



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0024706
(43) 공개일자 2009년03월09일

(51) Int. Cl.

C12N 15/52 (2006.01) C12N 15/74 (2006.01)
C12N 15/09 (2006.01) C07K 14/31 (2006.01)

(21) 출원번호 10-2008-7030491

(22) 출원일자 2008년12월15일

심사청구일자 없음

번역문제출일자 2008년12월15일

(86) 국제출원번호 PCT/US2007/012361

국제출원일자 2007년05월21일

(87) 국제공개번호 WO 2007/139870

국제공개일자 2007년12월06일

(30) 우선권주장

60/808,099 2006년05월23일 미국(US)

60/813,856 2006년06월14일 미국(US)

(71) 출원인

더 스크립스 리서치 인스티튜트

미국 캘리포니아주 92037 라 콜라 노스 토리 파
인 로드 10550

(72) 발명자

왕 지양윤

미국 캘리포니아주 92122 샌 디에고 아파트먼트
6323 팔밀라 드라이브 7675

크시 지양밍

미국 캘리포니아주 샌디에고 #2310 코스타 베르드
불바드 8730

술츠 피터 지

미국 캘리포니아주 92037 라 콜라 라 콜라 란쵸
로드 1650

(74) 대리인

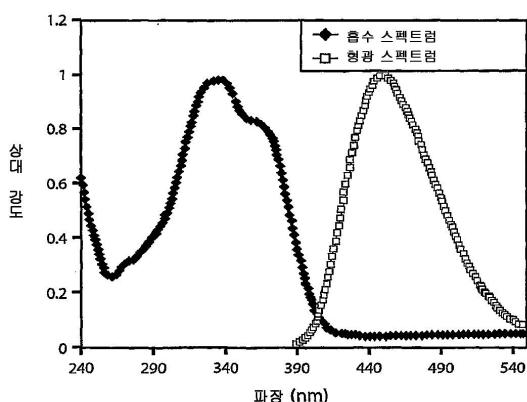
김성기, 김진희

전체 청구항 수 : 총 40 항

(54) 유전자적으로 코딩된 형광 쿠마린 아미노산

(57) 요 약

본 발명은 *E. coli*(*E. coli*)와 같은 진정세균 숙주 세포에서 생성되는 단백질 내로 쿠마린 비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신을 도입할 수 있는, tRNA 및 아미노아실-tRNA 합성효소의 오르소고날 쌍에 관한 것이다. 본 발명은 예를 들어, 단 비제한적으로, 신규 오르소고날 합성효소, 그 신규 합성효소의 동정 및 제조 방법, 비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신을 함유하는 단백질의 생성 방법, 및 관련 번역 시스템을 제공한다.

대 표 도 - 도2

특허청구의 범위

청구항 1

- (a) L-(7-하이드록시쿠마린-4-일)에틸글리신 비천연 아미노산인 제1 비천연 아미노산;
 - (b) 제1 오르소고날 아미노아실-tRNA 합성효소(O-RS); 및
 - (c) 제1 오르소고날 tRNA(O-tRNA)를 포함하는 번역 시스템으로서,
- 상기 제1 O-RS는 상기 제1 O-tRNA를 상기 제1 비천연 아미노산으로 우선적으로 아미노아실화하는 번역 시스템.

청구항 2

제1항에 있어서, 상기 제1 O-RS가 메타노코커스 잔나스키이(*Methanococcus jannaschii*) 아미노아실-tRNA 합성효소로부터 유래되는 번역 시스템.

청구항 3

제1항에 있어서, 상기 제1 O-RS가 야생형 메타노코커스 잔나스키이 티로실-tRNA 합성효소로부터 유래되는 번역 시스템.

청구항 4

제1항에 있어서, 상기 제1 O-RS가 서열 번호 4에 나타낸 아미노산 서열 및 이의 보존적 변이체를 포함하는 번역 시스템.

청구항 5

제1항에 있어서, 상기 제1 O-tRNA가 앰버 억제자 tRNA인 번역 시스템.

청구항 6

제1항에 있어서, 상기 제1 O-tRNA가 서열 번호 1에 나타낸 폴리뉴클레오티드 서열을 포함하거나 상기 서열에 의해 코딩되는 번역 시스템.

청구항 7

제1항에 있어서,

관심 단백질을 코딩하는 핵산을 추가로 포함하고,

상기 핵산은 상기 제1 O-tRNA에 의해 인식되는 하나 이상의 셀렉터 코돈을 포함하는 번역 시스템.

청구항 8

제7항에 있어서,

제2 O-RS 및 제2 O-tRNA를 추가로 포함하고,

상기 제2 O-RS는 제2 O-tRNA를 제1 비천연 아미노산과 상이한 제2 비천연 아미노산으로 우선적으로 아미노아실화하며, 상기 제2 O-tRNA는 제1 O-tRNA에 의해 인식되는 셀렉터 코돈과 상이한 셀렉터 코돈을 인식하는 번역 시스템.

청구항 9

제1항에 있어서, 상기 제1 비천연 아미노산, 상기 제1 O-RS, 및 상기 제1 O-tRNA를 포함한 숙주 세포를 포함하는 번역 시스템.

청구항 10

제9항에 있어서, 상기 숙주 세포가 진정세균 세포인 번역 시스템.

청구항 11

제10항에 있어서, 상기 진정세균 세포가 *E. coli* 세포인 번역 시스템.

청구항 12

제9항에 있어서, 상기 숙주 세포가 상기 제1 O-RS를 코딩하는 폴리뉴클레오티드를 포함하는 번역 시스템.

청구항 13

제12항에 있어서, 상기 폴리뉴클레오티드가 서열 번호 5에 나타낸 뉴클레오티드 서열을 포함하는 번역 시스템.

청구항 14

제9항에 있어서, 상기 숙주 세포가 상기 제1 O-tRNA를 코딩하는 폴리뉴클레오티드를 포함하는 번역 시스템.

청구항 15

번역 시스템 내에서 선택된 위치에 비천연 아미노산을 포함하는 단백질을 생성시키는 방법으로서,

- (a) (i) L-(7-히드록시쿠마린-4-일)에틸글리신 비천연 아미노산인 제1 비천연 아미노산;
- (ii) 제1 오르소고날 아미노아실-tRNA 합성효소(O-RS);
- (iii) 제1 오르소고날 tRNA(O-tRNA)(여기에서, 상기 제1 O-RS는 상기 제1 O-tRNA를 상기 비천연 아미노산으로 우선적으로 아미노아실화함); 및
- (iv) 상기 단백질을 코딩하는 핵산(여기에서, 상기 핵산은 상기 제1 O-tRNA에 의해 인식되는 하나 이상의 셀렉터 코돈을 포함함)

을 포함하는 번역 시스템을 제공하는 단계; 및

- (b) 상기 셀렉터 코돈에 대한 반응으로 상기 단백질의 번역 중에 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 번역 시스템의 제공 단계가 상기 O-RS를 코딩하는 폴리뉴클레오티드를 제공하는 것을 포함하는 방법.

청구항 17

제15항에 있어서, 상기 번역 시스템의 제공 단계가 메타노코커스 잔나스키이 아미노아실-tRNA 합성효소로부터 유래된 O-RS를 제공하는 것을 포함하는 방법.

청구항 18

제15항에 있어서, 상기 번역 시스템의 제공 단계가 야생형 메타노코커스 잔나스키이 티로실-tRNA 합성효소로부터 유래된 O-RS를 제공하는 것을 포함하는 방법.

청구항 19

제15항에 있어서, 상기 번역 시스템의 제공 단계가 서열 번호 4에 나타낸 아미노산 서열 및 이의 보존적 변이체를 포함하는 O-RS를 제공하는 것을 포함하는 방법.

청구항 20

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

청구항 21

제20항에 있어서, 상기 선택 단계가 부위-지정 돌연변이유발 후에 복수개의 생성된 아미노아실-tRNA 합성효소 분자를 포함하는 푸울로부터 상기 O-RS에 대해 양성 선택 또는 음성 선택을 수행하는 것을 포함하는 방법.

청구항 22

제15항에 있어서, 상기 번역 시스템의 제공 단계가 상기 O-tRNA를 코딩하는 폴리뉴클레오티드를 제공하는 것을 포함하는 방법.

청구항 23

제15항에 있어서, 상기 번역 시스템의 제공 단계가 앰버 억제자 tRNA인 O-tRNA를 제공하는 것을 포함하는 방법.

청구항 24

제15항에 있어서, 상기 번역 시스템의 제공 단계가 서열 번호 1에 나타낸 폴리뉴클레오티드 서열을 포함하거나 상기 서열에 의해 코딩되는 O-tRNA를 제공하는 것을 포함하는 방법.

청구항 25

제15항에 있어서, 상기 번역 시스템의 제공 단계가 앰버 셀렉터 코돈을 포함하는 핵산을 제공하는 것을 포함하는 방법.

청구항 26

제15항에 있어서, 상기 단백질이 상기 제1 비천연 아미노산과 상이한 제2 비천연 아미노산을 추가로 포함하고, 상기 번역 시스템이 제2 O-RS 및 제2 O-tRNA를 추가로 포함하며, 제2 O-RS가 제2 O-tRNA를 제1 비천연 아미노산과 상이한 제2 비천연 아미노산으로 우선적으로 아미노아실화하고, 제2 O-tRNA가 제1 O-tRNA에 의해 인식되는 셀렉터 코돈과 상이한 핵산 내의 셀렉터 코돈을 인식하는 방법.

청구항 27

제15항에 있어서, 상기 번역 시스템의 제공 단계가 숙주 세포를 제공하는 것을 포함하고, 상기 숙주 세포가 상기 제1 비천연 아미노산, 상기 제1 O-RS, 상기 제1 O-tRNA 및 상기 핵산을 포함하며, 상기 도입 단계가 상기 숙주 세포를 배양하는 것을 포함하는 방법.

청구항 28

제27항에 있어서, 상기 숙주 세포의 제공은 진정세균 숙주 세포를 제공하는 것을 포함하는 방법.

청구항 29

제28항에 있어서, 상기 진정세균 숙주 세포의 제공은 E. 콜라이 숙주 세포를 제공하는 것을 포함하는 방법.

청구항 30

제27항에 있어서, 상기 숙주 세포의 제공 단계가 상기 O-RS를 코딩하는 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 것을 포함하는 방법.

청구항 31

제30항에 있어서, 상기 O-RS를 코딩하는 폴리뉴클레오티드를 포함하는 숙주 세포의 제공은 서열 번호 5에 나타낸 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드를 포함하는 숙주 세포를 제공하는 것을 포함하는 방법.

청구항 32

제15항에 있어서, 상기 번역 시스템의 제공 단계가 세포 추출물을 제공하는 것을 포함하는 방법.

청구항 33

서열 번호 4에 나타낸 아미노산 서열 및 이의 보존적 변이체를 포함한 폴리펩티드를 포함하는 조성물로서, 보존적 변이체 폴리펩티드가 동족체 오르소고날 tRNA(O-tRNA)를, 상기 O-tRNA, 상기 비천연 아미노산, 및 서열 번호 4의 아미노산 서열을 포함하는 아미노아실-tRNA 합성효소를 포함하는 번역 시스템에 대해 관찰되는 효율의 50% 이상인 효율로, 비천연 아미노산으로 아미노아실화하는 조성물.

청구항 34

제33항의 폴리펩티드를 코딩하는 폴리뉴클레오티드.

청구항 35

제34항에 있어서, 서열 번호 5의 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드.

청구항 36

제33항에 있어서, 폴리펩티드를 포함하는 세포인 조성물.

청구항 37

서열 번호 5에 나타낸 뉴클레오티드 서열을 포함한 폴리뉴클레오티드를 포함하는 조성물.

청구항 38

제37항의 폴리뉴클레오티드를 포함하는 벡터.

청구항 39

제37항의 폴리뉴클레오티드를 포함하는 발현 벡터.

청구항 40

제37항의 폴리뉴클레오티드를 포함한 벡터를 포함하는 세포.

명세서

기술 분야

<1> 관련 출원에 대한 상호 참조

<2> 이 출원은 2006년 5월 23일에 출원된 미국 특허 가출원 일련 번호 제60/808,099호; 및 2006년 6월 14일에 출원된 미국 특허 가출원 일련 번호 제60/813,856호에 대한 우선권 및 이익을 주장하며, 이들의 개시내용 모두는 모든 목적을 위해 전체적으로 본원에 참조 인용된다.

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

<4> 본 발명은 허가 번호 제ER46051호 하에 에너지부로부터의 정부 지원으로 이루어졌다. 당 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 tRNA, 오르소고날 아미노아실-tRNA 합성효소, 및 이들의 쌍의 제조 및 이용을 위한 조성물 및 방법에 관한 것이다. 본 발명은 또한 그러한 쌍을 이용한 세포 내 단백질의 생성 방법 및 이 방법에 의해 제조된 단백질에 관한 것이다.

배경 기술

<7> 형광은 그것의 고감도 및 취급 안전성으로 인해 생물과학에서 가장 중요한 검출 신호들 중 하나가 되었다. 형광 공명 에너지 전이(FRET) 또는 형광 분극화와 같은 공정으로 인해, 바이오분자 결합 발생, 이동 또는 배좌 변화의 실시간 분석이 가능하게 된다. 형광 프로브로 단백질을 선택적으로 변형시키는 능력은 단백질 구조 및 기능의 시험관내 및 생체내 연구 모두를 크게 촉진하였다(Hermanson(1996) in *Bioconjugate Techniques*, Academic

Press: San Diego; 및 Tsien(1998) *Annu. Rev. Biochem.*, 67:509-544).

<8>

생체내 단백질을 연구하기 위한 현 형광 방법은 종종 녹색 형광 단백질(GFP)과 같은 큰 형광 단백질을 갖는 융합 작제물에 의존한다. 이 프로브는 단백질 발현, 국소화 및 이분자 상호작용의 연구에 유용한 것으로 입증되었으나, 그것의 큰 크기는 유의적 구조 불안을 초래할 수 있다. GFP 융합은 또한 표적 단백질의 C- 또는 N-말단에 한정되고, 비교적 환경에 대해 민감하지 않다(Tsien(1998) *Annu. Rev. Biochem.*, 67:509-44). GFP는 또한 적당한 신호를 달성하기 위한 많은 전사체를 필요로 하고, 그것의 폴딩 및 형광체 성숙을 위해 자체 시간을 필요로 한다.

<9>

화학 방법을 또한 사용하여, 구조적 불안을 최소화하는 각종 작은 합성 형광체로 단백질을 선택적으로 변형시킬 수 있다(Hermanson, in *Bioconjugate Techniques*, Academic Press: San Diego(1996); Martin et al., *Nat. Biotech.*, 23:1308-1314(2005); Keppler et al., *Nat. Biotech.*, 21:86-89(2003); Lin et al., *J. Am. Chem. Sci.*, 128:4542-3(2006)). 그러나, 이 기법들은 일반적으로 단리된 단백질 상의 특유하게 반응성인 표면 접근성 잔기에 한정되거나(예를 들어, 말레이미드 유도체를 이용한 시스테인의 변형; 문헌 [Hermanson in *Bioconjugate Techniques*(1996) Academic Press: San Diego] 참조), 약한 위치선택성을 나타내거나, 세포독성이거나, 염료 결합 단백질 모티프의 도입을 필요로 한다.

<10>

화학적으로 미스아실화된 tRNA를 이용하는 생합성 표지 방법(Mendel et al., *Annu. Rev. Biophys. Biomol. Struct.* (1995) 24:435-462; Hohsaka et al., *FEBS Lett.* (2004) 560:173-177)이 실험적으로 입증되었으나, 단백질의 한정된 수율을 제공하고, 전형적으로 단지 시험관내로만 수행된다.

<11>

생물 유기체 내 직접적으로 단백질 내 한정된 위치에 유전자적으로 코딩된 형광 아미노산을 도입함으로써, 단백질의 형광에 있어 많은 한계점을 극복하게 된다(Wang and Schultz, *Angew. Chem. Int. Ed.* (2005) 44:34- 66; Wang et al., *Annu. Rev. Biophys. Biomol. Struct.*, (2006) 35:225-249). 형광 아미노산의 부위-특이적 도입은 숙주 단백질에 대한 불안을 최소화하고, 더욱 큰 정확도로의 형광 공명 에너지 전이(FRET)의 측정을 허용한다 (Truong and Ikura, *Curr. Opin. Struct. Bio.* 2001, 11:573-578). 또한, 형광 아미노산의 사용은 각 아미노산 위치의 국소 환경의 탐침을 허용하고, 단백질 내 형광 아미노산의 위치를 변화시킴으로써 다른 세포내 성분과의 상호작용을 매개하는 잔기를 정확히 나타내도록 할 것이다. 이는 또한 한 단백질 분자가 정상적으로 하나 초과의 트립토판 잔기를 함유하고, 형광 프로브를 이용한 단백질의 특이적 화학 표지가 문제가 있기 때문에, 특히 단일 분자 시스템에서(Lipman et al., *Science*(2003) 301:1233-1235), 단백질 폴딩을 연구하는 데 매우 유용할 것이다(Lakowicz, *Principles of Fluorescence Spectroscopy* Ed. 2; Kluwer Academic/Plenum Publishers: New York, 1999).

<12>

당업계에 필요한 것은, 단백질 구조 및 기능을 연구하기 위한 목적으로, 형광 비천연 아미노산을 단백질 내로 도입하기 위한 신규 기법이다. 원핵세포 유기체 및 진핵세포 유기체 모두에서 단백질 내로 다양한 비천연 아미노산을 생체내 부위-특이적으로 도입하기 위한 일반 방법론이 개발되었다. 이 방법은 생체내 폴리펩티드 번역 중에 한정된 위치에 원하는 비천연 아미노산을 삽입하기 위해 적당한 셀렉터 코돈을 인식하는 오르소고날 단백질 번역 성분에 기초한다. 이 방법은 셀렉터 코돈을 인식하는 오르소고날 tRNA(O-tRNA)를 이용하는데, 여기에서 상응하는 특정 오르소고날 아미노아실-tRNA 합성효소(O-RS)는 O-tRNA를 비천연 아미노산으로 충전한다. 이 성분들은 숙주 세포 내 내인성 tRNA, RS, 아미노산 또는 코돈 중 어느 것과도 상호작용하지 않는다(즉, 그것은 오르소고날성이어야 함). 그러한 오르소고날 tRNA-RS 쌍의 사용으로, 상당수의 구조적으로 다양한 비천연 아미노산들을 유전자적으로 코딩할 수 있게 되었다.

<13>

하나 이상의 비천연 아미노산을 포함하는 단백질의 제조를 위해 적당한 오르소고날 번역 시스템을 이용하는 수행이, 오르소고날 번역 시스템의 일반 생성 방법과 같이 당업계에 일반적으로 공지되어 있다. 예를 들어, 국제 특허 공개 공보 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"); 2004년 7월 7일에 출원된 제WO 2005/019415호; 2004년 7월 7일에 출원된 제WO 2005/007870호; 2004년 7월 7일에 출원된 제WO 2005/007624호; 및 2005년 10월 27일에 출원된 제WO 2006/110182호(발명의 명칭: "ORTHOGONAL TRANSLATION COMPONENTS FOR THE IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")를 참조한다. 이 출원들은 각기 전체적으로 본원에 참조 인용된다. 비천연 아미노산을 도입하는 오르소고날 번역 시스템 및 이의 생성 방법, 및 용도에 대한 추가 논의에 대해, 문헌들 [Wang and Schultz, "Expanding the Genetic Code", *Chem. Commun. (Comb.)* 1:1-11(2002); Wang and Schultz, "Expanding the Genetic Code",

Angewandte Chemie Int. Ed., 44(1):34-66(2005); Xie and Schultz, "An Expanding Genetic Code", *Methods* 36(3):227-238(2005); Xie and Schultz, "Adding Amino Acids to the Genetic Repertoire", *Curr. Opinion in Chemical Biology* 9(6):548-554(2005); Wang et al., "Expanding the Genetic Code", *Annu. Rev. Biophys. Biomol. Struct.*, 35:225-249(2006); 및 Xie and Schultz, "a Chemical toolkit for Proteins - an expanded Genetic Code", *Nat. Rev. Mol. Cell. Biol.*, 7(10):775-782(2006)]를 또한 참조하고, 이들의 내용은 각기 전체적으로 본원에 참조 인용된다.

<14> 한정된 위치에 도입될 수 있는 형광 비천연 아미노산을 단백질 내로 도입하는 오르소고날 번역 성분의 개발이 당업계에 필요하다. 본원에 기재된 본 발명은 하기 개시내용을 검토할 때 명백해지는 바와 같이, 상기 필요 및 기타 필요를 이행한다.

발명의 상세한 설명

<15>

발명의 개요

<16> 본 발명은 생체내(예를 들어, 숙주 세포 내) 셀렉터 코돈, 예를 들어 앰버 종결 코돈에 대한 반응으로 쿠마린 비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신을 성장하는 폴리펩티드 사슬 내에 도입하기 위한 조성물 및 방법을 제공한다. 이 조성물은 숙주 세포 번역 기구와 상호작용하지 않는, 오르소고날-tRNA(0-tRNA) 및 오르소고날 아미노아실-tRNA 합성효소(0-RS)의 쌍을 포함한다. 즉, 0-tRNA는 내인성 숙주 세포 아미노아실-tRNA 합성효소에 의해 (천연 또는 비천연) 아미노산으로 충전되지 않는다(또는 유의적 수준으로 충전되지 않음). 유사하게, 본 발명에 의해 제공되는 0-RS는 유의적이거나 검출가능한 수준까지 임의의 내인성 tRNA를 (천연 또는 비천연) 아미노산으로 충전하지 않는다. 이 신규 조성물은 번역 도입된 L-(7-히드록시쿠마린-4-일)에틸글리신을 갖는 단백질의 대량 생산을 허용한다. 이 형광 표지 단백질은 매우 다양한 용도들을 가진다.

<17> 일부 측면에서, 본 발명은 번역 시스템을 제공한다. 이 시스템은 제1 오르소고날 아미노아실-tRNA 합성효소(0-RS), 제1 오르소고날 tRNA(0-tRNA), 및 L-(7-히드록시쿠마린-4-일)에틸글리신인 제1 비천연 아미노산을 포함하는데, 여기에서 제1 0-RS는 제1 0-tRNA를 제1 비천연 아미노산으로 우선적으로 아미노아실화한다.

<18> 번역 시스템은 각종 공급원들로부터 유래된 성분들을 사용할 수 있다. 한 실시양태에서, 제1 0-RS는 메타노코커스 잔나스키이(*Methanococcus jannaschii*) 아미노아실-tRNA 합성효소, 예를 들어 야생형 메타노코커스 잔나스키이 티로실-tRNA 합성효소로부터 유래된다. 시스템 내에 사용되는 0-RS는 서열 번호 4의 아미노산 서열, 및 그 서열의 보존적 변이체를 포함할 수 있다. 일부 실시양태에서, 0-tRNA는 앰버 억제자 tRNA이다. 일부 실시양태에서, 0-tRNA는 서열 번호 1을 포함하거나 그 서열에 의해 코딩된다.

<19> 일부 측면에서, 번역 시스템은 관심 단백질을 코딩하는 핵산을 추가로 포함하는데, 여기에서 핵산은 0-tRNA에 의해 인식되는 하나 이상의 셀렉터 코돈을 가진다.

<20> 일부 측면에서, 번역 시스템은 제2 비천연 아미노산을 이용하는 제2 오르소고날 쌍(즉, 제2 0-RS 및 제2 0-tRNA)을 도입하며, 이에 따라 그 시스템은 이제 폴리펩티드 내 상이한 선택 부위에 2개 이상의 상이한 비천연 아미노산을 도입할 수 있다. 이 이중 시스템에서, 제2 0-RS는 제2 0-tRNA를 제1 비천연 아미노산과 상이한 제2 비천연 아미노산으로 우선적으로 아미노아실화하고, 제2 0-tRNA는 제1 0-tRNA에 의해 인식되는 셀렉터 코돈과 상이한 셀렉터 코돈을 인식한다.

<21> 일부 실시양태에서, 번역 시스템은 숙주 세포 내에 있다(또한, 그 숙주 세포를 포함함). 사용되는 숙주 세포는 0-RS 및 0-tRNA가 그것의 숙주 세포 환경에서 자체의 오르소고날성을 보유하는 한, 특별히 제한되지 않는다. 숙주 세포는 진정세균 세포, 예컨대 *E. coli*일 수 있다. 숙주 세포는 0-RS 또는 0-tRNA를 비롯한, 번역 시스템의 성분을 코딩하는 하나 이상의 폴리뉴클레오티드를 포함할 수 있다. 일부 실시양태에서, 0-RS를 코딩하는 폴리뉴클레오티드는 서열 번호 5의 뉴클레오티드 서열을 포함한다.

<22> 본 발명은 또한 선택된 위치에 하나 이상의 비천연 아미노산을 갖는 단백질의 생성 방법을 제공한다. 이 방법은 상기 기재된 번역 시스템을 이용한다. 일반적으로, 이 방법은 (i) 쿠마린 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신인 제1 비천연 아미노산; (ii) 제1 오르소고날 아미노아실-tRNA 합성효소(0-RS); (iii) 제1 오르소고날 tRNA(0-tRNA)(여기에서, 0-RS는 0-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화함); 및 (iv) 단백질을 코딩하는 핵산(여기에서, 핵산은 제1 0-tRNA에 의해 인식되는 하나 이상의 셀렉터 코돈을 포함함)을 포함하는 번역 시스템을 제공하는 단계로 시작된다. 그 방법은 이어서 셀렉터 코돈에 대한 반응으로 단백질의 번역 중에 단백질 내 선택된 위치에 비천연 아미노산을 도입함으로써, 선택된 위치에 비천연 아미노산을 포함하는 단백

질을 생성시키는 단계를 포함한다.

- <23> 이 방법은 각종 시약 및 단계를 이용하여 폭넓게 적용될 수 있다. 일부 실시양태에서, O-RS를 코딩하는 폴리뉴클레오티드가 제공된다. 일부 실시양태에서, 메타노코커스 잔나스키이 아미노아실-tRNA 합성효소로부터 유래된 O-RS이 제공되는데, 예를 들어 야생형 메타노코커스 잔나스키이 티로실-tRNA 합성효소가 제공될 수 있다. 일부 실시양태에서, 제공 단계는 서열 번호 4의 아미노산 서열 및 이의 보존적 변이체를 포함하는 O-RS를 제공하는 것을 포함한다.
- <24> 이 방법들의 일부 실시양태에서, 번역 시스템의 제공 단계는 야생형 아미노아실-tRNA 합성효소의 아미노산 결합 포켓을 부위-지정 돌연변이유발에 의해 돌연변이화하고, O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는 생성된 O-RS를 선택하는 것을 포함한다. 선택 단계는 부위-지정 돌연변이유발 후에 생성된 아미노아실-tRNA 합성효소 분자의 푸울로부터 O-RS에 대해 양성 선택 또는 음성 선택을 수행하는 것을 포함할 수 있다. 일부 실시양태에서, 제공 단계는 O-tRNA를 코딩하는 폴리뉴클레오티드, 예를 들어 앰버 역제자 tRNA인 O-tRNA, 또는 서열 번호 1의 폴리뉴클레오티드를 포함하거나 이에 의해 코딩되는 O-tRNA를 제공한다. 이 방법에서, 제공 단계는 또한 번역 시스템에 의해 이용되는 앰버 셀렉터 코돈을 포함하는 핵산을 제공할 수 있다.
- <25> 이 방법은 또한 하나 초과의 비천연 아미노산을 단백질 내로 도입하기 위해 변형될 수 있다. 그 방법에서, 제2 오르소고날 번역 시스템은 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.
- <26> 비천연 아미노산을 갖는 단백질의 생성 방법도 또한 숙주 세포의 영역 내에서 수행될 수 있다. 이 경우들에서, 비천연 아미노산, O-RS, O-tRNA, 및 단백질을 코딩하는 하나 이상의 셀렉터 코돈을 갖는 핵산을 포함하는 숙주 세포가 제공되는데, 여기에서 숙주 세포의 배양은 비천연 아미노산의 도입을 초래할 수 있다. 일부 실시양태에서, 제공 단계는 진정세균 숙주 세포(예를 들어, *E. coli*)를 제공하는 것을 포함한다. 일부 실시양태에서, 제공 단계는 O-RS를 코딩하는 폴리뉴클레오티드를 함유하는 숙주 세포를 제공하는 것을 포함한다. 예를 들어, O-RS를 코딩하는 폴리뉴클레오티드는 서열 번호 5의 뉴클레오티드 서열을 포함할 수 있다. 일부 실시양태에서, 번역 시스템의 제공 단계는 세포 추출물을 제공함으로써 달성된다.
- <27> 본 발명은 또한 핵산 및 단백질을 비롯한 각종 조성물을 제공한다. 조성물의 성질은, 특정된 핵산 또는 단백질을 포함하는 것을 제외하고는 특별히 제한되지 않는다. 본 발명의 조성물은 임의의 성질을 갖는 임의의 수의 추가 성분을 포함할 수 있다.
- <28> 예를 들어, 본 발명은 O-RS 폴리펩티드를 포함한 조성물을 제공하는데, 여기에서 폴리펩티드는 서열 번호 4의 아미노산 서열 또는 이의 보존적 변이체를 포함하고, 보존적 변이체 폴리펩티드는 동족체 오르소고날 tRNA(O-tRNA)를 비천연 아미노산으로, O-tRNA, 비천연 아미노산, 및 서열 번호 4의 아미노산 서열을 포함하는 아미노아실-tRNA 합성효소를 포함하는 번역 시스템에 대해 관찰되는 효율의 50% 이상인 효율로, 아미노아실화한다. 본 발명은 또한 상기 이 폴리펩티드들 중 임의의 것을 코딩하는 폴리뉴클레오티드를 제공한다. 일부 실시양태에서, 이 폴리뉴클레오티드들은 서열 번호 5의 뉴클레오티드 서열을 포함할 수 있다. 일부 실시양태에서, 폴리펩티드는 세포 내에 있다.
- <29> 본 발명은 또한 서열 번호 5의 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드 조성물을 제공한다. 일부 실시양태에서, 본 발명은 폴리뉴클레오티드를 포함하는 백터, 예를 들어 발현 백터를 제공한다. 일부 실시양태에서, 본 발명은 상기 백터를 포함하는 세포를 제공한다.
- <30> **정의**
- <31> 본 발명을 상세히 기술하기 전에, 본 발명은 특정 생물학적 시스템에 한정되지 않고, 이 시스템은 물론 다양화 할 수 있음을 이해하도록 한다. 또한, 본원에 사용된 용어들은 특정 실시양태를 단지 기술하고자 하는 목적을 위한 것으로서, 제한적이지 않도록 의도됨을 이해하도록 한다. 본 명세서 및 첨부된 특허청구범위에 사용되는 단수 형태(영문 관사 "a", "an" 및 "the"에 해당함)는 그 내용이 달리 명백히 달리 표시되는 것이 아닌 한, 복수 표현을 포함한다. 따라서, 예를 들어 "세포"라는 표현은 2개 이상의 세포의 조합물을 포함하고; "폴리뉴클레오티드"라는 표현은 실질적으로 그 폴리뉴클레오티드의 많은 카피(copy)를 포함한다.
- <32> 여기에서 및 본 명세서의 하기 나머지 부분에서 달리 정의하지 않는 한, 본원에서 사용된 모든 기술적 용어 및

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

<33>

오르소고날(orthogonal): 본원에서 사용되는 용어 "오르소고날"은 세포 또는 번역 시스템에 대한 내인성을 가진 상응하는 문자에 비해 감소된 효율을 가진 세포의 내인성 성분들과 함께 기능하거나, 세포의 내인성 성분들과 함께 기능하지 못하는 문자(예컨대, 오르소고날 tRNA(O-tRNA) 및/또는 오르소고날 아미노아실-tRNA 합성효소(O-RS))를 지칭한다. tRNA 및 아미노아실-tRNA 합성효소와 관련하여, 오르소고날은 내인성 tRNA 합성효소와 함께 기능하는 내인성 tRNA에 비해 내인성 tRNA 합성효소와 함께 기능하는 오르소고날 tRNA, 또는 내인성 tRNA와 함께 기능하는 내인성 tRNA 합성효소에 비해, 내인성 tRNA와 함께 기능하는 오르소고날 아미노아실-tRNA 합성효소의 불능 또는 감소된 효율, 예를 들어 20% 미만의 효율, 10% 미만의 효율, 5% 미만의 효율, 또는 1% 미만의 효율을 지칭한다. 오르소고날 문자는 세포 내에서 기능적으로 정상인 내인성 상보적 문자를 갖지 않는다. 예를 들어, 세포 내에서 오르소고날 tRNA는 내인성 RS에 의한 내인성 tRNA의 아미노아실화에 비해, 감소된 또는 심지어 0% 효율로 세포의 임의의 내인성 RS에 의해 아미노아실화된다. 또 다른 예에서, 오르소고날 RS는 내인성 RS에 의한 내인성 tRNA의 아미노아실화에 비해 감소된 또는 심지어 0% 효율로 관심 세포의 임의의 내인성 tRNA를 아미노아실화한다. 제1 오르소고날 문자와 함께 기능하는 제2 오르소고날 문자를 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 일정한 효율(예컨대, 45% 효율, 50% 효율, 60% 효율, 70% 효율, 75% 효율, 80% 효율, 90% 효율, 95% 효율 또는 99% 이상의 효율)로 세포 내에서 함께 기능하는 도입된 상보적 성분들을 포함한다.

