



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2013년04월10일  
(11) 등록번호 10-1252835  
(24) 등록일자 2013년04월03일

(51) 국제특허분류(Int. Cl.)  
C12N 9/16 (2006.01) C12P 21/02 (2006.01)  
(21) 출원번호 10-2007-7016502  
(22) 출원일자(국제) 2005년12월01일  
심사청구일자 2010년11월08일  
(85) 번역문제출일자 2007년07월19일  
(65) 공개번호 10-2007-0092287  
(43) 공개일자 2007년09월12일  
(86) 국제출원번호 PCT/US2005/043603  
(87) 국제공개번호 WO 2006/068802  
국제공개일자 2006년06월29일  
(30) 우선권주장  
60/639,146 2004년12월22일 미국(US)  
(56) 선행기술조사문헌  
W02002085923 A1  
W02004035605 A1

(73) 특허권자  
암브룩스, 인코포레이티드  
미국 캘리포니아주 92037 라호야 노스 터레이 파  
인스 로드 10975 스위트 100  
(72) 발명자  
파울셀 앤드류  
미국 98115 워싱턴주 시애틀 38번가 애비뉴 엔이  
6549  
조 호승  
미국 92122 캘리포니아주 산디에고 아파트#107 피  
오르테라스 5225  
(74) 대리인  
김성기, 김진희

전체 청구항 수 : 총 19 항

심사관 : 김남경

(54) 발명의 명칭 아미노아실-tRNA 합성효소의 조성물 및 그것의 용도

(57) 요약

본 발명은 오르소고날 tRNA, 오르소고날 아미노아실-tRNA 합성효소, 및 tRNA/합성효소의 오르소고날 쌍을 포함하  
는 단백질 생합성 기구의 성분들을 제조하기 위한 조성물 및 방법을 제공한다. 이 오르소고날 쌍의 동정 방법이  
또한 이 오르소고날 쌍을 이용한 단백질의 생산 방법과 함께 제공된다.

**특허청구의 범위**

**청구항 1**

삭제

**청구항 2**

삭제

**청구항 3**

삭제

**청구항 4**

삭제

**청구항 5**

삭제

**청구항 6**

삭제

**청구항 7**

삭제

**청구항 8**

삭제

**청구항 9**

삭제

**청구항 10**

삭제

**청구항 11**

삭제

**청구항 12**

삭제

**청구항 13**

삭제

**청구항 14**

삭제

**청구항 15**

삭제

**청구항 16**

삭제

#### 청구항 17

삭제

#### 청구항 18

삭제

#### 청구항 19

번역 시스템을 포함하는 단리된 세포로서, 상기 번역 시스템은 서열 번호 5의 아미노아실 tRNA 합성 효소(RS)를 코딩하는 DNA 서열을 포함하는 것인 세포.

#### 청구항 20

제19항에 있어서, 세포가 진핵생물 세포인 세포.

#### 청구항 21

제20항에 있어서, 진핵생물 세포가 효모 세포인 세포.

#### 청구항 22

제20항에 있어서, 진핵생물 세포가 진균류 세포인 세포.

#### 청구항 23

제20항에 있어서, 진핵생물 세포가 포유류 세포인 세포.

#### 청구항 24

제20항에 있어서, 진핵생물 세포가 곤충 세포인 세포.

#### 청구항 25

제20항에 있어서, 진핵생물 세포가 식물 세포인 세포.

#### 청구항 26

제19항에 있어서, 세포가 비진핵생물 세포인 세포.

#### 청구항 27

제26항에 있어서, 비진핵생물 세포가 대장균(*E. coli*) 세포인 세포.

#### 청구항 28

제19항에 있어서, RS에 의해 아미노아실화된 tRNA, 및 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 추가로 포함하고, 폴리뉴클레오티드는 tRNA에 의해 인식되는 셀렉터 코돈을 포함하는 것인 세포.

#### 청구항 29

제28항에 있어서, 상기 관심 폴리펩티드가 인간 성장 호르몬인 세포.

#### 청구항 30

삭제

#### 청구항 31

삭제

**청구항 32**

삭제

**청구항 33**

삭제

**청구항 34**

삭제

**청구항 35**

삭제

**청구항 36**

서열 번호 5의 RS를 코딩하는 DNA 서열을 포함하는 벡터.

**청구항 37**

제36항에 있어서, 벡터는 플라스미드, 코스미드, 파지 또는 바이러스를 포함하는 것인 벡터.

**청구항 38**

제36항에 있어서, 벡터가 발현 벡터인 벡터.

**청구항 39**

제36항의 벡터를 포함하는 세포.

**청구항 40**

특정 위치에서 선택된 아미노산을 이용하여 세포 내에서 폴리펩티드를 생산하는 방법으로서, 상기 방법은

배지 내에서, 하나 이상의 셀렉터 코돈을 포함하고 폴리펩티드를 코딩하는 핵산을 포함하는 세포를 성장시키는 단계; 및

선택된 아미노산을 제공하는 단계

를 포함하고, 세포는

세포 내에서 기능하는 오르소고날(orthogonal) RS(O-RS)를 코딩하는 DNA 서열(여기에서, 상기 RS는 서열 번호 5의 아미노산 서열로 이루어짐); 및

셀렉터 코돈을 인식하는 오르소고날 tRNA(O-tRNA)(여기에서, 상기 O-RS는 선택된 아미노산으로 O-tRNA를 아미노아실화함)

를 추가로 포함하는 것인 방법.

**청구항 41**

제40항에 있어서, 상기 선택된 아미노산이 파라-아세틸페닐알라닌인 방법.

**청구항 42**

삭제

**청구항 43**

삭제

**청구항 44**

제40항에 있어서, 상기 폴리펩티드가 인간 성장 호르몬인 방법.

#### 청구항 45

서열 번호 5의 아미노산 서열을 코딩하는 단리된 폴리뉴클레오타이드 서열 및 이의 상보적 폴리뉴클레오타이드 서열.

### 명세서

### 기술분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은 본원에 전체적으로 인용되는 미국 가출원 특허 일련 제60/639,146호(2004년 12월 22일 출원)에 대해 우선권을 주장한다.

[0003] 기술분야

[0004] 본 발명은 번역 생화학의 분야에 속한다. 본 발명은 아미노아실-tRNA 합성효소의 생산 방법, 및 그것의 조성물, 및 상기 조성물의 용도에 관한 것이다. 본 발명은 또한 그러한 아미노아실-tRNA 합성효소 및 관련 조성물을 이용한 세포 내 단백질의 생산 방법에 관한 것이다.

### 배경기술

[0005] 세균에서 인간에 이르는 모든 공지된 유기체의 유전 코드는 동일한 20가지 통상의 아미노산을 코딩한다. 동일한 20가지 천연 아미노산의 상이한 조합들은 광합성에서 신호 전달 및 면역 반응에 이르는 실질적으로 모든 복잡한 생명 공정들을 수행하는 단백질을 형성하며, 과학자들은 단백질 구조 및 기능을 연구하고 개질하기 위해, 단백질의 유전 코드 및 아미노산 서열 모두를 조작하고자 시도하였다. 그러나, 단백질을 20가지 유전적으로 코딩된 표준 구성요소[드물 예외로서, 셀레노시스테인(예컨대, [A. Bock et al. (1991), Molecular Microbiology 5:515-20] 참고) 및 피롤리신(예컨대, [G. Srinivasan, et al., (2002), Science 296:1459-62] 참고)]로 제한하는 유전 코드에 의해 부과되는 한계점을 해소하기가 곤란하였다.

[0006] 이 한계점을 해소하기 위해 약간의 진전이 이루어졌으나, 이 진전은 제한되었고, 단백질 구조 및 기능을 합당히 조절하는 능력은 여전히 미숙한 단계이다. 예를 들어, 화학자들은 소분자의 구조를 합성하고 조작하기 위한 방법 및 기법을 개발하였다(예컨대, [E. J. Corey, & X.-M. Cheng, The Logic of Chemical Synthesis (Wiley-Interscience, New York, 1995)] 참고). 총체적 합성(예컨대, [B. Merrifield (1986), Science 232:341-7 (1986)] 참고), 및 반-합성 방법(예컨대, [D. Y. Jackson et al. (1994) Science 266:243-7; 및 P. E. Dawson, & S. B. Kent, (2000), Annual Review of Biochemistry 69:923-60] 참고)는 펩티드 및 소형 단백질을 합성할 수 있도록 하였으나, 이 방법들은 10 킬로달톤(kDa) 초과 단백질로 유용성을 제한하였다. 돌연변이유발 방법은 강력하나, 제한된 수의 구조적 변화에 제한된다. 수많은 경우들에서, 단백질 전반에 걸쳐 통상의 아미노산의 밀접한 구조적 유사체를 경쟁적으로 혼입할 수 있게 되었다. 예컨대, [R. Furter (1998), Protein Science 7:419-26; K. Kirshenbaum, et al., (2002), ChemBioChem 3:235-7; 및 V. Doring et al., (2001), Science 292:501-4]를 참고한다. 화학적 펩티드 결합 및 본연의(native) 화학적 결합이 각기 본원에 참고로 인용되는, 미국 특허 제6,184,344호, 미국 특허출원 공보 제2004/0138412호, 미국 특허출원 공보 제2003/0208046호, WO 02/098902, 및 WO 03/042235에 기재되어 있다. [Lu et al., Mol Cell. 2001 Oct; 8(4):759-69]는, 단백질을 비천연 아미노산을 함유하는 합성 펩티드에 화학적으로 결합하는 방법(발현 단백질의 결합)을 기재하고 있다.

[0007] 초기 연구는 대장균의 번역 기구가 통상의 20가지 아미노산과 구조적으로 유사한 아미노산을 수용함을 입증하였다. [Hortin, G. 및 Boime, I. (1983) Methods Enzymol. 96:777-784]를 참고한다. 이 연구는 내인성 대장균 합성효소의 특이성을 이완하여, 그것이 비천연 아미노산 및 그것의 동족 천연 아미노산을 활성화함으로써 더 연장되었다. 또한, 도메인 편집에 있어서의 돌연변이를 또한 이용하여 내인성 합성효소의 기질 범주를 연장할 수 있는 것으로 나타났다. [Doring, V., et al., (2001) Science 292:501-504]를 참고한다. 그러나, 이 기법들은 유전 코드를 확장하기보다 유전 코드를 기록하는데 제한되고, 통상의 20가지 아미노산들 중 하나의 비천연 아미노산으로의 치환도를 다양하게 한다.

[0008] 이후, 화학적으로 아미노아실화된 오르소고날(orthogonal) 앰버 역제자 tRNA를 시험관 내 전사/번역 반응에 부가함으로써, 비천연 아미노산을 시험관 내 단백질로 부위-특이적으로 혼입할 수 있는 것으로 나타났다. 예컨대,

[Noren, C. J., et al. (1989) Science 244:182-188; Bain, J. D., et al. (1989) J. Am. Chem. Soc. 111:8013- 8014; Dougherty, D. A. (2000) Curr. Opin. Chem. Biol. 4, 645-652; Cornish, V. W., et al. (1995) Angew. Chem., Int. Ed. 34: 621-633; J. A. Ellman, et al. (1992), Science 255:197- 200; 및 D. Mendel, et al. (1995), Annual Review of Biophysics and Biomolecular Structure 24:435-462]를 참고한다. 이 연구들은 리보솜 및 번역 인자들이 매우 많은 수의 비천연 아미노산들, 심지어 비통상적 구조의 그러한 아미노산들과 상용적임을 보여준다. 안타깝게도, tRNA의 화학적 아미노아실화가 곤란하고, 이 공정의 화학양론적 성질은 발생될 수 있는 단백질의 양을 심각히 제한하였다.

[0009] 비천연 아미노산을 세포 내에 미세주사하였다. 예를 들어, 화학적으로 잘못 아실화된 테트라히메나 썬모필라 (*Tetrahymena thermophila*) tRNA(예컨대, [M.E. Saks, et al. (1996), *An engineered Tetrahymena tRNA<sup>Gln</sup> for in vivo incorporation of unnatural amino acids into proteins by nonsense suppression*, J. Biol. Chem. 271:23169-23175] 참고), 및 관련 mRNA를 미세주입함으로써, 비천연 아미노산을 제노푸스(*Xenopus*) 난모세포 내 니코틴성 아세틸콜린 수용체에 도입하였다(예컨대, [M. W. Nowak, et al. (1998), *In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system*, Method Enzymol. 293:504-529]). 또한, [D.A. Dougherty (2000), *Unnatural amino acids as probes of protein structure and function*, Curr. Opin. Chem. Biol. 4:645-652 및 M. W. Nowak, P. C. Kearney, J. R. Sampson, M. E. Saks, C. G. Labarca, S. K. Silverman, W. G. Zhong, J. Thorson, J. N. Abelson, N. Davidson, P. G. Schultz, D. A. Dougherty 및 H. A. Lester, Science, 268:439 (1995)]를 참고한다. 제노푸스 난모세포에 시험관 내 제작한 2개 RNA 중, 즉 관심 아미노산 위치에서 UAG 종결 코돈을 갖는 표적 단백질을 코딩하는 mRNA, 및 원하는 비천연 아미노산으로 아미노아실화된 앰버 억제자 tRNA를 동시 주입하였다. 이어서, 난모세포의 번역 기구를 UAG에 의해 특정화된 위치에 비천연 아미노산을 주입한다. 안타깝게도, 이 방법은 미세주입될 수 있는 세포 내 단백질에 제한되고, 관련 tRNA가 시험관 내 화학적으로 아실화되고 재아실화될 수 없기 때문에, 단백질의 수율이 매우 낮다.

[0010] 이 한계점을 극복하기 위해, 새로운 성분들, 예컨대 오르소고날 tRNA, 오르소고날 아미노아실-tRNA 합성효소 및 그들의 쌍을 원핵생물 *에شري키아 콜라이*(*Escherichia coli*)(대장균)(예컨대, [L. Wang, et al., (2001), Science 292:498-500] 참고), 및 비유전적으로 코딩된 아미노산을 단백질에 생체 내 혼입할 수 있도록 한 진핵생물 *사카로마이세스 세레비지아에*(*Saccharomyces cerevisiae*)(*S. cerevisiae*)(예컨대, [J. Chin et al., Science 301:964-7 (2003)] 참고)의 단백질 생합성 기구에 첨가하였다. 광친화성 표지 및 광중합성 아미노산, 광가교성 아미노산(예컨대, [Chin, J. W., et al. (2002) Proc. Natl. Acad. Sci. U.S.A. 99:11020-11024; 및 Chin, J. W., et al., (2002) J. Am. Chem. Soc. 124:9026-9027] 참고), 케토 아미노산(예컨대, [Wang, L., et al., (2003) Proc. Natl. Acad. Sci. U.S.A. 100:56-61, 및 Zhang, Z. et al., Biochem. 42(22):6735-6746 (2003)] 참고), 중원자 함유 아미노산, 및 글리코실화된 아미노산을 포함한 신규 화학적, 물리적 또는 생물학적 성질을 갖는 수많은 새로운 아미노산들이 상기 방법을 이용하여, 예컨대 앰버 코돈(TAG)에 대한 반응으로 대장균 및 효모 내 단백질로 효율적으로 또한 고 적합도로 혼입되었다. 예컨대, [J. W. Chin, & P. G. Schultz, (2002), ChemBioChem 3(11): 1135-1137, 및 L. Wang, & P. G. Schultz, (2002), Chem. Comm. 1:1-11]을 참고한다.

[0011] 수가지 기타 오르소고날 쌍들이 보고되었다. *S. 세레비지아에* tRNA 및 합성효소로부터 유래된, 글루타밀(예컨대, [Liu, D. R., 및 Schultz, P. G.(1999) Proc. Natl. Acad. Sci. U.S.A. 96:4780-4785] 참고), 아스파르틸(예컨대, [Pastrnak, M., et al., (2000) Helv. Chim. Acta 83:2277-2286] 참고), 및 티로실(예컨대, [Ohno, S., et al. (1998) J. Biochem.(Tokyo, Jpn.) 124:1065-1068; 및 [Kowal, A. K., et al., (2001) Proc. Natl. Acad. Sci. U.S.A. 98:2268-2273] 참고) 시스템들이 대장균 내 비천연 아미노산으로의 잠정적 혼입에 대해 기재되었다. 대장균 글루타미닐(예컨대, [Kowal, A. K., et al., (2001) Proc. Natl. Acad. Sci. U.S.A. 98:2268-2273] 참고) 및 티로실(예컨대, [Edwards, H., 및 Schimmel, P.(1990) Mol. Cell. Biol. 10:1633-1641] 참고) 합성효소로부터 유래된 시스템이 *S. 세레비지아에*에 사용됨에 대해 기재되었다. 대장균 티로실 시스템이 포유동물 세포에서 생체 내 3-요오도-L-티로신의 혼입을 위해 사용되었다. [Sakamoto, K., et al., (2002) Nucleic Acids Res. 30:4692-4699]를 참고한다. 전형적으로, 이 시스템들은 앰버 종결 코돈을 이용하였다. 유전 코드를 더욱 확장하기 위해, 향상되고/되거나 부가적인 생합성 기구 성분, 예컨대 아미노아실-tRNA 합성효소의 개발이 필요하다. 하기 개시내용을 검토해볼 때 자명하게 되는 바, 본 발명은 상기 및 기타 필요를 이행한다.

## 발명의 상세한 설명

[0012]

### 발명의 개요

[0013]

유전 코드를 확장하기 위해, 본 발명은 오르소고날 아미노아실-tRNA 합성효소의 생산을 위한 조성물 및 방법을 제공한다. 본 발명의 아미노아실-tRNA 합성효소는 비천연적으로 코딩된 아미노산으로 tRNA를 아미노아실화한다. 이 번역 성분을 사용하여, tRNA에 의해 인식되는 셀렉터 코돈에 대한 반응으로 (핵산 번역 중에) 성장 폴리펩티드 사슬 내 특정 위치에 선택된 아미노산을 혼입할 수 있다.

[0014]

특정 위치에 선택된 아미노산을 갖는 세포 내 단백질의 생산 방법도 또한 본 발명의 한 특징이다. 예를 들어, 방법은 한 적절한 매질 내에, 하나 이상의 셀렉터 코돈을 포함하고 단백질을 코딩하는 핵산을 포함하는 세포를 성장시키는 단계; 및 선택된 아미노산을 제공하는 단계를 포함한다. 세포는 세포 내에서 기능하고 셀렉터 코돈을 인식하는 오르소고날 tRNA(O-tRNA), 및 선택된 아미노산으로 O-tRNA를 우선적으로 아미노아실화하는 오르소고날 아미노아실-tRNA 합성효소(O-RS)를 추가로 포함한다. 전형적으로, O-tRNA는 동족 합성효소의 존재 하에 억제 활성을 가진다. 이 방법에 의해 생산된 단백질도 또한 본 발명의 한 특징이다.

## 실시예

[0266]

하기 실시예는 청구하는 발명을 예시하기 위해 제공되되는 것으로서, 청구 발명을 제한하려는 것이 아니다. 당업자는 청구된 발명의 범주를 벗어나지 않으면서 변경될 수 있는, 다양한 비결정적 매개변수들을 인식할 것이다.

[0267]

### 실시예 1:

[0268]

### 파라-아세틸페닐알라닌에 대한 아미노아실-tRNA 합성효소 선택

[0269]

2개 DNA 라이브러리를 비천연적으로 코딩된 아미노산인 파라-아세틸페닐알라닌에 대한 아미노아실-tRNA 합성효소에 대해 선별(screening)하였다. 이 라이브러리는 pBK 플라스미드에서 *메타노코커스 잔네쉬*로부터의 티로실 tRNA 합성효소 유전자 내 6개 돌연변이로 구성되었다.

[0270]

5회 교대 실험의 선택, 즉 3회 양성, 2회 음성 선택으로 구성된 선택 절차를 수행하였다. 라이브러리를 1:1 비로 조합하였고, 양성 선택 세포주(양성 선택 플라스미드를 갖는 GeneHog, pREP)로 전기천공하였고, 적절한 항생제 및 비천연적으로 코딩된 아미노산 파라-아세틸페닐알라닌(pAF)을 갖는 최소 배지 플레이트(GMML) 상에 도말하였다. 플레이트를 약 40시간 동안 37°C에서 인큐베이션하였고, 이 시점에서 세포를 찰과도말(scraping)에 의해 수확하였다. DNA를 키아젠(Qiagen) 미니-프랩(Mini-Prep) 절차를 이용하여 추출한 후, 아가로스 겔로 정제하여, 라이브러리 플라스미드 DNA를 분리하였다.

[0271]

이어서, 이 DNA를 음성 선택 세포주(음성 선택 플라스미드 pBAD 유도체를 갖는 GeneHog)에 전기천공하였다. 이 형질전환체를 비천연적으로 코딩된 아미노산(pAF)없이 적절한 항생제가 있는 LB 플레이트 상에 도말하였다. 약 17시간 후에, 이 세포를 찰과도말에 의해 수확하였고, 플라스미드 DNA를 키아젠 미니-프랩 절차 및 아가로스 겔 정제를 이용하여 정제하였다.

[0272]

후속 실험(round)들의 선택 실험을 전기천공, 도말, 수확 및 DNA 정제의 동일한 방법을 이용하여 수행하였다. 마지막(제5회) 선택 실험에서, 최소 배지 플레이트 상에 도말된, 형질전환된 양성 선택 세포의 일련의 희석을 수행하였다. 이어서, 개별 콜로니를 선택하여, 하룻밤 동안 96 웰에서 성장시켰다. 이어서, 이 블록을 비천연 아미노산 pAF가 존재 또는 부재하는, 다양한 농도의 클로람페니콜(양성 선택 항생제)이 있는 최소 배지 상에 반 복 도말하였다. 37°C에서 약 40시간 동안 성장시킨 후, 플레이트를 시각적으로 비교하여, 어떤 콜로니가 최대의 클로람페니콜 농도에서 성장하였는지를 결정하였으나, 비천연적으로 코딩된 아미노산 pAF의 부재 하에서는 성장하지 않았거나 잘 성장하지 않았다. 이 기준을 만족한 콜로니를 하룻밤 동안 성장시켰다. DNA를 미니-프랩 및 아가로스 겔 정제에 의해 배양물로부터 분리하였고, 서열결정하였다.

[0273]

pAF에 대한 이 선택으로부터, 13개 클론이 독특한 아미노산 서열을 가지는 것으로 나타났고, 후속하여 특징화하여, pAF-tRNA 합성효소의 적합도 및 가공성을 결정하였다.

[0274]

이 합성효소를 특징화하기 위해, 작은 스케일의 앰버 억제제를 수행하였고, 이에 비천연적으로 코딩된 아미노산 pAF이 폴리펩티드에 혼입된 것으로 나타났으며, 결과를 SDS-PAGE에 의해 가시화하였다. 단일 콜로니를 선택하여 LB 브로쓰에서 하룻밤 동안 성장시켰고, 이를 이어서 50 mL의 LB를 접종하기 위해 사용하였다. 세포를 0.3-



0.4의 OD로 성장시켰고, 이 시점에서 1.5 mL 분취량(aliquot)을 유도전 시점으로 하였고, 배양물을 2개 플라스크로 나누었다. 1 mM pAF를 한 분량(split)에 첨가하였고, 양자 모두를 30분 동안 성장시켰다. 30분 동안 성장시킨 후, 양 배양물(+/- pAF)를 0.2% L-아라비노스로 유도하였고, 4.5시간 동안 성장시켰으며, OD<sub>600</sub>을 기록하였다. 이어서, 1.5 mL 분취량을 SDS-PAGE 분석용 +/- pAF 플라스크에 취했다.

[0275] 1.5 mL 분취량(유도전, + pAF, -pAF)을 10분 동안 10,000 xg으로 원심분리하여, 세포를 펠렛화하였다. 이어서, 세포를, 수확 시 그것의 OD<sub>600</sub>에 대한 상대량으로 비례적 세균 단백질 추출 시약(BPER, 피어스(Pierce))에 현탁시켰다. DNase I를 용해된 세포에 첨가하여, 20분 동안 4℃에서 인큐베이션하였다. 이어서, 샘플을 환원제 및 로딩 염색과 조합하여, 30분 동안 MES 완충액에서 4-12% 비스-트리스 상에 실행시켰다. 겔을 10분 동안 2회 DI H<sub>2</sub>O로 세정하였고, 쿠마씨 블루 염료로 염색하였다. +/- pAF 밴드를 AF의 혼입을 초래하는 pAF-tRNA RS의 적합도에 대해 비교하였고, + pAF 밴드를 이전 선택된 pAF-tRNA RS와 비교하였다.

[0276] RS의 가공성을 체크하기 위해, C-H6 S4am 미오글로빈(S4am-Myo)을 함유하는 플라스미드를 이용하여 동일한 절차를 수행하였다. 이어서, S4am Myo를 IMAC에 의해 정제하고, 단백질 서열결정을 행하여, pAF 혼입량을 결정하였다.

[0277] 이 선택으로부터 동정된 pAF-tRNA RS 중, 한 합성효소(E9)가 pAF를 효율적으로 혼입한 것으로 나타났고, 95% 초과 효율로 pAF를 S4am-Myo에 혼입시켰다. 아미노산 서열결정에 의해 혼입을 결정하였고, 한편 SDS-PAGE 겔 상의 단백질 밴드를 비교함으로써 가공성을 나타냈다. E9에 대한 뉴클레오타이드 서열이 서열 번호 4에 나와 있고, E9의 아미노산 서열이 서열 번호 5에 나와 있다.

[0278] E9와 유사한 활성을 갖는 부가적 돌연변이를 동정하였고, 서열 번호 17에 나와 있는 아미노산 서열을 가진다.

## [0279] 실시예 2:

### [0280] tRNA 돌연변이유발

[0281] tRNA J17의 3개 돌연변이를 발생시켰다. 야생형 J17의 DNA 서열이 서열 번호 8로 나와 있고, 미국 특허출원 공보 제2003/0108885호에서는 서열 번호 1로, 또한 US 제2003/0082575호에서는 서열 번호 1(각각 미국 특허 출원 일련 번호 제10/126,931호 및 제10/126,927호)로 나와 있고, 이들 모두는 본원에 참고로 인용된다. J17 tRNA는 도 1에 나와 있는 TΨC 줄기에서 U51 :G63 동요 쌍을 가진다.

[0282] 3개 J17 돌연변이(F12, F13 및 F14)를 발생시켜, TΨC 줄기의 위치 51 및 63에서 와트슨-크릭(Watson-Crick) 염기 쌍을 발생시켰다. 중첩 PCR에 의해 돌연변이유발을 수행하였고, 최종 구축물을, 아미노아실 tRNA 합성효소 E9을 코딩하는 폴리뉴클레오타이드 서열(서열 번호 4) 및 앰버 코돈 치환을 갖는 인간 성장 호르몬(hGH)을 코딩하는 폴리뉴클레오타이드 서열(서열 번호 16)을 포함하는 pET19 플라스미드 내 EcoRI 및 NdeI 부위에 클로닝하였다. hGH의 발현은 T7 프로모터의 조절을 받는다.

[0283] 중첩 PCR에 대해 2개 단편을 발생시켰다. 프라이머 연장에 의해 제1 단편을 수득하였다. 3개 돌연변이 각각을 발생시키는데 사용된 전방 프라이머의 서열은 하기와 같았다:

[0284] GTAACGCTGAATTCCTGGCGGTAGTTTCAGCAGGGCAGAACGGCGGACTCTAAATCCGCATGGCGC(FTam11; 서열 번호 9).

[0285] F12 돌연변이(51C:63G)를 발생시키기 위해, 하기 역방 프라이머를 사용하였다:

[0286] GATCTGCAGTGGTCCGGCGGGCCGATTGTAACCGCGCCATGCGGATTTAGAGTCCGCCGTTCTGC(FTam12; 서열 번호 10).

[0287] F13 돌연변이(51U:63A)를 발생시키기 위해, 하기 역방 프라이머를 사용하였다:

[0288] GATCTGCAGTGGTCCGGCGGGCTGGATTGTAACAGCGCCATGCGGATTTAGAGTCCGCCGTTCTGC(FTam13; 서열 번호 11).

[0289] F14 돌연변이(51A:63U)를 발생시키기 위해, 하기 역방 프라이머를 사용하였다:

[0290] GATCTGCAGTGGTCCGGCGGGCAGGATTGTAACCTGCGCCATGCGGATTTAGAGTCCGCCGTTCTGC(FTam14; 서열 번호 12).

[0291] 제2 단편을 발생시키기 위해, J17 tRNA에 대한 폴리뉴클레오타이드 서열(서열 번호 8), tRNA 합성효소 E9를 코딩하는 폴리뉴클레오타이드 서열(서열 번호 4), 및 앰버 코돈 치환을 갖는 인간 성장 호르몬을 코딩하는 폴리뉴클레오타이드 서열(서열 번호 16)을 포함하는 플라스미드 pET19 J17 E9 hGH를 하기 프라이머 세트를 이용한 증폭용 주형으로 사용하였다: CGCCGACCACTGCAGATCCTTAGCGAAAGCTAAGGATTTTTTTAAGC (전방 프라이머; FTam15; 서열 번호 13) 및 CAAATTCGTCCATATGGGATTCC(FTam16; 서열 번호 14). 전방 프라이머를 사용하여, 서열을 tRNA의 3' 말단에



서 플라스미드의 Nde I 부위로 연장시켰다. 수득된 산물을 겔 정제하였다.

[0292] 중첩 PCR의 최종 단계는 전방 프라이머 GTAACGCTGAATCCCGGCG(FTam17, 서열 번호 15), 역방 프라이머 FTam16 (서열 번호 14), 제1 단편 및 제2 단편을 포함하였다. 조립된 산물을 EcoR I 및 Nde I로 소화시켜, EcoR I 및 Nde I로 소화된 플라스미드 pET19 J17 E9 hGH에 결합시켰다. 각 구축물의 서열을 서열결정으로 확인하였고, 각각의 J17 돌연변이 tRNA의 DNA 서열이 서열 번호 1(F12), 서열 번호 2(F13), 및 서열 번호 3(F14)으로 나와 있다. 상응하는 역방 프라이머를 따라 tRNA를 명명하였다.

#### [0293] 단백질 발현

[0294] tRNA를 코딩하는 플라스미드(J17, F12, F13 또는 F14)를 화학적 수단에 의해 대장균 균주 1 및 균주 2의 세균 숙주 세포로 형질전환하여, 50 ug/ml 카르베니실린(carbenicillin)을 갖는 LB 아가 플레이트 상에 도말하였다. 플레이트를 하룻밤 동안 37℃에서 인큐베이션하였다. 각 tRNA에 대해, 단일 콜로니를 선택하여, 50 ug/ml 카르베니실린을 갖는 1 ml 2x YT 내 37℃에서 하룻밤 배양을 시작하였다. 이 1 ml 배양을 이용하여, 37℃에서 50 ug/ml 카르베니실린을 갖는 2개 10 ml 2xYT 배양액을 접종하였다. 1개 10 ml 배양액에 4 mM 파라-아세틸페닐알라닌을 보충하였다. OD<sub>600</sub>=0.7에서, 0.4 mM IPTG로 hGH 발현을 유도하였다. 250 rpm으로 4시간 동안 37℃에서 세포를 배양한 후, 세포를 5분 동안 5000 x g에서 원심분리하여 수확하였다. 세포를 5 ug/ml DNase I가 보충된 B-PER 시약(피어스, 미국 일리노이즈주 로크포스 소재)으로 용해시켰다. 4-12% SDS PAGE에 의해 총 세포 용해물을 분석하였다.

[0295] 도 2는 SDS PAGE에 의한 대장균 균주 1의 총 세포 용해물의 분석을 나타낸다. J17 또는 J17 돌연변이(F12, F13, F14) tRNA 및 아미노아실 tRNA 합성효소 E9를 이용하여, 인간 성장 호르몬 내 선택적 코돈의 역제를 수행하였다. J17 돌연변이를 보유한 세포는 J17를 보유한 세포보다 약간 더 늦게 성장하였다. 4 mM 파라-아세틸페닐알라닌의 부재 하에 tRNA 돌연변이에 대해 SDS-PAGE에 의해 전장 hGH 산물이 관찰되지 않았다. 4 mM 파라-아세틸페닐알라닌의 존재 하에, 각각의 tRNA 돌연변이를 갖는 전장 산물을 생산하였고, 이는 이 tRNA 돌연변이-RS E9 쌍이 대장균 기전에 대해 오르소고날성임을 나타낸다. SDS-PAGE에 기초하여, J17 돌연변이의 억제된 hGH 수율은 대장균 균주 1에서의 J17의 그 수율보다 대략 1.5-2배 더 높았다.

[0296] 1개 J17 돌연변이인 F13를 도 3에 나와 있는 앰버 억제에 대해 대장균 균주 2의 세균 세포주에서 더욱 시험하였다. 대장균 균주 2에서, 앰버 억제 및 발현 수율은 대장균 균주 1에서의 수율에 비해 감소되었다. 파라-아세틸페닐알라닌의 부재 하에, SDS-PAGE에 의해 전장 hGH 산물이 관찰되지 않았다. 4 mM 파라-아세틸페닐알라닌의 존재 하에, 양 tRNA에 대해 전장 hGH가 관찰되었다. SDS-PAGE에 기초하여, F13의 억제된 hGH 수율은 J17의 수율보다 약 3배 더 높았다.

[0297] J17 및 F13를 비교하는 발효 실행(round)을 대략 1.5 L의 최종 체적으로 수행하였다. J17 tRNA를 코딩하는 플라스미드 및 F13 tRNA를 코딩하는 플라스미드를 각각 대장균 균주 1로 형질전환하였다. 각각에 대한 최종 세포 밀도는 대략 190 g 습한 세포/L였다. hGH 역가는 J17 클론의 경우 347 mg/L이었고, F13 클론의 경우에는 542 mg/L이었다.

[0298] 상기 발명은 발명의 내용을 보다 명료히 하고 이해시키기 위해 다소 상세히 기재되었으나, 본 개시내용을 읽음으로써 당업자는 본 발명의 범주를 벗어나지 않는 한, 형태 및 상세내용에 있어 각종 변화가 가해질 수 있음이 명백해질 것이다. 예를 들어, 상기 모든 기법 및 장치들은 각종 조합들로 사용될 수 있다. 본원에 언급된 모든 공보, 특허, 특허 출원, 및/또는 기타 문헌들은, 마치 각 개별 공보, 특허, 특허 출원, 및/또는 기타 문헌이 모든 목적을 위해 인용되는 것으로 개별적으로 나타난 것과 동일한 정도로 모든 목적을 위해 전체적으로 참고로 인용된다.

