴클레오티드 서열(들)을 포함하는 미생물에서 아세트산염에 대한 에탄올 생산 비율 또는 에탄올 생성을 증가시키는 방법으로서, 당해 방법은, 청구항 16의 아세틸 조효소 A 환원효소의 활성을 증가시키거나 상기 뉴클레오티드 서열의 과다발현이나 미달발현을 유발시키기 위해 상기 뉴클레오티드 서열의 프로모터 영역을 변형, 복제 또는 하향 조절함을 포함한다.
- [0017] 본 발명은 다양한 양태와 아래 제시한 상세 설명을 기술하는 부수적인 도면을 통해 설명하였다. 구체적인 양태로 발명을 한정하기 위해 도면을 선택한 것이 아니라 설명과 이해를 위해 선택한 것이다. 상세 설명과 도면은 첨부한 청구 및 균등물에 의해 정의되는 발명의 범위를 제한하기보다는 단지 발명에 대한 실례를 든 것이다. 도면은 임의의 척도로 작성되었다. 상기 측면 및 기타 현재 발명에 대한 다른 부수적 장점은 첨부한 도면과 연관된 상세한 설명을 통해 좀 더 쉽게 이해될 것이다.

과제의 해결 수단

- [0018] 본 발명은 CO, CO₂, H₂, 또는 그 혼합물을 포함하는 합성가스로부터 에탄올과 산을 생성하는 초산형성 클로스트리듐 미생물을 암호화하는 새로운 유전자 서열에 관한 것이다.
- [0019] Wood-Ljungdahl 경로를 통해 C₂-C₆ 알코올과 산을 생성하는 초산형성 클로스트리듐의 여러 종: *C. ragsdahlei*, *C. ljungdahlii*, *C. carboxydivorans* 및 *C. autoethanogenum*에 대한 특성이 기술되어 있다. 관심 대상인 효소의 게놈 일부의 위치를 특정하기 위해 이들 미생물 중 세 가지의 게놈을 서열결정하였다.
- [0020] Wood-Ljungdahl 대사 경로 및 *C. ragsdahlei* 게놈에서 발견한 에탄올 합성에서 효소를 암호화하는 유전자가 표 1에 제시되어 있다. 첫 번째 열은 각 유전자와 연관이 있는 경로를 나타낸 것이다. 두 번째 열에 표시된 유전자 ID 번호는 도 1과 도 2에 표시된 Wood-Ljungdahl 경로 내 대사 반응에 효소가 연관되어 있음을 보여주는 번호와 일치한다.

표 1

표 1. 대사 공학 실험에 사용된 *Clostridium ragsdalei* 유전자

경로	유전자 ID	유전자 이름	EC 번호	ORF ID	복제 ID	설명
Wood-Ljungdahl	1	일산화탄소 탈수소효소	1.2.2.4	RCCC00183	CODH_1	일산화탄소 산화
	2			RCCC01175	CODH_2	일산화탄소 산화
	3			RCCC01176	CODH_3	일산화탄소 산화
	4			RCCC02026	CODH_4	일산화탄소 산화
	5			RCCC03874	CODH_5	일산화탄소 산화
	6	일산화탄소 탈수소효소 / 아세틸-CoA 합성효소	1.2.99.2	RCCC03862	cooS/acsA	2기능성 CODH/ACS 효소, 탄소 고정
	7	포름산 탈수소효소	1.2.1.2	RCCC00874	FDH_1	메틸 가지 사슬 탄소 고정
	8			RCCC03324	FDH_2	
	9	포름일테트라하이드로엽산 합성효소	6.3.4.3	RCCC03872	FTHFS	메틸 가지 사슬 탄소 고정
	10	메틸렌테트라하이드로엽산 사이클로화제	3.5.4.9	RCCC03870	MEC	메틸 가지 사슬 탄소 고정
	11	메틸렌테트라하이드로엽산 탈수소효소	1.5.1.5	RCCC03870	MED	메틸 가지 사슬 탄소 고정
	12	메틸렌테트라하이드로엽산 환원효소	1.5.1.20	RCCC03868	MER	메틸 가지 사슬 탄소 고정
	13	메틸전달효소	2.1.1.13	RCCC03863	acsE	메틸 가지 사슬 탄소 고정
	14	코리노이드/철-황 단백질	1.2.99.2	RCCC03864	acsC	CODH/ACS 복합체 일부, 큰 소단위체
	15	코리노이드/철-황 단백질	1.2.99.2	RCCC03865	acsD	CODH/ACS 복합체 일부, 작은 큰 소단위체
16	아세트키나제	2.7.2.1	RCCC01717	ACK	아세트산염 생성	
17	인-아세틸기 전이효소	2.3.1.8	RCCC01718	PTA	아세트산염 생성	

[0021]

에탄올 및 아세트산염 생성	18		1.2.7.5	RCCC00020	AOR_1	아세트산염을 아세트알데히드로 환원
	19	링스텐 함유 알데히드 페리독신 산화환원효소	1.2.7.5	RCCC00030	AOR_2	아세트산염을 아세트알데히드로 환원
	20		1.2.7.5	RCCC01183	AOR_3	아세트산염을 아세트알데히드로 환원
	21	아세틸-CoA 환원효소	1.1.1.2	RCCC02715	ADH_1	아연 함유, NADPH-의존 아세틸-CoA 환원효소
	22	알코올 탈수소효소	1.1.1.1	RCCC01356	ADH_2	2개의 다중서열정렬(pfam) 도메인: FeADH 및 ALDH, AdhE
	23		1.1.1.1	RCCC01357	ADH_3	2개의 다중서열정렬(pfam) 도메인: FeADH 및 ALDH, AdhE
	24		1.1.1.1	RCCC01358	ADH_4	2개의 다중서열정렬(pfam) 도메인: FeADH 및 ALDH, AdhE, 단편(76aa)
	25		1.1.1.1	RCCC03300	ADH_5	1개의 다중서열정렬(pfam) 도메인: FeADH
	26		1.1.1.1	RCCC03712	ADH_6	1개의 다중서열정렬(pfam) 도메인: FeADH
	27		1.1.1.1	RCCC04095	ADH_7	1개의 다중서열정렬(pfam) 도메인: FeADH
28	1.1.1.1		RCCC00004	ADH_8	짧은 사슬 ADH, 다중 복제	
29	1.1.1.1		RCCC01567	ADH_9	짧은 사슬 ADH, 다중 복제	
30	1.1.1.1	RCCC02765	ADH_10	짧은 사슬 ADH, 다중 복제		

[0022]

	31		1.-.-.-	RCCC02240	ADH_11	짧은 사슬 ADH, 다중 복제
	32	알데히드 탈수소효소	1.2.1.10	RCCC03290	ALDH_1	아세틸화
	33		1.2.1.10	RCCC04101	ALDH_2	아세틸화
	34		1.2.1.10	RCCC04114	ALDH_3	아세틸화
수소화 효소	35	수소화효소	1.12.7.2	RCCC00038	HYD_1	철만 해당, H2 생산
	36		1.12.7.2	RCCC00882	HYD_2	철만 해당, 큰 소단위체, H2 생산
	37		1.12.7.2	RCCC01252	HYD_3	철만 해당, H2 생산
	38		1.12.7.2	RCCC01504	HYD_4	철만 해당, H2 생산
	39		1.12.7.2	RCCC02997	HYD_5	Ni-Fe 큰 소단위체, H2 산화
전자 운반체	40	페리독신		RCCC00086		
	41		RCCC00301			
	42		RCCC00336			
	43		RCCC01168			
	44		RCCC01415			
	45		RCCC01825			
	46		RCCC02435			
	47		RCCC02890			
	48		RCCC03063			
	49		RCCC03726			
	50		RCCC04003			
51	RCCC04147					
전자 전달	52	이황화 피리딘뉴클레 오티드 산화환원효소		RCCC02615		글루타민산염 합성효소 작은 사슬, 그 옆에 큰 사슬은 아님
	53		RCCC02028		cooF 및 cooS 옆, 아마도 환원된 피리딘 보조인자 생성에 중요.	

[0023]

	54			RCCC03071		NADH 탈수소효소, 오페론의 일부가 아님
	55	막관련 전자 전달 FeS 단백질, cooF		RCCC02027	cooF	4번 유전자와 53번 유전자 사이

[0024]

[0025]

C. *ljungdahlii* 게놈의 서열 분석을 수행하였다. Wood-Ljungdahl 경로, 에탄올 및 아세트산염 생산과 전자 전달 시 게놈 안에서 효소를 암호화하는 유전자를 확인하고 위치를 결정하였다. 그 결과를 표 2에 제시하였다.

표 2

표 2. 대사 공학 실험에 사용된 *Clostridium ljungdahlii* 유전자

경로	유전자 ID	유전자 이름	EC 번호	ORF ID	복제 ID	설명
Wood-Ljungdah 1	1	일산화탄소 탈수소효소	1.2.2.4	RCCD00983	CODH_1	일산화탄소 산화
	2			RCCD00984	CODH_2	일산화탄소 산화
	3			RCCD01489	CODH_3	일산화탄소 산화
	4			RCCD04299	CODH_4	일산화탄소 산화
	5	일산화탄소 탈수소효소 / 아세틸-CoA 합성효소	1.2.99.2	RCCD00972	CODH_ACS	2기능성 CODH/ACS 효소, 탄소 고정
	6	포름산 탈수소효소	1.2.1.2	RCCD01275	FDH_1	메틸 가지 탄소 고정
	7			RCCD01472	FDH_2	
	8	포름일테트라히드로엽산 합성효소	6.3.4.3	RCCD00982	FTHFS	메틸 가지 탄소 고정
	9	메틸렌테트라히드로엽산 사이클로리드물라제	3.5.4.9	RCCD00980	MEC	메틸 가지 탄소 고정
	10	메틸렌테트라히드로엽산 탈수소효소	1.5.1.5	RCCD00980	MED	메틸 가지 사슬 탄소 고정
	11	메틸렌테트라히드로엽산 환원효소	1.5.1.20	RCCD00978	MER	메틸 가지 사슬 탄소 고정
	12	메틸전달효소	2.1.1.13	RCCD00973	MET	메틸 가지 사슬 탄소 고정

[0026]

