



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년05월30일
(11) 등록번호 10-1736484
(24) 등록일자 2017년05월10일

- (51) 국제특허분류(Int. Cl.)
C12N 9/02 (2006.01) C12N 15/53 (2006.01)
C12P 3/00 (2006.01)
- (21) 출원번호 10-2013-7034431(분할)
- (22) 출원일자(국제) 2009년09월07일
심사청구일자 2013년12월26일
- (85) 번역문제출일자 2013년12월26일
- (65) 공개번호 10-2014-0022083
- (43) 공개일자 2014년02월21일
- (62) 원출원 특허 10-2010-7013071
원출원일자(국제) 2009년09월07일
- (86) 국제출원번호 PCT/KR2009/005060
- (87) 국제공개번호 WO 2010/027233
국제공개일자 2010년03월11일
- (30) 우선권주장
1020080087794 2008년09월05일 대한민국(KR)
1020080087806 2008년09월05일 대한민국(KR)
- (56) 선행기술조사문헌
FEMS Microbiol Lett., Vol. 282, No. 1, pp. 8-14 (2008.05.)
Mar Biotechnol (NY)., Vol. 9, No. 4, pp. 450-458 (2007.7-8.)

- (73) 특허권자
한국해양과학기술원
경기도 안산시 상록구 해안로 787 (사동)
- (72) 발명자
이정현
경기 성남시 분당구 불정로 219, 116동 101호 (정자동, 한솔마을청구아파트)
강성균
경기 안산시 상록구 감골로 59, 702동 1502호 (사동, 상록수타운월드아파트)
(뒷면에 계속)
- (74) 대리인
윤여강

전체 청구항 수 : 총 6 항

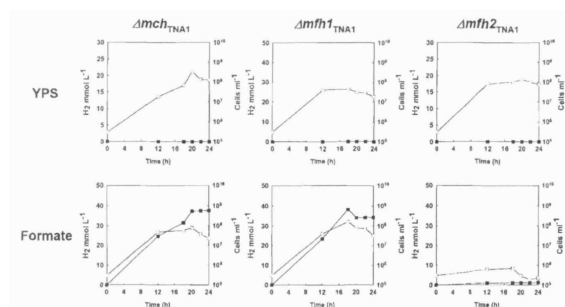
심사관 : 퇴_김남경

(54) 발명의 명칭 일산화탄소를 이용해 수소를 생산해 낼 수 있는 *Thermococcus* spp.로부터 분리된 신규한 수소화효소, 이를 암호화하는 유전자 및 그 유전자를 갖는 미생물을 이용하여 수소를 생산하는 방법

(57) 요약

본 발명은 *Thermococcus* 속에 속하는 고호열성 신균주로부터 분리한 신규한 수소화효소(hydrogenase), 이를 암호화하는 유전자 및 이들을 이용하여 수소를 생산하는 방법에 관한 것이다. 본 발명의 수소생산방법에 따르면 상기 균주를 특정 배양조건으로 배양하는 것만으로 많은 양의 수소를 생산할 수 있으므로, 기존의 수소생산방법에 비하여 경제적이고 효율적이며, 고온에서도 수소를 생산할 수 있는 장점이 있다.

대표도 - 도14



(72) 발명자

이현숙

경기 안산시 상록구 감골로 59, 702동 1502호 (사동, 상록수타운월드아파트)

김상진

경기 안산시 상록구 감골로 59, 707동 904호 (사동, 상록수타운월드아파트)

권개경

경기 안산시 상록구 성호로5길 7, 2층 (일동)

차선신

경기도 안산시 단원구 광덕2로 121 507동 506호 (고잔동, 푸르지오아파트)

전정호

경기 안산시 상록구 성안2길 34, 301호 (사동)

임형순

서울 영등포구 국제금융로 108-6, B동 806호 (여의도동, 진주아파트)

김윤재

경기 안산시 상록구 성안2길 34, 301호 (사동)

배승섭

경기 안산시 상록구 천문4길 19, 402호 (사동)

임재규

경기 시흥시 봉우재로36번길 11, 203호 (정왕동)

김민식

경기 안산시 상록구 장화2길 26, 202호 (사동)

김태완

경기 안산시 상록구 삼리2길 17, 203호 (사동)

류지영

경기 안산시 상록구 장화2길 26, 102호 (사동)

정재연

서울 관악구 인현12길 46-2, 103동 706호 (봉천동, 은천아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호	1525002978
부처명	해양수산부
연구관리전문기관	한국해양과학기술진흥원
연구사업명	해양생명공학기술개발
연구과제명	해양 초고온 고세균 이용 바이오수소 생산기술 개발
기여율	1/1
주관기관	한국해양과학기술원
연구기간	2013.07.01 ~ 2014.06.30

명세서

청구범위

청구항 1

서열번호 5의 아미노산 서열로 이루어진 수소화효소(hydrogenase).

청구항 2

제 1 항의 수소화효소를 암호화하는 유전자.

청구항 3

제 2 항에 있어서, 상기 유전자는 서열번호 16의 염기 서열로 이루어진 것을 특징으로 하는 수소화효소를 암호화하는 유전자.

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

서열번호 21 및 서열번호 16 (CODH-MCH-MNH3 클러스터를 구성하는 유전자)의 유전자가 작동가능하도록 연결된 재조합백터.

청구항 9

제 8 항의 재조합백터로 형질전환된 숙주세포.

청구항 10

제 9 항의 숙주세포를 이용하여 수소를 생산하는 방법에 있어서,

배양용기에 배지를 만드는 단계; 상기 배양용기의 가스층에 일산화탄소를 공급하는 단계;

상기 숙주세포를 상기 배양용기에서 배양하는 단계; 및

상기 배양용기로부터 수소를 추출하는 단계를 포함하는 것을 특징으로 하는 방법.

발명의 설명

기술분야

본 발명은 *Thermococcus* 속에 속하는 신균주로부터 발견한 새로운 수소화효소 (hydrogenase), 이를 암호화하는 유전자, 및 이들을 이용하여 수소를 생산하는 방법에 관한 것이다.

배경기술

[0001]

- [0002] 수소에너지는 중량당 발열량이 석유보다 3배 이상 높으면서도, 이산화탄소, NOx, SOx 등 환경에 악영향을 미칠 수 있는 물질들을 배출하지 않아, 장차 화석에너지를 대체할 에너지로써 각광받고 있다.
- [0003] 종래부터 사용된 수소 생산 방법에는 물의 전기분해, 천연가스나 나프타의 열분해 (thermal-cracking) 또는 증기 개질법 (steam reforming) 등이 있다. 그러나 이러한 방법들은 다시 화석연료를 사용하여 고온, 고압 조건을 만들어야 하는 문제가 있으며, 일산화탄소를 포함한 혼합가스를 발생시키므로 그러한 가스로부터 일산화탄소를 제거하여야 하는 어려운 문제를 발생시킨다.
- [0004] 반면 미생물을 이용한 생물학적 수소 생산 방법은 별도의 에너지를 투입하여 고온, 고압 조건을 만들 필요가 없고, 생성된 가스에 일산화탄소를 포함하지 않는다는 장점이 있다. 이러한 생물학적 수소생산방법은 크게 광합성 미생물을 이용하는 것과 비-광합성 미생물(주로 혐기성 미생물)을 이용하는 것으로 나뉘볼 수 있다. 전자에 속하는 예로 대한민국 등록특허 제10-0680624호 "높은 염분농도에서 수소생성능이 우수한 광합성 세균 로도박터 스페로이데스 균주를 이용한 수소생산 방법" 등이 있다.
- [0005] 그러나 빛을 에너지원으로 사용하는 광합성 세균들의 고농도 배양기술이 아직 충분히 개발되어 있지 않으며, 종래의 광합성 세균들은 높은 분압의 기질이 있을 경우 기질저해가 심하다는 단점이 있다. 또한, 이들은 빛이 존재하는 경우에만 수소생성능이 지속될 수 있다는 문제점이 있다.
- [0006] 따라서, 유기 탄소를 이용하여 수소를 생산할 수 있는 미생물들을 이용하여 수소를 생산하려는 시도가 지속적으로 이루어지고 있으며, 그 예로 대한민국 등록특허 제10-0315663호 "사이트로박터속 균주 Y19 및 이에 의한 수소 생산", 대한민국 등록특허 제10-0315662호 "로도슈도모나스 팔루스트리스 P4 및 이에 의한 수소생산" 등이 있다.
- [0007] 본 발명자들은 고호열성 신균주 *Thermococcus onnurineus* NA1 (기탁번호: KCTC 10859BP)의 신규한 단백질 및 이를 암호화하는 유전자들에 대해 2008년 9월 5일 자로 특허출원하였으며(대한민국 특허출원 제10-2008-0087794호), 본 발명은 이들 특허출원된 단백질 및 유전자들 중 특히 수소생성과 관련된 것들에 대한 것이다. 본 발명자들은 상기 균주의 수소생성능력과 관련된 실험을 한 결과 상기 균주가 고온의 환경에서도 많은 양의 수소를 생성함을 발견하였고, 특히 일산화탄소(CO) 또는 포름산염(formate)을 첨가한 배양조건에서 많이 발현되는 신규한 수소화효소들을 발견하고 본 발명을 완성하게 되었다.

발명의 내용

해결하려는 과제

- [0008] 본 발명의 목적은 고온의 환경에서도 수소를 생산할 수 있는 고호열성 *Thermococcus* spp.의 수소화효소 및 이를 암호화하는 유전자를 제공하고, 이를 이용하여 효율적으로 수소를 생산하는 방법을 제공하는 것에 있다.

과제의 해결 수단

- [0009] 상기 목적을 달성하기 위한 수단으로 본 발명은 혐기적 배양조건에서 수소를 생산할 수 있는 고호열성 *Thermococcus* spp. 균주가 갖고 있는 수소화효소(hydrogenase) 및 이를 암호화하는 유전자를 제공한다. 또한 본 발명은 상기 균주를 배양하여 수소를 생산하는 방법 및 상기 유전자를 이용하여 수소를 생산하는 방법을 제공한다.
- [0010] 제 1 양태로, 본 발명은 고호열성 신균주 *Thermococcus onnurineus* NA1 (기탁번호: KCTC 10859BP)이 생산하는 수소화효소를 제공한다. *T. onnurineus* NA1 은 8종류의 신규한 수소화효소 유전자 클러스터를 갖고 있으며, 이에 속하는 수소화효소에 대한 아미노산 서열은 서열번호 1 내지 서열번호 8과 같다.
- [0011] 제 2 양태로, 본 발명은 상기 아미노산을 암호화하는 유전자를 제공한다. 이에 제한되는 것은 아니나, 바람직하게 상기 유전자는 서열번호 12 내지 서열번호 19이다(상기 서열번호 1 내지 서열번호 8의 아미노산이 각각 서열번호 12 내지 19의 유전자에 대응됨).
- [0012] 제 3 양태로, 본 발명은 *Thermococcus* spp. 를 배양하여 수소를 생산하는 방법을 제공한다. 상기 방법은 1)배양 용기에 배지를 만드는 단계, 2) *Thermococcus* spp.를 상기 배양용기에서 배양하는 단계, 및 3) 상기 배양용기로부터 수소를 추출하는 단계를 포함하는 것을 특징으로 한다. 상기 *Thermococcus* spp.는 바람직하게는 *Thermococcus onnurineus* NA1(기탁번호: KCTC 10859BP)이다.
- [0013] 추가로, 상기 배지는 일산화탄소, 포름산염 및 녹말로 이루어진 그룹에서 선택되는 어느 하나 이상의 물질이 첨

가된 배지가 될 수 있다. 또한, 상기 배양은 80℃의 고온에서 이루어질 수 있으며, 혐기적 조건에서 이루어질 수 있다.

- [0014] 제 4 양태로, 본 발명은 서열번호 9 내지 서열번호 11로 이루어진 군에서 선택되는 어느 하나 이상의 아미노산 서열로 이루어진 탈수소화효소(dehydrogenase)를 제공한다.
- [0015] 제 5 양태로, 본 발명은 상기 탈수소화효소를 암호화하는 유전자를 제공하며, 바람직하게 이는 서열번호 20 내지 서열번호 22이다(상기 서열번호 9 내지 11의 아미노산이 각각 서열번호 20 내지 22에 대응됨).
- [0016] 제 6 양태로, 본 발명은 *T. onnurineus* NA1의 CODH-MCH-MNH3 수소화효소 클러스터를 구성하는 유전자가 모두 작동가능하도록 연결된 재조합백터를 제공한다. 바람직하게, 이에 제한되는 것은 아니나, 상기 유전자는 **서열번호 21**(CODH 탈수소화효소) 및 **서열번호 16**(MCH 수소화효소)의 유전자를 포함한다. 추가로, 본 발명은 상기 재조합백터로 형질전환된 숙주세포를 제공한다.
- [0017] 또한, 본 발명은 배양용기에 배지를 만드는 단계; 상기 배양용기의 가스층에 일산화탄소를 공급하는 단계; 상기 형질전환체를 상기 배양용기에서 배양하는 단계; 및 상기 배양용기로부터 수소를 추출하는 단계를 포함하는 것을 특징으로 하는 상기 형질전환체를 이용하여 수소를 생산하는 방법을 제공한다.
- [0018] 제 7 양태로, 본 발명은 *T. onnurineus* NA1의 FDH2-MFH2-MNH2 수소화효소 클러스터를 구성하는 유전자가 모두 작동가능하도록 연결된 재조합백터를 제공한다. 바람직하게, 이에 제한되는 것은 아니나, 상기 유전자는 **서열번호 22**(FDH2 탈수소화효소) 및 **서열번호 18**(MFH2 수소화효소)의 유전자를 포함한다. 추가로, 본 발명은 상기 재조합백터로 형질전환된 숙주세포를 제공한다.
- [0019] 또한, 본 발명은 배양용기에 포름산염을 포함하는 배지를 만드는 단계; 상기 형질전환체를 상기 배양용기에서 배양하는 단계; 및 상기 배양용기로부터 수소를 추출하는 단계를 포함하는 것을 특징으로 하는 상기 형질전환체를 이용하여 수소를 생산하는 방법을 제공한다.
- [0020] 제 8 양태로, 본 발명은 *T. onnurineus* NA1의 FDH1-MFH1-MNH1 수소화효소 클러스터를 구성하는 유전자가 모두 작동가능하도록 연결된 재조합백터를 제공한다. 바람직하게, 이에 제한되는 것은 아니나, 상기 유전자는 **서열번호 20**(FDH1 탈수소화효소) 및 **서열번호 13**(MFH1 수소화효소)의 유전자를 포함한다. 추가로, 본 발명은 상기 재조합백터로 형질전환된 숙주세포를 제공한다.
- [0021] 또한, 본 발명은 배양용기에 녹말을 포함하는 배지를 만드는 단계; 상기 형질전환체를 상기 배양용기에서 배양하는 단계; 및 상기 배양용기로부터 수소를 추출하는 단계를 포함하는 것을 특징으로 하는 상기 형질전환체를 이용하여 수소를 생산하는 방법을 제공한다.

발명의 효과

- [0022] 본 발명에 따른 수소생산방법은 종래의 화학적 생산 방법과 달리 고온, 고압 조건을 필요로 하지 않고, 상온, 상압 조건에서 수소를 발생시킬 수 있으며, 유해한 부산물을 발생시키지 않는다는 장점이 있다. 또한, 미생물을 이용하여 수소를 생산하는 종래의 기술과 비교하더라도 고순도의 수소를 고효율로 생산할 수 있고, 고온 조건에서도 수소를 생산할 수 있는 장점이 있다.
- [0023] 따라서 석유정제공정 등에서 배출되는 고온의 일산화탄소를 별도의 냉각과정 없이 바로 포획하여 수소생성에 활용할 수 있는 경제적 이점이 있으며, 공기정화의 측면에서도 유용하게 이용될 수 있다.

도면의 간단한 설명

- [0024] **도 1**은 *Thermococcales*, strains, *T. onnurineus* NA1 (NA1), *T. kodakaraensis*, *P. furiosus*, 및 *P. abyssi*의 프로테옴(proteome)의 공통영역 및 특이영역을 나타내는 벤다이어그램이다. 상기 균주들에 대한 단백질 세트는 NCBI의 RefSeq collection 에서 구하였다.
- 도 2의 A**는 *T. onnurineus* NA1 의 8개 수소화효소 유전자 클러스터의 대표도이다. A, B, C, D; 막 내 수소화효소(membrane-bound hydrogenases) 및 세포질 내 NiFe-수소화효소(cytoplasmic NiFe-hydrogenases). S1, S2, S3; *T. onnurineus* NA1. 유전자들에 특이적인 수소화효소들을 COG 기능 카테고리(functional categories)에 따라 채색하였다. TON_0051-0055는 서열번호 1부터 5, TON_0486-0498은 서열번호 35부터 47, TON_0533-0544는 서열번호 48부터 59, TON_1583-1595는 서열번호 100부터 112, TON_0261-0289는 서열번호 6부터 34, TON_1016-1031은 서열번호 60부터 75, TON_1559-1582는 서열번호 76부터 99를 나타낸다.

도 2의 B는 *T. onnurineus* NA1의 지놈 상에서 3-모듈 유전자 클러스터를 갖는 3개의 수소화효소 유전자 클러스터(*fdh1-mfh1-mnh1*, *fdh2-mfh2-mnh2* 및 *codh-mch-mnh*)의 유전자 구성을 나타낸 것이다. 같은 서브클러스터에 속하는 유전자들은 같은 색으로 표시하였다.

도 3은 31개의 고세균 유전체의 수소화효소 유전자 클러스터의 분포 및 보존 패턴을 나타낸다. 파란색 꺾쇠(아래에서 첫 번째, 세 번째, 다섯 번째 꺾쇠)로 표시한 부분은 31개의 고세균 유전체의 CDS들에 대해 낮은(<25%) 유사도를 보인 CDS들을 나타낸다. 검은색으로 표시한 부분은 *P. abyssi*의 수소화효소 4에 유사한 CDS들을 나타낸다. 1번부터 31번에 해당하는 고세균 유전자는 다음과 같다.

[표 8]

번호	유전자	번호	유전자
1	<i>Aeropyrum pernix</i>	17	<i>Pyrococcus abyssi</i>
2	<i>Pyrobaculum aerophilum</i>	18	<i>Pyrococcus furiosus</i>
3	<i>Sulfolobus acidocaldarius</i> DSM 639	19	<i>Pyrococcus horikoshii</i>
4	<i>Sulfolobus solfataricus</i>	20	<i>Thermococcus kodakaraensis</i> KOD1
5	<i>Sulfolobus tokodaii</i>	21	<i>Archaeoglobus fulgidus</i>
6	<i>Haloarcula marismortui</i> ATCC 43049	22	<i>Methanosarcina barkeri</i> fusaro
7	<i>Natronomonas pharaonis</i>	23	<i>Methanosarcina mazei</i>
8	<i>Halobacterium</i> sp	24	<i>Methanosarcina acetivorans</i>
9	<i>Halocquadratum walsbyi</i>	25	<i>Methanospirillum hungatei</i> JF-1
10	<i>Methanococcoides burtonii</i> DSM 6242	26	<i>Methanobacterium thermoautotrophicum</i>
11	<i>Picrophilus torridus</i> DSM 9790	27	<i>Methanococcus jannaschii</i>
12	<i>Thermoplasma acidophilum</i>	28	<i>Methanococcus maripaludis</i> S2
13	<i>Thermoplasma volcanium</i>	29	<i>Methanospiraera stadtmanae</i>
14	<i>Methanosaeta thermophila</i> PT	30	<i>Methanovirus kandleri</i>
15	<i>Pyrobaculum islandicum</i> DSM 4184	31	<i>Nanoarchaeum equitans</i>
16	<i>Thermofilum pendens</i> Hrk 5		

도 4 및 도 5는 CODH와 F420 수소화효소 단백질들의 알파 서브유닛(α subunit)을 비교한 것이다. 도 4는 CODH의 계통수(Phylogenetic tree)이며, 도 5는 F420 수소화효소 알파 서브유닛의 계통수이다. 계통수상 각각의 단백질의 동족체(homologue)들은 NCBI nr 데이터베이스에서 얻었다.

도 6은 CO에 의존하는 *T. onnurineus* NA1의 성장분석표이다. *T. onnurineus* NA1을 CO(2, 정삼각형), 황(3, 네모) 또는 양자 모두(4, 역삼각형) 첨가된 배지 1에서 키웠다. 대조군으로 첨가물이 없는 배지(1, 원)와 YPS 배지(C)에서 배양한 것을 표시하였다. 필터 상의 DAPI-염색된 세포들을 형광현미경으로 직접 세었다. a, 다양한 효모추출물(yeast extract, YE) 농도의 배지 조성에 따른 효과. b, 다른 보충물들이 첨가된 배지 1에서 *T. onnurineus* NA1의 성장곡선. c, CODH 유전자의 전사(transcription) 분석.

도 7은 YPS(A) 또는 CO-함유 배지(B)에서 *T. onnurineus* NA1의 성장 및 수소 생성을 보여준다. 열린 원은 성장을 나타내며, 닫힌 원은 수소생성을 나타냄.

도 8은 YPS(A) 또는 CO-함유 배지(B)에서 *T. onnurineus* NA1의 수소화효소 유전자 클러스터의 발현을 마이크로어레이(A) 및 RT-PCR(B)로 분석한 것이다. (A)는 *T. onnurineus* NA1의 8번의 수소화효소 유전자 클러스터의 마이크로어레이 분석이다. CO에서의 mRNA 변화의 계층적(hierarchical) 클러스터링과 대조군으로서 YPS 성장 조건에서의 것을 비교하였다. 상향조절(up-regulation) 및 하향조절(down-regulation)은 각각 붉은 색과 녹색으로 표시하였다. 성장조건들은 클러스터링의 상단에 표시하였다. 클러스터링의 오른쪽에, *codh-mch-mnh3*, *fdh1-mfh1-mnh1* 또는 *fdh2-mfh2-mnh2* 각각의 ORF들을 막대로 표시하였다. YE: yeast extract. (B)는 mbh(TON_1593), mbx(TON_0489), frh(TON_1560), sulf I(TON_0534), mch(TON_1023), mfh2(TON_1569), 또는 mfh1(TON_0276) 수소화효소의 큰 서브유닛 각각의 CO 또는 YPS 조건에서의 정량적 RT-PCR 분석 결과이다. *T. onnurineus* NA1의 *cha*, 샤페로닌(chaperonine)-암호화 유전자, 가 대조군으로서 발현 수준을 표준화하는데 사용되었다.

도 9는 mch(TON_1023) 및 mfh2(TON_1569) 수소화효소 각각의 큰 서브유닛의 표적 유전자 파괴(Targeted gene

disruption)를 나타낸다. (A)는 *T. onnurineus* NA1의 *codh-mch-mnh3* 및 *fdh2-mfh2-mnh2* 클러스터 각각의 유전자 구성이다. P_{gdh}, *T. kodakaraensis* KOD1의 글루타메이트 탈수소화효소 유전자의 5'-상류 지역의 프로모터; hmg_{Pfu}, *Pyrococcus furiosus*의 3-히드록시-메틸글루타릴 코엔자임 A 환원효소 유전자(3-hydroxy-methylglutaryl coenzyme A reductase gene). (B)는 PCR에 의한 유전자 결손의 확인을 나타낸다. 왼쪽 패널은 과발현 카세트(P_{gdh}-hmg_{Pfu}) 영역에 대한 프라이머를 이용한 PCR 증폭을 나타낸다. 오른쪽 패널은 *mch*_{TNA1} 또는 *mfh2*_{TNA1} 수소화효소 각각의 큰 서브유닛 영역에 대한 프라이머들을 이용한 PCR 증폭을 나타낸다. M, 크기 마커(1kb ladder); W, 야생형; lane 1-2, 각각의 돌연변이 균주.

도 10은 YPS 또는 CO-함유 배지에서 Δmch_{TNA1} 또는 $\Delta mfh2_{TNA1}$ 돌연변이 균주의 성장 및 수소생성을 나타낸다. 열린 원, 성장; 닫힌 원; 수소 생성.

도 11은 YPS(A) 또는 포름산염-함유 배지(B)에서 *T. onnurineus* NA1의 성장 및 수소 생성을 보여준다. 열린 원은 성장을 나타내며, 닫힌 원은 수소생성을 나타냄.

도 12는 YPS(A) 또는 포름산염-함유 배지(B)에서 *T. onnurineus* NA1의 수소화효소 유전자 클러스터의 발현을 마이크로어레이(A) 및 RT-PCR(B)로 분석한 것이다. (A)는 *T. onnurineus* NA1의 8 반복의 수소화효소 유전자 클러스터의 마이크로어레이 분석이다. 외부 포름산염상에서의 mRNA 변화의 계층적(hierarchical) 클러스터링과 대조구로서 YPS 성장 조건에서의 것을 비교하였다. 상향조절(up-regulation) 및 하향조절(down-regulation)은 각각 붉은 색과 녹색으로 표시하였다. 성장조건들은 클러스터링의 상단에 표시하였다. 클러스터링의 오른쪽에, *mbh*, *codh-mch-mnh3*, *fdh1-mfh1-mnh1* 또는 *fdh2-mfh2-mnh2* 각각의 ORF들을 막대로 표시하였다. YE: yeast extract. (B)는 *mbh* (TON_1593), *mbx* (TON_0489), *frh* (TON_1560), *sulf I* (TON_0534), *mch* (TON_1023), *mfh2* (TON_1569), 또는 *mfh1* (TON_0276) 수소화효소의 큰 서브유닛 각각의 포름산염 또는 YPS 조건에서의 정량적 RT-PCR 분석 결과이다. *T. onnurineus* NA1의 *cha*, 샤페론(chaperonine)-암호화 유전자, 가 대조구로서 발현수준을 표준화하는데 사용되었다.

도 13은 *mfh1* (TON_0276) 및 *mfh2* (TON_1569) 수소화효소 각각의 큰 서브유닛의 표적 유전자 파괴(Targeted gene disruption)를 나타낸다. (A)는 *T. onnurineus* NA1의 *fdh1-mfh1-mnh1* 및 *fdh2-mfh2-mnh2* 클러스터 각각의 유전자 구성이다. P_{gdh}, *T. kodakaraensis* KOD1의 글루타메이트 탈수소화효소 유전자의 5'-상류 지역의 프로모터; hmg_{Pfu}, *Pyrococcus furiosus*의 3-히드록시-메틸글루타릴 코엔자임 A 환원효소 유전자(3-hydroxy-methylglutaryl coenzyme A reductase gene). (B)는 PCR에 의한 유전자 결손의 확인을 나타낸다. 왼쪽 패널은 과발현 카세트(P_{gdh}-hmg_{Pfu}) 영역에 대한 프라이머를 이용한 PCR 증폭을 나타낸다. 오른쪽 패널은 *mfh1*_{TNA1} 또는 *mfh2*_{TNA1} 수소화효소 각각의 큰 서브유닛 영역에 대한 프라이머들을 이용한 PCR 증폭을 나타낸다. M, 크기 마커(1kb ladder); W, 야생형; lane 1-2, 각각의 돌연변이 균주.

도 14는 YPS 또는 포름산염-함유 배지에서 Δmch_{TNA1} , $\Delta mfh1_{TNA1}$ 또는 $\Delta mfh2_{TNA1}$ 돌연변이 균주의 성장 및 수소생성을 나타낸다. 열린 원, 성장; 닫힌 원; 수소 생성.

발명을 실시하기 위한 구체적인 내용

[0025] 제 1 양태로, 본 발명은 혐기성 조건에서 수소를 생산하는 고호열성 신균주 *Thermococcus onnurineus* NA1 (기탁 번호: KCTC 10859BP)이 생산하는 수소화효소를 제공한다. 상기 균주는 이스트 마누스 바신(East Manus Basin)에 있는 파크마누스 지역의 심해 열수 분출구로부터 분리한 것으로, 대한민국 생명공학연구원 생물자원센터(KCTC)에 2005년 10월 7일 기탁하여, 2005년 10월 20일에 KCTC 10859BP의 기탁번호를 부여받았다. 상기 균주의 특징 및 배양방법 등에 대해서는 본 출원의 기초출원인 대한민국 특허출원 제10-2007-0127255호에 기재되어 있다.

[0026] *T. onnurineus* NA1은 8종류의 신규한 수소화효소 유전자 클러스터를 갖고 있다. 상기 수소화효소는 수소 분자(H₂)의 대사에 관련된 핵심적 효소로서, $2H^+ + 2e^- \leftrightarrow H_2$ 의 가역적 반응에 촉매 작용을 한다. 바람직하게 상기 클러스터에 속하는 수소화효소는 서열번호 1 내지 서열번호 8로 이루어진 군에서 선택된 어느 하나 이상의 아미노산 서열을 포함하는 단백질 및 이의 기능적 동등물을 제공한다. "기능적 동등물"에는 상기 단백질들에 있어 그 일부 또는 전부가 치환되거나, 아미노산의 일부가 결실 또는 부가된 아미노산 서열 변형체가 포함된다. 아미노산의 치환은 바람직하게는 보존적 치환이다. 천연에 존재하는 아미노산의 보존적 치환의 예는 다음과 같다; 지방족 아미노산(Gly, Ala, Pro), 소수성 아미노산(Ile, Leu, Val), 방향족 아미노산(Phe, Tyr, Trp), 산성 아

미노산(Asp, Glu), 염기성 아미노산(His, Lys, Arg, Gln, Asn) 및 황함유 아미노산(Cys, Met). 아미노산의 결실은 바람직하게는 수소화효소의 활성에 직접 관여하지 않는 부분에 위치한다.

- [0027] 제 2 양태로, 본 발명은 상기 아미노산을 암호화하는 유전자를 제공한다. 이에 제한되는 것은 아니나, 바람직하게 상기 유전자는 서열번호 12 내지 서열번호 19이다(상기 서열번호 1 내지 서열번호 8의 아미노산이 각각 서열번호 12 내지 19의 유전자에 대응됨).
- [0028] 제 3 양태로, 본 발명은 *Thermococcus* spp. 를 배양하여 수소를 생산하는 방법을 제공한다. 상기 방법은 1) 배양 용기에 배지를 만드는 단계, 2) *Thermococcus* spp.를 상기 배양용기에서 배양하는 단계, 및 3) 상기 배양용기로부터 수소를 추출하는 단계를 포함하는 것을 특징으로 한다. 상기 *Thermococcus* spp.는 바람직하게는 *Thermococcus onnurineus* NA1(기탁번호: KCTC 10859BP)이다.
- [0029] 추가로, 상기 배지는 일산화탄소, 포름산염 및 녹말로 이루어진 그룹에서 선택되는 어느 하나 이상의 물질이 첨가된 배지가 될 수 있다. 또한, 상기 배양은 80°C의 고온에서 이루어질 수 있으며, 혐기적 조건에서 이루어질 수 있다.
- [0030] 제 4 양태로, 본 발명은 서열번호 9 내지 서열번호 11로 이루어진 군에서 선택되는 어느 하나의 아미노산 서열로 이루어진 탈수소화효소(dehydrogenase)를 제공한다. 상기 탈수소화효소, Fdh1(서열번호 20), Fdh2(서열번호 22) 및 CODH(서열번호 21) 각각은 수소화효소 MFH1, MFH2 및 MCH 수소화효소와 함께 클러스터를 이루어 기능할 수 있다.
- [0031] 제 5 양태로, 본 발명은 상기 탈수소화효소를 암호화하는 유전자를 제공하며, 바람직하게 이는 서열번호 20 내지 서열번호 22이다(상기 서열번호 9 내지 11의 아미노산이 각각 서열번호 20 내지 22에 대응됨).
- [0032] 제 6 양태로, 본 발명은 *T. onnurineus* NA1의 CODH-MCH-MNH3 수소화효소 클러스터를 구성하는 유전자가 모두 작동가능하도록 연결된 재조합백터를 제공한다. 바람직하게, 이에 제한되는 것은 아니나, 상기 유전자는 **서열번호 21**(CODH 탈수소화효소) 및 **서열번호 16**(MCH 수소화효소)의 유전자를 포함한다. 상기 "백터"는 또 다른 핵산을 그것에 결합시켜 이송시킬 수 있는 핵산 분자를 의미한다. 상기 백터에 의해 운반되는 각 재조합형 유전자에 의해 암호화되는 단백질을 합성시킬 수 있는 "발현백터"로 플라스미드, 코스미드 또는 파아지 등을 이용할 수 있다. 바람직한 백터는 그것이 결합된 핵산을 자기 복제 및 발현시킬 수 있는 백터이다.
- [0033] 추가로, 본 발명은 상기 재조합백터로 형질전환된 숙주세포를 제공한다. 상기 재조합 백터는 원핵생물, 곰팡이, 식물, 동물세포 등의 세포를 형질전환시켜 높은 효율로 수소를 생산시킬 수 있는 형질전환 세포를 만드는 데 이용될 수 있다. '형질전환'이란 용어는 외래 DNA 또는 RNA가 세포에 흡수되어 세포의 유전형이 변화되는 것을 말한다. 상기 숙주세포를 만들기 위해 각 세포에 따른 공지의 형질전환 방법을 이용할 수 있다.
- [0034] 또한, 본 발명은 배양용기에 배지를 만드는 단계; 상기 배양용기의 가스층에 일산화탄소를 공급하는 단계; 상기 형질전환체를 상기 배양용기에서 배양하는 단계; 및 상기 배양용기로부터 수소를 추출하는 단계를 포함하는 것을 특징으로 하는 상기 형질전환체를 이용하여 수소를 생산하는 방법을 제공한다.
- [0035] 제 7 양태로, 본 발명은 *T. onnurineus* NA1의 FDH2-MFH2-MNH2 수소화효소 클러스터를 구성하는 유전자가 모두 작동가능하도록 연결된 재조합백터를 제공한다. 바람직하게, 이에 제한되는 것은 아니나, 상기 유전자는 **서열번호 22**(FDH2 탈수소화효소) 및 **서열번호 18**(MFH2 수소화효소)의 유전자를 포함한다. 추가로, 본 발명은 상기 재조합백터로 형질전환된 숙주세포를 제공한다.
- [0036] 상기 "백터", "형질전환" 및 "숙주세포"에 관한 사항은 상기 제 6 양태에 기술된 바와 같다.
- [0037] 또한, 본 발명은 배양용기에 포름산염을 포함하는 배지를 만드는 단계; 상기 형질전환체를 상기 배양용기에서 배양하는 단계; 및 상기 배양용기로부터 수소를 추출하는 단계를 포함하는 것을 특징으로 하는 상기 형질전환체를 이용하여 수소를 생산하는 방법을 제공한다.
- [0038] 제 8 양태로, 본 발명은 *T. onnurineus* NA1의 FDH1-MFH1-MNH1 수소화효소 클러스터를 구성하는 유전자가 모두 작동가능하도록 연결된 재조합백터를 제공한다. 바람직하게, 이에 제한되는 것은 아니나, 상기 유전자는 **서열번호 20**(FDH1 탈수소화효소) 및 **서열번호 13**(MFH1 수소화효소)의 유전자를 포함한다. 추가로, 본 발명은 상기 재조합백터로 형질전환된 숙주세포를 제공한다.
- [0039] 상기 "백터", "형질전환" 및 "숙주세포"에 관한 사항은 상기 제 6 양태에 기술된 바와 같다.
- [0040] 또한, 본 발명은 배양용기에 녹말을 포함하는 배지를 만드는 단계; 상기 형질전환체를 상기 배양용기에서 배양

하는 단계; 및 상기 배양용기로부터 수소를 추출하는 단계를 포함하는 것을 특징으로 하는 상기 형질전환체를 이용하여 수소를 생산하는 방법을 제공한다.

[0041] [발명의 실시를 위한 형태]

[0042] 이하, 본 발명을 실시예에 의하여 더욱 상세하게 설명한다. 단, 하기 실시예는 본 발명을 예시하는 것으로 본 발명의 내용이 실시예에 의해 한정되는 것은 아니다.

[0043] 실시예 1. *Thermococcus onnurineus* NA1 균주의 수소화효소 유전자 분석

[0044] (1) 실험방법

[0045] 1) 배양 조건

[0046] 일반적인 배양조건으로서, 80℃의 혐기성 환경에서 YPS(yeast extract-peptone-sulfur) 배지에 세포들을 배양시켰다(Holden et al. 2001). 생리적 특성들에 대한 시험은 1.0 ml trace element mixture, 1 ml 비타민 용액 (Balch, W. E., G. E. Fox, L. J. Magrum, C. R. Woese, and R. S. Wolfe. 1979. Methanogens: reevaluation of a unique biological group. Microbiol. Rev. 43:260-296.), NaCl(30 g l⁻¹) 및 효모추출물(yeast extract) (0.5 g l⁻¹)을 추가한 변형배지 1(Sokolova, T. G., C. Jeanthon, N. A Kostrikin, N. A. Chernyh, A. V. Lebedinsky, E. Stackebrandt, and E. A. Bonch-Osmolovskaya. 2004. The first evidence of anaerobic CO oxidation coupled with H₂ production by a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. Extremophiles 8:317-323.)을 이용하여 수행하였다. pH는 NaOH를 사용하여 8.0에 맞췄다. 혐기적 조건에서 만든 배지를 25-ml 혈청병(serum bottle)에 넣고, 나머지 가스층(15ml)을 N₂/CO₂ (80:20, 1 bar) 또는 100% CO로 채웠다. 포름산염(formate)이나 녹말(starch)로 배양시키는 경우에는, 가압멸균 (autoclaving)시키기 전에 각각 10 g l⁻¹의 포름산염나트륨(Sigma)이나 5 g l⁻¹의 가용성 녹말(Sigma)을 배지에 첨가하였다. 생리적 시험을 위한 배양은 모두 80 ℃에서 2일간 수행하였다.

[0047] 2) 유전자 서열분석

[0048] *T. onnurineus* NA1의 유전체 서열은 총 지놈 샷건 시퀀싱(whole genome shotgun sequencing) 및 피로시퀀싱(pyrosequencing)을 이용하여 결정하였다. 캐필러리 시퀀싱(capillary sequencing)을 위하여, 2-에서 3-kb 크기의 삽입 라이브러리(insert library)(11,028 클론), 40-kb의 삽입 라이브러리(1,870 클론), 및 35-kb의 삽입 라이브러리(288 클론)를 제작하고, ABI 3730XL sequencer (Applied Biosystems)를 이용하여 서열 분석하였다. 피로시퀀싱을 위하여, GS-20 sequencer (454 Life Sciences)를 이용하여 581,990개의 DNA 단편들에 대해 서열 분석하였다. 양 서열분석기에 의해 생성된 컨티그(contig)들을 조합하고, 클론 워킹(clone walking) 및 PCR 시퀀싱으로 시퀀싱 간격(sequencing gap)을 메웠다. 단백질질을 암호화하는 ORF 및 RNA 유전자들은 Glimmer 3.0 (University of Maryland), GSFinder, 및 RBSFinder의 조합과 manual ORF fitting process를 통하여 예측하였다. 모든 ORF들을 결정한 후에, NCBI nr 단백질 데이터베이스, KEGG (Kyoto Encyclopedia of Genes and Genomes) 및 COGs (Clusters of Orthologous Groups of proteins) 데이터베이스에 대해 BLASTp 검색(Tatusova, R. L., D. A. Natale, I. V. Garkavtsev, T. A. Tatusova, U. T. Shankavaram, B. S. Rao, B. Kiryutin, M. Y. Galperin, N. D. Fedorova, and E. V. Koonin. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 29:22-28.)을 이용하여 단백질 서열을 더 분석하였다. tRNA-암호화 부위는 tRNA scan-SE를 이용하여 예측하였다(Lowe, T. M., and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25:955-964.).

[0049] 3) 단백질 분석

[0050] *T. onnurineus* NA1 세포들을 4% SDS 및 4 mM EDTA가 포함된 100 mM Tris-HCl 버퍼(buffer)(pH 6.8)에 부유시키고, 10분간 끓인 후, 22,000g 로 20분간 원심분리시켰다. 12% SDS-PAGE를 이용하여 세포 용해물(lysate)을 분리시켰고, 분자 크기에 기초하여 30개의 분획(fraction)들을 얻었다. 그 후 트립신(trypsin) (Promega, US A)을 이용하여 그것들을 인-겔 분해시켰고(in-gel digestion)(Kim, Y. H., K. Cho, S. H. Yun, J. Y. Kim, K. H. Kwon, J. S. Yoo, and S. I. Kim. 2006. Analysis of aromatic catabolic pathways in *Pseudomonas putida* KT 2440 by combined proteomic approach: 2-DE/MS and cleavable ICAT analysis. Proteomics 6:1301-1318.), 질량분석계(Thermo Finnigan LTQ IT)로 분석하기 위해 트립신으로 분해시킨 물질들을 0.5% 트리플루오로아세트

산 용액(trifluoroacetic acid solution)에 녹였다. 펩티드 동정은 시퀀스트 프로그램(Sequest program) (Thermo Finnigan, San Jose, CA)을 이용하여 수행하였다.

- [0051] 4) 총(total) RNA 분리 및 RT-PCR 분석
- [0052] 50-ml의 *T. onnurineus* NA1 배양액을 N₂/CO₂ (80:20, 1 bar) 또는 100% CO 기체상(gas phase)에서 다양한 농도의 효모추출물(yeast extract)을 첨가한 변형 배지 1에서 반-지수성장단계(mid-exponential growth phase)까지 키웠다. 6,000 x g 로 30분간 원심분리하여 세포들을 얻었다. 500μl의 트리졸 시약(Trizol reagent) (Invitrogen, Carlsbad, CA)를 첨가한 50μl의 50mM Tris-HCl 버퍼(buffer)(pH 7.5)에 펠렛(pellet)을 재현탁시켰다. 그 세포들을 냉동과 해동과정을 통해 용해시키고, 200μl의 클로로포름(chloroform)을 이용하여 시료들을 추출하였다. 총 RNA를 포함하는 수용액상(aqueous phase)은 에탄올 침전을 통해 더 처리한 후, 증류수에 재현탁시켰다. RNA의 농도 및 완전성(integrity)은 0.8% 아가로스 젤 분석과 260 및 280nm 에서의 흡광도 측정을 통해 결정하였다. 제품 사용설명서에 따른 방법으로 SuperScript™ II 역전사효소(Reverse Transcriptase)(Invitrogen)를 이용하여 역전사 및 PCR 증폭을 수행하였다. 대조군으로 CODH (carbon monooxide dehydrogenase) 및 Hsp60 (chaperonine) 의 증폭을 위해 다음 두 세트의 프라이머들을 사용하였다.
- [0053] CODH gene (forward, 5'-GGACCATGTAGAATCGAYCCGTTY-3'(서열번호 23) 및 reverse, 5'-TTCRTTCCGGTACAGCA-3'(서열번호 24))
- [0054] Hsp60 gene (forward, 5'-ATGGCACAGCTTAGTGGACAG-3'(서열번호 25) 및 reverse, 5'-CAAGGATTCCTGGGCTTCTC-3'(서열번호 26)).
- [0055] 5) 컴퓨터 분석
- [0056] 아미노산 서열의 상동성 검색은 NCBI 의 non-redundant protein 데이터베이스에 대해 BLAST 프로그램을 사용하여 수행하였다. 그룹 4 수소화효소의 L1 신호(C[GS][ILV]C[AGNS]xxH, x는 어떤 아미노산을 표시)를 가진 단백질들을 위한 모티프 검색은 NCBI 의 non-redundant protein 데이터베이스에 대해 ProteinFinder 프로그램 (Ensoltek, Korea)을 사용하여 수행하였다. 단백질들에 대한 다중서열정렬(Multiple sequence alignment) 은 ClustalW 프로그램(Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.)을 사용하여 수행하였으며, 계통학적 트리(phylogenetic tree)는 Molecular Evolutionary Genetics Analysis (Mega 4.1) 소프트웨어(Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596-1599.)를 사용하여 구축하였다. 16S rRNA 서열의 계통학적 트리는 Ribosomal 데이터베이스 사이트 (<http://rdp.cme.msu.edu/index.jsp>)에서 얻은 프리얼라인(prealigned) 서열들을 사용하여 만들었다.
- [0057] 6) 서열로고의 생성
- [0058] 로고 표시는 관련 서열들에 의해 공유되는 주어진 모티프들 각각의 위치와 관련된 정보를 시각화하기 위해 사용되었다. 그래픽 표시에서, 각각의 위치의 전체적인 높이는 그 위치(bits로 표시)에서의 보존도와 관련이 있는 반면, 어떤 위치내 심볼들의 상대적인 크기는 그것들의 상대적인 빈도들을 나타낸다. 로고 분석은 Berkeley Structural Genomics Center (<http://weblogo.berkeley.edu/>)에서 수행되었다.
- [0059] (2) 분석결과
- [0060] 1) *T. onnurineus* NA1 유전자의 일반적 특징
- [0061] 열수분출공(hydrothermal vent) 지역에서 *Thermococcus* spp. 가 성공적인 우점종이 된 이유를 밝히기 위해, *T. onnurineus* NA1의 유전체 서열을 임의적 총 지놈 샷건 시퀀싱(whole genome shotgun sequencing) 및 피로시퀀싱(pyrosequencing)을 조합하여 결정하였다. 그 결과, *T. onnurineus* NA1 은 별도의 염색체 외 구성요소(exochromosomal element)들이 없으며, 하나의 환상 염색체(circular chromosome)(1,847,607 bp)를 갖고 있음을 밝혔고, 총 1,976 개의 암호화 DNA 서열들(coding DNA sequences, CDSs)을 동정하였다(표 1 및 도 1). 1104 개의 CDS(55.8%) 들은 상동성(homology) 및 도메인(domain) 검색을 통해 어노테이션(annotation) 할 수 있으나, 나머지 872개의 CDS 들의 기능은 1차 구조로부터 예측할 수 없었다. 유전체(genome-wide) 스케일로 단백질 유사도를 검색하였을 때, *T. onnurineus* NA1 단백질의 82.7%가 *Thermococcales* 의 것과 유사성을 보였다.

[0062] [표 1]

[0063] *T. kodakaraensis* KOD1 (KOD1) 및 다른 *Pyrococcus* 균주들과 *T. onnurineus* NA1 (NA1) 유전체의 일반적인 특징.

	NA1	KOD1	<i>P. abyssi</i>	<i>P. barikoshii</i>	<i>P. furiosus</i>
유전자 크기 (bp)	1,847,607	2,088,737	1,765,118	1,738,505	1,908,256
단백질 압축	90.1%	92.1%	91.1%	91.2%	92.5%
화 부위					
GC 함량	51.0%	52.0%	44.7%	41.9%	40.8%
CDSs*	1976	2306	1784	2064	2065
tRNAs	46	46	46	46	46
rRNAs	5S(x2), 7S, 16S, 32S	5S(x2), 7S, 16S, 32S	5S(x2), 7S, 16S, 32S	5S(x2), 7S, 16S, 32S	5S(x2), 7S, 16S, 32S

* 균주들의 단백질 세트는 NCBI의 RefSeq collection 에서 얻었다.

[0064]

[0065] 2) 수소화효소 클러스터(cluster)

[0066]

T. onnurineus NA1의 유전체에서 수소화효소(hydrogenase) 및 그와 관련된 단백질들의 비율이 높은 것을 발견하였으며(5.5%), 이는 CO 탈수소효소(dehydrogenase) 및 포름산염 탈수소효소(formate dehydrogenase)를 포함하는 산화환원효소들과 관련된 환원력의 보존 내지 재순환의 증가를 나타낸다.

