



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0006686
(43) 공개일자 2024년01월15일

- (51) 국제특허분류(Int. Cl.) (71) 출원인
C12N 9/18 (2006.01) C08J 11/10 (2006.01) **삼사라 예코 피티와이 리미티드**
- (52) CPC특허분류 (72) 발명자
C12N 9/18 (2013.01) 호주, 2065, 뉴사우스웨일즈, 702/28-34 클라크 스트리트 크로우스 네스트
C08J 11/105 (2013.01)
- (21) 출원번호 10-2023-7043010 **손더스, 제이크 웨슬러**
- (22) 출원일자(국제) 2022년05월13일 호주, 2065, 뉴사우스웨일즈, 크로우스 네스트, 28-34 클라크 스트리트, 유닛 702
심사청구일자 없음
- (85) 번역문제출일자 2023년12월12일 **스펜스, 매튜**
- (86) 국제출원번호 PCT/AU2022/050455 호주, 2065, 뉴사우스웨일즈, 크로우스 네스트, 28-34 클라크 스트리트, 유닛 702
- (87) 국제공개번호 WO 2022/236377 **(뒷면에 계속)**
- 국제공개일자 2022년11월17일
- (30) 우선권주장 (74) 대리인
2021901431 2021년05월13일 오스트레일리아(AU) **김경교, 양용**

전체 청구항 수 : 총 48 항

(54) 발명의 명칭 효소 변이체 및 이의 용도

(57) 요약

본 발명은 일반적으로 모노-(2-하이드록시에틸)테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드에 관한 것으로, 보다 구체적으로는, (i) 서열번호 1과 적어도 70%의 서열 동일성을 가지고 (ii) 서열번호 1의 아미노산 위치 156 내지 396, 398 내지 410, 및 425 내지 603에 상응하는 위치로 이루어진 그룹으로부터 선택된 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하는 폴리펩티드에 관한 것이다.

대표도 - 도1a

```

>MHETase_wild-type (WT; SEQ ID NO:1)
GGGSTPLPLQQQPQQEPPPPVPLASRAACEALKDGNNDMVWPNAAATVVEAAWRDAA
PATAAALPEHCEVSGIAKRTGIDGYPYEKFLRMPAEWNGRPFMEGSGTNGSLSA
ATCSGGGQIASALSRGFVAVATDSEGHDAVNDPDLGTVAFGLDQARLDMGVNSYDQ
VTQAGAAVARYPGRADKSYFGCSSEGRGEMLSQRPSHYDGVAGAPQYLPKAGI
SGAWTTQSLAPAAVGLDAQVPLNKSFSDADLHLSQALGTCALDGLADGIVDNYRA
CQAAPATAANPANGALQCCVGAKTADCLSPVQVTAKRAMAGPVNSAGTFLYRWAWD
AGMSLGGSTYINQWRSWILGSPFNSSANNAQRVSGFSARSWLVDFATPEPMPMTQVAAR
MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAARFRDGGKMLIYHMSDAAFSALD
TADYERLGAAMPAAAGFARLFLVRGMNHCSSGPTDRFDMFLTVAVVERGEAPDQISA
WSGTPQYVGAARTPLCPYQIARYKSGSDINTEANFACAAPP
>MHETase_Consensus95_AminoAcid (SEQ ID NO:2)
GGGSTPLPLQQQPQQEPPPPVPLASRAACEALKDGNNDMVWPNAAATVVEAAWRDAA
PATAAALPEHCEVSGIAKRTGIDGYPYEKFLRMPAEWNGRPFMEGSGTNGSLSA
ATCSGGGQIASALSRGFVAVATDSEGHDAVNDPDLGTVAFGLDQARLDMGVNSYDQ
VTQAGAAVARYPGRADKSYFGCSSEGRGEMLSQRPSHYDGVAGAPQYLPKAGI
SGAWTTQSLAPAAVGLDAQVPLNKSFSDADLHLSQALGTCALDGLADGIVDNYRA
CQAAPATAANPANGALQCCVGAKTADCLSPVQVTAKRAMAGPVNSAGTFLYRWAWD
AGMSLGGSTYINQWRSWILGSPFNSSANNAQRVSGFSARSWLVDFATPEPMPMTQVAAR
MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAARFRDGGKMLIYHMSDAAFSALD
TADYERLGAAMPAAAGFARLFLVRGMNHCSSGPTDRFDMFLTVAVVERGEAPDQISA
WSGTPQYVGAARTPLCPYQIARYKSGSDINTEANFACAAPP
>MHETase_Consensus95_Y197N_AminoAcid (SEQ ID NO:3)
GGGSTPLPLQQQPQQEPPPPVPLASRAACEALKDGNNDMVWPNAAATVVEAAWRDAA
PATAAALPEHCEVSGIAKRTGIDGYPYEKFLRMPAEWNGRPFMEGSGTNGSLSA
ATCSGGGQIASALSRGFVAVATDSEGHDAVNDPDLGTVAFGLDQARLDMGVNSYDQ
VTQAGAAVARYPGRADKSYFGCSSEGRGEMLSQRPSHYDGVAGAPQYLPKAGI
SGAWTTQSLAPAAVGLDAQVPLNKSFSDADLHLSQALGTCALDGLADGIVDNYRA
CQAAPATAANPANGALQCCVGAKTADCLSPVQVTAKRAMAGPVNSAGTFLYRWAWD
AGMSLGGSTYINQWRSWILGSPFNSSANNAQRVSGFSARSWLVDFATPEPMPMTQVAAR
MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAARFRDGGKMLIYHMSDAAFSALD
TADYERLGAAMPAAAGFARLFLVRGMNHCSSGPTDRFDMFLTVAVVERGEAPDQISA
WSGTPQYVGAARTPLCPYQIARYKSGSDINTEANFACAAPP
>MHETase_Consensus95_197T89N_AminoAcid (SEQ ID NO:4)
GGGSTPLPLQQQPQQEPPPPVPLASRAACEALKDGNNDMVWPNAAATVVEAAWRDAA
PATAAALPEHCEVSGIAKRTGIDGYPYEKFLRMPAEWNGRPFMEGSGTNGSLSA
ATCSGGGQIASALSRGFVAVATDSEGHDAVNDPDLGTVAFGLDQARLDMGVNSYDQ
VTQAGAAVARYPGRADKSYFGCSSEGRGEMLSQRPSHYDGVAGAPQYLPKAGI
SGAWTTQSLAPAAVGLDAQVPLNKSFSDADLHLSQALGTCALDGLADGIVDNYRA
CQAAPATAANPANGALQCCVGAKTADCLSPVQVTAKRAMAGPVNSAGTFLYRWAWD
AGMSLGGSTYINQWRSWILGSPFNSSANNAQRVSGFSARSWLVDFATPEPMPMTQVAAR
MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAARFRDGGKMLIYHMSDAAFSALD
TADYERLGAAMPAAAGFARLFLVRGMNHCSSGPTDRFDMFLTVAVVERGEAPDQISA
WSGTPQYVGAARTPLCPYQIARYKSGSDINTEANFACAAPP
>MHETase_Consensus95_197T89N_AminoAcid (SEQ ID NO:5)
GGGSTPLPLQQQPQQEPPPPVPLASRAACEALKDGNNDMVWPNAAATVVEAAWRDPA
PATAAALPEHCEVSGIAKRTGIDGYPYEKFLRMPAEWNGRPFMEGSGTNGSLSA
ATCSGGGQIASALSRGFVAVATDSEGHDAVNDPDLGTVAFGLDQARLDMGVNSYDQ
VTQAGAAVARYPGRADKSYFGCSSEGRGEMLSQRPSHYDGVAGAPQYLPKAGI
SGAWTTQSLAPAAVGLDAQVPLNKSFSDADLHLSQALGTCALDGLADGIVDNYRA
CQAAPATAANPANGALQCCVGAKTADCLSPVQVTAKRAMAGPVNSAGTFLYRWAWD
AGMSLGGSTYINQWRSWILGSPFNSSANNAQRVSGFSARSWLVDFATPEPMPMTQVAAR
MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAARFRDGGKMLIYHMSDAAFSALD
TADYERLGAAMPAAAGFARLFLVRGMNHCSSGPTDRFDMFLTVAVVERGEAPDQISA
WSGTPQYVGAARTPLCPYQIARYKSGSDINTEANFACAAPP

```

(52) CPC특허분류

C12Y 301/01 (2013.01)

(72) 발명자

봉사우디, 베네사

호주, 2065, 뉴사우스웨일즈, 크로우스 네스트,
28-34 클라크 스트리트, 유닛 702

담리, 아담 마이클

호주, 2065, 뉴사우스웨일즈, 크로우스 네스트,
28-34 클라크 스트리트, 유닛 702

잭슨, 콜린 존

호주, 2065, 뉴사우스웨일즈, 크로우스 네스트,
28-34 클라크 스트리트, 유닛 702

명세서

청구범위

청구항 1

모노-(2-하이드록시에틸)테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드로서, (i) 서열번호 1과 적어도 70%의 서열 동일성을 가지고 (ii) 서열번호 1의 아미노산 위치 156 내지 396, 398 내지 410, 및 425 내지 603에 상응하는 위치로 이루어진 그룹으로부터 선택된 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하는, 폴리펩티드.

청구항 2

제1항에 있어서, 상기 폴리펩티드가 하기로 이루어진 그룹으로부터 선택되는 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하는, 폴리펩티드:

- (i) 서열번호 1의 아미노산 위치 156에 상응하는 위치;
- (ii) 서열번호 1의 아미노산 위치 159에 상응하는 위치;
- (iii) 서열번호 1의 아미노산 위치 192에 상응하는 위치; 및
- (iv) 서열번호 1의 아미노산 위치 503에 상응하는 위치.

청구항 3

제1항에 있어서, 상기 폴리펩티드가 서열번호 1의 아미노산 위치 252에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 11과 상이한 아미노산 서열을 포함하는, 폴리펩티드.

청구항 4

제2항 또는 제3항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 156에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 5

제2항 내지 제4항 중 어느 한 항에 있어서, 서열번호 1의 아미노산 위치 156에 상응하는 위치에서의 상기 아미노산 치환이 N156G, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 6

제2항 내지 제5항 중 어느 한 항에 있어서, 상기 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 159에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 7

제2항 내지 제6항 중 어느 한 항에 있어서, 서열번호 1의 아미노산 위치 159에 상응하는 위치에서의 상기 아미노산 치환이 T159V, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 8

제2항 내지 제7항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 252에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 9

제2항 내지 제8항 중 어느 한 항에 있어서, 서열번호 1의 아미노산 위치 156에 상응하는 위치에서의 상기 아미노산 치환이 Y252F, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 10

제2항 내지 제9항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 11

제2항 내지 제10항 중 어느 한 항에 있어서, 서열번호 1의 아미노산 위치 503에 상응하는 위치에서의 상기 아미노산 치환이 Y503W, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 12

제2항 내지 제11항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 159, 252 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 13

제2항 내지 제12항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 159, 192, 252 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 14

제2항 내지 제13항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 159, 192 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 15

제2항 내지 제14항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 156, 159 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 16

제2항 내지 제15항 중 어느 한 항에 있어서, 상기 아미노산 치환이 T159V, Y252F 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환인, 폴리펩티드.

청구항 17

제2항 내지 제11항, 및 제13항 중 어느 한 항에 있어서, 상기 아미노산 치환이 T159V, M192Y, Y252F 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환인, 폴리펩티드.

청구항 18

제2항 내지 제11항, 및 제14항 중 어느 한 항에 있어서, 상기 아미노산 치환이 N156G, M192Y 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환인, 폴리펩티드.

청구항 19

제2항 내지 제11항, 및 제15항 중 어느 한 항에 있어서, 상기 아미노산 치환이 N156G, T159V, M192Y 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환인, 폴리펩티드.

청구항 20

제1항에 있어서, 상기 폴리펩티드가 하기로 이루어진 그룹으로부터 선택되는 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하는, 폴리펩티드:

- (i) 서열번호 1의 아미노산 위치 156에 상응하는 위치;
- (ii) 서열번호 1의 아미노산 위치 159에 상응하는 위치;
- (iii) 서열번호 1의 아미노산 위치 196에 상응하는 위치;
- (iv) 서열번호 1의 아미노산 위치 197에 상응하는 위치;

- (v) 서열번호 1의 아미노산 위치 260에 상응하는 위치;
- (vi) 서열번호 1의 아미노산 위치 264에 상응하는 위치;
- (vii) 서열번호 1의 아미노산 위치 267에 상응하는 위치;
- (viii) 서열번호 1의 아미노산 위치 286에 상응하는 위치; 및
- (ix) 서열번호 1의 아미노산 위치 503에 상응하는 위치.

청구항 21

제20항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 196에 상응하는 위치에 서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 22

제20항 또는 제21항에 있어서, 서열번호 1의 아미노산 위치 196에 상응하는 위치에서의 상기 아미노산 치환이 S196A, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 23

제20항 내지 제22항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 197에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 24

제20항 내지 제23항 중 어느 한 항에 있어서, 서열번호 1의 아미노산 위치 197에 상응하는 위치에서의 상기 아미노산 치환이 Y197V, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 25

제20항 내지 제24항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 260에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 26

제20항 내지 제25항 중 어느 한 항에 있어서, 서열번호 1의 아미노산 위치 260에 상응하는 위치에서의 상기 아미노산 치환이 S260A, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 27

제20항 내지 제26항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 264에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 28

제20항 내지 제27항 중 어느 한 항에 있어서, 서열번호 1의 아미노산 위치 264에 상응하는 위치에서의 상기 아미노산 치환이 S264L, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 29

제20항 내지 제28항 중 어느 한 항에 있어서, 서열번호 1의 아미노산 위치 267에 상응하는 위치에서의 상기 아미노산 치환이 S267A, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 30

제20항 내지 제29항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 286에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 31

제20항 내지 제30항 중 어느 한 항에 있어서, 서열번호 1의 아미노산 위치 286에 상응하는 위치에서의 상기 아

미노산 치환이 S286A, 또는 이의 보존적 아미노산 치환인, 폴리펩티드.

청구항 32

제20항 내지 제31항 중 어느 한 항에 있어서, 상기 폴리펩티드의 상기 아미노산 서열은 서열번호 1의 아미노산 위치 156 및 159에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한, 폴리펩티드.

청구항 33

제32항에 있어서, 상기 아미노산 치환이 N156G 및 T159V 또는 임의의 전술한 것의 보존적 아미노산 치환인, 폴리펩티드.

청구항 34

모노-(2-하이드록시에틸)테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드로서, (i) 서열번호 1과 적어도 70%의 서열 동일성을 가지고 (ii) MHETase의 폴리에스테르 기질과 접촉하지 않는 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하며, 폴리펩티드의 MHETase 활성은 서열번호 1의 MHETase의 MHETase 활성보다 더 큰, 폴리펩티드.

청구항 35

제1항 내지 제34항 중 어느 한 항에 있어서, 서열번호 1의 MHETase 대비,

- (i) 숙주 세포 시스템 내 증가된 재조합 발현;
- (ii) 증가된 전 세포 활성; 및
- (iii) 증가된 열안정성 중 하나 이상을 더 포함하는, 폴리펩티드.

청구항 36

제1항 내지 제35항 중 어느 한 항의 폴리펩티드를 포함하는 조성물.

청구항 37

제1항 내지 제36항 중 어느 한 항의 폴리펩티드를 암호화하는 핵산 서열을 포함하는 폴리뉴클레오티드.

청구항 38

제37항에 있어서, 상기 핵산 서열이 서열번호 37-72 및 78-82로 이루어진 그룹으로부터 선택되는, 폴리뉴클레오티드.

청구항 39

제37항 또는 제38항의 폴리뉴클레오티드를 포함하는 발현 벡터.

청구항 40

제37항 또는 제38항의 폴리뉴클레오티드, 또는 제39항의 발현 벡터를 포함하는 숙주 세포.

청구항 41

- (a) 제37항 또는 제38항의 폴리뉴클레오티드를 제공하는 단계;
- (b) 숙주 세포 배양물에서 핵산 서열을 발현시켜 폴리펩티드를 생산하는 단계; 및
- (c) 상기 숙주 세포 배양물로부터 (b)에서 생산된 폴리펩티드를 수집하는 단계를 포함하는, MHETase 활성을 갖는 폴리펩티드의 생산 방법.

청구항 42

모노-(2-하이드록시에틸)테레프탈레이트를 테레프탈레이트 및 에틸렌 글리콜로 전환하는데 충분한 조건 하에서, 모노-(2-하이드록시에틸)테레프탈레이트를 제1항 내지 제35항 중 어느 한 항의 폴리펩티드, 제36항의 조성물 또

는 제40항의 숙주 세포에 노출시키는 단계를 포함하는, 모노-(2-하이드록시에틸)테레프탈레이트의 가수분해 방법.

청구항 43

플라스틱 제품을 제1항 내지 제35항 중 어느 한 항의 폴리펩티드, 제36항의 조성물 또는 제40항의 숙주 세포에 노출시키는 단계를 포함하는, 폴리에스테르를 포함하는 플라스틱 제품의 분해 방법.

청구항 44

제43항에 있어서, 상기 폴리에스테르는 폴리락트산(PLA), 폴리트리메틸렌 테레프탈레이트(PTT), 폴리부틸렌 테레프탈레이트(PBT), 폴리에틸렌 아이소소르바이드 테레프탈레이트(PEIT), 폴리에틸렌 테레프탈레이트(PET), 폴리하이드록시알카노에이트(PHA), 폴리부틸렌 숙시네이트(PBS), 폴리부틸렌 숙시네이트 아디페이트(PBSA), 폴리부틸렌 아디페이트 테레프탈레이트(PBAT), 폴리에틸렌 푸라노에이트(PEF), 폴리카프로락톤(PCL), 폴리(에틸렌아디페이트)(PEA), 폴리(글리콜산)(PGA), 폴리(락틱-co-글리콜산)(PLGA) 및 임의의 전술한 것의 조합으로 이루어진 그룹으로부터 선택되는, 방법.

청구항 45

제44항에 있어서, 상기 폴리에스테르가 폴리에틸렌 테레프탈레이트(PET)인, 방법.

청구항 46

제45항에 있어서,

(a) PETase가 PET의 전환을 촉매하여 모노-(2-하이드록시에틸)테레프탈레이트(MHET)를 생성하는데 충분한 조건 하에서, PET를 폴리에틸렌 테레프탈레이트 에스테라제(PETase)에 노출시키는 단계; 및

(b) 동시에 또는 순차적으로, 상기 폴리펩티드가 MHET의 가수분해를 촉매하여 테레프탈레이트 및 에틸렌 글리콜을 생산하는데 충분한 조건 하에서, 단계 (a)에서 생성된 MHET를 제1항 내지 제35항 중 어느 한 항의 폴리펩티드, 제36항의 조성물 또는 제40항의 숙주 세포에 노출시키는 단계를 포함하는, 방법.

청구항 47

제46항에 있어서, 단계 (b)에서 생산된 테레프탈레이트 및/또는 에틸렌 글리콜을 회수하는 단계를 더 포함하는, 방법.

청구항 48

제47항의 방법에 의해 회수된 상기 테레프탈레이트 및/또는 상기 에틸렌 글리콜을 포함하는 조성물.

발명의 설명

기술 분야

[0001] 본 발명은 신규 효소, 더욱 상세하게는 모노-(2-하이드록시에틸)테레프탈산의 에스테르 결합을 가수분해하는 재조합 효소 및 이의 용도에 관한 것이다.

배경 기술

[0002] 본 명세서에 인용된 임의의 특허 또는 특허 출원을 포함한 모든 참고문헌은 본 발명의 완전한 이해가 가능하도록 본원에 참고로 포함된다. 그럼에도 불구하고, 이러한 참고문헌은 이들 문서 중 어느 하나가 호주 또는 기타 국가에서 해당 기술 분야의 일반 지식의 일부를 형성한다는 사실을 인정하는 것으로 해석되어서는 안 된다.

[0003] 글로벌 산업화는 환경에 심각한 영향을 미쳤으며, 이 중 특히 플라스틱 및 플라스틱 제품에 대한 제조 및 의존도가 증가했다. 제조 및 폐기를 포함하여 플라스틱에 대한 적합하고 환경적으로 지속 가능한 대안을 찾으려는 노력이 증가하는 반면, 이러한 제품은 여전히 심각한 문제로 남아 있으며 대부분의 환경 오염 물질에 기여한다. 이러한 문제의 주요 원인 중 하나는 폴리에틸렌 테레프탈레이트(PET)와 이의 폐기물로, 매년 전 세계적으로 수백만 톤이 생산된다. 이 문제의 환경적 유의성은 적어도 부분적으로는 플라스틱, 특히 PET 기반 제품이 자연에

서 쉽게 분해되지 않는, 화학적 특성에 기인한다.

[0004] 플라스틱 폐기물 문제를 처리하기 위한 접근법은 일반적으로 소각, 매립 처리 및 기계적 분해를 포함한다. 그러나 이들 접근 방식은 또한 환경에 심각한 영향을 준다. 예를 들어, 플라스틱을 소각하면 대기 중으로 방출되는 잠재적으로 유해한 부산물이 생산되고; 매립지에서 플라스틱의 분해 속도는 일반적으로 매우 느리고 독성 물질이 지하수로 침출될 위험이 있으며; 기계적 분해는 상대적으로 비용이 비싸고 부산물 사용이 제한되는 경우가 많다.

[0005] 보다 최근에는 플라스틱의 생물학적(효소적) 분해가 플라스틱 폐기물 축적을 감소시키기 위한 대안으로 고려되어왔다. 이 접근법에는 PET를 모노-2-하이드록시에틸 테레프탈레이트(MHET)로 가수분해하는 것을 촉매하는 에스테라제 계열의 효소인 PETase의 사용을 포함한다. 또한 하이드록시에틸 테레프탈레이트 가수분해효소(MHETase)를 활용하여 MHET를 테레프탈레이트와 에틸렌 글리콜로 가수분해할 수 있다. 이러한 가수분해물은 플라스틱을 포함한 새로운 제품의 제조를 위한 재료로 적절하게 재활용할 수 있다.

[0006] 플라스틱의 효소 분해는 플라스틱 폐기물 및 이의 처리에 대한 환경적 영향을 완화시키는 매력적인 대안이지만, 일반적인 산업 숙주 균주에서 상대적 인 비효율성, 느린 효소 분해 속도 및 낮은 수준의 효소 발현으로 인해 아직 널리 채택되지는 않았다. 따라서, 플라스틱의 효소 분해를 위한 개선된 방법 및 시약에 대한 시급한 필요성이 존재한다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0007] 본 명세서에 개시된 국면에서, (i) 서열번호 1과 적어도 70%의 서열 동일성을 가지고 (ii) 서열번호 1의 아미노산 위치 156 내지 396, 398 내지 410, 및 425 내지 603에 상응하는 위치로 이루어진 그룹으로부터 선택된 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하는, 모노-(2-하이드록시에틸)테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드가 제공된다.

[0008] 다른 국면에서, 하기로 이루어진 그룹으로부터 선택되는 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하는 폴리펩티드가 제공된다:

[0009] (i) 서열번호 1의 아미노산 위치 156에 상응하는 위치;

[0010] (ii) 서열번호 1의 아미노산 위치 159에 상응하는 위치;

[0011] (iii) 서열번호 1의 아미노산 위치 192에 상응하는 위치; 및

[0012] (iv) 서열번호 1의 아미노산 위치 503에 상응하는 위치.

[0013] 하나의 실시양태에서, 폴리펩티드는 서열번호 1의 아미노산 위치 159, 252 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다. 다른 실시양태에서, 아미노산 치환은 T159V, Y252F 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환이다.

[0014] 다른 실시양태에서, 폴리펩티드는 서열번호 1의 아미노산 위치 159, 192, 252 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다. 다른 실시양태에서, 아미노산 치환은 T159V, M192Y, Y252F 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환이다.

[0015] 다른 국면에서, 모노-(2-하이드록시에틸)테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드로서, (a) 서열번호 1과 적어도 70%의 서열 동일성을 가지고, (b) 하기로 이루어진 그룹으로부터 선택되는 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하는 폴리펩티드가 제공된다:

[0016] (i) 서열번호 1의 아미노산 위치 156에 상응하는 위치;

[0017] (ii) 서열번호 1의 아미노산 위치 159에 상응하는 위치;

[0018] (iii) 서열번호 1의 아미노산 위치 196에 상응하는 위치;

- [0019] (iv) 서열번호 1의 아미노산 위치 197에 상응하는 위치;
- [0020] (v) 서열번호 1의 아미노산 위치 260에 상응하는 위치;
- [0021] (vi) 서열번호 1의 아미노산 위치 264에 상응하는 위치;
- [0022] (vii) 서열번호 1의 아미노산 위치 267에 상응하는 위치;
- [0023] (viii) 서열번호 1의 아미노산 위치 286에 상응하는 위치; 및
- [0024] (ix) 서열번호 1의 아미노산 위치 503에 상응하는 위치.
- [0025] 일부 실시양태에서, 본원에 개시된 폴리펩티드는, 서열번호 1의 MHETase 대비, (i) 숙주 세포 시스템 내 증가된 제조합 발현; (ii) 증가된 전체 세포 활성; 및 (iii) 증가된 열안정성 중 하나 이상을 더 포함한다.
- [0026] 본 개시내용은 또한 본원에 기술된 바와 같은 폴리펩티드를 포함하는 조성물로 확장된다.
- [0027] 본 개시내용은 또한 본원에 기술된 폴리펩티드를 암호화하는 핵산 서열로 확장된다. 하나의 실시양태에서, 폴리펩티드를 암호화하는 핵산 서열은 서열번호 37-72 및 79-84로부터 선택된다. 바람직한 실시양태에서, 폴리펩티드를 암호화하는 핵산 서열은 서열번호 79-84로부터 선택된다.
- [0028] 본 개시내용은 또한 본원에 기술된 핵산 서열을 포함하는 발현 벡터로 확장된다.
- [0029] 본 개시내용은 또한 본원에 기술된 핵산 서열 또는 발현 벡터를 포함하는 숙주 세포로 확장된다.
- [0030] 다른 국면에서, 본 개시내용은 (a) 본 명세서에 기술된 핵산 서열을 제공하는 단계; (b) 숙주 세포 배양물에서 핵산 서열을 발현시켜 폴리펩티드를 생산하는 단계; 및 (c) 상기 숙주 세포 배양물로부터 (b)에서 생산된 폴리펩티드를 수집하는 단계를 포함하는, MHETase 활성을 갖는 폴리펩티드의 생산 방법을 제공한다.
- [0031] 또 다른 국면에서, 모노-(2-하이드록시에틸)테레프탈레이트를 테레프탈레이트 및 에틸렌 글리콜로 전환하는데 충분한 조건 하에서, 모노-(2-하이드록시에틸)테레프탈레이트를 본원에 기술된 폴리펩티드, 조성물 또는 숙주 세포에 노출시키는 단계를 포함하는, 모노-(2-하이드록시에틸)테레프탈레이트의 가수분해 방법이 제공된다.
- [0032] 본 개시내용은 또한 플라스틱 제품을 본원에 기술된 폴리펩티드, 조성물 또는 숙주 세포에 노출시키는 단계를 포함하는, 폴리에스테르를 포함하는 플라스틱 제품의 분해 방법으로 확장된다.
- [0033] 다른 실시양태에서, 본원에 기술된 방법은
- [0034] (i) PETase가 PET의 전환을 촉매하여 모노-(2-하이드록시에틸)테레프탈레이트(MHET)를 생성하는데 충분한 조건 하에서, PET를 폴리에틸렌 테레프탈레이트 에스테라제(PETase)에 노출시키는 단계; 및
- [0035] (ii) 동시에 또는 순차적으로, 폴리펩티드가 MHET의 가수분해를 촉매하여 테레프탈레이트 및 에틸렌 글리콜을 생산하는데 충분한 조건 하에서, 단계 (a)에서 생성된 MHET를 본원에 기술된 폴리펩티드, 조성물 또는 숙주 세포에 노출시키는 단계를 포함한다.
- [0036] 본 개시내용은 또한 본원에 개시된 방법에 의해 회수된 테레프탈레이트 및/또는 에틸렌 글리콜을 포함하는 조성물로 확장된다.
- [0037] 다른 국면에서, 본원에 기술된 폴리펩티드를 발현하도록 유전적으로 변형된 숙주 세포가 제공된다.

도면의 간단한 설명

- [0038] 도 1은 야생형(WT) MHETase(서열번호 1) 및 상이한 컨센서스 설계(서열번호 2 내지 36)의 아미노산 서열을 도시한다.
- 도 2는 야생형(WT) MHETase(서열번호 37) 및 상이한 컨센서스 설계(서열번호 38-72)의 핵산 서열을 도시한다.
- 도 3은 MHET 유사체(1-나프틸 테레프탈레이트)에 대한 전세포 현탁액 내 MHETase 변이체($dA_{465}/dt(\text{분}^{-1})$)의 활성을 도시한다.
- 도 4는 SDS-PAGE 겔 전기영동 및 NTA-Atto550 염색(Sigma)에 의해 가용성 세포 용해물 내 야생형 MHETase(WT; 서열번호 1) 및 점 돌연변이를 포함하는 MHETase 변이체(MHETase 변이체 N156G+I159V를 포함함)의 발현 수준을 도시한다.

도 5는 222 nm (Y축)에서의 원편광 이색성(circular dichroism) 및 20 내지 90°C (X축) 범위의 온도에 의해 결정되는, 정제된 야생형 MHETase(WT) 및 점 돌연변이 N156G+T159V, N156G+T159V+Y197V 및 N156G+T159V+YY503W를 포함하는 MHETase 변이체의 열안정성을 도시한다.

도 6은 각 돌연변이 유발 라운드에서 테스트된 모든 MHETase 변이체에 대한 전세포 현탁액 FastBlue 검정 결과를 도시한다. 막대 높이는 각 변이체($n \geq 2$, 개별 측정값 표시)에 대해 측정된 평균 활성(dA_{465}/dt (분⁻¹))을 나타내고 오차 막대는 측정의 표준 오차 평균을 나타낸다. 강조 표시된 막대는 다음 돌연변이 유발 라운드에서 모체로 사용된 변이체를 나타낸다.

도 7은 ATTO550을 사용하여 염색하고 UV 투과조명 하에서 이미지화한 각 라운드에서 선택된 MHETase 변이체의 SDS-PAGE 겔을 도시한다. MHETase 변이체의 예상 크기(약 64kDa)가 표시된다.

도 8은 선택된 MHETase 변이체의 크기 배제 크로마토그램을 도시한다.

도 9는 본 명세서에 기술된 발색 검정을 사용하여 수득한, 선택된 MHETase 변이체에 대한 미카엘리스-멘텐(Michaelis-Menten) 플롯을 도시한다. 각 점은 각각이 6 nM MHETase 및 4 mM 패스트 블루 B 염과 함께 배양된 3회 기술 반복실험에서의 평균 초기 반응 속도를 나타낸다. 오차 막대는 표준 오차 평균을 나타낸다.

도 10은 아세트산 나트륨 pH 5.1에서 222 nm에서의 원편광 이색성에 의해 측정된 3회 반복 실험에서의 MHETase 변이체의 열안정성을 도시한다. 데이터는 표준 오차 평균에 상응하는 오차 막대를 사용하여 2상태 언폴딩 모델(선)(two-state unfolding model (lines))에 피팅하였다.

도 11은 야생형 MHETase, 라운드 5 Y252F(R5)의 활성, 및 위치 192, 156, 159, 252 및 503에서 R5의 야생형 MHETase 동일성으로의 회귀를 비교하는 HPLC 검정을 도시한다.

도 12는 MHETase R5 회귀 돌연변이에 대한 전세포 현탁액 FastBlue 검정 결과를 도시한다. 돌연변이 V159T, Y192M, F252Y 및 W503Y는 MHETase R5 (라운드 5의 MHETase Y252F)를 배경으로 제조되었다. 막대 높이는 각 변이체($n \geq 2$)에 대해 측정된 평균 활성을 나타내고 오차 막대는 표준 오차 평균을 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0039] 다르게 정의되지 않는 한, 본 명세서에 사용된 모든 기술 및 과학 용어는 본 발명이 속하는 기술분야의 통상의 기술자가 일반적으로 이해하는 것과 동일한 의미를 갖는다. 본 명세서에 기술된 것과 유사하거나 동등한 임의의 방법 및 재료가 본 발명의 실시 또는 테스트에 사용될 수 있지만, 바람직한 방법 및 물질이 기술된다. 본 발명의 목적을 위해, 다음 용어를 하기에 정의한다.
- [0040] 관사 "하나(a)" 및 "하나의(an)"는 달리 명시적으로 언급되지 않는 한, 관사의 문법적 대상 중 하나 또는 하나 이상(즉, 적어도 하나)을 지칭하기 위해 본원에 사용된다. 예를 들어, "요소"는 하나의 요소 또는 하나 이상의 요소를 의미한다.
- [0041] 본원에 사용된 용어 "대략"은 참조 수량, 수준, 값, 치수, 크기 또는 양에 대한 10%만큼(예를 들어, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2% 또는 1%) 변하는 양, 수준, 값, 치수, 크기 또는 양을 지칭한다.
- [0042] 본 명세서 전반에 걸쳐, 문맥상 달리 요구되지 않는 한, "포함하다", "포함시키다" 및 "포함하는"이라는 단어는 명시된 단계 또는 요소 또는 단계나 요소의 그룹을 포함하는 것을 의미하지만 임의의 다른 단계 또는 요소 또는 단계나 요소의 그룹을 배제하는 것은 아닌 것으로 이해될 것이다.
- [0043] 본 개시내용은 모노-(2-하이드록시에틸)테레프탈레이트 가수분해효소(MHETase)의 아미노산 서열의 특정 변형이 MHETase 활성을 유리하게 향상시킬 수 있다는 본 발명자들의 예상치 못한 발견에 적어도 부분적으로 근거를 두고 있다. 특정 변형은 또한 숙주 세포에서 발현시 변형된 MHETase의 발현을 예기치 않게 향상시켰다. 특정 변형은 놀랍게도 숙주 세포에서 발현시 변형된 MHETase의 전 세포 활성을 향상시켰다. 특정 변형은 또한 예기치 않게 효소의 열안정성을 향상시켰다. 본 발명자들은 또한 야생형 MHETase 대비 MHETase 활성 및 발현을 향상시키기 위해 야생형 MHETase의 활성 부위 외부에 있는 아미노산 잔기에 치환이 이루어질 수 있다는 것을 예기치 않게 발견했다. 따라서, 본 명세서에 개시된 국면에서, 모노-(2-하이드록시에틸) 테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드로서, (i) 서열번호 1과 적어도 70%의 서열 동일성을 가지고 (ii) MHETase의 폴리에스테르 기질과 접촉하지 않는 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하며, 폴리펩티드의 MHETase 활성은 서열번호 1의 MHETase의 MHETase 활성보다 더 큰 폴리펩티드

가 제공된다.

- [0044] 본 명세서에 개시된 다른 국면에서, 모노-(2-하이드록시에틸)테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드로서, (i) 서열번호 1과 적어도 70%의 서열 동일성을 가지고 (ii) 서열번호 1의 아미노산 위치 156 내지 396, 398 내지 410 및 425 내지 603에 상응하는 위치로 이루어진 그룹으로부터 선택된 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하는, 폴리펩티드가 제공된다.
- [0045] "적어도 70%"는 폴리펩티드가 서열번호 1과 적어도 70%, 바람직하게는 적어도 70%, 바람직하게는 적어도 75%, 바람직하게는 적어도 80%, 바람직하게는 적어도 85%, 바람직하게는 적어도 90%, 바람직하게는 적어도 92%, 바람직하게는 적어도 94%, 바람직하게는 적어도 95%, 바람직하게는 적어도 96%, 바람직하게는 적어도 97%, 바람직하게는 적어도 98%, 또는 보다 바람직하게는 99% 서열 동일성을 공유한다는 것을 의미한다. 본 명세서에 기술된 폴리펩티드는 서열번호 1의 자연 발생(야생형) MHETase의 변이체이므로, 이러한 맥락에서 "적어도 70%"는 서열번호 1의 전체 서열(잔기 1-603 또는 잔기 18-603)에 걸쳐 100% 서열 동일성을 포함하지 않는 것으로 이해되어야 한다. 일부 실시양태에서, 본원에 기술된 바와 같이, 변형된 폴리펩티드가 서열번호 1의 MHETase 대비 더 큰 MHETase 활성을 갖는 한, 폴리펩티드는 예를 들면 N- 및/또는 C-말단에서 아미노산 삽입 및/또는 결손을 포함할 수 있다.
- [0046] 다른 국면에서, 모노-(2-하이드록시에틸)테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드로서, (a) 서열번호 1과 적어도 70%의 서열 동일성을 가지고, (b) 하기로 이루어진 그룹으로부터 선택되는 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하는, 폴리펩티드가 제공된다:
- [0047] (i) 서열번호 1의 아미노산 위치 156에 상응하는 위치;
- [0048] (ii) 서열번호 1의 아미노산 위치 159에 상응하는 위치;
- [0049] (iii) 서열번호 1의 아미노산 위치 192에 상응하는 위치; 및
- [0050] (iv) 서열번호 1의 아미노산 위치 503에 상응하는 위치.
- [0051] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 156에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 11과 상이하다.
- [0052] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 156에 상응하는 위치에서의 아미노산 치환은 N156G 또는 이의 보존적 아미노산 치환이다.
- [0053] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 159에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0054] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 159에 상응하는 위치에서의 아미노산 치환은 T159V 또는 이의 보존적 아미노산 치환이다.
- [0055] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 252에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0056] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 156에 상응하는 위치에서의 아미노산 치환은 Y252F 또는 이의 보존적 아미노산 치환이다.
- [0057] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0058] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 156에 상응하는 위치에서의 아미노산 치환은 Y503W 또는 이의 보존적 아미노산 치환이다.
- [0059] 특정 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 159, 252 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다. 바람직한 실시양태에서, 아미노산 치환은 T159V, Y252F 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환이다.
- [0060] 특정 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 159, 192, 252 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다. 바람직한 실시양태에서, 아미노산 치환은 T159V, M192Y, Y252F 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환이다.

- [0061] 특정 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 159, 192 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다. 바람직한 실시양태에서, 아미노산 치환은 T159V, M192Y 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환이다.
- [0062] 특정 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 156, 159 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다. 바람직한 실시양태에서, 아미노산 치환은 N156G, T159V, 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환이다.
- [0063] 특정 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 156, 및 159 및 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다. 바람직한 실시양태에서, 아미노산 치환은 N156G, T159V, 및 Y503W, 또는 임의의 전술한 것의 보존적 아미노산 치환이다.
- [0064] 다른 국면에서, 모노-(2-하이드록시에틸)테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드로서, 폴리펩티드는 (a) 서열번호 1과 적어도 70%의 서열 동일성을 가지고, (b) 하기로 이루어진 그룹으로부터 선택되는 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하며, 서열번호 1의 폴리펩티드 대비 더 큰 전 세포 MHETase 활성을 갖는 폴리펩티드를 제공한다:
- [0065] (i) 서열번호 1의 아미노산 위치 156에 상응하는 위치;
- [0066] (ii) 서열번호 1의 아미노산 위치 159에 상응하는 위치;
- [0067] (iii) 서열번호 1의 아미노산 위치 196에 상응하는 위치;
- [0068] (iv) 서열번호 1의 아미노산 위치 197에 상응하는 위치;
- [0069] (v) 서열번호 1의 아미노산 위치 260에 상응하는 위치;
- [0070] (vi) 서열번호 1의 아미노산 위치 264에 상응하는 위치;
- [0071] (vii) 서열번호 1의 아미노산 위치 267에 상응하는 위치;
- [0072] (viii) 서열번호 1의 아미노산 위치 286에 상응하는 위치; 및
- [0073] (ix) 서열번호 1의 아미노산 위치 503에 상응하는 위치.
- [0074] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 156에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0075] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 156에 상응하는 위치에서의 아미노산 치환은 N156G 또는 이의 보존적 아미노산 치환이다.
- [0076] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 159에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0077] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 159에 상응하는 위치에서의 아미노산 치환은 T159V 또는 이의 보존적 아미노산 치환이다.
- [0078] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 196에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0079] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 196에 상응하는 위치에서의 아미노산 치환은 S196A 또는 이의 보존적 아미노산 치환이다.
- [0080] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 197에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0081] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 197에 상응하는 위치에서의 아미노산 치환은 Y197V 또는 이의 보존적 아미노산 치환이다.
- [0082] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 260에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0083] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 260에 상응하는 위치에서의 아미노산 치환은 S260A 또는 이

의 보존적 아미노산 치환이다.

