

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0094812
(43) 공개일자 2009년09월08일

(51) Int. Cl.

C12N 15/52 (2006.01) C12N 5/10 (2006.01)
C12P 21/04 (2006.01) C07K 14/00 (2006.01)

(21) 출원번호 10-2009-7009346

(22) 출원일자 2007년10월17일
심사청구일자 없음

(85) 번역문제출일자 2009년05월06일

(86) 국제출원번호 PCT/US2007/022232

(87) 국제공개번호 WO 2008/073184
국제공개일자 2008년06월19일

(30) 우선권주장

60/853,008 2006년10월18일 미국(US)

60/923,458 2007년04월12일 미국(US)

(71) 출원인

더 스크립스 리서치 인스티튜트

미국 캘리포니아주 92037 라 졸라 노스 토리 파
인 로드 10550

(72) 발명자

리우 웬쉬

미국 텍사스주 77840 칼리지 스테이션 #2099 볼
스트리트 301

숄츠 피터 지

미국 캘리포니아주 92037 라 졸라 라 졸라 란초
로드 1650

(74) 대리인

김성기, 김진희

전체 청구항 수 : 총 29 항

(54) 포유동물 세포에서 비천연 아미노산의 단백질 내로의 유전적 도입

(57) 요약

본 발명은 비천연 아미노산을 포유동물 숙주 세포, 예를 들면 영장류 숙주 세포 및 설치류 숙주 세포에서 단백질 내로 도입할 수 있는 tRNA 및 아미노아실-tRNA 합성효소의 오르소고달 쌍에 관한 것이다. 본 발명은, 예를 들면 단 비제한적인 예로서 숙주 세포(예, 영장류 또는 설치류 세포), 진정세균 합성효소에서 유도된 오르소고달 아미노아실-tRNA 합성효소, 오르소고달 tRNA, 및 비천연 아미노산을 포함하는 번역 시스템을 제공한다. 본 발명은 또한 포유동물 숙주 세포 시스템에서 하나 이상의 비천연 아미노산을 포함하는 관심 단백질의 제조 방법에 관한 것이다.

특허청구의 범위

청구항 1

- (a) i) p-메톡시페닐알라닌 (pMpa);
- ii) p-아세틸페닐알라닌 (pApa);
- iii) p-벤조일페닐알라닌 (pBpa);
- iv) p-요오도페닐알라닌 (pIpa);
- v) p-아지도페닐알라닌 (pAzpa);
- vi) p-프로파르길옥시페닐알라닌 (pPpa);
- vii) α-아미노카프릴산;
- viii) o-니트로벤질시스테인 (o-NBC);
- ix) 1,5-단실알라닌; 및
- x) o-니트로벤질세린 (o-NBS)

으로 이루어진 군에서 선택된 제1 비천연 아미노산;

(b) 제1 오르소고날 tRNA (O-tRNA);

(c) 진정세균 아미노아실-tRNA 합성효소에서 유도된 제1 오르소고날 아미노아실-tRNA 합성효소(O-RS)(여기서, 상기 O-RS는 상기 O-tRNA를 상기 비천연 아미노산으로 우선적으로 아미노아실화함); 및

(d) (a), (b) 및 (c)를 포함하고, 설치류 세포 및 영장류 세포에서 선택되는 숙주 세포를 포함하는 번역 시스템.

청구항 2

제1항에 있어서, 상기 제1 O-RS는, 상기 O-tRNA, 상기 비천연 아미노산, 및 서열 번호 57~101에서 선택된 아미노산 서열을 포함하는 아미노아실-tRNA 합성효소를 포함하는 번역 시스템에 대해 관찰되는 효율의 50% 이상인 효율로 상기 제1 O-tRNA를 상기 제1 비천연 아미노산으로 우선적으로 아미노아실화하는 번역 시스템.

청구항 3

제1항에 있어서, 상기 제1 O-tRNA는 서열 번호 3에 제시된 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩되는 번역 시스템.

청구항 4

제1항에 있어서, 상기 제1 O-RS는 이.콜라이(*E. coli*) 아미노아실-tRNA 합성효소에서 유도되는 번역 시스템.

청구항 5

제1항에 있어서, 상기 제1 O-RS는 서열 번호 57~101에 제시된 아미노산 서열의 군에서 선택된 아미노산 서열, 및 이의 보존적 변이체를 포함하는 번역 시스템.

청구항 6

제1항에 있어서, 상기 숙주 세포는 상기 제1 O-tRNA에 의해 인식되는 하나 이상의 셀렉터 코돈을 포함하고, 관심 단백질을 코딩하는 핵산을 추가로 포함하는 번역 시스템.

청구항 7

제6항에 있어서, 상기 숙주 세포는 제1 비천연 아미노산과 상이한 제2 비천연 아미노산, 제2 O-RS 및 제2 O-tRNA를 추가로 포함하고, 여기서 제2 O-RS는 제2 O-tRNA를 제2 비천연 아미노산으로 우선적으로 아미노아실화하고, 제2 O-tRNA는 제1 O-tRNA에 의해 인식되는 셀렉터 코돈과 상이한 핵산에 의해 코딩되는 셀렉터 코돈을 인식

하는 번역 시스템.

청구항 8

제1항에 있어서, 상기 숙주 세포는 상기 제1 O-RS를 코딩하는 폴리뉴클레오티드를 포함하는 번역 시스템.

청구항 9

제8항에 있어서, 상기 폴리뉴클레오티드는 서열 번호 8~56에 제시된 뉴클레오티드 서열에서 선택된 뉴클레오티드 서열을 포함하는 번역 시스템.

청구항 10

제1항에 있어서, 상기 숙주 세포는 상기 제1 O-tRNA를 코딩하는 폴리뉴클레오티드를 포함하는 번역 시스템.

청구항 11

제1항에 있어서, 상기 설치류 숙주 세포는 래트 숙주 세포 및 마우스 숙주 세포에서 선택하는 번역 시스템.

청구항 12

제1항에 있어서, 상기 영장류 숙주 세포는 인간 숙주 세포, 침팬지 숙주 세포, 보노보 숙주 세포, 고릴라 숙주 세포, 오랑우탄 숙주 세포, 긴팔원숭이 숙주 세포, 짧은꼬리원숭이 숙주 세포, 타마린 숙주 세포 및 명주원숭이 숙주 세포로 이루어진 군에서 선택하는 번역 시스템.

청구항 13

제1항에 있어서, 상기 설치류 숙주 세포는 CHO 숙주 세포인 번역 시스템.

청구항 14

제1항에 있어서, 상기 영장류 숙주 세포는 293T 숙주 세포인 번역 시스템.

청구항 15

선택된 위치에 비천연 아미노산을 포함하는 단백질의 생성 방법으로서,

- (a) (i) A) p-메톡시페닐알라닌 (pMpa);
- B) p-아세틸페닐알라닌 (pApa);
- C) p-벤조일페닐알라닌 (pBpa);
- D) p-요오도페닐알라닌 (pIpa);
- E) p-아지도페닐알라닌 (pAzpa);
- F) p-프로파르길옥시페닐알라닌 (pPpa);
- G) α-아미노카프릴산;
- H) o-니트로벤질시스테인 (o-NBC);
- I) 1,5-단실알라닌; 및
- J) o-니트로벤질세린 (o-NBS)

으로 이루어진 군에서 선택된 제1 비천연 아미노산;

(ii) 제1 오르소고달 tRNA (O-tRNA);

(iii) 제1 오르소고달 아미노아실-tRNA 합성효소(O-RS)(여기서, 상기 제1 O-RS는 상기 제1 O-tRNA를 상기 비천연 아미노산으로 우선적으로 아미노아실화함);

(iv) 상기 단백질을 코딩하는 핵산(여기서, 상기 핵산은 상기 제1 O-tRNA에 의해 인식되는 하나 이상의 선택터 코돈을 포함하고, 상기 핵산 내 선택터 코돈의 위치는 상기 단백질 내 선택된 위치와 상응함); 및

(v) (i), (ii), (iii) 및 (iv)를 포함하고, 설치류 세포 및 영장류 세포에서 선택되는 숙주 세포를 제공하는 단계; 및

(b) 상기 비천연 아미노산, O-tRNA 및 O-RS의 존재 하에 상기 선택된 코돈에 대한 반응으로 상기 단백질의 번역 동안 상기 단백질 내 상기 선택된 위치에 상기 비천연 아미노산을 도입함으로써, 선택된 위치에 상기 비천연 아미노산을 포함하는 상기 단백질을 생성시키는 단계를 포함하는 방법.

청구항 16

제15항에 있어서, 상기 제1 O-RS는, 상기 O-tRNA, 상기 비천연 아미노산, 및 서열 번호 57~101의 아미노산 서열의 군에서 선택된 아미노산 서열을 포함하는 아미노아실-tRNA 합성효소를 포함하는 번역 시스템에 대해 관찰되는 효율의 50% 이상인 효율로 상기 제1 O-tRNA를 상기 제1 비천연 아미노산으로 우선적으로 아미노아실화하는 방법.

청구항 17

제15항에 있어서, 상기 제1 O-tRNA는 서열 번호 3에 제시된 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩되는 방법.

청구항 18

제15항에 있어서, 제1 O-RS의 제공은 상기 제1 O-RS를 코딩하는 폴리뉴클레오티드를 제공하는 것을 포함하는 방법.

청구항 19

제15항에 있어서, 상기 제1 O-RS는 이.콜라이 아미노아실-tRNA 합성효소에서 유도되는 방법.

청구항 20

제15항에 있어서, 상기 제1 O-RS는 서열 번호 57~101에 제시된 아미노산 서열의 군에서 선택된 아미노산 서열, 및 이의 보존적 변이체를 포함하는 방법.

청구항 21

제15항에 있어서, 상기 제1 O-RS의 제공은 상기 O-RS를 코딩하는 폴리뉴클레오티드를 제공하는 것을 포함하는 방법.

청구항 22

제21항에 있어서, 상기 폴리뉴클레오티드는 서열 번호 8~56에 제시된 뉴클레오티드 서열에서 선택된 뉴클레오티드 서열 세트를 포함하는 방법.

청구항 23

제15항에 있어서, 상기 제1 O-RS의 제공은 부위-지정 돌연변이유발에 의해 야생형 아미노아실-tRNA 합성효소의 아미노산 결합 포켓을 돌연변이화하고, 상기 제1 O-tRNA를 상기 제1 비천연 아미노산으로 우선적으로 아미노아실화하는 생성된 O-RS를 선택하는 것을 포함하는 방법.

청구항 24

제23항에 있어서, 상기 선택 단계는 부위-지정 돌연변이유발 후에 생성된 복수개의 아미노아실-tRNA 합성효소 분자를 포함하는 풀(pool)로부터 상기 제1 O-RS에 대해 양성 선택 및 음성 선택을 하는 것을 포함하는 방법.

청구항 25

제15항에 있어서, 상기 제1 O-tRNA의 제공은 상기 제1 O-tRNA를 코딩하는 폴리뉴클레오티드를 제공하는 것을 포함하는 방법.

청구항 26

제15항에 있어서, 상기 단백질은 상기 제1 비천연 아미노산과 상이한 제2 비천연 아미노산을 포함하고, 상기 방법은 상기 제2 비천연 아미노산, 제2 O-RS 및 제2 O-tRNA를 추가로 제공하며, 제2 O-RS는 제2 O-tRNA를 제2 비천연 아미노산으로 우선적으로 아미노아실화하고, 제2 O-tRNA는 제1 O-tRNA에 의해 인식되는 셀렉터 코돈과 상이한 핵산 내 셀렉터 코돈을 인식하는 방법.

청구항 27

제15항에 있어서, 상기 도입 단계는 상기 숙주 세포를 배양하는 것을 포함하는 방법.

청구항 28

제15항에 있어서, 제1 비천연 아미노산은 광케이징된 아미노산을 포함하고, 상기 방법은 숙주 세포를 UV 광에 노출시키는 것을 추가로 포함하는 방법.

청구항 29

제28항에 있어서, 숙주 세포를 UV 광에 노출시키는 것은 숙주 세포 내 단백질 활성의 공간적 제어를 제공하는 방법.

명세서

기술 분야

<1> 관련 출원의 상호 참조

<2> 본 출원은 미국 가출원 번호 60/853,008(2006년 10월 18일 출원); 및 가출원 번호 60/923,458(2007년 4월 12일 출원)을 우선권으로 주장하며, 이의 공개내용은 그 전문이 본원에 참고인용된다.

<3> 연방 지원 연구 및 개발 하에 이루어진 발명에 대한 권리에 관한 진술

<4> 본 발명은 허가 제GM62159호 하에 미국립보건원으로부터 정부 지원을 받아 이루어졌다. 미 정부는 본 발명에 대한 특정 권리를 가질 수 있다.

<5> 기술 분야

<6> 본 발명은 번역 생화학의 분야에 속한다. 본 발명은 비천연 아미노산을 단백질 내로 도입하는, 오르소고날(orthogonal) tRNA, 오르소고날 아미노아실-tRNA 합성효소, 및 이들의 쌍의 제조 및 사용을 위한 조성물 및 방법에 관한 것이다. 본 발명은 또한 상기와 같은 쌍을 이용하여 세포 내에 단백질을 생성시키는 방법, 및 그 방법에 의해 제조된 단백질에 관한 것이다.

배경 기술

<7> 포유동물 세포에서 단백질 내에 부위 특이적으로 비천연 아미노산을 도입하는 것에 대한 방법들이 이전에 기재된 바 있다. 화학적으로 아미노아실화된 억제자 tRNA를 CHO 세포 및 뉴런에 각각 미세주사하거나 전기침투시켰고, 일련의 비천연 아미노산으로 넌센스 앰버(nonsense amber) 돌연변이를 억제시키는데 사용하였다(Monahan 등(2003), "Site-specific incorporation of unnatural amino acids into receptors expressed in Mammalian cells," *Chem Biol* 10:573-580). 하지만, 화학량론적 시약으로 아미노아실화된 tRNA를 사용하면 생성될 수 있는 단백질 양이 심각하게 제한된다.

<8> 대안적으로, 숙주 tRNA, aaRS 또는 아미노산(오르소고날 tRNA/aaRS)과 교차 반응하지 않는 이중성 억제자 tRNA/aaRS 쌍을 조작하여 단백질에 선택적으로 비천연 아미노산을 도입시켰다. 예를 들어, Yokoyama 및 동료들은 바실러스 스테아로써모필러스 앰버 억제자 tRNA^{tyr}_{CUA}(BstRNA^{tyr}_{CUA}) 및 이.콜라이(*E. coli*) 티로실-tRNA 합성효소(EcTyrRS)를 변형하여 CHO 세포에서 단백질에 3-요오도-L-티로신을 도입하였다(Sakamoto 등(2002), "Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells," *Nucleic Acids Res* 30:4692-4699). 유사하게, Zhang 및 동료들은 오르소고날 바실러스 서브틸리스 억제자 tRNA/트립토판-tRNA 합성효소 쌍을 조작하여 적합도가 높은 포유동물 세포에서 단백질에 5-히드록시트립토판을 도입하였다(Zhang 등

(2004), "Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells," *Proc Natl Acad Sci USA* **101**:8882-8887).

<9> 하지만, 측쇄가 야생형 기질의 것과 유의적으로 상이한 아미노산을 아미노아실화하는 aaRS 변이체를 생성하기 위해 구조를 기초로 한 돌연변이화를 이용하는 것은 연역적으로 예측하기 어려운 다중 활성 부위 잔기의 돌연변이를 필요로 한다.(Zhang 등(2002), "Structure-based design of mutant *Methanococcus jannaschii* tyrosyl-tRNA synthetase for incorporation of O-methyl-L-tyrosine," *Proc Natl Acad Sci U S A* **99**:6579-6584; Turner 등(2005), "Structural characterization of a p-acetylphenylalanyl aminoacyl-tRNA synthetase," *J Am Chem Soc* **127**:14976-14977; Turner 등(2006), "Structural plasticity of an aminoacyl-tRNA synthetase active site," *Proc Natl Acad Sci USA* **103**:6483-6488). 또한, 조작된 돌연변이체는, 이의 동족체 tRNA^{Tyr}_{CUA}를 3-요오도-L-티로신으로 충전시키는 돌연변이체 aaRS를 이용하는 경우에서와 같이, 숙주 아미노산을 여전히 인식할 수 있다(Sakamoto 등(2002), "Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells," *Nucleic Acids Res* **30**:4692-4699; 및 Kiga 등(2002), "An engineered *Escherichia coli* tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system," *Proc Natl Acad Sci U S A* **99**:9715-9720).

<10> 대안적으로, 포유동물 세포에서 직접적으로 특이성이 변경된 aaRS를 진화시키려는 시도를 할 수 있다. 예를 들어, Wang 및 동료들은 최근 인간 B 세포주에서 체세포 초돌연변이를 이용하여 광안정성 및 원적외선 발광이 향상된 단량체 적색 형광 단백질을 직접적으로 진화시켰다(Wang 등(2004), "Evolution of new nonantibody proteins via iterative somatic hypermutation," *Proc Natl Acad Sci USA* **101**:16745-16749). 하지만, 체세포 초돌연변이는 전체 단백질에 무작위 돌연변이를 도입하며, 이는 기질 특이성이 변경된 변이체를 진화시키는 경우 활성 부위의 표적화된 돌연변이유발에 의해 생성되는 유전적 다양성보다 덜 효과적일 수 있다. 하지만, 후자는 포유동물 세포에서 안정한 거대 라이브러리를 발생시키는데 어려움이 있어 한정된다.

<11> 오르소고날 번역 기법

<12> 원핵생물 및 진핵생물 유기체 모두에 있어 단백질에 다양한 비천연 아미노산을 생체내 부위 특이적으로 도입하기 위한 일반적 방법론이 개발되었다. 이 방법은 생체내 폴리펩티드 번역 중에 소정의 위치에 있는 원하는 비천연 아미노산을 삽입하기 위해 적당한 셀렉터 코돈을 인식하는 오르소고날 단백질 번역 성분에 의존한다. 이 방법은 셀렉터 코돈(예, 닌센스 엠버 코돈)을 인식하는 오르소고날 tRNA(O-tRNA)를 이용하는데, 여기에서 상응하는 특이적 오르소고날 아미노아실-tRNA 합성효소(O-RS)는 특이적으로 O-tRNA를 비천연 아미노산으로 충전한다. 이 성분은 숙주 유기체 내의 내인성 tRNA 또는 RS 중 어느 것과도 상호 작용하지 않는다(즉, 조작된 tRNA 및 RS는 오르소고날이어야 함).

<13> 이.콜라이 숙주 시스템에서 이러한 기법을 사용하여, 작용성 엠버 및 프레임이동 억제자 tRNA/aaRS 쌍을 메타노코커스 잔나쉬이(*Methanococcus jannaschii*) tRNA^{Tyr}/TyrRS 쌍, 고세균 tRNA^{Glu}/피로코커스 호리코쉬이(*Pyrococcus horikoshii*) 글루타밀-tRNA 합성효소 쌍, 및 고세균 tRNA^{Lys}/피로코커스 호리코쉬이 리실-tRNA 합성효소 쌍으로부터 유도하였다. 사카로미세스 세레비시예(*Saccharomyces cerevisiae*; *S. cerevisiae*), 작용성 tRNA_{CUA}/aaRS 쌍을 상응한 이.콜라이 tRNA^{Tyr}/TyrRS 및 tRNA^{Leu}/류실-tRNA 합성효소 쌍으로부터 유도하였다. 양성 및 음성 선택의 조합을 이용한 이러한 억제자 tRNA/aaRS 쌍의 방향성 진화로 인해 이.콜라이 및 에스.세레비시예에서 다량의 다양한 비천연 아미노산을 효율적이고 고도로 선택적으로 생체내 도입할 수 있다. 이들은 형광성 글리코실화, 황산화, 금속 이온 결합 및 산화환원-활성 아미노산, 및 새로운 화학적, 광화학적 반응성을 갖는 아미노산을 포함한다. 이러한 방법론은 시험관내 및 생체내 단백질 구조 및 기능을 조사하고, 새롭게나 향상된 성질을 갖는 단백질, 예컨대 치료적 단백질을 생성하기 위한 강력한 도구를 제공한다. 포유동물 세포, 예컨대 영양류 및 설치류 세포주에 대한 이러한 방법론의 연장은 상기 기법의 유용성을 유의적으로 향상시킬 것이다.

<14> 하나 이상의 비천연 아미노산을 포함하는 단백질의 생체내 제조에 적당한 오르소고날 번역 시스템을 이용하는 실무가 일반적으로 당업계에 공지되어 있다. 예를 들어, 국제 공개 번호 제WO 2002/086075호(발명의 명칭: "METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGONAL tRNA-AMINOACYL-tRNA SYNTHETASE PAIRS"); 제WO 2002/085923호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS"); 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE"); 제WO 2005/019415호(2004년 7월 7일 출원); 제WO

2005/007870호(2004년 7월 7일 출원); 제WO 2005/007624호(2004년 7월 7일 출원); 및 제WO 2006/110182호(2005년 10월 27일 출원)(발명의 명칭: "ORTHOGONAL TRANSLATION COMPONENTS FOR THE VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")를 참조한다. 이들 출원은 각각 전문이 본원에 참조 인용된다.

<15> 오르소고날 번역 시스템의 추가적인 논의가 또한 예를 들어 문헌[Wang 등(2001), "Expanding the genetic code of Escherichia coli," *Science* 292:498-500; Wang 및 Schultz(2002), "Expanding the Genetic Code," *Chem. Commun. (Camb.)* 1:1-11; Alfonta 등(2003), "Site-specific Incorporation of a Redox-Active Amino Acid into Proteins," *J Am Chem Soc* 125:14662-14663; Santoro 등(2003), "An archaeobacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in Escherichia coli," *Nucleic Acids Res* 31:6700-6709; Chin 등(2003), "An expanded eukaryotic genetic code," *Science* 301, 964-967; Chin 등(2003), "Progress toward an expanded eukaryotic genetic code," *Chem Biol* 10, 511-519; 및 Wu 등(2004), "A genetically encoded photocaged amino acid," *J Am Chem Soc* 126, 14306-14307; Summerer 등(2006), "A Genetically Encoded Fluorescent Amino acid," *PNAS* 103(26):9785-9789; Anderson 등(2004), "An expanded genetic code with a functional quadruplet codon," *Proc Natl Acad Sci U S A* 101:7566-7571; Zhang 등(2004), "A new strategy for the synthesis of glycoproteins," *Science* 303, 371-373; Wang 및 Schultz "Expanding the Genetic Code," *Angewandte Chemie Int. Ed.*, 44(1):34-66 (2005); Xie 및 Schultz, "An Expanding Genetic Code," *Methods* 36(3):227-238 (2005); Xie 및 Schultz, "Adding Amino acid to the Genetic Repertoire," *Curr. Opinion in Chemical Biology* 9(6):548-554 (2005); Wang 등, "Expanding the Genetic Code," *Annu. Rev. Biophys. Biomol. Struct.*, 35:225-249 (2006); Xie 및 Schultz (2006), "A Chemical Toolkit for Proteins - an Expanded Genetic Code," *Nat. Rev. Mol. Cell Biol.*, 7(10):775-782; Summerer 등(2006), "A Genetically Encoded Fluorescent Amino acid," *Proc Natl Acad Sci U S A* 103, 9785-9789; Wang 등(2006), "A Genetically Encoded Fluorescent Amino Acid," *J Am Chem Soc* 128, 8738-8739; 및 Liu 및 Schultz (2006), "Recombinant Expression of Selectively Sulfated Proteins in E.coli," *Nat Biotechnol.*, 24(11):1436-1440]에서 발견된다. 이들 공개 문헌의 각 내용은 참조 인용된다.

<16> 목적하는 비천연 아미노산이 소정 위치에 도입되는 포유동물 세포 숙주 시스템, 예컨대 영장류 및 설치류 숙주 세포 시스템에서 단백질에 비천연 아미노산을 도입하는 향상된 오르소고날 번역 성분의 개발이 당업계에 필요한 실정이다. 포유동물 세포, 예컨대 설치류 세포 및 인간 세포에서 작용할 수 있는 오르소고날 번역 성분(예, 돌연변이체 아미노아실-tRNA 합성효소)을 선별하고 동정하는 방법의 향상이 당업계에 필요하다. 본원에 기재된 본 발명은 상기 필요 및 다른 필요를 충족하며, 이는 하기 내용을 검토할 때 자명하게 될 것이다.

발명의 상세한 설명

<17> **발명의 개요**

<18> 본 발명은 다양한 물리화학적, 생물학적 성질을 갖는 비천연 아미노산이 포유동물 세포, 예컨대 설치류 세포 및 영장류 세포에서 유전적으로 코딩되도록 하는 일반적인 접근법을 제공한다. 먼저 돌연변이체 에셰리키아 콜라이(*Escherichia coli*; *E. coli*) 아미노아실-tRNA 합성효소(RS)를 효모에서 진화시켜 관심 비천연 아미노산을 선택적으로 이용하도록 한다. 이후 진정세균, 예컨대 바실러스 스테아로써모필러스(*B. stearothermophilus*) 유래의 엠버 억제자 tRNA와 함께 돌연변이체 RS를 사용하여 엠버 넌센스 코돈에 대한 반응으로 포유동물 세포에서 단백질에 비천연 아미노산을 부위 특이적으로 도입한다. 10개의 비천연 아미노산(도 1에 제공된 비천연 아미노산)을, 2×10^7 세포 당 최대 1 μ g 단백질의 효율로 중국 햄스터 난소(CHO) 세포 또는 인간 293T 세포에서 발현되는 모델 단백질에 독립적으로 도입하였다. 질량 분석법은 비천연 아미노산을 위한 높은 번역적 적합도를 확인하였다. 이러한 방법론은 세포 연구를 위한 단백질로의 생물학적 프로브의 도입을 용이하게 할 수 있고, 궁극적으로 포유동물 세포, 예컨대 설치류 세포 및 영장류 세포에서 비천연 아미노산을 함유하는 치료적 단백질의 합성을 용이하게 할 수 있다.

<19> 본 발명은 생체내 숙주 세포, 예컨대 설치류 숙주 세포 또는 영장류 숙주 세포에서 셀렉터 코돈, 예를 들면 엠버 정지 코돈에 대한 반응으로 성장하는 폴리펩티드 쇠에 비천연 아미노산을 도입하기 위한 조성물 및 방법을 제공한다. 이러한 조성물은 숙주 세포 번역 기구(machinery)와 상호작용하지 않는, 오르소고날-tRNA(O-tRNA)와 오르소고날 아미노아실-tRNA 합성효소(O-RS)의 쌍과 마찬가지로 숙주 세포를 포함한다. 즉, O-tRNA는 내인성 숙주 세포 아미노아실-tRNA 합성효소에 의해 (천연 또는 비천연) 아미노산으로 충전되지 않는다(혹은 유의적 수준으로 충전되지 않는다). 유사하게, 본 발명에 의해 제공되는 O-RS는, 유의적 수준 또는 검출가능한 수준으로 어떠한 내인성 tRNA도 (천연 또는 비천연) 아미노산으로 충전하지 않는다. 이러한 신규 조성물은 포유동물 숙주

세포 시스템에서 번역적으로 도입된 비천연 아미노산을 갖는 다량의 단백질이 생성되도록 한다.

- <20> 일부 측면에서, 본 발명은 번역 시스템을 제공한다. 이러한 시스템은 (a) 포유동물 숙주 세포, 예컨대 설치류 또는 영장류 숙주 세포를 포함하고, 상기 숙주 세포는 (b) 제1 오르소고날 아미노아실-tRNA 합성효소(O-RS), (c) 제1 오르소고날 tRNA(O-tRNA), 및 (d) 제1 비천연 아미노산을 포함하는데, 여기에서 제1 O-RS는 제1 O-tRNA를 제1 비천연 아미노산으로 우선적으로 아미노아실화한다. 비천연 아미노산은, 비제한적인 예로서,
- <21> p-메톡시페닐알라닌 (pMpa);
- <22> p-아세틸페닐알라닌 (pApa);
- <23> p-벤조일페닐알라닌 (pBpa);
- <24> p-요오도페닐알라닌 (pIpa);
- <25> p-아지도페닐알라닌 (pAzpa);
- <26> p-프로파르길옥시페닐알라닌 (pPpa);
- <27> α-아미노카프릴산;
- <28> o-니트로벤질시스테인 (o-NBC);
- <29> 1,5-단실알라닌; 및
- <30> o-니트로벤질세린 (o-NBS)
- <31> 일 수 있다.
- <32> 일부 측면에서, O-RS는, 동일한 O-tRNA, 비천연 아미노산, 및 서열 번호 57~101에서 선택된 아미노산 서열을 포함하는 아미노아실-tRNA 합성효소를 포함하는 번역 시스템에 대해 관찰되는 효율의 50% 이상인 효율로 상기 비천연 아미노산으로 O-tRNA를 우선적으로 아미노아실화한다.
- <33> 번역 시스템은 각종 공급원에서 유도된 성분을 사용할 수 있다. 일 구체예에서, 제1 O-RS는 이.콜라이 아미노아실-tRNA 합성효소, 예컨대 야생형 이.콜라이 티로실 또는 류실-tRNA 합성효소에서 유도된다. 일부 구체예에서, O-tRNA는 서열 번호 3을 포함하거나 그 서열에 의해 코딩된다. 상기 시스템에 사용된 O-RS는 서열 번호 57~101의 아미노산 서열, 및 상기 서열의 보존적 변이체를 포함할 수 있다. 일부 측면에서, 숙주 세포는 O-RS 또는 O-tRNA를 포함한 번역 시스템의 성분을 코딩하는 하나 이상의 폴리뉴클레오티드를 포함할 수 있다. 일부 구체예에서, O-RS를 코딩하는 폴리뉴클레오티드는 서열 번호 8~56에서 선택된 뉴클레오티드 서열을 포함한다.
- <34> 일부 측면에서, 번역 시스템은, 관심 단백질을 코딩하는 핵산을 추가로 포함하는데, 여기에서 핵산은 O-tRNA에 의해 인식되는 하나 이상의 셀렉터 코돈을 포함한다.
- <35> 일부 측면에서, 번역 시스템(즉, 숙주 세포)은 제2 비천연 아미노산을 이용하는 제2 오르소고날 쌍(즉, 제2 O-RS 및 제2 O-tRNA)을 도입하고, 이에 따라 시스템은 이제 폴리펩티드 내의 상이한 선택된 부위에 둘 이상의 상이한 비천연 아미노산을 도입할 수 있다. 이 이중 시스템에서, 제2 O-RS는 제2 O-tRNA를 제1 비천연 아미노산과 상이한 제2 비천연 아미노산으로 우선적으로 아미노아실화하고, 제2 O-tRNA는 제1 O-tRNA에 의해 인식되는 셀렉터 코돈과 상이한 셀렉터 코돈을 인식한다.
- <36> 상기 번역 시스템에 사용된 포유동물 숙주 세포는, O-RS 및 O-tRNA가 이들의 숙주 세포 환경에서 그들의 오르소고날성을 보유하는 한, 특별히 한정되지 않는다. 숙주 세포는 마우스 세포 또는 래트 세포와 같은 설치류 숙주 세포일 수 있다. 숙주 세포가 영장류 세포인 경우, 세포는, 예를 들면 인간 숙주 세포, 침팬지 숙주 세포, 보노보 숙주 세포, 고릴라 숙주 세포, 오랑우탄 숙주 세포, 긴팔원숭이 숙주 세포, 짧은꼬리원숭이 숙주 세포, 타마린 숙주 세포 또는 명주원숭이 숙주 세포일 수 있다. 숙주 세포는, 예를 들어 CHO 세포 또는 인간 293T일 수 있다.
- <37> 다른 측면에서, 본 발명은 선택된 위치에서 하나 이상의 비천연 아미노산을 갖는 관심 단백질을 생성하는 방법을 제공한다. 이러한 방법은 상기 기재된 번역 시스템을 이용한다. 일반적으로, 이러한 방법은, (i) 제1 비천연 아미노산(예, 도 1의 비천연 아미노산); (ii) 제1 오르소고날 아미노아실-tRNA 합성효소(O-RS); (iii) 제1 오르소고날 tRNA(O-tRNA)(여기에서, O-RS는 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화함); 및, (iv) 관심 단백질을 코딩하는 핵산(여기에서, 핵산은 제1 O-tRNA에 의해 인식되는 하나 이상의 셀렉터 코돈을 포함하

고, 핵산 내 셀렉터 코돈의 위치는 관심 단백질 내 선택된 위치와 상응함)을 포함하는 번역 시스템의 제공 단계로부터 시작한다. 모든 이러한 시스템 성분들은 적당한 숙주 세포, 예컨대 설치류 세포 또는 영장류 세포에 포함된다. 단백질을 비천연 아미노산으로 생성하는 방법은 상기 기재된 번역 시스템의 모든 성분들의 존재 하에 숙주 세포를 일반적으로 배양하는 단계를 수반한다. 이러한 방법에서, 비천연 아미노산은, 비제한적인 예로서,

<38> p-메톡시페닐알라닌 (pMpa);

<39> p-아세틸페닐알라닌 (pApa);

<40> p-벤조일페닐알라닌 (pBpa);

<41> p-요오도페닐알라닌 (pIpa);

<42> p-아지도페닐알라닌 (pAzpa);

<43> p-프로파르길옥시페닐알라닌 (pPpa);

<44> α-아미노카프릴산;

<45> o-니트로벤질시스테인 (o-NBC);

<46> 1,5-단실알라닌; 및

<47> o-니트로벤질세린 (o-NBS)

<48> 에서 선택될 수 있다.

<49> 이러한 방법의 일부 측면에서, O-RS는, 동일한 O-tRNA, 비천연 아미노산, 및 서열 번호 57~101에서 선택된 아미노산 서열을 포함하는 아미노아실-tRNA 합성효소(O-RS)를 포함하는 번역 시스템에 대해 관찰되는 효율의 50% 이상인 효율로, O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화한다.

<50> 이러한 방법은 각종 시약을 이용하여 광범위하게 적용될 수 있다. 방법의 일부 구체예에서, O-tRNA는 서열 번호 3을 포함하거나 그 서열에 의해 코딩된다. 일부 구체예에서, O-RS 및/또는 O-tRNA를 코딩하는 폴리뉴클레오티드가 제공된다. 일부 구체예에서, O-RS는 야생형 이.콜라이 티로실 또는 류실-tRNA 합성효소와 같은 이.콜라이 아미노아실-tRNA 합성효소에서 유도된다. 일부 구체예에서, 제공 단계는 서열 번호 57~101에서 선택된 아미노산 서열, 및 이의 보존적 변이체를 포함하는 O-RS를 제공하는 것을 포함한다. 이러한 방법의 일부 측면에서, O-RS를 코딩하는 폴리뉴클레오티드가 제공된다. 예를 들어, O-RS를 코딩하는 폴리뉴클레오티드는 서열 번호 8~56에서 선택된 뉴클레오티드 서열을 포함할 수 있다.

<51> 이러한 방법의 일부 구체예에서, 제공 단계는 부위-지정 돌연변이유발에 의해 야생형 아미노아실-tRNA 합성효소의 아미노산 결합 포켓을 돌연변이화하고, O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는 수득된 O-RS를 선택하는 것을 수반한다. 선택 단계는 부위-지정 돌연변이유발 후에 수득된 아미노아실-tRNA 합성효소 분자의 풀(pool)로부터 O-RS를 양성으로 선택 또는 음성으로 선택하는 것을 포함할 수 있다.

<52> 이러한 방법은 또한 하나 이상의 비천연 아미노산을 단백질에 도입하기 위해 변형될 수 있다. 상기 방법에서, 제2 오르소고달 번역 시스템이 제1 번역 시스템과 함께 이용되는데, 여기에서 제2 시스템은 상이한 아미노산 및 셀렉터 코돈 특이성을 가진다. 예를 들어, 제공 단계는 제2 O-RS 및 제2 O-tRNA를 제공하는 것을 포함할 수 있는데, 여기에서 제2 O-RS는 제2 O-tRNA를 제1 비천연 아미노산과 상이한 제2 비천연 아미노산으로 우선적으로 아미노아실화하고, 제2 O-tRNA는 제1 O-tRNA에 의해 인식되는 셀렉터 코돈과 상이한 핵산 내 셀렉터 코돈을 인식한다.

<53> 정의

<54> 본 발명을 상세히 기술하기 전, 본 발명은 특별한 생물학적 시스템으로 한정되지 않으며, 그 생물학적 시스템도 물론 다양할 수 있다는 것을 이해하도록 한다. 또한, 본원에서 사용되는 용어는 특별한 구체예를 단지 기술하기 위한 것으로, 한정하기 위한 것이 아님을 이해하도록 한다. 본 명세서 및 첨부된 청구의 범위에 사용되는, 단수 형태는 내용상 명확히 달리 명시하지 않는 한 복수 형태의 언급된 대상을 포함한다. 따라서, 예를 들어 "세포"라는 언급은 둘 이상의 세포의 조합물을 포함하고; "폴리뉴클레오티드"라는 언급은 실질적으로 그 폴리뉴클레오티드의 많은 복사체들을 포함한다.

<55> 여기에서 및 본 명세서의 하기 나머지 부분에서 달리 정의하지 않는 한, 본원에서 사용된 모든 기술적 및 과학

적 용어들은 본 발명이 속하는 분야에서 통상의 기술을 가진 자에 의해 통상적으로 이해되는 바와 동일한 의미를 가진다.

<56> 오르소고날: 본원에서 사용되는 용어 "오르소고날"은 세포 또는 번역 시스템에 대한 내인성을 가진 상응하는 분자에 비해 감소된 효율을 가진 세포의 내인성 성분들과 함께 작용하거나, 세포의 내인성 성분들과 함께 작용하지 못하는 분자(예를 들어, 오르소고날 tRNA(O-tRNA) 및/또는 오르소고날 아미노아실-tRNA 합성효소(O-RS))를 지칭한다. tRNA 및 아미노아실-tRNA 합성효소의 문맥에서, 오르소고날은 내인성 tRNA 합성효소와 함께 작용하는 내인성 tRNA에 비해 내인성 tRNA 합성효소와 함께 작용하는 오르소고날 tRNA, 또는 내인성 tRNA와 함께 작용하는 내인성 tRNA 합성효소에 비해 내인성 tRNA와 함께 작용하는 오르소고날 아미노아실-tRNA 합성효소의 불능 또는 감소된 효율, 예컨대 20% 미만의 효율, 10% 미만의 효율, 5% 미만의 효율 또는 1% 미만의 효율을 의미한다. 오르소고날 분자는 세포 내에서 기능적으로 정상인 내인성 상보적 분자를 가지지 않는다. 예를 들어, 세포 내 오르소고날 tRNA는 내인성 RS에 의한 내인성 tRNA의 아미노아실화에 비해 감소되거나 심지어 0%인 효율로 세포의 임의의 내인성 RS에 의해 아미노아실화된다. 또 다른 예에서, 오르소고날 RS는 내인성 RS에 의한 내인성 tRNA의 아미노아실화에 비해 감소되거나 심지어 0%인 효율로 관심 세포의 임의의 내인성 tRNA를 아미노아실화한다. 제1 오르소고날 분자와 함께 작용하는 제2 오르소고날 분자를 세포 내로 도입할 수 있다. 예를 들어, 오르소고날 tRNA/RS 쌍은 대조군, 예컨대 상응하는 tRNA/RS 내인성 쌍 또는 활성 오르소고날 쌍(예컨대, 티로실 오르소고날 tRNA/RS 쌍)의 효율에 비해 일정 효율(예, 45% 효율, 50% 효율, 60% 효율, 70% 효율, 75% 효율, 80% 효율, 90% 효율, 95% 효율 또는 99% 이상의 효율)로 세포 내에서 함께 작용하는, 도입된 상보적 성분들을 포함한다.

<57> 오르소고날 티로실-tRNA: 본원에서 사용되는 오르소고날 티로실 tRNA(티로실-O-tRNA)는 관심 번역 시스템에 대해 오르소고날성이 있는 tRNA이며, 여기에서 tRNA는 (1) 천연 발생 티로실-tRNA와 동일하거나 실질적으로 유사하거나, (2) 천연 또는 인공 돌연변이유발에 의해 천연 발생 티로실-tRNA로부터 유도되거나, (3) (1) 또는 (2)의 야생형 또는 돌연변이체 티로실-tRNA의 서열을 취하는 임의의 과정에 의해 유도되거나, (4) 야생형 또는 돌연변이체 티로실-tRNA와 상동성을 나타내거나, (5) 티로실-tRNA 합성효소에 대한 기질로서 표시된 임의의 예시적 tRNA와 상동성을 나타내거나, (6) 티로실-tRNA 합성효소에 대한 기질로서 표시된 임의의 예시적 tRNA의 보존적 변이체이다. 티로실-tRNA는 아미노산으로 충전된 상태로 존재할 수 있거나 충전되지 않은 상태로 존재할 수 있다. "티로실-O-tRNA"는 동족체 합성효소에 의해 각각 티로신 외의 아미노산, 예컨대 비천연 아미노산으로 임의적으로 충전(아미노아실화)된다는 것도 이해하도록 한다. 사실상, 본 발명의 티로실-O-tRNA는 번역과정 동안에 선택적 코돈에 대한 반응으로 본질적으로 임의의 천연 또는 비천연 아미노산을 성장하는 폴리펩티드 내로 삽입시키는 데 유리하게 사용된다는 것을 인식할 것이다.

<58> 오르소고날 티로실 아미노산 합성효소: 본원에서 사용되는 오르소고날 티로실 아미노산 합성효소(티로실-O-RS)는 관심 번역 시스템 내에서 티로실-O-tRNA를 아미노산으로 우선적으로 아미노아실화하는 효소이다. 티로실-O-RS가 티로실-O-tRNA 상에 로딩하는 아미노산은 임의의 천연, 비천연 또는 인공 아미노산일 수 있고 본원에서 한정되지 않는다. 그 합성효소는 임의적으로 천연 발생 티로실 아미노산 합성효소와 동일하거나 상동성을 나타내거나, 본원에 제공된 O-RS로서 표시된 합성효소와 동일하거나 상동성을 나타낸다. 예를 들어, O-RS는 도 16의 티로실-O-RS의 보존적 변이체일 수 있고/있거나, 도 16(서열 번호 57~101)의 O-RS와 서열이 50%, 60%, 70%, 80%, 90%, 95%, 98% 또는 99% 이상 동일할 수 있다.

<59> 오르소고날 류실-tRNA: 본원에서 사용되는 오르소고날 류실-tRNA(류실-O-tRNA)는 관심 번역 시스템에 대해 오르소고날이 있는 tRNA인데, 여기에서 tRNA는 (1) 천연 발생 류실-tRNA와 동일하거나 실질적으로 유사하거나, (2) 천연 또는 인공 돌연변이유발에 의한 천연 발생 류실-tRNA에서 유도되거나, (3) (1) 또는 (2)의 야생형 또는 돌연변이체 류실-tRNA 서열을 고려하는 임의의 과정에 의해 유도되고, (4) 야생형 또는 돌연변이체 류실-tRNA와 상동성을 나타내거나, (5) 류실-tRNA 합성효소에 대한 기질로서 표시된 임의의 tRNA와 상동성을 나타내거나, (6) 류실-tRNA 합성효소에 대한 기질로서 표시된 임의의 예시적 tRNA의 보존적 변이체이다. 류실-tRNA는 아미노산으로 충전된 상태로 존재할 수 있거나, 충전되지 않은 상태로 존재할 수 있다. "류실-O-tRNA"는 동족체 합성효소에 의해 각각 류신 외의 아미노산, 예컨대 비천연 아미노산으로 임의적으로 충전(아미노아실화)된다는 것도 이해하도록 한다. 사실상, 본 발명의 류실-O-tRNA는 번역 과정 동안에 선택적 코돈에 대한 반응으로 본질적으로 임의의 천연 또는 비천연 아미노산을 성장하는 폴리펩티드 내로 삽입시키는데 유리하게 사용된다는 것을 인식할 것이다.

<60> 오르소고날 류실-아미노산 합성효소: 본원에서 사용되는 오르소고날 류실 아미노산 합성효소(류실-O-RS)는 관심 번역 시스템 내에서 류실-O-tRNA를 아미노산으로 우선적으로 아미노아실화하는 효소이다. 류실-O-RS가 류실-O-

tRNA 상에 로딩하는 아미노산은 임의의 천연, 비천연 또는 인공 아미노산일 수 있고 본원에서 한정되지 않는다. 그 합성효소는 임의적으로 천연 발생 류실 아미노산 합성효소와 동일하거나 상동성을 나타내거나, 본원에 제공된 O-RS로서 표시된 합성효소와 동일하거나 상동성을 나타낸다. 예를 들어, O-RS는 도 16의 류실-O-RS의 보존적 변이체일 수 있고/있거나, 도 16의 O-RS와 서열이 50%, 60%, 70%, 80%, 90%, 95%, 98% 또는 99% 이상 동일할 수 있다.

<61> 동족체(cognate): 용어 "동족체"는, 예를 들어 오르소고날 tRNA 및 오르소고날 아미노아실-tRNA 합성효소와 함께 작용하거나 서로에 대한 특이성의 일부 측면을 가지는 성분들을 지칭한다. 그 성분들은 또한 상보적인 것으로서 지칭될 수도 있다.

<62> 우선적으로 아미노아실화한다: 오르소고날 번역 시스템에 대하여 본원에서 사용되는 O-RS는 그 자신이 발현 시스템 내에서 임의의 내인성 tRNA를 충전시키는 것보다 더 효율적으로 O-tRNA를 아미노산으로 충전시킬 때 동족체 O-tRNA를 "우선적으로 아미노아실화한다". 즉, O-tRNA 및 임의의 소정의 내인성 tRNA가 대략적으로 동등한 몰 비로 번역 시스템에 존재하는 경우, O-RS는 내인성 tRNA를 충전시키는 것보다 더 자주 O-tRNA를 충전시킬 것이다. 바람직하게는, O-tRNA 및 내인성 tRNA가 동등한 몰 농도로 번역 시스템에 존재하는 경우, O-RS에 의해 충전된 내인성 tRNA에 대한 O-RS에 의해 충전된 O-tRNA의 상대적 비율은 바람직하게는 O-RS가 O-tRNA만을 충전시키거나 거의 O-tRNA만을 충전시킬 정도로 높다. O-tRNA 및 O-RS가 동등한 몰 농도로 존재하는 경우, O-RS에 의해 충전되는 O-tRNA와 내인성 tRNA 사이의 상대적 비율은 1:1 초과, 바람직하게는 약 2:1 이상, 보다 바람직하게는 5:1, 보다 더 바람직하게는 10:1, 더욱 더 바람직하게는 20:1, 더욱 더 바람직하게는 50:1, 보다 더 바람직하게는 75:1, 더욱 더 바람직하게는 95:1, 98:1, 99:1, 100:1, 500:1, 1,000:1, 5,000:1 또는 그 이상이다.

<63> O-RS는 (a) O-RS가 내인성 tRNA에 비해 O-tRNA를 우선적으로 아미노아실화하는 경우, 및 (b) O-RS가 O-tRNA를 임의의 천연 아미노산으로 아미노아실화하는 것에 비해 아미노아실화가 비천연 아미노산에 대해 특이적인 경우, "O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화한다". 즉, 비천연 아미노산 및 천연 아미노산이 O-RS 및 O-tRNA를 포함하는 동물 양으로 번역 시스템에 존재하는 경우, O-RS는 천연 아미노산보다 더 빈번하게 비천연 아미노산을 O-tRNA에 로딩할 것이다. 바람직하게는, 천연 아미노산으로 충전된 O-tRNA에 대한 비천연 아미노산으로 충전된 O-tRNA의 상대적 비율은 높다. 보다 바람직하게는, O-RS는 O-tRNA만을, 또는 거의 O-tRNA만을 비천연 아미노산으로 충전한다. 천연 아미노산 및 비천연 아미노산이 동물 농도로 번역 시스템에 존재하는 경우, 비천연 아미노산을 사용한 O-tRNA의 충전과 천연 아미노산을 사용한 O-tRNA의 충전 사이의 상대적 비율은 1:1 초과, 바람직하게는 약 2:1 이상, 보다 바람직하게는 5:1, 보다 더 바람직하게는 10:1, 더욱 더 바람직하게는 20:1, 더욱 더 바람직하게는 50:1, 더욱 더 바람직하게는 75:1, 더욱 더 바람직하게는 95:1, 98:1, 99:1, 100:1, 500:1, 1,000:1, 5,000:1 또는 그 이상이다.

<64> 셀렉터 코돈: 용어 "셀렉터 코돈"은 번역과정에서 O-tRNA에 의해서는 인식되나 내인성 tRNA에 의해서는 인식되지 않는 코돈을 지칭한다. O-tRNA 안티코돈 루프는 mRNA 상의 셀렉터 코돈을 인식하여 그의 아미노산, 예컨대 비천연 아미노산을 폴리펩티드 내의 인식된 부위 내로 도입한다. 셀렉터 코돈은, 예를 들어 넌센스 코돈, 예컨대 정지 코돈, 예를 들어 엠버, 오키(ochre) 및 오팔(opal) 코돈; 4개 이상의 염기로 이루어진 코돈; 희귀 코돈; 천연 또는 비천연 염기 쌍 등으로부터 유도된 코돈 등을 포함할 수 있다.

<65> 억제자(suppressor) tRNA: 억제자 tRNA는 전형적으로 폴리펩티드의 번역 과정 중에 예를 들어 정지 코돈(즉, "리드 쓰루")에 대한 반응으로 아미노산을 도입함으로써 소정의 번역 시스템 내의 메신저 RNA(mRNA)의 관독을 변경시키는 tRNA이다. 일부 측면에서, 본 발명의 셀렉터 코돈은 억제자 코돈, 예컨대 정지 코돈(예, 엠버, 오키 또는 오팔 코돈), 4 염기 코돈, 희귀 코돈 등이다.

<66> 억제 활성: 본원에서 사용되는 용어 "억제 활성"은 일반적으로, 다른 경우라면 번역의 종료 또는 오역(mistranslation)(예, 프레임 이동)을 초래하게 되는 코돈(예, 엠버 코돈 또는 4개 이상의 염기로 이루어진 코돈인 셀렉터 코돈)의 번역 리드 쓰루를 허용하는 tRNA(예, 억제자 tRNA)의 능력을 지칭한다. 억제자 tRNA의 억제 활성은 제2 억제자 tRNA에 비해, 또는 대조군 시스템, 예컨대 O-RS가 결여된 대조군 시스템에 비해, 관찰된 번역 리드 쓰루의 활성율(%)로서 표시될 수 있다.

<67> 본 발명은 억제 활성을 정량화할 수 있는 다양한 방법을 제공한다. 관심 셀렉터 코돈(예, 엠버 코돈)에 대한 특별한 O-tRNA 및 O-RS의 억제율은 그 자신을 코딩하는 핵산 내에 셀렉터 코돈을 포함하는 소정의 발현된 시험 마커(예, LacZ)가 관심 번역 시스템에서 나타내는 활성율(%)을 지칭하는데, 여기에서 관심 번역 시스템은 양성 대조군 구축물(construct)에 비해 O-RS 및 O-tRNA를 포함하고, 여기에서 양성 대조군은 O-tRNA, O-RS 및 셀렉터 코돈을 가지지 않는다. 따라서, 예를 들어 셀렉터 코돈을 가지지 않은 활성 양성 대조군 마커 구축물이 소정의

번역 시스템에서 해당 마커 검정에 대한 유닛으로 X의 관찰된 활성을 나타내는 경우, 셀렉터 코돈을 포함하는 시험 구축물의 억제율(%)은 시험 마커 구축물이 O-tRNA 및 O-RS를 포함하는 번역 시스템에서 발현된다는 점을 제외하고 양성 대조군 마커가 발현된 환경 조건과 본질적으로 동일한 환경 조건 하에서 시험 마커 구축물이 나타내는 X의 활성율(%)이다. 전형적으로, 시험 마커를 발현하는 번역 시스템은 또한 O-RS 및 O-tRNA에 의해 인식되는 아미노산도 포함한다. 임의적으로, 억제율(%) 측정은 O-tRNA 및/또는 O-RS에 의해 인식되는 O-tRNA, O-RS 및/또는 관련 아미노산을 포함하지 않는 시스템에서 시험 마커와 동일한 셀렉터 코돈을 포함하는 "배경" 또는 "음성" 대조군 마커 구축물과 시험 마커의 비교에 의해 달성된다. 이 음성 대조군은 관심 번역 시스템에서 마커로부터의 배경 신호 효과를 고려하여 억제율 측정을 표준화하는 데 유용하다.

<68> 억제 효율은 당업계에 공지된 다수의 검정 중 임의의 검정에 의해 결정될 수 있다. 예를 들어, β -갈락토시다제 리포터 검정이 이용될 수 있는데, 예를 들어 유도체화된 lacZ 플라스미드(이 때, 구축물은 lacZ 핵산 서열 내에 셀렉터 코돈을 가짐)가 본 발명의 O-tRNA를 포함하는 플라스미드와 함께 적절한 유기체(예, 오르소고날 성분들이 사용될 수 있는 유기체)로부터 유도된 세포 내로 도입된다. 동족체 합성효소도 또한 (발현된 경우에는 동족체 합성효소를 코딩하는 폴리펩티드 또는 폴리뉴클레오티드로서) 도입될 수 있다. 세포는 소정의 밀도, 예를 들어 약 0.5의 OD₆₀₀까지 배지 중에서 성장시키고, β -갈락토시다제 검정을, 예컨대 BetaFluor™ β -갈락토시다제 분석 키트[노바젠(Novagen)]를 사용하여 수행한다. 억제율(%)은 비교 대조군에 대한 상대적인 샘플 활성율로서, 예를 들어 셀렉터 코돈보다 원하는 위치에서 상응하는 센스 코돈을 가진 유도체화된 lacZ 구축물로부터 관찰된 값으로서 계산될 수 있다.

<69> 번역 시스템: 용어 "번역 시스템"은 아미노산을 성장하는 폴리펩티드 채(단백질) 내로 도입하는 성분을 지칭한다. 번역 시스템의 성분은, 예를 들어 리보솜, tRNA, 합성효소, mRNA 등을 포함할 수 있다. 본 발명의 O-tRNA 및/또는 O-RS는 시험관내 또는 생체내 번역 시스템, 예를 들어 포유동물 세포, 예컨대 설치류 세포 또는 영장류 세포에 첨가될 수 있거나 이들의 일부일 수 있다.

<70> 비천연 아미노산: 본원에서 사용되는 용어 "비천연 아미노산"은 20종의 공통된 천연 발생 아미노산 중 하나가 아닌 임의의 아미노산, 변형된 아미노산 및/또는 아미노산 유사체 또는 셀레노 시스테인 또는 피롤리신을 지칭한다. 예를 들어, 도 1에 제공된 비천연 아미노산을 참조한다.

<71> ~로부터 유도된: 본원에서 사용되는 용어 "~로부터 유도된"은 특별한 분자 또는 유기체로부터 단리된 형태이거나, 특별한 분자 또는 유기체 또는 특별한 분자 또는 유기체의 정보를 사용하여 만든 성분을 지칭한다. 예를 들어, 제2 폴리펩티드로부터 유도된 폴리펩티드는 제2 폴리펩티드의 아미노산 서열과 동일하거나 실질적으로 유사한 아미노산 서열을 포함할 수 있다. 폴리펩티드의 경우, 유도된 종은, 예를 들어 천연 발생 돌연변이유발, 인공 지정 돌연변이유발 또는 인공 랜덤 돌연변이유발에 의해 획득될 수 있다. 폴리펩티드를 유도하는 데 이용되는 돌연변이유발은 의도적으로 지정될 수 있거나 의도적으로 무작위적일 수 있거나, 각각의 조합일 수 있다. 제1 폴리펩티드로부터 유도된 상이한 폴리펩티드를 생성하기 위한 폴리펩티드의 돌연변이유발은 (예를 들어, 중합효소 부정확성에 의해 야기된) 무작위적 사건일 수 있고, 유도된 폴리펩티드는, 예를 들어 본원에서 논의된 바와 같은, 적절한 선별 방법에 의해 동정될 수 있다. 폴리펩티드의 돌연변이유발은 전형적으로 그 폴리펩티드를 코딩하는 폴리뉴클레오티드의 조작을 수반한다.

<72> 양성 선택(selection) 또는 선별(screening) 마커: 본원에서 사용되는 용어 "양성 선택 또는 선별 마커"는 존재하는 경우, 예컨대 발현되거나 활성화 등이 일어난 경우 특징을 갖는 세포, 예를 들어 양성 선택 마커를 가진 세포를 그 특징을 가지지 않는 세포로부터 구별시켜주는 마커를 지칭한다.

<73> 음성 선택 또는 선별 마커: 본원에서 사용되는 용어 "음성 선택 또는 선별 마커"는 존재하는 경우, 예컨대 발현되거나 활성화 등이 일어난 경우 (예를 들어, 선택된 성질 또는 특징을 갖는 세포와 비교할 때) 상기 성질 또는 특징을 가지지 않는 세포를 구별시켜주는 마커를 지칭한다.

<74> 리포터: 본원에서 사용되는 용어 "리포터"는 관심 시스템의 표적 성분들을 동정하고/하거나 선택하는 데 사용될 수 있는 성분을 지칭한다. 예를 들어, 리포터는 항생제 내성 또는 감수성을 부여하는 단백질, 예컨대 효소(예, β -락타마제, 클로람페니콜 아세틸트랜스퍼라제(CAT) 등), 형광 선별 마커(예, 녹색 형광 단백질(예, (GFP), YFP, EGFP, RFP 등)), 발광 마커(예, 초파리 루시페라제 단백질), 친화성 기초 선별 마커, 또는 양성 또는 음성 선택가능한 마커 유전자, 예컨대 lacZ, β -gal/lacZ(β -갈락토시다제), ADH(알코올 탈수소효소), his3, ura3, leu2, lys2 등을 포함할 수 있다.

<75> 진핵생물: 본원에서 사용되는 용어 "진핵생물"은 진핵생물 계에 속하는 유기체를 지칭한다. 일반적으로, 진핵생

물은 그들의 전형적인 다세포 조직화(단, 오로지 다세포 조직화로만 한정되지 않음; 예컨대, 효모), 막-결합 핵 및 다른 막-결합 소기관의 존재, 선형 유전 물질(예를 들어, 선형 염색체), 오페론의 부재, 인트론, 메세지 캡핑(message capping) 및 폴리-A mRNA의 존재, 및 다른 생화학적 특징, 예컨대 구별가능한 리보솜 구조에 의해 원핵생물로부터 구별될 수 있다. 진핵생물은, 예를 들어 동물(예, 포유동물, 곤충, 파충류, 조류 등), 섬모충, 식물(예, 단자엽, 쌍자엽, 해조류 등), 진균류, 효모, 편모충, 미포자충(microsporidia), 원핵생물 등을 포함한다.

<76> 원핵생물: 본원에서 사용되는 용어 "원핵생물"은 모네라 계(Kingdom Monera)(원핵생물계로도 지칭됨)에 속하는 유기체를 지칭한다. 일반적으로, 원핵생물은 그들의 단세포 조직화, 발아 또는 분열(fission)에 의한 무성 생식, 막-결합 핵 또는 다른 막-결합 소기관의 부재, 원핵 염색체, 오페론의 존재, 인트론, 메세지 캡핑 및 폴리-A mRNA의 부재, 및 다른 생화학적 특징, 예컨대 구별가능한 리보솜 구조에 의해 진핵생물로부터 구별될 수 있다. 원핵생물에는 아계 진정세균 및 고세균(Archaea)(종종 "원시세균(Archaeobacteria)"으로 불림)이 포함된다. 역사적으로 "세균"으로 간주되었다 하더라도, 시아노박테리아(청녹색 해조류) 및 마이코플라스마는 종종 모네라계 하에 별도로 분류된다.

<77> 진정세균: 본원에서 사용되는 용어 "진정세균" 및 "세균"은 고세균으로부터 구별될 수 있는 원핵생물을 지칭한다. 유사하게, 고세균은 진정세균으로부터 구별될 수 있는 원핵생물을 지칭한다. 진정세균과 고세균은 다수의 형태학적 및 생화학적 기준에 의해 구별될 수 있다. 예를 들어, 리보솜 RNA 서열에서의 차이, RNA 중합효소 구조, 인트론의 존재 또는 부재, 항생제 감수성, 세포 벽 펩티도글리칸 및 다른 세포 벽 성분들의 존재 또는 부재, 막 지질의 분지 구조 대 비분지 구조, 및 히스톤 및 히스톤-유사 단백질의 존재/부재를 이용하여 유기체를 진정세균 또는 고세균으로 분류한다.

<78> 진정세균의 예에는 에세리키아 콜라이, 씨머스 씨모필러스, 바실러스 서브틸리스(*Bacillus subtilis*) 및 바실러스 스테아로씨모필러스(*Bacillus stearothermophilus*)가 포함된다. 고세균의 예에는 메타노코커스 잔나쉬이(*Methanococcus jannaschii*)(Mj), 메타노사르시나 메이제이(*Methanosarcina mazei*)(Mm), 메타노박테리움 씨모 오토트로피쿰(*Methanobacterium thermoautotrophicum*)(Mt), 메타노코커스 마리팔루디스(*Methanococcus maripaludis*), 메타노파이러스 칸들레리(*Methanopyrus kandleri*), 할로박테리움, 예컨대 할로페락스 볼케니이(*Haloferax volcanii*) 및 할로박테리움 종 NRC-1, 아카에오글로버스 풀기더스(*Archaeoglobus fulgidus*)(Af), 파이로코커스 푸리오서스(*Pyrococcus furiosus*)(Pf), 파이로코커스 호리코쉬이(Ph), 파이로바쿨럼 애로필럼(*Pyrobaculum aerophilum*), 파이로코커스 애비시시(*Pyrococcus abyssi*), 설펠로버스 솔파타리커스(*Sulfolobus solfataricus*)(Ss), 설펠로버스 토크다이이(*Sulfolobus tokodaii*), 애우로파이럼 페르닉스(*Aeuryopyrum pernix*)(Ap), 씨모플라스마 애시도필럼(*Thermoplasma acidophilum*) 및 씨모플라스마 볼케이눔(*Thermoplasma volcanium*)이 포함된다.

<79> 보존적 변이체: 본원에서 사용되는, 번역 성분의 영역에서의 용어 "보존적 변이체"는 기능적으로 기준 성분과 유사한 작용을 하는 번역 성분, 예를 들어 보존적 변이체 O-tRNA 또는 보존적 변이체 O-RS를 지칭하고, 기준 O-tRNA 또는 O-RS와 비교할 때 서열 내에 변이를 가진 보존적 변이체는, 예를 들어 O-tRNA 또는 O-RS와 유사하다. 예를 들어, O-RS 또는 이 O-RS의 보존적 변이체는 동족체 O-tRNA를 비천연 아미노산, 예를 들어 도 1의 비천연 아미노산으로 아미노아실화할 것이다. 본 실시예에서, O-RS와 보존적 변이체 O-RS는 동일한 아미노산 서열을 가지지 않는다. 보존적 변이체가 동족체의 상응하는 O-tRNA 또는 O-RS와 여전히 상보적인(예를 들어, 그것과 작용하는) 한, 그 보존적 변이체는 서열 내에 예를 들어, 1개의 변이, 2개의 변이, 3개의 변이, 4개의 변이 또는 5개 이상의 변이를 가질 수 있다.

<80> 일부 구체예에서, 보존적 변이체 O-RS는 그 자신의 기원인 O-RS에 비해 하나 이상의 보존적 아미노산 치환을 포함한다. 일부 구체예에서, 보존적 변이체 O-RS는 그 자신의 기원인 O-RS에 비해 하나 이상의 보존적 아미노산 치환을 포함할 뿐만 아니라, O-RS의 생물학적 활성을 보유하는데, 예를 들어 보존적 변이체 O-RS는 그 자신의 기원인 모 O-RS의 생물학적 활성의 10% 이상, 또는 대안적으로 20% 이상, 30% 이상 또는 40% 이상을 보유한다. 일부 바람직한 구체예에서, 보존적 변이체 O-RS는 그 자신의 기원인 모 O-RS 분자의 생물학적 활성의 50% 이상을 보유한다. 보존적 변이체 O-RS의 보존적 아미노산 치환은 아미노산 결합 포켓을 포함한 O-RS의 임의의 도메인에서 일어날 수 있다.

<81> 선택 또는 선별제: 본원에서 사용되는 용어 "선택 또는 선별제"는 존재하는 경우, 집단으로부터 특정 성분들의 선택/선별을 가능하게 하는 시약을 지칭한다. 예를 들어, 선택 또는 선별제는, 예를 들어 영양소, 항생제, 광과장, 항체, 발현된 폴리뉴클레오티드 등일 수 있으나, 이들로 한정되지 않는다. 선택제는, 예를 들어 농도, 강

도 등에 의해 다양해질 수 있다.

<82> ~에 대한 반응으로: 본원에서 사용되는 용어 "~에 대한 반응으로"는 본 발명의 O-tRNA가 선택적 코돈을 인식하고 그 tRNA에 결합된 비천연 아미노산이 성장하는 폴리펩티드 쇠 내로 도입되는 것을 매개하는 과정을 지칭한다.

<83> 코딩하다: 본원에서 사용되는 용어 "코딩하다"는 중합체성 거대분자 또는 서열 스트링(string) 내의 정보를 사용하여 제1 분자 또는 서열 스트링과 상이한 제2 분자 또는 서열 스트링을 생성시키는 임의의 과정을 지칭한다. 본원에서 사용되는 상기 용어는 광범위하게 사용되며, 다양하게 응용될 수 있다. 일부 측면에서, 용어 "코딩하다"는, 이중-가닥 DNA 분자의 한 가닥이 DNA-의존성 DNA 중합효소에 의해 새로이 합성된 상보적 시스터 가닥을 코딩하기 위한 주형으로서 사용되는, 반-보존적 DNA 복제 과정을 지칭한다.

<84> 또 다른 측면에서, 용어 "코딩하다"는 한 분자 내의 정보를 사용하여 제1 분자와 상이한 화학적 성질을 가진 제2 분자를 생성시키는 임의의 과정을 지칭한다. 예를 들어, DNA 분자는 (예를 들어, DNA-의존성 RNA 중합효소를 도입하는 전사 과정에 의해) RNA 분자를 코딩할 수 있다. 또한, RNA 분자는 번역 과정에서와 같이 폴리펩티드를 코딩할 수 있다. 번역 과정을 기술하는 데 사용되는 용어 "코딩하다"는 아미노산을 코딩하는 3 염기 코돈으로 확장된다. 일부 측면에서, RNA 분자는, 예를 들어 RNA-의존성 DNA 중합효소를 도입하는 역전사 과정에 의해 DNA 분자를 코딩할 수 있다. 또 다른 측면에서, DNA 분자는 폴리펩티드를 코딩할 수 있는데, 여기에서 이 경우에서 사용되는 "코딩하다"는 전사 과정 및 번역 과정 양자 모두를 수반한다는 것을 이해할 것이다.

<85> 발명의 상세한 설명

<86> 본 발명은 포유동물 세포 숙주 시스템에서 다양한 물리화학적, 생물학적 성질을 갖는 비천연 아미노산으로 적당한 오르소고날 억제자 tRNA를 충전하는 능력을 갖는 돌연변이체 아미노아실-tRNA 합성효소(RS)를 신속하게 선별 및 단리할 수 있는 일반적인 방법론을 제공하며, 또한 상기 선별 방법에서는 포유동물 세포에서의 느리고 기술적으로 어려운 선별이 필요 없다. 이러한 방법에서, 돌연변이체 에세리키아 콜라이 아미노아실-tRNA 합성효소(RS)를 효모에서 진화시켜 관심 비천연 아미노산을 선택적으로 이용하게 한다. 이후 바실러스 스테아로써모필러스(*B. stearothermophilus*) 유래의 엠버 억제자 tRNA와 함께 동일한 돌연변이체 RS를 사용하여 선택적 코돈(예, 엠버 넌센스 코돈)에 대한 반응으로 포유동물 세포에서 관심 단백질에 비천연 아미노산을 부위 특이적으로 도입한다. 단백질로의 비천연 아미노산의 도입은, 비천연 아미노산의 도입을 신호하는 선택적 코돈을 포함하도록 관심 단백질을 코딩하는 폴리뉴클레오티드를 처리함으로써 임의의 목적하는 위치에서 일어나도록 프로그램화된다.

<87> 본 발명은 또한 신규 번역 시스템을 제공하는데, 여기에서 상기 시스템은 여기에 고려되지 않았던 숙주 세포를 포함한다. 예를 들면, 본 발명의 번역 시스템은 (i) 비천연 아미노산, (ii) 오르소고날 tRNA(O-tRNA), (iii) 진정세균의 아미노아실-tRNA 합성효소에서 유도된 오르소고날 아미노아실-tRNA 합성효소(O-RS), 전형적으로 이.콜라이 RS, 및 (iv) 이러한 성분들을 함유하는 포유동물 숙주 세포(여기서, 포유동물 숙주 세포는 설치류 세포(예, 마우스 세포 또는 래트 세포), 또는 영장류 세포, 즉 원숭이류(유인원, 인간 및 뉴 월드 명키를 포함), 뉴 월드 명키 및 원원류(pro-simian)(예, 여우원숭이)를 공통적으로 언급하는 군의 임의의 것에 속하는 임의의 세포일 수 있음)를 포함한다. 인간 세포는 본 발명에서 특히 사용된다. 하지만, 본 발명에서 상기 나열된 숙주 세포로 한정하려는 의도가 아니며, 다른 숙주 세포도 또한 본 발명에 사용한다.

<88> 본 발명은 또한 선택된 위치에서 하나 이상의 비천연 아미노산을 갖는 단백질의 생성 방법을 제공하는데, 여기에서 상기 방법은 상기 기재된 번역 시스템을 사용한다. 이 방법은, 단백질 번역 동안 선택적 코돈의 부위에서 비천연 아미노산의 도입을 프로그래밍화하여 비천연 아미노산을 도입한다.

