



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0090633  
(43) 공개일자 2014년07월17일

(51) 국제특허분류(Int. Cl.)

C12Q 1/68 (2006.01)

(21) 출원번호 10-2014-7013195

(22) 출원일자(국제) 2012년10월18일

심사청구일자 없음

(85) 번역문제출일자 2014년05월16일

(86) 국제출원번호 PCT/GB2012/052579

(87) 국제공개번호 WO 2013/057495

국제공개일자 2013년04월25일

(30) 우선권주장

61/549,998 2011년10월21일 미국(US)

61/599,244 2012년02월15일 미국(US)

(71) 출원인

옥스포드 나노포어 테크놀로지스 리미티드

영국 옥스포드 옥스4 4지아이 옥스포드 사이언스 파크 로버트 로빈슨 애비뉴 4 에드먼드 카트라이트 하우스

(72) 발명자

모이지, 루쓰

영국 옥스4 4지아이 옥스포드 옥스포드셔 옥스포드 사이언스 파크 로버트 로빈슨 애비뉴 4 에드먼드 카트라이트 하우스 옥스포드 나노포어 테크놀로지스 리미티드

헤론, 앤드류 존

영국 옥스4 4지아이 옥스포드 옥스포드셔 옥스포드 사이언스 파크 로버트 로빈슨 애비뉴 4 에드먼드 카트라이트 하우스 옥스포드 나노포어 테크놀로지스 리미티드

(74) 대리인

위혜숙, 양영준

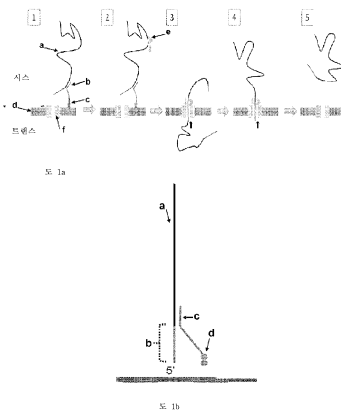
전체 청구항 수 : 총 43 항

(54) 발명의 명칭 포어 및 HEL308 헬리카제를 사용하여 표적 폴리뉴클레오티드를 특성화하는 방법

(57) 요약

본 발명은 표적 폴리뉴클레오티드를 특성화하는 새로운 방법에 관한 것이다. 이러한 방법은 포어(pore) 및 Hel308 헬리카제(helicase) 또는 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터를 사용한다. 헬리카제 또는 분자 모터는 포어를 통한 표적 폴리뉴클레오티드의 이동을 제어한다.

대표도



## 특허청구의 범위

### 청구항 1

(a) 표적 폴리뉴클레오타이드를 막횡단 포어(pore) 및 He1308 헬리카제(helicase)와 접촉시켜, 헬리카제가 포어를 통한 표적 폴리뉴클레오타이드의 이동을 제어하고 표적 뉴클레오타이드 내의 뉴클레오타이드들이 포어와 상호작용하게 하는 단계; 및

(b) 하나 이상의 상호작용 동안 표적 폴리뉴클레오타이드의 하나 이상의 특성을 측정하고, 이에 의해 표적 폴리뉴클레오타이드를 특성화하는 단계

를 포함하는, 표적 폴리뉴클레오타이드를 특성화하는 방법.

### 청구항 2

제1항에 있어서, 하나 이상의 특성이 (i) 표적 폴리뉴클레오타이드의 길이, (ii) 표적 폴리뉴클레오타이드의 신원(identity), (iii) 표적 폴리뉴클레오타이드의 서열, (iv) 표적 폴리뉴클레오타이드의 2차 구조, 및 (v) 표적 폴리뉴클레오타이드가 변형되었는지 또는 그렇지 않은지 여부로부터 선택되는 것인 방법.

### 청구항 3

제2항에 있어서, 표적 폴리뉴클레오타이드가 메틸화에 의해, 산화에 의해, 손상에 의해, 하나 이상의 단백질로, 또는 하나 이상의 표지, 태그(tag) 또는 스페이스(spacer)로 변형되는 것인 방법.

### 청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 표적 폴리뉴클레오타이드의 하나 이상의 특성이 전기 측정 및/또는 광학 측정에 의해 측정되는 것인 방법.

### 청구항 5

제4항에 있어서, 전기 측정이 전류 측정, 임피던스(impedance) 측정, 터널링(tunnelling) 측정, 또는 전계 효과 트랜지스터(FET) 측정인 방법.

### 청구항 6

제1항에 있어서,

(a) 표적 폴리뉴클레오타이드를 막횡단 포어 및 He1308 헬리카제와 접촉시켜, 헬리카제가 포어를 통한 표적 폴리뉴클레오타이드의 이동을 제어하고 표적 뉴클레오타이드 내의 뉴클레오타이드들이 포어와 상호작용하게 하는 단계; 및

(b) 하나 이상의 상호작용 동안 포어를 통과하는 전류를 측정하여 표적 폴리뉴클레오타이드의 하나 이상의 특성을 측정하고, 이에 의해 표적 폴리뉴클레오타이드를 특성화하는 단계

를 포함하는 방법.

### 청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 포어를 가로질러 전압을 인가하여 포어와 헬리카제 사이에 복합체를 형성시키는 단계를 추가로 포함하는 방법.

### 청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 폴리뉴클레오타이드의 적어도 일부분이 이중 가닥인 방법.

### 청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서, 포어가 막횡단 단백질 포어 또는 고체 상태 포어인 방법.

### 청구항 10

제9항에 있어서, 막횡단 단백질 포어가 용혈소, 류코시딘(leukocidin), 미코박테리움 스메그마티스(*Mycobacterium smegmatis*) 포린(porin) A (MspA), 외막 포린 F (OmpF), 외막 포린 G (OmpG), 외막 포스포리파제(phospholipase) A, 네이세리아(*Neisseria*) 오토트랜스포터(autotransporter) 지단백질 (NalP) 및 WZA로부터 선택되는 것인 방법.

#### 청구항 11

제10항에 있어서, 막횡단 단백질이 (a) 8개의 동일한, 서열 2에 제시된 바와 같은 서브유닛으로 형성되거나, 또는 (b) 7개의 서브유닛 중 1개 이상이 전체 서열에 걸친 아미노산 동일성을 기초로 서열 2에 대해 50% 이상의 상동성을 가지며 포어 활성을 보유하는 그의 변이체인 방법.

#### 청구항 12

제10항에 있어서, 막횡단 단백질이 (a) 7개의 동일한, 서열 4에 제시된 바와 같은 서브유닛으로 형성된  $\alpha$ -용혈소, 또는 (b) 7개의 서브유닛 중 1개 이상이 전체 서열에 걸친 아미노산 동일성을 기초로 서열 4에 대해 50% 이상의 상동성을 가지며 포어 활성을 보유하는 그의 변이체인 방법.

#### 청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서, He1308 헬리카제가 아미노산 모티프 Q-X1-X2-G-R-A-G-R (서열 8)을 포함하며, 여기서 X1은 C, M 또는 L이고, X2는 임의의 아미노산 잔기인 방법.

#### 청구항 14

제13항에 있어서, X2가 A, F, M, C, V, L, I, S, T 또는 P인 방법.

#### 청구항 15

제1항 내지 제14항 중 어느 한 항에 있어서, He1308 헬리카제가 표 4 또는 5에 제시된 헬리카제 중 하나 또는 그의 변이체인 방법.

#### 청구항 16

제15항에 있어서, He1308 헬리카제가 (a) 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 어느 하나에 제시된 서열, 또는 (b) 전체 서열에 걸친 아미노산 동일성을 기초로 관련 서열에 대해 40% 이상의 상동성을 가지며 헬리카제 활성을 보유하는 그의 변이체를 포함하는 것인 방법.

#### 청구항 17

제16항에 있어서, He1308 헬리카제가 (a) 서열 10 또는 33에 제시된 서열, 또는 (b) 전체 서열에 걸친 아미노산 동일성을 기초로 서열 10 또는 33에 대해 40% 이상의 상동성을 가지며 헬리카제 활성을 보유하는 그의 변이체를 포함하는 것인 방법.

#### 청구항 18

제17항에 있어서, He1308 헬리카제가 서열 10의 잔기 20 내지 211 또는 20 내지 727에 대해 아미노산 동일성을 기초로 70% 이상의 상동성을 갖는 서열을 포함하는 것인 방법.

#### 청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서, He1308 헬리카제가 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 것인 방법.

#### 청구항 20

제1항 내지 제19항 중 어느 한 항에 있어서, 0.3 M 이상의 염 농도를 사용하여 수행되고, 염이 임의로 KCl인 방법.

#### 청구항 21

제20항에 있어서, 염 농도가 1.0 M 이상인 방법.

#### 청구항 22

포어와 He1308 헬리카제 사이에 복합체를 형성시키고, 이에 의해 표적 폴리뉴클레오티드를 특성화하기 위한 센서를 형성시키는 것을 포함하는, 표적 폴리뉴클레오티드를 특성화하기 위한 센서를 형성시키는 방법.

#### 청구항 23

제22항에 있어서, 복합체가 (a) 포어와 헬리카제를 표적 폴리뉴클레오티드의 존재 하에 접촉시키고 (a) 포어를 가로질러 전위를 인가함으로써 형성되는 것인 방법.

#### 청구항 24

제23항에 있어서, 전위가 전압 전위 또는 화학 전위인 방법.

#### 청구항 25

제22항에 있어서, 복합체가 포어를 헬리카제에 공유적으로 부착시킴으로써 형성되는 것인 방법.

#### 청구항 26

포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하기 위한, He1308 헬리카제의 용도.

#### 청구항 27

(a) 포어 및 (b) He1308 헬리카제를 포함하는, 표적 폴리뉴클레오티드를 특성화하기 위한 키트.

#### 청구항 28

다수의 포어 및 다수의 He1308 헬리카제를 포함하는, 샘플 내의 표적 폴리뉴클레오티드를 특성화하기 위한 분석 장치.

#### 청구항 29

제28항에 있어서,

다수의 포어를 지지할 수 있고, 포어 및 헬리카제를 사용하여 폴리뉴클레오티드 특성화를 수행하도록 작동할 수 있는 센서 기구;

특성화를 수행하기 위해 물질을 보유하기 위한 하나 이상의 저장소;

물질을 하나 이상의 저장소로부터 센서 기구로 제어가능하게 공급하도록 구성된 유체공학 시스템; 및

각각의 샘플을 수용하기 위한 다수의 용기

를 포함하며, 유체공학 시스템은 샘플을 용기로부터 센서 기구로 선택적으로 공급하도록 구성된 것인, 분석 장치.

#### 청구항 30

(a) 표적 폴리뉴클레오티드를 막횡단 포어, 및 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터와 접촉시켜, 분자 모터가 포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하고 표적 뉴클레오티드 내의 뉴클레오티드들이 포어와 상호작용하게 하는 단계; 및

(b) 하나 이상의 상호작용 동안 표적 폴리뉴클레오티드의 하나 이상의 특성을 측정하고, 이에 의해 표적 폴리뉴클레오티드를 특성화하는 단계

를 포함하는, 표적 폴리뉴클레오티드를 특성화하는 방법.

#### 청구항 31

제30항에 있어서, 분자 모터가 헬리카제 또는 Hel308 헬리카제인 방법.

#### 청구항 32

제30항 또는 제31항에 있어서, 제2항 내지 제21항 중 어느 한 항에 정의된 바와 같은 방법.

#### 청구항 33

제30항 내지 제32항 중 어느 한 항에 있어서, 표적 폴리뉴클레오티드가 5000개 이상의 뉴클레오티드를 포함하는 것인 방법.

#### 청구항 34

제30항 내지 제33항 중 어느 한 항에 있어서, 분자 모터가 (a) 서열 22, 33 또는 52에 제시된 서열, 또는 (b) 전체 서열에 걸친 아미노산 동일성을 기초로 서열 22, 33 또는 52에 대해 40% 이상의 상동성을 가지며 헬리카제 활성을 보유하는 그의 변이체를 포함하는 것인 방법.

#### 청구항 35

포어와, 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터 사이에 복합체를 형성시키고, 이에 의해 표적 폴리뉴클레오티드를 특성화하기 위한 센서를 형성시키는 것을 포함하는, 표적 폴리뉴클레오티드를 특성화하기 위한 센서를 형성시키는 방법.

#### 청구항 36

포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하기 위한, 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터의 용도.

#### 청구항 37

(a) 포어 및 (b) 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터를 포함하는, 표적 폴리뉴클레오티드를 특성화하기 위한 키트.

#### 청구항 38

다수의 포어, 및 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 다수의 분자 모터를 포함하는, 샘플 내의 표적 폴리뉴클레오티드를 특성화하기 위한 분석 장치.

#### 청구항 39

(a) 표적 폴리뉴클레오티드를 막횡단 포어 및 0.3 M 이상의 염 농도를 견딜 수 있는 헬리카제와 접촉시켜, 헬리카제가 포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하고 표적 뉴클레오티드 내의 뉴클레오티드들이 포어와 상호작용하게 하는 단계; 및

(b) 하나 이상의 상호작용 동안 표적 폴리뉴클레오티드의 하나 이상의 특성을 측정하고, 이에 의해 표적 폴리뉴클레오티드를 특성화하는 단계

를 포함하며, 0.3 M 이상의 염 농도를 사용하여 수행되는, 표적 폴리뉴클레오티드를 특성화하는 방법.

#### 청구항 40

포어와 0.3 M 이상의 염 농도를 견딜 수 있는 헬리카제 사이에 복합체를 형성시키고, 이에 의해 표적 폴리뉴클레오티드를 특성화하기 위한 센서를 형성시키는 것을 포함하는, 표적 폴리뉴클레오티드를 특성화하기 위한 센서를 형성시키는 방법.

#### 청구항 41

포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하기 위한, 0.3 M 이상의 염 농도를 견딜 수 있는 헬리카제의 용도.

#### 청구항 42

(a) 포어 및 (b) 0.3 M 이상의 염 농도를 견딜 수 있는 헬리카제를 포함하는, 표적 폴리뉴클레오티드를 특성화하기 위한 키트.

#### 청구항 43

다수의 포어, 및 0.3 M 이상의 염 농도를 견딜 수 있는 다수의 헬리카제를 포함하는, 샘플 내의 표적 폴리뉴클레오티드를 특성화하기 위한 분석 장치.

### 명세서

#### 기술 분야

##### [0001] 발명의 분야

[0002] 본 발명은 표적 폴리뉴클레오티드를 특성화하는 새로운 방법에 관한 것이다. 이러한 방법은 포어(pore) 및 He1308 헬리카제(helicase) 또는 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터를 사용한다. 헬리카제 또는 분자 모터는 포어를 통한 표적 폴리뉴클레오티드의 이동을 제어한다.

#### 배경 기술

##### [0003] 발명의 배경

[0004] 현재, 광범위한 용도에 걸쳐 신속하고 저렴한 폴리뉴클레오티드 (예를 들어 DNA 또는 RNA) 서열분석 및 확인 기술이 요구된다. 기존의 기술은 느리고 고가인데, 이는 주로 이들이 다량의 폴리뉴클레오티드를 생산하기 위해 증폭 기술에 의존하고, 신호 검출을 위해 높은 품질의 전문 형광 화학물질을 필요로 하기 때문이다.

[0005] 막횡단 포어 (나노포어(nanopore))는 중합체 및 다양한 소형 분자에 대한 직접적인 전기 바이오센서로서의 잠재력이 크다. 특히, 최근에는 잠재적인 DNA 서열분석 기술로서의 나노포어에 초점이 맞춰졌다.

[0006] 나노포어를 가로질러 전위가 인가되면, 분석물, 예컨대 뉴클레오티드가 특정 기간 동안 일시적으로 배럴(barrel) 내에 체류할 때 전류 흐름에 변화가 있다. 뉴클레오티드의 나노포어 검출은 공지된 서명 및 기간의 전류 변화를 제공한다. "가닥 서열분석(Strand Sequencing)" 방법에서, 단일 폴리뉴클레오티드 가닥이 포어를 통과하고, 뉴클레오티드의 신원(identity)이 유래된다. 가닥 서열분석은 포어를 통한 폴리뉴클레오티드의 이동을 제어하는 뉴클레오티드 취급 단백질의 사용을 수반할 수 있다.

#### 발명의 내용

##### [0007] 발명의 개요

[0008] 본 발명자들은 He1308 헬리카제가, 특히 전위, 예컨대 전압이 인가되는 경우에, 포어를 통한 폴리뉴클레오티드의 이동을 제어할 수 있다는 것을 입증하였다. 헬리카제는 인가된 전압으로부터 초래되는 전기장에 대항하여 또는 이러한 전기장과 함께 표적 폴리뉴클레오티드를 제어 및 단계적 방식으로 이동시킬 수 있다. 놀랍게도, 이러한 헬리카제는 폴리뉴클레오티드를 특성화하는데, 특히 가닥 서열분석을 사용하여 이의 서열을 결정하는데 유리한 높은 염 농도에서 기능할 수 있다. 이는 하기에서 더욱 상세하게 논의된다.

[0009] 따라서, 본 발명은

[0010] (a) 표적 폴리뉴클레오티드를 막횡단 포어 및 He1308 헬리카제와 접촉시켜, 헬리카제가 포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하고 표적 뉴클레오티드 내의 뉴클레오티드들이 포어와 상호작용하게 하는 단계; 및

[0011] (b) 하나 이상의 상호작용 동안 표적 폴리뉴클레오티드의 하나 이상의 특성을 측정하고, 이에 의해 표적 폴리뉴클레오티드를 특성화하는 단계

[0012] 를 포함하는, 표적 폴리뉴클레오티드를 특성화하는 방법을 제공한다.

[0013] 본 발명은 하기를 또한 제공한다:

[0014] - 포어와 He1308 헬리카제 사이에 복합체를 형성시키고, 이에 의해 표적 폴리뉴클레오티드를 특성화하기 위한 센서를 형성시키는 것을 포함하는, 표적 폴리뉴클레오티드를 특성화하기 위한 센서를 형성시키는 방법;

[0015] - 포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하기 위한, He1308 헬리카제의 용도;

- [0016] - (a) 포어 및 (b) He1308 헬리카제를 포함하는, 표적 폴리뉴클레오타이드를 특성화하기 위한 키트; 및
- [0017] - 다수의 포어 및 다수의 He1308 헬리카제를 포함하는, 샘플 내의 표적 폴리뉴클레오타이드를 특성화하기 위한 분석 장치.
- [0018] 본 발명자들은 내부 뉴클레오타이드에서의 표적 폴리뉴클레오타이드에 결합할 수 있는 분자 모터가, 특히 전위, 예컨대 전압이 인가되는 경우에, 포어를 통한 폴리뉴클레오타이드의 이동을 제어할 수 있다는 것을 또한 입증하였다. 이러한 모터는 인가된 전압으로부터 초래되는 전기장에 대항하여 또는 이러한 전기장과 함께 표적 폴리뉴클레오타이드를 제어 및 단계적 방식으로 이동시킬 수 있다. 놀랍게도, 이러한 모터가 본 발명의 방법에서 사용되는 경우에, 나노포어를 통한 표적 폴리뉴클레오타이드의 전체 가닥의 이동을 제어하는 것이 가능하다. 이는 폴리뉴클레오타이드를 특성화하는데, 특히 가닥 서열분석을 사용하여 이의 서열을 결정하는데 유리하다.
- [0019] 따라서, 본 발명은
- [0020] (a) 표적 폴리뉴클레오타이드를 막횡단 포어, 및 내부 뉴클레오타이드에서의 표적 폴리뉴클레오타이드에 결합할 수 있는 분자 모터와 접촉시켜, 분자 모터가 포어를 통한 표적 폴리뉴클레오타이드의 이동을 제어하고 표적 뉴클레오타이드 내의 뉴클레오타이드들이 포어와 상호작용하게 하는 단계; 및
- [0021] (b) 하나 이상의 상호작용 동안 표적 폴리뉴클레오타이드의 하나 이상의 특성을 측정하고, 이에 의해 표적 폴리뉴클레오타이드를 특성화하는 단계
- [0022] 를 포함하는, 표적 폴리뉴클레오타이드를 특성화하는 방법을 또한 제공한다.

## 도면의 간단한 설명

### [0023] 도면 설명

도 1. a) 나노포어를 통한 DNA 이동을 제어하기 위한 헬리카제의 사용의 예시적인 개략도. 1) 콜레스테롤-태그 (c)를 함유하는 프라이머 (b)가 어닐링(annealing)되어 있는 ssDNA 기질 (a)이 이중층 (d)의 시스(cis) 측에 첨가된다. 콜레스테롤 태그가 이중층에 결합하여, 이중층 표면에서 기질을 강화시킨다. 2) 시스 구획에 첨가된 헬리카제 (e)가 DNA에 결합한다. 2가 금속 이온 및 NTP 기질의 존재 하에, 헬리카제가 DNA를 따라 이동한다. 3) 인가된 전압 하에, DNA 기질이 DNA 상의 리더(leader) 섹션을 통해 나노포어 (f)에 의해 포획된다. DNA에 결합된 헬리카제가 포어의 상부와 접촉하여 추가의 제어되지 않은 DNA 전위를 방지할 때까지, 인가된 전위의 힘 하에 DNA가 포어를 통해 잡아당겨 진다. 이러한 프로세스 동안, dsDNA 섹션 (예컨대 프라이머)이 제거된다. 3' → 5' 방향으로 DNA를 따라 헬리카제가 이동하는 것은 꿰여 있는 DNA를 인가된 장에 대항하여 포어 밖으로 잡아당긴다. 4) 헬리카제가 DNA를 나노포어 밖으로 잡아당겨, 이를 시스 구획으로 되돌린다. 나노포어를 통과하는 DNA의 마지막 섹션은 5'-리더이다. 5) 헬리카제가 DNA를 나노포어 밖으로 이동시킬 때, 이는 시스 구획으로 되돌려 놓여진다. b) 실시예에서 사용된 DNA 기질 디자인 (a = 50T 리더 (b)가 있는 DNA의 400량체 가닥, c = 프라이머, d = 콜레스테롤-태그).

도 2. 헬리카제가 제어 방식으로 나노포어를 통해 DNA를 이동시킬 수 있어서, DNA가 나노포어를 통해 이동함에 따라 전류에서의 단계적 변화를 일으킨다. 예시적인 헬리카제-DNA 사례 (상부 섹션에서 소형 화살표로 지시됨) (180 mV, 400 mM KCl, Hepes pH 8.0, 0.15 nM 400량체 DNA, 100 nM He1308 Mbu, 1 mM DTT, 1 mM ATP, 1 mM MgCl<sub>2</sub>). 상부) He1308 400량체 DNA 사례의 전류 (y축, pA) 대 시간 (x축, s) 취득의 섹션. 개방-포어 전류는 ~180 pA이다 (A로 표시됨). 인가된 전위 (+180 mV)의 힘 하에 DNA가 나노포어에 의해 포획된다. 효소가 부착된 DNA는 긴 차단 (이러한 조건에서 ~60 pA)을 초래하고, 이는 효소가 포어를 통해 DNA를 이동시키기에 따라 전류에서의 단계적 변화를 나타낸다. 중간) 중간 섹션은 DNA 사례들 (1) 중 하나의 확대도이고, DNA-효소 포획, DNA가 포어를 통해 잡아당겨짐에 따른 단계적 전류 변화, 및 나노포어를 떠나기 전에 특징적인 긴 폴리T 수준으로 끝나는 것을 나타낸다. 하부) DNA가 나노포어를 통해 이동함에 따른 전류에서의 단계적 변화의 확대도.

도 3. DNA가 나노포어를 통과함에 따른 일관적인 패턴의 전류 변이를 초래하는, 헬리카제-제어 DNA 이동 (도 3a 및 3b에 대해 y축 = 전류 (pA), x축 = 시간 (s)). 폴리T 수준으로 끝나는 4개의 전형적인 DNA 사례로부터의 마지막 ~80 전류 변이의 예. 4개의 예 (3a에서의 2개 및 3b에서의 2개)는 일관적인 패턴의 전류 변이가 관찰된다는 것을 도해한다.

도 4. 증가된 염 농도가 포어 전류를 증가시키고, 더 큰 DNA 구별 범위를 제공한다 (범위 = DNA 전류 변이에 걸친 최소 전류 대 최대 전류). 400 mM, 1 M, 및 2 M KCl에서의 예시적인 헬리카제-DNA 사례 (도 4a-c에 대해

y축 = 전류 (pA), x축 = 시간 (s), 180 mV, Hepes pH 8.0, 0.15 nM 400량체 DNA 서열 59 및 60, 100 nM He1308 Mbu, 1 mM DTT, 1 mM ATP, 1 mM MgCl<sub>2</sub>)가 도 4 및 4c에서 제시된다. 상부 트레이스(trace)는 폴리T 수준으로 끝나는 전체 사례를 나타내고 (I-개방이 A로 지시됨), 하부 트레이스는 각각의 사례의 마지막 10초의 줌(zoom) 섹션을 나타내며, 이때 일정한 y축 전류 척도는 150 pA이다. 염 농도를 400 mM KCl에서 2M KCl로 증가시키는 것은 개방-포어 전류에서의 ~350% 증가 (~180 pA → ~850 pA의 I-개방), 및 구별 범위에서의 ~200% 증가 (~25 pA → ~75 pA)에 이른다. 도 4d는 염 농도의 함수로서의 DNA 구별 범위의 플롯이다 (y축 = 범위 (pA), x축 = 염 (mM)).

도 5. 헬리카제 (a)가 2가지 이상의 작동 방식으로 DNA (b)의 이동을 제어할 수 있다. 헬리카제는 3'-5' 방향으로 DNA를 따라서 이동하지만, 나노포어 (c) 내의 DNA의 배향 (DNA의 어느 쪽 끝부분이 포획되는지에 좌우됨)은 효소가 인가된 장에 대항하여 DNA를 나노포어 밖으로 이동시키는데 (도 5b) 또는 인가된 장과 함께 DNA를 나노포어 내로 이동시키는데 (도 5a) 사용될 수 있다는 것을 의미한다. 도 5b) DNA의 5' 끝부분이 포획된 경우, 헬리카제는 전압에 의해 인가된 장의 방향에 대항하여 작업하여, DNA가 시스 챔버(chamber)로 되돌려 사출될 때까지 레인 DNA를 나노포어의 밖으로 잡아당긴다. 오른쪽은 인가된 장에 대항하여 5'을 아래로 하여 나아가는 He1308로부터의 예시적인 DNA-헬리카제 사례가다 (y축 = 전류 (pA), x축 = 시간 (s)). 도 5a) DNA가 나노포어에서 3'을 아래로 하여 포획된 경우, DNA가 포어를 통해 완전히 전위되어 이중층의 트랜스 측 상에서 떠날 때까지 효소는 DNA를 장의 방향으로 나노포어 내로 이동시킨다. 오른쪽은 인가된 장과 함께 3'을 아래로 하여 나아가는 He1308로부터의 예시적인 DNA-헬리카제 사례가다 (y축 = 전류 (pA), x축 = 시간 (s)). DNA의 5'을 아래로 하는 배향과 3'을 아래로 하는 배향 간에 전류 트레이스가 다르다.

도 6. 효소 활성을 테스트하기 위한 형광 검정법. a) 맞춤형 형광 기질을 사용하여, 혼성화된 dsDNA를 전치시키는 헬리카제 (a)의 능력을 검정하였다. 1) 형광 기질 가닥 (100 nM 최종)에 3' ssDNA 오버행(overhang), 및 혼성화된 dsDNA의 염기 40개의 섹션이 있다. 상부의 주 가닥 (b)은 5' 끝부분에 카르복시플루오레세인 염기 (c)가 있고, 혼성화된 상보물 (d)은 3' 끝부분에 블랙-홀 퀸처(black-hole quencher) (BHQ-1) 염기 (e)가 있다. 혼성화되었을 때, 플루오레세인으로부터의 형광이 국소적인 BHQ-1에 의해 퀸칭(quenching)되고, 기질이 본질적으로 비-형광성이다. 1 μM의 포획 가닥 (f) (형광 기질의 더 짧은 가닥 (d)에 상보성임)이 검정법에 포함된다. 2) 제시된 바와 같이, ATP (1 mM) 및 MgCl<sub>2</sub> (5 mM)의 존재 하에, 기질에 첨가된 헬리카제 (100 nM)가 형광 기질의 3' 꼬리에 결합하고, 주 가닥을 따라 이동하고, 상보적인 가닥을 전치시킨다. 3) BHQ-1이 있는 상보적인 가닥이 완전히 전치되면, 주 가닥 상의 플루오레세인이 형광을 발한다. 4) 과량의 포획 가닥이 상보적인 DNA에 우선적으로 어닐링하여, 초기 기질이 다시 어닐링되는 것 및 형광 상실을 방지한다. b) 400 mM 내지 2 M의 상이한 농도의 KCl (x축, mM)을 함유하는 완충제 용액 (10 mM Hepes pH 8.0, 1 mM ATP, 5 mM MgCl<sub>2</sub>, 100 nM 형광 기질 DNA, 1 μM 포획 DNA)에서의 활성의 초기 비율의 그래프 (y축, 상대 활성).

도 7은 상이한 He1308 헬리카제들을 사용하는 헬리카제-제어 DNA 사례의 예를 나타낸다 (도 7a-c에 대해, y축 = 전류 (pA), x축 = 시간 (분), 180 mV, Hepes pH 8.0, 0.15 nM 400량체 DNA 서열 59 및 60, 100 nM He1308, 1 mM DTT, 1 mM ATP, 1 mM MgCl<sub>2</sub>): He1308 Mhu (a), He1308 Mok (b) 및 He1308 Mma (c). 이들은 폴리T 수준으로 끝난 MspA 나노포어를 통한 DNA 제어 이동의 전형적인 예를 나타낸다.

도 8. 헬리카제 내부 결합 활성을 테스트하기 위한 형광 검정법. A) 맞춤형 형광 기질을 사용하여, 천연 3' 끝부분이 결여된 DNA에 결합하고, 이어서 혼성화된 dsDNA를 전치시키게 하는 헬리카제의 능력을 검정하였다. 형광 기질 가닥 (50 nM 최종)에 3' ssDNA 오버행, 및 혼성화된 dsDNA의 염기 40개의 섹션이 있다. 상부의 주 가닥 (a)이 3' 끝부분에서, 또는 내부에서 오버행과 dsDNA 사이의 접합부에서 (음성 대조군으로서), 4개의 연속적인 비-DNA-유래 트리메틸렌 글리콜 스페이서 ("스페이서 9" 기로 지칭되고, b로 표지됨)로 변형된다. 또한, 상부의 주 가닥은 5' 끝부분에 카르복시플루오레세인 염기 (c)가 있고, 혼성화된 상보물 (d)은 3' 끝부분에 블랙-홀 퀸처 (BHQ-1) 염기 (e)가 있다. 혼성화되었을 때, 플루오레세인으로부터의 형광이 국소적인 BHQ-1에 의해 퀸칭되고, 기질이 본질적으로 비-형광성이다. 포획 가닥 (1 μM, f) (형광 기질의 더 짧은 가닥 (d)에 상보성임)이 검정법에 포함된다. B) ATP (1 mM) 및 MgCl<sub>2</sub> (1 mM)의 존재 하에, 3'-말단 "스페이서 9" 기를 함유하는 기질에 첨가된 He1308 헬리카제 상동체 (20 nM, g)가 형광 기질의 ssDNA 오버행에 결합할 수 있고, 주 가닥을 따라 이동할 수 있으며, 상보적 가닥을 전치시킬 수 있다. C) BHQ-1이 있는 상보적인 가닥이 완전히 전치되면, 주 가닥 상의 플루오레세인이 형광을 발한다. D) 과량의 포획 가닥이 상보적인 DNA에 우선적으로 어닐링하여, 초기 기질이 다시 어닐링되는 것 및 형광 상실을 방지한다.



도 9는 400 mM NaCl, 10 mM Hepes, pH 8.0, 1 mM ATP, 1 mM MgCl<sub>2</sub>, 50 nM 형광 기질 DNA, 1 μM 포획 DNA에서의 3'-미변형 DNA 및 3'-"스페이스 9" DNA를 비교하는 Hel308-매개 dsDNA 턴오버(turnover)의 상대적인 비율을 나타낸다 (y축 = 상대적인 "3'-Sp9" 활성 (% wrt 천연 3'), x축 = a (Mbu), b (Csy), C (Tga), d (Mma), e (Mhu), f (Min), g (Mig), h (Mmaz), i (Mac), j (Mok), k (Mth), l (Mba), m (Mzh)).

도 10. 실시예 5에서 사용된 나노포어를 통한 DNA 이동을 제어하기 위한 헬리카제의 사용의 개략도. A) 콜레스테롤-태그 (z로 표지됨)가 부착되어 있는 프라이머 (서열 69 (y로 표지됨))가 어닐링되어 있는 DNA 기질 (서열 67 (w로 표지됨) 및 68 (x로 표지됨))이 이중층의 시스 측에 첨가된다. 콜레스테롤 태그가 이중층에 결합하여, 이중층 표면에서 기질을 강화시킨다. 시스 구획에 첨가된 헬리카제 (1로 표지됨)가 서열 67의 4 bp 리더에 결합한다. B) 인가된 전압 하에, DNA 기질이 DNA 상의 5' 리더 섹션을 통해 나노포어에 의해 포획되고, 이는 서열 69를 벗겨낸다. C) 인가된 장의 힘 하에, 결합된 헬리카제 (1)가 포어의 상부와 접촉하여 추가의 제어되지 않은 전위를 방지할 때까지, DNA가 포어 내로 잡아당겨 진다. 이러한 프로세스에서, 안티센스(antisense)가닥인 서열 68이 이러한 DNA 가닥으로부터 벗겨진다. D) 2가 금속 이온 및 NTP 기질의 존재 하에, 포어 상부의 헬리카제 (1)가 DNA를 따라 이동하고, 포어를 통한 DNA의 전위를 제어한다. 3' → 5' 방향으로 DNA를 따라 헬리카제가 이동하는 것은 꿰여 있는 DNA를 인가된 장에 대항하여 포어 밖으로 잡아당긴다. 시스 측의 노출된 단일 가닥 DNA (이러한 경우 3')는 추가적인 헬리카제들 (2-4)이 말단 뉴클레오티드에서 또는 내부 뉴클레오티드에서 결합하는 것에 이용될 수 있다. E) 포어에 있는 헬리카제 (1)가 DNA로부터 탈착되면, DNA 상의 다음 헬리카제 (2)가 포어에 도달할 때까지 DNA가 장에 의해 포어 내로 잡아당겨 진다. 포어에 있는 헬리카제가 DNA를 나노포어 밖으로 잡아당겨, 이를 시스 구획으로 되돌린다. 나노포어를 통과하는 DNA의 마지막 섹션은 5'-리더이다. F) 헬리카제가 DNA를 나노포어 밖으로 이동시킬 때, 이는 시스 구획으로 되돌려 놓쳐진다. 화살표는 DNA 이동 방향을 가리킨다.

도 11은 각각의 헬리카제 사례 동안 Hel308 헬리카제 상동체 Mbu가 MspA 포어를 통한 DNA 가닥의 전위 (x축)를 제어함에 따라 어떻게 900량체의 나노포어 내의 DNA 영역의 위치 (y축)가 변화하였는지를 가리키는 데이터 플롯이다 (y축 = 900량체에서의 위치, x축 = 지수). A-C는 대략적으로 가닥의 시작 부분부터 가닥의 끝부분까지의 전체 DNA 가닥의 전형적인 전위 사례를 나타내는 반면 (폴리T 리더를 통해 벗어남), 사례 D는 불완전한 DNA 전위의 예를 나타내고, 이때 효소 분리는 DNA가 가닥의 끝부분에 이르지 못한다는 것을 의미한다. 슬립(slip) (예컨대 점선 원으로 강조된 대형 슬립)은 서열이 가닥 내의 이전의 지점으로 다시 떨어지는 것을 가리키고, 효소 분리의 결과이다. 효소가 분리된 경우, DNA는 또 다른 효소가 추가로 가닥을 따라 포어와 접촉할 때까지 장의 힘 하에 나노포어 내로 다시 잡아당겨질 것이고, 그 후 헬리카제 이동이 계속된다.

도 12는 Hel308 헬리카제 상동체 Tga가 MspA 포어를 통한 DNA 가닥의 전위를 제어함에 따라 어떻게 900량체의 위치가 변화하였는지를 가리키는 데이터 플롯을 나타낸다 (y축 = 900량체에서의 위치, x축 = 지수). 사례 A-D는 전체 DNA 가닥의 전위를 나타낸다.

도 13은 Hel308 Mbu 헬리카제 (서열 10)의 효소 프로세스성(processivity)을 Hel 308 Mok 헬리카제 (서열 29)의 것에 비교하는데 사용된 형광 검정법을 나타낸다. 맞춤형 형광 기질을 사용하여, 혼성화된 dsDNA를 전치시키는 헬리카제의 능력을 검정하였다. 형광 기질 (50 nM 최종)에 3' ssDNA 오버행, 및 혼성화된 dsDNA의 염기쌍 80개 (a1) 및 33개 (a2)의 섹션이 있다 (섹션 A, 서열 70). 하부의 주 "주형" 가닥 (1)이 이의 3' 오버행에 인접한 80 nt "차단제" 가닥 (2, 서열 71), 및 33 nt 형광 프로브 (3, 서열 72) (이의 5' 및 3' 끝부분에서 각각 카르복시플루오레세인 (FAM) (4) 및 블랙-홀 켄처 (BHQ-1) (5) 염기로 표지됨)에 혼성화된다. 혼성화되었을 때, FAM이 BHQ-1에서 멀고, 기질이 본질적으로 형광성이다. 섹션 B에 제시된 바와 같이, ATP (1 mM) 및 MgCl<sub>2</sub> (10 mM)의 존재 하에 (ATP 및 MgCl<sub>2</sub>의 첨가가 각각 7 및 8로 지시됨), 헬리카제 (6, 20 nM)가 기질의 3' 오버행 (서열 70)에 결합하고, 하부 가닥을 따라 이동하고, 80 nt 차단제 가닥 (서열 71)을 전치시키기 시작한다. 프로세스성이면, 헬리카제가 형광 프로브 (섹션 C, 서열 72, 이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨)도 전치시킨다. 형광 프로브는 이의 5' 및 3' 끝부분이 자가-상보적인 방식으로 디자인되고, 따라서 전치되면 동역학적으로 안정적인 헤어핀을 형성하여, 프로브가 주형 가닥에 다시 어닐링하는 것을 방지한다 (섹션 D). 헤어핀 생성물의 형성 시, FAM이 BHQ-1에 인접하게 되고, 이의 형광성이 켄칭된다. 80량체 "차단제" (서열 71) 및 형광 (서열 72, 이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨) 가닥을 전치시킬 수 있는 프로세스성 효소는 따라서 경시적으로 형광 감소에 이를 것이다. 그러나, 효소의 프로세스성이 80 nt 미만이면, 이는 형광 가닥 (서열 72, 이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀

켄처 (BHQ-1)로 표지됨)을 전치시킬 수 없을 것이고, 따라서 "차단제" 가닥 (서열 71)이 하부의 주 가닥에 다시 어닐링할 것이다 (섹션 E).

도 14는 대조군 목적으로 또한 사용된 추가적인 맞춤 형광 기질을 나타낸다. 음성 대조군으로서 사용된 기질은 도 3에 기술된 것과 동일하지만 3' 오버행이 결여된다 (섹션 A, (서열 71 (1로 표지됨), 72 (2로 표지됨) (이의 5' 끝부분에서 카르복시플루오레세인 (FAM) (3)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1) (4)로 표지됨) 및 73 (5로 표지됨, 80 bp 섹션 (a1) 및 33 bp 섹션 (a2)로 이루어짐)). 도 3에 기술된 것과 유사하지만 염 기쌍 80개의 섹션이 결여된 기질 (서열 72 (2로 표지됨) (이의 5' 끝부분에서 카르복시플루오레세인 (FAM) (3)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1) (4)로 표지됨) 및 74 (6으로 표지됨, 28 bp 섹션 (a3)으로 이루어짐))이 활성이지만 반드시 프로세스성이지는 않은 헬리카제에 대한 양성 대조군으로서 사용되었다 (섹션 B).

도 15는 완충 용액 (400 mM NaCl, 10 mM Hepes pH 8.0, 1 mM ATP, 10 mM MgCl<sub>2</sub>, 50 nM 형광 기질 DNA (서열 70, 71 및 72 (이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨)에서 도 13에 제시된 프로세스성 기질에 대해 He1308 Mbu 헬리카제 (서열 10, 흰색 원) 및 He1 308 Mok 헬리카제 (서열 29, 흑색 삼각형)를 테스트했을 때의 시간-의존적 형광 변화의 그래프 (y축 = 형광 (임의 단위), x축 = 시간 (분))를 나타낸다. He1308 Mok이 나타난 형광 감소는 He1308 Mbu (서열 10)에 비교했을 때 이러한 복합체의 증가된 프로세스성을 나타낸다.

도 16은 완충 용액 (400 mM NaCl, 10 mM Hepes pH 8.0, 1 mM ATP, 10 mM MgCl<sub>2</sub>, 50 nM 형광 기질 DNA (서열 72 (이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨) 및 74))에서 양성 대조군 프로세스성 기질 (도 14 섹션 B에서 제시됨, 서열 72 (이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨) 및 74)에 대해 He1308 Mbu 헬리카제 (서열 10, 흰색 원) 및 He1 308 Mok 헬리카제 (서열 29, 흑색 삼각형)를 테스트했을 때의 시간-의존적 형광 변화의 그래프 (y축 = 형광 (임의 단위), x축 = 시간 (분))를 나타낸다. 이러한 양성 대조군은 양쪽 모두의 샘플에 대한 형광 감소가 나타내는 바와 같이, 양쪽 헬리카제 모두가 실제로 활성이었음을 입증하였다.

## 서열 목록 설명

서열 1은 MS-B1 돌연변이체 MspA 단량체를 코딩하는 코돈-최적화 폴리뉴클레오티드 서열을 나타낸다. 이러한 돌연변이체는 신호 서열이 결여되고, 하기의 돌연변이를 포함한다: D90N, D91N, D93N, D118R, D134R 및 E139K.

서열 2는 MspA 단량체의 MS-B1 돌연변이체의 성숙 형태의 아미노산 서열을 나타낸다. 이러한 돌연변이체는 신호 서열이 결여되고, 하기의 돌연변이를 포함한다: D90N, D91N, D93N, D118R, D134R 및 E139K.

서열 3은 α-용혈소-E111N/K147N (α-HL-NN; 문헌 [Stoddart et al., PNAS, 2009; 106(19): 7702-7707])의 1개의 서브유닛을 코딩하는 폴리뉴클레오티드 서열을 나타낸다.

서열 4는 α-HL-NN의 1개의 서브유닛의 아미노산 서열을 나타낸다.

서열 5 내지 7은 MspB, C 및 D의 아미노산 서열을 나타낸다.

서열 8은 He1308 모티프의 아미노산 서열을 나타낸다.

서열 9는 연장된 He1308 모티프의 아미노산 서열을 나타낸다.

서열 10 내지 58은 표 5의 He1308 헬리카제 및 모티프의 아미노산 서열을 나타낸다.

서열 59 내지 74는 실시예에서 사용된 서열을 나타낸다.

서열 75는 90면(공보 57면)부터의 정렬에서의 He1308 Dth의 서열을 나타낸다.

서열 76은 90면(공보 57면)부터의 정렬에서의 He1308 Mmar의 서열을 나타낸다.

서열 77은 90면(공보 57면)부터의 정렬에서의 He1308 Nth의 서열을 나타낸다.

서열 78은 90면(공보 57면)부터의 정렬에서의 컨센서스(consensus) 서열을 나타낸다.

## 발명을 실시하기 위한 구체적인 내용

### 발명의 상세한 설명

[0024]

- [0025] 개시된 생성물 및 방법의 여러 용도가 업계에서의 특정 요구에 맞춰질 수 있다는 것을 이해하여야 한다. 본원에서 사용된 전문 용어는 단지 본 발명의 특정 실시양태를 기술하기 위한 것이고, 제한적인 것으로 의도되지 않는다는 것을 또한 이해하여야 한다.
- [0026] 또한, 본 명세서 및 첨부된 청구항에서 사용되는 경우에, 단수형 형태 "a", "an", 및 "the"는 문맥적으로 명확하게 달리 지시되지 않는 한 복수의 지시 대상을 포함한다. 따라서, 예를 들어 "포어"에 대한 언급은 2개 이상의 이같은 포어를 포함하고, "헬리카제"에 대한 언급은 2개 이상의 이같은 헬리카제를 포함하며, "폴리뉴클레오타이드"에 대한 언급은 2개 이상의 이같은 폴리뉴클레오타이드를 포함하는 것 등이다.
- [0027] 본원에서 인용된 모든 간행물, 특허 및 특허 출원은, 상기이든 하기이든, 본원에 전문이 참조로 포함된다.
- [0028] 본 발명의 He1308 방법
- [0029] 본 발명은 표적 폴리뉴클레오타이드를 특성화하는 방법을 제공한다. 이러한 방법은 표적 폴리뉴클레오타이드를 막 횡단 포어 및 He1308 헬리카제와 접촉시켜, 헬리카제가 포어를 통한 표적 폴리뉴클레오타이드의 이동을 제어하고 표적 뉴클레오타이드 내의 뉴클레오타이드들이 포어와 상호작용하게 하는 단계를 포함한다. 그 후, 표적 폴리뉴클레오타이드의 하나 이상의 특성이 업계에 공지된 표준 방법을 사용하여 측정된다. 단계 (a) 및 (b)는 바람직하게는 포어를 가로질러 전위가 인가되면서 수행된다. 하기에서 더욱 상세하게 논의되는 바와 같이, 인가된 전위는 전형적으로 포어와 헬리카제 사이의 복합체의 형성을 초래한다. 인가된 전위는 전압 전위일 수 있다. 대안적으로, 인가된 전위는 화학 전위일 수 있다. 이의 예는 지질 막을 가로지르는 염 구배를 사용하는 것이다. 문헌 [Holden et al., J Am Chem Soc. 2007 Jul 11;129(27):8650-5]에 염 구배가 개시되어 있다.
- [0030] 일부 예에서, 하나 이상의 상호작용 동안 포어를 통과하는 전류가 표적 폴리뉴클레오타이드의 서열을 결정하는데 사용된다. 이것이 가닥 서열분석이다.
- [0031] 이러한 방법은 여러 장점이 있다. 첫째, 놀랍게도 본 발명자들은 He1308 헬리카제의 염 내성이 놀랍게도 높고, 따라서 본 발명의 방법이 높은 염 농도에서 수행될 수 있다는 것을 나타냈다. 가닥 서열분석의 정황에서, 전하 운반체, 예컨대 염이 전압 오프셋(offset)을 인가하여 표적 폴리뉴클레오타이드를 포획하고 전위시키기 위한 전도성 용액을 생성시키는데, 그리고 폴리뉴클레오타이드가 포어를 통과할 때 생성된 서열-의존적 전류 변화를 측정하는데 필요하다. 측정 신호가 염의 농도에 의존적이기 때문에, 취득된 신호의 규모를 증가시키기 위해 높은 염 농도를 사용하는 것이 유리하다. 높은 염 농도는 높은 신호 대 노이즈(noise) 비를 제공하고, 정상적인 전류 변동의 배경값에 대비하여 뉴클레오타이드의 존재를 지시하는 전류가 확인되게 한다. 가닥 서열분석을 위해, 100 mM을 초과하는 염 농도가 이상적이고, 1 M 이상의 염 농도가 바람직하다. 놀랍게도 본 발명자들은 He1308 헬리카제가 예를 들어 2 M만큼 높은 염 농도에서 효과적으로 기능할 것임을 나타냈다.
- [0032] 둘째, 전압이 인가될 때, 놀랍게도 He1308 헬리카제가 표적 폴리뉴클레오타이드를 2가지 방향으로, 즉 인가된 전압으로부터 초래되는 장과 함께 또는 이에 대항하여 이동시킬 수 있다. 따라서, 본 발명의 방법이 2개의 바람직한 양식 중 하나로 수행될 수 있다. 표적 폴리뉴클레오타이드가 포어를 통해 이동하는 방향, 즉 장의 방향 또는 장에 반대되는 방향에 의존적으로 상이한 신호가 수득된다. 이는 하기에서 더욱 상세하게 논의된다.
- [0033] 셋째, 전형적으로 He1308 헬리카제는 표적 폴리뉴클레오타이드를 한번에 1개의 뉴클레오타이드로 포어를 통해 이동시킨다. 따라서 He1308 헬리카제는 단일-염기 래칫(ratchet)처럼 기능할 수 있다. 이는 표적 폴리뉴클레오타이드의 서열분석 시 당연히 유리한데, 표적 폴리뉴클레오타이드 내의 전부는 아니더라도 실질적으로 모든 뉴클레오타이드가 포어를 사용하여 확인될 수 있기 때문이다.
- [0034] 넷째, He1308 헬리카제는 단일 가닥 폴리뉴클레오타이드 및 이중 가닥 폴리뉴클레오타이드의 이동을 제어할 수 있다. 이는 다양한 상이한 표적 폴리뉴클레오타이드가 본 발명에 따라 특성화될 수 있다는 것을 의미한다.
- [0035] 다섯째, He1308 헬리카제는 인가된 전압으로부터 초래되는 장에 대해 매우 저항성인 것으로 보인다. 본 발명자들은 "언지핑(unzipping)" 조건 하에 폴리뉴클레오타이드의 이동을 거의 보지 못했다. 이는 인가된 전압으로부터 초래되는 장에 대항하여 폴리뉴클레오타이드를 이동시킬 때 원치않는 "역행" 이동으로부터의 복잡화가 없다는 것을 의미하기 때문에, 이것이 중요하다.
- [0036] 여섯째, He1308 헬리카제는 생산이 용이하고, 취급이 용이하다. 따라서 이의 사용은 수월하고 덜 고가인 서열분석 방법에 기여하였다.
- [0037] 본 발명의 방법은 표적 폴리뉴클레오타이드를 특성화하기 위한 것이다. 폴리뉴클레오타이드, 예컨대 핵산은 2개 이상의 뉴클레오타이드를 포함하는 거대분자이다. 폴리뉴클레오타이드 또는 핵산은 임의의 뉴클레오타이드의 임의의 조

합을 포함할 수 있다. 뉴클레오타이드는 천연 발생이거나 또는 인공일 수 있다. 표적 폴리뉴클레오타이드 내의 하나 이상의 뉴클레오타이드가 산화 또는 메틸화될 수 있다. 표적 폴리뉴클레오타이드 내의 하나 이상의 뉴클레오타이드가 손상될 수 있다. 표적 폴리뉴클레오타이드 내의 하나 이상의 뉴클레오타이드가, 예를 들어 표지 또는 태그로, 변형될 수 있다. 표적 폴리뉴클레오타이드는 하나 이상의 스페이서를 포함할 수 있다.

[0038] 뉴클레오타이드는 전형적으로 핵염기, 당 및 하나 이상의 포스페이트 기를 함유한다. 핵염기는 전형적으로 헤테로시클릭이다. 핵염기에는 퓨린 및 피리미딘, 더욱 구체적으로는 아데닌, 구아닌, 티민, 우라실 및 시토신이 포함되지만, 이에 한정되지는 않는다. 당은 전형적으로 펜토스 당이다. 뉴클레오타이드 당에는 리보스 및 데옥시리보스가 포함되지만, 이에 한정되지는 않는다. 뉴클레오타이드는 전형적으로 리보뉴클레오타이드 또는 데옥시리보뉴클레오타이드이다. 뉴클레오타이드는 전형적으로 모노포스페이트, 디포스페이트 또는 트리포스페이트를 함유한다. 포스페이트는 뉴클레오타이드의 5' 또는 3' 측면 상에 부착될 수 있다.

[0039] 뉴클레오타이드에는 아데노신 모노포스페이트 (AMP), 구아노신 모노포스페이트 (GMP), 티미딘 모노포스페이트 (TMP), 우리딘 모노포스페이트 (UMP), 시티딘 모노포스페이트 (CMP), 시클릭 아데노신 모노포스페이트 (cAMP), 시클릭 구아노신 모노포스페이트 (cGMP), 데옥시아데노신 모노포스페이트 (dAMP), 데옥시구아노신 모노포스페이트 (dGMP), 데옥시티미딘 모노포스페이트 (dTMP), 데옥시우리딘 모노포스페이트 (dUMP) 및 데옥시시티딘 모노포스페이트 (dCMP)가 포함되지만, 이에 한정되지는 않는다. 뉴클레오타이드는 바람직하게는 AMP, TMP, GMP, CMP, UMP, dAMP, dTMP, dGMP 또는 dCMP로부터 선택된다.

[0040] 뉴클레오타이드는 무염기성일 수 있다 (즉, 핵염기가 결여됨).

[0041] 폴리뉴클레오타이드는 단일 가닥 또는 이중 가닥일 수 있다. 바람직하게는, 폴리뉴클레오타이드의 적어도 일부가 이중 가닥이다.

[0042] 폴리뉴클레오타이드는 핵산, 예컨대 데옥시리보핵산 (DNA) 또는 리보핵산 (RNA)일 수 있다. 표적 폴리뉴클레오타이드는 1개의 DNA 가닥에 혼성화된 1개의 RNA 가닥을 포함할 수 있다. 폴리뉴클레오타이드는 업계에 공지된 임의의 합성 핵산, 예컨대 펩티드 핵산 (PNA), 글리세롤 핵산 (GNA), 트레오스 핵산 (TNA), 잠금 핵산 (LNA), 또는 뉴클레오타이드 측쇄가 있는 기타 합성 중합체일 수 있다.

[0043] 이러한 방법을 사용하여 표적 폴리뉴클레오타이드 전체가 또는 이의 일부만이 특성화될 수 있다. 표적 폴리뉴클레오타이드는 임의의 길이일 수 있다. 예를 들어, 폴리뉴클레오타이드는 길이가 뉴클레오타이드 쌍 10개 이상, 50개 이상, 100개 이상, 150개 이상, 200개 이상, 250개 이상, 300개 이상, 400개 이상 또는 500개 이상일 수 있다. 폴리뉴클레오타이드는 길이가 뉴클레오타이드 쌍 1000개 이상, 뉴클레오타이드 쌍 5000개 이상, 또는 뉴클레오타이드 쌍 100000개 이상일 수 있다.

[0044] 임의의 적절한 샘플 내에 표적 폴리뉴클레오타이드가 존재한다. 전형적으로 본 발명은 표적 폴리뉴클레오타이드를 함유하는 것으로 공지되었거나 이를 함유하는 것으로 추정되는 샘플 상에서 수행된다. 대안적으로, 샘플 내의 존재가 공지되어 있거나 예상되는 하나 이상의 표적 폴리뉴클레오타이드의 신원을 확증하기 위해 샘플 상에서 본 발명이 수행될 수 있다.

[0045] 샘플은 생물학적 샘플일 수 있다. 임의의 생물 또는 미생물로부터 수득되거나 추출된 샘플 상에서 시험관 내에서 본 발명이 수행될 수 있다. 생물 또는 미생물은 전형적으로는 고세균, 원핵생물 또는 진핵생물이고, 전형적으로는 5계 중 하나에 속한다: 식물계, 동물계, 균계, 원핵생물계 및 원생생물계. 임의의 바이러스로부터 수득되거나 추출된 샘플 상에서 시험관 내에서 본 발명이 수행될 수 있다. 바람직하게는 샘플은 유체 샘플이다. 전형적으로 샘플은 환자의 체액을 포함한다. 샘플은 소변, 림프, 타액, 점액 또는 양수일 수 있지만, 바람직하게는 혈액, 혈장 또는 혈청이다. 전형적으로, 샘플은 인간 기원이지만, 대안적으로 샘플이 또 다른 포유동물로부터 예컨대 말, 소, 양 또는 돼지와 같이 상업적으로 사용되는 동물로부터의 것일 수 있거나 또는 대안적으로 고양이 또는 개와 같은 애완동물로부터의 것일 수 있다. 대안적으로, 상업 작물, 예컨대 곡물, 콩, 과일 또는 채소, 예를 들어 밀, 보리, 귀리, 캐놀라, 옥수수, 대두, 벼, 바나나, 사과, 토마토, 감자, 포도, 담배, 강낭콩, 렌즈콩, 사탕수수, 코코아, 면으로부터 식물 기원의 샘플이 전형적으로 수득된다.

[0046] 샘플은 비-생물학적 샘플일 수 있다. 비-생물학적 샘플은 바람직하게는 유체 샘플이다. 비-생물학적 샘플의 예로는 수술용 유체, 물 예컨대 식수, 해수 또는 강물, 및 실험실 테스트용 시약이 포함된다.

[0047] 전형적으로는, 예를 들어 원심분리에 의해 또는 원치 않는 분자 또는 세포, 예컨대 적혈구를 여과해내는 막을 통과시키는 것에 의해, 검정되기 전에 샘플이 프로세싱된다. 취득 시 바로 샘플을 측정할 수 있다. 또한 전형



적으로 샘플이 검정법 전에, 바람직하게는  $-70^{\circ}\text{C}$  미만에서, 보관될 수 있다.

- [0048] 막형단 포어는 인가된 전위에 의해 구동된 수화 이온이 막의 한쪽 측면에서 막의 다른 쪽 측면으로 유동하게 하는 구조이다.
- [0049] 임의의 막이 본 발명에 따라 사용될 수 있다. 적절한 막이 업계에 주지되어 있다. 바람직하게는 막은 양친매성 층이다. 양친매성 층은 친수성 및 친지성 성질 양쪽 모두를 지니는 양친매성 분자, 예컨대 인지질로부터 형성된 층이다. 양친매성 층은 단층 또는 이중층일 수 있다.
- [0050] 막은 바람직하게는 지질 이중층이다. 지질 이중층은 세포막의 모델이고, 광범위한 실험 연구를 위한 우수한 플랫폼으로서의 역할을 한다. 예를 들어, 지질 이중층이 단일 채널 기록에 의한 막 단백질의 시험관내 연구에 사용될 수 있다. 대안적으로, 지질 이중층이 광범위한 물질의 존재를 검출하기 위한 바이오센서로서 사용될 수 있다. 지질 이중층은 임의의 지질 이중층일 수 있다. 적절한 지질 이중층에는 평면형 지질 이중층, 지지형 이중층 또는 리포솜이 포함되지만, 이에 한정되지는 않는다. 지질 이중층은 바람직하게는 평면형 지질 이중층이다. 적절한 지질 이중층이 국제 출원 번호 PCT/GB08/000563 (WO 2008/102121로 공개됨), 국제 출원 번호 PCT/GB08/004127 (WO 2009/077734로 공개됨) 및 국제 출원 번호 PCT/GB2006/001057 (WO 2006/100484로 공개됨)에 개시되어 있다.
- [0051] 지질 이중층을 형성시키는 방법이 업계에 공지되어 있다. 적절한 방법이 실시예에서 개시된다. 지질 이중층은 지질 단층이 수성 용액/공기 계면에 대해 수직인 구멍의 어느 한쪽 측면을 지나 수성 용액/공기 계면 상에 보유되는 문헌 [Montal and Mueller, Proc. Natl. Acad. Sci. USA., 1972; 69: 3561-3566]의 방법에 의해 통상적으로 형성된다.
- [0052] 단백질 포어 삽입에 적절한 양호한 품질의 지질 이중층을 형성시키는 비용-효과적이고 비교적 수월한 방법이기 때문에 몬탈(Montal) & 뮐러(Mueller)의 방법은 대중적이다. 이중층을 형성시키는 다른 통상적인 방법에는 팁-침지(tip-dipping), 이중층 페인팅, 및 리포솜 이중층의 패치-클램핑(patch-clamping)이 포함된다.
- [0053] 바람직한 실시양태에서, 국제 출원 번호 PCT/GB08/004127 (WO 2009/077734로 공개됨)에 기술된 바와 같이 지질 이중층이 형성된다.
- [0054] 또 다른 바람직한 실시양태에서, 막은 고체 상태 층이다. 고체 상태 층은 이의 기원이 생물학적이거나 아니다. 달리 말하면, 고체 상태 층은 생물학적 환경 예컨대 생물 또는 세포로부터 유래 또는 단리되지 않거나, 또는 생물학적으로 입수가능한 구조의 합성에 의해 제작된 버전이 아니다. 마이크로전자 물질, 절연 물질 예컨대  $\text{Si}_3\text{N}_4$ ,  $\text{Al}_2\text{O}_3$ , 및  $\text{SiO}_2$ , 유기 및 무기 중합체 예컨대 폴리아미드, 플라스틱 예컨대 테플론(Teflon)® 또는 엘라스토머(elastomer) 예컨대 2-성분 부가-경화 실리콘 고무, 및 유리를 포함하지만 이에 한정되지 않는 유기 및 무기 물질 양쪽으로부터 고체 상태 층이 형성될 수 있다. 1원자 층, 예컨대 그래핀, 또는 단지 원자 몇 개 두께인 층으로부터 고체 상태 층이 형성될 수 있다. 적절한 그래핀 층이 국제 출원 번호 PCT/US2008/010637 (WO 2009/035647로 공개됨)에 개시되어 있다.
- [0055] 전형적으로, (i) 포어를 포함하는 인공 이중층, (ii) 포어를 포함하는, 단리된 천연 발생 지질 이중층, 또는 (iii) 포어가 삽입되어 있는 세포를 사용하여 방법이 수행된다. 바람직하게는, 인공 지질 이중층을 사용하여 방법이 수행된다. 이중층은 포어에 더하여 다른 막형단 및/또는 막내 단백질, 뿐만 아니라 다른 분자를 포함할 수 있다. 적절한 장치 및 조건이 하기에서 논의된다. 본 발명의 방법은 전형적으로 시험관 내에서 수행된다.
- [0056] 폴리뉴클레오티드가 막에 커플링될 수 있다. 이는 임의의 공지된 방법을 사용하여 행해질 수 있다. 막이 양친매성 층, 예컨대 지질 이중층 (상기에서 상세하게 논의된 바와 같음)인 경우, 바람직하게는 폴리뉴클레오티드가 막 내에 존재하는 폴리펩티드 또는 막 내에 존재하는 소수성 앵커(anchor)를 통해 막에 커플링된다. 소수성 앵커는 바람직하게는 지질, 지방산, 스테롤, 탄소 나노튜브(nanotube) 또는 아미노산이다.
- [0057] 폴리뉴클레오티드가 막에 직접적으로 커플링될 수 있다. 바람직하게는 링커를 통해 폴리뉴클레오티드가 막에 커플링된다. 바람직한 링커에는 중합체, 예컨대 폴리뉴클레오티드, 폴리에틸렌 글리콜(PEG) 및 폴리펩티드가 포함되지만, 이에 한정되지는 않는다. 폴리뉴클레오티드가 막에 직접적으로 커플링되면, 막과 헬리카제 사이의 거리로 인해 특성화 실험이 폴리뉴클레오티드의 끝부분까지 계속될 수 없기 때문에 일부 데이터가 손실될 것이다. 링커가 사용되면, 폴리뉴클레오티드가완전하게 프로세싱될 수 있다. 링커가 사용되면, 링커는 임의의 위치에서 폴리뉴클레오티드에 부착될 수 있다. 바람직하게는, 링커가 꼬리 중합체에서 폴리뉴클레오티드에 부착된다.

[0058] 커플링은 안정적이거나 또는 일시적일 수 있다. 특정 용도에 대해, 커플링의 일시적인 성질이 선호된다. 안정적인 커플링 분자가 폴리뉴클레오티드의 5' 또는 3' 끝부분에 직접적으로 부착되면, 이중층과 헬리카제의 활성화 부위 사이의 거리로 인해 특성화 실행이 폴리뉴클레오티드의 끝부분까지 계속될 수 없기 때문에 일부 데이터가 손실될 것이다. 커플링이 일시적이면, 커플링된 끝부분이 무작위로 이중층으로부터 자유롭게 되었을 때, 폴리뉴클레오티드가 완전하게 프로세싱될 수 있다. 막과 안정적인 또는 일시적인 연결을 형성하는 화학 기가 하기에서 더욱 상세하게 논의된다. 콜레스테롤 또는 지방 아실 쇠를 사용하여 양친매성 층 또는 지질 이중층에 폴리뉴클레오티드가 일시적으로 커플링될 수 있다. 탄소 원자 6개 내지 30개 길이의 임의의 지방 아실 쇠, 예컨대 헥사데칸산이 사용될 수 있다.

[0059] 바람직한 실시양태에서, 폴리뉴클레오티드가 지질 이중층에 커플링된다. 폴리뉴클레오티드를 합성 지질 이중층에 커플링하는 것이 기존에 다양한 상이한 테더링(tethering) 전략으로 수행되었다. 이들이 하기 표 1에서 요약된다.

표 1

부착 기	커플링 유형	참고문헌
티올	안정적	Yoshina-Ishii, C. and S. G. Boxer (2003). "Arrays of mobile tethered vesicles on supported lipid bilayers." <i>J Am Chem Soc</i> <b>125</b> (13): 3696-7.
비오틴	안정적	Nikolov, V., R. Lipowsky, et al. (2007). "Behavior of giant vesicles with anchored DNA molecules." <i>Biophys J</i> <b>92</b> (12): 4356-68
콜레스테롤	일시적	Pfeiffer, I. and F. Hook (2004). "Bivalent cholesterol-based coupling of oligonucleotides to lipid membrane assemblies." <i>J Am Chem Soc</i> <b>126</b> (33): 10224-5
지질	안정적	van Lengerich, B., R. J. Rawle, et al. "Covalent attachment of lipid vesicles to a fluid-supported bilayer allows observation of DNA-mediated vesicle interactions." <i>Langmuir</i> <b>26</b> (11): 8666-72

[0060]

[0061] 반응성 기, 예컨대 티올, 콜레스테롤, 지질 및 비오틴 기의 부가에 대해서 쉽게 상용성인 변형된 포스포르아미다이트를 사용하여 합성 반응에서 폴리뉴클레오티드가 관능화될 수 있다. 이러한 상이한 부착 화학물질들은 폴리뉴클레오티드의 부착을 위한 한 벌의 선택권을 제공한다. 각각의 상이한 변형 기는 폴리뉴클레오티드를 약간씩 상이한 방식으로 테더링시키고, 커플링이 항상 영구적이지는 않아서, 이중층에 대한 폴리뉴클레오티드의 상이한 체류 시간을 제공한다. 일시적인 커플링의 장점이 상기에 논의되어 있다.

[0062] 반응성 기가 표적 폴리뉴클레오티드에 부가될 수 있는 한, 폴리뉴클레오티드의 커플링이 다수의 다른 수단에 의해 달성될 수도 있다. 반응성 기를 DNA의 한쪽 끝부분에 부가하는 것이 기존에 보고되었다. 폴리뉴클레오티드 키나제(kinase) 및 ATPγS를 사용하여 ssDNA의 5'에 티올 기가 부가될 수 있다 (문헌 [Grant, G. P. and P. Z. Qin (2007). "A facile method for attaching nitroxide spin labels at the 5' terminus of nucleic acids." *Nucleic Acids Res* **35**(10): e77]). 변형된 올리고뉴클레오티드를 ssDNA의 3'에 혼입시키도록 말단 트랜스퍼라제(transferase)를 사용하여 더욱 다양한 선택의 화학 기, 예컨대 비오틴, 티올 및 형광단이 부가될 수 있다 (문헌 [Kumar, A., P. Tchen, et al. (1988). "Nonradioactive labeling of synthetic oligonucleotide probes with terminal deoxynucleotidyl transferase." *Anal Biochem* **169**(2): 376-82])).

[0063] 대안적으로, 부착이 혼성화를 통해 달성될 수 있도록, 반응성 기가 이중층 내에 이미 커플링된 것에 대해 상보적인 짧은 DNA 조각의 부가인 것으로 간주될 수 있다. T4 RNA 리가제(ligase) I을 사용하여 짧은 ssDNA 조각의 라이게이션이 보고되어 있다 (문헌 [Troutt, A. B., M. G. McHeyzer-Williams, et al. (1992). "Ligation-anchored PCR: a simple amplification technique with single-sided specificity." *Proc Natl Acad Sci U S A* **89**(20): 9823-5])). 대안적으로, ssDNA 또는 dsDNA가 천연 dsDNA에 라이게이션될 수 있고, 그 후 2개의 가닥이 열 또는 화학 변성에 의해 분리될 수 있다. 천연 dsDNA에 대해, ssDNA 조각을 듀플렉스(duplex)의 한쪽 또는 양쪽 끝부분에, 또는 dsDNA를 한쪽 또는 양쪽 끝부분에 부가하는 것이 가능하다. 그 후, 듀플렉스가 용융되면, 각각의 단일 가닥에 ssDNA가 라이게이션에 사용되었으면 5' 또는 3' 변형이, dsDNA가 라이게이션에 사용되었으면 5' 끝부분, 3' 끝부분 또는 양쪽 모두에서의 변형이 있을 것이다. 폴리뉴클레오티드가 합성 가닥이면, 폴리뉴클레오티드의 화학적 합성 동안에 커플링 화학이 혼입될 수 있다. 예를 들어, 반응성 기가 부착되어 있는 프라이머를 사용하여 폴리뉴클레오티드를 합성할 수 있다.