- [0084] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 264에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0085] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 264에 상응하는 위치에서의 아미노산 치환은 S264L 또는 이의 보존적 아미노산 치환이다.
- [0086] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 267에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0087] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 267에 상응하는 위치에서의 아미노산 치환은 S267A 또는 이의 보존적 아미노산 치환이다.
- [0088] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 286에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0089] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 286에 상응하는 위치에서의 아미노산 치환은 S286A 또는 이의 보존적 아미노산 치환이다.
- [0090] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 503에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0091] 하나의 실시양태에서, 서열번호 1의 아미노산 위치 503에 상응하는 위치에서의 아미노산 치환은 Y503W 또는 이의 보존적 아미노산 치환이다.
- [0092] 본 개시내용은 또한 본 명세서에 기술된 바와 같이 서열번호 1의 위치에 상응하는 2개 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 등) 위치에서의 아미노산 치환의 조합을 고려한다. 하나의 실시양태에서, 폴리펩티드는 본 명세서에 기술된 바와 같이 서열번호 1의 위치에 상응하는 적어도 2개, 바람직하게는 적어도 3개, 바람직하게는 적어도 4개, 바람직하게는 적어도 5개, 바람직하게는 적어도 6개, 바람직하게는 적어도 7개, 바람직하게는 적어도 8개, 바람직하게는 적어도 9개, 또는 더 바람직하게는 적어도 10개의 위치에서의 아미노산 치환의 조합을 포함한다.
- [0093] 하나의 실시양태에서, 폴리펩티드의 아미노산 서열은 서열번호 1의 아미노산 위치 156, 159 및 197에 상응하는 위치에서의 아미노산 치환에 의해 서열번호 1과 상이하다.
- [0094] 하나의 실시양태에서, 아미노산 치환은 N156G, T159V 및 Y197V, 또는 임의의 전술한 것의 보존적 아미노산 치환이다.
- [0095] 본 개시내용은 또한 본원에 기술된 바와 같은 폴리펩티드를 포함하는 조성물로 확장된다.
- [0096] 본 개시내용은 또한 본원에 기술된 폴리펩티드를 암호화하는 핵산 서열로 확장된다.
- [0097] 본 개시내용은 또한 본원에 기술된 핵산 서열을 포함하는 발현 벡터로 확장된다.
- [0098] 본 개시내용은 또한 본원에 기술된 핵산 서열 또는 발현 벡터를 포함하는 숙주 세포로 확장된다.
- [0099] 다른 국면에서, 본 개시내용은 (a) 본 명세서에 기술된 핵산 서열을 제공하는 단계; (b) 숙주 세포 배양물에서 핵산 서열을 발현시켜 폴리펩티드를 생산하는 단계; 및 (c) 상기 숙주 세포 배양물로부터 (b)에서 생산된 폴리펩티드를 수집하는 단계를 포함하는, MHETase 활성을 갖는 폴리펩티드의 생산 방법을 제공한다.
- [0100] 또 다른 국면에서, 모노-(2-하이드록시에틸)테레프탈레이트를 테레프탈레이트 및 에틸렌 글리콜로 전환하는데 충분한 조건 하에서, 모노-(2-하이드록시에틸)테레프탈레이트를 본 명세서에 기술된 폴리펩티드, 조성물 또는 숙주 세포에 노출시키는 단계를 포함하는, 모노-(2-하이드록시에틸)테레프탈레이트의 가수분해 방법을 제공한다.
- [0101] 본 개시내용은 또한 플라스틱 제품을 본 명세서에 기술된 폴리펩티드, 조성물 또는 숙주 세포에 노출시키는 단계를 포함하는, 폴리에스테르를 포함하는 플라스틱 제품의 분해 방법으로 확장된다.
- [0102] 하나의 실시양태에서, 폴리에스테르는 폴리락트산(PLA), 폴리트리메틸렌 테레프탈레이트(PTT), 폴리부틸렌 테레프탈레이트(PBT), 폴리에틸렌 아이소소르바이트 테레프탈레이트(PEIT), 폴리에틸렌 테레프탈레이트(PET), 폴리하이드록시알카노에이트(PHA), 폴리부틸렌 숙시네이트(PBS), 폴리부틸렌 숙시네이트 아디페이트(PBSA), 폴리부

틸렌 아디페이트 테레프탈레이트(PBAT), 폴리에틸렌 푸라노에이트(PEF), 폴리카프로락톤(PCL), 폴리에틸렌아디페이트(PEA), 폴리글리콜산(PGA), 폴리(락틱-co-글리콜산)(PLGA) 및 임의의 전술한 것의 조합으로 이루어진 그룹으로부터 선택된다. 하나의 실시양태에서, 폴리에스테르는 폴리에틸렌 테레프탈레이트(PET)이다.

- [0103] 다른 실시양태에서, 본 명세서에 기술된 방법은
- [0104] (a) PETase가 PET의 전환을 촉매하여 모노-(2-하이드록시에틸)테레프탈레이트(MHET)를 생성하는데 충분한 조건 하에서 PET를 폴리에틸렌 테레프탈레이트 에스테라제(PETase)에 노출시키는 단계; 및
- [0105] (b) 동시에 또는 순차적으로, 폴리펩티드가 MHET의 가수분해를 촉매하여 테레프탈레이트 및 에틸렌 글리콜을 생산하는데 충분한 조건 하에서, 단계 (a)에서 생성된 MHET를 제1항 내지 제21항 중 어느 한 항의 폴리펩티드, 제22항의 조성물 또는 제25항의 숙주 세포에 노출시키는 단계를 포함한다.
- [0106] 하나의 실시양태에서, 상기 방법은 단계 (b)에서 생산된 테레프탈레이트 및/또는 에틸렌 글리콜을 회수하는 단계를 더 포함한다.
- [0107] 본 개시내용은 또한 본 명세서에 개시된 방법에 의해 회수한 테레프탈레이트 및/또는 에틸렌 글리콜을 포함하는 조성물로 확장된다.
- [0108] 다른 국면에서, 본 명세서에 기술된 폴리펩티드를 발현하도록 유전적으로 변형된 숙주 세포가 제공된다.
- [0109] 본 명세서에서, "펩티드", "폴리펩티드", "단백질", "효소"라는 용어는 상기 사슬을 형성하는 아미노산의 수에 관계없이 펩티드 결합으로 연결된 아미노산 사슬을 지칭하는 것으로 이해되어야 한다. 아미노산은 일반적으로 다음 명명법에 따라 한 글자 또는 세 글자 코드로 표시한다: A: 알라닌(Ala); C: 시스테인(Cys); D: 아스파르트산(Asp); E: 글루탐산(Glu); F: 페닐알라닌(Phe); G: 글리신(Gly); H: 히스티딘(His); I: 이소류신(Ile); K: 라이신(Lys); L: 류신(Leu); M: 메티오닌(Met); N: 아스파라긴(Asn); P: 프롤린(Pro); Q: 글루타민(Gln); R: 아르기닌(Arg); S: 세린(Ser); T: 트레오닌(Thr); V: 발린(Val); W: 트립토판(Trp) 및 Y: 티로신(Tyr).
- [0110] "가수분해효소"라는 용어는 더 짧은 펩티드를 생성하기 위해 펩티드 또는 단백질 내 펩티드 결합의 가수분해를 촉매하는 효소 명명법에 따라 EC 3으로 분류되는 가수분해효소 부류에 속하는 효소를 지칭한다.
- [0111] "야생형" 또는 "모(parent)"라는 용어는 폴리펩티드의 자연 발생 이소형(isoform)을 나타내기 위해 본원에서 상호교환적으로 사용되는데; 즉, 자연에 나타나는 것과 마찬가지로이다. 본 발명에서, 야생형 폴리펩티드 서열번호 1(EC 3.1.1.102; UniProt 수탁번호 A0A0K8P8E7)에 제시된 아미노산 서열을 갖는 모노-(2-하이드록시에틸)테레프탈산 가수분해효소를 지칭한다.
- [0112] Palm 등 (문헌 참조: 2019, Nat. Comms. 10:1717)이 언급한 바와 같이, 최근 발견된 폴리에틸렌 테레프탈레이트 (PET)를 특이적으로 분해하는 두 가지 박테리아 효소는 환경적으로 부담이 되는 폴리에스테르 함유 제품에 대한 유망한 해결방안을 제시한다. 첫째로, 구조적으로 잘 특성화된 α/β -가수분해효소 폴드 효소(α/β -hydrolase fold enzyme)인 이데오넬라 사카이엔시스(*Ideonella sakaiensis*) PETase는 PET를 모노-(2-하이드록시에틸)테레프탈레이트(MHET)로 전환한다. 제2 핵심 효소인 MHETase는 MHET를 PET 추출물 테레프탈레이트와 에틸렌 글리콜로 가수분해한다[(Palm 등 (2019, Nat. Comm., 10:1717), Sagong 등(2020, ACS Catal. 10:4805)), 및 Yoshida 등(2020, Science, 352(6278):1196)].
- [0113] 야생형 MHETase의 아미노산 및 핵산 서열은 당업자에게 친숙할 것이며, 이의 예로는 서열번호 1을 포함한다.
- [0114] 본 명세서에서, 용어 "돌연변이" 및 "변이체"는 서열번호 1로부터 유래된 아미노산 서열을 포함하고, 하나 이상의(예를 들어, 여러) 위치에서 서열번호 1의 폴리펩티드 대비 증가된 MHETase 활성을 갖는 변형 또는 변경(예를 들어, 치환, 삽입 및/또는 결손)을 더 포함하는 폴리펩티드를 지칭하기 위해 상호교환적으로 사용될 수 있다. 이러한 변이체는 당업계에 잘 알려진 다양한 기술에 의해 수득될 수 있으며, 이의 예로는 위치-지정 돌연변이, 무작위 돌연변이 및 합성 올리고뉴클레오티드 작제를 포함한다. 아미노산 잔기 또는 위치와 관련하여 본원에 사용된 용어 "변형", "변경", "치환" 등은 일반적으로 특정 위치의 아미노산이 야생형 또는 모 폴리펩티드의 아미노산과 비교하여 변형되었음을 의미한다.
- [0115] 적합한 치환에는 아미노산 잔기를 자연 발생 표준 20개 아미노산 잔기, 희귀 자연 발생 아미노산 잔기(예를 들면, 하이드록시프롤린, 하이드록시리신, 알로하이드록시리신, 6-N-메틸리신, N-에틸글리신, N-메틸글리신, N-에틸아스파라긴, 알로-이소류신, N-메틸이소류신, N-메틸발린, 피로글루타민, 아미노부티르산, 오르니틴, 노르류신, 노르발린), 및 종종 합성으로 만들어진 비-자연 발생 아미노산 잔기(예를 들면, 사이클로헥실-알라닌)로부

터 선택된 다른 것으로 대체하는 것을 포함할 수 있다. 바람직하게는, 치환은 아미노산 잔기를 자연 발생 표준 20개 아미노산 잔기(G, P, A, V, L, I, M, C, F, Y, W, H, K, R, Q, N, E, D, S 및 T)로부터 선택된 다른 아미노산 잔기로 대체하는 것을 포함한다. 변형 또는 변경은 본 명세서에서 다음 용어를 사용하여 확인될 수 있다: Y197V는 모 폴리펩티드 서열의 위치 197에 있는 아미노산 잔기 티로신(Y)이 발린(V)으로 치환되었음을 나타낸다. Y197V/I/M은 모 서열의 위치 197에 있는 아미노산 잔기 티로신(Y)이 다음 아미노산: 발린(V), 이소류신(I) 또는 메티오닌(M) 중 하나로 대체될 수 있음을 나타낸다. 치환은 보존적 또는 비-보존적 치환일 수 있다. 보존적 치환의 예는 당업자에게 친숙할 것이며, 이의 예로는 염기성 아미노산(아르기닌, 리신 및 히스티딘), 산성 아미노산(글루탐산 및 아스파르트산), 극성 아미노산(글루타민, 아스파라긴 및 트레오닌), 소수성 아미노산(메티오닌, 류신, 이소류신, 시스테인 및 발린), 방향족 아미노산(페닐알라닌, 트립토판 및 티로신) 및 작은 아미노산(글리신, 알라닌 및 세린)의 그룹 내의 치환을 포함한다.

[0116] 달리 명시하지 않는 한, 본 출원에 개시된 위치는 서열번호 1에 제시된 아미노산 서열을 참조하여 번호가 매겨진다. 이에 관하여, 아미노산 위치와 관련하여 사용되는 용어 "상응하는"은 해당 위치가 서열번호 1에 제시된 서열의 등과 또는 상응하는 위치와 정렬될 때 폴리펩티드 서열의 아미노산 위치를 의미하도록 의도된다.

[0117] 본 명세서에 사용된 용어 "서열 동일성" 또는 "동일성"은 2개의 폴리펩티드 서열 사이의 일치(동일한 아미노산 잔기)의 수(또는 백분율%로 표현된 분획)를 지칭한다. 바람직한 실시양태에서, 서열 동일성은 서열 갭을 최소화 하면서 오버랩(overlap) 및 동일성을 최대화하도록 정렬될 때 서열을 비교함으로써 결정된다. 서열 동일성은 두 서열의 길이에 따라 당업자에게 공지된 다수의 수학적 전역 또는 국소 정렬 알고리즘 중 임의의 것을 사용하여 결정될 수 있다. 유사한 길이의 서열은 전체 길이에 걸쳐 서열을 최적으로 정렬하는 전역 정렬 알고리즘(예를 들어, Needleman 및 Wunsch 알고리즘; Needleman 및 Wunsch, 1970)을 사용하여 정렬될 수 있는 반면, 실질적으로 다른 길이의 서열은 바람직하게는 국소 정렬 알고리즘[예를 들어, Smith and Waterman 알고리즘(Smith 및 Waterman, 1981) 또는 Altschul 알고리즘(Altschul 등., 1997; Altschul 등., 2005)]을 사용하여 정렬된다. % 아미노산 서열 동일성을 결정하기 위한 정렬은 당업자가 이용할 수 있는 임의의 수단에 의해 달성될 수 있으며, 이의 예로는 사이트(참조: <http://blast.ncbi.nlm.nih.gov/> 또는 <http://www.ebi.ac.uk/Tools/emboss/>)에서 이용 가능한 데이터와 같이 공개적으로 이용 가능한 컴퓨터 소프트웨어를 포함한다. 당업자는 비교되는 서열의 전체 길이에 걸쳐 최대 정렬을 달성하는 데 필요한 임의의 알고리즘을 포함하여 정렬을 측정하기 위한 적절한 매개변수를 쉽게 결정할 수 있다. 본원에 사용된 바와 같이, % 서열 동일성은 전형적으로 두 서열의 최적의 전역 정렬을 생성하는 쌍방식 서열 정렬[예를 들어, 니들만-분쉬(Needleman-Wunsch) 알고리즘을 사용함]를 사용하여 생성된 값을 지칭하며, 여기서 모든 검색 매개변수는 디폴트 값(예를 들어 점수매김 매트릭스(Scoring matrix) = BLOSUM62, 갭 오픈(Gap open) = 10, 갭 확장(Gap open) = 0.5, 엔드 갭 페널티(End gap penalty) = false, 엔드 갭 오픈(End gap open) = 10, 및 엔드 갭 확장(End gap extend) = 0.5)으로 설정된다.

[0118] 본 명세서에 사용된 용어 "재조합"은 일반적으로 유전공학에 의해 생산된 핵산 작제물, 벡터, 폴리펩티드 또는 세포를 지칭한다.

[0119] 본 명세서에 사용된 용어 "발현"은 전형적으로 전사, 전사후 변형, 번역, 번역후 변형 및 분비와 같은 폴리펩티드 생산에 관여하는 임의의 단계를 지칭한다.

[0120] "발현 카세트"라는 용어는 암호화 영역, 적합하게는 암호화 영역이 작동 가능하게 연결된 조절 영역을 포함하는 핵산 작제물을 의미한다.

[0121] "발현 벡터"라는 용어는 일반적으로 발현 카세트를 포함하는 DNA 또는 RNA 분자를 의미한다. 발현 벡터는 선형 또는 원형 이중 가닥 DNA 분자일 수 있다.

[0122] 본 명세서에 사용된 용어 "중합체"는 일반적으로 공유 화학 결합에 의해 연결된 다중 단량체(반복 단위)로 구성된 구조를 갖는 화학적 화합물 또는 화합물의 혼합물을 지칭한다. 본 발명의 내용에서, 용어 중합체는 단일 유형의 반복 단위(즉, 단독중합체) 또는 다른 반복 단위의 혼합물(즉, 공중합체 또는 이형중합체)로 구성된 천연 또는 합성 중합체를 포함한다.

[0123] 본 명세서에 사용된 용어 "폴리에스테르 함유 물질", "폴리에스테르 함유 제품" 등은 결정성, 반결정성 또는 완전한 무정형 형태의 하나 이상의 폴리에스테르를 포함하는 플라스틱 제품과 같은 제품을 지칭하는 것으로 이해되어야 한다. 폴리에스테르 함유 물질은 적어도 하나의 폴리에스테르를 함유하는 플라스틱 시트, 튜브, 막대, 프로필, 형상, 필름, 대규모 블록, 섬유, 옷감 등과 같은 적어도 하나의 플라스틱 물질로 만들어진 임의의 품목을 지칭할 수 있으며, 가소제, 광물 또는 유기 충전제와 같은 다른 물질이나 첨가제일 수도 있다. 하나의 실시

양태에서, 폴리에스테르 함유 물질은 적어도 하나의 폴리에스테르 함유 섬유를 포함하는 옷감 또는 직물이다. 다른 실시양태에서, 폴리에스테르 함유 물질은 플라스틱 제품 제조에 적합한 용융 또는 고체 상태의 플라스틱 화합물 또는 플라스틱 제형이다.

[0124] 적합한 폴리에스테르는 당업자에게 친숙할 것이며, 이의 예로는 폴리락트산(PLA), 폴리에틸렌 테레프탈레이트(PET), 폴리트리메틸렌 테레프탈레이트(PTT), 폴리부틸렌 테레프탈레이트(PBT), 폴리에틸렌 아이소소르바이드 테레프탈레이트(PEIT), 폴리하이드록시알카노에이트(PHA), 폴리부틸렌 숙시네이트(PBS), 폴리부틸렌 숙시네이트 아디페이트(PBSA), 폴리부틸렌 아디페이트 테레프탈레이트(PBAT), 폴리에틸렌 푸라노에이트(PEF), 폴리카프로락톤(PCL) 및 폴리(에틸렌 아디페이트)(PEA)를 포함한다. 따라서, 하나의 실시양태에서, 폴리에스테르는 폴리락트산(PLA), 폴리에틸렌 테레프탈레이트(PET), 폴리트리메틸렌 테레프탈레이트(PTT), 폴리부틸렌 테레프탈레이트(PBT), 폴리에틸렌 아이소소르바이드 테레프탈레이트(PEIT), 폴리하이드록시알카노에이트(PHA), 폴리부틸렌 숙시네이트(PBS), 폴리부틸렌 숙시네이트 아디페이트(PBSA), 폴리부틸렌 아디페이트 테레프탈레이트(PBAT), 폴리에틸렌 푸라노에이트(PEF), 폴리카프로락톤(PCL) 및 폴리(에틸렌 아디페이트)(PEA), 및 임의의 전술한 것의 조합으로 이루어진 그룹으로부터 선택된다.

[0125] 본 명세서에서 언급한 바와 같이, 본 발명자들은 서열번호 1의 모 폴리펩티드 대비 증가된 MHETase 활성을 나타내는 서열번호 1의 자연 발생 MHETase의 변이체를 개발하였다. MHETase 변이체는 또한 재조합 숙주 세포 시스템에서 강화된 발현을 나타낸다. 보다 구체적으로, 본 발명자들은 산업 공정에 사용하기 위한 우수한 특성을 갖는 신규한 MHETase를 개발하였다. 분해성 플라스틱 제품의 산업적 생산이 수행될 수 있고/또는 플라스틱 제품의 환경적 분해가 이루어질 수 있는 조건에서 가수분해효소, 특히 MHETase의 활성을 개선할 목적으로, 본 발명자들은 당해 모 가수분해효소와 비교했을 때 예기치 않게 더 높은 MHETase 활성을 나타내는 서열번호 1의 야생형 MHETase로부터 유래된 신규한 MHETase를 개발했다. 본 명세서에 개시된 MHETase 변이체는 플라스틱 제품, 특히 PET를 함유한 제품의 분해에 특히 적합하다. 또한, 본 발명자들은 놀랍게도 단백질의 구조에서 폴리에스테르 기질과 접촉하도록 의도되지 않은 아미노산 잔기가 유리하게 변형되어 MHETase 활성을 향상시킬 수 있다는 것을 발견했다. 따라서, 본 명세서에 개시된 다른 국면에서, 모노-(2-하이드록시에틸) 테레프탈레이트 가수분해효소(MHETase) 활성을 갖는 폴리펩티드로서, 폴리펩티드는 (i) 서열번호 1과 적어도 70%의 서열 동일성을 가지고 (ii) MHETase의 폴리에스테르 기질과 접촉하지 않는 하나 또는 그 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 상이한 아미노산 서열을 포함하며, 폴리펩티드의 MHETase 활성은 서열번호 1의 MHETase의 MHETase 활성보다 더 큰 폴리펩티드가 제공된다. 이와 관련하여, 용어 "접촉"은 전형적으로 서열번호 1의 MHETase의 아미노산 잔기와 이의 폴리에스테르 기질에 의해 이루어진 직접적인 접촉을 지칭한다. 폴리에스테르 기질과 접촉하는 서열번호 1의 아미노산 잔기는 당업자에게 친숙할 것이다. 이들 잔기는 또한 문헌(참조: Sagong 등 (2020, ACS Catal. 10:4805)에 기재되어 있으며, 서열번호 1의 R411, S416 및 F424를 포함한다. 하나의 실시양태에서, 폴리펩티드는 서열번호 1의 MHETase의 활성 부위 외부에 있는 하나 이상의 위치에서의 아미노산 치환에 의해 서열번호 1과 다른 아미노산 서열을 포함한다. "활성 부위"라는 용어는 전형적으로 폴리에스테르 기질 (즉, MHE T)와 접촉하여 가수분해할 수 있는 서열번호 1의 영역을 지칭한다. 서열번호 1의 MHETase의 활성 부위 외부에 있는 서열번호 1의 아미노산 위치는 당업자에게 친숙할 것이다.

[0126] 본 개시내용과 관련하여, 증가된 또는 강화된 MHETase 활성에 대한 참조는 서열번호 1의 MHETase 대비 MHET를 가수분해하는 폴리펩티드의 증가된 능력; 서열번호 1의 MHETase 대비 숙주 세포 시스템에서 증가된 재조합 발현; 서열번호 1의 MHETase 대비 증가된 전 세포 활성; 서열번호 1의 MHETase 대비 증가된 열안정성 중 하나 이상을 포함할 수 있다.

[0127] 하나의 실시양태에서, 본 명세서에 개시된 폴리펩티드는 서열번호 1의 MHETase 대비, 숙주 세포 시스템에서 MHETase의 증가된 재조합 발현을 포함한다.

[0128] 다른 실시양태에서, 본 명세서에 개시된 폴리펩티드는 서열번호 1의 MHETase 대비 증가된 열안정성을 포함한다.

[0129] 다른 실시양태에서, 본 명세서에 개시된 폴리펩티드는 서열번호 1의 MHETase 대비 증가된 전 세포 활성을 포함한다. 본 명세서에 사용된 용어 전 세포 활성은 전형적으로 숙주 세포 시스템에서 발현될 때 MHETase가 MHET를 가수분해하는 능력을 지칭한다.

[0130] 추가의 실시양태에서, 본 명세서에 개시된 폴리펩티드는 서열번호 1의 MHETase 대비 증가된 열안정성 및 증가된 전 세포 활성을 포함한다.

[0131] 추가의 실시양태에서, 본 명세서에 개시된 폴리펩티드는 서열번호 1의 MHETase 대비, 숙주 세포 시스템에서

MHETase의 증가된 재조합 발현, 증가된 열안정성 및 증가된 전 세포 활성을 포함한다.

- [0132] 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드의 MHETase 활성은 서열번호 1의 MHETase의 MHETase 활성과 유사하다. 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드의 MHETase 활성은 서열번호 1의 MHETase 대비 적어도 약 5%, 바람직하게는 적어도 약 10%, 바람직하게는 적어도 약 20%, 바람직하게는 적어도 약 30%, 바람직하게는 적어도 약 40%, 바람직하게는 적어도 약 50%, 바람직하게는 적어도 약 100%, 바람직하게는 적어도 약 200%, 바람직하게는 적어도 약 300%, 바람직하게는 적어도 약 400%, 바람직하게는 적어도 약 500%, 바람직하게는 적어도 약 600%, 바람직하게는 적어도 약 700%, 바람직하게는 적어도 약 800%, 바람직하게는 적어도 약 900%, 또는 보다 바람직하게는 적어도 약 1,000% 이상만큼 증가된다. 폴리펩티드의 MHETase 활성을 결정하거나 측정하는 적합한 방법은 당업자에게 친숙할 것이며, 이의 예는 본 명세서에 기술되어 있다. 다른 예는 Palm 등 (문헌: 2019, Nat. Comm., 10:1717), Sagong 등 (2020, ACS Catal. 10:4805) 및 Yoshida 등 (2020, Science, 352(6278):1196)에 기재되어 있으며, 이의 내용은 전문이 참조로 본 명세서에 포함된다. 하나의 실시양태에서, MHETase 활성은, 1-나프틸 테레프탈레이트(INT)를 기질로 사용하여 발색 검정에 의해 결정된 경우 서열번호 1의 MHETase 대비 적어도 약 5%, 바람직하게는 적어도 약 10%, 바람직하게는 적어도 약 20%, 바람직하게는 적어도 약 30%, 바람직하게는 적어도 약 40%, 바람직하게는 적어도 약 50%, 바람직하게는 적어도 약 100%, 바람직하게는 적어도 약 200%, 바람직하게는 적어도 약 300%, 바람직하게는 적어도 약 400%, 바람직하게는 적어도 약 500%, 바람직하게는 적어도 약 600%, 바람직하게는 적어도 약 700%, 바람직하게는 적어도 약 800%, 바람직하게는 적어도 약 900%, 또는 보다 바람직하게는 적어도 약 1,000% 이상만큼 증가된다.
- [0133] MHETase 활성은 비교 대상(예를 들어, 서열번호 1의 MHETase)의 MHETase 활성에 대한 절대값 또는 상대적인 값으로 할당될 수 있다. 하나의 실시양태에서, MHETase 활성은 적합한 온도, pH 및 완충액의 조건 하에서 시간당 및 효소 mg당 방출되는 단량체 및/또는 올리고머의 속도 (예를 들어, mg 단위)로서 측정된다.
- [0134] MHETase 활성은 정제된 효소를 사용하여 측정하거나 분석할 수 있다. 대안적으로, MHETase 활성은 숙주 세포 시스템에서 재조합적으로 발현될 때 효소의 활성(본 명세서에서는 세포 촉매 활성 또는 전 세포 활성으로도 지칭됨)의 함수로서 측정될 수 있다.
- [0135] 유리하게는, 본 명세서에 기술된 폴리펩티드는, 서열번호 1의 MHETase 대비 숙주 세포에서 적어도 약 5%, 바람직하게는 적어도 약 10%, 바람직하게는 적어도 약 20%, 바람직하게는 적어도 약 30%, 바람직하게는 적어도 약 40%, 바람직하게는 적어도 약 50%, 바람직하게는 적어도 약 100%, 바람직하게는 적어도 약 200%, 바람직하게는 적어도 약 300%, 바람직하게는 적어도 약 400%, 바람직하게는 적어도 약 500%, 바람직하게는 적어도 약 600%, 바람직하게는 적어도 약 700%, 바람직하게는 적어도 약 800%, 바람직하게는 적어도 약 900%, 보다 바람직하게는 적어도 약 1,000% 증가되거나 또는 강화된 재조합 발현을 나타낸다.
- [0136] 유리하게는, 본 명세서에 기술된 폴리펩티드는 적어도 약 10°C 내지 약 60°C, 바람직하게는 약 20°C 내지 약 680°C, 바람직하게는 약 30°C 내지 약 60°C, 보다 바람직하게는 약 40°C 내지 약 60°C, 더욱 더 바람직하게는 약 40°C 내지 약 50°C, 훨씬 더 바람직하게는 약 45°C의 온도 범위에서 MHETase 활성을 나타낸다. 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는 약 10°C 내지 약 60°C, 바람직하게는 약 20°C 내지 약 60°C, 바람직하게는 약 30°C 내지 약 60°C, 보다 바람직하게는 약 40°C 내지 약 60°C, 더욱 더 바람직하게는 약 40°C 내지 약 50°C, 보다 더 바람직하게는 약 45°C의 온도에서 MHETase 활성을 나타낸다. 하나의 실시양태에서, MHETase 활성은 약 40°C 내지 약 60°C, 바람직하게는 약 40°C 내지 약 50°C, 더욱 더 바람직하게는 약 45°C의 온도에서 측정가능하다. 다른 특정 실시 양태에서, 폴리에스테르 분해 활성은 자연 환경의 평균 온도(주위 온도)에 상응하는 약 10°C 내지 약 30°C, 바람직하게는 약 15°C 내지 약 28°C의 온도에서 여전히 측정 가능하다.
- [0137] 하나의 실시양태에서, 폴리펩티드는, 약 10°C 내지 약 60°C, 바람직하게는 약 20°C 내지 약 60°C, 바람직하게는 약 30°C 내지 약 60°C, 보다 바람직하게는 약 40°C 내지 약 60°C, 더욱 더 바람직하게는 약 40°C 내지 약 50°C, 또는 보다 더 바람직하게는 약 45°C의 온도에서, 동일한 온도에서의 서열번호 1의 MHETase 활성 대비, 적어도 약 5%, 바람직하게는 적어도 약 10%, 바람직하게는 적어도 약 20%, 바람직하게는 적어도 약 30%, 바람직하게는 적어도 약 40%, 바람직하게는 적어도 약 50%, 바람직하게는 적어도 약 100%, 바람직하게는 적어도 약 200%, 바람직하게는 적어도 약 300%, 바람직하게는 적어도 약 400%, 바람직하게는 적어도 약 500%, 바람직하게는 적어도 약 600%, 바람직하게는 적어도 약 700%, 바람직하게는 적어도 약 800%, 바람직하게는 적어도 약 900%, 또는 보다 바람직하게는 적어도 약 1,000% 이상의 MHETase 활성을 포함한다.
- [0138] 다른 특정 실시양태에서, 본 명세서에 기술된 폴리펩티드는 서열번호 1의 폴리펩티드와 비교했을 때 약 10°C 내지 약 60°C, 바람직하게는 약 20°C 내지 약 60°C, 바람직하게는 약 30°C 내지 약 60°C, 바람직하게는 약 40°C

내지 약 60℃, 바람직하게는 약 40℃ 내지 약 50℃, 또는 보다 바람직하게는 약 45℃의 온도에서 증가된 MHETase 활성을 갖는다. 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는, 약 20℃ 내지 약 6045℃에서, 동일한 온도에서의 서열번호 1의 MHETase 대비, 적어도 약 5%, 바람직하게는 적어도 약 10%, 바람직하게는 적어도 약 20%, 바람직하게는 적어도 약 30%, 바람직하게는 적어도 약 40%, 바람직하게는 적어도 약 50%, 바람직하게는 적어도 약 100%, 바람직하게는 적어도 약 200%, 바람직하게는 적어도 약 300%, 바람직하게는 적어도 약 400%, 바람직하게는 적어도 약 500%, 바람직하게는 적어도 약 600%, 바람직하게는 적어도 약 700%, 바람직하게는 적어도 약 800%, 바람직하게는 적어도 약 900%, 보다 바람직하게는 적어도 약 1,000% 이상의 MHETase 활성을 갖는다.

[0139] 다른 실시양태에서, 본 명세서에 기술된 폴리펩티드는 서열번호 1의 폴리펩티드 대비 약 10℃ 내지 약 30℃, 바람직하게는 약 15℃ 내지 약 30℃, 보다 더 바람직하게는 약 20℃ 내지 약 30℃, 더욱 더 바람직하게는 약 28℃의 온도에서 증가된 MHETase 활성을 갖는다. 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는, 약 10℃ 내지 약 30℃의 온도에서, 동일한 온도에서의 서열번호 1의 MHETase 대비, 적어도 약 5%, 바람직하게는 적어도 약 10%, 바람직하게는 적어도 약 20%, 바람직하게는 적어도 약 30%, 바람직하게는 적어도 약 40%, 바람직하게는 적어도 약 50%, 바람직하게는 적어도 약 100%, 바람직하게는 적어도 약 200%, 바람직하게는 적어도 약 300%, 바람직하게는 적어도 약 400%, 바람직하게는 적어도 약 500%, 바람직하게는 적어도 약 600%, 바람직하게는 적어도 약 700%, 바람직하게는 적어도 약 800%, 바람직하게는 적어도 약 900%, 보다 바람직하게는 적어도 약 1,000% 이상의 MHETase 활성을 갖는다.

[0140] 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는 적어도 pH 5 내지 11의 범위, 바람직하게는 pH 6 내지 10의 범위, 더 바람직하게는 pH 6.5 내지 9의 범위, 훨씬 더 바람직하게는 pH 7 내지 8의 범위에서 측정가능한 MHETase 활성을 나타낸다.

[0141] 유리하게는, 본 명세서에 기술된 폴리펩티드의 열안정성은 서열번호 1의 폴리펩티드에 비해 크게 손상되지 않는다. 일부 실시양태에서, 본 명세서에 기술된 폴리펩티드의 열안정성은 서열번호 1의 폴리펩티드의 열안정성과 비교할 때 개선된다. 본 명세서에 사용된 용어 "향상된 열안정성" 또는 "증가된 열안정성"은 더 높은 온도, 보다 구체적으로 40℃ 내지 60℃ 사이의 온도에서 서열번호 1의 폴리펩티드에 비해 효소의 화학적 및/또는 물리적 구조의 변화에 저항하는 효소의 증가된 능력을 나타낸다. 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는 서열번호 1의 폴리펩티드 대비 40℃ 내지 60℃의 온도에서 증가된 반감기를 갖는다. 본 명세서에 기술된 폴리펩티드는 서열번호 1의 폴리펩티드와 비교하여 더 높거나 동등한 용융 온도(T_m)를 나타낼 수 있다. 일부 실시양태에서, 본 명세서에 기술된 폴리펩티드는 서열번호 1의 폴리펩티드 대비 40℃ 내지 68℃의 온도에서 향상된 열안정성을 나타낸다.

[0142] 폴리펩티드의 열안정성은 당업자에게 공지된 임의의 적합한 수단에 의해 평가될 수 있다. 예를 들어, 열안정성은 다양한 온도에서 배양한 후 폴리펩티드의 남아있는 MHETase 활성을 측정하여 평가할 수 있다. 다양한 온도에서 수 회의 MHETase-매개 가수분해를 수행하는 능력도 평가할 수 있다. 시차주사형광측정법(DSF)을 사용하여 폴리펩티드의 열안정성을 평가할 수도 있다. 원편광 이색성은 용융 온도(T_m)를 포함하여 본 명세서에 기술된 폴리펩티드의 열안정성을 측정하는 데에도 사용될 수 있다. "용융 온도(T_m)"라는 용어는 주어진 단백질이 상기 단백질의 50%가 변성되는 온도에 해당함을 의미하는 것으로 이해된다.

[0143] 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는 약 45℃ 내지 약 68℃, 바람직하게는 약 50℃ 내지 약 65℃, 바람직하게는 약 52℃ 내지 약 63℃의 용융 온도(T_m)를 나타낸다. 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는 서열번호 1의 폴리펩티드에 의해 나타나는 용융 온도(T_m)보다 낮은 용융 온도(T_m)를 나타낸다. 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는 서열번호 1의 폴리펩티드에 의해 나타나는 용융 온도(T_m)보다 높은 용융 온도(T_m)를 나타낸다. 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는 약 52℃ 내지 약 64℃, 바람직하게는 약 55℃ 내지 약 63℃, 및 보다 바람직하게는 약 63℃의 용융 온도(T_m)를 나타낸다.

[0144] 본 개시내용은 또한 본원에 기술된 MHETase 폴리펩티드를 암호화하는 핵산 서열을 포함하는 폴리뉴클레오티드로 확장된다. 본 명세서에 개시된 국면에서, 본 명세서에 기술된 폴리펩티드를 암호화하는 핵산 서열을 포함하는 폴리뉴클레오티드가 제공된다. 하나의 실시양태에서, 핵산 서열은 서열번호 38 내지 72 및 79 내지 84로 이루어진 그룹으로부터 선택된다. 다른 실시양태에서, 핵산 서열은 서열번호 79 내지 84로 이루어진 그룹으로부터 선택된다. 본 명세서에 사용된 용어 "핵산", "핵 서열", "폴리뉴클레오티드", "올리고뉴클레오티드" 및 "뉴클레오티드 서열"은 상호교환적으로 사용되며 데옥시리보뉴클레오티드 및/또는 리보뉴클레오티드의 서열을 지칭한다.

핵산은 DNA(cDNA 또는 gDNA), RNA 또는 이 둘의 혼합물일 수 있다. 이는 단일 가닥 형태일 수도 있고, 이중 형태 또는 이 둘의 혼합 형태일 수도 있다. 이는 재조합, 인공 및/또는 합성 기원일 수 있으며, 예를 들어 변형된 결합, 변형된 퓨린 또는 피리미딘 염기, 또는 변형된 당을 포함하는 변형된 뉴클레오티드를 포함할 수 있다. 본 발명의 핵산은 분리 또는 정제된 형태일 수 있으며, 당업계에 공지된 기술, 예를 들어 cDNA 라이브러리의 클로닝 및 발현, 증폭, 효소 합성 또는 재조합 기술에 의해 제조, 분리 및/또는 조작될 수 있다. 핵산은 예를 들어 문헌[참조: Belousov (1997) Nucleic Acids Res. 25:3440-3444]에 기술된 바와 같이 잘 알려진 화학적 합성 기술에 의해 시험관 내에서 합성될 수 있다.

[0145] 본 명세서에 개시된 핵산 서열은 적합하게는 코돈 최적화될 수 있다. 코돈 최적화를 위한 적합한 방법은 당업자에게 친숙할 것이며, 그 예시는 참고 매뉴얼 [참고: Sambrook 외 2001]에 기재되어 있다.

[0146] 본 명세서에 기술된 핵산 서열은 본 명세서에 기술된 폴리펩티드의 아미노산 서열로부터 적절하게 추론될 수 있으며, 코돈 사용은 핵산이 전사될 숙주 세포에 따라 조정될 수 있다.

[0147] 일부 실시양태에서, 본 명세서에 기술된 핵산 서열은 조절 영역, 즉 선택된 숙주 세포 또는 시스템 내 폴리펩티드의 발현을 유발하거나 조절하는 데 사용될 수 있는 프로모터, 인핸서, 사일런서, 종결자, 신호 펩티드 등과 같은 추가의 뉴클레오티드 서열을 적절하게 포함할 수 있다. 대안적으로, 또는 추가로, 본 명세서에 기술된 핵산 서열은 폴리펩티드 발현 및/또는 용해도를 조절하도록 사용될 수 있는 말토스 결합 단백질(MBP) 또는 글루타티온 S 트랜스퍼라제(GST)와 같은 용해 단백질질을 암호화하는 추가 뉴클레오티드 서열을 더 포함할 수 있다.