표 1

서열번호	표지	서열
1	F12 DNA	CCGGCGGTAGTTCAGCAGGGCAGAACGGCGGACTCTAAATCCGCA TGGCGCCGGTTCAAATCCGGCCCGCCGGACCA
2	F13 DNA	CCGGCGGTAGTTCAGCAGGGCAGAACGGCGGACTCTAAATCCGCA TGGCGCTGGTTCAAATCCAGCCCGCCGGACCA
3	F14 DNA	CCGGCGGTAGTTCAGCAGGGCAGAACGGCGGACTCTAAATCCGCA TGGCGCAGGTTCAAATCCTGCCCCCGCGACCA
4	E9 RS 핵산	ATGGACGAATTTGAAATGATAAAGAGAAAACACATCTGAAATTATC AGCGAGGAAGAGTTAAGAGAGGTTTTAAAAAAGATGAAAAATC TGCTGTTATAGGTTTGAACCAAGTGGTAAAAATACATTTAGGGCAT TATCTCCAAATAAAAAAGATGATTGATTACAAAAATGCTGGATTG ATATAATTATATATTTGGCTGATTTACACGCCTATTAAACCAGAA AGGAGAGTTGGATGAGATTAGAAAAATAGGAGATTATAACAAAA AAGTTTTTGAAGCAATGGGGTTAAAGGCAAAATATGTTTTATGGAA GTGAACGATGGTCTTGATAAGGATTATACACTGAATGTCTATAGATT GGCTTTAAAAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAAC TATAGCAAGAGAGGATGAAAATCCAAAGGTGCTGAAGTTATCTA TCCAATAATGCAGGTTAATGGGATTCATTATGAGGGCGTTGATGTT GCAGTTGGAGGGATGGAGCAGAGAAAAATACACATGTTAGCAAG GGAGCTTTTACCAAAAAAGGTTGTTTGTATTACAACCTGTCTTA ACGGGTTTGGATGGAGAAGGAAAGATGAGTTCTTCAAAGGGAAT TTTATAGCTGTTGATGACTCTCCAGAAGAGATTAGGGCTAAGATAA AGAAAGCATACTGCCAGCTGGAGTTGTTGAAGGAAATCCAATAA TGGAGATAGCTAAATACTTCCCTTGAATATCCTTTAACCATAAAAAAG GCCAGAAAAATTTGGTGGAGATTTGACAGTTAATAGCTATGAGGA GTTAGAGAGTTTATTTAAAAATAAGGAATTGCATCCAATGGATTTA AAAAATGCTGTAGCTGAAGAACTTATAAAGATTTTAGAGCCAATT AGAAAGAGATTATAA
5	E9 RS 아미노산	MDEFEMIKRNTSEIISEELREVLKKDEKSAVIGFEPGSKIHLGHYLQIK KMIDLQNAQFDIIYLAADLHAYLNQKGELDEIRKIGDYNKKVFAMGL KAKYVYVYGEHGLDKDYTLNVYRLALKTTTLKRARRSMELIAREENP KVAEVIYIMQVNGIHYEGVDVAVGGMEQRKIHMLARELLPKKVCCI HNPVLTGLDGEGKMSSSKGNFIAVDDSPFIRAKIKKAYCPAGVVEG NPIMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNKELHPMDL KNAVAEELIKILEPIRKRL
6	HL(TAG)3 tRNA DNA	CCCAGGGTAGCCAAGCTCGGCCAACGGCGACGGACTCTAAATCCG TTCTCGTAGGAGTTCGAGGGTTTCAATCCCTTCCCTGGGACCA
7	HL(TGA)1 tRNA DNA	GCGGGGGTTGCCGAGCCTGGCCAAAGGCGCCGACTTCAAATCCG GTCCCGTAGGGGTTCCGGGGTTCAAATCCCCGCCCCCGCACCA
8	J17 M. 갈대췌이 mtRNA <sup>Tyr</sup> <sub>CUA</sub> DNA	CCGGCGGTAGTTCAGCAGGGCAGAACGGCGGACTCTAAATCCGCA TGGCGCTGGTTCAAATCCGGCCCGCCGGACCA
9	FTam11	GTAACGCTGAATTCCTGGCGGTAGTTCAGCAGGGCAGAACGGCGG

[0299]

	프라이머	ACTCTAAATCCGCATGGCGC
10	FTam12 프라이머	GATCTGCAGTGGTCCGGCGGGCCGGATTGAACCGGCGCCATGCG GATTAGAGTCCGCCGTCTGCG
11	FTam13 프라이머	GATCTGCAGTGGTCCGGCGGGCTGGATTGAACAGCGCCATGCG GATTAGAGTCCGCCGTCTGCG
12	FTam14 프라이머	GATCTGCAGTGGTCCGGCGGGCAGGATTGAACCTGCGCCATGCG GATTAGAGTCCGCCGTCTGCG
13	FTam15 프라이머	CGCCGGACCACTGCAGATCCTTAGCGAAAGCTAAGGATTTTTTTA AGC
14	FTam16 프라이머	CAAAATTCGTCCATATGGGATTCC
15	FTam17 프라이머	GTAACGCTGAATTCCCGGCG
16	hGH (DNA)	ATGGGCCACCAACCAACCAACCACTTCCCAACCATCCCTTATCCA GGCTTTTTGACAACGCTATGCTCCGCGCCCATCGTCTGCACAGCT GGCCTTTGACACCTACCAGGAGTTTGAAGAAGCCTAGATCCCAA GGAACAGAAGTATTCATTCTCGAGAACCCCGAGACCTCCCTCTGT TTCTCAGAGTCTATTCCGACACCCCTCCAACAGGGAGGAAACACAA CAGAAATCCAACCTAGAGCTGCTCCGCATCTCCCTGCTGCTCATCC AGTCGTGGCTGGAGCCCGTGCAGTTCTCAGGAGTGTCTTCGCCAA CAGCTGGTGTACGGCGCCTCTGACAGCAACGTCTATGACCTCCTA AAGGACCTAGAGGAAGGCATCCAAACGCTGATGGGGAGGCTGGA AGATGGCAGCCCCGGACTGGGCAGATCTTCAAGCAGACCTACAG CAAGTTCGACACAAACTCACACAACGATGACGCACTACTCAAGAA CTACGGGCTGCTCTACTGCTTCAGGAAGGACATGGACAAGGTGCA GACATTCCTGCGCATCGTGCAGTGCCGCTCTGTGGAGGGCAGCTGT GGCTCTCAA
17	D286R E9의 돌연변이	MDEFEMIKRN TSEIISEEL REVLKKDEKS AVIGFEPSPGK IHLGHYLQIK KMDLQNA GF DIHYLADLH AYLNQKGELD EIRKIGDYNK KVFEAMGLKA KYVYGSEHGL DKDYTLNVYR LALKTTLKRA RRSMEIARE DENPKVAEVI YPIMQVNGIH YEGVDVAVGG MEQRKIHMLA RELLPKKVVC IHNPLVTGLD GEGKMSSSKG NFIAVDDSP EIRAKIKKAY CPAGVVEGNP IMEIAKYFLE YPLTIKRPEK FGGDLTVNSY EELESFLKNK ELHPMRLKNA VAEELKILE PIRKRL

[0300]

## 도면의 간단한 설명

[0015]

도 1 - TΨC 줄기 돌연변이 부위를 갖는 J17 tRNA의 클로버릴 구조가 나와 있다.

[0016]

도 2 - J17 또는 J17 돌연변이(F12, F13, F14) 및 E9 RS를 이용하여 인간 성장 호르몬 내 앰버 돌연변이의 억제  
가 나와 있다. 각 샘플에 대한 총 세포 용해물을 SDS PAGE에 의해 분석하였다.

[0017]

도 3 - F13 및 E9 RS를 이용하여 상이한 세포주에서 인간 성장 호르몬 내 앰버 돌연변이의 억제가 나와 있다.

[0018]

## 정의

[0019]

본 발명을 상세히 기술하기 전에, 본 발명이 특별한 생물학적 시스템에 제한되지 않고, 그 시스템은 물론 다양  
할 수 있음을 이해하도록 한다. 또한, 본원에 사용된 용어는 단지 특별한 실시양태를 기술하기 위한 것으로 본  
발명의 범주를 제한하지 않고, 본 발명의 범주는 단지 첨부된 특허청구범위에 의해서만 제한된다는 것도 이해하  
도록 한다. 본 명세서와 청구범위에서 사용된 단수 형태의 표현은 달리 명시하지 않는 한, 복수의 표현을 포함  
한다. 따라서, 예를 들어, "세포"라고 하는 것은 2종 이상의 세포의 조합물을 포함하고, 당업자에게 공지된 그  
의 등가물 등을 포함한다. "세균"이라는 지칭은 세균들의 혼합물 등을 포함한다.

[0020]

여기 및 명세서 내 나머지 부분에 있어 달리 정의되지 않으면, 본원에 사용되는 모든 기술 용어 및 과학 용어는  
본 발명이 속하는 당업자에게 통상적으로 이해되는 것과 동일한 의미를 가진다.

[0021]

본원에 언급된 모든 공보 및 특허는 예를 들어, 본 발명과 연관되어 사용될 수 있는 공보에 기재되는 구성 및  
방법을 기재하고 개시할 목적으로 본원에 참고로 인용된다. 본원에서 논의된 공보는 오직 본 출원의 출원일 전  
그의 개시 내용에 대해서만 제공된다. 본 명세서 중 어떤 것도 본 발명자들이 선행 발명에 의해 또는 임의의 다  
른 이유로 이러한 개시 내용을 선행할 권리가 없음을 용인하는 것으로 해석되어서는 안될 것이다.

[0022]

상동: 단백질 및/또는 단백질 서열이 공통된 조상 단백질 또는 단백질 서열로부터 천연적으로 또는 비천연적으  
로 유도될 때, "상동"이다. 유사하게, 핵산 및/또는 핵산 서열은 공통된 조상 핵산 또는 핵산 서열로부터 천연

적으로 또는 비천연적으로 유도될 때, "상동"이다. 예를 들어, 임의의 천연 발생 핵산이 임의의 이용가능한 돌연변이유발 방법에 의해 하나 이상의 선택된 코돈을 포함하도록 변형될 수 있다. 이 돌연변이유발된 핵산은 발현될 때, 하나 이상의 선택된 아미노산, 예컨대 비천연 아미노산을 포함하는 폴리펩티드를 코딩한다. 돌연변이화 공정은 물론 부가적으로 하나 이상의 표준 코돈을 변경함으로써 결과 돌연변이 단백질에서 하나 이상의 표준 아미노산을 변화시킬 수 있다. 하나 이상의 표준 아미노산은 비천연 아미노산 또는 천연 아미노산으로 변화될 수 있다. 상동성은 일반적으로 2개 이상의 핵산 또는 단백질 (또는 그것의 서열) 간의 서열 유사성으로부터 유추된다. 상동성을 구축하는데 유용한 서열 간의 정확한 유사도 %는 문제의 핵산 및 단백질에 따라 다양하나, 상동성을 구축하는데 25% 정도로 낮은 서열 유사도가 통상 이용된다. 상동성을 구축하기 위해, 보다 높은 수준의 서열 유사도, 예컨대 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%, 또는 그 이상을 사용할 수 있다. 서열 유사도 %를 결정하는 방법(예컨대, 디폴트 매개변수를 이용하는 BLASTP 및 BLASTN)이 본원에 기재되고, 일반적으로 이용가능하다.

[0023] 오르소고날(orthogonal): 본원에 사용되는 용어 "오르소고날"은 관심 시스템(예컨대, 번역 시스템, 예컨대 세포)에 의해 감소된 효율로 사용되는 분자(예컨대, 오르소고날 tRNA(O-tRNA) 및/또는 오르소고날 아미노아실 tRNA 합성효소(O-RS))를 지칭한다. 오르소고날성은 관심 번역 시스템에서의 오르소고날 tRNA 및/또는 오르소고날 RS의 기능의 불능, 또는 감소된 효율, 예컨대 20% 미만의 효율, 10% 미만의 효율, 5% 미만의 효율, 또는 예컨대 1% 미만의 효율을 지칭한다. 예를 들어, 관심 번역 시스템에서의 오르소고날 tRNA는, 내인성 RS에 의한 내인성 tRNA의 아미노아실화에 비해, 감소되거나 0인 효율로 관심 번역 시스템의 임의의 내인성 RS에 의해 아미노아실화된다. 또 다른 예로서, 오르소고날 RS는 내인성 RS에 의한 내인성 tRNA의 아미노아실화에 비해, 감소되거나 0인 효율로 관심 번역 시스템의 임의의 내인성 tRNA를 아미노아실화된다. 제2 오르소고날 분자는 세포에 도입되어, 제1 오르소고날 분자와 함께 기능할 수 있다. 예를 들어, 오르소고날 tRNA/RS 쌍은 상응하는 tRNA/RS 내인성 쌍의 효율 대비, 일정 효율(예컨대, 약 50% 효율, 약 60% 효율, 약 70% 효율, 약 75% 효율, 약 80% 효율, 약 85% 효율, 약 90% 효율, 약 95% 효율, 약 99%, 또는 그 이상의 효율)로 세포 내에서 함께 기능하는 상보적 도입 성분을 포함한다. "오르소고날성의 향상"은 출발 물질 또는 천연 발생 tRNA 또는 RS에 비해 증진된 오르소고날성을 지칭한다.

[0024] 동족(cognate): 용어 "동족"은 함께 기능하는 성분들, 예컨대 tRNA 및 아미노아실-tRNA 합성효소를 지칭한다. 그 성분들은 또한 상보적인 것으로 칭해질 수 있다.

[0025] 우선적으로 아미노아실화하다: 용어 "우선적으로 아미노아실화하다"는, O-RS가 O-tRNA의 발생에 사용되는 출발 물질 또는 천연 발생 tRNA를 아미노아실화하는 O-RS에 비해, 선택된 아미노산, 예컨대 비천연 아미노산으로 O-tRNA를 아미노아실화하는 효율, 예컨대 약 70% 효율, 약 75% 효율, 약 80% 효율, 약 85% 효율, 약 90% 효율, 약 95% 효율, 또는 약 99%, 또는 그 이상의 효율을 지칭한다. 이어서, 비천연 아미노산을 고 적합도로, 예컨대 소정의 선택된 코돈에 대해 약 70% 초과 효율, 소정의 선택된 코돈에 대해 약 75% 초과 효율, 소정의 선택된 코돈에 대해 약 80% 초과 효율, 소정의 선택된 코돈에 대해 약 85% 초과 효율, 소정의 선택된 코돈에 대해 약 90% 초과 효율, 소정의 선택된 코돈에 대해 약 95% 초과 효율, 또는 소정의 선택된 코돈에 대해 약 99% 초과 효율로 성장 폴리펩티드 사슬에 혼입된다.

[0026] 선택된 코돈: 용어 "선택된 코돈"은 번역 공정에서 O-tRNA에 의해 인식되나, 내인성 tRNA에 의해서는 인식되지 않는 코돈을 지칭한다. O-tRNA 안티코돈 루프는 mRNA 상의 선택된 코돈을 인식하고, 폴리펩티드 내에 상기 위치에 그것의 아미노산, 예컨대 선택된 아미노산, 예컨대 비천연 아미노산을 혼입한다. 선택된 코돈은 예컨대, 닌센스 코돈, 예컨대 앰버, 오키 및 오팔 코돈을 포함하나 이에 한정되지 않는 종결 코돈, 4개 이상의 염기 코돈, 회귀 코돈, 천연 또는 비천연 염기 쌍에서 유래된 코돈 등을 포함할 수 있으나 이에 한정되지 않는다. 소정의 시스템에 있어, 선택된 코돈은 또한 천연 3개의 염기 코돈 중 하나를 포함할 수 있고, 여기에서 내인성 시스템은 상기 천연 3개의 염기 코돈을 전혀 (혹은 거의) 사용하지 않는다. 예를 들어, 이는 천연 3개의 염기 코돈을 인식하는 tRNA가 결핍된 시스템, 및/또는 천연 3개의 염기 코돈이 회귀 코돈인 시스템을 포함한다.

[0027] 억제자 tRNA: 억제자 tRNA는 예컨대 선택된 코돈에 대한 반응으로 폴리펩티드 사슬에 아미노산을 혼입하기 위한 기전을 제공함으로써, 소정의 번역 시스템 내 메신저 RNA(mRNA)의 판독(reading)을 변경하는 rRNA이다. 예를 들어, 억제자 tRNA는 종결 코돈, 4개 염기 코돈, 또는 회귀 코돈을 포함하나 이에 한정되지 않는 코돈을 통해 판독할 수 있다.

[0028] 억제 활성: 용어 "억제 활성"은 선택된 코돈을 통한 tRNA, 예컨대 억제자 tRNA의 판독 능력을 지칭한다. 활성은 대조군(예컨대, 동족 합성효소 결핍) 대비로 관찰된 활성의 백분율로서 표시될 수 있다.



- [0029] 번역 시스템: 용어 "번역 시스템"은 천연 발생 아미노산을 성장 폴리펩티드 사슬(단백질)에 혼입하는데 필요한 성분을 지칭한다. 번역 시스템의 성분은 예컨대 리보솜, tRNA, 합성효소, mRNA 등을 포함할 수 있다. 본 발명의 성분을 시험관 내 또는 생체 내 번역 시스템에 부가할 수 있다. 번역 시스템의 예는 비진핵생물 세포, 예컨대 세균(예컨대, 대장균), 진핵생물 세포, 예컨대 효모 세포, 포유류 세포, 식물 세포, 조류 세포, 진균류 세포, 곤충 세포, 및 무세포 번역 시스템, 예컨대 세포 용해물 등을 포함하나 이에 한정되지 않는다.
- [0030] 번역 시스템은 세포내 또는 무세포 시스템일 수 있고, 원핵생물 또는 진핵생물 시스템일 수 있다. 세포내 번역 시스템은, 원하는 핵산 서열을 mRNA에 전사하고, mRNA를 번역할 수 있는 전세포 제제(whole cell preparation), 예컨대 투과화된 세포 또는 세포 배양물을 포함하나 이에 한정되지 않는다. 무세포 번역 시스템은 상업적으로 구입가능하고, 많은 상이한 유형 및 시스템이 공지되어 있다. 무세포 시스템의 예는, 원핵생물 용해물, 예컨대 에세리키아 콜라이 용해물, 및 진핵생물 용해물, 예컨대 밀 배아 추출물, 곤충 세포 용해물, 토끼 망상적혈구 용해물, 토끼 난모세포 용해물 및 인간 세포 용해물을 포함하나 이에 한정되지 않는다. 진핵세포 추출물 또는 용해물은 수득되는 단백질이 당화, 인산화, 또는 기타 방식으로 변형될 때 바람직할 수 있는데, 그 이유는 많은 그러한 변형들이 단지 진핵생물 시스템에서 가능하기 때문이다. 이 추출물 및 용해물 중 일부는 상업적으로 구입 가능하다(프로메가(Promega)(미국 위스콘신주 매디슨 소재); 스트라타진(Stratagene)(미국 캘리포니아주 라졸라 소재); 아머샴(Amersham)(미국 일리노이즈주 알링턴 하이츠 소재); 기브코(GIBCO)/BRL(미국 뉴욕주 그랜드 아일랜드 소재). 막 추출물, 예컨대 분비 단백질을 번역하는데 유용한, 소포체막 함유의 견치 채장 추출물도 또한 이용가능하다.
- [0031] 재구성된 번역 시스템이 또한 사용될 수 있다. 용해물, 또는 개시 인자-1(IF-1), IF-2, IF-3( $\alpha$  또는  $\beta$ ), 신장 인자 T(EF-Tu), 또는 종결 인자와 같은 정제 번역 인자로 보충된 용해물의 조합물과 함께, mRNA를 단백질로 번역하기 위해, 정제된 번역 인자의 혼합물이 또한 성공적으로 사용되었다. 무세포 시스템은 또한 본원에 특별히 참고로 인용되는 [Current Protocols in Molecular Biology(F. M. Ausubel et al. editors, Wiley Interscience, 1993)]에 기재된 바와 같이, DNA가 시스템에 도입되어, mRNA로 전사되며, mRNA는 번역되는 결합된 전사/번역 시스템일 수 있다. 진핵생물 전사 시스템에서 전사된 RNA는 이중핵 RNA(hnRNA) 또는 5'-말단 캡(7-메틸 구아노신) 및 3'-말단 폴리 A 꼬리를 갖는 성숙의 mRNA의 형태일 수 있고, 후자는 특정 번역 시스템에서 유리할 수 있다. 예를 들어, 캡핑된 mRNA는 망상적혈구 용해물 시스템에서 고효율로 번역된다.
- [0032] 선택된 아미노산: 용어 "선택된 아미노산"은 임의의 원하는 천연 발생 아미노산 또는 비천연 아미노산을 지칭한다. 본원에 사용되는 용어 "비천연 아미노산" 또는 "비천연적으로 코딩된 아미노산"은 20가지 통상의 천연 발생 아미노산들, 또는 셀레노시스테인 또는 피롤리신 중 하나가 아닌, 임의의 아미노산, 변형 아미노산, 및/또는 아미노산 유사체를 지칭한다. 용어 "비천연적으로 코딩된 아미노산" 및 "비천연 아미노산"과 함께 동의로 사용될 수 있는 다른 용어들로는 "비천연 아미노산," "비천연 발생 아미노산" 및 그들의 다양한 하이픈으로 결합된 형식 및 하이픈으로 결합되지 않은 형식이 있다. 또한, 용어 "비천연적으로 코딩된 아미노산"은 천연적으로 코딩된 아미노산(20개의 통상의 아미노산, 또는 피롤리신 및 셀레노시스테인을 포함하나 이에 한정되지 않음)의 변형(예컨대, 번역후 변형)에 의해 발생하지만, 그 자체로는 번역 복합체에 의해 성장 폴리펩티드 사슬에 천연적으로 혼입되지 않는 아미노산을 포함하나 이에 한정되지 않는다. 그러한 비천연 발생 아미노산의 예는 N-아세틸글루코사미닐-L-세린, N-아세틸글루코사미닐-L-트레오닌 및 O-포스포티로신을 포함하나 이에 한정되지 않는다.
- [0033] ~로부터 유도된: 본원에 사용되는 용어 "~로부터 유도된"은 특정화된 분자 또는 유기체로부터 단리되거나 그로부터의 정보를 이용하여 제조된 성분을 지칭한다.
- [0034] 양성 선택 또는 선별 마커: 본원에 사용되는 용어 "양성 선택 또는 선별 마커"는, 존재할 경우, 예컨대 발현화되거나 활성화되는 등의 경우에, 양성 선택 마커가 없는 세포로부터 양성 선택 마커가 있는 세포를 동정하도록 하는 마커를 지칭한다.
- [0035] 음성 선택 또는 선별 마커: 본원에 사용되는 용어 "음성 선택 또는 선별 마커"는, 존재할 경우, 예컨대 발현화되거나 활성화되는 등의 경우에, (예컨대, 원하는 성질을 갖는 세포와 대비하여) 원하는 성질을 가지지 않는 세포를 동정하도록 하는 마커를 지칭한다.
- [0036] 리포터: 본원에 사용되는 용어 "리포터"는 관심 시스템의 표적 성분을 선택하는데 사용될 수 있는 성분을 지칭한다. 예를 들어, 리포터는 항생제 내성 또는 감도를 부여하는 단백질, 예컨대 효소( $\beta$ -락타마제, 클로람페니콜 아세틸트랜스퍼라제(CAT) 등을 포함하나 이에 한정되지 않음), 형광 선별 마커(녹색 형광 단백질 (예컨대, GFP), YFP, EGFP, RFP를 포함하나 이에 한정되지 않음), 발광 마커(개뿔벌레 루시페라제 단백질을 포함하나 이에 한정되지 않음), 친화성 기재 선별 마커, 또는 양성 또는 음성 선택성 마커 유전자, 예컨대 lacZ,  $\beta$ -

gal/lacZ( $\beta$ -갈락토시다제), ADH(알코올 탈수소효소), his3, ura3, leu2, lys2 등을 포함하나 이에 한정되지 않음)를 포함할 수 있다.

[0037] **진핵생물**: 본원에 사용되는 용어 "진핵생물"은 진핵생물 계통 도메인에 속하는 유기체, 예컨대 동물(포유류, 곤충류, 양서류, 조류를 포함하나 이에 한정되지 않음), 섬모충, 식물류(단자엽식물, 쌍자엽식물, 조류를 포함하나 이에 한정되지 않음), 진균류, 효모, 편모충, 마포자충류, 원생생물 등을 지칭한다.

[0038] **비진핵생물**: 본원에 사용되는 용어 "비진핵생물"은 진핵생물이 아닌 유기체를 지칭한다. 예를 들어, 비진핵생물 유기체는 진정세균(에셰리키아 콜라이), 썬무스 썬모필러스(*Thermus thermophilus*), 바실러스 스테아로썬모필러스(*Bacillus stearothermophilus*), 슈도모나스 플루오레센스(*Pseudomonas fluorescens*), 슈도모나스 아에루기노사(*Pseudomonas aeruginosa*), 슈도모나스 푸티다(*Pseudomonas putida*)를 포함하나 이에 한정되지 않음) 계통 도메인, 또는 고세균(메타노코커스 잔네쉬이(*Methanococcus jannaschii*), 메타노박테리움 썬모아우트로피쿰(*Methanobacterium thermoautotrophicum*), 호염성 세균(*Halobacterium*), 예컨대 할로페락스 볼카니(*Haloferax volcanii*) 및 호염성 세균종 NRC-1, 아르카에오글로부스 풀기두스(*Archaeoglobus fulgidus*), 파이로코커스 푸리오서스(*Pyrococcus furiosus*), 파이로코커스 호리코쉬이(*Pyrococcus horikoshii*), 아에우로피쿰 페르닉스(*Aeuropyrum pernix*)를 포함하나 이에 한정되지 않음) 계통 도메인에 속할 수 있다.

[0039] **보존적 변이체**: 용어 "보존적 변이체"는, 보존적 변이체의 기초가 되는 성분, 예컨대 O-tRNA 또는 O-RS와 같이 작용적으로 수행하나, 서열에 있어 변이체를 갖는 번역 성분, 예컨대 보존적 변이체 O-tRNA 또는 보존적 변이체 O-RS를 지칭한다. 예를 들어, O-RS는 선택된 아미노산, 예컨대 비천연 아미노산으로 상보적 O-tRNA 또는 보존적 변이체 O-tRNA를 아미노아실화하나, O-tRNA 및 보존적 변이체 O-tRNA는 동일한 서열을 가지지 않는다. 유사하게, tRNA는 상보적 O-RS 또는 보존적 변이체 O-RS에 의해 11가지 선택된 아미노산, 예컨대 비천연 아미노산으로 아미노아실화될 것이나, O-RS 및 보존적 변이체 O-RS는 동일한 서열을 가지지 않는다. 보존적 변이체는, 그 보존적 변이체가 상응하는 O-tRNA 또는 O-RS에 상보적인 한, 서열에 있어 1개 변이체, 2개 변이체, 3개 변이체, 4개 변이체, 또는 5개 이상의 변이체를 가질 수 있다.

[0040] **선택 또는 선별제**: 본원에 사용되는 용어 "선택 또는 선별제"는, 존재할 경우, 모집단으로부터의 특정 성분을 선택/선별하도록 하는 제제를 지칭한다. 예를 들어, 선택 또는 선별제는 예컨대 영양소, 항생제, 광 파장, 항체, 발현된 폴리뉴클레오티드 등을 포함하나 이에 한정되지 않는다. 선택제는 예컨대 농도, 강도 등(단, 이에 한정되지 않음)에 의해 다양화될 수 있다.

[0041] 용어 "효율적으로 인식되지 않음"은 한 유기체로부터의 RS가 O-tRNA를 아미노아실화하는 효율이, 예컨대 약 10% 미만, 약 5%, 미만, 또는 약 1% 미만임을 지칭한다.

[0042] **발명의 상세한 설명**

[0043] 하나 이상의 선택된 아미노산, 예컨대 비천연 아미노산을 포함하는 단백질의 제조에 적당한 번역 시스템이 U.S. 특허 출원 제10/126,931호(발명의 명칭: "METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGONAL tRNA-AMINOACYL tRNA SYNTHETASE PAIRS") 및 제10/126,927호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")에 기재되어 있다. 또한, USSN 제10/825,867호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")을 참고한다. 이 출원들 각각은 전체적으로 본원에 참고로 인용된다. 그러한 번역 시스템은 일반적으로, 오르소고날 tRNA(O-tRNA), 오르소고날 아미노아실 tRNA 합성효소(O-RS), 및 선택된 아미노산, 예컨대 비천연 아미노산을 포함하고, 여기에서 O-RS는 선택된 아미노산으로 O-tRNA를 아미노아실화한다. 본 발명의 오르소고날 쌍은 O-tRNA, 예컨대 억제자 tRNA, 프레임쉬프트(frameshift) tRNA 등, 및 O-RS로 이루어진다. O-tRNA는 제1 셀렉터 코돈을 인식하고, 셀렉터 코돈에 대한 반응으로 동족 합성효소의 존재 하에 억제 활성을 가진다. 세포는 선택된 아미노산을 성장 폴리펩티드 사슬에 혼입하는 성분을 이용한다. 예를 들어, 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산도 또한 존재할 수 있고, 여기에서 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함한다. 번역 시스템은 또한 시험관 내 시스템일 수 있다. 본 발명의 RS 분자는 번역 시에 리보솜을 이용하는 시스템을 포함한 임의의 번역 시스템에서 유용하다.

[0044] 번역 시스템은 또한 무세포(시험관 내) 번역 시스템일 수 있다. mRNA를 주형으로 포함하거나(시험관 내 번역), DNA를 주형으로 포함할 수 있는(시험관 내 전사 및 번역 조합) 이 시스템에서, 시험관 내 합성은 리보솜에 의해 지정된다. 무세포 단백질 발현 시스템의 개발을 위해 상당한 노력이 가해졌다. 예를 들어, 본원에 참고로 인용되는 [Kim, D.M. 및 J.R. Swartz, *Biotechnology and Bioengineering*, 74: 309-316 (2001); Kim, D.M. 및 J.R. Swartz, *Biotechnology Letters*, 22, 1537-1542, (2000); Kim, D.M., 및 J.R. Swartz, *Biotechnology*

*Progress*, 16, 385-390, (2000); Kim, D.M., 및 J.R. Swartz, *Biotechnology and Bioengineering*, 66, 180-188 (1999); 및 Patnaik, R. 및 J.R. Swartz, *Biotechniques* 24, 862-868 (1998); U.S. 특허 제6,337,191호; U.S. 특허 공보 제2002/0081660호; WO 00/55353; 및 WO 90/05785를 참고한다. 적용될 수 있는 또 다른 접근법은 mRNA-펩티드 융합 기법을 포함한다. 예를 들어, [R. Roberts 및 J. Szostak, *Proc. Natl Acad. Sci.(USA)* 94:12297-12302(1997); A. Frankel, et al, *Chemistry & Biology* 10:1043-1050(2003)]을 참고한다. 이 접근법에서, 퓨로마이신에 연결된 mRNA 주형은 리보솜 상에서 펩티드로 번역된다. 하나 이상의 tRNA 분자가 변형된 경우, 비천연 아미노산이 또한 펩티드에 혼입될 수 있다. 마지막 mRNA 코돈이 판독된 경우, 퓨로마이신은 펩티드의 C-말단을 차지한다. 수득되는 mRNA-펩티드 접합물이 시험관 내 검정에서 유리한 성질을 가지는 것으로 나타나는 경우, 그것의 실체는 mRNA 서열로부터 용이하게 밝혀질 수 있다. 이 방식으로, 하나 이상의 비천연적으로 코딩된 아미노산을 포함하는 폴리펩티드의 라이브러리를 선별하여, 원하는 성질을 갖는 폴리펩티드를 동정할 수 있다. 보다 최근, 정제된 성분을 갖는 시험관 내 리보솜 번역은, 비천연적으로 코딩된 아미노산으로 치환된 펩티드가 합성되도록 하는 것으로 보고되었다. 예컨대 [A. Forster et al, *Proc. Natl Acad. Sci.(USA)* 100: 6353 (2003)]을 참고한다.

[0045] 특정 실시양태에서, 본 발명의 RS를 포함하는 대장균 세포는 그러한 번역 시스템을 포함한다. 예를 들어, 본 발명의 대장균 세포는 오르소고날 tRNA(O-tRNA)(여기에서, O-tRNA는 셀렉터 코돈에 대한 반응으로 동족 합성효소의 존재 하에 억제 활성을 가짐); 오르소고날 아미노아실-tRNA 합성효소(O-RS); 선택된 아미노산; 및 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산(여기에서, 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 셀렉터 코돈을 포함함)을 포함한다.

[0046] 본 발명은 또한 하나 초과 선택된 아미노산이 혼입되도록 하는 세포 내 다중 O-tRNA/O-RS 쌍을 특징으로 한다. 특정 실시양태에서, 세포는 부가적 상이한 O-tRNA/O-RS 쌍 및 제2의 선택된 아미노산을 추가로 포함할 수 있고, 여기에서 O-tRNA는 제2 셀렉터 코돈을 인식하고, O-RS는 제2의 선택된 아미노산으로 O-tRNA를 우선적으로 아미노아실화한다. 예를 들어, 세포는 예컨대, 메타노코커스 잔네쉬이의 티로실-tRNA 합성효소로부터 유래된 엠버 억제자 tRNA-아미노아실 tRNA 합성효소 쌍을 추가로 포함할 수 있다.

[0047] O-tRNA 및/또는 O-RS는 천연적으로 발생하거나, 다양한 유기체들로부터 예컨대 tRNA의 라이브러리 및/또는 RS의 라이브러리를 발생시키는 천연 발생 tRNA 및/또는 RS의 돌연변이화에 의해 유래될 수 있다. 예를 들어, 오르소고날 tRNA/아미노아실-tRNA 합성효소 쌍을 생성시키는 한 기법은 이중 tRNA/합성효소 쌍을, 예컨대 숙주 세포 이외의 공급원, 또는 다중 공급원로부터 숙주 세포에 이동하는 것으로 포함한다. 이중 합성효소 후보물질의 성질은 예컨대 그것이 임의의 숙주 세포 tRNA를 충전하지 않음을 포함하고, 이중 tRNA 후보물질의 성질은, 예컨대 그것이 숙주 세포 합성효소에 의해 아미노아실화되지 않음을 포함한다. 또한, 이중 tRNA는 모든 숙주 세포 합성효소에 대해 오르소고날성을 가진다.

[0048] 오르소고날 쌍을 발생시키는 제2 기법은 O-tRNA 또는 O-RS를 선별 및/또는 선택하도록 하는데 이용되는 돌연변이 라이브러리를 발생시키는 것을 포함한다. 이 기법들은 또한 조합될 수 있다.

[0049] 각종 실시양태들에서, O-tRNA 및 O-RS는 하나 이상의 유기체로부터 유래된다. 또 다른 실시양태에서, O-tRNA는 천연 발생되거나 제1 유기체로부터의 돌연변이화 천연 발생 tRNA로부터 유래되고, O-RS는 천연 발생되거나 제2 유기체로부터의 돌연변이화 천연 발생 RS로부터 유래된다. 한 실시양태에서, 제1 및 제2 유기체는 상이하다. 예를 들어, 오르소고날 쌍은 메타노박테리움 써모오토트로피쿰(*Methanobacterium thermoautotrophicum*)에서 유래된 tRNA 합성효소, 및 고세균 tRNA(예컨대, 호염성(*Halobacterium*) 세균종 NRC-1)에서 유래된 tRNA를 포함할 수 있다. 대안적으로, 제1 및 제2 유기체는 동일하다. 부가적 정보를 위해 본원에서의 단락(제목: "공급원 및 숙주 유기체(Sources and Host Organisms)")을 참고한다.

[0050] 본 발명의 특정 실시양태에서, 본 발명의 O-RS는 서열 번호 4에 나와 있는 폴리뉴클레오티드 서열, 또는 그것의 상보적 폴리뉴클레오티드 서열, 또는 그것의 보존적 변이체를 포함하거나 그에 의해 코딩된다. 특정 실시양태에서, O-RS는 서열 번호 5에 나와 있는 아미노산 서열을 포함한다. 또한 본원에서의 단락(제목: "핵산 및 폴리펩티드 서열 및 변이체(Nucleic Acid and Polypeptide Sequence and Variants)")을 참고한다.

[0051] 오르소고날 tRNA(O-tRNA)

[0052] 오르소고날 tRNA(O-tRNA)는, 예컨대 생체 내 또는 시험관 내, O-tRNA에 의해 인식되는 셀렉터 코돈을 포함하는 폴리뉴클레오티드에 의해 코딩되는 단백질로 선택된 아미노산을 혼입하는 것을 매개한다. O-tRNA는 화학적 또는 효소적 아미노아실화를 포함하나 이에 한정되지 않는 임의의 방법 또는 기법에 의해 원하는 아미노산으로 아미



노아실화될 수 있다. 아미노아실화된 O-tRNA는 번역 시스템에 직접 부가될 수 있다. O-tRNA는 시험관 내 또는 생체 내, 선택된 아미노산을 이용하여 본 발명의 RS에 의해 아미노아실화될 수 있다. 또한, RS는 O-RS일 수 있다. O-tRNA는 직접적으로, 또는 O-tRNA 또는 그것의 일부분을 코딩하는 폴리뉴클레오티드를 제공함으로써 번역 시스템(예컨대, 시험관 내 번역 성분, 또는 세포)에 제공될 수 있다. 예를 들어, O-tRNA 또는 그것의 일부분은 서열 번호 1, 2, 3에 나와 있는 폴리뉴클레오티드 서열, 또는 그것의 상보적 폴리뉴클레오티드 서열, 또는 그것의 보존적 변이체에 의해 코딩된다. O-RS는 직접적으로(예컨대, 서열 번호 5 또는 17, 또는 그것의 보존적 변이체), 또는 O-RS 또는 그것의 일부분을 코딩하는 폴리뉴클레오티드를 제공함으로써 번역 시스템(예컨대, 시험관 내 번역 성분, 또는 세포)에 제공될 수 있다. 예를 들어, O-RS 또는 그것의 일부분은 서열 번호 4에 나와 있는 폴리뉴클레오티드 서열, 아미노산 서열의 서열 번호 17을 코딩하는 폴리뉴클레오티드 서열, 또는 그것의 상보적 폴리뉴클레오티드 서열, 또는 그것의 보존적 변이체에 의해 코딩된다.

[0053] 본 발명의 O-tRNA는 셀렉터 코돈에 대한 반응으로 동족 합성효소의 존재 하에 억제 활성을 가진다. 억제 활성은 당업계에서 공지된 수많은 검정들 중 임의의 검정에 의해 결정될 수 있다. 예를 들어,  $\beta$ -갈락토시다제 리포터 검정이 이용될 수 있다. 프로모터의 조절 하에 *lacZ* 유전자를 발현하는 플라스미드의 유도체를 사용하며, 예컨대 여기에서 *lacZ*의 펩티드 VVLQRRDWEN의 Leu-25는, 티로신, 세린, 루이신 등에 대한 센스 코돈(대조군) 또는 셀렉터 코돈, 예컨대 TAG, TGA, AGGA 등의 코돈으로 대체된다. 유도체화된 *lacZ* 플라스미드는 본 발명의 O-tRNA를 포함하는 플라스미드와 함께 적절한 유기체(예컨대, 오르소고날성분이 사용될 수 있는 유기체)로부터의 세포에 도입된다. 동족 합성효소도 또한 (발현될 때 동족 합성효소를 코딩하는 폴리뉴클레오티드 또는 폴리펩티드로서) 도입될 수 있다. 세포를 원하는 밀도, 예컨대 약 0.5의 OD<sub>600</sub>으로 매질 내에 성장시키고,  $\beta$ -갈락토시다제 검정을, 예컨대 베타플루오르(BetaFluor)<sup>TM</sup>  $\beta$ -갈락토시다제 검정 키트(노바겐(Novagen))를 이용하여 수행한다. 억제율 %을, 예컨대 유도체화된 *lacZ* 구축물로부터 관찰되는 값인 필적하는 대조군 대비 샘플의 활성도 %로서 계산하고, 여기에서 구축물을 원하는 위치에 셀렉터 코돈이 아닌 상응하는 센스 코돈을 가진다.

[0054] 본 발명에 사용하기에 적당한 O-tRNA의 예는 미국 특허 출원 제10/126,931호, 제10/126,927호, 및 제10/825,867호에 개시된 O-tRNA 분자들 중 어느 하나이다. tRNA 분자에서, 티민(T)은 우라실(U)로 대체된다. 또한, 염기에 대한 부가적 변형이 존재할 수 있다. 본 발명은 또한 O-tRNA의 보존적 변이체를 포함한다. 예를 들어, O-tRNA의 보존적 변이체는 O-tRNA와 같이 기능하고 tRNA L-형상 구조를 유지하나, 동일한 서열을 가지지 않는(또한 야생형 tRNA 분자가 아님) 분자들을 포함한다. 또한, 본원에서의 단락(제목: "핵산 및 폴리펩티드 서열 및 변이체(Nucleic Acid and Polypeptide Sequence and Variants)")를 참고한다.