[0067]

수소화효소는 그들의 촉매금속중심(catalytic metal center)에 근거하여 [NiFe]-수소화효소, [FeFe]-수소화효소 및 [Fe]-수소화효소로 분류될 수 있다. 상기 각각의 수소화효소군 내에 보존된 특징적인 기능성 중심(functional core)에 근거할 때, 하나의 수소화효소를 제외한 *T. onnurineus* NA1 내의 모든 수소화효소는 [NiFe]-수소화효소에 속하는 것으로 볼 수 있다. Vignais *et al.*의 분류체계에 따르면, *T. onnurineus* NA1의 [NiFe]-수소화효소는 그룹 3 (하나의 F420-reducing 및 두 개의 NADP-reducing 수소화효소들) 또는 그룹 4 (4개의 막-관련 수소화효소)에 속한다(Silva, P.J., van den Ban, E.C., Wassink, H., Haaker, H., de Castro, B., Robb, F.T. and Hagen, W.R. (2000) Enzymes of hydrogen metabolism in *Pyrococcus furiosus*. Eur. J. Biochem. 267, 6541-6551). 그룹 4에 속하는 수소화효소들은 에너지 전환 수소화효소(energy-converting hydrogenases, Ech)라 명명되었으며, 박테리아 및 고세균 사이에 널리 분포되어 있다.

[0068]

수소화효소 대사의 분자적 기초를 이해하기 위하여 수소화효소 유전자 클러스터들을 비교분석 하였다. 상기 수소화효소들의 계통학적 분석 결과, 그룹 4 수소화효소들을 두 개의 서브그룹, 4a 및 4b로 분류할 수 있었고, 그룹 4b의 모든 서열들에 공통적인 한 쌍의 모티프(motif) 패턴을 발견할 수 있었다.

[0069]

도 2에 표시된 바와 같이, *Pyrococcus* spp. 및 *T. kodakaraensis* KOD1에서 보고된 두 개의 막 내 수소화효소(membrane-bound hydrogenase)인 Mbh(TON_1590-1595, 서열번호 8), Mbx(TON_0489-0486, 서열번호 3) 및 두 개의 세포질 내 NiFe-수소화효소(cytoplasmic NiFe-hydrogenase)인 Sulf-I(TON_0533-0544, 서열번호 4), Sulf-II(TON_0051-0055, 서열번호 1) 과 함께 *T. onnurineus* NA1 유전체에서 세 개와 추가적인 수소화효소 클러스터 FDH1-MFH1-MNH1 (Hyg4-I, S1:TON_0279-0274, MFH1: 서열번호 2), CODH-MCH-MNH3 (Hyg4-II, S2:TON_1021-1024, MCH: 서열번호 5), FDH2-MFH2-MNH2 (Hyg4-III, S3:TON_1565-1571, MFH2: 서열번호 7)와 Frh(TON_1559-1562, 서열번호 6)수소화효소를 찾았다. 31개의 고세균 유전체(archaeal genomes)의 CDS들로 수소화효소의 유전자 클러스터 분석을 한 결과 FDH1-MFH1-MNH1(Hyg4-I), CODH-MCH-MNH3(Hyg4-II) 및 FDH2-MFH2-MNH2(Hyg4-III)는 일차 서열이 특이적이었고, 이는 *P. abyssi* 및 *R. rubrum*의 수소화효소-4의 구성요소들과 유사도가 낮음을 보여준다(도 3). 추가적인 수소화효소들과 마찬가지로 FDH2-MFH2-MNH2(Hyg4-III) 클러스터(TON_1559-1582, 서열번호 76부터 99)는 유전체에서 F420 수소화효소(TON_1559-1561, 서열번호 76부터 78)의 α/β/γ 서브유닛(subunit)들을 포함하였다. F420 수소화효소의 α/β/γ 서브유닛은 특이적인 일차 서열을 가졌고, 이는 *Methanococcus maripaludis*의 조효소(coenzyme) F420-reducing hydrogenase (YP_001097176) (33%) 및 *Methanococcus maripaludis* S2의 조효소 F420-reducing hydrogenase (NP_987940) (33%)와의 유사성을 나타낸다(도 3 및 도 5). *Thermococcales*의 F420-hydrogenase에 상동성이 있는 것(homologue)들은 아직 보고되지 않았다.

[0070] 3) 3-모듈 유전자 클러스터들의 구성

[0071] [NiFe]-수소화효소의 그룹 4에 속하는 상기 세 개의 Ech 수소화효소들(MFH1, MFH2, MCH) 각각은 TON_266-TON_282, TON_1563-TON_1580 및 TON_1016-TON_1031 ORFs(open reading frames) 로 이루어진 큰 17- 또는 18-유전자 클러스터(*fdh1-mfh1-mnh1*, *fdh2-mfh2-mnh2* 및 *codh-mch-mnh3*)의 일부임이 밝혀졌다(도 2의 B). 클러스터 내의 ORF들은 세 개의 서브클러스터들로 나뉠 수 있다. 첫 번째 부분은 포름산염 탈수소화효소(formate dehydrogenases)(Fdh1(서열번호 9) 또는 Fdh2(서열번호 11)) 또는 일산화탄소 탈수소화효소(carbon monoxide dehydrogenase)(Codh(서열번호 10))와 같은 산화환원효소(oxidoreductase)들을 암호화하고 있다. 두 번째 파트는 5개에서 7개의 서브유닛들을 갖는 다중 막-결합 수소화효소들(multimeric membrane-bound hydrogenases)(MFH1, MFH2 또는 MCH)을 암호화한다. 마지막 부분은 Na⁺/H⁺ 안티포터(antiporter)와 같은 양이온/수소이온 안티포터들(cation/proton antiporters)을 암호화한다. 이러한 3 모듈 유전자 클러스터는 아직 보고된 바 없다.

[0072] 실시예 2. 가스 조성의 분석

[0073] (1) 분석방법

[0074] 수소가스는 HP-PLOT Molesieve column (Agilent) 및 TCD detector를 갖춘 gas chromatograph HP 5890 series II (Hewlett Packard)를 이용하여 측정하였다. 아르곤을 가스 운반체로 사용하였다. 수소가스를 정량하기 위하여 질소에 각각의 성분(CO, CO₂, H₂, CH₄ 및 O₂) 1%(w/w) 가 포함된 Gas calibration standard (Supleco) 를 사용하였다.

[0075] (2) 다양한 물질들을 이용한 수소 생성

[0076] 반복된 다수의 수소화효소가 다양한 환경에서 *T. onnurineus* NA1이 수소를 효율적으로 생산하게 하는지 알아보기 위하여, 다양한 에너지원을 이용하여 수소 생성율을 분석하였다(표 2). 그 결과, NA1 균주는 황이 없는 조건에서도 녹말, CO, 및 포름산염(formate)을 이용하여 수소를 만들 수 있었다. 특히 CO 및 포름산염은 수소를 효율적으로 생산하는데 매우 좋은 에너지원이었다.

[0077] [표 2]

[0078] 다양한 조건하에서의 *T. onnurineus* NA1 의 수소 생성 비교

매지	수소 생성(μmol/l)
M + CO	30.7
M + 포름산염(Formate)	49.7
M + 녹말	15.6

M. 변형 매지(modified medium) 1

[0079]

[0080] 회분식 배양 (batch culture)에서 NA1 균주의 수소 생산성은 *T. kodakaraensis* KOD1 및 *Pyrococcus furiosus*의 연속식 배양(continuous culture)에서 얻어지는 것과 유사하다. 고호열성 고세균(hyperthermophilic archaea)은 낮은 체적생산성(volumetric productivity)에도 불구하고, 중온성 세균의 발효(mesophilic bacterial fermentation)나 광합성 원핵생물(photobacteria)에 의한 수소생성에 비하여 높은 특이적 생산율 (specific production rate)을 보이며 고순도의 수소 생산 등과 같은 여러 장점을 갖고 있다. 본 명세서에 기재된 높은 수소 생성율은 배양조건 및 처리과정 설계의 최적화 또는 대사공학에 의해 더 향상될 여지가 많이 있다.

[0081] 실시예 3. *Thermococcus onnurineus* NA1에 의한 CO-의존적 H₂ 생성: CO-반응(responsive) 수소화효소의 동정

[0082] (1) CODH 유전자 클러스터 및 일산화탄소영양성장 (carboxydrotrophic growth)

[0083] 상술한 바와 같이, *T. onnurineus* NA1은 특이적인 유전자 클러스터를 갖고 있음을 발견했으며(CODH-MCH-MNH3), 이 유전자 클러스터는 수소화효소 (mch, TON_1021-1024, 서열번호 5)와 함께 잠정적인 전사조절인자 (TON_1016), CODH 부속 단백질(accessory protein)(CooC, TON_1019), CODH 촉매 서브유닛 (catalytic subunit)(CooS, TON_1018), 및 전자전달 단백질(CooF, TON_1017)을 암호화하는 염기서열로 이루어졌다(도 2의 B). 미생물의 일산화탄소(CO) 대사의 중심 효소인 CooS (TON_1018)는 *Methanosarcina acetivorans* C2A의 CODH (AAM06652) (60%) 및 *Methanosarcina mazei* Go1의 CODH (AAM29817) (59%)와 같은 CO-산화 메탄생성 고세균 (CO-oxidizing methanogenic archaea)의 CODH들과 상당한 유사도를 보였다(도 3 및 도 4). 이것은 두

기능을 갖는(bifunctional) CODH/ACS 효소의 아세틸-CoA 합성/절단 활성(acetyl-CoA synthesis/cleavage activity) 이 없는 단일 기능의 CODH (Bonam, D., L. Lehman, G. P. Roberts, and P. W. Ludden. 1989. Regulation of carbon monoxide dehydrogenase and hydrogenase in *Rhodospirillum rubrum*: effects of CO and oxygen on synthesis and activity. J. Bacteriol. 171:3102-3107. 및 Wu, M. Q. Ren, A. S. Durkin, S. C. Daugherty, L. M. Brinkac, R. J. Dodson, R. Madupu, S. A. Sullivan, J. F. Kolonay, W. C. Nelson, L. J. Tallon, K. M. Jones, L. E. Ulrich, J. M. Gonzalez, I. B. Zhulin, F. T. Robb, and J. A. Eisen. 2005. Life in hot carbon monoxide: the complete genome sequence of *Carboxydotherrmus hydrogenoformans* Z-2901. PLoS Genet. 1:e65.)인 것으로 보인다. Fox *et al.* (Fox, J. D., R. L. Kerby, G. P. Roberts, and P. W. Ludden. 1996. Characterization of the CO-induced, CO-tolerant hydrogenase from *Rhodospirillum rubrum* and the gene encoding the large subunit of the enzyme. J. Bacteriol. 178:1515-1524.) 에 따르면, *Rhodospirillum rubrum* 의 단일기능 CODH/수소화효소 복합체 (monofunctional CODH/hydrogenase complex) 는, 세포막을 가로질러 발생하는 수소이온 농도 구배로 에너지가 보존되는 CO-유발 양성자 호흡 (CO-driven proton respiration) 에 관여한다. 이러한 관점에서, CODH 클러스터가 CO를 산화시킴으로써 에너지 보존에 유사한 역할을 할 수 있다는 점을 설명하기 위해, 본 발명자들은 *T. onnurineus* NA1 이 CO를 이용할 수 있는지를 시험하였다. 그 결과 상기 균주는 실제로, YPS 배지에서보다는 성장률이 다소 떨어지더라도(도 6a), 기본배지(basal medium)에서 보다 배지 1(medium 1)에서 황이 있거나 없는 CO기체 환경에서 더 잘 자란다는 것을 발견했다(도 6a, b). CO기체 환경에서의 성장은 *CooS* 유전자의 전사와 관련되어 있으며, 그 유전자는 CO의 존재로 유도될 수 있다(도 6c). 황의 첨가는 *CooS* 유전자의 전사를 감소시킴을 주목할 필요가 있다. 이러한 결과들은 *T. onnurineus* NA1이 CO로부터 에너지를 발생시킨다는 가설을 뒷받침한다. 이하 상기 가설을 검증하기 위한 구체적인 실험방법 및 결과를 추가로 기술한다.

[0084] (2) 실험방법

[0085] 1) 배양조건

[0086] YPS(yeast extract-peptone-sulfur) 배지에서 80℃의 혐기성 조건으로 *T. onnurineus* NA1을 배양하였다. 돌연변이 균주의 성장 특성을 조사하기 위해, 기본배지로 30.0g/L 의 NaCl을 첨가한 변형배지 1을 사용하였다 (Uffen, R. L. 1976. Anaerobic growth of a Rhodopseudomonas species in the dark with carbon monoxide as sole carbon and energy substrate. Proc. Natl. Acad. Sci. USA 73:3298-3302). 상기 배지를 가압멸균시킨 후, 혐기성 챔버에서 1.0ml/L 의 비타민 용액(Balch, W. E., G. E. Fox, L. J. Magrum, C. R. Woese, and R. S. Wolfe. 1979. Methanogens: reevaluation of a unique biological group. Microbiol. Rev. 43:260-296) 및 0.5g/L의 효모추출물(yeast extract)을 변형배지 1에 첨가하였다. 상기 기본 배지에 1N NaOH을 첨가하여 pH를 8.0으로 맞췄다. 준비된 배지 30ml를 150ml 배양병(serum bottle)에 분배시켰고, 가스층(120ml)은 100% CO (10⁵ Pa)로 바꾸었다. 생리학적 실험을 위한 모든 배양은 80℃의 혐기성 조건에서 24시간 동안 이루어졌으며, 상기 실험들은 중복하여 수행하였다.

[0087] 2) RNA 추출 및 마이크로어레이 설계

[0088] 배양물들을 4℃에서 3,500×g 로 10분간 원심분리하고, 생산자의 프로토콜(Invitrogen, Carlsbad, California)에 따라 TRIZOL 시약으로 총(total) RNA를 추출하였다. 총 RNA 시료들은 NanoDrop 분광기(ND-1000, Thermo Scientific) 및 전기영동으로 정량 및 정성 분석하였다. 본 실험에 사용된 마이크로어레이는 Roche NimbleGen 마이크로어레이를 사용하였다. 간략하게, 총 31,994 개의 독립적인 60mer 올리고뉴클레오티드들을 설계하고, 광탈분해 (photo deprotection chemistry)를 이용하여, 인 시츄(in situ)로 합성하였다. 각각의 독립적인 올리고뉴클레오티드는 어레이 상에서 2번 반복되었다(총 ~72,000 features).

[0089] 3) cDNA 합성 및 혼성화(hybridization) 조건

[0090] 마이크로어레이 실험은 생산자의 프로토콜에 따라 실시하였다. 각각의 총 RNA 시료(5 μg)는 역전사효소, SuperScrip II (Invitrogen, Carlsbad, California)를 이용한 역전사 반응에 의해 시아닌(Cy5) 결합 dCTP (Amersham, Piscataway, NJ)로 표지되었다. 그 후 상기 표지된 cDNA 혼합물을 에탄올 침전법을 이용해 농축시켰다. 농축된 Cy5 표지 cDNAs 를 30μl 의 혼성화 용액(GenoCheck, Korea)에 현탁시켰다. 표지된 cDNA들을 마이크로어레이에 위치시킨 후, MAUI Mixer X4 hybridization chamber mixers (BioMicro systems, Salt Lake City, UT) 로 덮었다. 상기 슬라이드들을 MAUI 12-bay systems (BioMicro systems, Salt Lake City, UT)로 42° C에서 12시간 동안 혼성화시켰다. 상기 혼성화된 슬라이드들을 실온에서 2 X SSC, 0.1% SDS 에서 2분, 1X SSC 에서 3분, 그 후 0.2 X SSC 에서 2분 동안 세척하였다. 상기 슬라이드들을 말리기 위해 1,000 x g에서 20초간 원심분

리하였다.

- [0091] 4) 슬라이드 스캐닝(scanning), 표준화(normalization), 및 데이터 분석
- [0092] 어레이는 GenePix 4000B scanner (Molecular Devices Corporation, Union City, CA)를 사용하여 스캔하였고, 데이터는 NimbleScan 2.4 소프트웨어를 사용하여 추출하였다. 어레이 표준화는 메디안 폴리쉬 및 콰antil 표준화 (median polish and quantile normalization) 방법을 사용하여 수행하였다(Amaratunga, D., and J. Cabrera. 2001. Statistical analysis of viral microchip data. J. Am. Stat. Assoc. 96:1161-1170). 개별적인 프로브들에 대한 표준화된 발현 값은, Irizarry et al. (Karl, D. M. 1995. The microbiology of deep-sea hydrothermal vents. CRC Press, Boca Raton, FL)에 기술된 바와 같은 RMA (robust multi-array average) 방법을 사용하여, 주어진 ORF에 대한 발현 값을 얻는데 사용되었다. 마지막으로, 시료에서 특정 유전자를 위해 얻은 RMA-처리된 발현 값들(RMA calls)을 이용하여 n배 변화율(n-fold change ratios)(R)을 계산하였다. 데이터 분석은 GeneSpring GX 7.3.1 (Agilent technologies, CA) 를 이용하여 수행하였다. 폴드 변화 필터(Fold change filters)는 상향조절(up-regulated)되는 유전자들에 대해서는 적어도 대조구의 200%, 하향조절되는 유전자들에 대해서는 대조구의 50% 이하로 나타나야 한다는 조건을 포함시켰다. 데이터는 GeneSpring GX 7.3.1 (Agilent technologies, CA)을 사용한 실험들에서 유사하게 행동하는 유전자 그룹들 내에 클러스터시켰다. 유사한 패턴의 유전자를 분리하기 위하여 유클리드 거리(Euclidian distance) 및 평균결합법(average linkage)에 근거한 알고리즘을 사용하였다.
- [0093] 5) 정량적 RT-PCR
- [0094] *T. onnurineus* NA1 (Genbank accession number, CP000855) 의 유전체 서열로부터 유전자 특이적 프라이머들을 설계하였다. 프라이머 서열은 하기 표 3에 기재되어 있다.
- [0095] cDNA들은 역전사효소, SuperScrip II (Invitrogen, Carlsbad, California)를 사용하여 생산자의 프로토콜에 따라 350ng의 총 RNA로부터 합성하였다. PCR 반응은 *T1* thermocycler (Biometra)를 사용하여 *rTaq* (Takara) DNA 중합효소로 수행하였다. 상기 반응은 1차-사슬 반응(first-strand reaction) cDNA, 10 pmol 의 프라이머들, 250 μM dNTPs 및 생산자의 버퍼를 포함하는 50 μl 혼합물에서 이루어졌다. 94°C에서 2분간의 변성 (denaturation) 단계 후에, 변성 (30초, 94° C), 어닐링(annealing) (30초, 56° C) 및 신장반응(extension) (1분, 72° C)이 25 사이클 수행되었다. PCR 산물은 0.8% 아가로스젤 전기영동을 사용하여 분석하였다. 발현수준은 GelPro32.EXE v4.6 (Media Cybernetics, Inc.) 을 사용하여 측정하였다. *cha*로 명명된 샤페로닌 (Chaperonine)-암호화 유전자를 발현 수준을 표준화하기 위한 대조구로 사용하였다.
- [0096] [표 3]

[0097] 본 연구에 사용된 올리고뉴클레오티드

유전자	센스(5'→3')	안티-센스(5'→3')
유전자 파괴에 사용된 프라이머		
<i>mch</i>	cacgacatagactacgacacgg (서열번호 27)	ctgcttaactcctaagtcagg (서열번호 28)
<i>mcr</i>	gcgattcggatgataccggac (서열번호 29)	ccatccttcaccgaagactcg (서열번호 30)
<i>frh</i>	gtaagctcagcagatcagcgtg (서열번호 31)	gcaccacaacctatgcaagggcc (서열번호 32)
<i>sulfi</i>	gcagtcagcagagtcgagggg (서열번호 33)	gaggactctgctgataaggtcg (서열번호 34)
<i>mch</i>	ctaccgacgattgcccagaagg (서열번호 35)	ccttatatactgtgctctctccg (서열번호 36)
<i>mfh1</i>	gcgaccgctacgcaaccttcg (서열번호 37)	cttgctcagtcacgtagtagg (서열번호 38)
<i>mfh2</i>	gaccgagctcacctcgaatgc (서열번호 39)	gcagacctgctcgaagttagcc (서열번호 40)
유전자 파괴에 사용된 프라이머		
<i>Flk-mch</i>	cgttctcttcccttggcagggatata tc (서열번호 41)	gcaattgcttggactgccgaaagccaatgg (서열번호 42)
<i>Flk-mfh1</i>	gaagaatcgagagggccttatgactat cag (서열번호 43)	gctcctcgcttactcaagcattgacaaatgg (서열번호 44)
<i>Flk-mfh2</i>	ggactgcttctctcgcacggctcaata ttc (서열번호 45)	ggacgacttaagtcggctagcccttggc (서열번호 46)
<i>Ivs-mch</i>	aatttacccccaccatccccaaatcca ac (서열번호 47)	aatggggagctgaactactggcagggc (서열번호 48)
<i>Ivs-mfh1</i>	tggcccagggatttctctcaccgacagg (서열번호 49)	aattcaccaccaccacagcgtattatcagg (서열번호 50)
<i>Ivs-mfh2</i>	gagcaccacctcaccatcccagggagcta tc (서열번호 51)	gatggcctgacgctgaagtagcccttcgta (서열번호 52)
유전자 결손의 확인에 사용된 프라이머		
<i>P_{gdh}-hmg_{Pfu}</i>	gacgtagttttcgacaaaagacg (서열번호 53)	gctcaccagccaaaaccgaccaaac (서열번호 54)
<i>mch₂₀₁</i>	gcaatgtaccacatatcaactgcgatac (서열번호 55)	ccgataccgagttgaaaggagaaatctc (서열번호 56)
<i>mfh1₂₀₁</i>	tcagccaccccccttgccttctgt (서열번호 57)	atgagtcagcgtgtgtcgaggtg (서열번호 58)
<i>mfh2₂₀₁</i>	atgtctgaggttatcaagtttaacg (서열번호 59)	tgggaccttatggagagcttcttg (서열번호 60)

[0098]

[0099] 6) 표적 유전자 제거(Targeted gene disruption)

[0100]

T. onnurineus NA1의 생체 내(*in vivo*)에서의 수소화효소의 기능을 분석하기 위해, 고호열성 고세균(archaeon)인 *T. kodakaraensis* KOD1에 사용된 유전자 제거 시스템(gene disruption system)을 이용하여(Sapra, R., K. Bagramyan, and M. W. W. Adams. 2003. A simple energy-conserving system: Proton reduction coupled to proton translocation. Proc. Natl. Acad. Sci. USA 100:7545-7550), *mch* 또는 *mfh2* 수소화효소의 큰 서브유닛의 삽입 불활성 돌연변이(insertional inactivation mutant)들을 제작하였다. 구체적으로, *mch* 및 *mfh2* 수소화효소 각각의 큰 서브유닛(TON_023 및 TON_1569)의 인접지역(flanking region)을 포함하는 DNA 단편들을 Flk-*mch* 또는 Flk-*mfh2*에 대한 프라이머 세트들(표 3)을 사용하여 *T. onnurineus* NA1의 지놈 DNA로부터 증폭시켰다. 각각의 증폭된 단편들은 *Hinc*II 로 절단된 pUC118에 접합시켰다. 그 후, 주형(Flk-*mch*_pUC118 또는 Flk-*mfh2*_pUC118 재조합 플라스미드)으로부터 프라이머 세트(Ivs-*mch* 또는 Ivs-*mfh2*)(표 3)들을 사용하여 역(inverse) PCR을 수행하고, 이어서 P_{gdh}-hmg_{Pfu} 카세트에 접합시켰다(Sapra, R., K. Bagramyan, and M. W. W. Adams. 2003. A simple energy-conserving system: Proton reduction coupled to proton translocation. Proc. Natl. Acad. Sci. USA 100:7545-7550). 그 후, 상기 접합된 결과물을 *Escherichia coli* DH5α 세포들에 형질전환시켰다. 재조합 플라스미드(*mch*:P_{gdh}-hmg_{Pfu} 및 *mfh2*:P_{gdh}-hmg_{Pfu})들은 플라스미드 미니 키트(Qiagen, Hilden, Germany)로 준비하였다. 마지막으로, 상기 플라스미드들을 Sato et al.(Sato, T., T. Fukui, H. Atomi, and T. Imanaka. 2003. Targeted gene disruption by homologous recombination in the

hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Bacteriol.* 185:210-220., Sato, T., T. Fukui, H. Atomi, and T. Imanaka. 2005. Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. *Appl. Environ. Microbiol.* 71:3889-3899)의 약간 변형된 방법으로 *T. onnurineus* NA1에 형질전환시켰다. 형질전환체들은 10 μ M 심바스타틴(simvastatin)의 존재 하에 ASW-YT-S 배지에 의해 스크리닝되었고(Matsumi, R., K. Manabe, T. Fukui, H. Atomi, T. Imanaka. 2007. Disruption of a sugar transporter gene cluster in a hyperthermophilic archaeon using a host-marker system based on antibiotic resistance. *J. Bacteriol.* 189:2683-2691), 표적 유전자가 제거된 것으로 생각되는 후보군들은 PCR 증폭을 통해 표적 지역내 $P_{gdh-hmgPfu}$ 카세트의 존재여부로 제거여부를 확인할 수 있었다.

[0101] 7) 성장 및 수소 생성에 관한 동역학적 분석

[0102] 성장은 눈으로 직접 확인하였다. 시료들을 해염(Sea Salt) (30.0 g/L), 포르말린(2.5%), 및 4' -6' -디아미디노-2-페닐인돌(0.01%)를 포함하는 멸균수에 희석시키고(Sato, T., T. Fukui, H. Atomi, and T. Imanaka. 2005. Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. *Appl. Environ. Microbiol.* 71:3889-3899), 블랙 폴리카보네이트 막 필터로 거른 후(구멍 크기, 0.2 μ m; Whatman), 위상차현미경(Zeiss Axioplan)으로 썼다. 수소 가스는 HP-PLOT Molesieve column (Agilent) 및 TCD 검출기가 구비된 가스 크로마토그래프 HP 5890 series II (Hewlett Packard)를 사용하여 측정하였다. 아르곤이 가스 운반자로 사용되었다. 오븐 온도는 40 $^{\circ}$ C 였다. 분석을 위한 10 μ l의 가스 시료는 배양병의 부틸고무를 통해 가스-타이트 시린지(gas-tight syringe)로 꺼냈다. 검출된 수소 가스의 측정은 피크 면적을, 질소에 40% 수소를 포함하는 표준 가스를 사용한 희귀분석에 의해 수행된 보정 곡선과 비교함으로써 계산되었다.

[0103] (3) 실험결과

[0104] 1) *T. onnurineus* NA1 수소화효소의 *in silico* 분석

[0105] 앞선 *T. onnurineus* NA1의 유전체 분석은 8 반복의 수소화효소 유전자 클러스터의 존재를 보여주었으며(Porter, K. G. and Y. S. Feig. 1980. The use of DAPI for identifying and counting microflora. *Limnol. Oceanogr.* 25:943-948), 5개의 막-결합(membrane-bound) [NiFe]-수소화효소 (Mbh, TON_1583-1595; Mbx, TON_0486-0498; Mfh1, TON_0273-0278; Mfh2, TON_1566-1573; Mch, TON_1021-1024), 및 3개의 세포질내(cytoplasmic) [NiFe]-수소화효소 (Fru, TON_1559-1562; Sulf I, TON_0533-0544; Sulf II, TON_0051-0055)를 포함한다. 지놈 시퀀싱이 완료된 Thermococcales 균주들과의 수소화효소 유전자 클러스터 비교 분석 결과를 볼 때, Mfh1, Mfh2, 또는 Mch 클러스터들에 상동성이 있는 클러스터들은 매우 드물어 보이고, *T. barophilus* MP (Mfh1, Mch 상동유전자), *Thermococcus* sp. AM4 (Mfh1, Mch 상동유전자) (Unfinished sequence, GenBank accession number, ABXN00000000), 및 *T. gammatolerans* (Mfh1 및 Mfh2 상동성) (GenBank accession number, CP001398)와 같이 최근에 지놈 서열이 결정된 Thermococcales 균주들에서 발견된다. *T. onnurineus* NA1의 *fdh1-mfh1-mnh1* (Lee, H. S., S. G. Kang, S. S. Bae, J. K. Lim, Y. Cho, Y. J. Kim, J. H. Jeon, S.-S. Cha, K. K. Kwon, H.-T. Kim, C.-J. Park, H.-W. Lee, S. I. Kim, J. Chun, R. R. Colwell, S.-J. Kim, 및 J.-H. Lee. 2008. The complete genome sequence of *Thermococcus onnurineus* NA1 reveals a mixed heterotrophic and carboxydrotrophic metabolism. *J. Bacteriol.* 190:7491-7499.의 논문에서 Hyg4-I 으로 명명됨.), *fdh2-mfh2-mnh2* (Hyg4-III), 및 *codh-mch-mnh3* (Hyg4-II) 클러스터들의 서열분석은 상기 클러스터들이 각각 포름산염 탈 수소화효소(FDH) 및 CO 탈수소화효소(Codh)와 같은 산화환원효소들을 포함한다는 사실을 보여준다. 특히, CO 조건에서의 성장에 의해 나타나는 일산화탄소영양 대사는 CO를 산화시킴으로써 수소생성 경로에서 에너지를 제공하는 Codh-Mch-Mnh3의 기능적 역할을 암시한다.

[0106] 2) CO-유도(driven) 성장 조건하에서의 수소화효소 유전자의 발현

[0107] *T. onnurineus* NA1 이 CO 조건에서 성장하면서 수소를 생산할 수 있는지 알아보았다. 도 7에 나타난 바와 같이, YPS배지에서는 수소 생산을 탐지할 수 없었지만, CO가 첨가된 배지 1에서는 배양시간이 길어질수록 총 수소 가스와 세포 수가 증가하였다.

[0108] *그 후, 어떤 수소화효소가 일산화탄소영양 성장 동안에 수소 생성에 관여하는지 알아보기 위하여 CO가 포함된 성장조건 또는 복합배지(YPS)에서의 수소화효소 유전자들의 발현 수준을 분석하였다. 하기 표 4, 표 5 및 도 8A

에 나와 있듯이, *codh-mch-mnh3* 클러스터의 일부 ORF들(16개 중 10개의 ORF들)의 발현 수준은 CO를 포함한 성장 조건에서 YPS에 비해서 2배 이상 상향조절(up-regulated)되었다. 추가로, *fdh2-mfh2-mnh2* 클러스터 내 여러 ORF들(TON_1563, TON_1569-1571)의 발현 수준 역시 1g의 효모추출물을 포함한 CO 조건에서 상향조절되었다. *codh-mch-mnh3* 클러스터 내의 ORF들의 발현 수준 역시 효모추출물의 양에 따라 다양하다는 것을 주목할 필요가 있고, 이는 CO-유도 대사작용(CO-driven metabolism)에서 이화작용의 억제 또는 활성화에의 관련성을 암시한다(표 4 및 표 5). 반면에, 다른 수소화효소 유전자 클러스터들의 발현은 큰 변화가 없는 반면 *fdh1-mfh1-mnh1* 클러스터의 유전자들(29개 ORF들 중 20개의 유전자들)은 하향조절을 보였다. 수소화효소 각각의 큰 서브유닛에 대한 정량적 RT-PCR 데이터 역시 마이크로어레이 데이터와 일치하였다. *mch* 및 *mfh2* 수소화효소의 큰 서브유닛(TON_1023 및 TON_1569)의 발현은 2배 이상 증가한 반면(도 8B), *mfh1* 수소화효소의 큰 서브유닛(TON_0276)의 발현은 억제되었고, 다른 큰 서브유닛들(*mbh*, *mbx*, *frh*, 및 *sulf I*)은 양 쪽 조건에서 모두 일정하게 유지되었다. 이러한 결과들은 *codh-mch-mnh3* 또는 *fdh2-mfh2-mnh2* 클러스터들이 CO에 의해 유도될 수 있으며, 일산화탄소 영양 대사와 관련된 수소 생성 과정에 관련될 수 있다는 점을 암시한다.

[0109] 하기 표 4는 YPS 배지에서와 비교한 CO 성장 조건에서의 수소화효소 유전자 클러스터 각각의 ORF 발현 수준을 나타내며, 표 5는 *T. onnurineus* NA1 의 수소화효소 클러스터를 포함한 112개 ORF들의 계층적 클러스터링을 보여준다.

[0110] [표 4]

Gene cluster and ORFs	Annotation	M+CO+0.5g YE /YPS	M+CO+1g YE /YPS	M+CO+3g YE /YPS
<i>sufII</i> hydrogenase cluster				
TON_0051	hypothetical protein	1.273306	3.1454	2.569546
TON_0052	hydrogenase subunit alpha	1.085002	0.652365	0.550585
TON_0053	Sulfhydrogenase II, delta chain	0.896129	1.898148	1.337585
TON_0054	hydrogenase subunit gamma	0.606688	0.517724	0.630549
TON_0055	sulfhydrogenase II, beta chain	0.84527	1.44849	1.452254
<i>MhA-mfhI-mnhI</i> cluster				
TON_0261	hypA hydrogenase expression/formation protein	0.658409	0.78969	0.746964
TON_0262	ATPase involved in chromosome partitioning, ParA/MinD family, Mtp homolog	0.640277	0.776389	0.619674
TON_0263	hydrogenase maturation protease HycI	0.775719	1.531337	1.090123
TON_0264	molybdopterin-guanine dinucleotide biosynthesis protein A	0.538102	0.957335	0.826924
TON_0265	Nucleotidyltransferase, putative	0.421398	0.554433	0.563557
TON_0266	component F or formate hydrogen lyase	0.38336	0.338185	0.333459
TON_0267	Putative integral membrane protein, DUF68 family	0.296945	1.215972	0.838087
TON_0268	Putative integral membrane protein, DUF67 family	0.648077	0.424896	0.386235
TON_0269	Multisubunit Na ⁺ /H ⁺ antiporter, putative MnhB subunit	0.591411	0.571996	0.600235
TON_0270	hypothetical protein	0.827927	0.85379	0.581906
TON_0271	Na ⁺ /H ⁺ antiporter subunit	0.410417	0.328062	0.51901
TON_0272	Na ⁺ /H ⁺ antiporter MnhF subunit, putative	0.194663	0.411647	0.331549
TON_0273	hypothetical protein	0.269222	0.652722	0.482705
TON_0274	component I or formate hydrogen lyase	0.385799	0.627629	0.614487
TON_0275	formate hydrogen lyase subunit 6	0.313283	0.621333	0.511819
TON_0276	component O or formate hydrogen lyase	0.275138	0.724228	0.692331
TON_0277	component C or formate hydrogen lyase	0.460646	0.357836	0.387126
TON_0278	component B or formate hydrogen lyase	0.720779	0.623875	0.510273
TON_0279	NADH ubiquinone oxidoreductase	0.80929	0.515429	0.54289
TON_0280	oxidoreductase iron-sulfur protein	0.3913	0.690322	0.647104
TON_0281	fdhA formate dehydrogenase, alpha subunit	0.584576	0.446401	0.44946
TON_0282	putative transcriptional regulator	1.340444	1.195043	0.863294
TON_0283	hypC hydrogenase expression/formation protein	1.007318	1.853137	2.021914
TON_0284	367aa long hypothetical hydrogenase expression/formation protein hypD	1.036497	1.385466	1.182663
TON_0285	hypD hydrogenase expression/formation protein	0.904687	1.066448	0.758981
TON_0286	hydrogenase maturation protein HypF	0.685606	0.778307	0.719973
TON_0287	hydrogenase expression/formation protein HypE	0.851301	1.399569	1.257059
TON_0288	hypothetical protein	1.004008	1.964442	2.27641
TON_0289	cysteine desulfurase	0.666768	1.534991	1.148004
<i>mnh</i> hydrogenase cluster				
TON_0486	4Fe-4S cluster-binding subunit	0.821271	0.718277	0.504388
TON_0487	mnoD NADH dehydrogenase I, subunit D	0.606823	0.730138	0.517528
TON_0488	NADH dehydrogenase subunit	0.703506	0.618352	0.49939
TON_0489	mnoB NADH dehydrogenase I, subunit B	0.797159	0.618863	0.547462
TON_0490	NADH dehydrogenase subunit	0.372425	0.469467	0.278163

[0111]

TON_0491	617aa long hypothetical protein	0.459201	0.532164	0.388201
TON_0492	MfxH subunit	0.862715	0.484349	0.423606
TON_0493	Multisubunit Na ⁺ /H ⁺ antiporter, putative	0.70216	1.591408	1.138149
TON_0494	Multisubunit Na ⁺ /H ⁺ antiporter, putative MfxB subunit	0.765286	0.716126	0.711304
TON_0495	MfxD subunit	0.658086	0.817339	0.749402
TON_0496	MfxC subunit	0.609819	1.200655	0.995067
TON_0497	MfxB subunit	0.860304	0.889029	0.810902
TON_0498	Multisubunit Na ⁺ /H ⁺ antiporter, putative	0.675578	0.564247	0.483617
sdh1 hydrogenase cluster				
TON_0533	hydrogenase-specific maturation endopeptidase	0.977803	0.516234	0.73084
TON_0534	cytosolic NifH-hydrogenase, alpha subunit	1.283309	1.334781	1.071667
TON_0535	cytosolic NifH-hydrogenase, delta subunit	0.652099	1.080516	0.800955
TON_0536	hydrogenase (gamma subunit)	1.100959	1.192121	0.759086
TON_0537	Sulfhydrogenase beta subunit	1.025726	1.458113	1.160718
TON_0538	probable formate transporter	0.989789	1.113613	0.960108
TON_0539	probable formate dehydrogenase, alpha subunit	1.158997	1.791123	1.314525
TON_0540	oxidoreductase iron-sulfur protein	0.987581	1.525167	0.82485
TON_0541	4Fe-4S cluster-binding protein	1.112648	1.320761	0.855465
TON_0542	glutamate synthase beta chain-related oxidoreductase	0.842213	1.602828	0.936949
TON_0543	4Fe-4S cluster-binding protein	0.9185	1.257283	1.019564
TON_0544	alcohol dehydrogenase	0.87473	0.419323	0.605934
codB-mcb-asmA3 cluster				
TON_1016	putative transcriptional regulator, ModE family	1.187798	0.634394	0.731218
TON_1017	4Fe-4S ferredoxin, iron-sulfur binding domain	1.713085	4.46107	3.621394
TON_1018	carbon-monoxide dehydrogenase, catalytic subunit	1.474238	2.204944	1.785437
TON_1019	putative ATP-binding protein	0.824732	2.016924	0.974466
TON_1020	hypothetical protein	3.375546	9.366047	5.425704
TON_1021	component B or formate hydrogen lyase	1.112538	1.711122	1.250867
TON_1022	respiratory-chain NADH dehydrogenase, subunit 1	0.877363	1.428883	0.57715
TON_1023	component G or formate hydrogen lyase	1.927395	6.118015	4.270288
TON_1024	NADH dehydrogenase (ubiquinone), 20 kDa subunit	1.816315	3.578841	2.009973
TON_1025	Na ⁺ /H ⁺ antiporter MfxF subunit, putative	2.443938	3.389037	1.333503
TON_1026	Na ⁺ /H ⁺ antiporter subunit	1.337719	0.605839	0.623499
TON_1027	hypothetical protein	0.991048	1.43866	0.778446
TON_1028	Multisubunit Na ⁺ /H ⁺ antiporter, putative MfxB subunit	1.019583	1.203407	0.836452
TON_1029	Putative integral membrane protein, DUF67 family	2.648286	2.171138	1.123898
TON_1030	Putative integral membrane protein, DUF68 family	3.779798	4.956998	1.592196
TON_1031	component F or formate hydrogen lyase	1.030507	2.788588	1.982091
fth (F₄₂₀ reducing hydrogenase) hydrogenase cluster				
TON_1559	coenzyme F420 hydrogenase alpha subunit	0.695408	0.64447	0.726535
TON_1560	Coenzyme F420 hydrogenase gamma subunit	0.56769	0.715513	0.635236
TON_1561	4Fe-4S ferredoxin, iron-sulfur binding/Nitrite/sulfite reductase	0.824149	0.834127	0.807754
TON_1562	Formate dehydrogenase, subunit F4&D	0.908701	1.54082	1.468288
Fdh3-mfh2-asmA2 cluster				
TON_1563	Probable formate dehydrogenase, alpha subunit	1.008259	3.037865	3.330042
TON_1564	4Fe-4S cluster-binding protein	1.178705	1.221574	1.445098

[0112]

TON_1565	NADEH ubiquinone oxidoreductase	0.811937	1.787894	2.088693
TON_1566	NADEH dehydrogenase (quinone)	1.011779	0.736029	0.787627
TON_1567	component B or formate hydrogen lyase	1.091992	1.057076	1.184573
TON_1568	component C or formate hydrogen lyase	1.471267	1.365925	1.382301
TON_1569	component O or formate hydrogen lyase	1.204795	2.281342	1.780033
TON_1570	formate hydrogen lyase subunit 6	1.257649	2.427962	1.99559
TON_1571	component I or formate hydrogen lyase	1.030626	2.15871	1.806722
TON_1572	hypothetical protein	0.844948	0.504955	0.692272
TON_1573	probable formate transporter	1.451628	1.459657	1.158131
TON_1574	Na ⁺ /H ⁺ antiporter MnhF subunit, putative	1.53015	1.602287	1.244315
TON_1575	Na ⁺ /H ⁺ antiporter subunit	0.752784	0.646406	0.578467
TON_1576	hypothetical protein	1.168966	0.564328	0.60819
TON_1577	MnhK subunit Na ⁺ /H ⁺ antiporter, putative MnhB subunit	0.726511	0.655031	0.556386
TON_1578	Putative integral membrane protein, DUF67 family	0.989286	1.041599	0.900044
TON_1579	Putative integral membrane protein, DUF68 family	0.903028	0.812346	0.816319
TON_1580	component F or formate hydrogen lyase	1.015786	0.756283	0.529199
TON_1581	molybdopterin-guanine dinucleotide biosynthesis protein A	0.61468	1.708744	1.243367
TON_1582	hypothetical protein	0.753301	0.998059	0.952751
mnh hydrogenase cluster				
TON_1583	MnhB subunit	1.038442	0.519366	0.533773
TON_1584	MnhC subunit	0.947559	0.514932	0.5099
TON_1585	MnhD subunit	0.774604	1.015909	0.670497
TON_1586	MnhE subunit	1.101069	0.811262	0.659266
TON_1587	MnhF subunit	1.129566	0.811366	0.872572
TON_1588	MnhG subunit	1.027148	0.563686	0.510924
TON_1589	MnhH subunit	1.362042	0.948034	0.71984
TON_1590	MnhI subunit	1.334732	1.227655	0.904813
TON_1591	NiFe-hydrogenase small subunit	1.627315	1.158127	0.856929
TON_1592	NiFe-hydrogenase large subunit 1	0.940698	1.573248	0.956935
TON_1593	NiFe-hydrogenase large subunit 2	1.609483	2.286207	1.354849
TON_1594	MnhM subunit	0.982323	0.742211	0.458623
TON_1595	4Fe-4S cluster-binding subunit	1.589684	1.280298	0.491906

M, medium 1; CO, carbon monoxide; YE, yeast extract.