[0148] 본 명세서에서 언급된 바와 같이, 본 개시내용은 또한 적합한 숙주 세포에서 핵산 서열의 발현을 지시하는 하나 이상의 제어 서열에 임의로 작동가능하게 연결된 본 명세서에 기술된 핵산 서열을 포함하는 발현 벡터 및 발현 카세트로 확장된다. 전형적으로, 발현 벡터 또는 카세트는 전사 프로모터 및/또는 전사 종결자와 같은 제어 서열에 작동가능하게 연결된 본 명세서에 기술된 핵산 서열을 포함한다. 제어 서열은 본 명세서에 기술된 폴리펩티드를 암호화하는 핵산의 발현을 위해 숙주 세포 또는 시험관내 발현 시스템에 의해 인식되는 프로모터를 포함할 수 있다. 프로모터는 전형적으로 폴리펩티드의 발현을 매개하는 전사 제어 서열을 포함할 것이다. 프로모터는 돌연변이, 절단형 및 하이브리드 프로모터를 포함하여 숙주 세포에서 전사 활성을 나타내는 임의의 폴리뉴클레오티드일 수 있으며, 숙주 세포에 대해 상동성 또는 이종성인 세포외 또는 세포내 폴리펩티드를 암호화하는 유전자로부터 적합하게 수득할 수 있다. 제어 서열은 또한 전사를 종결시키기 위해 숙주 세포에 의해 인식되는 전사 종결자일 수 있다. 종결자는 전형적으로 폴리펩티드를 암호화하는 핵산의 3'-말단에 작동가능하게 연결된다. 숙주 세포에서 기능하는 임의의 종결자가 이와 관련하여 사용될 수 있다. 전형적으로, 발현 벡터 또는 카세트는 전사 프로모터 및 전사 종결자에 작동가능하게 연결된 본 명세서에 기술된 핵산 서열을 포함한다.

[0149] "벡터"라는 용어는 일반적으로 재조합 유전 물질을 숙주 세포로 전달하기 위한 운반체로 사용되는 DNA 분자를 지칭한다. 적합한 벡터는 플라스미드, 박테리오파지, 바이러스, 포스미드, 코스미드 및 인공 염색체를 포함한다. 벡터는 일반적으로 삽입물(이중 핵산 서열, 이식유전자) 및 벡터의 "골격(backbone)" 역할을 하는 더 큰 서열을 포함하는 DNA 서열이다. 유전정보를 숙주에게 전달하는 벡터의 목적은 일반적으로 표적 세포에서 삽입물을 분리, 증식 또는 발현하는 것이다. 발현 벡터(발현 작제물로도 지칭됨)는 표적 세포에서 이중 서열의 발현을 위해 특이적으로 조정되고, 일반적으로 폴리펩티드를 암호화하는 이중 서열의 발현을 유도하는 프로모터 서열을 갖는다.

[0150] 일반적으로, 발현 벡터에 사용되는 조절 요소는 전사 프로모터, 리보솜 결합 부위, 종결자 및 선택적으로 존재하는 오퍼레이터를 포함한다. 발현 벡터는 숙주 세포에서의 자율 복제를 위한 복제 기점, 선택 가능한 마커, 제한된 수의 유용한 제한 효소 부위 및 높은 복제수에 대한 가능성을 추가로 포함할 수 있다. 적합한 발현 벡터는 당업자에게 친숙할 것이며, 그 예시에는 클로닝 벡터, 변형된 클로닝 벡터, 플라스미드 및 바이러스가 포함된다. 다양한 숙주에서 적절한 수준의 폴리펩티드 발현을 제공할 수 있는 발현 벡터도 또한 당업계에 잘 알려져 있다. 벡터의 선택은 일반적으로 벡터가 도입될 숙주 세포와 벡터의 적합성에 좌우될 것이다. 하나의 실시양태에서, 벡터는 박테리아 발현 벡터 pET-28a(+)(서열번호 85)이다.

[0151] 본 개시내용은 또한 본원에 기술된 핵산 서열을 포함하는 숙주 세포로 확장된다. 숙주 세포는 일시적 또는 안정한 방식으로 형질전환, 형질감염 또는 형질도입될 수 있다. 핵산, 발현 카세트 또는 벡터는 핵산, 카세트 또는 벡터가 염색체 통합체 또는 자가 복제 염색체의 벡터로서 유지되도록 숙주 세포 내로 도입된다. 용어 "숙주 세포"는 복제 동안 발생하는 돌연변이로 인해 모 숙주 세포와 동일하지 않은 모 숙주 세포의 임의의 자손을 포함한다. 숙주 세포는 본 발명의 변이체, 예를 들어 원핵생물 또는 진핵생물의 생산에 유용한 임의의 세포일 수 있다. 원핵 숙주 세포는 임의의 그람 양성 또는 그람 음성 박테리아일 수 있다. 숙주 세포는 또한 효모, 진균, 포

유동물, 곤충 또는 식물 세포와 같은 진핵 세포일 수 있다. 특정 실시양태에서, 숙주 세포는 대장균 (*Escherichia coli*), 슈도모나스(*Pseudomonas*), 바실러스(*Bacillus*), 스트렙토마이세스(*Streptomyces*), 트리코테르마(*Trichoderma*), 아스퍼질러스(*Aspergillus*), 사카로마이세스(*Saccharomyces*), 피치아(*Pichia*), 테르무스(*Thermus*) 또는 야로위아(*Yarrowia*)의 그룹으로부터 선택된다.

- [0152] 본 발명에 따른 핵산, 발현 카세트 또는 발현 벡터는 당업자에게 공지된 임의의 적합한 방법에 의해 숙주 세포 내로 도입될 수 있으며, 그 예시에는 전기천공, 접합, 형질도입, 감응성 세포 형질전환, 원형질체 형질전환, 원형질체 융합, 생물학적 "유전자총" 형질전환, PEG 매개 형질전환, 지질 보조 형질전환 또는 형질감염, 화학적 매개 형질감염, 아세트산리튬 매개 형질전환 및 리포솜 매개 형질전환을 포함한다.
- [0153] 하나의 실시양태에서, 숙주 세포는 유전적으로 변형된 숙주 세포 또는 미생물이다. 이와 관련하여, 숙주 세포 또는 미생물은 발현되는 폴리펩티드의 발현 및/또는 활성을 강화시키기 위해 유전적으로 변형될 수 있다. 예를 들어, 본 명세서에 기술된 폴리펩티드는 해당 균주의 MHEase 활성을 개선 및/또는 증가시키기 위해 MHEase 활성을 가질 수 있는 것으로 이미 알려진 진균 또는 박테리아의 야생형 균주를 보완하는 데 사용될 수 있다.
- [0154] 본 개시내용은 또한 (a) 본 명세서에 기술된 바와 같은 핵산 서열을 제공하는 단계; (b) 숙주 세포 배양물에서 핵산 서열을 발현시켜 폴리펩티드를 생산하는 단계; 및 (c) 상기 숙주 세포 배양물로부터 (b)에서 생산된 폴리펩티드를 회수하는 단계를 포함하는, MHEase 활성을 갖는 폴리펩티드의 생산 방법으로 확장된다.
- [0155] 본 발명의 개시내용은 또한 (a) 본 발명의 핵산, 카세트 또는 벡터를 시험관내 발현 시스템과 접촉시키는 단계; (b) 생산된 폴리펩티드를 회수하는 단계를 포함하는, 본 명세서에 기술된 폴리펩티드를 생산하는 시험관내 방법으로 확장된다. 시험관내 발현 시스템은 당업자에게 잘 알려져 있으며 상업적으로 이용 가능하다.
- [0156] 적합한 숙주 세포는 당업자에게 친숙할 것이며, 이의 예로는 재조합 바실러스, 재조합 대장균, 재조합 슈도모나스, 재조합 아스퍼질러스, 재조합 트리코테르마, 재조합 스트렙토마이세스, 재조합 사카로마이세스, 재조합 피치아, 재조합 테르무스 또는 재조합 야로위아를 포함한다. 하나의 실시양태에서, 숙주 세포는 대장균이다.
- [0157] 숙주 세포는 당업자에게 공지된 방법을 사용하여 폴리펩티드 생산에 적합한 영양 배지에서 배양될 수 있다. 적합한 예로는 적절한 배지에서 효소가 발현 및/또는 분리될 수 있는 조건 하에서 수행되는 실험실 또는 산업용 발효기에서 진탕 플라스크 배양, 또는 소규모 또는 대규모 발효(연속식, 회분식, 유가식 또는 고체 상태 발효를 포함함)에 의한 숙주 세포 배양을 포함한다. 배양은 상업적 공급자로부터 또는 공개된 조성(예를 들면, ATCC: American Type Culture Collection의 카탈로그)에 따라 제조된 적절한 영양 배지 또는 세포 성장에 적합한 기타 배양 배지에서 일반적으로 수행된다. 폴리펩티드가 영양 배지로 발현 및/또는 분리되는 경우, 폴리펩티드는 세포/상등액 혼합물의 형태로 또는 조세포 용해물의 형태로 사용될 수 있다. 대안적으로, 폴리펩티드는 배양 상등액으로부터 직접 회수될 수 있다. 반대로, 폴리펩티드는 세포 용해물로부터 또는 숙주 세포막을 투과화한 후에 회수될 수 있다. 폴리펩티드는 당업자에게 공지된 임의의 적합한 방법, 예를 들면, 수집, 원심분리, 여과, 추출, 분무 건조, 증발 또는 침전을 포함하는 방법을 사용하여 회수될 수 있다. 임의로, 폴리펩티드는 열 쇼크(thermal shock), 크로마토그래피(예를 들면, 이온 교환, 친화성, 소수성, 크로마토포커싱 및 크기 배제), 전기영동 절차(예를 들면, 예비 등전점 포커싱), 차등 용해도(예를 들면, 황산암모늄 침전), SDS-PAGE, 또는 실질적으로 순수한 폴리펩티드를 얻기 위한 추출을 포함하지만, 이에 제한되지 않는 당업계에 공지된 다양한 절차에 의해 부분적으로 또는 전체적으로 정제될 수 있다.
- [0158] 폴리펩티드는 폴리에스테르를 함유한 플라스틱 제품과 같은 폴리에스테르 함유 물질의 분해 및/또는 재활용과 관련된 효소 반응을 촉매하기 위해 단독으로 또는 추가 효소(예를 들면, PETase)와 조합하여 정제된 형태로 사용될 수 있다. 본 명세서에 기술된 폴리펩티드는 가용성 형태이거나 고체상일 수 있다. 특히, 폴리펩티드는 세포막 또는 지질 소포에 결합될 수 있거나, 또는 예를 들어 비드, 컬럼, 플레이트 등의 형태로 유리, 플라스틱, 중합체, 필터, 멤브레인과 같은 합성 지지체에 결합될 수 있다.
- [0159] 본 개시내용은 또한 본 명세서에 기술된 폴리펩티드, 핵산 또는 숙주 세포를 포함하는 조성물로 확장된다.
- [0160] 상기 조성물은 액체 또는 건조형, 예를 들어 분말 형태일 수 있다. 일부 실시양태에서, 조성물은 동결건조물이다. 예를 들어, 조성물은 폴리펩티드, 핵산 및/또는 숙주 세포 및 임의의 부형제 및/또는 시약 등을 포함할 수 있다. 적합한 부형제로는 생화학에서 일반적으로 사용되는 완충제, pH 조절제, 벤조산나트륨, 소르빈산나트륨 또는 아스코르브산나트륨과 같은 방부제, 보존제, 보호제 또는 안정제, 예를 들면 전분, 텍스트린, 아라비아검, 염류, 당류, 예를 들면 소르비톨, 트레할로스 또는 락토스, 글리세롤, 폴리에틸렌글리콜, 폴리에틸렌 글리콜, 폴리프로필렌 글리콜, 프로필렌 글리콜, 칼슘과 같은 2가 이온, EDTA와 같은 격리제, 환원제 [(예를 들어, 베타

-메르캅토에탄올, 디티오프레이트, 아스코르브산, 트리스(2-카르복시에틸)포스핀)], 아미노산, 용매 또는 수용액과 같은 담체 등을 포함한다.

[0161] 하나의 실시양태에서, 조성물은 본 명세서에 기술된 폴리펩티드를 포함한다(폴리펩티드는 분리된 형태 또는 적어도 부분적으로 정제된 형태로 조성물에 존재할 수 있다). 하나의 실시양태에서, 조성물은 조성물 전체 중량의 약 0.1% 내지 약 99.9%, 바람직하게는 약 0.1% 내지 약 50%, 바람직하게는 약 0.1% 내지 약 30%, 바람직하게는 약 0.1% 내지 약 5%의 양으로 본 명세서에 기술된 폴리펩티드를 포함한다. 바람직한 실시양태에서, 조성물은 본 명세서에 기술된 폴리펩티드를 조성물의 총 중량의 약 0.1 내지 약 5 중량%의 양으로 포함한다. 다른 실시양태에서, 조성물은 본 명세서에 기술된 폴리펩티드를 조성물의 총 중량의 약 0.1 내지 약 0.2 중량%의 양으로 포함한다. 조성물 내 폴리펩티드의 양은 예를 들어 분해(가수분해)될 물질을 함유하는 폴리에스테르의 성질 및/또는 양 및/또는 조성물 내 임의의 추가 효소/폴리펩티드의 존재 또는 부재에 따라 당업자에 의해 적절하게 조절될 수 있다.

[0162] 본 명세서에 기술된 조성물은 MHETase에 제한되지 않고 효소 활성을 나타내는 추가적인 폴리펩티드(들)를 더 포함할 수 있다.

[0163] 하나의 실시양태에서, 본 명세서에 기술된 폴리펩티드는 폴리펩티드를 분해로부터 적절하게 안정화하거나 보호할 수 있는 부형제와 같은 하나 이상의 부형제와 함께 수성 매질에 용해된다. 예를 들어, 본 명세서에 기술된 폴리펩티드는 물에 용해된 다음, 글리세롤, 소르비톨, 텍스트린, 전분, 프로판디올과 같은 글리콜, 염 등과 같은 부형제와 혼합될 수 있다. 이후, 생성된 혼합물을 건조하여 분말을 수득할 수 있다. 이러한 혼합물을 건조하는 방법은 당업자에게 잘 알려져 있으며, 이에 제한되지는 않으나, 동결건조, 냉동-건조, 분무건조, 초임계 건조, 하향 기류 증발, 박층 증발, 원심 증발, 컨베이어 건조, 유동층 건조, 드립 건조 또는 이들의 조합을 포함한다.

[0164] 하나의 실시양태에서, 조성물은 본 명세서에 기술된 폴리펩티드를 발현하는 적어도 하나의 숙주 세포 또는 이의 추출물을 포함한다. "세포 추출물"은 세포 상등액, 세포 잔해, 세포벽, DNA 추출물, 효소 또는 효소 제제 또는 본질적으로 살아있는 세포가 없는 화학적, 물리적 및/또는 효소 처리에 의해 세포에서 유래된 모든 제제와 같은 세포로부터 수득한 모든 분획을 의미한다. 바람직한 추출물은 효소-활성 추출물이다. 조성물은 본 명세서에 기술된 폴리펩티드를 함유하는 하나 또는 여러 개의 숙주 세포 또는 이의 추출물, 및 선택적으로 하나 또는 여러 개의 추가 세포를 포함할 수 있다.

[0165] 본 명세서에서 언급된 바와 같이, 본 발명자들은 본 명세서에 기술된 폴리펩티드(MHETase 변이체)가 서열번호 1의 야생형 MHETase 대비 더 큰 MHETase 활성을 갖는다는 것을 놀랍게도 발견했다. 따라서, 모노-(2-하이드록시에틸)테레프탈레이트가 테레프탈레이트 및 에틸렌 글리콜로 전환하는데 충분한 조건 하에서 모노-(2-하이드록시에틸)테레프탈레이트를 본 명세서에 기술된 폴리펩티드, 조성물 또는 숙주 세포에 노출시키는 단계를 포함하는, 모노-(2-하이드록시에틸)테레프탈레이트의 가수분해 방법이 본 명세서에 개시되어 있다. 본 개시내용은 또한 플라스틱 제품을 본 명세서에 기술된 폴리펩티드, 조성물 또는 숙주 세포에 노출시키는 단계를 포함하는, 폴리에스테르를 포함하는 플라스틱 제품의 분해 방법으로 확장된다.

[0166] 본 개시내용은 호기성 또는 혐기성 조건에서 폴리에스테르를 분해하고/하거나 폴리에스테르로 제조되거나 폴리에스테르를 함유하는 플라스틱 제품으로서 폴리에스테르 함유 물질을 재활용하고/하거나 폴리에스테르를 함유하는 생분해성 플라스틱 제품을 생산하기 위해 본 명세서에 기술된 폴리펩티드, 조성물 또는 숙주 세포의 용도로 확장된다. 이들 방법 및 사용 방법은 PET를 포함하는 플라스틱 제품을 분해하는데 특히 유용하다.

[0167] 유리하게는, 폴리에스테르 함유 물질의 폴리에스테르(들)는 단량체 및/또는 올리고머까지 중합체된다. 하나의 실시양태에서, 적어도 하나의 폴리에스테르는 분해되어 재중합 가능한 단량체 및/또는 올리고머를 생성하며, 이는 추가 사용을 위해 유리하게 회복되거나 회수된다.

[0168] 하나의 실시양태에서, 폴리에스테르 함유 물질의 폴리에스테르(들)는 완전히 분해된다.

[0169] 본 명세서에 언급된 바와 같이, 플라스틱 제품은 폴리락트산(PLA), 폴리트리메틸렌 테레프탈레이트(PTT), 폴리부틸렌 테레프탈레이트(PBT), 폴리에틸렌 아이소소르바이드 테레프탈레이트(PEIT), 폴리에틸렌 테레프탈레이트(PET), 폴리하이드록시알카노에이트(PHA), 폴리부틸렌 숙시네이트(PBS), 폴리부틸렌 숙시네이트 아디페이트(PBSA), 폴리부틸렌 아디페이트 테레프탈레이트(PBAT), 폴리에틸렌 푸라노에이트(PEF), 폴리카프로락톤(PCL), 폴리(에틸렌 아디페이트)(PEA) 및 임의의 전술한 것의 조합으로 이루어진 그룹으로부터 선택되는 적어도 하나의 폴리에스테르를 포함할 수 있다.

- [0170] 폴리에스테르 함유 물질을 분해하는 데 필요한 시간은 폴리에스테르 함유 물질 자체(즉, 플라스틱 제품의 특성과 원산지, 구성, 모양 등), 사용된 폴리펩티드의 유형과 양, 뿐만 아니라 다양한 공정 매개변수(즉, 온도, pH, 추가 제제 등)에 따라 달라질 수 있다. 당업자는 폴리에스테르 함유 물질에 대한 공정 변수를 쉽게 조정할 수 있다.
- [0171] 유리하게는, 분해 공정은 약 10°C 내지 약 60°C, 바람직하게는 약 20°C 내지 약 60°C, 바람직하게는 약 30°C 내지 약 60°C, 보다 바람직하게는 약 40°C 내지 약 60°C, 더욱 더 바람직하게는 약 40°C 내지 약 50°C, 또는 훨씬 더 바람직하게는 약 45°C의 온도에서 실시된다. 온도는 전형적으로 폴리펩티드가 불활성화되고/되거나 재조합 미생물이 본 명세서에 기술된 폴리펩티드를 합성, 생산 또는 방출하지 않는 온도에 상응하는 비활성화 온도 이하로 유지된다. 하나의 실시양태에서, 온도는 폴리에스테르 함유 물질 중 폴리에스테르의 유리 전이 온도(Tg) 이하로 유지된다. 하나의 실시양태에서, 분해 공정 또는 방법은 약 10°C 내지 약 60°C, 바람직하게는 약 20°C 내지 약 60°C, 바람직하게는 약 30°C 내지 약 60°C, 보다 바람직하게는 약 40°C 내지 약 60°C, 더욱 더 바람직하게는 약 40°C 내지 약 50°C, 또는 훨씬 더 바람직하게는 약 45°C의 온도에서 실시된다. 이러한 공정 또는 방법은 폴리펩티드가 여러 번 사용될 수 있고/거나 재활용될 수 있는 온도에서 연속적인 방식으로 적절하게 실시될 수 있다.
- [0172] 유리하게는, 분해 공정 또는 방법은 5 내지 11 사이의 pH, 바람직하게는 6 내지 10 사이의 pH, 보다 바람직하게는 6.5 내지 9 사이의 pH, 더욱 더 바람직하게는 7 내지 8 사이의 pH에서 실시된다.
- [0173] 하나의 실시양태에서, 폴리에스테르 함유 물질은 폴리에스테르와 효소 사이의 접촉 표면을 증가시키기 위해 폴리에스테르 구조를 물리적으로 변화시키기 위해 폴리펩티드와 접촉되기 전에 전처리될 수 있다.
- [0174] 중합해체 또는 분해 공정 또는 방법으로부터 생성된 단량체는 순차적으로 또는 연속적으로 적절하게 회수될 수 있다. 출발 폴리에스테르 함유 물질에 따라 단일 유형의 단량체 또는 여러 가지 다른 유형의 단량체가 회수될 수 있다.
- [0175] 회수된 단량체는 임의의 적합한 정제 방법을 사용하여 추가로 정제되고 재중합 가능한 형태로 조절될 수 있다. 적합한 정제 방법의 예는 스트리핑 공정, 수용액 분리, 증기 선택적 응축, 바이오 공정 후 매질의 여과 및 농축, 분리, 증류, 진공 증발, 추출, 전기 투석, 흡착, 이온 교환, 침전, 결정화, 농축 및 산 첨가 탈수 및 침전, 나노여과, 산 촉매 처리, 반연속 모드 증류 또는 연속 모드 증류, 용매 추출, 증발 농축, 증발 결정화, 액체/액체 추출, 수소화, 공비 증류 공정, 흡착, 컬럼 크로마토그래피, 단순 진공 증류 및 미세여과를 결합 여부에 관계없이 포함한다.
- [0176] 재중합 가능한 단량체는 새로운 폴리에스테르를 합성하는데 사용될 수 있다. 유리하게는 동일한 성질의 폴리에스테르가 재중합된다. 그러나 예를 들어 새로운 공중합체를 합성하기 위해 회수된 단량체를 다른 단량체와 혼합하는 것이 가능하다. 대안적으로, 회수된 단량체는 관심 있는 새로운 화학적 화합물을 생산하기 위해 화학 중간체로 사용될 수 있다.
- [0177] 본 개시내용은 또한 폴리펩티드, 조성물 및/또는 상기 폴리펩티드를 발현하는 숙주 세포 또는 상기 폴리펩티드를 함유하는 이의 추출물을 포함하는 플라스틱 화합물로 확장된다.
- [0178] 본 개시내용은 또한 폴리펩티드, 조성물 및/또는 상기 폴리펩티드를 발현하는 숙주 세포 또는 상기 폴리펩티드를 함유하는 이의 추출물을 포함하는 마스터배치 조성물로 확장된다.
- [0179] 유리하게는, 본 명세서에 기술된 이러한 플라스틱 화합물 또는 마스터배치 조성물은 본 명세서에 기술된 폴리펩티드를 포함하는 폴리에스테르 함유 물질 및/또는 플라스틱 물품의 생산에 사용될 수 있다.
- [0180] 하나의 실시양태에서, 생성된 플라스틱 화합물, 마스터배치 조성물 또는 플라스틱 물품은 표준 EN 13432, 표준 ASTM D6400, OK 생분해 토양(Vincotte 라벨), OK 생분해 수(Vincotte 라벨), OK 퇴비(Vincotte 라벨), OK Home Compost(Vincotte 라벨)과 같은 관련 표준 및/또는 당업자에게 공지된 라벨 중 적어도 하나를 준수하는 생분해성 플라스틱 화합물, 마스터배치 조성물 또는 플라스틱 물품이다.
- [0181] 유리하게는, 폴리에스테르 함유 물질(즉, 플라스틱 화합물, 마스터배치 조성물 또는 플라스틱 제품)의 분해 공정은 10°C 내지 50°C, 바람직하게는 15°C 내지 40°C, 더 바람직하게는 20°C 내지 30°C, 더욱 바람직하게는 28°C, +/- 2°C의 온도에서 실시된다.
- [0182] 대안적으로, 폴리에스테르 함유 물질(즉, 플라스틱 화합물, 마스터배치 조성물 또는 플라스틱 제품)의 분해 공

정은 50℃ 내지 60℃, 보다 바람직하게는 55℃, +/- 2℃의 온도에서 실시된다.

[0183] 본 명세서에 개시된 MHETase 변이체는 산업적 적용을 포함하는 다양한 적용에 적합하며, 이의 예로는 세제, 사료 조성물(동물 사료용 포함), 옷감 생산, 전자제품 및 생물의학 적용의 첨가제로서 포함된다. 예를 들어, 본 명세서에 개시된 MHETase 변이체는 옷감 생산에 사용될 수 있으며, 여기서 이는 옷감 섬유의 특성을 적절하게 변형시키기 위한 엑소뉴클레아제로서 사용될 수 있다.

[0184] 본 발명의 일부 바람직한 국면을 예시하는 하기 실시예를 참조하여 본 발명을 기술할 것이다. 그러나, 본 발명에 대한 다음 기재의 특수성은 본 발명의 전술한 기재의 일반성을 대체하는 것은 아니라는 점을 이해해야 한다.

[0185] **실시예**

[0186] **물질 및 방법**

[0187] 재조합 하이드록시에틸 테레프탈레이트 가수분해효소(MHETase)의 작제, 발현 및 정제

[0188] *A. 컨센서스 설계 작제*

[0189] 야생형 I. 사카이엔시스 MHETase(서열번호 1)를 시드 서열로 사용하여 BLAST+에 의해 5,000개의 서열을 수집하였고, 10-5의 E-값 임계값을 사용하였다.

[0190] 서열번호 1 (UniProt 수탁번호 A0AOK8P8E7)

[0191] GGGSTPLPLPQQPPQEQEPPPPVPLASRAACEALKDGNMVMWPNAAATVVEVAAWRDAAPATASAAALPEHCEVSGAI AKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSAATGSI GGGQIASALSRNFATIA TDGGHDNAVNDNPDALGTVAFGLDPQARLDMGYNSYDQVTQAGKAAVARFYGRAADKSYF IG CSEGGREGMMLSQRFPSHYDGI VAGAPGYQLPKAGISGAWTTQSLAPAAVGLDAQVPLINKSFSADLHLLSQA ILGTCDALDGLADGIVDNYRACQAAFD PATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWDAGMSLGGTTPYNGWRSWWLGSFNSSANNAQRVSGFSARSWLVDFATP PEPMPMTQVAARMKFDFFDIDPLKIWATSGQFTQSSMDWHGATSTDLA AFRDRGGKMLIYHGMSDAAFSALDTADYYERLGAAMPAGAFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISAWSGTPGYFGVAARTRPLCPYPQIARYKSGDINTEANFACAAPP

[0192] 315개의 비-중복 서열은 MHETase와 높은 유사성을 보였으며 UniProt 데이터베이스에서 검색되었다. SignalP4.0을 사용하여 펩티드 수송 신호가 식별되고 결손되었다. 모든 서열은, 최종 선별된 정렬이 수동으로 정제되기 전에 PROMALS3D 라이브러리-기반 서열 정렬 알고리즘 및 사용 가능한 MHETase 구조(6QGB)를 사용하여 정렬되었다. 각 아미노산 위치의 정렬 컨센서스는 다양한 임계값(95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55% 및 50%)을 사용하여 작제되었다. N' 말단에서 처음 19개 아미노산이 결손된 MHETase 유전자(GenBank 수탁 번호 A0AOK8P8E7)의 절단된 코돈 최적화 버전 및 10개의 컨센서스 설계가 상업적으로 합성되어 pET-28a(+)(Twist Bioscience)에 클로닝되었다.

[0193] *B. 단백질 발현 및 정제*

[0194] 플라스미드를 전기감응성 대장균 SHuffle T7 Express 세포(New England Biolab)로 전기천공에 의해 형질전환시키고 100 µg/mL 암피실린이 보충된 LB (Lysogeny Broth) 환천에 도말하였으며 37℃에서 밤새 배양하였다. 단일 콜로니를 사용하여 100 µg/mL 암피실린(LBA)이 보충된 10 ml의 LB를 접종하고 30℃에서 밤새 배양했다. 이 배양물을 1L의 LBA에 첨가하고 OD600이 1.0에 도달할 때까지 30℃에서 배양했다. 최종 농도 1 mM의 이소프로필 β-D-1-티오갈락토피라노사이드를 첨가하고 배양물을 16시간 동안 18℃ 배양으로 이동시켰다.

[0195] 4℃에서 15분 동안 5000 x g로 원심분리하고 용해 완충액(500 mM NaCl, 30 mM 이미다졸, 0.5 mg/mL 라이소자임, 1%(v/v) TritonX-100, 1 U/ml 터보뉴클레아제(Sigma), 0.5 mM 디티오프레이톨 (DTT) 및 25 mM HEPES pH 7.5)에 재현탁하여 세포를 수확했다. 세포 현탁액을 50% 전력 및 5분 동안의 펄스 시간으로 초음파를 2회 처리하여 용해시키고, 4℃에서 45분 동안 32,000 x g에서 원심분리하여 가용성 세포 용해물을 불용성 세포 잔해로부터 분리했다. 용해물을 0.45 µm 공극 크기 필터를 통과시킨 후 용해 완충액으로 평형화된 5 mL HisTrap HP(GE Healthcare Life Sciences)를 사용하여 니켈 충전 IMAC로 정제시키고 용리 완충액(500 mM NaCl, 500 mM Imidazole, 0.5 mM 디티오프레이톨 (DTT) 및 25 mM HEPES pH 7.5)으로 용리시켰다. MHETase와 관련된 용리액을 수집하고, 농축하고, 0.2 µm 필터를 통해 여과시켰다. 여과된 산물은 SEC 완충액(150 mM NaCl, 25 mM HEPES pH 7.5)으로 평형화된 HiLoad 26/600 Superdex 200 (GE Healthcare Life Sciences)을 사용하여 추가로 정제하였다.

[0196] *C. 발색 검정*

- [0197] 발현된 단백질의 세포 현탁액 50 μ L를 반응 완충액(90 mM NaCl, 45 mM 인산나트륨 pH 7.5) 150 μ L에 희석했다. 100% (v/v) DMSO에 50 μ L의 10 mM INT(1-나프틸 테레프탈레이트) 및 5 mM 패스트 블루 B 염 염료를 첨가하여 반응을 개시하였다. 465 nm에서의 흡광도는 Epoch 마이크로플레이트 분광광도계(BioTek)를 사용하여 30분 동안 모니터링하였다.
- [0198] 동역학 측정을 위해 가용성 세포 용해물 대신 최종농도 7.5 nM로 균질한 MHETase를 1 mM에서 7.8 μ M까지 다양한 농도의 INT와 함께 사용했다. 패스트 블루 B 염의 농도는 일정하게 유지되었다. 1-나프톨의 1 mM 내지 7.8 μ M의 검정 곡선을 사용하여 흡광도를 산물 농도로 변환했다.
- [0199] D. SDS-PAGE Atto550
- [0200] 문헌 [참조: Raducanu 외 (2020, Journal of Biological Chemistry 295(34):12214-12223)]로부터 방법을 채택하였다. 1 mL의 세포 펠릿을 SEC 완충액에 희석된 1x BugBuster(Merck-Millipore)에 재현탁하고 실온에서 10분 동안 배양했다. 혼합물을 15,000 x g에서 10분 동안 원심분리하고 가용성 세포 용해물 5 μ L를 SDS-PAGE 겔에서 140V에서 60분 동안 실행하였다. 겔을 Milli-Q 물(MQ)에서 30초 동안 2회 마이크로파 처리한 다음, 고정 용액(MQ 내 40%(v/v) 메탄올, 10% (v/v) 아세트산)에서 2분 동안 마이크로파 처리했다. 이제 고정된 단백질 겔을 다시 MQ에서 10분 동안 마이크로파 처리하고, PBS 완충액 내에 NTA-Atto550 염료를 1 : 3000으로 희석한 어두운 환경의 진탕기에서 1시간 동안 배양했다. 이어서, 겔을 따뜻한 MQ가 담긴 용기로 이동시켜 30분 동안 추가 진탕시켰다. 그런 다음 DyLight 550 형광단 옵션을 사용하는 ChemiDoc MP 이미징 시스템 (BIO-RAD)을 사용하여 겔을 이미지화했다.
- [0201] E. 원편광 이색성 및 열 안정성
- [0202] MHETase에 대한 원편광 이색성 스펙트럼은 Applied Photophysics Chirscan Spectrometer의 1 mm 석영 큐벳에서 수행되었다. 균질한 효소를 25 mM 아세트산나트륨 pH 7.5에 0.2 mg/mL로 희석했다. CD 스펙트럼은 200 nm에서 260 nm 사이 20°C에서 측정되었으며 적응형 샘플링이 활성화된 상태에서 1 nm의 대역폭과 0.5초의 스캔 속도를 사용하였다. 스펙트럼을 3회 측정하고 결과값에서 완충액 블랭크를 차감하였다. 열 용융 평가를 위해 용액의 온도가 1°C/분의 속도로 20°C에서 90°C로 증가함에 따라 222 nm의 CD를 기록했다. T_m을 결정하기 위해 표준 S자형 곡선을 열 용융 데이터에 피팅하였다.
- [0203] F. MHETase에 대한 고성능 액체 크로마토그래피(HPLC) 활성 평가
- [0204] HPLC 분석은 문헌[참조: Palm 외 (2019)]로부터 채택하였다. 균질한 MHETase를 반응 완충액에서 최종 농도 7.5 nM (80 μ L)로 희석했다. 100% DMSO에 용해된 1 mM MHET 20 μ L를 첨가하여 반응을 개시했다. 100 μ L 켄칭(quenching) 완충액(160 mM 인산나트륨 pH 2)을 첨가하여 설정된 시점(0, 10, 30 및 60분) 후에 반응을 켄칭하고 10분 동안 80°C로 가열했다. 10 μ L 부피의 반응 혼합물을 Agilent ZORBAX SB-C18, 3.5 μ m, 4.6 X 150 mm 컬럼에 로딩했다. TPA와 MHET는 30°C에서 1mL/분의 유속을 사용하여 분리되었으며, 50% 인산염 완충액과 50% 아세트나이트릴에서 7분의 실행 시간에 걸쳐 평형화되었다. TPA와 MHET는 240 nm에서 검출되었으며 검정 곡선에 대해 정량화하였다.
- [0205] 실시예 1: 변이체 MHETase
- [0206] 컨센서스 기반 설계는 MHETase와 이와 가장 근접한 동류 (relatives)의 정렬 순서를 사용하여 수행하였다. 이로부터, 상이한 컨센서스 임계값(95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55% 및 50%)을 사용하여 다수의 상이한 조합 MHETase 서열을 작제하였다. 예를 들어, 95% 임계값의 컨센서스 설계는 정렬된 서열의 95%에서 관찰되지만 WT MHETase에서는 관찰되지 않는 모든 차이를 나타낸다. 야생형(WT) MHETase(서열번호 1) 및 상이한 컨센서스 설계(서열번호 2 내지 36)의 아미노산 서열이 도 1에 도시되어 있다. WT MHETase 및 상이한 컨센서스 설계(서열번호 37 내지 72)의 핵산 서열이 도 2에 도시되어 있다. 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55% 및 50%의 임계값으로 인한 다양한 컨센서스 설계는 본 명세서에서 각각 라운드 1 컨센서스(Round 1 Consensus) A, B, C, D, E, F, G, H, I 및 J로 지칭한다.
- [0207] MHETase의 활성/발현은 전체 세포 현탁액을 사용하여 위에서 설명한 발색 검정을 사용하여 측정하였다.
- [0208] 도 3에 도시된 바와 같이, 변이체 "라운드 1 컨센서스 A"(서열번호 73)는 WT MHETase 또는 다른 컨센서스 설계보다 더 많은 전 세포 활성을 나타냈다. 라운드 1 컨센서스 A는 WT와 비교했을 때 2개의 아미노산 치환, 즉, N156G 및 T159V를 포함하였다. 초기 연구로서, MHETase 활성 및/또는 발현을 더욱 안정화/개선하는 돌연변이를 확인하기 위해 다중 서열 정렬을 통해 확인된 다른 컨센서스 잔기를 기반으로 이 변이체에 점 돌연변이를 추가

했다. 나머지 변이체(돌연변이; 서열번호 3 내지 36)는 라운드 1 컨센서스 A 서열 대비 다음 점 돌연변이 중 하나를 포함한다(명명법은 야생형 MHETase, 서열번호 1의 아미노산 위치를 나타냄): T68V, A78P, E110A, M117L, E128Q, S131G, N134D, A161G, G156N, D191L, M192Y, S196A, Y197V, G204A, A207L, A216P, E226N, L234A, S235A, P255V, G258A, S260A, T264L, T265L, N284L, L295V, S296A, T355R, A377P, S463L, A493P, Y503W, E594A 및 N496S.

- [0209] 또한 도 3에 도시된 바와 같이, 라운드 1 컨센서스 A 서열에 삭제될 때 점 돌연변이 S196A, Y197V, S235A, P255V, S260A, S286A, Y503W는 대장균에서 발현시 훨씬 더 높은 전 세포 활성을 보여주었다.
- [0210] 가용성 세포 용해물을 SDS-PAGE를 통해 실행하고 NTA-Atto550으로 염색하여 효소의 발현 속도를 확인했다. MHETase는 겔에서 64kDa로 실행되며 도 4에 도시되어 있다. 이 겔은 보다 구체적인 염색 방법으로도 대부분의 MHETase 돌연변이(대조군을 포함함)가 보이지 않음을 보여준다. 이 결과는 낮은 가용성 단백질 발현과 일치한다. 라운드 1 컨센서스 A 및 변이체 B 백본에 점 돌연변이 S196A, Y197V, S235A, P255V, S260A, S286A 또는 Y503W를 포함하는 변이체는 야생형 (WT)에 비해 대장균에서 증가된 이중 발현을 보여주는 것으로 나타났다.
- [0211] 또한, 정제된 단백질 변이체를 사용하여, 선택된 생성된 변이체의 안정성을 테스트하였다. 도 5에 도시한 바와 같이, WT MHETase, 라운드 1 컨센서스 A 및 라운드 1 컨센서스 A + Y503W의 열 용융(T_m)은 모두 약 59°C인 반면, 라운드 1 컨센서스 A에 Y197V 점 돌연변이를 도입하면 T_m 이 52°C로 감소했다.
- [0212] 산업 응용에 있어서 야생형 MHETase 사용의 주요 제한사항 중 하나는 산업 균주에서 발현이 낮다는 점이다. 본 명세서에 개시된 변이체에 의해 입증된 바와 같이, 전 세포 활성(촉매 효율 및 활성 단백질의 발현 수율의 조합)의 증가는 생산 및 사용 비용을 유리하게 절감시킨다.
- [0213] 실시예 2: 활성이 향상된 엔지니어링 MHETase
- [0214] 열안정성이 향상된 변이체를 생성하기 위해 라운드 1 컨센서스 A(N156G 및 T159V 점 돌연변이 포함함)를 추가 엔지니어링을 위한 기반으로 선택하였으며, 라운드 1 컨센서스 A는 열안정성에 변화가 거의 없이 (도 5) 전 세포 MHETase 활성(도 3)이 개선된 것으로 나타났다. 이 새로운 조작 과정에서는 점 돌연변이를 라운드 1 컨센서스 A에 추가하여 돌연변이 라이브러리(라운드 2)를 생성하고 전 세포 활성에 대해 스크리닝하였다. 이 과정은 가장 유망한 변이체를 선택하고 여러 설계 라운드에 걸쳐 단일 점 돌연변이를 반복적으로 도입하는 작업을 수반한다.
- [0215] 라운드 2에서는 다음과 같은 점 돌연변이를 라운드 1 컨센서스 A 백본에 도입하였다(명명법은 야생형 MHETase, 서열번호 1의 아미노산 위치를 나타냄): T593D, P543A, Y503W, S463L, P449A, T355R, G301A, S296A, H293Q, S286A, N284L, S267A, T265L, T264L, S260A, P255V, S235A, K218R, A216P, A207L, Y197B, S196A, D191L, L190I, E110A, Y107Q, A99N, A81P, A78P, T68V.
- [0216] 라운드 2에서 선택된 최고의 성과를 나타내는 변이체는 "라운드 2 Y503W"(N156G, T159V 및 Y503W 점 돌연변이를 포함함, 서열번호 74)이다.
- [0217] 라운드 3에서, "라운드 2 Y503W에 도입된 점 돌연변이는 다음과 같다: E594A, A493P, A377P, S286A, S267A, T264L, S260A, S196A, M192Y.
- [0218] 라운드 3에서 선택된 최고의 성과를 나타내는 변이체는 "라운드 3 M192Y" (N156G, T159V, M192Y 및 Y503W 점 돌연변이를 포함함; 서열번호 75)이다.
- [0219] 라운드 4에서, 라운드 3 M192Y에 도입된 점 돌연변이는 다음과 같다: +N564, S561A, G534A, L486V, A469P, +P467, H467M, W398K, A377P, M361F, I357L, S288T, L282D, S286A, S267A, S260A, Y252F, M233V, G231A, E230H, V208I, Q202P, V200L, T162S, A161S, G156N, V159T, L137V, G130N, R114E, L112G, A79G (여기서 '+'는 특정 위치에 아미노산이 삽입된 것을 의미한다).
- [0220] 라운드 4에서 선택된 최고의 성과를 나타내는 변이체는 "라운드 4 G156N" (T159V, M192Y 및 Y503W 점 돌연변이를 포함함; 서열번호 76)이다.
- [0221] 라운드 5에서, 라운드 4 G156N에 도입된 점 돌연변이는 다음과 같다: N592E, I582V, G562R, R537Q, A494V, W466F, Q461E, M361F, I357L, N316D, I283L, Y252F, V246L, Y242F, G231A, V200L, G164A, I104V, E90A, E71T (여기서 '+'는 특정 위치에 아미노산이 삽입된 것을 의미한다).

