

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-537984

(P2004-537984A)

(43) 公表日 平成16年12月24日(2004. 12. 24)

(51) Int.Cl.⁷

C 1 2 N 15/09

C 1 2 N 9/00

C 1 2 Q 1/02

F I

C 1 2 N 15/00

C 1 2 N 9/00

C 1 2 Q 1/02

Z N A A

テーマコード (参考)

4 B O 2 4

4 B O 5 0

4 B O 6 3

審査請求 未請求 予備審査請求 未請求 (全 282 頁)

(21) 出願番号 特願2002-583590 (P2002-583590)
 (86) (22) 出願日 平成14年4月19日 (2002. 4. 19)
 (85) 翻訳文提出日 平成15年10月20日 (2003. 10. 20)
 (86) 国際出願番号 PCT/US2002/012635
 (87) 国際公開番号 W02002/086075
 (87) 国際公開日 平成14年10月31日 (2002. 10. 31)
 (31) 優先権主張番号 60/285, 030
 (32) 優先日 平成13年4月19日 (2001. 4. 19)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 60/355, 514
 (32) 優先日 平成14年2月6日 (2002. 2. 6)
 (33) 優先権主張国 米国 (US)

(71) 出願人 593052785
 ザ スクリップス リサーチ インスティ
 テュート
 アメリカ合衆国 カリフォルニア州 92
 037 ラ ジョラ ノース トーリー
 パインズ ロード 10550
 (74) 代理人 100064355
 弁理士 川原田 一穂
 (72) 発明者 ベーター・シュルツ
 アメリカ合衆国 カリフォルニア州 92
 037 ラ ジョラ ラ ジョラ ランチ
 ロード 1650

最終頁に続く

(54) 【発明の名称】 直交 tRNA-アミノアシル tRNA シンテターゼ対を生産するための方法及び組成物

(57) 【要約】

本発明は直交 tRNA と、直交アミノアシル tRNA シンテターゼと、直交 tRNA / シンテターゼ対を含むタンパク質生合成機構の成分を作製するための組成物と方法を提供する。直交対を同定するための方法も提供する。これらの成分は非天然アミノ酸をタンパク質に *in vivo* 組込むために使用することができる。

【特許請求の範囲】

【請求項 1】

直交アミノアシル tRNA シンテターゼ (O-RS) を含む組成物であって、O-RS が直交 tRNA (O-tRNA) を非天然アミノ酸で優先的にアミノアシル化する前記組成物。

【請求項 2】

O-RS が配列番号 4 ~ 34 及びその相補的ポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸を含む請求項 1 に記載の組成物。

【請求項 3】

O-RS が O-tRNA を非天然アミノ酸で *in vivo* アミノアシル化する請求項 1 に記載の組成物。 10

【請求項 4】

非天然アミノ酸が O-メチル-L-チロシン、L-3-(2-ナフチル)アラニン、3-メチルフェニルアラニン、O-4-アリル-L-チロシン、4-プロピル-L-チロシン、トリ-O-アセチル-GlcNAc β -セリン、L-Dopa、フッ素化フェニルアラニン、イソプロピル-L-フェニルアラニン、p-アジド-L-フェニルアラニン、p-アシル-L-フェニルアラニン、p-ベンゾイル-L-フェニルアラニン、L-ホスホセリン、ホスホセリン、ホスホチロシン、p-ヨードフェニルアラニン、p-ブロモフェニルアラニン、p-アミノ-L-フェニルアラニン、及びイソプロピル-L-フェニルアラニンから構成される群から選択される請求項 1 に記載の組成物。 20

【請求項 5】

非天然アミノ酸がチロシンアミノ酸の非天然類似体；グルタミンアミノ酸の非天然類似体；フェニルアラニンアミノ酸の非天然類似体；セリンアミノ酸の非天然類似体；スレオニンアミノ酸の非天然類似体；アルキル、アリール、アシル、アジド、シアノ、ハロ、ヒドラジン、ヒドラジド、ヒドロキシル、アルケニル、アルキニル、エーテル、チオール、スルホニル、セレノ、エステル、チオ酸、硼酸、ボロン酸、ホスホ、ホスホノ、ホスフィン、複素環、エノン、イミン、アルデヒド、ヒドロキシルアミン、ケトもしくはアミノ置換アミノ酸又はその任意組合せ；光架橋基をもつアミノ酸；スピンラベル付きアミノ酸；蛍光アミノ酸；新規官能基をもつアミノ酸；別の分子と共有又は非共有的に相互作用するアミノ酸；金属結合性アミノ酸；金属含有アミノ酸；放射性アミノ酸；フォトケージドアミノ酸、光異性化可能なアミノ酸；ビオチン又はビオチン類似体含有アミノ酸；グリコシル化又は糖鎖修飾アミノ酸；ケト含有アミノ酸；ポリエチレングリコールを含むアミノ酸；ポリエーテルを含むアミノ酸；重原子置換アミノ酸；化学分解性又は光分解性アミノ酸；延長側鎖をもつアミノ酸；毒性基を含むアミノ酸；糖鎖置換アミノ酸；糖鎖置換セリン；炭素結合糖含有アミノ酸；レドックス活性アミノ酸； α -ヒドロキシ含有酸；アミノチオ酸含有アミノ酸； α , α ジ置換アミノ酸； β アミノ酸；及びプロリン以外の環状アミノ酸から構成される群から選択される請求項 1 に記載の組成物。 30

【請求項 6】

O-RS の K_m と K_{cat} から構成される群から選択される 1 種以上の酵素特性が天然アミノ酸よりも非天然アミノ酸に対して改善又は強化されている請求項 1 に記載の組成物。 40

【請求項 7】

コーディングポリヌクレオチド配列によりコードされるアミノ酸配列を含むポリペプチドであって、コーディングポリヌクレオチド配列が、

- a) 配列番号 4 ~ 34 から選択されるコーディングポリヌクレオチド配列、
- b) 配列番号 35 ~ 66 から選択されるポリペプチドをコードするコーディングポリヌクレオチド配列、
- c) (a) 又は (b) のポリヌクレオチド配列の実質的に全長にわたって高ストリンジェント条件下でハイブリダイズするポリヌクレオチド配列、及び
- d) (a)、(b) 又は (c) の相補的配列から構成される群から選択される前記ポリペプチド。

【請求項 8】

コードされるポリペプチドが直交アミノアシル tRNA シンテターゼをコードする請求項 7 に記載のポリペプチド。

【請求項 9】

配列番号 35～66 から選択されるアミノ酸配列を含むポリペプチド。

【請求項 10】

a) 配列番号 1～3 から選択されるポリヌクレオチド配列又はその相補的ポリヌクレオチド配列、及び

b) ポリヌクレオチド配列 (a) の実質的に全長にわたって高ストリンジェント条件下でハイブリダイズするポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸。 10

【請求項 11】

ポリヌクレオチド配列が直交 tRNA を含む請求項 10 に記載の核酸。

【請求項 12】

ポリヌクレオチド配列が直交アミノアシル tRNA シンテターゼと相補対を形成する請求項 10 に記載の核酸。

【請求項 13】

直交アミノアシル tRNA シンテターゼのポリペプチド配列が配列番号 35～66 から構成される群から選択される請求項 12 に記載の核酸。

【請求項 14】

直交 tRNA (O-tRNA) を含む組成物であって、O-tRNA がセクターコドン を認識し、O-tRNA が直交アミノアシル tRNA シンテターゼにより非天然アミノ酸 で優先的にアミノアシル化される前記組成物。 20

【請求項 15】

O-tRNA が配列番号 1～3 及びその相補的ポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸を含む請求項 14 に記載の組成物。

【請求項 16】

セクターコドンがユニーク 3 塩基コドン、ナンセンスコドン、レアコドン、非天然コドン又は少なくとも 4 塩基コドンを含む請求項 14 に記載の組成物。

【請求項 17】

直交アミノアシル tRNA シンテターゼ (O-RS) を更に含む請求項 14 に記載の組成物。 30

【請求項 18】

O-tRNA と O-RS が相補的である請求項 17 に記載の組成物。

【請求項 19】

組成物が mutRNA Tyr-mutTyrRS 対を含む請求項 17 に記載の組成物。

【請求項 20】

組成物が mutRNA Tyr-SS12TyrRS 対を含む請求項 19 に記載の組成物。

【請求項 21】

組成物が mutRNA Leu-mutLeuRS 対を含む請求項 17 に記載の組成物。 40

【請求項 22】

組成物が mutRNA Thr-mutThrRS 対を含む請求項 17 に記載の組成物。

【請求項 23】

組成物が mutRNA Glu-mutGluRS 対を含む請求項 17 に記載の組成物。

【請求項 24】

O-tRNA と O-RS が少なくとも 1 種の生物に由来する天然 tRNA 及び RS の突然変異により誘導され、少なくとも 1 種の生物が原核生物である請求項 17 に記載の組成物。

