



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0049100
(43) 공개일자 2023년04월12일

(51) 국제특허분류(Int. Cl.) <i>C07K 14/31</i> (2006.01) <i>A61K 38/00</i> (2006.01) <i>C12N 15/10</i> (2017.01) <i>C12N 15/62</i> (2006.01) <i>C12N 9/22</i> (2006.01) <i>C12N 9/78</i> (2006.01) (52) CPC특허분류 <i>C07K 14/31</i> (2013.01) <i>C12N 15/102</i> (2013.01) (21) 출원번호 10-2023-7005335 (22) 출원일자(국제) 2021년07월15일 심사청구일자 없음 (85) 번역문제출일자 2023년02월15일 (86) 국제출원번호 PCT/US2021/041809 (87) 국제공개번호 WO 2022/015969 국제공개일자 2022년01월20일 (30) 우선권주장 63/052,175 2020년07월15일 미국(US)	(71) 출원인 라이프에디트 테라퓨틱스, 인크. 미국 27560 노스캐롤라이나주 모리스빌 에어포트 불러바드 507 스위트 101 (72) 발명자 보웬, 타이슨 디. 미국 27560 노스캐롤라이나 모리스빌 월튼스 크릭 로드 137 크롤리, 알렉산드라 브리너 미국 27511 노스캐롤라이나 캐리 스콧 플레이스 1220 일리치, 테드 디. 미국 27705 노스캐롤라이나 더럼 알лум우드 플레이 스 5 (74) 대리인 양영준, 이상남

전체 청구항 수 : 총 138 항

(54) 발명의 명칭 **우라실 안정화 단백질 및 그의 활성 단편 및 변이체 및 사용 방법**

(57) 요약

핵산의 표적화된 편집을 위한 우라실 안정화 폴리펩티드를 포함하는 조성물 및 방법이 제공된다. 조성물은 우라실 안정화 폴리펩티드를 포함한다. 또한, i) DNA-결합 폴리펩티드; ii) 데아미나제; 및 iii) 우라실 안정화 폴리펩티드 (USP)를 포함하는 융합 단백질이 제공된다. 융합 단백질은, 데아미나제에 융합되고 USP에 추가로 융합된 RNA-가이드된 뉴클레아제를, 임의로 가이드 RNA와의 복합체로 포함한다. 조성물은 또한 USP 또는 융합 단백질을 코딩하는 핵산 분자를 포함한다. USP 또는 융합 단백질을 코딩하는 핵산 분자를 포함하는 벡터 및 숙주 세포가 또한 제공된다.

(52) CPC특허분류

C12N 15/62 (2013.01)

C12N 9/22 (2013.01)

C12N 9/78 (2013.01)

C12Y 305/04001 (2013.01)

C12Y 305/04005 (2013.01)

C07K 2319/02 (2013.01)

C07K 2319/09 (2013.01)

C07K 2319/80 (2013.01)

C12N 2310/20 (2017.05)

명세서

청구범위

청구항 1

a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;

b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는

c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성

을 갖는 아미노산 서열을 포함하는 단리된 폴리펩티드로서,

여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖고, 상기 폴리펩티드는 이중 아미노산 서열을 추가로 포함하는 것인

단리된 폴리펩티드.

청구항 2

제1항에 있어서, 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 단리된 폴리펩티드.

청구항 3

비-자연 발생 제약상 허용되는 담체, 및

a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;

b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는

c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성

을 갖는 아미노산 서열을 포함하는 폴리펩티드를 포함하는 제약 조성물로서,

여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖는 것인

제약 조성물.

청구항 4

비-자연 발생 제약상 허용되는 담체, 및

a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;

b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는

c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성

을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산 분자를 포함하는 제약 조성물로서,

여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖는 것인

제약 조성물.

청구항 5

제3항 또는 제4항에 있어서, 폴리펩티드가 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 것인 제약 조성물.

청구항 6

제3항 내지 제5항 중 어느 한 항에 있어서, 플루오로피리미딘을 추가로 포함하는 제약 조성물.

청구항 7

- a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;
- b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는
- c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성

을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산 분자로서,

여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖고;

여기서 상기 핵산 분자는 상기 폴리뉴클레오티드에 작동가능하게 연결된 이중 프로모터를 추가로 포함하는 것인 핵산 분자.

청구항 8

제7항에 있어서, 폴리펩티드가 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 것인 핵산 분자.

청구항 9

플루오로피리미딘, 및

- a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;
- b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는
- c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성

을 갖는 아미노산 서열을 포함하는 폴리펩티드를 포함하는 조성물로서,

여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖는 것인

조성물.

청구항 10

플루오로피리미딘, 및

- a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;
- b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는
- c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성

을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 핵산 분자를 포함하는 조성물로서,

여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖는 것인

조성물.

청구항 11

제9항 또는 제10항에 있어서, 폴리펩티드가 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 것인 조성물.

청구항 12

(i) DNA-결합 폴리펩티드; (ii) 테아미나제; 및 (iii) 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하는 융합 단백질.

청구항 13

제12항에 있어서, USP가 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 것인 융합 단백질.

청구항 14

제12항 또는 제13항에 있어서, 테아미나제가 시티딘 테아미나제인 융합 단백질.

청구항 15

제14항에 있어서, 시티딘 데아미나제가 활성화-유도된 시티딘 데아미나제 (AID) 또는 데아미나제의 아포지단백질 B mRNA-편집 복합체 (APOBEC) 패밀리의 구성원인 융합 단백질.

청구항 16

제15항에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.

청구항 17

제12항 내지 제16항 중 어느 한 항에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 융합 단백질.

청구항 18

제12항 내지 제16항 중 어느 한 항에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 융합 단백질.

청구항 19

제18항에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 융합 단백질.

청구항 20

제19항에 있어서, RGN이 유형 II CRISPR-Cas 폴리펩티드인 융합 단백질.

청구항 21

제19항에 있어서, RGN이 유형 V CRISPR-Cas 폴리펩티드인 융합 단백질.

청구항 22

제19항에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.

청구항 23

제19항 내지 제22항 중 어느 한 항에 있어서, RGN이 RGN 니카제인 융합 단백질.

청구항 24

제19항 내지 제23항 중 어느 한 항에 있어서, RGN, 시티딘 데아미나제 및 USP를 포함하는 융합 단백질.

청구항 25

제24항에 있어서, 상기 RGN이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖고, 상기 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖고, 상기 USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 것인 융합 단백질.

청구항 26

제12항 내지 제25항 중 어느 한 항에 있어서, 적어도 하나의 핵 국재화 신호 (NLS)를 추가로 포함하는 융합 단백질.

청구항 27

(i) DNA-결합 폴리펩티드; (ii) 데아미나제; 및 (iii) 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하

는 융합 단백질을 코딩하는 폴리뉴클레오티드를 포함하는 핵산 분자로서,

여기서 USP는

a) 서열식별번호: 17-32 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖거나,

b) 서열식별번호: 17-32 중 어느 하나에 제시되거나,

c) 서열식별번호: 1-16에 대해 적어도 80% 동일한 아미노산 서열을 코딩하고, 서열식별번호: 33-39 중 어느 하나의 서열을 추가로 보유하거나,

d) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열에 대해 적어도 80% 동일한 아미노산 서열을 코딩하거나, 또는

e) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열을 코딩하는

뉴클레오티드 서열에 의해 코딩되는 것인

핵산 분자.

청구항 28

제27항에 있어서, 데아미나제가 시티딘 데아미나제인 핵산 분자.

청구항 29

제28항에 있어서, 시티딘 데아미나제가 활성화-유도된 시티딘 데아미나제 (AID) 또는 데아미나제의 아포지단백질 B mRNA-편집 복합체 (APOBEC) 패밀리의 구성원인 핵산 분자.

청구항 30

제29항에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.

청구항 31

제27항 내지 제30항 중 어느 한 항에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 핵산 분자.

청구항 32

제27항 내지 제30항 중 어느 한 항에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 핵산 분자.

청구항 33

제32항에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 핵산 분자.

청구항 34

제33항에 있어서, RGN이 유형 II CRISPR-Cas 폴리펩티드인 핵산 분자.

청구항 35

제33항에 있어서, RGN이 유형 V CRISPR-Cas 폴리펩티드인 핵산 분자.

청구항 36

제33항에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.

청구항 37

제33항 내지 제36항 중 어느 한 항에 있어서, RGN이 RGN 니카제인 핵산 분자.

청구항 38

제33항 내지 제36항 중 어느 한 항에 있어서, 융합 단백질이 RGN, 시티딘 데아미나제 및 USP를 포함하는 것인 핵산 분자.

청구항 39

제38항에 있어서, 상기 RGN이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖고, 상기 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖고, 상기 USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 것인 핵산 분자.

청구항 40

제27항 내지 제39항 중 어느 한 항에 있어서, 융합 단백질을 코딩하는 폴리뉴클레오티드가 그의 5' 말단에서 이 중 프로모터에 작동가능하게 연결된 것인 핵산 분자.

청구항 41

제27항 내지 제39항 중 어느 한 항에 있어서, 융합 단백질을 코딩하는 폴리뉴클레오티드가 그의 3' 말단에서 이 중 종결인자에 작동가능하게 연결된 것인 핵산 분자.

청구항 42

제29항 내지 제41항 중 어느 한 항에 있어서, 융합 단백질이 하나 이상의 핵 국제화 신호를 포함하는 것인 핵산 분자.

청구항 43

제27항 내지 제42항 중 어느 한 항에 있어서, 융합 단백질이 진핵 세포에서의 발현에 대해 코돈 최적화된 것인 핵산 분자.

청구항 44

제27항 내지 제43항 중 어느 한 항에 있어서, 융합 단백질이 원핵 세포에서의 발현에 대해 코돈 최적화된 것인 핵산 분자.

청구항 45

제27항 내지 제44항 중 어느 한 항의 핵산 분자를 포함하는 벡터.

청구항 46

제33항 내지 제39항 중 어느 한 항의 핵산 분자를 포함하고, 표적 서열에 혼성화할 수 있는 가이드 RNA (gRNA)를 코딩하는 적어도 하나의 뉴클레오티드 서열을 추가로 포함하는 벡터.

청구항 47

제46항에 있어서, gRNA가 단일 가이드 RNA인 벡터.

청구항 48

제46항에 있어서, gRNA가 이중 가이드 RNA인 벡터.

청구항 49

제12항 내지 제26항 중 어느 한 항의 융합 단백질을 포함하는 세포.

청구항 50

제18항 내지 제25항 중 어느 한 항의 융합 단백질을 포함하고, 가이드 RNA를 추가로 포함하는 세포.

청구항 51

제27항 내지 제44항 중 어느 한 항의 핵산 분자를 포함하는 세포.

청구항 52

제45항 내지 제48항 중 어느 한 항의 벡터를 포함하는 세포.

청구항 53

제49항 내지 제52항 중 어느 한 항에 있어서, 원핵 세포인 세포.

청구항 54

제49항 내지 제52항 중 어느 한 항에 있어서, 진핵 세포인 세포.

청구항 55

제54항에 있어서, 곤충, 조류 또는 포유동물 세포인 세포.

청구항 56

제54항에 있어서, 식물 또는 진균 세포인 세포.

청구항 57

제약상 허용되는 담체, 및 제7항, 제8항, 제27항 내지 제44항 중 어느 한 항의 핵산 분자, 제9항 내지 제11항 중 어느 한 항의 조성물, 제12항 내지 제26항 중 어느 한 항의 융합 단백질, 제45항 내지 제48항 중 어느 한 항의 벡터, 또는 제49항 내지 제55항 중 어느 한 항의 세포를 포함하는 제약 조성물.

청구항 58

제49항 내지 제56항 중 어느 한 항의 세포를 융합 단백질이 발현되는 조건 하에 배양하는 것을 포함하는, 융합 단백질을 제조하는 방법.

청구항 59

제27항 내지 제44항 중 어느 한 항의 핵산 분자 또는 제45항 내지 제48항 중 어느 한 항의 벡터를 세포에 도입하고, 세포를 융합 단백질이 발현되는 조건 하에 배양하는 것을 포함하는, 융합 단백질을 제조하는 방법.

청구항 60

제58항 또는 제59항에 있어서, 상기 융합 단백질을 정제하는 것을 추가로 포함하는 방법.

청구항 61

제33항 내지 제39항 중 어느 한 항의 핵산 분자 및 가이드 RNA를 코딩하는 발현 카세트를 포함하는 핵산 분자, 또는 제46항 내지 제48항 중 어느 한 항의 벡터를 세포에 도입하고, 세포를 융합 단백질 및 gRNA가 발현되어 RGN 융합 리보핵단백질 복합체를 형성하는 조건 하에 배양하는 것을 포함하는, RGN 융합 리보핵단백질 복합체를 제조하는 방법.

청구항 62

제61항에 있어서, 상기 RGN 융합 리보핵단백질 복합체를 정제하는 것을 추가로 포함하는 방법.

청구항 63

표적 DNA 서열을 포함하는 표적 DNA 분자를 변형시키기 위한 시스템으로서,

상기 시스템은

a) RNA-가이드된 뉴클레아제 폴리펩티드 (RGN), 시티딘 데아미나제, 및 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하며, 여기서 USP는 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 동일한 것인 융합 단

백질, 또는 상기 융합 단백질을 코딩하는 뉴클레오티드 서열; 및

b) 상기 표적 DNA 서열에 혼성화할 수 있는 하나 이상의 가이드 RNA, 또는 하나 이상의 가이드 RNA (gRNA)를 코딩하는 하나 이상의 뉴클레오티드 서열

을 포함하고;

여기서 하나 이상의 가이드 RNA를 코딩하고 융합 단백질을 코딩하는 상기 뉴클레오티드 서열은 각각 상기 뉴클레오티드 서열에 이중인 프로모터에 작동가능하게 연결되고;

여기서 하나 이상의 가이드 RNA는 융합 단백질과 복합체를 형성하여 상기 융합 단백질이 상기 표적 DNA 서열에 결합하고 표적 DNA 분자를 변형시키도록 지시할 수 있는 것인

시스템.

청구항 64

제63항에 있어서, 표적 DNA 서열이 RGN에 의해 인식되는 프로토스페이서 인접 모티프 (PAM)에 인접하여 위치하는 것인 시스템.

청구항 65

제63항 또는 제64항에 있어서, 표적 DNA 분자가 세포 내에 있는 것인 시스템.

청구항 66

제65항에 있어서, 세포가 진핵 세포인 시스템.

청구항 67

제66항에 있어서, 진핵 세포가 식물 세포인 시스템.

청구항 68

제66항에 있어서, 진핵 세포가 포유동물 세포인 시스템.

청구항 69

제66항에 있어서, 진핵 세포가 곤충 세포인 시스템.

청구항 70

제65항에 있어서, 세포가 원핵 세포인 시스템.

청구항 71

제63항 내지 제70항 중 어느 한 항에 있어서, 융합 단백질의 RGN이 유형 II CRISPR-Cas 폴리펩티드인 시스템.

청구항 72

제63항 내지 제70항 중 어느 한 항에 있어서, 융합 단백질의 RGN이 유형 V CRISPR-Cas 폴리펩티드인 시스템.

청구항 73

제63항 내지 제70항 중 어느 한 항에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80% 동일한 것인 시스템.

청구항 74

제63항 내지 제73항 중 어느 한 항에 있어서, 융합 단백질의 RGN이 RGN 니카제인 시스템.

청구항 75

제63항 내지 제74항 중 어느 한 항에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하

나에 대해 적어도 80% 동일한 것인 시스템.

청구항 76

제63항 내지 제75항 중 어느 한 항에 있어서, USP가 서열식별번호: 33-39 중 어느 하나의 서열을 포함하는 것인 시스템.

청구항 77

제63항 내지 제76항 중 어느 한 항에 있어서, RGN이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖고, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖고, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 것인 시스템.

청구항 78

제63항 내지 제77항 중 어느 한 항에 있어서, 융합 단백질이 하나 이상의 핵 국제화 신호를 포함하는 것인 시스템.

청구항 79

제63항 내지 제78항 중 어느 한 항에 있어서, 융합 단백질이 진핵 세포에서의 발현에 대해 코돈 최적화된 것인 시스템.

청구항 80

제63항 내지 제79항 중 어느 한 항에 있어서, 하나 이상의 가이드 RNA를 코딩하는 뉴클레오티드 서열 및 융합 단백질을 코딩하는 뉴클레오티드 서열이 하나의 벡터 상에 위치하는 것인 시스템.

청구항 81

표적 DNA 서열을 포함하는 표적 DNA 분자를 변형시키는 방법으로서, 상기 방법은 제63항 내지 제80항 중 어느 한 항에 따른 시스템을 상기 표적 DNA 분자 또는 표적 DNA 분자를 포함하는 세포에 전달하는 것을 포함하는 것인 방법.

청구항 82

제81항에 있어서, 상기 변형된 표적 DNA 분자가 표적 DNA 분자 내 적어도 1개의 뉴클레오티드의 C>T 돌연변이를 포함하는 것인 방법.

청구항 83

제81항에 있어서, 상기 변형된 표적 DNA 분자가 표적 DNA 서열 내에 적어도 하나의 뉴클레오티드의 C>T 돌연변이를 포함하는 것인 방법.

청구항 84

표적 서열을 포함하는 표적 DNA 분자를 변형시키는 방법으로서,

a) RGN-데아미나제-USP 리보뉴클레오티드 복합체의 형성에 적합한 조건 하에

i) 표적 DNA 서열에 혼성화할 수 있는 하나 이상의 가이드 RNA; 및

ii) RNA-가이드된 뉴클레아제 폴리펩티드 (RGN), 시티딘 데아미나제 및 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하며, 여기서 USP는 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 동일한 것인 융합 단백질

을 조합하여 시험관내에서 RGN-데아미나제-USP 리보뉴클레오티드 복합체를 조립하는 단계; 및

b) 상기 표적 DNA 분자 또는 상기 표적 DNA 분자를 포함하는 세포를 시험관내-조립된 RGN-데아미나제-USP 리보뉴클레오티드 복합체와 접촉시키는 단계

를 포함하고;

여기서 하나 이상의 가이드 RNA는 표적 DNA 서열에 혼성화하여, 상기 융합 단백질이 상기 표적 DNA 서열에 결합하도록 지시하고, 표적 DNA 분자의 변형이 일어나는 것인

방법.

청구항 85

제84항에 있어서, 상기 변형된 표적 DNA 분자가 표적 DNA 분자 내 적어도 1개의 뉴클레오타이드의 C>T 돌연변이를 포함하는 것인 방법.

청구항 86

제84항에 있어서, 상기 변형된 표적 DNA 분자가 표적 DNA 서열 내에 적어도 하나의 뉴클레오타이드의 C>T 돌연변이를 포함하는 것인 방법.

청구항 87

제84항 내지 제86항 중 어느 한 항에 있어서, 융합 단백질의 RGN이 유형 II CRISPR-Cas 폴리펩티드인 방법.

청구항 88

제84항 내지 제86항 중 어느 한 항에 있어서, 융합 단백질의 RGN이 유형 V CRISPR-Cas 폴리펩티드인 방법.

청구항 89

제84항 내지 제86항 중 어느 한 항에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80% 동일한 것인 방법.

청구항 90

제84항 내지 제89항 중 어느 한 항에 있어서, 융합 단백질의 RGN이 RGN 닉카제인 방법.

청구항 91

제84항 내지 제90항 중 어느 한 항에 있어서, 시티딘 테아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 동일한 것인 방법.

청구항 92

제84항 내지 제91항 중 어느 한 항에 있어서, USP가 서열식별번호: 33-39 중 어느 하나의 서열을 포함하는 것인 방법.

청구항 93

제84항 내지 제92항 중 어느 한 항에 있어서, RGN이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖고, 시티딘 테아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖고, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 것인 방법.

청구항 94

제84항 내지 제93항 중 어느 한 항에 있어서, 융합 단백질이 하나 이상의 핵 국제화 신호를 포함하는 것인 방법.

청구항 95

제84항 내지 제94항 중 어느 한 항에 있어서, 융합 단백질이 진핵 세포에서의 발현에 대해 코돈 최적화된 것인 방법.

청구항 96

제84항 내지 제95항 중 어느 한 항에 있어서, 상기 표적 DNA 서열이 프로토스페이서 인접 모티프 (PAM)에 인접하여 위치하는 것인 방법.

청구항 97

제84항 내지 제96항 중 어느 한 항에 있어서, 표적 DNA 분자가 세포 내에 있는 것인 방법.

청구항 98

제97항에 있어서, 세포가 진핵 세포인 방법.

청구항 99

제98항에 있어서, 진핵 세포가 식물 세포인 방법.

청구항 100

제98항에 있어서, 진핵 세포가 포유동물 세포인 방법.

청구항 101

제98항에 있어서, 진핵 세포가 곤충 세포인 방법.

청구항 102

제97항에 있어서, 세포가 원핵 세포인 방법.

청구항 103

제97항 내지 제102항 중 어느 한 항에 있어서, 상기 변형된 DNA 분자를 포함하는 세포를 선택하는 것을 추가로 포함하는 방법.

청구항 104

제103항의 방법에 따른 변형된 표적 DNA 서열을 포함하는 세포.

청구항 105

제104항에 있어서, 진핵 세포인 세포.

청구항 106

제105항에 있어서, 진핵 세포가 식물 세포인 세포.

청구항 107

제106항의 세포를 포함하는 식물.

청구항 108

제106항의 세포를 포함하는 종자.

청구항 109

제105항에 있어서, 진핵 세포가 포유동물 세포인 세포.

청구항 110

제105항에 있어서, 진핵 세포가 곤충 세포인 세포.

청구항 111

제104항에 있어서, 원핵 세포인 세포.

청구항 112

유전적 유전성 질환에 대한 원인 돌연변이에서의 교정으로 유전자 변형된 세포를 생성하는 방법으로서,
상기 방법은 세포에

a) RNA-가이드된 뉴클레아제 폴리펩티드 (RGN), 시티딘 데아미나제, 및 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하며, 여기서 USP는 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 동일한 것인 융합 단백질, 또는 상기 융합 단백질을 코딩하는 폴리뉴클레오티드로서, 여기서 융합 단백질을 코딩하는 상기 폴리뉴클레오티드는 세포에서 융합 단백질의 발현을 가능하게 하는 프로모터에 작동가능하게 연결된 것인 폴리뉴클레오티드; 및

b) 표적 DNA 서열에 혼성화할 수 있는 하나 이상의 가이드 RNA (gRNA), 또는 상기 gRNA를 코딩하는 폴리뉴클레오티드로서, 여기서 gRNA를 코딩하는 상기 폴리뉴클레오티드는 세포에서 gRNA의 발현을 가능하게 하는 프로모터에 작동가능하게 연결된 것인 폴리뉴클레오티드

를 도입하는 것을 포함하고;

이로써 융합 단백질 및 gRNA가 원인 돌연변이의 게놈 위치를 표적화하고 게놈 서열을 변형시켜 원인 돌연변이를 제거하는 것인

방법.

청구항 113

제112항에 있어서, 상기 융합 단백질의 RGN이 니카제인 방법.

청구항 114

제112항 또는 제113항에 있어서, USP가 서열식별번호: 33-39 중 어느 하나의 서열을 포함하는 것인 방법.

청구항 115

제112항 내지 제114항 중 어느 한 항에 있어서, 게놈 변형이 표적 DNA 서열 내에 적어도 하나의 뉴클레오티드의 C>T 돌연변이를 도입하는 것을 포함하는 것인 방법.

청구항 116

제112항 내지 제115항 중 어느 한 항에 있어서, 세포가 동물 세포인 방법.

청구항 117

제116항에 있어서, 동물 세포가 포유동물 세포인 방법.

청구항 118

제117항에 있어서, 세포가 개, 고양이, 마우스, 래트, 토끼, 말, 양, 염소, 소, 돼지 또는 인간으로부터 유래된 것인 방법.

청구항 119

제112항 내지 제118항 중 어느 한 항에 있어서, 원인 돌연변이의 교정이 정지 코돈을 도입하는 것을 포함하는 것인 방법.

청구항 120

a) (i) DNA-결합 폴리펩티드; 및 (ii) 데아미나제를 포함하는 융합 단백질; 또는 융합 단백질을 코딩하는 핵산 분자; 및

b) 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP);

또는 USP를 코딩하는 핵산 분자
를 포함하는 조성물.

청구항 121

제120항에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 조성물.

청구항 122

제120항 또는 제121항에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 조성물.

청구항 123

제120항 또는 제121항에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 조성물.

청구항 124

제123항에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 조성물.

청구항 125

제124항에 있어서, RGN이 RGN 니카제인 조성물.

청구항 126

융합 단백질을 코딩하는 핵산 분자 및 우라실 안정화 폴리펩티드 (USP)를 코딩하는 핵산 분자를 포함하는 벡터로서, 여기서 상기 융합 단백질은 DNA-결합 폴리펩티드 및 데아미나제를 포함하고, 상기 USP는 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 것인 벡터.

청구항 127

제126항에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 벡터.

청구항 128

제126항 또는 제127항에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 벡터.

청구항 129

제126항 또는 제127항에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 벡터.

청구항 130

제129항에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 벡터.

청구항 131

제130항에 있어서, RGN이 RGN 니카제인 벡터.

청구항 132

제126항 내지 제131항 중 어느 한 항의 벡터를 포함하는 세포.

청구항 133

a) (i) DNA-결합 폴리펩티드; 및 (ii) 데아미나제를 포함하는 융합 단백질; 또는 융합 단백질을 코딩하는 핵산 분자; 및

b) 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP); 또는 USP를 코딩하는 핵산 분자를 포함하는 세포.

청구항 134

제133항에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 세포.

청구항 135

제134항에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 세포.

청구항 136

제134항에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 세포.

청구항 137

제136항에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 세포.

청구항 138

제137항에 있어서, RGN이 RGN Nickase인 세포.

발명의 설명

기술 분야

<관련 출원에 대한 상호 참조>

본 출원은 2020년 7월 15일에 출원된 미국 가출원 번호 63/052,175를 우선권 주장하며, 이는 그 전문이 본원에 참조로 포함된다.

<서열 목록에 관한 진술>

본 출원과 연관된 서열 목록은 종이 사본 대신 ASCII 포맷으로 제공되며, 본 명세서에 참조로 포함된다. L103438_1220WO_0106_5_SL.txt로 명명된 ASCII 카피는 크기가 658,586 바이트이고, 2021년 7월 14일에 생성되었고, EFS-웹을 통해 전자적으로 제출되었다.

<기술분야>

본 발명은 분자 생물학 및 유전자 편집 분야에 관한 것이다.

배경 기술

표적화된 게놈 편집 또는 변형은 기초 및 응용 연구를 위한 중요한 도구가 되고 있다. 초기 방법은 뉴클레아제, 예컨대 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN의 조작을 수반하였고, 각각의 특정한 표적 서열에 대해 특이적인 조작된, 프로그래밍가능한, 서열-특이적인 DNA-결합 도메인을 갖는 키메라 뉴클레아제의 생성을 필요로 한다. RNA-가이드된 뉴클레아제 (RGN), 예컨대 CRISPR-Cas 박테리아 시스템의 클러스터링된 규칙적 간격의 짧은 회문식 반복부 (CRISPR)-연관 (Cas) 단백질은 특정 표적 서열과 특이적으로 혼성화하는 가이드 RNA와 뉴클레아제를 복합체화함으로써 특정 서열의 표적화를 허용한다. 표적-특이적인 가이드 RNA의 생성은 각각의 표적 서열에 대한 키메라 뉴클레아제를 생성하는 것에 비해 비용이 적게 들고 더 효율적이다. 이러한 RNA-가이드된 뉴클레아제를 사용하여, 오류-유발 비-상동 말단-연결 (NHEJ)을 통해 복구되는 서열-특이적인 이중-가닥 파손의 도입에 의해 게놈을 편집하여 특이적인 게놈 위치에 돌연변이를 도입시킬 수 있다.

추가적으로, RGN은 표적화된 DNA 편집 접근법에 유용하다. 핵산 서열의 표적화된 편집, 예를 들어 게놈 DNA에 특이적 변형의 도입을 가능하게 하는 표적화된 절단은 유전자 기능 및 유전자 발현을 연구하기 위한 고도로 미묘한 접근을 가능하게 한다. 이러한 표적화된 편집은 또한 인간에서 유전 질환을 표적화하기 위해 또는 작물 식물의 게놈에 농경학상 유익한 돌연변이를 도입하기 위해 배치될 수 있다. 게놈 편집 도구의 개발은 유전자

편집-기반 포유동물 치료제 및 농업생명공학에 대한 새로운 접근법을 제공한다.

발명의 내용

[0009] 표적 DNA 분자를 변형시키기 위한 조성물 및 방법이 제공된다. 조성물은 관심 표적 DNA 분자를 변형시키는 데 사용된다. 제공된 조성물은 우라실 안정화 폴리펩티드를 포함한다. 또한, DNA-결합 폴리펩티드, 데아미나제 폴리펩티드 및 우라실 안정화 폴리펩티드를 포함하는 융합 단백질이 제공된다. 제공된 조성물은 또한 우라실 안정화 폴리펩티드 또는 융합 단백질을 코딩하는 핵산 분자, 및 핵산 분자를 포함하는 벡터 및 숙주 세포를 포함한다. 본원에 개시된 방법은 관심 표적 DNA 분자 내의 관심 표적 서열에 결합하고, 관심 표적 DNA 분자를 변형시키는 것에 관한 것이다.

발명을 실시하기 위한 구체적인 내용

[0010] 본원에 제시된 본 발명의 여러 변형 및 다른 실시양태는 본 발명이 속한 관련 기술분야의 통상의 기술자에게 상기 설명에 제시된 교시의 이익을 갖는다는 것을 상기시킬 것이다. 따라서, 본 발명은 개시된 구체적 실시양태로 제한되지 않으며, 변형 및 다른 실시양태가 첨부된 청구범위의 범주 내에 포함되는 것으로 의도됨을 이해해야 한다. 구체적 용어가 본원에 사용되지만, 이들은 단지 일반적 및 설명적 의미로 사용되며, 제한의 목적을 위한 것이 아니다.

[0011] I. 개요

[0012] 본 개시내용은 DNA 분자 내의 우라실 잔기를 안정화시키는 우라실 안정화 폴리펩티드 (USP), 및 그를 코딩하는 핵산 분자를 제공한다.

[0013] 염기 편집으로도 지칭되는 표적화된 핵염기 편집은 2016년에 코모르(Komor) 등에 의해 변형된 RNA 가이드된 뉴클레아제 (SpCas9)에 작동가능하게 연결된 시토신 데아미나제 (rAPOBEC1)를 사용하여 개발되었다 (문헌 [Nature 533: 420-424]). 코모르 등에 의해 기재된 시스템에서, 가이드 RNA는 rAPOBEC1-Cas9 융합 단백질을 표적 DNA 서열로 가이드하고, 여기서 rAPOBEC1은 표적 시토신 (C)을 티민 (T)의 염기-쌍형성 특성을 갖는 우라실 (U)로 탈아미노화한다. 이러한 시스템을 사용하여, 표적화된 C>T 돌연변이를 DNA 분자에 도입할 수 있었다.

[0014] 생체내 rAPOBEC1-Cas9 융합체를 사용한 염기 편집의 주요 결점은 세포 우라실 DNA 글리코실라제 (UDG)가 U:G 헤테로듀플렉스 DNA를 인식하고, DNA로부터 우라실의 제거를 촉매하여 무염기성 부위를 남김으로써, indel (삽입 또는 결실) 형성이 또한 관찰되기는 하였지만 가장 흔한 결과로서 U:G 쌍의 C:G 쌍으로의 복구를 갖는 염기-결합 복구를 개시한다는 것이었다. 우라실 DNA 글리코실라제 억제제 (UGI)를 rAPOBEC1-Cas9 융합 단백질 상에 혼입시킴으로써, 우라실은 복제가 일어나고 목적하는 C>T 돌연변이를 도입하기에 충분히 길게 존재하도록 유지되었다.

[0015] 본 발명은 탈아미노화된 시토신에 의해 생성된 우라실을 안정화시킴으로써, 무염기성 부위의 생성이 방지될 수 있고 목적하는 C>T 돌연변이가 도입될 가능성이 더 크다는 것을 발견하였다. 이는 우라실 안정화 단백질 (우라실 안정화 폴리펩티드 또는 USP로도 지칭됨)의 확인에 의해 달성되었다.

[0016] 일부 실시양태에서, USP는 DNA-결합 폴리펩티드, 데아미나제 폴리펩티드 및 우라실 안정화 폴리펩티드를 포함하는 융합 단백질의 일부로서 제공된다. 일부 실시양태에서, DNA-결합 폴리펩티드는 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN이거나 또는 그로부터 유래된다. 일부 실시양태에서, DNA-결합 폴리펩티드는 가이드 RNA (gRNA로도 지칭됨)에 결합하는 RNA-가이드된 뉴클레아제, 예컨대 Cas9 폴리펩티드이며, 이는 다시 가닥 혼성화를 통해 표적 핵산 서열에 결합한다. 다른 실시양태에서, USP는 단독으로 제공된다.

[0017] 일부 실시양태에서, 데아미나제 폴리펩티드는 핵염기, 예컨대 예를 들어 시티딘을 탈아미노화할 수 있는 데아미나제 도메인일 수 있다. 데아미나제에 의한 핵염기의 탈아미노화는 각각의 잔기에서의 점 돌연변이를 유발할 수 있으며, 이는 본원에서 "핵산 편집" 또는 "염기 편집"으로 지칭된다. 따라서, RNA-가이드된 뉴클레아제 (RGN) 폴리펩티드 및 데아미나제 폴리펩티드를 포함하는 융합 단백질은 핵산 서열의 표적화된 편집에 사용될 수 있다.

[0018] 이러한 융합 단백질은 시험관내 DNA의 표적화된 편집, 예를 들어 돌연변이 세포의 생성에 유용하다. 이들 돌연변이 세포는 식물 또는 동물에 존재할 수 있다. 이러한 융합 단백질은 또한 표적화된 돌연변이의 도입, 예를 들어 생체의 포유동물 세포, 예를 들어 동일한 또는 또 다른 대상체에게 후속적으로 재도입되는 대상체로부터 수득된 세포에서의 유전적 결함의 교정; 및 생체내 표적화된 돌연변이의 도입, 예를 들어 포유동물 대상체 내

질환-연관 유전자에서의 유전적 결함의 교정 또는 탈활성화 돌연변이의 도입에 유용할 수 있다. 이러한 융합 단백질은 또한 식물 세포에서의 표적화된 돌연변이의 도입, 예를 들어 유익한 또는 농경학상 가치있는 형질 또는 대립유전자의 도입에 유용할 수 있다.

[0019] 용어 "단백질", "펩티드" 및 "폴리펩티드"는 본원에서 상호교환가능하게 사용되고, 펩티드 (아미드) 결합에 의해 함께 연결된 아미노산 잔기의 중합체를 지칭한다. 상기 용어는 임의의 크기, 구조 또는 기능의 단백질, 펩티드 또는 폴리펩티드를 지칭한다. 전형적으로, 단백질, 펩티드 또는 폴리펩티드는 적어도 3개의 아미노산 길이일 것이다. 단백질, 펩티드 또는 폴리펩티드는 개별 단백질 또는 단백질의 집합을 지칭할 수 있다. 단백질, 펩티드, 또는 폴리펩티드 내의 아미노산 중 하나 이상은, 예를 들어 화학 물질, 예컨대 탄수화물 기, 히드록실 기, 포스페이트 기, 파르네실 기, 이소파르네실 기, 지방산 기, 접합, 관능화, 또는 다른 변형을 위한 링커 등의 부가에 의해 변형될 수 있다. 단백질, 펩티드 또는 폴리펩티드는 또한 단일 분자일 수 있거나 또는 다중-분자 복합체일 수 있다. 단백질, 펩티드 또는 폴리펩티드는 단지 자연 발생 단백질 또는 펩티드의 단편일 수 있다. 단백질, 펩티드 또는 폴리펩티드는 자연 발생, 재조합 또는 합성, 또는 그의 임의의 조합일 수 있다.

[0020] 본원에 제공된 단백질 중 임의의 것은 관련 기술분야에 공지된 임의의 방법에 의해 생성될 수 있다. 예를 들어, 본원에 제공된 단백질은 펩티드 링커를 포함하는 융합 단백질에 특히 적합한 재조합 단백질 발현 및 정제를 통해 생성될 수 있다. 재조합 단백질 발현 및 정제를 위한 방법은 널리 공지되어 있고, 문헌 [Green and Sambrook, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012))] (그의 전체 내용은 본원에 참조로 포함됨)에 기재된 것들을 포함한다.

[0021] II. 우라실 안정화 단백질

[0022] 신규 우라실-안정화 폴리펩티드 (USP)는 서열식별번호: 1-16으로서 본원에 개시 및 제시된다. 본원에 기재된 USP는 DNA 분자에서 우라실을 안정화시키는 것이 요구되는 적용에 유용하다.

[0023] 본원에 사용된 용어 "우라실 안정화 단백질", "우라실 안정화 폴리펩티드" 및 "USP"는 우라실 안정화 활성을 갖는 폴리펩티드를 지칭한다. 본원에 사용된 용어 "우라실 안정화 활성"은 데아미나제에 의한 핵산 분자 내 적어도 1개의 시티딘, 데옥시시티딘 또는 시토신의 티미딘, 데옥시티미딘 또는 티민으로의 돌연변이율을 분자 (예를 들어, 우라실 안정화 폴리펩티드)의 부재 하의 데아미나제에 의한 돌연변이율과 비교하여 증가시키는 분자 (예를 들어, 폴리펩티드)의 능력을 지칭한다. 이론 또는 작용 메카니즘에 얽매이지는 않지만, 본원에 개시된 USP는 복제가 일어나고 목적하는 C>T 돌연변이를 도입하기에 충분한 기간 동안 시티딘, 데옥시시티딘 또는 시토신 염기의 탈아미노화를 통해 생성된 단일-가닥 DNA 중 우라실의 존재를 유지함으로써 기능할 수 있는 것으로 여겨진다. 우라실 안정화 활성은 우라실 DNA 글리코실라제의 억제, 염기 절제 복구 경로 또는 미스-매치 복구 메카니즘을 통해 발생할 수 있다.

[0024] 일부 실시양태에서, 우라실 안정화 활성을 보유하는 본원에 개시된 USP 또는 그의 활성 변이체 또는 단편은 데아미나제에 의한 핵산 분자 내 적어도 1개의 시티딘, 데옥시시티딘 또는 시토신의 티미딘, 데옥시티미딘 또는 티민으로의 돌연변이율을 USP의 부재 하의 데아미나제에 의한 돌연변이율과 비교하여 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 100%, 적어도 150%, 적어도 200% 또는 그 초과만큼 증가시킨다. 반대로, 데아미나제에 의한 핵산 분자 내 적어도 1개의 시티딘, 데옥시시티딘 또는 시토신의 티미딘, 데옥시티미딘 또는 티민 이외의 임의의 핵염기 (즉, 구아노신, 데옥시구아노신, 구아닌, 아데노신, 데옥시아데노신, 아데닌)로의 돌연변이율은 본원에 개시된 USP 또는 그의 활성 변이체 또는 단편에 의해 USP의 부재 하의 데아미나제에 의한 돌연변이율과 비교하여 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 그 초과만큼 감소된다. 시티딘, 데옥시시티딘 또는 시토신의 또 다른 핵염기로의 돌연변이율의 증가 또는 감소는 USP의 존재 또는 부재 하에 특정한 데아미나제의 돌연변이율을 특정한 핵염기와 비교함으로써 측정될 수 있다. 데아미나제가 DNA-결합 폴리펩티드와의 융합을 통해 핵산 분자의 특정 영역에 표적화된 실시양태에서, DNA-결합 폴리펩티드가 결합하는 표적 서열 내의 또는 그에 인접한 시티딘, 데옥시시티딘 또는 시토신의 돌연변이율은 폴리머라제 연쇄 반응 (PCR), 제한 단편 길이 다형성 (RFLP) 또는 (DNA) 서열분석을 포함한 관련 기술분야에 공지된 임의의 방법을 사용하여 측정될 수 있다.

[0025] 본원에 개시된 신규 USP 또는 우라실 안정화 활성을 보유하는 그의 활성 변이체 또는 단편은 데아미나제-DNA-결합 폴리펩티드 융합체의 일부로서 세포에 도입될 수 있고/거나, DNA-결합 폴리펩티드-데아미나제 융합체 또는

DNA-결합 폴리펩티드-데아미나제-USP 융합체와 공동-발현되어 표적 DNA 분자에 목적하는 C>T 돌연변이를 도입하는 효율을 증가시킬 수 있다. 우라실 안정화 활성을 보유하는 본원에 개시된 USP는 서열식별번호: 1-16 중 임의의 것의 아미노산 서열 또는 그의 변이체 또는 단편을 갖는다. 일부 실시양태에서, USP는 서열식별번호: 1-16 중 임의의 것의 아미노산 서열에 대해 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98% 또는 적어도 99% 동일성을 갖는 아미노산 서열을 갖는다. 특정한 실시양태에서, USP는 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함한다. 다른 실시양태에서, USP는 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성을 갖는 아미노산 서열을 포함한다. 또 다른 실시양태에서, USP는 서열식별번호: 6에 대해 적어도 82% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0026] III. 융합 단백질

[0027] 본 개시내용의 일부 측면은 DNA-결합 폴리펩티드 및 데아미나제 폴리펩티드, 및 일부 실시양태에서 USP 폴리펩티드를 포함하는 융합 단백질을 제공한다. 이러한 융합 단백질은 시험관내, 생체의 또는 생체내 DNA의 표적화된 편집에 유용하다.

[0028] 본원에 사용된 용어 "융합 단백질"은 적어도 2종의 상이한 단백질로부터의 단백질 도메인을 포함하는 하이브리드 폴리펩티드를 지칭한다. 융합 단백질은 상이한 도메인, 예를 들어 DNA-결합 도메인 및 데아미나제를 포함할 수 있다. 일부 실시양태에서, 융합 단백질은 핵산, 예를 들어 RNA와 복합체로 존재하거나 또는 이와 회합된다.

[0029] 데아미나제 폴리펩티드는 핵염기, 예컨대 예를 들어 시티딘을 탈아미노화할 수 있는 데아미나제 도메인을 포함한다. 데아미나제에 의한 핵염기의 탈아미노화는 각각의 잔기에서의 점 돌연변이를 유발할 수 있으며, 이는 본원에서 "핵산 편집" 또는 "염기 편집"으로 지칭된다. 따라서, RGN 폴리펩티드 변이체 또는 도메인 및 데아미나제 도메인을 포함하는 융합 단백질은 핵산 서열의 표적화된 편집에 사용될 수 있다. 일부 실시양태에서, 데아미나제는 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 그 초과로 동일한 아미노산 서열을 포함한다. 일부 실시양태에서, 데아미나제는 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 동일한 아미노산 서열을 포함한다. 일부 실시양태에서, 데아미나제는 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 동일한 아미노산 서열을 포함한다. 일부 실시양태에서, 데아미나제는 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 동일한 아미노산 서열을 포함한다. 일부 실시양태에서, 데아미나제는 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 동일한 아미노산 서열을 포함한다. 다른 실시양태에서, 데아미나제는 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 99% 동일한 아미노산 서열을 포함한다. 일부 구체적 실시양태에서, 데아미나제는 서열식별번호: 47, 48 및 76-94 중 어느 하나에 제시된 바와 같은 아미노산 서열을 포함한다.

[0030] 본원에 개시된 융합 단백질은 DNA-결합 폴리펩티드를 포함한다. 본원에 사용된 용어 "DNA-결합 폴리펩티드"는 DNA에 결합할 수 있는 임의의 폴리펩티드를 지칭한다. 특정 실시양태에서, 본원에 개시된 융합 단백질의 DNA-결합 폴리펩티드 부분은 이중-가닥 DNA에 결합한다. 특정한 실시양태에서, DNA-결합 폴리펩티드는 서열-특이적 방식으로 DNA에 결합한다. 본원에 사용된 용어 "서열-특이적" 또는 "서열-특이적 방식"은 특이적 뉴클레오티드 서열과의 선택적 상호작용을 지칭한다.

[0031] 2개의 폴리뉴클레오티드 서열은 2개의 서열이 엄격한 조건 하에 서로 혼성화할 때 실질적으로 상보성인 것으로 고려된다. 마찬가지로, DNA-결합 폴리펩티드는 DNA-결합 폴리펩티드가 엄격한 조건 하에 그의 서열에 결합하는 경우에 서열-특이적 방식으로 특정한 표적 서열에 결합하는 것으로 간주된다. "엄격한 조건" 또는 "엄격한 혼성화 조건"은 2개의 폴리뉴클레오티드 서열 (또는 폴리펩티드가 그의 특이적 표적 서열에 결합함)이 다른 서열보다 검출가능하게 더 큰 정도로 (예를 들어, 배경에 비해 적어도 2배) 서로 결합하도록 하는 의도된 조건이다. 엄격한 조건은 서열-의존적이고, 상이한 상황에서 상이할 것이다. 전형적으로, 엄격한 조건은 pH 7.0 내지 8.3에서 염 농도가 약 1.5 M Na 이온 미만, 전형적으로 약 0.01 내지 1.0 M Na 이온 농도 (또는 다른 염)이고, 온도가 짧은 서열 (예를 들어, 10 내지 50개의 뉴클레오티드)의 경우 적어도 약 30°C이고, 긴 서열 (예를 들어, 50개 초과 뉴클레오티드)의 경우 적어도 약 60°C인 조건일 것이다. 엄격한 조건은 또한 탈안정화제, 예컨대 포름아미드의 첨가로 달성될 수 있다. 예시적인 저 엄격도 조건은 37°C에서 30 내지 35% 포름아미드, 1 M

NaCl, 1% SDS (소듐 도데실 술페이트)의 완충제 용액을 사용한 혼성화, 및 50 내지 55℃에서 1X 내지 2X SSC (20X SSC = 3.0 M NaCl/0.3 M 시트르산삼나트륨)에서의 세척을 포함한다. 예시적인 중간 엄격도 조건은 37℃에서 40 내지 45% 포름아미드, 1.0 M NaCl, 1% SDS에서의 혼성화, 및 55 내지 60℃에서 0.5X 내지 1X SSC에서의 세척을 포함한다. 예시적인 고 엄격도 조건은 37℃에서 50% 포름아미드, 1 M NaCl, 1% SDS에서의 혼성화, 및 60 내지 65℃에서 0.1X SSC에서의 세척을 포함한다. 임의적으로, 세척 완충제는 약 0.1% 내지 약 1% SDS를 포함할 수 있다. 혼성화 지속시간은 일반적으로 약 24시간 미만, 통상적으로 약 4 내지 약 12시간이다. 세척 지속시간은 적어도 평형에 도달하기에 충분한 시간일 것이다.

[0032] T_m은 (규정된 이온 강도 및 pH 하에) 상보적 표적 서열의 50%가 매칭된 서열에 완벽하게 혼성화하는 온도이다. DNA-DNA 하이브리드의 경우, T_m은 문헌 [Meinkoth and Wahl (1984) Anal. Biochem. 138:267-284]의 방정식으로부터 근사화될 수 있다: $T_m = 81.5^{\circ}\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\% \text{ 형태}) - 500/L$; 여기서 M은 1가 양이온의 몰농도이고, %GC는 DNA 중 구아노신 및 시토신 뉴클레오티드의 백분율이고, % 형태는 혼성화 용액 중 포름아미드의 백분율이고, L은 염기 쌍 중 하이브리드의 길이이다. 일반적으로, 엄격한 조건은 규정된 이온 강도 및 pH에서 특이적 서열 및 그의 상보체에 대한 열 융점 (T_m)보다 약 5℃ 더 낮도록 선택된다. 그러나, 심하게 엄격한 조건은 열 융점 (T_m)보다 1, 2, 3, 또는 4℃ 더 낮은 혼성화 및/또는 세척을 이용할 수 있고; 중간 정도의 엄격한 조건은 열 융점 (T_m)보다 6, 7, 8, 9, 또는 10℃ 더 낮은 혼성화 및/또는 세척을 이용할 수 있고; 저 엄격도 조건은 열 융점 (T_m)보다 11, 12, 13, 14, 15, 또는 20℃ 더 낮은 혼성화 및/또는 세척을 이용할 수 있다. 방정식, 혼성화 및 세척 조성물, 및 목적하는 T_m을 사용하여, 관련 기술분야의 통상의 기술자는 혼성화 및/또는 세척 용액의 엄격도에서의 변동이 본질적으로 기재된다는 것을 이해할 것이다. 핵산의 혼성화에 대한 광범위한 지침은 문헌 [Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes, Part I, Chapter 2 (Elsevier, New York); and Ausubel et al., eds. (1995) Current Protocols in Molecular Biology, Chapter 2 (Greene Publishing and Wiley-Interscience, New York)]에서 발견된다. 문헌 [Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York)]을 참고한다.

[0033] 특정 실시양태에서, 서열-특이적 DNA-결합 폴리펩티드는 RNA-가이드된, DNA-결합 폴리펩티드 (RGDBP)이다. 본원에 사용된 용어 "RNA-가이드된, DNA-결합 폴리펩티드" 및 "RGDBP"는 연관 RNA 분자와 표적 DNA 서열의 혼성화를 통해 DNA에 결합할 수 있는 폴리펩티드를 지칭한다.

[0034] 일부 실시양태에서, 융합 단백질의 DNA-결합 폴리펩티드는 뉴클레아제, 예컨대 서열-특이적 뉴클레아제이다. 본원에 사용된 용어 "뉴클레아제"는 핵산 분자 내 뉴클레오티드 사이의 포스포디에스테르 결합의 절단을 촉매하는 효소를 지칭한다. 일부 실시양태에서, DNA-결합 폴리펩티드는 핵산 분자 내의 뉴클레오티드 사이의 포스포디에스테르 결합을 절단할 수 있는 엔도뉴클레아제인 반면에, 다른 실시양태에서, DNA-결합 폴리펩티드는 핵산 분자의 어느 한 말단 (5' 또는 3')에서 뉴클레오티드를 절단할 수 있는 엑소뉴클레아제이다. 일부 실시양태에서, 서열-특이적 뉴클레아제는 메가뉴클레아제, 아연 핑거 뉴클레아제, TAL-이펙터 DNA 결합 도메인-뉴클레아제 융합 단백질 (TALEN), 및 RNA-가이드된 뉴클레아제 (RGN), 또는 뉴클레아제 활성이 감소 또는 억제된 그의 변이체로 이루어진 군으로부터 선택된다.

[0035] 본원에 사용된 용어 "메가뉴클레아제" 또는 "귀소 엔도뉴클레아제"는 길이가 12 내지 40 bp인 이중-가닥 DNA 내의 인식 부위에 결합하는 엔도뉴클레아제를 지칭한다. 메가뉴클레아제의 비제한적 예는 보존된 아미노산 모티프 LAGLIDAG (서열식별번호: 75)를 포함하는 LAGLIDAG 패밀리에 속하는 것이다. 용어 "메가뉴클레아제"는 이량체 또는 단일-쇄 메가뉴클레아제를 지칭할 수 있다.

[0036] 본원에 사용된 용어 "아연 핑거 뉴클레아제" 또는 "ZFN"은 아연 핑거 DNA-결합 도메인 및 뉴클레아제 도메인을 포함하는 키메라 단백질을 지칭한다.

[0037] 본원에 사용된 용어 "TAL-이펙터 DNA 결합 도메인-뉴클레아제 융합 단백질" 또는 "TALEN"은 TAL 이펙터 DNA 결합 도메인 및 뉴클레아제 도메인을 포함하는 키메라 단백질을 지칭한다.

[0038] 본원에 사용된 용어 "RNA-가이드된 뉴클레아제" 또는 "RGN"은 뉴클레아제 활성을 갖는 RNA-가이드된, DNA-결합 폴리펩티드를 지칭한다. RGN은 "RNA-가이드된" 것으로 간주되는 데, 이는 가이드 RNA가 RNA-가이드된 뉴클레아제와 복합체를 형성하여 RNA-가이드된 뉴클레아제가 표적 서열에 결합하고, 일부 실시양태에서는 표적 서열에 단일-가닥 또는 이중-가닥 절단을 도입하도록 지시하기 때문이다.

[0039] 본원에 개시된 조성물 및 방법에 유용한 RGN의 비제한적 예는 각각 2021년 4월 23일 및 2021년 5월 11일에 출원

된 공개 번호 WO 2020/139783, WO 2019/236566, WO 2021/030344, WO/2021/138247, 및 출원 번호 PCT/US2021/028843 및 PCT/US2021/031794에 개시된 것들을 포함하며, 이들 각각은 그 전문이 본원에 참조로 포함된다. 일부 실시양태에서, 본원에 개시된 융합 단백질은 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 그 초과로 동일한 아미노산 서열을 포함하는 RGN을 포함한다. 일부 실시양태에서, 본원에 개시된 융합 단백질은 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80% 동일한 아미노산 서열을 갖는 RGN을 포함한다. 일부 실시양태에서, 본원에 개시된 융합 단백질은 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 85% 동일한 아미노산 서열을 갖는 RGN을 포함한다. 일부 실시양태에서, 본원에 개시된 융합 단백질은 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 90% 동일한 아미노산 서열을 갖는 RGN을 포함한다. 일부 실시양태에서, 본원에 개시된 융합 단백질은 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 95% 동일한 아미노산 서열을 갖는 RGN을 포함한다. 다른 실시양태에서, 본원에 개시된 융합 단백질은 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 99% 동일한 아미노산 서열을 갖는 RGN을 포함한다. 일부 구체적 실시양태에서, 본원에 개시된 융합 단백질은 서열식별번호: 40 및 95-142 중 어느 하나에 제시된 아미노산 서열을 갖는 RGN을 포함한다.

[0040] 본 발명에 따르면, 뉴클레아제-불활성 또는 "사멸"되도록 돌연변이된 RGN 단백질, 예를 들어 dCas9는 본원에서 RNA-가이드된, DNA-결합 폴리펩티드로 언급된다. 하나의 예시적인 적합한 뉴클레아제-불활성 Cas9 도메인은 D10A/H840A Cas9 도메인 돌연변이체이다 (예를 들어, 문헌 [Qi et al., Cell. 2013; 152(5): 1173-83] 참조, 그의 전체 내용은 본원에 참조로 포함됨). 추가적으로, 다른 공지된 RNA 가이드된 뉴클레아제 (RGN)의 적합한 뉴클레아제-불활성 Cas9 도메인이 결정될 수 있다 (예를 들어, 미국 특허 공개 번호 2019/0367949 (그의 전체 내용이 본원에 참조로 포함됨)에 개시된 RGN APG08290.1의 뉴클레아제-불활성 변이체).

[0041] 용어 "RGN 폴리펩티드"는 표적 뉴클레오티드 서열의 단일 가닥만을 절단하는 RGN 폴리펩티드를 포함하며, 이는 본원에서 니카제로 지칭된다. 이러한 RGN은 단일 기능성 뉴클레아제 도메인을 갖는다. RGN 니카제는 자연-발생 니카제일 수 있거나, 또는 추가의 뉴클레아제 도메인 내에서 돌연변이된 이중-가닥 핵산 분자의 두 가닥을 자연적으로 절단하여 이들 돌연변이된 도메인의 뉴클레아제 활성이 감소 또는 제거되어 니카제가 되도록 하는 RGN 단백질일 수 있다. 일부 실시양태에서, 융합 단백질의 니카제 RGN은 RGN이 핵산 듀플렉스의 비-염기 편집된, 표적 가닥 (PAM을 포함하고 gRNA에 염기 쌍형성된 가닥)만을 절단할 수 있게 하는 D10A 돌연변이 (예를 들어, nAPG07433.1 (서열식별번호: 41))를 포함한다. 일부 실시양태에서, 융합 단백질의 니카제 RGN은 서열식별번호: 40 및 95-142 중 어느 하나에서 D10A 돌연변이 또는 그의 등가 돌연변이를 포함한다. 일부 실시양태에서, 융합 단백질의 니카제 RGN은 H840A 돌연변이를 포함하며, 이는 RGN이 핵산 듀플렉스의 염기-편집된, 비-표적 가닥 (PAM을 포함하지 않고 gRNA에 염기 쌍형성되지 않은 가닥)만을 절단할 수 있게 한다. H840A 돌연변이 또는 등가 돌연변이를 포함하는 니카제 RGN은 불활성화된 HNH 도메인을 갖는다. D10A 돌연변이 또는 등가 돌연변이를 포함하는 니카제 RGN은 불활성화된 RuvC 도메인을 갖는다. 데아미나제는 비-표적 가닥에 작용한다. D10A 돌연변이 또는 등가 돌연변이를 포함하는 니카제는 불활성 RuvC 뉴클레아제 도메인을 가지며, DNA의 비-표적화된 가닥, 즉 염기 편집이 바람직한 가닥을 절단할 수 없다.

[0042] 다른 추가의 예시적인 적합한 뉴클레아제 불활성 Cas9 도메인은 D10A/D839A/H840A, 및 D10A/D839A/H840A/N863A 돌연변이 도메인을 포함하나 이에 제한되지는 않는다 (예를 들어, 문헌 [Mali et al., Nature Biotechnology. 2013; 31(9): 833-838] 참조, 그의 전체 내용은 본원에 참조로 포함됨). 니카제로 돌연변이된 추가의 적합한 RGN 단백질은 본 개시내용 및 관련 기술분야의 지식 (예컨대, 예를 들어 PCT 공개 번호 WO 2019/236566에 개시된 RGN)에 기초하여 관련 기술분야의 통상의 기술자에게 명백할 것이고, 본 개시내용의 범주 내에 있다.

[0043] 일부 실시양태에서, RGN 니카제 보유 니카제 활성은 nAPG07433.1 (서열식별번호: 41)에 대해 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 적어도 99.5% 동일한 아미노산 서열을 포함한다.

[0044] 돌연변이를 아미노산 서열에 도입시키기 위한 관련 기술분야에 공지된 임의의 방법, 예컨대 PCR-매개된 돌연변이 유발 및 부위-지정된 돌연변이유발은 니카제 또는 뉴클레아제-데드 RGN을 생성하기 위해 이용될 수 있다. 예를 들어, 미국 공개 번호 2014/0068797 및 미국 특허 번호 9,790,490을 참조하며; 이들 각각은 그 전문이 참조로 포함된다. RNA-가이드된 뉴클레아제 (RGN)는 게놈 내의 단일 부위의 표적화된 조작을 허용하고, 치료 및 연구 용도를 위한 유전자 표적화의 맥락에서 유용하다. 포유동물을 비롯한 다양한 유기체에서, RNA-가이드된 뉴클레아제는 비-상동 말단 연결 또는 상동 재조합을 자극함으로써 게놈 조작에 사용되었다. RGN은 클러스터링된

규칙적 간격의 짧은 회문식 반복부 (CRISPR) RNA-가이드된 뉴클레아제 시스템의 일부로서 가이드 RNA (gRNA)에 의해 표적 서열에 대해 지시된 RNA-가이드된 뉴클레아제인 CRISPR-Cas 단백질, 또는 그의 활성 변이체 또는 단편을 포함한다.

[0045] 본 개시내용의 일부 측면은 RNA-가이드된 DNA-결합 폴리펩티드, 데아미나제 폴리펩티드 및 USP를 포함하는 융합 단백질을 제공한다. 일부 실시양태에서, RNA-가이드된 DNA-결합 폴리펩티드는 RNA-가이드된 뉴클레아제이다. 추가 실시양태에서, RNA-가이드된 뉴클레아제는 자연-발생 CRISPR-Cas 단백질 또는 그의 활성 변이체 또는 단편이다. CRISPR-Cas 시스템은 클래스 I 또는 클래스 II 시스템으로 분류된다. 클래스 II 시스템은 단일 이펙터 뉴클레아제를 포함하고, 유형 II, V 및 VI를 포함한다. 각각의 클래스는 유형 (유형 I, II, III, IV, V, VI)으로 세분되고, 일부 유형은 하위유형 (예를 들어, 유형 II-A, 유형 II-B, 유형 II-C, 유형 V-A, 유형 V-B)으로 추가로 세분된다.

[0046] 특정 실시양태에서, CRISPR-Cas 단백질은 자연-발생 유형 II CRISPR-Cas 단백질 또는 그의 활성 변이체 또는 단편이다. 본원에 사용된 용어 "유형 II CRISPR-Cas 단백질", "유형 II CRISPR-Cas 이펙터 단백질" 또는 "Cas9"는 트랜스-활성화 RNA (tracrRNA)를 필요로 하고, 각각 이중-가닥 DNA 분자의 단일 가닥을 절단하는 것을 담당하는 2개의 뉴클레아제 도메인 (RuvC 및 HNH)을 포함하는 CRISPR-Cas 이펙터 단백질을 지칭한다.

[0047] 다른 실시양태에서, CRISPR-Cas 단백질은 자연-발생 유형 V CRISPR-Cas 단백질 또는 그의 활성 변이체 또는 단편이다. 본원에 사용된 용어 "유형 V CRISPR-Cas 단백질", "유형 V CRISPR-Cas 이펙터 단백질" 또는 "Cas12"는 dsDNA를 절단하고 단일 RuvC 뉴클레아제 도메인 또는 스플릿-RuvC 뉴클레아제 도메인을 포함하며 HNH 도메인이 결여된 CRISPR-Cas 이펙터 단백질을 지칭한다 (문헌 [Zetsche et al. 2015, Cell doi:10.1016/j.cell.2015.09.038; Shmakov et al. 2017, Nat Rev Microbiol doi:10.1038/nrmicro.2016.184; Yan et al. 2018, Science doi:10.1126/science.aav7271; Harrington et al. 2018, Science doi:10.1126/science.aav4294]). Cas12a는 또한 Cpf1로 지칭되고, tracrRNA를 필요로 하지 않지만, 다른 유형 V CRISPR-Cas 단백질, 예컨대 Cas12b는 tracrRNA를 필요로 한다는 것을 주목해야 한다. 대부분의 유형 V 이펙터는 또한 종종 PAM 요건 없이 ssDNA (단일-가닥 DNA)를 표적화할 수 있다 (문헌 [Zetsche et al. 2015; Yan et al. 2018; Harrington et al. 2018]). 용어 "유형 V CRISPR-Cas 단백질"은 분할 RuvC 뉴클레아제 도메인을 포함하는 고유한 RGN, 예컨대 2019년 12월 30일에 출원된 미국 가출원 번호 62/955,014에 개시된 것을 포괄하며, 그의 내용은 그 전문이 참조로 포함된다.

[0048] 또 다른 실시양태에서, CRISPR-Cas 단백질은 자연-발생 유형 VI CRISPR-Cas 단백질 또는 그의 활성 변이체 또는 단편이다. 본원에 사용된 용어 "유형 VI CRISPR-Cas 단백질", "유형 VI CRISPR-Cas 이펙터 단백질" 또는 "Cas13"은 tracrRNA를 필요로 하지 않고, RNA를 절단하는 2개의 HEPN 도메인을 포함하는 CRISPR-Cas 이펙터 단백질을 지칭한다. 용어 "가이드 RNA"는 표적 서열과 혼성화하기도 표적 뉴클레오티드 서열과 충분한 상보성을 가지며 연관 RGN과 표적 뉴클레오티드 서열의 서열-특이적 결합을 지시하는 뉴클레오티드 서열을 지칭한다. CRISPR-Cas RGN에 대해, 각각의 가이드 RNA는, RGN에 결합하고 RGN을 특정한 표적 뉴클레오티드 서열에 결합하도록 가이드할 수 있는 1개 이상의 RNA 분자 (일반적으로, 1 또는 2개)이고, RGN이 니카제 또는 뉴클레아제 활성을 갖는 경우에 또한 표적 뉴클레오티드 서열을 절단한다. 가이드 RNA는 CRISPR RNA (crRNA)를 포함하고, 일부 실시양태에서는 트랜스-활성화 CRISPR RNA (tracrRNA)를 포함한다.

[0049] CRISPR RNA는 스페이서 서열 및 CRISPR 반복 서열을 포함한다. "스페이서 서열"은 관심 표적 뉴클레오티드 서열과 직접 혼성화하는 뉴클레오티드 서열이다. 스페이서 서열은 관심 표적 서열과 완전히 또는 부분적으로 상보적일도록 조작된다. 다양한 실시양태에서, 스페이서 서열은 약 8개 뉴클레오티드 내지 약 30개 뉴클레오티드, 또는 그 초과를 포함할 수 있다. 예를 들어, 스페이서 서열은 약 8, 약 9, 약 10, 약 11, 약 12, 약 13, 약 14, 약 15, 약 16, 약 17, 약 18, 약 19, 약 20, 약 21, 약 22, 약 23, 약 24, 약 25, 약 26, 약 27, 약 28, 약 29, 약 30개 또는 그 초과 뉴클레오티드 길이일 수 있다. 일부 실시양태에서, 스페이서 서열은 약 10 내지 약 26개 뉴클레오티드 길이, 또는 약 12 내지 약 30개 뉴클레오티드 길이이다. 특정한 실시양태에서, 스페이서 서열은 약 30개 뉴클레오티드 길이이다. 일부 실시양태에서, 적합한 정렬 알고리즘을 이용하여 최적으로 정렬될 때 스페이서 서열과 그의 상응하는 표적 서열 사이의 상보성 정도는 약 50%, 약 60%, 약 70%, 약 75%, 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 또는 그 초과, 또는 대략 상기 값 초과이다. 특정한 실시양태에서, 스페이서 서열은 2차 구조를 함유하지 않으며, 이는 mFold (예를 들어, 문헌 [Zuker and Stiegler (1981) Nucleic Acids Res. 9:133-148] 참조) 및 RNAfold (예를 들어, 문헌 [Gruber et al. (2008) Cell 106(1):23-24] 참조)를 포함하나 이에 제한되지는 않는 관련 기술분야에 공지된 임의의 적합한

폴리뉴클레오티드 폴딩 알고리즘을 사용하여 예측될 수 있다.

- [0050] CRISPR RNA 반복 서열은 RGN 분자에 의해 인식되는 구조를 그 자체로 또는 혼성화된 tracrRNA와 함께 형성하는 뉴클레오티드 서열을 포함한다. 다양한 실시양태에서, CRISPR RNA 반복 서열은 약 8개 뉴클레오티드 내지 약 30개 뉴클레오티드, 또는 그 초과를 포함할 수 있다. 예를 들어, CRISPR 반복 서열은 약 8, 약 9, 약 10, 약 11, 약 12, 약 13, 약 14, 약 15, 약 16, 약 17, 약 18, 약 19, 약 20, 약 21, 약 22, 약 23, 약 24, 약 25, 약 26, 약 27, 약 28, 약 29, 약 30개, 또는 그 초과 뉴클레오티드 길이일 수 있다. 일부 실시양태에서, 적합한 정렬 알고리즘을 이용하여 최적으로 정렬될 때 CRISPR 반복 서열과 그의 상응하는 tracrRNA 서열 사이의 상보성 정도는 약 50%, 약 60%, 약 70%, 약 75%, 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 또는 그 초과, 또는 대략 상기 값 초과이다.
- [0051] 일부 실시양태에서, 가이드 RNA는 tracrRNA 분자를 추가로 포함한다. 트랜스-활성화 CRISPR RNA 또는 tracrRNA 분자는 crRNA의 CRISPR 반복 서열에 혼성화하기에 충분한 상보성을 갖는 영역을 포함하는 뉴클레오티드 서열을 포함하며, 이는 본원에서 안티-반복 영역으로 지칭된다. 일부 실시양태에서, tracrRNA 분자는 2차 구조를 갖는 영역 (예를 들어, 스템-루프)을 추가로 포함하거나, 또는 그의 상응하는 crRNA와의 혼성화 시 2차 구조를 형성한다. 특정한 실시양태에서, CRISPR 반복 서열에 완전히 또는 부분적으로 상보적인 tracrRNA의 영역은 분자의 5' 말단에 있고, tracrRNA의 3' 말단은 2차 구조를 포함한다. 이러한 2차 구조 영역은 일반적으로 안티-반복 서열과 인접하여 발견되는 몇몇 헤어핀 구조, 예컨대 넥서스 헤어핀을 포함한다. 종종 tracrRNA의 3' 말단에, 구조 및 개수가 다양할 수 있지만 종종 GC-풍부 Rho-비의존성 전사 종결인자 헤어핀에 이어서 3' 말단에 U의 문자열을 포함하는 말단 헤어핀이 존재한다. 예를 들어, 문헌 [Briner et al. (2014) Molecular Cell 56:333-339, Briner and Barrangou (2016) Cold Spring Harb Protoc; doi: 10.1101/pdb.top090902], 및 미국 공개 번호 2017/0275648을 참고하며, 이들 각각은 그 전문이 본원에 참조로 포함된다.
- [0052] 다양한 실시양태에서, CRISPR 반복 서열에 완전히 또는 부분적으로 상보적인 tracrRNA의 안티-반복 영역은 약 6개 뉴클레오티드 내지 약 30개 뉴클레오티드 또는 그 초과를 포함한다. 예를 들어, tracrRNA 안티-반복 서열과 CRISPR 반복 서열 사이의 염기 쌍형성 영역은 약 6, 약 7, 약 8, 약 9, 약 10, 약 11, 약 12, 약 13, 약 14, 약 15, 약 16, 약 17, 약 18, 약 19, 약 20, 약 21, 약 22, 약 23, 약 24, 약 25, 약 26, 약 27, 약 28, 약 29, 약 30개 또는 그 초과 뉴클레오티드 길이일 수 있다. 특정한 실시양태에서, CRISPR 반복 서열에 완전히 또는 부분적으로 상보적인 tracrRNA의 안티-반복 영역은 약 10개 뉴클레오티드 길이이다. 일부 실시양태에서, 적합한 정렬 알고리즘을 이용하여 최적으로 정렬될 때 CRISPR 반복 서열과 그의 상응하는 tracrRNA 안티-반복 서열 사이의 상보성 정도는 약 50%, 약 60%, 약 70%, 약 75%, 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 또는 그 초과, 또는 대략 상기 값 초과이다.
- [0053] 다양한 실시양태에서, 전체 tracrRNA는 약 60개 뉴클레오티드 내지 약 210개 초과 뉴클레오티드를 포함할 수 있다. 예를 들어, tracrRNA는 약 60, 약 65, 약 70, 약 75, 약 80, 약 85, 약 90, 약 95, 약 100, 약 105, 약 110, 약 115, 약 120, 약 125, 약 130, 약 135, 약 140, 약 150, 약 160, 약 170, 약 180, 약 190, 약 200, 약 210개 또는 그 초과 뉴클레오티드 길이일 수 있다. 특정한 실시양태에서, tracrRNA는 약 100 내지 약 201개 뉴클레오티드 길이, 예컨대 약 95, 약 96, 약 97, 약 98, 약 99, 약 100, 약 105, 약 106, 약 107, 약 108, 약 109, 및 약 100개 뉴클레오티드 길이이다.
- [0054] 가이드 RNA는 RNA-가이드된 뉴클레아제와 복합체를 형성하여, RNA-가이드된 뉴클레아제가 표적 서열에 결합하도록 지시하고, 표적 서열에 단일-가닥 또는 이중-가닥 파손을 도입한다. 표적 서열이 절단된 후에, 파손은 표적 서열의 DNA 서열이 복구 과정 동안에 변형되도록 복구될 수 있다. 본원은 숙주 세포의 DNA에서 표적 서열을 변형시키기 위해 데아미나제에 연결된 뉴클레아제 불활성 또는 니카제인 RNA-가이드된 뉴클레아제의 돌연변이 변이체를 사용하는 방법을 제공한다. 뉴클레아제 활성이 불활성화되거나 유의하게 감소된 RNA-가이드된 뉴클레아제의 돌연변이 변이체는 RNA-가이드된, DNA-결합 폴리펩티드로 지칭될 수 있는 데, 이는 폴리펩티드가 표적 서열에 결합할 수 있지만 반드시 절단할 필요는 없기 때문이다. 이중-가닥 핵산 분자의 단일 가닥만을 절단할 수 있는 RNA-가이드된 뉴클레아제는 본원에서 니카제로 지칭된다.
- [0055] 표적 뉴클레오티드 서열은 RGN에 의해 결합되고, RGN과 연관된 가이드 RNA와 혼성화한다. 이어서, 표적 서열은 폴리펩티드가 니카제로서의 활성을 포괄하는 뉴클레아제 활성을 보유하는 경우에 RGN에 의해 후속적으로 절단될 수 있다.

- [0056] 가이드 RNA는 단일 가이드 RNA 또는 이중-가이드 RNA 시스템일 수 있다. 단일 가이드 RNA는 RNA의 단일 분자 상에 crRNA 및 임의로 tracrRNA를 포함하는 반면에, 이중-가이드 RNA 시스템은 2개의 별개의 RNA 분자 상에 존재하는 crRNA 및 tracrRNA를 포함하며, 이들은 crRNA의 CRISPR 반복 서열의 적어도 일부분 및 crRNA의 CRISPR 반복 서열에 완전히 또는 부분적으로 상보적일 수 있는 tracrRNA의 적어도 일부분을 통해 서로 혼성화된다. 가이드 RNA가 단일 가이드 RNA인 이들 실시양태 중 일부에서, crRNA 및 임의로 tracrRNA는 링커 뉴클레오티드 서열에 의해 분리된다.
- [0057] 일반적으로, 링커 뉴클레오티드 서열은 링커 뉴클레오티드 서열 내의 또는 그의 뉴클레오티드를 포함하는 2차 구조의 형성을 피하기 위해 상보적 염기를 포함하지 않는 것이다. 일부 실시양태에서, crRNA와 tracrRNA 사이의 링커 뉴클레오티드 서열은 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12개 또는 그 초과 뉴클레오티드 길이이다. 특정한 실시양태에서, 단일 가이드 RNA의 링커 뉴클레오티드 서열은 적어도 4개 뉴클레오티드 길이이다.
- [0058] 특정한 실시양태에서, 가이드 RNA는 RNA 분자로서 표적 세포, 소기관 또는 배아에 도입될 수 있다. 가이드 RNA는 시험관내에서 전사되거나 또는 화학적으로 합성될 수 있다. 다른 실시양태에서, 가이드 RNA를 코딩하는 뉴클레오티드 서열은 세포, 소기관 또는 배아에 도입된다. 이들 실시양태 중 일부에서, 가이드 RNA를 코딩하는 뉴클레오티드 서열은 프로모터 (예를 들어, RNA 폴리머라제 III 프로모터)에 작동가능하게 연결된다. 프로모터는 천연 프로모터 또는 가이드 RNA-코딩 뉴클레오티드 서열에 대해 이중일 수 있다.
- [0059] 다양한 실시양태에서, 가이드 RNA는 본원에 기재된 바와 같이 리보핵단백질 복합체로서 표적 세포, 소기관 또는 배아에 도입될 수 있으며, 가이드 RNA는 RNA-가이드된 뉴클레아제 폴리펩티드에 결합된다.
- [0060] 가이드 RNA는 가이드 RNA의 표적 뉴클레오티드 서열에의 혼성화를 통해 연관된 RNA-가이드된 뉴클레아제를 특정한 관심 표적 뉴클레오티드 서열로 유도한다. 표적 뉴클레오티드 서열은 DNA, RNA, 또는 이들 둘의 조합물을 포함할 수 있고, 단일-가닥 또는 이중-가닥일 수 있다. 표적 뉴클레오티드 서열은 게놈 DNA (즉, 염색체 DNA), 플라스미드 DNA, 또는 RNA 분자 (예를 들어, 메신저 RNA, 리보솜 RNA, 전달 RNA, 마이크로 RNA, 소형 간섭 RNA)일 수 있다. 표적 뉴클레오티드 서열은 시험관내에서 또는 세포에서 RNA-가이드된 뉴클레아제에 의해 결합될 수 있다 (그리고 일부 실시양태에서 절단될 수 있다). RGN에 의해 표적화된 염색체 서열은 핵, 색소체 또는 미토콘드리아 염색체 서열일 수 있다. 일부 실시양태에서, 표적 뉴클레오티드 서열은 표적 게놈에서 고유하다.
- [0061] 일부 실시양태에서, 표적 뉴클레오티드 서열은 프로토스페이서 인접 모티프 (PAM)에 인접한다. PAM은 일반적으로 표적 뉴클레오티드 서열로부터 약 1 내지 약 10개 뉴클레오티드 (표적 뉴클레오티드 서열로부터 약 1, 약 2, 약 3, 약 4, 약 5, 약 6, 약 7, 약 8, 약 9, 또는 약 10개 뉴클레오티드 포함) 내에 있다. PAM은 표적 서열의 5' 또는 3'일 수 있다. 일부 실시양태에서, PAM은 표적 서열의 3'이다. 일반적으로, PAM은 약 2-6개 뉴클레오티드의 컨센서스 서열이지만, 특정한 실시양태에서, 1, 2, 3, 4, 5, 6, 7, 8, 9개 또는 그 초과 뉴클레오티드 길이일 수 있다.
- [0062] 그의 상응하는 PAM 서열을 인식할 때, RGN은 특이적인 절단 부위에서 표적 뉴클레오티드 서열을 절단할 수 있다. 본원에서 사용된 바와 같이, 절단 부위는 표적 뉴클레오티드 서열 내의 2개의 특정한 뉴클레오티드로 구성되며, 이들 사이에서 뉴클레오티드 서열이 RGN에 의해 절단된다. 절단 부위는 5' 또는 3' 방향으로 PAM으로부터 1번째 및 2번째, 2번째 및 3번째, 3번째 및 4번째, 4번째 및 5번째, 5번째 및 6번째, 6번째 및 7번째, 7번째 및 8번째, 또는 8번째 및 9번째 뉴클레오티드를 포함할 수 있다. RGN이 표적 뉴클레오티드 서열을 절단하여 엇갈린 말단을 생성할 수 있기 때문에, 일부 실시양태에서, 절단 부위는 폴리뉴클레오티드의 양성 (+) 가닥 상의 PAM으로부터 2개 뉴클레오티드의 거리 및 폴리뉴클레오티드의 음성 (-) 가닥 상의 PAM으로부터 2개 뉴클레오티드의 거리를 기반으로 하여 정의된다.
- [0063] RGN은 융합된 폴리펩티드, 폴리뉴클레오티드, 또는 소분자 페이로드를 특정 게놈 위치에 전달하기 위해 사용될 수 있다. 일부 실시양태에서, 뉴클레아제-불활성 또는 니카제 RGN은 데아미나제 및 또한 본 발명의 USP에 작동가능하게 연결된다.
- [0064] 본원에 사용된 용어 "데아미나제" 또는 "데아미나제 폴리펩티드"는 탈아미노화 반응을 촉매하는 폴리펩티드를 지칭한다. 데아미나제는 자연-발생 데아미나제 효소 또는 그의 활성 단편 또는 변이체일 수 있다. 일부 실시양태에서, 데아미나제는 시티딘 또는 데옥시시티딘의 각각 우라실 또는 데옥시우라실로의 가수분해성 탈아미노화를 촉매하는 시티딘 데아미나제이다. 시티딘 데아미나제는 DNA 또는 RNA에 작용할 수 있고, 전형적으로 단일-가닥 핵산 분자에 작용할 수 있다. 바람직한 실시양태에서, 표적 가닥에 니카제 활성을 갖는 RGN은 표적 가닥

을 Nicking하는 반면에, 상보적인 비-표적 가닥은 데아미나제에 의해 변형된다. 세포 DNA-복구 기구는 주형으로서 변형된 비-표적 가닥을 사용하여 Nicking된 표적 가닥을 복구함으로써 DNA에 돌연변이를 도입할 수 있다.

[0065] 일부 실시양태에서, 시티딘 데아미나제는 아포지단백질 B mRNA-편집 복합체 (APOBEC) 패밀리의 데아미나제이다. 이들 실시양태 중 일부에서, 데아미나제는 APOBEC1 패밀리의 데아미나제이다. 일부 실시양태에서, 시티딘 데아미나제는 활성화-유도된 시티딘 데아미나제 (AID)이다. 일부 실시양태에서, 데아미나제는 ACF1/ASE 데아미나제이다. 특정 실시양태에서, 데아미나제는 아데노신 데아미나제이다. 이들 실시양태 중 일부에서, 데아미나제는 ADAT 패밀리의 데아미나제이다. 추가의 적합한 데아미나제 효소 또는 도메인은 본 개시내용에 기초하여 관련 기술분야의 통상의 기술자에게 명백할 것이다.

[0066] 데아미나제 폴리펩티드의 하나의 예시적인 적합한 유형은, 예를 들어 APOBEC 패밀리의 시티딘 데아미나제이다. 시티딘 데아미나제 효소의 아포지단백질 B mRNA 편집 복합체 (APOBEC) 패밀리는 제어되고 유익한 방식으로 돌연변이유발을 개시하는 역할을 하는 11종의 단백질을 포괄한다 (문헌 [Conticello et al., 2008. Genome Biology, 9(6): 229]). 한 패밀리의 구성원인 활성화-유도된 시티딘 데아미나제 (AID)는 전사-의존성, 가닥-편향된 방식으로 ssDNA 내의 시티딘을 우라실로 전환시킴으로써 항체의 성숙을 담당한다 (문헌 [Reynaud et al., 2003. Nature Immunology, 4(7): 631-638]). 아포지단백질 B 편집 복합체 3 (APOBEC3) 효소는 역전사된 바이러스 ssDNA 내의 시티딘의 탈아미노화를 통해 특정 HIV-1 균주에 대한 인간 세포에 대한 보호를 제공한다 (문헌 [Bhagwat et al., 2004. DNA Repair (Amst), 3(1): 85-9]). 이들 단백질은 모두 촉매 활성을 위해 Zn²⁺-배위 모티프 (HisX-Glu-X₂₃₋₂₆-Pro-Cys-X₂₋₄-Cys) 및 결합된 물 분자를 필요로 한다. Glu 잔기는 탈아미노화 반응에서 친핵성 공격을 위해 물 분자를 수산화아염으로 활성화시키는 작용을 한다. 각각의 패밀리의 구성원은 hAID의 경우 WRC (여기서, W는 A 또는 T이고, R은 A 또는 G임) 내지 hAPOBEC3F의 경우 TTC의 범위로 그 자체의 특정 "핫스팟"에서 우선적으로 탈아미노화된다 (문헌 [Navaratnam et al., 2006. Intl J Hematol 83(3): 195-200]). APOBEC3G의 촉매 도메인의 최근 결정 구조는 6개의 α-나선에 의해 플랭킹된 5-가닥 β-시트 코어로 구성된 2차 구조를 밝혀냈으며, 이는 전체 패밀리에 걸쳐 보존되는 것으로 여겨진다 (문헌 [Holden et al., 2008. Nature 456(7218): 121-124]). 활성 중심 루프는 ssDNA 결합 및 "핫스팟" 동일성의 결정 둘 다를 담당하는 것으로 밝혀졌다 (문헌 [Chelico et al., 2009. J Biol Chem 284(41): 27761-27765]). 이들 효소의 과다발현은 게놈 불안정성 및 암과 연관되었고, 따라서 서열-특이적 표적화의 중요성을 강조한다 (문헌 [Pham et al., 2005. Biochem 44(8): 2703-2715]). 일부 실시양태에서, 데아미나제 폴리펩티드는 시티딘을 탈아미노화하여 우라실을 생성할 수 있는 데아미나제 폴리펩티드일 수 있다. 데아미나제에 의한 핵염기의 탈아미노화는 각각의 잔기에서 점 돌연변이를 유발하여 DNA 분자를 변형시킬 수 있다. 이러한 변형 작용은 본원에서 핵산 편집 또는 염기 편집으로도 지칭된다. 따라서, Cas9 변이체 또는 도메인, 데아미나제 도메인 및 USP 도메인을 포함하는 융합 단백질은 핵산 서열의 표적화된 편집에 사용될 수 있다.

[0067] 일부 실시양태에서, 본 발명의 데아미나제 및 USP에 융합된 뉴클레아제 불활성 RGN 또는 Nickase RGN은 핵산 분자 (즉, 표적 핵산 분자)의 특정한 위치에 표적화될 수 있고, 이는 일부 실시양태에서 특정한 게놈 유전자좌에서 목적하는 서열의 발현을 변경시킨다. 일부 실시양태에서, 표적 서열에 대한 융합 단백질의 결합은 뉴클레오타이드 염기의 탈아미노화를 유발하여, 하나의 뉴클레오타이드 염기에서 또 다른 것으로의 전환을 유발한다. 일부 실시양태에서, 표적 서열에 대한 이러한 융합 단백질의 결합은 표적 서열에 인접한 뉴클레오타이드 염기의 탈아미노화를 유발한다. 본원에 개시된 조성물 및 방법을 사용하여 탈아미노화되고 돌연변이된 표적 서열에 인접한 뉴클레오타이드 염기는 표적 핵산 분자 내의 표적 서열 (gRNA에 의해 결합됨)의 5' 또는 3' 말단으로부터 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 또는 100개 염기 쌍일 수 있다. 본 개시내용의 일부 측면은 (i) 뉴클레아제-불활성 또는 Nickase RGN 폴리펩티드; (ii) 데아미나제 폴리펩티드; 및 (iii) 우라실 안정화 폴리펩티드를 포함하는 융합 단백질을 제공한다.

[0068] 본 개시내용은 다양한 구성의 융합 단백질을 제공한다. 일부 실시양태에서, 데아미나제 폴리펩티드는 RGN 폴리펩티드의 N-말단에 융합된다. 일부 실시양태에서, 데아미나제 폴리펩티드는 RGN 폴리펩티드의 C-말단에 융합된다.

[0069] 일부 실시양태에서, USP 도메인, 데아미나제 도메인 및 RNA-가이드된, DNA-결합 폴리펩티드는 링커를 통해 서로 융합된다. 특정한 적용을 위한 데아미나제 활성에 대한 최적 길이를 달성하기 위해 융합 단백질의 3개의 기능적 도메인 사이에 다양한 링커 길이 및 가요성이 사용될 수 있다 (예를 들어, 형태 (GGGGS)_n 및 (G)_n의 매우 가요성인 링커로부터 형태 (EAAAK)_n 및 (XP)_n의 보다 강성인 링커에 이르는 범위임). 본원에 사용된 용어 "링커"

는 2개의 분자 또는 모이어티, 예를 들어 결합 도메인 및 뉴클레아제의 절단 도메인을 연결하는 화학적 기 또는 분자를 지칭한다. 일부 실시양태에서, 링커는 RNA 가이드된 뉴클레아제 및 데아미나제를 연결한다. 일부 실시양태에서, 링커는 dCas9 및 데아미나제를 연결한다. 전형적으로, 링커는 2개의 기, 분자 또는 다른 모이어티 사이에 위치하거나, 또는 그에 의해 플랭킹되고, 공유 결합을 통해 서로 연결되어 2개를 연결한다. 일부 실시양태에서, 링커는 아미노산 또는 복수의 아미노산 (예를 들어, 펩티드 또는 단백질)이다. 일부 실시양태에서, 링커는 유기 분자, 기, 중합체, 또는 화학적 모이어티이다. 일부 실시양태에서, 링커는 5-100개 아미노산 길이, 예를 들어 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150 또는 150-200개 아미노산 길이이다. 보다 길거나 짧은 링커가 또한 고려된다.

[0070] 일부 실시양태에서, 링커는 (GGGS)_n, (G)_n, (EAAAK)_n, 또는 (XP)_n 모티프, 또는 이들 중 임의의 것의 조합을 포함하며, 여기서 n은 독립적으로 1 내지 30의 정수이다. 일부 실시양태에서, n은 독립적으로 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 또는 30이거나, 또는 1개 초과인 링커 또는 1개 초과인 링커 모티프가 존재하는 경우에 그의 임의의 조합이다. 추가의 적합한 링커 모티프 및 링커 배위는 관련 기술분야의 통상의 기술자에게 명백할 것이다. 일부 실시양태에서, 적합한 링커 모티프 및 배위는 문헌 [Chen et al., 2013 (Adv Drug Deliv Rev. 65(10):1357-69, 그의 전체 내용은 본원에 참조로 포함됨)]에 기재된 것들을 포함한다. 추가의 적합한 링커 서열은 관련 기술분야의 통상의 기술자에게 명백할 것이다.

[0071] 일부 실시양태에서, 본원에 제공된 예시적인 융합 단백질의 일반적 아키텍처는 구조: [NH₂]-[데아미나제]-[RGN 폴리펩티드]-[USP]-[COOH]; [NH₂]-[USP]-[데아미나제]-[RGN 폴리펩티드]-[COOH]; [NH₂]-[USP]-[RGN 폴리펩티드]-[데아미나제]-[COOH]; [NH₂]-[RGN 폴리펩티드]-[데아미나제]-[USP]-[COOH]; [NH₂]-[RGN 폴리펩티드]-[USP 폴리펩티드]-[데아미나제 폴리펩티드]-[COOH]; 또는 [NH₂]-[데아미나제 폴리펩티드]-[USP 폴리펩티드]-[RGN 폴리펩티드]-[COOH]를 포함하며, 여기서 NH₂는 융합 단백질의 N-말단이고, COOH는 융합 단백질의 C-말단이다.