<34>

오르소고날 티로실-tRNA: 본원에서 사용되는 오르소고날 티로실 tRNA(티로실-O-tRNA)는, 관심 번역 시스템에 대해 오르소고날성인 tRNA이며, 여기에서 tRNA는 (1) 천연 발생 티로실-tRNA와 동일하거나 실질적으로 유사하거나, (2) 천연 또는 인공 돌연변이유발에 의해 천연 발생 티로실-tRNA로부터 유래되거나, (3) (1) 또는 (2)의 야생형 또는 돌연변이체 티로실-tRNA의 서열을 고려하는 임의의 과정에 의해 유도되거나, (4) 야생형 또는 돌연변이체 티로실-tRNA와 상동성이거나, (5) 도 9에서 티로실-tRNA 합성효소에 대한 기질로서 표시된 임의의 예시적 tRNA와 상동성을 나타내거나, (6) 도 9에서 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. 임의적으로 "티로실-O-tRNA"는 동족체 합성효소에 의해 티로신 외의 아미노산, 예를 들어 비천연 아미노산으로 충전(아미노아실화)된다는 것도 이해할 것이다. 사실상, 본 발명의 티로실-O-tRNA는 번역 중에 셀렉터 코돈에 대한 반응으로 본질적으로 임의의 천연 또는 인공 아미노산을 성장하는 폴리펩티드 내로 삽입시키는 데 유리하게 사용된다는 것을 인식할 것이다.

<35>

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 적재하는 아미노산은 임의의 천연, 비천연 또는 인공 아미노산일 수 있고, 본원에서 한정되지 않는다. 합성효소는 임의적으로 천연 발생 티로실 아미노산 합성효소와 동일하거나 상동성을 나타내거나, 도 9에서 O-RS로 표시된 합성효소와 동일하거나 상동성을 나타낸다. 예를 들어, O-RS는 도 9의 티로실-O-RS의 보존적 변이체일 수 있고/있거나, 서열에 있어 도 9의 O-RS와 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99% 또는 그 이상 동일할 수 있다.

<36>

동족체: 용어 "동족체"는, 예를 들어 오르소고날 tRNA 및 오르소고날 아미노아실-tRNA 합성효소와 함께 기능하거나, 상호에 대해 일부 특이성 측면을 가지는 성분들을 지칭한다. 그 성분들은 상보적인 것으로서 지칭될 수도 있다.

<37>

우선적으로 아미노아실화한다: 오르소고날 번역 시스템에 대하여, 본원에서 사용되는 O-RS는 그것이 발현 시스템 내에서 임의의 내인성 tRNA를 충전하는 것보다 더 효율적으로 O-tRNA를 아미노산으로 충전할 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 충전하는 것보다 더 자주 O-tRNA를 충전할 것이다. 바람직하게, 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 충전할 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 이상, 보다 더 바람직하게는 10:1 이상, 훨씬 더 바람직하게는 20:1 이상, 훨씬 더 바람직하게는 50:1 이상, 보다 더 바람직하게는 75:1 이상, 훨씬 더 바람직하게는 95:1, 98:1, 99:1, 100:1, 500:1, 1,000:1, 5,000:1, 또

는 그 이상이다.

<38> 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 번역 시스템에 동등한 몰 양으로 존재하는 경우, O-RS는 천연 아미노산보다 더 자주 비천연 아미노산을 O-tRNA에 적재할 것이다. 바람직하게, 천연 아미노산으로 충전된 O-tRNA에 대한 비천연 아미노산으로 충전된 O-tRNA의 상대적 비율이 높다. 보다 바람직하게, O-RS는 O-tRNA만 또는 거의 O-tRNA만을 비천연 아미노산으로 충전한다. 천연 아미노산 및 비천연 아미노산이 동등한 몰 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, 훨씬 더 바람직하게는 20:1, 훨씬 더 바람직하게는 50:1, 보다 더 바람직하게는 75:1, 훨씬 더 바람직하게는 95:1, 98:1, 99:1, 100:1, 500:1, 1,000:1, 5,000:1 또는 그 이상이다.

<39> **셀렉터 코돈:** 용어 "셀렉터 코돈"은 번역 과정에서 O-tRNA에 의해서는 인식되나 내인성 tRNA에 의해서는 인식되지 않는 코돈을 지칭한다. O-tRNA 안티코돈 루프(loop)는 mRNA 상의 셀렉터 코돈을 인식하여, 그것의 아미노산, 예컨대 비천연 아미노산을 폴리펩티드 내의 인식된 부위 새로 도입한다. 셀렉터 코돈은, 예를 들어 넌센스 코돈, 예컨대 종결 코돈, 예를 들어 앰버(amber), 오커(ochre) 및 오펠(opal) 코돈; 4개 이상의 염기로 된 코돈; 희귀 코돈; 천연 또는 비천연 염기 쌍으로부터 유래된 코돈 등을 포함할 수 있다.

<40> **억제자 tRNA:** 억제자 tRNA는, 전형적으로 폴리펩티드의 번역 중에 종결 코돈(즉, "통독")에 대한 반응으로 아미노산을 도입함으로써 소정의 번역 시스템 내의 메신저 RNA(mRNA)의 관독을 변경시키는 tRNA이다. 일부 측면에서, 본 발명의 셀렉터 코돈은 억제자 코돈, 예를 들어 종결 코돈(예를 들어, 앰버, 오커 또는 오펠 코돈), 4 염기 코돈, 희귀 코돈 등이다.

<41> **억제 활성:** 본원에서 사용되는 용어 "억제 활성"은, 일반적으로 다른 방식으로라면 번역의 종결 또는 오번역 (mistranslation)(예컨대, 프레임-변동)을 초래하게 될 코돈(예를 들어, 앰버 코돈 또는 4개 이상의 염기로 된 코돈인 셀렉터 코돈)을 통한 번역 통독을 허용하는 tRNA(예를 들어, 억제자 tRNA)의 능력을 지칭한다. 억제자 tRNA의 억제 활성은 제2 억제자 tRNA에 비해, 또는 대조군 시스템, 예컨대 O-RS를 갖지 않은 대조군 시스템에 비해, 관찰된 번역 통독 활성율(%)로서 표시될 수 있다.

<42> 본 발명은 억제 활성을 정량화할 수 있는 다양한 방법을 제공한다. 관심 셀렉터 코돈(예컨대, 앰버 코돈)에 대한 특정 O-tRNA 및 O-RS의 억제율(%)은 발현된 시험 마커를 코딩하는 핵산 내에 셀렉터 코돈을 포함하는 소정의 발현된 시험 마커(예컨대, LacZ)가 관심 번역 시스템에서 나타내는 활성율(%)을 지칭하는데, 여기에서 원하는 발현 시스템은 양성 대조군 작제물(construct)에 비해 O-RS 및 O-tRNA를 포함하는데, 여기에서 양성 대조군은 O-tRNA, O-RS 및 셀렉터 코돈을 갖지 않는다. 따라서, 예를 들어 셀렉터 코돈을 갖지 않은 활성 양성 대조군 마커 작제물이 소정의 번역 시스템에서 해당 마커 분석에 대한 유니트(unit)로 X의 관찰된 활성을 나타내는 경우, 셀렉터 코돈을 포함하는 시험 작제물의 억제율(%)은, 시험 마커 작제물이 O-tRNA 및 O-RS를 포함하는 번역 시스템에서 발현된다는 점을 제외하고 양성 대조군 마커가 발현된 환경 조건과 본질적으로 동일한 환경 조건 하에서 시험 마커 작제물이 나타내는 X의 백분율(%)이다. 전형적으로, 시험 마커를 발현하는 번역 시스템은 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 작제물과 시험 마커의 비교에 의해 달성된다. 이 음성 대조군은 관심 번역 시스템에서 마커로부터의 배경 신호 효과를 설명하기 위해 억제율 측정을 표준화하는 데 유용하다.

<43> 억제 효율은 당업계에 공지된 다수의 검정 중 임의의 검정에 의해 측정될 수 있다. 예를 들어, β -갈락토시다제 리포터 검정이 이용될 수 있는데, 예를 들어 유도체화된 lacZ 플라스미드(여기에서, 작제물은 lacZ 핵산 서열 내에 셀렉터 코돈을 가짐)가 본 발명의 O-tRNA를 포함하는 플라스미드와 함께 적절한 유기체(예컨대, 오르소고 날 성분들이 사용될 수 있는 유기체)로부터의 세포 내로 도입된다. 동족체 합성효소도 (또한, 발현될 때, 동족체 합성효소를 코딩하는 폴리펩티드 또는 폴리튜클레오티드로서) 도입될 수 있다. 세포는 원하는 밀도, 예를 들어 약 0.5의 OD₆₀₀까지 배지 중에서 성장시키고, β -갈락토시다제 검정을, 예컨대 베타플루오르(BetaFluor)TM β -갈락토시다제 분석 키트(노바젠(Novagen))를 사용하여 수행한다. 억제율(%)은 비교 대조군에 대한 상대적인 샘플의 백분율(%)로서, 예를 들어 셀렉터 코돈보다 오히려 소정의 위치에서 상응하는 센스 코돈을 가진 유도체화

된 lacZ 작제물로부터 관찰된 값으로서 계산될 수 있다.

- <44> **번역 시스템**: 용어 "번역 시스템"은, 아미노산을 성장하는 폴리펩티드 사슬(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 비진핵세포 세포, 예컨대 세균(예컨대, *E. 콜라 이*), 또는 진핵세포 세포, 예를 들어 효모 세포, 포유동물 세포, 식물 세포, 해조류 세포, 진균 세포, 곤충 세포 등에 첨가될 수 있거나 이들의 일부일 수 있다.
- <45> **비천연 아미노산**: 본원에서 사용되는 용어 "비천연 아미노산"은 20종의 공통된 천연 발생 아미노산 중 하나, 셀레노 시스테인 또는 피롤리신이 아닌, 임의의 아미노산, 변형된 아미노산 및/또는 아미노산 유사체를 지칭한다. 예를 들어, 비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신(7-히드록시쿠마린-에틸글리신으로도 기재됨: 도 1, 구조 1 참조)이 본 발명에서 사용될 수 있다.
- <46> **~로부터 유래된**: 본원에 사용되는 용어 "~로부터 유래된"은 특정된 문자 또는 유기체로부터 단리되거나, 특정된 문자 또는 유기체, 또는 특정한 문자 또는 유기체로부터의 정보를 사용하여 만든 성분을 지칭한다. 예를 들어, 제2 폴리펩티드로부터 유래된 폴리펩티드는 제2 폴리펩티드의 아미노산 서열과 동일하거나 실질적으로 유사한 아미노산 서열을 포함할 수 있다. 폴리펩티드의 경우, 유래된 종은, 예를 들어 천연 발생 돌연변이유발, 인공 지정 돌연변이유발 또는 인공 랜덤 돌연변이유발에 의해 수득될 수 있다. 폴리펩티드를 유도하는 데 이용되는 돌연변이유발은 의도적으로 지정될 수 있거나 의도적으로 무작위적일 수 있거나, 각각의 혼합식일 수 있다. 제1 폴리펩티드로부터 유래된 상이한 폴리펩티드를 생성하기 위한 폴리펩티드의 돌연변이유발은 (예를 들어, 중합효소 비신뢰성에 의해 유발되는) 무작위적 발생일 수 있고, 유래된 폴리펩티드는, 예를 들어 본원에서 논의된 바와 같은, 적절한 선별 방법에 의해 동정될 수 있다. 폴리펩티드의 돌연변이유발은 전형적으로 그 폴리펩티드를 코딩하는 폴리뉴클레오티드의 조작을 수반한다.
- <47> **양성 선택 또는 선별 마커**: 본원에서 사용되는 용어 "양성 선택 또는 선별 마커"는 존재하는 경우, 예컨대 발현되거나 활성화되는 경우 등의 특징을 갖는 세포, 예를 들어 양성 선택 마커를 가진 세포를 그 특징을 갖지 않는 세포로부터 동정하도록 하는 마커를 지칭한다.
- <48> **음성 선택 또는 선별 마커**: 본원에서 사용되는 용어 "음성 선택 또는 선별 마커"는 존재하는 경우, 예컨대 발현되거나 활성화되는 경우 등의(예를 들어, 선택된 성질 또는 특징을 갖는 세포에 비해) 상기 성질 또는 특징을 갖지 않는 세포를 동정하도록 하는 마커를 지칭한다.
- <49> **리포터**: 본원에서 사용되는 용어 "리포터"는 관심 시스템의 표적 성분을 동정하고/하거나 선별하는 데 사용될 수 있는 성분을 지칭한다. 예를 들어, 리포터는 항생제 내성 또는 민감성을 부여하는 단백질, 예컨대 효소(예를 들어, β -락타마제, 클로람페니콜 아세틸트랜스페라제(CAT) 등), 형광 선별 마커(예를 들어, 녹색 형광 단백질(예를 들어, GFP), YFP, EGFP, RFP 등), 발광 마커(예를 들어, 초파리 루시페라제 단백질), 친화성 기초 선별 마커, 또는 양성 또는 음성 선택 가능한 마커 유전자, 예컨대 lacZ, β -ga1/lacZ(β -갈락토시다제), ADH(알코올데히드로게나제), his3, ura3, leu2, lys2 등을 포함할 수 있다.
- <50> **진핵생물**: 본원에서 사용되는 용어 "진핵생물"은 진핵생물 계에 속하는 유기체를 지칭한다. 진핵생물은 일반적으로 그들의 전형적인 다세포 조직화(단, 오로지 다세포 조직화로만 한정되지 않음; 예컨대 효모), 막-결합 핵 및 다른 막-결합 소기관의 존재, 선형 유전 물질(예컨대, 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. 진핵생물에는 예를 들어 동물(예컨대, 포유동물, 곤충, 과충류, 조류 등), 섬모충, 식물(예컨대, 단자엽, 쌍자엽, 해조류 등), 진균, 효모, 편모충, 미포자충(microsporidia), 원생생물 등이 포함된다.
- <51> **원핵생물**: 본원에서 사용되는 용어 "원핵생물"은 모네라 계(Kingdom Monera)(원핵생물로도 지칭됨)에 속하는 유기체를 지칭한다. 원핵생물은 일반적으로 그것의 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 진정세균(Eubacteria) 및 고세균(Archaea)(종종 "원시세균(Archaebacteria)"으로 칭해짐)이 포함된다. 경우에 따라 시아노박테리아(청녹색 해조류) 및 마이코플라즈마로 모네라계 하에 별도 분류된다.
- <52> **세균**: 본원에서 사용되는 용어 "세균" 및 "진정세균"은 고세균으로부터 구별될 수 있는 원핵생물을 지칭한다. 유사하게, 고세균은 진정세균으로부터 구별될 수 있는 원핵생물을 지칭한다. 진정세균과 고세균은 다수의 형태

학적 및 생화학적 기준에 의해 구별될 수 있다. 예를 들어, 리보좀 RNA 서열에서의 차이, RNA 종합효소 구조, 인트론의 존재 또는 부재, 항생제 민감성, 세포 벽 웨პ티도글리칸 및 다른 세포 벽 성분들의 존재 또는 부재, 막지질의 분자 구조 대 비분자 구조, 및 히스톤 및 히스톤-유사 단백질의 존재/부재를 이용하여, 유기체를 진정세균 또는 고세균으로 분류한다.

<53> 진정세균의 예에는 *E. 콜라이*, *써모필러스(Thermus thermophilus)*, *바실러스 십틸리스(Bacillus subtilis)* 및 *바실러스 스테아로써모필러스(Bacillus stearothermophilus)*가 포함된다. 고세균의 예에는 *메타노코카스 잔나스키아(Mj)*, *메타노사르시나 마제이(Methanosarcina mazei)(Mm)*, *메타노박테리움 썬모오토트로피쿰(Methanobacterium thermoautotrophicum)(Mt)*, *메타노코카스 마리팔루디스(Methanococcus maripaludis)*, *메타노파이러스 칸들레리(Methanopyrus kandleri)*, *할로박테리움 예컨대 할로페락스 볼케니아이(Haloferax volcanii)* 및 *할로박테리움 종 NRC-1, 아카에오글로버스 풀기더스(Archaeoglobus fulgidus)(Af)*, *파이로코카스 푸리오서스(Pyrococcus furiosus)(Pf)*, *파이로코카스 호리코쉬이(Pyrococcus horikoshii)(Ph)*, *파이로바큘럼 애로필럼(Pyrobaculum aerophilum)*, *파이로코카스 애바이씨(Pyrococcus abyssi)*, *설풀로버스 솔파타리커스(Sulfolobus solfataricus)(Ss)*, *설풀로버스 토크다이이(Sulfolobus tokodaii)*, *애우로파이럼 페르닉스(Aeuropyrum pernix)(Ap)*, *써모플라스마 애시도필럼(Thermoplasma acidophilum)* 및 *써모플라스마 볼케늄(Thermoplasma volcanium)*이 포함된다.

<54> 보존적 변이체: 본원에서 사용되는, 번역 성분에 대한 내용에 있어서 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 그 O-RS의 보존적 변이체는 비천연 아미노산, 예를 들어 L-(7-히드록시쿠마린-4-일)에틸글리신 비천연 아미노산으로 동족체 O-tRNA를 아미노아실화할 것이다. 이 예에서, O-RS와 보존적 변이체 O-RS는 동일한 아미노산 서열을 갖지 않는다. 보존적 변이체가 상응하는 동족체 O-tRNA 또는 O-RS와 여전히 상보적인(예를 들어, 그것과 기능하는) 한, 상기 보존적 변이체는 서열 내에, 예를 들어 1개의 변이, 2개의 변이, 3개의 변이, 4개의 변이, 또는 5개 이상의 변이를 가질 수 있다.

<55> 일부 실시양태에서, 보존적 변이체 O-RS는 그 자신의 유래 기원인 O-RS에 비해 1개 이상의 보존적 아미노산 치환을 포함한다. 일부 실시양태에서, 보존적 변이체 O-RS는 그 자신의 유래 기원인 O-RS에 비해 1개 이상의 보존적 아미노산 치환을 포함하고, 또한 O-RS의 생물학적 활성을 보유하는데; 예를 들어 보존적 변이체 O-RS는 그 자신의 유래 기원인 모 O-RS의 생물학적 활성의 10% 이상, 또는 대안적으로 20% 이상, 30% 이상, 또는 40% 이상을 보유한다. 일부 바람직한 실시양태에서, 보존적 변이체 O-RS는 그 자신의 유래 기원인 모 O-RS 분자의 생물학적 활성의 50% 이상을 보유한다. 보존적 변이체 O-RS의 보존적 아미노산 치환은 아미노산 결합 포켓(pocket)를 포함하는 O-RS의 임의의 도메인에서 일어날 수 있다.

<56> 선택제 또는 선별제: 본원에서 사용되는 용어 "선택제 또는 선별제"는 존재하는 경우, 집단으로부터 특정 성분들의 선택/선별을 허용하는 시약을 지칭한다. 예를 들어, 선택제 또는 선별제는, 예를 들어 영양분, 항생제, 광파장, 항체, 발현된 폴리뉴클레오티드 등일 수 있으나, 이들로 한정되지 않는다. 선택제는, 예를 들어 농도, 강도 등에 의해 다양해질 수 있다.

<57> ~에 대한 반응으로: 본원에서 사용되는 용어 "~에 대한 반응으로"는 본 발명의 O-tRNA가 셀렉터 코돈을 인식하고, 상기 tRNA에 커플링된 비천연 아미노산이 성장하는 폴리펩티드 사슬 내로 도입되도록 하는 것을 매개하는 과정을 지칭한다.

<58> ~을 코딩한다: 본원에서 사용되는 용어 "~을 코딩한다"는 중합체성 거대분자 또는 서열 스트링(string) 내의 정보를 사용하여, 제1 문자 또는 서열 스트링과 상이한 제2 문자 또는 서열 스트링을 생성시키도록 하는 임의의 과정을 지칭한다. 본원에서 사용되는 그 용어는 광범위하게 사용되며, 다양한 용도를 가질 수 있다. 일부 측면에서, 용어 "~을 코딩한다"는 반-보존적 DNA 복제 과정을 가리키는데, 여기에서 이중-가닥 DNA 분자의 한 가닥은 DNA-의존성 DNA 중합효소에 의해 새로이 합성된 상보적 시스템 가닥을 코딩하기 위한 주형으로서 사용된다.

<59> 또 다른 측면에서, 용어 "~을 코딩한다"는 한 문자 내의 정보를 사용하여, 제1 문자와 상이한 화학적 성질을 가진 제2 문자를 생성시키도록 하는 임의의 과정을 지칭한다. 예를 들어, DNA 문자는 (예를 들어, DNA-의존성 RNA 중합효소를 도입하는 전사 과정에 의해) RNA 문자를 코딩할 수 있다. 또한, RNA 문자는 번역 과정에서와 같이 폴리펩티드를 코딩할 수 있다. 번역 과정을 기술하는 데 사용되는 용어 "~을 코딩한다"는 아미노산을 코딩하는 3 염기 코돈으로 확장된다. 일부 측면에서, RNA 문자는, DNA 문자를 예를 들어 RNA-의존성 DNA 중합효소를 사용하는 역전사 과정에 의해 코딩할 수 있다. 또 다른 측면에서, DNA 문자는 폴리펩티드를 코딩할 수 있는데, 여기

에서 이 경우에서 사용되는 "～을 코딩한다"는 전사 과정 및 번역 과정 양자 모두를 수반한다는 것을 이해할 것이다.

실시 예

<251> 하기 실시예는 청구된 본 발명을 설명하기 위해 제공되는 것이지 본 발명을 한정하기 위해 제공되는 것이 아니다. 당업자는 청구된 본 발명의 범주를 벗어나지 않으면서 변경될 수 있는 다양한 비임계 파라미터를 인식할 것이다. 본원에 기재된 실시예 및 실시양태는 단지 설명하기 위한 목적에 따른 것으로서, 이의 취지에서의 각종 변형 또는 변화가 당업자에게 제시될 것이며, 이는 본원 및 첨부된 특허청구범위의 범주의 취지 및 사상 내에 포함되는 것으로 한다.

실시예 1

일반 방법

<254> 모든 화학물질은 상업적 출처로부터 수득되었고, 추가 정제없이 사용하였다. ^1H NMR 스펙트럼을 테트라메틸실란 대비 보고된 화학 이동과 함께, 브루커(Bruker) AMX-400 다크 분광기(브루커 바이오스핀 게엠바하(Bruker BioSpin GmbH), 독일 라이슈테텐/칼스루헤 소재)에 기록하였다. 단백질 질량 스펙트럼은 질량 분광법들 위해 스 크립스 센터(Scripps Center for Mass Spectrometry)(미국 캘리포니아주 라졸라 소재)에서 습득되었다.

플라스미드 및 세포주

<256> 플라스미드 pBK-lib5는, 잔기 Tyr32, Leu65, Phe108, Gln109, Asp158 및 Leu162가 랜덤화된 *M. 잔나스키이* 티로신 tRNA 합성효소(*MjTyrRS*) 돌연변이체의 라이브러리를 코딩하고; 또한, His70를 Gly로 돌연변이화하였고, Ala67을 Ala 또는 Gly로 돌연변이화하였다. 플라스미드 pREP(2)/YC는 *MjtRNA_{CUA}^{Tyr}*, 잔기 112에 TAG 코돈을 갖는 클로람페니콜 아세틸트랜스퍼라제(CAT) 유전자, T7 프로모터의 조절 하의 GFP 유전자, 잔기 1 및 107에 TAG 코돈을 갖는 돌연변이체 T7 RNA 중합효소, 및 Tet^r 마커를 코딩한다. 플라스미드 pLWJ17B3은 lpp 프로모터 및 *rrnC* 종결자의 조절 하의 *MjtRNA_{CUA}^{Tyr}*, ara 프로모터의 조절 하의 바나제 유전자(잔기 2, 44 및 65에 3개의 앰버 코돈을 가짐), 및 Amp^r 마커를 코딩한다. 플라스미드 pBAD/JYAMB-4TAG-Myo는 아라비노즈 프로모터 및 *rrnB* 종결자를 갖는 돌연변이체 향유고래 미오글로빈 유전자, lpp 프로모터 및 *rrnC* 종결자를 갖는 *MjtRNA_{CUA}^{Tyr}*, 및 테트라사이클린 내성 마커를 코딩한다. *E. 콜라이* 균주 진호그® Fis, F⁻ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) 80*lacZ* Δ M15 Δ*lacX74* *recA1* *endA1* *araD139* Δ(*ara-leu*)7697 *galU* *galK* λ *rpsL*(Str^R) *nupG*, *fis*::Tn7(DHFR)를, 비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신 함유 단백질을 발현하는 연구를 위해 사용하였다.

실시예 2

비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신의 합성

<259> 비천연 쿠마린 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신(7-히드록시쿠마린-에틸글리신으로도 기재됨; 도 1, 구조 1 참조)을, 먼저 N-α-Cbz-L-글루탐산 α-벤질 에스테르를 측쇄 β-케토 에스테르로 전환한 후, 이를 메탄 술폰산 중 레소르시놀과 반응시켜(폰 페치만(von Pechmann) 반응; Brun et al., *Angew. Chem., Int. Ed.*, (2004) 43:3432-3436), 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신을 수득함으로써, 합성하였다. 이 단계는 이하 상세히 기재된다.

단계 1 - 에틸 마그네슘 말로네이트의 합성

<261> 1.5 M 염화마그네슘 및 3 M 에틸 칼륨 말로네이트의 교반 수용액에 5 체적의 이소프로판올을 첨가하였다. 여과후, 에틸 마그네슘 말로네이트를 진공 건조시켰다.

단계 2 - (2S)-2-벤질옥시카르보닐아미노-5-옥소-헵坦디오산 1-벤질 에스테르 7-에틸 에스테르(도 1. 구조 3)의 합성

<263> Z-Glu-Obz1(N-α-Cbz-L-글루탐산 α-벤질 에스테르)(5 g, 13.46 mmol)를 50 mL의 무수 THF에 실온에서 용해시켰다. N,N'-카르보닐 디이미다졸(1.1 eq)을 천천히 첨가한 후, 혼합물을 실온에서 2시간 동안 교반하였다. 에틸 마그네슘 말로네이트(0.55 eq)를 첨가하였고, 혼합물을 하룻밤 동안 실온에서 교반하였다. 생성물을 에테르로

추출하고, 10% NaHCO₃, 물 및 식염수로 세정하였다. **3**을 함유하는 잔류물을 실리카 겔 상의 플래쉬 크로마토그래피(아세트산에틸:헥산 50:50)로 정제하고, 회전 증발기에서 농축하여, 백색 고체를 40% 수율로 수득하였다. ¹H-NMR (400 MHz, CDCl₃): δ 1.26 (t, J= 7.2, 3H), 1.9-2.01 (m, 1H), 2.1- 2.28 (m, 1H), 2.5-2.7 (m, 2H), 3.37 (s, 2H), 4.18 (q, J = 7.2, 2H), 4.41 (s, 1H), 5.11 (s, 2H), 5.18 (s, 2H), 5.44 (d, J = 8, 1H), 7.25-7.41 (m, 10H). LC-MS (ESI) C₂₄H₂NO₇ (M+ H⁺)에 대한 계산치: 442.18, 관찰치: 442.0.