- [0064] 게놈 DNA의 절편들의 증폭을 위한 통상적인 기술은 폴리머라제 연쇄 반응 (PCR)을 사용하는 것이다. 여기에서, 2개의 합성 올리고뉴클레오타이드 프라이머를 사용하여, DNA의 동일한 절편의 다수의 카피가 생성될 수 있고, 이때 각각의 카피에 대해 듀플렉스 내의 각각의 가닥의 5'이 합성 폴리뉴클레오타이드일 것이다. 반응성 기, 예컨대 콜레스테롤, 티올, 비오틴 또는 지질이 있는 안티센스 프라이머를 사용함으로써, 증폭된 표적 DNA의 각각의 카피는 커플링을 위한 반응성 기를 함유할 것이다.
- [0065] 막횡단 포어는 바람직하게는 막횡단 단백질 포어이다. 막횡단 단백질 포어는 수화 이온, 예컨대 분석물이 막의 한쪽 측면에서 막의 다른 쪽 측면으로 유동하게 하는 폴리펩티드 또는 폴리펩티드 수집이다. 본 발명에서, 막횡단 단백질 포어는 인가된 전위에 의해 구동된 수화 이온이 막의 한쪽 측면에서 다른 쪽으로 유동하게 하는 포어를 형성할 수 있다. 바람직하게는, 막횡단 단백질 포어는 분석물 예컨대 뉴클레오타이드가 막, 예컨대 지질 이중층의 한쪽 측면에서 다른 쪽으로 유동하게 한다. 막횡단 단백질 포어는 폴리뉴클레오타이드, 예컨대 DNA 또는 RNA가 포어를 통해 이동하게 한다.
- [0066] 막횡단 단백질 포어는 단량체 또는 올리고머일 수 있다. 바람직하게는, 포어는 여러 개의 반복되는 서브유닛, 예컨대 6개, 7개 또는 8개의 서브유닛으로 이루어진다. 더욱 바람직하게는, 포어는 7량체 또는 8량체 포어이다.
- [0067] 전형적으로는 막횡단 단백질 포어는 이온이 이를 통과하여 유동할 수 있는 배럴 또는 채널을 포함한다. 포어의 서브유닛들은 전형적으로 중심 축을 둘러싸고, 가닥들을 막횡단  $\beta$  배럴 또는 채널 또는 막횡단  $\alpha$ -나선 다발 또는 채널에 제공한다.
- [0068] 전형적으로, 막횡단 단백질 포어의 배럴 또는 채널은 분석물, 예컨대 뉴클레오타이드, 폴리뉴클레오타이드 또는 핵산과의 상호작용을 용이하게 하는 아미노산을 포함한다. 이러한 아미노산들은 바람직하게는 배럴 또는 채널의 협착부 근처에 위치한다. 전형적으로 막횡단 단백질 포어는 하나 이상의 양성으로 하전된 아미노산, 예컨대 아르기닌, 리신 또는 히스티딘, 또는 방향족 아미노산, 예컨대 티로신 또는 트립토판을 포함한다. 전형적으로 이러한 아미노산들은 포어와 뉴클레오타이드, 폴리뉴클레오타이드 또는 핵산 간의 상호작용을 용이하게 한다.
- [0069] 본 발명에 따라 사용하기 위한 막횡단 단백질 포어는  $\beta$ -배럴 포어 또는  $\alpha$ -나선 다발 포어로부터 유래될 수 있다.  $\beta$ -배럴 포어는  $\beta$ -가닥으로부터 형성된 배럴 또는 채널을 포함한다. 적절한  $\beta$ -배럴 포어에는  $\beta$ -독소, 예컨대  $\alpha$ -용혈소, 탄저병 독소 및 류코시딘, 및 박테리아의 외막 단백질/포린(porin), 예컨대 마이코박테리움 스메그마티스(*Mycobacterium smegmatis*) 포린 (Msp), 예를 들어 MspA, 외막 포린 F (OmpF), 외막 포린 G (OmpG), 외막 포스포리파제(phospholipase) A 및 네이세리아(*Neisseria*) 오토트랜스포터(autotransporter) 지 단백질 (NalP)이 포함되지만, 이에 한정되지는 않는다.  $\alpha$ -나선 다발 포어는  $\alpha$ -나선으로부터 형성된 배럴 또는 채널을 포함한다. 적절한  $\alpha$ -나선 다발 포어에는 내막 단백질 및  $\alpha$  외막 단백질, 예컨대 WZA 및 ClyA 독소가 포함되지만, 이에 한정되지는 않는다. 막횡단 포어는 Msp로부터 또는  $\alpha$ -용혈소 ( $\alpha$ -HL)로부터 유래될 수 있다.
- [0070] 바람직하게는 막횡단 단백질 포어는 Msp로부터, 바람직하게는 MspA로부터 유래된다. 이같은 포어는 올리고머일 수 있고, 전형적으로는 Msp로부터 유래된 7개, 8개, 9개 또는 10개의 단량체를 포함한다. 포어는 동일한 단량체들을 포함하는 Msp로부터 유래된 동종-올리고머 포어일 수 있다. 대안적으로, 포어는 서로 상이한 하나 이상의 단량체를 포함하는 Msp로부터 유래된 이종-올리고머 포어일 수 있다. 바람직하게는, 포어는 MspA 또는 이의 상동체 또는 파라로그(paralog)로부터 유래된다.
- [0071] Msp로부터 유래된 단량체는 서열 2에 제시된 서열 또는 그의 변이체를 포함한다. 서열 2는 MspA 단량체의 MS-(B1)8 돌연변이체이다. 이는 하기의 돌연변이를 포함한다: D90N, D91N, D93N, D118R, D134R 및 E139K. 서열 2의 변이체는 아미노산 서열이 서열 2의 것과 다르고 포어를 형성하는 능력을 보유하는 폴리펩티드이다. 변이체가 포어를 형성하는 능력을 업계에 공지된 임의의 방법을 사용하여 검정할 수 있다. 예를 들어, 다른 적합한 서브유닛과 함께 변이체가 지질 이중층 내로 삽입될 수 있고, 올리고머화되어 포어를 형성하는 이의 능력을 결정할 수 있다. 서브유닛을 막, 예컨대 지질 이중층 내로 삽입하는 방법들이 업계에 공지되어 있다. 예를 들어, 지질 이중층 내로 확산되고, 지질 이중층에 결합하고 기능성 상태로 조립되는 것에 의해 삽입되도록, 지질 이중층을 함유하는 용액 내에 서브유닛들을 정제된 형태로 현탁시킬 수 있다. 대안적으로, 문헌 [M.A. Holden, H. Bayley. J. Am. Chem. Soc. 2005, 127, 6502-6503] 및 국제 출원 번호 PCT/GB2006/001057 (WO 2006/100484로 공개됨)에 기술된 "픽 & 플레이스(pick and place)" 방법을 사용하여 서브유닛들을 막 내로 직접적으로 삽입할 수 있다.

- [0072] 서열 2의 아미노산 서열의 전체 길이에 걸쳐, 바람직하게는 변이체는 아미노산 서열 동일성을 기초로 이러한 서열에 대해 적어도 50% 상동성일 것이다. 더욱 바람직하게는, 변이체는 전체 서열에 걸쳐 서열 2의 아미노산 서열에 대해 아미노산 동일성을 기초로 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 더욱 바람직하게는 적어도 95%, 97% 또는 99% 상동성일 수 있다. 100개 이상, 예를 들어 125개, 150개, 175개 또는 200개 또는 이를 초과하는 개수의 연속적인 아미노산의 신장물에 걸쳐 아미노산 동일성이 적어도 80%, 예를 들어 적어도 85%, 90% 또는 95%일 수 있다 ("강한 상동성").
- [0073] 업계에서의 표준 방법을 사용하여 상동성을 결정할 수 있다. 예를 들어, UWGCG 패키지는 상동성을 계산하는데 사용될 수 있는 BESTFIT 프로그램을 제공하고, 예를 들어 이는 이의 디폴트 설정으로 사용된다 (문헌 [Devereux et al (1984) Nucleic Acids Research 12, p387-395]). 예를 들어 문헌 [Altschul S. F. (1993) J Mol Evol 36:290-300]; [Altschul, S.F et al (1990) J Mol Biol 215:403-10]에 기술된 바와 같이, PILEUP 및 BLAST 알고리즘을 사용하여 상동성을 계산하거나 또는 서열들을 정렬할 수 있다 (예컨대, 등가의 잔기들 또는 상응하는 서열들을 확인함 (전형적으로는 이들의 디폴트 설정을 기초로 함)). BLAST 분석을 수행하기 위한 소프트웨어는 국립 생물공학 정보 센터(National Center for Biotechnology Information) (<http://www.ncbi.nlm.nih.gov/>)를 통해 공개적으로 입수가능하다.
- [0074] 서열 2는 MspA 단량체의 MS-(B1)8 돌연변이체이다. 이러한 변이체는 MspA와 비교하여 MspB, C 또는 D 단량체에서의 돌연변이 중 임의의 것을 포함할 수 있다. 성숙 형태의 MspB, C 및 D가 서열 5 내지 7에서 제시된다. 특히, 변이체는 MspB에 존재하는 하기의 치환을 포함할 수 있다: A138P. 변이체는 MspC에 존재하는 하기의 치환 중 하나 이상을 포함할 수 있다: A96G, N102E 및 A138P. 변이체는 MspD에 존재하는 하기의 돌연변이 중 하나 이상을 포함할 수 있다: G1 결실, L2V, E5Q, L8V, D13G, W21A, D22E, K47T, I49H, I68V, D91G, A96Q, N102D, S103T, V104I, S136K 및 G141A. 변이체는 Msp B, C 및 D로부터의 돌연변이 및 치환 중 하나 이상의 조합을 포함할 수 있다. 바람직하게는 변이체는 돌연변이 L88N을 포함한다. 서열 2의 변이체는 MS-B1의 모든 돌연변이에 더하여 돌연변이 L88N이 있고, MS-B2로 칭해진다. 본 발명에서 사용된 포어는 바람직하게는 MS-(B2)8이다.
- [0075] 예를 들어 1개, 2개, 3개, 4개, 5개, 10개, 20개 또는 30개의 치환까지, 상기 논의된 것들에 더하여 아미노산 치환이 서열 2의 아미노산 서열에 이루어질 수 있다. 보존적 치환은 아미노산을 유사한 화학 구조, 유사한 화학 성질 또는 유사한 측쇄 부피의 다른 아미노산으로 교체한다. 도입된 아미노산은 자신이 교체하는 아미노산과 극성, 친수성, 소수성, 염기성, 산성, 중성, 또는 전하가 유사할 수 있다. 대안적으로, 보존적 치환은 방향족 또는 지방족인 또 다른 아미노산을 기존의 방향족 또는 지방족 아미노산 대신에 도입할 수 있다. 보존적 아미노산 변화는 업계에 주지되어 있고, 하기 표 2에 정의된 바와 같이 20개의 주요 아미노산의 성질에 따라 선택될 수 있다. 아미노산들의 극성이 유사한 경우, 이는 표 3의 아미노산 측쇄에 대한 소수친수성(hydrophathy) 규모를 참고로 또한 결정될 수 있다.

## 표 2

아미노산의 화학적 성질

Ala	지방족, 소수성, 중성	Met	소수성, 중성
Cys	극성, 소수성, 중성	Asn	극성, 친수성, 중성
Asp	극성, 친수성, 하전됨 (-)	Pro	소수성, 중성
Glu	극성, 친수성, 하전됨 (-)	Gln	극성, 친수성, 중성
Phe	방향족, 소수성, 중성	Arg	극성, 친수성, 하전됨 (+)
Gly	지방족, 중성	Ser	극성, 친수성, 중성
His	방향족, 극성, 친수성, 하전됨 (+)	Thr	극성, 친수성, 중성
Ile	지방족, 소수성, 중성	Val	지방족, 소수성, 중성
Lys	극성, 친수성, 하전됨 (+)	Trp	방향족, 소수성, 중성
Leu	지방족, 소수성, 중성	Tyr	방향족, 극성, 소수성

[0076]



표 3

소수친수성 규모

측쇄	소수친수성
Ile	4.5
Val	4.2
Leu	3.8
Phe	2.8
Cys	2.5
Met	1.9
Ala	1.8
Gly	-0.4
Thr	-0.7
Ser	-0.8
Trp	-0.9
Tyr	-1.3
Pro	-1.6
His	-3.2
Glu	-3.5
Gln	-3.5
Asp	-3.5
Asn	-3.5
Lys	-3.9
Arg	-4.5

[0077]

[0078]

서열 2의 아미노산 서열의 하나 이상의 아미노산 잔기가 상기 기술된 폴리펩티드로부터 추가적으로 결실될 수 있다. 1개, 2개, 3개, 4개, 5개, 10개, 20개 또는 30개 또는 이를 초과하는 개수까지의 잔기가 결실될 수 있다.

[0079]

변이체는 서열 2의 단편을 포함할 수 있다. 이같은 단편은 포어 형성 활성을 보유한다. 단편은 길이가 적어도 아미노산 50개, 100개, 150개 또는 200개일 수 있다. 이같은 단편이 포어를 생산하는데 사용될 수 있다. 바람직하게는 단편은 서열 2의 포어 형성 도메인을 포함한다. 단편은 서열 2의 잔기 88, 90, 91, 105, 118 및 134 중 하나를 포함하여야 한다. 전형적으로, 단편은 서열 2의 잔기 88, 90, 91, 105, 118 및 134 모두를 포함한다.

[0080]

하나 이상의 아미노산이 대안적으로 또는 추가적으로 상기 기술된 폴리펩티드에 부가될 수 있다. 연장은 서열 2 또는 이의 폴리펩티드 변이체 또는 단편의 아미노산 서열의 아미노 말단 또는 카르복시 말단에서 제공될 수 있다. 연장은 꽤 짧을 수 있고, 예를 들어 아미노산 1개 내지 10개의 길이이다. 대안적으로, 연장이 더 길 수 있고, 예를 들어 아미노산 50개 또는 100개까지일 수 있다. 캐리어(carrier) 단백질이 본 발명에 따른 아미노산 서열에 융합될 수 있다. 기타 융합 단백질이 하기에서 더욱 상세하게 논의된다.

[0081]

상기 논의된 바와 같이, 변이체는 서열 2의 것과 아미노산 서열이 다르고 이의 포어 형성 능력을 보유하는 폴리펩티드이다. 전형적으로 변이체는 포어 형성을 담당하는 서열 2의 영역을 함유한다.  $\beta$ -배럴을 함유하는 Msp의 포어 형성 능력은 각각의 서브유닛 내의  $\beta$ -시트에 의해 제공된다. 전형적으로 서열 2의 변이체는  $\beta$ -시트를 형성하는 서열 2 내의 영역을 포함한다. 생성된 변이체가 포어를 형성하는 능력을 보유하는 한,  $\beta$ -시트를 형성하는 서열 2의 영역에 하나 이상의 변형이 이루어질 수 있다. 바람직하게는 서열 2의 변이체는 이의  $\alpha$ -나선 및/또는 루프 영역 내에 하나 이상의 변형, 예컨대 치환, 부가 또는 결실을 포함한다.

[0082]

Msp로부터 유래된 단량체가 이의 확인 또는 정제를 보조하기 위해, 예를 들어 히스티딘 잔기(hist 태그), 아스파르트산 잔기(asp 태그), 스트렙타비딘 태그 또는 플래그(flag) 태그의 부가에 의해, 또는 세포로부터의 이의 분비를 촉진하기 위한 신호 서열의 부가에 의해 (폴리펩티드가 이같은 서열을 천연적으로 함유하지 않는 경우), 변형될 수 있다. 유전학적 태그를 도입하는 것에 대한 대안은 태그를 포어 상의 천연 또는 조작 위치 상에서 화학적으로 반응시키는 것이다. 이의 예는 겔-시프트(shift) 시약을 포어의 외부 상에 조작된 시스테인에 반응시키는 것일 것이다. 이는 용혈소 이중-올리고머들을 분리하기 위한 방법으로서 입증되었다 (문헌 [Chem Biol. 1997 Jul; 4(7):497-505]).

[0083]

Msp로부터 유래된 단량체가 표식 표지로 표지될 수 있다. 표식 표지는 포어가 검출되게 하는 임의의 적절한 표

지일 수 있다. 적절한 표지에는 형광 분자, 방사성 동위원소, 예를 들어 <sup>125</sup>I, <sup>35</sup>S, 효소, 항체, 항원, 폴리뉴클레오티드 및 리간드 예컨대 비오틴이 포함되지만, 이에 한정되지는 않는다.

[0084] D-아미노산을 사용하여 Msp로부터 유래된 단량체가 생산될 수도 있다. 예를 들어, Msp로부터 유래된 단량체가 L-아미노산 및 D-아미노산의 혼합물을 포함할 수 있다. 이는 이같은 단백질을 생산하기 위한 업계에서 통상적이다.

[0085] Msp로부터 유래된 단량체는 뉴클레오티드 구별을 용이하게 하는 하나 이상의 특이적 변형을 함유한다. Msp로부터 유래된 단량체는 다른 비-특이적 변형을 함유할 수도 있고, 단 이들은 포어 형성을 방해하지 않는다. 다수의 비-특이적 측쇄 변형이 업계에 공지되어 있고, Msp로부터 유래된 단량체의 측쇄에 이루어질 수 있다. 이같은 변형에는, 예를 들어 알데히드와의 반응에 이어지는 NaBH<sub>4</sub>로의 환원에 의한 아미노산의 환원성 알킬화, 메틸아세티미데이트로의 아미딘화, 또는 아세트산 무수물로의 아실화가 포함된다.

[0086] 업계에 공지된 표준 방법을 사용하여 Msp로부터 유래된 단량체가 생산될 수 있다. 합성에 의해 또는 재조합 수단에 의해 Msp로부터 유래된 단량체가 제조될 수 있다. 예를 들어, 시험관내 번역 및 전사 (IVTT)에 의해 포어가 합성될 수 있다. 포어를 생산하는 적절한 방법이 국제 출원 번호 PCT/GB09/001690 (WO 2010/004273으로 공개됨), PCT/GB09/001679 (WO 2010/004265로 공개됨) 또는 PCT/GB10/000133 (WO 2010/086603으로 공개됨)에 논의되어 있다. 포어를 막 내로 삽입하는 방법이 논의된다.

[0087] 또한 바람직하게는, 막횡단 단백질 포어가 α-용혈소 (α-HL)로부터 유래된다. 야생형 α-HL 포어는 7개의 동일한 단량체 또는 서브유닛으로 형성된다 (즉, 이는 7량체성이다). α-용혈소-NN의 한 단량체 또는 서브유닛의 서열이 서열 4에서 제시된다. 바람직하게는 막횡단 단백질 포어는 7개의 단량체를 포함하고, 이들 각각은 서열 4에 제시된 서열 또는 그의 변이체를 포함한다. 서열 4의 아미노산 1, 7 내지 21, 31 내지 34, 45 내지 51, 63 내지 66, 72, 92 내지 97, 104 내지 111, 124 내지 136, 149 내지 153, 160 내지 164, 173 내지 206, 210 내지 213, 217, 218, 223 내지 228, 236 내지 242, 262 내지 265, 272 내지 274, 287 내지 290 및 294가 루프 영역을 형성한다. 서열 4의 잔기 113 및 147은 α-HL의 배럴 또는 채널의 협착부의 일부분을 형성한다.

[0088] 이같은 실시양태에서, 각각 서열 4에 제시된 서열 또는 그의 변이체를 포함하는 7개의 단백질 또는 단량체를 포함하는 포어가 본 발명의 방법에서 바람직하게 사용된다. 7개의 단백질은 동일할 수 있거나 (동종7량체), 또는 상이할 수 있다 (이종7량체).

[0089] 서열 4의 변이체는 서열 4의 것과 아미노산 서열이 다르고 이의 포어 형성 능력을 보유하는 단백질이다. 변이체가 포어를 형성하는 능력을 업계에 공지된 임의의 방법을 사용하여 검정할 수 있다. 예를 들어, 다른 적합한 서브유닛과 함께 변이체가 지질 이중층 내로 삽입될 수 있고, 올리고머화되어 포어를 형성하는 이의 능력을 결정할 수 있다. 서브유닛을 막, 예컨대 지질 이중층 내로 삽입하는 방법들이 업계에 공지되어 있다. 적절한 방법이 상기에서 논의된다.

[0090] 변이체는 헬리카제에 대한 공유 부착 또는 이와 상호작용을 용이하게 하는 변형을 포함할 수 있다. 바람직하게는 변이체는 헬리카제에 대한 부착을 용이하게 하는 하나 이상의 반응성 시스테인 잔기를 포함한다. 예를 들어, 변이체는 서열 4의 위치 8, 9, 17, 18, 19, 44, 45, 50, 51, 237, 239 및 287 중 하나 이상 및/또는 아미노 또는 카르복시 말단에 시스테인을 포함할 수 있다. 바람직한 변이체는 시스테인으로의 서열 4의 위치 8, 9, 17, 237, 239 및 287의 잔기의 치환 (A8C, T9C, N17C, K237C, S239C 또는 E287C)을 포함한다. 바람직하게는, 변이체는 국제 출원 번호 PCT/GB09/001690 (WO 2010/004273으로 공개됨), PCT/GB09/001679 (WO 2010/004265로 공개됨) 또는 PCT/GB10/000133 (WO 2010/086603으로 공개됨)에 기술된 변이체 중 어느 하나이다.

[0091] 변이체는 뉴클레오티드와의 임의의 상호작용을 용이하게 하는 변형을 또한 포함할 수 있다.

[0092] 변이체는 생물, 예를 들어 스탕필로코쿠스 (*Staphylococcus*) 박테리아가 천연적으로 발현하는 천연 발생 변이체일 수 있다. 대안적으로, 시험관 내에서 또는 재조합적으로 에스케리키아 콜라이 (*Escherichia coli*)와 같은 박테리아에 의해 변이체가 발현될 수 있다. 재조합 기술에 의해 생산된 비-천연 발생 변이체가 변이체에 또한 포함된다. 서열 4의 아미노산 서열의 전체 길이에 걸쳐, 바람직하게는 변이체는 아미노산 동일성을 기초로 이러한 서열에 대해 적어도 50% 상동성일 것이다. 더욱 바람직하게는, 변이체 폴리펩티드는 전체 서열에 걸쳐 서열 4의 아미노산 서열에 대해 아미노산 동일성을 기초로 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 더욱 바람직하게는 적어도 95%, 97% 또는 99% 상동성일 수 있다. 200개 이상, 예를 들어 230개, 250개, 270개 또는 280개 또는 이를 초과하는 개수의 연속적인 아

미노산의 신장물에 걸쳐 아미노산 동일성이 적어도 80%, 예를 들어 적어도 85%, 90% 또는 95%일 수 있다 ("강한 상동성"). 상기 논의된 바와 같이 상동성을 결정할 수 있다.

- [0093] 예를 들어 1개, 2개, 3개, 4개, 5개, 10개, 20개 또는 30개의 치환까지, 상기 논의된 것들에 더하여 아미노산 치환이 서열 4의 아미노산 서열에 이루어질 수 있다. 보존적 치환이 상기 논의된 바와 같이 이루어질 수 있다.
- [0094] 서열 4의 아미노산 서열의 하나 이상의 아미노산 잔기가 상기 기술된 폴리펩티드로부터 추가적으로 결실될 수 있다. 1개, 2개, 3개, 4개, 5개, 10개, 20개 또는 30개 또는 이를 초과하는 개수까지의 잔기가 결실될 수 있다.
- [0095] 변이체는 서열 4의 단편일 수 있다. 이같은 단편은 포어 형성 활성을 보유한다. 단편은 길이가 적어도 아미노산 50개, 100개, 150개 또는 200개일 수 있다. 바람직하게는 단편은 서열 4의 포어 형성 도메인을 포함한다. 전형적으로, 단편은 서열 4의 잔기 119, 121, 135, 113 및 139를 포함한다.
- [0096] 하나 이상의 아미노산이 대안적으로 또는 추가적으로 상기 기술된 폴리펩티드에 부가될 수 있다. 연장은 서열 4 또는 그의 변이체 또는 단편의 아미노산 서열의 아미노 말단 또는 카르복시 말단에서 제공될 수 있다. 연장은 꽤 짧을 수 있고, 예를 들어 아미노산 1개 내지 10개의 길이이다. 대안적으로, 연장이 더 길 수 있고, 예를 들어 아미노산 50개 또는 100개까지일 수 있다. 캐리어 단백질이 포어 또는 변이체에 융합될 수 있다.
- [0097] 상기 논의된 바와 같이, 서열 4의 변이체는 서열 4의 것과 아미노산 서열이 다르고 이의 포어 형성 능력을 보유하는 서브유닛이다. 전형적으로 변이체는 포어 형성을 담당하는 서열 4의 영역을 함유한다.  $\beta$ -배럴을 함유하는  $\alpha$ -HLL의 포어 형성 능력은 각각의 서브유닛 내의  $\beta$ -가닥에 의해 제공된다. 전형적으로 서열 4의 변이체는  $\beta$ -가닥을 형성하는 서열 4 내의 영역을 포함한다.  $\beta$ -가닥을 형성하는 서열 4의 아미노산이 상기에서 논의된다. 생성된 변이체가 포어를 형성하는 능력을 보유하는 한,  $\beta$ -가닥을 형성하는 서열 4의 영역에 하나 이상의 변형이 이루어질 수 있다. 서열 4의  $\beta$ -가닥 영역에 이루어질 수 있는 구체적인 변형이 상기에서 논의된다.
- [0098] 바람직하게는 서열 4의 변이체는 이의  $\alpha$ -나선 및/또는 루프 영역 내에 하나 이상의 변형, 예컨대 치환, 부가 또는 결실을 포함한다.  $\alpha$ -나선 및 루프를 형성하는 아미노산이 상기에서 논의된다.
- [0099] 변이체는 상기 논의된 바와 같이 이의 확인 또는 정제를 보조하도록 변형될 수 있다.
- [0100] Msp로부터 유래된 포어에 관하여 상기 논의된 바와 같이  $\alpha$ -HLL로부터 유래된 포어를 제조할 수 있다.
- [0101] 일부 실시양태에서, 막횡단 단백질 포어가 화학적으로 변형된다. 임의의 방식으로 임의의 부위에서 포어가 화학적으로 변형될 수 있다. 바람직하게는, 막횡단 단백질 포어가 하나 이상의 시스테인에 분자가 부착되는 것 (시스테인 연결), 하나 이상의 리신에 분자가 부착되는 것, 하나 이상의 비-천연 아미노산에 분자가 부착되는 것, 에피토프의 효소 변형, 또는 말단 변형에 의해 화학적으로 변형된다. 이같은 변형을 수행하는 적절한 방법이 업계에 주지되어 있다. 임의의 분자의 부착에 의해 막횡단 단백질 포어가 화학적으로 변형될 수 있다. 예를 들어, 염료 또는 형광단의 부착에 의해 포어가 화학적으로 변형될 수 있다.
- [0102] 포어 내의 임의의 개수의 단량체가 화학적으로 변형될 수 있다. 바람직하게는 하나 이상, 예컨대 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개 또는 10개의 단량체가 상기 논의된 바와 같이 화학적으로 변형된다.
- [0103] 인접 잔기의 변형에 의해 시스테인 잔기의 반응성이 강화될 수 있다. 예를 들어, 측면에 있는 아르기닌, 히스티딘 또는 리신 잔기의 염기성기가 시스테인 티올기의 pKa를 더욱 반응성인 S<sup>-</sup>기의 것으로 변화시킬 것이다. 티올 보호기 예컨대 dTNB에 의해 시스테인 잔기의 반응성이 보호될 수 있다. 링커가 부착되기 전에 포어의 하나 이상의 시스테인 잔기와 이를 반응시킬 수 있다.
- [0104] 분자 (이러한 분자로 포어가 화학적으로 변형됨)는 국제 출원 번호 PCT/GB09/001690 (WO 2010/004273으로 공개됨), PCT/GB09/001679 (WO 2010/004265로 공개됨) 또는 PCT/GB10/000133 (WO 2010/086603으로 공개됨)에 개시된 바와 같이 링커를 통해 부착될 수 있거나 또는 포어에 직접적으로 부착될 수 있다.
- [0105] 임의의 Hel308 헬리카제가 본 발명에 따라 사용될 수 있다. Hel308 헬리카제는 ski2-유사 헬리카제로 또한 공지되어 있고, 이러한 2가지 용어는 상호교환가능하게 사용될 수 있다.
- [0106] Hel308 헬리카제는 전형적으로 아미노산 모티프 Q-X1-X2-G-R-A-G-R (이하 Hel308 모티프로 칭해짐; 서열 8)를 포함한다. Hel308 모티프는 전형적으로 헬리카제 모티프 VI의 일부이다 (문헌 [Tuteja and Tuteja, Eur. J. Biochem. 271, 1849-1863 (2004)]). X1은 C, M 또는 L일 수 있다. 바람직하게는 X1은 C이다. X2는 임의의 아미노산 잔기일 수 있다. X2는 전형적으로 소수성 또는 중성 잔기이다. X2는 A, F, M, C, V, L, I, S, T, P

또는 R일 수 있다. 바람직하게는 X2는 A, F, M, C, V, L, I, S, T 또는 P이다. 더욱 바람직하게는 X2는 A, M 또는 L이다. 가장 바람직하게는 X2는 A 또는 M이다.

[0107] He1308 헬리카제는 바람직하게는 모티프 Q-X1-X2-G-R-A-G-R-P (이하 연장된 He1308 모티프로 칭해짐; 서열 9) [식 중, X1 및 X2는 상기 기술된 바와 같다]를 포함한다.

[0108] 가장 바람직한 He1308 모티프 및 연장된 He1308 모티프가 하기 표 5에서 제시된다. He1308 헬리카제는 이러한 바람직한 모티프들 중 임의의 것을 포함할 수 있다.

[0109] He1308 헬리카제는 바람직하게는 하기 표 4에서 제시된 헬리카제 또는 그의 변이체 중 하나이다.

표 4

바람직한 He1308 헬리카제	
등록	설명
NP_578406.1	ski2-유사 헬리카제 [피로코쿠스 푸리오수스(Pyrococcus furiosus) DSM 3638] >sp Q9V0A9.1 HELS_PYRAB 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >emb CA949795.1  DNA 헬리카제 [피로코쿠스 아비시 GE5]
	>pdb 2ZJ2 A 쇠 A, 형태 1의 고세균 Dna 헬리카제 Hjm 아포(Apo) 상태 >sp O59025.1 HELS_PYRHO 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >pdb 2ZJ8 A 쇠 A, 형태 2의 고세균 Dna 헬리카제 Hjm 아포 상태 >pdb 2ZJA A 쇠 A, 형태 2의 Ampcp와 복합체를 형성한 고세균 Dna 헬리카제 Hjm >dbj BAA32016.1  헬리카제 [피로코쿠스 푸리오수스] >gb AAL80801.1  헬리카제 [피로코쿠스 푸리오수스 DSM 3638]

[0110]

NP_126564.1	ski2-유사 헬리카제 [피로코쿠스 아비시(Pyrococcus abyssi) GE5] >sp Q9V0A9.1 HELS_PYRAB 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >emb CA949795.1  DNA 헬리카제 [피로코쿠스 아비시 GE5]
NP_143168.1	ski2-유사 헬리카제 [피로코쿠스 호리코시이(Pyrococcus horikoshii) OT3] >sp O59025.1 HELS_PYRHO 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >dbj BAA30383.1  아미노산 715개 길이의 가설 단백질 [피로코쿠스 호리코시이 OT3]
YP_004424773.1	ski2-유사 헬리카제 [피로코쿠스(Pyrococcus) 종 NA2] >gb AEC52769.1  ski2-유사 헬리카제 [피로코쿠스 종 NA2]
YP_004623750.1	ski2-유사 헬리카제 [피로코쿠스 야야노시이(Pyrococcus yayanosii) CH1] >gb AEH24478.1  ski2-유사 헬리카제 [피로코쿠스 야야노시이 CH1]
YP_002307730.1	ski2-유사 헬리카제 [썬모코쿠스 온누리네우스(Thermococcus onnurineus) NA1] >gb ACJ16833.1  DNA 헬리카제 [썬모코쿠스 온누리네우스 NA1]
YP_004763427.1	ski2-유사 헬리카제 [썬모코쿠스(Thermococcus) 종 4557] >gb AEK73750.1  ski2-유사 헬리카제 [썬모코쿠스 종 4557]
YP_002959236.1	ski2-유사 헬리카제 [썬모코쿠스 감마톨레란스(Thermococcus gammatolerans) EJ3] >gb ACS33372.1  ski2-유형 헬리카제, 추정 [썬모코쿠스 감마톨레란스 EJ3]
YP_004071709.1	ski2-유형 헬리카제 [썬모코쿠스 바로필루스(Thermococcus barophilus) MP] >gb ADT84486.1  추정 ski2-유형 헬리카제 [썬모코쿠스 바로필루스 MP]
YP_002994328.1	추정 ski2-유형 헬리카제 [썬모코쿠스 시비리쿠스(Thermococcus sibiricus) MM 739] >gb ACS89979.1  추정 ski2-유형 헬리카제 [썬모코쿠스 시비리쿠스 MM 739]
ZP_04875329.1	유형 III 제한 효소, res 서브유닛 패밀리 [아시둘리프로폰돌 보오네이(Aciduliprofundum boonei) T469] >gb EDY35111.1  유형 III 제한 효소, res 서브유닛 패밀리 [아시둘리프로폰돌 보오네이 T469]
YP_003436565.1	DEAD/DEAH 박스 헬리카제 [페로글로부스 플라시두스(Ferroglobus placidus) DSM 10642] >gb ADC66290.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [페로글로부스 플라시두스 DSM 10642]
YP_004485304.1	ski2-유형 헬리카제 [메타노토리스 이그네우스(Methanoterris igneus) Kol 5] >gb AEF97239.1  ski2-유형 헬리카제 [메타노토리스 이그네우스 Kol 5]
YP_004616424.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노살숨 질리나에(Methanosalsum zhilinae) DSM 4017] >gb AEH61205.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노살숨 질리나에 DSM 4017]
ZP_04873370.1	유형 III 제한 효소, res 서브유닛 패밀리 [아시둘리프로폰돌 보오네이 T469] >ref YP_003482774.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [아시둘리프로폰돌 보오네이 T469] >gb EDY36687.1  유형 III 제한 효소, res 서브유닛 패밀리 [아시둘리프로폰돌 보오네이 T469] >gb ADD08212.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [아시둘리프로폰돌 보오네이 T469]
YP_004342552.1	ski2-유형 헬리카제 [아르카에오글로부스 베네퍼쿠스(Archaeoglobus veneficus) SNP6] >gb AEA47837.1  ski2-유형 헬리카제 [아르카에오글로부스 베네퍼쿠스 SNP6]
NP_071282.1	SKI2-패밀리 헬리카제 [아르카에오글로부스 풀기두스(Archaeoglobus fulgidus) DSM 4304]
2P6R_A	쇠 A, 풀린 Dna와 복합체를 이룬 슈퍼패밀리 2 헬리카제 He1308의 결정 구조 >pdb 2P6U A 쇠 A, He1308 슈퍼패밀리 2 헬리카제의 아포 구조
YP_685308.1	ski2-유사 헬리카제 [미배양 메탄생성 고세균 RC-1] >sp Q0W6L1.1 HELS_UNCMA 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >emb CAJ35982.1  추정 ski2-유형 헬리카제 [미배양 메탄생성 고세균 RC-1]
YP_001048404.1	ski2-유사 헬리카제 [메타노쿨레우스 마리스니그리(Methanoculleus marisnigri) JR1] >gb ABN58422.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노쿨레우스 마리스니그리 JR1]
YP_919908.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [썬모필럼 펜덴스(Thermophilum pendens) Hrk 5] >gb ABL77905.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [썬모필럼 펜덴스 Hrk 5]
YP_843229.1	ski2-유사 헬리카제 [메타노사에타 썬모필라(Methanoseta thermophila) PT] >gb ABK14589.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노사에타 썬모필라 PT]
ZP_08045937.1	ski2-유사 헬리카제 [할라다프타투스 포시알로필루스(Haladaptatus paucihalophilus) DX253] >gb EFW90585.1  ski2-유사 헬리카제 [할라다프타투스 포시알로필루스 DX253]

[0111]

NP_280985.1	ski2-유사 헬리카제 [할로박테리움(Halobacterium) 종 NRC-1] >ref YP_001690117.1  ski2-유사 헬리카제 [할로박테리움 살리나룸(Halobacterium salinarum) R1] >sp Q9HNV6.1 HEL5_HALSA 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >sp BOR7Q2.1 HEL5_HALS3 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb AAG20465.1  DNA 복구 단백질 [할로박테리움 종 NRC-1] >emb CAP14771.1  추정 DNA 헬리카제 [할로박테리움 살리나룸 R1]
YP_003357840.1	홀리데이 접합부(Holliday junction) 이동 헬리카제 [메타노셀라 팔루디콜라(Methanocella paludicola) SANAЕ] >dbj BAI62857.1  홀리데이 접합부 이동 헬리카제 [메타노셀라 팔루디콜라 SANAЕ]
YP_003457479.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노칼도코쿠스(Methanocaldococcus) 종 FS406-22] >gb ADC68743.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노칼도코쿠스 종 FS406-22]
YP_003127632.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노칼도코쿠스 페르벤스(Methanocaldococcus fervens) AG86] >gb ACV24132.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노칼도코쿠스 페르벤스 AG86]
YP_003735335.1	ski2-유사 헬리카제 [할랄칼리코쿠스 제오토갈리(Halalkalicoccus jeotgali) B3] >gb ADJ13543.1  ski2-유사 헬리카제 [할랄칼리코쿠스 제오토갈리 B3]
YP_503885.1	ski2-유사 헬리카제 [메타노스피릴럼 훈가테이(Methanospirillum hungatei) JF-1] >gb ABD42166.1  DEAD/DEAH 박스 헬리카제-유사 단백질 [메타노스피릴럼 훈가테이 JF-1]
BAJ48115.1	헬리카제 [칸디다투스 칼디아르카에움 서브테라네움(Candidatus Caldarchaeum subterraneum)] >dbj BAJ48144.1  헬리카제 [칸디다투스 칼디아르카에움 서브테라네움] >dbj BAJ50919.1  헬리카제 [칸디다투스 칼디아르카에움 서브테라네움]
YP_001405615.1	ski2-유사 헬리카제 [칸디다투스 메타노레굴라 보오네이(Candidatus Methanoregula boonei) 6A8] >sp A7IB61.1 HEL5_METB6 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb ABS56972.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노레굴라 보오네이 6A8]
YP_306959.1	ski2-유사 헬리카제 [메타노사르시나 바르케리 균주 푸사로(Methanosarcina barkeri str. Fusaro)] >sp Q463K3.1 HEL5_METBF 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb AAZ72379.1  헬리카제 [메타노사르시나 바르케리 균주 푸사로]
YP_001031179.1	ski2-유사 헬리카제 [메타노코르푸스쿨룸 라브레아눔(Methanocorpusculum labreanum) Z] >gb ABN07912.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코르푸스쿨룸 라브레아눔 Z]
YP_003541733.1	DEAD/DEAH 박스 헬리카제 [메타노할로필루스 마히이(Methanohalophilus mahii) DSM 5219] >gb ADE36088.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노할로필루스 마히이 DSM 5219]
YP_004384692.1	추정 Ski2-유형 헬리카제 [메타노사에타 콘실리이(Methanosaeta concilii) GP6] >gb AEB68874.1  추정 Ski2-유형 헬리카제 [메타노사에타 콘실리이 GP6]
YP_003725904.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노할로비움 에베스티가툼(Methanohalobium evestigatum) Z-7303] >gb ADI73108.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노할로비움 에베스티가툼 Z-7303]
YP_003405271.1	DEAD/DEAH 박스 헬리카제 [할로테리케나 투르크메니카(Haloterrigena turkmenica) DSM 5511] >gb ADB62598.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로테리케나 투르크메니카 DSM 5511]
YP_004244914.1	DEAD/DEAH 박스 헬리카제 [불카니사에타 모우트노브스키아(Vulcanisaeta moutnovskia) 768-28] >gb ADY01412.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [불카니사에타 모우트노브스키아 768-28]
YP_001540156.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [칼디비르가 마퀀린겐시스(Caldivirga maquilgensis) IC-167] >sp ASMB76.1 HEL5_CALMQ 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb ABW01166.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [칼디비르가 마퀀린겐시스 IC-167]
NP_618094.1	ski2-유사 헬리카제 [메타노사르시나 아세티보란스(Methanosarcina acetivorans) C2A] >sp Q8TL39.1 HEL5_METAC 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb AAM06574.1  헬리카제 [메타노사르시나 아세티보란스 C2A]
YP_003900980.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [불카니사에타 디스트리부타(Vulcanisaeta distributa) DSM 14429] >gb ADN49929.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [불카니사에타 디스트리부타 DSM 14429]

[0112]



YP_003896003.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노플라누스 페트로레아리우스(Methanoplanus petrolearius) DSM 11571] >gb ADN37565.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노플라누스 페트로레아리우스 DSM 11571]
YP_003615773.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노칼도코쿠스 인페르누스(Methanocaldococcus infernus) ME] >gb ADG12809.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노칼도코쿠스 인페르누스 ME]
YP_183745.1	RNA 헬리카제 Ski2-유사 단백질 [써모코쿠스 코다카렌시스(Thermococcus kodakarensis) KOD1] >sp Q5JGV6.1 HELS_PYRKO 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제; 함유: 권장명: 전체 명칭 = 엔도뉴클레아제 (Endonuclease) PI-PkoHel; 별칭: 전체 명칭 = Pko Hel 인테인 >dbi BAD85521.1  RNA 헬리카제 Ski2 삼동체 [써모코쿠스 코다카렌시스 KOD1]
YP_001322557.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 만니엘리(Methanococcus vannielii) SB] >sp AGUN73.1 HELS_METVS 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb ABR53945.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코쿠스 만니엘리 SB]
YP_002467772.1	ski2-유사 헬리카제 [메타노스파에롤라 팔루스트리스(Methanosphaerula palustris) E1-9c] >gb ACL18049.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노스파에롤라 팔루스트리스 E1-9c]
YP_003480097.1	DEAD/DEAH 박스 헬리카제 [나트리알바 마가다이(Natrialba magadii) ATCC 43099] >gb ADD05535.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [나트리알바 마가다이 ATCC 43099]
YP_004577043.1	ski2-유형 헬리카제 [메타노써모코쿠스 오키나헨시스(Methanothermococcus okinawensis) IH1] >gb AEH07265.1  ski2-유형 헬리카제 [메타노써모코쿠스 오키나헨시스 IH1]
YP_004742641.1	슈퍼페밀리 II 헬리카제 [메타노코쿠스 마리팔루디스(Methanococcus maripaludis) X1] >gb AEK19898.1  슈퍼페밀리 II 헬리카제 [메타노코쿠스 마리팔루디스 X1]
NP_632449.1	ski2-유사 헬리카제 [메타노사르시나 마제이(Methanosarcina mazei) Go1] >sp Q8PZR7.1 HELS_METMA 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb AAM30121.1  헬리카제 [메타노사르시나 마제이 Go1]
YP_001097223.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 마리팔루디스 C5] >gb ABQ35008.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코쿠스 마리팔루디스 C5]
YP_004742247.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 마리팔루디스 X1] >gb AEK19504.1  DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 마리팔루디스 X1]
YP_004794766.1	ski2-유사 헬리카제 [할로아르쿨라 히스파니카(Haloarcula hispanica) ATCC 33960] >gb AEM55778.1  ski2-유사 헬리카제 [할로아르쿨라 히스파니카 ATCC 33960]
NP_988010.1	슈퍼페밀리 II 헬리카제 [메타노코쿠스 마리팔루디스 S2] >emb CAF30446.1  슈퍼페밀리 II 헬리카제 [메타노코쿠스 마리팔루디스 S2]
YP_565780.1	ski2-유사 헬리카제 [메타노코코이데스 부르토니(Methanococcoides burtonii) DSM 6242] >sp Q12WZ6.1 HELS_METBU 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb ABE52030.1  DEAD/DEAH 박스 헬리카제-유사 단백질 [메타노코코이데스 부르토니 DSM 6242]
YP_001549808.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 마리팔루디스 C6] >gb ABX02576.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코쿠스 마리팔루디스 C6]
YP_001548609.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 마리팔루디스 C6] >gb ABX01377.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코쿠스 마리팔루디스 C6]
YP_001329359.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 마리팔루디스 C7] >gb ABR65208.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코쿠스 마리팔루디스 C7]
YP_004595982.1	ski2-유형 헬리카제 [할로피게르 크사나두엔시스(Halopiger xanaduensis) SH-61] >gb AEH36103.1  ski2-유형 헬리카제 [할로피게르 크사나두엔시스 SH-61]
YP_656795.1	ski2-유사 헬리카제 [할로쿼드라툼 왈스비(Haloquadratum walsbyi) DSM 16790] >emb CAJ51138.1  ATP-의존적 DNA 헬리카제 [할로쿼드라툼 왈스비 DSM 16790]
CCC38992.1	ATP-의존적 DNA 헬리카제 Hel308 [할로쿼드라툼 왈스비 C23]
YP_004035272.1	슈퍼페밀리 ii 헬리카제 [할로지오메트리움 보린헨세(Halogeometricum borinquense) DSM 11551] >gb ADQ65833.1  슈퍼페밀리 II 헬리카제 [할로지오메트리움 보린헨세 DSM 11551]

[0113]

YP_137330.1	ski2-유사 헬리카제 [할로아르쿨라 마리스모르투이(Haloarcula marismortui) ATCC 43049] >sp Q5UYM9.1 HEL5_HALMA 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb AAV47624.1  추정 ski2-유형 헬리카제 [할로아르쿨라 마리스모르투이 ATCC 43049]
YP_001581577.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [니트로소푸밀루스 마리티무스(Nitrosopumilus maritimus) SCM1] >gb ABX12139.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [니트로소푸밀루스 마리티무스 SCM1]
EET90255.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [칸디다투스 미크라르카에움 아시도필루스(Candidatus Micrarchaeum acidophilum) ARMAN-2]
NP_376477.1	헬리카제 [솔폴로부스 토크다아이(Sulfolobus tokodaii) 균주 7] >sp Q974S1.1 HEL5_SULTO 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >dbj BAK54341.1  홀리데이 접합부 이동 헬리카제 [솔폴로부스 토크다아이 균주 7]
YP_001097792.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 마리퀴루디스 C5] >gb ABO35578.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코쿠스 마리퀴루디스 C5]
ZP_08667240.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [니트로소푸밀루스(Nitrosopumilus) 종 MY1] >gb EGP92972.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [니트로소푸밀루스 종 MY1]
YP_254972.1	DNA 헬리카제 [솔폴로부스 아시도칼다리우스(Sulfolobus acidocaldarius) DSM 639] >sp Q4JCO0.1 HEL5_SULAC 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb AAV79679.1  DNA 헬리카제 [솔폴로부스 아시도칼다리우스 DSM 639]
EF092533.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [칸디다투스 파르마르카에움 아시도필루스(Candidatus Parvarchaeum acidophilus) ARMAN-5]
YP_003176527.1	ski2-유사 헬리카제 [할로미크로비움 무코하타에이(Halomicrobium mukohataei) DSM 12286] >gb ACV46820.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로미크로비움 무코하타에이 DSM 12286]
EGD71904.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [칸디다투스 파르마르카에움 아시도필루스 ARMAN-5 '5-박식 FS']
YP_001040230.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [스타필로테르무스 마리누스(Staphylothermus marinus) F1] >gb ABN69322.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [스타필로테르무스 마리누스 F1]
ABZ07376.1	추정 DEAD/DEAH 박스 헬리카제 [미배양 해양 크렌코세균(crenarchaeote) HF4000 ANTW133M9]
YP_001097458.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 마리퀴루디스 C5] >gb ABO35243.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코쿠스 마리퀴루디스 C5]
ABZ08606.1	추정 DEAD/DEAH 박스 헬리카제 [미배양 해양 크렌코세균 HF4000 APKG3H9]
YP_325906.1	ski2-유사 헬리카제 [나트로노모나스 파라오니스(Natronomonas pharaonis) DSM 2160] >sp Q31U46.1 HEL5_NATPD 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >emb CA148337.1  ATP-의존적 DNA 헬리카제 1 [나트로노모나스 파라오니스 DSM 2160]
YP_930665.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [피로바쿨룸 아일랜드리쿰(Pyrobaculum islandicum) DSM 4184] >gb ABL88322.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [피로바쿨룸 아일랜드리쿰 DSM 4184]
YP_001435870.1	DEAD/DEAH 박스 헬리카제 [이그니코쿠스 호스피탈리스(Ignicoccus hospitalis) KIN4/1] >gb ABU82463.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [이그니코쿠스 호스피탈리스 KIN4/1]
YP_003668634.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [스타필로테르무스 헬레니쿠스(Staphylothermus hellenicus) DSM 12710] >gb AD131735.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [스타필로테르무스 헬레니쿠스 DSM 12710]
ZP_08558598.1	ski2-유사 헬리카제 [할로랍두스 티아마테아(Halorhabdus tiamatea) SARLAB] >gb EGM36528.1  ski2-유사 헬리카제 [할로랍두스 티아마테아 SARLAB]

[0114]

YP_002428409.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [테슬루로코쿠스 캄차트켄시스 (Desulfurococcus kamchatkensis) 1221n] >gb ACL11042.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [테슬루로코쿠스 캄차트켄시스 1221n]
YP_004336918.1	ATP-의존적, DNA 결합 헬리카제 [써모프로테우스 우조니엔시스(Thermoproteus uzoniensis) 768-20] >gb AE11606.1  ATP-의존적, DNA 결합 헬리카제 [써모프로테우스 우조니엔시스 768-20]
ZP_08257442.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [칸디다투스 니트로소아르카에움 림니아(Candidatus Nitrosoarchaeum limnia) SFB1] >gb EGG41989.1  DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [칸디다투스 니트로소아르카에움 림니아 SFB1]
YP_004459284.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [아시디아누스 호스피탈리스 (Acidianus hospitalis) W1] >gb AEE94986.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [아시디아누스 호스피탈리스 W1]
NP_558924.1	ATP-의존적, DNA 결합 헬리카제 [피로바쿨럼 아에로필룸(Pyrobaculum aerophilum) 균주 IM2] >gb AAL63106.1  ATP-의존적, DNA 결합 헬리카제 [피로바쿨럼 아에로필룸 균주 IM2]
YP_004409449.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메탈로스파에라 쿠프리나 (Metallosphaera cuprina) Ar-4] >gb AEB94965.1  DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메탈로스파에라 쿠프리나 Ar-4]
YP_003649556.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [써모스파에라 아그레간스 (Thermosphaera aggregans) DSM 11486] >gb ADG90604.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [써모스파에라 아그레간스 DSM 11486]
ZP_06387115.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 솔파타리쿠스 (Sulfolobus solfataricus) 98/2] >gb ACX90562.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 솔파타리쿠스 98/2]
2VA8_A	쇄 A, Dna 복구 헬리카제 Hel308 >db 2VA8 B 쇄 B, Dna 복구 헬리카제 Hel308 >emb CA085626.1  DNA 헬리카제 [술폴로부스 솔파타리쿠스]
YP_004809267.1	ski2-유형 헬리카제 [호염 고세균 DL31] >gb AEN06894.1  ski2-유형 헬리카제 [호염 고세균 DL31]
ADX84345.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 (Sulfolobus islandicus) REY15A] >gb ADX81629.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 HVE10/4]
YP_002828439.1	DEAD/DEAH 박스 헬리카제 [술폴로부스 아일랜드리쿠스 M.14.25] >ref YP_002842325.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 M.16.27] >gb ACP37141.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 M.14.25] >gb ACP54280.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 M.16.27]
YP_002913571.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 M.16.4] >gb ACR40903.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 M.16.4]
Q97VY9.1	권장명: 전체 명칭 = 추정 ski2-유형 헬리카제
YP_002841682.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 Y.N.15.51] >gb ACP49760.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 Y.N.15.51]
YP_002831080.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 L.S.2.15] >ref YP_003418425.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 L.D.8.5] >gb ACP34435.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 L.S.2.15] >gb ADB86055.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 L.D.8.5]
YP_001054984.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [피로바쿨럼 칼리디폰티스 (Pyrobaculum calidifontis) JCM 11548] >sp A3MSA1.1 HELS_PYRCJ 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb ABO07518.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [피로바쿨럼 칼리디폰티스 JCM 11548]
NP_343811.1	DNA 헬리카제 관련 단백질 [술폴로부스 솔파타리쿠스 P2] >ref YP_002836469.1  DEAD/DEAH 박스 헬리카제 [술폴로부스 아일랜드리쿠스 Y.G.57.14] >gb AAK42601.1  DNA 헬리카제 관련 단백질 [술폴로부스 솔파타리쿠스 P2] >gb ACP44547.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [술폴로부스 아일랜드리쿠스 Y.G.57.14]

[0115]



YP_001152379.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [피로바쿨룸 아르세나티쿰 (Pyrobaculum arsenaticum) DSM 13514] >gb ABP49727.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [피로바쿨룸 아르세나티쿰 DSM 13514]
YP_001191456.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메탈로스파에라 세둘라 (Metallosphaera sedula) DSM 5348] >gb ABP95532.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메탈로스파에라 세둘라 DSM 5348]
NP_147034.2	NP_147034.2 홀리데이 접합부 이동 헬리카제 [아에로피룸 페르닉스(Aeropyrum pernix) K1] >sp Q9VFQ8.2 HEL5_AERPE 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >dbj BAA79103.2  홀리데이 접합부 이동 헬리카제 [아에로피룸 페르닉스 K1]
YP_024158.1	ski2-유사 헬리카제 [피크로필루스 토리두스(Picrophilus torridus) DSM 9790] >gb AAT43965.1  UV-보호에서 수반되는 헬리카제 [피크로필루스 토리두스 DSM 9790]
YP_003816358.1	추정 ski2-유형 헬리카제 [아시딜로부스 사카로보란스(Acidilobus saccharovorans) 345-15] >gb ADL19327.1  추정 ski2-유형 헬리카제 [아시딜로부스 사카로보란스 345-15]
YP_003860265.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [이그니스파에라 아그레간스 (Ignisphaera aggregans) DSM 17230] >gb ADM28385.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [이그니스파에라 아그레간스 DSM 17230]
NP_394295.1	ski2-유사 헬리카제 [써모플라스마 아시도필룸(Thermoplasma acidophilum) DSM 1728] >sp Q9HJX7.1 HEL5_THEAC 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >emb CAC11964.1  DNA 헬리카제 관련 단백질 [써모플라스마 아시도필룸]
YP_876638.1	슈퍼페밀리 II 헬리카제 [세나르카에움 심비오움(Cenarchaeum symbiosum) Al] >gb ABK78334.1  슈퍼페밀리 II 헬리카제 [세나르카에움 심비오움 Al]
ZP_05571398.1	ski2-유사 헬리카제 [페로플라스마 아시다르마누스(Ferroplasma acidarmanus) fer1]
YP_004176252.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [데슬푸로코쿠스 무코수스 (Desulfurococcus mucosus) DSM 2162] >gb ADV64770.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [데슬푸로코쿠스 무코수스 DSM 2162]
YP_001737782.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [칸디다투스 코라르카에움 크립토폴룸(Candidatus Korarchaeum cryptofilum) OPF8] >gb ACB08099.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [칸디다투스 코라르카에움 크립토폴룸 OPF8]
EGQ40435.1	슈퍼페밀리 II 헬리카제 [칸디다투스 나노살리나룸(Candidatus Nanosalinarum) 중 J07AB56]
YP_002567343.1	ski2-유사 헬리카제 [할로루브룸 라쿠스프로퐁디(Halorubrum lacusprofundi) ATCC 49239] >gb ACM58273.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로루브룸 라쿠스프로퐁디 ATCC 49239]
YP_001793507.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [써모프로테우스 뉴트로필루스 (Thermoproteus neutrophilus) V24Sta] >gb ACB39061.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [써모프로테우스 뉴트로필루스 V24Sta]
YP_003534088.1	ATP-의존적 DNA 헬리카제 He1308a [할로페락스 볼카나이(Haloferax volcanii) DS2] >gb ADE04048.1  ATP-의존적 DNA 헬리카제 He1308a [할로페락스 볼카나이 DS2]
YP_004037165.1	슈퍼페밀리 ii 헬리카제 [할로지오메트리움 보린켄세 DSM 11551] >gb ADQ67720.1  슈퍼페밀리 II 헬리카제 [할로지오메트리움 보린켄세 DSM 11551]
NP_111333.1	ski2-유사 헬리카제 [써모플라스마 볼카나이움(Thermoplasma volcanium) GSS1] >sp Q97A12.1 HEL5_THEVO 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >dbj BAB59970.1  DNA 헬리카제 [써모플라스마 볼카나이움 GSS1]
YP_002565871.1	DEAD/DEAH 박스 헬리카제 [할로루브룸 라쿠스프로퐁디 ATCC 49239] >gb ACM58801.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로루브룸 라쿠스프로퐁디 ATCC 49239]
CCC39675.1	ATP-의존적 DNA 헬리카제 He1308 [할로퀴드라툼 왈스비이 C23]
YP_657401.1	ATP-의존적 DNA 헬리카제 [할로퀴드라툼 왈스비이 DSM 16790] >emb CAJ51759.1  ATP-의존적 DNA 헬리카제 [할로퀴드라툼 왈스비이 DSM 16790]
YP_003535028.1	ATP-의존적 DNA 헬리카제 He1308b [할로페락스 볼카나이 DS2] >gb ADE02398.1  ATP-의존적 DNA 헬리카제 He1308b [할로페락스 볼카나이 DS2]

[0116]

YP_003706863.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 볼타에 (Methanococcus voltae) A3] >gb AD135890.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코쿠스 볼타에 A3]
ABD17736.1	헬리카제 [메타노코쿠스 볼타에 PS]
NP_G13398.1	슈퍼페밀리 II 헬리카제 [메타노피루스 칸들레리(Methanopyrus kandleri) AV19] >gb AA01328.1  예상 슈퍼페밀리 II 헬리카제 [메타노피루스 칸들레리 AV19]
CBH38575.1	추정 ski2-유형 헬리카제 [미배양 고세균]
EEZ93258.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [칸디다투스 파르마르카에움 아시디필룸(Candidatus Parvarchaeum acidiphilum) ARMAN-4]
EG040350.1	슈퍼페밀리 II 헬리카제 [칸디다투스 나노살리나움 중 J07AB56]
YP_004004246.1	dead/deah 박스 헬리카제 도메인 단백질 [메타노테르무스 페르비두스(Methanothermus fervidus) DSM 2088] >gb ADP77484.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노테르무스 페르비두스 DSM 2088]
YP_003850109.1	헬리카제 [메타노세모박테르 마르부르겐시스 균주 마르부르크(Methanothermobacter marburgensis str. Marburg)] >gb ADL58796.1  예상 헬리카제 [메타노세모박테르 마르부르겐시스 균주 마르부르크]
YP_003424423.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노브레비박테르 루미난티움(Methanobrevibacter ruminantium) M1] >gb ADC47531.1  DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노브레비박테르 루미난티움 M1]
YP_004291107.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노박테리움(Methanobacterium) 중 AL-21] >gb ADZ10135.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노박테리움 중 AL-21]
YP_447162.1	헬리카제 [메타노스파에라 스타트르마나에(Methanosphaera stadtmanae) DSM 3091] >gb ABC56519.1  예상 헬리카제 [메타노스파에라 스타트르마나에 DSM 3091]
YP_004519549.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노박테리움 중 SWAN-1] >gb ABG17748.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노박테리움 중 SWAN-1]
NP_275949.1	DNA 헬리카제 관련 단백질 [메타노세모박테르 세모토티로피쿠스 균주 델타(Methanothermobacter thermotrophicus str. Delta) H] >sp O26901.1 HELS_METTH 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제 >gb AAB85310.1  DNA 헬리카제 관련 단백질 [메타노세모박테르 세모토티로피쿠스 균주 델타 H]
ZP_05975717.2	추정 Ski2-유형 헬리카제 [메타노브레비박테르 스미티이(Methanobrevibacter smithii) DSM 2374] >gb EFC93382.1  추정 Ski2-유형 헬리카제 [메타노브레비박테르 스미티이 DSM 2374]
ZP_03607647.1	가설 단백질 METSMIALI_00751 [메타노브레비박테르 스미티이 DSM 2375] >gb EEE41862.1  가설 단백질 METSMIALI_00751 [메타노브레비박테르 스미티이 DSM 2375]
YP_001273412.1	ATP-의존적 헬리카제 [메타노브레비박테르 스미티이 ATCC 35061] >gb ABQ87044.1  ATP-의존적 헬리카제 [메타노브레비박테르 스미티이 ATCC 35061]
YP_003247505.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노칼도코쿠스 볼카니우스(Methanocaldococcus vulcanius) M7] >gb ACX73023.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노칼도코쿠스 볼카니우스 M7]
NP_248116.1	SKI2 패밀리 헬리카제 [메타노칼도코쿠스 잔나쉬이(Methanocaldococcus jannaschii) DSM 2661] >sp Q58524.1 HELS_METJA 권장명: 전체 명칭 = 추정 ski2-유형 헬리카제; 함유: 권장명: 전체 명칭 = 엔도뉴클레아제 PI-MjaHel; 별칭: 전체 명칭 = Mja Hel 인테인; 별칭: 전체 명칭 = Mja Pep3 인테인 >gb AAB99126.1  추정 SKI2-패밀리 헬리카제 [메타노칼도코쿠스 잔나쉬이 DSM 2661]
YP_001324295.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메타노코쿠스 아에올리쿠스 난카이([Methanococcus aeolicus Nankai)-3] >gb ABR55683.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메타노코쿠스 아에올리쿠스 난카이-3]
YP_003536960.1	프리(pre)-mRNA 스플라이싱 헬리카제 [할로페락스 볼카나이 DS2] >gb ADE02332.1  프리-mRNA 스플라이싱 헬리카제 [할로페락스 볼카나이 DS2]

[0117]

YP_003131029.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로랍두스 우타헨시스(Halorhabdus utahensis) DSM 12940] >gb ACV12296.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로랍두스 우타헨시스 DSM 12940]
YP_002567151.1	DEAD/DEAH 박스 헬리카제 [할로루브룸 라쿠스프로폴디 ATCC 49239] >gb ACM58081.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로루브룸 라쿠스 프로폴디 ATCC 49239]
YP_004035351.1	슈퍼패밀리 ii 헬리카제 [할로지오메트리움 보란렌세 DSM 11551] >gb ADQ65912.1  슈퍼패밀리 II 헬리카제 [할로지오메트리움 보란렌세 DSM 11551]
YP_004808851.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [호염 고세균 DL31] >gb AEN06478.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [호염 고세균 DL31]
XP_002716686.1	예상: DNA 폴리머라제 세타 이소형(isoform) 1 [오리크톨라구스 쿠니쿨루스(Orvctolagus cuniculus)]
YP_656834.1	ATP-의존적 DNA 헬리카제 [할로퀴드라툼 왈스비이 DSM 16790] >emb CA51176.1  ATP-의존적 DNA 헬리카제 [할로퀴드라툼 왈스비이 DSM 16790]
XP_003248103.1	예상: DNA 폴리머라제 세타-유사 이소형 1 [아시르토시폰 피숨(Acyrtosiphon nisum)]
ABC72356.1	ATP-의존적 DNA 헬리카제 [할로퀴드라툼 왈스비이]
CCC39031.1	DEAD/DEAH 박스 헬리카제 [할로퀴드라툼 왈스비이 C23]
XP_001165150.2	예상: DNA 폴리머라제 세타 이소형 1 [판 트로글로디테스(Pan troglodytes)]
XP_003225852.1	예상: DNA 폴리머라제 세타-유사 [아놀리스 카롤리넨시스(Anolis carolinensis)]
XP_615375.3	예상: DNA 폴리머라제 세타 [보스 타우루스(Bos taurus)] >ref XP_002684835.1  예상: 폴리머라제 (DNA 지향성), 세타-유사 [보스 타우루스] >gb DAA33456.1  폴리머라제 (DNA 지향성), 세타-유사 [보스 타우루스]
XP_002813286.1	예상: 저품질 단백질: DNA 폴리머라제 세타-유사 [폰고 아벨리이(Pongo abelii)]
AA08421.2	DNA 폴리머라제 세타 [호모 사피엔스(Homo sapiens)]
EAW79510.1	폴리머라제 (DNA 지향성), 세타, 이소형 CRA_a [호모 사피엔스]
NP_955452.3	DNA 폴리머라제 세타 [호모 사피엔스] >sp O75417.2 DPOLQ_HUMAN 권장명: 전체 명칭 = DNA 폴리머라제 세타; 별칭: 전체 명칭 = DNA 폴리머라제 eta >gb AAI72289.1  폴리머라제 (DNA 지향성), 세타 [참성 폴리뉴클레오타이드]
NP_001099348.1	DNA 폴리머라제 세타 [라투스 노르베기쿠스(Rattus norvegicus)] >gb EDM11249.1  폴리머라제 (DNA 지향성), 세타 (예상), 이소형 CRA_a [라투스 노르베기쿠스]
XP_003341262.1	예상: 저품질 단백질: DNA 폴리머라제 세타-유사 [모노델피스 도메스티카(Monodelphis domestica)]
XP_001502374.3	예상: DNA 폴리머라제 세타 [에쿠우스 카발루스(Equus caballus)]
XP_545125.3	예상: 저품질 단백질: DNA 폴리머라제 세타 [카니스 루푸스 파밀리아리스(Canis lupus familiaris)]
XP_002928855.1	예상: 저품질 단백질: DNA 폴리머라제 세타-유사 [아일루로포다 멜라노레우카(Ailuropoda melanoleuca)]
NP_084253.1	DNA 폴리머라제 세타 이소형 1 [무스 무스쿨루스(Mus musculus)] >gb AAL77225.1  DNA 폴리머라제 세타 [무스 무스쿨루스] >gb EDK97951.1  폴리머라제 (DNA 지향성), 세타, 이소형 CRA_a [무스 무스쿨루스] >gb AAI38361.1  폴리머라제 (DNA 지향성), 세타 [무스 무스쿨루스] >gb AAI57901.1  폴리머라제 (DNA 지향성), 세타 [무스 무스쿨루스]
AAK39635.1	DNA 폴리머라제 세타 [호모 사피엔스]
AAK39838.1	DNA 폴리머라제 Q [무스 무스쿨루스]
XP_003412882.1	예상: DNA 폴리머라제 세타 [록소돈타 아프리카나(Loxodonta africana)]
YP_003735206.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [할랄칼리코루스 제오트칼리 B3] >gb ADJ13414.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할랄칼리코루스 제오트칼리 B3]
YP_004794841.1	프리-mRNA 스플라이싱 헬리카제 [할로아르콜라 히스파니카 ATCC 33960] >gb AEM55853.1  프리-mRNA 스플라이싱 헬리카제 [할로아르콜라 히스파니카 ATCC 33960]

[0118]