[0072] 본 개시내용의 일부 측면은 USP 도메인을 포함하지 않는 유사한 융합 단백질에 비해 증가된 C→T 핵염기 편집 효율을 갖는 데아미나제-RGN-USP 융합 단백질, 데아미나제-뉴클레아제 불활성 RGN-USP 융합 단백질 및 데아미나제-닉카제 RGN-USP 융합 단백질을 제공한다.

[0073] 일부 실시양태에서, 융합 단백질은 구조: [NH₂]-[데아미나제]-[뉴클레아제-불활성 RGN]-[USP]-[COOH]; [NH₂]-[데아미나제 폴리펩티드]-[USP]-[뉴클레아제-불활성 RGN]-[COOH]; [NH₂]-[USP]-[데아미나제]-[뉴클레아제-불활성 RGN]-[COOH]; [NH₂]-[USP]-[뉴클레아제-불활성 RGN]-[데아미나제]-[COOH]; [NH₂]-[뉴클레아제-불활성 RGN]-[데아미나제]-[USP]-[COOH]; 또는 [NH₂]-[뉴클레아제-불활성 RGN]-[USP]-[데아미나제]-[COOH]를 포함한다. "뉴클레아제-불활성 RGN"은 뉴클레아제-불활성으로 돌연변이된 임의의 CRISPR-Cas 단백질을 비롯한 임의의 RGN을 나타내는 것으로 이해되어야 한다. 또한, "USP"는 하나 이상의 USP 폴리펩티드를 나타내는 것으로 이해되어야 한다.

[0074] 다른 실시양태에서, 융합 단백질은 구조: [NH₂]-[데아미나제]-[RGN 닉카제]-[USP]-[COOH]; [NH₂]-[데아미나제]-[USP]-[RGN 닉카제]-[COOH]; [NH₂]-[USP]-[데아미나제]-[RGN 닉카제]-[COOH]; [NH₂]-[USP]-[RGN 닉카제]-[데아미나제]-[COOH]; [NH₂]-[RGN 닉카제]-[데아미나제]-[USP]-[COOH]; 또는 [NH₂]-[RGN 닉카제]-[USP]-[데아미나제]-[COOH]를 포함한다. "RGN 닉카제"는 닉카제로서 활성인 것으로 돌연변이된 임의의 CRISPR-Cas 단백질을 비롯한 임의의 RGN을 나타내는 것으로 이해되어야 한다. 또한, "USP"는 하나 이상의 USP 폴리펩티드를 나타내는 것으로 이해되어야 한다.

[0075] 일부 실시양태에서, 융합 단백질은 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 시티딘 데아미나제, 서열식별번호: 40, 41, 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 RGN (또는 그의 닉카제), 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 USP를 포함한다.

[0076] 일부 실시양태에서, 융합 단백질은 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 시티딘 데아미나제, 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 85% 서열 동일성을

갖는 RGN (또는 그의 닉카제), 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 USP를 포함한다.

[0077] 일부 실시양태에서, 융합 단백질은 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 시티딘 테아미나제, 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 RGN (또는 그의 닉카제), 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 USP를 포함한다.

[0078] 일부 실시양태에서, 융합 단백질은 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 시티딘 테아미나제, 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 RGN (또는 그의 닉카제), 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 USP를 포함한다.

[0079] 일부 실시양태에서, 융합 단백질은 서열식별번호: 47, 48 및 76-94 중 어느 하나에 제시된 아미노산 서열을 갖는 시티딘 테아미나제, 서열식별번호: 40, 41 및 95-142 중 어느 하나에 제시된 아미노산 서열을 갖는 RGN (또는 그의 닉카제), 및 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열을 갖는 USP를 포함한다.

[0080] 일부 실시양태에서, 상기 일반적 아키텍처에 사용된 "-"는 임의적인 링커 서열의 존재를 나타낸다. 일부 실시양태에서, 본원에 제공된 융합 단백질은 링커 서열을 포함하지 않는다. 일부 실시양태에서, 임의적인 링커 서열 중 적어도 하나가 존재한다.

[0081] 존재할 수 있는 다른 예시적인 특색은 국제화 서열, 예컨대 핵 국제화 서열, 세포질 국제화 서열, 유출 서열, 예컨대 핵 유출 서열, 또는 다른 국제화 서열, 뿐만 아니라 융합 단백질의 가용화, 정제 또는 검출에 유용한 서열 태그이다. 본원에 제공된 단백질 태그의 적합한 국제화 신호 서열 및 서열은 비오틴 카르복실라제 담체 단백질 (BCCP) 태그, myc-태그, 칼모듈린-태그, FLAG-태그, 헤마글루티닌 (HA)-태그, 히스티딘 태그 또는 His-태그로도 지칭되는 폴리히스티딘 태그, 말토스 결합 단백질 (MBP)-태그, nus-태그, 글루타티온-S-트랜스퍼라제 (GST)-태그, 녹색 형광 단백질 (GFP)-태그, 티오레독신-태그, S-태그, 소프태그 (예를 들어, 소프태그 1, 소프태그 3), 스트랩태그, 비오틴 리가제 태그, FIAsh 태그, V5 태그, 및 SBP-태그를 포함하나 이에 제한되지는 않는다. 추가의 적합한 서열은 관련 기술분야의 통상의 기술자에게 명백할 것이다.

[0082] 특정 실시양태에서, 본원에 개시된 융합 단백질은 융합 단백질의 세포 흡수를 용이하게 하는 적어도 하나의 세포-침투 도메인을 포함한다. 세포-관통 도메인은 관련 기술분야에 공지되어 있고, 일반적으로 양으로 하전된 아미노산 잔기 (즉, 다가양이온성 세포-관통 도메인), 교호 극성 아미노산 잔기 및 비-극성 아미노산 잔기 (즉, 양친매성 세포-관통 도메인), 또는 소수성 아미노산 잔기 (즉, 소수성 세포-관통 도메인)의 스트레치를 포함한다 (예를 들어, 문헌 [Milletti F. (2012) Drug Discov Today 17:850-860] 참조). 세포-투과 도메인의 비제한적인 예는 인간 면역결핍 바이러스 1로부터의 트랜스-활성화 전사 활성화제 (TAT)이다.

[0083] 일부 실시양태에서, 본원에 제공된 USP 또는 융합 단백질은 핵 국제화 서열 (NLS)을 추가로 포함한다. 핵 국제화 신호, 색소체 국제화 신호, 미토콘드리아 국제화 신호, 이중-표적화 국제화 신호, 및/또는 세포-침투 도메인은 아미노-말단 (N-말단), 카르복실-말단 (C-말단), 또는 융합 단백질의 내부 위치에 위치할 수 있다.

[0084] 일부 실시양태에서, NLS는 융합 단백질 또는 USP의 N-말단에 융합된다. 일부 실시양태에서, NLS는 융합 단백질 또는 USP의 C-말단에 융합된다. 일부 실시양태에서, NLS는 융합 단백질의 USP의 N-말단에 융합된다. 일부 실시양태에서, NLS는 융합 단백질의 USP의 C-말단에 융합된다. 일부 실시양태에서, NLS는 융합 단백질의 RGN 폴리펩티드의 N-말단에 융합된다. 일부 실시양태에서, NLS는 융합 단백질의 RGN 폴리펩티드의 C-말단에 융합된다. 일부 실시양태에서, NLS는 융합 단백질의 테아미나제 폴리펩티드의 N-말단에 융합된다. 일부 실시양태에서, NLS는 융합 단백질의 테아미나제 폴리펩티드의 C-말단에 융합된다. 일부 실시양태에서, NLS는 1개 이상의 링커를 통해 융합 단백질 또는 UPS에 융합된다. 일부 실시양태에서, NLS는 링커 없이 융합 단백질 또는 UPS에 융합된다. 일부 실시양태에서, NLS는 본원에 제공되거나 언급된 NLS 서열 중 어느 하나의 아미노산 서열을 포함한다. 일부 실시양태에서, NLS는 서열식별번호: 42 또는 서열식별번호: 45에 제시된 바와 같은 아미노산 서열을 포함한다.

[0085] 일부 실시양태에서, 본원에 제공된 바와 같은 융합 단백질은 우라실 안정화 단백질의 전장 서열, 예를 들어 서열식별번호: 1-16 중 어느 하나를 포함한다. 그러나, 다른 실시양태에서, 본원에 제공된 바와 같은 융합 단백질은 USP의 전장 서열을 포함하지 않고, 단지 그의 단편만을 포함한다. 예를 들어, 일부 실시양태에서, 본원에 제공된 융합 단백질은 RNA-가이드된, DNA-결합 도메인, 테아미나제 도메인, 및 USP의 활성 단편을 추가로 포함

한다.

- [0086] 일부 실시양태에서, 본 발명의 융합 단백질은 RGN, 데아미나제 및 USP를 포함하며, 여기서 USP는 서열식별번호: 1-16 중 임의의 것에 대해 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 동일한 아미노산 서열을 갖는다. 이러한 융합 단백질의 예는 본원의 실시예 섹션에 기재되어 있다.
- [0087] 일부 실시양태에서, 융합 단백질은 1개의 USP 폴리펩티드를 포함한다. 일부 실시양태에서, 융합 단백질은 직접 또는 링커를 통해 작동가능하게 연결된 적어도 2개의 USP 폴리펩티드를 포함한다. 일부 실시양태에서, 융합 단백질은 1개의 USP 폴리펩티드를 포함하고, 제2 USP 폴리펩티드는 융합 단백질과 공동-발현된다.
- [0088] 본 발명의 또 다른 실시양태는 단일 가이드로서 또는 이중 가이드 RNA (집합적으로 gRNA로 지칭됨)로서 융합 단백질 및 가이드 RNA를 포함하는 리보핵단백질 복합체이다.
- [0089] IV. 우라실 안정화 폴리펩티드, 융합 단백질, 및/또는 gRNA를 코딩하는 뉴클레오티드
- [0090] 본 개시내용은 본원에 개시된 우라실 안정화 폴리펩티드 (서열식별번호: 17-32)를 코딩하는 폴리뉴클레오티드를 제공한다. 본 개시내용은 데아미나제 및 DNA-결합 폴리펩티드, 예를 들어 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN을 포함하는 융합 단백질을 코딩하는 폴리뉴클레오티드를 추가로 제공한다. 본 개시내용은 USP, 데아미나제 도메인, 및 RNA-가이드된, DNA-결합 폴리펩티드를 포함하는 융합 단백질을 코딩하는 폴리뉴클레오티드를 추가로 제공한다. 이러한 RNA-가이드된, DNA-결합 폴리펩티드는 RGN 또는 RGN 변이체일 수 있다. 단백질 변이체는 뉴클레아제-불활성 또는 니카제일 수 있다. RGN은 CRISPR-Cas 단백질 또는 그의 활성 변이체 또는 단편일 수 있다. 서열식별번호: 40 및 41은 각각 RGN 및 니카제 RGN 변이체의 비제한적인 예이다. CRISPR-Cas 뉴클레아제의 예는 관련 기술분야에 널리 공지되어 있고, 유사한 상응하는 돌연변이는 또한 니카제 이거나 또는 뉴클레아제 불활성인 돌연변이 변이체를 생성할 수 있다.
- [0091] 본 발명의 한 실시양태는 본원에 기재된 RGN, 데아미나제 및 USP (서열식별번호: 1-16 또는 그의 변이체)를 포함하는 융합 단백질을 코딩하는 폴리뉴클레오티드를 제공한다. 일부 실시양태에서, 제2 폴리뉴클레오티드는 관심 뉴클레오티드 서열을 표적화하기 위해 RGN에 의해 요구되는 가이드 RNA를 코딩한다. 다른 실시양태에서, 가이드 RNA 및 융합 단백질은 동일한 폴리뉴클레오티드에 의해 코딩된다.
- [0092] 용어 "폴리뉴클레오티드"의 사용은 본 개시내용을 DNA를 포함하는 폴리뉴클레오티드로 제한하는 것으로 의도되지 않는다. 관련 기술분야의 통상의 기술자는 폴리뉴클레오티드가 리보뉴클레오티드 (RNA), 및 리보뉴클레오티드 및 데옥시리보뉴클레오티드의 조합을 포함할 수 있음을 알 것이다. 이러한 데옥시리보뉴클레오티드 및 리보뉴클레오티드는 자연 발생 분자 및 합성 유사체 둘 다를 포함한다. 본원에 개시된 폴리뉴클레오티드는 또한 단일-가닥 형태, 이중-가닥 형태, 스템- 및 -루프 구조 등을 포함하나 이에 제한되지 않는 서열의 모든 형태를 포괄한다.
- [0093] 본 발명의 한 실시양태는 서열식별번호: 17-32 중 임의의 것에 대해 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100% 동일한 서열을 포함하는 핵산 분자이며, 여기서 상기 핵산 분자는 우라실 안정화 활성을 갖는 USP를 코딩한다. 핵산 분자는 이중 프로모터 또는 종결인자를 추가로 포함할 수 있다. 핵산 분자는 융합 단백질을 코딩할 수 있으며, 여기서 코딩된 USP는 DNA-결합 폴리펩티드 및/또는 데아미나제에 작동가능하게 연결된다. 일부 실시양태에서, 핵산 분자는 융합 단백질을 코딩하고, 여기서 코딩된 USP는 RGN 및/또는 데아미나제에 작동가능하게 연결된다.
- [0094] 본 발명의 USP를 코딩하는 폴리뉴클레오티드를 포함하는 핵산 분자는 관심 유기체에서의 발현에 대해 코돈 최적화될 수 있다. "코돈-최적화된" 코딩 서열은 특정한 숙주 세포의 선호된 코돈 사용 빈도 또는 전사 조건을 모방하도록 설계된 그의 코돈 사용 빈도를 갖는 폴리뉴클레오티드 코딩 서열이다. 특정한 숙주 세포 또는 유기체에서의 발현은 번역된 아미노산 서열이 변화되지 않도록 하는 핵산 수준에서의 하나 이상의 코돈의 변경의 결과로서 증진된다. 핵산 분자는 전적으로 또는 부분적으로 코돈 최적화될 수 있다. 광범위한 유기체에 대한 선호도 정보를 제공하는 코돈 표 및 다른 참고문헌이 관련 기술분야에서 이용가능하다 (예를 들어, 식물-선호 코돈 용법의 논의에 대해서는 문헌 [Campbell and Gowri (1990) Plant Physiol. 92:1-11] 참조). 식물-선호 유전자를 합성하기 위한 방법은 관련 기술분야에서 이용가능하다. 예를 들어, 미국 특허 번호 5,380,831, 및 5,436,391, 및 문헌 [Murray et al. (1989) Nucleic Acids Res. 17:477-498] (본원에 참조로 포함됨)을 참고한다.

- [0095] 본원에 기재된 USP, 융합 단백질 및/또는 gRNA를 코딩하는 폴리뉴클레오티드는 관심 세포, 소기관, 배아 또는 유기체에서의 시험관내 발현 또는 발현을 위한 발현 카세트에 제공될 수 있다. 카세트는 본원에 제공된 USP 및/또는 USP, RNA-가이드된 DNA-결합 폴리펩티드 및 데아미나제를 포함하는 융합 단백질, 및/또는 gRNA를 코딩하는 폴리뉴클레오티드에 작동가능하게 연결된 5' 및 3' 조절 서열을 포함할 것이며, 이는 상기 폴리뉴클레오티드의 발현을 허용한다. 카세트는 유기체에 동시 형질전환되는 적어도 하나의 추가의 유전자 또는 유전자 요소를 추가로 함유할 수 있다. 추가의 유전자 또는 요소가 포함되는 경우, 성분들은 작동가능하게 연결된다. 용어 "작동가능하게 연결된"은 2개 이상의 요소 사이의 기능적 연결을 의미하는 것으로 의도된다. 예를 들어, 프로모터와 관심 코딩 영역 (예를 들어, USP, 데아미나제, RNA-가이드된 DNA-결합 폴리펩티드 및/또는 gRNA를 코딩하는 영역) 사이의 작동가능한 연결은 관심 코딩 영역의 발현을 허용하는 기능적 연결이다. 작동가능하게 연결된 요소는 인접하거나 인접하지 않을 수 있다. 2개의 단백질 코딩 영역의 결합을 지칭하기 위해 사용되는 경우, 작동가능하게 연결된이란 코딩 영역이 동일한 리딩 프레임 내에 있는 것으로 의도된다. 대안적으로, 추가의 유전자(들) 또는 요소(들)가 다중 발현 카세트 상에 제공될 수 있다. 예를 들어, 단독으로 또는 융합 단백질의 한 성분으로서 본원에 개시된 우라실 안정화 폴리펩티드를 코딩하는 뉴클레오티드 서열은 하나의 발현 카세트 상에 존재할 수 있는 반면에, gRNA를 코딩하는 뉴클레오티드 서열은 별개의 발현 카세트 상에 존재할 수 있다. 또 다른 예는 제1 발현 카세트 상의 본원에 개시된 USP 단독을 코딩하는 뉴클레오티드 서열, USP를 포함하는 융합 단백질을 코딩하는 제2 발현 카세트, 및 제3 발현 카세트 상의 gRNA를 코딩하는 뉴클레오티드 서열을 가질 수 있다. 이러한 발현 카세트에는 조절 영역의 전사 조절 하에 있는 폴리뉴클레오티드의 삽입을 위한 복수의 제한 부위 및/또는 재조합 부위가 제공된다. 선택 마커 유전자를 포함하는 발현 카세트가 또한 존재할 수 있다.
- [0096] 발현 카세트는 5'-3' 전사 방향으로 관심 유기체에서 기능성인 전사 (및 일부 실시양태에서, 번역) 개시 영역 (즉, 프로모터), 본 발명의 USP-코딩 폴리뉴클레오티드, 및 전사 (및 일부 실시양태에서, 번역) 종결 영역 (즉, 종결 영역)을 포함할 것이다. 본 발명의 프로모터는 숙주 세포에서 코딩 서열의 발현을 지시하거나 또는 유도할 수 있다. 조절 영역 (예를 들어, 프로모터, 전사 조절 영역, 및 번역 종결 영역)은 숙주 세포에 대해 또는 서로에 대해 내인성 또는 이종일 수 있다. 서열과 관련하여 본원에 사용된 "이종"은 외래 종으로부터 기원하거나, 또는 동일한 종으로부터 기원하는 경우에 의도적인 인간 개입에 의해 조성 및/또는 게놈 유전자좌에서 그의 천연 형태로부터 실질적으로 변형된 서열이다. 본원에 사용된 바와 같이, 키메라 유전자는 코딩 서열에 이종인 전사 개시 영역에 작동가능하게 연결된 코딩 서열을 포함한다.
- [0097] 편리한 종결 영역은 에이. 투메파시엔스(*A. tumefaciens*)의 Ti-플라스미드, 예컨대 옥토펜 신타제 및 노팔린 신타제 종결 영역으로부터 입수가능하다. 또한, 문헌 [Guerineau et al. (1991) Mol. Gen. Genet. 262:141-144; Proudfoot (1991) Cell 64:671-674; Sanfacon et al. (1991) Genes Dev. 5:141-149; Mogen et al. (1990) Plant Cell 2:1261-1272; Munroe et al. (1990) Gene 91:151-158; Ballas et al. (1989) Nucleic Acids Res. 17:7891-7903; and Joshi et al. (1987) Nucleic Acids Res. 15:9627-9639]을 참고한다.
- [0098] 추가의 조절 신호는 전사 개시 시작 부위, 오피레이터, 활성화제, 인핸서, 다른 조절 요소, 리보솜 결합 부위, 개시 코돈, 종결 신호 등을 포함하나 이에 제한되지는 않는다. 예를 들어, 미국 특허 번호 5,039,523 및 4,853,331; EPO 0480762A2; 문헌 [Sambrook et al. (1992) Molecular Cloning: A Laboratory Manual, ed. Maniatis et al. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.)], 이후 문헌 ["Sambrook 11"]; 문헌 [Davis et al., eds. (1980) Advanced Bacterial Genetics (Cold Spring Harbor Laboratory Press), Cold Spring Harbor, N.Y.], 및 여기에 인용된 참고문헌을 참고한다.
- [0099] 발현 카세트를 제조하기 위해, 적절한 배향으로, 적절한 경우 적절한 리딩 프레임으로 DNA 서열을 제공하도록 다양한 DNA 단편을 조작할 수 있다. 이를 위해, 어댑터 또는 링커를 사용하여 DNA 단편을 결합시킬 수 있거나, 또는 편리한 제한 부위, 불필요한 DNA의 제거, 제한 부위의 제거 등을 제공하기 위해 다른 조작이 수반될 수 있다. 이러한 목적을 위해, 시험관내 돌연변이유발, 프라이머 복구, 제한, 어닐링, 재치환, 예를 들어 전이 및 전환이 수반될 수 있다.
- [0100] 수많은 프로모터가 본 발명의 실시예에 사용될 수 있다. 프로모터는 원하는 결과를 기준으로 선택될 수 있다. 핵산은 구성적, 유도성, 성장 단계-특이적, 세포 유형-특이적, 조직-선택, 조직-특이적, 또는 관심 유기체에서의 발현을 위한 다른 프로모터와 조합될 수 있다. 예를 들어, WO 99/43838, 및 미국 특허 번호 8,575,425; 7,790,846; 8,147,856; 8,586,832; 7,772,369; 7,534,939; 6,072,050; 5,659,026; 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; 및 6,177,611에 제시된 프로모터를 참고

하며; 이들은 본원에 참조로 포함된다.

- [0101] 식물에서의 발현을 위해, 구성적 프로모터는 또한 CaMV 35S 프로모터 (문헌 [Ode11 et al. (1985) Nature 313:810-812]); 벼 액틴 (문헌 [McElroy et al. (1990) Plant Cell 2:163-171]); 유비퀴틴 (문헌 [Christensen et al. (1989) Plant Mol. Biol. 12:619-632 and Christensen et al. (1992) Plant Mol. Biol. 18:675-689]); pEMU (문헌 [Last et al. (1991) Theor. Appl. Genet. 81:581-588]); 및 MAS (문헌 [Velten et al. (1984) EMBO J. 3:2723-2730])를 포함한다.
- [0102] 유도성 프로모터의 예는 저산소 또는 저온 스트레스에 의해 유도가능한 Adh1 프로모터, 열 스트레스에 의해 유도가능한 Hsp70 프로모터, 둘 다 빛에 의해 유도가능한 PPKK 프로모터 및 캅카르복실라제 프로모터이다. 화학적으로 유도성인 프로모터, 예컨대 완화제 유도된 In2-2 프로모터 (미국 특허 번호 5,364,780), 옥신 유도되고 용단층 특이적이지만 캘러스에서도 활성인 Axig1 프로모터 (PCT US01/22169), 스테로이드-반응성 프로모터 (예를 들어, 에스트로겐 유도된 ERE 프로모터, 및 문헌 [Schena et al. (1991) Proc. Natl. Acad. Sci. USA 88:10421-10425 및 McNellis et al. (1998) Plant J. 14(2):247-257]에서의 글루코코르티코이드-유도성 프로모터 참조) 및 테트라시클린-유도성 및 테트라시클린-억제성 프로모터 (예를 들어, 문헌 [Gatz et al. (1991) Mol. Gen. Genet. 227:229-237], 및 미국 특허 번호 5,814,618 및 5,789,156 참조) (본원에 참조로 포함됨)가 또한 유용하다.
- [0103] 조직-특이적 또는 조직-선택 프로모터는 특정한 조직 내에서 발현 구축물의 발현을 표적화하는 데 이용될 수 있다. 특정한 실시양태에서, 조직-특이적 또는 조직-선택 프로모터는 식물 조직에서 활성이다. 식물에서 발달 제어 하에 있는 프로모터의 예는 특정 조직, 예컨대 잎, 뿌리, 과실, 종자 또는 꽃에서 전사를 우선적으로 개시하는 프로모터를 포함한다. "조직 특이적" 프로모터는 특정 조직에서만 전사를 개시하는 프로모터이다. 유전자의 구성적 발현과는 달리, 조직-특이적인 발현은 유전자 조절의 여러 상호작용 수준의 결과이다. 따라서, 상동성인 또는 밀접하게 관련된 식물 종으로부터의 프로모터가 특정한 조직에서 트랜스진의 효율적이고 신뢰할 수 있는 발현을 달성하기 위해 바람직하게 사용될 수 있다. 일부 실시양태에서, 발현은 조직-선택 프로모터를 포함한다. "조직 선택" 프로모터는 특정 조직에서 전사를 우선적으로 개시하지만 반드시 전체적으로 또는 단독으로 개시하는 것은 아닌 프로모터이다.
- [0104] 일부 실시양태에서, 본원에 기재된 USP를 코딩하는 핵산 분자는 세포 유형-특이적 프로모터를 포함한다. "세포 유형 특이적" 프로모터는 하나 이상의 기관의 특정 세포 유형에서 주로 발현을 유도하는 프로모터이다. 식물에서 기능성인 세포 유형 특이적 프로모터가 주로 활성일 수 있는 식물 세포의 일부 예는 예를 들어 BETL 세포, 뿌리, 잎에서의 맥관 세포, 자루 세포, 및 줄기 세포를 포함한다. 핵산 분자는 또한 세포 유형 선택 프로모터를 포함할 수 있다. "세포 유형 선택" 프로모터는 하나 이상의 기관의 특정 세포 유형에서 반드시 전적으로 또는 단독으로는 아니지만 주로 발현을 유도하는 프로모터이다. 식물에서 기능성인 세포 유형 선택 프로모터가 우선적으로 활성일 수 있는 식물 세포의 일부 예는 예를 들어 BETL 세포, 뿌리, 잎에서의 맥관 세포, 자루 세포, 및 줄기 세포를 포함한다.
- [0105] USP, 융합 단백질 및/또는 gRNA를 코딩하는 핵산 서열은 예를 들어 시험관내 mRNA 합성을 위해 파지 RNA 폴리머라제에 의해 인식되는 프로모터 서열에 작동가능하게 연결될 수 있다. 이러한 실시양태에서, 시험관내-전사된 RNA는 본원에 기재된 방법에 사용하기 위해 정제될 수 있다. 예를 들어, 프로모터 서열은 T7, T3, 또는 SP6 프로모터 서열 또는 T7, T3, 또는 SP6 프로모터 서열의 변이일 수 있다. 이러한 실시양태에서, 발현된 단백질 및/또는 RNA는 본원에 기재된 게놈 변형 방법에서 사용하기 위해 정제될 수 있다.
- [0106] 특정 실시양태에서, USP, 융합 단백질 및/또는 gRNA를 코딩하는 폴리뉴클레오타이드는 또한 폴리아데닐화 신호 (예를 들어, SV40 폴리A 신호 및 식물에서 기능성인 다른 신호) 및/또는 적어도 하나의 전사 종결 서열에 연결될 수 있다. 추가적으로, 데아미나제 또는 융합 단백질을 코딩하는 서열은 또한 본원의 다른 곳에 기재된 바와 같이, 단백질을 특정한 세포하 위치로 트래킹할 수 있는 적어도 하나의 핵 국제화 신호, 적어도 하나의 세포-침투 도메인 및/또는 적어도 하나의 신호 펩티드를 코딩하는 서열(들)에 연결될 수 있다.
- [0107] USP, 융합 단백질 및/또는 gRNA를 코딩하는 폴리뉴클레오타이드는 벡터 또는 다중 벡터에 존재할 수 있다. "벡터"는 핵산을 숙주 세포로 수송, 전달 또는 도입하기 위한 폴리뉴클레오타이드 조성물을 지칭한다. 적합한 벡터는 플라스미드 벡터, 파지미드, 코스미드, 인공/미니-염색체, 트랜스포손, 및 바이러스 벡터 (예를 들어, 렌티바이러스 벡터, 아데노-관련 바이러스 벡터, 바쿨로바이러스 벡터)를 포함한다. 벡터는 추가의 발현 제어 서열 (예를 들어, 인핸서 서열, 코작 서열, 폴리아데닐화 서열, 전사 종결 서열), 선택 마커 서열 (예를 들어, 항생제 저항성 유전자), 복제 기점 등을 포함할 수 있다. 추가의 정보는 문헌 ["Current Protocols in

Molecular Biology" Ausubel et al., John Wiley & Sons, New York, 2003 or "Molecular Cloning: A Laboratory Manual" Sambrook & Russell, Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 3rd edition, 2001]에서 확인할 수 있다.

[0108] 벡터는 또한 형질전환된 세포의 선택을 위해 선택 마커 유전자를 포함할 수 있다. 선택 마커 유전자는 형질전환된 세포 또는 조직의 선택을 위해 이용된다. 마커 유전자는 항생제 저항성을 코딩하는 유전자, 예컨대 네오마이신 포스포트랜스퍼라제 II (NEO) 및 히그로마이신 포스포트랜스퍼라제 (HPT)를 코딩하는 유전자, 뿐만 아니라 제초 화합물, 예컨대 글루포시네이트 암모늄, 브로모시닐, 이미다졸리논 및 2,4-디클로로페녹시아세테이트 (2,4-D)에 대한 저항성을 부여하는 유전자를 포함한다.

[0109] 일부 실시양태에서, RNA-가이드된 DNA-결합 폴리펩티드, 예컨대 RGN을 포함하는 융합 단백질을 코딩하는 서열을 포함하는 발현 카세트 또는 벡터는 gRNA를 코딩하는 서열을 추가로 포함할 수 있다. gRNA를 코딩하는 서열(들)은 관심 유기체 또는 숙주 세포에서의 gRNA의 발현을 위한 적어도 하나의 전사 제어 서열에 작동가능하게 연결될 수 있다. 예를 들어, gRNA를 코딩하는 폴리뉴클레오티드는 RNA 폴리머라제 III (Pol III)에 의해 인식되는 프로모터 서열에 작동가능하게 연결될 수 있다. 적합한 Pol III 프로모터의 예는 포유동물 U6, U3, H1, 및 7SL RNA 프로모터 및 벡터 U6 및 U3 프로모터를 포함하나 이에 제한되지는 않는다.

[0110] 나타낸 바와 같이, USP, 융합 단백질 및/또는 gRNA를 코딩하는 뉴클레오티드 서열을 포함하는 발현 구축물은 관심 유기체를 형질전환시키는 데 사용될 수 있다. 형질전환 방법은 뉴클레오티드 구축물을 관심 유기체에 도입하는 것을 수반한다. "도입시키는"은 구축물이 숙주 세포의 내부로 접근하게 하는 방식으로 뉴클레오티드 구축물을 숙주 세포에 도입시키는 것으로 의도된다. 본 발명의 방법은 뉴클레오티드 구축물을 숙주 유기체에 도입시키기 위해 특정한 방법을 필요로 하지 않으며, 단지 뉴클레오티드 구축물이 숙주 유기체의 적어도 하나의 세포의 내부로 접근하기만 하면 된다. 숙주 세포는 진핵 또는 원핵 세포일 수 있다. 특정한 실시양태에서, 진핵 숙주 세포는 식물 세포, 포유동물 세포 또는 곤충 세포이다. 안정한 형질전환 방법, 일시적인 형질전환 방법, 및 바이러스-매개된 방법을 비롯하여 이에 제한되지 않는, 뉴클레오티드 구축물을 식물 및 다른 숙주 세포에 도입시키는 방법은 관련 기술분야에 공지되어 있다.

[0111] 방법은 형질전환된 유기체, 예컨대 전체 식물, 뿐만 아니라 식물 기관(예를 들어, 잎, 줄기, 뿌리 등), 종자, 식물 세포, 번식체, 배아 및 그의 자손을 포함한 식물을 생성한다. 식물 세포는 분화 또는 미분화될 수 있다(예를 들어, 캘러스, 현탁 배양 세포, 원형질체, 잎 세포, 뿌리 세포, 체관부 세포, 화분).

[0112] "트랜스제닉 유기체" 또는 "형질전환된 유기체" 또는 "안정하게 형질전환된" 유기체 또는 세포 또는 조직은 본 발명의 데아미나제를 코딩하는 폴리뉴클레오티드가 혼입 또는 통합된 유기체를 지칭한다. 다른 외인성 또는 내인성 핵산 서열 또는 DNA 단편 또한 숙주 세포에 혼입될 수 있는 것으로 인지된다. 아그로박테리움- 및 바이올리스틱-매개 형질전환은 식물 세포의 형질전환을 위해 2가지 우세하게 이용되는 접근법으로 남아있다. 그러나, 숙주 세포의 형질전환은 감염, 형질감염, 미세주사, 전기천공, 미세투사, 바이올리스틱 또는 입자 충격, 전기천공, 실리카/탄소 섬유, 초음파 매개, PEG 매개, 인산칼슘 공동-침전, 다가양이온 DMSO 기술, DEAE 텍스트란 절차, 및 바이러스 매개, 리포솜 매개 등에 의해 수행될 수 있다. 데아미나제, 융합 단백질 및/또는 gRNA를 코딩하는 폴리뉴클레오티드의 바이러스-매개 도입은 레트로바이러스, 렌티바이러스, 아데노바이러스 및 아데노-관련 바이러스 매개 도입 및 발현, 뿐만 아니라 콜리모바이러스(예를 들어, 콜리플라워 모자이크 바이러스), 제미니 바이러스(예를 들어, 콩 골든 엘로우 모자이크 바이러스 또는 메이즈 줄무늬 바이러스) 및 RNA 식물 바이러스(예를 들어, 담배 모자이크 바이러스)의 사용을 포함한다.

[0113] 형질전환 프로토콜 뿐만 아니라 폴리펩티드 또는 폴리뉴클레오티드 서열을 식물에 도입하기 위한 프로토콜은 형질전환을 위해 표적화된 숙주 세포의 유형(예를 들어, 단자엽 또는 쌍자엽 식물 세포)에 따라 달라질 수 있다. 형질전환 방법은 관련 기술분야에 공지되어 있고, 미국 특허 번호: 8,575,425; 7,692,068; 8,802,934; 7,541,517에 제시된 것들이 포함되며, 이들 각각은 본원에 참조로 포함된다. 또한, 문헌 [Rakoczy-Trojanowska, M. (2002) Cell Mol Biol Lett. 7:849-858; Jones et al. (2005) Plant Methods 1:5; Rivera et al. (2012) Physics of Life Reviews 9:308-345; Bartlett et al. (2008) Plant Methods 4:1-12; Bates, G.W. (1999) Methods in Molecular Biology 111:359-366; Binns and Thomashow (1988) Annual Reviews in Microbiology 42:575-606; Christou, P. (1992) The Plant Journal 2:275-281; Christou, P. (1995) Euphytica 85:13-27; Tzfira et al. (2004) TRENDS in Genetics 20:375-383; Yao et al. (2006) Journal of Experimental Botany 57:3737-3746; Zupan and Zambryski (1995) Plant Physiology 107:1041-1047; Jones et al. (2005) Plant Methods 1:5]을 참고한다.

- [0114] 형질전환은 세포로 핵산의 안정적인 또는 일시적인 포함을 일으킬 수 있다. "안정한 형질전환"은 숙주 세포에 도입된 뉴클레오티드 구축물이 숙주 세포의 게놈에 통합되고, 그의 자손에 의해 유전될 수 있음을 의미하는 것으로 의도된다. "일시적인 형질전환"은, 폴리뉴클레오티드가 숙주 세포에 도입되고 숙주 세포의 게놈에 통합되지 않는 것을 의미하는 것으로 의도된다.
- [0115] 엽록체의 형질전환 방법은 관련 기술분야에 공지되어 있다. 예를 들어, 문헌 [Svab et al. (1990) Proc. Natl. Acad. Sci. USA 87:8526-8530; Svab and Maliga (1993) Proc. Natl. Acad. Sci. USA 90:913-917; Svab and Maliga (1993) EMBO J. 12:601-606]을 참조한다. 상기 방법은 선택 마커를 함유하는 DNA의 입자 총 전달, 및 상동 재조합을 통해 색소체 게놈으로 DNA의 표적화에 의존한다. 추가적으로, 색소체 형질전환은 핵-코딩된 및 색소체-지정된 RNA 폴리머라제의 조직-선택 발현에 의한 침묵 색소체-매개 트랜스진의 전사활성화에 의해 달성될 수 있다. 이러한 시스템은 문헌 [McBride et al. (1994) Proc. Natl. Acad. Sci. USA 91:7301-7305]에서 보고되었다.
- [0116] 형질전환된 세포는 통상적인 방식에 따라 트랜스제닉 유기체, 예컨대 식물로 성장할 수 있다. 예를 들어, 문헌 [McCormick et al. (1986) Plant Cell Reports 5:81-84]을 참조한다. 이어서, 이들 식물을 성장시킨 다음, 동일한 형질전환된 균주 또는 상이한 균주와 수분시킬 수 있고, 생성된 혼성체는 확인된 원하는 표현형 특징의 구성적 발현을 가질 수 있다. 원하는 표현형 특징의 발현이 안정하게 유지되고 유전된 다음, 종자가 수거되는 것을 보장하도록 2 세대 이상 성장시켜, 원하는 표현형 특징의 발현이 달성될 수 있도록 보장할 수 있다. 이러한 방식으로, 본 발명은 게놈에 안정하게 도입된 본 발명의 뉴클레오티드 구축물, 예를 들어 본 발명의 발현 카세트 갖는 형질전환된 종자 ("트랜스제닉 종자"로도 지칭됨)를 제공한다.
- [0117] 대안적으로, 형질전환된 세포는 유기체에 도입될 수 있다. 이들 세포는 유기체로부터 기원할 수 있고, 세포는 생체의 접근법으로 형질전환된다.
- [0118] 본원에 제공된 서열은 임의의 식물 중, 예컨대 비제한적으로 단자엽식물 및 쌍자엽식물의 형질전환을 위해 사용될 수 있다. 관심 식물의 예는 옥수수 (메이즈), 수수, 밀, 해바라기, 토마토, 십자화과, 페퍼, 감자, 목화, 벼, 대두, 사탕무, 사탕수수, 담배, 보리, 및 유지종자 평지, 브라시카 종, 알팔파, 호밀, 기장, 홍화, 땅콩, 고구마, 카사야, 커피, 코코넛, 파인애플, 시트러스 나무, 코코아, 차, 바나나, 아보카도, 무화과, 구아바, 망고, 올리브, 파파야, 캐슈, 마카다미아, 아몬드, 귀리, 채소, 관상식물, 및 침엽수를 포함하나 이에 제한되지는 않는다.
- [0119] 채소는 토마토, 상추, 녹두, 리마콩, 완두, 및 큐큐미스 속의 구성원, 예컨대 오이, 칸탈루프, 및 머스크 멜론을 포함하나 이에 제한되지는 않는다. 관상식물은 진달래, 수국, 히비스커스, 장미, 튜립, 수선화, 페튜니아, 카네이션, 포인세티아, 및 국화를 포함하나 이에 제한되지는 않는다. 바람직하게는, 본 발명의 식물은 작물 식물 (예를 들어, 메이즈, 수수, 밀, 해바라기, 토마토, 십자화과, 페퍼, 감자, 목화, 벼, 대두, 사탕무, 사탕수수, 담배, 보리, 유지종자 평지 등)이다.
- [0120] 본원에서 사용된 바와 같이, 용어 식물은 식물 세포, 식물 원형질체, 식물이 재생될 수 있는 식물 세포 조직 배양물, 식물 캘리(calli), 식물 덩어리, 및 식물 또는 식물의 일부분, 예컨대 배아, 화분, 배주, 종자, 잎, 꽃, 가지, 과실, 커널, 이삭, 콕, 겉껍질, 자루, 뿌리, 근단, 꽃밥 등에서 무손상인 식물 세포를 포함한다. 곡물은 종의 성장 또는 번식 이외의 목적으로 상업적 재배자에 의해 생산된 성숙 종자를 의미하는 것으로 의도된다. 재생 식물의 자손, 변이체 및 돌연변이체 또한 본 발명의 범위 내에 포함되나, 단 이들 부분은 도입된 폴리뉴클레오티드를 포함한다. 추가로, 예를 들어 대두박을 비롯한 본원에 개시된 서열을 보유하는 가공된 식물 생성물 또는 부산물이 제공된다.
- [0121] USP, 융합 단백질 및/또는 gRNA를 코딩하는 폴리뉴클레오티드는 동물 (예를 들어, 포유동물, 곤충, 어류, 조류 및 파충류), 진균, 아메바, 조류 및 효모를 포함하나 이에 제한되지는 않는 임의의 진핵 종을 형질전환시키는 데 사용될 수 있다. USP, 융합 단백질 및/또는 gRNA를 코딩하는 폴리뉴클레오티드는 또한 고세균 및 박테리아 (예를 들어, 바실루스(*Bacillus*) 종, 클레브시엘라(*Klebsiella*) 종, 스트렙토미세스(*Streptomyces*) 종, 리조비움(*Rhizobium*) 종, 에스케리키아(*Escherichia*) 종, 슈도모나스(*Pseudomonas*) 종, 살모넬라(*Salmonella*) 종, 시겔라(*Shigella*) 종, 비브리오(*Vibrio*) 종, 예르시니아(*Yersinia*) 종, 미코플라스마(*Mycoplasma*) 종, 아그로박테리움(*Agrobacterium*) 종 및 락토바실루스(*Lactobacillus*) 종)를 포함하나 이에 제한되지는 않는 임의의 원핵 종을 형질전환시키는 데 사용될 수 있다.
- [0122] 통상적인 바이러스 및 비-바이러스 기반 유전자 전달 방법을 사용하여 핵산을 포유동물 세포 또는 표적 조직에

도입할 수 있다. 이러한 방법은 본 발명의 융합 단백질 및 임의로 gRNA를 코딩하는 핵산을 배양물 또는 숙주 유기체 내의 세포에 투여하는 데 사용될 수 있다. 비-바이러스 벡터 전달 시스템은 DNA 플라스미드, RNA (예를 들어, 본원에 기재된 벡터의 전사체), 네이키드 핵산, 및 전달 비히클, 예컨대 리포솜과 복합체화된 핵산을 포함한다. 바이러스 벡터 전달 시스템은 세포로 전달 후에 에피솜 또는 통합 게놈을 갖는 DNA 및 RNA 바이러스를 포함한다. 비제한적 예는 콜리모나바이러스 (예를 들어, 콜리플라워 모자이크 바이러스), 제미니바이러스 (예를 들어, 콩 골든 엘로우 모자이크 바이러스 또는 메이즈 줄무늬 바이러스), 및 RNA 식물 바이러스 (예를 들어, 담배 모자이크 바이러스)를 이용하는 벡터를 포함한다. 유전자 요법 절차의 검토를 위해, [Anderson, Science 256: 808- 813 (1992); Nabel & Feigner, TIBTECH 11:211-217 (1993); Mitani & Caskey, TIBTECH 11:162-166 (1993); Dillon, TIBTECH 11:167-175 (1993); Miller, Nature 357:455-460 (1992); Van Brunt, Biotechnology 6(10): 1149-1154 (1988); Vigne, Restorative Neurology and Neuroscience 8:35-36 (1995); Kremer & Perricaudet, British Medical Bulletin 51(1):31-44 (1995); Haddada et al., in Current Topics in Microbiology and Immunology, Doerfler and Bohm (eds) (1995); and Yu et al., Gene Therapy 1:13-26 (1994)]를 참고한다.

[0123] 핵산의 비-바이러스 전달 방법은 리포펙션, 아그로박테리움-매개 형질전환, 뉴클레오펙션, 미세주사, 바이오리스틱, 비로솜, 리포솜, 이뮤노리포솜, 다가양이온 또는 지질: 핵산 접합체, 네이키드 DNA, 인공 비리온, 및 DNA의 작용제-증진된 흡수를 포함한다. 리포펙션은 예를 들어 미국 특허 번호 5,049,386, 4,946,787; 및 4,897,355에 기재되어 있고, 리포펙션 시약은 상업적으로 판매된다 (예를 들어, 트랜스펙탐(Transfectam)TM 및 리포펙틴(Lipofectin)TM). 폴리뉴클레오티드의 효율적인 수용체-인식 리포펙션에 적합한 양이온성 및 중성 지질은 프라이그너(Feigner), WO 91/17424; WO 91/16024의 것들을 포함한다. 전달은 세포 (예를 들어, 시험관내 또는 생체외 투여) 또는 표적 조직 (예를 들어, 생체내 투여)에 대한 것일 수 있다. 표적화된 리포솜, 예컨대 면역지질 복합체를 비롯한 지질:핵산 복합체의 제조는 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다 (예를 들어, 문헌 [Crystal, Science 270:404-410 (1995); Blaese et al., Cancer Gene Ther. 2:291- 297 (1995); Behr et al., Bioconjugate Chem. 5:382-389 (1994); Remy et al., Bioconjugate Chem. 5:647-654 (1994); Gao et al., Gene Therapy 2:710-722 (1995); Ahmad et al., Cancer Res. 52:4817-4820 (1992)]; 미국 특허 번호 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, 및 4,946,787 참조).

[0124] 핵산의 전달을 위한 RNA 또는 DNA 바이러스 기반 시스템의 사용은 바이러스를 신체 내의 특이적 세포에 표적화하고 바이러스 페이로드를 핵으로 트래픽킹하기 위한 고도로 진화된 과정을 이용한다. 바이러스 벡터는 환자에게 직접적으로 투여될 수 있거나 (생체내) 또는 이들을 이용하여 시험관내에서 세포를 치료할 수 있고, 임의로 변형된 세포가 환자에게 투여될 수 있다 (생체외). 통상적인 바이러스 기반 시스템은 유전자 전달을 위한 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-연관 및 단순 포진 바이러스 벡터를 포함할 수 있다. 숙주 게놈에서 통합은 레트로바이러스, 렌티바이러스, 및 아데노-연관된 바이러스 유전자 전달 방법에 의해 가능하며, 이는 종종 삽입된 트랜스진의 장기간 발현을 일으킨다. 추가적으로, 높은 형질도입 효율이 여러 상이한 세포 유형 및 표적 조직에서 관찰되었다.

[0125] 레트로바이러스의 형성은 외래 외피 단백질을 도입하고, 표적 세포의 잠재적인 표적 집단을 확장시킴으로써 변경될 수 있다. 렌티바이러스 벡터는 비-분열 세포를 형질도입하거나 또는 감염시킬 수 있고, 전형적으로 높은 바이러스 역가를 생성할 수 있는 레트로바이러스 벡터이다. 따라서, 레트로바이러스 유전자 전달 시스템의 선택은 표적 조직에 따라 좌우될 것이다. 레트로바이러스 벡터는 6-10 kb 이하의 외래 서열에 대해 패키징 능력을 갖는 시스-작용성 긴 말단 반복부로 구성된다. 최소 시스-작용성 LTR은 벡터의 복제 및 패키징에 충분하며, 이는 이어서 치료 유전자를 표적 세포에 통합시켜 영구적 트랜스진 발현을 제공하는 데 사용된다. 널리 사용되는 레트로바이러스 벡터는 뮌헨 백혈병 바이러스 (MuLV), 긴팔원숭이 유인원 백혈병 바이러스 (GaLV), 원숭이 면역 결핍 바이러스 (SIV), 인간 면역 결핍 바이러스 (HIV), 및 그의 조합을 기반으로 하는 것을 포함한다 (예를 들어, 문헌 [Buchscher et al., J. Virol. 66:2731-2739 (1992); Johann et al., J. Virol. 66:1635-1640 (1992); Sommerfelt et al., Virol. 176:58-59 (1990); Wilson et al., J. Virol. 63:2374-2378 (1989); Miller et al., J. Virol. 65:2220-2224 (1991)]; PCT/US94/05700 참조).

[0126] 일시적 발현이 바람직한 경우, 아데노바이러스 기반 시스템이 사용될 수 있다. 아데노바이러스 기반 벡터는 여러 세포 유형에서 매우 높은 형질도입 효율을 가질 수 있고, 세포 분열을 필요로 하지 않는다. 이러한 벡터를 사용하여, 높은 역가 및 발현 수준이 수득되었다. 이 벡터는 비교적 간단한 시스템으로 다량 생산될 수 있다. 아데노-관련 바이러스 ("AAV") 벡터는 또한, 예를 들어 핵산 및 캡시드의 시험관내 생성에서 및 생체내 및 생체

의 유전자 요법 절차를 위해 세포를 표적 핵산으로 형질도입하는 데 사용될 수 있다 (예를 들어, 문헌 [West et al., Virology 160:38-47 (1987)]; 미국 특허 번호 4,797,368; WO 93/24641; 문헌 [Katin, Human Gene Therapy 5:793-801 (1994); Muzyczka, J. Clin. Invest. 94:1351 (1994)] 참조). 재조합 AAV 벡터의 구축은 미국 특허 번호 5,173,414; 문헌 [Tratschin et al., Mol. Cell. Biol. 5:3251-3260 (1985); Tratschin, et al., Mol. Cell. Biol. 4:2072-2081 (1984); Hermonat & Muzyczka, PNAS 81:6466-6470 (1984); and Samulski et al., J. Virol. 63:03822-3828 (1989)]을 비롯한 다수의 간행물에 기재되어 있다. 패키징 세포는 전형적으로 숙주 세포를 감염시킬 수 있는 바이러스 입자를 형성하는 데 사용된다. 이러한 세포는 아데노바이러스를 패키징하는 293 세포, 및 레트로바이러스를 패키징하는 ψ J2 세포 또는 PA317 세포를 포함한다.

[0127] 유전자 요법에서 사용되는 바이러스 벡터는 통상적으로 핵산 벡터를 바이러스 입자로 패키징하는 세포주를 생성함으로써 생성된다. 벡터는 전형적으로 패키징 및 숙주로의 후속 통합에 요구되는 최소 바이러스 서열을 함유하고, 다른 바이러스 서열은 발현될 폴리뉴클레오티드(들)에 대한 발현 카세트에 의해 대체된다. 누락된 바이러스 기능은 전형적으로 패키징 세포주에 의해 트랜스로 공급된다. 예를 들어, 유전자 요법에서 사용되는 AAV 벡터는 전형적으로 패키징 및 숙주 게놈으로의 통합을 위해 필요한 AAV 게놈으로부터의 ITR 서열만을 갖는다. 바이러스 DNA는 다른 AAV 유전자, 즉 rep 및 cap를 코딩하는 헬퍼 플라스미드를 함유하지만 ITR 서열은 결여된 세포주에서 패키징된다.

[0128] 세포주는 또한 헬퍼로서 아데노바이러스로 감염될 수 있다. 헬퍼 바이러스는 헬퍼 플라스미드로부터 AAV 벡터의 복제 및 AAV 유전자의 발현을 촉진시킨다. 헬퍼 플라스미드는 ITR 서열의 결여로 인해 유의한 양으로 패키징되지 않는다. 아데노바이러스에 의한 오염은, 예를 들어 아데노바이러스가 AAV보다 더 감수성인 열 처리에 의해 감소될 수 있다. 세포로의 핵산의 추가의 전달 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있다. 예를 들어, 본원에 참조로 포함된 US20030087817을 참고한다.

[0129] 일부 실시양태에서, 숙주 세포는 본원에 기재된 하나 이상의 벡터로 일시적으로 또는 비-일시적으로 형질감염된다. 일부 실시양태에서, 세포는 대상체에서 자연적으로 발생하는 바와 같이 형질감염된다. 일부 실시양태에서, 형질감염된 세포를 대상체로부터 취한다. 일부 실시양태에서, 세포는 대상체로부터 취한 세포, 예컨대 세포주로부터 유래된다. 일부 실시양태에서, 세포 또는 세포주는 원핵이다. 다른 실시양태에서, 세포 또는 세포주는 진핵이다. 추가 실시양태에서, 세포 또는 세포주는 곤충, 조류, 식물 또는 진균 종으로부터 유래된다. 일부 실시양태에서, 세포 또는 세포주는 포유동물, 예컨대 예를 들어 인간, 원숭이, 마우스, 소, 돼지, 염소, 햄스터, 래트, 고양이 또는 개일 수 있다. 조직 배양을 위한 매우 다양한 세포주가 관련 기술분야에 공지되어 있다. 세포주의 예는 C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLaS3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEK293, HEK293T, MiaPaCell, Panel, PC-3, TF1, CTLL-2, CIR, Rat6, CV1, RPTE, A10, T24, 182, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, 145.01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, BS-C-1 원숭이 신장 상피, BALB/3T3 마우스 배아 섬유모세포, 3T3 Swiss, 3T3-L1, 132-d5 인간 태아 섬유모세포; 10.1 마우스 섬유모세포, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, BCP-1 세포, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr^{-/-}, COR-L23, COR-L23/CPR, COR-L235010, CORL23/R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepalclc7, HL-60, HMEC, HT-29, Jurkat, 1Y 세포, K562 세포, Ku812, KCL22, KG1, KYO1, LNCap, Ma-Me1 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCKII, MDCKII, MOR/0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, OPCN/OPCT 세포주, Peer, PNT-1A/PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 세포, Sf-9, SkBr3, T2, T-47D, T84, THP1 세포주, U373, U87, U937, VCaP, Vero 세포, WM39, WT-49, X63, YAC-1, YAR, 및 그의 트랜스제닉 품종을 포함하나 이에 제한되지는 않는다. 세포주는 관련 기술분야의 통상의 기술자에게 공지된 다양한 공급원으로부터 이용가능하다 (예를 들어, 아메리칸 타입 컬처 콜렉션 (ATCC) (버지니아주 마나사스) 참조).

[0130] 일부 실시양태에서, 본원에 기재된 하나 이상의 벡터로 형질감염된 세포를 사용하여, 하나 이상의 벡터-유래된 서열을 포함하는 새로운 세포주를 확립한다. 일부 실시양태에서, 본 발명의 융합 단백질 및 임의로 gRNA로, 또는 본 발명의 리보핵단백질 복합체로 일시적으로 형질감염되고 융합 단백질 또는 리보핵단백질 복합체의 활성을 통해 변형된 세포를 사용하여, 변형을 함유하지만 임의의 다른 외인성 서열이 결여된 세포를 포함하는 새로운 세포주를 확립한다. 일부 실시양태에서, 본원에 기재된 하나 이상의 벡터에 의해 일시적으로 또는 비-일시적으로

로 형질감염된 세포, 또는 이러한 세포로부터 유래된 세포주는 하나 이상의 시험 화합물을 평가하는데 사용된다.

[0131] 일부 실시양태에서, 본원에 기재된 하나 이상의 벡터를 사용하여 비-인간 트랜스제닉 동물 또는 트랜스제닉 식물을 생성한다. 일부 실시양태에서, 트랜스제닉 동물은 곤충이다. 추가 실시양태에서, 곤충은 곤충 해충, 예컨대 모기 또는 진드기이다. 다른 실시양태에서, 곤충은 식물 해충, 예컨대 옥수수 뿌리벌레 또는 가을 거엽벌레이다. 일부 실시양태에서, 트랜스제닉 동물은 조류, 예컨대 닭, 칠면조, 거위 또는 오리이다. 일부 실시양태에서, 트랜스제닉 동물은 포유동물, 예컨대 인간, 마우스, 래트, 햄스터, 원숭이, 유인원, 토끼, 돼지, 소, 말, 염소, 양, 고양이 또는 개이다.

[0132] V. 폴리펩티드 및 폴리뉴클레오티드의 변이체 및 단편

[0133] 본 개시내용은 그의 아미노산 서열이 서열식별번호: 1-16으로 제시된 자연-발생 (즉, 야생형) 우라실 안정화 폴리펩티드의 활성 변이체 및 단편, 및 그를 코딩하는 폴리뉴클레오티드를 제공한다.

[0134] 변이체 또는 단편의 활성이 관심 폴리뉴클레오티드 또는 폴리펩티드에 비해 변경될 수 있지만, 변이체 및 단편은 관심 폴리뉴클레오티드 또는 폴리펩티드의 기능을 보유해야 한다. 예를 들어, 변이체 또는 단편은 관심 폴리뉴클레오티드 또는 폴리펩티드와 비교하여 증가된 활성, 감소된 활성, 상이한 스펙트럼의 활성 또는 활성에서의 임의의 다른 변경을 가질 수 있다.

[0135] 자연-발생 USP의 단편 및 변이체, 예컨대 본원에 개시된 것들은 이들이 데아미나제 또는 그의 단편 및/또는 DNA-결합 폴리펩티드 또는 그의 단편을 추가로 포함하는 융합 단백질의 일부인 경우, 상기 융합 단백질이 USP 도메인을 포함하지 않는 유사한 융합 단백질과 비교하여 증가된 C→T 핵염기 편집 효율을 나타내도록 활성을 보유할 것이다.

[0136] 용어 "단편"은 본 발명의 폴리뉴클레오티드 또는 폴리펩티드 서열의 일부분을 지칭한다. "단편" 또는 "생물학적 활성 부분"은 생물학적 활성 (즉, 핵산 상의 데아미나제 활성)을 보유하기에 충분한 수의 인접 뉴클레오티드를 포함하는 폴리뉴클레오티드를 포함한다. "단편" 또는 "생물학적 활성 부분"은 생물학적 활성을 보유하기에 충분한 수의 인접 아미노산 잔기를 포함하는 폴리펩티드를 포함한다. USP의 단편은 대안적 하류 출발 부위의 사용으로 인해 전장 서열보다 더 짧은 것을 포함한다. USP의 생물학적 활성 부분은 예를 들어 서열식별번호: 1-16 중 임의의 것의 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110개 또는 그 초과인 인접 아미노산 잔기를 포함하는 폴리펩티드 또는 그의 변이체일 수 있다. 이러한 생물학적 활성 부분은 제조법 기술에 의해 제조되고 활성에 대해 평가될 수 있다.

[0137] 일반적으로, "변이체"는 실질적으로 유사한 서열을 의미하는 것으로 의도된다. 폴리뉴클레오티드의 경우, 변이체는 천연 폴리뉴클레오티드 내의 하나 이상의 내부 부위에서의 하나 이상의 뉴클레오티드의 결실 및/또는 부가 및/또는 천연 폴리뉴클레오티드 내의 하나 이상의 부위에서의 하나 이상의 뉴클레오티드의 치환을 포함한다. 본원에서 사용된 바와 같이, "본래의" 또는 "야생형" 폴리뉴클레오티드 또는 폴리펩티드는 각각 자연 발생 뉴클레오티드 서열 또는 아미노산 서열을 포함한다. 폴리뉴클레오티드의 경우, 보존적 변이체는 유전자 코드의 축퇴성 때문에 관심 유전자의 본래의 아미노산 서열을 코딩하는 이들 서열을 포함한다. 이들과 같은 자연 발생 대립유전자 변이체는 널리 공지된 분자 생물학 기술, 예를 들어 하기에서 설명되는 폴리머라제 연쇄 반응 (PCR) 및 혼성화 기술을 사용하여 확인될 수 있다. 변이체 폴리뉴클레오티드는 또한 합성에 의해 유래된 폴리뉴클레오티드, 예를 들어 부위-지정된 돌연변이유발을 이용하여 생성되었지만 관심 폴리펩티드 또는 폴리뉴클레오티드를 여전히 코딩하는 것들을 포함한다. 일반적으로, 본원에 개시된 특정한 폴리뉴클레오티드의 변이체는 본원에 기재된 서열 정렬 프로그램 및 파라미터에 의해 결정시 특정한 해당 폴리뉴클레오티드와 적어도 약 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과인 서열 동일성을 가질 것이다.

[0138] 본원에 개시된 특정한 폴리뉴클레오티드 (즉, 참조 폴리뉴클레오티드)의 변이체는 또한 변이체 폴리뉴클레오티드에 의해 코딩된 폴리펩티드와 참조 폴리뉴클레오티드에 의해 코딩된 폴리펩티드 사이의 퍼센트 서열 동일성의 비교에 의해 평가될 수 있다. 임의의 2개의 폴리펩티드 사이의 퍼센트 서열 동일성은 본원의 다른 곳에 기재된 서열 정렬 프로그램 및 파라미터를 사용하여 계산될 수 있다. 본원에 개시된 폴리뉴클레오티드의 임의의 주어진 쌍을 그들이 코딩하는 두 폴리펩티드에 의해 공유된 퍼센트 서열 동일성을 비교함으로써 평가하는 경우, 코딩된 두 폴리펩티드 사이의 퍼센트 서열 동일성은 적어도 약 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과인 서열 동일성이다.

- [0139] 특정한 실시양태에서, 본원에 개시된 폴리뉴클레오티드는 서열식별번호: 1-16 중 임의의 것의 아미노산 서열에 대해 적어도 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 그 초과와 동일성을 갖는 아미노산 서열을 포함하는 USP를 코딩한다.
- [0140] 본 발명의 우라실 안정화 폴리펩티드의 생물학적 활성 변이체는 적게는 약 1-15개 아미노산 잔기, 적게는 약 1-10개, 예컨대 약 6-10개, 적게는 5개, 적게는 4개, 적게는 3개, 적게는 2개 또는 적게는 1개 아미노산 잔기만큼 상이할 수 있다. 구체적 실시양태에서, 폴리펩티드는 폴리펩티드의 N 또는 C 말단으로부터 적어도 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60개 또는 그 초과의 아미노산의 결실을 포함할 수 있는 N-말단 또는 C-말단 절단을 포함할 수 있다.
- [0141] 변이체 단백질 및 폴리뉴클레오티드를 생성하는 변형이 본원에 제공된 USP에 이루어질 수 있는 것으로 인식된다. 인간에 의해 설계된 변화는 부위-지정 돌연변이유발 기술의 적용을 통해 도입될 수 있다. 대안적으로, 본원에 개시된 서열과 구조적으로 및/또는 기능적으로 관련이 있는 아직 공지되지 않거나 또는 아직 확인되지 않은 본래의 폴리뉴클레오티드 및/또는 폴리펩티드 또한 본 발명의 범위 내에 속하는 것으로 확인될 수 있다. 보존적 아미노산 치환은 우라실 안정화 폴리펩티드의 기능을 변경시키지 않는 비보존된 영역에서 이루어질 수 있다. 대안적으로, 우라실 안정화 폴리펩티드의 활성을 개선시키는 변형이 이루어질 수 있다.
- [0142] 변이체 폴리뉴클레오티드 및 단백질은 또한 돌연변이유발 및 재조합유발 절차, 예컨대 DNA 서플링으로부터 유래된 서열 및 단백질을 포함한다. 이러한 절차로, 본원에 개시된 하나 이상의 상이한 USP (예를 들어, 서열식별번호: 1-16)를 조작하여 목적하는 특성을 보유하는 새로운 USP를 생성한다. 이러한 방식으로, 재조합 폴리뉴클레오티드의 라이브러리는 실질적인 서열 동일성을 갖고 시험관내 또는 생체내에서 상동 재조합될 수 있는 서열 영역을 포함하는 관련 서열 폴리뉴클레오티드의 집단으로부터 생성된다. 예를 들어, 이러한 접근법을 사용하여, 관심 도메인을 코딩하는 서열 모티프는 본원에 제공된 USP 서열과 다른 후속적으로 확인된 USP 유전자 사이에 서플링되어 개선된 관심 특성, 예컨대 효소의 경우에 증가된 K_m 을 갖는 단백질을 코딩하는 새로운 유전자를 획득할 수 있다. 이러한 DNA 서플링을 위한 전략은 관련 기술분야에 공지되어 있다. 예를 들어, 문헌 [Stemmer (1994) Proc. Natl. Acad. Sci. USA 91:10747-10751; Stemmer (1994) Nature 370:389-391; Cramer et al. (1997) Nature Biotech. 15:436-438; Moore et al. (1997) J. Mol. Biol. 272:336-347; Zhang et al. (1997) Proc. Natl. Acad. Sci. USA 94:4504-4509; Cramer et al. (1998) Nature 391:288-291]; 및 미국 특허 번호 5,605,793 및 5,837,458을 참조한다. "서플링된" 핵산은 서플링 절차, 예컨대 본원에 제시된 임의의 서플링 절차에 의해 생성된 핵산이다. 서플링된 핵산은 예를 들어 인공적인, 임의적으로 재귀적인 방식으로 2종 이상의 핵산 (또는 문자열)을 (물리적으로 또는 가상으로) 재조합함으로써 생성된다. 일반적으로, 하나 이상의 스크리닝 단계를 서플링 과정에서 이용하여 관심 핵산을 확인하고; 이 스크리닝 단계는 임의의 재조합 단계 이전에 또는 이후에 수행될 수 있다. 일부 (그러나 전부는 아님) 서플링 실시양태에서, 스크리닝될 풀의 다양성을 증가시키기 위해 선택 이전에 다중 라운드의 재조합을 수행하는 것이 바람직하다. 재조합 및 선택의 전체 과정을 임의로 재귀적으로 반복한다. 문맥에 따라, 서플링은 재조합 및 선택의 전체 과정을 지칭할 수 있거나, 또는 다르게는 전체 과정의 재조합 부분을 간단히 지칭할 수 있다.
- [0143] 본원에 사용된 바와 같이, 2개의 폴리뉴클레오티드 또는 폴리펩티드 서열과 관련하여 "서열 동일성" 또는 "동일성"은 명시된 비교 윈도우에 걸쳐 최대 대응을 위해 정렬될 때 동일한 두 서열의 잔기를 지칭한다. 서열 동일성의 백분율이 단백질과 관련하여 사용되는 경우, 동일하지 않은 잔기 위치는 종종 보존적 아미노산 치환에 의해 상이하며, 여기서 아미노산 잔기는 유사한 화학적 특성 (예를 들어, 전하 또는 소수성)을 갖는 다른 아미노산 잔기로 치환되고, 따라서 분자의 기능적 특성을 변화시키지 않는 것으로 인식된다. 서열이 보존적 치환에서 상이할 때, 퍼센트 서열 동일성은 치환의 보존적 성질을 보정하기 위해 상향 조정될 수 있다. 이러한 보존적 치환에 의해 상이한 서열은 "서열 유사성" 또는 "유사성"을 갖는다고 한다. 이러한 조정을 위한 수단은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. 전형적으로, 이는 완전 미스매치보다는 부분 미스매치로서 보존적 치환을 점수화함으로써 서열 동일성 백분율을 증가시키는 것을 수반한다. 따라서, 예를 들어 동일한 아미노산에 1의 점수가 주어지고, 비-보존적 치환에 0의 점수가 주어지는 경우, 보존적 치환에는 0 내지 1의 점수가 주어진다. 보존적 치환의 점수화는 예를 들어 프로그램 PC/GENE (인텔리제네틱스(Intelligenetics), 캘리포니아주 마운틴 뷰)에서 구현되는 바와 같이 계산된다.
- [0144] 본원에 사용된 "서열 동일성의 백분율"은 비교 윈도우에 걸쳐 2개의 최적으로 정렬된 서열을 비교함으로써 결정된 값을 의미하며, 여기서 비교 윈도우 내의 폴리뉴클레오티드 서열의 부분은 2개의 서열의 최적 정렬을 위해

참조 서열 (부가 또는 결실을 포함하지 않음)과 비교하여 부가 또는 결실 (즉, 갭)을 포함할 수 있다. 백분율은 매칭된 위치의 개수를 획득하기 위해 두 서열에서 발생하는 동일한 핵산 염기 또는 아미노산 잔기의 위치의 개수를 결정하고, 매칭된 위치의 개수를 비교 윈도우의 총 위치의 개수로 나누고, 그 결과에 100을 곱하여 서열 동일성 백분율을 획득함으로써 계산된다.

[0145] 달리 언급되지 않는 한, 본원에 제공된 서열 동일성/유사성 값은 하기 파라미터를 사용하는 GAP 버전 10; 또는 그의 임의의 동등한 프로그램을 사용하여 획득된 값을 지칭한다: GAP 가중치 50 및 길이 가중치 3, 및 nws gapdna.cmp 점수화 매트릭스를 사용하는, 뉴클레오타이드 서열에 대한 % 동일성 및 % 유사성; GAP 가중치 8 및 길이 가중치 2, 및 BLOSUM62 점수화 매트릭스를 사용하는, 아미노산 서열에 대한 % 동일성 및 % 유사성. "동등한 프로그램"은 해당하는 임의의 2개의 서열에 대해 GAP 버전 10에 의해 생성된 상응하는 정렬과 비교 시 동일한 뉴클레오타이드 또는 아미노산 잔기 매치 및 동일한 퍼센트 서열 동일성을 갖는 정렬을 생성하는 임의의 서열 비교 프로그램으로 의도된다.

[0146] 2개의 서열은 이들이 서열의 쌍에 대해 가능한 최고 점수에 도달하도록 규정된 아미노산 치환 매트릭스 (예를 들어, BLOSUM62), 갭 존재 페널티 및 갭 연장 페널티를 사용하여 유사성 점수화를 위해 정렬될 때 "최적으로 정렬된다". 아미노산 치환 매트릭스 및 2개의 서열 사이의 유사성을 정량화하는 데 있어서의 그의 용도는 관련 기술분야에 널리 공지되어 있고, 예를 들어 문헌 [Dayhoff et al. (1978) "A model of evolutionary change in proteins." In "Atlas of Protein Sequence and Structure," Vol. 5, Suppl. 3 (ed. M. O. Dayhoff), pp. 345-352. Natl. Biomed. Res. Found., Washington, D.C. and Henikoff et al. (1992) Proc. Natl. Acad. Sci. USA 89:10915-10919]에 기재되어 있다. BLOSUM62 매트릭스는 종종 서열 정렬 프로토콜에서 디폴트 점수화 치환 매트릭스로서 사용된다. 갭 존재 페널티는 정렬된 서열 중 하나에 단일 아미노산 갭의 도입을 위해 부과되고, 갭 연장 페널티는 이미 개방된 갭에 삽입된 각각의 추가의 빈 아미노산 위치에 대해 부과된다. 정렬은 정렬이 시작되고 끝나는 각각의 서열의 아미노산 위치에 의해, 임의로 가능한 가장 높은 점수에 도달하도록 하나의 서열 또는 두 서열 모두에서 하나의 갭 또는 다중 갭의 삽입에 의해 정의된다. 최적의 정렬 및 점수화가 수동으로 달성될 수 있지만, 프로세스는 컴퓨터-구현 정렬 알고리즘, 예를 들어 문헌 [Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402]에 기재되고 미국 국립 생물 정보 센터 웹사이트 (www.ncbi.nlm.nih.gov)에서 공중이 이용가능하게 된 갭드 BLAST 2.0의 사용에 의해 용이해진다. 다중 정렬을 포함한 최적 정렬은, 예를 들어 www.ncbi.nlm.nih.gov를 통해 이용가능하고 문헌 [Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402]에 기재된 PSI-BLAST를 사용하여 제조될 수 있다.

[0147] 참조 서열에 대해 최적으로 정렬된 아미노산 서열과 관련하여, 아미노산 잔기는 정렬에서 잔기가 쌍을 형성한 참조 서열의 위치에 "상응한다". "위치"는 N-말단에 대한 그의 위치를 기반으로 하여 참조 서열에서 각각의 아미노산을 순차적으로 식별하는 번호로 표시된다. 최적의 정렬을 결정할 때 고려되어야 하는 결실, 삽입, 말단 절단, 융합 등으로 인해, 일반적으로 N-말단으로부터 간단히 카운팅함으로써 결정되는 시험 서열에서의 아미노산 잔기 개수는 참조 서열에서 그의 상응하는 위치의 번호와 반드시 동일하지는 않을 것이다. 예를 들어, 정렬된 시험 서열에서 결실이 있는 경우, 결실 부위에서는 참조 서열에서의 위치에 상응하는 아미노산이 없을 것이다. 정렬된 참조 서열에서 삽입이 있는 경우, 해당 삽입은 참조 서열에서 임의의 아미노산 위치에 상응하지 않을 것이다. 말단절단 또는 융합의 경우, 상응하는 서열 내의 임의의 아미노산에 상응하지 않는 참조 또는 정렬된 서열 내의 아미노산의 스트레치가 있을 수 있다.

[0148] VI. 항체

[0149] 서열식별번호: 1-16으로 제시된 아미노산 서열을 갖는 것을 비롯하여 본 발명의 USP, USP를 포함하는 융합 단백질, 또는 리보핵단백질에 대한 항체 또는 그의 활성 변이체 또는 단편이 또한 포함된다. 항체를 생성하는 방법은 관련 기술분야에 널리 공지되어 있다 (예를 들어, 문헌 [Harlow and Lane (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.]; 및 미국 특허 번호 4,196,265 참조). 이들 항체는 본원에 기재된 USP, 또는 USP를 포함하는 융합 단백질 또는 리보핵단백질의 검출 및 단리를 위한 키트에 사용될 수 있다. 따라서, 본 개시내용은 예를 들어 서열식별번호: 1-16 중 임의의 것에 대해 적어도 85% 동일성의 서열을 포함하는 폴리펩티드를 비롯하여 본원에 기재된 폴리펩티드 또는 리보핵단백질에 특이적으로 결합하는 항체를 포함하는 키트를 제공한다.

[0150] VII. 관심 표적 서열에 결합하기 위한 시스템 및 리보핵단백질 복합체 및 그의 제조 방법

[0151] 본 개시내용은 핵산 서열을 표적화하고 표적 핵산 서열을 변형시키는 시스템을 제공한다. 일부 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드, 예컨대 RGN, 및 gRNA는 관심 핵산 서열에 대한 리보핵단백질 복합체의 표

적화를 담당하고; 데아미나제 폴리펩티드는 C>U로부터 표적화된 핵산 서열의 변형을 담당하고; 우라실 안정화 폴리펩티드는 우라실이 DNA 분자에서 지속되도록 하여 목적하는 DNA 복구가 일어나도록 함으로써 C>T 돌연변이를 도입한다. 가이드 RNA는 관심 표적 서열에 혼성화하고, 또한 RNA-가이드된, DNA-결합 폴리펩티드와 복합체를 형성함으로써, RNA-가이드된, DNA-결합 폴리펩티드가 표적 서열에 결합하도록 지시한다. RNA-가이드된, DNA-결합 폴리펩티드는 3-도메인 융합 단백질의 한 도메인이고; 제2 도메인은 데아미나제이고, 제3 도메인은 본원에 기재된 USP이다. 일부 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드는 RGN, 예컨대 Cas9이다. RNA-가이드된, DNA-결합 폴리펩티드의 다른 예는 RGN, 예컨대 미국 특허 출원 공개 번호 2019/0367949 (그 전문이 본원에 참조로 포함됨)에 기재된 것들을 포함한다. 일부 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드는 유형 II CRISPR-Cas 폴리펩티드, 또는 그의 활성 변이체 또는 단편이다. 일부 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드는 유형 V CRISPR-Cas 폴리펩티드, 또는 그의 활성 변이체 또는 단편이다. 다른 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드는 유형 VI CRISPR-Cas 폴리펩티드이다. 다른 실시양태에서, 융합 단백질의 DNA-결합 도메인은 RNA 가이드, 예컨대 아연 핑거 뉴클레아제, TALEN 또는 메가뉴클레아제 폴리펩티드를 필요로 하지 않는다. 이들 실시양태 중 일부에서, 각각의 뉴클레아제 활성은 불활성화되었다. 추가 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드는 RGN의 아미노산 서열, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 활성 변이체 또는 단편, 예컨대 닉카제 APG07433.1 (서열식별번호: 41)에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함한다. 추가 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드는 RGN의 아미노산 서열, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 활성 변이체 또는 단편, 예컨대 닉카제 APG07433.1 (서열식별번호: 41)에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함한다. 추가 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드는 RGN의 아미노산 서열, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 활성 변이체 또는 단편, 예컨대 닉카제 APG07433.1 (서열식별번호: 41)에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함한다. 추가 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드는 RGN의 아미노산 서열, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 활성 변이체 또는 단편, 예컨대 닉카제 APG07433.1 (서열식별번호: 41)에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함한다. 추가 실시양태에서, RNA-가이드된, DNA-결합 폴리펩티드는 RGN의 아미노산 서열, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 활성 변이체 또는 단편, 예컨대 닉카제 APG07433.1 (서열식별번호: 41)을 포함한다.

[0152] 본원에 제공된 관심 표적 서열에 결합하기 위한 시스템은 적어도 하나의 단백질에 결합된 RNA의 적어도 하나의 분자인 리보핵단백질 복합체일 수 있다. 본원에 제공된 리보핵단백질 복합체는 RNA 성분으로서 적어도 하나의 가이드 RNA, 및 단백질 성분으로서 데아미나제, 본 발명의 USP, 및 RNA-가이드된, DNA-결합 폴리펩티드를 포함하는 융합 단백질을 포함한다. 리보핵단백질 복합체는, 융합 단백질 및 가이드 RNA를 코딩하는 폴리뉴클레오티드로 형질전환되고 융합 단백질 및 가이드 RNA의 발현을 허용하는 조건 하에 배양된 세포 또는 유기체로부터 정제될 수 있다. 따라서, USP, 융합 단백질, 또는 융합 단백질 리보핵단백질 복합체의 제조 방법이 제공된다. 이러한 방법은, USP 또는 융합 단백질 (및 일부 실시양태에서, 가이드 RNA)이 발현되는 조건 하에, USP를 코딩하는 뉴클레오티드 서열, 융합 단백질, 및 일부 실시양태에서 가이드 RNA를 코딩하는 뉴클레오티드 서열을 포함하는 세포를 배양하는 것을 포함한다. 이어서, USP, 융합 단백질 또는 융합 리보핵단백질은 배양된 세포의 용해물로부터 정제될 수 있다.

[0153] 생물학적 샘플의 용해물로부터 USP, 융합 단백질, 또는 융합 리보핵단백질 복합체를 정제하는 방법은 관련 기술 분야에 공지되어 있다 (예를 들어, 크기 배제 및/또는 친화성 크로마토그래피, 2D-PAGE, HPLC, 역상 크로마토그래피, 면역침전). 특정한 방법에서, USP 또는 융합 단백질은 재조합적으로 생성되고, 그의 정제를 보조하는 정제 태그, 예컨대 비제한적으로 글루타티온-S-트랜스퍼라제 (GST), 키틴 결합 단백질 (CBP), 말토스 결합 단백질, 티오레독신 (TRX), 폴리(NANP), 탠덤 친화도 정제 (TAP) 태그, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, 소프태그 1, 소프태그 3, 스트랩, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, 비오틴 카르복실 담체 단백질 (BCCP), 및 칼모듈린을 포함한다. 일반적으로, 태그부착된 USP, 융합 단백질, 또는 융합 리보핵단백질 복합체는 면역침전 또는 관련 기술분야에 공지된 다른 유사한 방법을 사용하여 정제된다.

[0154] "단리된" 또는 "정제된" 폴리펩티드, 또는 그의 생물학적 활성 부분은 그의 자연 발생 환경에서 발견되는 폴리펩티드에 일반적으로 동반되거나 또는 그와 상호작용하는 성분을 실질적으로 또는 본질적으로 함유하지 않는다. 따라서, 단리된 또는 정제된 폴리펩티드는 다른 세포 물질, 또는 재조합 기술에 의해 생성되는 경우 배양 배지를 실질적으로 함유하지 않거나, 또는 화학적으로 합성되는 경우 화학적 전구체 또는 다른 화학물질을 실질적으로 함유하지 않는다. 세포 물질을 실질적으로 함유하지 않는 단백질은 약 30%, 20%, 10%, 5% 또는 1% (건조 중량 기준) 미만의 오염 단백질을 갖는 단백질 제제를 포함한다. 본 발명의 단백질 또는 그의 생물학적 활성 부분이 재조합적으로 생성될 때, 최적으로 배양 배지는 약 30%, 20%, 10%, 5% 또는 1% (건조 중량 기준) 미만의

화학적 전구체 또는 관심 단백질이 아닌 화학물질을 나타낸다.

[0155] VIII. 표적 서열의 변형 방법

[0156] 본 개시내용은 관심 표적 핵산 분자 (예를 들어, 표적 DNA 분자)를 변형시키는 방법을 제공한다. 방법은 적어도 하나의 가이드 RNA 또는 그를 코딩하는 폴리뉴클레오티드, 및 본 발명의 USP, 데아미나제 및 RNA-가이드된, DNA-결합 폴리펩티드를 포함하는 적어도 하나의 융합 단백질 또는 그를 코딩하는 폴리뉴클레오티드를 포함하는 시스템을 표적 서열, 또는 표적 서열을 포함하는 세포, 소기관 또는 배아에 전달하는 것을 포함한다. 이들 실시양태 중 일부에서, 융합 단백질은 서열식별번호: 1-16의 아미노산 서열 중 어느 하나, 또는 그의 활성 변이체 또는 단편을 포함한다.

[0157] 일부 실시양태에서, 방법은 DNA 분자를 (a) USP, 데아미나제, 및 RNA-가이드된, DNA-결합 폴리펩티드, 예컨대 예를 들어 뉴클레아제-불활성 또는 니카제 Cas9 도메인을 포함하는 융합 단백질; 및 (b) (a)의 융합 단백질을 DNA 분자의 표적 뉴클레오티드 서열에 표적화하는 gRNA와 접촉시키는 것을 포함하며; 여기서 DNA 분자는 뉴클레오티드 염기의 탈아미노화에 적합한 조건 하에 및 유효량으로 융합 단백질 및 gRNA와 접촉된다. 일부 실시양태에서, 표적 DNA 분자는 질환 또는 장애와 연관된 서열을 포함하고, 여기서 뉴클레오티드 염기의 탈아미노화는 질환 또는 장애와 연관되지 않은 서열을 생성한다. 일부 실시양태에서, 질환 또는 장애는 동물에게 영향을 미친다. 추가 실시양태에서, 질환 또는 장애는 포유동물, 예컨대 인간, 소, 말, 개, 고양이, 염소, 양, 돼지, 원숭이, 래트, 마우스 또는 햄스터에 영향을 미친다. 일부 실시양태에서, 표적 DNA 서열은 작물 식물의 대립유전자에 존재하고, 여기서 관심 특성을 갖는 특정한 대립유전자는 더 적은 농경학적 가치를 갖는 식물을 생성한다. 뉴클레오티드 염기의 탈아미노화는 형질을 개선시키고 식물의 농경학적 가치를 증가시키는 대립유전자를 생성한다.

[0158] 일부 실시양태에서, 목적하는 돌연변이는 질환 또는 장애와 연관된 T→C 점 돌연변이를 포함하며, 여기서 돌연변이 C 염기의 탈아미노화는 질환 또는 장애와 연관되지 않은 서열을 생성한다. 일부 실시양태에서, 탈아미노화는 질환 또는 장애와 연관된 서열에서의 점 돌연변이를 교정한다.

[0159] 일부 실시양태에서, 질환 또는 장애와 연관된 서열은 단백질을 코딩하고, 여기서 탈아미노화는 질환 또는 장애와 연관된 서열에 정지 코돈을 도입시켜, 코딩된 단백질의 말단 절단을 일으킨다. 일부 실시양태에서, 접촉은 질환 또는 장애에 걸리기 쉽거나, 질환 또는 장애를 갖거나 또는 질환 또는 장애로 진단된 대상체에서 생체내에서 수행된다. 일부 실시양태에서, 질환 또는 장애는 개놈에서의 점 돌연변이 또는 단일-염기 돌연변이와 연관된 질환이다. 일부 실시양태에서, 질환은 유전 질환, 암, 대사 질환 또는 리소솜 축적 질환이다.

[0160] 따라서, 본원에 개시된 조성물 및 방법을 돌연변이되는 서열과 연관된 질환 또는 장애 (즉, 서열은 질환 또는 장애에 대한 원인이거나 또는 질환 또는 장애와 연관된 증상에 대한 원인임)의 치료에 사용하여 질환 또는 장애를 치료하거나 또는 질환 또는 장애와 연관된 증상을 감소시킬 수 있다. 본원에 사용된 용어 "치료하다" 또는 "치료"는 질환 또는 장애를 갖는 대상체에게 USP 또는 융합 단백질을 포함하는 본원에 개시된 제약 조성물을 투여하는 것을 지칭한다. 치료는 질환 또는 장애에 걸리기 쉬운 (예를 들어, 유전적으로 소인이 있는) 대상체에서 질환 또는 장애와 연관된 증상의 발병을 방지함으로써 예방적일 수 있다. 바람직한 치료 효과는 질환의 발생 또는 재발의 예방, 증상의 완화, 질환의 임의의 직접적 또는 간접적 병리학적 결과의 감소, 질환 진행 속도의 감소, 질환 상태의 호전 또는 완화, 및 완화 또는 개선된 예후를 포함하나 이에 제한되지는 않는다.

[0161] 제약 조성물은 활성 성분 (즉, USP 또는 융합 단백질 또는 그를 코딩하는 핵산 분자) 및 제약상 허용되는 담체를 포함하는 질환 또는 장애를 예방, 강도 감소, 치유 또는 달리 치료하는 데 사용되는 조성물이다. 제약상 허용되는 담체는 인간 또는 다른 척추동물에게 투여하기에 적합한 1종 이상의 상용성 고체 또는 액체 충전제, 희석제 또는 캡슐화 물질을 지칭한다. 일부 실시양태에서, 제약 조성물은 비-자연 발생인 제약상 허용되는 담체를 포함한다.

[0162] 본원에 개시된 방법에 사용되는 제약 조성물은 적합한 전달, 내성 등을 제공하는 적합한 담체, 부형제 및 다른 작용제와 함께 제제화될 수 있다. 다수의 적절한 제제가 관련 기술분야의 통상의 기술자에게 공지되어 있다. 비제한적 예는 멸균 희석제, 예컨대 주사용수, 염수 용액, 고정 오일, 폴리에틸렌 글리콜, 글리세린, 프로필렌 글리콜 또는 다른 합성 용매; 항박테리아제, 예컨대 벤질 알콜 또는 메틸 파라벤; 항산화제, 예컨대 아스코르브산 또는 중아황산나트륨; 킬레이트화제, 예컨대 에틸렌디아민테트라아세트산; 완충제, 예컨대 아세트이트, 시트레이트 또는 포스페이트 및 장성 조정제, 예컨대 염화나트륨 또는 텍스트로스를 포함한다. 정맥내로 투여되는 특정한 담체는 생리 염수 또는 포스페이트 완충 염수 (PBS)이다. 경구 또는 비경구 사용을 위한 제약 조성물은

활성 성분의 용량에 맞게 적합화된 단위 용량의 투여 형태로 제조될 수 있다. 단위 용량의 이러한 투여 형태는, 예를 들어 정제, 환제, 캡슐, 주사 (앰플), 좌제 등을 포함한다. 이들 조성물은 또한 보존제, 습윤제, 유화제 및 분산제를 비롯한 아주반트를 함유할 수 있다. 미생물 작용의 방지는 다양한 항박테리아제 및 항진균제, 예를 들어 파라벤, 클로로부탄올, 페놀, 소르브산 등에 의해 보장될 수 있다. 등장화제, 예를 들어 당, 염화나트륨 등을 포함하는 것이 또한 바람직할 수 있다. 주사가능한 제약 형태의 지속 흡수는 흡수를 지연시키는 작용제, 예를 들어 알루미늄 모노스테아레이트 및 젤라틴의 사용에 의해 이루어질 수 있다.

[0163] USP 또는 융합 단백질 또는 그를 코딩하는 핵산 분자 또는 그를 포함하는 세포를 포함하는 제약 조성물은 임의의 경로를 통해, 예컨대 경구로, 협측으로, 비경구로, 국소로, 흡입 또는 취입에 의해 (즉, 입을 통해 또는 코를 통해), 또는 직장으로 대상체에게 투여될 수 있다. 투여는, 예를 들어 1회, 복수회, 및/또는 1회 이상의 연장된 기간에 걸쳐 수행될 수 있다.

[0164] 본 발명의 제약 조성물의 유효량은 그의 목적 (예를 들어, 부분적 회복을 비롯한 예방 또는 회복, 또는 특정 순서에 의해 유발되는 장애 또는 질환의 예방 또는 저속화)을 달성하는 데 효과적인 임의의 양이다. 통상적으로 mg/kg으로 표현되는 유효량은 관련 기술분야의 통상의 기술자에 의해 전임상 및 임상 시험 동안 상용 방법에 의해 결정될 수 있다.

[0165] 방법이 가이드 RNA 및/또는 융합 단백질을 코딩하는 폴리뉴클레오티드를 전달하는 것을 포함하는 실시양태에서, 세포 또는 배아는 이어서 가이드 RNA 및/또는 융합 단백질이 발현되는 조건 하에 배양될 수 있다. 다양한 실시양태에서, 방법은 표적 서열을 gRNA 및 융합 단백질 (이는 본 발명의 USP, 데아미나제 및 RNA-가이드된 DNA-결합 폴리펩티드를 포함함)을 포함하는 리보핵단백질 복합체와 접촉시키는 것을 포함한다. 특정 실시양태에서, 방법은 표적 서열을 포함하는 세포, 소기관, 또는 배아에 본 발명의 리보핵단백질 복합체를 도입하는 것을 포함한다. 본 발명의 리보핵단백질 복합체는 본원에 기재된 바와 같이 생물학적 샘플로부터 정제되거나, 재조합적으로 생성되고 후속적으로 정제되거나, 또는 시험관내-조립된 것일 수 있다. 표적 서열 또는 세포 소기관 또는 배아와 접촉된 리보핵단백질 복합체가 시험관내에서 조립된 실시양태에서, 방법은 표적 서열, 세포, 소기관 또는 배아와의 접촉 전에 복합체의 시험관내 조립을 추가로 포함할 수 있다.