<264> 단계 3 - L-(7-히드록시쿠마린-4-일)에틸글리신(도 1. 구조 1)의 합성

<265> 전 단계로부터의 (2S)-2 벤질옥시카르보닐아미노-5-옥소-헵탄디오산 1-벤질 에스테르 7-에틸 에스테르(2.35 g, 5 mmol)를 메탄솔폰산(20 mL) 중의 레소르시놀(3 g, 25 mmol)에 천천히 첨가하고, 실온에서 2시간 동안 교반하였다. 이어서, 5 체적의 에테르를 혼합물에 첨가하고, -30°C로 냉각시켰다. 석출물을 냉 에테르로 세정하고, 물에 용해시키며, 여과하고, 동결건조시켰다. L-(7-히드록시쿠마린-4-일)에틸글리신을 역상 HPLC에 의해 정제하여, 갈색 고체를 50% 수율로 수득하였다(12분 과정에 걸쳐, YMC AA12S052503WT 칼럼, 12 mL/분 유속, 물 중 0% 내지 90% CH₃CN, 0.1% TFA (w/v)로부터). ¹H-NMR (DMSO): δ 2.06 (s, 2H), 2.70-2.90 (m, 2H), 3.82-4.0 (m, 1H), 6.06 (s, 1H), 6.75 (dd, J= 8.8, 2.4, 1H), 6.82 (d, J = 2.4, 1H), 7.62 (d, J = 8.8, 1H), 8.17 (s, 3H). LC-MS (ESI) C₁₃H₁₄NO₅(M+H⁺)에 대한 계산치 264.08 Da, 관찰치 264.0 Da.

실시예 3

7-히드록시쿠마린-에틸글리신에 대해 특이적인 돌연변이체 합성 효소의 유전적 선택

<266> E. 콜라이의 단백질 내 한정된 위치에 7-히드록시쿠마린-에틸글리신을 선택적으로 도입하기 위해, TAG 셀렉터 코돈에 대한 반응으로 7-히드록시쿠마린-에틸글리신을 특유하게 특정화하는 돌연변이체 메타노코커스 잔나스키이 티로실 앰버 역제자 tRNA (*MjtRNA*^{Tyr}_{CUA})/티로실-tRNA 합성효소(*MjTyrRS*) 쌍을 진화시켰다.

<267> 라이브러리 자체

<268> 비천연 아미노산 7-히드록시쿠마린-에틸글리신의 큰 쿠마린 측쇄를 수용하기 위해, *MjTyrRS* 라이브러리 pBK-lib5를 발생시켰고, 여기에서는 야생형 메타노코커스 잔나스키이 티로실-tRNA 합성효소(도 9; 서열 번호 2 및 3) His70을 Gly로 돌연변이화하였고, Ala67은 Ala 또는 Gly로 고정화되어, 활성 부위 크기를 증가시켰다. 결합된 티로신에 매우 근접하게 있는 6개 잔기, Tyr-32, Leu-65, Phe-108, Gln-109, Asp-158 및 Leu-162(Kobayashi et al., *Nat. Struct. Biol.*, (2003) 10:425-432; Zhang et al., *Protein Sci.*, (2005) 14:1340-1349)를 랜덤화하였다. pBK-lib5 라이브러리는 2×10⁹ TyrRS 독립적 클론으로 구성되었고, 표준 PCR 방법을 이용하여 자체되었다.

<269> 양성/음성 선택

<270> 후속하여, 이 라이브러리를 양성 및 음성 선택 모두의 시험에 적용하였다. 양성 선택에서, 세포 생존율은 pBK-lib 및 *MjtRNA*^{Tyr}_{CUA}로 동시 형질전환된 세포를 1 mM 비천연 아미노산(UAA) 및 클로람페니콜의 존재 하에 성장시킬 때, 클로람페니콜 아세틸 트랜스퍼라제(CAT) 유전자 내 허용 부위에 도입된 앰버 코돈의 역제에 의존적이다 (Xie and Schultz, *Methods*(2005) 36:227-238). 이어서, 양성 선택된 클론을, 허용 부위에 도입된 3개의 앰버 돌연변이를 갖는 독성 바나제 단백질을 코딩하는 유전자 및 *MjtRNA*^{Tyr}_{CUA}를 함유하는 세포 내로 형질전환한다. 이 세포를 UAA의 부재 하에 성장시켜, 내인성 아미노산을 이용하는 임의의 클론을 제거한다(음성 선택). 이 양성 및 음성 선택이 하기 상세히 기재된다.

<271> 양성 선택 - pREP(2)/YC 플라스미드 함유 *E. coli* DH10B를 양성 선택을 위한 숙주 균주로 사용하였다. 세포를 pBK-lib5 라이브러리로 형질전환하고, 1시간 동안 SOC에서 회수하며, 류신을 갖는 글리세롤 최소 배지(GMML)로 2회 세정한 후, 각각 각기 50 μg/mL, 60 μg/mL, 15 μg/mL 및 1 mM의 카나마이신, 클로람페니콜, 테트라사이클린 및 7-히드록시쿠마린-에틸글리신이 보충된 GMML-아가 플레이트 상에 도말하였다. 플레이트를 60시간 동안 37°C에서 인큐베이션하고, 생존 세포를 스크래핑하고, 플라스미드 DNA를 추출하고 겔 전기영동에 의해 정제하였다.

<272> 음성 선택 - pBK-lib5 DNA를 이어서 음성 선택 플라스미드 pLWJ17B3 함유 전기적격(electro-competent) 세포로

형질전환하고, 1시간 동안 SOC 중에 회수한 후, 0.2% 아라비노즈, 50 µg/mL 암피실린 및 50 µg/mL 카나마이신 함유의 LB-아가 플레이트 상에 도말하였다. 이어서, 플레이트를 8 내지 12시간 동안 37°C에서 인큐베이션하고, 생존 클론으로부터의 pBK-1ib5 DNA를 상기 기재된 바와 같이 추출하였다.

<275> 이어서, 라이브러리를 양성 선택의 후속 실행, 그에 이어 음성 선택, 및 마지막 양성 선택 실행(70 µg/mL의 클로람페니콜)에 전달하였다. 이 단계에서, 96개 개별 클론을 선택하고, 96-웰 플레이트 내 50 µL의 GMML에 혼탁하고, 2 세트의 GMML 플레이트 상에 반복-스포팅하였다. 한 세트의 GMML-아가 플레이트를, 1 mM 7-히드록시쿠마린-에틸글리신과 함께 60, 80, 100 및 120 µg/mL 농도의 테트라사이클린(15 µg/mL), 카나마이신(50 µg/mL) 및 클로람페니콜으로 보충하였다. 다른 세트의 플레이트도 동일하나, 단 7-히드록시쿠마린-에틸글리신을 함유하지 않았고, 사용된 클로람페니콜 농도는 0, 20, 40 및 60 µg/mL이었다. 37°C에서 60시간 인큐베이션한 후, 한 클론이 1 mM 7-히드록시쿠마린-에틸글리신의 존재 하에서는 100 µg/mL 클로람페니콜에서 생존하나, 7-히드록시쿠마린-에틸글리신의 부재 하에서는 20 µg/mL 클로람페니콜에서 생존하는 것으로 나타났다. 이에 따라, 양성 선택 3회 실행 및 음성 선택 2회 실행 후에, 이 라이브러리는, 7-히드록시쿠마린-에틸글리신의 존재 하에서는 100 µg/mL 클로람페니콜에서 성장하나 7-히드록시쿠마린-에틸글리신의 부재 하에서는 단지 20 µg/mL 클로람페니콜에서 성장하는 단일 클론을 발생시켰다.

<276> CouRS-D8로 칭해지는 이 클론을 시퀀싱하였고, 이는 하기 8개 돌연변이를 함유하는 것으로 나타났다: Tyr32Glu, Leu65His, Ala67Gly, His70Gly, Phe108Tyr, Gln109His, Asp158Gly 및 Leu162Gly(도 9 및 서열 번호 4 및 5 참조). 도 9에서의 서열 번호 4에 제공된 돌연변이체 합성효소 폴리펩티드 서열은 돌연변이체 위치를 굵은 체로 나타낸다. 돌연변이체 합성효소 단백질은 하기 표에 추가 요약되어 있다.

		아미노산 위치 (코돈)								
<i>Mj</i> 티로신 tRNA 합성효소	32	65	67	70	108	109	158	162	서열 번호:	
야생형	Tyr (TAC)	Leu (TTG)	Ala (GCT)	His (CAC)	Phe (TTC)	Gln (CAG)	Asp (GAT)	Leu (TTA)	2 (3)	
CouRS-D8 변이체	Glu (GAG)	His (CAT)	Gly (GGT)	Gly (GGC)	Tyr (TAT)	His (CAT)	Gly (GGG)	Gly (GGT)	4 (5)	

<277> <278> 4개 잔기가 글리신으로 돌연변이화되었고, 이는 주로 7-히드록시쿠마린-에틸글리신 잔기를 수용하기에 충분한 공간을 생성시키게 된다. 야생형 효소 내 결합된 티로신에 수소 결합하는 Tyr32 및 Asp158이 또한 돌연변이화되었고, 이는 티로신 쪽으로의 활성 손실과 일치한다(Kobayashi et al., *Nat. Struct. Biol.*, (2003) 10:425-432; Zhang et al., *Protein Sci.*, (2005) 14:1340-1349).

<279> 도 8은 돌연변이체 *Mj*TyrRS(CouRS-D8) 활성 부위의 예측되는 구조적 모델을 나타낸다. 야생형 *Mj*TyrRS 구조(pdb 코드 1J1U)을 이용하여 모델을 생성시켰다. 티로신 기질을 L-(7-히드록시쿠마린-4-일)에틸글리신으로 치환하였고, 하기 돌연변이가 나타났다: Tyr32Glu, Leu65His, Ala67Gly, His70Gly, Phe108Tyr, Gln109His, Asp158Gly 및 Leu162Gly. 이어서, 구조를 인사이트 II® 소프트웨어(어셀리즈 소프트웨어 인코포레이티드®, 미국 캘리포니아주 샌디에고 소재)를 이용한 에너지 최소화에 의해 최적화하였다.

실시예 4

7-히드록시쿠마린-에틸글리신을 함유하는 돌연변이체 미오글로빈 모델 단백질의 발현 및 특징분석

<280> <281> 7-히드록시쿠마린-에틸글리신의 단백질 내로의 도입의 효율 및 신뢰도를 결정하기 위해, 앰버 셀렉터 코돈을, C-말단 His₆ 텍을 함유하는 향유고래 미오글로빈 내 위치 4(Ser-4) 또는 위치 37(His-37)에서 치환하였다(도 3 참조). 단백질 발현을 최소 배지 또는 1 mM 7-히드록시쿠마린-에틸글리신으로 보충된 배지에서 성장한 *E. coli* 내의 선택된 합성효소 CouRS-D8 및 *Mj*tRNA^{Tyr}_{CUA}의 존재 하에서 수행하였다. 음성 대조군으로서, 7-히드록시쿠마린-에틸글리신의 부재 하에 단백질 발현을 또한 수행하였다. 이 단백질 생성 및 분석이 하기 상세히 기재된다.

돌연변이체 단백질 생성

<284> 플라스미드 pBAD/JYAMB-4TAG-Myo 또는 pBAD/JYAMB-37TAG-Myo를, 진호그®-Fis E. 콜라이 세포 내로 CouRS-D8을 발현하는 pBK 벡터로 독립적으로 동시 형질전환하였다. 세포를 카나마이신(50 µg/mL) 및 테트라사이클린(15 µg/mL)으로 보충된 2YT 배지(5 mL) 중에 증폭시켰다. 스타터 배양액(1 mL)을 사용하여, 적절한 항체 및 1 mM 7-히드록시쿠마린-에틸글리신으로 보충된 100 mL의 액체 GMML 또는 2YT를 접종하였다. 이어서, 세포를 37°C에서 0.5의 OD₆₀₀로 성장시켰고, 0.2% 아라비노즈 첨가에 의해 단백질 발현을 유도하였다. 추가 4 내지 12시간 후에, 세포를 원심분리에 의해 수거하였다. TAG4 및 TAG37 미오글로빈 돌연변이체를 본(native) 조건 하에 Ni-NTA 친화성 크로마토그래피에 의해 정제하였다. 이어서, 샘플을 탈염하였다. 각 배지 내의 TAG4 및 TAG37 돌연변이체 미오글로빈의 수율은 2 mg/L이었다. 이에 비해, 유사 조건 하에서의 야생형 미오글로빈의 수율은 5 mg/L이었다.

단백질 질량 분광 분석

<286> 정제된 TAG4 돌연변이체 미오글로빈 단백질을 ESI 질량 분광 분석에 적용하였다. 돌연변이체 미오글로빈의 ESI-질량 분광 분석 결과, 18512 Da의 계산된 질량 M_{avg}에 매우 일치하는 18511 Da의 평균 질량 M_{avg}가 관찰되었다(도 6 참조). 5%의 회선 역치가 사용될 때, ESI-질량 스펙트럼에서 천연 아미노산 도입에 상응하는 종이 나타나지 않았기 때문에, 7-히드록시쿠마린-에틸글리신에 대한 최소 95% 도입 순도가 단백질 질량 스펙트럼의 신호-대-노이즈로부터 수득되었다. 검출 역치는, 종의 풍부도가 주요 피크의 5% 초과인 경우에 단지 그 종을 검출할 수 있도록 설정된다. 단지 주요 피크만이 관찰되기 때문에, 돌연변이체 단백질은 >95% 순도를 가진다.

SDS-PAGE 해상법에 의한 단백질 분석

<288> SDS-PAGE에 의한 정제된 단백질의 분석은, 전체 길이의 단백질이 단지 7-히드록시쿠마린-에틸글리신의 존재 하에서만 발현되었음을 나타냈다(도 5 참조). 또한, 유전자적으로 코딩된 7-히드록시쿠마린-에틸글리신이 도입된 TAG4 돌연변이체 미오글로빈 단백질이 SDS-PAGE에서의 해상 후에 형광 밴드로서 나타나는데, 여기에서 유전자적으로 코딩된 7-히드록시쿠마린-에틸글리신이 없는 야생형 미오글로빈 단백질은 형광을 나타내지 않았다. 형광 영상화에는 스톰(STORM)® 형광 스캐너(몰레큘러 다이나믹스(Molecular Dynamics)/GE® 헬스케어 라이프 사이언시스(GE® Healthcare Life Sciences)(미국 뉴저지주 피스캐터웨이 소재)를 사용하였다.

실시예 5

유전자적으로 코딩된 7-히드록시쿠마린-에틸글리신의 단백질 접합을 위한 프로브로서의 용도

<290> 7-히드록시쿠마린-에틸글리신의 잠재적 용도들 중 하나를 입증하기 위해, 이 아미노산을 홀로미오글로빈의 요소-의존성 변성의 프로브로 사용하였다. 쿠마린 형광이 용매 극성에 대해 민감하기 때문에, 그것의 형광 강도는 7-히드록시쿠마린-에틸글리신과 근접한 단백질의 국소 언폴딩과 상관되게 된다. 미오글로빈 구조는 짧은 루프 및 턴에 의해 연결된 8개 나선(A 내지 H)으로 구성된다(Delli Castelli et al., *Protein Sci.*, (2002) 11:2273-2276). 나선 A 내의 Ser4 및 나선 C 내의 His37(도 3 참조; 양 잔기 모두는 대체로 용매에 노출되고, 다른 부근의 잔기와 유의적으로 상호작용하지 않음)은 각기 독립적으로 셀렉터 코돈 TAG으로 돌연변이화되어, 특정 오르소고날 O-tRNA/O-RS 쌍의 존재 하에서의 7-히드록시쿠마린-에틸글리신의 생체내 유전적 프로그래밍화 도입을 초래하였다.

<291> 도 4에 나와 있는 바와 같이, 원편광 이색성에 의해 모니터링되는, 야생형, Ser4 → 7-히드록시쿠마린-에틸글리신 및 His37 → 7-히드록시쿠마린-에틸글리신 홀로미오글로빈의 요소-유도 언폴딩 곡선이 실질적으로 동일하였고, 이는 7-히드록시쿠마린-에틸글리신의 나선 A 또는 나선 C로의 도입이 단백질 안정성을 유의적으로 불안하게 하지 않음을 제시한다. 2 M 요소 농도에서, 홀로미오글로빈 내 위치 4에서의 7-히드록시쿠마린-에틸글리신의 형광 강도는 30% 증가하는데(또한 2 M 내지 5 M 요소에서, 이 수준으로 유지됨), 이는 단백질의 이 영역이 무질서화됨을 가리킨다(Zhang et al., *Protein Sci.*, (2005) 14:1340-1349). 이와 대조적으로, 잔기 37에 7-히드록시쿠마린-에틸글리신을 갖는 돌연변이체 미오글로빈은 2 M 요소에서 형광 강도의 적은 변화를 나타내나, 3 M 요소에서는 유사한 형광 증가를 경험한다. 이 결과와 일치하게도, 과거 NMR 연구는, 요소 농도가 2.2 M 초과일 때, 이 영역에서의 짧은 범위 및 중간 범위의 NOE가 소실되는 것으로 나타내어지는 바대로, 나선 A 및 B이 대체로 무질서화되는 반면, 요소 농도가 3.0 M 초과일 때, 이는 나선 C, D 및 F가 후에 언폴딩됨을 나타냈다(Delli Castelli et al., *Protein Sci.*, (2002) 11:2273-2276). 이에 따라, 전체 구조 전반에 대해 평균화한 전반 변화를 보고하는 원편광 이색성과는 대조적으로, 아미노산 7-히드록시쿠마린-에틸글리신이 단백질 변화의 부

위-특이적 프로브인 것으로 보여진다.

<293> 도 7에 나와 있는 바와 같이, 7-히드록시쿠마린-에틸 글리신의 pH에 대한 민감도도 또한 시험하였다. 60 μM 쿠마리닐 아미노산 7-히드록시쿠마린-에틸글리신의 흡수 스펙트럼을 100 mM 인산나트륨 완충액 중에 측정하였다. pH가 증가함에 따라, 360 nM 흡광도가 증가하였다. 이 측정으로부터 7.8의 pKa가 계산되었다.

<294> 7-히드록시쿠마린-에틸글리신의 pH에 대한 민감도(도 7 참조) 및 용매 극성(Zinsli, J. Photochem(1974) 3:55-69)으로 인해, 그것은 시험관내 및 생체내 모두의 많은 생물학적 연구에 유용한 프로브가 될 것이다(Brun et al., Angew. Chem., Int. Ed., (2004) 43:3432-3436). 예를 들어, 아미노산 7-히드록시쿠마린-에틸글리신을 사용하여, 단백질 내 이분자 상호작용 또는 배좌 변화, 또는 막 결합 단백질의 위상기하학을 모니터링할 수 있다. 또한, 7-히드록시쿠마린-에틸글리신이 7.8의 pKa를 가지는데, 이는 쿠마린 링의 치환에 의해 계통적으로 변경될 수 있기 때문에(Sun et al., Bioorg. Med. Chem. Lett., (1998) 8:3107-3110), 그것은 세포기관 pH 및 pH 의존성 세포내 과정의 유용한 프로브가 될 것이다. 또한, 7-히드록시쿠마린은 여기 상태로 있을 때 강한 광산(photo-acid)임과 동시에(Cohen and Huppert, Phys. Chem. A(2001) 105:7157-7164), 단백질 내 양자 전달 과정의 연구를 촉진할 수 있다.

<295> 상기 발명은 명료성 및 이해를 목적으로 다소 상세하게 설명되었으나, 본 발명의 실제 범주에서 벗어나지 않고 형태 및 세부사항의 다양한 변화가 가능하다는 것은, 본 명세서를 읽은 당업자에게 명백해질 것이다. 본 출원에서 인용된 모든 공보, 특히, 특히 출원 및/또는 기타 문헌은, 각 개별 공보, 특히, 특히 출원 및/또는 기타 문헌이 모든 목적을 위해 참조 인용된 것으로 각각 제시된 바와 동일한 정도로, 모든 목적을 위해 본원에 참조 인용된다.

도면의 간단한 설명

<60> 도 1은 L-(7-히드록시쿠마린-4-일)에틸글리신(구조 1)의 화학적 합성의 개략도를 제공한다.

<61> 도 2는 L-(7-히드록시쿠마린-4-일)에틸글리신의 흡수 및 방출 스펙트럼을 제공한다. 양 스펙트럼 모두는 pH 7.4, 100 mM 인산나트륨 완충액에서 기록되었다. 이 pH에서, L-(7-히드록시쿠마린-4-일)에틸글리신은 폐놀 형태 및 폐놀레이트 형태 모두(대략 2:1 비)로 존재한다. 폐놀레이트 형태는 360 nM에서 17000의 소광 계수를 가지고, 0.63의 양자 수율을 가진다.

<62> 도 3은 향유고래 미오글로빈(pdb 코드 105M)의 구조를 제공한다. 잔기 Ser4 및 His37의 R-기가 나와 있다.

<63> 도 4는 야생형 및 돌연변이체 형태의 홀로미오글로빈의 요소-유도 언폴딩의 원편광 이색성 분석의 결과를 제공한다. 분석된 돌연변이체 형태는 Ser4 → 7-히드록시쿠마린-에틸글리신, 또는 His37 → 7-히드록시쿠마린-에틸글리신 돌연변이체 홀로미오글로빈이었다. 언폴딩을 7-히드록시쿠마린-에틸글리신 형광 강도 및 CD 모두를 이용하여 모니터링하였다. 양 실험 모두를, 나타낸 각종 농도의 요소를 이용하여, 100 mM 인산나트륨 완충액(pH 7.4)에서 수행하였다. 5 M 요소의 존재 하에, 양 미오글로빈 돌연변이체 모두는 0 M 요소에 비해 형광 신호의 30% 증가를 나타낸다. "폴딩 분율(fraction folded)"을 5 M 요소에서의 형광 신호 또는 몰 타원율을 완전 언폴딩 상태에 대해 정규화함으로써 계산한다.

<64> 도 5는 TAG4 돌연변이체 미오글로빈 발현의 SDS-PAGE 분석의 사진을 제공한다(흑색 화살표로 표시됨). 중심 패널은 1 mM L-(7-히드록시쿠마린-4-일)에틸글리신의 존재 및 부재 하에 축적된 단백질의 쿠마씨(coomassie)-염색 SDS-PAGE 해리를 보여준다. 우측 패널은 야생형 및 TAG4 돌연변이체 미오글로빈의 형광 이미지와 그에 이어 SDS-PAGE에 의한 해리를 보여준다. 좌측 패널은 염색된 분자 질량 마커를 보여준다.

<65> 도 6은 TAG4 돌연변이체 미오글로빈의 ESI-MS 스펙트럼을 제공한다. 삽입도는 디컨볼루션된 스펙트럼을 제공한다. 예상 질량: 18512 Da; 실측 질량: 18511 Da.

<66> 도 7은 100 mM 인산나트륨 완충액 중 60 μM 쿠마리닐 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신의 흡수 스펙트럼을 제공한다.

<67> 도 8은 돌연변이체 *MjTyrRS*(*CouRS-D8*) 활성 부위의 구조 모델을 제공한다. 야생형 *MjTyrRS* 구조(pdb 코드 1J1U)를 이용하여 모델을 생성시켰다. 티로신 기질을 L-(7-히드록시쿠마린-4-일)에틸글리신으로 치환하였고, 하기 돌연변이가 나타났다: Tyr32Glu, Leu65His, Ala67Gly, His70Gly, Phe108Tyr, Gln109His, Asp158Gly 및 Leu162Gly. 이어서, 구조를 인사이트(Insight) II[®] 소프트웨어(어셀리즈 소프트웨어 인코포레이티드(Accelrys

Software, Inc.)^⑧, 미국 캘리포니아주 샌디에고 소재)를 이용한 에너지 최소화에 의해 최적화하였다.

<68> 도 9는 각종 뉴클레오티드 및 아미노산 서열을 제공한다

발명의 상세한 설명

<70> 본 명세서는 셀렉터 코돈, 예를 들어 앰버 종결 코돈 TAG에 대한 반응으로 형광 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신(도 1, 구조 1 참조)을 *E. coli*의 단백질 내로의 생체내 선택적 도입을 허용하는 오르소고날 tRNA/아미노아실-tRNA 합성효소 쌍을 제공한다. 본 발명은 동족체 오르소고날 tRNA(O-tRNA)를 비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신으로 특이적으로 충전하는 신규 오르소고날 아미노아실-tRNA 합성효소(O-RS) 폴리펩티드를 제공한다.

<71> 7-히드록시쿠마린 부분은 처음에는 그것의 높은 형광 양자 수율, 비교적 높은 스톡스 이동(Stocke's shift)(도 2 참조), 작은 크기, pH에 대한 민감도(도 7 참조), 및 용매 극성으로 인해 연구되었다(Zinsli, J. *Photochem*(1974) 3:55-69).

<72> 일부 측면에서, 본 발명을 입증하기 위해(단, 본 발명에 한정되지 않음), 본원의 개시내용은, 비천연 아미노산 부분이 모델 단백질 내에 도입될 수 있음을 입증한다(실시예 4 및 5 참조). 비천연 아미노산의 도입이 임의의 특정 단백질에 한정되지 않는 것으로 한다. 본 개시내용으로부터, 비천연 형광 아미노산의 특정 관심 단백질 내로의 도입이 매우 다양한 목적들을 위해 유리하다는 것이 명료할 것이다.

<73> 본 개시내용은 셀렉터 코돈에 대한 반응으로 비천연 형광 아미노산(예를 들어, 도 1, 구조 1에 제공된 비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신)을 부위-특이적으로 도입하도록 진정세균 내에서 기능하는 신규 오르소고날 tRNA/아미노아실-tRNA 합성효소 쌍의 진화를 기술한다. 간략히, 본 발명자들은 *E. coli* 숙주 세포에서 억제자 tRNA를 비천연 아미노산으로 선택적으로 충전하는 *메타노코커스 잔나스키이* 티로실-tRNA 합성효소의 신규 돌연변이체를 동정하였다.

<74> 이 진화된 tRNA-합성효소 쌍을 사용하여, 비천연 형광 아미노산을 단백질 내로 부위-특이적으로 도입할 수 있다. 비천연 아미노산의 단백질 내로의 도입은, 비천연 아미노산의 도입을 신호하는 셀렉터 코돈을 함유하도록 관심 단백질을 코딩하는 폴리뉴클레오티드를 공학처리함으로써, 임의의 원하는 위치에서 일어나도록 프로그래밍될 수 있다.

<75> 본원에 개시된 발명은 진정세균, 예를 들어 *E. coli* 세포에서 단백질 내로의 L-(7-히드록시쿠마린-4-일)에틸글리신 비천연 아미노산의 도입 및 유전적 코딩을 위한 오르소고날 쌍을 제공하는데, 여기에서 오르소고날 성분은 숙주 세포의 번역 기구의 내인성 *E. coli* 성분과 교차 반응하지 않으나, 원하는 비천연 아미노산을 인식하고 그것을 셀렉터 코돈(예를 들어, 앰버 넌센스 코돈, TAG)에 대한 반응으로 단백질 내로 도입한다. 본 발명에 의해 제공되는 오르소고날 성분은, 진정세균 숙주 세포에서 오르소고날 쌍으로 기능하는, *메타노코커스 잔나스키이* 티로실-tRNA 합성효소로부터 유래된 오르소고날 아미노아실-tRNA 합성효소 및 돌연변이체 티로실 tRNACuA 앰버 억제자를 포함한다.

<76> 본 발명은 L-(7-히드록시쿠마린-4-일)에틸 글리신 비천연 아미노산을 단백질 내로 도입하기 위해 사용될 수 있는, 추가 오르소고날 tRNA-아미노아실-tRNA 합성효소 쌍, 예를 들어 O-tRNA/O-RS 쌍을 동정하고 생성시키기 위한 조성물 및 방법을 제공한다. 본 발명의 O-tRNA/O-RS 쌍은 폴리뉴클레오티드에 의해 코딩되는 단백질 내로의 7-히드록시쿠마린-에틸글리신의 도입을 매개할 수 있는데, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함한다. O-tRNA의 안티코돈 루프는 mRNA 상의 셀렉터 코돈을 인식하고, 폴리펩티드 내 상기 부위에 비천연 아미노산을 도입한다. 일반적으로, 본 발명의 오르소고날 아미노아실-tRNA 합성효소는 그것의 O-tRNA를 단지 하나의 특정 비천연 아미노산으로만 우선적으로 아미노아실화 (또는 충전)한다.

오르소고날 tRNA/아미노아실-tRNA 합성효소 기술

<78> 본 발명의 신규 조성물 및 방법의 이해는 오르소고날 tRNA 및 오르소고날 아미노아실-tRNA 합성효소 쌍과 관련된 활성의 이해를 필요로 한다. 추가의 반응성 비천연 아미노산을 유전 코드에 부가하기 위해, 숙주 번역 기구에서 효율적으로 기능할 수 있으나, 사용될 번역 시스템에 대한 내인성 합성효소 및 tRNA와 무관하게 기능한다는 의미에서 상기 번역 시스템에 대해 "오르소고날성"인 아미노아실-tRNA 합성효소 및 적절한 tRNA를 포함하는 신규 오르소고날 쌍이 필요하다. 오르소고날 쌍의 원하는 특징은, 임의의 내인성 tRNA에 의해 서로 해독되지 않는 특정한 코돈, 예를 들어 셀렉터 코돈만을 해독하거나 인식하는 tRNA, 및 그의 동족체 tRNA를 단지 하나의 특정 비천연 아미노산으로만 우선적으로 아미노아실화하는 (또는 "충전하는") 아미노아실-tRNA 합성효소를 포함한

다. 또한, O-tRNA는 전형적으로 내인성 합성효소에 의해 아미노아실화되지 않는다(또는 잘 아미노아실화되지 않거나 충전되지 않음). 예를 들어, *E. coli* 숙주 시스템에서, 오르소고날 쌍은, 예를 들어 *E. coli*에서 40종이 있는 임의의 내인성 tRNA 중 어떠한 tRNA와도 교차-반응하지 않는 아미노아실-tRNA 합성효소, 및, 예를 들어 *E. coli*에서 21종이 있는 내인성 합성효소들 중 어떠한 효소에 의해서도 아미노아실화되지 않는 오르소고날 tRNA를 포함할 것이다.