[0222] 라운드 5에서 선택된 최고의 성과를 나타내는 변이체는 "라운드 5 Y252F"(T159V, M192Y, Y252F 및 Y503W를 포함함; 서열번호 77)이고, 이는 야생형에 비해 전 세포 활성이 16배 증가한 것으로 나타났다(도 6).

[0223] 각 라운드에서 가장 좋은 변이체의 발현은 His-tag 특이적 형광 라벨 ATTO550으로 염색된 SDS-PAGE를 사용하여 정성적으로 측정하였다. WT MHETase의 발현은 배경 내인성(background endogenous) 대장균 단백질 발현에 비해 결정적으로 검출되기에는 너무 낮았다. 그러나 라운드 2의 엔지니어링을 통해 변이체 MHETase(라운드 2 Y503W) 재조합 발현 수준이 ATTO550 형광염료를 이용하여 육안으로 검출가능한 수준으로 향상되었다(도 7).

[0224] 본 명세서에 설명된 발색 검정은 정제된 단백질(도 8)을 사용하여 이들 효소의 동역학 매개변수를 결정했다. 모든 변이체의 촉매 효율(k_{cat}/K_M)은 야생형 MHETase의 촉매 효율과 비교적 일정하게 유지되었다. 도 9는 다양한 기질 농도에서 효소 작용(enzyme function) 속도의 미카엘리스 멘텐(Michaelis Menten) 플롯을 도시한다. 이들 데이터로부터 표 1에 나타난 바와 같이 k_{cat} 과 K_M 이라는 두 가지 상수가 획득되었다.

표 1

[0225]

선택된 MHETase 변이체의 미카엘리스 멘텐 동역학 매개변수			
효소 변이체	$K_M \pm SE$ (μM)	$k_{cat} \pm SE$ (초^{-1})	$k_{cat}/K_M \pm SE$ ($\mu M^{-1} \cdot \text{초}^{-1}$)
MHETase WT	21.41 \pm 3.63	88.69 \pm 4.34	4.14 \pm 0.90
라운드 1 컨센서스 A	18.55 \pm 3.19	73.00 \pm 3.47	3.94 \pm 0.86
라운드 2 Y503W	18.07 \pm 2.36	72.54 \pm 2.60	4.01 \pm 0.67
라운드 3 M192Y	10.24 \pm 1.30	33.96 \pm 0.99	3.32 \pm 0.52
라운드 4 G156N	11.31 \pm 1.81	39.59 \pm 1.50	3.50 \pm 0.69
라운드 5 Y252F	9.38 \pm 1.69	36.13 \pm 1.45	3.85 \pm 0.85

[0226] 본 명세서에 기술된 발색 검정을 사용하여 관찰된 데이터를 검증하기 위해, 천연 기질 MHET에 대한 야생형 MHETase 및 라운드 5의 최상의 변이체(라운드 5 Y252F; 서열번호 77)의 활성을 고성능 액체 크로마토그래피(HPLC)로 결정하였다. HPLC는 야생형 MHETase 또는 라운드 5 Y252F를 포함하는 효소 반응에서 시간 경과에 따른 테레프탈산(TPA) 및 MHET의 농도를 측정하기 위해 사용되었다. 데이터는 발색 검정을 사용하여 관찰된 k_{cat} 및 K_M 의 감소와 일치했으며 라운드 5 Y252F가 MHET를 TPA로 가수분해한다는 것을 확인하였다. (도 11 참조).

[0227] 정제된 단백질 변이체를 사용하여 컨센서스 설계에 대한 안정성 테스트도 수행하였다. 위치 192 및 252의 돌연변이는 WT MHETase에 비해 (T_m)의 열 안정성을 향상시킨다(도 10).

[0228] 전반적으로, 이러한 안정화 특성은 야생형 MHETase에 비해 라운드 5 Y252F의 개선된 발현 및 활성에 기여한다. 그러나 이러한 개선에도 불구하고 이 변이체의 특정 활성은 낮은 k_{cat} 에서 입증된 바와 같이 야생형에 비해 감소한다 (표 1).

[0229] 라운드 5 Y252F의 개별 돌연변이 중 일부를 야생형 잔기로 복귀시키기 위해 추가 엔지니어링을 수행하였다. 위치 156의 경우, 라운드 4에서 이 잔기를 야생형 동일성으로 복귀하는 영향을 조사하기 위해 돌연변이 N156G가 MHETase R5의 배경으로 다시 도입되었다. 패스트 블루 분석으로 측정된 전 세포 활성은 도 12에 도시된다.

[0230] 돌연변이 M192Y가 라운드 5 Y252F에서 야생형 잔기(M)로 복귀("R5-Y192M"; 서열번호 78)하는 경우, 라운드 5 Y252F에 비해 특정 활성이 약 3배 증가함을 관찰하였고(도 11), 서열번호 1의 야생형 MHETase의 활성과 더 유사한 수준으로 특정 활성을 복원하였다. 라운드 5 Y252F에서 다른 점 돌연변이의 복귀는 효소 활성을 감소시키거나 크게 변화시키지 않았다. MHETase 변이체의 열 안정성은 표 2에 도시되었다.

표 2

[0231]

야생형(WT) MHETase 및 선택된 MHETase 변이체의 열안정성	
효소	T_m ($^{\circ}C$)
MHETase WT	59.5
라운드 1 컨센서스 A 변이체	59.6
라운드 2 Y503W 변이체	57.9
라운드 3 M192Y 변이체	61.2

라운드 4 G156N 변이체	62.8
라운드 5 Y252F 변이체	62.6
라운드 5-Y192M 변이체	59.6

- [0232] 본원에 인용된 모든 특허, 특허 출원 및 간행물의 개시 내용은 전문이 참조로 본 명세서에 포함된다.
- [0233] 본 명세서에서 인용된 참고문헌이 본 출원에 대한 "선행 기술"로 이용 가능하다는 사실을 인정하는 것으로 해석되어서는 안 된다.
- [0234] 명세서 전반에 걸쳐, 본 발명을 임의의 하나의 실시예 또는 특정 특징들의 집합으로 제한하지 않고 본 발명의 바람직한 실시양태를 기술하는 것을 목적으로 한다. 따라서, 당업자는 본 개시 내용을 고려하여 본 발명의 범위를 벗어나지 않고 예시된 특정 실시양태에서 다양한 수정 및 변경이 이루어질 수 있음을 인식할 것이다. 이러한 모든 수정 및 변경은 첨부된 청구범위의 범위 내에 포함되도록 의도된다.

도면

도면1a

```

>MHETase_wild-type (WT; SEQ ID NO:1)
GGGSTPLPLPQQPPQEQPPPPVPLASRAACEALKDGNQDMVWPNAATVVEVAAWRDAA
PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
ATGSIGGGQIASALSRNFATIADGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLLSQRFP SHYD GIVAGAPGYQLPKAGI
SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
MMKFD FDI DPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
TADYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
WSGTPGYFGVAARTRPLCPYPQIARYKGGSDINTEANFACAAPP
>MHETase_Consensus95_AminoAcid (SEQ ID NO:2)
GGGSTPLPLPQQPPQEQPPPPVPLASRAACEALKDGNQDMVWPNAATVVEVAAWRDAA
PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLLSQRFP SHYD GIVAGAPGYQLPKAGI
SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
MMKFD FDI DPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
TADYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
WSGTPGYFGVAARTRPLCPYPQIARYKGGSDINTEANFACAAPP
>MHETase_Consensus95_Y197V_AminoAcid (SEQ ID NO:3)
GGGSTPLPLPQQPPQEQPPPPVPLASRAACEALKDGNQDMVWPNAATVVEVAAWRDAA
PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLLSQRFP SHYD GIVAGAPGYQLPKAGI
SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
MMKFD FDI DPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
TADYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
WSGTPGYFGVAARTRPLCPYPQIARYKGGSDINTEANFACAAPP
>MHETase_Consensus95_197_T68V_AminoAcid (SEQ ID NO:4)
GGGSTPLPLPQQPPQEQPPPPVPLASRAACEALKDGNQDMVWPNAATVVEVAAWRDAA
PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLLSQRFP SHYD GIVAGAPGYQLPKAGI
SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
MMKFD FDI DPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
TADYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
WSGTPGYFGVAARTRPLCPYPQIARYKGGSDINTEANFACAAPP
>MHETase_Consensus95_197_A78P_AminoAcid (SEQ ID NO:5)
GGGSTPLPLPQQPPQEQPPPPVPLASRAACEALKDGNQDMVWPNAATVVEVAAWRDPA
PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLLSQRFP SHYD GIVAGAPGYQLPKAGI
SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
MMKFD FDI DPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD

```

도면1b

TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGDINTEANFACAAPP
 >MHETase_Consensus95_197_E110A_AminoAcid (SEQ ID NO:6)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYAIKFLRLMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPAATANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGDINTEANFACAAPP
 >MHETase_Consensus95_197_M117L_AminoAcid (SEQ ID NO:7)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYAIKFLRLMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPAATANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGDINTEANFACAAPP
 >MHETase_Consensus95_197_E128Q_AminoAcid (SEQ ID NO:8)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYAIKFLRLMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPAATANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGDINTEANFACAAPP
 >MHETase_Consensus95_197_S131G_AminoAcid (SEQ ID NO:9)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYAIKFLRLMPAEWNGRFFMEGGGGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPAATANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGDINTEANFACAAPP
 >MHETase_Consensus95_197_N134D_AminoAcid (SEQ ID NO:10)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYAIKFLRLMPAEWNGRFFMEGGSGTDGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPAATANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD

도면1c

TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_G156N_AminoAcid (SEQ ID NO:11)
 GGGSTPLPLPQQPPQQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRNFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSDADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTD LAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_A161G_AminoAcid (SEQ ID NO:12)
 GGGSTPLPLPQQPPQQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSDADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTD LAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_D191L_AminoAcid (SEQ ID NO:13)
 GGGSTPLPLPQQPPQQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLLMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSDADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTD LAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_M192Y_AminoAcid (SEQ ID NO:14)
 GGGSTPLPLPQQPPQQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDYGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSDADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTD LAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_S196A_AminoAcid (SEQ ID NO:15)
 GGGSTPLPLPQQPPQQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNAYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSDADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTD LAAFDRGGKMILYHGMSDAAFSALD

도면1d

TADYYERLGAAMPGAAGFARFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_G204A_AminoAcid (SEQ ID NO:16)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAAKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_A207L_AminoAcid (SEQ ID NO:17)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKALVARFYGRAADKSYFIGCSEGGREGMMLSQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_A216P_AminoAcid (SEQ ID NO:18)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAPDKSYFIGCSEGGREGMMLSQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_E226N_AminoAcid (SEQ ID NO:19)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSNGGREGMMLSQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_T234L_AminoAcid (SEQ ID NO:20)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMASQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAGVGLDAQVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDVFATPPEPMPMTQVAAR
 MMKFDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFDRGGKMILYHGMSDAAFSALD

도면1e

TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_S235A_AminoAcid (SEQ ID NO:21)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMVMWPNAAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLAQRFPSHYDGIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_P255V_AminoAcid (SEQ ID NO:22)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMVMWPNAAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYDGIVAGAPGYQLVKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_G258A_AminoAcid (SEQ ID NO:23)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMVMWPNAAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYDGIVAGAPGYQLPKAAI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_S260A_AminoAcid (SEQ ID NO:24)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMVMWPNAAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYDGIVAGAPGYQLPKAGI
 AGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_S264L_AminoAcid (SEQ ID NO:25)
 GGGSTPLPLPQQQPPQEPPPPVPLASRAACEALKDGNMVMWPNAAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIEIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTVAFLDLPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYDGIVAGAPGYQLPKAGI
 SGAWLTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVDFATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD

도면1f

TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKGSVDINTEANFACAAPP
 >MHETase_Consensus95_197_T265L_AminoAcid (SEQ ID NO:26)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDNGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYEIKFRLRMPAEWNGRFFMEGGSNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQLSAPAAVGLDAQVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSSLGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKGSVDINTEANFACAAPP
 >MHETase_Consensus95_197_N284L_AminoAcid (SEQ ID NO:27)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDNGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYEIKFRLRMPAEWNGRFFMEGGSNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQLSAPAAVGLDAQVPLILKSFSDADLHLLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSSLGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKGSVDINTEANFACAAPP
 >MHETase_Consensus95_197_L295V_AminoAcid (SEQ ID NO:28)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDNGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYEIKFRLRMPAEWNGRFFMEGGSNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQLSAPAAVGLDAQVPLINKSFSADLHLVLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSSLGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKGSVDINTEANFACAAPP
 >MHETase_Consensus95_197_S296A_AminoAcid (SEQ ID NO:29)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDNGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYEIKFRLRMPAEWNGRFFMEGGSNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQLSAPAAVGLDAQVPLINKSFSADLHLLAQAILGTCDALDGLADGIVDNYRA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSSLGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPGYFGVAARTRPLCPYPQIARYKGSVDINTEANFACAAPP
 >MHETase_Consensus95_197_T355R_AminoAcid (SEQ ID NO:30)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDNGDMVWPNAATVVEVAAWRDAA
 PATASAAALPEHCEVSGAIAKRTGIDGYPYEIKFRLRMPAEWNGRFFMEGGSNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGTVAFLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQLSAPAAVGLDAQVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVRAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSSLGTTYNQGWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKDFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD

도면1g

TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPTYFVGAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_A377P_AminoAcid (SEQ ID NO:31)
 GGGSTPLPLPQQPPQEP PPPVPLASRAACEALKDGN GDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIKRTGIDGYPIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGT VAFGLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWPD
 AGMSGLSGT TYNQGWRSWWLGSFNSSANNAQRVSGFSARSWL VDFATPPEPMPMTQVAAR
 MMKFD FDI DPLKIWATSGQFTQSSMDWHGATSTD LAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPTYFVGAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_S463L_AminoAcid (SEQ ID NO:32)
 GGGSTPLPLPQQPPQEP PPPVPLASRAACEALKDGN GDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIKRTGIDGYPIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGT VAFGLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGT TYNQGWRSWWLGSFNSSANNAQRVSGFSARSWL VDFATPPEPMPMTQVAAR
 MMKFD FDI DPLKIWATSGQFTQSLMDWHGATSTD LAAFDRGGKMILYHGMSDAAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPTYFVGAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_A493P_AminoAcid (SEQ ID NO:33)
 GGGSTPLPLPQQPPQEP PPPVPLASRAACEALKDGN GDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIKRTGIDGYPIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGT VAFGLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGT TYNQGWRSWWLGSFNSSANNAQRVSGFSARSWL VDFATPPEPMPMTQVAAR
 MMKFD FDI DPLKIWATSGQFTQSSMDWHGATSTD LAAFDRGGKMILYHGMSDPAFSALD
 TADYYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPTYFVGAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_Y503W_AminoAcid (SEQ ID NO:34)
 GGGSTPLPLPQQPPQEP PPPVPLASRAACEALKDGN GDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIKRTGIDGYPIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGT VAFGLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGT TYNQGWRSWWLGSFNSSANNAQRVSGFSARSWL VDFATPPEPMPMTQVAAR
 MMKFD FDI DPLKIWATSGQFTQSSMDWHGATSTD LAAFDRGGKMILYHGMSDAAFSALD
 TADWYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPTYFVGAARTRPLCPYPQIARYKSGSDINTEANFACAAPP
 >MHETase_Consensus95_197_E594A_AminoAcid (SEQ ID NO:35)
 GGGSTPLPLPQQPPQEP PPPVPLASRAACEALKDGN GDMVWPNAATVVEVAWRDAA
 PATASAAALPEHCEVSGAIKRTGIDGYPIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALS RGF AVIATDGGHDNAVNDNPDALGT VAFGLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSQRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSADLHLLSQAILGTCDALDGLADGIVDNRYA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGT TYNQGWRSWWLGSFNSSANNAQRVSGFSARSWL VDFATPPEPMPMTQVAAR
 MMKFD FDI DPLKIWATSGQFTQSSMDWHGATSTD LAAFDRGGKMILYHGMSDAAFSALD

도면1h

TADYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPTYGFVGAARTRPLCPYPQIARYKSGSDINTAANFACAAPP
 >MHETase_Consensus95_197_N596S_AminoAcid (SEQ ID NO:36)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAACEALKDGNMGDMVWPNAATVVEVAAWRDA
 PATASAAALPEHCEVSGAIAKRTGIDGYPIKFRLRMPAEWNGRFFMEGGSGTNGSLSA
 ATGSIGGGQIASALSRGFAVIATDGGHDNAVNDNPDALGTAVFGLDPQARLDMGVNSYDQ
 VTQAGKAAVARFYGRAADKSYFIGCSEGGREGMMLSRFP SHYD GIVAGAPGYQLPKAGI
 SGAWTTQSLAPAAVGLDAQGVPLINKSFSDADLHLLSQAILGTCDALDGLADGIVDNYRA
 CQAAFDPATAANPANGQALQCVGAKTADCLSPVQVTAIKRAMAGPVNSAGTPLYNRWAWD
 AGMSGLSGTTYNQWRSWWLGSFNSSANNAQRVSGFSARSWLVD FATPPEPMPMTQVAAR
 MMKFDIDPLKIWATSGQFTQSSMDWHGATSTDLAAFRDRGGKMILYHGMSDAAFSALD
 TADYERLGAAMPGAAGFARLFLVPGMNHCSGGPGTDRFDMLTPLVAWVERGEAPDQISA
 WSGTPTYGFVGAARTRPLCPYPQIARYKSGSDINTEASFACAAPP
 > MHETase Round 1 Consensus A: (SEQ ID NO: 73)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAA
 CEALKDGNMGDMVWPNAATVVEVAAWRDAAPATASAAALPEHCEVSGAIAKRTGIDGYPIE
 IKFRLRMPAEWNGRFFMEGGSGTNGSLSAATGSIGGGQIASALSRGFAVIATDGGHDNAV
 NDNPDALGTAVFGLDPQARLDMGVNSYDQVTQAGKAAVARFYGRAADKSYFIGCSEGGRE
 GMMLSRFP SHYD GIVAGAPGYQLPKAGISGAWTTQSLAPAAVGLDAQGVPLINKSFSDA
 DLHLLSQAILGTCDALDGLADGIVDNYRACQAAFDPATAANPANGQALQCVGAKTADCL
 PVQVTAIKRAMAGPVNSAGTPLYNRWAWDAGMSGLSGTTYNQWRSWWLGSFNSSANNAQ
 RVSGFSARSWLVD FATPPEPMPMTQVAARMMKFDIDPLKIWATSGQFTQSSMDWHGAT
 STDLAAFRDRGGKMILYHGMSDAAFSALDTADYERLGAAMPGAAGFARLFLVPGMNHCS
 GGPGTDRFDMLTPLVAWVERGEAPDQISAWSGTPGYFGVAARTRPLCPYPQIARYKSGSD
 INTEANFACAAPP
 >MHETase_Round 2 Y503W (SEQ ID NO:74)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAA
 CEALKDGNMGDMVWPNAATVVEVAAWRDAAPATASAAALPEHCEVSGAIAKRTGIDGYPIE
 IKFRLRMPAEWNGRFFMEGGSGTNGSLSAATGSIGGGQIASALSRGFAVIATDGGHDNAV
 NDNPDALGTAVFGLDPQARLDMGVNSYDQVTQAGKAAVARFYGRAADKSYFIGCSEGGRE
 GMMLSRFP SHYD GIVAGAPGYQLPKAGISGAWTTQSLAPAAVGLDAQGVPLINKSFSDA
 DLHLLSQAILGTCDALDGLADGIVDNYRACQAAFDPATAANPANGQALQCVGAKTADCL
 PVQVTAIKRAMAGPVNSAGTPLYNRWAWDAGMSGLSGTTYNQWRSWWLGSFNSSANNAQ
 RVSGFSARSWLVD FATPPEPMPMTQVAARMMKFDIDPLKIWATSGQFTQSSMDWHGAT
 STDLAAFRDRGGKMILYHGMSDAAFSALDTADYERLGAAMPGAAGFARLFLVPGMNHCS
 GGPGTDRFDMLTPLVAWVERGEAPDQISAWSGTPGYFGVAARTRPLCPYPQIARYKSGSD
 INTEANFACAAPP
 >MHETase_Round 3 M192Y (SEQ ID NO:75)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAA
 CEALKDGNMGDMVWPNAATVVEVAAWRDAAPATASAAALPEHCEVSGAIAKRTGIDGYPIE
 IKFRLRMPAEWNGRFFMEGGSGTNGSLSAATGSIGGGQIASALSRGFAVIATDGGHDNAV
 NDNPDALGTAVFGLDPQARLDYGYNSYDQVTQAGKAAVARFYGRAADKSYFIGCSEGGRE
 GMMLSRFP SHYD GIVAGAPGYQLPKAGISGAWTTQSLAPAAVGLDAQGVPLINKSFSDA
 DLHLLSQAILGTCDALDGLADGIVDNYRACQAAFDPATAANPANGQALQCVGAKTADCL
 PVQVTAIKRAMAGPVNSAGTPLYNRWAWDAGMSGLSGTTYNQWRSWWLGSFNSSANNAQ
 RVSGFSARSWLVD FATPPEPMPMTQVAARMMKFDIDPLKIWATSGQFTQSSMDWHGAT
 STDLAAFRDRGGKMILYHGMSDAAFSALDTADYERLGAAMPGAAGFARLFLVPGMNHCS
 GGPGTDRFDMLTPLVAWVERGEAPDQISAWSGTPGYFGVAARTRPLCPYPQIARYKSGSD
 INTEANFACAAPP
 >MHETase_Round 4 G156N (SEQ ID NO:76)
 GGGSTPLPLPQQPPQEPPIPPVPLASRAA
 CEALKDGNMGDMVWPNAATVVEVAAWRDAAPATASAAALPEHCEVSGAIAKRTGIDGYPIE
 IKFRLRMPAEWNGRFFMEGGSGTNGSLSAATGSIGGGQIASALSRFAVIATDGGHDNAV
 NDNPDALGTAVFGLDPQARLDYGYNSYDQVTQAGKAAVARFYGRAADKSYFIGCSEGGRE
 GMMLSRFP SHYD GIVAGAPGYQLPKAGISGAWTTQSLAPAAVGLDAQGVPLINKSFSDA

도면1i

```
DLHLLSQAILGTCDALDGLADGIVDN YRACQAAFDPATAANPANGQALQCVGAKTADCLS
PVQVTAIKRAMAGPVNSAGTPLYNRWAWDAGMSGLSGTTYNQGWRSWWLGSFNSSANNAQ
RVSGFSARSWLVD FATPPEPMPMTQVAARMMKDFDIDPLKIWATSGQFTQSSMDWHGAT
STDLAAFDRDRGGKMILYHGMSDAAFSALDTADWYERLGAAMPGAAGFARLFLVPGMNHCS
GGPGTDRFDM LTPLVAVWERGEAPDQISAWSGTPGYFGVAARTRPLCPYPQIARYKSGSD
INTEANFACAAPP
>MHETase_Round 5 Y252F (SEQ ID NO:77)
GGGSTPLPLPQQQPPQEPPPPVPLASRAA
CEALKDGN GDMVWPNAATVVEVAAWRDAAPATASAAALPEHCEVSGAIKRTGIDGYPYE
IKFRLRMPAEWNGRFFMEGGSGTNGSLSAATGSIGGGQIASALSRNFAVIATDGGH DNAV
NDNPDALGTVAFLDLPQARLDYGYNSYDQVTQAGKAAVARFYGRAADKSYFIGCSEGGRE
GMMLSQRFP SHYD GIVAGAPGFQLPKAGISGAWTTQSLAPAAVGLDAQGVPLINKSFSDA
DLHLLSQAILGTCDALDGLADGIVDN YRACQAAFDPATAANPANGQALQCVGAKTADCLS
PVQVTAIKRAMAGPVNSAGTPLYNRWAWDAGMSGLSGTTYNQGWRSWWLGSFNSSANNAQ
RVSGFSARSWLVD FATPPEPMPMTQVAARMMKDFDIDPLKIWATSGQFTQSSMDWHGAT
STDLAAFDRDRGGKMILYHGMSDAAFSALDTADWYERLGAAMPGAAGFARLFLVPGMNHCS
GGPGTDRFDM LTPLVAVWERGEAPDQISAWSGTPGYFGVAARTRPLCPYPQIARYKSGSD
INTEANFACAAPP
>MHETase_R5-Y192M (SEQ ID NO:78)
GGGSTPLPLPQQQPPQEPPPPVPLASRAA
CEALKDGN GDMVWPNAATVVEVAAWRDAAPATASAAALPEHCEVSGAIKRTGIDGYPYE
IKFRLRMPAEWNGRFFMEGGSGTNGSLSAATGSIGGGQIASALSRNFAVIATDGGH DNAV
NDNPDALGTVAFLDLPQARLDYGYNSYDQVTQAGKAAVARFYGRAADKSYFIGCSEGGRE
GMMLSQRFP SHYD GIVAGAPGFQLPKAGISGAWTTQSLAPAAVGLDAQGVPLINKSFSDA
DLHLLSQAILGTCDALDGLADGIVDN YRACQAAFDPATAANPANGQALQCVGAKTADCLS
PVQVTAIKRAMAGPVNSAGTPLYNRWAWDAGMSGLSGTTYNQGWRSWWLGSFNSSANNAQ
RVSGFSARSWLVD FATPPEPMPMTQVAARMMKDFDIDPLKIWATSGQFTQSSMDWHGAT
STDLAAFDRDRGGKMILYHGMSDAAFSALDTADWYERLGAAMPGAAGFARLFLVPGMNHCS
GGPGTDRFDM LTPLVAVWERGEAPDQISAWSGTPGYFGVAARTRPLCPYPQIARYKSGSD
INTEANFACAAPP
```

도면2a

```

>MHETase_Wt_Nucleotide (SEQ ID NO:37)
GGTGGTGGTAGCACACCCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACC GTTGTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTAACTTTGAACC
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTGGTCTGGATCCGCAGGCACGCTGGATATGGGTTATAATAGCTATGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTAGGCAAGTATTGATTCCGGCAACAGCAGCAAAATCCGGCAAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCGGTTAATAGCGCAGGTACACCCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAAATAGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGAACCCAGCACCGATCTGGCAGCCTTTCTGAT
CTGGTGGTAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTCACTGGAT
ACAGCAGATTATTATGAACGCTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGTTGTCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGTTATTTGGTGTGGCGCACGTACCCGTCGCTGTGTCCGTAT
CCGAGATCGCACGTTATAAAGGTAGCGGTGATTAACACCGAAGCAATTTTGCATGT
GCAGCACCTCCGTAA

>MHETase_Consensus95_Nucleotide (SEQ ID NO:38)
GGTGGTGGTAGCACACCCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACC GTTGTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTGGTCTGGATCCGCAGGCACGCTGGATATGGGTTATAATAGCTATGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTAGGCAGCATTGATCCGGCAACAGCAGCAAAATCCGGCAAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCGGTTAATAGCGCAGGTACACCCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAAATAGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGAACCCAGCACCGATCTGGCAGCCTTTCTGAT
    
```

도면2b

```

CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTCATGGTGAACGTGGTGAAGCACCAGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTATTTTGGTGTGCCGCACGTACCCGTCCGCTGTGTCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_Y197V_Nucleotide (SEQ ID NO:39)
GGTGGTGGTAGCACACCCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAGTGGTGTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCCGTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCAGATTGCAAGCAGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGCATGATAATGCCGTTAATGATAATCCGGATGCAGTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCTGTCAGCAGATAAAAGC
TATTTTATCGTTGTAGCGAAGGTGGTCTGGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCCGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAAGTTT
ACCCAGATGACATGGATTGGCAGTGGTGCACCGCAGCAGCAGTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTGCATGGTGAACGTGGTGAAGCACCAGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTATTTTGGTGTGCCGCACGTACCCGTCCGCTGTGTCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_T68V_Nucleotide (SEQ ID NO:40)
GGTGGTGGTAGCACACCCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCCGTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCAGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCAGTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCTGTCAGCAGATAAAAGC
TATTTTATCGTTGTAGCGAAGGTGGTCTGGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCCGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
    
```

도면2c

```

TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTGATATTGACCCGCTGAAAATTTGGGCAACCGCGGTAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCGCAGCACCGATCTGGCAGCCTTTCTGTAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTGAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTATTGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGTTTTGCACGT
CTGTTTTCTGGTTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGGTCCGCACGTACCCGTCCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_A78P_Nucleotide (SEQ ID NO:41)
GGTGGTGGTAGCACACCGCTGCCGTCGCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTTCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATCCGGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAGC
TATTTTATCGGTTGTAGCGAAGGTGGTGTGTAAGGTATGATGCTGAGCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACAGAGCCTGGCACCGGCAGCCGTTGGCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGTATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTTGATCCGGCAACAGCAGCAAATCCGGCAAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTAACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTGATATTGACCCGCTGAAAATTTGGGCAACCGCGGTAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCGCAGCACCGATCTGGCAGCCTTTCTGTAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTGAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTATTGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGTTTTGCACGT
CTGTTTTCTGGTTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGGTCCGCACGTACCCGTCCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_E110A_Nucleotide (SEQ ID NO:42)
GGTGGTGGTAGCACACCGCTGCCGTCGCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTTCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGCGATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCTGGATCAG

```

도면2d

GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCGTGCAGCAGATAAAAGC
TATTTTATCGGTTGTAGCGAAGGTGGTCGTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCGGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATGATTTGACCCGCTGAAAATTTGGGCAACCAGCGGTGAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTGAACCTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGCTCCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGATGT
GCAGCACCTCCGTAA

>MHETase_Consensus95_197_M117L_Nucleotide (SEQ ID NO:43)

GGTGGTGGTAGCACACCCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCAGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCCTGTTGTTGAAGTTCAGCATGGCGTGCAGCA
CCGGCAACCAGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTCTGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGACGGCAGCTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCTGCAGCAGATAAAAGC
TATTTTATCGGTTGTAGCGAAGGTGGTCGTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCGGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATGATTTGACCCGCTGAAAATTTGGGCAACCAGCGGTGAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTGAACCTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGCTCCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGATGT
GCAGCACCTCCGTAA

>MHETase_Consensus95_197_E128Q_Nucleotide (SEQ ID NO:44)

GGTGGTGGTAGCACACCCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCAGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT

도면2e

GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
 CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
 AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTTCGTCTGCGTATGCCTGCA
 GAATGGAATGGTCGCTTTTTATGCAGGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
 GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
 ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
 GTTGCATTTGGTCTGGATCCGCAGGCACGCTCGGATATGGGTTATAATAGCGTGGATCAG
 GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCTGCAGCAGATAAAAAGC
 TATTTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
 AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
 AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
 GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
 CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
 TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
 TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
 GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGATAATCGTTGGGCATGGGAT
 GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
 GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
 TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
 ATGATGAAATTCGATTTTATGATTTGACCCGCTGAAAATTTGGGCAACCAGCGGTAGTTT
 ACCCAGAGTAGCATGGATTGGCATGGTGAACAGCACCGATCTGGCAGCCTTTTCGTGAT
 CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCATGGAT
 ACAGCAGATTATTGAACGCTCGGGTGCAGCAATGCCTGGTGCAGCAGTTTTGCACGT
 CTGTTTCTGGTCCGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
 ATGCTGACCCCTCTGTTGCATGGGTTGAACGTGGTGAACACCGGATCAGATTAGCGCA
 TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGCTCCGCTGTGTCCGTAT
 CCGCAGATGCACGTTATAAAGGTAGCGGTGATTAACACCGAAGCCAATTTGTCATGT
 GCAGCACCTCCGTAA

>MHETase_Consensus95_197_S131G_Nucleotide (SEQ ID NO:45)
 GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
 CCGCTCCGGTCCGCTGGCAAGCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
 GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
 CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
 AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTTCGTCTGCGTATGCCTGCA
 GAATGGAATGGTCGCTTTTTATGGAAGGTGGTGGCGGCACCAATGGTAGCCTGAGCGCA
 GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
 ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
 GTTGCATTTGGTCTGGATCCGCAGGCACGCTCGGATATGGGTTATAATAGCGTGGATCAG
 GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCTGCAGCAGATAAAAAGC
 TATTTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
 AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
 AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
 GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
 CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
 TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
 TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
 GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGATAATCGTTGGGCATGGGAT
 GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
 GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
 TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
 ATGATGAAATTCGATTTTATGATTTGACCCGCTGAAAATTTGGGCAACCAGCGGTAGTTT
 ACCCAGAGTAGCATGGATTGGCATGGTGAACAGCACCGATCTGGCAGCCTTTTCGTGAT
 CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCATGGAT
 ACAGCAGATTATTGAACGCTCGGGTGCAGCAATGCCTGGTGCAGCAGTTTTGCACGT
 CTGTTTCTGGTCCGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT

도면2f

```

ATGCTGACCCCTCGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTTGGTGTTCGCCGACGTACCCGTCCTGTCCTGAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_N134D_Nucleotide (SEQ ID NO:46)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTTCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTTGGCCGAATGCAGCAACCCTGTTGTAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCAATTTGGTCTGGATCCGCAGGCACGCTCGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAAGC
TATTTTATCGGTTGAGCGAAGGTGGTCTGTAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACACACAGAGCCTGGCACCAGCGCCTGGCCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTCTGAGTCCGGTTAGGTGACCGCAATTAACGC
GCAATGGCAGGTCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATGATTTGACCCGCTGAAAATTTGGGCAACCAGCGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCACCAGCACCGATCTGGCAGCCTTTCTGTGAT
CGTGGTGGTAAAATGATTTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTATGAACGCTCGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTTGGTGTTCGCCGACGTACCCGTCCTGTCCTGAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_G156N_Nucleotide (SEQ ID NO:47)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTTCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTTGGCCGAATGCAGCAACCCTGTTGTAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTAACTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCAATTTGGTCTGGATCCGCAGGCACGCTCGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAAGC
TATTTTATCGGTTGAGCGAAGGTGGTCTGTAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACACACAGAGCCTGGCACCAGCGCCTGGCCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTCTGAGTCCGGTTAGGTGACCGCAATTAACGC
GCAATGGCAGGTCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG

```

도면2g

```

GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGAACCCAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCGTTGCATGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGCTCCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGTCATGT
GCAGCACCTCCGTAA
>MHEase_Consensus95_197_A161G_Nucleotide (SEQ ID NO:48)
GGTGGTGGTAGCACACCCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCGAATGCAGCAACCCTGTTGAAGTTCAGCATGGCGTGATGCAGCA
CCGGCAACCGAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTAGCACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCATGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGCTCGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCAGTTTTTATGGTCTGCAGCAGATAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGTT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTCAGTCCGGTTCAGGTGACCCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGAACCCAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCGTTGCATGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGCTCCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGTCATGT
GCAGCACCTCCGTAA
>MHEase_Consensus95_197_D191L_Nucleotide (SEQ ID NO:49)
GGTGGTGGTAGCACACCCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCGAATGCAGCAACCCTGTTGAAGTTCAGCATGGCGTGATGCAGCA
CCGGCAACCGAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCATGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGCTCGTGTATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCAGTTTTTATGGTCTGCAGCAGATAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT

```

도면2h

```

AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTATTATGAACGCTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCGCACGTACCCGTCGGCTGTGTCGGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
GCAGCACCTCCGTAA

```

>MHETase_Consensus95_197_M192Y_Nucleotide (SEQ ID NO:50)

```

GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAAGTACCCGTTATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATTATGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
TATTTATCGGTTGTAGCGAAGGTGGTCTGGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCAACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTATTATGAACGCTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCGCACGTACCCGTCGGCTGTGTCGGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
GCAGCACCTCCGTAA

```

>MHETase_Consensus95_197_S196A_Nucleotide (SEQ ID NO:51)

```

GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCCGTTATGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA

```

도면2i

```

GAATGGAATGGTCGCTTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGCTGGATATGGGTTATAATGCGGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGC AAAAACCGCAGATTGCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTGCAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGTCAAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGGTGAACAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTATTGAACGCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCGTTGATGGTGAACGTTGAAAGCAGCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGCTCCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGTCATGT
GCAGCACCTCCGTAA
>MHEase_Consensus95_197_G204A_Nucleotide (SEQ ID NO:52)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTCAGCATGGCGTGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATGGC
AAACGTACCGGTATTGATGGTTATCCGATGAAATCAAATTTGCTCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGCTGGATATGGGTTATAATAGCTGGATCAG
GTTACCCAGGCAGCGAAAGCAGCAGTTGCACGTTTTTATGGTCTGTCAGCAGATAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGC AAAAACCGCAGATTGCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTGCAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGTCAAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGGTGAACAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTATTGAACGCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCGTTGATGGTGAACGTTGAAAGCAGCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGCTCCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGTCATGT

```

도면2j

```

GCAGCACCTCCGTAA
>MHETase Consensus95_197_A207L Nucleotide (SEQ ID NO:53)
GGTGGTGGTAGCACACCCGCTGCCGTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAAGATGGTAATGGT
GATATGGTTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCACTGGTTGCACGTTTTATGGTCGTGCAGCAGATAAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCCGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACACAGAGCCTGGCACCGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTATGATGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTTCTGAGTCCGGTTAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTGCAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATGATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCACCAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTATGAACGCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTCTGGTTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGTCGCTGTGTCCGTAT
CCGCAGATCGCAGTTATAAAAGTAGCGGTGATTAACACCGAAGCAATTTTGCATGT
GCAGCACCTCCGTAA
>MHETase Consensus95_197_A216P Nucleotide (SEQ ID NO:54)
GGTGGTGGTAGCACACCCGCTGCCGTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAAGATGGTAATGGT
GATATGGTTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCACTGGTTGCACGTTTTATGGTCGTGCACCGGATAAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCCGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACACAGAGCCTGGCACCGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTATGATGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTTCTGAGTCCGGTTAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTGCAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATGATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
    
```

도면2k

```

ACCCAGAGTAGCATGGATTGGCATGGTGAACAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGGCCGACGTACCCGCTCCGCTGTGTCCGTAT
CCGAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGCATGT
GCAGCACCTCCGTAA
>MHEase_Consensus95_197_E226N_Nucleotide (SEQ ID NO:55)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATGGC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCCGTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGACGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCTGTCAGCAGATAAAAGC
TATTTTATCGGTTGTAGCAACCGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTCTGAGTCCGGTTAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTGTATTTGACCCGCTGAAAATTTGGGCAACCGCGTCAAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGAACAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGTTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGGCCGACGTACCCGCTCCGCTGTGTCCGTAT
CCGAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGCATGT
GCAGCACCTCCGTAA
>MHEase_Consensus95_197_L234A_Nucleotide (SEQ ID NO:56)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATGGC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCCGTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGACGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCTGTCAGCAGATAAAAGC
TATTTTATCGGTTGTAGCGAAGGTGGTCTGTAAGGTATGATGGCGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA

```

도면21

TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
 TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCCGCAATFAAACGC
 GCAATGGCAGGTCCGGTAAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
 GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
 GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
 TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
 ATGATGAAATTCGATTTTATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
 ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCTGTGAT
 CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
 ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
 CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
 ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCAGGATCAGATTAGCGCA
 TGGTCTGGTACTCCGGGTTATTTGGTGTGGCCGACGTACCCGTCGGCTGTGTCCGTAT
 CCGCAGATCGCACGTTATAAAGGTAGCGGTGATTAACACCGAAGCCAATTTTGCATGT
 GCAGCACCTCCGTAA

>MHETase_Consensus95_197_S235A_Nucleotide (SEQ ID NO:57)

GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
 CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
 GATATGGTTTGGCCGAATGCAGCAACCCTTGTGAAGTTGCAGCATGGCGTGATGCAGCA
 CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
 AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCGTGCA
 GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
 GCCACCGGTAGCATTGGTGGTGGCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
 ATTGCAACCGATGGTGGTATGATAATGCCGTTAATGATAATCCGGATGCAGTGGGCACC
 GTTGCATTTGGTCTGGATCCGCAGGCAGTCTGGATATGGGTTATAATAGCGTGGATCAG
 GTTACCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGCAGCAGATAAAAGC
 TATTTTATCGGTTTGAAGGATGGTGGTGGTGAAGGTATGATGCTGGCGCAGCGTTTTCCG
 AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGATT
 AGTGGTGCATGGACCACACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
 GTTCCGCTGATTAACAAAAGCTTTAGTATGATGCCGATCTGCATCTGCTGAGTCAGGCAAT
 CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
 TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
 TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCCGCAATFAAACGC
 GCAATGGCAGGTCCGGTAAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
 GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
 GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
 TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
 ATGATGAAATTCGATTTTATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
 ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCTGTGAT
 CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
 ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
 CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
 ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCAGGATCAGATTAGCGCA
 TGGTCTGGTACTCCGGGTTATTTGGTGTGGCCGACGTACCCGTCGGCTGTGTCCGTAT
 CCGCAGATCGCACGTTATAAAGGTAGCGGTGATTAACACCGAAGCCAATTTTGCATGT
 GCAGCACCTCCGTAA

>MHETase_Consensus95_197_P255V_Nucleotide (SEQ ID NO:58)

GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
 CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
 GATATGGTTTGGCCGAATGCAGCAACCCTTGTGAAGTTGCAGCATGGCGTGATGCAGCA
 CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
 AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCGTGCA
 GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
 GCCACCGGTAGCATTGGTGGTGGCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
 ATTGCAACCGATGGTGGTATGATAATGCCGTTAATGATAATCCGGATGCAGTGGGCACC

도면 2m

GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
 GTTACCCAGGCAGGTAAAGCAGCAGTTGCACGTTTTTATGGTCTGTCAGCAGATAAAAGC
 TATTTTATCGGTTGTAGCGAAGGTGGTCTGTAAGGTATGATGCTGAGCCAGCGTTTTCCG
 AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGGTGAAAGCAGGTATT
 AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
 GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
 CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
 TGTACGGCAGCATTTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
 TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCACAATTAACGC
 GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCCGCTGTATAATCGTTGGGCATGGGAT
 GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
 GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
 TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
 ATGATGAAATTCGATTTTATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTACGTTT
 ACCCAGAGTAGCATGGATTGGCATGGTGAACCCAGCACCAGCTGGCAGCCTTTCTGTGAT
 CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
 ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
 CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
 ATGTCAGCCCTCTGGTTGCATGGTGAACCTGTTGAACCTGGTGAAGCACCAGGATCAGATTAGCGCA
 TGGTCTGGTACTCCGGGTTATTTGGTGTGCGCACGTACCCGCTCCGCTGTGTCCTGAT
 CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
 GCAGCACCTCCGTAA

>MHETase Consensus95_197_G258A_Nucleotide (SEQ ID NO:59)
 GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCAGCCT
 CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
 GATATGGTTTTGGCCGAATGCAGCAACCCTTGTGAAGTTGCAGCATGGCGTGATGCAGCA
 CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
 AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
 GAATGGAATGGTGCCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
 GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
 ATTGCAACCGATGGTGGTCAATGATAATGCCGTTAATGATAATCCGGATGCAGTGGGCACC
 GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
 GTTACCCAGGCAGGTAAAGCAGCAGTTGCACGTTTTTATGGTCTGTCAGCAGATAAAAGC
 TATTTTATCGGTTGTAGCGAAGGTGGTCTGTAAGGTATGATGCTGAGCCAGCGTTTTCCG
 AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGCCGAAAGCAGCGATT
 AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
 GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
 CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
 TGTACGGCAGCATTTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
 TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCACAATTAACGC
 GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCCGCTGTATAATCGTTGGGCATGGGAT
 GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
 GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
 TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
 ATGATGAAATTCGATTTTATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTACGTTT
 ACCCAGAGTAGCATGGATTGGCATGGTGAACCCAGCACCAGCTGGCAGCCTTTCTGTGAT
 CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
 ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
 CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
 ATGTCAGCCCTCTGGTTGCATGGTGAACCTGTTGAACCTGGTGAAGCACCAGGATCAGATTAGCGCA
 TGGTCTGGTACTCCGGGTTATTTGGTGTGCGCACGTACCCGCTCCGCTGTGTCCTGAT
 CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
 GCAGCACCTCCGTAA

>MHETase Consensus95_197_S260A_Nucleotide (SEQ ID NO:60)
 GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCAGCCT

도면2n

```

CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACC GTTGTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTCGTCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCTGCAGCAGATAAAAAGC
TATTTTATCGGTTGTAGCGAAGGTGGTCTGGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
GCGGGTGATGGACACACAGAGCCTGGCACCGGCAGCCGTTGGCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCTGTGCA
TGTAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTAAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTGATATTGACCCGCTGAAAATTTGGGCAACCGCGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGAACAGCAGCAGGATCTGGCAGCCTTTCTGTGAT
CGTGGTGGTAAAATGATCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCCTGGAT
ACAGCAGATTATTATGAACGCTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGTTGAACGTGGTGAAGCAGCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGGCGCACGTACCCGCTCCGCTGTGTCCGTAT
CCGCAGATCGCAGTTATAAAGGTAGCGGTGATTAACACCGAAGCCAATTTTGCATGT
GCAGCACCTCCGTAA

```

```

>MHEase_Consensus95_197_T264L_Nucleotide (SEQ ID NO:61)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACC GTTGTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTCGTCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTGCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTTATGGTCTGCAGCAGATAAAAAGC
TATTTTATCGGTTGTAGCGAAGGTGGTCTGGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGCTGACACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCTGTGCA
TGTAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTAAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATATTGACCCGCTGAAAATTTGGGCAACCGCGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGAACAGCAGCAGGATCTGGCAGCCTTTCTGTGAT
CGTGGTGGTAAAATGATTTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCCTGGAT
ACAGCAGATTATTATGAACGCTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT

```

도면20

```

CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGTCGGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_T265L_Nucleotide (SEQ ID NO:62)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCCGTTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCAATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGTAAAGCAGCAGTTGCACGTTTTTATGGTCTGCAGCAGATAAAAAGC
TATTTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCCTGCAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTCTGAAAGCTTTAGTGTAGTCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTGTATTTGATATTGACCCGCTGAAAATTTGGGCAACCGCGTCA
ATCCAGAGTAGCATGGATTGGCATGGTGCACACCGATCTGGCAGCCTTTCTGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCATGGAT
ACAGCAGATTATTAAGCCTGCTGGGTGCAGCAATGCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTTAATGATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGTCGGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_N284L_Nucleotide (SEQ ID NO:63)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCCGTTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCAATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGTAAAGCAGCAGTTGCACGTTTTTATGGTCTGCAGCAGATAAAAAGC
TATTTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTCTGAAAAGCTTTAGTGTAGTCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT

```

도면2p

GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGTTGGCGTAGTTGGTGGCTG
 GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
 TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
 ATGATGAAATTCGATTTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
 ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCGTGAT
 CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
 ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
 CTGTTTCTGGTTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
 ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
 TGGTCTGGTACTCCGGTTATTTGGTGTGCCGCACGTACCCGTCCGCTGTGTCCGTAT
 CCGCAGATCGCACGTTATAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
 GCAGCACCTCCGTAA

>MHETase_Consensus95_197_L295V_Nucleotide (SEQ ID NO:64)

GGTGGTGGTAGCACACCCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
 CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
 GATATGGTTTGGCCGAATGCAGCAACCCTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
 CCGGCAACCGAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
 AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTTCGTCTGCGTATGCCTGCA
 GAATGGAATGGTCCGCTTTTTATGGAAGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
 GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
 ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
 GTTGCATTTGGTCTGGATCCGCAGGCACGCTGGATATGGGTTATAATAGCGTGGATCAG
 GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGCAGCAGATAAAAAGC
 TATTTTATCGGTTGTAGCGAAGGTGGTGGTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
 AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
 AGTGGTGCATGGACCACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
 GTCCGCTGATTAACAAAAGCTTTAGTATGCGGATCTGCATCTGGTGAAGTCAAGCAATT
 CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
 TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAAGCCCTGCAG
 TGTGTTGGTGCAAAACCCGAGATTGTCTGAGTCCGGTTCAAGGTGACCGCAATTAACGC
 GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGATAAATCGTTGGCATGGGAT
 GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGTTGGCGTAGTTGGTGGCTG
 GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
 TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
 ATGATGAAATTCGATTTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
 ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCGTGAT
 CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
 ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
 CTGTTTCTGGTTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
 ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
 TGGTCTGGTACTCCGGTTATTTGGTGTGCCGCACGTACCCGTCCGCTGTGTCCGTAT
 CCGCAGATCGCACGTTATAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
 GCAGCACCTCCGTAA

>MHETase_Consensus95_197_S296A_Nucleotide (SEQ ID NO:65)

GGTGGTGGTAGCACACCCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
 CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
 GATATGGTTTGGCCGAATGCAGCAACCCTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
 CCGGCAACCGAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
 AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTTCGTCTGCGTATGCCTGCA
 GAATGGAATGGTCCGCTTTTTATGGAAGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
 GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
 ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
 GTTGCATTTGGTCTGGATCCGCAGGCACGCTGGATATGGGTTATAATAGCGTGGATCAG
 GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGCAGCAGATAAAAAGC
 TATTTTATCGGTTGTAGCGAAGGTGGTGGTGAAGGTATGATGCTGAGCCAGCGTTTTCCG

도면2q

```

AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGGCGCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATGATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGTCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATTAACACCGAAGCCAATTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_T355R_Nucleotide (SEQ ID NO:66)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGCTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGCAGCAGATAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGCTGAGTCCGGTTCAGGTGAGAGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTATGATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGTCGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATTAACACCGAAGCCAATTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_A377P_Nucleotide (SEQ ID NO:67)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC

```

도면2r

```

AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTCGTCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGTAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGCCGTGGGAT
GCAGGTATGAGCGGTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGAACAGCACCAGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCCTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCAGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGGCCGACGTACCCGCTCCGCTGTGTCCGAT
CCGAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCATTTTGCATGT
GCAGCACCTCCGTAA
>MHEase_Consensus95_197_S463L_Nucleotide (SEQ ID NO:68)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTGGCCGAATGCAGCAACCCGTTGTTGAAGTTGCAGCATGGCGTATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTCGTCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGTAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGCCATGGGAT
GCAGGTATGAGCGGTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTCTGATGGATTGGCATGGTGAACAGCACCAGATCTGGCAGCCTTTCGTGAT
CGTGGTGGTAAAATGATTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCCTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCAGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGGCCGACGTACCCGCTCCGCTGTGTCCGAT

```

도면2s

```

CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_A493P_Nucleotide (SEQ ID NO:69)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCAGGCAGGTAAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAAGC
TATTTTATCGGTTGAGCAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGCCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTCTGAGTCCGGTTAGGTGACCGCAATTAACCGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTGCAATTTGATATTGACGCGCTGAAAATTTGGGCAACCGCGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGAACACAGCACCGATCTGGCAGCCTTTCTGTGAT
CGTGGTGGTAAAATGATTTCTGTATCATGGTATGTGATGACAGCCTTTAGTGCAGTGGAT
ACAGCAGATTATTATGAACGTCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTTCCGGGTTGATGATGATGACGCGCTGAAAATTTGGGCAACCGCGTCAGTTT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCGGCACGTACCCGTCGCCGTGTGTCGGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_Y503W_Nucleotide (SEQ ID NO:70)
GGTGGTGGTAGCACACCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTCGCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCAGGCAGGTAAAGCAGCAGTTGCACGTTTTATGGTCTGTCAGCAGATAAAAAGC
TATTTTATCGGTTGAGCAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCGGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCGGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTCTGAGTCCGGTTAGGTGACCGCAATTAACCGC
GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGATAATCGTTGGGCATGGGAT
GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC

```

도면2t

```

ATGATGAAATTCGATTTTGATATTGACCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTTCGTGAT
CGTGGTGGTAAAATGATTTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTGGTATGAACGCTGCGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGTCGGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGAAGCCAATTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_E594A_Nucleotide (SEQ ID NO:71)
GGTGGTGGTAGCACACCCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTGCCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTCACGCTTTTTATGGTCTGCAGCAGATAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCCGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT
CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
TGTGTTGGTGCAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
GCAATGGCAGGTCCGGTAAATAGCGCAGGTACACCGCTGATAATCGTTGGGCATGGGAT
CAGGATATGACCGGTCTGACCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
TGGCTGGTTGATTTTGCACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
ATGATGAAATTCGATTTTGATTTGATCCGCTGAAAATTTGGGCAACCAGCGGTCAGTTT
ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTTCGTGAT
CGTGGTGGTAAAATGATTTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
ACAGCAGATTATTATGAACGCTGGGTGCAGCAATGCCTGGTGCAGCAGGTTTTGCACGT
CTGTTTCTGGTCCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
ATGCTGACCCCTCTGGTTGCATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGTCGGCTGTGTCCGTAT
CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACACCGCGGCCAATTTGCATGT
GCAGCACCTCCGTAA
>MHETase_Consensus95_197_N596S_Nucleotide (SEQ ID NO:72)
GGTGGTGGTAGCACACCCGCTGCCGCTGCCTCAGCAGCAGCCTCCGCAGCAAGAACCGCCT
CCGCTCCGGTCCGCTGGCAAGCCGTGCAGCATGTGAAGCACTGAAAGATGGTAATGGT
GATATGGTTTTGGCCGAATGCAGCAACCGTTGTTGAAGTTGCAGCATGGCGTGATGCAGCA
CCGGCAACCGCAAGCGCAGCAGCACTGCCGGAACATTGTGAAGTTAGCGGTGCAATTGCC
AAACGTACCGGTATTGATGGTTATCCGTATGAAATCAAATTTCTGCTGCGTATGCCTGCA
GAATGGAATGGTGCCTTTTTATGGAAGGTGGTAGCGGCACCAATGGTAGCCTGAGCGCA
GCCACCGGTAGCATTGGTGGTGGCCAGATTGCAAGCGCACTGAGCCGTGGTTTTGCAGTT
ATTGCAACCGATGGTGGTCATGATAATGCCGTTAATGATAATCCGGATGCACTGGGCACC
GTTGCATTTGGTCTGGATCCGCAGGCACGTCTGGATATGGGTTATAATAGCGTGGATCAG
GTTACCCAGGCAGGTAAGCAGCAGTTCACGCTTTTTATGGTCTGCAGCAGATAAAAGC
TATTTATCGGTTGTAGCGAAGGTGGTCTGAAGGTATGATGCTGAGCCAGCGTTTTCCG
AGCCATTATGATGGTATTGTTGCCGGTGCACCGGTTATCAGCTGCCGAAAGCAGGTATT
AGTGGTGCATGGACCACACAGAGCCTGGCACCAGCAGCCGTTGGCCTGGATGCACAGGGT
GTTCCGCTGATTAACAAAAGCTTTAGTGATGCCGATCTGCATCTGCTGAGTCAGGCAATT

```

도면2u

CTGGGCACCTGTGATGCCCTGGATGGTCTGGCAGATGGCATTGTTGATAACTATCGTGCA
 TGTCAGGCAGCATTGATCCGGCAACAGCAGCAAATCCGGCAAATGGTCAGGCCCTGCAG
 TGTGTTGGTGCAAAAACCGCAGATTGTCTGAGTCCGGTTCAGGTGACCGCAATTAACGC
 GCAATGGCAGGTCCGGTTAATAGCGCAGGTACACCGCTGTATAATCGTTGGGCATGGGAT
 GCAGGTATGAGCGGTCTGAGCGGTACAACCTATAATCAAGGTTGGCGTAGTTGGTGGCTG
 GGTAGCTTTAATAGCAGCGCAAATAATGCCAGCGTGTAGCGGTTTTAGCGCACGTAGC
 TGGCTGGTTGATTTTGAACCCCTCCGGAACCGATGCCGATGACACAGGTTGCAGCACGC
 ATGATGAAATTCGATTTTGTATTTGACCCCGCTGAAAATTTGGGCAACCAGCGGTAGTTT
 ACCCAGAGTAGCATGGATTGGCATGGTGCAACCAGCACCGATCTGGCAGCCTTTCTGTGAT
 CGTGGTGGTAAATGATTTCTGTATCATGGTATGTCAGATGCAGCCTTTAGTGCACCTGGAT
 ACAGCAGATTATTATGAACGCTCGGGTGACAGCAATGCTGGTGAGCAGGTTTTGCACGT
 CTGTTTCTGGTTCGGGTATGAATCATTGTAGCGGTGGTCCGGGTACAGATCGTTTTGAT
 ATGCTGACCCCTCTGGTTGATGGGTTGAACGTGGTGAAGCACCGGATCAGATTAGCGCA
 TGGTCTGGTACTCCGGGTTATTTGGTGTGCCGCACGTACCCGTCGGCTGTGTCGGTAT
 CCGCAGATCGCACGTTATAAAGGTAGCGGTGATATTAACCCGAAGCCAGCTTTGCATGT
 GCAGCACCTCCGTAA
 >MHETase_Round 1 Consensus A (SEQ ID NO:79)
 GGAGGCGGAAGTACTCCGCTCCCTCCGCGCAGCAGCAACCACACAGCAGGAGCCACCACCG
 CCTCCCGTGCCATTAGCTTACGCGCAGCTTGTGAAGCGCTGAAAGACGGCAACGGCGATATGG
 TTTGGCCGAACGCAGCAACAGTGGTTGAGGTTGCCGATGGCGCGATGCTGCGCCTGTACGG
 CCTCTGCAGCAGCGTTACCCGAACATTGCGAAGTGAAGTGGTCCATTGCGAAACGTACAGGCAT
 CGATGGTTATCCATGAGATCAAATCCGTCTCCGCATGCCCGCAGAATGGAACGGACGTTTTCT
 TTATGGAAGGAGGTTAGGTACAACCGGAAGTCTGAGTGCCGCTACAGGTAGTATTGGTGGTGG
 GCAGATTGCCAGCGCTTTGAGTCTGGGTTTCCGCTGATAGCCACCGACGGTGGTATGACAAT
 GCCGTTAATGACAATCCTGATGCTTGGGTAAGTGTAGCATTGGACTGGACCCGCAGGCAGGCTC
 TGGACATGGGTTACAACCTCTATGATCAGGTACACACAGGCCGGAAGGCGGCTGTGGCTCGCTT
 TTACGGCCGTGCGGCCGATAAATCCTATTTTATTGGCTGATGAGGAGGGCGGGAAGGCATG
 ATGCTGTCACAGCGGTTTCAAGCCACTACGACGGTATTGTGGCAGGTGCCCGGGCTACCAGT
 TACCTAAAGCCGGGATTTCTGGGGCGTGGACAACACAGTCTTTGCCAGCCGCGCTCGGTCT
 CGACGCTCAGGGTGTCCCGCTTATTAATAAATCATTACGCGATGCCGACTTACATCTGCTGAGCC
 AGGCAATTTTAGGTACATGTGATGCGCTGGATGGTCTGGCCGATGGCATAAGTAGACAATTATAG
 AGCCTGTCAAGCTGCCTTTGACCTGCTACCGCAGCCAATCCGGCTAACGGCCAAGCACTGCAA
 TGTGTTGGCGGAAAACCGGGATTGTCTGAGTCCAGTGCAAGTGACTGCGATAAAACGTGCAA
 TGGCGGGCCCGGTCAATAGTGTGGGACACCTTGTATAATCGCTGGGCTGGGACGCAGGAA
 TGTCGGGATTGAGTGGAAACACGTACAACCAGGGATGGCGTTCTGGTGGTGGGCTCGTTCAA
 CAGCAGCGCTAATAACGCTCAACGAGTGTACAGGTTTCAAGTGCAGCAGCTGGCTTGTGGATTTT
 GCCACCCCGCGGAACCAATGCCGATGACTCAGGTTGCCGCGCTATGATGAAGTTTGATTTG
 ATATTGATCCGTTGAAAATTTGGGCCACCTCAGGACAGTTTACGCAAAGTTTCGATGGACTGGCAC
 GGGGCGACATCAACTGATCTGGCAGCCTTTCGGGACAGAGGGGTAATAATGATCCTTTATCATG
 GTATGTCTGATGCCGATTCTCTGCGCTCGATACGGCCGACTACTACGAACGCTTAGGGGCTGC
 TATGCCCTGGTGCAGGCTGGGTTTGCACGTCTCTTTTGGTGCCTGGTATGAATCACTGTTCAAGGAG
 GCCCTGGCACCAGATAGATTTGATATGCTACCCCTCTGGTTGCTTGGGTAGAGAGAGGTGAAGC
 TCCGGATCAAATCAGTGTCTGGTCTGGGACCCCTGGCTACTTCGGTGTCTGCTGCCAGAACCCTG
 CCGCTGTGTCCATATCCCAAGATTGCGCGATATAAAGGGTCTGGGGATTAATACGGAAGCTAA
 TTTTGGCTGCGCGGCCGCCGCTAA
 >MHETase_Round 2 Y503W (SEQ ID NO:80)
 GGCGGTGGGTCAACACCGTTGCCGTTACTCAACAGCAGCCCCCGCAACAGGAACCGCCGCG
 CCACCTGTTCCGTTAGCATCAGTGCAGCATGTGAGGCCCTGAAAGATGGTAATGGCGATATGG
 TTTGGCCAAACCGCGCTACCGTTGTGGAGTGGCTGCTTGGCGTGACGCGGCTCCCGCAACAG
 CGTCTGCAGCCCGCTGCCAGAGCACTGCGAAGTTTCTGGTGAATTGCCAAGCGTACCGGAAT
 TGATGGTTATCCTTATGAAATTAAGTTCCTGCTGCGCATGCCCGCTGAATGGAATGGACGTTTTT
 TCATGGAGGGTGGCTCGGGTACAAATGGTTCTCTCCGCTGCGACAGGCAATCGGAGGCG
 GTCAAATCGCTCAGCGTTGCCAGAGGCTTTGCTGTTATTGCAACGGATGGTGGTATGATAAC
 GCAGTCAATGATAACCTGACGCACCTGGGAACGGTTGCGTTTTGGTTGGATCCTCAGGCAGGTT
 TAGATATGGGGTACAATCTTATGATCAGGTAACCTCAGGCAGGGAAGCCGCGCTTGTCTGTTTT

도면2v

TATGGCCGAGCTGCTGACAAGTCGTA CTTTATTGGGTGCAGCGAAGGTGGGCGAGAAGGCATG
 ATGTTGTCGCAACGATTTCCCACTGATTACGACGGGATTGTGGCGGGTGC GCCAGGTTACCAGC
 TGCCGAAAGCGGGTATCTCGGGTGTGGACAACCCAGAGTTTAGCGCCGGCGGCTGTAGGCT
 TAGATGCACAGGGTGTCCCACTGATTAATAAATCATTTTCTGATGCTGACCTCCATCTGTTATCGC
 AGGCGATCTCGGAACTGCGATGCGCTGGATGGCCTGGCAGACGGTATAGTCGATAATTACAG
 AGCTTGTCAAGCGGGCTTTGACCTGCGACCGCGGGCAACCCAGCGAACGGGCAGGCACTTCA
 ATGTGTTGGTGCAAAAACAGCGGATTGTCTGTCCCGGTACAAGTAACGGCAATCAAGCGTGCA
 ATGGCTGGTCCGGTGAATTCGGCCGGCACACCGCTGTACAATCGGTGGGCGTGGGACGCAGGT
 ATGTCAGGTTTATCCGGAACGACGTACAATCAAGGGTGGAGATCGTGGTGGCTGGGTTCTTTA
 ACTCTAGTGCAAAACAGCGCCAGCGCTCTGGAATTTCCGCGCGTTCCTGGCTGGTGCATTTT
 GCGACGCCACCCGAACCCATGCCGATGACCCAGGTAGCGGCACGCATGATGAAGTTTGATTTG
 ATATCGATCCCTCAAAATTTGGGCCACTAGTGGCCAGTTTACGCACTCAAGTATGGATTGGCAT
 GCGCCACGAGTACAGACTTGGCGGCTTTTCGCGACCGCGGGGTAAGATGATCTGTACCAC
 GGCATGAGCGATGCTGCATTTAGTGCCTCGATACAGCAGACTGGTATGAACGTCTTGGTGCA
 CTATGCCAGGTGCAGCAGGCTTTGCTCGCCTGTTCTTGGTTCCGGGGATGAACATTGCAGCGG
 TGGTCCGGGTACCGACCGCTTCGATATGCTGACGCCATTAGTAGCTTGGGTTGAACGTGGGGAA
 GCCCTGATCAAAATTAGTGCATGGTCCGGGCACACCTGGGATTTCCGGCTGGCGGCCCGGACG
 AGACCGCTTTGCCCTATCCCCAGATTGCCCGGTATAAAGGTAGTGGTGACATAAACACCGAAG
 CGAATTTGCGTGCGCCGCCCGCCCTGA

>MHETase_Round 3 M192Y (SEQ ID NO:81)

GGCGGTGGCTCCACCCCTTACCCTTGCCACAGCAGCAACCGCCGAGCAAGAACCCTCACCCC
 CTCGGTCCCTCGGCGTGCAGCGCAGCGTGC GAAGCCTTGAAGATGGTAATGGCGACATGG
 TTTGGCCAAACGCCGCCACCGTAGTTGAGGTAGCAGCCTGGCGGATGCCGCTCCAGCCACTGC
 CTCGGCCGCGCGCTCCGGAACATTGTGAAGTTAGTGGCGCAATTGCGAAACGCACAGGAATT
 GACGGCTATCCGTATGAGATCAAAATTCGGCTGCGTATGCCGGCAGAAATGGAACGGCCGGTCT
 TTATGGAAGGCGGGAGTGGAAACCAATGGATCCCTGTGAGTGCAGCTGCGACTGGATCCATAGCGGGT
 GGAGATAGCATCGGCTTTAGCGGGGTTTTCGGGTGATCGCGACCGATGGCGGGCATGACA
 ACGCAGTGAACGACAATCTGATGCGTTGGGCACCGTTGCTTTTGGTCTTGATCCGACGGCTCG
 GCTTGATTACGGGTACAATAGTTATGACCAGGTAACACAGGCTGGTAAAGCGGGCGGTGGCCCG
 TTCTATGGCAGAGCAGCCGACAATCGTATTTTATTGGCTGCTGTAAGGCGGTGCGGAAGGAA
 TGATGCTGTCTCAGCGATTTCCGCTCATTATGATGGGATTGTTGCCGGTGCSCCGGGGTATCAA
 CTTCCGAAAGCGGGTATCTCCGGTCTTGGACAACCTCAATCGCTGGCCCCCGCGGGTGGGTC
 TGGATGCTCAAGGCGTCCCTCTCATTAAATAGCTTTTATGATGCTGATTTGCAATTTGCTGTCCC
 AGGCCATTCTGGGAACGTGCGATGCTCTTGACGGTCTCGCCGATGGCATAGTTGACAACATCG
 CGCTGTGACGGCCGATTTGACCCAGCGACCGGGCTAATCCAGCGAATGGTCAGGCACTGCAA
 TGTTGGTGGCGAAAACAGCCGACTGCCTGTACCTGTTGAGGTAACAGCAATTAAGAGCGGA
 TGGCTGGCCAGTTAACTCAGCAGGCAACCCGCTGTATAATCGTTGGCATGGGATGCCGGTAT
 GTCAGGTTTATCCGGAACCTAGTATAATCAGGGATGGCGTTCTTGGTGGCTGGGGTCTTTAACT
 CTAGTGCAAAACATGCACAACGGGTGAGTGGTTTTAGTGCACGGAGCTGGCTTGTGATTTGCA
 AACCCCTCTGAGCCGATGCCATGACACAAGTTGCCGCTCGCATGATGAAATTCGATTTGACA
 TAGATCCGCTGAAGATTTGGGCGACCTCAGGTGAGTTTACCCAGTCTCAATGGATTGGCACGG
 CGCGACCTCGACTGACCTCGCCGATTTCTGACCGTGGAGGCAAAATGATTCTGTATCACGGT
 ATGAGCGATGCTGCCTTTCCGCGTTGGACACAGCGGATTTGGTACGAACGTTTGGGCGCTGCGA
 TGCCGGGCGCCGCGGGTTCGCCGATTATTCTCTGACCTGGAATGAATCACTGTTCCGGTGG
 GCCGGGTACAGACCGTTTCGATATGCTGACCCCGCTTGTAGCCTGGGTCGAGAGAGGGCGAAGC
 CCCGGACCAAAATCTGCGTGGTACGGTACTCCTGGTATTTTGGTGTGCGCCGCTCGTACACGTC
 CATTGTGCTCCGTACCCTCAGATTGCAGTTATAAAGGCGAGCGGCATATTAACACCGAAGCTAAT
 TTCGTTGTGCTGACCCACCTAA

>MHETase_Round 4 G156N (SEQ ID NO:82)

GGAGGTGGTAGCACGCCTTACCTCTGCCGACGAGCAGCCGCCTCAACAGGAACCACCCCG
 CCTCCGGTGCCGTTAGCATCCCGTGCGGCTGCGAAGCCCTGAAGGATGGAACGGTGCATG
 GTATGGCCAAACGAGCTACCGTGGTTCGAAGTTGCCGCTTGGCGGACGCTGCACCAGCCACC
 GCAAGCGCTGCCGCCCTGCCTGAACACTGCGAAGTGTCCGGTGCCTTGAACAAACGCACAGGCA
 TAGATGGTATCCCTATGAGATAAAGTTTCGCTGCGTATGCCTGCCGAGTGGAAATGGGCGTTTT
 TTTATGGAAGGTGGAAGCGGAACCAATGGCAGCCTGAGCGTGCAACAGGTTCAATCGGCGGT
 GGTCAAATTCCTCCGCCCTTCTCGCAATTTGCAAGTATTGCTACGGATGGCGGCCACGATAA
 TGCACTCAACGATAATCCAGACGCTCTGGGGACCGTTGCCCTTGGCCTGGATCCGCAAGCACGC

도면2w

```

TTAGACTATGGTTATAACAGTTACGATCAAGTCACACAAGCAGGTAAGCCGCTGTCGCCCCGCTT
TTATGGGCGTGCCGCCGACAATCTTATTTTATAGGATGCTCTGAAGGCGGACGCGAAGGTATG
ATGCTGTCCCAACGTTTTCCGTCGCACATGATGGAATTGTCGCTGGTGCCTGGTTATCAACT
GCCCAAAGCCGGCATAAGCGGTGCTTGGACGACGCAGTCCCTGGCGCCGCGCAGCAGTTGGATT
GGATGCTCAAGGCGTTCCACTGATCAATAAAAGCTTCAGCGATGCAGACCTGCACCTACTCAGTC
AGGCGATCTTGGGAACCTTGCAGCGCTGGATGGTTTGGCAGATGGAATTGGGATAATTACAG
AGCTTGTACAGGCGGCTTTGATCCAGCTACAGCCGCAAACCCGGCGAATGGACAGGCGCTGCA
GTGCGTAGGGGCCAAAACGGCAGATTGTCTGTCCCGGTTCAAGTAACAGCCATTAAGCGGGCC
ATGGCAGGTCCCCTAATCTGCAGGTACGCCCTGTACAACCGGTGGGCTGGGATGCGGGTA
TGTCCGGCCTCAGCGGAACAATATAACCAAGGTTGGCGATCTTGGTGGCTGGGTAGTTTCAA
CTCGAGTGCAAATAACGCACAGAGAGTGTCCGGTTTTAGTGCTCGGCTCGTTGGTTGATTTTG
CCACACCACCAGAACCTATGCCTATGACCAGGTGGCTGCGCGGATGATGAAGTTTGATTTTGA
TATTGACCCCTTAAGATCTGGGCAACGAGCGGCCAGTTTACACAGAGCAGCATGGATTGGCAT
GGCGCCACCTCTACCGATCTGGCGGCTTTTCTGTGACCGCGCGGAAAAATGATTTTATACCG
GCATGTCCGGATGCAGCCTTCCAGCGCCTTGATACTGCGGATTGGTATGAACGCTTAGGCGCAGC
TATGCTTGAGCTGCTGGTTTTGCGCAGACTGTTCTTGTGCCAGGAATGAATCATTGTAGTGCGC
GTCCCGTACTGATCGCTTGGACATGTTAACGCCCTGTTGGTGGCTGGGTTGAGCGTGGTGAAGC
CCCTGATCAAATCTCAGCGTGGTCTGGCACGCCAGGCTATTTCCGGTGTGACGCTGCACACGT
CCCCCTGTCCATACCCTCAGATAGCAAGATACAAAGGTAGTGGGGACATCAATACCGAGGCAA
ACTTCGCATGCGCGGCCCGCCCTGA
>MHETase_Round 5 Y252F (SEQ ID NO:83)
GGTGGCGGTAGTACCCCGCTGCCGCTGCCGACGAGCAGCCACCGCAACAGGAGCCACCTCCT
CCGCCGGTCCCTTTGGCATCCCGCGCAGCTTGTGAGGCTTTGAAAGACGGTAATGGAGACATGG
TTTGGCCGAATGCTGCGACTGTAGTGAAGTTGCGGCGTGGCGTGTATGCAGCACCCGGCAACTG
CGAGCGCCGCTGCGCTCCCTGAACATTGCGAGGTTAGTGGGGCTATTGCTAAGCGTACCGGGAT
TGATGGCTATCCATACGAGATTAATTTCTGTGCGCATGCCAGCGGAATGGAATGGTAGATTTT
TCATGGAAGGTGGTCCGGCACCAATGGCTCGCTCAGCGCTGCAACAGGCTCCATCGCGGGG
GTCAGATTGCGTACGCTTATCGCGTAATTTTGTGTTATCGCTACCGATGGCGGCCACGATAAT
GCAGTTAACGATAATCCGGACGCTTGGGACTGTTGCGTTTGGCTTAGATCCGCAGGCGCAGAC
TGGATTATGGGTATAATTCGTACGATCAGGTCACACAGGCGAGGTAAGCCGCGCTTGGCCGCTT
TTATGGGCGGGCTGCGGACAATCATATTTTATGGATGTTCCGGAAGGAGGCGGGGAAGGAATG
ATGTTGAGCCAACGATTTCTAGCCATTATGACGGCATTGTTGCCGGGCGCCGGGCTTCCAATT
GCCTAAAGCAGGAATTTAGGTGCGTGGACGACCCAGTCTCTGGCGCTGCGGCAGTAGGACT
GGACGCGCAGGAGTCCCTTTGATTAATAAATCCTTCTCTGATGCTGATTTGCATCTTTGTCCCA
AGCCATTTAGGCACCTGCGATGCGCTGGACGGCTTAGCAGATGGTATAGTTGATAATTATCGCG
CCTGCCAGGCGGCTTTTATCCCGCAACAGCAGCGAACCCTGCGAACGGTACGGCCCTGCAATG
TGTCGGTGCAAAAACCGCGGACTGTTTGTCTCCGGTGCAAGTGACTGCAATTAAGCGTGAATG
GCTGGACCGGTGAACTCCGCTGGAACACCCTTGTACAACCCTGGGCGTGGGATGCTGGCATG
AGCGGCTGTCTGGGACGACTTATAAACAGGGCTGGCGCTTTGGTGGTTAGGTTCAATTAATTC
CTCCGGAATAATGCCAGCGCTATCTGGTTTTCTGCCCGTTCTTGGTTGGTGGATTTTGCTA
CGCCGCCAGAACCGATGCCGATGACGCAAGTCGCGGCACGAATGATGAAGTTGACATTTGACAT
AGACCCGCTGAAAATTTGGGCAACAAGCGGTCAGTTTACGCAAAGCAGCATGGACTGGCACGGA
GCAACCAGCACTGATCTCGCGGCTTTAGAGATCGTGGGGAAAAATGATTTTATATCATGGCAT
GTCAGACGCGCTTTTCCGCTCTGGACACAGCCGACTGGTACGAACGGTGGGAGCAGCTATG
CCAGGGGCGCTGGTTTTGCACGCTTATTTCTGGTGCCAGGTATGAACCATTGTAGTGGGGGCC
CCGGAACCGATCGCTTTGATATGCTTACCCCGCTTGTAGCGTGGGTTGAACGTGGTGAAGCTCC
TGATCAAATTTCTGCTGGAGTGACGCTGGGTATTTCCGGCTGGCGGCGGCTACTCGACCA
CTGTGCCCTTACCCTCAAATTCGCGCTTATAAAGGTAGTGGTATTAACACAGAGGCGAAGCTT
TGCTGTGCTGCCCGCCCTAA
>MHETase_R5-Y192M (SEQ ID NO:84)
GGTGGCAGTACACCGCTTCCGCTTCCGCAACAGCAACCGCCTCAACAGGAACCTCCGCTCCTC
CCGTGCCTCTCGCCAGCGTGCAGCCTGCGAGGCGCTTAAAGACGGGAATGGTGACATGGTAT
GGCCAAACCGCGCCACGGTGGTTGAGGTTGCCGCTGGAGAGATGCTGCTCCGGCAACGGCAT
CCGCCGCTGCGTTGCCAGAACACTGTGAGGTATCGGGCGCCATAGCCAAACGTACCGGCATTGA
TGGCTATCCATACGAAATTAATTTAGACTCCGCATGCCAGCAGAGTGGAACGGACGTTTTTTTA
TGGAGGGTGGTTCTGGTACAAATGGCAGCTTAAAGTCCCGGACTGGCTCCATTGGCGGAGGGC
AGATTGCATCAGCGTTGTCTGTAATTTTGTGTTATCGCCACCGACGGGGGACATGACAACGCC