XP_416549.2	예상: DNA 폴리머라제 세타와 유사 [갈루스 갈루스( <i>Gallus gallus</i> )]
XP_003427319.1	예상: 헬리카제 POLQ-유사 이소형 2 [나소니아 비트리펜니스( <i>Nasonia vitripennis</i> )]
XP_003202748.1	예상: DNA 폴리머라제 세타-유사 [멜레아그리스 갈로파보( <i>Meleagris gallopavo</i> )]
XP_969311.1	예상: DNA 폴리머라제 세타와 유사 [트리볼리움 카스타네움( <i>Tribolium castaneum</i> )] >gb EEZ97532.1  가설 단백질 TcasGA2_TC011380 [트리볼리움 카스타네움]
ZP_08046037.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [할라다프루투스 포시알로필루스 DX253] >gb EFW90685.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할라다프루투스 포시알로필루스 DX253]
YP_461714.1	헬리카제 [신티트로푸스 아시디트로피쿠스( <i>Syntrophus aciditrophicus</i> ) SB] >gb ABC77546.1  헬리카제 [신티트로푸스 아시디트로피쿠스 SB]
YP_003176510.1	DEAD/DEAH 박스 헬리카제 [할로미크로비움 무코하타에이 DSM 12286] >gb ACV46803.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로미크로비움 무코하타에이 DSM 12286]
YP_137400.1	프리-mRNA 스플라이싱 헬리카제 [할로아르콜라 마리스모르투이 ATCC 43049] >gb AAV47694.1  프리-mRNA 스플라이싱 헬리카제 [할로아르콜라 마리스모르투이 ATCC 43049]
NP_001184156.1	폴리머라제 (DNA 지향성), 세타 [제노푸스 (실루라나) 트로피칼리스( <i>Xenopus (Silurana) tropicalis</i> )]
NP_280861.1	프리-mRNA 스플라이싱 헬리카제 [할로박테리움 중 NRC-1] >ref YP_001689987.1  ATP-의존적 DNA 헬리카제 [할로박테리움 살리나룸 R1] >gb AAG20341.1  프리-mRNA 스플라이싱 헬리카제 [할로박테리움 중 NRC-1] >emb CAP14641.1  ATP-의존적 DNA 헬리카제 [할로박테리움 살리나룸 R1]
YP_004595640.1	DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [할로피게르 크사나두엔시스 SH-6] >gb AEH35761.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로피게르 크사나두엔시스 SH-6]
XP_001521144.2	예상: DNA 폴리머라제 세타, 부분적 [오르니토린쿠스 아나티누스( <i>Ornithorhynchus anatinus</i> )]
XP_003261953.1	예상: DNA 폴리머라제 세타, 부분적 [노마스쿠스 레우코게니스( <i>Nomascus leucogenys</i> )]
XP_001358456.2	GA19301 [드로소필라 슈도옵스쿠라 슈도옵스쿠라( <i>Drosophila pseudoobscura pseudoobscura</i> )] >gb EAL27595.2  GA19301 [드로소필라 슈도옵스쿠라 슈도옵스쿠라]
ZP_08560003.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로람두스 티아마테아 SARLAB] >gb EGM34502.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로람두스 티아마테아 SARLAB]
XP_002187783.1	예상: 폴리머라제 (DNA 지향성), 세타와 유사 [타에나오키가 구타타( <i>Taeniopygia guttata</i> )]
XP_002112587.1	가설 단백질 TRIADDRAFT_25163 [트리코플락스 아드하에렌스( <i>Trichoplax adhaerens</i> )] >gb EDV24697.1  가설 단백질 TRIADDRAFT_25163 [트리코플락스 아드하에렌스]
YP_003405139.1	DEAD/DEAH 박스 헬리카제 [할로테리게나 투르크메니카 DSM 5511] >gb ADB62466.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로테리게나 투르크메니카 DSM 5511]
EGV92665.1	DNA 폴리머라제 세타 [크리세툴루스 그리세우스( <i>Cricetulus griseus</i> )]
CBY24305.1	무명 단백질 제품 [오이코플레우라 디오이카( <i>Oikopleura dioica</i> )]
YP_003130565.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로람두스 우타엔시스 DSM 12940] >gb ACV11832.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [할로람두스 우타엔시스 DSM 12940]
YP_003479811.1	DEAD/DEAH 박스 헬리카제 [나트리알바 마가다이 ATCC 43099] >gb ADD05249.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [나트리알바 마가다이 ATCC 43099]
EFB22383.1	가설 단백질 PANDA 000253 [아일루로포다 멜라노레우카]
YP_003357334.1	추정 ATP-의존적 헬리카제 [메타노셀라 플루디콜라 SANAE] >dbj BA162351.1  추정 ATP-의존적 헬리카제 [메타노셀라 플루디콜라 SANAE]
YP_325942.1	ATP-의존적 DNA 헬리카제 2 [나트로노모나스 파라오니스 DSM 2160] >emb CAI48373.2  ATP-의존적 DNA 헬리카제 2 [나트로노모나스 파라오니스 DSM 2160]
XP_002912509.1	예상: 지움진 단백질: 헬리카제 POLQ-유사 [아일루로포다 멜라노레우카]

[0119]

XP_002704678.1	예상: 헬리카제, POLQ-유사 [보스 타우루스]
CAE47762.2	인간 DNA-지향성 폴리머라제 세타 (POLQ)와 유사한 신규 단백질 [다니오 레리오 (Danio rerio)]
XP_003205636.1	예상: 헬리카제, POLQ-유사 [멜레아그리스 갈로파보]
XP_544959.2	예상: 헬리카제, POLQ-유사 [카니스 루푸스 파밀리아리스]
EFX86757.1	가설 단백질 DAPPUDRAFT_312857 [다프니아 폴렉스 (Daphnia pulex)]
YP_003389641.1	DEAD/DEAH 박스 헬리카제 [스피로소마 린구알레 (Spirosoma linguale) DSM 74] >gb ADB40842.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [스피로소마 린구알레 DSM 74]
XP_002602932.1	가설 단백질 BRAFLDRAFT_251779 [브란키오스토마 플로리다에 (Branchiostoma floridae)] >gb EEN58944.1  가설 단백질 BRAFLDRAFT_251779 [브란키오스토마 플로리다에]
YP_004144962.1	펩티다제(peptidase) C14 카스파제(caspase) 촉매 서브유닛 p20 [메소리조비움 시세리 생물변이형 비세롤라에 (Mesorhizobium ciceri biovar biserrulae) WSM1271] >ref YP_004614892.1  DEAD/DEAH 박스 헬리카제 도메인-함유 단백질 [메소리조비움 오포르투니즘 (Mesorhizobium opportunistum) WSM2075] >gb ADV14912.1  펩티다제 C14 카스파제 촉매 서브유닛 p20 [메소리조비움 시세리 생물변이형 비세롤라에 WSM1271] >gb AEH90798.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [메소리조비움 오포르투니즘 WSM2075]
XP_002124758.1	예상: DNA 폴리머라제 세타와 유사 [시오나 인테스티날리스 (Ciona intestinalis)]
XP_694437.5	예상: DNA 폴리머라제 세타 [다니오 레리오]
XP_420565.1	예상: DNA 헬리카제 HEL308와 유사 [갈루스 갈루스]
XP_003129397.1	예상: 헬리카제, POLQ-유사 [수스 스크로파 (Sus scrofa)]
EDL20278.1	mCG128467, 이소형 CRA b [무스 무스쿨루스]
XP_001517710.2	예상: 헬리카제, POLQ, 부분적 [오르니토리쿠스 아니티누스]
AAH82601.1	헬리카제, mus308-유사 (드로소필라 (Drosophila)) [무스 무스쿨루스]
XP_003384429.1	예상: DNA 폴리머라제 세타-유사 [알페데몬 퀸스란디카 (Amphimedon queenslandica)]
XP_003221282.1	예상: 헬리카제, POLQ-유사 [아틀리스 카롤리넨시스]
NP_524333.1	돌연변이원-감수성 308 [드로소필라 멜라노가스테르 (Drosophila melanogaster)] >gb AAB67306.1  Mus308 [드로소필라 멜라노가스테르] >gb AAF54858.1  돌연변이원-감수성 308 [드로소필라 멜라노가스테르] >gb ACH92234.1  F1Q3732p [드로소필라 멜라노가스테르]
AA333507.1	LP14642p [드로소필라 멜라노가스테르]
NP_001074576.1	헬리카제 POLQ-유사 [무스 무스쿨루스] >sp Q2VPA6.2 HELQ_MOUSE 권장명: 전체 명칭 = 헬리카제 POLQ-유사; 별칭: 전체 명칭 = Mus308-유사 헬리카제; 별칭: 전체 명칭 = POLQ-유사 헬리카제 >gb AAI09171.2  헬리카제, mus308-유사 (드로소필라) [무스 무스쿨루스]
YP_003523727.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [시테록시단스 리토포피쿠스 (Sideroxydans lithotrophicus) ES-1] >gb ADE11340.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [시테록시단스 리토포피쿠스 ES-1]
XP_002120889.1	예상: DNA 헬리카제 HEL308와 유사 [시오나 인테스티날리스]
XP_001892566.1	유형 III 제한 효소, res 서브유닛 패밀리 단백질 [브루기아 말라이어 (Brugia malayi)] >gb EDP38603.1  유형 III 제한 효소, res 서브유닛 패밀리 단백질 [브루기아 말라이어]
ABZ09232.1	추정 헬리카제 보존형 C-말단 도메인 단백질 [미메양 해양 크렌코세균 HF4000 APKG7F11]
XP_002814981.1	예상: 저분절 단백질: 헬리카제 POLQ-유사 [폰고 아벨리이]
XP_002717082.1	예상: DNA 헬리카제 HEL308 [오리크톨리쿠스 쿠니쿨루스]
XP_001104832.1	예상: 헬리카제, POLQ-유사 [마카 무라타 (Macaca mulatta)]
AAI85274.1	DNA 헬리카제 HEL308 [호모 사피엔스]
NP_598375.2	헬리카제 POLQ-유사 [호모 사피엔스] >gb EAX05934.1  DNA 헬리카제 HEL308, 이소형 CRA_a [호모 사피엔스] >gb AAI41525.1  헬리카제, POLQ-유사 [합성 폴리뉴클레오타이드]
Q8TDG4.2	권장명: 전체 명칭 = 헬리카제 POLQ-유사; 별칭: 전체 명칭 = Mus308-유사 헬리카제; 별칭: 전체 명칭 = POLQ-유사 헬리카제
XP_003265889.1	예상: 헬리카제, POLQ [노마스쿠스 레우코게니스]

[0120]

XP_002745688.1	예상: 헬리카제 POLQ-유사 [칼리트릭스 자루스(Callithrix jacchus)]
XP_003310356.1	예상: 저분질 단백질: 헬리카제 POLQ-유사 [콰 트로글로디테스]
NP_001014156.2	헬리카제, POLQ-유사 [라투스 노르베기쿠스] >ref XP_001060858.1  예상: 헬리카제, POLQ-유사 [라투스 노르베기쿠스] >gb EDL99554.1  rCG37823, 이소형 CRA_c [라투스 노르베기쿠스]
XP_001850567.1	ATP-의존적 DNA 헬리카제 MER3 [클렉스 퀸퀘파시아투스(Culex quinquefasciatus)] >gb EDS32308.1  ATP-의존적 DNA 헬리카제 MER3 [클렉스 퀸퀘파시아투스]
XP_003427318.1	예상: 헬리카제 POLQ-유사 이소형 1 [나소니아 비트리켄니스]
XP_003143912.1	가설 단백질 LOAG_08332 [로아 로아(Loa loa)] >gb EFO20157.1  가설 단백질 LOAG_08332 [로아 로아]
CAG11187.1	무명 단백질 제품 [테트라온 니그로비리디스(Tetraodon nigroviridis)]
XP_001111254.2	예상: DNA 폴리머라제 세타 이소형 2 [마카 물라타]
XP_003414242.1	예상: 헬리카제 POLQ [록소돈타 아프리카나]
XP_002681870.1	예상 단백질 [나에글레리아 그루베리(Naegleria gruberi)] >gb EFC49126.1  예상 단백질 [나에글레리아 그루베리]
EAX05935.1	DNA 헬리카제 HEL308, 이소형 CRA_b [호도 사피엔스]
AAH59917.1	Ascc3 단백질 [무스 무스쿨루스]
ZP_07082808.1	DEAD/DEAH 박스 헬리카제 도메인 단백질 [스핑크박테리움 스피리터보룸(Sphingobacterium spiritivorum) ATCC 33861] >gb EFK55937.1  DEAD/DEAH 박스 헬리카제 도메인 단백질 [스핑크박테리움 스피리터보룸 ATCC 33861]
XP_001494572.3	예상: 저분질 단백질: 헬리카제 POLQ-유사 [에쿠우스 카발루스]
XP_002714920.1	예상: 활성화 신호 코인테그레이터(cointegrator) 1 복합체 서브유닛 3 [오리 크로라쿠스 쿠니쿨루스]
XP_002598278.1	가설 단백질 BRAFLDRAFT_204526 [브란키오스토마 플로리다에] >gb EEN54290.1  가설 단백질 BRAFLDRAFT_204526 [브란키오스토마 플로리다에]
XP_001943294.1	예상: 헬리카제 POLQ-유사 이소형 1 [아시르토시온 피숨] >ref XP_003240510.1  예상: 헬리카제 POLQ-유사 이소형 2 [아시르토시온 피숨]
XP_002803889.1	예상: 활성화 신호 코인테그레이터 1 복합체 서브유닛 3-유사 [마카 물라타]
XP_001651546.1	DNA 폴리머라제 세타 [아에테스 아에집티(Aedes aegypti)] >gb EAT42599.1  DNA 폴리머라제 세타 [아에테스 아에집티]
CAA11679.1	RNA 헬리카제 [호도 사피엔스]
XP_002837795.1	가설 단백질 [투베르 멜라노스포룸(Tuber melanosporum) Mel128] >emb CAZ81986.1  무명 단백질 제품 [투베르 멜라노스포룸]
EGT47882.1	가설 단백질 CAEBREN_02542 [카에노랍티디스 브렌네리(Caenorhabditis brenneri)]
EDL99655.1	활성화 신호 코인테그레이터 1 복합체 서브유닛 3 (예상), 이소형 CRA_b [라투스 노르베기쿠스]
NP_932124.2	활성화 신호 코인테그레이터 1 복합체 서브유닛 3 [무스 무스쿨루스]
EDL05054.1	mCG119534 [무스 무스쿨루스]
gi1352115865	DEAD/DEAH 박스 헬리카제 도메인 단백질 [나트린에마 펠리루브룸(Natrinema pellirubrum) DSM 15624]
ZP_08963952.1	

[0121]

[0122]

더욱 바람직하게는, Hel308 헬리카제는 하기 표 5에 제시된 헬리카제 또는 그의 변이체 중 하나이다. 더욱 바람직하게는, Hel308 헬리카제는 표 5에 제시된 헬리카제 중 하나의 서열, 즉 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 하나, 또는 그의 변이체를 포함한다.

표 5

더욱 바람직한 Hel308 헬리카제 및 가장 바람직한 Hel308 모티프 및 연장된 Hel308 모티프

서열 번호	헬리카제	명칭	% 동일성	% 동일성	Hel308 모티프	연장된 Hel308 모티프
			Hel308 Pfu	Hel308 Mbu		
10	Hel308 Mbu	메타노코코이데스 부르토니이	37%	-	QMAGRAGR (서열 11)	QMAGRAGRP (서열 12)
13	Hel308 Pfu	피로코쿠스 푸리오수스 DSM 3638	-	37%	QMLGRAGR (서열 14)	QMLGRAGRP (서열 15)
16	Hel308 Hvo	할로벡투스 불카나이	34%	41%	QMMGRAGR (서열 17)	QMMGRAGRP (서열 18)
19	Hel308 Hla	할로루브룸 라루스프로판디	35%	42%	QMCGRAGR (서열 20)	QMCGRAGRP (서열 21)
22	Hel308 Csy	세나르카에움 심비오움	34%	34%	QLCGRAGR (서열 23)	QLCGRAGRP (서열 24)
25	Hel308 Sso	솔폴루부스 솔파타리쿠스	35%	33%	QMSGRAGR (서열 26)	QMSGRAGRP (서열 27)
28	Hel308 Mfr	메타노게니움 프리기둠	37%	44%	QMAGRAGR (서열 11)	QMAGRAGRP (서열 12)
29	Hel308 Mok	메타노썬코코쿠스 오키나헨시스	37%	34%	QCIGRAGR (서열 30)	QCIGRAGRP (서열 31)
32	Hel308 Mig	메타노토리스 이그네우스 Kol 5	40%	35%	QCIGRAGR (서열 30)	QCIGRAGRP (서열 31)
33	Hel308 Tga	썬코코쿠스 감마몰레란스 EJ3	60%	38%	QMMGRAGR (서열 17)	QMMGRAGRP (서열 18)
34	Hel308 Tba	썬코코쿠스 바르필루스 MP	57%	35%	QMIGRAGR (서열 35)	QMIGRAGRP (서열 36)
37	Hel308 Tsi	썬코코쿠스 시비리쿠스 MM 739	56%	35%	QMMGRAGR (서열 17)	QMMGRAGRP (서열 18)
38	Hel308 Mba	메타노사르시나 바르케리 군주 푸사로	39%	60%	QMAGRAGR (서열 11)	QMAGRAGRP (서열 12)
39	Hel308 Mac	메타노사르시나 아세티보란스	38%	60%	QMAGRAGR (서열 11)	QMAGRAGRP (서열 12)
40	Hel308 Mmah	메타노할로필루스 마하이 DSM 5219	38%	60%	QMAGRAGR (서열 11)	QMAGRAGRP (서열 12)

[0123]

41	Hel308 Mmaz	메타노사르시나 마제이	38%	60%	QMAGRAGR (서열 11)	QMAGRGRP (서열 12)
42	Hel308 Mth	메타노사에타 썬모필라 PT	39%	46%	QMAGRAGR (서열 11)	QMAGRGRP (서열 12)
43	Hel308 Mzh	메타노살숨 질리나에 DSM 4017	39%	57%	QMAGRAGR (서열 11)	QMAGRGRP (서열 12)
44	Hel308 Mev	메타노할로비움 에베스티가톨 Z-7303	38%	61%	QMAGRAGR (서열 11)	QMAGRGRP (서열 12)
45	Hel308 Mma	메타노코쿠스 마리팔루디스	36%	32%	QCIGRAGR (서열 30)	QCIGRGRP (서열 31)
46	Hel308 Nma	나트리알바 마가다이	37%	43%	QMMGRAGR (서열 17)	QMMGRGRP (서열 18)
47	Hel308 Mbo	메타노레콜라 보오네이 6A8	38%	45%	QMAGRAGR (서열 11)	QMAGRGRP (서열 12)
48	Hel308 Fac	페로플라스마 아시다르마누스	34%	32%	QMIGRAGR (서열 35)	QMIGRGRP (서열 36)
49	Hel308 Mfc	메타노칼도코쿠스 페르벤스 AG86	40%	35%	QCIGRAGR (서열 30)	QCIGRGRP (서열 31)
50	Hel308 Mja	메타노칼도코쿠스 잔나쉬이	24%	22%	QCIGRAGR (서열 30)	QCIGRGRP (서열 31)
51	Hel308 Min	메타노칼도코쿠스 인페르누스	41%	33%	QCIGRAGR (서열 30)	QCIGRGRP (서열 31)
52	Hel308 Mhu	메타노스피릴룸 훈가테이 JF-1	36%	40%	QMAGRAGR (서열 11)	QMAGRGRP (서열 12)
53	Hel308 Afu	아르카에오글로부스 풀기두스 DSM 4304	40%	40%	QMAGRAGR (서열 11)	QMAGRGRP (서열 12)
54	Hel308 Htu	할로테리게나 푸르크메니카	35%	43%	QMAGRAGR (서열 11)	QMMGRGRP (서열 12)
55	Hel308 Hpa	할라다프타투스 포시할로필루스 DX253	38%	45%	QMFGRAGR (서열 56)	QMFGRGRP (서열 57)
58	ski2-유사 헬리카제	할로박테리움 종 NRC-1	36.8%	42.0%	QMFGRAGR (서열 56)	QMFGRGRP (서열 57)

[0124]

[0125]

더욱 바람직하게는, Hel308 헬리카제는 (a) Hel308 Mbu의 서열 (즉 서열 10) 또는 그의 변이체, (b) Hel308 Pfu의 서열 (즉 서열 13) 또는 그의 변이체, (c) Hel308 Mok의 서열 (즉 서열 29) 또는 그의 변이체, (d) Hel308 Mma의 서열 (즉 서열 45) 또는 그의 변이체, (e) Hel308 Fac의 서열 (즉 서열 48) 또는 그의 변이체, 또는 (f) Hel308 Mhu의 서열 (즉 서열 52) 또는 그의 변이체를 포함한다. 더욱 바람직하게는, Hel308 헬리카제는 서열 10에 제시된 서열 또는 그의 변이체를 포함한다.

[0126]

더욱 바람직하게는, Hel308 헬리카제는 (a) Hel308 Tga의 서열 (즉 서열 33) 또는 그의 변이체, (b) Hel308 Csy의 서열 (즉 서열 22) 또는 그의 변이체, 또는 (c) Hel308 Mhu의 서열 (즉 서열 52) 또는 그의 변이체를 포함한다. 더욱 바람직하게는, Hel308 헬리카제는 서열 33에 제시된 서열 또는 그의 변이체를 포함한다.

[0127]

Hel308 헬리카제의 변이체는 아미노산 서열이 야생형 헬리카제의 것과 다르고 폴리뉴클레오티드 결합 활성을 보유하는 효소이다. 특히, 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 어느 하나의 변이체는 아미노산 서열이 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 어느 하나의 것과 다르고 폴리뉴클레오티드 결합 활성을 보유하는 효소이다. 서열 10 또는 33의 변이체는 아미노산 서열이 서열 10 또는 33의 것과 다르고 폴리뉴클레오티드 결합 활성을 보유하는 효소이다. 변이체는 헬리카제 활성을 보유한다. 변이체는 하기에서 논의되는 2가지 양식 중 하나 이상으로 작동하여야 한다. 바람직하게는, 변이체는 양쪽 방식으로 작동한다. 변이체는 헬리카제를 코딩하는 폴리뉴클레오티드의 취급을 용이하게 하고/하거나 높은 염 농도 및/또는 실온에서의 이의 활성을 용이하게 하는 변형을 포



함할 수 있다. 전형적으로, 변이체는 상기 논의된 He1308 모티프 또는 연장된 He1308 모티프 외부의 영역에서 야생형 헬리카제와 상이하다. 그러나, 변이체는 이러한 모티프(들) 내부의 변형을 포함할 수 있다.

[0128] 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 어느 하나, 예컨대 서열 10 또는 33의 아미노산 서열의 전체 길이에 걸쳐, 변이체는 바람직하게는 아미노산 동일성을 기초로 이러한 서열에 대해 적어도 30% 상동성일 것이다. 더욱 바람직하게는, 변이체 폴리펩티드는 전체 서열에 걸쳐 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 어느 하나, 예컨대 서열 10 또는 33의 아미노산 서열에 대해 아미노산 동일성을 기초로 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 더욱 바람직하게는 적어도 95%, 97% 또는 99% 상동성일 수 있다. 150개 이상, 예를 들어 200개, 300개, 400개, 500개, 600개, 700개, 800개, 900개 또는 1000개 또는 이를 초과하는 개수의 연속적인 아미노산의 신장물에 걸쳐 아미노산 동일성이 적어도 70%, 예를 들어 적어도 80%, 적어도 85%, 적어도 90%, 또는 적어도 95%일 수 있다 ("강한 상동성"). 상기 기술된 바와 같이 상동성이 결정된다. 변이체는 서열 2 및 4와 관련하여 상기에서 논의된 것 중 중 임의의 방식으로 야생형 서열과 상이할 수 있다.

[0129] 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 어느 하나의 변이체는 바람직하게는 관련된 야생형 서열의 He1308 모티프 또는 연장된 He1308 모티프를 포함한다. 예를 들어, 서열 10의 변이체는 바람직하게는 서열 10의 He1308 모티프 (QMAGRAGR; 서열 11) 또는 서열 10의 연장된 He1308 모티프 (QMAGRGRP; 서열 12)를 포함한다. 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 각각의 He1308 모티프 및 연장된 He1308 모티프가 표 5에서 제시된다. 그러나, 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 어느 하나의 변이체가 상이한 야생형 서열로부터의 He1308 모티프 또는 연장된 He1308 모티프를 포함할 수 있다. 예를 들어, 서열 10의 변이체가 서열 13의 He1308 모티프 (QMLGRAGR; 서열 14) 또는 서열 13의 연장된 He1308 모티프 (QMLGRGRP; 서열 15)를 포함할 수 있다. 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 어느 하나의 변이체는 표 5에 제시된 바람직한 모티프 중 어느 하나를 포함할 수 있다. 서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55 및 58 중 어느 하나의 변이체는 관련된 야생형 서열의 He1308 모티프 또는 연장된 He1308 모티프 내에 변형을 포함할 수도 있다. X1 및 X2에서의 적절한 변형은 이러한 2개의 모티프를 정의할 때 상기에서 논의되어 있다.

[0130] 서열 10의 변이체는 서열 10의 최초의 19개의 아미노산이 결여될 수 있고/있거나 서열 10의 마지막 33개의 아미노산이 결여될 수 있다. 바람직하게는 서열 10의 변이체는 서열 10의 아미노산 20 내지 211 또는 20 내지 727 과 아미노산 동일성을 기초로 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90% 또는 더욱 바람직하게는 적어도 95%, 적어도 97% 또는 적어도 99% 상동성인 서열을 포함한다.

[0131] 헬리카제가 포어에 공유적으로 부착될 수 있다. 바람직하게는, 헬리카제는 포어에 공유적으로 부착되지 않는다. 전형적으로, 포어 및 헬리카제에 전압을 인가하면, 표적 폴리뉴클레오티드를 서열분석할 수 있는 센서의 형성이 초래된다. 이는 하기에서 더욱 상세하게 논의된다.

[0132] 본원에서 기술된 단백질 중 임의의 것, 즉 막횡단 단백질 포어 또는 He1308 헬리카제는 이들의 확인 또는 정제를 보조하도록, 예를 들어 히스티딘 잔기 (his 태그), 아스파르트산 잔기 (asp 태그), 스트렙타비딘 태그, 플래그 태그, SUMO 태그, GST 태그 또는 MBP 태그의 부가에 의해, 또는 세포로부터 이들의 분비를 촉진하기 위한 신호 서열의 부가에 의해 (폴리펩티드가 이같은 서열을 천연적으로 함유하지 않는 경우), 변형될 수 있다. 유전학적 태그를 도입하는 것에 대한 대안은 태그를 포어 또는 헬리카제 상의 천연 또는 조작 위치 상에서 화학적으로 반응시키는 것이다. 이는 예를 들어 펩티드 시약을 포어의 외부 상에 조작된 시스템에 반응시키는 것일 것이다. 이는 용혈소 이중-올리고머들을 분리하기 위한 방법으로서 입증되었다 (문헌 [Chem Biol. 1997 Jul;4(7):497-505]).

[0133] 포어 및/또는 헬리카제가 표식 표지로 표지될 수 있다. 표식 표지는 포어가 검출되게 하는 임의의 적절한 표지일 수 있다. 적절한 표지에는 형광 분자, 방사성 동위원소, 예를 들어 <sup>125</sup>I, <sup>35</sup>S, 효소, 항체, 항원, 폴리뉴클레오티드 및 리간드 예컨대 비오틴이 포함되지만, 이에 한정되지는 않는다.

[0134] 합성에 의해 또는 재조합 수단에 의해 단백질이 제조될 수 있다. 예를 들어, 시험관내 번역 및 전사 (IVTT)에

의해 포어 및/또는 헬리카제가 합성될 수 있다. 포어 및/또는 헬리카제의 아미노산 서열이 비-천연 발생 아미노산을 포함하도록 또는 단백질의 안정성을 증가시키도록 변형될 수 있다. 단백질이 합성 수단에 의해 생산될 때, 생산 동안 이같은 아미노산들이 도입될 수 있다. 합성 또는 재조합 생산 후에 포어 및/또는 헬리카제가 변경될 수도 있다.

[0135] D-아미노산을 사용하여 포어 및/또는 헬리카제가 생산될 수도 있다. 예를 들어, 포어 또는 헬리카제가 L-아미노산 및 D-아미노산의 혼합물을 포함할 수 있다. 이는 이같은 단백질 또는 펩티드를 생산하기 위한 업계에서 통상적이다.

[0136] 포어 및/또는 헬리카제는 다른 비-특이적 변형을 함유할 수도 있고, 단 이들은 포어 형성 또는 헬리카제 기능을 방해하지 않는다. 다수의 비-특이적 측쇄 변형이 업계에 공지되어 있고, 이러한 단백질(들)의 측쇄에 이루어질 수 있다. 이같은 변형에는, 예를 들어 알데히드와의 반응에 이어지는  $\text{NaBH}_4$ 로의 환원에 의한 아미노산의 환원성 알킬화, 메틸아세티미데이트로의 아미딘화, 또는 아세트산 무수물로의 아실화가 포함된다.

[0137] 업계에 공지된 표준 방법을 사용하여 포어 및 헬리카제가 생산될 수 있다. 포어 또는 헬리카제를 코딩하는 폴리뉴클레오티드 서열이 업계에서의 표준 방법을 사용하여 유도 및 복제될 수 있다. 포어 또는 헬리카제를 코딩하는 폴리뉴클레오티드 서열이 업계에서의 표준 기술을 사용하여 박테리아 숙주 세포에서 발현될 수 있다. 포어 및/또는 헬리카제가 재조합 발현 벡터로부터 폴리펩티드의 계내 발현에 의해 세포에서 생산될 수 있다. 임의로 발현 벡터는 폴리펩티드의 발현을 제어하기 위한 유도성 프로모터를 보유한다. 이러한 방법들이 문헌 [Sambrook, J. and Russell, D. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY]에 기술되어 있다.

[0138] 단백질 생산 유기체로부터의 임의의 단백질 액체 크로마토그래피에 의한 정제 후에, 또는 재조합 발현 후에 대규모로 포어 및/또는 헬리카제가 생산될 수 있다. 전형적인 단백질 액체 크로마토그래피 시스템에는 FPLC, AKTA 시스템, 바이오-캐드(Bio-Cad) 시스템, 바이오-래드(Bio-Rad)의 바이오로직(BioLogic) 시스템 및 길슨(Gilson)의 HPLC 시스템이 포함된다.

[0139] 본 발명의 방법은 표적 폴리뉴클레오티드의 하나 이상의 특성을 측정하는 것을 수반한다. 이러한 방법은 표적 폴리뉴클레오티드의 2개, 3개, 4개 또는 5개 또는 이를 초과하는 개수의 특성을 측정하는 것을 수반할 수 있다. 하나 이상의 특성은 바람직하게는 (i) 표적 폴리뉴클레오티드의 길이, (ii) 표적 폴리뉴클레오티드의 신원, (iii) 표적 폴리뉴클레오티드의 서열, (iv) 표적 폴리뉴클레오티드의 2차 구조, 및 (v) 표적 폴리뉴클레오티드가 변형되었는지 또는 그렇지 않은지 여부로부터 선택된다. (i) 내지 (v)의 임의의 조합이 본 발명에 따라 측정될 수 있다.

[0140] (i)의 경우, 표적 폴리뉴클레오티드와 포어 사이의 상호작용의 수를 사용하여 표적 폴리뉴클레오티드의 길이가 측정될 수 있다.

[0141] (ii)의 경우, 다수의 방식으로 폴리뉴클레오티드의 신원이 측정될 수 있다. 표적 폴리뉴클레오티드의 서열을 측정하면서 또는 표적 폴리뉴클레오티드의 서열을 측정하지 않으면서 폴리뉴클레오티드의 신원이 측정될 수 있다. 전자는 직접적이다; 폴리뉴클레오티드가 서열분석되고, 이에 의해 확인된다. 후자는 여러 방식으로 행해질 수 있다. 예를 들어, 폴리뉴클레오티드 내의 특정 모티프의 존재가 측정될 수 있다 (폴리뉴클레오티드의 나머지 서열을 측정하지 않으면서). 대안적으로, 방법에서의 특정 전기 및/또는 광학 신호의 측정으로 표적 폴리뉴클레오티드가 특정 공급원으로부터 유래되는 것으로서 확인될 수 있다.

[0142] (iii)의 경우, 폴리뉴클레오티드의 서열이 기존에 기술된 바와 같이 결정될 수 있다. 적절한 서열분석 방법, 특히 전기 측정을 사용하는 것이 문헌 [Stoddart D et al., *Proc Natl Acad Sci*, 12;106(19):7702-7], [Lieberman KR et al, *J Am Chem Soc*. 2010;132(50):17961-72], 및 국제 출원 WO 2000/28312에 기술되어 있다.

[0143] (iv)의 경우, 2차 구조가 다양한 방식으로 측정될 수 있다. 예를 들어, 방법이 전기 측정을 수반하면, 체류 시간에서의 변화 또는 포어를 통과하여 흐르는 전류의 변화를 사용하여 2차 구조가 측정될 수 있다. 이는 단일-가닥 및 이중-가닥 폴리뉴클레오티드의 영역이 구별되도록 허용한다.

[0144] (v)의 경우, 임의의 변형의 존재 또는 부재가 측정될 수 있다. 바람직하게는, 방법은 표적 폴리뉴클레오티드가 메틸화에 의해, 산화에 의해, 손상에 의해, 하나 이상의 단백질로, 또는 하나 이상의 표지, 태그 또는 스페이서로 변형되었는지 또는 그렇지 않은지 여부를 결정하는 것을 포함한다. 특정 변형은 포어와의 특정 상호작용을 초래할 것이고, 하기 기술된 방법을 사용하여 이를 측정할 수 있다. 예를 들어, 메틸시토신이 이의 각각의 뉴

클레오티드와의 상호작용 동안 포어를 통해 흐르는 전류를 기초로 시토신과 구별될 수 있다.

- [0145] 다양한 상이한 유형의 측정이 이루어질 수 있다. 이는 비제한적으로 전기 측정 및 광학 측정을 포함한다. 가능한 전기 측정에는 전기 측정, 임피던스(impedance) 측정, 터널링(tunnelling) 측정 (문헌 [Ivanov AP et al., Nano Lett. 2011 Jan 12;11(1):279-85]), 및 FET 측정 (국제 출원 WO 2005/124888)이 포함된다. 광학 측정이 전기 측정과 조합될 수 있다 (문헌 [Soni GV et al., Rev Sci Instrum. 2010 Jan;81(1):014301]). 측정은 막횡단 전류 측정 예컨대 포어를 통해 흐르는 이온 전류의 측정일 수 있다.
- [0146] 문헌 [Stoddart D et al., Proc Natl Acad Sci, 12;106(19):7702-7], [Lieberman KR et al, J Am Chem Soc. 2010;132(50):17961-72], 및 국제 출원 WO-2000/28312에 기술된 바와 같이 표준 단일 채널 기록 장치를 사용하여 전기 측정이 이루어질 수 있다. 대안적으로, 예를 들어 국제 출원 WO-2009/077734 및 국제 출원 WO-2011/067559에 기술된 바와 같이, 다중 채널 시스템을 사용하여 전기 측정이 이루어질 수 있다.
- [0147] 바람직한 실시양태에서, 방법은 하기를 포함한다:
- [0148] (a) 표적 폴리뉴클레오티드를 막횡단 포어 및 He1308 헬리카제와 접촉시켜, 헬리카제가 포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하고 표적 뉴클레오티드 내의 뉴클레오티드들이 포어와 상호작용하게 하는 단계; 및
- [0149] (b) 하나 이상의 상호작용 동안 포어를 통과하는 전류를 측정하여 표적 폴리뉴클레오티드의 하나 이상의 특성을 측정하고, 이에 의해 표적 폴리뉴클레오티드를 특성화하는 단계.
- [0150] 포어가 막 내로 삽입되어 있는 막/포어 시스템을 연구하는데 적절한 임의의 장치를 사용하여 방법이 수행될 수 있다. 막횡단 포어 센싱에 적절한 임의의 장치를 사용하여 방법이 수행될 수 있다. 예를 들어, 장치는 수성 용액 및 챔버를 2개의 섹션으로 나누는 장벽을 포함하는 챔버를 포함한다. 장벽에는 포어를 함유하는 막이 형성되는 개구부가 있다.
- [0151] 국제 출원 번호 PCT/GB08/000562 (WO 2008/102120)에 기술된 장치를 사용하여 방법이 수행될 수 있다.
- [0152] 방법은 뉴클레오티드(들)과의 하나 이상의 상호작용 동안 포어를 통과하는 전류를 측정하는 것을 수반할 수 있다. 따라서, 전위를 인가할 수 있고 막 및 포어를 가로지르는 전기 신호를 측정할 수 있는 전기 회로를 장치가 또한 포함할 수 있다. 패치 클램프 또는 전압 클램프를 사용하여 방법이 수행될 수 있다. 바람직하게는, 방법은 전압 클램프의 사용을 수반한다.
- [0153] 본 발명의 방법은 뉴클레오티드와의 하나 이상의 상호작용 동안 포어를 통과하는 전류를 측정하는 것을 수반할 수 있다. 막횡단 단백질 포어를 통한 이온 전류를 측정하는데 적절한 조건이 업계에 공지되어 있고, 실시예에서 개시된다. 전형적으로, 막 및 포어를 가로질러 인가되는 전압으로 방법이 수행된다. 사용 전압은 전형적으로 +2 V 내지 -2 V, 전형적으로 -400 mV 내지 +400 mV이다. 바람직하게는 사용 전압은 하한이 -400 mV, -300 mV, -200 mV, -150 mV, -100 mV, -50 mV, -20mV 및 0 mV로부터 선택되고 상한이 독립적으로 +10 mV, + 20 mV, +50 mV, +100 mV, +150 mV, +200 mV, +300 mV 및 +400 mV로부터 선택되는 범위 내이다. 더욱 바람직하게는 사용 전압은 100 mV 내지 240 mV 범위이고, 가장 바람직하게는 120 mV 내지 220 mV의 범위이다. 증가된 인가 전위를 사용함으로써 포어에 의한 상이한 뉴클레오티드들 간의 구별을 증가시킬 수 있다.
- [0154] 전형적으로, 임의의 전하 운반체, 예컨대 금속 염, 예를 들어 알칼리 금속 염, 할라이드 염, 예를 들어 클로라이드 염, 예컨대 알칼리 금속 클로라이드 염의 존재 하에 방법이 수행된다. 전하 운반체는 이온성 액체 또는 유기 염, 예를 들어 테트라메틸 암모늄 클로라이드, 트리메틸페닐 암모늄 클로라이드, 페닐트리메틸 암모늄 클로라이드, 또는 1-에틸-3-메틸 이미다졸륨 클로라이드를 포함할 수 있다. 상기 논의된 예시적인 장치에서, 염은 챔버 내의 수성 용액 내에 존재한다. 염화칼륨 (KCl), 염화나트륨 (NaCl) 또는 염화세슘 (CsCl)이 전형적으로 사용된다. KCl이 바람직하다. 염 농도는 포화될 수 있다. 염 농도는 3M 이하일 수 있고, 전형적으로는 0.1 내지 2.5 M, 0.3 내지 1.9 M, 0.5 내지 1.8 M, 0.7 내지 1.7 M, 0.9 내지 1.6 M 또는 1 M 내지 1.4 M이다. 염 농도는 바람직하게는 150 mM 내지 1 M이다. 상기 논의된 바와 같이, He1308 헬리카제는 놀랍게도 높은 염 농도에서 작동한다. 바람직하게는, 적어도 0.3 M, 예컨대 적어도 0.4 M, 적어도 0.5 M, 적어도 0.6 M, 적어도 0.8 M, 적어도 1.0 M, 적어도 1.5 M, 적어도 2.0 M, 적어도 2.5 M 또는 적어도 3.0 M의 염 농도를 사용하여 방법이 수행된다. 높은 염 농도는 높은 신호 대 노이즈 비를 제공하고, 정상적인 전류 변동의 배경값에 대비하여 뉴클레오티드의 존재를 지시하는 전류가 확인되게 한다.
- [0155] 전형적으로, 완충제의 존재 하에 방법이 수행된다. 상기 논의된 예시적인 장치에서, 완충제는 챔버 내의 수성 용액 내에 존재한다. 임의의 완충제가 본 발명의 방법에서 사용될 수 있다. 전형적으로, 완충제는 HEPES이다.

또 다른 적절한 완충제는 트리스-HCl 완충제이다. 전형적으로, 4.0 내지 12.0, 4.5 내지 10.0, 5.0 내지 9.0, 5.5 내지 8.8, 6.0 내지 8.7 또는 7.0 내지 8.8 또는 7.5 내지 8.5의 pH에서 방법이 수행된다. 사용된 pH는 바람직하게는 약 7.5이다.

[0156] 0℃ 내지 100℃, 15℃ 내지 95℃, 16℃ 내지 90℃, 17℃ 내지 85℃, 18℃ 내지 80℃, 19℃ 내지 70℃, 또는 20℃ 내지 60℃에서 방법이 수행될 수 있다. 전형적으로는 실온에서 방법이 수행된다. 임의로는, 효소 기능을 지지하는 온도, 예컨대 약 37℃에서 방법이 수행된다.

[0157] 전형적으로, 유리 뉴클레오타이드 또는 유리 뉴클레오타이드 유사체 및 헬리카제의 작용을 용이하게 하는 효소 보조 인자의 존재 하에 방법이 수행된다. 유리 뉴클레오타이드는 하나 이상의 임의의 상기 논의된 개별적인 뉴클레오타이드일 수 있다. 유리 뉴클레오타이드는 아데노신 모노포스페이트 (AMP), 아데노신 디포스페이트 (ADP), 아데노신 트리포스페이트 (ATP), 구아노신 모노포스페이트 (GMP), 구아노신 디포스페이트 (GDP), 구아노신 트리포스페이트 (GTP), 티미딘 모노포스페이트 (TMP), 티미딘 디포스페이트 (TDP), 티미딘 트리포스페이트 (TTP), 우리딘 모노포스페이트 (UMP), 우리딘 디포스페이트 (UDP), 우리딘 트리포스페이트 (UTP), 시티딘 모노포스페이트 (CMP), 시티딘 디포스페이트 (CDP), 시티딘 트리포스페이트 (CTP), 시클릭 아데노신 모노포스페이트 (cAMP), 시클릭 구아노신 모노포스페이트 (cGMP), 데옥시아데노신 모노포스페이트 (dAMP), 데옥시아데노신 디포스페이트 (dADP), 데옥시아데노신 트리포스페이트 (dATP), 데옥시구아노신 모노포스페이트 (dGMP), 데옥시구아노신 디포스페이트 (dGDP), 데옥시구아노신 트리포스페이트 (dGTP), 데옥시티미딘 모노포스페이트 (dTMP), 데옥시티미딘 디포스페이트 (dTDP), 데옥시티미딘 트리포스페이트 (dTTP), 데옥시우리딘 모노포스페이트 (dUMP), 데옥시우리딘 디포스페이트 (dUDP), 데옥시우리딘 트리포스페이트 (dUTP), 데옥시시티딘 모노포스페이트 (dCMP), 데옥시시티딘 디포스페이트 (dCDP) 및 데옥시시티딘 트리포스페이트 (dCTP)를 포함하지만, 이에 한정되지는 않는다. 유리 뉴클레오타이드는 바람직하게는 AMP, TMP, GMP, CMP, UMP, dAMP, dTMP, dGMP 또는 dCMP로부터 선택된다. 유리 뉴클레오타이드는 바람직하게는 아데노신 트리포스페이트 (ATP)이다. 효소 보조인자는 헬리카제가 기능하게 하는 인자이다. 효소 보조인자는 바람직하게는 2가 금속 양이온이다. 2가 금속 양이온은 바람직하게는  $Mg^{2+}$ ,  $Mn^{2+}$ ,  $Ca^{2+}$  또는  $Co^{2+}$ 이다. 효소 보조인자는 가장 바람직하게는  $Mg^{2+}$ 이다.

[0158] 표적 폴리뉴클레오타이드는 임의의 순서로 He1308 헬리카제 및 포어와 접촉될 수 있다. 표적 폴리뉴클레오타이드가 He1308 헬리카제 및 포어와 접촉될 때, 표적 폴리뉴클레오타이드가 먼저 헬리카제와 복합체를 형성하는 것이 바람직하다. 포어를 가로질러 전압이 인가되면, 표적 폴리뉴클레오타이드/헬리카제 복합체가 포어와 복합체를 형성하고, 포어를 통한 폴리뉴클레오타이드의 이동을 제어한다.

[0159] 상기 논의된 바와 같이, He1308 헬리카제는 나노포어와 관련하여 2가지 방식으로 작동할 수 있다. 첫째로, He1308 헬리카제가 인가된 전압으로부터 초래되는 장과 함께 포어를 통해 표적 서열을 이동시키도록 He1308 헬리카제를 사용하여 방법이 바람직하게 수행된다. 이러한 방식에서, DNA의 3' 끝부분이 먼저 나노포어에 포획되고, 최종적으로 표적 서열이 이중층의 트랜스 측면으로 전위될 때까지 장과 함께 표적 서열이 나노포어를 통과하도록 효소가 DNA를 나노포어 내로 이동시킨다. 대안적으로, 효소가 인가된 전압으로부터 초래되는 장에 대항하여 포어를 통해 표적 서열을 이동시키도록 방법이 바람직하게 수행된다. 이러한 방식에서, DNA의 5' 끝부분이 먼저 나노포어에 포획되고, 최종적으로 이중층의 시스 측면으로 되돌려 사출될 때까지 표적 서열이 인가된 장에 대항하여 나노포어 밖으로 잡아당겨 지도록 효소가 DNA를 나노포어를 통해 이동시킨다.

[0160] 가장 바람직하게는, 본 발명의 방법은 MspA로부터 유래된 포어 및 서열 8 또는 10에 제시된 서열 또는 그의 변이체를 포함하는 헬리카제를 수반한다. MspA 및 서열 8 및 10과 관련하여 상기 논의된 실시양태들 중 임의의 것이 조합되어 사용될 수 있다.

[0161] 기타 방법

[0162] 본 발명은 표적 폴리뉴클레오타이드를 특성화하기 위한 센서를 형성시키는 방법을 또한 제공한다. 이러한 방법은 포어와 He1308 헬리카제 사이에 복합체를 형성시키는 단계를 포함한다. 포어와 헬리카제를 표적 폴리뉴클레오타이드의 존재 하에 접촉시키고, 그 후 포어를 가로질러 전위를 인가함으로써 복합체가 형성될 수 있다. 인가된 전위는 상기 기술된 바와 같이 화학 전위 또는 전압 전위일 수 있다. 대안적으로, 포어를 헬리카제에 공유적으로 부착시킴으로써 복합체가 형성될 수 있다. 공유 부착을 위한 방법이 업계에 공지되어 있고, 예를 들어 국제출원 번호 PCT/GB09/001679 (WO 2010/004265로 공개됨) 및 PCT/GB10/000133 (WO 2010/086603로 공개됨)에 개시되어 있다. 복합체는 표적 폴리뉴클레오타이드를 특성화하기 위한 센서이다. 바람직하게는, 방법은 Msp로부터 유래된 포어와 He1308 헬리카제 사이에 복합체를 형성시키는 단계를 포함한다. 본 발명의 방법과 관련하여 상



기에서 논의된 실시양태 중 임의의 것이 이러한 방법에 동등하게 적용된다.

[0163] 키트

[0164] 본 발명은 표적 폴리뉴클레오티드를 특성화하기 위한 키트를 또한 제공한다. 이러한 키트는 (a) 포어 및 (b) He1308 헬리카제를 포함한다. 본 발명의 방법과 관련하여 상기 논의된 실시양태 중 임의의 것이 이러한 키트에 동등하게 적용된다.

[0165] 키트는 막의 성분, 예컨대 지질 이중층을 형성하는데 필요한 인지질을 추가로 포함할 수 있다.

[0166] 본 발명의 키트는 상기 언급된 실시양태 중 임의의 것이 수행될 수 있게 하는 하나 이상의 다른 시약 또는 기기를 추가적으로 포함할 수 있다. 이같은 시약 또는 기기에는 하기의 것 중 하나 이상이 포함된다: 적절한 완충제(들) (수성 용액), 대상으로부터 샘플을 수득하는 수단 (예컨대 바늘을 포함하는 기기 또는 용기), 폴리뉴클레오티드를 증폭 및/또는 발현시키는 수단, 상기 정의된 바와 같은 막, 또는 전압 또는 패치 클램프 장치. 유체 샘플에 시약이 재현탁되도록, 시약이 건조 상태로 키트 내에 존재할 수 있다. 임의로, 키트는 본 발명의 방법에서 키트가 사용될 수 있게 하는 설명서 또는 어떤 환자를 위해 방법이 사용될 수 있는지에 관한 상세사항을 또한 포함할 수 있다. 키트는 임의로 뉴클레오티드를 포함할 수 있다.

[0167] 장치

[0168] 본 발명은 표적 폴리뉴클레오티드를 특성화하기 위한 장치를 또한 제공한다. 이러한 장치는 다수의 포어 및 다수의 He1308 헬리카제를 포함한다. 바람직하게는 장치는 본 발명의 방법을 수행하기 위한 설명서를 추가로 포함한다. 장치는 폴리뉴클레오티드 분석을 위한 임의의 통상적인 장치, 예컨대 어레이 또는 칩일 수 있다. 본 발명의 방법과 관련하여 상기 논의된 실시양태 중 임의의 것이 본 발명의 장치에 동등하게 적용가능하다.

[0169] 바람직하게는 장치는 본 발명의 방법을 수행하도록 설정된다.

[0170] 바람직하게는 장치는

[0171] 막 및 다수의 포어를 지지할 수 있고, 포어 및 헬리카제를 사용하여 폴리뉴클레오티드 특성화를 수행하도록 작동할 수 있는 센서 기구;

[0172] 특성화를 수행하기 위해 물질을 보유하기 위한 하나 이상의 저장소;

[0173] 물질을 하나 이상의 저장소로부터 센서 기구로 제어가능하게 공급하도록 구성된 유체공학 시스템; 및

[0174] 각각의 샘플을 수용하기 위한 다수의 용기

[0175] 를 포함하며, 유체공학 시스템은 샘플을 용기로부터 센서 기구로 선택적으로 공급하도록 구성된 것인, 분석 장치를 포함한다. 장치는 국제 출원 번호 PCT/GB08/004127 (WO 2009/077734로 공개됨), PCT/GB10/000789 (WO 2010/122293로 공개됨), 국제 출원 번호 PCT/GB10/002206 (아직 공개되지 않음) 또는 국제 출원 번호 PCT/US99/25679 (WO 00/28312로 공개됨)에 기술된 것들 중 임의의 것일 수 있다.

[0176] 내부 결합 분자 모터

[0177] 나노포어를 통한 중합체, 특히 폴리뉴클레오티드의 전위를 제어하기 위한 수단으로서 분자 모터가 통상적으로 사용된다. 놀랍게도, 본 발명자들은 분자 모터가 나노포어를 통한 폴리뉴클레오티드의 전위를 제어함에 따라 내부 뉴클레오티드, 즉 5' 또는 3' 말단 뉴클레오티드 이외의 위치에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터가 폴리뉴클레오티드의 증가된 판독 길이를 제공할 수 있다는 것을 발견하였다. 분자 모터의 제어 하에 나노포어를 통해 전체 폴리뉴클레오티드를 전위시키는 능력은 폴리뉴클레오티드의 특성, 예컨대 이의 서열이 공지된 방법에 비해 개선된 정확도 및 속도로 추정되게 한다. 이는 가닥 길이가 증가되고 분자 모터가 개선된 프로세스성과 함께 요구됨에 따라 더욱 중요해진다. 본 발명에서 사용되는 분자 모터는 뉴클레오티드 500개 이상, 예를 들어 뉴클레오티드 1000개, 5000개, 10000개 또는 20000개 또는 이를 초과하는 개수의 표적 폴리뉴클레오티드의 전위를 제어하는데 특히 효과적이다.

[0178] 따라서 본 발명은

[0179] (a) 표적 폴리뉴클레오티드를 막횡단 포어, 및 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터와 접촉시켜, 분자 모터가 포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하고 표적 뉴클레오티드 내의 뉴클레오티드들이 포어와 상호작용하게 하는 단계; 및

- [0180] (b) 하나 이상의 상호작용 동안 표적 폴리뉴클레오타이드의 하나 이상의 특성을 측정하고, 이에 의해 표적 폴리뉴클레오타이드를 특성화하는 단계
- [0181] 를 포함하는, 표적 폴리뉴클레오타이드를 특성화하는 방법을 제공한다.
- [0182] 본 발명의 He1308 방법과 관련하여 상기 논의된 실시양태 중 임의의 것이 본 발명의 이러한 방법에 동등하게 적용된다.
- [0183] 폴리뉴클레오타이드, 특히 뉴클레오타이드 500개 이상의 폴리뉴클레오타이드를 서열분석하는 것에서 발생하는 문제는 폴리뉴클레오타이드의 전위를 제어하는 분자 모터가 폴리뉴클레오타이드로부터 탈착될 수 있다는 것이다. 이는 폴리뉴클레오타이드가 인가된 장의 방향으로 포어를 통해 신속하게, 그리고 제어되지 않은 방식으로 잡아당겨 지도록 한다. 본 발명에서 사용되는 분자 모터의 여러 사례는 폴리뉴클레오타이드에 비교적 짧은 간격으로 결합하고, 따라서 추가적인 분자 모터가 포어와 결합되기 전에 포어를 통해 잡아당겨 질 수 있는 폴리뉴클레오타이드의 길이가 비교적 짧다.
- [0184] 내부 뉴클레오타이드는 표적 폴리뉴클레오타이드 내의 말단 뉴클레오타이드가 아닌 뉴클레오타이드이다. 예를 들어, 이는 3' 말단 뉴클레오타이드 또는 5' 말단 뉴클레오타이드가 아니다. 원형 폴리뉴클레오타이드 내의 모든 뉴클레오타이드는 내부 뉴클레오타이드이다.
- [0185] 일반적으로, 내부 뉴클레오타이드에서 결합할 수 있는 분자 모터는 말단 뉴클레오타이드에서 또한 결합할 수 있지만, 일부 분자 모터는 내부 뉴클레오타이드에서 결합하는 경향이 다른 것보다 클 것이다. 본 발명에서 사용하기에 적절한 분자 모터에 대해, 전형적으로 폴리뉴클레오타이드에 대한 그의 결합의 적어도 10%가 내부 뉴클레오타이드에서의 결합일 것이다. 전형적으로, 그의 결합의 적어도 20%, 적어도 30%, 적어도 40% 또는 적어도 50%가 내부 뉴클레오타이드에서의 결합일 것이다. 말단 뉴클레오타이드에서의 결합은 말단 뉴클레오타이드 및 인접한 내부 뉴클레오타이드 양쪽에 동시에 결합하는 것을 수반할 수 있다. 본 발명의 목적을 위해, 이는 내부 뉴클레오타이드에서의 표적 폴리뉴클레오타이드에 결합하는 것이 아니다. 달리 말하면, 본 발명에서 사용되는 분자 모터는 하나 이상의 인접한 내부 뉴클레오타이드와 조합하여 말단 뉴클레오타이드에 결합할 수 있는 것만은 아니다. 분자 모터는 동시에 말단 뉴클레오타이드에 결합하지 않으면서 내부 뉴클레오타이드에 결합할 수 있어야 한다.
- [0186] 내부 뉴클레오타이드에서 결합할 수 있는 분자 모터는 1개를 초과하는 내부 뉴클레오타이드에 결합할 수 있다. 전형적으로, 분자 모터는 적어도 2개의 내부 뉴클레오타이드, 예를 들어 적어도 3개, 적어도 4개, 적어도 5개, 적어도 10개 또는 적어도 15개의 내부 뉴클레오타이드에 결합한다. 전형적으로, 분자 모터는 적어도 2개의 인접한 내부 뉴클레오타이드, 예를 들어 적어도 3개, 적어도 4개, 적어도 5개, 적어도 10개 또는 적어도 15개의 인접한 내부 뉴클레오타이드에 결합한다. 적어도 2개의 내부 뉴클레오타이드는 인접할 수 있거나 또는 인접하지 않을 수 있다.
- [0187] 내부 뉴클레오타이드에서의 폴리뉴클레오타이드에 결합하는 분자 모터의 능력을 비교 검정법을 수행함으로써 결정할 수 있다. 대조군 폴리뉴클레오타이드 A에 결합하는 모터의 능력을, 동일하지만 말단 뉴클레오타이드에 차단기가 부착되어 있는 폴리뉴클레오타이드 (폴리뉴클레오타이드 B)에 결합하는 능력과 비교한다. 차단기는 가닥 B의 말단 뉴클레오타이드에서의 임의의 결합을 방지하고, 따라서 분자 모터의 내부 결합만 허용한다. 이러한 유형의 검정법의 예가 실시예 4에서 개시된다.
- [0188] 적절한 분자 모터가 업계에 주지되어 있고, 전형적으로 단일 및 이중 가닥 트랜스로카제(translocase), 예컨대 폴리머라제, 헬리카제, 토포이소머라제(topoisomerase), 리가제 및 뉴클레아제(nuclease), 예컨대 엑소뉴클레아제(exonuclease)를 포함하지만, 이에 한정되지는 않는다. 바람직하게는, 분자 모터는 헬리카제, 예를 들어 He1308 헬리카제이다. 내부 뉴클레오타이드에서 결합할 수 있는 He1308 헬리카제의 예는 He1308 Tga, He1308 Mhu 및 He1308 Csy를 포함하지만, 이에 한정되지는 않는다. 따라서, 분자 모터는 바람직하게는 (a) He1308 Tga의 서열 (즉 서열 33) 또는 그의 변이체 또는 (b) He1308 Csy의 서열 (즉 서열 22) 또는 그의 변이체 또는 (c) He1308 Mhu의 서열 (즉 서열 52) 또는 그의 변이체를 포함한다. 변이체는 전형적으로 전체 서열에 걸쳐 아미노산 동일성을 기초로 서열 33, 22 또는 52에 대해 적어도 40% 상동성이고, 헬리카제 활성을 보유한다. 추가적인 가능한 변이체가 하기에서 논의된다.
- [0189] 본 발명에서 사용되는 분자 모터는 상기에서 논의된 방법 중 임의의 것에 의해 제조될 수 있고, 상기 논의된 바와 같이 변형 또는 표지될 수 있다. 이러한 분자 모터는 본원에 기술된 방법에서 또는 본원에 기술된 장치의 일부분으로서 사용될 수 있다. 본 발명은 포어와, 내부 뉴클레오타이드에서의 표적 폴리뉴클레오타이드에 결합할 수 있는 분자 모터 사이에 복합체를 형성시키고, 이에 의해 표적 폴리뉴클레오타이드를 특성화하기 위한 센서를

형성시키는 것을 포함하는, 표적 폴리뉴클레오티드를 특성화하기 위한 센서를 형성시키는 방법을 추가로 제공한다. 본 발명은 포어를 통한 표적 폴리뉴클레오티드의 이동을 제어하기 위한, 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터의 용도를 또한 제공한다. 본 발명은 (a) 포어 및 (b) 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 분자 모터를 포함하는, 표적 폴리뉴클레오티드를 특성화하기 위한 키트를 또한 제공한다. 본 발명은 다수의 포어, 및 내부 뉴클레오티드에서의 표적 폴리뉴클레오티드에 결합할 수 있는 다수의 분자 모터를 포함하는, 샘플 내의 표적 폴리뉴클레오티드를 특성화하기 위한 분석 장치를 또한 제공한다.

[0190] 하기의 실시예들은 본 발명을 예시한다.

[0191] **실시예 1**

[0192] 본 실시예는 나노포어를 통한 무손상 DNA 가닥의 이동을 제어하기 위한 Hel308 헬리카제 (Hel308 MBu)의 사용을 예시한다. 본 실시예 전반에 걸쳐 사용된 일반적인 방법 및 기질이 도 1에서 제시되고, 도면 설명에서 기술된다.

[0193] **물질 및 방법**

[0194] PhiX174의 ~400 bp 단편을 증폭시키도록 프라이머를 디자인하였다. 이러한 프라이머들의 5'-끝부분 각각은 동중중합체성 신장물 또는 뉴클레오티드 10개의 동중중합체성 섹션의 반복 단위인 뉴클레오티드 50개의 비-상보성 영역을 포함하였다. 이들은 나노포어를 통한 가닥의 제어된 전위에 대한 식별자로서, 뿐만 아니라 전위의 방향성을 결정하는 역할을 한다. 또한, 정방향 프라이머의 5'-끝부분이 "캡핑(capping)"되어 4개의 2'-O-메틸-우라실 (mU) 뉴클레오티드를 포함하고, 역방향 프라이머의 5'-끝부분이 화학적으로 인산화되었다. 이러한 프라이머 변형은 람다 엑소뉴클레아제를 사용하여 안티센스 가닥만 우세하게 제어적으로 소화되게 한다. mU 캡핑은 센스 가닥을 뉴클레아제 소화로부터 보호하는 한편, 안티센스 가닥의 5'의 PO4는 이를 촉진한다. 따라서, 람다 엑소뉴클레아제와의 인큐베이션 후, 듀플렉스의 센스 가닥만 이체는 단일 가닥 DNA (ssDNA)로서 무손상으로 유지된다. 그 후, 생성된 ssDNA를 기존에 기술된 바와 같이 PAGE 정제하였다.

[0195] 본원에서 기술된 모든 실험에서 사용된 DNA 기질 디자인이 도 6에서 제시된다. 이러한 DNA 기질은 나노포어에 의한 포획을 보조하기 위한 50T 5'-리더가 있는, PhiX로부터의 ssDNA의 염기 400개의 섹션 (서열 59)로 이루어진다. 이중층의 표면 상에서 DNA를 강화시키고 따라서 포획 효율을 개선시키기 위해 3' 콜레스테롤 태그를 함유하는 프라이머 (서열 60)가 50T 리더 바로 뒤에서 이러한 가닥에 어닐링된다.

[0196] **완충 용액:** 400 mM-2 M KCl, 10 mM Hepes pH 8.0, 1 mM ATP, 1 mM MgCl<sub>2</sub>, 1 mM DTT

[0197] **나노포어:** 이.콜라이(*E.coli*) MS(B2)8 MspA ONLP3271 MS-(L88N/D90N/D91N/D93N/ D118R/D134R/E139K)8

[0198] **효소:** Hel308 Mbu (ONLP3302, ~7.7  $\mu$ M) 12.5  $\mu$ l → 100 nM 최종.

[0199] 1,2-디피타노일-글리세로-3-포스포콜린 지질 (아반티 폴라 리피즈(Avanti Polar Lipids)) 이중층 내에 삽입된 단일 MspA 나노포어로부터 전기 측정을 취득하였다. 몬탈-윌러 기술을 통해 (주문 제작된 델린(Delrin) 챔버에서) 20  $\mu$ m 두께의 PTFE 필름에서 ~100  $\mu$ m 직경의 개구부에 걸쳐 이중층이 형성되어, 2개의 1 ml 완충 용액이 분리되었다. 모든 실험은 언급된 완충 용액에서 수행되었다. 1440A 디지털라이저(digitizer)가 장착된 악소패치(Axopatch) 200B 증폭기 (몰레큘러 디바이시스(Molecular Devices))에서 단일-채널 전류를 측정하였다. 시스 구획 (나노포어 및 효소/DNA가 첨가됨)이 악소패치 헤드스테이지(headstage)의 바닥에 연결되고, 트랜스 구획이 헤드스테이지의 활성 전극에 연결되도록, Ag/AgCl 전극을 완충 용액에 연결하였다. 이중층에서 단일 포어가 달성된 후, DNA 폴리뉴클레오티드 및 헬리카제를 100  $\mu$ l의 완충제에 첨가하고, 5분 동안 예비-인큐베이션하였다 (DNA = 1.5 nM, 효소 = 1  $\mu$ M). 이러한 예비-인큐베이션 믹스를 전기생리학 챔버의 시스 구획 내의 900  $\mu$ l의 완충제에 첨가하여, MspA 나노포어에서의 헬리카제-DNA 복합체의 포획을 개시시켰다 (DNA = 0.15 nM, 효소 = 0.1  $\mu$ M의 최종 농도를 제공함). 필요하다면, 2가 금속 (1 mM MgCl<sub>2</sub>) 및 NTP (1 mM ATP)를 시스 구획에 첨가하여 헬리카제 ATPase 활성을 개시시켰다. +180 mV의 일정한 전위에서 실험을 수행하였다.

[0200] **결과 및 토의**

[0201] 도 1에 제시된 바와 같이 헬리카제-DNA 기질을 MspA 나노포어에 첨가하는 것은 도 2에 제시된 바와 같은 특징적인 전류 차단을 일으킨다. 헬리카제가 결합되지 않은 DNA는 나노포어와 일시적으로 상호작용하여 전류의 단기 차단을 일으킨다 (< 1 초). 헬리카제가 결합되고 활성인 (즉, ATPase 작용 하에 DNA 가닥을 따라 이동하는)

DNA는 도 2에 제시된 바와 같이 전류가 단계적으로 변화되면서 긴 특징적인 차단 수준을 일으킨다. 나노포어 내의 상이한 DNA 모티프들은 독특한 전류 차단 수준을 일으킨다.

[0202] 소정의 기질에 대해, 본 발명자들은 DNA 서열을 반영하는 전류 변이의 특징적인 패턴을 관찰하였다 (도 3에서의 예).

[0203] 도 1에서 제시된 실험에서, 가닥을 따라 있는 무작위 위치에서 DNA가 헬리카제로 포획됨에 따라 무작위 출발점으로부터 DNA 가닥이 서열분석된다. 그러나, 효소가 해리되지 않는 한, 가닥들은 모두 50T 리더에서 동일한 방식으로 끝날 것이다 (도 1). 도 2가 나타내는 바와 같이, 본 발명자들은 대부분의 가닥에 대해 동일한 특징적인 종결을 관찰하였고, 이때 긴 체류 시간 폴리T 수준으로 전류 변이가 종결된다 (도 3).

[0204] 염 허용성

[0205] 이러한 유형의 나노포어 가닥 서열분석 실험은 이온성 염을 필요로 한다. 이온성 염은 전압 오프셋을 인가하여 DNA를 포획하고 전위시키기 위한 전도성 용액을 생성시키는데, 그리고 DNA가 나노포어를 통과할 때 생성된 서열 의존적 전류 변화를 측정하는데 필요하다. 측정 신호가 이온 농도에 의존적이기 때문에, 취득된 신호의 규모를 증가시키기 위해 높은 농도의 이온성 염을 사용하는 것이 유리하다. 나노포어 서열분석을 위해, 100 mM KCl을 초과하는 염 농도가 이상적이고, 1 M KCl 이상의 염 농도가 바람직하다.

[0206] 그러나, 다수의 효소 (일부 헬리카제 및 DNA 모터 단백질 포함)는 높은 염 조건을 견디지 않는다. 높은 염 조건 하에, 효소가 언폴딩(unfolding)되거나 또는 구조적 완전성을 상실하거나, 또는 적절하게 기능하지 못한다. 공지 및 연구된 헬리카제들에 대한 현재의 문헌들은 거의 대부분의 헬리카제가 약 100 mM KCl/NaCl 초과의 염 농도에서 기능하지 못한다는 것을 나타내고, 400 mM KCl 이상의 조건에서 정확한 활성을 나타내는 헬리카제가 보고되지 않았다. 내염성 중으로부터의 유사한 효소들의 잠재적으로 호염성인 변이체가 존재하지만, 표준 발현 시스템 (예를 들어, 이. 콜라이)에서의 발현 및 정제가 매우 어렵다.

[0207] 놀랍게도, 본 발명자들은 Mbu로부터의 He1308이 매우 높은 KCl 수준까지 염 내성을 나타낸다는 것을 이러한 실시예에서 나타낸다. 본 발명자들은 형광 실험에서 또는 나노포어 실험에서 이러한 효소가 400 mM KCl 내지 2 M KCl의 염 농도에서 기능성을 유지한다는 것을 발견하였다 (도 4). 도 4는 도 1에 기술된 것과 동일한 시스템을 사용하여 수행된 400 mM KCl, 1 M KCl, 및 2 M KCl 염 조건에서의 He1308 Mbu DNA 사례를 나타낸다. 본 발명자들은 이러한 염 농도 범위에 걸쳐 유사한 이동을 관찰하였다. 염 농도가 증가됨에 따라, 본 발명자들은 고정된 전압에서 나노포어 (I-개방)를 통한 전류의 증가를 관찰하였다. 이는 용액의 전도성의 증가 및 인가된 장 하에서 나노포어를 통해 유동하는 이온 개수의 증가를 반영한다. 또한, 본 발명자들은 DNA 사례의 전류 수준에서의 구별의 최소 대 최대 범위의 증가도 관찰하였다 (도 4 확대도 및 하부 오른쪽 플롯 참조). 본 발명자들은 염 농도가 400 mM KCl에서 2M KCl로 증가됨에 따른 DNA 구별 범위에서의 ~200 % 증가를 관찰하였다 (하기 표 6).