【請求項 25】

少なくとも 1 種の生物が *Methanococcus jannaschii*、*Meth* 50

*anobacterium thermoautotrophicum*及び*Halobacterium*から構成される群から選択される請求項24に記載の組成物。

【請求項26】

O-tRNAとO-RSが少なくとも1種の生物に由来する天然tRNA及びRSの突然変異により誘導され、少なくとも1種の生物が真核生物である請求項17に記載の組成物。

【請求項27】

少なくとも1種の生物が酵母、哺乳動物、真菌、昆虫、植物及び原生生物から構成される群から選択される請求項26に記載の組成物。

【請求項28】

O-tRNAが第1の生物に由来する天然tRNAの突然変異により誘導され、O-RSが第2の生物に由来する天然RSの突然変異により誘導される請求項17に記載の組成物。

【請求項29】

O-tRNAとO-RSが少なくとも1種の生物から単離され、少なくとも1種の生物が原核生物である請求項17に記載の組成物。

【請求項30】

少なくとも1種の生物が*Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*及び*Halobacterium*から構成される群から選択される請求項29に記載の組成物。

【請求項31】

O-tRNAとO-RSが少なくとも1種の生物から単離され、少なくとも1種の生物が真核生物である請求項17に記載の組成物。

【請求項32】

少なくとも1種の生物が酵母、哺乳動物、真菌、昆虫、植物及び原生生物から構成される群から選択される請求項31に記載の組成物。

【請求項33】

O-tRNAが第1の生物から単離され、O-RSが第2の生物から単離される請求項17に記載の組成物。

【請求項34】

O-tRNAとO-RSの1種以上が1種以上のライブラリーから単離され、1種以上のライブラリーが1種以上の生物に由来するO-tRNA又はO-RSを含む請求項17に記載の組成物。

【請求項35】

1種以上の生物が原核生物又は真核生物を含む請求項34に記載の組成物。

【請求項36】

組成物が細胞に組込まれている請求項17に記載の組成物。

【請求項37】

組成物が*in vitro*翻訳系を含む請求項1に記載の組成物。

【請求項38】

少なくとも1種の組換え直交アミノアシルtRNAシンテターゼ(O-RS)の生産方法であって、

(a) 第1の生物に由来する少なくとも1種のアミノアシルtRNAシンテターゼ(RS)から誘導されるRSのライブラリーを作製する段階と、

(b) 非天然アミノ酸と天然アミノ酸の存在下に直交tRNA(O-tRNA)をアミノアシル化するメンバーをRSのライブラリーから選択又はスクリーニングすることにより活性RSのプールを提供する段階と、

(c) 非天然アミノ酸の不在下にO-tRNAを優先的にアミノアシル化する活性RSのプールを選択又はスクリーニングすることにより、O-tRNAを非天然アミノ酸で優先的にアミノアシル化する少なくとも1種の組換えO-RSを提供する段階を含む前記方法

10

20

30

40

50

。

【請求項 39】

第1の生物が原核生物である請求項38に記載の方法。

【請求項 40】

第1の生物が *Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum* 及び *Halobacterium* から構成される群から選択される請求項38に記載の方法。

【請求項 41】

少なくとも1種のRSが不活性RSである請求項38に記載の方法。

【請求項 42】

不活性RSが活性RSを突然変異させることにより作製される請求項41に記載の方法。

【請求項 43】

不活性RSがアラニンアミノ酸に突然変異した少なくとも約1、少なくとも約2、少なくとも約3、少なくとも約4、約5、少なくとも約6、又は少なくとも約10個以上のアミノ酸を含む請求項42に記載の方法。

【請求項 44】

RSのライブラリーが突然変異体RSのライブラリーを含む請求項38に記載の方法。

【請求項 45】

段階(a)がランダム突然変異を含む請求項44に記載の方法。

【請求項 46】

段階(b)が、

少なくとも1個のセクターコドンを含むポジティブ選択又はスクリーニングマーカースのライブラリーを複数の細胞に導入し、
選択又はスクリーニング物質の存在下に複数の細胞を増殖させ、
ポジティブ選択又はスクリーニングマーカの少なくとも1個のセクターコドンを抑圧することにより選択又はスクリーニング物質の存在下に生存するか又は特異的スクリーニング応答を示す細胞を同定し、活性RSのプールを含むポジティブ選択又はスクリーニング済み細胞のサブセットを提供することを含む請求項38に記載の方法。

【請求項 47】

選択又はスクリーニング物質濃度を変動させる請求項46に記載の方法。

【請求項 48】

少なくとも1個のセクターコドンがアンバーコドン、オーカーコドン又はオパール終止コドンを含む請求項46に記載の方法。

【請求項 49】

ポジティブ選択マーカークロラムフェニコールアセチルトランスフェラーゼ(CAT)遺伝子であり、少なくとも1個のセクターコドンがCAT遺伝子内のアンバー終止コドンである請求項46に記載の方法。

【請求項 50】

ポジティブ選択マーカークラクタマーゼ遺伝子であり、少なくとも1個のセクターコドンがクラクタマーゼ遺伝子内のアンバー終止コドンである請求項46に記載の方法。

【請求項 51】

ポジティブスクリーニングマーカークラ光又は発光スクリーニングマーカークを含む請求項46に記載の方法。

【請求項 52】

ポジティブスクリーニングマーカークが親和性に基づくスクリーニングマーカークを含む請求項46に記載の方法。

【請求項 53】

段階(c)が、

少なくとも1個のセクターコドンを含むネガティブ選択又はスクリーニングマーカークを段階(b)からの活性RSのプールと共に第2の生物の複数の細胞に導入し、

10

20

30

40

50

非天然アミノ酸と選択又はスクリーニング物質を加えた第1の培地で生存するか又は特異的スクリーニング応答を示すが、非天然アミノ酸と選択又はスクリーニング物質を加えない第2の培地では生存するか又は特異的スクリーニング応答を示すことができない細胞を同定することにより、少なくとも1種の組換えO-R Sをもつ生存又はスクリーニング済み細胞を提供することを含む請求項38に記載の方法。

【請求項54】

選択又はスクリーニング物質濃度を変動させる請求項53に記載の方法。

【請求項55】

第1の生物と第2の生物が異なる請求項53に記載の方法。

【請求項56】

第1の生物と第2の生物が原核生物、真核生物、哺乳動物、大腸菌、真菌、酵母、始原菌、真正細菌、植物、昆虫又は原生生物を含む請求項53に記載の方法。

【請求項57】

ネガティブ選択マーカーが少なくとも1個のセクターコドンを含むクロラムフェニコールアセチルトランスフェラーゼ(CAT)遺伝子を含む請求項53に記載の方法。

【請求項58】

スクリーニングマーカーが蛍光又は発光スクリーニングマーカーを含む請求項53に記載の方法。

【請求項59】

スクリーニングマーカーが親和性に基づくスクリーニングマーカーを含む請求項53に記載の方法。 20

【請求項60】

段階(c)が、
段階(b)からの活性RSのプールを単離し、
少なくとも1個のセクターコドンを含むネガティブ選択又はスクリーニングマーカー遺伝子と活性RSのプールを第2の生物の複数の細胞に導入し、
非天然アミノ酸を加えない第1の培地で生存するか又は特異的スクリーニング応答を示すが、非天然アミノ酸を加えた第2の培地では生存するか又は特異的スクリーニング応答を示すことができない細胞を同定することにより、非天然アミノ酸に特異的な少なくとも1種の組換えO-R Sをもつ生存又はスクリーニング済み細胞を提供することを含む請求項 30
38に記載の方法。