[0166] 본 발명의 정제된 또는 시험관내 조립된 리보핵단백질 복합체는 전기천공을 포함하고 이로 제한되지 않는 관련 기술분야에 공지된 임의의 방법을 사용하여 세포, 소기관, 또는 배아에 도입될 수 있다. 대안적으로, 본 발명의 USP, 데아미나제, 및 RNA-가이드된, DNA-결합 폴리펩티드를 포함하는 융합 단백질, 및 가이드 RNA를 코딩하거나 포함하는 폴리뉴클레오티드는 관련 기술분야에 공지된 임의의 방법 (예를 들어, 전기천공)을 사용하여 세포, 소기관 또는 배아에 도입될 수 있다.

[0167] 표적 서열, 또는 표적 서열을 포함하는 세포, 소기관 또는 배아로의 전달 또는 그와의 접촉 시, 가이드 RNA는 융합 단백질이 서열-특이적 방식으로 표적 서열에 결합하도록 지시한다. 표적 서열은 후속적으로 융합 단백질의 데아미나제 도메인 및 USP 도메인을 통해 변형될 수 있다. 일부 실시양태에서, 표적 서열에 대한 이러한 융합 단백질의 결합은 표적 서열에 인접한 뉴클레오티드의 변형을 유발한다. 표적 서열에 인접한 뉴클레오티드 염기는 표적 서열의 5' 또는 3' 말단으로부터 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 또는 100개 염기 쌍일 수 있다. 본 발명의 USP, 데아미나제 및 RNA-가이드된, DNA-결합 폴리펩티드를 포함하는 융합 단백질은 데아미나제 및 RNA-가이드된, DNA-결합 폴리펩티드를 단독으로 포함하는 융합 단백질과 비교하여 더 큰 효율로 표적화된 C>T 돌연변이를 도입할 수 있다.

[0168] 표적 서열에 대한 융합 단백질의 결합을 측정하는 방법은 관련 기술분야에 공지되어 있고, 염색질 면역침전 검정, 겔 이동성 변화 검정, DNA 풀-다운 검정, 리포터 검정, 마이크로플레이트 포획 및 검출 검정을 포함한다. 마찬가지로, 표적 서열의 절단 또는 변형을 측정하는 방법은 관련 기술분야에 공지되어 있고, 분해 산물의 검출을 용이하게 하기 위해 표적 서열에 대한 적절한 표지 (예를 들어, 방사성동위원소, 형광 물질)의 부착의 존재 또는 부재 하에, PCR, 서열분석 또는 겔 전기영동을 사용하여 절단을 확인하는 시험관내 또는 생체내 절단 검정을 포함한다. 대안적으로, Nicking 촉발 지수 증폭 반응 (NTEPAR) 검정을 사용할 수 있다 (예를 들어, 문헌 [Zhang et al. (2016) Chem. Sci. 7:4951-4957] 참조). 생체내 절단은 서베이어(Surveyor) 검정을 이용하여 평가될 수 있다 (문헌 [Guschin et al. (2010) Methods Mol Biol 649:247-256]).

[0169] 일부 실시양태에서, 방법은 1개 초과 가이드 RNA와 복합체화된, 융합 단백질의 일부로서의 RNA-결합, DNA-가이드된 도메인의 사용을 수반한다. 1개 초과 가이드 RNA는 단일 유전자의 상이한 영역을 표적화할 수 있거나 또는 다중 유전자를 표적화할 수 있다. 이러한 다중 표적화는 융합 단백질의 데아미나제 도메인이 핵산을 변형

시켜 관심 표적 핵산 분자 (예를 들어, 게놈)에 다중 돌연변이를 도입할 수 있게 한다. 융합 단백질의 USP 도메인은 목적하는 돌연변이의 도입의 효능을 증가시킨다.

[0170] 방법이 RNA-가이드된 뉴클레아제 (RGN), 예컨대 니카제 RGN (즉, 이중-가닥 폴리뉴클레오티드의 단일 가닥만을 절단할 수 있음, 예를 들어 nAPG07433.1 (서열식별번호: 41))의 사용을 수반하는 실시양태에서, 방법은 동일하거나 중첩되는 표적 서열을 표적화하고 폴리뉴클레오티드의 상이한 가닥을 절단하는 2개의 상이한 RGN 또는 RGN 변이체를 도입하는 것을 포함할 수 있다. 예를 들어, 이중-가닥 폴리뉴클레오티드의 양성 (+) 가닥만을 절단하는 RGN 니카제는 이중-가닥 폴리뉴클레오티드의 음성 (-) 가닥만을 절단하는 제2 RGN 니카제와 함께 도입될 수 있다. 대안적으로, 2종의 상이한 융합 단백질이 제공될 수 있으며, 여기서 각각의 융합 단백질은 상이한 PAM 인식 서열을 갖는 상이한 RGN을 포함하여, 보다 큰 다양성의 뉴클레오티드 서열이 돌연변이를 위해 표적화될 수 있다.

[0171] 관련 기술분야의 통상의 기술자는 본원에 개시된 임의의 방법을 이용하여 단일 표적 서열 또는 다중 표적 서열을 표적화할 수 있음을 이해할 것이다. 따라서, 방법은 단일 유전자 및/또는 다중 유전자 내의 다수의 별개의 서열을 표적화할 수 있는 다수의 별개의 가이드 RNA와 조합된 단일 RNA-가이드된, DNA-결합 폴리펩티드를 포함하는 융합 단백질의 사용을 포함한다. 이어서, 융합 단백질의 데아미나제 도메인은 각각의 표적화된 서열에 돌연변이를 도입할 것이다. 융합 단백질의 USP 도메인은 목적하는 돌연변이의 도입의 효능을 증가시킨다. 또한, 다수의 별개의 가이드 RNA가 다수의 별개의 RNA-가이드된, DNA 결합 폴리펩티드와 조합하여 도입되는 방법이 본원에 포괄된다. 이러한 RNA-가이드된, DNA-결합 폴리펩티드는 다중 RGN 또는 RGN 변이체일 수 있다. 이들 가이드 RNA 및 가이드 RNA/융합 단백질 시스템은 단일 유전자 및/또는 다중 유전자 내의 다수의 별개의 서열을 표적화할 수 있다.

[0172] IX. 폴리뉴클레오티드 유전자 변형을 포함하는 세포

[0173] 일부 실시양태는 본원에 제공된 융합 단백질을 사용하는 방법을 제공한다. 일부 실시양태에서, 융합 단백질은 표적 핵염기, 예를 들어 C 잔기를 탈아미노화함으로써 점 돌연변이를 표적 핵산 분자에 도입하는 데 사용된다. 일부 실시양태에서, 표적 핵염기의 탈아미노화는 유전적 결함의 교정, 예를 들어 유전자 산물에서의 기능 상실로 이어지는 점 돌연변이의 교정을 발생시킨다. 일부 실시양태에서, 유전적 결함은 질환 또는 장애, 예를 들어 리소솜 축적 장애 또는 대사 질환, 예컨대 예를 들어 제I형 당뇨병과 연관된다. 일부 실시양태에서, 본원에 제공된 방법은 탈활성화 점 돌연변이를 질환 또는 장애와 연관된 유전자 산물을 코딩하는 유전자 또는 대립유전자에 도입하는 데 사용된다. 예를 들어, 일부 실시양태에서, (예를 들어, 증식성 질환의 치료에서) 탈활성화 점 돌연변이를 종양유전자에 도입하기 위해 융합 단백질을 사용하는 방법이 본원에 제공된다. 탈활성화 돌연변이는, 일부 실시양태에서, 코딩 서열에서 조기 정지 코돈을 생성할 수 있으며, 이는 말단절단된 유전자 산물, 예를 들어 전장 단백질의 기능이 결여된 말단절단된 단백질의 발현을 유발한다. 일부 실시양태에서, 본원에 제공된 방법의 목적은 게놈 편집을 통해 기능장애 유전자의 기능을 회복시키는 것이다. 본원에 제공된 융합 단백질은, 시험관내에서 예를 들어 인간 세포 배양물에서 질환 연관 돌연변이를 교정함으로써 유전자 편집-기반 인간 치료제에 대해 검증될 수 있다. 본원에 제공된 융합 단백질, 예를 들어 본 발명의 RNA-가이드된, DNA-결합 도메인, 데아미나제 도메인 및 USP를 포함하는 융합 단백질은 임의의 단일 점 C>T 돌연변이를 교정하는 데 사용될 수 있다는 것이 관련 기술분야의 통상의 기술자에 의해 이해될 것이다. 돌연변이 C의 U로의 탈아미노화는 돌연변이의 교정으로 이어진다.

[0174] 일부 실시양태에서, 본 발명의 RNA-가이드된, DNA-결합 도메인, 데아미나제 도메인 및 USP를 포함하는 융합 단백질은 표적 유전자 또는 관심 유전자의 표적 영역에서 돌연변이를 생성하는 데 사용될 수 있다. 일부 실시양태에서, 본 발명의 융합 단백질은 표적 유전자 또는 관심 유전자의 표적 영역의 포화 돌연변이유발, 및 이어서 신규 돌연변이 및/또는 표현형을 확인하기 위한 고처리량 정방향 유전자 스크리닝에 사용될 수 있다. 다른 실시양태에서, 본원에 기재된 융합 단백질은 코딩 DNA 서열을 포함하거나 포함하지 않을 수 있는 표적화된 게놈 위치에서의 돌연변이를 생성시키는 데 사용될 수 있다. 상기 기재된 표적화된 돌연변이유발에 의해 생성된 세포주의 라이브러리는 또한 유전자 기능 또는 유전자 발현의 연구에 유용할 수 있다.

[0175] 본 발명의 융합 단백질은 또한 넌센스 돌연변이의 표적화된 삽입을 통해 녹-아웃 (KO) 세포주의 전체 라이브러리를 비롯한 KO 세포주를 효율적으로 생성하는 데 사용될 수 있다. 본 발명의 RNA-가이드된, DNA-결합 도메인, 데아미나제 도메인 및 USP를 포함하는 융합 단백질은 코딩 DNA 군주에 표적화되는 경우에 3개의 코돈 (CAA, CAG 및 CGA)을 정지 코돈 (TAG, TAA 또는 TGA)으로 전환시킬 수 있고, 비-코딩 DNA 군주에 표적화되는 경우에 TGG를 정지 코돈으로 전환시킬 수 있다. 문헌 [Billon et al. (2017, Mol Cell 67: 1068-1079; 본원에 참조로 포함

됨]]은 정지 코돈의 생성에 유용한 8종의 진핵 중에서 340만개 초과 가이드 RNA의 데이터베이스를 제공한다. 추가적으로, 문헌 [Kuscu et al. (2017, Nature Methods 14(7): 710-714; 본원에 참조로 포함됨)]은 인간 게놈에서 거의 17,000개의 유전자를 표적화하여 초기 정지 코돈을 도입할 수 있는 ~260,000개의 고유한 i-정지 sgRNA를 확인하였다. 일부 실시양태에서, KO 세포주는 진핵 세포이다. 다른 실시양태에서, KO 세포주는 원핵 세포이다. 일부 실시양태에서, 본 발명의 융합 단백질을 사용하여 생성된 KO 세포주는 인간 세포주이다. 다른 실시양태에서, KO 세포주는 포유동물 세포주, 예를 들어 마우스, 래트, 원숭이, 고양이, 개, 소, 돼지, 양 또는 말 세포주이다. 다른 실시양태에서, KO 세포주는 조류 세포주이다. 다른 실시양태에서, KO 세포주는 곤충 세포주이다. 다른 실시양태에서, KO 세포주는 미생물 세포주이다. 또 다른 실시양태에서, KO 세포주는 식물 세포주이다. 추가 실시양태에서, KO 세포주는 아라비도시스, 대두, 메이즈, 목화, 토마토, 감자 또는 콩 세포주이다. 추가 실시양태에서, 세포주는 식물 종자이다.

[0176] 일부 실시양태에서, 본원에 제공된 융합 단백질은 치료적 게놈 편집에 유용할 수 있다. 예를 들어, 본 발명의 RNA-가이드된, DNA-결합 도메인, 데아미나제 도메인 및 USP를 포함하는 융합 단백질은 PCSK9의 표적화된 넌센스 돌연변이 (전구단백질 컨버타제 서브틸리신/렉틴 유형 9)를 생성하는 데 사용될 수 있다. PCSK9는 지단백질 항상성에 수반되고, PCSK9를 차단하는 작용제는 혈액 중 저밀도 지단백질 입자 (LDL) 농도를 낮출 수 있다. 개체에서 PCSK9에서의 자연 발생 넌센스 변이체는 혈액 콜레스테롤 수준을 실질적으로 감소시키고 관상동맥 심장 질환에서의 위험을 감소시킨다 (문헌 [Cohen et al. (2006) N Engl J Med 354: 1264-1272]). 문헌 [Chadwick et al. (2017, Arteriscler Thromb Vasc Biol 37: 1741-1747; 본원에 참조로 포함됨)]은 이들이 표적화된 C>T 돌연변이를 PCSK9 유전자에 성공적으로 도입하여 마우스에서 PCSK9 단백질 수준 및 혈장 콜레스테롤 수준을 감소시킬 수 있다는 것을 발견하였다. 일부 실시양태에서, 유사한 접근법이 본 발명의 융합 단백질로 취해질 수 있다. 추가로, 관련 기술분야의 통상의 기술자는 본원에 개시된 융합 단백질이 다른 암, 및 다른 증식성 질환을 비롯한 암 이외의 질환과 연관된 다른 점 돌연변이 및 돌연변이를 교정하는 데 사용될 수 있음을 이해할 것이다.

[0177] 임의로 본원에 기재된 바와 같은 gRNA를 갖는 융합 단백질에 의해 매개되는 과정을 사용하여 변형된 관심 표적 핵산 분자를 포함하는 세포 및 유기체가 본원에 제공된다. 이들 실시양태 중 일부에서, 융합 단백질은 서열식별번호: 1-16 중 임의의 것의 아미노산 서열을 포함하는 USP, 또는 그의 활성 변이체 또는 단편을 포함한다. 일부 실시양태에서, 융합 단백질은 서열식별번호: 1-16 중 임의의 것에 대해 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98% 또는 적어도 99% 동일한 아미노산 서열을 포함하는 USP를 포함한다. 일부 실시양태에서, 융합 단백질은 데아미나제 및 RNA-가이드된, DNA-결합 폴리펩티드를 추가로 포함한다. 추가 실시양태에서, 융합 단백질은 데아미나제 및 RGN 또는 그의 변이체, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 닉카제 변이체 nAPG07433.1 (서열식별번호: 41)에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함한다. 추가 실시양태에서, 융합 단백질은 데아미나제 및 RGN 또는 그의 변이체, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 닉카제 변이체 nAPG07433.1 (서열식별번호: 41)에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함한다. 추가 실시양태에서, 융합 단백질은 데아미나제 및 RGN 또는 그의 변이체, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 닉카제 변이체 nAPG07433.1 (서열식별번호: 41)에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함한다. 추가 실시양태에서, 융합 단백질은 데아미나제 및 RGN 또는 그의 변이체, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 닉카제 변이체 nAPG07433.1 (서열식별번호: 41)에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함한다. 추가 실시양태에서, 융합 단백질은 데아미나제 및 RGN 또는 그의 변이체, 예컨대 APG07433.1 (서열식별번호: 40) 또는 그의 닉카제 변이체 nAPG07433.1 (서열식별번호: 41)을 포함한다. 다른 실시양태에서, 융합 단백질은 데아미나제 및 Cas9 또는 그의 변이체, 예컨대 예를 들어 dCas9 또는 닉카제 Cas9를 포함한다. 일부 실시양태에서, 융합 단백질은 유형 II CRISPR-Cas 폴리펩티드의 뉴클레아제-불활성 또는 닉카제 변이체를 포함한다. 다른 실시양태에서, 융합 단백질은 유형 V CRISPR-Cas 폴리펩티드의 뉴클레아제-불활성 또는 닉카제 변이체를 포함한다. 또 다른 실시양태에서, 융합 단백질은 유형 VI CRISPR-Cas 폴리펩티드의 뉴클레아제-불활성 또는 닉카제 변이체를 포함한다.

[0178] 변형된 세포는 진핵 (예를 들어, 포유동물, 식물, 곤충, 조류 세포) 또는 원핵일 수 있다. 또한, 본원에 기재된 바와 같은 융합 단백질을 이용하는 과정에 의해 변형된 적어도 하나의 뉴클레오티드 서열을 포함하는 소기관 및 배아가 제공된다. 유전자 변형된 세포, 유기체, 소기관 및 배아는 변형된 뉴클레오티드 서열에 대해 이형접합 또는 동형접합일 수 있다.

[0179] 융합 단백질의 데아미나제 도메인에 의해 도입된 돌연변이(들)는 변경된 발현 (상향-조절 또는 하향-조절), 불

활성화, 또는 변경된 단백질 산물 또는 통합된 서열의 발현을 발생시킬 수 있다. 돌연변이(들)가 유전자의 불활성화 또는 비-기능적 단백질 산물의 발현을 유발하는 경우, 유전자 변형된 세포, 유기체, 소기관 또는 배아는 "녹아웃"으로 지칭된다. 녹아웃 표현형은 결실 돌연변이 (즉, 적어도 1개의 뉴클레오타이드의 결실), 삽입 돌연변이 (즉, 적어도 1개의 뉴클레오타이드의 삽입), 또는 넌센스 돌연변이 (즉, 정지 코돈이 도입되도록 적어도 1개의 뉴클레오타이드의 치환)의 결과일 수 있다.

[0180] 다른 실시양태에서, 융합 단백질의 데아미나제 도메인에 의해 도입된 돌연변이(들)는 변이체 단백질 산물의 생성을 유발한다. 발현된 변이체 단백질 산물은 적어도 1개의 아미노산 치환 및/또는 적어도 1개의 아미노산의 부가 또는 결실을 가질 수 있다. 변이체 단백질 산물은 야생형 단백질과 비교 시 변형된 특징 또는 활성, 예컨대 비제한적으로 변경된 효소적 활성 또는 기질 특이성을 나타낼 수 있다.

[0181] 또 다른 실시양태에서, 융합 단백질의 데아미나제 도메인에 의해 도입된 돌연변이(들)는 단백질의 변경된 발현 패턴을 유발할 수 있다. 비제한적 예로서, 단백질 산물의 발현을 제어하는 조절 영역에서의 돌연변이(들)는 단백질 산물의 과다발현 또는 하향조절 또는 변경된 조직 또는 일시적 발현 패턴을 유발할 수 있다.

[0182] 본 개시내용의 일부 측면은 RNA-가이드된, DNA-결합 폴리펩티드, 예컨대 RGN 폴리펩티드, 예를 들어 뉴클레아제-불활성 Cas9 도메인, 및 본 발명의 데아미나제, 및 임의로 Cas9 도메인과 데아미나제 사이에 위치한 링커를 포함하는 융합 단백질을 포함하는 키트를 제공한다. 또한, 일부 실시양태에서, 키트는 적합한 시약, 완충제, 및/또는 예를 들어 시험관내 또는 생체내 DNA 또는 RNA 편집을 위한 융합 단백질의 사용에 대한 지침서를 포함한다. 일부 실시양태에서, 키트는 핵산 서열의 표적화된 편집을 위한 적합한 gRNA의 설계 및 사용에 관한 지침서를 포함한다.

[0183] X: USP의 추가의 적용

[0184] 본원에 기재된 USP는 또한 게놈 염기 편집을 넘어서는 유용성을 보유한다. 일반적으로, USP는 DNA 분자에서 우라실 핵염기를 안정화시키는 것이 요구되는 적용에 유용하다. 천연 과정을 통해 또는 인간의 손에 의해, 우라실은 반응성 산소 종, 이온화 방사선 및/또는 알킬화제에 의해 유발된 DNA 손상에 의해 게놈 DNA에 도입될 수 있다. DNA 복구의 메커니즘, 예컨대 염기 절제 복구 (BER) 경로에 대한 연구, 또는 DNA 복구 능력을 측정하는 연구는 우라실의 복구를 억제하기 위해 본 발명의 USP를 사용할 수 있다.

[0185] 추가적으로, 본원에 기재된 USP는 다양한 암의 치료에 유용할 수 있다. 예를 들어, 5-플루오로우라실 (5-FU) 및 그의 데옥시리보뉴클레오타이드 대사물 5-플루오로데옥시우리딘 (5-FdU)을 비롯한 플루오로피리미딘은 결장직 장암을 비롯한 다양한 고형 종양의 치료에 널리 사용되어 왔다. 5-FdU는 티미딜레이트 신타제의 억제를 통해 활성이며, 이는 결과적으로 우라실 및 5-FU 혼입을 세포의 게놈에 도입한다. 상기 기재된 바와 같이, 염기 복구 효소, 예컨대 UDG는 게놈 DNA에서 우라실 핵염기를 인식하고, 이를 제거한다. 문헌 [Yan et al. (2016; Oncotarget 7(37): 59299-59313)]은 UDG 고갈된 세포가 5-FdU 처리 후에 정지되고 지속적인 DNA 손상을 나타냄을 발견하였으며, 이는 우라실 및 5-FU의 제거에서의 UDG의 작용이 종양의 5-FdU 처리의 유효성에서 역할을 한다는 것을 나타낸다. 플루오로피리미딘과 조합된 본 발명의 USP의 전달은 종양의 이러한 치료의 유효성을 증진시킬 수 있다. 따라서, 본원에 개시된 USP 및 플루오로피리미딘을 포함하는 제약 조성물이, 이러한 제약 조성물의 유효량으로 암을 치료하는 방법과 함께 제공된다. 플루오로피리미딘은 카페시타빈, 카르포푸르, 독시플루리딘, 플루오로우라실, 5-플루오로데옥시우리딘 및 테가푸르를 포함하는 항암 항대사물의 한 부류이다.

[0186] 단수표현 (관사 "a" 및 "an")은 관사의 문법적 대상의 하나 또는 하나 초과 (즉, 적어도 하나)를 지칭하기 위해 본원에 사용된다. 예를 들어, "폴리펩티드"는 하나 이상의 폴리펩티드를 의미한다.

[0187] 본 명세서에 언급된 모든 공보 및 특허 출원은 본 개시내용이 속하는 기술분야의 관련 기술분야의 통상의 기술자의 수준을 나타낸다. 모든 공보 및 특허 출원은 각각의 개별 공보 또는 특허 출원이 구체적이고 개별적으로 참조로 포함되는 것으로 나타난 것과 동일한 정도로 본원에 참조로 포함된다.

[0188] 상기 발명은 이해의 명료함을 위해 예시 및 예로서 일부 상세히 기재되었지만, 첨부된 청구범위의 범주 내에서 특정 변화 및 변형이 실시될 수 있음이 명백할 것이다.

[0189] 비제한적인 실시양태에는 하기를 포함한다:

[0190] 1. a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;

[0191] b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는

- [0192] c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성
- [0193] 을 갖는 아미노산 서열을 포함하는 단리된 폴리펩티드로서,
- [0194] 여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖고, 상기 폴리펩티드는 이중 아미노산 서열을 추가로 포함하
- [0195] 는 것인 단리된 폴리펩티드.
- [0196] 2. 실시양태 1에 있어서, 아미노산 서열이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 것인 단리된 폴리펩티드.
- [0197] 3. 실시양태 1에 있어서, 아미노산 서열이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 것인 단리된 폴리펩티드.
- [0198] 4. 실시양태 1에 있어서, 아미노산 서열이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 것인 단리된 폴리펩티드.
- [0199] 5. 실시양태 1에 있어서, 아미노산 서열이 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 것인 단리된 폴리펩티드.
- [0200] 6. 실시양태 1에 있어서, 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 단리된 폴리펩티드.
- [0201] 7. 비-자연 발생 제약상 허용되는 담체, 및
- [0202] a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;
- [0203] b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는
- [0204] c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성
- [0205] 을 갖는 아미노산 서열을 포함하는 폴리펩티드를 포함하는 제약 조성물로서,
- [0206] 여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖는 것인
- [0207] 제약 조성물.
- [0208] 8. 실시양태 7에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 제약 조성물.
- [0209] 9. 실시양태 7에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 제약 조성물.
- [0210] 10. 실시양태 7에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 제약 조성물.
- [0211] 11. 실시양태 7에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 제약 조성물.
- [0212] 12. 비-자연 발생 제약상 허용되는 담체, 및
- [0213] a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;
- [0214] b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는
- [0215] c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성
- [0216] 을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산 분자를 포함하는 제약 조성물로서,
- [0217] 여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖는 것인
- [0218] 제약 조성물.
- [0219] 13. 실시양태 12에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 제약 조성물.

- [0220] 14. 실시양태 12에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 제약 조성물.
- [0221] 15. 실시양태 12에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 제약 조성물.
- [0222] 16. 실시양태 12에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 제약 조성물.
- [0223] 17. 실시양태 7 또는 12에 있어서, 폴리펩티드가 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 것인 제약 조성물.
- [0224] 18. 실시양태 7-17 중 어느 하나에 있어서, 플루오로피리미딘을 추가로 포함하는 제약 조성물.
- [0225] 19. a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;
- [0226] b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는
- [0227] c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성
- [0228] 을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산 분자로서,
- [0229] 여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖고;
- [0230] 여기서 상기 핵산 분자는 상기 폴리뉴클레오티드에 작동가능하게 연결된 이중 프로모터를 추가로 포함하는 것인
- [0231] 핵산 분자.
- [0232] 20. 실시양태 19에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0233] 21. 실시양태 19에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0234] 22. 실시양태 19에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0235] 23. 실시양태 19에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0236] 24. 실시양태 19에 있어서, 폴리펩티드가 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 것인 핵산 분자.
- [0237] 25. 플루오로피리미딘, 및
- [0238] a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;
- [0239] b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는
- [0240] c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성
- [0241] 을 갖는 아미노산 서열을 포함하는 폴리펩티드를 포함하는 조성물로서,
- [0242] 여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖는 것인
- [0243] 조성물.
- [0244] 26. 실시양태 25에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 조성물.
- [0245] 27. 실시양태 25에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 조성물.
- [0246] 28. 실시양태 25에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 조성물.
- [0247] 29. 실시양태 25에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 아

미노산 서열을 포함하는 것인 조성물.

- [0248] 30. 플루오로피리미딘, 및
- [0249] a) 서열식별번호: 1, 2, 4, 5, 및 7-15 중 어느 하나에 대해 적어도 80% 서열 동일성;
- [0250] b) 서열식별번호: 3 또는 16에 대해 적어도 81% 서열 동일성; 또는
- [0251] c) 서열식별번호: 6에 대해 적어도 82% 서열 동일성
- [0252] 을 갖는 아미노산 서열을 포함하는 폴리펩티드를 코딩하는 핵산 분자를 포함하는 조성물로서,
- [0253] 여기서 상기 폴리펩티드는 우라실 안정화 활성을 갖는 것인
- [0254] 조성물.
- [0255] 31. 실시양태 30에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 조성물.
- [0256] 32. 실시양태 30에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 조성물.
- [0257] 33. 실시양태 30에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 조성물.
- [0258] 34. 실시양태 30에 있어서, 폴리펩티드가 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 조성물.
- [0259] 35. 실시양태 25 또는 30에 있어서, 폴리펩티드가 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 것인 조성물.
- [0260] 36. (i) DNA-결합 폴리펩티드; (ii) 데아미나제; 및 (iii) 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하는 융합 단백질.
- [0261] 37. 실시양태 36에 있어서, 적어도 하나의 USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 것인 융합 단백질.
- [0262] 38. 실시양태 36에 있어서, 적어도 하나의 USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 것인 융합 단백질.
- [0263] 39. 실시양태 36에 있어서, 적어도 하나의 USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 것인 융합 단백질.
- [0264] 40. 실시양태 36에 있어서, 적어도 하나의 USP가 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 것인 융합 단백질.
- [0265] 41. 실시양태 36에 있어서, USP가 서열식별번호: 33-39 중 어느 하나의 서열을 갖는 것인 융합 단백질.
- [0266] 42. 실시양태 36 또는 41에 있어서, 데아미나제가 시티딘 데아미나제인 융합 단백질.
- [0267] 43. 실시양태 42에 있어서, 시티딘 데아미나제가 활성화-유도된 시티딘 데아미나제 (AID) 또는 데아미나제의 아포지단백질 B mRNA-편집 복합체 (APOBEC) 패밀리의 구성원인 융합 단백질.
- [0268] 44. 실시양태 43에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.
- [0269] 45. 실시양태 43에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.
- [0270] 46. 실시양태 43에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.
- [0271] 47. 실시양태 43에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.

- [0272] 48. 실시양태 43에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.
- [0273] 49. 실시양태 36-44 중 어느 하나에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 융합 단백질.
- [0274] 50. 실시양태 36-44 중 어느 하나에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 융합 단백질.
- [0275] 51. 실시양태 50에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 융합 단백질.
- [0276] 52. 실시양태 51에 있어서, RGN이 유형 II CRISPR-Cas 폴리펩티드인 융합 단백질.
- [0277] 53. 실시양태 51에 있어서, RGN이 유형 V CRISPR-Cas 폴리펩티드인 융합 단백질.
- [0278] 54. 실시양태 51에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.
- [0279] 55. 실시양태 51에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.
- [0280] 56. 실시양태 51에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.
- [0281] 57. 실시양태 51에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.
- [0282] 58. 실시양태 51에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 융합 단백질.
- [0283] 59. 실시양태 51-58 중 어느 하나에 있어서, RGN이 RGN 니카제인 융합 단백질.
- [0284] 60. 실시양태 51-59 중 어느 하나에 있어서, RGN, 시티딘 데아미나제 및 USP를 포함하는 융합 단백질.
- [0285] 61. 실시양태 60에 있어서, 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 USP를 포함하는 융합 단백질.
- [0286] 62. 실시양태 60에 있어서, 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 USP를 포함하는 융합 단백질.
- [0287] 63. 실시양태 60에 있어서, 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 USP를 포함하는 융합 단백질.
- [0288] 64. 실시양태 60에 있어서, 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 USP를 포함하는 융합 단백질.
- [0289] 65. 실시양태 60에 있어서, 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 100% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 100% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 USP를 포함하는 융합 단백질.
- [0290] 66. 실시양태 36-66 중 어느 하나에 있어서, 적어도 하나의 핵 국제화 신호 (NLS)를 추가로 포함하는 융합 단백질.
- [0291] 67. (i) DNA-결합 폴리펩티드; (ii) 데아미나제; 및 (iii) 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하는 융합 단백질을 코딩하는 폴리뉴클레오티드를 포함하는 핵산 분자로서,
- [0292] 여기서 USP는

- [0293] a) 서열식별번호: 17-32 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖거나,
- [0294] b) 서열식별번호: 17-32 중 어느 하나에 제시되거나,
- [0295] c) 서열식별번호: 1-16에 대해 적어도 80% 동일한 아미노산 서열을 코딩하고, 서열식별번호: 33-39 중 어느 하나의 서열을 추가로 보유하거나,
- [0296] d) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열에 대해 적어도 80% 동일한 아미노산 서열을 코딩하거나, 또는
- [0297] e) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열을 코딩하는
- [0298] 뉴클레오티드 서열에 의해 코딩되는 것인
- [0299] 핵산 분자.
- [0300] 68. 실시양태 67에 있어서, USP가
- [0301] a) 서열식별번호: 17-32 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖거나,
- [0302] b) 서열식별번호: 17-32 중 어느 하나에 제시되거나,
- [0303] c) 서열식별번호: 1-16에 대해 적어도 85% 동일한 아미노산 서열을 코딩하고, 서열식별번호: 33-39 중 어느 하나의 서열을 추가로 보유하거나,
- [0304] d) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열에 대해 적어도 85% 동일한 아미노산 서열을 코딩하거나, 또는
- [0305] e) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열을 코딩하는
- [0306] 뉴클레오티드 서열에 의해 코딩되는 것인 핵산 분자.
- [0307] 69. 실시양태 67에 있어서, USP가
- [0308] a) 서열식별번호: 17-32 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖거나,
- [0309] b) 서열식별번호: 17-32 중 어느 하나에 제시되거나,
- [0310] c) 서열식별번호: 1-16에 대해 적어도 90% 동일한 아미노산 서열을 코딩하고, 서열식별번호: 33-39 중 어느 하나의 서열을 추가로 보유하거나,
- [0311] d) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열에 대해 적어도 90% 동일한 아미노산 서열을 코딩하거나, 또는
- [0312] e) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열을 코딩하는
- [0313] 뉴클레오티드 서열에 의해 코딩되는 것인 핵산 분자.
- [0314] 70. 실시양태 67에 있어서, USP가
- [0315] a) 서열식별번호: 17-32 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖거나,
- [0316] b) 서열식별번호: 17-32 중 어느 하나에 제시되거나,
- [0317] c) 서열식별번호: 1-16에 대해 적어도 95% 동일한 아미노산 서열을 코딩하고, 서열식별번호: 33-39 중 어느 하나의 서열을 추가로 보유하거나,
- [0318] d) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열에 대해 적어도 95% 동일한 아미노산 서열을 코딩하거나, 또는
- [0319] e) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열을 코딩하는
- [0320] 뉴클레오티드 서열에 의해 코딩되는 것인 핵산 분자.
- [0321] 71. 실시양태 67에 있어서, USP가
- [0322] a) 서열식별번호: 17-32 중 어느 하나에 제시되거나,

- [0323] b) 서열식별번호: 1-16과 100% 동일한 아미노산 서열을 코딩하고, 서열식별번호: 33-39 중 어느 하나의 서열을 추가로 보유하거나, 또는
- [0324] c) 서열식별번호: 1-16 중 어느 하나에 제시된 아미노산 서열을 코딩하는
- [0325] 뉴클레오타이드 서열에 의해 코딩되는 것인 핵산 분자.
- [0326] 72. 실시양태 67에 있어서, 데아미나제가 시티딘 데아미나제인 핵산 분자.
- [0327] 73. 실시양태 72에 있어서, 시티딘 데아미나제가 활성화-유도된 시티딘 데아미나제 (AID) 또는 데아미나제의 아포지단백질 B mRNA-편집 복합체 (APOBEC) 패밀리의 구성원인 핵산 분자.
- [0328] 74. 실시양태 73에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0329] 75. 실시양태 73에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0330] 76. 실시양태 73에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0331] 77. 실시양태 73에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0332] 78. 실시양태 73에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0333] 79. 실시양태 67-78 중 어느 하나에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 핵산 분자.
- [0334] 80. 실시양태 67-78 중 어느 하나에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 핵산 분자.
- [0335] 81. 실시양태 80에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 핵산 분자.
- [0336] 82. 실시양태 81에 있어서, RGN이 유형 II CRISPR-Cas 폴리펩티드인 핵산 분자.
- [0337] 83. 실시양태 81에 있어서, RGN이 유형 V CRISPR-Cas 폴리펩티드인 핵산 분자.
- [0338] 84. 실시양태 81에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0339] 85. 실시양태 82에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0340] 86. 실시양태 81에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0341] 87. 실시양태 81에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0342] 88. 실시양태 81에 있어서, RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 것인 핵산 분자.
- [0343] 89. 실시양태 81-88 중 어느 하나에 있어서, RGN이 RGN Nickase인 핵산 분자.
- [0344] 90. 실시양태 81-89 중 어느 하나에 있어서, 융합 단백질이 RGN, 시티딘 데아미나제 및 USP를 포함하는 것인 핵산 분자.
- [0345] 91. 실시양태 90에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 USP를 포함하는

것인 핵산 분자.

- [0346] 92. 실시양태 90에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 USP를 포함하는 것인 핵산 분자.
- [0347] 93. 실시양태 90에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 USP를 포함하는 것인 핵산 분자.
- [0348] 94. 실시양태 90에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 USP를 포함하는 것인 핵산 분자.
- [0349] 95. 실시양태 90에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 100% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 100% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 USP를 포함하는 것인 핵산 분자.
- [0350] 96. 실시양태 67-95 중 어느 하나에 있어서, 융합 단백질을 코딩하는 폴리뉴클레오티드가 그의 5' 말단에서 이중 프로모터에 작동가능하게 연결된 것인 핵산 분자.
- [0351] 97. 실시양태 67-95 중 어느 하나에 있어서, 융합 단백질을 코딩하는 폴리뉴클레오티드가 그의 3' 말단에서 이중 종결인자에 작동가능하게 연결된 것인 핵산 분자.
- [0352] 98. 실시양태 67-97 중 어느 하나에 있어서, 융합 단백질이 하나 이상의 핵 국제화 신호를 포함하는 것인 핵산 분자.
- [0353] 99. 실시양태 67-98 중 어느 하나에 있어서, 융합 단백질이 진핵 세포에서의 발현에 대해 코돈 최적화된 것인 핵산 분자.
- [0354] 100. 실시양태 67-99 중 어느 하나에 있어서, 융합 단백질이 원핵 세포에서의 발현에 대해 코돈 최적화된 것인 핵산 분자.
- [0355] 101. 실시양태 67-100 중 어느 하나의 핵산 분자를 포함하는 벡터.
- [0356] 102. 실시양태 81-95 중 어느 하나의 핵산 분자를 포함하고, 표적 서열에 혼성화할 수 있는 가이드 RNA (gRNA)를 코딩하는 적어도 하나의 뉴클레오티드 서열을 추가로 포함하는 벡터.
- [0357] 103. 실시양태 102에 있어서, gRNA가 단일 가이드 RNA인 벡터.
- [0358] 104. 실시양태 102에 있어서, gRNA가 이중 가이드 RNA인 벡터.
- [0359] 105. 실시양태 36-66 중 어느 하나의 융합 단백질을 포함하는 세포.
- [0360] 106. 실시양태 51-66 중 어느 하나의 융합 단백질을 포함하고, 가이드 RNA를 추가로 포함하는 세포.
- [0361] 107. 실시양태 67-100 중 어느 하나의 핵산 분자를 포함하는 세포.
- [0362] 108. 실시양태 101-104 중 어느 하나의 벡터를 포함하는 세포.
- [0363] 109. 실시양태 105-108 중 어느 하나에 있어서, 원핵 세포인 세포.
- [0364] 110. 실시양태 105-108 중 어느 하나에 있어서, 진핵 세포인 세포.
- [0365] 111. 실시양태 110에 있어서, 곤충, 조류 또는 포유동물 세포인 세포.
- [0366] 112. 실시양태 110에 있어서, 식물 또는 진균 세포인 세포.
- [0367] 113. 제약상 허용되는 담체 및 실시양태 19-24 및 67-100 중 어느 하나의 핵산 분자, 실시양태 25-35 중 어느 하나의 조성물, 실시양태 36-66 중 어느 하나의 융합 단백질, 실시양태 101-104 중 어느 하나의 벡터, 또는 실

시양태 105-111 중 어느 하나의 세포를 포함하는 제약 조성물.

- [0368] 114. 실시양태 105-112 중 어느 하나의 세포를 융합 단백질이 발현되는 조건 하에 배양하는 것을 포함하는, 융합 단백질을 제조하는 방법.
- [0369] 115. 실시양태 67-100 중 어느 하나의 핵산 분자 또는 실시양태 101-104 중 어느 하나의 벡터를 세포에 도입하고, 세포를 융합 단백질이 발현되는 조건 하에 배양하는 것을 포함하는, 융합 단백질을 제조하는 방법.
- [0370] 116. 실시양태 114 또는 115에 있어서, 상기 융합 단백질을 정제하는 것을 추가로 포함하는 방법.
- [0371] 117. 실시양태 81-95 중 어느 하나의 핵산 분자 및 가이드 RNA를 코딩하는 발현 카세트를 포함하는 핵산 분자, 또는 실시양태 102-104 중 어느 하나의 벡터를 세포에 도입하고, 세포를 융합 단백질 및 gRNA가 발현되어 RGN 융합 리보핵단백질 복합체를 형성하는 조건 하에 배양하는 것을 포함하는, RGN 융합 리보핵단백질 복합체를 제조하는 방법.
- [0372] 118. 실시양태 117에 있어서, 상기 RGN 융합 리보핵단백질 복합체를 정제하는 것을 추가로 포함하는 방법.
- [0373] 119. 표적 DNA 서열을 포함하는 표적 DNA 분자를 변형시키기 위한 시스템으로서,
- [0374] 상기 시스템은
- [0375] a) RNA-가이드된 뉴클레아제 폴리펩티드 (RGN), 시티딘 데아미나제, 및 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하며, 여기서 USP는 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 동일한 것인 융합 단백질, 또는 상기 융합 단백질을 코딩하는 뉴클레오티드 서열; 및
- [0376] b) 상기 표적 DNA 서열에 혼성화할 수 있는 하나 이상의 가이드 RNA, 또는 하나 이상의 가이드 RNA (gRNA)를 코딩하는 하나 이상의 뉴클레오티드 서열
- [0377] 을 포함하고;
- [0378] 여기서 하나 이상의 가이드 RNA를 코딩하고 융합 단백질을 코딩하는 상기 뉴클레오티드 서열은 각각 상기 뉴클레오티드 서열에 이중인 프로모터에 작동가능하게 연결되고;
- [0379] 여기서 하나 이상의 가이드 RNA는 융합 단백질과 복합체를 형성하여, 상기 융합 단백질이 상기 표적 DNA 서열에 결합하고 표적 DNA 분자를 변형시키도록 지시할 수 있는 것인
- [0380] 시스템.
- [0381] 120. 실시양태 119에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 동일한 것인 시스템.
- [0382] 121. 실시양태 119에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 동일한 것인 시스템.
- [0383] 122. 실시양태 119에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 동일한 것인 시스템.
- [0384] 123. 실시양태 119에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 100% 동일한 것인 시스템.
- [0385] 124. 실시양태 119에 있어서, USP가 서열식별번호: 33-39 중 어느 하나에 제시된 서열을 포함하는 것인 시스템.
- [0386] 125. 실시양태 119-124 중 어느 하나에 있어서, 표적 DNA 서열이 RGN에 의해 인식되는 프로토스케이스 인접 모티프 (PAM)에 인접하여 위치하는 것인 시스템.
- [0387] 126. 실시양태 119-125 중 어느 하나에 있어서, 표적 DNA 분자가 세포 내에 있는 것인 시스템.
- [0388] 127. 실시양태 126에 있어서, 세포가 진핵 세포인 시스템.
- [0389] 128. 실시양태 127에 있어서, 진핵 세포가 식물 세포인 시스템.
- [0390] 129. 실시양태 127에 있어서, 진핵 세포가 포유동물 세포인 시스템.
- [0391] 130. 실시양태 127에 있어서, 진핵 세포가 곤충 세포인 시스템.
- [0392] 131. 실시양태 126에 있어서, 세포가 원핵 세포인 시스템.
- [0393] 132. 실시양태 119-131 중 어느 하나에 있어서, 융합 단백질의 RGN이 유형 II CRISPR-Cas 폴리펩티드인 시스템.
- [0394] 133. 실시양태 119-131 중 어느 하나에 있어서, 융합 단백질의 RGN이 유형 V CRISPR-Cas 폴리펩티드인 시스템.

- [0395] 134. 실시양태 119-131 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80% 동일한 것인 시스템.
- [0396] 135. 실시양태 119-131 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 85% 동일한 것인 시스템.
- [0397] 136. 실시양태 119-131 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 90% 동일한 것인 시스템.
- [0398] 137. 실시양태 119-131 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 95% 동일한 것인 시스템.
- [0399] 138. 실시양태 119-131 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 100% 동일한 것인 시스템.
- [0400] 139. 실시양태 119-138 중 어느 하나에 있어서, 융합 단백질의 RGN이 RGN 닉카제인 시스템.
- [0401] 140. 실시양태 119-139 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 동일한 것인 시스템.
- [0402] 141. 실시양태 119-139 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 동일한 것인 시스템.
- [0403] 142. 실시양태 119-139 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 동일한 것인 시스템.
- [0404] 143. 실시양태 119-139 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 동일한 것인 시스템.
- [0405] 144. 실시양태 119-139 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 100% 동일한 것인 시스템.
- [0406] 145. 실시양태 119-144 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41, 95 및 142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 USP를 포함하는 것인 시스템.
- [0407] 146. 실시양태 119-144 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41, 95 및 142 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 USP를 포함하는 것인 시스템.
- [0408] 147. 실시양태 119-144 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41, 95 및 142 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 USP를 포함하는 것인 시스템.
- [0409] 148. 실시양태 119-144 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41, 95 및 142 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 USP를 포함하는 것인 시스템.
- [0410] 149. 실시양태 119-144 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41, 95 및 142 중 어느 하나에 대해 100% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 100% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 USP를 포함하는 것인 시스템.
- [0411] 150. 실시양태 119-149 중 어느 하나에 있어서, 융합 단백질이 하나 이상의 핵 국제화 신호를 포함하는 것인 시스템.

- [0412] 151. 실시양태 119-150 중 어느 하나에 있어서, 융합 단백질이 진핵 세포에서의 발현에 대해 코돈 최적화된 것인 시스템.
- [0413] 152. 실시양태 119-151 중 어느 하나에 있어서, 하나 이상의 가이드 RNA를 코딩하는 뉴클레오타이드 서열 및 융합 단백질을 코딩하는 뉴클레오타이드 서열이 하나의 벡터 상에 위치하는 것인 시스템.
- [0414] 153. 실시양태 119-152 중 어느 하나에 따른 시스템을 표적 DNA 분자 또는 표적 DNA 분자를 포함하는 세포에 전달하는 것을 포함하는, 표적 DNA 서열을 포함하는 표적 DNA 분자를 변형시키는 방법.
- [0415] 154. 실시양태 153에 있어서, 상기 변형된 표적 DNA 분자가 표적 DNA 분자 내 적어도 하나의 뉴클레오타이드의 C>T 돌연변이를 포함하는 것인 방법.
- [0416] 155. 실시양태 153에 있어서, 상기 변형된 표적 DNA 분자가 표적 DNA 서열 내에 적어도 하나의 뉴클레오타이드의 C>T 돌연변이를 포함하는 것인 방법.
- [0417] 156. 표적 서열을 포함하는 표적 DNA 분자를 변형시키는 방법으로서,
- [0418] a) RGN-데아미나제-USP 리보뉴클레오타이드 복합체의 형성에 적합한 조건 하에
- [0419] i) 표적 DNA 서열에 혼성화할 수 있는 하나 이상의 가이드 RNA; 및
- [0420] ii) RNA-가이드된 뉴클레아제 폴리펩티드 (RGN), 시티딘 데아미나제 및 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하며, 여기서 USP는 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 동일한 것인 융합 단백질
- [0421] 을 조합하여 시험관내에서 RGN-데아미나제-USP 리보뉴클레오타이드 복합체를 조립하는 단계; 및
- [0422] b) 상기 표적 DNA 분자 또는 상기 표적 DNA 분자를 포함하는 세포를 시험관내-조립된 RGN-데아미나제-USP 리보뉴클레오타이드 복합체와 접촉시키는 단계
- [0423] 를 포함하고;
- [0424] 여기서 하나 이상의 가이드 RNA는 표적 DNA 서열에 혼성화하여, 상기 융합 단백질이 상기 표적 DNA 서열에 결합하도록 지시하고, 표적 DNA 분자의 변형이 일어나는 것인
- [0425] 방법.
- [0426] 157. 실시양태 156에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 동일한 것인 방법.
- [0427] 158. 실시양태 156에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 동일한 것인 방법.
- [0428] 159. 실시양태 156에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 동일한 것인 방법.
- [0429] 160. 실시양태 156에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 100% 동일한 것인 방법.
- [0430] 161. 실시양태 156에 있어서, USP가 서열식별번호: 33-39 중 어느 하나의 서열을 포함하는 것인 방법.
- [0431] 162. 실시양태 156-161 중 어느 하나에 있어서, 상기 변형된 표적 DNA 분자가 표적 DNA 분자 내 적어도 하나의 뉴클레오타이드의 C>T 돌연변이를 포함하는 것인 방법.
- [0432] 163. 실시양태 156-161 중 어느 하나에 있어서, 상기 변형된 표적 DNA 분자가 표적 DNA 서열 내에 적어도 하나의 뉴클레오타이드의 C>T 돌연변이를 포함하는 것인 방법.
- [0433] 164. 실시양태 156-163 중 어느 하나에 있어서, 융합 단백질의 RGN이 유형 II CRISPR-Cas 폴리펩티드인 방법.
- [0434] 165. 실시양태 156-163 중 어느 하나에 있어서, 융합 단백질의 RGN이 유형 V CRISPR-Cas 폴리펩티드인 방법.
- [0435] 166. 실시양태 156-163 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 80% 동일한 것인 방법.
- [0436] 167. 실시양태 156-163 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 85% 동일한 것인 방법.
- [0437] 168. 실시양태 156-163 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 90% 동일한 것인 방법.

- [0438] 169. 실시양태 156-163 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 적어도 95% 동일한 것인 방법.
- [0439] 170. 실시양태 156-163 중 어느 하나에 있어서, 융합 단백질의 RGN이 서열식별번호: 40 및 95-142 중 어느 하나에 대해 100% 동일한 것인 방법.
- [0440] 171. 실시양태 156-170 중 어느 하나에 있어서, 융합 단백질의 RGN이 RGN 니카제인 방법.
- [0441] 172. 실시양태 156-171 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 동일한 것인 방법.
- [0442] 173. 실시양태 156-171 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 동일한 것인 방법.
- [0443] 174. 실시양태 156-171 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 동일한 것인 방법.
- [0444] 175. 실시양태 156-171 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 동일한 것인 방법.
- [0445] 176. 실시양태 156-171 중 어느 하나에 있어서, 시티딘 데아미나제가 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 100% 동일한 것인 방법.
- [0446] 177. 실시양태 156-176 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 USP를 포함하는 것인 방법.
- [0447] 178. 실시양태 156-176 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 USP를 포함하는 것인 방법.
- [0448] 179. 실시양태 156-176 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 USP를 포함하는 것인 방법.
- [0449] 180. 실시양태 156-176 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 USP를 포함하는 것인 방법.
- [0450] 181. 실시양태 156-176 중 어느 하나에 있어서, 융합 단백질이 서열식별번호: 40, 41 및 95-142 중 어느 하나에 대해 100% 서열 동일성을 갖는 RGN, 서열식별번호: 47, 48 및 76-94 중 어느 하나에 대해 100% 서열 동일성을 갖는 시티딘 데아미나제, 및 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 USP를 포함하는 것인 방법.
- [0451] 182. 실시양태 156-181 중 어느 하나에 있어서, 융합 단백질이 하나 이상의 핵 국제화 신호를 포함하는 것인 방법.
- [0452] 183. 실시양태 156-182 중 어느 하나에 있어서, 융합 단백질이 진행 세포에서의 발현에 대해 코돈 최적화된 것인 방법.
- [0453] 184. 실시양태 156-183 중 어느 하나에 있어서, 상기 표적 DNA 서열이 프로토스페이서 인접 모티프 (PAM)에 인접하여 위치하는 것인 방법.
- [0454] 185. 실시양태 156-184 중 어느 하나에 있어서, 표적 DNA 분자가 세포 내에 있는 것인 방법.
- [0455] 186. 실시양태 185에 있어서, 세포가 진행 세포인 방법.

- [0456] 187. 실시양태 186에 있어서, 진핵 세포가 식물 세포인 방법.
- [0457] 188. 실시양태 186에 있어서, 진핵 세포가 포유동물 세포인 방법.
- [0458] 189. 실시양태 186에 있어서, 진핵 세포가 곤충 세포인 방법.
- [0459] 190. 실시양태 185에 있어서, 세포가 원핵 세포인 방법.
- [0460] 191. 실시양태 185-190 중 어느 하나에 있어서, 상기 변형된 DNA 분자를 포함하는 세포를 선택하는 것을 추가로 포함하는 방법.
- [0461] 192. 실시양태 191의 방법에 따른 변형된 표적 DNA 서열을 포함하는 세포.
- [0462] 193. 실시양태 192에 있어서, 진핵 세포인 세포.
- [0463] 194. 실시양태 193에 있어서, 진핵 세포가 식물 세포인 세포.
- [0464] 195. 실시양태 194의 세포를 포함하는 식물.
- [0465] 196. 실시양태 194의 세포를 포함하는 종자.
- [0466] 197. 실시양태 193에 있어서, 진핵 세포가 포유동물 세포인 세포.
- [0467] 198. 실시양태 193에 있어서, 진핵 세포가 곤충 세포인 세포.
- [0468] 199. 실시양태 192에 있어서, 원핵 세포인 세포.
- [0469] 200. 유전적 유전성 질환에 대한 원인 돌연변이에서의 교정으로 유전자 변형된 세포를 생성하는 방법으로서,
- [0470] 상기 방법은 세포에
- [0471] a) RNA-가이드된 뉴클레아제 폴리펩티드 (RGN), 시티딘 데아미나제, 및 적어도 하나의 우라실 안정화 폴리펩티드 (USP)를 포함하며, 여기서 USP는 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 동일한 것인 융합 단백질, 또는 상기 융합 단백질을 코딩하는 폴리뉴클레오티드로서, 여기서 융합 단백질을 코딩하는 상기 폴리뉴클레오티드는 세포에서 융합 단백질의 발현을 가능하게 하는 프로모터에 작동가능하게 연결된 것인 폴리뉴클레오티드; 및
- [0472] b) 표적 DNA 서열에 혼성화할 수 있는 하나 이상의 가이드 RNA (gRNA), 또는 상기 gRNA를 코딩하는 폴리뉴클레오티드로서, 여기서 gRNA를 코딩하는 상기 폴리뉴클레오티드는 세포에서 gRNA의 발현을 가능하게 하는 프로모터에 작동가능하게 연결된 것인 폴리뉴클레오티드
- [0473] 를 도입하는 것을 포함하고;
- [0474] 이로써 융합 단백질 및 gRNA가 원인 돌연변이의 게놈 위치를 표적화하고 게놈 서열을 변형시켜 원인 돌연변이를 제거하는 것인
- [0475] 방법.
- [0476] 201. 실시양태 200에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 동일한 것인 방법.
- [0477] 202. 실시양태 200에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 동일한 것인 방법.
- [0478] 203. 실시양태 200에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 동일한 것인 방법.
- [0479] 204. 실시양태 200에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 100% 동일한 것인 방법.
- [0480] 205. 실시양태 200에 있어서, 상기 융합 단백질의 RGN이 니카제인 방법.
- [0481] 206. 실시양태 200 또는 205에 있어서, USP가 서열식별번호: 33-39 중 어느 하나의 서열을 포함하는 것인 방법.
- [0482] 207. 실시양태 200-206 중 어느 하나에 있어서, 게놈 변형이 표적 DNA 서열 내에 적어도 하나의 뉴클레오티드의 C>T 돌연변이를 도입하는 것을 포함하는 것인 방법.
- [0483] 208. 실시양태 200-207 중 어느 하나에 있어서, 세포가 동물 세포인 방법.
- [0484] 209. 실시양태 208에 있어서, 동물 세포가 포유동물 세포인 방법.

- [0485] 210. 실시양태 209에 있어서, 세포가 개, 고양이, 마우스, 래트, 토끼, 말, 양, 염소, 소, 돼지 또는 인간으로부터 유래된 것인 방법.
- [0486] 211. 실시양태 200-210 중 어느 하나에 있어서, 원인 돌연변이의 교정이 정지 코돈을 도입하는 것을 포함하는 것인 방법.
- [0487] 212. a) (i) DNA-결합 폴리펩티드; 및 (ii) 데아미나제를 포함하는 융합 단백질; 또는 융합 단백질을 코딩하는 핵산 분자; 및
- [0488] b) 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP); 또는 USP를 코딩하는 핵산 분자
- [0489] 를 포함하는 조성물.
- [0490] 213. 실시양태 212에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 것인 조성물.
- [0491] 214. 실시양태 212에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 것인 조성물.
- [0492] 215. 실시양태 212에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 것인 조성물.
- [0493] 216. 실시양태 212에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 것인 조성물.
- [0494] 217. 실시양태 212에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 조성물.
- [0495] 218. 실시양태 212에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 조성물.
- [0496] 219. 실시양태 212에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 조성물.
- [0497] 220. 실시양태 212에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 조성물.
- [0498] 221. 실시양태 212에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 조성물.
- [0499] 222. 실시양태 212에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 조성물.
- [0500] 223. 실시양태 212에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 조성물.
- [0501] 224. 실시양태 223에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 조성물.
- [0502] 225. 실시양태 224에 있어서, RGN이 RGN 닉카제인 조성물.
- [0503] 226. 융합 단백질을 코딩하는 핵산 분자 및 우라실 안정화 폴리펩티드 (USP)를 코딩하는 핵산 분자를 포함하는 벡터로서, 여기서 상기 융합 단백질은 DNA-결합 폴리펩티드 및 데아미나제를 포함하고, 상기 USP는 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 것인 벡터.
- [0504] 227. 실시양태 226에 있어서, 상기 USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 것인 벡터.
- [0505] 228. 실시양태 226에 있어서, 상기 USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 것인 벡터.
- [0506] 229. 실시양태 226에 있어서, 상기 USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 것인 벡터.

- [0507] 230. 실시양태 226에 있어서, 상기 USP가 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 것인 벡터.
- [0508] 231. 실시양태 226에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 벡터.
- [0509] 232. 실시양태 226에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 벡터.
- [0510] 233. 실시양태 226에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 벡터.
- [0511] 234. 실시양태 226에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 벡터.
- [0512] 235. 실시양태 226에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 벡터.
- [0513] 236. 실시양태 226에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 벡터.
- [0514] 237. 실시양태 226에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 벡터.
- [0515] 238. 실시양태 237에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 벡터.
- [0516] 239. 실시양태 238에 있어서, RGN이 RGN 닉카제인 벡터.
- [0517] 240. 실시양태 226-239 중 어느 하나의 벡터를 포함하는 세포.
- [0518] 241. a) (i) DNA-결합 폴리펩티드; 및 (ii) 데아미나제를 포함하는 융합 단백질; 또는 융합 단백질을 코딩하는 핵산 분자; 및
- [0519] b) 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP); 또는 USP를 코딩하는 핵산 분자
- [0520] 를 포함하는 세포.
- [0521] 242. 실시양태 241에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 것인 세포.
- [0522] 243. 실시양태 241에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 것인 세포.
- [0523] 244. 실시양태 241에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 것인 세포.
- [0524] 245. 실시양태 241에 있어서, USP가 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는 것인 세포.
- [0525] 246. 실시양태 241에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 80% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 세포.
- [0526] 247. 실시양태 241에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 85% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 세포.
- [0527] 248. 실시양태 241에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 90% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 세포.
- [0528] 249. 실시양태 241에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 적어도 95% 서열 동일성을 갖는 우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 세포.
- [0529] 250. 실시양태 241에 있어서, 융합 단백질이 서열식별번호: 1-16 중 어느 하나에 대해 100% 서열 동일성을 갖는

우라실 안정화 폴리펩티드 (USP)를 추가로 포함하는 것인 세포.

- [0530] 251. 실시양태 241에 있어서, DNA-결합 폴리펩티드가 메가뉴클레아제, 아연 핑거 융합 단백질 또는 TALEN인 세포.
- [0531] 252. 실시양태 241에 있어서, DNA-결합 폴리펩티드가 RNA-가이드된, DNA-결합 폴리펩티드인 세포.
- [0532] 253. 실시양태 252에 있어서, RNA-가이드된, DNA-결합 폴리펩티드가 RNA-가이드된 뉴클레아제 폴리펩티드 (RGN)인 세포.
- [0533] 254. 실시양태 253에 있어서, RGN이 RGN 니카제인 세포.
- [0534] 하기 실시예는 예시로서 제한 없이 제공된다.
- [0535] <실험>
- [0536] 실시예 1: 우라실 안정화 폴리펩티드
- [0537] 본 발명의 우라실 안정화 폴리펩티드 (USP)에 대한 아미노산 서열은 표 1에 나타난 바와 같이 서열식별번호: 1-16으로서 제공된다. 개시된 모든 USP는 스타필로코쿠스 종으로부터의 것이고, 112 내지 116개 아미노산 범위의 길이이다. 본원에 기재된 USP의 폴리펩티드는 5.26의 예상 pI를 갖는 APG05963을 제외하고는, 3.86-4.55의 예상 pI로 16.4% 내지 20.7% 음전하를 보유한다.
- [0538] USP APG06351, APG03399, APG04638, APG09242, APG02463, APG04080, APG01791, APG04001 및 APG03327은 "KEGGNDHE"의 고유한 컨센서스 C-말단 서열 (서열식별번호: 33)을 공유한다. USP APG05198 및 APG05756은 "EKENYNNE"의 독특한 C-말단 서열 (서열식별번호: 34)을 공유한다. APG05963은 "EKEKHKNE" (서열식별번호: 35)의 고유한 C-말단 서열을 갖고; APG06702는 "DKGDDNHD" (서열식별번호: 36)의 고유한 C-말단 서열을 갖고; APG05316은 "QKGGQ" (서열식별번호: 37)의 고유한 C-말단 서열을 갖고; APG09230은 "KGENKYE" (서열식별번호: 38)의 고유한 C-말단 서열을 갖고; APG04100은 "KQGENNHE" (서열식별번호: 39)의 고유한 C-말단 서열을 갖는다.
- [0539] 표 1: 우라실 안정화 폴리펩티드

USP ID	서열 식별 번호
APG03399	1
APG06702	2
APG05198	3
APG04638	4
APG06351	5
APG05963	6
APG09230	7
APG04100	8
APG04001	9
APG02463	10
APG09242	11
APG04080	12
APG05316	13
APG03327	14
APG01791	15
APG05756	16

- [0540]
- [0541] 실시예 2: USP 융합 단백질은 포유동물 세포에서 증가된 염기 편집 활성을 나타냄
- [0542] RGN APG07433.1의 RuvC 도메인을 불활성화시킬 것으로 예측된 잔기 (서열식별번호: 40; PCT 공개 WO 2019/236566, 본원에 참조로 포함됨)를 확인하고, RGN을 니카제 변이체 (nAPG07433.1; 서열식별번호: 41)로 변형시켰다. 시티딘 데아미나제, 즉 APG09980 (서열식별번호: 47; PCT/US2019/068079 참조, 본원에 참조로 포함됨) 또는 APG07386CTD (서열식별번호: 48; PCT/US2019/068079 참조)를 포함하는 융합 단백질을 생성하였다. 표 1의 USP 중 3종, 즉 APG03399, APG06702 및 APG05198을 융합 단백질에서의 활성에 대해 검정하기 위해 선택하였다. 발현에 대해 코돈 최적화된 데아미나제, USP 및 nRGN 뉴클레오티드 서열을 N-말단 핵 국제화 태그를

갖는 융합 단백질로서 합성하고, p트위스트 CMV (트위스트 바이오사이언시스) 발현 플라스미드 내로 클로닝하였다. 본 발명의 USP가 결여된 융합 단백질은 아미노 말단에서 시작하여 SV40 NLS (서열식별번호: 42), 그의 C-말단 단부에 작동가능하게 연결된 3X FLAG 태그 (서열식별번호: 43), 그의 C-말단 단부에 작동가능하게 연결된 테아미나제, 그의 C-말단 단부에 작동가능하게 연결된 펩티드 링커 (서열식별번호: 44), 그의 C-말단 단부에 작동가능하게 연결된 nRGN (예를 들어, nAPG07433.1, 서열식별번호: 41), 최종적으로 그의 C-말단 단부에 작동가능하게 연결된 뉴클레오폴라스민 NLS (서열식별번호: 45)을 포함한다. 본 발명의 USP를 포함하는 융합 단백질은 아미노 말단에서 시작하여 SV40 NLS (서열식별번호: 42), 그의 C-말단 단부에 작동가능하게 연결된 3X FLAG 태그 (서열식별번호: 43), 그의 C-말단 단부에 작동가능하게 연결된 테아미나제, 그의 C-말단 단부에 작동가능하게 연결된 펩티드 링커 (서열식별번호: 44), 그의 C-말단 단부에 작동가능하게 연결된 nRGN (예를 들어, 서열식별번호: 41), 그의 C-말단 단부에 작동가능하게 연결된 제2 링커 서열 (서열식별번호: 46), 그의 C-말단 단부에 작동가능하게 연결된 본 발명의 USP, 최종적으로 그의 C-말단 단부에 작동가능하게 연결된 뉴클레오폴라스민 NLS (서열식별번호: 45)를 포함한다. 표 2는 생성되어 활성화에 대해 시험된 융합 단백질을 나타낸다. 모든 융합 단백질은 상기 기재된 바와 같이 적어도 1개의 NLS 및 3X FLAG 태그를 포함한다.

[0543] 표 2: C>T 편집에 대해 검정된 융합 단백질

융합 단백질	서열 식별 번호
APG09980-nAPG07433.1	49
APG09980-nAPG07433.1-APG03399	50
APG09980-nAPG07433.1-APG06702	51
APG09980-nAPG07433.1-APG05198	52
APG07386-CTD-nAPG07433.1	53
APG07386-CTD-nAPG07433.1-APG03399	54
APG07386-CTD-nAPG07433.1-APG06702	55
APG07386-CTD-nAPG07433.1-APG05198	56

[0544]

[0545] sgRNA를 코딩하는 발현 카세트를 포함하는 발현 플라스미드를 또한 생성하였다. 인간 게놈 표적 서열 및 융합 단백질을 상기 게놈 표적으로 가이드하기 위한 sgRNA 서열을 표 3에 나타낸다. 각각의 표적 서열에 대한 게놈 유전자좌가 또한 표시된다.

[0546] 표 3: 가이드 RNA 서열

sgRNA ID	표적 서열	sgRNA 서열	게놈 유전자좌
SGN000143	57	63	69
SGN000169	58	64	70
SGN000173	59	65	71
SGN000929	60	66	72
SGN000930	61	67	73
SGN001101	62	68	74

[0547]

[0548] 표 2에 나타낸 융합 단백질에 대한 코딩 서열을 포함하는 발현 카세트를 포함하는 플라스미드 500 ng 및 표 3에 나타낸 sgRNA를 코딩하는 발현 카세트를 포함하는 플라스미드 500 ng을 리포펙타민 2000시약 (라이프 테크놀로지스)을 사용하여 24-웰 플레이트에서 75-90% 전면성장률로 HEK293FT 세포 내로 공동-형질감염시켰다. 이어서, 세포를 37℃에서 72시간 동안 인큐베이션하였다. 인큐베이션 후, 게놈 DNA를 이어서 제조업체의 프로토콜에 따라 뉴클레오스핀 96 티슈 (마슈레이-나겔)를 사용하여 추출하였다. 표적화된 게놈 부위에 플랭킹된 게놈 영역을 PCR 증폭시키고, 생성물을 제조업체의 프로토콜에 따라 ZR-96 DNA 클린 앤드 컨센트레이터 (자이모 리서치)를 사용하여 정제하였다. 이어서, 정제된 PCR 생성물을 차세대 서열분석을 위해 일루미나 MiSeq (2x250) 상에 보냈다. 결과를 +30개 뉴클레오티드까지의 indel 형성 또는 특이적 시토신 돌연변이에 대해 분석하였고, 여기서 표 3에 기재된 표적 서열의 3' 말단의 마지막 뉴클레오티드는 +1이고, +30 뉴클레오티드는 서열식별번호: 57-62에 제시된 표적 서열에서 +1 뉴클레오티드로부터 29개 뉴클레오티드 상류 또는 5'이다 (서열식별번호: 57-62의 표적 서열은 각각 서열식별번호: 69-74에 제시된 게놈 유전자좌 서열 내의 소문자로 표시됨).

[0549] 표 4 내지 15는 표 2로부터의 융합 단백질 및 표 3으로부터의 가이드 RNA의 각각의 조합에 대한 시티딘 염기 편

집을 나타낸다. 융합 단백질은 그의 서열식별번호에 의해 확인된다. 표 4-15에서 시티딘 (C)의 넘버링은 이전 단락에서와 같고, 여기서 표적 서열의 3' 말단에서의 마지막 뉴클레오티드는 +1이고, 넘버링은 표적 서열 및 상응하는 게놈 유전자와 서열에서 3'에서 5' 방향으로 진행된다. 흥미롭게도, 동일한 가이드 RNA를 사용하여 데아미나제-nRGN 융합 단백질의 활성을 상응하는 데아미나제-nRGN + USP의 활성과 비교할 때, 시티딘의 목적하는 티미딘으로의 전환은 더 높고, 아데노신 또는 구아노신으로의 전환은 더 적다.

표 4: 데아미나제 APG09980 및 가이드 SGN000143을 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C10	C15	C20	C21	C22	C25
49	A	2.8	0.1	0	0.2	0.6	0.1
	G	1.1	0.1	0.1	0.1	0.2	0
	T	11	0.4	0.2	0.3	2	0.7
50	A	0	0	0	0	0	0
	G	0	0	0	0	0	0
	T	23.5	1.4	1.7	2	4.6	0.7
51	A	0.1	0	0	0	0	0
	G	0	0	0	0	0	0
	T	16.3	1.2	1.5	1.6	3.1	0.6
52	A	0.2	0	0	0	0	0
	G	0	0	0	0	0	0
	T	21.4	1.6	1.1	1.4	2.5	0.7

표 4의 결과는 위치 C10에서의 C>T 편집률이 USP를 갖는 샘플에서 증가하였음을 나타낸다.

표 5: 데아미나제 APG07386 및 가이드 SGN000143을 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C10	C15	C20	C21	C22	C25
53	A	0.1	0.4	0	0.5	1	0
	G	0	0.8	0	0	0.1	0
	T	0.5	4	0.3	0.7	3.8	0
54	A	0	0	0	0	0.1	0
	G	0	0	0	0	0	0
	T	1	4.5	1	1.7	3.1	0.1
55	A	0	0.1	0	0	0.1	0
	G	0	0	0	0	0	0
	T	1.6	9.7	2.5	3.2	5.8	0.2
56	A	0	0.1	0	0	0	0
	G	0	0	0	0	0	0
	T	1.9	10.2	2.6	3.1	7	0.2

표 5의 결과는 위치 C15에서의 C>T 편집률이 USP를 갖는 샘플에서 증가하였음을 나타낸다.

[0556] 표 6: 데아미나제 APG09980 및 가이드 SGN000169를 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C9	C13	C15	C18	C20	C23
49	A	0.5	0.8	3.6	0.3	1.3	1.3
	G	1.5	1.8	8.3	0	0.2	0.1
	T	1.5	23.9	16.3	0.3	8.6	4.9
50	A	0	0	0.6	0	0.1	0.2
	G	0.2	0	2.5	0	0	0
	T	7	38.1	35.8	0.5	16.6	13.4
51	A	0.2	0.3	1.3	0.2	0.3	0.4
	G	0.6	0.6	2.1	0.1	0.1	0
	T	9.2	42.4	40.8	0.8	20.4	16.6
52	A	0.8	1	2.4	0.7	0.6	1.5
	G	0.5	1.7	3.7	0.5	0.5	0.2
	T	8.5	37.8	35.7	1.2	17.1	14.9

[0557]

[0558] 표 6의 결과는 위치 C13 및 C15에서의 C>T 형성률이 USP의 첨가에 따라 증가하였음을 나타낸다.

[0559] 표 7: 데아미나제 APG07386 및 가이드 SGN000169를 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C9	C13	C15	C18	C20	C23
53	A	0.8	0.7	2	0	0.5	0.3
	G	5.4	0.8	10.2	0.1	0.1	0
	T	3.5	5.1	11.2	0.3	2.6	2.4
54	A	0.1	0.5	0.8	0	0.3	0.2
	G	0.7	0.2	2.5	0.2	0.1	0
	T	8.5	15.9	24.8	1.5	6.4	7.1
55	A	0.3	0.5	1.3	0.4	0.2	0.4
	G	2	0.5	3.5	0.1	0.2	0.1
	T	14.3	22.7	37.1	2	9.2	7.9
56	A	0.1	0.1	0.7	0	0.2	0
	G	0.7	0	2.9	0	0	0
	T	14.7	24	36.4	1.3	8.4	8.9

[0560]

[0561] 표 7의 결과는 위치 C13 및 C15에서의 C>T 형성률이 USP의 첨가에 따라 증가하였음을 나타낸다.

[0562] 표 8: 데아미나제 APG09980 및 가이드 SGN000930을 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C17	C19	C22
49	A	0.7	0.4	1.4
	G	34.5	0.4	0
	T	2.6	2.4	2.2
50	A	0.3	0.1	0.3
	G	8.6	0.3	0
	T	33	2.8	3.2
51	A	0.2	0.1	0.2
	G	5.1	0.1	0.1
	T	35.7	3.6	3.3
52	A	0.4	0.3	0.9
	G	11.2	0	0
	T	23.7	2.5	2.5

[0563]

[0564] 표 8의 결과는 위치 C17에서의 C>G 편집률이 C>T 변화에 유리한 USP를 갖는 모든 샘플에서 감소하였음을 나타낸다.

[0565] 표 9: 데아미나제 APG07386 및 가이드 SGN000930을 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C17	C19	C22
53	A	0.2	0	0.4
	G	14.7	0	0.4
	T	1.1	0.2	2.2
54	A	0	0	0
	G	1.4	0	0
	T	9.1	0.4	3.7
55	A	0.4	0.1	0.3
	G	2.6	0.2	0.2
	T	11.8	0.9	4.5
56	A	0.1	0.2	0
	G	2.7	0	0.1
	T	17.4	0.8	6.7

[0566]

[0567] 표 9의 결과는 위치 C17에서의 C>G 편집이 C>T 변화에 유리한 USP를 갖는 모든 샘플에서 감소하였음을 나타낸다.

[0568] 표 10: 데아미나제 APG09980 및 가이드 SGN00173을 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C1	C2	C3	C4	C7	C8	C10	C11	C17	C20
49	A	0	0	0.1	0.3	3.1	6.9	0.3	0	3.4	0.1
	G	0	0.1	0	0.2	0.3	1	0.1	0.4	3.4	0.1
	T	0.1	0	0	2.5	11.1	12.3	1.7	0.2	12.1	1.7
50	A	0	0.1	0	0.1	0.9	2.9	0	0.4	1.5	0
	G	0.1	0.1	0	0	0.5	0.8	0.4	0.3	1.9	0.2
	T	0.1	0.1	0.3	14	39	39.3	28.7	13	31.1	12
51	A	0	0	0	0	1.2	1.6	0	0	0.9	0
	G	0	0	0	0	0	1.4	0.1	0	2.2	0.1
	T	0	0	0.3	10.4	33.5	33.4	23.8	10.9	26.8	10.3
52	A	0	0	0	0	1.8	1.2	0.3	0	1.5	0
	G	0	0	0	0	0	0.8	0	0.5	2.7	0
	T	0	0	0.6	10.5	33.1	36.8	26.7	13.2	28.6	9.6

[0569]

[0570] 표 10의 결과는 위치 C4, C7, C8, C10, C11, C17 및 C20에서의 C>T 형성률이 USP의 첨가에 따라 증가하였음을 나타낸다.

[0571] 표 11: 데아미나제 APG07386 및 가이드 SGN00173을 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C1	C2	C3	C4	C7	C8	C10	C11	C17	C20
53	A	0	0	0	0.1	0.4	1.1	0	2.5	2.8	0
	G	0	0	0	0	0.1	0.3	0	0.5	2	0
	T	0	0	0	0.5	1.4	2.5	0	7	11.3	0.7
54	A	0	0	0	0	0	0.1	0	0.7	0.8	0
	G	0	0	0	0	0	0	0	0	0.6	0
	T	0	0	0	1.6	6.1	8.6	2.2	13.3	23	1.8
55	A	0	0	0	0	0.2	0	0	1	0.9	0
	G	0	0	0	0	0	0.3	0.1	0.1	0.5	0
	T	0	0	0.2	2.2	9.1	11.2	3.2	13	21.3	2
56	A	0	0	0	0	0	0	0	0.9	0.8	0
	G	0	0	0	0	0	0.1	0	0.2	0.5	0
	T	0	0	0.1	2.7	12.3	15	4.3	18.3	27.1	1.9

[0572]

[0573] 표 11의 결과는 위치 C7, C8, C11 및 C17에서의 C>T 형성률이 USP의 첨가에 따라 증가하였음을 나타낸다.

[0574] 표 12: 데아미나제 APG09980 및 가이드 SGN000929를 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C6	C9	C13	C23
49	A	0.2	1.1	0.2	5.2
	G	0.2	1.5	1.4	20.1
	T	1.2	5.2	1	7.4
50	A	0	0.1	0	1.3
	G	0	0.1	0	3.3
	T	2.2	9.4	3.2	30.8
51	A	0	0.2	0	0.8
	G	0	0.1	0	1.6
	T	1.5	9.2	3.5	34
52	A	0.5	0.6	0	1.8
	G	0.1	0.5	0	6.5
	T	1.6	8.6	2.7	22.3

[0575]

[0576] 표 12의 결과는 위치 C23에서의 C>T 형성률이 USP의 첨가에 따라 증가하였음을 나타낸다.

[0577] 표 13: 데아미나제 APG07386 및 가이드 SGN000929를 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C6	C9	C13	C23
53	A	1.5	0	0	1.5
	G	0.6	0	0.1	5.2
	T	5	0.3	0.1	2.1
54	A	0	0	0	0.1
	G	0	0	0	0.1
	T	4.2	0.3	0	4.6
55	A	0.2	0	0.1	0
	G	0	0	0	1.4
	T	10.4	0.9	0.3	8.2
56	A	0.1	0	0	0.2
	G	0.2	0	0	0.5
	T	10.8	0.6	0.4	8.5

[0578]

[0579] 표 13의 결과는 위치 C6 및 C23에서의 C>T 형성률이 USP의 첨가에 따라 증가하였음을 나타낸다.

[0580] 표 14: 데아미나제 APG09980 및 가이드 SGN001101을 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C2	C6	C7	C18
49	A	0.1	1.8	0.4	3.3
	G	0	2.8	0.1	18.5
	T	0.1	6.5	1.8	5.3
50	A	0.2	0.3	0.2	0.6
	G	0.1	0.1	0.2	2.1
	T	0.1	9.1	0.7	25.4
51	A	0.2	0.1	0.1	0.3
	G	0.1	0.1	0	2.4
	T	0.1	7.9	0.7	21.8
52	A	0.2	0.9	0.2	1.2
	G	0	0.8	0	4.8
	T	0.2	7.7	1.3	19.4

[0581]

[0582] 표 14의 결과는 위치 C18에서의 C>T 형성률이 USP의 첨가로 증가하였음을 나타낸다.

[0583] 표 15: 데아미나제 APG07386 및 가이드 SGN001101을 사용한 C>N 편집률

융합 단백질 (서열 식별 번호)		C2	C6	C7	C18
53	A	0.1	0	4	2.2
	G	0	0	2.1	11
	T	0	0	10.4	3.9
54	A	0.1	0	0.1	0
	G	0	0	0.1	0.6
	T	0	1.7	8.5	6.7
55	A	0.2	0.1	0.6	0.3
	G	0	0	0.5	1.4
	T	0	1.7	10.7	9.2
56	A	0.1	0	0.3	0.2
	G	0	0	0.4	1.4
	T	0	1.9	12.1	9.8

[0584]

[0585] 표 15의 결과는 위치 C18에서의 C>T 형성률이 USP의 첨가로 증가하였음을 나타낸다. C>G 전환율은 USP의 첨가로 위치 C18에서 감소하였다.

[0586] 표 16 및 17은 시험된 각각의 융합 단백질/가이드 조합에 대한 indel 형성률을 나타낸다. 융합 단백질은 서열 식별번호로 나타낸다. 데이터는 본원에 기재된 USP를 포함하는 융합 단백질이 시험된 모든 표적 게놈 위치에서 indel 형성률을 감소시켰음을 나타낸다.

[0587] 표 16: APG09980 및 USP를 사용한 삽입률 및 결실률

sgRNA ID	서열 식별 번호: 49	서열 식별 번호: 50	서열 식별 번호: 51	서열 식별 번호: 52
SGN000143	2.78	0	0	0.62
SGN000169	18.02	1.62	3.89	9.1
SGN000173	27.46	7.86	7.17	7.48
SGN000929	3.49	0.2	0.65	2.36
SGN000930	6.03	0.81	0.65	1.65
SGN001101	2.5	1.66	0.92	1.86

[0588]

[0589] 표 17: APG07386-CTD 및 USP를 사용한 삽입률 및 결실률

sgRNA ID	서열 식별 번호: 53	서열 식별 번호: 54	서열 식별 번호: 55	서열 식별 번호: 56
SGN000143	0.13	0.05	0	0.06
SGN000169	5.6	1.38	3.11	0.32
SGN000173	13.8	0.99	3.49	2.57
SGN000929	0.5	0	0.3	0.16
SGN000930	1.7	0	0.23	0.31
SGN001101	1.78	0	0.48	0.39

[0590]

[0591] 실시예 3: 상이한 전달 포맷의 시험

[0592] 염기 편집제가 상이한 포맷으로 전달할 수 있는지를 결정하기 위해, mRNA 전달을 1차 T-세포로 시험하였다. 정제된 CD3+ T-세포 또는 PBMC를 해동시키고, CD3/CD28 비드 (썬피셔)를 사용하여 3일 동안 활성화시킨 후, 론자 4D-뉴클레오펙터 X 유닛 및 뉴클레오펙터 스트림을 사용하여 뉴클레오펙션시켰다. P3 1차 세포 키트를 mRNA 및 RNP 전달 둘 다에 사용하였다. 세포를 각각 mRNA 및 RNP 전달을 위한 EO-115 및 EH-115 프로그램을 사용하여 형질감염시켰다. 뉴클레오펙션 후 4일 동안 IL-2, IL-7, 및 IL-15 (밀테니 바이오텍)를 함유하는 CTS 옵티마이저 T 세포 확장 배지 (썬피셔)에서 세포를 배양한 후, 뉴클레오스핀 조직 계층 DNA 단리 키트 (마슈레이나겔)를 사용하여 수거하였다.

[0593] 편집 부위를 둘러싸는 앰플리콘을 PCR에 의해 생성하고, 2x250bp 쌍형성된 말단 서열분석을 사용하는 일루미나 넥스테라 플랫폼을 사용하여 NGS 서열분석에 적용하였다. 각각의 샘플에 대한 전체 치환율을 계산함으로써 추정된 염기 편집률을 결정하였다. 시험된 각각의 가이드에 대한 샘플의 평균 및 수를 하기 표 18 및 19에 나타낸다.

[0594] APG09980-nAPG07433.1-APG03399 및 APG05840-nAPG07433.1-APG03399는 mRNA에 의해 전달되는 경우에 여러 표적으로서 높은 염기 편집률을 나타낸다. 염기 편집 구축물에서의 USP2의 혼입으로 인해, 높은 치환율에도 불구하고 매우 낮은 삽입-결실 형성율이 존재한다.