## 표 6

포어 전류 및 DNA 범위에 대한 염 농도 증가의 효과

염 (KCl) (M)	개방-포어 전류 (pA)	DNA 범위 (pA)
0.4	180	25
1.0	440	55
2.0	840	75

[0208]

[0209] 정방향 및 역방향 작동 방식

[0210] 대부분의 헬리카제는 단방향 방식으로 단일-가닥 폴리뉴클레오티드 기질을 따라 이동하여, 턴오버된 각각의 NTPase에 대해 특정 개수의 염기를 이동시킨다. 도 1은 케인 DNA를 나노포어 밖으로 잡아당기기 위해 이러한 이동을 사용하는 것을 도해하지만, 헬리카제 이동은 DNA를 제어된 방식으로 나노포어를 통해 공급하는 다른 방식으로 활용될 수 있다. 도 5는 기본적인 '정방향' 및 '역방향' 작동 방식을 도해한다. 정방향 양식에서, 인가된 장의 힘 하에 이동하는 것과 동일한 방향으로 헬리카제에 의해 DNA가 포어 내로 공급된다. 3'-5' 헬리카제인 He1308 Mbu에 대해, 이는 헬리카제가 나노포어의 상부와 접촉할 때까지 나노포어 내의 DNA의 3' 끝부분을



포획하는 것을 필요로 하고, 그 후 인가된 전위로부터의 장과 함께 헬리카제의 제어 하에 DNA가 나노포어 내로 공급되어, 최종적으로 이중층의 트랜스 측 상에서 빠져나간다. 역방향 방식은 DNA의 5' 끝부분을 포획하는 것을 필요로 하고, 그 후 인가된 전위로부터의 장에 대항하여 제인 DNA를 나노포어 밖으로 되돌려 잡아당기도록 헬리카제가 진행되어, 최종적으로 이를 이중층의 시스 측으로 사출시킨다. 도 5는 He1308 Mbu를 사용하는 이러한 2가지 작동 방식, 및 전형적인 예시적인 DNA 사례를 나타낸다.

## [0211] 실시예 2

[0212] 본 실시예는 효소 활성을 테스트하기 위한 형광 검정법을 사용하여 He1308 헬리카제 (He1308 Mbu)의 염 허용성을 예시한다.

[0213] 맞춤 형광 기질을 사용하여, 혼성화된 dsDNA를 전치시키는 헬리카제의 능력을 검정하였다 (도 6a). 도 6a의 1)에 제시된 바와 같이, 형광 기질 가닥 (100 nM 최종)에 3' ssDNA 오버행, 및 혼성화된 dsDNA의 염기 40개의 섹션이 있다. 상부의 주 가닥은 5' 끝부분에 카르복시플루오레세인 염기가 있고, 혼성화된 상보물은 3' 끝부분에 블랙-홀 켄처 (BHQ-1) 염기가 있다. 혼성화되었을 때, 플루오레세인으로부터의 형광이 국소적인 BHQ-1에 의해 켄칭되고, 기질이 본질적으로 비-형광성이다. 1  $\mu$ M의 포획 가닥 (형광 기질의 더 짧은 가닥에 상보성임)이 검정법에 포함된다. 2)에 제시된 바와 같이, ATP (1 mM) 및  $MgCl_2$  (5 mM)의 존재 하에, 기질에 첨가된 헬리카제 (100 nM)가 형광 기질의 3' 꼬리에 결합하고, 주 가닥을 따라 이동하고, 제시된 바와 같이 상보적인 가닥을 전치시킨다. 3)에 제시된 바와 같이, BHQ-1이 있는 상보적인 가닥이 완전히 전치되면, 주 가닥 상의 플루오레세인이 형광을 발한다. 4)에 제시된 바와 같이, 과량의 포획 가닥이 상보적인 DNA에 우선적으로 어닐링하여, 초기 기질이 다시 어닐링되는 것 및 형광 상실을 방지한다.

[0214] 기질 DNA: 5'FAM-서열 61 및 서열 62-BHQ1-3'. FAM = 카르복시플루오레세인 및 BHQ1 = 블랙 홀 켄처-1

[0215] 포획 DNA: 서열 62.

[0216] 도 6의 그래프는 400 mM 내지 2 M의 상이한 농도의 KCl을 함유하는 완충제 용액 (10 mM Hepes pH 8.0, 1 mM ATP, 5 mM  $MgCl_2$ , 100 nM 형광 기질 DNA, 1  $\mu$ M 포획 DNA)에서의 활성의 초기 비율을 나타낸다. 2 M에서 헬리카제가 작동한다.

## [0217] 실시예 3

[0218] 본 실시예에서, 3가지 상이한 He1308 헬리카제, 즉 He1308 Mhu (서열 52), He1308 Mok (서열 29) 및 He1308 Mma (서열 45)가 사용되었다. 모든 실험은 이전에 실시예 1에서 기술된 바와 같이 동일한 실험 조건 하에 수행되었다 (포어 = MspA B2, DNA = 400량체 서열 59 및 60, 완충제 = 400 mM KCl, 10 mM Hepes pH 8.0, 1 mM dtt, 1mM ATP, 0.1 mM  $MgCl_2$ ). 결과가 도 7에서 제시된다.

## [0219] 실시예 4

[0220] 본 실시예는 형광 검정법을 사용하여 다수의 He1308 헬리카제의 내부 결합 능력을 측정한다.

[0221] 맞춤 형광 기질을 사용하여, 천연 3' 끝부분이 결여된 DNA에서 시작하고, 이어서 혼성화된 dsDNA를 전치시키게 하는 헬리카제의 능력을 검정하였다 (도 8). 도 8의 섹션 A에 제시된 바와 같이, 형광 기질 가닥 (50 nM 최종)에 3' ssDNA 오버행, 및 혼성화된 dsDNA의 염기 40개의 섹션이 있다. 상부의 주 가닥이 3' 끝부분에서, 또는 내부에서 오버행과 dsDNA 사이의 접합부에서 (음성 대조군으로서), 4개의 연속적인 "스페이스 9" 기로 변형된다. 또한, 상부의 주 가닥은 5' 끝부분에 카르복시플루오레세인 염기가 있고, 혼성화된 상보물은 3' 끝부분에 블랙-홀 켄처 (BHQ-1) 염기가 있다. 혼성화되었을 때, 플루오레세인으로부터의 형광이 국소적인 BHQ-1에 의해 켄칭되고, 기질이 본질적으로 비-형광성이다. 포획 가닥 (1  $\mu$ M) (형광 기질의 더 짧은 가닥에 상보성임)이 검정법에 포함된다. 섹션 B에 제시된 바와 같이, ATP (1 mM) 및  $MgCl_2$  (1 mM)의 존재 하에, 3'-말단 "스페이스 9" 기를 함유하는 기질에 첨가된 He1308 헬리카제 상동체 (20 nM)가 형광 기질의 ssDNA 오버행에 결합할 수 있고, 주 가닥을 따라 이동할 수 있으며, 상보적 가닥을 전치시킬 수 있다. BHQ-1이 있는 상보적인 가닥이 완전히 전치되면 (섹션 C), 주 가닥 상의 플루오레세인이 형광을 발한다. 과량의 포획 가닥이 상보적인 DNA에 우선적으로 어닐링하여, 초기 기질이 다시 어닐링되는 것 및 형광 상실을 방지한다 (섹션 D).

[0222] 기질 DNA: 5' FAM이 있는 서열 63; 5' FAM 및 3' 스페이서 ((스페이스 9)<sub>4</sub>)가 있는 서열 63; 스페이서 ((스페이스 9)<sub>4</sub>)로 분리된 서열 64 (5' FAM가 있음) 및 65; 및 3' BHQ1이 있는 서열 62.

- [0223] 포획 DNA: 서열 66.
- [0224] 다수의 상이한 He1308 헬리카제 상동체를 이들의 중간-결합 능력에 대해 연구하였고, He1308 Mbu, He1308 Csy, He1308 Tga, He1308 Mma, He1308 Mhu, He1308 Min, He1308 Mig, He1308 Mmaz, He1308 Mac, He1308 Mok, He1308 Mth, He1308 Mba 및 He1308 Mzh가 이에 포함된다. 도 9의 그래프는 400 mM NaCl, 10 mM Hepes, pH 8.0, 1 mM ATP, 1 mM MgCl<sub>2</sub>, 50 nM 형광 기질 DNA, 1 μM 포획 DNA에서의 3'-미변형 DNA 및 3'-"스페이스 9" DNA를 비교하는 He1308-매개 dsDNA 턴오버의 상대적인 비율을 나타낸다. He1308 Csy, He1308 Tga, He1308 Mma, He1308 Mhu 및 He1308 Min을 포함하는 여러 He1308 상동체들이 상대 속도가 He1308-매개 dsDNA 턴오버의 20%를 초과하는 것으로 관찰되었다.
- [0225] **실시예 5**
- [0226] 본 실시예는 2개의 He1308 헬리카제인 He1308 MBu 및 He1308 Tga의 사용, 및 나노포어를 통한 긴 무손상 DNA 가닥 (900량체)의 이동을 제어하는 이들의 능력을 비교한다. 본 실시예 전반에 걸쳐 사용된 일반적인 방법 및 기질이 도 10에서 제시되고, 상기 도면 설명에서 기술된다.
- [0227] **물질 및 방법**
- [0228] 50-폴리T 5' 리더를 PhiX dsDNA의 염기 ~900개의 단편에 라이게이션시킴으로써 DNA가 생성되었다. 리더는 DNA가 이중층에 테더링되게 하도록 Chol-태그가 있는 서열 69가 혼성화되는 상보적인 섹션을 또한 함유한다. 마지막으로, PhiX dsDNA의 3' 끝부분을 AatII 소화 효소로 소화시켜, ACGT의 4 nt 3'-오버행이 산출되었다.
- [0229] 사용된 서열: 서열 67 - 5' 리더 및 테더를 포함하는 900량체 센스 가닥; 서열 68 - 안티-센스 마이너스 4 염기-쌍 리더 5'; 및 여러 스페이스 및 3' 끝부분의 Chol-태그가 있는 서열 69.
- [0230] **완충 용액:** 400 mM-2 NaCl, 10 mM 페로시아노칼륨, 10 mM 페리시아노화칼륨, 100 mM Hepes, pH 8.0, 1 mM ATP, 1 mM MgCl<sub>2</sub>,
- [0231] **나노포어:** MS-(B1- G75S-G77S-L88N-Q126R)8 (ONT Ref B2C)
- [0232] **효소:** He1308 Mbu 1000 nM 또는 He1308 Tga 400 nM 최종.
- [0233] 지질 이중층 내로 삽입된 단일 포어를 달성하도록 실시예 1에 기술된 바와 같이 전기 실험을 설정하였다. 이중층 내의 단일 포어를 달성한 후, ATP (1 mM) 및 MgCl<sub>2</sub> (1 mM)를 챔버에 첨가하였다. +140 mV에서의 대조군 기록을 2분 동안 실행하였다. 그 후, DNA 폴리뉴클레오티드 서열 67, 68 및 69 (DNA = 0.15 nM)를 첨가하고, DNA 사례를 관찰하였다. 마지막으로, He1308 헬리카제 (Mbu 1000 nM 또는 Tga, 400 nM)를 전기생리학 챔버의 시스 구획에 첨가하여, MspA 나노포어에서의 헬리카제-DNA 복합체의 포획을 개시시켰다. +140 mV의 일정한 전위에서 실험을 수행하였다.
- [0234] **결과 및 토의**
- [0235] 도 10에 제시된 바와 같이 헬리카제-DNA 기질을 MspA 나노포어에 첨가하는 것은 헬리카제가 포어를 통한 DNA의 전위를 제어함에 따라 특징적인 전류 차단을 일으킨다. 도 11은 He1308 헬리카제 상동체 Mbu가 MspA 포어를 통한 DNA 가닥의 전위를 제어함에 따라 어떻게 900량체의 위치가 변하였는지를 가리키는 예시적인 사례 트레이스를 나타낸다. 이러한 헬리카제는 DNA 전위를 매개하는 것으로 확인되었지만, 헬리카제가 DNA로부터 분리되었을 때, 외부에서 인가된 전위가 발휘하는 힘으로 인해 가닥이 포어를 통해 되돌아 이동하는 것으로 관찰되었다. He1308 헬리카제 상동체 Mbu의 경우, 헬리카제가 탈착될 때마다 900량체 가닥에서 다수의 위치 (약 100-200개의 염기)가 다시 슬립되었다. 이러한 급속한 위치 변화가 도 11에서 점선 원으로 지시된다. He1308 헬리카제 상동체 Mbu가 분자 모터로 사용된 이러한 실험에 대해, 검출된 모든 사례의 32%가 900량체 가닥 서열의 전체 길이를 관독한 것으로 확인되었다. 도 12는 He1308 헬리카제 상동체 Tga가 MspA 포어를 통한 DNA 가닥의 전위를 제어함에 따라 어떻게 900량체의 위치가 변하였는지를 가리키는 유사한 예시적인 사례 트레이스를 나타낸다. 이러한 효소는 Mbu 상동체보다 더 큰 내부 결합 경향을 나타냈는데, Tga 헬리카제가 탈착될 때 (도 12에서 흑색에서 회색으로의 변화로 지시됨), DNA 가닥이 비교적 작은 거리 (<50개의 염기)만큼 포어를 통해 되돌아 이동하였기 때문이다. He1308 헬리카제 상동체 Tga가 분자 모터로 사용된 이러한 실험에 대해, 검출된 모든 사례의 74%가 900량체 가닥 서열의 전체 길이를 관독한 것으로 확인되었다. 이는 Tga 헬리카제 상동체가 이의 증가된 내부 결합 경향으로 인해 Mbu 헬리카제 상동체와 비교하여 단일-가닥 DNA의 증가된 관독 길이를 제공할 수 있다는 것을 의미한다.

[0236] **실시예 6**

[0237] 본 실시예는 He1308 헬리카제 상동체 Tga를 사용함으로써 5 kb DNA 가닥의 전위를 제어하는 것이 가능하다는 것을 예시한다.

[0238] 실시예 5에 기술된 것과 유사한 실험 절차를 따랐다. He1308 Tga를 사용함으로써 MS-(B1- G75S-G77S-L88N-Q126R)8을 통한 전체 5 kb DNA 가닥의 제어된 전위를 검출하는 것이 가능하였음이 관찰되었다. He1308 Mbu를 사용한 동일한 실험에서는, 전체 5 kb 가닥의 전위를 검출하는 것이 가능하지 않았다.

[0239] **실시예 7**

[0240] 본 실시예는 형광 기반 검정법을 사용하여 He1308 Mbu 헬리카제 (서열 10)와 He1308 Mok (서열 29)의 효소 프로세스성을 비교한다.

[0241] 맞춤형 형광 기질을 사용하여, 혼성화된 dsDNA를 전치시키는 헬리카제의 능력을 검정하였다 (도 3). 형광 기질 (50 nM 최종)에 3' ssDNA 오버행, 및 혼성화된 dsDNA의 염기쌍 80개 및 33개의 섹션이 있다 (도 13 섹션 A, 서열 70). 하부의 주 "주형" 가닥이 이의 3' 오버행에 인접한 80 nt "차단제" 가닥 (서열 71), 및 33 nt 형광 프로브 (이의 5' 및 3' 끝부분에서 각각 카르복시플루오레세인 (FAM) 및 블랙-홀 켄처 (BHQ-1) 염기로 표지됨) (서열 72)에 혼성화된다. 혼성화되었을 때, FAM이 BHQ-1에서 멀고, 기질이 본질적으로 형광성이다. 도 13 섹션 B에 제시된 바와 같이, ATP (1 mM) 및  $MgCl_2$  (10 mM)의 존재 하에, 헬리카제 (20 nM)가 기질의 3' 오버행 (서열 70)에 결합하고, 하부 가닥을 따라 이동하고, 80 nt 차단제 가닥 (서열 71)을 전치시키기 시작한다. 프로세스성이면, 헬리카제가 형광 프로브 (서열 72, 이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨)도 전치시킨다 (도 13 섹션 C). 형광 프로브는 이의 5' 및 3' 끝부분이 자가-상보적인 방식으로 디자인되고, 따라서 전치되면 동역학적으로 안정적인 헤어핀을 형성하여, 프로브가 주형 가닥에 다시 어닐링하는 것을 방지한다 (도 13 섹션 D). 헤어핀 생성물의 형성 시, FAM이 BHQ-1에 인접하게 되고, 이의 형광성이 켄칭된다. 80량체 "차단제" (서열 71) 및 형광 (서열 72, 이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨) 가닥을 전치시킬 수 있는 프로세스성 효소는 따라서 경시적으로 형광 감소에 이를 것이다. 그러나, 효소의 프로세스성이 80 nt 미만이면, 이는 형광 가닥 (서열 72, 이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨)을 전치시킬 수 없을 것이고, 따라서 "차단제" 가닥 (서열 71)이 하부의 주 가닥 (서열 70)에 다시 어닐링할 것이다 (도 13 섹션 E).

[0242] 추가적인 맞춤형 형광 기질이 대조군 목적으로 또한 사용되었다. 음성 대조군으로 사용된 기질은 도 13에 기술된 것과 동일하지만 3' 오버행이 결여된다 (도 14 섹션 A, (서열 71, 72 (이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨) 및 73)). 도 13에 기술된 것과 유사하지만 염기쌍 80개의 섹션이 결여된 기질이 활성이지만 반드시 프로세스성이지는 않은 헬리카제에 대한 양성 대조군으로서 사용되었다 (도 14 섹션 B, (서열 72 (이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨) 및 74)).

[0243] 도 15는 완충 용액 (400 mM NaCl, 10 mM Hepes pH 8.0, 1 mM ATP, 10 mM  $MgCl_2$ , 50 nM 형광 기질 DNA (서열 70, 71 및 72 (이의 5' 끝부분에서 카르복시플루오레세인 (FAM)으로, 이의 3' 끝부분에서 블랙-홀 켄처 (BHQ-1)로 표지됨)에서 도 13에 제시된 프로세스성 기질에 대해 He1308 Mbu 헬리카제 (서열 10) 및 He1308 Mok 헬리카제 (서열 29)를 테스트했을 때의 시간-의존적 형광 변화의 그래프를 나타낸다. He1308 Mok이 나타낸 형광 감소는 He1308 Mbu (서열 10)에 비교했을 때 이러한 복합체의 증가된 프로세스성을 나타낸다. 도 16은 모든 샘플에 대한 형광 감소가 나타내는 바와 같이, 모든 헬리카제가 실제로 활성이었음을 입증하는 양성 대조군을 나타낸다.

서열 10, 13, 16, 19, 22, 25, 28, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54 및 55 (표 4)가 하기에서 정렬된다.

\* 아래의 숫자는 서열 번호를 가리킨다. "-"는 정렬 목적으로만 제시되고, 서열의 일부를 형성하지 않는다.

*		1	95
10 Hel308 Mbu	(1)	-----MIRELDIPDITIGFYEDSGIKELYPPOAALTEMGILF-KNNLLAIPFASGKTLIAELAMTK	
53 Hel308 Afu	(1)	-----MVEELIESISYAVGILKEGIELELPPOAEAVEKYS--GNLLAMPFAGKTLIAELAMVR	
22 Hel308 Csy	(1)	-----MRISLDIPPAIELEEGEGYKELYPPOAAAKAGLTD-GKSVYSAFPAAGKTLIAELAMIS	
75 Hel308 Dth	(1)	MPGVDELLQMGQGLQGLSTVAVKEIPAREAFPSGIEGLPPKQALTEGINDENFYTHOAAVNLVYK--GKSVYATFPAAGKSLIYVILE	
48 Hel308 Fac	(1)	-----MKLSITPSEFLKTVNDNDPTLYEHOBAAVAKREN--KNVYSAVTPAAGKTLIIGYISYD	
19 Hel308 Hla	(1)	-----MOPSGISGIPAGVGEALAEAGVAELYPPOAEAVEAGYAD-GEGLVAIVPFAAGKTLIAELAMIS	
55 Hel308 Hpa	(1)	-----MNVADITGIPGVPBEHFAQGITELIYPPOAEAVEAGTTE-GEGLVASIFPAAGKTLIAELAMIS	
54 Hel308 Hpa	(1)	-----MNVLEITGIPGATDHFPGGIELELPPOADAVEAGATD-GENLVAIVPFAAGKTLIAELAMIS	
16 Hel308 Hvu	(1)	-----MRTADITGIPGAIPEALNDGIELELPPOAEAVEAGLTD-GEGLVAIVPFAAGKTLIAELAMIS	
39 Hel308 Mac	(1)	-----MKIESLDLPBEVKRYENSGIPELPPOAEAVEKGILF-GENLLAIPFAAGKTLIAELAMTK	
38 Hel308 Mba	(1)	-----MKIESLDLPBEVKQFYNSGIMELYPPOAEAVEKGILF-GENLLAIPFAAGKTLIAELAMTK	
47 Hel308 Mbo	(1)	-----MQIODAIPPEPRAQYIGIGIRELPYPOAAVEKRGILD-GENLLAIPFAAGKTLIAELAMHR	
44 Hel308 Mey	(1)	-----METGKLEPEYVIOFYLDGIEKLELPPOAEAVEKRGILD-NNLLAIPFAAGKTLIASELAMTK	
49 Hel308 Mfe	(1)	-----MPTNKILIELIDGIELELRPPOKKALEKRGILD-KNNFLISIPFAAGKTLIIGEMALIN	
28 Hel308 Mfr	(1)	-----DISPRAFIQVYKDGIESIEXPQSGCIENGILD-GADILVAIPFAAGKTLIAELAMHA	
52 Hel308 Mfu	(1)	-----METASIPLPDSTIRACHAGITSIPPOAEKIEKGILF-GENLLISIPFAAGKTLIAELAMMS	
32 Hel308 Mlu	(1)	-----MQKYSHEVLEKENGIELELRPPQKRYEKGLNKEKNFLICIPFAAGKTLIAGEMALIN	
51 Hel308 Mln	(1)	-----MDELTKELGIELELRPPQKRALEGLIDKKNFLISIPFAAGKTVIAGEMALIN	
45 Hel308 Mma	(1)	-----MNVLDLKENKITELELRPPQKVIDEGLFDKTNFLICIPFAAGKTLIAGEMALIN	
40 Hel308 MmaH	(1)	-----MKIEELDLPSEBAIEVYLQAGIELELPPOADAVEKGLIQ-GENLLAIPFAAGKTLIAELAMTK	
76 Hel308 MmaH	(1)	-----MDVADLPQVPMWLPDHLNDGIELELPPOAEAVEAGTTE-GENLVAIVPFAAGKTLIAELAMTK	
41 Hel308 MmaZ	(1)	-----MKIESLDLPBEIKRYENSGIIELELPPOAEAVEKGILF-GENLLAIPFAAGKTLIAELAMTK	
29 Hel308 Mok	(1)	-----MMAMEYKENGITAELRPPQKVEGGLNKNKNFLICIPFAAGKTLIIGEMALIN	
42 Hel308 Mth	(1)	-----MLTIDILRLPEPSTIELELYAAGIDELYPPOAEALTEKGLD-GRNMVTSVPAAGKTLIAELMTR	
43 Hel308 Mzh	(1)	-----MNNINMLPEKRYKTYDGTGYDLYPPQREAVEKGLD-GENVYAIIPFAAGKTLIAELCMRK	
46 Hel308 Nma	(1)	-----MNVDELISGIPGASHFQGIIELELPPOAEAVEAGATE-GENLVAIVPFAAGKTLIAELMSLIS	
77 Hel308 Nth	(1)	-----MSETVYLISERQKTIEMGMDFTVPQKTIPIVANT-NDQVVASGPAAGKTEAVFLPLIS	
13 Hel308 Pfu	(1)	-----MNVDELR---VDRIRKSTIKENGIESFPVQAEALKSGILF-GRNALISIPFAAGKTLIAELAMVR	
25 Hel308 Pfu	(1)	-----MSLIELEMPPIEDIKLIPENVIELIKKGIKLNKPPQTEAVEKKGILF-GRNLLISFPAGKTLIAELMGIS	
34 Hel308 Tba	(1)	-----MUSTKPKAVKRFSPIG--YAAQVDELISFQVDRIRIKIRKIGISEFPVQAEALRSGVAN-GENLLAIPFAAGKTVIAELVMLH	
33 Hel308 Tga	(1)	-----MKVDELFP---VDRIRKAVLEKGIIELELPPOAEALKSGALE-GRNVYLAIPFAAGKTVIASEIYVNW	
37 Hel308 Tsi	(1)	-----MKMLKLSYINAFILGAMWMSKVDKLSIKGIVDRIRILBERGIELELPPOAEALKTEVYK-GRNVYLAIPFAAGKTVIAELVYN	
50 Hel308 Mja	(1)	-----MKKTIETLDGIVYELRPPQKALREGLIDKNNKFLISIPFAAGKTLIAGEMALIN	
78 권해시스	(1)	LP V L E GI ELYPQAEAVE GLID GRNLLAIPFAAGKTLIAELAML	

96

HeI308 Mbu (63) AIREGG-----KALIYVPLRALASEKTEREK-ELAP-----FGIKVGI5TGDDSDBRADWLVGNDIIIVATSEKTD5ILNRGTSWMD-----ETT  
HeI308 Afu (64) EAIKRG-----KALIYVPLRALAEKTEREK-RWEX-----IGRIJGI5TGDEKSDSEHIGODIIVTSEKTD5ILNRASWIK-----AVS  
HeI308 Csy (65) HLSRNR-----GKAYVLSPLRALAEKTEREKGIGIPL-GRPRVGVSTGDEKTEREGRSIRKNDIIVLTNEMSD5ILRRDPMW-----EVS  
HeI308 Dth (94) SIINDP-----ASALYIPELKAITDQUTSLEEPALRAGVHDVSAVGDGDPAPARIRSKPENIILITPMIDHASF5LPHRSMQKFSARAK-----YVG  
HeI308 Fac (60) TYIKRG-----KSMYIYPLSLAEKTESEI-SLRN-----LGKVTMSGIDVPSPEFVKNDIIVATSEKSD5MLRRDPIILN-----YGG  
HeI308 Hla (64) SIERGG-----KALIYVPLRALASEKTEREE-RWEE-----FGYTVGVSTGVN5SGDEWMLATRODIIVATSEKVD5ILNRGAPWD-----DIT  
HeI308 Hpa (64) SVARGG-----KALIYVPLRALASEKTEEE-BFEQ-----YGV5IGV5SGN5SGDEWMLASDIIIVATSEKVD5ILNRGAKWD-----DLS  
HeI308 Htu (64) AVORGG-----KALIYVPLRALASEKTEEE-AYEE-----FGYTVGVSTGVN5SGDEWMLATRODIIVATSEKVD5ILNRGADWT5-----ETT  
HeI308 Hvo (64) SVARGG-----KALIYVPLRALASEKTEREE-RWEE-----YGDVGVSIGV5SGDEWMLASDIIIVATSEKVD5ILNRGAPWD-----DIT  
HeI308 Mac (63) SVIAGG-----KALIYVPLRALASEKTRRRQ-DPFE-----LGIRVGI5TGDDVDRDEGGINDIIIVATSEKTD5ILNRETAMQ-----EIS  
HeI308 Mba (63) SIILAGG-----KALIYVPLRALASEKTRRRR-BFSE-----LGIRVGI5TGDDVDRDEGGINDIIIVATSEKTD5ILNRETAMQ-----EIS  
HeI308 Mbo (63) HIANGG-----KALIYVPLRALASEKTEEBG-NK-----GKRVGL5TGDDLRDRDDIAGKNDIIVATSEKVD5ILNRGAPWD-----DIT  
HeI308 Mew (63) SI5INGG-----KALIYVPLRALASEKTEREK-OFSS-----IGNVIGI5TGDDVSD5DEWMLGNSDIIIVATSEKAD5ILNRN5WAK-----DIT  
HeI308 Mfe (58) HILDEKNKPTNNKGFIYVPLKALASEKTEREKRYER-----YGRVAL5TGDD-DEDSILRHILITTAELKJDSILMRKPIDW-----DVS  
HeI308 Mlg (63) AIARGG-----MCYIYVPLRALASEKTEREK-OK-----GAEIGVAGDGVDOEKRLGNSDIIIVATSEKVD5ILNRGPWLS-----GYT  
HeI308 Mnu (63) RIAGGG-----KCIYIYVPLRALASEKTERE5-RKG-----YIRGVAL5TGDDRDYAGENDIIIVATSEKTD5ILNRKPWLS-----QTT  
HeI308 Mng (59) HILDEKNKPTNNKGFIYVPLKALASEKTEREKRYER-----YGRVAL5TGDD-DEKDS5NIIITTAELKJDSILMRHEDWN-----YVS  
HeI308 Mln (53) HILDD-----GKKGVYIYVPLKALASEKTEREKRYER-----FGVRVAL5TGDD-DEDEL5NDIITTAELKJDSILMRHGIKLS-----DIS  
HeI308 Mma (55) HILDEKNKLTGKRGFIYVPLKALASEKTEREKRYER-----YGRKVL5TGDD-IREKMSKPHIITITSEKJDSILMRHNVEN-----DVS  
HeI308 Mmh (64) AIKRG-----KALIYVPLRALASEKTRDR-BFES-----LGKTVAGI5TGDD5SGDEWMLGNSDIIIVATSEKTD5ILNR5PWAK-----DIT  
HeI308 Mmr (64) SVARGG-----KALIYVPLRALASEKRDADP-BFEQ-----YGLDIGV5GVN5SGDEWMLADQIIVATSEKVD5ILNRDAPWLE-----DIT  
HeI308 Mnz (63) SVANGG-----KALIYVPLRALASEKTRRRQ-BF5Y-----LGKRVGI5TGDDVDRDEGGINDIIIVATSEKTD5ILNRETAMQ-----EIS  
HeI308 Mnk (56) HILDEKNKPTNNKGFIYVPLKALASEKTEREKRYER-----YGRVAL5TGDD-DEKDLKGRDIIITTAELKJDSILMRKYEIK-----DIT  
HeI308 Mok (66) GALSCK-----RSYIYVPLRALASEKTERE55-RFSK-----LGIRVGI5TGDEPEKDEBLGKNDIIVATSEKAD5ILNRGASWR-----RIG  
HeI308 Mth (66) SIOMGG-----KCIYIYVPLRALASEKTSRRR-BFES-----LGKTVAGI5TGDD5SGDEWMLGNSDIIIVATSEKVD5ILNRN5WAK-----EIN  
HeI308 Mzh (63) AVORGG-----KALIYVPLRALASEKTEREED-AYEE-----FGYTVGVSTGVN5SGDEWMLATRODIIVATSEKVD5ILNRGADWT5-----DIT  
HeI308 Nma (64) QIEKDAT-----KDKILYI5PLKALINDQEPREIITIKCEKSY-IPHRHMGVDVGNQKKRQITKNPAGIILQITPESIE5LFINRNETNLYIL-----SDIE  
HeI308 Nth (63) RIITGG-----GKAYIYVPLKALASEKTEREÖQ-DWEK-----IGKRVAMA5TGDDVSD5DEWMLGKNDIIIVATSEKTD5ILNRHGSWIK-----DVK  
HeI308 Ptu (70) FILKNG-----GKAYIYVPLRALITNEKILTER-DWEI-----IGKRVAM5TGDDVSD5DEWMLGKNDIIITTAELKJDSILMRHREWI-----EYN  
HeI308 Sso (84) KLPTGG-----GKAYIYVPLKALAEKTEREK-RWED-----LGVRVAVTGGDVSDEWMLGKNDIIVATSEKTD5ILNRK5RWIR-----DVT  
HeI308 Tba (63) KLIOEG-----GKAYIYVPLKALAEKTEREK-REWE-----LGKRVATGGDVSDDWMLGKNDIIVATSEKTD5ILNRHGAWIN-----DVK  
HeI308 Tga (86) KILIREG-----GKTVYIYVPLKALAEKTEREK-RWEK-----LGRIAMTGGDVSDEWMLGKNDIIIVATSEKTD5ILNRK5PWIK-----DIN  
HeI308 Tst (56) HILDEKNKPTNNKGFIYVPLKALASEKTEREKRYER-----YGRVAL5TGDD-DEDSILRHILITTAELKJDSILMRKHDWIN-----DVS  
HeI308 Mja (96) IL GG KALIYVPLRALASEKTY EFR FE GVRGISTGDDV DEMWLG DIIVATSEKVD5ILNR WI DIT



[illegible]



Hel308 Mdu	(227)	LEK----	DDAVNLVLDITKAGQ----	CLVESSRRNCAGFAKTAS--	KVAKTIINDIMIKIAGIAEVES--	TGETDTAIVLANCIRKGV	380									
Hel308 Altu	(225)	TSRR----	VKEFEIVEECVANGQ----	CLVESTRGAKTAVKLSA--	ITAKYVEN-----	EGEKALIE--ENEGBMSRKIAECVAKGA										
Hel308 Csy	(233)	TGGG----	PAYDLAAEVSVAEGQ----	SLIFADTRARASIAAKASA--	VJPEAKGADAAKIAAAKTISS--	GGETKIAKTIAEIVAKGA										
Hel308 Dth	(273)	-----	AIRVIQKALELGLR-----	TIIVYQSRKMTIELIMMASQRAGRUKYISAYRAGFLPEQOREIEQKLAGELLAYVSTALELGI	-----	-----										
Hel308 Fac	(217)	-EKKHGRDEVSILKESTESGQ----	ALVFNSRRNAKVIQSMVN-----	-----	FFDQNDPEKLEIPDILFNEAQANMAHGV	-----										
Hel308 H1a	(229)	ERGE----	QOTALVADALDTEDEQGGSSIVFVNSRRNAESSARKLTD--	VTGPRLTDEBERDQJRELADEIRS--	GSDTDTASDLADAVEQGS	-----										
Hel308 Hpa	(229)	-GNE----	KETAIVBDTIEDGSS-----	SLVFNSRRNAAAKARLAD--	VTKTHLNDERROLDIADQIRD--	VSDTETSDIATATIEKGA										
Hel308 Htu	(229)	EGSE----	KOZALVBDIIRREGGS-----	SLVFNSRRNAAEAAKRLQ-----	VSSREITDERAEIADITRD--	DSDTETSDADLADCVBERGA										
Hel308 Hvo	(229)	GRGE----	KOTPALVADALEGDEGDQSSIVFVNSRRNAESAARMAD--	VTERYVTGDEERSDIAELAEITRD--	VSDTETSDILANAAVAKGA	-----										
Hel308 Mac	(227)	PTK----	DEAVNLVLDITREGQ----	CLVESSRRNCAGFAKKAAS--	AVKKTLSAEDREKLAGIADETIE--	NSETDTASIVLASCVRAGT										
Hel308 Mba	(227)	STK----	DEAVNLADITKRDQ-----	CLVESSRRNCAGFAKKAAS--	TYKKTLSAEDRNALAGIADETIE--	NSETDTSTNLAVCIRSGT										
Hel308 Mco	(224)	VSRN----	YDIDNLICDITIAEGQ----	CLVESSRRNAAEFAKKAAG--	ALKSEDA-----	ALAAEAERLIE--GPTMEMVKTIAECVAKGA										
Hel308 Mev	(227)	IVK----	DTAVNLVLDITDEGQ----	CLVESSRRNACAGFAKKAAS--	KVGSILDKGLAEINNIAAEEYLE--	TSDETETKEIASCIRGT										
Hel308 Mfe	(222)	RETKAINNDIYNLVDCVKDGGC----	CLVECONTRGAVNEAKLIN-----	LAKKLTJNEEKRLKEVAEELISTIEPTEMCRTIAECILNGS	-----	-----										
Hel308 Mfc	(220)	PAK----	FEDINILLDCEVADGQ----	CLVESSRRNAGYAKRAAT--	ALKCSHA-----	ALDSIAEKLEA--AAETDMGVLATCVAKGA										
Hel308 Mnu	(225)	KTK----	HDIDNLICDITIEEGQ----	CLVEFVSSRRNAGFAKKAAG--	ALKAGSP-----	DSKALAOELIR--LRDRDEGNVLADCVBERGA										
Hel308 Mlg	(223)	KELPNFSNNPMNLVDCVKEGGC----	CLVEFONSKRNGAIVSEAKLIN-----	LAKKYLISNKEYELOKLEELISTIDPPTETCKTIAECLEKGV	-----	-----										
Hel308 Mln	(211)	-----	VKEQDIKAEVVKDNGS-----	VIIIFCTKKKAENRALSHD-----	LSDLLKRSERKLEETSEELISLFDPPTELCKRLASCVRKGI	-----										
Hel308 Mma	(219)	KKIKQVSRNNLNDLVDSVEERKS-----	CLIFONSKRNAAVGEAKKH-----	LAKYLTREQHEINKLSEELISLIDRPVTECKALASCIQNGV	-----	-----										
Hel308 Mmah	(227)	RHK----	BDSNVIVDITIOGGQ----	CLVEFOSRRNCAGFAKKAAP--	AVGELIDRONNEIEEVAKEYLE--	NGETKLTETIAYCIRKGV										
Hel308 Mmat	(229)	QNNNE----	KQJALVBDITIEDGSS-----	TLVFVNSRRNAAEAAARLANS--	TVRPHLSTEEBDQJADIAEETIRD--	VSDTETSPDLADAVADGA										
Hel308 Mmz	(226)	PTK----	DEAVNLVLDITIEEGQ----	CLVEFOSRRNCAGFAKKAAS--	AVKKTLSNEDRETTIAGIADETIE--	NSETDVSSVLATCVASGT										
Hel308 Mok	(230)	RNR----	DPVNLVLDITVDGQ-----	CLIFONSKRNGAIVSEAKLIN-----	LAKKYLISPEDEISELRHKEEYLVLDNPTKTCODIAECIEKGV	-----										
Hel308 Mth	(227)	ESR----	DDAVNLVLDITVKRGQ----	CLVEFSSRRNCAGFAKKAAG--	WVSKTIDHDITQJLSQOEIGE--	AGETEIADVLSCVRQGV										
Hel308 Mzh	(229)	EAGE----	KOZALVBDIIRREGGS-----	SLVFNSRRNAAEAAARLANS--	VSSRELTAGEOINDAALATEIRE--	DSDETSDIADLADCVBERGA										
Hel308 Mna	(241)	-----	ID-LVQDLRELTKN-----	VHSLIFONSKRAEVEVETIYLINR--	LANREYNTIELIAHSSSIDDKER--	EYVEKTMANSKPSKYVTI										
Hel308 Nth	(226)	-----	WEELVYDALIRKRGK-----	ALIFVNMRRKARVALIELSK--	KVKSILTPELIRALNELADISLE--	ENPTNKRKLAAIRGV										
Hel308 Sso	(232)	TKRVHG--	DDAIITAYTSDISKNGQ----	ALVFVNSRRNAAEAAARLANS--	KYVRRLITPEARIRKELAEISLE--	SNPTNDKIKIEVYLVNGA										
Hel308 Tba	(247)	Q-----	WSSIVDAKTKRGKQ-----	ALVFVNSRRNAAEAAARLANS--	KYVRRLITPEARIRKELAEISLE--	SNPTNDKIKIEVYLVNGA										
Hel308 Tga	(226)	N-----	WSSIVDAKTKRGKQ-----	ALVFVNSRRNAAEAAARLANS--	KYVRRLITPEARIRKELAEISLE--	SNPTNDKIKIEVYLVNGA										
Hel308 Tsi	(249)	Q-----	WSSIVDAKTKRGKQ-----	ALVFVNSRRNAAEAAARLANS--	KYVRRLITPEARIRKELAEISLE--	SNPTNDKIKIEVYLVNGA										
Hel308 Mja	(225)	VDN-----	NDIVNLVDCVKEGGC-----	CLVEFONSKRNGAIVSEAKLIN-----	LAKKYLISNKEYELOKLEELISTIDPPTETCKTIAECILNGS	-----										
컨셉서스	(286)	LVLDITV	EGGQ	LVF	NSRRNAE	AKKIA	V	K	LT	E	L	IAEET	ETETS	LA	CV	KG

Hel308 Mbu	(307) AFHHAGLNSNH-----	-RKLYENGFRQNLIVISSTPTLAA-----	475
Hel308 Atu	(300) AFHHAGLINGQ-----	-RRVVEDAFRRGNIKVVAATPTLAA-----	
Hel308 Cay	(313) AFHHAGLINQDC-----	-RSVVEEFRSGRIKILIASTPTLAA-----	
Hel308 Dch	(353) DIGHDIDCLIVGYPGSWATMQBGRVGRSGRDSAINLIGHEBALDQYLINRPREFFSLEPESAVINPDNBSIMRHLVCAAEKPTALQEMMD		
Hel308 Fac	(290) MEHHAGLSNDQ-----	-RTIMEKLEKQGYIKILATPTLAA-----	
Hel308 Hla	(315) AFHHAGLRSED-----	-RARVEDAFRDRLIKICISATPTLAA-----	
Hel308 Hpa	(309) AFHHAGLASDH-----	-RSIVEDAFRDRLIVISATPTLAA-----	
Hel308 Htu	(310) AFHHAGLSSTQ-----	-RSIVEDAFRDRLIVISATPTLAA-----	
Hel308 Hvo	(315) AFHHAGLIAEH-----	-RTIVEDAFRDRLIKICATPTLAA-----	
Hel308 Hyc	(307) AFHHAGLTSPL-----	-RELIVETGREGYVKLISSTPTLAA-----	
Hel308 Mba	(307) AFHHAGLTPPL-----	-RELIVEDGFRAGRILKISSTPTLAA-----	
Hel308 Mbo	(300) AFHHAGLSRKE-----	-RSIVAEAFRRNLKICISSTPTLAA-----	
Hel308 Mev	(307) AFHHAGLNSAQ-----	-RKIVEDNFRNNKIKIVISSTPTLAA-----	
Hel308 Mfe	(307) AFHHAGLTYQH-----	-RKIVEDAFRRNLKIVICOTPTLSA-----	
Hel308 Mfr	(295) AFHHAGMNRMQ-----	-RTIVEGGFRDGFIKISSTPTLAA-----	
Hel308 Mhu	(300) AFHHAGLIRQE-----	-RTIIEEGFRNGYIEVIAATPTLAA-----	
Hel308 Mlg	(308) AFHHAGLTYEH-----	-RKIVEEGFRNNKIKIVICOTPTLSA-----	
Hel308 Min	(289) AFHHAGLTYEH-----	-RKIIIEKAFRRILKIVICSTTTLAF-----	
Hel308 Mma	(304) AFHHAGLTYKH-----	-RKIVEDGFRNNLIVICOTPTLSA-----	
Hel308 Mmah	(307) AFHHAGLINSAH-----	-RRIVEDAFRRNLIMICSTPTLAA-----	
Hel308 Mmar	(310) AFHHAGLSRGH-----	-RELIVEDAFRDRLVVCATPTLAA-----	
Hel308 Mmz	(307) AFHHAGLTPPL-----	-RELIVENGREGRIKILISSTPTLAA-----	
Hel308 Mck	(311) AFHHAGLTYEQ-----	-RKIVEEGFRKKLIRALICOTPTLSA-----	
Hel308 Mch	(302) AFHHAGLTPEQ-----	-RRLIIEGFRQVVVVIATCTPTLAA-----	
Hel308 Mzh	(307) AFHHAGLNSER-----	-RRMVEEGFRKNLIMISSTPTLAA-----	
Hel308 Nma	(310) AFHHAGLSSTQ-----	-RSIVEDAFRDRLIVISATPTLAA-----	
Hel308 Nth	(318) SSELGIDIGA-----	-IDYVVOIDDTFTVASLQGRSG-----	
Hel308 Pfu	(298) AFHHAGLGDE-----	-RVLVEENFRKGIKAVVATPTLSA-----	
Hel308 Sso	(316) AYHHAGLSKAL-----	-RDLIEEGFRQRIKIVVATPTLAA-----	
Hel308 Tpa	(320) AFHHAGLGRAE-----	-RTLIEDAFREGILIVITATPTLAA-----	
Hel308 Tga	(299) AFHHAGLSHYE-----	-RTLIEDAFREGILIVITATPTLSA-----	
Hel308 Tsi	(322) AFHHAGLGRTE-----	-RSLIEDAFREGILIVITATPTLSA-----	
Hel308 Mja	(305) AFHHAGLTYQH-----	-RKIVEDAFRRKLIVICOTPTLSA-----	
컨설팅스	(381) AFHHAGL-----	R IYVEDAFR LIRVI ATPTLAA-----	

Hel308 Mbu	(342)	-----GNTLPARVITIRSYRRFDS-NFG-----	-MQPIPVLE
Hel308 Atu	(335)	-----GNTLPARVIVRSIYRFDG-YSK-----	-RIKVSSE
Hel308 Cay	(348)	-----GNTLPARVIVSSVMRYNS-SSGM-----	-SEPISTILE
Hel308 Dch	(448)	NEAGKCIKSLKNDGELLASRDSFYITRARIPEHKDVLKGTGQTYNTFEHSTGEYLGENDGVPAFKETHPGAVYLMGETVYVQDLDETRAVYA	
Hel308 Fac	(325)	-----GNTLPARTVILRDITKFSQD--GY-----	-SKPISGLE
Hel308 H1a	(350)	-----GNTPSRRVIVDWMRRYDG-EFGG-----	-MKPLDVLE
Hel308 Hpa	(344)	-----GNTPSRRVIVDWMRRYDG-DIGG-----	-MQPLDVLE
Hel308 Htu	(345)	-----GNTPSRRVIVDWMRRFDP-SAGG-----	-MAPLDVLE
Hel308 Hvo	(350)	-----GNTPSRRVIVDWMQRYDG-DYGG-----	-MKPLDVLE
Hel308 Mac	(342)	-----GNTLPARVITIRSYRRYS-DSG-----	-MQPIPVLE
Hel308 Mda	(342)	-----GNTLPARVITIRNYRRYS-EDG-----	-MQPIPVLE
Hel308 Mdo	(335)	-----GNTLPARVILRDYLRFSQ-GEQ-----	-MQPIPVSE
Hel308 Mey	(342)	-----GNTLPARVIVRNRYKRYDP-NFG-----	-MQPIPVLD
Hel308 Mfe	(342)	-----GNTLPCRRALVDDITRYT--NRG-----	-MRYPIME
Hel308 Mit	(330)	-----GNTLPARVILRDYLRYSQ-GEQ-----	-MRPIPVRE
Hel308 Mtu	(335)	-----GNTLPARVILRDYLRFSQ-GLG-----	-WVPIPVGE
Hel308 Mfg	(343)	-----GNTLPCRRALVDDIMRFS--NGR-----	-MKPIPVME
Hel308 Mtn	(324)	-----GNTLPCRRVITISBLKRYT--RRG-----	-LTYPIVME
Hel308 Mma	(339)	-----GNTLPCRRALVDDIKRYS--QNG-----	-LVDIPVME
Hel308 Mmah	(342)	-----GNTLPARVITIRSYKRYDP-NAG-----	-MQPIPVLD
Hel308 Mmar	(345)	-----GNTPSRRVIVDWMRRYDG-SAGG-----	-MAPISVLE
Hel308 Mmaz	(342)	-----GNTLPARVILIRSYRRYS-DSG-----	-MQPIPVLE
Hel308 Mck	(346)	-----GNTMPCRRALIRDIKRFS--SRG-----	-YIPIPVME
Hel308 Mch	(337)	-----GNTLPARVITIRSYKRYEA-GLG-----	-TRPIPVME
Hel308 Mzh	(342)	-----GNTLPARVILIRSYKRYDP-NFG-----	-MKPIPVLE
Hel308 Mna	(345)	-----GNTPSRRVIVDWMRRFDP-SAGG-----	-MAPLDVLE
Hel308 Nth	(353)	-----RKGTNQVLOVYSTNDLSVQSLA-----	-VIDLILEK
Hel308 Pfu	(333)	-----GINTPAFRVITIRDIMRYS--DFG-----	-MERIPVIE
Hel308 Sso	(351)	-----GNTLPARTVILIGIYAFNKIKIAGY-----	-YDEIPVME
Hel308 Tpa	(355)	-----GNTLPSEFRVITIDTKRYS--TRG-----	-WSDIPVLE
Hel308 Tga	(334)	-----GNTLPSEFRVITIDTKRYA--GRG-----	-WTDIPVLE
Hel308 Tsi	(357)	-----GNTLPARVILRDYLRYS--NFG-----	-WVDIPVLE
Hel308 Mja	(340)	-----GNTLPCRRALVDDITKFSQD--KG-----	-MRYPIME
컨센서스	(476)	GNTLPARVITIRDYKRY G	M PIPVLE

476

570

He1308 Mbu (372) YK0MAGRAPHILDPYGEVLIATYDEF--AQIMENYVEADADIMSKIGTENALRTHVLSTIVNGFASSTR0ELFDFEGATFEAY0Q-DKXMLE  
He1308 Afu (363) YK0MAGRABRPEMDRGRGAIIVGRUR---EIAVKRYIGEPEPRTISKIGVETHLRFHSLIICDYGAKTIELEDEFDADTFEKKON--EISLS  
He1308 Cay (379) YK0LGRABRPQYDKSGEAIIVGVNAD---ELFDYIGGEPEPRTISAMVDDRALRHHVLSVTTSPGICKEDVTEPEFLIGG0QS-GE5TVK  
He1308 Dth (543) AKSEANYITRPTTEKYTEIVEV0ATRAIAGELCLGRLKYTEHVSAYEKRLVHQ0ARIGLIIPDLPEPLVFEYQ0KMFILDSQVRDVEDRRLHFM  
He1308 Fac (354) I00MIGRABRPYDKKGYGYIYA5PG---MLRAEYIGELPEYVSRMDSNSLIRFNVALISSGIATD0KGI0DFYGKTLA0N-DIDGYE  
He1308 Hla (381) VH0MGRABRPGLDPYGEAVILANDADTK--EELFEFYIWMADPEPVPYRSKLAEPALRTHVLATVASGFASSTRDGLSFLDNTLYATQTDDEGRLA  
He1308 Hpa (375) VH0MGRABRPGLDPYHGAVALIAKSHDEL--QELFDQYWMADPEPVPYRSKLAEPALRTHVLATVASGFAGTEEBELDPLERTIYATQTDDEGRLE  
He1308 Htu (376) VH0MGRABRPGLDPYGEAVILIAKSHDES--EELFDYIWMADPEPVPYRSKLAEPALRTHVLATVASGFAGTRGGLLEFLAATLYASQSSGARLE  
He1308 Hvo (381) VH0MGRABRPGLDPYGEAVILIAKADAR--DELFEFYIWMADPEPVPYRSKLAEPALRTHVLATVASGFAGTRGGLLEFLAATLYASQSSGARLE  
He1308 Mac (372) YK0MAGRABRPGLDPYGEAVILIAKSYEEL--LFLFEYITAGADIMSKIGTENALRTHVLSTISNGFARTYDELMDFLEATFFAF0Y-SNFGIS  
He1308 Mba (372) YK0MAGRABRPGLDPYGEAVILIAKSYEEL--VTLFEYITANADIMSKIGTENALRTHVLSTISNGFARTYDELMDFLEATFFAF0Y-SNFGIS  
He1308 Mbo (365) YK0MAGRABRPGLDPYGEAVILIAKAEQV--PELFEYIIEAEDVHSRLAEPALYTHVLSIVASGFAGTRGGLTEFMNBSFYVHEKQ0RLIH  
He1308 Mey (372) YK0MAGRABRPGLDPYGEVLSISHTYNEF--TDLDYRTIDAEPPDILSKIGTENALRTHVLSTIVNGFATRQ0QAVDFMSSSFFAY0Q-0KMSLI  
He1308 Mfe (371) I00CIGRABRPGLDPYGEGLIYAKNDR---DYLASQYVLQ0KPEPIYSKLSN0AVIRQ0LIGLIIATIEIRDEYDLEMTIRNTFYAYQYKNI0EVA  
He1308 Mfr (360) YK0MAGRABRPGLDPYGEAVILIAKTEYAV--NDLHEEYVAPEDVYSRGEGVLTAAHLSLIAFGARYSD0ELMAFLKTLIYAYQHTGRKALT  
He1308 Mtu (365) YH0MGRABRPGLDPYGEAVILIAKADAP5V--ERLFEFDIEAEKRVDS0CVDASICAHLISLIATGFADQ0ELASSMEKTFYFF0HPTKRSIP  
He1308 M4g (372) IH0CIGRABRPGLDPYGEGLIIFVNKR---DLERAEQYIEGKPEPIYSKLSN0AVIRQ0LIGLIIATRIEYDEYDLEMTIRNTFYAYQYKNI0EYK  
He1308 Min (353) V00CIGRABRPGLDPYGEGLIYAKDER---DYRALQ0LQ0KPEPIYSKLSND5VIRQ0LIGLIIATRYLDEYDLEMTIRNTFYAYQYKNI0EYK  
He1308 Mma (368) I00CIGRABRPGLDPYGEGLIYKNER---DAEKAYELIGSVENIYSKLSANQ0VLRHILGILSTGEIKDQ0NLVNMKNFTFYAHQ0GIN0AVIL  
He1308 Mma (372) YK0MAGRABRPGLDPYGEAVIIVKTYEEF--TDVLEKYISASADIMSKIGTENALRTHVLSTIASGFANCHREILITFLG5FFFAHQ-0S0MPE  
He1308 Mmat (376) VH0MGRABRPGLDPYGEAVILIAKSHDEY--DELFEKYIWMADPEPVPYRSKLAEPALRTHVLATVASGFARSKRGLLEFLBQTLIASQTD0SG0LE  
He1308 Mmaz (372) YK0MAGRABRPGLDPYGEAVILIAKSYEEF--VTLFEYIIEAGADIMSKIGTENALRTHVLSTISNGFARTREELMDFLAATFFAF0Y-SNFGIS  
He1308 Mck (375) IH0CIGRABRPGLDPYGEGLIYINNTEPDELINAKNYELIGNVEIYSKLSN0K0VLRTHMLGLITGDIKNKNDLEEFIKNTFYAYQYKNTKIL  
He1308 Mch (367) YK0MAGRABRPGLDPYGEVLSIMARSESL--Q0KIMDYHWGEPPDIMSKIASERALRTHVLATIASRFD5VDSLRMASTFYARQ0-DPSYIG  
He1308 Mth (372) YK0MAGRABRPGLDPYGEVLSIARSYDEF--MDIMENYVADPEPDIIMSKIGTENALRTHVLSTIVNGFARTYRGIMDFYKMTFFAYQ0-EASDILH  
He1308 Mna (376) VH0MGRABRPGLDPYGEAVILIAKSHDES--QELFDYIWMADPEPVPYRSKLAEPALRTHVLATIASGFATREGLLEFLAATLYASQSSGARLE  
He1308 Nth (385) WTEPATEYVLPDIIIFHOIISICHEANGVRDPLINDINKNAAFYKIKEDDINHVINMIENDFQ0LRNSAEILVGEGERLILNGKEFYAVNMT  
He1308 Pfu (362) VH0MGRABRPGLDPYGEGLIYISD0---PREYMANIITGKPEKLF5Q0LSNESNIR5Q0VALIATFGYSTVEELIKTISNTFYAYQ0RKDTYSLE  
He1308 Sso (383) YK0MGRABRPFPD0IGESIIVAD0KEV--DNYFKKYVLSDEVPLESKISERAFYTLGLISAEGLNSEK0LENFAYESILAK0L-----VD  
He1308 Tpa (384) I00MIGRABRPYDK0EGEAIIVAKTER---PEELMEKYISLDEAPKLF5MSLNDAPRSOVALITNFGVSPRELIGLEKTFYH0RKD0LEIE  
He1308 Tga (363) I00MGRABRPYDKY0GAIIVARDE---P0KIMEYIINGKPEKLF5MSLANQ0AFRSQVALITNFGIRSPPELVPRFLERFYAHQ0K0LSILE  
He1308 Tsi (386) I00MGRABRPYDIEG0AIIVIAKTER---PEDIMKRYVALGKPEKLF5MSLSN5ASFR5QVALITNFGVGNFKELVNFLERTFYAHQ0K0LSIALE  
He1308 Mja (369) I00CIGRABRPGLDPYGEGLIYAKNDR0Y---L0AY0ALQ0KPEPIYSKLSN0AVIRQ0LIGLIIATIEIRDEYDLEMTIRNTFYAYH0YGNIR0EVA  
He1308 Mja (571) I QM GRABRP LDPYGEAVILIAKS D EL E YI ADPE IMSKIA E ALRTHVLATIASGFA T ELIDFL TFYAYQ L

571

665

666

760

Hel1308 Mbu (464) EVINDCLEFLIDKAMVSET-B-----DIEDASKLFLRGTRLGSLVSMLYIDPLSGSKIVDGF  
 Hel1308 Afu (453) YELERVVRQLEMMGVVVAAH-----IAPTKGLSVLSRLYIDPLNGFTPHDYL  
 Hel1308 Csy (469) FSVAAVLRPLQEGMLGR-----GRLAATKMGRLVSRLYMDPMTAVTIRDAV  
 Hel1308 Dth (638) GGIHAHEHGLIGCMPLIITDRNDLGGIASFVHEQHLKG-----AVFIYDTPGGIGLCROAFELGDRVAPARANGILSCCENG  
 Hel1308 Fac (445) IAFESALYFKNDPFIETEN-----DIYSAKTFGRLLTSRLYIDPVSSLIAKCL  
 Hel1308 Fla (474) AVTDVLYLAVNDPIERDR-----GGSESLTATGIGHTVSRLYIDPMSAAEMIDGL  
 Hel1308 Hpa (468) TVTQVLDYLRNGFLERD-----RLRANGLGHRVSQLYIDPMSAAEIIIDGL  
 Hel1308 Hru (469) SYTDVLDYLRNDPIERSR--DDEADSGEDDGPFTSADILAQ-----QAAK-----REETLEATSLGHTVSRLYIDPMSAAEIVHGL  
 Hel1308 Hvo (474) QYTDRLVLYLRNGVFEFG-----ETIQATPVGHTVSRLYIDPMSAAEIIIDGL  
 Hel1308 Mac (464) VVDECLNPLRQEGMLEQDS-----DALISTMFGLVSRLYIDPLSAALIAKGL  
 Hel1308 Mba (464) TVNECLNPLRQEGMLEKD-----DALIPTSFGKLVSRLYIDPLSAARIAKGL  
 Hel1308 Mbo (458) RAIDELAQELITAEVYEV-----GEHIGATELGTVSRMYIDPRSAFAIVTTL  
 Hel1308 Mev (464) DVVDCLEFLQDNEMIKD-----DG--ER--LVATRLGQVISTLYIDPLSGALITIDKL  
 Hel1308 Mfe (463) KINNEVIRFLIEEK-----EMIDFIPTELGKRVAEIYIDPLSAKWMIDGL  
 Hel1308 Mfr (453) RLDDALGFLTEAEVMDL-----SGMLHATEVGDLSRLYIDPHSAEITTTAL  
 Hel1308 Mtu (458) RLVAADAIRFLTTAGVEER-----ENTISATRLGSLVSRLYINPCTARLITDGL  
 Hel1308 Mlg (464) RNIKEVINFLIEN-----DFIADYFPFKLGKRVSELYIDPLSAKIIIDGL  
 Hel1308 Min (445) KKIKELIEFLIEN-----FIKNPEVTPPLGKRVSELYIDPLSAKIMDNI  
 Hel1308 Mma (460) LVSEVVEFLERKLETTIKTKENKVRLEISDS-----S--NN--LVDSKETSPDLTPNPSNIEFRSKLGKRISELYIDPMSEIIIEEL  
 Hel1308 Mmah (464) ELLEDCLLFLKNEGMLEQD-N-----ET-----IRATELGKMLISSLYIDPLSASKIIRGL  
 Hel1308 Mmh (469) RVVDVLYLQDNDFLEIAG-----EIDATSLGHTVSRLYIDPMSAAEIVDGL  
 Hel1308 Mmz (464) AYVDECLDFLRREGMLEKD-----DALVSTVPFGKLVSRLYIDPLSAALIAKGL  
 Hel1308 Mox (470) ENIYELTNFLKNGFTLEINPRDENKDKSNNSHNKKNISNTNNSITKMLYLDNNNSLTITKSRHEEDVYINTEPLGKRVSELYIDPLSAEYIIDGL  
 Hel1308 Mzh (459) ETIASYLEFLVSDMIDD-----LTPPLGALVSRLYIDPLSAMWIDGL  
 Hel1308 Mzi (464) DYIEGCVRLIDNEMIIISD-S-----NDILPES--AFRSATGKLISMLYIDPLSGSLIMOGI  
 Hel1308 Nma (469) RYTDVLYSLERNDPIERSGPEDTLNSADASAFISADLADS-----DGDGSGCTTGQEDIEATSLGHTVSRLYIDPMSAAEIVHGL  
 Hel1308 Nth (480) QEEFEVRREGIRKIGSIDKS-----LVWSEGNIIITAGQMTIKINDIERDIIVAKA  
 Hel1308 Pfu (454) EKIRNLIYFLIEN-----EFIEISLEDKIRPLIGIRTAKLYIDPYAKMFKDKM  
 Hel1308 Pfu (471) VYFDRAIRMLLEHSTIKKE-----GNTPALNFGKRVADLYINFTYADITRKGL  
 Hel1308 Sso (476) GKAKSIVYFLIEN-----EFIDIDINDSFIALPFGIRTSQLYIDPLTAKKFKDAL  
 Hel1308 Tba (445) YKAKVYVFLIEN-----EFIDIDLEDRTIPLPFGKRTSQLYIDPLTAKKFKDAF  
 Hel1308 Tsi (478) GKAKSIVYFLIEN-----EFIDIDINDQEMPLPFGIRTSQLYIDPVYAKKFKDAF  
 Hel1308 Mja (461) KINNEVIRFLIENFT-----IDMPETELGKRVSELYIDPLSAKFTIIDGL  
 (666) I EVL FL N I I AT IG VS LYIDPLSA IIDGL

컨센서스  
 (666)

Hel308 Mbu	(520)	KDIGKSTGNNMSLEDKG-----	DDITVTDMTLHLVCSIPDMRQLY
Hel308 Afu	(501)	SMELIS-----	DIGALHLICRTPDMERLT
Hel308 Csy	(518)	GEASPGR-----	MHTIGFLHIVSECEMPPRF
Hel308 Dth	(719)	PGCHSPKCGSGNR-----	PIDKEAMHMLAVIAGERGE
Hel308 Fac	(494)	DLEFS-----	ELIYIYISKTPEMLTFN
Hel308 Hla	(527)	RSVARDADTGASAEADNG-EFVRTGADADASGGDEFGFTYRAGDDESGER----	ETENEETDEETEASEVTPLGLYHLISRPDMYELY
Hel308 Hpa	(516)	RDADG-----	KPTALGLYHLVSRPDMYQLY
Hel308 Htu	(547)	ERADER-----	PTALGLYQVLSRTPDMTELY
Hel308 Hvo	(523)	EWADHRTKRLALAGETPEKPTDRSESDSGGFQASSEMVDADGGGGEDVGANGDGDSDADGVEDTDRTYPTPLGLYHLVCRTPDMYQLY	
Hel308 Mac	(513)	REAGT-----	LTELTLHLVCSIPDMRLMY
Hel308 Mba	(512)	KGAKS-----	LSRLTLHLVCSIPDMRLIY
Hel308 Mbo	(507)	REQER-----	YADLGLIQLICTPDMPTLY
Hel308 Mey	(513)	KKADK-----	VTDMTMLHLICSTPDMRQLY
Hel308 Mfe	(508)	NEMENED-----	DIYIYIYISKTLEMMPYL
Hel308 Mfr	(502)	REEGE-----	LTDLALLQLCMTPEMTLY
Hel308 Mhu	(507)	KSCKT-----	PTLIGLHVICVSPDMQRLY
Hel308 Mig	(509)	KEMGNVNE-----	ELIYIYIISKTLEMMPYL
Hel308 Min	(490)	EYKDDLH-----	LIYIICKIEMKPLL
Hel308 Mma	(545)	HELKRCCKQDLR-----	SKIDQYLFYLI SKTNEMRPLL
Hel308 Mmah	(513)	EKTTH-----	VTDMTLIOLICSTPDMRRLIY
Hel308 Mmar	(518)	RDWERGASDSTSASGSPAD-----AQAEF-PANSGETTASFLADADESDADR-----	PDDISALGLYHLVSRTPDMYQLY
Hel308 Mmaz	(513)	REAGT-----	LTELTLHLHLICSTPDMRLMY
Hel308 Mok	(565)	KULHKTTLSNPKNM-----	ECYLIHLIYIISKTTEMQPYL
Hel308 Mth	(505)	RGIRR-----	PYVLTILHVTMTPDMEELF
Hel308 Mzh	(519)	KRADY-----	FEDITMMHLICSTPDMKNLY
Hel308 Nma	(555)	EDADER-----	PTALGLYQVLSRTPDMYELY
Hel308 Nth	(532)	VDGKPPK-----	YSGGGFLNPKI PERMKLIL
Hel308 Pfu	(504)	EEVYKDPN-----	PIGIFHLISLTPDITTPN
Hel308 Sso	(520)	EGHKAS-----	CEIAYIHLIAFTPDGPIVS
Hel308 Tba	(526)	PQIEENPN-----	PLGIFOLLASTPDGPIVS
Hel308 Tga	(505)	PAIEENPN-----	PFGLFOLLASTPDMAIT
Hel308 Tsi	(528)	EKLEKNPN-----	PLGIFOLLASTPDMSIRL
Hel308 Mja	(506)	EEMENEE-----	ELIYIYIISKTLEMMPYL
컨센서스	(761)		LGLHLIS TPDM LY

761

855



Hel308 Mbu	(563)	LRNTDYITVNEIYVAHSDEFH---ELPDKIKETNDYEMNGEVKTAMLLEW-----VTEVSAEDITRHFNVGEGDIALADTSEM	856	950
Hel308 Afu	(525)	VRKTDSWVEEAFPLRKELISY---YPSDFS-VEYDMFLSEVKTALCKDM-----IEKDEDEICAKYGIAFGDLRIIVETAEW		
Hel308 Csy	(545)	ALRKQDHEVAMMLAQRGELLIR---P-----YYSYECGRGLIALHR-----IGESPEAKIABDIKRESGVHRMVESSGM		
Hel308 Dth	(754)	AKRKDVSCRIETDGSMEIDSG-YTKSDQALPYALDIETRYSAQEVGWCNCRMGVSAVVPDSNRNQEVFTDQADLGSFLEDSLVVG		
Hel308 Fac	(517)	YRASDYETLEETLRHNISDFS-----ESMAAKTALILNEW-----INEVPINTIAETFGIGPGDIQAKASADW		
Hel308 Hla	(615)	IRSGDBREYTEL CYERETETLG--DVPSSEYEDVRFEDWLSAKTARLLEDM-----VNEVEDRITERYVGPDIIRGKVDTAEW		
Hel308 Hpa	(542)	IRSGDBREYTEL IAYEREPEFLG--HNPSEEDNAFEDWLSAKTARLLEDM-----ASELDEDRITERYVGPDIIRGKVDTAEW		
Hel308 Hpa	(542)	IRSGDBREYTEL IAYEREPEFLG--HNPSEEDNAFEDWLSAKTARLLEDM-----ASELDEDRITERYVGPDIIRGKVDTAEW		
Hel308 Htu	(573)	IRSGDBREYTEL CYERETETLG--DAPSEYEDRFDWLSAKTARLLEDM-----ADETDEETITDRYKIGPDLIRGKVDTAEW		
Hel308 Hvo	(618)	IRSGDBREYTEL CYERETETLG--RVPSSEYEDVAFEDWLSAKTARLLEDM-----VEVEDRITERYVGPDIIRGKVDTAEW		
Hel308 Mac	(538)	MRSQDYQDINDFVAHAKEFS---KVPSPNIVAYEWFTLSEVKTALLMDW-----IHEKPENEICLKFGEEDHTTADIAEW		
Hel308 Mbo	(537)	MRSQDYQDINDYVAHAKEFV---KVPSPDTEYEWFTLGEVKTALLMDW-----IHEKSENEICLKFGEEDHSHADIADIAEW		
Hel308 Mbo	(532)	AKNADIPALSMLEVRGADIW---LPP-PIDDDAETTYRAVKTALLMDW-----TDELSEKICERYGVGPDIINAVQGISW		
Hel308 Mbo	(538)	IRSKREYKINCYVTHDEYV---EYVNPFKSIEYEWFTLGEVKTALLINEM-----IDERTLDDITAEFGVGEDINALSDISEW		
Hel308 Mfe	(533)	RVKRSSE--LNLIDEMENIG---IKSFE---IEDLEAFKTAAMLNDW-----ISEVPEDEILKKYKIEPGLIRKYVENAVW		
Hel308 Mfr	(527)	VKKNDIGTLEKFFEHHEEFR---T--EFSYDEMEDFRSLKTAMLISDM-----TDEIGDDTICTRFGVGPDIINAVQGISW		
Hel308 Mhu	(532)	LKADDTQLLRFTLFKHHDDL---LPI-PEQGESEELMTSGUKTALVITDM-----ADEPSEGMIEERYGIGAGDLVNIYDSGM		
Hel308 Mlg	(536)	RVNSFEF--LDLITEMEAG---IYDRT---YDILAAFRNAKMLYDM-----INEVPEDEILKKYKIEPGLIRKYVEQAKW		
Hel308 Mln	(512)	RVYRKEE--ELIAEELINYE---I-FIS---YENIEEFTAKMLYDM-----INEVPEDEILKKYKIEPGLIRKYVEQAKW		
Hel308 Mma	(578)	RIRPNEE--LDLITELMDKMG---IKOYS---TENIEAFKNSKMPGDM-----VSEIPEEITIEKYGVGPGLIRKYVEQAKW		
Hel308 Mma	(538)	LRNRDYEILINDYVANHTEFT---EYVSPFKQIEYEWFTLSEVKTALLILEW-----INEKSLKETIVENYQVEGDYIASSDIAEW		
Hel308 Mma	(589)	IRSGDBREYEMELFEREETILG--PVPSEFEERFEDWLSAKTARLLEDM-----ATEVDEATITDRYGVGPDIIRGKVDTAEW		
Hel308 Mma	(538)	MRSQDYQDINDYVAHAKEFS---KVPSPNIVAYEWFTLGEVKTALLMDW-----IHEKPENEICLKFGEEDHSHADIADIAEW		
Hel308 Mma	(600)	RVYRKEE--NDLINDMTKIDIDVDVYIGIS---SENIEYKNAKLFYDM-----INEIPEEELILGYNIEPGILIRKYVEQAKW		
Hel308 Mth	(530)	VQQS--DNNWLEDFISEHSESLG---NEKN---FVWLIREVKTASMLDM-----INEVHEDRITERYVGPDIIRGKVDTAEW		
Hel308 Mth	(544)	MRSQDYQDINDYVAHAKEFS---SMSPFKMIEYEWFTLGEVKTALLMDW-----INEVPADICTKRYGIGEDIRMFSETAEW		
Hel308 Mth	(581)	IRSGDBREYTEL CYERETETLG--DAPSEYEDRFDWLSAKTARLLEDM-----ATEDEDEQITERYKIGPDLIRGKVDTAEW		
Hel308 Mma	(559)	CERNKFEFTDMAQNHLIEQR---IEEPSEYEDVSLYNALKVALIMDM-----PNERVIMNNGDILLETYTGKTIQOTLAW		
Hel308 Mth	(530)	YSKREFELEEEYEFEDRLYFDPPYISGDPYLERKFRFAFTALILMDW-----INEVEGEIIVERYSVGPDIIRGKVDTAEW		
Hel308 Pfu	(545)	VGRNBEELITELLDLCELL---IEEPSEYEDVSLYNALKVALIMDM-----MDEVDEDITILSKYNTSGDLINMVEITMDW		
Hel308 Sso	(552)	IKRKEQESYLDIAYEMEDYIYRSIPIWEDYE---FQKFTISEVKTAKLILDM-----INEVSEAKIIEAYGIDTGDLYRIITELADW		
Hel308 Tga	(531)	ARRREMEYLDIAYELDDKLYASIPYEDSR---FQFTLGQVKTAKLILDM-----INEVPEARIYETYSIDPGLYRIITELADW		
Hel308 Tsi	(554)	VKRKEOEDILDIAYEMEDYIYONIPWEDYK---FKFTLGQVKTAKLILDM-----INEVNDKILLETYEIDTGDLYRIITELADW		
Hel308 Mja	(531)	RVYNSSE--LNLIDEMDSLGIC-----SFETEDLEAKRYTAKMLYDM-----INEVPEDEILKKYKIEPGLIRKYVENAVW		
컨센서스	(856)	LR D E L E I E E F E FL VKTA LL DW I EV ED I ERYGIGPDL VE AEW		