	13	코리노이드/철-황 단백질	1.2.99.2	RCCD00974	COPL	CODH/ACS 복합체 일부, 큰 소단위체
	14	코리노이드/철-황 단백질	1.2.99.2	RCCD00975	COPS	CODH/ACS 복합체 일부, 작은 큰 소단위체
에탄올 및 아세트산염 생산	15	아세트키나제	2.7.2.1	RCCD02720	ACK	아세트산염 생산
	16	인-아세틸기전이 효소	2.3.1.8	RCCD02719	PTA	아세트산염 생산
	17	텅스텐 함유 알데히드 페리독신 산화환원효소	1.2.7.5	RCCD01679	AOR_1	아세트산염을 아세트알데히드로 환원
	18		1.2.7.5	RCCD01692	AOR_2	아세트산염을 아세트알데히드로 환원
	19	아세틸-CoA 환원효소	1.1.1.2	RCCD00257	ADH_1	아연 함유, NADPH-의존 아세틸-CoA 환원효소
	20		1.1.1.1	RCCD00167	ADH_2	2개의 다중서열정렬(pfam) 도메인: FeADH 및 ALDH, AdhE
	21		1.1.1.1	RCCD00168	ADH_3	2개의 다중서열정렬(pfam) 도메인: FeADH 및 ALDH, AdhE
	22		1.1.1.1	RCCD02628	ADH_5	1개의 다중서열정렬(pfam) 도메인: FeADH
	23		1.1.1.1	RCCD03350	ADH_7	1개의 다중서열정렬(pfam) 도메인: FeADH
	24	알코올 탈수소효소	1.-.-.-	RCCD00470	ADH_8	짧은 사슬 ADH, 다중 복제

[0027]

	25		1.-.-.-	RCCD01665	ADH_9	짧은 사슬 ADH, 다중 복제
	26		1.-.-.-	RCCD01767	ADH_10	짧은 사슬 ADH, 다중 복제
	27		1.-.-.-	RCCD02864	ADH_11	짧은 사슬 ADH, 다중 복제
	28		1.2.1.10	RCCD02636	ALDH_1	아세틸화
	29	알데히드 탈수소효소	1.2.1.10	RCCD03356	ALDH_2	아세틸화
	30		1.2.1.10	RCCD03368	ALDH_3	아세틸화
수소화 효소	31	수소화 효소	1.12.7.2	RCCD00346	HYD_1	Ni-Fe 큰 소단위체, H2 산화
	32		1.12.7.2	RCCD00938	HYD_2	Ni-Fe 작은 소단위체, H2 산화
	33		1.12.7.2	RCCD01283	HYD_3	철만 해당, 큰 소단위체, H2 생산
	34		1.12.7.2	RCCD01700	HYD_4	철만 해당, H2 생산
	35		1.12.7.2	RCCD02918	HYD_5	철만 해당, H2 생산
	36		1.12.7.2	RCCD04233	HYD_6	철만 해당, H2 생산
전자	37	페리독신		RCCD00424		
운반체	38			RCCD01226		
	39			RCCD01932		
	40			RCCD02185		
	41			RCCD02239		
	42			RCCD02268		
	43			RCCD02580		
	44			RCCD03406		
	45			RCCD03640		
	46			RCCD03676		
	47			RCCD04306		
전자	48	이황화 퍼리딘뉴클레오티드		RCCD00185		글루타민산염 합성효소 작은 사슬, 그 옆에 있는 큰 사슬은 아님

[0028]

	49	산화환원효소		RCCD01487		cooF 및 cooS 옆, 아마도 환원된 퍼리딘 보조인자 생성에 중요
	50			RCCD00433		NADH 탈수소효소, 오페론의 일부가 아님
	51	막관련 전자 전달 FeS 단백질, cooF		RCCD01488	cooF	3번 유전자와 49번 유전자 사이

[0029]

[0030] 이와 유사하게 *C. carboxydvorans*의 게놈을 서열결정하고 Wood-Ljungdahl 경로 및 에탄올과 아세트산염 합성 시의 효소를 암호화하는 유전자를 확인하고 위치를 결정하였다. 그 결과를 표 3에 제시하였다.

[0031] [표 3]

[0032] 대사공학 실험에 사용된 *C. carboxydivorans*의 유전자

경로	유전자 ID	유전자 이름	EC 번호	ORF ID	복제 ID	설명
Wood-Ljungdahl	1	일산화탄소 탈수소효소	1.2.2.4	RCCB04039	CODH_1	일산화탄소 산화
	2			RCCB00154	CODH_2	일산화탄소 산화
	3			RCCB02478	CODH_3	일산화탄소 산화
에탄올 및 아세트산 염생산	4			RCCB03963	CODH_4	일산화탄소 산화
	5			RCCB04038	CODH_5	일산화탄소 산화
	6	일산화탄소 탈수소효소 / 아세틸-CoA 합성효소	1.2.99.2	RCCB04293	CODH_ACS	2기능성 CODH/ACS 효소, 탄소 고정
	7	포름산 탈수소효소	1.2.1.2	RCCB05406	FDH_1	메틸 가지 탄소 고정
	8			RCCB01346	FDH_2	메틸 가지 탄소 고정
9	포름일테트라히드로엽산	6.3.4.3	RCCB04040	FTHFS	메틸 가지 탄소 고정	

[0033]

	합성효소				
10	메틸렌테트라 히드로엽산 사이클 로히드롤라제	3.5.4.9	RCCB0404 2	MEC	메틸 가지 탄소 고정
11	메틸렌테트라 히드로엽산 탈수소효소	1.5.1.5	RCCB0404 2	MED	메틸 가지 탄소 고정
12	메틸렌 테트라히드로 엽산 환원효소	1.5.1.2 0	RCCB0404 4	MER	메틸 가지 탄소 고정
13	메틸전달효소	2.1.1.1 3	RCCB0429 4	MET	메틸 가지 탄소 고정
14	코리노이드/철- 황 단백질	1.2.99. 2	RCCB0404 9	COPL	CODH/ACS 복합체 일부, 큰 소단위체
15	코리노이드/철- 황 단백질	1.2.99. 2	RCCB0404 7	COPS	CODH/ACS 복합체 일부, 작은 큰 소단위체
16	아세트키나제	2.7.2.1	RCCB0524 9	ACK	아세트산염 생산
17	인-아세트기전 이효소	2.3.1.8	RCCB0248 1	PTA	아세트산염 생산
18	텅스텐 함유 알데히드 페리독신 산화환원효소	1.2.7.5	RCCB0006 3	AOR_1	아세트산염을 아세트알데히드 로 환원
19	알코올 탈수소효소	1.1.1.2	RCCB0358 4	ADH_1	아연-ADH
20		1.1.1.1	RCCB0387 0	ADH_2	2개의 다중서열정렬(pf am) 도메인: FeADH 및 ALDH, AdhE
21		1.1.1.1	RCCB0567 5	ADH_3	잘림, AdhE
22		1.1.1.1	RCCB0095 8	ADH_5	1개의 다중서열정렬(pf am) 도메인: FeADH
23		1.1.1.1	RCCB0448 9	ADH_6	1개의

[0034]

						다중서열정렬(pf am) 도메인: FeADH
	24		1.1.1.1	RCCB0450 3	ADH_7	1개의 다중서열정렬(pf am) 도메인: FeADH
	25		1.-.-.	RCCB0246 5	ADH_9	짧은 사슬 ADH, 다중 복제
	26		1.-.-.	RCCB0555 1	ADH_10	짧은 사슬 ADH, 다중 복제
	27	알데히드 탈수소효소	1.2.1.1 0	RCCB0240 3	ALDH_1	아세틸화
	28		1.2.1.1 0	RCCB0256 1	ALDH_2	아세틸화
	29		1.2.1.1 0	RCCB0403 1	ALDH_3	아세틸화
	30		수소화효소	1.12.7. 2	RCCB0224 9	HYD_1
	31	1.12.7. 2		RCCB0131 9	HYD_2	철만 해당, H2 생산
	32	1.12.7. 2		RCCB0140 5	HYD_3	철만 해당, H2 생산
	33	1.12.7. 2		RCCB0151 6	HYD_4	철만 해당, 큰 소단위체, H2 산화
	34	1.12.7. 2		RCCB0348 3	HYD_5	철만 해당, H2 생산
	35	1.12.7. 2		RCCB0541 1	HYD_6	철만 해당, 큰 소단위체, H2 생산
	36	페리독신			RCCB0023 4	
	37			RCCB0034 5		
	38			RCCB0126 0		
	39			RCCB0133 4		
	40			RCCB0177 5		

[0035]

41			RCCB0196 0		
42			RCCB0197 2		
43			RCCB0261 8		
44			RCCB0263 8		
45			RCCB0283 6		
46			RCCB0285 3		
47			RCCB0302 3		
48			RCCB0319 1		
49			RCCB0327 8		
50			RCCB0345 2		
51			RCCB0359 6		
52			RCCB0376 2		
53			RCCB0397 2		
54			RCCB0416 5		
55			RCCB0438 3		
56			RCCB0457 1		
57			RCCB0458 5		
58			RCCB0578 0		
59			RCCB0597 5		
60			RCCB0630 4		
61			RCCB0630 5		
전자 전달	62	이황화 피리딘뉴클레 오티드 산화환원효소	RCCB0044 2		NADH 탈수소효소, 오페론의 일부가 아님
	63		RCCB0167 4		NADH 탈수소효소, 오페론의 일부가 아님

[0036]

64			RCCB0351 0		cooF 및 cooS 옆, 아마도 환원된 피리딘 보조인자 생성에 중요
65			RCCB0058 6		NADH 탈수소효소, 오페론의 일부가 아님
66			RCCB0479 5		NADH: 페리독신 산화환원효소, 오페론의 일부가 아님
67	막관련 전자 전달 FeS 단백질, cooF		RCCB0350 9	cooF	2번 유전자와 64번 유전자 사이

[0037]

[0038]

전자 전달 경로에서 효소를 암호화하는 유전자는 일산화탄소 탈수소효소 [효소 위원회 번호(EC 1.2.2.4)]를 포함한다. *C. ragsdalei* 및 *C. ljungdahlii* 에서는 5개의 별도의 ORF (open reading frame, 단백질 아미노산 배열로 번역되는 DNA 염기 서열) 서열을 발견하였으며, *C. carboxidivorans* 계놈에서는 일산화탄소 탈수소효

소에 대해 6개를 발견하였다.