[0595] 표 18: APG09980-nAPG07433.1-APG03399에 대한 평균 염기 편집률

융합 구축물	유전자 명칭	SGN	평균 % 치환	평균 % Indel	N
APG09980-nAPG07433.1-APG03399	유전자 1	SGN000754	23.32917436	1.128931	6
APG09980-nAPG07433.1-APG03399	유전자 1	SGN000755	59.37254849	7.1037823	4
APG09980-nAPG07433.1-APG03399	유전자 2	SGN001061	13.60100568	0.4214674	3
APG09980-nAPG07433.1-APG03399	유전자 2	SGN001062	26.9304354	4.3225871	4
APG09980-nAPG07433.1-APG03399	유전자 2	SGN001063	75.27761104	0.8163273	4
APG09980-nAPG07433.1-APG03399	유전자 2	SGN001064	72.94658862	1.0468487	3

[0596]

[0597] 표 19: APG05840-nAPG07433.1-APG03399에 대한 평균 염기 편집률

융합 구축물	유전자 명칭	SGN	평균 % 치환	평균 % Indel	N
APG05840-nAPG07433.1-APG03399	유전자 1	SGN000754	57.7775198	5.14624384	4
APG05840-nAPG07433.1-APG03399	유전자 1	SGN000755	68.352455	4.98538891	3
APG05840-nAPG07433.1-APG03399	유전자 2	SGN001061	14.6830209	0	2
APG05840-nAPG07433.1-APG03399	유전자 2	SGN001062	39.7312597	2.9887885	4
APG05840-nAPG07433.1-APG03399	유전자 2	SGN001063	70.4564399	0.25727852	4
APG05840-nAPG07433.1-APG03399	유전자 2	SGN001064	53.2112842	1.98008536	3

[0598]

서열 목록

SEQUENCE LISTING

<110> LifeEDIT, LLC

<120> URACIL STABILIZING PROTEINS AND ACTIVE FRAGMENTS AND VARIANTS
THEREOF AND METHODS OF USE

<130> L103438 1220WO (0106.5)

<140><141><150> 63/052,175

<151> 2020-07-15

<160> 142

<170> PatentIn version 3.5

<210> 1

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG03399

<400> 1

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe

1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala

20 25 30

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

35 40 45

His Leu Thr Leu Asn Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu

50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asp Gln Leu Ile Ser
 65 70 75 80
 Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp

85 90 95
 Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Gln Gly Glu
 100 105 110

Asn Asn His Glu
 115

<210> 2

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG06702

<400> 2

Met Thr Ile Glu Asn Glu Leu Lys Glu Tyr Ile Thr Asp Leu Phe Glu
 1 5 10 15
 Leu Pro Lys Asp Glu Gln Trp Leu Cys Glu Ala Val Asp Glu Val Ala

20 25 30
 Asp Asn Ile Leu Pro Asp Glu Tyr Val Thr Leu Gly Pro Leu Arg Asn
 35 40 45

Lys Thr Leu Glu Thr Phe Thr Tyr Tyr Ser Asn Thr Leu His Glu Arg
 50 55 60
 Asn Ile Tyr Pro Phe Ile Leu Tyr Tyr Gln Asn Arg Leu Ile Ala Ile

65 70 75 80
 Gly Tyr Ile Asp Glu Asn Tyr Glu Met Asp Phe Met Tyr Leu His Asn

85 90 95
 Thr Ala Ile Thr Leu Ile Asp Gln Arg Tyr Leu Ile Asp Lys Gly Asp
 100 105 110

Asp Asn His Asp
 115

<210> 3

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG05198

<400> 3

Met Thr Leu Glu Lys Gln Leu Lys Glu Tyr Ile Thr Asn Leu Phe Asn

1 5 10 15

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

20 25 30

Asp Asn Ile Leu Pro Asp Gln Tyr Ile Arg Leu Gly Pro Leu Thr Asn

35 40 45

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

50 55 60

His Ile Tyr Pro Phe Ile Leu Tyr Tyr Gln Lys Gln Leu Ile Ala Ile

65 70 75 80

Gly Tyr Ile Asp Glu Thr Asn Asp Met Asp Phe Leu Tyr Leu His Asn

85 90 95

Thr Val Met Pro Leu Leu Asp Gln Arg His Leu Leu Glu Lys Glu Asn

100 105 110

Tyr Asn Asn Glu

115

<210> 4

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG04638

<400> 4

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe

1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala

20 25 30

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

35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asn Glu Leu His Asp
50 55 60
Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80
Val Gly Tyr Leu Asn His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95
Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Arg Glu Gly Gly
100 105 110
Thr His His Glu
115
<210> 5
<211> 116
<212> PRT
<213> Staphylococcus spp
<223> APG06351
<400> 5
Met Lys Thr Ile Thr Gln Glu Leu Lys Gln Tyr Ile Thr His Leu Phe
1 5 10 15
Gln Leu Ser Asn Asn Glu Ala Trp Glu Cys Glu Ala Leu Glu Glu Ala
20 25 30
Ala Glu Asn Ile Leu Pro Glu Arg Phe Ile Asn Asp Ser Pro Leu Val
35 40 45
His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
50 55 60
Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80
Val Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95
Thr Gln Asn Ile Ile Ile Asp Glu Arg His Leu Leu Lys Gln Gly Gly
100 105 110
Gln Asp His Glu
115

<210> 6

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG05963

<400> 6

Met Thr Leu Glu Gln Gln Leu Lys His Tyr Ile Thr Asn Leu Phe Lys

1 5 10 15

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

20 25 30

Asp Asn Ile Leu Pro Asp Gln Tyr Val Arg Leu Gly Pro Leu Thr Asn

35 40 45

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

50 55 60

His Ile Tyr Pro Phe Ile Leu Tyr Tyr Gln Lys Gln Leu Ile Ala Ile

65 70 75 80

Gly Tyr Ile Asp Glu Thr Asn Asp Met Asp Phe Leu Tyr Leu His Asn

85 90 95

Thr Val Met Pro Leu Leu Asp Gln Arg His Leu Leu Glu Lys Glu Lys

100 105 110

His Lys Asn Glu

115

<210> 7

<211> 115

<212> PRT

<213> Staphylococcus spp

<223> APG09230

<400> 7

Met Lys Thr Gln Ile Asn Glu Leu Lys Ser Tyr Ile Thr Asn Leu Phe

1 5 10 15

Gln Leu Thr Ser Gln Glu Asn Trp Glu Cys Glu Ala Leu Asn Glu Ser

20 25 30

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

85 90 95
 Thr Gln Thr Thr Ile Ile Asp Glu Arg His Leu Leu Lys Gly Glu Asn
 100 105 110
 Lys Tyr Glu
 115

<210> 8

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG04100

<400> 8

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
 1 5 10 15
 Gln Leu Ser Asn Glu Glu Ser Trp Glu Cys Glu Val Leu Asp Glu Val

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

85 90 95
 Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Gln Gly Glu
 100 105 110

Asn Asn His Glu

115

<210> 9

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG04001

<400> 9

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe

1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala

20 25 30

Ala Glu Asn Ile Leu Pro Thr Arg Phe Val Asp His Thr Pro Leu Ala

35 40 45

His Leu Thr Leu Asp Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu

50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser

65 70 75 80

Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp

85 90 95

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

100 105 110

Asn Asp Arg Glu

115

<210> 10

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG02463

<400> 10

Met Gln Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe

1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala

20 25 30

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

35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu

50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser

65 70 75 80

Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp

85 90 95

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

100 105 110

Asn Asn His Glu

115

<210> 11

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG09242

<400> 11

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe

1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Asp Thr

20 25 30

Ala Glu Asn Ile Leu Pro Glu Cys Phe Ile Asn Asp Thr Pro Leu Ala

35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu

50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser

65 70 75 80

Val Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp

85 90 95
 Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Glu Gly Gly
 100 105 110

Asn Asp His Glu

115

<210> 12

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG04080

<400> 12

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
 1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala

20 25 30
 Ala Ala Asn Ile Leu Pro Glu Arg Phe Ile Asn Asn Ser Pro Leu Ala

35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu

50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser

65 70 75 80

Val Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp

85 90 95
 Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Glu Gly Gly
 100 105 110

Asn Asn His Glu

115

<210> 13

<211> 112

<212> PRT

<213> Staphylococcus spp

<223> APG05316

<400> 13

Met Thr Leu Ser Lys Gln Leu Lys Thr Tyr Ile Thr Glu Arg Phe Lys
1 5 10 15
Leu Asn Tyr Gln Glu Thr Trp Ala Cys Glu Thr Ile Asp Ala Val Ala
20 25 30
Glu Asp Val Leu Pro Glu Lys Tyr Ile Lys Asn Ser Pro Leu Glu His
35 40 45
Lys Ile Leu Asn Thr Phe Thr Tyr Tyr Asn Asp Glu Leu His Glu Ile
50 55 60
Ser Ile Tyr Pro Phe Leu Cys Tyr Leu Asp Lys Glu Leu Val Ala Ile
65 70 75 80
Gly Tyr Leu Asp Asn Phe Asp Leu Asp Phe Ile Phe Leu Asn Asp Thr
85 90 95
His Gln Val Ile Ile Asp Glu Arg Tyr Leu Leu Gln Lys Gly Gly Gln
100 105 110
<210> 14
<211> 113
<212> PRT
<213> Staphylococcus spp
<223> APG03327
<400> 14
Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15
Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Asp Ala
20 25 30
Ala Glu Asn Ile Leu Pro Thr Arg Phe Val Asp His Thr Pro Leu Ala
35 40 45
His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
50 55 60
Leu Ser Ile Tyr Pro Phe Leu Ile Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80
Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp

85 90 95

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

100 105 110

Glu

<210> 15

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG01791

<400> 15

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe

1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala

20 25 30

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

35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu

50 55 60

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

65 70 75 80

Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp

85 90 95

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

100 105 110

Asn Asp His Glu

115

<210> 16

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG05756

<400> 16

Met Thr Leu Glu Gln Gln Leu Lys His Tyr Ile Thr Asn Leu Phe Ser
 1 5 10 15
 Leu Pro Gln Asp Glu Lys Trp Glu Cys Glu Ser Ile Glu Glu Val Ala
 20 25 30
 Asp Asn Ile Leu Pro Asp Gln Tyr Val Arg Leu Gly Pro Leu Thr Asn
 35 40 45
 Lys Ile Leu His Thr Tyr Thr Tyr Tyr Ser Asp Lys Leu His Glu Arg
 50 55 60
 His Ile Tyr Pro Phe Ile Leu Tyr Tyr Gln Lys Gln Leu Ile Ala Ile
 65 70 75 80
 Gly Tyr Ile Asp Glu Ile Asn Asp Met Asp Phe Leu Tyr Leu His Asn
 85 90 95
 Thr Val Met Pro Leu Leu Asp Gln Arg His Leu Leu Glu Lys Glu Asn

100 105 110
 Tyr Asn Asn Glu
 115

<210> 17

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> APG03399 mammalian codon optimized sequence

<400> 17

atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
 aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccagagagg 120
 ttcataca acagcctgct ggcccacttg accctgaaca cctacaccta ctacaacgac 180
 gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacgacca gctgatcagc 240
 atcggtacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
 atcatcgacg agaggcacct gctgaagcag ggcgagaaca accacgag 348

<210> 18

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> APG06702 mammalian codon optimized sequence

<400> 18

atgaccatcg agaacgagct gaaggagtac atcaccgacc tgttcgagct gccaaggac	60
gagcagtggc tgtgcgaggc cgtggacgag gtggccgaca acatcctgcc cgacgagtac	120
gtgaccctgg gccccctgag gaacaagacc ctggagacct tcacctacta cagcaacacc	180

ctgcacgaga ggaacatcta ccccttcac ctgtactacc agaacaggct gatcgccatc	240
ggctacatcg acgagaacta cgagatggac ttcattgtacc tgcacaacac cgccatcacc	300
ctgatcgacc agaggtacct gatcgacaag ggcgacgaca accacgac	348

<210> 19

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> APG05198 mammalian codon optimized sequence

<400> 19

atgaccctgg agaagcagct gaaggagtac atcaccaacc tgttcaacct gccaaggac	60
gagaagtggg agtgcgagag catcgaggag gtggccgaca acatcctgcc cgaccagtac	120

atcaggctgg gccccctgac caacaagatc ctgcacacct acacctacta cagcgacacc	180
ctgcacgaga ggcacatcta ccccttcac ctgtactacc agaagcagct gatcgccatc	240
ggctacatcg acgagaccaa cgacatggac ttctgtacc tgcacaacac cgtgatgccc	300
ctgctggacc agaggcacct gctggagaag gagaactaca acaacgag	348

<210> 20

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> APG04638 mammalian codon optimized sequence

<400> 20

atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac	60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccagagg	120
ttcatcaacg acagccccct ggcccactg accctggaga cctacaccta ctacaacaac	180
gagctgcacg acctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc	240
gtgggtacc tgaaccactt cgacatggac ttctgtacc tgaccgacac caagaacacc	300

atcatcgacg agaggcacct gctgagggag ggcggcacc accacgag 348

<210> 21

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> APG06351 mammalian codon optimized sequence

<400>

> 21

atgaagacca tcaccagga gctgaagcag tacatcacc acctgttcca gctgagcaac 60

aacgaggcct gggagtgcga ggccctggag gagggcccg agaacatcct gcccagagg 120

ttcatcaacg acagcccct ggtgcacctg accctggaga cctacaccta ctacaacgac 180

gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgacgac 240

gtgggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac ccagaacatc 300

atcatcgacg agaggcacct gctgaagcag ggcggccagg accacgag 348

<210> 22

<211> 348

<212> DNA

<213>

Artificial Sequence

<220><223> APG05963 mammalian codon optimized sequence

<400> 22

atgaccctgg agcagcagct gaagcactac atcaccaacc tgttcaagct gcccaggac 60

gagaagtggg agtgcgagag catcgaggag gtggccgaca acatcctgcc cgaccagtac 120

gtgaggctgg gcccctgac caacaagatc ctgcacacct acacctacta cagcgacacc 180

ctgcacgaga ggcacatcta ccccttcac ctgtactacc agaagcagct gatcgccatc 240

ggctacatcg acgagaccaa cgacatggac ttcctgtacc tgcacaacac cgtgatgccc 300

ctgctggacc agaggcacct gctggagaag gagaagcaca agaacgag 348

<210> 23

<211> 345

<212> DNA

<213> Artificial Sequence

<220><223> APG09230 mammalian codon optimized sequence

<400> 23

atgaagacc agatcaacga gctgaagagc tacatcacca acctgttcca gctgaccagc 60

caggagaact gggagtgcga ggccctgaac gagagcgccg acaacatcct gcccagagg 120
 ttctgaacg acagcccct gagcaacagg atcctggaga cctacaccta ctacaacaac 180
 gagctgcacg agctgagcat ctacccttc ctgatgtact accagaacca gctgatcagc 240
 atcggtacc tggaccacta cgacatggac ttctgtacc tgaacgacac ccagaccacc 300

atcatcgacg agaggcacct gctgaaggcg gagaacaagt acgag 345

<210> 24
 <211> 348
 <212> DNA
 <213> Artificial Sequence
 <220><223> APG04100 mammalian codon optimized sequence
 <400> 24

atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
 gaggagagct gggagtgcga ggtgctggac gaggtggccg agaacatcct gccccccagg 120
 ttctgtgacg gcagcccct gaccacctg accctggaga cctacaccta ctacaacaac 180
 gagctgcacg acctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc 240

gtgggtacc tggaccactt cgacatggac ttctgtacc tgaccgacac caagaacacc 300
 atcatcgacg agaggcacct gctgaagcag ggcgagaaca accacgag 348

<210> 25
 <211> 348
 <212> DNA
 <213> Artificial Sequence
 <220><223> APG04001 mammalian codon optimized sequence
 <400> 25

atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
 aacgagacct gggagtgcga ggccctggag gaggccgccc agaacatcct gcccaccagg 120
 ttctgtgacc acaccccct ggcccacctg accctggaca cctacaccta ctacaacgac 180

gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc 240
 atcggtacc tggaccactt cgacatggac ttctgtacc tgaccgacac caagaacatc 300
 atcatcgacg agaggcacct gctgaggag ggcggcaacg acaggag 348

<210> 26
 <211> 348
 <212> DNA

<213> Artificial Sequence

<220><223> APG02463 mammalian codon optimized sequence

<400> 26

atgcagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60

aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccagagg 120

ttcataacg gcagccccct ggcccacctg accctggaga cctacaccta ctacaacgac 180

gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgacgacg 240

atcggtacc tggaccactt cgacatggac ttctgtacc tgaccgacac caagaacacc 300

atcatcgacg agaggcacct gctgaaggag ggcggcaaca accacgag 348

<210> 27

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> APG09242 mammalian codon optimized sequence

<400> 27

atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60

aacgagacct gggagtgcga ggccctggag gacaccgccg agaacatcct gcccaggtgc 120

ttcataacg acacccccct ggcccacctg accctggaga cctacaccta ctacaacgac 180

gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgacgacg 240

gtgggtacc tggaccactt cgacatggac ttctgtacc tgaccgacac caagaacacc 300

atcatcgacg agaggcacct gctgaaggag ggcggcaacg accacgag 348

<210> 28

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> APG04080 mammalian codon optimized sequence

<400>

> 28

atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60

aacgagacct gggagtgcga ggccctggag gaggccgccg ccaacatcct gcccagagg 120

ttcataacg acagccccct ggcccacctg accctggaga cctacaccta ctacaacgac 180

gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgacgacg 240

gtgggtacc tggaccactt cgacatggac ttctgtacc tgaccgacac caagaacacc 300

atcatcgacg agaggcacct gctgaaggag ggcggaaca accacgag 348

<210> 29

<211> 336

<212> DNA

<213>

Artificial Sequence

<220><223> APG05316 mammalian codon optimized sequence

<400> 29

atgacctga gcaagcagct gaagacctac atcaccgaga ggttcaagct gaactaccag 60

gagacctggg cctgcgagac catcgacgcc gtggccgagg acgtgctgcc cgagaagtac 120

atcaagaaca gccccctgga gcacaagatc ctgaacacct tcacctacta caacgacgag 180

ctgcacgaga tcagcatcta ccccttcttg tgctacctgg acaaggagct ggtggccatc 240

ggctacctgg acaacttcga cctggacttc atcttctga acgacacca ccaggatgac 300

atcgacgaga ggtacctgct gcagaagggc ggccag 336

<210> 30

<211> 339

<212> DNA

<213> Artificial Sequence

<220><223> APG03327 mammalian codon optimized sequence

<400> 30

atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60

aacgagacct gggagtgcga ggccctggag gacgccgccg agaacatcct gccaccagg 120

ttcgtggacc acacccccct ggcccacctg accctggaga cctacaccta ctacaacgac 180

gagctgcacg agctgagcat ctacccttc ctgatctacg ccaacaacca gctgacgac 240

atcggtacc tggaccactt cgacatggac ttctgtacc tgaccgacac caagaacacc 300

atcatcgacg agaggcacct gctgcagaag ggcaggag 339

<210> 31

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> APG01791 mammalian codon optimized sequence

<400> 31

atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60

aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gccaccagg 120
 ttcgtgggcc acacccccct ggcccactg accctggaga cctacaccta ctacaacgac 180
 gagctgcacg agctgagcat ctaccccttc ctgatgtacg ccaacgacca gctgatcagc 240

atcggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
 ctgatcgacg agaggcacct gctgaagcag ggcggaacg accacgag 348

<210> 32

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> APG05756 mammalian codon optimized sequence

<400> 32

atgacctgg agcagcagct gaagcactac atcaccaacc tgttcagcct gcccaggac 60
 gagaagtggg agtgcgagag catcgaggag gtggccgaca acatcctgcc cgaccagtac 120
 gtgaggctgg gcccctgac caacaagatc ctgcacacct acacctacta cagcgacaag 180

ctgcacgaga ggcacatcta ccccttcac cgtactacc agaagcagct gatcgccatc 240
 ggctacatcg acgagatcaa cgacatggac ttcctgtacc tgcacaacac cgtgatgccc 300
 ctgctggacc agaggcacct gctggagaag gagaactaca acaacgag 348

<210> 33

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Shared consensus C-terminus sequence

<400> 33

Lys Glu Gly Gly Asn Asp His Glu

1 5

<210> 34

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> APG05756 and APG05198 unique C-term seq

<400> 34

Glu Lys Glu Asn Tyr Asn Asn Glu

1 5

<210> 35

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> APG05963 unique C-term seq

<400> 35

Glu Lys Glu Lys His Lys Asn Glu

1 5

<210> 36

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> APG06702 unique C-terminal seq

<400> 36

Asp Lys Gly Asp Asp Asn His Asp

1 5

<210> 37

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> APG05316 unique C-terminal seq

<400> 37

Gln Lys Gly Gly Gln

1 5

<210> 38

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> APG09230 unique C-term seq

<400> 38

Lys Gly Glu Asn Lys Tyr Glu

1 5

<210> 39

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> APG04100 unique C-term seq

<400> 39

Lys Gln Gly Glu Asn Asn His Glu

1 5

<210> 40

<211> 1071

<212> PRT

<213> Artificial Sequence

<220><223> RGN APG07433.1

<400> 40

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser

1 5 10 15

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

20 25 30

Tyr Glu Lys Val Arg Ile Val Asp Gln Gly Val Arg Met Phe Asp Arg

35 40 45

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

50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys

65 70 75 80

Asn Ile Arg Asn Leu Leu Val Gln His Gly Val Ile Thr Gln Glu Glu

85 90 95

Leu Asp Ser Leu Tyr Pro Leu Ser Lys Lys Ser Met Asp Ile Trp Gly

100 105 110

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asn His Phe Glu Trp Ala

115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg

130 135 140

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

145 150 155 160
 Gln Leu Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
 165 170 175
 Trp Met Lys Asp Pro Asp Phe Ser Lys Tyr Asp Arg Lys Arg Asn Ser
 180 185 190

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

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

 Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln
 325 330 335
 Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
 340 345 350
 Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Asn Lys Val
 355 360 365
 Val Ala Asn Tyr Ser Glu Arg Thr Asn Glu Thr Phe Ser Thr Leu Asp
 370 375 380

 Tyr Asp Gly Ile Gly Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
 385 390 395 400

Ile Arg Ser Tyr Leu Lys Ser Ser His Asn Leu Pro Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Ser Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Leu Lys Ala Ile Asn His Val Leu Ser Ile Met Gln
435 440 445

Lys Gly Asn Thr Tyr Lys Glu Ala Val Asp Gln Leu Gly Tyr Asp Thr
450 455 460

Ser Gly Leu Lys Lys Glu Lys Arg Ser Lys Phe Leu Pro Pro Ile Ser
465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg His Gly Ser Pro His Ser Val
500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Asn His Asp Glu Arg Thr
515 520 525

Lys Ile Val Ser Ala Gln Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540

Ile Ser Ile Leu Ser Glu His Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575

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

Glu Arg Asn Gly Ala Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
610 615 620

Asp Glu Asn Arg Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
625 630 635 640

Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn

645 650 655
 Lys Phe Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Arg Ala Tyr
 660 665 670
 Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
 675 680 685
 Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
 690 695 700

 Phe Lys Glu Ala Glu Asp Asn Pro Arg Lys Arg Arg Val Gln Thr Val
 705 710 715 720
 Asn Gly Val Ile Thr Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys
 725 730 735
 Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
 740 745 750
 Ala Cys Thr Asp His His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln
 755 760 765

 Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
 770 775 780
 Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
 785 790 795 800
 Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Leu
 805 810 815
 Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Ala
 820 825 830

 Ala His Lys Gln Thr Ile Met Arg Lys Gly Gly Ile Asp Lys Lys Gly
 835 840 845
 Lys Thr Ile Ile Ile Glu Arg Leu His Leu Lys Asp Ile Lys Phe Asp
 850 855 860
 Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Met Ala Thr
 865 870 875 880
 Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Gly Lys Asn Ser Lys
 885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly

900 905 910

Asn Leu Ile Lys Arg Val Lys Val Glu Gly Gln Ala Lys Ser Phe Val

915 920 925

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

930 935 940

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

945 950 955 960

Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Ala Ser Ser

965 970 975

Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys

980 985 990

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

995 1000 1005

Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg

1055 1060 1065

Asn Phe His

1070

<210> 41

<211> 1071

<212> PRT

<213> Artificial Sequence

<220><223> nRGN APG07433.1

<400> 41

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

1 5 10 15

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

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

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

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

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

260	265	270
Gln Tyr Phe Ser Ala Leu Asp Gln Val Asn Arg Thr Arg Leu Gly Pro		
275	280	285
Asp Phe Gln Pro Phe Thr Lys Glu Gln Arg Glu Ile Ile Leu Asn Asn		
290	295	300
Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val		
305	310	315
		320
Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln		
325	330	335
Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu		
340	345	350
Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Asn Lys Val		
355	360	365
Val Ala Asn Tyr Ser Glu Arg Thr Asn Glu Thr Phe Ser Thr Leu Asp		
370	375	380
Tyr Asp Gly Ile Gly Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp		
385	390	395
Ile Arg Ser Tyr Leu Lys Ser Ser His Asn Leu Pro Lys Arg Cys Tyr		
405	410	415
Asp Asp Gln Leu Ile Glu Glu Leu Leu Ser Leu Ser Tyr Thr Lys Phe		
420	425	430
Gly His Leu Ser Leu Lys Ala Ile Asn His Val Leu Ser Ile Met Gln		
435	440	445
Lys Gly Asn Thr Tyr Lys Glu Ala Val Asp Gln Leu Gly Tyr Asp Thr		
450	455	460
Ser Gly Leu Lys Lys Glu Lys Arg Ser Lys Phe Leu Pro Pro Ile Ser		
465	470	475
Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg		
485	490	495
Lys Val Val Asn Ala Ile Ile Arg Arg His Gly Ser Pro His Ser Val		
500	505	510

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

 Ser Leu Lys Glu Ile Pro Ala Asp Thr Phe Phe Asn Glu Leu Lys Lys
 580 585 590
 Glu Arg Asn Gly Ala Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
 595 600 605
 Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
 610 615 620
 Asp Glu Asn Arg Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
 625 630 635 640

 Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
 645 650 655
 Lys Phe Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Arg Ala Tyr
 660 665 670
 Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
 675 680 685
 Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
 690 695 700

 Phe Lys Glu Ala Glu Asp Asn Pro Arg Lys Arg Arg Val Gln Thr Val
 705 710 715 720
 Asn Gly Val Ile Thr Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys
 725 730 735
 Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
 740 745 750
 Ala Cys Thr Asp His His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln

755 Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
 770 775 780
 Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
 785 790 795 800
 Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Leu
 805 810 815
 Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Ala
 820 825 830

 Ala His Lys Gln Thr Ile Met Arg Lys Gly Gly Ile Asp Lys Lys Gly
 835 840 845
 Lys Thr Ile Ile Ile Glu Arg Leu His Leu Lys Asp Ile Lys Phe Asp
 850 855 860
 Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Met Ala Thr
 865 870 875 880
 Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Gly Lys Asn Ser Lys
 885 890 895

 Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
 900 905 910
 Asn Leu Ile Lys Arg Val Lys Val Glu Gly Gln Ala Lys Ser Phe Val
 915 920 925
 Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
 930 935 940
 Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
 945 950 955 960

 Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Ala Ser Ser
 965 970 975
 Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
 980 985 990
 Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
 995 1000 1005

Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
 1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Lys Asn Glu Tyr
 1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val
 1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg
 1055 1060 1065

Asn Phe His
 1070

<210> 42

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> SV40 NLS

<

400> 42

Pro Lys Lys Lys Arg Lys Val

1 5

<210> 43

<211> 22

<212> PRT

<213> Artificial Sequence

<220><223> 3X FLAG tag

<400> 43

Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr

1 5 10 15

Lys Asp Asp Asp Asp Lys

20

<210> 44

<211> 32

<212> PRT

<213> Artificial Sequence

<220><223> Peptide linker 1

<400> 44

Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Ser Glu Thr Pro Gly Thr

1	5	10	15
Ser	Glu	Ser	Ala
Thr	Pro	Glu	Ser
Ser	Ser	Gly	Gly
Ser	Ser	Gly	Gly
Ser	Ser	Gly	Gly
Ser			
20	25	30	

<210> 45

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> Nucleoplasmin NLS

<400> 45

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys

1	5	10	15

<210> 46

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

><223> Peptide linker 2

<400> 46

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

1	5	10

<210> 47

<211> 230

<212> PRT

<213> Staphylococcus spp

<223> APG09980_cytidine deaminase

<400> 47

Met Ala Ala Gly Pro Ala Pro Glu Ala Arg Ser Leu Met Asp Glu Gln

1	5	10	15

Thr Phe Leu Asp Asn Phe Asn Asn Leu Lys Tyr Pro Arg Lys Thr Tyr

20	25	30

Leu Cys Tyr Glu Val Glu Leu Leu Val Gly Glu Asn His Ile Pro Leu

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

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

165 170 175
 Asn Leu Val Ala Arg Ser Glu Glu Leu Ser Gln Arg Leu Glu Gly Ile
 180 185 190
 Leu Gln Pro Ser Val Leu Val Phe Cys Trp Pro Ser Gln Val Ser Val
 195 200 205
 Thr Ala Ala His Ser Asp Ile Met Ser Gln Ala Ser Arg Ala Trp Glu
 210 215 220
 Lys Arg Arg Asp Pro Pro
 225 230

<210> 48

<211> 182

<212> PRT

<213> Artificial Sequence

<220><223> APG07386-CTD_cytidine deaminase

<400> 48

Met Glu Ala Thr Tyr Pro His Ile Phe Tyr Phe His Phe Lys Asn Leu

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

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

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

Leu Gln Glu Ile Leu Gln
 180

<210> 49

<211> 4158

<212> DNA

<213> Artificial Sequence

<220><223> APG09980_nAPG07433.1

<400> 49

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
 cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc 120
 agaagcctga tggacgagca gaccttctg gataacttca acaacctgaa gtaccccaga 180

aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac	240
gactacaagg gcttcgtgca caacgagggc ttgatatgg gcctggaacg gtgccacgcc	300
gagctgatct tectggaag aatggcctct tggaatctgg acaccgagct gagataccgg	360
atcacctgt tcatctcttg gtcccatgt cctgagtgcg ccgacgagct ggttaagttc	420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactggtat	480
cagggtacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg	540
accctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat	600
acacctttcc cccctggga caacctgttc gccagatccg aggaactgag ccagagactg	660
gaaggcatcc tgcagcctag cgtgctggtg tttctgtggc ctagccaagt gtccgtgacc	720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc	780
cccagcggcg gatctagcgg cggctctagc ggagtgaaa cccctggtac aagcgagtct	840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga	900
attggcctgg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcttggaac	960
aaagaccggg agagatacga gaagtcaga atcgtggatc aaggcgtgag aatgttcgac	1020
agagccgaga tgcccaagc aggcgccagc ttagctgaac ccagaagaat cgccagatcc	1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccgga cctgctggtg	1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt acccctgag caaaaagagc	1200
atggacatct ggggcattcg gctcgacggc ctggacagac tctcaatca tttcgagtgg	1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtga	1320
ctgaaagata cagagacagg caagtgctg agcagcatcc aactgaacga gaaacggctg	1380
agcttgtata gaacctggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat	1440
aggaagagaa atagcccaa cgaatactg ttcagcgtgt ctagagccga gctggaaaag	1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat	1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggaatgcc	1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattccaag	1680
gtacataca cttccaata cttctctgct ctggaccagg tgaatcggac cagactggga	1740
cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag	1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacggaag	1860
tggttggaat tggacgagac aattcagttc aagggcctga actacgaccc taacaggaa	1920

ctgaagaaga tcgagaagaa gcctttttatc aatctgaagg ccttctacga gatcaacaag	1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccaccctgga ctacgacggc	2040
atcggtacg cctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc	2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg	2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg	2220
cagaagggca atacctacaa ggaagccgtg gaccaactgg gctacgacac cagcggcctt	2280
aagaaggaga agagggtcaa gttcctgcca cctatttctg atgaaatcac gaatccaatc	2340
gtgaaaaggg cctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga	2400
tctcctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg	2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaagggcgc tatcagcatc	2520
ctgtctgagc acggcatcct gaaccttaca ggctacgaca tcgtgagata caaactgtgg	2580
aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc	2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc	2700
tacagccagt ccttcacga ctctaccac aacaaggtcc tgggttacag cgacgaaaac	2760
cggaaaaagg gcaacagaat cctttatacc tacttctctg aaaccaacaa ggattgggag	2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt	2880
ctgaagcggg ctatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac	2940
accagatacg cctctacctt cctgaagaac ttcatcgagc agaacctgca gttcaaggaa	3000
gccgaggaca accccagaaa aagacgggtg caaacctga acggcgttat caccgcccac	3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg	3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtac cgagtactat	3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttctatgcc ttgggaaggc	3240
ttccgggacg agctgctgag ccacttggct tctcagccta tcgccaagaa aatcagcgag	3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga	3360
agcattacag gcgtgctca taagcagacc atcatgcgga agggaggaat tgacaagaag	3420
ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagtctga cgagaacgga	3480
gatttcaaga tgggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga	3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc	3600
aaaaagggca caggtaacct gatcaagcgg gtgaagggtg aaggacaggc caagagcttt	3660

gtgaggaag tgaacggcgg agtggcccaa aatggcgatc tggtagagt tgatttgttt	3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag	3780
ctgccccaaa agtggtcgc cagctctaag ggctatgagc agtggctgac actggataac	3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggt caagggagat	3900
gaggatcggg tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag	3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac	4020
ttggagaagt acgaggtggg cgtgctgggc gatctaagac tggccggaa ggaaactcga	4080
agaaacttcc acagcggcgg aagcaaaaga cctgccgcta caaagaaggc cgccaggcc	4140
aagaaaaaga agtaatga	4158
<210> 50	
<211> 4530	
<212> DNA	
<213> Artificial Sequence	
<220><223> APG09980_nAPG07433.1_APG03399	
<400> 50	
atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat	60
cacgatatcg actacaagga cgacgatgat aagatggcgg ctggacctgc ccctgaagcc	120
agaagcctga tggacgagca gaccttcttg gataacttca acaacctgaa gtacccaga	180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac	240
gactacaagg gcttcgtgca caacgagggc tttgatatgg gcctggaacg gtgccacgcc	300
gagctgatct tcctggaag aatggcctct tggaatctgg acaccgagct gagataccgg	360
atcacctgtg tcattctctg gtcccatgt cctgagtgcg ccgacgagct ggttaagttc	420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactggtat	480
cagggtacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg	540
acctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat	600
acacctttcc cccctggga caacctggc gccagatccg aggaactgag ccagagactg	660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc	720
gccgcccaca gcgacatcat gagccaggcc tctagacct gggagaagcg gcgggacccc	780
cccagcggcg gatctagcgg cggctctagc ggagtgaaa cccctggtac aagcgagtct	840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga	900
attggcctgg ccatcgacac caacagcatc ggatggggcg tgatcgagct gtcctggaac	960
aaagaccggg agagatacga gaagtcaga atcgtggatc aaggcgtgag aatgttcgac	1020

agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc	1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccggaa cctgctggtg	1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt accccctgag caaaaagagc	1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg	1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtga	1320
ctgaaagata cagagacagg caagtgctg agcagcatcc aactgaacga gaaacggctg	1380
agcttgata gaaccgtggg cgagatgtgg atgaaggacc cgcacttctc taaatacgat	1440
aggaagagaa atagcccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaaaag	1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat	1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc	1620
atcctgaaca aggtcgata ctgctccctg ttgaaaggca aagaaagaag gattcccaag	1680
gctacataca ccttccaata cttctctgct ctggaccagg tgaatcggac cagactggga	1740
cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag	1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacgaag	1860
tggctggaat tggacgagac aattcagttc aagggcctga actacgaccc taacgaggaa	1920
ctgaagaaga tcgagaagaa gccttttatt aatctgaagg cttctacga gatcaacaag	1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccacctgga ctacgacggc	2040
atcggtacg ccctgacctg gtacaaaacc gacaaggata tccgcagcta cctgaagagc	2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg	2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg	2220
cagaaggga atacctaca ggaagccgtg gaccaactgg gctacgacac cagcggcctt	2280
aagaaggaga agaggtccaa gttcctgcca cctatttctg atgaaatcac gaatccaatc	2340
gtgaaaaggg ccctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga	2400
tctctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgacgg	2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaaggcgcc tatcagcatc	2520
ctgtctgagc acggcatcct gaacctaca ggctacgaca tcgtgagata caaactgtgg	2580
aaggagcagg gcgaacggtg gcctacagc ctgaaggaaa tccctgccga tacatttttc	2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc	2700
tacagccagt cttcatcga ctctaccac aacaaggtcc tggtgtacag cgacgaaaac	2760

cggaaaaagg gcaacagaat cccttatacc tacttctctgg aaaccaacaa ggattgggag 2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt 2880
ctgaagcggg cttatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac 2940
accagatacg ccctacctt cctgaagaac ttcacgagc agaacctgca gttcaaggaa 3000
gccgaggaca accccagaaa aagacgggtg caaacctga acggcgttat caccgcccac 3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgtatg 3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtgac cgagtactat 3180

cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcttatgcc ttgggaaggc 3240
ttccgggacg agctgctgag ccacttggct tctcagccta tcgccaagaa aatcagcgag 3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga 3360
agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag 3420
ggcaaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga 3480
gatttcaaga tggtagggcaa ggaacaggac atggccacat acgaagctat taaacagaga 3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc 3600

aaaaagggca caggtaacct gatcaagcgg gtgaagggtg aaggacaggc caagagcttt 3660
gtgaggggaag tgaacggcgg agtggcccaa aatggcgatc tggtagagtg tgatttggtt 3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag 3780
ctgccccaaa aggtggtcgc cagctctaag ggctatgagc agtggctgac actggataac 3840
agcttcacct ttaagttcag cctgtacct tatgatctgg tgcggctggt caaggagat 3900
gaggatcggg tctgtactt tggcaccttg gacatcgaca gcgacagact taacttcaag 3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac 4020

ttggagaagt acgaggtggg cgtgctgggc gatctaagac tggtagggaa ggaaactcga 4080
agaaacttcc acagtggagg cagcggcgga tctggaggca gcatgaaaac caccacacag 4140
gagctgaagc agtacatcac cagactgttc caactgagca acaacgagac atgggagtg 4200
gaggccctgg aagaggccgc cgagaacatc ctgcctgaga gattcatcaa caacagcctg 4260
ctggcccacc tgacactgaa cacctacacc tactacaacg atgagctgca cgagctttct 4320
atctaccctt tctgatgta cgccaacgac caactgatca gcatcggtc cctggaccac 4380
ttcgacatgg acttccctta tctgaccgac accaagaaca ccatcatcga cgagcggcac 4440

ctgctgaaac agggcgagaa caaccacgag agcggcgga gcaaaagacc tgccgctaca 4500
aagaaggccg gccaggccaa gaaaaagaag 4530

<210> 51

<211> 4530

<212> DNA

<213> Artificial Sequence

<220><223> APG09980_nAPG07433.1_APG06702

<400> 51

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat	60
cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc	120
agaagcctga tggacagcga gaccttcctg gataacttca acaacctgaa gtaccccaga	180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac	240
gactacaagg gcttcgtgca caacgagggc ttgatatgg gcctggaacg gtgccacgcc	300
gagctgatct tcttgaaaag aatggcctct tggaatctgg acaccgagct gagataccgg	360
atcacctgtg tcattctctg gtcccatgt cctgagtgcg ccgacgagct ggttaagttc	420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactggtat	480
cagggtacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg	540
acctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat	600
acacctttcc cccctggga caacctggtc gccagatccg aggaactgag ccagagactg	660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc	720
gccgcccaca gcgacatcat gagccaggcc tctagacct gggagaagcg gcgggacccc	780
cccagcggcg gatctagcgg cggtcttagc ggagtgaaa cccctggtac aagcgagtct	840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga	900
attggcctgg ccatcgacac caacagcatc ggatggggcg tgatcgagct gtcttggaa	960
aaagaccggg agagatacga gaagtcaga atcgtggatc aagcggtgag aatgttcgac	1020
agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc	1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccgga cctgctggtg	1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt accccctgag caaaaagagc	1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg	1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtga	1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg	1380
agcttgtata gaaccgtggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat	1440
aggaagagaa atagcccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaaaag	1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat	1560

ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc	1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattcccaag	1680
gctacataca ctttccaata cttctctgct ctggaccagg tgaatcggac cagactggga	1740
cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag	1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacgggaag	1860
tggctggaat tggacgagac aattcagttc aagggcctga actacgaccc taacgaggaa	1920
ctgaagaaga tcgagaagaa gccttttatac aatctgaagg ctttctacga gatcaacaag	1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccacctgga ctacgacggc	2040
atcggctacg cctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc	2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg	2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg	2220
cagaagggca atacctacaa ggaagccgtg gaccaactgg gctacgacac cagcggcctt	2280
aagaaggaga agaggtccaa gttcctgcca cctatttctg atgaaatcac gaatccaatc	2340
gtgaaaaggg cctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga	2400
tctcctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg	2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaagggcgc tatcagcatc	2520
ctgtctgagc acggcatcct gaaccttaca ggctacgaca tcgtgagata caaactgtgg	2580
aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc	2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc	2700
tacagccagt cttcatcga ctctaccac aacaagggtc tgggttacag cgacgaaaac	2760
cggaaaaagg gcaacagaat cctttatacc tacttcttgg aaaccaacaa ggattgggag	2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt	2880
ctgaagcggg cttatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac	2940
accagatacg cctctacctt cctgaagaac ttcacgagc agaacctgca gttcaaggaa	3000
gccgaggaca accccagaaa aagacgggtg caaacctga acggcggtat caccgcccac	3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg	3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtac cgagtactat	3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc	3240
ttccgggacg agctgctgag ccacttggct tctcagccta tcgccaagaa aatcagcgag	3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga	3360
agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag	3420

ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga 3480
gatttcaaga tgggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga 3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc 3600

aaaaagggca caggtaacct gatcaagcgg gtgaagggtg aaggacaggc caagagcttt 3660
gtgaggggaag tgaacggcgg agtggcccaa aatggcgatc tggtagagt tgatttgttt 3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag 3780
ctgccccaaa aggtggctgc cagctctaag ggctatgagc agtggctgac actggataac 3840
agcttcacct ttaagttcag cctgtacct tatgatctgg tgcggctggt caaggagat 3900
gaggatcggt tctgtactt tggcaccttg gacatcgaca gcgacagact taacttcaag 3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac 4020

ttggagaagt acgaggtggg cgtgctgggc gatctaagac tggccggaa ggaaactcga 4080
agaaacttcc acagtggagg cagcggcgga tctggaggca gcatgacat cgagaacgag 4140
ctgaaggaat acatcactga tctgttcgag ctgccaagg acgagcagt gctgtgcgag 4200
gccgtggagc aggtggccga taacatcctg cctgatgagt acgtgacctt gggccctctg 4260
agaaacaaga ccttggaac cttcacctac tacagcaaca cactgcacga gcggaacatc 4320
tacccttca tctgtacta ccagaacaga ctgatcgcca taggatatat cgacgaaaac 4380
tacgagatgg acttcatgta cctgcacaac accgcatca cctgatcga ccagagatac 4440

ctgattgata agggcgacga caaccacgac agcggcgga gcaaaagacc tgccgctaca 4500
aagaaggcgg gccaggccaa gaaaaagaag 4530

<210> 52
<211> 4530
<212> DNA
<213> Artificial Sequence
<220><223> APG09980_nAPG07433.1_APG05198
<400> 52

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggcgg ctggacctgc ccctgaagcc 120
agaagcctga tggacgagca gaccttcttg gataacttca acaacctgaa gtacccaga 180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac 240

gactacaagg gcttcgtgca caacgagggc tttgatatgg gcctggaacg gtgccacgcc 300
gagctgatct tcttggaag aatggcctct tggaatctgg acaccgagct gagataccgg 360

atcacctgtg tcatctcttg gtcccatgt cctgagtgcg ccgacgagct ggtaagttc	420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactggtat	480
cagggtacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg	540
acctgcacg agttcgagta ctgctggaac aattcgtgg accaccagca ggacgaagat	600
acacctttcc cccctggga caacctggc gccagatccg aggaactgag ccagagactg	660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc	720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc	780
cccagcggcg gatctagcgg cggctctagc ggagtgaaa cccctggtac aagcgagtct	840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga	900
attggccttg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcttggac	960
aaagaccggg agagatacga gaaggtcaga atcgtggatc aaggcgtgag aatgttcgac	1020
agagccgaga tgccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc	1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccgga cctgctggtg	1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt acccctgag caaaaagagc	1200
atggacatct ggggcatctg gctcgacggc ctggacagac tcctcaatca tttcgagtgg	1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtga	1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg	1380
agcttgata gaaccgtggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat	1440
aggaagagaa atagcccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaag	1500
gaaatcgtga cctgtttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat	1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc	1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattcccaag	1680
gtacataca ccttccaata cttctctgct ctggaccagg tgaatcggac cagactggga	1740
cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag	1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacggaag	1860
tggtggaat tggacgagac aattcagttc aagggcctga actacgaccc taacaggaa	1920
ctgaagaaga tcgagaagaa gccttttatc aatctgaagg cttctacga gatcaacaag	1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccacctgga ctacgacggc	2040
atcggctacg cctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc	2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg	2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg	2220

cagaaggga atacctaca ggaagccgtg gaccaactgg gctacgacac cagcgccctt	2280
aagaaggaga agaggtccaa gtctctgcca cctatttctg atgaaatcac gaatccaatc	2340
gtgaaaaggg cctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga	2400
tctctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg	2460
acaaagatcg tcagcgccca ggatgaaaaac tacaagaaaa acaaggcgcg tatcagcatc	2520
ctgtctgagc acggcatcct gaacctaca ggctacgaca tcgtgagata caaactgtgg	2580
aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc	2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc	2700
tacagccagt ccttcatcga ctctaccac aacaagggtc tgggttacag cgacgaaaac	2760
cggaaaaagg gcaacagaat cctttatacc tacttctctg aaaccaacaa ggattgggag	2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt	2880
ctgaagcggg ctatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac	2940
accagatacg cctctacctt cctgaagaac ttcacgagc agaacctgca gttcaaggaa	3000
gccgaggaca accccagaaa aagacgggtg caaacctga acggcggttat caccgcccac	3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg	3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtgc cgagtactat	3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc	3240
ttccgggacg agctgctgag ccacttggct tctcagccta tcgccaagaa aatcagcgag	3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga	3360
agcattacag gcgtgctca taagcagacc atcatgcgga agggaggaat tgacaagaag	3420
ggcaaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga	3480
gatttcaaga tgggtggcaa ggaacaggac atggccacat acgaagctat taaacagaga	3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc	3600
aaaaaggga caggtaacct gatcaagcgg gtgaaggtgg aaggacaggc caagagcttt	3660
gtgaggggaag tgaacggcgg agtggcccaa aatggcgatc tggttagagt tgatttgttt	3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag	3780
ctgccccaaa aggtggtcgc cagctctaag ggctatgagc agtggctgac actggataac	3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggt caaggagat	3900
gaggatcggg tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag	3960

gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac 4020

ttggagaagt acgaggtggg cgtgctgggc gatctaagac tgggtccgaa ggaaactcga 4080

agaaacttcc acagtggagg cagcggcgga tctggaggca gcatgaccct ggaaaagcaa 4140

ctgaaggaat acatcaccaa cctgttcaac ctgccaagg acgagaagt ggagtgcgag 4200

agcatcgagg aagtggccga caacatcctg cctgatcagt acatcagact gggccctctg 4260

accaacaaga tectgcatac atacacctac tacagcgaca ccctgcacga gcggcacatc 4320

tacccttca tcttttatta ccagaaaca ctgatcgcca tcggctacat cgacgagaca 4380

aacgacatgg acttctctga cctgcacaac accgtgatgc ctctgctgga ccagagacac 4440

ctgctggaag aagagaacta caacaacgag agcggcgga gcaaaagacc tgccgctaca 4500

aagaaggccg gccaggccaa gaaaaagaag 4530

<210> 53

<211> 4014

<212> DNA

<213> Artificial Sequence

<220><223> APG07386-CTD_nAPG07433.1

<400> 53

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60

cacgatatcg actacaagga cgacgatgat aagatggagg ccacctacc ccacatcttc 120

tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc 180

acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat 240

caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa 300

gatatactga gccctaacac cgactaccgg gtgacatggt acacaagctg gtcccatgt 360

ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaact gaagctggct 420

atctttgccg ctagactgta ctatttctgg gacctcact accagcagg cctgagaagc 480

ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg 540

gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaaggcct gaagtacaac 600

ttctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggtac tagcggcggc 660

tctagcggca gtgaaacccc tggtaacaagc gagtctgcca cacctgagag cagcggcgga 720

agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac 780

agcatcggat ggggcgtgat cgagctgtcc tggaacaaag accgggagag atacgagaag 840

gtcagaatcg tggatcaagg cgtgagaatg ttcgacagag ccgagatgcc caagacaggc 900

gccagcttag ctgaaccag aagaatcgcc agatccagca gacggagact gaatcgcaag 960

tcccagagaa agaaaaacat ccggaacctg ctggtgcaac acggcgtgat cacacaggag 1020

gaactggata gcctgtaccc cctgagcaaa aagagcatgg acatctgggg cattcggctc 1080

gacggcctgg acagactcct caatcatttc gagtgggcca gactgctgat ccacctggct 1140

cagagacggg gctttaagtc caacagaaag agtgaactga aagatacaga gacaggcaag 1200

gtgctgagca gcatccaact gaacgagaaa cggtgagct tgtatagaac cgtggcgag 1260

atgtggatga aggaccccg cttctctaaa tacgatagga agagaaatag cccaacgaa 1320

tacgtgttca gcgtgtctag agccgagctg gaaaaggaaa tegtaccct gttcgccgc 1380

cagcggagat tccagagccc ttacgccagc aaagatctgc aggagacata tctgcagatc 1440

tggaccacc aactgccttt cgccagcggc aatgccatcc tgaacaaggt cggatactgc 1500

tcctgttga aaggcaaaga aagaaggatt cccaaggcta catacacctt ccaatacttc 1560

tctgctctgg accagggtgaa tccgaccaga ctgggacctg atttcagcc cttcaccaag 1620

gagcaacggg aaattatctt gaacaacatg ttccagagga cagattacta caagaagaaa 1680

accatccccg aggtgacctt ctatgacata cggaagtggc tggaaattgga cgagacaatt 1740

cagttcaagg gcctgaacta cgacctaac gaggaactga agaagatcga gaagaagcct 1800

tttatcaatc tgaaggcctt ctacgagatc aacaagggtg tggccaacta cagcgaaaga 1860

accaacgaga ctttctccac cctggactac gacggcatcg gctacgcctt gaccgtgtac 1920

aaaaccgaca aggatatccg cagctacctg aagagcagtc acaacctacc taagagatgc 1980

tacgacgacc aactgatcga ggaactgctg agcctgagct acacaaagtt cggccacctg 2040

tcctgaaag ccatcaacca cgtgctgtct atcatgcaga agggcaatac ctacaaggaa 2100

gccgtggacc aactgggcta cgacaccagc ggccttaaga aggagaagag gtccaagttc 2160

ctgccaccta tttctgatga aatcacgaat ccaatcgtga aaagggcctt gacccaggcc 2220

agaaaagtgg tgaacgcat aattagaaga cacggatctc ctactccgt gcacatcgag 2280

ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cgcccaggat 2340

gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac 2400

cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc 2460

tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac 2520

ggcgccccta tccttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc 2580

taccacaaca aggtccttgt gtacagcgac gaaaaccgga aaaagggcaa cagaatccct 2640

tatacctact tcctggaaac caacaaggat tgggaggcct ttgagcggtg cgtgcggagc 2700

aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga 2760

gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgcttc taccttcctg 2820

aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga 2880

cgggtgcaaa ccgtgaacgg cgttatcacc gccaccttca gaaagcgggt gggcctggag 2940

aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca 3000

gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc 3060

gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac 3120

ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc 3180

ctggactaca tcttcgtgtc cagaatgcct aagagaagca ttacaggcgc tgctcataag 3240

cagaccatca tgcggaaggg aggaattgac aagaaggga aaacaatcat catcgacgg 3300

ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggt gggcaaggaa 3360

caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc 3420

aagaaggcct tcgagacccc tctgtacaag ccagcaaaa agggcacagg taacctgac 3480

aagcgggtga aggtggaagg acaggccaag agctttgtga gggaagtga cggcggagt 3540

gccccaaatg gcgatctggt tagagttgat ttgtttgaga aggatgataa gtactacatg 3600

gtccccatct acgtgccaga caccgtgtgt agcgagctgc ccaaaaaggt ggtcgccagc 3660

tctaagggt atgagcagtg gctgacactg gataacagct tcacctttaa gttcagcctg 3720

taccttatg atctggtgcg gctggtcaag ggagatgagg atcggttcct gtactttggc 3780

accttgga tgcacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag 3840

aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg 3900

ctgggcgac taagactggt ccggaaggaa actcgaagaa acttccacag cggcgggaagc 3960

aaaagacctg ccgtacaaa gaagccggc caggccaaga aaaagaagta atga 4014

<210> 54

<211> 4386

<212> DNA

<213> Artificial Sequence

<220><223> APG07386-CTD_nAPG07433.1_APG03399

<400> 54

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60

cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc 120

tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc 180

acaatgaaa tcataagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat	240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa	300
gatatcctga gccctaacac cgactaccgg gtgacatggt acacaagctg gtcccatgt	360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaactg gaagctggct	420
atctttgccg ctagactgta ctatttctgg gacctcact accagcaggg cctgagaagc	480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg	540
gaaaactttg tgtacaacgg cgacgagcca ttcaagcctt ggaaggcct gaagtacaac	600
ttctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggatc tagcggcggc	660
tctagcggca gtgaaacccc tggtaacaagc gactctgcca cacctgagag cagcggcggc	720
agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac	780
agcatcggat gggcggtgat cgagctgtcc tggaacaaag accgggagag atacgagaag	840
gtcagaatcg tggatcaagg cgtgagaatg ttgcacagag ccgagatgcc caagacaggc	900
gccagcttag ctgaaccagc aagaatcgcc agatccagca gacggagact gaatcgcaag	960
tcccagagaa agaaaaacat ccggaacctg ctggtgcaac acggcgtgat cacacaggag	1020
gaactggata gcctgtacct cctgagcaaa aagagcatgg acatctgggg cattcggtc	1080
gacggcctgg acagactcct caatcatttc gactgggcca gactgctgat ccacctggct	1140
cagagacggg gctttaagtc caacagaaaag agtgaactga aagatacaga gacaggcaag	1200
gtgctgagca gcatccaact gaacagaaaa cggtctgagct tgtatagaac cgtgggcgag	1260
atgtggatga aggacccga cttctctaaa tacgatagga agagaaatag ccccaacgaa	1320
tacgtgttca gcgtgtctag agccgagctg gaaaaggaaa tcgtgacctt gttcgccgcc	1380
cagcggagat tccagagccc ttacccagc aaagatctgc aggagacata tctgcagatc	1440
tggaccacc aactgccttt cgccagcggc aatgcatcc tgaacaaggt cggatactgc	1500
tccctgttga aaggcaaaga aagaaggatt cccaaggcta catacacctt ccaatacttc	1560
tctgctctgg accagtgaa tcggaccaga ctgggacctg atttcagcc cttaccaag	1620
gagcaacggg aaattatctt gaacaacatg ttccagagga cagattacta caagaagaaa	1680
accatccccg aggtgacctt ctatgacata cggaagtggc tggaattgga cgagacaatt	1740
cagttcaagg gcctgaacta cgacctaac gaggaactga agaagatcga gaagaagcct	1800
tttatcaatc tgaaggcctt ctacgagatc aacaaggtgg tggccaacta cagcgaaaga	1860
accaacgaga cttctccac cctggactac gacggcatcg gctacgcctt gaccgtgtac	1920

aaaaccgaca aggatatccg cagctacctg aagagcagtc acaacctacc taagagatgc	1980
tacgacgacc aactgatcga ggaactgctg agcctgagct acacaaagt t cggccacctg	2040
tccctgaaag ccatcaacca cgtgctgtct atcatgcaga agggcaatac ctacaaggaa	2100
gccgtggacc aactgggcta cgacaccagc ggccttaaga aggagaagag gtccaagttc	2160
ctgccaccta tttctgatga aatcacgaat ccaatcgtga aaagggccct gaccagggc	2220
agaaaagtgg tgaacgcat aattagaaga cacggtctc ctcactccgt gcacatcgag	2280
ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cggccaggat	2340
gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac	2400
cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc	2460
tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac	2520
ggcgccccta tcttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc	2580
taccacaaca aggtccttgt gtacagcgac gaaaaccgga aaaagggcaa cagaatccct	2640
tatacctact tcttggaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc	2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga	2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgctc taccttctg	2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga	2880
cgggtgcaaa ccgtgaacgg cgttatcacc gccacttca gaaagcgggt gggcctggag	2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca	3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc	3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac	3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc	3180
ctggactaca tcttctgtgc cagaatgcct aagagaagca ttacaggcgc tgctcataag	3240
cagaccatca tgcggaaggg aggaattgac aagaaggca aaacaatcat catcgaacgg	3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggt gggcaaggaa	3360
caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc	3420
aagaaggcct tcgagacccc tctgtacaag ccagcaaaa agggcacagg taacctgatc	3480
aagcgggtga aggtggaagg acaggccaag agctttgtga ggggaagtga cggcggagt	3540
gccccaaatg gcgatctggt tagagttgat ttgtttgaga aggatgataa gtactacatg	3600
gtccccatct acgtgccaga caccgtgtgt agcgagctgc caaaaaggt ggtcgccagc	3660

tctaagggt atgagcagt gctgacactg gataacagct tcacctttaa gttcagcctg 3720
 tacccttatg atctggtgcg gctggtcaag ggagatgagg atcggttcct gtactttggc 3780
 accctggaca tcgacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag 3840
 aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg 3900
 ctgggcgac taagactggt ccggaaggaa actcgaagaa acttccacag tggaggcagc 3960
 ggcggtatctg gaggcagcat gaaaaccacc acacaggagc tgaagcagta catcaccaga 4020
 ctgtttcaac tgagcaacaa cgagacatgg gagtgcgagg ccttgggaaga ggccgcccag 4080

aacatcctgc ctgagagatt catcaacaac agcctgctgg cccacctgac actgaacacc 4140
 tacacctact acaacgatga gctgcacgag ctttctatct accccttcct gatgtacgcc 4200
 aacgaccaac tgatcagcat cggctacctg gaccacttcg acatggactt cttttatctg 4260
 accgacacca agaaccat catcgacgag cggcacctgc tgaacaggcg cgagaacaac 4320
 cacgagagcg gcggaagcaa aagacctgcc gctacaaaga aggccggcca ggccaagaaa 4380
 aagaag 4386

<210> 55

<211> 4386

<212> DNA

<213> Artificial Sequence

<220><223> APG07386-CTD_nAPG07433.1_APG06702

<400> 55

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
 cacgatatcg actacaagga cgacgatgat aagatggagg ccacctacc ccacatcttc 120
 tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc 180
 acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat 240
 caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa 300
 gatatacctga gccctaacac cgactaccgg gtgacatggt acacaagctg gtcccatgt 360

ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaact gaagctggct 420
 atctttgccg ctagactgta ctatttctgg gacctcact accagcaggg cctgagaagc 480
 ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg 540
 gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaaggcct gaagtacaac 600
 ttctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggtac tagcggcggc 660
 tctagcggca gtgaaacccc tggtaaacg gagtctgcc cacctgagag cagcggcgga 720

agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac 780

 agcatcggat ggggcgtgat cgagctgtcc tggaacaaag accgggagag atacgagaag 840
 gtcagaatcg tggatcaagg cgtgagaatg ttcgacagag ccgagatgcc caagacaggc 900
 gccagcttag ctgaaccag aagaatcgcc agatccagca gacggagact gaatcgcaag 960
 tcccagagaa agaaaaacat ccggaacctg ctggtgcaac acggcgtgat cacacaggag 1020
 gaactggata gcctgtacc cctgagcaaa aagagcatgg acatctgggg cattcggctc 1080
 gacggcctgg acagactcct caatcatttc gactgggcca gactgctgat ccacctggct 1140
 cagagacggg gctttaagtc caacagaaag agtgaactga aagatacaga gacaggcaag 1200

 gtgctgagca gcatccaact gaacgagaaa cgctgagct tgtatagaac cgtgggcgag 1260
 atgtggatga aggaccccg cttctctaaa tacgatagga agagaaatag ccccaacgaa 1320
 tacgtgttca gcgtgtctag agccgagctg gaaaaggaaa tctgaccct gttcgccgcc 1380
 cagcggagat tccagagccc ttacgccagc aaagatctgc aggagacata tctgcagatc 1440
 tggaccacc aactgccttt cgccagcggc aatgccatcc tgaacaaggt cggatactgc 1500
 tccctgttga aaggcaaaga aagaaggatt cccaaggcta catacacctt ccaatacttc 1560
 tctgctctgg accaggtgaa tcggaccaga ctgggacctg atttccagcc cttaccaag 1620

 gagcaacggg aaattatctt gaacaacatg ttccagagga cagattacta caagaagaaa 1680
 accatcccc aggtgacctt ctatgacata cggaagtggc tggaaattgga cgagacaatt 1740
 cagttcaagg gcctgaacta cgaccctaac gaggaactga agaagatcga gaagaagcct 1800
 tttatcaatc tgaaggcctt ctacgagatc aacaagggtg tggccaacta cagcgaaaga 1860
 accaacgaga ccttctccac cctggactac gacggcatcg gctacgcct gaccgtgtac 1920
 aaaaccgaca aggatatccg cagctacctg aagagcagtc acaacctacc taagagatgc 1980
 tacgacgacc aactgatcga ggaactgctg agcctgagct acacaaagtt cggccacctg 2040

 tccctgaaag ccatcaacca cgtgctgtct atcatgcaga agggcaatac ctacaaggaa 2100
 gccgtggacc aactgggcta cgacaccagc ggccttaaga aggagaagag gtccaagtgc 2160
 ctgccaccta tttctgatga aatcacgaat ccaatcgtga aaagggcct gacccaggcc 2220
 agaaaagtgg tgaacgcat aattagaaga cacgcatctc ctactccgt gcacatcgag 2280
 ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cggccaggat 2340
 gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac 2400
 cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc 2460

tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac	2520
ggcgccccta tccttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc	2580
taccacaaca aggtccttgg gtacagcgac gaaaaccgga aaaagggcaa cagaatccct	2640
tatacctact tcctggaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc	2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga	2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgctc taccttctg	2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga	2880
cgggtgcaaa ccgtgaacgg cgttatcacc gcccacttca gaaagcggtg gggcctggag	2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca	3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc	3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac	3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc	3180
ctggactaca tcttcgtgc cagaatgcct aagagaagca ttacaggcgc tgctcataag	3240
cagaccatca tgcggaaggg aggaattgac aagaagggca aaacaatcat catcgaacgg	3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggt gggcaaggaa	3360
caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc	3420
aagaaggcct tcgagacccc tctgtacaag ccagcaaaa agggcacagg taacctgatc	3480
aagcgggtga aggtggaagg acaggccaag agctttgtga gggaagtga cggcggagt	3540
gccccaaatg gcatcttgg tagagttgat ttgtttgaga aggatgataa gtactacatg	3600
gtcccatct acgtgccaga caccgtgtgt agcgagctgc caaaaaggt ggtcgccagc	3660
tctaagggt atgagcagt gctgacactg gataacagct tcacctttaa gttcagcctg	3720
taccttatg atctggtcgg gctggtcaag ggagatgagg atcggttctt gtactttggc	3780
accctggaca tcgacacga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag	3840
aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg	3900
ctgggcgatc taagactggt ccggaaggaa actcgaagaa acttccacag tggaggcagc	3960
ggcggatctg gaggcagcat gaccatcgag aacgagctga aggaatacat cactgatctg	4020
ttcgagctgc ccaaggacga gcagtggctg tgcgaggccg tggacgaggt ggccgataac	4080
atctgcctg atgagtacgt gaccttgggc cctctgagaa acaagaccct ggaaaccttc	4140
acctactaca gcaacacact gcacgagcgg aacatctacc cttcatcct gtactaccag	4200
aacagactga tcgcatagg atatatcgac gaaaactacg agatggactt catgtacctg	4260
cacaacaccg ccatcacct gatcgaccag agatacctga ttgataaggg cgacgacaac	4320

cacgacagcg gcggaagcaa aagacctgcc gctacaaaga aggccggcca ggccaagaaa 4380
aagaag 4386

<210> 56
<211> 4386
<212> DNA
<213> Artificial Sequence
<220><223> APG07386-CTD_nAPG07433.1_AP05198
<400> 56

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc 120
tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc 180
acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat 240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa 300
gatatcctga gccctaacac cgactaccgg gtgacatggt acacaagctg gtcccatgt 360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaactg gaagctggct 420

atctttgccg ctagactgta ctatttctgg gacctcact accagcaggg cctgagaagc 480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg 540
gaaaactttg tgtacaacgg cgacagacca ttcaagccct ggaagggcct gaagtacaac 600
ttctgttcc tggacagcaa gctgcaggag attctgcaga gcggcgatc tagcggcggc 660
tctagcggca gtgaaacccc tggtaacaag gagtctgcca cacctgagag cagcggcggg 720
agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac 780
agcatcggat ggggcgtgat cgagctgtcc tggaacaaag accgggagag atacgagaag 840

gtcagaatcg tggatcaagg cgtgagaatg ttcgacagag ccgagatgcc caagacaggc 900
gccagcttag ctgaaccag aagaatcgcc agatccagca gacggagact gaatcgcaag 960
tcccagagaa agaaaaacat ccggaacctg ctggtgcaac acggcgtgat cacacaggag 1020
gaactggata gcctgtaccc cctgagcaaa aagagcatgg acatctgggg cattcggtc 1080
gacggcctgg acagactcct caatcatttc gactgggcca gactgctgat ccacctggct 1140
cagagacggg gctttaagtc caacagaaaag agtgaactga aagatacaga gacaggcaag 1200
gtgctgagca gcatccaact gaacgagaaa cggtgagct tgtatagaac cgtgggcgag 1260

atgtggatga aggaccccg cttctctaaa tacgatagga agagaaatag ccccaacgaa 1320
tacgtgttca gcgtgtctag agccgagctg gaaaaggaaa tcgtgaccct gttcgccgcc 1380

cagcggagat tccagagccc ttacgccagc aaagatctgc aggagacata tctgcagatc	1440
tggaccacc aactgccttt cgccagcggc aatgccatcc tgaacaaggt cggatactgc	1500
tccctgttga aaggcaaaga aagaaggatt cccaaggcta catacacctt ccaatacttc	1560
tctgctctgg accaggtgaa tccgaccaga ctgggacctg atttccagcc cttaccaag	1620
gagcaacggg aaattatctt gaacaacatg ttccagagga cagattacta caagaagaaa	1680
accatccccg aggtgacctt ctatgacata cggaagtggc tggaaattgga cgagacaatt	1740
cagttcaagg gcctgaacta cgaccctaac gaggaactga agaagatcga gaagaagcct	1800
tttatcaatc tgaaggcctt ctacgagatc aacaagggtg tggccaacta cagcgaaaga	1860
accaacgaga ccttctccac cctggactac gacggcatcg gctacgcctt gaccgtgtac	1920
aaaaccgaca aggatatccg cagctacctg aagagcagtc acaacctacc taagagatgc	1980
tacgacgacc aactgatcga ggaactgctg agcctgagct acacaaagtt cggccacctg	2040
tccctgaaag ccatcaacca cgtgctgtct atcatgcaga agggcaatac ctacaaggaa	2100
gccgtggacc aactgggcta cgacaccagc ggccttaaga aggagaagag gtccaagttc	2160
ctgccacctt tttctgatga aatcacgaat ccaatcgtga aaagggcctt gaccagggc	2220
agaaaagtgg tgaacgccat aattagaaga cacggatctc ctactccgt gcacatcgag	2280
ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cgcccaggat	2340
gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac	2400
cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc	2460
tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac	2520
ggcgccccct tcttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc	2580
taccacaaca aggtcctggt gtacagcgac gaaaaccgga aaaagggcaa cagaatccct	2640
tatacctact tcttgaaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc	2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga	2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatagcctc taccttctg	2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga	2880
cgggtgcaaa ccgtgaacgg cgttatcacc gccacttca gaaagcgggt gggcctggag	2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca	3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc	3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac	3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc	3180
ctggactaca tcttcgtgtc cagaatgcct aagagaagca ttacaggcgc tgctcataag	3240

cagaccatca tgcggaaggg aggaattgac aagaagggca aaacaatcat catcgaacgg	3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggt gggcaaggaa	3360
caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc	3420
aagaaggcct tcgagacccc tctgtacaag ccagcaaaa agggcacagg taacctgatc	3480
aagcgggtga aggtggaagg acaggccaag agctttgtga gggaagtga cggcggagt	3540
gccccaaatg gcgatctggt tagagttgat ttgtttgaga aggatgataa gtactacatg	3600
gtccccatct acgtgccaga caccgtgtgt agcgagctgc caaaaaggt ggtcgccagc	3660
tctaagggt atgagcagt gctgacactg gataacagct tcacctttaa gttcagcctg	3720
taccttatg atctggtcgc gctggtcaag ggagatgagg atcggttcct gtactttggc	3780
accctggaca tcgacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag	3840
aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg	3900
ctgggcgatc taagactggt ccggaaggaa actcgaagaa acttcacag tggaggcagc	3960
ggcgggatctg gaggcagcat gacctggaa aagcaactga aggaatacat caccaacctg	4020
ttcaacctgc ccaaggacga gaagtgggag tgcgagagca tcgaggaagt ggccgacaac	4080
atcctgcctg atcagtacat cagactgggc cctctgacca acaagatcct gcatacatc	4140
acctactaca gcgacacct gcacgagcgg cacatctacc ccttcacct ttattaccag	4200
aaacaactga tcgccatcgg ctacatcgac gagacaaacg acatggactt cctgtacctg	4260
cacaacaccg tgatgcctct gctggaccag agacacctgc tggaanaaga gaactacaac	4320
aacgagagcg gcggaagcaa aagacctgcc gctacaaaga aggccggcca ggccaagaaa	4380
aagaag	4386
<210> 57	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> SGN000143 target sequence	
<400> 57	
catggcagta cattagagca	20
<210> 58	
<211> 20	
<212> DNA	
<213> Artificial Sequence	

<220><223> SGN000169 target sequence
 <400> 58
 cacatctcga gcaagacgtt 20
 <210> 59
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> SGN000173 target sequence
 <400> 59
 cttctatagc ctccttcccc 20
 <210> 60
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> SGN000929 target sequence
 <400> 60
 atctggaggg aacttacagc atatg 25
 <210> 61
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> SGN000930 target sequence
 <400> 61
 gaacaactca aatggaaatg aatat 25
 <210> 62
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> SGN001101 target sequence
 <400> 62
 atattttctt taatggtgcc aggca 25
 <210> 63
 <211> 130
 <212> RNA

<213> Artificial Sequence

<220><223> SGN000143 sgRNA sequence

<400> 63

cauggcagua cauuagagca gucauaguuc cauuaaagcc aaaaguggcu uugauguuuc 60

uauaauaagg guuucgaccc guggcgucgg ggaucgccug ccuauugaaa ugggcuucuc 120

cccauuuuuu 130

<210> 64

<211> 130

<212> RNA

<213> Artificial Sequence

<220><223> SGN000169 sgRNA sequence

<400> 64

cacauucga gcaagacguu gucauaguuc cauuaaagcc aaaaguggcu uugauguuuc 60

uauaauaagg guuucgaccc guggcgucgg ggaucgccug ccuauugaaa ugggcuucuc 120

cccauuuuuu 130

<210> 65

<211> 130

<212> RNA

<213> Artificial Sequence

<220><223> SGN000173 sgRNA sequence

<400> 65

cuucuaagc cuccuucucc gucauaguuc cauuaaagcc aaaaguggcu uugauguuuc 60

uauaauaagg guuucgaccc guggcgucgg ggaucgccug ccuauugaaa ugggcuucuc 120

cccauuuuuu 130

<210> 66

<211> 135

<212> RNA

<213> Artificial Sequence

<220><223> SGN000929 sgRNA sequence

<400> 66

aucuggaggg aacuuacagc auauagguc auuuccauga aagccaaaag uggcuuugau 60

guuucuauga uaaggguuuc ggcccugggc gucggggauc gccugcccau uccgaugggc 120

uucuccccau uuauu 135

<210> 67

<211> 135

<212> RNA

<213> Artificial Sequence

<220><223> SGN000930 sgRNA sequence

<400> 67

gaacaacuca aauggaaaug aauaugucau aguuccauga aagccaaaag uggcuuugau 60

guuucuauga uaagggguuc ggcccguggc gucggggauc gccugcccau uccgaugggc 120

uucuccccau uuauu 135

<210> 68

<211> 130

<212> RNA

<213> Artificial Sequence

<220><223> SGN001101 sgRNA sequence

<400> 68

auauuuuuu uaaaggugcc aggcagucau aguuccauua aagccaaaag uggcuuugau 60

guuucuauga uaagggguuc gaccguggc gucggggauc gccugcccau ugaaaugggc 120

uucuccccau 130

<210> 69

<211> 200

<212> DNA

<213> Artificial Sequence

<220><223> SGN000143 genomic locus

<400> 69

gaagtcagca atcttcagct ctcccttgag ccctaagagc agattttctg gctttatgtc 60

tctgtgaate acctttcttc catggcagta cattagagca tctgccaact cctccatgat 120

ctgggagggg caacgacagg caatcagaca aggatggacc tccagctaca agcagcacac 180

cctcagcctc gagcccccta 200

<210> 70

<211> 200

<212> DNA

<213> Artificial Sequence

<220><223> SGN000169 genomic locus

<400> 70

tccagcaggt cagcaaagaa tttatagccc cccttgagca cacagagggc tacaatgtga 60

tggcctccca tctccttcat cacatctcga gcaagacgtt cagtcctaca gaaataaaat 120

caggaattta atagaaagtt tcatacatta aactttataa caaacacctc ttagtcatta 180

aacttcaca ccaacctggg 200

<210> 71

<211> 200

<212> DNA

<213> Artificial Sequence

<220><223> SGN000173 genomic locus

<400> 71

gtataactgg ccagccttta atacattgct gcttagagtc aaagcatgta cttagagttg 60

glatgattta tctttttggt cttctatagc ctcttccccc atcccatca gtcttaataca 120

gtcttggttac gttatgacta atctttgggg atttgtgcag aatgttattt tagataagca 180

aaaacgagca aaatagggga 200

<210> 72

<211> 200

<212> DNA

<213> Artificial Sequence

<220><223> SGN000929 genomic locus

<400> 72

gactctccca cccagacct cctagttctt gggcaggtga acaactacag cctttccatg 60

cctccttggt gagagatctg gaggaactt acagcatatg agcacctgct agggagagcc 120

cctggccctg gcaggaccga tacctcggt cctggacacc atttctattt actcaatata 180

cactgtttgt acaaagtaca 200

<210> 73

<211> 200

<212> DNA

<213> Artificial Sequence

<220><223> SGN000930 genomic locus

<400> 73

tttatcatca caatacctcg ggtccagatg ggccccttag tgagttccgc taccacagga 60
 aggatgactc aataagaaca actcaaatgg aatgaatat gggcccaggg tgggtgtgag 120
 gagtggcagg ggattcttga gagacaggac cattgcccac agcctatgtg agtactgttg 180
 ctgtgtctga aagaaggcaa 200

<210> 74

<211> 200

<212> DNA

<213> Artificial Sequence

<220><223> SGN001101 genomic locus

<400> 74

tttcttacct cttctagttg gcatgctttg atgacgcttc tgtatctata ttcacatag 60

gaaacaccaa agatgatatt ttctttaatg gtgccaggca taatccagga aaactgagaa 120
 cagaatgaaa ttcttcact gtgcttaatt ttaccctctg aaggtccag ttctccata 180
 atcaccatta gaagtgaagt 200

<210> 75

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> meganuclease family motif

<400> 75

Leu Ala Gly Leu Ile Asp Ala Asp Gly

1 5

<210> 76

<211> 182

<212> PRT

<213> Artificial Sequence

<220><223> Truncated APG09980_cystidine deaminase

<400> 76

Met Asp Glu Gln Thr Phe Leu Asp Asn Phe Asn Asn Leu Lys Tyr Pro

1 5 10 15

Arg Lys Thr Tyr Leu Cys Tyr Glu Val Glu Leu Leu Val Gly Glu Asn

20 25 30

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

35 40 45
Asp Met Gly Leu Glu Arg Cys His Ala Glu Leu Ile Phe Leu Glu Arg
50 55 60

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

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

<210> 77

<211> 201

<212> PRT

<213> Artificial Sequence

<220><223> APG05840_cytidine deaminase

<400> 77

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

Arg Asp Leu Met Asn Glu Gln Arg Asp Phe Leu Cys Asn Gln Ala Lys

50 55 60

Asn His Phe Ser Gly Ser Glu Asp His His Ala Glu Arg Cys Phe Leu

65 70 75 80

Asp Arg Ile Pro Ser Trp Gln Leu Asp Pro Ala Gln Thr Tyr Arg Val

85 90 95

Thr Cys Phe Ile Ser Trp Ser Pro Cys Phe Ser Cys Ala Gln Glu Val

100 105 110

Ala Glu Phe Leu His Glu Asn Pro His Val Asn Leu Arg Ile Phe Ala

115 120 125

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

130 135 140

Leu Gln Asn Ala Gly Ala Gln Val Ser Ile Met Thr Ser Glu Glu Phe

145 150 155 160

Gly His Cys Trp Asp Thr Phe Val Asp Arg Gln Gly His Pro Phe Gln

165 170 175

Pro Trp Glu Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu

180 185 190

Gln Ala Ile Leu Gln Asn Gln Gly Asn

195 200

<210> 78

<211> 185

<212> PRT

<213> Artificial Sequence

<220><223> Truncated APG05840_cytidine deaminase

<400> 78

Met Asp Pro His Met Phe Thr Gly Asn Phe Thr Asn Asn Pro Arg Val

1 5 10 15

Phe Gly Leu His Gln Thr Tyr Leu Cys Tyr Glu Val Lys Arg Gln Gly

20 25 30

Pro Asp Gly Thr Arg Asp Leu Met Asn Glu Gln Arg Asp Phe Leu Cys

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

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

 His Pro Phe Gln Pro Trp Glu Gly Leu Asp Glu His Ser Gln Ala Leu
 165 170 175
 Ser Gly Arg Leu Gln Ala Ile Leu Gln
 180 185

 <210> 79
 <211> 192
 <212> PRT
 <213> Artificial Sequence
 <220><223> APG00868_cytidine deaminase
 <400> 79
 Met Glu Pro Trp Arg Pro Ser Pro Arg Asn Pro Met Asp Arg Ile Asp
 1 5 10 15
 Pro Asn Thr Phe Arg Phe His Phe Pro Asn Leu Leu Tyr Ala Ser Gly

 20 25 30
 Arg Lys Leu Cys Tyr Leu Cys Phe Gln Val Glu Thr Gly Asp Tyr Phe
 35 40 45
 Ser Cys Asp Asp Ser Asp Arg Gly Val Phe Arg Asn Lys Val His Pro

50 55 60
 Trp Ala Arg Cys His Ala Glu Gln Cys Phe Leu Ser Trp Phe Arg Asp
 65 70 75 80
 Gln Tyr Pro Cys Arg Asp Glu Tyr Tyr Asn Val Thr Trp Phe Leu Ser

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

 145 150 155 160
 Asp Asn Phe Val Asp His Lys Gly Met Arg Phe Gln Arg Arg Asn Leu
 165 170 175
 Leu Lys Asp Tyr Asp Phe Leu Ala Ala Glu Leu Gln Glu Ile Leu Arg
 180 185 190
 <210> 80
 <211> 178
 <212> PRT
 <213> Artificial Sequence
 <220><223> Truncated APG00868_cytidine deaminase
 <400> 80
 Met Asp Pro Asn Thr Phe Arg Phe His Phe Pro Asn Leu Leu Tyr Ala

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

Arg Asp Gln Tyr Pro Cys Arg Asp Glu Tyr Tyr Asn Val Thr Trp Phe

65 70 75 80

Leu Ser Trp Ser Pro Cys Pro Thr Cys Ala Glu Glu Val Val Glu Phe

85 90 95

Leu Glu Glu Tyr Arg Asn Leu Thr Leu Ser Ile Phe Thr Ser Arg Leu

100 105 110

Tyr Tyr Phe Tyr His Pro Asn Tyr Gln Gln Gly Leu Arg Lys Leu Trp

115 120 125

Asp Ala Gly Val Gln Leu Asp Ile Met Ser Cys Asp Asp Phe Glu His

130 135 140

Cys Trp Asp Asn Phe Val Asp His Lys Gly Met Arg Phe Gln Arg Arg

145 150 155 160

Asn Leu Leu Lys Asp Tyr Asp Phe Leu Ala Ala Glu Leu Gln Glu Ile

165 170 175

Leu Arg

<210> 81

<211> 152

<212> PRT

<213> Artificial Sequence

<220><223> APG30125_cytidine deaminase

<400> 81

Met Ile Asn Leu Val Gln Thr Pro Tyr Asp Leu Asn Ser Gly Tyr Pro

1 5 10 15

Ile Val Arg Arg Thr Leu Glu Asp Lys Lys Lys Leu Val Lys His Glu

20 25 30

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

35 40 45

Gly Asn Ser Arg Tyr Ala Phe Gly Asn Ser Gln Met Gln Val Glu Met

50 55 60

Pro Pro Asn Ile Tyr Ala His Asn Trp Val Lys Leu His Gly Glu Met

65					70				75				80			
Thr	Ala	Leu	Val	Ala	Ala	Ile	Arg	Arg	Ile	Glu	Arg	Val	Asp	Ser	Thr	
				85				90				95				
Ser	Ala	Val	Leu	Pro	Ile	Thr	Ser	Ala	Tyr	Ile	Glu	Leu	Arg	Pro	Cys	
				100				105				110				
Glu	Ala	Ser	Cys	Leu	Pro	Ala	Leu	His	Asn	Met	Leu	Pro	Asp	Asn	Ile	
				115				120				125				
Thr	Val	Tyr	Phe	Ser	Phe	Leu	His	Pro	Thr	Gln	Val	Asp	Gln	Trp	Lys	
				130				135				140				
Gln	Ser	Ala	Arg	Ala	Leu	Cys	Ala									
145				150												
<210>	82															
<211>	152															
<212>	PRT															
<213>	Artificial Sequence															
<220><223>	APG30126_cytidine deaminase															
<400>	82															
Met	Pro	Val	Leu	Ile	Lys	Val	Pro	Tyr	Asp	Ile	Asn	Ser	Ala	Asn	Gly	
1				5				10				15				
Val	Val	Gln	Ala	Cys	Leu	Arg	Lys	Lys	Arg	Glu	Val	Val	Gln	Ser	Lys	
				20				25				30				
Asp	Asn	Gly	Gly	Ile	Thr	Gly	Ile	Gly	Ala	Gly	Ser	Cys	Cys	Ser	Phe	
35				40				45								
Val	Val	Tyr	Met	Lys	His	Gly	Gly	Asp	Val	Asp	Asn	Val	Phe	Gly	Asn	
50				55				60								
Ser	Arg	Ile	Arg	Ile	Pro	Phe	Lys	Val	Asn	Gly	Ile	Glu	Ile	Ala	Asn	
65				70				75				80				
Ala	Cys	Ala	His	Gly	Glu	Leu	Thr	Ala	Leu	Trp	Asn	Ala	Ile	Ala	Asp	
				85				90				95				
Glu	Pro	Ser	Ile	Pro	Thr	Ile	Leu	Ala	Met	Tyr	Ile	Glu	Met	Ser	Pro	
100				105				110								
Cys	Thr	Lys	Cys	Gln	Ser	Ala	Leu	Asp	Asn	Leu	Leu	Gln	Pro	Gly	Gln	

115 120 125
 Glu Ile Tyr Tyr Ser Phe Asp His Pro Gly Glu Val Lys Ala Trp Gln
 130 135 140
 Ala Ala Ala Lys His Leu Cys Ala
 145 150
 <210> 83
 <211> 148
 <212> PRT
 <213> Artificial Sequence

<220><223> APG30127_cytidine deaminsae
 <400> 83
 Met Ile Lys Thr Glu Tyr Ser Leu Ser Ser Gly Tyr Pro Ile Val Arg
 1 5 10 15
 Arg Thr Leu Glu Asp Lys Lys Asn Leu Val Lys Gln Pro Gly Phe Gly
 20 25 30
 Pro Glu Ser Cys Cys Ala Val Val Gln Tyr Arg Leu Arg Gly Asn Ile
 35 40 45
 Arg Tyr Ala Phe Gly Asn Ser Arg Met Gln Ile Ser Met Pro Pro Gly

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

115 120 125
 Ser Phe Asp His Pro Thr Gln Val Asp Glu Trp Lys Leu Arg Ala Asn
 130 135 140
 Glu Leu Cys Ser
 145
 <210> 84

<211> 147

<212> PRT

<213> Artificial Sequence

<220><223> APG30128_cytidine deaminase

<400> 84

Met Ile Lys Thr Glu Tyr Thr Leu Arg Ser Gly Tyr Pro Ile Val Arg

1 5 10 15

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

20 25 30

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

35 40 45

Arg Tyr Ala Phe Gly Asn Ser Arg Met Gln Val Ser Val Pro Pro Gly

50 55 60

Ile Tyr Thr Asn Asn Trp Val Arg Leu His Gly Glu Met Ala Ala Leu

65 70 75 80

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

85 90 95

Pro Ile Thr Ala Ala Tyr Ile Glu Leu Arg Pro Cys Glu Ala Asn Cys

100 105 110

Met Gln Ala Leu His Asn Ile Leu Pro Glu Asn Ala Asn Val Tyr Tyr

115 120 125

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

130 135 140

Glu Leu Cys

145

<210> 85

<211> 147

<212> PRT

<213> Artificial Sequence

<220><223> APG30129_cytidine deaminase

<400> 85

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

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

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

 115 120 125
 Ser Phe Asp His Pro Thr Gln Leu Asp Glu Trp Lys Leu Arg Ala His
 130 135 140
 Glu Leu Cys
 145
 <210> 86
 <211> 147
 <212> PRT
 <213> Artificial Sequence
 <220><223> APG30130_cytidine deaminase
 <400> 86
 Met Ile Gln Leu Val Gln Ala Glu Tyr Ser Ile Lys Ser Gly Tyr Pro
 1 5 10 15
 Ile Val Arg Arg Thr Leu Glu Asp Lys Lys Lys Leu Ile Glu Lys Pro

 20 25 30
 Gly Phe Gly Pro Glu Ser Cys Cys Ala Thr Ile Glu Tyr Gln Leu Arg
 35 40 45
 Gly Ser Thr Arg Tyr Ala Phe Gly Asn Ser Gln Met Lys Met Glu Met

50 55 60
 Pro Pro Asp Ile Tyr Thr His Asn Trp Val Lys Leu His Ala Glu Met
 65 70 75 80
 Ala Ala Leu Val Ala Ala Ile Arg Arg Ile Glu Arg Phe Asp Ala Asp

 85 90 95
 Lys Glu Gln Val Pro Ile Thr Asn Val Tyr Ile Glu Leu Arg Pro Cys
 100 105 110
 Glu Ala Asn Cys Met Gln Ala Leu Gln Asn Ile Leu Pro Asp Gly Thr
 115 120 125
 Thr Val Tyr Tyr Ser Phe Leu His Pro Thr Glu Val Glu Glu Trp Lys
 130 135 140
 Arg Ser Ala
 145
 <210> 87
 <211> 235

<212> PRT
 <213> Artificial Sequence
 <220><223> APG05241_cytidine deaminase
 <400> 87

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

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

100 105 110
Ile Phe Ile Ser Arg Leu Tyr Trp His Met Asp Gln Gln His Arg Gln

115 120 125
Gly Leu Lys Glu Leu Val His Ser Gly Val Thr Ile Gln Ile Met Ser

130 135 140
Tyr Ser Glu Tyr His Tyr Cys Trp Arg Asn Phe Val Asp Tyr Pro Gln

145 150 155 160
Gly Glu Glu Asp Tyr Trp Pro Lys Tyr Pro Tyr Leu Trp Ile Met Leu

165 170 175
Tyr Val Leu Glu Leu His Cys Ile Ile Leu Gly Leu Pro Pro Cys Leu

180 185 190
Lys Ile Ser Gly Ser His Ser Asn Gln Leu Ala Leu Phe Ser Leu Asp

195 200 205
Leu Gln Asp Cys His Tyr Gln Lys Ile Pro Tyr Asn Val Leu Val Ala

210 215 220
Thr Gly Leu Val Gln Pro Phe Val Thr Trp Arg

225 230 235

<210> 88

<211> 262

<212> PRT

<213> Artificial Sequence

<220><223> APG07280_cytidine deaminase

<400> 88

Met His His Ser Ala Arg Leu Pro Pro Asn Cys Ile Val Ser Arg Tyr

1 5 10 15

Ala Asn Ala Pro Trp Thr Val Leu Pro Leu Pro Leu Pro Pro Thr Glu

20 25 30

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

35 40 45

Glu Phe Glu Ala Phe Phe Asn Pro Gln Glu Leu Arg Arg Glu Ala Cys

50 55 60

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

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

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

Phe Val Phe Pro Trp Lys
260

<210> 89

<211> 211

<212> PRT

<213> Artificial Sequence

<220><223> APG09260_cytidine deaminase

<400> 89

Met Glu Asp Phe His Cys Val Cys Val Tyr Cys Pro Leu Gln Val Ser
1 5 10 15
Val Met Ser Asp Tyr Gln Ala Ala Ser Ser Ser Pro Asp Cys Pro Gln
20 25 30
Cys Pro Gln His Ile Leu Thr Gly Ala Glu Ala Ala Val Ser Tyr Ser
35 40 45
Gln Phe Cys Glu Thr Phe Gly Phe His Met Gly Pro Ser Gly Ala Arg
50 55 60
Ala Leu Leu Leu Phe Tyr Glu Leu Trp Gly Pro Ser Gly Thr Leu Val
65 70 75 80
Gln Arg Gly Gln Ala Ser Asn Leu Leu Glu Val Cys Glu Glu Val Leu
85 90 95
Tyr Ser Asn Tyr Met Pro Cys Gln Glu Cys Ser Gln Thr Leu Ile Ser
100 105 110
Phe Leu Leu Arg Tyr Pro Trp Val Arg Leu Asp Leu Leu Phe Ser Gln
115 120 125
Leu Tyr His Thr Ala Pro Ser Gln Thr His Ser Leu Asp Asn Gln Thr
130 135 140
Gly Leu Arg Ser Leu Ala Val Leu Thr Leu Ser Pro Asn Ser Gly Ala
145 150 155 160
Ala Trp Gly His Leu Leu Arg Cys Phe Val Arg Asp Val Pro Pro Ser
165 170 175
Ala Leu Gln Leu Pro Leu Leu Pro Glu Arg Val Glu Ala Asp Arg Val
180 185 190
Asn Ala Ile His Ile Ser Ala Thr Thr Gly Ile Gly Pro Ala Phe Trp
195 200 205
Thr Ser His
210
<210> 90
<211> 234
<212> PRT
<213> Artificial Sequence

<220><223> APG08360_cytidine deaminase

<400> 90

Met Asp Pro Gln Arg Leu Arg Gln Trp Pro Gly Pro Gly Pro Ala Ser

1 5 10 15

Arg Gly Gly Tyr Gly Gln Arg Pro Arg Ile Arg Asn Pro Glu Glu Trp

20 25 30

Phe His Glu Leu Ser Pro Arg Thr Phe Ser Phe His Phe Arg Asn Leu

35 40 45

Arg Phe Ala Ser Gly Arg Asn Arg Ser Tyr Ile Cys Cys Gln Val Glu

50 55 60

Gly Lys Asn Cys Phe Phe Gln Gly Ile Phe Gln Asn Gln Lys Lys Arg

65 70 75 80

His Ala Glu Ile Arg Phe Ile Asp Lys Ile Asn Ser Leu Asn Leu Asp

85 90 95

Gln Asn Gln Cys Tyr Arg Ile Ile Cys Tyr Val Thr Trp Ser Pro Cys

100 105 110

His Asn Cys Ala Lys Glu Leu Val Asp Phe Ile Ser Asn Arg His His

115 120 125

Leu Ser Leu Gln Leu Phe Ala Ser Arg Leu Tyr Phe His Trp Val Arg

130 135 140

Cys Tyr Gln Arg Gly Leu Gln Arg Leu Gln Ala Asn Arg Val Ser Val

145 150 155 160

Ala Val Met Lys Gly Pro Glu Phe Lys Asp Cys Trp Glu Lys Phe Val

165 170 175

Asp His Gln Gly Glu Ser Phe Pro Ser Trp Glu Lys Leu Glu Gln Tyr

180 185 190

Ser Glu Ser Ile Ser Arg Arg Leu Ser Arg Ile Leu Arg Phe Ala Asn

195 200 205

Gln Asn Asn Leu Glu Asp Ser Phe Arg Asp Leu Arg Leu Gly Ser Pro

210 215 220

Ser Pro Ser Ser Ser Arg Ser Asp Ser Arg

225 230

<210> 91

<211> 395

<212> PRT

<213> Artificial Sequence

<220><223> APG07386_cytidine deaminase

<400> 91

Met Asn Pro Gln Ile Arg Asn Pro Met Glu Gly Met Asp Arg His Ala

1 5 10 15

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

20 25 30

Thr Trp Leu Cys Tyr Glu Val Lys Ile Arg Lys Asp Pro Ser Lys Leu

35 40 45

Pro Trp Asp Thr Gly Val Phe Arg Gly Gln Val Arg Pro Lys Leu Gln

50 55 60

Ser Asn Arg Arg Tyr Glu Leu Ser Asn Trp Glu Cys Arg Lys His Val

65 70 75 80

Tyr Phe Gln Pro Gln Tyr His Ala Glu Met Cys Phe Leu Ser Trp Phe

85 90 95

Cys Gly Asn Gln Leu Pro Ala His Lys Arg Phe Gln Ile Thr Trp Phe

100 105 110

Val Ser Trp Thr Pro Cys Pro Asp Cys Val Ala Lys Val Thr Glu Phe

115 120 125

Leu Ala Glu His Pro Asn Val Thr Leu Thr Ile Ser Val Ala Arg Leu

130 135 140

Tyr Tyr Tyr Arg Gly Lys Asp Trp Arg Arg Ala Leu Cys Arg Leu His

145 150 155 160

Gln Ala Gly Ala Arg Val Lys Ile Met Asp Tyr Glu Glu Phe Ala Tyr

165 170 175

Cys Trp Glu Asn Phe Val Tyr Asn Glu Gly Gln Ser Phe Met Pro Trp

180 185 190

Asp Lys Phe Asp Asp Asn Tyr Ala Phe Leu His His Lys Leu Lys Glu

195 200 205
Ile Leu Arg Asn Pro Met Glu Ala Thr Tyr Pro His Ile Phe Tyr Phe

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

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

340 345 350
Met Gly Tyr Lys Asp Phe Lys Tyr Cys Trp Glu Asn Phe Val Tyr Asn
355 360 365
Gly Asp Glu Pro Phe Lys Pro Trp Lys Gly Leu Lys Tyr Asn Phe Leu
370 375 380
Phe Leu Asp Ser Lys Leu Gln Glu Ile Leu Gln
385 390 395

<210> 92

<211> 213

<212> PRT

<213> Artificial Sequence

<220><223> APG07386-NTD_cytidine deaminase

<400> 92

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

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

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

Asp Lys Phe Asp Asp Asn Tyr Ala Phe Leu His His Lys Leu Lys Glu
 195 200 205
 Ile Leu Arg Asn Pro
 210

<210> 93

<211> 267

<212> PRT

<213> Artificial Sequence

<220><223> APG09688_cytidine deaminase

<400> 93

Met Gly Thr Pro Lys Asp Thr Lys Ser Arg His Leu Ala Gly Arg Cys

1 5 10 15

His Glu Leu His Asn Ser Arg Gly Arg Ser Pro Glu Ser Asp Thr Met

20 25 30

Ala Ser Glu Lys Gly Pro Ser Asn Lys Asp Tyr Thr Leu Arg Arg Arg

35 40 45

Ile Glu Pro Trp Glu Phe Glu Val Phe Phe Asp Pro Gln Glu Leu Arg

50 55 60

Lys Glu Ala Cys Leu Leu Tyr Glu Ile Lys Trp Gly Ala Ser Ser Lys

65 70 75 80

Thr Trp Arg Ser Ser Gly Lys Asn Thr Thr Asn His Val Glu Val Asn

85 90 95

Phe Leu Glu Lys Leu Thr Ser Glu Gly Arg Leu Gly Pro Ser Thr Cys

100 105 110

Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Trp Glu Cys Ser

115 120 125

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

130 135 140

Ile Phe Val Ala Arg Leu Phe Gln His Met Asp Arg Arg Asn Arg Gln

145 150 155 160

Gly Leu Lys Asp Leu Val Thr Ser Gly Val Thr Val Gln Val Met Ser

165 170 175

Val Ser Glu Tyr Cys Tyr Cys Trp Glu Asn Phe Val Asn Tyr Pro Pro

180 185 190

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

195 200 205

Tyr Ala Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys Leu

210 215 220

Lys Ile Ser Arg Arg His Gln Lys Gln Leu Thr Phe Phe Ser Leu Thr

225 230 235 240
Pro Gln Tyr Cys His Tyr Lys Met Ile Pro Pro Tyr Ile Leu Leu Ala

 245 250 255
Thr Gly Leu Leu Gln Pro Ser Val Pro Trp Arg

 260 265

<210> 94

<211> 373

<212> PRT

<213> Artificial Sequence

<220><223> APG02316_cytidine deaminase

<400> 94

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

1 5 10 15
Ile Ser Asn Leu Lys Asp Leu Lys Asn Ala Leu Lys Asn Glu Asn Tyr

 20 25 30
Asn Ile Lys Ala Asn Asp Lys Glu Lys Phe Lys Asp Glu Val Lys Lys

 35 40 45
Val Phe Asn Ile Asn Asp Asp Ile Phe Glu Arg Leu Asp Lys Cys Leu

 50 55 60
Asn Arg Asp Ile Thr Tyr Lys Val Asp Asn Val Glu Asp Phe Ile Asp

65 70 75 80
Tyr Ile Lys Lys Ile Met Ile Phe Glu Asp Lys His Glu Ile Ile Cys

 85 90 95
Glu Lys Leu Lys Arg Ile Lys Lys Leu Tyr Ile Asn Arg Glu Glu Tyr

 100 105 110
Glu Arg Glu Lys Ser Thr Arg Asp Asn Val Glu His Ile Ile Glu Val

 115 120 125
Ile Glu Lys Thr Lys Glu Asn Val Ser Arg Lys Ile Ser Leu Glu Glu

 130 135 140
Leu Glu Arg Leu Glu Ile Leu Glu Glu Glu Leu Glu Asp Lys Tyr Leu

145 150 155 160
Phe Ala Lys Asp Ile Glu Phe Leu Lys Lys Met Ile Leu Gly Asn Cys

165 170 175
Lys Asn Val Ile Glu Ser Tyr Asn Glu Lys Thr Lys Ile Lys Thr Leu
180 185 190
Lys Met Lys Ile Pro Lys Glu Ile Asn Tyr Ser Tyr Ile Lys Ala Lys

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

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

325 330 335
Leu Tyr Thr Lys Trp Glu Pro Cys Pro Ser Cys Tyr Phe Val Ile Ser
340 345 350
Gln Phe Cys Glu Lys Tyr Pro Asn Ile Lys Val Glu Val Lys Tyr Asn
355 360 365
Lys Lys Tyr Gly Glu

370

<210> 95

<211> 1060

<212> PRT

<213> Bacillus sp.