951 1045

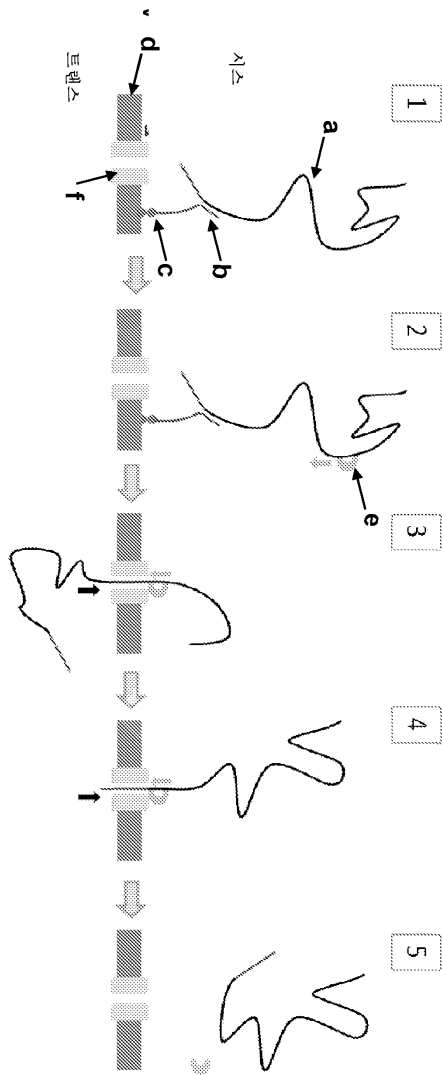
Hel308 Mdu (640) IMHAAKLAELIGVEYSS-----HAYSLERIRYSGJLIMELVIGIRGVYRARKLYNAGTYS-----YAKLKAGADISVLSKLYGP-  
Hel308 Atu (600) LSNAMNRIAEVVG-N--T-----SVSGLTERIKHGVKKELELVIRIHHIGVYRARKLYNAGIRN-----AEDIVHREKVASLIGRG-  
Hel308 Csy (614) ILRCIWEISKHÖERPDLIG-----EIDVIRSRVAYIKAKELPVPISIKIGIRVSRRLFRGIGK-----PGDIAVPYERLSRVEGIG  
Hel308 Dth (848) FNILKFDYVVOGLSDYDFSSLPITDMLRETEARIGHRLSDIHARHTLIGNKSANGMALKWKKEGELDKTIVEYCRQDVSVTRDLYLFRGDKGY  
Hel308 Fac (584) IYSYLRLSGMPDKENEN-----NILHINIRIKEGVKEELIRIIEIPQVGVRGRRLYNNGFS-----IDDIANARVEDISRLIGFS  
Hel308 H1a (693) ILRAETTLADYVEVDGVVV-----AVREARKRIEYGRELLIDLAGVNVGRKRARBLFEAGIET-----RADLRADKAVVLGALGR  
Hel308 Hpa (620) LINAERLAAELQORDAEGIPSAITTTAVREARKRYEVGREELIDLAGVNVGRKRARBLFEAGIES-----RADLRADKSVVLGALGR  
Hel308 Htu (651) ILGAESLAAEIDSEWTV-----AVREARARVEHGVGELELVYSGVGGRKRARBLYDAGIE-----PADLRADKGIYLSVLKG-  
Hel308 Hvo (696) ILGAERLATELD---LDSVY-----AVREARKRYEYGRELLIDLAGVGVGGRKRARBLFEAGVET-----RADLRADKRVYLAALGR  
Hel308 Mac (615) IMHATQULARLIDLKGA-----EAAELEKRIHYGAGPELMDLIDIGIRGVYRARKLYGAGFYS-----TADLAGATPEKYAALVGP-  
Hel308 Mba (614) IMHVTQQLAGLIDLKGA-----EAAELEKRIHYGAAPELIDLINIGIRGVYRARKLYEAGFYS-----SAALEAVDPEKYAALLGP-  
Hel308 Mbo (608) ILHATSQULARMPVPEKPYG-----QIADCEICMKNIGIRELLPIVRLKIGIRVYRARKLYENNIGTIS-----PEELSRHKEDLVKILIGS-  
Hel308 Mew (615) IMHSAVNIANTIDLDAD-----KAQELEKRIHHGVNKKDLIQVYSINIGIRVYRARKLYNAGIYS-----VSDIKNTKLIHISNYLGR-  
Hel308 Mfe (601) IMHAKEMAKIIGKN-----SEIPEKLEIRLEYGAKEDITEILNVYKIGIRVYRARKLYNAGIRN-----VEDIINNPK--VASIIG  
Hel308 Mtr (601) LHAQSRILAVVAPEHRD-----AVEETTLARHGIRELLPIVAVKIGIGVYRARKLYENNIGTG-----PELLAADPSPVGHIVG-  
Hel308 Mnu (608) ILHGTERRIYSVEMPEMSQ-----VVKTLGVYHVGKSELPIVALNNIGVYRARKLYNAGIYPD-----PEAVARAGLSTIARLIGE-  
Hel308 Mng (604) MYSTKEIKAKLINRN-----IDTISKLEIRLEYGAKEDITEILKIRYGVYRARKLYDAGIRS-----VEDIINNPK--VASIIG  
Hel308 Mtn (579) LYSLSKEIKAKLINKEVP-----NLEIRLEYGAKELIIEILIKIRYIGVYRARKLYSAGIRS-----REDIINKPK--VANILIG  
Hel308 Mma (646) MYSTKEIKAKLIHLNSE-----IKSILKMEVRIEYGAKEELTEILNVNVGIRSRKLYDAGIRS-----KIEINNPK--IIEILIG  
Hel308 Mmah (615) IMHATQRIASRINPOLET-----ECAALEKRIHYGAGSELTEIIVNPGYRARKLYFKGYS-----ROKLAFADEKOLAALVGP-  
Hel308 Mmar (667) ILGAESLASSEVDLDAAR-----AISEARIRVEHGVREELVDLAGVGVGGRKRARBLFQAGITD-----RAQJLRADKAVYLAALGR  
Hel308 Mmaz (615) IMHVTQQLAGLIDLKGA-----EASELEKRIHYGAPELMDLIDISVGVYRARKLYEAGFYS-----TAEIASAPEHLAVLGP-  
Hel308 Mok (674) MISHAKEIFNINLIDNKV-----IQCJNDLEIRMEYAKODITEILKIKHIGYRARKLYNAGIRN-----ANDIINNOK--IINILIG  
Hel308 Mch (600) IMSAHLRISGKMDLGYTY-----IABRALRIHYGAGPELQJLEILKIGIRVYRARKLYOAGYRS-----IEDIKRADKSTSEILGP-  
Hel308 Mzh (621) IMHATSRLSGLIKVSEASE-----KSKELEKRIYSINSELVNIYALKIGIGVYRARKLYENGYS-----IDDKRADPLKLSKIVGS-  
Hel308 Mna (659) ILGAESLASEIDSEWAV-----AVREARARVEHGVGELELVYSGVIGIRVYRARKLYAAGIE-----PALIRSADKGVILHILKIG-  
Hel308 Nth (618) ILRSNVNVIKEIDIGERIN-----IEGGIDLPQVLODKEMDMPREVLIDFTLEGEKFSKSPYLPE-----KOLQDKMHTAHLVDIEGVK  
Hel308 Ptu (610) LYSLSKEIKAVIG-AVE-----IDVYETTLARVKYKIRRELLIPMQPLPIVGRRRARALYNSGFRS-----IEDISQARPEELIKTEGIG  
Hel308 Sso (621) LITSAYHLISRELIKLNHAD-----KIRIIRIARVPGIKKELELVYQISGVGRKRARLYNNGIFE-----IGDVVAMPDKVNNLIGQK-  
Hel308 Tpa (629) IMYSLIEIKLVNAGGE-----TKYVRIHLIRIKHGVRELELIEVLEIMPIGRARALYNAGFKN-----VNDIVAKPSELIAVEGIG  
Hel308 Tga (608) IMYSLIEIKLVLEPPEE-----IINYIRDLHLIRHGVRELELVYRLNIGGRARALYNAGFRS-----VEAIANAKPAELIAVEGIG  
Hel308 Tsi (651) IMYSLIEIKLVDPKPE-----VLDFTKULIRYKHGVRELELITLPMIGGRARALYNAGFKG-----IDDIVAKASELILKVEGIG  
Hel308 Mja (599) IMHAKELIYKIGSSDI-----PEKLEIRLEYGAKEDITEILSIRYIGVYRARKLYNAGIRS-----IEDIINNPK--VASIIG  
컨센서스 (951) LAKKL LAKKL L EL IRI YGVKELIEIV IR IGRVYRARKLY AGIRS DL A L ILG

He1308 Mdu	(717)	-KVANNILSIGVANDKHNNSAPISSTLTD	TLIDRNOQKTENDFO
He1308 Afu	(674)	-IAERVVEGISVKSINLPESAAALEHHHHH	
He1308 Csy	(693)	ATLANNIKSQLRKG	
He1308 Dth	(943)	ILFKNRACKRVIRIPVSMQDTAFV	
He1308 Fac	(662)	TRIAODIIEANGLNNRKYR	
He1308 Hla	(774)	ERTARILIEAHGREDBSMDVPRDKSAAATRG	-ASDEGEGQASIGFR
He1308 Hpa	(706)	KRTMENTILENVRQDPEISDVEADAT	-AA
He1308 Htu	(728)	EKTEMENTILENAGRDPDMDGVEPADQGPVGAATNGSSGSGBTBTRGADAAESDSSSIGF	-TSATATNDGQOQSIGFE
He1308 Hvo	(774)	KRTAENTILEAAGRDBSMDAIVEDDAPDAVPDDA	-G
He1308 Mac	(692)	KIAERIFRQJGREAVSEISDBERLKS	-PETAKERADQOASIGFEGS
He1308 Mba	(691)	KIADRIFKOJRGSGTSGIIASEPPEKS	-PYSGOKTISDV
He1308 Mbo	(685)	-GIAQVLEQJHPEBKDTGKKEPPSGKNTN	-PG-QSILFHRG
He1308 Mev	(691)	-KTAKVLEQJGVPEDNQOIDEPEISIKSY	-SGNDQOQKTENDF
He1308 Mfe	(675)	EKITKILIEDG	-IKRQ
He1308 Mfr	(678)	-KTAEISII	QKLIIF
He1308 Mhu	(685)	-GIAQVVIDETTVGRKSGIHSSDDYQOQT	PE-LIITDIPGIGKMAEKIUNAGIITVSDLIITADEVILLSDV
He1308 Mlg	(678)	EKIAKILIGELG	-MKRQ
He1308 Min	(650)	EKISKRIPEELG	-VRIGQ
He1308 Mma	(724)	EKIGKILIGERG	-MKRQ
He1308 Mmah	(692)	KIAOKTISYIGRETDGNGVYEPETLENK	QTLINFN
He1308 Mmar	(692)	KRTAENTILENAGHDPBMEGVEPAPDVSYDNDGADG	-ASA
He1308 Mmaz	(692)	KITERIFRQJGREAVSEFSDIEPELKG	-SSDQRTISDY
He1308 Mok	(752)	EKIAKILISELGVDTFRQ	-MLISI
He1308 Mth	(677)	-KIAEGVISQJK	-EPGISA
He1308 Mzh	(699)	-KISQKILIKQJIDVIDISLTKERQSDTVP	-E
He1308 Nma	(736)	EKTEMENTILENAGREPEMDGVEPIPEVGGSGGSSNSSGSPBNADATADDDQSSIGDF	-P
He1308 Nth	(699)	TYLEKKRIKEIKL	
He1308 Pfu	(689)	KVTVYEAJFKFLGNKVKISE	KPKRSTLUDYFLNS
He1308 Sso	(699)	-IGERVQVEARILNRFH	
He1308 Tba	(709)	KVYLERIYRHHGVGLPLTKNIKDPDKPEKPKRP	KPKRGTLDFYK
He1308 Tga	(688)	AKIJDGJYRHLGIEKRYTE	-EK
He1308 Tsi	(711)	IGVLEKTIYQHGVLPTNE	-KK
He1308 Mja	(673)	EKIAKILIDELGVYFGQOKISFSGGSAMSHPDPEKGGSGGSGGSAMSHQPEK	-KL
	(1046)	KIAEKIL	IG
			TL
			F

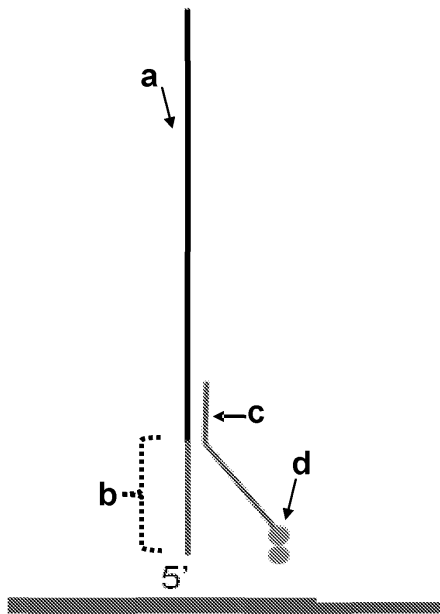
	1141	1186
Hel308 Mbu	(761)	-----
Hel308 Afu	(703)	-----
Hel308 Csy	(708)	-----
Hel308 Dch	(967)	-----
Hel308 Fac	(682)	-----
Hel308 Hla	(825)	-----
Hel308 Hpa	(753)	-----
Hel308 Htu	(792)	-----
Hel308 Hvo	(830)	-----
Hel308 Mac	(731)	-----
Hel308 Mba	(730)	-----
Hel308 Mbo	(724)	-----
Hel308 Mev	(734)	-----
Hel308 Mfe	(697)	-----
Hel308 Mfr	(685)	-----
Hel308 Mhu	(754)	IGAARAKVLAFLNSSEKENSDDKTEIIPDTOKIRGQSSWEDPGC
Hel308 Mig	(700)	-----
Hel308 Min	(671)	-----
Hel308 Mma	(748)	-----
Hel308 Mmah	(730)	-----
Hel308 Mmah	(730)	-----
Hel308 Mmar	(800)	-----
Hel308 Mmaz	(731)	-----
Hel308 Mok	(776)	-----
Hel308 Mzh	(694)	-----
Hel308 Mzh	(740)	-----
Hel308 Mma	(800)	-----
Hel308 Nth	(712)	-----
Hel308 Pfu	(721)	-----
Hel308 Sso	(716)	-----
Hel308 Tda	(756)	-----
Hel308 Tga	(721)	-----
Hel308 Tsi	(745)	-----
Hel308 Mja	(730)	-----
퀀센서스	(1141)	-----

도면

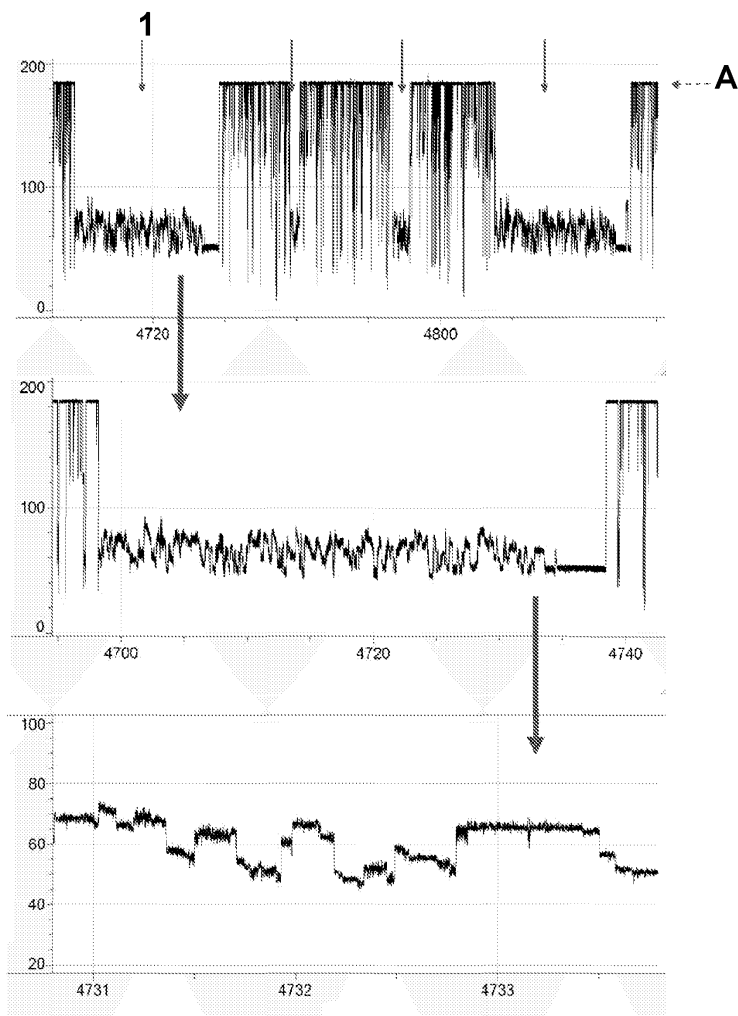
도면1a



도면1b

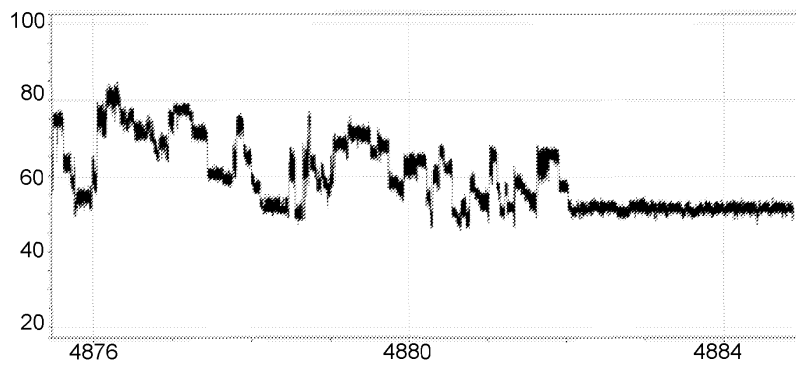
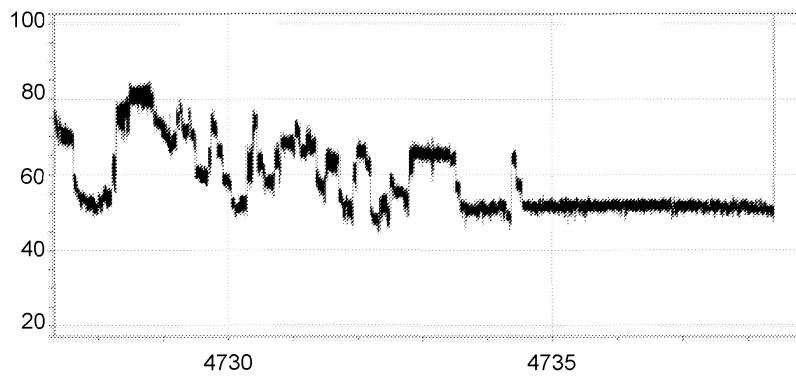


도면2

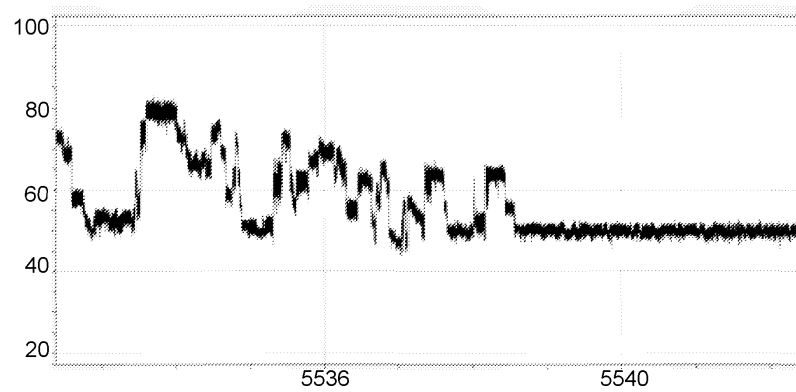
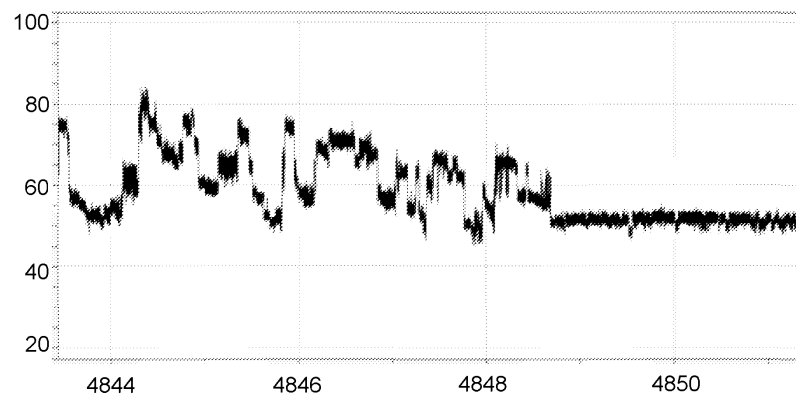




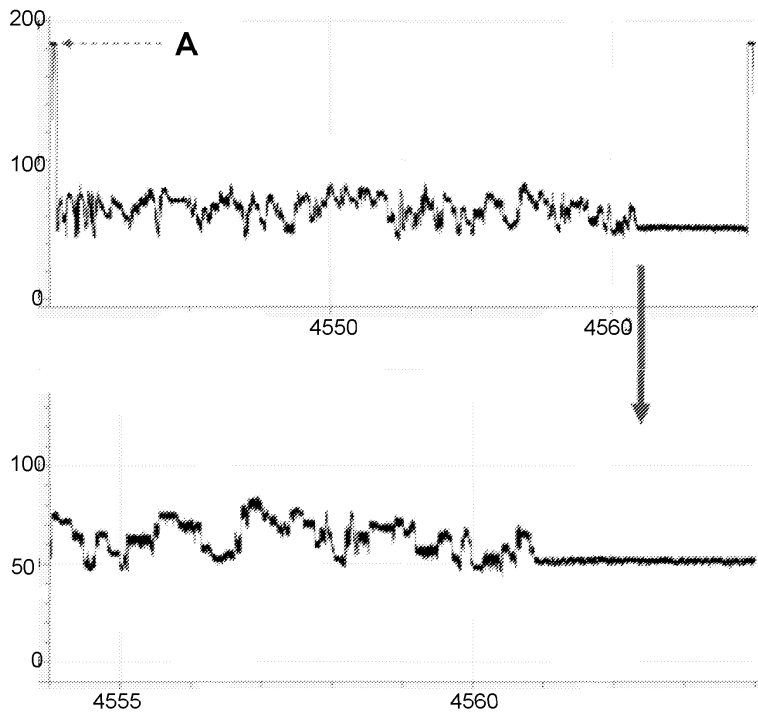
도면3a



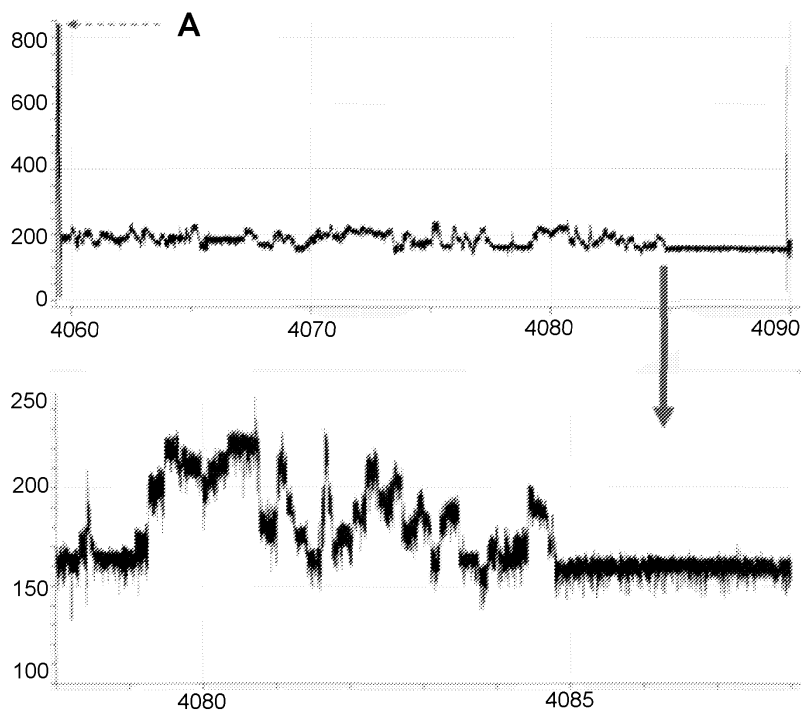
도면3b



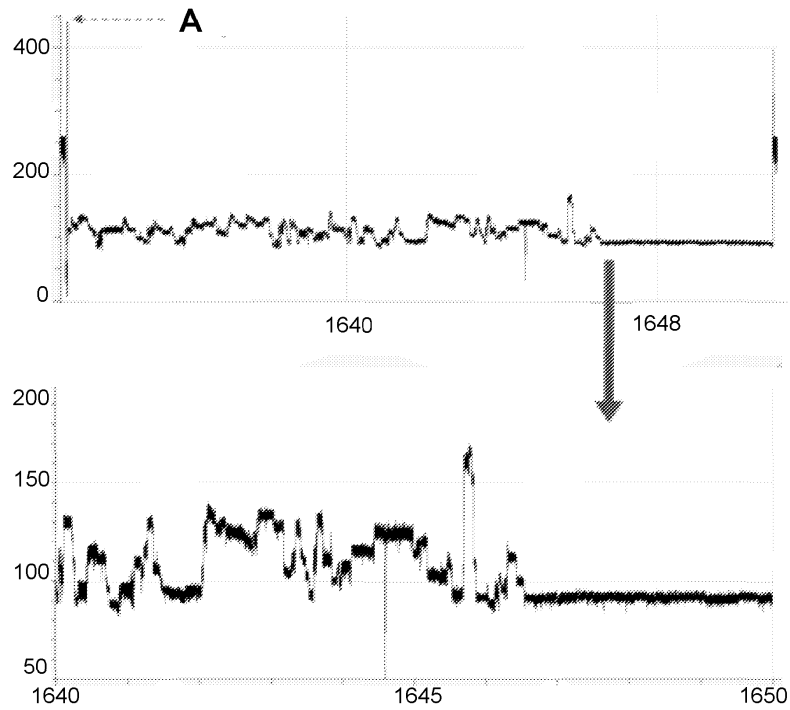
도면4a



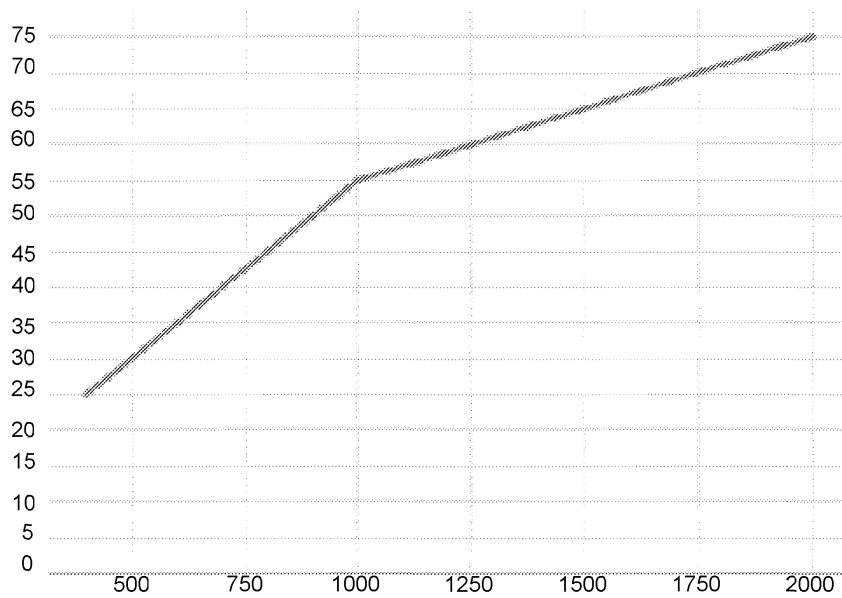
도면4b



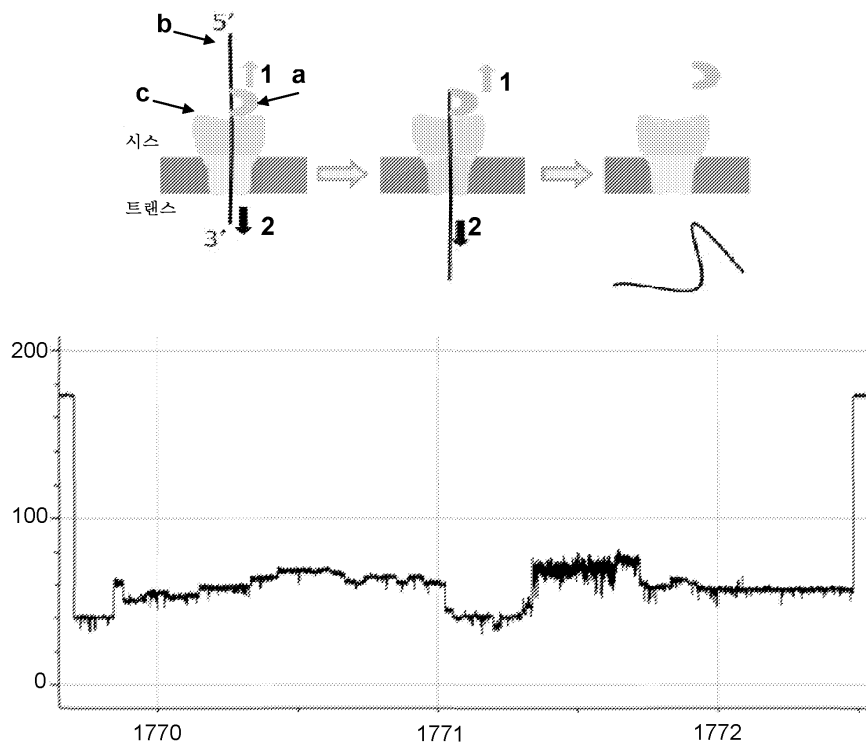
도면4c



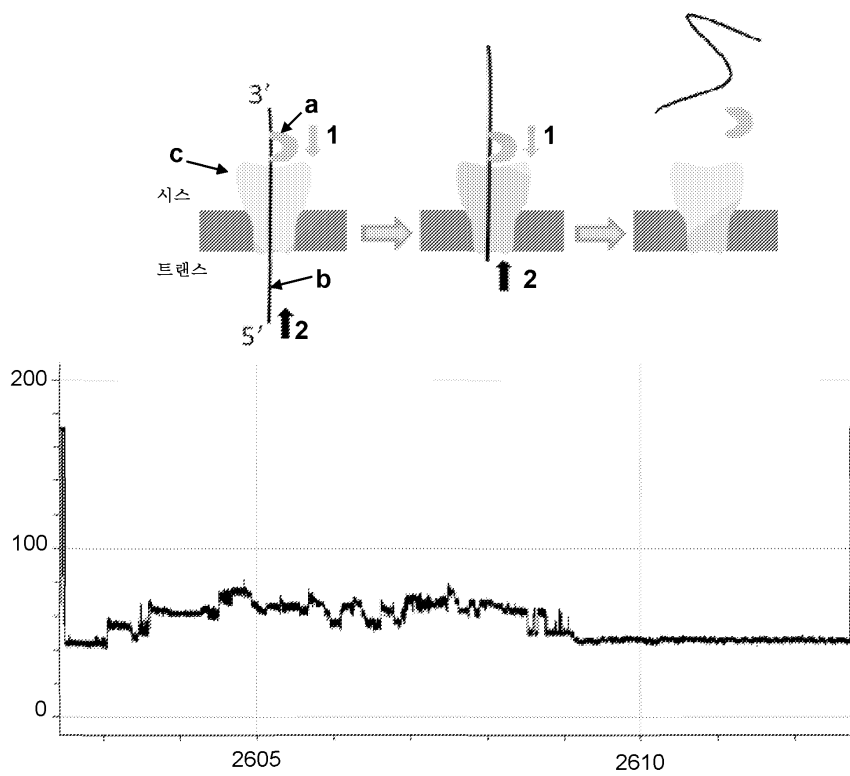
도면4d



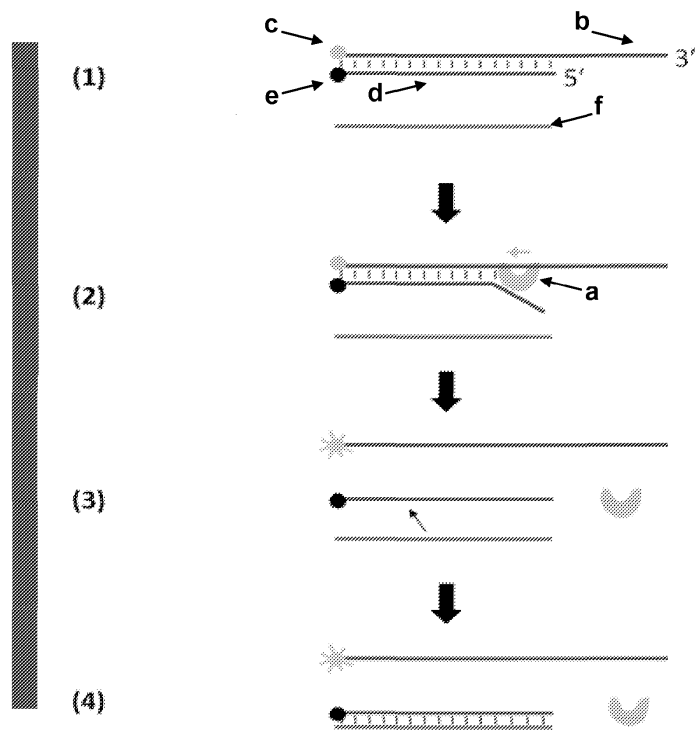
도면5a



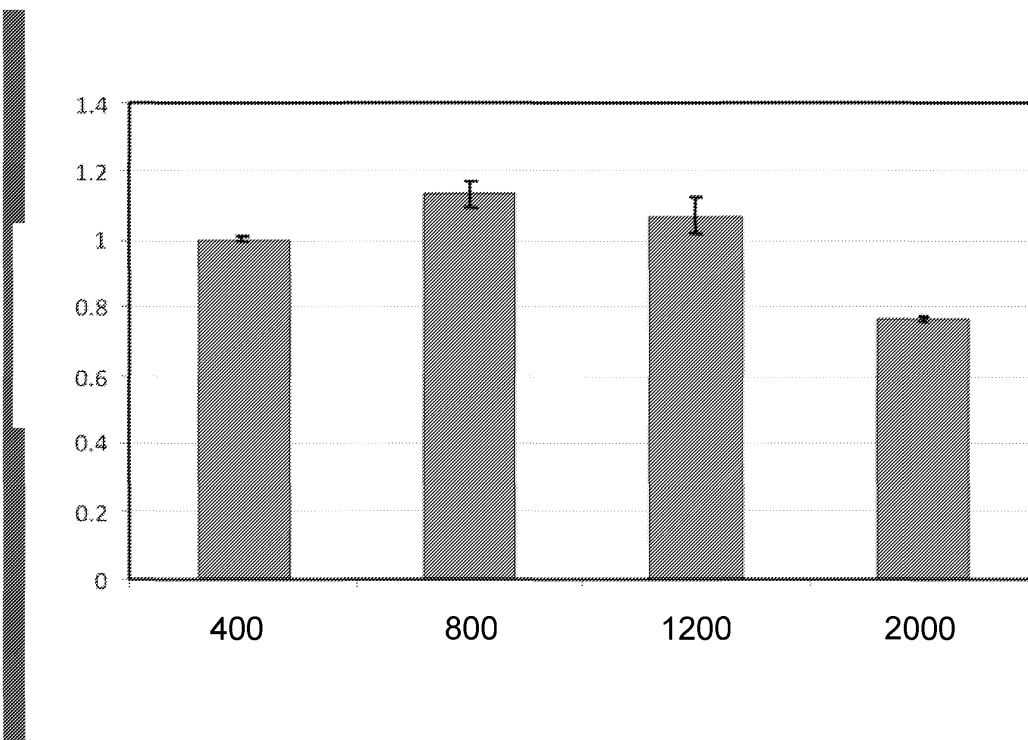
도면5b



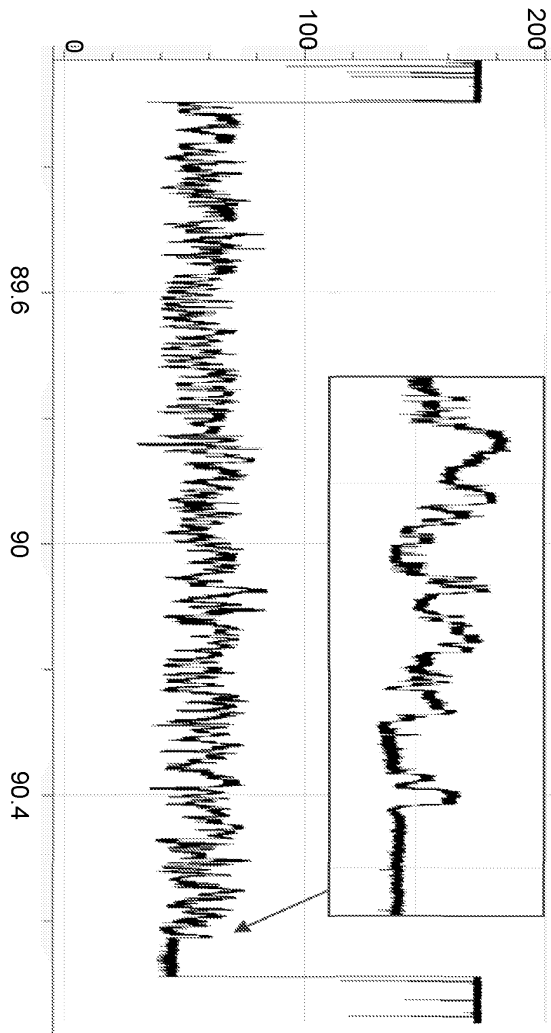
도면6a



도면6b

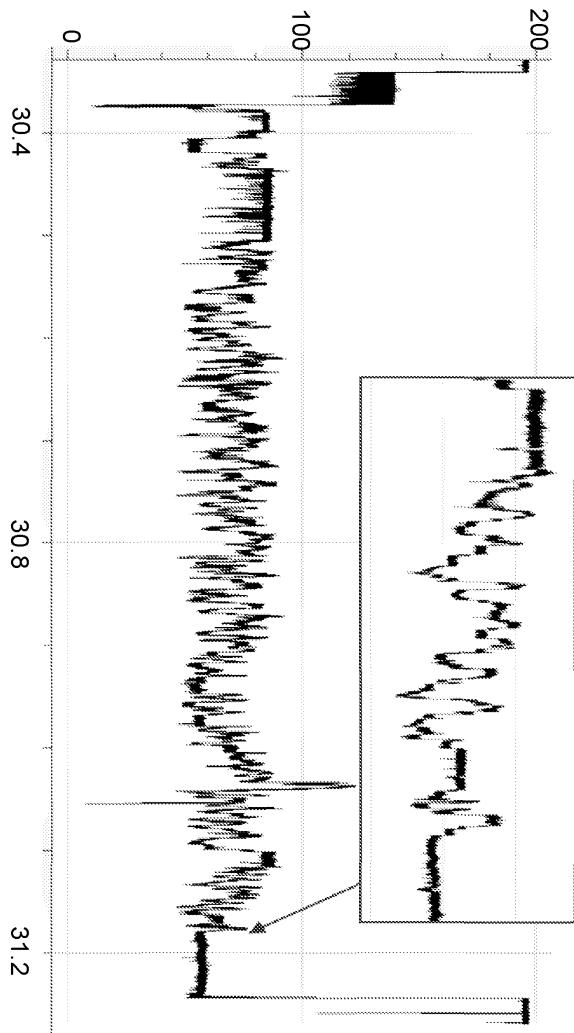


도면7a

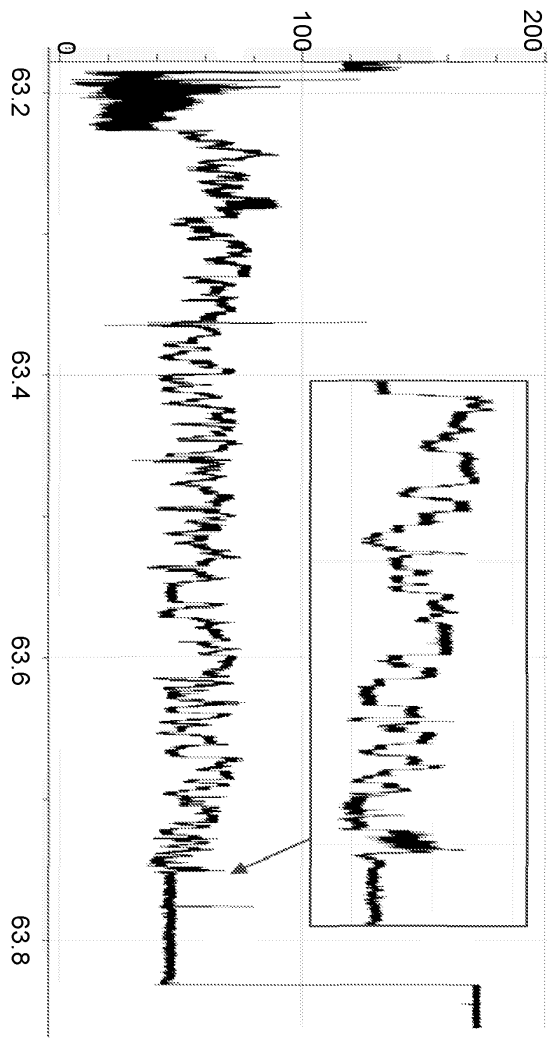




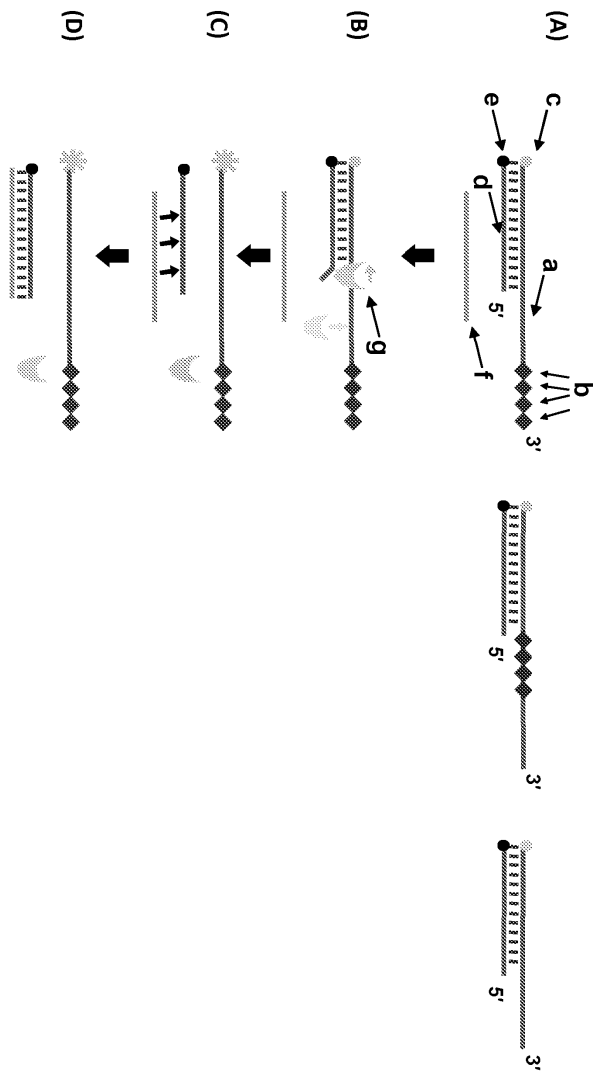
도면7b



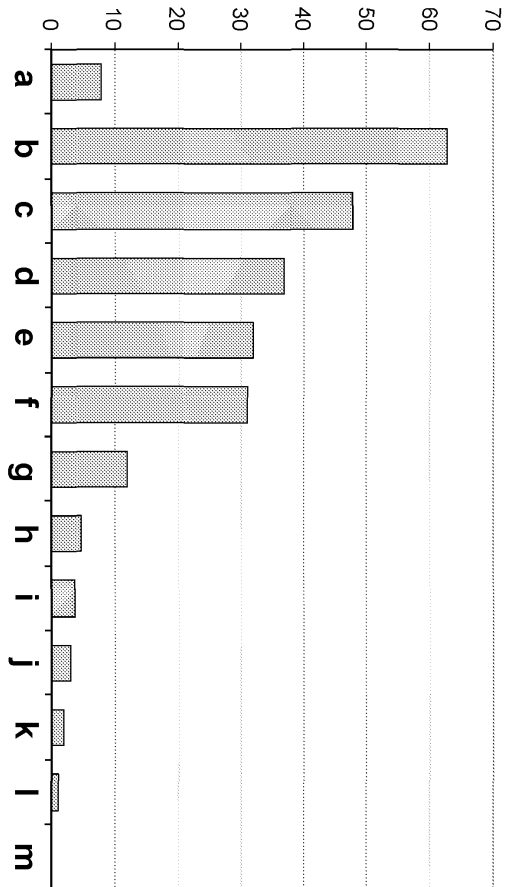
도면7c



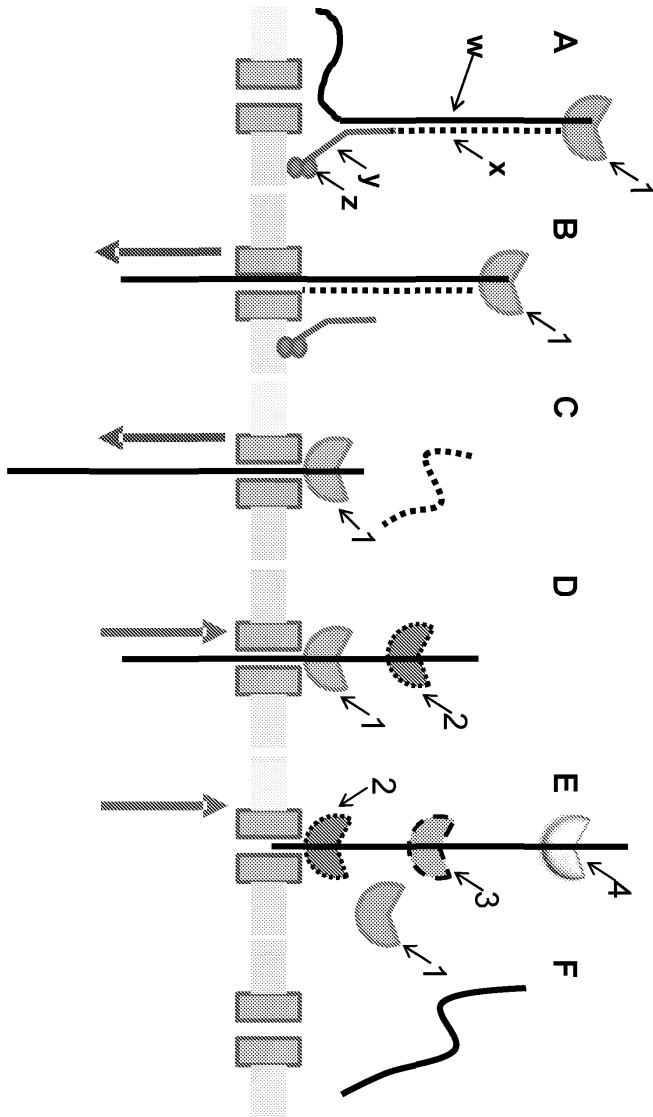
도면8



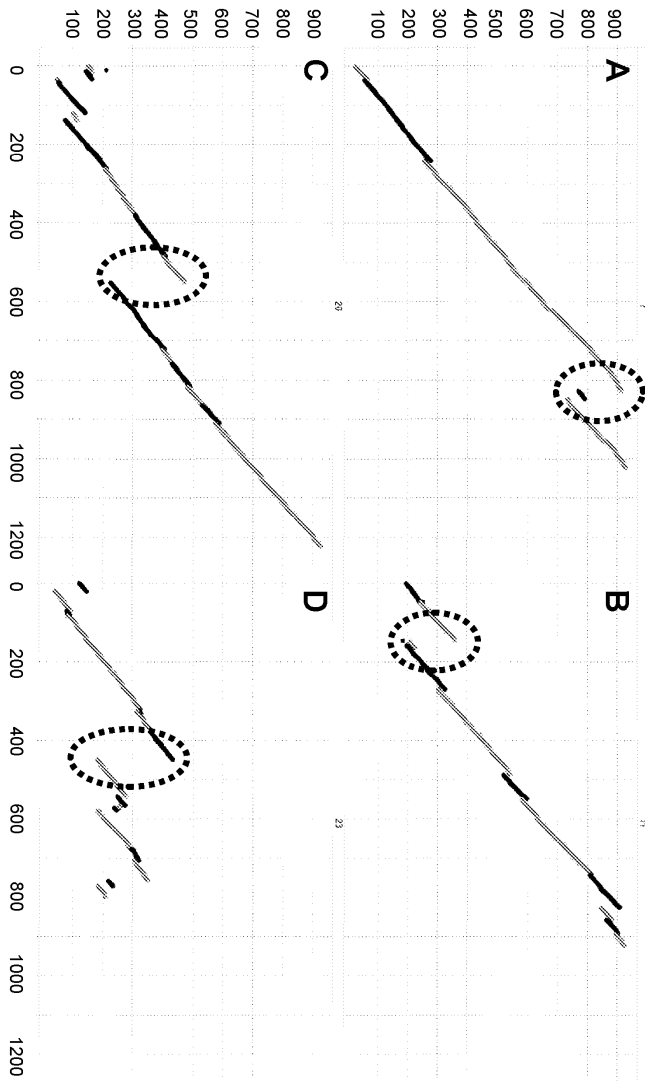
도면9



도면10

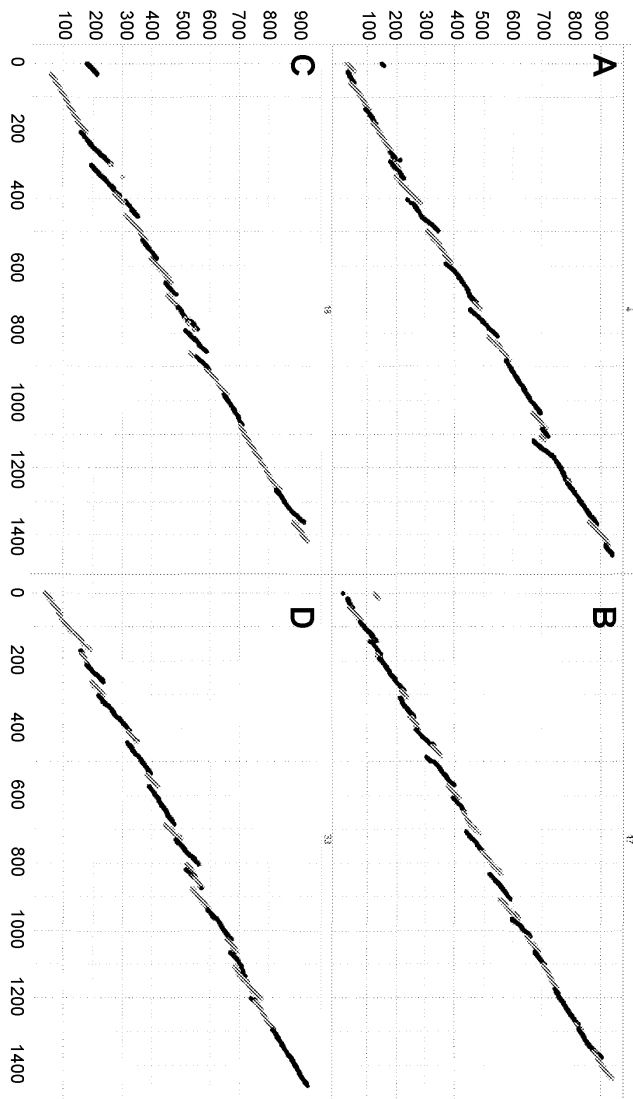


도면11

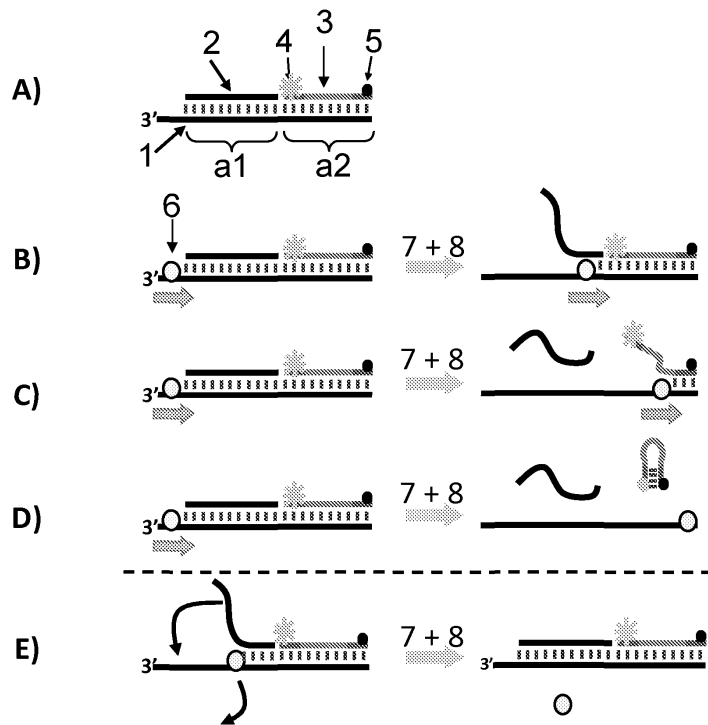




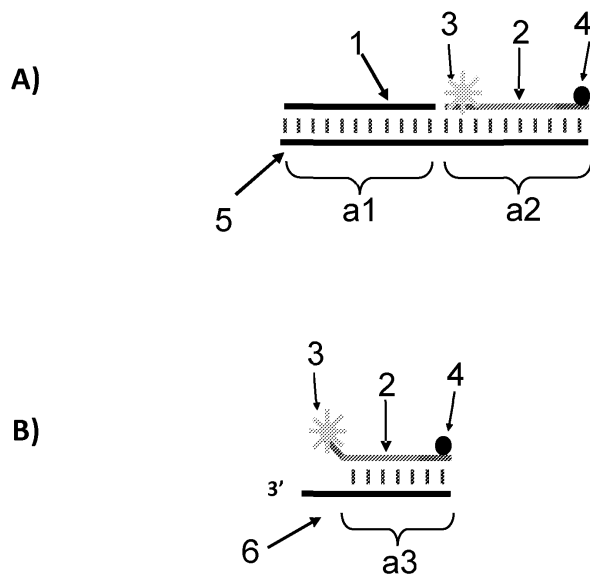
도면12



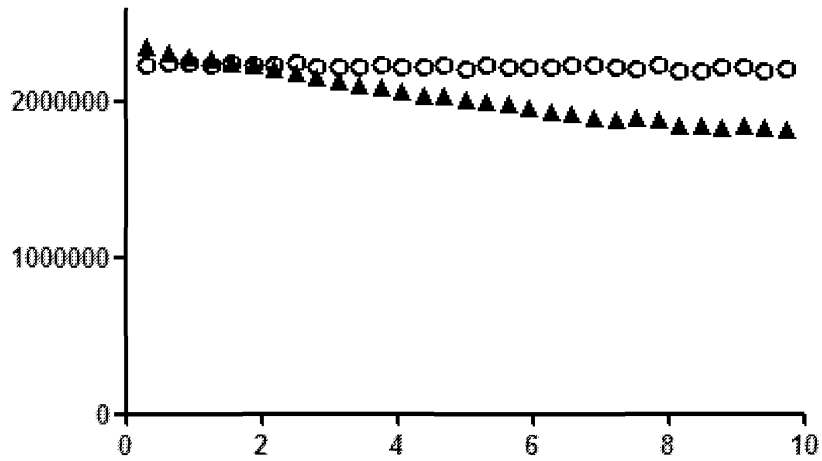
도면13



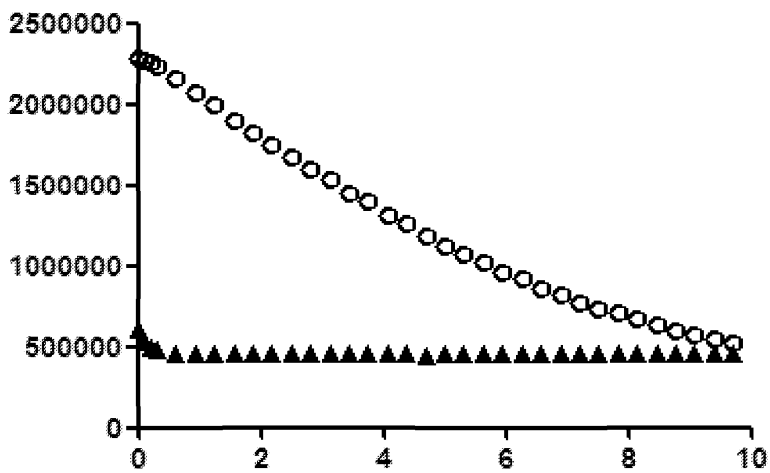
도면14



도면15



도면16



## 서열 목록

### SEQUENCE LISTING

<110> OXFORD NANOPORE TECHNOLOGIES LIMITED

<120> ENZYME METHOD

<130> N115020B

<150> US 61/549,998

<151> 2011-10-21

<150> US 61/599,244

<151> 2012-02-15

<160> 78

<170> PatentIn version 3.5

<210> 1

<211> 558

<212> DNA

<213> Artificial sequence

<220><223> Mycobacterium smegmatis porin A mutant

(D90N/D91N/D93N/D118R/D134R/E139K)

<400> 1

```
atgggtctgg ataatgaact gagcctggtg gacggtcaag atcgtaccct gacggtgcaa      60

caatgggata cctttctgaa tggcgttttt ccgctggatc gtaatcgct gacccgtgaa      120
tggtttcatt ccggtcgcgc aaaatatatc gtcgcaggcc cgggtgctga cgaattcgaa      180
ggcacgctgg aactgggtta tcagattggc ttcccggtgt cactgggcgt tggatatcaac      240
ttctctgata ccacgccgaa tattctgata aacaatggta acattaccgc accgccgttt      300
ggcctgaaca gcgtgattac gccgaacctg ttcccggtgt ttagcatctc tgcccgtctg      360
ggcaatggtc cgggcattca agaagtggca acctttagtg tgcgcgtttc cggcgctaaa      420
ggcgggtgtcg cgggtgtctaa cgcccacggt accgttacgg gcgcggccgg cgggtgtcctg      480

ctgcgtccgt tcgcgcgcct gattgcctct accggcgaca gcgttacgac ctatggcgaa      540
ccgtggaata tgaactaa                                         558
```

<210> 2

<211> 184

<212> PRT

<213> Artificial sequence

<220><223> Mycobacterium smegmatis porin A mutant

(D90N/D91N/D93N/D118R/D134R/E139K)

<400> 2

```
Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu
1           5           10           15

Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp

           20           25           30

Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr
           35           40           45

Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu
           50           55           60
```

Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe  
 65                      70                      75                      80  
 Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asn Asn Gly Asn Ile Thr Ala

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

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

<210> 3

<211> 885

<212> DNA

<213> Artificial sequence

<220><223> alpha-hemolysin mutant E111N/K147N

<400> 3

atggcagatt ctgatattaa tattaacc ggtactacag atattggaag caatactaca	60
gtaaaaacag gtgatttagt cacttatgat aaagaaatg gcatgcacaa aaaagtattt	120
tatagtttta tcgatgataa aaatcacaaat aaaaaactgc tagttattag aacaaaaggt	180
accattgctg gtcaatatag agtttatagc gaagaaggtg ctaacaaaag tggtttagcc	240
tggccttcag cctttaaggt acagttgcaa ctacctgata atgaagtagc tcaaatatct	300
gattactatc caagaaattc gattgataca aaaaactata tgagtacttt aacttatgga	360
ttcaacggta atgttactgg tgatgataca ggaaaaattg gcggccttat tggtgcaaat	420
gtttcgattg gtcataact gaactatgtt caacctgatt tcaaaacaat tttagagagc	480
ccaactgata aaaaagtagg ctggaaagtg atatttaaca atatggtgaa tcaaaattgg	540

ggaccatacg atcgagattc ttggaacccg gtatatggca atcaactttt catgaaaact 600  
 agaaatggtt ctatgaaagc agcagataac ttccttgatc ctaacaaagc aagttctcta 660  
 ttatcttcag ggttttcacc agacttcgct acagttatta ctatggatag aaaagcatcc 720  
 aaacaacaaa caaatataga tgtaatatac gaacgagttc gtgatgatta ccaattgcat 780  
 tggacttcaa caaatggaa aggtaccaat actaaagata aatggacaga tcgttcttca 840  
 gaaagatata aaatcgattg ggaaaaagaa gaaatgacaa attaa 885

<210> 4

<211> 293

<212> PRT

<213> Artificial sequence

<220><223> alpha-hemolysin mutant E111N/K147N

<400> 4

Ala Asp Ser Asp Ile Asn Ile Lys Thr Gly Thr Thr Asp Ile Gly Ser

1 5 10 15

Asn Thr Thr Val Lys Thr Gly Asp Leu Val Thr Tyr Asp Lys Glu Asn

20 25 30

Gly Met His Lys Lys Val Phe Tyr Ser Phe Ile Asp Asp Lys Asn His

35 40 45

Asn Lys Lys Leu Leu Val Ile Arg Thr Lys Gly Thr Ile Ala Gly Gln

50 55 60

Tyr Arg Val Tyr Ser Glu Glu Gly Ala Asn Lys Ser Gly Leu Ala Trp

65 70 75 80

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

85 90 95

Gln Ile Ser Asp Tyr Tyr Pro Arg Asn Ser Ile Asp Thr Lys Asn Tyr

100 105 110

Met Ser Thr Leu Thr Tyr Gly Phe Asn Gly Asn Val Thr Gly Asp Asp

115 120 125

Thr Gly Lys Ile Gly Gly Leu Ile Gly Ala Asn Val Ser Ile Gly His

130 135 140

Thr Leu Asn Tyr Val Gln Pro Asp Phe Lys Thr Ile Leu Glu Ser Pro

145 150 155 160



Thr Asp Lys Lys Val Gly Trp Lys Val Ile Phe Asn Asn Met Val Asn  
165 170 175  
Gln Asn Trp Gly Pro Tyr Asp Arg Asp Ser Trp Asn Pro Val Tyr Gly  
180 185 190  
Asn Gln Leu Phe Met Lys Thr Arg Asn Gly Ser Met Lys Ala Ala Asp  
195 200 205  
Asn Phe Leu Asp Pro Asn Lys Ala Ser Ser Leu Leu Ser Ser Gly Phe  
210 215 220  
Ser Pro Asp Phe Ala Thr Val Ile Thr Met Asp Arg Lys Ala Ser Lys  
225 230 235 240  
Gln Gln Thr Asn Ile Asp Val Ile Tyr Glu Arg Val Arg Asp Asp Tyr  
245 250 255  
Gln Leu His Trp Thr Ser Thr Asn Trp Lys Gly Thr Asn Thr Lys Asp  
260 265 270  
Lys Trp Thr Asp Arg Ser Ser Glu Arg Tyr Lys Ile Asp Trp Glu Lys  
275 280 285  
Glu Glu Met Thr Asn  
290  
<210> 5  
<211> 184  
<212> PRT  
<213> Mycobacterium smegmatis  
<400> 5  
Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu  
1 5 10 15  
Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp  
20 25 30  
Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr  
35 40 45  
Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu  
50 55 60

Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe

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

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

<210> 6

<211> 184

<212> PRT

<213> Mycobacterium smegmatis

<400> 6

Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu

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

65                      70                      75                      80

Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Gly  
85 90 95  
Pro Pro Phe Gly Leu Glu Ser Val Ile Thr Pro Asn Leu Phe Pro Gly  
100 105 110  
Val Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val  
115 120 125  
Ala Thr Phe Ser Val Asp Val Ser Gly Pro Ala Gly Gly Val Ala Val  
130 135 140  
Ser Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu  
145 150 155 160  
Arg Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr  
165 170 175  
Tyr Gly Glu Pro Trp Asn Met Asn  
180  
<210> 7  
<211> 183  
<212> PRT  
<213> *Mycobacterium smegmatis*  
<400> 7  
Val Asp Asn Gln Leu Ser Val Val Asp Gly Gln Gly Arg Thr Leu Thr  
1 5 10 15  
Val Gln Gln Ala Glu Thr Phe Leu Asn Gly Val Phe Pro Leu Asp Arg  
20 25 30  
Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Thr Tyr His  
35 40 45  
Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu Gly  
50 55 60  
Tyr Gln Val Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe Ser  
65 70 75 80  
Tyr Thr Thr Pro Asn Ile Leu Ile Asp Gly Gly Asp Ile Thr Gln Pro  
85 90 95  
Pro Phe Gly Leu Asp Thr Ile Ile Thr Pro Asn Leu Phe Pro Gly Val