【請求項61】

少なくとも1個のセクターコドンが2個以上のセクターコドンを含む請求項60に記載の方法。

【請求項62】

第1の生物と第2の生物が異なる請求項60に記載の方法。

【請求項63】

第1の生物と第2の生物が原核生物、真核生物、哺乳動物、大腸菌、真菌、酵母、始原菌、真正細菌、植物、昆虫又は原生生物を含む請求項60に記載の方法。

【請求項64】

ネガティブ選択マーカーが少なくとも1個のセクターコドンを含むリボヌクレアーゼバルナーゼ遺伝子を含む請求項60に記載の方法。 40

【請求項65】

スクリーニングマーカーが蛍光又は発光スクリーニングマーカーを含む請求項60に記載の方法。

【請求項66】

スクリーニングマーカーが親和性に基づくスクリーニングマーカーを含む請求項60に記載の方法。

【請求項67】

段階(b)、(c)又は段階(b)及び(c)の両方が選択又はスクリーニングストリン 50

ジェンシーを変動させることを含む請求項 38 に記載の方法。

【請求項 68】

更に、

(d) 少なくとも 1 種の組換え O-R S を単離する段階と、

(e) 少なくとも 1 種の組換え O-R S から誘導され、1 組の突然変異 O-R S を含む第 2 組の O-R S を作製する段階と、

(f) O-tRNA を優先的にアミノアシル化することが可能な突然変異 O-R S が得られるまで段階 (b) 及び (c) を反復する段階を含む請求項 38 に記載の方法。

【請求項 69】

更に段階 (d) ~ (f) を少なくとも約 2 回反復することを含む請求項 68 に記載の方法 10

【請求項 70】

段階 (e) がランダム突然変異誘発、部位特異的突然変異誘発、組換え又はその任意組合せを含む請求項 68 に記載の方法。

【請求項 71】

段階 (b)、(c) 又は段階 (b) 及び (c) の両方がレポーターを使用することを含み、レポーターが蛍光活性化細胞ソーティング (FACS) により検出されるか又はレポーターが発光により検出される請求項 38 に記載の方法。

【請求項 72】

段階 (b)、(c) 又は段階 (b) 及び (c) の両方がレポーターを使用することを含み、レポーターが細胞表面又はファージディスプレイに提示される請求項 38 に記載の方法 20

【請求項 73】

非天然アミノ酸が O-メチル-L-チロシン、L-3-(2-ナフチル)アラニン、3-メチルフェニルアラニン、O-4-アリル-L-チロシン、4-プロピル-L-チロシン、トリ-O-アセチル-GlcNAcβ-セリン、L-Dopa、フッ素化フェニルアラニン、イソプロピル-L-フェニルアラニン、p-アジド-L-フェニルアラニン、p-アシル-L-フェニルアラニン、p-ベンゾイル-L-フェニルアラニン、L-ホスホセリン、ホスホノセリン、ホスホノチロシン、p-ヨードフェニルアラニン、p-ブromoフェニルアラニン、p-アミノ-L-フェニルアラニン、及びイソプロピル-L-フェニルアラニンから構成される群から選択される請求項 38 に記載の方法。 30

【請求項 74】

請求項 38 に記載の方法により生産された少なくとも 1 組の組換え O-R S。

【請求項 75】

組換え直交 tRNA (O-tRNA) の生産方法であって、

(a) 第 1 の生物に由来する少なくとも 1 種の tRNA から誘導される tRNA のライブラリーを作製する段階と、

(b) 第 1 の生物に由来する R S の不在下に第 2 の生物に由来するアミノアシル tRNA シンテターゼ (R S) によりアミノアシル化される tRNA のライブラリーを選択又はスクリーニングすることにより tRNA のプールを提供する段階と、 40

(c) 導入した直交 R S (O-R S) によりアミノアシル化されるメンバーを tRNA のプールから選択又はスクリーニングすることにより、少なくとも 1 個のセクターコドンを認識し、第 2 の生物に由来する R S により効率的に認識されず、O-R S により優先的にアミノアシル化される少なくとも 1 種の組換え O-tRNA を提供する段階を含む前記方法。

【請求項 76】

少なくとも 1 種の tRNA がサプレッサー tRNA である請求項 75 に記載の方法。

【請求項 77】

セクターコドンが天然もしくは非天然塩基のユニーク 3 塩基コドン、ナンセンスコドン、レアコドン、非天然コドン又は少なくとも 4 塩基のコドンを含む請求項 75 に記載の方 50

法。

【請求項 78】

セクターコドンがアンバーコドン、オーカーコドン又はオパール終止コドンを含む請求項 75 に記載の方法。

【請求項 79】

少なくとも 1 種の組換え O-tRNA が改善された直交性をもつ請求項 75 に記載の方法。

【請求項 80】

第 1 の生物が原核生物である請求項 75 に記載の方法。

【請求項 81】

原核生物が *Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum* 及び *Halobacterium* から構成される群から選択される請求項 80 に記載の方法。

【請求項 82】

第 1 の生物と第 2 の生物が原核生物、真核生物、哺乳動物、大腸菌、真菌、酵母、始原菌、真正細菌、植物、昆虫又は原生生物を含む請求項 75 に記載の方法。

【請求項 83】

第 1 の生物と第 2 の生物が異なる請求項 75 に記載の方法。

【請求項 84】

組換え tRNA が非天然アミノ酸によりアミノアシル化され、非天然アミノ酸が天然又は遺伝子操作により *in vivo* 生合成される請求項 75 に記載の方法。

【請求項 85】

非天然アミノ酸を第 1 又は第 2 の生物の 1 種以上の増殖培地に添加する請求項 75 に記載の方法。

【請求項 86】

段階 (b) が、

少なくとも 1 個のセクターコドンを含む毒性マーカー遺伝子又は毒性もしくは静的物質の生産を誘導する遺伝子又は生物に必須の遺伝子と突然変異体 tRNA のライブラリーを第 2 の生物に由来する複数の細胞に導入し、

少なくとも 1 種の直交 tRNA 又は非機能的 tRNA を含む tRNA のプールを含む生存細胞を選択することを含む請求項 75 に記載の方法。

【請求項 87】

生存細胞の選択が比較細胞密度アッセイを含む請求項 86 に記載の方法。

【請求項 88】

セクターコドンが 2 個以上のセクターコドンを含む請求項 86 に記載の方法。

【請求項 89】

マーカー遺伝子がリボヌクレアーゼバルナーゼ遺伝子であり、リボヌクレアーゼバルナーゼ遺伝子が少なくとも 1 種のアンバーコドンを含む請求項 86 に記載の方法。

【請求項 90】

リボヌクレアーゼバルナーゼ遺伝子が 2 種以上のアンバーコドンを含む請求項 89 に記載の方法。

【請求項 91】

段階 (c) が、

少なくとも 1 個のセクターコドンを含む薬剤耐性遺伝子又は生物に必須の遺伝子又は毒性物質の解毒を誘導する遺伝子を含むポジティブ選択又はスクリーニングマーカー遺伝子と、O-RS と、tRNA のプールを第 2 の生物に由来する複数の細胞に導入し、

選択又はスクリーニング物質の存在下に増殖した生存又はスクリーニング済み細胞を同定することにより、O-RS によりアミノアシル化され、少なくとも 1 個のセクターコドンに応答してポジティブマーカー遺伝子によりコードされる翻訳産物にアミノ酸を挿入する少なくとも 1 種の組換え tRNA をもつ細胞プールを提供することを含む請求項 75 に

10

20

30

40

50

記載の方法。

【請求項 9 2】

選択物質が抗生物質である請求項 9 1 に記載の方法。

【請求項 9 3】

選択又はスクリーニング物質濃度を変動させる請求項 9 1 に記載の方法。

【請求項 9 4】

薬剤耐性遺伝子が β ラクタマーゼ遺伝子である請求項 9 1 に記載の方法。

【請求項 9 5】

β ラクタマーゼ遺伝子が少なくとも 1 個のアンバー終止コドンを含む請求項 9 4 に記載の方法。

10

【請求項 9 6】

請求項 7 5 に記載の方法により生産された少なくとも 1 種の組換え O-tRNA。

【請求項 9 7】

少なくとも 1 種の特異的 O-tRNA/O-RS 対の生産方法であって、

(a) 第 1 の生物に由来する少なくとも 1 種の tRNA から誘導される tRNA のライブラリーを作製する段階と、

(b) 第 1 の生物に由来する RS の不在下で第 2 の生物に由来するアミノアシル tRNA シンテターゼ (RS) によりアミノアシル化される tRNA のライブラリーをネガティブスクリーニング又は選択することにより、tRNA のプールを提供する段階と、

(c) 導入した直交 RS (O-RS) によりアミノアシル化されるメンバーを tRNA のプールから選択又はスクリーニングすることにより、セクターコドンを認識し、第 2 の生物に由来する RS により効率的に認識されず、O-RS により優先的にアミノアシル化される少なくとも 1 種の組換え O-tRNA を提供する段階と、