<79>

하나 이상의 비천연 아미노산을 포함하는 단백질의 제조를 위해 적당한 오르소고날 번역 시스템의 일반 원리가 오르소고날 번역 시스템의 일반 생성 방법과 같이 당업계에 일반적으로 공지되어 있다. 예를 들어, 국제 특허 공개 공보 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"); 2004년 7월 7일에 출원된 제WO 2005/019415호; 2004년 7월 7일에 출원된 제WO 2005/007870호; 2004년 7월 7일에 출원된 제WO 2005/007624호; 2005년 10월 27일에 출원된 제WO 2006/110182호(발명의 명칭: "ORTHOGONAL TRANSLATION COMPONENTS FOR THE IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS"); 및 2007년 3월 7일에 출원된 미국 실용신안 출원 일련 번호 제11/715,672호(발명의 명칭: "SYSTEMS FOR THE EXPRESSION OF ORTHOGONAL TRANSLATION COMPONENTS IN EUBACTERIAL HOST CELLS")를 참조한다. 이 출원들은 각기 전체적으로 본원에 참조 인용된다. 비천연 아미노산을 도입하는 오르소고날 번역 시스템 및 이의 생성 방법, 및 용도에 대한 논의에 대해, 문헌들 [Wang and Schultz, "Expanding the Genetic Code", *Chem. Commun. (Comb.)* 1:1-11(2002); Wang and Schultz, "Expanding the Genetic Code", *Angewandte Chemie Int. Ed.*, 44(1):34-66(2005); Xie and Schultz, "An Expanding Genetic Code", *Methods* 36(3):227-238(2005); Xie and Schultz, "Adding Amino Acids to the Genetic Repertoire", *Curr. Opinion in Chemical Biology* 9(6):548-554(2005); Wang et al., "Expanding the Genetic Code", *Annu. Rev. Biophys. Biomol. Struct.*, 35:225-249(2006); 및 Xie and Schultz, "a Chemical toolkit for Proteins - an expanded Genetic Code", *Nat. Rev. Mol. Cell. Biol.*, 7(10):775-782(2006)]를 또한 참조하고, 이들의 내용은 각기 전체적으로 본원에 참조 인용된다.

<80>

오르소고날 번역 시스템

<81>

오르소고날 번역 시스템은 일반적으로 오르소고날 tRNA(O-tRNA), 오르소고날 아미노아실 tRNA 합성효소(O-RS) 및 비천연 아미노산을 포함하는 세포(*E. coli*와 같은 원핵세포일 수 있음)를 포함하는데, 여기에서 O-RS는 O-tRNA를 비천연 아미노산으로 아미노아실화한다. 본 발명의 오르소고날 쌍은 O-tRNA, 예를 들어 억제자 tRNA, 프레임변동 tRNA 등, 및 동족체 O-RS를 포함할 수 있다. 본 발명의 오르소고날 시스템은 전형적으로 숙주 세포와 관련하여 또는 숙주 세포 없이, O-tRNA/O-RS 쌍을 포함할 수 있다. 다성분 시스템에 부가하여, 본 발명은 또한 신규 개별 성분, 예를 들어 신규 오르소고날 아미노아실-tRNA 합성효소 폴리펩티드(예를 들어, 서열 번호 4), 및 그 폴리펩티드를 코딩하는 폴리뉴클레오티드(예를 들어, 서열 번호 5)를 제공한다.

<82>

일반적으로, 오르소고날 쌍이 셀렉터 코돈을 인식하고 이 셀렉터 코돈에 대한 반응으로 아미노산을 적재하는 경우, 오르소고날 쌍은 셀렉터 코돈을 "억제"한다고 칭해진다. 즉, 번역 시스템(예를 들어, 세포)의 내인성 기구에 의해 인식되지 않는 셀렉터 코돈은 통상 번역되지 않고, 이로써 다른 방식으로라면 핵산으로부터 번역될 폴리펩티드의 생성이 차단되게 된다. 오르소고날 쌍 시스템에서, O-RS는 O-tRNA를 특정 비천연 아미노산으로 아미노아실화한다. 충전된 O-tRNA는 셀렉터 코돈을 인식하고, 셀렉터 코돈에 의해 유발되는 번역 블록을 억제한다.

<83>

일부 측면에서, 본 발명의 O-tRNA는 셀렉터 코돈을 인식하고, 본원의 서열목록에 나타낸 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩되는 O-tRNA의 억제 효율에 비해 셀렉터 코돈에 대한 반응으로 동족체 합성효소의 존재 하에, 적어도 예를 들어 약 45%, 50%, 60%, 75%, 80% 또는 90% 이상의 억제 효율을 가진다.

<84>

일부 실시양태에서, O-RS 및 O-tRNA가 함께 있는 경우의 억제 효율은 O-RS가 없는 경우의 O-tRNA의 억제 효율보다, 예를 들어 약 5배, 10배, 15배, 20배 또는 25배 이상 더 크다. 일부 측면에서, O-RS 및 O-tRNA가 함께 있는 경우의 억제 효율은 본원의 서열목록에 오르소고날 합성효소 쌍의 억제 효율의 적어도 예를 들어 약 35%, 40%, 45%, 50%, 60%, 75%, 80% 또는 90% 이상이다.

<85>

숙주 세포는, 예를 들어 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산을 통해, 성장하는 폴리펩티드 사슬 내로 비천연 아미노산을 도입하기 위해 O-tRNA/O-RS 쌍을 사용하는데, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함한다. 일부 바람직한 측면에서, 세포는 하나 이상의 추가 O-tRNA/O-RS 쌍을 포함할 수 있는데, 여기에서 추가 O-tRNA는 추가 O-RS에 의해 상이한 비천연 아미노산이 적재된다. 예를 들어, O-tRNA들 중 하나는 4 염기 코돈을 인식할 수 있고, 나머지는 종결 코돈을 인식할 수 있다. 대

안적으로, 다수의 상이한 종결 코돈 또는 다수의 상이한 4 염기 코돈이 동일한 코딩 핵산에 사용될 수 있다.

<86> 전술된 바와 같이, 일부 실시양태에서, 하나 초과의 비천연 아미노산을 폴리펩티드 내로 도입하는 다수의 O-tRNA/O-RS 쌍이 세포 또는 다른 번역 시스템 내에 존재한다. 예를 들어, 세포는 추가의 상이한 O-tRNA/O-RS 쌍 및 제2 비천연 아미노산을 추가로 포함할 수 있는데, 여기에서 이 추가 O-tRNA는 제2 셀렉터 코돈을 인식하고, 이 추가 O-RS는 O-tRNA를 제2 비천연 아미노산으로 우선적으로 아미노아실화한다. 예를 들어, O-tRNA/O-RS 쌍(여기에서, O-tRNA는, 예를 들어 앰버 셀렉터 코돈을 인식함)을 포함하는 세포는 제2 오르소고날 쌍을 추가로 포함할 수 있는데, 여기에서 제2 O-tRNA는 상이한 셀렉터 코돈, 예를 들어 오펠 코돈, 4 염기 코돈 등을 인식한다. 바람직하게, 상이한 셀렉터 코돈의 인식을 용이하게 할 수 있는 상이한 오르소고날 쌍이 상이한 공급원으로부터 유래된다.

<87> 특정 실시양태에서, 시스템은 오르소고날 tRNA(O-tRNA), 오르소고날 아미노아실-tRNA 합성효소(O-RS), 비천연 아미노산, 및 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산을 포함하는, *E. coli* 세포와 같은 세포를 포함하는데, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함한다. 번역 시스템은 또한 본원에 기재된 바와 같은, O-tRNA/O-RS 쌍 및 비천연 아미노산과 조합되는 무세포 시스템, 예를 들어 임의의 각종 시판되는 "시험관내" 전사/번역 시스템 중 임의의 시스템일 수도 있다.

<88> O-tRNA 및/또는 O-RS는 천연 발생되거나, 예를 들어 다양한 유기체들 중 임의의 유기체로부터 유래된 tRNA의 라이브러리 및/또는 RS의 라이브러리를 발생시키고/시키거나, 다양한 이용가능한 돌연변이유발 기법 중 임의의 기법을 이용하여, 천연 발생 tRNA 및/또는 RS를 돌연변이시킴으로써 유도될 수 있다. 예를 들어, 오르소고날 tRNA/아미노아실-tRNA 합성효소 쌍을 생성하기 위한 한 기법은, 예를 들어 숙주 세포 외의 한 공급원 또는 다수의 공급원으로부터 유래된 (숙주 세포에 대한) 이종성 tRNA/합성효소 쌍을 숙주 세포 내로 이입하는 단계를 수반한다. 이종성 합성효소 후보의 성질은, 예를 들어 임의의 숙주 세포 tRNA를 충전시키지 못하는 성질을 포함하며, 이종성 tRNA 후보의 성질에는, 예를 들어 임의의 숙주 세포 합성효소에 의해 아미노아실화되지 못하는 성질이 포함된다. 또한, 이종성 tRNA는 모든 숙주 세포 합성효소에 대해 오르소고날성이다. 오르소고날 쌍을 생성하기 위한 제2 기법은 O-tRNA 또는 O-RS를 선택 및/또는 선택할 돌연변이체 라이브러리를 생성시키는 단계를 수반한다. 이 기법들은 조합될 수도 있다.

오르소고날 tRNA(O-tRNA)

<90> 본 발명의 오르소고날 tRNA(O-tRNA)는 바람직하게 생체내 또는 시험관내 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함하는 폴리뉴클레오티드에 의해 코딩되는 단백질 내로의 비천연 아미노산의 도입을 매개한다. 특정 실시양태에서, 본 발명의 O-tRNA는 본원의 서열목록에서 O-tRNA 서열에 나타낸 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩되는 O-tRNA에 비해, 셀렉터 코돈에 대한 반응으로 동족체 합성효소의 존재 하에, 적어도 예를 들어 약 45%, 50%, 60%, 75%, 80% 또는 90% 이상의 억제 효율을 가진다.

<91> 억제 효율은 당업계에 공지된 다수의 검정 중 임의의 검정에 의해 측정될 수 있다. 예를 들어, β -갈락토시다제 리포터 검정을 이용할 수 있는데, 예를 들어 (작제물이 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 (또한 발현될 때 상기 동족체 합성효소를 코딩하는 폴리펩티드 또는 폴리뉴클레오티드로서) 도입될 수 있다. 세포 배지 중에서 원하는 밀도, 예를 들어 약 0.5의 OD₆₀₀까지 성장되고, β -갈락토시다제 검정은, 예컨대 베타플루오르™ β -갈락토시다제 분석 키트(노바젠)를 사용하여 수행한다. 억제율(%)은 비교 대조군에 대한 상대적인 샘플의 백분율(%)로서, 예를 들어 셀렉터 코돈보다 오히려 원하는 위치에서 상응하는 센스 코돈을 가진 유도체화된 lacZ 작제물로부터 관찰된 값으로서 계산될 수 있다.

<92> 본 발명의 O-tRNA의 예는 본원의 서열목록에 나와 있고, 예를 들어 도 9 및 서열 번호 1을 참조한다. 본원의 개시내용은 또한 추가 균등 O-tRNA 종의 설계를 위한 지침을 제공한다. O-RS mRNA 또는 O-tRNA 문자와 같은 RNA 문자에서, 티민(T)은 소정의 서열 또는 이의 상보체에 대해 우라실(U)로 치환된다(또는 코딩 DNA의 경우 이와 반대로 우라실(U)이 티민으로 치환됨). 대체로 기능적으로 균등한 문자를 생성시키기 위해 염기에 대한 추가 변형이 존재할 수도 있다.

<93> 본 발명은 또한 본원의 특정 O-tRNA에 상응하는 O-tRNA의 보존적 변이를 포괄한다. 예를 들어, O-tRNA의 보존적 변이에는, 예를 들어 본원의 서열목록에 나타낸 특정한 O-tRNA와 유사하게 기능하며 적절한 자가-상보성에 의해 tRNA L-형상의 구조를 유지하지만, 예를 들어 본원의 서열목록, 또는 도 9에 기재된 것들과 동일한 서열을 갖지

않는, 또한 바람직하게 야생형 tRNA 분자 외의 것인 분자들이 포함된다.

- <94> 0-tRNA를 포함하는 조성물은 오르소고날 아미노아실-tRNA 합성효소(0-RS)를 추가로 포함할 수 있는데, 여기에서 0-RS는 0-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화한다. 특정 실시양태에서, 0-tRNA를 포함하는 조성물은 (예를 들어, 시험관내 또는 생체내) 번역 시스템을 추가로 포함할 수 있다. 관심 폴리펩티드를 코딩하며, 0-tRNA에 의해 인식되는 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 포함하는 핵산, 또는 이들 중 하나 이상의 조합물이 세포 내에 존재할 수도 있다.
- <95> 오르소고날 tRNA(0-tRNA)의 생성 방법도 또한 본 발명의 한 특성이다. 그 방법에 의해 생성된 0-tRNA도 또한 본 발명의 한 특성이다. 본 발명의 특정 실시양태에서, 0-tRNA는 돌연변이체의 라이브러리를 발생시킴으로써 생성될 수 있다. 돌연변이체 tRNA의 라이브러리는 당업계에 공지된 다양한 돌연변이유발 기법을 이용하여 발생될 수 있다. 예를 들어, 돌연변이체 tRNA는, 예를 들어 서열 번호 1의 0-tRNA의 부위-특이적 돌연변이유발, 랜덤 점 돌연변이유발, 상동성 재조합, DNA 셔플링(shuffling) 또는 다른 반복적 돌연변이유발 방법, 키메라 작제, 또는 이들의 임의의 조합에 의해 발생될 수 있다.
- <96> 추가 돌연변이가 tRNA의 원하는 루프 또는 영역, 예를 들어 안티코돈 루프, 수용체 출기(stem), D 아암(arm) 또는 루프, 가변성 루프, TPC 아암 또는 루프, tRNA 분자의 다른 영역 또는 이들의 조합에서, 특정한 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 돌연변이시키는 단계를 포함한다. 이 방법은 추가 서열을 0-tRNA에 부가하는 단계를 추가로 포함할 수 있다. 전형적으로, 0-tRNA는 원하는 RS에 대한 자체의 친화성을 보존하면서, 출발 물질, 예를 들어 복수개의 tRNA 서열에 비해 원하는 유기체에 대한 개선된 오르소고날성을 가진다.
- <97> 본 방법은 임의적으로, 특정한 유기체에 대해 오르소고날성인 것으로 보이는, 0-tRNA, 0-RS 및/또는 이들의 쌍에 대한 잠재적 후보를 결정하기 위해, tRNA 및/또는 아미노아실-tRNA 합성효소의 서열 유사성 (및/또는 추정된 상동성)을 분석하는 단계를 포함한다. 당업계에 공지되어 있으며 본원에 기재되어 있는 컴퓨터 프로그램이 분석에 사용될 수 있는데, 예를 들어 BLAST 및 파일업(pileup) 프로그램이 사용될 수 있다. 한 예에서, *E. coli*에서 사용하기 위한 잠재적인 오르소고날 번역 성분들을 선택하기 위해, 진정세균 유기체에 대한 높은 서열 유사성을 나타내지 않는 합성효소 및/또는 tRNA를 선택한다.
- <98> 전형적으로, 0-tRNA는 제1 종의 세포로 이루어진 집단을, 예를 들어 음성 선택함으로써 수득되는데, 여기에서 세포는 복수개의 잠재적 0-tRNA의 구성원을 포함한다. 음성 선택은 세포에 대해 내인성인 아미노아실-tRNA 합성효소(RS)에 의해 아미노아실화되는 잠재적인 0-tRNA로 이루어진 라이브러리의 구성원을 포함하는 세포를 제거한다. 이는 제1 종의 세포에 대해 오르소고날성인 tRNA로 이루어진 푸울을 제공한다.
- <99> 특정 실시양태에서, 음성 선택에서, 셀렉터 코돈(들)은 음성 선택 마커, 예를 들어 항생제 내성을 부여하는 효소, 예를 들어 β -락타마제, 비필수 위치(예를 들어, 여전히 기능성 바나제(barnase)를 생산하는 위치)에 검출 가능한 생성물을 부여하는 효소, 예컨대 β -갈اكتоз이다제, 클로람페니콜 아세틸 트랜스퍼라제(CAT), 예를 들어 바나제와 같은 독성 산물을 코딩하는 폴리뉴클레오티드 내로 도입된다. 선별/선택은 임의적으로 선택제(예를 들어, 앰피실린과 같은 항생제)의 존재 하에 세포 집단을 성장시킴으로써 수행된다. 한 실시양태에서, 선택제의 농도는 다양화된다.
- <100> 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(예컨대, 종결) 또는 프레임변동 돌연변이에 기초한 선택 시스템이 사용된다. 예를 들어, 특정한 위치(예를 들어, A184)에서 셀렉터 코돈을 가진 폴리뉴클레오티드 변이체, 예를 들어 *bla* 변이체가 작제된다. 세포, 예컨대 세균은 이 폴리뉴클레오티드로 형질전환된다. 내인성 *E. coli* 합성효소에 의해 효율적으로 충전될 수 없는 오르소고날 tRNA의 경우, 항생제 내성, 예컨대 앰피실린 내성은 플라스미드로 형질전환되지 않은 세균에 대한 항생제 내성과 비슷하거나 이보다 더 낮게 된다. tRNA가 오르소고날 tRNA가 아닌 경우, 또는 tRNA를 충전할 수 있는 이종 합성효소가 상기 시스템 내에서 동시-발현되는 경우, 보다 높은 수준의 항생제, 예컨대 앰피실린의 내성이 관찰된다. 플라스미드로 형질전환되지 않은 세포와 거의 동등한 항생제 농도를 가진 LB 아가 플레이트 상에서 성장할 수 없는 세포, 예를 들어 세균이 선택된다.
- <101> 독성 물질(예컨대, 리보뉴클레아제 또는 바나제)의 경우, 복수개의 잠재적 tRNA의 구성원이 내인성 숙주, 예컨대 *에셀리키아 콜라이* 합성효소에 의해 아미노아실화되는 경우(즉, 상기 숙주, 예컨대 *에셀리키아 콜라이* 합성

효소에 대한 오르소고날성이 아닌 경우), 셀렉터 코돈은 억제되고, 생성된 독성 폴리뉴클레오티드 생성물을 세포 사멸을 초래한다. 오르소고날 tRNA 또는 비-기능성 tRNA를 보유하는 세포가 생존한다.

<102> 한 실시양태에서, 원하는 유기체에 대해 오르소고날성인 tRNA로 이루어진 푸울은 셀렉터 코돈은, 이어서 예를 들어 β -락타마제 유전자와 같은 약물 내성 유전자에 의해 코딩되는 양성 선택 마커에 위치되어 있는 양성 선택에 적용된다. 양성 선택은 세포에 대해 오르소고날성인 tRNA의 푸울의 구성원을 코딩하거나 포함하는 폴리뉴클레오티드, 양성 선택 마커를 코딩하는 폴리뉴클레오티드, 및 동족체 RS를 코딩하는 폴리뉴클레오티드를 포함하는 세포에서 수행된다. 특정 실시양태에서, 제2 세포 집단은 음성 선택에 의해 제거되지 않은 세포를 포함한다. 폴리뉴클레오티드는 세포 내에서 발현되고, 세포는 선택제, 예컨대 앰피실린의 존재 하에 성장한다. 이어서, tRNA는 동시 발현된 동족체 합성효소에 의해 아미노아실화되며 이 셀렉터 코돈에 대한 반응으로 아미노산을 삽입하는 그들의 능력에 대해 선택된다. 전형적으로, 이 세포들은 비-기능성 tRNA(들), 또는 관심 합성효소에 의해 효율적으로 인식될 수 없는 tRNA를 보유하는 세포에 비해 억제 효율의 증진을 보인다. 비-기능성 tRNA, 또는 관심 합성효소에 의해 효율적으로 인식되지 않는 tRNA를 보유하는 세포는 항생제에 대한 민감성을 보인다. 그러므로, (i) 내인성 숙주, 예를 들어 에쉐리키아 콜라이 합성효소에 대한 기질이 아니고; (ii) 관심 합성효소에 의해 아미노아실화될 수 있으며; (iii) 번역에 있어서 작용성을 나타내는 tRNA가 상기 두 선택에서 생존한다.

<103> 따라서, 동일한 마커가 그 자신이 선별되는 부문에 따라 양성 마커 또는 음성 마커일 수 있다. 즉, 마커는 그 자신에 대해 선별되는 경우에는 양성 마커이지만, 그 자신에 반하여 선별되는 경우에는 음성 마커이다.

<104> 선택, 예를 들어 양성 선택, 음성 선택, 또는 양성 선택 및 음성 선택 양자 모두의 엄격도는 임의적으로 선택 엄격도를 달리하는 것을 포함한다. 예를 들어, 바나제는 매우 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, 선택제 또는 선별제의 농도(예를 들어, 앰피실린 농도)는 다양화된다. 본 발명의 일부 측면에서, 엄격도는 원하는 활성이 초기 실행 동안 낮을 수 있기 때문에 다양하다. 그러므로, 엄격도가 보다 낮은 선택 기준이 초기 실행에서 적용되고, 엄격도가 보다 높은 기준이 선별의 후기 실행에서 적용된다. 특정 실시양태에서, 음성 선택, 양성 선택, 또는 음성 선택과 양성 선택 양자 모두가 다수 반복될 수 있다. 다수의 상이한 음성 선택 마커, 양성 선택 마커, 또는 음성 선택 마커와 양성 선택 마커 양자 모두가 사용될 수 있다. 특정 실시양태에서, 양성 선택 마커와 음성 선택 마커는 동일할 수 있다.

<105> 다른 유형의 선택/선별이 본 발명에서 오르소고날 번역 성분, 예를 들어 O-tRNA, O-RS, 및 셀렉터 코돈에 대한 반응으로 비천연 아미노산을 적재하는 O-tRNA/O-RS 쌍을 생성시키는 데 사용될 수 있다. 예를 들어, 음성 선택 마커, 양성 선택 마커, 또는 음성 선택 마커와 양성 선택 마커 양자 모두가 적절한 반응물의 존재 하에 형광을 나타내거나 발광 반응을 촉진하는 마커를 포함할 수 있다. 또 다른 실시양태에서, 마커의 생성물은 형광-활성화 세포 분류(FACS) 또는 발광에 의해 검출된다. 임의적으로, 마커는 친화성에 기초한 선별 마커를 포함한다. 또한, 문헌 [Francisco et al., (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*. 90:10444-10448]를 참조한다.

<106> 재조합 오르소고날 tRNA를 생성시키기 위한 추가 방법은, 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 2004년 7월 7일에 출원된 제WO 2005/019415호에서 찾아볼 수 있다. 또한, 문헌들 [Forster et al., (2003) "Programming peptidomimetic synthetases by translating genetic codes designed de novo", *PNAS* 100(11):6353-6357; 및 Feng et al., (2003), "Expanding tRNA recognition of a tRNA synthetase by a single amino acid change", *PNAS* 100(10): 5676-5681]을 참조한다.

오르소고날 아미노아실-tRNA 합성효소(O-RS)

<108> 본 발명의 O-RS는 시험관내 또는 생체내 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, 예시적 O-RS는 서열 번호 4에 나타낸 아미노산 서열 또는 이의 보존적 변이체를 포함한다. 또 다른 예에서, O-RS 또는 이의 일부는 본원의 서열목록 또는 실시예에서의 서열을 포함하는, 아미노산을 코딩하는 폴리뉴클레오티드 서열 또는 이들의 상보적 폴리뉴클레오티드 서열에 의해 코딩된다. 예를 들어, 서열 번호 5의 폴리뉴클레오티드를 참조한다.

- <109> O-tRNA와 함께 사용하기 위한 오르소고날 아미노아실-tRNA 합성효소(O-RS), 예를 들어 O-RS를 동정하는 방법도 본 발명의 한 특성이다. 예를 들어, 본 방법은 제1 종의 세포로 이루어진 접단을 선택, 예를 들어 양성 선택에 적용하는 단계를 포함하는데, 여기에서 세포는 개별적으로 1) 복수개의 아미노아실-tRNA 합성효소(RS)의 구성원(예를 들어, 복수개의 RS에는 돌연변이체 RS, 상기 제1 종 외의 종으로부터 유래된 RS, 또는 돌연변이체 RS 및 제1 종 외의 종으로부터 유래된 RS 양자 모두가 포함될 수 있음); 2) 오르소고날 tRNA(O-tRNA)(예를 들어, 종마다 상이함); 및 3) (예를 들어, 양성) 선별 마커를 코딩하며 하나 이상의 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 포함한다. 세포는 복수개의 RS 구성원을 갖지 않거나 감소된 양의 복수개의 RS 구성원을 가진 세포에 비해 억제 효율의 상승을 보여주는 세포에 대해 선택 또는 선별된다. 억제 효율은 당업계에 공지되어 있으며 본원에 기재된 바와 같은 기법에 의해 측정될 수 있다. 억제 효율이 증진된 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 선별함으로써, O-tRNA와 함께 사용하기 위한 효율적인 (최적화된) 오르소고날 아미노아실-tRNA 합성효소를 제공한다. 이 방법에 의해 동정된 O-RS도 본 발명의 한 특성이다.
- <110> 다수의 검정 중 임의의 검정을 사용하여 아미노아실화를 결정할 수 있다. 이 검정들은 시험관내 또는 생체내 수행될 수 있다. 예를 들어, 시험관내 아미노아실화 검정은, 예를 들어 문헌 [Hoben and Soil(1985), *Methods Enzymol.* 113:55-59]에 기재되어 있다. 아미노아실화는 또한 리포터를 오르소고날 변역 성분들과 함께 사용하고 단백질을 코딩하는, 하나 이상의 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 발현하는 세포 내에서 상기 리포터를 검출함으로써 결정될 수도 있다. 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.
- <111> 동정된 O-RS는, 단지 원하는 비천연 아미노산만이 O-tRNA에 충전되고 공통된 20종의 아미노산들 중 어떠한 아미노산도 O-tRNA에 충전되지 않도록, 합성효소의 기질 특이성을 변경하도록 더 조작될 수 있다. 비천연 아미노산에 대한 기질 특이성을 가진 오르소고날 아미노아실-tRNA 합성효소를 생성시키는 방법은, 예를 들어 합성효소 내의 활성 부위, 상기 합성효소 내의 편집 기작 부위, 합성효소의 상이한 도메인들의 조합에 의한 상이한 부위 등에서 상기 합성효소를 돌연변이시키는 단계, 및 선택 과정을 적용하는 단계를 포함한다. 양성 선택 후, 음성 선택이 이어지는 조합에 기초한 기법이 이용된다. 양성 선택에서, 양성 마커의 비필수 위치(들) 내로 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 충전하는 활성 합성효소를 코딩한다. 음성 선택에서, 음성 마커의 비필수 위치(들) 내로 도입된 셀렉터 코돈의 억제는 비천연 아미노산 특이성을 가진 합성효소를 제거한다. 음성 선택 및 양성 선택 양자 모두에서의 생존 세포는 오르소고날 억제자 tRNA를 단지 비천연 아미노산으로만 아미노아실화(충전)하는 합성효소를 코딩한다. 이어서, 이 합성효소는 추가 돌연변이유발, 예를 들어 DNA 셔플링 또는 다른 반복적 돌연변이유발 방법에 적용될 수 있다.
- <112> 돌연변이체 O-RS의 라이브러리는 당업계에 공지된 다양한 돌연변이유발 기법을 이용하여 생성될 수 있다. 예를 들어, 돌연변이체 RS는 부위-특이적 돌연변이유발, 랜덤 점 돌연변이유발, 상동성 재조합, DNA 셔플링, 또는 다른 반복적 돌연변이유발 방법, 키메라 작제 또는 이의 임의의 조합에 의해 생성될 수 있다. 예를 들어, 돌연변이체 RS의 라이브러리는 2개 이상의 다른, 예컨대 보다 작고 보다 덜 다양한 "하위-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 et al.); 미국 특허 제5,756,316호(Schallenberger et al.); 미국 특허 제5,783,431호(Petersen et al.); 미국 특허 제5,824,485호(Thompson et al.); 및 미국 특허 제5,958,672호(Short et al.) 참조)가 임의적으로 작제되어, 오르소고날 쌍에 대해 선별된다는 것을 주목해야 한다.
- <113> 일단 상기 합성효소가 양성 및 음성 선택/선별 기법에 적용되면, 이 합성효소를 이어서 추가 돌연변이유발할 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 폴리뉴클레오티드 세트를 상기 핵산으로부터 생성시킬 수 있고; 이 개개의 단계를 또는 이 단계들의 조합을, O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는 돌연변이 O-RS가 수득될 때까지 반복할 수 있다. 본 발명의 일부 측면에서, 상기 단

계들은 수회, 예를 들어 2회 이상 수행된다.

<114> 부가적 선택/선별 업격도 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 업격도는 달라질 수 있다. 이것은, 예를 들어 사용되는 선택제/선별제의 양을 변화시키는 것 등을 포함할 수 있다. 양성 및/또는 음성 선택의 추가 실행을 수행할 수도 있다. 선택 또는 선별은 아미노산 투과성의 하나 이상의 변화, 번역 효율의 변화, 번역 신뢰도의 변화 등을 포함할 수도 있다. 전형적으로, 하나 이상의 변화는 단백질을 생성하기 위해 오르소고날 tRNA-tRNA 합성효소를 사용하는 유기체 내에 하나 이상의 유전자에서의 돌연변이에 기초한 것이다.