<89> 본 발명은 적당한 진정세균, 예컨대 비. 스테아로써모필러스에서 유도된 엠버 억제자와 함께 돌연변이체 RS(본래 에스. 세레비시에에서 진화함)를 사용하여 포유동물 숙주 세포, 예컨대 설치류 세포 및 영장류 세포에서 단백질에 광범위하게 다양한 비천연 아미노산을 도입시키는 일반적인 접근법을 제공한다. 원리의 검증으로서, 이러한 방법론은 중국 햄스터 난소(CHO) 및 인간 293T 숙주 세포에서 탁월한 적합도 및 양호한 효율을 가지고 엠버 넌센스 코돈에 대한 반응으로 모델 단백질에 총 10개의 상이한 비천연 아미노산을 부위 특이적으로 도입하는데 성공적으로 사용되었다. 예를 들어, 비천연 아미노산은 2×10^7 세포 당 최대 1 μ g 단백질의 효율로 CHO 세포에서 발현되는 녹색 형광 단백질(GFP)에 독립적으로 도입되었다. 질량 분석법은 비천연 아미노산에 대한 높은 번역적 적합도를 확인하였다. 이러한 방법론은 세포 연구를 위해 단백질로의 생물학적 프로브의 도입을 용이하게 할 것이며 궁극적으로 포유동물 세포에서 비천연 아미노산을 함유하는 치료적 단백질의 합성을 용이하게 할 수 있다.

- <90> 일부 측면에서, 본 발명을 (비제한적으로) 입증하기 위해, 본원의 내용은 비천연 아미노산 부분이 포유동물 숙주 세포에서 다양한 모델 단백질에 도입될 수 있음을 입증한다. 비천연 아미노산의 도입이 임의의 특정 관심 단백질로 한정시키려는 의도는 아니다. 본 내용에서, 특정 관심 단백질에 다양한 비천연 아미노산의 도입은 광범위하게 다양한 단백질과 광범위하게 다양한 목적에 유리하다는 것이 명백할 것이다.
- <91> 효모와 O-tRNA에서 진화된 돌연변이체 이.콜라이 아미노아실-tRNA 합성효소(O-RS)의 발현을 제어하는 유전자는 포유동물 숙주 세포(예, 설치류 또는 인간 세포)에서 즉각적으로 사용될 수 없다는 점에 유의한다. 이러한 유전자들은 포유동물 세포 배경에서 이들의 발현을 추진하기 위해 적당한 조절 성분들을 이용하여 재조작되어야 한다. O-tRNA 및 O-RS 유전자와 관련된 효모 전사 조절 성분들은 원래 포유동물 세포에서 사용하는데 비효율적일 (또는 기껏해야 최적이지 아닐) 것이다. 당업자는 이 점을 인식하고 있다. 포유동물 세포 유전자 발현 성분들은 이들과 관련된 유전자 서열을 최적화(예, 최대화)시킬 목적으로 재조합 구축물에서 사용하기에 잘 특징화되어 있거나 쉽게 이용 가능하다.
- <92> 본 발명에서 사용하기 위한 비천연 아미노산
- <93> 본 발명은 포유동물 숙주세포 시스템, 예컨대 영장류 및 설치류 숙주 세포 시스템에서 비천연 아미노산을 포함하는 폴리펩티드를 생산하는 능력을 갖는 오르소고날 번역 시스템을 제공한다. 이러한 번역 시스템에 사용된 비천연 아미노산은 특별히 한정되지 않는다. 일반적으로, 본 발명에서 사용하기 위한 비천연 아미노산은, 이.콜라이 유도된 돌연변이체 아미노아실-tRNA 합성효소에 의해 이용되어 사카로미세스 세레비시에 (에스.세레비시에) 효모 숙주 세포 시스템에서 해당 오르소고날 tRNA를 충전할 수 있는 것이다. 예를 들어, 본원에 사용된 바와 같이, 하기 비천연 아미노산은 본 발명에 사용한다:
- <94> p-메톡시페닐알라닌 (pMpa)
- <95> p-아세틸페닐알라닌 (pApa)
- <96> p-벤조일페닐알라닌 (pBpa)
- <97> p-요오도페닐알라닌 (pIpa)
- <98> p-아지도페닐알라닌 (pAzpa)
- <99> p-프로파르길옥시페닐알라닌 (pPpa)
- <100> 이러한 6개의 비천연 아미노산의 구조는 도 1에 도시된다. 이러한 아미노산은 또한 대안적인 명칭과 다양한 약자에 의해 공지될 수 있으며, 이는 또한 도 1에 도시된다. 실시예에 제시된 바와 같이, 이러한 6개의 비천연 아미노산은, 효모 숙주 세포 시스템을 이용하여 원래 선별되고 단리되었던 이.콜라이 유도된 O-RS 종을 이용하여 인간 및 설치류 숙주 세포에서 녹색 형광 단백질(GFP)에 독립적으로 도입되었다.
- <101> 상기 열거된 것 이외에, 여전히 다른 비천연 아미노산 또한 본 발명에 사용된다. 하기 열거된 비천연 아미노산은 효모 숙주 세포 시스템에서 O-tRNA를 충전시키기 위해 이.콜라이 유도된 오르소고날 RS에 의해 이전에 성공적으로 사용된 바 있다. 그러한 비천연 아미노산은
- <102> α-아미노카프릴산
- <103> o-니트로벤질시스테인 (o-NBC)
- <104> 1,5-단실알라닌
- <105> o-니트로벤질세린
- <106> 을 포함한다.
- <107> 이러한 비천연 아미노산의 구조 및 이의 다양한 대안적인 명칭 및 약자는 도 1에 도시된다.
- <108> 효모 숙주 세포에서 모든 그러한 비천연 아미노산의 성공적 사용은, 예를 들어 문헌[Chin 등(2003), "An expanded eukaryotic genetic code," *Science* **301**:964-967; Deiters 등(2003), "Adding amino acid with novel reactivity to the genetic code of *Saccharomyces cerevisiae*," *J Am Chem Soc* **125**:11782-11783; Summerer 등(2006), "A Genetically Encoded Fluorescent Amino Acid," *PNAS* **103**(26):9785-9789; 제WO 2005/003294호(Deiters 등), "UNNATURAL REACTIVE AMINO ACID GENETIC CODE ADDITIONS,"(2004년 4월 16일 출원); 제WO 2006/034410호(Deiters 등), "ADDING PHOTOREGULATED AMINO ACID TO THE GENETIC CODE,"(2005년 9

월 21일 출원); 및 제WO 2006/110182호(2005년 10월 27일 출원),(발명의 명칭: "ORTHOGONAL TRANSLATION COMPONENTS FOR THE VIVO INCORPORATION OF UNNATURAL AMINO ACID")에 기재된다.

<109> 본 발명에 사용하기 위한 숙주 세포

<110> 본 발명은 신규 번역 시스템을 제공하는데, 여기에서 시스템은 여기에 고려되지 않았던 숙주 세포를 이용한다. 예를 들면, 본 발명의 번역 시스템은 (i) 비천연 아미노산, (ii) 오르소고날 tRNA(O-tRNA), (iii) 이.콜라이 아미노아실-tRNA 합성효소에서 유도된 오르소고날 아미노아실-tRNA 합성효소(O-RS), 및 (iv) 이러한 성분들을 함유하는 포유동물 숙주 세포(여기서, 포유동물 숙주 세포는 설치류 세포(예, 마우스 세포 또는 래트 세포), 또는 영장류 세포, 즉 원숭이류(유인원, 인간 및 뉴 월드 명키를 포함), 뉴 월드 명키 및 원원류(pro-simian)(예, 여우원숭이)를 공통적으로 언급하는 군의 임의의 것에 속하는 임의의 세포일 수 있음)를 포함한다. 인간 세포는 본 발명에서 특히 사용된다. 본 발명에 사용되는 동물 세포주는 동물에서 유도된 임의의 조직에서 유도될 수 있다. 본 발명은 상기 열거된 숙주 세포에 한정되도록 하는 것은 아니지만, 다른 숙주 세포들 또한 본 발명에 사용한다.

<111> 본원에 사용된 바와 같이, 용어 "숙주 세포"는 또한 "세포주"로 공통으로 지칭되고, 이 경우에는 "숙주 세포주"로 복수의 개별 세포들을 포함한다.

<112> 하나 이상의 비천연 아미노산을 함유하는 폴리펩티드의 생산을 위해 사용되는 숙주 세포는 임의의 유형의 숙주 세포일 수 있지만, 일부 숙주 세포는 특정 용도를 갖는다. 예를 들어, 일부 숙주 세포 유형을 사용하여 재조합 치료적 폴리펩티드의 생성시 사용하는 폴리펩티드를 생성할 수 있다.

<113> 숙주 세포가 설치류 세포인 경우, 설치류 세포의 유형은 특별히 한정되지 않는다. 하기 실시예에 기재된 바와 같이, 중국 햄스터 난소(CHO) 세포를 숙주 세포로 사용할 수 있다. CHO 세포는 광범위하게 맞춰지며 재조합 단백질 발현을 위해 잘 확립된 시스템이다. 하지만, 본 발명을 이러한 설치류 세포 유형에 한정시키려는 의도가 아니다. 사실상, 당업자는 재조합 단백질 생산을 위한 숙주 세포로 사용될 수 있는 설치류 세포주의 수많은 어레이를 알고 있다.

<114> 설치류 세포주는 포유동물 목 설치목(Rodentia)에서 유도된 임의의 세포주를 포함하고, 비제한적으로, 트루 마우스, 래트, 햄스터, 줄다람쥐, 비버 및 다람쥐를 포함한다. 설치류 세포주는 일반적인 생쥐, 무스 무스쿨러스(*Mus musculus*), 및 이의 임의의, 모든 종속에서 유도된 세포주를 포함한다. 유사하게, 설치류 세포주는 래트, 예컨대 속 라투스(*Rattus*)의 임의의 종, 예를 들어 라투스 노르베지쿠스(*Rattus norvegicus*)(노르웨이 쥐) 및 라투스 라투스(*Rattus rattus*)(검은 쥐), 및 이의 임의의, 모든 종속에서 유도된 세포주를 포함한다.

<115> 본 발명의 숙주 세포는 영장류 세포일 수 있다. 영장류 세포의 유형은 특별히 한정되지 않는다. 하기 실시예에 기재된 바와 같이, 인간 태아의 신장 세포(또한 HEK 세포, HEK 293 또는 293으로 공지됨)는 숙주 세포로 사용될 수 있다. 293 세포는 광범위하게 맞춰질 수 있고 재조합 단백질 조합을 위해 잘 확립된 시스템이다. 하지만, 본 발명은 이러한 인간 세포 유형에 한정하려는 의도가 아니다. 사실상, 당업자는 재조합 단백질 생성을 위한 숙주 세포로 사용될 수 있는 인간 및 다른 영장류 세포주의 다양한 어레이를 알고 있다.

<116> 영장류 세포주는 포유동물 목 영장목에서 유도된 임의의 세포주를 포함하고, 비제한적으로 (i) 인간을 포함한 올드 월드 명키 영장류를 포함하는 원숭이류, (ii) 뉴 월드 명키 및 (iii) 원원류(예, 여우원숭이)를 포함한다. 영장류 세포주는 임의의 영장류 종으로부터 유도될 수 있다. 예를 들면, 한정되지는 않지만, 영장류 세포주는 침팬지, 비비, 명주원숭이, 짧은꼬리원숭이, 및 아프리카 녹색 원숭이(예, COS-1 세포주)에서 유도될 수 있다.

<117> 하기 표는 본 발명에 사용하는 세포주에서 유도된 영장류 종의 규명되지 않은 목록을 제공한다.

종	일반 명칭
판 트로글로디테스 스크웨인푸르티이 (<i>Pan troglodytes schweinfurthii</i>)	동양 침팬지
판 트로글로디테스(<i>Pan troglodytes</i>)	침팬지
판 파니스쿠스(<i>Pan paniscus</i>)	보노보
판 트로글로디테스 베루스 (<i>Pan troglodytes verus</i>)	서양 침팬지
고릴라 고릴라(<i>Gorilla gorilla</i>)	서양 저지 고릴라

폰고 피그마에우스 아벨리이 (<i>Pongo pygmaeus abelii</i>)	수마트라 오랑우탄
폰고 피그마에우스 피그마에우스 (<i>Pongo pygmaeus pygmaeus</i>)	보루네오 오랑우탄
힐로바테스 종(<i>Hylobates sp.</i>)	긴팔원숭이 종
마카카 아르크토이데스(<i>Macaca arctoides</i>)	스텝프-꼬리원숭이
마카카 파시쿨라리스(<i>Macaca fascicularis</i>)	긴꼬리 원숭이
세르코피테쿠스 미티스 알보구라리스 (<i>Cercopithecus mitis albogularis</i>)	파랑 원숭이
마카카 파시쿨라리스(<i>Macaca fascicularis</i>)	긴 꼬리 원숭이
마카카 물라타(<i>Macaca mulatta</i>)	붉은털 원숭이
마카카 실바누스(<i>Macaca sylvanus</i>)	꼬리없는 원숭이
판 트로글로디테스 베루스 (<i>Pan troglodytes verus</i>)	서양 침팬지
파피오 하마드리아스 아누비스 (<i>Papio hamadryas anubis</i>)	올리브 비비
파피오 종(<i>Papio sp.</i>)	비비 종
사귀누스 오에디푸스(<i>Saguinus oedipus</i>)	목화머리 타마린
칼리스틱스 자쿠스(<i>Callithrix jacchus</i>)	일반적인 마모셋
판 파니스쿠스(<i>Pan paniscus</i>)	보노보
호모 사피엔스(<i>Homo sapiens</i>)	인간

<119> 이러한 영장류 중에서 유도된 세포주는, 예를 들어 [Research Infrastructure to Promote Primate Molecular Biology](INPRIMAT; 스페인 바르셀로나) 및 [Biomedical Primate Research Centre](BPRC; 네덜란드 리즈윅)에 대한 웹사이트 상에서 찾을 수 있다.

<120> 재조합 폴리펩티드의 생성을 위한 인간 세포주의 사용은 잘 확립되어 있고 본 발명의 범위 내에 있다. 인간 숙주 세포주는 임의의 공급원에서 유도될 수 있고, 임의의 조직에서 유도될 수 있다. 재조합 단백질의 생성에 적당한 인간 세포주는 광범위하게 공지되어 있으며, 당업자에게 잘 공지된 다수의 시판사 및 사적 공급원에서 입수 가능하다.

<121> 또한, 다른 다양한 포유동물 중에서 유도된 숙주 세포주는 또한 본 발명의 범위 내에 있다. 이들은 소과(소), 돼지과(돼지), 양과(양), 개과(개), 말과(말) 및 염소과(염소) 세포주를 포함한다. 이러한 중에서 유도된 세포주는 당업계에 공지되어 있고, 보통의 당업자가 쉽게 입수가능하다.

<122> 설치류 세포주, 영장류 세포주, 및 다른 포유동물 군에서 유도된 다수의 다른 세포주, 및 이러한 동물 세포주의 배양을 위한 재료 및 방법은 당업계에 잘 확립되어 있고, 여러 공급원으로부터 광범위하게 입수 가능하다. 예를 들어, 입수가능한 세포주의 폭넓은 목록을 위한 American Type Culture Collection(ATCC; 버지니아주 마나사스) 참조. 세포 배양 배지는 GIBCO® 상표(Invitrogen™) 세포 배양 배지 및 Cellgro®의 Mediatech, Inc.사 배양 배지와 같은 각종 공급원으로부터 입수가능하다. 또한, 새로운 세포주 (제1 세포주 또는 확립된 세포주)의 새로운 확립은 또한 하나 이상의 비천연 아미노산을 갖는 재조합 단백질의 생성을 위한 숙주 세포와 같이 본 발명을 위해 사용된다.

<123> 본 발명의 사용을 위한 O-RS

<124> 본 발명은 신규 오르소고날 번역 시스템을 제공하는데, 여기에서 시스템은, 이전에 고려하지 않았던 포유동물 숙주 세포, 예를 들어 영장류 및 설치류 숙주 세포 시스템에서 비천연 아미노산을 포함하는 폴리펩티드를 생성하는 능력을 갖는다. 본 발명의 번역 시스템은 (i) 비천연 아미노산, (ii) 오르소고날 tRNA(O-tRNA), (iii) 이.콜라이의 아미노아실-tRNA 합성효소에서 유도된 오르소고날 아미노아실-tRNA 합성효소(O-RS), 및 (iv) 이러한 성분들을 함유하는 포유동물 숙주 세포(여기서, 포유동물 숙주 세포는 설치류 세포(예, 마우스 세포 또는 래트 세포), 또는 영장류 세포(예, 인간 세포 또는 마우스 세포)일 수 있음)를 포함한다. 본 발명은 또한 선택된 위치에서 하나 이상의 비천연 아미노산을 갖는 단백질을 생성하는 방법을 제공하는데, 여기서 방법은 상기 기재된 번역 시스템을 사용한다. 단백질 번역 동안 셀렉터 코돈에서 비천연 아미노산의 도입을 프로그래밍함으로써

방법은 포유동물 숙주 세포에서 비천연 아미노산의 도입을 초래한다.

<125> 본 발명에 사용하기 위한 O-RS 종은 특별히 한정되지는 않지만, 하기 성질을 일반적으로 갖는다. 본 발명에 사용하기 위한 O-RS는 일반적으로 이.콜라이 아미노아실-tRNA 합성효소, 예컨대 야생형 이.콜라이 류실-tRNA 합성효소(Ec LeuRS; 서열 번호 4에서 제공된 아미노산 서열, 및 서열 번호 5에서 제공된 폴리뉴클레오티드 서열) 또는 야생형 이.콜라이 티로실-tRNA 합성효소(Ec TyrRS; 서열 번호 6에서 제공된 아미노산 서열, 및 서열 번호 7에서 제공된 폴리뉴클레오티드 서열)에서 유도된다. 본 발명에 사용하기 위한 돌연변이체 O-RS 종(예, 서열 번호 4 또는 6에서 유도된 합성효소)은 본래 사카로미세스 세레비시에 (에스.세레비시에) 효모 숙주 세포 시스템에서 진화되고 선택된다.

<126> 본 발명에 사용하기 위한 그러한 다량의 O-RS 종은 당업계에 공지된다. 에스.세레비시에에서, 작용성 tRNA_{CUA}/RS 쌍은 상응한 이.콜라이 tRNA^{Tyr}/TyrRS 및 tRNA^{Leu}/류실-tRNA 합성효소 쌍에서 성공적으로 유도되었다. 예를 들면, 문헌[Chin 등, "An Expanded Eukaryotic Genetic Code," *Science* **301**, 964-967 (2003); Chin 등, "Progress Toward an Expanded Eukaryotic Genetic Code," *Chem Biol* **10**, 511-519 (2003); Deiters 등(2003), "Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*," *J Am Chem Soc* **125**:11782-11783; Summerer 등(2006), "A Genetically Encoded Fluorescent Amino Acid," *PNAS* **103**(26):9785-9789; 및 Wu 등, "A Genetically Encoded Photocaged Amino Acid," *J Am Chem Soc* **126**, 14306-14307 (2004)]를 참조한다. 또한 [제WO 2005/003294호(Deiters 등), "UNNATURAL REACTIVE AMINO ACID GENETIC CODE ADDITIONS,"(2004년 4월 16일 출원); 제WO 2006/034410호(Deiters 등), "ADDING PHOTOREGULATED AMINO ACID TO THE GENETIC CODE,"(2005년 9월 21일 출원); 및 제WO 2006/110182호(2005년 10월 27일 출원), (발명의 명칭: "ORTHOGONAL TRANSLATION COMPONENTS FOR THE VIVO INCORPORATION OF UNNATURAL AMINO ACID")]]를 참조한다. 이러한 문헌들 각각은 본원에 그 전문이 참고인용된다. 당업계에서 이러한 교시는 또한 본 발명에 사용하기 위한 추가의 O-RS 종의 구축을 위한 충분한 지침을 제공한다.

<127> 본 발명에 사용하기 위한 O-RS 종의 예는 도 16에 도시된다. O-RS를 코딩하는 폴리뉴클레오티드 서열은 서열 번호 8~56에 제공된다. 해당 O-RS 폴리펩티드 서열은 서열 번호 57~101에 제공된다. 전장 O-RS가 본 발명에 작용하는데 RS가 반드시 요구되지는 않는다는 것을 유념한다. RS 활성 부위 단편만의 발현은 도 16의 서열 일부에서 도시된 바와 같이 본 발명에 사용하기에 충분할 수 있다. 도 16의 O-RS 서열 각각은 본원에 인용된 문헌들에 기재된 바와 같이 당업계에 공지된다.

<128> 본 발명은 도 16에 도시된 O-RS의 사용을 한정하려는 의도가 아니다. 본 발명에 사용하기 위해 본원에서 확인된 임의의 O-RS 종은 유도를 위한 템플릿이, 예를 들어 존재하는 O-RS 종 내에 보존적 아미노산 치환기를 제조함으로써 본 발명 내에서 사용될 수 있는 여전히 다른 O-RS 종인 것과 같은 작용을 할 수 있다. 새롭게 유도된 O-RS가 비천연 아미노산으로 해당 O-tRNA를 우선적으로 아미노아실화하는 생물학적 활성을 보유하는 조건으로 허용되는 치환기의 수는 한정되지 않는다.

<129> O-RS의 생물학적 활성은 본 발명에 사용하도록 제시되는 또다른 "기준" O-RS의 생물학적 활성의 백분율로서 제시될 수 있다. 예를 들어, 본 발명에 사용되는 O-RS는 O-tRNA, 비천연 아미노산, 및 공지된 서열의 아미노아실-tRNA 합성효소(RS)를 포함하는 "기준" 번역 시스템에 대해 관찰되는 효율의 50% 이상인 효율로 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는 O-RS를 포함할 수 있다. 임계값 활성의 선택은 임의적이며, "50%"의 값은 예로서만 제공된다. 사실상, 본원에 사용되는 O-RS는 "기준" 번역 시스템에 대해 관찰되는 효율의 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% 또는 99%인 효율로 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는 O-RS를 포함할 수 있다.

<130> 본원에 사용되기 위한 O-tRNA 종

<131> 본원에 사용되기 위한 돌연변이체 O-RS 종은 본래 효모 숙주 세포 시스템에서 진화되고 선택된다. 효모 숙주 세포 시스템에서 진화된 O-RS 종은 비천연 아미노산을 이용하여 O-RS에 의해 우선적으로 아미노아실화된 적당한 O-tRNA와 함께 작용한다. 예를 들어, 효모 숙주 시스템에서 작용하는 억제자 O-tRNA 종은, 예를 들어 서열 번호 1 또는 2의 tRNA일 수 있다.

<132> 대조적으로, 본 발명의 번역 시스템은 포유동물 숙주 세포(예, 설치류 세포 또는 영장류 세포)에서 오르소고날인 상이한 O-tRNA를 사용한다. 예를 들어, 본 발명에 사용하기 위한 O-tRNA는, 서열 번호 3에 제공된 바와 같이, 바실러스 스테아로써모필러스 엠버 억제자 티로실-tRNA_{CUA}(Bs-tRNA^{Tyr}CUA)일 수 있다. 본 발명은 tRNA 서열의 변이체로서 이러한 일 tRNA 종으로 한정하려는 의도가 아니며, 마찬가지로 다른 유기체(예, 다른 진정세균

중)에서 유도된 tRNA 종으로서 바실러스 스테아로써모필러스에서 유도된 다른 tRNA 또한 고려된다.

<133> 오르소고날 tRNA/아미노아실-tRNA 합성효소 기술

<134> 본 발명의 신규 조성물 및 방법의 이해는 오르소고날 tRNA 및 오르소고날 아미노아실-tRNA 합성효소 쌍과 관련된 활성의 이해를 필요로 한다. 부가적 비천연 아미노산을 유전 코드에 부가하기 위해, 숙주 번역 기구에서 효율적으로 작용할 수 있으면서도, 당해 번역 시스템에 대한 내인성 합성효소 및 tRNA와 무관하게 작용한다는 의미에서, 당해 번역 시스템에 대해 "오르소고날" 성질을 갖는 아미노아실-tRNA 합성효소 및 적절한 tRNA를 포함하는 신규 오르소고날 쌍이 필요하다. 오르소고날 쌍의 원하는 특징은 임의의 내인성 tRNA에 의해 해독되지 않는 특정한 코돈, 예를 들어 셀렉터 코돈만을 해독하거나 인식하는 tRNA, 및 그의 동족체 tRNA를 1개의 특정한 비천연 아미노산으로만 우선적으로 아미노아실화하는 (또는 "충전시키는") 아미노아실-tRNA 합성효소를 포함한다. 또한, O-tRNA는 전형적으로 내인성 합성효소에 의해 아미노아실화되지 않는다 (또는 아미노아실화가 잘 되지 않는다, 즉 충전이 잘되지 않는다). 예를 들어, 이. 콜라이 숙주 시스템에서 오르소고날 쌍은, 예를 들어 이. 콜라이에서 40종이 존재하는 임의의 내인성 tRNA와 교차-반응하지 않는 아미노아실-tRNA 합성효소, 및 예를 들어, 이. 콜라이에서 21종이 존재하는 임의의 내인성 합성효소에 의해 아미노아실화되지 않는 오르소고날 tRNA를 포함할 것이다.

<135> 하나 이상의 비천연 아미노산을 포함하는 단백질의 제조에 적당한 오르소고날 번역 시스템의 일반 원칙이, 오르소고날 번역 시스템의 생성을 위한 일반 방법과 같이, 당업계에 공지되어 있다. 예를 들어, 국제 특허 공개 공보 제WO 2002/086075호(발명의 명칭: "METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGONAL tRNA-AMINO ACYL-tRNA SYNTHETASE PAIRS"); 제WO 2002/085923호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS"); 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE"); 제WO 2005/019415호(2004년 7월 7일 출원); 제WO 2005/007870호(2004년 7월 7일 출원); 제WO 2005/007624호(2004년 7월 7일 출원); 국제 출원 번호 PCT/US2005/039210호(2005년 10월 27일 출원)(발명의 명칭: "ORTHOGONAL TRANSLATION COMPONENTS FOR THE IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS"); 및 미국 가출원 번호 60/783,497(2006년 3월 17일 출원)(발명의 명칭: "SYSTEMS FOR THE EXPRESSION OF ORTHOGONAL TRANSLATION COMPONENTS IN EUBACTERIAL HOST CELLS")를 참조한다. 이 공보들 각각은 전체적으로 본원에 참조 인용된다. 비천연 아미노산이 도입된 오르소고날 번역 시스템 및 이의 생성 및 사용 방법의 논의를 위해, 또한 문헌 [Wang 및 Schultz, "Expanding the Genetic Code", *Angewandte Chemie Int. Ed.*, 44(1):34-66(2005)], 문헌 [Xie 및 Schultz, "An Expanding Genetic Code", *Methods* 36(3):227-238(2005)]; 문헌 [Xie 및 Schultz, "Adding Amino Acids to the Genetic Repertoire", *Curr. Opinion in Chemical Biology* 9(6): 548-554(2005)]; 및 문헌 [Wang 등, "Expanding the Genetic Code", *Anna. Rev. Biophys. Biomol. Struct.*, 35:225-249(2006)]를 참조하고; 이들 내용은 각기 전체적으로 본원에 참조 인용된다.

<136> 오르소고날 번역 시스템

<137> 오르소고날 번역 시스템은 일반적으로 오르소고날 tRNA(O-tRNA), 오르소고날 아미노아실 tRNA 합성효소(O-RS) 및 비천연 아미노산을 포함하는 (설치류 세포 또는 영장류 세포와 같은 포유동물 세포일 수 있는) 세포를 포함하는데, 여기에서 O-RS는 O-tRNA를 비천연 아미노산으로 아미노아실화한다. 본 발명의 오르소고날 쌍은 O-tRNA, 예를 들어 억제자 tRNA, 프레임이동(frameshift) tRNA 등, 및 동족체 O-RS를 포함한다. 본 발명의 오르소고날 시스템은 전형적으로, 숙주 세포와 관련하여, 또는 숙주 세포의 부재 하에 O-tRNA/O-RS 쌍을 포함할 수 있다. 다중 성분 시스템에 부가하여, 본 발명은 또한 신규 개별 성분, 예를 들어 신규 오르소고날 아미노아실-tRNA 합성효소 폴리펩티드(예를 들어, 서열 번호 57~101), 및 그 폴리펩티드를 코딩하는 폴리뉴클레오티드(예를 들어, 서열 번호 8~56)를 제공한다.

<138> 일반적으로, 오르소고날 쌍이 셀렉터 코돈을 인식하고 이 셀렉터 코돈에 대한 반응으로 아미노산을 로딩하는 경우, 오르소고날 쌍은 셀렉터 코돈을 "억제"한다고 언급된다. 즉, 번역 시스템(예, 세포)의 내인성 기구에 의해 인식되지 않는 셀렉터 코돈은 통상적으로 충전되지 않고, 이는 다른 방식으로서는 핵산으로부터 번역될 폴리펩티드의 생성을 차단할 수 있다. 오르소고날 쌍 시스템에서, O-RS는 O-tRNA를 특정 비천연 아미노산으로 아미노아실화한다. 충전된 O-tRNA는 셀렉터 코돈을 인식하고, 셀렉터 코돈에 의해 유발되는 번역 블록을 억제한다.

<139> 일부 측면에서, 본 발명의 O-tRNA는 셀렉터 코돈을 인식하고, 본원의 서열목록에 나와 있는 바와 같은 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩되는 O-tRNA의 억제 효율과 비교할 때 셀렉터 코돈에 대한 반응으로 동족체 합성효소의 존재 하에 적어도 대략 예를 들어, 45%, 50%, 60%, 75%, 80% 또는 90% 이상의 억제 효율을 포함한다.

- <140> 일부 구체예에서, O-RS 및 O-tRNA가 함께 있는 경우의 억제 효율은 O-RS가 없는 경우의 O-tRNA의 억제 효율보다 대략 예를 들어, 5배, 10배, 15배, 20배 또는 25배 이상 더 높다. 일부 측면에서, O-RS 및 O-tRNA가 함께 있는 경우의 억제 효율은 본원의 서열목록에 나와 있는 바와 같은 오르소고날 합성효소 쌍의 억제 효율의 적어도 대략 예를 들어, 35%, 40%, 45%, 50%, 60%, 75%, 80% 또는 90% 이상이다.
- <141> 숙주 세포는 비천연 아미노산을 성장하는 폴리펩티드 채로, 예를 들어 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산을 통해 도입하기 위해 O-tRNA/O-RS 쌍을 사용하는데, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함한다. 특정 바람직한 측면에서, 세포는 하나 이상의 부가적 O-tRNA/O-RS 쌍을 포함할 수 있는데, 여기에서 부가적 O-tRNA가 상이한 비천연 아미노산을 갖는 부가적 O-RS에 의해 로딩된다. 예를 들어, O-tRNA 중 하나는 4 염기 코돈을 인식할 수 있고, 다른 O-tRNA는 정지 코돈을 인식할 수 있다. 대안적으로, 다수의 상이한 정지 코돈 또는 다수의 상이한 4 염기 코돈이 동일한 코딩 핵산 내에 사용될 수 있다.
- <142> 언급된 바와 같이, 일부 구체예에서, 세포 또는 기타 번역 시스템 내에는 하나 초과와 비천연 아미노산을 폴리펩티드 내로 도입하도록 하는 다수의 O-tRNA/O-RS 쌍이 존재한다. 예를 들어, 세포는 부가적인 상이한 O-tRNA/O-RS 쌍 및 제2 비천연 아미노산을 포함할 수 있는데, 여기에서 이 부가적 O-tRNA는 제2 셀렉터 코돈을 인식하고, 이 부가적 O-RS는 O-tRNA를 제2 비천연 아미노산으로 우선적으로 아미노아실화한다. 예를 들어, O-tRNA/O-RS 쌍(여기에서, O-tRNA는 예를 들어 엠버 셀렉터 코돈을 인식함)을 포함하는 세포는 제2 오르소고날 쌍을 추가로 포함할 수 있는데, 여기에서 제2 O-tRNA는 상이한 셀렉터 코돈, 예를 들어 오팔 코돈, 4 염기 코돈 등을 인식한다. 바람직하게는, 상이한 오르소고날 쌍은 상이한 공급원으로부터 유도되며, 이는 상이한 셀렉터 코돈의 인식을 용이하게 할 수 있다.
- <143> 특정 구체예에서, 시스템은 오르소고날 tRNA(O-tRNA), 오르소고날 아미노아실-tRNA 합성효소(O-RS), 비천연 아미노산, 및 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산을 포함한, 비제한적으로 설치류 및 영장류 세포를 포함하는 포유동물 세포와 같은 세포를 포함하는데, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함한다. 번역 시스템은 본원에 기재된 바와 같은 O-tRNA/O-RS 쌍 및 비천연 아미노산과 조합된 무세포 시스템, 예를 들어 다양한 시중 입수가능한 "시험관내" 전사/번역 시스템들 중 임의의 시스템일 수도 있다.
- <144> O-tRNA 및/또는 O-RS는 천연적으로 발생될 수 있거나, 예를 들어 각종 유기체 중 임의의 유기체로부터 유도된 tRNA의 라이브러리 및/또는 RS의 라이브러리를 발생시키고/시키거나 다양한 이용가능한 돌연변이유발 기법 중 임의의 기법을 이용하여 천연 발생 tRNA 및/또는 RS를 돌연변이화함으로써 유도될 수 있다. 예를 들어, 오르소고날 tRNA/아미노아실-tRNA 합성효소 쌍의 생성을 위한 한 기법은, 예를 들어 숙주 세포 외의 한 공급원 또는 다수의 공급원으로부터 유도된 (숙주 세포에 대한) 이중성 tRNA/합성효소 쌍을 숙주 세포 내로 이입시키는 단계를 포함한다. 그 이중성 합성효소 후보의 성질은, 예를 들어 임의의 숙주 세포 tRNA를 충전하지 않는 성질을 포함하며, 이중성 tRNA 후보물질의 성질은, 예를 들어 그것이 어떠한 의 숙주 세포 합성효소에 의해서도 아미노아실화되지 않는다는 성질을 포함한다. 또한, 이중성 tRNA는 모든 숙주 세포 합성효소에 대해 오르소고날성을 가진다. 오르소고날 쌍을 생성하기 위한 제2 기법은 O-tRNA 또는 O-RS를 선별하고/하거나 선택하는 기초가 되는 돌연변이체 라이브러리를 생성시키는 단계를 포함한다. 이 기법들은 조합될 수도 있다.
- <145> 오르소고날 tRNA(O-tRNA)
- <146> 본 발명의 오르소고날 tRNA(O-tRNA)는 바람직하게, 예를 들어 생체내 또는 시험관내, O-tRNA에 의해 인식되는 셀렉터 코돈을 포함하는 폴리뉴클레오티드에 의해 코딩되는 단백질 내로의 비천연 아미노산의 도입을 매개한다. 특정 구체예에서, 본 발명의 O-tRNA는 본원의 서열목록에서 O-tRNA 서열에 기재된 바와 같은 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩되는 O-tRNA와 비교 시, 셀렉터 코돈에 대한 반응으로 동족체 합성효소의 존재 하에 적어도 대략 예를 들어, 45%, 50%, 60%, 75%, 80% 또는 90% 이상의 억제 효율을 포함한다.
- <147> 억제 효율은 당업계에 공지된 다수의 검정들 중 임의의 검정에 의해 측정될 수 있다. 예를 들어, β -갈락토시다제 리포터 검정을 이용할 수 있는데, 예를 들어 (구축물이 lacZ 핵산 서열 내에 셀렉터 코돈을 가지는 경우) 유도체화된 lacZ 플라스미드가 본 발명의 O-tRNA를 포함하는 플라스미드와 함께 적절한 유기체(예를 들어, 오르소고날 성분들이 사용될 수 있는 유기체)로부터의 세포 내로 도입된다. 동족체 합성효소도 (발현된 때, 그 동족체 합성효소를 코딩하는 폴리펩티드 또는 폴리뉴클레오티드로서) 도입될 수 있다. 세포가 배지 중에서 원하는 밀도, 예를 들어 약 0.5의 OD₆₀₀까지 성장하고, β -갈락토시다제 검정을, 예컨대 베타플루오르(BetaFluor)TM β -

갈락토시다제 분석 키트[노바젠(Novagen)]를 사용하여 수행한다. 억제율(%)은 비교 대조군에 대한 상대적인 샘플 활성율(%)로서, 예를 들어 셀렉터 코돈이 아닌 원하는 위치에 상응하는 센스 코돈을 가지는, 유도체화된 lacZ 구축물로부터 관찰된 값으로서 계산될 수 있다.

<148> 본 발명에서 사용하기 위한 O-tRNA의 예가 본원의 서열목록에 기재되어 있는데, 예를 들어 도 16 및 서열 번호 3을 참조한다. 또한, 본원의 개시 내용은 부가적 동등 O-tRNA 종의 설계에 대한 지침을 제공한다. RNA 분자, 예컨대 O-RS mRNA 또는 O-tRNA 분자에서, 티민(T)은 소정의 서열 또는 이의 상보체에 대해 우라실(U)로 치환된다 (또는 코딩 DNA의 경우, 이와 반대로 우라실(U)이 티민으로 치환된다). 대체로 기능성에 있어 동등한 분자를 발생시키기 위해 염기에 대한 부가적 변형이 존재할 수도 있다.

<149> 본 발명은 또한 본원의 특별한 O-tRNA에 상응하는 O-tRNA의 보존적 변이를 포괄한다. 예를 들어, O-tRNA의 보존적 변이에는, 예를 들어 본원의 서열목록에 기재된 바와 같이, 특별한 O-tRNA와 유사하게 기능하며 적절한 자가-상보성에 의해 tRNA L-형상의 구조를 유지하지만, 예를 들어 본원의 서열목록 또는 도 16에 기재된 것들과 동일한 서열을 가지지 않는 분자들로서, 바람직하게는 야생형 tRNA 분자 이외의 분자들이 포함된다.

<150> O-tRNA를 포함하는 조성물은 오르소고날 아미노아실-tRNA 합성효소(O-RS)를 추가로 포함할 수 있는데, 여기에서 ORS는 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화한다. 특정 구체예에서, O-tRNA를 포함하는 조성물은 (예를 들어, 시험관내 또는 생체내) 번역 시스템을 추가로 포함할 수 있다. 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드로서 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 포함하는 핵산 또는 이들 중 하나 이상의 조합이 세포 내에 존재할 수도 있다.

<151> 오르소고날 tRNA(O-tRNA)의 생성 방법도 또한 본 발명의 한 특성이다. 그 방법에 의해 생성된 O-tRNA도 또한 본 발명의 한 특성이다. 본 발명의 특정 구체예에서, O-tRNA는 돌연변이체의 라이브러리를 발생시킴으로써 생성될 수 있다. 돌연변이체 tRNA의 라이브러리는 당업계에 공지되어 있는 각종 돌연변이유발 기법들에 의해 발생될 수 있다. 예를 들어, 돌연변이체 tRNA는, 예를 들어 서열 번호 3의 O-tRNA의 부위-특이적 돌연변이유발, 랜덤 점 돌연변이유발, 상동성 재조합, DNA 셔플링(shuffling) 또는 다른 반복적 돌연변이유발법, 키메라 구축, 또는 이들의 임의의 조합에 의해 발생될 수 있다.

<152> 부가적 돌연변이는 tRNA의 원하는 루프 또는 영역, 예를 들어 안티코돈 루프, 수용체 줄기(stem), D 암(arm) 또는 루프, 가변성 루프, TPC 암 또는 루프, tRNA 분자의 다른 영역 또는 이들의 조합에서 특별한 위치(들), 예를 들어 비보존적 위치(들), 보존적 위치, 무작위 위치(들), 또는 이들의 조합 위치에서 도입될 수 있다. 전형적으로, tRNA 내의 돌연변이는 셀렉터 코돈의 인식을 허용하도록 돌연변이체 tRNA의 라이브러리의 각 구성원의 안티코돈 루프를 돌연변이화하는 단계를 포함한다. 방법은 추가 서열을 O-tRNA에 부가하는 단계를 추가로 포함할 수 있다. 전형적으로, O-tRNA는 원하는 RS에 대한 그의 친화성을 보존하면서 출발 물질, 예를 들어 복수개의 tRNA 서열에 비해 원하는 유기체에 대한 개선된 오르소고날성을 보유한다.

<153> 방법은 임의적으로 tRNA 및/또는 아미노아실-tRNA 합성효소의 서열의 유사성 (및/또는 추정된 상동성)을 분석하여, 특정 유기체에 대해 오르소고날성을 가지는 것으로 보이는 O-tRNA, O-RS 및/또는 이들의 쌍에 대한 잠재적 후보를 결정하는 단계를 포함한다. 당업계에 공지되어 있으며 본원에 기재되어 있는 컴퓨터 프로그램이 상기 분석에 사용될 수 있는데, 예를 들어 BLAST 및 파일업(pileup) 프로그램이 사용될 수 있다. 한 예에서, 포유동물 세포, 예컨대 설치류 세포 및 영장류 세포에서 사용하기 위한 잠재적인 오르소고날 번역 성분들을 선택하기 위해, 숙주 유기체와의 높은 서열 유사성을 나타내지 않는 합성효소 및/또는 tRNA를 선택한다.

<154> 전형적으로, O-tRNA는 제1 종의 세포로 이루어진 집단을 예를 들어, 음성 선택함으로써 획득되는데, 여기에서 세포는 복수개의 잠재적 O-tRNA의 구성원을 포함한다. 그 음성 선택은 세포에 대해 내인성인 아미노아실-tRNA 합성효소(RS)에 의해 아미노아실화되는 잠재적 O-tRNA로 이루어진 라이브러리의 구성원을 포함하는 세포를 제거한다. 이는 제1 종의 세포에 대해 오르소고날성을 갖는 tRNA의 풀을 제공한다.

<155> 특정 구체예에서, 음성 선택에서, 셀렉터 코돈(들)은 음성 선택 마커, 예를 들어 항생제 내성을 부여하는 효소, 예컨대 β-락타마제, 검출가능한 생성물을 부여하는 효소, 예컨대 β-갈락토시다제, 클로람페니콜 아세틸트랜스퍼라제(CAT), 예를 들어 비필수 위치(예를 들어, 여전히 기능성 바나제(barnase)를 생산하는 위치)에 있는 바나제 등과 같은 독성 물질 등을 코딩하는 폴리뉴클레오티드 내로 도입된다. 선별/선택은 임의적으로 선별제(예를 들어, 앰피실린과 같은 항생제)의 존재 하에 세포 집단을 성장시킴으로써 수행된다. 한 구체예에서, 선택제의 농도는 다양하다.

<156> 예를 들어, 억제자 tRNA의 활성을 측정하기 위해, 셀렉터 코돈의 생체내 억제에 기초한, 예를 들어 음성 선택

마커를 코딩하는 폴리뉴클레오티드, 예를 들어 β -락타마제(*bla*)에 대한 유전자 내로 도입된 넌센스(예를 들어, 정지) 또는 프레임이동 돌연변이에 기초한 선택 시스템이 사용된다. 예를 들어, 특별한 위치(예를 들어, A184)에서 셀렉터 코돈을 가진 폴리뉴클레오티드 변이체, 예를 들어 *bla* 변이체가 구축된다. 세포, 예컨대 세균은 이 폴리뉴클레오티드로 형질전환된다. 내인성 이. 콜라이 합성효소에 의해 효율적으로 충전될 수 없는 오르소고날 tRNA의 경우, 항생제 내성, 예컨대 앰피실린 내성은 플라스미드로 형질전환되지 않은 세균에 대한 항생제 내성과 비슷하거나 이보다 더 낮아야 한다. tRNA가 오르소고날 tRNA가 아닌 경우, 또는 tRNA를 충전시킬 수 있는 이중 합성효소가 상기 시스템 내에서 동시-발현되는 경우, 보다 높은 수준의 항생제, 예컨대 앰피실린 내성이 관찰된다. 플라스미드로 형질전환되지 않은 세포와 거의 동등한 항생제 농도를 가진 LB 아가 플레이트 상에서 성장할 수 없는 세포, 예를 들어 세균이 선택된다.

<157> 독성 물질(예를 들어, 리보뉴클레아제 또는 바나제)의 경우, 복수개의 잠재적 tRNA의 구성원이 내인성 숙주, 예컨대 에세리키아 콜라이 합성효소에 의해 아미노아실화되는 경우(즉, 상기 숙주, 예컨대 에세리키아 콜라이 합성효소에 대한 오르소고날 tRNA가 아닌 경우), 셀렉터 코돈은 억제되고, 생성된 독성 폴리뉴클레오티드 생성물은 세포 사멸을 초래한다. 오르소고날 tRNA 또는 비-기능성 tRNA를 보유하는 세포는 생존한다.

<158> 한 구체예에서, 원하는 유기체에 대해 오르소고날성을 가지는 tRNA의 풀은, 셀렉터 코돈이 예를 들어, β -락타마제 유전자와 같은 약물 내성 유전자에 의해 코딩되는 양성 선택 마커에 위치되어 있는 양성 선택에 적용된다. 양성 선택은, 세포에 대해 오르소고날성을 가지는 tRNA의 풀의 구성원을 코딩하거나 포함하는 폴리뉴클레오티드, 양성 선택 마커를 코딩하는 폴리뉴클레오티드, 및 동족체 RS를 코딩하는 폴리뉴클레오티드를 포함하는 세포에 대해 수행된다. 특정 구체예에서, 제2 세포 집단은 음성 선택에 의해 제거되지 않은 세포를 포함한다. 폴리뉴클레오티드는 세포 내에서 발현되고, 세포는 선택제, 예컨대 앰피실린의 존재 하에 성장된다. 그 후, tRNA는 동시 발현된 동족체 합성효소에 의해 아미노아실화되며 이 셀렉터 코돈에 대한 반응으로 아미노산을 삽입하는 그것의 능력에 대해 선택된다. 전형적으로, 이 세포들은 비-기능성 tRNA(들), 또는 관심 합성효소에 의해 효율적으로 인식될 수 없는 tRNA를 보유하는 세포에 비해 억제 효율에서의 상승을 보인다. 비-기능성 tRNA, 또는 관심 합성효소에 의해 효율적으로 인식되지 않는 tRNA를 보유하는 세포는 항생제에 대해 민감하다. 그러므로, (i) 내인성 숙주, 예를 들어 에세리키아 콜라이 합성효소에 대한 기질이 아니고; (ii) 관심 합성효소에 의해 아미노아실화될 수 있으며; (iii) 번역에 있어서 기능성을 나타내는 tRNA는 상기 양 선택 모두에서 생존한다.

<159> 따라서, 동일한 마커가 그 자신이 선별되는 배경에 따라 양성 마커 또는 음성 마커일 수 있다. 즉, 마커는 그 자신에 대해 선별되는 경우에는 양성 마커이지만, 그 자신에 반하여 선별되는 경우에는 음성 마커이다.

<160> 전술된 방법에서 선별, 예를 들어 양성 선택, 음성 선택 또는 양성 선택 및 음성 선택 양자 모두의 엄격도는 임의적으로 선택 엄격도를 달리하는 것을 포함한다. 예를 들어, 바나제는 극히 독성이 강한 단백질이기 때문에, 음성 선택의 엄격도는 상이한 수의 셀렉터 코돈을 바나제 유전자 내로 도입하고/하거나 유도가능한 프로모터를 사용함으로써 조절될 수 있다. 또 다른 예에서, 선택 또는 선별제의 농도(예를 들어, 앰피실린 농도)는 다양하다. 본 발명의 일부 측면에서, 엄격도는 원하는 활성이 초기 순환 동안 낮을 수 있기 때문에 다양하다. 그러므로, 엄격도가 보다 낮은 선택 기준이 초기 수행에서 적용되고, 엄격도가 보다 높은 기준이 선별의 후기 수행에서 적용된다. 특정 구체예에서, 음성 선택, 양성 선택 또는 음성 선택 및 양성 선택 양자 모두가 수회 반복될 수 있다. 다수의 상이한 음성 선택 마커, 양성 선택 마커 또는 음성 선택 마커 및 양성 선택 마커 양자 모두가 사용될 수 있다. 특정 구체예에서, 양성 선택 마커와 음성 선택 마커는 동일할 수 있다.

<161> 다른 유형의 선택/선별이 본 발명에서 오르소고날 번역 성분, 예를 들어 O-tRNA, O-RS, 및 셀렉터 코돈에 반응하여 비천연 아미노산을 로딩하는 O-tRNA/O-RS 쌍을 생성시키는 데 사용될 수 있다. 예를 들어, 음성 선택 마커, 양성 선택 마커 또는 음성 선택 마커 및 양성 선택 마커 양자 모두가 적절한 반응물의 존재 하에 형광을 나타내거나 발광 반응을 촉진하는 마커를 포함할 수 있다. 또 다른 구체예에서, 마커의 생성물은 형광-활성화 세포 분류(FACS) 또는 발광에 의해 검출된다. 임의적으로, 그 마커는 친화성에 기초한 선별 마커를 포함한다. 또한, 문헌 [Francisco, J. A. 등(1993) *Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface. Proc Natl Acad Sci U S A.* 90: 10444-8]을 참조한다.

<162> 재조합 오르소고날 tRNA를 생성시키기 위한 부가적 방법을, 예를 들어 국제 특허 공개 공보 제WO 2002/086075호(발명의 명칭: "METHODS AND COMPOSITIONS FOR THE PRODUCTION OF ORTHOGONAL tRNA AMINOACYL-tRNA SYNTHETASE PAIRS"), 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE") 및 제WO

2005/019415호(2004년 7월 7일 출원)에서 찾아볼 수 있다. 또한, 문헌 [Forster 등(2003), *Programming peptidomimetic synthetases by translating genetic codes designed de novo*, PNAS 100(11):6353-6357]; 및 문헌 [Feng 등(2003), *Expanding tRNA recognition of a tRNA synthetase by a single amino acid change*, PNAS 100(10): 5676-5681]을 참조한다.

<163> 오르소고날 아미노아실-tRNA 합성효소(O-RS)

<164> 본 발명의 O-RS는 시험관내 또는 생체내에서 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화한다. 본 발명의 O-RS는 O-RS를 포함하는 폴리펩티드 및/또는 O-RS 또는 이의 일부를 코딩하는 폴리뉴클레오티드에 의해 번역 시스템, 예를 들어 세포에 제공될 수 있다. 예를 들어, 한 예의 O-RS는 서열 번호 57~101에서 선택된 아미노산 서열, 또는 이의 보존적 변이체를 포함한다. 또 다른 예에서, O-RS 또는 이의 일부는 본원의 서열목록 또는 실시예에 기재된 서열을 포함하는, 아미노산을 코딩하는 폴리뉴클레오티드 서열 또는 이들의 상보적 폴리뉴클레오티드 서열에 의해 코딩된다. 예를 들어, 서열 번호 8~56의 폴리뉴클레오티드를 참조한다.

<165> O-tRNA와 함께 사용하기 위한 오르소고날 아미노아실-tRNA 합성효소(O-RS), 예를 들어 O-RS를 동정하는 방법도 또한 본 발명의 한 특성이다. 예를 들어, 그 방법은 제1 종의 세포로 이루어진 집단을 선택, 예를 들어 양성 선택하는 단계를 포함하는데, 여기에서 세포는 1) 복수개의 아미노아실-tRNA 합성효소(RS)의 구성원(예를 들어, 복수개의 RS에는 돌연변이체 RS, 상기 제1 종 외의 종으로부터 유도된 RS, 또는 돌연변이체 RS 및 상기 제1 종 외의 종으로부터 유도된 RS 양자 모두가 포함될 수 있다); 2) (예를 들어, 하나 이상의 종으로부터의) 오르소고날 tRNA(O-tRNA); 및 3) (예를 들어, 양성) 선택 마커를 코딩하며 하나 이상의 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 개별적으로 포함한다. 세포는 복수개의 RS 구성원을 가지지 않거나 감소된 양의 복수개의 RS 구성원을 가진 세포에 비해 억제 효율의 상승을 보여주는 세포에 대해 선택되거나 선별된다. 억제 효율은 당업계에 공지되어 있으며 본원에 기재된 바와 같은 기법에 의해 측정될 수 있다. 억제 효율이 상승된 세포는 O-tRNA를 아미노아실화하는 활성 RS를 포함한다. 제1 종으로부터의 tRNA로 이루어진 제1 세트의 활성 RS에 의한 (시험관내 또는 생체내) 아미노아실화 수준은 제2 종으로부터의 tRNA로 이루어진 제2 세트의 활성 RS에 의한 (시험관내 또는 생체내) 아미노아실화 수준과 비교된다. 아미노아실화 수준은 검출가능한 물질(예를 들어, 표지된 비천연 아미노산)에 의해 결정될 수 있다. 전형적으로, 제1 tRNA 세트에 비해 제2 tRNA 세트를 더 효율적으로 아미노아실화하는 활성 RS를 선택함으로써, O-tRNA와 함께 사용할 효율적인 (최적화된) 오르소고날 아미노아실-tRNA 합성효소를 제공한다. 이 방법에 의해 동정된 O-RS도 또한 본 발명의 한 특성이다.

<166> 아미노아실화를 결정하기 위해 다수의 검정들 중 임의의 검정을 이용할 수 있다. 이 검정은 시험관내 또는 생체내에서 수행될 수 있다. 예를 들어, 시험관내 아미노아실화 검정은, 예를 들어 문헌 [Hoben 및 So11(1985), Methods Enzymol. 113:55-59]에 기재되어 있다. 아미노아실화는 리포터를 오르소고날 번역 성분들과 함께 사용하고 단백질을 코딩하는, 하나 이상의 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 발현하는 세포 내에서 상기 리포터를 검출함으로써 결정될 수도 있다. 국제 특허 공개 공보 제WO 2002/085923호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS"); 및 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")도 또한 참조한다.

<167> 동정된 O-RS는 원하는 비천연 아미노산만이 O-tRNA에 충전되고 공통된 20종의 아미노산들 중 임의의 아미노산은 O-tRNA에 충전되지 않도록 합성효소의 기질 특이성이 변경되도록 더 조작될 수 있다. 비천연 아미노산에 대한 기질 특이성을 가진 오르소고날 아미노아실-tRNA 합성효소를 발생시키는 방법은, 예를 들어 상기 합성효소 내의 활성 부위, 상기 합성효소 내의 편집(editing) 기작 부위, 합성효소의 상이한 도메인들의 조합에 의한 상이한 부위 등에서 상기 합성효소를 돌연변이화하는 단계, 및 선택 과정을 적용하는 단계를 포함한다. 양성 선택 후 음성 선택이 이어지는 조합에 기초한 기법이 이용된다. 양성 선택에서, 양성 마커의 비필수 위치(들)에 도입된 셀렉터 코돈의 억제는 세포가 양성 선택 압력 하에서 생존할 수 있게 한다. 따라서, 생존 세포는 천연 아미노산 및 비천연 아미노산 양자 모두의 존재 하에서 오르소고날 억제자 tRNA를 천연 아미노산 또는 비천연 아미노산으로 충전시키는 활성 합성효소를 코딩한다. 음성 선택에서, 음성 마커의 비필수 위치(들)에 도입된 셀렉터 코돈의 억제는 비천연 아미노산 특이성을 가진 합성효소를 제거한다. 음성 선택 및 양성 선택 양자 모두에서의 생존 세포는 오르소고날 억제자 tRNA를 비천연 아미노산으로만 아미노아실화하는(충전시키는) 합성효소를 코딩한다. 이어서, 이 합성효소는 부가적 돌연변이유발, 예를 들어 DNA 서플링 또는 다른 반복적 돌연변이유발법으로 돌연변이시킬 수 있다.

<168> 돌연변이체 O-RS의 라이브러리는 당업계에 공지된 다양한 돌연변이유발 기법을 이용하여 발생시킬 수 있다. 예를 들어, 돌연변이체 RS는 부위-특이적 돌연변이유발, 랜덤 점 돌연변이유발, 상동성 재조합, DNA 서플링 또는

다른 반복적 돌연변이유발법, 키메라 구축 또는 이의 임의의 조합에 의해 생성될 수 있다. 예를 들어, 돌연변이체 RS로 이루어진 라이브러리는 2개 이상의 다른, 예컨대 보다 작고 보다 덜 다양한 "서브-라이브러리"로부터 생성될 수 있다. RS의 키메라 라이브러리도 본 발명에 포함된다. 다양한 유기체(예를 들어, 진정세균 또는 고세균과 같은 미생물)로부터의 tRNA 합성효소로 이루어진 라이브러리, 예컨대 천연 다양성을 포함하는 라이브러리 [예를 들어, 미국 특허 제6,238,884호(Short 등); 미국 특허 제5,756,316호(Schallenger 등); 미국 특허 제5,783,431호(Petersen 등); 미국 특허 제5,824,485호(Thompson 등); 및 미국 특허 제5,958,672호(Short 등) 참조]가 임의적으로 구축되어 오르소고날 쌍에 대해 선별된다는 것을 주목하도록 한다.

<169> 일단 합성효소가 양성 및 음성 선택/선별 기법에 적용되면, 이 합성효소를 더 돌연변이화할 수 있다. 예를 들어, O-RS를 코딩하는 핵산을 단리할 수 있고; (예를 들어, 랜덤 돌연변이유발, 부위-특이적 돌연변이유발, 재조합 또는 이들의 임의의 조합에 의해) 돌연변이된 O-RS를 코딩하는 폴리뉴클레오타이드 세트를 상기 핵산으로부터 발생시킬 수 있고; 이 개별 단계들 또는 이 단계들의 조합을, O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는 돌연변이 O-RS가 수득될 때까지 반복할 수 있다. 본 발명의 일부 측면에서, 그 단계들은 수회, 예를 들어 2회 이상 수행된다.

<170> 부가적 선택/선별 엄격도 수준도 또한 O-tRNA, O-RS 또는 이들의 쌍을 생성시키기 위한 본 발명의 방법에서 이용될 수 있다. O-RS를 생성시키기 위한 본 방법의 한 단계 또는 두 단계에 대한 선택 또는 선별 엄격도는 달라질 수 있다. 이것은, 예를 들어 사용된 선택제/선별제의 양을 변화시키는 것 등을 포함할 수 있다. 부가적 회수의 양성 및/또는 음성 선택을 또한 수행할 수도 있다. 선택 또는 선별은 아미노산 투과성의 변화, 번역 효율에서의 변화, 번역 신뢰도 변화 등 중 하나 이상을 포함할 수도 있다. 전형적으로, 하나 이상의 변화는 단백질을 생성하기 위해 오르소고날 tRNA-tRNA 합성효소를 사용하는 유기체 내 하나 이상의 유전자 내 돌연변이에 기초한다.

<171> O-RS의 생성 및 합성효소의 기질 특이성의 변경에 대한 부가적 일반 상세 내용을 국제 특허 공개 공보 제 WO2002/086075호(발명의 명칭: "METHODS AND COMPOSITIONS FOR THE PRODUCTION OF ORTHOGONAL tRNA AMINOACYL-tRNA SYNTHETASE PAIRS") 및 제 WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")에서 찾아볼 수 있다. 또한, 문헌 [Wang 및 Schultz, "Expanding the Genetic Code", *Angewandte Chemie Int. Ed.*, 44(1):34-66(2005)]도 또한 참조하고, 이의 내용은 전체적으로 본원에 참조 인용된다.

<172> 공급원 및 숙주 유기체

<173> 본 발명의 오르소고날 번역 성분(O-tRNA 및 O-RS)은, 임의의 다른 종으로부터의 숙주 번역 시스템에서 사용하기 위한 임의의 유기체 (또는 유기체의 조합물)로부터 유도될 수 있는데, 단 O-tRNA/O-RS 성분 및 숙주 시스템은 오르소고날 방식으로 작동한다. 오르소고날 쌍으로부터의 O-tRNA 및 O-RS는 동일한 유기체로부터 유도될 필요는 없다. 일부 측면에서, 오르소고날 성분들 포유동물 숙주 시스템(예, 설치류 또는 영장류 숙주 시스템)에서 사용하기 위한 이.콜라이 유전자(즉, 진균세균)로부터 유도된다.

<174> 예를 들어, 오르소고날 O-tRNA는 고세균 유기체, 예를 들어 원시세균(archaebacterium), 예컨대 메타노코커스 잔나쉬이, 메타노박테리움 썬모토티트로피쿰, 할로박테리움(*Halobacterium*), 예컨대 할로페락스 볼케니이 및 할로박테리움 종 NRC-1, 아카에오글로버스 풀기더스, 파이로코커스 푸리오서스, 파이로코커스 호리코쉬이(*Pyrococcus horikoshii*), 애우로파이럼 페르닉스, 메타노코커스 마리팔루디스, 메타노파이러스 칸들레리, 메타노사르시나 메이제이(Mm), 파이로바쿨럼 애로필럼, 파이로코커스 애비스시, 설폴로버스 솔파타리커스(Ss), 설폴로버스 토코다이이, 썬모플라스마 애시도필럼, 썬모플라스마 볼케이눔 등, 또는 진정세균, 예컨대 에셰리키아 콜라이, 썬머스 썬모필러스, 바실러스 셉틸리스, 바실러스 스테아로썬모필러스 등으로부터 유도될 수 있는 반면, 오르소고날 O-RS는 유기체 또는 유기체들의 조합물, 예를 들어 원시세균, 예컨대 메타노코커스 잔나쉬이, 메타노박테리움 썬모토티트로피쿰, 할로박테리움, 예컨대 할로페락스 볼케니이 및 할로박테리움 종 NRC-1, 아카에오글로버스 풀기더스, 파이로코커스 푸리오서스, 파이로코커스 호리코쉬이, 애우로파이럼 페르닉스, 메타노코커스 마리팔루디스, 메타노파이러스 칸들레리, 메타노사르시나 메이제이, 파이로바쿨럼 애로필럼, 파이로코커스 애비스시, 설폴로버스 솔파타리커스, 설폴로버스 토코다이이, 썬모플라스마 애시도필럼, 썬모플라스마 볼케이눔 등, 또는 진정세균, 예컨대 에셰리키아 콜라이, 썬머스 썬모필러스, 바실러스 셉틸리스, 바실러스 스테아로썬모필러스 등으로부터 유도될 수 있다. 한 구체예에서, 진핵생물 공급원, 예를 들어 식물, 해조류, 원생동물, 진균류, 효모, 동물(예를 들어, 포유동물, 곤충, 절지동물 등) 등도 O-tRNA 및 O-RS의 공급원으로서 사용될 수 있다.

<175> O-tRNA/O-RS 쌍의 개별 성분은 동일한 유기체 또는 상이한 유기체로부터 유도될 수 있다. 한 구체예에서, O-

tRNA/O-RS 쌍은 동일한 유기체로부터 유도된다. 대안으로, O-tRNA/O-RS 쌍의 O-tRNA 및 O-RS는 상이한 유기체로부터 유도된다.

- <176> O-tRNA, O-RS 또는 O-tRNA/O-RS 쌍은 생체내 또는 시험관내에서 선택 또는 선별되고/되거나, 세포 예를 들어, 포유동물 세포, 예컨대 설치류 세포 또는 영장류 세포 내에서 사용되어, 비천연 아미노산을 가진 폴리펩티드를 생성시킬 수 있다. 사용되는 포유동물 숙주 세포는 특별히 한정되지 않는다. 본 발명의 번역 성분들을 포함하는 포유동물 세포의 조성물도 본 발명의 한 특성이다.
- <177> 또한, 또 다른 종에서 사용하기 위해 한 종에서 O-tRNA 및/또는 O-RS를 선별하는 것에 대해, 국제 특허 공개 공보 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")(2004년 4월 16일 출원)를 참조한다.
- <178> (예를 들어, O-RS, O-tRNA 및 비천연 아미노산을 포함하는) 오르소고달 번역 시스템이 배양된 숙주 세포를 사용하여, 비천연 아미노산을 가진 단백질을 생성시킬 수 있다 하더라도, 본 발명의 오르소고달 번역 시스템이 온전한 생존가능한 숙주 세포를 필요로 하는 것으로 의도되지 않는다. 예를 들어, 오르소고달 번역 시스템은 세포 추출물의 존재 하에 무세포 시스템을 사용할 수 있다. 사실상, 단백질의 생성을 위한 무세포 시험관내 전사/번역 시스템의 사용은 잘 확립된 기법이다. 본원에 기재된 오르소고달 번역 시스템 성분들을 사용하여 비천연 아미노산을 가진 단백질을 생성시키기 위한 이들 시험관내 시스템의 적용은 본 발명의 범위 내에 속한다.
- <179> 선택터 코돈
- <180> 본 발명의 선택터 코돈은 단백질 생합성 기구의 유전 코돈 골격을 확장시킨다. 예를 들어, 선택터 코돈은, 예를 들어 독특한 3 염기 코돈, 넌센스 코돈, 예컨대 정지 코돈, 예를 들어 옴버 코돈(UAG) 또는 오팔 코돈(UGA), 비천연 코돈, 4개 이상의 염기로 이루어진 코돈, 회귀 코돈 등을 포함한다. 다수의 선택터 코돈, 예를 들어 하나 이상, 두 개 이상, 세 개 이상 등의 선택터 코돈이 원하는 유전자 내로 도입될 수 있다. 상이한 선택터 코돈을 사용함으로써, 이들 상이한 선택터 코돈을 사용하는 다수의 비천연 아미노산, 예를 들어 하나 이상의 비천연 아미노산의 동시적인 부위-특이적 도입을 허용하는 다수의 오르소고달 tRNA/합성효소 쌍을 사용할 수 있다.
- <181> 한 구체예에서, 본 방법은 생체내 세포에서 비천연 아미노산을 도입하기 위한 정지 코돈인 선택터 코돈의 사용을 수반한다. 예를 들어, 정지 코돈을 인식하며 ORS에 의해 비천연 아미노산으로 아미노아실화되는 O-tRNA가 제조된다. 이 O-tRNA는 천연 발생 숙주의 아미노아실-tRNA 합성효소에 의해 인식되지 않는다. 통상적인 부위-지정 돌연변이유발을 이용하여 관심 폴리펩티드를 코딩하는 폴리뉴클레오타이드 내의 관심 부위에 정지 코돈을 도입할 수 있다. 예를 들어, 문헌 [Sayers, J.R. 등(1988), *5',3' Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis*, *Nucleic Acids Res.*, 791-802]을 참조한다. O-RS, O-tRNA, 및 관심 폴리펩티드를 코딩하는 핵산이 예를 들어, 생체내에서 조합되는 경우, 비천연 아미노산이 정지 코돈에 대한 반응으로 도입됨으로써 특정 부위에서 비천연 아미노산을 함유하는 폴리펩티드를 제공한다. 본 발명의 한 구체예에서, 선택터 코돈으로서 사용된 정지 코돈은 옴버 코돈 UAG 및/또는 오팔 코돈 UGA이다. UAG 및 UGA 양자 모두가 선택터 코돈으로서 사용되는 유전 코드는 가장 풍부한 종결 신호인 오키 넌센스 코돈 UAA를 보존하면서 22종 아미노산을 코딩할 수 있다.
- <182> 생체내 비천연 아미노산의 도입은 숙주 세포에게 유의한 악영향을 주지 않으면서 수행될 수 있다. 예를 들어, 에세리키아 콜라이와 같은 비-진핵세포에서, UAG 코돈에 대한 억제 효율은 O-tRNA, 예를 들어 옴버 억제자 tRNA와 방출 인자 1(RF1)(이것은 UAG 코돈에 결합하고 리보솜으로부터의 성장하는 펩티드의 방출을 개시함) 사이의 경쟁에 달려 있기 때문에, 억제 효율은, 예를 들어 O-tRNA, 예컨대 억제자 tRNA의 발현 수준을 증가시키거나 RF1 결핍 균주를 사용함으로써 조절할 수 있다. 진핵세포에서, UAG 코돈에 대한 억제 효율은 O-tRNA, 예를 들어 옴버 억제자 tRNA와 진핵생물 방출 인자(예를 들어, eRF)(이것은 정지 코돈에 결합하고 리보솜으로부터의 성장하는 펩티드의 방출을 개시함) 사이의 경쟁에 달려 있기 때문에, 억제 효율은, 예를 들어 O-tRNA, 예컨대 억제자 tRNA의 발현 수준을 증가시킴으로써 조절할 수 있다. 또한, 부가적 화합물, 예를 들어 디티오프레이톨(DTT)과 같은 환원제도 존재할 수 있다.
- <183> 비천연 아미노산도 회귀 코돈으로 코딩될 수 있다. 예를 들어, 시험관내 단백질 합성 반응에서 아르기닌 농도가 감소된 경우, 회귀 아르기닌 코돈인 AGG는 알라닌으로 아실화된 합성 tRNA에 의한 Ala의 삽입에 효율적인 것으로 입증되었다. 예를 들어, 문헌 [Ma 등, *Biochemistry*, 32:7939(1993)]을 참조한다. 이 경우, 합성 tRNA는 에세리키아 콜라이에서 소수의 종으로서 존재하는 천연 발생 tRNA^{Arg}와 경쟁한다. 추가로, 몇몇 유기체는 모든 3염기 염기 코돈을 사용하지는 않는다. 마이크로코커스 루테우스(*Micrococcus luteus*)에서 사용되지 않는 코돈인 AGA

는 시험관내 전사/번역 추출물 중에서의 아미노산의 삽입을 위해 사용되고 있다. 예를 들어, 문헌 [Kowal 및 Oliver, *Nucl. Acid. Res.*, 25:4685(1997)]을 참조한다. 본 발명의 성분들은 생체내 이 희귀 코돈들을 사용하여 제조될 수 있다.