[0055] O-tRNA를 포함하는 조성물은 또한 오르소고날 아미노아실-tRNA 합성효소(O-RS)를 추가로 포함할 수 있고, 여기에서 O-RS는 선택된 아미노산(예컨대, 비천연 아미노산)으로 O-tRNA를 우선적으로 아미노아실화한다. 특정 실시양태에서, O-tRNA를 포함하는 조성물은 번역 시스템(예컨대, 시험관 내 또는 생체 내 번역 시스템)을 추가로 포함할 수 있다. 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함하는 핵산(여기에서, 폴리뉴클레오티드는 O-tRNA에 의해 인식되는 하나 이상의 셀렉터 코돈, 또는 이들 중 하나 이상의 조합물을 포함함)이 또한 세포 또는 기타 번역 시스템 내에 존재할 수 있다. 또한, 본원에서의 단락(제목: "오르소고날 아미노아실-tRNA 합성효소(O-RS)(Orthogonal Aminoacyl-tRNA Synthetases(O-RS))")를 참고한다.

[0056] 오르소고날 tRNA(O-tRNA), 예컨대 O-tRNA의 생산 방법도 또한 본 발명의 한 특징이다. tRNA, 예컨대 상기 방법에 의해 생산된 O-tRNA도 또한 본 발명의 한 특징이다.

[0057] 오르소고날 tRNA의 생산 방법은, 셀렉터 코돈(예컨대, 앰버 코돈, 오팔 코돈, 4 염기 코돈 등)을 인식하도록 하는 tRNA의 각 풀의 안티코돈 루프를 돌연변이화함으로써, 복수개의 잠정적 O-tRNA를 제공하고; 복수개의 잠정적 O-tRNA의 구성원의 2차 구조를 분석하여, 2차 구조 내 비정규 염기쌍을 동정하고, 임의적으로 비정규 염기 쌍을 돌연변이화하는 것(예컨대, 비정규 염기 쌍은 정규 염기 쌍으로 돌연변이화됨)을 포함한다. 비정규 염기 쌍은 2차 구조의 줄기 영역 내에 위치할 수 있다. O-tRNA는 하나 이상의 특성 또는 활성의 향상, 예컨대 출발 물질, 예컨대 복수개의 tRNA 서열 대비, 원하는 유기체에 대한 오르소고날성의 향상을 가지면서, 한편 원하는 RS에 대한 친화성을 유지할 수 있다.

[0058] 대안적으로, O-tRNA는 번역에 영향을 주거나 번역 기구의 성분인 하나 이상의 분자와의 상호 작용 또는 그에 대한 결합 친화성을 조절하기 위해 공지된 tRNA를 돌연변이화함으로써 개발될 수 있다. 그러한 성분은 신장 인자를 포함하나 이에 한정되지 않는다. 세균 신장 인자 EF-Tu는 단백질 합성에서의 신장 단계에서 핵심적 역할을 한다. tRNA 합성효소에 의해 tRNA를 아미노아실화한 후에, EF-Tu는 아미노아실화된 tRNA에 결합하여, 그것을 리보솜의 A 부위로 보낸다. 충전된 아미노산과 tRNA 사이의 에스테르 결합은 EF-Tu와 아미노아실화된 tRNA 사이의

결합으로 인해 자발적 가수분해로부터 보호된다. Stortchevoi 등은, TΨC 줄기에서의 대장균 개시 tRNA<sup>fMet</sup> U50:G64 동요(wobble) 염기쌍의 돌연변이를 조사하였는데, 그 이유는 신장에 있어 tRNA의 활성을 막는 2차적 음성 결정자인 것으로 나타났기 때문이다. 이 염기 쌍이 추정컨대 EF-Tu.GTP와 아미노아실화된 tRNA 사이의 약화된 상호작용으로 인해 다(JBC 2003 278(20):17672-17679). 또한, LaRiviere 등은 [Science 2001 Oct 5;294(5540): 165-8]에서 EF-Tu에 대한 전반적 결합 친화성에 대한 아미노산 및 tRNA 체의 열역학적 기여를 기재하였다. 그들은 tRNA 체와 아미노산의 기여가 상호 독립적이며, tRNA가 올바르게 아실화될 때 그것들을 상호 상보함을 보여주었다. EF-Tu.GTP와 비천연 아미노산으로 아미노아실화된 tRNA 사이의 상호작용의 변경은 리보솜의 A 부위에 tRNA를 로딩하는 효율에 영향을 줄 수 있다. 잠정적 돌연변이 부위는 또한 tRNA와 번역 기구의 다른 성분, 예컨대 EF-Tu 사이의 착물의 결정 구조를 분석함으로써 밝혀질 수 있다. 예를 들어, Nissen 등은 EF-Tu.GTP가 효모 페닐알라닐-전이 RNA(Phe-tRNA)의 TΨC 줄기의 인산염 골격에 직접 결합함을 보여주었다(Science 1995 270(5241):1464-1472).

[0059] 방법은 임의적으로 tRNA 및/또는 아미노아실-tRNA 합성효소의 서열 상동성을 분석하여, 특정 유기체에 대해 오르소고날성인 것으로 보이는 O-tRNA, O-RS 및/또는 그들의 쌍에 대한 잠정적 후보물질을 결정하는 것을 포함한다. 당업계에 공지되고 본원에 기재된 컴퓨터 프로그램을 분석을 위해 사용할 수 있다. 원핵생물 유기체에 사용하기 위한 잠정적 오르소고날 번역 성분을 선택하기 위해, 한 예로서, 원핵생물 유기체에 특이적 상동성을 나타내지 않는 합성효소 및/또는 tRNA를 선택할 수 있다.

[0060] tRNA의 풀을 또한 일치 기법(consensus strategy)에 의해 생산할 수 있다. 예를 들어, 복수개의 tRNA 서열을 정렬하고; 일치 서열을 결정하며; 전체 일치 서열 중 적어도 일부, 대부분, 또는 전부를 이용하여 tRNA의 라이브러리를 발생시킴으로써 tRNA의 풀을 생산한다. 예를 들어, 일치 서열을 컴퓨터 프로그램, 예컨대 GCG 프로그램 *파일업(pileup)*을 이용하여 컴파일(compile)할 수 있다. 임의적으로, 프로그램에 의해 결정된 축퇴 위치를 그 위치들에서 가장 빈번한 염기로 바꾼다. 일치 서열을 이용하여 당업계에 공지된 기법을 이용하여 라이브러리를 합성한다. 예를 들어, tRNA 유전자의 각 부위가 90% 일치 서열 및 10% 다른 3개 염기의 혼합물로 된 도핑 혼합물로서 합성될 수 있는 올리고뉴클레오타이드의 중첩 연장을 사용하여, 일치 서열에 기초한 라이브러리를 제공할 수 있다. 다른 혼합물, 예컨대 75% 일치 서열 및 25% 다른 3개 염기의 혼합물; 80% 일치 서열 및 20% 다른 3개 염기의 혼합물; 95% 일치 서열 및 5% 다른 3개 염기의 혼합물 등을 또한 사용할 수 있다.

[0061] 돌연변이 tRNA의 라이브러리를 당업계에 공지된 각종 돌연변이유발 기법을 이용하여 발생시킬 수 있다. 예를 들어, 돌연변이 tRNA는 부위-특이적 돌연변이, 무작위 점 돌연변이, 상동 재조합, DNA 셔플링(shuffling) 또는 기타 재귀적 돌연변이유발, 키메라 구성, 또는 이들의 임의의 조합에 의해 발생될 수 있다.

[0062] 부가적 돌연변이는 tRNA의 원하는 루프 또는 영역, 예컨대 안티코돈 루프, 어셉터 줄기, D 팔(arm) 또는 루프, 가변성 루프, TΨC 팔 또는 루프, tRNA 분자의 다른 영역, 또는 이들의 조합 내 특정 위치(들), 예컨대 비보존적 위치(들) 또는 보존적 위치(들), 또는 무작위 위치(들), 이들의 조합에서 도입될 수 있다. 돌연변이는 줄기 영역에 매칭된 염기 쌍을 포함할 수 있다.

[0063] 전형적으로, O-tRNA는 제1 종의 세포 집단을 음성 선택에 적용함으로써 획득되고, 여기에서 세포는 복수개의 잠정적 O-tRNA의 한 구성원을 포함한다. 음성 선택은 세포에 대해 내인성인 아미노아실-tRNA 합성효소(RS)에 의해 아미노아실화된 복수개의 잠정적 O-tRNA의 한 구성원을 포함하는 세포를 제거한다. 이로써, 제1 종의 세포에 대해 오르소고날성인 O-tRNA의 풀이 제공된다.

[0064] 음성 선택의 특정 실시양태에서, 셀렉터 코돈(들)은 비필수 위치 등에서, 음성 선택 마커, 예컨대 항생제 내성을 부여하는 효소, 예컨대 β-락타마제, 검출가능한 산물을 부여하는 효소, 예컨대 β-갈락토시다제, 클로람페니콜 아세틸트랜스퍼라제(CAT), 예컨대 독성 산물, 예컨대 바나제(barnase)를 코딩하는 폴리뉴클레오타이드에 도입된다. 선별/선택은 선택제(예컨대, 암피실린과 같은 항생제)의 존재 하에 세포 집단을 성장시킴으로써 행해질 수 있다. 한 실시양태에서, 선택제의 농도가 변화된다.

[0065] 예를 들어, 억제자 tRNA의 활성을 측정하기 위해, 음성 선택 마커를 코딩하는 폴리뉴클레오타이드, 예컨대 β-락타마제(bla)를 위한 유전자에 도입된 넌센스 또는 프레임쉬프트 돌연변이와 같은 셀렉터 코돈의 생체 내 억제에 기초한 선택 시스템을 사용한다. 예를 들어, 특정 위치에 예컨대 TAG, AGGA 및 TGA를 갖는 폴리뉴클레오타이드 변이체, 예컨대 bla 변이체가 구축될 수 있다. 세포, 예컨대 세균은 이 폴리뉴클레오타이드를 이용하여 형질전환되고, 내인성 대장균 합성효소에 의해 효율적으로 충전될 수 없는 오르소고날 tRNA의 경우, 항생제 내성, 예컨대 암피실린 내성은 플라스미드 없이 형질전환된 세균의 상기 내성과 유사하거나 그 미만이다. tRNA가 오르소고날성이 아닌 경우, 또는 tRNA를 충전할 수 있는 이중 합성효소가 시스템 내에 동시 발현되는 경우, 높은 수준의

항생제 내성, 예컨대 암피실린 내성이 관찰된다. 플라스미드 없이 형질전환된 세포와 대략 동등한 항생제 농도로 LB 아가 플레이트 상에서 성장할 수 없는 세포, 예컨대 세균이 선택된다.

[0066] 독성 산물(예컨대, 리보뉴클레아제 바나제)의 경우, 복수개의 잠정적 tRNA의 한 구성원이 내인성 숙주, 예컨대 에셰리키아 콜라이 합성효소에 의해 아미노아실화될 때(즉, 그것이 숙주, 예컨대 에셰리키아 콜라이 합성효소에 대해 오르소고날성이 아님), 셀렉터 코돈이 억제되고, 생산된 독성 폴리뉴클레오티드 산물은 세포 사멸을 초래한다. 오르소고날 tRNA 또는 비작용성 tRNA를 제공하는(harboring) 세포가 생존한다.

[0067] 한 실시양태에서, 원하는 유기체에 대해 오르소고날성인 tRNA의 풀을 이어서 양성 선택에 적용하고, 여기에서 셀렉터 코돈을, 예컨대 약물 내성 유전자, 예컨대  $\beta$ -락타마제 유전자에 의해 코딩된 양성 선택 마커에 둔다. 양성 선택을 tRNA의 풀의 한 구성원을 코딩하거나 그것을 포함하는 폴리뉴클레오티드, 양성 선택 마커를 코딩하는 폴리뉴클레오티드, 및 동족 RS를 코딩하는 폴리뉴클레오티드를 포함하는 세포에 대해 수행한다. 이 폴리뉴클레오티드는 세포에서 발현되고, 세포는 선택제, 예컨대 암피실린의 존재 하에 성장된다. 이어서, 동시 발현된 동족 합성효소에 의해 아미노아실화하고, 이 셀렉터 코돈에 대한 반응으로 아미노산을 삽입하는 능력을 갖는 tRNA를 선택한다. 전형적으로, 이 세포는 비작용성 tRNA, 또는 관심 합성효소에 의해 효율적으로 인식될 수 없는 tRNA를 제공하는 세포에 비해 증진된 억제 효율을 나타낸다. 비작용성 tRNA, 또는 관심 합성효소에 의해 효율적으로 인식될 수 없는 tRNA를 제공하는 세포는 항생제에 대해 민감하다. 그러므로, (i) 내인성 숙주, 예컨대 에셰리키아 콜라이, 합성효소에 대한 기질이 아니고; (ii) 관심 합성효소에 의해 아미노아실화될 수 있으며; (iii) 번역에 있어 작용성인 tRNA이 양 선택에서 생존한다.

[0068] 상기 방법들에 있어 선택, 예컨대 양성 선택, 음성 선택, 또는 양성 및 음성 선택 모두의 엄격성은 임의적으로 변화될 수 있다. 예를 들어, 바나제가 극히 독성인 단백질이기 때문에, 음성 선택의 엄격성은 상이한 수의 셀렉터 코돈을 바나제 유전자에 도입하고/하거나 유동성 프로모터를 이용함으로써 조절될 수 있다. 또 다른 예에서, 선택 또는 선별제(예컨대, 암피실린)의 농도가 변화된다. 한 측면에서, 원하는 활성이 초기 실행(round) 동안 낮을 수 있기 때문에, 엄격성이 변화된다. 따라서, 덜 엄격한 선택 기준이 초기 실행 동안 적용되고, 더 엄격한 기준은 더 나중의 실행 단계에서 적용된다. 특정 실시양태에서, 음성 선택, 양성 선택, 또는 양성 및 음성 선택 모두는 다수회 반복될 수 있다. 다수의 상이한 음성 선택 마커, 양성 선택 마커, 또는 양성 및 음성 선택 마커 모두가 사용될 수 있다. 특정 실시양태에서, 양성 및 음성 선택 마커가 동일할 수 있다.

[0069] 다른 유형의 선택/선별을 오르소고날 번역 성분, 예컨대 O-tRNA, O-RS, 및 O-tRNA/O-RS 쌍을 생산하기 위해 본 발명에 이용될 수 있다. 예를 들어, 음성 선택 마커, 양성 선택 마커, 또는 양성 및 음성 선택 마커 모두가 적당한 반응물의 존재 하에 발광 반응을 촉매하거나 형광화하는 마커를 포함한다. 또 다른 실시양태에서, 마커의 산물은 형광 활성화 세포 분류법(FACS) 또는 발광에 의해 검출된다. 임의적으로, 마커는 친화성 기재 선별 마커를 포함한다. [Francisco, J. A., et al. (1993) *Production and fluorescence-activated cell sorting of Escherichia coli expressing afunctional antibody fragment on the external surface. Proc. Natl. Acad. Sci. U.S.A.* 90:10444-8]을 참고한다.

[0070] 재조합 오르소고날 tRNA의 생산을 위한 부가적 방법을 예컨대 미국 특허 출원 제10/126,931호(발명의 명칭: "METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGONAL tRNA-AMINOACYL tRNA SYNTHETASE PAIRS") 및 제10/126,127호(발명의 명칭: "In vivo Incorporation of Unnatural Amino Acids") 및 USSN 제10/825,867호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")에서 찾아볼 수 있다. 또한, [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 참고한다.

[0071] tRNA는 화학적 또는 효소적 아미노아실화를 포함하나 이에 한정되지 않는 임의의 방법 또는 기법에 의해, 원하는 아미노산으로 아미노아실화될 수 있다.

[0072] 아미노아실화는 아미노아실 tRNA 합성효소에 의해, 또는 리보자임을 포함하나 이에 한정되지 않는 다른 효소적 분자에 의해 달성될 수 있다. 용어 "리보자임"은 "촉매적 RNA"와 상호 혼용된다. Cech 및 동료 연구원들(Cech, 1987, *Science*, 236:1532-1539; McCorkle et al., 1987, *Concepts Biochem.* 64:221-226)은 촉매(리보자임)로서 작용할 수 천연 발생 RNA의 존재를 입증하였다. 그러나, 이 천연 RNA 촉매는 단지 절단 및 스플라이싱을 위한 리보핵산 기질에 대해 작용하는 것으로 나타났으나, 최근 리보자임의 인위적 진화의 발달은 촉매의 영역을 각종 화학 반응들로 확장시켰다. 연구는 자체의 (2')3'-말단에서의 아미노아실-RNA 결합을 촉매할 수 있는 RNA 분자(Illangakekare et al., 1995 *Science* 267:643-647), 및 한 RNA 분자에서 다른 한 RNA 분자로 아미노산을



이동시킬 수 있는 RNA 분자(Lohse et al., 1996, Nature 381:442-444)를 규명하였다.

- [0073] 본원에 참고로 인용되는 미국 특허 출원 공보 제2003/0228593호는 리보자임의 구축 방법, 및 그것의, 천연적으로 코딩된 아미노산 및 비천연적으로 코딩된 아미노산을 이용한 tRNA의 아미노아실화에 있어서의 용도를 기재한다. 리보자임을 포함하나 이에 한정되지 않는, tRNA를 아미노아실화할 수 있는 효소적 분자의 기질-고정화 형태는 아미노아실화 산물의 효율적 친화성 정제를 가능하게 할 수 있다. 적당한 기질의 예는 아가로스, 세파로스 및 자기 비이드를 포함한다. 아미노아실화를 위한 리보자임의 기질-고정화 형태의 생산 및 용도가, 본원에 참고로 인용되는 [Chemistry and Biology 2003, 10:1077-1084] 및 미국 특허 출원 공보 제2003/0228593호에 기재된다.
- [0074] 화학적 아미노아실화 방법은, 아미노아실화에 있어 합성효소의 사용을 피하기 위한, 본원에 참고로 인용되는 Hecht 및 동료 연구원들에 의해 도입되는 방법들[Hecht, S. M. Ace. Chem. Res. 1992, 25, 545; Heckler, T. G.; Roesser, J. R.; Xu, C; Chang, P.; Hecht, S. M. Biochemistry 1988, 27, 7254; Hecht, S. M.; Alford, B. L.; Kuroda, Y.; Kitano, S. J. Biol. Chem. 1978, 253, 4517] 및 Schultz, Chamberlin, Dougherty 등 (Cornish, V. W.; Mendel, D.; Schultz, P. G. Angew. Chem. Int. Ed. Engl. 1995, 34, 621; Robertson, S. A.; Ellman, J. A.; Schultz, P. G. J. Am. Chem. Soc. 1991, 113, 2722; Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C; Schultz, P. G. Science 1989, 244, 182; Bain, J. D.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R. J. Am. Chem. Soc. 1989, 111, 8013; Bain, J. D. et al. Nature 1992, 356, 537; Gallivan, J. P.; Lester, H. A.; Dougherty, D. A. Chem. Biol. 1997, 4, 740; Turcatti, et al. J. Biol. Chem. 1996, 271, 19991; Nowak, M. W. et al. Science, 1995, 268, 439; Saks, M. E. et al. J. Biol. Chem. 1996, 271, 23169; Hohsaka, T. et al. J. Am. Chem. Soc. 1999, 121, 34)]을 포함하나 이에 한정되지 않는다. 상기 방법 및 기타 화학적 아미노아실화 방법이 tRNA 분자의 아미노아실화를 위해 사용될 수 있다.
- [0075] 화학적으로 변형된 아미노아실-tRNA를 이용하는 생합성 방법을 사용하여 수가지 생체물리적 프로브를 시험관 내에서 합성된 단백질에 혼입시켰다. 하기 공보 및 여기에 인용된 참고 문헌을 참고한다: [Brunner, J. *New Photolabeling and crosslinking methods*, Annu. Rev Biochem, 62: 483-514(1993); 및 Krieg, U. C., Walter, P., Hohnson, A. E. *Photocrosslinking of the signal sequence of nascent preprolactin of the 54-kilodaton polypeptide of the signal recognition particle*, Proc. Natl. Acad. Sci, 83(22): 8604-8608(1986)].
- [0076] 과거, 화학적으로 아미노아실화된 억제자 tRNA를 원하는 앰버 넌센스 돌연변이를 갖는 유전자로 프로그래밍된 단백질 합성 반응에 첨가함으로써, 비천연 아미노산이 시험관 내에서 부위-특이적으로 혼입될 수 있다는 점이 밝혀졌다. 이러한 접근법을 사용하여, 특정 아미노산에 대해 영양요구성인 균주를 사용하여 다수의 통상의 20가지 아미노산을 근접한 구조적 동족체, 예를 들어 페닐알라닌의 경우에는 플루오로페닐알라닌으로 치환시킬 수 있다. 예를 들어, [Noren, C. J., Anthony-Cahill, Griffith, M. C., Schultz, P. G. *A general method for site-specific incorporation of unnatural amino acid into proteins*, Science, 244: 182-188(1989); M. W. Nowak, et al., Science 268: 439-42(1995); Bain, J. D., Glabe, C. G., Dix, T. A., Chamberlin, A. R., Diala, E. S. *Biosynthetic site-specific Incorporation of a nonnatural amino acid into a polypeptide*, J. Am Chem Soc, 111: 8013-8014(1989); N. Budisa et al., FASEB J. 13:41-51(1999); Ellman, J. A., Mendel, D., Anthony-Cahill, S., Noren, C. J., Schultz, P. G. *Biosynthetic method for introducing unnatural amino acid site-specifically into proteins*, Methods in Enz., 301-336(1992); 및 Mendel, D., Cornish, V. W. & Schultz, P. G. *Site-Directed Mutagenesis with an Expanded Genetic Code*, Annu RevBiophys. Biomol Struct. 24, 435-62(1995)]을 참고한다.
- [0077] 예를 들어, 종결 코돈 UAG를 인식하고, 비천연 아미노산으로 화학적으로 아미노아실화된 억제자 tRNA를 제조하였다. 통상적인 부위-지정 돌연변이유발을 사용하여, 단백질 유전자의 관심 부위에서 종결 코돈 TAG를 도입하였다. 예를 들어, [Sayers, J.R., Schmidt, W. Eckstein, F. 5'-3' *Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis*, Nucleic Acids Res, 16(3): 791-802(1988)]을 참고한다. 아실화된 억제자 tRNA 및 돌연변이체 유전자가 시험관 내 전사/번역 시스템에서 합쳐진 경우, 특수 위치에서 그 아미노산을 함유하는 단백질을 제공하는 UAG 코돈에 대한 반응으로 비천연 아미노산이 혼입되었다. [<sup>3</sup>H]-Phe를 사용한 실험 및 α-히드록시산을 사용한 실험은 오직 원하는 아미노산이 UAG 코돈에 의해 특정화된 위치에 혼입되고, 이 아미노산이 단백질 중의 임의의 다른 부위에서는 혼입되지 않는다는 점을 입증하였다. 예를 들어, [Noren, et al, 이하 동문: Kobayashi et al., (2003) Nature Structural Biology 10(6): 425-432; 및 Ellman, J. A.,

Mendel, D., Schultz, P. G. *Site-specific incorporation of novel backbone structures into proteins*, *Science*, 255(5041): 197-200(1992)]을 참고한다.

- [0078] 촉매 RNA를 발생시키는 방법은 무작위화 리보자임 서열의 분리된 풀을 발생시키고, 풀에 지정 진화를 수행하며, 바람직한 아미노아실화 활성을 위한 풀을 선별하며, 요망되는 아미노아실화 활성을 나타내는 리보자임의 서열을 선택하는 것을 포함할 수 있다.
- [0079] 리보자임은 아실화 활성을 촉진하는 모티프 및/또는 영역, 예컨대 GGU 모티프 및 U-풍부 영역을 포함할 수 있다. 예를 들어, U-풍부 영역은 아미노산 기질의 인식을 용이하게 할 수 있고, GGU-모티프는 tRNA의 3' 말단과 염기 쌍을 형성할 수 있음이 보고되었다. 이와 함께, GGU-모티프 및 U-풍부 영역은 아미노산 및 tRNA를 함께 동시에 인식할 수 있도록 촉진하고, 이로써 tRNA의 3' 말단의 아미노아실화를 촉진한다.
- [0080] 리보자임은 tRNA<sup>Asn</sup><sub>cccg</sub>과 접합된 부분 무작위화 r24mini를 이용한 시험관 내 선택에 의해, 또한 그 후 활성 클론에서 나타나는 일치 서열의 계통적 공학처리에 의해 발생될 수 있다. 이 방법에 의해 수득되는 한 예시적 리보자임은 "Fx3 리보자임"으로 칭해지고, 이는 내용이 본원에 참고로 인용되는 미국 특허출원 공보 제2003/0228593호에 기재되어 있으며, 동족 비천연 아미노산으로 충전된 각종 아미노아실-tRNA의 합성을 위한 다방면 촉매로서 작용한다.
- [0081] 기질 상에서의 고정화를 사용하여, 아미노아실화된 tRNA의 효율적 진화성 정제를 가능하게 할 수 있다. 적당한 기질의 예는 아가로스, 세파로스 및 자기 비이드를 포함하나 이에 한정되지 않는다. 리보자임은 RNA의 화학적 구조를 이용함으로써 수지 상에 고정될 수 있으며, 예컨대 RNA의 리보스 상의 3'-시스-디올은 과요오드산염으로 산화되어 상응하는 디알데히드를 생성시켜, 수지 상에서의 RNA의 고정을 촉진한다. 환원 아민화가 수지와 리보자임 간의 상호작용으로 비가역적 연결기를 만드는 저비용의 히드라지드 수지를 포함한 각종 유형의 수지들이 사용될 수 있다. 아미노아실-tRNA의 합성은 상기 작동(on)-칼럼 아미노아실화 기법에 의해 상당히 촉진될 수 있다. [Kourouklis et al. Methods 2005; 36:239-4]는 칼럼 기재 아미노아실화 시스템을 기재한다.
- [0082] 아미노아실화된 tRNA의 단리는 다양한 방식으로 달성될 수 있다. 한 적당한 방법은 10 mM EDTA를 갖는 아세트산 나트륨 용액과 같은 완충액, 50 mM N-(2-히드록시에틸)피페라진-N'-(3-프로판술포산) 함유 완충액, 12.5 mM KCl(pH 7.0), 10 mM EDTA, 또는 단순히 EDTA 완충수(pH 7.0)를 갖는 칼럼으로부터 아미노아실화된 tRNA를 용리하는 것이다.
- [0083] 아미노아실화된 tRNA를 번역 반응에 부가하여, tRNA가 번역 반응에 의해 생산된 폴리펩티드에서 선택 위치에 아미노아실화하는데 이용되는 아미노산을 혼입할 수 있다. 아미노아실화된 tRNA가 사용될 수 있는 번역 시스템의 예는 세포 용해물을 포함할 수 있으나 이에 한정되지 않는다. 세포 용해물은 투입 mRNA로부터 폴리펩티드를 생체 내 번역하는데 필요한 반응 성분들을 제공한다. 그러한 반응 성분의 예는 리보솜형 단백질, rRNA, 아미노산, tRNA, GTP, ATP, 번역 개시 및 신장 인자, 및 번역 관련 부가적 인자들을 포함하나 이에 한정되지 않는다. 부가적으로, 번역 시스템은 배치 번역 또는 구획화 번역일 수 있다. 배치 번역 시스템은 단일 구획에서 반응 성분들을 조합하나, 구획화 번역 시스템은 번역 효율을 억제할 수 있는 반응 산물로부터 번역 반응 성분을 분리한다. 그러한 번역 시스템은 상업적으로 구입 가능하다.
- [0084] 또한, 결합된 전사/번역 시스템이 사용될 수 있다. 결합된 전사/번역 시스템은 투입 DNA의 상응 mRNA로의 양 번역을 허용하고, 상기 mRNA는 이후 반응 성분에 의해 번역되게 된다. 상업적으로 구입 가능한 결합된 전사/번역의 예는 급속 번역 시스템(RTS, 로쉐 인코포레이티드(Roche Inc.))이다. 그 시스템은 리보솜 및 번역 인자와 같은 번역 성분을 제공하기 위한 대장균 용해물을 함유하는 혼합물을 포함한다. 부가적으로, RNA 중합효소는 번역 시 사용하기 위한 mRNA 주형으로의 유입 DNA의 전사를 위해 포함된다. RTS는 공급/배출 구획과 전사/번역 구획을 포함한 반응 구획들 사이에 삽입된 막에 의해 반응 성분을 구획화하는 것을 이용할 수 있다.
- [0085] tRNA의 아미노아실화는 트랜스퍼라제, 중합효소, 촉매 항체, 다작용 단백질 등을 포함하나 이에 한정되지 않는 기타 제제에 의해 수행될 수 있다.
- [0086] 오르소고날 아미노아실-tRNA 합성효소(O-RS)
- [0087] 본 발명의 O-RS는 시험관 내 또는 생체 내에서, 선택된 아미노산으로 O-tRNA를 우선적으로 아미노아실화한다. 본 발명의 O-RS는, O-RS를 포함하는 폴리펩티드 및/또는 O-RS 또는 그것의 일부분을 코딩하는 폴리펩티드에 의해 번역 시스템(예컨대, 시험관 내 번역 성분, 또는 세포)에 제공될 수 있다. 예를 들어, O-RS, 또는 그것의 일부분은 서열 번호 4에 나와 있는 폴리뉴클레오티드 서열, 또는 그것의 상보적 폴리뉴클레오티드 서열, 또는 그

것의 보존적 변이체에 의해 코딩된다. 본 발명의 O-RS는 본원에 기재된 것들을 포함하나 이에 한정되지 않는 수많은 상이한 O-tRNA 분자들을 아미노아실화할 수 있다.

[0088] 오르소고날 아미노아실-tRNA 합성효소(RS), 예컨대 O-tRNA, 예컨대 O-tRNA와 함께 사용하기 위한 O-RS를 동정하는 방법도 또한 본 발명의 한 특징이다. 예를 들어, 방법은 제1 종의 세포 집단을 양성 선택에 적용하는 것을 포함하고, 여기에서 세포는 각기 1) 복수개의 아미노아실-tRNA 합성효소(RS)의 한 구성원(여기에서, 복수개의 RS는 돌연변이 RS, 제1 종 이외의 종으로부터 유래된 RS, 또는 돌연변이 RS 및 제1 종 이외의 종으로부터 유래된 RS 제1 종의 양자 모두; 2) 제2 종으로부터의 오르소고날 tRNA(O-tRNA); 및 3) 양성 선택 마커를 코딩하고, 하나 이상의 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 포함한다. 복수개의 RS의 구성원이 결핍되거나 그 양이 감소된 세포에 비해, 증진된 억제 효율을 나타내는 세포를 선택하거나 선별한다. 억제 효율이 증진된 세포는 O-tRNA를 아미노아실화하는 활성 RS를 포함한다. 제1 종으로부터의 제1 세트의 tRNA의 활성 RS에 의한 (시험관 내 또는 생체 내) 아미노아실화 수준을, 제2 종으로부터의 제2 세트의 tRNA의 활성 RS에 의한 (시험관 내 또는 생체 내) 아미노아실화 수준과 비교한다. 아미노아실화 수준은 검출가능한 물질(예컨대, 표지화된 아미노산 또는 비천연 아미노산)에 의해 결정될 수 있다. 제1 세트의 tRNA에 비해 제2 세트의 tRNA를 더욱 효율적으로 아미노아실화하는 활성 RS를 선택하고, 이로써 O-tRNA와 함께 사용하기 위한 오르소고날 아미노아실-tRNA 합성효소가 제공된다. O-RS, 예컨대 상기 방법에 의해 동정되는 O-RS도 또한 본 발명의 한 특징이다.

[0089] 수많은 검정들 중 임의의 검정을 이용하여, 아미노아실화를 결정할 수 있다. 이 검정은 시험관 내 또는 생체 내 수행될 수 있다. 예를 들어, 시험관 내 아미노아실화 검정이 예컨대 [Hoben, P., 및 Soil, D. (1985) *Methods Enzymol.* 113:55-59] 및 미국 특허 출원 공보 제2003/0228593호에 기재되어 있다. 아미노아실화는 또한 오르소고날 번역 성분과 함께 리포터를 사용하고, 단백질을 코딩하는 하나 이상의 셀렉터 코돈을 포함하는 폴리뉴클레오티드를 발현하는 세포 내 리포터를 검출함으로써 결정될 수 있다. 또한, 미국 특허 출원 제10/126,927호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS") 및 USSN 제10/825,867호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")를 참고한다.

[0090] 동정된 O-RS를 추가 조작하여, 통상의 20가지 아미노산 중 임의의 것이 아닌, 오직 원하는 비천연 아미노산만을 O-tRNA에 충전하는 합성 효소의 기질 특이성을 변경시킬 수 있다. 비천연 아미노산에 대한 기질 특이성을 갖는 오르소고날 아미노아실 tRNA 합성효소를 발생시키는 방법은, 합성효소를 예컨대 합성효소 내 활성 부위, 합성효소 내 편집 기전 부위, 합성효소의 상이한 도메인의 조합에 의한 상이한 부위 등에서 돌연변이화하고, 선택 공정을 적용하는 것을 포함한다. 양성 선택 후에 음성 선택이 조합되는 것에 기초하는 기법이 사용된다. 양성 선택에서, 양성 마커의 비필수 위치(들)에 도입된 셀렉터 코돈의 양성 억제는 세포가 양성 선택 압력 하에 생존하도록 한다. 천연 및 비천연 아미노산의 양자 모두의 존재 하에, 생존자는 그에 따라 천연 또는 비천연 아미노산으로 오르소고날 억제자 tRNA를 충전하는 활성 합성효소를 코딩한다. 음성 선택에서, 음성 마커의 비필수 위치(들)에 도입된 셀렉터 코돈의 억제는 천연 아미노산 특이성을 갖는 합성효소를 제거한다. 음성 및 양성 선택의 생존자는 비천연 아미노산만으로 오르소고날 억제자 tRNA를 아미노아실화(충전)하는 합성 효소를 코딩한다. 이어서, 이 합성효소는 추가적 돌연변이유발, 예컨대 DNA 서플링 또는 기타 재귀적 돌연변이유발 방법에 적용될 수 있다.

[0091] 돌연변이 O-RS의 라이브러리를 당업계에 공지된 각종 돌연변이유발 기법을 이용하여 발생시킬 수 있다. 예를 들어, 돌연변이 O-RS는 부위-특이적 돌연변이, 무작위 점 돌연변이, 상동 재조합, DNA 서플링 또는 기타 재귀적 돌연변이유발, 키메라 구성, 또는 이들의 임의의 조합에 의해 발생될 수 있다. 예를 들어, 돌연변이 RS의 라이브러리는 2개 이상의 다른, 예컨대 보다 작거나, 보다 덜 다양한 "서브라이브러리"로부터 생산될 수 있다. RS의 키메라 라이브러리도 또한 본 발명에 포함된다. 각종 유기체(예컨대, 진정세균 또는 고세균과 같은 미생물)로부터의 tRNA 합성효소의 라이브러리, 예컨대 천연 다양성을 갖는 라이브러리(예컨대, 미국 특허 제6,238,884호 (Short et al); 미국 특허 제5,756,316호(Schallenger et al); 미국 특허 제5,783,431호(Petersen et al); 미국 특허 제5,824,485호(Thompson et al); 미국 특허 제5,958,672호(Short et al) 참고)가 임의적으로 구축되고, 오르소고날 쌍에 대해 선별됨을 주목하도록 한다.

[0092] 일단 합성효소를 양성 및 음성 선택/선별 기법에 적용하면, 이에 이 합성효소를 추가 돌연변이유발에 적용할 수 있다. 예를 들어, O-RS를 코딩하는 핵산을 분리할 수 있고; 돌연변이화된 O-RS를 (예컨대, 무작위 돌연변이유발, 부위-특이적 돌연변이유발, 재조합 또는 이들의 임의의 조합에 의해) 코딩하는 한 세트의 폴리뉴클레오티드를 핵산으로부터 발생될 수 있고; 이 개별 단계 또는 이 단계들의 조합을, 비천연 아미노산으로 O-tRNA를 우선적으로 아미노아실화하는 돌연변이화된 O-RS가 수득될 때까지 반복할 수 있다. 본 발명의 한 측면에

서, 단계들을 다수회, 예컨대 2회 이상 수행한다.

[0093] 부가적 선택/선별 엄격성의 수준을 또한 본 발명의 O-tRNA, O-RS, 또는 그들의 쌍의 생산 방법에 이용할 수 있다. 선택 또는 선별 엄격성은 O-RS의 생산 방법의 한 단계 또는 양 단계 모두에서 다양화될 수 있다. 이는 예컨대 선택/선별제의 사용량 등의 다양화를 포함할 수 있다. 부가적 실행 회수의 양성 및/또는 음성 선택을 또한 수행할 수 있다. 선택 또는 선별은 또한 예컨대 아미노산 투과도의 변화, 번역 효율의 변화, 번역 신뢰도의 변화 등을 포함하는 하나 이상의 양성 또는 음성 선택 또는 선별을 포함할 수 있다. 전형적으로, 하나 이상의 변화는 오르소고날 tRNA-tRNA 합성효소 쌍을 사용하여 단백질을 생산하는 유기체 내 하나 이상의 유전자에서의 돌연변이에 기초한다.