<115> O-RS의 생성 및 합성효소의 기질 특이성의 변경에 대한 일반적인 추가 상세 설명은 국제 특허 공개 제WO2002/086075호(발명의 명칭: "METHODS AND COMPOSITIONS FOR THE PRODUCTION OF ORTHOGONAL tRNA AMINOACYL-tRNA SYNTHETASE PAIRS") 및 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")에서 찾아볼 수 있다. 그 내용이 전체적으로 참조 인용되는 문헌 [Wang and Schultz, "Expanding the Genetic Code", *Angewandte Chemie Int. Ed.*, 44(1):34-66(2005)]도 또한 참조한다.

공급원 및 숙주 유기체

<117> 임의의 다른 종으로부터의 숙주 번역 시스템에서 사용하기 위해, 본 발명의 오르소고날 번역 성분(O-tRNA 및 O-RS)은 임의의 유기체(또는 유기체의 조합물)로부터 유래될 수 있으나, 단 O-tRNA/O-RS 성분 및 숙주 시스템은 오르소고날 방식으로 작동한다. 오르소고날 쌍으로부터의 O-tRNA 및 O-RS는 동일한 유기체로부터 유래될 필요는 없다. 일부 측면에서, 진정세균 숙주 시스템에서 사용하기 위해, 오르소고날 성분은 고세균 유전자(즉, 원시세균)로부터 유래된다.

<118> 예를 들어, 오르소고날 O-tRNA는 고세균 유기체, 예를 들어 원시세균, 예컨대 메타노코커스 잔나스키아, 메타노박테리움 씨모오토토트로피컴, 할로박테리움, 예컨대 할로페락스 볼케니아이 및 할로박테리움 종 NRC-1, 아카에오글로버스 풀기더스, 파이로코커스 푸리오서스, 파이로코커스 호리코쉬이, 애우로파이럼 페르닉스, 메타노코커스 마리팔루디스, 메타노파이러스 칸들레리, 메타노사르시나 마제이(Mm), 파이로바콜럼 애로필럼, 파이로코커스 애바이씨, 설플로버스 솔파타리커스(Ss), 설플로버스 토코다이아, 씨모플라스마 애시도필럼, 씨모플라스마 볼케늄 등, 또는 진정세균, 예컨대 E. 콜라이, 씨머스 씨모필러스, 바실러스 섭틸리스, 바실러스 스테아로씨모필러스 등으로부터 유래될 수 있는 반면, 오르소고날 O-RS는 유기체 또는 유기체의 조합물, 예를 들어 원시세균, 예컨대 메타노코커스 잔나스키아, 메타노박테리움 씨모오토토트로피컴, 할로박테리움, 예컨대 할로페락스 볼케니아이 및 할로박테리움 종 NRC-1, 아카에오글로버스 풀기더스, 파이로코커스 푸리오서스, 파이로코커스 호리코쉬이, 애우로파이럼 페르닉스, 메타노코커스 마리팔루디스, 메타노파이러스 칸들레리, 메타노사르시나 마제이, 파이로바콜럼 애로필럼, 파이로코커스 애바이씨, 설플로버스 솔파타리커스, 설플로버스 토코다이아, 씨모플라스마 애시도필럼, 씨모플라스마 볼케늄 등, 또는 진정세균, 예컨대 에쉐리키아 콜라이, 씨머스 씨모필러스, 바실러스 섭틸리스, 바실러스 스테아로씨모필러스 등으로부터 유래될 수 있다. 한 실시양태에서, 진핵세포 공급원, 예를 들어 식물, 해조류, 원생동물, 진균, 효모, 동물(예를 들어, 포유동물, 곤충, 절지동물 등) 등도 O-tRNA 및 O-RS의 공급원으로서 사용될 수 있다.

<119> O-tRNA/O-RS 쌍의 개개의 성분들은 동일한 유기체 또는 상이한 유기체로부터 유래될 수 있다. 한 실시양태에서, O-tRNA/O-RS 쌍은 동일한 유기체로부터 유래된다. 대안적으로, O-tRNA/O-RS 쌍의 O-tRNA 및 O-RS는 상이한 유기체로부터 유래된다.

<120> O-tRNA, O-RS 또는 O-tRNA/O-RS 쌍은 생체내 또는 사험관내 선택 또는 선별되고/되거나, 세포, 예를 들어 진정세균 세포 내에서 비천연 아미노산을 가진 폴리펩티드를 제조하는 데 사용될 수 있다. 사용되는 진정세균 세포는, 예를 들어 에쉐리키아 콜라이, 씨머스 씨모필러스, 바실러스 섭틸리스, 바실러스 스테아로씨모필러스 등이나, 이들로 한정되지 않는다. 본 발명의 번역 성분들을 포함하는 진정세균 세포의 조성물도 본 발명의 한 특성이다.

<121> 또 다른 종에서 사용하기 위한, 한 종에서 O-tRNA 및/또는 O-RS를 선별하는 것에 대해 2004년 4월 16일에 출원된 국제 특허 공개 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")를 참조한다.

<122> 오르소고날 번역 시스템(예를 들어, O-RS, O-tRNA 및 비천연 아미노산을 포함함)이 배양된 숙주 세포를 사용하여 비천연 아미노산을 가진 단백질을 생성시킬 수 있다 하더라도, 본 발명의 오르소고날 번역 시스템이 온전한

생존가능한 숙주 세포를 필요로 한다는 것은 아니다. 예를 들어, 오르소고날 번역 시스템은 세포 추출물의 존재 하에 무세포 시스템을 사용할 수 있다. 사실상, 단백질의 제조를 위한 무세포 시험관내 전사/번역 시스템의 사용은 잘 확립된 기법이다. 본원에 기재된 오르소고날 번역 시스템 성분들을 사용하여 비천연 아미노산을 가진 단백질을 제조하기 위한 이들 시험관내 시스템의 적합화는 본 발명의 범주 내에 속한다.

<123> 셀렉터 코돈

<124> 본 발명의 셀렉터 코돈은 단백질 생합성 기구의 유전 코돈 골격을 확장시킨다. 예를 들어, 셀렉터 코돈은 예를 들어, 특유의 3 염기 코돈, 넌센스 코돈, 예컨대 종결 코돈, 예를 들어 앰버 코돈(UAG) 또는 오펠 코돈(UGA), 비천연 코돈, 4개 이상의 염기로 된 코돈, 희귀 코돈 등을 포함한다. 다수의 셀렉터 코돈, 예를 들어 하나 이상, 두 개 이상, 세 개 초과 등의 셀렉터 코돈이 원하는 유전자 내로 도입될 수 있다. 상이한 셀렉터 코돈을 사용함으로써, 이들 상이한 셀렉터 코돈을 사용하는 다수의 비천연 아미노산, 예를 들어 하나 이상의 비천연 아미노산의 동시적인 부위-특이적 도입을 허용하는 다수의 오르소고날 RNA/합성효소 쌍을 사용할 수 있다.

<125> 한 실시양태에서, 본 방법은 세포에서 생체내 비천연 아미노산을 도입하기 위한 종결 코돈인 셀렉터 코돈의 사용을 수반한다. 예를 들어, 종결 코돈을 인식하며 O-RS에 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 폴리뉴클레오티드 내의 관심 부위에 종결 코돈을 도입 할 수 있다. 예를 들어, 문헌 [Sayers, et al., (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 생체내 조합되는 경우, 비천연 아미노산은 종결 코돈에 대한 반응으로 도입됨으로써 특정된 부위에 비천연 아미노산을 함유하는 폴리펩티드를 제공한다. 본 발명의 한 실시양태에서, 셀렉터 코돈으로서 사용된 종결 코돈은 앰버 코돈 UAG 및/또는 오펠 코돈 UGA이다. UAG 및 UGA 양자 모두가 셀렉터 코돈으로서 사용되는 유전 코드는 가장 풍부한 종결 신호인 오키 넌센스 코돈 UAA를 보존하면서 22종의 아미노산을 코딩할 수 있다.

<126> 생체내 비천연 아미노산의 도입은 숙주 세포에게 유의한 불안을 주지 않으면서 수행될 수 있다. 예를 들어, 에쉐리키아 콜라이와 같은 비진핵세포에서, UAG 코돈에 대한 억제 효율은 O-tRNA, 예를 들어 앰버 억제자 tRNA와 방출 인자 1(RF1)(이것은 UAG 코돈에 결합하고 리보솜으로부터의 성장하는 웨პ티드의 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, 앰버 억제자 tRNA와 진핵생물 방출 인자(예를 들어, eRF)(이것은 종결 코돈에 결합하고 리보솜으로부터의 성장하는 웨პ티드의 방출을 개시함) 사이의 경쟁에 달려 있기 때문에, 억제 효율은 예를 들어, O-tRNA, 예컨대 억제자 tRNA의 발현 수준을 증가시킴으로써 조절할 수 있다. 또한, 추가 화합물, 예를 들어 디티오쓰레이톨(DTT)과 같은 환원제도 존재할 수 있다.

<127> 비천연 아미노산은 또한 희귀 코돈으로 코딩될 수 있다. 예를 들어, 시험관내 단백질 합성 반응에서 아르기닌 농도가 감소된 경우, 희귀 아르기닌 코돈인 AGG는 알라닌으로 아실화된 합성 tRNA에 의한 Ala의 삽입에 효율적인 것으로 입증되었다. 예를 들어, 문헌 [Ma et al.(1993), *Biochemistry*, 32:7939]을 참조한다. 이 경우, 합성 tRNA는 에쉐리키아 콜라이에서 소수의 종으로서 존재하는 천연 발생 tRNA^{Arg}와 경쟁한다. 추가로, 일부 유기체는 모든 3 염기 코돈을 사용하지는 않는다. 마이크로코커스 루테우스(*Micrococcus luteus*)에서 사용되지 않는 코돈인 AGA는 시험관내 전사/번역 추출물 중에서의 아미노산의 삽입을 위해 사용되고 있다. 예를 들어, 문헌 [Kowal and Oliver(1997), *NucI. Acid. Res.*, 25:4685]을 참조한다. 본 발명의 성분들은 생체내 이 희귀 코돈들을 사용하도록 제조될 수 있다.

<128> 셀렉터 코돈은 또한 확장된 코돈, 예를 들어 4개 이상의 염기로 된 코돈, 예컨대 4개, 5개 또는 6개 이상의 염기로 된 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 프레임변동 억제에 기초한 확장된 코돈을 사용하는 것을 포함한다. 4개 이상의 염기로 된 코돈은 예를 들어, 하나 또는 다수의 비천연 아미노산을 동일한 단백질 내로 삽입시킬 수 있다. 다른 실시양태에서, 안티코돈 루프는 예를 들어, 4개 이상의 염기로 된 코돈, 5개 이상의 염기로 된 코돈 또는 6개 이상의 염기로 된 코돈을 해독할 수 있다. 256개의 가능한 4 염기 코돈이 존재하기 때문에, 4개 이상의 염기로 된 코돈을 사용하여 다수의 비천연 아미노산을 동일한 세포 내에서 코딩할 수 있다. 문헌들 [Anderson et al., (2002), "Exploring the Limits of Codon and Anticodon Size", *Chemistry and Biology*, 9:237-244; 및 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.

<129> 예를 들어, 4 염기 코돈은 시험관내 생합성 방법을 이용하여 비천연 아미노산을 단백질 내로 도입하는 데 사용되어 왔다. 예를 들어, 문헌들 [Ma et al., (1993) *Biochemistry*, 32:7939; 및 Hohsaka et al., (1999) *J. Am. Chem. Soc.* 121:34]을 참조한다. CGGG 및 AGGU는 화학적으로 아실화된 2개의 프레임변동 억제자 tRNA를 사용하여 시험관내 2-나프틸알라닌 및 리신의 NBD 유도체를 스트렙타비딘 내로 동시 내로 도입하는 데 사용되었다. 예를 들어, 문헌 [Hohsaka et al., (1999) *J. Am. Chem. Soc.*, 121:12194]을 참조한다. 생체내 연구에서, 무어와 그의 동료들(Moore et al.)은 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 et al., (2000) *J. Mol. Biol.*, 298:195]을 참조한다. 한 실시양태에서, 다른 원치 않는 부위에서 미스센스(missense) 판독 및 프레임변동 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, 제WO 2005/019415호; 제WO 2005/007870호; 및 제WO 2005/07624호를 참조한다. 그 내용이 전체적으로 참조 인용되는 문헌 [Wang and Schultz, "Expanding the Genetic Code", *Angewandte Chemie Int. Ed.*, 44(1):34-66(2005)]도 또한 참조한다. 하기 실시예는 앰버 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 된 코돈도 사용할 수 있다.

<130> 소정의 시스템에 있어서, 셀렉터 코돈은 또한 천연 3 염기 코돈들 중 하나를 포함할 수 있는데, 여기에서 내인성 시스템은 천연 염기 코돈을 사용하지 (또는 거의 사용하지) 않는다. 예를 들어, 이는 천연 3 염기 코돈을 인식하는 tRNA를 갖지 않은 시스템, 및/또는 3 염기 코돈이 희귀 코돈인 시스템을 포함한다.

<131> 셀렉터 코돈은 임의적으로 비천연 염기 쌍을 포함한다. 이러한 비천연 염기 쌍은 기존 유전자 알파벳을 더 확장시킨다. 한 여분의 염기 쌍은 3 염기 코돈의 수를 64개에서 125개로 증가시킨다. 세 번째 염기 쌍의 성질은 안정한 선택적 염기 짹형성, 신뢰도가 높은 중합효소에 의한 DNA 내로의 효율적인 효소적 도입, 및 초기(nascent) 비천연 염기 쌍의 합성 후 효율적인 연속적 프라이머 연장을 포함한다. 본 방법 및 조성물에서 사용될 수 있는 비천연 염기 쌍의 설명은 예를 들어 문헌 [Hirao, et al., (2002) An unnatural base pair for incorporating amino acid analogues into protein, *Nature Biotechnology*. 20:177-182]을 포함한다. 다른 관련 문헌은 하기에 나열되어 있다.

<132> 생체내 사용의 경우, 비천연 뉴클레오시드는 막 투과가능하고, 인산화되어 상응하는 삼인산염을 형성한다. 또한, 증가된 유전 정보는 안정하고, 세포내 효소에 의해 파괴되지 않는다. 베너(Benner) 및 다른 사람들에 의해 이루어진 과거 연구는 정규 와슨-크릭(Watson-Crick) 쌍에서의 수소 결합 패턴과 상이한 수소 결합 패턴을 이용하였는데, 이중 가장 주목할 만한 예는 이소-C:이소-G 쌍이다. 예를 들어, 문헌들 [Switzer et al., (1989) *J. Am. Chem. Soc.*, 111: 8322; 및 Piccirilli et al., (1990) *Nature*, 343: 33; Kool, (2000) *Curr. Opin. Chem. Biol.*, 4: 602]을 참조한다. 일반적으로, 이들 염기는 천연 염기와 어느 정도 미스페어링(mispairing)하며, 효소적으로 복제될 수 없다. 쿠(Kool) 및 그의 동료들은 염기들 사이의 소수성 팩킹(packing) 상호작용이 수소 결합을 대체하여 염기 쌍의 형성을 유도할 수 있다는 점을 입증하였다. 문헌들 [Kool(2000) *Curr. Opin. Chem. Biol.*, 4: 602; 및 Guckian and Kool, (1998) *Angew. Chem. Int. Ed. Engl.*, 36, 2825]을 참조할 수 있다. 모든 상기 요건을 만족시키는 비천연 염기 쌍을 개발하려는 노력으로, 슬츠(Schultz), 로메스버그(Romesberg) 및 이들 동료들은 일련의 비천연 소수성 염기들을 체계적으로 합성하여 연구하였다. PICS:PICS 자가-쌍(self-pair)은 천연 염기 쌍보다 안정한 것으로 밝혀졌고, 에쉐리키아 콜라이 DNA 중합효소 I의 클레나우(Klenow) 단편(KF)에 의해 DNA 내로 효율적으로 도입될 수 있다. 예를 들어, 문헌들 [McMinn et al., (1999) *J. Am. Chem. Soc.*, 121: 11586; 및 Ogawa et al., (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 et al., (2000) *J. Am. Chem. Soc.* 122: 8803]을 참조한다. 그러나, 양쪽 염기 양자 모두 추가 복제를 위한 사슬 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 et al., (2001) *J. Am. Chem. Soc.*, 123: 7439]을 참조한다. 또한, 새로운 메탈로염기 쌍인 Dipic:Py도 개발되었으며, 상기 염기 쌍은 Cu(I)와의 결합 시 안정한 쌍을 형성한다. 문헌 [Meggers et al., (2000) *J. Am. Chem. Soc.*, 122:10714]을 참조

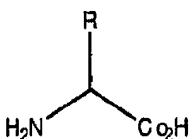
할 수 있다. 확장된 코돈 및 비천연 코돈은 본질적으로 천연 코돈에 대해 오르소고날성이기 때문에, 본 발명의 방법은 이 성질을 이용하여 이들에 대한 오르소고날 tRNA를 생성시킬 수 있다.

<133> 번역 우회(translational bypassing) 시스템을 또한 사용하여, 비천연 아미노산을 원하는 폴리펩티드 내로 도입할 수 있다. 번역 우회 시스템에서, 큰 서열은 유전자 내로 도입되지만, 단백질로 번역되지 않는다. 그 서열은 리보솜이 서열 상에서 호핑하여(hopping), 삽입 부위의 다운스트림(downstream)에서 번역을 재개하도록 유도하는 신호(cue)로서 기능하는 구조를 포함한다.

비천연 아미노산

<135> 본원에서 사용되는 비천연 아미노산은 셀레노시스테인 및/또는 피롤리신, 및 하기 20종의 유전자적으로 코딩된 알파-아미노산을 제외한 임의의 아미노산, 변형된 아미노산 또는 아미노산 유사체를 지칭한다: 알라닌, 아르기닌, 아스파라긴, 아스파르트산, 시스테인, 글루타민, 글루탐산, 글리신, 히스티딘, 이소류신, 류신, 리신, 메티오닌, 페닐알라닌, 프롤린, 세린, 트레오닌, 트립토판, 티로신 및 발린. 알파-아미노산의 일반 구조는 하기 화학식 I로 표시된다:

<136> 화학식 I



<137> <138> 비천연 아미노산은 전형적으로 화학식 I을 갖는 임의의 구조이며, 상기 식 중에서, R기는 20종의 천연 아미노산에서 사용되는 치환기 이외의 임의의 치환기이다. 20종의 천연 아미노산의 구조에 대해서는 예를 들어, 문헌 [Biochemistry by L. Stryer, 3rd ed. 1988, Freeman and Company, New York]을 참조한다. 본 발명의 비천연 아미노산은 상기 20종의 알파-아미노산 이외의 천연 발생 화합물일 수 있다는 것을 주목한다.

<139> 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.

<140> 본원에서 특히 관심을 끄는 것은 비천연 쿠마린 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신(7-히드록시쿠마린-에틸글리신으로도 기재됨; 도 1, 구조 1 참조)이다. 7-히드록시쿠마린-에틸글리신 비천연 아미노산 외에, 다른 비천연 아미노산을, 본 발명에 의해 제공되는 오르소고날 쌍과 함께, 예를 들어 적절한 제2 O-RS/O-tRNA 쌍을 이용하여, 관심 폴리펩티드 내로 동시에 도입될 수 있다. 그러한 많은 추가 비천연 아미노산 및 적당한 오르소고날 쌍이 알려져 있다. 본 개시내용 및 그 안에 인용된 참조문헌을 참조한다. 예를 들어, 문헌 [Wang and Schultz, "Expanding the Genetic Code", Angewandte Chemie Int. Ed., 44(1):34-66(2005); Xie and Schultz, "An Expanding Genetic Code", Methods 36(3):227-238(2005); Xie and Schultz, "Adding Amino Acids to the Genetic Repertoire", Curr. Opinion in Chemical Biology 9(6):548~554(2005); 및 Wang et al., "Expanding the Genetic Code", Ahnu. Rev. Biophys. Biomol. Struct., 35:225-249(2006)]를 참조하고; 이들의 내용은 각각 전체적으로 본원에 참조 인용된다.

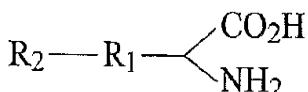
<141> 비천연 쿠마린 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신(도 1, 구조 1 도시)이 본원에 기재된 실시예에서 주된 관심 대상이나, 본 발명이 그 구조에 염격히 한정되지 않도록 한다. 실제로, 형광의 원칙 특징을 보유하는 각종 용이하게 유도되는, 구조적 관련 유사체도 용이하게 생성될 수 있고, 또한 본 발명의 아미노아실-tRNA 합성효소(예를 들어, 서열 번호 4의 O-RS)에 의해 특이적으로 인식된다. 이 관련 형광 유사체도 본 발명의 범주내에 속하는 것으로 한다.

<142> 다른 비천연 아미노산에 있어, 예전대 화학식 I에서의 R은 알킬-, 아릴-, 아실-, 히드라진, 시아노-, 할로-, 히드라지드, 알케닐, 에테르, 보레이트, 보로네이트, 포스포, 포스포노, 포스핀, 엔온, 이민, 알데히드, 히드록실아민, 아민 등 또는 이의 임의의 조합을 임의적으로 포함한다. 기타 관심 비천연 아미노산은 광활성 가교기를 포함하는 아미노산, 스핀 표지된 아미노산, 형광 아미노산, 금속 결합 아미노산, 금속 함유 아미노산, 방사성 아미노산, 신규 작용기를 갖는 아미노산, 다른 분자들과 공유 또는 비공유 상호작용하는 아미노산, 광케이지형(photocaged) 및/또는 광이성체화가능한 아미노산, 비오틴 또는 비오틴 유사체 함유 아미노산, 케토 함유 아미-

노산, 글리코실화된 아미노산, 아미노 측쇄에 결합된 삽카라이드 부분, 폴리에틸렌 글리콜 또는 폴리에테르를 포함하는 아미노산, 중원자로 치환된 아미노산, 화학적으로 분해가능하거나 광분해가능한 아미노산, 천연 아미노산에 비해 신장된 측쇄를 갖는 아미노산(예를 들어, 폴리에테르 또는 장쇄 탄화수소, 예컨대 약 5개 초파의 탄소, 약 10개 초파의 탄소 등), 탄소 결합된 당 함유 아미노산, 아미노 티오산 함유 아미노산, 및 하나 이상의 독성 부분을 포함하는 아미노산을 포함하나, 이에 한정되지 않는다.

<143> 또 다른 측면에서, 본 발명은 하기 화학식 IV에 의해 도시되는 일반 구조를 갖는 비천연 아미노산을 제공한다:

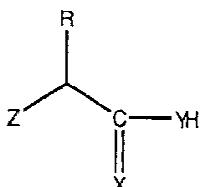
<144> IV



<145> <146> 이 구조를 갖는 비천연 아미노산은 전형적으로, R^1 은 20종의 천연 아미노산(예를 들어, 티로신 또는 페닐알라닌) 중 하나에 사용되는 치환기이고, R^2 는 치환기인 임의의 구조이다. 이에 따라, 이 유형의 비천연 아미노산을 천연 아미노산 유도체로 볼 수 있다.

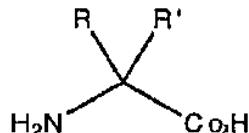
<147> <147> 도 1, 구조 1에 나와 있는 것과 같은 쿠마린 측쇄를 함유하는 비천연 아미노산에 부가하여, 비천연 아미노산은 또한 임의적으로, 예를 들어 화학식 II 및 III의 구조에 의해 도시되는 것과 같은 변형된 골격 구조를 포함할 수 있다:

<148> II



<149>

<150> III



<151> <152> (식 중에서, 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 I을 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 카르복실기에서의 치환을 포함한다. 이러한 유형의 비천연 아미노산은, 예컨대 20종의 공통된 천연 아미노산에 상응하는 측쇄 또는 비천연 측쇄를 갖는 α-히드록시산, α-티오산, α-아미노티오카르복실레이트를 포함하나 이들로 한정되지 않는다. 또한, α-탄소에서의 치환은 임의적으로 L, D 또는 α-α-이치환 아미노산, 예를 들면 D-글루타메이트, D-알라닌, D-메틸-O-티로신, 아미노부티르산 등을 포함한다. 다른 구조적 대체물은 시클릭 아미노산, 예를 들면 프롤린 유사체 뿐만 아니라, 3, 4, 6, 7, 8 및 9원 고리 프롤린 유사체, β 및 γ 아미노산 예를 들어, 치환된 β-알라닌 및 γ-아미노 부티르산을 포함한다.

<153> 일부 측면에서, 본 발명은 L-구조의 비천연 아미노산을 이용한다. 그러나, 본 발명이 L-구조의 비천연 아미노산의 사용으로 한정되지 않는 것으로 한다. 이 비천연 아미노산들의 D-거울상이성질체도 본 발명에서 사용될 수 있는 것으로 구성된다.

<154> 본 발명에 사용되는 비천연 아미노산은 쿠마린 비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신에 염격히 한정되지 않는다. 당업자는, 천연 발생 아미노산의 매우 다양한 비천연 유도체들이 용이하게 유도되어짐을 인식 할 것이다. 비제한적인 예로서, 티로신 유래의 비천연 유사체가 용이하게 생성된다. 티로신 유사체에는 파라-치환 티로신, 오르토-치환 티로신 및 메타-치환 티로신이 포함되는데, 여기에서 치환 티로신은 알키닐 기, 아세틸

기, 벤조일 기, 아미노 기, 히드라진, 히드록시아민, 티올 기, 카르복시 기, 이소프로필 기, 메틸 기, C₆-C₂₀ 칙체 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, 시클 유도체 및 아미드 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 알카닐 기, 히드록시 기, 메톡시 기, 메틸 기, 알릴 기, 알데히드, 니트로, 티올 기 또는 케토 기 등을 포함한다. 비천연 아미노산의 구체적인 예로는 p-에틸티오카르보닐-L-페닐알라닌, p-(3-옥소부타노일)-L-페닐알라닌, 1,5-단실-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 있다. 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), 불소화된 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 등도 있다. 각종 비천연 아미노산의 구조가 본원에 인용된 참조문헌들에 개시되어 있다. 국제 특허 공개 공보 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE"); 및 2005년 10월 27일에 출원된 제WO 2006/110182호(발명의 명칭: "ORTHOGONAL TRANSLATION COMPONENTS FOR THE IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")도 또한 참조한다.

<155> 비천연 아미노산의 화학적 합성

<156> 상기 제공되어 있는 비천연 아미노산의 대다수는 예를 들어, 시그마(Sigma, 미국 소재) 또는 알드리히(Aldrich, 미국 위스콘신주 밀워키 소재)로부터 시판된다. 시판되지 않는 비천연 아미노산은 임의적으로 다양한 문헌에 제공된 바대로 합성되거나, 당업자에게 공지된 표준 방법을 사용하여 합성된다. 유기 합성 기법에 대해, 예를 들어 문헌들 [Organic Chemistry by Fessenden and Fessenden, (1982, Second Edition, Willard Grant Press, Boston Mass.); Advanced Organic Chemistry by March (Third Edition, 1985, Wiley and Sons, New York); 및 Advanced Organic Chemistry by Carey and Sundberg (Third Edition, Parts A and B, 1990, Plenum Press, New York)]을 참조한다. 비천연 아미노산의 합성을 기술하는 추가 문헌은 예를 들어, 제WO2002/085923호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS"), 문헌들 [Matsoukas et al., (1995) *J. Med. Chem.*, 38, 4660-4669; King and Kidd, (1949) "A New Synthesis of Glutamine and of γ-Dipeptides of Glutamic Acid from Phthylated Intermediates", *J. Chem. Soc.*, 3315-3319; Friedman and Chatterji (1959) "Synthesis of Derivatives of Glutamine as Model Substrates for Anti-Tumor Agents", *J. Am. Chem. Soc.*, 81, 3750-3752; Craig et al., (1988) "Absolute Configuration of the Enantiomers of 7-Chloro-4 [[4-(diethylamino)-1-methylbutyl]amino]quinoline (Chloroquine)", *J. Org. Chem.*, 53, 1167-1170; Azoulay et al., (1991) "Glutamine analogues as Potential Antimalarials", *Eur. J. Med. Chem.* 26, 201-5; Koskinen and Rapoport(1989) "Synthesis of 4-Substituted Prolines as Conformationally Constrained Amino Acid Analogues", *J. Org. Chem.* 54, 1859-1866; Christie and Rapoport(1985) "Synthesis of Optically Pure Pipecolates from L-Asparagine. Application to the Total Synthesis of (+)-Apovincamine through Amino Acid Decarbonylation and Iminium Ion Cyclization", *J. Org. Chem.*, 1989: 1859-1866; Barton et al., (1987) "Synthesis of Novel α-Amino-Acids and Derivatives Using Radical Chemistry: Synthesis of L- and D- α-Amino-Adipic Acids, L- α-aminopinic Acid and Appropriate Unsaturated Derivatives", *Tetrahedron Lett.* 43: 4297-4308; 및 Subasinghe et al., (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, 2003년 12월 22일에 출원된 국제 특허 공개 공보 제WO2004/058946호(발명의 명칭: "PROTEIN ARRAYS")를 참조한다.

<157> 비천연 아미노산의 세포내 흡수(cellular uptake)

<158> 세포에 의한 비천연 아미노산 흡수는 예를 들어, 단백질 내로 도입하기 위해 비천연 아미노산을 설계하고 선택하는 경우 전형적으로 고려되는 한 문제이다. 예를 들어, α-아미노산의 높은 충전 밀도는 이 화합물이 세포를 투과할 수 없을 수 있음을 제시한다. 천연 아미노산은 종종 다양한 아미노산 특이성 수준을 나타내는 일군의 단

백질-기재 수송 시스템을 통해 세포 내로 흡수된다. 비천연 아미노산이 세포에 의해 흡수될 경우, 어떤 비천연 아미노산이 세포에 의해 흡수되는지를 평가하는 신속한 선별을 수행할 수 있다. 예를 들어, 2003년 12월 22일에 출원된 국제 특허 공개 제WO 2004/058946호(발명의 명칭: "PROTEIN ARRAYS"); 및 문헌 [Liu and Schultz(1999) "Progress toward the evolution of an organism with an expanded genetic code", *PNAS* 96: 4780-4785]에 기재된 독성 검정을 참조한다. 흡수는 다양한 검정으로 용이하게 분석되지만, 세포내 흡수 경로를 따라 흡수될 수 있는 비천연 아미노산의 설계에 대한 대안은 생체내 아미노산을 생성시키기 위한 생합성 경로를 제공하는 것이다.