20

(d) 第 3 の生物に由来する少なくとも 1 種のアミノアシル tRNA シンテターゼ (RS) から誘導される RS のライブラリーを作製する段階と、

(e) 非天然アミノ酸と天然アミノ酸の存在下に少なくとも 1 種の組換え O-tRNA を優先的にアミノアシル化するメンバーを RS のライブラリーから選択又はスクリーニングすることにより、活性 RS のプールを提供する段階と、

(f) 非天然アミノ酸の不在下に少なくとも 1 種の組換え O-tRNA を優先的にアミノアシル化する活性 RS のプールをネガティブ選択又はスクリーニングすることにより、非天然アミノ酸に特異的な少なくとも 1 種の組換え O-RS と少なくとも 1 種の組換え O-tRNA を含む少なくとも 1 種の特異的 O-tRNA/O-RS 対を提供する段階を含む前記方法。

30

【請求項 9 8】

請求項 9 7 に記載の方法により生産された特異的 O-tRNA/O-RS 対。

【請求項 9 9】

少なくとも 1 種の特異的 O-tRNA/O-RS 対が mutRNA Tyr-mutTyr RS 対である請求項 9 7 に記載の方法。

【請求項 1 0 0】

第 1 の生物と第 3 の生物が同一である請求項 9 7 に記載の方法。

40

【請求項 1 0 1】

第 1 の生物と第 3 の生物が *Methanococcus jannaschii* である請求項 9 7 に記載の方法。

【請求項 1 0 2】

第 2 の生物の in vivo 翻訳系で使用するための直交 tRNA-tRNA シンテターゼ対の同定方法であって、

マーカー遺伝子と、tRNA と、第 1 の生物から単離又は誘導されるアミノアシル tRNA シンテターゼ (RS) を第 2 の生物に由来する第 1 組の細胞に導入する段階と、

マーカー遺伝子と tRNA を第 2 の生物に由来する複製組の細胞に導入する段階と、

第 1 組と複製組の細胞を選択又はスクリーニング物質の存在下に増殖させ、第 1 組では生

50

存するか又は特異的スクリーニング応答を示すが、複製組では生存するか又は特異的スクリーニング応答を示すことができない細胞を選択又はスクリーニングすることにより、第2の生物の *in vivo* 翻訳系で使用するための直交 tRNA-tRNA シンテターゼ対を含む生存又はスクリーニング済み細胞を提供する段階を含む前記方法。

【請求項103】

比較と選択又はスクリーニングが *in vivo* 相補アッセイを含む請求項102に記載の方法。

【請求項104】

選択又はスクリーニング物質濃度を変動させる請求項102に記載の方法。

【請求項105】

第1の生物が原核生物である請求項102に記載の方法。

【請求項106】

第2の生物が原核生物である請求項102に記載の方法。

【請求項107】

第1の生物と第2の生物が異なる請求項102に記載の方法。

【請求項108】

第1の生物が *Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum* 及び *Halobacterium* から構成される群から選択される請求項102に記載の方法。

【請求項109】

第2の生物が大腸菌である請求項102に記載の方法。

【請求項110】

選択又はスクリーニングがアミノ酸透過性の変化、翻訳効率の変化及び翻訳忠実度の変化から構成される群から選択される1種以上のポジティブ又はネガティブ選択又はスクリーニングを含み、1種以上の変化がタンパク質を生産するために直交 tRNA-tRNA シンテターゼ対を使用する生物における1個以上の遺伝子の突然変異に基づく請求項38、75、97又は102に記載の方法。

【請求項111】

選択又はスクリーニングが1種以上の選択遺伝子内又は1種以上のスクリーニング遺伝子内の少なくとも2個のセクターコドンの選択又はスクリーニングを含む請求項38、75、97又は102に記載の方法。

【請求項112】

少なくとも2個のセクターコドンが同一選択遺伝子又は同一スクリーニング遺伝子に含まれる請求項111に記載の方法。

【請求項113】

少なくとも2個のセクターコドンが異なる選択又はスクリーニング遺伝子に含まれる請求項111に記載の方法。

【請求項114】

少なくとも2個のセクターコドンが異なるセクターコドンを含む請求項111に記載の方法。

【請求項115】

少なくとも2個のセクターコドンが同一セクターコドンを含む請求項111に記載の方法。

ために使用することができる。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、2001年4月19日に提出された米国仮特許出願第60/285,030号、及び2002年2月6日に提出された米国特許出願第60/355,514号の優先権

10

20

30

40

50

を主張し、それらの明細書の全体をここで援用する。

【0002】

連邦政府の後援を受けた研究開発の下で為された発明の権利に関する声明

本発明は、海軍研究事務所からの補助金交付第6502573号及び国立研究所からの補助金交付第GM2159号の下で米国政府の支援により為された。米国政府はこの発明に一定の権利を有する。

【0003】

技術分野

本発明は翻訳生化学の分野に関する。特に、本発明は突然変異直交 tRNA、突然変異直交アミノアシル tRNA シンテターゼ及びその対の生産方法に関する。本発明は非天然アミノ酸をタンパク質に *in vivo* 組込むために使用する直交対を同定するための方法と関連組成物も提供する。

10

【背景技術】

【0004】

タンパク質は光合成からシグナル伝達や免疫応答に至るまでの複雑な生命現象のほぼ全てを行っている。これらの複雑な機能を理解し、制御するためには、タンパク質の構造と機能の関係の理解を深めることが必要である。

【0005】

化合物の機能特性を変えるためにほぼ任意構造を変化させることができる小有機分子合成と異なり、タンパク質合成は20種の天然アミノ酸によりコードされる変化に限られる。細菌からヒトに至る全公知生物の遺伝コードは同一の20種の共通アミノ酸をコードする。これらのアミノ酸はタンパク質の翻訳後修飾（例えばグリコシル化、リン酸化又は酸化）、又は頻度は低い、例えばセレノシステインの場合のようにアミノアシル化サプレッサー tRNA の酵素修飾により改変することができる。それにもかかわらず、これらの僅か20種の単純な構成単位から合成されるポリペプチドが複雑な生命現象の全てを行っている。

20

【0006】

部位特異的及びランダム突然変異誘発はタンパク質の特定アミノ酸を他の19種の共通アミノ酸の任意のもので置換でき、タンパク質の構造と機能の関係を解明するための重要なツールになっている。これらの方法により、安定性、触媒活性及び結合特異性等の特性を強化したタンパク質を作製することが可能になっている。しかし、タンパク質置換は大半が単純な官能基をもつ20種の共通アミノ酸に限られている Knowles, J. R. *Tinkering with enzymes: what are we learning? Science*, 236 (4806) 1252-1258 (1987); 及び Zoller, M. J., Smith, M. *Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, Methods Enzymol*, 154: 468-500 (1983) 参照。新規な生物学的、化学的又は物理的特性をもつ付加アミノ酸を含むように遺伝コードを拡張することにより、例えば天然タンパク質のように20種の共通アミノ酸に含まれるアミノ酸のみから構成されるタンパク質と比較してタンパク質の性質（例えば寸法、酸性度、求核性、水素結合、疎水性）を改変することができる。

30

40

【0007】

非天然アミノ酸をタンパク質に導入するために数種のストラテジーが利用されている。最初の実験は Lys、Cys 及び Tyr 等の反応性側鎖をアミノ酸に付加した（例えばリジン

【数1】

N^ε—

アセチル—リジンに変換)。化学合成も非天然アミノ酸を導入する簡単な方法であるが、日常的固相ペプチド合成は一般に100残基未満の小ペプチド又はタンパク質に限られている。ペプチドフラグメントの酵素ライゲーションと天然化学ライゲーションの最近の開発に伴い、より大きいタンパク質を生産することが可能であるが、方法の大規模化は容易ではない。例えばP. E. DawsonとS. B. H. Kent, Ann. Rev. Biochem., 69:923 (2000) 参照。タンパク質生合成に有用な*in vitro*抽出物に所望非天然アミノ酸で化学的にアシル化したサプレッサートRNAを加える一般的な*in vitro*生合成法を使用して100種を上回る非天然アミノ酸がほぼ任意寸法の各種タンパク質に部位特異的に組込まれている。例えばV. W. Cornish, D. Mendel and 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); 及び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) 参照。タンパク質安定性、タンパク質フォールディング、酵素メカニズム及びシグナル伝達の研究のために多様な官能基がタンパク質に導入されている。これらの研究はタンパク質生合成機構が多様なアミノ酸側鎖を許容することを立証しているが、方法は技術的に困難であり、突然変異タンパク質収率は低い。