```

도면2x

GTTAACGATAACCCGTGACGCCCTGGGCACTGTTGCCCTTTGGTCTTGATCCCCAGGCGCGTTTAGA
 TATGGGCTATAATTCTTACGATCAGGTTACACAGGCCGGCAAGGCTGCAGTAGCACGGTTCTAC
 GGTAGAGCCGCAGATAAATCATATTTTATAGTTGCTCTGAGGGTGGGCGTGAAGGTATGATGC
 TTTCTCAGCGCTTCCCTCCATTATGATGGAATAGTGGCTGGTCTCCAGGCTTCCAGCTGCCA
 AAAGCCGGCATCTCAGGCGCATGGACCACGCAAGCCTCGCTCCTGCCGAGTTGGTTTGGATG
 CGCAGGGCGTGCCGTTAATCAATAAGTCTTTCTCTGATGCCGATTTGCACCTGCTGAGCCAAGCA
 ATTTTAGGAAGTTCGATGCCCTGGATGGGTTGGCGGACGGTATAGTCGATAATTATCGCGCAT
 GTCAGGCCGCTTTCGATCCGGCAACCGCAGCTAATCCCGCAATGGCCAGGCGCTGCAGTGTGT
 AGGTGCAAAAACAGCTGACTGTCTCAGTCTGTCCAGGTAAGTCTATTAAGCGCGCATGGCA
 GGTCTGTGAACCTCAGCAGGCACGCCGTTATATAATCGCTGGGCCGTTGGATGCGGGGATGTCC
 GGGCTGTGAGTACCATTATAATCAGGGTTGGAGAAGCTGGTGGCTGGGATCATTAAACAGTA
 GCGCTAACAAATGCACAGCGCTCTCGGATTTAGCGCTCGCTCCTGGTTAGTAGATTTTGTACT
 CCCCAGAAACCGATGCCGATGACGCAGGTTGCTGCTCGAATGATGAAGTTTACTTTGATATCG
 ACCCTCTGAAGATTTGGGCGACAGTGGTCACTTTACACAGTCTCAATGGATTGGCATGGCGC
 TACGCTACGGACTTAGCAGCTTTCCGATCGCGGTGGGAAAATGATACTGTATCACGGCATG
 TCAGATGCGGCGTTTAGCGCGTTGGATACGGCTGATTGGTACGAACGTTTAGGTGCAGCTATGC
 CGGGTGCAGCAGGTTTGCAGCGGCTGTTTTGGTTCCCGGTATGAATCACTGTTCCGGGAGTCC
 GGGTACTGATCGCTTTGATATGCTTACCCCGCTTGTAGCCTGGGTGGAACGTGGAGAGGCCCCG
 GATCAGATCTCCGCTGGAGCGGTACCCCGGGTATTTTGGAGTCGCGGACGGACTCGCCCTC
 TGTGCCATATCTCAGATTGCTAGATATAAAGGTAGCGGCGACATAAATACAGAAGCCAACCTT
 GCATGCGCAGCACCATGA
 > bacterial expression vector pET-28a (+); (SEQ ID NO:85)
 ATCCGGATATAGTTCCTCTTTCAGCAAAAACCCCTCAAGACCCGTTTAGAGGCCCAAGGGGT
 TATGCTAGTTATTGC
 TCAGCGGTGGCAGCAGCCAACTCAGCTTCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGT
 GGTGGTGGTGGTGGTCTC
 GAGTGCAGCCGCAAGCTTGTGACGAGCTCGAATTCGGATCCGCGACCCATTTGCTGTCCACC
 AGTCATGCTAGCCATA
 TGGCTGCCGCGCGCACAGGCCGCTGCTGTGATGATGATGATGATGGCTGCTGCCATGGTAT
 ATCTCTTCTTAAAGT
 TAAACAAAATTATTTAGAGGGGAATTGTTATCCGCTCACAAATCCCTTATAGTGAGTCGTATTA
 ATTTGCGGGGATCG
 AGATCTCGATCCTCTACGCCGGACGCATCGTGCCCGGCATCACCGGCCCCACAGGTGCGGTTG
 CTGGCGCTATATCGCC
 GACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTCGGCGTG
 GGTATGGTGGCAGGCC
 CGTGGCCGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCTTGCAGCGCGCGGTGCT
 CAACGGCCTCAACCTAC
 TACTGGGCTGCTTCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGAGATCCCGACACCATCG
 AATGGCGCAAAACCTT
 TCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACAGTA
 ACGTTATACGATGTCG
 CAGAGTATGCCGGTGTCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCACGTTTC
 TGCGAAAACGCGGGAA
 AAAGTGAAGCGGGCATGGCGGAGCTGAATTACATTCACCCGCGTGGCACAACAACCTGGCG
 GGCAACAGTCGTTGCT
 GATTGGCGTTGCCACCTCAGTCTGGCCCTGCACGCGCCGTCGCAAAATTGTCGCGGCGATTAA
 TCTCGCGCGATCAAC
 TGGGTGCCAGCGTGGTGGTGTGATGGTGAAGCAAGCGGCGTGAAGCCTGTAAGCGGGCGG
 TGCACAATCTTCTCGCG
 CAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAG
 CTGCTGCACTAATGT
 TCCGCGTTATTTCTTGATGTCTTGACCAGACACCCATCAACAGTATTATTTCTCCCATGAAGA
 CGGTACGCGACTGG
 GCGTGGAGCATCTGGTGCATTGGGTACCAGCAATCGCGCTGTTAGCGGGGCCATTAAAGTTC
 TGCTCGGGCGGCTCTG

도면2y

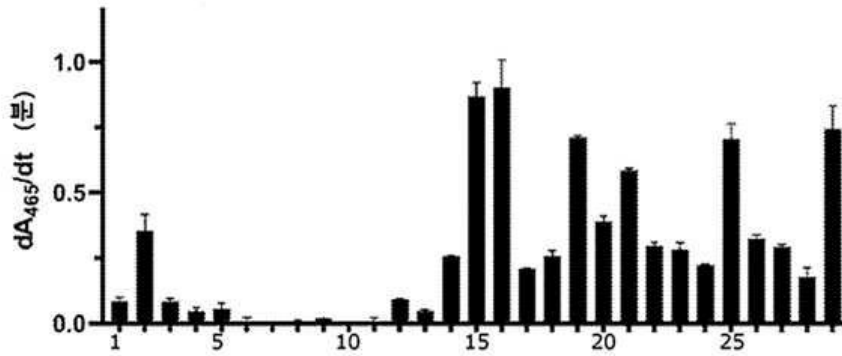
CGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAG
GCGACTGGAGTGCCAT
GTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTCCCCTGCGATGCTGGTTG
CCAACGATCAGATGG
CGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGC GGATATCTCGGTAG
TGGGATACGACGATACC
GAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCA
AACCAGCGTGGACCG
CTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCTCTCACTGGTG
AAAAGAAAACCAACC
TGGCGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACG
ACAGGTTTCCGACTG
GAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAAGTTAGTCACTCATTAGGCACCGGGATCT
CGACCGATGCCCTTGA
GAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTAT
GACTGTCTTCTTATC
ATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCATTTTCGGCGAGGACCGCTTTCGCT
GGAGCGGACGATGAT
CGGCCTGTCGCTTGC GGATTTCGGAATCTTGCACGCCCTCGCTCAAGCCTTCGTCAGTGGTCCC
GCCACCAAACGTTTTG
GCGAGAAGCAGGCCATTATCGCCGGCATGGCGGGCCCCACGGGTGCGCATGATCGTCTCCTGT
CGTTGAGGACCCGGCTA
GGCTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGAGCGAACGTGAAGC
GACTGCTGCTGCAAAAAC
GTCTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTCCGTGTTTCGTAAGTCTGGAAACGC
GGAAGTCAGCGCCCT
GCACCATATGTTCCGGATCTGCATCGCAGGATGCTGCTGGTACCCTGTGGAACACCTACATCT
GTATTAACGAAGCGC
TGGCATTGACCTGAGTGATTTTTCTCTGGTCCCGCCGCATCCATACCGCCAGTTGTTACCCTC
ACAACGTTCCAGTAA
CCGGGCATGTTTATCATCAGTAACCCGATCGTGAGCATCTCTCTGTTTCATCGGTATCATTAC
CCCCATGAACAGAA
ATCCCCCTTACACGGAGGCATCAGTGACCAAAACAGGAAAAAACCGCCCTTAACATGGCCCCGTTT
ATCAGAAGCCAGACA
TTAACGCTTCTGGAGAAAACCAACGAGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGC
TTCACGACCACGCTGA
TGAGCTTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAG
CTCCCGGAGACGGTCA
CAGCTTGTCTGAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGGTCAGCGGGTGTG
GCGGGTGTGCGGGCGCA
GCCATGACCCAGTCACGTAGCGATAGCGGAGTGATACTGGCTTAACTATGCGGCATCAGAGCA
GATTGTAAGAGAGTG
CACCATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTC
CGCTTCCCTCGCTCACT
GACTCGCTGCGCTCGGTCTGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATAC
GGTTATCCACAGAATC
AGGGGATAACGCGAGGAAAGCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAA
GGCCGCGTTGCTGGCGT
TTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCG
AAACCCGACAGGACTA
TAAAGATAACAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCTGCCGCT
TACCGGATACCTGTC
CGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCG
TGTAGGTCGTTGCT
CCAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA
TCGCTTGAGTCCAAC

도면2z

CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
 ATGTAGGCGGTGCTAC
 AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC
 TGCTGAAGCCAGTTA
 CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCCTGGTAGCGGTGGTTT
 TTTTGTTCGCAAGCAG
 CAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGC
 TCAGTGGAACGAAAA
 CTCACGTTAAGGGATTTTGGTCATGAACAATAAACTGTCTGCTTACATAAACAGTAATACAAGG
 GGTGTTATGAGCCAT
 ATCAACGGGAAACGTCTTGCTCTAGGCCGCGATTAATTCACATGGATGCTGATTTATATGG
 GTATAAATGGGCTCG
 CGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAG
 TTGTTCTGAAACATG
 GCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAACTGGCTGACGGAATT
 TATGCCTCTCCGACC
 ATCAAGCATTTTATCCGTACTCTGATGATGCATGGTTACTCACCCTGCGATCCCCGGGAAAAAC
 AGCATTCCAGGTATT
 AGAAGAATATCCTGATTACAGGTGAAAAATATTGTTGATGCGCTGGCAGTGTTCTGCGCCGGTTGC
 ATTCGATTCCTGTTT
 GTAATTGTCCTTTTAAACAGCGATCGCGTATTTTCGCTCGCTCAGGCGCAATCACGAATGAATAAC
 GGTGGTTGATGCG
 AGTGATTTTGTGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAACT
 TTTGCCATTCTCACC
 GGATTACGTCGTCATCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAAT
 AGGTTGTATTGATG
 TTGGACGAGTCGGAATCGCAGACCATAACAGGATCTTGCCATCTATGGAAGTGCCTCGGTGA
 GTTTTCTCCTTCATTA
 CAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTG
 ATGCTCGATGAGTT
 TTTCTAAGAATTAATTCATGAGCGGATACATTTTGAATGTATTTAGAAAAATAACAATAGGGG
 TTCCGCGCACATTT
 CCCCAGAAAGTGCCACCTAAATTGTAAGCGTTAATATTTTGTAAAATTCGCGTTAAATTTTTGTT
 AAATCAGCTCATT
 TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAAGAATAGACCGAGATAGGGTT
 GAGTGTGTTCCAGT
 TTGGAACAAGAGTCCACTATTAAGAAGCTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATC
 AGGGCGATGGCCAC
 TACGTGAACCATCACCTAATCAAGTTTTTTGGGGTTCGAGGTGCCGTAAGCACTAAATCGGAAC
 CCTAAAGGGAGCCCC
 CGATTTAGAGCTTGACGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAA
 GGAGCGGGCGCTAGGGC
 GCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCTTAATGCGCCGCTA
 CAGGGCGCTCCCAT
 CGCCA

도면3

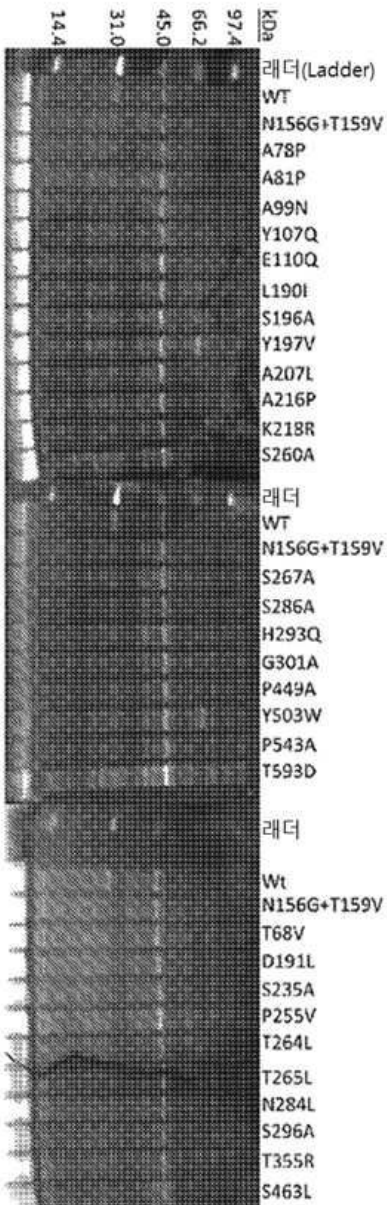
MHETase 컨센서스 설계 및 이들 변이체 중 하나에 대한 점(point) 돌연변이의 세포 현탁액 검정



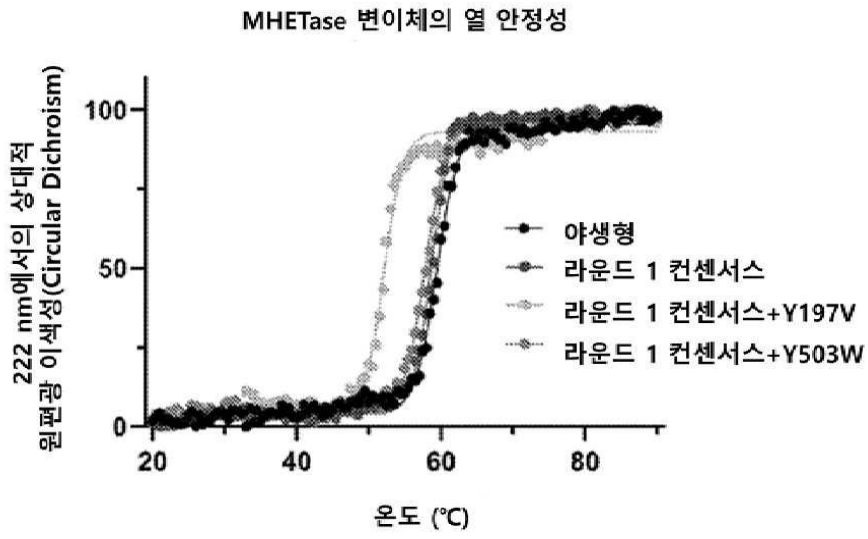
좌측에서 우측:

- 1: WT MHETase
- 2: Round 1 Consensus A
- 3: Round 1 Consensus B
- 4: Round 1 Consensus C
- 5: Round 1 Consensus D
- 6: Round 1 Consensus E
- 7: Round 1 Consensus F
- 8: Round 1 Consensus G
- 9: Round 1 Consensus H
- 10: Round 1 Consensus I
- 11: Round 1 Consensus J
- 12: Round 1 Consensus A + T68V
- 13: Round 1 Consensus A + E110A
- 14: Round 1 Consensus A + D191L
- 15: Round 1 Consensus A + S196A
- 16: Round 1 Consensus A + Y197V
- 17: Round 1 Consensus A + A207L
- 18: Round 1 Consensus A + A216P
- 19: Round 1 Consensus A + S235A
- 20: Round 1 Consensus A + P255V
- 21: Round 1 Consensus A + S260A
- 22: Round 1 Consensus A + T264L
- 23: Round 1 Consensus A + T265L
- 24: Round 1 Consensus A + N284L
- 25: Round 1 Consensus A + S286A
- 26: Round 1 Consensus A + S296A
- 27: Round 1 Consensus A + T355R
- 28: Round 1 Consensus A + S463L
- 29: Round 1 Consensus A + Y503W

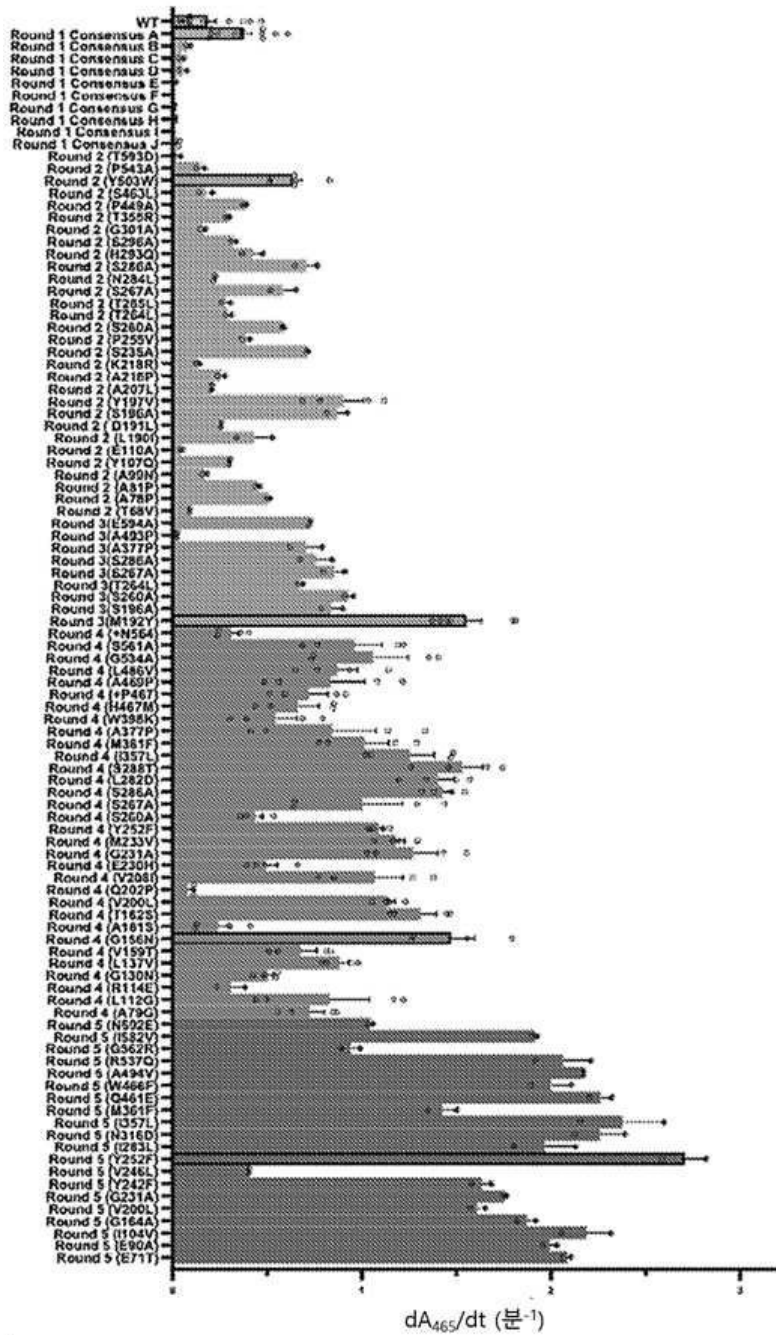
도면4



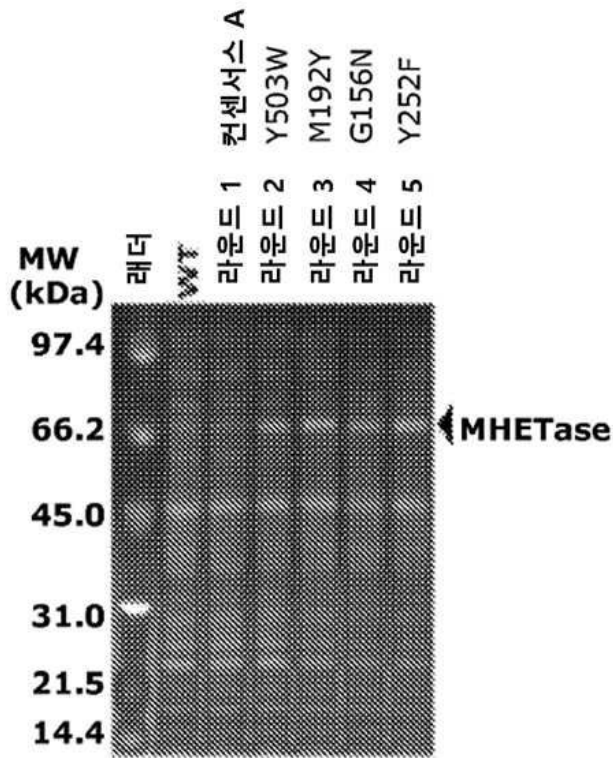
도면5



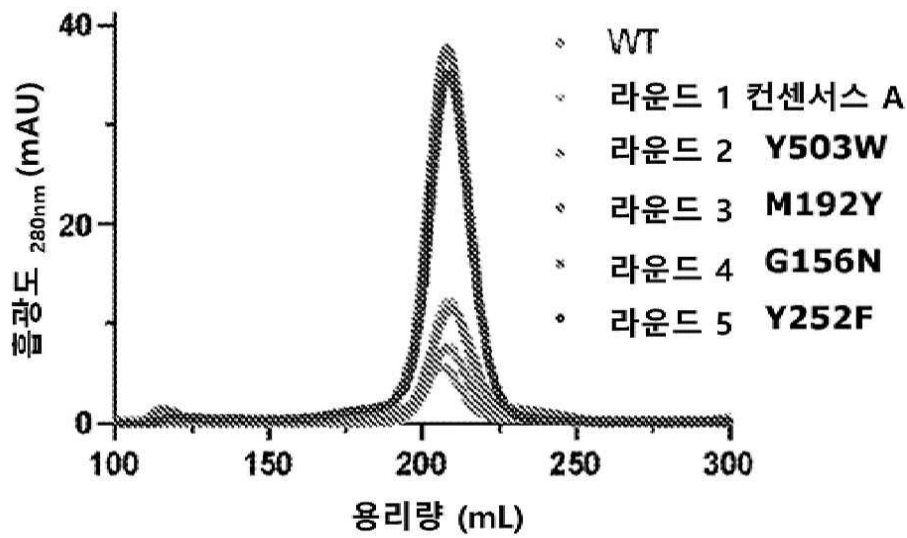
도면6



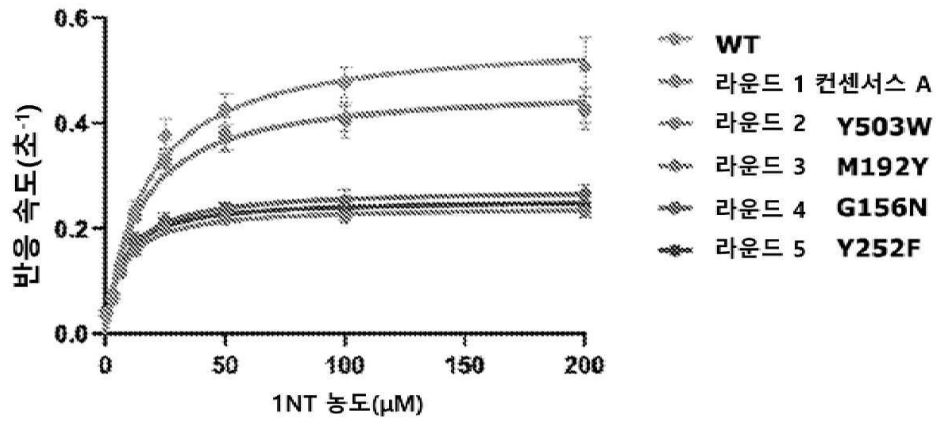
도면7



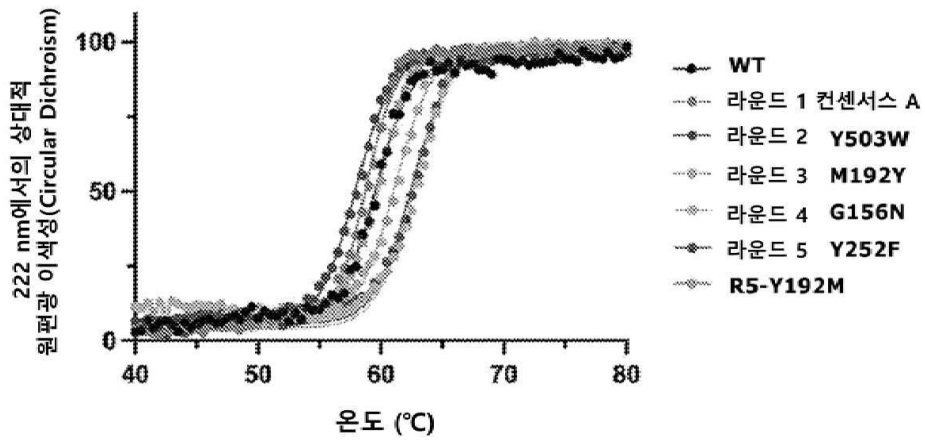
도면8



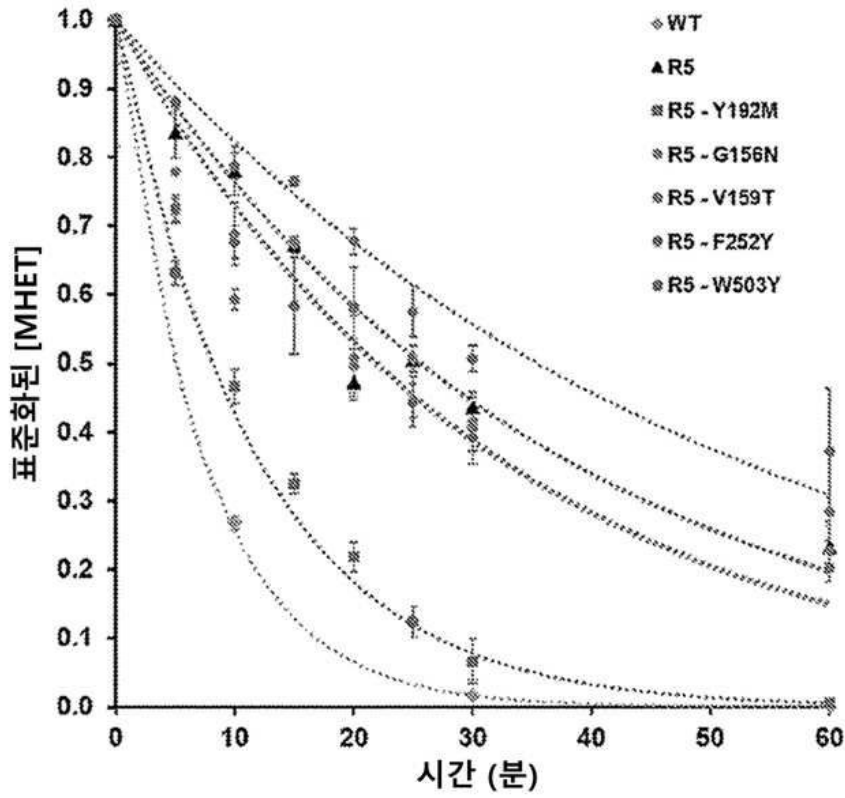
도면9



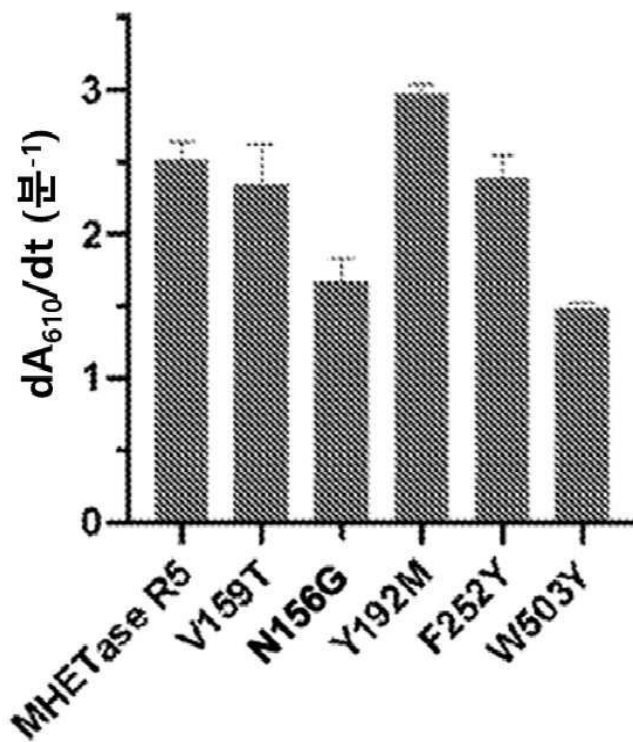
도면10



도면11



도면12



서열목록

SEQUENCE LISTING

<110> Samsara Recycling Pty Limited

<120> ENZYME VARIANTS AND USES THEREOF

<130> 35576771/AXT

<150> 2021901431

<151> 2021-05-13

<160> 85

<170> PatentIn version 3.5

<210> 1

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 1

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80

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

85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly

115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Asn Phe Ala Thr Ile Ala Thr Asp

130 135 140
 Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr
 145 150 155 160

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

 Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240
 Ser Gly Ala Trp Thr Thr Gln Ser Leu Ala Pro Ala Ala Val Gly Leu
 245 250 255
 Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270
 Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

 Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300
 Phe Asp Pro Ala Thr Ala Ala Asn Pro Ala Asn Gly Gln Ala Leu Gln
 305 310 315 320
 Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335
 Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

 Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365
 Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 2

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 2

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala
 65 70 75 80

Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser
 100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp
 130 135 140

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

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Tyr Asn
 165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe
 180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly
 195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp
 210 215 220

Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240

Ser Gly Ala Trp Thr Thr Gln Ser Leu Ala Pro Ala Ala Val Gly Leu

245 250 255
 Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 3
 <211> 584
 <212> PRT
 <213> Artificial Sequence

<220><223> Synthetic
 <400> 3

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala
 65 70 75 80

Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser
 100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp
 130 135 140

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

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn
 165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe
 180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly
 195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp
 210 215 220

Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240

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

Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly

355 360 365
 Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 4

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 4

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

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

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

 Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe
 180 185 190
 Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly
 195 200 205
 Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp
 210 215 220

Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240

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

Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp

465 470 475 480
 Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

 Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510
 Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525
 Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540
 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575
 Asn Phe Ala Cys Ala Ala Pro Pro
 580
 <210> 5
 <211> 584
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 5
 Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15
 Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

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

	85	90	95
Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser			
	100	105	110
Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly			
	115	120	125
Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp			
	130	135	140
Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr			
145	150	155	160
Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn			
	165	170	175
Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe			
	180	185	190
Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly			
	195	200	205
Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp			
	210	215	220
Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile			
225	230	235	240
Ser Gly Ala Trp Thr Thr Gln Ser Leu Ala Pro Ala Ala Val Gly Leu			
	245	250	255
Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp			
	260	265	270
Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp			
	275	280	285
Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala			
	290	295	300
Phe Asp Pro Ala Thr Ala Ala Asn Pro Ala Asn Gly Gln Ala Leu Gln			
305	310	315	320
Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr			
	325	330	335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro

580

<210> 6

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 6

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80

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

85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly

115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp

130 135 140

Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr

145 150 155 160

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn

165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe

180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly

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

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

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

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

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 7

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 7

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

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

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

 Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400
 Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415
 Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430
 Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

 Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525
 Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540
 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr

545 550 555 560
 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 8

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 8

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

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

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125
 Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp
 130 135 140
 Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr
 145 150 155 160
 Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn

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

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 9

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 9

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp

515 520 525
 Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560
 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 10

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 10

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala
 65 70 75 80

Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser
 100 105 110

Gly Thr Asp Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp

130 135 140

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

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

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

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

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 11

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 11

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

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

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

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

245	250	255	
Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp			
260	265	270	
Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp			
275	280	285	
Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala			
290	295	300	
Phe Asp Pro Ala Thr Ala Ala Asn Pro Ala Asn Gly Gln Ala Leu Gln			
305	310	315	320
Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr			
325	330	335	
Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro			
340	345	350	
Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly			
355	360	365	
Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn			
370	375	380	
Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser			
385	390	395	400
Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln			
405	410	415	
Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys			
420	425	430	
Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His			
435	440	445	
Gly Ala Thr Ser Thr Asp Leu Ala Ala Phe Arg Asp Arg Gly Gly Lys			
450	455	460	
Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp			
465	470	475	480
Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala			
485	490	495	

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 12

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 12

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala
 65 70 75 80

Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

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

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 13

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 13

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80

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

85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly

115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp

130 135 140

Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr

145 150 155 160

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Leu Met Gly Val Asn

165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe

180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly

195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 14

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 14

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80
 Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95
 Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser
 100 105 110
 Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125
 Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp
 130 135 140

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

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

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala

565 570 575
 Asn Phe Ala Cys Ala Ala Pro Pro
 580
 <210> 15
 <211> 584
 <212> PRT
 <213> Artificial Sequence

 <220><223> Synthetic
 <400> 15
 Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15
 Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30
 Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45
 Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

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

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

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

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 16

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 16

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

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

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr

530 535 540
 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 17

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 17

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15
 Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

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

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

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

 Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240
 Ser Gly Ala Trp Thr Thr Gln Ser Leu Ala Pro Ala Ala Val Gly Leu
 245 250 255
 Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270
 Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

 Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300
 Phe Asp Pro Ala Thr Ala Ala Asn Pro Ala Asn Gly Gln Ala Leu Gln
 305 310 315 320
 Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335
 Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

 Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365
 Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380
 Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 18

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 18

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

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

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

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

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

Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly

500 505 510
 Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540
 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560
 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575
 Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 19
 <211> 584
 <212> PRT
 <213> Artificial Sequence

<220><223> Synthetic

<400> 19
 Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15
 Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30
 Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45
 Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

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

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 20

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 20

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80

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

85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly

115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp

130 135 140

Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr

145 150 155 160

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn

165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe

180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly

195 200 205

Gly Arg Glu Gly Met Met Ala Ser Gln Arg Phe Pro Ser His Tyr Asp

210 215 220

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

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 21

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 21

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala
 65 70 75 80

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

	85	90	95
Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser			
	100	105	110
Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly			
	115	120	125
Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp			
	130	135	140
Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr			
145	150	155	160
Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn			
	165	170	175
Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe			
	180	185	190
Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly			
	195	200	205
Gly Arg Glu Gly Met Met Leu Ala Gln Arg Phe Pro Ser His Tyr Asp			
	210	215	220
Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile			
225	230	235	240
Ser Gly Ala Trp Thr Thr Gln Ser Leu Ala Pro Ala Ala Val Gly Leu			
	245	250	255
Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp			
	260	265	270
Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp			
	275	280	285
Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala			
	290	295	300
Phe Asp Pro Ala Thr Ala Ala Asn Pro Ala Asn Gly Gln Ala Leu Gln			
305	310	315	320
Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr			
	325	330	335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro

580

<210> 22

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 22

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80

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

85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly

115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp

130 135 140

Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr

145 150 155 160

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn

165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe

180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 23

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 23

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

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

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

 Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400
 Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415
 Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430
 Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

 Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525
 Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540
 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr

545 550 555 560
 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 24

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 24

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

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

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125
 Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp
 130 135 140
 Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr
 145 150 155 160
 Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn

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

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 25

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 25

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

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

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp

515 520 525
 Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560
 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 26

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 26

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala
 65 70 75 80

Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser
 100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp

130 135 140

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

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn
 165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe
 180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly
 195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp
 210 215 220

Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240

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

Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 27

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 27

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

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

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

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

245	250	255
Asp Ala Gln Gly Val Pro Leu Ile Leu Lys Ser Phe Ser Asp Ala Asp		
260	265	270
Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp		
275	280	285
Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala		
290	295	300
Phe Asp Pro Ala Thr Ala Ala Asn Pro Ala Asn Gly Gln Ala Leu Gln		
305	310	315
Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr		
325	330	335
Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro		
340	345	350
Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly		
355	360	365
Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn		
370	375	380
Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser		
385	390	395
Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln		
405	410	415
Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys		
420	425	430
Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His		
435	440	445
Gly Ala Thr Ser Thr Asp Leu Ala Ala Phe Arg Asp Arg Gly Gly Lys		
450	455	460
Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp		
465	470	475
Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala		
485	490	495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 28

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 28

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala
 65 70 75 80

Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

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

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 29

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 29

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80

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

 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

 100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly

 115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp

 130 135 140

Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr

145 150 155 160

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn

 165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe

 180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly

 195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 30

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 30

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80

Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser
 100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp
 130 135 140

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

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn
 165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe
 180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly
 195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp
 210 215 220

Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240

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

Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Arg
 325 330 335

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

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

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

 Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540
 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560
 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala

180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly
 195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp
 210 215 220

Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240

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

Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Pro Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 32

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 32

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Leu Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr

530 535 540
 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575
 Asn Phe Ala Cys Ala Ala Pro Pro
 580
 <210> 33
 <211> 584
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 33
 Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15
 Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

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

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

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

 Gly Ile Val Ala Gly Ala Pro Gly Tyr Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240
 Ser Gly Ala Trp Thr Thr Gln Ser Leu Ala Pro Ala Ala Val Gly Leu
 245 250 255
 Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270
 Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

 Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300
 Phe Asp Pro Ala Thr Ala Ala Asn Pro Ala Asn Gly Gln Ala Leu Gln
 305 310 315 320
 Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335
 Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

 Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365
 Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380
 Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Pro Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 34

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 34

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

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

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

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

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

Asp Ala Gln Gly Val Pro Leu Ile Asn Lys Ser Phe Ser Asp Ala Asp
 260 265 270

Leu His Leu Leu Ser Gln Ala Ile Leu Gly Thr Cys Asp Ala Leu Asp
 275 280 285

Gly Leu Ala Asp Gly Ile Val Asp Asn Tyr Arg Ala Cys Gln Ala Ala
 290 295 300

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

Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335

Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350

Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Trp Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly

500 505 510
 Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540
 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560
 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575
 Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 35
 <211> 584
 <212> PRT
 <213> Artificial Sequence

<220><223> Synthetic

<400> 35
 Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15
 Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30
 Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45
 Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

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

Thr Thr Tyr Asn Gln Gly Trp Arg Ser Trp Trp Leu Gly Ser Phe Asn
 370 375 380

Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser
 385 390 395 400

Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415

Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430

Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Ala Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 36

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 36

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80

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

 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

 100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly

 115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp

 130 135 140

Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr

145 150 155 160

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Val Asn

 165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe

 180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly

 195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp

 210 215 220

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

Thr Ala Asp Tyr Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Ser Phe Ala Cys Ala Ala Pro Pro
 580

<210> 37

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 37

ggtagtggtgta gacaccgct gccctgcct cagcagcagc ctccgagca agaaccgct 60

ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tgtaaatggt 120

gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180

ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240

aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgcga 360

gccaccggta gcattgtgtg tggccagatt gcaagcgcac tgagccgtaa ctttgcaacc 420

attgcaaccg atggtgtgca tgataatgcc gttaatgata atccggatgc actgggcacc 480

gttgcatgtt gctctgatcc gcaggcacgt ctggatatgg gttataatag ctatgatcag 540

gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600

tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660

agccattatg atggtattgt tgcgggtgca ccgggttatc agctgccgaa agcaggtatt 720
 agtgggtgcat ggaccacaca gacacctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggctcg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag 960
 tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020

gcaatggcag gtccggttaa tagcgcaggf acaccgctgt ataactgttg ggcatgggat 1080
 gcaggatga gcggtctgag cggtacaacc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccgga ccgatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggta aatgattct gtatcatggt atgtcagatg cagccttag tgcactggat 1440

acagcagatt attatgaacg tctgggtgca gcaatgectg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacce ctctggttgc atgggttgaa cgtggtgaag caccggatca gattagcga 1620
 tggctctggta ctccgggtta ttttgggtgt gccgcacgta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggf gatattaaca ccgaagccaa ttttgcatgt 1740
 gcagcacctc cgtaa 1755

<210> 38

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 38

ggtggtggta gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgcct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcga 360

gccaccgta gcattggtgg tggccagatt gcaagcgac tgagccgtgg ttttgcagtt 420

attgcaaccg atggtggca tgataatgcc gttaatgata atccggatgc actgggcacc 480

gttgcatittg gctctggatcc gcaggcacgt ctggatatgg gttataatag ctatgatcag 540

gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600

tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660

agccattatg atggtattgt tgcgggtgca ccgggttadc agctgccgaa agcaggtatt 720

agtggatgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780

gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840

ctgggcacct gtgatgcctt ggatggtctg gcagatggca ttgttgataa ctatctgca 900

tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag 960

tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020

gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgtt ggcatgggat 1080

gcaggtatga gcggtctgag cggtaacaacc tataatcaag gttggcgtag ttggtggctg 1140

ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200

tggctggttg attttgcaac ccctccgga ccgatgccga tgacacaggt tgcagcacgc 1260

atgatgaaat tcgatttga tattgaccgc ctgaaaattt gggcaaccag cggtcagttt 1320

accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttctgat 1380

cgtggtggtg aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440

acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500

ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560

atgctgacce ctctggttgc atgggttga cgtggtgaag caccggatca gattagcga 1620

tggctggtgta ctccgggtta ttttgggtgt gccgcagta cccgtccgct gtgtccgtat 1680

ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagcaa ttttgcattg 1740

gcagcacctc cgtaa 1755

<210> 39

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 39

ggtggtggtg gcacaccgct gccctgcct cagcagcagc ctccgagca agaaccgcct 60

ccgcctccgg ttccgctggc aagccctgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatatggttt ggccgaatgc agcagtggtt gttgaagttg cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccc gaacattgtg aagtagcgg tgcaattgcc 240