비천연 아미노산의 생합성

아미노산 및 다른 화합물들을 생성시키기 위한 많은 생합성 경로들이 세포에 이미 존재한다. 특정한 비천연 아미노산에 대한 생합성 방법이 자연계, 예컨대 세포에 존재하지 않을 수 있으나, 본 발명은 이러한 방법을 제공한다. 예를 들어, 비천연 아미노산에 대한 생합성 경로는 임의적으로, 새로운 효소를 첨가하거나 기존 숙주 세포 경로를 변경함으로써 숙주 세포에서 생성된다. 추가의 새로운 효소로는 임의적으로 천연 발생 효소 또는 인공적으로 진화된 효소가 있다. 예를 들어, (국제 특허 공개 제WO 2002/085923호의 실시예에서 제공된 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. 상기 유전자는 세포에서 발현되는 경우, 원하는 화합물을 합성하는 효소적 경로를 제공한다. 임의적으로 첨가되는 효소 유형의 예는 하기 실시예에서 제공된다. 추가 효소의 서열은 예를 들어, 유전자은행(Genbank)에서 찾아볼 수 있다. 또한, 인공적으로 진화된 효소는 임의적으로 동일한 방식으로 세포 내로 첨가된다. 이 방식으로, 세포내 기구 및 세포내 자원을 비천연 아미노산의 생성을 위해 조작한다.

사실상, 시험관내 또는 생체내 생합성 경로, 기존 경로의 개발 또는 비천연 아미노산의 생성에 사용하기 위한 신규 효소를 제조하는 다양한 방법들 중 임의의 방법이 이용될 수 있다. 효소 및 다른 생합성 경로 성분들을 진화시키기 위한 많은 이용가능한 방법들을 본 발명의 방법에 적용하여, 비천연 아미노산을 생성시킬 수 있다(또는 사실상, 새로운 기질 특이성 또는 다른 원하는 활성을 가진 합성효소를 개발할 수 있음). 예를 들어, 임의적으로 DNA 셔플링을 이용하여 시험관내 또는 생체내 비천연 아미노산의 생성 (또는 신규 합성효소의 생성)을 위한 신규 효소 및/또는 이 효소의 경로를 개발한다. 예를 들어, 문헌들 [Stemmer(1994), "Rapid evolution of a protein in vitro by DNA shuffling", *Nature* 370(4):389-391; 및 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. 관련된 방법들이 관련된(예를 들어, 상동성) 유전자 족(family)을 셔플링하여 원하는 특징을 가진 효소들을 신속히 진화시킨다. 이러한 "족 유전자 셔플링" 방법의 예는 문헌 [Cramer et al., (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 성분이든 합성효소이든 무관하게) 신규 효소는 예를 들어, 문헌 [Ostermeier et al., (1999) "A combinatorial approach to hybrid enzymes independent of DNA homology", *Nature Biotech* 17: 1205]에 기재된 바와 같이 "흔성 효소의 생성을 위한 점진적 말단절단(truncation)"("ITCHY")으로서 공지된 DNA 재조합 방법을 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, 문헌들 [Ostermeier et al., (1999) "Combinatorial Protein Engineering by Incremental Truncation", *Proc. Natl. Acad. Sci. USA*, 96: 3562-67, 및 Ostermeier et al., (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 촉매하는 능력에 대해 선택된 효소 또는 다른 경로 변이체 라이브러리를 제조하기 위해 기하급수적이고 동시적인 돌연변이유발을 이용한다. 이 방법에서, 원하는 서열 내의 잔기들로 이루어진 작은 군을 동시에 랜덤화하여 각각의 변경된 위치에서 기능성 단백질을 생성시키는 아미노산들을 동정한다. 비천연 아미노산(또는 신규 합성효소)의 생성을 위한 신규 효소를 제조하기 위해 본 발명에 적합화될 수 있는 이러한 절차들의 예는 문헌 [Delegrave and Youvan(1993) *Biotechnology Research* 11:1548-1552]에서 찾아볼 수 있다. 또 다른 방법에서, 효소 및/또는 경로 성분 공학 처리를 위한 도핑되거나(doped), 축퇴성인 올리고뉴클레오티드를 사용하는 랜덤 또는 세미-랜덤 돌연변이유발을 사용할 수 있는데, 예를 들어, 문헌들 [Arkin and Youvan(1992) "Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis" *Biotechnology* 10:297-300; 또는 Reidhaar-Olson et al., (1991) "Random mutagenesis of protein sequences using oligonucleotide cassettes" *Methods Enzymol.* 208:564-86]의 일반적인 돌연변이유발 방법을 이용할 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 효소 및/또는 경로 성분들은 (예를 들어, 생체내 비천연 아미노산의 생성을 위한) 하나 이상의 합성효소 또는 생합성 경로 기능을 수행하는 능력에 대해 선별될 수 있다. 예를 들어, 쇼트(Short)의 제WO 00/46344호(발명의 명칭: "NON-STOCHASTIC GENERATION OF GENETIC VACCINES AND ENZYMES")를 참조한다.

<162> 이러한 돌연변이유발 방법에 대한 대안법은 유기체의 전체 게놈을 재조합하는 단계, 및 생성된 자손(progeny)을 특정한 경로 기능(종종 "전체 게놈 셔플링"으로 지칭됨)에 대해 선택하는 단계를 포함한다. 이 방법은 예를 들어, 비천연 아미노산(또는 이의 중간체)을 생성시키는 능력에 대한 유기체(예를 들어, *E. coli* 또는 다른 세포)의 게놈 재조합 및 선택에 의해 본 발명에 적용될 수 있다. 예를 들어, 문헌들 [Patnaik et al., (2002) "Genome shuffling of *lactobacillus* for improved acid tolerance" *Nature Biotechnology*, 20(7): 707-712; 및 Zhang et al., (2002) "Genome shuffling leads to rapid phenotypic improvement in bacteria" *Nature*, February 7, 415(6872): 644-646]에 교시된 방법은, 생체내 비천연 아미노산을 생성시키기 위해 세포내의 기준 및/또는 신규 경로를 진화시키기 위한 경로 설계에 적용될 수 있다.

<163> 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 문헌의 예로는 문헌들 [Nakamura and White(2003) "Metabolic engineering for the microbial production of 1,3-propanediol", *Curr. Opin. Biotechnol.* 14(5):454-9; Berry et al., (2002) "Application of Metabolic engineering to improve both the production and use of Biotech Indigo", *J. Industrial Microbiology and Biotechnology* 28:127-133; Banta et al., (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; Selivonova et al., (2001) "Rapid Evolution of Novel Traits in Microorganisms" *Applied and Environmental Microbiology*, 67:3645] 및 많은 다른 문헌들이 있다.

<164> 사용된 방법과 관계없이, 전형적으로 본 발명의 공학처리된 생합성 경로를 이용하여 생성된 비천연 아미노산은 효율적인 단백질 생합성에 충분한 농도, 예를 들어 세포내 본래 양으로 생성되나, 다른 세포내 아미노산의 농도에 유의한 영향을 미치거나 세포내 자원들을 고갈시킬 정도로 제조되는 것은 아니다. 이 방식으로 생체내 생성되는 전형적인 농도는 약 10 mM 내지 약 0.05 mM이다. 일단 세포가 특정한 경로에 필요한 효소를 제조하도록 공학처리되고 비천연 아미노산이 생성되면, 임의적으로 생체내 선택을 이용하여, 리보솜에 의한 단백질 합성 및 세포 성장 양자 모두를 위한 비천연 아미노산의 생성을 더 최적화한다.

비천연 아미노산의 도입을 위한 오르소고날 성분

<166> 본 발명은 예를 들어 생체내, 셀렉터 코돈, 예를 들어 앰버 종결 코돈, 넌센스 코돈, 4개 이상의 염기로 된 코돈 등에 대한 반응으로 비천연 아미노산 L-(7-히드록시쿠마린-4-일)에틸글리신(도 1, 구조 1 참조)을 성장하는 폴리펩티드 사슬 내로 도입하기 위한 오르소고날 성분을 생성시키기 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 사슬 내로 도입하는 데 사용될 수 있다.

<167> 본 발명의 조성물은 오르소고날 아미노아실-tRNA 합성효소(O-RS)를 포함하는데, 여기에서 O-RS는 O-tRNA를 7-히드록시쿠마린-에틸글리신으로 우선적으로 아미노아실화한다. 특정 실시양태에서, O-RS는 서열 번호 4를 포함한 아미노산 서열 및 이의 보존적 변이체를 포함한다. 본 발명의 특정 실시양태에서, O-RS는 임의의 내인성 tRNA보다 O-tRNA를 특정한 비천연 아미노산으로 우세하게 아미노아실화하는데, 여기에서 O-RS는 O-tRNA를 선호하며, 비천연 아미노산으로 충전된 내인성 tRNA에 대한 동일한 비천연 아미노산으로 충전된 O-tRNA의 비율은 1:1 초파이고, 보다 바람직하게 O-RS는 O-tRNA만을 충전시키거나 거의 O-tRNA만을 충전한다.

<168> O-RS를 포함하는 조성물은 임의적으로 오르소고날 tRNA(O-tRNA)를 추가로 포함할 수 있는데, 여기에서 O-tRNA는 셀렉터 코돈을 인식한다. 전형적으로, 본 발명의 O-tRNA는 서열목록(예를 들어, 서열 번호 1) 및 본원의 실시예에 나타낸 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩되는 O-tRNA의 억제 효율에 비해, 셀렉터 코돈에 대한 반응으로 동족체 합성효소의 존재 하에서 예를 들어, 적어도 약 45%, 50%, 60%, 75%, 80% 또는 90% 이상의 억제 효율을 포함한다. 한 실시양태에서, O-RS 및 O-tRNA가 함께 있는 경우의 억제 효율은 O-RS가 없는 경우의 O-tRNA의 억제 효율보다 예를 들어, 5배, 10배, 15배, 20배 또는 25배 이상 더 크다. 일부 측면에서, O-RS 및 O-tRNA가 함께 있는 경우의 억제 효율은 *메타노코카스* *잔나스키아*로부터 유래된 오르소고날 티로실-tRNA 합성효소의 억제 효율의 45% 이상이다.

- <169> O-tRNA를 포함하는 조성물은 임의적으로 세포(예를 들어, 진정세균 세균, 예컨대 *E. 콜라이* 세포 등, 또는 진핵 세포 세포, 예컨대 효모 세포), 및/또는 번역 시스템을 포함할 수 있다.
- <170> 번역 시스템을 포함하는 세포(예를 들어, 진정세균 세포 또는 효모 세포)도 또한 본 발명에 의해 제공되는데, 여기에서 번역 시스템은 오르소고날-tRNA(O-tRNA); 오르소고날 아미노아실-tRNA 합성효소(O-RS); 및 7-히드록시 쿠마린-에틸글리신 비천연 아미노산을 포함한다. 전형적으로, O-RS는 임의의 내인성 tRNA보다 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화하는데, 여기에서 O-RS는 O-tRNA를 선호하며, 비천연 아미노산으로 충전된 내인성 tRNA에 대한 비천연 아미노산으로 충전된 O-tRNA의 비율은 1:1 초과이고, 보다 바람직하게는 O-RS는 O-tRNA만을 충전시키거나 거의 O-tRNA만을 충전한다. O-tRNA는 제1 셀렉터 코돈을 인식하고, O-RS는 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화한다. 한 실시양태에서, O-tRNA는 서열 번호 1에 나타낸 폴리뉴클레오티드 서열 또는 이의 상보적 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩된다. 한 실시양태에서, O-RS는 서열 번호 4에 나타낸 아미노산 서열 및 이의 보존적 변이체를 포함한다.
- <171> 본 발명의 세포는 임의적으로 추가의 상이한 O-tRNA/O-RS 쌍 및 제2 비천연 아미노산을 추가로 포함할 수 있는데, 여기에서 이 O-tRNA는 제2 셀렉터 코돈을 인식하고, 이 O-RS는 상응하는 O-tRNA를, 제1 비천연 아미노산과 상이한 제2 비천연 아미노산으로 우선적으로 아미노아실화한다. 임의적으로, 본 발명의 세포는 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산을 포함하는데, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함한다.
- <172> 특정 실시양태에서, 본 발명의 세포는, 오르소고날-tRNA(O-tRNA), 오르소고날 아미노아실-tRNA 합성효소(O-RS), 비천연 아미노산, 및 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산을 포함하는 진정세균 세포(예를 들어, *E. 콜라이*)인데, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함한다. 본 발명의 특정 실시양태에서, O-RS는 O-RS가 임의의 내인성 tRNA를 아미노아실화하는 효율보다 큰 효율로 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실화한다.
- <173> 본 발명의 특정 실시양태에서, 본 발명의 O-tRNA는 서열목록(예를 들어, 서열 번호 1) 및 본원의 실시예에 나타낸 폴리뉴클레오티드 서열 또는 이의 상보적 폴리뉴클레오티드 서열을 포함하거나 그 서열에 의해 코딩된다. 본 발명의 특정 실시양태에서, O-RS는 서열목록에 나타낸 아미노산 서열 또는 이의 보존적 변이체를 포함한다. 한 실시양태에서, O-RS 또는 이의 일부는 서열목록 또는 본원의 실시예에 나타낸, 아미노산을 코딩하는 폴리뉴클레오티드 서열 또는 이의 상보적 폴리뉴클레오티드 서열에 의해 코딩된다.
- <174> 본 발명의 O-tRNA 및/또는 O-RS는 임의의 각종 유기체(예를 들어, 진핵세포 및/또는 비진핵세포 유기체)로부터 유래될 수 있다.
- <175> 폴리뉴클레오티드도 또한 본 발명의 한 특성이다. 본 발명의 폴리뉴클레오티드(예를 들어, 서열 번호 5)는 본원의 서열목록에 나타낸, 폴리펩티드를 코딩하는 뉴클레오티드 서열을 포함하는 인공(예를 들어, 천연 발생이 아닌, 인간에 의해 제조된) 폴리뉴클레오티드를 포함하고/하거나 그 폴리뉴클레오티드 서열에 대해 상보적이다. 본 발명의 폴리뉴클레오티드는 또한 핵산의 실질적으로 전체 길이에 대해, 매우 엄격한 조건 하에서, 상기 폴리뉴클레오티드에 혼성화하는 핵산을 포함할 수도 있다. 본 발명의 폴리뉴클레오티드는 또한 천연 발생 tRNA의 서열과 예를 들어, 75% 이상, 80% 이상, 90% 이상, 95% 이상, 98% 이상 또는 그 이상 동일한 폴리뉴클레오티드 또는 이에 상응하는 코딩 핵산(단, 본 발명의 폴리뉴클레오티드는 천연 발생 tRNA 또는 이의 상응하는 코딩 핵산이 아님)을 포함하는데, 여기에서 tRNA는 셀렉터 코돈, 예를 들어 4개 염기 코돈을 인식한다. 상기 것들 중 임의의 것에 대해 예를 들어 80% 이상, 90% 이상, 95% 이상, 98% 이상 또는 그 이상 동일한 인공 폴리뉴클레오티드 및/또는 상기 것들 중 임의의 것의 보존적 변이체를 포함하는 폴리뉴클레오티드도 또한 본 발명의 폴리뉴클레오티드에 포함된다.
- <176> 본 발명의 폴리뉴클레오티드를 포함하는 백터도 또한 본 발명의 한 특성이다. 예를 들어, 본 발명의 백터에는 플라스미드, 코스미드, 파지, 바이러스, 발현 백터 등이 포함될 수 있다. 본 발명의 백터를 포함하는 세포도 또한 본 발명의 한 특성이다.
- <177> O-tRNA/O-RS 쌍의 성분의 생성 방법도 또한 본 발명의 특성이다. 이 방법에 의해 생성된 성분도 또한 본 발명의 특성이다. 예를 들어, 세포에 대해 오르소고날성인 하나 이상의 tRNA(O-tRNA)의 생성 방법은, 돌연변이체 tRNA의 라이브러리를 발생시키는 단계; 돌연변이체 tRNA의 라이브러리의 각 구성원의 안티코돈 루프를 돌연변이화하여 셀렉터 코돈이 인식되도록 함으로써, 잠재적 O-tRNA의 라이브러리를 제공하고, 제1 종의 세포의 제1 집단을 음성 선택에 적용하는 단계를 포함하는데, 여기에서 세포는 잠재적 O-tRNA의 라이브러리의 구성원을 포함한다.

음성 선택은 세포에 대해 오르소고날성인 아미노아실-tRNA 합성효소(RS)에 의해 아미노아실화되는 잠재적 O-tRNA의 라이브러리의 구성원을 포함하는 세포를 제거한다. 이는 제1 종의 세포에 대해 오르소고날성인 tRNA의 푸울을 제공하고, 이로써 하나 이상의 O-tRNA를 제공한다. 본 발명의 방법에 의해 생성된 O-tRNA도 또한 제공된다.

<178> 특정 실시양태에서, 본 방법은 제1 종의 세포의 제2 집단을 양성 선택에 적용하는 것을 추가로 포함하는데, 여기에서 세포는 제1 종의 세포에 대해 오르소고날성인 tRNA의 푸울의 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 보여줌으로써 O-tRNA를 제공하는 세포에 대해 선택 또는 선별한다. 특정 실시양태에서, 세포의 제2 집단은 음성 선택에 의해 제거되지 않은 세포를 포함한다.

<179> O-tRNA를 비천연 아미노산으로 충전하는 오르소고날-아미노아실-tRNA 합성효소의 동정 방법도 또한 제공된다. 예를 들어, 본 방법은 제1 종의 세포의 집단을 선택에 적용하는 것을 포함하는데, 여기에서 세포는 각기 1) 복수개의 아미노아실-tRNA 합성효소(RS)의 구성원, (예를 들어, 복수개의 RS는 돌연변이체 RS, 제1 종 외의 종으로부터 유래된 RS, 또는 유래된 돌연변이체 RS와 제1 종 외의 종으로부터 RS의 양자 모두); 2) 오르소고날-tRNA(O-tRNA)(예를 들어, 하나 이상의 종으로부터의 것); 및 3) 양성 선택 마커를 코딩하고 하나 이상의 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 포함한다.

<180> 세포(예를 들어, 숙주 세포)는, 복수개의 RS의 구성원의 양이 감소되거나 그것이 없는 세포에 비해 억제 효율의 증진을 나타내는 세포에 대해 선택 또는 선별된다. 이 선택/선별된 세포는 O-tRNA를 아미노아실화하는 활성 RS를 포함한다. 본 방법에 의해 동정되는 오르소고날 아미노아실-tRNA 합성효소도 또한 본 발명의 한 특성이다.

<181> 선택된 위치에 비천연 아미노산을 갖는 세포(예를 들어, 진정세균 세포, 예컨대 *E. coli* 세포 등, 또는 효모 세포) 내 단백질의 생성 방법도 또한 본 발명의 한 특성이다. 예를 들어, 본 방법은 적절한 매질 내에 세포를 성장시키는 것을 포함하는데, 여기에서 세포는, 하나 이상의 셀렉터 코돈을 포함하고 단백질을 코딩하는 핵산을 포함하여, 비천연 아미노산을 제공하고 비천연 아미노산을 하나 이상의 셀렉터 코돈을 갖는 핵산의 번역 중에 단백질 내 특정된 위치 내로 도입함으로써, 단백질을 생성시킨다. 세포는, 세포 내에서 기능하고 셀렉터 코돈을 인식하는 오르소고날-tRNA(O-tRNA); 및 O-tRNA를 비천연 아미노산으로 우선적으로 아미노아실하는 오르소고날 아미노아실-tRNA 합성효소(O-RS)를 추가로 포함한다. 이 방법에 의해 생성된 단백질도 또한 본 발명의 한 특성이다.

<182> 본 발명은 또한 7-히드록시쿠마린-에틸글리신을 포함한 단백질을 포함한 조성물을 제공한다. 특정 실시양태에서, 단백질은 공지된 단백질, 예를 들어 치료 단백질, 진단 단백질, 산업용 효소, 또는 이의 일부의 아미노산 서열과 75% 이상 동일한 아미노산 서열을 포함한다. 임의적으로, 조성물은 약학적으로 허용가능한 담체를 포함한다.

핵산 및 폴리펩티드 서열 및 변이체

<184> 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 또는 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 9, 예를 들어 서열 번호 4 및 5 참조). 그러나, 당업자는 본 발명이 본원에, 예를 들어 실시예 및 서열목록에 개시된 서열에 한정되지 않음을 인식할 것이다. 당업자는 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.

<185> O-tRNA를 7-히드록시쿠마린-에틸글리신 비천연 아미노산으로 아미노아실화하는 오르소고날 합성효소 종(O-RS)의 작제 및 분석이 실시예 3 내지 5에 기재되어 있다. 이 실시예들은 비천연 아미노산 7-히드록시쿠마린-에틸글리신을 도입할 수 있는 O-RS 종의 작제 및 분석을 기재한다.

<186> 본 발명은 폴리펩티드(O-RS) 및 폴리뉴클레오티드, 예를 들어 O-tRNA, O-RS 또는 이의 일부를 코딩하는 폴리뉴클레오티드, 아미노아실-tRNA 합성효소 클론을 단리하는 데 사용되는 올리고뉴클레오티드 등을 제공한다. 본 발명의 폴리뉴클레오티드는 본 발명의 관심 단백질 또는 폴리펩티드를 하나 이상의 셀렉터 코돈으로 코딩하는 것들을 포함한다. 또한, 본 발명의 폴리뉴클레오티드는 예를 들어 서열 번호 5에 나타낸 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드, 및 이의 폴리뉴클레오티드 서열에 대해 상보적이거나 그 서열을 코딩하는 폴리뉴클레오티드를 포함한다. 본 발명의 폴리뉴클레오티드는 또한 서열 번호 4를 포함하는 O-RS 아미노산 서열을 코딩하

는 임의의 폴리뉴클레오티드를 포함한다. 유사하게, 고도의 염격한 조건 하에 실질적으로 전체 길이의 핵산 서열에서 상기 나타낸 폴리뉴클레오티드에 혼성화하는 (또한 천연 발생 뉴클레오티드가 아닌) 인공 핵산은 본 발명의 폴리뉴클레오티드이다. 한 실시양태에서, 조성물은 본 발명의 폴리펩티드와 부형제(예를 들어, 완충액, 물, 약학적으로 허용가능한 부형제 등)를 포함한다. 또한, 본 발명은 본 발명의 폴리펩티드와 특이적으로 면역 반응성인 항체 또는 항혈청을 제공한다. 인공 폴리뉴클레오티드는 인간에 의해 제조되고 천연 발생적인 것이 아닌 폴리뉴클레오티드이다.

<187> 본 발명의 폴리뉴클레오티드는 또한 천연 발생 tRNA의 서열과 예를 들어 75% 이상, 80% 이상, 90% 이상, 95% 이상, 98% 이상 동일한(단, 천연 발생 폴리뉴클레오티드는 아닌) 인공 폴리뉴클레오티드를 포함한다. 폴리뉴클레오티드는 또한 천연 발생 tRNA와 예를 들어 75% 이상, 80% 이상, 90% 이상, 95% 이상, 98% 이상, 또는 그 이상 동일한 (그러나, 100% 동일하지는 않은) 인공 폴리뉴클레오티드를 포함한다.

<188> 특정 실시양태에서, 백터(예를 들어, 플라스미드, 코스미드, 파지, 바이러스 등)는 본 발명의 폴리뉴클레오티드를 포함한다. 한 실시양태에서, 백터는 발현 백터이다. 또 다른 실시양태에서, 발현 백터는 본 발명의 폴리뉴클레오티드 중 하나 이상에 작동 가능하게 결합된 프로모터를 포함한다. 또 다른 실시양태에서, 세포는 본 발명의 폴리뉴클레오티드를 포함하는 백터를 포함한다.

<189> 당업자는 또한 개시된 서열의 여러 변이체가 본 발명에 포함된다는 것을 인식할 것이다. 예를 들어, 기능적으로 동일한 서열을 산출하는 개시된 서열의 보존적 변이체가 본 발명에 포함된다. 하나 이상의 개시된 서열에 혼성화하는, 핵산 폴리뉴클레오티드 서열의 변이체는 본 발명에 포함되는 것으로 본다. 예를 들어, 표준 서열 비교 방법에 의해 결정되는 본원에 개시된 서열의 특유의 부분서열도 본 발명에 포함된다.

보존적 변이체

<191> 유전 코드의 축퇴성 때문에, "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 작제물과 매우 유사한 것으로서 용이하게 확인된다. 각 개시된 서열의 이러한 보존적 변이체는 본 발명의 한 특성이다.

<192> 특정 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 전형적으로는 4%, 2% 또는 1% 미만)을 변경시키거나, 부가하거나 결실시키는 개개의 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 전형적으로는 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.

<193> 기능적으로 유사한 아미노산을 제공하는 보존적 치환 표는 당업계에 공지되어 있는데, 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 표에는 유사한 화학적 성질의 천연 아미노산을 함유하는 예시적 군이 기재되어 있는데, 여기에서 군 내의 치환은 "보존적 치환"이다.

보존적 아미노산 치환

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	리신 아르기닌 히스티딘	아스파르테이트 글루타메이트

<196> 핵산 혼성화

<197> 비교 혼성화를 이용하여, 본원의 핵산의 보존적 변이체를 포함한, 본 발명에서 사용될 수 있는 핵산을 동정할 수 있고, 이 비교 혼성화 방법은 본 발명의 핵산을 구별하는 한 바람직한 방법이다. 또한, 높은 염격도, 매우 높은 염격도 및 극도로 매우 높은 염격도의 조건 하에서 서열 번호 5로 나타낸 핵산에 혼성화되는 표적 핵산도 본 발명의 한 특성이다. 그러한 핵산의 예에는 소정의 핵산 서열에 비해 하나 또는 소수의 침묵 또는 보존적 핵산 치환을 가진 핵산들이 포함된다

<198> 시험 핵산은 그것이 완벽하게 매칭되는(matched) 상보적 표적에 대한 프로브에도 50% 이상 혼성화되는 경우, 즉 완벽하게 매칭되는 프로브가 매칭되지 않은 표적 핵산들 중 임의의 핵산과의 혼성화에 대해 관찰된 것보다 약 5 배 내지 10배 이상 더 높은 신호 대 노이즈(noise) 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.

<199> 핵산은 이들이, 전형적으로 용액 중에서, 결합될 때 "혼성화된다". 핵산은 특징이 잘 규명되어 있는 다양한 물리화학적 힘, 예컨대 수소 결합, 용매 배제, 염기 적층 등으로 인해 혼성화된다. 핵산의 혼성화에 대한 심의 지침은 문헌들 [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 *Current Protocols in Molecular Biology*, Ausubel et al., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (2006년 개정 증보판); Hames and Higgins (1995) *Gene Probes 1*, IRL Press at Oxford University Press, Oxford, England, 및 Hames and Higgins (1995) *Gene Probes 2*, IRL Press at Oxford University Press, Oxford, England]은 올리고뉴클레오티드를 비롯하여 DNA 및 RNA의 합성, 표지화, 검출 및 정량화에 대한 상세한 설명을 제공한다.

<200> 서던 또는 노던 블롯에서 필터 상에 100개 초과의 상보적 잔기를 가진 상보적 핵산의 혼성화를 위한 염격한 혼성화 조건의 예는, 1 mg의 헤파린을 함유하는 42°C의 50% 포르말린이며, 여기에서 혼성화는 하룻밤 동안 수행된다. 염격한 세척 조건의 한 예는 15분 동안 65°C에서 0.2×SSC를 사용한 세척이다(SSC 완충제에 대한 설명에 대해서는 문헌 [Sambrook et al., *Molecular Cloning - A Laboratory Manual* (3rd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001] 참조). 종종, 높은 염격도 세척은 배경 프로브 신호를 제거하기 위해 낮은 염격도 세척 후에 행한다. 예시적인 낮은 염격도 세척은 15분 동안 40°C에서 2×SSC를 사용한 세척이다. 일반적으로, 특정한 혼성화 검정에서 관련없는 프로브에 대해 관찰된 신호 대 노이즈 비의 5 배 (또는 더 높은) 신호 대 노이즈 비는 특이적 혼성화의 검출을 표시한다.

<201> 핵산 혼성화 실험(예컨대, 서던 혼성화 및 노던 혼성화)와 관련하여 "엄격한 혼성화 세척 조건"은 서열 의존적이고, 상이한 환경 파라미터 하에 상이하다. 핵산의 혼성화에 대한 심의 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 Hames and Higgins (1995) *Gene Probes 1*, IRL Press at Oxford University Press, Oxford, England, and Hames and Higgins (1995) *Gene Probes 2*, IRL Press at Oxford University Press, Oxford, England]에서 찾아볼 수 있다. 염격한 혼성화 및 세척 조건은 임의의 시험 핵산에 대해 실험적으로 용이하게 결정될 수 있다. 예를 들어, 염격한 혼성화 및 세척 조건의 결정에 있어서, 선택된 기준 세트를 충족시킬 때까지(예를 들어, 혼성화 또는 세척에 있어, 온도 증가, 염 농도 감소, 세제 농도 증가 및/또는 유기 용매, 예컨대 포르말린 농도의 증가에 의해) 혼성화 및 세척 조건을 점차 강화한다. 예를 들어, 높은 염격도의 혼성화 및 세척 조건에서, 매칭되지 않은 표적과 프로브의 혼성화에 대해 관찰되는 신호 대 노이즈 비보다 5배 이상 더 높은 신호 대 노이즈 비로 프로브가 완벽하게 매칭되는 상보적 표적에 결합할 때까지 혼성화 및 세척 조건을 점차 강화한다.

<202> "매우 염격한" 조건은 특정한 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 염격한" 혼성화 및 세척 조건은 한정된 이온 강도 및 pH에서 특정 서열에 대한 T_m 보다 약 5°C 정도 낮게 선택된다.