[0113]

[0114]

[표 5]

Gene cluster	ORFs	Annotation	M+CD+8.5g YE /YPS	M+CD+1g YE /YPS	M+CD+5g YE /YPS
<i>fdh1-nyf1-msh1</i>	TON_0272	Na ⁺ /H ⁺ antiporter MshF subunit, putative	0.194663	0.411647	0.331549
<i>fdh1-nyf1-msh1</i>	TON_0267	Putative integral membrane protein, DUF68 family	0.296945	1.215972	0.838087
<i>fdh1-nyf1-msh1</i>	TON_0273	hypothetical protein	0.269222	0.652722	0.482705
<i>fdh1-nyf1-msh1</i>	TON_0275	formate hydrogen lyase subunit 6	0.313283	0.621333	0.511819
<i>fdh1-nyf1-msh1</i>	TON_0276	component G or formate hydrogen lyase	0.275138	0.724228	0.692331
<i>fdh1-nyf1-msh1</i>	TON_0274	component I or formate hydrogen lyase	0.385799	0.627629	0.614487
<i>fdh1-nyf1-msh1</i>	TON_0280	oxidoreductase iron-sulfur protein	0.3913	0.690322	0.647104
<i>fdh1-nyf1-msh1</i>	TON_0271	Na ⁺ /H ⁺ antiporter subunit	0.410417	0.328062	0.51901
<i>fdh1-nyf1-msh1</i>	TON_0266	component F or formate hydrogen lyase	0.38336	0.338185	0.333459
<i>fdh1-nyf1-msh1</i>	TON_0277	component C or formate hydrogen lyase	0.460646	0.357836	0.387126
<i>fdh1-nyf1-msh1</i>	TON_0268	Putative integral membrane protein, DUF67 family	0.648077	0.424896	0.86235
<i>fdh1-nyf1-msh1</i>	TON_0281	fdhA formate dehydrogenase, alpha subunit	0.584576	0.446401	0.44946
<i>fdh1-nyf1-msh1</i>	TON_0265	Nucleotidyltransferase, putative	0.421398	0.554433	0.503557
<i>mbx</i>	TON_0491	617aa long hypothetical protein	0.459201	0.532164	0.388201
<i>mbx</i>	TON_0490	NADH dehydrogenase subunit	0.372425	0.469467	0.278163
<i>fdh1-nyf1-msh1</i>	TON_0261	hypA hydrogenase expression/formation protein	0.658409	0.78949	0.746964
<i>mbx</i>	TON_0493	MbxD subunit	0.658086	0.817339	0.749402
<i>fdh1-nyf1-msh1</i>	TON_0286	hydrogenase maturation protein HypF	0.685606	0.778307	0.719973
<i>mbx</i>	TON_0494	Multisubunit Na ⁺ /H ⁺ antiporter, putative MshB subunit	0.765286	0.716126	0.711304
<i>frh</i>	TON_1559	coenzyme F420 hydrogenase alpha subunit	0.695408	0.64447	0.726535
<i>mbx</i>	TON_0487	naoD NADH dehydrogenase I, subunit D	0.606823	0.730138	0.517528
<i>fdh1-nyf1-msh1</i>	TON_0262	ATPase involved in chromosome partitioning, Mrp homolog	0.640277	0.776389	0.619674
<i>frh</i>	TON_1560	Coenzyme F420 hydrogenase gamma subunit	0.56769	0.715513	0.635236
<i>nyf1</i>	TON_0054	hydrogenase subunit gamma	0.606088	0.517724	0.630549
<i>fdh1-nyf1-msh1</i>	TON_0269	Multisubunit Na ⁺ /H ⁺ antiporter, putative MshB subunit	0.591411	0.571996	0.600235
<i>mbx</i>	TON_0486	4Fe-4S cluster-binding subunit	0.821271	0.718277	0.504388
<i>fdh1-nyf1-msh1</i>	TON_0278	component E or formate hydrogen lyase	0.720779	0.623875	0.510273
<i>mbx</i>	TON_0488	NADH dehydrogenase subunit	0.703506	0.618352	0.49939
<i>mbx</i>	TON_0498	Multisubunit Na ⁺ /H ⁺ antiporter, putative	0.675578	0.564247	0.483617
<i>mbx</i>	TON_0489	naoB NADH dehydrogenase I, subunit B	0.797159	0.618863	0.547462
<i>fdh2-nyf2-msh2</i>	TON_1575	Na ⁺ /H ⁺ antiporter subunit	0.752784	0.646406	0.578467
<i>fdh2-nyf2-msh2</i>	TON_1577	Multisubunit Na ⁺ /H ⁺ antiporter, putative MshB subunit	0.726511	0.655031	0.556386
<i>nyf1</i>	TON_0544	alcohol dehydrogenase	0.87473	0.419323	0.603934
<i>nyf1</i>	TON_0533	hydrogenase-specific maturation endopeptidase	0.977803	0.516234	0.73084
<i>fdh2-nyf2-msh2</i>	TON_1572	hypothetical protein	0.844948	0.504955	0.692272
<i>mbx</i>	TON_0492	MbxH subunit	0.862715	0.484349	0.423606
<i>fdh1-nyf1-msh1</i>	TON_0279	NADH ubiquinone oxidoreductase	0.80929	0.515429	0.54289
<i>mbh</i>	TON_1584	MbhC subunit	0.947559	0.514932	0.5099
<i>mbh</i>	TON_1583	MbhB subunit	1.038442	0.519366	0.533773
<i>mbh</i>	TON_1588	MbhG subunit	1.027148	0.563686	0.510924
<i>nyf1</i>	TON_0052	hydrogenase subunit alpha	1.085002	0.652365	0.550585

[0115]

<i>fdh2-afh2-afh2</i>	TON_1580	component F or formate hydrogen lyase	1.015786	0.756283	0.529199
<i>mbh</i>	TON_1594	MbhM subunit	0.982323	0.742211	0.458623
<i>codH-mch-afh3</i>	TON_1016	putative transcriptional regulator, ModE family	1.187798	0.634394	0.731218
<i>codH-mch-afh3</i>	TON_1026	Na ⁺ /H ⁺ antiporter subunit	1.337719	0.605839	0.623499
<i>fdh2-afh2-afh2</i>	TON_1576	hypothetical protein	1.168966	0.564328	0.60819
<i>codH-mch-afh3</i>	TON_1022	respiratory-chain NADH dehydrogenase, subunit 1	0.877363	1.628883	0.57715
<i>fdh1-afh1-afh1</i>	TON_0264	molybdopterin-guanine dinucleotide biosynthesis protein A	0.538102	0.957335	0.826924
<i>sufl1</i>	TON_0535	cytosolic NiFe-hydrogenase, delta subunit	0.652099	1.000516	0.800955
<i>mbx</i>	TON_0496	MbxC subunit	0.609819	1.200655	0.995067
<i>fdh2-afh2-afh2</i>	TON_1582	hypothetical protein	0.753301	0.998059	0.952751
<i>fdh1-afh1-afh1</i>	TON_0270	hypothetical protein	0.827927	0.85379	0.581906
<i>fdh1-afh1-afh1</i>	TON_0285	hypD hydrogenase expression/formation protein	0.904687	1.066448	0.758981
<i>mbh</i>	TON_1585	MbhD subunit	0.774604	1.015909	0.670497
<i>mbx</i>	TON_0497	MbxB subunit	0.860304	0.889029	0.810902
<i>fh</i>	TON_1561	4Fe-4S ferredoxin, iron-sulfur binding Nitrite/sulfite reductase	0.824149	0.834127	0.807754
<i>fdh2-afh2-afh2</i>	TON_1579	Putative integral membrane protein, DUF68 family	0.903028	0.812346	0.816319
<i>fdh2-afh2-afh2</i>	TON_1566	NADH dehydrogenase (quinone)	1.011779	0.736029	0.787627
<i>mbh</i>	TON_1587	MbhF subunit	1.129566	0.811366	0.872572
<i>mbh</i>	TON_1586	MbhE subunit	1.101049	0.811262	0.639266
<i>sufl1</i>	TON_0051	hypothetical protein	1.273306	3.1454	2.569546
<i>fdh2-afh2-afh2</i>	TON_1563	Probable formate dehydrogenase, alpha subunit	1.008259	3.037865	3.330042
<i>fdh1-afh1-afh1</i>	TON_0283	hypC hydrogenase expression/formation protein	1.007318	1.853137	2.021914
<i>fdh1-afh1-afh1</i>	TON_0288	hypothetical protein	1.004008	1.964442	2.27641
<i>fdh2-afh2-afh2</i>	TON_1571	component I or formate hydrogen lyase	1.030626	2.15871	1.806722
<i>fdh2-afh2-afh2</i>	TON_1565	NADH ubiquinone oxidoreductase	0.811937	1.787894	2.048693
<i>codH-mch-afh3</i>	TON_1031	component F or formate hydrogen lyase	1.030507	2.788588	1.982091
<i>codH-mch-afh3</i>	TON_1018	carbon-monooxide dehydrogenase, catalytic subunit	1.474238	2.204944	1.785437
<i>fdh2-afh2-afh2</i>	TON_1569	component G or formate hydrogen lyase	1.204795	2.281342	1.780033
<i>fdh2-afh2-afh2</i>	TON_1570	formate hydrogen lyase subunit 6	1.257649	2.427962	1.99359
<i>fdh1-afh1-afh1</i>	TON_0287	hydrogenase expression/formation protein HypE	0.851301	1.399569	1.257059
<i>sufl2</i>	TON_0053	sulfhydrogenase II, beta chain	0.84527	1.44889	1.452254
<i>fh</i>	TON_1562	Formate dehydrogenase, subunit FdhD	0.908701	1.54082	1.468288
<i>sufl2</i>	TON_0053	Sulfhydrogenase II, delta chain	0.896129	1.898148	1.337585
<i>sufl1</i>	TON_0539	probable formate dehydrogenase, alpha subunit	1.158997	1.791123	1.314325
<i>codH-mch-afh3</i>	TON_1021	component H or formate hydrogen lyase	1.112538	1.711122	1.250867
<i>fdh1-afh1-afh1</i>	TON_0263	hydrogenase maturation protease HycI	0.775719	1.531337	1.090123
<i>fdh1-afh1-afh1</i>	TON_0289	cytosine desulfurase	0.666768	1.534991	1.148004
<i>mbx</i>	TON_0493	Mbx subunit Na ⁺ /H ⁺ antiporter, putative	0.70216	1.591408	1.138149
<i>fdh2-afh2-afh2</i>	TON_1581	molybdopterin-guanine dinucleotide biosynthesis protein A	0.61468	1.708744	1.243367
<i>codH-mch-afh3</i>	TON_1019	putative ATP-binding protein	0.824732	2.016924	0.974466
<i>sufl1</i>	TON_0542	glutamate synthase beta chain-related oxidoreductase	0.842213	1.602828	0.934649
<i>mbh</i>	TON_1592	NiFe-hydrogenase large subunit 1	0.940698	1.573248	0.956935
<i>fdh1-afh1-afh1</i>	TON_0284	367aa long hypothetical hydrogenase protein hypD	1.036497	1.385466	1.182663
<i>sufl1</i>	TON_0537	Sulfhydrogenase beta subunit	1.025726	1.458113	1.160718

[0116]

<i>szj1</i>	TON_0534	cytosolic NiFe-hydrogenase, alpha subunit	1.283309	1.334781	1.071667
<i>fdh2-mfh2-mnh2</i>	TON_1564	4Fe-4S cluster-binding protein	1.178705	1.221574	1.445098
<i>fdh2-mfh2-mnh2</i>	TON_1567	component B or formate hydrogen lyase	1.091992	1.057076	1.184573
<i>fdh2-mfh2-mnh2</i>	TON_1568	component C or formate hydrogen lyase	1.471267	1.365925	1.382301
<i>fdh2-mfh2-mnh2</i>	TON_1573	probable formate transporter	1.451628	1.459657	1.158131
<i>fdh2-mfh2-mnh2</i>	TON_1574	Na ⁺ /H ⁺ antiporter MnhF subunit, putative	1.53015	1.602287	1.244315
<i>szj1</i>	TON_0541	4Fe-4S cluster-binding protein	1.112648	1.328761	0.855465
<i>szj1</i>	TON_0536	hydrogenase (gamma subunit)	1.100959	1.192121	0.759086
<i>codh-mch-mnh3</i>	TON_1028	Multisubunit Na ⁺ /H ⁺ antiporter putative MnhB subunit	1.019583	1.203407	0.836452
<i>szj1</i>	TON_0540	oxidoreductase iron-sulfur protein	0.987581	1.525167	0.82485
<i>codh-mch-mnh3</i>	TON_1027	hypothetical protein	0.991048	1.43866	0.778446
<i>szj1</i>	TON_0543	4Fe-4S cluster-binding protein	0.9185	1.257283	1.019564
<i>szj1</i>	TON_0538	probable formate transporter	0.989789	1.113613	0.960108
<i>fdh2-mfh2-mnh2</i>	TON_1578	Putative integral membrane protein, DUF67 family	0.989286	1.061599	0.900044
<i>fdh1-mfh1-mnh1</i>	TON_0282	putative transcriptional regulator	1.340444	1.195043	0.883294
<i>mnh1</i>	TON_1590	Mnh1 subunit	1.334732	1.227655	0.904813
<i>mnh1</i>	TON_1591	NiFe-hydrogenase small subunit	1.627315	1.158127	0.856929
<i>mnh1</i>	TON_1589	MnhH subunit	1.362042	0.948034	0.71984
<i>mnh1</i>	TON_1595	4Fe-4S cluster-binding subunit	1.589644	1.280298	0.491906
<i>codh-mch-mnh3</i>	TON_1030	Putative integral membrane protein, DUF64 family	3.779798	4.956998	1.592196
<i>codh-mch-mnh3</i>	TON_1024	NADEI dehydrogenase (ubiquinone), 20 kDa subunit	1.816315	3.578441	2.009973
<i>codh-mch-mnh3</i>	TON_1025	Na ⁺ /H ⁺ antiporter MnhF subunit, putative	2.443938	3.389037	1.333503
<i>codh-mch-mnh3</i>	TON_1029	Putative integral membrane protein, DUF67 family	2.648286	2.171138	1.123898
<i>mnh1</i>	TON_1593	NiFe-hydrogenase large subunit 2	1.609483	2.286207	1.354849
<i>codh-mch-mnh3</i>	TON_1020	hypothetical protein	3.375546	9.366047	5.425704
<i>codh-mch-mnh3</i>	TON_1017	4Fe-4S ferredoxin, iron-sulfur binding domain	1.713085	4.46107	3.621394
<i>codh-mch-mnh3</i>	TON_1023	component G or formate hydrogen lyase	1.927395	6.118015	4.270288

[0117]

[0118]

[0119]

3) 유전자 파괴 및 돌연변이들의 형질분석(Gene disruption and phenotype analysis of disruption mutants)

전사체(transcriptome) 분석은 *codh*와 가깝게 클러스터된 *mch* 수소화효소가(도 9A) *T. onnurineus* NA1에서 일산화탄소영양 수소 생성(carboxydrotrophic hydrogenogenesis) 과정에 중요한 역할을 할 수 있다는 사실을 암시한다. 하지만, *fdh2-mfh2-mnh2* 클러스터의 상호조절 및 다른 수소화효소 mRNA의 많은 반복수는 *codh-mch-mnh3*가 단독으로 상기 과정에 관여하는 것인지 아니면 다른 수소화효소들이 탈수소화효소(dehydrogenases)와 복합체를 형성하거나 탈수소화효소와 연합하여 FADH₂나 NADH와 같은 전자운반체를 재활용함으로써 *mch*에 대해 대체적 경로가 될 수 있는지에 관하여 의문을 불러일으킨다. 따라서 본 발명자들은 *mch* 또는 *mfh2* 수소화효소 각각의 큰 서브유닛의 결손(disruption) 돌연변이들을 만들었다(Matsumi, R., K. Manabe, T. Fukui, H. Atomi, T. Imanaka. 2007. Disruption of a sugar transporter gene cluster in a hyperthermophilic archaeon using a host-marker system based on antibiotic resistance. J. Bacteriol. 189:2683-2691). 제작 방법은 도 9A에 도시되어 있다. 표적 영역에서의 상동재조합(homologous recombination)에 의한 P_{gdh}-*hmg*_{P_{fu}} 카세트의 삽입 불활성화 및 그에 따른 *hmg*-CoA 유전자의 과발현으로 Mch 또는 Mfh2 수소화효소 유전자 클러스터 각각의 큰 서브유닛이 파괴(disruption)된다. 표적 유전자 파괴는, 10 μM 심바스타틴(simvastatin)이 첨가된 YPS 배지에서의 생존으로 후보군들을 선별한 후에, PCR 증폭을 통한 P_{gdh}-*hmg*_{P_{fu}} 카세트의 존재로 확인되었다(도 9B). 파괴 후보군들(Δ*mch*_{TNA1} 및 Δ*mfh2*_{TNA1})에서 P_{gdh}-*hmg*_{P_{fu}}가 증폭될 수 있는 반면, *mch* 또는 *mfh2*의 큰 서브유닛의 증폭은 실패했다. 이러한 결과들은 *T. kodakaraensis* KOD1 (Sapra, R., K. Bagramyan, and M. W. W. Adams. 2003. A simple energy-conserving system: Proton reduction coupled to proton translocation. Proc. Natl. Acad. Sci. USA 100:7545-7550)에 보고된 유전자 파괴 시스템이 *T. onnurineus* NA1에도 적용될 수 있음을 보여준다.

[0120]

YPS 배지에서 결손(disruption) 돌연변이 Δ*mch*_{TNA1} 및 Δ*mfh2*_{TNA1} 을 얻을 수 있었으므로, Mch 또는 Mfh2가 YPS 배지의 대사 소비에 필수적이지 않다는 것을 알 수 있다. 도 10에서 알 수 있듯이, 돌연변이 균주들 Δ*mch*_{TNA1} 및 Δ*mfh2*_{TNA1}의 성장 및 형태 변화는 YPS 배지 상에서 상기 유전자들이 반드시 필수적인 것은 아니라는 것을 확인시켜준다. 추가적으로, CO 성장 조건에서 Δ*mfh2*_{TNA1} 균주는 야생형 균주와 비슷한 수준으로 성장하고 수소를 생성하였다(도 7 및 10). 반면에, Δ*mch*_{TNA1} 돌연변이는 CO 성장 조건에서 성장할 수 없었으며, 수소도 생산하지 못했다(도 10). 이는 Mch 큰 서브유닛의 부재가 CO 하에서 *T. onnurineus* NA1의 일산화탄소영양 생존능력을 완전히 파괴한다는 것을 보여준다. 이 결과들을 함께 고려하면, CO가 기질로 공급되었을 경우 Mch 수소화효소만이 성장 및 수소 생성에 관여한다고 볼 수 있다.

- [0121] 실시예 4. *Thermococcus onnurineus* NA1에 의한 포름산염(formate)-의존적 H₂ 생성: 포름산염-반응(responsive) 수소화효소의 동정
- [0122] (1) 실험방법
- [0123] 1) 배양조건
- [0124] YPS(yeast extract-peptone-sulfur) 배지에서 80℃의 혐기성 조건으로 *T. onnurineus* NA1을 배양하였다. 돌연변이 균주의 성장 특성을 조사하기 위해, 기본배지로 30.0g/L의 NaCl을 첨가한 변형배지 1을 사용하였다(Uffen, R. L. 1976. Anaerobic growth of a Rhodopseudomonas species in the dark with carbon monoxide as sole carbon and energy substrate. Proc. Natl. Acad. Sci. USA 73:3298-3302). 상기 배지를 가압멸균시킨 후, 혐기성 챔버에서 1.0ml/L의 비타민 용액(Balch, W. E., G. E. Fox, L. J. Magrum, C. R. Woese, and R. S. Wolfe. 1979. Methanogens: reevaluation of a unique biological group. Microbiol. Rev. 43:260-296) 및 1.0g/L의 효모추출물(yeast extract)을 첨가하였다. 상기 기본 배지에 1N NaOH 및 10.0g/L의 포름산염 나트륨(Sigma)을 첨가하여 pH를 8.0으로 맞췄다. 준비된 배지 30ml를 150ml 배양병(serum bottle)에 분배시켰고, 가스충(120ml)은 80% N₂ - 20% CO₂ (10⁵ Pa)로 하였다. 생리학적인 실험을 위한 모든 배양은 80℃의 혐기성 조건에서 24시간 동안 이루어졌으며, 상기 실험들은 중복하여 수행하였다.
- [0125] 2) RNA 추출 및 마이크로어레이 설계
- [0126] 상기 배양물들을 4℃에서 3,500×g로 10분간 원심분리하고, 생산자의 프로토콜(Invitrogen, Carlsbad, California)에 따라 TRIZOL 시약으로 총(total) RNA를 추출하였다. 총 RNA 시료들은 NanoDrop 분광기(ND-1000, Thermo Scientific) 및 전기영동으로 정량 및 정성 분석하였다. 본 실험에 사용된 마이크로어레이는 Roche NimbleGen 마이크로어레이를 사용하였다. 간략하게, 총 31,994 개의 독립적인 60mer 올리고뉴클레오티드들을 설계하고, 광 탈보호(photo deprotection chemistry)를 이용하여, 인 시츄(in situ)로 합성하였다. 각각의 독립적인 올리고뉴클레오티드는 어레이 상에서 2번 반복되었다(총 ~72,000 features).
- [0127] 3) cDNA 합성 및 혼성화(hybridization) 조건
- [0128] 마이크로어레이는 생산자의 프로토콜에 따라 실시하였다. 각각의 총 RNA 시료(5 μg)는 역전사효소, SuperScrip II (Invitrogen, Carlsbad, California)를 이용한 역전사 반응에 의해 시아닌(Cy5) 결합 dCTP (Amersham, Piscataway, NJ)로 표지되었다. 그 후 상기 표지된 cDNA 혼합물을 에탄올 침전법을 이용해 농축시켰다. 농축된 Cy5 표지 cDNAs를 30 μl의 혼성화 용액(GenoCheck, Korea)에 현탁시켰다. 표지된 cDNA들을 마이크로어레이에 위치시킨 후, MAUI Mixer X4 hybridization chamber mixers (BioMicro systems, Salt Lake City, UT)로 덮었다. 상기 슬라이드들을 MAUI 12-bay systems (BioMicro systems, Salt Lake City, UT)로 42° C에서 12시간 동안 혼성화시켰다. 상기 혼성화된 슬라이드들을 실온에서 2 X SSC, 0.1% SDS에서 2분, 1X SSC에서 3분, 그 후 0.2 X SSC에서 2분 동안 세척하였다. 상기 슬라이드들을 말리기 위해 1,000 x g에서 20초간 원심분리하였다.
- [0129] 4) 슬라이드 스캐닝(scanning), 표준화(normalization), 및 데이터 분석
- [0130] 어레이는 GenePix 4000B scanner (Molecular Devices Corporation, Union City, CA)를 사용하여 스캔하였고, 데이터는 NimbleScan 2.4 소프트웨어를 사용하여 추출하였다. 어레이 표준화는 메디안 폴리쉬 및 콰antil 표준화(median polish and quantile normalization) 방법을 사용하여 수행하였다(Amaratunga, D., and J. Cabrera. 2001. Statistical analysis of viral microchip data. J. Am. Stat. Assoc. 96:1161-1170). 개별적인 프로브들에 대한 표준화된 발현 값은, Irizarry *et al.* (Karl, D. M. 1995. The microbiology of deep-sea hydrothermal vents. CRC Press, Boca Raton, FL)에 기술된 바와 같은 RMA (robust multi-array average) 방법을 사용하여, 주어진 ORF에 대한 발현 값을 얻는데 사용되었다. 마지막으로, 시료에서 특정 유전자를 위해 얻은 RMA-처리된 발현 값들(RMA calls)을 이용하여 n배 변화율(n-fold change ratios)(R)을 계산하였다. 데이터 분석은 GeneSpring GX 7.3.1 (Agilent technologies, CA)를 이용하여 수행하였다. 폴드 변화 필터(Fold change filters)는 상향조절(up-regulated)되는 유전자들에 대해서는 적어도 대조구의 200%, 하향조절되는 유전자들에 대해서는 대조구의 50% 이하로 나타나야 한다는 조건을 포함시켰다. 데이터는 GeneSpring GX 7.3.1 (Agilent technologies, CA)을 사용한 실험들에서 유사하게 행동하는 유전자 그룹들 내에 클러스터시켰다. 유사한 패턴의 유전자를 분리하기 위하여 유클리드 거리(Euclidian distance) 및 평균결합법(average linkage)에 근거한 알고리즘을 사용하였다.

- [0131] 5) 정량적 RT-PCR
- [0132] *T. onnurineus* NA1 (Genbank accession number, CP000855) 의 유전체 서열로부터 유전자 특이적 프라이머들을 설계하였다(상기 표 3 참조).
- [0133] cDNA들은 역전사효소, SuperScrip II (Invitrogen, Carlsbad, California)를 사용하여 생산자의 프로토콜에 따라 350ng의 총 RNA로부터 합성하였다. PCR 반응은 *T1* thermocycler (Biometra)를 사용하여 *rTaq* (Takara) DNA 중합효소로 수행하였다. 상기 반응은 1차-사슬 반응(first-strand reaction) cDNA, 10 pmol 의 프라이머들, 200 μM dNTPs 및 생산자의 버퍼를 포함하는 50 μl 혼합물에서 이루어졌다. 94°C에서 2분간의 변성(denaturation) 단계 후에, 변성(30초, 94° C), 어닐링(annealing) (30초, 56° C) 및 신장반응(extension) (1분, 72° C)이 25 사이클 수행되었다. PCR 산물은 0.8% 아가로스젤 전기영동을 사용하여 분석하였다. 발현수준은 GelPro32.EXE v4.6 (Media Cybernetics, Inc.) 을 사용하여 측정하였다. Cha로 명명된 샤페로닌(Chaperonine)-암호화 유전자를 발현 수준을 표준화하기 위한 대조구로 사용하였다.
- [0134] 6) 표적 유전자 제거(Targeted gene disruption)
- [0135] *T. onnurineus* NA1의 생체 내(*in vivo*)에서의 수소화효소의 기능을 분석하기 위해, 고호열성 고세균(archaeon)인 *T. kodakaraensis* KOD1 에 사용된 유전자 제거 시스템(gene disruption system)을 이용하여(Sapra, R., K. Bagramyan, and M. W. W. Adams. 2003. A simple energy-conserving system: Proton reduction coupled to proton translocation. Proc. Natl. Acad. Sci. USA 100:7545-7550), *Mfh* 수소화효소의 큰 서브유닛의 삽입 불활성 돌연변이(insertional inactivation mutant)들을 제작하였다.
- [0136] 구체적으로, *mfh1* 및 *mfh2* 수소화효소의 각각의 큰 서브유닛(TON_0276 및 TON_1569)의 인접지역(flanking region)을 포함하는 DNA 단편들을 *Flk-mfh1* 또는 *Flk-mfh2*에 대한 프라이머 세트들(표 3)을 사용하여 *T. onnurineus* NA1의 지놈 DNA로부터 증폭시켰다. 각각의 증폭된 단편들은 *HincII* 로 절단하여 pUC118에 접합시켰다. 그 후, 주형(*Flk-Mfh1*_pUC118 또는 *Flk-mfh2*_pUC118 재조합 플라스미드)으로부터 프라이머 세트(*Ivs-Mfh1* 또는 *Ivs-mfh2*) (표 3)들을 사용하여 역(inverse) PCR을 수행하고, 이어서 $P_{gdh-hmgP_{fu}}$ 카세트에 접합시켰다(Sapra, R., K. Bagramyan, and M. W. W. Adams. 2003. A simple energy-conserving system: Proton reduction coupled to proton translocation. Proc. Natl. Acad. Sci. USA 100:7545-7550). 그 후, 상기 접합된 결과물을 *Escherichia coli* DH5α 세포들에 형질전환시켰다. 재조합 플라스미드(*Mfh1::P_{gdh-hmgP_{fu}}* 및 *mfh2::P_{gdh-hmgP_{fu}}*)들은 플라스미드 미니 키트(Qiagen, Hilden, Germany)로 준비하였다. 마지막으로, 상기 플라스미드들을 Sato et al.(Sato, T., T. Fukui, H. Atomi, and T. Imanaka. 2003. Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. J. Bacteriol. 185:210-220., Sato, T., T. Fukui, H. Atomi, and T. Imanaka. 2005. Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. Appl. Environ. Microbiol. 71:3889-3899)의 약간 변형된 방법으로 *T. onnurineus* NA1에 형질전환시켰다. 형질전환체들은 10 μM 심바스타틴(simvastatin)의 존재 하에 ASW-YT-S 배지에 의해 스크리닝되었고, 표적 유전자의 제거여부는 PCR 증폭을 통해 표적 지역내 $P_{gdh-hmgP_{fu}}$ 카세트의 존재여부로 확인할 수 있었다.
- [0137] 7) 성장 및 수소 생산에 관한 동역학적 분석
- [0138] 성장은 눈으로 직접 확인하였다. 시료들을 해염(Sea Salt) (30.0 g/L), 포르말린(2.5%), 및 4' -6' -디아미디노-2-페닐인돌(0.01%)을 포함하는 멸균수에 희석시키고(Sato, T., T. Fukui, H. Atomi, and T. Imanaka. 2005. Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. Appl. Environ. Microbiol. 71:3889-3899), 블랙 폴리카보네이트 막 필터로 거른 후(구멍 크기, 0.2μm; Whatman), 위상차현미경(Zeiss Axioplan)으로 썼다. 수소 가스는 HP-PLOT Molesieve column (Agilent) 및 TCD 검출기가 구비된 가스 크로마토그래프 HP 5890 series II (Hewlett Packard)를 사용하여 측정하였다. 아르곤이 가스 운반자로 사용되었다. 오븐 온도는 40°C 였다. 분석을 위한 10μl의 가스 시료는 배양병의 부틸고무를 통해 가스-타이트 시린지(gas-tight syringe)로 꺼냈다. 검출된 수소 가스의 측정은 피크 면적을, 질소에 40% 수소를 포함하는 표준 가스를 사용한 희귀분석에 의해 수행된 보정 곡선과 비교함으로써 계산되었다.
- [0139] (2) 실험결과

[0140] 1) *T. onnurineus* NA1 내의 수소화효소

[0141] 앞선 *T. onnurineus* NA1의 유전체 분석은 8 반복의 수소화효소 유전자 클러스터의 존재를 보여주었으며(Porter, K. G. and Y. S. Feig. 1980. The use of DAPI for identifying and counting microflora. Limnol. Oceanogr. 25:943-948), 5개의 막-결합(membrane-bound) [NiFe]-수소화효소들(Mbh, TON_1583-1595; Mbx, TON_0486-0498; Mfh1, TON_0273-0278; Mfh2, TON_1566-1573; Mch, TON_1021-1024), 및 3개의 세포질내(cytoplasmic) [NiFe]-수소화효소들(Fru, TON_1559-1562; Sulf I, TON_0533-0544; Sulf II, TON_0051-0055)을 포함한다. 지놈 시퀀싱이 완료된 Thermococcales 균주들과의 수소화효소 유전자 클러스터의 비교 분석을 통해 Mbh, Mbx, Sulf I 및 Sulf II 클러스터들이 다른 Thermococcales 균주들에서 많이 발견된다는 것을 알 수 있었다(Fox, J. D., R. L. Kerby, G. P. Roberts, and P. W. Ludden. 1996. Characterization of the CO-induced, CO-tolerant hydrogenase from Rhodospirillum rubrum and the gene encoding the large subunit of the enzyme. J. Bacteriol. 178:1515-1524. 등 참조). 반면에, Mfh1, Mfh2 또는 Mch 클러스터들에 상동성이 있는 클러스터들은 매우 드물었고, *T. barophilus* MP (Mfh1, Mch homologues) 및 *Thermococcus* sp. AM4 (Mfh1, Mch homologues)와 같이 최근에 지놈 서열이 결정된 Thermococcales 균주들에서 발견된다. 특히, *T. onnurineus* NA1의 Mfh1 및 Mfh2 의 비교 분석은 양 클러스터의 유전자 구성이 *T. litoralis* (Wu, M., Q. Ren, A. S. Durkin, S. C. Daugherty, L. M Brinkac, R. J. Dodson, R. Madupu, S. A. Sullivan, J. F. Kolonay, W. C. Nelson, L. J. Tallon, K. M. Jones, L. E. Ulrich, J. M. Gonzalez, I. B. Zhulin, F. T. Robb, and J. A. Eisen. 2005. Life in hot carbon monoxide: the complete genome sequence of Carboxydotherrmus hydrogenoformans Z-2901. PLoS Genet. 1:e65), *T. barophilus* MP (Unfinished sequence, GenBank accession number, ABSF00000000), *Thermococcus* sp. AM4 (Unfinished sequence, GenBank accession number, ABXN00000000) 및 *Pyrococcus abyssi* (GenBank accession number, AL096836) 에서 발견된 것과 매우 유사하다는 것을 보여준다. 상기 유전자 클러스터들은 포름산염 탈수소화효소 (FDH) 유닛, 에너지-전환 수소화효소 (ECH) 유닛 및 H⁺/Na⁺ 역수송체 (Mnh) 유닛으로 구성된다. 막-결합 수소화효소와 연관된 FDHs 에 더하여, 막-결합 [NiFe]-수소화효소에 클러스터되지 않은 2개 이상의 FDH 복제물을 유전체에서 발견하였다.

[0142] 2) 포름산염 성장 조건에서의 수소화효소 유전자들의 발현

[0143] 도 11에 나타낸 바와 같이, *T. onnurineus* NA1 은 외부 포름산염을 기질로 하여 성장할 수 있으며, 상기 성장은 수소 생성과 밀접한 상관관계가 있다. 수소 가스의 양과 세포 수는 포름산염을 첨가한 배지 1에서 배양 시간이 길어질수록 증가하였다. 포름산염-포함 배지에서 수소생성과 관련된 수소화효소를 동정하기 위하여, 포름산염 배지에서 키웠을 때의 수소화효소 유전자 클러스터의 mRNA 발현 수준과 복합 배지(YPS)에서 키웠을 때의 발현 수준을 마이크로어레이 분석을 이용하여 비교하였다. 하기 표 6, 표 7 및 도 12A 에 나타낸 바와 같이, fdh2-mfh2-mnh2 클러스터(앞서, Hyg4-III 로 명명) 에서 20개의 ORF들 중 10개의 ORF들의 mRNA 발현 수준은 YPS에 비해서 포름산염 배지에서 상당히 (2배 이상) 상향조절(up-regulated)되었다. 게다가, codh-mch-mnh3 클러스터(앞서, Hyg4-II 로 명명)에서 6개의 ORF들의 발현 수준 역시 3.0g의 효모추출물을 포함한 포름산염 배지에서만 상향조절되었다. 또한, fdh2-mfh2-mnh2 클러스터의 ORF 발현수준은 효모추출물의 양에 따라 다양했으며, 이는 포름산염-유도 대사작용(formate-driven metabolism)에서 이화작용의 억제 또는 활성에의 관련성을 암시한다(표 6 및 표 7 참조). 반면에, 다른 수소화효소 유전자 클러스터들의 ORF 발현 수준들은 상대적으로 일정하거나(mbx, frh 및 mch) 억제된 수준(mbh, sulf I 및 mfh1)으로 나타났다. 흥미롭게도, fdh1-mfh1-mnh1 클러스터(앞서, Hyg4-I 로 명명) 내의 대부분의 유전자들(29개 ORF들 중 25개 유전자들)은 하향조절되었고, 이는 외부의 포름산염이 fdh2-mfh2-mnh2 는 상향조절시키는 반면, fdh1-mfh1-mnh1의 발현은 억제할 수 있다는 것을 시사한다. 수소화효소 유전자 클러스터 각각의 큰 서브유닛에 대한 정량적 RT-PCR 결과 역시 마이크로어레이 데이터와 일치하였다. mfh2 수소화효소의 큰 서브유닛 (TON_1569)은 YPS에서 보다 포름산염-첨가 배지에서 적어도 2 사이클 빨리 증폭되었다(도 12B). 이러한 결과들은 fdh2-mfh2-mnh2 클러스터가 포름산염에 반응할 수 있다는 사실을 암시한다. fdh2-mfh2-mnh2 클러스터만이 가능한 포름산염 트랜스포터(putative formate transporter) (TON_1573)로 어노테이션된 유전자를 가졌다는 점을 주목할 필요가 있으며, 포름산염을 유입시키는 기능과 관련될 수 있다.

[0144] 하기 표 6은 YPS 배지에서와 비교한 포름산염 포함 배지에서의 수소화효소 유전자 클러스터 각각의 ORF 발현 수준을 보여주며, 표 7은 *T. onnurineus* NA1 의 수소화효소 클러스터를 포함한 112개 ORF들의 계층적 클러스터링을 나타낸다.

[0145]

[표 6]

Gene cluster and ORFs	Annotation	Mt+Formate+0.5g YE /YPS	Mt+Formate+3g YE /YPS
mgH hydrogenase cluster			
TON_0051	hypothetical protein	2.102325	1.996031
TON_0052	hydrogenase subunit alpha	1.203768	0.802758
TON_0053	Sulfhydrogenase II, delta chain	1.065323	1.367881
TON_0054	hydrogenase subunit gamma	0.737077	0.849314
TON_0055	sulfhydrogenase II, beta chain	0.867926	0.808351
/Mh1-mgH1-mmh1 cluster			
TON_0261	hypA hydrogenase expression/formation protein	0.727489	0.848653
TON_0262	ATPase involved in chromosome partitioning, ParA/MinD family, Mrp homolog	0.810405	0.474952
TON_0263	hydrogenase maturation protease HycI	0.929162	1.474294
TON_0264	molybdopterin-guanine dinucleotide biosynthesis protein A	0.483945	0.66368
TON_0265	Nucleosidyltransferase, putative	0.316634	0.585297
TON_0266	component F or formate hydrogen lyase	0.318627	0.458616
TON_0267	Putative integral membrane protein, DUF68 family	0.450192	1.592257
TON_0268	Putative integral membrane protein, DUF67 family	0.575413	0.478743
TON_0269	Multisubunit Na ⁺ /H ⁺ antiporter, putative MmhB subunit	0.575645	0.850024
TON_0270	hypothetical protein	0.754546	0.615634
TON_0271	Na ⁺ /H ⁺ antiporter subunit	0.658151	0.60018
TON_0272	Na ⁺ /H ⁺ antiporter MmhF subunit, putative	0.265439	0.481402
TON_0273	hypothetical protein	0.251284	0.533794
TON_0274	component I or formate hydrogen lyase	0.769351	0.759056
TON_0275	formate hydrogen lyase subunit 6	0.542026	0.622878
TON_0276	component G or formate hydrogen lyase	0.358736	0.446038
TON_0277	component C or formate hydrogen lyase	0.741095	0.623746
TON_0278	component B or formate hydrogen lyase	0.741675	0.678797
TON_0279	NADH ubiquinone oxidoreductase	0.817573	0.613605
TON_0280	oxidoreductase iron-sulfur protein	0.77206	0.704646
TON_0281	fdhA formate dehydrogenase, alpha subunit	0.366492	0.363246
TON_0282	putative transcriptional regulator	1.517388	1.556418
TON_0283	hypC hydrogenase expression/formation protein	1.601742	1.508267
TON_0284	367aa long hypothetical hydrogenase expression/formation protein hypD	1.092018	0.699213
TON_0285	hypD hydrogenase expression/formation protein	1.045483	0.489843
TON_0286	hydrogenase maturation protein HypF	0.529731	0.665571
TON_0287	hydrogenase expression/formation protein HypE	0.770546	0.834498
TON_0288	hypothetical protein	0.798203	2.193867
TON_0289	cysteine desulfurase	0.972411	0.894363
msh hydrogenase cluster			
TON_0486	4Fe-4S cluster-binding subunit	0.912618	0.560754
TON_0487	msoD NADH dehydrogenase I, subunit D	0.947824	0.64502
TON_0488	NADH dehydrogenase subunit	0.74442	0.331352
TON_0489	msoB NADH dehydrogenase I, subunit B	0.761859	0.54662
TON_0490	NADH dehydrogenase subunit	0.514819	0.256304

[0146]

TON_0491	617aa long hypothetical protein	0.445818	0.43555
TON_0492	MlxH subunit	0.786868	0.427415
TON_0493	Multisubunit Na ⁺ /H ⁺ antiporter, putative	0.67054	0.631248
TON_0494	Multisubunit Na ⁺ /H ⁺ antiporter, putative MnhB subunit	0.682815	1.042934
TON_0495	MlxD subunit	0.886482	0.784067
TON_0496	MlxC subunit	0.74959	0.923091
TON_0497	MlxB subunit	0.905817	0.796591
TON_0498	Multisubunit Na ⁺ /H ⁺ antiporter, putative	0.841502	0.512737
aa171 hydrogenase cluster			
TON_0533	hydrogenase-specific saturation endopeptidase	0.736449	1.021127
TON_0534	cytosolic NiFe-hydrogenase, alpha subunit	0.650022	0.677196
TON_0535	cytosolic NiFe-hydrogenase, delta subunit	0.741138	0.667442
TON_0536	hydrogenase (gamma subunit)	0.576788	0.558566
TON_0537	Sulfhydrogenase beta subunit	0.624364	0.629266
TON_0538	probable formate transporter	0.714736	0.690887
TON_0539	probable formate dehydrogenase, alpha subunit	0.852292	1.341748
TON_0540	oxidoreductase iron-sulfur protein	1.063573	0.561844
TON_0541	4Fe-4S cluster-binding protein	0.811281	0.78328
TON_0542	glutamate synthase beta chain-related oxidoreductase	0.812768	0.926217
TON_0543	4Fe-4S cluster-binding protein	0.983984	1.160664
TON_0544	alcohol dehydrogenase	1.029043	0.862322
caa85-aaa8-mnh3 cluster			
TON_1016	putative transcriptional regulator, ModE family	1.219563	1.440116
TON_1017	4Fe-4S ferredoxin, iron-sulfur binding domain	1.537494	2.172151
TON_1018	carbon-monoxide dehydrogenase, catalytic subunit	1.420438	1.274753
TON_1019	putative ATP-binding protein	0.698035	0.930905
TON_1020	hypothetical protein	1.971922	3.839171
TON_1021	component B or formate hydrogen lyase	0.972559	1.151527
TON_1022	respiratory-chain NADH dehydrogenase, subunit 1	1.771992	1.106361
TON_1023	component G or formate hydrogen lyase	1.863412	4.043911
TON_1024	NADH dehydrogenase (ubiquinone), 20 kDa subunit	1.353224	3.47523
TON_1025	Na ⁺ /H ⁺ antiporter MnhF subunit, putative	0.923065	2.35011
TON_1026	Na ⁺ /H ⁺ antiporter subunit	0.964696	0.621115
TON_1027	hypothetical protein	0.932259	0.97379
TON_1028	Multisubunit Na ⁺ /H ⁺ antiporter, putative MnhB subunit	0.916306	1.019604
TON_1029	Putative integral membrane protein, DUF67 family	1.207926	0.928939
TON_1030	Putative integral membrane protein, DUF68 family	0.712484	2.238387
TON_1031	component F or formate hydrogen lyase	0.955361	2.246413
β⁸⁸ (F₄₂₀ reducing hydrogenase) hydrogenase cluster			
TON_1559	coenzyme F420 hydrogenase alpha subunit	1.183388	0.530475
TON_1560	Coenzyme F420 hydrogenase gamma subunit	1.415197	1.018573
TON_1561	4Fe-4S ferredoxin, iron-sulfur binding; Nitrite/sulfite reductase	1.355412	1.07094
TON_1562	Formate dehydrogenase, subunit FdhD	1.00278	0.973292
Fdh3-aa33-aa32 cluster			
TON_1563	Probable formate dehydrogenase, alpha subunit	4.251498	2.533032
TON_1564	4Fe-4S cluster-binding protein	3.005736	1.806787

[0147]

TON_1565	NADH ubiquinone oxidoreductase	2.572148	1.499477
TON_1566	NADH dehydrogenase (quinone)	1.807319	0.972802
TON_1567	component B or formate hydrogen lyase	3.431877	1.536872
TON_1568	component C or formate hydrogen lyase	2.936933	1.620744
TON_1569	component G or formate hydrogen lyase	3.631759	1.932512
TON_1570	formate hydrogen lyase subunit 6	2.69592	2.206621
TON_1571	component I or formate hydrogen lyase	2.061174	1.799012
TON_1572	hypothetical protein	1.247346	1.037087
TON_1573	probable formate transporter	2.315563	1.228123
TON_1574	Na ⁺ /H ⁺ antiporter MnhF subunit, putative	3.337396	2.209763
TON_1575	Na ⁺ /H ⁺ antiporter subunit	1.446144	0.801728
TON_1576	hypothetical protein	1.538306	0.954677
TON_1577	Multisubunit Na ⁺ /H ⁺ antiporter, putative MnhB subunit	1.43188	1.026239
TON_1578	Putative integral membrane protein, DUF67 family	1.656445	1.198853
TON_1579	Putative integral membrane protein, DUF68 family	1.382536	1.135341
TON_1580	component F or formate hydrogen lyase	1.804806	1.031391
TON_1581	methyltetrahydropterin-guanine dinucleotide biosynthesis protein A	1.42751	1.994937
TON_1582	hypothetical protein	1.086162	0.6947
nitH hydrogenase cluster			
TON_1583	MhhB subunit	0.67072	0.367439
TON_1584	MhhC subunit	0.715606	0.384515
TON_1585	MhhD subunit	0.649744	0.488634
TON_1586	MhhE subunit	0.579611	0.376248
TON_1587	MhhF subunit	0.660332	0.683508
TON_1588	MhhG subunit	0.667072	0.521087
TON_1589	MhhH subunit	0.439735	0.343741
TON_1590	MhhI subunit	0.355357	0.42006
TON_1591	NiFe-hydrogenase small subunit	0.278923	0.289092
TON_1592	NiFe-hydrogenase large subunit 1	0.497784	0.493123
TON_1593	NiFe-hydrogenase large subunit 2	0.355	0.852904
TON_1594	MhhM subunit	0.373076	0.289584
TON_1595	4Fe-4S cluster-binding subunit	0.421286	0.597661

M, medium 1; CO, carbon monoxide; YE, yeast extract.