<184> 셀렉터 코돈은 또한 확장된 코돈, 예를 들어 4개 이상의 염기로 이루어진 코돈, 예컨대 4개, 5개 또는 6개 이상의 염기로 이루어진 코돈을 포함할 수도 있다. 4 염기 코돈의 예에는, 예를 들어 AGGA, CUAG, UAGA, CCCU 등이 포함된다. 5 염기 코돈의 예에는, 예를 들어 AGGAC, CCCCU, CCCUC, CUAGA, CUACU, UAGGC 등이 포함된다. 본 발명의 방법은 프레임이동 억제에 기초한 확장된 코돈을 사용하는 것을 포함한다. 4개 이상의 염기로 이루어진 코돈은, 예를 들어 하나 또는 다수의 비천연 아미노산을 동일한 단백질 내로 삽입시킬 수 있다. 다른 구체예에서, 안티코돈 루프는, 예를 들어 4개 이상의 염기로 이루어진 코돈, 5개 이상의 염기로 이루어진 코돈 또는 6개 이상의 염기로 이루어진 코돈을 해독할 수 있다. 256개의 가능한 4 염기 코돈이 존재하기 때문에, 4개 이상의 염기로 이루어진 코돈을 사용하여 다수의 비천연 아미노산을 동일한 세포 내에서 코딩할 수 있다. 문헌 [Anderson 등(2002), "Exploring the Limits of Codon and Anticodon Size," *Chemistry and Biology*, 9:237-244]; 및 문헌 [Magliery(2001), "Expanding the Genetic Code: Selection of Efficient Suppressors of Four-base Codons and Identification of "Shifty" Four-base Codons with a Library Approach in *Escherichia coli*," *J. Mol. Biol.*, 307: 755-769]을 참조한다.

<185> 예를 들어, 4 염기 코돈은 시험관내 생합성 방법을 이용하여 비천연 아미노산을 단백질 내로 도입하는 데 사용되어 왔다. 예를 들어, 문헌 [Ma 등(1993), *Biochemistry*, 32:7939]; 및 문헌 [Hohsaka 등(1999), *J. Am. Chem. Soc.*, 121:34]을 참조한다. CGGG 및 AGGU는 화학적으로 아실화된 2개의 프레임이동 억제자 tRNA를 사용하여 시험관내에서 2-나프틸알라닌 및 라이신의 NBD 유도체를 스트렙타비딘 내로 동시에 도입하는 데 사용되었다. 예를 들어, 문헌 [Hohsaka 등(1999), *J. Am. Chem. Soc.*, 121:12194]을 참조한다. 생체내 연구에서, Moore 등은 UAGN 코돈(N은 U, A, G 또는 C일 수 있음)을 억제하는 NCUA 안티코돈을 가진 tRNA^{Leu} 유도체의 능력을 조사하여, 4 염기 코돈인 UAGA가 0 또는 -1 프레임으로 거의 해독되지 않으면서 UCUA 안티코돈을 가진 tRNA^{Leu}에 의해 13 내지 26%의 효율로 해독될 수 있다는 것을 발견하였다. 문헌 [Moore 등(2000), *J. Mol. Biol.*, 298:195]을 참조한다. 한 구체예에서, 다른 원치 않는 부위에서 미스센스(missense) 판독 및 프레임이동 억제를 감소시킬 수 있는, 희귀 코돈 또는 넌센스 코돈에 기초한 확장된 코돈이 본 발명에서 사용될 수 있다. 4 염기 코돈이 다양한 오르소고날 시스템에서 셀렉터 코돈으로서 사용되어 왔다. 예를 들어, 국제 특허 공개 공보 제WO 2005/019415호; 제WO 2005/007870호; 및 제WO 2005/07624호를 참조한다. 또한, 문헌 [Wang 및 Schultz, "Expanding the Genetic Code", *Angewandte Chemie Int. Ed.*, 44(1):34-66(2005)]를 참조하고, 이의 내용은 전체적으로 본원에 참조 인용된다. 하기 실시예는 엠버 셀렉터 코돈을 사용하지만, 다양한 비천연 아미노산 O-RS에 대해 전술된 돌연변이와 유사한 돌연변이를 포함하도록 개질된 4 염기 O-tRNA 및 합성효소를 포함하도록 본원의 실시예를 변형시킴으로써 4개 이상의 염기로 이루어진 코돈도 사용할 수 있다.

<186> 소정의 시스템에 있어서, 셀렉터 코돈은 천연 3 염기 코돈들 중 하나를 포함할 수도 있는데, 여기에서 내인성 시스템은 천연 염기 코돈을 사용하지 (또는 거의 사용하지) 않는다. 예를 들어, 이는 천연 3 염기 코돈을 인식하는 tRNA를 가지지 않은 시스템, 및/또는 3 염기 코돈이 희귀 코돈인 시스템을 포함한다.

<187> 셀렉터 코돈은 임의적으로 비천연 염기 쌍을 포함한다. 그러한 비천연 염기 쌍은 기존 유전자 알파벳을 더 확장시킨다. 한 예분의 염기 쌍은 3 염기 코돈의 수를 64개에서 125개로 증가시킨다. 세 번째 염기 쌍의 성질은 안정한 선택적 염기쌍 형성, 중합효소에 의한 높은 신뢰도의 DNA 내로의 효율적인 효소적 도입, 및 초기(nascent) 비천연 염기 쌍의 합성 후 효율적인 연속적 프라이머 연장을 포함한다. 방법 및 조성물에 적용될 수 있는 비천연 염기 쌍의 설명은, 예를 들어 [Hirao 등(2002) *An unnatural base pair for incorporating amino acid analogues into protein*, *Nature Biotechnology*, 20:177-182]을 포함한다. 또한 문헌 [Wu, Y. 등, (2002) *J. Am. Chem. Soc.*, 124:14626-14630]을 참조한다. 다른 관련 발행물이 하기에 열거되어 있다.

<188> 생체내에서 사용하는 경우, 비천연 뉴클레오타이드는 막 투과성을 나타내며, 인산화되어 상응하는 삼인산염을 형성한다. 또한, 증가된 유전 정보는 안정하고 세포내 효소에 의해 파괴되지 않는다. 베너(Benner) 및 다른 사람들에게 의해 이루어진 과거 노력은, 정규 왓슨-크릭(Watson-Crick) 쌍에서의 수소 결합 패턴과 상이한 수소 결합 패턴을 이용하였는데, 이중 가장 주목할 만한 예는 이소-C:이소-G 쌍이다. 예를 들어, 문헌 [Switzer 등(1989), *J. Am. Chem. Soc.*, 111:8322]; 및 문헌 [Piccirilli 등(1990) *Nature*, 343:33; Kool, (2000) *Curr. Opin. Chem. Biol.*, 4:602]을 참조한다. 일반적으로, 이들 염기는 천연 염기와 어느 정도 잘못된 염기쌍을 형성하며, 효소적으로 복제될 수 없다. 쿨(Kool) 및 그의 동료들은 염기들 사이의 소수성 팩킹(packaging) 상호작용이 수소

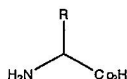
결합을 대체하여 염기 쌍의 형성을 유도할 수 있다는 점을 입증하였다. 문헌 [Kool, (2000) *Curr. Opin. Chem. Biol.*, 4:602]; 및 문헌 [Guckian 및 Kool(1998), *Angew. Chem. Int. Ed. Engl.*, 36, 2825]을 참조할 수 있다. 모든 상기 요건을 만족시키는 비천연 염기 쌍을 개발하려는 노력으로, 슈츠(Schultz), 로메스버그(Romesberg) 및 동료들은 일련의 비천연 소수성 염기들을 체계적으로 합성하여 연구하였다. PICS:PICS 자가-쌍(self-pair)은 천연 염기 쌍보다 안정한 것으로 밝혀졌고, 에세리키아 콜라이 DNA 중합효소 I의 클레노우(Klenow) 단편(KF)에 의해 DNA 내로 효율적으로 도입될 수 있다. 예를 들어, 문헌 [McMinn 등(1999) *J. Am. Chem. Soc.*, 121:11586]; 및 문헌 [Ogawa 등(2000) *J. Am. Chem. Soc.*, 122:3274]을 참조한다. 3MN:3MN 자가-쌍은 생물학적 기능에 충분한 효율 및 선택성을 갖는 KF에 의해 합성될 수 있다. 예를 들어, 문헌 [Ogawa 등(2000) *J. Am. Chem. Soc.*, 122:8803]을 참조한다. 그러나, 염기 양자 모두 부가적 복제를 위한 쇠 종결자로 작용한다. 최근, 돌연변이 DNA 중합효소는 PICS 자가-쌍을 복제하는 데 사용될 수 있는 것으로 밝혀졌다. 또한, 7AI 자가-쌍도 복제될 수 있다. 예를 들어, 문헌 [Tae 등(2001) *J. Am. Chem. Soc.*, 123:7439]을 참조한다. 또한, 새로운 메탈로염기 쌍인 Dipic:Py도 개발되었으며, 이는 염기 쌍은 Cu(II)와의 결합 시 안정한 쌍을 형성한다. 문헌 [Meggers 등(2000) *J. Am. Chem. Soc.*, 122:10714]을 참조할 수 있다. 확장된 코돈 및 비천연 코돈은 본질적으로 천연 코돈에 대해 오르소고날하기 때문에, 본 발명의 방법은 이 성질을 이용하여 이들에 대한 오르소고날 tRNA를 발생시킬 수 있다.

<189> 또한, 번역 우회(translational bypassing) 시스템을 사용하여 관심 폴리펩티드로 비천연 아미노산을 도입할 수 있다. 번역 우회 시스템에서, 큰 서열은 유전자 내로 삽입되지만, 단백질로 번역되지는 않는다. 서열은 리보솜이 서열 상에서 호핑하여(hopping), 삽입 부위의 다운스트림(downstream)에서 번역을 재개하도록 유도하는 신호(cue)로서 작용하는 구조를 포함한다.

<190> 비천연 아미노산

<191> 본원에서 사용되는 비천연 아미노산은 셀레노시스테인 및/또는 피롤라이신, 및 하기 20종의 유전적으로 코딩된 알파-아미노산을 제외한 임의의 아미노산, 변형된 아미노산 또는 아미노산 유사체를 지칭한다: 알라닌, 아르기닌, 아스파라긴, 아스파르트산, 시스테인, 글루타민, 글루탐산, 글리신, 히스티딘, 이소류신, 류신, 라이신, 메티오닌, 페닐알라닌, 프롤린, 세린, 트레오닌, 트립토판, 티로신 및 발린. 알파-아미노산의 일반 구조는 하기 화학식 I로 표시된다:

화학식 I



<192>

<193> 비천연 아미노산은 전형적으로 화학식 I(상기 식 중, R기는 20종의 천연 아미노산에서 사용되는 치환기 이외의 임의의 치환기임)을 갖는 임의의 구조이다. 20종의 천연 아미노산의 구조에 대해서는, 예를 들어 문헌 [Biochemistry, L. Stryer, 제3판, 1988, Freeman and Company, New York]을 참조한다. 본 발명의 비천연 아미노산이 상기 20종의 알파-아미노산 외의 천연 발생 화합물일 수 있다는 것을 주목한다.

<194> 본 발명의 비천연 아미노산은 전형적으로 측쇄에서 천연 아미노산과 상이하기 때문에, 비천연 아미노산은 천연 발생 단백질에서 형성되는 방식과 동일한 방식으로 다른 아미노산, 예를 들어 천연 또는 비천연 아미노산과 아미드 결합을 형성한다. 그러나, 비천연 아미노산은 이들을 천연 아미노산으로부터 구별시켜 주는 측쇄 기를 가진다.

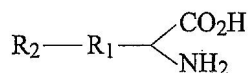
<195> 도 1에 나와 있는 비천연 아미노산이 본원에 기재된 실시예에서 일차적 관심이 있는 것이나, 본 발명은 그 구조에 엄격히 한정되는 것은 아니다. 실제로, 이들로부터 유도된 구조(즉, 도 1에 제시된 구조)의 원리 특성을 보유하고, 또한 본원의 아미노아실-tRNA 합성효소 기준(예를 들어, 서열 번호 57~101의 O-RS)에 의해 특이적으로 인식되는, 각종의 용이하게 유도되고 구조적으로 관련된 유사체도 바로 생성될 수 있다. 이 관련된 아미노산 유사체들도 본 발명의 범주 내에 속하는 것으로 의도된다.

<196> 다른 비천연 아미노산에 있어서, 예를 들어 화학식 I에서의 R은 임의적으로 알킬-, 아릴-, 아실-, 히드라진, 시아노-, 할로-, 히드라지드, 알케닐, 에테르, 붕산염, 보론산염, 포스포, 포스포노, 포스핀, 에논, 이민, 에스테르, 히드록실아민, 아민 등, 또는 이들의 임의의 조합물을 포함한다. 다른 비천연 관심 아미노산은 광활성화가 가능한 가교 결합제를 포함하는 아미노산, 스핀-표지 아미노산, 형광 아미노산, 금속 결합 아미노산, 금속 함유

아미노산, 방사능 아미노산, 새로운 작용기를 갖는 아미노산, 다른 분자와 공유적으로 또는 비공유적으로 상호 작용하는 아미노산, 포토케이징되고/되거나 광이성질체화가능한 아미노산, 바이오틴 또는 바이오틴 유사체를 포함하는 아미노산, 케토 함유 아미노산, 글리코실화된 아미노산, 아미노산 측쇄에 부착된 당류 잔기, 폴리에틸렌 글리콜 또는 폴리에테르를 포함하는 아미노산, 중원자로 치환된 아미노산, 화학적으로 절단될 수 있고/있거나 광절단될 수 있는 아미노산, 천연 아미노산에 비해 연장된 측쇄를 갖는 아미노산(예를 들어, 탄소수 약 5 초과, 약 10 초과 등의 장쇄 탄화수소 또는 폴리에테르), 탄소-결합 당 함유 아미노산, 아미노 티오산 함유 아미노산, 및 하나 이상의 독성 잔기를 포함하는 아미노산을 포함하나 이들로 한정되지 않는다.

<197> 또 다른 측면에서, 본 발명은 하기 화학식 IV로 표시되는 일반 구조를 가진 비천연 아미노산을 제공한다:

화학식 IV

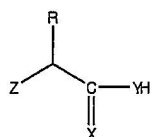


<198>

<199> 이 구조를 갖는 비천연 아미노산은 전형적으로, R₁이 20종의 천연 아미노산 중 하나(예를 들어, 티로신 또는 페닐알라닌)에서 사용된 치환기이고, R₂가 치환기인 임의의 구조이다. 따라서, 이러한 유형의 비천연 아미노산은 천연 아미노산 유도체로서 간주될 수 있다.

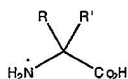
<200> 도 1에 제시된 비천연 아미노산 구조에 부가하여, 비천연 아미노산은 또한 임의적으로 예를 들어, 하기 화학식 II 및 화학식 III으로 표시되는 변형된 골격 구조를 포함할 수도 있다:

화학식 II



<201>

화학식 III



<202>

<203> (상기 식에서, Z는 전형적으로 OH, NH₂, SH, NH-R' 또는 S-R'를 포함하고; X 및 Y는 동일하거나 상이할 수 있으며 전형적으로 S 또는 O를 포함하고, R 및 R'는 임의적으로 동일하거나 상이할 수 있으며 전형적으로 화학식 I 을 갖는 비천연 아미노산에 대해 전술한 R 기에 대한 치환기들의 동일한 목록 및 수소로부터 선택됨). 예를 들어, 본 발명의 비천연 아미노산은 임의적으로 화학식 II 및 화학식 III으로 표시되는 아미노 또는 카르복실 기에서의 치환을 포함한다. 그러한 종류의 비천연 아미노산은, 예컨대 20종의 공통된 천연 아미노산에 상응하는 측쇄 또는 비천연 측쇄를 갖는 α-히드록시산, α-티오산, α-아미노티오카르복실산염을 포함하나 이들로 한정되지 않는다. 또한, α-탄소에서의 치환은 임의적으로 L, D 또는 α-α-이치환 아미노산, 예를 들어 D-글루탐산염, D-알라닌, D-메틸-O-티로신, 아미노부티르산 등을 포함한다. 다른 구조적 대체물은 시클릭 아미노산, 예를 들어 프롤린 유사체 뿐만 아니라, 3, 4, 6, 7, 8 및 9원 고리 프롤린 유사체, β 및 γ 아미노산, 예컨대 치환된 β-알라닌 및 γ-아미노 부티르산을 포함한다.

<204> 일부 측면에서, 본 발명은 L-구조형태의 비천연 아미노산을 이용한다. 그러나, 본 발명이 L-구조형태의 비천연 아미노산의 사용으로 한정되는 것으로 의도되지 않는다. 이 비천연 아미노산들의 D-거울상이성질체도 본 발명에서 사용될 수 있음이 구상된다.

<205> 본 발명에서 사용될 수 있는 비천연 아미노산은 도 1에 나와 있는 비천연 아미노산에 엄격히 한정되지 않는다. 당업자라면 천연 발생 아미노산의 매우 각종의 비천연 유사체들이 용이하게 유도됨을 인식할 것이다. 예를 들어, 단 비제한적 예로서 페닐알라닌 또는 티로신으로부터 유도된 비천연 아미노산이 용이하게 생성된다. 티로

신 유사체에는, 예를 들어 파라-치환 티로신, 오르토-치환 티로신 및 메타-치환 티로신이 포함되고, 여기에서 치환 티로신은 알킬닐 기, 아세틸 기, 벤조일 기, 아미노 기, 히드라진, 히드록시아민, 티올 기, 카르복시 기, 이소프로필 기, 메틸 기, C₆-C₂₀ 직쇄 또는 분지쇄 탄화수소, 포화 또는 불포화 탄화수소, O-메틸 기, 폴리에테르 기, 니트로 기 등을 포함한다. 또한, 다-치환 아릴 고리도 구상된다. 본 발명의 글루타민 유사체는 α-히드록시 유도체, γ-치환 유도체, 시클릭 유도체 및 아마이드 치환 글루타민 유도체를 포함하나 이들로 한정되지 않는다. 예시적인 페닐알라닌 유사체는 파라-치환 페닐알라닌, 오르토-치환 페닐알라닌, 및 메타-치환 페닐알라닌을 포함하나, 이들로 한정되지 않으며, 여기에서 치환기는 알킬닐 기, 히드록시 기, 메톡시 기, 메틸 기, 알릴 기, 알데히드, 니트로, 티올 기 또는 케토 기 등 등을 포함한다. 비천연 아미노산의 구체예에는 설포티로신, p-에틸티오카르보닐-L-페닐알라닌, p-(3-옥소부타노일)-L-페닐알라닌, 1,5-단일-알라닌, 7-아미노-쿠마린 아미노산, 7-히드록시-쿠마린 아미노산, 니트로벤질-세린, O-(2-니트로벤질)-L-티로신, p-카르복시메틸-L-페닐알라닌, p-시아노-L-페닐알라닌, m-시아노-L-페닐알라닌, 비페닐알라닌, 3-아미노-L-티로신, 비페리딜 알라닌, p-(2-아미노-1-히드록시에틸)-L-페닐알라닌, p-이소프로필티오카르보닐-L-페닐알라닌, 3-니트로-L-티로신 및 p-니트로-L-페닐알라닌이 포함된다. 또한, p-프로파르길옥시페닐알라닌, 3,4-디히드록시-L-페닐알라닌(DHP), 3,4,6-트리히드록시-L-페닐알라닌, 3,4,5-트리히드록시-L-페닐알라닌, 4-니트로-페닐알라닌, p-아세틸-L-페닐알라닌, O-메틸-L-티로신, L-3-(2-나프틸)알라닌, 3-메틸-페닐알라닌, O-4-알릴-L-티로신, 4-프로필-L-티로신, 3-니트로-티로신, 3-티올-티로신, 트리-O-아세틸-GlcNAc β-세린, L-도파(L-Dopa), 불화 페닐알라닌, 이소프로필-L-페닐알라닌, p-아지도-L-페닐알라닌, p-아실-L-페닐알라닌, p-벤조일-L-페닐알라닌, L-포스포세린, 포스포노세린, 포스포노티로신, p-요오도-페닐알라닌, p-브로모페닐알라닌, p-아미노-L-페닐알라닌, 이소프로필-L-페닐알라닌 등이 포함된다. 각종 비천연 아미노산의 구조가 본원에 인용된 참조문헌에 개시되어 있다, 또한 공개된 국제 출원 제 WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE"), 및 국제 출원 PCT/US2005/039210호(발명의 명칭: "ORTHOGONAL TRANSLATION COMPONENTS FOR THE IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")(2005년 10월 27일 출원)를 참조한다.

<206> 비천연 아미노산의 화학적 합성

<207> 전술한 비천연 아미노산의 대다수는, 예를 들어 시그마(Sigma)(미국 소재) 또는 알드리히(Aldrich)(미국 위스콘신주 밀워키 소재)로부터 상업적으로 입수가 가능하다. 상업적으로 입수가 가능하지 않은 비천연 아미노산은 임의적으로 다양한 문헌에 기재된 바와 같이 합성되거나 당업자에게 공지된 표준 방법을 사용하여 합성된다. 유기 합성 기법에 대해, 예를 들어 문헌 [*Organic Chemistry*, Fessenden and Fessenden, (1982, 제2판, Willard Grant Press, Boston Mass.)]; 문헌 [*Advanced Organic Chemistry*, March(제3판, 1985, Wiley and Sons, New York)]; 및 문헌 [*Advanced Organic Chemistry*, Carey 및 Sundberg(제3판, A 및 B 파트, 1990, Plenum Press, New York)]을 참조한다. 비천연 아미노산의 합성을 기술하는 부가적 문헌에는, 예를 들어 국제 특허 공개 공보 제 WO 2002/085923호(발명의 명칭: "In vivo incorporation of Unnatural Amino Acids"), 문헌 [Matsoukas 등 (1995) *J. Med. Chem.*, 38, 4660-4669]; 문헌 [King 및 Kidd(1949), A New Synthesis of 글루타민 and of γ-Dipeptides of Glutamic Acid from Phthylated Internediate, *J. Chem. Soc.*, 3315-3319]; 문헌 [Friedman 및 Chatterrji(1959), Synthesis of Derivatives of 글루타민 as Model Substrates for Anti-Tumor Agents, *J. Am. Chem. Soc.*, 81, 3750-3752]; 문헌 [Craig 등(1988), Absolute Configuration of the Enantiomers of 7-Chloro-4 [[4-(diethylamino)-1-methylbutyl]amino]quinoline(Chloroquine), *J. Org. Chem.*, 53, 1167-1170]; 문헌 [Azoulay 등(1991), 글루타민 analogues as Potential Antimalarials, *Eur. J. Med. Chem.*, 26, 201-5]; 문헌 [Koskinen 및 Rapoport(1989), Synthesis of 4-Substituted Prolines as Conformationally Constrained Amino Acid Analogues, *J. Org. Chem.*, 54, 1859-1866]; 문헌 [Christie 및 Rapoport(1985), Synthesis of Optically Pure Pípecolates from L-Asparagine. Application to the Total Synthesis of (+)-Apovincamine through Amino Acid Decarbonylation and Iminium Ion Cyclization, *J. Org. Chem.*, 1989:1859-1866]; 문헌 [Barton 등(1987), Synthesis of Novel α-Amino-Acids and Derivatives Using Radical Chemistry: Synthesis of L- and D-α-Amino-Adipic Acids, L-α-aminopinelic Acid and Appropriate Unsaturated Derivatives, *Tetrahedron Lett.*, 43:4297-4308]; 및 문헌 [Subasinghe 등(1992), Quisqualic acid analogues: synthesis of betaheterocyclic 2-aminopropanoic acid derivatives and their activity at a novel quisqualate-sensitized site, *J. Med. Chem.*, 35: 4602-7]을 포함한다. 또한, 국제 특허 공개 공보 제 WO 2004/058946호(발명의 명칭: "PROTEIN ARRAYS", 2003년 12월 22일 출원)를 참조한다.

<208> 비천연 아미노산의 세포내 흡수

<209> 세포에 의한 비천연 아미노산 흡수는, 예를 들어 단백질 내로 도입하기 위해 비천연 아미노산을 설계하고 선택

할 때, 전형적으로 고려되는 한 사항이다. 예를 들어, α -아미노산의 높은 전하 밀도는 이 화합물이 세포를 투과하기 어렵다는 것을 제시한다. 천연 아미노산은 종종 다양한 아미노산 특이성 수준을 나타내는 일군의 단백질-기계 수송 시스템을 통해 세포 내로 흡수된다. 비천연 아미노산이 세포에 의해 흡수된다면 어떤 비천연 아미노산이 세포에 의해 흡수되는지를 평가하는 신속 선별을 수행할 수 있다. 예를 들어, 국제 특허 공개 공보 제WO 2004/058946호(발명의 명칭: "PROTEIN ARRAYS")(2003년 12월 22일 출원); 및 문헌 [Liu 및 Schultz(1999), *Progress toward the evolution of an organism with an expanded genetic code*, PNAS 96: 4780-4785]에 기재된 독성 분석을 참조한다. 흡수는 다양한 검정으로 용이하게 분석되지만, 세포내 흡수 경로를 따라 흡수될 수 있는 비천연 아미노산의 설계에 대한 대안은 생체내에서 아미노산을 생성시키는 생합성 경로를 제공하는 것이다.

<210> 비천연 아미노산의 생합성

<211> 아미노산 및 다른 화합물들을 제조하기 위한 많은 생합성 경로들이 세포에 이미 존재한다. 특별한 비천연 아미노산에 대한 생합성 방법이 자연계, 예를 들어 세포에 존재하지 않을 수 있으나, 본 발명은 그러한 방법을 제공한다. 예를 들어, 비천연 아미노산에 대한 생합성 경로는 임의적으로 새로운 효소를 첨가하거나 기존 숙주 세포 경로를 변형시킴으로써 숙주 세포에서 발생된다. 새로운 부가적 효소로는 임의적으로 천연 발생 효소 또는 인공적으로 유도된 효소가 있다. 예를 들어, (제WO 2002/085923호의 실시예에서 제시된 바와 같은) p-아미노페닐알라닌의 생합성은 다른 유기체로부터의 공지된 효소의 조합물의 첨가에 의존한다. 이 효소들에 대한 유전자는 유전자를 포함하는 플라스미드를 사용하여 세포를 형질전환시킴으로써 세포 내로 도입될 수 있다. 그 유전자는 세포에서 발현될 때, 원하는 화합물을 합성하는 효소적 경로를 제공한다. 임의적으로 첨가되는 효소 종류의 예는 하기 실시예에서 제공된다. 부가적 효소의 서열은, 예를 들어 유전자은행(Genbank)에서 찾아볼 수 있다. 또한, 인공적으로 유도된 효소는 임의적으로 동일한 방식으로 세포 내로 첨가된다. 이 방식으로, 세포내 기구 및 세포내 자원을 조작하여, 비천연 아미노산을 생성시킨다.

<212> 사실상, 시험관내 또는 생체내에서 생합성 경로, 기존 경로의 개발 또는 비천연 아미노산의 제조에 사용하기 위한 신규 효소를 제조하는 다양한 방법들 중 임의의 방법이 이용될 수 있다. 효소 및 다른 생합성 경로 성분들을 개발하기 위한 많은 이용가능한 방법들을 본 발명의 방법에 적용하여, 비천연 아미노산을 생성시킬 수 있다(또는, 사실상 새로운 기질 특이성 또는 기타 관심 활성을 가진 합성효소를 개발할 수 있다). 예를 들어, 경우에 따라 DNA 서플링을 이용하여 시험관내 또는 생체내에서 비천연 아미노산의 생성 (또는 신규 합성효소의 생성)을 위한 신규 효소 및/또는 이 효소의 경로를 개발한다. 예를 들어, 문헌 [Stemmer(1994), *Rapid evolution of a protein in vitro by DNA shuffling*, *Nature* 370(4):389-391]; 및 문헌 [Stemmer(1994), *DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution*, *Proc. Natl. Acad. Sci. USA.*, 91:10747-10751]을 참조한다. 관련 방법들은 관련(예를 들어, 상동성) 유전자 패밀리(family)를 서플링하여 원하는 특징을 가진 효소들을 신속히 개발한다. 그러한 "패밀리 유전자 서플링" 방법의 예를 문헌 [Cramer 등(1998), *DNA shuffling of a family of genes from diverse species accelerates directed evolution*, *Nature*, 391(6664): 288-291]에서 찾아볼 수 있다. (생합성 경로 성분 또는 합성효소 여부를 불문한) 신규 효소는, 예를 들어 문헌 [Ostermeier 등(1999), "A combinatorial approach to hybrid enzymes independent of DNA homology", *Nature Biotech* 17: 1205]에 기재된 바와 같이 "혼성 효소의 생성을 위한 점진적 절두(truncation)"("ITCHY")로서 알려진 DNA 재조합 절차를 이용하여 발생시킬 수도 있다. 이 방법은 하나 이상의 시험관내 또는 생체내 재조합 방법에 대한 기질로서 작용할 수 있는 효소 또는 다른 경로 변이체 라이브러리를 발생시키는 데 이용될 수도 있다. 예를 들어, 문헌 [Ostermeier 등(1999), "Combinatorial Protein Engineering by Incremental Truncation", *Proc. Natl. Acad. Sci. USA*, 96: 3562-67], 및 문헌 [Ostermeier 등(1999), "Incremental Truncation as a Strategy in the Engineering of Novel Biocatalysts", *Biological and Medicinal Chemistry*, 7: 2139-44]도 또한 참조한다. 또 다른 방법은, 예를 들어 비천연 아미노산 (또는 신규 합성효소)의 제조에 대한 생합성 반응을 촉진하는 능력에 대해 선택된 효소 또는 다른 경로 변이체 라이브러리를 제조하기 위해 기하급수적이고 동시적인 돌연변이유발을 이용한다. 이 방법에서, 관심 서열 내의 잔기들로 이루어진 작은 군을 동시에 무작위화하여 각각의 변경된 위치에서 기능성 단백질을 생성시키는 아미노산들을 동정한다. 비천연 아미노산(또는 신규 합성효소)의 생성을 위한 신규 효소를 제조하기 위해 본 발명에 적용될 수 있는 그러한 절차의 예를 문헌 [Delegrave 및 Youvan(1993), *Biotechnology Research* 11:1548-1552]에서 찾아볼 수 있다. 또 다른 방법에서, 효소 및/또는 경로 성분 공학처리를 위한 도핑된(doped) 또는 축퇴성 올리고뉴클레오타이드를 사용하는 랜덤 또는 세미-랜덤 돌연변이유발, 예를 들어 문헌 [Arkin 및 Youvan(1992), "Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis", *Biotechnology* 10:297-300]; 또는 문헌 [Reidhaar-Olson 등(1991), "Random mutagenesis of

protein sequences using oligonucleotide cassettes", *Methods Enzymol.* 208:564-86]의 일반적인 돌연변이유발법의 사용에 의해 적용될 수 있다. 폴리뉴클레오티드 재조립(reassembly) 및 부위-포화 돌연변이유발을 이용하는, 종종 "비-확률적" 돌연변이유발로 지칭되는 또 다른 방법을 이용하여, 효소 및/또는 경로 성분들을 생성시킬 수 있고, 이어서 상기 효소 및/또는 경로 성분들은 (예를 들어, 생체내에서 비천연 아미노산의 생성을 위한) 하나 이상의 합성효소 또는 생합성 경로 기능을 수행하는 능력에 대해 선별될 수 있다. 예를 들어, Short의 제WO 00/46344호(발명의 명칭: "NON-STOCHASTIC GENERATION OF GENETIC VACCINES AND ENZYMES")를 참조한다.

<213> 그러한 돌연변이유발법에 대한 대안은 유기체의 전체 게놈을 재조합하고, 생성된 자손(progeny)을 특정한 경로 기능(종종 "전체 게놈 셔플링"으로 지칭됨)에 대한 선택하는 단계를 포함한다. 이 방법은, 예를 들어 비천연 아미노산(또는 이의 중간체)를 생성시키는 능력에 대한 유기체(예를 들어, 이. 콜라이 또는 다른 세포)의 게놈 재조합 및 선택에 의해 본 발명에 적용될 수 있다. 예를 들어, 문헌 [Patnaik 등(2002), "Genome shuffling of lactobacillus for improved acid tolerance" *Nature Biotechnology* 20(7): 707-712]; 및 문헌 [Zhang 등(2002), "Genome shuffling leads to rapid phenotypic improvement in bacteria" *Nature*, February 7, 415(6872): 644-646]에 교시된 방법은 생체내에서 비천연 아미노산을 생성시키기 위해 세포내의 기존 및/또는 신규 경로를 개발하기 위한 경로 설계에 적용될 수 있다.

<214> 예를 들어 원하는 화합물의 제조를 위한 유기체 및 대사 경로 공학처리를 위한 다른 기법도 이용가능하며, 비천연 아미노산의 생성에 적용될 수도 있다. 유용한 경로 공학처리 방법을 교시하는 발행물의 예에는, 문헌 [Nakamura 및 White(2003), "Metabolic engineering for the microbial production of 1,3-propanediol" *Curr. Opin. Biotechnol.* 14(5):454-9; Berry 등(2002) "Application of Metabolic Engineering to improve both the production and use of Biotech Indigo" *J. Industrial Microbiology and Biotechnology* 28:127-133]; 문헌 [Banta 등(2002) "Optimizing an artificial metabolic pathway: Engineering the cofactor specificity of Corynebacterium 2,5-diketo-D-gluconic acid reductase for use in vitamin C biosynthesis" *Biochemistry.* 41(20), 6226-36]; 문헌 [Selivonova 등(2001), "Rapid Evolution of Novel Traits in Microorganisms" *Applied and Environmental Microbiology*, 67:3645)], 및 많은 다른 문헌들이 포함된다.

<215> 사용 방법과 무관하게, 전형적으로 본 발명의 공학처리된 생합성 경로를 이용하여 생성된 비천연 아미노산은 효율적인 단백질 생합성에 충분한 농도, 예를 들어 세포내 본래 양으로 제조되나, 다른 세포내 아미노산의 농도에 유의한 영향을 미치거나 세포내 자원들을 고갈시킬 정도로 제조되는 것은 아니다. 이 방식으로 생체내에서 제조된 전형적인 농도는 약 10 mM 내지 약 0.05 mM이다. 일단 세포가 특별한 경로에 필요한 효소를 제조하도록 공학처리되고 비천연 아미노산이 생성되면, 임의적으로 생체내 선택을 이용하여, 리보솜에 의한 단백질 합성 및 세포 성장 양자 모두를 위한 비천연 아미노산의 제조를 더 최적화시킨다.

<216> 비천연 아미노산의 도입을 위한 오르소고날 성분

<217> 본 발명은, 예를 들어 생체내에서 셀렉터 코돈, 예를 들어 엠버 정지 코돈, 넌센스 코돈, 4개 이상의 염기로 이루어진 코돈 등에 대한 반응으로 도 1에 도시된 비천연 아미노산을 성장하는 폴리펩티드 쇠대로 도입하기 위한 오르소고날 성분의 생성을 위한 조성물 및 방법을 제공한다. 예를 들어, 본 발명은 오르소고날-tRNA (O-tRNA), 오르소고날 아미노아실-tRNA 합성효소(O-RS) 및 이들의 쌍을 제공한다. 이 쌍은 비천연 아미노산을 성장하는 폴리펩티드 쇠대로 도입하는 데 사용될 수 있다.

<218> 본 발명의 조성물은 오르소고날 아미노아실-tRNA 합성효소(O-RS)를 포함하는데, 여기에서 O-RS는 O-tRNA를 도 1의 비천연 아미노산으로 우선적으로 아미노아실화한다. 특정 구체예에서, O-RS는 서열 번호 57~101에서 선택된 아미노산 서열, 이의 보존적 변이체를 포함한다. 본 발명의 특정 구체예에서, O-RS는 임의의 내인성 tRNA에 비해 O-tRNA를 특별한 비천연 아미노산으로 우선적으로 아미노아실화하고, 여기에서 O-RS는 O-tRNA에 대해 선호적이고, 비천연 아미노산으로 충전된 O-tRNA:동일한 비천연 아미노산으로 충전된 내인성 tRNA의 비가 1:1 초과이고, 더욱 바람직하게는 O-RS가 O-tRNA만을 충전하거나, 거의 O-tRNA만을 충전한다.

<219> O-RS를 포함하는 조성물은 오르소고날 tRNA(O-tRNA)을 임의적으로 추가로 포함할 수 있는데, 여기에서 O-tRNA는 셀렉터 코돈을 인식한다. 전형적으로, 본 발명의 O-tRNA는 서열목록에 나와 있는 바와 같은 폴리뉴클레오티드 서열(예, 서열 번호 3)을 포함하거나 그 서열에 의해 코딩되는 O-tRNA의 억제 효율에 비해 셀렉터 코돈에 대한 반응으로 동족체 합성효소의 존재 하에 적어도 대략, 예를 들어 45%, 50%, 60%, 75%, 80% 또는 90% 이상의 억제 효율을 포함한다. 한 구체예에서, O-RS 및 O-tRNA가 함께 있는 경우의 억제 효율은 O-RS가 없는 경우, O-tRNA의 억제 효율보다 예를 들어, 5배, 10배, 15배, 20배 또는 25배 이상 더 높다. 일부 측면에서, O-RS 및 O-tRNA가 함께 있는 경우의 억제 효율은 이.콜라이로부터 유도된 오르소고날 티로실-tRNA 합성효소의 억제 효율의 45% 이

상이다.

- <220> O-tRNA를 포함하는 조성물은 임의적으로 숙주 세포(예를 들어, 포유동물 세포, 예컨대 설치류 세포 또는 영장류 세포), 및/또는 번역 시스템을 포함할 수 있다.
- <221> 번역 시스템을 포함하는 세포(예, 설치류 세포 또는 영장류 세포)도 또한 본 발명에 의해 제공되는데, 여기에서 번역 시스템은 오르소고날-tRNA(O-tRNA); 오르소고날 아미노아실-tRNA 합성효소(O-RS); 및 도 1에 도시된 비천연 아미노산을 포함한다. 전형적으로, O-RS는 임의의 내인성 tRNA에 비해, O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는데, 여기에서 O-RS는 tRNA에 대해 선호적이고, 비천연 아미노산으로 충전된 O-tRNA:동일한 비천연 아미노산으로 충전된 내인성 tRNA의 비가 1:1 초과이고, 더욱 바람직하게는 O-RS가 O-tRNA만을 충전하거나, 거의 O-tRNA만을 충전한다. O-tRNA는 제1 선택터 코돈을 인식하고, O-RS는 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화한다. 한 구체예에서, O-tRNA는 서열 번호 3에 나와 있는 폴리뉴클레오티드 서열 또는 이의 상보적 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩된다. 한 구체예에서, O-RS는 서열 번호 57-101에 나와 있는 아미노산 서열, 또는 이의 보존적 변이체를 포함한다.
- <222> 본 발명의 세포는 부가적 상이한 O-tRNA/O-RS 쌍 및 제2 비천연 아미노산을 임의적으로 추가로 포함할 수 있는데, 여기에서 이 O-tRNA는 제2 선택터 코돈을 인식하고, 이 O-RS는 상응하는 O-tRNA를 제2 비천연 아미노산으로 우선적으로 아미노아실화하며, 여기에서 제2 아미노산은 제1 비천연 아미노산과 상이하다. 임의적으로, 본 발명의 세포는 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함한 핵산을 포함하는데, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 선택터 코돈을 포함한다.
- <223> 특정 구체예에서, 본 발명의 세포는, 오르소고날-tRNA(O-tRNA), 오르소고날 아미노아실-tRNA 합성효소(O-RS), 비천연 아미노산, 및 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함한 핵산을 포함하는 포유동물 세포(예, 설치류 세포 또는 영장류 세포)인데, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 선택터 코돈을 포함한다. 본 발명의 특정 구체예에서, O-RS는, O-RS가 임의의 내인성 tRNA를 아미노아실화하는 효율을 초과하는 효율로, O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화한다.
- <224> 본 발명의 특정 구체예에서, 본 발명의 O-tRNA는 본원의 서열목록(예, 서열 번호 3) 또는 실시예에 나와 있는 바와 같은 폴리뉴클레오티드 서열, 또는 이의 상보적인 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩된다. 본 발명의 특정 구체예에서, O-RS는 서열목록에 나와 있는 아미노산 서열 또는 이의 보존적 변이체를 포함한다. 한 구체예에서, O-RS 또는 이의 일부는 본원의 서열목록 또는 실시예에 나와 있는 폴리뉴클레오티드 서열, 또는 이의 상보적 폴리뉴클레오티드 서열에 의해 코딩된다.
- <225> 본 발명의 O-tRNA 및/또는 O-RS는 각종 유기체(예, 진핵생물 및/또는 진핵생물의 유기체) 중 임의의 유기체로부터 유도될 수 있다.
- <226> 폴리뉴클레오티드도 또한 본 발명의 한 특성이다. 본 발명의 폴리뉴클레오티드(예를 들어, 서열 번호 8~56)는 본원의 서열목록에 나와 있는 폴리펩티드를 코딩하는 뉴클레오티드 서열을 포함하는 인공(예를 들어, 수동-제조되고, 천연적으로 발생되지 않은) 폴리뉴클레오티드를 포함하고/하거나, 상기 서열에 대해 상보적이다. 본 발명의 폴리뉴클레오티드는 또한 고도로 엄격한 조건 하에 실질적으로 핵산의 전체 길이에 걸쳐 상기 폴리뉴클레오티드에 혼성화되는 핵산을 포함할 수도 있다. 본 발명의 폴리뉴클레오티드는 또한 천연 발생 tRNA 또는 이에 상응하는 코딩 핵산의 서열과 예를 들어 75% 이상, 80% 이상, 90% 이상, 95% 이상, 98% 이상, 또는 그 이상으로 동일한 (단, 천연 발생 tRNA 또는 이에 상응하는 코딩 핵산이 아닌) 폴리뉴클레오티드를 포함하는데, 여기에서 tRNA는 선택터 코돈, 예를 들어 4 염기 코돈을 인식한다. 상기 서열 및/또는 상기 서열의 보존적 변이체를 포함하는 폴리뉴클레오티드 중 임의의 것과 예를 들어 80% 이상, 90% 이상, 95% 이상, 98% 이상 또는 그 이상 동일한 인공 폴리뉴클레오티드도 또한 본 발명의 폴리뉴클레오티드 내에 포함된다.
- <227> 본 발명의 폴리뉴클레오티드를 포함하는 벡터도 또한 본 발명의 한 특성이다. 예를 들어, 본 발명의 벡터는 플라스미드, 코스미드, 파지, 바이러스, 발현 벡터 등을 포함할 수 있다. 본 발명의 벡터를 포함한 세포도 또한 본 발명의 한 특성이다.
- <228> O-tRNA/O-RS 쌍의 생성 방법도 또한 본 발명의 특성이다. 이 방법에 의해 생성된 성분도 또한 본 발명의 특성이다. 예를 들어, 세포에 대해 오르소고날성을 갖는 하나 이상의 tRNA(O-tRNA)의 생성 방법은, 돌연변이체 tRNA의 라이브러리를 발생시키는 단계; 돌연변이체 tRNA의 라이브러리의 각 구성원의 안티코돈 루프를 돌연변이화하여 선택터 코돈을 인식하도록 함으로써 잠재적 O-tRNA의 라이브러리를 제공하는 단계; 및 제1 종의 세포의 제1 집단을 음성 선택하는 단계를 포함하고, 여기에서 세포는 잠재적 O-tRNA의 라이브러리의 한 구성원을 포함

한다. 음성 선택은 세포에 대해 내인성인 아미노아실-tRNA 합성효소(RS)에 의해 아미노아실화되는 잠재적 O-tRNA의 라이브러리의 구성원을 포함하는 세포를 제거한다. 이는 제1 종의 세포에 대해 오르소고날성을 갖는 tRNA의 풀을 제공함으로써, 하나 이상의 O-tRNA를 제공한다. 본 발명의 방법에 의해 생성되는 O-tRNA도 또한 제공된다.

<229> 특정 구체예에서, 그 방법은 제1 종의 세포의 제2 집단을 양성 선택하는 것을 추가로 포함하는데, 여기에서 세포는 제1 종의 세포에 대해 오르소고날성인 tRNA의 풀의 구성원, 동족체 아미노아실-tRNA 합성효소 및 양성 선택 마커를 포함한다. 양성 선택을 이용하여, 세포를 동족체 아미노아실-tRNA 합성효소에 의해 아미노아실화되고 양성 선택 마커의 존재 하에 원하는 반응을 나타내는 tRNA의 풀의 구성원을 포함하는 상기 세포에 대해 선택 또는 선별함으로써, O-tRNA를 제공한다. 특정 구체예에서, 세포의 제2 집단은 음성 선택에 의해 제거되지 않은 세포를 포함한다.

<230> O-tRNA를 비천연 아미노산으로 충전하는 오르소고날-아미노아실-tRNA 합성효소의 동정 방법도 또한 공지되어 있다. 예를 들어, 그 방법은 제1 종의 세포의 집단을 선택하는 단계를 포함하는데, 여기에서 세포는 각기 1) 복수개의 아미노아실-tRNA 합성효소(RS)의 구성원(예를 들어, 복수개의 RS에는 돌연변이체 RS, 상기 제1 종 외의 종으로부터 유도된 RS, 또는 돌연변이체 RS 및 상기 제1 종 외의 종으로부터 유도된 RS 양자 모두가 포함될 수 있음); 2) (예를 들어, 하나 이상의 종으로부터의) 오르소고날 tRNA(O-tRNA); 및 3) 양성 선택 마커를 코딩하며 하나 이상의 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 포함한다.

<231> 세포(예, 숙주 세포)는, 복수개의 RS 구성원을 가지지 않거나 감소된 양으로 가지는 세포에 비해 억제 효율의 증진을 나타내는 세포에 대해 선택 또는 선별된다. 이 선택/선별된 세포는 O-tRNA를 아미노아실화하는 활성 RS를 포함한다. 그 방법에 의해 동정된 오르소고날 아미노아실-tRNA 합성효소도 또한 본 발명의 한 특성이다.

<232> 선택된 위치에 비천연 아미노산을 갖는 숙주 세포(예를 들어, 포유동물 세포, 예컨대 설치류 세포 또는 영장류 세포 등) 내 단백질의 생성 방법도 또한 본 발명의 한 특성이다. 예를 들어, 그 방법은 하나 이상의 셀렉터 코돈을 포함하면서 단백질을 코딩하는 핵산을 포함하는 세포를 적절한 배지에서 성장시키는 단계, 비천연 아미노산을 제공하는 단계, 및 하나 이상의 셀렉터 코돈으로 핵산의 번역 과정 중에 단백질 내 특정된 위치에 비천연 아미노산을 도입함으로써 단백질을 생성시키는 단계를 포함한다. 세포는, 세포 내에서 기능하면서 셀렉터 코돈을 인식하는 오르소고날-tRNA(O-tRNA); 및 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는 오르소고날 아미노아실-tRNA 합성효소(O-RS)를 추가로 포함한다. 이 방법에 의해 생성된 단백질도 또한 본 발명의 한 특성이다.

<233> 본 발명은 또한 비천연 아미노산을 포함한 단백질을 포함하는 조성물을 제공한다. 특정 구체예에서, 단백질은 공지된 단백질, 예를 들어 치료 단백질, 진단 단백질, 산업용 효소, 또는 이들의 일부의 아미노산 서열과 75% 이상 동일한 아미노산 서열을 포함한다. 임의적으로, 조성물은 약학적으로 허용가능한 담체를 포함한다.

<234> 핵산 및 폴리펩티드 서열 및 변이체

<235> 본원에 기재된 바와 같이, 본 발명은 예를 들어 O-tRNA 및 O-RS를 코딩하는 폴리뉴클레오티드 서열, 및 폴리펩티드 아미노산 서열, 예를 들어 O-RS, 및 예를 들어 상기 폴리뉴클레오티드 및 폴리펩티드 서열을 포함하는 조성물, 시스템 및 방법을 제공한다. 상기 서열, 예를 들어 O-tRNA 및 O-RS 아미노산 및 뉴클레오티드 서열의 예가 본원에 개시되어 있다(도 16, 예컨대 서열 번호 3 및 8~101 참조). 그러나, 당업자라면 본 발명이 본원에, 예를 들어 실시예 및 서열목록에 구체적으로 언급된 서열로 한정되지 않는다는 것을 인식할 것이다. 당업자라면 본 발명이 또한 본원에 기재된 기능을 가진 많은 관련 서열, 예를 들어 본원에 개시된 O-RS의 보존적 변이체를 코딩하는 폴리뉴클레오티드 및 폴리펩티드를 제공한다는 것을 인식할 것이다. 효모 숙주 시스템에서 O-tRNA를 비천연 아미노산으로 아미노아실화시킬 수 있는 오르소고날 합성효소 종(O-RS)의 구축 및 분석에 대한 일반적인 방법론은 당업계에 공지되어 있다.

<236> 본 발명은 폴리펩티드(O-RS) 및 폴리뉴클레오티드, 예를 들어 O-tRNA, O-RS 또는 이의 부분을 코딩하는 폴리뉴클레오티드, 아미노아실-tRNA 합성효소 클론을 단리하기 위해 사용되는 올리고뉴클레오티드 등을 제공한다. 본 발명의 폴리뉴클레오티드는 하나 이상의 셀렉터 코돈을 이용하여 본 발명의 관심 단백질 또는 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함한다. 또한, 본 발명의 폴리뉴클레오티드에는 예를 들어 서열 번호 8~56에 나와 있는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드, 및 그것의 폴리뉴클레오티드 서열에 상보적이거나 그 서열을 코딩하는 폴리뉴클레오티드가 포함된다. 본 발명의 폴리뉴클레오티드에는 또한 서열 번호 57~101을 포함하는 O-RS 아미노산 서열을 코딩하는 임의의 폴리뉴클레오티드가 포함된다. 유사하게, 고도로 엄격한 조건 하

에 실질적으로 핵산의 전체 길이에 걸쳐 상기 폴리뉴클레오티드에 혼성화되는 (또한, 천연 발생 폴리뉴클레오티드가 아닌) 인공 핵산이 본 발명의 폴리뉴클레오티드이다. 한 구체예에서, 조성물은 본 발명의 폴리펩티드 및 부형제(예를 들어, 완충액, 물, 약학적으로 허용가능한 부형제 등)를 포함한다. 본 발명은 또한 본 발명의 폴리펩티드와 특이적으로 면역반응성을 갖는 항체 또는 항원형을 제공한다. 인공 폴리뉴클레오티드는 인간에 의해 제조되고, 천연적으로 발생되지 않은 폴리뉴클레오티드이다.

<237> 본 발명의 폴리뉴클레오티드는 또한 천연 발생 tRNA의 서열과 예를 들어 75% 이상, 80% 이상, 90% 이상, 95% 이상, 98% 이상, 또는 그 이상으로 동일한(단, 천연 발생 tRNA이 아닌) 인공 폴리뉴클레오티드를 포함한다. 폴리뉴클레오티드는 또한 천연 발생 tRNA의 서열과 예를 들어 75% 이상, 80% 이상, 90% 이상, 95% 이상, 98% 이상, 또는 그 이상으로 동일한(단, 100% 동일하지는 않은) 인공 폴리뉴클레오티드를 포함한다.

<238> 특정 구체예에서, 벡터(예를 들어, 플라스미드, 코스미드, 파지, 바이러스 등)는 본 발명의 폴리뉴클레오티드를 포함한다. 한 구체예에서, 벡터는 발현 벡터이다. 또 다른 구체예에서, 발현 벡터는 본 발명의 하나 이상의 폴리뉴클레오티드에 작동가능하게 연결된 프로모터를 포함한다. 또 다른 구체예에서, 세포는 본 발명의 폴리뉴클레오티드를 포함하는 벡터를 포함한다.

<239> 당업자라면 개시된 서열들의 많은 변이체도 본 발명에서 사용될 수 있다는 것을 인식할 것이다. 예를 들어, 기능적으로 동일한 서열을 생성시키는 개시된 서열의 보존적 변이체가 본 발명에 포함된다. 하나 이상의 개시된 서열에 혼성화되는, 핵산 폴리뉴클레오티드 서열의 변이체도 본 발명에 포함되는 것으로 간주된다. 예를 들어 표준 서열 비교 기법에 의해 결정되는, 본원에 개시된 서열의 독특한 하위서열도 또한 본 발명에 포함된다.

<240> 보존적 변이체

<241> 유전 코드의 축퇴성(degeneracy)으로 인해, "침묵 치환"(즉, 코딩된 폴리펩티드에서의 변경을 초래하지 않는 핵산 서열에서의 치환)은 아미노산 서열을 코딩하는 모든 핵산 서열의 내포된 특징이다. 유사하게, 아미노산 서열 내의 하나 또는 한정된 수의 아미노산이 고도로 유사한 성질을 가진 상이한 아미노산으로 치환되는 "보존적 아미노산 치환"도 개시된 구축물과 매우 유사한 것으로서 용이하게 확인된다. 각 개시된 서열의 그러한 보존적 변이체도 본 발명의 한 특징이다.

<242> 특별한 핵산 서열의 "보존적 변이체"는 동일하거나 본질적으로 동일한 아미노산 서열을 코딩하는 핵산을 지칭하거나, 핵산이 아미노산 서열을 코딩하지 않는 경우, 본질적으로 동일한 서열을 지칭한다. 당업자는 코딩된 서열 내에서 단일 아미노산 또는 적은 비율의 아미노산(전형적으로 5% 미만, 보다 더 전형적으로 4%, 2% 또는 1% 미만)을 변경시키거나, 부가하거나 결실시키는 개별 치환, 결실 또는 부가가 아미노산의 결실, 아미노산의 부가, 또는 화학적으로 유사한 아미노산으로의 아미노산의 치환을 초래하는 "보존적으로 변형된 변이"라는 것을 인식할 것이다. 따라서, 본 발명의 열거된 폴리펩티드 서열의 "보존적 변이"는 폴리펩티드 서열의 적은 비율, 전형적으로 5% 미만, 보다 전형적으로 2% 또는 1% 미만을 동일한 보존적 치환 기의 아미노산으로 치환시키는 것을 포함한다. 마지막으로, 핵산 분자의 코딩된 활성을 변경시키지 않는 서열의 부가, 예를 들어 비기능적 서열의 부가는 기본 핵산의 보존적 변이이다.

<243> 기능적으로 유사한 아미노산을 제공하는 보존적 치환 표가 당업계에 잘 공지되어 있고, 여기에서 아미노산 잔기는 유사한 화학적 성질(예를 들어, 방향족 측쇄 또는 양으로 하전된 측쇄)을 가진 또 다른 아미노산 잔기로 치환되므로, 폴리펩티드 분자의 기능성을 실질적으로 변화시키지 않는다. 하기 표는 유사한 화학적 성질의 천연 아미노산을 함유하는 예시적 군을 나타내는데, 여기에서 군 내의 치환은 "보존적 치환"이다.

<244> 보존적 아미노산 치환

<245> 비극성 및/또는 지방족 측쇄	극성 비하전 측쇄	방향족 측쇄	양으로 하전된 측쇄	음으로 하전된 측쇄
글리신 알라닌 발린 류신 이소류신 프롤린	세린 트레오닌 시스테인 메티오닌 아스파라긴 글루타민	페닐알라닌 티로신 트립토판	라이신 아르기닌 히스티딘	아스파르트산염 글루탐산염

<246> 핵산 혼성화

- <247> 비교 혼성화를 이용하여, 본원의 핵산의 보존적 변이체를 포함한, 본 발명의 핵산을 동정할 수 있고, 이 비교 혼성화 방법은 본 발명의 핵산을 구별하는 바람직한 방법이다. 또한, 높은 엄격도의 조건, 매우 높은 엄격도의 건조 및 극도로 매우 높은 엄격도의 조건 하에서, 서열 번호 8~56로 표시되는 핵산에 혼성화되는 표적 핵산도 본 발명의 특성이다. 그러한 핵산의 예에는 소정의 핵산 서열에 비해 하나 또는 몇개의 침묵 또는 보존적 핵산 치환을 가진 핵산이 포함된다.
- <248> 시험 핵산은, 그 자신이 완벽하게 매칭되는(matched) 상보적 표적에 대한 프로브에 50% 이상 혼성화될 때, 즉 완벽하게 매칭되는 프로브가 매칭되지 않은 표적 핵산들 중 임의의 핵산과의 혼성화에 대해 관찰된 것보다 약 $5 \times$ 내지 $10 \times$ 이상 더 높은 신호 대 노이즈 비로 완벽하게 매칭되는 상보적 표적에 결합하는 조건 하에서 시험 핵산이 표적과 프로브의 혼성화보다 절반 이상 더 높은 신호 대 노이즈 비로 혼성화되는 경우, 프로브 핵산에 특이적으로 혼성화된다고 말한다.
- <249> 핵산은 이들이 전형적으로 용액 중에서 결합될 때 "혼성화된다". 핵산은 특징이 잘 규명되어 있는 다양한 물리 화학적 힘, 예컨대 수소 결합, 용매 배제, 염기 적층 등으로 인해 혼성화된다. 핵산의 혼성화에 대한 광범위한 지침은 문헌 [Tijssen(1993), *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays", (Elsevier, New York)]뿐만 아니라, 문헌 [Current Protocols in Molecular Biology, Ausubel 등, eds., *Current Protocols*, Greene Publishing Associates, Inc.와 John Wiley & Sons, Inc.의 합작 회사(2004년 증보 개정판)]에서도 찾아볼 수 있고, 문헌 [Hames 및 Higgins(1995), *Gene Probes* 2, IRL Press at Oxford University Press, Oxford, England]은 올리고뉴클레오타이드를 비롯한 DNA 및 RNA의 합성, 표지, 검출 및 정량화에 대한 상세한 설명을 제공한다.
- <250> 서던(Southern) 또는 노던(Northern) 블롯에서 필터 상에 100개 이상의 상보적 잔기를 가진 상보적 핵산의 혼성화를 위한 엄격한 혼성화 조건의 한 예는 42℃에서 1 mg의 해파린을 함유하는 50% 포르말린이며, 여기에서 혼성화는 밤새 수행된다. 엄격한 세척 조건의 예는 15분 동안 65℃에서 $0.2 \times$ SSC를 사용한 세척이다(SSC 완충액에 대한 설명에 대해서는 문헌 [Sambrook 등, *Molecular Cloning A Laboratory Manual*(제3판), 제1-3권, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001] 참조). 종종, 높은 엄격도 세척 전에는 배경 프로브 신호를 제거하기 위해 낮은 엄격도 세척이 선행된다. 예시적인 낮은 엄격도 세척은 15분 동안 40℃에서 $2 \times$ SSC를 사용한 세척이다. 일반적으로, 특별한 혼성화 검정에서 관련없는 프로브에 대해 관찰된 신호 대 노이즈 비의 5배 (또는 그 이상) 신호 대 노이즈 비는 특이적 혼성화의 검출을 표시한다.
- <251> 핵산 혼성화 실험(예를 들어, 서던 혼성화 및 노던 혼성화)와 관련하여 "엄격한 혼성화 세척 조건"은 서열 의존적이고, 상이한 환경 파라미터 하에서 상이하다. 핵산의 혼성화에 대한 광범위 지침은 상기 문헌 [Tijssen(1993), *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays", (Elsevier, New York)], 및 문헌 [Hames 및 Higgins(1995), *Gene Probes* 2, IRL Press at Oxford University Press, Oxford, England]에서 찾아볼 수 있다. 엄격한 혼성화 및 세척 조건은 임의의 시험 핵산에 대해 실험적으로 용이하게 결정될 수 있다. 예를 들어, 엄격한 혼성화 및 세척 조건의 결정에 있어서, 선택된 기준 세트를 충족시킬 때까지 (예를 들어, 혼성화 또는 세척에서 온도 증가, 염 농도 감소, 세제 농도 증가 및/또는 유기 용매, 예컨대 포르말린 농도의 증가에 의해) 혼성화 및 세척 조건을 점점 강화시킨다. 예를 들어, 높은 엄격도의 혼성화 및 세척 조건에서, 매칭되지 않은 표적과 프로브의 혼성화에 대해 관찰되는 신호 대 노이즈 비보다 $5 \times$ 이상 더 높은 신호 대 노이즈 비로 프로브가 완벽하게 매칭되는 상보적 표적에 결합할 때까지 혼성화 및 세척 조건을 점점 강화시킨다.
- <252> "매우 엄격한" 조건은 특별한 프로브에 대한 열적 용점(T_m)과 동등하도록 선택된다. T_m 은 시험 서열의 50%가 (한정된 이온 강도 및 pH 하에) 완벽하게 매칭되는 프로브에 혼성화하는 온도이다. 본 발명의 목적을 위해, 일반적으로, "매우 엄격한" 혼성화 및 세척 조건은 한정된 이온 강도 및 pH에서 특정 서열에 대한 T_m 보다 약 5℃ 정도 낮도록 선택된다.
- <253> "매우 높은 엄격도" 혼성화 및 세척 조건은 완벽하게 매칭되는 상보적 표적 핵산과 프로브의 결합에 대한 신호 대 노이즈 비가 임의의 매칭되지 않은 표적 핵산과의 혼성화에 대해 관찰된 신호 대 노이즈 비보다 10배 이상 더 높아질 때까지 혼성화 및 세척 조건의 엄격도가 증가되는 조건이다. 그러한 조건 하에서 완벽하게 매칭되는 상보적 표적 핵산의 1/2 이상의 신호 대 노이즈 비로 프로브에 혼성화하는 표적 핵산은 매우 높은 엄격도 조건

하에서 프로브에 결합한다고 말한다.

- <254> 유사하게, 더욱 더 높은 엄격도 수준은 관련 혼성화 검정의 혼성화 및/또는 세척 조건을 점점 강화시킴으로써 결정될 수 있다. 예를 들어, 완벽하게 매칭되는 상보적 표적 핵산과 프로브의 결합에 대한 신호 대 노이즈 비가 임의의 매칭되지 않는 표적 핵산과의 혼성화에 대해 관찰된 신호 대 노이즈 비보다 적어도 10배, 20배, 50배, 100배, 또는 500배 또는 그 이상 높아질 때까지 혼성화 및 세척 조건의 엄격도를 증가시킨다. 그러한 조건 하에서 완벽하게 매칭되는 상보적 표적 핵산에 대한 신호 대 노이즈 비의 1/2 이상인 신호 대 노이즈 비로 프로브에 혼성화하는 표적 핵산은 극도로 높은 엄격도 조건 하에서 프로브에 결합한다고 말한다.
- <255> 엄격한 조건 하에서 서로 혼성화하지 않는 핵산은 이들이 코딩하는 폴리펩티드가 실질적으로 동일한 경우, 여전히 실질적으로 동일하다. 이는, 예를 들어 유전 코드에 의해 허용되는 최대 코돈 축퇴성을 사용하여 핵산의 복사체를 생성할 때 일어난다.
- <256> 독특한 하위서열
- <257> 일부 측면에서, 본 발명은 본원에 개시된 0-tRNA 및 0-RS의 서열로부터 선택된 핵산 내의 독특한 하위서열을 포함하는 핵산을 제공한다. 독특한 하위서열은 임의의 공지된 0-tRNA 또는 0-RS 핵산 서열에 상응하는 핵산과 비교 시에 독특하다. 정렬은, 예를 들어 디폴트 파라미터로 설정된 BLAST를 사용하여 수행할 수 있다. 임의의 독특한 하위서열은, 예를 들어 본 발명의 핵산 또는 관련 핵산을 확인하기 위한 프로브로서 유용하다.
- <258> 유사하게, 본 발명은 본원에 개시된 0-RS의 서열로부터 선택된 폴리펩티드 내의 독특한 하위서열을 포함하는 폴리펩티드를 포함한다. 여기에서, 독특한 하위서열은 임의의 공지된 폴리펩티드 서열에 상응하는 폴리펩티드와 비교 시에 독특하다.
- <259> 본 발명은 또한 엄격한 조건 하에서 0-RS의 서열로부터 선택된 폴리펩티드 내의 독특한 하위서열을 코딩하는 독특한 코딩 올리고뉴클레오타이드에 혼성화하는 표적 핵산을 제공하며, 여기에서 독특한 하위서열은 임의의 대조군 폴리펩티드(예를 들어 돌연변이에 의해 유도된 본 발명의 합성효소의 모(parental) 서열)에 상응하는 폴리펩티드와 비교 시에 독특하다. 독특한 서열은 전술한 바와 같이 결정된다.
- <260> 서열 비교, 동일성 및 상동성
- <261> 2개 이상의 핵산 또는 폴리펩티드 서열과 관련하여 용어 "동일한" 또는 "동일성(%)"은, 최대 상응성에 대해 비교되고 정렬될 때, 하기 서열 비교 알고리즘 (또는 당업자가 입수할 수 있는 다른 알고리즘) 중 하나를 이용하거나 육안 조사를 이용하여 측정한 때, 동일하거나 또는 특정화 비율(%)의 아미노산 잔기 또는 뉴클레오타이드를 가지는 2개 이상의 서열 또는 하위서열을 의미한다.
- <262> 2개의 핵산 또는 폴리펩티드(예를 들어, 0-tRNA 또는 0-RS를 코딩하는 DNA, 또는 0-RS의 아미노산 서열)와 관련하여 문구 "실질적으로 동일한"은 최대 상응성에 대해 비교되고 정렬된 경우, 서열 비교 알고리즘을 이용하거나 육안 조사를 이용하여 측정할 때 약 60%, 약 80%, 약 90 내지 95%, 약 98% 또는 약 99% 이상, 또는 그 이상의 뉴클레오타이드 또는 아미노산 잔기 동일성을 가지는 2개 이상의 서열 또는 하위서열을 의미한다. 그러한 "실질적으로 동일한" 서열은 전형적으로 실제 모 서열에 대한 언급 없이 "상동성"을 갖는 것으로 간주된다. 바람직하게는, "실질적 동일성"은 길이가 약 50개 이상의 잔기로 이루어진 서열의 영역에 걸쳐, 더 바람직하게는 약 100개 이상의 잔기로 이루어진 영역에 걸쳐 존재하며, 가장 바람직하게는 서열은 약 150개 이상의 잔기 또는 비교될 2개 서열의 전체 길이에 걸쳐 실질적으로 동일하다.
- <263> 단백질 및/또는 단백질 서열은, 공통된 모 단백질 또는 단백질 서열로부터 자연적으로 또는 인공적으로 유도되는 경우, "상동성"을 가진다. 유사하게, 핵산 및/또는 핵산 서열은 공통된 모 핵산 또는 핵산 서열로부터 자연적으로 또는 인공적으로 유도되는 경우, 상동성을 가진다. 예를 들어, 임의의 천연 발생 핵산은 하나 이상의 선택적 코돈을 포함하도록 임의의 이용가능한 돌연변이유발법에 의해 변형될 수 있다. 발현되는 경우, 이 돌연변이 핵산은 하나 이상의 비천연 아미노산을 포함하는 폴리펩티드를 코딩한다. 물론 돌연변이 과정은 하나 이상의 표준 코돈을 추가로 변경하여, 생성된 돌연변이 단백질 내에서도 하나 이상의 표준 아미노산을 변경시킬 수 있다. 상동성은 일반적으로 2개 이상의 핵산 또는 단백질 (또는 이의 서열) 사이의 서열 유사성으로부터 추정된다. 상동성을 입증하는 데 유용한 서열 사이의 정확한 유사성(%)은 해당 핵산 및 단백질에 따라 다르지만, 관용적으로 25% 정도의 낮은 서열 유사성이 상동성을 확립하는 데 사용된다. 보다 높은 수준의 서열 유사성(예를 들어, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99% 또는 그 이상) 또한 상동성을 확립하는 데 사용될 수 있다. 서열 유사성(%)을 결정하기 위한 방법(예를 들어, 디폴트 파라미터를 사용하는 BLASTP 및

BLASTN)은 본원에 기재되어 있으며 일반적으로 입수가 가능하다.