<223> APG00969_RGN

<400> 95

Met Lys Gln Thr Ser Tyr Ser Leu Gly Leu Asp Ile Gly Ile Ala Ser

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

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

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

195 200 205
Ser Gly Arg Arg Lys Tyr Tyr Glu Gly Pro Gly Gly Glu Ser Lys Tyr
210 215 220
Gly Trp Lys Gly Asp Ile Glu Lys Trp Tyr Glu Gly Leu Met Gly Lys
225 230 235 240

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

485 490 495
 Ser Leu Asn Glu Lys Asn Lys Gln Ile Asn Asp Glu Val Ile Glu Lys
 500 505 510
 Leu Asn Ala Ser Asn His Arg Asp Asn Lys Gly Met Phe Asn Lys Val

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

 580 585 590
 Tyr Gln Tyr Leu Thr Ser Ala Asp Lys Thr Ile Thr Tyr Glu Lys Phe
 595 600 605
 Lys Ala Asn Ile Thr Gln Leu Ala Lys Ser Asn His Lys Ile Ser Lys
 610 615 620
 Lys Lys Leu Asp Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe His
 625 630 635 640
 Ile Lys Lys Glu Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala

 645 650 655
 Thr Arg Ser Leu Ile Asn Leu Leu Lys Tyr Tyr Phe Ser Glu Lys Asp
 660 665 670
 Ile Asn Val Lys Val Lys Ser Ile Asn Gly Ser Phe Thr Asp Tyr Leu
 675 680 685
 Arg Lys Leu Trp Asn Phe Pro Lys Asp Arg Glu Phe Tyr His Lys His
 690 695 700
 His Ala Glu Asp Ala Leu Ile Ile Ala Met Ala Asn Lys Ile Phe Thr

 705 710 715 720
 Thr Arg Lys Ile Phe Lys Glu Gln Asn Ser Val Phe Ser Asp Glu Gln
 725 730 735

Ile Leu Asp Gly Glu Val Thr Asn Ile Leu Ser Asp Asp Gln Phe Gln
740 745 750

Ala Glu Phe Thr Glu Lys Phe Tyr Lys Val Gln Ala Ile Lys Lys Tyr
755 760 765

Asp Lys Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Gln
770 775 780

Leu Phe Asp Asp Thr Leu Tyr Ser Thr Arg Glu Phe Glu Gly Glu Glu
785 790 795 800

Tyr Tyr Ile Gly Lys Ile Lys Asp Ile Tyr Asn Leu Lys Asp Lys Arg
805 810 815

Leu Lys Lys Ile Phe Thr Lys Ser Pro Glu Lys Ile Leu Met Tyr Gln
820 825 830

His Asp Ser Gln Thr Phe Lys Lys Leu Lys Gln Ile Met Arg Ser Tyr
835 840 845

Glu Asp Glu Val Asn Pro Leu Ala Lys Tyr His Lys Glu Thr Gly Glu
850 855 860

Tyr Leu Arg Lys Glu Cys Lys Lys Gly Asn Gly Pro Ile Val Lys Ser
865 870 875 880

Leu Lys Tyr Arg Val Thr Lys Leu Gly Val His Lys Asp Ile Thr His
885 890 895

Lys Tyr Glu Asn Ser Lys Asn Lys Val Val Ile Leu Ser Leu Lys Pro
900 905 910

Phe Arg Met Asp Val Phe Lys Glu Asn Gly Val Tyr Lys Phe Ile Thr
915 920 925

Ile Arg Tyr Cys Asp Leu Lys Glu Thr Val Asn Ser Tyr Thr Ile Ser
930 935 940

Glu His Leu Tyr Lys Ala Lys Leu Lys Ala Lys Asp Ile Lys Ser Met
945 950 955 960

Asp Ser Phe Lys Trp Ser Phe Tyr Lys Asn Asp Leu Leu Glu Tyr Asn
965 970 975

Gly Glu Leu Cys Thr Phe Lys Gly Val Asn Asp Asp Lys Lys Asn Lys

980 985 990
 Ile Glu Val Asn Trp Val Glu Lys Asn Phe Ala Ile Tyr Ala Glu Lys
 995 1000 1005
 Lys Asn Leu Lys Ser Lys Gln Leu Val Lys Ser Ile Thr Lys Ser
 1010 1015 1020
 Thr Val Lys Ser Leu Leu Lys Tyr Thr Thr Asp Ile Leu Gly Asn

1025 1030 1035
 Arg Tyr Pro Val Arg Asn Glu Lys Leu Lys Leu Met Ile Arg Lys
 1040 1045 1050
 Gln Thr Phe Arg Gly Asp Leu
 1055 1060

<210> 96
 <211> 1090
 <212> PRT
 <213> Rhizobium sp.
 <223> APG03128_RGN
 <400> 96

Met Asn Asn Leu Leu Tyr Ser Phe Asp Ile Gly Thr Asn Ser Ile Gly
 1 5 10 15
 Trp Cys Val Phe Ala Leu Asp Glu Val Gly Asp Pro Cys Arg Ile Val

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

85 90 95
 Arg Lys Thr Gly Asp Arg Glu Gly Gly Asp Glu Gly Phe Asn Pro Tyr
 100 105 110
 Ala Leu Arg Ala Arg Ala Leu Gly Glu Lys Leu Pro Pro Phe Tyr Ile

115 120 125
 Gly Arg Ala Leu Phe His Leu Gly Gln Arg Arg Gly Phe Lys Ser Asn
 130 135 140
 Arg Lys Thr Asp Arg Lys Asp Asn Asp Lys Gly Lys Ile Ala Leu Gly

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

 210 215 220
 Ala Val His His Pro Gln Leu Leu Thr Ser Glu Arg Arg Ala His Leu
 225 230 235 240
 Phe Arg Val Met Phe Tyr Gln Arg Pro Leu Lys Lys Pro Leu Val Gly
 245 250 255
 Arg Cys Ser Phe Asn Pro Ala Glu Ala Arg Leu Ser Arg Ala His Pro
 260 265 270
 Leu Phe Gln Glu Phe Arg Leu Tyr Lys Glu Val Asn Asp Leu Glu Val

 275 280 285
 Val Leu Pro Asp Gln Ser His Arg Lys Leu Thr Leu Asp Glu Arg Asn
 290 295 300
 Ala Leu Val Ala Lys Leu Lys Ser Ser Arg Lys Ala Ser Phe Ser Val
 305 310 315 320
 Leu Arg Arg Thr Leu Lys Leu Thr Pro Asp Leu Ala Phe Asn Lys Glu
 325 330 335
 Ser Glu Ala Arg Lys Asp Leu Leu Gly Asp Glu Ile Asn Ser Ala Leu

 340 345 350
 Ala Asp Ala Lys Met Phe Gly Val Arg Trp Gly Gly Phe Pro Arg Ala
 355 360 365

Arg Gln Trp Glu Ile Ile Thr His Leu Lys Glu Glu Glu Asn Pro Ala
370 375 380

Arg Leu Ser Asp Trp Leu Lys Ser Glu Phe Gly Leu Asp Asp Glu Arg
385 390 395 400

Val Val Ala Ile Ala Asn Ile Ala Leu Pro Glu Gly Tyr Gly Arg Leu
405 410 415

Gly Glu Thr Ala Leu Ala Ser Met Leu Glu Glu Met Lys Thr Ala Val
420 425 430

Ile Pro Glu Ser Glu Ala Ala Lys Arg Cys Gly Tyr Asp His Ser Asn
435 440 445

Leu Ala Lys Glu Gln Asp Glu Gly Leu Asp Ile Leu Pro Glu Tyr Gln
450 455 460

Glu Ile Leu Glu Arg His Ile Pro Pro Gly Thr Gly Asp Pro Asp Asp
465 470 475 480

Ile Tyr Asp Ile Arg Lys Gly Arg Ile Thr Asn Pro Thr Val His Ile
485 490 495

Gly Leu Asn Gln Leu Arg Arg Val Val Asn Ala Leu Ile Lys Arg His
500 505 510

Gly Lys Pro Arg His Ile Val Val Glu Leu Ala Arg Asp Leu Gln Leu
515 520 525

Ser Glu Lys Gln Lys Ala Asp Val Asn Arg Arg Ile Ala Lys Asn Thr
530 535 540

Arg Glu Ala Glu Gly Arg Ser Gln Lys Leu Ile Glu Met Gly Gln Leu
545 550 555 560

Asp Thr Gly Tyr Asn Arg Leu Leu Leu Lys Leu Trp Glu Glu Leu Asn
565 570 575

Gln Asp Lys Pro Glu Asp Arg Val Cys Ile Tyr Ser Gly Lys Pro Ile
580 585 590

Gly Ile Ala Met Leu Phe Ser Gly Glu Val Asp Ile Asp His Ile Leu
595 600 605

Pro Trp Ser Lys Thr Leu Asp Asp Ser Gln Ala Asn Lys Leu Leu Cys

610 615 620
 Leu Lys Ser Ala Asn Arg Gln Lys Arg Asn Arg Thr Pro Ser Asp Val
 625 630 635 640
 Pro Glu Trp Arg Asp Cys Tyr Glu Glu Val Leu Ala Arg Ala Ala Arg
 645 650 655
 Leu Pro Lys Asn Lys Arg Trp Arg Phe Ala Ala Asp Ala Met Gln Gln

 660 665 670
 Phe Glu Ala Glu Gly Gly Phe Leu Ala Arg Gln Leu Thr Asp Thr Gln
 675 680 685
 Tyr Leu Ser Arg Met Ala Phe Glu Tyr Leu Ser Ala Leu Phe Pro Ser
 690 695 700
 Glu Glu Ala Asp Lys Trp Gly Glu Leu Arg Gln Arg Lys Arg Val His
 705 710 715 720
 Val Val Pro Gly Arg Leu Thr Glu Leu Leu Arg Arg Asn Trp Gly Leu

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

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

 850 855 860

Val Val Arg Arg Lys Pro Phe Met Ala Leu Glu Val Lys Asp Ile Ala
865 870 875 880
Ser Ile Arg Asp Thr Glu Leu Gln Ser Ala Leu Tyr Asp Ala Ile Gly
885 890 895
Ser Leu Thr Glu Lys Lys Ala Leu Gln Glu Ala Leu Val Arg Phe Arg
900 905 910
Asp Arg His Pro Gln Phe Lys Gly Ile Arg Arg Val Arg Val Leu Glu
915 920 925
Ala Leu Ser Val Ile Pro Ile Val Asp Asn Asn Gly Lys Ala Tyr Lys
930 935 940
Gly Tyr Lys Gly Asp Ala Asn Tyr Arg Tyr Glu Val Trp Glu Thr Leu
945 950 955 960
Asp Gly Arg Trp His Thr Glu Val Val Ser Met Phe Asp Ala His Gln
965 970 975
Pro Gly Trp Gln Ser Pro Phe His Arg Gln His Pro Ala Ala Arg Arg
980 985 990
Val Leu Lys Leu Gln Gln Asn Asp Met Val Ala Tyr Glu His Pro Ala
995 1000 1005
Asp Gly Tyr Thr Ile Ala Arg Val Val Lys Phe Ser Asn Asp Lys
1010 1015 1020
Arg Ile Tyr Phe Ala Ser His Arg Glu Ser Gly Ser Leu Lys Ala
1025 1030 1035
Arg Asp Ala Asp Lys Gly Asp Pro Phe Thr Tyr Phe Ala Lys Ala
1040 1045 1050
Asn Asn Gly Leu Arg Asp Ile Lys Cys Arg Gln Val Arg Ile Asp
1055 1060 1065
Ala Ala Gly Arg Val Phe Asp Pro Gly Pro Gln Asp Arg Glu Ala
1070 1075 1080
Arg Gln Ser Lys Gly Arg Ala
1085 1090
<210> 97
<211> 1108

<212> PRT

<213> Brevibacillus sp.

<223> APG09748_RGN

<400> 97

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

1 5 10 15

Glu Ala Gln Asn Leu Arg Lys Gly Ile Trp Arg Thr His Arg Leu Leu

20 25 30

Asn Glu Gly Val Ala Tyr Tyr Met Lys Met Leu Leu Leu Phe Arg Gln

35 40 45

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

50 55 60

Arg His Ile Arg Glu Gln Gln Gln Lys Asn Arg Ala Asp Lys Asn Thr

65 70 75 80

Gln Ala Leu Pro Leu Asp Lys Ala Phe Ala Ala Leu Arg Gln Leu Tyr

85 90 95

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

100 105 110

Ile Ser Arg Lys Phe Leu Ser Pro Leu Val Asp Pro Asn Ser Glu Gly

115 120 125

Gly Lys Gly Thr Ser Lys Ala Gly Ala Lys Pro Thr Trp Gln Lys Lys

130 135 140

Lys Glu Ala Asn Asp Pro Thr Trp Glu Gln Asp Tyr Glu Lys Trp Lys

145 150 155 160

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

165 170 175

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

180 185 190

Asp Ile Ala Trp Leu Pro Leu Gln Ser Asn Gln Phe Val Arg Thr Trp

195 200 205

Asp Arg Asp Met Leu Gln Gln Ala Ile Glu Arg Leu Leu Ser Trp Glu

210 215 220
 Ser Trp Asn Lys Arg Val Gln Glu Glu Tyr Ser Lys Leu Gln Glu Lys
 225 230 235 240
 Met Thr Gln Leu Asn Glu Gln Leu Glu Gly Gly Gln Glu Trp Ile Ser
 245 250 255
 Leu Leu Glu Gln Tyr Glu Glu Gln Arg Glu Gln Glu Leu Ile Glu Asn

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

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

 385 390 395 400
 Ala Asp Asn Pro Arg Ser Arg Arg Phe Val Thr Phe Ser Gln Leu Ile
 405 410 415
 Trp Pro Asn Glu Ser Gly Trp Met Glu Lys Gln Asp Val Glu Val Glu
 420 425 430
 Leu Ala Leu Ser Lys Gln Phe Tyr Gln Gln Val Thr Leu Gln Lys Asn
 435 440 445
 Asp Lys Gly Lys Gln Glu Ile Glu Phe Lys Asp Lys Gly Ser Gly Ser

 450 455 460

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

705 710 715 720
 Gly Lys Val Val Gln Ala Trp Arg Lys Arg Phe Asp Ala Asp Glu Arg
 725 730 735
 Lys Gly Ile Ala Gly Leu Ser Met Trp Ser Ile Glu Glu Leu Asp Ser
 740 745 750
 Leu Arg Lys Leu Leu Ile Ser Trp Ser Arg Arg Thr Arg Asn Pro Arg
 755 760 765
 Glu Ile Asn Cys Phe Glu Gln Gly His Thr Ser His Gln Arg Leu Leu

 770 775 780
 Thr His Ile Gln Asn Val Lys Glu Asp Arg Leu Lys Gln Leu Ser His
 785 790 795 800
 Ala Ile Val Met Thr Ala Leu Gly Tyr Val Tyr Asp Glu Lys Lys Leu
 805 810 815
 Glu Trp Phe Ala Lys Tyr Pro Ala Cys Gln Val Ile Leu Phe Glu Asn
 820 825 830
 Leu Ser Gln Tyr Arg Ser Asn Met Asp Arg Ser Thr Lys Glu Asn Ser

 835 840 845
 Thr Leu Met Lys Trp Ala His Arg Ser Ile Pro Lys Tyr Val His Met
 850 855 860
 Gln Ala Glu Pro Tyr Gly Ile Gln Ile Gly Asp Val Arg Ala Glu Tyr
 865 870 875 880
 Ser Ser Arg Phe His Ala Lys Thr Gly Thr Pro Gly Ile Arg Cys Lys
 885 890 895
 Met Val Ser Gly His Asp Leu Gln Gly Arg Arg Phe Glu Asn Leu Gln

 900 905 910
 Lys Arg Leu Ile Ser Glu Gln Phe Leu Thr Glu Glu Gln Val Lys Gln
 915 920 925
 Leu Arg Pro Gly Asp Ile Val Pro Asp Asp Ser Gly Glu Trp Phe Met
 930 935 940
 Thr Leu Ser Asp Gly Ser Glu Gly Lys Glu Val Val Phe Leu Gln Ala
 945 950 955 960

Asp Ile Asn Ala Ala Gln Asn Leu Gln Lys Arg Phe Trp Gln Arg Tyr

965

970

975

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

980

985

990

Glu Tyr Leu Ile Pro Lys Thr Lys Ser Val Gln Ala Lys Leu Gly Lys

995

1000

1005

Gly Leu Phe Val Lys Lys Thr Asp Thr Val Met Lys Asp Val Tyr

1010

1015

1020

Val Trp Asp Ser Gln Ala Lys Leu Lys Gly Lys Thr Thr Phe Thr

1025

1030

1035

Glu Glu Ser Glu Ser Pro Glu Gln Leu Glu Asp Phe Gln Glu Ile

1040

1045

1050

Ile Glu Glu Ala Glu Glu Ala Lys Gly Thr Tyr Arg Thr Leu Phe

1055

1060

1065

Arg Asp Pro Ser Gly Val Phe Phe Pro Glu Phe Val Trp Ser Thr

1070

1075

1080

Gln Lys Asp Phe Trp Ser Glu Val Lys Arg Arg Leu Tyr Gly Lys

1085

1090

1095

Leu Arg Glu Arg Phe Leu Met Lys Thr Arg

1100

1105

<210> 98

<211> 1429

<212> PRT

<213> Chryseobacterium sp.

<223> APG00771_RGN

<400> 98

Met Thr Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly

1

5

10

15

Trp Ala Leu Ile Lys Gln Asp Phe Glu Asn Lys Gln Gly Glu Ile Leu

20

25

30

Gly Met Gly Ser Arg Ile Ile Pro Met Asp Ala Gly Asp Met Gly Lys

35	40	45
Phe Ala Glu Gly Ala Thr Val Ser Lys Thr Ala Asp Arg Thr Lys Tyr		
50	55	60
Arg Ser Val Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu Arg		
65	70	75
80		
Leu His Arg Val Leu His Leu Leu Asn Phe Leu Pro Gln His Tyr Ala		
85	90	95
Ser Gln Ile Asp Phe Glu Lys Arg Phe Gly Lys Phe Lys Val Glu Thr		
100	105	110
Glu Pro Lys Leu Ala Trp Lys Asn Ile Glu Gly Lys Phe Ser Phe Leu		
115	120	125
Phe Gln Asn Ser Phe Asn Glu Met Leu Glu Asp Phe Lys Ala Asn Glu		
130	135	140
Gln Asp Leu Lys Ile Pro Tyr Asp Trp Thr Ile Tyr Phe Leu Arg Lys		
145	150	155
160		
Lys Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp Ile Leu		
165	170	175
Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu Glu		
180	185	190
Glu Glu Glu Asn Pro Asn Lys Leu Val Glu Phe Tyr Ser Leu Lys Val		
195	200	205
Val Asp Val Leu Ala Asp Glu Pro Gln Lys Gly Lys Ser Glu Ile Trp		
210	215	220
Tyr Ser Leu Val Leu Lys Asn Gly Trp Ile Tyr Arg Arg Ala Ser Lys		
225	230	235
240		
Ile Pro Leu Phe Glu Trp Lys Asp Lys Thr Arg Asp Phe Ile Val Thr		
245	250	255
Thr Asp Leu Asn Asp Asp Gly Ser Ile Lys Lys Asp Lys Asp Gly Asn		
260	265	270
Glu Lys Arg Ser Phe Arg Ala Pro Gly Glu Asn Asp Trp Thr Leu Val		
275	280	285

Lys Lys Lys Thr Glu Gln Glu Ile Glu Gln Ser His Lys Thr Val Gly
 290 295 300
 Thr Tyr Ile Tyr Glu Thr Leu Leu Leu Asn Pro Lys Gln Lys Ile Lys
 305 310 315 320
 Gly Lys Leu Val Arg Thr Ile Glu Arg Lys Phe Tyr Lys Asp Glu Leu
 325 330 335
 Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe His Gln Glu Leu Gln Asn
 340 345 350

 Asp Asp Leu Tyr Asn Asp Cys Ile Arg Glu Leu Tyr Arg Asn Asn Glu
 355 360 365
 Ala His Gln Leu Thr Leu Ser Lys Lys Asp Phe Val His Leu Leu Met
 370 375 380
 Asp Asp Leu Ile Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser Ser
 385 390 395 400
 Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr Lys Asp Glu Asn Gly
 405 410 415

 Ile Glu His Thr Gln Tyr Leu Lys Ala Ile Pro Lys Ser Asn Pro Tyr
 420 425 430
 Tyr Gln Glu Phe Arg Leu Trp Gln Trp Ile Tyr Asn Leu Asn Ile Tyr
 435 440 445
 Arg Lys Asp Asp Glu Val Asn Val Thr Lys Asp Phe Leu Asn Thr Thr
 450 455 460
 Lys Glu Phe Glu Asn Leu Phe Glu Phe Leu Asn Ser Arg Lys Glu Val
 465 470 475 480

 Asp Gln Lys Ala Leu Leu Lys Tyr Phe Lys Leu Asn Glu Lys Thr His
 485 490 495
 Arg Trp Asn Phe Val Glu Asp Lys Lys Tyr Pro Cys Asn Glu Thr Lys
 500 505 510
 Thr Met Ile Ser Ser Arg Leu Asp Lys Val Glu Asn Ile Ser Asp Asp
 515 520 525
 Phe Leu Thr Arg Asn Ile Glu Gln Lys Ile Trp His Ile Ile Tyr Ser

530 535 540
 Val Asn Asp Lys Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Arg
 545 550 555 560
 Lys His His Leu Asp Glu Thr Ser Phe Phe Glu Ala Phe Arg Lys Phe
 565 570 575
 Pro Pro Phe Lys Ser Glu Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys
 580 585 590
 Lys Leu Leu Pro Leu Met Arg Leu Gly Lys Tyr Trp Asn Tyr Ala Asp
 595 600 605

 Ile Asp Lys Tyr Ser Arg Glu Arg Ile Gln Lys Ile Ile Thr Gly Glu
 610 615 620
 Tyr Asp Glu Asn Ile Lys Asp Lys Ile Arg Glu Lys Ala Ile His Leu
 625 630 635 640
 Thr Lys Glu Asn Asp Phe Gln Gly Leu Gln Leu Trp Leu Ala Gln Tyr
 645 650 655
 Leu Val Tyr Gly Arg His Ser Glu Ala Ser Met Ile Gly Lys Trp Asn
 660 665 670

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

 Asp Arg Lys Lys Leu Thr Ser Gln Ile Ser Glu Asn Glu Asn Thr Asn
 740 745 750
 Leu Arg Ile Lys Ala Leu Leu Ala Glu Met Ile Asn Asp His Ser Val
 755 760 765
 Glu Asn Val Arg Pro Phe Ser Pro Met Gln Gln Glu Ile Leu Lys Ile
 770 775 780

Tyr Glu Asp Gly Val Leu Asn Ser Asp Ile Glu Ile Glu Asp Asp Ile
785 790 795 800

Leu Lys Ile Ser Lys Thr Ala Gln Pro Ser Pro Ser Asp Leu Lys Arg
805 810 815

Tyr Lys Leu Trp Leu Glu Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln
820 825 830

Ile Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His
835 840 845

Ile Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser Phe Ser Asn Lys Val
850 855 860

Ile Cys Glu Ser Ala Val Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu
865 870 875 880

Glu Phe Ile Lys Lys Leu Gly Gly Thr Val Val Glu Cys Gly Asn Gly
885 890 895

Lys Ser Val Thr Val Leu Lys Glu Asp Gln Tyr Glu Asp Phe Val Lys
900 905 910

Lys His Tyr Ala Asn Asn Arg Gly Lys Arg Asn Lys Leu Leu Met Glu
915 920 925

Asp Ile Pro Glu Lys Met Ile Glu Arg Gln Leu Asn Asp Thr Arg Tyr
930 935 940

Ile Ser Lys Tyr Ile Ser Gly Val Leu Ser Asn Ile Val Arg Val Glu
945 950 955 960

Asp Gly Ser Asp Glu Gly Ile Asn Ser Lys Asn Ile Val Pro Gly Asn
965 970 975

Gly Lys Ile Thr Ser Gln Leu Lys Gln Asp Trp Gly Leu Asn Asp Val
980 985 990

Trp Asn Asp Leu Ile Leu Pro Arg Phe Glu Arg Met Asn Gln Leu Thr
995 1000 1005

Asn Ser Thr Asp Phe Thr Ala Trp Asn Glu Asn Tyr Gln Lys Phe
1010 1015 1020

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

1025	1030	1035
Arg Ile Asp His Arg His His	Ala Leu Asp Ala Leu	Val Ile Ala
1040	1045	1050
Cys Ala Thr Lys Asp His Val	Asn Leu Leu Asn Asn	Gln Ser Ala
1055	1060	1065
Lys Ser Asp Thr Lys Arg Tyr	Asp Leu Lys Lys Lys	Leu Met Lys
1070	1075	1080
Phe Glu Lys Val Val Tyr His	His Thr Gln Thr Gly	Glu Lys Ile
1085	1090	1095
Glu Arg Glu Ile Pro Arg Gln	Phe Leu Lys Pro Trp	Glu Asn Phe
1100	1105	1110
Thr Val Asp Ala Lys His Asn	Leu Glu Ser Ile Ile	Val Ser Phe
1115	1120	1125
Lys Gln Asn Leu Arg Val Ile	Asn Lys Ala Thr Asn	Tyr Tyr Glu
1130	1135	1140
Lys Tyr Val Glu Lys Asp Gly	Thr Lys Asn Lys Glu	Arg Val Glu
1145	1150	1155
Gln Thr Gly Thr Asn Trp Ala	Ile Arg Lys Pro Met	His Lys Asp
1160	1165	1170
Thr Val Ser Gly Lys Val Asp	Leu Pro Trp Val Lys	Val Pro Lys
1175	1180	1185
Gly Lys Ile Leu Thr Ala Thr	Arg Lys Ser Leu Asp	Thr Ser Phe
1190	1195	1200
Asp Leu Lys Ser Ile Gly Ser	Ile Thr Asp Thr Gly	Ile Gln Lys
1205	1210	1215
Ile Leu Lys Asn Tyr Leu Ala	Phe Lys Asp Gly Asn	Pro Glu Met
1220	1225	1230
Ala Phe Ser Pro Glu Gly Ile	Asp Asp Leu Asn Lys	Asn Ile Glu
1235	1240	1245
Lys Tyr Asn Asp Gly Lys Pro	His Gln Pro Ile Asn	Lys Val Arg
1250	1255	1260

Val Phe Glu Leu Gly Ser Lys Phe Gln Val Gly Gln Thr Gly Asn
1265 1270 1275

Lys Lys Asp Lys Tyr Val Glu Ala Ala Lys Gly Thr Asn Leu Phe
1280 1285 1290

Phe Ala Val Tyr Glu Asp Glu Lys Gly Lys Arg Ser Tyr Glu Thr
1295 1300 1305

Ile Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu Ser
1310 1315 1320

Val Val Asp Leu Lys Asp Thr Asn Asp Phe Tyr Leu Cys Pro Asn
1325 1330 1335

Asp Leu Val Tyr Ile Pro Ser Gly Asp Glu Leu Glu Asn Val Asn
1340 1345 1350

Asn Val Asp Phe Lys Asp Ile Lys Lys Asp Lys Asn Glu Arg Ile
1355 1360 1365

Tyr Lys Val Val Ser Phe Ser Gly Ser Gln Ile Phe Phe Val Arg
1370 1375 1380

Gln Asp Ile Ala Thr Phe Ile Val Asn Lys Ala Glu Leu Ser Thr
1385 1390 1395

Leu Asn Lys Met Glu Arg Ala Ile Asp Gly Ser Met Ile Lys Glu
1400 1405 1410

Asn Cys Ile Lys Leu Asn Ile Asp Arg Leu Gly Asn Ile Ser Lys
1415 1420 1425

Ala

<210> 99

<211> 1112

<212> PRT

<213> Bacillus sp.

<223> APG02789_RGN

<400> 99

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

1

5

10

15

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

260 265 270
 Phe Ala Pro Asn Asp Gly Tyr Phe Ile Thr Ser Arg Gln Ile Arg Gly
 275 280 285
 Trp Asp Arg Val Tyr Glu Lys Trp Ser Lys Leu Pro Glu Ser Ala Ser

 290 295 300
 Pro Glu Glu Leu Trp Lys Val Val Ala Glu Gln Gln Asn Lys Met Ser
 305 310 315 320
 Glu Gly Phe Gly Asp Pro Lys Val Phe Ser Phe Leu Ala Asn Arg Glu
 325 330 335
 Asn Arg Asp Ile Trp Arg Gly His Ser Glu Arg Ile Tyr His Ile Ala
 340 345 350
 Ala Tyr Asn Gly Leu Gln Lys Lys Leu Ser Arg Thr Lys Glu Gln Ala

 355 360 365
 Thr Phe Thr Leu Pro Asp Ala Ile Glu His Pro Leu Trp Ile Arg Tyr
 370 375 380
 Glu Ser Pro Gly Gly Thr Asn Leu Asn Leu Phe Lys Leu Glu Glu Lys
 385 390 395 400
 Gln Lys Lys Asn Tyr Tyr Val Thr Leu Ser Lys Ile Ile Trp Pro Ser
 405 410 415
 Glu Glu Lys Trp Ile Glu Lys Glu Asn Ile Glu Ile Pro Leu Ala Pro

 420 425 430
 Ser Ile Gln Phe Asn Arg Gln Ile Lys Leu Lys Gln His Val Lys Gly
 435 440 445
 Lys Gln Glu Ile Ser Phe Ser Asp Tyr Ser Ser Arg Ile Ser Leu Asp
 450 455 460
 Gly Val Leu Gly Gly Ser Arg Ile Gln Phe Asn Arg Lys Tyr Ile Lys
 465 470 475 480
 Asn His Lys Glu Leu Leu Gly Glu Gly Asp Ile Gly Pro Val Phe Phe

 485 490 495
 Asn Leu Val Val Asp Val Ala Pro Leu Gln Glu Thr Arg Asn Gly Arg
 500 505 510

Leu Gln Ser Pro Ile Gly Lys Ala Leu Lys Val Ile Ser Ser Asp Phe
 515 520 525
 Ser Lys Val Ile Asp Tyr Lys Pro Lys Glu Leu Met Asp Trp Met Asn
 530 535 540
 Thr Gly Ser Ala Ser Asn Ser Phe Gly Val Ala Ser Leu Leu Glu Gly

 545 550 555 560
 Met Arg Val Met Ser Ile Asp Met Gly Gln Arg Thr Ser Ala Ser Val
 565 570 575
 Ser Ile Phe Glu Val Val Lys Glu Leu Pro Lys Asp Gln Glu Gln Lys
 580 585 590
 Leu Phe Tyr Ser Ile Asn Asp Thr Glu Leu Phe Ala Ile His Lys Arg
 595 600 605
 Ser Phe Leu Leu Asn Leu Pro Gly Glu Val Val Thr Lys Asn Asn Lys

 610 615 620
 Gln Gln Arg Gln Glu Arg Arg Lys Lys Arg Gln Phe Val Arg Ser Gln
 625 630 635 640
 Ile Arg Met Leu Ala Asn Val Leu Arg Leu Glu Thr Lys Lys Thr Pro
 645 650 655
 Asp Glu Arg Lys Lys Ala Ile His Lys Leu Met Glu Ile Val Gln Ser
 660 665 670
 Tyr Asp Ser Trp Thr Ala Ser Gln Lys Glu Val Trp Glu Lys Glu Leu

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

 740 745 750
 Leu Ile Ser Trp Ser Lys Arg Ser Arg Thr Pro Gly Glu Ala Asn Arg

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

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

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

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

 995 1000 1005

Tyr Phe Gly Lys Gly Ser Phe Val Lys Asn Asn Thr Glu Gln Glu
 1010 1015 1020
 Val Tyr Lys Trp Glu Lys Ser Glu Lys Met Lys Ile Lys Thr Asp
 1025 1030 1035
 Thr Thr Phe Asp Leu Gln Asp Leu Asp Gly Phe Glu Asp Ile Ser
 1040 1045 1050
 Lys Thr Ile Glu Leu Ala Gln Glu Gln Gln Lys Lys Tyr Leu Thr

1055 1060 1065
 Met Phe Arg Asp Pro Ser Gly Tyr Phe Phe Asn Asn Glu Thr Trp
 1070 1075 1080
 Arg Pro Gln Lys Glu Tyr Trp Ser Ile Val Asn Asn Ile Ile Lys
 1085 1090 1095
 Ser Cys Leu Lys Lys Lys Ile Leu Ser Asn Lys Val Glu Leu
 1100 1105 1110

<210> 100

<211> 1070

<212> PRT

<213> Bacillus sp.

<223> APG05083.1_RGN

<400> 100

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

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
 65 70 75 80
 Asp Ile Arg Asn Leu Leu Val Gln His Glu Ile Ile Ser Gln Lys Glu
 85 90 95

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

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

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

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

 Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
 325 330 335
 Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu

340 345 350
 Leu Lys Ser Phe Ile Asn Leu Lys Pro Phe Tyr Glu Ile Lys Lys Val
 355 360 365
 Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Ala Phe Ser Thr Leu Asp
 370 375 380

 Tyr Asp Thr Phe Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
 385 390 395 400
 Ile Arg Ser Tyr Leu Lys Lys Ser Asn Asn Leu Ser Lys Cys Cys Tyr
 405 410 415
 Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe
 420 425 430
 Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
 435 440 445

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

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

 Ser Leu Lys Gln Ile Pro Ala Asn Thr Phe Phe Asn Glu Met Lys Lys
 580 585 590

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

 Glu Gly Asn Lys Asp Trp Glu Ser Phe Glu Ser Tyr Val Arg Leu Asn
 645 650 655
 Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr
 660 665 670
 Leu Pro Arg Glu Ser Asn Met Ile Lys Glu Arg His Leu Asn Asp Thr
 675 680 685
 Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys
 690 695 700

 Phe Lys Glu Val Glu Gly Ser Thr Arg Lys Lys His Val Gln Thr Val
 705 710 715 720
 Asn Gly Ile Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
 725 730 735
 Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
 740 745 750
 Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln
 755 760 765

 Ile Lys Glu Ser Asn Lys Ser Ile Arg Lys Pro Tyr Phe Pro Leu Pro
 770 775 780
 Trp Val Gly Phe Arg Glu Glu Ile Leu Ser His Leu Ala Arg Gln Pro
 785 790 795 800
 Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Phe
 805 810 815
 Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ala
 820 825 830

 Ala His Glu Gln Thr Ile Met Lys Lys Gly Gly Ile Asp Lys Lys Gly

835 840 845
 Lys Thr Ile Ile Ile Lys Arg Val Tyr Leu Lys Asp Ile Lys Phe Asp
 850 855 860
 Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
 865 870 875 880
 Tyr Glu Ala Ile Lys Gln Arg Tyr Ile Glu Tyr Gly Lys Glu Ser Lys
 885 890 895

 Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
 900 905 910
 Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
 915 920 925
 Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
 930 935 940
 Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Ile Pro Ile Tyr Val
 945 950 955 960

 Met Asp Thr Val His Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
 965 970 975
 Lys Gly Tyr Glu Gln Trp Leu Thr Ile Asp Asn Ser Phe Thr Phe Lys
 980 985 990
 Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asn Glu
 995 1000 1005
 Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
 1010 1015 1020

 Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Asn
 1025 1030 1035
 Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Val
 1040 1045 1050
 Gly Val Leu Gly Asp Leu Arg Phe Val Arg Gln Glu Ile Arg Lys
 1055 1060 1065
 Asn Phe
 1070
 <210> 101

<211> 1071

<212> PRT

<213> Bacillus sp.

<223> APG07433.1_RGN

<400> 101

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser

1 5 10 15

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

20 25 30

Tyr Glu Lys Val Arg Ile Val Asp Gln Gly Val Arg Met Phe Asp Arg

35 40 45

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

50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys

65 70 75 80

Asn Ile Arg Asn Leu Leu Val Gln His Gly Val Ile Thr Gln Glu Glu

85 90 95

Leu Asp Ser Leu Tyr Pro Leu Ser Lys Lys Ser Met Asp Ile Trp Gly

100 105 110

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asn His Phe Glu Trp Ala

115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg

130 135 140

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

145 150 155 160

Gln Leu Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met

165 170 175

Trp Met Lys Asp Pro Asp Phe Ser Lys Tyr Asp Arg Lys Arg Asn Ser

180 185 190

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

195 200 205

Ile Val Thr Leu Phe Ala Ala Gln Arg Arg Phe Gln Ser Pro Tyr Ala
210 215 220

Ser Lys Asp Leu Gln Glu Thr Tyr Leu Gln Ile Trp Thr His Gln Leu
225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Leu Asn Lys Val Gly Tyr Cys Ser
245 250 255

Leu Leu Lys Gly Lys Glu Arg Arg Ile Pro Lys Ala Thr Tyr Thr Phe
260 265 270

Gln Tyr Phe Ser Ala Leu Asp Gln Val Asn Arg Thr Arg Leu Gly Pro
275 280 285

Asp Phe Gln Pro Phe Thr Lys Glu Gln Arg Glu Ile Ile Leu Asn Asn
290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320

Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln
325 330 335

Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350

Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Asn Lys Val
355 360 365

Val Ala Asn Tyr Ser Glu Arg Thr Asn Glu Thr Phe Ser Thr Leu Asp
370 375 380

Tyr Asp Gly Ile Gly Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400

Ile Arg Ser Tyr Leu Lys Ser Ser His Asn Leu Pro Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Ser Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Leu Lys Ala Ile Asn His Val Leu Ser Ile Met Gln
435 440 445

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

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

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

 Ser Leu Lys Glu Ile Pro Ala Asp Thr Phe Phe Asn Glu Leu Lys Lys
 580 585 590
 Glu Arg Asn Gly Ala Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
 595 600 605
 Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
 610 615 620
 Asp Glu Asn Arg Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
 625 630 635 640

 Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
 645 650 655
 Lys Phe Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Arg Ala Tyr
 660 665 670
 Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
 675 680 685
 Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
 690 695 700

Phe Lys Glu Ala Glu Asp Asn Pro Arg Lys Arg Arg Val Gln Thr Val
 705 710 715 720
 Asn Gly Val Ile Thr Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys
 725 730 735
 Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
 740 745 750
 Ala Cys Thr Asp His His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln
 755 760 765

 Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
 770 775 780
 Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
 785 790 795 800
 Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Leu
 805 810 815
 Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Ala
 820 825 830

 Ala His Lys Gln Thr Ile Met Arg Lys Gly Gly Ile Asp Lys Lys Gly
 835 840 845
 Lys Thr Ile Ile Ile Glu Arg Leu His Leu Lys Asp Ile Lys Phe Asp
 850 855 860
 Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Met Ala Thr
 865 870 875 880
 Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Gly Lys Asn Ser Lys
 885 890 895

 Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
 900 905 910
 Asn Leu Ile Lys Arg Val Lys Val Glu Gly Gln Ala Lys Ser Phe Val
 915 920 925
 Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
 930 935 940
 Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val

945 950 955 960

 Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Ala Ser Ser
 965 970 975
 Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
 980 985 990
 Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
 995 1000 1005
 Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
 1010 1015 1020

 Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Lys Asn Glu Tyr
 1025 1030 1035
 Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val
 1040 1045 1050
 Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg
 1055 1060 1065
 Asn Phe His
 1070
 <210> 102
 <211> 1071
 <212> PRT
 <213> Bacillus sp.
 <223> APG07513.1_RGN

 <400
 > 102
 Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
 1 5 10 15
 Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Gln
 20 25 30
 Tyr Glu Lys Thr Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
 35 40 45
 Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
 50 55 60

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

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

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

Leu Leu Lys Gly Lys Glu Lys Arg Val Pro Lys Ala Thr Tyr Thr Phe
260 265 270
Gln Tyr Phe Asn Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
275 280 285
Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Ile Ile Leu Asp Lys
290 295 300
Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val

305						310						315						320
Thr	Tyr	Tyr	Asp	Ile	Arg	Lys	Trp	Leu	Ala	Leu	Asp	Glu	Thr	Ile	Gln			
					325						330						335	
Phe	Lys	Gly	Leu	Thr	Tyr	Asp	Pro	Asn	Glu	Glu	Leu	Lys	Lys	Ile	Glu			
					340						345						350	
Met	Lys	Pro	Phe	Ile	Asn	Leu	Lys	Pro	Phe	Tyr	Glu	Ile	Lys	Lys	Val			
					355						360						365	
Val	Thr	Asn	Tyr	Ala	Lys	Lys	Thr	Asn	Glu	Val	Phe	Ser	Ala	Leu	Asp			
					370						375						380	
Tyr	Asp	Thr	Val	Ala	Tyr	Ala	Leu	Thr	Val	Tyr	Lys	Thr	Asp	Lys	Asp			
					385						390						395	
Ile	Arg	Ser	Tyr	Leu	Lys	Arg	Ser	Asn	Asn	Leu	Ser	Lys	Arg	Cys	Tyr			
					405						410						415	
Asp	Asp	Gln	Leu	Ile	Glu	Glu	Leu	Leu	Thr	Leu	Ser	Tyr	Thr	Lys	Phe			
					420						425						430	
Gly	His	Leu	Ser	Phe	Lys	Ala	Ile	Asn	His	Val	Leu	Pro	Ile	Met	Gln			
					435						440						445	
Glu	Gly	Arg	Thr	Tyr	Gln	Glu	Ala	Ile	His	Gln	Leu	Gly	Tyr	Asp	Thr			
					450						455						460	
Thr	Asn	Leu	Lys	Lys	Glu	Asn	Arg	Ser	Met	Phe	Leu	Pro	Ile	Ile	Pro			
					465						470						475	
Asp	Glu	Ile	Thr	Asn	Pro	Ile	Val	Lys	Arg	Ala	Leu	Thr	Gln	Ala	Arg			
					485						490						495	
Lys	Val	Val	Asn	Ala	Ile	Ile	Arg	Arg	Tyr	Gly	Ser	Pro	Asn	Ser	Val			
					500						505						510	
His	Ile	Glu	Leu	Ala	Arg	Glu	Leu	Ser	Lys	Ser	His	Asp	Glu	Arg	Lys			
					515						520						525	
Lys	Ile	Met	Thr	Ala	His	Asp	Glu	Asn	Tyr	Lys	Lys	Asn	Lys	Gly	Ala			
					530						535						540	
Val	Ser	Ile	Leu	Ile	Asp	Asn	Gly	Ile	Leu	Asn	Pro	Thr	Gly	Tyr	Asp			
					545						550						555	
																	560	

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575

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

Glu Arg Ser Gly Pro Pro Val Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Gly
610 615 620

Asp Glu Asn Gln Lys Lys Gly Asn Arg Ile Pro Tyr Thr Phe Phe Ser
625 630 635 640

Glu Glu Asp Lys Glu Trp Glu Ser Phe Glu Ser Tyr Val Arg Ser Asn
645 650 655

Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr
660 665 670

Leu Pro Arg Glu Ser Asn Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685

Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys
690 695 700

Phe Lys Glu Ala Val Gly Ile Thr Arg Lys Lys Tyr Val Gln Thr Val
705 710 715 720

Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735

Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750

Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln
755 760 765

Ile Lys Glu Gly Asn Lys Ser Ile Lys Lys Pro Tyr Phe Pro Leu Pro
770 775 780

Trp Met Gly Phe Arg Glu Glu Ile Leu Ser His Leu Glu Ser Gln Pro
785 790 795 800

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

805 810 815
 Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ser
 820 825 830

 Ala His Asp Gln Thr Val Met Lys Lys Gly Asp Ile Asp Lys Lys Gly
 835 840 845
 Lys Thr Ile Ile Ile Lys Arg Val His Leu Lys Asp Ile Lys Phe Asp
 850 855 860
 Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
 865 870 875 880
 Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu Tyr Arg Lys Glu Ser Lys
 885 890 895

 Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
 900 905 910
 Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
 915 920 925
 Arg Glu Ile Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
 930 935 940
 Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Val Pro Ile Tyr Val
 945 950 955 960

 Val Asp Thr Val Arg Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
 965 970 975
 Lys Gly Tyr Glu Gln Trp Leu Ser Ile Asp Asn Ser Phe Thr Phe Lys
 980 985 990
 Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
 995 1000 1005
 Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
 1010 1015 1020

 Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Tyr
 1025 1030 1035
 Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Ile
 1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Gln Glu Thr Arg Lys
1055 1060 1065

Ile Phe Lys
1070

<210> 103
<211> 1072
<212> PRT
<213> Bacillus sp.
<223> APG08290.1_RGN

<400
> 103

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

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

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

Trp Ile Lys Asp Ala Glu Phe Ser Lys Tyr Asp Arg Arg Arg Asn Ser
180 185 190

Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Asp Leu Glu Lys Glu
195 200 205

Ile Val Thr Leu Phe Glu Ala Gln Arg Lys Phe Gln Ser Ser Tyr Ala
210 215 220

Ser Lys Asn Leu Gln Glu Thr Tyr Leu Gln Ile Trp Ala His Gln Leu
225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Leu Asn Lys Val Gly Tyr Cys Ser
245 250 255

Leu Leu Lys Gly Lys Glu Arg Arg Ile Pro Lys Ala Thr Tyr Thr Phe
260 265 270

Gln Tyr Phe Ser Ala Leu Asp Gln Val Asn Arg Thr Arg Leu Gly Pro
275 280 285

Asp Phe Gln Pro Phe Thr Gln Glu Gln Lys Glu Ile Ile Leu Asp Lys
290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320

Ser Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln
325 330 335

Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350

Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Lys Lys Val
355 360 365

Val Ala Asn Tyr Ala Glu Arg Thr Asn Glu Ala Phe Ser Thr Leu Asp
370 375 380

Tyr Asp Ala Ile Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400

Ile Arg Ser Tyr Leu Lys Lys Ser Asn Asn Leu Ser Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Phe Thr Leu Ser Tyr Thr Lys Phe

420 425 430
 Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
 435 440 445

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

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

 Ser Leu Lys Glu Ile Pro Pro Asp Thr Phe Phe Asn Glu Leu Lys Lys
 580 585 590
 Glu Arg Asn Gly Ser Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
 595 600 605
 Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
 610 615 620
 Asp Glu Asn Arg Asn Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
 625 630 635 640

 Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
 645 650 655
 Lys Leu Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Lys Thr Tyr
 660 665 670

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

 Phe Lys Glu Val Glu Val Asn Leu Arg Lys Lys Arg Val Gln Thr Val
 705 710 715 720
 Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
 725 730 735
 Asn Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
 740 745 750
 Ala Cys Thr Asp His His Met Val Thr Arg Ile Thr Glu Tyr Tyr Gln
 755 760 765

 Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
 770 775 780
 Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
 785 790 795 800
 Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Ser
 805 810 815
 Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ala
 820 825 830

 Ala His Asp Gln Thr Ile Arg Arg Lys Gly Gly Ile Asp Lys Lys Gly
 835 840 845
 Lys Thr Ile Ile Ile Lys Arg Val Arg Leu Lys Asp Ile Lys Phe Asp
 850 855 860
 Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
 865 870 875 880
 Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Arg Lys Asn Ser Lys
 885 890 895

 Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
 900 905 910
 Asn Leu Ile Lys Arg Val Lys Ile Glu Gly Gln Thr Lys Ala Phe Val

915 920 925
 Arg Glu Val Asn Gly Gly Val Ala Gln Asn Ser Asp Leu Val Arg Val
 930 935 940
 Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
 945 950 955 960

 Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Lys Ser Gly
 965 970 975
 Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
 980 985 990
 Ser Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asn Glu
 995 1000 1005
 Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
 1010 1015 1020

 Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Asn Glu Tyr
 1025 1030 1035
 Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Val
 1040 1045 1050
 Gly Val Leu Gly Asp Leu Arg Leu Val Lys Gln Glu Thr Arg Arg
 1055 1060 1065
 Ile Phe Asn Arg
 1070
 <210> 104
 <211> 1337
 <212> PRT
 <213> Enterococcus sp.
 <223> APG05459.1_RGN

 <400> 104
 Met Lys Lys Asp Tyr Val Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
 1 5 10 15
 Gly Trp Ala Val Met Thr Glu Asp Tyr Gln Leu Val Lys Lys Lys Met
 20 25 30
 Pro Ile Tyr Gly Asn Thr Glu Lys Lys Lys Ile Lys Lys Asn Phe Trp

35 40 45
 Gly Val Arg Leu Phe Glu Glu Gly His Thr Ala Glu Asp Arg Arg Leu
 50 55 60

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

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

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

 Asp Phe Lys Lys Val Phe Gly Leu Glu Glu Glu Ala Lys Ile Thr Tyr
 260 265 270
 Ala Ser Glu Ser Tyr Glu Glu Asp Leu Glu Gly Ile Leu Ala Lys Val
 275 280 285

Gly Asp Glu Tyr Ser Asp Val Phe Leu Ala Ala Lys Asn Val Tyr Asp
290 295 300

Ala Val Glu Leu Ser Thr Ile Leu Ala Asp Ser Asp Lys Lys Ser His
305 310 315 320

Ala Lys Leu Ser Ser Ser Met Ile Val Arg Phe Thr Glu His Gln Glu
325 330 335

Asp Leu Lys Lys Phe Lys Arg Phe Ile Arg Glu Asn Cys Pro Asp Glu
340 345 350

Tyr Asp Asn Leu Phe Lys Asn Glu Gln Lys Asp Gly Tyr Ala Gly Tyr
355 360 365

Ile Ala His Ala Gly Lys Val Ser Gln Leu Lys Phe Tyr Gln Tyr Val
370 375 380

Lys Lys Ile Ile Gln Asp Ile Ala Gly Ala Glu Tyr Phe Leu Glu Lys
385 390 395 400

Ile Ala Gln Glu Asn Phe Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly
405 410 415

Val Ile Pro His Gln Ile His Leu Ala Glu Leu Gln Ala Ile Ile His
420 425 430

Arg Gln Ala Ala Tyr Tyr Pro Phe Leu Lys Glu Asn Gln Glu Lys Ile
435 440 445

Glu Gln Leu Val Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ser
450 455 460

Lys Gly Asp Ala Ser Thr Phe Ala Trp Leu Lys Arg Gln Ser Glu Glu
465 470 475 480

Pro Ile Arg Pro Trp Asn Leu Gln Glu Thr Val Asp Leu Asp Gln Ser
485 490 495

Ala Thr Ala Phe Ile Glu Arg Met Thr Asn Phe Asp Thr Tyr Leu Pro
500 505 510

Ser Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Lys Phe Met
515 520 525

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

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

 Arg Asn Glu Tyr Asn Thr Glu Ile Val Thr Leu Ser Gly Leu Glu Glu
 580 585 590
 Asp Gln Phe Asn Ala Ser Phe Ser Thr Tyr Gln Asp Leu Leu Lys Cys
 595 600 605
 Gly Leu Thr Arg Ala Glu Leu Asp His Pro Asp Asn Ala Glu Lys Leu
 610 615 620
 Glu Asp Ile Ile Lys Ile Leu Thr Ile Phe Glu Asp Arg Gln Arg Ile
 625 630 635 640

 Arg Thr Gln Leu Ser Thr Phe Lys Gly Gln Phe Ser Ala Glu Val Leu
 645 650 655
 Lys Lys Leu Glu Arg Lys His Tyr Thr Gly Trp Gly Arg Leu Ser Lys
 660 665 670
 Lys Leu Ile Asn Gly Ile Tyr Asp Lys Glu Ser Gly Lys Thr Ile Leu
 675 680 685
 Asp Tyr Leu Ile Lys Asp Asp Gly Val Ser Lys His Tyr Asn Arg Asn
 690 695 700

 Phe Met Gln Leu Ile Asn Asp Ser Gln Leu Ser Phe Lys Asn Ala Ile
 705 710 715 720
 Gln Lys Ala Gln Ser Ser Glu His Glu Glu Thr Leu Ser Glu Thr Val
 725 730 735
 Asn Glu Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Tyr Gln Ser
 740 745 750
 Leu Lys Ile Val Asp Glu Leu Val Ala Ile Met Gly Tyr Ala Pro Lys
 755 760 765

 Arg Ile Val Val Glu Met Ala Arg Glu Asn Gln Thr Thr Ser Thr Gly
 770 775 780

Lys Arg Arg Ser Ile Gln Arg Leu Lys Ile Val Glu Lys Ala Met Ala

785 790 795 800

Glu Ile Gly Ser Asn Leu Leu Lys Glu Gln Pro Thr Thr Asn Glu Gln

805 810 815

Leu Arg Asp Thr Arg Leu Phe Leu Tyr Tyr Met Gln Asn Gly Lys Asp

820 825 830

Met Tyr Thr Gly Asp Glu Leu Ser Leu His Arg Leu Ser His Tyr Asp

835 840 845

Ile Asp His Ile Ile Pro Gln Ser Phe Met Lys Asp Asp Ser Leu Asp

850 855 860

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

865 870 875 880

Val Pro Ser Lys Glu Val Val Lys Asp Met Lys Ala Tyr Trp Glu Lys

885 890 895

Leu Tyr Ala Ala Gly Leu Ile Ser Gln Arg Lys Phe Gln Arg Leu Thr

900 905 910

Lys Gly Glu Gln Gly Gly Leu Thr Leu Glu Asp Lys Ala His Phe Ile

915 920 925

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

930 935 940

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

945 950 955 960

Ile Ile Thr Leu Lys Ala Ser Leu Thr Ser Gln Phe Arg Ser Ile Phe

965 970 975

Gly Leu Tyr Lys Val Arg Glu Val Asn Asp Tyr His His Gly Gln Asp

980 985 990

Ala Tyr Leu Asn Cys Val Val Ala Thr Thr Leu Leu Lys Val Tyr Pro

995 1000 1005

Asn Leu Ala Pro Glu Phe Val Tyr Gly Glu Tyr Pro Lys Phe Gln

1010 1015 1020

Ala Phe Lys Glu Asn Lys Ala Thr Ala Lys Thr Ile Ile Tyr Thr

1025	1030	1035
Asn Leu Met Arg Phe Phe Thr	Glu Asp Glu Pro Arg	Phe Met Lys
1040	1045	1050
Asp Gly Glu Ile Leu Trp Ser	Asn Ser Tyr Leu Lys	Asn Ile Lys
1055	1060	1065
Lys Glu Leu Asn Tyr His Gln	Met Asn Ile Val Lys	Lys Val Glu
1070	1075	1080
Val Gln Lys Gly Gly Phe Ser	Lys Glu Ser Ile Lys	Pro Lys Gly
1085	1090	1095
Pro Ser Asn Lys Leu Ile Pro	Val Lys Asn Gly Leu	Asp Pro Gln
1100	1105	1110
Lys Tyr Gly Gly Phe Asp Ser	Pro Val Val Ala Tyr	Thr Val Leu
1115	1120	1125
Phe Thr His Glu Lys Gly Lys	Lys Pro Leu Ile Lys	Gln Glu Ile
1130	1135	1140
Leu Gly Ile Thr Ile Met Glu	Lys Thr Arg Phe Glu	Gln Asn Pro
1145	1150	1155
Ile Leu Phe Leu Glu Glu Lys	Gly Phe Leu Arg Pro	Arg Val Leu
1160	1165	1170
Met Lys Leu Pro Lys Tyr Thr	Leu Tyr Glu Phe Pro	Glu Gly Arg
1175	1180	1185
Arg Arg Leu Leu Ala Ser Ala	Lys Glu Ala Gln Lys	Gly Asn Gln
1190	1195	1200
Met Val Leu Pro Glu His Leu	Leu Thr Leu Leu Tyr	His Ala Lys
1205	1210	1215
Gln Cys Leu Leu Pro Asn Gln	Ser Glu Ser Leu Ala	Tyr Val Glu
1220	1225	1230
Gln His Gln Pro Glu Phe Gln	Glu Ile Leu Glu Arg	Val Val Asp
1235	1240	1245
Phe Ala Glu Val His Thr Leu	Ala Lys Ser Lys Val	Gln Gln Ile
1250	1255	1260

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

Ser Thr Thr Gly Leu Tyr Glu Thr Arg Arg Lys Val Val Asp
 1325 1330 1335

<210> 105

<211> 1436

<212> PRT

<213> Enterococcus sp.

<223> APG04583.1_RGN

<400> 105

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

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

100 105 110
 Glu Pro Lys Leu Ala Phe Asp Ser Glu Phe Leu Phe Lys Asp Ser Phe
 115 120 125

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

370 375 380
 His Leu Phe Leu Asn Asp Ile Ile Phe Tyr Gln Arg Pro Leu Arg Ser
 385 390 395 400
 Gln Lys Ser Ser Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Ser Lys
 405 410 415
 Asn Glu Asn Gly Glu Glu Val Ile His Arg Leu Lys Val Ile Ala Lys

 420 425 430
 Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Leu Leu Gln Trp Val Gln Asn
 435 440 445
 Leu Ala Ile Tyr Thr Lys Asp Asp Asp Lys Asn Val Thr Asn Glu Phe
 450 455 460
 Leu Lys Ser Thr Gln Asp Trp Glu Asp Leu Leu Arg Trp Leu His Ser
 465 470 475 480
 Lys Lys Glu Ile Lys Gln Asp Ala Leu Ile Lys Phe Leu Ile Glu Lys

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

 545 550 555 560
 Asn Asp Lys Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Asn Lys
 565 570 575
 Lys Gly Leu Asp Glu Val Thr Phe Val Glu Ala Phe Lys Lys Phe Pro
 580 585 590
 Pro Phe Lys Ser Asp Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys Lys
 595 600 605
 Leu Leu Pro Leu Met Arg Phe Gly Thr Gln Trp Asn Trp Asp Asn Ile

 610 615 620

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

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

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

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

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

 930 935 940
 Ile Pro Glu Lys Met Ile Glu Arg Gln Leu Asn Asp Thr Arg Tyr Ile
 945 950 955 960
 Ser Lys Tyr Ile Thr Gln Val Leu Ser Asn Ile Val Arg Asp Asp Lys
 965 970 975
 Glu Gly Ser Lys Asp Asp Gly Val Asn Ser Lys Asn Ile Val Pro Gly
 980 985 990
 Asn Gly Lys Ile Thr Thr Arg Leu Lys Gln Asp Trp Gly Leu Asn Asp

 995 1000 1005
 Val Trp Asn Asp Leu Val Leu Pro Arg Phe Glu Arg Met Asn Thr
 1010 1015 1020
 Leu Thr Asn Ser Asn Asp Phe Thr Ser Lys Asn Thr His Gly Lys
 1025 1030 1035
 Thr Ile Pro Thr Val Pro Ile Glu Leu Ser Lys Gly Phe Ser Lys
 1040 1045 1050
 Lys Arg Ile Asp His Arg His His Ala Met Asp Ala Leu Val Ile

 1055 1060 1065
 Ala Cys Ala Thr Arg Asp His Val Asn Leu Leu Asn Asn Glu Ser
 1070 1075 1080
 Ser Lys Ser Asp Thr Lys Arg Tyr Asp Leu Asn Arg Lys Leu Arg
 1085 1090 1095
 Lys Tyr Glu Lys Val Ala Tyr Asn Asp Pro Lys Thr Gly Glu Arg
 1100 1105 1110

Ile Glu Lys Glu Val Pro Lys Asp Phe Ile Lys Pro Trp Glu Thr

1115

1120

1125

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

1130

1135

1140

Phe Lys Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr

1145

1150

1155

Glu Lys Ile Glu Asn Gly Lys Lys Val Lys Val Glu Gln Lys Gly

1160

1165

1170

Ile Asn Trp Ala Val Arg Lys Ala Leu His Lys Glu Thr Val Ser

1175

1180

1185

Gly Gln Val His Leu Asp Arg Ile Lys Val Ala Lys Gly Lys Ile

1190

1195

1200

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

1205

1210

1215

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

1220

1225

1230

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

1235

1240

1245

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

1250

1255

1260

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

1265

1270

1275

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

1280

1285

1290

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

1295

1300

1305

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

1310

1315

1320

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

1325

1330

1335

Glu Thr Asn Glu Lys Gly Lys Gln Leu Leu Phe Thr Leu Ser Pro

1340 1345 1350
Asn Asp Leu Val Tyr Val Pro Glu Glu Gly Val Phe Asp Glu Asn

1355 1360 1365
Asn Ile Lys Val Asp Arg Ile Tyr Lys Val Val Ser Phe Ser Thr

1370 1375 1380
Tyr Gln Cys Phe Phe Val Arg Asn Asp Val Ser Thr Ser Val Val

1385 1390 1395
Asn Lys Val Glu Tyr Ser Ala Leu Asn Lys Met Glu Lys Ser Ile

1400 1405 1410
Asp Asn Ile Met Ile Lys Glu Asn Cys Val Lys Leu Asn Val Asp

1415 1420 1425
Arg Leu Gly Lys Ile Ser Lys Ala

1430 1435

<210> 106

<211> 1456

<212> PRT

<213> Enterococcus sp.

<223> APG01688.1_RGN

<400> 106

Met Met Ile Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile

1 5 10 15

Gly Trp Ala Leu Ile Lys Gln Asp Phe Glu Asn Lys His Gly Glu Ile

20 25 30

Leu Gly Met Gly Ser Arg Ile Ile Pro Met Ser Gln Asp Ile Leu Gly

35 40 45
Asp Phe Gly Lys Gly Asn Ser Ile Ser Gln Thr Ala Asp Arg Thr Lys

50 55 60

Tyr Arg Ser Val Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu

65 70 75 80

Arg Leu His Arg Val Leu His Leu Leu Asn Phe Leu Pro Gln His Tyr

85 90 95

Ala Ser Gln Ile Asp Phe Glu Lys Lys Phe Gly Lys Phe Lys Ser Glu

100 105 110

Thr Glu Pro Lys Leu Ala Trp Glu Asn Trp Gly Gly Lys Phe Ser Phe

115 120 125

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

130 135 140

Gly Gln Gly Leu Lys Ile Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu Arg

145 150 155 160

Lys Lys Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp Ile

165 170 175

Leu Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu

180 185 190

Glu Glu Glu Glu Asn Pro Asn Lys Leu Val Glu Phe Tyr Ser Leu Lys

195 200 205

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

210 215 220

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

225 230 235 240

Lys Ile Pro Leu Phe Asp Trp Lys Asp Lys Thr Arg Asp Phe Ile Val

245 250 255

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

260 265 270

Asn Glu Lys Arg Ser Phe Arg Ala Pro Ser Glu Asn Asp Trp Thr Leu

275 280 285

Val Lys Lys Lys Thr Glu Gln Glu Ile Asp Gln Ser His Lys Thr Val

290 295 300

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

305 310 315 320

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

325 330 335

Leu Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe His Gln Glu Leu Lys

340 345 350
 Asn Asp Asp Leu Tyr Asn Asp Cys Ile Arg Glu Leu Tyr Arg Asn Asn

 355 360 365
 Glu Ala His Gln Leu Thr Leu Ser Lys Lys Asp Phe Val His Leu Leu
 370 375 380
 Met Asp Asp Leu Ile Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser
 385 390 395 400
 Ser Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr Lys Asp Glu Asn
 405 410 415
 Gly Ile Glu His Thr Gln Tyr Leu Lys Ala Ile Pro Lys Ser Asn Pro

 420 425 430
 Tyr Tyr Gln Glu Phe Arg Leu Trp Gln Trp Met Tyr Asn Leu Asn Ile
 435 440 445
 Tyr Arg Lys Asp Asp Glu Ala Asn Val Thr Lys Glu Phe Leu Asn Thr
 450 455 460
 Asn Lys Asp Phe Glu Ser Leu Phe Glu Phe Leu Asn Asn Arg Lys Glu
 465 470 475 480
 Ile Glu Gln Lys Pro Leu Ile Lys Phe Leu Leu Glu Gln Lys Asp Ile

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

 545 550 555 560
 Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Thr Arg Asn Asp Leu
 565 570 575
 Asp Glu Asn Ser Phe Ile Glu Ala Phe Lys Lys Phe Ser Pro Phe Lys
 580 585 590

Ser Glu Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro
595 600 605

Leu Met Arg Leu Gly Lys Tyr Trp Tyr Glu Asp Glu Ile Val Lys His

610 615 620

Ser Asp Ile Tyr Phe Lys Asn Ile Glu Asn Leu Leu Gly Asp Phe Ser
625 630 635 640

Asn Arg Asp Lys Lys Ile Ser Glu Glu Asp Lys Glu Lys Trp Asn Lys
645 650 655

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

Glu Ile Asp Leu Phe Gln Gly Leu Arg Leu His Ile Ala Gln Tyr Leu

675 680 685

Val Tyr Gly Arg His Ser Glu Ala Ser Met Ile Gly Lys Trp Asn Ser
690 695 700

Ala Glu Asp Leu Glu Glu Phe Leu Lys Asp Phe Lys Gln His Ser Leu
705 710 715 720

Arg Asn Pro Ile Val Glu Gln Val Ile Thr Glu Thr Leu Arg Val Val
725 730 735

Lys Asp Ile Trp Leu Lys Tyr Gly Asn Gly Ala Lys Asp Phe Phe Asn

740 745 750

Glu Ile His Ile Glu Leu Gly Arg Glu Met Lys Leu Pro Ala Asp Asp
755 760 765

Arg Lys Lys Leu Thr Asn Gln Ile Ser Glu Asn Glu Asn Thr Asn Phe
770 775 780

Arg Ile Lys Ala Leu Leu Ala Glu Met Met Asn Asp Ser Ser Val Glu
785 790 795 800

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

805 810 815

Glu Asp Asp Val Leu Lys Ser Asp Ile Glu Ile Glu Asp Asp Ile Leu
820 825 830

Lys Ile Ser Lys Thr Ala Gln Pro Ser Pro Ser Asp Leu Lys Arg Tyr

835 840 845
 Lys Leu Trp Leu Glu Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln Ile
 850 855 860
 Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile

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

 930 935 940
 His Tyr Ala Asn Asn Gln Gly Lys Arg Asn Lys Leu Leu Met Glu Asp
 945 950 955 960
 Ile Pro Glu Lys Met Ile Glu Arg Gln Met Asn Asp Thr Arg Tyr Ile
 965 970 975
 Ser Lys Tyr Ile Ser Gly Val Leu Ser Asn Ile Val Arg Val Glu Asp
 980 985 990
 Gly Ser Asp Glu Gly Val Asn Ser Lys Asn Ile Val Pro Gly Asn Gly

 995 1000 1005
 Lys Ile Thr Thr Gln Leu Lys Gln Asp Trp Gly Leu Asn Asp Val
 1010 1015 1020
 Trp Asn Asp Leu Ile Leu Pro Arg Phe Glu Arg Met Asn Gln Leu
 1025 1030 1035
 Thr Asn Ser Lys Val Phe Thr Ala Trp Asn Glu Asn Tyr Gln Lys
 1040 1045 1050
 Phe Leu Pro Thr Val Pro Ile Glu Tyr Ser Lys Gly Phe Ser Lys

 1055 1060 1065
 Lys Arg Ile Asp His Arg His His Ala Leu Asp Ala Leu Val Ile
 1070 1075 1080

Ala Cys	Ala Thr	Lys Asp	His Val	Asn Leu	Leu Asn	Asn Gln	Ser
1085			1090			1095	
Ala Lys	Ser Asp	Thr Lys	Arg Tyr	Asp Leu	Lys Lys	Lys Ser	Met
1100			1105			1110	
Lys Phe	Glu Lys	Val Val	Tyr Asn	Asp Ala	Lys Thr	Gly Glu	Lys
1115			1120			1125	
Ile Glu	Arg Glu	Val Pro	Lys Gln	Phe Leu	Lys Pro	Trp Glu	Asn
1130			1135			1140	
Phe Thr	Leu Asp	Val Lys	His Asn	Leu Glu	Thr Ile	Ile Val	Ser
1145			1150			1155	
Phe Lys	Gln Asn	Leu Arg	Val Ile	Asn Lys	Ala Thr	Asn Tyr	Tyr
1160			1165			1170	
Glu Lys	Tyr Val	Glu Lys	Asp Gly	Thr Lys	Asn Lys	Glu Arg	Val
1175			1180			1185	
Glu Gln	Thr Gly	Thr Asn	Trp Ala	Ile Arg	Lys Pro	Met His	Lys
1190			1195			1200	
Asp Thr	Val Ser	Gly Lys	Val Asp	Leu Pro	Trp Val	Lys Val	Pro
1205			1210			1215	
Lys Gly	Lys Ile	Leu Thr	Ala Thr	Arg Lys	Ser Leu	Asp Ser	Ser
1220			1225			1230	
Phe Asp	Leu Lys	Ser Ile	Gly Ser	Ile Thr	Asp Thr	Gly Ile	Gln
1235			1240			1245	
Lys Ile	Leu Lys	Asn Tyr	Leu Ala	Phe Lys	Asp Gly	Asn Pro	Glu
1250			1255			1260	
Leu Ala	Phe Ser	Pro Glu	Gly Ile	Asp Asp	Leu Asn	Lys Asn	Ile
1265			1270			1275	
Glu Lys	Tyr Asn	Asp Gly	Lys Pro	His Gln	Pro Ile	Asn Lys	Val
1280			1285			1290	
Arg Val	Phe Glu	Leu Gly	Ser Lys	Phe Gln	Val Gly	Gln Ser	Gly
1295			1300			1305	
Asn Lys	Lys Asp	Lys Tyr	Val Glu	Ala Ala	Lys Gly	Thr Asn	Leu

1310 1315 1320
Phe Phe Ala Val Tyr Glu Asp Glu Lys Gly Lys Arg Asn Tyr Glu
1325 1330 1335
Thr Ile Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu
1340 1345 1350
Ser Val Val Asp Leu Lys Gly Thr Asn Asp Phe Tyr Leu Cys Pro

1355 1360 1365
Asn Asp Phe Val Tyr Ile Pro Ser Gly Asp Glu Leu Glu Asn Ile
1370 1375 1380
Asn Asn Val Asp Phe Lys Asp Ile Lys Lys Glu Ile Asn Glu Arg
1385 1390 1395
Ile Tyr Lys Val Val Ser Phe Thr Gly Asn Arg Leu Ser Cys Ile
1400 1405 1410
Pro Tyr Met Val Ala Thr Thr Ile Val Asn Lys Leu Glu Phe Thr

1415 1420 1425
Gln Leu Asn Lys Ile Glu Phe Thr Lys Glu Lys Glu Ile Cys Ile
1430 1435 1440
Lys Leu Asn Val Asp Arg Leu Gly Asn Ile Ser Lys Ala
1445 1450 1455

<210> 107
<211> 1071
<212> PRT
<213> Bacillus sp.
<223> APG05733.1_RGN
<400> 107

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Gln
20 25 30
Tyr Glu Lys Thr Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45
Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile

50 55 60
 Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
 65 70 75 80

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

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

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

 Gln Tyr Phe Asn Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
 275 280 285
 Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Ile Ile Leu Asp Lys
 290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
 305 310 315 320
 Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
 325 330 335

Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
 340 345 350

Met Lys Pro Phe Ile Asn Leu Lys Pro Phe Tyr Glu Ile Lys Lys Val
 355 360 365

Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Val Phe Ser Ala Leu Asp
 370 375 380

Tyr Asp Thr Val Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
 385 390 395 400

Ile Arg Ser Tyr Leu Lys Arg Ser Asn Asn Leu Ser Lys Arg Cys Tyr
 405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe
 420 425 430

Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
 435 440 445

Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Thr
 450 455 460

Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Ile Ile Pro
 465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
 485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
 500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
 515 520 525

Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
 530 535 540

Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp

545 550 555 560
 Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
 565 570 575
 Ser Leu Lys Lys Ile Pro Ala Asn Thr Phe Phe Asn Glu Leu Lys Lys
 580 585 590

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

 Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr
 660 665 670
 Leu Pro Arg Glu Ser Asn Leu Ile Lys Glu Arg His Leu Asn Asp Thr
 675 680 685
 Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys
 690 695 700
 Phe Lys Glu Ala Val Gly Ile Thr Arg Lys Lys Tyr Val Gln Thr Val
 705 710 715 720

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

 Trp Met Gly Phe Arg Glu Glu Ile Leu Ser His Leu Glu Ser Gln Pro
 785 790 795 800

Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Ser
805 810 815
Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ser
820 825 830
Ala His Asp Gln Thr Val Met Lys Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845

Lys Thr Ile Ile Ile Lys Arg Val His Leu Lys Asp Ile Lys Phe Asp
850 855 860
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880
Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu Tyr Arg Lys Lys Ser Lys
885 890 895
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
900 905 910

Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
915 920 925
Arg Glu Ile Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940
Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960
Val Asp Thr Val Arg Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
965 970 975

Lys Gly Tyr Glu Gln Trp Leu Ser Ile Asp Asn Ser Phe Thr Phe Lys
980 985 990
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005
Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
1010 1015 1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Tyr
1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Ile

1040 1045 1050
 Gly Val Leu Gly Asp Leu Arg Leu Val Arg Gln Glu Thr Arg Lys
 1055 1060 1065
 Ile Phe Lys
 1070
 <210> 108
 <211> 1403
 <212> PRT
 <213> Chryseobacterium sp.
 <223> APG06207.1_RGN
 <400> 108
 Met Val Glu Arg Ile Leu Gly Leu Asp Leu Gly Val Ser Ser Ile Gly
 1 5 10 15

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

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

 Val Gln Ser Asp Asp Glu Gly Lys Ala Ile Asp Gly Met Ser Val Ala
 145 150 155 160
 Lys Glu Leu Tyr Glu Lys Asn Tyr Thr Pro Gly Gln Tyr Val Tyr Asp

165 170 175
 Leu Leu Lys Asn Gly Lys Lys Gln Ile Pro Asp Phe Tyr Arg Ser Asp
 180 185 190
 Leu Gln Asn Glu Leu Asp Arg Val Trp Glu Phe Gln Lys Gln Tyr His
 195 200 205

 Pro Asp Ile Leu Thr Asp Glu Tyr Lys Lys Glu Leu Glu Gly Lys Gly
 210 215 220
 Gln Arg Ala Thr Ala Ala Ile Phe Leu Asn Arg Tyr Lys Ile Tyr Thr
 225 230 235 240
 Ala Asp Asn Lys Gly Thr Arg Glu Glu Lys Arg Phe Gln Ala Tyr Gln
 245 250 255
 Trp Arg Val Glu Ala Leu Ser Gln Pro Leu Asp Ile Lys Glu Val Ala
 260 265 270

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

 Glu Lys Ile Trp Glu Thr Gln Ala Lys Phe His Ser Val Leu Thr Thr
 340 345 350
 Glu Leu Lys Glu Asp Val Arg Asp Ile Ile Ile Phe Tyr Gln Arg Lys
 355 360 365
 Leu Lys Ser Gln Lys Gly Leu Ile Ser Phe Cys Glu Phe Glu Ser Lys
 370 375 380
 Glu Ile Val Ile Glu Glu Asn Gly Lys Lys Lys Arg Lys Lys Val Gly
 385 390 395 400

 Leu Lys Val Val Pro Lys Ser Ser Pro Leu Phe Gln Glu Phe Lys Ile
 405 410 415

Trp Gln Val Leu Asn Asn Ile Glu Ile Lys Asn Glu Asp Gly Arg Arg
420 425 430

Tyr Leu Asp Gln Glu Glu Lys Glu Leu Leu Phe Asn Glu Leu Asn Cys
435 440 445

Lys Gly Asn Leu Thr Ser Gln Lys Cys Leu Glu Phe Leu Gly Phe Lys
450 455 460

Ser Lys Glu Thr Lys Ile Asn Phe Lys Val Ile Glu Gly Asn Arg Thr
465 470 475 480

Asn Glu Lys Leu Tyr Asp Ala Phe Leu Lys Ile Leu Glu Leu Lys Gly
485 490 495

Tyr Asn Val Phe Glu Leu Leu Lys Ile Lys Glu Asp Arg Asp Glu Ala
500 505 510

Lys Leu Ser Glu Leu Lys Ala Ser Ala Asp Glu Ile Lys Arg Met Val
515 520 525

Lys Glu Ile Phe Asn Ala Asn Leu Ile Asn Thr Ser Ile Leu Asp Phe
530 535 540

Asn Ala Glu Leu Glu Gly Lys Asp Phe Glu Asn Gln Ala Ser Tyr Gln
545 550 555 560

Phe Trp His Leu Ile Tyr Ser Tyr Glu Gly Asp Asn Ser Pro Ser Gly
565 570 575

Asn Glu Lys Leu Tyr Glu Leu Leu Glu Asn Lys Phe Gly Phe Lys Lys
580 585 590

Glu His Ser Lys Ile Leu Ser Thr Ile Val Phe Pro Gln Glu Tyr Gly
595 600 605

Ser Leu Ser Ala Lys Ala Met Arg Arg Ile Tyr Pro Tyr Ile Lys Asp
610 615 620

Asn Lys Tyr Ser Glu Ala Cys Leu Leu Ala Gly Tyr Asn His Ser Lys
625 630 635 640

Gln Ser Leu Thr Lys Glu Glu Leu Glu Asn Arg Lys Leu Lys Glu Gln
645 650 655

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

660	665	670	
Ile Leu Asn Gln Met Ile Asn Leu Val Asn Ala Ile Ile Ser Glu His			
675	680	685	
Gly Lys Leu Asp Gly Val Arg Ile Glu Leu Ala Arg Glu Leu Lys Lys			
690	695	700	
Ser Ala Gln Glu Arg Glu Asp Met Thr Lys Ser Ile Asn Glu Ala Thr			
705	710	715	720
Ile Leu His Gln Lys Tyr Ala Thr Val Leu Arg Gln Glu Phe Gly Val			
725	730	735	
Val Asn Pro Ser Arg Asn Asp Ile Ile Arg Tyr Lys Leu Tyr Met Glu			
740	745	750	
Leu Ala Gly Asn Gly Tyr Lys Asp Leu Tyr Thr Asn Val Lys Ile Glu			
755	760	765	
Lys Glu Asn Ile Phe Thr Asp Lys Tyr Asp Ile Asp His Ile Ile Pro			
770	775	780	
Gln Ser Arg Phe Phe Asp Asp Ser Phe Ser Asn Lys Val Leu Val Pro			
785	790	795	800
Arg Gln Ala Asn Leu Asp Lys Gly Asn Leu Thr Gly Tyr Asp Phe Met			
805	810	815	
Ser Asn Lys Gly Ala Glu Arg Glu Glu His Phe Leu Asn Val Ile Lys			
820	825	830	
Asp Leu Leu Glu Asn Gly Ser Ile Ser Lys Ala Lys Tyr Glu Lys Leu			
835	840	845	
Lys Lys Lys Gly Ile Glu Ile Gly Asp Gly Phe Ile Glu Arg Asp Leu			
850	855	860	
Arg Asp Thr Gln Tyr Ile Ala Arg Lys Ala Lys Glu Ile Leu Phe Glu			
865	870	875	880
Ile Thr Asn Ser Val Ile Ser Thr Ser Gly Arg Ile Thr Asp Lys Leu			
885	890	895	
Arg Glu Asp Trp Asp Leu Val Asn Thr Met Lys Glu Leu Asn Leu Asp			
900	905	910	

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

915 920 925

Glu Glu Lys Glu Val Ile Pro Asp Trp Thr Lys Arg Asn Asp His Arg

930 935 940

His His Ala Met Asp Ala Leu Thr Val Ala Phe Thr Thr Arg Asn His

945 950 955 960

Ile Gln Tyr Leu Asn Tyr Leu Asn Ala Arg Lys Asp Glu Lys His Lys

965 970 975

Glu His Gln Asn Ile Tyr Ala Ile Glu Asn Ile Ile Thr Glu Ile Ile

980 985 990

Glu Lys Lys Asn Gly Ser Lys Ser Arg Arg Phe Lys Thr Pro Met Asn

995 1000 1005

Arg Phe Arg His Glu Ala Lys Gln His Leu Lys Glu Val Leu Val

1010 1015 1020

Ser His Lys Thr Lys Asn Lys Val Val Thr Ser Asn Ile Asn Lys

1025 1030 1035

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

1040 1045 1050

Arg Gly Gln Leu His Lys Glu Thr Val Tyr Gly Ser Lys Lys Phe

1055 1060 1065

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

1070 1075 1080

Thr Ile Leu Met Val Val Asn Pro Ile His Arg Asn Ala Leu Leu

1085 1090 1095

Asn Arg Leu Lys Glu Tyr Gly Asn Asp Pro Lys Lys Ala Phe Ser

1100 1105 1110

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

1115 1120 1125

Lys Ile Glu Val Leu Pro Glu Lys Val Thr Leu Ser Trp Phe Glu

1130 1135 1140

Thr Gly Tyr Thr Ile Arg Lys Ala Val Thr Pro Asp Asn Phe Lys

1145	1150	1155
Asp Tyr Lys Asn Leu Glu Lys	Ile Thr Asp Met Gly	Ile Lys Lys
1160	1165	1170
Ile Leu Lys Glu Arg Leu Asp	Gln Phe Lys Gly Asn	Ala Lys Glu
1175	1180	1185
Ala Phe Ser Asn Leu Asp Lys	Glu Pro Ile Trp Leu	Asn Glu Glu
1190	1195	1200
Lys Gly Ile Ala Ile Lys Thr	Val Thr Ile Thr Gly	Val Ser Asn
1205	1210	1215
Ala Glu Ser Leu His Phe Lys	Lys Asp His Leu Gly	Lys Glu Ile
1220	1225	1230
Leu Asp Glu Asn Gly Asn Lys	Ile Pro Val Asp Phe	Val Ser Thr
1235	1240	1245
Gly Asn Asn His His Leu Ala	Ile Tyr Leu Asn Glu	Glu Glu Lys
1250	1255	1260
Leu Asp Asp Lys Met Val Thr	Phe Tyr Glu Ala Val	Leu Arg Val
1265	1270	1275
Asn Gln Gly Leu Pro Val Ile	Asp Lys Asn Tyr Asn	Arg Glu Lys
1280	1285	1290
Gly Tyr Lys Phe Leu Met Thr	Leu Lys Gln Asn Glu	Met Phe Val
1295	1300	1305
Phe Pro Asn Glu Glu Phe Asp	Pro Ser Glu Ile Asp	Leu Leu Asp
1310	1315	1320
Glu Lys Asn Leu Glu Gln Ile	Ser Gln Asn Leu Phe	Arg Val Gln
1325	1330	1335
Lys Ile Ser Lys Val Gly Tyr	Gly Asn Ser Phe Ile	Arg Asp Phe
1340	1345	1350
Val Phe Arg His His Leu Glu	Thr Thr Val Glu Asp	Arg Lys Glu
1355	1360	1365
Leu Arg Asn Thr Thr Tyr Ile	Gln Leu Lys Ser Leu	Glu Gly Leu
1370	1375	1380

Arg Asn Ile Val Lys Val Arg Leu Asn His Leu Gly Lys Ile Val
1385 1390 1395

Gln Ile Gly Glu Tyr
1400

<210> 109

<211> 1404

<212> PRT

<213> Sphingobacterium sp.