[0094] 예컨대 O-RS, O-tRNA 및 O-tRNA/O-RS 쌍에 대해 본 발명에서 다른 유형의 선택을 또한 이용할 수 있다. 양성 선택 마커는 성장을 위한 영양 보충물을 제공하는 산물을 포함하나 이에 한정되지 않는 다양한 분자들 중 임의의 것일 수 있고, 선택을 영양적 보충물이 결여된 매질에서 수행할 수 있다. 양성 선택 마커를 코딩하는 폴리뉴클레오티드의 예는 예컨대 세포의 아미노산 보조영양의 보충에 기초한 리포터 유전자, his3 유전자(예컨대, 여기에서 his3 유전자는 3-아미노트리아졸(3-AT)을 제공함으로써 검출되는 이미다졸 인산염 테히드라타제를 코딩함), ura3 유전자, leu2 유전자, lys2 유전자, lacZ 유전자, adh 유전자 등을 포함하나 이에 한정되지 않는다. 예컨대 [G. M. Kishore, & D. M. Shah (1988), Amino acid biosynthesis inhibitors as herbicides, Annual Review of Biochemistry 57:627-663]를 참고한다. 한 실시양태에서, lacZ 생산은 오르토-니트로페닐-β-D-갈락토피라노시드(ONPG) 가수분해에 의해 검출된다. 예컨대 [I. G. Serebriiskii, & E. A. Golemis, (2000), Uses of lacZ to study gene function: evaluation of beta-galactosidase assays employed in the yeast two-hybrid system, Analytical Biochemistry 285:1-15]를 참고한다. 부가적 양성 선택 마커는, 예컨대 루시페라제, 녹색 형광 단백질(GFP), YFP, EGFP, RFP, 항생제 내성 유전자의 산물(예컨대, 클로람페니콜 아세틸 트랜스퍼라제(CAT)), 전사 조절자 단백질(예컨대, GAL4) 등을 포함한다. 임의적으로, 양성 선택 마커를 코딩하는 폴리뉴클레오티드는 셀렉터 코돈을 포함한다.

[0095] 양성 선택 마커를 코딩하는 폴리뉴클레오티드는 반응 요소(response element)에 작용적으로 연결될 수 있다. 반응 요소로부터의 전사를 조절하는 전사 조절자 단백질을 코딩하고, 하나 이상의 셀렉터 코돈을 포함하는 한 부가적 폴리뉴클레오티드도 또한 존재할 수 있다. 비천연 아미노산으로 아미노아실화된 O-tRNA에 의한 전사 조절자 단백질로의 비천연 아미노산의 혼입은 양성 선택 마커를 코딩하는 폴리뉴클레오티드(예컨대, 리포터 유전자)의 전사를 초래한다. 임의적으로, 셀렉터 코돈은 전사 조절자 단백질의 DNA 결합 도메인을 코딩하는 폴리뉴클레오티드의 일부에, 또는 그것의 실질적 부근에 위치한다.

[0096] 음성 선택 마커를 코딩하는 폴리뉴클레오티드가 또한 반응 요소에 작용적으로 연결될 수 있고, 그 반응 요소로부터의 전사는 전사 조절자 단백질에 의해 매개된다. 예컨대 [A. J. DeMaggio, et al., (2000), The yeast split-hybrid system, Method Enzymol. 328:128-137; H. M. Shih, et al. (1996), A positive genetic selection for disrupting protein-protein interactions: identification of CREB mutations that prevent association with the coactivator CBP, Proc. Natl. Acad. Sci. U.S.A. 93:13896-13901; M. Vidal, et al. (1996), Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system, Proc. Natl. Acad. Sci. U.S.A. 93:10321- 10326; 및 M. Vidal, et al. (1996), Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions (Proc. Natl. Acad. Sci. U.S.A. 93:10315-10320)]를 참고한다. 천연 아미노산으로 아미노아실화된 O-tRNA에 의한 전사 조절자 단백질로의 천연 아미노산의 혼입은 음성 선택 마커의 전사를 초래한다. 임의적으로, 음성 선택 마커는 셀렉터 코돈을 포함한다. 본 발명의 양성 선택 마커 및/또는 음성 선택 마커는 2개 이상의 셀렉터 코돈을 포함할 수 있고, 이의 각각 또는 양자 모두는 2개 이상의 상이한 셀렉터 코돈 또는 2개 이상의 동일한 셀렉터 코돈을 포함할 수 있다.

[0097] 전사 조절자 단백질은 핵산 서열(예컨대, 반응 요소)에 (직접적으로 또는 간접적으로) 결합하고, 반응 요소에 작용적으로 연결된 서열의 전사를 조절하는 분자이다. 전사 조절자 단백질은 전사 활성화자 단백질(예컨대, GAL4, 핵 호르몬 수용체, API, CREB, LEF/tcf 부류 구성원, SMAD, VP16, SP1 등), 전사 억제자 단백질(예컨대, 핵 호르몬 수용체, Groucho/tle 부류, 잉그레일드(Engrailed) 부류 등), 또는 환경에 따라 양 활성화 모두를 가질 수 있는 단백질(예컨대, LEF/tcf, 호모박스(homobox) 단백질 등)일 수 있다. 반응 요소는 전형적으로 전사 조절자 단백질에 의해 인식되는 핵산 서열, 또는 전사 조절자 단백질과 함께 작용하는 부가적 체제이다.

[0098] 전사 조절자 단백질의 또 다른 예는 전사 활성화자 단백질, GAL4이다. 예컨대 [A. Laughon, et al. (1984),



Identification of two proteins encoded by the *Saccharomyces cerevisiae* GAL4 gene, *Molecular & Cellular Biology* 4:268- 275; A. Laughon, & R. F. Gesteland (1984), Primary structure of the *Saccharomyces cerevisiae* GAL4 gene, *Molecular & Cellular Biology* 4:260-267; L. Keegan, et al. (1986), Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein, *Science* 231:699-704; 및 M. Ptashne (1988), How eukaryotic transcriptional activators work, *Nature* 335:683-689]를 참고한다. 이 881 아미노산 단백질의 N-말단 147 아미노산은 DNA 서열에 특이적으로 결합하는 DNA 결합 도메인(DBD)을 형성한다. 예컨대, [M. Carey, et al. (1989), An amino-terminal fragment of GAL4 binds DNA as a dimer, *J. Mol. Biol.* 209:423-432; 및 E. Giniger, et al. (1985), Specific DNA binding of GAL4, a positive regulatory protein of yeast, *Cell* 40:767-774]를 참고한다. DBD는 간접 단백질 서열에 의해, DNA에 결합될 때 전사를 활성화할 수 있는 C-말단 113 아미노산 활성화 도메인(AD)에 연결된다. 예컨대, [J. Ma, & M. Ptashne (1987), Deletion analysis of GAL4 defines two transcriptional activating segments, *Cell* 48:847-853; 및 J. Ma, & M. Ptashne (1987), The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80, *Cell* 50:137-142]를 참고한다. 예컨대, GAL4의 N-말단 DBD 및 그것의 C-말단 AD 모두를 함유하는 단일 폴리펩티드의 N-말단 DBD 측으로 앰버 코돈을 두으로써, O-tRNA/O-RS 쌍에 의한 앰버 억제제는 GAL4에 의한 전사 활성화에 연결될 수 있다. GAL4 활성화 리포터 유전자를 사용하여, 그 유전자로 양성 및 음성 선택 모두를 수행할 수 있다.

[0099] 음성 선택을 위해 사용되는 매질은 음성 선택 마커에 의해 검출가능한 물질로 전환되는 선택 또는 선별제를 포함할 수 있다. 본 발명의 한 측면에서, 검출가능한 물질은 독성 물질이다. 음성 선택 마커를 코딩하는 폴리뉴클레오티드는 예컨대 *ura3* 유전자일 수 있다. 예를 들어, URA3 리포터를 GAL4 DNA 결합 부위를 함유하는 프로모터의 조절 하에 둘 수 있다. 음성 선택 마커를, 예컨대 셀렉터 코돈으로 GAL4를 코딩하는 폴리뉴클레오티드의 번역에 의해 생산할 때, GAL4는 URA3의 전사를 활성화한다. 음성 선택은, *ura3* 유전자의 유전자 산물에 의해 검출가능한 물질(예컨대, 세포를 사멸시키는 독성 물질)로 전환되는 5-플루오로오로트산(5-fluoroorotic acid)(5-FOA)를 포함하는 매질 상에서 달성된다. 예컨대 [J. D. Boeke, et al. (1984), Boeke, et al. (1984), A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoroorotic acid resistance, *Molecular & General Genetics* 197:345-346; M. Vidal, et al. (1996), Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system., *Proc. Natl. Acad. Sci. U.S.A.* 93:10321-10326; 및 M. Vidal, et al. (1996), Reverse two-hybrid and one-hybrid systems to detect dissociation of protein- protein and DNA-protein interactions., *Proc. Natl. Acad. Sci. U.S.A.* 93:10315-10320]을 참고한다.

[0100] 양성 선택 마커에 대해, 음성 선택 마커는 또한 다양한 분자들 중 임의의 것일 수 있다. 양성 선택 마커 및/또는 음성 선택 마커는 적당한 반응물의 존재 하에 발광 반응을 촉매하거나 형광화하는 폴리펩티드일 수 있다. 예를 들어, 음성 선택 마커는 예컨대 루시페라제, 녹색 형광 단백질(GFP), YFP, EGFP, RFP, 항생제 내성 유전자의 산물(예컨대, 클로람페니콜 아세틸트랜스퍼라제(CAT)), *lacZ* 유전자의 산물, 전사 조절자 단백질 등을 포함하나 이에 한정되지 않는다. 양성 선택 마커 및/또는 음성 선택 마커는 형광 활성화 세포 분류법(FACS) 또는 발광에 의해 검출될 수 있다. 양성 선택 마커 및/또는 음성 선택 마커는 친화성 기재 선별 마커를 포함할 수 있다. 동일한 폴리뉴클레오티드는 양성 선택 마커 및 음성 선택 마커 모두를 코딩할 수 있다. 예를 들어, 양성 선택 단계, 음성 선택 단계, 또는 양성 및 음성 선택 단계 모두는 리포터의 사용을 포함할 수 있고, 여기에서 리포터는 형광 활성화 세포 분류법(FACS)에 의해 검출된다. 예를 들어, 양성 선택은 먼저 양성 선택 마커, 예컨대 클로람페니콜 아세틸트랜스퍼라제(CAT) 유전자(여기에서, CAT 유전자는 셀렉터 코돈, 예컨대 앰버 종결 코돈을 CAT 유전자 내에 포함함)를 이용하여 수행된 후, 그에 이어서 음성 마커, 예컨대 T7 RNA 중합효소 유전자 내의 위치에서 셀렉터 코돈(들), 예컨대 2개 이상의 셀렉터 코돈을 억제할 수 없음에 기초하는 음성 선택 선별을 수행한다. 양성 선택 마커 및 음성 선택 마커는 동일한 벡터, 예컨대 플라스미드에서 발견될 수 있다. 음성 마커의 발현은 리포터, 예컨대 녹색 형광 단백질(GFP)의 발현을 유도한다. 선택 및 선별 엄격성이 다양화될 수 있고, 예컨대 리포터의 형광화에 필요한 광 강도가 다양화될 수 있다. 양성 선택은 FACS에 의해 선별되는, 양성 선택 마커로서의 리포터를 이용하여 수행될 수 있고, 그에 이어서 음성 마커, 예컨대 바나제 유전자 내의 위치에서 셀렉터 코돈(들), 예컨대 2개 이상의 셀렉터 코돈을 억제할 수 없음에 기초하는 음성 선택 선별을 수행한다.

[0101] 임의적으로, 리포터는 세포 표면, 예컨대 과지 표시 등에서 표시된다. 세포-표면 표시, 예컨대 OmpA-기재 세포-표면 표시 시스템은 에舍利키아 콜라이 세포의 표면 상에서 특별한 에피토프, 예컨대 세포외막 포린(porin) OmpA에 융합된 소아마비 바이러스 C3 펩티드의 발현에 의존한다. 에피토프는 단백질 메시지 내 셀렉터 코돈이 번역 중에 억제될 때에만 세포 표면 상에 표시된다. 이어서, 표시된 펩티드는 라이브러리 내 돌연변이 아미노산

실-tRNA 합성효소 중 하나에 의해 인식되는 아미노산을 함유하고, 상응하는 합성효소 유전자를 함유하는 세포는 특정 비천연 아미노산을 함유하는 펩티드에 대한 항체를 이용하여 단리될 수 있다. OmpA-기재 세포-표면 표시 시스템은 파지 표시에 대안으로서 Georgiou 등에 의해 개발되고 최적화되었다. [Francisco, J. A., Campbell, R., Iverson, B. L. & Georgiou, G. Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface. Proc. Natl. Acad. Sci. U S A 90:10444-8 (1993)]을 참고한다.

[0102] 본 발명의 다른 실시양태는 시험관 내 선택 단계들 중 하나 이상을 수행함을 포함한다. 이어서, 선택된 성분, 예컨대 합성효소 및/또는 tRNA는 생체 내 비천연 아미노산의 생체 내 혼입에 사용하기 위해 세포에 도입될 수 있다.

[0103] O-RS의 생산 및 합성효소의 기질 특이성 변형에 대한 부가적 상세 내용은, 본원에 각기 참고로 인용되는 미국 특허 출원 제10/126,931호(발명의 명칭: "METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGONAL tRNA-AMINOACYL tRNA SYNTHETASE PAIRS") 및 USSN 제10/825,867호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")에서 찾아볼 수 있다. O-RS의 생산에 대한 부가적 상세 내용은, 본원에 각기 참고로 인용되는 [Hamano-Takaku et al., (2000) A mutant Escherichia coli Tyrosyl-tRNA Synthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently than Tyrosine, Journal of Biological Chemistry, 275(51):40324-40328; Kiga et al. (2002), An engineered Escherichia coli tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system, PNAS 99(15): 9715-9723; 및 Francklyn et al., (2002), Aminoacyl-tRNA synthetases: Versatile players in the changing theater of translation; RNA, 8:1363-1372]에서 찾아볼 수 있다.

[0104] 공급원 및 숙주 유기체

[0105] 본 발명의 번역 성분은 전형적으로 비진핵생물 유기체로부터 유래된다. 예를 들어, 오르소고날 O-tRNA는 비진핵 생물 유기체, 예를 들어 고세균, 예컨대 메타노코커스 잔네쉬이, 메타노박테리움 썬모아우트로피쿰, 호염성 세균, 예컨대 할로페락스 볼카니 및 호염성 세균종 NRC-1, 아르카에오글로부스 풀기두스, 파이로코커스 푸리오서스, 파이로코커스 호리코쉬이, 아에우로피룸 페르닉스 등, 또는 진정세균, 예컨대 에셰리키아 콜라이, 썬무스 썬모필러스, 바실러스 스테아로썬모필러스 등으로부터 유래되고, 한편 오르소고날 O-RS는 비진핵생물 유기체, 예컨대 메타노코커스 잔네쉬이, 메타노박테리움 썬모아우트로피쿰, 호염성 세균, 예컨대 할로페락스 볼카니 및 호염성 세균종 NRC-1, 아르카에오글로부스 풀기두스, 파이로코커스 푸리오서스, 파이로코커스 호리코쉬이, 아에우로피룸 페르닉스 등, 또는 진정세균, 예컨대 에셰리키아 콜라이, 썬무스 썬모필러스, 바실러스 스테아로썬모필러스 등으로부터 유래된다. 한 실시양태에서, 식물, 해조류, 원생생물, 진균류, 효모, 동물(예컨대, 포유류, 곤충류, 절지류 등) 등을 포함하나 이에 한정되지 않는 진핵생물 공급원을 이용할 수 있다.

[0106] O-tRNA/O-RS 쌍의 개별 성분은 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. 예를 들어, O-tRNA는 예컨대 호염성 세균종 NRC-1로부터 유래되고, O-RS는 예컨대 메타노박테리움 썬모아우트로피쿰으로부터 유래된다.

[0107] O-tRNA, O-RS 또는 O-tRNA/O-RS 쌍은 생체 내 또는 시험관 내 선택 또는 선별되고/되거나, 세포, 예컨대 비진핵 생물 세포(예컨대, 대장균 세포), 또는 진핵생물 세포에서 사용되어, 선택된 아미노산(예컨대, 비천연 아미노산)을 갖는 폴리펩티드를 생산할 수 있다. 비진핵생물 세포는 다양한 공급원, 예컨대 메타노코커스 잔네쉬이, 메타노박테리움 썬모아우트로피쿰, 호염성 세균, 예컨대 할로페락스 볼카니 및 호염성 세균종 NRC-1, 아르카에오글로부스 풀기두스, 파이로코커스 푸리오서스, 파이로코커스 호리코쉬이, 아에우로피룸 페르닉스를 포함하나 이에 한정되지 않는 고세균 계통 도메인, 또는 에셰리키아 콜라이, 썬무스 썬모필러스, 바실러스 스테아로썬모필러스, 슈도모나스 플루오레센스, 슈도모나스 아에루기노사, 슈도모나스 푸티다 등을 포함하나 이에 한정되지 않는 진정세균 계통 도메인으로부터 유래될 수 있다. 진핵생물 세포는 식물류(예컨대, 단자엽식물 또는 쌍자엽식물과 같은 복합 식물), 해조류, 원생생물, 진균류, 효모(사카로마이세스 세레비지아에를 포함하나 이에 한정되지 않음), 동물류(포유류, 곤충류, 절지류 등을 포함하나 이에 한정되지 않음) 등을 포함하나 이에 한정되지 않는 다양한 공급원으로부터 유래될 수 있다. 본 발명의 번역 성분을 갖는 세포의 조성물도 또한 본 발명의 한 특징이다. 또한, 또 다른 종에 사용하기 위한 어느 한 종에서의 O-tRNA 및/또는 O-RS를 선별하는 것에 대해, USSN 제10/825,867호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE")를 참고한다.

- [0108] 숙주 세포 내 선택된 아미노산을 갖는 관심 폴리펩티드를 발현하기 위해, 관심 폴리펩티드를 코딩하는 폴리뉴클레오티드를, 전사를 지시하는 강한 프로모터, 전사/번역 종결자, 및 단백질을 코딩하는 핵산의 경우, 번역 개시를 위한 리보솜 결합 부위를 함유하는 발현 벡터에 서브클로닝(subcloning)할 수 있다. 적당한 세균 프로모터는 당업계에 공지되어 있고, 예컨대 [Sambrook et al. 및 Ausubel et al.]에 기재되어 있다.
- [0109] 관심 폴리펩티드 발현용 세균 발현 시스템은 대장균, 바실러스 종, 슈도모나스 플루오레센스, 슈도모나스 아에루기노사, 슈도모나스 푸티다 및 살모넬라를 포함하나 이에 한정되지 않는 것들에서 입수가능하다([Palva et al., Gene 22: 229-235(1983); Mosbach et al., Nature 302: 543-545(1983)] 참고). 이러한 발현 시스템용 키트는 상업적으로 구입가능하다. 포유류 세포, 효모 및 곤충 세포용 진핵생물 발현 시스템은 당업계에 공지되어 있고, 또한 상업적으로도 입수가능하다.
- [0110] 본 발명의 tRNA 및/또는 RS, 및/또는 관심 폴리펩티드는 예를 들어, 효모, 곤충 세포, 포유류 세포 및 세균을 비롯한 임의의 수의 적당한 발현 시스템에서 이용되고/되거나 발현될 수 있다. 이하 예시적인 발현 시스템의 설명이 제공된다.
- [0111] 효모 본원에 사용된 용어 "효모"는 관심 폴리펩티드를 발현시킬 수 있는 임의의 다양한 효모를 포함한다. 이러한 효모는 자낭포자형성(ascosporogenous) 효모(엔도마이세탈레스(*Endomycetales*)), 담자포자(basidiosporogenous) 효모, 및 불완전 진균류(*Fungi imperfecti*)(블라스토마이세테스(*Blastomycetes*)) 군에 속하는 효모를 포함하나 이에 한정되지 않는다. 자낭포자형성 효모는 2개의 부류, 스퍼모포르타세아에(*Spermophthoraceae*) 및 사카로마이세타세아에(*Saccharomycetaceae*)로 나뉜다. 후자는 4개의 아부류, 쉬조사카로마이코이데에(*Schizosaccharomycoidae*)(예컨대, 쉬조사카로마이세스(*Schizosaccharomyces*) 속), 나드소니오이데에(*Nadsonioideae*), 리포마이코이데에(*Lipomycoideae*) 및 사카로마이코이데에(*Saccharomycoidae*)(예컨대, 피치아(*Pichia*) 속, 클루이베로마이세스(*Kluyveromyces*) 속 및 사카로마이세스(*Saccharomyces*) 속)를 포함한다. 담자포자 효모는 류코스포리디움(*Leucosporidium*) 속, 로도스포리디움(*Rhodospiridium*) 속, 스포리디오볼루스(*Sporidiobolus*) 속, 필로바시디움(*Filobasidium*) 속 및 필로바시디엘라(*Filobasidiella*) 속을 포함한다. 불완전 진균류(블라스토마이세테스) 군에 속하는 효모는 2개의 부류, 스포로볼로마이세타세아에(*Sporobolomycetaceae*)(예컨대, 스포로볼로마이세스(*Sporobolomyces*) 속 및 불레라(*Bullera*) 속) 및 크립토크카세아에(*Cryptococcaceae*)(예컨대, 칸디다(*Candida*) 속)로 나뉜다.
- [0112] 본 발명에 사용하기에 특히 유익한 것으로는 *P. 파스토리스*(*P. pastoris*), *P. 길러리몬디*(*P. guillerimondii*), *S. 세레비시아에*(*S. cerevisiae*), *S. 칼스버그렌시스*(*S. carlsbergensis*), *S. 디아스타티쿠스*(*S. diastaticus*), *S. 더글라시*(*S. douglasii*), *S. 클루이베리*(*S. kluyveri*), *S. 노르벤시스*(*S. norbensis*), *S. 오비포르미스*(*S. oviformis*), *K. 락티스*(*K. lactis*), *K. 프라길리스*(*K. fragilis*), *C. 알비칸스*(*C. albicans*), *C. 말토사*(*C. maltosa*) 및 *H. 폴리모르파*(*H. polymorpha*)를 포함하나 이에 한정되지 않는 피치아 속, 클루이베로마이세스 속, 사카로마이세스 속, 쉬조사카로마이세스 속, 한세놀라(*Hansenula*) 속, 토룰롭시스(*Torulopsis*) 속 및 칸디다 속에 속하는 종이 있다. 효모는 일반적으로 캘리포니아 대학(University of California)(미국 캘리포니아주 버클리 소재) 생체물리학 및 의료 물리학부(Department of Biophysics and Medical Physics) 이스트 제네틱 스톡 센터(Yeast Genetic Stock Center) 및 아메리칸 타입 컬처 콜렉션(American Type Culture Collection; "ATCC")(미국 버지니아주 만나사스 소재)을 포함하나 이에 한정되지 않은 다양한 공급원으로부터 입수가능하다.
- [0113] 용어 "효모 숙주" 또는 "효모 숙주 세포"는 재조합 벡터 또는 다른 전이 DNA에 대한 수용물(recipient)로 사용될 수 있거나 사용된 효모를 포함한다. 이 용어는 재조합 벡터 또는 다른 전이 DNA를 수용한 원래의 효모 숙주 세포의 후손을 포함한다. 단일 모 세포의 후손이 우발적 돌연변이 또는 의도적 돌연변이로 인해 원래의 모체와 형상 또는 총 DNA 보체가 완전히 동일하지 않을 수 있는 것으로 이해된다. 관련 성질, 예를 들어 관심 폴리펩티드를 코딩하는 뉴클레오티드의 존재에 의해 특성화되는 모체와 충분히 유사한 모 세포의 후손은 이 정의에 의해 의도되는 후손에 포함된다.
- [0114] 염색체의 레플리콘(extrachromosomal replicon) 또는 혼입 벡터를 비롯한 발현 벡터 및 형질전환 벡터가 많은 효모 숙주로의 형질전환을 위해 개발되었다. 예를 들어, *S. 세레비시아에*([Sikorski et al., GENETICS (1998) 112: 19; Ito et al., J. BACTERIOL. (1983) 153: 163; Hinnen et al., PROC. NATL. ACAD. SCI. USA (1978) 75: 1929] 참고); *C. 알비칸스*([Kurtz et al., MOL. CELL. BIOL. (1986) 6: 142] 참고); *C. 말토사*([Kunze et al., J. BASICMICROBIOL. (1985) 25: 141] 참고); *H. 폴리모르파*([Gleeson et al., J. GEN. MICROBIOL. (1986) 132: 3459; Roggenkamp et al., MOL. GEN. GENET. (1986) 202: 302] 참고); *K. 프라길리스*([Das et al., J. BACTERIOL. (1984) 158: 1165] 참고); *K. 락티스*([De Louvencourt et al., J. BACTERIOL. (1983)

154: 737; Van den Berg et al., BIOTECHNOLOGY(NY) (1990) 8: 135] 참고); P. 길러리몬디([Kunze et al., J. BASIC MICROBIOL. (1985) 25: 141] 참고); P. 파스토리스(미국 특허 제5,324,639호, 제4,929,555호 및 제 4,837,148호; 문헌[Cregg et al., MOL. CELL. BIOL. (1985) 5: 3376] 참고); 쉬조사카로마이세스 폼베 (*Schizosaccharomyces pombe*) ([Beach et al., NATURE (1981) 300: 706] 참고); 및 Y. 리폴리티카 (*lipolytica*); A. 니들란스(*nidulans*) ([Ballance et al., BIOCHEM. BIOPHYS. RES. COMMUN. (1983) 112: 284-89; Tilburn et al., GENE (1983) 26: 205-221; 및 Yelton et al., PROC. NATL. ACAD. SCI. USA (1984) 81: 1470-74] 참고); A. 나이저(*niger*) ([Kelly and Hynes, EMBO J. (1985) 4: 475479] 참고); T. 레시아(*reesia*) (EP 0 244 234); 및 섬사상 진균류, 예를 들어 뉴로스포라(*Neurospora*), 페니실리움(*Penicillium*), 톨리포클라 디움(*Tolypocladium*) (WO 91/00357)(상기 각각의 문헌은 본원에 참고 문헌으로 인용됨)에 대한 발현 벡터가 개 발되었다.

[0115] 효모 벡터에 대한 조절 서열은 당업자에게 공지되어 있고, 이는 알코올 데히드로게나제(ADH)(EP 0 284 044); 에 놀라제; 글루코키나제; 포도당-6-인산염 이소머라제; 글리세르알데히드-3-인산염-데히드로게나제(GAP 또는 GAPDH); 헥소키나제; 포스포프룩토키나제[Hitzeman et al., J. BIOL. CHEM. (1980) 255: 2073]; 3-포스포글리 세레이트 뮤타제; 및 피루베이트 키나제(PyK) (EP 0 329 203)와 같은 유전자로부터의 프로모터 영역을 포함하나 이에 한정되지 않는다. 또한, 산 포스파타제를 코딩하는 효모 PH05 유전자도 유용한 프로모터 서열을 제공할 수 있다 ([Myanohara et al., PROC. NATL. ACAD. SCI. USA (1983) 80: 1] 참고). 효모 숙주에 사용될 수 있는 다 른 적당한 프로모터 서열은 3-포스포글리세레이트 키나제; 및 다른 해당 효소, 예를 들어 피루베이트 데카르복 실라제, 트리오스인산염 이소머라제, 및 포스포글루코스 이소머라제[Holland et al., BIOCHEMISTRY (1978) 17: 4900; Hess et al., J. ADV. ENZYME REG. (1968) 7: 149]에 대한 프로모터를 포함할 수 있다. 전사가 성장 조 건에 의해 조절되는 추가적인 이점을 갖는 유도성 효모 프로모터는 알코올 데히드로게나제 2; 이소시토크롬 C; 산 포스파타제; 메탈로티오네인(metallothionein); 글리세르알데히드-3-인산염 데히드로게나제; 질소 기전과 관 련된 분해성 효소 및 말토스 및 갈락토스 이용을 담당하는 효소에 대한 프로모터 영역을 포함할 수 있다. 효모 발현에 사용하기에 적당한 벡터 및 프로모터는 EP 0 073 657에 추가로 기재되어 있다.

[0116] 효모 인핸서(enhancer)를 또한 효모 프로모터와 함께 사용할 수 있다. 부가적으로, 합성 프로모터도 또한 효모 프로모터로서 기능할 수 있다. 예를 들어, 효모 프로모터의 상류 활성화 서열(UAS)이 또 다른 효모 프로모터의 전사 활성화 영역과 결합되어, 합성 혼성 프로모터를 생산할 수 있다. 이러한 혼성 프로모터의 예는 GAP 전사 활성화 영역에 결합된 ADH 조절 서열을 포함한다. 본원에 참고로 인용되는 미국 특허 제4,880,734호 및 제 4,876,197호를 참고한다. 혼성 프로모터의 다른 예는 해당 효소 유전자의 전사 활성화 영역, 예를 들어 GAP 또 는 PyK와 조합된 ADH2, GAL4, GAL10 또는 PH05 유전자로 된 조절 서열로 이루어진 프로모터를 포함한다. EP 0 164 556을 참고한다. 또한, 효모 프로모터는 효모 RNA 중합효소에 결합하고 전사를 개시할 수 있는 능력을 갖는 비효모 기원의 천연 발생 프로모터를 포함할 수 있다.

[0117] 효모 발현 벡터의 일부를 포함할 수 있는 다른 조절 요소는 예를 들어, GAPDH 또는 에놀라제 유전자로부터의 종 결자를 포함한다 ([Holland et al., J. BIOL. CHEM. (1981) 256: 1385] 참고). 또한, 2  $\mu$  플라스미드 기원으 로부터의 복제 기점은 효모에 적합하다. 효모에 사용하기에 적당한 선택 유전자로는 효모 플라스미드에 존재하 는 *trp1* 유전자가 있다. [Tschemper et al., GENE (1980) 10: 157; Kingsman et al., GENE (1979) 7: 141]을 참고한다. 이러한 *trp1* 유전자는 트립토판에서 성장하는 능력이 결여된 효모의 돌연변이체 군주에 선별 마커를 제공한다. 유사하게, Leu2-결핍 효모 군주(ATCC 20,622 또는 38,626)는 Leu2 유전자를 갖는 공지된 플라스미드 에 의해 보충된다.

[0118] 외인성 DNA를 효모 숙주에 도입시키는 방법은 당업자에게 공지되어 있고, 전형적으로 구상 원시 세포 (spheroplast)의 형질전환 또는 알칼리 양이온으로 처리된 본연 그대로의 효모 숙주 세포의 형질전환을 포함하 나 이에 한정되지 않는다. 예를 들어, 효모의 형질전환은 [Hsiao et al., PROC. NATL. ACAD. SCI. USA (1979) 76: 3829, 및 Van Solingen et al., J. BACT. (1977) 130: 946]에 기재된 방법에 따라 수행될 수 있다. 그러나, 예를 들어, 핵 주사, 전기천공법 또는 원형질체 융합에 의해 DNA를 세포에 도입시키는 다른 방법도 또 한 일반적으로 [SAMBROOK ET AL., MOLECULAR CLONING: A LAB. MANUAL (2001)]에 기재된 바와 같이 사용될 수 있다. 이어서, 효모 숙주 세포를 당업자에게 공지된 표준 기술을 사용하여 배양할 수 있다.

[0119] 효모 숙주 세포에서 이종 단백질을 발현시키는 다른 방법이 당업자에게 공지되어 있다. 일반적으로, 미국 특허 공보 20020055169, 미국 특허 제6,361,969호, 제6,312,923호, 제6,183,985호, 제6,083,723호, 제6,017,731호, 제5,674,706호, 제5,629,203호, 제5,602,034호 및 제5,089,398호; 재심사된 미국 특허 제RE37,343호 및 제 RE35,749호; PCT 공개 특허 출원 WO 99/078621, WO 98/37208, 및 WO 98/26080; 유럽 특허 출원 EP 0 946 736,



EP 0 732 403, EP 0 480 480, EP 0 460 071, EP 0 340 986, EP 0 329 203, EP 0 324 274, 및 EP 0 164 556을 참고한다. 또한, 각각 본원에 참고로 인용되는 [Gellissen et al., ANTONIE VAN LEEUWENHOEK (1992) 62 (1-2): 79-93; Romanes et al., YEAST (1992) 8 (6): 423-488; Goeddel, METHODS IN ENZYMOLOGY (1990) 185: 3-7]을 참고한다.

[0120] 효모 숙주 균주는 당업자에게 공지되어 있는 표준 공급물 배치 발효 방법을 사용하여 증폭 단계 동안 발효기에 서 성장할 수 있다. 발효 방법은 특별한 효모 숙주의 탄소 이용 경로 또는 발현 조절 방식에서의 차이를 유발하 도록 적합될 수 있다. 예를 들어, 사카로마이세스 효모 숙주의 발효는 단일 포도당 공급물, 복합체 질소 공급원 (예컨대, 카세인 가수분해물) 및 복합 비타민 보충물을 필요로 할 수 있다. 대조적으로, 메틸요구성 효모 P. 파 스토리스는 글리세롤, 메탄올 및 미량의 무기질 공급물을 필요로 할 수 있지만, 최적의 성장 및 발현을 위해 서는 오직 단순한 암모늄(질소) 염만을 필요로 할 수 있다. 예를 들어, 본원에 각기 참고로 인용되는, 미국 특허 제5,324,639호; [Elliott et al., J. PROTEIN CHEM. (1990) 9: 95]; 및 [Fieschko et al., BIOTECH. BIOENG. (1987) 29: 1113]을 참고한다.

[0121] 그러나, 이러한 발효 방법은 사용되는 효모 숙주 균주와는 독립적인 특정한 공통된 특징을 가질 수 있다. 예를 들어, 성장 제한 영양소, 전형적으로 탄소를 증폭기 동안 발효기에 첨가하여 최대 성장을 유발시킬 수 있다. 또 한, 발효 방법은 일반적으로 적당한 양의 탄소, 질소, 기저 염, 인, 및 다른 소수의 영양분 (비타민, 미량의 무 기질 및 염 등)을 함유하도록 설계된 발효 배지를 사용한다. 피치아를 사용하는데 적당한 발효 배지의 예는 본 원에 참고로 인용되는 미국 특허 제5,324,639호 및 제5,231,178호에 기재되어 있다.

[0122] 바큇로바이러스(Baculovirus)-감염 곤충 세포 용어 "곤충 숙주" 또는 "곤충 숙주 세포"는 재조합 벡터 또는 다 른 전이 DNA에 대한 수용물로 사용될 수 있거나 사용된 곤충을 지칭한다. 이 용어는 트랜스펙션된 원래의 곤충 숙주 세포의 후손을 포함한다. 단일 모 세포의 후손이 우발적 돌연변이 또는 의도적 돌연변이로 인해 원래의 모 체와 형상 또는 총 DNA 보체가 완전히 동일하지 않을 수 있을 것으로 이해된다. 관련 성질, 예를 들어 관심 폴 리펩티드를 코딩하는 뉴클레오티드의 존재에 의해 특성화되는 모체와 충분히 유사한 모 세포의 후손은 이 정의 에 의해 의도되는 후손에 포함된다.

[0123] 관심 폴리펩티드의 발현에 적당한 곤충 세포의 선택은 당업자에게 공지되어 있다. 아에데스 아에집티(*Aedes aegypti*), 붕박스 모리(*Bombyx mori*), 드로소필라 멜라노가스터(*Drosophila melanogaster*), 스포도프테라 프루 기페르다(*Spodoptera frugiperda*) 및 트리초플루시아 니(*Trichoplusia ni*)를 비롯한 수가지 곤충 종은 당 분야 에서 잘 기재되어 있고, 상업적으로 구입가능하다. 발현을 위해 곤충 숙주를 선택함에 있어서, 적당한 숙주는 그 중에서도, 우수한 분비 용량, 낮은 단백질분해 활성 및 전체적인 강인성을 갖는 것으로 보이는 것들을 포함 할 수 있다. 일반적으로, 곤충은 캘리포니아 대학(미국 캘리포니아주 버클리 소재) 생체물리학 및 의료 물리학 부 인섹트 제네틱 스톡 센터(Insect Genetic Stock Center) 및 아메리칸 타입 컬처 콜렉션("ATCC") (미국 버지 니아주 만나사스 소재)을 포함하나 이에 한정되지 않은 다양한 공급원으로부터 입수가능하다.

[0124] 일반적으로, 바큇로바이러스-감염 곤충 발현 시스템의 성분은 전이 벡터, 통상 바큇로바이러스 게놈의 단편 및 발현될 이중 유전자의 삽입에 편리한 제한 부위 양자 모두를 함유하는 세균 플라스미드; 전이 벡터 중에서 바큇 로바이러스-특이적 단편에 대해 동종성인 서열을 갖는 야생형 바큇로바이러스(이는 바큇로바이러스 게놈으로 이 중 유전자의 동족 재조합을 가능케 함); 및 적당한 곤충 숙주 세포 및 성장 배지를 포함한다. 벡터를 구성하고, 세포를 트랜스펙션시키고, 플라크(plaque)를 선택하고, 배양물 중에서 세포를 성장시키는 것 등에 사용되는 물 질, 방법 및 기술은 당업계에 공지되어 있고, 이러한 기술들을 기재하는 편람도 입수가능하다.

[0125] 이중 유전자를 전이 벡터에 삽입시킨 후, 벡터 및 야생형 바이러스 게놈을 곤충 숙주 세포로 트랜스펙션시키며, 여기에서 벡터 및 바이러스 게놈은 재조합된다. 패키징된 재조합 바이러스를 발현시키고, 재조합 플라크를 동정 하고 정제한다. 바큇로바이러스/곤충 세포 발현 시스템용 물질 및 방법은 예를 들어, 인비트로젠 코퍼레이션 (Invitrogen Corp.) (미국 캘리포니아주 칼사버드 소재)으로부터 키트 형태로 상업적으로 구입가능하다. 이러한 기술들은 일반적으로 당업자에게 공지되어 있고, 본원에 참고로 인용되는 [SUMMERS AND SMITH, TEXAS AGRICULTURAL EXPERIMENT STATION BULLETIN No. 1555 (1987)]에 충분히 기재되어 있다. 또한, [RICHARDSON, 39 METHODS IN MOLECULAR BIOLOGY: BACULOVIRUS EXPRESSION PROTOCOLS (1995); AUSUBEL ET AL., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY 16.9-16.11 (1994); KING AND POSSEE, THE BACULOVIRUS SYSTEM: A LABORATORY GUIDE (1992); 및 O'REILLY ET AL., BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL (1992)]을 참고한다.