- <264> 서열 비교 및 상동성 결정을 위해, 전형적으로 한 서열은 시험 서열과 비교되는 기준 서열로 작용한다. 서열 비교 알고리즘을 사용하는 경우, 시험 서열 및 기준 서열을 컴퓨터에 입력하고, 필요하다면 하위서열 좌표를 지정하고, 서열 알고리즘 프로그램 파라미터를 지정한다. 그 후, 서열 비교 알고리즘은 지정된 프로그램 파라미터를 기초로 하여 기준 서열에 대한 시험 서열(들)의 서열 동일성(%)을 계산한다.
- <265> 비교를 위한 서열의 최적 정렬은, 예를 들어 문헌 [Smith 및 Waterman, *Adv. Appl. Math.* 2:482(1981)]의 국소 상동성 알고리즘, 문헌 [Needleman 및 Wunsch, *J. Mol. Biol.* 48:443(1970)]의 상동성 정렬 알고리즘, 문헌 (Pearson and Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444(1988))의 유사성 조사 방법, 이들 알고리즘의 자동화 실행(위스콘신주 매디슨 사이언스 드라이브 575 지네틱스 컴퓨터 그룹, 위스콘신 지네틱스 소프트웨어 패키지의 GAP, BESTFIT, FASTA 및 TFASTA), 또는 육안 조사(일반적으로 문헌 [*Current Protocols in Molecular Biology*, Ausubel 등, ed., *Current Protocols*, Breene Publishing Associates, Inc.와 John Wiley & Sons, Inc.의 합작 회사(2006년 증보 개정판)] 참조)로 수행할 수 있다.
- <266> 서열 동일성(%) 및 서열 유사성(%)을 결정하는 데 적합한 알고리즘의 한 예는 문헌 [Altschul 등(1990), *J. Mol. Biol.* 215:403-410]에 기재되어 있는 BLAST 알고리즘이다. BLAST 분석을 수행하기 위한 소프트웨어는 미국 국립생명공학 정보센터(National Center for Biotechnology Information) 웹사이트를 통해 공개적으로 입수가 가능하다. 이 알고리즘은 먼저 데이터베이스 서열에서 동일한 길이의 워드와 정렬될 때 일정한 양성 값 역치 점수 T와 일치하거나 이를 만족시키는 길이 W의 짧은 워드를 질의(query) 서열에서 동정함으로써 높은 점수를 기록하는 서열 쌍(HSP)을 확인하는 단계를 수반한다. T는 인접 워드 점수 역치로서 언급된다(상기 문헌 [Altschul 등, *J. Mol. Biol.* 215:403-410(1990)] 참조). 그러한 초기 인접 워드 히트(hit)는 이들을 포함하는 더 긴 HSP를 찾는 선별을 개시하기 위한 시드(seed)로 작용한다. 이어서, 워드 히트는 누적 정렬 점수가 증가될 수 있는 한 각 서열을 따라 양 방향으로 확장된다. 뉴클레오타이드 서열에 대한 누적 점수는 파라미터 M(1쌍의 매칭 잔기에 대한 보상 점수; 항상 > 0) 및 N(미스매칭된 잔기에 대한 벌점; 항상 < 0)을 사용하여 계산한다. 아미노산 서열의 경우, 점수화 매트릭스를 이용하여 누적 점수를 계산한다. 각 방향에서 워드 히트의 확장은, 누적 정렬 점수가 그의 최대 달성 값으로부터 X 양만큼 떨어졌을 때; 음의 점수를 기록한 하나 이상의 잔기 정렬의 축적으로 인해 누적 점수가 0 이하로 떨어졌을 때; 또는 어느 한 서열이 말단에 도달했을 때 정지된다. BLAST 알고리즘 파라미터 W, T 및 X는 정렬의 감수성 및 속도를 결정한다. (뉴클레오타이드 서열의 경우) BLASTN 프로그램은 디폴트로서 워드 길이(W) 11, 기대값(E) 10, 컷오프 100, M=5, N=-4, 및 양 가닥의 비교를 사용한다. 아미노산 서열의 경우, BLASTP 프로그램은 디폴트로서 워드 길이(W) 3, 기대 값(E) 10, 및 BLOSUM62 점수화 매트릭스를 사용한다(문헌 [Henikoff 및 Henikoff(1989), *Proc. Natl. Acad. Sci. USA* 89:10915] 참조).
- <267> 서열 동일성(%)을 계산하는 것 외에, BLAST 알고리즘은 2개 서열 사이의 유사성의 통계학적 분석도 수행한다(예를 들어, 문헌 [Karlín 및 Altschul(1993), *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787] 참조). BLAST 알고리즘에 의해 제공되는 유사성의 한 척도는 2개의 뉴클레오타이드 또는 아미노산 서열 사이에 매칭이 우연히 일어날 확률을 표시하는 최소 합계 확률(P(N))이다. 예를 들어, 기준 핵산과 시험 핵산의 비교에서 최소 합계 확률이 약 0.1 미만, 더 바람직하게는 약 0.01 미만, 가장 바람직하게는 약 0.001 미만일 때, 핵산은 기준 서열과 유사한 것으로 간주된다.
- <268> 돌연변이유발 및 다른 분자생물학 기법
- <269> 본 발명에 사용되는 본 발명의 폴리뉴클레오타이드 및 폴리펩티드는 분자생물학적 기법을 이용하여 조작될 수 있다. 분자생물학적 기법을 설명하는 일반 문헌에는 문헌 [Berger 및 Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology*, 제152권(1987), Academic Press, Inc., San Diego, CA]; 문헌 [Sambrook 등, *Molecular Cloning - A Laboratory Manual*(제3판), 제1-3권, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001], 및 문헌 [*Current Protocols in Molecular Biology*, Ausubel 등, eds., *Current Protocols*, Greene Publishing Associates, Inc.와 John Wiley & Sons, Inc.의 합작 회사(2006년 증보 개정판)]이 포함된다. 그 문헌에는, 예를 들어 다른 비천연 아미노산을 포함하는 단백질의 생산을 위한 셀렉터 코돈, 오르소고날 tRNA, 오르소고날 합성효소 및 이들의 쌍을 포함하는 유전자의 발생과 관련된 돌연변이유발법, 벡터의 사용, 프로모터 및 다른 다수의 관련 주제가 기재되어 있다.
- <270> 예를 들어, tRNA 분자를 돌연변이화하기 위해, tRNA의 라이브러리를 생성시키기 위해, 합성효소 라이브러리를 생성시키기 위해, 관심 단백질 또는 폴리펩티드 내에 비천연 아미노산을 코딩하는 셀렉터 코돈을 삽입하기 위해, 본 발명에서 다양한 유형들의 돌연변이유발법을 이용할 수 있다. 이 돌연변이유발법에는 부위-지정 돌연

변이유발, 랜덤 점 돌연변이유발, 상동성 재조합, DNA 서폴링 또는 다른 반복적 돌연변이유발법, 키메라 구축, 우라실을 함유하는 주형을 사용한 돌연변이유발, 올리고뉴클레오타이드-지정 돌연변이유발, 포스포로티오에이트-변형 DNA 돌연변이유발, 갭 이중 DNA를 사용한 돌연변이유발 등, 또는 이들의 임의의 조합이 포함된다. 다른 적절한 방법에는 점 mismatch 복구, 복구-결함 숙주 균주를 사용한 돌연변이유발, 제한-선택 및 제한-정제, 결실 돌연변이유발, 전체 유전자 합성에 의한 돌연변이유발, 이중 가닥 절단 복구 등이 포함된다. 예를 들어, 키메라 구축물을 사용한 돌연변이유발 또한 본 발명에 포함된다. 한 구체예에서, 돌연변이유발법은 천연 발생 분자 또는 변경되거나 돌연변이된 천연 발생 분자의 공지된 정보, 예컨대 서열, 서열 비교, 물성, 결정 구조 등에 의해 결정될 수 있다.

<271>

숙주 세포는 본 발명의 폴리뉴클레오타이드 또는 폴리뉴클레오타이드를 포함하는 구축물, 예컨대 클로닝 벡터 또는 발현 벡터일 수 있는 벡터로 유전적으로 공학처리(예, 형질전환, 형질도입 또는 형질감염)된다. 예를 들어, 오르소고날 tRNA, 오르소고날 tRNA 합성효소, 및 유도체화될 단백질에 대한 코딩 영역은 원하는 숙주 세포에서 기능성을 갖는 유전자 발현 조절 요소에 작동가능하게 연결된다. 전형적인 벡터는 전사 및 번역 종결자, 전사 및 번역 개시 서열, 및 특별한 표적 핵산의 발현 조절에 유용한 프로모터를 포함한다. 임의적으로, 그 벡터는 하나 이상의 독립적인 종결자, 진핵생물 또는 원핵생물, 또는 이들 양자 모두에서 카세트의 복제를 허용하는 서열(예, 셔틀 벡터), 및 원핵세포 시스템 및 진핵세포 시스템 양자 모두를 위한 선택 마커를 포함하는 일반 발현 카세트를 포함한다. 벡터는 원핵생물, 진핵생물, 또는 바람직하게는 이들 양자 모두에서 복제 및/또는 일체화에 적합하다. 문헌 [Gilman 및 Smith, *Gene* 8:81(1979)]; 문헌 [Roberts 등, *Nature*, 328:731(1987)]; 문헌 [Schneider, B 등, *Protein Expr. Purif.* 6435:10(1995)]; 문헌 [*Current Protocols in Molecular Biology*, Ausubel 등, eds., *Current Protocols*, Greene Publishing Associates, Inc.와 John Wiley & Sons, Inc.의 합작 회사(2004년 증보 개정판)]; 문헌 [Sambrook 등, *Molecular Cloning - A Laboratory Manual*(제3판), 제1-3권, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001]; 및 문헌 [Berger 및 Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology*, 제152권(1987), Academic Press, Inc., San Diego, CA]을 참조한다. 벡터는, 예를 들어 플라스미드, 세균, 바이러스, 네이키드(naked) 폴리뉴클레오타이드, 또는 접합된 폴리뉴클레오타이드의 형태일 수 있다. 벡터는 전기천공법(문헌 [Fromm 등(1985), *Proc. Natl. Acad. Sci. USA* 82, 5824]), 바이러스 벡터에 의한 감염, 소형 비드 또는 입자로 이루어진 매트릭스 내에 또는 표면에 핵산을 가진 소립자에 의한 고속 탄도(ballistic) 침투(문헌 [Klein 등, *Nature* 327, 70-73(1987)]) 등을 비롯한 표준 방법에 의해 세포 및/또는 미생물 내로 도입된다.

<272>

클로닝에 유용한 세균 및 박테리오파지는 당업자에게 널리 공지되어 있고, 각종 공급원으로부터 입수가 가능하다. 예를 들어, 미국 유형 배양 수집(ATCC; 버지니아주 매나세스) 및 ATCC에 의해 발행된 문헌 [*The ATCC Catalogue of Bacteria and Bacteriophag*(1996) Gherna 등(편저)] 참고. 시퀀싱, 클로닝 및 분자생물학의 다른 측면에 대한 부가적 기본 절차 및 기초적 이론 고려사항도 또한 문헌 [Sambrook 등, *Molecular Cloning - A Laboratory Manual*(제3판), 제1-3권, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001]; 문헌 [*Current Protocols in Molecular Biology*, Ausubel 등, eds., *Current Protocols*, Greene Publishing Associates, Inc.와 John Wiley & Sons, Inc.의 합작 회사(2004년 증보 개정판)]; 및 문헌 [Watson 등(1992) *Recombinant DNA* 제2판 Scientific American Books, NY)에서 찾아볼 수 있다. 또한, 본질적으로 임의의 핵산(및 표준 핵산이든 아니면 비표준 핵산이든 사실상 임의의 표지된 핵산)은 다양한 상업적 공급처(예를 들어, 미드랜드 서티파이드 리에이전트 컴퍼니(Midland Certified Reagent Company)(미국 텍사스주 미드랜드 소재), 더 그레이트 아메리칸 진 컴퍼니(The Great American Gene Company)(미국 캘리포니아주 라모나 소재), 익스프레스젠 인코포레이티드(ExpressGen Inc.)(미국 일리노이주 시카고 소재), 오페론 테크놀로지스 인코포레이티드(Operon Technologies Inc.)(미국 캘리포니아주 알라메다 소재) 및 다른 많은 공급처) 중 임의의 공급처로부터 맞춤 주문받거나 또는 일반 주문받을 수 있다.

<273>

공학처리된 숙주 세포는, 예를 들어 선별 단계, 프로모터의 활성화 또는 형질전환체의 선택과 같은 공정에 적절하도록 변형된 통상적인 영양 배지 중에서 배양할 수 있다. 이 세포는 임의적으로 형질전환 유기체로 배양될 수 있다. 예를 들어, 세포 단리 및 배양(예를 들어, 후속 핵산 단리)에 유용한 다른 참고문헌에는 문헌 [Freshney(1994) *Culture of Animal Cells, a Manual of Basic Technique*, 제3판, Wiley-Liss, New York] 및 이 문헌에 인용된 참고문헌 [Payne 등(1992) *Plant Cell and Tissue Culture in Liquid Systems* John Wiley & Sons, Inc. New York, NY]; 문헌 [Gamborg 및 Phillips(편저)(1995) *Plant Cell, Tissue and Organ Culture*]; 문헌 [Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York)], 및 [Atlas 및 Parks(편저), *The Handbook of Microbiological Media*(1993) CRC Press, Boca Raton, FL]이 포함된다.

다.

<274>

관심 단백질 및 폴리펩티드

<275>

특정된 위치에 비천연 아미노산을 이용하여 세포 내 단백질을 생성시키는 방법도 또한 본 발명의 한 특성이다. 예를 들어, 방법은 하나 이상의 선택적 코돈을 포함하고 단백질을 코딩하는 핵산을 포함한 세포를 적절한 배지에서 성장시키는 단계; 및 비천연 아미노산을 제공하는 단계를 포함하는데, 여기에서 세포는 세포에서 기능하고 선택적 코돈을 인식하는 오르소고날-tRNA(O-tRNA); 및 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는 오르소고날 아미노아실-tRNA 합성효소(O-RS)를 추가로 포함한다. 이 방법에 의해 생성된 단백질도 또한 본 발명의 한 특성이다.

<276>

특정 구체예에서, O-RS는 발현 시스템 내에서 임의의 내인성 tRNA에 비해 동족체 O-tRNA의 아미노아실화에 대해 더 선호성을 가진다. O-tRNA 및 O-RS가 동등한 몰 농도로 존재하는 경우, O-RS에 의해 충전되는 O-tRNA와 내재성 tRNA 사이의 상대적 비는 1:1 초과, 보다 바람직하게는 약 2:1 이상, 보다 더 바람직하게는 5:1, 더욱 더 바람직하게는 10:1, 보다 더 바람직하게는 20:1, 더욱 더 바람직하게는 50:1, 보다 더 바람직하게는 75:1, 더욱 더 바람직하게는 95:1, 98:1, 99:1, 100:1, 500:1, 1,000:1, 5,000:1 또는 그 이상이다.

<277>

본 발명은 또한 비천연 아미노산을 포함한 단백질을 포함하는 조성물을 제공한다. 특정 구체예에서, 단백질은 치료 단백질, 진단 단백질, 산업용 효소 또는 이의 일부의 아미노산 서열과 75% 이상 동일한 아미노산 서열을 포함한다.

<278>

본 발명의 조성물 및 본 발명의 방법에 의해 제조된 조성물은 임의적으로 세포 내에 있다. 이에 따라 본 발명의 O-tRNA/O-RS 쌍 또는 개별 성분은 숙주 시스템의 번역 기구에 사용될 수 있고, 이로써 비천연 아미노산이 단백질 내로 도입되게 된다. 국제 특허 공개 공보 제WO 2004/094593호(2004년 4월 16일 출원)(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE"), 및 제WO 2002/085923호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINOS")가 이 방법을 기재하고 있고, 본원에 참조 인용된다. 예를 들어, O-tRNA/O-RS 쌍이 숙주, 예를 들어 설치류 또는 영장류 세포 내로 도입될 때, 그 쌍은 생체내 선택적 코돈에 대한 반응으로 비천연 아미노산(예, 도 1의 비천연 아미노산)이 단백질 내로 도입되게 한다. 시스템에 첨가된 비천연 아미노산은 성장 배지에 외인적으로 첨가될 수 있는 합성 아미노산, 예컨대 페닐알라닌 또는 티로신의 유도체일 수 있다. 임의적으로, 본 발명의 조성물은 시험관내 번역 시스템 또는 생체내 시스템(들) 내에 있을 수 있다.

<279>

본 발명의 세포는 큰 유용량의 비천연 아미노산을 포함하는 단백질을 합성하는 능력을 제공한다. 일부 측면에서, 조성물은 임의적으로 예를 들어 10 마이크로그램 이상, 50 마이크로그램 이상, 75 마이크로그램 이상, 100 마이크로그램 이상, 200 마이크로그램 이상, 250 마이크로그램 이상, 500 마이크로그램 이상, 1 밀리그램 이상, 10 밀리그램 이상, 또는 그 이상의 양, 또는 생체내 단백질 생성 방법(재조합 단백질 생성 및 정제에 대한 상세 내용이 본원에 나와 있음)으로 달성될 수 있는 양의 비천연 아미노산을 포함한 단백질을 포함한다. 또 다른 측면에서, 단백질은 임의적으로, 예를 들어 세포 용매물, 완충액, 약학적 완충액 또는 기타 액체 현탁액 내(예를 들어, 예컨대 약 1 nL 내지 약 100 L의 임의의 체적), 조성물 내에 예를 들어 리터 당 10 마이크로그램 이상의 단백질, 리터 당 50 마이크로그램 이상의 단백질 리터 당 75 마이크로그램 이상의 단백질, 리터 당 100 마이크로그램 이상의 단백질, 리터 당 200 마이크로그램 이상의 단백질, 리터 당 250 마이크로그램 이상의 단백질, 리터 당 500 마이크로그램 이상의 단백질, 리터 당 1 밀리그램 이상의 단백질, 또는 리터 당 10 밀리그램 이상의 단백질의 농도로 존재한다. 예를 들어 하나 이상의 비천연 아미노산을 포함하는 세포 내에서 다량(예를 들어, 다른 방법, 예를 들어 시험관내 번역으로 전형적으로 가능한 양을 초과하는 양)의 단백질을 생성시키는 것도 본 발명의 특성이다.

<280>

비천연 아미노산의 도입은 단백질 구조 및/또는 기능의 변화를 조정하기 위해 예를 들어, 크기, 산도, 친핵성, 수소 결합, 소수성, 단백질분해효소 표적 부위의 접근가능성, (예를 들어, 단백질 어레이의 경우) 잔기에 대한 표적, 표지 또는 반응성 기의 도입 등을 변화시키기 위해 수행될 수 있다. 비천연 아미노산을 포함하는 단백질은 개선된 또는 심지어 완전히 새로운 촉매 활성 또는 물성을 가질 수 있다. 예를 들어, 독성, 생체분포, 구조적 성질, 분광학적 성질, 화학적 성질 및/또는 광화학적 성질, 촉매 능력, 반감기(예를 들어, 혈청 반감기), 예를 들어 공유 Ehe는 비공유적으로 다른 분자와 반응하는 능력 등의 성질을, 임의적으로 비천연 아미노산을 단백질 내에 포함시킴으로써 변형시킨다. 하나 이상의 비천연 아미노산을 포함하는 단백질을 포함하는 조성물은 예를 들어, 새로운 치료, 진단, 촉매 효소, 산업용 효소, 결합 단백질(예, 항체), 및 예를 들어, 단백질 구조 및 기능 연구에 유용하다. 예를 들어, 문헌 [Dougherty, (2000) *Unnatural Amino Acids as Probe of Protein Structure and Function*, *Current Opinion in Chemical Biology*, 4:645-652]을 참조한다.

- <281> 본 발명의 일부 측면에서, 조성물은 1개 이상, 예를 들어 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6개 이상, 7개 이상, 8개 이상, 9개 이상 또는 10개 이상 또는 그 이상의 비천연 아미노산을 가진 하나 이상의 단백질을 포함한다. 그 비천연 아미노산은 동일하거나 상이할 수 있는데, 예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개 또는 그 이상의 상이한 비천연 아미노산을 포함하는 단백질 내에 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개 이상 또는 그 이상의 상이한 부위가 존재할 수 있다. 또 다른 측면에서, 조성물은 단백질에 존재하는 (전부는 아닌) 1개 이상의 특별한 아미노산이 비천연 아미노산인 단백질을 포함한다. 1개 이상의 비천연 아미노산을 가진 소정의 단백질의 경우, 비천연 아미노산은 동일하거나 상이할 수 있다(예를 들어, 단백질은 2개 이상의 상이한 유형의 비천연 아미노산을 포함하거나, 2개 이상의 동일한 비천연 아미노산을 포함할 수 있다). 2개 이상의 비천연 아미노산을 가진 소정의 단백질의 경우, 비천연 아미노산은 동일하거나 상이하거나, 하나 이상의 상이한 비천연 아미노산과 동일한 종류의 다수의 비천연 아미노산의 조합물일 수 있다.
- <282> 비천연 아미노산을 포함하는, 본질적으로 임의의 단백질 (또는 이의 일부) (및 예를 들어 1개 이상의 선택적 코돈을 포함하는, 임의의 상응하는 코딩 핵산)은 본원의 조성물 및 방법을 사용하여 생성될 수 있다. 공지된 수많은 단백질을 동정하기 위해 어떠한 시도도 하지 않고, 예를 들어 관련 번역 시스템 내에 1개 이상의 적절한 선택적 코돈이 포함되도록 임의의 이용가능한 돌연변이 방법을 조합함으로써 임의의 공지된 단백질이 하나 이상의 비천연 아미노산을 포함하도록 변형시킬 수 있다. 공지된 단백질에 대한 통상의 서열 저장 기관에는 유전자은행, EMBL, DDBJ 및 NCBI가 포함된다. 다른 저장 기관은 인터넷 검색에 의해 쉽게 찾을 수 있다.
- <283> 전형적으로, 단백질은 임의의 이용가능한 단백질(예를 들어, 치료 단백질, 진단 단백질, 산업용 효소 또는 이의 일부 등)과 예를 들어 60% 이상, 70% 이상, 75% 이상, 80% 이상, 90% 이상, 95% 이상, 99% 이상 또는 그 이상 동일하며, 하나 이상의 비천연 아미노산을 포함한다. 하나 이상의 비천연 아미노산을 포함하도록 변형될 수 있는 치료 단백질, 진단 단백질 및 다른 단백질의 예를 2004년 4월 16일 출원된 국제특허공개 공보 제 WO2004/094593호(2004년 4월 16일 출원)(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE"); 및 제WO 2002/085923호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")에서 찾아볼 수 있지만, 이들로 한정되는 것은 아니다. 1개 이상의 비천연 아미노산을 포함하도록 변형될 수 있는 치료 단백질, 진단 단백질 및 다른 단백질의 예에는 예를 들어, 허루딘, 알파-1 항트립신, 엔지오스타틴, 항응혈성 인자, 항체(항체에 대한 상세한 설명이 이하에 제공됨), 아폴리포단백질, 아포단백질, 심방성 나트륨이뇨 인자, 심방성 나트륨이뇨 폴리펩티드, 심방성 펩티드, C-X-C 케모카인(예를 들어, T39765, NAP-2, ENA-78, Gro-a, Gro-b, Gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG), 칼시토닌, CC 케모카인(예를 들어, 단백질 세포 화학유인물질 단백질-1, 단백질 세포 화학유인물질 단백질-2, 단백질 세포 화학유인물질 단백질-3, 단백질 세포 염증성 단백질-1 알파, 단백질 세포 염증성 단백질-1 베타, RANTES, I309, R83915, R91733, HCC1, T58847, D31065, T64262), CD40 리간드, C-키트 리간드, 카스페이스, 콜라겐, 콜로니 자극 인자(CSF), 보체 인자 5a, 보체 억제제, 보체 수용체 1, 사이토카인(예를 들어, 상피 호중구 활성화 펩티드-78, GRO α /MGSa, GRO β , GRO γ , MIP-1 α , MIP-1 δ , MCP-1), 상피 성장 인자(EGF), 에리트로포이에틴("EPO"), 엑스폴리에이팅(Exfoliating) 독소 A 및 B, 인자 IX, 인자 VII, 인자 VIII, 인자 X, 섬유아세포 성장 인자(FGF), 피브리노겐, 피브로넥틴, G-CSF, GM-CSF, 글루코세레브로시다제, 고나도트로핀(Gonadotropin), 성장 인자, 헤지호그(Hedgehog) 단백질(예를 들어, Sonic, Indian, Desert), 헤모글로빈, 간세포 성장 인자(HGF), 허루딘, 인간 혈청 알부민, 인슐린, 인슐린-유사 성장 인자(IGF), 인터페론(예를 들어, IFN- α , IFN- β , IFN- γ), 인터루킨(예를 들어, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12 등), 케라티노사이트 성장 인자(KGF), 락토페린, 백혈병 억제 인자, 루시페라제, 뉴르투린(Neurturin), 호중구 억제 인자(NIF), 온코스타틴 M, 골생성 단백질, 부갑상선 호르몬, PD-ECSF, PDGF, 펩티드 호르몬(예를 들어, 인간 성장 호르몬), 플레이오토로핀, 프로카스페이스-3, 프로카스페이스-9, 단백질 A, 단백질 G, 발열원성 외독소 A, B 및 C, 렉신(Relaxin), 레닌(Renin), SCF, 가용성 보체 수용체 I, 가용성 I-CAM 1, 가용성 인터루킨 수용체(IL-1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15), 가용성 TNF 수용체, 소마토메딘(Somatostatin), 소마토스타틴(Somatostatin), 소마토포트로핀(Somatotropin), 스트렙토키나제, 수퍼항원, 즉 스태필로코컬 장독소(SEA, SEB, SEC1, SEC2, SEC3, SED, SEE), 수퍼옥사이드 디스무타제(SOD), 독성 쇼크증후군 독소(TSST-1), 티모신(Thymosin) 알파 1, 조직 플라스미노겐 활성화자, 종양 괴사 인자 베타(TNF 베타), 종양 괴사 인자 수용체(TNFR), 종양 괴사 인자-알파(TNF 알파), 혈관 내피세포 성장 인자(VEGF), 유로키나제 및 다른 많은 것들이 포함되나, 이에 한정되지 않는다.
- <284> 비천연 아미노산을 생체내에서 단백질 내로 도입하기 위한 본원에 기재된 조성물 및 방법을 사용하여 제조될 수 있는 단백질의 한 부류는 전사 조절자 또는 이의 일부를 포함한다. 전사 조절자의 예에는 세포 성장, 분화, 조절 등을 조절하는 유전자 및 전사 조절자 단백질이 포함된다. 전사 조절자는 원핵생물, 바이러스 및 진핵생물

(예컨대, 진균류, 식물, 효모, 곤충, 및 포유동물을 비롯한 동물을 포함함)에서 발견되며, 이는 광범위한 치료 표적을 제공한다. 발현 및 전사 활성화자는 많은 기작을 통해, 예를 들어 수용체에 결합하고, 신호 전달 캐스케이드를 자극하고, 전사 인자의 발현을 조절하고, 프로모터 및 인핸서에 결합하고, 프로모터 및 인핸서에 결합하는 단백질에 결합하고, DNA를 풀고, pre-mRNA를 스플라이싱하고, RNA를 폴리아데닐화하고, RNA를 분해시킴으로써 전사를 조절한다는 것이 인식될 것이다.

<285> 본 발명의 단백질의 한 부류(예를 들어, 하나 이상의 비천연 아미노산을 가지는 단백질)는 생물학적 활성 단백질, 예컨대 히루딘, 사이토카인, 염증성 분자, 성장 인자, 이들의 수용체 및 종양유전자 생성물, 예컨대 인터루킨(예, IL-1, IL-2, IL-8 등), 인터페론, FGF, IGF-I, IGF-II, FGF, PDGF, TNF, TGF- α , TGF- β , EGF, KGF, SCF/c-키트, CD40L/CD40, VLA-4/VCAM-1, ICAM-1/LFA-1 및 히알루린/CD44; 신호 전달 분자 및 상응하는 종양유전자 생성물, 예컨대 Mos, Ras, Raf 및 Met; 및 전사 활성화자 및 억제자, 예컨대 p53, Tat, Fos, Myc, Jun, Myb, Rel; 및 스테로이드 호르몬 수용체, 예컨대 에스트로겐, 프로게스테론, 테스토스테론, 알도스테론, LDL 수용체 리간드 및 코르티코스테론에 대한 수용체를 포함한다.

<286> 본 발명에 의해 하나 이상의 비천연 아미노산을 가지는 효소(예를 들어, 산업용 효소) 또는 이의 일부가 제공된다. 효소의 예에는, 예를 들어 아미다제(amidase), 아미노산 라세마제(racemase), 아실라제(acylase), 데할로게나제(dehalogenase), 디옥시게나제(dioxygenase), 디아릴프로판 퍼옥시다제, 에피머라제(epimerase), 에폭시드 히드롤라제(hydrolase), 에스터라제, 이소머라제, 키나제, 글루코스 이소머라제, 글리코시다제, 글리코실 트랜스퍼라제, 할로퍼옥시다제, 모노옥시게나제(예, p450), 리파제, 리그닌 퍼옥시다제, 니트릴 히드라타제(hydratase), 니트릴라제(nitrilase), 단백질분해효소, 포스파타제, 셉틸리신, 트랜스아미나제 및 뉴클레아제가 포함되나, 이에 한정되지 않는다.

<287> 이들 단백질의 대다수는 시판되고 있으며(예를 들어, 시그마 바이오사이언시스(Sigma BioSciences) 참조), 상응하는 단백질 서열 및 유전자, 및 전형적으로 이들의 많은 변이체가 공지되어 있다(예, 유전자는행 참조). 임의의 이들 단백질은 본 발명에 따라 하나 이상의 비천연 아미노산의 삽입에 의해 변형되어, 예를 들어 그의 1개 이상의 관심 치료적, 진단적 또는 효소적 성질이 변경될 수 있다. 치료적으로 관련된 성질의 예에는 혈청 반감기, 저장 수명 반감기, 안정성, 면역원성, 치료 활성, (예를 들어, 리포터 기(예를 들어, 표지 또는 표지 결합 부위)를 비천연 아미노산에 포함시켜 얻는) 검출용이성, LD₅₀의 감소 또는 다른 부작용, 위관을 통해 체내로 들어가는 능력(예를 들어, 경구 이용가능성) 등을 포함한다. 관련된 진단적 성질의 예에는 저장 수명 반감기, 안정성, 진단 활성, 검출용이성 등이 포함된다. 관련 효소적 특성의 예에는 저장 수명 반감기, 안정성, 효소적 활성, 생산 용량 등이 포함된다.

<288> 각종 다른 단백질도 또한 본 발명의 조성물 및 방법을 이용하여 하나 이상의 비천연 아미노산을 포함하도록 변형될 수 있다. 예를 들어, 본 발명은 비천연 아미노산을 갖는 하나 이상의 백신 단백질, 예컨대 감염성 진균류, 예컨대 아스페길러스(*Aspergillus*), 칸디다(*Candida*) 종; 세균, 특히 병원성 세균에 대한 모델로 사용되는 이. 콜라이 뿐만 아니라, 스태필로코커스(*Staphylococcus*)(예를 들어, 아루레우스(*aureus*)), 또는 스트렙토코커스(*Streptococcus*)(예를 들어, 뉴모니아(*pneumoniae*))와 같이 의학적으로 중요한 세균; 포자류(예를 들어, 플라스모디아(*Plasmodia*)), 근족충류(예를 들어, 엔타모에바(*Entamoeba*)) 및 편모충(트리파노소마(*Trypanosoma*), 레이쉬마니아(*Leishmania*), 트리코모나스(*Trichomonas*), 지아르디아(*Giardia*) 등)와 같은 원생동물; (+) RNA 바이러스(예를 들어, 폭스바이러스, 예컨대 백신시아(*vaccina*); 피코르나바이러스, 예컨대 폴리오(*polio*); 토가바이러스, 예컨대 루벨라(*rubella*); 플라비바이러스, 예컨대 HCV; 및 코로나바이러스를 포함함), (-) RNA 바이러스(예컨대, 래브도바이러스, 예컨대 VSV; 파라믹소바이러스, 예컨대 RSV; 오소믹소바이러스, 예컨대 인플루엔자; 변아바이러스; 및 아레나바이러스), dsDNA 바이러스(예를 들어, 레오바이러스), RNA로부터 DNA를 합성하는 바이러스, 즉 레트로바이러스, 예컨대 HIV 및 HTLV, 및 DNA로부터 RNA를 합성하는 일부 바이러스, 예컨대 B형 간염과 같은 바이러스로부터 유도된 단백질 내의 하나 이상의 천연 아미노산을 치환시키는 것을 포함할 수 있다.

<289> 곤충 내성 단백질(예, Cry 단백질), 전분 및 지질 생성 효소, 식물 및 곤충 독소, 독소-내성 단백질, 미코톡신 해독 단백질, 식물 성장 효소(예를 들어, 리볼로스 1,5-비스포스페이트 카르복실라제/옥시게나제 "RUBISCO"), 리폭시게나제(LOX), 및 포스포에놀피루베이트(PEP) 카르복실라제와 같이 농업 관련 단백질도 또한 비천연 아미노산 변형에 적합한 표적이다.

<290> 특정 구체예에서, 본 발명의 방법 및/또는 조성물 내의 관심 단백질 또는 폴리펩티드 (또는 이의 일부)가 핵산에 코딩된다. 전형적으로, 핵산은 하나 이상의 선택터 코돈, 2개 이상의 선택터 코돈, 3개 이상의 선택터 코돈, 4개 이상의 선택터 코돈, 5개 이상의 선택터 코돈, 6개 이상의 선택터 코돈, 7개 이상의 선택터 코돈, 8개 이상

의 선택터 코돈, 9개 이상의 선택터 코돈, 10개 또는 그 이상의 선택터 코돈을 포함한다.

<291> 관심 단백질 또는 폴리펩티드를 코딩하는 유전자는 당업자에게 잘 공지되어 있으면서 본 명세서의 "돌연변이유발 및 기타 분자생물학 기법(Mutagenesis and Other Molecular Biology Techniques)"에 기재된 방법을 사용하여, 비천연 아미노산의 도입을 위한 하나 이상의 선택터 코돈을 포함하도록 돌연변이화될 수 있다. 예를 들어, 관심 단백질에 대한 핵산은 하나 이상의 비천연 아미노산의 삽입을 위해 제공된 하나 이상의 선택터 코돈을 포함하도록 돌연변이화한다. 본 발명은, 예를 들어 하나 이상의 비천연 아미노산을 포함하는 임의의 단백질의 임의의 그러한 변이체, 예를 들어 돌연변이체를 포함한다. 유사하게, 본 발명은 상응하는 핵산, 즉 하나 이상의 비천연 아미노산을 코딩하는 하나 이상의 선택터 코돈을 가진 임의의 핵산을 포함한다.

<292> 비천연 아미노산을 포함하는 단백질을 제조하기 위해, 오르소고달 tRNA/RS 쌍을 통해 비천연 아미노산을 생체 내에서 도입하기에 적합한 숙주 세포 및 유기체를 사용할 수 있다. 숙주 세포는 오르소고달 tRNA를 발현하는 하나 이상의 벡터, 오르소고달 tRNA 합성효소, 및 유도체화된 단백질을 코딩하는 벡터로 유전적으로 공학처리된다(예를 들어, 형질전환, 형질도입 또는 형질감염된다). 이 성분들 각각은 동일한 벡터 상에 존재할 수 있거나, 별도의 벡터 상에 존재할 수 있거나, 또는 2개의 성분은 한 벡터 상에 존재하고 제3 성분은 제2 벡터 상에 존재할 수 있다. 상기 벡터는, 예를 들어 플라스미드, 세균, 바이러스, 네이키드 폴리뉴클레오티드 또는 접합된 폴리뉴클레오티드의 형태일 수 있다.

<293> 면역반응성에 의한 폴리펩티드의 정의

<294> 본 발명의 폴리펩티드가 각종 새로운 폴리펩티드 서열(예를 들어, 본원의 번역 시스템에서 합성된 단백질의 경우, 비천연 아미노산을 포함하는 폴리펩티드, 또는 예를 들어 신규 합성효소의 경우, 표준 아미노산의 신규 서열을 포함하는 폴리펩티드)을 제공하기 때문에, 그 폴리펩티드는, 예를 들어 면역학적 분석에서 인식될 수 있는 새로운 구조적 특성도 제공한다. 본 발명의 폴리펩티드에 특이적으로 결합하는 항혈청의 생성뿐만 아니라, 항혈청에 결합하는 폴리펩티드는 본 발명의 한 특성이다. 본원에서 사용되는 용어 "항체"는 실질적으로 면역글로불린 유전자 또는 면역글로불린 유전자들에 의해 코딩되는 폴리펩티드, 또는 피분석물(항원)에 특이적으로 결합하여 인식하는 상기 폴리펩티드의 단편을 포함하나, 이에 한정되는 것은 아니다. 예로는 다중클론 항체, 단일클론 항체, 키메라 항체 및 단쇄 항체 등이 있다. Fab 단편, 및 파지 디스플레이를 비롯한 발현 라이브러리에 의해 생성된 단편을 비롯한 면역글로불린 단편 또한 본원에서 사용되는 용어 "항체"에 포함된다. 예를 들어, 항체 구조 및 용어에 대해서는 문헌 [Paul, Fundamental Immunology, 제4판, 1999, Raven Press, New York]을 참조한다.

<295> 면역분석에서 사용하기 위한 항혈청을 생성하기 위해, 본원에 기재된 바와 같이 하나 이상의 면역원성 폴리펩티드를 생성시키고 정제한다. 예를 들어, 재조합 단백질은 재조합 세포에서 생성될 수 있다. 사실상 유전적으로 동일하기 때문에 결과가 더 재현가능하므로 본 분석에서 사용되는 근친교배 마우스 종을 표준 아주반트(adjuvant)(예를 들어, 프로인트(Freund's) 아주반트) 및 표준 마우스 면역화 프로토콜과 함께 사용된 면역원성 단백질(들)로 면역화한다(예를 들어, 항체 생성, 면역분석 포맷, 및 구체적인 면역반응성을 측정하는 데 이용될 수 있는 조건에 대한 표준 설명을 위해, 문헌 [Harlow 및 Lane(1988) Antibodies. A Laboratory Manual, Cold Spring Harbor Publications, New York]을 참조한다). 단백질, 항체, 항혈청 등에 대한 부가적 상세한 설명은 국제 특허 공개 공보 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE"); 제2002/085923호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS"); 제WO 2004/035605호(발명의 명칭: "GLYCOPROTEIN SYNTHESIS"); 및 제WO 2004/058946호(발명의 명칭: "PROTEIN ARRAYS")에서 찾을 수 있다.

<296> O-tRNA, O-RS, 및 O-tRNA/O-RS 쌍의 용도

<297> 본 발명의 조성물 및 본 발명의 방법에 의해 제조된 조성물은 임의적으로 세포 내에 있다. 이에 따라 본 발명의 O-tRNA/O-RS 쌍 또는 개별 성분은 숙주 시스템의 번역 기구에 사용될 수 있고, 이로써 비천연 아미노산이 단백질 내로 도입되게 된다. 국제 특허 공개 공보 제WO 2002/085923호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINOS"가 이 방법을 기재하고 있고, 본원에 참조 인용된다. 예를 들어, O-tRNA/O-RS 쌍이 숙주, 예를 들어 예셀리키아 콜라이 또는 효모 내로 도입될 때, 그 쌍은 생체내 선택터 코돈, 예를 들어 엠버 년센스 코돈에 대한 반응으로, 비천연 아미노산이 단백질, 예를 들어 미오글로빈 시험 단백질 또는 치료 단백질 내로 도입되도록 한다. 임의적으로, 본 발명의 조성물은 시험관내 번역 시스템 또는 세포내 생체내 시스템(들) 내에 있을 수 있다. 비천연 아미노산을 갖는 단백질은 광범위한 용도들 중 임의의 용도에 사용될 수 있다. 예를 들어, 단백질 내로 도입된 비천연 부분은 광범위한 변형들 중 임의의 변형, 예를 들어 표지 또는 염료와 같은 소분자 및/또는 생체분자를 갖는 다른 단백질과 가교되도록 하기 위한 표적으로 작용할 수 있다. 이러한 변형으로, 비

천연 아미노산의 도입은 향상된 치료 단백질을 초래할 수 있고, 효소의 촉매 기능을 변경 또는 향상시키기 위해 사용될 수 있다. 일부 측면에서, 단백질 내 비천연 아미노산의 도입 및 후속 변형은 단백질 구조, 다른 단백질과의 상호작용 등에 대한 연구를 도모할 수 있다.

<298> 키트

<299> 키트도 또한 본 발명의 한 특성이다. 예를 들어, 설치류 또는 영장류 세포 내 하나 이상의 비천연 아미노산을 포함하는 단백질을 제조하기 위한 키트가 제공되는데, 여기에서 키트는 O-tRNA를 코딩하는 폴리뉴클레오타이드 서열, 및/또는 O-RS를 코딩하는 폴리뉴클레오타이드 서열, 및/또는 O-RS 폴리펩티드를 보유하기 위한 하나 이상의 용기를 포함한다. 한 구체예에서, 키트는 비천연 아미노산을 추가로 포함한다. 또 다른 구체예에서, 키트는 비천연 아미노산을 포함하는 단백질을 생성시키기 위한 지시 물질을 추가로 포함한다.

실시예

<317> 하기 실시예는 청구된 본 발명을 설명하기 위해 제공되는 것이나 본 발명을 한정하기 위해 제공된 것은 아니다. 당업자는 청구된 본 발명의 범위를 벗어나지 않으면서 변경될 수 있는 다양한 비임계적 변수를 인식할 것이다. 본원에 기재된 실시예 및 구체예는 단지 설명하기 위한 목적으로만 제공된 것이고, 이러한 취지의 각종 변형 또는 변화가 당업자에게 제안되고 이는 본원의 사상 및 취지와 첨부된 특허청구범위의 범주 내에 속함을 이해하도록 한다.

<318> 실시예 1

<319> 포유동물의 번역 성분에 의한 포유동물 세포에서 년센스 억제 기술 제한

<320> 포유동물 세포에서 비천연 아미노산을 유전적으로 코딩하기 위한 한 전략은 에쉐리키아 콜라이 또는 사카로미세스 세레비시에 숙주 시스템에서 발생한 존재하는 돌연변이체 tRNA/aaRS 쌍에 맞추는 것이다. 이.콜라이에서 수용점 줄기(acceptor stem)에 가변 암 및 G1:C72 염기쌍을 포함하는 tRNA^{Tyr} 동일성 성분이 포유동물 세포에서의 것과 완전히 별개이기 때문에(Wang 및 Schultz (2004), "Expanding the genetic code," *Angew Chem Int Ed Engl* **44**:34-66; 및 Bonnefond 등(2005), "Evolution of the tRNA(Tyr)/TyrRS aminoacylation systems," *Biochimie* **87**:873-883), 불리하게도 세균에서 사용된 tRNA/aaRS 쌍은 진핵 세포와 오르소고날성일 것 같지 않다. 대조적으로, 에스.세레비시에 및 포유동물 세포의 tRNA는 유사하게 처리되고 동일한 동일성 성분을 갖는다(Bonnefond 등(2005), "Evolution of the tRNA(Tyr)/TyrRS aminoacylation systems," *Biochimie* **87**:873-883). 또한, 에스.세레비시에의 번역적 기구가 고등의 진핵생물의 것과 상동하기 때문에, 에스.세레비시에에서 진화된 변형된 오르소고날 tRNA/aaRS 쌍을 포유동물 세포로 전달할 수 있는 것이 가능하다. 사실상, Yokoyama 및 동료들은 에스.세레비시에에서 진화된 EcTyrRS 변이체를 사용하여 p-벤조일-L-페닐알라닌(pBpa)을 허용함으로써 CHO 세포에서 광반응성 아미노산을 인간 Grb2 단백질로 도입시켰다(Hino 등(2005), "Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid," *Nat Methods* **2**, 201-206). 불리하게도, 작용적으로 활성인 이.콜라이 tRNA는 포유동물 세포에서 잘 발현되지 않아서, 돌연변이체 단백질의 수율을 심각하게 제한한다(Sakamoto 등(2002) "Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells," *Nucleic Acids Res* **30**:4692-4699).

<321> 실시예 2

<322> 포유동물 세포에서 신규 년센스 억제 시스템의 생성

<323> 상기 기재된 오르소고날 tRNA 발현시 제한을 극복하기 위해, 본 발명은 포유동물 세포에 사용될 효모에서 진화되는 돌연변이체 억제자 tRNA/aaRS 쌍을 전달할 수 있는 일반적 시스템을 제공한다. 상기 방법 및 신규 오르소고날 시스템은 설치류 CHO 숙주 세포 및 영장류(인간) 293T 숙주 세포 둘다에서 다수의 비천연 아미노산으로 녹색 형광 단백질(GFP)로 효율적으로 도입시키는 것을 입증한다. 이러한 2개의 세포 유형이 실험적으로 사용되고 있지만, 본 발명은 다양한 종 유래의 포유동물 숙주 세포에 널리 적용가능하다.

<324> 신규 포유동물 년센스 억제 시스템의 생성

<325> 진핵생물에서 tRNA는 2개의 보존된 유전자내 전사 제어 성분, 즉, A 및 B 박스를 인식하는 RNA 중합효소 III에 의해 전사된다(Sprague, *Transcription of Eukaryotic tRNA Genes*, AMS Press, Washington, DC; 1994). 이.콜라이 tRNA^{Tyr}만이 B 박스 성분을 갖고, EcTyrRS에 의해 인식되지 않는 비작용성 tRNA를 초래한다는 것을 제시한 바 있다(Sakamoto 등, "Site-specific incorporation of an unnatural amino acid into proteins in

mammalian cells," *Nucleic Acids Res* **30**:4692-4699 (2002)). 이.콜라이 tRNA^{Tyr}과 달리, 바실러스 스테아로씨 모필러스 유래의 tRNA^{Tyr} (유사한 동일성 성분을 갖고 여전히 EcTyrRS에 의해 충전됨; Bedouelle, "Recognition of tRNA(Tyr) by tyrosyl-tRNA synthetase," *Biochimie* **72**, 589-598 (1990))은 천연 발생 내부 A 및 B 박스를 갖는다. 따라서, EcTyrRS와 함께 상기 tRNA는 포유동물 세포에서 오르소고날 tRNA/aaRS 쌍으로 작용할 수 있다 (Sakamoto 등(2002), "Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells," *Nucleic Acids Res* **30**:4692-4699).

<326> 엠버 억제자 오르소고날 tRNA의 구축 및 발현

<327> 엠버 억제자 BstRNA^{Tyr}_{CUA}를 제공하기 위해, BstRNA^{Tyr}의 트리뉴클레오티드 안티코돈은 C(34)UA으로 충전되었다. 원핵생물 tRNA^{Tyr}의 G34가 유일한 TyrRS의 약한 동일성 성분이기 때문에(Hou 및 Schimmel (1989), "Modeling with in vitro kinetic parameters for the elaboration of transfer RNA identity in vivo," *Biochemistry* **28**:4942-4947), G34C 돌연변이체는 BstRNA^{Tyr}_{CUA}가 EcTyrRS에 결합하는 것에 유의적인 영향을 미치지 말아야 한다. 또한, 넨센스 엠버 억제는 포유동물 계놈(호모 사피엔스에서 TAG, 23%; TAA, 30%; TGA, 47%)에서 TAG 정지 코돈의 낮은 발생률로 인해 오판 또는 오우커(ochre) 억제보다 포유동물 세포에서의 내성이 더 양호해야 한다. 또한 진핵생물에서 BstRNA^{Tyr}_{CUA} 유전자의 억제가 5' 측접 서열에 의존하고 있기 때문에, 인간 tRNA^{Tyr}의 5' 측접 서열은 BstRNA^{Tyr}_{CUA}에 첨가되어 포유동물 세포에서 이의 전사를 향상시킨다. BstRNA^{Tyr}_{CUA}의 전사를 추가로 증가시키기 위해, BstRNA^{Tyr}_{CUA} 유전자의 3개의 직렬 반복을 함유하는 유전자 클러스터를 구축하고 이러한 유전자 클러스터를 pUC18 플라스미드에 삽입시켜 pUC18-3BstRNA^{Tyr}_{CUA}를 제공한다. pUC18-3BstRNA^{Tyr}_{CUA}로 형질감염된 CHO 세포로부터 단리된 총 tRNA의 노던 블롯 분석은 BstRNA^{Tyr}_{CUA}의 단 하나의 복사체를 함유하는 pUC18 플라스미드로 형질감염된 세포보다 BstRNA^{Tyr}_{CUA}의 2배 더 높은 농도를 제시한다.

<328> 오르소고날 합성효소의 발현

<329> 다음으로, 야생형 EcTyrRS 유전자를 포유동물 발현 벡터 pcDNA4/TO/myc-His A에 삽입하여 pcDNA4-EcTyrRS를 제공하는데, 여기에서 발현은 테트라사이클린(Tet)-조절된 CMV 프로모터에 의해 제어된다. 유도가능한 발현의 사용은 포유동물 세포에서 EcTyrRS의 이중성 발현으로 인해 가능하면 보다 낮은 독성을 의도하였다.

<330> 작용성 EcTyrRS/BstRNA^{Tyr}_{CUA} 오르소고날 쌍의 입증

<331> 모델 단백질 녹색 형광 단백질 GFP(GFP37TAG)의 위치 Y37에서 엠버 코돈 돌연변이를 효율적으로 억제하기 위해 생성된 억제자 BstRNA^{Tyr}_{CUA}/EcTyrRS 쌍의 능력은 포유동물 세포에서 분석되었다. Y37이 GFP의 표면에 위치하고 형광체의 말단에 위치하기 때문에, 상기 위치에서 비천연 아미노산의 도입은 단백질의 폴딩 및 형광성에 영향을 줄 것으로 기대되지 않는다(Ormo 등(1996), "Crystal structure of the Aequorea victoria green fluorescent protein," *Science* **273**:1392-1395). 또한, GFP37TAG의 엠버 억제가 전장 GFP의 발현을 초래하여, 엠버 억제 효능의 신속한 정량 분석을 제공한다. 돌연변이체 GFP37TAG 유전자는 pcDNA4/TO/myc-His A에 삽입되어 pcDNA4-GFP37TAG를 제공하는데, 여기에서 GFP는 C 말단에서 Tet-조절된 프로모터의 제어 하에 myc 및 6×His 에피토프에 융합된다.

<332> GFP37TAG의 넨센스 엠버 코돈을 억제하기 위한 BstRNA^{Tyr}_{CUA}/EcTyrRS 쌍의 능력은 Invitrogen™ T-REx™ 설치류 CHO 및 인간 293 세포 둘다에서 조사되었다. 세포주 둘다 테트라사이클린 억제자를 구조적으로 발현하고 pcDNA4/TO/myc-His A 플라스미드를 사용한 Tet-조절된 단백질 발현에 적당하다. 세포들은 80~90% 포화도로 성장하였고(Invitrogen™ T-REx™ CHO 세포는 10% FBS, 1% Pen-Strep, L-글루타민 2 mM, 및 블라스티시딘 10 μg/ml를 함유하는 F12 배지에서 성장시켰음; Invitrogen™ T-REx™ 293 세포를 10% FBS, 1% Pen-Strep, L-글루타민 2 mM, 및 블라스티시딘 5 μg/ml를 함유하는 DMEM에서 성장시켰음) Roche Applied Science에서 FuGENE® 6을 사용하여 2 × 10⁶ 세포 당 3 μg 플라스미드에 일시적으로 형질감염시켰다(pUC18-3BstRNA^{Tyr}_{CUA} 0.5 μg,

pcDNA4-EcTyrRS 0.5 μ g, 및 pcDNA4-GFP37TAG 2 μ g; pUC18-3BstRNA^{Tyr}_{CUA} 또는 pcDNA4-EcTyrRS는 이들이 존재하지 않는 경우 pUC18과 동일한 양으로 대체되었음). 감염 6시간 후 테트라사이클린 1 μ g/ml를 첨가함으로써 단백질 발현을 유도하고, 세포들을 2일 동안 37°C에서 성장시켰다. 두 세포주에서, 전장 GFP 발현은 BstRNA^{Tyr}_{CUA} 및 EcTyrRS 유전자 모두의 존재에 의존한다(도 13A 및 13B 참조). pcDNA4-GFP37TAG 및 pcDNA4-EcTyrRS 또는 pUC18-3BstRNA^{Tyr}_{CUA} 단독에 형질감염시키는 경우, 세포는 녹색 형광을 제시하지 않았다. 이러한 실험은 BstRNA^{Tyr}_{CUA}가 EcTyrRS에 의해서만 충전되며 BstRNA^{Tyr}_{CUA}/EcTyrRS 쌍이 효율적으로 작용하여 영장류 세포 및 설치류 세포에서 넌센스 엠버 코돈을 억제할 수 있음을 입증한다.

실시예 3

6개의 비천연 아미노산을 이용한 시스템 보편성의 입증

상기 오르소고날 시스템의 일반성을 측정하였다. 상기 테스트 시스템에서, Invitrogen™ T-REx™ CHO 및 293 세포주 둘다에서 각종 상응한 비천연 아미노산을 GFP37TAG에 도입시키기 위한 BstRNA^{Tyr}_{CUA} 및 6개의 EcTyrRS 변이체의 능력을 측정하였다. 6개의 EcTyrRS 변이체는 에스.세레비시에서 이전에 진화되어 하기 비천연 아미노산을 코딩하였다:

- p-메톡시-L-페닐알라닌 (pMpa)
- p-아세틸-L-페닐알라닌 (pApa)
- p-벤조일-L-페닐알라닌 (pBpa)
- p-요오도-L-페닐알라닌 (pIpa)
- p-아지도-L-페닐알라닌 (pAzpa)
- p-프로파르길옥시페닐알라닌 (pPpa)

이러한 비천연 아미노산의 구조는 도 1에 도시된다; 구조 1 내지 6. 합성효소 변이체의 선택 및 단리는, 예를 들어 문헌[Chin 등(2003), "An expanded eukaryotic genetic code," *Science* **301**:964-967; Deiters 등(2003), "Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*," *J Am Chem Soc* **125**:11782-11783; 제WO 2005/003294호(Deiters 등), "UNNATURAL REACTIVE AMINO ACID GENETIC CODE ADDITIONS,"(2004년 4월 16일 출원); 및 제WO 2006/034410호(Deiters 등), "ADDING PHOTOREGULATED AMINO ACID TO THE GENETIC CODE,"(2005년 9월 21일 출원)]에 기재된다.

6개의 EcTyrRS 유전자 각각은 pcDNA4/TO/myc-His A에 삽입되어 pcDNA4-EcTyrRS 유도체를 제공하고, T-REx™ CHO 및 293 세포에서 aaRS의 발현은 웨스턴 블롯 분석에 의해 증명되었다(도 14). 모든 6개의 EcTyrRS 변이체들은 T-REx™ CHO 및 T-REx™ 293 세포주 둘다에서 유사한 발현 수준을 제시한다.

이후 GFP37TAG에서 엠버 코돈을 억제하기 위한 BstRNA^{Tyr}_{CUA}와 함께 6개의 EcTyrRS 변이체의 능력은 성장 배지에서 비천연 아미노산의 존재 및 부재 하에 테스트되었다. 플라스미드를 이용한 세포의 일시적 형질감염은 야생형 EcTyrRS에서 기재한 바와 같이 수행되었다. 감염 6시간 후, 배지를 테트라사이클린 1 μ g/ml를 함유하는 새로운 배지로 대체하고 해당 비천연 아미노산으로 보충한다(pMpa, 10 mM; pApa, 10 mM; pBpa, 1 mM; pIpa, 8 mM; pAzpa, 5 mM; or pPpa, 1 mM). 이후 수확 전 T-REx™ CHO 세포주를 위해서는 24시간 동안, T-REx™ 293 세포주를 위해서는 48시간 동안 세포를 성장시켰다. T-REx™ CHO 및 293 세포 둘다에서, 전장 GFP37TAG의 발현은 도 2 및 도 3에 제시된 바와 같이 성장 배지에서 비천연 아미노산의 존재에 의존하였다. 비천연 아미노산의 부재 하에서는, GFP 발현이 검출되지 않았으며(< 1%), 이는 EcTyrRS 변이체가 높은 적합도로 BstRNA^{Tyr}_{CUA}를 이들의 동족체 비천연 아미노산으로 특이적으로 충전되었다. 아마도 pIpa의 세포 독성으로 인하여 T-REx™ CHO 세포에서는 pIpa를 함유하는 GFP가 낮은 발현 수준으로 관찰되었고, T-REx™ 293 세포에서는 pIpa를 함유하는 GFP 발현이 검출되지 않았다(8mM pIpa의 존재 하에, T-REx™ 293 세포는 6기간 내에 사망함).

실시예 4

<347>

p-메톡시-L-페닐알라닌을 도입하기 위한

<348>

오르소고날 성분의 발현에 대한 향상된 플라스미드 발현 시스템

<349>

p-메톡시-L-페닐알라닌-tRNA 합성효소(pMpaRS)의 이중성 발현이 사용된 성장 조건 하에 어떠한 명백한 세포 독성을 제시하지 않았기 때문에, pMpaRS 유전자를 함유하는 플라스미드를 변형시켜 (a) BstRNA^{Tyr}_{CUA} 유전자의 3개의 직렬 반복, 및 (b) pMpaRS 유전자(도 4A에서 pSWAN-pMpaRS)를 코딩하였다.

<350>

pMpaRS의 효율적이고 연속적인 발현을 위해 비조절된 CMV 프로모터 직후에 pMpaRS 유전자를 삽입하였다. BstRNA^{Tyr}_{CUA} 유전자의 3개의 직렬 반복 및 GFP37TAG 유전자 둘다를 함유하는 또다른 플라스미드(도 4B에서 pSWAN-GFP37TAG)를 또한 구축하였다. GFP37TAG를 코딩하는 유전자를 Tet-조절된 CMV 프로모터 후에 삽입하여 포유동물 세포에서 내인성 엠버 억제로 유발되는 넨센스 엠버 코돈의 잠재적인 리드-쓰루를 최소화하였다. 이후 억제 효율에 대해 이러한 2개의 플라스미드를 분석하였다. 플라스미드 모두 BstRNA^{Tyr}_{CUA} 유전자의 3개의 직렬 반복을 함유하기 때문에, 억제 수준을 증가시키기 위해 2개 플라스미드의 비율을 바꾸는 것은 세포에 형질감염된 BstRNA^{Tyr}_{CUA} 유전자의 총량을 변화시키는 것으로 기대하기 어렵다. 최적화된 형질감염 조건(2×10^6 세포 당 pSWAN-pMpaRS 0.5 μ g 및 pSWAN-GFP37TAG 2.5 μ g) 하에, 억제 수준은 3개의 플라스미드 pUC18-3BstRNA^{Tyr}_{CUA}, pcDNA4-pMpaRS, 및 pcDNA4-GFP37TAG를 이용한 것 보다 대략 2배 높다(억제 수준은 형과 세포의 수로 측정하였음). 유도 후 1~2일 동안 세포를 성장시키고 수확 전까지 생존시켰다. pMpa 함유 돌연변이체 GFP(GFP-pMpa)의 수율 또한 정량하였다. 성장 배지에서 10 mM pMpa의 존재 하에, T-RExTM CHO 세포 접착성 2×10^7 로부터 돌연변이체 GFP 1 μ g을 얻을 수 있다.

<351>

실시예 5

<352>

질량 분광기 및 적합도 분석

<353>

GFP-pMpa를 추가로 특징화하기 위해, pSWAN-pMpaRS 및 pSWAN-GFP37TAG 둘다에 일시적으로 형질감염된 T-RExTM CHO 세포에서 돌연변이체 단백질을 발현시켰다. 항-myc 항체 아가로스 컬럼을 이용하여 단백질을 정제한 후 SDS-PAGE로 분리시켰다. SDS-PAGE 겔로부터의 GFP-pMpa 밴드를 트립신으로 분해시키고 나노스케일 역상 액체 크로마토그래피/질량 분석/질량 분석(나노-RP LC/MS/MS)으로 분석하였다. 동시에, pcDNA4-GFP로 형질감염된 세포로부터 야생형 GFP를 정제하고 동일한 분석을 실시하였다.

<354>

돌연변이체 단백질로부터 FSVSGEGEGDATY^{*}GK(Y^{*}는 티로신 또는 pMpa를 가리킴)(도 5 및 서열 번호 103 참조) 단편을 함유하는 트립신 Y37의 직렬 질량 스펙트럼은 강력한 pMpa 피크를 제시하여, pMpa의 효율적 도입을 나타낸다(도 6 및 도 7 참조). Y^{*} 함유 이온(y₃ 내지 y₁₄) 모두는 야생형 GFP와 비교하였을 때 14 Da의 질량 이동을 가지고, 이는 티로신과 pMpa 사이의 질량 차이점과 정확하게 일치한다. 도 7에서 b13 이온의 동일한 질량 이동의 관찰과 함께 y 이온 시리즈의 질량 이동은 GFP의 위치 37로 pMpa 도입 부위를 명확하게 배정한다.

<355>

또한 티로신 함유 펩티드의 피크에 대한 pMpa 함유 펩티드의 MS 피크의 비율을 얻었다. FSVSGEGEGDATY^{*}GK (도 5 및 서열 번호 103 참조)의 전구체 이온의 단일 이온 크로마토그램의 통합은 pMpa의 도입에 대해 높은 적합도를 시사한다(> 99.9%). 직렬 MS 데이터에서 3개의 가장 풍부한 단편 이온(y₉, y₁₁, 및 y₁₂)을 모니터링하는 것을 기초로 하는 추정치는 균일하게 보다 우수한 순도(99.93%)를 시사한다. 모 단백질의 질량을 획득하기 위해, 항-myc 항체 컬럼에 의해 정제된 돌연변이체 단백질에 ESI TOF-MS를 이용한 분석을 실시하였다. N 말단 메티오닌이 결손된 아세틸화된 돌연변이체 단백질의 이론적 질량은 29,696.00 Da이고, 이는 도 8에서 하전 상태 탈소용돌이화된 ESI TOF-MS 스펙트럼에서 관찰되는 29,696.0 Da의 주요 성분과 탁월하게 일치한다. 29,654.0 Da에서의 더 작은 피치는 N 말단 아세틸화가 결핍된 GFP-pMpa의 질량에 배정된다. 야생형 GFP 신호 또는 다중 pMpas의 도입을 가리키는 신호가 검출되지 않았다. 이러한 결과는 GFP의 위치 37에서 pMpa의 선택적 도입을 추가로 확인한다.

<356>

실시예 6

<357>

pApa, pBpa, pIpa, pAzpa 및 pPpa를 도입하는

<358>

오르소고날 성분들의 발현을 위한 향상된 플라스미드 발현 시스템

- <359> 플라스미드는 다양한 합성효소 변이체를 코딩하는 유전자를 하위클로닝함으로써 다른 pSWAN-EcTyrRS 변이체 2개 플라스미드 시스템을 제공하도록 재조작되었다. 이러한 합성효소 변이체는 다음과 같다:
- <360> p-아세틸-L-페닐알라닌-tRNA 합성효소 (pApaRS)
- <361> p-벤조일-L-페닐알라닌-tRNA 합성효소 (pBpaRS)
- <362> p-요오도-L-페닐알라닌-tRNA 합성효소 (pIpaRS)
- <363> p-아지도-L-페닐알라닌-tRNA 합성효소 (pAzpaRS)
- <364> p-프로파르길옥시페닐알라닌-tRNA 합성효소 (pPpaRS)
- <365> pSWAN-EcTyrRS 변이체 및 pSWAN-GFP37TAG (2×10^6 세포 당 pSWAN-EcTyrRS 변이체 0.5 μ g 및 pSWAN-GFP37TAG 2.5 μ g)를 이용한 T-RExTM CHO 및 293 세포의 일시적 형질감염은 모두 (pIpa 특이적 변이체의 예상을 가지고) 3개의 플라스미드 pUC18-3BstRNA^{Tyr}_{CUA}, pcDNA4-EcTyrRS 변이체, 및 pcDNA4-GFP37TAG보다 대략 2배 높은 억제 수준을 제공한다. 또한 pApa, pBpa, pAzpa 또는 pPpa 함유 돌연변이체 GFP는 pSWAN-EcTyrRS 변이체 및 pSWAN-GFP37TAG로 형질감염된 T-RExTM CHO 세포에서 발현되고 1~2일 동안 상응한 비천연 아미노산(pApa, 10 mM; pBpa, 1 mM; pAzpa, 5 mM; 또는 pPpa, 1 mM)을 함유한 매질에서 성장시키고, 항-myc 항체 컬럼을 이용하여 정제하였다. pApa 및 pAzpa를 함유하는 돌연변이체 단백질의 수율은 pMpa에 대한 것(2×10^7 세포 당 $\sim 1 \mu$ g)과 근접하였고, pBpa 및 pPpa 함유 돌연변이 단백질의 수율은 다소 낮았다(2×10^7 세포 당 $\sim 0.7 \mu$ g). T-RExTM CHO 세포에서 EcTyrRS 변이체의 발현 수준이 유사하기 때문에, pBpa 및 pPpa를 함유하는 돌연변이체 단백질의 더 낮은 수율은 배지에서 비천연 아미노산의 낮은 농도로 인한 것일 수 있다.
- <366> 또한 FSVSGEGEGDATY^{*}GK (도 5 및 서열 번호 103)의 트립신 단편의 직렬 MS를 얻기 위해 나노-RP LC/MS/MS를 이용하여 돌연변이체 GFP 단백질에 분석을 실시하였다.
- <367> 직렬 MS 데이터(도 9~12)는 GFP의 위치 37에서 비천연 아미노산의 도입을 명백하게 제시한다. pAzpa 함유 돌연변이체 GFP는 약하게만 검출가능하였다. 데이터에 근접한 관찰은 p-아미노페닐알라닌(pAmpa) 함유 펩티드의 존재 대신 밝혀졌고(도 15), 이는 아지도 기의 화학적 반응성 및 광-불안정성을 고려하는 것이 놀라운 일이 아니다(pAmpa는 pAzpa 함유 펩티드의 MS 분석에서 이전에 관찰된 바 있음; Chin 등(2003), "Progress toward an expanded eukaryotic genetic code," *Chem Biol* 10, 511-519). pBpa 샘플에서 미량의 야생형 펩티드를 확인하였지만, pBpa/티로신 비율의 정확한 정량을 위해서는 신호가 너무 약하였다. pApa, pAzpa, 및 pPpa 샘플에서 야생형 신호가 검출되지 않았다. 모든 5개의 돌연변이체 단백질로부터의 데이터는 비천연 아미노산의 도입의 높은 선택성과 적합도를 가리킨다.

<368> 실시예 7

<369> 논의/결론

- <370> 포유동물 계통에서, 옴버 정지 코돈의 발생률은, 이.콜라이(7%)와 비교하였을 때 더 높다(인간에서 23%). 따라서, 주요 단백질이 올바르게 종료되지 않는 경우 옴버 억제가 세포에 독성이 될 수 있다. Yokoyama 및 동료들은, 동족체 tRNA를 3-요오도-L-티로신으로 충전하는 돌연변이체 EcTyrRS의 유도가능한 발현이 3-요오도-L-티로신의 부재 하에 돌연변이체 aaRS에 의해 내인성 티로신의 배경 도입으로부터 초래되는 가능한 세포 독성을 최소화시키는 것을 제시하였다. 본원에 기재된 본 발명은 이러한 문제에 대한 해답을 제공한다. 모든 EcTyrRS 변이체들은, 2단계의 양성/음성 선택 계획을 이용하여 에스.세레비시에서 이전에 진화되는데, 이는 내인성 아미노산을 도입하는 aaRS 변이체를 제거한다. pSWAN-EcTyrRS 변이체 및 pSWAN-GFP37TAG를 이용한 세포의 공형질감염은 비천연 아미노산의 부재 하에 GFP37TAG에서 넌센스 옴버 코돈의 관찰가능한 리드-쓰루를 초래하지 않으며, 이는 tRNA/aaRS 쌍이 비천연 아미노산의 부재 하에 천연 TAG 정지 코돈을 효율적으로 억제하지 않는다는 것을 시사한다. 따라서 tRNA 및 aaRS 단백질이 안정하게 발현하는 세포주는 비천연 아미노산의 부재 하에 생존해야 한다. 관찰되는 내인성 옴버 억제가 없기 때문에, 비천연 아미노산을 함유하는 표적 단백질의 발현 또한 유도가 필요하지 않다. tRNA, aaRS, 및 표적 단백질 유전자를 유지하는 안정한 세포주의 생성은, 세포가 비천연 아미노산으로 보충되는 경우 비천연 아미노산을 함유하는 표적 단백질의 효율적인 생산을 허용할 것이다. 최근 기재된 원리의 증명 실험에서, 비천연 아미노산의 첨가 3일 후 숙주 포유동물 세포가 생존하였고, 이는 재조합 단백질 발현에 충분한 것으로 보인다.

<371> 이러한 연구는 비천연 아미노산, 1,5-다니슬라닌, o-니트로벤질시스테인 및 α-아미노카프릴산을 비롯한 다른 비천연 아미노산에 성공적으로 확장되었다. 본 발명은 본원에 기재된 세포주 또는 비천연 아미노산으로 한정하려는 의도가 아니다.

<372> 실시예 8

<373> 포유동물 세포에서 광조절된 단백질 서열

<374> 청구된 번역 시스템 및 방법의 일부 구체예에서, 비천연 아미노산은 광조절된 아미노산, 예컨대 o-니트로벤질세린, o-니트로벤질시스테인, α-아미노카프릴산 등이다. 광조절된 아미노산(예, 광색성, 광분해성, 광이성질화 등)을 이용하여, 예를 들어 효소, 수용체, 이온 채널 등의 활성을 직접 조절함으로써, 또는 다양한 신호 분자의 세포내 농도를 조절함으로써 공간적으로, 일시적으로 각종 생물학적 과정을 제어할 수 있다. 예를 들어, 문헌 [Shigeri 등, *Pharmacol. Therapeut.*, 2001, 91:85+; Curley, 등, *Pharmacol. Therapeut.*, 1999, 82:347+; Curley, 등, *Curr. Op. Chem. Bio.*, 1999, 3:84+; "Caged Compounds" *Methods in Enzymology*, Marriott, G., Ed, Academic Press, NY, 1998, V. 291; Adams, 등, *Annu. Rev. Physiol.*, 1993, 55:755+; 및 Bochet, 등, *J. Chem. Soc., Perkin 1*, 2002, 125+] 참조.

<375> 이.콜라이 tRNA^{Leu}에서 유도된 엠버 억제자 tRNA의 전사 수준을 증가시키기 위해, 인간 개시자 tRNA^{Met}의 5'-측접 서열 및 3'-측접 서열의 제어 하에 억제자 tRNA를 넣었다. o-니트로벤질시스테인에 특이적인 돌연변이체 이.콜라이 류실-tRNA 합성효소와 함께 과전사된 엠버 억제자 tRNA를, 중국 햄스터 난소 세포 및 인간 293T 세포에서 상기 비천연 아미노산을 인간 프로카스페이스-3 활성 부위에 도입시켰다. o-니트로벤질시스테인의 존재 하에 단백질의 발현 후 상기 세포에 UV 광을 5분 동안 노출시켜 세포 사멸을 촉발시키는 것은, UV 광이 비천연 시스템 잔기로부터의 니트로벤질 기의 분해를 유도하고 세포 사멸을 촉발하는 프로카스페이스-3의 활성을 회복시킨다는 것을 제시한다.

<376> 이러한 실험은, 생존 포유동물 세포 내 단백질에서 o-니트로벤질시스테인으로 활성 부위 시스템의 선택적 치환을 이용하여, 예를 들어 UV 광을 이용함으로써 단백질을 공간적으로 제어할 수 있음을 확인한다. 반응성 부위가 상이한 단백질에 대한 유사한 결과를 얻을 수 있다(예를 들어, 세린 프로테아제의 활성 부위 잔기 내 o-니트로벤질세린의 이용). 광케이지되거나 광조절된 아미노산과 관련된 추가의 상세한 사항은 Dieters 등에 의한 문헌[PCT 공개 제WO 2006/034410호(2006년 3월 30일 출원)(발명의 명칭 "Adding Photoregulated Amino Acids to the Genetic Code")]에서 찾을 수 있으며, 이는 본원에 그 전문이 참고인용된다.