[0148]

[0149]

[표 7]

Gene cluster and ORFs	Annotation	M+Formate+0.5g YE / YPS	M+Formate+3g YE / YPS	Gene cluster and ORFs
<i>fdh1-ufh1-msh1</i>	TON_0262	ATase involved in chromosome partitioning, Mry homolog	0.810405	0.474952
<i>Mtx</i>	TON_0498	Multisubunit Na ⁺ /H ⁺ antiporter, putative	0.841502	0.512737
<i>Mdx</i>	TON_0489	mshB NADH dehydrogenase I, subunit B	0.761839	0.54662
<i>Mdx</i>	TON_0492	MshH subunit	0.786868	0.427415
<i>Mdx</i>	TON_0486	4Fe-4S cluster-binding subunit	0.912618	0.560734
<i>Mdx</i>	TON_0487	mshD NADH dehydrogenase I, subunit D	0.947824	0.64502
<i>codh-mch-msh3</i>	TON_1026	Na ⁺ /H ⁺ antiporter subunit	0.964696	0.621115
<i>fdh1-ufh1-msh1</i>	TON_0285	hypD hydrogenase expression/formation protein	1.045483	0.489843
<i>suf1</i>	TON_0540	oxidoreductase iron-sulfur protein	1.063573	0.561844
<i>Fvh</i>	TON_1559	coenzyme F420 hydrogenase alpha subunit	1.183388	0.530475
<i>fdh1-ufh1-msh1</i>	TON_0269	Multisubunit Na ⁺ /H ⁺ antiporter, putative MshB subunit	0.575645	0.850024
<i>fdh1-ufh1-msh1</i>	TON_0270	hypothetical protein	0.754546	0.615634
<i>fdh1-ufh1-msh1</i>	TON_0277	component C or formate hydrogen lyase	0.741093	0.623746
<i>fdh1-ufh1-msh1</i>	TON_0279	NADH ubiquinone oxidoreductase	0.817573	0.613605
<i>fdh1-ufh1-msh1</i>	TON_0280	oxidoreductase iron-sulfur protein	0.77206	0.704646
<i>fdh1-ufh1-msh1</i>	TON_0278	component B or formate hydrogen lyase	0.741673	0.678797
<i>suf1</i>	TON_0538	cytosolic NiFe-hydrogenase, delta subunit	0.741138	0.667442
<i>suf1</i>	TON_0538	probable formate transporter	0.714736	0.690887
<i>fdh1-ufh1-msh1</i>	TON_0271	Na ⁺ /H ⁺ antiporter subunit	0.658151	0.60018
<i>msh</i>	TON_0493	Multisubunit Na ⁺ /H ⁺ antiporter, putative	0.67054	0.631248
<i>suf1</i>	TON_0537	Sulfhydrogenase beta subunit	0.624364	0.629266
<i>suf1</i>	TON_0534	cytosolic NiFe-hydrogenase, alpha subunit	0.650022	0.677196
<i>msh</i>	TON_1587	MshF subunit	0.660332	0.683508
<i>fdh1-ufh1-msh1</i>	TON_0264	molybdopterin-guanine dinucleotide biosynthesis protein A	0.483943	0.66368
<i>fdh1-ufh1-msh1</i>	TON_0273	formate hydrogen lyase subunit 6	0.542026	0.622878
<i>fdh1-ufh1-msh1</i>	TON_0286	hydrogenase maturation protein HypF	0.529731	0.665571
<i>msh</i>	TON_1595	4Fe-4S cluster-binding subunit	0.421286	0.597661
<i>fdh1-ufh1-msh1</i>	TON_0268	Putative integral membrane protein, DUF67 family	0.375413	0.478743
<i>msh</i>	TON_1592	NiFe-hydrogenase large subunit I	0.497784	0.493123
<i>suf1</i>	TON_0536	hydrogenase (gamma subunit)	0.576788	0.558566
<i>msh</i>	TON_1585	MshD subunit	0.649744	0.488634
<i>msh</i>	TON_1588	MshG subunit	0.667072	0.521087
<i>suf2</i>	TON_0053	sulfhydrogenase II, beta chain	0.867926	0.808351
<i>msh</i>	TON_0495	MshD subunit	0.886482	0.784067
<i>msh</i>	TON_0497	MshB subunit	0.905817	0.796591
<i>fdh1-ufh1-msh1</i>	TON_0274	component I or formate hydrogen lyase	0.769351	0.759056
<i>suf1</i>	TON_0541	4Fe-4S cluster-binding protein	0.811281	0.78328
<i>suf2</i>	TON_0054	hydrogenase subunit gamma	0.737077	0.849314
<i>fdh1-ufh1-msh1</i>	TON_0261	hypA hydrogenase expression/formation protein	0.727489	0.848653
<i>fdh1-ufh1-msh1</i>	TON_0287	hydrogenase expression/formation protein HypE	0.770546	0.834498
<i>msh</i>	TON_0496	MshC subunit	0.74959	0.923091

[0150]

<i>codH-mch-mnh3</i>	TON_1019	putative ATP-binding protein	0.698035	0.930505
<i>sdj1</i>	TON_0542	glutamate synthase beta chain-related oxidoreductase	0.812768	0.926217
<i>mhx</i>	TON_0494	Multisubunit Na ⁺ /H ⁺ antiporter, putative MnhB subunit	0.682815	1.042934
<i>sdj1</i>	TON_0533	hydrogenase-specific maturation endopeptidase	0.736449	1.021127
<i>sdj1</i>	TON_0543	4Fe-4S cluster-binding protein	0.983984	1.160664
<i>codH-mch-mnh2</i>	TON_1021	component B or formate hydrogen lyase	0.972559	1.151527
<i>fdh1-mfh1-mnh1</i>	TON_0289	cysteine desulfurase	0.972411	0.894363
<i>sdj1</i>	TON_0544	alcohol dehydrogenase	1.029043	0.862232
<i>codH-mch-mnh2</i>	TON_1027	hypothetical protein	0.932259	0.97379
<i>codH-mch-mnh3</i>	TON_1028	Multisubunit Na ⁺ /H ⁺ antiporter, putative MnhB subunit	0.916306	1.019604
<i>frh</i>	TON_1562	Formate dehydrogenase, subunit FdhD	1.00278	0.973292
<i>sdj1</i>	TON_0052	hydrogenase subunit alpha	1.203768	0.802758
<i>fdh1-mfh1-mnh1</i>	TON_0284	367 aa long hypothetical hydrogenase protein hypD	1.092018	0.699213
<i>fdh2-mfh2-mnh2</i>	TON_1582	hypothetical protein	1.086162	0.6947
<i>fdh2-mfh2-mnh2</i>	TON_1575	Na ⁺ /H ⁺ antiporter subunit	1.446144	0.801728
<i>fdh1-mfh1-mnh1</i>	TON_0267	Putative integral membrane protein, DUF68 family	0.450192	1.592257
<i>mnh</i>	TON_1593	NiFe-hydrogenase large subunit 2	0.355	0.852904
<i>fdh1-mfh1-mnh1</i>	TON_0265	Nucleotidyltransferase, putative	0.316634	0.585297
<i>Fdh1-mfh1-mnh1</i>	TON_0272	Na ⁺ /H ⁺ antiporter MnhF subunit, putative	0.263459	0.481402
<i>Fdh1-mfh1-mnh1</i>	TON_0273	hypothetical protein	0.251284	0.533794
<i>Fdh1-mfh1-mnh1</i>	TON_0266	component F or formate hydrogen lyase	0.318627	0.458616
<i>Fdh1-mfh1-mnh1</i>	TON_0276	component G or formate hydrogen lyase	0.358736	0.446038
<i>mhx</i>	TON_0491	617 aa long hypothetical protein	0.445818	0.43555
<i>fdh1-mfh1-mnh1</i>	TON_0281	FdhA formate dehydrogenase, alpha subunit	0.366492	0.363246
<i>mnh</i>	TON_1589	MnhH subunit	0.439733	0.343741
<i>mnh</i>	TON_1594	MnhM subunit	0.373076	0.289584
<i>mnh</i>	TON_1591	NiFe-hydrogenase small subunit	0.278923	0.289092
<i>mhx</i>	TON_0490	NADEH dehydrogenase subunit	0.514819	0.256304
<i>mhx</i>	TON_0488	NADEH dehydrogenase subunit	0.74442	0.331352
<i>mnh</i>	TON_1583	MnhB subunit	0.67072	0.367439
<i>mnh</i>	TON_1584	MnhC subunit	0.715606	0.384515
<i>mnh</i>	TON_1586	MnhE subunit	0.579611	0.376248
<i>mnh</i>	TON_1590	MnhI subunit	0.555557	0.42006
<i>fdh1-mfh1-mnh1</i>	TON_0288	hypothetical protein	0.798203	2.193867
<i>codH-mch-mnh3</i>	TON_1030	Putative integral membrane protein, DUF68 family	0.712484	2.238387
<i>codH-mch-mnh3</i>	TON_1025	Na ⁺ /H ⁺ antiporter MnhF subunit, putative	0.923065	2.35011
<i>codH-mch-mnh3</i>	TON_1031	component F or formate hydrogen lyase	0.955361	2.246413
<i>sdj1</i>	TON_0051	hypothetical protein	2.102325	1.996031
<i>fdh2-mfh2-mnh2</i>	TON_1571	component I or formate hydrogen lyase	2.061174	1.799012
<i>codH-mch-mnh3</i>	TON_1017	4Fe-4S ferredoxin, iron-sulfur binding domain	1.537494	2.172151
<i>fdh2-mfh2-mnh2</i>	TON_1581	molybdopterin-guanine dinucleotide biosynthesis protein A	1.42751	1.994937
<i>fdh1-mfh1-mnh1</i>	TON_0263	hydrogenase maturation protease HycI	0.929162	1.474294
<i>sdj1</i>	TON_0339	probable formate dehydrogenase, alpha subunit	0.832292	1.341748
<i>sdj1</i>	TON_0053	SsdHydrogenase II, delta chain	1.065323	1.367881

[0151]

<i>codh-mch-mnh3</i>	TON_1016	putative transcriptional regulator, ModE family	1.219563	1.440116
<i>fdh1-mfh1-mnh1</i>	TON_0282	putative transcriptional regulator	1.517388	1.556418
<i>fdh1-mfh1-mnh1</i>	TON_0283	hycC hydrogenase expression/formation protein	1.601742	1.508267
<i>codh-mch-mnh3</i>	TON_1018	carbon-monoxide dehydrogenase, catalytic subunit	1.420438	1.274753
<i>codh-mch-mnh3</i>	TON_1022	respiratory-chain NADH dehydrogenase, subunit 1	1.771932	1.186961
<i>fdh2-mfh2-mnh2</i>	TON_1578	Putative integral membrane protein, DUF67 family	1.656445	1.198853
<i>codh-mch-mnh3</i>	TON_1029	Putative integral membrane protein, DUF67 family	1.207926	0.928939
<i>fdh2-mfh2-mnh2</i>	TON_1572	hypothetical protein	1.247346	1.037087
<i>frh</i>	TON_1560	Coenzyme F420 hydrogenase gamma subunit	1.415197	1.018373
<i>fdh2-mfh2-mnh2</i>	TON_1577	Methionine Na ⁺ /H ⁺ antiporter, putative MnhB subunit	1.43188	1.026139
<i>frh</i>	TON_1561	4Fe-4S ferredoxin, iron-sulfur binding; Nitrite/sulfite reductase	1.355412	1.07094
<i>fdh2-mfh2-mnh2</i>	TON_1579	Putative integral membrane protein, DUF68 family	1.382336	1.135341
<i>fdh2-mfh2-mnh2</i>	TON_1566	NADH dehydrogenase (quinone)	1.807319	0.972802
<i>fdh2-mfh2-mnh2</i>	TON_1580	component F or formate hydrogen lyase	1.804806	1.031391
<i>fdh2-mfh2-mnh2</i>	TON_1576	hypothetical protein	1.538306	0.954677
<i>codh-mch-mnh3</i>	TON_1020	hypothetical protein	1.971922	3.839171
<i>codh-mch-mnh3</i>	TON_1023	component G or formate hydrogen lyase	1.863412	4.045911
<i>codh-mch-mnh3</i>	TON_1024	NADH dehydrogenase (ubiquinone), 20 kDa subunit	1.353224	3.47523
<i>fdh2-mfh2-mnh2</i>	TON_1563	Probable formate dehydrogenase, alpha subunit	4.251498	2.533032
<i>fdh2-mfh2-mnh2</i>	TON_1564	4Fe-4S cluster-binding protein	3.005736	1.806787
<i>fdh2-mfh2-mnh2</i>	TON_1568	component C or formate hydrogen lyase	2.936933	1.620744
<i>fdh2-mfh2-mnh2</i>	TON_1567	component B or formate hydrogen lyase	3.431877	1.536872
<i>fdh2-mfh2-mnh2</i>	TON_1570	formate hydrogen lyase subunit 6	2.69592	2.206621
<i>fdh2-mfh2-mnh2</i>	TON_1569	component G or formate hydrogen lyase	3.631759	1.932512
<i>fdh2-mfh2-mnh2</i>	TON_1574	Na ⁺ /H ⁺ antiporter MnhF subunit, putative	3.337596	2.209765
<i>fdh2-mfh2-mnh2</i>	TON_1565	Na ⁺ /H ⁺ antiporter subunit	1.446144	0.801728
<i>fdh2-mfh2-mnh2</i>	TON_1573	probable formate transporter	2.315563	1.228123

[0152]

[0153]

[0154]

3) 유전자 파괴 및 돌연변이들의 형질분석(Gene disruption and phenotype analysis of disruption mutants)

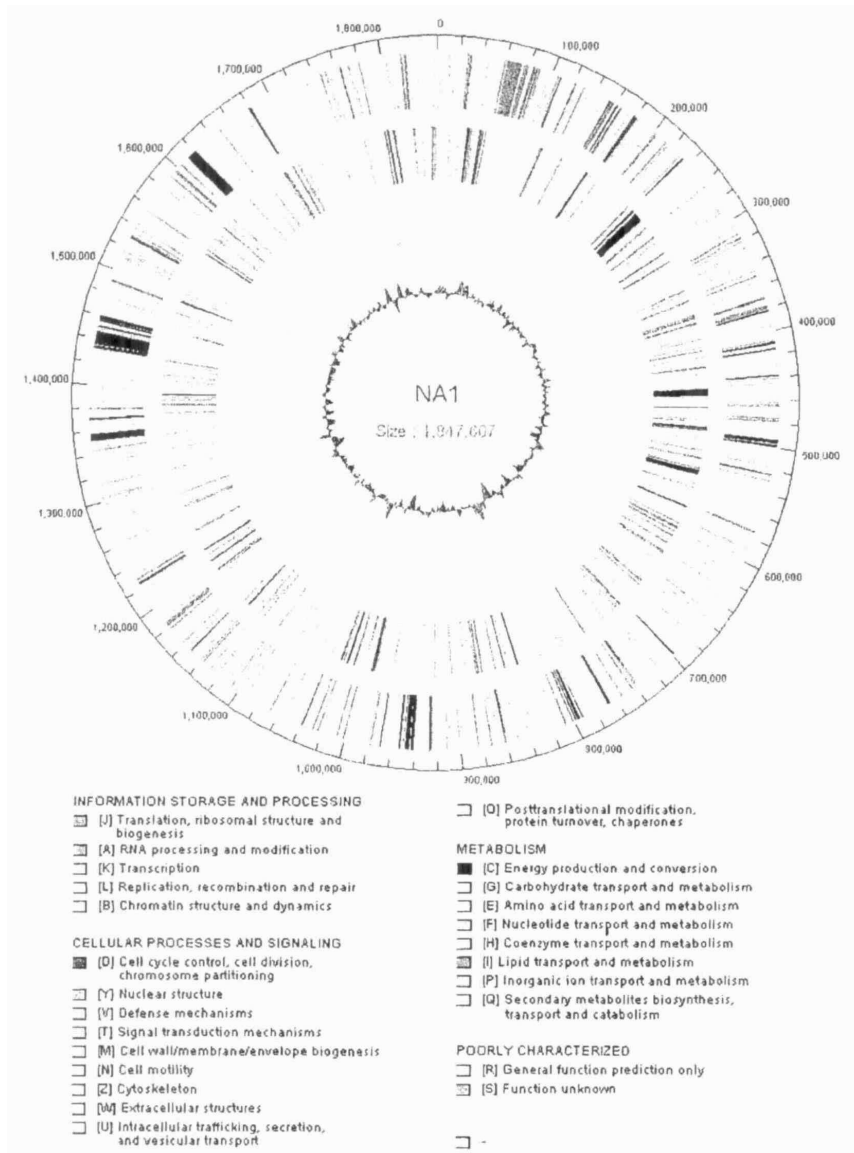
전사체(transcriptome) 분석은 Mfh2 가 *T. onnurineus* NA1에서 외부 포름산염을 이용한 수소 생성에 중요한 역할을 할 수 있다는 사실을 암시한다. 하지만, 포름산염 포함 배지에서의 *codh-mch-mnh3* 클러스터의 여러 ORF들의 상향조절 또는 Mch 및 Frh 와 같은 다른 수소화효소 유전자들의 높은 발현수준은 여전히 Mfh2가 단독으로 상기 과정에 관여하는 것인지 아니면 다른 수소화효소들이 포름산염 탈수소화효소(formate dehydrogenases)와 복합체를 형성하거나 FADH₂나 NADH와 같은 전자운반체를 재활용함으로써 Mfh2에 대해 대체적 경로가 될 수 있는지에 관하여 의문을 불러일으킨다. 따라서 본 발명자들은 Mfh1 (TON_0276) 또는 Mfh2 (TON_1569) 수소화효소 각각의 큰 서브유닛의 파괴(disruption) 돌연변이들을 만들었다. 제작 방법은 도 13A에 도시되어 있다. 표적 영역에서의 상동재조합(homologous recombination)에 의한 P_{gdh-hmgP_{fu}} 카세트의 삽입 불활성화 및 그에 따른 hmg-CoA 유전자의 과발현으로 Mfh1 또는 Mfh2 수소화효소 유전자 클러스터 각각의 큰 서브유닛이 파괴(disruption)된다. 표적 유전자 파괴는, 10 μM 심바스타틴(simvastatin)이 첨가된 YPS 배지에서의 생존으로 후보군들을 선별한 후에, PCR 증폭을 통한 P_{gdh-hmgP_{fu}} 카세트의 존재로 확인되었다(도 13B). 예를 들어, 파괴 후보군들(Δmfh1_{NA1} 및 Δmfh2_{NA1})에서는 P_{gdh-hmgP_{fu}}가 증폭될 수 있는 반면, Mfh1 또는 Mfh2의 큰 서브유닛의 증폭은 실패했다. 이러한 결과들은 표적 유전자 파괴가 원하는 유전자 위치에서 정확하게 일어났으며, *T. kodakaraensis* KOD1 (Sapra, R., K. Bagramyan, and M. W. W. Adams. 2003. A simple energy-conserving system: Proton reduction coupled to proton translocation. Proc. Natl. Acad. Sci. USA 100:7545-7550)에 보고된 유전자 파괴 시스템이 *T. onnurineus* NA1에도 적용될 수 있음을 보여준다.

[0155]

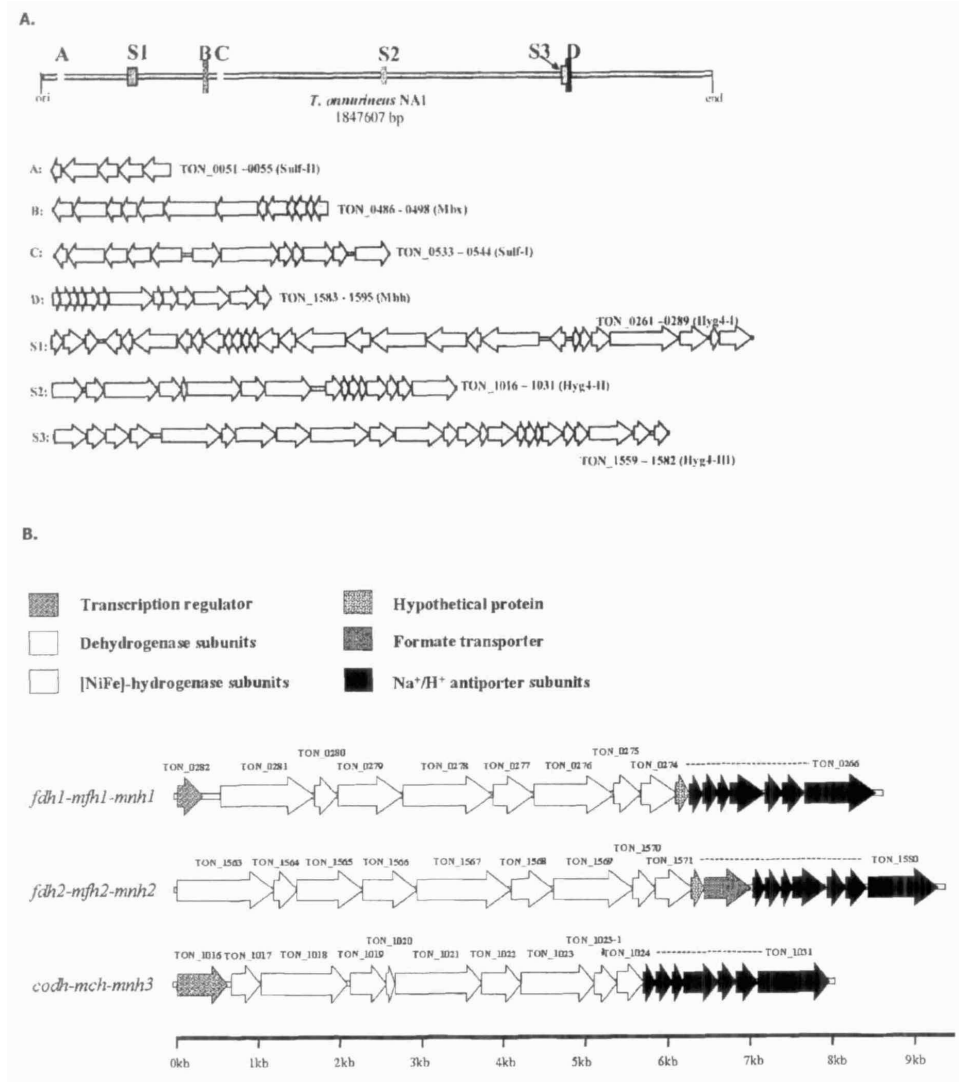
YPS 배지에서 결손(disruption) 돌연변이 Δmfh1_{NA1} 및 Δmfh2_{NA1} 을 얻을 수 있었으므로, mfh1 및 mfh2 가 YPS 배지의 대사 소비에 필수적이지 않다는 것을 알 수 있다. Mfh1, Mfh2 및 Mch 수소화효소의 돌연변이 균주들(Δmfh1_{NA1}, Δmfh2_{NA1} 및 Δmch_{NA1})의 성장 및 형태 변화는 YPS 배지 상에서 상기 유전자들이 반드시 필수적인 것은 아니라는 것을 보여준다(도 14). 추가적으로, 포름산염이 포함된 배지에서 Δmfh1_{NA1} 및 Δmch_{NA1} 균주들의 성장 및 수소생성은 야생형 균주와 비슷한 수준이었다(도 11 및 14). 이는 상기 유전자 클러스터들이 외부 포름산염이 투입된 조건에서의 성장과 무관하다는 것을 보여준다. 반면에, Δmfh2_{NA1} 돌연변이는 성장할 수 없으며, 수소를 생산할 수도 없었다(도 14). 이 결과들을 함께 고려하면, 포름산염이 기질로 공급되었을 경우 Mfh2 수소화효소만이 성장 및 수소 생성에 관여한다는 것을 보여준다.

도면

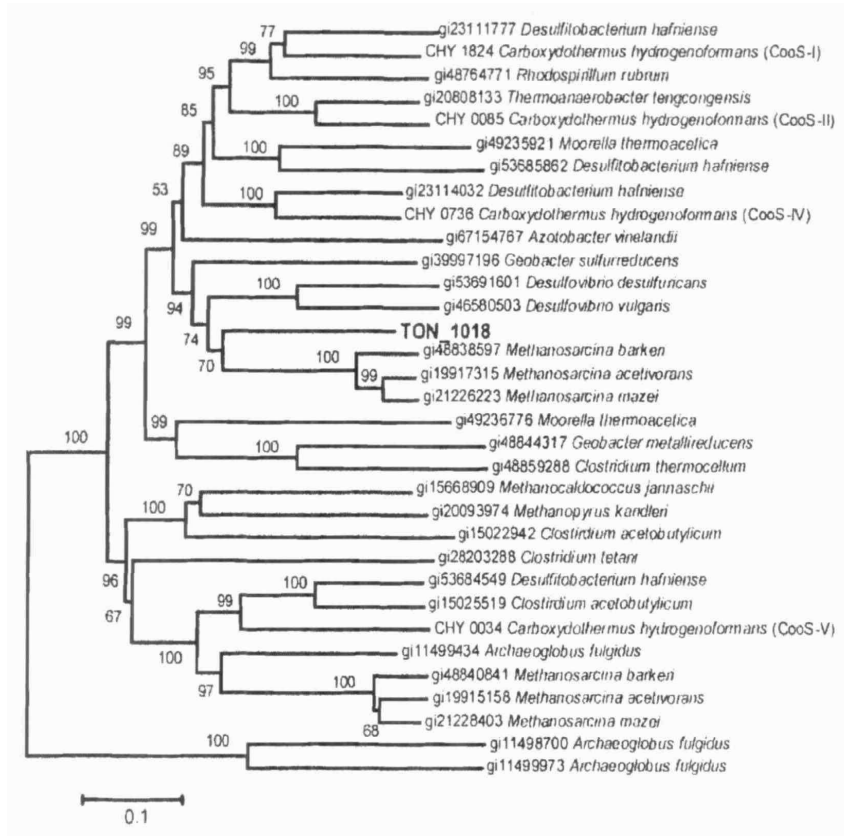
도면1



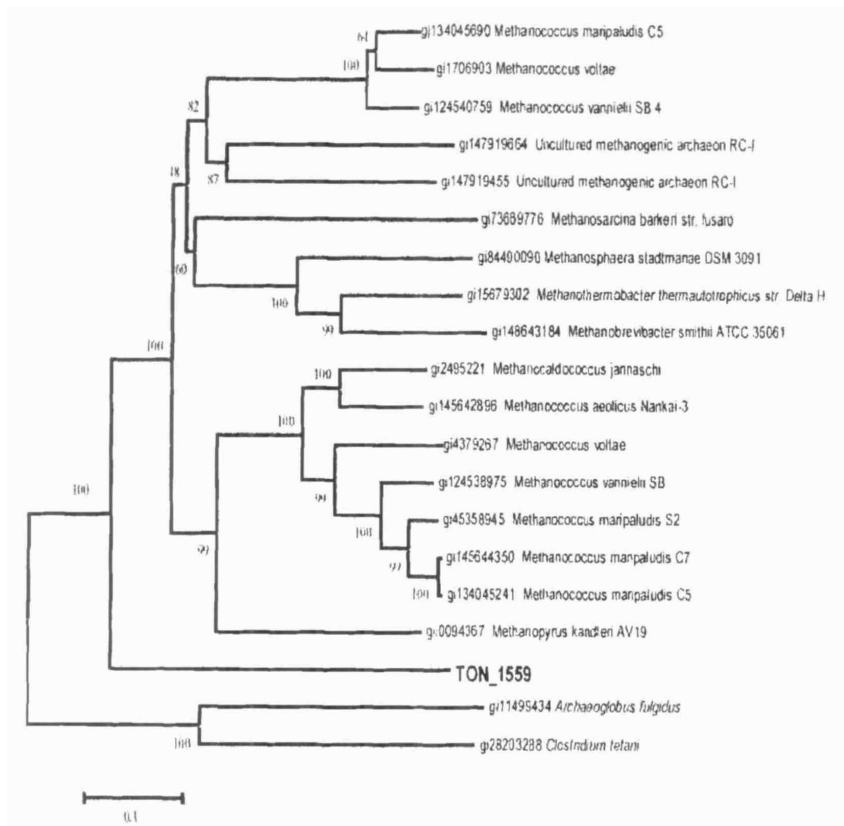
도면2



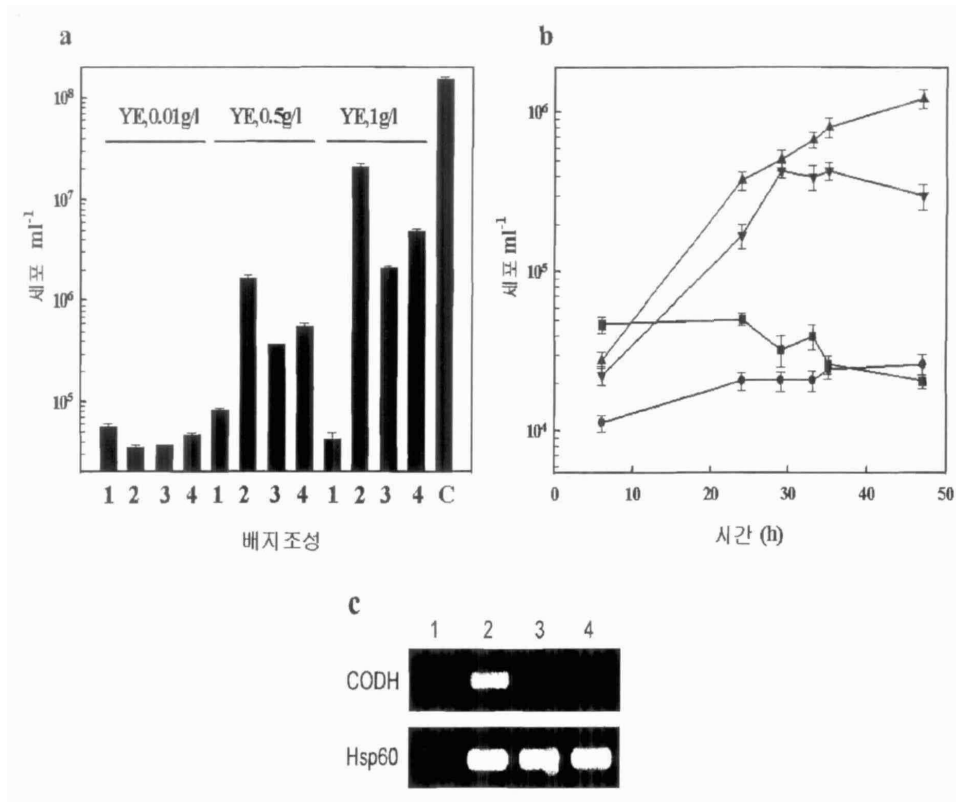
도면4



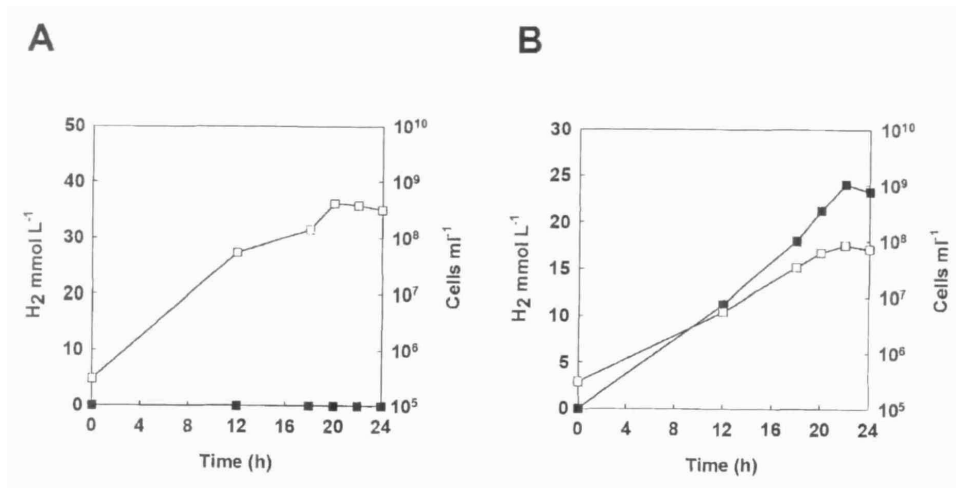
도면5



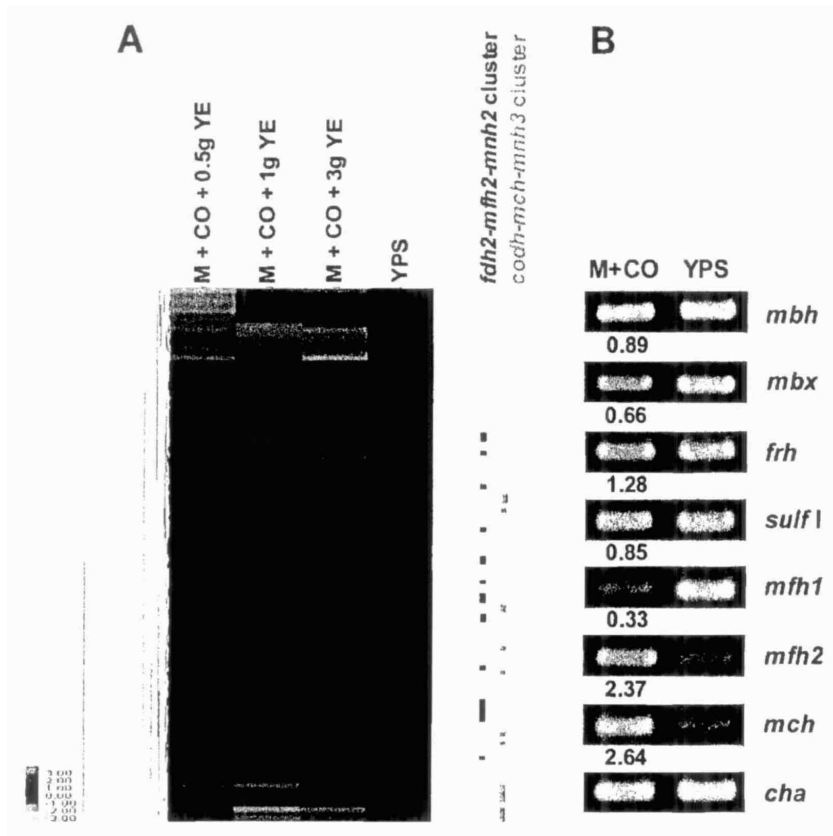
도면6



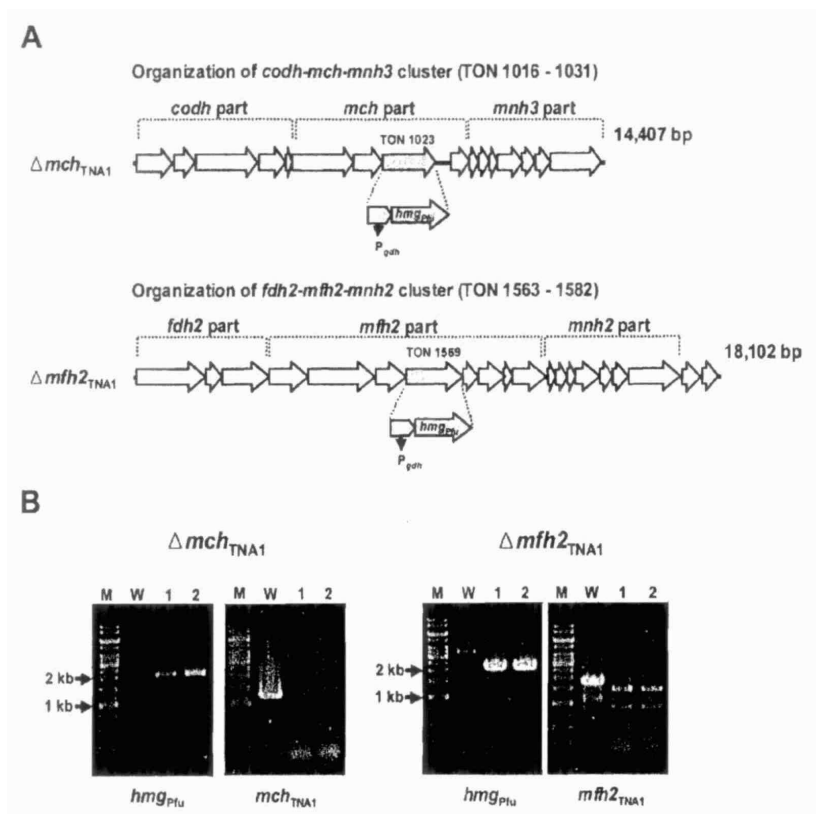
도면7



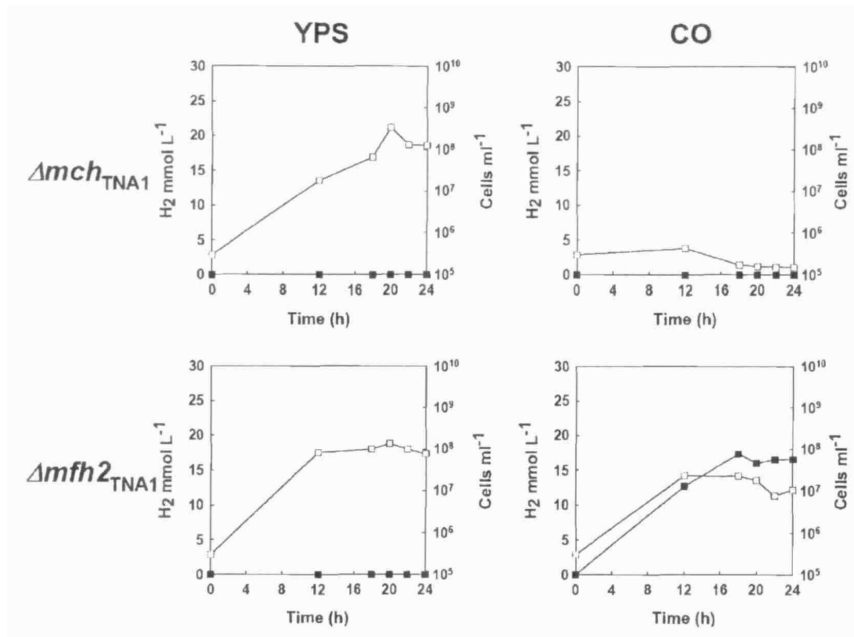
도면8



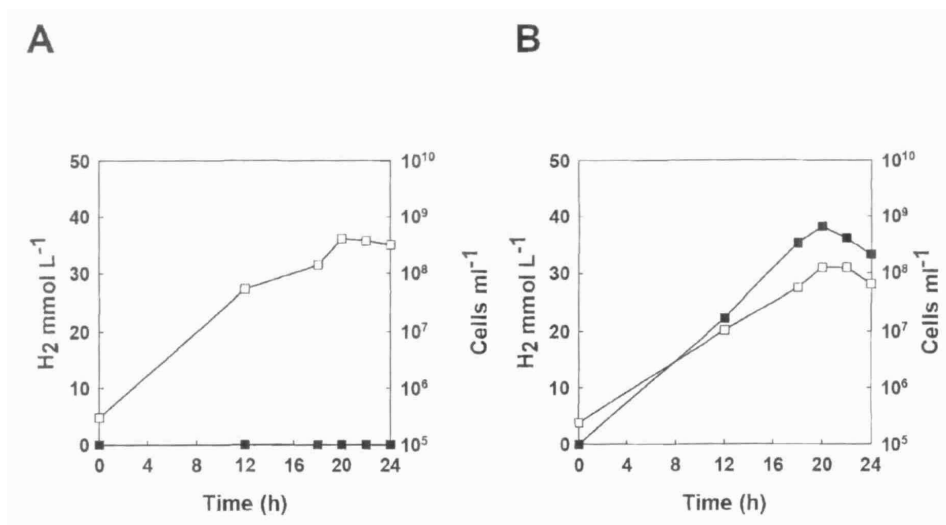
도면9



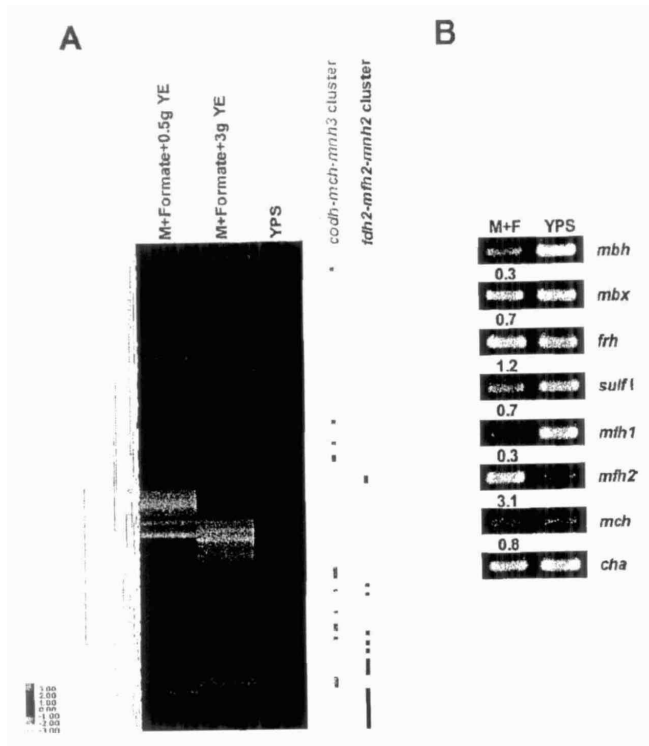
도면10



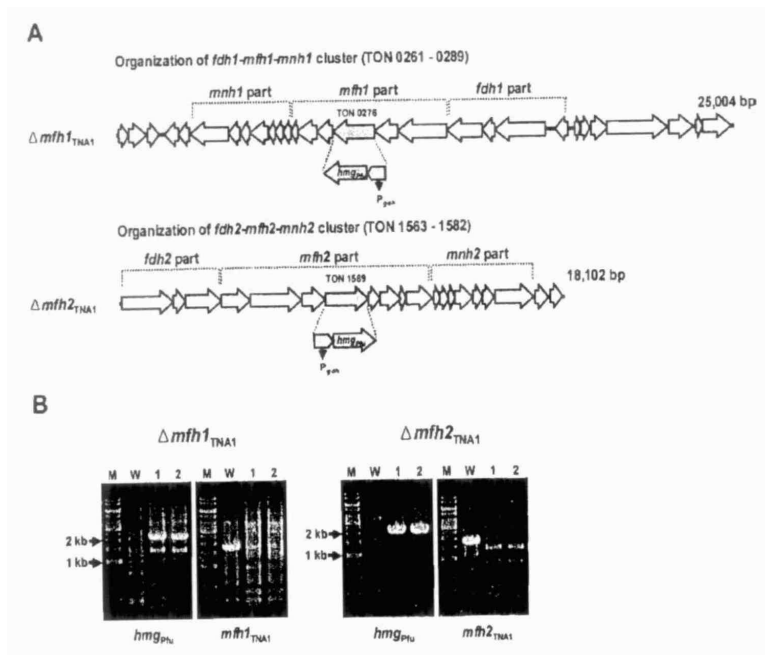
도면11



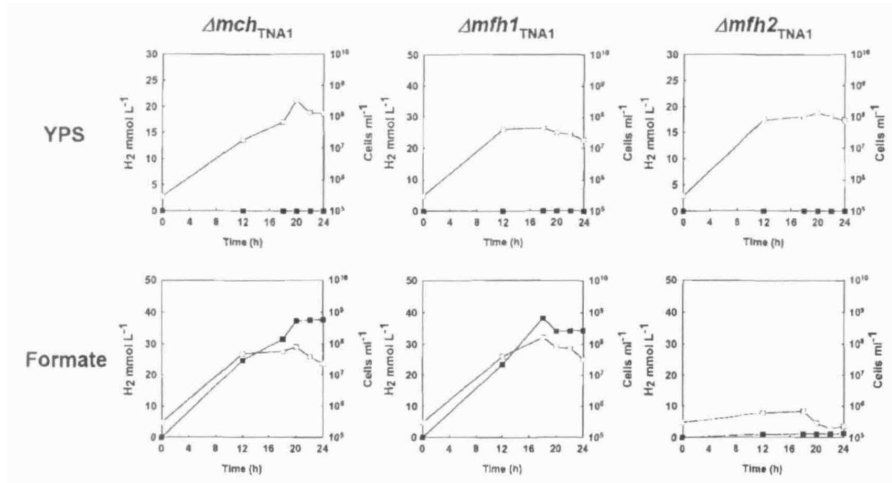
도면12



도면13



도면14



서열 목록

- <110> Korean Ocean Research & Development Institute
- <120> Novel hydrogenases of Sequence number 7 purified from Thermococcus spp. by using formate, Genes encoding them, and Methods for producing hydrogen using microorganism having the genes
- <130> PP090052KR-2
- <150> KR 10-2008-0087806
- <151> 2008-09-05
- <150> KR 10-2008-0087794
- <151> 2008-09-05
- <160> 60
- <170> KopatentIn 1.71
- <210> 1
- <211> 1393
- <212> PRT
- <213> Thermococcus onnurineus NA1

- <400> 1

Met Gly Phe Leu Ser Arg Leu Phe Gly Gly Lys Lys Glu Thr Asp Thr

1 5 10 15

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

20 25 30

Glu Lys Val Phe His Ile Met Gly Arg Glu Thr Leu Val Gly Thr Val
 35 40 45
 Glu Arg Gly Val Ile Tyr Pro Gly Tyr Lys Val Lys Gly Lys Lys Ala
 50 55 60
 Ala Val Ile Tyr Arg Ile Glu Lys Gly Arg Lys Ala Val Asp Phe Val
 65 70 75 80
 Val Asp Gly Asp Lys Ala Ala Leu Ile Leu Glu Gly Ile Thr Lys Ala
 85 90 95
 Glu Glu Gly Asp Thr Leu Glu Val Tyr Gln Ser Met Ile Ile Glu Leu
 100 105 110
 Arg Glu Phe Thr Arg Val Glu Gly Asn Gly Lys Ala Glu Ile Val Val
 115 120 125
 Glu Asn Gly Glu Val Lys Asp Val Arg Leu Lys Ile Val Glu Gly Pro
 130 135 140
 Arg Phe Phe Glu Leu Leu Thr Leu Gly Arg His Tyr Tyr Asp Val Pro
 145 150 155 160
 Asp Leu Glu Ala Arg Ile Cys Ala Ile Cys Tyr Leu Ser His Ser Val
 165 170 175
 Ala Ser Val Leu Gly Ile Glu Lys Ala Phe Gly Val Glu Val Ser Glu
 180 185 190
 Glu Ile Gln Leu Leu Arg Glu Leu Gly Leu Ile Gly Glu Leu Leu Glu
 195 200 205
 Ser His Ala Leu His Leu Tyr Leu Leu Val Ala Pro Asp Val Phe Gly
 210 215 220
 Tyr Pro Asp Ala Ile Arg Met Ala Thr Lys His Gly Glu Leu Val Lys
 225 230 235 240
 Glu Gly Leu Ala Leu Lys Ala Phe Gly Asn Ser Ile Arg Glu Leu Ile
 245 250 255
 Gly Gly Arg Glu Ile His Gly Ile Asn Val Lys Pro Gly Gly Phe Gly
 260 265 270
 Arg Tyr Pro Thr Val Glu Glu Leu Glu Asn Ile Glu Arg Glu Ser Gly

Leu Ala Val Phe Glu Leu Thr Asp Cys Gly Gly Cys Ala Leu Asn Ile
 530 535 540
 Leu Phe Leu Tyr Glu Lys Leu Phe Asp Leu Leu Glu Phe Tyr Glu Ile
 545 550 555 560
 Thr Glu Phe His Met Ala Thr Ser Leu Ser Glu Gly Ser His Tyr Asp
 565 570 575

 Val Ala Leu Val Thr Gly Thr Val Ser Ser Gln Arg Asp Leu Ala Leu
 580 585 590
 Leu Lys Glu Ala Arg Asn His Ser Asp Tyr Leu Ile Ala Leu Gly Thr
 595 600 605
 Cys Ala Thr His Gly Ser Val Gln Ala Ser Val Glu Gly Ser Ile Arg
 610 615 620
 Glu Lys Leu Lys Arg Val Tyr Gly Asp Glu Gly Asn Pro Met Arg Ala
 625 630 635 640
 Leu Asp Ser Lys Pro Val Val Glu Tyr Val Ala Val Asp Phe Ala Leu
 645 650 655
 Pro Gly Cys Pro Tyr Asp Lys Asn Glu Val Tyr Gln Val Leu Met Asp
 660 665 670
 Ile Ala Lys Gly Ile Glu Pro Val Lys Lys Asp Tyr Pro Val Cys Val
 675 680 685
 Glu Cys Lys Leu Asn Glu Tyr Glu Cys Val Leu Val Lys Lys Gly Leu
 690 695 700
 Pro Cys Leu Gly Pro Ile Thr Tyr Gly Gly Cys Asn Ala Ala Cys Ile
 705 710 715 720

 Arg Ser Gly Leu Gly Cys Ile Gly Cys Arg Gly Pro Leu Pro Gly Glu
 725 730 735
 Val Asn Pro Ala Ser Glu Tyr Glu Ile Leu Lys Asp Leu Gly Tyr Asp
 740 745 750
 Asp Asp Tyr Ile Leu Arg Lys Phe Lys Thr Phe Ala Arg Trp Glu Pro
 755 760 765
 Met Ser Glu Asn Pro His Gln Thr Tyr Asp Ala Arg Ile Leu Glu Val
 770 775 780

Lys Asp Leu Thr Pro Arg Glu Lys Leu Phe Thr Leu Arg Phe Leu Asp
 785 790 795 800
 Pro Glu Ile Gly Glu His Phe Thr Phe Lys Pro Gly Gln Phe Val Ile
 805 810 815
 Val Asp Ile Arg Gly Phe Gly Glu Phe Pro Ile Ser Leu Cys Ser Ser
 820 825 830
 Pro Thr Arg Lys Gly Tyr Ile Gln Leu Cys Ile Arg Lys Ala Gly Arg
 835 840 845
 Met Thr Lys Phe Ile His Gln Met Lys Glu Gly Glu Val Val Gly Ile
 850 855 860

 Arg Gly Pro Tyr Gly Asn Gly Phe Pro Met Glu Lys Met Glu Gly Ser
 865 870 875 880
 Asn Leu Leu Leu Val Ala Gly Gly Leu Gly Met Ala Pro Leu Arg Ser
 885 890 895
 Val Leu Trp Tyr Ala Ile Asp Thr Gly Lys Tyr Glu His Val Trp Leu
 900 905 910
 Leu Tyr Gly Thr Lys Ala Tyr Glu Asp Ile Leu Phe Arg Asp Glu Ile
 915 920 925
 Ile His Leu Leu Lys His Gly Asp Ala Val Gly Cys Ser Val Lys Leu

 930 935 940
 Ala Tyr Glu Val Glu Ser Pro Ser Cys Ile Tyr Leu Glu Arg Gly Phe
 945 950 955 960
 Phe Asp Arg Val Cys Lys Gly Val Val Thr Asp Leu Phe Arg Gly Glu
 965 970 975
 Glu Phe Asp Val Asp Lys Ala Tyr Ala Leu Ile Cys Gly Pro Pro Val
 980 985 990
 Met Tyr Arg Phe Val Ile Lys Glu Leu Leu Asp Arg Lys Leu Ser Pro
 995 1000 1005

 Gly Arg Ile Tyr Met Thr Leu Glu Arg Arg Met Arg Cys Gly Ile Gly
 1010 1015 1020
 Lys Cys Gly His Cys Ile Val Gly Thr Ser Thr Ser Ile Lys Tyr Val

1025 1030 1035 1040
 Cys Lys Asp Gly Pro Val Phe Thr Tyr Trp Asp Ala Leu Ser Thr Arg
 1045 1050 1055
 Gly Leu Ile Leu Arg Tyr Val Lys Leu Ser Ser Glu Asn Phe Ser Ser
 1060 1065 1070
 Phe Phe Glu Ser Leu Arg Asn Trp Gly Lys Val Tyr Ala Pro Ile Lys

 1075 1080 1085
 Arg Gly Ser Ile Tyr Thr Phe Gln Glu Val His Glu Leu Gly Glu Ile
 1090 1095 1100
 Glu Leu Asn Tyr Thr Arg Thr Met Leu Pro Pro Lys Lys Phe Phe Val
 1105 1110 1115 1120
 Arg Pro Arg Asp Glu Ile Leu Arg Leu Lys Asn Gly Arg Trp Glu Asn
 1125 1130 1135
 Gly Thr Asp Ala Glu Pro Ile Val Leu Phe Gly Leu His Ser Cys Asp
 1140 1145 1150