[0126] 실제로, 바큇로바이러스/곤충 세포 발현 시스템을 사용한 다양한 이중 단백질의 생산이 당업계에 공지되어

있다. 예를 들어, 본원에 참고로 인용되는 미국 특허 제6,368,825호, 제6,342,216호, 제6,338,846호, 제6,261,805호, 제6,245,528호, 제6,225,060호, 제6,183,987호, 제6,168,932호, 제6,126,944호, 제6,096,304호, 제6,013,433호, 제5,965,393호, 제5,939,285호, 제5,891,676호, 제5,871,986호, 제5,861,279호, 제5,858,368호, 제5,843,733호, 제5,762,939호, 제5,753,220호, 제5,605,827호, 제5,583,023호, 제5,571,709호, 제5,516,657호, 제5,290,686호; WO 02/06305, WO 01/90390, WO 01/27301, WO 01/05956, WO 00/55345, WO 00/20032 WO 99/51721, WO 99/45130, WO 99/31257, WO 99/10515, WO 99/09193, WO 97/26332, WO 96/29400, WO 96/25496, WO 96/06161, WO 95/20672, WO 93/03173, WO 92/16619, WO 92/02628, WO 92/01801, WO 90/14428, WO 90/10078, WO 90/02566, WO 90/02186, WO 90/01556, WO 89/01038, WO 89/01037, WO 88/07082를 참고한다.

[0127] 바콜로바이러스/곤충 세포 발현 시스템에 유용한 벡터는 당업계에 공지되어 있고, 예를 들어, 헬퍼-독립적 바이러스 발현 벡터인 바콜로바이러스 오토그라파칼리포르니카(*Autographa californica*) 핵 다각체병(polyhedrosis) 바이러스(AcNPV)로부터 유도된 곤충 발현 및 전이 벡터를 포함한다. 통상, 이 시스템으로부터 유도된 바이러스 발현 벡터는 강한 바이러스 다각체 유전자 프로모터를 사용하여 이중 유전자를 발현시킨다. 일반적으로, [O'Reilly ET AL., BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL (1992)]을 참고한다.

[0128] 외래 유전자를 바콜로바이러스 게놈에 삽입시키기 전에, 프로모터, 리더 (leader) (필요에 따라), 관심 코딩 서열 및 전사 종결 서열을 포함하는 상기한 성분을 전형적으로 중간 전도 구축물(intermediate transplacement construct) (전이 벡터)로 조립시킨다. 종종, 중간 전도 구축물은 숙주, 예를 들어 세균에서 안정하게 유지될 수 있는 레플리콘, 예를 들어 염색체 외부 요소(예컨대, 플라스미드)에서 유지된다. 레플리콘은 복제 시스템을 갖기 때문에, 이를 클로닝 및 증폭에 적합한 숙주에서 유지되게 한다. 더욱 구체적으로, 플라스미드는 다각체 폴리아데닐화 신호([Miller et al., ANN. REV. MICROBIOL. (1988) 42: 177] 참고) 및 대장균 중에서의 선택 및 증식을 위한 원핵생물 암피실린-내성(amp) 유전자 및 복제 기점을 함유할 수 있다.

[0129] 외래 유전자를 AcNPV로 도입시키는데 통상 사용되는 한 전이 벡터는 pAc373이다. 또한, 예를 들어, 다각체 출발 코돈을 ATG로부터 ATT로 변경시키고, ATT로부터의 하류에 BamHI 클로닝 부위 32 염기쌍을 도입하는 pVL985를 비롯한 당업자에게 공지된 많은 다른 벡터들이 설계되었다. [Luckow and Summers, VIROLOGY 170:31 (1989)]을 참고할 수 있다. 다른 상업적으로 구입가능한 벡터는 예를 들어, PBlueBac4.5/V5-His; pBlueBacHis2; pMelBac; 및 pBlueBac4.5(인비트로젠 코퍼레이션, 캘리포니아주 칼스버드 소재)를 포함한다.

[0130] 이중 유전자의 삽입 후, 전이 벡터 및 야생형 바콜로바이러스 게놈은 곤충 세포 숙주로 함께 트랜스펙션된다. 바콜로바이러스 바이러스에서 이중 DNA를 원하는 부위로 도입시키는 방법이 당업계에 공지되어 있다. [SUMMERS AND SMITH, TEXAS AGRICULTURAL EXPERIMENT STATION BULLETIN No. 1555 (1987); Smith et al., MOL. CELL. BIOL. (1983) 3: 2156; Luckow and Summers, VIROLOGY (1989) 170:31]을 참고한다. 예를 들어, 동종 이중 교차 재조합에 의해 유전자, 예를 들어 다각체 유전자로 삽입될 수 있고, 또한 원하는 바콜로바이러스 유전자로 유전 공학처리되는 제한 효소 부위로 삽입될 수 있다. [Miller et al., BIOESSAYS (1989) 11(4): 91]을 참고한다.

[0131] 트랜스펙션은 전기천공법에 의해 달성될 수 있다. [TROTTER AND WOOD, 39 METHODS IN MOLECULAR BIOLOGY (1995); Mann and King, J. GEN. VIROL. (1989) 70: 3501]을 참고한다. 대안적으로, 재조합 발현 벡터 및 바콜로바이러스를 사용하여 곤충 세포를 트랜스펙션시키는데 리포솜을 사용할 수 있다. 예를 들어, [Liebman et al., BIOTECHNIQUES (1999) 26 (1): 36; Graves et al., BIOCHEMISTRY (1998) 37: 6050; Nomura et al., J. BIOL. CHEM. (1998) 273 (22): 13570; Schmidt et al., PROTEIN EXPRESSION AND PURIFICATION (1998) 12: 323; Siffert et al., Nature GENETICS (1998) 18: 45; TILKINS ET AL., CELL BIOLOGY: A LABORATORY HANDBOOK 145-154 (1998); Cai et al., PROTEIN EXPRESSION AND PURIFICATION (1997) 10: 263; Dolphin et al., NATURE GENETICS (1997) 17: 491; Kost et al., GENE (1997) 190: 139; Jakobsson et al., J. BIOL. CHEM. (1996) 271: 22203; Rowles et al., J. BIOL. CHEM. (1996) 271 (37): 22376; Reversey et al., J. BIOL. CHEM. (1996) 271 (39): 23607-10; Stanley et al., J. BIOL. CHEM. (1995) 270: 4121; Sisk et al., J. VIROL. (1994) 68 (2): 766; 및 Peng et al., BIOTECHNIQUES (1993) 14(2); 274]을 참고한다. 상업적으로 구입가능한 리포솜은 예를 들어, 셀펙틴(Cellfectin)<sup>®</sup> 및 리포펙틴(Lipofectin)<sup>®</sup>(인비트로젠 코퍼레이션, 미국 캘리포니아주 칼스버드 소재)를 포함한다. 또한, 인산칼슘 트랜스펙션도 사용할 수 있다. [TROTTER AND WOOD, 39 METHODS IN MOLECULAR BIOLOGY (1995); Kitts, NAR (1990) 18 (19): 5667; 및 Mann and King, J. GEN. VIROL. (1989) 70: 3501]을 참고한다.

[0132] 바콜로바이러스 발현 벡터는 통상 바콜로바이러스 프로모터를 포함한다. 바콜로바이러스 프로모터는 바콜로바이

러스 RNA 중합효소를 결합시킬 수 있고, 코딩 서열(예컨대, 구조적 유전자)의 mRNA로의 하류(3') 전사를 개시할 수 있는 임의의 DNA 서열이다. 프로모터는 통상 코딩 서열의 5' 말단에 근접하게 위치하는 전사 개시 영역을 갖게 된다. 전형적으로, 이 전사 개시 영역은 RNA 중합효소 결합 부위 및 전사 개시 부위를 포함한다. 또한, 바콜로바이러스 프로모터는 존재하는 경우, 통상 구조적 유전자와 멀리 떨어져 위치하는 인핸서로 불리는 제2 도메인을 가질 수 있다. 더욱이, 발현은 조절되거나, 구성적일 수 있다.

[0133] 감염 주기의 후기에서 풍부하게 전사되는 구조적 유전자는 특히 유용한 프로모터 서열을 제공한다. 그 예는 바콜로바이러스 다각체 단백질을 코딩하는 유전자[FRIESEN ET AL., The Regulation of Baculovirus Gene Expression in THE MOLECULAR BIOLOGY OF BACULOVIRUSES (1986); EP 0 127 839 및 0 155 476] 및 p10 단백질을 코딩하는 유전자[Vlak et al., J. GEN. VIROL. (1988) 69: 765]로부터 유도된 서열을 포함한다.

[0134] 새로이 형성된 바콜로바이러스 발현 벡터를 감염성 재조합 바콜로바이러스로 패키징하고, 후속하여 성장한 플라스크를 당업자에게 공지되어 있는 기술에 의해 정제한다. [Miller et al., BIOESSAYS (1989) 11(4): 91; SUMMERS AND SMITH, TEXAS AGRICULTURAL EXPERIMENT STATION BULLETIN No. 1555 (1987)]을 참고한다.

[0135] 재조합 바콜로바이러스 발현 벡터는 수가지 곤충 세포의 감염을 위해 개발되었다. 예를 들어, 그 중에서도, 아에데스 아에기프(ATCC 번호 CCL-125), 봄빅스 모리(ATCC 번호 CRL-8910), 드로소필라 멜라노가스터(ATCC 번호 1963), 스포도프테라 프루기페르다 및 트리초폴루시아 니에 대한 재조합 바콜로바이러스가 개발되었다. [Wright, NATURE (1986) 321: 718; Carbonell et al., J. VIROL. (1985) 56: 153; Smith et al., MOL. CELL. BIOL. (1983) 3: 2156]을 참고한다. 일반적으로, [Fraser et al., IN VITRO CELL. DEV. BIOL. (1989) 25: 225]을 참고한다. 더욱 구체적으로, 통상적으로 바콜로바이러스 발현 벡터 시스템에 사용되는 세포주는 Sf9(스포도프테라 프루기페르다) (ATCC 번호 CRL-1711), Sf21(스포도프테라 프루기페르다) (인비트로젠 코퍼레이션, 분류 번호 11497-013(미국 캘리포니아주 칼스버드 소재)), Tri-368(트리초폴루시아 니 (*Trichopulsia ni*)) 및 하이-화이브(High-Five)<sup>TM</sup> BTI-TN-5B1-4(트리초폴루시아 니)를 포함하나 이에 한정되지 않는다.

[0136] 바콜로바이러스/발현 중에서 이중 폴리펩티드의 직접 및 융합 발현 모두를 위한 세포 및 배양 배지가 상업적으로 구입가능하고, 세포 배양 기술이 일반적으로 당업자에게 공지되어 있다.

[0137] 대장균, 슈도모나스 종, 및 다른 원핵생물 세균 발현 기술은 당업자에 공지되어 있다. 세균 숙주에 사용하기 위해 매우 다양한 벡터들이 이용가능하다. 벡터는 단일 카피, 또는 낮거나 높은 다중카피 (multicopy) 벡터일 수 있다. 벡터는 클로닝 및/또는 발현에 작용할 수 있다. 벡터, 많은 벡터들의 상업적 이용가능성, 및 심지어 벡터 및 그들의 제한 지도 및 특징을 기재하는 편람에 관한 충분한 문헌이 존재한다는 점을 고려할 때, 본원에서 광범위한 논의는 불필요하다. 공지된 바와 같이, 벡터는 정상적으로 선택을 가능케 하는 마커를 포함하며, 여기에서 마커는 세포독성제 내성, 자가영양성(prototrophy) 또는 면역성을 제공할 수 있다. 흔히, 상이한 특성들을 제공하는 복수개의 마커가 존재한다.

[0138] 세균 프로모터는 세균 RNA 중합효소를 결합시킬 수 있고 코딩 서열(예컨대, 구조적 유전자)의 mRNA로의 하류(3') 전사를 개시할 수 있는 임의의 DNA 서열이다. 프로모터는 통상 코딩 서열의 5' 말단에 근접하게 위치하는 전사 개시 영역을 갖게 된다. 이 전사 개시 영역은 전형적으로 RNA 중합효소 결합 부위 및 전사 개시 부위를 포함한다. 또한, 세균 프로모터는 RNA 합성이 시작되는 인접 RNA 중합효소 결합 부위와 중첩될 수 있는 오퍼레이터(operator)로 불리는 제2 도메인을 가질 수 있다. 유전자 리프레서(repressor) 단백질이 오퍼레이터에 결합함으로써 특이적 유전자의 전사를 억제할 수 있음에 따라 오퍼레이터는 음성 조절된 (유도성) 전사를 일으킨다. 구성적 발현은 음성 조절 요소, 예를 들어 오퍼레이터의 부재 하에서 발생할 수 있다. 또한, 양성 조절은 존재하는 경우, 통상 RNA 중합효소 결합 서열에 근접(5')한 유전자 활성화제 단백질 결합 서열에 의해 수행될 수 있다. 유전자 활성화제(activator) 단백질은 에셰리키아 콜라이(대장균)에서 lac 오페론의 초기 전사를 돕는 이화산물 활성화제 단백질(CAP)이다[Raibaud et al., ANNU. REV. GENET. (1984) 18: 173]. 따라서, 조절된 발현은 양성 또는 음성일 수 있으며, 이로써 전사를 증진시키거나 감소시킬 수 있다.

[0139] 대사 경로 효소를 코딩하는 서열은 특히 유용한 프로모터 서열을 제공한다. 그 예는 당 대사 효소, 예를 들어 갈락토스, 락토스(lac)[Chang et al., NATURE (1977) 198: 1056], 및 말토스로부터 유도된 프로모터 서열을 포함한다. 추가적인 예는 생합성 효소, 예를 들어 트립토판(trp)으로부터 유도된 프로모터 서열을 포함한다 (본원에 참고로 인용되는 [Goeddel et al., Nuc. ACIDSRES. (1980) 8: 4057; Yelverton et al., NUCL. ACIDS RES. (1981) 9: 731]; 미국 특허 제4,738,921호; EP 공보 036 776 및 121 775 참고). 또한, β-갈락토시다제(bla) 프로모터 시스템(본원에 참고로 인용되는 [Weissmann (1981) "The cloning of interferon and other mistakes." In Interferon 3(Ed. I. Gresser)]), 박테리오파지 람다 PL(본원에 참고로 인용되는 [Shimatake et



al., NATURE (1981) 292: 128] 참고), 및 T5(본원에 참고로 인용되는 미국 특허 제4,689,406호 참고) 프로모터 시스템도 유용한 프로모터 서열을 제공한다. 강한 프로모터, 예컨대 T7 프로모터를 사용하여, 높은 수준으로 관심 폴리펩티드를 유도할 수 있다. 이러한 벡터의 예는 당업계에 공지되어 있고, 이는 노바젠(Novagen)으로부터의 pET29 계열, 및 본원에 참고로 인용되는 WO 99/05297에 기재된 pPOP 벡터를 포함한다. 이러한 발현 시스템은 숙주 세포 생존 능력 또는 성장 매개변수를 약화시키지 않으면서, 숙주 내 관심 폴리펩티드를 높은 수준으로 생성시킨다. pET19(노바젠)가 당업계에 공지된 또 다른 벡터이다.

[0140] 또한, 자연계에서 발생하지 않는 합성 프로모터도 세균 프로모터로서 기능한다. 예를 들어, 한 세균 또는 박테리오파지 프로모터의 전사 활성화 서열을 또 다른 세균 또는 박테리오파지 프로모터의 오프론 서열과 결합시켜, 합성 혼성 프로모터를 생산시킬 수 있다(본원에 참고로 인용되는 미국 특허 제4,551,433호 참고). 예를 들어, tac 프로모터는 trp 프로모터 및 lac 억제자에 의해 조절되는 lac 오프론 서열 양자 모두를 포함하는 혼성 trp-lac 프로모터이다 ([Amam et al., GENE (1983) 25: 167; de Boer et al., PROC. NATL. ACAD. SCI. (1983) 80: 21] 참고). 또한, 세균 프로모터는 세균 RNA 중합효소를 결합시킬 수 있고 전사를 개시할 수 있는 능력을 갖는 비세균 기원의 천연 발생 프로모터를 포함할 수 있다. 또한, 비세균 기원의 천연 발생 프로모터를 상용가능한 RNA 중합효소와 결합하여, 원핵생물에서 일부 유전자를 높은 수준으로 발현시킬 수 있다. 박테리오파지 T7 RNA 중합효소/프로모터 시스템은 결합된 프로모터 시스템의 일례이다[Studier et al., J. MOL. BIOL. (1986) 189: 113; Tabor et al., Proc Natl. Acad. Sci. (1985) 82: 1074]. 게다가, 혼성 프로모터는 또한 박테리오파지 프로모터 및 대장균 오퍼레이터 영역을 포함할 수 있다(유럽 특허 공보 제267 851호 참고).

[0141] 기능 프로모터 서열 외에도, 효과적인 리보솜 결합 부위는 또한 원핵생물에서의 외래 유전자의 발현에 유용하다. 대장균에서, 리보솜 결합 부위는 샤인-달가노(Shine-Dalgarno; SD) 서열로 불리고, 개시 코돈(ATG) 및 개시 코돈의 3-11 뉴클레오티드 상류에 위치한 길이가 3-9개의 뉴클레오티드인 서열을 포함한다[Shine et al., NATURE (1975) 254: 34]. 이러한 SD 서열은 대장균 16SrRNA의 3' 말단과 SD 서열 사이의 염기 짝지움에 의해 mRNA의 리보솜에 대한 결합을 촉진시키는 것으로 사료된다[Steitz et al. "Genetic signals and nucleotide sequences in messenger RNA", In Biological Regulation and Development: Gene Expression (Ed. R. F. Goldberger, 1979)]. 약한 리보솜 결합 부위를 사용하여 진핵생물 유전자 및 원핵생물 유전자를 발현시키는 것에 대해서는, [Sambrook et al. "Expression of cloned genes in Escherichia coli", Molecular Cloning: A Laboratory Manual, 1989]을 참고한다.

[0142] 용어 "세균 숙주" 또는 "세균 숙주 세포"는 재조합 벡터 또는 다른 전이 DNA에 대한 수용물에 사용될 수 있거나 사용된 세균을 지칭한다. 이 용어는 트랜스펙션된 원래의 세균 숙주 세포의 후손을 포함한다. 단일 모 세포의 후손이 우발적 돌연변이 또는 의도적 돌연변이로 인해 원래의 모체와 형상 또는 총 DNA 보체가 완전히 동일하지 않을 수 있을 것으로 이해된다. 관련 성질, 예를 들어 hIFN 폴리펩티드를 코딩하는 뉴클레오티드의 존재에 의해 특성화되는 모체와 충분히 유사한 모 세포의 후손은 이 정의에 의해 의도되는 후손에 포함된다.

[0143] 폴리펩티드의 발현에 적당한 숙주 세균의 선택은 당업자에게 공지되어 있다. 발현을 위해 세균 숙주를 선택함에 있어서, 적당한 숙주는 그 중에서도, 우수한 내포체(inclusion body), 낮은 단백질분해 활성 및 전체적인 강인성을 갖는 것으로 보이는 것들을 포함할 수 있다. 세균 숙주는 일반적으로 캘리포니아 대학(미국 캘리포니아주 버클리 소재) 생체물리학 및 의료 물리학부 인섹트 제네틱 스톡 센터(Bacterial Genetic Stock Center) 및 아메리칸 타입 컬처 콜렉션(ATCC) (미국 버지니아주 만나사스 소재)을 포함하나 이에 한정되지 않은 다양한 공급원으로부터 입수가 가능하다. 공업적/제약상 발효는 일반적으로 K 균주(예컨대, W3110)로부터 유도된 세균 또는 B 균주 (예컨대, BL21)로부터 유도된 세균을 사용한다. 이러한 균주들은 그들의 성장 매개변수가 널리 공지되어 있고 강건하기 때문에 특히 유용하다. 또한, 이러한 균주는 비병원성이며, 이 점은 안정성 및 환경상의 이유로 상업적으로 중요하다. 적당한 대장균 숙주의 다른 예는 BL21, DH10B 또는 이의 유도체를 포함하나 이에 한정되지 않는다. 본 발명의 방법의 또 다른 실시양태에서, 대장균 숙주는 OMP- 및 LON-를 포함하나 이에 한정되지 않은 프로테아제 마이너스(minus) 균주이다. 숙주 세포 균주는 슈도모나스 플루오레센스, 슈도모나스 아에루기노사, 및 슈도모나스 푸티다 종을 포함하나 이에 한정되지 않은 슈도모나스이다. 균주 MB101로 지정된 슈도모나스 플루오레센스 생물 변이형 1은 재조합 생성에 유용한 것으로 공지되어 있고, 치료적 단백질의 생산 공정에 이용 가능하다. 슈도모나스 발현 시스템의 예는 숙주 균주로서 다우 케미칼 캄파니(Dow Chemical Company) (미국 미시간주 미드랜드 소재, 월드 와이드 웹 dow.com 상에서 입수가 가능)에 의해 입수가 가능한 시스템을 포함한다. 본원에 인용되는 미국 특허 제4,755,465호 및 제4,859,600호는 hGH 생산을 위한 숙주 세포로서 슈도모나스 균주를 사용하는 것을 기재하고 있다.

[0144] 일단 재조합 숙주 세포 균주를 확립하면(즉, 발현 구축물은 숙주 세포로 도입되었고, 적당한 발현 구축물을 갖

는 숙주 세포가 단리되면), 재조합 숙주 세포 균주를 관심 폴리펩티드 생산에 적당한 조건 하에서 배양한다. 당업자에게 자명한 바와 같이, 재조합 숙주 세포 균주의 배양 방법은 사용되는 발현 구축물의 성질 및 숙주 세포의 실체에 따라 좌우되게 된다. 재조합 숙주 균주는 정상적으로 당업계에 공지된 방법을 사용하여 배양된다. 재조합 숙주 세포는 전형적으로 탄소, 질소 및 무기 염의 동화가능한 (assimilatable) 공급원, 및 임의적으로는 비타민, 아미노산, 성장 인자 및 당업계에 공지된 다른 단백질 배양 보충물을 함유하는 액체 배지 중에서 배양된다. 임의적으로, 숙주 세포 배양용 액체 배지는 바람직하지 못한 미생물의 성장을 방지하는 항생제 또는 항진균제, 및/또는 발현 벡터를 함유하는 숙주 세포의 선택을 위한 항생제를 포함하나 이에 한정되지 않는 화합물을 함유할 수 있다.

[0145] 재조합 숙주 세포는 배치 형식 또는 연속 형식으로 배양될 수 있으며, 배치 형식 또는 연속 형식으로 세포를 수확하거나(관심 폴리펩티드가 세포 내 축적되는 경우), 배양 상청액을 수확한다. 원핵생물 숙주 세포에서 생산하는 경우, 배치 배양 및 세포 수확이 바람직하다.

[0146] 셀렉터 코돈

[0147] 본 발명의 셀렉터 코돈은 단백질 생합성 세포 기전의 유전자 코돈 골격 (framework)을 확장시킨다. 예를 들어, 셀렉터 코돈은 독특한 3개의 염기 코돈, 넨센스 코돈, 예컨대 앰버 코돈(UAG) 또는 오팔 코돈(UGA)을 포함하나 이에 한정되지 않는 종결 코돈, 비천연 코돈, 4개(또는 그 이상 개수)의 염기 코돈, 희귀 코돈 등을 포함한다. 수많은 셀렉터 코돈은 원하는 유전자, 또는 폴리뉴클레오티드, 예컨대 1개 이상, 2개 이상, 3개 이상 등에 도입될 수 있다.

[0148] 한 실시양태에서, 방법은 선택된 아미노산, 예컨대 비천연 아미노산을 생체 내 혼입시키기 위해 종결 코돈인 셀렉터 코돈을 사용하는 것을 포함한다. 예를 들어, 종결 코돈을 인지하는 O-tRNA가 생산되고, 선택된 아미노산을 사용하여 O-RS에 의해 아미노아실화된다. 이 O-tRNA는 천연 발생 숙주의 아미노아실-tRNA 합성효소에 의해 인식되지 않는다. 통상적인 부위-지정 돌연변이유발을 사용하여, 관심 폴리펩티드의 관심 부위에 종결 코돈을 도입할 수 있다. 예를 들어, [Sayers, J.R., et al. (1988), *5',3' Exonuclease in a phosphorothioate-based oligonucleotide-directed mutagenesis. Nucleic Acids Res.*, 16: 791-802]을 참고한다. 관심 폴리펩티드를 코딩하는 O-RS, O-tRNA 및 핵산이 예컨대 생체 내에서 조합되는 경우, 선택된 아미노산은 종결 코돈에 대한 반응으로, 혼입되어 특수화된 위치에서 선택된 아미노산, 예컨대 비천연 아미노산을 함유하는 폴리펩티드를 제공한다. 본 발명의 한 실시양태에서, 셀렉터 코돈으로 사용되는 종결 코돈은 앰버 코돈, UAG, 및/또는 오팔 코돈, UGA이다. 예를 들어, 앰버 코돈을 인식하는 O-tRNA의 한 예에 대해 서열 번호 6을 참고하고, 오팔 코돈을 인식하는 O-tRNA의 한 예에 대해 서열 번호 7을 참고한다. UAG 및 UGA의 양자 모두가 셀렉터 코돈으로 사용되는 유전 코드는 가장 풍부한 말단 신호인 오키 넨센스 코돈, UAA를 보존하면서 22개 아미노산을 코딩할 수 있다.

[0149] 생체 내 선택된 아미노산, 예컨대 비천연 아미노산의 혼입은 숙주 세포의 큰 혼란(perturbation) 없이 수행될 수 있다. 예를 들어, UAG 코돈에 대한 억제 효율은 앰버 억제자 tRNA를 포함하나 이에 한정되지 않은 O-tRNA와 UAG 코돈에 결합하여, 리보솜으로부터 성장 펩티드의 방출을 개시하는 방출 인자(RF1) 사이의 경쟁에 의존하기 때문에, 억제 효율은 O-tRNA, 예컨대 억제자 tRNA의 발현 수준을 증가시키거나, RF1 결핍 균주를 이용함으로써 조절될 수 있다. 진핵생물 세포에서, UAG 코돈에 대한 억제 효율이 O-tRNA, 예컨대 앰버 억제자 tRNA와 (종결 코돈에 결합하고, 리보솜으로부터 성장 펩티드의 방출을 개시하는) 진핵생물 방출 인자(예컨대, eRF) 사이의 경쟁에 의존하기 때문에, 억제 효율은 예를 들어 O-tRNA, 예컨대 억제자 tRNA의 발현 수준을 증가시킴으로써 조절될 수 있다.

[0150] 비천연 아미노산은 또한 희귀 코돈으로 코딩될 수 있다. 예를 들어, 시험관 내 단백질 합성 반응에서의 아르기닌 농도가 감소하면, 희귀 아르기닌 코돈인 AGG가 알라닌으로 아실화된 합성 tRNA에 의해 Ala가 삽입되는데 효율적인 것으로 입증되었다. 예컨대 [Ma et al., *Biochemistry*, 32:7939 (1993)]을 참고한다. 이 경우, 합성 tRNA는 *에세리키아 콜라이* 내 소수종으로 존재하는 천연 발생 tRNAArg와 경쟁한다. 일부 유기체는 모든 삼중 코돈을 이용하지 않는다. *마이크로코커스 루테우스(Micrococcus luteus)* 내의 비지정 코돈 AGA는 시험관 내 전사/번역 추출물 내 아미노산의 삽입을 위해 이용되었다. 예컨대 [Kowal and Oliver, *Nucl. Acid. Res.*, 25:4685 (1997)]을 참고한다. 생체 내 이 희귀 코돈을 이용하는 본 발명의 성분들이 발생될 수 있다.

[0151] 셀렉터 코돈은 또한 4개 이상의 염기 코돈, 예를 들어, 4개, 5개, 6개 이상의 염기 코돈을 포함하나 이에 한정되지 않은 연장된 코돈을 포함한다. 4개의 염기 코돈의 예는 AGGA, CUAG, UAGA, CCCU 등을 포함하나 이에 한정되지 않는다. 5개의 염기 코돈의 예는 AGGAC, CCCC, CCCUC, CUAGA, CUACU, UAGGC 등을 포함하나 이에 한정되지 않는다. 한 특성은 프레임쉬프트(frameshift) 억제를 기초로 하여 연장된 코돈을 사용하는 것을 포함할 수

있다. 4개 이상의 염기 코돈은 비천연 아미노산을 포함하나 이에 한정되지 않은 하나 또는 다수의 선택된 아미노산을 동일한 단백질에 삽입시킬 수 있다. 예를 들어, 돌연변이화된 O-tRNA, 예컨대 안티코돈 루프, 예를 들어 CU(X)<sub>n</sub> XXXAA 서열(여기에서, n=1)을 갖는 특수 프레임쉬프트 억제자 tRNA의 존재 하에서, 4개 이상의 염기 코돈을 단일 아미노산으로서 판독한다. 4개 염기 코돈을 인식하는 O-tRNA에 대해 예를 들어, PCT/US04/22061의 서열 번호 6 및 12를 참고한다. 다른 실시양태에서, 안티코돈 루프는 4개 이상의 염기 코돈, 5개 이상의 염기 코돈, 또는 6개 이상의 염기 코돈 또는 그 이상을 포함하나 이에 한정되지 않은 것들을 디코딩(decoding)할 수 있다. 256개의 가능한 4개 염기 코돈이 존재하기 때문에, 다수의 비천연 아미노산은 4개 이상의 염기 코돈을 사용하여 동일한 세포에서 코딩될 수 있다. [Anderson et al., (2002) *Exploring the Limits of Codon and Anticodon Size, Chemistry and Biology*, 9: 237-244; 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 참고한다.

[0152] 예를 들어, 4-염기 코돈은 시험관 내 생합성 방법을 사용하여 비천연 아미노산을 단백질에 혼입시키는데 사용되었다. 예컨대, [Ma et al. (1993) *Biochemistry*, 32: 7939; 및 Holsaka et al. (1999) *J. Am. Chem. Soc.*, 121: 34]을 참고한다. CGGG 및 AGGU는 2-나프틸알라닌 및 리신의 NBD 유도체를 2개의 화학적으로 아실화된 프레임쉬프트억제자 tRNA를 사용하여 동시에 스트렙타비딘(streptavidin)에 시험관 내 혼입시키는데 사용되었다. 예를 들어, [Holsaka et al. (1999) *J. Am. Chem. Soc.*, 121: 12194]을 참고한다. 생체 내 연구에서, Moore 등은 UAGN 코돈 (N은 U, A, G 또는 C일 수 있음)을 억제하는 NCUA 안티코돈을 갖는 tRNA<sup>Leu</sup> 유도체의 능력을 조사하였고, 4중체 UAGA가UCUA 안티코돈을 갖는 tRNA<sup>Leu</sup>에 의해, 0 또는 -1 프레임(frame)에서 약간 디코딩하면서 13 내지 26%의 효율로 디코딩될 수 있다는 것을 발견하였다. [Moore et al., (2000) *J. Mol. Biol.*, 298: 195]을 참고한다. 한 실시양태에서, 희귀 코돈 또는 넌센스 코돈을 기초로 하는 연장된 코돈이 본 발명에 사용될 수 있으며, 이는 다른 원치 않는 부위에서의 미스센스 리드쓰루(missense readthrough) 및 프레임쉬프트 억제를 감소시킬 수 있다.

[0153] 소정의 시스템에 대해, 셀렉터 코돈은 또한 3개의 천연 염기 코돈 중 하나를 포함할 수 있으며, 여기에서 내인성 시스템은 천연 염기 코돈을 전혀 (또는 거의) 사용하지 않는다. 예를 들어, 이는 3개의 천연 염기 코돈을 인식하는 tRNA가 결여된 시스템 및/또는 3개의 염기 코돈이 희귀 코돈인 시스템을 포함한다.

[0154] 셀렉터 코돈은 임의적으로 비천연 염기쌍을 포함한다. 이러한 비천연 염기쌍은 현존하는 유전자 알파벳을 추가로 연장시킨다. 한 여분의 염기쌍은 64 내지 125의 3중 코돈의 수를 증가시킨다. 세 번째 염기쌍의 성질은 안정한 선택적 염기쌍화, 중합효소에 의해 고 적합도로의 DNA로의 효과적인 효소 혼입, 및 초기 (nascent) 비천연 염기쌍의 합성 후 효과적인 연속되는 프라이머 연장을 포함한다. 본 방법 및 조성물에 적합될 수 있는 비천연 염기쌍의 설명은 예를 들어, [Hirao, et al., (2002) *An unnatural base pair for incorporating amino acid analogues into protein*, *Nature Biotechnology*, 20: 177-182]을 포함한다. 또한 [Wu, Y., et al., (2002) *J. Am. Chem. Soc.*, 124:14626-14630]을 참고한다. 다른 관련 공보는 하기에 열거된다.

[0155] 생체 내 사용에 대해, 비천연 뉴클레오타이드는 막 투과성을 갖고, 인산화되어 상응하는 삼인산염을 형성한다. 또한, 증가된 유전 정보는 안정하고, 세포 효소에 의해 파괴되지 않는다. Benner 등에 의해 이루어진 기존 노력에서, 정규 왓슨-크릭(Watson-Crick) 쌍의 경우와 상이한 수소 결합 패턴을 이용하였고, 이 중 가장 주목할만한 예가 이소-C:이소-G 쌍이다. 예를 들어, [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 참고한다. 일반적으로, 이들 염기는 어느 정도 천연 염기와 잘못 쌍을 이루고, 효소에 의해 복제될 수 없다. Kool 및 그의 동료들은 염기들 사이의 소수성 패킹(packing) 상호작용이 수소 결합을 대체하여 염기쌍의 형성을 유도할 수 있다는 점을 입증하였다. [Kool, (2000) *Curr. Opin. Chem. Biol.*, 4: 602; 및 Guckian and Kool (1998) *Angew. Chem. Int. Ed. Engl.*, 36, 2825]을 참고한다. 모든 상기 요건을 만족시키는 비천연 염기쌍을 개발하려는 노력으로, 쉘츠(Schultz), 로메스버그(Romesberg) 및 동료들은 일련의 비천연 소수성 염기들을 체계적으로 합성하고 연구하였다. PICS:PICS 자가 쌍(self-pair)은 천연 염기쌍보다 안정한 것으로 발견되었으며, 에세리키아 콜라이 DNA 중합효소 I(KF)의 클레노우(Klenow) 단편에 의해 DNA로 효율적으로 혼입될 수 있다. 예를 들어, [McMinn et al. (1999) *J. Am. Chem. Soc.*, 121: 11586; 및 Ogawa et al., (2000) *J. Am. Chem. Soc.*, 122: 3274]을 참고한다. 3MN:3MN 자가 쌍은 생물학적 기능에 충분한 효율 및 선택도를 사용하여 KF에 의해 합성될 수 있다. 예를 들어, [Ogawa et al., (2000) *J. Am. Chem. Soc.*, 122: 8803]을 참고한다. 그러나, 양자 모두의 염기는 추가 복제시 사슬 종결자로 작용한다. 최근, 돌연변이 DNA 중합효소는 PICS 자가 쌍을 복제하는데 사용될 수 있는 것으로 밝혀졌다. 또한, 7AI 자가 쌍도 복제될 수 있다. 예를 들어, [Tae et al., (2001) *J. Am. Chem. Soc.*, 123:

7439]을 참고한다. 또한, 신규한 금속성 염기쌍, Dipic:Py도 개발되었으며, 이는 Cu (II) 결합시 안정한 쌍을 형성한다. [Meggers et al., (2000) *J. Am. Chem. Soc.*, 122: 10714]을 참고한다. 연장된 코돈 및 비천연 코돈은 본질적으로 천연 코돈에 오르소고달하기 때문에, 본 발명의 방법은 이 성질을 이용하여 이들에 대한 오르소고달 tRNA를 발생시킬 수 있다.

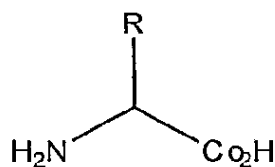
[0156] 번역 우회(translational bypassing) 시스템을 또한 사용하여, 원하는 폴리펩티드에서 선택된 아미노산, 예컨대 비천연 아미노산을 혼입시킬 수 있다. 번역 우회 시스템에서, 대형 서열은 유전자로 혼입되지만, 단백질로 번역되지 않는다. 이 서열은 리보솜을 유도하여 서열을 뛰어넘고(hop over), 삽입의 번역 하류를 재개하는 계기로서 작용하는 구조를 가진다.