<377> 실시예 9

<378> 재료 및 방법: 포유동물 세포 형질감염 및 웨스턴 블롯 분석

<379> T-REXTM CHO 및 T-REXTM 293 세포(InvitrogenTM) 둘다 테트라사이클린 억제자를 구조적으로 발현하고, 이는 pcDNA4/TO/myc-His A 플라스미드로 삽입된 유전자의 발현을 조절한다. T-REXTM CHO 세포는 37°C에서 5% CO₂의 습한 분위기 하에 F-12 (InvitrogenTM), 10% FBS (InvitrogenTM), 1% Pen-Strep (InvitrogenTM), 2 mM L-글루타민 (InvitrogenTM), 및 10 μg/ml 블라스티시딘 (InvitrogenTM) 중에 성장시키고; T-REXTM 293 세포는 37°C에서 5% CO₂의 습한 분위기 하에 Gibco D-MEM 배지 (InvitrogenTM), 10% FBS, 1% Pen-Strep, 2 mM L-글루타민, 및 5 μg/ml 블라스티시딘 중에 성장시킨다. Costar[®] 6 웰 세포 배양 클러스터에서 80~90% 합류성으로 세포를 성장시킨 후 FuGENE[®] 6 (Roche Applied Science)(9 μl FuGENE[®] + 3 μg 플라스미드)을 이용하여 플라스미드에 형질감염시켰다. 형질감염 6시간 후, 테트라사이클린 1 μg/ml를 함유하는 새로운 배지로 배지를 대체하였다.

<380> 해당 비천연 아미노산의 존재 하에 EcTyrRS 변이체의 엠버 억제를 테스트하기 위해, 또한 비천연 아미노산을 새로운 배지에 첨가하였다. 배지에서 비천연 아미노산의 농도는 pMpa 10 mM(Bachem, Inc), pApa 10 mM(Synchem, Inc); pBpa 1 mM(Bachem, Inc), pIpa 8 mM(Bachem, Inc), pAzpa 5 mM(Bachem, Inc), 및 pPpa 1 mM이다. 키랄 순수 pPpa는 이전에 기재된 바와 같이 합성되었다(Deiters 등(2003), "Adding amino acid with novel reactivity to the genetic code of *Saccharomyces cerevisiae*," *J Am Chem Soc* **125**:11782-11783). 테트라사이클린 첨가 후 24시간 동안 T-REXTM CHO 세포를 성장시킨 후 수확하였고; 48시간 항온처리 후 T-REXTM 293 세포를 수확하였다.

<381> 웨스턴 블롯 분석을 위해, 프로테아제 억제자 콕테일(Sigma)의 1:100 희석으로 RIPA 완충액(Upstate

Biotechnology/Millipore) 중에 수확된 세포를 용해시켰다. 변성된 조건 하에 SDS-PAGE로 상청액을 분할하고 0.45 μ m 니트로셀룰로스 막(Invitrogen™)으로 분주하였다. T-REx™ CHO 세포를 위해, 막 위에 고정된 단백질은 1차 항체로서 항-myc 항체(Invitrogen™; 1:5000 희석) 및 2차 항체로서 항 마우스 IgG-HRP(Invitrogen™)를 이용하여 조사하였다. 이후 화학발광은 PIERCE ECL 웨스턴 블로팅 기질을 이용하여 검출되었다. T-REx™ 293 세포를 위해, 막은 항-His-HRP(Invitrogen™; 1:5000 희석)로 조사된 후 PIERCE ECL 웨스턴 블로팅 기질을 이용하여 검출되었다.

<382>

재료 및 방법: 세균 세포 형질감염

<383>

플라스미드를 클로닝, 유지, 및 증폭시키는데 탐 10개의 이.콜라이 세포(Invitrogen™)를 사용하였다. Pfx 높은 적합도의 DNA 중합효소(Invitrogen™)를 중합 연쇄 반응(PCR)에 사용하였다. 형질감염 시약으로 FuGENE® 6(Roche Applied Science)을 사용하였다. 모든 플라스미드를 시퀀싱으로 증명하였다.

<384>

실시예 10

<385>

재료 및 방법: 플라스미드 구축물

<386>

pUC18-3BstRNA^{Tyr}_{CUA}

<387>

4개의 올리고데옥시뉴클레오타이드(Integrated DNA Technologies, Inc)를 어닐링함으로써 BstRNA^{Tyr}_{CUA} 유전자를 구축하였다. 상기 유전자는 인간 tRNA^{Tyr} 유전자 AGCGCTCCGGTTTTCTGTGCTGAACCTCAGGGGACGCCGACACGTACACGTC(서열 번호 104)의 3'-CCA 및 5'-측접 서열이 결합된 상응한 tRNA 서열로 이루어진다.

<388>

2개의 제한효소 부위 - 5' 말단에서의 EcoRI 및 3' 말단에서의 BamHI를 합성 DNA 중복부위(duplex)에 도입시킨 후, pUC18에 삽입하여 pUC18-BstRNA^{Tyr}_{CUA}를 제공하였다. 이후 pUC-BstRNA^{Tyr}_{CUA}에서 BstRNA^{Tyr}_{CUA} 유전자를 PCR로 증폭시켰다. 5'-말단에서 BglII 제한효소 부위 및 3' 말단에서 BamHI 부위를 함유하는 증폭된 BstRNA^{Tyr}_{CUA}를 pUC18-BstRNA^{Tyr}_{CUA}의 BamHI 제한효소 부위에 삽입하여 pUC18-2BstRNA^{Tyr}_{CUA}를 제공하였다. 플라스미드 pUC18-2BstRNA^{Tyr}_{CUA}는 BstRNA^{Tyr}_{CUA}의 또다른 복사체를 도입시키는데 사용했던 하나의 BamHI 제한효소 부위만을 포함하여 pUC18-3BstRNA^{Tyr}_{CUA}를 제공한다. EcTyrRS 및 이의 6개의 변이체들은 모두 pEcTyrRS/tRNA_{CUA}로부터 PCR로 증폭된 후(Chin 등(2003), "An expanded eukaryotic genetic code," *Science* **301**:964-967) pcDNA4/TO/myc-His A 플라스미드(Invitrogen™)의 XbaI 및 ApaI 부위에 삽입하여 pcDNA4-EcTyrRS 벡터를 제공하였다. 2개의 테트라사이클린 조작자 서열 및 CMV 프로모터 이후에 EcTyrRS 유전자를 위치시키는데, 이는 Invitrogen™ T-REx™ 세포주에서 테트라사이클린 조절된 유전자 발현을 부여한다. 이들은 또한 웨스턴 블롯 분석을 위해 c-myc 에피토프에 연결되었다. 6개의 EcTyrRS 변이체에서 돌연변이는 다음과 같다:

<389>

pMpaRS: Y37V/D182S/F183M

<390>

pApaRS: Y37I/D182G/F183M/L186A

<391>

pBpaRS: Y37G/D182G/F183Y/L186M

<392>

pIpaRS: Y37I/D183S/F183M

<393>

pAzpaRS: Y37L/D182S/F183M/L186A

<394>

pPpaRS: Y37S/D182T/F183M/L186V

<395>

pcDNA4-EcTyrRS 및 pcDNA4-GFP37TAG

<396>

향상된 야생형 GFP 유전자를 벡터 pcDNA4/TO/myc-His A의 XbaI 및 ApaI 부위에 삽입함으로써 pcDNA4-GFP를 생성하였다. pcDNA4-GFP에서, 발현된 단백질의 웨스턴 블롯 분석 및 친화력 정제를 위해 C 말단에서 myc 에피토프 및 6 × His 태그를 이용하여 GFP 유전자를 결합시켰다. Y37의 3중 코돈이 TAG 옴버 코돈에 돌연변이를 시키는 플라스미드 pcDNA4-GFP37TAG는 QuikChange® 부위-지정 돌연변이유발 키트(Stratagene)에 의해 발생되었다.

<397>

pSWAN-EcTyrRS 변이체 및 pSWAN-GFP37TAG

<398> pSWAN-EcTyrRS 변이체를 구축하기 위해, BamHI 및 EcoRI 제한효소를 이용하여 pUC18-3BstRNA^{Tyr}_{CUA}로부터 3BstRNA^{Tyr}_{CUA}를 단리시키고 pcDNA3.1/hygro(+)(Invitrogen™)의 BglII 및 MfeI 부위에 삽입시켜 pcDNA-3BstRNA^{Tyr}_{CUA}를 제공하였다. 6개의 EcTyrRS 변이체를 코딩하는 유전자를 pcDNA-3BstRNA^{Tyr}_{CUA}의 XbaI 및 ApaI 부위에 삽입시켜 pSWAN-EcTyrRS 변이체를 제공하였다. pSWAN-GFP37TAG를 생성하기 위해, pSWAN-EcTyrRS 벡터를 증폭시키고 MluI 및 PciI 제한 효소로 분해하였다. 분해된 단편을 아가로스 겔 전기영동으로 분리시키고 3BstRNA^{Tyr}_{CUA} 함유 단편을 QIAquick[®] 겔 정제 키트(QIAGEN[®])로 정제시켰다. CMV 프로모터에서 BGH polyA 부위의 pcDNA4-GFP37TAG 단편을 PCR로 증폭시키고, MluI 및 PciI 제한 효소로 분해한 후, 3BstRNA^{Tyr}_{CUA} 함유 pSWAN-EcTyrRS 단편으로 결합시켜 pSWAN-GFP37TAG를 제공하였다(∼4000 bp).

<399> **실시예 11**

<400> **재료 및 방법: 단백질 발현 및 정제**

<401> pSWAN-EcTyrRS 변이체 및 pSWAN-GFP37TAG를 사용하여 비천연 아미노산 함유 돌연변이체 GFP를 발현하였다. T-REx™ CHO 세포(Invitrogen™)를 75 cm² 조직 배양 플라스크(BD Biosciences)에서 80∼90%의 포화도로 성장시키고 60 μl의 FuGENE[®] 6을 이용하여 pSWAN-EcTyrRS 2 μg 및 pSWAN-GFP37TAG 10 μg으로 형질감염시켰다. 밤새 항온처리 후, 비천연 아미노산 및 1 μg/ml 테트라사이클린을 함유하는 새로운 배지(pMpa, 10 mM; pApa, 10 mM; pBpa, 1 mM; pAzpa, 5 mM; and pPpa, 1 mM)로 배지를 교환하였다. 이후 수확 전에 37℃에서 1∼2일 동안 세포를 성장시킨 후, RIPA 용해 완충액으로 용해시켰다. 항-myc 항체 아가로스 컬럼(Sigma) 상에 로딩하기 전에 PBS 완충액으로 세포 용해물로부터의 상청액을 투석하고 평형시켰다. PBS 완충액 15 컬럼 부피로 상기 컬럼을 세척한 후 0.1 M 수산화암모늄으로 용출하였다. 0.1 M 아세트산으로 정제된 단백질을 중화시키고 SDS-PAGE로 분석하였다. 질량 분광적 분석을 위해 GFP에 상응한 밴드를 절취하였다. 수율을 측정하고 비손상 단백질 질량 스펙트럼을 획득하기 위해, 용출 후 ∼0.1 mg/ml로 단백질을 농축시켰다.

<402> **실시예 12**

<403> **재료 및 방법: 단백질 가수분해**

<404> 친화력 정제된 단백질(야생형 및 pBpa 샘플을 위해 수행됨)의 단백질 가수분해는 100 mM 트리에틸암모늄 중탄산염 완충액(pH 8.5)에서 추정된 단백질 농도를 기초로 10:1(중량/중량)의 기질/효소 비율을 이용하여 시퀀싱-등급 변형된 트립신(Promega)과 함께 37℃에서 단백질을 밤새 항온처리함으로써 수행되었다. 인-겔 분해(pMpa, pAzpa, pApa, 및 pPpa 샘플을 위해 수행됨)는 변형된 EMBL 절차에 따라 수행되었다(EMBL의 웹사이트 참조; 독일 하이델베르크 소재; Bioanalytical Research Group). 간단히 말해서, 겔 슬라이스를 대략 1 mm × 1 mm의 사면체로 절취하고, 물로 5분 동안 세척하고, 아세토니트릴로 15분 동안 세척하고, 진공 원심분리기로 건조시킨 후 50 mM 트리에틸암모늄 중탄산염 완충액 (pH 8.5) 중에서 시퀀싱-등급 변형된 트립신 12.5 ng/μl로 30분 동안 0℃에서 재수화시켰다. 과잉 완충액을 제거하고 50 mM 트리에틸암모늄 중탄산염 완충액(pH 8.5) 20 μl를 겔 슬라이스에 첨가한 후 밤새 37℃에서 항온처리하였다. 프로토콜에 기재된 바와 같이 겔 슬라이스로부터 펩티드를 추출하였다.

<405> **실시예 13**

<406> **재료 및 방법: 나노-RP LC/MS/MS**

<407> HPLC 펌프, 자동채수기(Agilent Technologies; 캘리포니아주 팔로 알토 소재), 및 LTQ Orbitrap 하이브리드 질량 분광계(ThermoElectron, 캘리포니아주 산 호세 소재)를 이용하여 나노-RP LC/MS/MS를 수행하였다. 대략 2 μl/분의 유속에서 4 cm의 5-μm 모니터 C18 입자(Column Engineering; 캘리포니아 온타리오 소재)로 패키징된 75 μm 내경 전치컬럼 상에 압력탄을 이용하여 트립신 분해물을 로딩하였다. 이후 전치 컬럼을 HPLC 펌프에 연결시키고 용매 A로 수분 동안 세척 후 집적 이미터 팁이 장착된 분석용 컬럼(360 μm O.D. x 75 μm 내경; 5 μm C18의 10 cm, ∼5-μm 팁)을 일렬로 연결하였다. 크로마토그래피 프로파일은 40분에 100% 용매 A (0.1% 수성 아세트산)에서 50% 용매 B (아세토니트릴 중 0.1% 아세트산)였고; 분석용 컬럼을 통한 유속은 약 100 nl/분이었다.

<408> 데이터 의존 실험의 경우, 질량 분광계를 프로그램화하여 전체 스캔 이온 트랩 스펙트럼보다 고해상도 오르비트랩 스캔을 첫번째로 기록하였다(m/z 500-2,000). 이는 이온 트랩 스캔 및 최종적으로 2개의 표적된 MS/MS 스캔

의 측발된 10 데이타 의존 MS/MS 스캔(상대 충돌 에너지 = 35%; 3-Da 분리 윈도우)을 따른다. 제1 표적된 MS/MS 스캔을 설정하여 예측된 비천연 아미노산 변형된 펩티드(FSVSGEGEGDATY^{*}GK; 서열 번호 103)의 이중 하전된 전구체 이온을 단리시키고 단편화하였다.

<409> 제2 표적된 MS/MS 스캔은 항상 야생형 펩티드(FSVSGEGEGDATY^{*}GK; 서열 번호 102)의 이중 하전된 전구체 이온을 m/z 752.3에서 단리시키고 단편화하였다. 단백질 동정 및 다양한 변형으로 비천연 아미노산을 이용한 표적 펩티드를 함유하는 스캔을 찾기 위해 MASCOT (Matrixscience, 영국 런던 소재)를 이용하여 MSDB 데이터베이스에 대한 자료를 조사하였다. GFP-pMpa의 Y^{*}37에서 적합도를 추정하기 위해, 단일 이온 크로마토그램(야생형에 대해 m/z 752.3 및 pMpa에 대해 m/z 759.3에서 이중 하전된 전구체 이온) 및 선택된 이온 크로마토그램(y₉, y₁₁, 및 y₁₂)을 Xcalibur 소프트웨어 패키지(ThermoElectron, 캘리포니아 산 호세 소재)를 이용하여 집적되었다.

<410> 자동채수기 및 QTOF2 질량 분광계가 장착된 측정관 LC로 이루어진 자동화된 LC/MS 시스템(Waters, 메사추세츠주 밀포드 소재) 상에서 비손상 단백질 질량 스펙트럼을 획득하였다. 수중 0.1% 아세트산으로 탈염하기 위해 역상 단백질 캡트랩(Michrom Bioresources, 캘리포니아주 어번 소재) 상에 GFP-pMpa (0.1 mg/ml)를 로딩하고 질량 분광계의 ESI 공급원으로 5 μl/분에서 80% 아세토니트릴/0.1% 아세트산으로 용출하였다. MassLynx(Waters, 메사추세츠주 밀포드 소재) 소프트웨어 패키지를 이용하여 MaxEnt1 알고리즘으로 스펙트럼의 합계, 평탄화 및 탈소용돌이화를 수행하였다.

<411> **실시예 14**

<412> **포유동물 전사를 위한 프로모터**

<413> 포유동물 세포에서 tRNA 전사를 위한 강한 프로모터는 이.콜라이 억제자 tRNA^{Leu5}(CUA)에 대한 5' 측접 서열 GATCCGACCGTGTGC TTGGCAGAAC (서열 번호 105) 및 이.콜라이 억제자 tRNA^{Leu5}(CUA)에 대한 3' 측접 서열 GTCCTTTTITG (서열 번호 106)를 포함한다.

<414> * * *

<415> 상기와 같은 본 발명은 보다 명료하게 하고 이해를 돕기 위한 목적으로 다소 상세히 기재되었으나, 본 발명의 진정한 범주를 벗어나지 않는 한, 형태 및 세부사항에서의 다양한 변화가 가해질 수 있음이 본 개시내용을 읽은 당업자에게 명백할 것이다. 모든 공보, 특허, 특허출원, 및/또는 본원에서 인용된 기타 문헌은 마치 개별 공보, 특허, 특허출원 및/또는 기타 문헌이 모든 목적을 위해 참조 인용되는 것으로 개별적으로 나타내어지는 것과 같은 동일한 정도로 모든 목적을 위해 완전히 참조 인용된다.

도면의 간단한 설명

<300> **도 1**은 비천연 아미노산 p-메톡시페닐알라닌(pMpa), p-아세틸페닐알라닌(pApa), p-벤조일페닐알라닌(pBpa), p-요오도페닐알라닌(pIpa), p-아지도페닐알라닌(pAzpa), p-프로파르길옥시페닐알라닌(pPpa), α-아미노카프릴산, o-니트로벤질시스테인(o-NBC), 1,5-단실알라닌 및 o-니트로벤질세린(o-NBS)의 화학적 구조 및 명칭을 도시한다. 이러한 비천연 아미노산 중 대부분은 다양한 대안적 명칭을 이용하여 지칭되었고, 이 또한 도면에 제시한다.

<301> **도 2**는 오르소고날 BstRNA^{tyr}_{CUA} 및 EcTyrRS 쌍을 이용하여 Invitrogen™ T-REx™ CHO 세포에서 전장 GFP 발현의 웨스턴 블롯 분석을 도시한다. 플라스미드 pcDNA4-GFP37TAG, pcDNA4-EcTyrRS 변이체 및 pUC18-3BstRNA^{tyr}_{CUA}를 이용하여 상기 세포를 공형질감염시키고 비천연 아미노산의 존재 또는 부재 하에 성장시킨다. 제1 라인은 일시적으로 pcDNA4-GFP로 형질감염되었던 T-rex CHO 세포에서의 야생형 GFP 발현이다. 제2 라인은 BstRNA^{tyr}_{CUA}와 함께 야생형 EcTyrRS에 의해 억제된 GFP37TAG의 전장 발현이다. 그 다음 라인은 비천연 아미노산의 존재 또는 부재 하에 BstRNA^{tyr}_{CUA}와 함께 EcTyrRS 변이체에 의해 억제된 GFP37TAG의 전장 발현이다. pMpaRS는 pMpa-tRNA 합성효소 등을 나타낸다. 각 반응을 위한 세포 용해물 40 μg 분액(10 μg 세포 용해물을 제1 라인에 로딩하였음)을 항-c-myc 항체로 분석하였다. 단백질 발현을 위한 비천연 아미노산의 농도는 pMpa 10 mM, pApa 10 mM, pBpa 1 mM, pIpa 8 mM, pAzpa 5 mM 및 pPpa 1 mM이었다.

<302> **도 3**은 **도 2**에 기재된 바와 같이 오르소고날 BstRNA^{tyr}_{CUA} 및 EcTyrRS 쌍을 이용하여 Invitrogen™ T-REx™ 293

세포에서 전장 GFP 발현의 웨스턴 블롯 분석을 도시한다. 각 반응을 위한 세포 용해물 20 μ g 분액(대조군 반응 5 μ g)을 항-His-HRP 항체로 분석하였다. 시약, 비천연 아미노산의 농도 및 라인 명칭은 도 2에 기재된 바와 동일하였다.

- <303> 도 4A는 pSWAN-pMpaRS의 플라스미드 지도를 도시한다.
- <304> 도 4B는 pSWAN-GFP37TAG의 플라스미드 지도를 도시한다.
- <305> 도 5는 야생형 GFP 및 GFP-pMpa 펩티드의 아미노산 서열을 도시한다. Y*는 야생형 GFP 내 티로신 또는 GFP-pMpa 내 pMpa를 의미한다. 펩티드 FSVSGEGEGDATY*GK(서열 번호 103)의 단편화 동안 발생된 y형 및 b형 이온을 도시한다.
- <306> 도 6은 도 5에 제공된 야생형 GFP 펩티드 FSVSGEGEGDATY*GK(서열 번호 103)의 주석이 달린 직렬 MS 스펙트럼을 도시한다. 비교 및 명패함을 위해, 풍부한 Y*-이온 시리즈만이 위치 37에서 명백하게 pMpa에 위치하는 b₁₃ 이온과 마찬가지로 주석이 달렸다. 항-myc 친화성 컬럼으로 단백질을 정제하고, SDS-PAGE로 분석하였다. MS 분석에 대한 GFP 밴드를 절취하였다.
- <307> 도 7은 GFP-pMpa에서 유도된 FSVSGEGEGDATY*GK(서열 번호 103) 펩티드의 주석이 달린 직렬 MS 스펙트럼을 도시한다. Y*-이온 시리즈의 직렬 MS 스펙트럼은 14 Da의 예상된 질량 이동을 제시한다.
- <308> 도 8은 친화력 정제된 GFP-pMpa의 비손상 단백질 ESI TOF-MS 스펙트럼을 도시한다. 탈소용돌이화된 충전 pAza 펩티드에 대한 신호 강도가 꽤 약하고(도 11 참조) 야생형 GFP의 단편과 pAza 함유 돌연변이체 GFP의 단편 간의 검출가능한 질량인 1 Da 질량의 차이점으로 인해 상태의 엔벨롭(도면 삽입에 도시됨)은 29696 Da의 일 주요 성분을 도시하며, 이는 실험적 오류 내에서 변형된 단백질 마이너스 N 말단 메티오닌 및 플러스 아세틸화(이론적 질량 29696.52)의 예상 질량과 일치한다. 29654 Da에서 더 작은 피쳐는 GFP-pMpa 마이너스 N 말단 메티오닌으로 배정된다. 2개의 양쪽 밴드는 비특이적 변형으로 인한 것으로 보인다.
- <309> 도 9는 pAa 함유 돌연변이체 GFP의 직렬 MS 스펙트럼을 도시한다. Y*-이온은 야생형 GFP의 스펙트럼에서 동일한 이온과 관련하여 26 Da 질량 이동을 제시한다.
- <310> 도 10은 pBpa 함유 돌연변이체 GFP의 직렬 MS 스펙트럼을 도시한다. 돌연변이체 GFP와 야생형 GFP 사이의 86 Da의 특징적 질량 이동이 명백히 관찰된다.
- <311> 도 11은 pAza 함유 돌연변이체 GFP의 직렬 MS 스펙트럼을 도시한다. 대부분의 pAza는 pAmpa를 붕괴한다(p-아미노페닐알라닌; 트립신 pAmpa 함유 단편의 직렬 MS 스펙트럼에 대한 도 15 참조). 하지만, 야생형 GFP의 신호와 비교하였을 때 Y*-이온의 25 Da의 특징적 질량 이동은 여전히 명백하게 관찰된다.
- <312> 도 12는 pPpa 함유 돌연변이체 GFP의 직렬 MS 스펙트럼을 도시한다. pPpa 단편의 신호는 배경에 의해 크게 가라앉게 된다. 이럼에도 불구하고, MSDB에 대한 스펙트럼의 데이터베이스 조사에 의한 탁월한 판별로서 여전히 확인된다. 보다 강한 Y*-이온은 38 Da의 특징적 질량 이동으로 관찰된다.
- <313> 도 13A 및 13B는 엠버 역제가 Invitrogen™ T-REX™ CHO 및 Invitrogen™ T-REX™ 293 세포에서 EcTyrRS 및 BstRNA^{tyr}_{CUA} 유전자 모두에 의존한다는 것을 도시한다. 도 13A는 상이한 구축물에 일시적으로 형질감염되는 Invitrogen™ T-REX™ CHO 세포의 녹색 형광 화상의 판별을 도시한다. 제1 판넬에서, 세포는 pcDNA4-GFP로 형질감염되었다. 제2 판넬에서, 세포는 pcDNA4-EcTyrRS 및 pcDNA4-GFP37TAG로 형질감염되었다. 제3 판넬에서, 세포는 pUC18-3BstRNA^{tyr}_{CUA} 및 pcDNA4-GFP37TAG로 형질감염되었다. 제4 판넬에서, 세포는 pUC18-3BstRNA^{tyr}_{CUA}, pcDNA4-EcTyrRS 및 pcDNA4-GFP37TAG로 형질감염되었다. 도 13B는 상이한 구축물에 일시적으로 형질감염되는 Invitrogen™ T-REX™ 293 세포의 녹색 형광 화상의 판별을 도시한다. 다양한 판넬에 사용되는 구축물은 도 13A에 도시된 것과 일치하였다.
- <314> 도 14는 Invitrogen™ T-REX™ CHO 및 293 세포에서 6개의 EcTyrRS 변이체의 발현의 웨스턴 블롯 분석을 도시한다. 각 반응을 위한 세포 용해물 40 μ g 분액을 항-c-myc 항체로 분석하였다. EcTyrRS 변이체의 발현은 활성 부위 돌연변이에 의해 영향을 받지 않는다. 변이체는 T-REX™ CHO 및 T-REX™ 293 세포주 모두에서 동일하게 발

현된다.

<315> 도 15는 pAzpa 함유 돌연변이체 GFP 유래의 펩티드 FSVSGE^{*}GEGDATY^{*}GK(서열 번호 103)의 주석이 달린 직렬 MS 스펙트럼을 도시한다. Y^{*}는 pAmpa(p-아미노페닐알라닌)를 의미한다. 확실히 하기 위해 Y^{*}-이온 시리즈에만 주석을 달았다. pAzpa 펩티드에 대한 신호 강도가 꽤 약하고(도 11 참조) 야생형 GFP의 단편과 pAzpa 함유 돌연변이체 GFP의 단편 간의 검출가능한 질량인 1 Da 질량의 차이점으로 인해 대부분의 pAzpa는 pAmpa를 붕괴한다.

<316> 도 16은 뉴클레오타이드 및 아미노산 서열을 제공한다.

도면

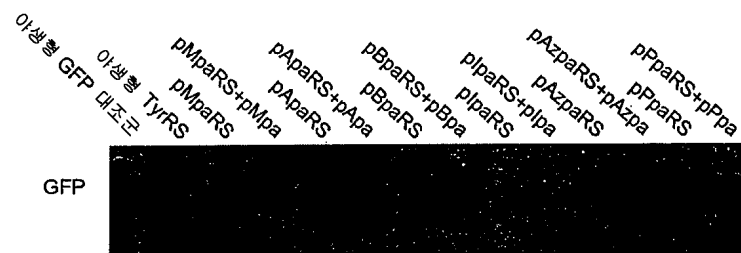
도면1A

구조 번호	명칭(약자)	구조
1	p-메톡시-L-페닐알라닌 (pMpa) O-메틸-L-티로신 O-Me-L-티로신 O-MeY 으로도 명명됨	
2	p-아세틸-L-페닐알라닌 (pApa) p-아실-L-페닐알라닌 p-아세틸Phe 으로도 명명됨	
3	p-벤조일-L-페닐알라닌 (pBpa) p-벤조페논 p-벤조일Phe 으로도 명명됨	
4	p-요오도-L-페닐알라닌 (pIpa) p-요오도-티로신 으로도 명명됨	
5	p-아지도-L-페닐알라닌 (pAzpa) p-아지도Phe 으로도 명명됨	

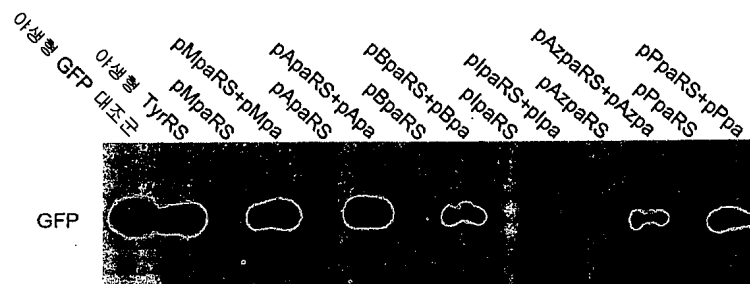
도면1B

구조 번호	명칭(약자)	구조
6	p-프로파르길옥시페닐알라닌 (pPpa) IUPAC 학명에 따라 2-아미노-3-[4-(프로프 -2-인일옥시)페닐] -프로피온산 pPR pPRO-Phe으로도 명명됨	
7	α-아미노카프릴산 알파-아미노카프릴산 2-아미노카프릴산 2-아미노옥탄산 또는 "C8"로도 명명됨	
8	o-니트로벤질시스테인 (o-NBC 또는 nbC)	
9	1,5-단실알라닌 IUPAC 학명에 따라 2- 아미노-3-(5-디메틸아미노 -나프탈렌-1-설폰닐아미노) -프로피온산으로도 명명됨	
10	o-니트로벤질세린 (o-NBS 또는 nbs)	

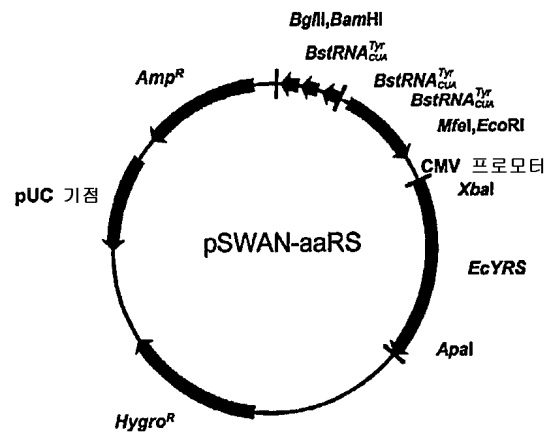
도면2



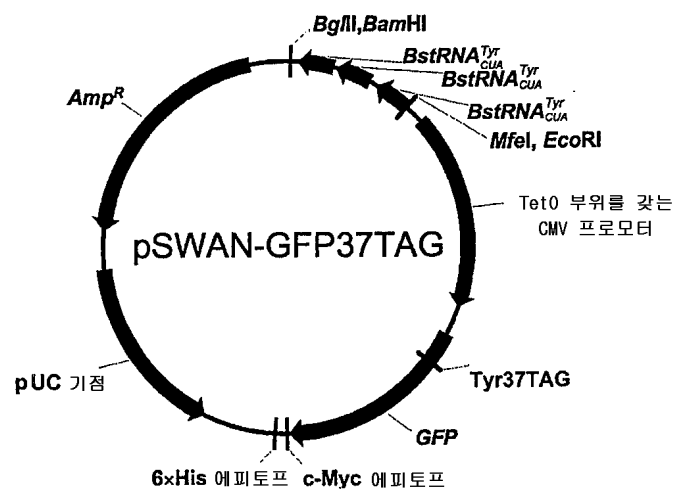
도면3



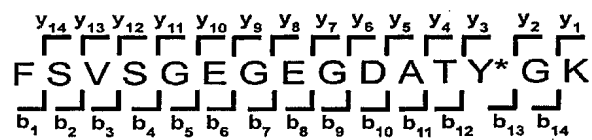
도면4A



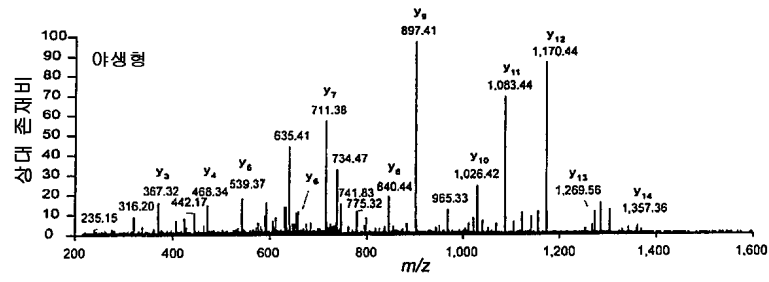
도면4B



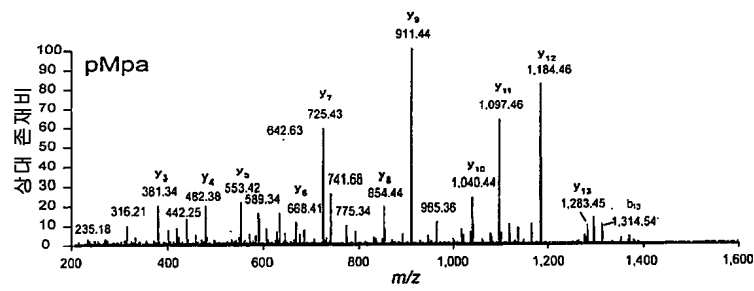
도면5



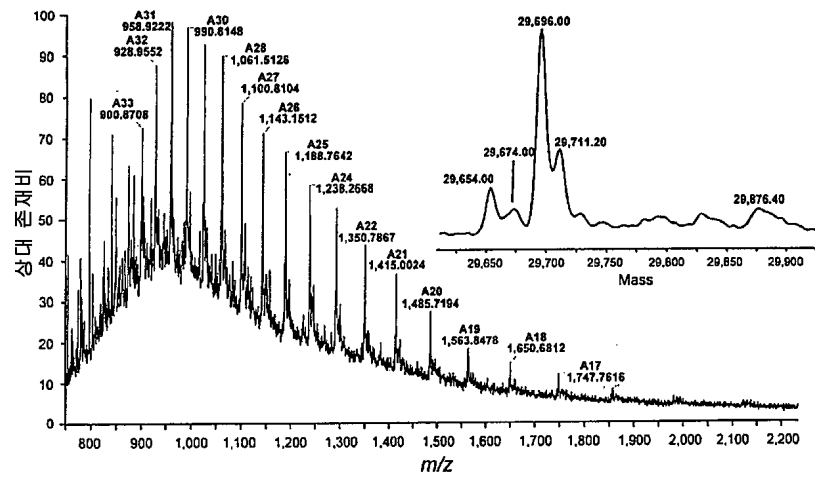
도면6



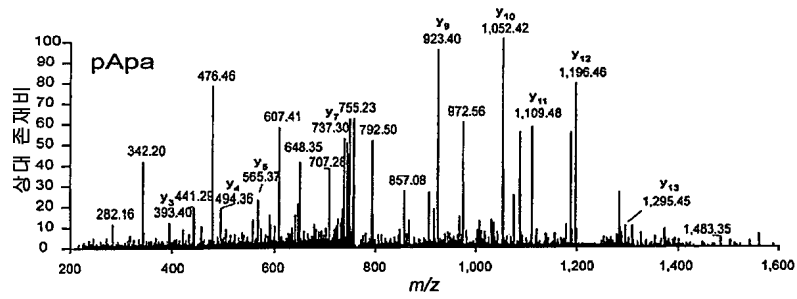
도면7



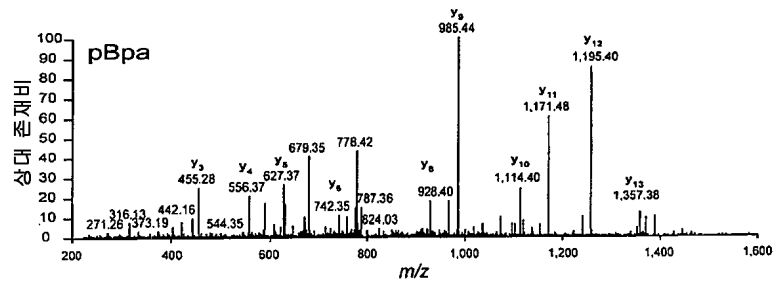
도면8



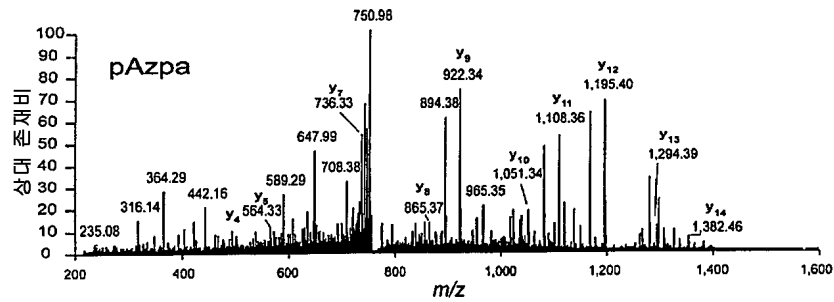
도면9



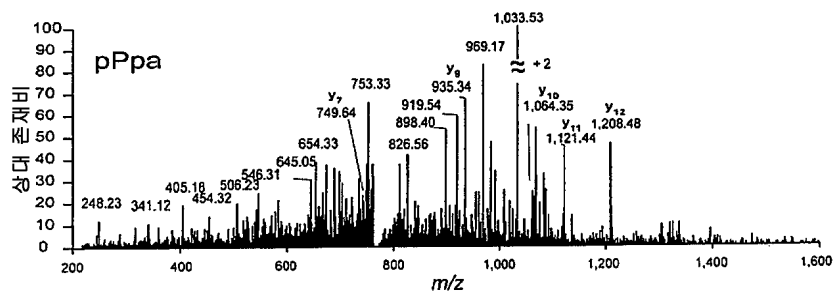
도면10



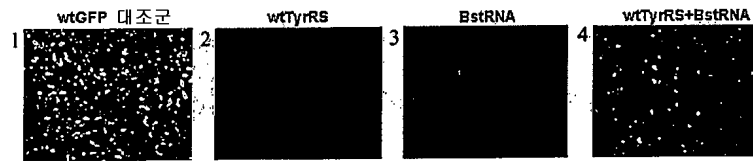
도면11



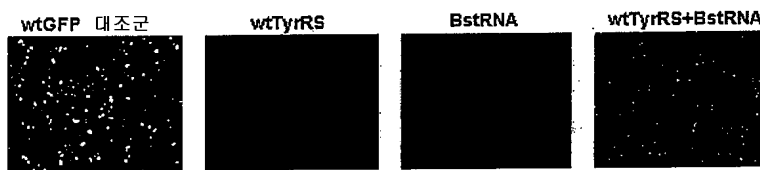
도면12



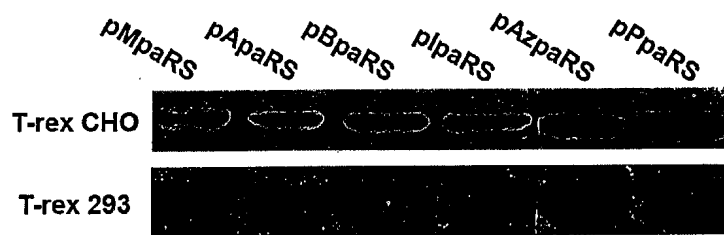
도면13A



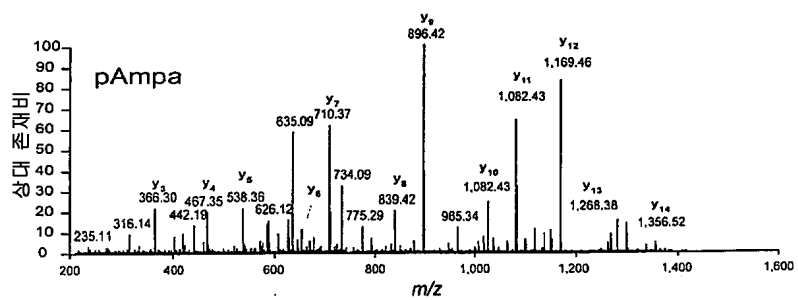
도면13B



도면14



도면15



도면16A

뉴클레오타이드 및 아미노산 서열

서열 번호 및 설명	서열
서열번호 1 이. 콜리 억제자 tRNA ^{Leu5} _{CUA}	GCCCCGAUGGUGGAAUCGUGAGACACAAGGGAUUCUAAAUCCUCGCGGUUCGCGCUGU GCGGGUUCAGUCCCGUCCGCGGUACCA
서열번호 2 이. 콜리 억제자 티로실-tRNA (<i>Ec</i> -tRNA ^{Tyr})	GGTGGGGTTCGCCAGCGGCCAAAGGGAGCAGACTCTAAATCTGCCGTCATCGACTTCGA AGGTTTCAATCCTTCCCCCACCACCA
서열번호 3 바실러스 스테아로서모필러스 염배 억제자 티로실-tRNA(CUA) (<i>Bs</i> -tRNA ^{Tyr} _{CUA})	GGAGGGGTAGCGAAGTGGCTAAACGCGCGGACTCTAAATCCGCTCCCTTTGGGTTCGG CGGTTTCAATCCGTCCTCCCTTCCA
서열번호 4 야생형 이. 콜리 류실-tRNA 합성효소 (<i>Ec</i> -LeuRS) 아미노산 서열	MQEQYRPEEIESKVQLHWDEKRTFEVTEDESKEKYCYLSMLPYPSGRLHMHGVRNYTTIG DVIARYQRLMKGNVLQPIGWDAFGLPAEGAAVKNNTPAPWPTYDNIAYMKNQKMLGFG YDWSRELATCTPEYRWQKFFTELYKKGLVYKKTSAVNWCPNDQTVLANEQVIDGCCW RCDTIVERKEIPQWFIKITAYADELLNDLKLHWPDTVTKMQRNWIGRSEGVETFNV NDYDNTLTVYTTTTPDPTFMGCTYLAVAGHPLAQKAAENNELAAFIDECRTKVAEAM ATMEKKGVDTGFKAVHPLTGEIIPVWAANFVLMYEGTGAVMAVPGHDQRDYEFASKYGL NTPKPVILAADGSEPDLSQALTEKGVLFNSGEFNGLDHEAAFNIAIDKLTAAMGVGERKV NYRLRDWQVSRQRYWGAPIPMVLEDTVMPTPDDQLPVILPEDVMDGITSPIKADPE WAKTTVNGMPALRETDFTFMESWYIYARCTCPQYKEGMLDSEAAANYLVPDIYIGGI EHAIMHLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVYVGENGERNWSVPD AIVERDEKGRIVKAKDAAGHELVTYTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVGEGANRFLKRVWKLVEHTAKGDVAALNVDALTENQKALRRDVHKT IAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPTDEGQDRALMQEALLAVVRLNLFPTPH ICTFLWQELKGGIDIDNAPWPADEKAMVEDSTLVVQVQNGKVRKITVPVDATBEEQVR ERAGQEHIVAKYLDGVTVRKVIYVPGKLLNLVVG

도면16B

서열 번호 및 설명	서열
서열번호 5 야생형 이. 콜리 류실-tRNA 합성효소 (<i>Ec</i> -LeuRS) 뉴클레오타이드 서열	ATGCAAGAGCAATACCGCCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTGAAGTAACCGAAGACGAGAGCAAGAGAAAGTATTACTGGCTGTCTA TGCTTCCCTTATCCTTCTGGTTCGACTACACATGGGCCACGTACGTAACTACACCATTCGGT GACGTGATCGCCCGCTACCGCATATGCTGGGCAAAACGTCCTGCAGCCGATCGGCTG GGACGCGTTTGGTCTGCTCTGCGGAAGGCGCGCGGTGAAACCAACACCGCTCCCGCAC CGTGGACGTACGACAAACATCGCGTATATGAAAAACCGAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACCGCCGGAATACCTACCGTTGGGAACAGAA ATTCTTACCGGAGCTGTATAAAAAAGGCTGGTATATAAGAAAGACTTCTCGGCTCAACT GGTGGCCGGAACGACGACGACCGTACTGGCGAACGAACAGTTATCGACGCGCTGCTGCTGG CGCTGCGATACCAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCAATGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAAGTGGTCACTGGCCAGACACCGTTA AAACCATGCAGCGTAACGTGATCGGTCGTTCCGAAGGCGTGGAGATCACTTCAACGTT AAGCAGTATGACAACACGCTGACCGTTTACACTACCGCCCGGACACCTTTATGGGTTG TACCTTACCTGGCGGTACGTGCGGGTTCATCCGCTGGCGCAGAAAGCGCGGAAATATC CTGAACCTGGCGGCTTTTATGACGAATGCCGTAAACCAAAAGTTGCGGAAGCTGAAATG GCGACGATGGAGAAAAAGGCGTCGATCTGGCTTTTAAAGCGGTTCACCCATTAACGGG CGAAGAAATTCCTGCTTTGGGCGACAAACTTCGTATTTGAGTACGGCAGCGGCGCAG TTATGGCGGTACCGGGGCGACGACGCGCGACTACGAGTTTGGCTCTAAATACGGCTG AACATCAAAACGCTTTTCTTGGCAGCTGACGGCTCTGAGCCAGATCTTTCTCAGCAAGC CCTGACTGAAAAAGGCGTGTCTTCAACTCTGGCGAGTTCAACGGCTTTGACCATGAAG CGGCTTCAACGCCATCGCCGATAAAGTACTGCTGAGTGGGCGTTGGCGAGCGTAAAGTG AACTACCGCCCTGCGCGACTGCGGTGTTTCCCGTCAGCGTTACTGGGGCGCGCGATTC GATGGTGACGCTGGAAGACGGTACCGTAAATGCCGACCCCGGACGACGCTGCCGGTGA TCCTGCCGGAAGATGTGTAATGGACGGCATTACCGCCGATTAAGAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGGTATGCCAGCAGTTCGCTGAAACCGACACTTTCGACAC CTTTATGGAGTCTCTTGGTACTATGCGCGCTACACTTGGCCGAGTACAAAGAAAGGTA TGCTGGATTTCCGAAGCGGCTAACTACTGGCTGCGCGTGGATATCTCATATTTGGTGGTATT GAACACGCCATTTATGCACCTGCTCTACTTCCGCTTCTTCCACAACTGATGCGGTGATGC AGGCATGTTGAATCTGACGAACGACGCAAAAGTGTCTGTGTGAGGATGAGTGGTCTGG CAGATGCTTCTACTATGTTGGCGAAAAAGGCGAACGTAACCTGGGTTTCCCGGTTGAT GCTATCTGTTGAACGTCAGCAGAAAGGCGTATCGTGAAAGCGAAAGATGCGCGAGGCCA TGAACCTGGTTTATACCGCATGAGCAAAATGTCCAAGTCAAGAACAAACGGTATCGAC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTTCGCTGTTTATGATGTTTGGT TCTCCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCGCTT CCTGAAACGCTGCTGGAAACTGGTTTACGAGCACACAGCAAAAGGTTGATGTTGCGGAC TGAACGTTGATGCGCTGACTGAAAAACAGAAAGCGCTGCGTTCGATGTCATTAAGACG ATCGCTAAAGTGACCGATGATATCGGCGCTGCTGAGACCTTCAACACCGCAATTCGCGG GATTTATGGAGTGTATGAACAACTGGCGAAAGCAGCAACCGATGGCGAGCAGGACCGCG CTTGTATGACGAAAGCACTGCTGGCGTTGTCTGATGCTTAACCGCTTACCCCGGCTG ATCTGCTTCAACGCTGTGGCAGGAACGAAAGGCGAAGGCGATATCGACAAACGCGCGGTG GCCGGTGTGTCAGCAAAAGCGATGGTGAAGACTTCAACGCTGCTGCTGCTGAGGTTA ACGGTAAAGTCCGTGCCAAATACCGTTCCGGTGGACGCAACGGAAGAACAGGTTCGC GAACGTGCTGGCAGGAACATCTGGTAGCAAAATATCTTGATGGCTTACTGTACGTAA AGTGATTTACGTACAGGTAACCTCTCACTTGGTCTGTTGGCTAA
서열번호 6 이. 콜리 야생형 TyrRS (합성효소)	MASSNLIKQLQERGLVAVQVTEDEALAEALAQGPALYCGFDPDADSLHLGHLVPLCLLK RFQQAQGHKEVALVGGATGLIDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNGNMNVLFLRLDIKHPVSNQMINKEAVKQRLNRDQGISFTFESYNL LQGYDFACLNKQYGVVLQIGGSDQWGNITSGIDLTRHLHQNVFGLTVPLITKADGTFK GKTEGGAVALDPKITSYKIFYQWINTADADVYRFLKFFTFMSTEEINALDEEDKNSGK APRAQYVLAQVTRLVHGEELQAARITECLFSGSLSALSEADFEQLAQDGVPMVEME KGADLMQALVDSSELQPSRGQARTIASNAITTINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLTCWK

도면16C

서열 번호 및 설명	서열
<p>서열번호 7</p> <p>이.폴리 아생형 TyrRS (합성효소) 폴리뉴클레오티드</p>	<p>ATGGCAAGCAGTAACCTTGATTAAACAATTGCAAGAGCGGGGGCTGGTAGCCAGGTGAC GGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCGCTCTATTGCGGGT TCGATCCTACCGCTGACAGCTTGCAATTGGGGCATCTTGTTCATTGTTATGCCGTGAAA CGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGCGCGACGGGTCTGAT TGCGGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTTTCAGG AGTGGGTGGCAAAATCCGTAAAGCAGGTTGCCCGCTTCCCTCGATTTTCGACTGTGGAGAA AACTCTGCTATCGCGGCGAACAACATATGACTGGTTCGGCAATATGAATGTGTGACCTT CCTGCGCGATATTGGCAAAACATCTCTCCGTTAACCCAGATGATCAACAAAGAAGCGGTTA AGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTG TTGAGGGTTATGACTTCGCTGTCTGAACAAACAGTACGGTGTGGTGTGCAAAATTGG TGGTCTGACCAAGTGGGTAAACATCACTTCTGGTATCGACCTGACCCGTCGCTGCATC AGAATCAGGTGTTTGGCCTGACCGTTCCCGTGATCACTAAAGCAGATGGCACAATAATT GGTAAACCTGAAGGCGGCGCAGTCTGGTTGGATCCGAAGAAAACAGCCCGTACAAATT CTACCACTTCTGGATCAACACTGCGGATGCCGACGTTTACCGCTTCTCGAAGTTCTTCA CCTTTATGAGCATTGAAGAGATCAACGCCCTTGAAGAAGAGATAAAAAACAGCGGTAAA GCACCGCGCGCCAGTATGTACTGGCGGAGCAGGTGACTCGTCTGGTTACCGGTGAAGA AGGTTTACAGCGCGCAAAACGTATTACCGAATGCCCTGTTACGCGGTCTTTGAGTGCGC TGAGTGAAGCGGACTTCGAACAGCTGGCGCAGGACGGCTACCGATGGTTGAGATGGAA AAGGGCGCAGACCTGATGCAAGGCACTGGTCGATTCTGAACGCAACCTTCCCGTGGTCA GGCAGCTAAACTATCGCTTCCAATGCCATCACCATTACGGGTGAAAAACAGTCCGATC CTGAATACCTTTTAAAGAAGAAGATCGTCTGTTTGGTCTGTTTACCTTACTGCGTCGC GGTAAAAAGAATTACTGTCTGATTTCGTGGAATAA</p>
<p>서열번호 8</p> <p>pOMeTyrRS-1 합성효소 폴리뉴클레오티드</p>	<p>ATGGCAAGCAGTAACCTTGATTAAACAATTGCAAGAGCGGGGGCTGGTAGCCAGGTGAC GGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTCGTGTGGGCT TCGATCCTACCGCTGACAGCTTGCAATTGGGGCATCTTGTTCATTGTTATGCCGTGAAA CGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGCGCGACGGGTCTGAT TGCGGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAAACTGTTTCAGG AGTGGGTGGCAAAATCCGTAAAGCAGGTTGCCCGCTTCCCTCGATTTTCGACTGTGGAGAA AACTCTGCTATCGCGGCGAATAATTATGACTGGTTCGGCAATATGAATGTGTGACCTT CCTGCGCGATATTGGCAAAACATCTCTCCGTTAACCCAGATGATCAACAAAGAAGCGGTTA AGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTG CTGAGGGTTATAGTATGGCCTGTTTGAACAAACAGTACGGTGTGGTGTGCAAAATTGG TGGTCTGACCAAGTGGGTAAACATCACTTCTGGTATCGACCTGACCCGTCGCTGCATC AGAATCAGGTGTTTGGCCTGACCGTTCCCGTGATCACTAAAGCAGATGGCACAATAATT GGTAAACCTGAAGGCGGCGCAGTCTGGTTGGATCCGAAGAAAACAGCCCGTACAAATT CTACCACTTCTGGATCAACACTGCGGATGCCGACGTTTACCGCTTCTCGAAGTTCTTCA CCTTTATGAGCATTGAAGAGATCAACGCCCTTGAAGAAGAGATAAAAAACAGCGGTAAA GCACCGCGCGCCAGTATGTACTGGCGGAGCAGGTGACTCGTCTGGTTACCGGTGAAGA AGGTTTACAGCGCGCAAAACGTATTACCGAATGCCCTGTTACGCGGTCTTTGAGTGCGC TGAGTGAAGCGGACTTCGAACAGCTGGCGCAGGACGGCTACCGATGGTTGAGATGGAA AAGGGCGCAGACCTGATGCAAGGCACTGGTCGATTCTGAACGCAACCTTCCCGTGGTCA GGCAGCTAAACTATCGCTTCCAATGCCATCACCATTACGGGTGAAAAACAGTCCGATC CTGAATACCTTTTAAAGAAGAAGATCGTCTGTTTGGTCTGTTTACCTTACTGCGTCGC GGTAAAAAGAATTACTGTCTGATTTCGTGGAATAA</p>

도면16D

서열 번호 및 설명	서열
<p>서열번호 9</p> <p>pOMeTyrRS-2 합성효소 폴리뉴클레오티드</p>	<p>ATGGCAAGCAGTAACCTTGATTAAACAATTGCAAGAGCGGGGGCTGGTAGCCACAGGTGAC GGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCGATCGCACTCAGTTGTGGCT TCGATCCTACCGCTGACAGCTTGCATTTGGGGCATCTTGTTCATTTGTTATGCTGAAA CGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGGACGGGCTCGAT TGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTTCAGG AGTGGGTGGACAAAATCCGTAGCAGGTGCCCCGTTCCTCGATTTCGACTGTGGAGAA AACTCTGCTATCGCGGCCAATAATTATGACTGGTTCAGCAATATGAATGTGCTGACCTT CCTGCGCGATATTGGCAAACTTCTCCGTTAACCCAGATGATCAACAAAGAGCGGTTA AGCAGCGTCTCAACCGTGAAGATCAGGGGATTTTCGTTCACTGAGTTTTCCTACAACTG CTGCAGGGTTATACGTATGCTTGTCTGAACAAACAGTACGGTGTGGTGTGCTGCAAAATGG TGGTTCGACAGTGGGGTAACATCACTTCTGGTATCGACCTGACCCGCTGCTGTCGATC AGAATCAGGTGTTTGGCCTGACCGTTCCGCTGATCACTAAAGCAGATGGCACCAAAATTT GGTAAACTGAAGCGGCGCAGTCTGGTTGGATCCGAAGAAACAGCCCGTACAAATTT CTACCACTTCTGGATCAACACTGCGGATGCCGAGCTTACCGCTTCTTGAAGTTCTTCA CCTTTATGAGCATGAAGAGATCAACGCCCTGGGAAGAAGAATAAACAGCGGTAA GCACCGCGCGCCAGTATGTAAGTGGCGGAGCAGGTGACTGCTGTTTACGGTGAAGA AGGTTTACAGGCGGCAAAACGTATTACCGAATGCCTGTTTACGCGGTTCTTTGAGTGC TGAGTGAAGCGGACTTCGAACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAA AAGGGCGCAGACCTGATGCAAGCACTGGTCGATTCTGAACCTGCAACCTTCCCGTGGTCA GGCAGTAAACTATCGCCCTCCAATGCCATCACCATTAAACGGTGAAAAACAGTCCGATC CTGAATACTTCTTAAAGAAGAAGATCGTCTGTTTGGTCTGTTTACCTTACTGCGTCGC GGTAAAAAGAATTACTGCTGATTGCTGGAATAA</p>
<p>서열번호 10</p> <p>pOMeTyrRS-3 합성효소 폴리뉴클레오티드</p>	<p>ATGGCAAGCAGTAACCTTGATTAAACAATTGCAAGAGCGGGGGCTGGTAGCCACAGGTGAC GGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCGATCGCACTCAGTTGTGGCT TCGATCCTACCGCTGACAGCTTGCATTTGGGGCATCTTGTTCATTTGTTATGCTGAAA CGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGGACGGGCTCGAT TGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTTCAGG AGTGGGTGGACAAAATCCGTAGCAGGTGCCCCGTTCCTCGATTTCGACTGTGGAGAA AACTCTGCTATCGCGGCCAATAATTATGACTGGTTCGGCAATATGAATGTGCTGACCTT CCTGCGCGATATTGGCAAACTTCTCCGTTAACCCAGATGATCAACAAAGAGCGGTTA AGCAGCGTCTCAACCGTGAAGATCAGGGGATTTTCGTTCACTGAGTTTTCCTACAACTG CTGCAGGGTTATAGTATGGCCTGTTTGAACAAACAGTACGGTGTGGTGTGCTGCAAAATGG TGGTTCGACAGTGGGGTAACATCACTTCTGGTATCGACCTGACCCGCTGCTGTCGATC AGAATCAGGTGTTTGGCCTGACCGTTCCGCTGATCACTAAAGCAGATGGCACCAAAATTT GGTAAACTGAAGCGGCGCAGTCTGGTTGGATCCGAAGAAACAGCCCGTACAAATTT CTACCACTTCTGGATCAACACTGCGGATGCCGAGCTTACCGCTTCTTGAAGTTCTTCA CCTTTATGAGCATGAAGAGATCAACGCCCTGGGAAGAAGAATAAACAGCGGTAA GCACCGCGCGCCAGTATGTAAGTGGCGGAGCAGGTGACTGCTGTTTACGGTGAAGA AGGTTTACAGGCGGCAAAACGTATTACCGAATGCCTGTTTACGCGGTTCTTTGAGTGC TGAGTGAAGCGGACTTCGAACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAA AAGGGCGCAGACCTGATGCAAGCACTGGTCGATTCTGAACCTGCAACCTTCCCGTGGTCA GGCAGTAAACTATCGCCCTCCAATGCCATCACCATTAAACGGTGAAAAACAGTCCGATC CTGAATACTTCTTAAAGAAGAAGATCGTCTGTTTGGTCTGTTTACCTTACTGCGTCGC GGTAAAAAGAATTACTGCTGATTGCTGGAATAA</p>

도면16E

서열 번호 및 설명	서열
서열번호 11 pOMeTyrRS-4 합성효소 폴리뉴클레오티드	ATGGCAAGCAGTAACCTTGATTAAACAATTGCAAGAGCGGGGGCTGGTAGCCAGGTGAC GGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTCGTGTGGCT TCGATCCTACCGCTGACAGCTTGCAATTGGGGCATCTTGTTCATTGTTATGCTGAAA CGCTTCCAGCAGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCAGCGGTCTGAT TGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTTACAGG AGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGCTTCTCGATTTGCACTGTGGAGAA AACTCTGCTATCGCGGCCAATAATTATGACTGGTTCCGCAATATGAATGTGCTGACCTT CCTGCGGATATTGGCAACACTTCTCCGTTAACAGATGATCAACAAAGAGCGGTTA AGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTCTTACAACCTG CTGCAGGGTTATAGTATGGCTGTTTGAACAACAGTACGGTGTGGTGTGCAAAATTGG TGGTCTGACCAAGTGGGTAAACATCACTTCTGGTATCGACCTGACCCGCTGCTGCAATC AGAATCAGGTGTTTGGCTGACCGTTCCGCTGATCACATAAGCAGATGGCACCATAATT GGTAAACTGAAGGCGCGCAGCTCTGGTTGGATCCGAAGAAAACAGCCCGTACAAATT CTACCACTTCTGGATCAACACTGCGGATGCCGACGTTTACCCTTCTGAACTTCTTCA CCTTTATGAGCATGAAGAGATCAACGCCCTGGAAGAAGAAGATAAAAACAGCGGTAAA GCACCGCGCGCCAGTATGTACTGGCGGAGCAGGTGACTCGTCTGGTTACCGGTGAAGA AGGTTTACAGGCGGCAACAGTATTACCGAATGCCGTGTTACGCGGTCTTTGAGTGCCG TGAGTGAAGCGGACTTCGAACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAA AAGGGCGCAGACCTGATGCAAGCAGTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAA GGCACGTAAAACATCGCCCTCCAAATGCCATCACCATTAAACGGTGAAAAACAGTCCGATC CTGAATACTTCTTTAAAGAAGAAGATCGTCTGTTTGGTCTGTTTACCTTACTGCGTCCG GGTAAAAAGAAATTACTGTCTGATTGCTGGAAATAA
서열번호 12 pOMeTyrRS-5 합성효소 폴리뉴클레오티드	ATGGCAAGCAGTAACCTTGATTAAACAATTGCAAGAGCGGGGGCTGGTAGCCAGGTGAC GGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCCGATCGCACTCACGTGTGGCT TCGATCCTACCGCTGACAGCTTGCAATTGGGGCATCTTGTTCATTGTTATGCTGAAA CGCTTCCAGCAGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCAGCGGTCTGAT TGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTTACAGG AGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGCTTCTCGATTTGCACTGTGGAGAA AACTCTGCTATCGCGGCCAATAATTATGACTGGTTCCGCAATATGAATGTGCTGACCTT CCTGCGGATATTGGCAACACTTCTCCGTTAACAGATGATCAACAAAGAGCGGTTA AGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTCTTACAGCCCTG CTGCAGGGTTATACGATGGCTGTCTGAACAACAGTACGGTGTGGTGTGCAAAATTGG TGGTCTGACCAAGTGGGTAAACATCACTTCTGGTATCGACCTGACCCGCTGCTGCAATC AGAATCAGGTGTTTGGCTGACCGTTCCGCTGATCACTAAAGCAGATGGCACCATAATT GGTAAACTGAAGCGCGCAGTCTGGTTGGATCCGAAGAAAACAGCCCGTACAAATT CTACCACTTCTGGATCAACACTGCGGATGCCGACGTTTACCCTTCTTGAAGTTCTTCA CCTTTATGAGCATGAAGAGATCAACGCCCTGGAAGAAGAAGATAAAAACAGCGGTAAA GCACCGCGCGCCAGTATGTACTGGCGGAGCAGGTGACTCGTCTGGTTACCGGTGAAGA AGGTTTACAGGCGGCAAAACGTATTACCGAATGCCGTGTTACGCGGTCTTTGAGTGCCG TGAGTGAAGCGGACTTCGAACAGCTGGCGCAGGACGGCGTACCGATGGTTGAGATGGAA AAGGGCGCAGACCTGATGCAAGCAGTGGTCTGATTTGAACTGCAACCTTCCCGTGGTCA GGCACGTAAAACATCGCCCTCCAAATGCCATCACCATTAAACGGTGAAAAACAGTCCGATC CTGAATACTTCTTTAAAGAAGAAGATCGTCTGTTTGGTCTGTTTACCTTACTGCGTCCG GGTAAAAAGAAATTACTGTCTGATTGCTGGAAATAA
서열번호 13 pOMeTyrRS-6 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTGACAGCTTGCAATTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGGACGGGCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTGAGGAGTGGGTGGCAAAAATCCGTAAGCAGGTTGCCCGCT TCTTCGATTTCGACTGTGGAGAAAACCTGCTATCGCGGCCAATAATTATGACTGGTTC AGCAATATGAATGTGCTGACCTTCTGCGGATATTGGCAACACTTCTCCGTTAACCA GATGATCAACAAAGAGCGGTAAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACTGAGTTTCTTACAACCTGCTGCAGGGTTATACGATGCTGTGCAACAAACAG TACGGTGTG

도면16F

서열 번호 및 설명	서열
<p>서열 번호 14</p> <p>pOMeTyrRS-7 (활성 부위) 합성효소 폴리뉴클레오티드</p>	<p>CGGGGGCTGGTACCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTC AGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAAAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTTCGT TCACTGAGTTTTCCTACAACCTGCTGACGGGTATACGTATGCCTGTCTGAACAAACAG TACGGTGTG</p>
<p>서열 번호 15</p> <p>pOMeTyrRS-8 (활성 부위) 합성효소 폴리뉴클레오티드</p>	<p>CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTC AGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAAAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTTCGT TCACTGAGTTTTCCTACAACCTGCTGACGGGTATACGTATGCCTGTCTGAACAAACAG TACGGTGTG</p>
<p>서열 번호 16</p> <p>pOMeTyrRS-9 (활성 부위) 합성효소 폴리뉴클레오티드</p>	<p>CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTC GGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAAAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTTCGT TCACTGAGTTTTCCTACAACCTGCTGACGGGTATTCGTATGCCTGTCTGCAACAAACAG TACGGTGTG</p>
<p>서열 번호 17</p> <p>pOMeTyrRS-10 (활성 부위) 합성효소 폴리뉴클레오티드</p>	<p>CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCACTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTC AGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAAAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTTCGT TCACTGAGTTTTCCTACAACCTGCTGACGGGTATACGTATGCCTGTCTGCAACAAACAG TACGGTGTG</p>
<p>서열 번호 18</p> <p>pOMeTyrRS-11 (활성 부위) 합성효소 폴리뉴클레오티드</p>	<p>CGGGGGCTGGTACCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCCTTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTC GGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAAAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTTCGT TCACTGAGTTTTCCTACAACCTGCTGACGGGTATTCGTATGCCTGTCTGCAACAAACAG TACGGTGTG</p>

도면16G

서열 번호 및 설명	서열
서열 번호 19 pOMeTyrRS-12 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCGTGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGCT TCCTCGATTTCGACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTC GGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACTTCTCCGTTAAACCA GATGATCAACAAAGAGCGGTAAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACCTGAGTTTTCCTACAACCTGCTGCGAGGGTTATAGTATTCGCTGTTTGAACAAACAG TACGGTGTG
서열 번호 20 pOMeTyrRS-13 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTACCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCGTGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGCT TCCTCGATTTCGACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTC GGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACTTCTCCGTTAAACCA GATGATCAACAAAGAGCGGTAAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACCTGAGTTTTCCTACAACCTGCTGCGAGGGTTATAGTATTCGCTGTTTGAACAAACAG TACGGTGTG
서열 번호 21 pOMeTyrRS-14 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCGTGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGCT TCCTCGATTTCGACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTC GGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACTTCTCCGTTAAACCA GATGATCAACAAAGAGCGGTAAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACCTGAGTTTTCCTACAACCTGCTGCGAGGGTTATAGTATTCGCTGTTTGAACAAACAG TACGGTGTG
서열 번호 22 p-아세틸-Phe-1 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCATTGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTGGGGCATC TTGTTCCATTGTTATGCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGCT TCCTCGATTTCGACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTC GGCAATATGAATGTGCTGACCTTCCTGCGCGATATTGGCAAACTTCTCCGTTAAACCA GATGATCAACAAAGAGCGGTAAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACCTGAGTTTTCCTACAACCTGCTGCGAGGGTTATAGTATTCGCTGTTTGAACAAACAG TACGGTGTGTTGCTGCAATTGGTGGTCTGACCAATGGGGTAACATCACCTTCGTGTTAT CGACCTGACCCGTCGTTGCTGATCAGAATCAGGTG
서열 번호 23 p벤조페논-1 (활성 부위) 합성효소 폴리뉴클레오티드	CAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCGATCGCACTCGG TTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTGGGGCATCTTGTTCCATTGTTAT GCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGTTAGGCGGCGCGACG GGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACT GTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGCTTCCTCGATTTCGACT GTGGAGAAAACCTCTGCTATCGCGGCCAATAATTATGACTGGTTCGGCAATATGAATGTG CTGACCTTCCTGCGCGATATTGGCAAACTTCTCCGTTAACCAGATGATCAACAAAGA AGCGGTAAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACCTGAGTTTTCCT ACAACCTGCTGCGAGGGTTATGGTTTTCGCTGTTTGAACAAACAGTACGGTGTGGTGTG CAAAATTGGTGGTTCGACCAAGTGGGGTAACATCACTTCTGGTATCGACCTGACCCGCTG TCTGCATCAGAATCAGGTG