 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgcga 360
 gccaccggta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgagtt 420
 attgcaaccg atggtggca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtcctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggctcg gaaggtatga tgctgagcca gcgttttccg 660

 agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720
 agtgggtgcat ggaccacaca gacacctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggacact gtgatgccct ggatggctctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataactgttg ggcatgggat 1080

 gcaggtatga gcggtctgag cggtacaacc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccgga ccgatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcattgattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggta aaatgattct gtatcatggt atgtcagatg cagccttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgacgt 1500

 ctgtttctgg ttccgggtat gaatcattgt agcggtggtc cgggtacaga tcgttttgat 1560
 atgctgacce ctctggttgc atgggttga cgtggtgaag caccggatca gattagcga 1620
 tggctggtga ctccgggtta ttttgggtt gccgcagta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740
 gcagcacctc cgtaa 1755

<210> 40

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 40

```

ggtaggtgta gcaacaccgt gcccgctgct cagcagcagc ctccgcagca agaaccgcct      60

ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt      120
gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca      180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc      240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca      300
gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgcga      360
gccaccggta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt      420
attgcaaccg atggtggta tgataatgcc gttaatgata atccggatgc actgggcacc      480

gttgcatattg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag      540
gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc      600
tattttatcg gtttagcga aggtggctgt gaaggtatga tgctgagcca gcgttttccg      660
agccattatg atggtattgt tgcgggtgca ccgggttata agctcccga agcaggtatt      720
agtggtgcat ggaccacaca gagectggca ccggcagccg ttggcctgga tgcacagggt      780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt      840
ctgggcacct gtgatgccct ggatggctctg gcagatggca ttgttgataa ctatcgtgca      900

tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag      960
tgtgttggtg caaaaaccg agattgtctg agtccggttc aggtgaccgc aattaaacgc     1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcatgggat     1080
gcaggtatga gcggtctgag cggtaacaac tataatcaag gttggcgtag ttggtggctg     1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc     1200
tggtcgttg attttgcaac ccctccgga ccgatgccga tgacacaggt tgcagcacgc     1260
atgatgaaat tcgattttga tattgaccgc ctgaaaattt gggcaaccag cggtcagttt     1320

accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat     1380
cgtggtggtg aaatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat     1440
acagcagatt attatgaacg tctgggtgca gcaatgcctg gtcagcagg ttttgcacgt     1500
ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat     1560
atgctgacc ctctggttgc atgggttga cgtggtgaag caccggatca gattagcga     1620

```

tggctctggta ctccgggtta ttttgggtgt gccgcacgta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcatgt 1740

 gcagcacctc cgtaa 1755
 <210> 41
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 41
 ggtggtggta gcacaccgt gccgctgcct cagcagcagc ctccgcagca agaaccgcct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tgtaaatggt 120
 gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatccggca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

 gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgc 360
 gccaccggta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
 attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggctgt gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttadc agctgccgaa agcaggtatt 720

 agtggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgcctt ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcatgggat 1080
 gcaggtatga gcggtctgag cggtaacaac tataatcaag gttggcgtag ttggtggctg 1140

 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccggaa ccgatgccga tgacacaggt tgcagcagc 1260
 atgatgaaat tcgattttga tattgacctg ctgaaaattt gggcaaccag cggtcagttt 1320

acccagagta gcatggattg gcatggtgca acccagcaccg atctggcagc ctttcgtgat 1380
 cgtgggtgga aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcggtggtc cgggtacaga tcgttttgat 1560

 atgctgacce ctctggttgc atgggttgaa cgtgggtgaag caccggatca gattagcga 1620
 tggctcggta ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggf gatattaaca ccgaagccaa ttttgcacgt 1740
 gcagcacctc cgtaa 1755
 <210> 42
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 42
 ggtgggtgga gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgcct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120

 gatatggttt ggccgaatgc agcaaccgtt gttgaagtgg cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gcgatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcga 360
 gccaccggta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
 attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540

 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggctgt gaaggatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttadc agctgccgaa agcaggtatt 720
 agtgggtcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gtccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960

 tgtgttgggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaacgc 1020

gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcatgggat 1080
gcaggtatga gcggtctgag cggtaacaac tataatcaag gttggcgtag ttggtggctg 1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
tggctggttg attttgcaac ccctccggaa ccgatgccga tgacacaggt tgcagcacgc 1260
atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380

cgtggtggta aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
ctgtttctgg ttccgggtat gaatcattgt agcggtggtc cgggtacaga tcgttttgat 1560
atgctgacc ctctggttgc atgggttga cgtggtgaag caccggatca gattagcga 1620
tggctggtta ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgtat 1680
ccgcagatcg cacgttataa aggtagcgggt gatattaaca ccgaagccaa ttttgcagtg 1740
gcagcacctc cgtaa 1755

<210> 43

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 43

ggtggtggta gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgcct 60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tctgcctgca 300
gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcga 360

gccaccggtg gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
gttgcathtt gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660
agccattatg atggtattgt tgccggtgca ccgggttatac agctgccgaa agcaggtatt 720

agtggatgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780

 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggctcg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcagggt acaccgctgt ataatcgttg ggcatgggat 1080
 gcaggataga gcggtctgag cgggtacaacc tataatcaag gttggcgtag ttggtagctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200

 tggctggttg attttgcaac ccctccgga ccgatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgatttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtgggtgta aaatgattct gtatcatggt atgtcagatg cagccttttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggc cgggtacaga tcgttttgat 1560
 atgctgacc ctcgtgttc atgggttga cgtggtgaag caccggatca gattagcga 1620

 tggctcggta ctccgggtta ttttgggtt gccgcagta cccgtccgct gtgtccgat 1680
 ccgcagatcg cagttataa aggtagcgggt gatattaaca ccgaagccaa ttttgcattg 1740
 gcagcacctc cgtaa 1755

 <210> 44
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 44

 ggtgggtgta gcacaccgct gccgctgect cagcagcagc ctccgcagca agaaccgct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatatggttt ggccgaatgc agcaaccgtt gttgaagtgg cagcatggcg tgatgcagca 180

 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatgcagggt gtagcggca ccaatggtag cctgagcgea 360
 gccaccggtg gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420

attgcaaccg atggtggcca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600

tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgccgggtgca ccgggttadc agctccgaa agcaggtatt 720
 agtggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggctctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020

gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcatgggat 1080
 gcaggatga gcggtctgag cggtaacaac tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccgaa ccgatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgattttga tattgacctg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtgggtgta aaatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440

acagcagatt attatgaacg tctgggtgca gcaatgcctg gtcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggc cgggtacaga tcgttttgat 1560
 atgctgacct ctctggttgc atgggttgaa cgtggtgaag caccggatca gattagcga 1620
 tggctggtga ctccgggtta ttttgggtgt gccgcagta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcgggt gatattaaca ccgaagccaa ttttgcattg 1740
 gcagcacctc cgtaa 1755

<210> 45

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 45

ggtgggtgta gcacaccgct gccgctgct cagcagcagc ctccgagca agaaccgct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120

gatatggttt ggccgaatgc agcaaccgtt gttgaagttag cagcatggcg tgatgcagca 180
ccggcaaccg caagcgcagc agcactgccc gaacattgtg aagtagcgg tgcaattgcc 240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
gaatggaatg gtcgcttttt tatggaaggt ggtggcggca ccaatggtag cctgagcgcga 360
gccaccgcta gcattggtag tggccagatt gcaagcgcac tgagccgtgg ttttcagtt 420

attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
gttgcatattg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
tattttatcg gtttagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660
agccattatg atggtattgt tgcgggtgca ccgggttacc agctgccgaa agcaggtatt 720
agtgtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840

ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgag 960
tgtgttgggt caaaaaccg agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgtt ggcattggat 1080
gcaggtatga gcggtctgag cggtacaacc tataatcaag gttggcgtag ttggtggctg 1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc 1200
tggtctggtg attttgcaac cctccggaa ccgatgccga tgacacaggt tgcagcagc 1260

atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
cgtggtggtg aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttcacgt 1500
ctgtttctgg ttccgggtat gaatcattgt agcggtagtc cgggtacaga tcgttttgat 1560
atgctgacce ctctggttgc atgggtttaa cgtggtgaag caccggatca gattagcga 1620
tggctggtg ctccgggtta ttttgggtt gccgcagta cccgtccgct gtgtccgtat 1680

ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740
gcagcacctc cgtaa 1755

<210> 46

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 46

```

ggtggtgta gcacaccgt gccgctgcct cagcagcagc ctccgcagca agaaccgcct      60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt      120
gatatggttt ggccgaatgc agcaaccgtt gttgaagtgg cagcatggcg tgatgcagca      180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc      240

aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca      300
gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccgatggtag cctgagcgcga      360
gccaccggta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgagtt      420
attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc      480
gttgcaattg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag      540
gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc      600
tattttatcg gttgtagcga aggtggtcgt gaaggatga tgctgagcca gcgttttccg      660

agccattatg atggtattgt tgcgggtgca ccgggttadc agctgccgaa agcaggtatt      720
agtggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt      780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt      840
ctgggcacct gtgatgcctt ggatggtctg gcagatggca ttgtgataa ctatctgca      900
tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag      960
tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc     1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg gcatgggat     1080

gcaggtatga gcggtctgag cgttacaacc tataatcaag gttggcgtag ttggtggctg     1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc     1200
tggtctggtt attttgcaac cctccggaa ccgatgccga tgacacaggt tgcagcacgc     1260
atgatgaaat tegatttga tattgaccgg ctgaaaattt gggcaaccag cggtcagttt     1320
accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat     1380
cgtggtggtg aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat     1440
acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt     1500

ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat     1560
atgctgacce ctctggttgc atgggttgaa cgtggtgaag caccggatca gattagcgea     1620
tggtctggta ctccgggtta ttttgggtgt gccgcagta cccgtccgct gtgtccgtat     1680

```

ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740
gcagcacctc cgttaa 1755
<210> 47
<211> 1755
<212> DNA
<213> Artificial Sequence
<220><223> Synthetic
<400> 47
ggtaggtgta gcacaccgct gccgctgctt cagcagcagc ctccgcagca agaaccgct 60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
gatatggttt ggccgaatgc agcaaccgtt gttgaagtgg cagcatggcg tgatgcagca 180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
aacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgctg tatgcctgca 300
gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgcga 360
gccaccgcta gcattggtgg tggccagatt gcaagcgcac tgagccgtaa ctttgcagtt 420
attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660
agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720
agtgtgcat ggaccacaca gacctggca ccggcagccg ttggcctgga tgcacagggt 780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
ctgggcacct gtgatgcctt ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag 960
tgtgttggtg caaaaaccg agattgtctg agtccggttc aggtgaccgc aattaaaccg 1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataactgttg ggcattggat 1080
gcaggtatga gcggtctgag cggtacaacc tataatcaag gttggcgtag ttggtggctg 1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc 1200
tggctggttg attttgcaac ccctccgaa ccgatgccga tgacacaggt tgcagcacgc 1260
atgatgaaat tcgatttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
accagagta gcatggattg gcatggtgca accagaccg atctggcagc ctttctgat 1380

cgtgggtgga aaatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggc cgggtacaga tcgttttgat 1560
 atgctgacce ctctggttgc atgggttgaa cgtgggtgaag caccggatca gattagcgca 1620
 tggctggtga ctccgggtta ttttgggtgt gccgcacgta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcacgt 1740

gcagcacctc cgtaa 1755

<210> 48

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 48

ggtgggtgga gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatatggttt ggccgaatgc agcaaccgtt gtfgaagttg cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagtagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgcga 360
 gccaccggtg gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
 attagcaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtcctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720

agtggatgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgcctt ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccgggtta tagcgcaggt acaccgctgt ataacgttg ggcattggat 1080

gcaggtatga gcggtctgag cggtagaacc tataatcaag gttggcgtag ttggtggctg 1140

ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200

tggctggttg attttgcaac ccctccgga ccatgcccga tgacacaggt tgcagcagc 1260

atgatgaaat tcgattttga tattgacctg ctgaaaattt gggcaaccag cggtcagttt 1320

accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380

cgtgggtgta aaatgattct gtatcatggt atgtcagatg cagccttttag tgcactggat 1440

acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgacgt 1500

ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560

atgctgacct ctctggttgc atgggttga cgtgggtgaag caccggatca gattagcga 1620

tggtctggta ctccgggtta ttttgggtt gccgcagta cccgtccgct gtgtccgat 1680

ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740

gcagcacctc cgtaa 1755

<210> 49

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 49

ggtgggtgta gcacaccgt gccctgcct cagcagcagc ctccgcagca agaaccgct 60

ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120

gatatggttt ggccgaatgc agcaaccgtt gttgaagtgt cagcatggcg tgatgcagca 180

ccggcaaccg caagcgcagc agcaactgcc gaacattgtg aagttagcgg tgcaattgcc 240

aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcga 360

gccaccgta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgagtt 420

attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480

gttgcatttg gtctggatcc gcaggcacgt ctgctgatgg gttataatag cgtggatcag 540

gttaccagcagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600

tattttatcg gttgtagcga aggtggctgt gaaggtatga tgctgagcca gcgttttccg 660

agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720

agtgggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780

gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
ctgggcacct gtgatgcctt ggatggctctg gcagatggca ttgttgataa ctatcgtgca 900
tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag 960

tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcattggat 1080
gcaggataga gcggtctgag cggtaacaacc tataatcaag gttggcgtag ttggtggctg 1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
tggctggttg attttgcaac ccctccggaa ccgatgccga tgacacaggt tgcagcacgc 1260
atgatgaaat tcgattttga tattgaccgc ctgaaaattt gggcaaccag cggtcagttt 1320
accagagta gcatggattg gcatggtgca accagaccgc atctggcagc ctttcgtgat 1380

cggtgggtga aaatgattct gtatcatggt atgtcagatg cagccttttag tgcactggat 1440
acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
ctgtttctgg ttccgggtat gaatcattgt agcgggtggc cgggtacaga tcgttttgat 1560
atgctgacce ctctggttgc atgggtttaa cgtggtgaag caccggatca gattagcga 1620
tggctctgga ctccgggtta ttttgggtgt gccgcacgta cccgtccgct gtgtccgtat 1680
ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740
gcagcacctc cgtaa 1755

<210> 50

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 50

ggtggtgga gcaaccgct gccctgcct cagcagcagc ctccgcagca agaaccgct 60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
gatattggtt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcga 360

gccaccgga gcatggtggc tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480

gttgcatTTg gtctggatcc gcaggcacgt ctggattatg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggctcg gaaggatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttattc agctgccgaa agcaggtatt 720
 agtgggtgcat ggaccacaca gacccctggca ccggcagccg ttggcctgga tgcacagggt 780

gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggctcg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggc acaccgctgt ataactgttg ggcatgggat 1080
 gcaggtatga gcggctctgag cggatcaacc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200

tggctggttg attttgcaac ccctccgga ccatgcccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tegatTTTga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggta aaatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
 acagcagatt attatgaac tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacc ctctggttgc atgggttgaa cgtggtgaag caccggatca gattagcga 1620

tggtctggta ctccgggtta ttttgggtt gccgcagta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cagttataa aggtagcggc gatattaaca ccgaagcaa ttttgcattg 1740
 gcagcacctc cgtaa 1755

<210> 51
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 51

ggtggtggta gcacaccgct gccctgcct cagcagcagc ctccgagca agaaccgct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatattggtt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180

ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgca 360
 gccaccggtg gcattgggtg tggccagatt gcaagcgcac tgagccgtgg ttttgagtt 420
 attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtcctggatcc gcaggcacgt ctggatatgg gttataatgc ggtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600

tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttadc agctgccgaa agcaggtatt 720
 agtgggtgat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccctgta ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgcctt ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag 960
 tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020

gcaatggcag gtccggttaa tagcgcaggc acaccgctgt ataatcgtt ggcatgggat 1080
 gcaggataga gcggtctgag cggtaacaacc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccgaa ccgatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgatttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggta aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440

acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacce ctctggttgc atgggttgaa cgtggtgaag caccggatca gattagcgca 1620
 tggctggtgta ctccgggtta ttttggtgtt gccgcagta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggc gatattaaca ccgaagccaa ttttgcatgt 1740
 gcagcacctc cgtaa 1755

<210> 52

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 52

```

ggtggtgta gcacaccgt gccctgcct cagcagcagc ctccgcagca agaaccgcct      60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt    120
gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca    180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc    240
aacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca    300
gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgc    360
gccaccgcta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt   420

attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc    480
gttgcathtt gctcggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag    540
gttaccagg cagcgaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc    600
tattttatcg gttgtagcga aggtggctgt gaaggtatga tgctgagcca gcgttttccg   660
agccattatg atggtattgt tgcgggtgca ccgggttacc agctccgaa agcaggtatt    720
agtggatgat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt    780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt   840

ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca    900
tgtcaggcag cattgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag    960
tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc  1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcatgggat  1080
gcaggtatga gcggtctgag cggtagaacc tataatcaag gttggcgtag ttggtggctg  1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc  1200
tggtggttg attttgaac ccctccgaa ccgatgccga tgacacaggt tgcagcacgc    1260

atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt   1320
accagagta gcattgattg gcatggtgca accagcaccg atctggcagc ctttcgtgat   1380
cgtggtggtg aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat   1440
acagcagatt attatgaac tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt   1500
ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat  1560
atgctgacc ctctggttgc atgggtttaa cgtggtgaag caccggatca gattagcga    1620
tggctggtg ctccgggtta ttttgggtt gccgcagta cccgtccgct gtgtccgtat    1680

ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagcaa ttttgcagtt   1740

```

gcagcacctc cgtaa 1755

<210> 53

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 53

ggtggtggtgta gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgct 60

ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120

gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180

ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240

aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgcga 360

gccaccggtg gcatgtggggc tggccagatt gcaagcgcac tgagccgtgg ttttgagtt 420

attgcaaccg atggtggca tgataatgcc gttaatgata atccgatgc actgggcacc 480

gttgcatctg gtcctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540

gttaccagg caggtaaagc actggttgca cgtttttatg gtcgtgcagc agataaaagc 600

tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660

agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720

agtgggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780

gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840

ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900

tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960

tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020

gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcattggat 1080

gcaggtatga gcggtctgag cggtaacaacc tataatcaag gttggcgtag ttggtggctg 1140

ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200

tggtcgtttg attttgcaac ccctccgga cccatgccga tgacacaggt tgcagcacgc 1260

atgatgaaat tcgattttga tattgaccgc ctgaaaattt gggcaaccag cggtcagttt 1320

accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380

cggtggtgta aaatgattct gtatcatggt atgtcagatg cagccttttag tgcactggat 1440

acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500

ctgtttctgg ttccgggtat gaatcattgt agcgggtggc cgggtacaga tcgttttgat 1560

atgctgacce ctctggttgc atgggttga cgtgggtgaag caccggatca gattagcgc 1620

tggctctgga ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgcat 1680

ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740

gcagcacctc cgtaa 1755

<210> 54

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 54

ggtggtgga gcacaccgt gccctgcct cagcagcagc ctccgcagca agaaccgct 60

ccgcctccgg ttccgctggc aagccctgca gcatgtgaag cactgaaaga tggtaatggt 120

gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180

ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240

aaacgtaccg gtattgatgg ttatccgcat gaaatcaaat ttcgtctgcg tatgcctgca 300

gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgc 360

gccaccggtg gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420

attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480

gttgcatctg gctcggatcc gcagcagcgt ctggatatgg gttataatag cgtggatcag 540

gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcacc ggataaaagc 600

tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660

agccattatg atggtattgt tgccgggtgca ccgggttatt agctgccgaa agcaggtatt 720

agtgggtgat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacaggg 780

gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840

ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900

tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960

tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020

gcaatggcag gtccggttaa tagcgcaggt acaccctgtg ataactgttg ggcatgggat 1080

gcaggtatga gcggtctgag cggtaacaac tataatcaag gttggcgtag ttggtggctg 1140

ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccgga ccatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgattttga tattgacctg ctgaaaattt gggcaaccag cggtcagttt 1320

 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggta aatgattct gtatcatggt atgtcagatg cagccttttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacce ctctggttgc atgggtttaa cgtggtgaag caccgatca gattagcga 1620
 tggctggtta ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagcaa ttttgcattg 1740

 gcagcacctc cgtaa 1755
 <210> 55
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 55

 ggtggtggta gcacaccgt gccgctgcct cagcagcagc ctccgagca agaaccgct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatatggttt ggccgaatgc agcaaccgtt gttgaagtgt cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

 gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgc 360
 gccaccggta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
 attgcaaccg atggtggca tgataatgcc gttaatgata atccgatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gtttagcaa cgggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720

 agtggatgat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840

ctgggcacct gtgatgcctt ggatggctct gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctcagc 960
 tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcatgggat 1080
 gcaggtatga gcggtctgag cgttacaacc tataatcaag gttggcgtag ttggtggctg 1140

ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccgga ccatgcccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggta aaatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560

atgctgacce ctctggttgc atgggttgaa cgtggtgaag caccggatca gattagcgc 1620
 tggctcggta ctccgggtta ttttgggtgt gccgcagta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagcaa ttttgcagtg 1740
 gcagcacctc cgtaa 1755

<210> 56
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 56

ggtggtggta gcacaccgt gccgctgcct cagcagcagc ctccgcagca agaaccgcct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tgtaaatggt 120

gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgc 360
 gccaccggtg gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgagtt 420
 attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540

gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
tattttatcg gttgtagcga aggtggtcgt gaaggtatga tggcgagcca gcgttttccg 660
agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720
agtgggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
ctgggcacct gtgatgcctt ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960

tgtgttgggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcatgggat 1080
gcaggtatga gcggtctgag cgggtacaacc tataatcaag gttggcgtag ttgggtgctg 1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
tggctggttg attttgcaac ccctccgga ccgatgccga tgacacaggt tgcagcacgc 1260
atgatgaaat tcgatttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
accagagta gcatggatg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380

cgtgggtgta aaatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
atgctgacc ctcgtggtgc atgggttga cgtggtgaag caccggatca gattagcga 1620
tggctggtg ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgtat 1680
ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740
gcagcacctc cgtaa 1755

<210> 57

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 57

gggtgggtgta gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgcct 60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
gatatggttt ggccgaatgc agcaaccgtt gttgaagtgt cagcatggcg tgatgcagca 180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgctg tatgcctgca 300

gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgcga 360

 gccaccggtg gcattgggtg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
 attgcaaccg atggtggca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggctgt gaaggtatga tgctggcgca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttadc agctgccgaa agcaggtatt 720
 agtgggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780

 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggctctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgtt ggcattggat 1080
 gcaggtatga gcggtctgag cggtacaacc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200

 tggctggttg attttgcaac ccctccgga ccatgcca tgacacaggt tgcagcagc 1260
 atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggtg aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
 acagcagatt atiatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacce ctctggttgc atgggttga cgtggtgaag caccggatca gattagcgc 1620

 tggctggtg ctccgggtta ttttgggtt gccgcagta cccgtccgct gtgtccgtat 1680
 ccgcatagc cagttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740
 gcagcacctc cgtaa 1755

<210> 58

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 58

ggtgggtgta gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgcct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tgtaaatggt 120
 gatatggttt ggccgaatgc agcaaccgtt gttgaagtgg cagcatggcg tgatgcagca 180

 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgcga 360
 gccaccggta gcattgggtg tggccagatt gcaagcgcac tgagccgtgg ttttgagtt 420
 attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600

 tattttatcg gttgtagcga aggtggctgt gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttadc agctggtgaa agcaggtatt 720
 agtgggtgat ggaccacaca gacccctggca ccggcagccg ttggcctgga tgcacagggt 780
 gtcccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgcctt ggatggctctg gcagatggca ttgttgataa ctatctgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag 960
 tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020

 gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg gccatgggat 1080
 gcaggtatga gcggtctgag ccgtacaacc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac cctccggaa ccgatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgattttga tattgaccgc ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatgattg gcatggtgca accagaccg atctggcagc ctttctgat 1380
 cgtggtggtg aatgattct gtatcatggt atgtcagatg cagccttttag tgcactggat 1440

 acagcagatt atiatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacc ctcggttgc atgggttgaa cgtggtgaag caccggatca gattagcgca 1620
 tggctctgta ctccgggtta ttttggtgtt gccgcagta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagccaa ttttgcatgt 1740
 gcagcacctc cgtaa 1755

<210> 59

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 59

```

ggtaggtgta gcaaccgct gccgctgct cagcagcagc ctccgcagca agaaccgct      60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt    120
gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca    180
ccggcaaccg caagcgcagc agcaactgcc gaacattgtg aagttagcgg tgcaattgcc    240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca    300
gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgc    360
gccaccgcta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt    420

attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc    480
gttgcatitt gctcggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag    540
gttaccaggc caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc    600
tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg    660
agccattatg atggtattgt tgcgggtgca ccgggttata agctgccgaa agcagcagatt    720
agtggatgca ggaccacaca gacacctgca ccggcagccg ttggcctgga tgcacagggt    780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt    840

ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca    900
tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctcagc    960
tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaaccg    1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgtt ggcatgggat    1080
gcaggtatga gcggtctgag cggtacaacc tataatcaag gttggcgtag ttggtggctg    1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc    1200
tggctggttg atttgcaac ccctccgga ccgatgccga tgacacaggt tgcagcagc    1260

atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt    1320
accagagta gcatggattg gcatggtgca accagaccg atctggcagc ctttcgtgat    1380
cgtggtggtg aatgattct gtatcatggt atgtcagatg cagccttag tgcactggat    1440
acagcagatt attatgaacg tctgggtgca gcaatgectg gtgcagcagg ttttgcacgt    1500
ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat    1560

```

atgctgacc ctctggttgc atgggttgaa cgtggtgaag caccgatca gattagcga 1620
 tggctctgga ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgtat 1680

ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740
 gcagcacctc cgtaa 1755

<210> 60
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 60

ggtggtgga gcaaccgct gccctgcct cagcagcagc ctccgcagca agaaccgct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatattggtt ggccgaatgc agcaaccgtt gttgaagtgc cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240

aaacgtaccg gattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaagggt ggtagcggca ccaatggtag cctgagcgcga 360
 gccaccgcta gcaattggtg tggccagatt gcaagcgcac tgagccgtgg ttttgcattt 420
 attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtcctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660

agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720
 gcgggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgcctt ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
 tgcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtcgggttaa tagcgcaggt acaccgctgt ataactgttg ggcatgggat 1080

gcaggtatga gcggtctgag cggtaacaacc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggtttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccggaa ccgatgccga tgacacaggt tgcagcacgc 1260

atgatgaaat tcgattttga tattgaccgc ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggta aatgatctct gtatcatggt atgtcagatg cagccttttag tgcaactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500

 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacce ctctggttgc atgggttgaa cgtgggtgaag caccggatca gattagcgc 1620
 tggctctggta ctccgggtta ttttgggtgt gccgcacgta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcacgt 1740
 gcagcacctc cgtaa 1755
 <210> 61
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 61
 ggtggtggta gcacaccgt gccgtgcct cagcagcagc ctccgcagca agaaccgcct 60

 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgc 360
 gccaccggtg gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
 attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480

 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttadc agctcccga agcaggtatt 720
 agtggtgcat ggetgacaca gagectggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900

 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960

tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcattggat 1080
gcaggtatga gcggtctgag cggtagaacc tataatcaag gttggcgtag ttggtggctg 1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
tggctggttg attttgcaac ccctccgga cccgatgccga tgacacaggt tgcagcacgc 1260
atgatgaaat tcgattttga tattgaccgc ctgaaaattt gggcaaccag cggtcagttt 1320

accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
cgtggtggtta aatgattct gtatcatggt atgtcagatg cagccttttag tgcactggat 1440
acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
atgctgacc ctctggttc atgggttgaa cgtggtgaag caccgatca gattagcga 1620
tggctctgta ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgat 1680
ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagcaa ttttgcattg 1740

gcagcacctc cgtaa 1755

<210> 62

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 62

ggtggtggtta gcacaccgct gccctgcct cagcagcagc ctccgagca agaaccgct 60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
gatattggttt ggccgaatgc agcaaccgtt gttgaagttag cagcatggcg tgatgcagca 180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggttag cctgagcga 360
gccaccggtta gcattggttg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
gttgcatctg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660

agccattatg atggtattgt tgcgggtgca ccgggttatc agctgccgaa agcaggtatt 720

agtggatgcat ggaccctgca gacacctgca ccggcagccg ttggcctgga tgcacagggt 780

gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840

ctgggcacct gtgatgcctt ggatggctcg gcagatggca ttgttgataa ctatcgtgca 900

tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960

tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020

gcaatggcag gtccggttaa tagcgcaggf acaccgctgt ataatcgtt ggcatgggat 1080

gcaggtatga gcggtctgag cggtacaacc tataatcaag gttggcgtag ttggtggctg 1140

ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc 1200

tggctggttg attttgcaac ccctccggaa ccgatgccga tgacacaggt tgcagcacgc 1260

atgatgaaat tegtatttga tattgaccgc ctgaaaattt gggcaaccag cggtcagttt 1320

accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380

ctggttggtg aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440

acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500

ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560

atgctgacc ctctggttgc atgggtttaa cgtggtgaag caccggatca gattagcgc 1620

tggctctggtg ctccgggtta ttttgggtgt gccgcacgta cccgtccgct gtgtccgtat 1680

ccgcagatcg cacgttataa aggtagcggf gatattaaca ccgaagccaa ttttgcattg 1740

gcagcacctc cgtaa 1755

<210> 63

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 63

ggtggtggtg gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgct 60

ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120

gatattggtt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180

ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240

aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

gaaatggaat gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgc 360

gccaccgta gcattggtgg tggccagatt gcaagcgac tgagccgtgg ttttgagtt 420
 attgcaaccg atggtgtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540

gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggtcgt gaaggatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720
 agtgggtgat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccctga ttctgaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgcct ggatggtctg gcagatggca ttgttgataa ctatctgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag 960

tgtgttggtg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcattggat 1080
 gcaggtatga gcggtctgag cggtaacaacc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccgga ccgatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgatttga tattgaccg ctgaaaattt gggcaaccag cgtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttctgat 1380

cgtggtggtg aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacce ctctggttc atgggttgaa cgtggtgaag caccggatca gattagcga 1620
 tggctggtg ctccgggtta ttttggtgtt gccgcagta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagcaa ttttgatgt 1740
 gcagcacctc cgtaa 1755

<210> 64

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 64

ggtggtggta gcacaccgct gccctgcct cagcagcagc ctccgcagca agaaccgcct 60

ccgcctccgg ttccgctggc aagccctgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatatggttt ggccgaatgc agcaaccgtt gttgaagtig cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagtagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgc 360

 gccaccgta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
 attgcaaccg atggtggca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcathtt gctcggatcc gcagcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggctgt gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgccgggtgca ccgggttctc agctccgaa agcaggtatt 720
 agtgggtgat ggaccacaca gacccctggca ccggcagccg ttggcctgga tgcacagggt 780

 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctggtgag tcaggcaatt 840
 ctgggcaacct gtgatccctt ggatggctctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttggcg caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggt acaccctgtg ataactgttg ggcatgggat 1080
 gcaggtatga gcggtctgag cggtaaaccc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200

 tggctggttg attttgcaac ccctccgga ccatgcca tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggta aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacc ctctggttgc atgggttga cgtggtgaag caccggatca gattagcgc 1620

 tggctcggta ctccgggtta ttttgggtt gccgcagta cccgtccgt gtgtccgtat 1680
 ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740
 gcagcacctc cgtaa 1755

<210> 65

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 65

```

ggtggtggta gcacaccgt gccctgcct cagcagcagc ctccgcagca agaaccgcct      60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt      120
gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca      180

ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc      240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca      300
gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgcga      360
gccaccggta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt      420
attgcaaccg atggtggta tgataatgcc gttaatgata atccggatgc actgggcacc      480
gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag      540
gttaccaggg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc      600

tattttatcg gttgtagcga aggtggctcg gaaggtatga tgctgagcca gcgttttccg      660
agccattatg atggtattgt tgcgggtgca ccgggttata agctgccgaa agcaggtatt      720
agtgggtgcat ggaccacaca gagectggca ccggcagccg ttggcctgga tgcacagggt      780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctggc gcaggcaatt      840
ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca      900
tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag      960
tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc     1020

gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcatgggat     1080
gcaggtatga gcggtctgag cggtaacaac tataatcaag gttggcgtag ttggtggctg     1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgtaa gcggttttag cgcacgtagc     1200
tggtcgtttg attttgcaac ccctccgga cccatgccga tgacacaggt tgcagcacgc     1260
atgatgaaat tcgattttga tattgaccgg ctgaaaattt gggcaaccag cggtcagttt     1320
accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat     1380
cgtggtggta aaatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat     1440

acagcagatt attatgaacg tctgggtgca gcaatgcctg gtcagcagc ttttgcacgt     1500
ctgtttctgg ttccgggtat gaatcattgt agcgggtggc cgggtacaga tcgttttgat     1560
atgctgacce ctctggttgc atgggtttaa cgtggtgaag caccggatca gattagcga     1620

```

tggctctggta ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740
 gcagcacctc cgtaa 1755
 <210> 66
 <211> 1755
 <212> DNA
 <213> Artificial Sequence

 <220><223> Synthetic
 <400> 66
 ggtggtggta gcacaccgt gccctgcct cagcagcagc ctccgcagca agaaccgct 60
 ccgcctccgg ttccgtggc aagccgtgca gcatgtgaag cactgaaaga tgtaaatggt 120
 gatattggtt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgc 360
 gccaccggta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420

 attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggctcg gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggtatc agctgccga agcaggtatt 720
 agtggatgat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840

 ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgagagc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataactgttg ggcattggat 1080
 gcaggtatga gcggtctgag cgttacaacc tataatcaag gttggcgtag ttggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
 tggctggttg attttgaac ccctccggaa ccgatgccga tgacacaggt tgcagcagc 1260

 atgatgaaat tcgattttga tattgaccgg ctgaaaattt gggcaaccag cggtcagttt 1320

accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtgggtgga aaatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggc cgggtacaga tcgttttgat 1560
 atgctgacce ctctggttgc atgggttga cgtggtgaag caccggatca gattagcga 1620
 tggctcgga ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgat 1680

ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagcaa ttttgcattg 1740
 gcagcacctc cgtaa 1755

<210> 67

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 67

ggtgggtgga gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgcct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatattggtt ggccgaatgc agcaaccgtt gttgaagtgc cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240

aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
 gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcga 360
 gccaccgga gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
 attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660

agccattatg atggtattgt tgcgggtgca ccgggttadc agctgccgaa agcaggtatt 720
 agtgggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gtcccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggtctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020

gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg gccgtgggat 1080

gcaggtatga gcggtctgag cgttacaacc tataatcaag gttggcgtag ttggtggctg 1140

ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200

tggctggttg attttgcaac ccctccgga ccatgcccga tgacacaggt tgcagcacgc 1260

atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320

accagagta gcatggattg gcatggtgca accagcaccc atctggcagc ctttcgtgat 1380

cgtggtggta aaatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440

acagcagatt attatgaacg tctgggtgca gcaatgctg gtgcagcagg ttttgcacgt 1500

ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560

atgctgacce ctctggttgc atgggttgaa cgtggtgaag caccggatca gattagcgc 1620

tggctctgga ctccgggtta ttttgggtgt gccgcacgta cccgtccgct gtgtccgtat 1680

ccgcagatcg cagttataa aggtagcggg gatattaaca ccgaagccaa ttttgcattg 1740

gcagcacctc cgtaa 1755

<210> 68

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 68

ggtggtggta gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgcct 60

ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120

gatatggttt ggccgaatgc agcaaccgtt gttgaagtgg cagcatggcg tgatgcagca 180

ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240

aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgc 360

gccaccggta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420

attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480

gttgcatctg gctctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540

gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600

tattttatcg gttgtagcga aggtgggtcgt gaaggtatga tgctgagcca gcgttttccg 660

agccattatg atggtattgt tgccgggtgca ccgggttadc agctgccgaa agcaggtatt 720

agtggatgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggctctg gcagatggca ttgttgataa ctatcgtgca 900

 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataacgttg ggcatgggat 1080
 gcaggtatga gcggtctgag cgggtacaacc tataatcaag gttggcgtag ttgggtggctg 1140
 ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccggaa ccgatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgattttga tattgacccg ctgaaaattt gggcaaccag cggtcagttt 1320

 acccagagtc tgatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtgggtgta aaatgattct gtatcatggt atgtcagatg cagccttttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat 1560
 atgctgacc ctcctggttc atgggttgaa cgtgggtgaag caccggatca gattagcgc 1620
 tggctggtg ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcagtt 1740

 gcagcacctc cgtaa 1755
 <210> 69
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 69
 ggtgggtgta gcacaccgct gccgetgect cagcagcagc ctccgcagca agaaccgcct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
 gatatggttt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180
 ccggcaaccg caagcgcagc agcaactgcc gaacattgtg aagttagcgg tgcaattgcc 240
 aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300

 gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgc 360
 gccaccggta gcattgggtg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420

attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc 480
 gttgcatttg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
 gttaccacagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
 tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg 660
 agccattatg atggtattgt tgcgggtgca ccgggttadc agctgccgaa agcaggtatt 720

agtggtgcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
 gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
 ctgggcacct gtgatgccct ggatggctctg gcagatggca ttgttgataa ctatcgtgca 900
 tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960
 tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc 1020
 gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcatgggat 1080
 gcaggtatga gcggtctgag cggtaacaacc tataatcaag gttggcgtag ttggtggctg 1140

ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc 1200
 tggctggttg attttgcaac ccctccggaa ccgatgccga tgacacaggt tgcagcacgc 1260
 atgatgaaat tcgattttga tattgacgcg ctgaaaattt gggcaaccag cggtcagttt 1320
 acccagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttcgtgat 1380
 cgtggtggtg aaatgattct gtatcatggt atgtcagatg cagcctttag tgcaactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggc cgggtacaga tcgttttgat 1560

atgctgacc ctctggttgc atgggttgaa cgtggtgaag caccggatca gattagcga 1620
 tggctggtg ctccgggtta ttttgggtgt gccgcagta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccaa ttttgcatgt 1740
 gcagcacctc cgtaa 1755

<210> 70
 <211> 1755
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 70

ggtggtggtg gcacaccgct gccgctgcct cagcagcagc ctccgcagca agaaccgct 60
 ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120

gatatggttt ggccgaatgc agcaaccgtt gttgaagtgg cagcatggcg tgatgcagca 180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagtagcgg tgcaattgcc 240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgcga 360
gccaccggta gcattgggtg tggccagatt gcaagcgcac tgagccgtgg ttttcagtt 420
attgcaaccg atggtggfca tgataatgcc gttaatgata atccggatgc actgggcacc 480
gttgcatthg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540

gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600
tattttatcg gtttagcga aggtggctgt gaaggatga tgctgagcca gcgttttccg 660
agccattatg atggtattgt tgcgggtgca ccgggttadc agctgccgaa agcaggtatt 720
agtggfcat ggaccacaca gagcctggca ccggcagccg ttggcctgga tgcacagggt 780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
ctgggcacct gtgatgccct ggatggctg gcagatggca ttgtgataa ctatcgtgca 900
tgtcaggcag cattgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag 960

tgtgttggf caaaaaccg agattgtctg agtccggttc aggtgaccg aattaacgc 1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgtt ggcattggat 1080
gcaggatga gcggtctgag cggtacaacc tataatcaag gttggcgtag ttggtggctg 1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc 1200
tggctggtg attttgcaac cctccggaa ccgatgccga tgacacaggt tgcagcagc 1260
atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cggtcagttt 1320
accagagta gcatggattg gcatggtgca accagaccg atctggcagc ctttcgtgat 1380

cgtggfggta aatgattct gtatcatggt atgtcagatg cagcctttag tgcactggat 1440
acagcagatt gglatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttcacgt 1500
ctgtttctgg ttccgggtat gaatcattgt agcggfggtc cgggtacaga tcgttttgat 1560
atgctgacce ctctggttgc atgggttga cftggfgaag caccggatca gattagcgc 1620
tggctctgga ctccgggtta ttttggfgt gccgcagta cccgtccgct gftccgtat 1680
ccgcagatcg cacgttataa aggtagcggf gatattaaca ccgaagccaa ttttgcattg 1740
gcagcacctc cgtaa 1755