【0008】

50年以上前に天然アミノ酸の多数の類似体は細菌の増殖を阻害することが発見された。これらのアミノ酸類似体の存在下で生産したタンパク質を分析すると、各種程度までその天然対応物を置換していることが分かった。例えば、M. H. Richmond, Bacteriol. Rev. 26:398 (1962) 参照。これは正しいアミノ酸とそのコグネイトtRNAの結合に関与する酵素であるアミノアシル—tRNAシンテターゼが類似体に対応する天然アミノ酸から厳密に区別できないためである。例えば、ノルロイシンはメチオニル—tRNAシンテターゼにより負荷され、p-フルオロフェニルアラニンはフェニルアラニン—tRNAシンテターゼにより負荷される。D. B. Cowie, G. N. Cohen, E. T. Bolton and H. DeRobinchon-Szulmajst, Biochim. Biophys. Acta, 1959, 34:39 (1959); 及びR. MunierとG. N. Cohen, Biochim. Biophys. Acta, 1959, 31:378 (1959) 参照。

【0009】

その後、種々の野生型シンテターゼを利用するために選択圧組込みと呼ばれる*in vivo*方法が開発された。例えばN. Budisa, C. Minks, S. Alefelder, W. Wenger, F. M. Dong, L. Moroder and R. Huber, FASEB J., 13:41 (1999) 参照。この方法では特定天然アミノ酸を細胞に供給する該代謝経路を遮断する栄養要求株を使用し、ターゲット遺伝子の転写を抑制しながら制限濃度の天然アミノ酸を含む最少培地で培養する。定常増殖期が開始したら天然アミノ酸の供給を停止して非天然アミノ酸類似体に交換する。組換えタンパク質

の発現を誘導すると、非天然類似体を含むタンパク質が蓄積する。例えば、このストラテジーを使用して、*o*、*m*及び*p*-フルオロフェニルアラニンがタンパク質に組込まれ、紫外線スペクトルで容易に識別可能な2つの特徴的な肩を示しており（例えばC. Minks, R. Huber, L. Moroder and N. Budisa, Anal. Biochem., 284:29 (2000) 参照）、また、トリフルオロメチオンinを使用してバクテリオファージT4リゾチームのメチオンinを置換し、キトオリゴ糖リガンドとのその相互作用が¹⁹F NMRにより試験されており（例えばH. Duewel, E. Darb, V. Robinson and J. F. Honek, Biochemistry, 36:3404 (1997)）、更にロイシンの代わりにトリフルオロロイシンを挿入してロイシンジッパータンパク質の熱及び化学安定性を高めている。例えばY. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado and D. A. Tirrell, Angew. Chem. Int. Ed. Engl., 40:1494 (2001) 参照。更に、セレノメチオンinとテルロメチオンinを各種組換えタンパク質に組込んでX線結晶解析で位相決定を容易にしている。例えばW. A. Hendrickson, J. R. Horton and D. M. Lemaster, EMBO J., 9:1665 (1990)；J. O. Boles, K. Lewinski, M. Kunkle, J. D. Odom, B. Dunlap, L. Lebioda and M. Hatada, Nat. Struct. Biol., 1:283 (1994)；N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann and R. Huber, Eur. J. Biochem., 230:788 (1995)；及びN. Budisa, W. Karnbrock, S. Steinbacher, A. Humm, L. Prade, T. Neufeind, L. Moroder and R. Huber, J. Mol. Biol., 270:616 (1997) 参照。アルケン又はアルキン官能基をもつメチオンin類似体も効率的に挿入され、化学的手段によりタンパク質の付加修飾が可能である。例えばJ. C. M. van HestとD. A. Tirrell, FEBS Lett., 424:68 (1998)；J. C. M. van Hest, K. L. Kiick and D. A. Tirrell, J. Am. Chem. Soc., 122:1282 (2000)；及びK. L. KiickとD. A. Tirrell, Tetrahedron, 56:9487 (2000) 参照。

10

20

30

【0010】

この方法の成功は一般にタンパク質翻訳の忠実度を確保するためには高い選択性を必要とするアミノアシル-tRNAシンテターゼによる非天然アミノ酸類似体の認識に依存している。従って、この方法により接近可能な化学官能基の範囲は限られている。例えば、チアプロリンは定量的にタンパク質に組込むことができるが、オキサプロリンやセレノプロリンはできない。N. Budisa, C. Minks, F. J. Medrano, J. Lutz, R. Huber and L. Moroder, Proc. Natl. Acad. Sci. U S A, 95:455 (1998) 参照。この方法の範囲を拡大する1つの方法はアミノアシル-tRNAシンテターゼの基質特異性を緩和することであり、少数の例で成功している。例えば、大腸菌フェニルアラニル-tRNAシンテターゼ (PheRS) でAla²⁹⁴をGlyに置換すると、基質結合ポケットの寸法が増し、その結果、tRNA^{Phe}は*p*-Cl-フェニルアラニン (*p*-Cl-Phe) によりアシル化される。M. Ibba, P. Kast and H. Hennecke, Biochemistry, 33:7107 (1994) 参照。この突然変異体PheRSを大腸菌株に導入すると、フェニルアラニンの代わりに*p*-Cl-フェニルアラニン又は*p*-Br-フェニルアラニンを組込むことができる。例えばM. IbbaとH. Hennecke, FEBS Lett., 364:272 (1995)；及びN. Sharma, R. Furrer, P. Kast and D. A. Tirrell, FEBS Lett., 467:37 (2000) 参照。同様に、大腸菌チロシル-tRNAシンテターゼのアミノ酸結合部位の近傍の点突然変異Phe130Serの結果、アザチロシンをチロシンよりも効率

40

50

的に組込めることが示されている。F. Hamano-Takaku, T. Iwama, S. Saito-Yano, K. Takaku, Y. Monden, M. Kitabatake, D. Soll and S. Nishimura, J. Biol. Chem., 275:40324 (2000) 参照。

【0011】

アミノアシル化の忠実性は非コグネイト中間体及び生成物の基質識別とブルーフリーディングの両レベルで維持される。従って、非天然アミノ酸をタンパク質に *in vivo* 組込む代替戦略はブルーフリーディングメカニズムをもつシンテターゼの改変である。これらのシンテターゼはコグネイト天然アミノ酸に構造的に類似するアミノ酸を識別することができず、従って活性化することができない。このエラーは別部位で修正され、tRNAから誤って負荷されたアミノ酸を脱アシル化し、タンパク質翻訳の忠実度を維持する。シンテターゼのブルーフリーディング活性が機能しなくなると、誤って活性化された構造類似体が編集機能から抜け落ちて組込まれる恐れがある。このアプローチはバリル-tRNAシンテターゼ (ValRS) で最近立証されている。V. Doring, H. D. Mootz, L. A. Nangle, T. L. Hendrickson, V. de Crecy-Lagard, P. Schimmel and P. Marliere, Science, 292:501 (2001) 参照。ValRSがtRNA ValをCys、Thr又はアミノ酪酸 (Abu) で誤ってアミノアシル化すると、これらの非コグネイトアミノ酸は編集ドメインにより加水分解される。大腸菌染色体のランダム突然変異誘発後に、ValRSの編集部位に突然変異をもつ突然変異体大腸菌株が選択された。この編集欠損ValRSはtRNA Valに誤ってCysを負荷する。AbuはCysに立体的に似ている (AbuではCysの-SH基を-CH₃で置換) ので、この突然変異体大腸菌株をAbuの存在下に増殖させると、突然変異体ValRSはAbuもタンパク質に組込む。質量分析によると、天然タンパク質ではバリンの約24%が各バリン位でAbuにより置換されている。

10

20

30

【0012】

上記方法の少なくとも1つの重大な欠点は、タンパク質中の特定天然アミノ酸に対応する全部位を置換することである。また、天然及び非天然アミノ酸の組込み度も一定せず、細胞の内側でコグネイト天然アミノ酸を完全に枯渇させることは困難であるので、定量的置換が達成されることは稀である。類似体の多重部位組込みは毒性を生じるので、これらの戦略は生きた細胞では突然変異体タンパク質を試験し難いという欠点もある。最後に、ゲノム内の全部位で置換できなければならないという理由から、この方法は一般に共通アミノ酸の近似構造類似体にしか適用できない。