도면16H

서열 번호 및 설명	서열
서열 번호 24 p벤조페논-2 (활성 부위) 합성효소 폴리뉴클레오타이드	GCCTTAGCAGAGCGACTGGCGCAAGGCCGATCGCACTCGGGTGTGGCTTCGATCCTAC CGCTGACAGCTTGCATTGGGGCATCTGTTCATTGTTATGCCTGAAACGCTTCCAGC AGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGGACGGGTCTGATTGGCGACCCG AGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTTAGGAGTGGGTGGA CAAAATCCGTAAGCAGGTTGCCCGTTCTCGATTTCGACTGTGGAGAAAACCTCGCTA TCGCGGCCAATAATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCTCGCGGAT ATTGGCAAACTCTCCGTTAACCAGATGATCAACAAAGAACGGTTAAGCAGCGTCT CAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTCTTACAACCTGCTGCAGGGTT ATGTTTATGCCTGTATGAACAAACAGTACGGTGTGGTGTGCAAAATGGTGGTTCTGAC CAGTGGGGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGT G
서열 번호 25 p아지도Phe-1 (활성 부위) 합성효소 폴리뉴클레오타이드	GGGCTGGTAGCCAGGTGACGGACGNAGAAGCGTTAGCAGAGCGACTGGCGCAAGGCC GATCGCACTCCTTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTGGGGCATCTTG TTCCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGTA GGCGGCGCAGCGGTCTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAA CACCGAAGAACTGTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGTTC TCGATTTTCGACTGTGGAGAAAACCTGCTATCGCGGCAATAATATGACTGGTTCGGC AATATGAATGTGCTGACCTTCTGCGCGATATGGCAAACTTCTCCGTTAACCAGAT GATCAACAAAGAACGGTTAAGCAGCGCTCAACCGTGAAGATCAGGGGATTTCGTTC CTGAGTTTCTTACAACCTGCTGCAGGGTTATCTATGGCCTGTGCGAACAACAGTAC GGTGTGGTGTGCAAAATGGTGGTTCTGACCAAGTGGGGTAACATCACTTCTGGTATCGA CCTGACCCGTCGTCTGCATCANAATCANGTG
서열 번호 26 p아지도Phe-2 (활성 부위) 합성효소 폴리뉴클레오타이드	TTAGCAGAGCGACTGGCGCAAGGCCGATCGCACTCGTTTGTGGCTTCGATCCTACCGC TGACAGCTTGCATTGGGGCATCTGTTCATTGTTATGCCTGAAACGCTTCCAGCAGG CGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGGACGGGTCTGATTGGCGACCCGAGC TTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTTAGGAGTGGGTGGACAA AATCCGTAAGCAGGTTGCCCGTTCTCGATTTCGACTGTGGAGAAAACCTCGCTATCG CGGCCAATAATTATGACTGGTTCCGCAATATGAATGTGCTGACCTTCTGCGCGATATT GGCAAACTTCTCCGTTAACCAGATGATCAACAAAGAACGGTTAAGCAGCGCTCTCAA CCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTCTTACAACCTGCTGCAGGGTTATT CTGCGGCTGTGCGAACAACAGTACGGTGTGGTGTGCAAAATGGTGGTTCTGACCA TGGGGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
서열 번호 27 p아지도Phe-3 (활성 부위) 합성효소 폴리뉴클레오타이드	GACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCGATCGCACTCCTGTGTGGCTT CGATCCTACCGCTGACAGCTTGCATTGGGGCATCTGTTCATTGTTATGCCTGAAAC GCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGGACGGGTCTGATT GGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTTACAGGA GTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGTTCTCGATTTCGACTGTGGAGAAA ACTCTGCTATCGCGGCCAATAATTATGACTGGTTCCGCAATATGAATGTGCTGACCTTC CTGCGGATATGTGGCAAACTTCTCCGTTAACCAGATGATCAACAAANAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTCTTACAACCTGC TGCAGGGTTATTTCGGCTGCCTGTGCGAACAACAGTACGGNGNGNGCTGCAAAATGGN GGTTCTGACCAAGGGGGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCTGCATCA AAATCAGGTG
서열 번호 28 p아지도Phe-4 (활성 부위) 합성효소 폴리뉴클레오타이드	GCCTTAGCAGAGCGACTGGCGCAAGGCCGATCGCACTCGTTTGTGGCTTCGATCCTAC CGCTGACAGCTTGCATTGGGGCATCTGTTCATTGTTATGCCTGAAACGCTTCCAGC AGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGGACGGGTCTGATTGGCGACCCG AGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTTAGGAGTGGGTGGA CAAAATCCGTAAGCAGGTTGCCCGTTCTCGATTTCGACTGTGGAGAAAACCTCGCTA TCGCGGCCAATAATTATGACTGGTTCGGCAATATGAATGTGCTGACCTTCTGCGCGAT ATTGGCAAACTTCTCCGTTAACCAGATGATCAACAAAGAACGGTTAAGCAGCGTCT CAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTCTTACAACCTGCTGCAGGGTT ATAGTGCGGCTGTGTTAACAACAGTACGGTGTGGTGTGCAAAATGGTGGTTCTGAC CAGTGGGGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCTGCATCAGAATCANGT G

도면16I

서열 번호 및 설명	서열
서열 번호 29 p아지도Phe-5 (활성 부위) 합성효소 폴리뉴클레오티드	GACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCGATCGCACTCATTGTGGGCTT CGATCCTACCGCTGACAGCTTGCATTGGGGCATCTTGTTCATTGTTATGCCTGAAAC GCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGCAGCGGTTGATT GGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTTTCAGGA GTGGGTGGCAAAATCCGTAAAGCAGGTTGCCCGTTCCTCGATTTCGACTGTGGAGAAA ACTCTGCTATCGCGGCAATGATTATGACTGGTTCGGCAATATGAATGTGTGACCTTC CTGCGCGATATTGGCAAACTCTCCGTTAACCAGATGATCAACAAGAAGCGGTTAA GCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCCTACAACCTGC TGCAGGGTTATAATTTCCTGTGTGAACAAACAGTACGGTGTGGTGTGCAAAATGGT GGTCTGACCACTGGGGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCATCA GAATCAGGTG
서열 번호 30 p아지도Phe-6 (활성 부위) 합성효소 폴리뉴클레오티드	CGACTGGCGCAAGGCCGATCGCACTACGTTGTGGCTTCGATCCTACCGCTGACAGCTT GCATTGGGGCATCTTGTTCATTGTTATGCCTGAAACGCTTCCAGCAGGCGGGCCACA AGCCGGTTGCGCTGGTAGGCGGCGCAGCGGTTGATTGGCGACCCGAGCTTCAAAGCT GCCGAGCGTAAGCTGAACACCGAAGAACTGTTTCAGGAGTGGGTGGCAAAATCCGTAA GCAGGTTGCCCGCTTCTCGATTTCGACTGTGGAGAAACTCTGCTATCGCGGCAATA ATTATGACTGGTTCGCAATATGAATGTGTGACCTTCTTCGCGGATATTGGCAACAC TTCCTCCGTAAACAGATGATCAACAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGA TCAGGGGATTTCGTTCACTGAGTTTTCCTACAATCTGCTGCAAGGTTATTTCGGTGTCT GTCCTTAACAACAGTACGGTGTGGTGTGCAAAATGGTGGTTCTGACCACTGGGGTAAC ATCACTTCTGGTATCGACCTGACCCGTCGCTGCATCAGAATCAGGTG
서열 번호 31 pPR-EcRS-1 (프로파르길 옥시페닐 알라닌 합성효소) (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTANCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCAGTCGCACTCGGGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTGGGGCATC TTGTTCCATTGTTATGCCGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCAGCGGTTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGCAAAATCCGTAAAGCAGGTTGCCCGCT TCCTCGATTTCGACTGTGGAGAAACTCTGCTATCGCGGCAATAAATTATGACTGGTTC GGCAATATGAATGTGTGACCTTCTCTGCGGATATTGGCAAACTTCTCCGTTAAACCA GATGATCAACAAGAAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACTGAGTTTTCCTACAACCTGCTGCAAGGTTATTCTATGGCTTGTGTAACAAACAG TACGGTGTGGTGTGCAAAATGGTGGTTCTGACCACTGGGGTAACATCACTTCTGGTAT CGACCTGANCCGTCGCTGCATCAGAATCAGGTG
서열 번호 32 pPR-EcRS -2 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCAGTCGCACTCACGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTGGGGCATC TTGTTCCATTGTTATGCCGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCAGCGGTTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGCAAAATCCGTAAAGCAGGTTGCCCGCT TCCTCGATTTCGACTGTGGAGAAACTCTGCTATCGCGGCAATAAATTATGACTGGTTC GGCAATATGAATGTGTGACCTTCTCTGCGGATATTGGCAAACTTCTCCGTTAAACCA GATGATCAACAAGAAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACTGAGTTTTCCTACAACCTGCTGCAAGGTTATTTCGCTGCTGCTTAAACAAACAG TACGGTGTGGTGTGCAAAATGGTGGTTCTGACCACTGGGGTAACATCACTTCTGGTAT CGAACCTGANCCGTCGCTGCATCAAAATCAAGTG
서열 번호 33 pPR-EcRS -3 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTACCCCAAGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCAGTCGCACTCTCTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTGGGGCATC TTGTTCCATTGTTATGCCGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCAGCGGTTGATTGGCGACCCGAGCTTCAAAGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGCAAAATCCGTAAAGCAGGTTGCCCGCT TCCTCGATTTCGACTGTGGAGAAACTCTGCTATCGCGGCAATAAATTATGACTGGTTC GGCAATATGAATGTGTGACCTTCTCTGCGGATATTGGCAAACTTCTCCGTTAAACCA GATGATCAACAAGAAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACTGAGTTTTCCTACAACCTGCTGCAAGGTTATTTCGCTGCTGCTTAAACAAACAG TACGGTGTGGTGTGCAAAATGGTGGTTCTGACCACTGGGGTAACATCACTTCTGGTAT CGACCTGACCCGTCGCTGCATCAGAATCAGGTG

도면16J

서열 번호 및 설명	서열
서열 번호 34 pPR-EcRS-4 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCGCGTGGCGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAGGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATATGACTGGTTC GGCAATATGAATGTGCTGACCTTCTCGCGCATATTGGCAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACGTAGTTTTCTTACAACCTGCTGCAAGGTTATTCTTATGCTGTCTTAACAAACAG TACGGTGTGGTGTGCAAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTAT CGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
서열 번호 35 pPR-EcRS-5 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCGCGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAGGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATATGACTGGTTC GGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACGTAGTTTTCTTACAACCTGCTGCAAGGTTATACGATGGCTGTTGTAACAAACAG TACGGTGTGGTGTGCAAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTAT CGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
서열 번호 36 pPR-EcRS-6 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCAGGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAGGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATATGACTGGTTC GGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCGCTGAGTTTTCTTACAACCTGCTGCAAGGTTATACGTTTGCCTGTATGAACAAACAG TACGGTGTGGTGTGCAAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTAT CGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
서열 번호 37 pPR-EcRS-7 (활성 부위) 합성효소 폴리뉴클레오티드	GTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGGCCGATCGCACTCAGGTG TGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATCTTGTTCCATTGTTATGCC TGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTGGTAGGCGGCGGACGGGT CTGATTGGCGACCCGAGCTTCAAGGCTGCCGAGCGTAAGCTGAACACCGAAGAACTGTT TCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGTTCCTCGATTTGCACTGTG GAGAAAACCTCTGCTATCGCGGCCAATAATATGACTGGTTCGGCAATATGAATGTGCTG ACCTTCTGCGCGATATTGGCAACACTTCTCCGTTAACCCAGATGATCAACAAAGAAGC GGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGTTCACTGAGTTTTCTTACA ATCTGCTGCAAGGTTATTTCGGTGCCTGCTTAAACAAACAGTACGGTGTGGTGTGCA ATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTATCGACCTGACCCGTCGTCT GCATCAGAATCAGGTG
서열 번호 38 pPR-EcRS-8 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCGTTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAGGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATATGACTGGTTC GGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACGTAGTTTTCTTACAACCTGCTGCAAGGTTATTCTGATGGCTGTACGAACAAACAG TACGGTGTGGTGTGCAAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTAT CGACCTGACCCGTCGTCTGCATCAGAATCAGGTG

도면16K

서열 번호 및 설명	서열
서열 번호 39 pPR-EcRS-9 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCAGTTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAGGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATCTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATATGACTGGTTC GGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACGTAGTTTTCTTACAACCTGCTGCAAGGTTATAGTTTGCCTGTCTGAACAAACAG TACGGTGTGGTGTGCAAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTAT CGACCTGACCCGTCGTCTGCATCAGAATCAGGTG
서열 번호 40 pPR-EcRS-10 (활성 부위) 합성효소 폴리뉴클레오티드	CGGGGGCTGGTAGCCAGGTGACGGACGAGGAAGCGTTAGCAGAGCGACTGGCGCAAGG CCCGATCGCACTCAGGTGTGGCTTCGATCCTACCGCTGACAGCTTGCATTTGGGGCATC TTGTTCCATTGTTATGCCCTGAAACGCTTCCAGCAGGCGGGCCACAAGCCGGTTGCGCTG GTAGGCGGCGCGACGGGTCTGATTGGCGACCCGAGCTTCAAGGCTGCCGAGCGTAAGCT GAACACCGAAGAACTGTTTCAGGAGTGGGTGGACAAAATCCGTAAGCAGGTTGCCCGGT TCCTCGATTTGCACTGTGGAGAAAACCTCTGCTATCGCGGCCAATAATATGACTGGTTC GGCAATATGAATGTGCTGACCTTCTGCGCGATATTGGCAACACTTCTCCGTTAACCA GATGATCAACAAAGAAGCGGTTAAGCAGCGTCTCAACCGTGAAGATCAGGGGATTTCGT TCACGTAGTTTTCTTACAACCTGCTGCAAGGTTATACGTTTGCCTGTACTAACAACAG TACGGTGTGGTGTGCAAAATTGGTGGTTCTGACCAGTGGGGTAACATCACTTCTGGTAT CGACCTGACCCGTCGTCTGCATCAGAATCAGGTG

도면16L

서열 번호 및 설명	서열
서열 번호 41 1,5-단실알라닌 아미노아실-tRNA 합성효소 클론 B8 뉴클레오타이드 서열	ATGGAAGAGCAATACCGCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTTGAAGTAACCGAAGAGAGAGCAAAGAGAGTATTACTGCCGTGTCTG CTAATCCCTATCCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCCGCTACCAGCGTATGCTGGGCAAAACGTCCTGCAGCCGATCGGGTG GGACGCGTTTGGTCTGCTGCGGAAGGCGCGGCGGTGAAAAACAACACCGCTCCGGCAC CGTGGACGTACGACAACATCGCGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCCGGAATACTACCGTTGGGAACAGAA ATTCTTTCACCGAGCTGTATAAAAAAGGCTTGGTATATAAGAGACTTCTGCGGTCAACT GGTGTCCGAACGACGACCGTACTGGCGAACGAAACAAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGCAGCGTAACTGGATCGGTCTTCCGAAGCGTGGAGATCACCTTCAACGTT AACGACTATGACAACGCTGACCGTTTACACTACCCGCCCGGACACCTTTATGGGTG TACCTTACCTGGCGGTAGCTGCGGGTCATCCGCTGGCGCAGAAAGCGCGGAAAAATAATC CTGAATGGCGGCTTTATGACGAATGCCGTAAACCAAAAGTTGCCGAAGCTGAAATG CGCAGCATGGAGAAAAAGGCGTGCATACTGGCTTTAAAGCGGTTACCCATTAACGGG CGAAGAAATTCCTTTGGGCGAGCAAACTTCGTATTGATGGAGTACGGCAGCGGCGCAG TTATGGCGGTACCGGGGACGACCGAGCGCGACTACGAGTTTGCCTCTAAATACGGCTG AACATCAAACCGGTATCTTGGCAGCTGACGGCTCTGAGCCAGATCTTTCTCAGCAAGC CCTGACTGAAAAAGGCGTGTCTTCAACTCTGGCGAGTTCAACGGTCTTGACCATGAAG CGGCCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGTG AACTACCGCTGCGCGACTGGGGTGTTCCTCGTCAGCGTTACTGGGCGCGCCGATTC GATGGTGACTCTAGAAGACGGTACCCTAATGCCGACCCGGACGACAGCTGCCGGTGA TCCTGCCGGAAGATGTGGTAATGGACGGCATTACAGCCCGATTAAGCAGATCCGGAG TGGCGCAAACTACCGTTAACGGTATGCCAGCACTGCGTGAAACCGACACTTTTGACAC CTTTATGGAGTCTGCTGGATTATGCGCGCTACACTTGCCCGCAGTACAAGAAGGTA TGCTGGATTCCGAAGCGGCTAACTACTGCTGCCGCTGGATATCGGTATTGGTGGTATT GAACACGCCATATGACGCTGCTCTACTTCCGCTTCTTCCACAACTGATGCGTGATGC AGGCATGGTGAACCTTGACGAACAGCGAAACAGTTGCTGTCTAGGGTATGGTGTCTGG CAGATGCCCTTCTACTATGTTGGCGAAAACGGCGAACGTAACCTGGGTTTCCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGGCCGTATCGTGAAAGCGAAAGATGCGGCGAGGCCA TGAACCTGGTTTATACCGCATGAGCAAAATGTCCAAGTCGAAGAACACCGGTATCGACC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTTCTGCTGTTTATGATGTTTGTCT TCTCCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAAGGGGCTAACCGCTT CCTGAAACGTGTCTGGAACCTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAATCAGAAAGCGCTGCGTCGCGATGTGCATAAAAACG ATCGCTAAAGTGACCGATGATATCGGCCGTGCTCAGACCTTCAACACCGCAATTGCGGC GATTATGGAGCTGATGAACAACTGGCGAAAGCACCAACCGATGGCGAGCAGGATCGCG CTCTGATGCAGGAAGCACTGCTGGCCGTTGTCCGTATGCTTAAACCGTTTACCCCGCAC ATCTGCTTCAACGCTGTGGCAGGAACGAAAGGCGAAGGCGATATCGACAACCGCGCTG GCCGGTTGCTGACGAAAAAGCGATGGTGGAAAGACTCCACGCTGGTGTGGTGCAGGTTA ACGGTAAAGTCCGTGCCAAAAATCACCGTTCCGGTGGACGCAACGGGAAGACAGGTTACG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTACGTACAGGTAACTCCTCAATCTGGTCTGTGGCTAA

도면16M

서열 번호 및 설명	서열
서열 번호 42 1,5-단실알라닌 아미노아실-tRNA 합성효소 T252A 뉴클레오타이드 서열	ATGGAAGAGCAATACCGCCCGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTTGAAGTAACCGAAGACGAGAGCAAAAGAGAATATTACTGCCTGCTCG CTAATCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCCGCTACCGAGCTATGCTGGGCAAAACGCTCTGCAGCCGATCGGCTG GGACGCGTTTGGTCTGCCTGCGGAAGCGCGGCGGTGAAAAACAACACCGCTCCGCGCAC CGTGGACGTACGACAACATCGCTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCCGGAATACTACCGTTGGGAACAGAA ATCTTCCACCGAGCTGTATAAAAAAGCGCTGGTATATAAGAAAGACTTCTGCGGTCAACT GGTGTCCGAACGACGACCGTACTGGCGAACGAACAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGACGCGTAACCTGGATCGGTCTGTCCGAAGCGTGGAGATCACCTTCAACGTT AACGACTATGACAACACGCTGACCGTTTACACTACCGCCCGGACGCGTTTATGGGTTG TACCTACCTGGCGGTAGCTGCGGTCATCCGCTGGCGCAGAAAGCGCGGAAAAATAATC CTGAACCTGGCGGCTTTATGACGAATGCCGTAACACCAAAGTTGCCGAAGCTGAAATG GCGACGATGGAGAAAAAGGCGTCGATACTGGCTTTAAAGCGGTTACCCATTAACGGG CGAAGAAATTCCTGTTTGGCGACAACTTCTGATTGATGGAGTACGGCACGGCGCAG TTATGGCGGTACCGGGGACGACACCGCGACTACGAGTTTGCTCTAAATACGGCGCTG AACATCAAACCGGTTATCTGGCAGCTGACGGCTCTGAGCCAGATCTTCTCAGCAAGC CCTGACTGAAAAAGCGGTGCTGTCAACTCTGGCGAGTTCAACGGCTTTCACCATGAAG CGGCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGTG AACTACCGCTGCGCGACTGGGGTGTTCCTCGTCAGCGTTACTGGGGCGCGCCGATTC GATGGTGACTCTAGAAGACGGTACCGTAATGCCGACCCCGGACGACGCTCCCGGTGA TCCTGCCGGAAGATGTGGTAATGGACGGCATTACAGCCCGATTAAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGGTATGCCAGCAGCTGCGTGAAACCGACACTTTCGACAC CTTTATGGAGTCTGCTGGAATTTATGCCGCGTACACTTGCCCGCAGTACAAGAAGGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCCGCTGGATATCGGTATTGGTGGTATT GAACACGCCATTATGACGCTGCTCTACTTCCGCTTCTTCCACAACATGATGCGTGATGC AGGCATGGTGAACCTCTGACGAACAGCGAAACAGTTGCTGTCTCAGGATATGGTCTGG CAGATGCCCTTCTACTATGTTGGCGAAAAACGGCGAACGTAACCTGGGTTTCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGCCGATCTGTGAAAGCGAAAGATCGCGCAGGCCA TGAACCTGGTTTATACCGGATGAGCAAAATGTCCAAGTCGAAGAACAACGGTATCGACC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTTCTGTTTATGATGTTTGGCT TCTCCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCGCTT CCTGAAACGTGTCTGGAACTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAT TGAACGTTGATGCGCTGACTGAAATCAGAAAGCGCTGCGTTCGCGATGTGATATAAACG ATCGCTAAAGTGACCGATGATATCGCCCGTCTGTCAGACCTTCAACACCGCAATTCGCGC GATTATGGAGCTGATGAACAACTGGCGAAAGCACCACCGATGGCGAGCAGGATCCGC CTCTGATGACGGAAGCACTGCTGGCCGTTGTCCGTATGCTTAAACCGTTTACCCCGCAC ATCTGCTTCAACGCTGTGGCAGGAACGTAAGGCGAAGGCGATATCGACAACGCGCCGTG GCCGGTTGCTGACGAAAAAGCGATGGTGAAGACTCCACGCTGGTCTGGTGGCAGGTTA ACGGTAAAGTCCGTGCCAAAAATCACCGTTCCGGTGGACGCAACGGAAGAACAGGTTACG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTACGTACAGGTAACTCCTCAATCTGGTCTGGCTAAGCGGCC

도면16N

서열 번호 및 설명	서열
서열 번호 43 1, 5-단실알라닌 아미노아실-tRNA 람성효소 V338A 뉴클레오타이드 서열	ATGGAAGAGCAATACCGCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCCGACATTTGAAGTAACCGAAGACGAGAGCAAAGAGATATTACTGCCCTGCTG CTAATCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCGCTACCGCTATGCTGGGCAAAACGTCCTGCAGCCGATCGGCTG GGACGCGTTTGGTCTGCCGTAAGGCGCGGCGGTGAAAAACAACACCGCTCCGGCAC CGTGGACGTACGACAACATCGGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCGGAATATACCGTTGGGAACAGAA ATTCTTACCCGAGCTGTATAAAAAAGGCTGGTATATAAGAACTTCTGGGTCAACT GGTGTCGAACGACGACCGTACTGGCGAACGAACAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCCGCAGTGGTTATCAAAATCACTGC TTACGCTGACGAGCTGCTAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGACGCGTAACGTGATCGGTCTGTCGAAAGGCTGGAGATCACTTTCAACGTT AACGACTATGACAAACGCTGACCGTTTACACTACCGCCCGGACACCTTTATGGGTG TACCTACCTGGCGGTAGCTGCGGCTCATCCGCTGGCGCAGAAAGCGCGGAAAAATATC CTGAACCTGGCGGCTTTATTGACGAATGCCGTAAACCAAAGTTGCCGAAGCTGAAATG GCGACGATGGAGAAAAAGCGCTCGATACTGGCTTTAAAGCGGTTACCCATTAAACGGG CGAAGAAATTCCTGTTGGGACGAACTTCGTATTGATGGAGTACGGCAGCGGCGCAG TTATGGCGGCGCGGGGACGACCGCGGCTACGAGTTTGCCTCTAAATACGGCTG AACATCAAAACCGGTTATCTGGCAGCTGACGGCTCTGAGCCAGATCTTTCTCAGCAAGC CCTGACTGAAAAAGCGGTGCTGTCAACTCTGGCGAGTTCAACGGCTTTGACCATGAAG CGGCCCTTCAACCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGGCTAAAGTG AACTACCGCTGCGGCTGCGGCTGTTTCCCGTCAGCGTTACTGGGCGCGCCGATTCG GATGGTGACTCTAGAAGACGGTACCGTAATGCCGACCCCGACGACCGCTGCCGCTGA TCCTGCCGGAAGATGGTAAATGGACGGCATTACCGCCGATTAAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGGTATGCCGACACTGCGTGAAACCGGACACTTTGACAC CTTTATGGAGTCTGCTGGATTATGCGCGCTACACTTGCCCGCAGTACAAAGAAAGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCCGCTGGATATCGGTATTGGTGGTATT GAACACGCCATTATGACGCTGCTCTACTTCCGCTTCTCCACAACTGATGCGTGATGC AGCCATGGTGAATCTGACGAACGAGCAACAGTTGCTGTGTCAGGGTATGGTGGCTGG CAGATGCCCTTCTACTATGTTGGCGAAAAACGGCGAACGTAACGGGTTTCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGCCGTATCGTGAAACGGAAGATGCGGCGAGGCCA TGAACCTGGTTTATACCGCATGAGCAAAATGTCCAAGTCGAAGAACCAACGGTATCGACC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTTCTGCTGTTTATGATGTTTGGCT TCTCCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGGGAAGGGGCTAACCGCTT CCTGAAACGTGCTGGAACTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAAAATCAGAAAGCGCTGCTGCGGATGTGCATAAAACG ATCGCTAAAGTGACCGATGATATCGGCCGCTGCTCAGACCTTCAACACCGCAATTGCGGC GATTTAGGAGCTGATGAACAACTGGCGAAAGCACCACCGATGGCGAGCAGGATCGCG CTCTGATGACGGAAGCACTGCTGGCCGTTGTCCGTATGCTTAACCCGTTACCCCGCAC ATCTGCTTACGCTGTGGCAGGAATGAAAGGCGAAGGCGATATCGACAACGCGCCGTG GCCGTTGCTGACGAAAAAGCGATGGTGAAGACTCCACGCTGCTGCTGGTGCAGGTTA ACGGTAAAGTCCGTGCCAAAATCACCGTTCCGGTGGACGCAACGGAAGAACAGGTTCCG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTTACGTACCGGTAACCTCTCAATCTGGTCTGTTGGCTAAGCGGCC

도면160

서열 번호 및 설명	서열
<p>서열 번호 44</p> <p>o-니트로벤질세린 아미노아실-tRNA 합성효소 클론 G2-6 뉴클레오타이드 서열</p>	<p>ATCTCGAAGCACACGAAACTTTTCCTTCCCTTCATTACGCACACTACTCTCTAATGAG CAACGGTATACGGCTTCCTTCCAGTTACTTGAATTTGAAATAAAAAAAGTTTGCTGT CTTGCTATCAAGTATAAATAGACCTGCAATTATTAATCTTTTGTTTCCTGTCATTGTT CTCGTTCCCTTTCCTTCTTCTTTTTCCTGACCAATATTTCAAGCTATACCAAGCAT ACAATCAACTGAATTCAGTATGGAAGAGCAATACCGCCCGGAAGAGATAGAATCCAAG TACAGCTTCAATGGGATGAGAAGCGCAGATTTGAAGTAACCGAAGACGAGGGCAAGAG AAGTATTACTGCCCTGCTTGGTTCGCCCTATCCTTCTGGTCTGACTACACATGGGCCAGT ACGTAATACACCATCGGTGACGTGATCGCCCGCTACAGCGTATGCTGGGCAAAACG TCCTGCAGCGGATCGGCTGGGACGGGTTGGTCTGCCGCGGAAGCGCGCGCTGAAA AACACACCCGCTCCGGCACCCTGGACGTACGACAACATCGGTATATGAAAAACAGCT CAAAATGCTGGGCTTTGGTTATGACTGGAGCCGAGCTGGCAACCTGTACGCCGGAAT ACTACCGTTGGGAACAGAAATTCCTACCGAGCTGTATAAAAAAGGCGTGGTATATAAG AAGACTTCTCGGTTCACTGGTGTCCGAACGACGAGCCGCTACTGGCGAAGCAAGAT TATCGAGCGCTGCTGCTGGCGCTGCGATACCAAGTTGAACGTAAAGAGATCCCGCAGT GGTTTATCAAAATCACTGCTTACGCTGACGAGCTGCTCAACGATCTGGATAAATGGAT CACTGGCCAGACCCGTTAAACCATGACAGCTAATGGATCGGTCTCCGAAGGCGT GGAGATCACCTTCAACGTTAACGACTATGACAACACGCTGACCGTTTACGCTTCCCGCC CGGACACCTTTATGGGTTGTACCTACCTGGCGGTAGCTGCGGCTCATCCGCTGCGCGAG AAAGCGCGCGAAATAATCTGAACTGGCGGCTTTATTTGACGAATGCCGTACACCAAA AGTTGCCGAAGCTGAAATGGCGACGATGGAGAAAAAGGCGTGCATCTGCTTTAAAG CGGTTACCCATTAAACGGGCGAAGAAATTCGCCGTTTGGGCGAGCAACTTCGTTATGATG GAGTACGGCAGCGGCGCAGTTATGGCGGTACCGGGGCGACGACGCGCGACTACGAGTT TGCTCTAAATACGGCTGAACATCAAACCGGTTATCCTGGCAGCTGACGGCTCTGAGC CAGATCTTTCTAGCAAGCCCTGACTGAAAAAGGCGTGTGTTCAACTCTGGCGAGTTT AACGGCTTTGACCATGAAGCGGCTTCAACGCCATCGCCGATAAACTGACTGCGATGGG CGTTGGCGAGCGTAAAGTGAACCTACCGCTCGCGACTGGGGTGTTCCTCGTACGCGTT ACTGGGGCGCGCGGATTCGATGGTGAATCTAGAAAGCGGTACCGTAATGCGCACCCCG GACGACGAGCTGCGCGTGATCTGCGCGAAGATGTGTAATGACGCGATTAACAGCCCG GATTAAAGCAGATCCGGAGTGGGCGAAACCTACCGTTAACGGTATGCCGACACTGCGTG AAACCGACACTTTCGACACCTTTATGGAGTCTGCTGGATTATGCGCGCTACACTTGC CCGAGTACAAAAGGATGCTGGATTCCGAAGCGGCTAACTACTGCTGCGCGTGGAT TATCGGATTTGGTGGTATTGAACACGCCATTATGGGGCTGCTTACTTCCGCTTCTTCC ACAACTGATGCTGATGTCAGGCAATGGTGAACCTCTGACGAACGAGCAAAAGTTGCTG TGTCAGGATATGGTCTGCGAGATGCTTCTACTATGTTGGCGAAAACGGCGAACGTAA CTGGGTTTCCCGGTTGATGCTATCGTTGAACGTGACGAGAAAGGCGGTATCGTGAAG CGAAAGATGCGCGAGGCCATGAACCTGGTTTATACCGGCATAAGCAAAATGTCCAAGTCG AAGAACAACGGTATCGACCCGAGGTGATGGTTGAACGTACGGCGCGGACACCGTTCCG TCTGTTTATGATGTTTGCCTCTCCGGCTGATGACTCTCGAATGGCAGGAATCCGCTG TGGAAAGGGCTAACCCTTCTGAAACGTGCTGGAACTGGTTTACGAGCACACAGCA AAAGGTGATGTTGCGGCACTGAACGTTGATGCGCTGACTGAAAATCAGAAAGCGCTGCG TCGCGATGTGCATAAAACGATCGCTAAAGTGACCGATGATATCGGCCGCTGCTCAGACCT TCAACACCGCAATTCGCGGATTTATGGAGTGTGAACAACTGGCGAAAGCACCACCAAC GATGGCGAGCAGGATCGCGCTCTGATGCAAGGAAGCACTGCTGGCGGTTGCTCCGATGCT TAACCGCTTACCCCGCACAATCTGCTTACGCTGTTGGCAGGAACTGAAAGGCGAAGGCG ATATCGACAACCGCGCTGGCGGTTGCTGACGAAAAGCGATGGTGAAGACTCCACG CTGGTCTGGTGCAGGTTAACGGTAAAGTCCGTGCCAAATCACCGTTCCCGTGGAGCG AACGGAAGACAGGTTTCGCGAAGCTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTG ATGGCGTTACTGTACGTAAAGTGATTTACGTACAGGTAAACTCCTCAATGTGGTCTG GGCTAAGCGGCC</p>

도면16P

서열 번호 및 설명	서열
서열 번호 45 a-아미노카프릴산 아미노아실-tRNA 합성효소 단리원- 107 핵산 서열 (야생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)	ATGCAAGAGCAATACCGCCCGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTTGAAGTAACCGAAGACGAGAGCAAGAGAGTATTACTGCCCTGCTG CTGCGCCCTATCCTTCTGGTTCGACTACACATGGGCCACGTACGTAACACACCATCGGT GACGTGATCGCCCGCTACCGAGCTATGCTGGGCAAAACGTCCTGCAGCCGATCGGCTG GGACGCGTTTGGTCTGCCGGAAGGCGGCGGCGTGAACAAACACACCGCTCCGGCAC CGTGACGTAGGACCAACATCGCGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCCGAATACTACCGTTGGGAACAGAA ATCTTTCACCGAGCTGTATAAAAAGGCTGGTATATAAGAAGACTTCTGCGGTCAACT GGTGCCCCGAACGACGACCGGTACTGGCGAACGAACAAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCTATGCAGCGTAACCTGGATCGGTCGTTCCGAAGGCGTGAGAGATCACCTTCAACGTT AACGACTATGACAAACGCTGACCGTTTACACTACCCGCCCGGACACCTTTATGGGTTG TACCTACCTGGCGGTAGCTGCCGGTTCATCCGCTGGCGCAGAAAGCGGCGGAAATATC CTGAACTGGCGGCTTTATTGACGAATGCCGTAAACACCAAGTTGCCGAAGCTGAAATG GCGAGCATGGAGAAAAAGGCGTCGATACTGGCTTTAAAGCGGTTACCCATTAACGGG CGAAGAAATTCCTGTTGGGCAGCAAACTTCGTATTGATGGAGTACGGCAGCGGCGCAG TTATGGCGGTACCGGGCACGACACGCGGACTACGAGTTTGCCCTCAAAATACGGCCTG AACATCAAAACCGGTTATCTCGCAGCTGACGGCTCTGAGCCAGATCTTTCTCAGCAAGC CCTGACTGAAAAAGGCGTGCTGTTCAACTCTGGCGAGTTCAACGGTCTTGACCATGAAG CGGCCCTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGTG AACTACCGCTGCGGACTGGGGTGTTCCTCGTCAGCGTTACTGGGGCGCGCGATTC GATGGTGACGCTGGAAGACGCTACCGTAATGCCGACCCCGGACGACGCTGCCCGTGA TCCTGCCGGAAGATGTGGTAATGGACGGCATTACGAGCCGATTAAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGSTATGCCAGCACTGCGTGAACCGACACTTTCGAGAG CTTTATGGAGTCTCTCGGCTTATGCGCGCTACACTTGCCCGCAGTACAAAGAGGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCGCGTGATATCGTTATTGGTGGTATT GAACACGCCATTATGGGGCTGCTCTACTTCCGCTTCTTCCACAACTGATCGTGTATGC AGGCATGGTGAACCTTGACGAACACGCGAAACAGTTGCTGTGTACGGGTATGGTGTG CAGATGCCCTTCTACTATGTTGGCGAAAAACGCGAACGTAACCTGGGTTTCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGGCGGTATCGTGAAAGCGAAAGATGCGGCAGGCCA TGAACCTGGTTTATACCGCATGAGCAAAATGTCCAAAGTCGAAGAACAACGGTATCGACC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTTCTGTTGTTATGATGTTTGTCT TCTCTCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGCTAACCGCTT CCTGAAACGTGTCTGGAACTGGTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAATCAGAAAGCGCTGCGTCCGATGTGATATAAAGC ATCGCTAAAGTGACCGATGATATCGGCCGTCGTCAGACCTTCAACACCGCAATTCGCGC GATTATGGAGCTGATGAACAACTGGCGAAAGCACCAACCGATGGCGAGCAGGATCGCG CTCTGCTCAGGAAGCACTGCTGGCCGTTGTCCGTATGCTTAACCCGTTACCCCGCAC ATCTGCTTCAGCTGTGGCAGGAACGAAAGGCGAAGGCGATATCGACAACGCGCCGTG GCCGTTGCTGACGAAAAAGCGATGGTGAAGACTCCACGCTGGTCTGGTGCAGGTTA ACGGTAAAGTCCGTGCCAAATCACCGTTCCGGTGGACGCAACGGAAGAACAGGTTCCG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTACGTACAGGTAACTCCTCAATCTGGTCTGTGGCTAA

도면16Q

서열 번호 및 설명	서열
<p>서열 번호 46</p> <p>a-아미노카프릴산</p> <p>아미노아실-tRNA</p> <p>합성 효소 단리된-</p> <p>168(C8-RS)</p> <p>핵산 서열</p> <p>(야생형 이. 콜리</p> <p>류실 tRNA-</p> <p>합성 효소에서</p> <p>유도됨)</p>	<p>ATGCAAGAGCAATACCGCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA</p> <p>GAAGCGCACATTGAAGTAACCGAAGACGAGAGCAAGAGAGATATTACTGCCTGTCTG</p> <p>TGATGCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT</p> <p>GACGTGATCGCCGCTACCAAGCTATGCTGGGCAAAAACGTCTGCAGCCGATCGGGTG</p> <p>GGACGCGTTTGGTCTGCCTGCGGAAGGCGCGGCGGTGAAAAACAACACCGCTCCGGGCAC</p> <p>CGTGGACGTACGACAACATCGCGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT</p> <p>TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCCGGAATACTACCGTTGGGAACAGAA</p> <p>ATTCTTACCCGAGCTGTATAAAAAGGCCCTGGTATATAAGAAGACTTTCGCGGTCAACT</p> <p>GGTGGCCGAACGACGACCGTACTGGCGAACGAACAAGTTATCGACGGCTGCTGCTGG</p> <p>CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCCGAGTGGTTTATCAAAATCACTGC</p> <p>TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA</p> <p>AAACCATGCAGCGTAACCTGGATCGGTCTGTCCGAAGCGTGGAGATCACTTCAACGTT</p> <p>AACGACTATGACAACAGCTGACCGTTTACACTACCCGCCCGGACACCTTTATGGGTTG</p> <p>TACCTACCTGCGCGTAGCTGCGGGTCACTCCGCTGGCGCAGAAAGCGCGGAAAAATAATC</p> <p>CTGAACTGGCGGCTTTTATGACGAATGCCGTAAACACCAAAGTTGCCGAAGCTGAAATG</p> <p>GCGAGATGGAGAAAAAAGCGCTCGATACTGGCTTTAAAGCGGTTCAACCATTAACGGG</p> <p>CGAAGAAATTCCCGTTTGGGCGACAAACTTCGTATTGATGGAGTACGGCACGGCGCAG</p> <p>TTATGSCGGTACCGGGGACGACGACGCGGACTACGAGTTTGCCTCTAAATACGGCTG</p> <p>AACATCAAACCGGTTATCCTGGCAGCTGACGGCTCTGAGCCAGATCTTCTCAGCAAGC</p> <p>CCTGACTGAAAAAGGCGTGTCTCAACTCTGGCGAGTTCAACGGTCTTGACCATGAAG</p> <p>CGGCCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGTG</p> <p>AACTACCGCTGCGCGACTGGGGTGTTCCTCGTCAGCGTTACTGGGGCGCGCCGATTCC</p> <p>GATGGTGACGCTGGAAAGACGGTACCGTAATGCCGACCCGGACGACGCTGCCGGTGA</p> <p>TCCTGCCGGAAGATGTGGTAATGGACGGCATTACACGCCGATTAAGCAGATCCGGAG</p> <p>TGGGCGAAACTACCGTTAACGGTATGCCAGCACTGCGTGAAACCGACACTTTGCACAC</p> <p>CTTTATGGAGTCTCTCGCTGTATGCGCGCTACACTTGCCCCGAGTACAAAGAGGTA</p> <p>TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCCGCTGGATATCCTGATTGGTGGTATT</p> <p>GAACACGCCATTATGGGGCTGCTCTACTTCCGCTTCTTCCACAACTGATGCGTGATGC</p> <p>AGGCATGGTGAACCTCTGACGAACGACGAAACAGTTGCTGTGTCAGGGTATGGTGGTGG</p> <p>CAGATGCCCTTCTACTATGTTGGCGAAAACGGCGAACGTAACTGGGTTTCCCGGTTGAT</p> <p>GCTATCGTTGAACGTGACGAGAAAGGCCGTATCGTGAAAGCGAAAGATGCGCGAGGCCA</p> <p>TGAACTGGTTTATACCGGCATGAGCAAAATGTCCAAGTCGAAGAACACCGGTATCGACC</p> <p>CGCAGTGATGGTTGAACGTTACGGCGCGGACACCGTTCTGCTGTGTTATGATGTTTGGCT</p> <p>TCTCGGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCGCTT</p> <p>CCTGAAACGTGTCTGGAACCTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAT</p> <p>TGAACGTTGATGCGCTGACTGAAAAATCAGAAAGCGCTGCCGTCCGATGTGCATAAAACG</p> <p>ATCGCTAAAGTGACCGATGATATCGGCCGCTGTCAGACCTTCAACACCGCAATTGCGGC</p> <p>GATTATGGAGCTGATGAACAACTGGCGAAAGCACCAACCGATGGCGAGCAGGATCCGC</p> <p>CTCTGATGCAGGAAGCACTGCTGGCCGTTGTCCGTATGCTTAACCGTTTACCCCGCAC</p> <p>ATCTGCTTCACGCTGTGGCAGGAACGAAAGGCGAAGGCATATCGACAACGCGCCGTG</p> <p>GCCGGTTGCTGACGAAAAAGCGATGGTGAAGACTCCACGCTGGTCTGGTGCAGGTTA</p> <p>ACGGTAAAGTCCGTGCCAAAAATCACGTTCCGGTGGACGCAACGGAAGAACAGGTTCCG</p> <p>GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA</p> <p>AGTGATTACGTACCAAGGTAACCTCTCAATCTGGTCTGGCTAA</p>

도면16R

서열 번호 및 설명	서열
서열 번호 47 a-아미노카프릴산 아미노아실-tRNA 합성효소 단리된- 2F2 핵산 서열 (아생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)	ATGCAAGAGCAATACCGCCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTTGAAGTAACCGAAGACGAGAGCAAAGAGATATTACTGCCTGTCTC ATCCTCCCTATCCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCCGCTACCGAGCTATGCTGGGCAAAACGTCTGCAGCCGATCGGGTG GGACGCGTTTGGTCTGCTCGGGAAGGCGCGGCGGTGAAAAACAACCCGATCCGGGCAC CGTGGACGTACGACAACTCGCGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACCGCGGAATACTACCGTTGGGAACAGAA ATTCTTCCCGAGCTGTATAAAAAAGGCTTGGTATATAAGAAGACTTCTCGCGTCAACT GGTGCCCGAACGACGACCGTACTGGCGAACGAACAAGTTATCGAGCGCTGCTGCTGG CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGCAGCGTAACCTGGATCGGTCTTCCGAAGCGTGGAGATCACCTTCAACGTT AACGACTATGACAAACGCTGACCGTTTACACTACCGCCCGGACACCTTTATGGGTG TACCTACCTGGCGGTAGCTGCGGGTTCATCCGCTGGCGCAGAAAGCGCGGAAAAATAATC CTGAACCTGGCGGCTTTATGACGAATGCCGTAAACACCAAAGTTGCCGAAGCTGAAATG GCGACGATGGAGAAAAAGGCGTCGATACTGGCTTTAAAGCGGTTACCCCATTAACGGG CGAAGAAATTTCCGTTTGGGCAGCAAACTTCGTATTGATGGATACGGCACGGGCGCAG TTATGGCGGTACCGGGGCACGACCGAGCGGACTACGAGTTTGCCTCTAAATACGGCTG AACATCAAACCGGTTATCCTGGCAGCTGACGGCTCTGAGCCAGATCTTCTCAGCAAGC CCTGACTGAAAAAGGCGTGTCTTCAACTCTGGCGAGTTCAACGGTCTTGACCATGAAG CGGCCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGGCTAAAGTG AACTACCGCTGCGCGACTGGGGTGTTCCTCGTCAGCGTTACTTGGGCGCGCCGATTC GATGGTGACGCTGGGAAGACGGTACCGTAATGCCGACCCGGACGACAGCTGCCGGTGA TCCTGCCGGAAGATGTGGTAATGGACGGCATTACAGCCCGATTAAGCAGATCCGGAG TGGCGCAAACTACCGTTAACGGTATGCCAGCACTGCGTGAAACCGACACTTTCGACAC CTTTATGGAGTCCTCCTGGCGCTATGCCGCTACACTTGCCTCCAGTACAAGAGGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCCGCTGGATATCATGATTGGTGGTATT GAACACGCCATTATGGGTCTGCTCTACTTCCGCTTCTTCCACAACTGATGCGTGATGC AGGCATGTTGAACCTTGACGAACCGAGCAACAGTTGCTGTGTCAGGGTATGGTGGTGG CAGATGCCCTTCTACTATGTTGGCGAAAACGGCGAACGTAACCTGGGTTTCCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGGCGGTATCGTGAAAGCGAAAGATGCGGCAGGCCA TGAACCTGGTTTATACCGGCATGAGCAAAATGTCCAACTCGAAGAACAAACGGTATCGACC CGCAGGTGATGGTTGAACGTACGGCGCGGACACCGTTCGTCTGTTTATGATGTTTGTCT TCTCCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCCGCTT CCTGAAACGTGTCTGAACTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAAATCAGAAAGCGCTGCGTCCGATGTGCATATAAACG ATCGCTAAAGTGACCGATGATATCGGCCGTCGTGACACTTCAACACCGCAATTGCGGC GATTATGGAGCTGATGAACAACTGGCGAAAGCACCAACCGATGGCGAGCAGGATCCCG CTCTGATGACGGAAGCACTGCTGGCCGTGTGTCGTATGCTTAACCCGTTACCCCGCAC ATCTGCTTACGCTGTGGCAGGAACGAAAGGCGAAGGCGATATCGACAAACGCCGCTG GCCGGTTGCTGACGAAAAAGCGATGGTGAAGACTCCACGCTGGTCTGGTGGTGCAGGTTA ACGGTAAAGTCCGTGCCAAAAATCACCGTTCCGGTGGACGCAACCGGAAGAACAGGTTCCG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTACGTACAGGTAACTCCTCAATCTGGTCTGTGGCTAA

도면16S

서열 번호 및 설명	서열
서열 번호 48 a-아미노카프릴산 아미노아실-tRNA 합성효소 단리된- 2F5 핵산 서열 (야생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)	ATGCAAGAGCAATACCGCCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCCGACATTTGAAGTAACCGAAGACGAGAGCAAGAGAGTATTACTGCGCTGTCTG TGTATCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACACACCATCGGT GAGGTGATCGCCCGCTACGAGCGTATGCTGGGCAAAACGTCCTGCGCCGATCGGCTG GGACGGCTTTGGTCTGCCCTGCGGAAGGCGCGGCGGTGAAAAACAACACCCGCTCCGGCAC CGTGGACGTACGACAACATCGCGTATATGAAAAACGAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCCGGAATACTACCGTTGGGAACAGAA ATTCTTCAACCGAGCTGTATAAAAAAGGCCGTGGTATATAAGAAGACTTCTGCGGTCAACT GGTGCCCGAACGACGAGCCGTAAGTGGCGAACGAACAAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCGCGAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGCGAGCGTAACGTGATCGGTCTGTTCCGAAGGCGTGGAGATCACTTCAACGTT AACGACTATGACCAACGCTGACCGTTTACACTACCCGCCCGGACACCTTTATGGGTG TACCTACCTGGCGGTAGCTGCGGGTCATCCGCTGGCGCAGAAAGCGCGGAAAAATATC CTGAACCTGGCGGCTTTTATTGACGAATGCCGTAACACCAAAGTTGCCGAAGCTGAAATG GCGACGATGGAGAAAAAGGCGTCGATCTGCTTTTAAAGCGGTTACCCCATTAACGGG CGAAGAAATTCCTGTTTGGGCAGCAAACTTCGTATTGATGGAGTACGGCACGGGCGCAG TTATGGCGGTACCGGGGCACGACGCGCGACTACGAGTTTGCCTCTAAATACGGCCCTG AACATCAAAACCGGTTATCTTGGCAGCTGACGGCTCTGAGCCAGATCTTTCTCAGCAAGC CCTGACTGAAAAAGCGGTGCTGTCTCAACTCTGGCGAGTTCAACGGTCTTGACCATGAAG CGGCTTTCAACGCCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGTG AACTACCGCTGCGCGACTGGGGTGTTCCTCCGTCAGCGTTACTGGGCGCGCGGATTC GATGGTGACGCTGGGAAGACGGTACCGTAATGCGCGACCCCGGACGACGAGTCCGGGTGA TCCTGCCCGAAGATGTGGTAATGGACGGCATTACAGCCCGATTAAAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGGTATGCCAGCACTGCGTGAAACCGACACTTTTCGACAC CTTTATGGAGTCTCTTGGCTGTATGCGCGCTACACTTGCCCGCAGTACAAAGAAGGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCGCGGTGGAATCTCTGATGGTGGTATT GAACACGCCATTATGGGTCTGCTCTACTTCCGCTTCTTCCCAAACTGATGCGTGATGC AGGCATGGTGAACCTGACGAACGAGCAACAGTGTCTGTGTCAGGGTATGGTGCTGG CAGATGCCCTTCTACTATGTGGCGAAAAAGCGCAACGTAACCTGGGTTTCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGCCGTATCGTGAAAGCGAAAGATGCGGCGAGGCCA TGAACCTGGTTTATACCGCATGAGCAAAATGTCCAGTCGAAGAACACCGGTATTCGACC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTTTCGTCTGTTTATGATGTTTTCCT TCCTCCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCGCTT CCTGAAACGTGTCTGGAACCTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAAAACGAAAGCGCTGCGTCGCGATGTGCTATAAACG ATCGCTAAAGTGACCGATGATATCGGCCGTGCTCAGACCTTCAACACCGCAATTGCGGC GATTATGGAGCTGATGAACAACTGGCGAAAGCACCAACCGATGGCGAGCAGGATCGCG CTCTGATGCGAGGAGCACTGCTGGCGGTTGTCCGTATGCTTAACCGGTTACCCCGCAC ATCTGCTTCACGCTGTGGCAGGAACGAAAGCGAAGCGGATATCGACAAACGCGCGGTG GCCGGTTGCTGACGAAAAAGCGATGGTGAAGACTCCACGCTGGTCTGGTGCAGGTTA ACGGTAAGTCCGTGCCAAATCACCCTTCCGGTGGACGCAACGGAAGAACAGGTTCCG GACGCTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTTACGTACCGGTAACCTCCTCAATCTGGTCTGTTGGCTAA

도면16T

서열 번호 및 설명	서열
<p>서열 번호 49</p> <p>0-메틸 티로신 아미노아실-tRNA 합성효소 단리된 -3A7 (0MeYRS) 핵산 서열 (야생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>ATGCAAGAGCAATACCGCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTTGAAGTAACGGAAGACGAGAGCAAAGAGATATTACTGCCTGTCTT TGGAGCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCCGCTACCAAGCGTATGCTGGGCAAAACGCTCCTGCAGCCGATCGGCTG GGACGCGTTTGGTCTGCCCTGCGGAAGGCGGGCGGTGAAAAACAACACCGCTCCGGCAC CGTGGACGTACGACCAACATCGGTAATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCGGGAATACTACCGTTGGGAACAGAA ATCTTTCACCGAGCTGTATAAAAAAGGCTTGGTATATAAGAAGACTTCTGCGGTCAACT GGTCCCCGAACGACGACCGTACTGGCGAACGAACAGTTATCGAGCGCTGCTGCTGG CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGCAGCGTAACCTGGATCGGTCTTCCGAAGGCGTGGAGATCACCTTCAACGTT AACGACTATGACAACACGCTGACCGTTTACACTACCGCCCGGACACCTTTATGGGTG TACCTACCTGGCGGTAGCTGCGGTCATCCGCTGGCGCAGAAAGCGCGGAAAAATAATC CTGAACTGGCGGCTTTATGACGAATGCCATAACACCAAAGTTGCCGAAGCTGAAATG GCGAGCATGGAGAAAAAGGCGTCCGATCTGGCTTTAAAGCGGTTCACCCATTACCGGG CGAAAGAAATTCCTGTTTGGCGACGAACTTCGTATTGATGGAGTACGGCACGGGCGCAG TTATGGCGGTACCGGGGACGACCGAGCGGACTACGAGTTTGCCTCTAAATACGSCCTG AACATCAAACCGGTTATCTGGCAGCTGACGGCTCTGAGCCAGATCTTCTCAGCAAGC CCTGACTGAAAAAGCGTGTCTTCAACTCTGGCGAGTTCAACGGTCTTGACCATGAAG CGGCCCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGTG AACTACCGCTGCGCGACTGGGGTGTTCCTCGTCAGCGTTACTTGGGGCGCGCGGATTC GATGGTGACGCTGGAAGACGGTACCGTAATGCCGACCCCGGACGACCGCTGCCCGTGA TCTTCCCGGAAGATGTGGTAATGGACGGCATTACGAGCCGATTAAGCAGATCCGGAG TGCGCGAAACTACCGTTAACGGTATGCCAGCACTGCGTGAAACCGACACTTTCGACAC CTTTATGGAGTCTTCTGGCGTTATCGCGCTACACTTCCCGCAGTACAAAGAGGTA TGCTGGATTCCGAAGCGCTAACTACTGGCTGCCCGTGGATATCGCTATTGGTGGTATT GAACACGCCATTATGGGTCTGCTCTACTTCCGCTTCTTCCACAAACTGATGCGTGATGC AGGCATGTGTAACCTCTGACGAACCGGAAACAGTTGCTGTCTCAGGGTATGGTGTGG CAGATGCCCTTCTACTATGTTGGCGAAAACGGCGAACGTAACGGGTTTCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGGCCGTATCGTGAAGCGAAAGATGCGGCGAGGCCA TGAAGTGGTTTATACCGCATGAGCAAAATGTCCAGTCAAGAACAAACGGTATCGACC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTTCGTCTGTTTATGATGTTTGTCT TCTCCGGCTGATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCGCTT CCTGAAACGTGTCTGGAACCTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAATCAGAAAGCGCTGCGTCCGATGTGCATATAAACG ATCGCTAAAGTGACCGATGATATCGGCCGTCGTCAGACCTTCAACACCGCAATTGCGGC GATTATGGAGCTGATGAACAACTGGCGAAAGCACCAACCGATGGCGAGCAGSATCGCG CTCTGATGACGGAAGCACTGCTGGCCGTTGTCCGTATGCTTAACCCGTTACCCCGCAC ATCTGCTTACGCTGTGGCAGGAACGAAAGGCGAAGGCGATATCGACAACCGCGCTG GCCGGTGTGACGAAAAAGCGATGGTGAAGACTCCACGCTGGTCTGGTGCAGGTTA ACGGTAAAGTCCGTGCCAAATCACCCTTCCGGTGGACGCAACGGAAGAACAGGTTACG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTTACGTACCAGGTAACCTCCTCAATCTGGTCTGTGGCTAA</p>

도면16U

서열 번호 및 설명	서열
서열 번호 50 O-메틸 티로신 아미노아실 -tRNA 합성효소 단리된-3A2 핵산 서열 (야생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)	ATGCAAGAGCAATACCGCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTGAAAGTAACCGAAGACGAGAGCAAGAGAAAGTATTACTGCGCTGTCTA TGGAGCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCCGCTACCGAGCTATGCTGGGCAAAAACGTCCTGCAGCCGATCGCGTG GGACGCGTTTGGTCTGCCTGCGGAAGGCGCGCGGTGAAAAACAACACCGCTCCGGCAC CGTGGACGTACGACAACTACCGGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGAGCTGGCAACCTGTACGCCGGAATACTACCGTTGGGAACAGAA ATCTTTCACCGAGCTGTATAAAAAGGCTGGTATATAAGAAGACTTCTCGGTCACCT GGTGCCCCGAACGACGACCGTACTGGCGAAGCAAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGCAGCGTAACCTGGATCGGTCGTTCCGAAGGCGTGGAGATCACCTTCAACGTT AACGCTATGACAAACGCTGACCGTTTACACTACCCGCCCGGACACCTTTATGGGTTG TACCTACCTGGCGGTAGCTGCGGGTCACTCCGCTGGCGCAGAAAGCGGCGGAAAAATATC CTGAACCTGGCGGCTTTATTGACGAATGCCGTAAACCAAAGTTGCCGAAGCTGAAATG GCGACGATGGAGAAAAAGCGCTCGATACTGGCTTTAAAGCGGTTCAACCTTAACGGG CGAAGAAATTCCCGTTTGGGCGACAACTTCGTTATGATGGAGTACGGCAGGCGCGCAG TTATGGCGGTACCGGGGACGACGACGCGGACTACGAGTTTGGCTCTAAATACGGCTG AACATCAAAACCGGTTATCCTGGCAGCTGACGGCTCTGAGCCAGATCTTTCTCAGCAAGC CCTGACTGAAAAAGGCGTGCTGTCAACTCTGGCGAGTTCAACGGCTTTGACCATGAAG CGGCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAGTG AACTACCGCTGCGCGACTGGGGTGTTCCTCGTCAGCGTTACTGGGCGCGCGGATTC GATGGTGACGCTGGAAGACGGTACCGTAATGCCGACCCCGGACGACGAGCTGCCGCTGA TCCTGCCGGAAGATGTGGTAATGGACGGCATTACAGCCCGATTAAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGGTATGCCAGCACTGCCGTGAAACCGACACTTTCGACAC CTTTATGGAGTCCCTCGCGGTTATGCGCGCTACACTTGCCTCGCAGTACAAAGAAGGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCCGTTGGATATCTTTATGTGTGTTAT GAACACGCCATTATGGGGCTGCTACTTCCGCTTCTTCCACAACTGATGCGTGATGC AGGCATGTGTAACCTCTGACGAACAGCGAAACAGTTGCTGTGTCAGGGTATGGTGTGG CAGATGCCCTTCTACTATGTTGGCGAAAAAGGCGAACGTAACCTGGGTTTCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGGCGGTATCGTGAAAGCGAAAGATGCCGCGAGGCA TGAACGTGGTTTATACCGGCATGAGCAAAATGTCCAAGTCGAAGAACACCGGTATCGACC CGCAGGTGATGGTTGAACGTACCGCGCGGACACCGTTCTGCTGTTTATGATGTTTGGT TCTCGCGTGATATGACTCTCGAATGGCAGGAATCCGCTGTGGAAGGGGCTAACCGCTT CCTGAAACGTGCTTGGAACTGGTTTACGAGCACACAGCAAAAGGTGATGTGCGGCAC TGAACGTTGATGCGCTGACTGAAAAACAGAAAGCGCTGCGTCGCGATGTGCATAAAACG ATCGCTAAAGTGACCGATGATATCGGCGCTGCTCAGACCTTCAACACCGCAATTTCGGC GATTATGGAGCTGATGAACAACTGGCGAAAGCACCACCGATGGCGAGCAGGATCGCG CTCTGATGCAGGAAGCACTGCTGGCCGTGTCCGTATGCTTAAACCGTTTCAACCCGCAC ATCTGCTTCAACGCTGTGGCAGGAACGAAAGGCGAAGGCGATATCGACAAACGCGCGTG GCCGTTGCTGACGAAAAAGCGATGGTGGAAAGACTCCACGCTGGTGTGTCAGGTTA ACGGTAAAGTCCGTGCCAAATCACCCTTCCGTTGGACGCAACGGAAGAACAGGTTCCG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTGATGGCGTTACTGTACGTAA AGTGATTTACGTACCAGGTAACTCCTCAATCTGGTCTGGCTAA

도면16V

서열 번호 및 설명	서열
서열 번호 51 0-메틸 티로신 아미노아실 -tRNA 합성효소 단리된-3F11 핵산 서열 (야생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)	ATGCAAGAGCAATACCGCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTGAAAGTAACCGAAGACGAGAGCAAGAGAGTATTACTGCCTGCTT TGGAGCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCCGCTACCAAGCTATGCTGGGCAAAAACGCTCTGCAGCCGATCGGCTG GGACGCGTTTGGTCTGCTTGGGAAGGCGGGCGGTGAAAAACAACACGCTCCCGGCAC CGTGGACGTACGACAACTACCGGTATATGAAAAACGAGCTCAAAATGCTGGGCTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCCGGAATACTACCGTTGGGAACAGAA ATTCTTCAACCGAGCTGTATAAAAAAGCCCTGGTATATAAGAAGACTTCTGCGGTCAACT GGTGCCCGAACGACGACAGCCGTAAGTGGCGAACGAACAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAAGTTGAACGTAAAGAGATCCGCGAGTGGTTTATCAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGACGCGTAAGTGGATCGGTCTGTTCCGAAGCGTGGAGATCACTTCAACGTT AACGACTATGACAAACGCTGACCGTTTACACTACCCGCCCGGACACCTTTATGGGTG TACCTACCTGGCGGTAGCTGCGGGTCATCCGTGGCGCAGAAAGCGCGGAAATATATC CTGAACCTGGCGCCTTTATTGACGAATGCCGTAAACCAAAAGTTGCCGAAGCTGAAATG GCGACGATGGAGAAAAAGCGCTCGATACTGGCTTTAAAGCGGTTCAACCATTAACGGG CGAAGAAATTCCTGTTTGGGCGACAACTTCGTATTGATGGAGTACGGCACGCGCGCAG TTTATGGCGTACCGGGGCGACGACGCGGACTACGAGTTTGCCTCTAAATACGGCCTG AACATCAAACCGGTTATCCTGGCAGCTGACGCTCTGAGCCAGATCTTCTCAGCAAGC CCTGACTGAAAAAGCGGTGCTGTCAACTCTGGCGAGTTCAACGGTCTTGACCATGAAG CGGCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCTAAAGTG AACTACCGCTGCGGACTGGGGTGTTCCTCGTCAGCGTTACTGGGGCGCGCCGATTCC GATGGTGACGCTGGAAGACGGTACCGTAATGCCGACCCCGGACGACCGCTGCGGGTGA TCCTGCCGGAAGATGTGTTAATGGACGGCATTACGACCCGATTAAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGATATGCCAGCACTGCGTGAAACCGACACTTTGACAC CTTTATGGAGTCCCTCTGGCGTTATGCGCGCTACACTTGGCCGAGTACAAAGAGGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGGCGGTGGATATCTGTATTGTTGGTATT GAACACGCCATTATGGGCTGCTCTACTTCCGCTTCTTCCACAAACTGATGCGTGATGC AGGCAATGGTGAACCTCTGACGAACGAGCGAAACAGTTGCTGTGTCAGGGTATGGTGTGG CAGATGCCCTTCTACTATGTTGGCGAAAAAGCGGAAACGTAACCTGGGTTTCCCGGTGAT GCTATCGTTGAACGTGACGAGAAAGCCGTTATCGTGAAGCGAAAGATGCGGACGGCCA TGAACCTGGTTTATACCGGCATGAGCAAAATGTCCAAGTCAAGAAACAACGGTATCGACC CGCAGGTGATGGTTGAACGTACCGCGCGGACACCGTTCTGTTGTTATGATGTTTGGCT TCTCCGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCGCTT CCTGAAACCTGTCTGGAACCTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAAAATCAGAAAGCGCTGCGTGCATGTGATATAAACG ATCGCTAAAGTGACCGATGATATCGGCGCTCGTCAGACCTTCAACACCGCAATTCGCGGC GATTATGGAGTGATGAACAACTGGCGAAAGCACCACCGATGGCGAGCAGGATCCGCG CTCTGATGCAGGAAGCACTGCTGGCCGTTGTCGGTATGCTTAACCGTTTACCCCGCAC ATCTGCTTCAACGCTGTGGCAGGAACGAAAGGCGAAGGCGATATCGACAACCGCCGCTG GCCGGTTGCTGACGAAAAAGCGATGGTGGAAAGACTCCACGCTGGTCTGGTGAGGTTA ACGGTAAAGTCCGTCGCAAAATCACCCTTCCGGTGGACGCAACGGAAGAACAGGTTCCG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTACGTACCGGTAACCTCTCAATCTGGTCTGGGTAA

도면16W

서열 번호 및 설명	서열
서열 번호 52 0-메틸 티로신 아미노아실-tRNA 합성효소 단리된 -3E7 핵산 서열 (야생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)	ATGCAAGAGCAATACCGCCCGAAGAGATAGAATCCAAGTACAGCTTCATTGGGATGA GAAGCGCACATTGAAAGTAACCGAAGACGAGAGCAAGAGAAGTATTACAGCTGCTGCTT TTGAGCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCCGCTACCGAGCGTATGCTGGGCAAAAACGTCTGCAGCCGATCGGCTG GGACGCGTTTGGTCTGCCTGCGGAAGGCGGCGGTGAAAAACAACACCGCTCCGGCAC CGTGGACGTACGACAACATCGCGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCCGGAATACTACCGTTGGGAACAGAA ATTCTTACCGAGCTGTATAAAAAGGCTTGGTATATAAGAAAGACTTCTGCGGTCAACT GGTGCCCGAAGCAGACCGCTACTGGCGAAGCAACAAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGCGAGCTAACTGGATCGGTGCTTCCGAAGGCGTGGAGATCACTTCAACGT AACGATATGACAACACGCTGACCGCTTACACTACCGCCCGGACACCTTTATGGGTTG TACCTACCTGGCGGTAGCTGCGGGTCATCCGCTGGCGCAGAAAGCGCGGAAAAATAATC CTGAACCTGGCGGCTTTATTTGACGAATGCCGTAAACCAAGTTGCCGAAGCTGAAATG GCGACGATGGAGAAAAAGGCGTCGATACCTGGCTTTAAAGCGGTTACCCATTAACGGG CGAAGAAATTCCTGTTGGGCGAGCAAACTTCGTATTGATGGAGTACGGCACGGGCGCAG TTTATGGCGGTACCGGGCAGACACGCGGACTACGAGTTTGCCTCTAAATACGGCCTG AACATCAAAACCGGTTATCCTGGCAGCTGACGGCTCTGAGCCAGATCTTTCTCAGCAAGC CCTGACTGAAAAAGGCGTGTCTCAACTCTGGCGAGTTCAACGGCTTTGACCATGAAG CGGCCCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGTG AACTACCGCTGCGCGACTGGGGTGTTCCTGTCAGCGTTACTGGGCGCGCGGATTC GATGGTGACGCTGGAAGACGGTACCGTAATGCCGACCCCGGACGACGAGTCCCGGTGA TCCTGCGGAAAGATGTGGTAAATGGACGGCATTACAGCCCGATTAAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGGTATGCCAGCACTGCGTGAAACCGACACTTTCCGACAC CTTTATGGAGTCCCTCGCGCTTATGCGCGCTACACTTGGCCCGAGTACAAAGAAAGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCCGCTGGATATCAGATTTGGTGTATT GAACACGCCATTATGGGTCTGCTCTACTTCCGCTTCTCCACAACATGATGGGTGATGC AGGCATGGTGAACCTCTGACGAACGAGCAAAACAGTTGCTGTGTCAGGGTATGGTCTGG CAGATGCCCTTCTACTATGTTGGCGAAAAACGGCGAACGTAACTGGGTTTCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGGCCGTATCGTGAAAGCGAAAGATGCGCGAGGCCA TGAACGTGGTTTATACCGGCATGAGCAAAATGTCCAAGTCAAGAAACAACGGTATCGACC CGCAGGTGATGGTTGAACGTACCGCGCGGACACCGTTCTGTTTATGATGTTTGGCT TCTCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCGCTT CCTGAAACGTGTCTGGAAACTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAAAATCAGAAAGCGCTGCGTCGCGATGTGCATAAAAACG ATCGCTAAAGTGACCGATGATATCGGCCGTCTGTAGACCTTCAACACCGCAATTGCGGC GATTTATGGAGCTGATGAACAACTGGCGAAAGCACCAACCGATGGCGAGCAGGATCCCG CTCTGATGCAGGAAGCACTGCTGGCCGTTGTCGGTATGCTTAACCCGTTACCCCGCAC ATCTGCTTCAACGCTGTGGCAGGAACGAAAGCGGAGCGGATATCGACAAACGCGCGGTG GCCGGTTGCTGACGAAAAAGCGATGGTGGAAAGACTCCACGCTGGTCTGGTGCAGGTTA ACGGTAAAGTCCGTGCCAAATCACCCTTCCGGTGGACGCAACGGAAGAACAGGTTCCG GAACGTGCTGGCGAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTTACGTACCAGGTAACCTCCTCAATCTGGTCTGGCTAA

도면16X

서열 번호 및 설명	서열
서열 번호 53 o-니트로벤질 시스테인 아미노아실 -tRNA 합성효소 단리원-1A3 핵산 서열 (야생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)	ATGCAAGAGCAATACCGCCCGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTTGAAGTAACCGAAGACGAGAGCAAGAGAGTATTACTGCCGTGCTG GGGAGCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACACACCATCGGT GACGTGATCGCCCGCTACCAGCGTATGCTGGGCAAAAACGTCCTGCAGCCGATCGGGTG GGACGCGTTTGGTCTGCCCTGCGGAAGGCGCGCGGTGAAAAACAACACCGGTCCGGGCAC CGTGGACGTACGACAACATCGCGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCCGAGCTGGCAACCTGTACGCCGGAATACTACCGTTGGGAACAGAA ATTCTTCACCGAGCTGTATAAAAAAGGCTTGGTATATAAGAAAGACTTCTGCCGTCAACT GGTGCCCGAAGCAGACGACCGTACTGGCGAAGCAACAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAGTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGGTA AAACCATGCAGCGTAACCTGGATCGGTCTGTTCCGAAGGCGTGGAGATCACCTTCAACGTT AACGACTATGACAACAGCTGACCGTTTACACTACCGCCCGGACACCTTTATGGGTTG TACCTTACCTGGCGGTAGCTGCGGGTCTCCGCTGGCGCAGAAAGCGCGGAAAAATAATC CTGAATCGCGGCCCTTTATGACGAATGCCGTAAACCAAAAGTTGCCGAAGCTGAAATG CGGACGATGGAGAAAAAGGCGTCTGATCTGCTTTTAAAGCGGTTTACCCATTAAACGGG CGAAGAAATTCCCGTTTGGGCAGCAAACTTCGTATTGATGGAGTACGGCAGCGGCCGAG TTATGGCGGTACCGGGGACGACCGCGGACTACGAGTTTGCCCTCAAATACGGCCTG AACATCAAAACCGGTTATCTGGCAGCTGACGGCTCTGAGCCAGATCTTCTCAGCAAGC CCTGACTGAAAAAGGCGTCTGTTCAACTCTGGCGAGTTCAACGGCTTGACCATGAAG CGGCCCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGTG AACTACCGCCTGCGCGACTGGGCTGTTCCCGTCAGCGTTACTTGGGCGCGCCGATTCC GATGGTGACGCTGGAAGACGGTACCGTAATGCCGACCCGGACGACCGCTGCCGGTGA TCCTGCCGGAAGATGTGGTAATGGACGGCATTACAGCCCGATTAAAGCAGATCCGGAG TGGCGCAAACTACCGTTAACGGTATGCCAGCACTGCGTGAAACCGACACTTTCGACAC CTTTATGGAGTCTCTTGGCGGTATGCCGCGTACACTTGGCCGAGTACAAGAGGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCCGCTGGATATCCTGATTGGTGGTATT GAACACGCCATTATGGGTCTGCTCTACTTCCGCTTCTTCCACAACTGATCGGTGATGC AGGCATGGTGAACCTCTGACGAACCGGAAACAGTTGCTGTGTCAGGGTATGGTGTGCG CAGATGCCCTTCTACTATGTTGGCGAAAACGGCGAACGTAACGGGTTTCCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGGCCGTATCGTGAAAGCGAAAGATGCCGCGAGGCCA TGAACCTGGTTTATACCGGCATGAGCAAAATGTCCAAGTCGAAGAACAAACGGTATCGACC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTCTGCTGTTTATGATGTTTGTCT TCTCGGCTGATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCCGCTT CCTGAAACGTGTCTGGAACCTGGTTTACGAGCACACGAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAAAATCAGAAAGCGCTGCGTCCGATGTGCATATAAACG ATCGCTAAAGTGACCGATGATATCGGCCGCTGCTGAGACCTTCAACACCGCAATTGCCGGC GATTATGAGCTGATGAACAACTGGCGAAAGCAACCAACCGATGCGCAGCAGGATCGCG CTCTGATGCAGGAAGCACTGCTGGCCGTTGTCCGTAATGCTTAACCCGTTCAACCCGCAC ATCTGCTTCACGCTGTGGCAGGAACGTAAGGGCGAAGGCATATCGACAACCGCCGCTG GCCGGTTGCTGACGAAAAAGCATGGTGAAGACTCCACGCTGGTCTGGTGCAGGTTA ACGGTAAAGTCCGTGCCAAATCACCCTTCCGGTGGACGCAACGGGAAGAACAGGTTCCG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTTAA AGTGATTTACGTACCGGTAACTCCTCAATCTGGTCTGGCTAA