[0157] 선택된 아미노산 및 비천연 아미노산

[0158] 본원에 사용되는 선택된 아미노산은 임의의 원하는 천연 발생 아미노산 또는 비천연 아미노산을 지칭한다. 천연 발생 아미노산은 20가지 유전적으로 코딩된 알파-아미노산들, 즉 알라닌, 아르기닌, 아스파라긴, 아스파르트산, 시스테인, 글루타민, 글루탐산, 글리신, 히스티딘, 이소류신, 류신, 리신, 메티오닌, 페닐알라닌, 프롤린, 세린, 트레오닌, 트립토판, 티로신 및 발린 중 임의의 것을 포함한다. 한 실시양태에서, 선택된 아미노산은 고 적합도로, 예컨대 소정의 셀렉터 코돈에 대해서는 약 70% 초과 효율, 소정의 셀렉터 코돈에 대해서는 75% 초과 효율, 소정의 셀렉터 코돈에 대해서는 약 80% 초과 효율, 소정의 셀렉터 코돈에 대해서는 약 85% 초과 효율, 소정의 셀렉터 코돈에 대해서는 약 90% 초과 효율, 소정의 셀렉터 코돈에 대해서는 약 95% 초과 효율, 또는 소정의 셀렉터 코돈에 대해서는 약 99% 초과, 또는 그 이상의 효율로 성장 폴리펩티드 사슬에 혼입된다.

[0159] 본원에 사용되는 비천연 아미노산은 셀레노시스테인 및/또는 피롤리신, 및 하기 20가지 유전적으로 코딩된 알파-아미노산들 이외의 임의의 아미노산, 변형 아미노산, 또는 아미노산 유사체를 지칭한다: 알라닌, 아르기닌, 아스파라긴, 아스파르트산, 시스테인, 글루타민, 글루탐산, 글리신, 히스티딘, 이소류신, 류신, 리신, 메티오닌, 페닐알라닌, 프롤린, 세린, 트레오닌, 트립토판, 티로신, 발린. 알파-아미노산의 일반 구조는 화학식 I로 도시된다:

### 화학식 I



[0160]

[0161] 비천연 아미노산은 전형적으로 R 기가 20가지 천연 아미노산에 사용된 치환기 이외의 임의의 치환기인 화학식 I을 갖는 임의의 구조이다. 20가지 천연 아미노산의 구조에 대해, 예컨대 [Biochemistry, L. Stryer 저, 제3판, 1988, Freeman and Company, New York]을 참고한다. 본 발명의 비천연 아미노산은 상기 20가지 알파-아미노산 이외의 상기 천연 발생 화합물일 수 있음을 주목한다.

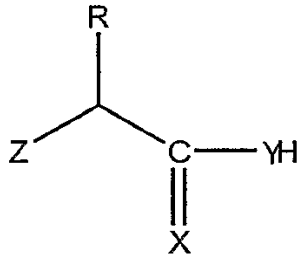
[0162] 본 발명의 비천연 아미노산은 전형적으로 측쇄의 구조에서만 천연 아미노산과 상이하고, 비천연 아미노산은 천연 발생 단백질에서 형성되게 되는 동일한 방식으로, 천연 또는 비천연(이를 포함하나 이에 한정되지는 않음)의 다른 아미노산과 아미드 결합을 형성한다. 그러나, 비천연 아미노산은 천연 아미노산과 구분되도록 하는 측쇄기를 가진다. 예를 들어, 화학식 I 내의 R은 알킬-, 아릴-, 아실-, 케토-, 아지도-, 히드록실-, 히드라진, 시아노-, 할로-, 히드라지드, 알케닐, 알킬닐, 에테르, 티올, 셀레노-, 술폰-, 보레이트, 보로네이트, 포스포, 포스포노, 포스핀, 헤테로시클릭, 에논(enone), 이민, 알데히드, 에스테르, 티오산, 히드록실아민, 아민 등, 또는 이들의 임의의 조합을 포함할 수 있다. 다른 비천연 발생 아미노산은 광활성화가능한 가교제를 포함하는 아미노산, 스핀 표지화된 아미노산, 형광 아미노산, 금속 결합 아미노산, 금속 함유 아미노산, 방사능 아미노산, 신규한 작용기를 갖는 아미노산, 다른 분자와 공유적으로 또는 비공유적으로 상호작용하는 아미노산, 포토케이지되고/되거나 광이성화가능한 아미노산, 비오틴 또는 비오틴 유사체를 포함하는 아미노산, 글리코실화된 아미노산, 예컨대 당 치환 세린, 기타 탄수화물 변형 아미노산, 케토 함유 아미노산, 폴리에틸렌 글리콜 또는 폴리에테르를 포함하는 아미노산, 중 원자로 치환된 아미노산, 화학 분절성 및/또는 광분절성 아미노산, 약 5개 초과 또는 약 10개 초과 탄소를 포함하나 이에 한정되지 않은 장쇄 탄화수소 또는 폴리에테르를 포함하나 이에 한정되지 않은 천연 아미노산에 비해 연장된 측쇄를 갖는 아미노산, 탄소가 연결된 당 함유 아미노산, 산화환원 활성 아



미노산, 아미노 티오산 함유 아미노산, 및 하나 이상의 독성 부분을 포함하는 아미노산을 포함하나 이에 한정되지 않는다. 또한, 본원에 참고로 인용되는 미국 특허 출원 공보 제2003/0082575호 및 제2003/0108885호를 참고한다. 비천연 아미노산은 예컨대 고체 지지체에 단백질을 연결시키기 위해 사용되는 광활성화가능한 가교제를 가질 수 있다. 비천연 아미노산은 아미노산 측쇄에 부착된 당류(saccharide) 부분을 가질 수 있다.

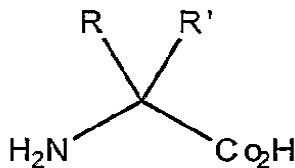
[0163] 신규한 측쇄를 함유하는 비천연 아미노산 외에도, 비천연 아미노산은 또한 변형된 골격 구조로서, 예컨대 하기 화학식 II 및 화학식 III의 구조에 의해 도시되는 구조를 포함한다:

### 화학식 II



[0164]

### 화학식 III



[0165]

[0166] (식 중, Z는 전형적으로 OH, NH<sub>2</sub>, SH, NH-R' 또는 S-R'를 포함하고, 동일하거나 상이할 수 있는 X 및 Y는 전형적으로 S 또는 O를 포함하고, 임의적으로 동일하거나 상이할 수 있는 R 및 R'는 전형적으로 화학식 I을 갖는 비천연 아미노산에 대해 상기한 R 기에 대한 구성성분으로 된 동일한 열거 성분들, 및 수소로부터 선택된다). 예를 들어, 비천연 아미노산은 화학식 II 및 화학식 III에 도시되는 바와 같이 아미노 또는 카르복실기에서의 치환을 포함할 수 있다. 이 유형의 비천연 아미노산은 예컨대 공통의 20개의 천연 아미노산에 상응하는 측쇄 또는 비천연 측쇄를 갖는, α-히드록시 산, α-티오산, α-아미노티오카르복실레이트를 포함하나 이에 한정되지 않는다. 또한, α-탄소에서의 치환은 임의적으로 L, D 또는 α-α-이치환 아미노산, 예를 들어 D-글루타메이트, D-알라닌, D-메틸-0-티로신, 아미노부티르산 등을 포함하나 이에 한정되지 않는다. 다른 구조적 대체물은 시클릭 아미노산, 예를 들어 프롤린 유사체, 및 3, 4, 6, 7, 8 및 9원 고리 프롤린 유사체, β 및 γ 아미노산, 예를 들어 치환 β-알라닌 및 γ-아미노 부티르산을 포함한다.

[0167] 많은 비천연 아미노산은 천연 아미노산, 예를 들어 티로신, 글루타민, 페닐알라닌 등을 기초로 한다. 티로신 유사체는 파라-치환된 티로신, 오르토-치환된 티로신 및 메타 치환된 티로신을 포함하고, 여기에서 치환된 티로신은 케토기(아세틸 기를 포함하나 이에 한정되지 않음), 벤조일기, 아미노기, 히드라진, 히드록시아민, 티올기, 카르복시기, 이소프로필기, 메틸기, C<sub>6</sub>-C<sub>20</sub> 직쇄형 사슬 또는 분지형 탄화수소, 포화된 또는 불포화된 탄화수소, 0-메틸기, 폴리에테르기, 니트로기, 알킬닐기 등을 포함한다. 또한, 다중 치환된 아릴 고리도 고려된다. 글루타민 유사체는 α-히드록시 유도체, γ-치환된 유도체, 시클릭 유도체 및 아미드 치환된 글루타민 유도체를 포함하나 이에 한정되지 않는다. 페닐알라닌 유사체의 예는 파라-치환된 페닐알라닌, 오르토-치환된 페닐알라닌, 및 메타-치환된 페닐알라닌을 포함하나 이에 한정되지 않으며, 여기에서 치환기는 히드록시기, 메톡시기, 메틸기, 알릴기, 알데히드, 아지도, 요오도, 브로모, 케토기(아세틸 기를 포함하나 이에 한정되지 않음) 등을 포함하나 이에 한정되지 않은 것들을 포함한다. 본 발명에 사용하기 적합할 수 있는 비천연 아미노산의 구체적인 예는 p-아세틸-L-페닐알라닌, p-프로파르길-페닐알라닌, 0-메틸-L-티로신, L-3-(2-나프틸)알라닌, 3-메틸-페닐알라닌, 0-4-알릴-L-티로신, 4-프로필-L-티로신, 트리-0-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 및 p-프로파르길옥시-페닐알라닌 등을 포함하나 이에 한정되지 않는다. 다양한 비천연 아미노산의 구조의 예는 예를 들어, WO 2002/085923(발명의 명칭: "In vivo incorporation of unnatural

amino acids")에 제공되어 있다. 또한, 추가적인 메티오닌 유사체에 대해서는 [Kiick et al., (2002) *Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation*, PNAS 99: 19-24]을 참고한다.

[0168] 아미노 말단에서 폴리펩티드로 혼입되는 비천연 아미노산은  $\alpha$ -아미노산 중에 정상적으로 존재하는  $\text{NH}_2$  기와 상이한 제2 반응성 기 및 20개의 천연 아미노산에서 사용된 것 이외의 임의의 치환기인 R 기로 구성될 수 있다(화학식 I 참고). 유사한 비천연 아미노산은  $\alpha$ -아미노산 중에 정상적으로 존재하는  $\text{COOH}$  기와 상이한 제2 반응성 기를 갖는 카르복실 말단에서 혼입될 수 있다(화학식 I 참고).

[0169] 본 발명의 비천연 아미노산은 20가지 천연 아미노산에서 이용불가능한 부가적 특성을 제공하도록 선택되거나 설계될 수 있다. 예를 들어, 비천연 아미노산은 예컨대 그 아미노산이 혼입되게 되는 단백질의 생물학적 성질을 변형하도록 임의적으로 설계되거나 선택될 수 있다. 예를 들어 비천연 아미노산을 단백질에 포함시킴으로써 하기 성질들이 임의적으로 변형될 수 있다: 독성, 생체내분포, 용해도, 안정성, 예컨대 열, 가수분해, 산화, 효소 분해에 대한 내성, 화학적 및/또는 광화학적 성질, 산화환원 포텐셜, 반감기, 다른 분자와 예컨대 공유적으로 또는 비공유적으로 반응할 수 있는 능력 등.

[0170] 다양한 비천연 아미노산의 구조들이 예를 들어, 본원에 참고로 인용되는 WO 2002/085923(발명의 명칭: "In vivo incorporation of unnatural amino acids")의 도 16, 17, 18, 19, 26 및 29에 제공되어 있다. 그 예는 본 발명의 tRNA에 부착될 수 있는 아미노산에 결코 제한되는 것을 의미하지는 않는다.

[0171] 비천연 아미노산의 한 이점은, 그것이 부가적 분자를 부가하기 위해 사용될 수 있는 부가적 화학적 부분을 제공한다는 것이다. 이 변형은 진핵생물 또는 비진핵생물 세포에서 생체 내, 또는 시험관 내 이루어질 수 있다. 따라서, 특정 실시양태에서, 번역후 변형은 비천연 아미노산을 통해 이루어진다. 폴리펩티드 내의 비천연 아미노산은, 특정 반응성 기에 적당한 것으로 당업자에게 공지되어 있는 화학 방법론을 사용하여 제1 반응성 기를 포함하는 하나 이상의 비천연 아미노산에 대해 제2 반응기를 포함하는, 표지, 염료, 중합체, 수용성 중합체, 폴리 에틸렌 글리콜의 유도체, 광가교제, 방사성 핵종, 세포독성 화합물, 약물, 친화성 표지, 광친화성 표지, 반응성 화합물, 수지, 제2 단백질 또는 폴리펩티드 또는 폴리펩티드 유사체, 항체 또는 항체 단편, 금속 킬레이트제, 보조인자, 지방산, 탄수화물, 폴리뉴클레오티드, DNA, RNA, 안티센스 폴리뉴클레오티드, 당류, 수용성 덴드리머, 시클로덱스트린, 억제성 리보핵산, 생물소재, 나노입자, 스핀 표지, 형광발색단, 금속 함유 부분, 방사능 부분, 신규한 작용기, 다른 분자와 공유적으로 또는 비공유적으로 상호작용하는 기, 포토케이징된 부분, 화학선 복사 여기성 부분, 광이성화가능한 부분, 비오틴, 비오틴의 유도체, 비오틴 유사체, 중 원자가 혼입된 부분, 화학 분절성 기, 광분절성 기, 신장된 측쇄, 탄소-연결 당, 산화환원 활성제, 아미노 티오산, 독성 부분, 동위 원소로 표지화된 부분, 생물물리학적 프로브, 형광 기, 화학발광기, 전자 밀집 기, 자성 기, 층간삽입성 기, 발색단, 에너지 전달제, 생물학적 활성제, 검출가능한 표지, 소분자, 양자점(quantum dot), 나노트랜스미터(nanotransmitter), 또는 상기 또는 임의의 다른 바람직한 화합물 또는 물질의 임의의 조합물을 포함하나 이에 한정되지 않는 폴리펩티드에 또 다른 분자를 부착시키는 것을 포함한다.

[0172] 예를 들어, 번역후 변형은 친핵성-친전자성 반응을 통해 이루어질 수 있다. 현재 단백질의 선택적 변형에 사용되는 대부분의 반응은  $\alpha$ -할로케톤과 히스티딘 또는 시스테인 측쇄와의 반응을 포함하나 이에 한정되지 않는 친핵성 반응 파트너와 친전자성 반응 파트너 사이의 공유 결합 형성을 포함한다. 이러한 경우에서의 선택도는 단백질 중의 친핵성 잔기의 접근성 및 수에 의해 결정된다. 본 발명의 단백질에서, 다른 더욱 선택적인 반응들, 예를 들어 시험관 내 및 생체 내 비천연 케토-아미노산과 히드라이드 또는 아미노옥시 화합물의 반응을 사용할 수 있다. 예컨대, [Cornish, et al. (1996) J. Am. Chem. Soc., 118: 8150-8151; Mahal, et al. (1997) Science, 276:1125-1128; Wang, et al., (2001) Science 292:498-500; Chin, et al., (2002) J. Am. Chem. Soc. 124:9026-9027; Chin, et al., (2002) Proc. Natl. Acad. Sci. 99: 11020-11024; Wang, et al., (2003) Proc. Natl. Acad. Sci. 100: 56-61; Zhang, et al., (2003) Biochemistry, 42:6735-6746; 및 Chin, et al., (2003) Science, 301: 964-7]를 참고하고, 이들 모두는 본원에 참고로 인용된다. 이는 형광 발색단, 가교제, 당류 유도체 및 세포독성 분자를 비롯한 시약으로 된 숙주를 사용하여 거의 모든 단백질을 선택적으로 표지화할 수 있게 한다. 또한, 본원에 참고로 인용되는 미국 특허 제6,927,042호(발명의 명칭: "Glycoprotein synthesis")를 참고한다. 또한, 아지도 아미노산을 통한(단, 이에 한정되지 않음) 번역후 변형은 스타우딩저 결합(Staudinger ligation) (트리아릴포스핀 시약을 사용하는 것을 포함하나 이에 한정되지 않음)을 통해 이루어질 수 있다. 예를 들어, [Kiick et al., (2002) *Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation*, PNAS 99: 19-24]을 참고한다.

[0173] 비천연 아미노산의 화학적 합성

[0174] 많은 비천연 아미노산들이, 예컨대, 시그마-알드리히(Sigma-Aldrich)(미국 미조리주 세인트루이스 소재), 노바바이오켄(Novabiochem)(EMD 바이오사이언스사 지사, 독일 담슈타트 소재) 또는 펩테크(Peptech)(미국 메사추세츠주 버링톤 소재)로부터 상업적으로 구입가능하다. 시중 입수가능하지 않은 것들은 임의적으로 본원에 제공된 채로 합성되거나, 당업자에게 공지된 표준 방법을 이용하여 합성된다. 유기 합성 기법에 대해, 예컨대 [Organic Chemistry, Fessenden and Fessenden(1982)(제2판, Willard Grant Press, Boston Mass); Advanced Organic Chemistry, March(제3판, 1985, Wiley and Sons, New York); 및 Advanced Organic Chemistry, Carey 및 Sundberg(제3판, 파트 A 및 B, 1990, Plenum Press, New York)]을 참고한다. 비천연 아미노산의 합성을 기술하는 부가적 공보는 예컨대 WO 2002/085923(발명의 명칭: "In vivo incorporation of unnatural amino acids"; Matsoukas et al. (1995) J. Med. Chem., 38, 4660-4669; King, F.E. & Kidd, D. A. A. (1949) *A New Synthesis of Glutamine and of  $\gamma$ -Dipeptides of Glutamic Acid from Phthylated Intermediates*. J. Chem. Soc., 3315-3319; Friedman, O.M. & Chatterji, R. (1959) *Synthesis of Derivatives of Glutamine as Model Substrates for Anti-Tumor Agents*. L Am. Chem. Soc., 81, 3750-3752; Craig, J.C. 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, M., Vilmon, M. & Frappier, F. (1991) *Glutamine analogues as Potential Antimalarials*, Eur. J. Med. Chem., 26, 201-5; Koskinen, A.M.P. & Rapoport, H. (1989) *Synthesis of 4-Substituted Prolines as Conformationally Constrained Amino Acid Analogues*. J. Org. Chem., 54, 1859-1866; Christie, B.D. & Rapoport, H. (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., 50:1239-1246; Barton et al. (1987) *Synthesis of Novel  $\alpha$ -Amino-Acids and Derivatives Using Radical Chemistry: Synthesis of L- and D- $\alpha$ -Amino-Adipic Acids, L- $\alpha$ -aminopimelic Acid and Appropriate Unsaturated Derivatives*. Tetrahedron 43:4297-4308; 및 Subasinghe et al. (1992) *Quisqualic acid analogues: synthesis of beta-heterocyclic 2-aminopropanoic acid derivatives and their activity at a novel quisqualate-sensitized site*. J. Med. Chem., 35:4602-7)을 포함한다. 또한, 본원에 참고로 인용되는 미국 특허 공보 US 제 2004/0198637호(발명의 명칭: "Protein Arrays")를 참고한다.

[0175] 비천연 아미노산의 세포내 흡수

[0176] 세포에 의한 비천연 아미노산 흡수는, 예컨대 단백질로의 혼입을 위해 비천연 아미노산을 설계하고 선택할 때 전형적으로 고려되는 하나의 관건이다. 예를 들어,  $\alpha$ -아미노산의 고전하 밀도는 이 화합물이 세포 투과가 어려울 수 있음을 제시한다. 천연 아미노산은 단백질-기재 수송 시스템의 수집을 통해 세포로 흡수된다. 비천연 아미노산이 있는 경우, 그것이 세포에 의해 흡수되는지를 평가하는 급속 스크린이 행해질 수 있다. 예를 들어, 독성 검정에 대해서는 예컨대 US 제2004/0198637호(발명의 명칭: "Protein Arrays"); 및 Liu, D.R. & Schultz, P.G. (1999) *Progress toward the evolution of an organism with an expanded genetic code*. PNAS United States 96: 4780-4785]을 참고한다. 흡수가 각종 검정을 이용하여 용이하게 분석되나, 세포내 흡수 경로에 수용적인 비천연 아미노산을 설계함에 있어 대안법은, 생체 내 아미노산을 생산하는 생합성 경로를 제공하는 것이다.

[0177] 비천연 아미노산의 생합성

[0178] 많은 생합성 경로가 아미노산 및 기타 화합물의 제조를 위해 세포 내에 이미 존재한다. 특별한 비천연 아미노산을 위한 생합성 방법이 세포 내를 비제한적 예로 포함하는 자연에 존재하지 않을 수 있으나, 본 발명은 그러한 방법을 제공한다. 예를 들어, 비천연 아미노산을 위한 생합성 경로는 임의적으로 새 효소를 첨가하거나, 현존 숙주 세포 경로를 변형시킴으로써 숙주 세포 내에 발생될 수 있다. 부가적 새 효소는 임의적으로 천연 발생 효소 또는 인공 진화 효소이다. (예를 들어, WO 2002/085923(발명의 명칭: "In vivo incorporation of unnatural amino acids")에서 한 예로 제시된) p-아미노페닐알라닌의 생합성은 다른 유기체로부터의 공지 효소들의 조합물을 첨가하는 것에 의존한다. 이 효소들을 위한 유전자는 유전자를 포함하는 플라스미드를 이용하여 세포를 형질 변환시킴으로써 세포에 도입될 수 있다. 유전자는 세포 내에 발현될 때, 요망되는 화합물을 합성하기 위한 효소 경로를 제공한다. 임의적으로 첨가되는 효소의 유형의 예가 이하의 예들에서 제공된다. 부가적 효소 서열이 예를 들어, 유전자은행(Genbank)에 나와 있다. 인공 진화된 효소는 또한 동일한 방식으로 세포에 임의적으로 첨가된다. 이 방식으로, 세포 기전 및 세포의 리소스를 조작하여, 비천연 아미노산을 생산한다.



- [0179] 생합성 경로에서의 사용, 또는 현존 경로의 진화를 위한 신규 효소를 생산하기 위해 각종 방법들이 이용가능하다. 예를 들어, 맥시젠 인코포레이티드(Maxygen, Inc.)(월드 와이드 웹 maxygen.com에서 이용가능함)에 의해 개발된(단, 이에 제한되지 않음) 재귀적 재조합을 임의적으로 이용하여, 신규 효소 및 경로를 개발한다. 예컨대, [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]를 참고한다. 유사하게, 지넨코어(Genencor)(월드 와이드 웹 genencor.com에서 이용가능함)에 의해 개발된 디자인패스(DesignPath)<sup>TM</sup>를 임의적으로 대사 경로로 공학처리하기 위해, 세포 내 O-메틸-L-티로신을 생성시키기 위한 경로를 공학처리한다(단, 이에 제한되지 않음). 이 기술은 작용적 유전체학, 및 분자 진화 및 설계를 통해 확인된 유전자들(단, 이에 제한되지 않음)을 포함하나 이에 한정되지 않는 새 유전자들의 조합을 이용하여 숙주 유기체 내의 현존 경로를 재구축한다. 다이벌사 코포레이션(Diversa Corporation)(월드 와이드 웹 diversa.com에서 이용가능함)은 또한 새 경로를 발생시키기 위한(단, 이에 제한되지 않음) 유전자의 라이브러리 및 유전자 경로를 급속히 선별하는 기술을 제공한다.
- [0180] 전형적으로, 본 발명의 공학처리된 생합성 경로를 이용하여 생산된 비천연 아미노산은, 천연 세포내 양을 비제한적 예로 포함하는, 효율적인 단백질 생합성을 위해 충분한 농도로, 다만 다른 아미노산의 농도에 영향을 주거나, 세포 리소스를 소진하는 정도로는 아닌 농도로 생산된다. 이러한 식으로 생체 내 생산된 전형적 농도는 약 10 mM 내지 약 0.05 mM이다. 일단 세포가 특정 경로 및 비천연 아미노산에 요망되는 효소를 생산하기 위해 사용되는 유전자를 포함하는 플라스미드를 이용하여 형질전환되면, 생체 내 선택을 임의적으로 사용하여, 리보솜내 단백질 합성 및 세포 성장 모두를 위한 비천연 아미노산의 생산을 더욱 최적화한다.
- [0181] 핵산 및 폴리펩티드 서열 및 변이체
- [0182] 상기 및 하기에 기재된 바와 같이, 본 발명은 핵산 폴리뉴클레오티드 서열 및 폴리펩티드 아미노산 서열, 예컨대 tRNA 및 RS, 및 예컨대 상기 서열을 포함하는 조성물 및 방법을 제공한다. 상기 서열, 예컨대 tRNA 및 RS의 예가 본원에 개시된다. 그러나, 당업자는 본 발명이 본원, 예컨대 실시예에 개시된 서열에 한정되지 않음을 인식할 것이다. 당업자는 본 발명이 또한 본원에 기재된 기능을 갖는, 예컨대 O-tRNA 또는 O-RS를 코딩하는 많은 관련 및 비관련 서열을 제공함을 인식할 것이다.
- [0183] 본 발명은 폴리펩티드(O-RS) 및 폴리뉴클레오티드, 예컨대 O-tRNA, O-RS 또는 이의 일부를 코딩하는 폴리뉴클레오티드, 아미노아실-tRNA 합성효소 클론을 단리하는데 사용되는 올리고뉴클레오티드 등을 제공한다. 본 발명의 폴리뉴클레오티드는 하나 이상의 선택적 코돈을 갖는 본 발명의 관심 단백질 또는 폴리펩티드를 코딩하는 것들을 포함한다. 또한, 본 발명의 폴리뉴클레오티드는 예컨대 서열 번호 4에 나와 있는 뉴클레오티드 서열을 포함하는 폴리뉴클레오티드; 이의 폴리뉴클레오티드 서열에 상보적이거나 그것을 코딩하는 폴리뉴클레오티드, 또는 이의 보존적 변이체를 포함한다. 본 발명의 폴리뉴클레오티드는 또한 본 발명의 폴리펩티드를 코딩하는 폴리뉴클레오티드를 포함한다. 유사하게, 핵산의 실질적으로 전체 길이에서 고도로 엄격한 조건 하에서 상기 나타난 폴리뉴클레오티드에 혼성화된 핵산은 본 발명의 폴리뉴클레오티드이다. 한 실시양태에서, 조성물은 본 발명의 폴리펩티드 및 부형제(예컨대, 완충액, 물, 약제학적으로 허용가능한 부형제 등)를 포함한다. 또한, 본 발명의 폴리펩티드는 예컨대 서열 번호 5에 나와 있는 아미노산 서열을 포함하는 폴리펩티드; 이의 폴리펩티드 서열에 상보적이거나 그것을 코딩하는 폴리펩티드, 또는 이의 보존적 변이체를 포함한다. 본 발명은 또한 본 발명의 폴리펩티드에 특이적으로 면역반응성인 항체 또는 항혈청을 제공한다.
- [0184] 특정 실시양태에서, 벡터(예컨대, 플라스미드, 코스미드, 파지, 세균, 바이러스, 나형(naked) 폴리뉴클레오티드, 접합된 폴리뉴클레오티드 등)는 본 발명의 폴리뉴클레오티드를 포함한다. 한 실시양태에서, 벡터는 발현 벡터이다. 또 다른 실시양태에서, 발현 벡터는 본 발명의 폴리뉴클레오티드들 중 하나 이상에 작용적으로 연결된 프로모터를 포함한다. 또 다른 실시양태에서, 세포는 본 발명의 폴리뉴클레오티드를 포함하는 벡터를 포함한다.
- [0185] 당업자는 또한 개시된 서열의 많은 변이체들이 본 발명에 포함됨을 인식할 것이다. 예를 들어, 작용적으로 동일한 서열을 산출하는 개시된 서열의 보존적 변이체는 본 발명에 포함된다. 핵산 폴리뉴클레오티드 서열의 변이체로서, 하나 이상의 개시된 서열에 혼성화하는 변이체는 본 발명에 포함되는 것으로 간주된다. 예컨대 표준 서열 비교 기법에 의해 결정되는, 본원에 개시된 서열의 독특한 하위서열도 또한 본 발명에 포함된다.
- [0186] 보존적 변이체

- [0187] 유전 코드의 퇴화로 인해, "침묵 치환"(즉, 코딩된 폴리펩티드에서의 변형을 초래하지 않는 핵산 서열에서의 치환)은 아미노산을 코딩하는 모든 핵산 서열의 함축된 특성이다. 유사하게, 아미노산 서열 내 하나 또는 몇 개의 아미노산에서의 매우 유사한 성질을 갖는 다른 아미노산으로의 "보존적 아미노산 치환"은 또한 개시된 구축물과 매우 유사한 것으로 용이하게 확인된다. 각 개시된 서열의 그러한 보존적 변이체는 본 발명의 한 특징이다.
- [0188] 특별한 핵산 서열의 "보존적 변이체"는 동일하거나 본질적으로 동일한 아미노산 서열을 코딩하는 핵산을, 또는 핵산이 아미노산 서열을 코딩하지 않는 경우에는, 본질적으로 동일한 서열을 지칭한다. 당업자는 코딩된 서열에서의 단일 아미노산 또는 적은 백분율의 아미노산을 변경, 부가 또는 결실시키는 개별적 치환, 결실 또는 부가가 "보존적으로 변형된 변이체" 또는 "보존적으로 변형된 변이체"임을 인식할 것이며, 여기에서 변경은 아미노산의 결실, 아미노산의 부가, 또는 아미노산의 화학적으로 유사한 아미노산으로의 치환을 초래한다. 따라서, 본 발명의 열거된 폴리펩티드 서열의 "보존적 변이체"는, 폴리펩티드 서열 내 아미노산의 동일한 보존적 치환기의 보존적으로 선택된 아미노산으로의, 적은 백분율, 전형적으로는 5% 미만, 더욱 전형적으로는 4%, 2% 또는 1% 미만의 치환을 포함한다. 핵산 분자의 코딩된 활성을 변경시키지 않는 서열의 부가, 예컨대 비작용성 서열의 부가는 기본 핵산의 보존적 변이체이다.
- [0189] 작용적으로 유사한 아미노산을 제공하는 보존적 치환 표는 당업자에게 공지되어 있다. 하기 8개의 군은 각각 서로 보존적 치환인 아미노산을 함유한다.
- [0190] 1) 알라닌(A), 글리신(G);
- [0191] 2) 아스파르트산(D), 글루탐산(E);
- [0192] 3) 아스파라긴(N), 글루타민(Q);
- [0193] 4) 아르기닌(R), 리신(K);
- [0194] 5) 이소류신(I), 류신(L), 메티오닌(M), 발린(V);
- [0195] 6) 페닐알라닌(F), 티로신(Y), 트립토판(W);
- [0196] 7) 세린(S), 트레오닌(T); 및
- [0197] 8) 시스테인(C), 메티오닌(M)
- [0198] (예컨대, [Creighton, *Proteins: Structures and Molecular Properties* (W H Freeman & Co.; 제2판 (December 1993))] 참고)
- [0199] 핵산 혼성화
- [0200] 비교 혼성화를 이용하여, 본 발명의 핵산의 보존적 변이체를 포함하는, 본 발명의 핵산, 예컨대 서열 번호 4를 동정할 수 있고, 이 비교 혼성화 방법은 본 발명의 핵산을 구별하는 한 바람직한 방법이다. 또한, 고도, 초고도, 및/또는 초초고도 엄격성 조건 하에서 서열 번호 4로 표시되는 핵산에 혼성화되는 표적 핵산은 본 발명의 한 특징이다. 그러한 핵산의 예는 소정의 핵산 서열에 비해 1개 또는 몇 개의 침묵 또는 보존적 핵산 치환을 갖는 것들을 포함한다.
- [0201] 시험 핵산은, 완전히 매칭되는 상보적 표적에 대한 혼성화 대비, 프로브에 대해 1/2 이상으로 혼성화할 때, 즉 매칭되지 않는 표적 핵산 중 임의의 것에 대한 혼성화에서 관찰되는 비의 약 5x 내지 10x 이상인 신호 대 노이즈 비로 완전히 매칭되는 프로브가 완전히 매칭되는 상보적 표적에 결합하도록 하는 조건 하에서, 표적에 대한 프로브의 혼성화의 1/2 이상인 신호 대 노이즈 비를 가질 때, 프로브 핵산에 특이적으로 혼성화한다고 일컬어진다.
- [0202] 핵산은 그것이 전형적으로 용매 중 연합할 때, "혼성화"한다. 핵산은 다양한 잘 특징화된 물리화학적 힘, 예컨대 수소 결합, 용매 배제, 염기 적층 등으로 인해 혼성화한다. "엄격한 혼성화 조건"이란 어구는 당업계에 공지된 바와 같이, 낮은 이온 강도와 고온인 조건을 의미한다. 전형적으로, 엄격한 조건하에서 프로브는 복잡한 핵산 혼합물(전 세포 내 또는 라이브러리 DNA 또는 RNA를 포함하나 이에 한정되지 않음) 중에 존재하는 상기 프로브의 표적 서열에 혼성화될 것이지만, 복잡한 혼합물 중에 존재하는 다른 서열과는 혼성화되지 않는다. 핵산의 혼성화에 관한 자세한 지침은 [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 Ausubel, et al,

Current Protocols in Molecular Biology (1995)]에 나와 있다. [Hames and Higgins (1995) Gene Probes 1 IRL Press at Oxford University Press, Oxford, England, (Hames and Higgins 1) 및 Hames and Higgins (1995) Gene Probes 2 IRL Press at Oxford University Press, Oxford, England (Hames and Higgins 2)]는 올리고뉴클레오타이드를 포함하는 DNA 및 RNA 의 합성, 표지화, 검출 및 정량화에 대한 상세 내용을 제공한다. 일반적으로, 엄격한 조건은 한정된 이온 강도 pH에서 특정 서열에 대한 열 융점 온도( $T_m$ )보다 약 5 내지 10°C 낮게 선택된다.  $T_m$ 은 표적에 상보적인 프로브의 50%가 평형인 표적 서열에 혼성화하는 (한정된 이온 강도, pH 및 핵산 농도 하에서의) 온도이다[즉, 표적 서열이  $T_m$ 에서 과량 존재할 때, 프로브의 50%가 평형 상태로 점유됨]. 엄격한 조건이란, pH가 7.0~8.3에서 염 농도가 약 1.0 M 나트륨 이온 미만, 통상적으로는 약 0.01~1.0 M 나트륨 이온(또는 기타 염) 미만이고, 길이가 짧은 프로브(10 내지 50개 뉴클레오타이드의 프로브를 포함하나 이에 한정되지 않음)에 대해서는 온도가 약 30°C 이상이고, 길이가 긴 프로브(50개 이상의 뉴클레오타이드 프로브를 포함하나 이에 한정되지 않음)에 대해서는 온도가 약 60°C 이상인 조건일 수 있다. 엄격한 조건은 또한 탈 안정화 제제(destabilizing agent), 예를 들어 포름아미드를 첨가함으로써 달성할 수 있다. 선택적 또는 특이적 혼성화에 있어서, 양성 시그널은 배경의 2배 이상, 임의로는 배경 혼성화의 10배일 수 있다. 대표적인 엄격한 혼성화 조건은 다음과 같을 수 있다: 50% 포름아미드, 5X SSC, 및 1% SDS, 42°C에서 인큐베이션, 또는 5X SSC, 1% SDS, 65°C에서 인큐베이션, 0.2 X SSC에서 세정, 그리고 0.1% SDS, 65°C. 이때, 세정은 5, 15, 30, 60 또는 120분 이상 동안 수행될 수 있다.

[0203] 써던 또는 노던 블로트 내 필터 상에 100개 초과 상보적 잔기를 가지는 상보적 핵산의 혼성화를 위한 엄격한 혼성화 조건의 한 예는, 42°C에서의 1 mg의 헤파린을 갖는 50% 포르말린이고, 여기에서 혼성화는 하룻밤 동안 수행된다. 엄격한 세정 조건은 15분 동안 65°C에서 0.2x SSC 세정하는 것이다(SSC 완충액의 대한 설명을 위해 [Sambrook, et al, *Molecular Cloning, A Laboratory Manual*(제3판, 2001)]를 참고한다). 종종 고 엄격성 세정 전에, 저 엄격성 세정이 선행되어, 배경 프로브 신호를 제거한다. 한 예로서의 저 엄격성 세정은 15분 동안 40°C에서 2x SSC를 이용하여 세정하는 것이다. 일반적으로, 특별한 혼성화 검정에서 비관련 프로브에 대해 관찰되는 것보다 5x (또는 그 이상) 더 높은 신호 대 노이즈 비는 특정 혼성화의 검출을 가리킨다.

[0204] 서던(Southern) 및 노던(Northern) 혼성화와 같은 핵산 혼성화 실험의 맥락에서의 "엄격한 혼성화 세정 조건"은 서열 의존적이고, 상이한 환경 매개변수 하에서 상이하다. 보다 긴 서열은 특별히 보다 높은 온도에서 혼성화한다. 핵산의 혼성화에 대한 자세한 지침이 [Tijssen (1993), 이하 동문] 및 [Hames 및 Higgins, 1 및 2]에 나와 있다. 엄격한 혼성화 및 세정 조건은 임의의 시험 핵산에 대해 실험적으로 용이하게 결정될 수 있다. 예를 들어, 매우 엄격한 혼성화 및 세정 조건을 결정할 때, 선택된 기준 세트가 충족될 때까지, 혼성화 및 세정 조건을 (예컨대, 혼성화 또는 세정 시에 온도 증가, 염 농도 감소, 세제 농도 증가, 및/또는 포르말린과 같은 유기 용매의 농도 증가에 의해) 점차적으로 증가시킨다. 예를 들어, 매칭되지 않은 표적에 대한 프로브의 혼성화에서 관찰되는 것의 5x 이상인 신호 대 노이즈 비로 프로브가 완전히 매칭되는 상보적 표적에 결합할 때까지, 혼성화 및 세정 조건을 점차적으로 증가시킨다.

[0205] "매우 엄격한" 조건은 특별한 프로브에 대한 열 융점 온도( $T_m$ )과 동등하도록 선택된다.  $T_m$ 은 시험 서열의 50%가 완전히 매칭되는 표적에 혼성화하는 (한정된 이온 강도 및 pH 하에서의) 온도이다. 본 발명의 목적 상, 일반적으로, "매우 엄격한" 혼성화 및 세정 조건은 한정된 이온 강도 및 pH에서 특정 서열에 대한  $T_m$ 보다 약 5°C 더 낮도록 선택된다.