- [0039] 도 3은 일산화탄소 탈수소효소 오페론 300의 다이어그램이다. 오페론 300 내 유전자 순서는 초산형성 클로스트리듐 3개 종 모두에서 고도로 보존되었으며, 일산화탄소 탈수소효소(cooS)(유전자 ID 4, 표 1, 2, 및 3), 그 뒤에 막관련 전자 전달 FeS 단백질(cooF) (유전자 ID 55, 표 1; 유전자 ID 51, 표 2; 유전자 ID 67, 표 3), 및 다시 그 뒤에 페리독신 산화환원효소(FOR)를 암호화하는 유전자들을 포함한다.
- [0040] 초산형성 클로스트리듐 3개 종에 대하여 도 3의 오페론에서 찾은 유전자 서열에 대한 비교를 수행하였다. cooS 유전자는 *c. ragsdalei*와 *c. ljungdahlii* 사이 98%의 동일성을, *c. carboxydivorans*와 *c. ragsdahlii* 사이 84%의 동일성을, *c. carboxydivorans*와 *c. ljungdahlii* 사이 85%의 동일성을 가지고 있었다. cooF 유전자는 *c. ragsdalei*와 *c. ljungdahlii* 사이 98%의 동일성을, *c. carboxydivorans*와 *c. ragsdalei* 사이 80%의 동일성을, *c. carboxydivorans*와 *c. ljungdahlii* 사이 81%의 동일성을 가지고 있었다. FOR 유전자는 *c. ragsdalei*와 *c. ljungdahlii* 사이 97%의 동일성을, *c. carboxydivorans*와 *c. ragsdalei* 사이 77%의 동일성을, *c. carboxydivorans*와 *c. ljungdahlii* 사이 77%의 동일성을 가지고 있었다.
- [0041] 초산형성 클로스트리듐 각 종의 게놈에서 6개 수소화효소(EC 1.12.7.2) ORF 서열을 발견하였다.
- [0042] 12개의 페리독신 생물합성 유전자(유전자 ID 40-51)를 *C. ragsdalei* 게놈에서 발견하였다. 11개의 페리독신 생물합성 유전자(유전자 ID 37-47, 표 2)를 *C. ljungdahlii*에서 발견하였으며 26개는(유전자 ID 36-61, 표 3) *C. carboxidivorans*에서 발견하였다.
- [0043] 페리독신 산화환원효소를 암호화하는 3개 유전자는 페리독신과 니코틴아미드 보조인자 결합 도메인을 모두 포함하는 *C. ragsdalei* 게놈에서 발견하였다. 이 유전자의 ORF 서열 ID 번호(표 1)는 다음과 같다: RCCC02615, RCCC02028 및 RCCC03071. 대사 공학을 위한 주요 유전자, RCCC02028은 도 3에서도 표시된 cooS/cooF 오페론의 일부이다. 이와 유사하게 페리독신 산화환원효소(FOR)를 암호화하는 3가지 유전자가 *C. ljungdahlii* 게놈에서 발견되었다. 이들 유전자 각각이 페리독신 및 보조 인자 결합 도메인 둘다를 암호화한다. 이 유전자의 ORF 서열 ID 번호는 다음과 같다: RCCD00185, RCCD01847 및 RCCD00433(표 2). 주요 유전자 RCCD01847은 도 3에 표시된 cooF/cooS 오페론의 일부이다.
- [0044] *C. carboxidivorans* 게놈에서 페리독신과 보조인자 결합 도메인을 모두 포함한 5개 유전자를 찾을 수 있었다. 이 유전자의 ORF 서열 ID 번호(표 3)는 다음과 같다: RCCB00442, RCCB01674, RCCB03510, RCCB00586 및 RCCB04795. 전자 흐름을 조절할 가능성이 있는 주요 유전자는 cooF/cooS 오페론의 일부인(도 3) RCCB03510이다.
- [0045] *C. ragsdalei* 및 *C. ljungdahlii*에서 AR을 암호화하는 유전자 (유전자 ID 21, 표 1, 유전자 ID 19, 표 2)를 서열결정하였다. *C. ragsdalei* 및 *C. ljungdahlii*에서 아세틸 CoA 환원효소 유전자의 유전자 보존 정도가 높은 것이 관찰되었다. 더 나아가 두 미생물 모두에서 효소는 높은 상동성을 나타냈다. *c. ragsdalei* 및 *c. ljungdahlii*에서 아세틸 CoA 유전자의 서열을 비교하여 97.82%의 동일성을 갖는 것을 확인하였다.
- [0046] 더 나아가 아세틸 CoA 환원효를 암호화하는 유전자(프로모터 포함)의 기능을 테스트하였다. PCR을 통해 유전자 영역을 증폭하고 서플 백터인 pCOS52로 전달하여 EcoRI 부위로 연결하여 pCOS54를 형성하였다. 백터는 전체 아세틸-CoA 환원효소 유전자와 그 프로모터를 고복제수 플라스미드에 포함하고 있다. pCOS52는 pCOS54와 동일한 백본 백터를 포함하지만 AR 유전자는 없다. pCOS52는 *E. coli*의 유전자 발현을 결정하기 위한 기능 분석에서 대조 플라스미드로 사용되어 클로스트리듐의 유전자 기능을 확인하였다. 그 결과, 아세틸 CoA 환원효소 유전자의 기능이 확인되었다.
- [0047] 기능 분석은 기질로서 NADPH와 아세톤을 함유한 반응완충용액에 대해 주어진 시점에서 수확한 추가 세포로 구성하였다. 378nm에서 분광광도 활동(NADPH를 NADP+로 전환)을 측정하였으며 표준 곡선과 비교하여 총 활동 레벨을 결정하였다. 특이적 활동은 OD 측정치 1에서 317mg/g의 건조 세포 중량을 사용하여 결정하였다.
- [0048] PTA-ACK 오페론(유전자 ID 16-17, 표 1 및 3, 유전자 ID 15-16, 표 2) 및 그 프로모터를 암호화하는 유전자를 *C. ragsdalei*, *C. ljungdahlii* 및 *C. carboxydivorans*에서 서열결정하였다. 오페론의 기능을 확인하고 오페론의 하향 조절이 아세트산염에 대한 에탄올 생산 비율을 증가시키는 것이 입증되었다. 하향 조절은 부위 지시된 돌연변이를 통한 프로모터 변형을 거쳐 2-유전자 오페론 전사 발현을 감소시킴과 관련이 있다. 이와 같은 하향 조절은 mRNA 감소로 이어져 단백질 생성이 줄어들게 되고 이에 따라 균주의 아세트산염 생산 능력이 줄어들게 된다. 이와 같은 하향 조절은 실시예 2에 설명된 방법에 따라 이루어진다.
- [0049] 또한 초산형성 클로스트리듐 3개 종에 대하여 PTA-ACK 오페론에서 발견한 유전자 서열에 대한 비교를 수행하였다. PTA 유전자는 *c. ragsdalei*와 *c. ljungdahlii* 사이 97%의 동일성을 갖고 있었으며, *c. carboxydivorans*

와 *c. ragsdalei* 사이 78%의 동일성을, *c. ljungdahlii*와 *c. carboxydivorans* 사이 79%의 동일성을 갖고 있었다.

[0050] ACK 유전자는 *c. ragsdalei*와 *c. ljungdahlii* 사이 96%의 동일성을 갖고 있었으며, *c. carboxydivorans*와 *c. ragsdahlii* 사이 78%의 동일성을, *c. carboxydivorans*와 *c. ljungdahlii* 사이 77%의 동일성을 갖고 있었다.

[0051] *C. ragsdahlei*에서 에탄올의 생성을 촉진하기 위한 주요 유전자는 다음과 같다: 서열번호 1 (유전자 ID 번호 4, 55, 53, 표 1), 실험적으로 결정되는 프로모터 영역을 포함한, 일산화탄소 탈수소효소 *cooS*, 전자 전달 단백질 *cooF* 및 NADH 의존 페리독신 산화환원효소(FOR)에 대한 유전자 서열;

[0052] 서열번호 2 (유전자 ID 번호 17, 16, 표 1), 실험적으로 결정되는 프로모터 영역을 포함한, ACK 및 PTA에 대한 유전자 서열;

[0053] 서열번호 3 (유전자 ID 번호 6, 표 1), 실험적으로 결정되는 프로모터 영역을 포함한 아세틸 CoA 환원효소의 유전자 서열;

[0054] 서열 목록

[0055] *C. ragsdalei* 유전자 서열 (표 1)

[0056] >서열번호 1: (*cooS*, *cooF*, NADH:페리독신 산화환원효소 오페론(STOP 포함), 유전자 ID 번호 4, 55, 53)

```
TATTATATCAATATAGAATAATTTTCAATCAAATAAGAATTATTTTATATTTT
ATATTGACAAGGAAACCGAAAAGGTTTATATTATTGTTATTGGATAACAATT
ATTTTTTAGTTAGTTGTAAGTAAATAAATAGTATTAATTAATACTATTTAAA
CTATTACAGTTTTTTGATTCTTAGTATAAGTATTCTTAGTATCTTTAGCACTTAG
AATACGTTATCCTTTAGGAGAATAATCCTAATCAGTAATTTTAATAATTTAAT
AGTATACTTAAATAGTATAGTTTGGAGGTTTTATTATGTCAAATAACAAAATT
TGTAAGTCAGCAGATAAGGTAAGTCTTGGAAAAGTTTATAGGTTCTCTAGATGGTGT
AGAACTTCTCATCATAGGGTAGAAAAGCCAAAGTGTTAAATGTGGTTTTGGT
CAGCTAGGAGTCTGCTGTAGACTCTGTGCAAACGGTCCCTGCAGAATAACAC
CTAAAGCTCCAAGAGGAGTATGTGGTGCTAGTGCTGATACCATGGTTGCAAG
AACTTTCTTAGAGCTGTAGCTGCCGGCAGTGGATGTTATATCCATATAGTCG
AAAATACAGCTAGAAACGTAAAATCAGTAGGTGAAACCGGCGGAGAGATAA
AAGGAATGAATGCTCTCAACACCTAGCAGAAAAACTTGGTATAACAGAATC
TGACCCACATAAAAAAGCTGTACTAGTAGCTGTGCCGTATTAAGGACTTAT
ACAAACCAAAATTCGAAAAAATGGAAGTTATAAATAAATTAGCTTATGCACC
TAGACTAGAAAATTGGAACAAATTAATATAATGCCTGGCGGTGCAAAATCA
GAAGTTTTTGGATGGTGTAGTAAAACTTCTACAAATCTAAACAGCGACCCTGT
AGATATGCTTCTAAATTGTTTAAAACTTGGAAATATCCACTGGGATTTACGGAC
TTACCCTTACAAATTTATTAATGACATAATTTTAGGTGAACCTGCTATAAGA
CCTGCAAAAGTTGGTTTTAAAGTTGTAGATACGGATTATATAAATTTGATGAT
AACAGGCCACCAGCACTCCATGATTGCCACCTTCAAGAAGAAGCTTGTAAAA
```

[0057]