도면16Y

서열 번호 및 설명	서열
<p>서열 번호 54</p> <p>o-니트로벤질 시스테인 아미노아실-tRNA 함성효소 단리된 -3A12 렉산 서열 (야생형 이.콜리 류실 tRNA -함성효소에서 유도됨)</p>	<p>ATGCAAGAGCAATACCGCCCGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTTGAAGTAACCGAAGACGAGAGCAAAGAGATTTACTGCTGTCTG GTTGGCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCCGCTACCAGCGTATGCTGGGCAAAAACGCTCTGCAGCCGATCGGCTG GGACGCGTTTGGTCTGCTTGGGAAGGCGGCGGCTGAAAAACAACCGCTCCGCGCAC CGTGGACGTACGACAACATCGCGTATATGAAAAACCAAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCGGAAATACTACCGTTGGGAACAGAA ATTCTTACCGAGCTGTATAAAAAAGGCTGGTATATAAGAGACTTCTGCGGTCAACT GGTGGCCGACGACGACCGGTACTGGCGAAGCAACAGTTATCGACGGCTGCTGCTGG CGCTGGGATACCAAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGCAGCGTAACCTGGATCGGTCTGTTCCGAAGGCGTGGAGATCACCTTCAACGTT AACGACTATGACAAACGCTGACCGTTTACACTACCGCCCGGACACCTTTATGGGTG TACCTACCTGGCGTAGCTGCGGCTCATCCGCTGGCGCAGAAAGCGCGGAAAAATAATC CTGAAGTGGCGGCTTTATGACGAATGCCGTAAACCAAAGTTGCCGAAGCTGAAATG GCGACGATGGAGAAAAAGGCGTCGATACTGGCTTTAAAGCGGTTACCCATTAAACGGG CGAAGAAATTCCTGTTGGGCGAGCAAACTTCGTATTGATGGAGTACGGCAGGCGCGCAG TTATGGCGGTACCGGGGACGACGACGCGGACTACGAGTTTGCCTCTAAATACGGCCTG AACATCAAAACCGGTTATCTTGGCAGCTGACGGCTCTGAGCGAGATCTTCTCAGCAAGC CCTGACTGAAAAAGGCGGTGCTGTTCAACTCTGGCGAGTTCAACGGTCTTGACCATGAAG CGGCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGCTAAAGTG AACTACCGCTGCGCGACTGGGGTGTTCCTCGTCAGCGTTACTGGGGCGCGCCGATTC GATGTTGACGCTGGAAGACGGTACCGTAATCCGACCCCGGACGACGCTGCCCGTGA TCTTCCCGGAAGATGTGGTAATGGACGGCATTACAGCCGATTAAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGGTATGCCAGCACTGCGTGAACCGACACTTTTCGAGAC CTTTATGGAGTCTCTCGGCTTATGCGCGCTACACTTGCCTCGAGTACAAAGAGGTA TGCTGGATTCGAAGCGGCTAACTACTGGCTGCCCGTGGATATCTTTATGGTGGTATT GAACACGCCATTATGGGTCTGCTCTACTTCCGCTTCTTCCACAAACTGATGCGTGATGC AGGCATGGTGAACCTTGACGAACGAGCGAAACAGTTGCTGTGTCAGGGTATGGTGTGG CAGATGCCCTTCTACTATGTTGGCGAAAACGGCGAACGTAACTGGGTTTCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAAGCCGATCGTGAAAGCGAAAGATGCGGCGAGGCCA TGAACCTGGTTTATACCGCATGAGCAAAATGTCCAGTCAAGAACCAACGGTATCGACC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTCTGTTGTTATGATGTTGCT TCTCCGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGCTAACCGCTT CCTGAAACGTGTCTGGAACCTGGTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAAAATCAGAAAGCGCTGCGTCCGATGTGCTATAAACG ATCGTAAAGTGACCGATGATATCGGCCGTGTCAGACCTTCAACACCGCAATTGCGGC GATTATGGAGCTGATGAACAACTGGCGAAAGCAACCGATGGCGAGCAGGATCGCG CTCTGATGCAGGAAGCACTGCTGGCCGTTGTCCGTATGCTTAACCCGTTACCCCGCAC ATCTGCTTCAACGCTGTGGCAGGAATGAAAGGCGGAAGGCGATATCGAACACGCGCGTG GCCGTTGCTGACGAAAAAGCGATGGTGAAGACTCCACGCTGGTCTGGTGCAGGTTA ACGGTAAAGTCCGTGCCAAATCACCGTTCCGGTGGACGCAACGGAAGAACAGGTTGCG GACCTGCTGCCAGGAACATCTGGTAGCAAAATATCTTGATGGGCTTACTGTACGTAA AGTGATTTACGTACCAGGTAACCTCCTCAATCTGGTCTGTGGCTAA</p>

도면162

서열 번호 및 설명	서열
서열 번호 55 o-니트로벤질 시스테인 아미노아실-tRNA 합성효소 단리된 -3H11 (nbCRS) 렉산 서열 (야생형 이. 폴리 류실 tRNA- 합성효소에서 유도됨)	ATGCAAGAGCAATACCGCCCGGAAGAGATAGAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTTGAAGTAACCGAAGACGAGAGCAAAAGAGAGTATTACTGCCTGTCTT GGTGCGCCCTATCCTTCCTGGTTCGACTACACATGGGCCACGTACGTAACTACACCATCGGT GACGTGATCGCCCGCTACCGAGCTATGCTGGGCAAAAACGTCTGCAGCCGATCGGGTG GGACGCGTTTGGTCTGCTTGGGAAGGCGGGCGGTGAAAAACAACACCGCTCCGGGCAC CGTGGACGTACGACAACATCGCGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCGGAATACTACCGTTGGGAACAGAA ATTCTTACCGGAGCTGTATAAAAAAGGCTTGTATATAAGAGACTTCTGCGGTCAACT GGTGCCCGAACGACGACCGTACTGGCGAACGAACAGTTATCGACGGCTGCTGCTGG CGCTGCGATACCAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCACTGC TTACGCTGACGAGCTGCTAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGCAGCGTAACGTGATCGGTCTGTTCCGAAGCGTGGAGATCACTTTCAACGTT AACGACTATGACAACAGCTGACCGTTTACACTACCGCCCGGACACCTTTATGGGTG TACCTACCTGGCGGTAGCTGCGGGTCATCCGCTGGCGCAGAAAGCGCGGAAATAATC CTGAACCTGGCGGCTTTTATGACGAATGCCGTAAACACCAAGTTGCCGAAGCTGAAATG GCGACGATGGAGAAAAAGGCGTCGATCTGGCTTTAAAGCGGTTCACCATTAACGGG CGAAGAAATCCCGTTTGGGCGAGCAAACTTGTATGATGGAGTACGGCAGCGGCGAG TTATGGCGGTACCGGGCACGACCGCGACTACGAGTTTGCTCTAAATACCGGCTG AACATCAAAACCGGTTATCTGGCAGCTGACGCTCTGAGCCAGATCTTCTCAGCAAGC CCTGACTGAAAAAGGCTGCTGTTCACACTCTGGCGAGTTCAACGGTCTTGACCATGAAG CGGCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGTG AACTACCGCTGCGCGACTGGGGTGTTTCCCGTCAGCGTTACTGGGCGCGCCGATTC GATGGTGACGCTGGAAGACGGTACCGTAATGCCGACCCCGGACGACGCTGCCCGTGA TCCTGCCGGAAGATGTGGTAATGGACGGCATTACAGCCCGATTAAGCAGATCCGGAG TGGGCGAAACTACCGTTAACGGTATGCCAGCACTGCGTGAAACCGGACACTTTGACAC CTTTATGGAGTCTTCTGGATTATGCGCGCTACACTTGCCCGCAGTACAAGAGAGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCGGCTGGATATCGCGATTGGTGGTATT GAACACGCCATTATGGGGCTGCTCTACTTCCGCTTCTTCCACAACTGATGCGTGATGC AGGCATGTTGAACCTCTGACGAACAGCGAAACAGTTGCTGTGTACGGGTATGGGTGCTGG CAGATGCCCTTCTACTATGTTGGCGAAAACGGCGAACGTAACCTGGGTTTCCCGGTTGAT GCTATCGTTGAACGTGACGAGAAAGGCGGTATCGTGAAGCGAAAGATGCGGCAGGCCA TGAACGTGGTTTATACCGCATGAGCAAAATGTCCAAGTCGAAGAACACGGGTATCGACC CGCAGGTGATGGTTGAACGTACGCGCGGACACCGTTCGTCTGTTTATGATGTTTGTCT TCTCCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGGCTAACCGCTT CCTGAAACGTGCTGGAACCTGGTTTACGAGCACACAGCAAAAGGTGATGTTGCGGCAC TGAACGTTGATGCGCTGACTGAAAATCAGAAAGCGCTGCGTCGCGATGTGCATAAAACG ATCGCTAAAGTGACCGATGATATCGGCCGTCGTGAGACCTTCAACACCGCAATTGCGGC GATTATGGAGTGATGAACAACTGGCGAAAGCACCAACCGATGGCGAGCAGGATCGCG CTCTGATGACGGAAGCACTCTCGCCGTTGTCCGTATGCTTAAACCGCTTCAACCGGCAC ATCTGCTTCAAGCTGTGGCAGGAACCTGAAAGGCGAAGCGGATATCGACAACGCGCGTG GCCGGTTGCTGACGAAAAAGCGATGGTGAAGACTCCACGCTGGTCTGGTGCAGGTTA ACGGTAAAGTCCGTGCCAAAAATCACCGTTCCGGTGGACGCAACGGAAAGAACAGGTTCCG GAACGTGCTGGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCGTTACTGTACGTAA AGTGATTTACGTACAGGTAACTCCTCAATCTGGTCTGTGGCTAA

도면16AA

서열 번호 및 설명	서열
<p>서열 번호 56</p> <p>o-니트로벤질 시스테인 아미노아실-tRNA 합성효소 단리된 -4E1 핵산 서열 (야생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>ATGCAAGAGCAATACCGCCCGGAAGAGATAGAAATCCAAAGTACAGCTTCATTGGGATGA GAAGCGCACATTTGAAGTAACCGAAGACGAGAGCAAGAGAAGTATTACTGCCTGTCTG GTACGCCCTATCCTTCTGGTCGACTACACATGGGCCACGTACGTAACCTACACCATCGGT GACGTGATCGCCCGCTACACAGCGTATGCTGGGCAAAACGTCCTGCAGCCGATCGGGT GGACCGCTTTCGTCTGCTTCCGGAAGGCGCGCGGTGAAAAACAACACCGCTCCGCGCAC CGTGGACGTACGACAACATCGCGTATATGAAAAACAGCTCAAAATGCTGGGCTTTGGT TATGACTGGAGCCGCGAGCTGGCAACCTGTACGCCGGAATACCTACCGTTGGGAACAGAA ATTCTTCACCGAGCTGTATAAAAAAGGCTTGGTATATAAGAAGACTTTCGCGGTCAACT GGTCCCCGAACGACGACCGCTACTGGCGAAGCAACAAAGTTATCGACCGCTCTGCTGTG CGCTGCCGATACCAAGTTGAACGTAAAGAGATCCCGCAGTGGTTTATCAAAATCAATGC TTACGCTGACGAGCTGCTCAACGATCTGGATAAACTGGATCACTGGCCAGACACCGTTA AAACCATGCAGCGTAACCTGGATCGGTCTGTCCGAAGCGTGGAGATCACCTTCAACGTT AACGACTATGACAACACGCTGACCGTTTACACTACCCGCCCGGACACCTTTATGGGTGT TACCTACTTGGCGGTAGCTGCCGGTCTATCCGCTGGCGCAGAAAGCGCGGAAATATATC CTGAACTGGCGGCTTTTATGACGAATGCCGTAAACCAAAAGTTGCCGAAGCTGAAATC GCGACGATGGAGAAAAAGGCGTCTGATACTGGCTTTAAAGCGGTTACCCATTAAACGGG CGAAGAAATTCCTGTTTGGCGAGCAAACTTCGTATTGATGGAGTACGGCAGCGGCGCAG TTATGGCGGTACCGGGGACGACCGAGCGGACTACGAGTTTGCCTCTAAATACGSCCTG AACATCAAAACCGGTTATCTTGGCAGCTGACGGCTCTGAGCCAGATCTTCTCAGCAACG CCTGACTGAAAAAGGCGTGTCTTCAACTCTGGCGAGTTCAACGGTCTTGAACATGAAG CGGCTTCAACGCCATCGCCGATAAACTGACTGCGATGGGCGTTGGCGAGCGTAAAGT AACTACCGCTTGGCGGACTGGGGTGTTCCTCGTCAGCGTTACTGGGGCGCGCCGATCC GATGGTGACGCTGGAAGACGGTACCGTAATGCGGACCCCGGACGACGCTGCGGGTGA TCTTCCCGGAAGATGTGGTAATGGACGGCATTACAGCCCGATTAAAGCAGATCCGGAG TGGGCGAAAACTACCGTTAACGGTATGCCAGCACTGCGTGAAACCGACACTTTCGACAC CTTTATGGAGTCTCTCTGGTGGTATGCGCGCTACACTTGCCTCGAGTACAAAGAGGTA TGCTGGATTCCGAAGCGGCTAACTACTGGCTGCCGCTGGATATCTTTATGGTGGTATT GAACACGCCATTATGGGTCTGCTCTACTTCCGCTTCTTCCAAACTGATGCGTGATGC AGGCATGGTGAACCTGACGAACCGAGCAACAGTTGCTGTGTACGGGTATGGTGCTGG CAGATGCTTCTACTATGTGTGGCGAAAACGGCGAACGTAACCTGGGTTTCCCGGTTGAT GCTATGCTTGAACGTGACGAGAAAGCGGCTATCGTGAAGCGAAGATGCGCGAGGCCA TGAACCTGGTTTATACCGGCATGAGCAAAATGTCCAAGTCGAAGAACAACGGTATCGACC CGCAGGTGATGGTTGAACGTTACGGCGCGGACACCGTTCGCTGTGTTATGATGTTTGGT TCTCCGGCTGATATGACTCTCGAATGGCAGGAATCCGGTGTGGAAGGGCTTAACCGCTT CCTGAAACGTGTCTGGAACTGGTTACGAGCACACAGCAAAAGGTGATGTTGCCGCAC TGAACGTTGATGCGCTGACTGAAAATCAGAAAGCGCTGCTGCGGATGTGCATAAAACG ATCGCTAAAGTGACCGATGATATCGGCCGTCGTCAGACCTTCAACACCGCAATTGCGGC GATTATGGAGCTGATGAACAACTGGCGAAAGCACCACCGATGGCGAGCAGGATCGCG CTCTGATGCAGGAAGCACTGCTGGCCGTTGTCCGATGCTTAAACCGTTACCCCGCAC ATCTGCTTACGCTGTGGCAGGAACGAAAGGCGAAGGCGATACGACAACCGCCCGTG GCCGCTTCTGACGAAAAAGCGATGGTGGAAAGACTCCACGCTGGTCTGGTGCGAGGTTA ACGGTAAAGTCCGTGCCAAATCACCCTTCCGCTGGACGCAACGGAAGAACAGGTTCGC GAACGTGCTGCCAGGAACATCTGGTAGCAAAATATCTTGATGGCTTACTGTACGTAA AGTGATTACGTACCGGTAACCTCTCAATCTGGTCTGGCTAA</p>
<p>서열 번호 57</p> <p>OMeTyrRS-1 합성효소 아미노산(aa)</p>	<p>MASSNLIKQLQERGLVAQVTDEEALAEERLAQGPIALVCGFDPADSLHLGHLVPLLLCK RFQQAGHKPVALVGGATGLIGDPSFKAAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLFRLDRIKHFVSNQMINKEAVKQRLNREDQGISFTFEFSYNL LQGYSMACLNKQYGVVLQIGGSDOWNITSGIDLTRRLHQNQVFLTVPLITKADGTFK GKTEGGAVWLDPKKTSPYKFYQFWINTADADVYRFLKFFTFMSIEINALSEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITELFSGSLSALSEADFQLAQDGVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITNGEKQSDPEYFFKEDRLFRFTLLRR GKKNYCLICWK</p>

도면16BB

서열 번호 및 설명	서열
서열 번호 58 OMeTyrRS-2 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKKQRLNREDQGISFTFESYNL LQGYTMACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 59 OMeTyrRS-3 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKKQRLNREDQGISFTFESYNL LQGYTYACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 60 OMeTyrRS-4 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKKQRLNREDQGISFTFESYNL LQGYSMACSNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 61 OMeTyrRS-5 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKKQRLNREDQGISFTFESYNL LQGYSMACANKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 62 OMeTyrRS-6 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKKQRLNREDQGISFTFESYNL LQGYRMACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 63 p-아세틸PheRS-1 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKKQRLNREDQGISFTFESYNL LQGYGMACANKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 64 p-아세틸PheRS-2 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKKQRLNREDQGISFTFESYNL LQGYGMACANKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK

도면16CC

서열 번호 및 설명	서열
서열 번호 65 p-벤조일 PheRS-1 합성 효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALCGFDPTADSLHLGHLVPLCLK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKEAVKQRLNREDQGISFTFESYNL LQGYGFACANKQYGVVLQIGGSDQWGNITSGIDLTRLHQNVFGLTVPLITKADGTYKF GKTEGGAVWLDPKKTSFYKFYQFWINTADADVYRFLKFFTFMSIEEINAL EEDKNSGK APRAQYVLAEQVTRLVHGEEGLQAAKRITECLFSGSL SALSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 66 p-벤조일 PheRS-2 합성 효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALCGFDPTADSLHLGHLVPLCLK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKEAVKQRLNREDQGISFTFESYNL LQGYGYACMNKQYGVVLQIGGSDQWGNITSGIDLTRLHQNVFGLTVPLITKADGTYKF GKTEGGAVWLDPKKTSFYKFYQFWINTADADVYRFLKFFTFMSIEEINAL EEDKNSGK APRAQYVLAEQVTRLVHGEEGLQAAKRITECLFSGSL SALSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 67 p-아지도PheRS-1 합성 효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALCGFDPTADSLHLGHLVPLCLK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKEAVKQRLNREDQGISFTFESYNL LQGYSMACANKQYGVVLQIGGSDQWGNITSGIDLTRLHQNVFGLTVPLITKADGTYKF GKTEGGAVWLDPKKTSFYKFYQFWINTADADVYRFLKFFTFMSIEEINAL EEDKNSGK APRAQYVLAEQVTRLVHGEEGLQAAKRITECLFSGSL SALSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 68 p-아지도PheRS-2 합성 효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALCGFDPTADSLHLGHLVPLCLK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKEAVKQRLNREDQGISFTFESYNL LQGYSAACANKQYGVVLQIGGSDQWGNITSGIDLTRLHQNVFGLTVPLITKADGTYKF GKTEGGAVWLDPKKTSFYKFYQFWINTADADVYRFLKFFTFMSIEEINAL EEDKNSGK APRAQYVLAEQVTRLVHGEEGLQAAKRITECLFSGSL SALSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 69 p-아지도PheRS-3 합성 효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALCGFDPTADSLHLGHLVPLCLK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKEAVKQRLNREDQGISFTFESYNL LQGYSAACANKQYGVVLQIGGSDQWGNITSGIDLTRLHQNVFGLTVPLITKADGTYKF GKTEGGAVWLDPKKTSFYKFYQFWINTADADVYRFLKFFTFMSIEEINAL EEDKNSGK APRAQYVLAEQVTRLVHGEEGLQAAKRITECLFSGSL SALSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 70 p-아지도PheRS-4 합성 효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALCGFDPTADSLHLGHLVPLCLK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKEAVKQRLNREDQGISFTFESYNL LQGYSAACVKNQYGVVLQIGGSDQWGNITSGIDLTRLHQNVFGLTVPLITKADGTYKF GKTEGGAVWLDPKKTSFYKFYQFWINTADADVYRFLKFFTFMSIEEINAL EEDKNSGK APRAQYVLAEQVTRLVHGEEGLQAAKRITECLFSGSL SALSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 71 p-아지도PheRS-5 합성 효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALCGFDPTADSLHLGHLVPLCLK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFRLDIGKHFSVNQMINKEAVKQRLNREDQGISFTFESYNL LQGYNFACVKNQYGVVLQIGGSDQWGNITSGIDLTRLHQNVFGLTVPLITKADGTYKF GKTEGGAVWLDPKKTSFYKFYQFWINTADADVYRFLKFFTFMSIEEINAL EEDKNSGK APRAQYVLAEQVTRLVHGEEGLQAAKRITECLFSGSL SALSEADFEQLAQDGVPMVME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK

도면16DD

서열 번호 및 설명	서열
서열 번호 72 p-아지도PheRS-6 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFSVNQMINKEAVKQRLNREDQGISFTFEFSYNL LQGYSAACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNVFGLTVPLITKADGTFK GKTEGGAVWLDPKKTSFYKFWINTADADVRFKFFTFMSIEEINALAEEDKNSGK APRAQYVLAQVTRLVHGEEGLQAAKRITTECLFSGSLALSSEADFEQLAQDGVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 73 pPR-EcRS-1 합성효소 p-프로파르길옥시 -페닐알라닌 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFSVNQMINKEAVKQRLNREDQGISFTFEFSYNL LQGYSAACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNVFGLTVPLITKADGTFK GKTEGGAVWLDPKKTSFYKFWINTADADVRFKFFTFMSIEEINALAEEDKNSGK APRAQYVLAQVTRLVHGEEGLQAAKRITTECLFSGSLALSSEADFEQLAQDGVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 74 pPR-EcRS-2 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALTCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFSVNQMINKEAVKQRLNREDQGISFTFEFSYNL LQGYSAACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNVFGLTVPLITKADGTFK GKTEGGAVWLDPKKTSFYKFWINTADADVRFKFFTFMSIEEINALAEEDKNSGK APRAQYVLAQVTRLVHGEEGLQAAKRITTECLFSGSLALSSEADFEQLAQDGVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 75 pPR-EcRS-3 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFSVNQMINKEAVKQRLNREDQGISFTFEFSYNL LQGYTMACVKNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNVFGLTVPLITKADGTFK GKTEGGAVWLDPKKTSFYKFWINTADADVRFKFFTFMSIEEINALAEEDKNSGK APRAQYVLAQVTRLVHGEEGLQAAKRITTECLFSGSLALSSEADFEQLAQDGVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 76 pPR-EcRS-4 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFSVNQMINKEAVKQRLNREDQGISFTFEFSYNL LQGYSAACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNVFGLTVPLITKADGTFK GKTEGGAVWLDPKKTSFYKFWINTADADVRFKFFTFMSIEEINALAEEDKNSGK APRAQYVLAQVTRLVHGEEGLQAAKRITTECLFSGSLALSSEADFEQLAQDGVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 77 pPR-EcRS-5 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPIALCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFSVNQMINKEAVKQRLNREDQGISFTFEFSYNL LQGYTMACVKNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNVFGLTVPLITKADGTFK GKTEGGAVWLDPKKTSFYKFWINTADADVRFKFFTFMSIEEINALAEEDKNSGK APRAQYVLAQVTRLVHGEEGLQAAKRITTECLFSGSLALSSEADFEQLAQDGVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK

도면16EE

서열 번호 및 설명	서열
서열 번호 78 pPR-EcRS-6 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALTCGFDPTADSLHLGHLVPLLLCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFVNQMINKKQRLNREDQGISFTFESYNL LQGYTFACMNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSALSEADFEQLAQDQVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 79 pPR-EcRS-7 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALTCGFDPTADSLHLGHLVPLLLCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFVNQMINKKQRLNREDQGISFTFESYNL LQGYSVACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSALSEADFEQLAQDQVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 80 pPR-EcRS-8 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALTCGFDPTADSLHLGHLVPLLLCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFVNQMINKKQRLNREDQGISFTFESYNL LQGYSMACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSALSEADFEQLAQDQVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 81 pPR-EcRS-9 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALTCGFDPTADSLHLGHLVPLLLCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFVNQMINKKQRLNREDQGISFTFESYNL LQGYSFACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSALSEADFEQLAQDQVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 82 pPR-EcRS-10 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALTCGFDPTADSLHLGHLVPLLLCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFVNQMINKKQRLNREDQGISFTFESYNL LQGYTFACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSALSEADFEQLAQDQVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 83 p-요오도PheRS-1 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALTCGFDPTADSLHLGHLVPLLLCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFVNQMINKKQRLNREDQGISFTFESYNL LQGYSYACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSALSEADFEQLAQDQVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK
서열 번호 84 p-요오도PheRS-2 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDEEALAERLAQGPITALTCGFDPTADSLHLGHLVPLLLCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLNTEETVQEWVDKIRKQVAPFLDFDCGE NSAIAANNYDWFNMNVLTFLRDIGKHFVNQMINKKQRLNREDQGISFTFESYNL LQGYSMACLNKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSFYKQFWINTADADVYRFLKFTFMSIEEINALEEEDKNSGK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSALSEADFEQLAQDQVPMVEME KGADLMQALVDSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFGRFTLLRR GKKNYCLICWK

도면16FF

서열 번호 및 설명	서열
서열 번호 85 p-요오도PheRS-3 합성효소 아미노산(aa)	MASSNLIKQLQERGLVAQVTDDEALAEERLAQGPTALVCGFDPTADSLHLGHLVPLLCCK RFQQAGHKPVALVGGATGLIGDPSFKAERKLANTEETVQEWVDKIRKQVAPRLDFDCGE NSAIAANNYDWFNMNVLTLFLRDIGKHFVSNQMINKEAVKQRLNRDQGISFTEFSYNL LQGYSMACANKQYGVVLQIGGSDQWGNITSGIDLTRRLHQNQVFGTLVPLITKADGTFK GKTEGGAVWLDPKKTSYKPYQFWINTADADVYRFLKFFTFMSIEEINALDEEDKNSCK APRAQYVLAQVTRLVHGEGLQAAKRITTECLFSGSLSEADFEQLAQDQGVPMVEME KGADLMQALVDSSELQPSRGQARKTIASNAITINGEKQSDPEYFFKEEDRLFRGFTLLRR GKKNYCLICWK
서열 번호 86 a-아미노카프릴산 아미노아실- tRNA 합성효소 단리된-107 아미노산 서열 (야생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)	MQEQYRPEEIESKVQLHWDEKRTFEVTEDESKEKYCYLSAAYPSGRLHMGHVRNYTIG DVIARYQRLMGNVLQPIGWDAFGLPAEGAAVKNNTPAPWTDNIAYMKNQKMLGFG YDWSRELATCTPEYYRWEQKFFTELYKKGLVYKKTSAVNWCPNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTMRNWIIGRSEGEVITFNV NDYDNTLTVYTTRPDTFMGCTYLAVAAGHPLAQKAAENNPAAAFIDECRNTKVAEAE ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYEYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQQALTEKGVLFNSGEFNGLDHEAAFNAIADKLTAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVLTEDGTVMPTPDDQLFVILPEDVVMGITSPIKADPE WAKTTVNGMPALRETDFTDFMESSWYARYTCPOYKEGMLDSEAAANYWLPVDIIGGI EHAIMGLLYFRFFHKLMRDAGMVNSDEPAKQLLCQGMVLADAFYVYVGENGERNNVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMMFA SPADMTLEWQESGVEGANRFLKRVLVYEHTAKGDVAALNVDALENQKALRRDVHKT IAKVTDIGRRQTFNTAIAAIMELMNKLAKAPTDEQDRALMQEALLAVVRMLNPFTH ICFTLWQELKGEGLDIDNAPWVPADEKAMVEDSTLVVVQVNGKVRAKITVPVDATTEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG
서열 번호 87 a-아미노카프릴산 아미노아실- tRNA 합성효소 단리된-108(C8-RS) 아미노산 서열 (야생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)	MQEQYRPEEIESKVQLHWDEKRTFEVTEDESKEKYCYLSVMPYPSGRLHMGHVRNYTIG DVIARYQRLMGNVLQPIGWDAFGLPAEGAAVKNNTPAPWTDNIAYMKNQKMLGFG YDWSRELATCTPEYYRWEQKFFTELYKKGLVYKKTSAVNWCPNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTMRNWIIGRSEGEVITFNV NDYDNTLTVYTTRPDTFMGCTYLAVAAGHPLAQKAAENNPAAAFIDECRNTKVAEAE ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYEYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQQALTEKGVLFNSGEFNGLDHEAAFNAIADKLTAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVLTEDGTVMPTPDDQLFVILPEDVVMGITSPIKADPE WAKTTVNGMPALRETDFTDFMESSWYARYTCPOYKEGMLDSEAAANYWLPVDIIGGI EHAIMGLLYFRFFHKLMRDAGMVNSDEPAKQLLCQGMVLADAFYVYVGENGERNNVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMMFA SPADMTLEWQESGVEGANRFLKRVLVYEHTAKGDVAALNVDALENQKALRRDVHKT IAKVTDIGRRQTFNTAIAAIMELMNKLAKAPTDEQDRALMQEALLAVVRMLNPFTH ICFTLWQELKGEGLDIDNAPWVPADEKAMVEDSTLVVVQVNGKVRAKITVPVDATTEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG
서열 번호 88 a-아미노카프릴산 아미노아실- tRNA 합성효소 단리된-2F2 아미노산 서열 (야생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)	MQEQYRPEEIESKVQLHWDEKRTFEVTEDESKEKYCYLSHPPYPSGRLHMGHVRNYTIG DVIARYQRLMGNVLQPIGWDAFGLPAEGAAVKNNTPAPWTDNIAYMKNQKMLGFG YDWSRELATCTPEYYRWEQKFFTELYKKGLVYKKTSAVNWCPNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTMRNWIIGRSEGEVITFNV NDYDNTLTVYTTRPDTFMGCTYLAVAAGHPLAQKAAENNPAAAFIDECRNTKVAEAE ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYEYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQQALTEKGVLFNSGEFNGLDHEAAFNAIADKLTAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVLTEDGTVMPTPDDQLFVILPEDVVMGITSPIKADPE WAKTTVNGMPALRETDFTDFMESSWYARYTCPOYKEGMLDSEAAANYWLPVDIIGGI EHAIMGLLYFRFFHKLMRDAGMVNSDEPAKQLLCQGMVLADAFYVYVGENGERNNVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMMFA SPADMTLEWQESGVEGANRFLKRVLVYEHTAKGDVAALNVDALENQKALRRDVHKT IAKVTDIGRRQTFNTAIAAIMELMNKLAKAPTDEQDRALMQEALLAVVRMLNPFTH ICFTLWQELKGEGLDIDNAPWVPADEKAMVEDSTLVVVQVNGKVRAKITVPVDATTEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG

도면16GG

서열 번호 및 설명	서열
<p>서열 번호 89</p> <p>a-아미노카프릴산 아미노아실 -tRNA 합성효소 단리원-2F5 아미노산 서열 (야생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MQEQYRPEETESKVLHWDEKRTFEVTEDESKEKYCYLSVYPYPSGRLHMGHVRNYTIG DVIARYQRMGLGKNVLQPIGWDAFGLPAEGAAVKNNTPAPWTDNIAVMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFTELYKKGLVYKTSAVNNWCPNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKDLDHWPDTVKTQMRNWIGRSEGVETITFNV NDYDNTLTVYTTTRPDTFMGCTYLAVAAGHPLAQKAAENNPDLAAFDICRNKTKVAEAE ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQQALTEKGVLFNSGEFNGLDHEAAFNADKLTAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVLEDTGTVMPPTDDQLPVILPEDVVMGDTSPKADPE WAKTTVNGMPALRETDFTDFMESSWLYARYTCPOYKEGMLDSEAAANYWLPVDILIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVVGNGERNWVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDALENQKALRRDVHKT IAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPTDEQDRALMQEALLAVVRMLNPFPTH ICFTLWQELKGECDIDNAPWPVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATBEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG</p>
<p>서열 번호 90</p> <p>O-메틸 티로신 아미노아실-tRNA 합성효소 (OmeYRS) 단리원-3A7 아미노산 서열 (야생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MQEQYRPEETESKVLHWDEKRTFEVTEDESKEKYCYLSLEPYPSGRLHMGHVRNYTIG DVIARYQRMGLGKNVLQPIGWDAFGLPAEGAAVKNNTPAPWTDNIAVMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFTELYKKGLVYKTSAVNNWCPNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKDLDHWPDTVKTQMRNWIGRSEGVETITFNV NDYDNTLTVYTTTRPDTFMGCTYLAVAAGHPLAQKAAENNPDLAAFDICRNKTKVAEAE ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQQALTEKGVLFNSGEFNGLDHEAAFNADKLTAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVLEDTGTVMPPTDDQLPVILPEDVVMGDTSPKADPE WAKTTVNGMPALRETDFTDFMESSWLYARYTCPOYKEGMLDSEAAANYWLPVDIAIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVVGNGERNWVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDALENQKALRRDVHKT IAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPTDEQDRALMQEALLAVVRMLNPFPTH ICFTLWQELKGECDIDNAPWPVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATBEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG</p>
<p>서열 번호 91</p> <p>O-메틸 티로신 아미노아실 -tRNA 합성효소 단리원-3A2 아미노산 서열 (야생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MQEQYRPEETESKVLHWDEKRTFEVTEDESKEKYCYLSMEPYPSGRLHMGHVRNYTIG DVIARYQRMGLGKNVLQPIGWDAFGLPAEGAAVKNNTPAPWTDNIAVMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFTELYKKGLVYKTSAVNNWCPNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKDLDHWPDTVKTQMRNWIGRSEGVETITFNV NDYDNTLTVYTTTRPDTFMGCTYLAVAAGHPLAQKAAENNPDLAAFDICRNKTKVAEAE ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQQALTEKGVLFNSGEFNGLDHEAAFNADKLTAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVLEDTGTVMPPTDDQLPVILPEDVVMGDTSPKADPE WAKTTVNGMPALRETDFTDFMESSWLYARYTCPOYKEGMLDSEAAANYWLPVDIFIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVVGNGERNWVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDALENQKALRRDVHKT IAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPTDEQDRALMQEALLAVVRMLNPFPTH ICFTLWQELKGECDIDNAPWPVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATBEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG</p>

도면16HH

서열 번호 및 설명	서열
<p>서열 번호 92</p> <p>0-메틸 티로신 아미노아실-tRNA 합성효소 단리된 -3F11 아미노산 서열 (아생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MQEQYRPEEIESKVQLHWDEKRTFEVTEDESKEKYCYLSLEPYPSGRLHMGHVRNYTIG DVIARYQRMGLGKNVLQPIGWDAFGLPAEGAAVKNNTPAPWPTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFTELYKKGLVYKKTSAVNWCPNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTQQRNWIGRSEGVEITFNV NDYDNTLTVYTTTPDTFMGCTYLAVAAGHPLAQKAAENNPAAAFIDECRNTKVAEAEAM ATMEKKGVDTGFKAVHPLTGEELPVWAANFVLMYEGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQALTEKGVLFNSGEFNGLDHEAAFNADKLTAAGVGGERKV NYRLRDWGVSRQRYWGAPIPMVTELDGTVMPTDDQLPVILPEDVVMGDTSPKADPE WAKTTVNGMPALRETDFTDFMESSWRVARYTCPOYKEGMLDSEAANYWLPVDICIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVVGENGERNWVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDALENQKALRRDVHKT IAKVTDIGRRQTFNTAIAAIMELMNKLAKAPTDGEQDRALMQEALLAVVRMLNPFTH ICFTLWQELKGECDIDNAPWVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIVPGKLLNLVVG</p>
<p>서열 번호 93</p> <p>0-메틸 티로신 아미노아실 -tRNA 합성효소 단리된-3E7 아미노산 서열 (아생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MQEQYRPEEIESKVQLHWDEKRTFEVTEDESKEKYCYLSFEYPYPSGRLHMGHVRNYTIG DVIARYQRMGLGKNVLQPIGWDAFGLPAEGAAVKNNTPAPWPTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFTELYKKGLVYKKTSAVNWCPNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTQQRNWIGRSEGVEITFNV NDYDNTLTVYTTTPDTFMGCTYLAVAAGHPLAQKAAENNPAAAFIDECRNTKVAEAEAM ATMEKKGVDTGFKAVHPLTGEELPVWAANFVLMYEGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQALTEKGVLFNSGEFNGLDHEAAFNADKLTAAGVGGERKV NYRLRDWGVSRQRYWGAPIPMVTELDGTVMPTDDQLPVILPEDVVMGDTSPKADPE WAKTTVNGMPALRETDFTDFMESSWRVARYTCPOYKEGMLDSEAANYWLPVDITIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVVGENGERNWVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDALENQKALRRDVHKT IAKVTDIGRRQTFNTAIAAIMELMNKLAKAPTDGEQDRALMQEALLAVVRMLNPFTH ICFTLWQELKGECDIDNAPWVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIVPGKLLNLVVG</p>
<p>서열 번호 94</p> <p>0-니트로벤질 시스테인 아미노아실-tRNA 합성효소 단리된 -1A3 아미노산 서열 (아생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MQEQYRPEEIESKVQLHWDEKRTFEVTEDESKEKYCYLSGEPYPSGRLHMGHVRNYTIG DVIARYQRMGLGKNVLQPIGWDAFGLPAEGAAVKNNTPAPWPTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFTELYKKGLVYKKTSAVNWCPNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTQQRNWIGRSEGVEITFNV NDYDNTLTVYTTTPDTFMGCTYLAVAAGHPLAQKAAENNPAAAFIDECRNTKVAEAEAM ATMEKKGVDTGFKAVHPLTGEELPVWAANFVLMYEGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQALTEKGVLFNSGEFNGLDHEAAFNADKLTAAGVGGERKV NYRLRDWGVSRQRYWGAPIPMVTELDGTVMPTDDQLPVILPEDVVMGDTSPKADPE WAKTTVNGMPALRETDFTDFMESSWRVARYTCPOYKEGMLDSEAANYWLPVDILIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVVGENGERNWVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDALENQKALRRDVHKT IAKVTDIGRRQTFNTAIAAIMELMNKLAKAPTDGEQDRALMQEALLAVVRMLNPFTH ICFTLWQELKGECDIDNAPWVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIVPGKLLNLVVG</p>

도면1611

서열 번호 및 설명	서열
<p>서열 번호 95</p> <p>o-니트로벤질 시스테인</p> <p>아미노아실-tRNA</p> <p>-3A12 아미노산 서열 (야생형 이. 콜리 류실 tRNA-합성 효소에서 유도됨)</p>	<p>MQEYRPEEIESKVLHWDEKRTFEVTEDESKEKYYCLSGWPYPGRLHMGHVRNYTIG DVIARYQRLGKNVLQPIGWDAFGLPAEGAAVNNTAPAPWTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFTELYKKGLVYKKTSAVNWCNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTMRNWIGRSEGVETITFNV NDYDNTLTIVYTRPDTFMGCTYLAVAAGHPLAQKAAENNELAAAFIDECRNTKVAEAEAM ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQALTEKGVLFNSGEFNGLDHEAAFNADKLTAAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVLEDTGTVMPTDDQLPVILPEDVVMGITSPIKADPE WAKTTVNGMPALRETDFTFMESWAYARYTCPOYKEGMLDSEAAANYWLVPDILIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVVGENGERNVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDALTENQKALRRDVHKT IAKVTDIGRRQTFNTAIAAIMELMNKLAKAPTDGQDRALMQEALLAVVRMLNPFTPH ICFTLWQELKGEIDIDNAPWPVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG</p>
<p>서열 번호 96</p> <p>o-니트로벤질 시스테인</p> <p>아미노아실-tRNA</p> <p>합성효소 (nbCRS) 단리된-3H11</p> <p>아미노산 서열 (야생형 이. 콜리 류실 tRNA-합성 효소에서 유도됨)</p>	<p>MQEYRPEEIESKVLHWDEKRTFEVTEDESKEKYYCLSWSPYPSGRLHMGHVRNYTIG DVIARYQRLGKNVLQPIGWDAFGLPAEGAAVNNTAPAPWTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFTELYKKGLVYKKTSAVNWCNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTMRNWIGRSEGVETITFNV NDYDNTLTIVYTRPDTFMGCTYLAVAAGHPLAQKAAENNELAAAFIDECRNTKVAEAEAM ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQALTEKGVLFNSGEFNGLDHEAAFNADKLTAAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVLEDTGTVMPTDDQLPVILPEDVVMGITSPIKADPE WAKTTVNGMPALRETDFTFMESWIYARYTCPOYKEGMLDSEAAANYWLVPDIAIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVVGENGERNVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDALTENQKALRRDVHKT IAKVTDIGRRQTFNTAIAAIMELMNKLAKAPTDGQDRALMQEALLAVVRMLNPFTPH ICFTLWQELKGEIDIDNAPWPVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG</p>
<p>서열 번호 97</p> <p>o-니트로벤질 시스테인</p> <p>아미노아실-tRNA</p> <p>합성효소 단리된 -4E1 아미노산 서열 (야생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MQEYRPEEIESKVLHWDEKRTFEVTEDESKEKYYCLSGTPYPGRLHMGHVRNYTIG DVIARYQRLGKNVLQPIGWDAFGLPAEGAAVNNTAPAPWTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFTELYKKGLVYKKTSAVNWCNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTMRNWIGRSEGVETITFNV NDYDNTLTIVYTRPDTFMGCTYLAVAAGHPLAQKAAENNELAAAFIDECRNTKVAEAEAM ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQALTEKGVLFNSGEFNGLDHEAAFNADKLTAAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVLEDTGTVMPTDDQLPVILPEDVVMGITSPIKADPE WAKTTVNGMPALRETDFTFMESWYARYTCPOYKEGMLDSEAAANYWLVPDILIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVVGENGERNVSPVD AIVERDEKGRIVKAKDAAGHELVTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMFMA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDALTENQKALRRDVHKT IAKVTDIGRRQTFNTAIAAIMELMNKLAKAPTDGQDRALMQEALLAVVRMLNPFTPH ICFTLWQELKGEIDIDNAPWPVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG</p>

도면16JJ

서열 번호 및 설명	서열
<p>서열 번호 98</p> <p>1,5-단실알라닌 아미노아실-tRNA 합성효소 클론 88 아미노산 서열 (아생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MEEQYRPEEIESKVLHWDEKRTFEVTEDESKEKYCLSANPYPSGRLHMGHVRNYTIG DVIARYQRLGKGNVLQPIGWDAFGLPAEGAAVKNNNTAPAPWTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFFTELYKKGLVYKKTSAVNNWCNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLDKLDHWPDTVKTMRNWIGRSEGVETITFNV NDYDNTLTVYTTTRPDTFMGCTYLAVAAGHPLAQKAAENNPELAFAFIDECRNTKVAEAE ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQQALTEKGVLFNSGEFNGLDHEAAFNIAADKLTAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVTLEDGTVMPTPDDQLFVILPEDVVMGDITSPKADPE WAKTTVNGMPALRETDTFDTFMESCWIYARYTCPOYKEGMLDSEAANYWL PVDIGIGGI EHAIMTLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVYVGENGERNWNVSPVD AIVERDEKGRIVKAKDAAGHELVTYTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMMFA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDAL TENQKALRRDVHKT IAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPT DGEQDRALMQEALLAVVRMLNFFT ICFTLWQELKGEGLDIDNAPWPFVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG</p>
<p>서열 번호 99</p> <p>1,5-단실알라닌 아미노아실-tRNA 합성효소 T252A 아미노산 서열 (아생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MEEQYRPEEIESKVLHWDEKRTFEVTEDESKEKYCLSANPYPSGRLHMGHVRNYTIG DVIARYQRLGKGNVLQPIGWDAFGLPAEGAAVKNNNTAPAPWTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFFTELYKKGLVYKKTSAVNNWCNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLDKLDHWPDTVKTMRNWIGRSEGVETITFNV NDYDNTLTVYTTTRPDTFMGCTYLAVAAGHPLAQKAAENNPELAFAFIDECRNTKVAEAE ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQQALTEKGVLFNSGEFNGLDHEAAFNIAADKLTAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVTLEDGTVMPTPDDQLFVILPEDVVMGDITSPKADPE WAKTTVNGMPALRETDTFDTFMESCWIYARYTCPOYKEGMLDSEAANYWL PVDIGIGGI EHAIMTLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVYVGENGERNWNVSPVD AIVERDEKGRIVKAKDAAGHELVTYTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMMFA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDAL TENQKALRRDVHKT IAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPT DGEQDRALMQEALLAVVRMLNFFT ICFTLWQELKGEGLDIDNAPWPFVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG</p>
<p>서열 번호 100</p> <p>1,5-단실알라닌 아미노아실-tRNA 합성효소 V338A 아미노산 서열 (아생형 이. 콜리 류실 tRNA- 합성효소에서 유도됨)</p>	<p>MEEQYRPEEIESKVLHWDEKRTFEVTEDESKEKYCLSANPYPSGRLHMGHVRNYTIG DVIARYQRLGKGNVLQPIGWDAFGLPAEGAAVKNNNTAPAPWTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFFTELYKKGLVYKKTSAVNNWCNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLDKLDHWPDTVKTMRNWIGRSEGVETITFNV NDYDNTLTVYTTTRPDTFMGCTYLAVAAGHPLAQKAAENNPELAFAFIDECRNTKVAEAE ATMEKKGVDTGFKAVHPLTGEEIPVWAANFVLMYGTGAVMAVPGHDQRDYEFASKYGL NIKPVILAADGSEPDLSQQALTEKGVLFNSGEFNGLDHEAAFNIAADKLTAMGVGERKV NYRLRDWGVSRQRYWGAPIPMVTLEDGTVMPTPDDQLFVILPEDVVMGDITSPKADPE WAKTTVNGMPALRETDTFDTFMESCWIYARYTCPOYKEGMLDSEAANYWL PVDIGIGGI EHAIMTLLYFRFFHKLMDAGMVNSDEPAKQLLCQGMVLADAFYVYVGENGERNWNVSPVD AIVERDEKGRIVKAKDAAGHELVTYTGMSKMSKSKNNGIDPQVMVERYGADTVRLFMMFA SPADMTLEWQESGVEGANRFLKRVWKLVEHTAKGDVAALNVDAL TENQKALRRDVHKT IAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPT DGEQDRALMQEALLAVVRMLNFFT ICFTLWQELKGEGLDIDNAPWPFVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG</p>

도면16KK

서열 번호 및 설명	서열
서열 번호 101 o-니트로벤질세린 아미노아실-tRNA 합성효소 클론 G2-6 아미노산 서열 (야생형 이.콜리 류실 tRNA- 합성효소에서 유도됨)	MEEQYRPEEIESKVLHWDEKRTFEVTEDEGKEKYYCLSWSPYPSGRLHMGHVRNYTIG DVIARYQRMLGKNVLQPIGWDAFGLPAEGAAVKNNTPAPWTYDNIAYMKNQLKMLGFG YDWSRELATCTPEYYRWEQKFFTELYKKGLVYKTSAVNWCNDQTVLANEQVIDGCCW RCDTKVERKEIPQWFIKITAYADELLNDLKLHWPDTVKTMRNWIWRSEGEVEITFNV NDYDNTLTVYASRPDTFMGCTYLAVAAGHPLAQKAAENNPDLAAFIDECRTNKVAEAM ATMEKKGVDTFGKAVHPLTGEEIPVWAANFVLEMEYGTGAVMAVPGHDQDYEFASKYGL NIKPVILAADGSEPDLSQALTEKGLFNSGEFNGLDHEAFAFNADKLAMGVGERKV NYRLRDWGVSRQRYNGAPIPMVLTEDGTVMPPTDDQLPVILPEDVVMGDTSPISKADPE WAKTTVNGMPALRETDTFDTFMESCWIYARYTC PQYKEGMLDSEAANYWLPVDIAIGGI EHAIMGLLYFRFFHKLMDAGMVNSDEPAKQLCQGMVLADAFYYVGENGERNNWSPVD AIVERDEKGRIVKAKDAAGHELVTGISKMSKSKNNGIDPQVMVERYGADTVRLFMFA SPADMTLEWQESGVEGANRFLKRAWKLVEHTAKGDVAALNVDALTENQKALRRDVHKT IAKVTDDIGRRQTFNTAIAAIMELMNKLAKAPTDGEQDRALMQEALLAVVRMLNPFTPH ICFTLWQELKGEIDINAPWPFVADEKAMVEDSTLVVVQVNGKVRAKITVPVDATEEQVR ERAGQEHVAKYLDGVTVRKVIYVPGKLLNLVVG
서열 번호 102 GFP 펩티드 (여기서, Y는 야생 형 티로신 아미노산 위치를 가리킴)	FSVSGEGEGDATY ^Y GK
서열 번호 103 GFP 펩티드 (여기서, Y*는 비천연 아미노산 위치를 가리킴)	FSVSGEGEGDATY* ^Y GK
서열 번호 104 인간 tRNA ^{Tyr} 유전자의 5'-측점 서열	AGCGCTCCGGT ^T TTTCTGTGCTGAACCTCAGGGGACGCCGACACACGTCACCGTC
서열 번호 105 포유동물 세포 내 이.콜리 억제자 tRNA ^{Leu5} CUA 전사를 위한 5'-측점 서열	gatccgaccgtgtgcttggcagaac
서열 번호 106 포유동물 세포 내 이.콜리 억제자 tRNA ^{Leu5} CUA 전사를 위한 3'-측점 서열	gtccttttttttg

서열 목록

SEQUENCE LISTING

<110> The Scripps Research Institute
LIU, WENSHE
SCHULTZ, PETER G.

<120> GENETIC INCORPORATION OF UNNATURAL AMINO ACIDS INTO
PROTEINS IN
MAMMALIAN CELLS

<130> 54-002320US

<140> US 11/975,351

<141> 2007-10-17

<160> 106

<170> PatentIn version 3.4

<210> 1

<211> 87

<212> RNA

<213> Artificial

<220>

<223> mutant suppressor tRNA Leu5 CUA

<400> 1

gcccggauagg uggaauccgu agacacaagg gauucuaauu cccucggcgu ucgcgcugug 60

cggguucaag ucccgcuccg gguacca 87

<210> 2

<211> 85

<212> DNA

<213> Escherichia coli

<400> 2

ggtggggttc ccgagcggcc aaaggagca gactctaaat ctgccgtcat cgacttcgaa 60

ggttcgaatc ctccccccac cacca 85

<210> 3

<211> 82

<212> DNA

<213> Artificial

<220>

<223> mutant suppressor tyrosyl-tRNA(CUA)

<400> 3

ggaggggtag cgaagtggct aaacgcggcg gactctaaat ccgctccctt tgggttcggc 60

ggttcgaatc cgtccccctc ca 82

<210> 4

<211> 860

<212> PRT

<213> Escherichia coli

<400> 4

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Met Leu Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr
450 455 460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Ser Trp Tyr Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Tyr Ile
515 520 525

Gly Gly Ile Glu His Ala Ile Met His Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile

755	760	765	
Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn			
770	775	780	
Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr			
785	790	795	800
Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val			
805	810	815	
Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu			
820	825	830	
His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile			
835	840	845	
Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly			
850	855	860	
<210> 5			
<211> 2583			
<212> DNA			
<213> Escherichia coli			
<400> 5			
atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag	60		
aagcgcatat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctatg	120		
cttcctatc cttctggtcg actacacatg ggccacgtac gtaactacac catcggtgac	180		
gtgatcgccc gctaccagca tatgctgggc aaaaacgtcc tgcagccgat cggctgggac	240		
gcgtttggtc tgccatgcga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg	300		
acgtacgaca acatgcgta tatgaaaaac cagctcaaaa tgctgggctt tggttatgac	360		

tggagccg agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagtgt aacgtaaaga gatccgcag tggtttatca aaatcaatgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgcag 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgtga ccgtttacac taccgccccg gacaccttta tgggttgtag ctacctggcg 780

gtacgtgcgg gtcacccgtt ggccgagaaa gcggcggaataa ataactctga actggcgccc 840

tttattgacg aatgccgtaa caccaaagt t gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcggtt caccattaa cgggcgaaga aattcccggtt 960

tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020

cacgaccagc gcgactacga gtttgcctct aaatacggcc tgaacatcaa accggttattc 1080

ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cggtggcgag cgtaaagtga actaccgcct gcgcgactgg 1260

ggtgtttccc gtcagcgta ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgttgtaatg 1380

gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgtaaacggt 1440

atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcttc ctggtactat 1500

gcgcgctaca cttgcccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560

tggctgccgg tggatatcta cattggtggt attgaacacg ccattatgca cctgctctac 1620
 ttccgcttct tccacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680
 aaacagttgc tgggtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740
 ggcgaaacgta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800
 atcgtgaaag cgaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860
 tccaagtcga agaacaacgg tatcgaccg caggtgatgg ttgaacgtta cggcgcgac 1920
 accgttcgtc tgtttatgat gtttcttct cggctgata tgactctga atggcaggaa 1980
 tccggtgtgg aaggggctaa ccgcttctg aaactgtct ggaaactggt ttacgagcac 2040
 acagcaaaag gtgatgtgc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100
 ctgctgcgc atgtcataa aacgatcgt aaagtaccg atgatatcg ccgtcgtcag 2160
 acctcaaca ccgaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220
 accgatggcg agcaggaccg cgctctgatg caggaagcac tgctggccgt tgtccgtatg 2280
 cttaacccgt tcacccgca catctgcttc acgtgtggc aggaactgaa aggcgaaggc 2340
 gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400
 ctggtcgtgg tgcaggtaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460
 acggaagaac aggttcgca acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520
 ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tctcaatct ggtcgttggc 2580
 taa 2583

<210> 6

<211> 424

<212> PRT

<213> Escherichia coli

<400> 6

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Tyr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Asp Phe Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 7
<211> 1275
<212> DNA
<213> Escherichia coli

<400> 7
atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60

gacgaggaag cgtagcaga gcgactggcg caaggccga tcgcgtcta ttgcggcttc 120

gatcctaccg ctgacagctt gcatttgggg catcttggtc cattgttatg cctgaaacgc 180

ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240

gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300

gtggacaaaa tccgtaagca ggttgccccg ttctctgatt tcgactgtgg agaaaactct 360

gctatcgagg cgaacaacta tgactgggtc ggcaatatga atgtgctgac ctctctgcgc 420

gatattggca aacacttctc cgtaaccag atgatcaaca aagaagcggg taagcagcgt 480

ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gttgcagggt 540

tatgacttcg cctgtctgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600

cagtgggta acatcacttc tggatcgcac ctgacccgtc gtctgcatca gaatcaggtg 660

tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720

ggcggcgag tctggttgga tccgaagaaa accagcccggt acaaattcta ccagtctctg 780

atcaacactg cggatgccga cgtttaccgc ttctgaagt tcttcacctt tatgagcatt 840

gaagagatca acgcccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900

tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 960

aaacgtatta ccgaatgcct gttcagcgggt tctttgagtg cgctgagtga agcggacttc 1020

gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080

caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc 1140

tccaatgcca tcaccattaa cggtgaaaaa cagtccgac ctgaatactt ctttaaagaa 1200

gaagatcgtc tgtttggctg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260

atttgctgga aataa 1275

<210> 8
 <211> 1275
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 8
 atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60

gacgaggaag cgtagcaga gcgactggcg caaggcccgga tcgcactcgt gtgtggcttc 120

gacctaccg ctgacagctt gcatttgggg catcttgctt cattgttatg cctgaaacgc 180

ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc	240
gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg	300
gtggacaaaa tccgtaagca ggttgccccg ttctctgatt tcgactgtgg agaaaactct	360
gctatcgcg ccaataatta tgactgggtc ggcaatatga atgtgctgac ctctctgcgc	420
gatattggca aacacttctc cgtaaccag atgatcaaca aagaagcggg taagcagcgt	480
ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt	540
tatagtatgg cctgtttgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac	600
cagtggggta acatcacttc tggatcgac ctgaccgctc gtctgcatca gaatcagggtg	660
tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa	720
ggcggcgag tctggttga tccgaagaaa accagcccggt acaaattcta ccagttctgg	780
atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt	840
gaagagatca acgcccctga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag	900
tatgtactgg cggagcagggt gactcgtctg gttcacgggtg aagaaggttt acaggcggca	960
aaacgtatta ccgaatgcct gttcagcgggt tctttgagtg cgctgagtga agcggacttc	1020
gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg	1080
caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc	1140
tccaatgcca tcaccattaa cggtgaaaaa cagtccgac ctgaatactt ctttaaagaa	1200
gaagatcgtc tgtttggctg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg	1260
atttgcctga aataa	1275

<210> 9
 <211> 1275
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 9
 atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60
 gacgaggaag cgtagcaga gcgactggcg caaggcccga tcgcactcac ttgtggcttc 120
 gatcctaccg ctgacagctt gcatttgggg catcttgctt cattgttatg cctgaaacgc 180
 ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240
 gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300
 gtggacaaaa tccgtaagca gggtgccccg ttctctgatt tcgactgtgg agaaaactct 360
 gctatcgcg ccaataatta tgactggctt agcaatatga atgtgctgac cttcctgcgc 420
 gatattggca aacacttctc cgtaaccag atgatcaaca aagaagcggg taagcagcgt 480
 ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 540
 tatacgtatg cctgtctgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600
 cagtggggta acatcacttc tggatcgcac ctgaccgctc gtctgcatca gaatcaggtg 660
 tttagcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720
 ggcggcgcag tctggttggg tccgaagaaa accagcccggt acaaattcta ccagttctgg 780
 atcaacactg cggatgccga cgtttaccgc ttctgaagt tcttcacctt tatgagcatt 840
 gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900

tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 960

aaacgtatta ccgaatgcct gttcagcgtt tctttgagtg cgctgagtga agcggacttc 1020

gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080

caggcactgg tcgattctga actgcaacct tcccgtggtc aggcacgtaa aactatcgcc 1140

tccaatgcc a tcaccattaa cggtgaaaaa cagtccgatc ctgaatactt ctttaagaa 1200

gaagatcgtc tgittggctg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260

atttgctgga aataa 1275

<210> 10
 <211> 1275
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 10
 atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtgacg 60

gacgaggaag cgtagcaga gcgactggcg caaggccga tcgcactcgt gtgtggcttc 120

gatcctaccg ctgacagctt gcatttgggg catcttggtc cattgttatg cctgaaacgc 180

ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240

gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300

gtggacaaaa tccgtaagca ggttgccccg ttctctgatt tcgactgtgg agaaaactct 360

gctatcgagg ccaataatta tgactgggtc ggcaatatga atgtgctgac ctctctgcgc 420

gatattggca aacacttctc cgtaaccag atgatcaaca aagaagcgtg taagcagcgt 480

ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 540

tatagtatgg cctgtttgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600

cagtggggta acatcacttc tggtaicgac ctgacccgtc gtcigcatca gaatcagggtg 660

tttggcctga ccgttccgt gatactaaa gcagatggca ccaaatttgg taaaactgaa 720

ggcggcgag tctggttga tccgaagaaa accagcccgt acaaattcta ccagtcttg 780

atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt 840

gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900

tatgtactgg cggagcagg gactcgtctg gttcacggtg aagaagttt acaggcgga 960

aaacgtatta ccgaatgcct gttcagcggg tctttgagtg cgctgagtga agcggacttc 1020

gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080

caggcactgg tcgattctga actgcaacct tcccgtggc aggcacgtaa aactatgcc 1140

tccaatgcca tcaccattaa cggtgaaaaa cagtccgatc ctgaatactt ctttaaagaa 1200

gaagatcgtc tgtttggtcg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260

atttgctgga aataa 1275

<210> 11
 <211> 1275
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 11
 atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtagc 60

gacgaggaag cgtagcaga gcgactggcg caaggcccga tcgcactcgt gtgtggcttc 120

gacctaccg ctgacagctt gcatttgggg calcttgctt cattgttatg cctgaaacgc 180

ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240

gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300

gtggacaaaa tccgtaagca ggttccccg ttcctcgatt tcgactgtgg agaaaactct 360

gctatcgcg ccaataatta tgactggctt ggcaatatga atgtgctgac ctctctgcgc 420

gatattggca aacacttctc cgtaaccag atgatcaaca aagaagcggg taagcagcgt 480

ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 540

tatagtatgg cctgtttgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600

cagtggggta acatcacttc tggtaicgac ctgaccgctc gtctgcatca gaatcagggt 660

tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720

ggcggcgag tctggttga tccgaagaaa accagcccgt acaaattcta ccagtcttg 780

atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt 840

gaagagatca acgcccigga agaagaagat aaaaacagcg gtaaacgacc gcgcgcccag 900

tatgtactgg cggagcagg gactcgtctg gttcacggtg aagaagttt acaggcggca 960

aaacgtatta ccgaatgcct gttcagcggg tctttgagtgc gctgagtga agcggacttc 1020

gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaaggcg gc agacctgatg 1080

caggcactgg tcgattctga actgcaacct tcccgtgggc aggcacgtaa aactatcgcc 1140

tccaatgcca tcaccattaa cggtgaaaaa cagtccgac ctgaatactt ctttaaagaa 1200

gaagatcgtc tgtttggctg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260

atttgcctgga aataa 1275

<210> 12
<211> 1275
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 12
atggcaagca gtaacttgat taaacaattg caagagcggg ggctggtagc ccaggtagac 60

gacgaggaag cgtagcaga gcgactggcg caaggcccg tgcactcac gtgtggcttc 120

gatcctaccg ctgacagctt gcatttgggg catcttgctt cattgttatg cctgaaacgc 180

ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 240

gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 300

gtggacaaaa tccgtaagca ggttgcccg ttcctcgatt tcgactgtgg agaaaactct 360

gctatcgcg ccaataatta tgactggctt ggcaatatga atgtgctgac ctccctgcgc 420

gatattggca aacacttctc cgtaaccag atgatcaaca aagaagcggg taagcagcgt 480

ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacagcct gctgcagggt 540

tatacgatgg cctgtctgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 600

cagtggggtg acatcacttc tggatcgac ctgaccgctc gtctgcatca gaatcagggtg 660

tttggcctga ccgttccgct gatcactaaa gcagatggca ccaaatttgg taaaactgaa 720

ggcggcgag tctggttga tccgaagaaa accagcccg acaaattcta ccagttctgg 780

atcaacactg cggatgccga cgtttaccgc ttcctgaagt tcttcacctt tatgagcatt 840

gaagagatca acgccctgga agaagaagat aaaaacagcg gtaaagcacc gcgcgcccag 900

tatgtactgg cggagcaggt gactcgtctg gttcacggtg aagaaggttt acaggcggca 960

aaacgtatta ccgaatgcct gttcagcggg tctttgagtg cgctgagtga agcggacttc 1020

gaacagctgg cgcaggacgg cgtaccgatg gttgagatgg aaaagggcgc agacctgatg 1080

caggcactgg tcgattctga actgcaacct tcccgtgggc aggcacgtaa aactatcgcc 1140

tccaatgcc a tcaccattaa cggtgaaaaa cagtcgatc ctgaatactt ctttaaagaa 1200

gaagatcgtc tgtttggtcg ttttacctta ctgcgtcgcg gtaaaaagaa ttactgtctg 1260

atttgctgga aataa 1275

<210> 13
 <211> 540
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 13
 cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatgcac tcacttgigg cttegatcct accgctgaca gcttgcatth ggggcatctt 120

gttcattgt tatgcctgaa acgcttcag caggcgggcc acaagccggt tgcgctggtg 180

ggcggcgcca cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcagcaat 360

atgaatgtgc tgaccttctt gcgcgatatt ggcaaacact tctccgttaa ccagatgatc 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttctaca acctgctgca gggttatacg tatgcctgtc tgaacaaaca gtacgggtgtg 540

<210> 14
<211> 540
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 14
cgggggctgg taccccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60

ccgatcgac tcaattgtgg ctctgatcct accgtgaca gcttgcatTT ggggcatctt 120

gttcattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggtg 180

ggcggcgcgga cgggtctgat tggcgaccgc agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcttc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcagcaat 360

atgaatgtgc tgaccttctt gcgcgataat ggcaaacact tctccgttaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttctaca acctgctgca gggttatacg tatgcctgtc tgaacaaaca gtacgggtgtg 540

<210> 15
<211> 540
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 15
 cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

 ccgatcgac tcacttggg cttcgaccc accgctgaca gcttgcattt ggggcatctt 120

 gtccattgt tatgcctgaa acgcttcag caggcgggcc acaagccgtg tgcgctgga 180

 ggcgcgcgga cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

 accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

 gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcagcaat 360

 atgaatgtgc tgaccttctt gcgcgatatt ggcaaact tctccgttaa ccagatgac 420

 aacaaagaag cggtaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

 ttttctaca acctgctgca gggttatacg tatgcctgtc tgaacaaaca gtacggtgtg 540

<210> 16
 <211> 540
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 16
 cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

 ccgatcgac tcacttggg cttcgaccc accgctgaca gcttgcattt ggggcatctt 120

 gtccattgt tatgcctgaa acgcttcag caggcgggcc acaagccgtg tgcgctgga 180

 ggcgcgcgga cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

 accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

 gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttcct gcgcgataatt ggcaaact tctccgttaa ccagatgac 420

aacaaagaag cggtaagca gcgtctcaac cgtgaagac aggggatttc gttcactgag 480

ttttctaca acctgctgca gggttattcg tatgcctgtg cgaacaaaca gtacggtgtg 540

<210> 17
<211> 540
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 17
cgggggctgg tagcccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60

ccgatcgac tcaattgtgg cttcgatcct accgtgaca gcttgcatth ggggcatctt 120

gttccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180

ggcggcgca cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcagcaat 360

atgaatgtgc tgaccttcct gcgcgataatt ggcaaact tctccgttaa ccagatgac 420

aacaaagaag cggtaagca gcgtctcaac cgtgaagac aggggatttc gttcactgag 480

ttttctaca acctgctgca gggttatacg tatgcctgtc tgaacaaaca gtacggtgtg 540

<210> 18
<211> 540
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 18
 cgggggctgg taccccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

 ccgatcgac tcctttgtgg cttcgaccc accgctgaca gcttgcatctt ggggcatctt 120

 gtccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggtta 180

 ggccggcgga cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

 accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

 gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gtccgcaat 360

 atgaatgtgc tgaccttctt gcgcgataat ggcaaact tctccgttaa ccagatgac 420

 aacaaagaag cggtaagca gcgtctcaac cgtgaagatc aggggatttc gtccactgag 480

 ttttctaca acctgctgca gggttattct attgcctgtt cgaacaaaca gtacggtgtg 540

<210> 19
 <211> 540
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 19
 cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

 ccgatcgac tcgtgtgtgg cttcgaccc accgctgaca gcttgcatctt ggggcatctt 120

 gtccattgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggtta 180

 ggccggcgga cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

 accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgtaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttcctaca acctgctgca gggttatagt attgcctgtt tgaacaaaca gtacggtgtg 540

<210> 20
 <211> 540
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 20
 cgggggctgg taccccaggt gacggacgag gaagcgttag cagagcgact ggcgcaaggc 60

ccgatcgac tcgtgtgtgg cttegatect accgtgaca gcttgcattt ggggcatctt 120

gttccattgt taigcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180

ggcggcgcca cgggtctgat tggcgaccgc agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcagggttc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttcct gcgcgatatt ggcaaacact tctccgtaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttcctaca acctgctgca gggttatagt attgcctgtt tgaacaaaca gtacggtgtg 540