【0013】

固相合成及び半合成法により新規アミノ酸を含む多数の小タンパク質の合成も行われている。例えば以下の刊行物とその引用文献を参照されたい。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., Bohm, H. Studies on polypeptides. XXXVI. The effect of pyrazole-imidazole replacements on the S-protein activating potency of an S-peptide fragments, J. Am. Chem., 5914-5919 (1966); Kaiser, E. T. Synthetic approaches to biologically active peptides and proteins including enzymes, Acc. Chem. Res., 47-54 (1989); Nakatsuka, T., Sasaki, T., Kaiser, E. T. Peptide segment coupling catalyzed by the semisynthetic enzyme thiosubstilis in, J. Am. Chem. Soc., 3808-3810 (1987); Schnol

40

50

zer, M., Kent, S. B. H. Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease, Science, 221-225 (1992); Chaiken, I. M. Semisynthetic peptides and proteins, CRC Crit. Rev. Biochem., 255-301 (1981); Offord, R. E. Protein engineering by chemical means? Protein Eng., 151-157 (1987); 及び 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, 243 (1994)。

10

【0014】

補因子、スピンラベル及びオリゴヌクレオチド等の各種非天然側鎖をタンパク質に *in vitro* 導入するために化学修飾が使用されている。例えば Corey, D. R., Schultz, P. G. Generation of a hybrid sequence-specific single-stranded deoxyribonuclease, Science, 283 (4832): 1401-1403 (1987); Kaiser, E. T., Lawrence D. S., Rokita, S. E. The chemical modification of enzymatic specificity, 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-substitisin, J. Biol. Chem., 243 (24): 6392-6401 (1968); Polgar, L. B., M. L. A new enzyme containing a synthetically formed active site. Thiol-substitisin. J. Am. Chem. Soc., 88: 3153-3154 (1966); 及び Pollack, S. J., Nakayama, G. Schultz, P. G. Introduction of nucleophiles and spectroscopic probes into antibody combining sites, Science, 242 (4881): 1038-1040 (1988) 参照。

20

30

【0015】

あるいは、数種の生物学的プローブを *in vitro* 合成タンパク質に組込むために化学修飾アミノアシル-tRNAを利用する生合成法が使用されている。以下の刊行物とその引用文献参照。Brunner, J. New Photolabeling and crosslinking methods, Annu. Rev. Biochem., 62: 483-514 (1993); 及び Krieg, U. C., Walter, P., Hohnson, A. E. Photocrosslinking of the signal sequence of nascent preprolactin of the 54-kilodalton polypeptide of the signal recognition particle, Proc. Natl. Acad. Sci., 83 (22): 8604-8608 (1986)。

40

【0016】

所望アンバーナンセンス突然変異を含む遺伝子でプログラムされたタンパク質合成反応に化学的にアミノアシル化したサプレッサ-tRNAを加えることにより非天然アミノ酸をタンパク質に部位特異的に *in vitro* 組込むことができることは従来示されている。これらのアプローチを使用すると、特定アミノ酸の栄養要求性株を使用して20種の共通アミノ酸の多数のものを近似構造相同体（例えばフェニルアラニンをフルオロフェニル

50

アラニン)で置換することができる。例えばNoren, C. J., Anthony-Cahill, Griffith, M. C., Schultz, P. G. A general method for site-specific incorporation of unnatural amino acids into proteins, Science, 244:182-188 (1989); M. W. Nowakら, 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 non-natural amino acid into a polypeptide, J. Am Chem Soc, 111:8013-8014 (1989); N. Budisaら, 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 acids site-specifically into proteins. Methods in Enz, 301-336 (1992); 及びMendel, D., Cornish, V. W. & Schultz, P. G. Site-Directed Mutagenesis with an Expanded Genetic Code, Annu. Rev. Biophys. Biomol. Struct. 24, 435-62 (1995) 参照。

10

20

【0017】

例えば、終止コドンUAGを認識するサプレッサートRNAが作製され、非天然アミノ酸で化学的にアミノアシル化された。タンパク質遺伝子の所期部位に終止コドンTAGを導入するために慣用部位特異的突然変異誘発が使用された。例えばSayers, J. R., Schmidt, W. Eckstein, F. 5', 3' Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis, Nucleic Acids Res, 16(3):791-802 (1988) 参照。アシル化サプレッサートRNAと突然変異体遺伝子をin vitro転写／翻訳系で組合せると、UAGコドンに応答して非天然アミノ酸が組込まれ、このアミノ酸を特定位置に含むタンパク質が得られた。[³H]-Pheを使用する実験とα-ヒドロキシ酸を使用する実験によると、所望アミノ酸のみがUAGコドンにより指定される位置に組込まれ、このアミノ酸はタンパク質の他のどの位置にも組込まれないことが判明した。例えばNorenら, 前出; 及びEllman, J. A., Mendel, D., Schultz, P. G. Site-specific incorporation of novel backbone structures into proteins, Science, 197-200 (1992) 参照。

30

40

【0018】

一般に、これらのin vitroアプローチはアミノ酸の部位特異的組込みが困難であることや、アミノ酸が20種の共通アミノ酸の単純な誘導體でなければならないことや、大きいタンパク質又はペプチドフラグメントの合成に伴う固有の問題などの限界がある。

【0019】

非天然アミノ酸をタンパク質に組込むためにマイクロインジェクション法も使用されている。例えば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 and H. A. Lester, Science, 268:439 (1995); 及びD. A. Dougherty, Curr. Opin. Chem. Biol., 4:645 (2000) 参照。in vitro作製した2種のRNA種即ち所期アミノ酸位置にUAG終止コドンをもつターゲットタンパク質をコードするmRNAと、所望非天然アミノ酸でアミノアシル化したアンバーサプレ

50

ッサー tRNA をアフリカツメガエル卵母細胞に同時に注射している。こうすると、卵母細胞の翻訳機構は UAG により指定される位置に非天然アミノ酸を挿入する。この方法により一体的膜タンパク質の *in vivo* 構造－機能研究が可能になったが、一般に *in vitro* 発現系には応用できない。例えば、蛍光アミノ酸をタキキニンニューロキニン-2 受容体に組込んで蛍光共鳴エネルギー移動により距離を測定したり（例えば G. Turcatti, K. Nemeth, M. D. Edgerton, U. Meseth, F. Talabot, M. Peitsch, J. Knowles, H. Vogel and A. Chollet, J. Biol. Chem., 271:19991 (1996) 参照）、ピオチン化アミノ酸を組込んでイオンチャンネルで表面残基を同定したり（例えば J. P. Gallivan, H. A. Lester and D. A. Dougherty, Chem. Biol., 4:739 (1997) 参照）、ケージドチロシン類似体を使用してイオンチャンネルの共焦点変化をリアルタイムでモニターしたり（例えば J. C. Miller, S. K. Silverman, P. M. England, D. A. Dougherty and H. A. Lester, Neuron, 20:619 (1998) 参照）、 α ヒドロキシアミノ酸を使用してイオンチャンネルのバックボーンを改変し、その開閉メカニズムを探索する（例えば P. M. England, Y. Zhang, D. A. Dougherty and H. A. Lester, Cell, 96:89 (1999) ; 及び T. Lu, A. Y. Ting, J. Mainland, L. Y. Jan, P. G. Schultz and J. Yang, Nat. Neurosci., 4:239 (2001) 参照）等の方法が挙げられる。

10

20

【0020】

しかし、マイクロインジェクション法には、例えばサプレッサー tRNA を非天然アミノ酸で化学的に *in vitro* アミノアシル化しなければならないことや、アシル化 tRNA が翻訳中に化学量論的試薬として消費され、再生できないなどの欠点がある。この欠点により、抑圧効率が悪く、タンパク質収率が低いので、電気生理学的測定等の高感度技術で突然変異体タンパク質を定量することが必要である。更に、この方法はマイクロインジェクションが可能な細胞にしか適用できない。

【0021】

非天然アミノ酸をタンパク質に直接 *in vivo* 組込むことができるならば、突然変異体タンパク質の収率が上がり、技術的に容易であり、細胞や場合によっては生きた生物で突然変異体タンパク質を試験することが可能になり、これらの突然変異体タンパク質を治療処置に使用できる等の利点がある。各種寸法、酸性度、求核性、疎水性等の特性をもつ非天然アミノ酸をタンパク質に組込むことができるならば、タンパク質の構造を合理的且つ体系的に操作し、タンパク質機能を探索すると共に新規特性をもつ新規タンパク質又は生物を創製する可能性が著しく広がる。しかし、タンパク質翻訳の忠実度を高くするために必要な tRNA－シンテターゼ相互作用の複雑な性質によりその方法は困難である。