 Met His Gly Leu Lys Ile Leu Asp Lys Val Tyr Leu Asp Glu Pro Ala
 1155 1160 1165
 Asp Pro Tyr Tyr Lys Ala Arg Arg Glu Lys Thr Phe Ile Val Gly Ile
 1170 1175 1180
 Ser Cys Met Pro Asp Glu Tyr Cys Phe Cys Lys Ser Leu Gly Thr His
 1185 1190 1195 1200
 Phe Ala Met Asp Gly Phe Asp Leu Phe Leu His Glu Leu Pro Asp Gly
 1205 1210 1215
 Trp Leu Val Arg Ile Gly Ser Val Arg Gly His Glu Val Val Trp Glu

 1220 1225 1230
 Asn Gly Glu Leu Phe Glu Glu Val Thr Asp Glu Asp Leu Lys His Phe
 1235 1240 1245
 Lys Glu Phe Glu Glu Arg Arg Ala Asn Ala Phe Gln Lys Glu Ile Pro
 1250 1255 1260
 Gln Glu Gly Leu Ala Asp Met Leu Asp Leu Ala Tyr Asn Ser Pro Val
 1265 1270 1275 1280
 Trp Lys Glu Tyr Ala Glu Ile Cys Leu Ala Cys Gly Asn Cys Asn Met

Ala Val Asp Tyr Met Ser Glu Arg Asn Lys Gln His Pro Val Tyr Ser
 100 105 110
 Asp Lys Gly Arg Phe Tyr Ala Trp Met Val Ile Phe Val Gly Ala Thr
 115 120 125
 Leu Ala Phe Ile Tyr Ser Ser Thr Thr Leu Gln Leu Leu Ile Phe Phe
 130 135 140
 Glu Ile Met Gly Leu Ala Cys Trp Gly Val Val Gly Tyr Tyr Lys Gly
 145 150 155 160
 Pro Lys Ala Glu Arg Ala Ala Tyr Lys Ala Leu Leu Val Pro Asn Phe
 165 170 175
 Gly Ala Met Val Gly Leu Tyr Thr Thr Val Gly Ile Gly Ile Leu Lys
 180 185 190
 Leu His Asp Leu Ser Ile Tyr Ala Leu Gln Asn Leu Asn Asp Glu Leu
 195 200 205
 Lys Leu Leu Val Phe Leu Gly Val Met Val Ala Ala Phe Thr Lys Ser
 210 215 220
 Ala Gln Phe Pro Leu Tyr Ser Trp Leu Pro Asp Ala Met Ala Ala Pro
 225 230 235 240
 Thr Pro Ala Ser Ala Phe Leu His Gly Ala Ala Met Val Glu Met Gly
 245 250 255
 Val Tyr Leu Leu Ala Arg Val Thr Gln Phe Met Gln Pro Ile Pro Glu
 260 265 270
 Thr Ala Phe Tyr Val Met Leu Val Phe Val Ser Leu Thr Leu Leu Ile
 275 280 285
 Ala Ile Leu Tyr Tyr Pro Leu Gln Lys Asp Ala Lys Arg Leu Leu Ala
 290 295 300
 Tyr Ser Thr Ile Ala Glu Ala Gly Val Met Tyr Val Gly Val Leu Tyr
 305 310 315 320
 Ala Val Leu Gly Ser Val Tyr Gly Leu Gln Ala Ala Met Phe Gln Leu
 325 330 335
 Ala Asn His Ala Phe Val Lys Gly Leu Ala Phe Leu Thr Ala Gly Thr

Ala Met Pro Tyr Ser Ser Leu Phe Thr Val Ile Gly Ala Leu Ser Leu
 850 855 860

Ala Gly Val Pro Leu Phe Ser Gly Phe Leu Ser Lys Trp Met Ile Tyr
 865 870 875 880

Gln Ala Gly Tyr Tyr Ser Gly Ile Gly Leu Phe Val Phe Gly Ser Val
 885 890 895

Met Ala Val Phe Met Ser Ala Val Thr Leu Ala Tyr Ser Leu Lys Leu
 900 905 910

Tyr Thr Ser Ala Phe Gly Gly Glu Pro Asn Glu Arg Thr Glu Asn Ala
 915 920 925

Arg Glu Val Pro Ser Gly Met Leu Leu Gly Glu Gly Ile Ile Ala Leu
 930 935 940

Thr Ser Leu Ala Val Gly Ile Leu Pro Ala Ile Ala Tyr Pro Ile Leu
 945 950 955 960

Thr Ile Ser Leu Asn Gly Gly Asp Val Thr Val Thr Met Gly Ser Ile
 965 970 975

Ser Thr Asp Phe Glu Tyr Phe Ser Pro Ile Ala Leu Leu Leu Ala Val
 980 985 990

Ser Phe Ile Ala Val Ala Ser Tyr Phe Val Phe Arg Pro Lys Thr Thr
 995 1000 1005

Asn Val Lys Pro Trp Asn Thr Gly Ala Leu Phe Leu Pro Glu Glu Arg
 1010 1015 1020

Tyr Gly Ala Lys Ala Arg Asp Tyr Tyr Arg Gln Tyr Phe Thr Glu Met
 1025 1030 1035 1040

Glu Gly Leu Tyr Lys Leu Gly Ser Ala Ala Gly Lys Val Gly Arg Val
 1045 1050 1055

Leu Leu Ser Ala Leu Met Ser Val Tyr Leu Val Leu Ala Arg Gly Leu
 1060 1065 1070

Val Tyr Thr Gly Arg Glu Lys Lys Arg Ser Phe Thr Leu Asp Glu Leu
 1075 1080 1085

Arg His Arg Thr Val Arg Tyr Leu Asp Glu Ala Phe Phe Ala Pro Met
 1090 1095 1100

Met Asp Leu Leu Lys Asn Ile Ala Val Leu Ala Ala Gly Ile Ser Val

1105 1110 1115 1120

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

1125 1130 1135

Leu Ala Leu Leu Val Leu Met Asp Tyr Val Ser Ile Ile Ala Ala Pro

1140 1145 1150

Ile Val Leu Phe Leu Leu Pro Pro Phe Leu Asp Gly Ile Gly Arg Arg

1155 1160 1165

Ile Lys Ala Arg Ile Gln Tyr Arg Arg Gly Pro Pro Ile Met Gln Thr

1170 1175 1180

Phe Tyr Asp Leu Glu Lys Leu Leu Lys Leu Pro Ser Val Leu Pro Thr

1185 1190 1195 1200

Glu Gly Pro Ile Phe Arg Leu Ala Pro Tyr Ile Ala Leu Ala Ser Ala

1205 1210 1215

Ile Ala Gly Gly Leu Met Leu Pro Phe Gly Ser Glu Pro Val Leu Ala

1220 1225 1230

Phe Gly Lys Ser Leu Ile Val Phe Phe Tyr Val Met Ala Met Val Ser

1235 1240 1245

Val Val Met Ile Leu Ala Ala Phe Ser Val Gln Asn Ala Phe Ser His

1250 1255 1260

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

1265 1270 1275 1280

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

1285 1290 1295

Val Ala Glu Met Pro Phe Ser Ala Asn Leu Ser Leu Ser Val Ala Leu

1300 1305 1310

Ala Tyr Ile Leu Leu Ala Tyr Ala Val Tyr Val Glu Gly Gly Phe Val

1315 1320 1325

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

1330 1335 1340

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

1345 1350 1355 1360
 Val Lys Arg Val Val Leu Leu Trp Leu Leu Ala Ser Met Ile Val Ile
 1365 1370 1375
 Pro Ala Met Arg Ser Leu Gly Ile Thr Ser Ser Met Ala Leu Leu Val
 1380 1385 1390
 Ala Gln Leu Val Val Thr Phe Leu Leu Tyr Ser Leu Ala Val Ala Val

 1395 1400 1405
 Glu Ala Ala Asn Ala Arg Leu Arg Ile Asp Gln Ala Val Ser Leu Asn
 1410 1415 1420
 Lys Lys Val Phe Leu Met Ser Leu Ala Val Leu Ile Ile Ala Leu Val
 1425 1430 1435 1440
 Gly Trp Met Glu Cys Ser Val Cys Ala Gly Gly Cys Arg Ser Ala Glu
 1445 1450 1455
 Val Glu Asp Val Leu Glu Asp Gly His Leu Lys Glu Phe Val Glu Lys
 1460 1465 1470

 Phe Arg Gly Ala Ile Phe Glu Cys Lys Lys Leu Thr Arg Asn Gln Tyr
 1475 1480 1485
 Leu Phe Ile Val Asp Arg Glu Ala Leu Pro Glu Met Val Leu His Trp
 1490 1495 1500
 His Asn His Ser Glu Leu Lys Glu Thr His Phe Ser Met Gly Thr Gly
 1505 1510 1515 1520
 Thr Asp Glu Arg Asn Ile Ala Gly Lys Phe Thr Tyr Ala Pro Val Ile
 1525 1530 1535
 Asn Val Ala Val Glu Pro Gly Asn Gly Glu Arg Asn Tyr Trp Val Ile

 1540 1545 1550
 Leu Lys Ala Tyr Leu Asp Glu Asp Asn Pro Glu Phe Pro Ser Ile Ala
 1555 1560 1565
 Ala Lys Leu Pro Ala Ala Leu Trp Ala Glu Arg Glu Val Tyr Asp Leu
 1570 1575 1580
 Leu Gly Phe Asn Pro Lys Gly His Pro Asp Leu Arg Arg Leu Val Leu
 1585 1590 1595 1600
 Pro Glu Asp Trp Pro Glu Gly Val Tyr Pro Leu Arg Lys Asp His Asp

Asn Ile Asp Ile Arg Arg Asp Gln Pro Phe Ala Ala Tyr Lys Asp Leu
 1860 1865 1870
 Asp Phe Lys Val Pro Val Tyr Lys Glu Gly Asp Val Leu Ala Arg Phe
 1875 1880 1885
 Leu Ile Arg Met Asp Glu Val Leu Glu Ser Ile Trp Ile Ile Glu Gln
 1890 1895 1900

 Ala Ile Asp Gln Met Pro Gly Gly Asp Val Phe Val Pro Ile Gly Glu
 1905 1910 1915 1920
 Leu Pro Glu Tyr Glu Glu Ala Leu Gly Tyr Ser Glu Ala Pro Arg Gly
 1925 1930 1935
 Glu Val Ile His Tyr Val Met Thr Asp Lys Lys Asn Lys Val Tyr Arg
 1940 1945 1950
 Trp Lys Val Arg Ala Pro Thr Tyr Asn Asn Leu Pro Ala Val Pro Glu
 1955 1960 1965
 Met Leu Lys Gly Tyr Ser Val Ala Asp Ala Pro Leu Ile Ile Ala Ser

 1970 1975 1980
 Ile Asp Pro Cys Tyr Ser Cys Thr Glu Arg Val Gln Ile Val Asp Val
 1985 1990 1995 2000
 Glu Thr Gly Lys Ala Gln Thr Leu Asn Glu Gln Gln Phe Asn Met Leu
 2005 2010 2015
 Ser Ile Gln Lys Gly Lys Gly Val Ala Met Ala Gln Ala Ile Ser Phe
 2020 2025 2030
 Thr Asp Arg Leu Lys Phe Trp Lys Arg Pro Glu Glu Asp Val Lys Lys
 2035 2040 2045

 Ala Pro Val Thr Thr Ser Tyr Pro Phe Val Asp Ile Glu Lys Pro Pro
 2050 2055 2060
 Glu Tyr Arg Gly Ile Pro Arg Ile Asp Pro His Leu Cys Ile Gly Cys
 2065 2070 2075 2080
 Gly Ala Cys Val Arg Ala Cys Pro Pro Asp Ala Leu Thr Ile Glu Trp
 2085 2090 2095
 Asp Phe Glu Asn Gly Arg Lys Arg Ile Val Phe Asn Ala Ala Arg Cys
 2100 2105 2110

Ile Arg Cys His Arg Cys Val Glu Val Cys Pro Thr Gly Ala Met Gln

2115 2120 2125
 Gly Thr Thr Arg Phe Glu Ile Ala Thr Pro Asn Lys Glu Asp Leu Ile
 2130 2135 2140
 Glu Val Val Asp His Lys Leu Tyr Arg Cys Pro Arg Cys Gly Arg Tyr
 2145 2150 2155 2160
 Glu Glu Phe Thr Glu Arg Gln Ile Gly Lys Met Phe Gln Ile Leu Pro
 2165 2170 2175
 Glu Glu Val Ile Asp Gln His Gly Ile Ala Glu Arg Ala Phe Leu Cys
 2180 2185 2190

Arg Glu Cys Arg Met Glu Glu Ser Ala Lys Thr Leu Ala Val Gln Gly
 2195 2200 2205
 Pro Tyr Ala Asp Ser Leu Leu Leu Ser Leu Tyr Pro Arg Gly Ser Lys
 2210 2215 2220
 Val Met Gly Glu Arg Arg Met Ser Gly Leu Lys Ser Val Trp Val Phe
 2225 2230 2235 2240
 His Val Asp Ser Gly Ser Cys Asn Gly Cys Asp Ile Glu Ile Leu Asp
 2245 2250 2255
 Val Leu Thr Pro Tyr Tyr Asp Ala Glu Arg Leu Gly Ile Lys Leu Val

2260 2265 2270
 Pro Ser Pro Arg His Ala Asp Ala Leu Leu Val Ser Gly Pro Leu Thr
 2275 2280 2285
 Arg Gln Thr Tyr Tyr Ala Val Lys Ala Ala Tyr Glu Ala Met Pro Pro
 2290 2295 2300
 Lys Pro Arg Ile Val Val Ala Ile Gly Thr Cys Ala Ser Ser Gly Gly
 2305 2310 2315 2320
 Ile Phe Tyr Asn Gly Tyr Pro Ile Tyr Asn Pro Asn Pro Glu Arg Gly
 2325 2330 2335

Ser Asp Arg Leu Arg Thr Gly Gly Ile Glu Val Leu Leu Ala Glu Tyr
 2340 2345 2350
 Gly Lys Lys Pro Asp Met Tyr Ile Pro Gly Cys Pro Pro Ser Pro Glu

2355 2360 2365
 Glu Ile Leu Tyr Gly Leu Ala Gln Leu Leu Gly Leu Lys Glu Lys Lys
 2370 2375 2380
 Met Lys Gly Glu Tyr Tyr Tyr Ala Asp Glu Ile Glu Phe Val Leu Pro
 2385 2390 2395 2400
 Glu Arg Pro Ile Glu Glu Arg Ile Tyr Leu Thr Leu Arg Glu Ser Leu

 2405 2410 2415
 Arg Arg Val Val Gly Tyr Phe Asp Arg Glu Lys Val Leu Glu Asp Phe
 2420 2425 2430
 Met Ala Leu Val Glu Lys Ala Gln Glu Ser Glu Asn Pro Arg Glu Arg
 2435 2440 2445
 Leu His Glu Leu Val Ile Gly Tyr Phe Leu Arg Glu Lys Asp Ser Arg
 2450 2455 2460
 Val Lys Phe Ala Ile Arg Phe Leu Glu Asn Glu Tyr Trp Arg Leu Lys
 2465 2470 2475 2480

 Asp Ala Tyr Glu Lys Arg His Leu Ala Leu Val Lys Ala Gly Val Arg
 2485 2490 2495
 <210> 3
 <211> 990
 <212> PRT
 <213> Thermococcus onnurineus NA1
 <400> 3
 Met Val Asp Trp Arg Leu Phe Glu Pro Leu Phe Asn Tyr Ala Arg Lys
 1 5 10 15
 Lys Ser Leu Trp Ile Val Ser Phe Cys Thr Gly Cys Gly Gly Ile Glu
 20 25 30
 Met Pro Pro Leu Met Thr Ser Arg Tyr Asp Leu Glu Arg Phe Gly Met

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

Ala Pro Asp Pro Lys Tyr Val Leu Ala His Gly Ser Cys Pro Leu Asn
85 90 95

Gly Gly Ile Tyr Trp Asp Ala Tyr Asn Ala Ile Lys His Leu Asp Lys
100 105 110

Tyr Ile Pro Val Asp Val Val Ile Ala Gly Cys Met Pro Arg Pro Glu
115 120 125

Ala Val Met Asp Gly Ile Gln Lys Ile Met Glu Met Ile Glu Asn Gly
130 135 140

Thr Ala Asp Gly Trp Lys Arg Tyr Lys Glu Asn Tyr Glu Trp Tyr Lys
145 150 155 160

Lys Asn Gln Asp Glu Leu Phe Gly Glu Gly Trp Arg Glu Arg Glu Ala
165 170 175

Arg Arg Trp Ile Pro Trp Leu Val Asp Lys Lys Lys Glu Glu Met Gly
180 185 190

Glu Val Lys Trp Glu Arg Glu Gln Met Leu Val Asp Lys Ile Leu Glu
195 200 205

Lys Ala Pro Tyr Ala Glu Gly Lys Val Arg Arg Glu Arg Arg Ile Glu
210 215 220

Phe Ser Ile Pro Ala Asp Arg Ile Arg Asp Phe Leu Met Leu Leu Arg
225 230 235 240

Asp Asn Asp Phe Glu Leu Met Leu Gln Ile Thr Thr Val Asp Trp Pro
245 250 255

Asn Asp Gly Glu Leu Glu Leu Ile Tyr Gln Met Trp Ser Val Thr His
260 265 270

Arg Thr His Ala Met Val Arg Thr Arg Ile Pro Arg Asp Leu Asp Lys
275 280 285

Ala Arg Met Pro Thr Val Lys Asp Ile Tyr Pro Val Ala Glu Thr Tyr
290 295 300

Glu Arg Asp Ala His Asp Phe Phe Gly Val Tyr Phe Glu Gly Asn Glu
305 310 315 320

Lys Met Glu Met Pro Trp Ile Leu Asp Asp Thr Glu Gln Gly Leu Phe

Ile Glu Met Thr Trp Ile Glu Val Asp Gly Lys Met Glu Lys Arg Pro
 835 840 845
 His Pro Lys Ile Asp Tyr Gly Arg Cys Thr Phe Cys Glu Phe Cys Val
 850 855 860
 Asp Val Cys Pro Pro Gly Ala Leu Gly Phe Ile Glu Asn Tyr Ile Leu
 865 870 875 880
 Thr Thr Glu Trp Lys Asp Glu Glu Leu Glu Leu Phe Asp Trp Val Pro
 885 890 895
 Ile His Pro Asp Lys Phe Arg Glu Ile Asn Glu Lys Phe Pro Asp Tyr
 900 905 910
 Arg Phe Pro Val Glu Lys Ile Glu Phe Asn Lys Glu Thr Lys Glu Val
 915 920 925
 Thr Tyr Tyr Leu Arg Asp Gly Glu Val Met Lys Phe Lys Ile Leu Gly
 930 935 940
 Tyr Gly Ile Arg Pro Pro Lys Pro Pro Thr Lys Pro Ala Gln Lys Ala
 945 950 955 960
 Ala Ala Lys Ala Ala Glu Lys Asn Asp Thr Lys Pro Val Glu Lys Pro
 965 970 975

 Thr Glu Lys Lys Glu Ala Gly Lys Ile Glu Glu Lys Lys Glu
 980 985 990
 <210> 4
 <211> 3724
 <212> PRT
 <213> Thermococcus onnurineus NA1
 <400> 4
 Met Glu Thr Leu Ile Leu Ala Leu Gly Asn Glu Leu Met Lys Asp Asp
 1 5 10 15
 Gly Val Gly Leu Lys Val Gly Arg Leu Leu Ala Glu Lys Gly Tyr Asn
 20 25 30
 Val Leu Glu Val Gly Thr Asp Ile Phe Met Leu Gln Ser His Tyr Ser
 35 40 45
 Gly Glu Glu Arg Leu Ile Ile Ile Asp Ala Ile Leu Ser Glu Lys Phe

305 310 315 320
 Leu Pro Asp Lys Thr Thr Leu Glu Lys Met Lys Lys Arg Leu Gln Glu

 325 330 335
 Ala Leu Pro Leu Ala Glu Tyr Thr Phe Glu Leu Phe Ser Lys Leu Glu
 340 345 350
 Gln Tyr Glu Glu Val Glu Gly Pro Ile Ile His Met Ala Val Arg Pro
 355 360 365
 Arg Gly Asp Val Tyr Gly Ile Tyr Gly Asp Ala Ile Ser Val Ser Asp
 370 375 380
 Gly Phe Glu Phe Pro Ser Glu Gly Tyr Lys Lys His Met Val Glu Arg
 385 390 395 400

 Val Val Glu His Ser Phe Ala Lys His Ser Phe Tyr Asn Gly Glu Lys
 405 410 415
 Pro Phe Met Thr Gly Ala Ile Ser Arg Val Val Asn His Ala Asp Lys
 420 425 430
 Leu Tyr Gly Arg Ala Lys Glu Leu Tyr Glu Ser His Lys Asp Leu Leu
 435 440 445
 Arg Pro Thr Asn Pro Phe Ala Asn Asn Leu Ala Gln Ala Leu Glu Leu
 450 455 460
 Val Tyr Phe Met Glu Arg Gly Ile Asp Leu Ile Asp Glu Ala Leu Ala

 465 470 475 480
 Lys Trp Pro Ile Arg Pro Arg Asp Glu Val Asp Val Lys Asp Gly Phe
 485 490 495
 Gly Val Ser Thr Thr Glu Ala Pro Arg Gly Ile Leu Val Tyr Ala Leu
 500 505 510
 Glu Val Lys Asp Gly Arg Val Ala Tyr Ala Asp Ile Ile Thr Pro Thr
 515 520 525
 Ala Phe Asn Leu Ala Met Met Glu Val His Val Arg Met Met Ala Glu
 530 535 540

 Lys His Tyr Asn Asp Asp Pro Glu Arg Leu Lys Leu Leu Ala Glu Met
 545 550 555 560

Val Val Arg Ala Tyr Asp Pro Cys Ile Ser Cys Ser Val His Val Ala
 565 570 575

Arg Leu Met Glu Gly Lys Val Arg Ile Gly Phe Tyr Ala Leu Thr Ser
 580 585 590

Cys Tyr Gly Cys Gln Leu Arg Phe Ala Met Met Asp Glu Ile Leu Gln
 595 600 605

Leu Leu Pro Asn Ala Glu Ile Val Cys Trp Tyr Met Leu Asp Arg Asp
 610 615 620

Ser Ser Glu Asp Glu Pro Val Asp Ile Ala Phe Ile Glu Gly Ser Val
 625 630 635 640

Ser Thr Glu Glu Glu Val Glu Leu Val Lys Lys Ile Arg Glu Asn Ala
 645 650 655

Lys Ile Val Val Ala Val Gly Ala Cys Ala Thr Gln Gly Gly Val Gln
 660 665 670

Ser Trp Glu Lys Asp Lys Ser Leu Glu Glu Leu Trp Lys Ala Val Tyr
 675 680 685

Gly Asp Gly Lys Val Lys Phe Glu Pro Lys Met Ala Glu Pro Leu Glu
 690 695 700

Asn Tyr Ile Lys Val Asp Tyr Arg Ile Tyr Gly Cys Pro Pro Glu Lys
 705 710 715 720

Lys Asp Phe Ile Tyr Ala Ile Gly Thr Phe Leu Val Gly Ser Trp Pro
 725 730 735

Glu Asp Ile Asp Tyr Pro Val Cys Leu Glu Cys Arg Leu Lys Gly Asn
 740 745 750

Thr Cys Ile Leu Ile Glu Lys Gly Glu Pro Cys Leu Gly Pro Ile Thr
 755 760 765

Arg Ala Gly Cys Asp Ala Arg Cys Pro Ser Tyr Gly Ile Ala Cys Ile
 770 775 780

Gly Cys Arg Gly Ala Ile Gly Tyr Asp Val Ala Trp Phe Asp Ser Leu
 785 790 795 800

Ala Arg Thr Phe Lys Glu Lys Gly Leu Thr Lys Glu Glu Ile Leu Glu
 805 810 815

Arg Met Lys Ile Phe Asn Ala His Asn Pro Lys Leu Glu Glu Met Val
 820 825 830

Asp Lys Val Phe Gln Phe Gln Gly Val Lys Glu Met Asn Glu Ala His
 835 840 845

Val Cys Met Cys His Asp Asn Pro Tyr Ala Leu Asp Arg Val Lys Val
 850 855 860

Leu Arg Val Tyr Arg Leu Thr Glu Thr Glu Lys Leu Phe Leu Phe Arg
 865 870 875 880

Phe Glu Asp Gln Glu Ile Ala Glu Asn Trp Thr Phe Lys Pro Gly Gln
 885 890 895

Phe Val Gln Leu Thr Ile Pro Gly Val Gly Glu Val Pro Ile Ser Ile
 900 905 910

Cys Ser Ser Pro Met Lys Arg Gly Phe Phe Glu Leu Cys Ile Arg Lys
 915 920 925

Ala Gly Arg Val Thr Thr Val Val His Lys Leu Lys Pro Gly Asp Thr
 930 935 940

Val Leu Val Arg Gly Pro Tyr Gly Asn Gly Phe Pro Val Asp Glu Trp
 945 950 955 960

Glu Gly Met Asp Leu Leu Leu Ile Ala Ala Gly Leu Gly Thr Ala Pro
 965 970 975

Leu Arg Ser Val Phe Leu Tyr Ala Met Asp Asn Arg Trp Lys Tyr Gly
 980 985 990

Asn Ile Thr Phe Ile Asn Thr Ala Arg Tyr Gly Lys Asp Leu Leu Phe
 995 1000 1005

Tyr Lys Glu Leu Glu Ala Met Lys Asp Leu Ala Glu Ala Glu Asn Val
 1010 1015 1020

Gln Ile Ile Gln Ser Val Thr Arg Asp Pro Asp Trp Pro Gly Arg His
 1025 1030 1035 1040

Gly Arg Pro Gln Lys Phe Ile Val Glu Ala Asn Thr Asn Pro Lys Asn
 1045 1050 1055

Thr Ala Ile Ala Ile Cys Gly Pro Pro Arg Met Tyr Lys Ala Val Phe

1060 1065 1070
 Glu Ala Leu Ile Asn Tyr Gly Tyr Arg Pro Glu Asn Ile Tyr Val Thr
 1075 1080 1085
 Leu Glu Arg Lys Met Lys Cys Gly Ile Gly Lys Cys Gly His Cys Asn
 1090 1095 1100
 Val Gly Thr Ser Thr Ser Trp Lys Tyr Val Cys Arg Asp Gly Pro Val
 1105 1110 1115 1120

 Phe Thr Tyr Phe Asp Ile Val Ser Thr Pro Gly Leu Leu Asp Met Arg
 1125 1130 1135
 Tyr Val Lys Leu Pro Lys Glu Asn Thr Tyr Glu Phe Leu Glu Arg Leu
 1140 1145 1150
 Lys Asn Leu Gly Lys Leu Tyr Ala Pro Val Lys Ile Ser Asp Gln Phe
 1155 1160 1165
 Tyr Asp Phe Arg Glu Ile Asp Asp Val Arg Lys Ile Glu Phe Asn Tyr
 1170 1175 1180
 Thr Arg Thr Leu Met Pro Pro Lys Lys Phe Phe Phe Ala Pro Arg Glu

 1185 1190 1195 1200
 Lys Met Phe Glu Phe Ser Ile Ser Lys Ala Glu Tyr Arg Glu Val Ile
 1205 1210 1215
 Pro Glu Val Glu Pro Phe Val Leu Phe Gly Leu His Ala Cys Asp Ile
 1220 1225 1230
 Tyr Gly Leu Lys Ile Leu Asp Ser Val Tyr Leu Asp Glu Tyr Pro Asp
 1235 1240 1245
 Lys Tyr Tyr Lys Val Arg Arg Glu Lys Gly Ile Ile Ile Gly Ile Ser
 1250 1255 1260

 Cys Met Pro Asp Glu Tyr Cys Phe Cys Asn Leu Leu Arg Thr Asp Phe
 1265 1270 1275 1280
 Glu His Asp Gly Phe Asp Leu Phe Phe His Glu Leu Pro Asp Gly Trp
 1285 1290 1295
 Leu Ile Arg Ile Gly Thr Pro Thr Gly His Arg Ile Val Asp Lys Asn
 1300 1305 1310
 Ile Lys Leu Phe Thr Glu Val Ala Gln Glu Asp Ile Cys Asn Phe Arg

1315 1320 1325
 Glu Phe Glu Arg Lys Arg Ala Gln Ala Phe Arg Tyr His Glu Glu Trp

1330 1335 1340
 Asp Asn Ile His Tyr Leu Leu Glu Leu Glu Met Glu His Pro Leu Trp
 1345 1350 1355 1360

Glu Lys Glu Ala Glu Lys Cys Phe Ala Cys Gly Asn Cys Ser Thr Val
 1365 1370 1375

Cys Pro Thr Cys Arg Cys Tyr Glu Val Gln Asp Ile Val Asn Leu Asp
 1380 1385 1390

Gly Asp Thr Gly Tyr Arg Glu Arg Arg Trp Asp Ser Cys Lys Phe Arg
 1395 1400 1405

Ser His Gly Leu Val Ala Gly Gly His Asn Phe Arg Pro Thr Lys Lys
 1410 1415 1420

Asp Arg Phe Ile Asn Arg Tyr Leu Cys Lys Met Ser Phe His Trp Thr
 1425 1430 1435 1440

Leu Gly Ile Asn Phe Cys Val Gly Cys Gly Arg Cys Thr Ala Phe Cys
 1445 1450 1455

Pro Ala Gly Ile Asp Phe Val Lys Asn Leu Arg Ile Ile Ala Gly Leu
 1460 1465 1470

Glu Asp Ala Ser Cys Pro Ser Lys Leu Ser Glu Glu Ile Pro Lys Lys

1475 1480 1485
 Gly Phe Ala Tyr Ala Asn Asn Ile Arg Gly Glu Asp Ile Met Ala Gln

1490 1495 1500
 Asn Asn Ser Leu Val Leu Tyr Asp Val His Glu Thr Val Asp Val Cys

1505 1510 1515 1520
 Ser Asn Val Gly Cys Val Lys Thr Lys Ala Thr Pro Ser Arg Leu Leu

1525 1530 1535
 Phe Ala Gly Phe Met Ala Gly Ala Tyr Ile Ala Phe Gly Phe Ile Phe

1540 1545 1550

Ala Ile Val Ala Ser Ala Ser Phe His Pro Lys Leu Gly Thr Phe Pro
 1555 1560 1565

Asn Leu Ser Leu Phe Lys Leu Leu Leu Gly Ala Val Phe Pro Val Gly
 1570 1575 1580
 Leu Ile Ala Val Leu Leu Gly Gly Ala Asp Leu Trp Thr Gly Asn Ala
 1585 1590 1595 1600
 His Ile Val Thr Leu Ser Lys Met Thr Gly Arg Ala Ser Val Lys Asp
 1605 1610 1615
 Val Leu Tyr Asn Trp Ile Gly Ser Tyr Thr Gly Asn Phe Val Gly Ser
 1620 1625 1630
 Val Phe Leu Ala Phe Leu Ala Val Tyr Gly Thr Gly Leu Met Ala Gly
 1635 1640 1645
 Gly Leu Phe Lys Asp Val Leu Ile Gly Ile Gly Asn Tyr Lys Val Ala
 1650 1655 1660
 Leu Thr Pro Trp Lys Ala Leu Trp Leu Gly Ile Gly Cys Asn Trp Leu
 1665 1670 1675 1680
 Val Asn Val Ala Ile Trp Leu Tyr Ile Arg Ala Lys Asp Thr Ala Gly
 1685 1690 1695

 Lys Val Ile Val Thr Trp Phe Pro Ile Phe Ala Phe Val Ala Ile Gly
 1700 1705 1710
 Phe Glu His Ser Ile Ala Asn Met Trp Ala Ile Ser Ala Ser Ile Phe
 1715 1720 1725
 Ala Ser Asp Gly Ala Ile Ser Trp Val Gln Phe Phe His Asn Ile Ile
 1730 1735 1740
 Pro Val Thr Ile Gly Asn Ala Ile Gly Gly Phe Leu Phe Val Gly Phe
 1745 1750 1755 1760
 Tyr His Trp Tyr Leu Ala Asp Gly Arg Asn Ala Ile Lys Glu Leu Ile
 1765 1770 1775
 Asp Phe Val Glu Val Leu Ala Leu Phe Val Phe Ile Met Val Leu Ile
 1780 1785 1790
 Pro Ala Gly Ile Ala Tyr Ala Leu Ser Gly Leu Gly Asn Ile Ala Thr
 1795 1800 1805
 Trp Leu Val Pro Leu Ile Ile Ser Val Tyr Gly Val Val Met Thr Tyr
 1810 1815 1820

Leu Val Arg Arg Ala Leu Met Glu Glu Phe Lys Ile Gly Leu Cys Pro
 1825 1830 1835 1840

Tyr Cys Gly Met Gly Cys Arg Phe Tyr Ile Lys Thr Leu Asn Gly Gln
 1845 1850 1855

Pro Ile Gly Ile Glu Pro Tyr Pro Gly Gly Val Asn Glu Gly Lys Leu
 1860 1865 1870

Cys Pro Lys Gly Val Ala Ala Val Asp Phe Leu Arg His Lys Asp Arg
 1875 1880 1885

Leu Lys Lys Pro Leu Lys Arg Thr Glu Asn Gly Phe Val Glu Ile Ser
 1890 1895 1900

Trp Glu Gln Ala Ile Lys Glu Ile Ala Glu Lys Leu Leu Glu Ile Arg

 1905 1910 1915 1920

Glu Lys Tyr Gly Pro Asp Thr Leu Gly Phe Phe Ser Ser Ala Arg Cys
 1925 1930 1935

Ser Asn Glu Glu Asn Tyr Leu Leu Gln Lys Ile Ala Arg Leu Leu Gly
 1940 1945 1950

Thr Asn Asn Val Asp His Cys Ala Arg Leu Cys His Ala Ser Thr Val
 1955 1960 1965

Val Gly Leu Ala Gln Thr Val Gly Ala Ala Ala Gln Ser Gly Ser Tyr
 1970 1975 1980

Thr Asp Ile Pro Lys Ala Lys Val Leu Leu Ile Trp Gly Tyr Asn Pro
 1985 1990 1995 2000

Ser Glu Thr His Pro Val Leu Met Arg Tyr Ile Leu Arg Ala Arg Asp
 2005 2010 2015

Asn Gly Ala Lys Ile Ile Val Val Asp Pro Arg Lys Thr Arg Thr Val
 2020 2025 2030

Trp Phe Ala Asp Met His Leu Gln Leu Lys Pro Gly Thr Asp Ile Val
 2035 2040 2045

Leu Ala Asn Ala Met Met His Val Ile Ile Glu Glu Arg Leu Tyr Asp

 2050 2055 2060

Arg Glu Phe Ile Met Asn Arg Thr Lys Gly Phe Glu Lys Leu Ile Ala

2065 2070 2075 2080
 Ala Val Gln Lys Tyr Thr Pro Glu Tyr Ala Glu Glu Ile Thr Gly Val
 2085 2090 2095
 Pro Ala Lys Leu Ile Arg Glu Ala Ala Ile Thr Phe Ala Thr Ala Gly
 2100 2105 2110
 Arg Gly Ile Val Met Trp Ala Met Gly Leu Thr Gln His Val Thr Gly
 2115 2120 2125

 Ala Ala Asn Val Lys Ala Leu Ala Asp Leu Ala Leu Ile Cys Gly Tyr
 2130 2135 2140
 Val Gly Arg Glu Gly Thr Gly Leu Phe Pro Met Arg Gly Gln Asn Asn
 2145 2150 2155 2160
 Val Gln Gly Ala Cys Asp Met Ala Ala Leu Pro Asn Val Phe Pro Gly
 2165 2170 2175
 Tyr Gln Lys Val Thr Asp Asp Glu Lys Arg Lys His Val Ala Glu Ile
 2180 2185 2190
 Trp Gly Val Glu Asp Leu Pro Ser Lys Pro Gly Leu Thr Ile Pro Glu

 2195 2200 2205
 Met Ile Asp Ala Ala Ala Lys Gly Glu Leu Lys Ala Leu Tyr Ile Met
 2210 2215 2220
 Gly Glu Asn Pro Val Met Ser Asp Pro Asn Thr Lys His Val Ile Glu
 2225 2230 2235 2240
 Ala Leu Lys Asn Leu Glu Leu Leu Val Val Gln Asp Ile Phe Leu Thr
 2245 2250 2255
 Glu Thr Ala Glu Leu Ala His Tyr Val Leu Pro Ala Ala Ala Tyr Ala
 2260 2265 2270

 Glu Lys Glu Gly Ser Phe Thr Ala Ser Glu Arg Arg Val Gln Trp Asn
 2275 2280 2285
 Phe Lys Ala Ile Glu Pro Pro Gly Glu Ala Lys Pro Asp Trp Glu Ile
 2290 2295 2300
 Leu Thr Met Leu Gly Lys Ala Leu Gly Leu Pro Lys Phe Asp Tyr Ser
 2305 2310 2315 2320
 Asp Val Glu Asp Ile Thr Arg Glu Ile Thr Leu Val Ala Pro Gln Tyr

Asp Gly Ala Val Val Ile Asp Glu Ser Lys Cys Ile Gly Cys Tyr Met
 2580 2585 2590
 Cys Ser Ala Val Cys Pro Tyr Ala Ile Pro Ile Val Asp Pro Ile Lys
 2595 2600 2605
 Glu Leu Ala Val Lys Cys Asp Leu Cys Ala Glu Arg Arg Lys Glu Gly
 2610 2615 2620
 Arg Asp Pro Leu Cys Ala Ala Val Cys Pro Thr Asp Ala Ile Ile Tyr
 2625 2630 2635 2640
 Ala Asp Leu Asn Glu Leu Met Glu Glu Lys Arg Arg Arg Lys Ala Glu
 2645 2650 2655
 Arg Ile Val Glu Ala Gln Arg Lys Ala Val Glu Thr Leu Ala Tyr Phe
 2660 2665 2670
 Gly Val Leu Lys Val Glu Leu Cys Val Gly Cys Gly Val Cys Ala Lys
 2675 2680 2685
 Ala Cys Pro His Ser Ala Ile Ser Val Phe Glu Asp Ser Val Arg Arg
 2690 2695 2700
 Ile Val Phe Asp Pro Lys Lys Cys Glu Glu Cys Ser Phe Glu Cys Asn
 2705 2710 2715 2720
 Glu Ala Cys Pro Thr Gly Ala Leu Glu Gly Lys Ser Asp Lys Arg Glu
 2725 2730 2735
 Leu Val Phe Glu Phe Ala Tyr Cys Ala Ile Cys Gly Lys Arg Leu Asn
 2740 2745 2750
 Ile Val Lys Glu Glu Ala Glu Tyr Leu Ala Lys Lys Leu Ile Glu Leu
 2755 2760 2765
 Gly Glu Asn Pro Glu Ile Ala Phe Leu Cys Asp Asp Cys Lys Arg Lys
 2770 2775 2780
 Arg Leu Phe Gly Val Ala Asn Lys Tyr Glu Ala Tyr Leu Gly Met Ser
 2785 2790 2795 2800
 Gly Met Arg Phe Ala Phe Leu Cys Arg Glu Arg Pro Glu Pro Thr Gly
 2805 2810 2815
 Lys Lys Ile Ala Val Ile Gly Ala Gly Pro Ala Gly Leu Ala Ala Thr
 2820 2825 2830

3075 3080 3085
 Phe Ala Glu Glu Ile Asp Ile Ala Val Asp Lys Lys Gly Arg Ile Val
 3090 3095 3100
 Val Asp Asn Arg His Met Thr Ser Arg Glu Gly Val Phe Ala Ala Gly
 3105 3110 3115 3120
 Asp Val Val Leu Gly Pro Ser Lys Val Gly Lys Ala Val Lys Asp Gly
 3125 3130 3135

 Leu Tyr Ala Ala Glu Ala Met His Met Trp Leu Met Gly Arg Met Thr
 3140 3145 3150
 Arg Arg Ile Leu His Val Asp Tyr Ser Leu Cys Ile Gly Cys Glu Thr
 3155 3160 3165
 Cys Glu Ala Val Cys Asp Phe Leu His Gly Gly Lys Pro Asn Ile Arg
 3170 3175 3180
 Ile Tyr Tyr Thr Val Thr Gly Leu Pro Ile Pro Ile Asn Cys Arg His
 3185 3190 3195 3200
 Cys Glu Arg Ala Pro Cys Met Asp Val Cys Pro Ala Gly Ala Ile Tyr

 3205 3210 3215
 Arg Asp Ser Asp Gly Ala Ile Ile Ile Asn Pro Asp Lys Cys Ile Gly
 3220 3225 3230
 Cys Tyr Met Cys Leu Ala Val Cys Pro Phe Gly Val Pro Ser Phe Asp
 3235 3240 3245
 Val Lys Thr Lys Ala Val Thr Lys Cys Asp Met Cys Ala Asp Arg Arg
 3250 3255 3260
 Arg Leu Gly Met Glu Pro Ala Cys Ala Glu Met Cys Pro Ala Glu Ala
 3265 3270 3275 3280

 Ile Phe Phe Gly Lys Pro Glu Glu Val Glu Asp Arg Ile Arg Arg Arg
 3285 3290 3295
 Thr Ala Glu Arg Ile Ala Arg Glu Arg Ile Ala Ala Val Asp Met Glu
 3300 3305 3310
 Gly Val Gly Arg Met Leu Met Leu Trp Glu Ser Gln Ile Pro Ile Asn
 3315 3320 3325
 Gln Val Phe Glu Leu Arg Cys Arg Ser Met Thr Tyr Phe Gly Val Gly

3330 3335 3340
 Ala Ile Asn Lys Phe Tyr Asp Ile Ala Lys Asp Leu Lys Glu Asn Arg

 3345 3350 3355 3360
 Gly Ile Thr Lys Val Ile Leu Val Thr Gly Lys Ser Ser Tyr Lys Lys
 3365 3370 3375
 Cys Gly Ala Trp Asp Val Val Lys Pro Ala Leu Glu Glu Tyr Gly Ile
 3380 3385 3390
 Glu Tyr Val His Tyr Asp Lys Val Gly Pro Asn Pro Thr Val Asp Met
 3395 3400 3405
 Ile Asp Glu Ala Thr Gln Leu Gly Lys Glu Phe Gly Ala Gln Ala Val
 3410 3415 3420

 Ile Gly Ile Gly Gly Gly Ser Pro Ile Asp Ser Ala Lys Ser Val Ala
 3425 3430 3435 3440
 Ile Leu Leu Glu Tyr Thr Asp Lys Thr Ala Arg Asp Leu Tyr Glu Leu
 3445 3450 3455
 Lys Phe Thr Pro Thr Lys Ala Lys Pro Ile Ile Ala Val Asn Thr Thr
 3460 3465 3470
 His Gly Thr Gly Thr Glu Val Asp Arg Phe Ala Val Ala Ser Ile Pro
 3475 3480 3485
 Glu Lys Glu Tyr Lys Pro Ala Ile Ala Tyr Asp Cys Ile Tyr Pro Leu

 3490 3495 3500
 Tyr Ser Ile Asp Asp Pro Ala Leu Met Thr Lys Leu Pro Ala Asp Gln
 3505 3510 3515 3520
 Thr Arg Tyr Val Thr Ile Asp Ala Leu Asn His Ile Thr Glu Ala Ala
 3525 3530 3535
 Thr Thr Lys Phe Ala Ser Pro Tyr Ser Ile Leu Leu Ala Gln Glu Thr
 3540 3545 3550
 Ala Arg Leu Ile Phe Asp Tyr Leu Pro Glu Ala Leu Ala His Pro Asp
 3555 3560 3565

 Asn Leu Gln Ala Arg Tyr Tyr Leu Leu Tyr Ala Ser Ala Ile Ala Gly
 3570 3575 3580

Ile Ser Phe Asp Asn Gly Leu Leu His Phe Thr His Ala Leu Glu His
 3585 3590 3595 3600

Pro Leu Ser Ala Val Lys Pro Asp Leu Pro His Gly Leu Gly Leu Ala
 3605 3610 3615

Met Leu Leu Pro Ala Val Ile Lys His Ile Tyr Pro Ala Thr Ala Arg
 3620 3625 3630

Ile Leu Ala Glu Val Tyr Arg Pro Leu Val Pro Glu Ala Lys Gly Val
 3635 3640 3645

Pro Gly Glu Ala Glu Leu Val Ala Lys Lys Val Glu Glu Trp Leu Phe
 3650 3655 3660

Asn Ile Gly Ile Thr Gln Lys Leu Ile Asp Val Gly Phe Thr Glu Glu
 3665 3670 3675 3680

Asp Val Asp Lys Leu Ala Glu Leu Ala Met Thr Thr Pro Ser Leu Asp
 3685 3690 3695

Leu Leu Leu Ser Leu Ala Pro Ile Glu Ala Thr Lys Glu Thr Val Ala
 3700 3705 3710

Ala Ile Tyr Arg Asp Ser Leu Tyr Pro Leu Asn Lys
 3715 3720

<210> 5

<211> 1658

<212> PRT

<213> Thermococcus onnurineus NA1

<400> 5

Met Glu Val Ile Phe Leu Phe Ile Val Ile Ile Leu Ser Val Ala Ser
 1 5 10 15

Phe Ile Gly Val Phe Ser Arg Ser Ala Ile Leu Thr Lys Leu Val Asn
 20 25 30

Ala Leu Ser Ala Leu Gly Ser Leu Thr Ile Ala Tyr Ala Gly Ile Val
 35 40 45

Gly Leu Lys Glu Ser Val Glu Leu Asn Ile Thr Leu Leu His Leu Lys
 50 55 60

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

Lys Pro Ile Glu Lys Glu Arg Glu His Lys Lys Asp Val Met Glu Ile
 1075 1080 1085
 Pro Leu Gly Pro Tyr His Pro Ser Leu His Glu Pro Glu Tyr Phe Glu
 1090 1095 1100
 Leu Tyr Val Lys Gly Asp Lys Val Val Asp Ala Glu Tyr Arg Gly Phe
 1105 1110 1115 1120

 His Ile His Arg Gly Met Glu Lys Leu Ala Glu Ser Arg Met Thr Ile
 1125 1130 1135
 Asn Gln Ile Pro Phe Leu Ala Glu Arg Ile Cys Gly Ile Cys Gly Cys
 1140 1145 1150
 Thr His Ser Ala Ala Tyr Cys Gln Ala Val Glu Asp Ala Ala Gly Ile
 1155 1160 1165
 Tyr Val Pro Glu Arg Ala Gln Tyr Ile Arg Thr Ile Met Leu Glu Val
 1170 1175 1180
 Glu Arg Ile His Ser His Leu Leu Trp Phe Gly Val Val Cys His Leu

 1185 1190 1195 1200
 Leu Gly Phe Asp Ser Gly Phe Met His Ile Trp Arg Ala Arg Glu Tyr
 1205 1210 1215
 Ile Met Asp Ile Ala Glu Leu Ile Thr Gly Asn Arg Lys Thr Tyr Gly
 1220 1225 1230
 Ile Asn Ile Val Gly Gly Val Arg Arg Asp Ile Thr Glu Asp Lys Lys
 1235 1240 1245
 Glu Lys Thr Leu Lys Leu Leu Asp Met Val Glu Lys Glu Ser Arg Glu
 1250 1255 1260