<223> APG01647.1_RGN

<400> 109

Met Met Lys Arg Ile Leu Gly Leu Asp Leu Gly Thr Ser Ser Ile Gly

1 5 10 15

Trp Ala Tyr Val Leu Glu Ala Gln Asp Glu Asn Glu Ile Ser Gln Ile

20 25 30

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

35 40 45

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

50 55 60

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

65 70 75 80

Asn Leu Ile Glu Val Leu Lys Arg Glu Lys Trp Ile Asp Asp Lys Ser

85 90 95

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

100 105 110

Arg Ala Ala Ala Ala Ser Lys Glu Ile Lys Leu Glu Glu Leu Ala Arg

115 120 125

Val Leu Leu Met Ile Asn Lys Lys Arg Gly Tyr Lys Ser Ser Arg Lys

130 135 140

Val Asn Lys Glu Asp Asp Gly Gln Leu Val Asp Gly Met Glu Val Ala

145 150 155 160

Lys Lys Leu Tyr His Glu Asn Leu Thr Pro Gly Glu Phe Val Tyr Arg

165 170 175
 Leu Leu Lys Glu Gly Lys Lys Lys Ile Pro Asp Phe Tyr Arg Ser Asp
 180 185 190
 Leu Gln Arg Glu Leu Asp Leu Ile Trp Asp Phe Gln Phe Lys Tyr Tyr
 195 200 205
 Pro Asp Leu Leu Thr Gln Asp Leu Lys Glu Glu Leu Lys Gly Lys Gly
 210 215 220
 Leu Arg Ala Thr Ser Ala Ile Phe Trp Ser Lys Tyr Lys Phe Asn Thr
 225 230 235 240

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

 Thr Val Gly Gln Phe Leu Phe Asn Gln Leu Gln His Asn Thr His Ala
 305 310 315 320
 Arg Val Lys Asn Gln Val Phe Tyr Arg Gln Asp Tyr Leu Asp Glu Phe
 325 330 335
 Glu Lys Ile Trp Ser Val Gln Ser Gln Tyr His Gln Gly Leu Thr Glu
 340 345 350
 Gln Leu Lys Glu Glu Ile Arg Asp Ile Ile Ile Phe Tyr Gln Arg Lys
 355 360 365

 Leu Lys Ser Gln Lys Gly Leu Ile Ser Phe Cys Glu Phe Glu Gln His
 370 375 380
 Glu Val Val Ile Asn Gly Lys Asn Lys Ile Val Gly Leu Arg Val Ala
 385 390 395 400
 Pro Lys Ser Ser Pro Ile Phe Gln Glu Phe Lys Ile Trp Gln Gln Ile
 405 410 415

Asn Asn Val Lys Leu Arg Asn Lys Ile Thr Lys Glu Val Tyr Asn Leu
420 425 430

Ala Glu Glu Gln Lys Arg Ser Leu Phe Glu Thr Leu Asn Leu Lys Gly
435 440 445

Lys Leu Ser Ala Tyr Gln Val Leu Asn Leu Ile Glu Val Lys Pro Lys
450 455 460

Glu Trp Glu Leu Asn Tyr Thr Glu Leu Glu Gly Asn Ser Thr Asn Leu
465 470 475 480

Ala Leu Tyr Asn Ala Tyr Leu Asp Ile Leu Asp Ile Glu Gly Tyr Asp
485 490 495

Val Arg Ser Glu Leu Lys Ile Lys Leu Asn Lys Asp Glu Ile Thr Leu
500 505 510

Ser Asp Leu Asp Ile Pro Val Ser Glu Ile Lys Gly Met Ile Arg Arg
515 520 525

Ile Phe Gln His Leu Gly Ile Asn Thr Ser Ile Leu Asp Phe Asn Ala
530 535 540

Gln Met Gln Gly Asp Asp Phe Glu Lys Gln Leu Ser Tyr Gln Leu Trp
545 550 555 560

His Leu Leu Tyr Ser Tyr Glu Glu Asp Asn Ser Lys Thr Gly Met Glu
565 570 575

Arg Leu Tyr Asn Gln Leu Asn Leu Lys Phe Gly Phe Ser Leu Asp Gln
580 585 590

Ala Lys Ile Phe Gly Lys Ile Ala Leu Gln Glu Asp Tyr Gly Asn Leu
595 600 605

Ser Thr Lys Ala Ile Arg Lys Ile Tyr Pro Tyr Ile Gln Asp Val Glu
610 615 620

Tyr Ser Val Ala Cys Lys Gln Ala Gly Tyr Asn His Ser Lys Ser Ser
625 630 635 640

Leu Thr Lys Glu Gln Leu Thr Asn Arg Val Leu Lys Asp His Leu Asp
645 650 655

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

660 665 670
 Asn Gln Met Val Asn Val Val Asn Thr Leu Ile Glu Thr Glu Asn Asp
 675 680 685

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

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

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

 Gln Tyr Ile Ala Lys Lys Ala Lys Ala Met Leu Phe Glu Ile Ala Pro
 885 890 895
 Ser Val Ile Ser Thr Ser Gly Ser Val Thr Asp Arg Leu Arg Glu Asp
 900 905 910

Trp Gly Leu Val Ser Val Met Lys Glu Leu Asn Leu Pro Lys Phe Lys

915 920 925

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

930 935 940

Glu Val Ile Lys Asp Trp Ser Lys Arg Asn Asp His Arg His His Ala

945 950 955 960

Met Asp Ala Leu Thr Val Ala Phe Thr Lys His Ser His Ile His Tyr

965 970 975

Leu Asn His Leu Asn Ala Arg Lys Asn Glu Lys Ser Glu Phe Phe Ser

980 985 990

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

995 1000 1005

Leu Gly Asn Arg Lys Arg Leu Phe Lys Glu Pro Ile Pro His Phe

1010 1015 1020

Arg Ser Ile Ala Lys Glu His Leu Glu Ser Val Leu Val Ser His

1025 1030 1035

Lys Ala Lys Asn Lys Val Val Thr Lys Asn Arg Asn Lys Ile Asp

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

Glu Glu Lys Ile Ser Ala Lys Phe Asp Ala Ala Thr Ile Ala Met

1085 1090 1095

Val Ser Asn Pro Arg Tyr Arg Ser Ala Leu Leu Asp Arg Leu Ser

1100 1105 1110

Ala Tyr Gly Asn Asp Pro Lys Lys Ala Phe Thr Gly Lys Asn Ser

1115 1120 1125

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

1130 1135 1140

Val Pro Glu Lys Val Lys Leu Val Trp Leu Glu Glu Asp Tyr Thr

1145	1150	1155
Ile Arg Lys Glu Ile Gly Pro	Asp Leu Lys Ile Glu	Lys Val Ile
1160	1165	1170
Asp Gln Gly Val Lys Arg Ile	Leu Glu Gln Arg Leu	Thr Glu Tyr
1175	1180	1185
Gly Gly Asp Ser Lys Lys Ala	Phe Ser Asp Leu Asp	Asn Asn Pro
1190	1195	1200
Ile Trp Leu Asn Lys Glu Lys	Gly Ile Ala Ile Lys	Arg Val Thr
1205	1210	1215
Ile Ser Gly Val Lys Asn Ala	Glu Ala Leu His Ile	Lys Lys Asp
1220	1225	1230
His Leu Gly Gln Val Leu Lys	Asp Lys Asp Gly Lys	Pro Met Pro
1235	1240	1245
Val Asp Phe Val Ser Thr Gly	Asn Asn His His Val	Ala Ile Tyr
1250	1255	1260
Glu Asp Asp Asn Gly Ala Leu	Gln Glu Arg Ile Val	Pro Phe Tyr
1265	1270	1275
Glu Ala Val Ala Arg Val Asn	Asn Gly Phe Pro Ile	Ile Asp Lys
1280	1285	1290
His Tyr Asn Gln His Leu Gly	Trp Lys Phe Leu Phe	Ser Met Lys
1295	1300	1305
Gln Asn Glu Met Phe Leu Phe	Pro Ser Glu Glu Phe	Asp Pro Lys
1310	1315	1320
Gln Ile Asn Leu Leu Asp Arg	Ser Asn Tyr Arg Leu	Ile Ser Lys
1325	1330	1335
Asn Leu Phe Arg Val Gln Lys	Phe Gly Glu Leu Ser	Lys Ser Gly
1340	1345	1350
Phe Trp Phe Arg His His Leu	Glu Thr Ser Val Asp	Val Asn Lys
1355	1360	1365
Ser Leu Arg Ser Ile Ser Tyr	Phe Asp Phe Tyr Ser	Lys Asp Phe
1370	1375	1380

Met Lys Asn Ile Ala Lys Val Arg Ile Asn His Leu Gly Glu Ile
1385 1390 1395

Ile His Val Gly Glu Tyr
1400

<210> 110

<211> 1453

<212> PRT

<213> Chryseobacterium sp.

<223> APG08032.1_RGN

<400> 110

Met Met Ile Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile
1 5 10 15

Gly Trp Ala Leu Ile Lys Gln Asp Phe Glu Asn Lys Tyr Gly Glu Ile
20 25 30

Leu Gly Met Gly Ser Arg Ile Ile Pro Met Asp Gln Lys Val Leu Gly
35 40 45

Glu Phe Gly Ser Gly Asn Ser Val Ser Gln Thr Ala Asp Arg Thr Lys
50 55 60

Tyr Arg Gly Val Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu
65 70 75 80

Arg Leu His Arg Val Leu Asn Val Leu Asn Phe Leu Pro Ile His Tyr
85 90 95

Ala Ser Gln Ile Asp Phe Glu Lys Arg Phe Gly Lys Phe Arg Glu Glu
100 105 110

Ile Glu Pro Lys Leu Val Tyr Asn Asn Asp Gly Phe Ile Phe Lys Asn
115 120 125

Ser Phe Glu Glu Met Leu Ser Asp Phe Arg Lys His Gln Pro Gln Ile
130 135 140

Leu Glu Asn Asp Lys Lys Ile Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu
145 150 155 160

Arg Lys Lys Ala Leu Thr Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp
165 170 175

Ile Leu Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly
180 185 190

Glu Asp Ile Glu Glu Glu Lys Asp Lys Thr Phe Val Arg Leu Lys Val
195 200 205

Glu Lys Ile Val Asp Ser Gly Asp Asn Val Lys Gly Lys Ile Leu Tyr
210 215 220

Asp Val Tyr Phe Asp Asn Gly Trp Lys Tyr Asp Lys Gln Ile Val Lys
225 230 235 240

Thr Glu Asp Trp Ile Glu Arg Val Lys Glu Phe Ile Val Thr Glu Ser
245 250 255

Phe Leu Lys Asn Gly Asp Ile Lys Arg Thr Tyr Lys Ala Val Asp Ser
260 265 270

Glu Lys Asp Trp Ile Ala Ile Lys Thr Lys Thr Glu Gln Glu Ile Glu
275 280 285

Lys Ser His Lys Thr Val Gly Thr Tyr Ile Tyr Glu Thr Ile Leu His
290 295 300

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

Lys Phe Tyr Lys Asp Glu Leu Arg Gln Ile Leu Glu Lys Gln Lys Glu
325 330 335

Phe His Pro Glu Leu Gln Asn Asp Asp Leu Tyr Ser Asp Cys Val Arg
340 345 350

Glu Leu Tyr Arg Asn Asn Glu Ala His Gln Leu Thr Leu Ser Lys Lys
355 360 365

Asp Phe Val His Leu Phe Ile Glu Asp Ile Ile Phe Tyr Gln Arg Pro
370 375 380

Leu Arg Ser Gln Lys Ser Ser Ile Ser Asn Cys Thr Leu Glu Tyr Arg
385 390 395 400

Arg Tyr Lys Asp Lys His Gly Val Glu His Ile Gln Tyr Leu Lys Ala
405 410 415

Ile Pro Lys Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Ile Trp Gln Trp

420 425 430
 Met Tyr Asn Leu Asn Ile Tyr Lys Arg Asp Asp Asp Thr Asn Val Thr
 435 440 445
 Lys Glu Phe Leu Asn Ser Thr Glu Asp Phe Glu Asn Leu Phe Asp Phe
 450 455 460

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

 Ile Trp His Ile Ile Tyr Ser Val Asn Asp Lys Ile Glu Tyr Glu Lys
 530 535 540
 Ala Leu Lys Ser Phe Ala Asn Lys Asn Asn Leu Asp Glu Val Ser Phe
 545 550 555 560
 Phe Glu Ala Phe Lys Lys Phe Pro Pro Phe Lys Asn Glu Tyr Gly Ser
 565 570 575
 Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu Met Arg Val Gly
 580 585 590

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

 Gln Lys Leu Arg Leu His Ile Ala Gln Tyr Leu Val Tyr Gly Arg His
 660 665 670

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

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

Ser Pro Met Gln Gln Glu Ile Leu Lys Ile Tyr Glu Asp Gly Val Leu
785 790 795 800
Asn Ser Gly Ile Glu Ile Glu Asp Glu Tyr Leu Lys Ile Ser Lys Thr
805 810 815
Ala Gln Pro Ser Ser Ser Asp Leu Lys Arg Tyr Lys Leu Trp Leu Glu
820 825 830
Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln Ile Ile Pro Leu Asn Lys
835 840 845

Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile Ile Pro Lys Ser Arg
850 855 860
Tyr Phe Asp Asp Ser Phe Ser Asn Lys Ile Ile Cys Glu Ser Ala Val
865 870 875 880
Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu Gly Phe Ile Lys Gln Phe
885 890 895
Ala Gly Ala Ile Val Glu Leu Gly Phe Gly Lys Asn Val Lys Val Phe
900 905 910

Glu Ile Glu Glu Tyr Glu Asp Phe Val Lys Lys His Tyr Ala Asn Asn

915	920	925	
Arg Gly Lys Arg Asn Lys Leu Leu Leu Glu Glu Ile Pro Glu Lys Met			
930	935	940	
Ile Glu Arg Gln Leu Asn Asp Thr Arg His Ile Ser Lys Tyr Ile Ser			
945	950	955	960
Ser Val Leu Ser Asn Ile Val Arg Val Glu Asp Gly Thr Asp Glu Gly			
965	970	975	
Val Asn Ser Lys Asn Ile Val Pro Gly Asn Gly Lys Ile Thr Ile Gln			
980	985	990	
Leu Lys Gln Asp Trp Gly Leu Asn Asp Val Trp Asn Asp Leu Ile Leu			
995	1000	1005	
Pro Arg Phe Glu Arg Met Asn Gln Leu Thr Asn Ser Thr Asp Phe			
1010	1015	1020	
Thr Ala Trp Asn Lys Asn His Gln Lys Tyr Leu Pro Thr Val Pro			
1025	1030	1035	
Ile Glu Phe Ser Lys Gly Phe Ser Lys Lys Arg Ile Asp His Arg			
1040	1045	1050	
His His Ala Leu Asp Ala Leu Val Ile Ala Cys Thr Thr Lys Asp			
1055	1060	1065	
His Val Asn Leu Leu Asn Asn Gln Ser Ala Lys Ser Asp Thr Lys			
1070	1075	1080	
Arg Tyr Asp Leu Lys Lys Lys Leu Met Lys Phe Glu Lys Val Val			
1085	1090	1095	
Tyr Asn His Thr Gln Thr Gly Glu Lys Ile Glu Arg Asp Ile Pro			
1100	1105	1110	
Lys Gln Phe Leu Lys Pro Trp Glu Thr Phe Thr Ile Asp Ala Lys			
1115	1120	1125	
Asn Arg Leu Glu Thr Ile Ile Val Ser Phe Lys Gln Asn Leu Arg			
1130	1135	1140	
Val Ile Asn Lys Ala Thr Asn His Tyr Glu Lys Tyr Val Glu Lys			
1145	1150	1155	

Glu Gly	Val Lys Met Lys Glu	Arg Val Glu Gln Thr	Gly Thr Asn
1160	1165	1170	
Trp Ala	Ile Arg Lys Ser Leu	His Glu Glu Thr Val	Ser Gly Lys
1175	1180	1185	
Ile Thr	Leu Ser Trp Val Glu	Thr Ser Lys Gly Glu	Phe Ile Thr
1190	1195	1200	
Ala Thr	Arg Lys Pro Leu Asp	Ser Ser Phe Thr Phe	Glu Lys Ile
1205	1210	1215	
Asn Lys	Ile Thr Asp Thr Gly	Ile Gln Lys Ile Leu	Asn Asn Tyr
1220	1225	1230	
Leu Glu	Ala Lys Asp Asn Asn	Pro Glu Leu Ala Phe	Ser Ala Glu
1235	1240	1245	
Gly Ile	Glu Asp Leu Asn Lys	Asn Ile Glu Lys Tyr	Asn Asp Gly
1250	1255	1260	
Lys Ser	His Gln Pro Ile Ile	Lys Ile Arg Leu Tyr	Glu Lys Gly
1265	1270	1275	
Lys Gly	Arg Phe Val Leu Gly	Gln Thr Gly Asn Lys	Ile Asn Lys
1280	1285	1290	
Tyr Val	Gln Gly Ser Pro Asn	Leu Phe Phe Ala Ile	Tyr Lys Asp
1295	1300	1305	
Glu Asn	Gly Lys Lys Ile Phe	Glu Ser Ile Arg Leu	Asp Ile Val
1310	1315	1320	
Ile Glu	Arg Leu Lys Gln Gly	Leu Gln Ala Ile Pro	Glu Thr Asn
1325	1330	1335	
Gln Asn	Gly Val Ser Leu Tyr	Gln Ser Leu Ser Pro	Leu Asp Leu
1340	1345	1350	
Val Tyr	Ile Pro Thr Glu Tyr	Glu Leu Glu Ser Pro	His Ile Leu
1355	1360	1365	
Asp Phe	Ser Lys Leu Asn Lys	Thr Gln Ile Thr Arg	Leu Tyr Asn
1370	1375	1380	
Thr Asn	Asp Phe Ser Gly Val	Thr Ala Tyr Phe Ser	Gln Asn Ser

1385 1390 1395

Phe Ala Lys His Ile Tyr Pro Lys Glu Met Asp Leu Ser Trp Asn

1400 1405 1410

Glu Lys Lys Gln Lys Leu Ser Gly Ser Phe Asp Ser Lys Thr Ala

1415 1420 1425

Ser Tyr Asn Asn Thr Ser Ile Lys Asp Ile Phe Ile Lys Val Lys

1430 1435 1440

Val Asp Arg Leu Gly Asn Ile Ser Lys Ala

1445 1450

<210> 111

<211> 1093

<212>

> PRT

<213> Brucella sp.

<223> APG05712.1_RGN

<400> 111

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

1 5 10 15

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

20 25 30

Gln Ser Glu Met Ala Gly Arg Asp Pro Gln Ser Lys Ala Ser Leu Ala

35 40 45

Val Ala Arg Arg Asp Ala Arg Gly Ala Arg Arg Arg Arg Asp Arg Tyr

50 55 60

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

65 70 75 80

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

85 90 95

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

100 105 110

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

115 120 125

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

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

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

 500 505 510
 Arg Lys Tyr Gly Lys Pro Asn Lys Ile Ala Ile Glu Val Gly Arg Glu
 515 520 525
 Leu Lys Leu Asn Glu Lys Gln Arg Asn Glu Val Asn Arg Glu Ile Gly
 530 535 540
 Gln Asn Thr Arg Ala Ala Met Ala Arg Gly Gln Gln Leu Val Glu Ile
 545 550 555 560
 Phe Lys Gln Pro Asn Thr Gly Tyr Asn Arg Leu Arg Leu Glu Leu Trp

 565 570 575
 Glu Asp Leu Asn Arg Glu Gln Pro Leu Lys Arg Leu Cys Thr Tyr Cys
 580 585 590
 Gly Lys Ala Ile Ala Ala His Met Leu Phe Asn Gly Glu Thr Asp Ile
 595 600 605
 Asp His Ile Leu Pro Tyr Ser Lys Thr Leu Asp Asp Ser Lys Ala Asn
 610 615 620

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

625 630 635 640

Pro Ala Asn Val Leu Gln Trp Arg Asp His Tyr Gly Glu Ile Leu Ala

645 650 655

Arg Ala Thr Ala Leu Pro Lys Asn Lys Gln Trp Arg Phe Ala Glu Asp

660 665 670

Ala Met Thr Arg Tyr Glu Ala Glu Gly Gly Phe Leu Ala Arg Gln Leu

675 680 685

Thr Asp Met Gln Tyr Ile Ser Arg Leu Ala Leu Thr Tyr Leu Ala His

690 695 700

Leu Tyr Asp Tyr Glu Glu Pro Asp Leu Asp Gly Val Tyr Lys Arg His

705 710 715 720

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

725 730 735

Gln Trp Ala Leu Asn Glu Leu Leu His Gly His Asn Leu Ala Gly Gly

740 745 750

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

755 760 765

Asp Ala Ile Val Ile Ala Cys Thr Ser Gln Ser Leu Ile Asn Arg Leu

770 775 780

Ser Thr Ala Ala Gly Glu Ala Glu Glu Arg Gly Ala Ala Arg Val Val

785 790 795 800

Glu Arg Ile Asp Pro Pro Trp Pro Ser Phe Arg Glu Asp Val Arg Glu

805 810 815

Ala Val Asn Ala Ile Val Val Ser His Lys Pro Asp His Gly Thr Ala

820 825 830

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

835 840 845

His Asn Asp Thr Ala Tyr Gly Glu Thr Gly Glu Lys Asp His Asn Gly

850 855 860

Asn Asn Leu Val Val Arg Arg Ile Ala Ile Ser Asp Ile Lys Arg Ser

865					870					875					880												
Ala	Asp	Ile	Met	Lys	Ile	Arg	Thr	Asn	Ala	His	Gly	His	Ser	Glu	Leu												
				885								890								895							
Arg	Asp	Arg	Leu	Tyr	Glu	Ala	Thr	Arg	Asp	Leu	Glu	Gly	Lys	Ala	Phe												
				900								905								910							
Glu	Gln	Ala	Val	Thr	Ala	Phe	Val	Lys	His	Asp	Ala	Lys	Phe	Lys	Gly												
				915								920								925							
Ile	Arg	His	Val	Arg	Val	Thr	Glu	Val	Gln	Asn	Pro	Val	Trp	Ile	Thr												
				930								935								940							
His	Gly	Gly	Gly	Lys	Tyr	Lys	Lys	Gly	Tyr	Leu	Pro	Gly	Gly	Asn	Asp												
945								950								955								960			
Arg	Phe	Asp	Val	Trp	Glu	Leu	Pro	Asp	Gly	Lys	Trp	Asp	Ala	Glu	Val												
				965								970								975							
Val	Thr	Thr	Phe	Asp	Ala	His	Arg	Pro	Asp	Phe	Thr	Pro	Arg	Met	Arg												
				980								985								990							
Ile	Glu	His	His	Asn	Ala	Arg	Lys	Ile	Met	Ser	Leu	Lys	Lys	Gly	Asp												
				995								1000								1005							
Met	Ile	Ala	Tyr	Asp	Asp	Pro	Asp	Ser	Gly	Lys	Arg	Val	Ile	Ala													
1010								1015								1020											
Ile	Val	Arg	Lys	Phe	Asp	Gln	Arg	Asn	Lys	Gln	Leu	Tyr	Leu	Asp													
1025								1030								1035											
Pro	His	Asn	Glu	Ala	Gly	Asn	Leu	Asp	Gln	Arg	Glu	Lys	Glu	Lys													
1040								1045								1050											
Thr	Tyr	Lys	Pro	Leu	Arg	Pro	Met	Pro	Asn	Pro	Leu	Lys	Lys	Tyr													
1055								1060								1065											
Arg	Pro	Arg	Gln	Val	Arg	Val	Asp	Glu	Ile	Gly	Gln	Val	Phe	Asp													
1070								1075								1080											
Pro	Gly	Pro	Trp	Trp	Glu	Lys	Arg	Ser	Asp																		
1085								1090																			
<210>	112																										
<211>	1054																										

<212> PRT

<213> Staphylococcus sp.

<223> APG01658.1_RGN

<400> 112

Met Lys Glu Lys Tyr Ile Leu Gly Leu Asp Leu Gly Ile Thr Ser Val

1 5 10 15

Gly Tyr Gly Ile Ile Asn Phe Glu Thr Lys Lys Ile Ile Asp Ala Gly

20 25 30

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

35 40 45

Ser Lys Arg Gly Ser Arg Arg Leu Lys Arg Arg Arg Ile His Arg Leu

50 55 60

Glu Arg Val Lys Leu Leu Leu Thr Glu Tyr Asp Leu Ile Asn Lys Glu

65 70 75 80

Gln Ile Pro Thr Ser Asn Asn Pro Tyr Gln Ile Arg Val Lys Gly Leu

85 90 95

Ser Glu Ile Leu Ser Lys Asp Glu Leu Ala Ile Ala Leu Leu His Leu

100 105 110

Ala Lys Arg Arg Gly Ile His Asn Ile Asn Val Ser Ser Glu Asp Glu

115 120 125

Asp Ala Ser Asn Glu Leu Ser Thr Lys Glu Gln Ile Asn Arg Asn Asn

130 135 140

Lys Leu Leu Lys Asp Lys Tyr Val Cys Glu Val Gln Leu Gln Arg Leu

145 150 155 160

Lys Glu Gly Gln Ile Arg Gly Glu Lys Asn Arg Phe Lys Thr Thr Asp

165 170 175

Ile Leu Lys Glu Ile Asp Gln Leu Leu Lys Val Gln Lys Asp Tyr His

180 185 190

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

195 200 205

Arg Arg Glu Tyr Phe Glu Gly Pro Gly Gln Gly Ser Pro Phe Gly Trp

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

 Val Phe Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu
 290 295 300
 Ile Gly Val Asn Pro Glu Asp Ile Lys Gly Tyr Arg Ile Thr Lys Ser
 305 310 315 320
 Gly Thr Pro Gln Phe Thr Glu Phe Lys Leu Tyr His Asp Leu Lys Ser
 325 330 335
 Ile Val Phe Asp Lys Ser Ile Leu Glu Asn Glu Ala Ile Leu Asp Gln
 340 345 350

 Ile Ala Glu Ile Leu Thr Ile Tyr Gln Asp Glu Gln Ser Ile Lys Glu
 355 360 365
 Glu Leu Asn Lys Leu Pro Glu Ile Leu Asn Glu Gln Asp Lys Ala Glu
 370 375 380
 Ile Ala Lys Leu Ile Gly Tyr Asn Gly Thr His Arg Leu Ser Leu Lys
 385 390 395 400
 Cys Ile His Leu Ile Asn Glu Glu Leu Trp Gln Thr Ser Arg Asn Gln
 405 410 415

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

Lys Val Ile Glu Lys Tyr Gly Ile Pro Glu Asp Ile Ile Ile Glu Leu
465 470 475 480

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

Gln Lys Lys Asn Glu Ala Thr Arg Lys Arg Ile Asn Glu Ile Ile Gly
500 505 510

Gln Thr Gly Asn Gln Asn Ala Lys Arg Ile Val Glu Lys Ile Arg Leu
515 520 525

His Asp Gln Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ser Ile Ala
530 535 540

Leu Met Asp Leu Leu Asn Asn Pro Gln Asn Tyr Glu Val Asp His Ile
545 550 555 560

Ile Pro Arg Ser Val Ala Phe Asp Asn Ser Ile His Asn Lys Val Leu
565 570 575

Val Lys Gln Ile Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Tyr Gln
580 585 590

Tyr Leu Asn Ser Ser Asp Ala Lys Leu Ser Tyr Asn Gln Phe Lys Gln
595 600 605

His Ile Leu Asn Leu Ser Lys Ser Lys Asp Arg Ile Ser Lys Lys Lys
610 615 620

Lys Asp Tyr Leu Leu Glu Glu Arg Asp Ile Asn Lys Phe Glu Val Gln
625 630 635 640

Lys Glu Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg
645 650 655

Glu Leu Thr Ser Tyr Leu Lys Ala Tyr Phe Ser Ala Asn Asn Met Asp
660 665 670

Val Lys Val Lys Thr Ile Asn Gly Ser Phe Thr Asn His Leu Arg Lys
675 680 685

Val Trp Arg Phe Asp Lys Tyr Arg Asn His Gly Tyr Lys His His Ala
690 695 700

Glu Asp Ala Leu Ile Ile Ala Asn Ala Asp Phe Leu Phe Lys Glu Asn

705 710 715 720
 Lys Lys Leu Gln Asn Ala Asn Lys Ile Leu Glu Lys Pro Thr Ile Glu
 725 730 735

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

 Val Gln Thr Ile Thr Asp Ile Tyr Gly Lys Asp Asn Thr Asn Leu Lys
 805 810 815
 Lys Gln Phe Asn Lys Asn Pro Glu Lys Phe Leu Met Tyr Gln Asn Asp
 820 825 830
 Pro Lys Thr Phe Glu Lys Leu Ser Ile Ile Met Lys Gln Tyr Ser Asp
 835 840 845
 Glu Lys Asn Pro Leu Ala Lys Tyr Tyr Glu Glu Thr Gly Glu Tyr Leu
 850 855 860

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

 Tyr Leu Asn Val Phe Lys Lys Asp Asn Tyr Tyr Tyr Ile Pro Lys Asp
 930 935 940
 Lys Tyr Gln Glu Leu Lys Glu Lys Lys Lys Ile Lys Asp Thr Asp Gln
 945 950 955 960

Phe Ile Ala Ser Phe Tyr Lys Asn Asp Leu Ile Lys Leu Asn Gly Asp
965 970 975
Leu Tyr Lys Ile Ile Gly Val Asn Ser Asp Asp Arg Asn Ile Ile Glu
980 985 990
Leu Asp Tyr Tyr Asp Ile Lys Tyr Lys Asp Tyr Cys Glu Ile Asn Asn
995 1000 1005
Ile Lys Gly Glu Pro Arg Ile Lys Lys Thr Ile Gly Lys Lys Thr
1010 1015 1020
Glu Ser Ile Glu Lys Phe Thr Thr Asp Val Leu Gly Asn Leu Tyr
1025 1030 1035
Leu His Ser Thr Glu Lys Ala Pro Gln Leu Ile Phe Lys Arg Gly
1040 1045 1050

Leu

<210> 113
<211> 1076
<212> PRT
<213> Bacillus sp.
<223> APG06498.1_RGN
<400> 113

Met Ser Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15
Ile Gly Trp Gly Val Ile Glu Leu Phe Trp Asn Lys Asp Arg Glu Arg
20 25 30
Tyr Glu Lys Val Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45
Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60
Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80
Glu Ile Arg Asn Leu Leu Val Gln His Gly Val Ile Thr Gln Glu Glu
85 90 95

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

340 345 350
Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Lys Lys Val
355 360 365
Val Ala Asn Tyr Ala Glu Arg Thr Asn Glu Ala Phe Ser Thr Leu Asp

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

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

500 505 510
His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
515 520 525
Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540
Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560
Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr

565 570 575
Ser Leu Lys Glu Ile Pro Pro Asp Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590

Glu Arg Asn Gly Ser Ser Ile Leu Glu Val Asp His Ile Leu Pro Tyr
 595 600 605
 Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
 610 615 620
 Asp Glu Asn Arg Asn Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
 625 630 635 640
 Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
 645 650 655
 Lys Leu Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Lys Ala Tyr
 660 665 670
 Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
 675 680 685
 Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
 690 695 700
 Phe Lys Glu Val Glu Val Asn Leu Arg Lys Lys Arg Val Gln Thr Val
 705 710 715 720
 Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
 725 730 735
 Asn Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
 740 745 750
 Ala Cys Thr Asp His His Met Val Thr Arg Ile Thr Glu Tyr Tyr Gln
 755 760 765
 Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
 770 775 780
 Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
 785 790 795 800
 Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Ser
 805 810 815
 Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ala
 820 825 830
 Ala His Asn Gln Met Ile Arg Arg Lys Gly Gly Ile Asp Lys Lys Gly

835	840	845	
Lys Thr Ile Ile Ile Lys Arg Val Arg Leu Lys Asp Ile Lys Phe Asp			
850	855	860	
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr			
865	870	875	880
Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Asp His Gly Lys Asn Leu Lys			
885	890	895	
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly			
900	905	910	
Asn Leu Ile Lys Arg Val Lys Ile Glu Gly Gln Thr Lys Ala Phe Val			
915	920	925	
Arg Glu Val Asn Gly Gly Val Ala Gln Asn Ser Asp Leu Val Arg Val			
930	935	940	
Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val			
945	950	955	960
Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Lys Ser Gly			
965	970	975	
Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys			
980	985	990	
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu			
995	1000	1005	
Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg			
1010	1015	1020	
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Asn Glu Tyr			
1025	1030	1035	
Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val			
1040	1045	1050	
Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg			
1055	1060	1065	
Asn Phe His Glu Ile Lys Ile Asn			
1070	1075		

<210> 114

<211> 1108

<212> PRT

<213> Brevibacillus sp.

<223> APG09106.1_RNG

<400> 114

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

1 5 10 15

Glu Ala Gln Asn Leu Arg Lys Gly Ile Trp Arg Thr His Arg Leu Leu

20 25 30

Asn Glu Gly Val Ala Tyr Tyr Met Lys Met Leu Leu Leu Phe Arg Gln

35 40 45

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

50 55 60

Arg His Ile Arg Glu Gln Gln Gln Lys Asn Arg Ala Asp Lys Asn Thr

65 70 75 80

Gln Ala Leu Pro Leu Asp Lys Ala Phe Ala Ala Leu Arg Gln Leu Tyr

85 90 95

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

100 105 110

Ile Ser Arg Lys Phe Leu Ser Pro Leu Val Asp Pro Asn Ser Glu Gly

115 120 125

Gly Lys Gly Thr Ser Lys Ala Gly Ala Lys Pro Thr Trp Gln Lys Lys

130 135 140

Lys Glu Ala Asn Asp Pro Thr Trp Glu Gln Asp Tyr Glu Lys Trp Lys

145 150 155 160

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

165 170 175

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

180 185 190

Asp Ile Ala Trp Leu Pro Leu Gln Ser Asn Gln Phe Val Arg Thr Trp

195 200 205

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

450 455 460
 Thr Phe Ser Gly His Leu Gly Gly Ala Lys Leu Gln Leu Glu Arg Gly
 465 470 475 480
 Asp Leu Glu Lys Glu Glu Lys Asp Phe Glu Gly Gly Glu Ile Gly Ser
 485 490 495
 Val Tyr Leu Asn Ile Val Ile Asp Phe Glu Pro Leu Gln Glu Val Lys

 500 505 510
 Asn Gly Arg Leu Gln Ser Pro Tyr Gly Gln Val Leu Gln Leu Val Arg
 515 520 525
 Arg Pro Asn Glu Phe Pro Lys Val Thr Thr Tyr Lys Ser Glu Glu Leu
 530 535 540
 Val Glu Trp Met Lys Ala Ser Gln Asn His Ser Ser Gly Val Glu Ser
 545 550 555 560
 Leu Glu Ser Gly Phe Arg Val Met Ser Ile Asp Leu Gly Leu Arg Thr

 565 570 575
 Ala Ala Ala Thr Ser Ile Phe Ser Val Glu Glu Ser Asn Asp Ala Asn
 580 585 590
 Ala Ala Gly Phe Ser Tyr Trp Ile Glu Gly Thr Pro Leu Val Ala Val
 595 600 605
 His Lys Arg Ser Tyr Met Leu Lys Leu Pro Gly Glu Gln Val Glu Lys
 610 615 620
 Gln Val Arg Glu Lys Arg Asp Glu Arg Gln Asp Gln Gln Arg Arg Val

 625 630 635 640
 Arg Phe Gln Ile Arg Ile Leu Ser Gln Val Ile Arg Met Ala Lys Lys
 645 650 655
 Gln Asn Arg Glu Arg Ala Asp Glu Leu Asp His Leu Ser Gln Ala Leu
 660 665 670
 Glu Lys Gln Lys Ser Leu Leu Asp Gln Thr Asp Arg Thr Phe Trp Asn
 675 680 685
 Gly Ile Val Cys Asp Leu Thr Asp Ala Leu Arg Glu Lys Glu Gly Gly

 690 695 700

Trp Glu Gln Ala Val Val Gln Ile His Arg Lys Ala Glu Glu His Val
705 710 715 720
Gly Lys Val Val Gln Ala Trp Arg Lys Arg Phe Asp Ala Asp Glu Arg
725 730 735
Lys Gly Ile Ala Gly Leu Ser Met Trp Ser Ile Glu Glu Leu Asp Ser
740 745 750
Leu Arg Lys Leu Leu Ile Ser Trp Ser Arg Arg Thr Arg Asn Pro Gln
755 760 765
Glu Ile Asn Arg Phe Glu Gln Gly His Thr Ser His Gln Arg Leu Leu
770 775 780
Thr His Ile Gln Asn Val Lys Glu Asp Arg Leu Lys Gln Leu Ser His
785 790 795 800
Ala Ile Val Met Thr Ala Leu Gly Tyr Val Tyr Asp Glu Lys Lys Leu
805 810 815
Glu Trp Phe Ala Lys Tyr Pro Ala Cys Gln Val Ile Leu Phe Glu Asn
820 825 830
Leu Ser Gln Tyr Arg Ser His Met Asp Arg Ser Thr Lys Glu Asn Ser
835 840 845
Thr Leu Met Lys Trp Ala His Arg Ser Ile Pro Lys Tyr Val His Met
850 855 860
Gln Ala Glu Pro Tyr Gly Ile Gln Ile Gly Asp Val Arg Ala Glu Tyr
865 870 875 880
Ser Ser Arg Phe His Ala Lys Thr Gly Thr Pro Gly Ile Arg Cys Lys
885 890 895
Met Val Lys Gly Gln Glu Leu Gln Gly Lys Arg Phe Glu Asn Leu Gln
900 905 910
Lys Arg Leu Val Ser Glu Gln Phe Leu Thr Glu Glu Gln Val Lys Gln
915 920 925
Leu Arg Pro Gly Asp Ile Val Pro Asp Asp Ser Gly Glu Trp Phe Met
930 935 940
Thr Leu Ser Asp Gly Ser Glu Gly Lys Glu Val Val Phe Leu Gln Ala

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

1010 1015 1020
 Val Trp Asp Ser Gln Ala Lys Leu Lys Gly Lys Thr Thr Phe Thr
 1025 1030 1035
 Glu Glu Ser Glu Ser Pro Glu Gln Leu Glu Asp Phe Gln Glu Ile
 1040 1045 1050
 Ile Glu Glu Ala Glu Glu Ala Lys Gly Thr Tyr Arg Thr Leu Phe
 1055 1060 1065
 Arg Asp Pro Ser Gly Val Phe Phe Pro Glu Phe Val Trp Asn Thr

1070 1075 1080
 Gln Lys Asp Phe Trp Ser Glu Val Lys Arg Arg Leu Tyr Gly Lys
 1085 1090 1095
 Leu Arg Glu Arg Phe Leu Met Lys Thr Arg
 1100 1105

<210> 115

<211> 1150

<212> PRT

<213> Enterococcus sp.

<223> APG09882.1_RGN

<400> 115

Met Tyr Ser Ile Gly Leu Asp Leu Gly Ile Ser Ser Val Gly Trp Ser
 1 5 10 15

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

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

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

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

 Val Gln Pro Gly Ile Ala Glu Glu Ala Gly Leu Ile Tyr Arg Lys Arg
 210 215 220
 Pro Tyr Tyr His Gly Pro Gly Asn Glu Ala Asn Asn Ser Pro Tyr Gly
 225 230 235 240
 Arg Trp Ser Asp Phe Lys Lys Thr Gly Gln Pro Ala Thr Asn Ile Phe
 245 250 255
 Asp Lys Leu Ile Gly Lys Asp Phe Gln Gly Glu Leu Arg Ala Ser Gly
 260 265 270

 Leu Ser Leu Ser Ala Gln Gln Tyr Asn Leu Leu Asn Asp Leu Thr Asn
 275 280 285

Leu Lys Ile Asp Gly Glu Ile Pro Leu Ser Pro Glu Gln Lys Glu Tyr
 290 295 300
 Ile Leu Ala Glu Leu Met Thr Lys Glu Phe Thr Arg Phe Gly Val Asn
 305 310 315 320
 Asp Val Val Lys Leu Leu Gly Val Lys Lys Glu Arg Leu Ser Gly Trp
 325 330 335

 Arg Leu Asp Lys Lys Gly Lys Pro Glu Ile His Thr Leu Lys Gly Tyr
 340 345 350
 Arg Asn Trp Arg Lys Ile Phe Ala Glu Ser Gly Ile Asp Leu Ala Thr
 355 360 365
 Leu Pro Thr Glu Thr Ile Asp Cys Leu Ala Lys Val Leu Thr Leu Asn
 370 375 380
 Thr Glu Arg Glu Gly Ile Glu Asn Thr Leu Ala Phe Glu Leu Ser Glu
 385 390 395 400

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

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

 Glu Leu His Ala Lys Asn Ser Gln Arg Lys Asn Asp Ser Gln Ser Tyr

530 535 540
 Phe Met Gln Lys Ser Gly Trp Ser Gln Glu Lys Phe Gln Thr Thr Ile
 545 550 555 560
 Gln Lys Asn Arg Arg Phe Leu Ala Lys Leu Leu Tyr Tyr Tyr Glu Gln
 565 570 575
 Asp Gly Ile Cys Ala Tyr Thr Gly Leu Ser Ile Ser Pro Glu Leu Leu
 580 585 590

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

 Trp Val Glu Ser Cys His Phe Ser His Lys Lys Glu Asn Asn Leu Leu
 660 665 670
 Glu Thr Arg Asn Ile Phe Asp Ser Glu Gln Val Glu Lys Phe Leu Ala
 675 680 685
 Arg Asn Leu Asn Asp Thr Arg Tyr Ala Ser Arg Leu Val Leu Asn Thr
 690 695 700
 Leu Gln Ser Phe Phe Ala Asn Gln Glu Thr Lys Val Arg Val Val Asn
 705 710 715 720

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

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

 Arg Glu Lys Thr Glu Val Lys Thr Leu Lys Ser Gly Lys Gln Lys Ile
 850 855 860
 Thr Thr Asp Glu Tyr Thr Ile Gly Lys Ile Lys Asp Ile Tyr Thr Val
 865 870 875 880
 Asp Gly Trp Glu Ala Phe Lys Lys Lys Gln Asp Lys Leu Leu Met Lys
 885 890 895
 Asp Leu Asp Glu Lys Thr Tyr Glu Arg Leu Leu Ser Ile Ala Glu Thr
 900 905 910

 Thr Pro Asp Phe Gln Glu Val Glu Glu Lys Asn Gly Lys Val Lys Arg
 915 920 925
 Val Lys Arg Ser Pro Phe Ala Val Tyr Cys Glu Glu Asn Asp Ile Pro
 930 935 940
 Ala Ile Arg Lys Tyr Ala Lys Lys Asn Asn Gly Pro Leu Ile Arg Ser
 945 950 955 960
 Leu Lys Tyr Tyr Asp Gly Lys Leu Asn Lys His Ile Asn Ile Thr Lys
 965 970 975

 Asp Ser Gln Gly Arg Pro Val Glu Lys Thr Lys Asn Gly Arg Lys Val
 980 985 990
 Thr Leu Gln Ser Leu Lys Pro Tyr Arg Tyr Asp Ile Tyr Gln Asp Leu
 995 1000 1005
 Glu Thr Lys Ala Tyr Tyr Thr Val Gln Leu Tyr Tyr Ser Asp Leu
 1010 1015 1020
 Arg Phe Val Glu Gly Lys Tyr Gly Ile Thr Glu Lys Glu Tyr Met

1025 1030 1035

Lys Lys Val Ala Glu Gln Thr Lys Gly Gln Val Val Arg Phe Cys

1040 1045 1050

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

1055 1060 1065

Ser Gln Cys Tyr Asp Val Arg Phe Tyr Asn Phe Gln Ser Ala Asn

1070 1075 1080

Ser Ile Asn Phe Lys Gly Leu Glu Gln Glu Met Met Pro Ala Glu

1085 1090 1095

Asn Gln Phe Lys Gln Lys Pro Tyr Asn Asn Gly Ala Ile Asn Leu

1100 1105 1110

Asn Ile Ala Lys Tyr Gly Lys Glu Gly Lys Lys Leu Arg Lys Phe

1115 1120 1125

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

1130 1135 1140

Glu Pro Lys Asn Ile Ile Lys

1145 1150

<210> 116

<211> 1408

<212> PRT

<213> Sphingobacterium sp.

<223> APG02675.1_RGN

<400> 116

Met Lys Arg Ile Leu Gly Leu Asp Leu Gly Thr Ser Ser Ile Gly Trp

1 5 10 15

Ala Tyr Val Leu Glu Ala Gln Asp Glu Asn Glu Ile Ser Lys Ile Glu

20 25 30

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

35 40 45

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

50 55 60

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

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

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

 Arg Lys Asp Gly Leu Thr Lys Gln Leu Ala Lys Glu Glu Val Ala Tyr
 260 265 270
 Ile Ile Ala Glu Ile Asn Gly Gln Leu Ala Asn Ser Ser Gly Tyr Leu
 275 280 285
 Gly Ala Ile Ser Asp Arg Ser Lys Glu Leu Tyr Phe Asn Lys Gln Thr
 290 295 300
 Val Gly Gln Phe Leu Phe Asn Gln Leu Gln Gln Asn Thr His Ala Arg

305	310							315							320				
Leu	Lys	Asn	Gln	Val	Phe	Tyr	Arg	Gln	Asp	Tyr	Leu	Asp	Glu	Phe	Glu				
325				330				335											
Lys	Ile	Trp	Cys	Ile	Gln	Ser	Gln	Phe	His	Gln	Ala	Leu	Thr	Glu	Gln				
340				345				350											
Leu	Lys	Glu	Glu	Ile	Arg	Asp	Ile	Ile	Ile	Phe	Tyr	Gln	Arg	Lys	Leu				
355				360				365											
Lys	Ser	Gln	Lys	Gly	Leu	Ile	Ser	Phe	Cys	Glu	Phe	Glu	Gln	His	Glu				
370				375				380											
Val	Val	Val	Asn	Gly	Lys	Asn	Lys	Val	Val	Gly	Leu	Arg	Val	Ala	Pro				
385				390				395				400							
Lys	Ser	Ser	Pro	Ile	Phe	Gln	Glu	Phe	Lys	Ile	Trp	Gln	Gln	Leu	Asn				
405				410				415											
Asn	Val	Lys	Leu	Arg	Asn	Asn	Ile	Thr	Lys	Glu	Val	Tyr	Asn	Leu	Ala				
420				425				430											
Glu	Glu	Gln	Lys	Lys	Leu	Leu	Phe	Glu	Thr	Leu	Asn	Leu	Lys	Gly	Lys				
435				440				445											
Leu	Ser	Ala	Tyr	Gln	Val	Leu	Ser	Leu	Ile	Glu	Gly	Lys	Pro	Lys	Glu				
450				455				460											
Trp	Glu	Leu	Asn	Tyr	Thr	Glu	Leu	Glu	Gly	Asn	Ser	Thr	Asn	Ser	Ala				
465				470				475				480							
Leu	Tyr	Asn	Ala	Tyr	Leu	Asn	Ile	Leu	Asp	Ile	Glu	Gly	Tyr	Asp	Val				
485				490				495											
Arg	Ser	Glu	Leu	Lys	Ile	Lys	Leu	Asn	Lys	Asp	Glu	Ile	Thr	Leu	Ser				
500				505				510											
Asp	Leu	Asp	Ile	Pro	Val	Ser	Glu	Ile	Lys	Asp	Met	Ile	Arg	Arg	Ile				
515				520				525											
Phe	Gln	His	Leu	Gly	Ile	Asn	Thr	Ser	Ile	Leu	Asp	Phe	Asp	Ala	Gln				
530				535				540											
Met	Gln	Gly	Asp	Asp	Phe	Glu	Lys	Gln	Leu	Ser	Tyr	Gln	Leu	Trp	His				
545				550				555				560							

Leu Leu Tyr Ser Tyr Glu Glu Asp Asp Ser Lys Thr Gly Met Asp Arg
 565 570 575

Leu Tyr Asn Gln Leu Asn Leu Lys Phe Gly Phe Ser Leu Asp Gln Ala
 580 585 590

Lys Phe Ile Gly Lys Ile Ala Leu Gln Asp Asp Tyr Gly Asn Leu Ser
 595 600 605

Thr Lys Ala Ile Arg Gln Ile Tyr Pro Tyr Ile Gln Asp Ala Glu Tyr
 610 615 620

Ser Ala Ala Cys Lys Leu Ala Gly Tyr Asn His Ser Lys Ser Ser Leu
 625 630 635 640

Thr Lys Glu Gln Leu Thr Asn Arg Val Leu Lys Asp His Leu Asp Ile
 645 650 655

Leu Pro Lys Asn Ser Leu Arg Asn Pro Val Val Glu Lys Ile Leu Asn
 660 665 670

Gln Met Val Asn Val Val Asn Thr Leu Ile Asp Thr Glu Asn Asp Lys
 675 680 685

Leu Ile Lys Glu Gly Lys Asn Ala Asp Phe Arg Phe Asp Glu Ile Arg
 690 695 700

Ile Glu Leu Ala Arg Glu Leu Lys Lys Asn Ala Lys Glu Arg Glu Glu
 705 710 715 720

Leu Thr Lys Ala Ile Asn Thr Ser Lys Glu Gln His Glu Lys Ile Ile
 725 730 735

Lys Ile Leu Gln Thr Glu Asp Gly Ile Lys Asn Pro Thr Arg Asn Asp
 740 745 750

Ile Ile Arg Phe Lys Leu Tyr Gln Glu Leu Lys Asn Asn Gly Tyr Lys
 755 760 765

Asn Leu Tyr Thr Asn Glu Tyr Ile Gln Arg Lys Asp Leu Phe Thr Asn
 770 775 780

Ile Tyr Asp Ile Asp His Ile Val Pro Gln Ser Arg Leu Phe Asp Asp
 785 790 795 800

Ser Phe Ser Asn Lys Val Leu Val Pro Arg Asn Ile Asn Ile Glu Lys

	805	810	815
Gly Asn Gln Thr Ala Phe Asp Tyr Val Arg Thr Lys Phe Gly Glu Asp			
	820	825	830
Gly Ile Glu Ala Tyr Glu Ala Arg Val Glu Arg Leu Phe Asn Leu Lys			
	835	840	845
Glu Glu Gly Val Ser Arg Ser Lys Tyr Lys Lys Leu Leu Met Arg Gly			
	850	855	860
Ser Asp Ile Gly Glu Gly Phe Ile Glu Arg Asp Leu Arg Asp Ser Gln			
865	870	875	880
Tyr Ile Ala Lys Lys Ala Lys Ala Met Leu Phe Glu Ile Thr Pro Ser			
	885	890	895
Val Ile Ser Thr Ser Gly Ser Val Thr Asp Arg Leu Arg Glu Asp Trp			
	900	905	910
Gly Leu Val Ser Val Met Lys Glu Leu Asn Leu Pro Lys Phe Lys Ala			
	915	920	925
Val Gly Leu Thr Glu Tyr Leu Glu Thr Lys Asp Gly Asn Arg Lys Glu			
	930	935	940
Val Ile Lys Asp Trp Ser Lys Arg Asn Asp His Arg His His Ala Met			
945	950	955	960
Asp Ala Leu Thr Val Ala Phe Thr Lys His Ser His Ile Gln Tyr Leu			
	965	970	975
Asn His Leu Asn Ala Arg Lys Asn Glu Lys Ser Glu Phe Phe Ser Ser			
	980	985	990
Ile Lys Ala Ile Glu Ile Lys Glu Thr Tyr Val Glu Arg Asp Asp Leu			
	995	1000	1005
Gly Asn Arg Lys Arg Leu Phe Lys Glu Pro Ile Ala His Phe Arg			
1010	1015	1020	
Thr Ile Val Lys Glu His Leu Glu Arg Val Leu Val Ser His Lys			
	1025	1030	1035
Ala Lys Asn Lys Val Val Thr Lys Asn Arg Asn Lys Ile Asp Gly			
	1040	1045	1050

Lys Lys	Glu Ala Gln Glu Val	Leu Thr Pro Arg Gly	Gln Leu His
1055	1060	1065	
Lys Glu	Thr Ile Tyr Gly Lys	Ile Leu Gln Tyr Ala	Ser Lys Glu
1070	1075	1080	
Glu Lys	Ile Ser Ala Lys Phe	Asp Ala Ala Thr Ile	Ala Met Val
1085	1090	1095	
Ser Asn	Pro Arg Tyr Arg Ser	Ala Leu Leu Asp Arg	Leu Ile Glu
1100	1105	1110	
Tyr Gly	Asn Asp Pro Lys Lys	Ala Phe Thr Gly Lys	Asn Ser Pro
1115	1120	1125	
Ser Lys	Ser Pro Ile Tyr Leu	Asp Gln Gln Arg Gln	Leu Ala Val
1130	1135	1140	
Pro Glu	Lys Val Lys Leu Val	Trp Leu Glu Ala Asp	Tyr Thr Ile
1145	1150	1155	
Arg Lys	Glu Ile Gly Pro Asp	Leu Lys Ile Glu Lys	Val Ile Asp
1160	1165	1170	
Gln Gly	Val Lys Arg Ile Leu	Lys Gln Arg Leu Ser	Val Tyr Asp
1175	1180	1185	
Gly Asp	Thr Lys Lys Ala Phe	Ser Asp Leu Asp Asn	Asn Pro Ile
1190	1195	1200	
Trp Leu	Asn Glu Glu Lys Gly	Ile Ala Ile Lys Arg	Val Thr Ile
1205	1210	1215	
Ser Gly	Val Lys Asn Ala Glu	Ala Leu His Thr Lys	Lys Asp His
1220	1225	1230	
Leu Gly	Gln Val Leu Glu Asp	Lys Asp Gly Asn Thr	Met Pro Val
1235	1240	1245	
Asp Phe	Val Ser Thr Gly Asn	Asn His His Val Ala	Ile Tyr Glu
1250	1255	1260	
Asp Ser	Glu Gly Asn Leu Tyr	Asp Arg Val Val Ser	Phe Tyr Glu
1265	1270	1275	
Ala Val	Glu Arg Ile Asn Gln	Asn Leu Pro Ile Val	Asp Arg His

1280 1285 1290
Tyr Lys Gln Glu Glu Gly Trp Lys Phe Leu Phe Ser Met Lys Gln
1295 1300 1305
Asn Glu Leu Phe Val Phe Pro Ser Ile Asn Phe Asp Pro Lys Glu
1310 1315 1320

Ile Asp Leu Leu Asp Lys Lys Asn Tyr Lys Glu Val Ser Lys His
1325 1330 1335
Leu Phe Arg Met Gln Lys Phe Ser Lys Val Glu Tyr Gly Asn Ser
1340 1345 1350
Ala Val Arg Asp Tyr Val Phe Arg His His Leu Glu Thr Ser Ile
1355 1360 1365
Ile Asp Thr Lys Glu Leu Arg Asp Ile Ala Tyr Lys Val Phe Lys
1370 1375 1380

Ser Ile Gly Glu Phe Lys Ser Phe Val Lys Ile Arg Thr Asn His
1385 1390 1395
Leu Gly Glu Ile Val His Leu Gly Glu Tyr
1400 1405

<210> 117

<211> 1369

<212> PRT

<213> Chryseobacterium sp.

<223> APG01405.1_RGN

<400> 117

Ala Arg Thr Asp Tyr Arg Gly Ile Arg Arg Leu Arg Glu Arg Phe Leu
1 5 10 15
Leu Arg Arg Glu Arg Leu His Arg Ile Leu Asn Val Leu Asn Phe Leu

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

65 70 75 80
 Phe Lys Ala His Gly Gln Asp Leu Lys Val Pro Tyr Asp Trp Thr Ile

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

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

 210 215 220
 Asp Trp Thr Leu Val Lys Thr Lys Thr Glu Gln Glu Ile Asp Arg Ser
 225 230 235 240
 Arg Lys Thr Val Gly Thr Tyr Ile Tyr Glu Thr Leu Leu Gln Asn Pro
 245 250 255
 Lys Gln Lys Ile Lys Gly Lys Leu Val Arg Thr Ile Glu Arg Lys Phe
 260 265 270
 Tyr Lys Glu Glu Leu Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe His

 275 280 285
 Tyr Glu Leu Gln Ser Asp Asp Leu Tyr Asp Asp Cys Ile Arg Glu Leu
 290 295 300
 Tyr Arg Asn Asn Glu Ala His Gln Leu Thr Leu Ser Lys Lys Asp Phe
 305 310 315 320

Val His Leu Phe Met Glu Asp Ile Ile Phe Tyr Gln Arg Pro Leu Lys
325 330 335

Ser Gln Lys Ser Ser Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr
340 345 350

Lys Asp Glu Asn Gly Val Glu His Thr Gln Phe Leu Lys Ala Ile Pro
355 360 365

Lys Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Ile Trp Gln Trp Ile Phe
370 375 380

Asn Leu Asn Ile Tyr Lys Lys Asp Asn Asp Asp Asn Val Thr Lys Glu
385 390 395 400

Phe Leu Ser Thr Thr Glu Asp Phe Glu Asn Leu Phe Glu Phe Leu Asn
405 410 415

Asn Arg Lys Glu Ile Asp Gln Lys Ala Leu Leu Lys His Phe Lys Leu
420 425 430

Asn Glu Lys Thr His Arg Trp Lys Tyr Val Glu Asp Lys Lys Tyr Pro
435 440 445

Cys Asn Glu Thr Lys Thr Met Ile Ser Glu Arg Leu Lys Lys Val Glu
450 455 460

Asn Ile Ser Asp Asp Phe Leu Thr Arg Gly Met Glu Gln Lys Ile Trp
465 470 475 480

His Ile Ile Tyr Ser Val Asn Asp Lys Thr Glu Tyr Glu Lys Ala Leu
485 490 495

Lys Ser Phe Ala Glu Lys Asn Asn Leu Asp Glu Asn Ser Phe Phe Glu
500 505 510

Ala Phe Arg Lys Phe Pro Pro Phe Lys Ser Glu Tyr Gly Ser Phe Ser
515 520 525

Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu Met Arg Leu Gly Lys Tyr
530 535 540

Trp Ser Tyr Ala Asn Ile Asp Leu Tyr Ser Lys Asn Arg Ile Gln Lys
545 550 555 560

Ile Ile Thr Gly Glu Phe Asp Glu Asn Ile Lys Asp Arg Val Arg Glu

565 570 575
 Lys Ala Ile His Leu Thr Ala Glu Asn Asp Phe Gln Gly Leu Gln Leu
 580 585 590
 Trp Leu Ala Gln Tyr Ile Val Tyr Gly Arg His Ser Glu Ala Thr Met

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

 660 665 670
 Lys Gln Thr Lys Asp Glu Arg Ala Asn Ala Thr Lys Thr Ile Thr Glu
 675 680 685
 Asn Glu Asn Thr Asn Leu Arg Ile Lys Ala Leu Leu Ala Glu Met Met
 690 695 700
 Asn Asp His Ser Val Glu Asn Val Arg Pro Tyr Ser Pro Met Gln Gln
 705 710 715 720
 Glu Ile Leu Lys Ile Tyr Glu Asp Gly Ile Leu Lys Ser Asp Ile Glu

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

 785 790 795 800
 Leu Ser Asn Lys Ile Ile Cys Glu Ser Ala Val Asn Lys Leu Lys Asp
 805 810 815

Asn Tyr Ile Gly Leu Gly Phe Ile Lys Gln Phe Gly Gly Thr Ile Ile
 820 825 830
 Glu Cys Gly Phe Gly Lys Arg Val Lys Val Phe Lys Ala Glu Glu Tyr
 835 840 845
 Glu Glu Phe Val Lys Lys His Tyr Ala Asn Asn Arg Gly Lys Arg Asn
 850 855 860
 Lys Leu Leu Leu Glu Asp Ile Pro Glu Lys Met Ile Glu Arg Gln Met
 865 870 875 880
 Asn Asp Thr Arg His Ile Ser Lys Tyr Ile Ser Gly Ile Leu Ser His
 885 890 895
 Ile Val Arg Val Glu Asp Gly Thr Asp Glu Gly Val Asn Ser Lys Asn
 900 905 910
 Val Ile Pro Gly Asn Gly Lys Ile Thr Thr Gln Leu Arg Gln Asp Trp
 915 920 925
 Gly Leu Asn Asp Val Trp Asn Glu Leu Val Leu Pro Arg Phe Glu Arg
 930 935 940
 Met Asn Gln Leu Thr Asn Ser Thr Asp Phe Thr Ser Trp Asn Glu Asn
 945 950 955 960
 His Gln Lys Tyr Leu Pro Thr Val Pro Val Glu Phe Ser Lys Gly Phe
 965 970 975
 Ser Lys Lys Arg Ile Asp His Arg His His Ala Leu Asp Ala Leu Val
 980 985 990
 Ile Ala Cys Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser
 995 1000 1005
 Ala Lys Ser Asp Thr Lys Arg Tyr Asp Leu Lys Lys Lys Leu Met
 1010 1015 1020
 Lys Phe Glu Lys Thr Val Tyr Lys Asp Pro Gln Thr Glu Lys Arg
 1025 1030 1035
 Ile Glu Arg Glu Val Pro Lys Tyr Phe Leu Lys Pro Trp Glu Thr
 1040 1045 1050
 Phe Thr Val Asp Ala Lys Asn Lys Leu Glu Thr Ile Ile Val Ser

1055	1060	1065
Phe Lys Gln Asn Leu Arg Val	Ile Asn Lys Ala Thr	Asn Tyr Tyr
1070	1075	1080
Glu Lys Tyr Val Asp Lys Asp	Gly Val Lys Thr Lys	Glu Arg Val
1085	1090	1095
Glu Gln Thr Gly Thr Asn Trp	Ala Ile Arg Lys Pro	Met His Lys

1100	1105	1110
Glu Thr Val Ser Gly Lys Ile	Asp Leu Pro Trp Val	Lys Val Pro
1115	1120	1125
Lys Gly Lys Ile Leu Thr Ala	Thr Arg Lys Ser Leu	Asp Thr Ser
1130	1135	1140
Phe Asp Leu Lys Ala Ile Ala	Ser Ile Thr Asp Thr	Gly Ile Gln
1145	1150	1155
Lys Ile Leu Lys Asn Tyr Leu	Glu Phe Lys Glu Ser	Pro Glu Leu

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

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

1280	1285	1290

Ser Ile Glu Ile Lys Asn Ile Thr Lys Glu Lys Ser Glu Arg Phe
 1295 1300 1305
 Tyr Lys Val Val Ser Phe Ser Gly Ser Gln Ile Phe Phe Val Arg
 1310 1315 1320
 His Asp Ile Ala Val Ser Ile Val Asn Lys Gly Glu Phe Ser Thr
 1325 1330 1335
 Leu Asn Lys Met Glu Arg Ala Ile Asp Gly Ser Met Val Lys Glu

1340 1345 1350
 Ser Cys Ile Lys Leu Lys Ile Asp Arg Leu Gly Asn Val Leu Lys
 1355 1360 1365
 Ala

<210> 118

<211> 1440

<212> PRT

<213> Chryseobacterium sp.

<223> APG06250.1_RGN

<400> 118

Met Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly Trp
 1 5 10 15
 Ala Leu Val Lys Gln Asp Phe Lys Asn Lys Gln Gly Glu Ile Leu Gly

20 25 30
 Met Gly Thr Arg Ile Ile Pro Met Ser Gln Asp Ile Leu Gly Asp Phe
 35 40 45

Gly Lys Gly Asn Ser Val Ser Gln Thr Ala Glu Arg Thr Lys Tyr Arg
 50 55 60

Ser Ala Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu Arg Leu
 65 70 75 80

His Arg Val Leu Asn Ile Leu Asn Phe Leu Pro Glu His Tyr Ala Ser

85 90 95
 Gln Ile Asp Phe Glu Lys Arg Phe Gly Lys Phe Lys Val Glu Thr Glu
 100 105 110

Pro Lys Leu Ala Trp Lys Asn Ile Glu Gly Gln Phe Ser Phe Leu Phe
115 120 125

Gln Asn Ser Phe Asn Glu Met Leu Glu Asp Phe Lys Ala Asn Ala Gln
130 135 140

Asp Leu Arg Ile Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu Arg Lys Lys

145 150 155 160

Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu Leu Val Trp Ile Leu Leu
165 170 175

Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu Glu Glu
180 185 190

Glu Glu Asn Pro Asn Lys Leu Val Glu Leu Tyr Ser Leu Lys Val Val
195 200 205

Asp Val Phe Val Asp Glu Pro Gln Lys Gly Lys Ser Asp Ile Trp Tyr

210 215 220

Ser Leu Val Leu Glu Asn Gly Trp Ile Tyr Arg Arg Ala Ser Lys Ile
225 230 235 240

Ser Leu Phe Asp Trp Lys Asp Lys Ile Arg Asp Phe Ile Val Thr Thr
245 250 255

Asp Leu Asn Asp Asp Gly Ser Val Lys Thr Asp Lys Asp Gly Asn Glu
260 265 270

Lys Arg Ser Phe Arg Ala Pro Gly Glu Asn Asp Trp Thr Leu Val Lys

275 280 285

Lys Lys Thr Glu Gln Glu Ile Glu Gln Ser His Lys Thr Val Gly Met
290 295 300

Tyr Ile Tyr Glu Thr Leu Arg Ala Asn Pro Lys Gln Lys Ile Lys Gly
305 310 315 320

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

Gln Ile Leu Glu Lys Gln Lys Glu Phe His Gln Glu Leu Gln Asn Asp

340 345 350

Asp Leu Tyr Asn Asp Cys Ile Arg Glu Leu Tyr Arg Asn Asn Glu Ala

355 360 365
 His Gln Leu Thr Leu Ser Lys Lys Asp Phe Val His Leu Phe Ile Glu
 370 375 380
 Asp Ile Val Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser Ser Ile
 385 390 395 400
 Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr Lys Asp Glu Asn Gly Ala

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

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

 530 535 540
 Asn Asp Lys Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Leu Lys
 545 550 555 560
 His Asn Leu Asp Glu Asn Ser Phe Phe Glu Ala Phe Arg Lys Phe Pro
 565 570 575
 Pro Phe Lys Ser Glu Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys Lys
 580 585 590
 Leu Leu Pro Leu Met Arg Leu Gly Lys Tyr Trp Asp Tyr Ala Asn Ile

 595 600 605

Asp Gln Phe Ser Lys Gly Arg Ile Gln Lys Ile Ile Asn Gly Glu Tyr
 610 615 620
 Asp Glu Asn Ile Lys Asp Arg Val Arg Glu Lys Ala Val His Leu Thr
 625 630 635 640
 Ser Glu Asn Asp Phe Gln Gly Leu Gln Leu Trp Leu Ala Gln Tyr Ile
 645 650 655
 Val Tyr Gly Arg His Ser Glu Ala Ser Ile Ala Gly Lys Trp Asn Ser
 660 665 670
 Ala Asp Asp Leu Glu Glu Phe Leu Lys Asp Phe Lys His His Ser Leu
 675 680 685
 Arg Asn Pro Ile Val Glu Gln Val Ile Thr Glu Thr Leu Arg Val Val
 690 695 700
 Lys Asp Ile Trp Leu Lys Tyr Gly Lys Gly Ala Lys Asp Phe Phe Asn
 705 710 715 720
 Glu Ile His Ile Glu Leu Gly Arg Glu Met Lys Leu Pro Ala Asp Asp
 725 730 735
 Arg Lys Lys Leu Thr Asn Gln Ile Thr Glu Asn Glu Asn Thr Asn Leu
 740 745 750
 Arg Ile Lys Ala Leu Leu Ala Glu Met Met Asn Asp His Gly Val Glu
 755 760 765
 Asn Val Arg Pro Phe Ser Pro Gln Gln Gln Glu Ile Leu Lys Ile Tyr
 770 775 780
 Glu Asp Gly Val Leu Asn Ser Asp Ile Glu Ile Glu Asp Glu Tyr Leu
 785 790 795 800
 Lys Ile Ser Lys Thr Ala Gln Pro Ser Pro Ser Asp Leu Lys Arg Tyr
 805 810 815
 Lys Leu Trp Leu Glu Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln Ile
 820 825 830
 Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile
 835 840 845
 Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser Phe Ser Asn Lys Ile Ile

850 855 860
 Cys Glu Ser Ala Val Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu Gly
 865 870 875 880
 Phe Ile Lys Gln Phe Gly Gly Thr Ile Ile Glu Leu Gly Phe Gly Lys
 885 890 895
 Ser Val Lys Val Phe Asp Ile Asp Glu Tyr Glu Asp Phe Val Lys Lys
 900 905 910
 His Tyr Ala Asn Asn Arg Ser Lys Arg Asn Lys Leu Leu Leu Glu Asp

 915 920 925
 Ile Pro Glu Lys Met Ile Glu Arg Gln Ile Asn Asp Thr Arg Tyr Ile
 930 935 940
 Ser Lys Tyr Ile Ser Gly Ile Leu Ser Asn Ile Val Arg Leu Glu Asp
 945 950 955 960
 Gly Ser Asp Glu Gly Ile Asn Ser Lys Asn Ile Val Pro Gly Asn Gly
 965 970 975
 Lys Ile Thr Thr Gln Leu Lys Lys Asp Trp Gly Leu Asn Asp Val Trp

 980 985 990
 Asn Asp Leu Ile Leu Pro Arg Phe Lys Arg Met Asn Gln Leu Thr Asn
 995 1000 1005
 Ser Thr Asp Phe Ile Ala Trp Asn Glu Asn Phe Gln Lys Phe Leu
 1010 1015 1020
 Pro Thr Val Pro Ile Glu Tyr Ser Lys Gly Phe Ser Lys Lys Arg
 1025 1030 1035
 Ile Asp His Arg His His Ala Leu Asp Ala Leu Val Ile Ala Cys

 1040 1045 1050
 Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser Ala Lys
 1055 1060 1065
 Ser Glu Thr Lys Arg Tyr Asp Leu Lys Lys Lys Leu Met Lys Phe
 1070 1075 1080
 Glu Arg Val Val Tyr Leu His Thr Gln Thr Gly Glu Lys Ile Glu
 1085 1090 1095

Arg Glu Val Pro Lys His Phe Leu Lys Pro Trp Glu Asn Phe Thr

1100

1105

1110

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

1115

1120

1125

Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr Glu Lys

1130

1135

1140

Tyr Ala Glu Asn Gly Gly Ala Arg Asn Lys Ala Arg Val Glu Gln

1145

1150

1155

Lys Gly Ile Asn Trp Ala Ile Arg Lys Pro Met His Lys Asp Thr

1160

1165

1170

Val Ser Gly Lys Val Asp Leu Pro Trp Val Lys Val Pro Lys Gly

1175

1180

1185

Lys Ile Leu Thr Ala Thr Arg Lys Ser Leu Asp Thr Ser Phe Asp

1190

1195

1200

Leu Lys Ser Ile Gly Ser Ile Thr Asp Thr Gly Ile Gln Lys Ile

1205

1210

1215

Leu Lys Asn Tyr Leu Met Leu Lys Asp Gly Asn Pro Glu Leu Ala

1220

1225

1230

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

1235

1240

1245

Tyr Asn Asp Gly Lys Pro His Gln Pro Ile Asn Lys Val Arg Val

1250

1255

1260

Phe Glu Leu Gly Ser Lys Phe Gln Ile Gly Gln Ala Gly Asn Lys

1265

1270

1275

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

1280

1285

1290

Ala Val Tyr Glu Asp Lys Asn Gly Lys Arg Ser Tyr Glu Thr Ile

1295

1300

1305

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

1310

1315

1320

Val Asp Leu Arg Ser Thr Asn Asp Phe Tyr Leu Cys Pro Asn Asp

1325 1330 1335
Leu Val Tyr Ile Leu Ser Asp Asp Glu Arg Gly Asn Met Thr Asn

1340 1345 1350
Asn Asp Phe Glu Asn Leu Ser Asp Glu Gln Val Lys Arg Ile Tyr

1355 1360 1365
Lys Phe Val Ser Cys Thr Gly Gly Glu Gly His Phe Ile Pro Tyr

1370 1375 1380
Ala Asn Ala Thr Glu Ile Ile Lys Asn Glu Asn Gly Thr Asn Ser

1385 1390 1395
Lys Ser Glu Arg Met Gln Asn Phe Tyr Asp Gly Ser Cys Met Leu

1400 1405 1410
Asp Lys Asn Ser Lys Pro Ile Met Ile Lys Glu Asn Cys Val Lys

1415 1420 1425
Leu Lys Val Asp Arg Leu Gly Asn Ile Ser Lys Ile

1430 1435 1440

<210> 119

<211> 1071

<212> PRT

<213> Bacillus sp.

<223> APG06877.1_RGN

<400> 119

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Gln
20 25 30

Tyr Glu Lys Thr Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45

Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80

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

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

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

 Gln Tyr Phe Asn Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
 275 280 285
 Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Ile Ile Leu Asp Lys
 290 295 300
 Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
 305 310 315 320
 Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln

				325				330				335			
Phe	Lys	Gly	Leu	Thr	Tyr	Asp	Pro	Asn	Glu	Glu	Leu	Lys	Lys	Ile	Glu
340				345				350							
Met	Lys	Pro	Phe	Ile	Asn	Leu	Lys	Pro	Phe	Tyr	Glu	Ile	Lys	Lys	Val
355				360				365							
Val	Thr	Asn	Tyr	Ala	Lys	Lys	Thr	Asn	Glu	Val	Phe	Ser	Ala	Leu	Asp
370				375				380							
Tyr	Asp	Thr	Val	Ala	Tyr	Ala	Leu	Thr	Val	Tyr	Lys	Thr	Asp	Lys	Asp
385				390				395				400			
Ile	Arg	Ser	Tyr	Leu	Lys	Arg	Ser	Asn	Asn	Leu	Ser	Lys	Arg	Cys	Tyr
405				410				415							
Asp	Asp	Gln	Leu	Ile	Glu	Glu	Leu	Leu	Thr	Leu	Ser	Tyr	Thr	Lys	Phe
420				425				430							
Gly	His	Leu	Ser	Phe	Lys	Ala	Ile	Asn	His	Val	Leu	Pro	Ile	Met	Gln
435				440				445							
Glu	Gly	Arg	Thr	Tyr	Gln	Glu	Ala	Ile	His	Gln	Leu	Gly	Tyr	Asp	Thr
450				455				460							
Thr	Asn	Leu	Lys	Lys	Glu	Asn	Arg	Ser	Met	Phe	Leu	Pro	Ile	Ile	Pro
465				470				475				480			
Asp	Glu	Ile	Thr	Asn	Pro	Ile	Val	Lys	Arg	Ala	Leu	Thr	Gln	Ala	Arg
485				490				495							
Lys	Val	Val	Asn	Ala	Ile	Ile	Arg	Arg	Tyr	Gly	Ser	Pro	Asn	Ser	Val
500				505				510							
His	Ile	Glu	Leu	Ala	Arg	Glu	Leu	Ser	Lys	Ser	His	Asp	Glu	Arg	Lys
515				520				525							
Lys	Ile	Met	Thr	Ala	His	Asp	Glu	Asn	Tyr	Lys	Lys	Asn	Lys	Gly	Ala
530				535				540							
Ile	Ser	Ile	Leu	Ile	Glu	Asn	Gly	Ile	Leu	Asn	Pro	Thr	Gly	Tyr	Asp
545				550				555				560			
Ile	Val	Arg	Tyr	Lys	Leu	Trp	Lys	Glu	Gln	Gly	Glu	Arg	Cys	Ala	Tyr
565				570				575							

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

Glu Arg Ser Gly Pro Pro Val Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Gly
610 615 620

Asp Glu Asn Gln Lys Lys Gly Asn Arg Ile Pro Tyr Thr Phe Phe Ser
625 630 635 640

Glu Glu Asp Lys Glu Trp Glu Ser Phe Glu Ser Tyr Val Arg Ser Asn
645 650 655

Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr
660 665 670

Leu Pro Arg Glu Ser Asn Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685

Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys
690 695 700

Phe Lys Glu Ala Val Gly Ile Thr Arg Lys Lys Tyr Val Gln Thr Val
705 710 715 720

Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735

Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750

Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln
755 760 765

Ile Lys Glu Gly Asn Lys Ser Ile Lys Lys Pro Tyr Phe Pro Leu Pro
770 775 780

Trp Met Gly Phe Arg Glu Glu Ile Leu Ser His Leu Glu Ser Gln Pro
785 790 795 800

Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Ser
805 810 815

Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ser

820	825	830
Ala His Asp Gln Thr Val Met Lys Lys Gly Gly Ile Asp Lys Lys Gly		
835	840	845
Lys Thr Ile Ile Ile Lys Arg Val His Leu Lys Asp Ile Lys Phe Asp		
850	855	860
Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr		
865	870	875
Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu Tyr Arg Lys Glu Ser Lys		
885	890	895
Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly		
900	905	910
Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val		
915	920	925
Arg Glu Ile Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val		
930	935	940
Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Val Pro Ile Tyr Val		
945	950	955
Val Asp Thr Val Arg Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser		
965	970	975
Lys Gly Tyr Glu Gln Trp Leu Ser Ile Asp Asn Ser Phe Thr Phe Lys		
980	985	990
Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu		
995	1000	1005
Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg		
1010	1015	1020
Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Tyr		
1025	1030	1035
Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Ile		
1040	1045	1050
Gly Val Leu Gly Asp Leu Arg Leu Val Arg Gln Glu Thr Arg Lys		
1055	1060	1065

Ile Phe Lys

1070

<210> 120

<211> 1091

<212> PRT

<213> Rhizobium sp.

<223> APG09053.1_RGN

<400> 120

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

1 5 10 15

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

20 25 30

Ile Asp Ile Gly Thr Arg Ile Tyr Ala Asp Gly Arg Asp Pro Gln Ser

35 40 45

Lys Ala Ser Leu Ala Val Ala Arg Arg Glu Ala Arg Ala Met Ser Arg

50 55 60

Arg Arg Asp Arg Tyr Val Asn Arg Arg Lys Ala Val Leu Arg Thr Leu

65 70 75 80

Thr Glu Tyr Gly Leu Met Pro Ala Glu Ala Ser Ala Arg Gln Ala Leu

85 90 95

Ile Ala Glu Thr Ser Asp Arg Asn Gly Ala Ala Gly Glu Ala Ser Ala

100 105 110

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

115 120 125

Tyr Trp Ile Gly Arg Val Leu Phe His Leu Asn Gln Arg Arg Gly Phe

130 135 140

Lys Ser Asn Arg Lys Thr Asp Arg Lys Asp Asn Glu Lys Gly Val Ile

145 150 155 160

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

165 170 175

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

180 185 190

Thr Arg Leu Arg Ala Gly Ser Asp Ala Phe Glu Gly Asp Gly Tyr Ala
195 200 205

Phe Tyr Pro Glu Arg Ser Leu Leu Glu Ala Glu Phe Arg Glu Ile Trp
210 215 220

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

Asp His Leu Phe Gln Val Met Phe Tyr Gln Arg Pro Leu Lys Lys Pro
245 250 255

Arg Val Gly Lys Cys Ala Phe Asn Pro Leu Glu Glu Arg Leu Ala Lys
260 265 270

Ala His Pro Leu Phe Gln Glu Phe Arg Leu Tyr Lys Glu Val Asn Glu
275 280 285

Leu Glu Leu Val Leu Pro Asp His Ser His Lys Lys Leu Asp Ile Glu
290 295 300

Gln Arg Asn Ala Leu Val Ile Leu Leu Arg Ala Asn Arg Glu Val Thr
305 310 315 320

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

Asn Lys Glu Ser Glu Ser Arg Glu Lys Leu Arg Gly Asp Glu Val His
340 345 350

Ser Val Leu Ala Asp Lys Lys Arg Phe Gly Pro Gln Trp Gly Ala Leu
355 360 365

Ser Arg Glu Arg Gln Trp Gln Ile Ile Gln Thr Leu Lys Asp Glu Glu
370 375 380

Asn Pro Glu Lys Leu Phe Gly Trp Leu Lys Glu Glu Phe Gly Phe Glu
385 390 395 400

Gly Glu Lys Ala Asp Ala Ile Ala Asn Ala Pro Leu Pro Glu Gly Tyr
405 410 415

Gly Arg Leu Gly Glu Thr Ala Leu Ser Ser Met Leu Asp Glu Met Lys
420 425 430

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

435 440 445
 His Ser Lys Leu Gly Glu Asp Arg Glu Glu Gly Glu Ala Phe Leu Pro
 450 455 460

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

 Leu Gln Leu Ser Glu Lys Gln Lys Ala Glu Ala Asn Arg Ala Ile Ala
 530 535 540
 Lys Asn Thr Arg Glu Ala Ala Ala Arg Ser Lys Lys Leu Gly Glu Met
 545 550 555 560
 Gly Gln Leu Asp Thr Gly Tyr Asn Arg Leu Leu Leu Lys Leu Trp Glu
 565 570 575
 Glu Leu Asn Pro Ser Lys Pro Glu Asp Arg Val Cys Ile Tyr Ser Gly
 580 585 590

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

 Ala Ala Arg Leu Ala Pro Asn Lys Arg Arg Arg Phe Ala Pro Asn Ala
 660 665 670
 Met Ala Glu Phe Glu Ala Glu Gly Gly Phe Leu Ala Arg Gln Leu Thr
 675 680 685

Asp Thr Gln Tyr Leu Ser Arg Met Ala Arg Glu Tyr Leu Ser Cys Leu
690 695 700
Phe Pro Ser Ala Glu Ile Asp Arg Arg Gly Glu Arg Lys Lys Lys Ile
705 710 715 720

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

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

His Asn Asp Thr Ala Tyr Gly Phe Gly Thr Asp Ala Ser Gly Asn Pro
850 855 860
Val Ala Val Arg Arg Lys Leu Phe Thr Ser Leu Glu Ala Lys Asp Ile
865 870 875 880
Pro Met Ile Arg Asp Pro Glu Leu Gln Ala Glu Leu Tyr Ala Ala Ile
885 890 895
Asp Gly Leu Asp Asp Arg Arg Ala Leu Gln Asp Ala Leu Leu Ser Phe
900 905 910

Arg Arg Thr His Pro Lys Phe Lys Gly Ile Arg Arg Val Arg Met Ala
915 920 925
Glu Thr Leu Ser Phe Ile Pro Ile Arg Asp Ser Gln Gly Asn Val Tyr

930 935 940
 Lys Gly Tyr Lys Gly Asp Ala Asn Tyr Arg Tyr Asp Val Trp Glu Thr
 945 950 955 960
 Leu Asp Gly Lys Trp His Ala Asp Val Val Thr Met Phe Asn Ala His
 965 970 975

 Gln Pro Asp Trp Arg Ser Pro Val His Gln Glu His Pro Thr Ala Arg
 980 985 990
 Arg Val Leu Arg Leu Gln Gln Asn Asp Met Val Ala Tyr Glu His Pro
 995 1000 1005
 Asn Asp Gly Tyr Thr Ile Ala Arg Val Val Lys Phe Asn Thr Ala
 1010 1015 1020
 Gly Ile Val Tyr Phe Ala Ser His Arg Glu Ser Gly Ser Leu Lys
 1025 1030 1035

 Ala Arg Asp Ala Asp Lys Gln Asp Pro Phe Lys Tyr Phe Ser Lys
 1040 1045 1050
 Ser Ala Ala Gly Leu Lys Asp Ile Gln Cys Arg Gln Ile Arg Ile
 1055 1060 1065
 Asp Ala Ala Gly Arg Val Phe Asp Pro Gly Pro Gln Asp Arg Ala
 1070 1075 1080
 Ser Lys Ser Thr Arg Lys Thr Asn
 1085 1090

 <210> 121
 <211> 1401
 <212> PRT
 <213> Myroides sp.

 <223> APG04293.1_RGN
 <400> 121
 Met Lys Arg Ile Leu Gly Ile Asp Leu Gly Thr Thr Ser Ile Gly Trp
 1 5 10 15
 Ala Tyr Val His Glu Ala Glu Asn Glu Asn Glu Ile Ser Ser Ile Glu
 20 25 30
 Lys Leu Gly Val Arg Val Asn Pro Leu Thr Val Asp Glu Gln Ile Asn

35 40 45
 Phe Glu Lys Gly Lys Pro Ile Thr Thr Asn Ala Thr Arg Thr Leu Phe
 50 55 60

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

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

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

 Arg Ser Asp Ala Leu Ile Lys Gln Leu Asp Lys Glu Ile Val Ala Tyr
 260 265 270
 Val Ile Ser Glu Ile Asn Gly Gln Ile Asn Asn Ser Ser Gly Tyr Leu
 275 280 285

Gly Ala Ile Ser Asp Arg Ser Lys Glu Leu Tyr Phe Asn Gln Leu Thr
290 295 300

Val Gly Gln Tyr Leu His Arg Gln Leu Ile Ala Asn Pro His Thr Arg
305 310 315 320

Leu Lys Ser Gln Val Phe Tyr Arg Gln Asp Tyr Leu Asp Glu Phe Glu
325 330 335

Lys Ile Trp Glu Thr Gln Ser Lys Phe Tyr Pro Lys Glu Leu Thr Thr
340 345 350

Val Leu Lys Asn Glu Ile Arg Asp Thr Val Ile Phe Tyr Gln Arg Lys
355 360 365

Leu Lys Ser Gln Lys Gly Leu Ile Ser Phe Cys Glu Phe Glu Gln Glu
370 375 380

Gln Lys Ile Ile Asn Gly Lys Thr Lys Thr Ile Gly His Arg Val Ile
385 390 395 400

Pro Lys Ser Ser Pro Leu Phe Gln Glu Phe Lys Ile Trp Gln Gln Leu
405 410 415

His Asn Val Val Leu Arg Asn Lys Lys Thr Asn Val Ile Thr Pro Leu
420 425 430

Ala Glu Glu Gln Lys Ser His Leu Phe Glu Glu Leu Asn Leu Lys Gly
435 440 445

Lys Leu Ser Ser Thr Gln Ile Leu Lys Leu Ile Glu Asp Lys Pro Lys
450 455 460

Asp Trp Glu Leu Asn Tyr Ser Glu Leu Glu Gly Asn Asn Thr Asn Lys
465 470 475 480

Ala Leu Tyr Asn Ala Tyr Leu Asp Ile Leu Asp Leu Glu Gly Tyr Asp
485 490 495

Ile Arg Asp Glu Leu Lys Ile Lys Leu Asn Lys Asp Asp Ile Glu Leu
500 505 510

Ser Asp Leu Asp Val Asn Ala Ser Glu Ile Lys Glu Met Ile Phe Ser
515 520 525

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

530 535 540
 Thr Leu Lys Asp Lys Ala Phe Glu Lys Gln Ala Ser Tyr Gln Leu Trp
 545 550 555 560
 His Leu Leu Tyr Ser Tyr Glu Glu Asp Asn Ser Pro Thr Gly Leu Asp
 565 570 575

 Arg Leu His Asn Leu Leu Gln Lys Lys Phe Asn Phe Thr Leu Asp Gln
 580 585 590
 Ala Lys Leu Val Gly Asn Val Leu Phe Gln Asp Asp Tyr Gly Asn Leu
 595 600 605
 Ser Ser Lys Ala Ile Lys Asn Ile Phe Pro Tyr Ile Thr Asp Asn Asn
 610 615 620
 Tyr Ser Thr Ala Cys Glu Leu Ala Gly Tyr Lys His Ser Lys His Ser
 625 630 635 640

 Leu Thr Lys Glu Glu Asn Glu Lys Arg Glu Leu Lys Ser Arg Leu Thr
 645 650 655
 Ile Leu Thr Lys Asn Ser Leu Arg Asn Pro Val Val Glu Lys Ile Leu
 660 665 670
 Asn Gln Met Ile Asn Val Ile Asn Thr Leu Ile Glu Ser Glu Asn Asp
 675 680 685
 Lys Leu Val Ala Gln Gly Lys Glu Pro Asn Phe Gln Phe Asp Glu Ile
 690 695 700

 Arg Ile Glu Leu Ala Arg Glu Leu Lys Lys Asn Ala Lys Glu Arg Glu
 705 710 715 720
 Glu Leu Thr Lys Ser Met Thr Gln Gly Lys Thr Asn His Glu Lys Ile
 725 730 735
 Ile Lys Ile Leu Gln Lys Glu Asp Gly Ile Lys Asn Pro Thr Arg Asn
 740 745 750
 Asp Ile Thr Arg Phe Lys Leu Tyr Ser Glu Leu Lys Asn Asn Gly Tyr
 755 760 765

 Lys Asp Leu Tyr Thr Asn Glu Tyr Ile Glu Arg Lys Asp Ile Phe Ser
 770 775 780

Lys Glu Tyr Asp Ile Glu His Ile Ile Pro Gln Ser Lys Leu Phe Asp
 785 790 795 800
 Asp Ser Phe Ser Asn Lys Thr Leu Val Arg Arg Asn Val Asn Leu Lys
 805 810 815
 Lys Gly Asn Gln Thr Ala Tyr Asp Phe Ile Leu Ser Glu Tyr Gly Gln
 820 825 830

 Glu Lys Ala Asn Glu Phe Glu Thr Arg Ile Thr Asn Leu Tyr Ser Leu
 835 840 845
 Gly Lys Asp Glu Gly Ile Ser Lys Ser Lys Tyr Lys Lys Leu Leu Met
 850 855 860
 Gln Glu Ser Glu Ile Gly Lys Gly Phe Ile Glu Arg Asp Leu Arg Glu
 865 870 875 880
 Thr Gln Tyr Ile Ala Lys Lys Ala Lys Ser Leu Leu Leu Glu Ile Thr
 885 890 895

 Arg Asn Val Val Ser Thr Ser Gly Gly Ile Thr Asp Arg Leu Arg Glu
 900 905 910
 Asp Trp Gly Leu Val Asn Val Met Lys Glu Leu Asn Leu Lys Lys Phe
 915 920 925
 Arg Asp Ala Gly Leu Thr Glu Phe Val Glu Met Lys Asp Gly Asn Lys
 930 935 940
 Lys Glu Val Ile Ile Asp Trp Thr Lys Arg Asn Asp His Arg His His
 945 950 955 960

 Ala Met Asp Ala Leu Thr Ile Ala Phe Thr Lys His Asn His Ile Gln
 965 970 975
 Tyr Leu Asn His Leu Asn Ala Arg Lys Asn Thr Lys Asp Glu Leu His
 980 985 990
 Thr Asn Ile Ile Ala Ile Glu Lys Lys Glu Thr Thr Ile Ile Thr Asp
 995 1000 1005
 Glu Arg Gly Asn Lys Lys Arg Ile Phe Lys Glu Pro Ile Pro Asn
 1010 1015 1020

 Phe Arg Asn Ile Ala Lys Ser His Leu Glu Ala Ile Leu Val Ser

1025	1030	1035
His Lys Ala Lys Asn Lys Val	Val Thr Lys Asn Ile	Asn Lys Ile
1040	1045	1050
Ser Gly Lys Lys Ile Gly Gln	Gln Thr Leu Thr Pro	Arg Gly Gln
1055	1060	1065
Leu His Lys Glu Thr Val Tyr	Gly Lys Ile Arg Gln	Tyr Val Ser
1070	1075	1080
Lys Glu Glu Lys Val Gly Pro	Lys Phe Thr Lys Glu	Val Ile Glu
1085	1090	1095
Leu Val Ser Asn Pro Thr Tyr	Arg Asp Leu Leu Leu	Lys Arg Leu
1100	1105	1110
Gln Glu Asn Asn Asn Asp Pro	Lys Lys Ala Phe Gly	Gly Lys Asn
1115	1120	1125
Ala Leu Ser Lys Ser Pro Ile	Tyr Ile Asn Leu Glu	Lys Asn Ile
1130	1135	1140
Ile Val Pro Glu Val Val Lys	Leu Val Trp Leu Glu	Asp Asp Tyr
1145	1150	1155
Thr Ile Arg Lys Glu Ile Ser	Pro Glu Leu Lys Ile	Asp Lys Val
1160	1165	1170
Val Asp Glu Gly Val Lys Arg	Ile Leu Gln Asn Arg	Leu Ala Leu
1175	1180	1185
Tyr Asn Gly Asp Ala Lys Ile	Ala Phe Ser Asp Leu	Asp Asn Asn
1190	1195	1200
Pro Ile Trp Leu Ser Lys Asp	Lys Gly Ile Thr Ile	Lys Arg Val
1205	1210	1215
Thr Leu Ser Gly Val Lys Asn	Ala Thr Ala Leu His	Thr Lys Lys
1220	1225	1230
Asp His Leu Gly Gln Pro Ile	Leu Asp His Lys Gly	Gln Glu Ile
1235	1240	1245
Pro Val Asp Phe Ile Ser Thr	Gly Asn Asn His His	Val Ala Ile
1250	1255	1260

Tyr Glu Asp Lys Asn Gly Lys Leu Gln Glu Ser Val Val Ser His
1265 1270 1275

Phe Glu Ala Val Glu Arg Val Asn Gln Gln Leu Pro Ile Ile Asp
1280 1285 1290

Lys Thr Phe Asn Gln His Leu Gly Trp Lys Phe Leu Phe Thr Met
1295 1300 1305

Lys Gln Asn Glu Leu Phe Leu Phe Pro Ser Asp Asp Phe Asn Pro
1310 1315 1320

Asn Glu Ile Asp Leu Phe Asp Thr Arg Asn Tyr Ala Val Ile Ser
1325 1330 1335

Lys His Leu Phe Arg Ala Gln Lys Leu Ser Thr Lys Asp Tyr Ser
1340 1345 1350

Phe Arg His His Leu Glu Thr Asn Val Glu Asp Ile Lys Asp Leu
1355 1360 1365

Lys Thr Ile Thr Trp Arg Arg Glu Gly Leu Asn Gly Ile Gln Asn
1370 1375 1380

Ile Val Lys Ile Arg Thr Asn His Leu Gly Glu Ile Val His Ile
1385 1390 1395

Gly Glu Tyr
1400

<210> 122

<211> 1432

<212> PRT

<213> Chryseobacterium sp.