30

【発明の開示】

【発明が解決しようとする課題】

【0022】

従って、細胞の生合成機構をより効率的且つ有効に改変できるように方法を改善することが必要である。以下の開示から自明の通り、本発明はこれらの必要及び他の必要に取組むものである。

40

【課題を解決するための手段】

【0023】

本発明は直交 tRNA－アミノアシル tRNA シンテターゼ対と前記対の個々の成分を含むタンパク質生合成機構で使用される成分の組成物を提供する。非天然アミノ酸を使用することができる直交 tRNA、直交アミノアシル tRNA シンテターゼ及びその対を作製及び選択するための方法も提供する。本発明の組成物は新規直交 tRNA－アミノアシル tRNA シンテターゼ対（例えば *mutRNA* Tyr－*mut* Tyr RS 対、*mutRNA* Leu－*mut* Leu RS 対、*mutRNA* Thr－*mut* Thr RS 対、*mutRNA*

50

A G l u - m u t G l u R S 対等)を含む。新規直交対は非天然アミノ酸をポリペプチドに *in vivo* 組込むために使用することができる。本発明の他の態様は直交対の選択を含む。

本発明の組成物は直交アミノアシル tRNA シンテターゼ (O-RS) を含み、O-RS は場合により *in vivo* において直交 tRNA (O-tRNA) を非天然アミノ酸で優先的にアミノアシル化する。1 態様において、O-RS は配列番号 4 ~ 34 (表 5 参照) 及びその相補的ポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸を含む。別の態様において、O-RS は天然アミノ酸 (例えば 20 種の公知アミノ酸の 1 種) よりも非天然アミノ酸に対して改善又は強化された酵素特性を持ち、例えば K_m が高いか低く、 k_{cat} が高いか低く、 k_{cat} / K_m が高いか低い等である。

10

【0024】

本発明の非天然アミノ酸は多様な物質を含む。例えば、場合により (限定しないが)、O-メチル-L-チロシン、L-3-(2-ナフチル)アラニン、3-メチルフェニルアラニン、O-4-アリル-L-チロシン、4-プロピル-L-チロシン、トリ-O-アセチル-GlcNAc β -セリン、L-Dopa、フッ素化フェニルアラニン、イソプロピル-L-フェニルアラニン、p-アジド-L-フェニルアラニン、p-アシル-L-フェニルアラニン、p-ベンゾイル-L-フェニルアラニン、L-ホスホセリン、ホスホノセリン、ホスホチロシン、p-ヨードフェニルアラニン、p-ブロモフェニルアラニン、p-アミノ-L-フェニルアラニン、及びイソプロピル-L-フェニルアラニン等の分子が挙げられる。更に、他の例としては (限定しないが) チロシニアミノ酸の非天然類似体；グルタミニアミノ酸の非天然類似体；フェニルアラニンアミノ酸の非天然類似体；セリンアミノ酸の非天然類似体；スレオニンアミノ酸の非天然類似体；アルキル、アリアル、アシル、アジド、シアノ、ハロ、ヒドラジン、ヒドラジド、ヒドロキシル、アルケニル、アルキニル、エーテル、チオール、スルホニル、セレノ、エステル、チオ酸、硼酸、ボロン酸、ホスホ、ホスホノ、ホスフィン、複素環、エノン、イミン、アルデヒド、ヒドロキシルアミン、ケトもしくはアミノ置換アミノ酸又はその任意組合せ；光架橋基をもつアミノ酸；スピララベル付きアミノ酸；蛍光アミノ酸：新規官能基をもつアミノ酸；別の分子と共有又は非共有的に相互作用するアミノ酸；金属結合性アミノ酸；金属含有アミノ酸；放射性アミノ酸；フォトケージドアミノ酸；光異性化可能なアミノ酸；ビオチン又はビオチン類似体含有アミノ酸；グリコシル化又は糖鎖修飾アミノ酸；ケト含有アミノ酸；ポリエチレングリコールを含むアミノ酸；ポリエーテルを含むアミノ酸；重原子置換アミノ酸；化学分解性又は光分解性アミノ酸；延長側鎖をもつアミノ酸；毒性基を含むアミノ酸；糖鎖置換アミノ酸 (例えば糖鎖置換セリン等)；炭素結合糖含有アミノ酸；レドックス活性アミノ酸； α -ヒドロキシ含有酸；アミノチオ酸含有アミノ酸； α , α ジ置換アミノ酸； β アミノ酸；及びプロリン以外の環状アミノ酸が挙げられる。

20

30

【0025】

本発明は配列番号 4 ~ 34 (配列については表 5 参照) から選択されるコーディングポリヌクレオチド配列、配列番号 35 ~ 66 から選択されるポリペプチドをコードするコーディングポリヌクレオチド配列、前記ポリヌクレオチド配列の実質的に全長にわたって高ストリンジェント条件下でハイブリダイズするポリヌクレオチド配列、及び前記配列の任意のものの相補的配列から選択されるコーディングポリヌクレオチド配列によりコードされるアミノ酸配列を含むポリペプチドも含む。更に、前記ポリペプチドは場合により直交アミノアシル tRNA シンテターゼ及び／又は配列番号 35 ~ 66 から選択されるアミノ酸配列をコードする。

40

【0026】

本発明は配列番号 1 ~ 3 から選択されるポリヌクレオチド配列 (又はその相補的ポリヌクレオチド配列)、及び前記ポリヌクレオチド配列の実質的に全長にわたって高ストリンジェント条件下でハイブリダイズするポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸も含む。前記核酸はポリヌクレオチド配列が直交 t R

50

NAを含むか及び／又はポリヌクレオチド配列が（場合により配列番号35～66の配列をもつものから選択される）直交アミノアシルtRNAシンテターゼと相補対を形成する態様も含む。

【0027】

本発明は直交tRNA（O-tRNA）の組成物も含み、O-tRNAはセクターコドン認識し、O-tRNAは直交アミノアシルtRNAシンテターゼにより非天然アミノ酸で優先的にアミノアシル化される。1態様では、O-tRNAは配列番号1～3（表5参照）及びその相補的ポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸を含む。

【0028】

本発明のセクターコドンはタンパク質生合成機構の遺伝コドン枠を拡張する。例えば、セクターコドンとしては例えば（天然又は非天然塩基から構成される）ユニーク3塩基コドン、ナンセンスコドン（例えばアンバーコドン又はオパールコドン等の終止コドン）、非天然コドン、レアコドン、少なくとも4塩基を含むコドン、少なくとも5塩基を含むコドン、少なくとも6塩基を含むコドン等が挙げられる。

【0029】

1態様において、O-tRNAは（場合により組成物内に）直交アミノアシルtRNAシンテターゼ（O-RS）を含むことができ、例えばO-tRNAとO-RSは相補的である（例えばO-tRNA/O-RS対）。1態様において、対は例えばmutRNA Tyr-mutTyrRS対（例えばmutRNA Tyr-SS12TyrRS対）、mutRNA Leu-mutLeuRS対、mutRNA Thr-mutThrRS対、mutRNA Glu-mutGluRS対等を含む。別の態様において、対は大腸菌由来mutRNA Gln-mutGlnRS、酵母由来mutRNA Asp-mutAspRS又は酵母由来mutRNA PheCUA-mutPheRS以外のものであり、これらの対は本発明の対の特性をもたない。

【0030】

O-tRNAとO-RSは多様な生物に由来する天然tRNA及びRSの突然変異により誘導することができる。1態様では、O-tRNAとO-RSは少なくとも1種の生物に由来し、生物は例えば*Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*、*Halobacterium*、*Escherichia coli*、*A. fulgidus*、*P. furiosus*、*P. horikoshii*、*A. pernix*、*T. thermophilus*等の原核生物である。場合により、生物は例えば植物（例えば単子葉植物又は双子葉植物等の複雑な植物）、藻類、真菌類（例えば酵母等）、動物（例えば哺乳動物、昆虫、節足動物等）、昆虫、原生生物等の真核生物である。場合により、O-tRNAは第1の生物に由来する天然tRNAの突然変異により誘導され、O-RSは第2の生物に由来する天然RSの突然変異により誘導される。1態様では、O-tRNAとO-RSは突然変異tRNAと突然変異RSから誘導することができる。