100                      105                      110  
 Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val Ala  
 115                      120                      125  
 Thr Phe Ser Val Asp Val Lys Gly Ala Lys Gly Ala Val Ala Val Ser  
  
 130                      135                      140  
 Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu Arg  
 145                      150                      155                      160  
 Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr Tyr  
 165                      170                      175  
 Gly Glu Pro Trp Asn Met Asn  
 180

<210> 8  
 <211> 8  
 <212> PRT  
 <213> Artificial sequence  
 <220><223> Amino acid sequence of the Hel308 motif

<220>  
 ><221> MISC\_FEATURE  
 <222> (2)..(2)  
 <223> Xaa = C, M or L  
 <220><221> MISC\_FEATURE  
 <222> (3)..(3)  
 <223> Xaa = any amino acid  
 <400> 8

Gln Xaa Xaa Gly Arg Ala Gly Arg

1                      5  
 <210> 9  
 <211> 9  
 <212> PRT  
 <213> Artificial sequence  
 <220><223> Amino acid sequence of the extended Hel308 motif  
 <220><221> MISC\_FEATURE  
 <222> (2)..(2)  
 <223> Xaa = C, M or L

<220><221> MISC\_FEATURE

<222> (3)..(3)

<223> Xaa = any amino acid

<400> 9

Gln Xaa Xaa Gly Arg Ala Gly Arg Pro

1 5

<210> 10

<211> 760

<212> PRT

<213> Methanococcoides burtonii

<400> 10

Met Met Ile Arg Glu Leu Asp Ile Pro Arg Asp Ile Ile Gly Phe Tyr

1 5 10 15

Glu Asp Ser Gly Ile Lys Glu Leu Tyr Pro Pro Gln Ala Glu Ala Ile

20 25 30

Glu Met Gly Leu Leu Glu Lys Lys Asn Leu Leu Ala Ala Ile Pro Thr

35 40 45

Ala Ser Gly Lys Thr Leu Leu Ala Glu Leu Ala Met Ile Lys Ala Ile

50 55 60

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

65 70 75 80

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

85 90 95

Val Gly Ile Ser Thr Gly Asp Leu Asp Ser Arg Ala Asp Trp Leu Gly

100 105 110

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

115 120 125

Arg Asn Gly Thr Ser Trp Met Asp Glu Ile Thr Thr Val Val Val Asp

130 135 140

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

145 150 155 160

Thr Ile Thr Lys Leu Met Arg Leu Asn Pro Asp Val Gln Val Val Ala

					165						170						175		
Leu	Ser	Ala	Thr	Val	Gly	Asn	Ala	Arg	Glu	Met	Ala	Asp	Trp	Leu	Gly				
					180						185						190		
Ala	Ala	Leu	Val	Leu	Ser	Glu	Trp	Arg	Pro	Thr	Asp	Leu	His	Glu	Gly				
					195						200						205		
Val	Leu	Phe	Gly	Asp	Ala	Ile	Asn	Phe	Pro	Gly	Ser	Gln	Lys	Lys	Ile				
					210						215						220		
Asp	Arg	Leu	Glu	Lys	Asp	Asp	Ala	Val	Asn	Leu	Val	Leu	Asp	Thr	Ile				
225					230					235					240				
Lys	Ala	Glu	Gly	Gln	Cys	Leu	Val	Phe	Glu	Ser	Ser	Arg	Arg	Asn	Cys				
					245						250						255		
Ala	Gly	Phe	Ala	Lys	Thr	Ala	Ser	Ser	Lys	Val	Ala	Lys	Ile	Leu	Asp				
					260						265						270		
Asn	Asp	Ile	Met	Ile	Lys	Leu	Ala	Gly	Ile	Ala	Glu	Glu	Val	Glu	Ser				
					275						280						285		
Thr	Gly	Glu	Thr	Asp	Thr	Ala	Ile	Val	Leu	Ala	Asn	Cys	Ile	Arg	Lys				
290					295					300									
Gly	Val	Ala	Phe	His	His	Ala	Gly	Leu	Asn	Ser	Asn	His	Arg	Lys	Leu				
305					310					315					320				
Val	Glu	Asn	Gly	Phe	Arg	Gln	Asn	Leu	Ile	Lys	Val	Ile	Ser	Ser	Thr				
					325						330						335		
Pro	Thr	Leu	Ala	Ala	Gly	Leu	Asn	Leu	Pro	Ala	Arg	Arg	Val	Ile	Ile				
					340						345						350		
Arg	Ser	Tyr	Arg	Arg	Phe	Asp	Ser	Asn	Phe	Gly	Met	Gln	Pro	Ile	Pro				
355					360					365									
Val	Leu	Glu	Tyr	Lys	Gln	Met	Ala	Gly	Arg	Ala	Gly	Arg	Pro	His	Leu				
370					375					380									
Asp	Pro	Tyr	Gly	Glu	Ser	Val	Leu	Leu	Ala	Lys	Thr	Tyr	Asp	Glu	Phe				
385					390					395					400				
Ala	Gln	Leu	Met	Glu	Asn	Tyr	Val	Glu	Ala	Asp	Ala	Glu	Asp	Ile	Trp				
					405						410						415		



Ser Lys Leu Gly Thr Glu Asn Ala Leu Arg Thr His Val Leu Ser Thr  
420 425 430

Ile Val Asn Gly Phe Ala Ser Thr Arg Gln Glu Leu Phe Asp Phe Phe  
435 440 445

Gly Ala Thr Phe Phe Ala Tyr Gln Gln Asp Lys Trp Met Leu Glu Glu  
450 455 460

Val Ile Asn Asp Cys Leu Glu Phe Leu Ile Asp Lys Ala Met Val Ser  
465 470 475 480

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

Arg Leu Gly Ser Leu Val Ser Met Leu Tyr Ile Asp Pro Leu Ser Gly  
500 505 510

Ser Lys Ile Val Asp Gly Phe Lys Asp Ile Gly Lys Ser Thr Gly Gly  
515 520 525

Asn Met Gly Ser Leu Glu Asp Asp Lys Gly Asp Asp Ile Thr Val Thr  
530 535 540

Asp Met Thr Leu Leu His Leu Val Cys Ser Thr Pro Asp Met Arg Gln  
545 550 555 560

Leu Tyr Leu Arg Asn Thr Asp Tyr Thr Ile Val Asn Glu Tyr Ile Val  
565 570 575

Ala His Ser Asp Glu Phe His Glu Ile Pro Asp Lys Leu Lys Glu Thr  
580 585 590

Asp Tyr Glu Trp Phe Met Gly Glu Val Lys Thr Ala Met Leu Leu Glu  
595 600 605

Glu Trp Val Thr Glu Val Ser Ala Glu Asp Ile Thr Arg His Phe Asn  
610 615 620

Val Gly Glu Gly Asp Ile His Ala Leu Ala Asp Thr Ser Glu Trp Leu  
625 630 635 640

Met His Ala Ala Ala Lys Leu Ala Glu Leu Leu Gly Val Glu Tyr Ser  
645 650 655

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

660 665 670  
 Asp Leu Met Glu Leu Val Gly Ile Arg Gly Val Gly Arg Val Arg Ala  
 675 680 685

Arg Lys Leu Tyr Asn Ala Gly Phe Val Ser Val Ala Lys Leu Lys Gly  
 690 695 700  
 Ala Asp Ile Ser Val Leu Ser Lys Leu Val Gly Pro Lys Val Ala Tyr  
 705 710 715 720  
 Asn Ile Leu Ser Gly Ile Gly Val Arg Val Asn Asp Lys His Phe Asn  
 725 730 735  
 Ser Ala Pro Ile Ser Ser Asn Thr Leu Asp Thr Leu Leu Asp Lys Asn  
 740 745 750

Gln Lys Thr Phe Asn Asp Phe Gln  
 755 760

<210> 11

<211> 8

<212> PRT

<213> Artificial sequence

<220><223> Exemplary Hel308 motif

<400> 11

Gln Met Ala Gly Arg Ala Gly Arg

1 5

<210> 12

<211> 9

<212> PRT

<213> Artificial sequence

<220><223> Exemplary extended Hel308 motif

<400> 12

Gln Met Ala Gly Arg Ala Gly Arg Pro

1 5

<210> 13

<211> 720

<212> PRT

<213> Pyrococcus furiosus

<400> 13

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

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

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

Leu Ile Val Ser Asp Trp Arg Pro Val Lys Leu Arg Arg Gly Val Phe  
                   195                    200                    205  
 Tyr Gln Gly Phe Val Thr Trp Glu Asp Gly Ser Ile Asp Arg Phe Ser  
                   210                    215                    220  
 Ser Trp Glu Glu Leu Val Tyr Asp Ala Ile Arg Lys Lys Lys Gly Ala

225                      230                      235                      240  
 Leu Ile Phe Val Asn Met Arg Arg Lys Ala Glu Arg Val Ala Leu Glu  
                          245                      250                      255  
  
 Leu Ser Lys Lys Val Lys Ser Leu Leu Thr Lys Pro Glu Ile Arg Ala  
                          260                      265                      270  
 Leu Asn Glu Leu Ala Asp Ser Leu Glu Glu Asn Pro Thr Asn Glu Lys  
                          275                      280                      285  
 Leu Ala Lys Ala Ile Arg Gly Gly Val Ala Phe His His Ala Gly Leu  
                          290                      295                      300  
 Gly Arg Asp Glu Arg Val Leu Val Glu Glu Asn Phe Arg Lys Gly Ile  
 305                      310                      315                      320  
  
 Ile Lys Ala Val Val Ala Thr Pro Thr Leu Ser Ala Gly Ile Asn Thr  
                          325                      330                      335  
 Pro Ala Phe Arg Val Ile Ile Arg Asp Ile Trp Arg Tyr Ser Asp Phe  
                          340                      345                      350  
 Gly Met Glu Arg Ile Pro Ile Ile Glu Val His Gln Met Leu Gly Arg  
                          355                      360                      365  
 Ala Gly Arg Pro Lys Tyr Asp Glu Val Gly Glu Gly Ile Ile Val Ser  
                          370                      375                      380  
  
 Thr Ser Asp Asp Pro Arg Glu Val Met Asn His Tyr Ile Phe Gly Lys  
 385                      390                      395                      400  
 Pro Glu Lys Leu Phe Ser Gln Leu Ser Asn Glu Ser Asn Leu Arg Ser  
                          405                      410                      415  
 Gln Val Leu Ala Leu Ile Ala Thr Phe Gly Tyr Ser Thr Val Glu Glu  
                          420                      425                      430  
 Ile Leu Lys Phe Ile Ser Asn Thr Phe Tyr Ala Tyr Gln Arg Lys Asp  
                          435                      440                      445  
  
 Thr Tyr Ser Leu Glu Glu Lys Ile Arg Asn Ile Leu Tyr Phe Leu Leu  
                          450                      455                      460  
 Glu Asn Glu Phe Ile Glu Ile Ser Leu Glu Asp Lys Ile Arg Pro Leu  
 465                      470                      475                      480

Ser Leu Gly Ile Arg Thr Ala Lys Leu Tyr Ile Asp Pro Tyr Thr Ala  
485 490 495

Lys Met Phe Lys Asp Lys Met Glu Glu Val Val Lys Asp Pro Asn Pro  
500 505 510

Ile Gly Ile Phe His Leu Ile Ser Leu Thr Pro Asp Ile Thr Pro Phe  
515 520 525

Asn Tyr Ser Lys Arg Glu Phe Glu Arg Leu Glu Glu Glu Tyr Tyr Glu  
530 535 540

Phe Lys Asp Arg Leu Tyr Phe Asp Asp Pro Tyr Ile Ser Gly Tyr Asp  
545 550 555 560

Pro Tyr Leu Glu Arg Lys Phe Phe Arg Ala Phe Lys Thr Ala Leu Val  
565 570 575

Leu Leu Ala Trp Ile Asn Glu Val Pro Glu Gly Glu Ile Val Glu Lys  
580 585 590

Tyr Ser Val Glu Pro Gly Asp Ile Tyr Arg Ile Val Glu Thr Ala Glu  
595 600 605

Trp Leu Val Tyr Ser Leu Lys Glu Ile Ala Lys Val Leu Gly Ala Tyr  
610 615 620

Glu Ile Val Asp Tyr Leu Glu Thr Leu Arg Val Arg Val Lys Tyr Gly  
625 630 635 640

Ile Arg Glu Glu Leu Ile Pro Leu Met Gln Leu Pro Leu Val Gly Arg  
645 650 655

Arg Arg Ala Arg Ala Leu Tyr Asn Ser Gly Phe Arg Ser Ile Glu Asp  
660 665 670

Ile Ser Gln Ala Arg Pro Glu Glu Leu Leu Lys Ile Glu Gly Ile Gly  
675 680 685

Val Lys Thr Val Glu Ala Ile Phe Lys Phe Leu Gly Lys Asn Val Lys  
690 695 700

Ile Ser Glu Lys Pro Arg Lys Ser Thr Leu Asp Tyr Phe Leu Lys Ser  
705 710 715 720

<210> 14

<211> 8

<212> PRT

<213> Artificial sequence

<220><223> Exemplary Hel308 motif

<400> 14

Gln Met Leu Gly Arg Ala Gly Arg

1 5

<210> 15

<211> 9

<212> PRT

<213> Artificial sequence

<220><223> Exemplary extended Hel308 motif

<400> 15

Gln Met Leu Gly Arg Ala Gly Arg Pro

1 5

<210> 16

<211> 829

<212> PRT

<213> Haloferax volcanii

<400> 16

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

1 5 10 15

Leu Arg Asp Glu Gly Ile Glu Glu Leu Tyr Pro Pro Gln Ala Glu Ala

20 25 30

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

35 40 45

Thr Ala Ser Gly Lys Thr Leu Ile Ala Glu Leu Ala Met Leu Ser Ser

50 55 60

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

65 70 75 80

Ala Ser Glu Lys Lys Ala Glu Phe Glu Arg Trp Glu Glu Tyr Gly Ile

85 90 95

Asp Val Gly Val Ser Thr Gly Asn Tyr Glu Ser Asp Gly Glu Trp Leu



100                      105                      110  
 Ser Ser Arg Asp Ile Ile Val Ala Thr Ser Glu Lys Val Asp Ser Leu  
  
 115                      120                      125  
 Val Arg Asn Asn Ala Ala Trp Met Asp Gln Leu Thr Cys Val Val Ala  
 130                      135                      140  
 Asp Glu Val His Leu Val Asp Asp Arg His Arg Gly Pro Thr Leu Glu  
 145                      150                      155                      160  
 Val Thr Leu Ala Lys Leu Arg Arg Leu Asn Thr Asn Leu Gln Val Val  
 165                      170                      175  
 Ala Leu Ser Ala Thr Val Gly Asn Ala Gly Val Val Ser Asp Trp Leu  
  
 180                      185                      190  
 Asp Ala Glu Leu Val Lys Ser Asp Trp Arg Pro Ile Asp Leu Lys Met  
 195                      200                      205  
 Gly Val His Tyr Gly Asn Ala Val Ser Phe Ala Asp Gly Ser Gln Arg  
 210                      215                      220  
 Glu Val Pro Val Gly Arg Gly Glu Arg Gln Thr Pro Ala Leu Val Ala  
 225                      230                      235                      240  
 Asp Ala Leu Glu Gly Asp Gly Glu Gly Asp Gln Gly Ser Ser Leu Val  
  
 245                      250                      255  
 Phe Val Asn Ser Arg Arg Asn Ala Glu Ser Ala Ala Arg Arg Met Ala  
 260                      265                      270  
 Asp Val Thr Glu Arg Tyr Val Thr Gly Asp Glu Arg Ser Asp Leu Ala  
 275                      280                      285  
 Glu Leu Ala Ala Glu Ile Arg Asp Val Ser Asp Thr Glu Thr Ser Asp  
 290                      295                      300  
 Asp Leu Ala Asn Ala Val Ala Lys Gly Ala Ala Phe His His Ala Gly  
  
 305                      310                      315                      320  
 Leu Ala Ala Glu His Arg Thr Leu Val Glu Asp Ala Phe Arg Asp Arg  
 325                      330                      335  
 Leu Ile Lys Cys Ile Cys Ala Thr Pro Thr Leu Ala Ala Gly Val Asn  
 340                      345                      350

Thr Pro Ser Arg Arg Val Val Val Arg Asp Trp Gln Arg Tyr Asp Gly  
355 360 365  
Asp Tyr Gly Gly Met Lys Pro Leu Asp Val Leu Glu Val His Gln Met  
370 375 380  
Met Gly Arg Ala Gly Arg Pro Gly Leu Asp Pro Tyr Gly Glu Ala Val  
385 390 395 400  
Leu Leu Ala Lys Asp Ala Asp Ala Arg Asp Glu Leu Phe Glu Arg Tyr  
405 410 415  
Ile Trp Ala Asp Ala Glu Asp Val Arg Ser Lys Leu Ala Ala Glu Pro  
420 425 430  
Ala Leu Arg Thr His Leu Leu Ala Thr Val Ala Ser Gly Phe Ala His  
435 440 445  
Thr Arg Glu Gly Leu Leu Glu Phe Leu Asp Gln Thr Leu Tyr Ala Thr  
450 455 460  
Gln Thr Asp Asp Pro Glu Arg Leu Gly Gln Val Thr Asp Arg Val Leu  
465 470 475 480  
Asp Tyr Leu Glu Val Asn Gly Phe Val Glu Phe Glu Gly Glu Thr Ile  
485 490 495  
Gln Ala Thr Pro Val Gly His Thr Val Ser Arg Leu Tyr Leu Asp Pro  
500 505 510  
Met Ser Ala Ala Glu Ile Ile Asp Gly Leu Glu Trp Ala Ala Asp His  
515 520 525  
Arg Thr Glu Lys Leu Arg Ala Leu Ala Gly Glu Thr Pro Glu Lys Pro  
530 535 540  
Thr Arg Asp Arg Ser Glu Ser Asp Glu Ser Gly Gly Phe Gln Arg Ala  
545 550 555 560  
Ser Glu Met Val Ala Asp Asp Gly Asp Gly Gly Gly Glu Asp Gly  
565 570 575  
Val Gly Ala Asn Gly Asp Gly Asp Ser Asp Asp Ala Asp Gly Val Glu  
580 585 590  
Thr Asp Arg Thr Tyr Pro Thr Pro Leu Gly Leu Tyr His Leu Val Cys

595                      600                      605  
 Arg Thr Pro Asp Met Tyr Gln Leu Tyr Leu Lys Ser Gly Asp Arg Glu  
 610                      615                      620  
 Thr Tyr Thr Glu Leu Cys Tyr Glu Arg Glu Pro Glu Phe Leu Gly Arg  
  
 625                      630                      635                      640  
 Val Pro Ser Glu Tyr Glu Asp Val Ala Phe Glu Asp Trp Leu Ser Ala  
 645                      650                      655  
 Leu Lys Thr Ala Lys Leu Leu Glu Asp Trp Val Gly Glu Val Asp Glu  
 660                      665                      670  
 Asp Arg Ile Thr Glu Arg Tyr Gly Val Gly Pro Gly Asp Ile Arg Gly  
 675                      680                      685  
 Lys Val Glu Thr Ser Glu Trp Leu Leu Gly Ala Ala Glu Arg Leu Ala  
  
 690                      695                      700  
 Thr Glu Leu Asp Leu Asp Ser Val Tyr Ala Val Arg Glu Ala Lys Lys  
 705                      710                      715                      720  
 Arg Val Glu Tyr Gly Val Arg Glu Glu Leu Leu Asp Leu Ala Gly Val  
 725                      730                      735  
 Arg Gly Val Gly Arg Lys Arg Ala Arg Arg Leu Phe Glu Ala Gly Val  
 740                      745                      750  
 Glu Thr Arg Ala Asp Leu Arg Glu Ala Asp Lys Pro Arg Val Leu Ala  
  
 755                      760                      765  
 Ala Leu Arg Gly Arg Arg Lys Thr Ala Glu Asn Ile Leu Glu Ala Ala  
 770                      775                      780  
 Gly Arg Lys Asp Pro Ser Met Asp Ala Val Asp Glu Asp Asp Ala Pro  
 785                      790                      795                      800  
 Asp Asp Ala Val Pro Asp Asp Ala Gly Phe Glu Thr Ala Lys Glu Arg  
 805                      810                      815  
 Ala Asp Gln Gln Ala Ser Leu Gly Asp Phe Glu Gly Ser  
  
 820                      825

<210> 17

<211> 8

<212> PRT

<213> Artificial sequence

<220><223> Exemplary Hel308 motif

<400> 17

Gln Met Met Gly Arg Ala Gly Arg

1 5

<210> 18

<211> 9

<212> PRT

<213> Artificial sequence

<220><223> Exemplary extended Hel308 motif

<400> 18

Gln Met Met Gly Arg Ala Gly Arg Pro

1 5

<210> 19

<211> 824

<212> PRT

<213> Halorubrum lacusprofundi

<400>

> 19

Met Gln Pro Ser Ser Leu Ser Gly Leu Pro Ala Gly Val Gly Glu Ala

1 5 10 15

Leu Glu Ala Glu Gly Val Ala Glu Leu Tyr Pro Pro Gln Glu Ala Ala

20 25 30

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

35 40 45

Thr Ala Ser Gly Lys Thr Leu Ile Ala Glu Leu Ala Met Leu Ser Ser

50 55 60

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

65 70 75 80

Ala Ser Glu Lys Lys Thr Glu Phe Glu Arg Trp Glu Glu Phe Gly Val

85 90 95

Thr Val Gly Val Ser Thr Gly Asn Tyr Glu Ser Asp Gly Glu Trp Leu

100 105 110

Ala Thr Arg Asp Ile Ile Val Ala Thr Ser Glu Lys Val Asp Ser Leu  
115 120 125

Ile Arg Asn Gly Ala Pro Trp Ile Asp Asp Leu Thr Cys Val Val Ser  
130 135 140

Asp Glu Val His Leu Val Asp Asp Pro Asn Arg Gly Pro Thr Leu Glu  
145 150 155 160

Val Thr Leu Ala Lys Leu Arg Lys Val Asn Pro Gly Leu Gln Thr Val  
165 170 175

Ala Leu Ser Ala Thr Val Gly Asn Ala Asp Val Ile Ala Glu Trp Leu  
180 185 190

Asp Ala Glu Leu Val Glu Ser Asp Trp Arg Pro Ile Asp Leu Arg Met  
195 200 205

Gly Val His Phe Gly Asn Ala Ile Asp Phe Ala Asp Gly Ser Lys Arg  
210 215 220

Glu Val Pro Val Glu Arg Gly Glu Asp Gln Thr Ala Arg Leu Val Ala  
225 230 235 240

Asp Ala Leu Asp Thr Glu Glu Asp Gly Gln Gly Gly Ser Ser Leu Val  
245 250 255

Phe Val Asn Ser Arg Arg Asn Ala Glu Ser Ser Ala Arg Lys Leu Thr  
260 265 270

Asp Val Thr Gly Pro Arg Leu Thr Asp Asp Glu Arg Asp Gln Leu Arg  
275 280 285

Glu Leu Ala Asp Glu Ile Arg Ser Gly Ser Asp Thr Asp Thr Ala Ser  
290 295 300

Asp Leu Ala Asp Ala Val Glu Gln Gly Ser Ala Phe His His Ala Gly  
305 310 315 320

Leu Arg Ser Glu Asp Arg Ala Arg Val Glu Asp Ala Phe Arg Asp Arg  
325 330 335

Leu Ile Lys Cys Ile Ser Ala Thr Pro Thr Leu Ala Ala Gly Val Asn  
340 345 350

Thr Pro Ala Arg Arg Val Ile Val Arg Asp Trp Arg Arg Tyr Asp Gly

355                      360                      365  
 Glu Phe Gly Gly Met Lys Pro Leu Asp Val Leu Glu Val His Gln Met  
 370                      375                      380  
  
 Cys Gly Arg Ala Gly Arg Pro Gly Leu Asp Pro Tyr Gly Glu Ala Val  
 385                      390                      395                      400  
 Leu Leu Ala Asn Asp Ala Asp Thr Lys Glu Glu Leu Phe Glu Arg Tyr  
 405                      410                      415  
 Leu Trp Ala Asp Pro Glu Pro Val Arg Ser Lys Leu Ala Ala Glu Pro  
 420                      425                      430  
 Ala Leu Arg Thr His Val Leu Ala Thr Val Ala Ser Gly Phe Ala Ser  
 435                      440                      445  
  
 Thr Arg Asp Gly Leu Leu Ser Phe Leu Asp Asn Thr Leu Tyr Ala Thr  
 450                      455                      460  
 Gln Thr Asp Asp Glu Gly Arg Leu Ala Ala Val Thr Asp Thr Val Leu  
 465                      470                      475                      480  
 Asp Tyr Leu Ala Val Asn Asp Phe Ile Glu Arg Asp Arg Asp Gly Gly  
 485                      490                      495  
 Ser Glu Ser Leu Thr Ala Thr Gly Ile Gly His Thr Val Ser Arg Leu  
 500                      505                      510  
  
 Tyr Leu Asp Pro Met Ser Ala Ala Glu Met Ile Asp Gly Leu Arg Ser  
 515                      520                      525  
 Val Ala Arg Asp Ala Ala Asp Thr Gly Ala Ser Ala Glu Ala Asp Asn  
 530                      535                      540  
 Gly Glu Phe Val Arg Thr Gly Asp Ala Asp Asp Ala Ser Gly Gly Asp  
 545                      550                      555                      560  
 Glu Pro Gly Phe Gly Thr Tyr Thr Arg Ala Gly Asp Asp Glu Ser Gly  
 565                      570                      575  
  
 Glu Arg Glu Thr Glu Asn Glu Glu Thr Asp Glu Glu Glu Thr Glu Ala  
 580                      585                      590  
 Ser Glu Val Thr Pro Leu Gly Leu Tyr His Leu Ile Ser Arg Thr Pro  
 595                      600                      605

Asp Met Tyr Glu Leu Tyr Leu Lys Ser Gly Asp Arg Glu Thr Tyr Thr  
 610 615 620  
 Glu Leu Cys Tyr Glu Arg Glu Thr Glu Phe Leu Gly Asp Val Pro Ser  
 625 630 635 640  
  
 Glu Tyr Glu Asp Val Arg Phe Glu Asp Trp Leu Ala Ser Leu Lys Thr  
 645 650 655  
 Ala Arg Leu Leu Glu Asp Trp Val Asn Glu Val Asp Glu Asp Arg Ile  
 660 665 670  
 Thr Glu Arg Tyr Gly Val Gly Pro Gly Asp Ile Arg Gly Lys Val Asp  
 675 680 685  
 Thr Ala Glu Trp Leu Leu Arg Ala Ala Glu Thr Leu Ala Arg Asp Val  
 690 695 700  
  
 Glu Gly Val Asp Gly Asp Val Val Val Ala Val Arg Glu Ala Arg Lys  
 705 710 715 720  
 Arg Ile Glu Tyr Gly Val Arg Glu Glu Leu Leu Asp Leu Ala Gly Val  
 725 730 735  
 Arg Asn Val Gly Arg Lys Arg Ala Arg Arg Leu Phe Glu Ala Gly Ile  
 740 745 750  
 Glu Thr Arg Ala Asp Leu Arg Glu Ala Asp Lys Ala Val Val Leu Gly  
 755 760 765  
  
 Ala Leu Arg Gly Arg Glu Arg Thr Ala Glu Arg Ile Leu Glu His Ala  
 770 775 780  
 Gly Arg Glu Asp Pro Ser Met Asp Asp Val Arg Pro Asp Lys Ser Ala  
 785 790 795 800  
 Ser Ala Ala Ala Thr Ala Gly Ser Ala Ser Asp Glu Asp Gly Glu Gly  
 805 810 815  
 Gln Ala Ser Leu Gly Asp Phe Arg  
 820  
 <210> 20  
 <211> 8  
 <212> PRT



<213> Artificial sequence

<220><223> Exemplary Hel308 motif

<400> 20

Gln Met Cys Gly Arg Ala Gly Arg

1 5

<210> 21

<211> 9

<212> PRT

<213> Artificial sequence

<220><223> Exemplary extended Hel308 motif

<400> 21

Gln Met Cys Gly Arg Ala Gly Arg Pro

1 5

<210> 22

<211> 707

<212> PRT

<213> Cenarchaeum symbiosum

<400> 22

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

1 5 10 15

Glu Gly Glu Gly Tyr Lys Lys Leu Tyr Pro Pro Gln Ala Ala Ala Ala

20 25 30

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

35 40 45

Ala Ser Gly Lys Thr Leu Ile Ala Ala Ile Ala Met Ile Ser His Leu

50 55 60

Ser Arg Asn Arg Gly Lys Ala Val Tyr Leu Ser Pro Leu Arg Ala Leu

65 70 75 80

Ala Ala Glu Lys Phe Ala Glu Phe Gly Lys Ile Gly Gly Ile Pro Leu

85 90 95

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

100 105 110

Gly Arg Ser Leu Gly Asn Asn Asp Ile Leu Val Leu Thr Asn Glu Arg  
115 120 125

Met Asp Ser Leu Ile Arg Arg Arg Pro Asp Trp Met Asp Glu Val Gly  
130 135 140

Leu Val Ile Ala Asp Glu Ile His Leu Ile Gly Asp Arg Ser Arg Gly  
145 150 155 160

Pro Thr Leu Glu Met Val Leu Thr Lys Leu Arg Gly Leu Arg Ser Ser  
165 170 175

Pro Gln Val Val Ala Leu Ser Ala Thr Ile Ser Asn Ala Asp Glu Ile  
180 185 190

Ala Gly Trp Leu Asp Cys Thr Leu Val His Ser Thr Trp Arg Pro Val  
195 200 205

Pro Leu Ser Glu Gly Val Tyr Gln Asp Gly Glu Val Ala Met Gly Asp  
210 215 220

Gly Ser Arg His Glu Val Ala Ala Thr Gly Gly Gly Pro Ala Val Asp  
225 230 235 240

Leu Ala Ala Glu Ser Val Ala Glu Gly Gly Gln Ser Leu Ile Phe Ala  
245 250 255

Asp Thr Arg Ala Arg Ser Ala Ser Leu Ala Ala Lys Ala Ser Ala Val  
260 265 270

Ile Pro Glu Ala Lys Gly Ala Asp Ala Ala Lys Leu Ala Ala Ala Ala  
275 280 285

Lys Lys Ile Ile Ser Ser Gly Gly Glu Thr Lys Leu Ala Lys Thr Leu  
290 295 300

Ala Glu Leu Val Glu Lys Gly Ala Ala Phe His His Ala Gly Leu Asn  
305 310 315 320

Gln Asp Cys Arg Ser Val Val Glu Glu Glu Phe Arg Ser Gly Arg Ile  
325 330 335

Arg Leu Leu Ala Ser Thr Pro Thr Leu Ala Ala Gly Val Asn Leu Pro  
340 345 350

Ala Arg Arg Val Val Ile Ser Ser Val Met Arg Tyr Asn Ser Ser Ser

355	360	365
Gly Met Ser Glu Pro Ile Ser Ile Leu Glu Tyr Lys Gln Leu Cys Gly		
370	375	380
Arg Ala Gly Arg Pro Gln Tyr Asp Lys Ser Gly Glu Ala Ile Val Val		
385	390	395
		400
Gly Gly Val Asn Ala Asp Glu Ile Phe Asp Arg Tyr Ile Gly Gly Glu		
405	410	415
Pro Glu Pro Ile Arg Ser Ala Met Val Asp Asp Arg Ala Leu Arg Ile		
420	425	430
His Val Leu Ser Leu Val Thr Thr Ser Pro Gly Ile Lys Glu Asp Asp		
435	440	445
Val Thr Glu Phe Phe Leu Gly Thr Leu Gly Gly Gln Gln Ser Gly Glu		
450	455	460
Ser Thr Val Lys Phe Ser Val Ala Val Ala Leu Arg Phe Leu Gln Glu		
465	470	475
		480
Glu Gly Met Leu Gly Arg Arg Gly Gly Arg Leu Ala Ala Thr Lys Met		
485	490	495
Gly Arg Leu Val Ser Arg Leu Tyr Met Asp Pro Met Thr Ala Val Thr		
500	505	510
Leu Arg Asp Ala Val Gly Glu Ala Ser Pro Gly Arg Met His Thr Leu		
515	520	525
Gly Phe Leu His Leu Val Ser Glu Cys Ser Glu Phe Met Pro Arg Phe		
530	535	540
Ala Leu Arg Gln Lys Asp His Glu Val Ala Glu Met Met Leu Glu Ala		
545	550	555
		560
Gly Arg Gly Glu Leu Leu Arg Pro Val Tyr Ser Tyr Glu Cys Gly Arg		
565	570	575
Gly Leu Leu Ala Leu His Arg Trp Ile Gly Glu Ser Pro Glu Ala Lys		
580	585	590
Leu Ala Glu Asp Leu Lys Phe Glu Ser Gly Asp Val His Arg Met Val		
595	600	605

Glu Ser Ser Gly Trp Leu Leu Arg Cys Ile Trp Glu Ile Ser Lys His  
 610 615 620  
 Gln Glu Arg Pro Asp Leu Leu Gly Glu Leu Asp Val Leu Arg Ser Arg  
 625 630 635 640  
 Val Ala Tyr Gly Ile Lys Ala Glu Leu Val Pro Leu Val Ser Ile Lys  
 645 650 655

Gly Ile Gly Arg Val Arg Ser Arg Arg Leu Phe Arg Gly Gly Ile Lys  
 660 665 670  
 Gly Pro Gly Asp Leu Ala Ala Val Pro Val Glu Arg Leu Ser Arg Val  
 675 680 685  
 Glu Gly Ile Gly Ala Thr Leu Ala Asn Asn Ile Lys Ser Gln Leu Arg  
 690 695 700

Lys Gly Gly

705

<210> 23

<211> 8

<212> PRT

<213> Artificial sequence

<220><223> Exemplary Hel308 motif

<400> 23

Gln Leu Cys Gly Arg Ala Gly Arg

1 5

<210> 24

<211> 9

<212> PRT

<213> Artificial sequence

<220><223> Exemplary extended Hel308 motif

<400> 24

Gln Leu Cys Gly Arg Ala Gly Arg Pro

1 5

<210> 25

<211> 715

<212> PRT

<213> Sulfolobus solfataricus

<400> 25

Met Ser Leu Glu Leu Glu Trp Met Pro Ile Glu Asp Leu Lys Leu Pro

1 5 10 15

Ser Asn Val Ile Glu Ile Ile Lys Lys Arg Gly Ile Lys Lys Leu Asn

20 25 30

Pro Pro Gln Thr Glu Ala Val Lys Lys Gly Leu Leu Glu Gly Asn Arg

35 40 45

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

50 55 60

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

65 70 75 80

Val Thr Pro Leu Arg Ala Leu Thr Asn Glu Lys Tyr Leu Thr Phe Lys

85 90 95

Asp Trp Glu Leu Ile Gly Phe Lys Val Ala Met Thr Ser Gly Asp Tyr

100 105 110

Asp Thr Asp Asp Ala Trp Leu Lys Asn Tyr Asp Ile Ile Ile Thr Thr

115 120 125

Tyr Glu Lys Leu Asp Ser Leu Trp Arg His Arg Pro Glu Trp Leu Asn

130 135 140

Glu Val Asn Tyr Phe Val Leu Asp Glu Leu His Tyr Leu Asn Asp Pro

145 150 155 160

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

165 170 175

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

180 185 190

Lys Trp Leu Gly Ala Glu Pro Val Ala Thr Asn Trp Arg Pro Val Pro

195 200 205

Leu Ile Glu Gly Val Ile Tyr Pro Glu Arg Lys Lys Lys Glu Tyr Asn

210 215 220

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

225                      230                      235                      240  
 Ile Ile Ala Tyr Thr Leu Asp Ser Leu Ser Lys Asn Gly Gln Val Leu  
                          245                      250                      255  
 Val Phe Arg Asn Ser Arg Lys Met Ala Glu Ser Thr Ala Leu Lys Ile  
                          260                      265                      270  
 Ala Asn Tyr Met Asn Phe Val Ser Leu Asp Glu Asn Ala Leu Ser Glu  
  
                          275                      280                      285  
 Ile Leu Lys Gln Leu Asp Asp Ile Glu Glu Gly Gly Ser Asp Glu Lys  
                          290                      295                      300  
 Glu Leu Leu Lys Ser Leu Ile Ser Lys Gly Val Ala Tyr His His Ala  
 305                      310                      315                      320  
 Gly Leu Ser Lys Ala Leu Arg Asp Leu Ile Glu Glu Gly Phe Arg Gln  
                          325                      330                      335  
 Arg Lys Ile Lys Val Ile Val Ala Thr Pro Thr Leu Ala Ala Gly Val  
  
                          340                      345                      350  
 Asn Leu Pro Ala Arg Thr Val Ile Ile Gly Asp Ile Tyr Arg Phe Asn  
                          355                      360                      365  
 Lys Lys Ile Ala Gly Tyr Tyr Asp Glu Ile Pro Ile Met Glu Tyr Lys  
                          370                      375                      380  
 Gln Met Ser Gly Arg Ala Gly Arg Pro Gly Phe Asp Gln Ile Gly Glu  
 385                      390                      395                      400  
 Ser Ile Val Val Val Arg Asp Lys Glu Asp Val Asp Arg Val Phe Lys  
  
                          405                      410                      415  
 Lys Tyr Val Leu Ser Asp Val Glu Pro Ile Glu Ser Lys Leu Gly Ser  
                          420                      425                      430  
 Glu Arg Ala Phe Tyr Thr Phe Leu Leu Gly Ile Leu Ser Ala Glu Gly  
                          435                      440                      445  
 Asn Leu Ser Glu Lys Gln Leu Glu Asn Phe Ala Tyr Glu Ser Leu Leu  
                          450                      455                      460  
 Ala Lys Gln Leu Val Asp Val Tyr Phe Asp Arg Ala Ile Arg Trp Leu  
  
 465                      470                      475                      480

Leu Glu His Ser Phe Ile Lys Glu Glu Gly Asn Thr Phe Ala Leu Thr  
 485 490 495  
 Asn Phe Gly Lys Arg Val Ala Asp Leu Tyr Ile Asn Pro Phe Thr Ala  
 500 505 510  
 Asp Ile Ile Arg Lys Gly Leu Glu Gly His Lys Ala Ser Cys Glu Leu  
 515 520 525  
 Ala Tyr Leu His Leu Leu Ala Phe Thr Pro Asp Gly Pro Leu Val Ser  
 530 535 540  
 Val Gly Arg Asn Glu Glu Glu Glu Leu Ile Glu Leu Leu Glu Asp Leu  
 545 550 555 560  
 Asp Cys Glu Leu Leu Ile Glu Glu Pro Tyr Glu Glu Asp Glu Tyr Ser  
 565 570 575  
 Leu Tyr Ile Asn Ala Leu Lys Val Ala Leu Ile Met Lys Asp Trp Met  
 580 585 590  
 Asp Glu Val Asp Glu Asp Thr Ile Leu Ser Lys Tyr Asn Ile Gly Ser  
 595 600 605  
 Gly Asp Leu Arg Asn Met Val Glu Thr Met Asp Trp Leu Thr Tyr Ser  
 610 615 620  
 Ala Tyr His Leu Ser Arg Glu Leu Lys Leu Asn Glu His Ala Asp Lys  
 625 630 635 640  
 Leu Arg Ile Leu Asn Leu Arg Val Arg Asp Gly Ile Lys Glu Glu Leu  
 645 650 655  
 Leu Glu Leu Val Gln Ile Ser Gly Val Gly Arg Lys Arg Ala Arg Leu  
 660 665 670  
 Leu Tyr Asn Asn Gly Ile Lys Glu Leu Gly Asp Val Val Met Asn Pro  
 675 680 685  
 Asp Lys Val Lys Asn Leu Leu Gly Gln Lys Leu Gly Glu Lys Val Val  
 690 695 700  
 Gln Glu Ala Ala Arg Leu Leu Asn Arg Phe His  
 705 710 715

<210> 26

<211> 8

<212> PRT

<213> Artificial sequence

<220><223> Exemplary Hel308 motif

<400> 26

Gln Met Ser Gly Arg Ala Gly Arg

1 5

<210> 27

<211> 9

<212> PRT

<213> Artificial sequence

<220><223> Exemplary extended Hel308 motif

<400> 27

Gln Met Ser Gly Arg Ala Gly Arg Pro

1 5

<210> 28

<211> 685

<212> PRT

<213> Methanogenium frigidum

<400> 28

Met Asp Leu Ser Leu Pro Lys Ala Phe Ile Gln Tyr Tyr Lys Asp Lys

1 5 10 15

Gly Ile Glu Ser Leu Tyr Pro Pro Gln Ser Glu Cys Ile Glu Asn Gly

20 25 30

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

35 40 45

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

50 55 60

Gly Met Cys Leu Tyr Ile Val Pro Leu Lys Ala Leu Ala Thr Glu Lys

65 70 75 80

Ala Gln Glu Phe Lys Gly Lys Gly Ala Glu Ile Gly Val Ala Thr Gly

85 90 95

Asp Tyr Asp Gln Lys Glu Lys Arg Leu Gly Ser Asn Asp Ile Val Ile



100	105	110	
Ala Thr Ser Glu Lys Val Asp Ser Leu Leu Arg Asn Gly Val Pro Trp			
115	120	125	
Leu Ser Gln Val Thr Cys Leu Val Val Asp Glu Val His Leu Ile Asp			
130	135	140	
Asp Glu Ser Arg Gly Pro Thr Leu Glu Met Val Ile Thr Lys Leu Arg			
145	150	155	160
His Ala Ser Pro Asp Met Gln Val Ile Gly Leu Ser Ala Thr Ile Gly			
165	170	175	
Asn Pro Lys Glu Leu Ala Gly Trp Leu Gly Ala Asp Leu Ile Thr Ser			
180	185	190	
Asp Trp Arg Pro Val Asp Leu Arg Glu Gly Ile Cys Tyr His Asn Thr			
195	200	205	
Ile Tyr Phe Asp Asn Glu Asp Lys Glu Ile Pro Ala Pro Ala Lys Thr			
210	215	220	
Glu Asp Ile Asn Leu Leu Leu Asp Cys Val Ala Asp Gly Gly Gln Cys			
225	230	235	240
Leu Val Phe Val Ser Ser Arg Arg Asn Ala Glu Gly Tyr Ala Lys Arg			
245	250	255	
Ala Ala Thr Ala Leu Lys Cys Ser His Ala Ala Leu Asp Ser Ile Ala			
260	265	270	
Glu Lys Leu Glu Ala Ala Ala Glu Thr Asp Met Gly Arg Val Leu Ala			
275	280	285	
Thr Cys Val Lys Lys Gly Ala Ala Phe His His Ala Gly Met Asn Arg			
290	295	300	
Met Gln Arg Thr Leu Val Glu Gly Gly Phe Arg Asp Gly Phe Ile Lys			
305	310	315	320
Ser Ile Ser Ser Thr Pro Thr Leu Ala Ala Gly Leu Asn Leu Pro Ala			
325	330	335	
Arg Arg Val Ile Ile Arg Asp Tyr Leu Arg Tyr Ser Gly Gly Glu Gly			
340	345	350	

Met Arg Pro Ile Pro Val Arg Glu Tyr Arg Gln Met Ala Gly Arg Ala  
355 360 365

Gly Arg Pro His Leu Asp Pro Tyr Gly Glu Ala Ile Leu Ile Ala Lys  
370 375 380

Thr Glu Tyr Ala Val Asn Asp Leu His Glu Glu Tyr Val Glu Ala Pro  
385 390 395 400

Asp Glu Asp Val Thr Ser Arg Cys Gly Glu Lys Gly Val Leu Thr Ala  
405 410 415

His Ile Leu Ser Leu Ile Ala Thr Gly Tyr Ala Arg Ser Tyr Asp Glu  
420 425 430

Leu Met Ala Phe Leu Glu Lys Thr Leu Tyr Ala Tyr Gln His Thr Gly  
435 440 445

Lys Lys Ala Leu Thr Arg Thr Leu Asp Asp Ala Leu Gly Phe Leu Thr  
450 455 460

Glu Ala Glu Met Val Thr Asp Leu Ser Gly Met Leu His Ala Thr Glu  
465 470 475 480

Tyr Gly Asp Leu Thr Ser Arg Leu Tyr Ile Asp Pro His Ser Ala Glu  
485 490 495

Ile Ile Thr Thr Ala Leu Arg Glu Glu Gly Glu Leu Thr Asp Leu Ala  
500 505 510

Leu Leu Gln Leu Leu Cys Met Thr Pro Asp Met Phe Thr Leu Tyr Val  
515 520 525

Lys Lys Asn Asp Leu Gly Thr Leu Glu Lys Phe Phe Phe Glu His Glu  
530 535 540

Glu Glu Phe Arg Thr Glu Phe Ser Tyr Asp Glu Met Glu Asp Phe Phe  
545 550 555 560

Arg Ser Leu Lys Thr Ala Met Leu Leu Ser Asp Trp Thr Asp Glu Ile  
565 570 575

Gly Asp Asp Thr Ile Cys Thr Arg Phe Gly Val Gly Pro Gly Asp Ile  
580 585 590

Phe Asn Ala Val Gln Gly Ile Ser Trp Leu Leu His Ala Ser Gly Arg

595 600 605  
 Leu Ala Arg Leu Val Ala Pro Glu His Arg Asp Ala Val Glu Glu Thr  
 610 615 620  
 Thr Leu Arg Val Arg His Gly Ile Arg Arg Glu Leu Ile Pro Leu Val  
 625 630 635 640  
 Arg Val Lys Gly Ile Gly Arg Val Arg Ala Arg Arg Leu Phe Asn Asn  
 645 650 655  
 Gly Ile Thr Gly Pro Glu Leu Leu Ala Ala Ala Asp Pro Ser Val Val

660 665 670  
 Gly His Ile Val Gly Gly Lys Thr Ala Glu Ser Ile Ile  
 675 680 685

<210> 29

<211> 775

<212> PRT

<213> Methanothermococcus okinawensis

<400> 29

Met Leu Met Leu Met Glu Val Leu Lys Glu Asn Gly Ile Ala Glu Leu  
 1 5 10 15  
 Arg Pro Pro Gln Lys Lys Val Val Glu Gly Gly Leu Leu Asn Lys Asn  
 20 25 30

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

Ile Ala Leu Ser Ile Gly Asp Phe Asp Glu Lys Glu Asp Leu Lys Gly  
 100 105 110  
 Tyr Asp Leu Ile Ile Thr Thr Ala Glu Lys Leu Asp Ser Leu Ile Arg  
 115 120 125

His Lys Val Glu Trp Ile Lys Asp Ile Ser Val Val Val Ile Asp Glu  
 130 135 140  
 Ile His Leu Ile Gly Asp Glu Ser Arg Gly Gly Thr Leu Glu Val Leu  
 145 150 155 160  
  
 Leu Thr Lys Leu Lys Thr Lys Lys Thr Ile Gln Ile Ile Gly Leu Ser  
 165 170 175  
 Ala Thr Ile Gly Asn Pro Glu Glu Leu Ala Lys Trp Leu Asn Ala Glu  
 180 185 190  
 Leu Ile Val Asp Glu Trp Arg Pro Val Lys Leu Lys Lys Gly Ile Gly  
 195 200 205  
 Tyr Gly Asn Lys Ile Met Phe Ile Asp Asp Asn Gly Asn Thr Ile Asn  
 210 215 220  
  
 Glu Val Ile Val Asp Glu Ile Ser Lys Asn Asn Met Phe Asn Leu Val  
 225 230 235 240  
 Val Asp Ser Ile Leu Lys Asp Gly Ser Cys Ile Ile Phe Cys Asn Ser  
 245 250 255  
 Lys Arg Gly Ala Val Gly Glu Ala Lys Lys Leu Asn Leu Lys Lys Tyr  
 260 265 270  
 Leu Ser Pro Asp Glu Ile Ser Glu Leu Arg His Leu Lys Glu Glu Val  
 275 280 285  
  
 Leu Ser Val Leu Asp Asn Pro Thr Lys Thr Cys Lys Asp Leu Ala Glu  
 290 295 300  
 Cys Ile Glu Lys Gly Val Ala Phe His His Ala Gly Leu Thr Tyr Glu  
 305 310 315 320  
 Gln Arg Lys Ile Val Glu Glu Gly Phe Arg Lys Lys Leu Ile Lys Ala  
 325 330 335  
 Ile Cys Cys Thr Pro Thr Leu Ser Ala Gly Ile Asn Met Pro Cys Arg  
 340 345 350  
  
 Arg Ala Ile Ile Arg Asp Leu Lys Arg Phe Ser Ser Arg Gly Tyr Ile  
 355 360 365  
 Pro Ile Pro Lys Met Glu Ile His Gln Cys Ile Gly Arg Ala Gly Arg

370                      375                      380  
 Pro Asn Leu Asp Pro Tyr Gly Glu Gly Ile Ile Tyr Ile Asn Asn Thr  
 385                      390                      395                      400  
 Glu Asn Pro Glu Leu Ile Glu Asn Ala Lys Asn Tyr Leu Ile Gly Asn  
                          405                      410                      415  
  
 Val Glu Glu Ile Tyr Ser Lys Leu Ser Asn Gln Lys Val Leu Arg Thr  
                          420                      425                      430  
 His Met Leu Gly Leu Ile Thr Thr Gly Asp Ile Lys Asn Lys Asn Asp  
                          435                      440                      445  
 Leu Glu Glu Phe Ile Lys Asn Thr Phe Tyr Ala Tyr Gln Tyr Gln Asn  
                          450                      455                      460  
 Thr Lys Lys Ile Leu Glu Asn Ile Tyr Glu Ile Thr Asn Phe Leu Glu  
 465                      470                      475                      480  
  
 Lys Asn Gly Phe Ile Glu Leu Asn Tyr Arg Arg Asp Glu Asn Lys Asp  
                          485                      490                      495  
 Lys Ser Asn Asn Ser His Asn Asn Lys Lys Asn Ile Ser Asn Thr Asn  
                          500                      505                      510  
 Asn Ser Ile Lys Met Leu Val Leu Asp Asn Asn Asn Ser Leu Thr Ile  
                          515                      520                      525  
 Lys Ser Arg His Glu Glu Asp Val Tyr Tyr Asn Ile Thr Pro Leu Gly  
                          530                      535                      540  
  
 Lys Lys Val Ser Glu Leu Tyr Ile Asp Pro Leu Ser Ala Glu Tyr Ile  
 545                      550                      555                      560  
 Ile Asp Gly Leu Lys Asn Leu His Lys Lys Thr Leu Ser Asn Pro Lys  
                          565                      570                      575  
 Asn Met Glu Cys Tyr Ile Leu His Ile Leu Tyr Ile Ile Ser Lys Thr  
                          580                      585                      590  
 Thr Glu Met Gln Pro Val Leu Arg Val Arg Arg Lys Glu Glu Asn Asp  
                          595                      600                      605  
  
 Leu Ile Asn Asp Met Ile Lys Leu Asp Ile Asp Val Asp Asp Val Ile  
                          610                      615                      620

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

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

Asn Asp Ile Ile Asn Asn Gln Lys Asn Ile Ile Asn Leu Leu Gly Glu  
740 745 750  
Lys Ile Ala Arg Lys Ile Leu Ser Glu Leu Gly Val Asp Thr Lys Phe  
755 760 765  
Gly Gln Met Arg Leu Ser Ile  
770 775

<210> 30

<211> 8

<212> PRT

<213> Artificial sequence

<220><223> Exemplary Hel308 motif

<400> 30

Gln Cys Ile Gly Arg Ala Gly Arg

1 5

<210

> 31

<211> 9

<212> PRT

<213> Artificial sequence

<220><223> Exemplary extended Hel308 motif

<400> 31

Gln Cys Ile Gly Arg Ala Gly Arg Pro

1 5

<210> 32

<211> 699

<212> PRT

<213> *Methanotorris igneus*

<400> 32

Met Gln Lys Tyr Ser His Val Phe Glu Val Leu Lys Glu Asn Gly Ile

1 5 10 15

Lys Glu Leu Arg Pro Pro Gln Lys Lys Val Ile Glu Lys Gly Leu Leu

20 25 30

Asn Lys Glu Lys Asn Phe Leu Ile Cys Ile Pro Thr Ala Ser Gly Lys

35 40 45

Thr Leu Ile Gly Glu Met Ala Leu Ile Asn His Leu Leu Asp Glu Asn

50 55 60

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

65 70 75 80

Leu Ala Ser Glu Lys Tyr Glu Glu Phe Lys Arg Lys Tyr Glu Lys Tyr

85 90 95

Gly Leu Lys Val Ala Leu Ser Ile Gly Asp Tyr Asp Glu Lys Glu Asp

100 105 110

Leu Ser Ser Tyr Asn Ile Ile Ile Thr Thr Ala Glu Lys Leu Asp Ser

115 120 125

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

130 135 140

Val Asp Glu Ile His Met Ile Asn Asp Glu Lys Arg Gly Gly Thr Leu

145 150 155 160

Glu Val Leu Leu Thr Lys Leu Lys Asn Leu Asp Val Gln Ile Ile Gly

165 170 175

Leu Ser Ala Thr Ile Gly Asn Pro Glu Glu Leu Ala Glu Trp Leu Asn

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



Ala Thr Arg Glu Ile Glu Asn Glu Phe Asp Leu Ile Ser Phe Ile Lys  
435 440 445

Asn Thr Phe Tyr Ala His Gln Tyr Gly Asn Leu Gly Gly Val Leu Arg  
450 455 460

Asn Ile Lys Glu Val Ile Asn Phe Leu Glu Glu Asn Asp Phe Ile Ala  
465 470 475 480

Asp Tyr Phe Pro Thr Lys Leu Gly Lys Arg Val Ser Glu Leu Tyr Ile  
485 490 495

Asp Pro Leu Ser Ala Lys Ile Ile Ile Asp Gly Leu Lys Glu Met Gly  
500 505 510

Asn Val Asp Asn Glu Glu Leu Tyr Tyr Leu Tyr Leu Ile Ser Lys Thr  
515 520 525

Leu Glu Met Met Pro Leu Leu Arg Val Asn Ser Phe Glu Glu Leu Asp  
530 535 540

Leu Ile Leu Glu Met Glu Glu Ala Gly Ile Tyr Asp Arg Thr Tyr Asp  
545 550 555 560

Asp Leu Ala Ala Phe Lys Asn Ala Lys Met Leu Tyr Asp Trp Ile Asn  
565 570 575

Glu Val Pro Glu Asp Glu Ile Leu Lys Lys Tyr Lys Ile Glu Pro Gly  
580 585 590

Ile Leu Arg Tyr Lys Val Glu Gln Ala Lys Trp Met Ile Tyr Ser Thr  
595 600 605

Lys Glu Ile Ala Lys Leu Leu Asn Arg Asn Ile Asp Thr Leu Ser Lys  
610 615 620

Leu Glu Ile Arg Leu Glu Tyr Gly Ala Lys Glu Asp Ile Ile Glu Leu  
625 630 635 640

Leu Lys Ile Lys Tyr Val Gly Arg Ala Arg Ala Arg Lys Leu Tyr Asp  
645 650 655

Ala Gly Ile Arg Ser Val Glu Asp Ile Ile Asn Asn Pro Lys Lys Val  
660 665 670

Ala Ser Leu Leu Gly Glu Lys Ile Ala Lys Lys Ile Leu Gly Glu Leu

675                      680                      685  
 Gly Met Lys Phe Gly Gln Gln Thr Leu Gln Ile  
 690                      695  
 <210> 33  
 <211> 720  
 <212> PRT  
 <213> Thermococcus gammatolerans  
 <400> 33  
 Met Lys Val Asp Glu Leu Pro Val Asp Glu Arg Leu Lys Ala Val Leu  
 1                      5                      10                      15  
 Lys Glu Arg Gly Ile Glu Glu Leu Tyr Pro Pro Gln Ala Glu Ala Leu  
 20                      25                      30  
  
 Lys Ser Gly Ala Leu Glu Gly Arg Asn Leu Val Leu Ala Ile Pro Thr  
 35                      40                      45  
 Ala Ser Gly Lys Thr Leu Val Ser Glu Ile Val Met Val Asn Lys Leu  
 50                      55                      60  
 Ile Gln Glu Gly Gly Lys Ala Val Tyr Leu Val Pro Leu Lys Ala Leu  
 65                      70                      75                      80  
 Ala Glu Glu Lys Tyr Arg Glu Phe Lys Glu Trp Glu Lys Leu Gly Leu  
 85                      90                      95  
  
 Lys Val Ala Ala Thr Thr Gly Asp Tyr Asp Ser Thr Asp Asp Trp Leu  
 100                      105                      110  
 Gly Arg Tyr Asp Ile Ile Val Ala Thr Ala Glu Lys Phe Asp Ser Leu  
 115                      120                      125  
 Leu Arg His Gly Ala Arg Trp Ile Asn Asp Val Lys Leu Val Val Ala  
 130                      135                      140  
 Asp Glu Val His Leu Ile Gly Ser Tyr Asp Arg Gly Ala Thr Leu Glu  
 145                      150                      155                      160  
  
 Met Ile Leu Thr His Met Leu Gly Arg Ala Gln Ile Leu Ala Leu Ser  
 165                      170                      175  
 Ala Thr Val Gly Asn Ala Glu Glu Leu Ala Glu Trp Leu Asp Ala Ser  
 180                      185                      190

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

435                      440                      445  
 Asp Leu Ser Ser Leu Glu Tyr Lys Ala Lys Glu Val Val Tyr Phe Leu  
 450                      455                      460  
 Ile Glu Asn Glu Phe Ile Asp Leu Asp Leu Glu Asp Arg Phe Ile Pro  
 465                      470                      475                      480  
  
 Leu Pro Phe Gly Lys Arg Thr Ser Gln Leu Tyr Ile Asp Pro Leu Thr  
 485                      490                      495  
 Ala Lys Lys Phe Lys Asp Ala Phe Pro Ala Ile Glu Arg Asn Pro Asn  
 500                      505                      510  
 Pro Phe Gly Ile Phe Gln Leu Ile Ala Ser Thr Pro Asp Met Ala Thr  
 515                      520                      525  
 Leu Thr Ala Arg Arg Arg Glu Met Glu Asp Tyr Leu Asp Leu Ala Tyr  
 530                      535                      540  
  
 Glu Leu Glu Asp Lys Leu Tyr Ala Ser Ile Pro Tyr Tyr Glu Asp Ser  
 545                      550                      555                      560  
 Arg Phe Gln Gly Phe Leu Gly Gln Val Lys Thr Ala Lys Val Leu Leu  
 565                      570                      575  
 Asp Trp Ile Asn Glu Val Pro Glu Ala Arg Ile Tyr Glu Thr Tyr Ser  
 580                      585                      590  
 Ile Asp Pro Gly Asp Leu Tyr Arg Leu Leu Glu Leu Ala Asp Trp Leu  
 595                      600                      605  
  
 Met Tyr Ser Leu Ile Glu Leu Tyr Lys Leu Phe Glu Pro Lys Glu Glu  
 610                      615                      620  
 Ile Leu Asn Tyr Leu Arg Asp Leu His Leu Arg Leu Arg His Gly Val  
 625                      630                      635                      640  
 Arg Glu Glu Leu Leu Glu Leu Val Arg Leu Pro Asn Ile Gly Arg Lys  
 645                      650                      655  
 Arg Ala Arg Ala Leu Tyr Asn Ala Gly Phe Arg Ser Val Glu Ala Ile  
 660                      665                      670  
  
 Ala Asn Ala Lys Pro Ala Glu Leu Leu Ala Val Glu Gly Ile Gly Ala  
 675                      680                      685

Lys Ile Leu Asp Gly Ile Tyr Arg His Leu Gly Ile Glu Lys Arg Val  
 690 695 700  
 Thr Glu Glu Lys Pro Lys Arg Lys Gly Thr Leu Glu Asp Phe Leu Arg  
 705 710 715 720  
 <210> 34  
 <211> 755  
 <212> PRT  
 <213> Thermococcus barophilus  
 <400> 34  
 Met Leu Ser Thr Lys Pro Lys Ala Tyr Lys Arg Phe Ser Pro Ile Gly  
  
 1 5 10 15  
 Tyr Ala Met Gln Val Asp Glu Leu Ser Lys Phe Gly Val Asp Glu Arg  
 20 25 30  
 Ile Ile Arg Lys Ile Lys Glu Arg Gly Ile Ser Glu Phe Tyr Pro Pro  
 35 40 45  
 Gln Ala Glu Ala Leu Arg Ser Gly Val Leu Asn Gly Glu Asn Leu Leu  
 50 55 60  
 Leu Ala Ile Pro Thr Ala Ser Gly Lys Thr Leu Val Ala Glu Ile Val  
  
 65 70 75 80  
 Met Leu His Lys Leu Phe Thr Gly Gly Gly Lys Ala Val Tyr Leu Val  
 85 90 95  
 Pro Leu Lys Ala Leu Ala Glu Glu Lys Tyr Arg Glu Phe Lys Thr Trp  
 100 105 110  
 Glu Asp Leu Gly Val Arg Val Ala Val Thr Thr Gly Asp Tyr Asp Ser  
 115 120 125  
 Ser Glu Glu Trp Leu Gly Lys Tyr Asp Ile Ile Ile Ala Thr Ser Glu  
  
 130 135 140  
 Lys Phe Asp Ser Leu Leu Arg His Lys Ser Arg Trp Ile Arg Asp Val  
 145 150 155 160  
 Thr Leu Ile Val Ala Asp Glu Ile His Leu Leu Gly Ser Tyr Asp Arg  
 165 170 175  
 Gly Ala Thr Leu Glu Met Ile Leu Ser His Met Leu Gly Lys Ala Gln

180 185 190  
Ile Leu Gly Leu Ser Ala Thr Val Gly Asn Ala Glu Glu Leu Ala Glu

195 200 205  
Trp Leu Asn Ala Lys Leu Val Val Ser Asp Trp Arg Pro Val Lys Leu

210 215 220  
Arg Lys Gly Val Phe Ala His Gly Gln Leu Ile Trp Glu Asp Gly Lys

225 230 235 240  
Val Asp Lys Phe Pro Pro Gln Trp Asp Ser Leu Val Ile Asp Ala Val

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

260 265 270  
Glu Lys Glu Ala Gly Met Leu Gly Lys Lys Val Arg Arg Leu Leu Thr

275 280 285  
Lys Pro Glu Ala Arg Arg Leu Lys Glu Leu Ala Glu Ser Leu Glu Ser

290 295 300  
Asn Pro Thr Asn Asp Lys Leu Lys Glu Val Leu Val Asn Gly Ala Ala

305 310 315 320  
Phe His His Ala Gly Leu Gly Arg Ala Glu Arg Thr Leu Ile Glu Asp

325 330 335  
Ala Phe Arg Glu Gly Leu Ile Lys Val Leu Thr Ala Thr Pro Thr Leu

340 345 350  
Ala Met Gly Val Asn Leu Pro Ser Phe Arg Val Ile Ile Arg Asp Thr

355 360 365  
Lys Arg Tyr Ser Thr Phe Gly Trp Ser Asp Ile Pro Val Leu Glu Ile

370 375 380  
Gln Gln Met Ile Gly Arg Ala Gly Arg Pro Lys Tyr Asp Lys Glu Gly

385 390 395 400  
Glu Ala Ile Ile Val Ala Lys Thr Glu Lys Pro Glu Glu Leu Met Glu

405 410 415  
Lys Tyr Ile Phe Gly Lys Pro Glu Lys Leu Phe Ser Met Leu Ser Asn

420 425 430

Asp Ala Ala Phe Arg Ser Gln Val Leu Ala Leu Ile Thr Asn Phe Gly  
 435 440 445  
 Val Glu Ser Phe Arg Glu Leu Ile Gly Phe Leu Glu Lys Thr Phe Tyr  
 450 455 460  
 Tyr His Gln Arg Lys Asp Leu Glu Ile Leu Glu Gly Lys Ala Lys Ser  
 465 470 475 480  
 Ile Val Tyr Phe Leu Leu Glu Asn Glu Phe Ile Asp Ile Asp Leu Asn  
 485 490 495  
 Asp Ser Phe Ile Ala Leu Pro Phe Gly Ile Arg Thr Ser Gln Leu Tyr  
 500 505 510  
 Leu Asp Pro Leu Thr Ala Lys Lys Phe Lys Asp Ala Leu Pro Gln Ile  
 515 520 525  
 Glu Glu Asn Pro Asn Pro Leu Gly Ile Phe Gln Leu Leu Ala Ser Thr  
 530 535 540  
 Pro Asp Met Gly Thr Leu Ser Ile Lys Arg Lys Glu Gln Glu Ser Tyr  
 545 550 555 560  
 Leu Asp Tyr Ala Tyr Glu Met Glu Asp Tyr Leu Tyr Arg Ser Ile Pro  
 565 570 575  
 Tyr Trp Glu Asp Tyr Glu Phe Gln Lys Phe Leu Ser Glu Val Lys Thr  
 580 585 590  
 Ala Lys Leu Leu Leu Asp Trp Ile Asn Glu Val Ser Glu Ala Lys Leu  
 595 600 605  
 Ile Glu Ala Tyr Gly Ile Asp Thr Gly Asp Leu Tyr Arg Ile Ile Glu  
 610 615 620  
 Leu Ala Asp Trp Leu Met Tyr Ser Leu Ile Glu Leu Ala Lys Val Leu  
 625 630 635 640  
 Asn Ala Gly Gly Glu Thr Ile Lys Tyr Leu Arg Arg Leu His Leu Arg  
 645 650 655  
 Leu Lys His Gly Val Arg Glu Glu Leu Leu Glu Leu Val Glu Leu Pro  
 660 665 670  
 Met Ile Gly Arg Arg Arg Ala Arg Ala Leu Tyr Asn Ala Gly Phe Lys

675 680 685  
 Asn Val Asn Asp Ile Val Lys Ala Lys Pro Ser Glu Leu Leu Ala Val  
 690 695 700  
 Glu Gly Ile Gly Val Lys Val Leu Glu Arg Ile Tyr Arg His Phe Gly

705 710 715 720  
 Val Glu Leu Pro Leu Leu Lys Asn Ile Lys Asp Pro Asp Lys Pro Glu  
 725 730 735  
 Asp Lys Pro Lys Glu Lys Pro Lys Pro Lys Lys Gly Thr Leu Asp Tyr  
 740 745 750  
 Phe Leu Lys  
 755

<210> 35

<211> 8

<212> PRT

<213> Artificial sequence

<220><223> Exemplary Hel308 motif

<400> 35

Gln Met Ile Gly Arg Ala Gly Arg

1 5

<210> 36

<211> 9

<212> PRT

<213> Artificial sequence

<220><223> Exemplary extended Hel308 motif

<400> 36

Gln Met Ile Gly Arg Ala Gly Arg Pro

1 5

<210> 37

<211> 744

<212> PRT

<213> Thermococcus sibiricus

<400> 37

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



1	5	10	15
Met Val Met Ser Met Lys Val Asp Glu Leu Lys Ser Leu Gly Val Asp			
	20	25	30
Glu Arg Ile Leu Arg Leu Leu Arg Glu Arg Gly Ile Glu Glu Leu Tyr			
	35	40	45
Pro Pro Gln Ala Asp Ala Leu Lys Thr Glu Val Leu Lys Gly Lys Asn			
	50	55	60
Leu Val Leu Ala Ile Pro Thr Ala Ser Gly Lys Thr Leu Val Ala Glu			
65	70	75	80
Ile Val Met Ile Asn Lys Ile Leu Arg Glu Gly Gly Lys Thr Val Tyr			
	85	90	95
Leu Val Pro Leu Lys Ala Leu Ala Glu Glu Lys Tyr Lys Glu Phe Lys			
	100	105	110
Phe Trp Glu Lys Leu Gly Ile Arg Ile Ala Met Thr Thr Gly Asp Tyr			
	115	120	125
Asp Ser Thr Glu Glu Trp Leu Gly Lys Tyr Asp Ile Ile Ile Ala Thr			
	130	135	140
Ser Glu Lys Phe Asp Ser Leu Leu Arg His Lys Ser Pro Trp Ile Lys			
145	150	155	160
Asp Ile Asn Leu Val Ile Ala Asp Glu Ile His Leu Leu Gly Ser Tyr			
	165	170	175
Asp Arg Gly Ala Thr Leu Glu Met Ile Leu Ala His Leu Asp Asp Lys			
	180	185	190
Ala Gln Ile Leu Gly Leu Ser Ala Thr Val Gly Asn Ala Glu Glu Val			
	195	200	205
Ala Glu Trp Leu Asn Ala Asp Leu Val Met Ser Glu Trp Arg Pro Val			
	210	215	220
Ala Leu Arg Lys Gly Val Phe Tyr His Gly Glu Leu Phe Trp Glu Asp			
225	230	235	240
Gly Ser Ile Glu Arg Phe Pro Thr Gln Trp Asp Ser Leu Val Ile Asp			
	245	250	255