<223> APG01308.1_RGN

<400> 122

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

1 5 10 15

Trp Ala Leu Ile Asn Gln Asp Phe Glu Asn Lys Gln Gly Lys Ile Leu

20 25 30

Gly Met Gly Ser Arg Ile Ile Pro Met Ser Gln Asp Ile Leu Gly Asp

35 40 45

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

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

 Lys Lys Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp Ile
 165 170 175
 Leu Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu
 180 185 190
 Asp Phe Glu Glu Glu Lys Asp Lys Met Phe Val Arg Leu Lys Val Glu
 195 200 205
 Lys Ile Ile Asp Ser Gly Asp Asn Val Lys Gly Lys Ile Leu Tyr Asp
 210 215 220

 Val Tyr Phe Glu Asn Gly Trp Lys Tyr Asp Lys Gln Ile Val Lys Thr
 225 230 235 240
 Glu Asp Trp Ile Glu Arg Val Lys Glu Phe Ile Val Thr Glu Ser Phe
 245 250 255
 Leu Lys Asn Gly Asp Ile Lys Arg Thr Phe Lys Ala Val Asp Ser Glu
 260 265 270
 Lys Asp Trp Ile Ala Ile Lys Thr Lys Thr Glu Gln Glu Ile Asp Lys
 275 280 285

 Ser His Lys Thr Val Gly Val Tyr Ile Tyr Glu Thr Leu Leu Gln Asn

290 295 300
 Pro Lys Gln Lys Ile Lys Gly Lys Leu Val Arg Thr Ile Glu Arg Lys
 305 310 315 320
 Phe Tyr Lys Asp Glu Leu Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe
 325 330 335
 His Gln Glu Leu Lys Asn Asp Asp Leu Tyr Asn Asp Cys Val Arg Glu
 340 345 350

 Leu Tyr Arg Asn Asn Glu Ala His Gln Leu Thr Leu Ser Lys Lys Glu
 355 360 365
 Phe Val His Leu Leu Met Glu Asp Ile Ile Phe Tyr Gln Arg Pro Leu
 370 375 380
 Arg Ser Gln Lys Ser Ser Ile Ser Asn Cys Ser Leu Glu Phe Arg Lys
 385 390 395 400
 Tyr Lys Asp Glu Asn Gly Val Glu His Ile Gln Tyr Leu Lys Ala Val
 405 410 415

 Pro Lys Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Ile Trp Gln Trp Ile
 420 425 430
 Phe Asn Leu Asn Ile Tyr Lys Arg Asp Asp Glu Gln Asn Pro Val Thr
 435 440 445
 Thr Glu Phe Leu Asn Thr Thr Thr Asp Ile Glu Asn Leu Phe Glu Phe
 450 455 460
 Leu Asn Asn Arg Lys Glu Val Asp Gln Lys Ala Leu Leu Lys His Phe
 465 470 475 480

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

Ala Leu Lys Ser Phe Ala Lys Lys Asn Asn Leu Asp Glu Asn Ser Phe
545 550 555 560
Phe Glu Ala Phe Lys Lys Phe Pro Pro Phe Lys Ser Glu Tyr Gly Ser
565 570 575
Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu Met Arg Leu Gly
580 585 590
Lys Tyr Trp Asn Cys Glu Asn Ile Ser Asp Asp Ser Lys Glu Arg Ile
595 600 605

Gln Lys Ile Ile Asn Gly Glu Tyr Asp Glu Asn Ile Lys Asp Arg Val
610 615 620
Arg Glu Lys Ala Ile His Leu Thr Ser Glu Asn Asn Phe Gln Gly Leu
625 630 635 640
Gln Leu Trp Leu Ala Gln Tyr Val Val Tyr Asp Arg His Ser Glu Ala
645 650 655
Ser Met Ile Gly Lys Trp Asn Ser Ala Asp Asp Leu Glu Glu Phe Leu
660 665 670

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

Thr Glu Asn Glu Asn Thr Asn Leu Arg Ile Lys Ala Leu Leu Ala Glu
740 745 750
Met Met Asn Asp Asn Ser Val Glu Asn Val Arg Pro Tyr Ser Pro Met
755 760 765
Gln Gln Glu Ile Leu Lys Ile Tyr Glu Asp Gly Val Leu Asn Ser Gly
770 775 780
Ile Glu Ile Glu Asp Glu Tyr Leu Lys Ile Ser Lys Thr Ala Gln Pro

785 790 795 800

Ser Pro Ser Asp Leu Lys Arg Tyr Lys Leu Trp Leu Glu Gln Arg Tyr

805 810 815

Lys Ser Pro Tyr Thr Gly Gln Ile Ile Pro Leu Asn Lys Leu Phe Thr

820 825 830

Pro Glu Tyr Glu Ile Glu His Ile Ile Pro Gln Ser Arg Tyr Phe Asp

835 840 845

Asp Ser Phe Ser Asn Lys Ile Ile Cys Glu Ser Ala Val Asn Lys Leu

850 855 860

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

865 870 875 880

Ile Val Glu Cys Gly Leu Gly Lys Asn Val Lys Val Phe Glu Val Asn

885 890 895

Glu Tyr Glu Asp Phe Val Lys Lys His Tyr Ala Asn Asn Arg Gly Lys

900 905 910

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

915 920 925

Gln Leu Asn Asp Thr Arg His Ile Ser Lys Tyr Ile Ser Gly Val Leu

930 935 940

Ser Asn Ile Val Arg Val Glu Asp Gly Ser Asp Glu Gly Val Asn Ser

945 950 955 960

Lys Asn Ile Val Pro Gly Asn Gly Lys Ile Thr Thr Gln Leu Lys Gln

965 970 975

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

980 985 990

Glu Arg Met Asn Gln Leu Thr Asn Ser Thr Asp Phe Thr Ala Trp Asn

995 1000 1005

Glu Asn His Gln Lys Phe Leu Pro Thr Val Pro Ile Glu Phe Ser

1010 1015 1020

Lys Gly Phe Ser Lys Lys Arg Ile Asp His Arg His His Ala Leu

1025 1030 1035

Asp Ala	Leu Val	Ile Ala	Cys	Ala Thr	Lys Asp	His	Ile Asn	Leu
1040			1045				1050	
Leu Asn	Asn Gln	Ser Ala	Lys	Ser Asp	Thr Lys	Arg	Tyr Asp	Leu
1055			1060				1065	
Lys Lys	Lys Leu	Met Lys	Phe	Glu Lys	Gly Val	Tyr	Asn His	Pro
1070			1075				1080	
Gln Thr	Gly Glu	Arg Ile	Gln	Arg Asp	Val Pro	Lys	Gln Phe	Leu
1085			1090				1095	
Lys Pro	Trp Glu	Ser Phe	Thr	Ile Asp	Ala Lys	Asn	Asn Leu	Asp
1100			1105				1110	
Lys Ile	Ile Ile	Ser Phe	Lys	Gln Asn	Leu Arg	Val	Ile Asn	Lys
1115			1120				1125	
Ala Thr	Asn Tyr	Tyr Glu	Lys	Tyr Val	Glu Lys	Asn	Gly Ile	Lys
1130			1135				1140	
Thr Lys	Glu Arg	Val Glu	Gln	Thr Gly	Thr Asn	Trp	Ala Ile	Arg
1145			1150				1155	
Lys Pro	Met His	Lys Glu	Thr	Val Ser	Gly Ile	Val	Asn Leu	Pro
1160			1165				1170	
Trp Val	Lys Val	Pro Lys	Gly	Lys Ile	Leu Thr	Ala	Thr Arg	Lys
1175			1180				1185	
Ser Leu	Asp Thr	Thr Phe	Asp	Leu Lys	Ser Ile	Asn	Ser Ile	Thr
1190			1195				1200	
Asp Thr	Gly Ile	Gln Lys	Ile	Leu Arg	Asn Tyr	Leu	Glu Phe	Lys
1205			1210				1215	
Gly Ser	Pro Glu	Leu Ala	Phe	Ser Pro	Glu Gly	Ile	Glu Asp	Met
1220			1225				1230	
Asn Lys	Asn Ile	Glu Lys	Tyr	Asn Asp	Gly Lys	Leu	His Gln	Pro
1235			1240				1245	
Ile Asn	Lys Val	Arg Val	Phe	Glu Leu	Gly Ser	Lys	Phe Gln	Val
1250			1255				1260	
Gly Gln	Thr Gly	Asn Lys	Lys	Asp Lys	Tyr Val	Glu	Ala Ala	Lys

1265 1270 1275
 Gly Thr Asn Leu Phe Phe Ala Val Tyr Glu Asp Lys Asn Gly Lys
 1280 1285 1290

Arg Asn Tyr Glu Thr Ile Pro Leu Asn Ile Val Ile Glu Arg Gln
 1295 1300 1305
 Lys Gln Gly Leu Leu Ala Cys Pro Glu Ser Asn Glu Lys Gly Glu
 1310 1315 1320
 Lys Leu Leu Phe Gln Leu Ser Pro Asn Asp Phe Val Tyr Leu Thr
 1325 1330 1335
 Thr Glu Glu Glu Asn Asp Asn Ala Thr Pro Ile Asn Phe Ser Leu
 1340 1345 1350

Leu Ser Lys Glu Gln Ile Asn Asn Leu Tyr Lys Ile Val Ser Phe
 1355 1360 1365
 Thr Gly Asn Arg Leu Tyr Gly Ile Pro Ile Cys Val Ala Thr Thr
 1370 1375 1380
 Ile Val Asn Lys Ala Glu Tyr Thr Gln Leu Asn Lys Ile Glu Phe
 1385 1390 1395
 Thr Lys Glu Lys Asp Leu Leu Leu Lys Leu Asn Val Asn Arg Leu
 1400 1405 1410

Gly Asp Val Lys Thr Phe Thr Ala Asn Asp Ile Arg Lys Ile Phe
 1415 1420 1425

Asn Arg Gln Glu
 1430

<210> 123

<211> 1092

<212> PRT

<213> Bacillus sp.

<223> APG06646.1_RGN

<400> 123

Met Arg Tyr Ser Ile Gly Leu Asp Ile Gly Thr Thr Ser Ile Gly Asn
 1 5 10 15

Ala Val Ile Asn Lys Asp Leu Gln Arg Phe Glu His Leu Gly Val Arg

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

Ile Leu Ser Thr Thr Ala Pro Ser Arg Pro Leu Thr Asn Glu Glu Arg
275 280 285

Lys Ser Ile Leu Ser Ser Leu Phe Ser Lys Lys Glu Val Lys Tyr Lys
290 295 300

Glu Leu Arg Lys Leu Leu Lys Leu Thr Asp Asp Gln Arg Phe Asn Glu
305 310 315 320

Ile Tyr Tyr Ser Pro Asp Glu Thr Ile Glu Lys Thr Glu Asn Arg Thr
325 330 335

Phe Leu Ser Leu Glu Ser Gln Tyr Lys Ile Lys Lys Ile Ile Glu Lys
340 345 350

Thr Glu Ser Lys Asn Met Gln Ser Ser Tyr His Pro Ile Asp Tyr Asp
355 360 365

Thr Ile Gly Tyr Ala Leu Thr Val Phe Lys Asp Asp Lys Asp Ile Gln
370 375 380

His Tyr Leu Gln Asn Ser Tyr Ile Asp Ser Lys Gly Lys Ala Ile Pro
385 390 395 400

Asn Met Ala Asn Arg Glu Tyr Asn Leu Glu Leu Ile Glu Glu Leu Leu
405 410 415

Gly Leu Ser Phe Ala Lys Phe Gly His Leu Ser Leu Lys Ala Leu Asn
420 425 430

Asn Ile Leu Pro Tyr Met Glu Glu Gly Glu Pro Tyr His Ile Ala Cys
435 440 445

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

Arg Leu Leu Pro Pro Ile Pro Val Asp Glu Ile Pro Asn Pro Val Val
465 470 475 480

Val Arg Ala Leu Thr Gln Val Arg Lys Val Leu Asn Ser Ile Ile Lys
485 490 495

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

Ser Lys Pro Phe Lys Glu Arg Lys Ser Leu Glu Arg Glu Phe Asn Glu

515 520 525
 Asn Arg Gln Ile Asn Glu Lys Ala Lys Ala His Ile Ser Glu Leu Tyr
 530 535 540

 Arg Ile Pro Asn Asp Pro Arg Pro His Asp Ile Leu Lys Phe Lys Leu
 545 550 555 560
 Trp Asn Glu Gln Asn Gly Ile Cys Pro Tyr Ser Leu Lys Pro Ile Ser
 565 570 575
 Ile Glu Tyr Leu Phe Asn Ile Gly Tyr Ala Glu Val Asp His Ile Ile
 580 585 590
 Pro Tyr Ser Arg Ser Phe Asp Asp Ser Asn Gly Asn Lys Val Leu Val
 595 600 605

 Leu Thr Arg Glu Asn Gln Asn Lys Leu Asn Arg Thr Pro Tyr Glu Trp
 610 615 620
 Phe Gly His Glu Glu Asn Arg Trp Glu Asp Phe Val Ser Phe Ile Arg
 625 630 635 640
 Thr Met Lys Val Gly Lys Lys Lys Lys Asn Met Leu Leu Lys Lys Asn
 645 650 655
 Phe Asp Glu Glu Gln Glu Glu Gln Ile Leu Ser Arg Asn Leu Asn Asp
 660 665 670

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

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

Glu Pro Trp Arg Asn Phe Ile Lys Glu Leu Glu Ala Arg Met Ile Gln
770 775 780

Asp Pro Glu Lys Met Lys Arg Ala Leu Glu Ser Leu Glu Leu Glu Thr
785 790 795 800

Tyr Gly Glu Ile Phe Leu Asn Lys Leu Lys Pro Ile Phe Val Ser Arg
805 810 815

Met Pro Lys His Ser Ile Lys Gly Gln Ile His Glu Glu Thr Ile Arg
820 825 830

Arg Val Arg Gly Phe Thr Glu Glu Gly Phe Leu Val Thr Val Lys Lys
835 840 845

Thr Arg Leu Asp Gln Ile Pro Phe Asp Lys Asn Gly Asp Phe Pro Met
850 855 860

Tyr Gly Lys Glu Thr Asp Ile Lys Thr Tyr Met Ala Ile Lys Gln Arg
865 870 875 880

Tyr Leu Glu Tyr Gly Gln Asp Lys Gln Lys Ala Phe Ala Val Pro Leu
885 890 895

Arg Lys Pro Ser Lys Asn Pro Lys Asn Ala Pro Ile Val Arg Ser Val
900 905 910

Lys Ile Glu Gly Lys Ala Asn Arg Val Val Met Leu Asp Asp Lys Ala
915 920 925

Ala Ala Asp Asn Ala Ser Ile Val Arg Thr Glu Val Phe Arg His Lys
930 935 940

Lys Thr Gly Glu Tyr Tyr Leu Thr Pro Val Tyr Val Ala Asp Ile Leu
945 950 955 960

Ser Asn Lys Ile Pro Asp Arg Leu Ile Thr Ile Lys Lys Ser Tyr Ser
965 970 975

Asp Trp Asp Arg Ile Thr Asp Glu His Glu Tyr Leu Phe Ser Leu Tyr
980 985 990

Asn Asn Asp Leu Val Lys Ile Ile Leu Pro Lys Glu Lys Glu Thr Lys
995 1000 1005

Lys Tyr Thr Gly Gly Asn His Leu Trp Gln Glu Gly Phe Phe Tyr

1010 1015 1020
 Phe Lys Gly Val Asp Ser Ser Asn Ala Gly Ile Lys Ile Ile Asn
 1025 1030 1035
 His Leu Asn Ser Phe Glu Ala Arg Ile Gly Thr Lys Arg Leu Ile
 1040 1045 1050

 Ala Phe Glu Lys Tyr Gln Val Asn Pro Leu Gly Glu Ile Asn Lys
 1055 1060 1065
 Val His Gly Glu Lys Arg Pro Gly Glu Leu Leu Asn Lys Glu Glu
 1070 1075 1080
 Ile Lys Glu Asn Arg Lys Asn Ile Ser
 1085 1090
 <210> 124
 <211> 1735
 <212> PRT
 <213> Pedobacter sp.
 <223> APG06622_RGN
 <400> 124
 Met Thr Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly
 1 5 10 15

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

 Arg Leu His Leu Val Leu Asn Leu Leu Glu Ala Leu Pro Asn His Tyr
 85 90 95
 Lys Ile Glu Val Asp Phe Glu Arg Asn Asp Lys Arg Cys Gly Gln Phe
 100 105 110
 Lys Glu Gly Lys Glu Pro Lys Met Ala Tyr Leu Pro Thr Lys Asn Leu

115	120	125	
Glu Asn Lys Phe Asp Phe Tyr Phe Glu Glu Ala Phe Asn Glu Met Ile			
130	135	140	
Glu Asp Leu Gln Lys Val Asn Pro Glu Ile Lys Asn Glu Lys Lys Lys			
145	150	155	160
Arg Val Pro Lys Asp Trp Thr Ile Tyr Phe Leu Arg Asn Lys Ala Met			
165	170	175	
Ser Gln Lys Val Ser Leu Glu Glu Leu Ala Trp Val Leu Leu Ser Tyr			
180	185	190	
Asn Gln Lys Arg Gly Thr Asp Ser Glu Glu Ile Glu Asn Glu Asp Glu			
195	200	205	
Lys Ser Glu Ile Lys Glu Gln Ile Asp Leu Lys Val Ile Lys Ala Glu			
210	215	220	
Lys Lys Val Asp Val Asn Gly Ala Phe Phe Glu Ile Thr Leu Asn Asp			
225	230	235	240
Ala Glu Asn Phe Lys Tyr Lys Glu Tyr Thr Thr Lys Gln Leu Thr Phe			
245	250	255	
Glu Asp Asp Val Lys Glu Val Thr Lys Ile Leu Lys Leu Asp Glu Ala			
260	265	270	
Gly Asn Ile Ile Asn Glu Lys Thr Ala Tyr Gln Ile Ser Asp Leu Tyr			
275	280	285	
Asn Leu Thr Leu Ala Asn Ile Glu His Thr Asp Ile Asn Asp Lys Lys			
290	295	300	
Ala Lys His Lys Tyr Asp Phe Ile Tyr Thr Asn Gly Trp Cys Thr Glu			
305	310	315	320
Lys Lys Lys Glu Asn Phe Asp Val Thr Tyr Gln Lys Leu Glu Lys Gln			
325	330	335	
Ile Lys Glu Glu Ser Lys Val Ile Thr Glu Leu Phe Val Val Ser Asn			
340	345	350	
Lys Tyr Asp Phe Ala Gly Lys Pro Glu Ser Ile Ile Pro Ser Ile Lys			
355	360	365	

Leu Pro Asp Phe Asn Ser Glu Gly Ser Lys Asp Trp Thr Leu Leu Lys
 370 375 380
 Lys Lys Thr Glu Lys Glu Ile Val Lys Ser Glu Lys Thr Leu Ser Gly
 385 390 395 400

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

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

 Val Tyr Lys Lys Ala Ile Phe Ala Ser His Pro Leu Phe Gln Glu Phe
 530 535 540
 Arg Ile Trp Asp Lys Ile His Asn Ile Lys Leu Ile Gln Leu Glu Ser
 545 550 555 560
 Lys Asp Ala Asp Gly Asn Thr Gln Thr Asn Val Asp Ile Thr Ser Gln
 565 570 575
 Tyr Phe Lys Pro Glu Asn Tyr Arg Ala Leu Phe Asn His Phe Asn Gly
 580 585 590

 Arg Ser Thr Val Ala Ile Asp Asp Phe Leu Lys Phe Cys Glu Glu Gln
 595 600 605
 Phe Lys Leu Gly Ile Gly Lys Lys Gly Glu Arg Lys Ile Leu Trp Asn

610 Tyr Pro Val Glu Glu Glu Phe Lys Gly Asn Glu Thr Arg Lys Gly Phe
 625 630 635 640
 Glu Asn Arg Phe Arg Lys Cys Gly Phe Asp Gly Phe Asp Asp Phe Met
 645 650 655

 Thr Gln Gln Lys Glu Ile Glu Leu Trp His Tyr Leu Tyr Ser Val Ser
 660 665 670
 Ser Ala Glu Arg Lys Lys Val Ser Gln Val Lys Asp Ser Ser Asp Pro
 675 680 685
 Lys Phe Gly Lys Thr Gly Ile Trp Asn Phe Phe Thr Lys Tyr Phe Lys
 690 695 700
 Asp Glu Asn Ile Ser Leu Glu Thr Leu Glu Lys Leu Cys Arg Asp Phe
 705 710 715 720

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

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

 Ser Pro Thr Lys Ile Arg Glu Glu Ile Thr Lys Glu Leu Lys His His
 850 855 860

Ser Leu Asn Asn Pro Thr Ala Glu Lys Val Leu Lys Glu Thr Leu Lys
865 870 875 880
Val Val Ala Asp Ile Trp Glu Ile Tyr Gly Asn Ser Glu Glu Lys Phe
885 890 895
Phe Asn Glu Ile His Leu Glu Val Ala Arg Glu Leu Gln Lys Ser Asn
900 905 910
Gln Glu Lys Lys Asp Ile Ile Glu Arg Gln Lys Asn Asn Arg Ala Glu
915 920 925
Asn Asn Arg Ile Arg Asn Ile Leu Glu Glu Phe Leu Ser Ser Ser Pro
930 935 940
Tyr Asn Ala Lys Arg Gly Asn Gln Asp His Phe Glu Arg Leu Lys Ile
945 950 955 960
Val Glu Asp Gly Ala Lys Val Arg Ser Tyr Ser Asp Lys Glu Phe Tyr
965 970 975
Ser Asn Asn Lys Asp Lys Phe Ser Lys Ile Asp Ile Glu Val Ile Leu
980 985 990
Lys Lys Lys Lys Ile Thr Gln Ser Asp Phe Glu Lys Tyr Lys Leu Trp
995 1000 1005
Ile Glu Gln Gly Tyr Lys Ser Pro Tyr Thr Asp Gln Ile Ile Lys
1010 1015 1020
Leu Thr Asp Leu Phe Asp Gly Asn Lys Tyr Asn Val Asp His Ile
1025 1030 1035
Phe Ser Arg Ala Ala Ile Thr Asn Asp Ser Leu Asn Asn Lys Val
1040 1045 1050
Val Cys Glu Ala Val Ile Asn Arg Phe Lys Ser Tyr Lys Thr Gly
1055 1060 1065
Arg Glu Phe Ile Ala Gln Phe Gly Gly Lys Thr His Thr Ile Phe
1070 1075 1080
Asp Asp Lys Leu Asn Lys Asn Val Thr Phe Ser Leu Ile Asp Glu
1085 1090 1095
Asp Ala Tyr Val Ser Leu Val Lys Ser Gln Phe Lys Asp Ala Lys

1100	1105	1110
Arg Leu Ile Leu Leu Ser Lys	Glu Val Pro Lys Ala	Phe Thr Glu
1115	1120	1125
Ser Gln Leu Asn Asn Ala Lys	Tyr Ile Ala Arg Lys	Ala Met Glu
1130	1135	1140
Leu Leu Ser His Ile Val Arg	Glu Glu Gly Glu Val	Glu Phe Arg
1145	1150	1155
Ser Lys Asn Val Leu Pro Val	Ser Gly Ala Ile Thr	Asp Arg Leu
1160	1165	1170
Lys Lys Glu Trp Lys Ile Asn	Gln Ile Trp Thr Glu	Leu Leu Lys
1175	1180	1185
Pro Arg Phe Glu Arg Leu Asn	Lys Ile His Asn Ser	Glu Ala Phe
1190	1195	1200
Gly Ser Tyr Ala Ile Ser Lys	Ser Gly His Gln Tyr	Phe Asp Ile
1205	1210	1215
Asn Thr Lys Tyr Ile Leu Gly	Glu Asn Asp Lys Phe	Asp Leu Lys
1220	1225	1230
Arg Leu Asp His Arg His His	Ala Leu Asp Ala Leu	Ile Ile Ala
1235	1240	1245
Leu Cys Thr Asp Asn His Val	Gln Tyr Phe Asn Asn	Ile Asn Ser
1250	1255	1260
Gly Ile Thr Asn Lys Lys Lys	Gly Lys Met Glu Ala	Ile Lys Lys
1265	1270	1275
Gln Arg Ala Gly Ile Lys Arg	Gln Ile Met Tyr Ser	Glu Lys Asp
1280	1285	1290
Lys Glu Asn Pro Asn Glu Lys	Val Trp Arg Tyr Met	Leu Pro Gly
1295	1300	1305
Ser Phe Arg Asn Lys Glu Ser	Val Asn Asn Glu Lys	Asp Ser Val
1310	1315	1320
Val Asp Leu Ser Trp Glu Asn	Asn Tyr Ser Ala Lys	Glu Ser Asn
1325	1330	1335

Asp Tyr	Arg Lys Thr	Ile Leu	Glu Cys Leu	Asp Asn	Cys Ile Val
1340		1345		1350	
Ser Phe	Lys Asn Asp Phe	Lys Ile Val	Ser Lys Ser	Ser Asn Lys	
1355		1360		1365	
Tyr Glu	Ser Tyr Tyr Asp	Glu Asn Gly	Asn Leu Arg	Leu Asp Ser	
1370		1375		1380	
Asn Gly	Lys Pro Gln Lys	Glu Phe Ile	Ser Gln Gln	Asn Asn Asn	
1385		1390		1395	
Gln Lys	His Trp Ser Val	Arg Lys Pro	Leu His Thr	Asp Asn Pro	
1400		1405		1410	
Ser Ala	Glu Ile Thr Leu	Gln Phe Asp	Arg Leu Lys	Leu Ile Asp	
1415		1420		1425	
Asn Leu	Gly Lys Val Asp	Leu Ile Ile	Glu Glu Lys	Ile Lys Asp	
1430		1435		1440	
Ala Val	Lys Glu Ile Leu	Gly Met His	Asn Asn Lys	Ile Gly Lys	
1445		1450		1455	
Ala Glu	Ser Tyr Leu Lys	Lys Asn Pro	Leu Ile Leu	Glu Glu Lys	
1460		1465		1470	
Lys Ile	Glu Tyr Thr Asp	Phe Lys Ile	Ser Asn Lys	Lys Tyr Arg	
1475		1480		1485	
Lys Arg	Gln Pro Ile Thr	Asn Leu Ser	Val Arg Thr	Gly Thr Gly	
1490		1495		1500	
Ala Leu	Pro Thr Tyr Asp	Ser Val Val	Asp Arg Ile	Asn Lys Val	
1505		1510		1515	
Ser Asp	Phe Val Leu Arg	Asn Glu Leu	Leu Gln His	Leu Lys Ile	
1520		1525		1530	
Asn Asp	Asn Asp Ile Asp	Lys Ala Phe	Ser Ile Glu	Gly Ile Asn	
1535		1540		1545	
Glu Phe	Asn Ser Asn Arg	Lys Ile Pro	Val Phe Arg	Leu Pro Ile	
1550		1555		1560	
Ser Glu	Ser Ala Thr Leu	Lys Phe Lys	Ile Gly His	Lys Asn Ser	

1565	1570	1575
Asn Ser His Lys Trp Val Glu Thr Gly Gly Asn Phe His Phe Arg		
1580	1585	1590
Ile Ala Asp Glu Glu Asn Ser Arg Asp Tyr Glu Thr Ile Pro Leu		
1595	1600	1605
Arg Ile Ala Ile Glu Met Glu Lys Asp Asn Ile Ile Arg Arg Gln		
1610	1615	1620
Ile Val Asp Asn Glu Asn Glu Ile Val Leu Ser Pro Asn Asp Leu		
1625	1630	1635
Val Tyr Val Pro Leu Glu Lys Asn Glu Ile Val Asp Phe Asn Asn		
1640	1645	1650
Leu Thr Ile Gln Gln Ile Lys Arg Ile Tyr Lys Phe Val Phe Ser		
1655	1660	1665
Thr Lys Lys Glu Ala His Phe Ile Pro Tyr Ser Ser Ser Ala Glu		
1670	1675	1680
Ile Ile Lys Asn Glu Asn Gly Thr Asn Gly Tyr Ser Glu Arg Thr		
1685	1690	1695
Lys Asp Phe Phe Asp Gly Asn Val Leu Gln Asp Asp Lys Gly Lys		
1700	1705	1710
Pro Ile Gln Val Lys Ser Val Cys Trp Lys Leu Lys Val Asp Arg		
1715	1720	1725
Leu Gly Asn Ile Ser Lys Ala		
1730	1735	
<210> 125		
<211> 1224		
<212> PRT		
<213> Chitinophaga sp.		
<223> APG02787_RGN		
<400> 125		
Met Lys Lys Leu Ile Val Gly Leu Asp Leu Gly Val Asn Ser Ile Gly		
1	5	10
		15

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

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

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

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

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

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

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

 Gly His Ile Phe Ser Asp Ala Met Ser Leu Ala Gly Tyr Asn His Ser
 465 470 475 480
 Ala Ser Leu Thr His Glu Glu Asn Leu Lys Arg Gln Leu Lys Gly Ser
 485 490 495
 Leu Pro Leu Leu Lys Lys Asn Ser Leu Arg Gln Pro Val Val Glu Lys
 500 505 510

Ile Leu Asn Gln Met Ile Gly Val Met Asn Ala Ile Ile Lys Glu Trp
515 520 525

Gly Arg Pro Asp Glu Ile Arg Val Glu Leu Ala Arg Glu Leu Lys Gln
530 535 540

Ser Ala Glu Glu Arg Asn Lys Ala Phe Ala Ala Leu Ser Lys Arg Glu
545 550 555 560

Arg Glu Gly Glu Ile Ile Thr Lys Arg Ile Glu Ala Glu Phe Arg Gln
565 570 575

Phe Gly Ile Arg Ala Thr Arg Asn Asn Ile Leu Lys Trp Arg Leu Phe
580 585 590

His Glu Ile Ser Asn Asp Glu Ala Arg Leu Asn Ala Thr Cys Val Tyr
595 600 605

Cys Gly His Pro Phe Gly Ile Thr Asp Ala Leu Arg Gly Ser Ser Val
610 615 620

Asp Ile Glu His Ile Ile Pro Lys Ser Lys Leu Phe Asp Asp Ser Gln
625 630 635 640

Gly Asn Lys Thr Leu Val His Arg Arg Cys Asn Glu Asp Lys Gly Asp
645 650 655

Arg Thr Ala Phe Asp Phe Met Ser Ala Lys Gly Glu Ser Val Leu Ala
660 665 670

Ser Tyr Ile Glu Thr Val Asp Arg Leu Tyr Gln Ser Asn Val Ile Ser
675 680 685

Arg His Lys Arg Asp Arg Leu Leu Thr Thr Gly Asp Asp Ile Pro Lys
690 695 700

Asp Phe Ile Glu Arg Gln Leu Arg Glu Thr Gln Tyr Ile Ala Arg Lys
705 710 715 720

Ala Lys Glu Ile Leu Glu Gln Ile Cys Tyr Asn Val Trp Ser Thr Ser
725 730 735

Gly Gln Ile Thr Glu Tyr Leu Arg Arg Ile Trp Gly Trp Asn Asp Val
740 745 750

Leu Met Asn Leu Gln Leu Pro Val Tyr Arg Asp Leu Gly Leu Thr Glu

755	760	765	
Trp Lys Glu Tyr Thr Asp Arg Asp Gly Asn Leu His Arg Gln Glu Val			
770	775	780	
Ile Lys Asp Trp Ser Lys Arg Asn Asp His Arg His His Ala Ile Asp			
785	790	795	800
Ala Leu Thr Ile Ala Cys Thr Arg Gln Gly Phe Ile Gln Arg Ile Asn			
805	810	815	
Arg Leu Asn Ala Thr Asn Asn Arg Glu Glu Met Arg Ala Ser Leu Asn			
820	825	830	
Gly Asn Tyr Asp Pro Lys Lys Asn Leu Leu Glu Asn Tyr Ile Phe Ala			
835	840	845	
Leu Lys Pro Phe Ser Thr Lys Gln Val Glu Val Ala Ala Ala Gln Ile			
850	855	860	
Leu Ile Ser Phe Lys Pro Gly Lys Arg Val Ala Thr Leu Ser Lys Leu			
865	870	875	880
Lys Ala Lys Gly Lys Asn Ser Gly Thr Gly Val Ile Val Pro Arg Gly			
885	890	895	
Pro Leu Ser Glu Glu Ser Val Tyr Gly Arg Val Arg Leu Met Asp Ser			
900	905	910	
Lys Lys Gln Pro Val Arg Tyr Leu Phe Glu His Pro Glu Leu Ile Phe			
915	920	925	
Lys Pro Tyr Ile Arg Val Leu Val Glu Glu Arg Leu Ala Ser Tyr Ser			
930	935	940	
Gln Asp Leu Lys Lys Ala Leu Ala Ser Leu Lys Lys Asp Pro Ile Tyr			
945	950	955	960
Leu Asp Gly Thr Lys Thr Ile Ile Leu Glu Trp Gly Ser Cys Tyr Lys			
965	970	975	
Asp Glu Val Val Ile Lys Tyr Pro Leu Thr Asp Ile Lys Ala Lys Asp			
980	985	990	
Thr Gln Phe Ile Ile Asp Gly His Ile Arg Ala Ile Val Asn Glu Arg			
995	1000	1005	

Leu Ser	Gln Tyr	Lys Gly	Asn Glu	Lys Lys	Ala Phe	Glu His	Pro
1010			1015			1020	
Leu Tyr	Ser Asp	Ala Ala	Cys Thr	Lys Glu	Ile Arg	Ser Val	Arg
1025			1030			1035	
Cys Leu	Thr Gly	Leu Thr	Ala Val	Val Ser	Val Lys	Lys Asp	Glu
1040			1045			1050	
Ser Gly	Arg Glu	Ile Gly	Phe Val	Lys Pro	Gly Asn	Asn His	His
1055			1060			1065	
Ile Ala	Leu Tyr	Thr Asp	Ala Glu	Gly Gln	Lys Arg	Glu His	Leu
1070			1075			1080	
Cys Thr	Phe Trp	His Ala	Val Glu	Arg Lys	Lys Trp	Gly Met	Pro
1085			1090			1095	
Ala Val	Ile Thr	Gln Pro	Gly Val	Ile Trp	Asp Asn	Ile Leu	Gln
1100			1105			1110	
Lys Gln	Asp Asn	Phe Pro	Gln Ala	Phe Leu	Glu Lys	Leu Pro	Leu
1115			1120			1125	
Glu Ala	Trp Thr	Phe Asp	Lys Ser	Leu Gln	Gln Asn	Glu Met	Phe
1130			1135			1140	
Ile Val	Gly Leu	Thr Ile	Glu Glu	Val Gln	Thr Cys	Ile Asp	Ser
1145			1150			1155	
Phe Asp	Tyr Ala	Arg Leu	Ser Lys	His Leu	Tyr Arg	Ile Gln	Lys
1160			1165			1170	
Leu Ser	Ser Thr	Gly Tyr	Lys Ile	Val Leu	Arg His	His Leu	Ala
1175			1180			1185	
Thr Thr	Ile Glu	His Thr	Arg Asp	Met Ile	Gly Ile	Ala Ser	Pro
1190			1195			1200	
Ser Ser	Phe Thr	Gly Ile	Lys Val	Tyr Val	Ser Lys	Thr Gly	Lys
1205			1210			1215	
Ile Ser	Met Pro	His Asp					
1220							

<210> 126

<211> 1457

<212> PRT

<213> Mucilaginibacter sp.

<223> APG06248_RGN

<400> 126

Met Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly Trp

1 5 10 15

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

20 25 30

Thr Gly Ser Arg Ile Ile Pro Met Ser Gln Asp Ile Lys Asp Glu Tyr

35 40 45

Gly Lys Gly Asn Ser Ile Ser Gln Thr Ala Asp Arg Thr Arg Phe Arg

50 55 60

Ser Leu Arg Arg Leu Arg Glu Arg His Leu Leu Arg Arg Glu Arg Leu

65 70 75 80

His Arg Ile Leu Asn Ile Leu Gly Phe Leu Pro Glu His Tyr Ala Asp

85 90 95

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

100 105 110

Pro Lys Leu Ala Tyr Arg Asn Asn Glu Phe Ile Phe Lys Ala Ser Phe

115 120 125

Glu Glu Met Leu Ile Asp Phe Arg Glu Gln Gln Pro Glu Met Leu Gln

130 135 140

Gly Asp Lys Lys Val Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu Arg Lys

145 150 155 160

Lys Ala Leu Thr Gln Lys Val Glu Lys Ala Glu Leu Ala Trp Leu Ile

165 170 175

Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu Glu

180 185 190

Gln Gln Glu Asn Pro Asn Lys Lys Val Glu Phe His Ser Leu Lys Val

195 200 205

Val Asp Val Ile Ala Asp Glu Thr Gln Arg Gly Lys Ala Glu Thr Trp

210 215 220
 Tyr Ser Leu Ile Leu Glu Asn Gly Trp Val Tyr Arg Arg Ser Ser Lys
 225 230 235 240

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

 Ala Tyr Ile Tyr Asp His Leu Leu Val Ala Pro Asp Gln Lys Ile Lys
 305 310 315 320
 Gly Lys Leu Val Arg Thr Ile Glu Arg Arg Phe Tyr Lys Thr Glu Leu
 325 330 335
 Lys Ala Ile Leu Ala Lys Gln Gln Asn Phe His Ala Glu Leu Gln Asn
 340 345 350
 Lys Asp Leu Phe Leu Asp Cys Val Arg Glu Leu Tyr Arg His Asn Leu
 355 360 365

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

 Tyr Gln Glu Phe Arg Leu Trp Gln Trp Ile Tyr Asn Leu Gly Ile Tyr
 435 440 445
 Gln Lys Glu Asp Asp Thr Asn Val Thr Ala Asp Leu Ile Lys Gly Ile
 450 455 460

Gln Gly Tyr Gln Asp Leu Phe Asp Phe Leu Asn Ser Lys Lys Glu Ile
 465 470 475 480
 Lys Gln Asp Thr Leu Ile Lys Phe Leu Leu Ser Gln Thr Gly Leu Lys
 485 490 495

Gly Lys Glu Leu Asn Lys Ala Ala Leu Thr Leu Arg Trp Asn Phe Val
 500 505 510
 Glu Asp Lys Thr Tyr Pro Cys Asn Glu Thr Gly Thr Glu Ile Arg Ala
 515 520 525
 Arg Leu Ala Lys Val Ala Tyr Ile Pro Glu Asn Phe Leu Thr Ala Glu
 530 535 540
 Lys Ala Leu Ser Leu Trp His Ile Ile Tyr Ser Val Asn Asp Lys Lys
 545 550 555 560

Glu Tyr Glu Leu Ala Leu Arg Lys Phe Ala Lys Lys His Gln Ile Asp
 565 570 575
 Glu Glu Thr Phe Phe Ile Asn Phe Gln Lys Phe Pro Pro Phe Gln Ser
 580 585 590
 Glu Tyr Gly Ala Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu
 595 600 605
 Ile Arg Leu Gly Lys Tyr Trp His Trp Glu Ala Ile Asp Pro Lys Ser
 610 615 620

Gln Glu Arg Ile Glu Lys Ile Leu Ser Gly Glu Ala Asp Gln Asn Ile
 625 630 635 640
 Lys Glu Arg Ile Arg Glu Lys Ala Met His Leu Thr Arg Glu Glu His
 645 650 655
 Phe Gln Gly Leu Gln Leu Trp Leu Ala Asn Tyr Ile Val Tyr Asp Arg
 660 665 670
 His Ala Glu Ala Ala Asn Pro Asp Lys Trp Ser Thr Val Ser Asp Leu
 675 680 685

Glu Asp Tyr Leu Lys Glu Leu Lys Gln His Ser Leu Arg Asn Pro Ile
 690 695 700
 Val Glu Gln Val Val Thr Glu Thr Leu Arg Val Val Lys Asp Ile Trp

705 710 715 720
 Leu Thr Tyr Gly Asn Gly Ala Lys Asn Phe Phe Asp Glu Ile His Ile
 725 730 735
 Glu Leu Ala Arg Glu Met Lys Asn Thr Met Glu Glu Arg Lys Arg Leu
 740 745 750

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

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

 Ile Asn Lys Leu Lys Asp Lys Gln Leu Gly Leu Glu Phe Ile Lys Asn
 885 890 895
 His Ala Gly Thr Ile Val Glu Leu Gly Leu Gly Lys Thr Ala Arg Val
 900 905 910
 Phe Asp Glu Glu Arg Tyr Gln Asp Phe Val Lys Ala Asn Tyr Thr Lys
 915 920 925
 Asn Pro Gly Lys Arg Ser Lys Leu Met Leu Glu Glu Ile Pro Asp Lys
 930 935 940

 Met Ile Ala Arg Gln Leu Asn Asp Thr Arg Tyr Ile Ser Lys Phe Ile
 945 950 955 960

Thr Ser Val Leu Ser Asn Ile Val Arg Ser Glu Thr Lys Asp Asp Gly
965 970 975
Val Asn Ser Lys Asn Val Leu Leu Cys Asn Gly Arg Ile Thr Ser Val
980 985 990
Leu Arg Gln Asp Trp Gly Met Asn Asp Ile Trp Asn Glu Leu Ile Leu
995 1000 1005

Pro Arg Phe Glu Arg Leu Asn Gln Ile Thr Asp Ser Thr Val Phe
1010 1015 1020
Thr Ala Trp Asn Glu Asn His Gln Lys Phe Leu Pro Thr Val Pro
1025 1030 1035
Met Gly Phe Ser Lys Gly Phe Gln Arg Lys Arg Ile Asp His Arg
1040 1045 1050
His His Ala Met Asp Ala Leu Ile Ile Ala Ser Cys Thr Arg Asp
1055 1060 1065

His Ile Asn Leu Leu Asn Asn Glu Tyr Ala Lys Ser Ser Lys Arg
1070 1075 1080
Phe Asp Leu Asn Arg Lys Leu Arg Lys Phe Glu Lys Val Ala Tyr
1085 1090 1095
Lys Ser Pro Val Asn Gly Gln Gln Ile Asp Lys Asp Ile Pro Thr
1100 1105 1110
Val Phe Ile Lys Pro Trp Gln Arg Phe Thr Ile Asp Ala Lys Cys
1115 1120 1125

Ser Leu Glu Asp Ile Ile Val Ser Phe Lys Gln Asn Leu Arg Val
1130 1135 1140
Ile Asn Lys Thr Thr Asn Tyr Phe Asp Gly Tyr Lys Asp Glu His
1145 1150 1155
Gly Ile Val Asn Phe Asp Asn Lys Gly Ala Pro Val Lys Gln Gln
1160 1165 1170
Arg Lys Gln Thr Lys Gly Asp Ser Trp Ala Ile Arg Lys Pro Met
1175 1180 1185

His Lys Asp Phe Val Tyr Gly Lys Val Asp Leu Lys His Ile Lys

1190	1195	1200
Ile Pro Lys Gly Lys Phe Ala Thr Ala Ile Arg Lys Ser Leu Asp		
1205	1210	1215
Thr Thr Phe Asp Gln Lys Lys Ile Gln Thr Ile Thr Asp Thr Gly		
1220	1225	1230
Val Gln Lys Ile Leu Ile Asn Tyr Leu Gln Ser Lys Asp Asn Lys		
1235	1240	1245
Ala Glu Leu Ala Phe Ser Pro Glu Gly Ile Glu Asp Met Asn Gln		
1250	1255	1260
Asn Ile Arg Ala Tyr Asn Asp Gly Lys Leu His Gln Pro Ile Lys		
1265	1270	1275
Lys Val Arg Val Phe Glu Ile Gly Ser Arg Phe Pro Leu Gly Glu		
1280	1285	1290
Thr Gly Asn Lys Lys Thr Lys Tyr Val Glu Thr Ala Lys Gly Thr		
1295	1300	1305
Asn Leu Phe Phe Ala Ile Tyr Gln His Lys Asn Lys Arg Asn Tyr		
1310	1315	1320
Glu Thr Val Pro Leu Asn Thr Val Ile Glu Arg Gln Lys Gln His		
1325	1330	1335
Leu Ser Pro Val Pro Thr Glu Asp Asp Lys Gly Gly Thr Leu Leu		
1340	1345	1350
Phe Tyr Leu Ser Pro Asn Asp Leu Val Tyr Val Pro Thr Pro Asp		
1355	1360	1365
Glu Leu Gly Asn Asn Leu Gly Gly Leu Asp Ile Arg Asn Leu Asn		
1370	1375	1380
Lys Glu Gln Arg Arg Arg Ile Tyr Lys Val Glu Lys Ala Ser Gly		
1385	1390	1395
Val Glu Cys Tyr Phe Ile Arg Gln Asp Ile Ala Tyr Leu Ile Lys		
1400	1405	1410
Ser Tyr Asp Ala Thr Thr Lys Ile Gly Glu Leu Glu Ser Gln Asn		
1415	1420	1425

Lys Leu Gln Thr Thr Met Ser Ser Asp Arg Leu Lys Ile Ala Asp

1430 1435 1440

Val Cys Ile Lys Leu Lys Ile Asp Arg Leu Gly Ile Ser Arg

1445 1450 1455

<210> 127

<211> 1662

<212> PRT

<213> Acidovorax sp.

<223> APG06007_RGN

<400> 127

Met Lys Lys Ile Leu Gly Leu Asp Ile Gly Thr Thr Ser Ile Gly Tyr

1 5 10 15

Ser Ile Ile Lys Val Ser Asp Glu Lys Asn Glu Asn Ile Ile Ser Asp

20 25 30

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

35 40 45

Glu Ala Ile Gly Val Arg Ile Ile Asn Gln Asp Thr Asp Arg Phe Asp

50 55 60

Arg Gly Leu Lys Leu Asn Asp Pro Lys Gly Ser Thr Leu Thr Pro Thr

65 70 75 80

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

85 90 95

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

100 105 110

Thr Pro Asp Glu Ser Tyr Phe Thr Asn Lys Lys Gly Lys Arg Gly Lys

115 120 125

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

130 135 140

Arg Glu Lys Ile Ser Leu Ser Glu Leu Ala Arg Val Leu Leu His Leu

145 150 155 160

Asn Gln Leu Arg Gly Tyr Ser Ser Asp Arg Phe Ala Asn Asp Glu Glu

165 170 175

Thr Lys Phe Asp Tyr Tyr Leu Thr Glu Val Val Glu Ala Asp Trp Glu
180 185 190

Asn Arg Thr Ala Val Phe Glu Lys Gly Ser Thr Glu Asp Val Lys Tyr
195 200 205

Tyr Gln Phe Lys Ala Lys Leu Arg Phe Val Glu Pro Tyr Val Ile Val
210 215 220

Glu Asp Asp Glu Lys Thr Ile Leu Tyr Glu Val Glu Gly Tyr Ile Phe
225 230 235 240

Lys Lys Glu Leu Asp Phe Glu Val Gly Asp Phe Ile Thr Ile Lys Ala
245 250 255

Pro Glu Phe Lys Gln Asn Lys Lys Gly Lys Val Val Ile Glu Glu Tyr
260 265 270

Tyr Lys Ile Thr Pro Thr Lys Pro Asp Ala Thr Asp Trp Lys Tyr Lys
275 280 285

Tyr Gln Thr Leu His Lys Asn Leu Ser Asp Trp Cys Lys Ala Gly Gly
290 295 300

Thr Val Gly Ser Tyr Phe Tyr Lys Asn Phe Tyr Glu Leu Thr Thr Leu
305 310 315 320

Ser Arg Ile Arg Asn Asn Val Val Asn Arg Glu Trp Tyr Glu Asn Glu
325 330 335

Phe Asp Ala Ile Trp Thr Leu Gln Tyr Glu Phe Tyr Lys Asp Phe Phe
340 345 350

Glu Lys Ile Asp Val Ser Glu Leu Val Lys Gly Val Phe Lys Asp Tyr
355 360 365

Ser Ala Val Leu Asn Asp Val Ile Ser Lys Ser Gly Asn Lys Ala Gln
370 375 380

Leu Gln Thr Leu Ile Lys Asp Lys Ile Ile Phe Tyr Gln Arg Pro Trp
385 390 395 400

Gln Gln Ala Lys Ser Lys Gly Arg Cys Glu Phe Glu Lys Ile Lys Val
405 410 415

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

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

 465 470 475 480
 Lys Leu Thr Gly Lys Thr Ile Glu Lys Val Lys Glu Leu Leu Phe Lys
 485 490 495
 Lys Leu Gln Gly Ser Lys Thr Leu Ser Trp Arg Thr Phe Val Lys Glu
 500 505 510
 Glu Leu Gly Leu Asp Thr Tyr Asp Glu Ile Glu Glu Lys Arg Lys Ser
 515 520 525
 Lys Val Ser Lys Lys Gln Gly Leu Asp Ile Glu Thr Gly Glu Ile Val

 530 535 540
 Asp Ser Phe Phe Ser Val Asn Phe Arg Lys Arg Lys Arg Asp Gly Thr
 545 550 555 560
 Phe Asp Asp Ile Lys Leu Lys Gly Asn Thr Thr Lys Ala Glu Leu Gln
 565 570 575
 Asn Ile Leu Lys Glu Lys Arg Ala Glu Trp Phe Asp Thr Leu His Asn
 580 585 590
 Gln Asn Lys Lys Val Thr Asn Leu Gln Leu Leu Trp Glu Ile Ile Tyr

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

 660 665 670

Ala Lys Val Asp Glu Leu Ile Ser Leu Asn Thr Ser Glu Thr Asp Val
675 680 685

Asp Ala Asp Asp Lys Leu Asn Gly Leu Lys Asp Phe Val Thr Asp Arg
690 695 700

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

Tyr Leu Asn Tyr Trp Glu Ala Ala Ala Ile Val Tyr Gly Ser His Ser
725 730 735

Ser Lys Lys Ser Ser Glu Thr Lys Ile Met Glu Arg Val Lys Asn His
740 745 750

Ser Met Asn Asn Pro Val Val Glu Lys Ile Val Asn Glu Ala Ile Ser
755 760 765

Ile Val Asn Leu Leu Lys Glu Lys Tyr Gly Phe Asp Glu Ile Arg Ile
770 775 780

Glu Leu Ser Arg Glu Leu Arg Ala Ser Ala Asp Glu Arg Gln Gln Met
785 790 795 800

Trp Glu Gly Met Ser Asn Gly Ala Ala Lys Asn Glu Trp Ala Lys Lys
805 810 815

Met Leu Arg Glu Ile Lys Gln Ala Met Leu Asp Asn Asn Arg Asp Ile
820 825 830

Ser Tyr Leu Asp Thr Glu Ile Ser Asn Ser Asn Leu Asp Lys Ile Lys
835 840 845

Ile Ile Glu Asp Val Val Lys Tyr Gln Lys Pro Val Glu Phe Lys Thr
850 855 860

Lys Glu Lys Glu Tyr Lys Leu Thr Glu Pro Thr Lys Ala Glu Val Thr
865 870 875 880

Lys Tyr Leu Leu Trp Leu Glu Gln Asn Phe Lys Cys Pro Tyr Thr Asn
885 890 895

Gln Pro Ile Pro Leu Thr Asp Val Phe Ala Arg Gly Lys Val Val Glu
900 905 910

Ile Glu His Ile Ile Pro Arg Glu Arg Tyr Tyr Ser Asn Ala Tyr Ala

915	920	925	
Asn Lys Val Ile Thr Trp Arg Glu Val Asn Gln Glu Lys Gly Asn Lys			
930	935	940	
Thr Ala Tyr Glu Phe Ile Val Ser Lys Arg Asn Pro Ala Gly Thr Val			
945	950	955	960
Gln Val Gly Lys Lys Glu Ile Ser Leu Val Ser Ala Ser Gln Trp Glu			
965	970	975	
Asn His Val Lys Glu Met Phe Pro Lys Ser Gly Lys Arg Thr Asn Leu			
980	985	990	
Leu Arg Lys Glu Ile Pro Glu Asp Pro Ile Glu Arg Thr Leu Lys Glu			
995	1000	1005	
Asn Gln Tyr Ile Asn Lys Lys Leu Lys Glu Lys Leu Gly Glu Leu			
1010	1015	1020	
Val Gly Ser Asp Lys Val Trp Val Thr Ser Gly Ala Ile Thr Asp			
1025	1030	1035	
Met Leu Arg Glu Lys Trp Lys Leu Asn Asp Ile Trp Lys Glu Leu			
1040	1045	1050	
Val Gln Glu Arg Phe Leu Lys Phe Asn Asn Gly Glu Val Lys Asp			
1055	1060	1065	
Arg Leu Asn Leu Lys Tyr Thr Thr Lys Tyr Tyr Asn Asn Lys Thr			
1070	1075	1080	
Asn Gln Tyr Glu Asp Val Glu Lys Phe Ala Gly Met Ser Lys Arg			
1085	1090	1095	
Ile Asp His Arg His His Ala Leu Asp Ala Ile Ile Ile Ala Cys			
1100	1105	1110	
Thr Lys Gln Ser His Ile Gln Tyr Ile Asn Asn Leu Asn Ala Ile			
1115	1120	1125	
Asn Ser Ala Asp Ile Glu Asn Asp Asp Asn Lys Lys Thr Lys Tyr			
1130	1135	1140	
Lys Gly Ile Lys Thr Asp Val Cys Ile Gly Asn Ser Ser Ser Lys			
1145	1150	1155	

Phe Lys Phe Pro Trp Asp Glu Asn Lys Phe Ile Pro Glu Val Lys

1160

1165

1170

Gln Ala Leu Gln Asn Ile Leu Ile Ser His Lys Asn Ser Asn Val

1175

1180

1185

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

1190

1195

1200

Gly Lys Val Ala Ser Val Arg Gly Glu Leu His Leu Gln Thr Asn

1205

1210

1215

Tyr Ala Lys Lys Lys Tyr Phe Glu Asn Glu Lys Ile Asp Ile Val

1220

1225

1230

Lys Leu Ile Pro Leu Leu Phe Lys Lys Lys Phe Glu Asn Gln Asn

1235

1240

1245

Gln Thr Leu Val Arg Phe Lys Ser Phe Asp Glu Ile Ile Lys Glu

1250

1255

1260

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

1265

1270

1275

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

1280

1285

1290

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

1295

1300

1305

Ile His Glu Lys Ser Gly Glu Lys Met Thr Trp Leu Ser Thr Phe

1310

1315

1320

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

1325

1330

1335

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

1340

1345

1350

Arg Arg Ile Gly Gln Tyr Arg Leu Asn Tyr Ile Asn Lys Lys Ile

1355

1360

1365

Ile Glu Ile Lys Ala Leu Lys Leu Pro Phe Lys Glu Glu Glu Asn

1370

1375

1380

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

1385 1390 1395
Ile Tyr Glu Leu Arg Ile Lys Gln Lys Asp Gly Lys Tyr His Trp

1400 1405 1410
Val Glu Ile Lys Asp Val Gln His Ser Asp Leu Glu Asn Ile Glu

1415 1420 1425
Tyr Ala Lys Gln Glu Thr Thr Ile Ala Val Lys Asn Lys Ile Lys

1430 1435 1440
Asp Phe Ser Leu Asp Glu Leu Lys Glu Arg Tyr Ile Ser Asn Pro

1445 1450 1455
Ile Phe Leu Ser Glu Lys Pro Thr Pro Val Lys Lys Ala Arg Gln

1460 1465 1470
Lys Ser Phe Phe Gln Asp Leu Tyr Gln Leu Asn Asp Gly Arg Tyr

1475 1480 1485
Val Tyr Ser Arg Asp Val Phe Met Thr Tyr Val Phe Val Pro Glu

1490 1495 1500
Asp Ser Asn Arg Lys Asp Ala Lys Arg Glu Leu Glu Phe Leu Lys

1505 1510 1515
Phe Phe Asp Ala Val Lys Ile Val Thr Leu Glu Lys Pro Glu Lys

1520 1525 1530
Ile Asp Tyr Gln Lys Leu Ile Lys Lys Glu Asp His Lys Leu Leu

1535 1540 1545
Phe Thr Leu Ser Lys Asn Asp Ile Val Tyr Ile Pro Glu Glu Ile

1550 1555 1560
Leu Thr Asp Glu Gln Ile Ala Glu Ile Asn Trp Asn Asp Lys Arg

1565 1570 1575
Asp Ile Val Pro Lys Leu Phe Val Val Lys Glu Met Thr Pro Ser

1580 1585 1590
Arg Asn Glu Ile Val Phe Gln His Leu Tyr Lys Ser Asp Ser Ile

1595 1600 1605
Arg Val Asn Gly Asp Glu Ala Lys Asn Ile Leu Asn Thr Glu Lys

1610 1615 1620

Leu Glu Glu Gln Ile Lys Tyr Gly Asp Thr Asn Met Trp Lys Arg
1625 1630 1635

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

1640 1645 1650

Tyr Trp Glu Asn Asn Lys Ile Val Leu

1655 1660

<210> 128

<211> 1075

<212> PRT

<213> Bacillus sp.

<223> APG02874_RGN

<400> 128

Met Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp

1 5 10 15

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

20 25 30

Gln Gly Ile Val Asp Ala Gly Val Arg Met Phe Asp Lys Ala Glu Ile

35 40 45

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

50 55 60

Ser Arg Lys Arg Leu Arg Arg Arg Ser Gln Arg Lys Lys Ala Ile Arg

65 70 75 80

Tyr Leu Ile Ile Leu His Glu Leu Leu Ser Glu Asp Gln Leu Gln Gln

85 90 95

Leu Tyr Pro Leu Gln Lys Glu Ser Ile Asp Ile Trp Asp Ile Arg Leu

100 105 110

Glu Ser Leu Asp Arg Val Leu Asn Gln Thr Glu Trp Ala Arg Leu Leu

115 120 125

Ile His Leu Gly Gln Arg Arg Gly Phe Lys Ser Asn Arg Lys Ser Glu

130 135 140

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

145 150 155 160

Gln Lys Ala Leu Gln Gln Tyr Arg Thr Val Ala Glu Met Trp Met Lys

165 170 175

Asp Ala Lys Phe Ser Gln Phe Asp Lys Lys Arg Asn Ala Pro Gly Glu

180 185 190

Tyr Leu Phe Ser Val Ser Arg Tyr Asp Leu Gln Asn Glu Ile Gln Thr

195 200 205

Ile Phe Glu Ser Gln Arg Asn Leu Gly Ser Pro Tyr Ala Thr Thr Asp

210 215 220

Leu Glu Glu Glu Tyr Ile Ser Val Trp Ser Lys Gln Leu Pro Phe Ala

225 230 235 240

Ser Gly Glu Asp Ile Leu Lys Lys Val Gly Tyr Cys Ser Leu Glu Arg

245 250 255

Lys Glu Arg Arg Ile Pro Lys Ala Thr Tyr Thr Phe Gln Tyr Phe Ile

260 265 270

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

275 280 285

Leu Thr Thr Glu Glu Arg Ser Ile Leu Ile Glu Lys Ile Leu Thr Arg

290 295 300

Asn Asp Tyr Phe Ser Lys Lys Ser Leu Pro Asp Leu Thr Tyr Gly Asp

305 310 315 320

Leu Arg Lys Trp Leu Ala Leu Asp Ser Ser Leu Leu Phe Lys Asp Leu

325 330 335

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

340 345 350

Ala Asn Leu Lys Ser Leu Tyr Glu Leu Arg Lys Val Val Glu Thr Ser

355 360 365

Ser Asn Gly Gln Ser Phe Glu Arg Ile Asp Tyr Asp Thr Phe Gly Tyr

370 375 380

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

385 390 395 400

Lys Pro Ser Asn Leu Ala His Lys Val Tyr Glu Glu Asp Leu Ile Glu

405 410 415
Arg Leu Leu Ser Leu Ser Tyr Glu Lys Phe Gly His Leu Ser Leu Lys

420 425 430
Ala Ile Gln Thr Ile Leu Pro Tyr Met Glu Asp Gly Met Met Tyr Thr

435 440 445
Glu Ala Val Gln Ala Ala Gly Tyr Asp Ile Thr Gly Leu Gln Ser Ala

450 455 460
Lys Arg Thr Leu Leu Leu Pro Thr Ile Pro Ala Asp Ile Thr Asn Pro

465 470 475 480
Val Val Arg Arg Ala Leu Thr Gln Thr Arg Lys Val Val Asn Ala Ile

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

500 505 510
Glu Leu Ser Lys Asp His Gln Glu Arg Gln Gln Ile Leu Lys Asn Gln

515 520 525
Gln Ala Asn Tyr Glu Arg Asn Lys Gly Ala Ile Ser Val Leu Val Glu

530 535 540
Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp Ile Val Arg Tyr Lys Leu

545 550 555 560
Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr Ser Leu Lys Val Ile Pro

565 570 575
Pro Ser Val Phe Phe Gly Glu Leu Lys Arg Glu Arg Asn His Ile Pro

580 585 590
Thr Leu Glu Val Asp His Ile Leu Pro Tyr Ser Gln Ser Phe Met Asp

595 600 605
Asn Tyr His Asn Lys Val Leu Val Tyr Ser Asp Glu Asn Gln Lys Lys

610 615 620
Gly Asn Arg Ile Pro Ser Gln Tyr Leu Met Ala Val Ser Pro Glu Lys

625 630 635 640
Trp Thr Lys Phe Glu Glu Tyr Val Ser Ser Asn Lys Gly Phe Ser Lys

645 650 655

Gln Lys Arg Gln Tyr Leu Leu Lys Lys Asp Tyr Ser Ser Arg Glu Arg
660 665 670
Asp Ile Val Lys Glu Arg His Leu Val Asp Thr Arg Tyr Ala Thr Arg
675 680 685
Phe Met Lys Asn Phe Ile Glu Lys Thr Leu Leu Phe Lys Glu Ser Arg
690 695 700
Ser Ser Ser Leu Thr Lys Arg Val Gln Thr Val Asn Gly Lys Ile Thr
705 710 715 720
Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys Glu Arg Thr Glu Thr
725 730 735
Tyr Leu His His Ala Leu Asp Ala Ile Val Val Ala Cys Thr Asp Gln
740 745 750
His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln Glu Arg Glu Arg Asn
755 760 765
Val Arg Gln Pro Tyr Phe Pro Trp Pro Trp Glu Gly Phe Arg Asp Glu
770 775 780
Leu Leu Asn Tyr Leu Asn Arg Gln Pro Asp Ser Leu Glu Ile Ser Gln
785 790 795 800
Ser Ile Lys Gln Asn Leu Phe Leu Arg Asp Tyr Leu Met Val Ser Arg
805 810 815
Met Pro Lys His Ser Val Thr Gly Met Ser His Lys Gln Thr Ile Arg
820 825 830
Ser Lys Ala Cys Val Asp Glu Lys Gly Lys Val Lys Thr Thr Lys Arg
835 840 845
Val Leu Leu Gln Asp Ile Lys Phe Asp Ser Asn Gly Asp Phe Lys Met
850 855 860
Ile Gly Lys Glu Gln Asp Met Ala Thr Tyr Asn Ala Ile Lys Glu Arg
865 870 875 880
Phe Leu Ser Ser Gly Gly Asn Val Lys Ser Ala Phe Thr Glu Pro Leu
885 890 895
His Lys Pro Ser Lys Lys Gly Lys Gly Asn Val Ile Arg Lys Val Thr

900 905 910
 Val Glu Asp Gln Thr Lys Ser Phe Val Arg Asp Val Asn Gly Gly Val
 915 920 925
 Ala Glu Asn Gly Ser Leu Val Arg Ile Asp Leu Phe Glu Lys Glu Gly

 930 935 940
 Met Tyr Tyr Val Val Pro Ile Tyr Val Gly Asp Thr Val Leu Lys Gln
 945 950 955 960
 Leu Pro Ile Lys Ile Ser Ala Ser Asn Arg Gly Tyr Tyr Lys Trp Val
 965 970 975
 Glu Leu Asp Glu Asn Tyr Arg Phe Lys Phe Ser Leu Tyr Ala His Asp
 980 985 990
 Leu Ile Arg Ile Val Thr Lys Asp Glu Asp Arg Phe Phe Tyr Phe Ser

 995 1000 1005
 Tyr Phe Asp Asn Asp Gly Asn Arg Ile Lys Phe Leu Asn Ile Asn
 1010 1015 1020
 Ser Pro Thr Gly Lys Asn Glu Asn Arg Tyr Gly Val Arg Ser Ile
 1025 1030 1035
 Lys Leu Ile Glu Lys Tyr Asn Val Ser Thr Leu Gly Asp Ile Ser
 1040 1045 1050
 Leu Val Lys Ser Glu Glu Arg Gln Leu Phe Gln Gly Met Lys Lys

 1055 1060 1065
 Lys Val Leu Asn Ile Ile Asn
 1070 1075
 <210> 129
 <211> 1068
 <212> PRT
 <213> Bacillus sp.
 <223> APG03850_RGN
 <400> 129
 Met Lys Tyr Val Leu Gly Leu Asp Ile Gly Ile Ala Ser Cys Gly Trp
 1 5 10 15
 Ala Val Ile Asn Gln Glu Lys His Arg Ile Glu Asp Leu Gly Val Arg

20 25 30
 Ile Phe Asp Lys Ala Glu Asn Pro Lys Asp Gly Lys Ser Leu Ala Thr

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

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

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

 225 230 235 240
 Cys Thr Leu Glu Pro Lys Glu Lys Arg Val Pro Lys Ala Ser Trp Ser
 245 250 255
 Phe Gln Tyr Phe Val Ala Leu Gln Thr Ile Asn Asn Leu Arg Leu Ile
 260 265 270

Asn Lys Asp Arg Ile Glu Glu Leu Ser Phe Glu Glu Lys Asn Gln Ile
275 280 285

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

Lys Leu Leu Ser Ile Pro Asn Glu Phe His Phe Asn Asn Leu Leu Tyr
305 310 315 320

Ser Ala Asp Thr Val Asp Thr Ala Val Glu Asn Lys Lys Cys Ile Glu
325 330 335

Phe Lys Glu Tyr His Ser Ile Asn Lys Leu Tyr Lys Gln Ile Tyr Gly
340 345 350

Lys Ser Val Pro Asn Leu Leu Pro Ile Asp Tyr Asp Thr Ile Ala Cys
355 360 365

Gly Leu Thr Ile Phe Lys Asp Asp Lys Asp Ile Leu Ala Tyr Leu Gln
370 375 380

Asn Lys Tyr Val Asn Ala Lys Gly Lys Pro Ile Ser Asn Leu Ala Lys
385 390 395 400

Lys Thr Tyr Asp Asp Thr Phe Ile Gln Ala Leu Leu Thr Leu Asn Phe
405 410 415

Ser Lys Met Gly His Leu Ser Phe Lys Ala Leu Lys Asn Ile Ile Pro
420 425 430

Phe Leu Glu Glu Gly Leu Ser Tyr Asp Lys Ala Cys Glu Lys Ala Gly
435 440 445

Tyr Asn Phe Lys Gly Thr Ser His Ala Glu Lys Thr Lys Tyr Leu Pro
450 455 460

Val Ile Pro Gln Asn Thr Asn Pro Val Val His Arg Ala Leu Ser Gln
465 470 475 480

Thr Arg Lys Val Ile Asn Ala Ile Ile Lys Lys Tyr Gly Ser Pro Ser
485 490 495

Ala Ile His Ile Glu Thr Ala Arg Glu Leu Ser Lys Thr Phe Gln Glu
500 505 510

Arg Lys Glu Ile Asp Ser Met Tyr Gln Asp Asn Ser Lys Lys Asn Glu

515 520 525
His Ala Ile His Lys Leu Lys Glu Leu Gly Leu Ile Asn Pro Ser Gly
530 535 540
Ile Asn Ile Val Lys Phe Lys Leu Trp Asn Glu Gln Asp Gly Arg Cys

545 550 555 560
Met Tyr Ser Gly Lys Tyr Ile Glu Pro His Arg Leu Phe Glu Glu Gly
565 570 575
Tyr Thr Glu Val Asp His Ile Leu Pro Tyr Ser Arg Ser Leu Asp Asp
580 585 590
Ser Tyr Asn Asn Lys Ala Leu Thr Leu Gly Ile Glu Asn Gln Arg Lys
595 600 605
Gly Asn Lys Thr Pro Tyr Glu Tyr Ile Gly Lys Thr Ser Ile Trp His

610 615 620
Glu Phe Glu Thr Arg Val Gln Ser Asn Lys Arg Ile Asn Lys Lys Lys
625 630 635 640
Gln Gln Lys Leu Leu Leu Gln Tyr Phe Ser Tyr Thr Arg Glu Gln Glu
645 650 655
Phe Ile Lys Arg Asn Leu Asn Asp Thr Arg Tyr Ala Thr Ile Tyr Leu
660 665 670
Ser Thr Leu Ile Gln Gln His Leu Ile Phe Ser Glu Ser Ser Arg Lys

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

740 745 750
Lys Arg Met Gln Phe Pro Glu Pro Trp Glu Gly Phe Arg Leu Glu Leu
755 760 765

Glu Ala Arg Ile Ser Pro Asn Thr Gln Gln Tyr Leu Lys Arg Ile Leu
 770 775 780
 Phe Lys Asn Tyr Ala Asp Val Asn Leu Ser Glu Ile Lys Pro Ile Phe
 785 790 795 800
 Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Glu Leu His Gln Glu
 805 810 815
 Thr Ile Arg Lys Leu Ile Gly Tyr Asn Glu Lys Gly Lys Val Leu Thr
 820 825 830
 Ala Ile Lys Thr Lys Leu Glu Asp Ile Pro Phe Asp Ala Asn Gly Asp
 835 840 845
 Phe Pro Met Tyr Gly Lys Glu Thr Asp Leu Tyr Thr Tyr Asn Ala Ile
 850 855 860
 Lys Glu Arg Tyr Leu Ser His Lys Lys Asp Lys Arg Lys Ser Phe Gln
 865 870 875 880
 Asp Pro Leu Tyr Lys Pro Thr Lys Ser Gly Glu Ile Gly Pro Leu Ile
 885 890 895
 Lys Ser Ile Lys Ile Met Asp Thr Arg Thr Ile Val Asn Pro Val Asn
 900 905 910
 Gln Gly Lys Gly Val Val Tyr Asn Ser Lys Ile Ala Arg Thr Asp Val
 915 920 925
 Phe Lys Lys Asp Glu Lys Tyr Tyr Leu Ile Pro Ile Tyr Thr Ile Asp
 930 935 940
 Leu Leu Lys Asn Ile Leu Pro Gln Lys Ala Ile Thr Ala Gly Lys Gly
 945 950 955 960
 Tyr Glu Asp Trp Thr Thr Ile Asp Pro Ser Phe Thr Phe Leu Phe Ser
 965 970 975
 Leu Phe Pro Asn Asp Leu Ile Gln Ile Val Pro Ser Lys Asn Lys Thr
 980 985 990
 Ile Lys Ala Arg Thr Thr Val Ser Lys Lys Glu Val Leu Leu Pro Ser
 995 1000 1005
 Leu Thr Gly Tyr Phe Lys Gly Val His Ser Gly Thr Ala Gly Ile

1010 1015 1020
 Thr Val Glu Thr His Asp Gly Ser Val Ile Ala Asn Val Gly Ser
 1025 1030 1035
 Lys Gln Leu Leu Leu Phe Glu Lys Tyr Gln Val Asp Val Leu Gly
 1040 1045 1050
 His Tyr Thr Lys Ile Lys Glu Glu Lys Arg Ile Gly Met Val Ile

 1055 1060 1065
 <210> 130
 <211> 1081
 <212> PRT
 <213> Bacillus sp.
 <223> APG07553_RGN
 <400> 130
 Met Gln Tyr Val Leu Gly Leu Asp Ile Gly Ile Ala Ser Cys Gly Trp
 1 5 10 15
 Ala Val Ile Asn Gln Glu Lys Glu Arg Ile Glu Asp Leu Gly Val Arg
 20 25 30
 Ile Phe Asp Lys Ala Glu Asn Pro Lys Asp Gly Lys Ser Leu Ala Ala
 35 40 45

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

 Glu Phe Thr Arg Val Leu Ile His Leu Ala Lys Arg Arg Gly Phe Lys
 115 120 125
 Ser Asn Ser Lys Glu Thr Thr Leu Ser Glu Asn Gly Gln Ile Leu Glu
 130 135 140
 Ser Ile Ser Glu Asn Gln Gln Ile Met Glu Gln Lys Asn Tyr Lys Thr

145 150 155 160
 Ile Gly Glu Met Ile Leu Lys Asp Lys Lys Phe Glu Asn His Lys Arg
 165 170 175

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

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

 Lys Leu Leu Ser Ile Pro Asn Glu Leu His Phe Asn Ser Leu Leu Tyr
 305 310 315 320
 Ser Ala Asp Thr Val Asp Thr Thr Val Glu Asn Arg Lys Cys Ile Glu
 325 330 335
 Leu Lys Glu Tyr His Ser Ile Asn Lys Val Tyr Lys Gln Ile Tyr Gly
 340 345 350
 Lys Asn Ala Leu Asn Leu Leu Pro Ile Asp Tyr Asp Thr Ile Ala Tyr
 355 360 365

 Gly Leu Thr Ile Phe Lys Asp Asp Lys Asp Ile Leu Glu His Leu Lys
 370 375 380
 Asn Lys Tyr Val Asn Ala Lys Gly Lys Pro Ile Asn Asn Leu Ala Lys
 385 390 395 400

Lys Thr Tyr Asp Asp Thr Phe Ile Gln Ala Leu Leu Thr Leu Asn Phe
 405 410 415
 Ser Lys Met Gly His Leu Ser Phe Lys Ala Leu Lys Asn Ile Ile Pro
 420 425 430

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

 Ala Ile His Ile Glu Thr Ala Arg Glu Leu Ser Lys Thr Phe Gln Glu
 500 505 510
 Arg Lys Glu Ile Asp Ser Met Tyr Gln Asp Asn Ser Lys Lys Asn Glu
 515 520 525
 His Ala Ile His Lys Leu Lys Glu Leu Gly Asn Ile Asn Pro Ser Gly
 530 535 540
 Ile Asn Ile Val Lys Phe Lys Leu Trp Asn Glu Gln Asp Gly Lys Cys
 545 550 555 560

 Met Tyr Ser Gly Lys Tyr Ile Glu Pro His Arg Leu Phe Glu Glu Gly
 565 570 575
 Tyr Thr Glu Val Asp His Ile Leu Pro Tyr Ser Arg Ser Leu Asp Asp
 580 585 590
 Ser Tyr Asn Asn Lys Thr Leu Thr Leu Gly Ile Glu Asn Gln Arg Lys
 595 600 605
 Gly Asn Lys Thr Pro Tyr Glu Tyr Met Gly Asn Thr Ser Ile Trp Asp
 610 615 620

 Glu Tyr Glu Ile Arg Val Gln Ser Asn Lys Lys Ile Asn Lys Lys Lys
 625 630 635 640
 Gln Gln Lys Leu Leu Leu Gln His Phe Ser Tyr Ala Arg Glu Gln Glu

645	650	655	
Phe Ile Glu Arg Asn Leu Asn Asp Thr Arg Tyr Ala Thr Ile Tyr Leu			
660	665	670	
Ser Ser Leu Ile Gln Gln His Leu Ile Phe Ser Glu Ser Ser Arg Lys			
675	680	685	
Lys Lys Val His Thr Val Ser Gly Ile Ile Thr Ser His Leu Arg Ser			
690	695	700	
Arg Trp Gly Phe Asn Lys Asp Arg Lys Glu Gly His Ile His His Ala			
705	710	715	720
Leu Asp Ala Val Ile Val Ala Val Thr Ser Asp His Met Ile Gln Arg			
725	730	735	
Val Thr Lys Tyr Tyr Lys Leu Lys Glu Leu Asn Arg Asn Leu Gln Ala			
740	745	750	
Lys Arg Met Pro Phe Pro Glu Pro Trp Glu Gly Phe Arg Leu Glu Leu			
755	760	765	
Glu Ala Arg Ile Ser Pro Asn Thr Gln Gln Tyr Leu Lys Gly Leu Arg			
770	775	780	
Phe Lys Asn Tyr Ala Asp Val Asn Leu Cys Glu Ile Lys Pro Ile Phe			
785	790	795	800
Val Ser Lys Met Pro Lys Arg Ser Ile Thr Gly Glu Leu His Gln Glu			
805	810	815	
Thr Ile Arg Lys Phe Ile Gly Tyr Asn Glu Lys Gly Lys Val Leu Thr			
820	825	830	
Ala Ile Lys Thr Lys Leu Glu Asp Ile Pro Phe Asp Ala Asn Gly Asp			
835	840	845	
Phe Pro Met Tyr Gly Lys Glu Thr Asp Leu Tyr Thr Tyr Asn Ser Ile			
850	855	860	
Lys Glu Arg Tyr Leu Ser His Lys Lys Asp Lys Arg Lys Ser Phe Gln			
865	870	875	880
Glu Pro Leu Tyr Lys Pro Thr Lys Ser Gly Gly Ile Gly Pro Leu Ile			
885	890	895	

Lys Ser Ile Lys Ile Met Asp Thr Arg Thr Ile Val Asn Pro Val Asn
 900 905 910

Gln Gly Lys Gly Val Val Tyr Asn Ser Lys Ile Ala Arg Thr Asp Val
 915 920 925

Phe Lys Lys Asp Asp Lys Tyr Tyr Leu Ile Pro Ile Tyr Thr Ile Asp
 930 935 940

Leu Met Lys Asn Ile Leu Pro Gln Lys Ala Ile Thr Ala Gly Lys Gly
 945 950 955 960

Tyr Glu Asp Trp Ile Thr Ile Asp His Ser Phe Thr Phe Leu Phe Ser
 965 970 975

Leu Phe Pro Asn Asp Leu Ile Lys Ile Val Pro Ser Lys Asn Lys Glu
 980 985 990

Ile Lys Ala Arg Ser Thr Ser Ser Lys Lys Glu Ile Leu Leu Pro Ser
 995 1000 1005

Leu Ile Gly Tyr Phe Lys Ser Val His Ser Gly Thr Ala Gly Ile
 1010 1015 1020

Thr Val Glu Ser His Asp Gly Arg Phe Ile Ala Asn Val Gly Ser
 1025 1030 1035

Lys Gln Leu Leu Leu Phe Glu Lys Tyr Gln Val Asp Val Leu Gly
 1040 1045 1050

His Tyr Thr Lys Ile Lys Glu Glu Lys Arg Ile Gly Met Ala Thr
 1055 1060 1065

Cys Asn Asp Asn Lys Lys Ser Thr Ala Phe Gly Ser Leu
 1070 1075 1080

<210> 131

<211> 1404

<212> PRT

<213> Chryseobacterium sp.