 Val Leu Asp Asn Ile Ala Glu Met Lys Glu Leu Arg Glu Arg Met Glu
 1265 1270 1275 1280
 Gly Val Gly Val Leu Pro Lys Lys Glu Ala Arg Glu Ile Gly Val Val
 1285 1290 1295
 Gly Pro Met Ala Arg Ser Ser Gly Ile Asp Thr Asp Val Arg Arg Asp
 1300 1305 1310
 His Pro Tyr Ala Ala Tyr Lys Asp Leu Asp Phe Lys Val Pro Val Tyr
 1315 1320 1325

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

1330 1335 1340

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

1345 1350 1355 1360

Gly Glu Leu Ile Asn Asp Glu Tyr Glu Ile Pro Pro Phe Lys Leu Gly

1365 1370 1375

Ile Gly Val Thr Glu Ala Pro Arg Gly Glu Asn Ile His Ala Val Ile

1380 1385 1390

Thr Trp Gly Glu Asn Met Ile Tyr Arg Trp His Pro Arg Ala Ala Thr

1395 1400 1405

Tyr Asn Asn Leu Pro Ala Val Pro Ile Met Leu Arg Gly Asn Asp Val

1410 1415 1420

Ala Asp Ala Pro Leu Ile Ile Ala Ser Ile Asp Pro Cys Phe Ser Cys

1425 1430 1435 1440

Thr Asp His Val Ser Ile Ile Asp Ser Glu Ser Gly Lys Ile Leu Trp

1445 1450 1455

Arg Gly Pro Leu Lys Glu Gly Val Arg Arg Val Met Val Lys Asn Ser

1460 1465 1470

Leu Trp Val Phe His Leu Asn Ser Gly Ser Cys Asn Gly Cys Asp Ile

1475 1480 1485

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

1490 1495 1500

Ile Lys Leu Val Gly Ser Pro Arg His Ala Asp Ala Ile Ala Phe Thr

1505 1510 1515 1520

Gly Pro Ile Thr Arg Glu Cys Leu Pro Lys Val Ile Asp Ala Leu Lys

1525 1530 1535

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

1540 1545 1550

Gly Gly Gly Ile Trp Tyr Asp Thr Tyr Ser Val Ile Gly Gly Val Lys

1555 1560 1565

Glu Leu Tyr Arg Ile Leu Lys Glu Glu Tyr Asn Met Glu Pro Pro Ala

1570 1575 1580
 Thr Val Phe Ile Pro Gly Cys Pro Pro Lys Pro Glu Ala Ile Ile Tyr
 1585 1590 1595 1600
 Gly Val Ala Val Ala Ser Gly Met Leu Glu Ser Lys Gln Lys Lys Thr
 1605 1610 1615
 Val Tyr Val Glu Pro Glu Glu Ser Val Ala Asn Glu Lys Leu Met Ile

 1620 1625 1630
 Ala Glu Leu Ile Ser Glu Thr Glu Lys Thr Arg His Phe Met Pro Gly
 1635 1640 1645
 Ile Val Ile Arg Gly Val Glu Asp Glu Pro
 1650 1655
 <210> 6
 <211> 1145
 <212> PRT
 <213> Thermococcus onnurineus NA1
 <400> 6
 Leu Ser Glu Ile Thr Leu Asn Lys Val Cys Arg Ile Ala Gly Glu Ala
 1 5 10 15
 Lys Leu Val Leu Tyr Glu Glu Asn Gly Thr Val Gln Asp Ala Leu Phe

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

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

Met Asp Phe His Ala Lys Ile Ser Asp Tyr Leu Leu Lys Met Gly Gly
 130 135 140

Ala Ala Thr His Pro Pro Asn Leu Thr Val Gly Gly Met Phe Ser Val
 145 150 155 160

Pro Lys Trp Ser Val Phe Asn Asn Leu Lys Ala Arg Leu Pro Lys Leu
 165 170 175

Thr Gly Gln Trp Glu Glu Ile Ala His Leu Leu Thr Asp Glu Asp Ile
 180 185 190

Gln Thr Glu Val Ala Asp Glu Leu Arg Glu Lys Lys Ala Glu Asn Asn
 195 200 205

Tyr Leu Val Ser Ser Leu Phe Tyr Gly Asp Arg Phe Asn Ile Asn Ala
 210 215 220

Glu Arg Ile Glu Thr Met Pro Tyr Tyr Glu Tyr Arg Lys Asp Asn Pro
 225 230 235 240

His Ser Lys Glu Ser Thr Thr Leu Ile Ala Phe Tyr Gly Gly Glu Lys
 245 250 255

Val Glu Ala Gly Pro Arg Ala Arg Met Lys Val Tyr Arg Glu Phe Thr
 260 265 270

Asp Ser Ser Leu Tyr Gly Leu His Thr Ala Arg Val Gln Asp Thr Thr
 275 280 285

Leu Ala Leu Ile Arg Leu Glu Glu Ile Leu Asp Ser Ile Lys Met Asp
 290 295 300

Glu Pro Phe Arg Thr Lys Asn Ile Val Phe Gly Pro Gly Lys Gly Val
 305 310 315 320

Gly Val Tyr Glu Ala Pro Arg Gly Thr Leu Ile His Leu Ile Glu Leu
 325 330 335

Gly Asp Glu Gly Arg Val Val Ser Ser Lys Ile Ile Val Pro Thr Met
 340 345 350

Phe Asn Ile Pro Val Met Glu Glu Met Ala Lys Gly Leu Ser Val Lys
 355 360 365

Ala Ala Glu Ala Val Met Arg Leu Tyr Asp Pro Cys Ile Pro Cys Thr

Leu Leu Gly Asn Val Phe Gly Ile Tyr Leu Ala Arg Ala Thr Asp Glu
 625 630 635 640
 Glu Ile Leu Lys Arg Lys Val Ala Ser Gly Gly Ala Val Thr Ala Leu
 645 650 655
 Leu Ala Tyr Ala Leu Glu Lys Gly Leu Ile Asp Gly Val Val Thr Ala
 660 665 670

 Lys Arg Thr Glu Gly Leu Glu Gly Gln Ala Val Val Ala Arg Thr Arg
 675 680 685
 Glu Glu Leu Leu Glu Thr Ala Gly Asn Lys Trp Ser Ile Val Pro Phe
 690 695 700
 Ala Ser Arg Met Lys Ala Lys Ile Glu Glu Glu Asp Leu Lys Asn Val
 705 710 715 720
 Ala Val Val Cys Leu Pro Cys Gln Ala Gln Phe Phe Gly Gln Met Arg
 725 730 735
 Asp Phe Pro Leu Leu Glu Ser Asp Phe Gly Glu Arg Ile Lys Tyr Ile

 740 745 750
 Val Ser Leu Phe Cys Ile Gly Thr Phe Ala Phe Glu Ala Phe Leu Asn
 755 760 765
 Tyr Leu Arg Met Lys His Gly Ile Met Ala Gln Asp Ile Lys Asp Ile
 770 775 780
 Val Leu Lys Gly Asp Phe Leu Glu Ile Tyr His Gly Asp Ser Val Leu
 785 790 795 800
 Ser Leu Pro Ile Lys Glu Val Tyr Ser Tyr Leu Gln Ala Gly Cys Leu
 805 810 815

 Val Cys Thr Asp Tyr Thr Gly Thr Trp Ser Asp Ile Ser Ala Gly Phe
 820 825 830
 Val Glu Ser Glu Arg Gly Trp Thr Val Leu Ile Thr Arg Asn Leu Lys
 835 840 845
 Ala Glu Glu Leu Val Lys Ser Ala Glu Lys Asp Gly Tyr Ile Glu Leu
 850 855 860
 Arg Asp Gly Ser His Val Met Gly Glu Val Leu Lys Ala Ala Arg Glu
 865 870 875 880

Lys Leu Ala Arg Ala Gln Lys Asn Met Met Tyr Leu Leu Leu Ile Lys

885 890 895

Lys Val Lys Ile Leu Lys Trp Gln Asp Gly Leu Val Pro Thr Glu Asp

900 905 910

Tyr Ile Cys Val Glu Glu Thr Phe Glu Ile Phe Ala Val His Glu Lys

915 920 925

Asp Glu Glu Phe Leu Ala Glu Leu Pro Ala Ser Pro Asn Gln Leu Lys

930 935 940

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

945 950 955 960

Asp Ile Val Asp Val Trp Val Glu Gly Lys Glu Ile Tyr Val Lys Leu

965 970 975

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

980 985 990

Gly Asp Pro Tyr Arg Met Lys Glu Gly Arg Ile Leu Ser Arg Lys Gly

995 1000 1005

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

1010 1015 1020

Met Thr Thr Leu Ala Glu Thr Trp Arg Lys Thr Gly Gly Thr His Trp

1025 1030 1035 1040

Ala Ala Leu Phe Asp Leu Asn Ala Asn Val Val Ala Phe Ser Glu Asp

1045 1050 1055

Ile Gly Arg His Asn Ala Val Asp Lys Val Val Gly Tyr Ala Val Leu

1060 1065 1070

Asn Gly Leu Asp Leu Glu Arg Leu Ile Leu Ala Ser Ser Gly Arg Met

1075 1080 1085

Pro Tyr Gly Met Val Arg Lys Ala Val Asn Ala Gly Ile Pro Val Val

1090 1095 1100

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

1105 1110 1115 1120

His Gly Val Thr Leu Ile Gly Phe Ala Arg Gly Arg Arg Phe Asn Val

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

Met Ala Val Leu Ser Met Leu Ala Ala Ile Phe Ser Tyr Ala Leu Gln
 690 695 700

Asp Asp Val Lys Arg Leu Phe Ala Trp Ser Thr Ile Asp Asn Met Gly
 705 710 715 720

Trp Met Tyr Leu Leu Ile Leu Ala Gly Leu Leu Gly Val Ser Gly Val
 725 730 735

Glu Lys Gly Val Asp Tyr Tyr Val Val Ala His Gly Leu Ala Lys Ala
 740 745 750

Ala Ala Phe Ile Ser Thr Gly Ala Leu Leu Tyr Val Phe Gly Thr Arg
 755 760 765

Ser Leu Lys Lys Ala Lys Gly Met Met Asn Thr Asp Ser Leu Thr Ala
 770 775 780

Gly Leu Met Met Ala Ser Ile Phe Ala Leu Glu Gly Val Pro Pro Phe
 785 790 795 800

Asn Leu Phe Met Asn Lys Leu Asn Val Ile Lys Thr Leu Leu Thr Val
 805 810 815

Ser Pro Ala Leu Ala Tyr Phe Thr Ala Leu Glu Trp Val Ile Ala Phe
 820 825 830

Ile Leu Phe Leu Arg Val Val His Ala Tyr Ile Leu Ser Glu Gly Glu
 835 840 845

Pro Glu Ala Lys Arg Lys Leu Ala Gly Ser Ile Ala Leu Ser Val Ile
 850 855 860

Val Leu Leu Ile Leu Ser Met Val Ser Gln Phe Val Cys Asp Tyr Ile
 865 870 875 880

Trp Val Arg Trp Met Glu Gly Leu Phe Thr Leu Ala Val Ile Leu Tyr
 885 890 895

Phe Leu Ser Ile Pro Ala Ala Leu Ala Leu Lys Arg Ser Phe Lys Ala
 900 905 910

Ser Ile Ser Ile Gly His Ile Leu Thr Ala Leu Ala Ser Ile Ala Leu
 915 920 925

Leu Ala Phe Thr Phe Val Ser Ile Pro Asp Ile Leu Ser Gly Lys Ala
 930 935 940

Ile Glu Phe Thr Tyr Asp Leu Gly Val Ala Gln Ile Pro Phe Gln Ile

945 950 955 960

Asp Gly Leu Ser Leu Ile Met Cys Phe Ile Phe Gly Ala Leu Gly Leu

 965 970 975

Ala Ala Ser Ile Tyr Ser Pro Arg Tyr Met Ala Ile Tyr Glu Lys Ser

 980 985 990

Gly Arg Gly Trp Met Tyr Ile Thr Ile Tyr Ser Val Phe Met Leu Ser

 995 1000 1005

Met Ile Leu Ile Val Thr Ile Ala Asn Met Phe Trp Phe Ile Phe Leu

1010 1015 1020

Trp Glu Val Met Thr Phe Thr Ser Tyr Leu Leu Thr Ile Trp Glu Ser

1025 1030 1035 1040

Asp Lys Glu Asp Val Arg Lys Ala Gly Trp Lys Tyr Phe Val Thr Met

 1045 1050 1055

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

 1060 1065 1070

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

 1075 1080 1085

Leu Lys Leu Ser Pro Val Phe Tyr Ala Leu Phe Leu Ile Gly Phe Gly

1090 1095 1100

Ser Lys Ser Gly Val Val Pro Leu His Phe Trp Ala Pro Glu Ala Tyr

1105 1110 1115 1120

Thr Val Ala Pro Ser Asn Val Ser Ala Leu Met Ala Gly Ala Leu Glu

 1125 1130 1135

Lys Val Ala Val Tyr Ala Leu Ile Arg Thr Thr Cys Phe Ile Met Lys

 1140 1145 1150

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

 1155 1160 1165

Leu Thr Val Gly Thr Leu Tyr Ala Leu Lys Gln Thr Asp Ala Lys Arg

 1170 1175 1180

Leu Leu Ala Tyr His Ser Ile Gly Gln Ile Gly Tyr Ile Trp Leu Gly

1185 1190 1195 1200
 Met Gly Val Gly Ile Val Phe Ile Ala Arg Gly Asp Met Tyr Ser Ala
 1205 1210 1215
 Phe Gly Ala Ile Ala Leu Ala Ser Ser Leu Tyr His Leu Val Asn His
 1220 1225 1230
 Thr Phe Phe Lys Gly Leu Leu Phe Leu Ser Thr Gly Ser Ile Phe Tyr
 1235 1240 1245
 Arg Thr Arg Ser Arg Asp Leu Asn Gln Leu Arg Gly Leu Ala Lys Leu
 1250 1255 1260
 Met Pro Phe Thr Ala Leu Phe Thr Phe Ile Ala Ala Met Ser Ile Ala
 1265 1270 1275 1280
 Gly Thr Pro Pro Phe Asn Gly Phe Met Ser Lys Trp Met Ile Tyr Gln
 1285 1290 1295
 Ser Thr Phe Leu Ser Gly Asn Gly Leu Ile Val Phe Phe Gly Val Met
 1300 1305 1310

 Ala Leu Phe Ile Ser Ala Ala Thr Leu Ala Ser Phe Ile Lys Phe Tyr
 1315 1320 1325
 Thr Thr Ala Phe Gly Gly Glu Pro Thr Glu Phe Thr Lys Asp Ala Glu
 1330 1335 1340
 Glu Val Pro Ser Pro Met Leu Ile Ala Lys Gly Phe Leu Ala Ser Leu
 1345 1350 1355 1360
 Cys Ile Leu Leu Gly Leu Val Pro Ser Leu Ile Leu Pro Ile Leu Leu
 1365 1370 1375
 Ser Pro Gly Ala Ala Leu Ala Gly Ile Asp Val Ser Gly Leu Met Asp
 1380 1385 1390
 Thr Asn Tyr Trp Leu Val Thr Ile Lys Ala Pro Leu Met Pro Thr Gly
 1395 1400 1405
 Ala Glu Ser Tyr Phe Lys Pro Leu Leu Phe Ala Thr Leu Phe Gly Val
 1410 1415 1420
 Ile Phe Leu Gly Met Tyr Leu Leu Phe Pro Ile Ser Lys Lys Thr Tyr
 1425 1430 1435 1440
 Arg Pro Trp Thr Leu Gly Glu Pro Val Ala Met Glu His Tyr Lys Phe

Tyr Ser Asn Ala Gly Ala Asn Arg Glu Met Met Leu Ile Leu Ser Val
 1700 1705 1710
 Glu Pro Ile Leu Gly Ile Ala Ile Gly Ile Leu Ala Leu Lys Thr His
 1715 1720 1725
 Ser Leu Ser Val Ser Gly Ile Pro Leu Asn Leu Ser Leu Thr Pro Ser
 1730 1735 1740

 Val Val Leu Ala Phe Ile Phe Leu Ala Tyr Ala Val Tyr Thr Glu Cys
 1745 1750 1755 1760
 Ala Phe Ile Pro Phe Asp Ile Ala Glu Ala Glu Thr Glu Ile Leu Glu
 1765 1770 1775
 Gly Pro Leu Val Glu Tyr Ser Gly Lys Leu Leu Gly Ile Phe Lys Trp
 1780 1785 1790
 Ala Met Leu Ile Lys Arg Val Ala Leu Ile Trp Leu Phe Ala Ser Phe
 1795 1800 1805
 Ile Val Ile Pro Val Met Lys Gly Phe Val Asp Ile Thr Thr Pro Tyr

 1810 1815 1820
 Gly Gly Ala Val Thr Leu Ala Ala Gln Leu Val Leu Leu Val Val Phe
 1825 1830 1835 1840
 Tyr Val Met Ser Ala Ile Ile Glu Ser Thr Thr Ala Arg Met Lys Val
 1845 1850 1855
 Ile Gln Ala Ile Arg Gln Asn Thr Val Ile Phe Leu Ala Gly Ile Val
 1860 1865 1870
 Ala Leu Val Ile Ala Ser Leu Gly Trp Met Ser Glu Val Ile Lys Phe
 1875 1880 1885

 Asn Glu Ala Leu Lys Lys Lys Arg Val His Arg Gly Asp Glu Lys Ala
 1890 1895 1900
 Lys Val Thr Arg Glu Tyr Leu Asp Glu Ile Ile Glu Lys Phe Gly Glu
 1905 1910 1915 1920
 Lys Ile Arg Asp Val Lys Gln Ala Ala Tyr Asn Gln Trp Ile Ile Thr
 1925 1930 1935
 Val Glu Arg Glu Asp Leu Pro Glu Ile Val Leu Tyr Phe Leu Asn His
 1940 1945 1950

Pro Glu Trp Lys Glu Thr Gln Leu Ser Ser Met Val Ala Thr Asp Glu

1955 1960 1965

Arg Pro Leu Asn Gly Lys Phe Ser Ile Thr Tyr Trp Leu Ser Val Asn

1970 1975 1980

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

1985 1990 1995 2000

Asp Asp Pro Arg Phe Thr Ser Ile Ala Ala Lys His Arg Gly Ala Asn

2005 2010 2015

Trp Tyr Glu Arg Glu Ala Met Glu Met Leu Gly Leu Thr Ala Glu Gly

2020 2025 2030

His Pro Asp Pro Arg Arg Leu Val Leu Pro Asp Asp Trp Pro Ser Cys

2035 2040 2045

Val Tyr Pro Leu Arg Lys Asp Phe His Tyr Ser Asn Ser Pro Pro Gly

2050 2055 2060

Glu Lys Phe Tyr Pro Tyr Lys Glu Pro Lys Lys Asp Glu Ile Val Val

2065 2070 2075 2080

Pro Tyr Gly Pro Tyr His Val Ala Leu Glu Glu Ala Ala His Phe Arg

2085 2090 2095

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

2100 2105 2110

Tyr Ala His Arg Gly Ile Glu Lys Ile Ser Glu Gly Arg Leu Thr Tyr

2115 2120 2125

Asp Gln Val Cys Phe Ile Ala Glu Arg Ile Cys Gly Ile Cys Gly Cys

2130 2135 2140

Thr His Ser Thr Ala Tyr Cys Gln Ala Val Glu Asn Ala Gly Gly Ile

2145 2150 2155 2160

Glu Val Pro Glu Arg Ala Glu Tyr Ile Arg Thr Ile Val Leu Glu Ile

2165 2170 2175

Glu Arg Leu His Ser His Leu Leu Asn Phe Gly Ile Val Ser His Leu

2180 2185 2190

Val Gly Tyr Asp Tyr Gly Phe Met Lys Ala Trp Arg Ile Arg Glu His

2195 2200 2205
 Val Met Trp Leu Ala Glu Arg Leu Thr Gly Asn Arg Lys Thr Tyr Gly
 2210 2215 2220
 Met Leu Leu Val Gly Gly Val Arg Arg Asp Leu Leu Glu Tyr Arg Lys
 2225 2230 2235 2240
 Ser Leu Ile Glu Asp Val Leu Lys Lys Ile Lys Thr Glu Phe Ser Glu

 2245 2250 2255
 Leu Val Asp Glu Ala Ile Ser Thr Ser Thr Phe Val Lys Arg Leu Glu
 2260 2265 2270
 Gly Val Gly Val Leu Pro Tyr Lys Val Ala Lys Glu Trp Asp Val Asp
 2275 2280 2285
 Gly Pro Leu Gly Arg Gly Ser Gly Arg Asp Phe Asp Val Arg Arg Asp
 2290 2295 2300
 His Pro Tyr Ala Ala Tyr Lys Tyr Leu Asp Phe Lys Val Pro Val Tyr
 2305 2310 2315 2320

 Lys Glu Gly Asp Val Leu Ala Arg Ala Leu Val Arg Ile Glu Glu Val
 2325 2330 2335
 Phe Glu Ser Ile Trp Ile Ile Glu Gln Ala Leu Asp Gln Met Pro Gly
 2340 2345 2350
 Gly Asp Ile Leu Ala Glu Tyr Lys Glu Ile Pro Pro Tyr Ser Glu Ala
 2355 2360 2365
 Ile Gly Met Thr Glu Ala Pro Arg Gly Glu Asn Ile His Tyr Val Met
 2370 2375 2380
 Thr Gly Glu Asn Asn Lys Val Tyr Arg Tyr Arg Ala Arg Ala Ala Thr

 2385 2390 2395 2400
 Tyr Asn Asn Leu Pro Ala Val Pro Asp Met Met Arg Gly Tyr Thr Ile
 2405 2410 2415
 Ala Asp Ala Pro Leu Ile Val Ala Ser Ile Asp Pro Cys Tyr Ser Cys
 2420 2425 2430
 Thr Glu Arg Val Gln Val Val Asp Val Glu Ser Gly Lys Val Arg Val
 2435 2440 2445
 Leu Ser Glu Thr Glu Phe Asn Lys Leu Ser Ile Lys Ala Ser Arg Arg

2450 2455 2460
 Val Met Ala Val Thr Leu Lys Tyr Pro Phe Val Lys Leu Glu Ala Pro
 2465 2470 2475 2480
 Pro Glu Tyr Arg Gly Ile Pro Gln Ile Asp Ala Thr Leu Cys Ile Gly
 2485 2490 2495
 Cys Gly Ala Cys Val Asn Ala Cys Pro Pro Asp Ala Leu Leu Arg Ile
 2500 2505 2510
 Asp Asp Tyr Asn Arg Gly Val Arg Glu Ile Val Leu Asp Val Gly Arg
 2515 2520 2525
 Cys Ile Arg Cys Ala Arg Cys Glu Glu Val Cys Pro Thr Gly Ala Ile

 2530 2535 2540
 Lys Leu Thr Asn Leu Phe Glu Ala Ala Ser Pro Asp Arg Met Asp His
 2545 2550 2555 2560
 Val Glu Val Val Arg Leu Arg Leu Val Lys Cys Lys Asn Cys Gly Arg
 2565 2570 2575
 Tyr Ala Asp Phe Thr Glu Arg Gln Val Arg Lys Ala Leu Gln Ile Leu
 2580 2585 2590
 Pro Glu Glu Ile Ile Glu Arg Glu Ala Leu Glu Glu Lys Val Trp Ile
 2595 2600 2605

 Cys Arg Asp Cys Arg Arg Lys Gly Thr Val Asp Gly Thr Ile Glu Ala
 2610 2615 2620
 Ser Lys Glu Val Val Leu Met Ser Gly Lys Pro Lys Leu Arg Ser Ile
 2625 2630 2635 2640
 Trp Val Phe His Leu Asn Thr Gly Ser Cys Asn Gly Cys Asp Ile Glu
 2645 2650 2655
 Ile Ile Asp Val Leu Thr Pro Phe Tyr Asp Val Glu Arg Phe Gly Ile
 2660 2665 2670
 Lys Leu Val Gly Ser Pro Arg His Ala His Ala Leu Leu Val Ser Gly

 2675 2680 2685
 Pro Leu Thr Arg Gln Ala Tyr Tyr Gly Ala Lys Glu Thr Ile Lys Ala
 2690 2695 2700

Met Pro Pro Glu Pro Arg Val Ile Val Ala Ile Gly Thr Cys Thr Cys
 2705 2710 2715 2720
 Ser Gly Gly Ile Phe Tyr Asn Gly Tyr Pro Val Tyr Arg Arg Pro Glu
 2725 2730 2735
 Ser Gly Arg Glu Gly Ser Glu Tyr Pro Arg Arg Gly Gly Ile Ala Glu
 2740 2745 2750

Leu Ile Ala Asp Leu Arg Asp Glu Gly Glu Lys Val Gly Pro Val Ile
 2755 2760 2765

Tyr Ile Pro Gly Cys Pro Pro Arg Pro Glu Glu Ile Ile Tyr Gly Ile
 2770 2775 2780

Ala Gln Leu Val Gly Leu Val Glu Lys Lys Leu Ser Tyr Gln Glu Tyr
 2785 2790 2795 2800

Ser Asp Glu Leu Val Pro Phe Lys Leu Pro Glu Gly Pro Leu Glu Glu
 2805 2810 2815

Arg Ile Arg Leu Thr Leu Met Glu Arg Leu Arg His Leu Val Gly Tyr
 2820 2825 2830

Leu Asp Arg Glu Lys Ile Leu Glu Asp Phe Met Gly Leu Val Lys Glu
 2835 2840 2845

Ala Glu Lys Ser Glu Asn Pro Arg Glu Glu Leu Ala Arg Leu Val Lys
 2850 2855 2860

Asp Tyr Ala Ala Lys Cys Gly Asp Val Arg Leu Gly Phe Cys Met Met
 2865 2870 2875 2880

Leu Leu Glu Arg Glu Tyr Trp Arg Val Lys Asp Ala Leu Asp Ala Gly
 2885 2890 2895

Lys Glu Phe Val Tyr Trp Val
 2900

<210> 8

<211> 1387

<212> PRT

<213> Thermococcus onnurineus NA1

<400> 8

Met Phe Gly Tyr Trp Asp Ala Leu Tyr Phe Val Phe Ile Phe Ile Ile

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

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

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

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

Pro Gly Cys Pro Pro Arg Pro Glu Ala Ile Leu Tyr Gly Val Val Leu
 260 265 270
 Ala Leu Glu Lys Leu Ala Lys Ile Leu Lys Gly Glu Val Pro Glu Gly
 275 280 285
 Glu Glu Met Ala Asp Asp Asn Arg Ile Met Glu Asn Val Asp Asn Val
 290 295 300
 Arg Glu Pro Thr Lys Glu Asp Thr Val Ala Glu Thr Ile Lys Ser Arg
 305 310 315 320
 Phe Pro Asn Ala His Val Glu Ile Arg Glu Asn Lys Trp Gly Arg Lys
 325 330 335

 Arg Val Trp Val Ile Val Pro Arg Glu Asp Tyr Lys Ala Leu Met Lys
 340 345 350
 Phe Leu Leu Glu Leu Asp Pro Glu Ala His Tyr Ser Ile Gly Ile Glu
 355 360 365
 Gln Asp Tyr Gly Glu Glu Ile Gly Tyr Met Ser His Ile Leu Leu His
 370 375 380
 Tyr Asp Asn Ala Pro Ala Val Ser Leu Leu Val Asp Val Arg Val Pro
 385 390 395 400
 Lys Asp Asp Pro Val Ile Pro Asp Ile Ser Asp Ile Phe Pro Ile Ala
 405 410 415
 Leu Gln Tyr Glu Arg Glu Ala Ala Glu Met Met Gly Ile Val Phe Glu
 420 425 430
 Gly Ile Pro Asp Ser Arg Arg Leu Phe Leu Pro Asp Asp Phe Pro Glu
 435 440 445
 Gly Ile Tyr Pro Leu Arg Leu Asp Glu Lys Gly Ile Pro Glu Glu Ile
 450 455 460
 Val Lys Asn Ala Gly His Pro Tyr Tyr Leu Lys Gly Gly Asp Lys Met
 465 470 475 480

 Thr Lys Lys Val Glu Tyr Trp Ile Lys Ile Pro Phe Gly Pro Ile His
 485 490 495
 Pro Gly Leu Glu Glu Pro Glu Lys Phe Ile Leu Thr Leu Asp Gly Glu

Val Ile Val Phe Ile Tyr Leu Leu Thr Leu Ile Gly Phe Phe Lys Ile
 1010 1015 1020

Leu Gly Gly Ile Ser Ser Gly Ser Pro Tyr Ala Lys Ile Gly Ala Ala
 1025 1030 1035 1040

Arg Glu Ala Ala Ile Met Val Ser Arg Glu Pro Ala Met Met Leu Ala
 1045 1050 1055

Leu Phe Ala Ile Ile Trp Arg Leu Gly Lys Leu Gly Val Asn Lys Pro
 1060 1065 1070

Phe Ser Met Glu Val Phe Tyr Gln Tyr Asn Ile Trp Glu Ile Gly Thr
 1075 1080 1085

Pro Leu Ser Leu Ile Gly Ala Val Ile Leu Leu Tyr Val Phe Val Ile
 1090 1095 1100

Trp Leu Ala Ser Glu Ile Glu Val Gly Tyr Phe Asn Ile Pro Asp Ala
 1105 1110 1115 1120

Glu Glu Glu Ile Ala Glu Gly Leu Leu Ala Glu Tyr Ser Gly Arg Tyr
 1125 1130 1135

Leu Ala Leu Leu Lys Leu Thr Lys Ala Leu Lys Thr Tyr Ile Ala Ala
 1140 1145 1150

Ser Leu Val Val Ala Ile Phe Phe Pro Trp Gly Ile Ala Asp Tyr Phe
 1155 1160 1165

Asn Leu Thr Gly Leu Pro Ala Asn Val Val Asn Leu Leu Phe His Thr
 1170 1175 1180

Leu Lys Val Phe Ile Leu Leu Phe Ala Val Gln Ser Val Phe Arg Ala
 1185 1190 1195 1200

Thr Thr Gly Arg Leu Lys Ile Thr Gln Ala Val Asp Phe Leu Trp Lys
 1205 1210 1215

Asn Val Phe Leu Ala Ser Leu Ile Gly Thr Leu Leu Ile Ala Met Glu
 1220 1225 1230

Val Ile Met Val Arg Leu Ser Pro Leu Ile Pro Thr Val Leu Arg Asn
 1235 1240 1245

Met Phe Lys Lys Pro Ala Thr Asn Leu Phe Pro Ala Thr Glu Pro Val
 1250 1255 1260

Pro Val Pro Asp Asn Phe Arg Gly Gln Leu Lys Tyr Asn Val Asp Lys

1265 1270 1275 1280

Cys Val Gly Cys Arg Met Cys Val Thr Val Cys Pro Ala Gly Val Phe

 1285 1290 1295

Val Phe Leu Pro Glu Ile Arg Lys Val Ala Leu Trp Thr Ala Arg Cys

 1300 1305 1310

Val Tyr Cys Ser Gln Cys Val Asp Val Cys Pro Thr Ala Ala Leu Gln

 1315 1320 1325

Met Ser Asp Glu Phe Leu Leu Ala Ser Tyr Asn Asn Tyr Asp Asp Lys

 1330 1335 1340

Phe Ile Pro Leu Lys Pro Glu Lys Val Glu Glu Ile Lys Lys Lys Leu

1345 1350 1355 1360

Glu Glu Gln Lys Lys Ala Lys Ala Ala Ala Ala Ala Lys Lys Ala Met

 1365 1370 1375

Glu Lys Lys Glu Ala Gly Lys Glu Ala Lys Lys

 1380 1385

<210> 9

<211> 854

<212> PRT

<213> Thermococcus onnurineus NA1

<400> 9

Met Gly Met Ala Glu Lys Arg Ile Ser Val Val Cys Pro Trp Cys Ser

 1 5 10 15

Val Gly Cys Arg Phe Tyr Ile Val Asn Val Asn Gly Tyr Pro Lys Lys

 20 25 30

Ile Glu Phe Asp Tyr Asp His Asp Ile Arg Asn His Gly Lys Leu Cys

 35 40 45

Pro Lys Gly Val Ala Ala Phe Gln His Leu Arg His Pro Asp Arg Leu

 50 55 60

Lys Lys Pro Leu Lys Arg Val Gly Glu Arg Gly Glu Gly Lys Phe Lys

 65 70 75 80

Pro Gly Ala His Asn Gly Met Gly Ala Gly Leu Met Gly Ile Gly Pro
 340 345 350
 His Glu Leu Pro Gly Arg Phe Pro Leu His Ala Glu Glu His Lys Arg
 355 360 365

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

 435 440 445
 Ala Ala Trp Phe Glu Lys Asp Gly Thr Ala Ile Ser Phe Glu Arg Arg
 450 455 460
 Val Gln Arg Ser Phe Lys Ala Ala Asp Ala Pro Gly Glu Ala Lys Pro
 465 470 475 480
 Asp Trp Glu Ile Leu Val Met Leu Ala Lys Glu Leu Gly Phe Gly Glu
 485 490 495
 Tyr Phe Asn Tyr Ser Asp Ala Asp Asp Ile Leu Arg Glu Ile Asn Arg
 500 505 510

 Ile Ile Pro Pro Leu Ala Gly Ala Thr Pro Glu Arg Leu Lys Lys Asn
 515 520 525
 Leu Lys Gly Cys Met Ile Pro Cys Pro Asp Glu Asn Thr Glu Val Pro
 530 535 540
 Arg Leu Phe Val Gln Gly Phe Leu Thr Pro Asn Gly Lys Ala Gln Leu
 545 550 555 560
 Ile Pro Val Glu Tyr Lys Glu Pro Gly Glu Val Pro Asp Glu Glu Tyr
 565 570 575
 Pro Phe Trp Leu Thr Asn Tyr Arg Phe Val Gly His Phe His Thr Gly

Arg Glu Ile Cys Leu Ser Met His Arg Thr His Ile Gly Cys Asp Ala
 195 200 205
 Asp Pro Ala Ser Leu Leu Leu His Gly Val Arg Thr Ala Leu Ala Asp
 210 215 220
 Gly Trp Cys Gly Ser Met Met Ala Thr Tyr Leu Ser Asp Ile Leu Phe
 225 230 235 240
 Gly Thr Pro Lys Pro Ile Lys Ser Leu Ala Asn Leu Gly Val Leu Lys
 245 250 255
 Glu Asp Met Val Asn Ile Ile Val His Gly His Asn Pro Ile Leu Ser
 260 265 270
 Met Lys Ile Ala Glu Ile Ala Gln Ser Glu Glu Met Gln Lys Leu Ala
 275 280 285

 Glu Gln Tyr Gly Ala Lys Gly Ile Asn Val Ala Gly Met Cys Cys Thr
 290 295 300
 Gly Asn Glu Val Leu Ser Arg Met Gly Val Gln Val Ala Gly Asn Phe
 305 310 315 320
 Leu Met Gln Glu Leu Ala Ile Ile Thr Gly Ala Val Glu Ala Val Ile
 325 330 335
 Val Asp Tyr Gln Cys Leu Met Pro Ser Leu Val Asp Val Ala Ser Cys
 340 345 350
 Tyr His Thr Lys Ile Ile Thr Thr Glu Pro Lys Ala Arg Ile Pro Gly
 355 360 365
 Ala Ile His Val Glu Phe Glu Pro Glu Lys Ala Asp Glu Ile Ala Lys
 370 375 380
 Glu Ile Ile Lys Ile Ala Ile Glu Asn Tyr Lys Asn Arg Val Pro Ala
 385 390 395 400
 Lys Val Tyr Ile Pro Glu His Lys Met Glu Leu Val Ala Gly Phe Ser
 405 410 415
 Val Glu Ala Ile Leu Glu Ala Leu Gly Gly Thr Leu Glu Pro Leu Ile
 420 425 430

 Lys Ala Leu Gln Asp Gly Thr Ile Lys Gly Ile Val Gly Ile Val Gly

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

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

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

 Ile Asn Gly Leu Ile His Val Ile Ile Lys Glu Arg Leu Tyr Asp Glu
 225 230 235 240
 Lys Phe Val Arg Glu Arg Cys Val Gly Phe Asp Glu Val Val Ala Ala
 245 250 255
 Val Glu Lys Phe Thr Pro Glu Phe Val Glu Lys Val Thr Gly Val Pro
 260 265 270

Ala Glu Leu Ile Ile Glu Ala Ala Arg Thr Phe Ala Thr Ala Gly Lys
 275 280 285

Gly Val Ile Thr Trp Ala Met Gly Leu Thr Gln His Thr His Gly Thr
 290 295 300

Glu Asn Val Lys Leu Leu Gly Thr Leu Ala Ala Ile Cys Gly Tyr Gln
 305 310 315 320

Gly Lys Glu Gly Ala Gly Cys Ser Pro Met Arg Gly Gln Asn Asn Val
 325 330 335

Gln Gly Ala Cys Asp Met Ala Ala Leu Pro Asn Val Phe Pro Gly Tyr
 340 345 350

Gln Ala Val Thr Asp Pro Glu Lys Arg Lys Phe Phe Glu Glu Phe Trp
 355 360 365

Gly Val Glu Leu Ser Gly Glu Val Gly Leu Thr Thr Val Glu Ala Ala
 370 375 380

Tyr Ala Ala Asp Lys Gly Lys Val Lys Ala Tyr Tyr Val Met Gly Glu
 385 390 395 400

Asn Pro Val Ile Ser Glu Ala Asn Ala Asn His Val Met His Thr Leu
 405 410 415

Glu Lys Leu Glu Phe Met Val Val Gln Asp Ile Val Pro Thr Pro Thr
 420 425 430

Met Glu Tyr Ala Asp Ile Val Leu Pro Ala Ala Ala Met Leu Glu Asn
 435 440 445

Glu Gly Ser Leu Thr Asn Thr Glu Arg Arg Val Gln Trp Ser Phe Gln
 450 455 460

Ala Val Lys Pro Pro Gly Glu Ala Arg Pro Asp Trp Trp Ile Leu Ser
 465 470 475 480

Glu Val Gly Lys Ala Ile Gly Phe Asp Lys Thr Gly Ser Gly Gly Phe
 485 490 495

Val Tyr Asn Asp Ala Ala Asp Val Leu Arg Glu Ile Asn Ala Cys Thr
 500 505 510

Pro Gln Tyr Arg Gly Ile Thr Pro Glu Arg Leu Lys Glu Asn Leu Ala

770	775	780	
Leu Leu Ala Pro Gln Lys Cys Ile Gly Cys Leu Met Cys Gly Ile Val			
785	790	795	800
Cys Pro Phe Gly Ile Pro Glu Leu Asp Leu Ile Asn Lys Ile Met Gly			
	805	810	815
Lys Cys Asp Leu Cys Ala His Arg Arg Ala Glu Gly Lys Leu Pro Ala			
	820	825	830
Cys Val Glu Thr Cys Pro Thr Asp Ala Leu Ile Tyr Gly Asp Phe Asn			
	835	840	845
Glu Ile Val Lys Lys Arg Arg Glu Lys Phe Thr Glu Lys Thr Ile Glu			
	850	855	860
Leu Ala Lys Thr Ala Glu Arg Ile Pro Leu Thr Gly Val			
865	870	875	
<210>	12		
<211>	4194		
<212>	DNA		
<213>	Thermococcus onnurineus NA1		
<400>	12		
atgggattcc taagcagact gttcgggtggg aagaagggaaa ccgacactga ggaaattcag			60
atagtctcga ggaagccagt tggaaagtgc cacgttgaga aggtcttcca catcatggga			120
cgcgaaacc tcgtggggac cgtggagaga ggagtaatct atccaggcta caaggtaag			180
ggaaagaaag ccgcagtgat ctacaggata gagaaaggta gaaaggctgt ggattttgtc			240
gtcgacggag acaaggctgc actcattctc gagggaatta ctaaagcgga agagggagat			300
acccttgaag tctatcagtc gtgaatgata atcgagttga gggagttcac gagagttgaa			360
ggcaacggca aagccgagat agttgtcgag aacggtgaag ttaaggacgt caggctcaaa			420
atcgttgaag ggccgcgctt cttcgagctg ctgactttgg gaaggcatta ctatgacgtt			480
ccagacttgg aagcggagat atgcgccata tgctacctct cacacagcgt cgctcggtt			540
ctgggcatag agaaggcctt cggagttgaa gtttcggagg agatccagct cctaagagag			600
ctcggcctca tcggtgaatt actggagagc cagcactcc acctgtacct gctcgttgct			660
ccagacgttt tcggttatcc tgacgcgata agaatggcta cgaagcacgg ggagctcgtc			720
aaggaggggc ttgcactgaa ggccttcggc aacagtataa gggaactcat tggaggaagg			780

gagatccacg gcataaacgt taaacccggt ggatttggca ggtatccgac ggttgaggaa 840
cttgaaaaca tcgagagga gagtggggcc ctcttagac tcgcaaggag agcggtagg 900
ctattcgctt cgtttagacc ctacggcgaa aaggcgggac acttcgtcgc gacggacggc 960
tacctatggg gcgacaagct gatttccgat gaggatggct ctttccacta caccgagaga 1020
atagaggaac gtcactggt ttacagcttc gcaaagcaga gccgctataa gggtagagcc 1080
ttcttcgctt gcgcactgcc gaggtcctg ctcaaagcag agatgctgac acccacagcg 1140

aagaggctct tcgaggagca cagggaaaag ctgccaccg gttacgtcag ctacaacaac 1200
ctcgctcagg ccatagagct cgtctacgcg cttgaaaggg cgggagagat agcaaagaaa 1260
ctcctcgaca agggcataaa gggtagaaaac gttcccgttg aagtcaaaga aggcgagggc 1320
atagggtacg tggaaagcgc taggggtgtc ctaatacatc actaccgcat tgactctggg 1380
ggcaaaatcg cctactcgaa catcataacg cccacggctt taaaccacgc tatgatggag 1440
gcaagcctgt tcaaggaagc gagaaaactc tacggagaga cggacgagac ggtactcgtc 1500
cagaggctgg aggaaacggt tagagccttt gatccgtgca tttctgttc agtgcacatc 1560

gtgaagcttt agatgatgga caagetcaag ttggccgtct tcgagcttac cgactcgggc 1620
ggctgtgcbc tgaatattct cttcctctac gagaagctgt ttgacctgct cgagtctac 1680
gagataacgg agttccacat ggcgaccagc ctaagcgagg ggagccacta cgactggcc 1740
ctcgtaaccg gaacggtctc aagccagcgc gacctagcgc tcctcaagga ggcaagaaac 1800
cactccgact acctcatagc cctcggaacc tgcgcaacgc acggctcggg tcaggctagc 1860
gtcgagggga gcataagggga gaagctgaag agggctctatg gagatgaggg caaccgatg 1920
agggcgctgg actcgaagcc cgtcgttag tacgtcgccg ttgatttcgc cctcccaggc 1980

tgtccctacg acaaaaacga ggtatatcag gttctgatgg acattgcaa aggcattgag 2040
ccggtaaaga aagactacc cgtctgcgtc gagtgcaagc tcaacgaata cgaatgtgtt 2100
ctcgtgaaga agggcctccc ctgcctcggc ccaataacct acggcggctg caacgctgct 2160
tgtatacgtc cgggctggg atgcataggc tgtcgccggc cgttccccgg cgaggtgaat 2220
cctgcaagtg agtacgagat actcaaggat ctgggctacg atgatgacta catcctcagg 2280
aagttcaaga ccttcgcgag gtgggagcca tgaatgagcg agaatccaca tcaaacttac 2340
gatgcgcgca ttctggaagt gaaggaccta acaccaggg agaagctctt cacgctccgc 2400

ttcttagacc cggaaattgg cgaacacttc acattcaagc cggccagtt cgtcatcgtc 2460
gatatacggg gcttcggtga gttcccata agcctctgct cctaccaac gagaaaagga 2520
tacattcagc tctgcatcag aaaagccgga aggatgacca agttcatcca tcagatgaaa 2580
gagggagaag tggtaggcat ccgcccggcc tacggcaacg gcttccgat ggagaaaatg 2640

gagggtcga atctactcct ggtcgccggt ggactcggta tggcaccct cgcctcggtt 2700
ctctggtacg cgatagacac cggaaagtac gaggcagctt ggctcctcta cggcaccaaa 2760
gcctacgagg acatactctt ccgacgacgag ataatccacc tgctgaagca cggcagacgcg 2820

gttggctgca gcgtaaagct cgcctatgag gtcgaaagcc cctcgtgcat ctacctgag 2880
cggggcttct tcgacagggt gtgcaagggt gtcggtaccg acctcttccg cggggaggag 2940
ttcgacgtcg acaaggctta cgccctcctc tgtgggcccgc cggttatgta ccgcttcgtc 3000
atcaaggagc tcctagacag gaaactctcg ccgggcagga tatacatgac cctcagaggg 3060
cgcatgcgct gcggaatagg caagtgcggc cactgtatag tgggaacgag cacctccata 3120
aagtacgtct gcaaggacgg ccccgctctc acatactggg atgctctctc cacgaggggg 3180
ttgatatgat tgagatatgt aaaactatca tctgagaact ttagctcatt tttgaaatc 3240

ctaaggaatt ggggcaaagt ctacgctccc atcaaaagag gaagcattta cacattccaa 3300
gaagttcacg agctaggaga gatagaactc aactatacaa ggacaatgct acctccaaaa 3360
aagttcttcg tgaggccaag ggacgaaatc cttcgctga agaacggtcg ctgggaaaat 3420
ggaaccgacg cagagccgat agttctcttc ggcctccact cctgcgatat gcacgggctc 3480
aagattctcg ataaggtcta tctcgacgag cccgccgacc cgtactaaa ggccgagcgc 3540
gagaaaacct tcatagttgg gataagctgc atgcccgacg agtattgctt ctgcaagagt 3600
ctcggcacgc accttgccat ggacggcttt gacctatcc tgcacgagct tcccagcga 3660

tggctcgtca ggataggaag tgtgagggga cagcaggtag tctgggagaa cggtagctc 3720
ttcgaggagg tgaccgacga ggacttgaag cacttcaagg agttcgagga gaggcgcgca 3780
aatgcgttcc agaaggagat cccgcaggaa ggactcgcag acatgctcga tttggcctac 3840
aacagcccgg tctggaagga gtaccccagc atatgcctgg cctgcggcaa ctgcaacatg 3900
gtctgtccta cctgcccgtg ctacaggtc tgcgataact ggatcagcgc ctacgacgcc 3960
gtcagagaga gacgctacga ctctgcttt atggagaacc acggactggt tgccggaggc 4020
cacaacttea ggccaactag actcgacaga ttcaggcaca gatactactg caagagctac 4080

ttcgatccct cctcgggtta taactcgtga ggttgtgaa ggtgtgacga gttctgcccg 4140
gcgaagatag agcacgtcaa ggttcttgag gaggtcaggg ggtcgtgag atga 4194