Ala Leu Lys Lys Gly Lys Gln Ala Leu Val Phe Val Asn Thr Arg Arg  
260 265 270

Ser Ala Glu Lys Glu Ala Leu Leu Leu Ala Gly Lys Ile Gln Arg Phe  
275 280 285

Leu Thr Lys Pro Glu Glu Arg Lys Leu Lys Gln Leu Ala Asp Gly Leu  
290 295 300

Asp Thr Thr Pro Thr Asn Gln Lys Leu Lys Glu Ala Leu Thr Lys Gly  
305 310 315 320

Val Ala Phe His His Ala Gly Leu Gly Arg Thr Glu Arg Ser Ile Ile  
325 330 335

Glu Asp Ala Phe Arg Glu Gly Leu Ile Lys Val Ile Thr Ala Thr Pro  
340 345 350

Thr Leu Ser Ala Gly Val Asn Leu Pro Ala Tyr Arg Val Ile Ile Arg  
355 360 365

Asp Thr Lys Arg Tyr Ser Asn Phe Gly Trp Val Asp Ile Pro Val Leu  
370 375 380

Glu Ile Gln Gln Met Met Gly Arg Ala Gly Arg Pro Lys Tyr Asp Ile  
385 390 395 400

Glu Gly Gln Ala Ile Ile Ile Ala Lys Thr Glu Lys Pro Glu Asp Leu  
405 410 415

Met Lys Arg Tyr Val Leu Gly Lys Pro Glu Lys Leu Phe Ser Met Leu  
420 425 430

Ser Asn Glu Ala Ser Phe Arg Ser Gln Val Leu Ala Leu Ile Thr Asn  
435 440 445

Phe Gly Val Gly Asn Phe Lys Glu Leu Val Asn Phe Leu Glu Arg Thr  
450 455 460

Phe Tyr Tyr His Gln Arg Lys Asn Leu Glu Ala Leu Glu Gly Lys Ala  
465 470 475 480

Lys Ser Ile Val Tyr Phe Leu Phe Glu Asn Glu Phe Ile Asp Ile Asp  
485 490 495

Leu Asn Asp Gln Phe Met Pro Leu Pro Leu Gly Ile Arg Thr Ser Gln

500                      505                      510  
 Leu Tyr Leu Asp Pro Val Thr Ala Lys Lys Phe Lys Asp Ala Phe Glu  
 515                      520                      525  
 Lys Leu Glu Lys Asn Pro Asn Pro Leu Gly Ile Phe Gln Leu Leu Ala  
 530                      535                      540  
  
 Ser Thr Pro Asp Met Ser Ser Leu Arg Val Lys Arg Lys Glu Gln Glu  
 545                      550                      555                      560  
 Asp Leu Leu Asp Tyr Ala Tyr Glu Met Glu Glu Tyr Leu Tyr Gln Asn  
 565                      570                      575  
 Ile Pro Tyr Trp Glu Asp Tyr Lys Phe Glu Lys Phe Leu Gly Glu Thr  
 580                      585                      590  
 Lys Thr Ala Lys Leu Leu Leu Asp Trp Ile Asn Glu Val Asn Asp Val  
 595                      600                      605  
  
 Lys Ile Leu Glu Thr Tyr Glu Ile Asp Thr Gly Asp Leu Tyr Arg Ile  
 610                      615                      620  
 Leu Glu Leu Val Asp Trp Leu Met Tyr Ser Leu Ile Glu Leu Tyr Lys  
 625                      630                      635                      640  
 Leu Phe Asp Pro Lys Pro Glu Val Leu Asp Phe Leu Lys Lys Leu His  
 645                      650                      655  
 Ile Arg Val Lys His Gly Val Arg Glu Glu Leu Leu Glu Leu Ile Thr  
 660                      665                      670  
  
 Leu Pro Met Ile Gly Arg Lys Arg Ala Arg Ala Leu Tyr Asn Ala Gly  
 675                      680                      685  
 Phe Lys Gly Ile Asp Asp Ile Val Arg Ala Lys Ala Ser Glu Leu Leu  
 690                      695                      700  
 Lys Val Glu Gly Ile Gly Ile Gly Val Ile Glu Lys Ile Tyr Gln His  
 705                      710                      715                      720  
 Phe Gly Val Glu Leu Pro Thr Asn Glu Lys Lys Lys Lys Val Lys Lys  
 725                      730                      735  
  
 Gly Thr Leu Asp Glu Phe Phe Lys  
 740

<210> 38

<211> 729

<212> PRT

<213> Methanosarcina barkeri fusaro

<400> 38

Met Lys Ile Glu Ser Leu Asp Leu Pro Asp Glu Val Lys Gln Phe Tyr

1 5 10 15

Leu Asn Ser Gly Ile Met Glu Leu Tyr Pro Pro Gln Ala Glu Ala Val

20 25 30

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

35 40 45

Ala Ser Gly Lys Thr Leu Leu Ala Glu Leu Ala Met Leu Lys Ser Ile

50 55 60

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

65 70 75 80

Ser Glu Lys Phe Arg Arg Phe Arg Glu Phe Ser Glu Leu Gly Ile Arg

85 90 95

Val Gly Ile Ser Thr Gly Asp Tyr Asp Leu Arg Asp Glu Gly Leu Gly

100 105 110

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

115 120 125

Arg Asn Glu Thr Val Trp Met Gln Glu Ile Ser Val Val Val Ala Asp

130 135 140

Glu Val His Leu Ile Asp Ser Pro Asp Arg Gly Pro Thr Leu Glu Val

145 150 155 160

Thr Leu Ala Lys Leu Arg Lys Met Asn Pro Ser Cys Gln Ile Leu Ala

165 170 175

Leu Ser Ala Thr Val Gly Asn Ala Asp Glu Leu Ala Val Trp Leu Glu

180 185 190

Ala Glu Leu Val Val Ser Glu Trp Arg Pro Thr Glu Leu Leu Glu Gly

195 200 205

Val Phe Phe Asn Gly Thr Phe Tyr Cys Lys Asp Arg Glu Lys Thr Val

210                      215                      220  
 Glu Gln Ser Thr Lys Asp Glu Ala Val Asn Leu Ala Leu Asp Thr Leu  
 225                      230                      235                      240  
  
 Lys Lys Asp Gly Gln Cys Leu Val Phe Glu Ser Ser Arg Lys Asn Cys  
                          245                      250                      255  
 Met Ala Phe Ala Lys Lys Ala Ala Ser Thr Val Lys Lys Thr Leu Ser  
                          260                      265                      270  
 Ala Glu Asp Arg Asn Ala Leu Ala Gly Ile Ala Asp Glu Ile Leu Glu  
                          275                      280                      285  
 Asn Ser Glu Thr Asp Thr Ser Thr Asn Leu Ala Val Cys Ile Arg Ser  
                          290                      295                      300  
  
 Gly Thr Ala Phe His His Ala Gly Leu Thr Thr Pro Leu Arg Glu Leu  
 305                      310                      315                      320  
 Val Glu Asp Gly Phe Arg Ala Gly Arg Ile Lys Leu Ile Ser Ser Thr  
                          325                      330                      335  
 Pro Thr Leu Ala Ala Gly Leu Asn Leu Pro Ala Arg Arg Val Ile Ile  
                          340                      345                      350  
 Arg Asn Tyr Arg Arg Tyr Ser Ser Glu Asp Gly Met Gln Pro Ile Pro  
                          355                      360                      365  
  
 Val Leu Glu Tyr Lys Gln Met Ala Gly Arg Ala Gly Arg Pro Arg Leu  
                          370                      375                      380  
 Asp Pro Tyr Gly Glu Ala Val Leu Val Ala Lys Ser Tyr Lys Glu Phe  
 385                      390                      395                      400  
 Val Phe Leu Phe Glu Asn Tyr Ile Glu Ala Asn Ala Glu Asp Ile Trp  
                          405                      410                      415  
 Ser Lys Leu Gly Thr Glu Asn Ala Leu Arg Thr His Val Leu Ser Thr  
                          420                      425                      430  
  
 Ile Ser Asn Gly Phe Ala Arg Thr Tyr Asp Glu Leu Met Asp Phe Leu  
                          435                      440                      445  
 Glu Ala Thr Phe Phe Ala Phe Gln Tyr Ser Asn Phe Gly Leu Ser Thr  
                          450                      455                      460

Val Val Asn Glu Cys Leu Asn Phe Leu Arg Gln Glu Gly Met Leu Glu  
465 470 475 480

Lys Asp Asp Ala Leu Ile Pro Thr Ser Phe Gly Lys Leu Val Ser Arg  
485 490 495

Leu Tyr Ile Asp Pro Leu Ser Ala Ala Arg Ile Ala Lys Gly Leu Lys  
500 505 510

Gly Ala Lys Ser Leu Ser Glu Leu Thr Leu Leu His Leu Val Cys Ser  
515 520 525

Thr Pro Asp Met Arg Leu Leu Tyr Met Arg Ser His Asp Tyr Gln Asp  
530 535 540

Ile Asn Asp Tyr Val Met Ala His Ala Ser Glu Phe Val Lys Val Pro  
545 550 555 560

Ser Pro Phe Asp Thr Thr Glu Tyr Glu Trp Phe Leu Gly Glu Val Lys  
565 570 575

Thr Ser Leu Leu Leu Leu Asp Trp Ile His Glu Lys Ser Glu Asn Glu  
580 585 590

Ile Cys Leu Lys Phe Gly Thr Gly Glu Gly Asp Ile His Ser Ile Ala  
595 600 605

Asp Ile Ala Glu Trp Ile Met His Val Thr Ser Gln Leu Ala Gly Leu  
610 615 620

Leu Asp Leu Lys Gly Ala Arg Glu Ala Ala Glu Leu Glu Lys Arg Ile  
625 630 635 640

His Tyr Gly Ala Ala Pro Glu Leu Ile Asp Leu Leu Asn Ile Arg Gly  
645 650 655

Ile Gly Arg Val Arg Ala Arg Lys Leu Tyr Glu Ala Gly Phe Lys Ser  
660 665 670

Ser Ala Glu Leu Ala Glu Val Asp Pro Glu Lys Val Ala Ala Leu Leu  
675 680 685

Gly Pro Lys Ile Ala Asp Arg Ile Phe Lys Gln Ile Arg Gly Arg Gly  
690 695 700

Thr Ser Ser Gly Ile Ile Ala Ser Glu Pro Pro Glu Lys Ser Pro Tyr

705	710					715					720				
Ser	Gly	Gln	Lys	Thr	Ile	Ser	Asp	Tyr							
725															
<210>	39														
<211>	730														
<212>	PRT														
<213>	Methanosarcina acetivorans														
<400>	39														
Met	Lys	Ile	Glu	Ser	Leu	Asp	Leu	Pro	Asp	Glu	Val	Lys	Arg	Phe	Tyr
1				5					10					15	
Glu	Asn	Ser	Gly	Ile	Pro	Glu	Leu	Tyr	Pro	Pro	Gln	Ala	Glu	Ala	Val
			20					25					30		
Glu	Lys	Gly	Leu	Leu	Glu	Gly	Lys	Asn	Leu	Leu	Ala	Ala	Ile	Pro	Thr
		35					40					45			
Ala	Ser	Gly	Lys	Thr	Leu	Leu	Ala	Glu	Leu	Ala	Met	Leu	Lys	Ser	Val
	50					55					60				
Leu	Ala	Gly	Gly	Lys	Ala	Leu	Tyr	Ile	Val	Pro	Leu	Arg	Ala	Leu	Ala
65					70					75					80
Ser	Glu	Lys	Phe	Arg	Arg	Phe	Gln	Asp	Phe	Ser	Glu	Leu	Gly	Ile	Arg
				85					90					95	
Val	Gly	Ile	Ser	Thr	Gly	Asp	Tyr	Asp	Arg	Arg	Asp	Glu	Gly	Leu	Gly
			100					105					110		
Ile	Asn	Asp	Ile	Ile	Val	Ala	Thr	Ser	Glu	Lys	Thr	Asp	Ser	Leu	Leu
		115						120				125			
Arg	Asn	Glu	Thr	Ala	Trp	Met	Gln	Glu	Ile	Ser	Val	Val	Val	Val	Asp
	130					135					140				
Glu	Val	His	Leu	Ile	Asp	Ser	Ala	Asp	Arg	Gly	Pro	Thr	Leu	Glu	Val
145					150					155					160
Thr	Leu	Ala	Lys	Leu	Arg	Lys	Met	Asn	Pro	Phe	Cys	Gln	Ile	Leu	Ala
				165					170				175		
Leu	Ser	Ala	Thr	Val	Gly	Asn	Ala	Asp	Glu	Leu	Ala	Ala	Trp	Leu	Asp
			180					185					190		

Ala Glu Leu Val Leu Ser Glu Trp Arg Pro Thr Asp Leu Met Glu Gly  
195 200 205

Val Phe Phe Asp Gly Thr Phe Phe Cys Lys Asp Lys Glu Lys Leu Ile  
210 215 220

Glu Gln Pro Thr Lys Asp Glu Ala Ile Asn Leu Val Leu Asp Thr Leu  
225 230 235 240

Arg Glu Gly Gly Gln Cys Leu Val Phe Glu Ser Ser Arg Lys Asn Cys  
245 250 255

Met Gly Phe Ala Lys Lys Ala Thr Ser Ala Val Lys Lys Thr Leu Ser  
260 265 270

Ala Glu Asp Lys Glu Lys Leu Ala Gly Ile Ala Asp Glu Ile Leu Glu  
275 280 285

Asn Ser Glu Thr Asp Thr Ala Ser Val Leu Ala Ser Cys Val Arg Ala  
290 295 300

Gly Thr Ala Phe His His Ala Gly Leu Thr Ser Pro Leu Arg Glu Leu  
305 310 315 320

Val Glu Thr Gly Phe Arg Glu Gly Tyr Val Lys Leu Ile Ser Ser Thr  
325 330 335

Pro Thr Leu Ala Ala Gly Leu Asn Leu Pro Ala Arg Arg Val Ile Ile  
340 345 350

Arg Ser Tyr Arg Arg Tyr Ser Ser Asp Ser Gly Met Gln Pro Ile Pro  
355 360 365

Val Leu Glu Tyr Lys Gln Met Ala Gly Arg Ala Gly Arg Pro Arg Leu  
370 375 380

Asp Pro Tyr Gly Glu Ala Val Leu Leu Ala Lys Ser Tyr Glu Glu Leu  
385 390 395 400

Leu Phe Leu Phe Glu Lys Tyr Ile Glu Ala Gly Ala Glu Asp Ile Trp  
405 410 415

Ser Lys Leu Gly Thr Glu Asn Ala Leu Arg Thr His Val Leu Ser Thr  
420 425 430

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



435                      440                      445  
 Glu Ala Thr Phe Phe Ala Tyr Gln Tyr Ser Asn Phe Gly Leu Ser Val  
 450                      455                      460  
  
 Val Val Asp Glu Cys Leu Asn Phe Leu Arg Gln Glu Gly Met Leu Glu  
 465                      470                      475                      480  
 Gln Asp Ser Asp Ala Leu Ile Ser Thr Met Phe Gly Lys Leu Val Ser  
 485                      490                      495  
 Arg Leu Tyr Ile Asp Pro Leu Ser Ala Ala Leu Ile Ala Lys Gly Leu  
 500                      505                      510  
 Arg Glu Ala Gly Thr Leu Thr Glu Leu Thr Leu Leu His Leu Val Cys  
 515                      520                      525  
  
 Ser Thr Pro Asp Met Arg Leu Met Tyr Met Arg Ser Gln Asp Tyr Gln  
 530                      535                      540  
 Asp Ile Asn Asp Phe Val Met Ala His Ala Glu Glu Phe Ser Lys Val  
 545                      550                      555                      560  
 Pro Ser Pro Phe Asn Ile Val Glu Tyr Glu Trp Phe Leu Ser Glu Val  
 565                      570                      575  
 Lys Thr Ser Leu Leu Leu Met Asp Trp Ile His Glu Lys Pro Glu Asn  
 580                      585                      590  
  
 Glu Ile Cys Leu Lys Phe Gly Thr Gly Glu Gly Asp Ile His Thr Thr  
 595                      600                      605  
 Ala Asp Ile Ala Glu Trp Ile Met His Val Ala Thr Gln Leu Ala Arg  
 610                      615                      620  
 Leu Leu Asp Leu Lys Gly Ala Lys Glu Ala Ala Glu Leu Glu Lys Arg  
 625                      630                      635                      640  
 Ile His Tyr Gly Ala Gly Pro Glu Leu Met Asp Leu Leu Asp Ile Arg  
 645                      650                      655  
  
 Gly Ile Gly Arg Val Arg Ala Arg Lys Leu Tyr Gly Ala Gly Phe Lys  
 660                      665                      670  
 Ser Thr Ala Asp Leu Ala Gly Ala Thr Pro Glu Lys Val Ala Ala Leu  
 675                      680                      685

Val Gly Pro Lys Ile Ala Glu Arg Ile Phe Arg Gln Ile Gly Arg Arg  
690 695 700

Glu Ala Val Ser Glu Ile Ser Asp Ser Glu Arg Leu Glu Lys Ser Ser  
705 710 715 720

Gln Asp Gly Gln Ser Thr Ile Ser Asp Phe  
725 730

<210> 40  
<211> 729  
<212> PRT  
<213> Methanohalophilus mahii  
<400> 40

Met Lys Ile Glu Glu Leu Asp Leu Pro Ser Glu Ala Ile Glu Val Tyr  
1 5 10 15

Leu Gln Ala Gly Ile Glu Glu Leu Tyr Pro Pro Gln Ala Asp Ala Val  
20 25 30

Glu Lys Gly Leu Leu Gln Gly Glu Asn Leu Leu Ala Ala Ile Pro Thr  
35 40 45

Ala Ser Gly Lys Thr Leu Leu Ala Glu Met Ala Met Leu Lys Ala Ile  
50 55 60

Lys Lys Gly Gly Lys Ala Leu Tyr Ile Val Pro Leu Arg Ala Leu Ala  
65 70 75 80

Ser Glu Lys Phe Arg Asp Phe Lys Arg Phe Glu Ser Leu Gly Ile Lys  
85 90 95

Thr Ala Ile Ser Thr Gly Asp Phe Asp Ser Arg Asp Glu Trp Leu Gly  
100 105 110

Ser Asn Asp Ile Ile Val Ala Thr Ser Glu Lys Thr Asp Ser Leu Leu  
115 120 125

Arg Asn Ser Thr Pro Trp Met Lys Asp Ile Thr Ala Val Ile Val Asp  
130 135 140

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

Thr Leu Ala Lys Leu Lys Arg Leu Asn Pro Gly Ala Gln Val Val Ala

165	170	175
Leu Ser Ala Thr Val Gly Asn Ala Met Glu Ile Ala Gln Trp Leu Glu		
180	185	190
Ala Lys Leu Val Leu Ser Glu Trp Arg Pro Thr Tyr Leu His Glu Gly		
195	200	205
Ile Phe Tyr Gly Asp Ala Ile Asn Phe Asp Glu Asp Gln Thr Phe Ile		
210	215	220
Glu Arg Arg His Lys Glu Asp Ser Val Asn Leu Val Ile Asp Thr Val		
225	230	235
		240
Ile Gln Gly Gly Gln Cys Leu Val Phe Asp Ser Ser Arg Arg Asn Cys		
245	250	255
Val Gly Phe Ala Lys Lys Cys Ala Pro Ala Val Gly Glu Leu Leu Asp		
260	265	270
Arg Gln Asn Arg Asn Glu Leu Glu Glu Val Ala Lys Glu Val Leu Glu		
275	280	285
Asn Gly Glu Thr Lys Leu Thr Glu Thr Leu Ala Tyr Cys Ile Lys Lys		
290	295	300
Gly Val Ala Phe His His Ala Gly Leu Asn Ser Ala His Arg Arg Ile		
305	310	315
		320
Val Glu Asp Ala Phe Arg Asn Asn Leu Ile Lys Met Ile Cys Ser Thr		
325	330	335
Pro Thr Leu Ala Ala Gly Leu Asn Leu Pro Ala Arg Arg Val Ile Ile		
340	345	350
Arg Ser Tyr Lys Arg Tyr Asp Pro Asn Ala Gly Met Gln Pro Ile Pro		
355	360	365
Val Leu Asp Tyr Lys Gln Met Ala Gly Arg Ala Gly Arg Pro His Leu		
370	375	380
Asp Pro Tyr Gly Glu Ala Val Val Ile Val Lys Thr Tyr Glu Glu Phe		
385	390	395
		400
Thr Asp Val Leu Glu Arg Tyr Ile Ser Ala Ser Ala Glu Asp Ile Trp		
405	410	415

Ser Lys Leu Gly Thr Glu Asn Ala Leu Arg Thr His Ile Leu Ser Thr  
420 425 430

Ile Ala Ser Gly Phe Ala Asn Cys His Arg Glu Ile Leu Thr Phe Leu  
435 440 445

Gly Ser Thr Phe Phe Ala His Gln Gln Gln Ser Trp Asn Phe Glu Glu  
450 455 460

Leu Leu Glu Asp Cys Leu Ile Phe Leu Lys Asn Glu Gly Met Leu Glu  
465 470 475 480

Gln Asp Asn Glu Thr Ile Arg Ala Thr Glu Leu Gly Lys Met Ile Ser  
485 490 495

Ser Leu Tyr Ile Asp Pro Leu Ser Ala Ser Lys Ile Ile Arg Gly Leu  
500 505 510

Glu Lys Thr Thr His Val Thr Asp Met Thr Leu Leu Gln Leu Ile Cys  
515 520 525

Ser Thr Pro Asp Met Arg Leu Leu Tyr Leu Arg Asn Arg Asp Tyr Glu  
530 535 540

Ile Ile Asn Asp Tyr Val Met Asn His Thr Glu Glu Phe Ile Glu Val  
545 550 555 560

Pro Ser Pro Phe Lys Gln Ile Glu Tyr Glu Trp Phe Leu Ser Glu Val  
565 570 575

Lys Thr Ala Leu Leu Leu Leu Glu Trp Ile Asn Glu Lys Ser Leu Glu  
580 585 590

Lys Ile Val Glu Asn Tyr Gln Val Gly Glu Gly Asp Ile Tyr Ala Ser  
595 600 605

Ser Asp Ile Ala Glu Trp Leu Met His Ala Thr Gln Arg Ile Ala Ser  
610 615 620

Arg Ile Asn Pro Gln Leu Glu Thr Glu Cys Ala Lys Leu Glu Lys Arg  
625 630 635 640

Ile His Tyr Gly Ala Gly Ser Glu Leu Ile Glu Leu Val Glu Ile Pro  
645 650 655

Asn Val Gly Arg Ala Arg Ala Arg Lys Leu Phe Lys Lys Gly Tyr Arg

660                      665                      670  
 Ser Arg Gln Lys Leu Ala Thr Ala Asp Glu Lys Gln Leu Ala Gly Ile  
 675                      680                      685  
  
 Val Gly Pro Lys Ile Ala Gln Lys Ile Leu Ser Tyr Leu Gly Arg Glu  
 690                      695                      700  
 Thr Asp Ser Asn Gly Tyr Val Glu Pro Glu Thr Leu Glu Asn Lys Lys  
 705                      710                      715                      720  
 Gln Gln Lys Thr Phe Gln Asp Phe Ile  
 725  
 <210> 41  
 <211> 730  
 <212> PRT  
 <213> Methanosarcina mazei  
 <400> 41  
 Met Lys Ile Glu Ser Leu Asp Leu Pro Asp Glu Ile Lys Arg Phe Tyr  
 1                      5                      10                      15  
  
 Glu Asn Ser Gly Ile Leu Glu Leu Tyr Pro Pro Gln Ala Glu Ala Val  
 20                      25                      30  
 Glu Lys Gly Leu Leu Glu Gly Lys Asn Leu Leu Ala Ala Ile Pro Thr  
 35                      40                      45  
 Ala Ser Gly Lys Thr Leu Leu Ala Glu Leu Ala Met Leu Lys Ser Val  
 50                      55                      60  
 Leu Asn Gly Gly Lys Ala Leu Tyr Ile Val Pro Leu Arg Ala Leu Ala  
 65                      70                      75                      80  
  
 Ser Glu Lys Phe Arg Arg Phe Gln Glu Phe Ser Val Leu Gly Met Arg  
 85                      90                      95  
 Val Gly Ile Ser Thr Gly Asp Tyr Asp Arg Arg Asp Glu Gly Leu Gly  
 100                      105                      110  
 Ile Asn Asp Ile Ile Val Ala Thr Ser Glu Lys Thr Asp Ser Leu Leu  
 115                      120                      125  
 Arg Asn Glu Thr Ala Trp Met Gln Glu Ile Ser Val Val Val Ala Asp

130                      135                      140  
 Glu Val His Leu Ile Asp Ser Pro Asp Arg Gly Pro Thr Leu Glu Ile  
 145                      150                      155                      160  
 Thr Leu Ser Lys Leu Arg Arg Met Asn Pro Ser Cys Gln Val Leu Ala  
 165                      170                      175  
 Leu Ser Ala Thr Val Gly Asn Ala Asp Glu Leu Ala Ala Trp Leu Asp  
 180                      185                      190  
 Ala Glu Leu Val Leu Ser Glu Trp Arg Pro Thr Asp Leu Met Glu Gly  
 195                      200                      205  
  
 Val Phe Tyr Asn Gly Ile Phe Tyr Cys Lys Asp Lys Glu Lys Pro Val  
 210                      215                      220  
 Gly Gln Pro Thr Lys Asp Glu Ala Val Asn Leu Val Leu Asp Thr Ile  
 225                      230                      235                      240  
 Lys Glu Gly Gly Gln Cys Leu Val Phe Glu Ser Ser Arg Lys Asn Cys  
 245                      250                      255  
 Met Gly Phe Ala Lys Lys Ala Val Ser Ala Val Lys Lys Thr Leu Ser  
 260                      265                      270  
  
 Asn Glu Asp Arg Glu Thr Leu Ala Gly Ile Ala Asp Glu Ile Ile Glu  
 275                      280                      285  
 Asn Ser Glu Thr Asp Val Ser Ser Val Leu Ala Thr Cys Val Arg Ser  
 290                      295                      300  
 Gly Thr Ala Phe His His Ala Gly Leu Thr Thr Pro Leu Arg Glu Leu  
 305                      310                      315                      320  
 Val Glu Asn Gly Phe Arg Glu Gly Arg Ile Lys Ile Ile Ser Ser Thr  
 325                      330                      335  
  
 Pro Thr Leu Ala Ala Gly Leu Asn Leu Pro Ala Arg Arg Val Ile Ile  
 340                      345                      350  
 Arg Ser Tyr Arg Arg Tyr Ser Ser Asp Ser Gly Met Gln Pro Ile Pro  
 355                      360                      365  
 Val Leu Glu Tyr Lys Gln Met Ala Gly Arg Ala Gly Arg Pro Arg Leu  
 370                      375                      380

Asp Pro Tyr Gly Glu Ala Val Leu Leu Ala Lys Ser Tyr Glu Glu Phe  
385 390 395 400

Val Phe Leu Phe Glu Lys Tyr Ile Glu Ala Gly Ala Glu Asp Ile Trp  
405 410 415

Ser Lys Leu Gly Thr Glu Asn Ala Leu Arg Thr His Ile Leu Ser Thr  
420 425 430

Ile Ser Asn Gly Phe Ala Arg Thr Arg Glu Glu Leu Met Asp Phe Leu  
435 440 445

Glu Ala Thr Phe Phe Ala Phe Gln Tyr Ser Asn Phe Gly Leu Ser Ala  
450 455 460

Val Val Asp Glu Cys Leu Asp Phe Leu Arg Arg Glu Gly Met Leu Glu  
465 470 475 480

Lys Asp Pro Asp Ala Leu Val Ser Thr Val Phe Gly Lys Leu Val Ser  
485 490 495

Arg Leu Tyr Ile Asp Pro Leu Ser Ala Ala Leu Ile Ala Lys Gly Leu  
500 505 510

Arg Glu Ala Gly Thr Leu Thr Glu Leu Thr Leu Leu His Leu Ile Cys  
515 520 525

Ser Thr Pro Asp Met Arg Leu Met Tyr Met Arg Ser Gln Asp Tyr Gln  
530 535 540

Glu Val Asn Asp Tyr Val Met Ala His Ala Gly Glu Phe Ser Lys Val  
545 550 555 560

Pro Asn Pro Phe Asn Ile Ala Glu Tyr Glu Trp Phe Leu Gly Glu Val  
565 570 575

Lys Thr Ser Leu Leu Leu Met Asp Trp Ile His Glu Lys Pro Glu Asn  
580 585 590

Glu Ile Cys Leu Lys Phe Gly Ile Gly Glu Gly Asp Ile His Ala Thr  
595 600 605

Ala Asp Ile Ala Glu Trp Ile Met His Val Thr Ala Gln Leu Ala Gly  
610 615 620

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

625                      630                      635                      640  
 Ile Arg Tyr Gly Ala Ala Pro Glu Leu Met Asp Leu Leu Asp Ile Arg  
                          645                      650                      655  
  
 Ser Val Gly Arg Val Arg Ala Arg Lys Leu Tyr Glu Ala Gly Phe Lys  
                          660                      665                      670  
 Ser Thr Ala Glu Leu Ala Ala Ala Ser Pro Glu His Ile Ala Val Leu  
                          675                      680                      685  
 Val Gly Pro Lys Ile Thr Glu Arg Ile Phe Lys Gln Ile Gly Arg Arg  
                          690                      695                      700  
 Glu Ala Val Ser Glu Phe Ser Asp Ile Glu Pro Leu Glu Lys Gly Ser  
 705                      710                      715                      720

Ser Asp Gly Gln Arg Thr Ile Ser Asp Tyr  
                          725                      730

<210> 42

<211> 693

<212> PRT

<213> Methanosaeta thermophila

<400> 42

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

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



100	105	110	
Arg Leu Gly Arg Asn Asp Ile Ile Ile Ala Thr Ser Glu Lys Ala Asp			
115	120	125	
Ser Leu Ile Arg Asn Gly Ala Ser Trp Val Arg Arg Ile Gly Val Leu			
130	135	140	
Val Val Asp Glu Ile His Leu Leu Asp Ser Ala Asn Arg Gly Pro Thr			
145	150	155	160
Leu Glu Met Thr Met Thr Lys Leu Met His Leu Asn Pro Glu Met Gln			
165	170	175	
Val Ile Gly Leu Ser Ala Thr Ile Ala Asn Gly Arg Glu Ile Ala Asp			
180	185	190	
Trp Ile Lys Gly Glu Ile Val Ser Ser Asp Trp Arg Pro Val Arg Leu			
195	200	205	
Arg Glu Gly Val Leu Leu Glu Asp Arg Leu Val Phe Pro Asp Gly Glu			
210	215	220	
Ile Gln Leu Glu Asn Arg Asn Arg Asp Pro Val Leu Asn Leu Val Leu			
225	230	235	240
Asp Thr Val Asp Gln Gly Gly Gln Met Leu Ile Phe Glu Ser Thr Arg			
245	250	255	
Arg Asn Ala Glu Ser Met Ala Lys Lys Val Ser Gly Ala Leu Gln Glu			
260	265	270	
Ser Gly Glu Thr Ile Glu Leu Ala Glu Arg Leu Ser Gly Glu Gly Lys			
275	280	285	
Thr Ala Lys Lys Leu Ala Met Cys Leu Arg His Gly Ala Ala Phe His			
290	295	300	
His Ala Gly Leu Leu Pro Glu Gln Arg Arg Leu Ile Glu Leu Gly Phe			
305	310	315	320
Arg Gln Asn Val Val Lys Val Ile Ala Cys Thr Pro Thr Leu Ala Ala			
325	330	335	
Gly Leu Asn Leu Pro Ala Arg Arg Val Leu Ile Arg Ser Tyr Lys Arg			
340	345	350	

Tyr Glu Ala Gly Leu Gly Thr Arg Pro Ile Pro Val Met Glu Tyr Arg  
 355 360 365  
  
 Gln Met Ala Gly Arg Ala Gly Arg Pro Gly Leu Asp Pro Tyr Gly Glu  
 370 375 380  
 Ser Leu Ile Met Ala Arg Ser Glu Ser Glu Leu Gln Lys Leu Met Asp  
 385 390 395 400  
 His Tyr Val Met Gly Glu Pro Glu Asp Ile Trp Ser Lys Leu Ala Ser  
 405 410 415  
 Glu Arg Ala Leu Arg Thr His Val Leu Ala Thr Ile Ala Ser Arg Phe  
 420 425 430  
  
 Ala Asp Ser Val Asp Ser Leu Ser Arg Leu Met Ala Ser Thr Phe Tyr  
 435 440 445  
 Ala Arg Gln Gln Asp Pro Ser Tyr Leu Gly Glu Thr Ile Ala Ser Val  
 450 455 460  
 Leu Glu Phe Leu Val Arg Ser Asp Met Ile Asp Lys Asp Leu Thr Pro  
 465 470 475 480  
 Thr Pro Leu Gly Ala Leu Val Ser Arg Leu Tyr Ile Asp Pro Leu Ser  
 485 490 495  
  
 Ala Met Val Met Ile Gln Glu Ile Arg Gly Ile Arg Arg Pro Thr Val  
 500 505 510  
 Leu Thr Leu Leu His Val Ile Thr Met Thr Pro Asp Met Glu Leu Leu  
 515 520 525  
 Phe Val Gln Gln Ser Asp Asn Trp Leu Glu Asp Phe Ile Ser Glu His  
 530 535 540  
 Ser Ser Glu Leu Gly Asn Glu Lys Asn Phe Asp Trp Leu Leu Arg Glu  
 545 550 555 560  
  
 Val Lys Thr Ala Ser Met Leu Met Asp Trp Ile Asn Glu Val His Glu  
 565 570 575  
 Asp Arg Ile Glu Asp Arg Tyr Ser Ile Ser Pro Gly Asp Leu Val Arg  
 580 585 590  
 Ile Ala Glu Thr Ala Glu Trp Leu Met Ser Ala Leu His Arg Ile Ser

595 600 605  
Lys His Met Asp Leu Gly Val Thr Tyr Leu Ala Glu Arg Leu Ala Leu  
610 615 620

Arg Ile His Tyr Gly Ala Gly Asp Glu Leu Leu Gln Leu Leu Glu Leu  
625 630 635 640

Lys Gly Ile Gly Arg Val Arg Ala Arg Lys Leu Tyr Gln Ala Gly Tyr  
645 650 655

Arg Ser Leu Glu Asp Leu Lys Ala Ala Asp Lys Ser Thr Leu Ser Glu  
660 665 670

Ile Leu Gly Pro Lys Ile Ala Glu Gly Val Ile Ser Gln Leu Lys Glu  
675 680 685

Pro Gly Val Ser Ala  
690

<210> 43

<211> 739

<212> PRT

<213> Methanosalsum zhilinae

<400> 43

Met Asn Ile Asn Asn Leu Asn Leu Pro Glu Lys Val Lys Lys Tyr Tyr  
1 5 10 15

Thr Asp Thr Gly Ile Val Asp Leu Tyr Pro Pro Gln Arg Glu Ala Val  
20 25 30

Asp Lys Gly Leu Leu Asp Gly Glu Asn Ile Val Ala Ala Ile Pro Thr  
35 40 45

Ala Ser Gly Lys Thr Leu Leu Ala Glu Leu Cys Met Leu Lys Ser Ile  
50 55 60

Gly Met Gly Gly Lys Cys Leu Tyr Ile Val Pro Leu Lys Ala Leu Ala  
65 70 75 80

Ser Glu Lys Tyr Ser Arg Phe Arg Glu Phe Glu Ser Leu Gly Ile Lys  
85 90 95

Val Gly Ile Ala Thr Gly Asp Leu Asp Ser Arg Glu Glu Trp Leu Gly

100	105	110	
Lys Asn Asp Ile Ile Ile Ala Thr Ser Glu Lys Val Asp Ser Leu Leu			
115	120	125	
Arg Asn Glu Ser Ser Trp Met Lys Glu Ile Asn Thr Val Val Ala Asp			
130	135	140	
Glu Val His Leu Leu Asn Ser Val Asn Arg Gly Pro Thr Leu Glu Ile			
145	150	155	160
Thr Leu Ala Lys Leu Ile His Leu Asn Pro Gly Ser Gln Ile Ile Ala			
165	170	175	
Leu Ser Ala Thr Ile Gly Asn Pro Glu Asp Ile Ala Gly Trp Leu Gly			
180	185	190	
Ala Arg Leu Val Val Ser Glu Trp Arg Pro Thr Asp Leu Tyr Glu Gly			
195	200	205	
Ile Leu Leu Asp Gly Leu Leu His Ile Gly Asn Ile Lys Lys Asp Ile			
210	215	220	
Gln Asp Glu Ser Arg Asp Asp Ala Val Asn Leu Val Ile Asp Thr Val			
225	230	235	240
Lys Asp Lys Gly Gln Cys Leu Val Phe Glu Ser Ser Arg Arg Asn Cys			
245	250	255	
Met Gly Phe Ala Lys Lys Ala Gly Lys Trp Val Ser Lys Ile Leu Asp			
260	265	270	
Glu His Asp Thr Ile Gln Leu Lys Ser Leu Ser Gln Glu Ile Gly Glu			
275	280	285	
Ala Gly Glu Thr Glu Ile Ala Asp Val Leu Ser Arg Cys Val Arg Gln			
290	295	300	
Gly Val Ala Phe His His Ala Gly Leu Asn Ser Glu His Arg Arg Met			
305	310	315	320
Val Glu Glu Gly Phe Arg Lys Asn Leu Ile Lys Met Ile Ser Ser Thr			
325	330	335	
Pro Thr Leu Ala Ala Gly Leu Asn Leu Pro Ala Arg Arg Val Ile Ile			
340	345	350	

Arg Ser Tyr Lys Arg Tyr Asp Pro Asn Phe Gly Met Lys Pro Ile Pro  
355 360 365

Val Leu Glu Tyr Lys Gln Met Ala Gly Arg Ala Gly Arg Pro His Leu  
370 375 380

Asp Pro Tyr Gly Glu Ser Val Leu Ile Ala Arg Ser Tyr Asp Glu Phe  
385 390 395 400

Met Asp Ile Met Glu Asn Tyr Val Asn Ala Asp Pro Glu Asp Ile Trp  
405 410 415

Ser Lys Leu Gly Thr Glu Asn Ala Leu Arg Thr His Val Leu Ser Thr  
420 425 430

Ile Val Asn Gly Phe Ala Tyr Thr Tyr Arg Gly Leu Met Asp Phe Val  
435 440 445

Lys Met Thr Phe Phe Ala Tyr Gln Lys Glu Ala Ser Asp Leu His Asp  
450 455 460

Val Ile Glu Glu Cys Val Arg Phe Leu Ile Asp Asn Glu Met Ile Ile  
465 470 475 480

Ser Asp Ser Asn Asp Ile Leu Pro Glu Ser Ala Phe Arg Ser Thr Ala  
485 490 495

Thr Gly Lys Leu Ile Ser Met Leu Tyr Ile Asp Pro Leu Ser Gly Ser  
500 505 510

Leu Ile Met Asp Gly Ile Arg Lys Ala Asp Tyr Phe Glu Asp Ile Thr  
515 520 525

Met Met His Leu Ile Cys Ser Thr Pro Asp Met Lys Asn Leu Tyr Met  
530 535 540

Arg Ser Ser Asp Tyr Glu Asn Val Asn Met Tyr Val Leu Gln Asn Lys  
545 550 555 560

Asp Lys Phe Ile Ser Met Pro Ser Pro Phe Lys Met Ile Glu Tyr Glu  
565 570 575

Trp Phe Leu Gly Glu Val Lys Thr Ala Leu Leu Leu Leu Asp Trp Ile  
580 585 590

Asn Glu Val Pro Ala Asp Asp Ile Cys Lys Lys Tyr Gly Ile Gly Glu

595                      600                      605  
 Gly Asp Ile Arg Met Phe Ser Glu Thr Ala Val Trp Leu Met His Ala  
 610                      615                      620  
  
 Thr Ser Arg Leu Ser Gly Leu Leu Lys Val Ser Glu Ala Ser Glu Lys  
 625                      630                      635                      640  
 Ser Lys Glu Leu Glu Lys Arg Leu Ser Tyr Gly Ile Asn Ser Glu Leu  
 645                      650                      655  
 Val Asn Ile Val Ala Leu Lys Gly Ile Gly Arg Val Arg Ala Arg Lys  
 660                      665                      670  
 Ile Tyr Glu Asn Gly Tyr Arg Ser Ile Asp Asp Leu Lys Lys Ala Asp  
 675                      680                      685  
  
 Pro Leu Lys Leu Ser Lys Ile Val Gly Ser Lys Ile Ser Gln Lys Ile  
 690                      695                      700  
 Leu Lys Gln Leu Asp Ile Asp Val Asp Ile Ser Glu Ile Lys Glu Lys  
 705                      710                      715                      720  
 Asp Ser Asp Thr Val Pro Glu Pro Glu Ser Ser Gln Lys Thr Ile Ser  
 725                      730                      735  
 Asp Phe Thr

<210> 44

<211> 733

<212> PRT

<213> Methanohalobium evestigatum

<400> 44

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

50	55	60	
Ser Asn Gly Gly Lys Cys Leu Tyr Ile Val Pro Leu Arg Ala Leu Ala			
65	70	75	80
Ser Glu Lys Phe Glu Arg Phe Lys Gln Phe Ser Ser Ile Gly Val Asn			
	85	90	95
Ile Gly Ile Ser Thr Gly Asp Phe Asp Ser Thr Asp Glu Trp Leu Gly			
	100	105	110
Ser Asn Asp Ile Ile Val Ala Thr Ser Glu Lys Ala Asp Ser Leu Leu			
	115	120	125
Arg Asn Glu Thr Ser Trp Met Lys Asp Ile Thr Thr Ile Val Val Asp			
	130	135	140
Glu Ile His Leu Leu Asp Ser Ala Asp Arg Gly Pro Thr Leu Glu Ile			
145	150	155	160
Thr Ile Ala Lys Leu Leu Arg Leu Asn Pro Asn Ser Gln Ile Ile Gly			
	165	170	175
Leu Ser Ala Thr Ile Gly Asn Ala Glu Glu Ile Ala Gly Trp Leu Asp			
	180	185	190
Ala Glu Leu Val Gln Ser Gln Trp Arg Pro Ile Glu Leu Tyr Glu Gly			
	195	200	205
Val Phe Leu Glu Asp Asn Ile Asn Phe Lys Gln Ser Gln Lys Pro Ile			
	210	215	220
Lys Asn Ile Val Lys Asp Thr Ala Val Asn Leu Val Leu Asp Thr Ile			
225	230	235	240
Asp Glu Asn Gly Gln Cys Leu Val Phe Glu Ser Ser Arg Arg Asn Cys			
	245	250	255
Ala Gly Phe Ala Lys Lys Ala Lys Ser Lys Val Gly Lys Ser Leu Asp			
	260	265	270
Lys Gly Leu Leu Ala Glu Leu Asn Asn Ile Ala Glu Glu Val Leu Glu			
	275	280	285
Thr Ser Asp Thr Glu Thr Thr Lys Glu Leu Ala Ser Cys Ile Lys Arg			
	290	295	300

Gly Thr Ala Phe His His Ala Gly Leu Asn Ser Ala Gln Arg Lys Ile  
305 310 315 320

Val Glu Asp Asn Phe Arg Asn Asn Lys Ile Lys Val Ile Ser Ser Thr  
325 330 335

Pro Thr Leu Ala Ala Gly Leu Asn Leu Pro Ala Arg Arg Val Ile Val  
340 345 350

Arg Asn Tyr Lys Arg Tyr Asp Pro Asn Phe Gly Met Gln Pro Ile Pro  
355 360 365

Val Leu Asp Tyr Lys Gln Met Ala Gly Arg Ala Gly Arg Pro Ser Leu  
370 375 380

Asp Pro Tyr Gly Glu Ser Val Leu Ile Ser His Thr Tyr Asn Glu Phe  
385 390 395 400

Thr Asp Leu Leu Asp Arg Tyr Ile Asp Ala Glu Pro Glu Asp Ile Leu  
405 410 415

Ser Lys Leu Gly Thr Glu Asn Ala Leu Arg Thr His Val Leu Ser Thr  
420 425 430

Ile Val Asn Gly Phe Ala Thr Thr Arg Gln Gly Met Val Asp Phe Met  
435 440 445

Gly Ser Ser Phe Phe Ala Tyr Gln Gln Gln Lys Trp Ser Leu Ile Asp  
450 455 460

Val Val Asp Asp Cys Ile Glu Phe Leu Gln Asp Asn Glu Met Ile Lys  
465 470 475 480

Asp Asp Gly Glu Arg Leu Tyr Ala Thr Arg Leu Gly Gln Val Ile Ser  
485 490 495

Thr Leu Tyr Ile Asp Pro Leu Ser Gly Ala Ile Ile Ile Asp Lys Leu  
500 505 510

Lys Lys Ala Asp Lys Val Thr Asp Met Thr Met Leu His Ile Ile Cys  
515 520 525

Ser Thr Pro Asp Met Arg Gln Leu Tyr Leu Arg Ser Lys Glu Tyr Glu  
530 535 540

Lys Ile Asn Glu Tyr Val Met Thr His Ser Asp Glu Phe Val Glu Val



545                      550                      555                      560  
 Pro Asn Pro Phe Lys Ser Ile Glu Tyr Glu Trp Phe Leu Gly Glu Val  
                          565                      570                      575  
  
 Lys Thr Ala Leu Leu Ile Asn Glu Trp Ile Asp Glu Lys Thr Leu Asp  
                          580                      585                      590  
 Asp Ile Thr Ala Glu Phe Gly Val Gly Glu Gly Asp Ile Asn Ala Leu  
                          595                      600                      605  
 Ser Asp Ile Ser Glu Trp Leu Met His Ser Ala Val Asn Leu Ala Asn  
                          610                      615                      620  
 Leu Thr Asp Leu Asp Ala Asp Lys Ala Gln Glu Leu Glu Lys Arg Ile  
 625                      630                      635                      640  
  
 His His Gly Val Asn Lys Asp Leu Ile Gln Leu Val Ser Ile Ser Asn  
                          645                      650                      655  
 Ile Gly Arg Val Arg Ala Arg Lys Leu Tyr Glu Ala Gly Ile Gln Ser  
                          660                      665                      670  
 Val Ser Asp Ile Lys Asn Thr Lys Leu His Ile Leu Ser Asn Tyr Leu  
                          675                      680                      685  
 Gly Arg Lys Thr Ala Tyr Lys Val Leu Glu Gln Leu Gly Val Glu Pro  
                          690                      695                      700  
  
 Glu Asp Asn Gln Gln Ile Asp Glu Glu Pro Glu Ser Ile Lys Ser Tyr  
 705                      710                      715                      720  
 Ser Gly Asn Asp Gln Gly Gln Lys Thr Phe Asn Asp Phe  
                          725                      730  
  
 <210> 45  
 <211> 747  
 <212> PRT  
 <213> Methanococcus maripaludis  
 <400> 45  
 Met His Val Leu Asp Leu Leu Lys Glu Asn Lys Ile Thr Glu Leu Arg  
 1                      5                      10                      15  
 Pro Pro Gln Lys Lys Val Ile Asp Glu Gly Leu Phe Asp Lys Thr Lys

20	25	30	
Asn Phe Leu Ile Cys Ile Pro Thr Ala Ser Gly Lys Thr Leu Ile Gly			
35	40	45	
Glu Met Ala Leu Leu Asn His Ile Leu Asp Glu Asn Lys Asn Leu Thr			
50	55	60	
Gly Lys Lys Gly Leu Phe Ile Val Pro Leu Lys Ala Leu Ala Asn Glu			
65	70	75	80
Lys Phe Asp Glu Phe Arg Glu Lys Tyr Glu Lys Tyr Gly Ile Lys Val			
85	90	95	
Gly Leu Ser Ile Gly Asp Phe Asp Thr Lys Glu Asn Leu Ser Lys Phe			
100	105	110	
His Ile Ile Ile Thr Thr Ser Glu Lys Leu Asp Ser Leu Met Arg His			
115	120	125	
Asn Val Glu Trp Ile Asn Asp Val Ser Leu Ala Val Ile Asp Glu Ile			
130	135	140	
His Leu Ile Gly Asp Asn Glu Arg Gly Gly Thr Leu Glu Val Ile Leu			
145	150	155	160
Thr Lys Leu Lys Asn Leu Asn Ala Gln Ile Val Gly Leu Ser Ala Thr			
165	170	175	
Ile Gly Asn Pro Glu Glu Leu Ser Asn Trp Leu Asn Ala Lys Leu Ile			
180	185	190	
Val Asp Gly Trp Arg Pro Val Glu Leu Lys Lys Gly Ile Tyr Phe Glu			
195	200	205	
Asn Glu Leu Glu Phe Leu Lys Asn Pro Ala Lys Lys Ile Lys Gln Val			
210	215	220	
Ser Arg Asn Asn Leu Thr Asp Leu Ile Val Asp Ser Val Glu Glu Lys			
225	230	235	240
Gly Ser Cys Leu Ile Phe Cys Asn Ser Lys Arg Asn Ala Val Gly Glu			
245	250	255	
Ala Lys Lys His Asn Leu Ala Lys Tyr Leu Thr Arg Thr Glu Gln His			
260	265	270	

Glu Leu Asn Lys Leu Ser Glu Glu Ile Leu Ser Ile Leu Asp Arg Pro

275 280 285

Val Glu Thr Cys Lys Ala Leu Ser Lys Cys Ile Gln Asn Gly Val Ala

290 295 300

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

305 310 315 320

Gly Phe Arg Asn Arg Leu Ile Lys Val Ile Cys Cys Thr Pro Thr Leu

325 330 335

Ser Ala Gly Leu Asn Leu Pro Cys Arg Arg Ala Ile Val Arg Asp Ile

340 345 350

Lys Arg Tyr Ser Gln Asn Gly Leu Val Asp Ile Pro Arg Met Glu Ile

355 360 365

Gln Gln Cys Ile Gly Arg Ala Gly Arg Pro Gly Leu Asp Pro Tyr Gly

370 375 380

Glu Gly Ile Ile Tyr Ile Lys Asn Glu Arg Asp Ala Glu Lys Ala Tyr

385 390 395 400

Glu Ile Leu Thr Gly Ser Val Glu Asn Ile Tyr Ser Lys Leu Ala Asn

405 410 415

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

420 425 430

Ile Lys Asp Gly Gln Asn Leu Val Asn Phe Met Lys Asn Thr Phe Tyr

435 440 445

Ala His Gln Phe Gly Asn Ile Gly Ala Val Leu Leu Asn Val Ser Glu

450 455 460

Val Val Glu Phe Leu Glu Lys Asn Lys Phe Leu Glu Thr Thr Ile His

465 470 475 480

Lys Lys Thr Glu Asn Lys Val Arg Glu Leu Ser Phe Asp Ser Ser Asn

485 490 495

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

500 505 510

Asn Ser Asn Ile Glu Phe Arg Ser Thr Lys Leu Gly Lys Arg Ile Ser

515 520 525  
Glu Leu Tyr Ile Asp Pro Met Ser Ser Glu Ile Ile Ile Glu Glu Leu

530 535 540  
His Glu Leu Lys Lys Lys Cys Asp Gln Leu Asp Arg Ser Lys Ile Asp  
545 550 555 560

Gln Tyr Leu Phe Tyr Leu Ile Ser Lys Thr Asn Glu Met Arg Pro Leu  
565 570 575

Leu Arg Ile Arg Pro Asn Glu Glu Leu Asp Leu Ile Leu Glu Met Asp  
580 585 590

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

595 600 605  
Asn Ser Lys Met Phe Cys Asp Trp Val Ser Glu Ile Pro Glu Glu Ile

610 615 620  
Ile Leu Glu Lys Tyr Gly Val Glu Pro Gly Ile Leu Arg Tyr Lys Val

625 630 635 640  
Glu Gln Ala Lys Trp Met Ile Tyr Ser Thr Lys Glu Ile Ala Lys Leu

645 650 655  
Ile His Leu Asp Asn Ser Glu Ile Tyr Lys Ser Leu Leu Lys Met Glu

660 665 670  
Val Arg Ile Glu Tyr Gly Ala Lys Glu Glu Leu Ile Glu Leu Leu Asn

675 680 685  
Val Lys Asn Val Gly Arg Ile Arg Ser Arg Lys Leu Tyr Asp Ala Gly

690 695 700  
Ile Arg Ser Lys Ile Glu Ile Asn Lys Asn Pro Glu Lys Ile Leu Glu

705 710 715 720  
Leu Phe Gly Glu Lys Ile Gly Lys Lys Ile Leu Gly Glu His Gly Met

725 730 735  
Lys Tyr Gly Gln Gln Thr Leu Leu Asn Phe Asn

740 745

<210> 46

<211> 799

<212> PRT

<213> Natrialba magadii

<400> 46

Met Asn Val Glu Glu Leu Ser Gly Leu Pro Pro Gly Ala Arg Ser His

1 5 10 15

Phe Gln Glu Gln Gly Ile Glu Glu Leu Tyr Pro Pro Gln Ala Glu Ala

20 25 30

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

35 40 45

Thr Ala Ser Gly Lys Thr Met Ile Ala Ala Leu Ser Met Leu Ser Ala

50 55 60

Val Gln Arg Gly Gly Lys Ala Leu Tyr Ile Val Pro Leu Arg Ala Leu

65 70 75 80

Ala Ser Glu Lys Lys Ala Glu Phe Asp Ala Tyr Glu Glu Phe Gly Val

85 90 95

Thr Thr Gly Val Ala Thr Gly Asn Tyr Glu Ser Thr Ser Glu Trp Leu

100 105 110

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

115 120 125

Val Arg Asn Gly Ala Asp Trp Leu Ser Asp Leu Thr Cys Val Val Ser

130 135 140

Asp Glu Val His Leu Ile Asp Asp Arg Asn Arg Gly Pro Thr Leu Glu

145 150 155 160

Val Thr Leu Ala Lys Leu Arg Arg Leu Asn Pro Gln Leu Gln Val Val

165 170 175

Ala Leu Ser Ala Thr Val Gly Asn Ala Asp Glu Leu Ala Asp Trp Leu

180 185 190

Asp Ala Glu Leu Val Asp Thr Asp Trp Arg Pro Ile Asp Leu Gln Met

195 200 205

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

210 215 220

Glu Val Pro Val Glu Ala Gly Glu Lys Gln Glu Ala Ala Leu Val Arg

225 230 235 240

Asp Ile Leu Gln Glu Gly Gly Ser Ser Leu Val Phe Val Asn Ser Arg

245 250 255

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

260 265 270

Glu Leu Thr Ala Gly Glu Gln Asn Asp Leu Ala Ala Leu Ala Thr Glu

275 280 285

Ile Arg Glu Asp Ser Asp Thr Glu Thr Ser Gln Asp Leu Ala Asp Cys

290 295 300

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

305 310 315 320

Arg Ser Leu Val Glu Asp Ala Phe Arg Asp Arg Leu Leu Lys Val Ile

325 330 335

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

340 345 350

Val Ile Val Arg Asp Trp Arg Arg Phe Asp Pro Ser Ala Gly Gly Met

355 360 365

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

370 375 380

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

385 390 395 400

His Asp Glu Ser Gln Glu Leu Phe Asp Arg Tyr Val Trp Ala Asp Pro

405 410 415

Glu Pro Val Arg Ser Lys Leu Ala Ala Glu Pro Ala Leu Arg Thr His

420 425 430

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

435 440 445

Leu Glu Phe Leu Glu Ala Thr Leu Tyr Ala Ser Gln Ser Ser Glu Gly

450 455 460

Gly Arg Leu Glu Arg Val Thr Asp Asp Val Leu Ser Tyr Leu Glu Arg

465                      470                      475                      480  
 Asn Asp Phe Ile Glu Arg Ser Gly Gly Pro Glu Asp Thr Leu Asn Ser  
  
                          485                      490                      495  
 Glu Ala Asp Ala Ala Ser Ala Phe Thr Ser Ala Ala Asp Leu Ala Asp  
                          500                      505                      510  
 Ser Asp Gly Gly Asp Ser Gly Gly Thr Thr Gly Gln Glu Glu Asp Leu  
                          515                      520                      525  
 Glu Ala Thr Ser Leu Gly His Thr Val Ser Arg Leu Tyr Leu Asp Pro  
                          530                      535                      540  
 Met Ser Ala Ala Glu Ile Val His Gly Leu Glu Asp Ala Asp Glu Arg  
  
 545                      550                      555                      560  
 Pro Thr Ala Leu Gly Leu Tyr Gln Leu Val Ser Arg Thr Pro Asp Met  
                          565                      570                      575  
 Tyr Glu Leu Tyr Leu Arg Ser Gly Glu Asp Glu Lys Phe Gly Glu Leu  
                          580                      585                      590  
 Tyr Tyr Glu Arg Glu Arg Glu Leu Leu Gly Asp Ala Pro Ser Glu Phe  
                          595                      600                      605  
 Glu Glu Glu Arg Phe Glu Asp Trp Leu Ala Ala Leu Lys Thr Gly Lys  
  
 610                      615                      620  
 Leu Leu Glu Asp Trp Ala Thr Glu Asp Asp Glu Glu Gln Ile Thr Glu  
 625                      630                      635                      640  
 Arg Tyr Lys Ile Gly Pro Gly Asp Leu Arg Gly Lys Val Asp Thr Ala  
                          645                      650                      655  
 Glu Trp Leu Leu Gly Ala Ala Glu Ser Leu Ala Ser Glu Ile Asp Ser  
                          660                      665                      670  
 Glu Trp Ala Val Ala Val Arg Glu Ala Arg Ala Arg Val Glu His Gly  
  
                          675                      680                      685  
 Val Gly Glu Glu Leu Leu Glu Leu Val Ser Val Ser Gly Ile Gly Arg  
                          690                      695                      700  
 Lys Arg Ala Arg Arg Leu Tyr Ala Ala Gly Ile Glu Glu Pro Ala Ala  
 705                      710                      715                      720

Leu Arg Ser Ala Asp Lys Gly Val Ile Leu His Val Leu Lys Gly Glu  
 725 730 735  
 Lys Thr Ala Glu Asn Ile Leu Glu Asn Ala Gly Arg Glu Glu Pro Ser  
 740 745 750  
 Met Asp Gly Val Glu Pro Ile Pro Val Glu Gly Gly Ser Gly Ser Gly  
 755 760 765  
 Ser Ser Asn Ser Ser Gly Ser Ser Glu Pro Asn Ala Asp Ala Asn Ala  
 770 775 780  
 Thr Glu Asp Asp Ala Asp Asp Asn Gln Ser Ser Leu Gly Asp Phe  
 785 790 795  
 <210> 47  
 <211> 723  
 <212> PRT  
 <213> Methanoregula boonei  
  
 <400> 47  
 Met Gln Ile Gln Asp Leu Ala Ile Pro Glu Pro Leu Arg Gln Gln Tyr  
 1 5 10 15  
 Leu Gly Leu Gly Ile Arg Glu Leu Tyr Pro Pro Gln Ala Ala Cys Val  
 20 25 30  
 Glu Arg Gly Leu Leu Asp Gly Lys Asn Leu Leu Val Ala Ile Pro Thr  
 35 40 45  
 Ala Ser Gly Lys Thr Leu Ile Ala Glu Met Ala Met His Arg His Ile  
 50 55 60  
  
 Ala Asn Gly Gly Lys Cys Leu Tyr Ile Val Pro Leu Lys Ala Leu Ala  
 65 70 75 80  
 Ser Glu Lys Tyr Glu Glu Phe Gly Asn Lys Gly Val Lys Val Gly Leu  
 85 90 95  
 Ser Thr Gly Asp Leu Asp Arg Arg Asp Asp Ala Leu Gly Lys Asn Asp  
 100 105 110  
 Ile Ile Val Ala Thr Ser Glu Lys Val Asp Ser Leu Leu Arg Asn Gly  
 115 120 125



Ala Arg Trp Ile Pro Asp Ile Thr Leu Val Val Ile Asp Glu Ile His  
130 135 140

Leu Ile Asp Ser Pro Asp Arg Gly Pro Thr Leu Glu Met Val Ile Ala  
145 150 155 160

Lys Met Arg Ser Lys Asn Pro Gly Met Gln Leu Ile Gly Leu Ser Ala  
165 170 175

Thr Ile Gly Asn Pro Lys Val Leu Ala Gly Trp Leu Asp Ala Glu Leu  
180 185 190

Val Thr Ser Ser Trp Arg Pro Val Asp Leu Arg Gln Gly Val Phe Tyr  
195 200 205

Asp Asn Arg Ile Gln Phe Ala Glu Arg Met Arg Pro Val Lys Gln Val  
210 215 220

Ser Lys Asn Tyr Asp Asp Leu Asn Leu Cys Leu Asp Thr Ile Ala Glu  
225 230 235 240

Gly Gly Gln Cys Leu Val Phe Val Ser Ser Arg Arg Asn Ala Glu Ala  
245 250 255

Phe Ala Lys Arg Ala Ala Gly Ala Ile Lys Ser Glu Asp Ala Ala Leu  
260 265 270

Ala Ala Cys Ala Glu Arg Leu Leu Glu Gly Thr Pro Thr Glu Met Val  
275 280 285

Lys Thr Leu Ala Ala Cys Val Ala Lys Gly Ala Ala Phe His His Ala  
290 295 300

Gly Leu Ser Arg Lys Glu Arg Ser Ile Val Glu Glu Ala Phe Arg Lys  
305 310 315 320

Asn Leu Leu Lys Cys Ile Ser Ser Thr Pro Thr Leu Ala Ala Gly Leu  
325 330 335

Asn Leu Pro Ala Arg Arg Val Ile Ile Arg Asp Tyr Leu Arg Phe Ser  
340 345 350

Ala Gly Glu Gly Met Gln Pro Ile Pro Val Ser Glu Tyr Arg Gln Met  
355 360 365

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

370                      375                      380  
 Leu Ile Ala Lys Glu Ala Glu Gln Val Pro Glu Leu Phe Glu Val Tyr  
 385                      390                      395                      400  
 Ile Glu Ala Glu Ala Glu Asp Val His Ser Arg Ile Ala Glu Pro Thr  
                     405                      410                      415  
 Ala Leu Tyr Thr His Val Leu Ser Leu Val Ala Ser Gly Phe Ala Gly  
                     420                      425                      430  
 Thr Arg Gly Glu Leu Thr Glu Phe Met Asn Arg Ser Phe Tyr Val His  
                     435                      440                      445  
  
 Glu His Lys Gln Gly Arg Leu Ile His Arg Ala Ile Asp Glu Ala Leu  
                     450                      455                      460  
 Gln Phe Leu Ile Thr Ala Glu Met Val Val Glu Val Gly Glu His Ile  
 465                      470                      475                      480  
 Gly Ala Thr Glu Leu Gly Thr Leu Val Ser Arg Met Tyr Ile Asp Pro  
                     485                      490                      495  
 Arg Ser Ala Phe Ala Ile Val Thr Thr Leu Arg Glu Gln Glu Lys Tyr  
                     500                      505                      510  
  
 Ala Asp Leu Gly Leu Ile Gln Leu Ile Cys Thr Thr Pro Asp Met Pro  
                     515                      520                      525  
 Thr Leu Tyr Ala Lys Asn Ala Asp Leu Pro Ala Leu Ser Arg Met Leu  
                     530                      535                      540  
 Glu Val Arg Gly Ala Asp Ile Trp Leu Pro Pro Pro Leu Asp Asp Asp  
 545                      550                      555                      560  
 Ala Ala Glu Thr Tyr Tyr Arg Ala Val Lys Thr Ala Met Leu Leu Ser  
                     565                      570                      575  
  
 Asp Trp Thr Asp Glu Leu Ser Glu Glu Lys Ile Cys Glu Arg Tyr Gly  
                     580                      585                      590  
 Val Gly Pro Gly Asp Val Phe Gly Met Val Glu Asn Ile Asn Trp Leu  
                     595                      600                      605  
 Leu His Ala Thr Ser Gln Leu Ala Arg Met Phe Val Pro Lys Phe Tyr  
                     610                      615                      620

Gly Gln Ile Ala Asp Cys Glu Ile Cys Met Lys Asn Gly Ile Arg Arg  
625 630 635 640

Glu Leu Leu Pro Leu Val Arg Leu Arg Gly Ile Gly Arg Val Arg Ala  
645 650 655

Arg Arg Leu Phe Asn Asn Gly Ile Thr Ser Pro Glu Glu Leu Ser Arg  
660 665 670

His Lys Lys Glu Asp Leu Val Lys Ile Leu Gly Ser Gly Ile Ala Glu  
675 680 685

Gln Val Leu Glu Gln Leu His Pro Ser Lys Asp Thr Gly Lys Lys Glu  
690 695 700

Pro Pro Ser Gly Asp Lys Asn Thr Asn Pro Gly Gln Ser Thr Leu Phe  
705 710 715 720

His Phe Gly

<210> 48

<211> 681

<212> PRT

<213> Ferroplasma acidarmanus

<400> 48

Met Lys Leu Ser Glu Ile Thr Pro Ser Glu Phe Leu Lys Val Thr Asp  
1 5 10 15

Asn Asn Asp Phe Thr Leu Tyr Glu His Gln Glu Glu Ala Val Ala Lys  
20 25 30

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

35 40 45

Lys Thr Leu Ile Gly Tyr Ile Ser Ile Tyr Asp Thr Tyr Leu Lys Gly  
50 55 60

Lys Lys Ser Met Tyr Ile Val Pro Leu Arg Ser Leu Ala Met Glu Lys  
65 70 75 80

Phe Ser Glu Leu Leu Ser Leu Arg Asn Leu Gly Val Lys Val Thr Met  
85 90 95

Ser Ile Gly Asp Tyr Asp Val Pro Pro Ser Phe Val Lys Asn Tyr Asp

100 105 110

Val Ile Ile Ala Thr Ser Glu Arg Ala Asp Ser Met Leu His Arg Asp

115 120 125

Pro Asp Ile Leu Asn Tyr Phe Gly Leu Val Ile Ile Asp Glu Ile His

130 135 140

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

145 150 155 160

Ser Leu Leu Tyr Leu Asn Pro Glu Ile Leu Leu Leu Gly Leu Ser Ala

165 170 175

Thr Val Ser Asn Ile Gln Glu Ile Ala Glu Trp Met Asn Ala Glu Thr

180 185 190

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

195 200 205

Lys Gly Asn Leu Ile Thr Asp Gly Glu Lys Lys His Leu Gly Arg Asp

210 215 220

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

225 230 235 240

Leu Val Phe Arg Asn Ser Arg Arg Asn Ala Glu Lys Tyr Ala Gln Ser

245 250 255

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

260 265 270

Pro Pro Asp Leu Phe Asn Glu Ala Gln Ala Asn Met Val Ala His Gly

275 280 285

Val Met Phe His His Ala Gly Leu Ser Asn Asp Gln Arg Thr Met Ile

290 295 300

Glu Lys Leu Phe Lys Gln Gly Tyr Ile Lys Ile Leu Thr Ala Thr Pro

305 310 315 320

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

325 330 335

Asp Ile Thr Arg Phe Ser Asp Gly Tyr Ser Lys Pro Ile Ser Gly Ile

340                      345                      350  
 Glu Ile Gln Gln Met Ile Gly Arg Ala Gly Arg Pro Lys Tyr Asp Lys  
  
 355                      360                      365  
 Lys Gly Tyr Gly Tyr Ile Tyr Ala Ala Ser Pro Gly Met Leu Arg Val  
 370                      375                      380  
 Ala Glu Gly Tyr Leu Thr Gly Glu Leu Glu Pro Val Ile Ser Arg Met  
 385                      390                      395                      400  
 Asp Ser Asn Ser Leu Ile Arg Phe Asn Val Leu Ala Leu Ile Ser Ser  
 405                      410                      415  
 Gly Ile Ala Thr Asp Leu Lys Gly Ile Gln Asp Phe Tyr Gly Lys Thr  
  