<223> APG03031_RGN

<400> 131

Met Lys Lys Thr Leu Gly Ile Asp Leu Gly Thr Asn Ser Ile Gly Trp
 1 5 10 15

Val Ile Arg Asn Ala Asp Leu Leu Asp Asn Gln Ile Glu Asp Phe Gly
20 25 30

Val Ile Thr Phe Glu Lys Gly Val Ala Ser Glu Lys Gly Asn Glu Phe
35 40 45

Pro Lys Val Gln Lys Arg Thr Glu Ser Arg Gly Lys Arg Arg Asn Tyr
50 55 60

Gln Ser Glu Lys Tyr Arg Lys Phe Ala Leu Leu Ala Phe Leu Ile Gln
65 70 75 80

Lys Asn Met Cys Pro Leu Thr Met Asp Glu Phe Cys Asn Trp Lys Asp
85 90 95

Tyr Arg Lys Gly Lys Lys Arg Glu Tyr Pro Gln Ser Glu Glu Phe Leu
100 105 110

Asn Trp Ile Arg Phe Asp Phe Asn Gly Asp Gly Leu Pro Asp Phe His
115 120 125

Leu Phe Gly Lys Glu Lys Asp Asp Ser Tyr Tyr Val Phe Arg Ala Phe
130 135 140

Ala Ala Asp Glu Asn His Lys Asn Val Phe Asp Asn Asn Pro His Ile
145 150 155 160

Leu Gly Arg Val Leu Tyr Gln Met Val Gln Arg Arg Gly Phe Lys Gly
165 170 175

Arg Asp Glu Glu Glu Ala Lys Thr Met Leu Val Gly Ser Asp Lys Asn
180 185 190

Gly Thr Lys Gly Arg Asn Asp Ile Glu Glu Tyr Ile Glu Lys Tyr Lys
195 200 205

Thr Leu Gly Ala Ala Leu Tyr Tyr His Gln Lys Asp Phe Gly Gly Arg
210 215 220

Ile Lys Gln Arg Tyr Asn Leu Arg Lys Asp Asn Glu Asn Glu Leu Lys
225 230 235 240

Glu Ile Cys Lys Ile His Gln Leu Ser Gln Glu Asp Tyr Glu Lys Leu
245 250 255

Trp Lys Ala Ile Ile Trp Gln Arg Pro Leu Arg Thr Gln Lys Gly Leu

260 265 270
Val Gly Asn Cys Ile Tyr Glu Lys Asn Lys Lys Arg Val Ala Val Ser
275 280 285
His Pro Leu Tyr Glu Glu Tyr Arg Thr Trp Val Phe Ile Asn Asn Leu

290 295 300
Asn Ile Ile Pro Pro Phe Gly Glu Glu Lys Gln Asn Tyr Ile Ser Glu
305 310 315 320
Lys Ile Tyr Pro Leu Phe Tyr Lys Ser Lys Pro Asp Phe Glu Leu Ser
325 330 335
Asp Ile Asp Lys Gln Leu Lys Lys Asp Gly Ala Lys Arg Thr Ser Lys
340 345 350
His Asn Asp Lys Thr Lys Val Leu Ser Val Lys Leu Leu Lys Gln Phe

355 360 365
Gln Asp Val Phe Gly Glu Asp Trp Lys Gln Lys Leu Gln Trp Asp Ile
370 375 380
Asn Gly Asp Arg Asn Ser Gln Pro Glu Lys Lys Glu Lys Ser Leu Tyr
385 390 395 400
Thr Phe Glu Asp Leu Trp His Ile Leu Leu Thr Phe Asp Gly Gln Glu
405 410 415
Asn Leu Lys Lys Phe Ala Leu Glu Lys Leu Asn Leu Asp Asp Glu Lys

420 425 430
Ala Ile Lys Phe Ser Lys Ile Lys Leu Gln Gln Gly Tyr Ala Thr Leu
435 440 445
Ser Ile Ser Ala Ile Lys Lys Ile Leu Pro Tyr Leu Gln Arg Gly Phe
450 455 460
Leu Tyr Ser His Ala Val Tyr Met Ala Asn Leu Tyr Lys Met Leu Gly
465 470 475 480
Glu Thr Thr Ile Ser Glu Thr Leu Thr Asn His Phe Ser Asp Glu Ile

485 490 495
Leu Lys Ile Leu Glu Asn Asp Asp Phe Asn Arg Lys Asn Ser Ala Ile
500 505 510

Ile Asn Ser Leu Ile Ser Glu Met Leu Asn Asp Asp Asn Arg Tyr Tyr
 515 520 525
 Ile Ala Asp Asn Arg Asn Leu Asp Lys Thr Glu Glu Gln Ile Val Leu
 530 535 540
 Lys Ser Ile Thr Asp Asn Tyr Gly Leu Glu Thr Trp Gly Asn Phe Asp

 545 550 555 560
 Glu Glu Lys Lys Met Glu Ile Ile Asp Phe Val Ser Ser Lys Tyr Leu
 565 570 575
 Glu Phe Leu Lys Lys Thr Tyr Thr Glu Lys Gly Asn Leu Phe Ile Gln
 580 585 590
 Pro Glu Arg Ile His Asp Lys Ile Phe Asn Phe Leu Lys Glu Thr Tyr
 595 600 605
 His Val Ser Glu Asp Lys Ile Lys Tyr Leu Trp His Pro Ser Glu Gln

 610 615 620
 Glu Asn Tyr Ala Lys Ala Asp Glu Tyr Phe Glu Tyr Ser Ile Gly Ser
 625 630 635 640
 Lys Lys Tyr Tyr Val Lys Glu Glu Ser Thr Gln Asn Phe Leu Ser Arg
 645 650 655
 Asn Pro Asn Ala Glu Phe Asp Gly Arg Gln Leu Lys Leu Leu Gly Asn
 660 665 670
 Pro Glu Pro Leu Ser Lys Gly Phe Lys Asn Pro Met Ala Leu Lys Ser

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

 740 745 750
 Phe Pro Asn Ile Asn Leu Ile Asp Glu Asn Thr Leu Ile Arg Lys Ile

755 760 765
 Arg Leu Trp His Glu Gln Asn Lys Ile Cys Leu Tyr Thr Gly Lys Thr
 770 775 780
 Ile Pro Phe Ser Glu Leu Ile Ile Gly Asn Lys Tyr Asp Phe Glu His
 785 790 795 800
 Thr Ile Pro Ala Ser Ile Ser Phe Asp Asn Glu Leu Lys Asn Leu Thr

 805 810 815
 Ile Ser Asp Ser Ser Tyr Asn Arg Leu Tyr Lys Gly Lys Lys Phe Pro
 820 825 830
 Thr Gln Leu Ser Asn Tyr Asp Ser Glu Gln Thr Ile Asn Gly Glu Ser
 835 840 845
 Cys Asn Ser Ile Leu Arg Asn Ile Glu Phe Ile Phe Gly Glu Arg Ile
 850 855 860
 Val Glu Tyr Lys Glu Ile Lys Gly Lys Thr Glu Lys Ile Val Lys Trp

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

 930 935 940
 Arg Asp Thr Gln Ile Ile Thr Lys Tyr Ala Leu Pro Tyr Phe Lys Thr
 945 950 955 960
 Leu Phe Lys Arg Val Ser Val Glu Lys Gly Gly Val Thr Asp Ile Phe
 965 970 975
 Lys Lys Val Tyr Lys Val Gln Ser Arg Asn Ser Lys Lys Asp Arg Ser
 980 985 990
 Val His Ser His His Ala Gln Asp Ala Ala Ile Leu Thr Leu Ile Pro

 995 1000 1005

Asn Ala	Phe His Arg Glu Arg	Ile Ile Lys Ala Tyr	Glu Asn Glu
1010	1015	1020	
Ile Asp	Asn Arg Thr Gly Lys	Thr Tyr His Glu Gln	Pro Leu Asp
1025	1030	1035	
Trp Glu	Asn Phe Ser Glu Lys	Tyr Ile Leu Glu Leu	Gln Asp Lys
1040	1045	1050	
Val Leu	Ile Asn Asn Leu Thr	Asp Asn Arg Thr Ile	Thr Gln Thr
1055	1060	1065	
Tyr Lys	Thr Val Arg Lys Arg	Gly Lys Val Val Tyr	Asp Ile Asp
1070	1075	1080	
Glu Asn	Gly Asn Ser Leu Lys	Arg Val Ser Lys Gly	Asp Thr Ile
1085	1090	1095	
Arg Gly	Gln Leu His Gly Glu	Thr Phe Tyr Gly Ala	Ile Lys Gln
1100	1105	1110	
Pro Val	Arg Asp Glu Asp Asn	Lys Ile Leu Phe Asp	Glu Asn Lys
1115	1120	1125	
Lys Met	Ile Leu Lys Asp Glu	Ile Tyr Leu Ala Val	Arg Lys Pro
1130	1135	1140	
Leu Val	Tyr Lys Lys Asp Ala	Asn Ser Pro Gly Phe	Lys Thr Leu
1145	1150	1155	
Asp Glu	Val Glu Lys Val Ile	Val Asp Lys Gly Leu	Phe Lys Met
1160	1165	1170	
Ile Lys	Lys Gln Val Glu Glu	Ser Asp Phe Lys Thr	Ala Leu Ile
1175	1180	1185	
Asp Gly	Val Tyr Met Leu Asn	Glu Ser Gly Glu Lys	Val Asn Lys
1190	1195	1200	
Ile Arg	Arg Ile Arg Cys Phe	Glu Asn Gly Leu Lys	Tyr Thr Thr
1205	1210	1215	
Ala Ile	Lys Val His Glu His	Ser Phe Val Ser His	Lys Glu Tyr
1220	1225	1230	
Lys Gln	Ser Thr Leu Ala Thr	Asn Gly Glu Asn Thr	Tyr Cys Leu

1235 1240 1245
Phe Tyr Lys Asn Asp Lys Gly Lys Ala Met Lys Ile Leu Ser Ile
1250 1255 1260
Val Asp Leu Ala Glu Leu Lys Leu Lys Asn Ile Gln Asn Leu Tyr
1265 1270 1275
Asn Glu Pro Glu Phe Ser Ser Phe Glu Val Gly Lys Gly Lys Asn
1280 1285 1290
Lys Tyr Lys Ile Pro Leu Tyr Ser Val Leu Lys Ser Gly Asp Lys

1295 1300 1305
Val Leu Phe Tyr Lys Glu Asn Ile Tyr Glu Leu Lys Asp Leu Glu
1310 1315 1320
Gln Lys Glu Leu Ser Asp Arg Met Phe Lys Met Tyr Gln Phe Glu
1325 1330 1335
Lys Asp Gly Arg Ile Lys Phe Arg His His Leu Ala Ala Gly Ile
1340 1345 1350
Asp Thr Glu Leu Lys Lys Glu Asn Pro Glu Asn Ser Ser Val Asn

1355 1360 1365
Phe Glu Glu Lys Gln Val Phe Leu Arg Leu Ser Gln Gly Gln Trp
1370 1375 1380
Asn Phe Ala Ile Asp Gly Lys Asp Phe Glu Ile Ser Leu Asp Gly
1385 1390 1395
Lys Ile Asn Trp Lys Phe

1400

<210> 132

<211> 1092

<212> PRT

<213> Bacillus sp.

<223> APG09208_RGN

<400> 132

Met Arg Tyr Ser Ile Gly Leu Asp Ile Gly Thr Thr Ser Ile Gly Ser

1 5 10 15

Ala Val Ile Asn Asn Asp Leu Gln Arg Phe Glu His Leu Gly Val Arg

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

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

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

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

 260 265 270

Ile Leu Ser Thr Thr Thr Pro Ser Arg Pro Leu Thr Asn Glu Glu Arg
275 280 285
Asp Ser Ile Leu Asn Ser Leu Phe Ser Lys Lys Glu Val Lys Tyr Lys
290 295 300
Glu Leu Arg Lys Met Leu Lys Leu Thr Asp Asp Gln Arg Phe Asn Glu
305 310 315 320
Ile Tyr Tyr Ser Pro Asp Glu Thr Ile Glu Lys Thr Glu Asn Arg Val
325 330 335
Phe Leu Ser Ser Glu Ser Gln Tyr Lys Ile Lys Lys Ile Ile Glu Lys
340 345 350
Thr Glu Ser Lys Asn Met Leu Ser Thr Tyr Cys Pro Ile Asp Tyr Asp
355 360 365
Thr Ile Gly Tyr Ala Leu Thr Val Phe Lys Asp Asp Lys Asp Ile Gln
370 375 380
His Tyr Leu Gln Asn Arg Tyr Val Asp Ser Lys Gly Lys Ala Ile Ser
385 390 395 400
Asn Met Ala Asn Arg Glu Tyr Asn Leu Glu Leu Ile Glu Glu Leu Leu
405 410 415
Gly Leu Ser Phe Ser Lys Phe Gly His Leu Ser Leu Lys Ala Leu Asn
420 425 430
Asn Ile Leu Pro Tyr Met Glu Glu Gly Glu Pro Tyr His Ile Ala Cys
435 440 445
Glu Met Ala Gly Tyr Gln Phe Ser Gln Arg Leu Thr Lys Glu Lys His
450 455 460
Arg Leu Leu Pro Pro Ile Pro Val Asp Glu Ile Pro Asn Pro Val Val
465 470 475 480
Val Arg Ala Leu Thr Gln Val Arg Lys Val Leu Asn Ser Ile Ile Lys
485 490 495
Lys Tyr Gly Pro Pro Ser Asp Ile Tyr Ile Glu Leu Ala Arg Glu Met
500 505 510
Gly Lys Pro Phe Lys Glu Arg Lys Ser Leu Glu Arg Glu Phe Asn Glu

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

 580 585 590
 Pro Tyr Ser Arg Ser Phe Asp Asp Ser Asn Gly Asn Lys Val Leu Val
 595 600 605
 Leu Thr Asn Glu Asn Arg Asn Lys Leu Asn Arg Thr Pro Tyr Glu Trp
 610 615 620
 Phe Gly His Asp Glu Asn Arg Trp Gly Asp Phe Val Ser Phe Val Arg
 625 630 635 640
 Thr Met Lys Val Gly Lys Lys Lys Lys Asp Met Leu Leu Lys Lys Asn

 645 650 655
 Phe Asp Glu Glu Gln Glu Glu Gln Ile Leu Ser Arg Asn Leu Asn Asp
 660 665 670
 Thr Arg Tyr Ile Thr Arg Tyr Ile Lys Ser Phe Ile Glu Asp Asn Leu
 675 680 685
 Glu Phe Arg Thr Glu Glu Asn Lys Lys Gln Tyr Val His Thr Val Asn
 690 695 700
 Gly Gly Tyr Thr Ser Leu Met Arg Lys Arg Trp Gly Leu Asn Lys Asp

 705 710 715 720
 Arg Arg Gly Asn Asp Leu His His Ala Val Asp Ala Ala Ile Ile Ala
 725 730 735
 Val Ser Leu Pro Phe Lys His Lys Val Asn Thr Tyr Phe Lys Arg Gln
 740 745 750
 Glu Ile Gly Leu Ser Asn Leu Ser Asn Asn Lys Lys Asp Ile Phe Pro
 755 760 765

Glu Pro Trp Arg Asn Phe Ile Lys Glu Leu Glu Ala Arg Met Ile Gln

770

775

780

Asp Pro Glu Lys Met Lys Thr Ala Leu Glu Ser Leu Glu Leu Glu Thr

785

790

795

800

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

805

810

815

Met Pro Lys His Ser Ile Lys Gly Gln Ile His Glu Glu Thr Ile Arg

820

825

830

Arg Val Arg Gly Leu Thr Glu Glu Gly Phe Leu Val Thr Val Lys Lys

835

840

845

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

850

855

860

Tyr Gly Lys Glu Thr Asp Ile Lys Thr Tyr Met Ala Ile Lys Gln Arg

865

870

875

880

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

885

890

895

Arg Lys Pro Ser Lys Asn Pro Glu Asn Ala Pro Ile Val Arg Ser Val

900

905

910

Lys Ile Glu Gly Lys Ala Asn Arg Val Val Met Leu Asp Asp Lys Ala

915

920

925

Ala Ala Asp Asn Ala Ser Ile Ala Arg Thr Glu Val Phe Arg His Lys

930

935

940

Glu Thr Gly Glu Tyr Tyr Val Thr Pro Val Tyr Val Ala Asp Ile Val

945

950

955

960

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

965

970

975

Glu Trp Asp Arg Ile Thr Asp Glu Tyr Glu Tyr Leu Phe Ser Leu Tyr

980

985

990

Asn Asn Asp Leu Ile Lys Ile Val Leu Pro Lys Glu Lys Glu Thr Lys

995

1000

1005

Lys Tyr Thr Gly Gly Asn His Leu Trp Gln Glu Gly Phe Phe Tyr

1010 1015 1020
Phe Lys Gly Val Asp Ser Ser Asn Ala Ser Ile Lys Ile Ile Asp

1025 1030 1035
His Leu Asn Ser Phe Glu Val Arg Ile Gly Ile Arg Arg Leu Ile

1040 1045 1050
Ala Phe Glu Lys Tyr Gln Val Asn Pro Leu Gly Glu Ile Asn Lys

1055 1060 1065
Val His Gly Glu Lys Arg Pro Gly Glu Leu Leu Asn Glu Glu Glu

1070 1075 1080
Ile Lys Glu Asn Arg Lys Asn Ile Ser

1085 1090

<210> 133

<211> 1150

<212> PRT

<213> Enterococcus sp.

<223> APG05586_RGN

<400> 133

Met Tyr Ser Ile Gly Leu Asp Leu Gly Ile Ser Ser Val Gly Trp Ser

1 5 10 15

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

20 25 30

Phe Ser Ala Lys Asn Ser Glu Lys Asn Leu Glu Arg Arg Thr Ser Arg

35 40 45

Gly Ala Arg Arg Leu Ile Arg Arg Lys Thr Asn Arg Leu Lys Asp Ala

50 55 60

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

65 70 75 80

Asn Val Cys Pro Tyr Gln Leu Arg Val Lys Gly Leu Thr Glu Gly Leu

85 90 95

Thr Lys Gly Glu Leu Tyr Lys Val Val Leu His Ile Val Lys Lys Arg

100 105 110

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

115 120 125

Gln Asp Tyr Lys Glu Gln Val Arg Lys Asn Ala Gln Leu Leu Thr Lys

130 135 140

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

145 150 155 160

Val Lys Thr Gly Ile Asn Gly Gln Gly His Tyr Gln Leu Asn Val Phe

165 170 175

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

180 185 190

Gln Ala Leu Tyr Pro Asn Glu Leu Thr Asp Asp Trp Ile Ala Leu Phe

195 200 205

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

210 215 220

Pro Tyr Tyr His Gly Pro Gly Asn Glu Ala Asn Asn Ser Pro Tyr Gly

225 230 235 240

Arg Trp Ser Asp Phe Gln Lys Thr Gly Gln Pro Ala Ala Asn Ile Phe

245 250 255

Asp Lys Leu Ile Gly Lys Asp Phe Gln Gly Glu Leu Arg Ala Ser Gly

260 265 270

Leu Ser Leu Ser Ala Gln Gln Tyr Asn Leu Leu Asn Asp Leu Thr Asn

275 280 285

Leu Lys Ile Asp Gly Glu Val Ser Leu Ser Pro Glu Gln Lys Glu Phe

290 295 300

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

305 310 315 320

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

325 330 335

Arg Leu Asp Lys Lys Gly Lys Pro Glu Ile His Thr Leu Lys Gly Tyr

340 345 350

Arg Asn Trp Arg Lys Ile Phe Ala Glu Ala Gly Ile Asp Leu Ala Thr

355 360 365
Leu Pro Thr Glu Thr Ile Asp Cys Leu Ala Lys Val Leu Thr Leu Asn

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

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

500 505 510
Glu Met Pro Arg Asp Asp Asn Glu Glu Glu Arg Lys Arg Ile Lys
515 520 525
Glu Leu Gln Thr Lys Asn Ser Gln Arg Lys Asn Asp Ser Gln Gln Tyr
530 535 540
Phe Leu Gln Lys Ser Gly Trp Ser Gln Glu Lys Phe Gln Ala Thr Ile
545 550 555 560
His Lys Asn Arg Arg Phe Leu Ala Lys Leu Leu Tyr Tyr Phe Glu Gln

565 570 575
Asp Gly Val Cys Ala Tyr Thr Gly Asn Pro Ile Ser Pro Glu Leu Leu
580 585 590
Val Ser Asp Ser Thr Glu Ile Asp His Ile Ile Pro Ile Ser Ile Ser
595 600 605

Leu Asp Asp Ser Ile Asn Asn Lys Val Leu Val Leu Ser His Ala Asn
 610 615 620
 Gln Val Lys Gly Gln Gln Thr Pro Tyr Asp Ala Arg Met Ala Gly Ala

 625 630 635 640
 Phe Asn Lys Ile Asn Gly Lys Phe Ser Asn Trp Asp Glu Tyr Gln Lys
 645 650 655
 Trp Val Glu Ser Arg Pro Phe Ser Arg Lys Lys Val Asn Asn Leu Leu
 660 665 670
 Glu Thr Arg Asn Ile Phe Asp Ser Glu Gln Val Gln Lys Phe Leu Ser
 675 680 685
 Arg Asn Leu Asn Asp Thr Arg Tyr Ala Ser Arg Leu Val Leu Asn Thr

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

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

 820 825 830
 Asp Arg Lys Ala Asn Arg Lys Leu Ser Asp Ala Thr Ile Tyr Ser Val
 835 840 845
 Arg Glu Lys Thr Glu Val Lys Thr Leu Lys Ser Gly Lys Glu Lys Ile

850 855 860
 Thr Thr Asp Glu Tyr Thr Ile Gly Lys Ile Lys Asp Ile Tyr Thr Val
 865 870 875 880
 Asp Gly Trp Glu Ala Phe Lys Lys Lys Gln Asp Lys Leu Leu Met Lys

 885 890 895
 Glu Phe Asp Glu Lys Thr Tyr Glu Leu Leu Val Thr Ile Ala Ala Thr
 900 905 910
 Thr Pro Asp Phe Gln Glu Val Glu Glu Lys Asn Gly Lys Val Lys Arg
 915 920 925
 Val Lys Arg Ser Pro Phe Ala Val Tyr Cys Glu Glu Asn Gly Ile Pro
 930 935 940
 Ala Ile Arg Lys Tyr Ala Lys Lys Asn Asn Gly Pro Val Ile Arg Ser

 945 950 955 960
 Leu Lys Tyr Tyr Asp Gly Lys Leu Asn Lys His Ile Asn Ile Thr Lys
 965 970 975
 Asp Glu Lys Gly Arg Pro Val Glu Gln Thr Lys Asn Gly Arg Lys Val
 980 985 990
 Thr Leu Gln Ser Leu Lys Pro Tyr Arg Tyr Asp Ile Tyr Gln Asp Leu
 995 1000 1005
 Glu Thr Lys Ala Tyr Tyr Thr Val Gln Leu Tyr Tyr Ser Asp Leu

 1010 1015 1020
 Arg Phe Val Glu Gly Glu Tyr Gly Ile Thr Glu Lys Glu Tyr Met
 1025 1030 1035
 Lys Lys Val Ala Glu Gln Thr Lys Gly Gln Val Val Arg Phe Cys
 1040 1045 1050
 Phe Ser Leu Gln Lys Asn Asp Gly Leu Glu Ile Glu Trp Lys Asp
 1055 1060 1065
 Ser Gln Arg Tyr Asp Val Arg Phe Tyr Asn Phe Gln Ser Ala Asn

 1070 1075 1080
 Ser Ile Asn Phe Lys Gly Leu Glu Gln Glu Met Ile Pro Ala Glu
 1085 1090 1095

Asn Gln Phe Lys Gln Lys Pro Tyr Asn Asn Gly Ala Ile Asn Leu
1100 1105 1110

Asn Ile Ala Lys Tyr Gly Lys Glu Gly Lys Lys Leu Arg Lys Phe
1115 1120 1125

Asn Thr Asp Ile Leu Gly Lys Lys His His Leu Ser Tyr Glu Lys

1130 1135 1140

Glu Pro Lys Asn Ile Ile Lys

1145 1150

<210> 134

<211> 1150

<212> PRT

<213> Enterococcus sp.

<223> APG08770_RGN

<400> 134

Met Tyr Ser Ile Gly Leu Asp Leu Gly Ile Ser Ser Val Gly Trp Ser

1 5 10 15

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

20 25 30

Phe Ser Ala Lys Asn Ser Glu Lys Asn Leu Glu Arg Arg Thr Ser Arg

35 40 45

Gly Ala Arg Arg Leu Ile Arg Arg Lys Thr Asn Arg Leu Lys Asp Ala

50 55 60

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

65 70 75 80

Asn Val Cys Pro Tyr Gln Leu Arg Val Lys Gly Leu Thr Glu Gly Leu

85 90 95

Thr Lys Gly Glu Leu Tyr Lys Val Val Leu His Ile Val Lys Lys Arg

100 105 110

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

115 120 125

Gln Asp Tyr Lys Glu Gln Val Arg Lys Asn Ala Gln Leu Leu Thr Lys

130 135 140

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

385 390 395 400
 Leu Ala Glu Pro Val Lys Ser Leu Val Leu Asp His Tyr Lys Glu Leu
 405 410 415
 Ser Gln Ser Ile Ser Thr Gln Ala Trp His Arg Phe Ser Leu Lys Thr

 420 425 430
 Leu His Leu Leu Ile Pro Glu Leu Ile Lys Ser Thr Ser Glu Gln Asn
 435 440 445
 Thr Leu Leu Glu Gln Phe Gln Leu Lys Ala Gly Val Arg Lys Arg Tyr
 450 455 460
 Ser Asp Tyr Lys Lys Leu Pro Thr Lys Glu Val Leu Ala Glu Ile Tyr
 465 470 475 480
 Asn Pro Thr Val Asn Lys Thr Val Ser Gln Ala Phe Lys Val Met Asp

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

 545 550 555 560
 His Lys Asn Arg Arg Phe Leu Ala Lys Leu Leu Tyr Tyr Phe Glu Gln
 565 570 575
 Asp Gly Val Cys Ala Tyr Thr Gly Asn Pro Ile Ser Pro Glu Leu Leu
 580 585 590
 Val Ser Asp Ser Thr Glu Ile Asp His Ile Ile Pro Ile Ser Ile Ser
 595 600 605
 Leu Asp Asp Ser Ile Asn Asn Lys Val Leu Val Leu Ser His Ala Asn

 610 615 620
 Gln Val Lys Gly Gln Gln Thr Pro Tyr Asp Ala Trp Met Ala Gly Ala
 625 630 635 640

Phe Lys Lys Ile Asn Gly Lys Phe Ser Asn Trp Asp Glu Tyr Gln Lys
 645 650 655
 Trp Val Glu Ser Arg Pro Phe Ser Arg Lys Lys Ala Asn Asn Leu Leu
 660 665 670
 Glu Thr Arg Asn Ile Phe Asp Ser Glu Gln Val Gln Asn Phe Leu Ser
 675 680 685
 Arg Asn Leu Asn Asp Thr Arg Tyr Ala Ser Arg Leu Val Leu Asn Thr
 690 695 700
 Leu Gln Ser Phe Phe Glu Asn Gln Asp Thr Ile Val Arg Val Val Asn
 705 710 715 720
 Gly Ser Phe Thr His Thr Leu Arg Lys Lys Trp Gly Ala Asp Leu Asp
 725 730 735
 Lys Thr Arg Glu Thr His His His His Ala Ile Asp Ala Ser Leu Cys
 740 745 750
 Ala Val Thr Pro Phe Val Lys Val Ser Arg Tyr His Tyr Ala Val Asn
 755 760 765
 Glu Glu Thr Gly Glu Lys Phe Met Arg Glu Ile Asp Val Glu Thr Gly
 770 775 780
 Glu Ile Leu Asp Glu Ile Pro Tyr Arg Glu Tyr Lys Lys Ala Lys His
 785 790 795 800
 Tyr Glu Arg Lys Thr Tyr Gln Val Lys Trp Ser Asn Phe Arg Glu Gln
 805 810 815
 Leu Lys Pro Leu Thr Ile His Pro Lys Ile Lys Phe Ser His Gln Val
 820 825 830
 Asp Arg Lys Ala Asn Arg Lys Leu Ser Asp Ala Thr Ile Tyr Ser Val
 835 840 845
 Arg Glu Lys Thr Glu Val Lys Thr Leu Lys Ser Gly Lys Glu Lys Ile
 850 855 860
 Thr Thr Asp Glu Tyr Thr Ile Gly Lys Ile Lys Asp Ile Tyr Thr Val
 865 870 875 880
 Asp Gly Trp Glu Ala Phe Lys Lys Lys Gln Asp Lys Leu Leu Met Lys

885 890 895
 Glu Phe Asp Glu Lys Thr Tyr Glu Leu Leu Val Thr Ile Ala Ala Thr
 900 905 910
 Thr Pro Asp Phe Gln Glu Val Glu Glu Lys Asn Gly Lys Val Lys Arg
 915 920 925
 Val Lys Arg Ser Pro Phe Ala Val Tyr Cys Glu Glu Asn Gly Ile Pro

 930 935 940
 Ala Ile Arg Lys Tyr Ala Lys Lys Asn Asn Gly Pro Val Ile Arg Ser
 945 950 955 960
 Leu Lys Tyr Tyr Asp Gly Lys Leu Asn Lys His Ile Asn Ile Thr Lys
 965 970 975
 Asp Glu Lys Gly Arg Pro Val Glu Gln Thr Lys Asn Gly Arg Lys Val
 980 985 990
 Thr Leu Gln Ser Leu Lys Pro Tyr Arg Tyr Asp Ile Tyr Gln Asp Leu

 995 1000 1005
 Glu Thr Lys Ala Tyr Tyr Thr Val Gln Leu Tyr Tyr Ser Asp Leu
 1010 1015 1020
 Arg Phe Val Glu Gly Glu Tyr Gly Ile Thr Glu Lys Glu Tyr Met
 1025 1030 1035
 Lys Lys Val Ala Glu Gln Thr Lys Gly Gln Val Val Arg Phe Cys
 1040 1045 1050
 Phe Ser Leu Gln Lys Asn Asp Gly Leu Glu Ile Glu Trp Lys Asp

 1055 1060 1065
 Ser Gln Arg Tyr Asp Val Arg Phe Tyr Asn Phe Gln Ser Ala Asn
 1070 1075 1080
 Ser Ile Asn Phe Lys Gly Leu Glu Gln Glu Met Met Pro Ala Glu
 1085 1090 1095
 Asn Gln Phe Lys Gln Lys Pro Tyr Asn Asn Gly Ala Ile Asn Leu
 1100 1105 1110
 Asn Ile Ala Lys Tyr Gly Lys Glu Gly Asn Lys Leu Arg Lys Phe

 1115 1120 1125

Asn Thr Asp Ile Leu Gly Lys Lys His His Leu Ser Tyr Glu Lys

1130 1135 1140

Glu Pro Lys Asn Ile Ile Lys

1145 1150

<210> 135

<211> 1053

<212> PRT

<213> Staphylococcus sp.

<223> APG08167_RGN

<400> 135

Met Lys Arg His Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val

1 5 10 15

Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Glu Val Ile Asp Ala Gly

20 25 30

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

35 40 45

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Lys His Arg Leu

50 55 60

Gln Ser Val Lys Lys Leu Leu Phe Asp Tyr Asp Leu Leu Thr Asn His

65 70 75 80

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

85 90 95

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

100 105 110

Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr

115 120 125

Arg Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala

130 135 140

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Thr

145 150 155 160

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

165 170 175

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Glu Tyr His Gln
180 185 190

Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg
195 200 205

Arg Thr Tyr Tyr Glu Gly Pro Gly Lys Gly Ser Pro Phe Gly Trp Lys
210 215 220

Asp Ile Lys Glu Trp Tyr Glu Met Leu Ile Gly His Cys Thr Tyr Phe
225 230 235 240

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
245 250 255

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Ser Arg Asp Asp Asn
260 265 270

Glu Lys Leu Ala Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
275 280 285

Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ser Lys Glu Ile Phe
290 295 300

Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
305 310 315 320

Pro Glu Phe Thr Lys Phe Lys Ile Tyr His Asp Ile Lys Gly Ile Lys
325 330 335

Glu Arg Lys Glu Ile Ile Glu Asn Ala Glu Val Leu Asp Gln Ile Ala
340 345 350

Glu Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
355 360 365

Ala Lys Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
370 375 380

Lys Leu Thr Gly Tyr Thr Gly Thr His Ser Leu Ser Leu Lys Ala Ile
385 390 395 400

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Ser Asp Asn Gln Met Ala
405 410 415

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Met Asp Leu Ser Gln

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

 465 470 475 480
 Glu Lys Asn Ser Lys Asp Ala Gln Lys Leu Ile Asn Glu Met Gln Lys
 485 490 495
 Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
 500 505 510
 Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asn
 515 520 525
 Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu

 530 535 540
 Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
 545 550 555 560
 Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
 565 570 575
 Gln Glu Glu Asn Ser Lys Lys Asp Asn Arg Thr Pro Phe Gln Tyr Leu
 580 585 590
 Ser Ser Ser Asp Ser Lys Met Asp Tyr Glu Thr Phe Lys Lys His Ile

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

 660 665 670

Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
675 680 685

Lys Phe Lys Lys Glu Arg Asn Gln Gly Tyr Lys His His Ala Glu Asp
690 695 700

Ala Leu Ile Ile Ala Asn Ala Asp Tyr Val Phe Lys Glu Trp Lys Lys
705 710 715 720

Leu Asp Arg Ala Asn Lys Val Met Glu Asn Gln Thr Val Glu Glu Asn
725 730 735

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
740 745 750

Ile Phe Ile Thr Pro His Gln Ile Gln His Ile Lys Asp Phe Lys Asp
755 760 765

Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Ser Asn Arg Glu Leu Val
770 775 780

Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
785 790 795 800

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
805 810 815

Lys Lys Leu Met Asn Gln Ser Pro Glu Lys Leu Leu Met Tyr His His
820 825 830

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
835 840 845

Gly Glu Lys Asn Pro Leu Tyr Lys Tyr His Glu Glu Thr Gly Asn Tyr
850 855 860

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
865 870 875 880

Lys Tyr Tyr Gly Asn Lys Leu Asn Val His Leu Asp Ile Thr Asp Asp
885 890 895

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Phe
900 905 910

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

915 920 925
 Lys Asn Leu Asp Val Ile Lys Lys Glu Asp Tyr Tyr Glu Val Asn Ser
 930 935 940
 Lys Cys Tyr Lys Glu Ala Lys Lys Leu Lys Gln Ile Ser Asp Gln Ala
 945 950 955 960
 Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Val Asn Gly
 965 970 975
 Glu Leu Tyr Arg Ile Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile

980 985 990
 Glu Val Asn Met Val Asp Ile Thr Tyr Lys Glu Tyr Leu Glu Thr Met
 995 1000 1005
 Asn Asp Lys Arg Thr Pro Arg Ile Phe Lys Ala Ile Thr Gly Lys
 1010 1015 1020
 Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu
 1025 1030 1035
 Tyr Glu Val Lys Ser Lys Gln Asn Pro Gln Met Ile Met Lys Arg

1040 1045 1050
 <210> 136
 <211> 1052
 <212> PRT
 <213> Staphylococcus sp.
 <223> APG01604_RGN
 <400> 136
 Met Val Thr Lys Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val
 1 5 10 15
 Gly Tyr Gly Ile Ile Asn Tyr Glu Asp Lys Thr Ile Ile Asp Ala Gly
 20 25 30
 Val Arg Leu Phe Pro Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
 35 40 45
 Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Ile His Arg Leu
 50 55 60
 Asp Arg Ile Lys Gln Leu Leu Ser Glu Tyr Asn Leu Val Asp Leu Asp

65					70				75				80			
Asn	Ile	Pro	Gln	Ser	Pro	Ser	Pro	Tyr	Glu	Ile	Arg	Val	Lys	Gly	Leu	
				85				90				95				
Arg	Glu	Glu	Leu	Ser	Lys	Asp	Glu	Leu	Val	Ile	Ala	Leu	Leu	His	Ile	
				100				105				110				
Ala	Lys	Arg	Arg	Gly	Ile	His	Asn	Val	Glu	Ala	Val	Asp	Glu	Thr	Gln	
				115				120				125				
Asp	Glu	Gly	Asn	Glu	Leu	Ser	Thr	Lys	Glu	Gln	Leu	Ala	Lys	Asn	Asn	
				130				135				140				
Asn	Leu	Leu	Lys	Asp	Lys	Tyr	Val	Cys	Glu	Leu	Leu	Leu	Glu	Arg	Leu	
				145				150				155				
				160												
Lys	Asp	Gly	Lys	Val	Arg	Gly	Glu	Lys	Asn	Arg	Phe	Lys	Thr	Thr	Asp	
				165				170				175				
Ile	Ile	Lys	Glu	Val	Lys	Gln	Leu	Leu	Glu	Thr	Gln	Lys	Glu	Ala	His	
				180				185				190				
Gln	Leu	Asp	Asp	Asp	Phe	Ile	Asn	Arg	Tyr	Ile	Asp	Leu	Ile	Glu	Thr	
				195				200				205				
Arg	Arg	Glu	Tyr	Phe	Glu	Gly	Pro	Gly	Lys	Gly	Ser	Pro	Phe	Gly	Trp	
				210				215				220				
Gly	Gly	Asp	Leu	Lys	Lys	Trp	Tyr	Glu	Thr	Leu	Met	Gly	His	Cys	Thr	
				225				230				235				
				240												
Tyr	Phe	Pro	Asn	Glu	Leu	Arg	Ser	Val	Lys	Tyr	Ser	Tyr	Ser	Ala	Asp	
				245				250				255				
Leu	Phe	Asn	Ala	Leu	Asn	Asp	Leu	Asn	Asn	Leu	Val	Ile	Gln	Arg	Glu	
				260				265				270				
Gly	Asn	Ser	Lys	Leu	Glu	Tyr	His	Glu	Lys	Tyr	His	Ile	Ile	Glu	Asn	
				275				280				285				
Val	Phe	Lys	Gln	Lys	Lys	Lys	Pro	Thr	Leu	Lys	Gln	Ile	Ala	Asn	Glu	
				290				295				300				
Ile	Gly	Val	Ser	Pro	Asp	Asp	Ile	Lys	Gly	Phe	Arg	Ile	Asn	Lys	Ser	
				305				310				315				
				320												

Gly Lys Glu Thr Phe Thr Glu Phe Lys Leu Tyr His Asp Leu Lys Lys
325 330 335

Val Leu Ile Asp Gln Ser Ile Leu Glu Asn Val Gln Leu Leu Asp Gln
340 345 350

Ile Ala Glu Ile Leu Thr Ile Tyr Gln Asp Lys Glu Ser Ile Ile Asn
355 360 365

Glu Leu Asn Gln Leu Ser Glu Ile Ile Asn Glu Gln Asp Lys Glu Ser
370 375 380

Ile Ser Asn Leu Ser Gly Tyr Asn Gly Thr His Arg Leu Ser Leu Lys
385 390 395 400

Cys Ile Asn Leu Val Ile Glu Glu Leu Trp His Thr Ser Arg Asn Gln
405 410 415

Met Glu Ile Phe Thr Tyr Leu Asn Ile Lys Pro Lys Lys Ile Asp Leu
420 425 430

Ala Lys Thr Asn Lys Ile Pro Lys Asn Met Ile Asp Glu Phe Ile Leu
435 440 445

Ser Pro Val Val Lys Arg Thr Phe Gly Gln Ala Ile Asn Val Ile Asn
450 455 460

Lys Val Ile Glu Lys Tyr Gly Val Pro Glu Asp Ile Ile Ile Glu Leu
465 470 475 480

Ala Arg Glu Ser Asn Ser Lys Asp Lys Gln Lys Phe Ile Asn Ser Leu
485 490 495

Gln Lys Lys Asn Glu Thr Thr Arg Lys Arg Ile Asn Glu Ile Ile Gly
500 505 510

Gln Tyr Gly Asn Gln Asn Ala Lys Arg Leu Val Glu Lys Ile Arg Leu
515 520 525

His Asp Glu Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ser Ile Pro
530 535 540

Leu Glu Asp Leu Ile Asn Asn Pro Gln Tyr Tyr Glu Val Asp His Ile
545 550 555 560

Ile Pro Arg Ser Val Ser Phe Asp Asn Ser Tyr Gln Asn Lys Val Leu

565	570	575
Val Lys Gln Thr Glu Asn Ser Lys Lys Ser Asn Arg Thr Pro Tyr Gln		
580	585	590
Tyr Phe Asn Ser Gly Glu Thr Lys Leu Ser Tyr Asn Gln Phe Lys Gln		
595	600	605
His Val Leu Asn Leu Ser Lys Ser Lys Asp Arg Ile Ser Lys Lys Lys		
610	615	620
Lys Glu Tyr Leu Leu Glu Glu Arg Asp Ile Asn Lys Tyr Glu Val Gln		
625	630	635
Lys Glu Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg		
645	650	655
Glu Leu Thr Asn Tyr Leu Lys Ala Tyr Phe Ser Ala Asn Asp Met Asp		
660	665	670
Val Lys Val Lys Thr Ile Asn Gly Ser Phe Thr Asp Tyr Leu Arg Lys		
675	680	685
Val Trp Lys Phe Lys Lys Glu Arg Asn His Gly Tyr Lys His His Ala		
690	695	700
Glu Asp Ala Leu Ile Ile Ala Asn Ala Asp Phe Leu Phe Lys Glu Asn		
705	710	715
Lys Lys Leu Lys Lys Ala Asn Ala Ile Leu Glu Gln Pro Ser Leu Asp		
725	730	735
Asn Gly Lys Ser Asp Ala Thr Val Glu Asn Asp Asn Glu Tyr Val Glu		
740	745	750
Thr Phe Ser Ile Pro Lys Gln Val Asn Asp Ile Lys Glu Phe Arg Asp		
755	760	765
Phe Lys Phe Ser His Arg Val Asp Lys Lys Pro Asn Arg Gln Leu Ile		
770	775	780
Asn Asp Thr Leu Tyr Ser Thr Arg Lys Ile Glu Asn His Thr Phe Ile		
785	790	795
Val Ser Pro Ile Thr Asn Ile Tyr Ser Lys Asp Asn Asp Glu Leu Lys		
805	810	815

Lys Lys Phe Asn Lys Asn Pro Glu Lys Phe Leu Met Tyr Gln His Asp

820 825 830

Pro Lys Thr Phe Glu Lys Leu Glu Val Ile Met Lys Gln Tyr Ala Asn

835 840 845

Glu Lys Asn Pro Leu Ala Lys Tyr His Glu Glu Thr Gly Glu Tyr Leu

850 855 860

Thr Lys Tyr Ser Lys Lys Asn Asn Gly Pro Ile Val Lys Thr Ile Lys

865 870 875 880

Val Leu Gly Asp Lys Val Gly Lys His Leu Asp Val Thr His Lys Tyr

885 890 895

Lys Tyr Ser Asn Ser Lys Ile Val Lys Lys Thr Ile Asn Pro Tyr Arg

900 905 910

Phe Asp Val Tyr Leu Thr Asp Lys Gly Tyr Lys Phe Ile Thr Ile Ser

915 920 925

Tyr Leu Asp Val Leu Lys Lys Asp Asn Tyr Tyr Tyr Ile Leu Lys Glu

930 935 940

Lys Tyr Glu Glu Leu Lys Ile Lys Lys Ser Ile Ser Asp Thr Asp Gln

945 950 955 960

Phe Ile Gly Ser Phe Tyr Tyr Asn Asp Leu Ile Lys Ile Asn Asp Gln

965 970 975

Ile Phe Lys Val Val Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu

980 985 990

Leu Asp Leu Leu Asp Ile Ser Tyr Lys Glu Tyr Cys Lys Ile Asn Asn

995 1000 1005

Ile Lys Thr Asn Arg Ile Ile Lys Gly Ile Thr Lys Lys Ile Thr

1010 1015 1020

Asn Ile Glu Lys Phe Ser Thr Asp Val Leu Gly Asn Leu Tyr Lys

1025 1030 1035

Ala His Ser Asn His Pro Gln Leu Ile Phe Lys Gln Arg Asp

1040 1045 1050

<210> 137

<211> 1129

<212> PRT

<213> Streptococcus sp.

<223> APG03021_RGN

<400> 137

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

1 5 10 15

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

20 25 30

Ser Arg Leu Phe Ser Ala Ala Asn Ala Glu Asn Asn Ala Glu Arg Arg

35 40 45

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

50 55 60

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

65 70 75 80

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

85 90 95

Glu Gln Leu Thr Asn Glu Glu Leu Phe Ala Ala Leu Arg Thr Ile Ala

100 105 110

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

115 120 125

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

130 135 140

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

145 150 155 160

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

165 170 175

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

180 185 190

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

195 200 205

Glu Phe Ile Glu Asp Tyr Ile Glu Ile Leu Thr Gln Lys Arg Lys Tyr

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

 260 265 270
 Thr Ala Gln Glu Phe Asn Phe Leu Asn Asp Leu Asn Asn Leu Lys Val
 275 280 285
 Pro Thr Glu Thr Gly Lys Leu Ser Thr Glu Gln Lys Glu Tyr Leu Val
 290 295 300
 Asp Phe Ala Lys Lys Ser Lys Ala Leu Gly Ala Ser Lys Leu Leu Lys
 305 310 315 320
 Glu Ile Ala Lys Ile Val Asp Cys Ser Val Asp Asp Ile Lys Gly Tyr

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

 385 390 395 400
 Thr Glu Glu Gln Ile Ser Glu Ile Val Gln Ile Arg Lys Asn Gln Ser
 405 410 415
 Ser Ala Phe Asn Lys Gly Trp His Ser Phe Ser Ala Lys Leu Met Asn
 420 425 430
 Glu Leu Ile Pro Glu Leu Tyr Val Thr Ser Glu Glu Gln Met Thr Ile
 435 440 445
 Leu Thr Arg Leu Glu Lys Phe Lys Val Asn Lys Lys Ser Ser Lys Asn

 450 455 460

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

705 710 715 720
 Gln Leu Arg Arg Lys Trp Asn Ile Asp Lys Ser Arg Glu Thr Tyr His
 725 730 735
 His His Ala Val Asp Ala Leu Ile Ile Ala Ala Ser Ser Gln Leu Lys
 740 745 750
 Leu Trp Gln Lys Gln Glu Asn Pro Met Phe Glu Ser Tyr Gly Glu Asn
 755 760 765
 Gln Val Val Asn Lys Glu Thr Gly Glu Ile Leu Ser Ile Ser Asp Asp

 770 775 780
 Lys Tyr Lys Glu Leu Val Phe Gln Pro Pro Tyr Gln Gly Phe Val Asn
 785 790 795 800
 Thr Ile Ser Ser Lys Gly Phe Glu Asp Glu Ile Leu Phe Ser Tyr Gln
 805 810 815
 Val Asp Ser Lys Phe Asn Arg Lys Val Ser Asp Ala Thr Ile Tyr Ser
 820 825 830
 Thr Arg Lys Ala Lys Leu Gly Lys Asp Lys Lys Asp Glu Thr Tyr Val

 835 840 845
 Leu Gly Lys Ile Lys Asp Ile Tyr Ser Gln Asp Gly Phe Asp Thr Phe
 850 855 860
 Ile Lys Arg Tyr Lys Lys Asp Lys Thr Gln Phe Leu Met Tyr Gln Lys
 865 870 875 880
 Asp Pro Leu Thr Trp Glu Asn Val Ile Glu Val Ile Leu Arg Asp Tyr
 885 890 895
 Pro Ser Glu Lys Leu Ser Glu Asp Gly Lys Lys Thr Val Lys Cys Asn

 900 905 910
 Pro Phe Glu Glu Tyr Arg Arg Glu Asn Gly Leu Ile Cys Lys Tyr Ser
 915 920 925
 Lys Lys Gly Asn Gly Thr Pro Ile Lys Ser Leu Lys Tyr Tyr Asp Lys
 930 935 940
 Lys Leu Gly Asn Cys Ile Asp Ile Thr Pro Glu Lys Ser Lys Asn Arg
 945 950 955 960

Val Val Leu Arg Gln Ile Ser Pro Trp Arg Ala Asp Ile Tyr Phe Asn

965 970 975
 Leu Glu Thr Leu Lys Tyr Glu Leu Met Gly Leu Lys Tyr Ser Asp Leu
 980 985 990
 Ser Phe Glu Lys Gly Thr Gly Lys Tyr His Ile Ser Gln Glu Lys Tyr
 995 1000 1005
 Asp Ala Ile Arg Glu Lys Glu Gly Ile Gly Lys Lys Ser Glu Phe
 1010 1015 1020
 Lys Phe Thr Leu Tyr Arg Asn Asp Leu Ile Leu Ile Lys Asp Thr

1025 1030 1035
 Leu Asn Asn Cys Glu Arg Met Leu Arg Phe Gly Ser Lys Asn Asp
 1040 1045 1050
 Thr Ser Lys His Tyr Val Glu Leu Lys Pro Leu Glu Lys Gly Thr
 1055 1060 1065
 Phe Asp Ser Glu Glu Glu Ile Leu Pro Val Leu Gly Lys Val Ala
 1070 1075 1080
 Lys Ser Gly Gln Phe Ile Lys Gly Leu Asn Lys Thr Asn Ile Ser

1085 1090 1095
 Ile Tyr Lys Val Arg Thr Asp Val Leu Gly Asn Lys Phe Phe Ile
 1100 1105 1110
 Lys Lys Glu Gly Asp Lys Pro Lys Leu Asp Phe Lys Asn Asn Asn
 1115 1120 1125
 Lys

<210> 138

<211> 1359

<212> PRT

<213> *Pediococcus* sp.

<223> APG06015_RGN

<400> 138

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

1 5 10 15

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

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

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

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

Leu Asp Ala Ile Leu Lys Leu Asp Val Glu Glu Ser Lys Asp Trp Ser
260 265 270

Ile Ser Leu Asn Asp Glu Glu Ile Asp Ser Ile Leu Asp Asn Leu Thr
275 280 285

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

Leu Tyr Ser His Ile Ala Leu Asn Glu Ile Val Pro Asn Gly Gln Ser
305 310 315 320

Leu Ser Lys Ser Met Met Asp Lys Tyr Asp Lys His His Ala Asp Leu
325 330 335

Asp Val Leu Lys Lys Val Ile Ser Asn Met Asp Asp Arg Lys Lys Ala
340 345 350

Lys Ser Leu Lys Asn Ile Tyr Asn Gln Tyr Val Gly Lys Thr Asn Asp
355 360 365

Lys Val Leu Asp Lys Asp Glu Phe Tyr Lys Gln Ile Gln Lys Asn Leu
370 375 380

Asp Lys Ser Glu Asp Ala Met Lys Ile Val Asn Glu Ile Glu Leu Asp
385 390 395 400

Gln Phe Met Pro Lys Gln Arg Thr Ser Gln Asn Gly Val Ile Pro His
405 410 415

Gln Leu His Gln Lys Glu Leu Asp Glu Ile Ile Glu Asn Gln Lys Gln
420 425 430

Tyr Tyr Pro Phe Leu Ala Glu Pro Asn Pro Asn Glu Lys Arg Lys Pro
435 440 445

His Ala Gln Phe Lys Leu Asp Glu Leu Ile Ala Phe Lys Ile Pro Tyr
450 455 460

Tyr Val Gly Pro Leu Ile Thr Lys Glu Glu Gln Gln Ala Gln Ser Gly
465 470 475 480

Ala Lys Phe Ala Trp Met Lys Arg Lys Gln Asp Gly Val Ile Thr Pro
485 490 495

Trp Asn Phe Asp Glu Lys Val Asp Arg Met Ala Ser Ala Asn Glu Phe

500 505 510
 Ile Arg Arg Met Thr Thr Lys Asp Thr Tyr Leu Leu Gly Glu Asp Val
 515 520 525

Leu Pro Asp Glu Ser Leu Ile Tyr Gln Lys Phe Lys Val Leu Asn Glu
 530 535 540
 Leu Asn Asn Val Lys Val Asn Asp Lys Lys Leu Thr Val Ala Asp Lys
 545 550 555 560
 Gln Asp Ile Phe Asn Asp Leu Phe Lys Lys Gln Lys Thr Val Ser Val
 565 570 575
 Ser Lys Leu Gln Lys Tyr Phe Val Thr Glu Lys His Tyr Leu Thr Glu
 580 585 590

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

Trp Leu Thr Glu Lys Gln Ile Lys Ala Ile Arg Arg Tyr Arg Gly Trp
 660 665 670
 Gly Arg Leu Ser Lys Lys Leu Leu Val Asp Leu Arg Asn Asn Glu Gly
 675 680 685
 Lys Ser Ile Leu Asp Glu Leu Trp Arg Thr Asn Asp Asn Phe Met Gln
 690 695 700
 Ile Gln Ala Arg Glu Glu Phe Ala Lys Ser Ile Val Glu Ala Asn Gln
 705 710 715 720

Lys Leu Met Asn Leu Gly Gly Ala Ala Ser Val Gln Asn Thr Val Glu
 725 730 735
 Ser Thr Leu Glu Asp Ala Tyr Thr Ser Pro Gln Asn Lys Lys Ala Ile
 740 745 750

Arg Gln Val Ile Lys Val Val Glu Asp Ile Val Lys Ala Val Gly Tyr

755 760 765

Ala Pro Glu Lys Ile Thr Ile Glu Phe Thr Arg Gly Ala Asp Lys Asn

770 775 780

Ser Arg Arg Thr Gln Asn Arg Gln Lys Gln Ile Met Glu Val Tyr Lys

785 790 795 800

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

805 810 815

Asn Glu Glu Lys Leu Thr Asp Lys Leu Tyr Leu Tyr Phe Thr Gln Leu

820 825 830

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

835 840 845

Asn Tyr Asp Ile Asp His Ile Leu Pro Gln Ala Phe Ile Lys Asp Asp

850 855 860

Ser Leu Asp Asn Arg Val Leu Thr Ser Arg Asp Leu Asn Asn Gly Lys

865 870 875 880

Ser Asp Ser Val Pro Val Lys Gln Phe Gly Ala Asn Met Lys Ser Phe

885 890 895

Trp Met Arg Leu Gln Ala His Gly Leu Ile Ser Lys Arg Lys Leu Asn

900 905 910

Asn Leu Met Thr Asp Pro Asp Ser Ile Gly Lys Tyr Thr Met Gln Gly

915 920 925

Phe Val Arg Arg Gln Leu Val Glu Thr Ser Gln Val Ile Lys Leu Thr

930 935 940

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

945 950 955 960

Ile Pro Ala Lys Leu Thr His Gln Met Arg Glu Lys Phe Asn Leu Tyr

965 970 975

Lys Val Arg Glu Val Asn Asp Tyr His His Ala Phe Asp Ala Tyr Leu

980 985 990

Thr Thr Phe Val Gly Asn Tyr Leu Phe Lys Arg Tyr Pro Lys Leu Arg

995	1000	1005
Pro Tyr Phe Val Tyr Gly Asp	Phe Lys Met Thr Asp	Asn Ala Leu
1010	1015	1020
Lys Gly Met Arg Arg Phe Asn	Phe Leu His Asp Leu	Lys Asp Asp
1025	1030	1035
Glu Val Leu Val Asp Asn Glu	Thr Gly Glu Val Leu	Trp Glu Gly
1040	1045	1050
Gln Lys Ser Ile Glu Glu Leu	Lys Lys Ile Tyr Gly	Tyr Lys Phe
1055	1060	1065
Met Leu Thr Thr His Glu Ala	Tyr Ile Asn Asn Gly	Ala Met Tyr
1070	1075	1080
Asp Gln Thr Val Tyr Pro Ala	Lys Ser Asn Lys Lys	Leu Ile Pro
1085	1090	1095
Ser Lys Asn Asn Arg Pro Thr	Asn Ile Tyr Gly Gly	Phe Ser Gly
1100	1105	1110
Asn Lys Asp Ala Tyr Leu Ala	Ile Val Lys Ile Leu	Gly Lys Lys
1115	1120	1125
Gly Asp Thr Tyr Lys Val Val	Gly Val Pro Arg Arg	Phe Ser Asp
1130	1135	1140
Glu Leu Leu Asn Ala Tyr Asn	Asp Asp Tyr Ser Glu	Tyr Leu Asn
1145	1150	1155
Lys Leu His Asp Val Leu Val	Thr Gln Phe Thr Lys	Lys Gln Arg
1160	1165	1170
Asn Lys Lys Thr Gly Lys Ile	Ile Lys Lys Ile Glu	Asn Phe Glu
1175	1180	1185
Ile Val Val Gly Lys Val Asn	Tyr Arg Gln Arg Ile	Ile Asp Arg
1190	1195	1200
Asn Gln Ser Phe Met Leu Gly	Ser Ser Thr Tyr Lys	Tyr Asn Thr
1205	1210	1215
Lys Gln Leu Val Leu Thr Glu	Ser Ser Met Lys Ile	Leu Ser Asn
1220	1225	1230

Phe Lys His Ala Ser Asp Asp Asn Leu Asn Leu Val Tyr Asp Glu
 1235 1240 1245
 Ile Leu Glu Lys Val Asn Glu Ser Phe Glu Leu Tyr Asp Lys Asn
 1250 1255 1260
 Gly Phe Arg Lys Lys Leu Asn Asn Asn Ile Asp Lys Phe Arg Lys
 1265 1270 1275

 Leu Pro Cys Lys Ser Ile Phe Glu Lys Asn Lys Leu Ile Ser Val
 1280 1285 1290
 Gly Lys Cys Glu Ile Leu Phe Gln Ile Leu Asn Gly Leu His Ala
 1295 1300 1305
 Asn Ala Thr Met Gly Asp Leu Lys Gly Ile Gly Phe Ser Thr Pro
 1310 1315 1320
 Phe Gly Phe Met Gln Asp Ala Ser Gly Ile Lys Leu Asn Lys Asp
 1325 1330 1335

 Ser Met Leu Leu Tyr Gln Ser Pro Thr Gly Leu Phe Glu Arg Lys
 1340 1345 1350
 Val Lys Leu Ser Asp Leu
 1355
 <210> 139
 <211> 1488
 <212> PRT
 <213> Weissella sp.
 <223> APG09344_RGN
 <400> 139
 Met Ala Lys Gln Lys Tyr Thr Val Gly Leu Asp Ile Gly Thr Asn Ser
 1 5 10 15
 Val Gly Trp Ala Val Ile Asn Lys Asp Phe Lys Leu Val Ser Gly Lys
 20 25 30

 Lys His Ile Asn Asp Asn Gly Lys Gln Lys Arg Ser Arg Thr Asn Leu
 35 40 45
 Trp Gly Val Arg Leu Phe Asp Ala Ala Asp Thr Ala Ala Asp Arg Arg
 50 55 60

Leu Lys Arg Gly Met Arg Arg Arg Ile Ala Arg Arg Lys Glu Arg Leu
 65 70 75 80
 Asn Tyr Leu Arg Gly Ile Phe Gln Ala Asp Ile Leu Ala Phe Asp Asp
 85 90 95

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

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

 Glu Ala Ser Ser Phe Lys Phe Gly Leu Ala Asp Val Asp Ala Asp Gln
 225 230 235 240
 Ala Asn Ala Ile Leu Lys Asn Lys Gln Trp Ser Ala Ser Lys Lys Ala
 245 250 255
 Tyr Glu Leu Asn Thr Leu Tyr Glu Ile Ile Gln Asp Thr Val Tyr Arg
 260 265 270
 Arg Glu Asn Glu Thr Gln Ala Val Ala Tyr Asp Ser Leu Thr Val Lys
 275 280 285

 Gln Gln Glu Lys Trp Leu Ala Glu Lys Gln Lys Gln Val Lys Ala Phe
 290 295 300
 Leu Thr Gly Ile Val Gly Asn Thr Ile Ser Pro Lys Asp Ile Phe Ala

305 310 315 320
 Asn Lys Asp Tyr Asp Arg Lys Gln Asn Glu Asp Phe Pro Glu Lys Ile
 325 330 335
 Lys Tyr Ser Asn Glu Asn Phe Glu Glu Gln Ile Gly Glu Leu Glu Lys
 340 345 350

 Tyr Leu Asp Glu Thr Glu Val Ala Ala Ile Val Ala Gly Lys Asp Val
 355 360 365
 Tyr Glu Ser Leu Val Leu Ser Asn Ile Leu Thr Lys Ser Thr Leu Ser
 370 375 380
 Gly Ser Met Ile Glu Lys Tyr Asn Val His Gly Thr Gln Leu Lys Ser
 385 390 395 400
 Leu Lys Ser Phe Ser Arg Lys Val Ser Ala Glu Phe Tyr Asp Lys Leu
 405 410 415

 Phe Asn Val Val Lys Asp Glu Lys Gly Lys Pro Gln Ile Gly Val Tyr
 420 425 430
 Ala Gln Tyr Ile Asp Gly Val Gly Asn Pro Ala Lys Arg Leu Ala Arg
 435 440 445
 Glu Asp Phe Tyr Asp Ala Leu Lys Lys Val Phe Glu Ser Glu Phe Lys
 450 455 460
 Gly Leu Thr Phe Pro Ala Gly Glu Lys Gly Ile Asp Phe Thr Lys Thr
 465 470 475 480

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

 Tyr Glu Asp Glu Ser Gly Gln Thr Gln Val Arg Ser Glu Tyr Lys Ile
 545 550 555 560

Gln Val Leu Phe Lys Phe Arg Ile Pro Tyr Tyr Val Gly Thr Leu Ala
565 570 575

Lys Asn Ser Gly Trp Val Asn Glu Asp Gly Lys Lys Leu Val Ala Arg
580 585 590

Asn Val Ala Ala Lys Asn Ser Trp Val Val Arg Asn Ser Asp Glu Arg
595 600 605

Val Thr Pro Trp Asn Phe Glu Gln Val Ile Asn Lys Glu Ala Ser Ser
610 615 620

Ile Asn Phe Ile Glu Arg Met Thr Asn Phe Asp Thr Tyr Leu Pro Asn
625 630 635 640

Glu Lys Val Leu Pro Lys Asn Ser Leu Leu Tyr Gln Glu Phe Ile Val
645 650 655

Asn Asn Glu Leu Ile Ser Ser Gly Tyr Tyr Leu Asn Gly Arg Lys Glu
660 665 670

Tyr Phe Leu Pro Glu Gln Arg Gln Arg Ile Ile Thr Arg Leu Phe Lys
675 680 685

Arg Tyr Arg Lys Val Ser Ala Lys Gln Met Leu Asp Phe Leu Lys Asn
690 695 700

Glu Tyr Gln Ile Lys Leu Asp Asp Asp Pro Lys Lys Leu Phe Gly Leu
705 710 715 720

Asp Thr Phe Val Lys Val Pro Ser Tyr Asn Gly Ser Leu Ser Thr Val
725 730 735

Val Asp Leu Thr Asn Lys Met Gly Ile Ser Glu Asp Leu Ile Ala Asn
740 745 750

Tyr Pro Glu Arg Phe Glu Glu Ile Ile Lys Trp Gln Thr Ile Phe Glu
755 760 765

Asp Lys Lys Ile Leu Lys Lys Thr Ile Arg Glu Val Asn His Asn Lys
770 775 780

Trp Ser Gly Met Leu Thr Ala Glu Gln Val Asn Lys Leu Ser Lys Arg
785 790 795 800

His Tyr Thr Gly Trp Gly Arg Leu Ser Gln Lys Leu Leu Asp Gly Val

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

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

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

Thr Tyr Ile Lys Asp Asp Ser Leu Asp Asn Lys Val Leu Val Thr Gln
 995 1000 1005
 Arg Asp Asn Gln Asn Lys Gly Gly Asp Thr Pro Ser His Thr Val
 1010 1015 1020
 Val Thr Arg Met Asn Ser Tyr Trp Asp Met Leu Ala Lys Asn Gly
 1025 1030 1035
 Gln Val Ser Pro Arg Lys Leu Ala Asn Leu Lys Lys Gly Thr Ile
 1040 1045 1050

Thr	Leu	Lys	Gln	Arg	Glu	Gly	Phe	Ile	Asn	Arg	Gln	Leu	Val	Glu
1055						1060					1065			
Asn	Arg	Gln	Ile	Ile	Lys	His	Val	Ala	Asn	Ile	Leu	Thr	Asn	Tyr
1070						1075					1080			
Tyr	Thr	Gly	Thr	Asp	Thr	Leu	Val	Leu	Thr	Pro	Lys	Ser	Gly	Leu
1085						1090					1095			
Thr	Ser	Gln	Leu	Arg	Ser	Gly	Leu	Ile	Tyr	Glu	Leu	Asn	Pro	Ala
1100						1105					1110			
Phe	Asp	Glu	Ala	Lys	Ala	Leu	Asp	Arg	Gly	Lys	Ala	Tyr	Gln	Val
1115						1120					1125			
Glu	Arg	Tyr	Thr	Thr	Val	Lys	Leu	His	Asp	Arg	Phe	Val	Lys	Asn
1130						1135					1140			
Arg	Lys	Leu	Asn	Asp	Tyr	His	His	Ala	His	Asp	Ala	Tyr	Leu	Asn
1145						1150					1155			
Ala	Phe	Val	Ala	Gln	Tyr	Val	Tyr	Gln	Glu	His	Pro	Glu	Trp	Lys
1160						1165					1170			
Asn	Ala	Trp	Val	Tyr	Gly	Lys	Tyr	Pro	Arg	Asn	Gly	Gln	Ala	Asp
1175						1180					1185			
Phe	Gly	Lys	Trp	Ala	Thr	Gln	Arg	Lys	Gln	Lys	Ser	Leu	Gln	Leu
1190						1195					1200			
Leu	Ser	Ser	Met	Ala	Asn	Asp	Val	Trp	Asn	Leu	Glu	Asp	Pro	Asp
1205						1210					1215			
Thr	Met	Glu	Lys	Thr	Ile	Leu	Asn	Arg	Asp	Glu	Thr	Phe	Glu	Gln
1220						1225					1230			
Met	Arg	Gln	Thr	Leu	Cys	Tyr	Arg	Asn	Ile	Asn	Ile	Val	Lys	Lys
1235						1240					1245			
Leu	Glu	Thr	Gln	Thr	Gly	Lys	Phe	Gly	Asp	Glu	Ser	Val	Tyr	Lys
1250						1255					1260			
Lys	Gly	Asn	Lys	Ala	Asp	His	Tyr	Ser	Leu	Gly	Leu	Lys	Arg	Lys
1265						1270					1275			
Tyr	Pro	Pro	Gln	Lys	Tyr	Gly	Gly	Thr	Lys	Gly	Ala	Ile	Ser	Ala

1280	1285	1290
Val Thr Ala Leu Val Lys Asp Asn Ser Gly Arg Val Ile Pro Val		
1295	1300	1305
Ser Ile Ser Ala Ser Asp Tyr Asp Ser Tyr Gln Asn Ala Ser Asp		
1310	1315	1320
Lys Leu Ala Trp Leu Leu Met Arg Lys Lys Asn Ile Ala Glu Ile		
1325	1330	1335
Leu Val Ala Ser Leu Pro Lys Tyr Thr Lys Tyr Glu Leu Pro Ser		
1340	1345	1350
Gly Ala Leu Arg Leu Leu Ala Ser Tyr Gln Glu Ala Gln Ser Ala		
1355	1360	1365
Val Glu Leu Pro Met Leu Asn Leu Gly Leu Met Asp Asp Glu Ala		
1370	1375	1380
Lys Val Leu Asn Leu Tyr Asp Gln Leu Ala Gln Phe Ile Ala Asp		
1385	1390	1395
Asn Lys Leu Phe Thr Asp Lys Lys Ile Ala Leu Leu Lys Gly Glu		
1400	1405	1410
Met Arg Val Ala Phe Glu Gly Leu Glu Glu Phe Lys Asp Lys Gln		
1415	1420	1425
Lys Val Ile Asp Glu Leu Leu Gly Val Thr Asn Gly Ser Asn Gln		
1430	1435	1440
Gly Leu Gln Ala Leu Asn Ser Ile Gly Leu Gly Thr Ser Asn Gln		
1445	1450	1455
Arg Leu Lys Ser Gly Asn Thr Met Thr Asn Gly Val Thr Leu Ile		
1460	1465	1470
Asn Glu Ser Val Thr Gly Leu Tyr Ser Thr Arg Thr Thr Tyr Asn		
1475	1480	1485
<210> 140		
<211> 1337		
<212> PRT		
<213> Enterococcus sp.		

<223> APG07991_RGN

<400> 140

Met Lys Lys Asp Tyr Val Ile Gly Leu Asp Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Met Thr Glu Asp Tyr Gln Leu Val Lys Lys Lys Met

20 25 30

Pro Ile Tyr Gly Asn Thr Glu Lys Lys Lys Ile Lys Lys Asn Phe Trp

35 40 45

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

50 55 60

Lys Arg Thr Ala Arg Arg Arg Ile Ser Arg Arg Arg Asn Arg Leu Arg

65 70 75 80

Tyr Leu Gln Ala Phe Phe Glu Glu Ala Met Thr Ala Leu Asp Glu Asn

85 90 95

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

100 105 110

Trp His Arg His Pro Ile Phe Ala Lys Leu Glu Asp Glu Val Ala Tyr

115 120 125

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

130 135 140

Ser Ser Glu Gln Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145 150 155 160

Ile Val Lys Tyr Arg Gly His Phe Leu Ile Glu Gly Lys Leu Ser Thr

165 170 175

Glu Asn Ile Ser Val Lys Glu Gln Phe Gln Gln Phe Met Ile Ile Tyr

180 185 190

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

195 200 205

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

210 215 220

Thr Lys Lys Ser Glu Lys Val Leu Gln Gln Phe Pro Gln Glu Lys Ala

225 230 235 240
 Asn Gly Leu Phe Gly Gln Phe Leu Lys Leu Met Val Gly Asn Lys Ala
 245 250 255
 Asp Phe Lys Lys Val Phe Gly Leu Glu Glu Glu Ala Lys Ile Thr Tyr
 260 265 270
 Ala Ser Glu Ser Tyr Glu Glu Asp Leu Glu Gly Ile Leu Ala Lys Val
 275 280 285
 Gly Asp Glu Tyr Ser Asp Val Phe Leu Ala Ala Lys Asn Val Tyr Asp

 290 295 300
 Ala Val Glu Leu Ser Thr Ile Leu Ala Asp Ser Asp Lys Lys Ser His
 305 310 315 320
 Ala Lys Leu Ser Ser Ser Met Ile Val Arg Phe Thr Glu His Gln Glu
 325 330 335
 Asp Leu Lys Asn Phe Lys Arg Phe Ile Arg Glu Asn Cys Pro Asp Glu
 340 345 350
 Tyr Asp Asn Leu Phe Lys Asn Glu Gln Lys Asp Gly Tyr Ala Gly Tyr

 355 360 365
 Ile Ala His Ala Gly Lys Val Ser Gln Leu Lys Phe Tyr Gln Tyr Val
 370 375 380
 Lys Lys Ile Ile Gln Asp Ile Ala Gly Ala Glu Tyr Phe Leu Glu Lys
 385 390 395 400
 Ile Ala Gln Glu Asn Phe Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly
 405 410 415
 Val Ile Pro His Gln Ile His Leu Ala Glu Leu Gln Ala Ile Ile His

 420 425 430
 Arg Gln Ala Ala Tyr Tyr Pro Phe Leu Lys Glu Asn Gln Lys Lys Ile
 435 440 445
 Glu Gln Leu Val Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ser
 450 455 460
 Lys Gly Asp Ala Ser Thr Phe Ala Trp Leu Lys Arg Gln Ser Glu Glu
 465 470 475 480

Pro Ile Arg Pro Trp Asn Leu Gln Glu Thr Val Asp Leu Asp Gln Ser

485 490 495

Ala Thr Ala Phe Ile Glu Arg Met Thr Asn Phe Asp Thr Tyr Leu Pro

500 505 510

Ser Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Lys Phe Met

515 520 525

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

530 535 540

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

545 550 555 560

Phe Lys Thr Arg Arg Lys Val Lys Lys Lys Asp Ile Ile Gln Phe Tyr

565 570 575

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

580 585 590

Asp Gln Phe Asn Ala Ser Phe Ser Thr Tyr Gln Asp Leu Leu Lys Cys

595 600 605

Gly Leu Thr Arg Ala Glu Leu Asp His Pro Asp Asn Ala Glu Lys Leu

610 615 620

Glu Asp Ile Ile Lys Ile Leu Thr Ile Phe Glu Asp Arg Gln Arg Ile

625 630 635 640

Arg Thr Gln Leu Ser Thr Phe Lys Gly Gln Phe Ser Glu Glu Val Leu

645 650 655

Lys Lys Leu Glu Arg Lys His Tyr Thr Gly Trp Gly Arg Leu Ser Lys

660 665 670

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

675 680 685

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

690 695 700

Phe Met Gln Leu Ile Asn Asp Ser Gln Leu Ser Phe Lys Asn Ala Ile

705 710 715 720

Gln Lys Ala Gln Ser Ser Glu His Glu Glu Thr Leu Ser Glu Thr Val

725 730 735
Asn Glu Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Tyr Gln Ser

740 745 750
Leu Lys Ile Val Asp Glu Leu Val Ala Ile Met Gly Tyr Ala Pro Lys

755 760 765
Arg Ile Val Val Glu Met Ala Arg Glu Asn Gln Thr Thr Ser Thr Gly

770 775 780
Lys Arg Arg Ser Ile Gln Arg Leu Lys Ile Val Glu Lys Ala Met Ala

785 790 795 800
Glu Ile Gly Ser Asn Leu Leu Lys Glu Gln Pro Thr Thr Asn Glu Gln

805 810 815
Leu Arg Asp Thr Arg Leu Phe Leu Tyr Tyr Met Gln Asn Gly Lys Asp

820 825 830
Met Tyr Thr Gly Asp Glu Leu Ser Leu His Arg Leu Ser His Tyr Asp

835 840 845
Ile Asp His Ile Ile Pro Gln Ser Phe Met Lys Asp Asp Ser Leu Asp

850 855 860
Asn Leu Val Leu Val Gly Ser Thr Glu Asn Arg Gly Lys Ser Asp Asp

865 870 875 880
Val Pro Ser Lys Glu Val Val Lys Lys Met Lys Ala Tyr Trp Glu Lys

885 890 895
Leu Tyr Ala Ala Gly Leu Ile Ser Gln Arg Lys Phe Gln Arg Leu Thr

900 905 910
Lys Gly Glu Gln Gly Gly Leu Thr Leu Glu Asp Lys Ala His Phe Ile

915 920 925
Gln Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys Asn Val Ala Gly

930 935 940
Ile Leu Asp Gln Arg Tyr Asn Ala Lys Ser Lys Glu Lys Lys Val Gln

945 950 955 960
Ile Ile Thr Leu Lys Ala Ser Leu Thr Ser Gln Phe Arg Ser Ile Phe

965 970 975

Gly Leu Tyr Lys Val Arg Glu Val Asn Asp Tyr His His Gly Gln Asp
980 985 990

Ala Tyr Leu Asn Cys Val Val Ala Thr Thr Leu Leu Lys Val Tyr Pro
995 1000 1005

Asn Leu Ala Pro Glu Phe Val Tyr Gly Glu Tyr Pro Lys Phe Gln
1010 1015 1020

Ala Phe Lys Glu Asn Lys Ala Thr Ala Lys Ala Ile Ile Tyr Thr
1025 1030 1035

Asn Leu Leu Arg Phe Phe Thr Glu Asp Glu Pro Arg Phe Thr Lys
1040 1045 1050

Asp Gly Glu Ile Leu Trp Ser Asn Ser Tyr Leu Lys Thr Ile Lys
1055 1060 1065

Lys Glu Leu Asn Tyr His Gln Met Asn Ile Val Lys Lys Val Glu
1070 1075 1080

Val Gln Lys Gly Gly Phe Ser Lys Glu Ser Ile Lys Pro Lys Gly
1085 1090 1095

Pro Ser Asn Lys Leu Ile Pro Val Lys Asn Gly Leu Asp Pro Gln
1100 1105 1110

Lys Tyr Gly Gly Phe Asp Ser Pro Val Val Ala Tyr Thr Val Leu
1115 1120 1125

Phe Thr His Glu Lys Gly Lys Lys Pro Leu Ile Lys Gln Glu Ile
1130 1135 1140

Leu Gly Ile Thr Ile Met Glu Lys Thr Arg Phe Glu Gln Asn Pro
1145 1150 1155

Ile Leu Phe Leu Glu Glu Lys Gly Phe Leu Arg Pro Arg Val Leu
1160 1165 1170

Met Lys Leu Pro Lys Tyr Thr Leu Tyr Glu Phe Pro Glu Gly Arg
1175 1180 1185

Arg Arg Leu Leu Ala Ser Ala Lys Glu Ala Gln Lys Gly Asn Gln
1190 1195 1200

Met Val Leu Pro Glu His Leu Leu Thr Leu Leu Tyr His Ala Lys

1205 1210 1215
Gln Cys Leu Leu Pro Asn Gln Ser Glu Ser Leu Ala Tyr Val Glu
1220 1225 1230
Gln His Gln Pro Glu Phe Gln Glu Ile Leu Glu Arg Val Val Asp

1235 1240 1245
Phe Ala Glu Val His Thr Leu Ala Lys Ser Lys Val Gln Gln Ile
1250 1255 1260
Val Lys Leu Phe Glu Ala Asn Gln Thr Ala Asp Val Lys Glu Ile
1265 1270 1275
Ala Ala Ser Phe Ile Gln Leu Met Gln Phe Asn Ala Met Gly Ala
1280 1285 1290
Pro Ser Thr Phe Lys Phe Phe Gln Lys Asp Ile Glu Arg Ala Arg

1295 1300 1305
Tyr Thr Ser Ile Lys Glu Ile Phe Asp Ala Thr Ile Ile Tyr Gln
1310 1315 1320
Ser Thr Thr Gly Leu Tyr Glu Thr Arg Arg Lys Val Val Asp
1325 1330 1335

<210> 141

<211> 1340

<212> PRT

<213> Enterococcus sp.

<223> APG01868_RGN

<400> 141

Met Lys Lys Glu Tyr Thr Ile Gly Leu Asp Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ser Val Leu Thr Asp Asp Tyr Arg Leu Val Ser Lys Lys Met

20 25 30

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

35 40 45

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

50 55 60

Lys Arg Thr Ala Arg Arg Arg Leu Ala Arg Arg Arg Gln Arg Ile Leu

65				70				75				80			
Glu	Leu	Gln	Lys	Ile	Phe	Ala	Pro	Glu	Ile	Leu	Lys	Ile	Asp	Glu	His
				85				90				95			
Phe	Phe	Ala	Arg	Leu	Asn	Glu	Ser	Phe	Leu	Val	Pro	Asp	Glu	Lys	Lys
100				105				110							
Gln	Ser	Arg	His	Pro	Val	Phe	Ala	Thr	Ile	Lys	Gln	Glu	Lys	Ser	Tyr
115				120				125							
His	Gln	Thr	Tyr	Pro	Thr	Ile	Tyr	His	Leu	Arg	Gln	Ala	Leu	Ala	Asp
130				135				140							
Ser	Ser	Glu	Lys	Ala	Asp	Ile	Arg	Leu	Val	Tyr	Leu	Ala	Met	Ala	His
145				150				155				160			
Leu	Leu	Lys	Tyr	Arg	Gly	His	Phe	Leu	Ile	Glu	Gly	Glu	Leu	Asn	Thr
165				170				175							
Glu	Asn	Ser	Ser	Val	Thr	Glu	Thr	Phe	Arg	Gln	Phe	Leu	Ser	Thr	Tyr
180				185				190							
Asn	Gln	Gln	Phe	Ser	Glu	Ala	Gly	Asp	Lys	Gln	Thr	Glu	Lys	Leu	Asp
195				200				205							
Glu	Ala	Val	Asp	Cys	Ser	Phe	Val	Phe	Thr	Glu	Lys	Met	Ser	Lys	Thr
210				215				220							
Lys	Lys	Ala	Glu	Thr	Leu	Leu	Lys	Tyr	Phe	Pro	His	Glu	Lys	Ser	Asn
225				230				235				240			
Gly	Tyr	Leu	Ser	Gln	Phe	Ile	Lys	Leu	Met	Val	Gly	Asn	Gln	Gly	Asn
245				250				255							
Phe	Lys	Asn	Val	Phe	Gly	Leu	Glu	Glu	Glu	Ala	Lys	Leu	Gln	Phe	Ser
260				265				270							
Lys	Glu	Thr	Tyr	Glu	Glu	Asp	Leu	Gln	Glu	Leu	Leu	Glu	Lys	Ile	Gly
275				280				285							
Asp	Asp	Tyr	Ile	Asp	Leu	Phe	Val	Gln	Ala	Lys	Asn	Val	Tyr	Asp	Ala
290				295				300							
Val	Leu	Leu	Ser	Glu	Ile	Leu	Ser	Asp	Ser	Thr	Lys	Asn	Thr	Arg	Ala
305				310				315				320			

Lys Leu Ser Ala Gly Met Ile Arg Arg Tyr Asp Ala His Lys Glu Asp
 325 330 335

Leu Val Leu Leu Lys Arg Phe Val Lys Glu Asn Leu Pro Lys Lys Tyr
 340 345 350

Arg Ala Phe Phe Gly Asp Asn Ser Val Asn Gly Tyr Ala Gly Tyr Ile
 355 360 365

Glu Gly His Ala Thr Gln Glu Asp Phe Tyr Lys Phe Val Lys Lys Glu
 370 375 380

Leu Thr Gly Ile Arg Gly Ser Glu Val Phe Leu Thr Lys Ile Glu Gln
 385 390 395 400

Glu Asn Phe Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Val Ile Pro
 405 410 415

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

Lys His Tyr Pro Phe Leu Lys Glu Glu Gln Glu Lys Leu Glu Ser Leu
 435 440 445

Leu Thr Phe Lys Ile Pro Tyr Tyr Val Gly Pro Leu Ala Lys Lys Gln
 450 455 460

Glu Asn Ser Pro Phe Ala Trp Leu Ile Arg Lys Ser Glu Glu Lys Ile
 465 470 475 480

Lys Pro Trp Asn Leu Pro Glu Ile Val Asp Met Glu Asp Ser Ala Val
 485 490 495

Arg Phe Ile Glu Arg Met Ile Asn Thr Asp Met Tyr Met Pro His Asn
 500 505 510

Lys Val Leu Pro Lys Asn Ser Leu Leu Tyr Gln Lys Phe Ser Ile Tyr
 515 520 525

Asn Glu Leu Thr Lys Val Arg Tyr Gln Asp Glu Arg Gly Gln Met Asn
 530 535 540

Tyr Phe Ser Ser Ile Glu Lys Lys Glu Ile Phe His Glu Leu Phe Glu
 545 550 555 560

Lys Asn Arg Lys Val Thr Lys Lys Asp Leu Gln Glu Phe Leu Tyr Leu

565 570 575
 Lys Tyr Asp Ile Lys His Ala Glu Leu Ser Gly Ile Glu Lys Ala Phe
 580 585 590

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

 Ile Lys Lys Leu Thr Lys Lys His Tyr Thr Gly Trp Gly Arg Leu Ser
 660 665 670
 Lys Glu Leu Ile Gln Gly Ile Arg Asp Lys Gln Ser Asn Lys Thr Ile
 675 680 685
 Leu Asp Tyr Leu Ile Asn Asp Asp Asp Phe Pro His His Arg Asn Arg
 690 695 700
 Asn Phe Met Gln Leu Ile Asn Asp Asp Ser Leu Ser Phe Lys Lys Glu
 705 710 715 720

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

 Gly Leu Lys Ser Ser Arg Pro Arg Leu Lys Ala Leu Glu Glu Ser Leu
 785 790 795 800
 Lys Asp Phe Gly Ser Gln Leu Leu Lys Glu Tyr Pro Thr Asp Asn Ser
 805 810 815

Ser Leu Gln Lys Asp Arg Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg
820 825 830

Asp Met Tyr Thr Gly Ala Pro Leu Asp Ile His Arg Leu Ser Asp Tyr
835 840 845

Asp Ile Asp His Ile Ile Pro Arg Ser Phe Thr Thr Asp Asn Ser Ile
850 855 860

Asp Asn Lys Val Leu Val Ser Ser Lys Glu Asn Arg Leu Lys Lys Asp
865 870 875 880

Asp Val Pro Ser Glu Lys Val Val Lys Lys Met Arg Ser Phe Trp Tyr
885 890 895

Asp Leu Tyr Ser Ser Lys Leu Ile Ser Lys Arg Lys Leu Asp Asn Leu
900 905 910

Thr Lys Ile Lys Leu Thr Glu Glu Asp Lys Ala Gly Phe Ile Lys Arg
915 920 925

Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gly Ile Leu
930 935 940

His His Arg Phe Asn Lys Ala Glu Asp Thr Asn Glu Pro Ile Arg Lys
945 950 955 960

Val Arg Ile Ile Thr Leu Lys Ser Ala Leu Val Ser Gln Phe Arg Asn
965 970 975

Arg Phe Gly Ile Tyr Lys Val Arg Glu Ile Asn Glu Tyr His His Ala
980 985 990

His Asp Ala Tyr Leu Asn Gly Val Ile Ala Leu Ala Leu Leu Lys Lys
995 1000 1005

Tyr Pro Gln Leu Ala Pro Glu Phe Val Tyr Gly Glu Tyr Leu Lys
1010 1015 1020

Phe Asn Ala His Lys Ala Asn Lys Ala Thr Val Lys Lys Glu Phe
1025 1030 1035

Tyr Ser Asn Ile Met Lys Phe Phe Glu Ser Asp Thr Pro Val Cys
1040 1045 1050

Asp Glu Asn Gly Glu Ile Phe Trp Asp Lys Ser Lys Ser Ile Ala

1055	1060	1065
Gln Val Lys Lys Val Ile Asn His His His Met Asn Ile Val Lys		
1070	1075	1080
Lys Thr Glu Ile Gln Lys Gly Gly Phe Ser Lys Glu Thr Val Glu		
1085	1090	1095
Pro Lys Lys Asp Ser Ser Lys Leu Leu Pro Arg Lys Asn Asn Trp		
1100	1105	1110
Asp Pro Ala Lys Tyr Gly Gly Leu Gly Ser Pro Asn Val Ala Tyr		
1115	1120	1125
Thr Val Ala Phe Thr Tyr Glu Lys Gly Lys Ala Arg Lys Arg Thr		
1130	1135	1140
Asn Ala Leu Glu Gly Ile Thr Ile Met Glu Arg Glu Ala Phe Glu		
1145	1150	1155
Gln Ser Pro Val Leu Phe Leu Lys Asn Lys Gly Tyr Glu Gln Ala		
1160	1165	1170
Glu Ile Glu Met Lys Leu Pro Lys Tyr Ala Leu Phe Glu Leu Glu		
1175	1180	1185
Asn Gly Arg Lys Arg Met Val Ala Ser Asn Lys Glu Ala Gln Lys		
1190	1195	1200
Ala Asn Ser Phe Leu Leu Pro Glu His Leu Val Thr Leu Leu Tyr		
1205	1210	1215
His Ala Lys Gln Tyr Asp Glu Ile Ser His Lys Glu Ser Phe Asp		
1220	1225	1230
Tyr Val Asn Glu His His Lys Glu Phe Ser Glu Val Phe Ala Arg		
1235	1240	1245
Val Leu Glu Phe Ala Gly Lys Tyr Thr Leu Ala Glu Lys Asn Ile		
1250	1255	1260
Glu Lys Leu Glu Lys Ile Tyr Lys Glu Asn Gln Thr Asp Asp Leu		
1265	1270	1275
Ala Lys Leu Ala Ser Ser Phe Val Asn Leu Met Gln Phe Asn Ala		
1280	1285	1290

Met Gly Ala Pro Ala Asp Phe Lys Phe Phe Asp Val Thr Ile Pro
 1295 1300 1305
 Arg Lys Arg Tyr Thr Ser Leu Thr Glu Ile Trp Gln Ser Thr Ile
 1310 1315 1320
 Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Arg Met
 1325 1330 1335

Gly Lys

1340

<210> 142

<211> 1327

<212> PRT

<213> Enterococcus sp.

<223> APG02998_RGN

<400> 142

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

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

Asn Glu Arg His Pro Ile Phe Gly Asn Val Ala Glu Glu Val Ala Tyr
 115 120 125
 His Glu Gln Phe Arg Thr Ile Tyr Arg Leu Arg Glu Val Leu Ala Asn
 130 135 140

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

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

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

 Asp Ser Tyr Leu Pro Met Glu Lys Val Leu Pro Lys His Ser Leu Leu
 500 505 510
 Tyr Glu Lys Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Arg Tyr Met
 515 520 525
 Asp Asp Asn Gly Arg Val Gln Asn Phe Ser Ser Lys Glu Lys Gln Gln
 530 535 540
 Ile Val Asn Asp Leu Phe Lys Gln Gln Arg Lys Val Lys Lys Glu Ile
 545 550 555 560

 Leu Glu Ala Phe Ile Lys Asn Glu Tyr Gly Ile Glu Asn Pro Lys Val
 565 570 575
 Glu Gly Ile Glu Lys Ala Phe Asn Ala Ser Leu Gly Thr Tyr His Asp
 580 585 590
 Leu Val Lys Leu Gly Ile Arg Pro Ala Leu Phe Glu Gln Pro Glu Tyr
 595 600 605
 Glu Gln Gln Phe Glu Gln Ile Val Lys Ile Leu Thr Val Phe Glu Asp
 610 615 620

 Arg Lys Met Arg Arg Lys Gln Leu Glu Gln Phe Ser Asp Leu Leu Thr
 625 630 635 640

Glu Asp Glu Gln Lys Gln Leu Glu Arg Lys His Tyr Lys Gly Trp Gly
 645 650 655
 Arg Leu Ser Ala Lys Leu Ile His Gly Ile Val Asp Gln Lys Thr Gln
 660 665 670
 Lys Thr Ile Leu Asp Tyr Leu Ile Asp Asp Asp Asp Leu Pro Arg Asn
 675 680 685

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

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

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

 Ala Thr Trp Glu Arg Leu Tyr Ala Ser Asn Leu Ile Ser Lys Lys Lys

885	890	895
Leu Asn Asn Leu Met Lys Ala Val Asn Gly Gly Leu Thr Glu Glu Asp		
900	905	910
Lys Ala Tyr Phe Ile Gln Arg Gln Leu Val Glu Thr Arg Gln Ile Thr		
915	920	925
Lys Asn Val Ala Gln Ile Leu His Gln Lys Tyr Asn Asn Glu Gln Ser		
930	935	940
Ser Glu Lys Pro Val Arg Val Val Thr Leu Lys Ser Ala Leu Ala Ser		
945	950	955
Gln Phe Arg Lys Asp Phe Ser Leu Tyr Lys Ile Arg Glu Leu Asn Asp		
965	970	975
Tyr His His Ala Gln Asp Ala Tyr Leu Asn Gly Val Ile Ala Gln Ala		
980	985	990
Leu Leu Lys Val Tyr Pro Lys Leu Glu Pro Glu Phe Val Tyr Gly Glu		
995	1000	1005
Tyr Gln Lys Val Ser Ile Arg Ala Leu Asn Lys Ala Thr Ala Lys		
1010	1015	1020
Lys Glu Thr Tyr Ser Asn Ile Met Lys Phe Phe Thr Asn Asp Glu		
1025	1030	1035
Val Leu Thr Asn Lys Glu Thr Gly Glu Val Leu Trp Asn Arg Gln		
1040	1045	1050
Arg Asp Ile Lys Thr Ile Lys Lys Val Met Asn Tyr His Gln Met		
1055	1060	1065
Asn Ile Val Lys Lys Val Glu Ile Gln Thr Gly Arg Phe Thr Lys		
1070	1075	1080
Glu Ser Ile Gln Pro Lys Gly Pro Ser Lys Lys Leu Ile Ala Arg		
1085	1090	1095
Lys Asn Asn Trp Ala Pro Val Asn Tyr Gly Gly Phe Asp Ser Pro		
1100	1105	1110
Thr Val Ala Tyr Ser Val Ile Ile Thr His Glu Lys Gly Lys Ala		
1115	1120	1125

Ala Lys	Val Val Gln Gln Leu	Val Gly Ile Lys Ile	Leu Glu Arg
1130	1135	1140	
Gln Ala	Phe Glu Gln Asn Glu	Val Ala Phe Leu Glu	Gly Lys Gly
1145	1150	1155	
Phe Ile	His Pro Lys Val Gln	Leu Lys Leu Pro Lys	Tyr Ser Leu
1160	1165	1170	
Tyr Gln	Phe Ala Asp Gly Arg	Arg Arg Leu Leu Ala	Ser Ala Asp
1175	1180	1185	
Glu Ala	Gln Lys Gly Asn Gln	Met Val Leu Pro Thr	Glu Leu Ile
1190	1195	1200	
Glu Leu	Leu Tyr His Ala Lys	His Val Asn Asp Ser	Ser Gly Glu
1205	1210	1215	
Ser Leu	Thr Tyr Leu Glu Asn	His Arg Thr Lys Phe	Ala Glu Leu
1220	1225	1230	
Leu Glu	Val Ile Leu Gln Phe	Thr Asn Arg Tyr Val	Gly Ala Glu
1235	1240	1245	
Asp Arg	Gln Lys Thr Val Arg	Asp Leu Tyr Glu Lys	Asn Gln Asp
1250	1255	1260	
Ala Asp	Met Lys Glu Leu Ala	Pro Ser Phe Val Gln	Leu Leu Gln
1265	1270	1275	
Leu Asn	Lys Gln Gly Ala Pro	Ala Ser Phe Lys Phe	Phe Gly Glu
1280	1285	1290	
Thr Ile	Glu Arg Ser Arg Tyr	Lys Asn Met Thr Glu	Leu Asn Lys
1295	1300	1305	
Ala Thr	Phe Ile Asn Gln Ser	Ile Thr Gly Leu Tyr	Glu Thr Gln
1310	1315	1320	
Arg Arg	Leu Val		
1325			