CCTGAAGCTGTAAAAAAGCCCAAGCAGTTGGTGCTAAAGGATTCAAAGTACTAG
 TTGGATGTACCTGTGTCGGACAGGATTTACAGTTAAGAGGTAAATACTATACT
 GATGTTTTCTCCGGTCATGCAGGAAATAACTTTACAAGTGAAGCCTTAATAGC
 AACTGGAGGTATAGATGCAATAGTATCTGAATTTAACTGTACTCTTCCCTGGCA
 TCGAGCCAATAGCTGATAAGTTCATGGTTAAAATGATATGCCTAGATGACGT
 TTCTAAAAAATCAAATGCAGAATATGTAGAATACTCTTTTAAAGATAGAGAA
 AAAATAAGCAACCATGTTATAGATACGGCTATTGAAAGTTATAAGGAAAGAA
 GATCTAAAGTTACAATGAATATTCCTAAAAACCATGGCTTTGATGACGTCATA
 ACAGGTGTAAGTGAAGGTTCCCTTAAAAATCCTTCTTAGGCGGAAGTTGGAAAC
 CTCTTGTAGACTTAATTGCTGCTGGAAAAATTAAGGTGTTGCTGGAATAGTA
 GGTTGTTCAAACCTAACTGCCAAAGGTCACGATGTATTTACAGTAGAACTTAC
 AAAAGAAGTCTATAAAGAGAAATATAATTGTAAGTTCTGCAAGTTGTTCAAGT
 GGTGGACTTGAAAATGTAGGACTTATGTCTCCAGGAGCTGCTGAACTTGCAAG
 GAGATAGCTTAAAAGAAGTATGTAAGAGCCTAGGTATACCACCTGTAAGTAAA
 TTTTGGTCCATGTCTTGTCTATTGGAAGATTGGAATTTGTAGCAAAAGAAGTAA
 CAGAATACCTAAAAATAGATATTCACAGCTTCCACTTGTGCTTTCTGCACCT
 CAATGGCTTGAAGAACAAGCATTGGCAGATGGAAGTTTTGGTCTTGCCCTTG
 GATTACCATTACCTTGCTATATCTCTTTTATTGGTGGAAAGCAAAGTGGTA
 ACAAAAAGTTTTATGTGAAGATATGGAAAATCTAACAGGGCGCAAGCTTATAA
 TAGAAGACGATGTAATAAAAGCTGCAGATAAATTAGAAGAAACCATACTTGC
 AAGAAGGAAAAGCTTAGGTCTTAATTAATGAAAAGAATAATGATAAATAA
 GGATTTATGTACCGGATGCTTAAATTGTAAGTTTACTTTAGCTTGTATGGCAGAACA
 ATGAAAATGGGAAATCTTTTTATGATCTGGATCTCAGCAATAAATTTCTTGAA
 AGTAGAAATCATATATCTAAAGATGATAATGGAAACAAGCTTCCCTATATTTT
 GCCGTCAGTGTGACGAACCTGAGTGCCTAATGACATGTATGAGCGGTGCCAT
 GACTAAAGATCCTGAAACTGGTATAGTATCCTATGATGAGCATAAATGTGCC
 AGCTGCTTTATGTGCGTCATGTCTGCTTATGGAGTATTGAAACCAGATAC
 TCAGACCAAAAGTAAAGTAGTTAAATGTGACCTGTGTGGTGACAGAGATACA
 CCTAGATGCGTTGAAAATTGTCCAACAGAAGCAATTTATATTGAAAAGGAGG
 CAGATCTCCTATGAATGAGTGGTTTAAACAATAAAAAATATTTTTTACACAAAA
 TATGTAATAATAGGAGCCAGTGTCTGCTGGAATAAATGCTGCTAAAAGTTTAA
 GAAAGTTAGATAAATCCTCCAAAATAACTATTATTTCAAAGGATGATGCAGT
 TTATTTCAAGATGTATACTCCACAAAGTACTTGAGGGAAGTAGAAATTTAGAT
 ACCATAAATTTTGTAGATTCTGATTTCTTTGAAAAAATAATATAGAATGGAT
 AAAAGATGCAGATGTAAGCAATATTGATATTGACAAGAAAAAAGTCTTACTT
 CAAGACAACAGCAGCTTCAAATTTGACAAGCTCCTTATAGCTTCTGGTGCTTC
 CTCCTTTATCCCCCAGTTAAAAAATTAAGAGAAGCTAAAGGAGTGTACTCCC
 TTAGAAATTTTGAAGATGTAAGTCTGCTATACAAGACAACTTAAAAACGCAAA
 ACAAGTGGTAATACTTGGTGCAGGTCTTGTAGGAATTGATGCACTTTTAGGTC
 TTATGGTGAAAAATATAAAGATTTTCAAGTTGTAGAAATGGGAGATAGGATTCT
 CCCCCTTCAACTGGACAAAAGTGCATCCACTATATATGAAAAGTTGTTAAAA
 GAAAAAGGTATAGATGTCTTTACTTTCAGTTAAATTGGAAGAGGTAGTTTTAA
 ATAAAGACGGAAGTAAAGTAAAGCAGTACTATCAAATTTCAACTTCTATAGA
 TTGCGATATGATAATAGTTGCTGCTGGTGTAGACCAAAATGTAAGCTTTATAA
 AAGACAGCAGGATAAAAGTTGAAAAGGCATTGTCATAGACAAACATTGTA
 AAACCACTGTAGATAATATATATGCTGCAGGAGATGTTACTTTTACTGCTCCT

[0058]

ATATGGCCTATAGCTGTAAAGCAGGGAATAACTGCTGCTTACAACATGGTAG
 GTATAAATAGAGAATTACATGACACTTTTGGCATGAAGAAGTCAATGAATTT
 ATTTAACCTTCCATGCGTATCCCTTGGTAATGTAATATAGCAGATGAAAGTT
 ATGCTGTTGATACATTAGAAGGAGATGGAGTTTATCAAAAAATAGTTCACAA
 AGATGGAGTAATCTACGGTGCCTTCTAGTTGGAGATATATCTTACTGCGGCG
 TACTAGGATATCTCATAAAAAATAAAGTAAATATAAGCAATATCCATAAAAA
 TATTTTTGACATAGATTATTCTGATTTTTACAATGTTGAAGAAGATGGACAAT
 ATAGTTATCAATTGAGGTAA

[0059]

[0060] 서열번호 2: (PTA-ACK 오페론(STOP 포함), 유전자 ID 번호 17, 16)

GCATACTGATTGATTATTTATTTGAAAATGCCTAAGTAAAATATATACATATT
 ATAACAATAAAATAAGTATTAGTGTAGGATTTTTAAATAGAGTATCTATTTTC
 AGATTAAATTTTTACTTATTTGATTTACATTGTATAATATTGAGTAAAGTATTG
 ACTAGTAAAATTTTGTGATACTTTAATCTGTGAAATTTCTTAGCAAAAGTTAT
 ATTTTTGAATAATTTTTATTGAAAAATACAACATAAAAAGGATTATAGTATAAG
 TGTGTGTAATTTTGTGTTAAATTTAAAGGGAGGAAATAAACATGAAATTGAT
 GGAAAAAATTTGGAATAAGGCAAAGGAGACAAAAAAAAGATTGTCTTAGC
 TGAAGGAGAAGAAGAAAGAACTCTTCAAGCTTGTGAAAAAATAATTAAAGA
 AGGTATTGCAAATTTAATCCTTGTAGGGAATGAAAAGGTAATAGAGGAGAAG
 GCATCAAATAGGCGTAAGTTTAAATGGAGCAGAAATAGTAGATCCAGAAA
 CCTCGGATAAACTAAAAAATATGCAGATGCTTTTTATGAATTGAGAAAGAA
 GAAGGGAATAACACCAGAAAAAGCGGATAAAATAGTAAGAGATCCAATATA
 TTTTGCTACGATGATGGTTAAGCTTGGAGATGCAGATGGATTGGTTTCAGGTG
 CAGTGCATACTACAGGTGATCTTTTGAGACCAGGACTTCAAATAGTAAAGAC
 AGCTCCAGGTACATCAGTAGTTTCCAGCACATTTATAATGGAAGTACCAAATT
 GTGAATATGGTGACAATGGTGTACTTCTATTTGCTGATTGTGCTGTAATCCA
 TGCCCAGATAGTGATCAATTGGCTTCAATTGCAATAAGTACAGCAGAACTG
 CAAAGAACTTATGTGGAATGGATCCAAAAGTAGCAATGCTTTCATTTTCTACT
 AAGGGAAGTGCAAAACACGAATTAGTAGATAAAGTTAGAAATGCTGTAGAA
 ATTGCCAAAAAGCTAAACCAGATTTAAGTTTGGACGGAGAATTACAATTAG
 ATGCCTCTATCGTAGAAAAGGTTGCAAGTTTAAAGGCTCCTGAAAGTGAAGT
 AGCAGGAAAAGCAAATGTACTTGTATTTCCAGATCTCCAAGCAGGAAATATA
 GGTATAAACTTGTTCAAAGATTTGCAAAAGCTGATGCTATAGGACCTGTATG
 CCAGGGATTTGCAAACCTATAAATGATTTGTCAAGAGGATGTAACCTCCGAT
 GATATAGTAAATGTAGTAGCTGTAACAGCAGTTCAGGCACAAGCTCAAAAGT
 AAATGAAAATATTAGTAGTAAACTGTGGAAGTTCATCTTTAAAAATATCAACTT
 ATTGATATGAAAGATGAAAGCGTTGTGGCAAAAGGACTTGTAGAAAGAATA
 GGAGCAGAAGGTTCAAGTTTAAACACATAAAGTTAACGGAGAAAAGTTTGTTA
 CAGAGCAGCCAATGGAAGATCATAAAGTTGCTATACAATTAGTATTAATGC
 TCTTGTAGATAAAAAACATGGTGTATAAAGATATGTCAGAAATATCTGCT
 GTAGGGCATAAGATTTTGCATGGTGGAAAAAATATGCGGCATCCATTCTTA
 TTGATGACAATGTAATGAAAGCAATAGAAGAATGTATTCCATTAGGACCATT
 ACATAATCCAGCTAATATAATGGGAATAGATGCTTGTAAAAAACTAATGCCA
 AATACTCCAATGGTAGCAGTATTTGATACAGCATTTTCATCAGACAATGCCAG
 ATTATGCTTATACTTATGCAATACCTTATGATATATCTGAAAAGTATGATATC
 AGAAAATATGGTTTTTCATGGAACCTTTCATAGATTTCGTTTCAATTGAAGCAGC

[0061]

CAAGTTGTTAAAGAAAGATCCAAAAGATCTTAAGCTAATAACTTGTCATTTA
 GGAAATGGAGCTAGTATATGTGCAGTAAACCAGGGAAAAGCAGTAGATACA
 ACTATGGGACTTACTCCCCTTGCAGGACTTGTAAATGGGAACTAGATGTGGTG
 ATATAGATCCAGCTATAATACCATTTGTAATGAAAAGAACAGGTATGTCTGT
 AGATGAAATGGATACTTTAATGAACAAAAAGTCAGGAATACTTGGAGTATCA
 GGAGTAAGCAGCGATTTTAGAGATGTAGAAGAAGCTGCAAATTCAGGAAAT
 GATAGAGCAAAACTTGCATTAATATGTATTATCACAAAGTTAAATCTTTCAT
 AGGAGCTTATGTTGCAGTTTTAAATGGAGCAGATGCTATAATATTTACAGCA
 GGACTTGGAGAAAATTCAGCTACTAGCAGATCTGCTATATGTAAGGGATTAA
 GCTATTTTGGAAATTAATAGATGAAGAAAAGAATAAGAAAAGGGGAGAAG
 CACTAGAAATAAGCACACCTGATTCAAAGATAAAAGTATTAGTAATTCCTAC
 AAATGAAGAACTTATGATAGCTAGGGATACAAAAGAAATAGTTGAAAATAA
 ATAA

[0062]