<210> 13
<211> 7506
<212> DNA
<213> Thermococcus onnurineus NA1
<400> 13

atgaacgcct ctcccttcat tatactttt ttgatccccc tgctcctcgg tccactcctg 60
ttcaaattag acggtagaaa ggccgatgta ttcatgctca tcgccgttgt gtcttccttc 120
ctggctaate ttgtgggagt cctcgaatac ctgaaagtgc gtggcgctca tcatatcgtt 180
tacctcgaaa ctcttcctc cggtgaggtc tacggcgta taatcgacce aatgagcgtt 240

ttggtcggtt ttgtcgtgag cttggctggc gtgctgttcc ttctctacgc ggtggactac 300
atgagcgaga gaaacaagca gcaccccgtc tactctgata agggcaggtt ctatgcttgg 360
atggtcatct tcgttggagc tacgctggca ttcatatact cctccacgac gcttcagctg 420
ctcatattct tcgagataat gggactcgcc tgctggggtg tcgttgggta ttataagggc 480
ccaaaagccg agagggtgc atacaaggcc ctgcttgtgc cgaacttcgg tgccatggtg 540
ggcctctaca ccaccgttgg cattggcatc ctcaagctcc atgattttag catctatgcg 600
ctccagaacc tgaatgatga gctcaagctt ctgctgttcc ttggcgtaat ggttgcggcc 660

ttaccaaga gcgcccagtt cccgctctat tcatggcttc cggatgcaat ggcggcgccc 720
acacctgctt ccgcttttct ccacgggtct gcaatggttg aaatgggcgt ttacctgctc 780
gccagggtea cccagttcat gcaaccgatt ccggagacag ctttctacgt tatgctcgtc 840
ttcgtgtcgc taactttact catagcaatt ctctactacc cgctccagaa ggacccaag 900
agactccttg cttattcaac catagcagag gcaggagtga tgtacgttgg cgtgctctat 960
gccgtgcttg gcctctcta tggctcctcag gcggccatgt tccagctggc taaccacgt 1020
ttcgtcaagg gtcttgcctt cctcaccgcg ggaaccttca gttacgcttt tggaacgctc 1080

gacatggaga agattagggg cctcggaaaag ctcgctccgg tcgttgggtc aagctggttc 1140
ttagcccttc tcggcctggc tggagtctct ccgctcggcc tgttcttcag caaggcgtat 1200
ctcttcatga acgcgtcttc aataaccagc tgggttggct ggattccgct ctctctagt 1260
ttggccgatg ccacggtttt ccttgcggta tctctcggat ggattaagag gatggtatc 1320
agcgagcccc tccaggagag tgcagaagt tccccgctga tgcgctttgt cctcgtagtc 1380
ctaatagtcc tgtccatcgt tgcgccgttc ctaagcgtga agctcgtgac tcagataggg 1440
ttcatggggt gaatgatgga aattecaatc gcgctctact cactctcagc gatttccggc 1500

ctgattggag actttaagcg gagcattaag atttcaagcg tcctctcagc catagcatcc 1560
ctatcccttc tgggcatagc tgcccacgcc ctgtccaggg ggcttcccgt tcaggagagc 1620
tttttgggca ttcccctaat catagacagc ctctccctcc cgcttctgtt catcatagcc 1680
ctgctcagcc ttgtggtttc agtgtattcc atttcatata tggaagtcca cagagatacc 1740
ggaagaccac tggcgtacac catctctac ggcacgttcg tgctgtcgat tgtattcgtg 1800
gctctgacgt caaacctgct ctggttcgtc ttcttctggg agctgatgac cctaacttcc 1860

ttcatcttcg tgagctggag ggagcaggac gctggaatta aatacctcct cacgatgcag 1920

 ctgccaaca cggigccct ctctgtggcc ctggcataa tctactccgc cactggaage 1980
 ttcagcgttg attacgccac gcttagggag gttgcatctt cctttctcc agtccagctc 2040
 aagctgctct acgcgatgtt cctcgtgacg ttctttgcaa aatctggaag cgtgcccttc 2100
 cagttctggg tgcccgatgc atacgaagcc gctcccagca atatagcctc gctgatggcc 2160
 ggcgatcatgg agaagatggc ggtttacggt ctgataagge tctctgcaa cgccctgcca 2220
 tgcagtgaag gcattgggta cgttctcgtt atcgtcggca tacttaccat gaccttcgga 2280
 accctctacg ccctcagaga gactcacgca aagaggctcc tcgcttactc aagcgttgga 2340

 caaatgggct acatctggtt cgcggtgggc atgggcatga tcttctgac gatgggcatg 2400
 gagagcctgg cttacctggc ctctctgcc ggagtcttc actcctcaa tcacacactc 2460
 ttcaaggggc tgcctttct catctcgggc aacttcgagt actccgccgg aaccgctgac 2520
 ctcaacgagc ttggtggttt gaggagggca atgccgtact cgtcgtcttt caccgtcata 2580
 ggtgcgctct ccctcgtgg agtgccctc ttcagcgtt tctctcaa gtggatgatt 2640
 taccaggcgg gctactactc tggaatcggc ctctctctt ttggctccgt aatggcgggtg 2700
 tttatgagcg ccgtaacctt ggcatattcg ctcaagctct acacctctgc ctttgggggc 2760

 gaaccgaacg agagaactga gaacccagg gaagtcccg cgggtatgct cctcgtgag 2820
 ggaattattg ccttaacttc acttgccgtt ggaatacttc cggctattgc ttaccgata 2880
 ttaacgattt cattgaatgg cggcgacgtc accgttaca tgggctcgat atccactgac 2940
 tttagtact tctcgcaat agccctgctc cttgcggttt cattcattgc ggttgcctca 3000
 tacttctctc tcaggccaaa gacgaccaat gtcaaacctt ggaacactgg agcgcttttc 3060
 ctgccggagg agaggtatgg agcgaaggcc agggactatt acaggcagta cttaccgag 3120
 atggagggcc tctacaagct tggaaagcgc gctggcaagg tcggaagggt ccttctctct 3180

 gctctgatgt ccgtctacct cgttctgcc aggggcctcg tctacaccgg cagggagaag 3240
 aagcgtctct tcaccttga cgagcttcgc caccgcaccg tcaggtacct ggacgaggca 3300
 ttcttcgcbc cgatgatgga tctactcaa aacatcggcg tgctggcagc gggcatctcg 3360
 gigtccatgg acgagctctt cctggcttca atgctgacca cggtgataat actcgactc 3420
 cttgtgtgtg gaatggacta cgtaagcatt atcgtctctc cgatcgtctt ctctctctt 3480
 ccaccgttcc ttgacggaat agggagaagg ataaaggcga ggattcagta caggagagga 3540
 ccgcctataa tgcagacgtt ctacacctc gaaaagcttc tcaagctgcc gtcagtctt 3600

ccaactgagg gcccaatctt caggctggcc ccgtacatag ccctggcatc tgccattgcc 3660
ggcggcctaa tgcttcctt cggaagcgag ccggtgttgg cttttgaaa gacccata 3720
gtgttcttct acgtcatggc gatggtcagc gtagtgatga tacttgctgc tttctccgtc 3780
cagaacgcgt tcctcacat aggtggacac agagaggtca tgctgatact ctcgattgag 3840
ccagtctgg ccgtcgtctt cgggtcctg gcattcaagc ttggaacgct caacgtcgt 3900
gagatgcctt tcagtgetaa cctctcgtt tccgttgccc tagcttacat ctgctggct 3960
tacgcggtct acgttgaggg cggattcgtt ccatttgaca tagctgaggc agaaaccgaa 4020

gtaatcgggg gcccgctcac cgagtacagc ggaaggtcc tcggagtctt caagtacgcc 4080
ctgctcgtca agagggttgt cctgctctgg ctgctggcgt ctatgattgt gattcccgcc 4140
atgaggtctc tcggtataac aagctcaatg gactgctcg tcgccagct ggtcgttacg 4200
tttctgcttt attcgttgc cgtggcgtt gaggctgcaa acgccgcct gaggatcgac 4260
caggcggttt cccttaaca gaaggtcttc ctgatgtccc ttgctgtcct gataatagcg 4320
ctggtggggt ggtgaatgga gtgcagcgtg tgtgcgggtg gatgcagatc ggctgaagtt 4380
gaggacgtcc ttgaggatgg tcatctaaag gaattcgtgg agaagtttag gggagcgtac 4440

ttcagtgca agaagctgac gaggaaccag tacctgttca tcgttgatag ggaggcactt 4500
ccggagatgg tctccactg gcacaacat tccgagctaa aagaaacca ctctcgtatg 4560
ggaacaggaa ccgatgagag gaacatcgcc ggaagtca cctacgtcc gtaataaac 4620
gttgccgttg agcctgaaa cggggagagg aactactggg ttattctgaa gcctacctc 4680
gacgaggaca acccggagt cccctccata gccgcgaagc ttccagcagc cctctggcg 4740
gagaggaag tctatgatc gcttgcttc aacccaaag gccatcccga cctgaggagg 4800
ctcgtcctgc cggaggactg gccggagggt gttaccgc tcaggaagga ccatgactac 4860

aaggcctgc cgatggatac gccaaagtgc tactacaagc ccgggcccgc cgacacaatg 4920
acggttccga ttggtccgta ccacctggcg ctcgacgagc cggcccactt caggatattc 4980
gtcaaggggg aaacggtggt tgacgttgc taccgaggct tctactcca cagggaatc 5040
gagaagatag gagaggaag actgacctac aatcaggtgc tctcatagc cgagagaata 5100
tgtggaatct gtggcttcca gcactcgac agctacccc aggcggttga aaacatagcc 5160
ggcgttgaat tccccgagag ggccatgtac ataaggacga taatgctgga gatagaaagg 5220
attactccc acatgctctg ggccggtgtt gcggtcacc tgacgggctt tgacacggga 5280

ttcatgcac cttggcgcgt cagagagcct gttatgtggc tcgacagag gctcacagga 5340
aacaggaaga cctacggaat caacatcgtc ggaggagtta ggagggactt cctcgactac 5400
cgcaaggaga tgataatgga gaagatcaag gagctcagga ggcaggtcga agagttcatc 5460

gaaatagcga ccggtacggc aaccttcgtc aagagggccg agggggttgg aattctgccg 5520
 tacaagtggt ccaaggctta ctcagtcctt ggtccgaac gaaggccag tgggaggaac 5580
 attgacatta gaaggatca gccgttcgca gcatacaagg atttgactt caaggttcca 5640
 gtctacaagg agggcgactt cttggcaagg ttctcatca ggatggacga ggtgctcgag 5700

 agcatctgga taatagagca ggccattgac cagatgccgg gaggagactt cttcgtgccg 5760
 ataggggagc ttccggagta tgaagaggcc ctaggctaca gtgaagctcc aaggggcgaa 5820
 gtcattccact acgtcatgac tgacaagaag aacaaggtct accgctggaa ggttagagcc 5880
 ccgacctaca acaaccttc agctgttccg gagatgctca agggctacag cgttgccgat 5940
 gccccgctca tcatagcgag catagatccg tgctactcct gtacggagag gtttcagata 6000
 gtggacgttg agaccgaaa ggcccagacc ctgaacgagc agcagttcaa catgctctca 6060
 atacagaagg gcaagggggt ggccctgaatg gccccaggca tttccttcac cgacaggctc 6120

 aagtcttggga agcgaccaga ggaggactt aagaaggctc ccgtcacgac ttcttatcct 6180
 tttgttgata tcgaaaagcc gccggaatat aggggcatac ctcgcataga tcctcacctc 6240
 tgcattgggtt gtggagcctg tgttagggcc tgtccaccgg acgctctcac gatagagtgg 6300
 gacttcgaga acgggagaa gaggatagtc ttcaacgccg cgcgctgcat aaggtgccac 6360
 cgctgcgtcg aggtttgtcc aaccggtgct atgcaggca caacgaggtt cgagatagcg 6420
 acgccgaaca agaggacct catcgaggtc gttgaccaca agctctacag atgcccgctc 6480
 tgtggcggtt acgaggagtt caccgagagg cagataggga agatgttcca gattctgccg 6540

 gaggaagtca ttgaccagca cggcatagct gagagggtt ttctctgcag ggagtgcagg 6600
 atggaggaga gcgccaagac cttggcgggt caaggccct atgcggatag ccttctcctt 6660
 tcctctate cgagaggctc aaaggtgatg ggtgagagga gatgaatgag cgggttgaag 6720
 tccgtttggg tcttccactg tgacagtggg agctgcaacg gctgcgacat agagatactc 6780
 gacgtgctca cgcctatta cgacccgag aggcttggga taaagctcgt gccgagtcca 6840
 agacatgccg atgccctcct cgtttcagge ccactcacga ggcagactta ctacgtgctc 6900
 aaagcagcct acgaggcgat gccgccgaag ccgaggatag ttgtggccat aggcacctgc 6960

 gcgtccagtg gtggtatatt ctacaacggt tacccaatct acaaccgaa ccctgagagg 7020
 ggaagcgaca ggctcaggac ggggtggaata gaggtccttt tggcggagta cgggaaaaag 7080
 cccgacatgt acattccagg atgtccaccg agtccggagg agatactata tgggctggcc 7140
 cagctcctcg gcctgaagga gaagaagatg aaggcgagat actactatgc agacgagatt 7200
 gagttcgttc ttccagagag acccatcgag gagaggattt acctgacgct cagagaatec 7260
 ctgaggcgcg tcgtgggta cttcgacagg gagaaggttc tcgaggactt catggccctc 7320

gtggaaaagg ctcaggagag cgagaacccg agggagaggc tccacgagct agtcatcgga 7380

tacttctga gggagaagga ttcccgtgtg aagttcgcga taagttcct cgaaacgag 7440

tactggaggt tgaaggatgc ctacgaaaag aggcacctgg cacttgtaa agctggtgta 7500

cgtaa 7506

<210> 14

<211> 2982

<212> DNA

<213> Thermococcus onnurineus NA1

<400> 14

atggtggact ggagactctt tgaaccttc ttaattatg cgagaaagaa gacccittgg 60

attgtgtcgt tctgtaccgg atgcggcggg atagagatgc cgcccctcat gacgtcaagg 120

tatgacctag agcgattcgg tatgataccg gacccgagtc caagacagta cgaccttcc 180

ctcatcacgg gctacgttac gccaaagacc ctcaaaagaa tcataatcac ctacgaaatg 240

gcacctgate caaagtacgt cctggcgcac ggtcctgcc cgctcaacgg tggaatctac 300

tgggacgctt acaacgctat caagcacctc gacaagtaca tcccggtcga tgtcgtcata 360

gccggatgca tgccgcgccc agaggcagtc atggacggaa tccagaagat aatggagatg 420

attgagaacg gaacagcaga cggatggaag aggtacaagg agaactacga gtggtacaag 480

aagaaccagg acgagctctt cggcgaagga tggcgtgaga gggaaagccag aaggtggatc 540

ccatggctgg tggacaagaa aaaggaggag tgaatggggg aagttaagt ggagagagag 600

cagatgctcg ttgacaaaat ccttgaaaaa gccccctacg ccgagggcaa ggtgcggcgc 660

gaacggagga ttgattcag cattccggca gacaggataa gggacttctt catgctgctc 720

agggataacg acttcgagct catgctccag ataacgaccg tcgactggcc caacgacggt 780

gagcttgagc ttatctatca gatgtggagc gtgaccaca gaaccacgc catggtcaga 840

acacggatc cgagggatct cgataaggca aggatgcca cctcaagga tatctacct 900

gtggctgaga cctacgagag ggacccccac gacttctttg gacttactt cgaggcaat 960

gagaagatgg agatgccgtg gatcctcgac gataccgagc aggggctctt cccgcacaga 1020

aaggacttcg acatgctgac ctactgtaag aagaagtaca agctgctcga caggttcgat 1080

gaggataagg acaactacgt gatttgaatg gtttcacaga atgagctcat tcgggaagcg 1140

agagaaaatg ggatggatct gctcccaatc gataaggaca cttacgagtt gttctttggc 1200

ccacgcaca tggtactga gaacttcagc ataatctca agatggacgg tcacagggtt 1260

gtgaaggcta tagccaacc cggcttctc cacaggggtc ttgagaagct cgccgagtac 1320

aggccgtggc acacgaacat agcgcctcctc cttagaatct gtgttccaga gccagacgtc 1380
cccaggcaaa tatactcaat ggccgttgat gagataatcg gctgggaggt tccagagagg 1440

gctcagtgga ttagaacaac cgctcctcga atggcgaggg tttccgcata tetgttctgg 1500
ataatgggtc tcagcttcaa gctcgggtgc tacactgccg gtcagtgggc tgctgcctac 1560
agggagagge tgatggccct cttcgagcaa ctgaccgtg ccagggtcta tcacatatac 1620
accatccccg gcggtgtcag gagggacatt ccgggcgaca agtggcttcg ccagctcaag 1680
gacaccgtcg agtacatcag gagcaagtta tcagacttcg acaaccttgt cttcgagaac 1740
tacgttgccc acaggaggtt tgagggaatt ggagtgatgg acaagaagtt tgcctcggc 1800
gaaggcgtca ctgggcaaaa cctcagagcc accggcgctt cctacgacgt gaggagggca 1860

gatccatacc tgctctatcc agagctcgac ttcgaggttc ccgtcctgaa ggagggcgat 1920
gccctcgcga ggcttttgat aaggcgcttc gagcttgagc aggatcttta catcctcgac 1980
cagctcctcg agatgggacc gccgagcggg ccgtataagg ttgaagatcc caagctcaag 2040
aatctcccga ggtttaaggt tccggctgga gatgcatttg cccacgtgga atcaacgaag 2100
ggcgactttg tgccctatgt cgctcagtgt ggaagcaca agccgtacag agtgcagata 2160
aggggcccga gtatagccca cggagtcagg gttctcgagc agctcttggt tggagcaaga 2220
atagccgacg tccccgtgat attgatgagc cttgacaact gtcaccaga cattgacagg 2280

tgaatggagg ttgattttaa ggtcggccca gaggagaaag tcaggaagaa gccatcattc 2340
atcaagccct ggatgggctt caagtacctc ttcaagaagc ccgttactat caagataccc 2400
tacgagaggg tacagatagc taaggactac aggggatcc acaccctaga ctggaagaag 2460
tgtgtcggct gtaacttctg cggccagata tgtccggcga gggcaataga gatgacctgg 2520
atagaagtgg atggcaagat ggagaagagg cccatccaa agatagacta tggcaggtgt 2580
accttctgtg agttctgtgt cgacgtctgt ccacctggag cgctgggctt catcgagaac 2640
tacatcctca ccaccgagtg gaaggacgag gagctggagc tctttgactg ggttccaatc 2700

catccagaca agttcagga gataaacgag aagttccccg actacagatt cccggtggag 2760
aagatagagt tcaacaagga aacgaaggag gtcacctact acctgagaga cggagaggtc 2820
atgaagtcca agatactcgg ctacggcatc agaccgccga agccaccgac aaagcctgct 2880
cagaaggcag ccgcaaaagc agcggaaaag aatgatacca agcctgttga aaagccact 2940
gagaagaagg aagctgggaa gatagaagaa aagaaagaat ga 2982

<210> 15
<211> 11208
<212> DNA

<213> Thermococcus onnurineus NA1

<400> 15

atggagacc taatcctcgc cctcggcaac gaacttatga aggatgatgg tgtcggcctg 60

aaagtccgca ggcttttggc ggaaaaaggt tacaacgttc tagaagttgg cacagacata 120

ttcatgctcc agagccatta cagtggagag gaaaggctca taatcatcga cgccatactt 180

agcgaaaagt tcaagccggg agaaataatc cacgtgagcg gcggtgaagt tttcgagaag 240

ctgaaagctg aaattaggag tgcgcacttc atgggggcaa tagacgggct caaactcctc 300

atggcactcg atgaaaggct tgccaacgtt gaaatccact tcattggtat tgttgctaag 360

gaaatcaacc tgggtatgga actgaccgag gaagttaggg aagcccttcc gaaggetgtt 420

gagctcgttg aagaattagt caaaaaataa atgaagaacc tctaccttcc aatcacctgt 480

gatcacatag cgcgtgtaga gggcaagggc ggcgttgaaa tagtcgttgg tgatgacggg 540

gtcaaggagg tcaagctcaa catcattgag ggtccgaggt tctttgagc gataaccata 600

ggcaagaagc tcgacgagc tttggccgtt tatcctagga tatgctcctt ctgttcagca 660

gcccacaagc tcaactgccg cgaggcggct gaaaaggcag ttggctttga ggtacgcgag 720

gagatacagg ccctcagga ggttctctac attggagaca tgatagagag ccacgcctc 780

caactctacc tctcgtcct gccagattac atgggctact ccaaccgct caaatgctt 840

gacaagtaca agaaagagat caacattgcc cttgacctta agaacctcgg aagctggatg 900

atggacgagc ttggtgcaag ggccatccac caggaaaacg tcgtcatggg tggttttggc 960

aagctccccg acaagactac gctggagaag atgaagaaga ggctccagga ggcccttcca 1020

cttgcagaat acaccttcca gctgttctcc aagcttgagc agtacgagga agtcgagggg 1080

ccgataatcc atatggccgt caggccgagg ggagacgtct acggcatata cggcgacgcc 1140

ataagcgtca gcgacggttt cgagttcccg agcggaggct acaagaagca catggttgag 1200

cgcgttgg agcacagctt cgccaagcac agcttctaca acggggagaa gcccttcatg 1260

acgggagcca tatcgcgcgt tgtcaacat gctgataagc tctacggaag ggcaaaggag 1320

ctctacgaga gccacaagga cttactcagg ccaaccaacc ccttcgcaa caacctgcc 1380

caggcgttg agttagtga cttcatggag aggggcatcg acctatcga cgaggcctc 1440

gctaaatggc cgataaggcc gagggacgag gtagatgtaa aggacggtt cgggtgcagc 1500

accaccgagg caccgcgtgg aatcctcgtc tacgccctag aagtcaagga cggaagagta 1560

gcttacgcgg atatcataac gccacagcc tccaatctcg ccatgatgga ggttcatgta 1620

cgcatgatgg ctgagaagca ctacaacgac gatcctgaga ggctgaagct ccttgcctgag 1680

atggtcgtca gggcctacga cccgtgcatc tctgttcgg tgcacgtggc gaggctttag 1740

atggagggga aggttcgtat aggtttttac gctctcacct catgctacgg ctgtcagctc 1800

cggttcgcca tgatggacga gatacttcag ctgctcccga acgctgaaat cgtctgctgg 1860

tacatgcttg accgtgacag cagtgaagat gagccccgctg atatagcctt catcgaagga 1920

agcgtttcaa cacaggaaga ggttgaactc gtcaagaaga tacgcgagaa cgccaagata 1980

gtcgttgtag ttggggcgtg tgcaaccag ggtggggttc agagctggga gaaggataag 2040

agcctcgaag agctatggaa ggccgtttac ggcgatggga aggttaagtt cgagccgaag 2100

atggccgaac ccctcgaaaa ctacatcaaa gttgactacc gcatctacgg ttgccacct 2160

gagaagaagg acttcatcta tgccataggt accttctctg ttggttcttg gccagaggac 2220

atcgattacc cggctctgtct tgagtgcaga ctgaaggta acacatgcat tctcatcgag 2280

aagggcgagc cgtgccttgg cccgataacg agagcgggct gtgacgcaag atgcccgagc 2340

tacggaatag catgcatagg atgtagaggg gcgataggct acgacgttgc atggttcgac 2400

tcacttgcca ggacgttcaa ggaaaaggc ctaccaagg aagaaatcct cgaaagaatg 2460

aagattttca acgcacacaa tccaaagctg gaggagatgg tcgataaggt ctttcagttt 2520

caggggtga aagaatgaat gaacgaggct cacgtctgca tgtgtcacga caatccctac 2580

gcccttgaca ggttcaaggt tctcagagtg taccgtttaa cagaaacaga gaagctattt 2640

ctgttcaggt ttgaagatca agagatagcc gagaactgga cctttaagcc aggacagttc 2700

gttcagctta ccataccgg tgttgcgaa gtcccataa gcatctgctc gtctcccatg 2760

aagagaggat tctttgagct ctgtatcaga aaggcgggaa gagtaccac agttgttcac 2820

aagctcaagc cgggcgacac tgtccttgtc cgcgggccgt acggaacgg cttccctgtc 2880

gatgagtggg aagggatgga cctactctc atcgcccgag gattaggaac ageccccactc 2940

aggagcgtct tctctacgc catggacaac cgtggaagt acggaacat aaccttcac 3000

aacaccgcc gttacgaaa ggacctctc ttctacaagg aacttgaagc catgaaggat 3060

ctcgagagg ccgagaactg ccagataatt cagagcgtaa ccagagatcc tgactggccc 3120

ggcagacatg gaagcctca gaagttcata gtcgaggcca ataccaacc gaagaacacc 3180

gcatagcta tctgcgttcc gccgaggatg tataaggcgg tcttcgagc gctcatcaat 3240

tacggttacc gccctgagaa catctatgta acgctggaaa ggaaaatgaa gtgtggaata 3300

ggcaagtgcg gccactgtaa cgtcggaacg agtactagct ggaagtacgt ctgcagagat 3360

ggaccggtt tcaactactt cgacatagta tcaacgccg gactgctgga ctgaatgaga 3420

tacgtcaagc ttccaaaaga aaatacatat gaatttttag aaagactaaa aaacctcgga 3480
 aagctgtacg ccccaagtga aatttccgat cagttctatg atttagaga aattgacgac 3540
 gtcagaaaga ttgaattcaa ctacacgaga accctgatgc cgccgaagaa gttcttcttc 3600
 gcaccaaggg agaagatggt cgagttcagc atctcaaaag cagagtatag ggaagtaatc 3660
 cccgaagtcg agcccttcgt cctcttcggt ctccacgctt gcgacatcta cggcttgaag 3720
 atactcgaca gcgtatatct ggatgagtac ccagacaagt actacaaagt caggcgtgaa 3780
 aaaggcataa tcatcgggat aagctgtatg cccgacgagt actgcttctg caacctgctc 3840

 aggacggact tcgagcacga cggctttgat ctgttcttcc acgagctccc tgacggctgg 3900
 ctgataagga tagggacccc caccggtcat aggatagtcg acaagaacat caagctcttc 3960
 actgaagtcg cacaggagga catctgcaac ttcagagagt tcgagaggaa gcgcgcccag 4020
 gcttttaggt atcatgagga gtgggacaac atccactacc tcctcgagct ggagatggag 4080
 caccctctct gggagaaaga ggccgagaag tgcttcgctt gcggcaactg cagcacggtg 4140
 tgtccacct gccgctgcta tgaggttcag gacatcgtca acctcgacgg agacacggga 4200
 tacagggaaa ggcgctggga ctcgtgtaag ttcaggagcc acggactggt cgcgggcccgc 4260

 cacaacttca ggccgacgaa gaaggaccgc ttcataaacc gctatctctg taagatgtcc 4320
 ttccactgga cacttggaat taacttctgt gtgggctgtg gaagatgtac tgccttctgt 4380
 ccggcgggca ttgatttcgt gaagaacctc agaattatag ctggattgga ggatgcatcc 4440
 tgcccgtcaa agctgagcga ggaaattcca aagaaagggt ttgcatatgc caacaacatt 4500
 agaggtgaag acatatgaat ggcacagaat aattcactcg tgctgtatga tgttcatgag 4560
 accgtggatg tigtctcaa cgttggctgt gttaagacca aggccactcc atcaaggctg 4620
 ctctttgceg gtttcatggc tggctcatal atagcctttg gattcatttt tgccatagtc 4680

 gctagtgcaa gcttccatcc aaagetgggc actttcccaa acctatctct cttcaagctg 4740
 cttctgggtg cagttttccc agtcggtctc atagccgtcc ttcttggcgg tgcggacctc 4800
 tggacgggca acgcccata agtaaacactt tcgaaaatga cgggcagggc gagcgttaag 4860
 gatgtgctct acaactggat cggcagctac acaggcaatt tcgtaggctc ggtcttcttg 4920
 gcattcttgg cagtttacgg aacggggctc atggcagggt gtttgttcaa ggacgttctg 4980
 ataggcattg gcaactaaa agtggcgctc accccatgga aggccctctg gctgggaata 5040
 ggctgtaact ggcttgtgaa cgtggcgata tggctctaca ttcgcgcaa agacactgcc 5100

 gggaaggtaa tcgtaacctg gttcccgatc ttcgcttctg ttgcataggg ttttgagcac 5160
 agcatagcca atatgtgggc cataagcgcc agcatatttg cctcgacggg tgcgataagc 5220
 tgggtccagt tcttcacaa cataatcca gtcacgatag gaaatgcat cggaggcttc 5280

ctctttgtgg gcttctacca ctggtagctc gctgacggta gaaatgcat taaagagctg 5340
 attgactttg tcgaggctgt ggcactcttc gcttttatca tgggtcttat cccagcggga 5400
 atagcctacg ccctcagcgg tctcggaaac attgccacat ggcttgtgcc actcatcata 5460
 agtgtctatg gaggttgat gacgtattta gtaaggagag cgctgtgaat ggaggagttt 5520

 aagattggcc tgtgccata ctgtgggatg gggtagcagg tttacataaa gactctaac 5580
 gggcagccca taggaataga gccgtatccc ggtgggttta atgaaggaaa gctctgtcca 5640
 aagggtgtcg ccgccgttga cttcctcaga cacaaagata ggctgaaaaa gccgctcaag 5700
 agaactgaaa acggcttctg cgagataagc tgggaacagg cgataaagga gattgtctgaa 5760
 aagcttctgg agatacgcga gaagtacggg ccggtatcgt taggcttctt ctcaagtgcc 5820
 cgttgttcca acgaggagaa ctacctctg cagaaaatag cccgccttct gggcaccaac 5880
 aacgtcgacc actgcgcgag gctctgtcac gcctcaacgg tcgtcggctt tgctcagacg 5940

 gttggcgtg ccgctcagag cggctcctac acggacatac ccaaggctaa ggtactcctg 6000
 atatgggatg acaaccgctc agaaaccac ccggttctca tgcgctacat cctccgcgcg 6060
 agggacaacg gggccaagat aatcgtctga gatccgagga agacgaggac tgtctggttc 6120
 gccgatatgc acctccagct taagcctgga acggacatag tcctagccaa cgccatgatg 6180
 cacgtcatca ttgaagaaag gctctatgac agggagtcca tcatgaaccg gacgaagggc 6240
 tttgagaagc teatagcagc tgtccagaag tacacgccag aatacggcga ggaataaacc 6300
 ggtgttcccg ccaagctcat cagagaagcc gctataacct ttgctactgc cggacggggc 6360

 atcgtgatgt gggcaatggg actgacgcag cacgtcactg gggcggccaa cgttaaggcc 6420
 ctcgctgacg tggctctgat ctgtggctac gtcggaagag aaggaaacagg tctcttcccg 6480
 atgcgcggtc agaacaatgt tcaggagaca tgtgacatgg cagccttgcc aaacgtcttt 6540
 ccagctatc agaaggtaac tgacgacgag aagaggaagc acgtggcggg aatttggggc 6600
 gttgaagatc tgcctcgaag gccgggcctt actattccag agatgattga tgcggctgct 6660
 aaaggcgagt tgaaggcact ctacataatg ggcgagaatc cggctcatgag cgatccgaac 6720
 acgaagcagc ttatcgaggc tctcaagaac ctggaacttc tcgttgttca ggatatatc 6780

 ctaccgaaa cggccgagct ggctcactac gtgctcccag cagccgcata cgccgagaag 6840
 gaaggatcat tcaccgcgag cgagaggcgc gtcagtgga acttcaagcc gattgagccg 6900
 ccaggagaag ccaaacccga ctgggagata ctgacgatgc ttggaaggc tctcggcctg 6960
 ccaaagtctg actactcaga cgttgaagat attacgagg agataaccct cgttctcgg 7020
 cagtaccgtg ggataacccc cgagaggctc aagcgagagg ttatgggtgt gcagtggccg 7080
 tgcccagcgc aggatcatcc tggaaacccg aggctgcacg tcgagcgctt cgccaccccc 7140

gacggaaagg ccaacataat ccccgtagag ttcaagccac ctgcagaaga gcccgatgag 7200

gagtacccat tcatactgac gacattccgc atcgtcggcc agtaccacac actcacgatg 7260
agtaacagga gtgaaagctt gaagaagcgc tgggccagcc cgtacgccca gataagtccg 7320
gaagatgcaa agaagctggg tatacaggat ggtgaaatga taaggatagt tacgagacgt 7380
ggaagctaca cctgcagggc ggtcgttact gaagatgtct cggaaggggt gatcgcagtt 7440
ccgtggcact ggggggccaata taaactcagc aacgatgtcc tcgatccaga agcaaagatt 7500
cccagactga aggtggccgc atgtagggtg gagaagattg gggggtgctg aatggagaaa 7560
aagctgttca taaacctcgg gcgctgcatt gcctgccgcg cctgagaggt ggctgtgag 7620

aaggagcagc gaatttcatt catcacggtc tatgagtca gggacatagc ggttccctc 7680
aactgccgc actgtgagaa ggctccgtgt atcgaagtct gcccagcga ggccatctat 7740
cgcgacgaag atggcgagc tgtgatagac gattccaagt gtatcggtg ctacatgtgt 7800
tcggccgtct gccctacgc gattccgata gttgaccga taaaggagct ggctgtgaag 7860
tgtgacctat gtccgaaag aaggaaggag ggcagagatc cgctctgcgc tgcggtctgt 7920
cccaccgatg cgataatcta cgctgacctc aacgagctga tggaaagaaa gaggaggcgc 7980
aaggccgagc gcatcgtcga agcccagagg aaggcggtcg aaacgctcgc ctacttcggg 8040

tgagtgtga aggtggagct ctgtgtgggg tgtggggttt gtgcaaagc ctgccccac 8100
tcggccattt cagttttga agatagtgtg aggaggatag tcttcgacc gaagaaatgc 8160
gaagaatgct cctttgagtg caacgaagcc tgcccaacgg gggcgctgga agggaagtca 8220
gacaaaaggg agctggtctt tgagtttgc tactgtgcca tctgcgggaa aaggctcaac 8280
atcgtgaagg aagaagccga atatcttgca aaaaagctga ttgagctggg tgaaaacct 8340
gagattgctt ttctctgta tgaactgcaag aggaaaagc tgtttggcgt tgccaacaaa 8400
tatgaggctt acctggggtg aatgagcggg atgaggtttg cgttctctgt tagggaaaga 8460

ccagaaccaa ctgggaagaa gatagccgtt atcggagccg gaccggcagg cttggcggca 8520
actggctacc ttgtctgtca ggtcatgag gttcatgttt acgacaagtt gccggagcct 8580
ggaggattaa tgctatttgg cataccagag ttcaggattc caatataccg cgtagagaa 8640
ggctatgaag aattagaag ggtctacaat gtcaagtct tcaccagaac caaagtgtac 8700
ttcgggaatc tggaaggaga atcaggagac gatttcgttg agaacagggt agacttcaag 8760
gaactcgtgg agaagtacga tgcggtacta atagcaacag gaactggaa gtgctggatt 8820
ccaaacattg agggagcggg gcttgagggt gtcttcccgg ccctcgagta tctctttagg 8880

ataaagagcg ccaagctcgg ccacatggat tggggcaagg tcacaccagt ggagggcaag 8940
 aaagtgctgg tcgttggtgc cggccacaca gccgtcgatg ccgcattgga gagcgttctc 9000
 ctcgagcggg ataaggtgta cctcagttac cgcaggacga taagggaggc tcctgcgggg 9060
 gcctacgaga ttaacctcct ccagcagagg ggtgtgaagt ggctggagag gacgatgccg 9120
 gtcaggataa taggtgagaa cggaaaggtc agagctgtgg agctggtgaa gaccaagctc 9180
 agtgaacccg acgagagcgg caggaggaga cctgttccaa tagaaggttc gaacttcag 9240
 atagacgtgg attatgtcat cttcgccgtc ggtcagagcc ccaactccacc cttcgagaa 9300

 gagatcgata tagccgtcga taagaaggc aggatcgtag ttgataacag gcacatgacg 9360
 agcagggag gtgttttcgc cgcaggagac gtcgttttag gccctcaaa ggttggtgaa 9420
 gctgttaagg atggcctgta tgctgctgag gccatgcaca tgtggctgat ggggaggtga 9480
 atgacgagga gaatcctca cgttgattac agcctttgta ttggctgtga gacgtgtgag 9540
 gcagtctgtg acttcctcca cggctggcaag cccaacataa ggatttacta cactgtcacc 9600
 ggacttcga ttccaataaa ctgccccac tgcgagaggg caccctgtat ggacgtctgt 9660
 cctgcaggtg caatttaccg cgacagcgac ggagccatca taataaacc tgacaagtgc 9720

 ataggctgct acatgtgtct tgccgtctgt ccatttggcg tgccgagctt tgacgtcaag 9780
 actaaggcag tcacgaagtg cgatatgtgt gccgacagga gaaggcttgg catggaacct 9840
 gcctgcgccg agatgtgtcc cgcagaggca atattctttg gaaagcccga agaggtcgag 9900
 gacaggataa gacgcaggac tgccgagagg atagcacgcg agaggatagc tgccgtagac 9960
 atggaaggtg ttgggaggat gctttaaatg ttgtgggagt cccagatccc cataaatcag 10020
 gtgttcgaac ttcgctgcag atccatgaca tacttcggtg ttggagccat taacaagttc 10080
 tacgacatag ccaaggatct taaggaaaac cgcggcataa ccaaggtcat tctcgtcact 10140

 ggaaagagtt catacaagaa gtgtggcgcc tgggacgttg tcaagcccgc cettgaggag 10200
 tacggtattg agtacgtcca ctacgataag gtcggcccaa acccaaccgt tgatatgac 10260
 gatgaggcca ccagctcgg taaagagttt ggagcccagg ctgttattgg cattggtggt 10320
 ggaagtccaa ttgacagcgc caagagcgtt gcgattctgc tggagtacac cgacaagact 10380
 gccagggacc ttacgagct taagttcacc ccaacgaagg ccaagccat catagccgtg 10440
 aacacgactc acggaactgg aactgaagtc gacaggttcg cagtggcttc gattccggag 10500
 aaggagtaca agecggccat agcctacgac tgtatctacc cactctactc catcgacgac 10560

 ccggcgctta tgacaaagct tccgcccac cagactcgt acgtaacat tgatgcctt 10620
 aaccacatca cagagccgc caccaccaag ttgctagcc cttactcaat acttctcgc 10680
 caggagaccg ccaggctgat attcgactac ctcccagagg ccttggctca cccggacaac 10740

ctgcaggcca ggiactacct gctctacgcc tccgcgatag cgggtatata cttcgacaac 10800
 ggctctctcc acttcaccca cgctcttgag caccgcctca gcgccgtaaa gccagacctt 10860
 ccacacggac ttggcctcgc catgctctctg ccggcagtta tcaagcacat ctacceggcc 10920
 accgccagga tactcgcgca ggtctacagg cactctgttc cagaagctaa gggagtcccg 10980

 ggagaggctg aactcgtcgc aaagaaggtt gaggagtggc tcttcaacat cggatcact 11040
 cagaagctga ttgacgttgg ctctactgag gaggacgttg ataagctcgc cgaactcgcc 11100
 atgaccacc caagcctcga cctgttgctt tcaactgccc cgattgaggc taccaaggag 11160
 accgtcgcgg ccatctaccg cgactcactc taccgcctta acaagtga 11208

 <210> 16
 <211> 4986
 <212> DNA
 <213> Thermococcus onnurineus NA1
 <400> 16

 atggaggtaa tttttctttt tattgtcatt attctgtcag ttgcatcttt cattggagtc 60
 ttttcgagga gcgcaatfff aacgaagtta gtaaattgctc tctccgctct tggctcattg 120

 acaatagcct atgccgggat ttaggtctt aaagagagcg ttgaattaaa taccactctg 180
 ttacatctga aatcggattc cattatcaat gcgtttctca ctctaacact caaagtcgac 240
 ccgctgtcag gctttttcat gataatactc ggaattctgg gcttctgtac atcggtttat 300
 ggtattgcat acttagacat gtataaagga gacaagagac tatatgcctt caactatccc 360
 ctcttctcgc tcttcatgtt ccttgttctc gtctcatgga atctcttgtg gttcgttgtg 420
 ttttgggaac tgatgactct cttctcccag ttctgttag cgtttgaaag gaacgagaag 480
 actctcattg cgacctcaa gtacttctgc atgactaaag ccgcagcaga ctttatgctg 540

 atagccatag tattggtact cataacaatc tctggcggag gtgattacga taccctctcg 600
 tcccagctcg taaactatff ccgctctcat cctctggaga tgtatcttgt aagtgtgga 660
 ttcatgatcg gacttgggtg caagccgcc cttgtgccgt tccacgatg gettccagac 720
 gcatacgtgg aggaccaag taacgtctcg tcattgtca gtggggccat ggaaaaaatg 780
 ccagtgtata tgatgttccg ctttttctg agtttcacc cactaacccc taatattggt 840
 ctactcatag cactgttcgg aacattaacc ctgttttttg gtacgatgta cgcactaaag 900
 caaacagact caaagccct actggcctac catagtgtcg gccagatagg ttacgttctc 960

 ttgccccttg gggcagggat atatctctt tccaaggggt acaccacatt cggagctctt 1020
 gcccttatgg catccctatt ccacgtctc aaccatgcat tcttcaaggg actgctcttc 1080

ctaacagcag gctcgatcct ttatagaact ggaagcaggg atttggacca cttgggagga 1140
 ctagecagat ttatgccgat aacggcattt gctgcaactga taggttctct ttccatagct 1200
 gggatgccac cattcaacgg cttttgttagc aagtggatga tatacgtttc aacacttccg 1260
 actccgactc tcgtttccct gtttggggcc ctageactat tcataagcgc tgtaacaacc 1320
 gcacttttcg ttaagtactt cacttccatc tttgtaagac cgcctgcca ggagataacg 1380

 gtcaaagaag tcccagtatc aatgtgggcg tcccagtga ttctcgagt tctttgtga 1440
 atttttggtg tttatccagc attgccactg gaagcaatct caaaagcggg tgactcagta 1500
 ggcgtgacca ctccatcaat cacggtcttt cccggtctca tagtgtccga cggatttga 1560
 aacatagctc ctctggctct cctggatttc tccggagctc tgaccgcagt gttactagcc 1620
 attttcccat acaaaatcag tcttccggtg tggacaactg gcacgagacg gtccctggcc 1680
 atgaggcttc cagcgagctc atactatgcc tcctttgagg aagaattcga ggatgtttat 1740
 agctggggag aatggtgtgt atgtaccacg aaaagactat gggacgccac aaaagccgtc 1800

 ttgtccaact ttgaggaagt atccttcgac ttggacaaga tgatgactgg agcttggcta 1860
 atgctcctta tactccttac aatactcggg ggtgttctgt tatgaatgaa tgcagtttat 1920
 gctgccctca atctaacttt catagtactc tttgctccgt tattagacgg aatcgagagg 1980
 aaagtcaaag caagacttca gtcaagacaa gggccgccgt taatccagac gtggcttgat 2040
 ttattaaage tcttcagaag gccaaacgtc aggccatagg agtccgtaag atggctgttc 2100
 gaaccagcac cagcaatagc gcttgtatct gtattggcgg cgtccctgtt catcccatca 2160
 ctgcttctg gctctttaga cacatggggg gatataatcg ccttcatata cctctcaacg 2220

 ctctcagccg tcgcatagc tctcggagcg ttctcaactg gaagtcata tgcccaaata 2280
 ggatcccaca gagaagttc aatcataatg gcagaggaat tctccttggc ttttatagtg 2340
 gccgcaactc cagcatccag tggaggtctc tcgttctcgc gacttttccc cctccaacta 2400
 aaagtatcta ccataacagg tgctctggca ttcgcagtta tggcatacgt cgcgggagcc 2460
 agaatcccat ttgacgtcgc tgaagccgaa ccagagatag tcgagggtcc cttcatagag 2520
 ttcagcggga aaggcctggg aatgttaaag ctctcaatct acgtgaaacg getacttctc 2580
 accacgatac tctgaactt cttcctacc caagatggca cagtgagagt actagtctac 2640

 gtcattggac tagtcatcat atcagttgtt tacgcgtcaa ttgaagccca ctatggaaga 2700
 ttcaggacta aagacgccgc cagattcctc aagcgttttg caatagttgg aatcctaagt 2760
 tggattttgg gagtgggtgg gtggtaaatg gtatttgata tcctcaaagg atgtaaata 2820
 ctggagcaca atgataagat gacagtcgcc gaggtcggcg ccagcaatat acgggagatt 2880
 gcaagggcgt tattcgagag gggttattac tactctagtg gcattggagt agacgaacgg 2940

cccataaacg ggaggtttgc aatgtaccac atattcaact gcgatacaga gggaagatat 3000
 gtggttctca agataacatc ccccgaaggg agccttgagg tacctcaat aaccctgtt 3060

 atcaagggtg ctgaatggtc agagagagaa gccatggaca tgctcggcat agttttcagt 3120
 gggcatccaa agcccgaag gcttattcta ccggacgatt ggccagaagg agtctatccc 3180
 ttgagaaaag actttccta caacaaaaag ctccaccgt caaaacccat agaaaaagaa 3240
 agggagcaca aaaaagacgt catggagata ccctgggac catacatcc ctcccttcac 3300
 gaaccagagt attttgagct ctatgttaa ggagacaaag tcgtagatgc ggaatacagg 3360
 ggatttaca tccatagggg aatggagaag ctgctgaat cacgaatgac aataaaccaa 3420
 atcccatcc tcgaggag gatatgtgga atctgtggt gcaccattc cgccgatac 3480

 tgtcaggcag ttgaggatgc tgctggcatc tacgttccgg agagagcaca gtatataagg 3540
 acaataatgc ttgaggtaga gagaattcac agccacctcc tttggttcgg gtttgtatgc 3600
 catcttctcg gctttgacag cggcttcatg cacatctgga gggctagaga atacataatg 3660
 gacatagcag agctcataac cggcaacaga aagacctacg gaataaatat tgtcggaggg 3720
 gtaagaagag acatcacgga ggataaaaag gaaaaaacgc tgaaacttct ggacatggtt 3780
 gaaaaagaga gcagggaagt acttgataac atcgctgaga tgaaggagct cagagaaagg 3840
 atggagggtg tcggagtct accgaagaaa gaggccaggg agataggtgt ggtcggccc 3900

 atggccagga gctctgggat tgatactgat gtaaggcgag accatcccta cgcggcttac 3960
 aaggacttgg acttcaaagt cccggtttac aaagaagggg acgtttttgc aaggttcctc 4020
 gtaagatacg aagaaatfff tgagagcttc aatatgataa gacaggccct ggaaaatag 4080
 cctccaggag aactgataaa tgacgaatat gagattctc cattcaaact cggatcggga 4140
 gtactgaag ctccacgtgg ggagaacatc cagccgtga taacatgggg agagaacatg 4200
 atttaccgt ggcatecaag agctgcaacc tataacaacc tcccggcgt tectataatg 4260
 ctgagagga atgacgtgc agatgctccc ctaataatag ccagcataga cccatgcttc 4320

 tctgtacag accagctctc aataattgat tctgaaagcg ggaagattct tggagaggg 4380
 ccgcttaagg agggcgtgag gaggtctga atggtaaaaa atagtctatg ggttttcat 4440
 ctcaactccg gctcgtgtaa cggctgtgat atagatcc taaatctt tgcaccacga 4500
 aacgatgtt aaagactcgg gataaagctc gttggctctc ccagacatgc agacccata 4560
 gcatttacc gaccaattac aaggaggtgt ctgccaagg ttattgacgc tctgaaagcg 4620
 gttccggagc caaagtgtt tctggccata ggagcgtgcg cctgtggagg gggcatatgg 4680