<210> 21
 <211> 540
 <212> DNA
 <213> Artificial

<220>

<223> mutant synthetase

<400> 21

cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatcgac tctggtgtgg cttcgaccc accgctgaca gcttgcatctt ggggcatctt 120

gttcatttgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180

ggcggcgca cgggtctgat tggcgaccg agcttcaagg ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaatt gttatgactg gttcggcaat 360

atgaatgtgc tgaccttctt gcgcgataat ggcaaact tctccgttaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gtccactgag 480

ttttctaca acctgctgca gggttatatg cgtgcctgtg agaacaaaca gtacggtgtg 540

<210> 22

<211> 624

<212> DNA

<213> Artificial

<220>

<223> mutant synthetase

<400> 22

cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatcgac tcatttgtgg cttcgaccc accgctgaca gcttgcatctt ggggcatctt 120

gttcatttgt tatgcctgaa acgcttccag caggcgggcc acaagccggt tgcgctggta 180

ggcggcgca cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttcct gcgcgataatt ggcaaact tctccgtaa ccagatgac 420

aacaaagaag cggtaagca gcgtctcaac cgtgaaggtc aggggatttc gttcactgag 480

tttctctaca acctgctgca gggttatggt atggcctgtg ctaacaaaca gtacgggtgtg 540

gtgctgcaaa ttggtgggtc tgaccaatgg gtaacatca cttctggtat cgacctgacc 600

cgctgctgc atcagaatca ggtg 624

<210> 23
 <211> 609
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 23
 caggtgacgg acgaggaagc gtttagcagag cgactggcgc aaggcccgat cgcactcgg 60

tgtggcttcg atcctaccgc tgacagcttg catttggggc atcttgttcc attgttatgc 120

ctgaaacgct tccagcaggc gggccacaag ccggttgcgc tggtaggcgg gcgcacgggt 180

ctgattggcg acccgagctt caaagctgcc gagcgtaacg tgaacaccga agaaactgtt 240

caggagtggg tggacaaaat ccgtaagcag gttgccccgt tcctcgattt cgactgtgga 300

gaaaactctg ctatcgcggc caataattat gactggttcg gcaatatgaa tgtgctgacc 360

ttctcgcg atattggcaa acattctcc gttaccaga tgatcaaca agaagcggtt 420

aagcagcgtc tcaaccgtga agatcagggg atttcgttca ctgagttttc ctacaacctg 480

ctgcagggtt atggttttgc ctgtttgaac aaacagtacg gtgtgggtgt gcaaattggt 540

ggtttctgacc agtggggtaa catcacttct ggtatcgacc tgacccgtcg tctgcatcag 600

aatcagggtg 609

<210> 24
<211> 591
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 24
gcgttagcag agcgactggc gcaaggcccg atcgactcg ggtgtggctt cgatcctacc 60

gctgacagct tgcatTTGGG gcacttTgtt ccattgttat gcctgaaacg cttccagcag 120

gcgggccaca agccggttgc gctgtaggc ggcgcgacgg gtctgattgg cgacccgagc 180

ttcaaagctg ccgagcgtaa gctgaacacc gaagaaactg ttcaggagtg ggtggacaaa 240

atccgtaagc aggttgcgcc gttcctcgat ttcgactgtg gagaaaactc tgctatcgcg 300

gccataatt atgactggtt cggcaatatg aatgtgctga cttcctgcg cgatattggc 360

aaacacttct ccgttaacca gatgatcaac aaagaagcgg ttaagcagcg tctcaaccgt 420

gaagatcagg ggatttcgtt cactgagttt tcctacaacc tgctgcaggg ttatggttat 480

gcctgtatga acaaacagta cgggtgtggtg ctgcaaattg gtggttctga ccagtggggt 540

aacatcactt ctggtatcga cctgacccgt cgtctgcac agaatcaggt g 591

<210> 25
<211> 621
<212> DNA
<213> Artificial

<220>

<223> mutant synthetase

<220>

<221> misc_feature

<222> (26)..(26)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (612)..(612)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (618)..(618)

<223> n is a, c, g, or t

<400> 25

gggctggttag cccaggtgac ggacgnagaa gcgttagcag agcgactggc gcaaggcccg 60

atcgcactcc ttigtggctt cgatcctacc gctgacagct tgcatittggg gcatcttgtt 120

ccattgttat gcctgaaacg cttccagcag gcgggccaca agccggttgc gctggtaggc 180

ggcgcgacgg gtctgattgg cgacccgagc ttcaaagctg ccgagcgtaa gctgaacacc 240

gaagaaactg ttcaggagtg ggtggacaaa atccgtaagc aggttgcccc gttcctcgat 300

ttcgactgtg gagaaaactc tgctatcgcg gccataatt atgactggtt cggcaatatg 360

aatgtgctga ctttctcgcg cgatattggc aaacacttct ccgttaacca gatgatcaac 420

aaagaagcgg ttaagcagcg tctcaaccgt gaagatcagg ggatttcgtt cactgagttt 480

tcctacaacc tgctgcaggg ttattctatg gcctgtgcga acaaacagta cgggtgtggtg 540

ctgcaaattg gtggttctga ccagtggggt aacatcactt ctggtatcga cctgacccgt 600

cgctctgcac anaatcangt g 621

<210> 26

<211> 588
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 26
 ttagcagagc gactggcgca aggcccgatc gcactcgttt gtggttcga tcctaccgt 60

 gacagcttgc atttggggca tcttgttcca ttgttatgcc tgaaacgctt ccagcaggcg 120

 ggccacaagc cgtttgcgct ggtaggcggc gcgacgggtc tgattggcga cccgagcttc 180

 aaagctgccg agcgtaagct gaacaccgaa gaaactgttc aggagtgggt ggacaaaatc 240

 cgtaagcagg ttccccgtt cctcgatttc gactgtggag aaaactctgc tatcgcggcc 300

 aataattatg actggttcgg caatatgaat gtgctgacct tctgcgcga tattggcaaa 360

 cacttctccg ttaaccagat gatcaacaaa gaagcggtta agcagcgtct caaccgtgaa 420

 gatcagggga ttctgttac tgagttttcc tacaacctgc tgcagggtta ttctgcggcc 480

 tgtgcgaaca aacagtacgg tgtgtgctg caaatgggtg gttctgacca gtggggtaac 540

 atcacttctg gtatcgacct gaccgcgtgt ctgcatcaga atcaggtg 588

<210> 27
 <211> 600
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<220>
 <221> misc_feature
 <222> (403)..(403)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (513)..(513)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (515)..(515)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (518)..(518)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (531)..(531)
 <223> n is a, c, g, or t

<400> 27
 gacgaggaag cgtagcaga gcgactggcg caaggcccg tgcactcct gtgtggcttc 60

 gatectaccg ctgacagctt gcatttgggg catcttggtc cattgttatg cctgaaacgc 120

 ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 180

 gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 240

 gtggacaaaa tccgtaagca gggtgccccg ttctctgatt tcgactgtgg agaaaactct 300

 gctatcgcg ccaataatta tgactgggtc ggcaatatga atgtgctgac cttcctgcgc 360

 gatattggca aacacttctc cgtaaccag atgatcaaca aanaagcggg taagcagcgt 420

 ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 480

 tattcggctg cctgtgcgaa caaacagtac gngngngnc tgcaaattgg nggttctgac 540

 caggggggta acatcacttc tggatcgcac ctgaccgctc gtctgcatca aaatcagggt 600

<210> 28

<211> 591
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<220>
<221> misc_feature
<222> (588)..(588)
<223> n is a, c, g, or t

<400> 28
gcgttagcag agcgactggc gcaaggcccg atcgactcg tttgtggctt cgatcctacc 60

gctgacagct tgcatttggg gcatcttggt ccattgttgt gcctgaaacg cttccagcag 120

gcggggccaca agccggttgc gctggtaggc ggcgcgacgg gtctgattgg cgacccgagc 180

ttcaaagctg ccgagcgtaa gctgaacacc gaagaaactg ttcaggagtg ggtggacaaa 240

atccgtaagc aggttgcccc gttcctcgat ttcgactgtg gagaaaactc tgctatcgcg 300

gccataaatt atgactggtt cggcaatatg aatgtgctga ctttcttgcg cgatattggc 360

aaacacttct ccgttaacca gatgatcaac aaagaagcgg ttaagcagcg tctcaaccgt 420

gaagatcagg ggatttcgtt cactgagttt tctacaacc tgctgcaggg ttatagtgcg 480

gcctgtgtta acaaacagta cgggtgtgtg ctgcaaattg gtggttctga ccagtggggt 540

aacatcactt ctggtatcga cctgacccgt cgtctgcatc agaatcangt g 591

<210> 29
<211> 600
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 29
gacgaggaag cgttagcaga gcgactggcg caaggcccg tgcactcat ttgtgcttc 60

gatcctaccg ctgacagctt gcatttgggg catcttgctt cattgttatg cctgaaacgc 120

ttccagcagg cgggccacaa gccggttgcg ctggtaggcg gcgcgacggg tctgattggc 180

gacccgagct tcaaagctgc cgagcgtaag ctgaacaccg aagaaactgt tcaggagtgg 240

gtggacaaaa tccgtaagca gggtgccccg ttcctcgatt tcgactgtgg agaaaactct 300

gctatcgagg ccaatgatta tgactggctt ggcaatatga atgtgctgac ctctctgcgc 360

gatattggca aacacttctc cgtaaccag atgatcaaca aagaagcggg taagcagcgt 420

ctcaaccgtg aagatcaggg gatttcgttc actgagtttt cctacaacct gctgcagggt 480

tataattttg cctgtgtgaa caaacagtac ggtgtggtgc tgcaaattgg tggttctgac 540

cagtggggtg acatcacttc tggatcgac ctgaccgctc gtctgcatca gaatcaggtg 600

<210> 30
<211> 579
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 30
cgactggcgc aaggcccgat cgacttcacg tgtggcttcg atcctaccgc tgacagcttg 60

catttggggc atcttgctcc attgttatgc ctgaaacgct tccagcaggc gggccacaag 120

ccggttgcgc tggtaggcgg cgcgacgggt ctgattggcg acccgagctt caaagctgcc 180

gagcgtaagc tgaacaccga agaaactgtt caggagtggg tggacaaaat ccgtaagcag 240

gttggccccg tctctgattt cgactgtgga gaaaactctg ctatcgcggc caataattat 300

gactggttcg gcaatatgaa tgtgtctgacc ttcttgcgcg atattggcaa acattctcc 360

gttaaccaga tgatcaacaa agaagcgggt aagcagcgtc tcaaccgtga agatcagggg 420

atttcgttca ctgagttttc ctacaatctg ctgcaggggt attcggctgc ctgtcttaac 480

aaacagtacg gtgtgggtgct gcaaattggt ggttctgacc agtggggtaa catcacttct 540

ggtatcgacc tgaccctcgc tctgcatcag aatcaggtg 579

<210> 31
 <211> 624
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<220>
 <221> misc_feature
 <222> (13)..(13)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (599)..(599)
 <223> n is a, c, g, or t

<400> 31
 cgggggctgg tanccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatcgac tcgggtgtgg ctctgacct accgtgaca gcttgcatctt ggggcatctt 120

gttcattgt tatgctgaa acgtctccag caggcgggcc acaagccggt tgcgctggta 180

ggcggcgca cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttctt gcgcgatatt ggcaaacact tctccgttaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttctaca acctgctgca gggttattct atggcctgtt tgaacaaaca gtacgggtgtg 540

gtgctgcaaa ttggtgggtc tgaccagtgg ggtaacatca cttctgggat cgacctganc 600

cgtcgtctgc atcagaatca ggtg 624

<210> 32
 <211> 625
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<220>
 <221> misc_feature
 <222> (600)..(600)
 <223> n is a, c, g, or t

<400> 32
 cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatcgac tcacgtgtgg cttcgatcct accgtgaca gcttgcatth ggggcatctt 120

gttcattgt tatgcctgaa acgtctccag caggcgggcc acaagccgtt tgcgttggtg 180

ggcggcgca cggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttctt gcgcgatatt ggcaaacact tctccgttaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttcctaca atctgctgca gggttattcg gctgcctgtc ttaacaaaca gtacggtgtg 540

gtgctgcaaa ttggtgggtc tgaccagtgg ggtaacatca cttctggat cgaacctgan 600

ccgtcgtctg catcaaaatc aagtg 625

<210> 33

<211> 624

<212> DNA

<213> Artificial

<220>

<223> mutant synthetase

<400> 33

cgggggctgg taccceaagt gacggacgag gaaacgttag cagagcgact ggcgcaaggc 60

ccgatcgac tctcttggg cttcgatcct accgtgaca gcttgcattt ggggcatctt 120

gttccattgt tatgcctgaa acgcttccag caggcaggcc acaagccggt tgcgctggta 180

ggcggcgca cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcagggttc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttctt gcgcgataat ggcaaacact tctccgttaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttcctaca acctgctgca gggttatacg atggcctgtg tgaacaaaca gtacggtgtg 540

gtgctgcaaa ttggtgggtc tgaccagtgg ggtaacatca cttctggat cgacctgacc 600

cgctcgtctgc atcagaatca ggtg 624

<210> 34

<211> 624

<212> DNA

<213> Artificial

<220>

<223> mutant synthetase

<400> 34

cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatgcac tcgctgcgg cttecatcct accgtgaca gcttgcattt ggggcatctt 120

gttcattgt tatgcctgaa acgtctccag caggcgggcc acaagccggt tgcgctggtta 180

ggcggcgcca cgggtctgat tggcgaccg agcttcaagg ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gticggcaat 360

atgaatgtgc tgaccttcct gcgcgatatt ggcaaact tctccgttaa ccagatgac 420

aacaaagaag cggtaagca gcgtctcaac cgtgaagatc aggggatttc gticactgag 480

ttttctaca acctgctgca gggttattct tatgcctgtc ttaacaaaca gtacggtgtg 540

gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600

cgtcgtctgc atcagaatca ggtg 624

<210> 35

<211> 624

<212> DNA

<213> Artificial

<220>

<223> mutant synthetase

<400> 35

cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatgcac tcgctgtgg cttecatcct accgtgaca gcttgcattt ggggcatctt 120

gttccattgt tatgctgaa acgttccag caggcgggcc acaagccgtg tgcgttgta 180

ggcggcgca cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttctt gcgcgataat ggcaaacact tctccgttaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

tttcttaca acctgctgca gggttatacg atggcctgtt gtaacaaaca gtacggtgtg 540

gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600

cgtcgtctgc atcagaatca ggtg 624

<210> 36
 <211> 624
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 36
 cgggggctgg tacccaagt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatcgac tcacgtgtgg cttcgacct accgtgaca gcttgcatth ggggcatctt 120

gttccattgt tatgctgaa acgttccag caggcgggcc acaagccgtg tgcgttgta 180

ggcggcgca cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttctc gcgcgatatt ggcaaacact tctccgttaa ccagatgac 420

aacaaagaag cggtaaagca gcgtctcaac cgtgaagatc aggggatttc gttcgtgag 480

ttttctaca acctgctgca gggttatacg ttgcctgta tgaacaaaca gtacggtgtg 540

gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600

cgtcgtctgc atcagaatca ggtg 624

<210> 37
 <211> 606
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 37
 gtgacggacg aggaagcgtt agcagagcga ctggcgcaag gcccgatcgc actcacgtgt 60

ggcttcgacg ctaccgctga cagcttgcgt ttggggcgcg ttgttccatt gttatgcctg 120

aaacgcttcc agcaggcggg ccacaagccg gttgcgctgg taggcggcgc gacgggtctg 180

attggcgacc cgagcttcaa agctgccgag cgtaagctga acaccgaaga aactgttcag 240

gagtgggtgg acaaaatccg taagcagggt gccccgttcc tcgatttcga ctgtggagaa 300

aactctgcta tcgcggccaa taattatgac tggttcggca atatgaatgt gctgaccttc 360

ctgcgcgata ttggcaaaaca cttctccgtt aaccagatga tcaacaaaga agcggttaag 420

cagcgtctca accgtgaaga tcaggggatt tcgttccactg agttttccta caatctgctg 480

cagggttatt cggtgcctg tcttaacaaa cagtacggtg tggctgctga aattggtggt 540

tctgaccagt ggggtaacat cacttctggt atcgacctga cccgtcgtct gcatcagaat 600

caggtg 606

<210> 38
<211> 624
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 38
cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatcgac tcgtttgtgg cttcgatcct accgtgaca gcttgcatth ggggcatctt 120

gttccattgt tatgcctgaa acgcttccag caggcggggc acaagccggt tgcgctggta 180

ggcggcgcca cgggtctgat tggcgaccgc agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttctt gcgcgataat ggcaaacact tctccgttaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttctaca acctgctgca gggttattcg atggcctgta cgaacaaaca gtacggtgtg 540

gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600

cgtcgtctgc atcagaatca ggtg 624

<210> 39
<211> 624
<212> DNA
<213> Artificial

<220>

<223> mutant synthetase

<220>

<221> misc_feature

<222> (13)..(13)

<223> n is a, c, g, or t

<400> 39

cgggggctgg tancccaagt gacggacggg gaagcgtag cagagcgact ggcgcaaggc 60

ccgatcgac tcagttgtgg cttegatcct accgtgaca gcttgcattt ggggcatctt 120

gttcattgt tatgcctgaa acgttccag caggcgggcc acaagccgtg tgcgctggtg 180

ggcggcgcca cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatctcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcgcaat 360

atgaatgtgc tgaccttctt gcgcgataat ggcaaact tctccgttaa ccagatgac 420

aacaaagaag cggttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

ttttctaca acctgctgca ggggttatag tttgcctgtc tgaacaaaca gtacggtgtg 540

gtgctgcaaa ttggtggttc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600

cgtcgtctgc atcagaatca ggtg 624

<210> 40

<211> 624

<212> DNA

<213> Artificial

<220>

<223> mutant synthetase

<400> 40

cgggggctgg tagcccaggt gacggacgag gaagcgtag cagagcgact ggcgcaaggc 60

ccgatcgac tcacgtgtgg cttegtatcct accgtgaca gcttgcatTT ggggcatctt 120

gttccattgt tatgcctgaa acgtttccag caggcgggcc acaagccggt tgcgttgga 180

ggcggcgcca cgggtctgat tggcgaccg agcttcaaag ctgccgagcg taagctgaac 240

accgaagaaa ctgttcagga gtgggtggac aaaatccgta agcaggttgc cccgttcctc 300

gatttcgact gtggagaaaa ctctgctatc gcggccaata attatgactg gttcggcaat 360

atgaatgtgc tgaccttctt gcgcgataatt ggcaaact tctccgttaa ccagatgac 420

aacaaagaag cgtttaagca gcgtctcaac cgtgaagatc aggggatttc gttcactgag 480

tttctctaca acctgctgca ggggttatacg ttgcctgta ctaacaaaca gtacggtgtg 540

gtgctgcaaa ttggtgggtc tgaccagtgg ggtaacatca cttctggtat cgacctgacc 600

cgtcgtctgc atcagaatca ggtg 624

<210> 41
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 41
 atggaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctgct 120

aatccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgtctggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgcgggctt tggttatgac 360

tggagccgag agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctgggtatat aagaagactt ctgcgggtcaa ctgggtgtccg 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatccgcag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aaccatgcag 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgctga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780

gtagctgcgg gtcacccgtt ggccgagaaa gcggcggaata ataactctga actggcgccc 840

tttattgacg aatgccgtta caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcgtt caccatttaa cggcggaaga aattccgtt 960

tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020

cacgaccagc gcgactaga gtttgcctct aaatacggcc tgaacatcaa accggtatc 1080

ctggcagctg acggctctga gccagatcct tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cggtggcgag cgtaaagtga actaccgcct gcgcgactgg 1260

ggtgtttccc gtcagcgtta ctggggcgcg ccgattccga tggtagctct agaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380

gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440

atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcctg ctggatttat 1500

gcgcgttaca cttgcccgca gtacaaagaa ggatgctgg attccgaagc ggctaactac 1560

tggctgccgg tggatatcgg tattggtggt attgaacacg ccattatgac gctgctctac 1620

ttccgcttct tcacaaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680

aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740

ggcgaacgta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800

atcgtgaaag cgaagaatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tcgaagtcca agaacaacgg tatcgaccgg caggtgatgg ttgaacttta cggcgcggac 1920

accgttcgtc tgtttatgat gtttcttct cgggtgata tgactctcga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgcttctg aaactgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtcataa aacgatcgt aaagtgaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgtatg 2280

cttaaccgt tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgca acgtctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tectcaatct ggtcgttggc 2580

taa 2583

<210> 42
 <211> 2589
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 42
 atggaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

 aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctgct 120

 aatccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

 gtgatcgccc gctaccagcg tatgtctggc aaaaacgtcc tgcagccgat cggctgggac 240

 gcgtttggtc tgcctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

 acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgcctgggctt tggttatgac 360

 tggagccgcg agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

 accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcgggtcaa ctggtgtccg 480

 aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

 accaaagttg aacgtaaaga gatcccgag tggtttatca aaatcactgc ttacgtgac 600

 gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aaccatgcag 660

 cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

 aacacgtga ccgtttacac tacccgcccg gacgcgttta tgggttgtac ctacctggcg 780

 gtagctgcgg gtcatccgct ggcgagaaa gcggcgaaa ataatcctga actggcgccc 840

 ttatttgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcggtt caccattaa cgggcgaaga aattcccgtt 960
 tgggcagcaa acttcgtatt gatggagtac ggcacgggcg cagttatggc ggtaccgggg 1020
 cagcaccagc gcgactacga gtttgcctct aaatacggcc tgaacatcaa accggttacc 1080
 ctggcagctg acggctctga gccagatctt tctcagcaag cctgactga aaaaggcgtg 1140
 ctgttcaact ctggcgagtt caacggctct gaccatgaag cggccttcaa cgccatcgcc 1200
 gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgcct gcgcgactgg 1260
 ggtgtttccc gtcagcgta ctggggcgcg ccgattccga tggtagctct agaagacggt 1320
 accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380
 gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440
 atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcttg ctggatttat 1500
 gcgcgctaca ctgcccgcga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560
 tggctgccgg tggatatcgg tattggtggt attgaacacg ccattatgac gctgctctac 1620
 ttccgcttct tcacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680
 aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740
 ggcgaaacta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800
 atcgtgaaag cgaaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860
 tccaagtga agaacaacgg tatcgaccg caggtgatgg ttgaacttta cggcgcgagc 1920
 accgttcgtc tgtttatgat gtttgcctct ccggtgata tgactctga atggcaggaa 1980
 tccggtgtgg aaggggctaa ccgcttcttg aaactgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttgc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtgcataa aacgatcgct aaagtgaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgaattgc ggcgattatg gagctgatga acaaaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgtatg 2280

cttaacccgt tcaccccgca catctgcttc acgtctgtggc aggaactgaa aggcgaaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggctgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgcga acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tctcaatct ggtcgttggc 2580

taagcggcc 2589

<210> 43
 <211> 2589
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 43
 atggaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctgct 120

aatccctatc cttctggtcg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgtctggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgttttgte tgcctgcgga aggcgcggcg gtgaaaaaca acaccgctcc ggcaccgtgg 300

acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgcctggcctt tggttatgac 360

tggagccgag agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgtccg 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatccgcag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgcag 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgctga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780

gtagctgcgg gtcacccgtt ggccgagaaa gcggcggaata ataactctga actggcgccc 840

tttattgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcgtt caccattaa cgggcgaaga aattccgtt 960

tgggcagcaa acttcgtatt gatggagtac ggcacggcgg cagttatggc ggcgccgggg 1020

cacgaccagc gcgactacga gtttcctctt aaatacggcc tgaacatcaa accggttatt 1080

ctggcagctg acggctctga gccagatctt tctcagcaag cctgactga aaaaggcgtg 1140

ctgttaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgctt gcgcgactgg 1260

ggtgtttccc gtcagcgtta ctggggcgcg ccgattccga tggtagctct agaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380

gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgtaaacggt 1440

atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcttg ctggatttat 1500

gcgcgctaca cttgcccgca gtacaaagaa ggatgctgg attccgaagc ggctaactac 1560

tggtgccgg tggatatcg tattggtggt attgaacacg ccattatgac gctgctctac 1620

ttccgcttct tccacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680

aaacagtgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740

ggcgaacgta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800

atcgtgaaag cgaagaatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tccaagtcga agaacaacgg tatcgaccg caggtgatgg ttgaacgtta cggcgcgac 1920

accgttcgtc tgtttatgat gtttcttct cggctgata tgactctga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgcttctg aaactgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgtgc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtcataa aacgatcgt aaagtgaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgatg 2280

cttaaccgt tcaccccgca catctgcttc acgctgtggc aggaactgaa aggcgaagc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgca acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tectcaatct ggtcgttggc 2580

taagcggcc 2589

<210> 44
 <211> 2844
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 44
 atctcgaagc acacgaaact ttttccttcc ttcattcacg cactactc tetaatgagc 60

 aacggtatac ggccttcctt ccagttactt gaatttgaaa taaaaaaaag ttgctgtct 120

 tgctatcaag tataaataga cctgcaatta ttaatctttt gtttcctcgt cattgttctc 180

 gtcccttttc ttctttgttt ctttttctgc acaatatitc aagctatacc aagcatacaa 240

 tcaactgaat tcagtatgga agagcaatac cgcccgaag agatagaatc caaagtacag 300

 cticattggg atgagaagcg cacatttgaa gtaaccgaag acgagggcaa agagaagtat 360

 tactgcctgt ctiggtcgcc ctatccttct ggtcgactac acatgggcca cgtacgtaac 420

 tacaccatcg gtgacgtgat cgcccgtac cagcgtatgc tgggcaaaaa cgtcctgcag 480

 ccgatcggct gggacgcgtt tggctctgcct gcggaaggcg cggcggtgaa aaacaacacc 540

 gctccggcac cgtggacgta cgacaacatc gcgtatatga aaaaccagct caaatgctg 600

 ggcttttggtt atgactggag ccgcgagctg gcaacctgta cgccggaata ctaccgttgg 660

 gaacagaaat tcttcaccga gctgtataaa aaaggcctgg tatataagaa gacttctgcg 720

 gtcaactggt gtccgaacga ccagaccgta ctggcgaacg aacaagttat cgacggctgc 780

 tgctggcgct gcgatacaa agttgaacgt aaagagatcc cgcagtgggt tatcaaaatc 840

 actgcttacg ctgacgagct gctcaacgat ctggataaac tggatcactg gccagacacc 900

gttaaaacca tgcagcgtaa ctggatcggc cgttccgaag gcgtggagat caccttcaac 960

gttaacgact atgacaacac gctgaccgtt tacgtttccc gcccgacac ctttatgggt 1020

tgtacctacc tggcggtagc tgcgggtcat ccgtggcgc agaaagcggc ggaaaataat 1080

cctgaactgg cggcctttat tgacgaatgc cgtaacacca aagttgccga agctgaaatg 1140

gcgacgatgg agaaaaaagg cgtcgatact ggctttaaag cggttcaccc attaacgggc 1200

gaagaaattc ccgtttgggc agcaaacttc gtattgatgg agtacggcac gggcgcagtt 1260

atggcggtag cggggcacga ccagcgcgac tacgagtttg cctctaaata cggcctgaac 1320

atcaaaccgg ttatcctggc agctgacggc tctgagccag atctttctca gcaagccctg 1380

actgaaaaag gcgtgctgtt caactctggc gagttcaacg gtcttgacca tgaagcgcc 1440

ttcaacgcca tcgccataa actgactgcg atgggcgttg gcgagcgtaa agtgaactac 1500

cgctgcgcg actgggggtg ttcccgtagc cgttactggg gcgcgccgat tccgatggtg 1560

actctagaag acggtaccgt aatgccgacc ccggacgacc agctgccggt gatcctgccg 1620

gaagatgtgg taatggacgg cattaccagc ccgattaaag cagatccgga gtgggcgaaa 1680

actaccgtta acggtatgcc agcactgcgt gaaaccgaca ctttcgacac ctttatggag 1740

tcctgtgga tttatgcgcg ctacacttgc ccgcagtaca aagaaggtat gctggattcc 1800

gaagcggcta actactggct gccggtggat atcgcgattg gtggtattga acacgccatt 1860

atggggctgc tctacttccg cttcttccac aaactgatgc gtgatgcagg catggtgaac 1920

tctgacgaac cagcgaaaca gttgctgtgt cagggtatgg tgctggcaga tgccttctac 1980

tatgttggcg aaaacggcga acgtaactgg gtttccccgg ttgatgctat cgttgaacgt 2040

gacgagaaag gccgtatcgt gaaagcgaaa gatgcggcag gccatgaact ggtttatacc 2100

ggcataagca aaatgtccaa gtcgaagaac aacggtatcg acccgaggt gatggttgaa 2160

cgttacggcg cggacacggt tcgtctgttt atgatgtttg cttctccggc tgatatgact 2220

ctcgaatggc aggaatccgg tgtggaagg gctaaccgct tcctgaaacg tgcctggaaa 2280

ctggtttacg agcacacagc aaaaggtgat gttgcggcac tgaacgttga tgcgctgact 2340

gaaaatcaga aagcgctgcg tcgcatgtg cataaaacga tcgctaaagt gaccgatgat 2400

atcgccgctc gtcagacctt caacaccgca attgcggcga ttatggagct gatgaacaaa 2460

ctggcgaaag caccaaccga tggcgagcag gatcgcgctc tgatgcagga agcactgctg 2520

gccgttgctc gtatgcttaa cccgttcacc ccgcacatct gcttcacgct gtggcaggaa 2580

ctgaaaggcg aaggcgatat cgacaacgcg ccgtggccgg ttgctgacga aaaagcgatg 2640

gtggaagact ccacgtggt cggtgtgcag gttaacggt aagtcgctc caaaatcacc 2700

gttccggtgg acgcaacgga agaacagggt cgcgaacgtg ctggccagga acatctggtg 2760

gcaaaatata ttgatggcgt tactgtacgt aaagtgattt acgtaccagg taaactctc 2820

aatctggtcg ttggctaagc ggcc 2844

<210> 45
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 45
 atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctgct 120

gcgccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgctgggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgcta tatgaaaaac cagctcaaaa tgcctggcctt tggttatgac 360

tggagccgag agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgcat 540

accaaagttg aacgtaaaga gatcccgag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgacg 660

cgtaactgga tcggtcgctc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgctga ccgtttacac taccgccccg gacaccttta tgggttgtag ctacctggcg 780

gtagctgcgg gtcacccgtt ggccgagaaa gcggcggaata ataactctga actggcggcc 840

tttatgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcggtt caccattaa cgggcgaaga aattcccgtt 960

tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttaggc ggtaccgggg 1020

cacgaccagc gcgactacga gtttgcctct aaatacggcc tgaacatcaa accggttacc 1080

ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggtctt gaccatgaag cgcccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cggtggcgag cgtaaagtga actaccgcct gcgcgactgg 1260

ggtgtttccc gtcagcgtaa ctggggcgcg ccgattccga tggtagacgt ggaagacggt 1320
 accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtgtaatg 1380
 gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440
 atgccagcac tgcgtgaaac cgacacttic gacacctta tggagtcttc ctggccttat 1500
 gcgcgtaca ctgccccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560
 tggctgccgg tggatatcgt tattgtgtgt attgaacacg ccattatggg gctgctctac 1620
 ttccgttct tccacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680
 aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740
 ggcgaaacgta actgggttic cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800
 atcgtgaaag cgaaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860
 tccaagtca agaacaacgg tatcgaccgg cagggtgatgg ttgaacgtta cggcgcgac 1920
 accgttcgtc tgittatgat gtttgcttct ccggtgata tgactctcga atggcaggaa 1980
 tccggtgtgg aaggggctaa ccgcttcctg aaacgtgtct ggaaactggt tiacgagcac 2040
 acagcaaaag gtgatgttc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100
 ctgcgtcgcg atgtgcataa aacgatcgct aaagtaccg atgatatcgg ccgtcgtcag 2160
 accttcaaca ccgcaattgc ggcgattatg gagctgatga acaaactggc gaaagcacca 2220
 accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgtatg 2280
 cttaacccgt tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaaaggc 2340
 gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400
 ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgcga acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tcttcaatct ggtcgttggc 2580

taa 2583

<210> 46
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 46
 atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctgtg 120

atgccctatc ctcttggtcg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgtctggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcttcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcggtg tatgaaaaac cagctcaaaa tgctgggctt tggttatgac 360

tggagccgag agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatcccgag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgacg 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgctga ccgtttacac taccgccccg gacaccttta tgggttgtag ctacctggcg 780
 gtagctgcgg gtcattccgtt ggcgagaaaa ggcgcggaaa ataattctga actggcggcc 840
 ttatttgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900
 aaaggcgtcg atactggctt taaagcggtt caccatttaa cgggcgaaga aattcccggt 960
 tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020
 cagcaccagc gcgactacga gtttgctctt aaatacggcc tgaacatcaa accggttacc 1080
 ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140
 ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200
 gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgcct gcgcgactgg 1260
 ggtgtttccc gtcagcgtaa ctggggcgcg ccgattccga tggtagcgtt ggaagacggt 1320
 accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380
 gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgtaaacggt 1440
 atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcctc ctggctgtat 1500
 gcgcgctaca ctgcccgcga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560
 tggctgccgg tggatatcct gattggtggt attgaacacg ccattatggg gctgctctac 1620
 ttccgcttct tccacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680
 aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740
 ggcgaaagta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800
 atcgtgaaag cgaaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tccaagtga agaacaacgg tatcgaccg caggtgatgg ttgaacgtta cggcgcgac 1920

accgttcgtc tgttatgat gtttcttct cggctgata tgactctga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgttcctg aaacgtgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttc ggcactgaac gttgatgcgc tgactgaaa tcagaaagcg 2100

ctgcgtcgcg atgtgcataa aacgatcgt aaagtaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgaattgc ggcgattatg gagctgatga acaaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgtatg 2280

cttaaccctg tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaagcg 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggctgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgcga acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tcctcaatct ggtcgttggc 2580

taa 2583

<210> 47
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 47
 atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctcat 120

cctccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgctgggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgcitgggctt tggttatgac 360

tggagccgcg agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatccgcag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgcag 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgtga ccgtttacac taccgccccg gacaccttta tgggttgtag ctacctggcg 780

gtagctgcgg gtcacccgtt ggccagaaaa gcggcgaaaa ataactctga actggcggcc 840

tttattgacg aatgccgtaa caccaaagt t gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcgtt caccattaa cgggcgaaga aattccgtt 960

tgggcagcaa acitcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020

cacgaccagc gcgactacga gtttgcctct aaatacggcc tgaacatcaa accggtatc 1080

ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cggtggcgag cgtaaagtga actaccgcct gcgcgactgg 1260

ggtgtttccc gtcagcgtta ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380

gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgtaaaccgt 1440

atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcttc ctgggcgtat 1500

gcgcgctaca ctgtcccgca gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560

tggctgccgg tggatatcat gattggtggt attgaacacg ccattatggg tctgctctac 1620

tccgcttct tccacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680

aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740

ggcgaacgta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800

atcgtgaaag cgaaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tccaagtca agaacaacgg tatcgaccgg caggatgatg ttgaacgtta cggcgcgac 1920

accgttcgtc tgittatgat gtttcttct cgggtgata tgactctga atggcaggaa 1980

tccgtgtgg aaggggctaa ccgcttctg aaactgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttc ggcactgaac gttgatgcgc tgactgaaa tcagaaagcg 2100

ctgcgtcgcg atgtgcataa aacgatcgct aaagtaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgcaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgtatg 2280

cttaaccgt tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgcga acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tcttcaatct ggtcgttggc 2580

taa 2583

<210> 48
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 48
 atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctgtg 120

tatccctatc ctcttggtcg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgtctggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcttcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgctga tatgaaaaac cagctcaaaa tgctgggctt tggttatgac 360

tggagccgag agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatcccgag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgacg 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgctga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780
 gtagctgcgg gtcattccgtt ggcgagaaaa ggcgcggaaa ataattctga actggcggcc 840
 ttatttgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900
 aaagcgctcg atactggctt taaagcggtt caccatttaa cgggcgaaga aattcccggtt 960
 tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020
 cagcaccagc gcgactacga gtttgctctt aaatacggcc tgaacatcaa accggttattc 1080
 ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaagcgctg 1140
 ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200
 gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgcct gcgcgactgg 1260
 ggtgtttccc gtcagcgta ctggggcgcg ccgattccga tggtagcgtt ggaagacggt 1320
 accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380
 gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440
 atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcttc ctggctgtat 1500
 gcgcgctaca ctgcccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560
 tggctgccgg tggatatcct gattggtggt attgaacacg ccattatggg tctgctctac 1620
 ttccgcttct tccacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680
 aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740
 ggcgaaagta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800
 atcgtgaaag cgaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860
 tccaagtga agaacaacgg tatcgaccgg caggtgatgg ttgaacgtta cggcgcgac 1920

accgttcgtc tgtttatgat gtttcttct cggctgata tgactctga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgcttcctg aaacgtgtct ggaaactggg ttacgagcac 2040

acagcaaaag gtgatgttgc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtgcataa aacgatcgtt aaagtaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgcaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgcctgatg 2280

cttaacccgt tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaagcg 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggctgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgcga acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tcctcaatct ggtcgttggc 2580

taa 2583

<210> 49
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 49
 atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctttg 120

gagccctatc ctcttggtcg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgtctgggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcttcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgctgggctt tggttatgac 360

tggagccgcg agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatccgcag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgcag 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgtga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780

gtagctgcgg gtcatccgct ggccagaaaa gcggcgaaaa ataactctga actggcggcc 840

tttattgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcggtt cacccattaa cgggcgaaga aattccggtt 960

tgggcagcaa acitcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020

cacgaccagc gcgactacga gtttgcctct aaatacggcc tgaacatcaa accggttacc 1080

ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cggtggcgag cgtaaagtga actaccgcct gcgcgactgg 1260

gggtgtttccc gtcagcgtaa ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtgtaatg 1380

gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440

atgccagcac tgcgtgaaac cgacaccttc gacaccttta tggagtcctc ctggcgttat 1500

gcgcgctaca ctigcccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560

tggctgccgg tggatatcgc tattggtggt attgaacacg ccattatggg tctgctctac 1620

ttccgcttct tccacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680

aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740

ggcgaacgta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800

atcgtgaaag cgaaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tccaagtcca agaacaacgg tatcgaccgg caggatgatg ttgaacgtta cggcgcgac 1920

accgttcgtc tgittatgat gtttcttct ccggtgata tgactctga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgcttctg aaactgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtcataa aacgatcgt aaagtaccg atgatcgg ccgtcgtcag 2160

accttaaca ccgaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgtatg 2280

cttaaccgt tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgca acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tctcaatct ggtcgttggc 2580

taa 2583

<210> 50

<211> 2583

<212> DNA

<213> Artificial

<220>

<223> mutant synthetase

<400> 50

atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctatg 120

gagccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgtctggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgctgggctt tggttatgac 360

tggagccgag agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctgcgat 540

accaaagttg aacgtaaaga gatcccgag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgacg 660

cgtaactgga tcggtcgctc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgctga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780

gtagctgctg gtcacccgctt ggccgagaaa ggcgcggaaa ataactctga actggcggcc 840

tttattgacg aatgccgtta caccaaagt gccaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcgtt caccattaa cgggcgaaga aattcccgtt 960

tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagtattggc ggtaccgggg 1020

cacgaccagc gcgactacga gtttgcctct aaatacggcc tgaacatcaa accggtatc 1080

ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgcct gcgcgactgg 1260

gggtgttccc gtcagcgta ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggaatg 1380

gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440

atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcttc ctggcggttat 1500

gcgcgctaca ctgcccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560

tggctgccgg tggatatctt tattggtggg attgaacacg ccattatggg gctgctctac 1620

ttccgcttct tccaaaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680

aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740

ggcgaacgta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800

atcgtgaaag cgaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tccaagtga agaacaacgg tatcgaccgg caggatgatg ttgaacgtta cggcgcgac 1920

accgttcgtc tgtttatgat gtttgcttct cggctgata tgactctga atggcaggaa 1980

tccgggtgtg aaggggctaa ccgcttcctg aaacgtgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttgc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtgcataa aacgatcgtt aaagtaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgcaattgc ggcgattatg gagctgatga acaaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgatg 2280

cttaacccgt tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaaaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgca acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tctcaatct ggtcgttggc 2580

taa 2583

<210> 51
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 51
 atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctttg 120

gagccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgtctgggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgcctgggctt tggttatgac 360

tggagccgcg agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcgggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatccgcgag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgcag 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgtga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780

gtagctgcgg gtcacccgtt ggccgagaaa gcggcggaaa ataactctga actggcggcc 840

tttattgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcggtt caccattaa cgggcgaaga aattcccgtt 960

tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020

cacgaccagc gcgactagca gtttgcctct aaatacggcc tgaacatcaa accggtatc 1080

ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgct gcgcgactgg 1260

ggtgtttccc gtcagcgta ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380

gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440

atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcttc ctggcggttat 1500

gcgcgctaca ctigcccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560

tggctgccgg tggatatctg tattggtggt attgaacacg ccattatggg tctgctctac 1620

ttccgcttct tccacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680

aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740

ggcgaacgta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800

atcgtgaaag cgaaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tccaagtcca agaacaacgg tatcgaccgg caggatgatg ttgaacgtta cggcgcggaac 1920

accgttcgtc tgtttatgat gtttcttct cgggtgata tgactctga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgcttctg aaactgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtgcataa aacgatcgct aaagtaccg atgatatcgg cgtcgtcag 2160

accttaaca ccgaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgtatg 2280

cttaaccctg tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgca acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tctcaatct ggtcgttggc 2580

taa 2583

<210> 52
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 52
 atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgcatat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctttt 120

gagccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgctgggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgcctgggctt tggttatgac 360

tggagccgag agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatcccgag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgacg 660

cgtaactgga tcggtcgctc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgtga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780

gtagctgcgg gtcatccgtt ggcgagaaa gcggcggaataa ataatcctga actggcggcc 840

tttattgacg aatgccgtta caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcgtt caccattaa cgggcgaaga aattcccgtt 960

tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020

cacgaccagc gcgactacga gtttcctct aaatacggcc tgaacatcaa accggttacc 1080

ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgcct gcgcgactgg 1260

ggtgtttccc gtcagcgta ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380

gacggcatta ccagccgat taaagcagat ccggagtggg cgaactac cttaacggt 1440

atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtctc ctggcgttat 1500

gcgcgctaca ctgcccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560

tggctgccgg tggatatcac gattggtggt attgaacacg ccattatggg tctgctctac 1620

ttccgcttct tccaaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680

aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740

ggcgaacgta actgggtttc cccggttgat gctatcgtt aacgtgacga gaaaggccgt 1800

atcgtgaaag cgaaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tccaagtca agaacaacgg tatcgaccg caggtgatgg ttgaacgtta cggcgcgac 1920

accgttcgtc tgtttatgat gtttcttct ccggtgata tgactctga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgcttcctg aaacgtgtct ggaaactggg ttacgagcac 2040

acagcaaaag gtgatgttgc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtgcataa aacgatcgct aaagtgaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgcaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgatg 2280

cttaacccgt tcaccccgca catctgcttc acgctgtggc aggaactgaa aggcgaagc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgcga acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tctcaatct ggtcgttggc 2580

taa 2583

<210> 53
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 53
 atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca tigggatgag 60

aagcgcatat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctggg 120

gagccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgtctggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgcctgggctt tggttatgac 360

tggagccgcg agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcgggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatccgcgag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aaccatgcag 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgtga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780

gtagctgcgg gtcateccgt ggccgagaaa gcggcggaata ataatcctga actggcggcc 840

tttattgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcggtt caccattaa cgggcgaaga aattcccgtt 960

tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagtatatgc ggtaccgggg 1020

cacgaccagc gcgactagc gtttgctctt aaatacggcc tgaacatcaa accggtatc 1080

ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggctctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgcct gcgcgactgg 1260

ggtgtttccc gtcagcgta ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380

gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440

atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcttc ctggcggtat 1500

gcgcgtaca ctigcccga gtacaaagaa ggatgctgg attccgaagc ggctaactac 1560

tggctgccgg tggatactct gattggtggt attgaacacg ccattatggg tctgctctac 1620

ttccgcttct tcacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680

aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740

ggcgaacgta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800

atcgtgaaag cgaaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tcgaagtcca agaacaacgg tatcgaccgg caggatgatg ttgaacttta cggcgcggaac 1920

accgttcgtc tgtttatgat gtttcttct cgggtgata tgactctga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgcttctg aaactgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtcataa aacgatcgt aaagtaccg atgatatcgg ccgtcgtcag 2160

accttaaca ccgcaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgatg 2280

cttaaccgt tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgca acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tctcaatct ggtcgttggc 2580

taa 2583

<210> 54
<211> 2583
<212> DNA
<213> Artificial

<220>
<223> mutant synthetase

<400> 54
atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtctggt 120

tggccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgtctggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgctcgga aggcgcgcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatcgcta tatgaaaaac cagctcaaaa tgctgggctt tggttatgac 360

tggagccgag agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc 480

aacgaccaga ccgtactggc gaacgaaca gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatcccgag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aacctgcag 660

cgtaactgga tcggtcgctc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgtga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780

gtagctgcgg gtcacccgt ggccgagaaa gcggcggaata ataactctga actggcgccc 840

ttattgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900
 aaaggcgtcg atactggctt taaagcggtt caccattaa cgggcgaaga aattcccgtt 960
 tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020
 cagcaccagc gcgactacga gtttcctct aaatacggcc tgaacatcaa accggttacc 1080
 ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140
 ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200
 gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgcct gcgcgactgg 1260
 ggtgtttccc gtcagcgta ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320
 accgtaatgc cgaccccgga cgaccagctg ccggtgacc tgccggaaga tgtggtaatg 1380
 gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440
 atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtctc ctgggttat 1500
 gcgcgctaca ctgcccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560
 tggctgccgg tggatacct tattggtggt attgaacacg ccattatggg tctgctctac 1620
 ttccgcttct tccaaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680
 aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740
 ggcgaaacgta actgggtttc cccggttgat gctatcgtt aacgtgacga gaaaggccgt 1800
 atcgtgaaag cgaaagatgc ggcaggccat gaactggtt ataccgcat gagcaaatg 1860
 tccaagtga agaacaacgg tatcgaccg caggtgatgg ttgaacgtta cggcgcgac 1920
 accgttcgtc tgtttatgat gtttcttct ccggtgata tgactctga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgcttcctg aaacgtgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttgc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtgcataa aacgatcgct aaagtgaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgcctgatg 2280

cttaaccctg tcaccccgca catctgcttc acgctgtggc aggaactgaa aggcgaaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgcga acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tctcaatct ggtcgttggc 2580

taa 2583

<210> 55
 <211> 2583
 <212> DNA
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 55
 atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag 60

aagcgcatat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtcttgg 120

tcgccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac 180

gtgatcgccc gctaccagcg tatgctgggc aaaaacgtcc tgcagccgat cggctgggac 240

gcgtttggtc tgcctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg 300

acgtacgaca acatgcgta tatgaaaaac cagctcaaaa tgcgggctt tggttatgac 360

tggagccgcg agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc 420

accgagctgt ataaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccg 480

aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat 540

accaaagttg aacgtaaaga gatccgcag tggtttatca aaatcactgc ttacgtgac 600

gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aaccatgcag 660

cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac 720

aacacgtga ccgtttacac taccgccccg gacaccttta tgggttgtac ctacctggcg 780

gtagctgcgg gtcattccgt ggccgagaaa gcggcggaataa ataactctga actggcgccc 840

tttattgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900

aaaggcgtcg atactggctt taaagcggtt caccattaa cgggcgaaga aattcccgtt 960

tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020

cacgaccagc gcgactacga gtttgcctct aaatacggcc tgaacatcaa accggtatc 1080

ctggcagctg acggtctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140

ctgttcaact ctggcgagtt caacggtctt gaccatgaag cggccttcaa cgccatcgcc 1200

gataaactga ctgcgatggg cggtggcgag cgtaaagtga actaccgcct gcgcgactgg 1260

ggtgtttccc gtcagcgta ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320

accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtggtaatg 1380

gacggcatta ccagcccgat taaagcagat ccggagtggg cgaaaactac cgttaacggt 1440

atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtcttc ctggatttat 1500

gcgcgttaca ctigcccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560

tggctgccgg tggatatcgc gattggtggt attgaacacg ccattatggg gctgctctac 1620

ttccgcttct tccacaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680

aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740

ggcgaacgta actgggtttc cccggttgat gctatcgttg aacgtgacga gaaaggccgt 1800

atcgtgaaag cgaaagatgc ggcaggccat gaactggttt ataccggcat gagcaaatg 1860

tccaagtcga agaacaacgg tatcgaccgg caggtgatgg ttgaacttta cggcgcggac 1920

accgttcgtc tgtttatgat gtttcttct cgggtgata tgactctcga atggcaggaa 1980

tccggtgtgg aaggggctaa ccgcttctg aaactgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttc ggcactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtgataa aacgatcgt aaagtaccg atgatatcgg ccgtcgtcag 2160

accttaaca ccgcaattgc ggcgattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgtatg 2280

cttaaccgt tcaccccgca catctgcttc acgtgtggc aggaactgaa aggcgaaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggtcgtgg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgca acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tctcaatct ggtcgttggc 2580

taa	2583
<210> 56	
<211> 2583	
<212> DNA	
<213> Artificial	
<220>	
<223> mutant synthetase	
<400> 56	
atgcaagagc aataccgccc ggaagagata gaatccaaag tacagcttca ttgggatgag	60
aagcgacat ttgaagtaac cgaagacgag agcaaagaga agtattactg cctgtcttgt	120
acgccctatc cttctggctg actacacatg ggccacgtac gtaactacac catcggtgac	180
gtgatcgccc gctaccagcg tatgtcgggc aaaaacgtcc tgcagccgat cggctgggac	240
gcgtttggtc tgctgcgga aggcgcggcg gtgaaaaaca acaccgtcc ggcaccgtgg	300
acgtacgaca acatcgcgta tatgaaaaac cagctcaaaa tgcctgggctt tggttatgac	360
tggagccgag agctggcaac ctgtacgccg gaatactacc gttgggaaca gaaattcttc	420
accgagctgt ataaaaaagg cctggtatat aagaagactt ctgcggtcaa ctggtgcccc	480
aacgaccaga ccgtactggc gaacgaacaa gttatcgacg gctgctgctg gcgctcgat	540
accaaagttg aacgtaaaga gatcccgag tggtttatca aaatcactgc ttacgtgac	600
gagctgctca acgatctgga taaactggat cactggccag acaccgttaa aaccatgcag	660
cgtaactgga tcggtcgttc cgaaggcgtg gagatcacct tcaacgttaa cgactatgac	720
aacacgtga ccgtttacac taccgcccc gacaccttta tgggttgtac ctacctggcg	780
gtagctgcgg gtcatecgct ggccgagaaa gcggcggaataa ataactctga actggcgccc	840

ttattgacg aatgccgtaa caccaaagtt gccgaagctg aaatggcgac gatggagaaa 900
 aaaggcgtcg atactggctt taaagcggtt caccattaa cgggcgaaga aattcccggt 960
 tgggcagcaa acttcgtatt gatggagtac ggcacggcg cagttatggc ggtaccgggg 1020
 cacgaccagc gcgactacga gtttgcctct aaatacggcc tgaacatcaa accggttatt 1080
 ctggcagctg acggctctga gccagatctt tctcagcaag ccctgactga aaaaggcgtg 1140
 ctgttcaact ctggcgagtt caacggctct gaccatgaag cggccttcaa cgccatcgcc 1200
 gataaactga ctgcgatggg cgttggcgag cgtaaagtga actaccgcct gcgcgactgg 1260
 ggtgtttccc gtcagcgtaa ctggggcgcg ccgattccga tggtagcgt ggaagacggt 1320
 accgtaatgc cgaccccgga cgaccagctg ccggtgatcc tgccggaaga tgtgtaatg 1380
 gacggcatta ccagccgat taaagcagat ccggagtggg cgaactac cttaacggt 1440
 atgccagcac tgcgtgaaac cgacactttc gacaccttta tggagtctc ctggtggtat 1500
 gcgcgctaca ctgcccga gtacaaagaa ggtatgctgg attccgaagc ggctaactac 1560
 tggctgccgg tggatatcct tattggtggt attgaacacg ccattatggg tctgctctac 1620
 ttccgcttct tccaaaact gatgcgtgat gcaggcatgg tgaactctga cgaaccagcg 1680
 aaacagttgc tgtgtcaggg tatggtgctg gcagatgcct tctactatgt tggcgaaaac 1740
 ggcgaaagta actgggtttc cccggttgat gctatcgtt aacgtgacga gaaaggccgt 1800
 atcgtgaaag cgaaagatgc ggcaggccat gaactggtt ataccggcat gagcaaatg 1860
 tccaagtga agaacaacgg tatcgaccg caggtgatgg ttgaacgtta cggcgcgac 1920
 accgttcgtc tgtttatgat gtttcttct ccggtgata tgactctga atggcaggaa 1980
 tccggtgtgg aaggggctaa ccgcttctg aaacgtgtct ggaaactggt ttacgagcac 2040

acagcaaaag gtgatgttgc ggcaactgaac gttgatgcgc tgactgaaaa tcagaaagcg 2100

ctgcgtcgcg atgtgcataa aacgatcgtt aaagtgaccg atgatatcgg ccgtcgtcag 2160

accttcaaca ccgaattgc ggcaattatg gagctgatga acaactggc gaaagcacca 2220

accgatggcg agcaggatcg cgctctgatg caggaagcac tgctggccgt tgtccgatg 2280

cttaaccctg tcaccccgca catctgcttc acgctgtggc aggaactgaa aggcgaaggc 2340

gatatcgaca acgcgccgtg gccggttgct gacgaaaaag cgatggtgga agactccacg 2400

ctggctcgtg tgcaggttaa cggtaaagtc cgtgccaaaa tcaccgttcc ggtggacgca 2460

acggaagaac aggttcgcga acgtgctggc caggaacatc tggtagcaaa atatcttgat 2520

ggcgttactg tacgtaaagt gatttacgta ccaggtaaac tectcaatct ggtcgttggc 2580

taa 2583

<210> 57
 <211> 424
 <212> PRT
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 57

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
 1 5 10 15

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

Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
 35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Met Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

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

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

<210> 58
 <211> 424
 <212> PRT
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 58

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
 1 5 10 15

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

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
 35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
 50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
 85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
 100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
 115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
 130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Thr Met Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 59
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 59

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Thr Tyr Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe

245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 60
<211> 424

<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 60

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Leu Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Met Ala Cys Ser Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 61
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 61

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Leu Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu

260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 62
<211> 424
<212> PRT
<213> Artificial

<220>

<223> mutant synthetase

<400> 62

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Arg Met Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 63
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 63

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Ile Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Gly Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu

275 280 285
 Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
 290 295 300
 Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
 305 310 315 320
 Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
 325 330 335
 Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
 340 345 350
 Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
 355 360 365
 Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
 370 375 380
 Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
 385 390 395 400
 Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
 405 410 415
 Asn Tyr Cys Leu Ile Cys Trp Lys
 420

<210> 64
 <211> 424
 <212> PRT
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 64

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Ile Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Gly Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Gly Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 65
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 65

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Gly Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Gly Phe Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 66
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 66

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

1					5					10					15				
Ala	Gln	Val	Thr	Asp	Glu	Glu	Ala	Leu	Ala	Glu	Arg	Leu	Ala	Gln	Gly				
			20						25						30				
Pro	Ile	Ala	Leu	Gly	Cys	Gly	Phe	Asp	Pro	Thr	Ala	Asp	Ser	Leu	His				
			35						40						45				
Leu	Gly	His	Leu	Val	Pro	Leu	Leu	Cys	Leu	Lys	Arg	Phe	Gln	Gln	Ala				
			50						55						60				
Gly	His	Lys	Pro	Val	Ala	Leu	Val	Gly	Gly	Ala	Thr	Gly	Leu	Ile	Gly				
			65						70						75				
Asp	Pro	Ser	Phe	Lys	Ala	Ala	Glu	Arg	Lys	Leu	Asn	Thr	Glu	Glu	Thr				
				85						90						95			
Val	Gln	Glu	Trp	Val	Asp	Lys	Ile	Arg	Lys	Gln	Val	Ala	Pro	Phe	Leu				
			100						105						110				
Asp	Phe	Asp	Cys	Gly	Glu	Asn	Ser	Ala	Ile	Ala	Ala	Asn	Asn	Tyr	Asp				
			115						120						125				
Trp	Phe	Gly	Asn	Met	Asn	Val	Leu	Thr	Phe	Leu	Arg	Asp	Ile	Gly	Lys				
			130						135						140				
His	Phe	Ser	Val	Asn	Gln	Met	Ile	Asn	Lys	Glu	Ala	Val	Lys	Gln	Arg				
			145						150						155				
Leu	Asn	Arg	Glu	Asp	Gln	Gly	Ile	Ser	Phe	Thr	Glu	Phe	Ser	Tyr	Asn				
				165						170						175			
Leu	Leu	Gln	Gly	Tyr	Gly	Tyr	Ala	Cys	Met	Asn	Lys	Gln	Tyr	Gly	Val				
			180						185						190				
Val	Leu	Gln	Ile	Gly	Gly	Ser	Asp	Gln	Trp	Gly	Asn	Ile	Thr	Ser	Gly				

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

385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
 405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>	67
<211>	424
<212>	PRT
<213>	Artificial

<220>
<223> mutant synthetase

<400> 67

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Leu Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 68
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 68

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Ala Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

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

405

410

415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 69

<211> 424

<212> PRT

<213> Artificial

<220>

<223> mutant synthetase

<400> 69

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Leu Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp

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

305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>	70
<211>	424
<212>	PRT
<213>	Artificial

<220>
<223> mutant synthetase

<400> 70

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Ala Ala Cys Val Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 71
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 71

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Ile Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asp Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Asn Phe Ala Cys Val Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 72
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 72

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His

35

40

45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Ala Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

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

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

<210> 73
 <211> 424
 <212> PRT
 <213> Artificial

<220>
 <223> mutant synthetase

<400> 73

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
 1 5 10 15

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

Pro Ile Ala Leu Gly Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
 35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
 50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
 85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
 100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
 115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
 130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Met Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 74
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 74

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Ala Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe

245

250

255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 75

<211> 424

<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 75

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Ser Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Thr Met Ala Cys Val Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 76
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 76

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Ala Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Tyr Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu

260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 77
<211> 424
<212> PRT
<213> Artificial

<220>

<223> mutant synthetase

<400> 77

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Ala Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Thr Met Ala Cys Cys Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 78
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 78

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Thr Phe Ala Cys Met Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu

275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 79
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 79

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Val Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 80
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 80

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Met Ala Cys Thr Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 81
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 81

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

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

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

385 390 395 400

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

 405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>	82
<211>	424
<212>	PRT
<213>	Artificial

<220>
<223> mutant synthetase

<400> 82

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Thr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Thr Phe Ala Cys Thr Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 83
<211> 424
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 83

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Tyr Ala Cys Leu Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

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

405

410

415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 84

<211> 424

<212> PRT

<213> Artificial

<220>

<223> mutant synthetase

<400> 84

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Ile Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp

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

305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210>	85
<211>	424
<212>	PRT
<213>	Artificial

<220>
<223> mutant synthetase

<400> 85

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val
1 5 10 15

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

Pro Ile Ala Leu Val Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His
35 40 45

Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gln Gln Ala
50 55 60

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

Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr
85 90 95

Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu
100 105 110

Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp
115 120 125

Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys
130 135 140

His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg
145 150 155 160

Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn
165 170 175

Leu Leu Gln Gly Tyr Ser Met Ala Cys Ala Asn Lys Gln Tyr Gly Val
180 185 190

Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly
195 200 205

Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr
210 215 220

Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu
225 230 235 240

Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe
245 250 255

Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu
260 265 270

Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu
275 280 285

Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala
290 295 300

Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala
305 310 315 320

Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser
325 330 335

Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu
340 345 350

Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu
355 360 365

Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile
370 375 380

Thr Ile Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu
385 390 395 400

Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys
420

<210> 86
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 86

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Ala Ala Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr
450 455 460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Ser Trp Pro Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Val Ile

515

520

525

Gly Gly Ile Glu His Ala Ile Met Gly Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln

705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210>	87
<211>	860
<212>	PRT
<213>	Artificial

<220>
<223> mutant synthetase

<400> 87

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Val Met Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

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

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

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 88
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 88

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser His Pro Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala

260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr

450

455

460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Ser Trp Ala Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Met Ile
515 520 525

Gly Gly Ile Glu His Ala Ile Met Gly Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 89
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 89

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Val Tyr Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr
450 455 460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Ser Trp Leu Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Leu Ile

515

520

525

Gly Gly Ile Glu His Ala Ile Met Gly Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln

705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210>	90
<211>	860
<212>	PRT
<213>	Artificial

<220>
<223> mutant synthetase

<400> 90

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Leu Glu Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

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

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

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 91
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 91

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Met Glu Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala

260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr

450

455

460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Ser Trp Arg Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Phe Ile
515 520 525

Gly Gly Ile Glu His Ala Ile Met Gly Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 92
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 92

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Leu Glu Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr
450 455 460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Ser Trp Arg Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Cys Ile

515

520

525

Gly Gly Ile Glu His Ala Ile Met Gly Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln

705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210>	93
<211>	860
<212>	PRT
<213>	Artificial

<220>
<223> mutant synthetase

<400> 93

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Phe Glu Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

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

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

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 94
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 94

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Gly Glu Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala

260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr

450

455

460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Ser Trp Arg Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Leu Ile
515 520 525

Gly Gly Ile Glu His Ala Ile Met Gly Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 95
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 95

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Gly Trp Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr
450 455 460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Ser Trp Ala Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Leu Ile

515

520

525

Gly Gly Ile Glu His Ala Ile Met Gly Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln

705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210>	96
<211>	860
<212>	PRT
<213>	Artificial

<220>
<223> mutant synthetase

<400> 96

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Trp Ser Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

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

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

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 97
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 97

Met Gln Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Gly Thr Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala

260 265 270

 Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
 275 280 285

 Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
 290 295 300

 Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
 305 310 315 320

 Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
 325 330 335

 Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
 340 345 350

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

 Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
 370 375 380

 Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
 385 390 395 400

 Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
 405 410 415

 Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
 420 425 430

 Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
 435 440 445

 Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr

450

455

460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Ser Trp Trp Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Leu Ile
515 520 525

Gly Gly Ile Glu His Ala Ile Met Gly Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 98
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 98

Met Glu Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Ala Asn Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr
450 455 460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Cys Trp Ile Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Gly Ile

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

705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210>	99
<211>	860
<212>	PRT
<213>	Artificial

<220>
<223> mutant synthetase

<400> 99

Met Glu Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Ala Asn Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Ala Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

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

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

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 100
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 100

Met Glu Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Ser Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Ala Asn Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Thr Thr Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala

260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Ala Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr

450

455

460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Cys Trp Ile Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Gly Ile
515 520 525

Gly Gly Ile Glu His Ala Ile Met Thr Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Met Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Val Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln
705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210> 101
<211> 860
<212> PRT
<213> Artificial

<220>
<223> mutant synthetase

<400> 101

Met Glu Glu Gln Tyr Arg Pro Glu Glu Ile Glu Ser Lys Val Gln Leu
1 5 10 15

His Trp Asp Glu Lys Arg Thr Phe Glu Val Thr Glu Asp Glu Gly Lys
20 25 30

Glu Lys Tyr Tyr Cys Leu Ser Trp Ser Pro Tyr Pro Ser Gly Arg Leu
35 40 45

His Met Gly His Val Arg Asn Tyr Thr Ile Gly Asp Val Ile Ala Arg
50 55 60

Tyr Gln Arg Met Leu Gly Lys Asn Val Leu Gln Pro Ile Gly Trp Asp
65 70 75 80

Ala Phe Gly Leu Pro Ala Glu Gly Ala Ala Val Lys Asn Asn Thr Ala
85 90 95

Pro Ala Pro Trp Thr Tyr Asp Asn Ile Ala Tyr Met Lys Asn Gln Leu
100 105 110

Lys Met Leu Gly Phe Gly Tyr Asp Trp Ser Arg Glu Leu Ala Thr Cys
115 120 125

Thr Pro Glu Tyr Tyr Arg Trp Glu Gln Lys Phe Phe Thr Glu Leu Tyr
130 135 140

Lys Lys Gly Leu Val Tyr Lys Lys Thr Ser Ala Val Asn Trp Cys Pro
145 150 155 160

Asn Asp Gln Thr Val Leu Ala Asn Glu Gln Val Ile Asp Gly Cys Cys
165 170 175

Trp Arg Cys Asp Thr Lys Val Glu Arg Lys Glu Ile Pro Gln Trp Phe
180 185 190

Ile Lys Ile Thr Ala Tyr Ala Asp Glu Leu Leu Asn Asp Leu Asp Lys
195 200 205

Leu Asp His Trp Pro Asp Thr Val Lys Thr Met Gln Arg Asn Trp Ile
210 215 220

Gly Arg Ser Glu Gly Val Glu Ile Thr Phe Asn Val Asn Asp Tyr Asp
225 230 235 240

Asn Thr Leu Thr Val Tyr Ala Ser Arg Pro Asp Thr Phe Met Gly Cys
245 250 255

Thr Tyr Leu Ala Val Ala Ala Gly His Pro Leu Ala Gln Lys Ala Ala
260 265 270

Glu Asn Asn Pro Glu Leu Ala Ala Phe Ile Asp Glu Cys Arg Asn Thr
275 280 285

Lys Val Ala Glu Ala Glu Met Ala Thr Met Glu Lys Lys Gly Val Asp
290 295 300

Thr Gly Phe Lys Ala Val His Pro Leu Thr Gly Glu Glu Ile Pro Val
305 310 315 320

Trp Ala Ala Asn Phe Val Leu Met Glu Tyr Gly Thr Gly Ala Val Met
325 330 335

Ala Val Pro Gly His Asp Gln Arg Asp Tyr Glu Phe Ala Ser Lys Tyr
340 345 350

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

Asp Leu Ser Gln Gln Ala Leu Thr Glu Lys Gly Val Leu Phe Asn Ser
370 375 380

Gly Glu Phe Asn Gly Leu Asp His Glu Ala Ala Phe Asn Ala Ile Ala
385 390 395 400

Asp Lys Leu Thr Ala Met Gly Val Gly Glu Arg Lys Val Asn Tyr Arg
405 410 415

Leu Arg Asp Trp Gly Val Ser Arg Gln Arg Tyr Trp Gly Ala Pro Ile
420 425 430

Pro Met Val Thr Leu Glu Asp Gly Thr Val Met Pro Thr Pro Asp Asp
435 440 445

Gln Leu Pro Val Ile Leu Pro Glu Asp Val Val Met Asp Gly Ile Thr
450 455 460

Ser Pro Ile Lys Ala Asp Pro Glu Trp Ala Lys Thr Thr Val Asn Gly
465 470 475 480

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

Cys Trp Ile Tyr Ala Arg Tyr Thr Cys Pro Gln Tyr Lys Glu Gly Met
500 505 510

Leu Asp Ser Glu Ala Ala Asn Tyr Trp Leu Pro Val Asp Ile Ala Ile

515

520

525

Gly Gly Ile Glu His Ala Ile Met Gly Leu Leu Tyr Phe Arg Phe Phe
530 535 540

His Lys Leu Met Arg Asp Ala Gly Met Val Asn Ser Asp Glu Pro Ala
545 550 555 560

Lys Gln Leu Leu Cys Gln Gly Met Val Leu Ala Asp Ala Phe Tyr Tyr
565 570 575

Val Gly Glu Asn Gly Glu Arg Asn Trp Val Ser Pro Val Asp Ala Ile
580 585 590

Val Glu Arg Asp Glu Lys Gly Arg Ile Val Lys Ala Lys Asp Ala Ala
595 600 605

Gly His Glu Leu Val Tyr Thr Gly Ile Ser Lys Met Ser Lys Ser Lys
610 615 620

Asn Asn Gly Ile Asp Pro Gln Val Met Val Glu Arg Tyr Gly Ala Asp
625 630 635 640

Thr Val Arg Leu Phe Met Met Phe Ala Ser Pro Ala Asp Met Thr Leu
645 650 655

Glu Trp Gln Glu Ser Gly Val Glu Gly Ala Asn Arg Phe Leu Lys Arg
660 665 670

Ala Trp Lys Leu Val Tyr Glu His Thr Ala Lys Gly Asp Val Ala Ala
675 680 685

Leu Asn Val Asp Ala Leu Thr Glu Asn Gln Lys Ala Leu Arg Arg Asp
690 695 700

Val His Lys Thr Ile Ala Lys Val Thr Asp Asp Ile Gly Arg Arg Gln

705 710 715 720

Thr Phe Asn Thr Ala Ile Ala Ala Ile Met Glu Leu Met Asn Lys Leu
725 730 735

Ala Lys Ala Pro Thr Asp Gly Glu Gln Asp Arg Ala Leu Met Gln Glu
740 745 750

Ala Leu Leu Ala Val Val Arg Met Leu Asn Pro Phe Thr Pro His Ile
755 760 765

Cys Phe Thr Leu Trp Gln Glu Leu Lys Gly Glu Gly Asp Ile Asp Asn
770 775 780

Ala Pro Trp Pro Val Ala Asp Glu Lys Ala Met Val Glu Asp Ser Thr
785 790 795 800

Leu Val Val Val Gln Val Asn Gly Lys Val Arg Ala Lys Ile Thr Val
805 810 815

Pro Val Asp Ala Thr Glu Glu Gln Val Arg Glu Arg Ala Gly Gln Glu
820 825 830

His Leu Val Ala Lys Tyr Leu Asp Gly Val Thr Val Arg Lys Val Ile
835 840 845

Tyr Val Pro Gly Lys Leu Leu Asn Leu Val Val Gly
850 855 860

<210>	102
<211>	15
<212>	PRT
<213>	Artificial

<220>
<223> tryptic peptide from enhanced GFP

<400> 102

Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys
1 5 10 15

<210> 103
<211> 15
<212> PRT
<213> Artificial

<220>
<223> tryptic peptide from enhanced GFP

<220>
<221> MISC_FEATURE
<222> (13)..(13)
<223> X is tyrosine, p-methoxy-L-phenylalanine,

p-acetyl-L-phenylalanine, p-benzoyl-L-phenylalanine,
p-iodo-L-phenylalanine, p-azido-L-phenylalanine, or
p-propargyloxyphenylalanine

<400> 103

Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Xaa Gly Lys
1 5 10 15

<210> 104
<211> 55
<212> DNA
<213> Homo sapiens

<400> 104
agcgcctcgg tttttctgtg ctgaacctca ggggacgccg acacacgtac acgtc 55

<210> 105
<211> 25
<212> DNA
<213> Artificial

<220>
<223> 5'-flanking sequence

<400> 105

gatccgaccg tgtgcttggc agaac

25

<210> 106

<211> 12

<212> DNA

<213> Artificial

<220>

<223> 3'-flanking sequence

<400> 106

gtcctttttt tg

12