[0063] 서열번호 3: (ORF RCCC02715, P11, NADPH-SADH (STOP 포함), 유전자 ID 번호 6)

```

ATGAAAGGTTTTGCAATGTTAGGTATTAACAAGTTAGGATGGATTGAAAAGA
AAAACCCAGTACCAGGTCCTTATGATGCGATTGTACATCCTCTAGCTGTATCC
CCATGTACATCAGATATACATACGGTTTTTTGAAGGAGCACTTGGTAATAGGG
AAAATATGATTTTAGGTCACGAAGCTGTAGGTGAAATAGCTGAAGTTGGCAG
TGAAGTTAAAGATTTTAAAGTTGGCGATAGAGTTATCGTACCATGCACAACA
CCTGACTGGAGATCCTTAGAAGTCCAAGCTGGTTTTTCAACAGCATTCAAACG
GTATGCTTGCAGGATGGAAGTTTTCCAATTTTAAAGACGGTGTATTTGCAGAT
TACTTTTACGTAAACGATGCAGATATGAATCTTGAATACTTCCAGATGAAAT
ACCTTTAGAAAGTGCAGTTATGATGACAGACATGATGACTACTGGTTTTTCATG
GGGCAGAACTTGCTGACATAAAAATGGGTTCCAGTGTGTCGTAATTGGTAT
AGGAGCTGTTGGATTAATGGGAATAGCCGGTTCCAAACTTCGAGGAGCAGGT
AGAATTATCGGTGTTGGAAGCAGACCCGTTTGTGTTGAAACAGCTAAATTTTA
TGGAGCAACTGATATTGTAATTATAAAAAATGGTGATATAGTTGAACAAATA
ATGGACTTAACTCATGGTAAAGGTGTAGACCGTGAATCATGGCAGGCGGTG
GTGCTGAAACACTAGCACAAGCAGTAACTATGGTTAAACCTGGCGGCGTAAT
TTCTAACATCAACTACCATGGAAGCGGTGATACTTTGCCAATACCTCGTGTTT
AATGGGGCTGCGGCATGGCTCACAAAACCTATAAGAGGAGGGTTATGTCCCGG
CGGACGTCTTAGAATGGAAATGCTAAGAGACCTTGTTCTATATAAACGTGTT
GATTTGAGCAAACCTTGTTACTCATGTATTTGATGGTGCAGAAAATATTGAAAA
GGCCCTTTTGCTTATGAAAAATAAGCCAAAAGATTTAATTAATCAGTAGTTA
CATTCTAA
    
```

[0064]

[0065] 상세한 게놈 정보를 사용하여 초산형성 클로스트리듐 미생물을 대사적으로 엔지니어링하고 이와 동시에 본원 발명에 따라 해당하는 아세트산염 및 부티르산염 형성 경로를 통한 탄소 및 전자 플럭스를 줄이거나 제거하면 서, 에탄올 및 부탄올의 생합성 경로를 통한 탄소 및 전자 플럭스를 증가시켰다. 이러한 목적을 위하여 경로 내 효소를 암호화하는 주요 유전자 활성을 조절하였다. 한 가지 양태에서 주요 알코올 생성 유전자 발현이 유전자 복제수 증가에 의해 증가하였다. 예를 들어 주요 일산화탄소 탈수소효소 오페론(도 3), 및 아세틸 CoA 환원효소와 알데히드 페리독신 산화환원효소를 포함한 관련된 전자 전달 단백질을 변형된 유기체의 게놈 안에서 복제하였다. 한 가지 양태에서 이들 복제물을 아세트산염 생산을 못하게 만들거나 감소시킨 균주 안에 넣어 에탄올 또는 부탄올 생성 경로로 전자가 들어갈 수 있도록 하였다. 또 다른 양태에서, 합성가스에서 성장했을 때 에탄올, 부탄올, 헥산을 및 해당하는 카르복실산 등 알코올과 산의 더 많은 복잡한 화합물을 생성하는 초산형성 클로스트리듐 균주에 녹아웃 전략을 적용하였다.

[0066]

한 가지 양태에서 클로닝 비히클로부터 초산형성 클로스트리듐 모균주로 초산형성 클로스트리듐 클로닝된 유전자를 전달하는 데 사용하게 되는 벡터는 표준 방법을 사용하여 제작하였다 (Sambrook 외, 1989). 서열 특이 프라이머를 이용한 복제의 신뢰성이 높은 증합효소 연쇄반응(PCR) 기술을 사용하여 분자 유전학 실험에 사용된 모든 표적 유전자를 증폭하였다. 유전자는 증폭된 다음 매개 복제 운반체로 하위 복제하고 후에 다중 성분 연결 반응으로 재조합하여 유전자 전달 실험에 사용될 원하는 재조합 벡터를 생산하였다. 벡터는 유전자 전달 실험을 성공적으로 수행하는 데 필요한 적절한 기능들을 갖추고 있으며 사용되는 방법에 따라 다양하다. 재조합 벡터를 수용자 초산형성 클로스트리듐에 전달하기 위해 다양한 방법이 사용된다. 여기에는 전기천공법, 양친 접합(bi-parental conjugation) 또는 삼자 접합(tri-parental conjugation), 리포솜 매개 형질전환 및 폴리에틸렌글리콜 매개 형질전환이 포함된다. 재조합 초산형성 클로스트리듐이 분리되어 DNA 통합시 얻은 특이 형질의 획득을 기반으로 한 분자 생물학 기술을 통해 확인된다.

도면의 간단한 설명

[0067] 도 1은 프로세스에 참여한 주요 효소 중 일부를 포함한 초산형성 클로스트리듐 내 합성가스 발효 동안 전자 흐름 경로를 묘사한 다이어그램이다.

도 2는 아세틸 CoA 생성 및 아세틸-CoA를 아세트산염과 에탄올로 효소 전환하기 위한 Wood-Ljungdahl (C₁) 경로를 설명하는 다이어그램이다.

도 3은 본원 발명에 따른 cooS, cooF 및 페리독신 산화환원효소(FOR)를 포함한 일산화탄소 탈수소효소(CODH) 오페론 중 하나의 위치를 보여주는 유전자 지도를 기술한 다이어그램이다.

도 4는 본 발명에 따라 C. ragsdalei [서열번호 4], C. ljungdahlii [서열번호 5] 및 Thermoanaerobactor

ethanolicus [서열번호 6]에서 NADPH 의존 2차 알코올 탈수소효소의 계층 구조의 배열을 보여주는 다이어그램이다.

도 5는 에탄올 합성을 위한 Wood-Ljungdahl 경로를 기술하고 아세테이트키나제(ack)와 인산아세틸기전이효소(pta) 암호화 유전자가 작용하지 못하게 만들거나 발명에 따른 인산아세틸기전이효소 및 아세테이트키나제 유전자 발현을 촉발하는 프로모터를 돌연변이 유발시키거나 교체시켜 아세트산염 생성을 조절함으로써 초산형성 클로스트리듐에서 아세트산염 생성을 특별히 감소시키거나 제거하는 전략을 보여 주는 다이어그램이다.

도 6은 에탄올 합성을 위한 Wood-Ljungdahl 경로에 대한 다이어그램이며 본 발명에 따른 아세테이트키나제 또는 인산아세틸기전이효소가 작용하지 못하게 만든 숙주에서 아세틸 CoA 환원효소의 과다발현에 의해 C. ragsdalei에서 에탄올 생산을 특별히 증가시키는 전략을 보여 준다.

도 7은 에탄올 합성을 위한 Wood-Ljungdahl 경로의 다이어그램이며, 본 발명에 따른, 아세트산염 생성 능력이 약화되고 NADPH-의존 알코올 탈수소효소 활성이 증가된 숙주 균주에서 알데히드 페리독신 산화환원효소(AOR)에 의해 초산생성 클로스트리듐에서 에탄올 생산을 증가시키는 전략을 보여준다.

도 8은, 본 발명에 따른, 부탄올 생성을 증가시키기 위한 전략을 보여주는 C. carboxidivorans에서의 부탄올 및 부티레이트 생합성 경로 및 아세틸-CoA의 부탄올 및 부티레이트로의 전환을 촉매화하는 해당 유전자의 다이어그램이다.

발명을 실시하기 위한 구체적인 내용

- [0068] 실시예 1
- [0069] 초산형성 클로스트리듐은 도 3에서 볼 수 있듯이 일산화탄소 탈수소효소 104(cooS, 유전자 ID 4, 표 1, 표 2, 표 3), 막관련 전자 전달 단백질(cooF) 및 페리독신 산화환원효소(FOR)로 이루어진 오픈론 300을 포함한다. 초산형성 클로스트리듐 내 일산화탄소 탈수소효소 104의 과다발현은 합성가스 성분으로부터 산화된 뉴클레오타이드 보조인자 NAD^+ 및 $NADP^+$ 로의 전자 흐름을 증가시키는 것으로 알려졌다. 환원된 뉴클레오타이드 보조인자의 증가된 레벨은 Wood-Ljungdahl 경로 100에서 중간 화합물의 생성을 자극한다.
- [0070] 한 가지 양태에서 오픈론 300을 일산화탄소 탈수소효소 유전자의 영역 200 뉴클레오타이드(nt) 상류와 페리독신 산화환원효소 유전자의 200nt 하류에 어닐링하도록 설계한 프라이머를 사용하는 롱(long)-PCR 기술을 통해 증폭하였다. 총 영역은 약 3.8 키로베이스 페어이다. 증폭된 DNA는 pGEM T easy (Promega, Madison, WI) 또는 pTOPO (Invitrogen, Carlsbad, CA) 등 PCR 산물을 연결하도록 특별히 설계된 적절한 플라스미드 벡터 안으로 직접 클로닝된다. PCR 산물 말단에는 엔지니어링된 제한 부위가 있어 후에 클로닝 단계를 쉽게 진행할 수 있다. 오픈론 300은 이미 클로닝된 염색체 C. ragsdalei 또는 기타 초산형성 클로스트리듐 DNA를 함유하는 벡터로 서브클로닝되어 중간 부위에서 염색체 통합을 수행할 수 있다.
- [0071] 실시예 2
- [0072] 카르복실산은 알코올과 전자를 뺏으려 경쟁하기 때문에 산의 생성을 감소시키면 더 많은 전자가 알코올 생산 경로로 흘러들어갈 CoA 중간 생성물에서 직접 알코올로 들어갈 수 있다. 초산형성 클로스트리듐은 아세틸 CoA를 아세틸 인산염으로 전환하는 인-아세틸기전이효소의 유전자(유전자 ID 17, 표 1 및 3, 유전자 ID 16, 표 2)와 아세틸 인산염 218을 아세트산염 214로 전환하는 아세테이트키나제(유전자 ID 16, 표 1)를 함유하고 있다. 한 가지 양태에서 양 효소의 전부 또는 일부를 제거하는 유전자 변형과 아세트산염의 생산을 못하게 하거나 감소시키기 위한 유전자 변형을 도 5에서와 같이 수행하였다.
- [0073] PCR 및 기타 표준 방법을 사용하여 두 가지 비인접한 DNA 조각을 포함한 재조합 벡터를 생성하였다. 재조합 벡터 유전자로 원래의 유전자를 교체한 경우 클로스트리듐 균주는 각각 도 5의 X 504와 X 502에서 보듯이 인산아세틸기전이효소 또는 아세테이트키나제의 활성을 포함하지 않는다.
- [0074] 인산아세틸기전이효소 508 및 아세테이트키나제 510의 유전자 발현 및 이후의 아세트산염 생산을 감소시키기 위하여 프로모터 영역을 변형시킨 일련의 재조합 벡터를 생성함으로써 공통 프로모터 영역, P* 506의 변형을 수행하였다. 부위 지시된 돌연변이를 통해 벡터 시리즈를 제작하였다.
- [0075] 또한 프로모터 영역의 부위 지시된 돌연변이를 통해 pta/ack 유전자를 포함하는 2 유전자 오픈론의 하향 조절을 수행하였다. RNA 폴리메라아제 결합의 감소는 pta/ack 프로모터의 전사 활성 감소로 이어지고 다시 단백질 활성의 감소로 이어졌다. 최종 결과는 더 낮은 속도로 중간 생성물이 생성된 이래 아세트산염의 생산이 감소

하였고 아세틸 CoA로부터 더 많은 탄소가 에탄올 생성에 사용되었다. 쉽게 정량화할 수 있는 리포터 그룹을 이용한 프로모터 프로브 분석을 개발하여 생체내에서 프로모터의 상대적인 프로모터 강도를 측정하였다. 200 개의 기본 쌍 영역에 대해 단일 손상부와 다중 손상부를 부여하는 부위 지시 돌연변이를 수행한 후 프로모터의 활성이 감소한 균주를 분리하여 분석 시 원래 프로모터의 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10% 및 0% 활성을 갖는 일련의 균주를 재조합 클로스트리듐 균주에서 분리하여 테스트하였다.