 420                      425                      430  
 Leu Leu Ala Ala Gln Asn Asp Ile Asp Gly Tyr Glu Leu Ala Phe Glu  
 435                      440                      445  
 Ser Ala Leu Tyr Phe Leu Lys Asp Asn Asp Phe Ile Thr Glu Glu Asn  
 450                      455                      460  
 Asp Ile Tyr Ser Ala Thr Lys Phe Gly Arg Leu Thr Ser Asp Leu Tyr  
 465                      470                      475                      480  
 Ile Asp Pro Val Ser Ser Leu Ile Leu Lys Lys Cys Leu Asp Leu Glu  
  
 485                      490                      495  
 Phe Ser Glu Glu Leu Tyr Leu Tyr Tyr Ile Ser Lys Thr Pro Asp Met  
 500                      505                      510  
 Leu Thr Phe Asn Tyr Arg Ala Ser Asp Tyr Glu Tyr Leu Glu Glu Phe  
 515                      520                      525  
 Leu Asp Arg His Asn Ile Ser Asp Phe Ser Glu Glu Ser Met Gly Ala  
 530                      535                      540  
 Ala Lys Thr Ala Ile Ile Leu Asn Glu Trp Ile Asn Glu Val Pro Ile  
  
 545                      550                      555                      560  
 Asn Thr Ile Ala Glu Thr Phe Gly Ile Gly Pro Gly Asp Ile Gln Ala  
 565                      570                      575  
 Lys Ala Ser Ser Ala Asp Trp Ile Ser Tyr Ser Leu Tyr Arg Leu Gly  
 580                      585                      590

Ser Met Phe Asp Lys Glu Asn Glu Asn Asn Leu Leu His Leu Asn Ile  
595 600 605

Arg Ile Lys Glu Gly Val Lys Glu Glu Ile Ile Arg Ile Ile Glu Ile

610 615 620

Pro Gln Val Gly Arg Val Arg Gly Arg Arg Leu Tyr Asn Asn Gly Phe  
625 630 635 640

Lys Ser Ile Asp Asp Ile Ala Asn Ala Arg Val Glu Asp Ile Ser Arg  
645 650 655

Ile Phe Gly Phe Ser Thr Lys Leu Ala Lys Asp Ile Ile Glu Asn Ala  
660 665 670

Gly Lys Leu Asn Asn Arg Tyr Tyr Arg  
675 680

<210> 49

<211> 696

<212> PRT

<213> Methanocaldococcus fervens

<400> 49

Met Pro Thr Asn Lys Ile Leu Glu Ile Leu Lys Asp Phe Gly Ile Glu  
1 5 10 15

Glu Leu Arg Pro Pro Gln Lys Lys Ala Leu Glu Lys Gly Leu Leu Asp  
20 25 30

Lys Asn Lys Asn Phe Leu Ile Ser Ile Pro Thr Ala Ser Gly Lys Thr  
35 40 45

Leu Ile Gly Glu Met Ala Leu Ile Asn His Leu Leu Asp Glu Asn Lys

50 55 60

Asn Pro Thr Asn Lys Lys Gly Ile Phe Ile Val Pro Leu Lys Ala Leu  
65 70 75 80

Ala Ser Glu Lys Tyr Glu Glu Phe Lys Asn Lys Tyr Glu Arg Tyr Gly  
85 90 95

Leu Arg Val Ala Leu Ser Ile Gly Asp Tyr Asp Glu Asp Glu Asp Leu  
100 105 110

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

115 120 125  
Trp Arg His Lys Ile Asp Trp Ile Asp Asp Val Ser Val Val Val Val

130 135 140  
Asp Glu Ile His Leu Ile Asn Asp Glu Ser Arg Gly Gly Thr Leu Glu

145 150 155 160  
Ile Leu Leu Thr Lys Leu Lys Lys Phe Asn Ile Gln Ile Ile Gly Leu

165 170 175  
Ser Ala Thr Ile Gly Asn Pro Glu Glu Leu Ala Asn Trp Leu Asn Ala

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

195 200 205  
Tyr Lys Asn Gly Ile Ile Glu Phe Ile Asn Gly Glu Asn Arg Glu Ile

210 215 220  
Lys Ala Ile Asn Asn Asn Asp Ile Tyr Asn Leu Val Val Asp Cys Val

225 230 235 240  
Lys Asp Gly Gly Cys Cys Ile Val Phe Cys Asn Thr Lys Arg Gly Ala

245 250 255  
Val Asn Glu Ala Lys Lys Leu Asn Leu Lys Lys Phe Leu Thr Asn Glu

260 265 270  
Glu Lys Arg Lys Leu Lys Glu Val Ala Glu Glu Ile Leu Ser Ile Leu

275 280 285  
Glu Pro Pro Thr Glu Met Cys Lys Thr Leu Ala Glu Cys Ile Leu Asn

290 295 300  
Gly Ser Ala Phe His His Ala Gly Leu Thr Tyr Gln His Arg Lys Ile

305 310 315 320  
Val Glu Asp Ala Phe Arg Asn Lys Leu Ile Lys Val Ile Cys Cys Thr

325 330 335  
Pro Thr Leu Ser Val Gly Leu Asn Leu Pro Cys Arg Arg Ala Ile Val

340 345 350  
Lys Asp Leu Thr Arg Tyr Thr Asn Arg Gly Met Arg Tyr Ile Pro Ile

355                      360                      365  
 Met Glu Ile Gln Gln Cys Ile Gly Arg Ala Gly Arg Leu Gly Leu Asp  
  
 370                      375                      380  
 Pro Tyr Gly Glu Gly Ile Ile Val Ala Lys Asn Asp Arg Asp Tyr Leu  
 385                      390                      395                      400  
 Arg Ser Tyr Gln Val Leu Thr Gln Lys Pro Glu Pro Ile Tyr Ser Lys  
 405                      410                      415  
 Leu Ser Asn Gln Ala Val Leu Arg Thr Gln Leu Leu Gly Leu Ile Ala  
 420                      425                      430  
 Thr Ile Glu Ile Arg Asp Glu Tyr Asp Leu Glu Trp Phe Ile Arg Asn  
  
 435                      440                      445  
 Thr Phe Tyr Ala Tyr Gln Tyr Gly Asn Leu Arg Glu Val Ala Lys Asn  
 450                      455                      460  
 Ile Asn Glu Val Ile Arg Phe Leu Glu Glu Lys Glu Phe Met Ile Asp  
 465                      470                      475                      480  
 Phe Ile Pro Thr Glu Leu Gly Lys Arg Val Ala Glu Leu Tyr Ile Asp  
 485                      490                      495  
 Pro Leu Ser Ala Lys Tyr Met Ile Asp Gly Leu Asn Glu Met Glu Asn  
  
 500                      505                      510  
 Glu Asp Asp Ile Tyr Tyr Leu Tyr Leu Ile Ser Lys Thr Leu Glu Met  
 515                      520                      525  
 Met Pro Asn Leu Arg Val Tyr Lys Ser Glu Glu Leu Asn Leu Ile Asp  
 530                      535                      540  
 Glu Met Glu Asn Leu Gly Ile Lys Ser Phe Glu Ile Glu Asp Leu Glu  
 545                      550                      555                      560  
 Ala Phe Lys Thr Ala Lys Met Leu Tyr Asp Trp Ile Ser Glu Val Pro  
  
 565                      570                      575  
 Glu Asp Glu Ile Leu Lys Lys Tyr Lys Ile Glu Pro Gly Ile Leu Arg  
 580                      585                      590  
 Tyr Lys Val Glu Asn Ala Val Trp Leu Met His Ala Leu Lys Glu Met  
 595                      600                      605



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

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

625 630 635 640

Lys Tyr Ile Gly Arg Val Arg Ala Arg Lys Leu Tyr Asn Ala Gly Ile  
645 650 655

Arg Asn Val Glu Asp Ile Ile Asn Asn Pro Ser Lys Val Ala Ser Ile  
660 665 670

Ile Gly Glu Lys Ile Thr Lys Lys Ile Leu Glu Asp Leu Gly Ile Lys  
675 680 685

Phe Gly Gln Gln Lys Leu Ile Phe  
690 695

<210> 50  
<211> 729  
<212> PRT  
<213> Methanocaldococcus jannaschii  
<400> 50

Met Asp Lys Ile Leu Glu Ile Leu Lys Asp Phe Gly Ile Val Glu Leu  
1 5 10 15

Arg Pro Pro Gln Lys Lys Ala Leu Glu Arg Gly Leu Leu Asp Lys Asn  
20 25 30

Lys Asn Phe Leu Ile Ser Ile Pro Thr Ala Ser Gly Lys Thr Leu Ile  
35 40 45

Gly Glu Met Ala Leu Ile Asn His Leu Leu Asp Gly Asn Lys Asn Pro

50 55 60

Thr Asn Lys Lys Gly Ile Phe Ile Val Pro Leu Lys Ala Leu Ala Ser  
65 70 75 80

Glu Lys Tyr Glu Glu Phe Lys Ser Lys Tyr Glu Arg Tyr Gly Leu Arg  
85 90 95

Ile Ala Leu Ser Ile Gly Asp Tyr Asp Glu Asp Glu Asp Leu Ser Lys  
100 105 110

Tyr His Leu Ile Ile Thr Thr Ala Glu Lys Leu Asp Ser Leu Trp Arg

115 120 125  
His Lys Ile Asp Trp Ile Asn Asp Val Ser Val Val Val Val Asp Glu

130 135 140  
Ile His Leu Ile Asn Asp Glu Thr Arg Gly Gly Thr Leu Glu Ile Leu

145 150 155 160  
Leu Thr Lys Leu Lys Glu Phe Asn Val Gln Ile Ile Gly Leu Ser Ala

165 170 175  
Thr Ile Gly Asn Pro Asp Glu Leu Ala Glu Trp Leu Asn Ala Glu Leu

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

195 200 205  
Asn Glu Ala Ile Glu Phe Ile Asn Gly Glu Ile Arg Glu Ile Lys Ala

210 215 220  
Val Asp Asn Asn Asp Ile Tyr Asn Leu Val Val Asp Cys Val Lys Glu

225 230 235 240  
Gly Gly Cys Cys Leu Val Phe Cys Asn Thr Lys Arg Asn Ala Val Asn

245 250 255  
Glu Ala Lys Lys Leu Asn Leu Lys Lys Phe Leu Thr Glu Glu Glu Lys

260 265 270  
Ile Arg Leu Lys Glu Ile Ala Glu Glu Ile Leu Ser Ile Leu Glu Pro

275 280 285  
Pro Thr Glu Met Cys Lys Thr Leu Ala Glu Cys Ile Leu Asn Gly Ser

290 295 300  
Ala Phe His His Ala Gly Leu Thr Tyr Gln His Arg Lys Ile Val Glu

305 310 315 320  
Asp Ala Phe Arg Lys Arg Leu Ile Lys Val Ile Cys Cys Thr Pro Thr

325 330 335  
Leu Ser Ala Gly Leu Asn Leu Pro Cys Arg Arg Ala Ile Val Lys Asp

340 345 350  
Leu Thr Arg Phe Thr Asn Lys Gly Met Arg Tyr Ile Pro Ile Met Glu

355 360 365  
Ile Gln Gln Cys Ile Gly Arg Ala Gly Arg Pro Gly Leu Asp Pro Tyr

370 375 380  
Gly Glu Gly Ile Ile Val Ala Lys Asn Asp Arg Asp Tyr Leu Arg Ala  
385 390 395 400

Tyr Gln Ala Leu Thr Gln Lys Pro Glu Pro Ile Tyr Ser Lys Leu Ser  
405 410 415

Asn Gln Ala Val Leu Arg Thr Gln Leu Leu Gly Leu Ile Ala Thr Gly  
420 425 430

Glu Ile Arg Asp Glu Tyr Asp Leu Glu Trp Phe Ile Arg Asn Thr Phe

435 440 445  
Tyr Ala His Gln Tyr Gly Asn Leu Arg Glu Val Ala Lys Asn Ile Asn  
450 455 460

Glu Val Ile Arg Phe Leu Glu Glu Asn Glu Phe Ile Ile Asp Phe Met  
465 470 475 480

Pro Thr Glu Leu Gly Lys Arg Val Ser Glu Leu Tyr Ile Asp Pro Leu  
485 490 495

Ser Ala Lys Phe Ile Ile Asp Gly Leu Glu Glu Met Glu Asn Glu Glu

500 505 510  
Glu Ile Tyr Tyr Leu Tyr Leu Ile Ser Lys Thr Leu Glu Met Met Pro  
515 520 525

Asn Leu Arg Val Tyr Asn Ser Glu Glu Leu Asn Leu Ile Asp Glu Met  
530 535 540

Asp Ser Leu Gly Ile Lys Ser Phe Glu Ile Glu Asp Leu Glu Ala Phe  
545 550 555 560

Lys Thr Ala Lys Met Leu Tyr Asp Trp Ile Asn Glu Val Pro Glu Asp

565 570 575  
Glu Ile Leu Lys Arg Tyr Lys Ile Glu Pro Gly Ile Leu Arg Tyr Lys  
580 585 590

Val Glu Asn Ala Val Trp Ile Met His Ala Leu Lys Glu Ile Ala Lys  
595 600 605

Leu Ile Gly Lys Ser Ser Asp Ile Pro Glu Lys Leu Glu Ile Arg Leu  
 610 615 620  
 Glu Tyr Gly Ala Lys Glu Asp Ile Ile Glu Leu Leu Ser Ile Lys Tyr  
  
 625 630 635 640  
 Ile Gly Arg Val Arg Ala Arg Lys Leu Tyr Asn Ala Gly Ile Arg Ser  
 645 650 655  
 Ile Glu Asp Ile Ile Asn Asn Pro Ser Lys Val Ala Ser Ile Ile Gly  
 660 665 670  
 Glu Lys Ile Ala Lys Lys Ile Leu Asp Glu Leu Gly Val Lys Phe Gly  
 675 680 685  
 Gln Gln Lys Leu Ser Phe Ser Gly Gly Ser Ala Trp Ser His Pro Gln  
  
 690 695 700  
 Phe Glu Lys Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Ser Ala Trp  
 705 710 715 720  
 Ser His Pro Gln Phe Glu Lys Lys Leu  
 725  
 <210> 51  
 <211> 670  
 <212> PRT  
 <213> Methanocaldococcus infernus  
 <400> 51  
 Met Asp Glu Ile Leu Lys Phe Leu Gly Ile Lys Glu Leu Arg Pro Pro  
 1 5 10 15  
 Gln Lys Lys Ala Leu Glu Leu Gly Ile Leu Asp Lys Lys Lys Asn Phe  
  
 20 25 30  
 Leu Ile Ser Ile Pro Thr Gly Ala Gly Lys Thr Val Ile Ala Glu Met  
 35 40 45  
 Ala Leu Ile Asn His Leu Leu Leu Asp Lys Gly Lys Lys Gly Val Tyr  
 50 55 60  
 Ile Val Pro Leu Lys Ala Leu Ala Ser Glu Lys Tyr Glu Glu Phe Lys  
 65 70 75 80

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

85 90 95

Tyr Asp Glu Asp Glu Asp Leu Glu Asn Tyr Asp Leu Ile Ile Thr Thr

100 105 110

Ala Glu Lys Phe Asp Ser Leu Trp Arg His Gly Ile Lys Leu Ser Asp

115 120 125

Ile Ser Val Val Val Val Asp Glu Ile His Val Ile Gly Asp Ser Glu

130 135 140

Arg Gly Gly Thr Leu Glu Val Leu Leu Thr Lys Leu Lys Glu Leu Asp

145 150 155 160

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

165 170 175

Ser Glu Trp Leu Asn Ala Glu Leu Leu Leu Asp Asn Trp Arg Pro Val

180 185 190

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

195 200 205

Gly Glu Val Lys Glu Cys Gln Asp Ile Val Lys Glu Val Val Lys Asp

210 215 220

Asn Gly Ser Val Ile Ile Phe Cys Pro Thr Lys Lys Lys Ala Glu Asn

225 230 235 240

Arg Ala Leu Ser Leu Asp Leu Ser Asp Leu Leu Lys Lys Ser Glu Lys

245 250 255

Arg Lys Leu Glu Glu Ile Ser Glu Glu Leu Leu Ser Leu Phe Asp Pro

260 265 270

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

275 280 285

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

290 295 300

Lys Ala Phe Arg Glu Arg Ile Leu Lys Val Ile Cys Ser Thr Thr Thr

305 310 315 320

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

325                                      330                                      335  
 Leu Lys Arg Tyr Thr Arg Arg Gly Leu Thr Tyr Ile Pro Ile Met Glu  
  
 340                                      345                                      350  
 Val Gln Gln Cys Ile Gly Arg Ala Gly Arg Pro Gly Leu Asp Glu Tyr  
 355                                      360                                      365  
 Gly Glu Gly Ile Leu Val Ala Lys Asp Glu Arg Asp Tyr Leu Arg Ala  
 370                                      375                                      380  
 Leu Gln Cys Leu Thr Gln Lys Pro Glu Pro Ile Tyr Ser Lys Leu Ser  
 385                                      390                                      395                                      400  
 Asn Asp Ser Val Leu Arg Thr Gln Ile Leu Gly Leu Ile Ala Thr Arg  
  
 405                                      410                                      415  
 Tyr Val Leu Asp Glu Tyr Asp Leu Glu Glu Phe Ile Lys Asn Thr Phe  
 420                                      425                                      430  
 Tyr Ala Tyr Gln Tyr Lys Asn Leu Asp Glu Ile Lys Lys Lys Ile Lys  
 435                                      440                                      445  
 Glu Ile Ile Glu Phe Leu Glu Asp Cys Asn Phe Ile Lys Asn Phe Glu  
 450                                      455                                      460  
 Val Thr Pro Leu Gly Lys Lys Val Ser Asn Leu Tyr Leu Asp Pro Leu  
  
 465                                      470                                      475                                      480  
 Ser Ala Lys Ile Met Ile Asp Asn Ile Glu Val Lys Asp Asp Leu His  
 485                                      490                                      495  
 Leu Leu Tyr Ile Leu Cys Lys Cys Ile Glu Met Lys Pro Leu Leu Arg  
 500                                      505                                      510  
 Val Tyr Arg Lys Glu Glu Glu Glu Leu Ala Glu Glu Leu Leu Asn Tyr  
 515                                      520                                      525  
 Glu Ile Phe Ile Ser Tyr Glu Asn Leu Glu Glu Phe Lys Thr Ala Lys  
  
 530                                      535                                      540  
 Met Leu Tyr Asp Trp Ile Asn Glu Val Pro Glu Asp Glu Ile Leu Lys  
 545                                      550                                      555                                      560  
 Thr Tyr Lys Val Glu Pro Gly Ile Leu Arg Tyr Lys Val Glu Val Ala  
 565                                      570                                      575

Lys Trp Leu Ser Tyr Ser Leu Lys Glu Ile Ala Lys Ile Leu Asn Lys  
 580 585 590  
 Glu Val Pro Asn Leu Glu Leu Arg Leu Glu Tyr Gly Ala Lys Glu Glu  
 595 600 605  
 Leu Leu Glu Leu Leu Lys Ile Lys Tyr Ile Gly Arg Val Arg Ala Arg  
 610 615 620  
 Lys Leu Tyr Ser Ala Gly Ile Arg Asn Arg Glu Asp Ile Ile Lys Asn  
 625 630 635 640  
 Pro Lys Lys Val Ala Asn Ile Leu Gly Glu Lys Ile Ser Lys Lys Ile  
 645 650 655  
 Phe Glu Glu Leu Gly Val Arg Tyr Gly Gln Gln Arg Leu Ile  
 660 665 670  
 <210> 52  
 <211> 799  
 <212> PRT  
 <213> Methanospirillum hungatei  
 <400> 52  
 Met Glu Ile Ala Ser Leu Pro Leu Pro Asp Ser Phe Ile Arg Ala Cys  
 1 5 10 15  
 His Ala Lys Gly Ile Arg Ser Leu Tyr Pro Pro Gln Ala Glu Cys Ile  
 20 25 30  
 Glu Lys Gly Leu Leu Glu Gly Lys Asn Leu Leu Ile Ser Ile Pro Thr  
 35 40 45  
 Ala Ser Gly Lys Thr Leu Leu Ala Glu Met Ala Met Trp Ser Arg Ile  
 50 55 60  
 Ala Ala Gly Gly Lys Cys Leu Tyr Ile Val Pro Leu Arg Ala Leu Ala  
 65 70 75 80  
 Ser Glu Lys Tyr Asp Glu Phe Ser Lys Lys Gly Val Ile Arg Val Gly  
 85 90 95  
 Ile Ala Thr Gly Asp Leu Asp Arg Thr Asp Ala Tyr Leu Gly Glu Asn  
 100 105 110

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



355 Ala Gly Arg Ala Gly Arg Pro His Leu Asp Pro Tyr Gly Glu Ala Val  
 370 375 380  
 Leu Leu Ala Lys Asp Ala Pro Ser Val Glu Arg Leu Phe Glu Thr Phe  
 385 390 395 400  
 Ile Asp Ala Glu Ala Glu Arg Val Asp Ser Gln Cys Val Asp Asp Ala  
 405 410 415  
 Ser Leu Cys Ala His Ile Leu Ser Leu Ile Ala Thr Gly Phe Ala His  
 420 425 430  
  
 Asp Gln Glu Ala Leu Ser Ser Phe Met Glu Arg Thr Phe Tyr Phe Phe  
 435 440 445  
 Gln His Pro Lys Thr Arg Ser Leu Pro Arg Leu Val Ala Asp Ala Ile  
 450 455 460  
 Arg Phe Leu Thr Thr Ala Gly Met Val Glu Glu Arg Glu Asn Thr Leu  
 465 470 475 480  
 Ser Ala Thr Arg Leu Gly Ser Leu Val Ser Arg Leu Tyr Leu Asn Pro  
 485 490 495  
  
 Cys Thr Ala Arg Leu Ile Leu Asp Ser Leu Lys Ser Cys Lys Thr Pro  
 500 505 510  
 Thr Leu Ile Gly Leu Leu His Val Ile Cys Val Ser Pro Asp Met Gln  
 515 520 525  
 Arg Leu Tyr Leu Lys Ala Ala Asp Thr Gln Leu Leu Arg Thr Phe Leu  
 530 535 540  
 Phe Lys His Lys Asp Asp Leu Ile Leu Pro Leu Pro Phe Glu Gln Glu  
 545 550 555 560  
  
 Glu Glu Glu Leu Trp Leu Ser Gly Leu Lys Thr Ala Leu Val Leu Thr  
 565 570 575  
 Asp Trp Ala Asp Glu Phe Ser Glu Gly Met Ile Glu Glu Arg Tyr Gly  
 580 585 590  
 Ile Gly Ala Gly Asp Leu Tyr Asn Ile Val Asp Ser Gly Lys Trp Leu  
 595 600 605

Leu His Gly Thr Glu Arg Leu Val Ser Val Glu Met Pro Glu Met Ser  
610 615 620

Gln Val Val Lys Thr Leu Ser Val Arg Val His His Gly Val Lys Ser  
625 630 635 640  
Glu Leu Leu Pro Leu Val Ala Leu Arg Asn Ile Gly Arg Val Arg Ala  
645 650 655  
Arg Thr Leu Tyr Asn Ala Gly Tyr Pro Asp Pro Glu Ala Val Ala Arg  
660 665 670  
Ala Gly Leu Ser Thr Ile Ala Arg Ile Ile Gly Glu Gly Ile Ala Arg  
675 680 685

Gln Val Ile Asp Glu Ile Thr Gly Val Lys Arg Ser Gly Ile His Ser  
690 695 700  
Ser Asp Asp Asp Tyr Gln Gln Lys Thr Pro Glu Leu Leu Thr Asp Ile  
705 710 715 720  
Pro Gly Ile Gly Lys Lys Met Ala Glu Lys Leu Gln Asn Ala Gly Ile  
725 730 735  
Ile Thr Val Ser Asp Leu Leu Thr Ala Asp Glu Val Leu Leu Ser Asp  
740 745 750

Val Leu Gly Ala Ala Arg Ala Arg Lys Val Leu Ala Phe Leu Ser Asn  
755 760 765  
Ser Glu Lys Glu Asn Ser Ser Ser Asp Lys Thr Glu Glu Ile Pro Asp  
770 775 780  
Thr Gln Lys Ile Arg Gly Gln Ser Ser Trp Glu Asp Phe Gly Cys  
785 790 795

<210> 53

<211> 702

<212> PRT

<213> Archaeoglobus fulgidus

<400> 53

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

1 5 10 15

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

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

Leu Ile Cys Arg Thr Pro Asp Met Glu Arg Leu Thr Val Arg Lys Thr

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

580 585 590  
 Ile Val Glu Thr Ala Glu Trp Leu Ser Asn Ala Met Asn Arg Ile Ala  
 595 600 605  
 Glu Glu Val Gly Asn Thr Ser Val Ser Gly Leu Thr Glu Arg Ile Lys  
 610 615 620  
 His Gly Val Lys Glu Glu Leu Leu Glu Leu Val Arg Ile Arg His Ile  
 625 630 635 640  
 Gly Arg Val Arg Ala Arg Lys Leu Tyr Asn Ala Gly Ile Arg Asn Ala

645 650 655  
 Glu Asp Ile Val Arg His Arg Glu Lys Val Ala Ser Leu Ile Gly Arg  
 660 665 670  
 Gly Ile Ala Glu Arg Val Val Glu Gly Ile Ser Val Lys Ser Leu Asn  
 675 680 685  
 Pro Glu Ser Ala Ala Ala Leu Glu His His His His His His  
 690 695 700

<210> 54

<211> 791

<212> PRT

<213> Haloterrigena turkmenica

<400> 54

Met Asn Leu Glu Glu Leu Thr Gly Leu Pro Pro Gly Ala Thr Asp His

1 5 10 15

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

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

Ala Ala Lys Arg Glu Glu Thr Leu Glu Ala Thr Ser Leu Gly His Thr  
515 520 525

Val Ser Arg Leu Tyr Leu Asp Pro Met Ser Ala Ala Glu Ile Val His  
530 535 540

Gly Leu Glu Arg Ala Asp Glu Arg Pro Thr Ala Leu Gly Leu Tyr Gln  
545 550 555 560

Leu Val Ser Arg Thr Pro Asp Met Tyr Glu Leu Tyr Leu Arg Ser Gly  
565 570 575

Glu Asp Glu Lys Phe Gly Glu Leu Phe Tyr Glu Arg Glu Thr Glu Leu  
580 585 590

Leu Gly Asp Ala Pro Ser Glu Tyr Glu Glu Asp Arg Phe Glu Asp Trp  
595 600 605

Leu Ala Ala Leu Lys Thr Gly Lys Leu Leu Glu Asp Trp Ala Asp Glu  
610 615 620

Thr Asp Glu Glu Thr Ile Thr Asp Arg Tyr Lys Ile Gly Pro Gly Asp  
625 630 635 640

Leu Arg Gly Lys Val Asp Thr Ala Glu Trp Leu Leu Gly Ala Ala Glu  
645 650 655

Ser Leu Ala Ala Glu Ile Asp Ser Glu Trp Thr Val Ala Val Arg Glu  
660 665 670

Ala Arg Ala Arg Val Glu His Gly Val Gly Glu Glu Leu Leu Glu Leu  
675 680 685

Val Ser Val Gly Gly Val Gly Arg Lys Arg Ala Arg Arg Leu Tyr Asp  
690 695 700

Ala Gly Ile Glu Glu Pro Ala Asp Leu Arg Ser Ala Asp Lys Gly Ile  
705 710 715 720

Val Leu Ser Val Leu Lys Gly Glu Lys Thr Ala Glu Asn Ile Leu Glu  
725 730 735

Asn Ala Gly Arg Glu Asp Pro Ser Met Asp Gly Val Glu Pro Ala Asp  
740 745 750

Gly Gly Pro Ala Val Gly Ala Ala Thr Asn Gly Ser Ser Gly Gly Ser



755			760			765									
Glu	Thr	Asp	Glu	Thr	Gly	Arg	Ala	Asp	Ala	Ala	Glu	Ser	Asp	Asp	Ser
770			775			780									
Gln	Ser	Ser	Leu	Gly	Asp	Phe									
785			790												
<210>			55												
<211>			752												
<212>			PRT												
<213>			Haladaptatus paucihalophilus												
<400>			55												
Met	Asn	Val	Ala	Asp	Leu	Thr	Gly	Leu	Pro	Asp	Gly	Val	Pro	Glu	His
1			5			10			15						
Phe	His	Ala	Gln	Gly	Ile	Glu	Glu	Leu	Tyr	Pro	Pro	Gln	Ala	Glu	Ala
20			25			30									
Val	Glu	Ala	Gly	Ile	Thr	Glu	Gly	Glu	Ser	Val	Val	Ala	Ser	Ile	Pro
35			40			45									
Thr	Ala	Ser	Gly	Lys	Thr	Phe	Ile	Ala	Glu	Leu	Ala	Met	Leu	Ser	Ser
50			55			60									
Val	Ala	Arg	Gly	Gly	Lys	Ala	Leu	Tyr	Ile	Val	Pro	Leu	Arg	Ala	Leu
65			70			75			80						
Ala	Ser	Glu	Lys	Lys	Glu	Glu	Phe	Glu	Glu	Phe	Glu	Gln	Tyr	Gly	Val
85			90			95									
Ser	Ile	Gly	Val	Ser	Thr	Gly	Asn	Tyr	Glu	Ser	Asp	Gly	Asp	Trp	Leu
100			105			110									
Ala	Ser	Arg	Asp	Ile	Ile	Val	Ala	Thr	Ser	Glu	Lys	Val	Asp	Ser	Leu
115			120			125									
Val	Arg	Asn	Gly	Ala	Lys	Trp	Ile	Asp	Asp	Leu	Ser	Cys	Val	Val	Ala
130			135			140									
Asp	Glu	Val	His	Leu	Val	Asn	Asp	Ala	His	Arg	Gly	Pro	Thr	Leu	Glu
145			150			155			160						
Val	Thr	Leu	Ala	Lys	Leu	Arg	Arg	Val	Asn	Pro	Asp	Leu	Gln	Thr	Val

165	170	175
Ala Leu Ser Ala Thr Val Gly Asn Ala Gly Glu Met Ala Asp Trp Leu		
180	185	190
Asp Ala Thr Leu Val Asp Ser Thr Trp Arg Pro Ile Asp Leu Arg Lys		
195	200	205
Gly Val Leu Tyr Gly Gln Ala Leu His Phe Asp Asp Gly Thr Gln Gln		
210	215	220
Glu Leu Ala Arg Gly Asn Glu Lys Glu Thr Ala Ala Leu Val Arg Asp		
225	230	235
Thr Leu Glu Asp Gly Gly Ser Ser Leu Val Phe Val Asn Ser Arg Arg		
245	250	255
Asn Ala Glu Ala Ala Ala Lys Arg Leu Ala Asp Val Thr Lys Thr His		
260	265	270
Leu Thr Asp Asp Glu Arg Arg Asp Leu Leu Asp Ile Ala Asp Gln Ile		
275	280	285
Arg Asp Val Ser Asp Thr Glu Thr Ser Asp Asp Leu Ala Thr Ala Ile		
290	295	300
Glu Lys Gly Ala Ala Phe His His Ala Gly Leu Ala Ser Asp His Arg		
305	310	315
Ser Leu Val Glu Asp Ala Phe Arg Asp Lys Leu Ile Lys Val Ile Ser		
325	330	335
Ala Thr Pro Thr Leu Ala Ala Gly Val Asn Thr Pro Ser Arg Arg Val		
340	345	350
Ile Val Arg Asp Trp Arg Arg Tyr Asp Gly Asp Ile Gly Gly Met Gln		
355	360	365
Pro Leu Asp Val Leu Glu Val His Gln Met Phe Gly Arg Ala Gly Arg		
370	375	380
Pro Gly Leu Asp Pro His Gly Glu Ala Val Leu Ile Ala Lys Ser His		
385	390	395
Asp Glu Leu Gln Glu Leu Phe Asp Gln Tyr Val Trp Ala Asp Pro Glu		
405	410	415

Pro Val His Ser Lys Leu Ala Ala Glu Pro Ala Leu Arg Thr His Ile  
420 425 430

Leu Ala Thr Val Ala Ser Gly Phe Ala Gly Thr Glu Glu Glu Leu Leu  
435 440 445

Asp Phe Leu Glu Arg Thr Leu Tyr Ala Thr Gln Thr Asp Glu Thr Gly  
450 455 460

Arg Leu Glu Thr Val Thr Gln His Val Leu Asp Tyr Leu Asp Arg Asn  
465 470 475 480

Gly Phe Leu Glu Arg Asp Asp Arg Leu Arg Ala Thr Gly Leu Gly His  
485 490 495

Arg Val Ser Gln Leu Tyr Leu Asp Pro Met Ser Ala Ala Glu Ile Ile  
500 505 510

Asp Gly Leu Arg Asp Ala Asp Gly Lys Pro Thr Ala Leu Gly Leu Tyr  
515 520 525

His Leu Val Ser Arg Thr Pro Asp Met Tyr Gln Leu Tyr Leu Arg Ser  
530 535 540

Gly Asp Arg Glu Arg Tyr Thr Glu Ile Ala Tyr Glu Arg Glu Pro Glu  
545 550 555 560

Phe Leu Gly His Met Pro Ser Glu Phe Glu Asp Asn Ala Phe Glu Asp  
565 570 575

Trp Leu Ser Ala Leu Lys Thr Ala Arg Leu Leu Glu Asp Trp Ala Ser  
580 585 590

Glu Leu Asp Glu Asp Arg Ile Thr Glu Arg Tyr Ala Ile Gly Pro Gly  
595 600 605

Asp Ile Arg Gly Lys Val Glu Thr Ala Gln Trp Leu Leu Asn Ala Ala  
610 615 620

Glu Arg Leu Ala Ala Glu Leu Gln Arg Asp Asp Ala Glu Gly Ile Pro  
625 630 635 640

Ser Ala Thr Thr Thr Ala Val Arg Glu Ala Arg Lys Arg Val Glu Tyr  
645 650 655

Gly Val Glu Glu Glu Leu Leu Asp Leu Ala Gly Val Arg Asn Val Gly

660 665 670

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

675 680 685

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

690 695 700

Arg Lys Lys Thr Ala Glu Asn Ile Leu Glu Asn Val Gly Arg Gln Asp

705 710 715 720

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

725 730 735

Ala Arg Ala Thr Asn Asp Gly Gly Gln Gln Ser Leu Gly Asp Phe Glu

740 745 750

<210> 56

<211> 8

<212> PRT

<213> Artificial sequence

<220><223> Exemplary Hel308 motif

<400> 56

Gln Met Phe Gly Arg Ala Gly Arg

1 5

<210> 57

<211> 9

<212> PRT

<213> Artificial sequence

<220><223> Exemplary extended Hel308 motif

<400> 57

Gln Met Phe Gly Arg Ala Gly Arg Pro

1 5

<210> 58

<211> 783

<212> PRT

<213> Halobacterium sp. NRC-1

<400> 58

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

245                                      250                                      255  
 Thr Leu Val Phe Val Asn Ser Arg Arg Asn Ala Glu Ala Ser Ala Arg  
 260                                      265                                      270  
 Arg Leu Ala Asp Val Thr Gly Asn Ala Leu Ser Ser Ala Glu Arg Glu  
 275                                      280                                      285  
 Arg Leu Ala Asp Ile Ala Ala Glu Ile Arg Gly Val Ser Asp Thr Glu  
 290                                      295                                      300  
 Thr Ser Asp Glu Leu Ala Asp Ala Val Ala Ser Gly Ala Ala Phe His  
  
 305                                      310                                      315                                      320  
 His Ala Gly Leu Ala Arg Glu His Arg Glu Leu Val Glu Glu Ala Phe  
 325                                      330                                      335  
 Arg Asp Arg Leu Val Lys Ala Val Ser Ala Thr Pro Thr Leu Ala Ala  
 340                                      345                                      350  
 Gly Val Asn Thr Pro Ala Arg Arg Val Val Val Arg Asp Trp Gln Arg  
 355                                      360                                      365  
 Tyr Asp Gly Thr Ala Gly Gly Met Gln Pro Leu Asp Val Leu Glu Val  
  
 370                                      375                                      380  
 His Gln Met Phe Gly Arg Ala Gly Arg Pro Gly Leu Asp Pro Tyr Gly  
 385                                      390                                      395                                      400  
 Glu Ala Val Leu Leu Ala Asn Ser His Asp Glu Leu Glu Glu Leu Phe  
 405                                      410                                      415  
 Asp Arg Tyr Val Tyr Ala Asp Pro Glu Pro Val Arg Ser Lys Leu Ala  
 420                                      425                                      430  
 Ala Glu Pro Ala Leu Arg Thr His Val Leu Ala Ala Ile Ala Thr Gly  
  
 435                                      440                                      445  
 Phe Thr Thr Thr Glu Asp Gly Leu His Glu Phe Leu Gly Gly Thr Leu  
 450                                      455                                      460  
 Tyr Ala Thr Gln Thr Asp Asp Thr Gly Arg Leu Arg Ser Val Thr Gly  
 465                                      470                                      475                                      480  
 Asp Val Leu Arg Tyr Leu Asp Arg Asn Gly Phe Val Glu Arg Asp Gly  
 485                                      490                                      495

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

500 505 510  
Val Asp Pro Met Ser Ala Ala Thr Ile Ile Asp Gly Leu Arg Asp Ala  
515 520 525  
Ala Arg Asp Ala Thr Glu Thr Asp Asp Glu Gly Ala Phe Arg Pro Ala  
530 535 540  
Ser Glu Leu Gly Asp Asp Ala Ala Leu Pro Ala Asp Ala Ser Val Glu  
545 550 555 560  
Pro Thr Pro Leu Gly Leu Tyr His Leu Val Ser Arg Thr Pro Asp Met

565 570 575  
Tyr Glu Leu Tyr Leu Arg Ser Gly Asp Arg Glu Gln Tyr Thr Glu Val  
580 585 590  
Ala Tyr Glu His Glu Asp Glu Leu Leu Gly Ala Thr Pro Arg Glu Glu  
595 600 605  
Gln Ala Glu Phe Glu Asp Trp Leu Ser Ala Leu Lys Thr Ala Arg Leu  
610 615 620  
Met Ala Asp Trp Ala Ser Glu Leu Asp Glu Glu Arg Ile Ala Glu Arg

625 630 635 640  
Tyr Asp Val Gly Pro Gly Asp Ile Arg Gly Lys Val Glu Thr Ala Glu  
645 650 655  
Trp Leu Leu Asn Ala Ala Glu Arg Leu Ala Gly Glu Leu Asp Val Glu  
660 665 670  
Cys Gly Pro Ala Val Arg Glu Ala Arg Lys Arg Val Gln Tyr Gly Val  
675 680 685  
Arg Glu Glu Leu Leu Gly Leu Ala Gly Val Arg Asn Val Gly Arg Lys

690 695 700  
Arg Ala Arg Arg Leu Tyr Asn Ala Gly Val Glu Ser Arg Ala Asp Leu  
705 710 715 720  
Arg Asn Ala Asp Lys Gly Val Val Leu Gly Ala Val Arg Gly Arg Ala  
725 730 735  
Ala Thr Ala Glu Arg Ile Leu Glu Thr Val Gly His Pro Asp Pro Gly

740 745 750  
Met Asp Gly Val Ala Ala Asp Thr Asp Ala Ala Pro Glu Ser Gly Gly

755 760 765  
Glu Ala Gly Gly Asp Glu Gly Gln Ala Ser Leu Gly Asp Phe Ser

770 775 780

<210> 59

<211> 450

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 59

tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt ggttggttct 60  
gttggtgctg atattgcttt tgatgccgac cctaaatttt ttgcctgttt ggttcgcttt 120  
gagtccttctt cggttccgac taccctcccg actgcctatg atgtttatcc ttggaatggt 180

cgccatgatg gtggttatta taccgtcaag gactgtgtga ctattgacgt ccttccccgt 240  
acgccgggca ataacgttta tgttggtttc atggttttgt ctaactttac cgctactaaa 300  
tgccgcggat tggtttcgct gaatcaggtt attaaagaga ttatttgtct ccageccactt 360  
aagtgaggtg atttatgttt ggtgctattg ctggcggtat tgcttctgct cttgctggtg 420  
gcgcatgtc taaattgttt ggaggcggtc 450

<210> 60

<211> 57

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 60

gcaatatcag caccaacaga aacaaccttt tttttttttt tttttttttt tttttttt 57

<210> 61

<211> 80

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 61



ctattctgtt tatgtttctt gtttgtagc cctattctgt ccccccccc acccccccc 60  
 acccccccc acccccccc 80  
 <210> 62  
 <211> 40  
 <212> DNA  
 <213> Artificial sequence  
 <220><223> Sequence used in the Examples  
 <400> 62  
 acagaatagg gctaacaaac aagaacata aacagaatag 40  
 <210> 63  
 <211> 60  
 <212> DNA  
 <  
 213> Artificial sequence  
 <220><223> Sequence used in the Examples  
 <400> 63  
 ctattctgtt tatgtttctt gtttgtagc cctattctgt ccccccccc acccccccc 60  
 <210> 64  
 <211> 40  
 <212> DNA  
 <213> Artificial sequence  
 <220><223> Sequence used in the Examples  
 <400> 64  
 ctattctgtt tatgtttctt gtttgtagc cctattctgt 40  
 <210> 65  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 <220><223> Sequence used in the Examples  
 <400> 65  
 ccccccccc acccccccc 20  
 <210> 66  
 <211> 40  
 <212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 66

ctattctgtt tatgtttctt gtttgttagc cctattctgt 40

<210> 67

<211> 974

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 67

tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt ggttgtttct 60

gttgggtgctg atattgctgt gttctatgtc ttattctgtg tatgtatctt gtctgttagc 120

cccgattgtt accggataat tcgagctcgg taccaccccc ggttgataat cagaaaagcc 180

ccaaaaacag gaagattgta taagcaaata tttaaattgt aaacgttaat attttgtaa 240

aattcgcgtt aaatTTTTgt taaatcagct catTTTTta ccaataggcc gaaatcggca 300

aaatccctta taaatcaaaa gaatagaccg agatagggtt gagtgttgtt ccagtttgga 360

acaagagtcc agtattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc 420

agggcgatgg ccactacgt gaaccatcac cctaatcaag tttttgggg tcgaggtgcc 480

gtaaagcact aaatcggaac cctaaaggga tgccccgatt tagagcttga cggggaaagc 540

cggcgaacgt ggcgagaaag gaagggaaga aagcgaaagg agcgggcgct agggcgctgg 600

caagtgtagc ggtcacgctg cgcgtaacca ccacaccgc cgcgcttaat gcgccgtac 660

agggcgctg gggatcctct agagtcgacc tgcaggcatg caagctatcc cgcaagaggc 720

ccggcagtag cggcataacc aagcctatgc ctacagcatc cagggtgacg gtgccgagga 780

tgacgatgag cgcattgtta gatttcatac acggtgcctg actgcgttag caatttaact 840

gtgataaact accgcattaa agctagctta tcgatgataa gctgtcaaac atgagaattc 900

ttgaagacga aagggcctcg tgatacgcct atttttatag gttaatgtca tgataataat 960

ggtttcttag acgt 974

<210> 68

<211> 893

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 68

ctaagaaacc attattatca tgacattaac ctataaaaat aggcgatatca cgaggccctt 60  
tcgtcttcaa gaattctcat gtttgacagc ttatcatcga taagctagct ttaatgcggt 120  
agtttatcac agttaaattg ctaacgcagt caggcacctg gtatgaaatc taacaatgcg 180  
ctcatcgta tcttcggcac cgtcacctg gatgctgtag gcataggctt ggttatgccg 240  
gtactgccgg gcctcttgcg ggatagcttg catgcctgca ggtcgactct agaggatccc 300  
cacgcgccct gtagcggcgc attaacgcgc gcgggtgtgg tggttacgcg cagcgtgacc 360

gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc ctttctcgcc 420  
acgttcgccg gctttcccg tcaagctcta aatcggggca tccctttagg gttccgattt 480  
agtgccttac ggcacctga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg 540  
ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatact 600  
ggactcttgi tccaaactgg aacaacactc aaccctatct cggctctattc ttttgattta 660  
taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt 720  
aacgcgaatt ttaacaaaat attaacgttt acaatttaaa tatttgctta tacaatcttc 780

ctgttttttg ggcttttctg attatcaacc ggggtgggta ccgagctcga attatccggt 840  
aacaatcggg gctaacagac aagatacata cacagaataa gacatagaac aca 893

<210> 69

<211> 28

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 69

gcaatatcag caccaacaga aacaacct 28

<210> 70

<211> 117

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 70

tcgtctctcc acaggtctca gcttgagcag cgaataaag aacattatga tcagtaggag 60

cactacgacc ttgtttctgg tgctcgtccg ggcgcccaaa gtggagcgag tgcccc 117

<210> 71

<211> 80

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 71

gcactcgctc cactttgggc gcccgacga gcaccagaac aaaggtcgta gtgctcctac 60

tgatcataat gttcttattt 80

<210> 72

<211> 33

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 72

tcgctgctca agctgagacc tgtggagcag cga 33

<210> 73

<211> 113

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 73

tcgctgctcc acaggtctca gcttgagcag cgaaaataag aacattatga tcagtaggag 60

cactacgacc ttgtttctgg tgctcgctcg ggcgcccaaa gtggagcgag tgc 113

<210> 74

<211> 28

<212> DNA

<213> Artificial sequence

<220><223> Sequence used in the Examples

<400> 74

tcgctgctcc acaggtctca gcttcccc 28

<210> 75

<211> 966

<212> PRT

<213> Desulfonatronospira thiodismutans

<400> 75

Met Pro Gly Val Asp Glu Leu Leu Gln Gln Met Gly Gln Gly Asp Leu

1 5 10 15

Gln Gly Leu Ser Thr Val Ala Val Lys Glu Ile Pro Ala Arg Glu Ala

20 25 30

Glu Phe Ser Gly Ile Glu Gly Leu Pro Pro Pro Leu Lys Gln Ala Leu

35 40 45

Thr Glu Ser Gly Ile Glu Asn Phe Tyr Thr His Gln Ala Arg Ala Val

50 55 60

Asn Leu Val Arg Lys Gly Arg Ser Val Val Thr Ala Thr Pro Thr Ala

65 70 75 80

Ser Gly Lys Ser Leu Ile Tyr Asn Ile Pro Val Leu Glu Ser Ile Ile

85 90 95

Asn Asp Pro Ala Ser Arg Ala Leu Tyr Leu Phe Pro Leu Lys Ala Leu

100 105 110

Thr Arg Asp Gln Leu Thr Ser Leu Glu Glu Phe Ala Arg Leu Leu Ala

115 120 125

Gly Lys Val His Val Asp Ser Ala Val Tyr Asp Gly Asp Thr Asp Pro

130 135 140

Gln Ala Arg Ala Arg Ile Arg Ser Lys Pro Pro Asn Ile Leu Leu Thr

145 150 155 160

Asn Pro Asp Met Leu His Arg Ser Phe Leu Pro Tyr His Arg Ser Trp

165 170 175

Gln Lys Phe Phe Ser Ala Leu Lys Tyr Ile Val Val Asp Glu Val His

180 185 190

Thr Tyr Arg Gly Val Met Gly Ser Asn Met Ala Trp Val Phe Arg Arg

195 200 205

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

210 215 220

Ser Ser Ala Thr Ile Ala Asn Pro Gly Gln Leu Cys Ser Ala Leu Thr

225 230 235 240

Gly His Glu Pro Glu Val Ile Gln Lys Gly Gly Ala Pro Ala Gly Lys

245 250 255

Lys His Phe Leu Leu Leu Asp Pro Glu Met Gln Gly Ala Ala Gln Ser

260 265 270

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

275 280 285

Val Tyr Thr Gln Ser Arg Lys Met Thr Glu Leu Ile Ala Met Trp Ala

290 295 300

Ser Gln Arg Ala Gly Arg Leu Lys Lys Tyr Ile Ser Ala Tyr Arg Ala

305 310 315 320

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

325 330 335

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

340 345 350

Asp Ile Gly His Leu Asp Leu Cys Leu Leu Val Gly Tyr Pro Gly Ser

355 360 365

Val Met Ala Thr Met Gln Arg Gly Gly Arg Val Gly Arg Ser Gly Arg

370 375 380

Asp Ser Ala Ile Met Leu Ile Gly His Glu Asp Ala Leu Asp Gln Tyr

385 390 395 400

Leu Leu Arg Asn Pro Arg Glu Phe Phe Ser Leu Glu Pro Glu Ser Ala

405 410 415

Val Ile Asn Pro Asp Asn Pro Ser Ile Met Arg Arg His Leu Val Cys

420 425 430

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

435 440 445

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

450 455 460

Ala Ser Arg Asp Arg Ser Phe Tyr Tyr Thr Arg Ala Arg Tyr Pro His

465 470 475 480

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

485                      490                      495  
 His Ser Thr Gly Glu Tyr Leu Gly Glu Val Asp Gly Val Arg Ala Phe  
  
 500                      505                      510  
 Lys Glu Thr His Pro Gly Ala Val Tyr Leu His Met Gly Glu Thr Tyr  
 515                      520                      525  
 Val Val Gln Asp Leu Asp Leu Glu Thr Phe Ala Val Tyr Ala Ala Lys  
 530                      535                      540  
 Ser Glu Ala Asn Tyr Tyr Thr Arg Pro Ile Thr Glu Lys Tyr Thr Glu  
 545                      550                      555                      560  
 Ile Val Glu Val Gln Ala Thr Arg Ala Thr Ala Ala Gly Glu Leu Cys  
  
 565                      570                      575  
 Leu Gly Arg Leu Lys Val Thr Glu His Val Ser Ala Tyr Glu Lys Arg  
 580                      585                      590  
 Leu Val Arg Gly Gln Ala Arg Ile Gly Leu Ile Pro Leu Asp Leu Pro  
 595                      600                      605  
 Pro Leu Val Phe Glu Thr Gln Gly Met Trp Phe Thr Leu Asp Ser Gln  
 610                      615                      620  
 Val Arg Arg Asp Val Glu Asp Arg Arg Leu His Phe Met Gly Gly Leu  
  
 625                      630                      635                      640  
 His Ala Leu Glu His Gly Leu Ile Gly Cys Met Pro Leu Ile Ile Leu  
 645                      650                      655  
 Thr Asp Arg Asn Asp Leu Gly Gly Ile Ala Ser Pro Val His Glu Gln  
 660                      665                      670  
 Leu His Lys Gly Ala Val Phe Ile Tyr Asp Gly Thr Pro Gly Gly Ile  
 675                      680                      685  
 Gly Leu Cys Arg Gln Ala Phe Glu Leu Gly Asp Arg Leu Val Ala Arg  
  
 690                      695                      700  
 Ala Met Gly Ile Leu Ser Ser Cys Thr Cys Glu Asn Gly Cys Pro Gly  
 705                      710                      715                      720  
 Cys Ile His Ser Pro Lys Cys Gly Ser Gly Asn Arg Pro Leu Asp Lys  
 725                      730                      735

Glu Ala Ala Met His Met Leu Ala Val Leu Ala Gly Glu Arg Cys Gly  
 740 745 750  
 Glu Ala Lys Arg Lys Asp Val Ser Cys Arg Ile Glu Thr Asp Glu Gly  
 755 760 765  
 Ser Met Glu Ile Asp Ser Gly Tyr Thr Lys Ser Asp Gln Ala Glu Leu  
 770 775 780  
 Pro Tyr Ala Val Leu Asp Ile Glu Thr Arg Tyr Ser Ala Gln Glu Val  
 785 790 795 800  
 Gly Gly Trp Gly Asn Cys His Arg Met Gly Val Ser Phe Ala Val Val  
 805 810 815  
 Phe Asp Ser Arg Asn Gln Glu Phe Val Thr Phe Asp Gln Glu Gln Ala  
 820 825 830  
 Ala Asp Leu Gly Ser Phe Leu Glu Asp Phe Ser Leu Val Val Gly Phe  
 835 840 845  
 Asn Leu Leu Lys Phe Asp Tyr Arg Val Leu Gln Gly Leu Ser Asp Tyr  
 850 855 860  
 Asp Phe Ser Ser Leu Pro Thr Leu Asp Met Leu Arg Glu Ile Glu Ala  
 865 870 875 880  
 Arg Leu Gly His Arg Leu Ser Leu Asp His Leu Ala Arg His Thr Leu  
 885 890 895  
 Gly Thr Asn Lys Ser Ala Asn Gly Leu Met Ala Leu Lys Trp Trp Lys  
 900 905 910  
 Glu Gly Glu Leu Asp Lys Ile Val Glu Tyr Cys Arg Gln Asp Val Ser  
 915 920 925  
 Val Thr Arg Asp Leu Tyr Leu Phe Gly Arg Asp Lys Gly Tyr Leu Leu  
 930 935 940  
 Phe Lys Asn Lys Ala Gly Lys Lys Val Arg Ile Pro Val Ser Trp Gln  
 945 950 955 960  
 Asp Thr Ala Phe Gln Val  
 965  
 <210> 76



<211> 799

<212> PRT

<213> Haloarcula marismortui

<400> 76

Met Asp Val Ala Asp Leu Pro Gly Val Pro Glu Trp Leu Pro Asp His

1 5 10 15

Leu Arg Asp Asp Gly Ile Glu Glu Leu Tyr Pro Pro Gln Ala Glu Ala

20 25 30

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

35 40 45

Thr Ala Ser Gly Lys Thr Leu Ile Ala Glu Leu Ala Met Leu Ser Ser

50 55 60

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

65 70 75 80

Ala Ser Glu Lys Gln Ala Asp Phe Glu Glu Phe Glu Gln Tyr Gly Leu

85 90 95

Asp Ile Gly Val Ser Thr Gly Asn Tyr Glu Ser Glu Gly Gly Trp Leu

100 105 110

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

115 120 125

Val Arg Asn Asp Ala Pro Trp Ile Glu Asp Leu Thr Cys Val Val Thr

130 135 140

Asp Glu Val His Leu Val Asp Asp Gly Glu Arg Gly Pro Thr Leu Glu

145 150 155 160

Val Thr Leu Ala Lys Leu Arg Arg Leu Asn Pro Asp Leu Gln Thr Val

165 170 175

Ala Leu Ser Ala Thr Ile Gly Asn Ala Glu Ala Leu Ala Thr Trp Leu

180 185 190

Asp Ala Gly Leu Val Asp Ser Asp Trp Arg Pro Ile Asp Leu Gln Lys

195 200 205

Gly Val His Tyr Gly Gln Ala Leu His Leu Glu Asp Gly Ser Gln Gln

210 215 220

Arg Leu Ser Val Gln Asn Asn Glu Lys Gln Thr Ala Ala Ile Val Arg

225 230 235 240

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

245 250 255

Arg Asn Ala Glu Ala Ala Ala Gly Arg Leu Ala Asn Thr Val Arg Pro

260 265 270

His Leu Ser Thr Glu Glu Arg Asp Gln Leu Ala Asp Ile Ala Glu Glu

275 280 285

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

290 295 300

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

305 310 315 320

Arg Glu Leu Val Glu Asp Ala Phe Arg Asp Arg Leu Val Lys Val Val

325 330 335

Cys Ala Thr Pro Thr Leu Ala Ala Gly Val Asn Thr Pro Ser Arg Arg

340 345 350

Val Val Val Arg Asp Trp Arg Arg Tyr Asp Gly Ser Ala Gly Gly Met

355 360 365

Ala Pro Leu Ser Val Leu Glu Val His Gln Met Met Gly Arg Ala Gly

370 375 380

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

385 390 395 400

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

405 410 415

Glu Pro Val Arg Ser Lys Leu Ala Ala Glu Pro Ala Leu Arg Thr His

420 425 430

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

435 440 445

Leu Glu Phe Leu Glu Gln Thr Leu Tyr Ala Ser Gln Thr Asp Asp Ser

450 455 460

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

465                      470                      475                      480  
 Asn Asp Phe Leu Glu Ile Glu Ala Gly Glu Leu Asp Ala Thr Ser Leu  
  
                          485                      490                      495  
 Gly His Thr Val Ser Arg Leu Tyr Leu Asp Pro Met Ser Ala Ala Glu  
                          500                      505                      510  
 Ile Val Asp Gly Leu Arg Asp Trp Glu Arg Gly Ala Ser Asp Ser Thr  
                          515                      520                      525  
 Ser Ala Ser Gly Ser Pro Ala Asp Ala Gln Ala Glu Pro Pro Ala Asn  
                          530                      535                      540  
 Ser Gly Phe Thr Thr Ala Ser Glu Leu Ala Glu Asp Ala Asp Glu Ser  
  
 545                      550                      555                      560  
 Asp Ala Asp Arg Asp Pro Asp Asp Ile Ser Ala Leu Gly Leu Tyr His  
                          565                      570                      575  
 Leu Val Ser Arg Thr Pro Asp Met Tyr Gln Leu Tyr Leu Arg Ser Gly  
                          580                      585                      590  
 Asp Arg Glu Glu Tyr Glu Met Glu Leu Phe Glu Arg Glu Glu Glu Leu  
                          595                      600                      605  
 Leu Gly Pro Thr Pro Ser Glu Phe Glu Glu Gly Arg Phe Glu Asp Trp  
  
 610                      615                      620  
 Leu Ser Ala Leu Lys Thr Ala Arg Leu Leu Glu Asp Trp Ala Thr Glu  
 625                      630                      635                      640  
 Val Asp Glu Ala Thr Ile Thr Asp Arg Tyr Gly Val Gly Pro Gly Asp  
                          645                      650                      655  
 Ile Arg Gly Lys Val Glu Thr Ala Gln Trp Leu Leu Gly Ala Ala Glu  
                          660                      665                      670  
 Ser Leu Ala Ser Glu Val Asp Leu Asp Ala Ala Arg Ala Ile Ser Glu  
  
                          675                      680                      685  
 Ala Arg Ile Arg Val Glu His Gly Val Arg Glu Glu Leu Val Asp Leu  
                          690                      695                      700  
 Ala Gly Val Arg Gly Val Gly Arg Lys Arg Ala Arg Arg Leu Phe Gln  
 705                      710                      715                      720

Ala Gly Ile Thr Asp Arg Ala Gln Leu Arg Asp Ala Asp Lys Ala Val  
725 730 735

Val Leu Ala Ala Leu Arg Gly Arg Arg Lys Thr Ala Glu Asn Val Leu

740 745 750

Glu Asn Ala Gly His Arg Asp Pro Ser Met Glu Gly Val Glu Pro Ala

755 760 765

Pro Asp Val Ser Val Asp Leu Asn Asp Gly Ala Asp Gly Asp Ala Ser

770 775 780

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

785 790 795

<210> 77

<211> 711

<212> PRT

<213> Natranaerobius thermophilus

<400> 77

Met Ser Glu Thr Phe Tyr Leu Leu Ser Glu Arg Met Gln Lys Lys Ile

1 5 10 15

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

20 25 30

Pro Ile Val Met Asn Thr Asn Lys Asp Val Val Val Ser Ser Gly Thr

35 40 45

Ala Ser Gly Lys Thr Glu Ala Val Phe Leu Pro Ile Leu Ser Gln Ile

50 55 60

Glu Lys Asp Ala Thr Lys Asp Leu Lys Ile Leu Tyr Ile Ser Pro Leu

65 70 75 80

Lys Ala Leu Ile Asn Asp Gln Phe Glu Arg Ile Ile Lys Leu Cys Glu

85 90 95

Lys Ser Tyr Ile Pro Ile His Arg Trp His Gly Asp Val Asn Gln Asn

100 105 110

Lys Lys Lys Gln Leu Thr Lys Asn Pro Ala Gly Ile Leu Gln Ile Thr

115 120 125

Pro Glu Ser Ile Glu Ser Leu Phe Ile Asn Arg Thr Asn Glu Leu Asn  
130 135 140

Tyr Ile Leu Ser Asp Ile Glu Phe Ile Ile Ile Asp Glu Leu His Ala  
145 150 155 160

Phe Leu Asp Asn Glu Arg Gly Val His Leu Arg Ser Leu Leu Ser Arg  
165 170 175

Leu Glu Asn Tyr Ile Lys Glu Lys Pro Arg Tyr Phe Ala Leu Ser Ala  
180 185 190

Thr Leu Asn Asn Phe Lys Leu Ile Lys Glu Trp Ile Asn Tyr Asn Asp  
195 200 205

Ile Lys Asn Val Glu Ile Ile Asp Ser Asn Glu Asp Asp Lys Asp Leu  
210 215 220

Leu Leu Ser Leu Met His Phe Asp Lys Gly Lys Asp Tyr Lys Lys Pro  
225 230 235 240

Ile Asp Leu Tyr Gln Asp Leu Arg Glu Leu Thr Lys Asn Val His Ser  
245 250 255

Leu Ile Phe Cys Asn Ser Arg Ala Glu Val Glu Glu Thr Thr Leu Tyr  
260 265 270

Leu Asn Arg Leu Ala Asn Arg Glu Val Asn Thr Glu Leu Tyr Leu Ala  
275 280 285

His His Ser Ser Ile Asp Lys Lys Glu Arg Glu Tyr Val Glu Lys Thr  
290 295 300

Met Ala Asn Ser Lys Ser Pro Lys Ser Val Val Thr Thr Ser Ser Leu  
305 310 315 320

Glu Leu Gly Ile Asp Ile Gly Ala Ile Asp Tyr Val Val Gln Ile Asp  
325 330 335

Asp Thr His Thr Val Ser Ser Leu Lys Gln Arg Leu Gly Arg Ser Gly  
340 345 350

Arg Lys Leu Gly Thr Asn Gln Val Leu Gln Val Tyr Ser Thr Thr Asn  
355 360 365

Asp Ser Leu Val Gln Ser Leu Ala Val Ile Asp Leu Leu Leu Glu Lys

370 Trp Ile Glu Pro Ala Thr Glu Tyr Pro Leu Pro Leu Asp Ile Leu Phe  
 385 390 395 400  
 His Gln Ile Ile Ser Ile Cys His Glu Ala Asn Gly Val Arg Leu Asp  
 405 410 415  
 Pro Leu Ile Asp Asn Ile Lys Ala Asn Ala Ala Phe Tyr Lys Leu Lys  
 420 425 430  
 Glu Glu Asp Ile Asn His Val Ile Asn Tyr Met Ile Glu Asn Asp Phe  
 435 440 445  
  
 Leu Gln Leu Ile Arg Asn Ser Ala Glu Leu Ile Val Gly Leu Glu Gly  
 450 455 460  
 Glu Arg Leu Leu Arg Gly Lys Glu Phe Tyr Ala Val Phe Met Thr Gln  
 465 470 475 480  
 Glu Glu Phe Glu Val Arg Glu Gly Ile Arg Lys Ile Gly Ser Ile Asp  
 485 490 495  
 Lys Ser Leu Met Val Ser Glu Gly Asp Asn Ile Ile Leu Ala Gly Gln  
 500 505 510  
  
 Leu Trp Thr Ile Lys Asn Ile Asp Ile Glu Arg Asp Ile Ile Tyr Val  
 515 520 525  
 Ala Lys Ala Val Asp Gly Lys Pro Pro Lys Tyr Ser Gly Gly Gly Phe  
 530 535 540  
 Ile Leu Asn Pro Lys Ile Pro Glu Arg Met His Lys Ile Leu Cys Glu  
 545 550 555 560  
 Arg Lys Asn Phe Glu Phe Ile Asp Asn Met Ala Gln Asn His Leu Glu  
 565 570 575  
  
 Glu Gln Arg Lys Pro Phe Glu Leu Tyr Asn Ile Lys Pro Asn Glu Arg  
 580 585 590  
 Val Ile Trp Asn Asn Gly Asp Glu Ile Leu Phe Glu Thr Tyr Thr Gly  
 595 600 605  
 Thr Lys Ile Phe Gln Thr Leu Ala Trp Ile Leu Arg Ser Tyr Asn Val  
 610 615 620

Asn Ile Lys Glu Ile Asp Gly Ile Gly Arg Ile Asn Ile Glu Gly Gly  
625 630 635 640

Ile Asp Leu Pro Gly Val Leu Gln Asp Ile Lys Glu Thr Asp Trp Arg  
645 650 655

Pro Glu Tyr Leu Leu Asp Phe Thr Leu Glu Gln Glu Lys Phe Lys Ser  
660 665 670

Lys Phe Ser Pro Tyr Leu Pro Lys Asp Leu Gln Asp Lys Met His Ile  
675 680 685

Ala His Leu Val Asp Ile Glu Gly Val Lys Thr Phe Leu Glu Asn Lys  
690 695 700

Lys Ile Lys Glu Ile Lys Leu

705 710

<210> 78

<211> 491

<212> PRT

<213> Artificial sequence

<220><223> Consensus sequence from alignment of Hel308 sequences

<400> 78

Leu Pro Val Leu Glu Gly Ile Glu Leu Tyr Pro Pro Gln Ala Glu Ala  
1 5 10 15

Val Glu Gly Leu Leu Asp Gly Lys Asn Leu Leu Ile Ala Ile Pro Thr  
20 25 30

Ala Ser Gly Lys Thr Leu Ile Ala Glu Leu Ala Met Leu Ile Leu Gly

35 40 45

Gly Lys Ala Leu Tyr Ile Val Pro Leu Arg Ala Leu Ala Ser Glu Lys  
50 55 60

Tyr Glu Phe Lys Phe Glu Gly Val Arg Val Gly Ile Ser Thr Gly Asp  
65 70 75 80

Tyr Asp Asp Glu Trp Leu Gly Asp Ile Ile Val Ala Thr Ser Glu Lys  
85 90 95

Val Asp Ser Leu Leu Arg Asn Trp Ile Asp Ile Thr Val Val Val Val

100	105	110	
Asp Glu Ile His Leu Ile Asp Arg Gly Pro Thr Leu Glu Val Leu Leu			
115	120	125	
Ala Lys Leu Arg Leu Asn Pro Leu Gln Ile Ile Ala Leu Ser Ala Thr			
130	135	140	
Ile Gly Asn Ala Glu Glu Leu Ala Glu Trp Leu Ala Glu Leu Val Val			
145	150	155	160
Ser Asp Trp Arg Pro Val Asp Leu Arg Gly Val Phe Tyr Leu Phe Asp			
	165	170	175
Ile Leu Val Leu Asp Thr Val Glu Gly Gly Gln Leu Val Phe Asn Ser			
180	185	190	
Arg Arg Asn Ala Glu Ala Lys Lys Leu Ala Val Lys Leu Thr Glu Leu			
195	200	205	
Leu Ala Glu Glu Ile Glu Thr Glu Thr Ser Leu Ala Cys Val Lys Gly			
210	215	220	
Ala Phe His His Ala Gly Leu Arg Leu Val Glu Asp Ala Phe Arg Leu			
225	230	235	240
Ile Lys Val Ile Ala Thr Pro Thr Leu Ala Ala Gly Leu Asn Leu Pro			
245	250	255	
Ala Arg Arg Val Ile Ile Arg Asp Tyr Lys Arg Tyr Gly Met Pro Ile			
260	265	270	
Pro Val Leu Glu Ile Gln Met Gly Arg Ala Gly Arg Pro Leu Asp Pro			
275	280	285	
Tyr Gly Glu Ala Val Leu Ile Ala Lys Ser Asp Glu Leu Glu Tyr Ile			
290	295	300	
Ala Asp Pro Glu Ile Trp Ser Lys Leu Ala Glu Ala Leu Arg Thr His			
305	310	315	320
Val Leu Ala Leu Ile Ala Ser Gly Phe Ala Thr Glu Leu Leu Asp Phe			
325	330	335	
Leu Thr Phe Tyr Ala Tyr Gln Leu Ile Glu Val Leu Phe Leu Asn Ile			
340	345	350	



Leu Ala Thr Leu Gly Val Ser Leu Tyr Ile Asp Pro Leu Ser Ala Ile

355 360 365

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

370 375 380

Leu Tyr Leu Arg Asp Glu Leu Glu Ile Glu Glu Phe Phe Glu Phe Leu

385 390 395 400

Val Lys Thr Ala Leu Leu Asp Trp Ile Glu Val Glu Asp Ile Glu Arg

405 410 415

Tyr Gly Ile Gly Pro Gly Asp Leu Val Glu Ala Glu Trp Leu Met His

420 425 430

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

435 440 445

Glu Glu Leu Leu Glu Leu Val Ile Arg Ile Gly Arg Val Arg Ala Arg

450 455 460

Lys Leu Tyr Ala Gly Ile Arg Ser Asp Leu Ala Leu Ile Leu Gly Lys

465 470 475 480

Ile Ala Glu Lys Ile Leu Leu Gly Thr Leu Phe

485 490