<203> "매우 높은 염격도의" 혼성화 및 세척 조건은 완벽하게 매칭되는 상보적 표적 핵산과 프로브의 결합에 대한 신

호 대 노이즈 비가 임의의 매칭되지 않은 표적 핵산과의 혼성화에 대해 관찰된 신호 대 노이즈 비보다 10배 이상 더 높아질 때까지 혼성화 및 세척 조건의 업격도가 증가되는 조건이다. 이러한 조건 하에서 완벽하게 매칭되는 상보적 표적 핵산의 1/2 이상의 신호 대 노이즈 비로 프로브에 혼성화하는 표적 핵산은 매우 높은 업격도 조건 하에서 프로브에 결합한다고 말한다.

<204> 유사하게, 훨씬 더 높은 업격도 수준은 관련 혼성화 검정의 혼성화 및/또는 세척 조건을 점차 강화시킴으로써 결정할 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 업격도를 증가시킨다. 이러한 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 업격도 조건 하에서 프로브에 결합한다고 말한다.

<205> 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 카페를 만들 때 일어난다.

특유의 부분서열(unique subsequence)

<207> 일부 측면에서, 본 발명은 본원에 개시된 0-tRNA 및 0-RS의 서열로부터 선택된 핵산 내의 특유의 부분서열을 포함하는 핵산을 제공한다. 특유의 부분서열은 공지된 임의의 0-tRNA 또는 0-RS 핵산 서열에 상응하는 핵산에 비해 특유하다. 정렬은 예를 들어, 디폴트 파라미터로 설정된 BLAST를 사용하여 수행할 수 있다. 임의의 특유의 부분서열이, 예를 들어 본 발명의 핵산을 동정하기 위한 프로브로서 유용하다.

<208> 유사하게, 본 발명은 본원에 개시된 0-RS의 서열로부터 선택된 폴리펩티드 내의 특유의 부분서열을 포함하는 폴리펩티드를 포함한다. 여기에서, 특유의 부분서열은 공지된 폴리펩티드 서열 중 임의의 것에 상응하는 폴리펩티드에 비해 특유하다.

<209> 본 발명은 또한 엄격한 조건 하에서 0-RS의 서열로부터 선택된 폴리펩티드 내의 특유의 부분서열을 코딩하는 특유의 코딩 오리고뉴클레오티드에 혼성화하는 표적 핵산을 제공하는데, 여기에서 특유의 부분서열은 임의의 대조 폴리펩티드(예를 들어 돌연변이에 의해, 본 발명의 합성효소가 유도되도록 하는 기원의 모서열)에 상응하는 폴리펩티드와 비교했을 때 특유하다. 특유의 서열은 전술한 바와 같이 결정된다.

서열 비교, 동일성 및 상동성

<211> 2개 이상의 핵산 또는 폴리펩티드 서열과 관련하여 용어 "동일한" 또는 "동일도(%)"은 최대 상응성에 대해 비교되고 정렬된 경우에, 하기 서열 비교 알고리즘(또는 당업자가 입수할 수 있는 다른 알고리즘) 중 하나 또는 육안 조사를 이용하여 측정할 때, 동일하거나 또는 일정 비율(%) 동일한 아미노산 잔기 또는 뉴클레오티드를 가지는 2개 이상의 서열 또는 부분서열을 의미한다.

<212> 2개의 핵산 또는 폴리펩티드(예컨대, 0-tRNA 또는 0-RS를 코딩하는 DNA, 또는 0-RS의 아미노산 서열)와 관련한, 어구 "실질적으로 동일한"은 최대 상응성에 대해 비교되고 정렬된 경우에, 서열 비교 알고리즘 또는 육안 조사를 이용하여 측정한 때에 적어도 약 60%, 약 80%, 약 90 내지 95%, 약 98% 또는 약 99%, 또는 그 이상의 뉴클레오티드 또는 아미노산 잔기 동일성을 가지는 2개 이상의 서열 또는 부분서열을 의미한다. 이러한 "실질적으로 동일한" 서열은 전형적으로 실제 모 서열에 대한 언급 없이 "상동성"을 갖는 것으로 간주된다. 바람직하게, "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.

<213> 단백질 및/또는 단백질 서열은 공통된 모 단백질 또는 단백질 서열로부터 천연적으로 또는 인공적으로 유래되는 경우 "상동성"을 가진다. 유사하게, 핵산 및/또는 핵산 서열은 공통된 모 핵산 또는 핵산 서열로부터 자연적으로 또는 인공적으로 유래되는 경우, 상동성을 가진다. 예를 들어, 임의의 천연 발생 핵산은 하나 이상의 셀렉터 코돈을 포함하도록 임의의 이용가능한 돌연변이유발 방법에 의해 변형될 수 있다. 발현되는 경우, 이 돌연변이화된 핵산은 하나 이상의 비천연 아미노산을 포함하는 폴리펩티드를 코딩한다. 물론 돌연변이화 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 유사도(%)는 해당 핵산 및 단백질에 따라 다르지만, 25% 정도의 낮은 서열 유사성이 상동성을 확립하는 데 통상 사용된다. 보다 높은 수준의 서열 유사성(예컨대,

30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% 또는 99%, 또는 그 이상)도 또한 상동성을 확립하는 데 사용될 수 있다. 서열 유사도(%)를 결정하기 위한 방법(예컨대, 디폴트 파라미터를 사용하는 BLASTP 및 BLASTN)은 본원에 기재되어 있으며 일반적으로 입수가능하다.

<214> 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 부분서열 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 동일도(%)을 계산한다.

<215> 비교를 위한 서열의 최적 정렬은 예를 들어, 문헌 [Smith and Waterman, *Adv. Appl. Math.* 2:482(1981)]의 국소적 상동성 알고리즘, 문헌 [Needleman and 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(미국 위스콘신주 매디슨 사이언스 드라이브 575 지네티스 컴퓨터 그룹, 위스콘신 지네티스 소프트웨어 패키지의 GAP, BESTFIT, FASTA 및 TFASTA), 또는 육안 조사(일반적으로 문헌 [*Current Protocols in Molecular Biology*, Ausubel et al., ed., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., 2006년 증보 개정판] 참조)에 의해 수행될 수 있다.

<216> 서열 동일도(%) 및 서열 유사도(%)를 결정하는 데 적합한 알고리즘의 한 예는 문헌 [Altschul et al., *J. Mol. Biol.* 215:403-410(1990)]에 기재되어 있는 BLAST 알고리즘이다. BLAST 분석을 수행하기 위한 소프트웨어는 국립생명공학 정보센터(National Center for Biotechnology Information) 웹사이트를 통해 공개적으로 입수가능하다. 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 동정하는 단계를 수반한다. T는 인접 워드 점수 역치로서 언급된다(문헌 [Altschul et al.(1990), *J. Mol. Biol.* 215:403-410] 참조). 이러한 초기 인접 워드 히트(hit)는 이들을 포함하는 더 긴 HSP를 찾는 연구를 개시하기 위한 시드(seed)로 작용한다. 이어서, 워드 히트는 누적 정렬 점수가 증가될 수 있는 한 각 서열을 따라 양 방향으로 확장된다. 뉴클레오티드 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 행렬을 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 민감성 및 속도를 결정한다. (뉴클레오티드 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 행렬을 사용한다(문헌 [Henikoff and Henikoff(1989) *Proc. Natl. Acad. Sci. USA* 89:10915] 참조).

<217> 서열 동일도(%)를 계산하는 것 외에, BLAST 알고리즘은 2개 서열 사이의 유사성의 통계학적 분석도 수행한다(예를 들어, 문헌 [Karlin and Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5877(1993)] 참조). BLAST 알고리즘에 의해 제공되는 유사도의 한 척도는 2개의 뉴클레오티드 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.

<218> 돌연변이유발 및 다른 분자생물학 기법

<219> 본 발명에 사용되는 본 발명의 폴리뉴클레오티드 및 폴리펩ти드는 분자생물학적 기법을 이용하여 조작할 수 있다. 분자생물학적 기법을 설명하는 일반 문헌은 문헌들 [Berger and Kimmel, "Guide to Molecular Cloning Techniques", *Methods in Enzymology* volume 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook et al., *Molecular Cloning - A Laboratory Manual* (3rd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001; 및 *Current Protocols in Molecular Biology*, Ausubel et al., eds., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (2006년 증보 개정판)]을 포함한다. 상기 문헌들에는 예를 들어, 다른 비천연 아미노산을 포함하는 단백질 제조를 위한 셀렉터 코돈, 오르소고날 tRNA, 오르소고날 합성효소 및 이들의 쌍을 포함하는, 유전자의 발생과 관련된 돌연변이유발 방법, 백터의 사용, 프로모터 및 다른 다수의 관련 주제가 기재되어 있다.

<220> 예를 들어, tRNA 분자를 돌연변이화하기 위해, tRNA의 라이브러리를 생성시키기 위해, 합성효소 라이브러리를

제조하기 위해, 비천연 아미노산을 코딩하는 셀렉터 코돈을 관심 단백질 또는 폴리펩티드 내로 삽입하기 위해, 본 발명에 각종 유형의 돌연변이유발 방법을 이용한다. 이에는 부위 지정, 랜덤 점 돌연변이유발, 상동성 재조합, DNA 셔플링 또는 다른 반복적 돌연변이유발 방법, 키메라 작제, 우라실을 함유하는 주형을 사용한 돌연변이 유발, 올리고뉴클레오티드-지정 돌연변이유발, 포스포로티오에이트-변형 DNA 돌연변이유발, 캡 이중 DNA를 사용한 돌연변이유발 등, 또는 이들의 임의의 조합법이 포함되나, 이에 한정되지 않는다. 다른 적절한 방법에는 점 미스매치 복구, 복구-결합 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 작제물을 사용한 돌연변이유발 또한 본 발명에 포함된다. 한 실시양태에서, 돌연변이유발은 천연 발생 분자 또는 변경되거나 돌연변이된 천연 발생 분자의 공지된 정보, 예컨대 서열, 서열 비교, 물성, 결정 구조 등에 의해 지정될 수 있다.

<221> 숙주 세포는 본 발명의 폴리뉴클레오티드, 또는 본 발명의 폴리뉴클레오티드를 포함하는 작제물, 예컨대 클로닝 벡터 또는 발현 벡터일 수 있는 본 발명의 벡터로 유전자적으로 공학처리(예컨대, 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 개시서열, 및 특정한 표적 핵산의 발현 조절에 유용한 프로모터를 포함한다. 벡터는 임의적으로 하나 이상의 독립적인 종결자 서열, 진핵생물 또는 원핵생물 또는 이를 양자 모두에서 카세트의 복제를 허용하는 서열(예컨대, 셔틀 벡터) 및 원핵세포 시스템 및 진핵세포 시스템 양자 모두를 위한 선택 마커를 포함하는 일반 발현 카세트를 포함한다. 벡터는 원핵생물, 진핵생물, 또는 바람직하게는 이를 양자 모두에서 복제 및/또는 삽입에 적합하다. 문헌들 [Giliman and Smith, *Gene* 8:81 (1979); Roberts, et al., *Nature*, 328:731(1987); Schneider et al., *Protein Expr. Purif.* 6435: 10(1995); Berger and Kimmel, "Guide to Molecular Cloning Techniques", *Methods in Enzymology*, volume 152, Academic Press, Inc., San Diego, CA; Sambrook et al., *Molecular Cloning - A Laboratory Manual* (3rd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001; 및 *Current Protocols in Molecular Biology*, Ausubel et al., eds., *Current Protocols*, joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (2006년 증보 개정판)]을 참조한다. 벡터는 예를 들어, 플라스미드, 세균, 바이러스, 네이키드(naked) 폴리뉴클레오티드, 또는 접합된 폴리뉴클레오티드의 형태일 수 있다. 벡터는 전기천 공법(문헌 [From et al., *Proc. Natl. Acad. Sci. USA* 82, 5824(1985)]), 바이러스 벡터에 의한 감염, 소형 비드 또는 입자로 이루어진 매트릭스 내에 또는 표면 상에 핵산을 가진 소립자에 의한 고속 탄도(ballistic) 침투(문헌 [Klein et al., *Nature* 327, 70-73(1987)]) 등을 비롯한 표준 방법에 의해 세포 및/또는 미생물 내로 도입된다.

<222> E. 콜라이 내에서 앰버 종결 코돈(UAG)에 대한 반응으로 비천연 아미노산을 단백질 내로 부위-특이적으로 도입하기 위한, 매우 효율적이고 다용도인 단일 플라스미드 시스템이 개발되었다. 이 새로운 시스템에서, 메타노코커스 잔나스카이 억제자 tRNATyr(CUA)와 티로실-tRNA 합성효소로 이루어진 쌍은 대부분의 E. 콜라이 발현 벡터와 상용가능한 단일 플라스미드 내에 코딩되어 있다. 최적 2차 구조 및 tRNA 프로세싱을 위해, proK 프로모터 및 종결자의 조절 하에 놓인 단일시스트론(monocistronic) tRNA 오페론을 작제하였다. 상기 합성효소에 대한 glnS 프로모터의 돌연변이화는 형태의 도입은 억제 효율 및 신뢰도 모두를 유의하게 증가시켰다. 억제 효율의 증가는 상기 합성효소 상의 특정한 돌연변이(D286R)에 의해서뿐만 아니라 tRNA 유전자의 다수 카페에 의해서도 수득되었다(Kobayashi et al.(2003), "Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion", *Nat. Struct. Biol.*, 10(6):425-432). 최적화된 시스템의 보편성은 또한 단백질 기능 및 구조 연구에 있어서의 특유의 유용성이 이미 입증된, 수가지 상이한 비천연 아미노산의 매우 효율적이고 정확한 도입에 의해서 입증되었다.

<223> 클로닝에 유용한 세균 및 박테리오파지의 카탈로그는 예를 들어 ATCC, 예컨대 ATCC에 의해 발행된 문헌 [*The ATCC Catalogue of Bacteria and Bacteriophage*(1996) Gherna et al.,(eds)]에 의해 제공된다. 시퀀싱, 클로닝 및 분자생물학의 다른 측면에 대한 추가 기본 절차 및 기초적 이론 연구 또한 문헌 [Sambrook et al., *Molecular Cloning - A Laboratory Manual* (3rd Ed.), Vol. 1- 3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001; *Current Protocols in Molecular Biology*, Ausubel et al., eds., *Current Protocols*, joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (2006년 개정 증보판); 및 Watson et al., (1992) *Recombinant DNA Second Edition*, 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(예컨대, 미드랜드 서터파이어 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.

<224> 공학처리된 숙주 세포는 예를 들어, 선별 단계, 프로모터의 활성화 또는 형질전환체의 선택과 같은 공정에 적절하도록 변형된 통상적인 영양 배지 중에서 배양할 수 있다. 상기 세포는 임의적으로 형질전환 유기체 내로 배양될 수 있다. 예를 들어, 세포 단리 및 배양(예컨대, 후속 핵산 단리)에 유용한 다른 참조문헌에는, 문헌들 [Freshney(1994) *Culture of Animal Cells, Manual of Basic Technique*, Third Edition, Wiley-Liss, New York, 및 그 문헌 내에 인용된 참조문헌; Payne et al.(1992) *Plant Cell and Tissue Culture in Liquid Systems*, John Wiley & Sons, Inc. New York, NY; Gamborg and Phillips(eds)(1995) *Plant Cell, Tissue and Organ Culture; Fundamental Methods*, Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York), 및 Atlas and Parks(eds) *The Handbook of Microbiological Media*(1993) CRC Press, Boca Raton, FL.]이 포함된다.

관심 단백질 및 폴리펩티드

<226> 특정된 위치에 비천연 아미노산을 갖는 단백질을 세포내 생성시키는 방법도 또한 본 발명의 한 특성이다. 예를 들어, 본 방법은 적절한 매질 내에, 하나 이상의 셀렉터 코돈을 포함하고 단백질을 코딩하는 세포를 성장시키는 단계; 비천연 아미노산을 제공하는 것을 포함하는데, 여기에서 세포는 세포 내에서 기능하고 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.

<227> 특정 실시양태에서, O-RS는 발현 시스템 내에서 임의의 내인성 tRNA에 비해 동족체 O-tRNA의 아미노아실화를 선호한다. O-tRNA 및 O-RS가 동등한 몰 농도로 존재하는 경우 O-RS에 의해 충전되는 내인성 tRNA와 O-tRNA 사이의 상대적 비는 1:1 초과, 보다 바람직하게는 약 2:1 이상, 보다 더 바람직하게는 5:1, 훨씬 더 바람직하게는 10:1, 보다 더 바람직하게는 20:1, 훨씬 더 바람직하게는 50:1, 보다 더 바람직하게는 75:1, 훨씬 더 바람직하게는 95:1, 98:1, 99:1, 100:1, 500:1, 1,000:1, 5,000:1 또는 그 이상이다.

<228> 본 발명은 또한 비천연 아미노산을 포함한 단백질을 포함하는 조성물을 제공한다. 특정 실시양태에서, 단백질은 치료 단백질, 진단 단백질, 산업용 효소 또는 이들의 일부의 아미노산 서열과 75% 이상 동일한 아미노산 서열을 포함한다.

<229> 본 발명의 조성물 및 본 발명의 방법에 의해 제조되는 조성물은 임의적으로 세포 내에 있다. 본 발명의 O-tRNA/O-RS 쌍 또는 개별 성분은 이어서 비천연 아미노산이 단백질에 도입되도록 하는 숙주 시스템의 번역 기구에 사용될 수 있다. 2004년 4월 16일에 출원된 국제 특허 공개 공보 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE", 및 제WO 2002/085923호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")는 이 방법을 기재하고, 이들은 본원에 참조 인용된다. 예를 들어, O-tRNA/O-RS 쌍이 숙주, 예를 들어 에세리키아 콜라이 세포에 도입될 때, 쌍은 셀렉터 코돈에 대한 반응으로 7-히드록시쿠마린-에틸글리신과 같은 비천연 아미노산이 단백질 내로 생체내 도입되도록 한다. 시스템에 첨가되는 비천연 아미노산은 성장 매질에 외인적으로 첨가될 수 있는 합성 아미노산, 예컨대 폐닐알라닌 또는 티로신의 유도체일 수 있다. 임의적으로, 본 발명의 조성물은 시험관내 번역 시스템 또는 생체내 시스템(들)에 존재할 수 있다.

<230> 본 발명의 세포는 큰 유용한 양으로 비천연 아미노산을 포함하는 단백질을 합성하는 능력을 제공한다. 한 측면에서, 조성물은 예를 들어 비천연 아미노산을 포함하는 단백질 10 µg 이상, 50 µg 이상, 75 µg 이상, 100 µg 이상, 200 µg 이상, 250 µg 이상, 500 µg 이상, 1 mg 이상, 10 mg 이상 또는 그 이상의 양, 또는 생체내 단백질 생성 방법으로 달성될 수 있는 양을 임의로 포함한다(제조합 단백질 생성 및 정제에 대한 상세한 내용이 본원에서 제공되어 있음). 다른 측면에서, 단백질은 예를 들어 10 µg/l 이상의 단백질, 50 µg/l 이상의 단백질, 75 µg/l 이상의 단백질, 100 µg/l 이상의 단백질, 200 µg/l 이상의 단백질, 250 µg/l 이상의 단백질, 500 µg/l 이상의 단백질, 1 mg/l 이상의 단백질, 또는 10 mg/l 이상의 단백질 또는 그 이상의 단백질의 농도로, 예를 들어 세포 용해물, 완충액, 약학 완충액, 또는 다른 액체 혼탁액(예를 들어, 대체로 약 1 nL 내지 약 100 L의 부피로) 중 조성물에 임의로 존재한다. 하나 이상의 비천연 아미노산을 포함한 세포 내 단백질의 다량 생산(예를 들어, 시험관내 번역과 같은 다른 방법을 사용하여 통상적으로 얻을 수 있는 양보다 더 큼)은 본 발명의 한 특성이다.

<231> 비천연 아미노산의 도입은 예를 들어, 크기, 산도, 친핵성, 수소 결합, 소수성, 프로테아제 표적 부위의

접근성, 부분(예를 들어, 단백질 어레이)에 대한 표적화, 표지 또는 반응성 기의 도입 등을 변화시키는, 예를 들어 단백질 구조 및/또는 기능의 변화에 맞추기 위해 수행할 수 있다. 비천연 아미노산을 포함하는 단백질은 강화된 또는 심지어 전적으로 새로운 촉매적 특성 또는 물리적 특성을 가질 수 있다. 예를 들어, 독성, 생분포, 구조적 특성, 분광학적 특성, 화학적 특성 및/또는 광화학적 특성, 촉매능, 반감기(예를 들어, 혈청 반감기), 다른 분자와 예를 들어, 공유 또는 비공유 방식으로 반응하는 능력 등과 같은 성질은, 비천연 아미노산을 단백질 내로 도입함으로써 임의 변형된다. 하나 이상의 비천연 아미노산을 포함하는 단백질을 포함하는 조성물은, 예를 들어 새로운 치료, 진단, 촉매 효소, 산업용 효소, 결합 단백질(예를 들어, 항체), 및 예를 들어 단백질 구조 및 기능 연구에 유용하다. 예를 들어, 문헌 [Dougherty, (2000) *Unnatural Amino Acids as Probes of Protein Structure and Function, Current Opinion in Chemical Biology*, 4:645-652]을 참조한다.

<232> 본 발명의 일부 측면에서, 조성물은 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6개 이상, 7개 이상, 8개 이상, 9개 이상, 또는 10개 이상 또는 그 이상 개수를 포함하나 이에 제한되지 않는 하나 이상의 비천연적으로 코딩된 아미노산을 갖는 하나 이상의 단백질을 포함한다. 비천연적으로 코딩된 아미노산은 동일하거나 상이할 수 있고, 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개, 또는 그 이상 개수의 상이한 비천연적으로 코딩된 아미노산을 포함하는 단백질 내의 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개, 또는 그 이상 개수의 상이한 부위가 있을 수 있으나, 이를 포함하나 이에 제한되지 않는다. 또 다른 측면에서, 조성물은 단백질 내에 존재하는 특별한 아미노산의 하나 이상, 단 전부보다는 적은 아미노산이 비천연적으로 코딩된 아미노산으로 치환된 단백질을 포함한다. 1개 초과의 비천연적으로 코딩된 아미노산을 갖는 소정의 단백질의 경우, 비천연적으로 코딩된 아미노산은 동일하거나 상이 할 수 있다(단백질은 2개 이상의 상이한 유형의 비천연적으로 코딩된 아미노산을 포함하거나, 동일한 비천연적으로 코딩된 아미노산 2개를 포함할 수 있으나, 이에 제한되지 않음). 2개 초과의 비천연적으로 코딩된 아미노산을 갖는 소정의 단백질의 경우, 비천연적으로 코딩된 아미노산은 동일하거나 상이하거나, 혹은 하나 이상의 상이한 비천연적으로 코딩된 아미노산을 갖는 동일 종류의 다중 비천연적으로 코딩된 아미노산의 조합물일 수 있다.

<233> 비천연 아미노산을 포함하는, 본질적으로 임의의 단백질 (또는 이의 일부)(및 예를 들어 하나 이상의 셀렉터 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 하나 이상의 적절한 셀렉터 코돈이 포함되도록 임의의 이용 가능한 돌연변이 방법을 조정함으로써 임의의 공지된 단백질이 하나 이상의 비천연 아미노산을 포함하도록 변형시킬 수 있다. 공지된 단백질에 대한 통상의 서열 저장 기관은 유전자은행, EMBL, DDBJ 및 NCBI를 포함한다. 다른 저장 기관은 인터넷을 선별하여 쉽게 찾을 수 있다.

<234> 전형적으로, 단백질은 임의의 이용 가능한 단백질(예컨대, 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 예는 2004년 4월 16일 출원된 국제특허공개 공보 제 WO2004/094593호(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, 알파-1 항트립신, 안지오스타틴, 항용혈성 인자, 항체(항체에 대한 상세한 설명은 하기에서 제공), 아포지단백질, 아포단백질, 심방성 나트륨이뇨 인자, 심방성 나트륨이뇨 폴리펩티드, 심방성 웨პ티드, C-X-C 케모카인(예컨대, T39765, NAP-2, ENA-78, Gro-a, Gro-b, Gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG), 칼시토닌, CC 케모카인(예컨대, 단핵세포 화학유인물질 단백질-1, 단핵세포 화학유인물질 단백질-2, 단핵세포 화학유인물질 단백질-3, 단핵세포 염증성 단백질-1 알파, 단핵세포 염증성 단백질-1 베타, RANTES, I309, R83915, R91733, HCC1, T58847, D31065, T64262), CD40 리간드, C-카이트 리간드, 콜라겐, 콜로니 자극 인자(CSF), 보체 인자 5a, 보체 억제제, 보체 수용체 1, 사이토카인(예컨대, 상피 호중구 활성화 웨პ티드-78, GRO α /MGSA, GRO β , GRO γ , MIP-1 α , MIP-1 β , MCP-1), 상피 성장 인자(EGF), 에리쓰로포이에틴("EPO"), 익스폴리에이팅(Exfoliating) 독소 A 및 B, 인자 IX, 인자 VII, 인자 VIII, 인자 X, 섬유아세포 성장 인자(FGF), 피브리노겐, 피브로넥틴, G-CSF, GM-CSF, 글루코세레бр로시다제, 성선 자극 호르몬, 성장 인자, 지호그(Hedgehog) 단백질(예컨대, 소닉(Sonic), 인디언(Indian), 디저트(Desert)), 헤모글로빈, 간세포 성장 인자(HGF), 히루딘, 인간 혈청 알부민, 인슐린, 인슐린-유사 성장 인자(IGF), 인터페론(예컨대, IFN- α , IFN- β , IFN- γ), 인터루킨(예컨대, 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 억제자, 루시퍼라제, 뉴르튜린(Neurturin), 호중구 억제 인자(NIF), 온코스타틴 M, 골생성 단백질, 부갑상선 호르몬, PD-ECSF, PDGF, 웨პ티드 호르몬(예컨대, 인간 성장 호르

몬), 플레이오토로핀, 단백질 A, 단백질 G, 발열원성 외독소 A, B 및 C, 렐랙신(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(Somatomedin), 소마토스타틴(Somatostatin), 소마토트로핀(Somatotropin), 스트렙토키나제, 수퍼항원, 즉 포도상구균 장독소(SEA, SEB, SEC1, SEC2, SEC3, SED, SEE), 수퍼옥사이드 디스뮤타제(SOD), 독성 쇼크 증후군 독소(TSST-1), 티모신(Thymosin) 알파 1, 조직 플라스미노겐 활성화제, 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 포함되지만, 이들로 한정되지 않는다.

<235> 본원에 기재된 비천연 아미노산을 생체내 도입하기 위한 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, 예를 들어 진균, 식물, 효모, 곤충, 및 포유동물을 비롯한 동물에서 발견되며, 이는 광범위한 치료 표적을 제공한다. 발현 및 전사 활성화제는 많은 기작에 의해, 예를 들어 수용체에 결합하고, 신호 전달도입 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 것은 인식되어 있을 것이다.

<236> 본 발명의 단백질의 한 부류(예컨대, 하나 이상의 비천연 아미노산을 가지는 단백질)는 생물학적 활성 단백질, 예컨대 사이토카인, 염증성 분자, 성장 인자, 이들의 수용체 및 종양유전자 생성물, 예를 들어 인터루킨(예컨대, 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 전달도입 분자 및 상응하는 종양유전자 생성물, 예컨대 Mos, Ras, Raf 및 Met; 및 전사 활성화제 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.

<237> 본 발명은 하나 이상의 비천연 아미노산을 가진 효소(예컨대, 산업용 효소) 또는 이의 일부를 제공한다. 효소의 예에는 예를 들어, 아미다제(amidase), 아미노산 라세마제(racemase), 아실라제(acylase), 데할로게나제(dehalogenase), 디옥시게나제(dioxygenase), 디아릴프로판 페옥시다제, 에피머라제(epimerase), 에폭시드 히드롤라제(hydrolase), 에스터라제, 이소머라제, 키나제, 글루코스 이소머라제, 글리코시다제, 글리코실 트랜스페라제, 할로페옥시다제, 모노옥시게나제(예컨대, p450), 리파제, 리그닌 페옥시다제, 니트릴 히드라타제(hydratase), 니트릴라제(nitrilase), 단백질분해효소, 포스파타제, 서브틸리신, 트랜스아미나제 및 뉴클레아제가 포함되지만, 이들로 한정되지 않는다.

<238> 이들 단백질의 대다수는 시판되고 있으며(예를 들어, 시그마 바이오사이언시스(Sigma BioSciences) 2002 카탈로그 및 가격 목록 참조), 상응하는 단백질 서열 및 유전자, 및 전형적으로 이들의 많은 변이체는 공지되어 있다(예를 들어, 유전자은행 참조). 임의의 이들 단백질은 본 발명에 따라 하나 이상의 비천연 아미노산의 삽입에 의해 변형되어, 예를 들어 그의 원하는 1개 이상의 치료적, 진단적 또는 효소적 성질이 변경될 수 있다. 치료적으로 관련된 성질의 예에는 혈청 반감기, 저장 수명 반감기, 안정성, 면역원성, 치료 활성, (예를 들어, 리포터기(예컨대, 표지 또는 표지 결합 부위)를 비천연 아미노산에 포함시킴에 의한) 검출용이성, LD₅₀의 감소 또는 다른 부작용, 위관을 통해 체내로 들어가는 능력(예컨대, 경구 이용가능성) 등을 포함한다. 관련된 진단적 성질의 예에는 저장 수명 반감기, 안정성, 진단 활성, 검출용이성 등이 포함된다. 관련된 효소적 특성의 예에는 저장 수명 반감기, 안정성, 효소적 활성, 생성 능력 등이 포함된다.