[0076] 실시예 3

[0077] 생체내에서, 102 및 도 5에서 지정된 아세틸 CoA 효소는 아세틸-CoA 102 또는 부틸-CoA 등 탄소 모이어티의 조효소 A(CoA) 형태를 직접 해당하는 알코올로 변환한다. 열역학적으로 CoA 형태에서 알코올로 직접 변환하려면 4개 전자의 전달이 필요하며 이것은 카르복실산을 해당 알코올로 2단계 변환하는 것과 비교하여 알코올을 생성하는 더 효율적인 방법이다. 예를 들어 도 6에서 볼 수 있듯이 2단계 변환은 아세트산염 214를 먼저 알데히드 형태(아세트알데히드, 604)로 변환한 다음 해당 알코올인 에탄올 216으로 변환해야 한다. 따라서 알코올 생산 증가를 위해, 또한 산에 대한 알코올 비율을 증가시켜 프로세스의 선택성을 높이기 위해 수직 화살표 602로 묘사한 AR의 활성을 증가시키는 것이 바람직하다.

[0078] 한 가지 양태에서 복제의 신뢰성이 높은 중합효소 연쇄반응(PCR) 기술을 사용하여 시험관내 유전자를 증폭하고 표준 분자 유전학 기술을 사용하여 염색체 내 중간 부위에 클로닝된 유전자를 삽입함으로써 초산형성 클로스트리듐 내 AR의 활성이 증가하였다. 박테리의 유전자를 교체한 후 염색체는 AR의 복제물 두 개를 포함하였다. 유전자 교체를 확인한 후 재조합 균주에 대한 유전자 발현 연구를 수행하여 모균주와 비교하였다.

[0079] 다른 양태에서 유사한 전략을 사용하여 adhE 타입 알코올 탈수소효소, 짧은 사슬 알코올-탈수소효소 및 1차 철 함유 알코올 탈수소효소의 활성을 증가시켰다.

[0080] 실시예 4

[0081] 어떠한 조건 하에서 클로스트리듐은 아테노신 삼인산염 생성물(ATP) 형태의 추가 에너지를 얻어 세포에서 아세틸-CoA 102로부터 아세트산염 214의 생산을 일시적으로 증가시켜야 한다. 순 반응은 아세틸-인산염을 통해 ADP + P_i로부터 1 ATP를 얻는다. 아세트산염 생성물은 아세트산 농도가 낮음에서 중간인 합성가스 발효 프로세스에 비해 유리하다. 세포에서 더 많은 에너지를 생성할 수 있게 하고 안정된 상태를 유지할 수 있게 하기 때문이다. 그러나 유리 아세트산이 너무 많을 경우 1차 ATP 생성원으로 사용되는 막관통 이온 구배가 분산되어 세포에 해를 끼치게 된다. 산업용도의 생산을 위하여는 아세트산염을 에탄올로 변환하여 에탄올 생산을 증가시키고 너무 많은 유리 아세트산이 축적될 가능성을 줄이는 것이 유리하다.

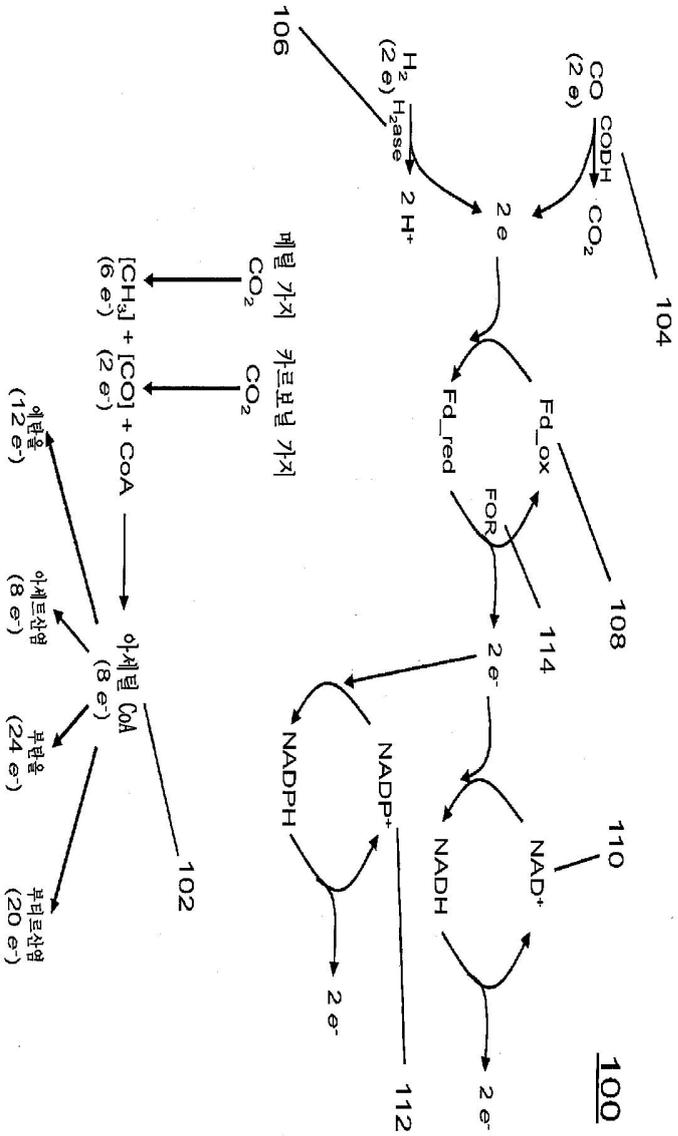
[0082] 한 가지 양태에서 알데히드 페리독신 산화환원효소와 AR 활성을 증가시킨 결과 이중 돌연변이 C. ragsdalei 균주 내 에탄올 생산은 10% ~ 40% 사이로 증가하였다.

[0083] 또 다른 양태에서 아세트산염에 대한 에탄올 생성 비율이 5~10배로 증가하였다. 그러나 미생물의 에너지 요구량을 충족시키기 위해 필요한 ATP 생산을 지원할 만큼 충분한 아세트산염이 형성되었다.

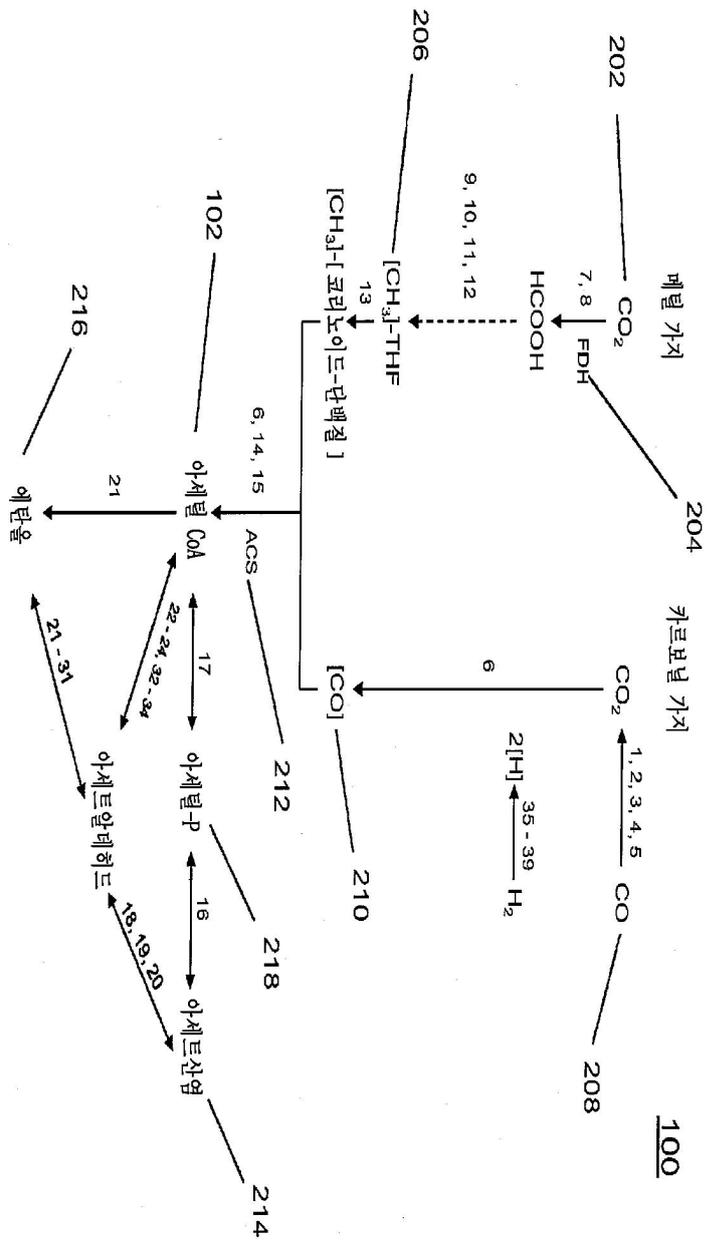
[0084] 본 발명은 특정 양태를 인용하여 설명한 반면, 발명의 정신과 범위에서 벗어나지 않은 범위에서 형태 및 세부 사항에서 변형과 다양과 변경이 이루어질 수 있음을 당업자는 이해할 것이다.