<210> 71

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 71

```

ggtggtgta gcacaccgt gccgctgcct cagcagcagc ctccgcagca agaaccgcct      60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt      120
gatatggttt ggccgaatgc agcaaccgtt gttgaagtgg cagcatggcg tgatgcagca      180
ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc      240
aaacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctcgc tatgcctgca      300
gaatggaatg gtcgcttttt tatggaaggt gtagcggca ccaatggtag cctgagcgc      360

gccaccgta gcattggtgg tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt      420
attgcaaccg atggtggtca tgataatgcc gttaatgata atccggatgc actgggcacc      480
gttgcaattg gtctggatcc gcaggcacgt ctggatatgg gttataatag cgtggatcag      540
gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc      600
tattttatcg gttgtagcga aggtggtcgt gaaggtatga tgctgagcca gcgttttccg      660
agccattatg atggtattgt tgcgggtgca ccgggttata agctgccgaa agcaggtatt      720
agtggatgat ggaccacaca gacccctggc ccggcagccg ttggcctgga tgcacagggt      780

gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt      840
ctgggcacct gtgatgcctt ggatggtctg gcagatggca ttgtgataa ctatcgtgca      900
tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggccctgcag      960
tgtgttgggt caaaaaccgc agattgtctg agtccggttc aggtgaccgc aattaaacgc     1020
gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataatcgttg ggcattggat     1080
gcaggtatga gcggtctgag cggtaacaac tataatcaag gttggcgtag ttggtggctg     1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgtta gcggttttag cgcacgtagc     1200

tggctggttg attttgcaac cctccggaa ccgatgccga tgacacaggt tgcagcacgc     1260
atgatgaaat tegattttga tattgaccgg ctgaaaattt gggcaaccag cggtcagttt     1320
accagagta gcatggattg gcatggtgca accagaccg atctggcagc ctttcgtgat     1380
cgtggtgta aatgatctt gtatcatggt atgtcagatg cagccttttag tgcactggat     1440
acagcagatt attatgaacg tctgggtgca gcaatgcctg gtgcagcagg ttttgcacgt     1500
ctgtttctgg ttccgggtat gaatcattgt agcgggtggtc cgggtacaga tcgttttgat     1560
atgctgacc ctcgtgttgc atgggttgaa cgtggtgaag caccggatca gattagcgc      1620

tggtctggta ctccgggtta ttttgggtt gccgcagta cccgtccgct gtgtccgtat     1680

```

ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgcggccaa ttttgcattg 1740
gcagcacctc cgttaa 1755
<210> 72
<211> 1755
<212> DNA
<213> Artificial Sequence
<220><223> Synthetic
<400> 72
ggtaggtgta gcacaccgt gccctgcct cagcagcagc ctccgcagca agaaccgct 60
ccgcctccgg ttccgctggc aagccgtgca gcatgtgaag cactgaaaga tggtaatggt 120
gatattggtt ggccgaatgc agcaaccgtt gttgaagttg cagcatggcg tgatgcagca 180

ccggcaaccg caagcgcagc agcactgccg gaacattgtg aagttagcgg tgcaattgcc 240
aacgtaccg gtattgatgg ttatccgtat gaaatcaaat ttcgtctgcg tatgcctgca 300
gaatggaatg gtcgcttttt tatggaaggt ggtagcggca ccaatggtag cctgagcgcga 360
gccaccggtg gcattggtag tggccagatt gcaagcgcac tgagccgtgg ttttgcagtt 420
attgcaaccg atggtggtag tgataatgcc gttaatgata atccggatgc actgggcacc 480
gttgcatttg gtcctggatc gcaggcacgt ctggatatgg gttataatag cgtggatcag 540
gttaccagg caggtaaagc agcagttgca cgtttttatg gtcgtgcagc agataaaagc 600

tattttatcg gttgtagcga aggtggctgt gaaggtatga tgctgagcca gcgttttccg 660
agccattatg atggtattgt tgcgggtgca ccgggttatac agctgccgaa agcaggtatt 720
agtggtagc gaccacaca gacccctggc ccggcagccg ttggcctgga tgcacaggtt 780
gttccgctga ttaacaaaag ctttagtgat gccgatctgc atctgctgag tcaggcaatt 840
ctgggcacct gtgatgcctt ggatggctct gcagatggca ttgttgataa ctatctgca 900
tgtcaggcag catttgatcc ggcaacagca gcaaatccgg caaatggtca ggcctgcag 960
tgtgttggtg caaaaaccg agattgtctg agtccggttc aggtgaccgc aattaaaccg 1020

gcaatggcag gtccggttaa tagcgcaggt acaccgctgt ataactgtt ggcatgggat 1080
gcaggtatga gcggtctgag cggtacaacc tataatcaag gttggcgtag ttggtggctg 1140
ggtagcttta atagcagcgc aaataatgcc cagcgtgta gcggttttag cgcacgtagc 1200
tggctggttg attttgcaac ccctccgaa ccgatgccga tgacacaggt tgcagcacgc 1260
atgatgaaat tcgattttga tattgaccg ctgaaaattt gggcaaccag cgttcagttt 1320
accagagta gcatggattg gcatggtgca accagcaccg atctggcagc ctttctgat 1380

cgtgggtgga aaatgattct gtatcatggt atgtcagatg cagccttttag tgcactggat 1440
 acagcagatt attatgaacg tctgggtgca gcaatgcctg gtcagcagg ttttgcacgt 1500
 ctgtttctgg ttccgggtat gaatcattgt agcgggtggc cgggtacaga tcgttttgat 1560
 atgctgacct ctctggttgc atgggttga cgtgggtgaag caccggatca gattagcgca 1620
 tggctctgga ctccgggtta ttttgggtt gccgcacgta cccgtccgct gtgtccgtat 1680
 ccgcagatcg cacgttataa aggtagcggg gatattaaca ccgaagccag ctttgcacgt 1740
 gcagcacctc cgtaa 1755

<210> 73

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 73

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

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

Gln Ile Ala Ser Ala Leu Ser Arg Gly Phe Ala Val Ile Ala Thr Asp
 130 135 140

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

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

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

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

 Ser Ser Ala Asn Asn Ala Gln Arg Val Ser Gly Phe Ser Ala Arg Ser

385 390 395 400
 Trp Leu Val Asp Phe Ala Thr Pro Pro Glu Pro Met Pro Met Thr Gln
 405 410 415
 Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430
 Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445

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

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

 Asn Phe Ala Cys Ala Ala Pro Pro
 580
 <210> 74
 <211> 584
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 74
 Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

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

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

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

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

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

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

 Phe Asp Pro Ala Thr Ala Ala Asn Pro Ala Asn Gly Gln Ala Leu Gln
 305 310 315 320
 Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335
 Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350
 Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

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

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

 Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly

500 505 510
 Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525
 Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540
 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575
 Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 75
 <211> 584
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 75

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15
 Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

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

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser
 100 105 110
 Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly

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

 Val Ala Ala Arg Met Met Lys Phe Asp Phe Asp Ile Asp Pro Leu Lys
 420 425 430
 Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445
 Gly Ala Thr Ser Thr Asp Leu Ala Ala Phe Arg Asp Arg Gly Gly Lys
 450 455 460
 Met Ile Leu Tyr His Gly Met Ser Asp Ala Ala Phe Ser Ala Leu Asp
 465 470 475 480

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

 Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560
 Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575
 Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 76

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 76

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala
 65 70 75 80

Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser
 100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Asn Phe Ala Val Ile Ala Thr Asp
 130 135 140

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

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Tyr Gly Tyr Asn
 165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe
 180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly
 195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp
 210 215 220

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

Thr Ala Asp Trp Tyr Glu Arg Leu Gly Ala Ala Met Pro Gly Ala Ala
 485 490 495

Gly Phe Ala Arg Leu Phe Leu Val Pro Gly Met Asn His Cys Ser Gly
 500 505 510

Gly Pro Gly Thr Asp Arg Phe Asp Met Leu Thr Pro Leu Val Ala Trp
 515 520 525

Val Glu Arg Gly Glu Ala Pro Asp Gln Ile Ser Ala Trp Ser Gly Thr
 530 535 540

Pro Gly Tyr Phe Gly Val Ala Ala Arg Thr Arg Pro Leu Cys Pro Tyr
 545 550 555 560

Pro Gln Ile Ala Arg Tyr Lys Gly Ser Gly Asp Ile Asn Thr Glu Ala
 565 570 575

Asn Phe Ala Cys Ala Ala Pro Pro
 580

<210> 77

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 77

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln
 1 5 10 15

Gln Glu Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys
 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala
 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala
 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala
 65 70 75 80

Lys Arg Thr Gly Ile Asp Gly Tyr Pro Tyr Glu Ile Lys Phe Arg Leu
 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser
 100 105 110
 Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly
 115 120 125

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

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

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

 Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335
 Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro

<210> 78

<211> 584

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 78

Gly Gly Gly Ser Thr Pro Leu Pro Leu Pro Gln Gln Gln Pro Pro Gln

1 5 10 15

Gln Glu Pro Pro Pro Pro Pro Val Pro Leu Ala Ser Arg Ala Ala Cys

 20 25 30

Glu Ala Leu Lys Asp Gly Asn Gly Asp Met Val Trp Pro Asn Ala Ala

 35 40 45

Thr Val Val Glu Val Ala Ala Trp Arg Asp Ala Ala Pro Ala Thr Ala

 50 55 60

Ser Ala Ala Ala Leu Pro Glu His Cys Glu Val Ser Gly Ala Ile Ala

65 70 75 80

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

 85 90 95

Arg Met Pro Ala Glu Trp Asn Gly Arg Phe Phe Met Glu Gly Gly Ser

 100 105 110

Gly Thr Asn Gly Ser Leu Ser Ala Ala Thr Gly Ser Ile Gly Gly Gly

 115 120 125

Gln Ile Ala Ser Ala Leu Ser Arg Asn Phe Ala Val Ile Ala Thr Asp

 130 135 140

Gly Gly His Asp Asn Ala Val Asn Asp Asn Pro Asp Ala Leu Gly Thr

145 150 155 160

Val Ala Phe Gly Leu Asp Pro Gln Ala Arg Leu Asp Met Gly Tyr Asn

 165 170 175

Ser Tyr Asp Gln Val Thr Gln Ala Gly Lys Ala Ala Val Ala Arg Phe

 180 185 190

Tyr Gly Arg Ala Ala Asp Lys Ser Tyr Phe Ile Gly Cys Ser Glu Gly

 195 200 205

Gly Arg Glu Gly Met Met Leu Ser Gln Arg Phe Pro Ser His Tyr Asp
 210 215 220
 Gly Ile Val Ala Gly Ala Pro Gly Phe Gln Leu Pro Lys Ala Gly Ile
 225 230 235 240

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

 Phe Asp Pro Ala Thr Ala Ala Asn Pro Ala Asn Gly Gln Ala Leu Gln
 305 310 315 320
 Cys Val Gly Ala Lys Thr Ala Asp Cys Leu Ser Pro Val Gln Val Thr
 325 330 335
 Ala Ile Lys Arg Ala Met Ala Gly Pro Val Asn Ser Ala Gly Thr Pro
 340 345 350
 Leu Tyr Asn Arg Trp Ala Trp Asp Ala Gly Met Ser Gly Leu Ser Gly
 355 360 365

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

 Ile Trp Ala Thr Ser Gly Gln Phe Thr Gln Ser Ser Met Asp Trp His
 435 440 445
 Gly Ala Thr Ser Thr Asp Leu Ala Ala Phe Arg Asp Arg Gly Gly Lys

gtcacacagg ccgggaaggc ggctgtggct cgcttttacg gccgtgcggc cgataaatcc 600

tatattattg gctgtagtga aggagggcgg gaaggcatga tgctgtcaca gcggtttcca 660

agccactacg acggtattgt ggcaggtgcc ccgggctacc agttaccta agccgggatt 720

tctggggcgt ggacaacaca gtctcttgcc ccagccgccg tcggtctcga cgctcaggtt 780

gtcccgtta ttaataaatc attcagcgt gccgacttac atctgctgag ccaggaatt 840

ttaggtacat gtgatgcgt ggatggctctg gccgatggca tagtagacaa ttatagagcc 900

tgtcaagctg ctttgacc tgctaccgca gccaatccgg ctaacggcca agcactgcaa 960

tgtgttggcg cgaaaaccgc ggattgtctg agtccagtgc aagtgactgc gataaaacgt 1020

gcaatggcgg gcccggtcaa tagtgctggg acaccctgt ataatcgctg ggcttgggac 1080

gcaggaatgt cgggattgag tgaaccacg tacaaccagg gatggcgttc ctggtggttg 1140

ggctcgttca acagcagcgc taataacgt caacgagtgt caggtttcag tgcgcgcagc 1200

tggcttggg attttgccac cccccgga ccaatgccga tgactcaggt tcccgcgcgt 1260

atgatgaagt ttgatttga tattgatccg ttgaaaattt gggccacctc aggacagttt 1320

acgcaaaagt cgatggactg gcacggggcg acatcaactg atctggcagc ctttcgggac 1380

agagggggta aaatgatcct ttatcatggt atgtctgatg ccgcattctc tgcgctcgat 1440

acggccgact actacgaacg cttaggggct gctatgctg gtgcggctgg gtttgcacgt 1500

ctcttttttg tgcttggtat gaatcactgt tcaggaggcc ctggcaccga tagatttgat 1560

atgctcacc cctcgttgc ttgggtagag agaggtgaag ctccggatca aatcagtct 1620

tggctcggga cccttggeta cttcgggtgc gctgccagaa cccgtccgct gtgtccatat 1680

cccagattg cgcatataa agggctctggg gatattaata cggaagctaa ttttgcgtgc 1740

gcggcggcgc cgtaa 1755

<210> 80

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 80

ggcgggtgggt caacaccgtt gccgttacct caacagcagc ccccgaaca ggaaccgccg 60

ccgccacctg ttccgttagc atcacgtgca gcatgtgagg ccctgaaaga tggtaatggc 120

gatatggttt ggccaaacgc cgctaccgtt gtggaggtgg ctgcttggcg tgaccggct 180

cccgaacag cgtctgcagc cgccctgcca gagcactgcg aagtttctgg tgcaattgcc 240

aagcgtaccg gaattgatgg ttatccttat gaaattaagt tccgtctgcg catgcccgct 300
aatggaatg gacgtttttt catggagggt ggctcgggta caaatggttc tctctccgct 360
gcgacaggca gtatcggagg cggtaaatac gcttcagcgt tgtccagagg ctttgcctgt 420

attgcaacgg atggtggcca tgataacgca gtcaatgata accctgacgc actgggaacg 480
gttgcgtttg gtttggatcc tcaggcacgt ttagatatgg ggtacaattc ttatgatcag 540
gtaactcagg cagggaaaagc cgccgttgcg cgtttttatg gccgagctgc tgacaagtgc 600
tactttattg gggtcagcga aggtgggcga gaaggcatga tgttgcgca acgatttccc 660
agtcattacg acgggattgt ggccgggtgcg ccaggttacc agctcccga agcgggtatc 720
tcgggtgctt ggacaacca gagtttagcg ccggcggctg taggcttaga tgcacagggt 780
gtcccattga ttaataaatc attttctgat gctgacctcc atctgttacc gcaggcgatc 840

ctcggaaact gcgatgcgct ggatggcctg gcagacggta tagtgcataa ttacagagct 900
tgtcaagcgg cgtttgacc tcgcaccgcg gccaacccag cgaacgggca ggcacttcaa 960
tgtgttggcg caaaaacagc ggattgtctg tccccgtac aagtaacggc aatcaagcgt 1020
gcaatggctg gtccggtaga ttcggccggc acaccgctgt acaatcggtg ggcgtgggac 1080
gcaggtatgt caggtttatc cggaacgacg tacaatcaag ggtggagatc gtggtggctg 1140
ggttccttta actctagtgc aaacaacgcc cagcgcgtct ctggattttc cgcgcgttct 1200
tggctggctg atttcgcgac gccaccgaa cccatgccga tgaccaggt agcggcacgc 1260

atgatgaagt ttgatttga tatcgatccc ctcaaaattt gggccactag tggccagttt 1320
acgcagtcaa gtatggattg gcatggcgcc acgagtacag acttggcggc ttttcgcgac 1380
cgcggggta agatgatcct gtaccacggc atgagcgatg ctgcatttag tgccttcgat 1440
acagcagact ggtatgaacg tcttggtgca gctatgccag gtgcagcagg ctttgcctgc 1500
ctgttcttgg ttccggggat gaaccattgc agcgggtggtc cgggtaccga ccgcttcgat 1560
atgctgacgc cattagtagc ttgggttga cgtggggaag cccctgatca aattagtga 1620
tggtcgggca cacctgggta tttcggcgtg gcggcccgga cgagaccgct ttgtccctat 1680

ccccagattg cccggtataa aggtagtggg gacataaaca ccgaagcga ttttgcgtgc 1740
gccgcgccgc cctga 1755
<210> 81
<211> 1755
<212> DNA
<213> Artificial Sequence
<220><223> Synthetic

<400> 81

ggcgggtggct ccaccccttt acccttgcca cagcagcaac cgcccgagca agaacctcca 60
 cccctccgg tccccctggc gtcgcgcgca gctgcaag ccttgaaaga tgtaaatggc 120
 gacatggttt ggccaaacgc cgccaccgta gttgaggtag cagcctggcg cgatgccgct 180
 ccagccactg cctcggccgc cgcgctcccg gaacattgtg aagttagtgg cgcaattgcg 240

 aaacgcacag gaattgacgg ctatccgtat gagatcaaat tccggctgcg tatgccgca 300
 gaatggaacg gccggttctt tatggaaggc gggagtggaa ccaatggatc cctgtcagct 360
 gcgactggat ccatagcgcg tgggcagata gcatcgctc ttagccgggg ttttgcggtg 420
 atcgcgaccg atggcgggca tgacaacgca gtgaacgaca atcctgatgc gttgggcacc 480
 gttgcttttg gtcttgatcc gcaggctcgg cttgattacg ggtacaatag ttatgaccag 540
 gtaacacagg ctggtaaagc ggcggtggcc cgttctatg gcagagcagc cgacaaatcg 600
 tatttcattg gctgctctga aggcggtcgc gaaggaatga tgctgtctca gcgatttccg 660

 tctcattatg atgggattgt tgccggtgcc cgggggtatc aacttccgaa agcgggtatc 720
 tccggtgctt ggacaactca atcgetggcc cccgcggcgg tgggtctgga tgctcaaggc 780
 gtccctctca ttaataagtc ttttagtgat gctgatttgc atttctgtc ccaggccatt 840
 ctgggaacgt gcgatgctct tgaccgtctc gccgatggca tagttgaca ctatcgcgc 900
 tgtcaggccg catttgacc agcgaccgcg gctaateccag cgaatggtca ggcaactgcaa 960
 tgtgttggcg caaaaacagc cgactgcctg tcacctgttc aggtaacagc aattaaaga 1020
 gcgatggctg gcccagttaa ctcagcaggc accccgctgt ataatcgttg ggcattggat 1080

 gccggtatgt caggtttatc cggaactacg tataatcagg gatggcgttc ttggtggctg 1140
 gggctcttta actctagtc aaacaatgca caacgggtga gtggttttag tgcacggagc 1200
 tggcttgttg attttgcaac ccctctgag ccgatgecta tgacacaagt tgccgctcgc 1260
 atgatgaaat tcgatttca catagatccg ctgaagattt gggcgacctc aggtcagttt 1320
 acccagtcgt caatggattg gcacggcgcg acctcgactg acctcgccgc atttcgtgac 1380
 cgtggaggca aatgattct gtatcacggt atgagcgatg ctgccttttc cgcgttggac 1440
 acagcggatt ggiacgaacg tttgggcgct gcgatgccgg gcgccccgg gttcgcccga 1500

 ttattctctg tacctggaat gaatcactgt tccggtgggc cgggtacaga cegtctgat 1560
 atgctgacc cgctttagc ctgggtcgag agaggcgaag ccccggacca aatatctgcg 1620
 tggtcaggta ctctggtta ttttgggtgc gccgctcgta cacgtccatt gtgtccgtac 1680
 cctcagattg cacgttataa aggcagcggc gatattaaca ccgaagctaa tttcgttgt 1740
 gctgcaccac cctaa 1755

<210> 82

<211> 1755

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 82

```

ggaggtagta gcacgccttt acctctgccg cagcagcagc cgcctcaaca ggaaccaccc      60

ccgcctccgg tgccttagc atcccgtgcg gcctgcgaag ccctgaagga tggaaacggt      120
gacatggtat ggcccaacgc agctaccgtg gtcgaagttag ccgcttgccg cgacgctgca      180
ccagccaccg caagcgtgct gcacctgctt gaacctgctg aagtgtccgg tgccattgca      240
aaacgcacag gcatagatgg ctatccctat gagataaagt ttgcctgctg tatgcctgcc      300
gagtggaaat ggcgtttttt tatggaaggt ggaagcggaa ccaatggcag cctgagcgtt      360
gcaacaggtt caatcgccgg tggtaaatt gcctccgcc tttctgcaa ttttgagtt      420
attgctacgg atggcggcca cgataatgca gtcaacgata atccagacgc tctggggacc      480

gttgcccttg gcctggatcc gcaagcacgc ttagactatg gttataacag ttacgatcaa      540
gtcacacaag caggtaaagc cgctgtcggc cgcttttatg ggctgcccgc cgacaaatct      600
tattttatag gatgctctga aggcggacgc gaaggtatga tgctgtccca acgttttccg      660
tcgcactatg atggaattgt cgctggtgct cctggttatc aactgcccaa agccggcata      720
agcggtgctt ggacgacgca gtccctggcg ccggcagcag ttggattgga tgctcaaggc      780
gtccactga tcaataaaaag cttcagcgat gcagacctgc acttactcag tcaggcgatc      840
ttgggaactt gcgacgcgct ggatggtttg gcagatggaa ttgtggataa ttacagagct      900

tgtcaggcgg cgtttgatcc agctacagcc gcaaaccggg cgaatggaca ggccctgcag      960
tgctaggggg ccaaacggc agattgtctg tccccggttc aagtaacagc cattaagcgg      1020
gccatggcag gtcccgtaa ttctgcaggt acgccctgtt acaaccggtg ggctgggat      1080
gcgggtatgt ccggcctcag ccgaacaaca tataaccaag gttggcgtatc ttggtggctg      1140
ggtagtttca actcgagtgc aaataacgca cagagagtgt ccggttttag tgctcgttcc      1200
tggttggttg attttgccac accaccagaa cctatgccta tgaccaggt ggctgcgagg      1260
atgatgaagt ttgatttcca tattgacccc ctaagatct gggcaacgag cggccagttt      1320

acacagagca gcatggattg gcatggcgcc acctctaccg atctggcggc ttttcgtgac      1380
cgcgccggaa aaatgatttt atatcacggc atgtcggatg cagccttcag cgcgcttgat      1440
actgcggatt ggtatgaacg cttaggcgca gctatgcctg gagctgctgg ttttgccaga      1500
    
```

ctgttccttg tgccaggaat gaatcattgt agtggcggtc ccggtactga tcgctttgac 1560
atgttaacgc ctctggtggc ctgggttgag cgtggtgaag cccctgatca aatctcagcg 1620
tggctcggca cgccaggcta tttcgggtgt gcagctcgca cacgtcccct ctgtccatac 1680
cctcagatag caagatacaa aggtagtggg gacatcaata ccgaggcaaa cttcgcatgc 1740

gcggccccgc cctga 1755

<210> 83
<211> 1755
<212> DNA
<213> Artificial Sequence
<220><223> Synthetic
<400> 83

ggtggcggta gtaccccgt gccgctgccg cagcagcagc caccgcaaca ggagccacct 60
cctccgccgg tcctttggc atcccgcga gcttgtgagg ctttgaaaga cggtaatgga 120
gacatggttt ggccgaatgc tgcgactgta gtggaagtgc cggcgtggcg tgatgcagca 180
ccggcaactg cgagcgcgcg tgcgctccct gaacattgcg aggttagtgg ggctattgct 240
aagcgtaccg ggattgatgg ctatccatac gagattaaat ttcgtctgcg catgccagcg 300

gaatggaatg gtagattttt catggaaggt ggtccggca ccaatggctc gctcagcgt 360
gcaacaggct ccatcggcgg ggtcagatt gcgtcagcct tatcgcgtaa ttttgctggt 420
atcgctaccg atggcggcca cgataatgca gttaacgata atccggacgc tctggggact 480
gttgcgtttg gcttagatcc gcagccaga ctggattatg ggtataattc gtacgatcag 540
gtcacacagg caggtaaagc cgccgttgc cgcttttatg ggcgggctgc ggacaaatca 600
tattttattg gatgttcgga aggaggccgg gaaggaatga tgttgagcca acgatttct 660
agccattatg acggcattgt tgccggggca ccgggcttcc aattgcctaa agcaggaatt 720

tcaggcgcgt ggacgacca gtctctggcg cctgcggcag taggactgga cgcgcaggga 780
gtccctttga ttaataaatc cttctctgat gctgatttgc atcttttgc ccaagccatt 840
ttaggcactt gcgatgcgt ggacggctta gcagatggta tagttgataa ttatcgcgcc 900
tgccaggcgg cttttgatcc cgcaacagca gcgaacctg cgaacggtca ggcctgcaa 960
tgtgtcggtg caaaaaccgc ggactgtttg tctccggtgc aagtactgc aattaagcgt 1020
gcaatggctg gaccggtgaa ctccgctgga acaccttgt acaaccgctg ggcgtgggat 1080
gctggcatga gcggcctgct tgggacgact tataaccagg gctggcgtc ttggtggtta 1140

ggttcattta attcctccgc gaataatgcc cagcgcgtat ctggtttttc tgcccgttct 1200

tggttggtgg attttgctac gccgccagaa ccgatgccga tgacgcaagt cgcggcacga 1260
 atgatgaagt tcgactttga catagacccg ctgaaaattht gggcaacaag cggtcagttt 1320
 acgcaaagca gcatggactg gcacggagca accagcactg atctcgcggc gtttagagat 1380
 cgtgggggaa aaatgatttt atatcatggc atgtcagacg ccgctttttc cgctctggac 1440
 acagccgact ggtacgaacg gctgggagca gctatgccag gggccgctgg ttttgacgc 1500
 ttatttctgg tgccaggat gaaccattgt agtgggggccc ccggaaccga tcgctttgat 1560

atgcttacc cgctttagc gtgggtttaa cgtggtgaag ctctgatca aatttctgcc 1620
 tggagtggta cgctgggta tttcgcgctg gcggcgcgta ctgaccact gtgcccttac 1680
 cctcaaattg cgcgttataa aggtagtggg gatattaaca cagaggcga ctttgctgt 1740
 gctgccccgc cctaa 1755

<210> 84

<211> 1752

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 84

ggtggcagta caccgcttc gcttccgcaa cagcaaccgc ctcaacagga acctccgct 60
 cctcccgtgc ctctgccag ccgtgcagcc tgcgaggcgc ttaaagacgg gaatggtgac 120

atggtatggc caaacgcggc cacggtggtt gaggttgccg cgtggagaga tgctgctccg 180
 gcaacggcat ccgccgctgc gttgccagaa cactgtgagg tatcgggcgc catagccaaa 240
 cgtaccggca ttgatggcta tccatacga attaaattta gactccgcat gccagcagag 300
 tggaaaggac gttttttat ggagggtggg tctggtacaa atggcagctt aagtgccgcg 360
 actggctcca ttggcggagg gcagattgca tcagcgttgt ctctgaattt tgctgttatc 420
 gccaccgacg ggggacatga caacccgctt aacgataacc ctgacgcctt gggcactggt 480
 gcctttggtc ttgatcccca ggcgcgttta gatatgggct ataattctta cgatcaggtt 540

acacaggccg gcaaggctgc agtagcacgg ttctacggta gagccgcaga taaatcatat 600
 tttataggtt gctctgaggg tgggcgtgaa ggtatgatgc tttctcagcg cttcccctcc 660
 cattatgatg gaatagtggc tgggtctcca ggcttccagc tgccaaaagc cggcatctca 720
 ggcgcatgga ccacgcaaag cctcgtcctt gcggcagttg gtttggatgc gcaggcctg 780
 ccgttaatca ataagtcttt ctctgatgcc gatttgcacc tgctgagcca agcaatttta 840
 ggaacttgcg atgccctgga tgggttggcg gacggtatag tcgataatta tcgcatgt 900

caggccgcct tcgatccggc aaccgcagct aatcccgcca atggccaggc gctgcagtgt 960

gtaggtgcaa aaacagctga ctgtctcagt cctgtccagg taactgctat taagcgcgcg 1020

atggcaggte ctgtgaactc agcaggcacg cgttatata atcctgggc ctgggatgcg 1080

gggatgtccg ggctgtcagg taccacttat aatcagggtt ggagaagctg gtggctggga 1140

tcatttaaca gtagcgctaa caatgcacag cgcgtctctg gatttagcgc tcgctcctgg 1200

ttagtagatt ttgctactcc cccagaaccg atgccgatga cgcaggttgc tgctegaatg 1260

atgaagtttg actttgatat cgaccctctg aagatttggg cgaccagtgg tcagtttaca 1320

cagtcctcaa tggattggca tggcgtacg tctacggact tagcagcctt tcgcatcgc 1380

ggtaggaaaa tgatactgta tcacggcatg tcagatgcgg cgttagcgc gttggatagc 1440

gctgattggt acgaacgttt aggtgcagct atgccgggtg cgcaggggtt tgcgcggctg 1500

tttttggttc ccggtatgaa tcaactgttc ggaggtccgg gtaactgatcg ctttgatatg 1560

cttaccctgc ttgtagcctg ggtggaacgt ggagaggccc cggatcagat ctccgctgg 1620

agcggtagcc cgggttattt tggagtgcg gcacggactc gccctctgtg cccatatcct 1680

cagattgcta gatataaagg tagcggcgac ataatacag aagccaactt tgcatacgea 1740

gcaccacat ga 1752

<210> 85

<211> 5365

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic

<400> 85

atccggatat agttctctct ttcagcaaaa aaccctcaa gaccggtta gaggecccaa 60

ggggttatgc tagttattgc tcagcgggtg cagcagccaa ctgagcttcc tttcgggctt 120

tgtttagcagc cggatctcag tgggtgtggt ggtggtgctc gagtgcggcc gcaagcttgt 180

cgacggagct cgaattcga tccgcgacc atttgctgtc caccagtcat getagccata 240

tggctgccgc gcggcaccag gccgctgctg tgatgatgat gatgatggct gctgcccatg 300

gtatatctcc ttcttaaagt taacaaaaat tatttctaga ggggaattgt tatccgctca 360

caattcccct atagtgagtc gtattaat tgcgggatcg agatctgat cctctacgcc 420

ggacgcacgc tggccggcat caccggcgc acaggtgcgg ttgctggcgc ctatatcgcc 480

gacatcaccg atggggaaga tcgggctcgc cacttcgggc tcatgagcgc ttgtttcgcc 540

gtgggtatgg tggcaggccc cgtggccggg ggactgttgg gcgccatctc cttgcatgca 600

ccattccttg cggcggcggg gctcaacggc ctcaacctac tactgggctg cttcctaattg 660
caggagtgcg ataagggaga gcgtcgagat cccggacacc atcgaatggc gcaaaacctt 720
tcgcggtatg gcatgatagc gcccgggaaga gagtcaattc aggggtgggga atgtgaaacc 780

agtaacgtta tacgatgtcg cagagtatgc cggtgtctct tatcagaccg tttcccgcgt 840
ggatgaaccag gccagccacg tttctgcaaa aacgcgggaa aaagtggaag cggcgatggc 900
ggagctgaat tacattccca accgcgtggc acaacaactg gcgggcaaac agtcgttgct 960
gattggcgtt gccacctcca gtctggccct gcacgcgccg tcgcaaattg tcgcgcgat 1020
taaatctcgc gccgatcaac tgggtgccag cgtgggtggg tcgatggtag aacgaagcgg 1080
cgtcgaagcc tgtaaagcgg cgggtcacaaa tcttctcgcg caacgcgtca gtgggctgat 1140
cattaactat ccgctggatg accaggatgc cattgctgtg gaagctgcct gcactaatgt 1200

tccggcggtta tttcttgatg tctctgacca gacacccatc aacagtatta ttttctccca 1260
tgaagacggt acgcgactgg gcgtggagca tctggctgca ttgggtcacc agcaaatcgc 1320
gctgttagcg ggcccattaa gttctgtctc ggcgctctg cgtctggctg gctggcataa 1380
atatctcact cgcaatcaaa ttcagccgat agcgggaacgg gaaggcgact ggagtccat 1440
gtcccgtttt caacaaacca tgcaaatgct gaatgagggc atcgttccca ctgcgatgct 1500
ggttgccaac gatcagatgg cgctgggcgc aatgcgcgcc attaccgagt ccgggctgcg 1560
cgttgggtgcg gatatctcgg tagtgggata cgacgatacc gaagacagct catgttatat 1620

cccgccgtta accaccatca aacaggattt tcgcctgctg gggcaaacca gcgtggaccg 1680
cttgcctgca ctctctcagg gccaggcggg gaagggcaat cagctgttgc cegtctcact 1740
ggtgaaaaga aaaaccacc tggcgcccaa tacgaaacc gcctctccc gcgcgttggc 1800
cgattcatta atgcagctgg cagcacaggt ttcccactg gaaagcgggc agtgagcgca 1860
acgcaattaa tgtaagttag ctcactcatt aggcaccggg atctcgacc atgcccttga 1920
gagccttcaa ccagtcagc tcttccggt gggcgcgggg catgactatc gtcgccgac 1980
ttatgactgt cttctttatc atgcaactcg taggacaggt gccggcagcg ctctgggtca 2040

ttttcggcga ggaccgcttt cgctggagcg cgacgatgat cggcctgtcg cttgcggtat 2100
tcggaatctt gcacgcctc gctcaagcct tcgtcactgg tcccgcacc aaacgtttcg 2160
gcgagaagca ggccattatc gccggcatgg cggccccacg ggtgcgcatg atcgtgctcc 2220
tgtcgttgag gaccggccta ggctggcggg gttgccttac tggtagcag aatgaatcac 2280
cgatacgcga gcgaacgiga agcgactgct gctgcaaac gtctgcgacc tgagcaaaaa 2340
catgaatggt cttcggtttc cgtgtttcgt aaagtctgga aacgcggaag tcagcgcct 2400

gcaccattat gttccggatc tgcacgcag gatgctgctg gctaccctgt ggaacaccta 2460

catctgtatt aacgaagcgc tggcattgac cctgagtgat ttttctctgg tcccgcgcga 2520

tccataaccgc cagttgttta ccctcacaac gttccagtaa ccgggcatgt tcatcatcag 2580

taaccctgat cgtgagcadc ctctctcgtt tcatcggat cattaccccc atgaacagaa 2640

atccccctta cacggaggca tcaatgacca aacaggaaaa aaccgcccctt aacatggccc 2700

gctttatcag aagccagaca ttaacgcttc tggagaaaact caacgagctg gacgcggatg 2760

aacaggcaga catctgtgaa tcgcttcacg accacgctga tgagctttac cgcagctgcc 2820

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtea 2880

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 2940

ttggcgggtg tcggggcgca gccatgacc agtcacgtag cgatagcgga gtgtatactg 3000

gcttaactat gcggcatcag agcagattgt actgagagtg caccattgcg gtgtgaaata 3060

ccgcacagat gcgtaaggag aaaataccgc atcaggcgt cttccgcttc ctgctcact 3120

gactcgtgc gctcggctgt tcggctgcgg cgagcggat cagctcactc aaaggcgta 3180

atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag 3240

caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc 3300

cctgacgagc atcacaaaaa tcgacgtca agtcagaggt ggcgaaacc gacaggacta 3360

taaagatacc aggcgtttcc ccctggaagc tcctcgtgc gctctcctgt tccgaccctg 3420

ccgcttaccg gatacctgac gcctttctc ccttcgggaa gcgtggcgct ttctcatagc 3480

tcacgtgta ggtatctcag ttccgtgtag gtcgttcgct ccaagctggg ctgtgtgcac 3540

gaacccccg ttacgccga ccgctgcgcc ttatccgta actatcgtct tgagtccaac 3600

ccgtaagac acgacttacc gccactggca gcgccactg gtaacaggat tagcagagcg 3660

aggtatgtag gcggtgctac agagtctctg aagtgggtgc ctaactacgg ctacactaga 3720

aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagtgggt 3780

agctcttgat ccggcaaaaa aaccaccgct ggtagcgggt gttttttgt ttgcaagcag 3840

cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatctttc tacgggtctt 3900

gacgctcagt ggaacgaaaa ctacagttaa gggattttg tcatgaaca taaaactgac 3960

tgcttacata aacagtaata caaggggtgt tatgagccat attcaacggg aaacgtcttg 4020

ctctaggccg cgattaaatt ccaacatgga tgctgattta tatgggtata aatgggctcg 4080

cgataatgac gggcaatcag gtgcgacaat ctatcgattg tatgggaagc ccgatgcgcc 4140

agagttgttt ctgaaacatg gcaaaggtag cgttgccaat gatgttacag atgagatggt 4200
 cagactaaac tggtgacgg aatttatgcc tcttccgacc atcaagcatt ttatccgtac 4260
 tcctgatgat gcatggttac tcaccactgc gatccccggg aaaacagcat tccaggtatt 4320
 agaagaatat ccigattcag gtgaaaatat tgttgatgcg ctggcagtgt tcctgcgccg 4380
 gttgcattcg attcctgttt gtaattgtcc ttttaacagc gatcgcgtat ttcgtctcgc 4440
 tcaggcgcaa tcacgaatga ataacggttt ggttgatgcg agtgattttg atgacgagcg 4500
 taatggctgg ccigtgaac aagtctggaa agaaatgcat aaacttttgc cattctcacc 4560

 ggattcagtc gtcactcatg gtgatttctc acttgataac cttatTTTTg acgaggggaa 4620
 attaataggt tgiattgatg ttggacgagt cggaatcgca gaccgatacc aggatcttgc 4680
 catcctatgg aactgcctcg gtgagttttc tccttcatta cagaaacggc tttttcaaaa 4740
 atatggtatt gataatcctg atatgaataa attgcagttt catttgatgc tcgatgagtt 4800
 tttctaagaa ttaattcatg agcggataca tatttgaatg tatttagaaa aataa caaaa 4860
 taggggttcc ggcacattt ccccgaaaag tgccacctaa attgtaagcg ttaatatTTT 4920
 gttaaaattc gcgttaaatt tttgttaaat cagctcattt tttaaccaat aggccgaaat 4980

 cggcaaaatc cctataaat caaaagaata gaccgagata gggttgagtg ttgttccagt 5040
 ttggaacaag agtccactat taaagaacgt ggactccaac gtcaaagggc gaaaaaccgt 5100
 ctatcagggc gatggccac tacgtgaacc atcacctaa tcaagTTTT tggggtcgag 5160
 gtgccgtaaa gactaaatc ggaaccctaa agggagcccc cgatttagag cttgacgggg 5220
 aaagccggcg aacgtggcga gaaaggaagg gaagaaagcg aaaggagcgg gcgctagggc 5280
 gctggcaagt gtagcggta cgctgcgctg aaccaccaca cccgccgcgc ttaatgcgcc 5340
 gctacagggc gcgtccatt cgcca 5365