[0206] "초고도 엄격성" 혼성화 및 세정 조건은, 혼성화 및 세정 조건의 엄격성이 완전히 매칭되는 상보적 표적 핵산에 대한 프로브의 결합을 위한 신호 대 노이즈 비가 매칭되지 않은 표적 핵산들 중 임의의 것에 대한 혼성화에서 관찰되는 비의 10x 이상이 될 때까지 증가되는 조건이다. 완전히 매칭되는 상보적 표적 핵산의 비의 1/2 이상인 신호 대 노이즈 비로 상기 조건 하에서 프로브에 혼성화하는 표적 핵산은 초고도 엄격성 조건 하에서 프로브에 결합한다고 일컬어진다.

[0207] 유사하게, 더 높은 엄격성 수준은 관련 혼성화 검정의 혼성화 및/또는 세정 조건을 점차적으로 증가시킴으로써 결정될 수 있다. 예를 들어, 완전히 매칭되는 상보적 표적 핵산에 대한 프로브의 결합을 위한 신호 대 노이즈 비가, 매칭되지 않은 표적 핵산들 중 임의의 것에 대한 혼성화에서 관찰되는 비의 적어도 10X, 20X, 50X, 100X, 또는 500X 또는 그 이상일 때까지, 혼성화 및 세정 조건의 엄격성을 증가시키는 조건이다. 완전히 매칭되는 상보적 표적 핵산의 비의 1/2 이상인 신호 대 노이즈 비로 상기 조건 하에서 프로브에 혼성화하는 표적 핵산은, 초초고도 엄격성 조건 하에서 프로브에 결합한다고 일컬어진다.

- [0208] 엄격한 조건 하에서 상호에 대해 혼성화하지 않는 핵산은, 그것이 코딩하는 폴리펩티드가 실질적으로 동일한 경우, 여전히 실질적으로 동일하다. 이는, 예컨대 유전 코드에 의해 허용되는 최대 코돈 퇴화를 이용하여 핵산의 카피가 생산될 때 일어난다.
- [0209] 독특한 하위서열
- [0210] 한 측면에서, 본 발명은 본원에 개시된 O-tRNA 및 O-RS의 서열로부터 선택되는 핵산 내 독특한 하위서열을 포함하는 핵산을 제공한다. 독특한 하위서열은 임의의 공지된 O-tRNA 또는 O-RS 핵산 서열에 상응하는 핵산에 비해 독특하다. 매개변수의 디폴트를 위해 예컨대 BLAST 세트를 이용하여 정렬을 수행할 수 있다. 임의의 독특한 하위서열은, 예컨대 본 발명의 핵산을 동정하기 위한 프로브로서 유용하다.
- [0211] 유사하게, 본 발명은 본원에 개시된 O-RS의 서열로부터 선택되는 폴리펩티드 내 독특한 하위서열을 포함하는 폴리펩티드를 포함한다. 여기에서, 독특한 하위서열은 공지된 폴리펩티드 서열들 중 임의의 것에 상응하는 폴리펩티드에 비해 독특하다.
- [0212] 본 발명은 또한 O-RS의 서열로부터 선택되는 폴리펩티드 내 독특한 하위서열을 코딩하는 독특한 코딩 올리고뉴클레오티드에 대해, 엄격한 조건 하에서 혼성화하는 표적 핵산을 제공하며, 여기에서 독특한 하위서열은 대조 폴리펩티드(예컨대, 본 발명의 합성효소가 예컨대 돌연변이에 의해 유도되도록 하는 모 서열) 중 임의의 것에 상응하는 폴리펩티드에 비해 독특하다. 독특한 서열은 상기 방식으로 결정된다.
- [0213] 서열 비교, 동일성 및 상동성
- [0214] 2개 이상의 핵산 또는 폴리펩티드 서열의 맥락에서 용어 "동일한" 또는 "동일성"은 하기 서열 비교 알고리즘(또는 당업자가 이용 가능한 기타 알고리즘) 중 하나를 사용하거나, 수동 정렬 및 시각적 조사에 의해 측정 시에, 비교 창 또는 지정된 영역에 대한 최대 상응에 대해 비교하고 정렬할 때, 동일한 아미노산 잔기 또는 뉴클레오티드의 특정 %를 가지거나, 동일한 2개 이상의 서열 또는 부분서열을 지칭한다.
- [0215] 2개의 핵산 또는 폴리펩티드(예컨대, O-tRNA 또는 O-RS를 코딩하는 DNA, 또는 O-RS의 아미노산 서열)의 맥락에서, 어구 "실질적으로 동일한"은 서열 비교 알고리즘(또는 당업자에 의해 이용가능한 기타 알고리즘)을 사용하거나, 수동 정렬 및 시각적 조사에 의해 측정 시에, 비교 창 또는 지정된 영역에 대한 최대 상응에 대해 비교하고 정렬할 때, 적어도 약 60%, 약 80%, 약 90 내지 95%, 약 98%, 약 99% 또는 그 이상의 뉴클레오티드 또는 아미노산 잔기 동일성을 가지는 2개 이상의 서열 또는 부분서열을 지칭한다. 그러한 "실질적으로 동일한" 서열은 전형적으로 실제 조상과 무관하게 "상동"인 것으로 간주된다. "실질적 동일성"은 약 50개 이상의 잔기 길이인 서열 영역, 약 100개 이상 잔기의 영역, 또는 약 150개 이상 잔기의 영역, 또는 비교하는 2개 서열의 전체 길이에서 존재할 수 있다.
- [0216] 서열 비교 및 상동성 결정을 위해, 전형적으로 하나의 서열은 시험 서열을 비교하기 위한 기준 서열로서 작용한다. 서열 비교 알고리즘을 사용하는 경우, 시험 서열 및 기준 서열을 컴퓨터에 입력하고, 필요에 따라 부분서열 좌표를 지정하며, 서열 알고리즘 프로그램 매개변수를 지정한다. 이어서, 서열 비교 알고리즘은 지정된 프로그램 매개변수를 기초로 하여 기준 서열에 대한 시험 서열(들)의 서열 동일성 %을 계산한다.
- [0217] 비교를 위한 서열의 정렬 방법은 당업자에게 공지되어 있다. 비교를 위한 서열의 최적의 정렬은 예컨대 [Smith and Waterman *Adv. Appl. Math.* 2: 482c (1981)]의 국소 상동성 알고리즘, [Needleman and Wunsch *J. Mol. Biol.* 48: 443 (1970)]의 상동성 정렬 알고리즘, [Pearson and Lipman *Proc. Nat'l. Acad. Sci. USA* 85: 2444 (1988)]의 유사성 방법 탐색, 이들 알고리즘의 컴퓨터화된 실행([Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI]에서의 GAP, BESTFIT, FASTA 및 TFASTA), 또는 수동 정렬 및 시각적 조사(예컨대, [Ausubel et al., *Current Protocols in Molecular Biology* (1995 부록)] 참고)에 의해 수행될 수 있다.
- [0218] 서열 동일성 및 서열 유사성을 결정하기에 적당한 알고리즘의 한 예로는 각각 [Altschul et al. (1977) *Nuc. Acids Res.* 25: 3389-3402, 및 Altschul et al. *J. Mol. Biol.* 215: 403-410 (1990)]에 기재되어 있는 BLAST 알고리즘 및 BLAST 2.0 알고리즘이 있다. BLAST 분석 수행용 소프트웨어는 월드 와이드 웹 [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov)에서 이용가능한 국립 생명공학 정보 센터(National Center for Biotechnology Information)를 통해 공개적으로 입수가능하다. 이 알고리즘은 먼저, 데이터베이스 서열 내 동일한 길이의 단어로 정렬될 때, 일부 양의 값 역치 스코어 T를 매칭시키거나 만족시키는, 질의 서열 내 길이 W의 짧은 단어를 확인함으로써 하이 스코어링 서열 쌍(HSP)을 확인하는 것을 포함한다. T는 이웃 단어 스코어 역치(neighborhood word score threshold)로 칭해진다



(Altschul et al, 이하 동문). 이 초기 이웃 단어 일치(word hit)는 그것을 함유하는 보다 긴 HSP를 찾기 위한 검색을 개시하는 중(seed)으로서 작용한다. 이어서, 단어 일치를 누적 정렬 스코어가 증가될 수 있는 한 각 서열을 따라 양 방향으로 연장시킨다. 뉴클레오타이드 서열에 대해 매개변수 M(매칭하는 잔기의 쌍에 대한 보상 스코어; 항상 > 0) 및 N(매칭하지 않는 잔기에 대한 페널티 스코어; 항상 < 0)을 이용하여 누적 스코어를 계산한다. 아미노산 서열에 대해, 스코어링 행렬을 이용하여, 누적 스코어를 계산한다. 누적 정렬 스코어가 최대 달성된 값에서 양 X만큼 떨어질 때; 1 이상의 음의 스코어링 잔기 정렬의 누적으로 인해 누적 스코어가 0 이하로 떨어질 때; 또는 어느 한 서열의 말단에 도달할 때, 각 방향으로의 단어 일치 연장을 중단한다. BLAST 알고리즘 매개변수 W, T 및 X는 정렬의 감도 및 속도를 결정한다. (뉴클레오타이드 서열의 대한) BLASTN 프로그램은 디폴트로서 단어 길이(wordlength; W) = 11, 기대값(E) = 10, 컷오프(cutoff) = 100, M = 5, N = -4 및 양쪽 가닥의 비교를 사용한다. 아미노산 서열에 대해서는, BLASTP 프로그램이 디폴트로서 단어 길이(W) = 3 및 기대값(E) = 10, 및 BLOSUM62 스코어링 매트릭스([Henikoff and Henikoff (1992) *Proc. Natl. Acad. Sci. USA* 89: 10915] 참고) 정렬(B) 50, 기대값(E) = 10, M = 5, N = -4 및 양쪽 가닥의 비교를 사용한다. 전형적으로, BLAST 알고리즘은 "저 복잡성 (complexity)" 필터를 오프하여(turned off) 수행된다.

[0219] 서열 동일성 %의 계산에 부가하여, BLAST 알고리즘은 2개의 서열 사이의 유사성을 통계적으로 분석한다(예컨대, [Karlin 및 Altschul *Proc. Natl. Acad. Sci. USA* 90:5773-5787 (1993)] 참고). BLAST 알고리즘에 의해 제공되는 유사성의 한 척도는 2개의 뉴클레오타이드 또는 아미노산 서열 사이의 매칭이 우연히 발생하는 확률의 지표를 제공하는 최소 합계 확률(P(N))이다. 예를 들어, 기준 핵산에 대해 시험 핵산을 비교할 때 최소 합계 확률이 약 0.2 미만, 약 0.01 미만, 약 0.001 미만인 경우, 핵산은 기준 서열과 유사한 것으로 간주될 수 있다.

[0220] 돌연변이유발 및 기타 분자 생물학적 기법

[0221] 본 발명에 사용되는, 본 발명의 폴리뉴클레오타이드 및 폴리펩티드는 분자 생물학적 기법을 이용하여 조작될 수 있다. 뉴클레오타이드 서열은 통상적인 방법에 따라 부위-지정 돌연변이유발에 의해 편리하게 변형될 수 있다. 대안적으로, 뉴클레오타이드 서열은 올리고뉴클레오타이드 합성기를 사용하는 것을 포함하나 이에 한정되지 않은 화학적 합성에 의해 제조될 수 있으며, 여기에서 올리고뉴클레오타이드는 원하는 폴리펩티드의 아미노산 서열에 기초하여 설계되며, 바람직하게는, 재조합 폴리펩티드가 생산되는 숙주 세포에서 유리한 코돈을 선택한다. 예를 들어, 원하는 폴리펩티드의 부분에 대한 수가지 소형 올리고뉴클레오타이드 코딩은 PCR, 절찰(ligation) 또는 절찰 연쇄 반응에 의해 합성되고, 조립될 수 있다. 예를 들어, 본원에 참고로 인용되는 [Barany, et al., *Proc. Natl. Acad. Sci.* 88: 189-193(1991)]; 및 미국 특허 제6,521,427호를 참고한다.

[0222] 본 발명은 재조합 유전학 분야의 통상의 기법을 사용한다. 본 발명에 사용되는 일반적 방법을 개시하는 기본서는 [Sambrook et al., *Molecular Cloning, A Laboratory Manual* (제3판 2001); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); 및 *Current Protocols in Molecular Biology* (Ausubel et al., eds., 1994)]을 포함한다.

[0223] 분자 생물학적 기술을 설명하는 일반 연구서는 [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 (2nd Ed.), Vol. 1-3*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook") 및 *Current Protocols in Molecular Biology*, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1999년 보충) ("Ausubel")]을 포함한다. 이러한 연구서들은, 예컨대 선택된 아미노산 (예컨대, 비천연 아미노산), 오르소고달 tRNA, 오르소고달 합성효소 및 그들의 쌍을 포함하는 단백질을 생산하기 위한 선택적 코돈을 포함하는 유전자 또는 폴리뉴클레오타이드의 발생과 관련된 돌연변이유발, 벡터의 사용, 프로모터 및 많은 다른 관련 주제를 기재하고 있다.

[0224] 다양한 유형의 돌연변이유발은, 신규 합성효소 또는 tRNA의 생성, tRNA의 라이브러리 생산, RS 분자의 돌연변이화, 합성효소의 라이브러리 생산, 선택적 코돈 생산, 관심 단백질 또는 폴리펩티드에서 선택된 아미노산을 코딩하는 선택적 코돈 삽입을 포함하나 이에 한정되지 않은 다양한 목적을 위해 본 발명에 사용된다. 이들은 부위-지정, 무작위 점 돌연변이유발, 동족 재조합, DNA 서플링 또는 다른 재귀적 돌연변이유발 방법, 키메라 구성, 주형을 함유하는 우라실을 사용한 돌연변이유발, 올리고뉴클레오타이드-지정 돌연변이유발, 포스포로티오에이트-변형 DNA 돌연변이유발, 갭이 있는 이중체 DNA를 사용한 돌연변이유발 등 또는 그의 임의의 조합을 포함하나 이에 한정되지 않는다. 추가적인 적당한 방법은 점 미스매치 수복, 수복-결핍 숙주 균주를 사용한 돌연변이유발, 제한-선택 및 제한-정제, 결실 돌연변이유발, 총 유전자 합성에 의한 돌연변이유발, 이중 가닥 파괴 수복 등을

포함한다. 키메라 구축물의 이용을 포함하나 이에 한정되지 않은 돌연변이유발도 본 발명에 포함된다. 한 실시양태에서, 서열, 서열 비교, 물리적 성질, 2차, 3차 또는 4차 구조, 결정 구조 등을 포함하나 이에 한정되지 않은 천연 발생 분자, 또는 변형 또는 돌연변이화된 천연 발생 분자의 공지된 정보에 의해 돌연변이유발을 주도할 수 있다.

[0225]

본원에서 발견되는 본문 및 예들은 이러한 절차를 기재한다. 하기 발행물 및 그 안에 언급된 참고문헌에 추가적 정보가 나와 있다: [Ling et al., *Approaches to DNA mutagenesis: an overview*, Anal Biochem. 254(2): 157-178 (1997); Dale et al., *Oligonucleotide-directed random mutagenesis using the phosphorothioate method*, Methods Mol. Biol. 57:369-374 (1996); Smith, *In vitro mutagenesis*, Ann. Rev. Genet. 19:423-462 (1985); Botstein & Shortle, *Strategies and applications of in vitro mutagenesis*, Science 229:1193-1201 (1985); Carter, *Site-directed mutagenesis*, Biochem. J. 237:1-7 (1986); Kunkel, *The efficiency of oligonucleotide directed mutagenesis*, in Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D.M.J. eds., Springer Verlag, Berlin) (1987); Kunkel, *Rapid and efficient site-specific mutagenesis without phenotypic selection*, Proc. Natl. Acad. Sci. USA 82:488-492 (1985); Kunkel et al., *Rapid and efficient site-specific mutagenesis without phenotypic selection*, Methods in Enzymol. 154, 367-382 (1987); Bass et al., *Mutant Trp repressors with new DNA-binding specificities*, Science 242:240-245 (1988); Zoller & Smith, *Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment*, Nucleic Acids Res. 10:6487-6500 (1982); Zoller & Smith, *Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors*, Methods in Enzymol. 100:468-500 (1983); Zoller & Smith, *Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template*, Methods in Enzymol. 154:329-350 (1987); Taylor et al., *The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA*, Nucl. Acids Res. 13: 8749-8764 (1985); Taylor et al., *The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA*, Nucl. Acids Res. 13: 8765-8785 (1985); Nakamaye & Eckstein, *Inhibition of restriction endonuclease Nd I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis*, Nucl. Acids Res. 14: 9679-9698 (1986); Sayers et al., *5'-3' Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis*, Nucl. Acids Res. 16:791-802 (1988); Sayers et al., *Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide* (1988) Nucl. Acids Res. 16: 803-814; Kramer et al., *The gapped duplex DNA approach to oligonucleotide-directed mutation construction*, Nucl. Acids Res. 12: 9441-9456 (1984); Kramer & Fritz *Oligonucleotide-directed construction of mutations via gapped duplex DNA*, Methods in Enzymol. 154:350-367 (1987); Kramer et al., *Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations*, Nucl. Acids Res. 16: 7207 (1988); Fritz et al., *Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro*, Nucl. Acids Res. 16: 6987-6999 (1988); Kramer et al., *Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of E. coli*, Cell 38:879-887 (1984); Carter et al., *Improved oligonucleotide site-directed mutagenesis using M13 vectors*, Nucl. Acids Res. 13: 4431-4443 (1985); Carter, *Improved oligonucleotide-directed mutagenesis using M13 vectors*, Methods in Enzymol. 154: 382-403 (1987); Eghtedarzadeh & Henikoff, *Use of oligonucleotides to generate large deletions*, Nucl. Acids Res. 14: 5115 (1986); Wells et al., *Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin*, Phil. Trans. R. Soc. Lond. A 317: 415-423 (1986); Nambiar et al., *Total synthesis and cloning of a gene coding for the ribonuclease S protein*, Science 223: 1299-1301 (1984); Sakmar and Khorana, *Total synthesis and expression of a gene for the alpha-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)*, Nucl. Acids Res. 14: 6361-6372 (1988); Wells et al., *Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites*, Gene 34:315-323 (1985); Grundstrom et al., *Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis*, Nucl. Acids Res. 13: 3305-3316 (1985); Mandecki, *Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis*, Proc. Natl. Acad. Sci. USA, 83:7177-7181 (1986); Arnold, *Protein*

engineering for unusual environments, *Current Opinion in Biotechnology* 4:450-455 (1993); Sieber, et al., *Nature Biotechnology*, 19:456-460 (2001); W. P. C. [0278] 68 Stemmer, *Nature* 370, 389-91 (1994); 및 I. A. Lorimer, I. Pastan, *Nucleic Acids Res.* 23, 3067-8 (1995)]. 상기 방법들 중 많은 방법에 대한 추가적인 상세 내용을, 다양한 돌연변이유발 방법을 이용한 타겟 문제점에 대한 유용한 조절을 또한 기재하고 있는 [Methods in Enzymology Volume 154]에서 찾아볼 수 있다.

[0226] 예컨대 tRNA 또는 합성효소를 돌연변이화하거나, tRNA 또는 RS를 변경시키는, 예컨대 본 발명의 돌연변이유발에 사용하기 위한 올리고뉴클레오타이드는 전형적으로 예컨대, [Needham-VanDevanter et al., *Nucleic Acids Res.* 12:6159-6168 (1984)]에 기재된 바와 같은 자동화 합성장치를 이용하여, [Beaucage and Caruthers, *Tetrahedron Letts.* 22(20):1859-1862 (1981)]에 기재된 고체상 포스포르아미다이트 트리에스테르 방법에 따라 화학적으로 합성된다.

[0227] 또한, 본질적으로 임의의 핵산을 다양한 상업적 공급원, 예를 들어 더 미들랜드 써티파이드 리에이전트 컴퍼니(The Midland Certified Reagent Company)(mcrcloligos.com), 더 그레이트 어메리칸 진 컴퍼니(The Great American Gene Company)(www.genco.com), 익스프레스젠 인코포레이티드(ExpressGen Inc.)(www.expressgen.com), 오페론 테크놀로지스 인코포레이티드(Operon Technologies Inc.)(미국 캘리포니아주 알라메다 소재) 및 기타 많은 회사들 중 임의의 회사로부터 고객 주문하거나 표준 주문할 수 있다.

[0228] 본 발명은 또한 오르소고날 tRNA/RS 쌍을 통한 비천연 아미노산의 생체 내 혼입을 위한 진핵생물 숙주 세포, 비진핵생물 숙주 세포 및 유기체에 관한 것이다. 숙주 세포는, 예를 들어, 클로닝 벡터 또는 발현 벡터일 수 있는 본 발명의 벡터를 포함하나 이에 한정되지 않는 본 발명의 폴리뉴클레오타이드를 포함하는 본 발명의 폴리뉴클레오타이드 또는 구축물을 사용하여 유전공학처리(형질전환, 형질도입 또는 트랜스펙션을 포함하나 이에 한정되지 않음)된다. 예를 들어, 오르소고날 tRNA, 오르소고날 tRNA 합성효소, 및 유도체화되는 단백질에 대한 코딩 영역은 원하는 숙주 세포 내에서 작용성인 유전자 발현 조절 요소에 작용적으로 연결된다. 벡터는 예를 들어, 플라스미드, 코스미드, 파지, 세균, 바이러스, 나형 폴리뉴클레오타이드 또는 접합된 폴리뉴클레오타이드의 형태일 수 있다. 전기천공법[From et al., *Proc. Natl. Acad. Sci. USA* 82, 5824 (1985)], 바이러스 벡터에 의한 감염, 작은 비드 또는 입자의 매트릭스 내 또는 표면 상에서 핵산을 사용하여 작은 입자에 의한 고속 탄동 침투[Klein et al., *Nature* 327, 70-73 (1987)] 등을 비롯한 표준 방법에 의해 벡터를 세포 및/또는 미생물에 도입시킨다.

[0229] 표적 핵산을 세포에 도입시키는 수가지 공지된 방법이 이용가능하며, 이들 중 임의의 방법이 본 발명에 사용될 수 있다. 이들은 수용 세포와 DNA를 함유하는 세균 원형질체의 융합, 전기천공법, 추진체 포격(projectile bombardment), 및 (이하에서 추가 논의되는) 바이러스 벡터를 사용한 감염 등을 포함한다. 세균 세포를 본 발명의 DNA 구축물을 함유하는 다수의 플라스미드를 증폭시키는데 사용할 수 있다. 세균은 로그 상태(log phase)로 성장하고, 세균 내 플라스미드는 당업계에 공지된 다양한 방법 (예컨대, [Sambrook] 참고)에 의해 분리될 수 있다. 또한, 세균으로부터 플라스미드를 정제하기 위한 수많은 키트가 상업적으로 구입가능하다(예컨대, 양자 모두 파마시아 바이오테크(Pharmacia Biotech)로부터 각기 입수되는 이지프렙(EasyPrep)<sup>TM</sup>, 플렉시프렙(FlexiPrep)<sup>TM</sup>; 스트라타진(Stratagene)으로부터 입수되는 스트라타클린(StrataClean)<sup>TM</sup>; 및 키아젠(Qiagen)으로부터 입수되는 쿼이프렙(QIAprep)<sup>TM</sup>). 이어서, 분리되고 정제된 플라스미드를 추가 조작하여, 세포 트랜스펙션에 사용되거나 관련된 벡터에 혼입시켜 유기체를 감염시키는 다른 플라스미드를 생산한다. 전형적인 벡터는 전사 및 번역 종결자, 전사 및 번역 개시 서열, 및 특정 표적 핵산의 발현의 조절에 유용한 프로모터를 함유한다. 벡터는 임의적으로 하나 이상의 독립적인 종결자 서열, 진핵생물 또는 원핵생물 또는 양자 모두 중에서 카세트의 복제를 허용하는 서열(서열 벡터를 포함하나 이에 한정되지 않음), 및 원핵생물 및 진핵생물 시스템 양자 모두에 대한 선별 마커를 함유하는 일반 발현 카세트를 포함한다. 벡터는 원핵생물, 진핵생물 또는 양자 모두에서 복제 및/또는 일체화에 적합하다. [Gilman & Smith, *Gene* 8: 81 (1979); Roberts, et al., *Nature*, 328: 731 (1987); Schneider, B., et al., *Protein Expr. Purif.* 6(1):10: 14 (1995); Ausubel, Sambrook, Berger(모두 이하 동문)]을 참고한다. 클로닝에 유용한 세균 및 박테리오파지(bacteriophage)의 카탈로그는 예를 들어, ATCC에 의해 예를 들어, ATCC에 의해 공개된 문헌 [*The ATCC Catalogue of Bacteria and Bacteriophage* (1992) Gherna et al. (eds)]에 제공되어 있다. 또한, 분자 생물학 및 기초적인 이론적 고려사항의 서열 결정, 클로닝 및 다른 측면에 대한 추가적인 기본 절차는 [Watson et al. (1992) *Recombinant DNA Second Edition Scientific American Books*, NY]에 나와있다. 또한, 필수적으로 임의의 핵산 (및 표준 또는 비표준이든 실질적으로 임의의 표지화된 핵산)을 다양한 상업적 공급원, 예를 들어 미들랜드 써티파이드 리에이전트 컴퍼니(미국 텍사스 미들랜드 소재, 월드 와이드 웹 상의 mcrcloligos.com에서 이용가능함), 더 그레이트 어메리칸 진 컴퍼니(미국

캘리포니아주 라모나 소재, 월드 와이드 웹 상의 genco.com에서 이용가능함), 익스프레스젠 인코포레이티드(미국 일리노이주 시카고 소재, 월드 와이드 웹 상의 expressgen.com에서 이용가능함), 오펜론 테크놀로지스 인코포레이티드(미국 캘리포니아주 알라메다 소재) 및 기타 많은 회사들 중 임의의 회사로부터 고객 주문하거나 표준 주문할 수 있다.

[0230] 유전공학처리된 숙주 세포는 예를 들어, 선별 단계, 프로모터 활성화 단계 또는 형질 전환체 선택 단계와 같은 활동들에 적당하도록 개질된 통상적 영양 배지 중에서 배양될 수 있다. 이러한 세포는 유전자도입(transgenic) 유기체 내에서 배양될 수도 있다. 예를 들어, 세포 분리 및 배양(예컨대, 후속 핵산 분리)에 관한 기타 유용한 참고 문헌으로서는 다음의 문헌들을 포함한다[Freshney (1994) Culture of Animal Cells, a Manual of Basic Technique, 제3판, Wiley-Liss, New York 및 여기에 인용된 참고 문헌들; Payne와 다수 (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].

[0231] 생체 내에서 비천연 아미노산을 단백질에 직접 혼입시키는 능력은 돌연변이 단백질의 고수율 수득, 기술적 용이성, 돌연변이 단백질의 세포 내 또는 생 유기체 내 연구 가능성 및 치료적 처치법에서의 돌연변이 단백질의 사용과 같은 매우 다양한 이점들을 제공한다. 다양한 크기, 산성도, 친핵성, 소수성 및 기타 특성을 갖는 비천연 아미노산을 단백질에 혼입하는 능력은, 단백질 기능을 탐침하고 새로운 성질을 갖는 신규의 단백질 또는 유기체를 생산하기 위하여, 단백질 구조를 합리적으로, 또한 체계적으로 조작하는 능력을 크게 확장할 수 있다.

[0232] 관심 단백질 및 폴리펩티드

[0233] 비천연 아미노산의 혼입은 단백질 구조 및/또는 기능의 변화를 조정하거나, 크기, 산도, 친핵성, 수소 결합, 소수성, 프로테아제 표적 부위의 접근성, 부분(단백질 어레이에 대한 경우를 포함하나 이에 한정되지 않음)에 대한 표적화, 생물학적 활성 분자의 첨가, 중합체의 부착, 방사성 핵종의 부착, 혈청 반감기의 조절, 조직 투과(예컨대, 종양)의 조절, 활성 수송의 조절, 조직, 세포 또는 기관 특이성 또는 분포의 조절, 면역원성의 조절, 프로테아제 내성의 조절 등을 포함하나 이에 한정되지 않는 다양한 목적을 위해 수행될 수 있다. 비천연 아미노산을 포함하는 단백질은 증진되거나 심지어 완전히 신규한 촉매적 성질 또는 생체물리적 성질을 가질 수 있다. 예를 들어, 비천연 아미노산을 단백질에 포함시킴으로써 하기 성질들이 임의적으로 변형된다: 독성, 생체 분포, 구조적 성질, 분광학적 성질, 화학적 및/또는 광화학적 성질, 촉매적 능력, 반감기(혈청 반감기를 포함하나 이에 한정되지 않음), 공유적으로 또는 비공유적으로(단, 이에 한정되지 않음) 반응할 수 있는 다른 분자와 반응할 수 있는 능력 등. 하나 이상의 비천연 아미노산을 포함하는 단백질을 포함하는 조성물은 신규한 치료제, 진단제, 촉매 작용 효소, 산업용 효소, 항체를 포함하나 이에 한정되지 않는 결합 단백질(단, 이에 한정되지 않음), 및 단백질 구조 및 기능의 연구(단, 이에 한정되지 않음)에 유용하다. 예를 들어, [Dougherty, (2000) Unnatural Amino Acids as Probes of Protein Structure and Function, Current Opinion in Chemical Biology, 4: 645-652]을 참고한다.

[0234] 단백질은 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6개 이상, 7개 이상, 8개 이상, 9개 이상 또는 10개 이상을 포함하나 이에 한정되지 않는, 1개 이상의 개수의 비천연 아미노산을 가질 수 있다. 비천연 아미노산은 동일하거나 상이할 수 있고, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개, 또는 그 이상 개수의 상이한 비천연 아미노산을 포함하는 단백질 중의 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개, 또는 그 이상 개수의 상이한 부위가 존재할 수 있는 경우를 포함하나 이에 한정되지 않는다. 단백질은, 단백질 중에 존재하는 특정 아미노산 중 하나 이상, 단 전부가 아닌 그보다 적은 수가 비천연 아미노산으로 치환될 수 있다. 1개 초과와 비천연 아미노산을 갖는 소정의 단백질의 경우, 비천연 아미노산은 동일하거나 상이할 수 있다(단백질은 2개 이상의 상이한 유형의 비천연 아미노산을 포함하거나, 2개의 동일한 비천연 아미노산을 포함할 수 있으나, 이에 한정되지 않음). 2개 초과와 비천연 아미노산을 갖는 소정의 단백질의 경우, 비천연 아미노산은 동일하거나 상이하거나, 또는 하나 이상의 상이한 비천연 아미노산과 동일한 종류로 된 다수의 비천연 아미노산의 조합물일 수 있다.

[0235] 하나 이상의 비천연 아미노산을 갖는 관심 단백질 또는 폴리펩티드를 진핵세포에서 생산함으로써, 단백질 또는 폴리펩티드는 전형적으로 번역후 진핵세포 변형을 포함하게 된다. 특정 실시양태에서, 단백질은 하나 이상의 비천연 아미노산, 및 진핵세포에 의해 생체 내 생성된 하나 이상의 번역후 변형을 포함하며, 여기에서 번역후 변형은 원핵세포에 의해서는 이루어지지 않는다. 예를 들어, 번역 후 변형은 아세틸화, 아실화, 지질-변형, 팔미토일화, 팔미테이트 첨가, 인산화, 당지질-연결 변형, 글리코실화 등을 포함하나 이에 한정되지 않는 것들을 포



함한다. 또 다른 측면에서, 번역 후 변형은 전구체(칼시토닌 전구체, 칼시토닌 유전자 관련 펩티드 전구체, 전 전구부갑상선 호르몬, 전전구인슐린(preproinsulin), 전구인슐린(proinsulin), 전전구-오피오멜라코르틴(preproopiomelanocortin), 전구-오피오멜라코르틴(pro-opiomelanocortin) 등을 포함하나 이에 한정되지 않음)의 단백질분해 처리, 다중 서브유닛 단백질 또는 거대분자 어셈블리의 조립, 세포 중의 또 다른 부위로(소기관, 예를 들어 소포체, 골지체, 핵, 리소좀, 퍼옥시좀, 미토콘드리아, 엽록체, 액포 등으로, 또는 분비성 경로를 통하는 것을 포함하나 이에 한정되지 않음) 번역시키는 것을 포함한다. 특정 실시양태에서, 단백질은 분비 또는 국소화 서열, 에피토프 태그, FLAG 태그, 폴리히스티딘 태그, GST 융합 등을 포함한다.

[0236] 특정 위치에서의 선택된 아미노산으로 세포 내 단백질을 생산하는 방법도 또한 본 발명의 한 특징이다. 예를 들어, 방법은 한 적절한 매질 내에서, 하나 이상의 셀렉터 코돈을 포함하고 단백질을 코딩하는 핵산을 포함하는 세포를 성장시키고; 선택된 아미노산을 제공하는 것(여기에서, 세포는 세포 내에서 기능하고 셀렉터 코돈을 인식하는 오르소고날 tRNA(O-tRNA); 및 선택된 아미노산으로 O-tRNA를 우선적으로 아미노아실화하는 오르소고날 아미노아실-tRNA 합성효소(O-RS)를 추가로 포함함)을 포함한다. 전형적으로, O-tRNA는 셀렉터 코돈에 대한 반응으로 동족 합성효소의 존재 하에 억제 활성을 포함한다. 이 방법에 의해 생산된 단백질도 또한 본 발명의 한 특징이다.

[0237] 본 발명의 조성물, 및 본 발명의 방법에 의해 제조된 조성물은 임의적으로 세포 내에 있다. 본 발명의 O-tRNA/O-RS 쌍 또는 개별 성분을 숙주 시스템의 번역 기구에 사용될 수 있고, 이에 따라 선택된 아미노산, 예컨대 비천연 아미노산이 단백질 내에 혼입되게 된다. 특허 출원 USSN 제10/825,867호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE) 및 제10/126,927호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")는 상기 공정을 기재하고 있고, 본원에 참고로 인용된다. 예를 들어, O-tRNA/O-RS 쌍이 숙주, 예컨대 *에셰리키아 콜라이*에 도입될 때, 쌍은 셀렉터 코돈에 대한 반응으로, 성장 매질에 외인적으로 첨가될 수 있는, 선택된 아미노산, 예를 들어 비천연 아미노산, 예컨대 합성 아미노산, 예컨대 류신 아미노산의 유도체가 단백질로 생체 내 혼입되도록 한다. 임의적으로, 본 발명의 조성물은 시험관 내 번역 시스템, 또는 생체 내 시스템(들) 내에 있을 수 있다.

[0238] 선택된 아미노산, 예컨대 비천연 아미노산 (및 예컨대 하나 이상의 셀렉터 코돈을 포함하는, 임의의 상응하는 코딩 핵산)을 포함하는 임의의 단백질 (또는 이의 부분)을 본원의 조성물 및 방법을 이용하여 생산할 수 있다. 임의의 폴리펩티드가 하나 이상의 선택된 아미노산을 혼입하는데 적당하다. 예컨대 임의의 이용가능한 돌연변이 방법을 조정하여 관련 번역 시스템에 하나 이상의 적절한 셀렉터 코돈을 포함하도록 함으로써, 하나 이상의 비천연 아미노산을 포함하도록 변형될 수 있는 수천만가지의 공지 단백질을 동정하기 위한 시도는 없다. 공지된 단백질을 위한 통상의 서열 정보저장소(repositories)는 유전자은행(GenBank) EMBL, DDBJ 및 NCBI를 포함한다. 인터넷 검색에 의해 기타 정보저장소도 용이하게 확인할 수 있다.

[0239] 전형적으로, 단백질은 임의의 이용가능한 단백질(예컨대, 치료 단백질, 진단 단백질, 산업적 효소 또는 이의 일부분 등)에 대해, 예컨대 60% 이상, 70% 이상, 75% 이상, 80% 이상, 90% 이상, 95% 이상, 또는 99% 이상, 또는 그 이상의 동일성을 가지고, 하나 이상의 선택된 아미노산을 포함한다. 하나 이상의 선택된 아미노산, 예컨대 비천연 아미노산을 포함하도록 변형될 수 있는 치료, 진단 및 기타 단백질의 예는 USSN 제10/825,867호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE") 및 미국 특허 출원 일련 번호 제10/126,927호(발명의 명칭: "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS")에서 찾아볼 수 있으나, 이들 문헌에 한정되지 않는다.

[0240] 특정 실시양태들에서, 본 발명의 방법 및/또는 조성물에서의 관심 단백질 또는 폴리펩티드 (또는 이의 일부분)는 핵산에 의해 코딩된다. 전형적으로, 핵산은 하나 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, 또는 10개 이상의 셀렉터 코돈을 포함한다.

[0241] 관심 단백질 또는 폴리펩티드를 코딩하는 유전자는 당업자에게 공지되고, "Mutagenesis and Other Molecular Biology Techniques" 하에 기재된 방법을 이용하여, 예컨대 선택된 아미노산, 예컨대 비천연 아미노산을 혼입하기 위한 하나 이상의 셀렉터 코돈을 포함하도록 돌연변이유발될 수 있다. 예를 들어, 관심 단백질을 위한 핵산은 하나 이상의 셀렉터 코돈을 포함하도록 돌연변이유발되어, 하나 이상의 선택된 아미노산, 예컨대 비천연 아미노산이 삽입되도록 한다. 본 발명은 예컨대 하나 이상의 선택된 아미노산을 포함하는, 임의의 그러한 변이체, 예컨대 돌연변이, 임의의 단백질의 양태를 포함한다. 유사하게, 본 발명은 상응하는 핵산, 즉 하나 이상의 선택된 아미노산을 코딩하는 하나 이상의 셀렉터 코돈을 갖는 임의의 핵산을 포함한다.