도면

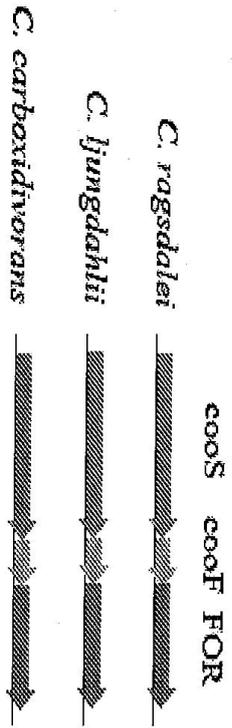
도면1



도면2

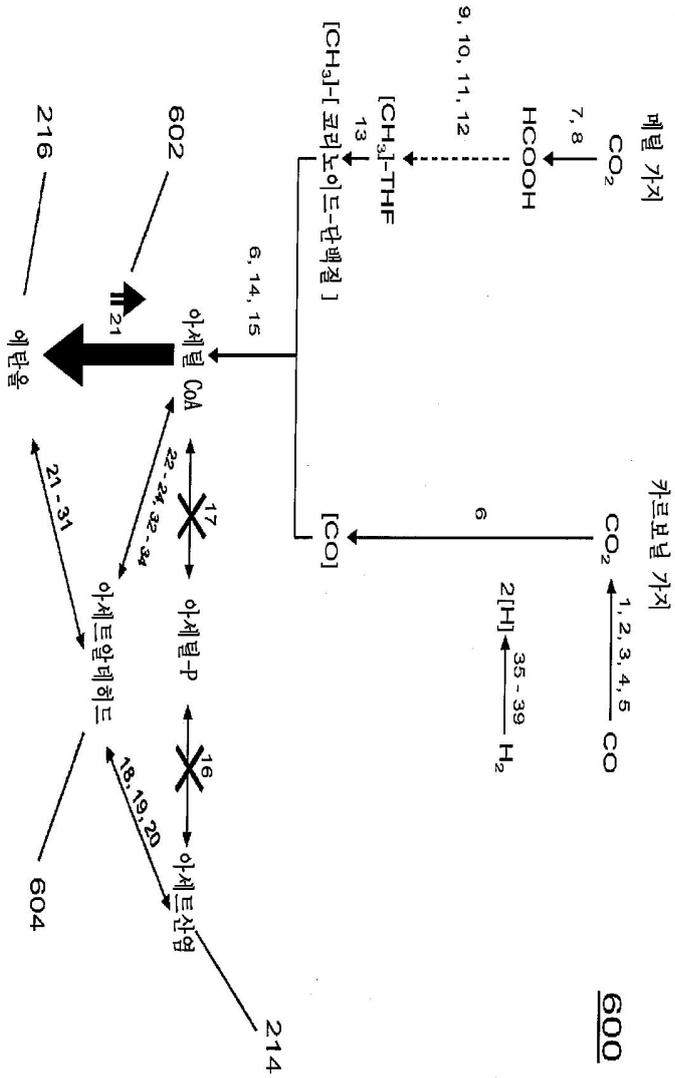


도면3

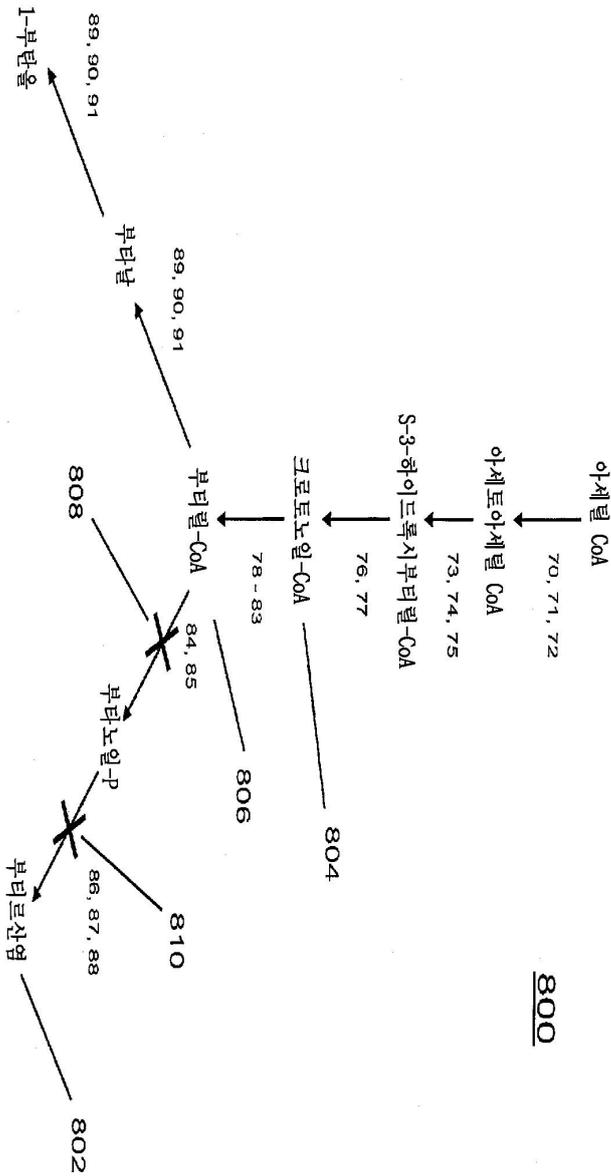


300

도면6



도면8



서열 목록

SEQUENCE LISTING

<110> COSKATA, INC.

<120> CLONING AND EXPRESSION OF THE GENES ENCODING KEY CLOSTRIDIAL CATALYZING MECHANISMS FOR SYNGAS TO ETHANOL PRODUCTION AND FUNCTIONAL CHARACTERIZATION THEREOF

<130> 11-214-US-DIV

<150> US 12/802,560

<151> 2010-06-09

<160> 6

<170> KopatentIn 1.71

<210> 1

<211> 3899

<212> DNA

<213> Clostridium ragsdalei

<400> 1

tattatatca atatagaata attttcaatc aaataagaat tttttatat tttatattga	60
caaggaaacc gaaaaggttt atattattgt tattggataa caattatfff ttagttagtt	120
gtacttgtaa ataaatagta ttaattaata ctattaaact attacagttt ttgattctta	180
gtataagtat tcttagtatac tttagcactt agaatacgtt atcctttagg agaataatcc	240
taatcagtaa ttttaataat ttaatagtat acttaaatag tatagtttgg aggttttatt	300
atgicaaata acaaaatftg taagtacgca gataaggtag ttgaaaagtt tataggttct	360
ctagatggtg tagaaacttc tcatcatagg gtagaaagcc aaagtgttaa atgtggtttt	420
ggtcagctag gagtctgctg tagactctgt gcaaacggtc cctgcagaat aacacctaaa	480
gctccaagag gagtatgtgg tgctagtct gataccatgg ttgcaagaaa ctttcttaga	540
gctgtagctg cggcagtggt atgttatatc catatagtcg aaaatacagc tagaaacgta	600
aaatcagtag gtgaaaccgg cggagagata aaaggaatga atgctctcaa caccctagca	660
gaaaaacttg gtataacaga atctgaccca cataaaaaag ctgtactagt agctgtgccg	720
tattaaagga cttatacaaa ccaaaatctg aaaaaatgga agttataaat aaattagctt	780
atgcacctag actagaaaat tggaaacaat taaatataat gcctggcgggt gcaaaatcag	840
aagtttttga tgggtgtagta aaaacttcta caaatctaaa cagcgacct gtatatatgc	900
ttctaaattg tttaaaactt ggaatatcca ctgggattta cggacttacc cttacaaatt	960
tattaaatga cataatftta ggtgaacctg ctataagacc tgcaaaagtt ggttttaag	1020
ttgtagatag ggattatata aatttgaatga taacaggcca ccagcactcc atgattgccc	1080
acctcaaga agaacttgta aaacctgaag ctgtaaaaaa agcccaagca gttggtgcta	1140
aaggattcaa actagttgga tgtacctgtg tcggacagga tttacagtta agaggtaaat	1200
actatactga tgttttctcc ggtcatgcag gaaataactt tacaagttaa gccttaatag	1260
caactggagg tatagatgca atagtatctg aatttaactg tactcttctt ggcatcgagc	1320
caatagctga taagttcatg gttaaaatga tatgcctaga tgacgtttct aaaaaatcaa	1380
atgcagaata tgtagaatac tcttttaag atagagaaaa aataagcaac catgttatag	1440
atacggctat tgaaagttat aaggaaagaa gatctaaagt tacaatgaat attcttaaaa	1500
accatggctt tgatgacgtc ataacaggtg taagtgaagg ttctttaaaa tcttcttag	1560

gcggaagtg gaaacctctt gtagacttaa ttgctgctgg aaaaattaaa ggtgtgctg 1620
 gaatagtagg ttgttcaaac ttaactgcc aaggtcacga tgtatttaca gtagaactta 1680
 caaaagaact cataaagaga aatataattg tactttctgc aggttgttca agtggtagac 1740
 ttgaaaatgt aggacttatg tctccaggag ctgctgaact tgcaggagat agcttaaaag 1800

aagtatgtaa gagcctaggt ataccacctg tactaaatth ttgtccatgt cttgctattg 1860
 gaagattgga aattgtagca aaagaactag cagaatacct aaaaatagat attccacagc 1920
 ttccacttgt gctttctgca cctcaatggc ttgaagaaca agcattggca gatggaagtt 1980
 ttggtcttgc ccttggatta ccaacttacc ttgctatata tcctttcatt ggtggaagca 2040
 aagtggtaac aaaagtttta tgtgaagata tggaaaatct aacagggcggc aagcttataa 2100
 tagaagacga tgtaataaaa gctgcagata aattagaaga aaccatactt gcaagaagga 2160
 aaagcttagg tcttaattaa atgaaaagaa taatgataaa taaggattta tgtaccggat 2220

gcttaaatg tactttagct tgtatggcag aacacaatga aaatgggaaa tctttttatg 2280
 atctggatct cagcaataaa tttcttgaag gtagaatca tatactaaa gatgataatg 2340
 gaaacaagct tcctatattt tgccgtcact gtgacgaacc tgagtgcgta atgacatgta 2400
 tgagcgggtc catgactaaa gatcctgaaa ctggatatag atcctatgat gagcataaat 2460
 gtgccagctg ctttatgtgc gtcatgtcct gtccctatgg agtattgaaa ccagatactc 2520
 agaccaaaag taaagtagtt aaatgtgacc tgtgtggtga cagagataca cctagatgcg 2580
 ttgaaaattg tccaacagaa gcaatttata ttgaaaagga ggcagatctc ctatgaatga 2640

gtggtttaac aataaaaata tttttcaca caaaatatgt aataatagga gccagtgctg 2700
 ctggaataaa tgctgctaaa actttaagaa agttagataa atcctccaaa ataactatta 2760
 tttcaaagga tgatgcagtt tattcaagat gtatactcca caaagtactt gagggaagta 2820
 gaaatttaga taccataaat tttgtagatt ctgatttctt tgaaaaaat aatatagaat 2880
 ggataaaaaga tgcagatgta agcaatattg atattgacaa gaaaaaagtc ttacttcaag 2940
 acaacagcag cttcaaattt gacaagctcc ttatagcttc tgggtgcttcc tcctttattc 3000
 ccccagttaa aaaattaaga gaagctaaag gagtgtactc ccttagaat tttgaagatg 3060