<239> 다양한 다른 단백질은 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 진균, 예컨대 아스페길러스(*Aspergillus*), 칸디다(*Candida*) 종; 세균, 특히 병원성 세균에 대한 모델로 사용되는 *E. 콜라*이뿐만 아니라, 스타필로코커스(*Staphylococcus*)(예컨대, *아우레우스(aureus)*), 또는 스트렙토코커스(*Streptococcus*)(예컨대, 뉴모니아에(*pneumoniae*))와 같이 의학적으로 중요한 세균; 포자류(예컨대, *플라스모디아(Plasmodia)*), 균족충류(예컨대, *엔타모에바(Entamoeba)*) 및 편모충(*트리파노소마(Trypanosoma)*, *레이시마니아(Leishmania)*, *트리코모나스(Trichomonas)*, *지아르디아(Giardia)* 등)과 같은 원생동물; (+) RNA 바이러스(예를 들어, 천연두 바이러스, 예컨대 백시니아(*vaccina*); 피코르나바이러스, 예컨대 폴리오(*polio*); 토가바이러스, 예컨대 루벨라(*rubeolla*); 플라비바이러스, 예컨대 HCV; 및 코로나바이러스를 포함함), (-) RNA 바이러스(예컨대, 래브도바이러스, 예컨대 VSV; 파라믹소빔세스, 예컨대 RSV; 오소믹소빔세스, 예컨대 인플루엔

자; 번야바이러스; 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 간염 바이러스와 같은 바이러스로부터 유래된 단백질 내의 하나 이상의 천연 아미노산을 치환시키는 것을 포함할 수 있다.

<240> 곤충 내성 단백질(예컨대, Cry 단백질), 전분 및 지질 생성 효소, 식물 및 곤충 독소, 독소-내성 단백질, 미코 톡신 해독 단백질, 식물 성장 효소(예컨대, 리불로스 1,5-비스포스페이트 카르복실라제/옥시게나제 "RUBISCO"), 리폭시게나제(LOX), 및 포스포에놀피루베이트(PEP) 카르복실라제와 같이 농업 관련 단백질도 또한 비천연 아미노산 변형에 적합한 표적이다.

<241> 특정 실시양태에서, 본 발명의 방법 및/또는 조성물에서 관심 단백질 또는 폴리펩티드 (또는 이의 일부)가 핵산에 의해 코딩된다. 전형적으로, 핵산은 1개 이상의 셀렉터 코돈, 2개 이상의 셀렉터 코돈, 3개 이상의 셀렉터 코돈, 4개 이상의 셀렉터 코돈, 5개 이상의 셀렉터 코돈, 6개 이상의 셀렉터 코돈, 7개 이상의 셀렉터 코돈, 8개 이상의 셀렉터 코돈, 9개 이상의 셀렉터 코돈, 10개 또는 그 이상 개수의 셀렉터 코돈을 포함한다.

<242> 관심 단백질 또는 폴리펩티드를 코딩하는 유전자는, 당업자에게 공지되어 있고 본 명세서의 "돌연변이유발 및 다른 문자생물학 기법(Mutagenesis and Other Molecular Biology Techniques)"에 기재된 방법을 사용하여, 비천연 아미노산의 도입을 위한 하나 이상의 셀렉터 코돈을 포함하도록 돌연변이화될 수 있다. 예를 들어, 관심 단백질에 대한 핵산은 하나 이상의 셀렉터 코돈을 포함하도록 돌연변이화되어, 하나 이상의 비천연 아미노산이 삽입되도록 한다. 본 발명은 예를 들어, 하나 이상의 비천연 아미노산을 포함하는 임의의 단백질의 임의의 상기와 같은 변이체, 예를 들어 돌연변이체, 변형태를 포함한다. 유사하게, 본 발명은 또한 상응하는 핵산, 즉 하나 이상의 비천연 아미노산을 코딩하는 하나 이상의 셀렉터 코돈을 가진 임의의 핵산을 포함한다.

<243> 비천연 아미노산을 포함하는 단백질을 제조하기 위해, 오르소고날 tRNA/RS 쌍을 통해 비천연 아미노산을 생체내 도입하도록 적합화된 숙주 세포 및 유기체를 사용할 수 있다. 숙주 세포는 오르소고날 tRNA 및 오르소고날 tRNA 합성효소를 발현하는 하나 이상의 벡터, 및 유도체화될 단백질을 코딩하는 벡터로 유전공학처리(예컨대, 형질전환, 형질도입 또는 형질감염)된다. 이 성분들 각각은 동일한 벡터 상에 존재할 수 있거나, 각기 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.

면역반응성에 의한 폴리펩티드의 정의

<245> 본 발명의 폴리펩티드가 다양한 새로운 폴리펩티드 서열(예컨대, 본원의 번역 시스템에서 합성된 단백질의 경우 비천연 아미노산, 또는 예를 들어 신규 합성효소의 경우 표준 아미노산의 신규 서열을 포함하는 폴리펩티드)을 제공하기 때문에, 폴리펩티드는 또한 예를 들어, 면역학적 검정에서 인식될 수 있는 새로운 구조적 특성도 제공한다. 본 발명의 폴리펩티드에 특이적으로 결합하는 항혈청의 생성뿐만 아니라, 상기 항혈청에 결합하는 폴리펩티드도 본 발명의 한 특성이다. 본원에 사용되는 용어 "항체"는 면역글로불린 유전자(들)에 의해 실질적으로 코딩되는 폴리펩티드, 또는 피분석물(항원)에 특이적으로 결합하여 인식하는 상기 폴리펩티드의 단편을 포함하지만, 이들로 한정되는 것은 아니다. 그 예로는 다클론 항체, 단클론 항체, 키메라 항체 및 단체 항체 등이 있다. Fab 단편, 및 과시 디스플레이를 비롯한 발현 라이브리리에 의해 생성된 단편을 비롯한 면역글로불린 단편도 또한 본원에 사용되는 용어 "항체"에 포함된다. 예를 들어, 항체 구조 및 용어에 대해서는 문헌 [Paul, *Fundamental Immunology*, 4th Ed., 1999, Raven Press, New York]을 참조한다.

<246> 면역검정에 사용하기 위한 항혈청을 생성하기 위해, 본원에 기재된 바와 같이 하나 이상의 면역원성 폴리펩티드를 제조하고 정제한다. 예를 들어, 재조합 단백질은 재조합 세포에서 생성될 수 있다. 근친교배 마우스 종(이 마우스는 사실상 유전자적으로 동일하기 때문에 결과가 더 재현가능하므로 본 검정에서 사용됨)을 표준 아주반트(adjuvant)(예컨대, 프로이드(Freund's) 아주반트) 및 표준 마우스 면역화 프로토콜과 함께 사용된 면역원성 단백질(들)로 면역화한다(예를 들어, 항체 생성, 면역검정 포맷, 및 특이적 면역반응성을 측정하는 데 이용될 수 있는 조건에 대한 표준 설명을 위해서는 문헌 [Harlow and Lane(1988) *Antibodies. A Laboratory Manual*, Cold Spring Harbor Publications, New York]을 참조한다). 단백질, 항체, 항혈청 등에 대한 추가 상세한 설명은 국제 특허 공개 공보 제WO 2004/094593호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE"); 제 2002/085923호(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 찾아볼 수 있다.

O-tRNA 및 O-RS, 및 O-tRNA/O-RS 쌍의 용도

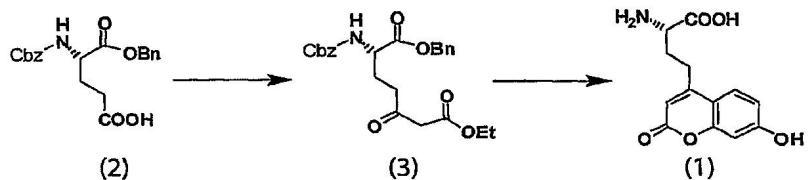
본 발명의 조성물 및 본 발명의 방법에 의해 제조되는 조성물은 임의적으로 세포 내에 있다. 본 발명의 O-tRNA/O-RS 쌍 또는 개별 성분은 이어서 비천연 아미노산이 단백질에 도입되도록 하는 숙주 시스템의 번역 기구에 사용될 수 있다. 국제 특허 공개 공보 제WO 2002/085923호(Schultz et al., 발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")는 이 방법을 기재하고, 이는 본원에 참조 인용된다. 예를 들어, O-tRNA/O-RS 쌍이 숙주, 예를 들어 에쉐리키아 콜라이 세포 또는 효모에 도입될 때, 쌍은 셀렉터 코돈, 예를 들어 엠버 넌센스 코돈에 대한 반응으로, 성장 매질에 외인적으로 침가될 수 있는 비천연 아미노산이 단백질 내, 예를 들어 미오글로빈 시험 단백질 또는 치료 단백질 내로 생체내 도입되게 된다. 임의적으로, 본 발명의 조성물은 시험관내 번역 시스템 또는 세포내 생체내 시스템(들) 내에 있을 수 있다. 비천연 아미노산을 갖는 단백질은 각종 범위의 용도들 중 임의의 용도에 사용될 수 있다. 예를 들어, 단백질 내에 도입된 비천연 부분은 넓은 범위의 변형들, 예를 들어 다른 단백질, 소분자, 예컨대 표지 또는 염료 및/또는 바이오분자와의 가교결합 중 임의의 것을 위한 표적으로서 작용할 수 있다. 이 변형으로써, 비천연 아미노산의 도입이 향상된 치료 단백질이 수득되도록 할 수 있고, 효소의 촉매 기능을 변경 또는 향상시키기 위해 사용될 수 있다. 일부 측면에서, 단백질 내로의 비천연 아미노산의 도입 및 후속 변형은 단백질 구조, 다른 단백질과의 상호작용 등에 대한 연구를 촉진할 수 있다.

키트

키트도 또한 본 발명의 한 특성이다. 예를 들어, 세포 내에 하나 이상의 비천연 아미노산을 포함하는 단백질을 생성시키기 위한 키트가 제공되는데, 여기에서 키트는 O-tRNA를 폴리뉴클레오티드 서열 및/또는 O-tRNA, 및/또는 O-RS를 코딩하는 폴리뉴클레오티드 서열, 및/또는 O-RS를 함유하는 하나 이상의 용기를 포함한다. 한 실시양태에서, 키트는 비천연 아미노산 7-히드록시쿠마린-에틸글리신을 추가로 포함한다. 또 다른 실시양태에서, 키트는 단백질 및/또는 숙주 세포를 생성시키기 위한 사용설명서를 추가로 포함한다.

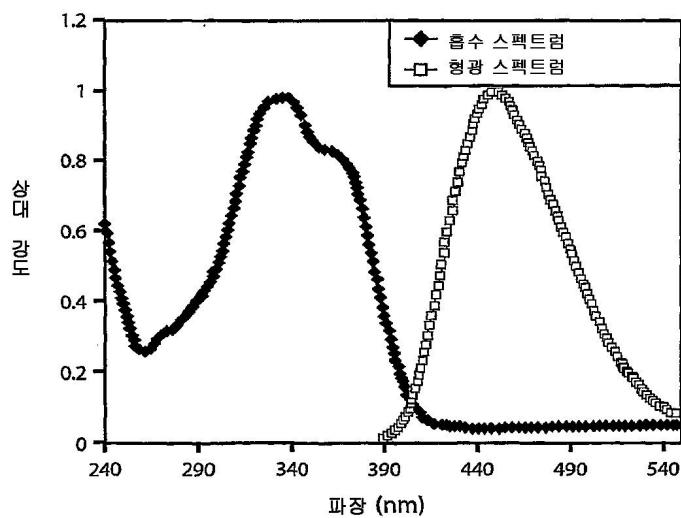
도면

도면1

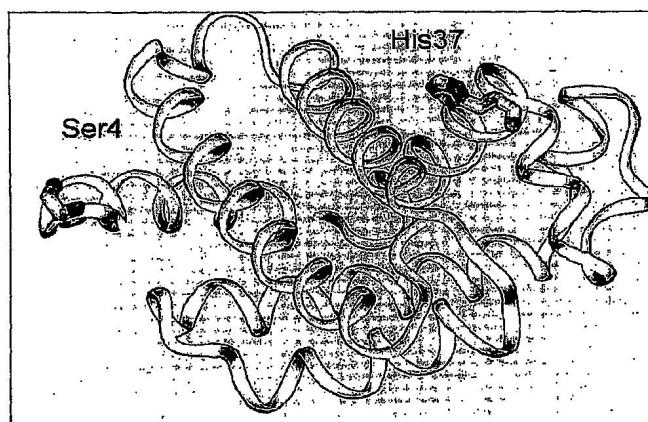


(1) $L = -(7 - \alpha)L$ 드록 시 쿠마린-4-일에 텔글리신
 (2) $N - \alpha - C_6H_5 - L - C_6H_5 - \alpha - C_6H_5$ 앤트리페인
 (3) $(2S)-2-\text{벤}[\text{질}]\text{옥시}[\text{카르보닐}] \text{아이} - 5 - \text{옥}[\text{소}]-\text{헵}[\text{탄}]\text{디}[\text{이}]\text{오}[\text{산}]$ 1-벤질 에스테르 7-에틸 에스테르

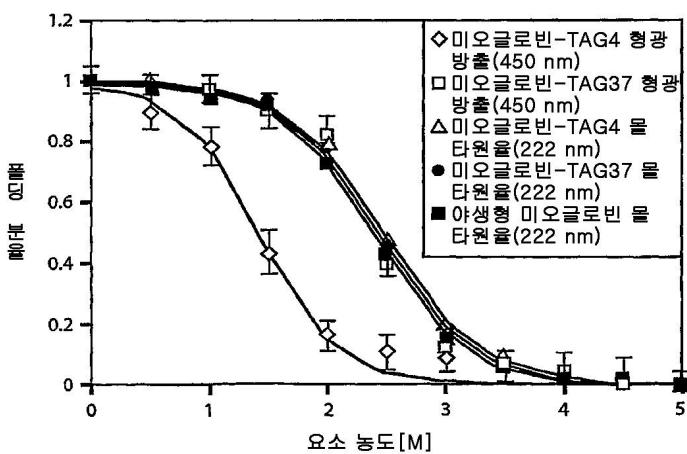
도면2



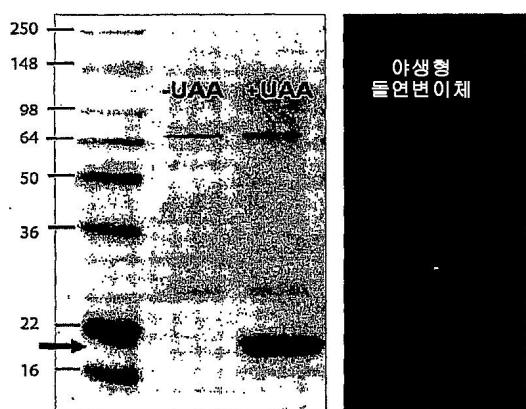
도면3



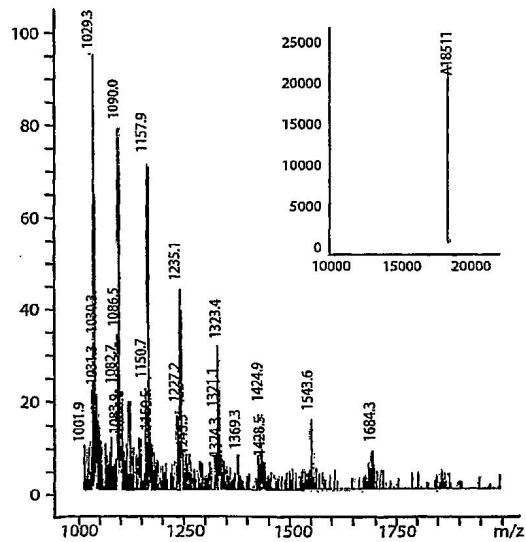
도면4



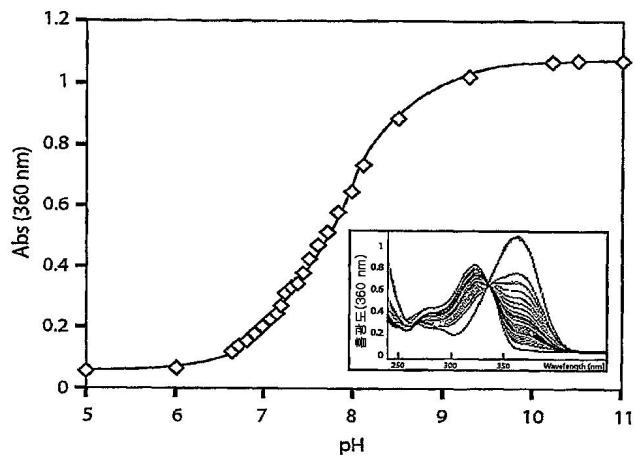
도면5



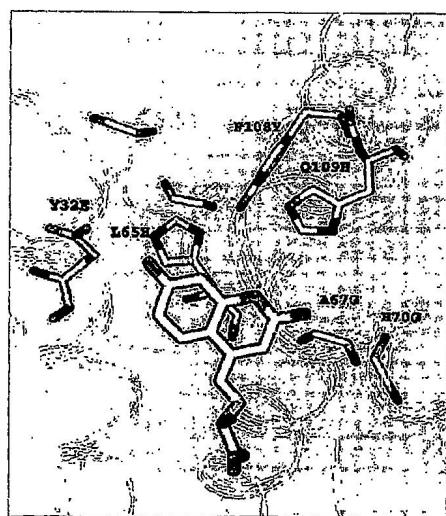
도면6



도면7



도면8



도면9

뉴클레오티드 및 아미노산 서열

서열 번호	설명	서열
1	메타노코커스 잔나스키이- 의 제자 티로실-tRNA CUA aka MjtRNA-Tyr(CUA) 또는 mutRNA _{Ty} _{CUA}	CCGGCGGUAGUUUCAGCAGGGCAGAACGGCGGACUCUAAAUCG CAUGGCGCUGGUCAAAUCGGCCCGGCCGACCA
2	아생형 메타노코커스 잔나스키이- 티로실-tRNA 합성효소 (MjtTyrRS) 아미노 산 서열	MDEFEMIKRNTSETIISSEELREVLKKDEKSAYIGFEPGKIH GHYLQIKKMLIDLQNAGFDIILHLGDLGAYLNQKGELDDEIRKIG DYNKVKFEAMGLKAKYVVGSEFOLDKDYTLNVYRLALKTTLKR ARRSMBELIAREDENPKVAEVIVPIMQVNIDHYLVDVAVGGM QRKIHMLAREELLPKKVVCIHNPVLTLGDGEKGKMSSSKGNFIAV DDSPPEIRAKIKKAYCPAGVVEGNPIMETAKYFLEYPLTIKR EKFPGDLTVNSYEELESLFKNKELEHPMDLKNAVAEELKILEP TRKRL
3	아생형 메타노코커스 잔나스키이 티로실-tRNA 합성효소 (MjtTyrRS) 뉴클레오티드 서열	ATGGACGAATTGTAAATGATAAAAGAGAACACATCTGAAATTATCAGC GAGGAAGAGGTTAACAGAGGGTTTAAAAAAAGTGAAAAATTCTGCCTAC ATAGGTTTTGAAACCAAAGTGTAAATACATTTAGGGCATTATCTCCAA ATAAAAAAAGATGATGTGATTTACAAAATGCTGGATTGTGATTAATTTATA TTGTTGGCTGATTACACCGCCTATTAAACCAGAAAAGGAGACTTGGAT GAGGTTAGBAAAATAGGGATAATTACAAAAGCTGGTTTGGAGGAA GGGTTAAAGGCAAAATATGTTATGGAAAGTGAATTCCAGCAGCTGGATAAG GATTATACACTGATGTCATAGATTTGGCTTTAAAAACTACCTTTAAA AGACCAAGAAGGAGTACGCTTAATAGCAAGAGAGGATGAAAATCCA AAGCTTCCTCAAGTTCATCCAAATGCGGTTAAAGTAAATTCATTTCTC TATTGAGGCTGATTTCAGGACTTGGAGTGGAGGATGAGCAGAGAAAATA CRACATGTTAGCAAGGSGAGCTTTDACCAAAAGGTTGTTGTTGTTTAC AACCTTCTTAACGGGTTTGGAGAGAGGAAAGATGAGTGTCTCA AAAGGGAATTTAAGCTGTTGAGACTCTCCAGAAGAGGATTTGGGCT AAGATAAAAGAAGCATACTGCCAGCTGGAGTTGTTGAGGAAATCCA ATAATGGAGATAGCTAAATCTCTGAAATTCCTTAACCCATAAAA AGGCCAGAAAAAATTGTTGGAGGATTTGACAGTTAAAGCTATGAGGAG TTAGAGACTTTATTTAAATAAGGAATGCACTCCAAATGGAAATTAAAAA AATGCTGTAGCTGAGAACCTTATAAGATTTAGAGCCAATTAGAAAG AGATTATAA
4	L-(7-히드록시 쿠마린-4-일) 에틸글리신 아미노아실-tRNA 합성효소 아미노산 서열 (아생형 메타노코커스 잔나스키이 티로실-tRNA 합성효소로부터 유래됨)	MDEFEMIKRNTSETIISSEELREVLKKDEKSAEIGFEPGKIH GHYLQIKKMLIDLQNAGFDIILHLGDLGAYLNQKGELDDEIRKIG DYNKVKFEAMGLKAKYVVGSEYHLGDLTVNLYRLALKTTLKR KRARRSMBELIAREDENPKVAEVIVPIMQVNIDHYLVDVAVG GMEQRKIHMLAREELLPKKVVCIHNPVLTLGDGEKGKMSSSKGNF IAVDDSPPEIRAKIKKAYCPAGVVEGNPIMETAKYFLEYPLTIKR EKFPGDLTVNSYEELESLFKNKELEHPMDLKNAVAEELKILEP TRKRL
5	L-(7-히드록시 쿠마린-4-일) 에틸글리신 아미노아실-tRNA 합성효소 뉴클레오티드 서열	ATGGACGAATTGTAAATGATAAAAGAGAACACATCTGAAATTATCAGC GAGGAAGAGGTTAACAGAGGGTTTAAAAAAAGTGAAAAATTCTGCCTAC ATAGGTTTTGAAACCAAAGTGTAAATACATTAGGGCATTATCTCCAA ATAAAAAAAGATGATGTGATTTACAAAATGCTGGATTGTGATTAATTTATA CATTTGGCTGATTTCAGGCTGTTAAACCAGAAAAGGAGACTTGGAT CGAGATGACAAAAGGGATAATTACAAAAGCTGGTTTGGAGGAA GGGTTAAAGGCAAAATATGTTATGGAAAGTGAATTCCAGCAGCTGGATAAG GATTATACACTGATGTCATAGATTTGGCTTTAAAAACTACCTTTAAA AGACCAAGAAGGAGTACGCTTAATAGCAAGAGAGGATGAAAATCCA AAGCTTCCTCAAGTTCATCCAAATGCGGTTAAAGTAAATTCATTTCTC TATTGAGGCTGATTTCAGGACTTGGAGGATGGAGGATGAGCAGAGAAAATA CRACATGTTAGCAAGGSGAGCTTTDACCAAAAGGTTGTTGTTGTTTAC AACCTTCTTAACGGGTTTGGAGAGAGGAAAGATGAGTGTCTCA AAAGGGAATTTAAGCTGTTGAGACTCTCCAGAAGAGGATTTGGGCT AAGATAAAAGAAGCATACTGCCAGCTGGAGTTGTTGAGGAAATCCA ATAATGGAGATAGCTAAATCTCTGAAATTCCTTAACCCATAAAA AGGCCAGAAAAAATTGTTGGAGGATTTGACAGTTAAAGCTATGAGGAG TTAGAGACTTTATTTAAATAAGGAATGCACTCCAAATGGAAATTAAAAA AATGCTGTAGCTGAGAACCTTATAAGATTTAGAGCCAATTAGAAAG AGATTATAA

서열 목록

SEQUENCE LISTING

<110> The Scripps Research Institute

Wang, Jiangyun

Xie, Jianming

Schultz, Peter G

<120> GENETICALLY ENCODED FLUORESCENT COUMARIN AMINO ACIDS

<130> 54-002020US

<140> US 11/805,247

<141> 2007-05-21

<160> 5

<170> PatentIn version 3.4

<210> 1
<211> 77
<212> RNA
<213> Artificial

<220>
<223> mutant suppressor tyrosyl-tRNA CUA

<400> 1
ccggcguag uucagcagg cagaacggcg gacucuaau ccgcauggcg cugguucaaa 60

uccggccgc cgacca 77

<210> 2
<211> 306
<212> PRT
<213> Methanococcus jannaschii

<400> 2

Met	Asp	Glu	Phe	Glu	Met	Ile	Lys	Arg	Asn	Thr	Ser	Glu	Ile	Ile	Ser
1															
					5							10			15

Glu	Glu	Glu	Leu	Arg	Glu	Val	Leu	Lys	Lys	Asp	Glu	Lys	Ser	Ala	Tyr
					20					25			30		

Ile	Gly	Phe	Glu	Pro	Ser	Gly	Lys	Ile	His	Leu	Gly	His	Tyr	Leu	Gln
					35				40			45			

Ile	Lys	Lys	Met	Ile	Asp	Leu	Gln	Asn	Ala	Gly	Phe	Asp	Ile	Ile	Ile
					50				55			60			

Leu	Leu	Ala	Asp	Leu	His	Ala	Tyr	Leu	Asn	Gln	Lys	Gly	Glu	Leu	Asp
					65				70			75		80	

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

85

90

95

Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Glu Phe Gln Leu Asp Lys
 100 105 110

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys
 115 120 125

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro
 130 135 140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Asp Ile His
 145 150 155 160

Tyr Leu Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile
 165 170 175

His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His
 180 185 190

Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser
 195 200 205

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

Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro
 225 230 235 240

Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys
 245 250 255

Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu
 260 265 270

Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys

275

280

285

Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys
 290 295 300

Arg Leu
 305

<210> 3
 <211> 918
 <212> DNA
 <213> Methanococcus jannaschii

<400> 3
 atggacgaat ttgaaatgt aaagagaaaac acatctgaaa ttatcagcga ggaagagtta 60

agagaggttt taaaaaaaga tgaaaaatct gcttacatag gtttgaaacc aagtggtaaa 120

atacatttag ggcattatct ccaaataaaa aagatgattt atttacaaaa tgctggattt 180

gatataatta tatttgttgc tgatttacac gcctatttaa accagaaagg agagttggat 240

gagattagaa aaataggaga ttataacaaa aaagttttt aagcaatggg gttaaaggca 300

aaatatgttt atggaagtga attccagctt gataaggatt atacactgaa tgtctataga 360

ttggcttaa aaactacctt aaaaagagca agaaggagta tggacttat agcaagagag 420

gatgaaaatc caaagggtgc tgaagttatc tatccaataa tgcaggtaa tgatattcat 480

tatttaggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540

agggagctt taccaaaaaa gttgtttgt attcacaacc ctgtcttaac gggttggat 600

ggagaaggaa agatgagttc ttcaaaaggaa aattttatag ctgttgatga ctctccagaa 660

gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720

ataatggaga tagctaaata cttccttcaa tatccttaa ccataaaaag gccagaaaaa 780

tttggtagat attgacagt taatagctat gaggagttag agagtttatt taaaaataag 840

gaattgcattt caatggattt aaaaaatgct gtagctgaag aacttataaa gatttagag 900

ccaattagaa agagatta 918

<210> 4

<211> 306

<212> PRT

<213> Artificial

<220>

<223> L-(7-hydroxycoumarin-4-yl) ethylglycine aminoacyl-tRNA synthetase

<400> 4

Met	Asp	Glu	Phe	Glu	Met	Ile	Lys	Arg	Asn	Thr	Ser	Glu	Ile	Ile	Ser
1					5					10			15		

Glu	Glu	Glu	Leu	Arg	Glu	Val	Leu	Lys	Lys	Asp	Glu	Lys	Ser	Ala	Glu
						20					25			30	

Ile	Gly	Phe	Glu	Pro	Ser	Gly	Lys	Ile	His	Leu	Gly	His	Tyr	Leu	Gln
							35					40		45	

Ile	Lys	Lys	Met	Ile	Asp	Leu	Gln	Asn	Ala	Gly	Phe	Asp	Ile	Ile	Ile
						50					55		60		

His	Leu	Gly	Asp	Leu	Gly	Ala	Tyr	Leu	Asn	Gln	Lys	Gly	Glu	Leu	Asp
							65					70		75	80

Glu	Ile	Arg	Lys	Ile	Gly	Asp	Tyr	Asn	Lys	Val	Phe	Glu	Ala	Met
							85					90		95

Gly	Leu	Lys	Ala	Lys	Tyr	Val	Tyr	Gly	Ser	Glu	Tyr	His	Leu	Asp	Lys
						100					105		110		

Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys
 115 120 125

Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro
 130 135 140

Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Gly Ile His
 145 150 155 160

Tyr Gly Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile
 165 170 175

His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His
 180 185 190

Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser
 195 200 205

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

Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro
 225 230 235 240

Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys
 245 250 255

Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu
 260 265 270

Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys
 275 280 285

Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys
 290 295 300

Arg Leu
305

<210> 5
<211> 921
<212> DNA
<213> Artificial

<220>
<223> L-(7-hydroxycoumarin-4-yl) ethylglycine aminoacyl-tRNA synthetase

<400> 5
atggacgaat ttgaaatgt aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60

agagaggttt taaaaaaaga tgaaaaatct gctgagatag gtttgaaacc aagtggtaaa 120

atacatttag ggcatttatct ccaaataaaa aagatgattt atttacaaaa tgctggattt 180

gatataatta tacatttggg tgatttaggc gcctatttaa accagaaagg agagttggat 240

gagattagaa aaataggaga ttataacaaa aaagttttt aagcaatggg gttaaggca 300

aaatatgttt atggaagtga atatcattt gataaggatt atacactgaa tgtctataga 360

ttggcttaa aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420

gatgaaaatc caaagggtgc tgaagttatc tatccaataa tgccagttaa tggattcat 480

tatgtggcg ttgatgtgc agttggaggg atggagcaga gaaaaataca catgttagca 540

agggagctt tacaaaaaaa ggttgttgtt attcacaacc ctgtcttaac gggtttggat 600

ggagaaggaa agatgagttc ttcaaaagg aattttatag ctgttgatga ctctccagaa 660

gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgtga aggaatcca 720

ataatggaga tagctaaata cttccttcaa tatccttaa ccataaaaag gccagaaaaa 780

tttggtagat attgacagt taatagctat gaggagttag agagtttatt taaaaataag 840

gaattgcattc caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900

ccaatttagaa agagattata a 921