tatgatactt aciccgtaat aggtgggtgtt aaagagctct acaggattct aaaagaagaa 4740

tacaacatgg agcctccgc gacggttttt atacctggct gtccccaaa gccagaggcc 4800

ataatctacg gtgtggctgt tgctagtggg atgctagagt caaacagaa gaagactgtc 4860

tatgtcgagc cggaggaatc tgtggcaaat gagaagctaa tgatcgccga gctcataagt 4920

gaaacagaaa agacaggca ctttatgccg ggaattgtca tcaggggggt tgaggatgag 4980

ccttga 4986

<210> 17

<211> 3447

<212> DNA

<213> Thermococcus onnurineus NA1

<400> 17

ttgagtgaga ttaccctcaa caaagtgtgc cgaatcgccg gtgaggcga gctcgtgctg 60

tatgaggaaa acggaacagt tcaagatgca ctcttcacg ccactgcccc aattagggga 120

tttgagaage ttgtagtagg caaaaateca cttttcgcgg tcgaagctgt catgagaatc 180

tgcggtctct gtcatgcgtc ccacggtata gcgatgagcg aggccataga aatgccatt 240

ggaataatac caccaagaaa cggaatactg atgagagaag ccctcggcct cgtgaacagg 300

attcagagcc acatgctcga gttctcatg gttgctgggg acctgctaata cgaggagaaa 360

agggaagaag ttctgttcca gctcatggac ttccacgcca agatcagcga ctacctctc 420

aagatgggag ggcagcaac acatcccca aacctaccg tgggaggaat gttctctgctc 480

cccaagtgga gcgtcttcaa caacctcaag gcacgccttc caaagctgac cgggcagtgg 540

gaagagatag cacatttgtc gaccgatgag gacatccaga cagaagttgc tgatgaactc 600

agggagaaga aagcggaaaa caactacctg gtaagcagcc tgttctacgg ggacaggttc 660

aacataaacg ccgagagaat tgagacaatg ccctattatg aatacagaaa ggacaacccc 720

cactcaaagg agtccaccac actcatagcc ttctacggtg gggaaaaggt tgaagctggc 780

ccaagggcaa ggatgaaagt ttaccgggag ttacagatt cttcctcta tggccttcac 840

accgcgaggg ttcaggatac aacgctggca ctcattaggc ttgaagaaat ccttgacagc 900

ataaagatgg acgagccgtt cagaacgaag aacatagttt tcgggccagg caagggtgtt 960

ggagtctacg aggcaccaag gggaacactc atccacttga tcgaacttgg agacgagggc 1020

agggtggttt cctccaagat aatcgtcccc acaatgttca acattcccgt gatggaggag 1080

atggcaaaag gtctgagcgt taaagcggcc gaggccgtta tgcgcctata tgacctatg 1140

attccatgta cgaccacgt tgtgaggttg ggggatgaa tggagaagct taaggttctt 1200

catgttgatg tagggggttg tgagggatgc aacgtcagta tcattcgcgc atatccaaag 1260
 ctcatggact tgatagagct cgacatatca tacctgcgga aggatgagtg taagctcgac 1320

 gagtacgacg tggcgataat aaccgggtgga gcatgtatga acgaaccaag gattcttgaa 1380
 gagctaaagg agataaggga aaaagctcac actgtggtgg ccttcggttc gtgtgcaacc 1440
 ttcagcggga tattgcgctt ctgcccgccg ggcaggagc caaggcccga ccacaggaac 1500
 ttccagccca taaacagcgt gattaaagt gattactcca tcccgggctg cccgccaaca 1560
 ccacagatgc tccagtcctt cttcaagttc tacatcaacg gtgacgagag aaggctgagg 1620
 ctcttcaagg tgagtgccga cataaagaag ctgagcggct ttgacctgat agacgatata 1680
 gtgcttacgg gcctctgcat aggttgtggt gcctgtgagc tgtcgtgcc gaccaacgca 1740

 atcaagctga tagacaagag gcctaacctc gttcaggaga agtgtatccg ctgcgcacc 1800
 tgctataata ggtgtccgcg cgcctcacag attctgtcca tgggtggtgc gagatgaatg 1860
 atgagcgttt cagaaaatct tttgggaaac gtctttgaa tttatcttgc gcgggcaacc 1920
 gatgaggaaa tactcaaaag aaaggttgc agcggcggtg cggttacagc cctcttagcc 1980
 tacgccctgg agaagggcct catagatggc gttgtaacgg ccaaaaggac agaggggttg 2040
 gagggtcagg ctgtagttgc gaggacaagg gaggagctcc ttgaaactgc cggaaacaag 2100
 tggagcatag tgccttcgc ctccaggatg aaggccaaga tagaggagga agacctaaag 2160

 aacgttgccg tggctgctc cccctgccag gccagttct tcggccagat gagggacttc 2220
 ccactctgg aaagcgattt cggagagagg ataaagtaca tcgttagtct cttctgcata 2280
 ggaacattcg cattcgaggc attctcaac tacctcagga tgaagcacgg cataatggcc 2340
 caggatatca aggacatagt ccttaagggg gacttctcag agatatacca cggcgattca 2400
 gtgctctcac tggcgataaa agaggtttac tcatactcc aagccggtg tctggtctgt 2460
 actgactaca ccggaacctg gagcgacatc tcggccggct tcgtggagag cgagagggga 2520
 tggactgtcc tcataacgcg caaccttaag gcagaagagc tcgttaagag cgccgagaag 2580

 gacggataca tagagctcgc cgacggctcc cacgtgatgg gagaagtctt caaagcggcc 2640
 agggaaaagc ttgcgagagc gcagaagaac atgatgtatc tgctctgatt gataaagaaa 2700
 gtgaaaattc tcaagtggca agatgggtc gttcccaccg aggactatat ctgcgtcgag 2760
 gagacctttg aaatcttcgc agtacacgaa aaggacgaag agtttctcgc cgaacttct 2820
 gcttcacca accagctgaa ggaacttga gccggattcg tcgtgtcgg aggctatgaa 2880
 agaccggagg acatagtga cgtatgggtt gagggcaagg agatttacgt gaagttgaag 2940

gataacccccg ccacgggcga gctggttgtg aaacacaccc cctgcggcga ccctacaga 3000

 atgaaggagg gcagaattct cagcaggaag ggcgaggaag tcaaataac ccccggcctc 3060
 gtattgaaga tatcctccac gatgacaacg ctggctgaga cgtggagaaa gacagggggc 3120
 acccactggg cggccctctt cgatttgaac gccaatgtcg ttgccttcag cgaggacata 3180
 ggcaggcaca acgccgtcga taaggtcgtg ggatacggcg tectcaacgg actcgacctt 3240
 gaaaggctta tectggcatc gagcggcagg atgccatagc gcatgtaag aaaggcagtc 3300
 aacgcgggca ttccagtagt ggtgacgaaa tcaccgccga cggacaaggg cgtggagctc 3360
 gccaggggagc acggggtaac cctaataggc ttcgcgaggg gaaggcgctt caacgtgtac 3420

 tccggggagc atcgattatt gttctaa 3447
 <210> 18
 <211> 8730
 <212> DNA
 <213> Thermococcus onnurineus NA1
 <400> 18

 atggaagagc tttttattct ttccttttca attccgctgg ttggaggcct tttactgttc 60
 aaactcgacg gtaaaagagc ggattacttc atgctcatca ctgtcatcct tgccacaata 120
 ctcaatctcg cgggagttta tgagttctat tctctggga tgcctactat acacaaaact 180
 ctcgtgagct ccacaacct cggtaggttc tacggtctct tgatagacc aatgagtgtg 240
 tgcgttggcc tggttgtgat aacagccgga ctgcttttta tgatctatgc aaaagactat 300

 atgagcccgg agaacaaga gcaccagtc tatgaggata agggtaggtt ttacgcttgg 360
 atggtgctct ttatcggagc aacactcgca ttcatttact cctcgtcggg tctccagctg 420
 ctgatattct tcgaaattat gagcctcgcc tgttggggtg tagcaggcta ttatggaagt 480
 aaaaaggcca aaagggcagc atataaagcg ttgcttgta ccaactttgg agcggtgata 540
 ggtctataca ccgcggtagg tataggaatc acacacctc acgacctgag catatttgca 600
 tattctggcc tcaatgacag ccttaaacctc gtcgttttta tcggggtcat gatagctgct 660
 tttaccaaga gtgccagtt tccgctttat tcatggctcc cagatgcaat ggtcgcccca 720

 acacctgctt cagccttctt ccacggtgca gcatggttg agatgggtgt ctacctgctg 780
 gccagattca tccagttcat gaatccata cccaaagaag gattttacgt catggcagcc 840
 ctcatcatcg ccaactcagat aatctgcatc ctgatgtatc ctctccagaa gagcgcgaag 900
 aggctgctcg cctactcaac aatagcagag tctggactga tgtacgtggc ccttgcgacg 960
 gcagtctctg ggttgcaggg gggacttcag gcttcaatgt tccagctatt caacatgcc 1020

tacatcaaag gtctcgcttt cctgacggct ggaacgttca gctacgcctt gggaaccctt	1080
gagatggaca ggataaaggg cctcattaaa tcccctgtgg tcggctacag ctggaccttt	1140
gcccctgcttg gtttagctgg cgttccacca ttccggcgtgt tcttcggaaa getgggaata	1200
ctcagcaatg ccaaggcaat ggaagagagc gtccctcatca ttgccatggt tgttttactt	1260
ctcctcgact cagcagtgtt cctcatgggtg tctctgaaga ggatacacgg catggcttcc	1320
agcgagggcg gagaagaagt cgagattaca ccaactgatga aggcctgtgat ggttatcctg	1380
ctcgtcctgg ccatgctggc cccgtacata gcgtatccac tgatcgtcaa agtgggggtgg	1440
tgaatgttcg acgtaacgt cactctttca ctcgatagaa ctgcagtgtt ctctgtactc	1500
aacgtcgcga tactcggcat agcggcgcta gttgcatcgt tcagatacat gaggatatac	1560
gagttcaaac caaagatacc ctactacca acgctcgcca tcttcatagt ctcgatgctt	1620
ctcatcccaa tgggtccagga ctggctcagt ttctcttcc tctgggagat aatgactctc	1680
gcctcatact tctgataat ctacgactgg ccggaggaga gcgtcaagaa ggccggctgg	1740
aagtacttcg tgaccatgca cctcttcgac acatctcccc tcatgctggc agtgactatg	1800
tactacgctt tccatggaac gttcaacttt ggggccataa cggagtacag caatgccata	1860
gtcgtctctt ttctcctggg atttgcggcc aaggccggcc tcttcccgt ccaactctgg	1920
ttaccggacg cccaccggc cgcaccgagc ccggtctcag ccttgatgag tgggtccatg	1980
gtcgaactcg gcctctacgg aaccatcagg gttctcaatg ctgtgggatg gagcgtcgca	2040
acctggatag tctatctcat cggcgtatg gcagtgtca gcatgctggc tgccatattc	2100
agctacgccc tccaggacga cgtcaagagg ctcttcgcat ggtccacat cgacaacatg	2160
ggctggatgt atctgctcat cctggcaggc ctctttggcg tttcaggagt ggagaagggc	2220
gttgactact acgtcgtggc acatggactc gcgaaggcgg cggcgttcat atcaacgggt	2280
gcccctctct acgttttcgg tacgaggagc ctgaagaagg caaaaggat gatgaacact	2340
gacagcctca ccgcgggact catgatggcc tcaatattcg ccctcgaagg tgttccgccc	2400
ttcaacctct tcatgaacaa gctcaatgtg ataaagactc tctgacggt cagccccggc	2460
ctggcatact tcacggcctt tgagtggttg atagcgttca tactgttctt cagagtggtt	2520
cacgcctaca tctcagcga aggtgaacca gaggccaaga gaaagcttgc tggaagcata	2580
gcccctctcc tgatagtctt gctcactctc tccatggtaa gccagtctgt ctgcgactac	2640
atatgggtga ggigtgaaat ggagggactc tttacgctcg ccgtcatcct gtatttctc	2700
tccataccg cagcgttagc cctcaaaagg agcttcaagg cttcaatcag cattggccac	2760

atactcacgg ctctagcctc catagctctg ttagcattta cctttgtgtc aataccagat 2820

 atcctcagcg ggaaggccat agaattcaca tatgacttag gagtggcca gattccgttc 2880
 cagattgatg ggcctctcgt gataatgtgc ttcactctcg gcgcctcgg acttgcagcg 2940
 tcaatatatt ccccagata catggcaatc tacgagaagt caggcagagg ctggatgtac 3000
 ataaccatat attcagtgtt tatgctctcc atgatactca tagtaacaat agccaacatg 3060
 ttctggttca ttttctctcg ggaggtcatg acgtttacat cgtacctct gacgatctgg 3120
 gaaagcgaca aagaggatgt cagaaaagcc ggctggaagt acttcgtgac catgcacata 3180
 gtgagcacac tgccactgat aatcgccctc gccctgctgt atgcagacgt tagctcaatc 3240

 gagggactta actttgagag tctagcggcc ttaaaactaa gcccagtatt ctacgcctc 3300
 ttctgatag gctttggaag caagtcaggc gttgtcccgc tgcacttctg ggccccggag 3360
 gcctatacgg tcgccccgag caacgtctcc gctctaattg ctggagcact ggagaaggtc 3420
 gcggtctatg cctgataag gactacatgc tttatcatga agccaaacga gactttcgga 3480
 tatgcagttg cctgcttgg aacagtaacc ctgacagttg gaacctcta cgcgttgaag 3540
 cagaccgatg ccaagagact tttggcctac cacagtatcg gccagatagg ctacatctgg 3600
 ctcgcatgg gcgttgggat agttttcata gccaggggag atatgtactc agccttcgga 3660

 gccatagccc tagcatcaag tctgtaccac ctcgttaacc acacgttctt caagggactg 3720
 ctcttctctg cgacgggctc aatattctac agaaccgca gcagggatct caaccagctg 3780
 agaggtctgg ctaaactgat gccctttacg gcgctcttca cattcatagc cgcaatgtca 3840
 atagctggaa ctctccgtt caacggcttt atgagcaagt ggatgattta tcagtcaacg 3900
 ttctctcgg gcaacggcct gatagtgttc tttggagtga tggcctctt cataagcgca 3960
 gcaacgtgg cttcattcat caagttctac acaaccgeat ttggaggaga acctactgag 4020
 tttacgaagg atgctgagga agttccatcc cctatgetca tcgccaaggg cttctctgct 4080

 tcactctgca tctctcttgg actggttcca agcctcatcc tgccgatact gctttcgcca 4140
 ggggcagccc tagccggtat agatgtctca ggactgatgg acacaaacta ctggcttgtc 4200
 acgattaaag ctccgcttat gccgacaggg gcagagagct acttcaaacc gctactcttt 4260
 gcgacactct tcggcgtgat cttctctggc atgtacctgc tcttcccaat ctcaaagaaa 4320
 acctacagac cctggacct cggtgagccc gtggcgatgg agcactacaa gttcaaggcg 4380
 ataaactact acgaaccctt cgaggagtac atccaccgc tctaccacac cggccacgtt 4440
 ctcagcgagt tcggatctgc cctgattggc gcagtcgcca atgcgtacgt ctcaacaaca 4500

agggctctcc acagagtatg cgattctata agcaagagtg tggccgggat cggaaaagag 4560
 tacgagaaga agtccccga agtctacctt gacgaatact tccttgcccc actggtcaag 4620
 atagtgaggg tctcaggagt gcttctagat gagggattca tgaggccaaa tgcagcgttc 4680
 acaatagccc tggtaactct ggcggttata cttgccctga tggtgctgtg aatgacgctc 4740
 gaaaaaattg cattcgcggc cctttcactg atgataatca tcctccttc gccctcctc 4800
 gacggaataa gcagaaagat caagctacc gtccaagaga ggcaggggcc ccccgcttc 4860
 cagacctact atgacctctc aagcctgctc tcaatggagc cgatccttc aacggacaga 4920

 ctgggcttcc tcatagctcc ctatgtggcc tttgcttcag cagtctcagc cgcctgctc 4980
 ctccccttcg ggaacttcgt cccagtggcc ttcacagggg acatcttcgt cttcctctac 5040
 gtgctggcga tattctcgat atcgatgatg atggcaggct tcctcgtgaa caacacctac 5100
 tcaaacgagg gtccaacag ggagatgatg ctcatcctca gcgtggagcc gatactggga 5160
 atagcgatag gcatactcgc gcttaagacc cactcgtca gcgtgagcgg aattccactc 5220
 aacctcagcc tcacacctc cgttgctctc gctttcatct tcctcgcta cgcctctat 5280
 actgagtgcg cttcataacc cttcgacata gccgaggccg aaacggagat acttgagggt 5340

 ccactcgtcg agtacagcgg gaagctgctt ggaatcttca agtgggcat gctgataaag 5400
 cgcgtagccc tgatatggct gttcgcgagc ttcatagtca ttccagtcac gaaggggttc 5460
 gtcgacatca cgacgccta cggtgggtgca gtaacgctcg cggcacagct ggtactcctg 5520
 gtggtcttct acgtcatgct ggccatcata gagtcaacga cggcccgtat gaagtaatc 5580
 caggccatca ggcagaacac ggtgatattc cttgcccggaa tagtcgcgct ggtgatagct 5640
 tccttgggat ggtgaatgct tgaagttatc aagtttaacg aggcctctgaa aaagaagcgc 5700
 gtacacaggg gagatgaaaa agccaaagta acgcgggagt acttggatga gattatcgag 5760

 aagttcgggg agaagataag ggacgtcaag caggccgctt acaaccagtg gattataacc 5820
 gtcgagaggg aagaccttc ggagatagtc ctctacttc tcaaccacc ggagtggaag 5880
 gagaccacgc tctcatcgat ggtggccacc gacgagaggc ccctaacgg caagttcagc 5940
 atcacctact ggctcagcgt taacggaaag gcgggtgact tctatctcgg cgtcagggtc 6000
 tacctgccgg aggacgacc gaggttcacc tcgatagcgg ccaagcacag gggcgcaac 6060
 tggtagcaga ggaagccat ggagatgctc ggcctcactg ccgaaggcca cccgaccg 6120
 agggcgctcg tccttcgga cgactggccg tctgctct acccgctcag gaaggacttc 6180

 cactactcga acagcccgc gggggagaag ttctaccct acaaggaacc gaagaaggac 6240
 gagatagtcg tcacctagg accgtatcat gtggccttg aagaggcagc aacttcagg 6300
 ctctatgta agggagaac cataacagac gttgactatc gcggcttcta cccccagg 6360

ggcatagaga agatatccga gggaaggcta acctacgacc aggtctgctt catagcggag 6420
agaatatgtg gaatctgctg ctgcacacac tccacagcct actgccaggc ggttgagaac 6480
gccggaggta tagaggttcc cgagagagcc gagtacatca ggacgatagt cctcgagata 6540
gagagactcc acagccacct gctcaacttt ggaatagtct cccacctcgt tggctacgac 6600

tacggcttca tgaagcctg gaggataagg gaggcacgtga tgtggctcgc gaaaggcta 6660
acgggcaaca gaaagaccta cggaatgctc cttgtcggcg gcgttaggag agaccttctg 6720
gagtacagaa aatccctgat agaagacgtc ctcaagaaga taaagaccga gttcagtgag 6780
ctcgtcgatg aggcaatctc aacgagcacc ttcgtgaagc gccttgaagg cgttggggtt 6840
ctgcctaca aggtcgccaa ggagtgggac gttgatggac cccttggcag gggctccgga 6900
agggacttct acgtgagaag ggaccaccg tacgcgcct acaagtacct cgacttcaag 6960
gtcccagtct acaaggaggg tgacgttctg gcaagggcc tcgtcagaat agaggaagtt 7020

ttcgagagca tctggataat agagcaggcc ctcgaccaga tgcccggagg agacattctg 7080
gcggagtaca aggagatacc cccgtactcg gaagcgatag gcatgactga ggcaccgagg 7140
ggcgagaaca ttactactg catgaccggc gagaacaaca aggtctacag gtacagggcc 7200
agggcggcaa cctacaaca cctgcccgtc gttcccgaca tgatgcgagg ctacaccata 7260
gccgacgcc cgctcatagt ggcgagcata gaccctgct actcctgtac ggagagagtt 7320
caggtagtcg acgtcgaag cggaagggtt agggttctca gcgagacgga gttcaacaag 7380
ctctccataa aggcctcaag gagggtctga atggccgtga cgctgaagta cccttctgtg 7440

aagcttgaag ccctccgga gtacagagga attccacaga tagacgcgac cctctgcata 7500
ggctgcggtg cctgcgttaa cgcctgtccg ccagatgcac tcctcaggat agacgactac 7560
aacagaggag ttagagaaat tgtcctcgat gtgggaaggt gcatccgctg tgetcgetgt 7620
gaggaggtct gtcccaccgg agcgatcaag ctcacgaacc tcttcgaggc cgcttcgccc 7680
gacaggatgg accacgtgga ggttgttagg ctcaggctcg tgaatgcaa aaactgcggc 7740
aggtacgcc acttactga gaggcaggtg agaaaggccc tccagattct ccccaggag 7800
atcatcgaag gggaagctct ggaagagaag gtctggatct gcagggactg caggaggaaa 7860

gggacagttg atggaacat agaagccagc aaggaggtgg ttctatgaat gagcggaaag 7920
ccgaagctcc gctccatag ggtcttccac ctcaacaccg gctcgtgcaa cggctgtgac 7980
atcgagataa tcgacgtct cacaccgttc tacgacgtcg agcgctttgg aatcaagcta 8040
gttggctcgc cgagacacgc tcatgcactc ctctctcgg gtcgctcac gagacaggcc 8100
tactacggcg ccaagagac cataaaggcg atgcctccgg agccaagggt aatagtcgcc 8160
atcggaacgt gcacctgtag cggagggata ttctacaacg gctatccagt ctacagaagg 8220

cccgagagcg gtagggaggg aagcgagtat ccacggaggg gaggtatagc ggagctcatc 8280

gctgacttga gggacgaggg cgagaaggtc ggtccgtca tctacatccc cggctgtcca 8340

ccgagaccgg aggagataat ctacggcata gcacagctcg tgggactcgt cgagaagaag 8400

ctcagctatc aggagtacag cgacgagctg gttcccttca agctcccaga ggggccgctg 8460

gaggagcgca tcaggctgac ccttatggag aggtcaggc acctcgtggg atacctcgac 8520

agggaaaaga tectcgagga tttcatgggg ctctgtaaag aggccgagaa gagcgagaat 8580

cccagggagg agctggccag gctcgtcaag gactacccg ccaaatgcgg ggacgttaga 8640

ctgggcttct gtatgatgct tctcgaaga gactactgga gggtaaaga tgccttggat 8700

gctggtaaag agttcgtata ttgggtttaa 8730

<210> 19

<211> 4179

<212> DNA

<213> Thermococcus onnurineus NA1

<400> 19

atgtttggct actgggacgc tctttacttt gtctttatct ttattatcgg cctcatcata 60

gcatggatgt tgaacgaatg ggccaagaag tctggaatgg gtacgagaga agctggtgat 120

ggcacgaaag tttcatcag tggtaggac ccggacaagg taatccccgg cttcgagcat 180

tacgaaggtt actacactgg caagaacgtc atgtggggtc tcacatacgc cctcaaagg 240

ttcttcgcc tectcaggaa cgagcacaca ggtctgctca ccgattacgt aagctacctg 300

ctcataacga cggctttcgt gctcggagtg atactgattt gggggtgaat gagcatcaag 360

gttcccgtc accagaatag aacaaacgga accacagtg agcgcgagat gctggagaag 420

agaatagccc agttgtgcc cttcatagga agatcacct gggatattca cgtaaacagt 480

ggaagctgca acggctgtga catcgaatc atagccgcc tgaccccgc ctatgatgcc 540

gagcgtttg gcgtcaagt cgtgggaagc ccaggcacg ctgatgttct cttagtaaca 600

gggccagtca cagaccagag cttgaaagg gtcaagcttg tctacgagca gacaccagac 660

ccgaagatag tcatagcagt gggatcgtgc cccactggcg gtagcgtgtt ctatgagagc 720

ccattacca atgcaccgt gagcaacatc attccggttg acgtctacgt gccagctgt 780

ccaccaaggc ccgaggccat actctacggc gtcgttttgg cccttgaaaa gctggctaaa 840

atcctgaaag gcgaagtcc ggagggtgaa gactgaatgg ctgatgataa cagaatcatg 900

gagaatgttg ataagttag agaaccaacc aaggaagata ctgtcgctga gaccataaag 960

agccgtttcc ccaatgctca cgtggagata cgggagaaca agtgggaag aaagcgcgtc 1020

tgggtgatcg ttccacgaga agactacaaa gcgctcatga agttcctcct tgaactcgac 1080
 ccagaggccc actattcgat aggaatagag caggactacg gggaagagat aggctatatg 1140

 agccacatcc tgetgcaacta cgacaatgct ccagcagtct cactgctcgt tgatgttaga 1200
 gtacccaaag acgatccagt aattcccgat atcagcgaca tcttcccgat agcactccag 1260
 tacgaaaggg aagccgctga gatgatgggc atagtcttcg aaggtatccc cgacagtaga 1320
 aggcttttcc ttccggacga ctteccagag ggtatctacc cgctcagact cgacgaaaaa 1380
 ggcataccag aagagattgt caagaacgcc ggacaccgt actacctgaa ggggggagat 1440
 aatgtaatga ccaagaaggt cgagtactgg ataaagatac cgttcggccc aattcatccc 1500
 ggcttagagg aacctgagaa gttcatactt acgctcgatg gcgaaaggat agtcaacgtt 1560

 gatgttaagc ttggctacaa tctacgtggc ctgcagtgga tagcatacag gagaaattac 1620
 gtccagataa tgtacctcgc ggagaggata tgtggtatct gttcgttctc ccacaaccac 1680
 acctacacca gagccgttga ggaagcggcc ggaatagaag tgccagagag ggctgagtac 1740
 atccgtgcca taataggcga gctcgagagg gttcactccc acctgcttaa ccttggtgtc 1800
 ctgggccacg acatagceta cgacacggtc cttcacctca catggctggc acgcgagagg 1860
 gtcattgatg ttcttgaagc catctcaggg aaccgcgtga actactcgat ggtaaccata 1920
 ggccgtgtga gaagagacat cgatgaaaaa gggaagcggc tcattcttga tatgataaag 1980

 tactacagga gcataatgcc tcagatagaa gaggttttcc tccacgacce aacctagaa 2040
 gcccgttga gggactgtgc ggtgataagc aagcgcgtcg cccttgagca ggggtgcagtg 2100
 ggaccgactg ccagagcttc cggctctaaag gtcgatgcca ggtggagtga gaggcttggg 2160
 gtttacctg acctaggagt taagccagtg atgccacagg acgttacggg agaaaaaccg 2220
 cacggtgatg tattcgacag ggcagccgta agaataggag aaatatacca gacctcgac 2280
 atgctcgaac acgcactaga ccagatgcca gagggtaaga taaagacatt cccaaaggac 2340
 aacatcttgg ttccaagct caagattatg gttgacggag agggaatcgg aaggtacgag 2400

 gctccacgtg gcgagctggt acactatgtt cgcggaaaga aaggctccga taaaccgctc 2460
 cgctggaaac caaggagacc aactttcccg aacctcttcg cagttgccaa ggtgtgaca 2520
 ggtgatcagg tggcagactt cgtgctggca gtggcctcga tagatccgtg cctgagctgt 2580
 acagacaggg ttgccgtagt acaggatgga aagaagagaa ttcttactga aaccgacctg 2640
 ctgagactct caataaagaa gacacgcgag ataaaccccg aagttaaagg cgaccaaca 2700
 ccggtcggct tcggctgctc gaggtgaatg gacgtaatgg cgaacatcat ttatccgta 2760

gcaggtttaa taggccttta cgctttcgtc tcaactggcat cgctcgtctg ggaaggtata 2820

gacagaaagc tcgtcgcaag gatgcagaga agggtaggac cgccgcttct ccagcccctc 2880

tatgacttct tcaagctagc gagcaaggaa acaataatcc ccaacacggc taactttatg 2940

tttagagccg cacctgtact cgccctggca acggccatag cactcctcgc ttacaccccg 3000

atgggctttg ctccactact cgcgagcaag ggagacgtca tcgttttcat atatctctc 3060

accctcattg gcttcttcaa gatactcggg ggcataagct caggaagccc ctacgcaaag 3120

ataggagctg caaggaagc agcaataatg gtttccagag agcctgcat gatgctggcc 3180

ctattcgcta taatatggcg tcttggaaaa ctggagtc acaagccatt cagcatggag 3240

gtctttacc agtacaacat ttgggaaata ggtacccgc tcagccttat aggtgccgta 3300

atcctccttt acgtcttctg catttggctg gcaagtgaaa tagaagtcgg atatttcaac 3360

ataccgatg cagaggagga gatagccgag ggactgctcg ccgagtacag cgggcgctac 3420

ctggccctgt taaagctcac gaagcactg aaaacttaca tagcagcatc gctcgtcgt 3480

gcaatattct tcccctgggg aatagcagat tacttcaacc tcaccggact tccagcaaac 3540

gtcgtaaacc tgccttcca tacaactcaag gtattcatac tgccttttgc tgtgcagagt 3600

gtcttcaggg ccaactacag cagactcaaa ataacgcagg cggttgactt cctctggaag 3660

aacgtcttct tagcttcgct catcggcaca ctcttatcg ccatggaggt gataatgtga 3720

gtgaggctct cccactcat ccccaccgig ctcagaaaca tgttcaaaaa gcctgccacc 3780

aacctcttct ctgcgactga accagtgccg gttccagata acttcagggg ccagctgaag 3840

tacaacgtgg acaagtgtgt cggctgcagg atgtgcgtca cagtctgtcc agccggcgtc 3900

ttcgtcttcc tacttgagat aaggaaggic gctctgtgga ccgctagatg tgtctactgc 3960

tcgcagtggt ttgacgtttg tccgaccgca gccctccaga tgagcgatga gttcctgctg 4020

gcaagctaca acaactacga cgacaagttc atcccgtca agcccgaaaa ggttgaagag 4080

ataaagaaga aactggagga gcaaaagaaa gcgaaggcag ctgcagctgc caagaaagcc 4140

atggagaaaa aggaagcagg gaaagaggcc aaaaagtga 4179

<210> 20

<211> 2568

<212> DNA

<213> Thermococcus onnurineus NA1

<400> 20

atgggtatgg cagaaaaagcg catatcagtg gtgtgtccgt ggtgttccgt tggttgtaga 60

ttttacatag taaacgtcaa tggctacca aagaagatcg agttcgacta cgaccacgac 120

atcaggaacc acggaagct ctgtccaaag ggtgtcgag ccttcagca tctcagcat 180
 ccagacaggc ttaaaaagcc ccttaagagg gttggcgaga ggggtgaagg caagtcaag 240

 gaaataagct ggaagaggc tattaaggaa atcgcacaga agctcagtga aatcaaggag 300
 aagtatggtt cggaggctct tgcttttctc ggaagtgaaa ggtgtccat agaggagaac 360
 tacgttcttc agaagctggc aagggtttg ggaaccaaca acattgaata tgtgttagg 420
 atgtgtcagt caacggctgt tgcaggtaag gggatggttc ttggacaccc cggctgacg 480
 aacccttcg aggacattct taaggccaaa gttatcgtcc tttggggata caatccagcc 540
 gcaactaatc cggcttcttc cggccagtac attgagaagg caattctcga caacaacgcc 600
 accctcattg tcgttgacc aagaaaaacg aagactgcca agtacgcaga cataacctg 660

 cagccatate ccggaaccga ccttgccatt gcgttgcta tgctcaactg cataatcacc 720
 gaggagctct acgataagga ctctgtggcg gagcgcgagg agggccttga ggagctgcc 780
 aagaccgtcg aaaagtacac tccagaatgg gctgagaagg tcagcggcgt tctgcccag 840
 ctcataagga aggccgcaat cacctttgca acggctggaa ctgccccct gctgacgaac 900
 gaggagtgta accgacgcg caacggaacg aggactgtta tggtatcac tgagatgatg 960
 gttctctgcg gctacttcgg aaaggagggc gtcattctg gagctatacc cggtgcccac 1020
 aacggtatgg gccttgtct aatgggtatt ggaccacacg aactgccagg aagattcccg 1080

 ctccacgccg aggagcaca gaggagaatt gaggaggcat ggggcttcaa gatcccagag 1140
 aagcttggaa tcaactactg tgaatgatt gatgcaatcc ttgagggcaa gctcaaggcc 1200
 ctctacgtea tgggaaccaa ccctgccaag gcccttcga acctcaagaa ggctgaggag 1260
 gcctttaaga acatcgagtt cctcgtcgtc caggatatct tccttactga gaccgcgaaa 1320
 tacgccgaca tagtcttcc agcgctgca tggtttgaga aggacggaac cgccataagc 1380
 ttcgagagaa gggttcagag gagctttaag gctgctgacg caccgggaga ggccaagcct 1440
 gactgggaaa tccttgttat gctcgtctaa gagctcggct ttggagagta cttcaactac 1500

 tctgatgcag acgacatcct gagagaaata aacagaatca ttccgccct tgctggcgcg 1560
 acaccgaga ggctcaagaa gaacctcaa ggctgtatga tacctgccc agacgagaac 1620
 actgaggttc cgaggctctt tgtccagggc ttctcagc caaacgaaa ggcccagctt 1680
 attctgtgg agtataaaga gcctggagaa gtccccgatg aggagtacc gttctggctc 1740
 accaactaca ggttcgttgg ccacttcac accggaacca tgagccacag gagcaagagc 1800
 ctgagcaaga ggtggccaga ggagtacatt gagatcaacg agaacgcgc gaagaggctc 1860

ggcataaagg acggcgacct cgtgagagtc gagaccagga gggcagcgct ggttctcagg 1920

 gccaaagtta caccgcatat cagggagggc gtcgttgcg cgccgtggca ctgggacttc 1980
 aactacctga ccacggacgt cctcgacgaa tacgccaaga tgccggagt gaagacggcc 2040
 gcgtgtagga tctccaaggt tgaggggtga atgagcaaaa agatatttat cgattttaag 2100
 cgctgcattg cctgtaagc ctgtgaagtc gcctgtgaaa tggagcacgg ggaagcgagg 2160
 attagggttt ttgagttccc cgatctgacc agcgtcgcct tcaactgccg ccaactgtgaa 2220
 aaggctccat gtatggaagt gtgtccagtt aacgcgctct ccaaggacga tgatggcgca 2280
 gtcgttctcg atcccctcaa gtgtatcggc tgtctcatgt gcggtctggc ctgtccattc 2340

 ggcatlccaa agatagacga gtacaacaag ataatggaca agtgcgacct ctgtgccac 2400
 aggagagccg aaggaaagct tctgcctgt gtctcagcgt gcccaactga ggcctcaag 2460
 tacggcgaca taaacgatgt cctctgggcc agagaaggaa agatagtcgc cgagcttaag 2520
 gacatcggcg acaggaccaa cgtcctcgag gcctacctca tcagatga 2568

 <210> 21
 <211> 1893
 <212> DNA
 <213> Thermococcus onnurineus NA1
 <400> 21

 atggccggaa agaaggttcc ctcaaagcaa gtctccataa ctccaggtgt tggaaagctt 60
 attgagaaag ccgaggagga tggggtcaag actgcctggc acagattttt ggagcagcag 120

 cctcagtgtg gattcggctc cttaggtgtc tgctgtaaga actgtacaat gggacatgt 180
 agaatcgate cgtttgggtg tggcccaact aaggagttt gtggtgcgga tgcagataca 240
 atagtagcaa ggaacattgt aagaatgata gcggctggta ctgccggtca cagcgatcac 300
 tcaagagatg tagtccatgt attcaagggc attgctgaag gaaagttaa ggactataaa 360
 ctaacagatg ttgaaaagct caaagagctg gctaagattc tgggtgtcga aacagagggc 420
 aagagcgaaa atgaaattgc attggaagtc gccacattc ttgagatgga gttcggaaaa 480
 caggatgagg agccagtaag attacttgca gcaacagcac caaagaagag gattaaggtc 540

 tgggagaagc taggagtctt accaagagcc atcgacaggg agatatgtct cagtatgcac 600
 agaaccaca taggctgtga tgcagaccct gcaagccttc tactgcatgg tgtgaggact 660
 gccctggccg acggctgggt cggctcaatg atggccactt atctgagcga cattctcttt 720
 ggaacaccaa agccgataaa gtcgctggcg aacctgggag tcttgaagga agacatggtc 780
 aacataatcg ttcacggcca caaccgatt ctctccatga aaatagcaga gattgcccag 840

agtgaagaga tgcagaagct tgcagagcag tacggagcaa agggaattaa cgttgctgga 900
 atgtgctgta ccggaacga agttctctca agaatgggag ttcaggctgc tggaaacttc 960

 ctaatgcaag agctggcgat tataactggt gcagttgagg ccgtgatagt tgactaccag 1020
 tgcctaatgc cctcattagt tgatgtcgct tcatgttacc aactaagat aataactact 1080
 gagccaaagg ctcgattcc gggagcaata cacgtcgaat ttgaacctga gaaagcggac 1140
 gagatcgcca aagagatcat caagattgca attgagaact ataagaacag agttccggca 1200
 aaagtctaca ttccagagca caagatggaa ttggttgctg gatttagtgt cgaggcaata 1260
 cttgaagccc ttggtggaac actggagccc ctcataaaag ccctccagga cggaacaata 1320
 aagggaatcg tcggaatcgt tggatgtaac aatccaaggg tcaagcagaa ctacggtcac 1380

 gtcaccttgg ccaaggagct catcaagagg gacatcctgg ttgttgaac tggttgctgg 1440
 ggaattgctg cagcaatgca tggattacta acccccgaag cagctgaaat ggccggtcca 1500
 gggctgaagg caglatgca agcgcctgga attccaccat gcctgcacat gggaagctgt 1560
 gttgactgtt cgagaatcct gctggtcttg agtgccttg ccaatgctct gaatgtgac 1620
 atttcagact tgccagtgc tggctctgct ccagaatgga tgagcgagaa ggcagtggca 1680
 ataggaacct acttcgttc aagcggcgtc ttcacgcact tgggagttat cccaccagtc 1740
 cttggaagcc agaaggttac caaactcctt acggatgaca tcgaggatct cttggaggg 1800

 aagttctacg ttgagacaga tccagtgaaa gcggcagaaa caatatacaa cgtgataatt 1860
 gagaagagga aaaaacttgg atggcccatc taa 1893

 <210> 22
 <211> 2637
 <212> DNA
 <213> Thermococcus onnurineus NA1
 <400> 22

 atgagtgaaa ggctcgtccc cgtggtctgc ccctactgtg gtgtagggtg caggctatac 60
 atcaggagtg ttgatggcta tcccgtagc atagaatagc ccaaggacat ccccaacatc 120
 tcaaacgaac tcggaagct ctgccctaaa ggcaacccg tcgttgagta cctcctcgca 180
 aaggacagc tcaagagacc cctcaaggcc aaggaacagg gcaagttcgt tgagataagc 240

 tggagcgagg caataaagga ggttccgag aggctcaagg cttatgcaa ggacgacccg 300
 aaccagctca tgttcttcgg ctctcgaga acattcaacg agcccaacta cctcgtccag 360
 aagctggcca ggatgctcgg caccaacaac gttgatcact gtgcaagget ctgtcatgca 420
 ccgaccgtca cgggtctcaa ggctgttttc ggtgccggcg caatgaccaa cacctacaag 480

gacattgaag aggcaaacgt catcttcatt atcggccaca actacgctga gacccacccg 540
gttggcttcc gctacgtcct taaggccaag gagagggcg ctaaggtcat agtcgctgac 600
ccgaggttca ccaggacggc ctggttcgcc gacatatcc tacagcacta cccgggaagc 660

gacattgcgc tgataaacgg tctcatccac gtcatcatca aggagcggct ctacgacgag 720
aagttcgtga gggagagatg cgttggcttc gatgaagttg tggcagccgt cgagaagttc 780
acacccgagt tcgtcgagaa ggtaaccggt gtcccggcgg aactcatcat tgaagctgca 840
agaaccttcg cgaccgcagg aaagggtgtc ataacctggg ccatgggtct aaccagcac 900
accacggaa ctgaaaacgt caagctcctt ggaacgctcg cggccatttg cggttatcag 960
ggcaaagaag gtgccggctg ttcccctaatg cgcggtcaga acaacgttca gggagcatgt 1020
gacatggcag ccctgccgaa cgtcttcca ggctatcagg ccgtcactga tctgaaaag 1080

aggaagtctt ttgaggagt ctgggggtgt gagctgagcg gcgaagttgg actgacaact 1140
gtggaggctg cctacgcggc cgacaagggc aaggtaaagg cctactatgt catgggtgag 1200
aaccagtcga taagcgaggc caacccaac cacgtgatgc acaccctcga gaagctcgag 1260
ttcatggtcg tccaggatat cgtcccagacc ccaactatgg agtatgcaga tatagttctg 1320
ccggcccgcg ccatgctcga gaacgagggt tctctgacca atacagagag gcgcgtgcag 1380
tggagcttcc aggcggtaaa accaccggga gaagcaaggc ccgactggtg gattcttagc 1440
gaggtcggta aggccatcgg ttttgacaag accggatccg gtggattcgt ctacaatgat 1500

gcagccgacg ttctcagga aatcaacgcc tgtactccgc agtatcggg tataactcca 1560
gagaggctca aggagaacct tgcaggactc cactggccgt gccaagcga ggaccatcca 1620
ggaacgaggg tctctacaa ggagaagttc ctactcca gcggaaggc caacctcgcg 1680
gccgttccgg agtacaagg accagtcgaa atgccggacg aagagtatcc gttcctcctt 1740
acgaccaca gatagctcgg aatgtaccac accgcaacca tgaccatgag gagctcgcga 1800
ctcaagaagc gctggccaga acccctcgcc gagatacacc cggatgacgc agtgaagctc 1860
ggaataaaga gtggagactg ggttaaggtc gtcacaagga gaggagcata tccgattaag 1920

gcaaaggtca cccgggctgt caagaagggc gtaatagctg tcccgtggca ctggggagca 1980
aacgtcctca ccaacgatgc cctcgaccg gtagcaaga taccggaac caaagcctgt 2040
gcctgtaatg tcgccaagat cacagaagaa gaggccagga agctcatgga gaaactcca 2100
ccactcatac ccaagattga ggtcgttagg ggtgaaatgg ctagaaagac cgtctttatt 2160
gacttttcaa agtgcatcga gtgccgcgcc tftgaggtag cttgagagcg gaaacacagt 2220
ggaatgtcat tcatcagcgt ctttgagtgg caggaaatgg ccgctatggc cctcaactgc 2280

cgccactgtg agaaggctcc ctgtgttgag gtctgtccaa ccaacgcct ctaccgcgac 2340

aaggatggag cagtcctgct cgctccacag aagtgtatcg gctgtctcat gtgcggcata 2400

gtctgtccct ttggaatacc cgagetgat ctcatacaaca agataatggg caaatgtgac 2460

ctctgcgcc acaggagagc cgaaggaaag cttccagcct gtgttgagac ctgtccaaca 2520

gatgtctca tctacggcga cttcaacgag atagtcaaga agagaaggga gaagtttacg 2580

gagaaaacca tagaactgc caaaactgca gagcgcatcc cgctgacggg ggtgtga 2637

<210> 23

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> CODH/F primer

<400> 23

ggaccatgta gaatcgaycc gtty 24

<210> 24

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> CODH/R primer

<400> 24

ttertttccg gtacagca 18

<210> 25

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Hsp60/F

<400> 25

atggcacagc ttagtgaca g 21

<210> 26

<211> 22

<212> DNA

<213>

> Artificial Sequence

<220><223> Hsp60/R

<400> 26
 caaggatttc ctgggctttc tc 22
 <210> 27
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> mbh/F primer
 <400> 27
 cacgacatag gctacgacac gg 22
 <210> 28
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> mbh/R primer
 <400> 28
 ctggcttaac tcctaggtca gg 22

 <210> 29
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> mbx/F primer
 <400> 29
 gcgattcggg atgataccgg ac 22
 <210> 30
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> mbx/R primer
 <400> 30
 ccatccttcg ccgaagagct cg 22
 <210> 31
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220><223> frh/F primer
 <400> 31
 gtaagctcga cgagtacgac gtg 23

<210> 32
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> frh/R primer
 <400> 32
 gcaccacaac ctatgcagag gcc 23

<210> 33
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> sulf1/F
 <400> 33
 gcagtacgag gaagtcgagg gg 22

<210> 34
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> sulf1/R primer
 <400> 34
 gagggcctcg tcgataaggt cg 22

<210> 35
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> mch/F primer
 <400> 35
 ctaccggacg attggccaga agg 23

<210> 36
 <211> 23

<212> DNA
 <213> Artificial Sequence
 <220><223> mch/R primer
 <400> 36
 ccttatatac tgtgctctct ccg 23
 <210> 37
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> mfh1/F
 <400> 37
 gcgaccggta cggcaacctt cg 22

 <210> 38
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> mfh1/R primer
 <400> 38
 cttgtcagtc atgacgtagt gg 22
 <210> 39
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> mfh2/F
 <400> 39
 gacccgaggt tcacctgat agc 23
 <210> 40
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> mfh2/R
 <400> 40
 gcagacctgg tcgtaggta gcc 23

<210> 41
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 41
 cgttgtcttt gccttgggg caggatata tc 32
 <210> 42
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 42
 ggcaattgct tggactgccg aaaagccaat ggc 33
 <210> 43
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 43
 gaagaaatcg cagagggcgc ctatgactat cag 33

 <210> 44
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 44
 gctcctcgct tactcaagcg ttggacaaat gg 32
 <210> 45
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 45

ggactgctct tcctgtcgac gggctcaata ttc 33

<210> 46

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 46

ggacgcactt aaagtcggcg tagccctttg cc 32

<210> 47

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 47

aattaccac cccaccactc caaaatcca ac 32

<210> 48

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 48

aatggggagg ctgaaactac tgggcaaggc 30

<210> 49

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 49

tggcccaggc gatttccttc accgacagg 29

<210> 50

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> primer
 <400> 50
 aattcaccac cccaccagcg ctattatcag g 31
 <210> 51
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 51
 gagcaccacc tcaccatccc agggaagcta tc 32
 <210> 52
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 52
 gatggccgtg acgctgaagt accccttcgt ga 32

 <210> 53
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 53
 gaacgtagt tttcgacaaa agacg 25
 <210> 54
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 54
 getcaccagc caaaaccgca ccage 25
 <210> 55
 <211> 29
 <212> DNA

<213> Artificial Sequence
 <220><223> primer
 <400> 55
 gcaatgtacc acatattcaa ctgcgatac 29

<210> 56
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 56
 ccgataccga gtttgaatgg aggaatctc 29

<210> 57
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 57
 tcaggccacc cccttgcctt tctgt 25

<210> 58
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 58
 atggagtgca gcgtgtgtgc gggtg 25

<210> 59
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 59
 atgtctgaag ttatcaagtt taacg 25
 <210> 60

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 60

tgaggccttt atggagagct tgttg

25

【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 10

【변경전】

상기 형질전환체를

【변경후】

상기 숙주세포를