taactgctat acaagacaaa cttaaaaacg caaaacaagt ggtaatactt ggtgcaggtc 3120
 ttgtaggaat tgatgcactt ttaggtctta tggtgaaaa tataaagatt tcagtgttag 3180
 aaatgggaga taggatctc ccccttcaac tggacaaaac tgcattccact atatatgaaa 3240
 agttgttaaa agaaaaaggt atagatgtct ttacttcagt taaattgaa gaggtagttt 3300
 taaataaaga cggaactgta agtaaagcag tactatcaaa ttcaacttct atagattgcg 3360
 atatgataat agttgctgct ggtgttagac caaatgtaag ctttataaaa gacagcagga 3420

taaaagttga aaaaggcatt gtcatagaca aacattgtaa aaccactgta gataatatat 3480

atgctgcagg agatgttact ttactgctc ctatatggcc tatagctgta aagcaggaa 3540

taactgctgc ttacaacatg gtaggtataa atagagaatt acatgacact tttggcatga 3600

agaactcaat gaatttatti aaccttccat gcgtatccct tggtaatgta aatatagcag 3660

atgaaagtta tgctgttgat acattagaag gagatggagt ttatcaaaaa atagttcaca 3720

aagatggagt aatctacggg gcacttctag ttggagatat atcttactgc ggcgtactag 3780

gatatctcat aaaaaataaa gtaaatataa gcaatatcca taaaaatatt tttgacatag 3840

attattctga tttttacaat gttgaagaag atggacaata tagttatcaa ttgaggtaa 3899

<210> 2

<211> 2506

<212> DNA

<213> Clostridium ragsdalei

<400> 2

gcatactgat tgattattta ttgaaaatg cctaagtaaa atatatacat attataacaa 60

taaaataagt attagtgtag gatttttaaa tagagtatct attttcagat taaattttta 120

cttatttgat ttacattgta taatattgag taaagtattg actagtaaaa ttttgtgata 180

ctttaatctg tgaattttct tagcaaaagt tataattttg aataattttt attgaaaaat 240

acaactaaaa aggattatag tataagtgtg tgtaattttg tgttaaattt aaaggaggga 300

aataaacatg aaattgatgg aaaaaattg gaataaggca aaggaagaca aaaaaagat 360

tgtcttagct gaaggagaag aagaaagaac tcttcaagct tgtgaaaaaa taattaaaga 420

aggatttgca aatttaatec ttgtagggaa tgaaaaggta atagaggaga aggcatcaaa 480

attaggcgta agtttaaatg gagcagaaat agtagatcca gaaacctcgg ataaactaaa 540

aaaatatgca gatgcttttt atgaattgag aaagaagaag ggaataacac cagaaaaagc 600

ggataaaaata gtaagagatc caatatattt tgctacgatg atggttaagc ttggagatgc 660

agatggattg gtttcagggt cagtgcatatc tacaggtgat cttttgagac caggacttca 720

aatagtaaag acagctccag gtacatcagt agtttccagc acatttataa tggaagtacc 780

aaattgtgaa tatggtgaca atgggtgact tctatttgct gattgtgctg taaatccatg 840

cccagatagt gatcaattgg cttcaattgc aataagtaca gcagaaactg caaagaactt 900

atgiggaatg gatccaaaag tagcaatgct ttcatTTTTCT actaaggga gtcgaaaaca 960

cgaattagta gataaagtta gaaatgctgt agaaattgcc aaaaaagcta aaccagattt 1020

aagtttgac ggagaattac aattagatgc ctctatcgta gaaaaggttg caagtttaaa 1080

ggctcctgaa agtgaagtag caggaaaagc aaatgtactt gtatttccag atctccaagc 1140
 aggaaatata gttataaac ttgttcaaag atttgcaaaa gctgatgcta taggacctgt 1200

 atgccagga tttgcaaac ctataaatga tttgtcaaga ggatgtaact ccgatgatat 1260
 agtaaatgta gtagctgtaa cagcagttca ggcacaagct caaaaagtaa tgaaaatatt 1320
 agtagtaaac tgtggaagtt catctttaa atatcaactt attgatatga aagatgaaag 1380
 cgttgtggca aaaggacttg tagaaagaat aggagcagaa ggttcagttt taacacataa 1440
 agttaacgga gaaaagtttg ttacagagca gccaatggaa gatcataaag ttgctataca 1500
 attagtatta aatgctcttg tagataaaaa acatggtgta ataaaagata tgtcagaaat 1560
 atctgctgta gggcatagag ttttgcattg tggaaaaaaaa tatgcggcat ccattcttat 1620

 tgatgacaat gtaatgaaag caatagaaga atgtattcca ttaggacat tacataatcc 1680
 agctaataa atgggaatag atgcttgtaa aaaactaatg ccaataactc caatggtagc 1740
 agtatttgat acagcatttc atcagacaat gccagattat gcttatactt atgcaatacc 1800
 ttatgatata tctgaaaagt atgatatcag aaaatatggt tttcatggaa cttctcatag 1860
 attcgtttca attgaagcag ccaagttggt aaagaaagat caaaagatc ttaagctaat 1920
 aacttgtcat ttaggaaatg gagctagat atgtgcagta aaccagggaa aagcagtaga 1980
 tacaactatg ggacttactc cccttgcagg acttgtaatg ggaactagat gtggtgatat 2040

 agatccagct ataataccat ttgtaatgaa aagaacaggt atgtctgtag atgaaatgga 2100
 tactttaatg aacaaaaagt caggaatact tggagtatca ggagtaagca gcgattttag 2160
 agatgtagaa gaagctgcaa attcaggaat tgatagagca aaacttgcatt taaatatgta 2220
 ttatcacaaa gttaaatctt tcataggagc ttatgttgca gttttaaatg gagcagatgc 2280
 tataatattt acagcaggac ttggagaaaa ttcagctact agcagatctg ctatatgtaa 2340
 gggattaagc tattttggaa ttaaaataga tgaagaaaag aataagaaaa ggggagaagc 2400
 actagaaata agcacacctg attcaaagat aaaagtatta gtaattccta caaatgaaga 2460

 acttatgata gctagggata caaaagaaat agttgaaaat aaataa 2506
 <210> 3
 <211> 1056
 <212> DNA
 <213> Clostridium ragsdalei
 <400> 3
 atgaaaggtt ttgcaatggt aggtattaac aagttaggat ggattgaaaa gaaaaacca 60
 gtaccaggtc cttatgatgc gattgtacat cctctagctg tatccccatg tacatcagat 120

atacatacgg tttttgaagg agcacttggg aatagggaaa atatgatttt aggtcacgaa 180
 gctgtaggtg aaatagctga agttggcagt gaagttaaag attttaaagt tggcgataga 240
 gttatcgtac catgcacaac acctgactgg agatccttag aagtccaagc tggttttcaa 300

cagcattcaa acggtatgct tgcaggatgg aagttttcca attttaaaga cgggtgatttt 360
 gcagattact ttcattgaaa cgatgcagat atgaatcttg caatacttcc agatgaaata 420
 cctttagaaa gtgcagttaa gatgacagac atgatgacta ctggttttca tggggcagaa 480
 cttgctgaca taaaaatggg ttccagtgtt gtcgtaattg gtataggagc tgttggatta 540
 atgggaatag ccggttccaa acttcgagga gcaggtagaa ttatcgggtg tggaaagcaga 600
 cccgtttgtg ttgaaacagc taaattttat ggagcaactg atattgtaaa ttataaaaat 660
 ggtgatatag ttgaacaaat aatggactta actcatggta aaggtgtaga ccgtgtaatc 720

atggcaggcg gtggtgctga aacactagca caagcagtaa ctatggttaa acctggcggc 780
 gtaatttcta acatcaacta ccatggaagc ggtgatactt tgccaatacc tcgtgttcaa 840
 tggggctgcg gcatggctca caaaactata agaggagggt tatgtcccgg cggacgtctt 900
 agaatggaaa tgctaagaga ccttgttcta tataaacgtg ttgatttgag caaacttgtt 960
 actcatgtat ttgatggtgc agaaaatatt gaaaaggccc ttttgcttat gaaaaataag 1020
 ccaaaagatt taattaaatc agtagttaca ttctaa 1056

<210> 4

<211> 351

<212> PRT

<213> Clostridium ragsdalei

<400> 4

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

Val Ile Val Pro Cys Thr Thr Pro Asp Trp Arg Ser Leu Glu Val Gln
 85 90 95

Ala Gly Phe Gln Gln His Ser Asn Gly Met Leu Ala Gly Trp Lys Phe
 100 105 110

Ser Asn Phe Lys Asp Gly Val Phe Ala Asp Tyr Phe His Val Asn Asp
 115 120 125

Ala Asp Met Asn Leu Ala Ile Leu Pro Asp Glu Ile Pro Leu Glu Ser
 130 135 140

Ala Val Met Met Thr Asp Met Met Thr Thr Gly Phe His Gly Ala Glu
 145 150 155 160

Leu Ala Asp Ile Lys Met Gly Ser Ser Val Val Val Ile Gly Ile Gly
 165 170 175

Ala Val Gly Leu Met Gly Ile Ala Gly Ser Lys Leu Arg Gly Ala Gly
 180 185 190

Arg Ile Ile Gly Val Gly Ser Arg Pro Val Cys Val Glu Thr Ala Lys
 195 200 205

Phe Tyr Gly Ala Thr Asp Ile Val Asn Tyr Lys Asn Gly Asp Ile Val
 210 215 220

Glu Gln Ile Met Asp Leu Thr His Gly Lys Gly Val Asp Arg Val Ile
 225 230 235 240

Met Ala Gly Gly Gly Ala Glu Thr Leu Ala Gln Ala Val Thr Met Val
 245 250 255

Lys Pro Gly Gly Val Ile Ser Asn Ile Asn Tyr His Gly Ser Gly Asp
 260 265 270

Thr Leu Pro Ile Pro Arg Val Gln Trp Gly Cys Gly Met Ala His Lys
 275 280 285

Thr Ile Arg Gly Gly Leu Cys Pro Gly Gly Arg Leu Arg Met Glu Met
 290 295 300

Leu Arg Asp Leu Val Leu Tyr Lys Arg Val Asp Leu Ser Lys Leu Val
 305 310 315 320

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

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

<210> 6

<211> 352

<212> PRT

<213> Thermoanaerobacter ethanolicus

<400> 6

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

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

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

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

 Ile Ala Gly Gly Asn Ala Asp Ile Met Ala Thr Ala Val Lys Ile Val
 245 250 255
 Lys Pro Gly Gly Thr Ile Ala Asn Val Asn Tyr Phe Gly Glu Gly Glu
 260 265 270
 Val Leu Pro Val Pro Arg Leu Glu Trp Gly Cys Gly Met Ala His Lys
 275 280 285
 Thr Ile Lys Gly Gly Leu Cys Pro Gly Gly Arg Leu Arg Met Glu Arg
 290 295 300

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