[0242] 선택된 아미노산을 포함하는 단백질을 제조하기 위해, 오르소고날 tRNA/RS 쌍을 통해 선택된 아미노산을 생체

내 혼입하기 위해 적합화된 숙주 세포 및 유기체를 사용할 수 있다. 숙주 세포는 오르소고날 tRNA, 오르소고날 tRNA 합성효소, 및 유도체화되는 단백질을 코딩하는 벡터를 발현하는 하나 이상의 벡터로 유전적으로 공학처리 (예컨대, 형질전환, 트랜스덕션 또는 트랜스펙션)된다. 각각의 이 성분은 동일한 벡터 상에 있을 수 있거나, 각각은 분리된 벡터 상에 있을 수 있으며, 양 성분은 한 벡터 상에 있고, 제3 성분이 제2 성분 상에 있을 수 있다. 벡터는 예를 들어, 플라스미드, 코스미드, 파지, 세균, 바이러스, 나형 폴리뉴클레오티드, 또는 접합된 폴리뉴클레오티드의 형태일 수 있다.

[0243] 교대 시스템

[0244] 비제조합 숙주 세포, 돌연변이유발된 숙주 세포 또는 무세포 시스템에서 비천연 아미노산을 단백질에 도입시키는데 수가지 기법들이 사용되었다. 반응성 측쇄, 예를 들어 Lys, Cys 및 Tyr을 사용한 아미노산의 유도체화하는 리신을 N<sup>2</sup>-아세틸-리신으로 전환시켰다. 화학적 합성은 또한 비천연 아미노산을 혼입하는 직접적 방법을 제공한다. 펩티드 단편의 효소적 결합, 및 본연의 화학적 결합의 최근 개발로, 더 큰 단백질을 제조하는 것이 가능하다. 예를 들어, [P. E. Dawson 및 S. B. H. Kent, *Annu. Rev. Biochem.*, 69: 923(2000)]을 참고한다. 화학적 펩티드 결합 및 본연의 화학적 결합이 본원에 참고로 인용되는 미국 특허 제6,184,344호, 미국 특허 공보 제2004/0138412호, 미국 특허 공보 제2003/0208046호, WO 02/098902, 및 WO 03/042235에 기재되어 있다. 원하는 비천연 아미노산으로 화학적으로 아실화된 억제제 tRNA를 단백질 생합성을 지지할 수 있는 시험관 내 추출물에 첨가하는 일반적 시험관 내 생합성 방법을 사용하여, 100개 초과인 비천연 아미노산을 실질적으로 임의의 크기의 다양한 단백질에 부위 특이적으로 혼입시켰다. 예를 들어, [V. W. Cornish, D. Mendel 및 P. G. Schultz, *Angew. Chem. Int. Ed. Engl.*, 1995, 34: 621(1995); C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, P. G. Schultz, *A general method for site-specific incorporation of unnatural amino acids into Proteins*, *Science* 244:182-188(1989); 및 J. D. Bain, C. G. Glabe, T. A. Dix, A. R. Chamberlin, E. S. Diala, *Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide*, *J. Am. Chem. Soc.* 111: 8013-8014(1989)]을 참조할 수 있다. 단백질 안정성, 단백질 접힘, 효소 기전 및 신호 전달의 연구를 위해 광범위한 범위의 작용기들을 단백질에 도입해왔다.

[0245] 선택적 압력 혼입으로 불리는 생체 내 방법을 개발하여 야생형 합성효소의 무차별 혼합(promiscuity)을 이용하였다. 예를 들어, [N. Budisa, C. Minks, S. Alefelder, W. Wenger, F. M. Dong, L. Moroder 및 R. Huber, *FASEB J.*, 13: 41(1999)]을 참고한다. 세포에 특정 천연 아미노산을 공급하는 관련 대사 경로가 비작동 (switched off) 영양요구성 균주는 제한된 농도의 천연 아미노산을 함유하는 최소 배지에서 성장하고, 한편 표적 유전자의 전사는 억제된다. 정제 성장기의 개시 시에, 천연 아미노산은 고갈되고, 비천연 아미노산 유사체로 대체된다. 제조합 단백질의 발현 유도는 비천연 유사체를 함유하는 단백질을 축적시키게 된다. 예를 들어, 이 기법을 사용하여, o, m 및 p-플루오로페닐알라닌은 단백질로 혼입되고, 용이하게 확인될 수 있는 UV 스펙트럼 중의 2개의 특징적인 숄더(shoulder)를 나타내며(예컨대, [C. Minks, R. Huber, L. Moroder 및 N. Budisa, *Anal. Biochem.*, 284: 29(2000)] 참고), 트리플루오로메티오닌을 박테리오파지 T4 리소자임(lysozyme)에서 메티오닌을 대체하는데 사용하여 <sup>19</sup>F NMR에 의해 그의 키토-이당류 리간드와의 상호작용을 연구하였고(예컨대, [H. Duewel, E. Daub, V. Robinson 및 J. F. Honek, *Biochemistry*, 36:3404(1997)] 참고), 트리플루오로류신을 류신 대신 혼입한 결과, 류신-지퍼(zipper) 단백질의 열 안정성 및 화학적 안정성이 증가되었다. 예컨대, [Y. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado 및 D. A. Tirrell, *Angew. Chem. Int. Ed. Engl.*, 40: 1494(2001)]을 참고한다. 더욱이, 셀레노메티오닌 및 텔루로메티오닌을 다양한 제조합 단백질에 혼입하여 X선 결정학에서의 위상의 해석을 용이하게 한다. 예를 들어, [W. A. Hendrickson, J.R. Horton 및 D. M. Lemaster, *EMBO J.*, 9: 1665(1990); J. O. Boles, K. Lewinski, M. Kunkle, J. D. Odom, B. Dunlap, L. Lebioda 및 M. Hatada, *Nat. Struct. Biol.*, 1: 283(1994); N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann 및 R. Huber, *Eur. J. Biochem.*, 230: 788(1995); 및 N. Budisa, W. Karnbrock, S. Steinbacher, A. Humm, L. Prade, T. Neufelnd, L. Moroder 및 R. Huber, *J. Mol. Biol.* 270: 616(1997)]을 참고한다. 또한, 알켄 또는 알킨 작용기를 갖는 메티오닌 유사체를 유효하게 혼입함으로써, 화학적 수단에 의해 단백질을 추가적으로 변형시켰다. 예를 들어, 본원에 참고로 인용되는 [J. C. M. van Hest 및 D. A. Tirrell, *FEBS Lett.*, 428: 68(1998); J. C. M. van Hest, K. L. Kiick 및 D. A. Tirrell, *J. Am. Chem. Soc.*, 122: 1282(2000); 및 K. L. Kiick 및 D. A. Tirrell, *Tetrahedron*, 56: 9487(2000)]; 미국 특허 제6,586,207호; 및 미국 특허 공보 제2002/0042097호를 참고한다.

[0246] 이 방법의 성공은 아미노아실-tRNA 합성효소에 의한 비천연 아미노산 유사체의 인식에 따라 좌우되며, 이는 일

반적으로 단백질 번역의 적합도를 보증하는 높은 선택성을 필요로 한다. 이 방법의 범위를 확장하는 한 방법은 아미노아실-tRNA 합성효소의 기질 특이성을 완화시키는 것이며, 이는 제한된 수의 경우들에서 달성되었다. 예를 들어, 에세리키아 콜라이 페닐알라닌-tRNA 합성효소(PheRS)에서 Gly에 의해 Ala<sup>294</sup>를 대체하는 경우 기질 결합 포켓의 크기를 증가시켜 p-Cl-페닐알라닌(p-Cl-Phe)에 의해 tRNA<sup>Phe</sup>를 아실화시킨다. [M. Ibba, P. Kast 및 H. Hennecke, *Biochemistry*, 33: 7107(1994)]을 참고한다. 이 돌연변이체 PheRS를 정박시키는 에세리키아 콜라이 균주는 페닐알라닌 대신 p-Cl-페닐알라닌 또는 p-Br-페닐알라닌을 혼입시킨다. 예를 들어, [M. Ibba 및 H. Hennecke, *FEBS Lett.*, 364: 272(1995); 및 N. Sharma, R. Furter, P. Kast 및 D. A. Tirrell, *FEBS Lett.*, 467: 37(2000)]을 참고한다. 유사하게, 에세리키아 콜라이 티로실-tRNA 합성효소의 아미노산 결합 부위 근처의 점 돌연변이 Phe130Ser는 아자티로신을 티로신보다 유효하게 혼입시키는 것으로 밝혀졌다. [F. Hamano-Takaku, T. Iwama, S. Saito-Yano, K. Takaku, Y. Monden, M. Kitabatake, D. Soll 및 S. Nishimura, *J. Biol. Chem.*, 275: 40324(2000)]을 참고한다.

[0247]

비천연 아미노산을 생체 내에서 단백질에 혼입시키는 또 다른 기법은 교정(proofreading) 기전을 갖는 합성효소를 변형시키는 것이다. 이러한 합성효소는 식별할 수 없기 때문에 동족 천연 아미노산과 구조적으로 유사한 아미노산을 활성화시킨다. 이러한 오류는 별도의 부위에서 정정되며, tRNA로부터 잘못 충전된 아미노산을 탈아실화하여 단백질 번역의 적합도를 유지시킨다. 상기 합성효소의 교정 활성이 불능인 경우, 잘못 활성화된 구조적 유사체는 편집(editing) 기능을 벗어나 혼입될 수 있다. 최근, 이 접근은 발릴-tRNA 합성효소(VaIRS)를 사용하여 입증되었다. [V. Doring, H. D. Mootz, L. A. Nangle, T. L. Hendrickson, V. de Crecy-Lagard, P. Schimmel 및 P. Marliere, *Science*, 292: 501(2001)]을 참고한다. VaIRS는 Cys, Thr 또는 아미노부티레이트(Abu)를 사용하여 tRNA<sup>Val</sup>을 잘못 아미노아실화할 수 있고, 이러한 비동족 아미노산들은 추후 수정 도메인에 의해 가수분해된다. 에세리키아 콜라이 염색체의 무작위 돌연변이 유발 후, VaIRS의 수정 부위에서 돌연변이를 갖는 돌연변이체 에세리키아 콜라이 균주를 선택하였다. 이러한 수정-결손 VaIRS는 tRNA<sup>Val</sup>을 Cys로 잘못 충전시킨다. Abu는 입체적으로 Cys와 유사하게 때문에(Cys의 -SH 기는 Abu 중에서 -CH<sub>3</sub>로 대체됨), 돌연변이체 VaIRS는 또한 이 돌연변이체 에세리키아 콜라이 균주가 Abu의 존재 하에서 성장하는 경우 Abu를 단백질에 혼입시킨다. 질량 분광 분석은 본래의 단백질의 각각의 발린 위치에서 발린 중 약 24%가 Abu로 대체됨을 나타내고 있다.

[0248]

또한, 고체상 합성 및 반합성 방법도 신규한 아미노산을 함유하는 다수의 단백질을 합성할 수 있게 하였다. 예를 들어, 하기와 같은 공보 및 여기에 인용된 참고 문헌을 참고한다: [Crick, F. J. C., Barrett, L. Brenner, S. Watts-Tobin, R. *General nature of the genetic code for Proteins*, *Nature*, 192: 1227-1232(1961); Hofmann, K., Bohn, H. *Studies on Polypeptides. XXXVI. The effect of pyrazole-imidazole replacements on the S-Protein activating potency of an S-peptide fragment*, *J. Am Chem*, 88(24): 5914-5919(1966); Kaiser, E. T. *Synthetic approaches to Biologically Active peptides and Proteins including enzymes*, *Acc Chem Res*, 22: 47-54(1989); Nakatsuka, T., Sasaki, T., Kaiser, E. T. *Peptide segment coupling catalyzed by the semisynthetic enzyme thiosubtilisin*, *J Am Chem Soc*, 3808-3810(1987); Schnolzer, M., Kent, S B H. *Constructing Proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease*, *Science*, 256(5054): 221-225(1992); Chaiken, I. M. *Semisynthetic peptides and Proteins*, *CRC Crit Rev Biochem*, 11(3): 255-301(1981); Offord, R. E. *Protein engineering by chemical means?* *Protein Eng.*, 1(3): 151-157(1987); 및 Jackson, D. Y., Burnier, J., Quan, C., Stanley, M., Tom, J., Wells, J. A. *A Designed Peptide Ligase for Total Synthesis of Ribonuclease A with Unnatural Catalytic Residues*, *Science*, 266(5183): 243(1994)].

[0249]

보조인자, 스핀 표지 및 올리고뉴클레오타이드를 비롯한 다양한 비천연 측쇄를 시험관 내에서 단백질에 도입시키는데 화학적 변형을 사용하였다. 예를 들어, [Corey, D. R., Schultz, P. G. *Generation of a hybrid sequence-specific single-stranded deoxyribonuclease*, *Science*, 238(4832): 1401-1403(1987); Kaiser, E. T., Lawrence D. S., Rokita, S. E. *The chemical modification of enzymatic specificity*, *Annu Rev Biochem*, 54: 565-595(1985); Kaiser, E. T., Lawrence, D. S. *Chemical mutation of enzyme active sites*, *Science*, 226(4674): 505-511(1984); Neet, K. E., Nanci A, Koshland, D. E. *Properties of thiol-subtilisin*, *J Biol. Chem*, 243(24): 6392-6401(1968); Polgar, L. et., M. L. Bender. *A new enzyme containing a synthetically formed active site. Thiol-subtilisin*, *J. Am Chem Soc*, 88: 3153-3154(1966); 및 Pollack, S. J., Nakayama, G. Schultz, P. G. *Introduction of nucleophiles and spectroscopic probes into antibody combining sites*, *Science*, 242(4881): 1038-1040(1988)]을 참고한다.

[0250] 면역반응성에 의한 폴리펩티드 정의

[0251] 본 발명의 폴리펩티드가 (예컨대, 본원의 번역 시스템에서 합성된 단백질의 경우에는, 선택된 아미노산(예컨대, 비천연 아미노산)을 포함하고, 예컨대 신규 합성효소, 표준 아미노산의 신규 서열을 포함하는 다양한 신규 폴리펩티드 서열을 제공하기 때문에, 폴리펩티드는 또한 예컨대 면역학적 검정으로 인식될 수 있는 신규 구조적 특성을 제공한다. 본 발명의 폴리펩티드에 특이적으로 결합하는 항혈청, 및 그러한 항혈청에 의해 결합되는 폴리펩티드의 생산은 본 발명의 한 특징이다. 본원에 사용되는 용어 "항체"는 분석물(항원)에 특이적으로 결합하고 인식하는, 면역글로불린 유전자 또는 면역글로불린 유전자 또는 그것의 단편에 의해 실질적으로 코딩되는 폴리펩티드를 포함하나 이에 한정되지 않는다. 그 예는 다클론성, 단클론성, 키메라 및 단일 사슬 항체 등을 포함한다. Fab 단편, 및 파지 표지를 포함하는 발현 라이브러리에 의해 생산되는 단편을 포함한, 면역글로불린의 단편도 또한 본원에 사용되는 용어 "항체"에 포함된다. 예컨대 [Paul, Fundamental Immunology, 4th Ed., 1999, Raven Press, New York, for antibody structure and terminology]를 포함한다.

[0252] 예를 들어, 본 발명은 아미노산 서열을 포함하는 면역원에 대해 생산된 항체 또는 항혈청에 특이적으로 면역반응성이거나 그것에 특이적으로 결합하는 본 발명의 tRNA 및/또는 RS를 이용하여 제조된 단백질 및 RS를 포함한다. 다른 동족체와의 교차 반응성을 제거하기 위해, 항체 또는 항혈청을 이용가능한 단백질, 예컨대 야생형 폴리펩티드, 예컨대 "조절" 폴리펩티드 차감한다. 야생형 단백질이 핵산에 상응하는 경우, 핵산에 의해 코딩되는 폴리펩티드가 생산되고, 항체/항혈청 차감 목적을 위해 사용된다.

[0253] 한 전형적인 형태로, 면역검정은 하나 이상의 폴리펩티드 또는 이의 실질적 하위서열(즉, 제공된 전장 서열의 약 30% 이상)에 대해 상승된 다클론성 항혈청을 사용한다. 단백질 유래의 잠정적 폴리펩티드 면역원의 세트는 이하 "면역원성 폴리펩티드"로 집합적으로 지칭된다. 수득되는 항혈청은 임의적으로 대조 합성효소 동족체에 대한 낮은 교차 반응성을 가지도록 선택되고, 임의의 그러한 교차 반응성은 면역검정에 다클론성 항혈청을 사용하기 전에, 대조 동족체들 중 하나 이상을 사용하여, 예컨대 면역흡수에 의해 제거된다.

[0254] 면역검정에 사용하기 위한 항혈청을 생산하기 위해, 본원에 기재된 바대로 면역원성 폴리펩티드들 중 하나 이상을 생산하여 정제한다. 예를 들어, 재조합 단백질은 재조합 세포에서 생산된다. (마우스의 실제 유전적 동일성으로 인해 결과가 더욱 재현가능하기 때문에 본 검정에 사용되는) 마우스의 교배 균주를, 표준 보조제, 예컨대 프로인트 보조제(Freund's adjuvant), 및 표준 마우스 면역화 프로토콜과 조합하여 면역원성 단백질(들)로 면역화한다(예컨대, 특이적 면역반응성을 결정하기 위해 사용될 수 있는 항체 생산, 면역검정 형태 및 조건의 표준 기재에 대해, [Harlow and Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York] 참고).

[0255] 항체에 대한 부가적 참고문헌 및 논의가 또한 본원에 나와 있고, 이는 면역반응성에 의해 폴리펩티드를 한정하기 위해 여기에 적용될 수 있다. 대안적으로, 본원에 개시된 서열로부터 유래된 하나 이상의 합성 또는 재조합 폴리펩티드는 운반체 단백질에 접합되어, 면역원으로 사용된다.

[0256] 다클론성 혈청을 수집하여, 면역검정, 예를 들어 고체 지지체 상에 고정화된 면역원성 단백질 중 하나 이상을 이용한 고체상 면역검정에 대해 적정한다.  $10^6$  이상의 역가를 갖는 다클론성 항혈청을 선택하고, 폴링하여, 대조 합성효소 폴리펩티드를 차감함으로써, 차감되고 폴링되며 적정된 다클론성 항혈청을 생산한다.

[0257] 차감되고 폴링되며 적정된 다클론성 항혈청을 비교 면역검정에서 대조 동족체에 대한 교차 반응성에 대해 시험한다. 이 비교 검정에서, 대조 합성효소 동족체에 대한 결합에 비해, 면역원성 단백질에 대한 적정된 다클론성 항혈청의 결합에 대해 적어도 약 5-10 배 더 높은 신호 대 노이즈 비를 초래하는, 차감되고 적정된 다클론성 항혈청에 대한 차별적 결합 조건을 결정한다. 즉, 결합 반응의 엄격성은 알부민 또는 비지방 건조 밀크와 같은 비특이적 경쟁자의 첨가, 및/또는 염 조건, 온도 등의 조정에 의해 조정될 수 있다. 이 결합 조건은, 시험 폴리펩티드(폴리펩티드는 면역원성 폴리펩티드 및/또는 대조 폴리펩티드와 비교됨)이, 폴링되고 차감된 다클론성 항혈청에 의해 특이적으로 결합되는지의 여부를 결정하기 위한 후속 검정에 사용된다.

[0258] 또 다른 예에서, 경쟁적 결합 형태에서 면역검정을 시험 폴리펩티드의 검출에 사용한다. 예를 들어, 나와 있는 바와 같이, 대조 폴리펩티드를 이용한 면역흡수에 의해, 교차 반응 항체를 폴링된 항혈청 혼합물에서 제거한다. 이어서, 면역원성 폴리펩티드(들)를, 차감되고 폴링된 항혈청에 노출된 고체 지지체에 고정화한다. 시험 단백질을 차감되고 폴링된 항혈청에 결합하기 위해 경쟁하도록 검정에 부가한다. 고정화된 단백질(들) 대비, 차감되고 폴링된 항혈청에 대해 경쟁하는 시험 단백질(들)의 능력을, 결합을 위해 경쟁하도록 검정에 부가된 면역원성 폴리펩티드(들)(면역원성 폴리펩티드는 폴링된 항혈청에 결합하기 위해 고정화된 면역원성 폴리펩티드와 효과적인



로 경쟁함)의 능력과 비교한다. 표준 계산법을 이용하여, 시험 단백질에 대한 교차 반응도 %를 계산한다.

[0259] 병행 검정에서, 폴링되고 차감된 항혈청에 결합하기 위해 경쟁하는 대조 단백질의 능력을, 임의적으로 항혈청에 결합하기 위해 경쟁하는 면역원성 폴리펩티드(들)의 능력과 비교하여 결정한다. 이에 다시, 표준 계산법을 이용하여, 대조 폴리펩티드에 대한 교차 반응도 %를 계산한다. 교차 반응도 %가 대조 폴리펩티드에 비해 시험 폴리펩티드의 교차 반응도 %의 적어도 5-10x인 경우, 및/또는 시험 폴리펩티드의 결합이 면역원성 폴리펩티드의 결합 범위 내인 경우, 시험 폴리펩티드는 폴링되고 차감된 항혈청에 특이적으로 결합하는 것으로 일컬어진다.

[0260] 일반적으로, 면역흡수되고 폴링된 항혈청을 본원에 기재된 경쟁적 결합 면역검정에 사용하여, 임의의 시험 폴리펩티드를 면역원성 및/또는 대조 폴리펩티드(들)와 비교한다. 이 비교를 행하기 위해, 면역원성, 시험 및 대조 폴리펩티드는 광범위한 농도 범위에서 각기 검정되고, 예컨대 면역화된 대조, 시험 또는 면역원성 단백질에 대한 차감된 항혈청의 결합을 50% 억제하는데 필요한 각 폴리펩티드의 양은 표준 기법을 이용하여 결정된다. 경쟁 검정에서 결합을 위해 필요한 시험 폴리펩티드의 양이, 필요한 면역원성 폴리펩티드의 양의 2배 미만인 경우, 시험 폴리펩티드는 생산된 항체 면역원성 단백질에 대해 발생된 항체에 특이적으로 결합하는 것으로 일컬어지고, 단 그 양은 대조 폴리펩티드에서의 양의 약 5-10x로 높다.

[0261] 한 부가적 특이성 결정으로서, 폴링된 항혈청은, 면역흡수에 사용되는 면역원성 폴리펩티드(들)에 대한 생산된 면역원성 폴리펩티드 차감의 폴링된 항혈청의 결합이 검출가능하지 않거나 거의 검출가능하지 않을 때까지, (대조 폴리펩티드보다는) 면역원성 폴리펩티드(들)로 임의적으로 완전히 면역흡수된다. 이 완전 면역흡수된 항혈청을 시험 폴리펩티드과의 반응성에 대해 시험한다. 반응성이 관찰되지 않거나 거의 관찰되지 않을 경우(즉, 면역원성 폴리펩티드에 대한 완전 면역흡수된 항혈청의 결합에 대해 관찰되는 신호 대 노이즈 비의 2x 이하인 경우), 시험 폴리펩티드는 면역원성 단백질에 의해 도출되는 항혈청에 의해 특이적으로 결합된다.

[0262] 단백질, 항체, 항혈청 등에 대한 부가적 상세 내용을 USSN 제10/825,867호(발명의 명칭: "EXPANDING THE EUKARYOTIC GENETIC CODE"), WO 2002/085923(발명의 명칭: "In vivo incorporation of unnatural amino acids"); 미국 특허 제6,927,042호(발명의 명칭: "Glycoprotein synthesis"); 및 미국 특허 출원 공보 제2004/0198637호(발명의 명칭: "Protein Arrays")(본원에 참고로 인용됨)에서 찾아볼 수 있다.

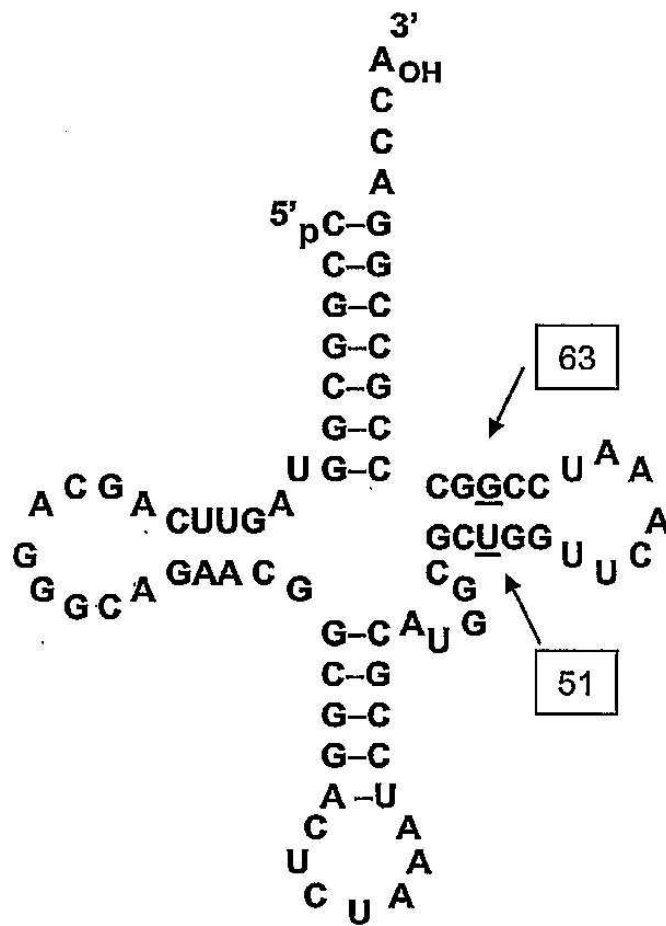
[0263] 키트

[0264] 키트도 또한 본 발명의 한 특징이다. 예를 들어, 세포에서 하나 이상의 선택된 아미노산, 예컨대 비천연 아미노산을 포함하는 단백질을 생산하기 위한 키트로서, O-tRNA를 코딩하는 폴리뉴클레오티드 서열, 및/또는 O-tRNA, 및/또는 O-RS를 코딩하는 폴리뉴클레오티드 서열, 및/또는 O-RS를 함유하는 용기를 포함하는 키트를 제공한다. 한 실시양태에서, 키트는 적어도 선택된 아미노산을 포함한다. 또 다른 실시양태에서, 키트는 본 발명의 아미노아실화된 tRNA를 포함한다. 또 다른 실시양태에서, 키트는 단백질을 생산하기 위한 지시 재료를 추가로 포함한다.

[0265] 한 부가적 예는, 무세포 번역 시스템에서 하나 이상의 선택된 아미노산, 예컨대 비천연 아미노산을 포함하는 단백질을 생산하기 위한 키트로서, O-tRNA를 코딩하는 폴리뉴클레오티드 서열, 및/또는 O-tRNA, 및/또는 O-RS를 코딩하는 폴리뉴클레오티드 서열, 및/또는 O-RS를 함유하는 용기를 포함하는 키트를 제공한다. 한 실시양태에서, 키트는 선택된 아미노산을 추가로 포함한다. 또 다른 실시양태에서, 키트는 본 발명의 아미노아실화된 tRNA를 포함한다. 또 다른 실시양태에서, 키트는 단백질을 생산하기 위한 지시 재료를 추가로 포함한다.

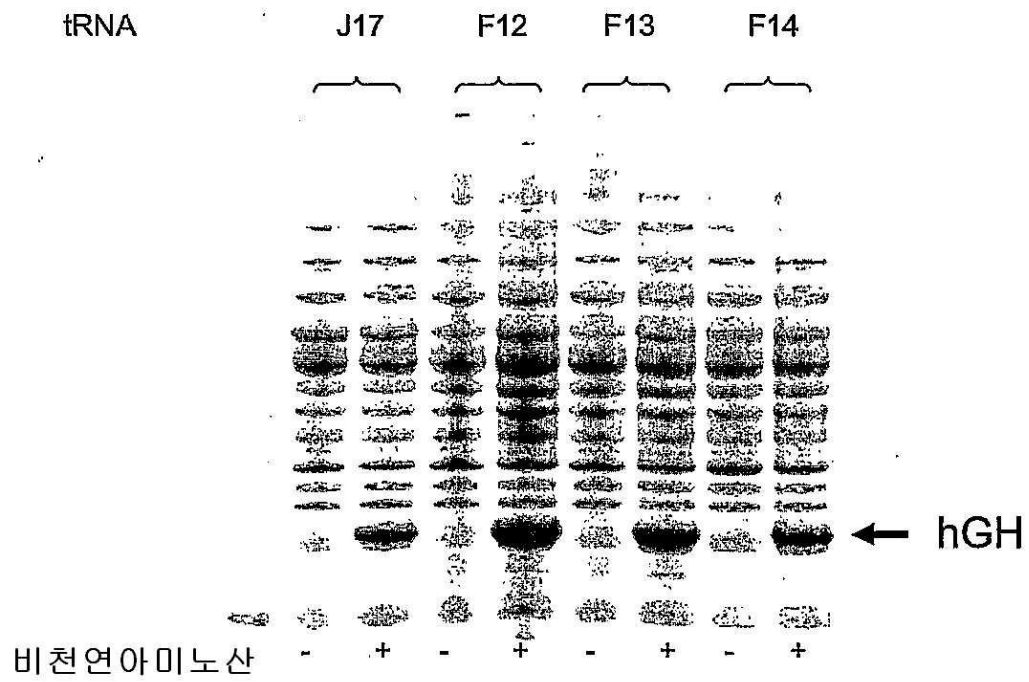
도면

도면1

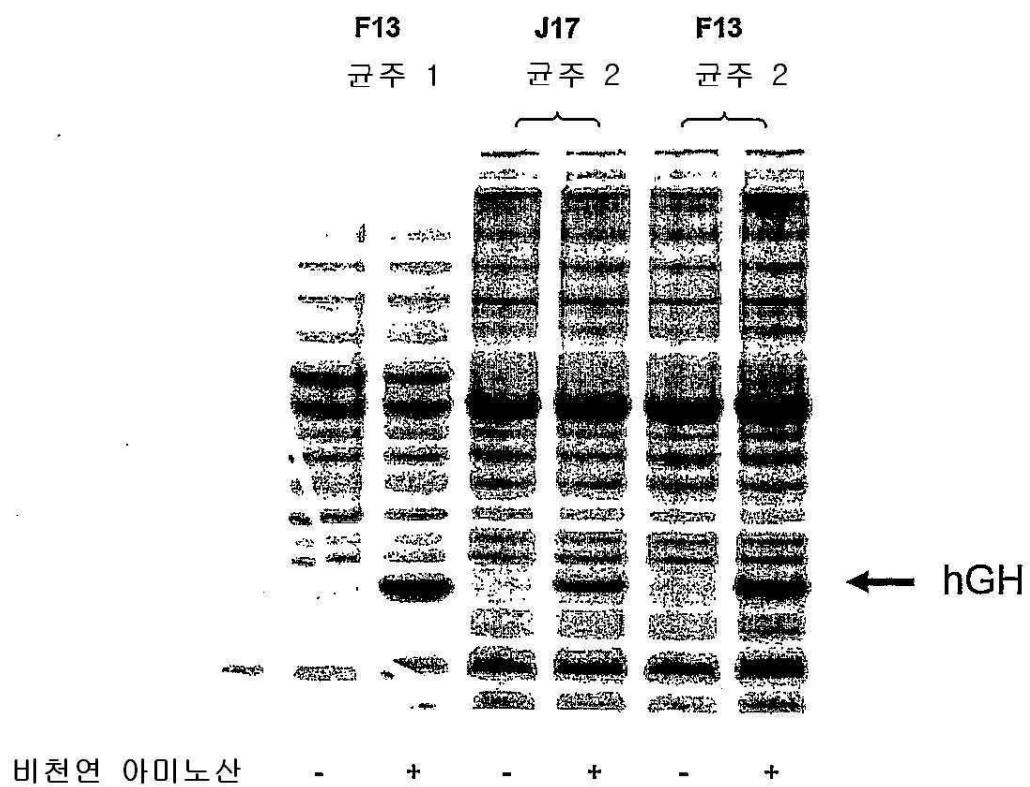


TΨC 줄기

도면2



도면3



서 열 목 록

SEQUENCE LISTING

<110> Paulsel, Andrew  
Cho, Ho S.

<120> Compositions of Aminoacyl-tRNA Synthetase and Uses Thereof

<130> AMBX-0070.00PCT

<150> US 60/639,146  
<151> 2004-12-22

<160> 17

<170> PatentIn version 3.3

<210> 1  
<211> 77  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Mutant tRNA derived from Methanococcus jannaschii tRNA

<400> 1  
ccggcggtag ttcagcaggg cagaacggcg gactctaaat ccgcatggcg ccggttcaaa 60

tccggcccg cggacca 77

<210> 2  
<211> 77  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Mutant tRNA derived from Methanococcus jannaschii tRNA

<400> 2  
ccggcggtag ttcagcaggg cagaacggcg gactctaaat ccgcatggcg ctggttcaaa 60



tccagcccg cggacca 77

<210> 3  
<211> 77  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Mutant tRNA derived from Methanococcus jannaschii tRNA

<400> 3  
ccggcggtag ttcagcaggg cagaacggcg gactctaaat ccgcatggcg caggttcaaa 60

tcctgcccc cggacca 77

<210> 4  
<211> 921  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Mutant synthetase derived from Methanococcus jannaschii synthetase

<400> 4  
atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60

agagaggttt taaaaaaaga tgaaaaatct gctgttatag gttttgaacc aagtggtaaa 120

atacatttag ggcattatct ccaaataaaa aagatgattg atttcaaaa tgctggattt 180

gatataatta tatatttggc tgatttacac gcctatttaa accagaaagg agagttggat 240

gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300

aaatatgttt atggaagtga acatggtctt gataaggatt atacactgaa tgtctataga 360

ttggctttaa aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420

gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcaggttaa tgggattcat 480

tatgagggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540

agggagcttt taccaaaaaa gggtgtttgt attcacaacc ctgtcttaac gggtttggat 600

ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660

gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720

ataatggaga tagctaaata cttccttgaa tatcctttaa ccataaaaag gccagaaaaa 780

tttgggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840

gaattgcac caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900

ccaattagaa agagattata a 921

<210> 5  
 <211> 306  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant synthetase derived from Methanococcus jannaschii synthetase

<400> 5

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 Val  
 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

Tyr 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 His Gly 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 Glu 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> 6  
<211> 88  
<212> DNA  
<213> Halobacterium sp. NRC-1

<400> 6  
cccagggtag ccaagctcgg ccaacggcga cggactctaa atccgttctc gtaggagttc 60

gagggttcga atcccttccc tgggacca 88

<210> 7  
<211> 88  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Orthogonal tRNA that recognizes an opal codon

<400> 7  
gcgggggttg ccgagcctgg ccaaaggcgc cggacttcaa atccggtccc gtaggggttc 60

cggggttcaa atccccgccc ccgcacca 88

<210> 8  
<211> 77  
<212> DNA  
<213> Methanococcus jannaschii



<400> 8  
ccggcggtag ttcagcaggg cagaacggcg gactctaaat ccgcatggcg ctggttcaaa 60

tccggccccg cggacca 77

<210> 9  
<211> 65  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic oligonucleotide primer

<400> 9  
gtaacgtga attcccgcg gtagttcagc agggcagaac ggccgactct aaatccgcat 60

ggcgc 65

<210> 10  
<211> 67  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic oligonucleotide primer

<400> 10  
gatctgcagt ggtccggcgg gccggatttg aaccggcgcc atgcggattt agagtccgcc 60

gttctgc 67

<210> 11  
<211> 67  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic oligonucleotide primer

<400> 11  
gatctgcagt ggtccggcgg gctggatttg aaccagcgcc atgcggattt agagtccgcc 60

gttctgc 67

<210> 12  
 <211> 67  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide primer

<400> 12  
 gatctgcagt ggtccggcgg gcaggatttg aacctgcgcc atgcggattt agagtccgcc 60

gttctgc 67

<210> 13  
 <211> 49  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide primer

<400> 13  
 cgccggacca ctgcagatcc ttagcgaaag ctaaggattt tttttaagc 49

<210> 14  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide primer

<400> 14  
 caaattcgtc catatgggat tcc 23

<210> 15  
 <211> 20  
 <212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide primer

<400> 15

gtaacgctga attcccggcg 20

<210> 16

<211> 600

<212> DNA

<213> Homo sapiens

<400> 16

atgggccacc accaccacca ccacttccca accattccct tatccaggct ttttgacaac 60

gctatgctcc ggccecatcg tctgcaccag ctggcctttg acacctacca ggagtttgaa 120

gaagcctaga tcccaaagga acagaagtat tcattcctgc agaaccceca gacctccctc 180

tgtttctcag agtctattcc gacaccctcc aacaggaggagg aaacacaaca gaaatccaac 240

ctagagctgc tccgcatctc cctgtgtctc atccagtcgt ggctggagcc cgtgcagtgc 300

ctcaggagtg ttttcgcca cagcctgggtg tacggcgctc ctgacagcaa cgtctatgac 360

ctcctaaagg acctagagga aggcacccaa acgctgatgg ggaggctgga agatggcagc 420

ccccggactg ggcagatctt caagcagacc tacagcaagt tcgacacaaa ctcacacaac 480

gatgacgcac tactcaagaa ctacgggctg ctctactgct tcaggaagga catggacaag 540

gtcgagacat tctgcgcat cgtgcagtgc cgctctgtgg agggcagctg tggcttctaa 600

<210> 17

<211> 306

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant synthetase derived from Methanococcus jannaschii synthetase

<400> 17

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 Val  
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

Tyr 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 His Gly 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 Glu 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 Arg 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