【0031】

O-tRNAとO-RSは場合により多様な生物から単離することもできる。1態様では、O-tRNAとO-RSは少なくとも1種の生物から単離され、生物は例えば*Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*、*Halobacterium*、*Escherichia coli*、*A. fulgidus*、*P. furiosus*、*P. horikoshii*、*A. pernix*、*T. thermophilus*等の原核生物である。場合により、生物は例えば植物（例えば単子葉植物又は双子葉植物等の複雑な植物）、藻類、真菌類（例えば酵母等）、動物（例えば哺乳動物、昆虫、節足動物等）、昆虫、原生生物等の真核生物である。場合により、O-tRNAは第1の生物に由来する天然tRNAから単離され、O-RSは第2の生物に由来する天然RSから単離される。1態様では、O-tRNAとO-RSは（場合により（原核生物及び／又は真核生物を含む）1種以上の生

10

20

30

40

50

物に由来する1種以上のO-tRNA及び／又はO-RSを含む) 1種以上のライブラリーから単離することができる。

【0032】

別の側面では、本発明の組成物は細胞に組込むことができる。場合により、本発明の組成物はin vitro翻訳系に組込むことができる。

【0033】

本発明は非天然アミノ酸を組込むために使用することができるタンパク質生合成機構の成分(例えばO-RS、O-tRNA及び直交O-tRNA/O-RS対)の作製方法も提供する。生物のin vivo翻訳系で使用する直交tRNA-tRNAシンテターゼ対の選択方法も提供する。前記方法で使用する非天然アミノ酸とセクターコードンについては上記及び下記に記載する。 10

【0034】

少なくとも1種の組換え直交アミノアシルtRNAシンテターゼ(O-RS)の生産方法は、(a)第1の生物(例えば*Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*、*Halobacterium*、*Escherichia coli*、*A. fulgidus*、*P. furiosus*、*P. horikoshii*、*A. pernix*、*T. thermophilus*等の原核生物)に由来する少なくとも1種のアミノアシルtRNAシンテターゼ(RS)から誘導される(場合により突然変異体)RSのライブラリーを作製する段階と、(b)非天然アミノ酸と天然アミノ酸の存在下に直交tRNA(O-tRNA)をアミノアシル化するメンバーをRS(場合により突然変異体RS)のライブラリーから選択(及び／又はスクリーニング)することにより活性(場合により突然変異体)RSのプールを提供する段階と、及び／又は(c)非天然アミノ酸の不在下にO-tRNAを優先的にアミノアシル化する活性RS(例えば突然変異体RS)のプールを(場合によりネガティブ選択により)選択することにより、O-tRNAを非天然アミノ酸で優先的にアミノアシル化する少なくとも1種の組換えO-RSを提供する段階を含む。前記方法により生産された組換えO-RSも本発明に含まれる。 20

【0035】

1態様では、RSは不活性RSである。不活性RSは活性RSを突然変異させることにより作製することができる。例えば、不活性RSは少なくとも約1、少なくとも約2、少なくとも約3、少なくとも約4、少なくとも約5、少なくとも約6、又は少なくとも約10個以上のアミノ酸を別のアミノ酸(例えばアラニン)に突然変異させることにより作製することができる。 30

【0036】

突然変異体RSのライブラリーは当分野で公知の種々の突然変異誘発技術を使用して作製することができる。例えば、突然変異体RSは部位特異的突然変異、ランダム突然変異、多様性を生じる組換え突然変異、キメラ構築物及び本明細書に記載するか又は当分野で公知の他の方法により作製することができる。

【0037】

1態様では、非天然アミノ酸と天然アミノ酸の存在下に活性なメンバー(例えば直交tRNA(O-tRNA)をアミノアシル化するメンバー)をRS(場合により突然変異体RS)のライブラリーから選択(及び／又はスクリーニング)する段階は、少なくとも1個のセクターコードン(例えばアンバー、オーカー又はオパールコードン)を含むポジティブ選択又はスクリーニングマーカー(例えば抗生物質耐性遺伝子等)と(場合により突然変異体)RSのライブラリーを複数の細胞に導入し、選択物質の存在下に複数の細胞を増殖させ、ポジティブ選択又はスクリーニングマーカー中で少なくとも1個のセクターコードンを抑圧することにより選択及び／又はスクリーニング物質の存在下に生存する(か又は特異的応答を示す)細胞を同定し、活性(場合により突然変異体)RSのプールを含むポジティブ選択済み細胞のサブセットを提供することを含む。場合により、選択及び／又はスクリーニング物質濃度を変動させてもよい。 40 50

【0038】

1 側面では、ポジティブ選択マーカーはクロラムフェニコールアセチルトランスフェラーゼ（CAT）遺伝子であり、少なくとも1個のセクターコドンはCAT遺伝子内のアンバー終止コドンである。場合により、ポジティブ選択マーカーはβラクタマーゼ遺伝子であり、少なくとも1個のセクターコドンはβラクタマーゼ遺伝子内のアンバー終止コドンである。別の側面では、ポジティブスクリーニングマーカーは蛍光もしくは発光スクリーニングマーカー又は親和性に基づくスクリーニングマーカー（例えば細胞表面マーカー）を含む。

【0039】

1 態様では、非天然アミノ酸の不在下にO⁻tRNAを優先的にアミノアシル化する活性RS（場合により突然変異体）のプールをネガティブ選択又はスクリーニングする段階は、少なくとも1個のセクターコドンを含むネガティブ選択又はスクリーニングマーカー（例えば抗生物質耐性遺伝子、例えばクロラムフェニコールアセチルトランスフェラーゼ（CAT）遺伝子）をポジティブ選択又はスクリーニングからの活性（場合により突然変異体）RSのプールと共に第2の生物の複数の細胞に導入し、非天然アミノ酸と選択又はスクリーニング物質を加えた第1の培地で生存するか又は特異的スクリーニング応答を示すが、非天然アミノ酸と選択又はスクリーニング物質を加えない第2の培地では生存するか又は特異的応答を示すことができない細胞を同定することにより、少なくとも1種の組換えO⁻RSをもつ生存細胞又はスクリーニング済み細胞を提供することを含む。例えば、ポジティブ選択及び／又はネガティブスクリーニングとして場合によりCAT同定プロトコールを使用して適当な組換えO⁻RSを決定する。例えば、場合により1種以上の非天然アミノ酸の存在下又は不在下に（少なくとも1個のセクターコドンを含む）CATを加えた増殖プレートでクローンのプールを複製する。こうして、非天然アミノ酸を加えたプレートでしか増殖しないコロニーが組換えO⁻RSを含むとみなす。1 側面では、選択（及び／又はスクリーニング）物質濃度を変動させる。所定側面では、第1の生物と第2の生物は異なる。即ち、第1及び／又は第2の生物は場合により原核生物、真核生物、哺乳動物、大腸菌、真菌、酵母、始原菌、真正細菌、植物、昆虫、原生生物等を含む。別の態様では、スクリーニングマーカーは蛍光もしくは発光スクリーニングマーカー又は親和性に基づくスクリーニングマーカーを含む。

10

20

【0040】

別の態様では、活性（場合により突然変異体）RSのプールをスクリーニング又は選択（例えばネガティブ選択）する段階は、ポジティブ選択段階（b）からの活性突然変異体RSのプールを単離し、少なくとも1個のセクターコドンを含むネガティブ選択又はスクリーニングマーカー（例えば少なくとも1個のセクターコドンを含む毒性マーカー遺伝子、例えばリボヌクレアーゼバルナーゼ遺伝子）と活性（場合により突然変異体）RSのプールを第2の生物の複数の細胞に導入し、非天然アミノ酸を加えない第1の培地で生存するか又は特異的スクリーニング応答を示すが、非天然アミノ酸を加えた第2の培地では生存するか又は特異的スクリーニング応答を示すことができない細胞を同定し、非天然アミノ酸に特異的な少なくとも1種の組換えO⁻RSをもつ生存又はスクリーニング済み細胞を提供することを含む。1 側面では、少なくとも1個のセクターコドンは約2個以上のセクターコドンを含む。このような態様は場合により、少なくとも1個のセクターコドンが2個以上のセクターコドンを含み、第1の生物と第2の生物が異なる（例えば各生物が場合により例えば原核生物、真核生物、哺乳動物、大腸菌、真菌、酵母、始原菌、真正細菌、植物、昆虫、原生生物等である）態様を含むことができる。更に、所定側面では、ネガティブ選択マーカーは（少なくとも1個のセクターコドンを含む）リボヌクレアーゼバルナーゼ遺伝子を含む。他の側面では、スクリーニングマーカーは場合により蛍光もしくは発光スクリーニングマーカー又は親和性に基づくスクリーニングマーカーを含む。本明細書に記載する態様では、スクリーニング及び／又は選択は場合によりスクリーニング及び／又は選択ストリンジェンシーの変動を含む。

30

40

【0041】

50

1 態様では、少なくとも 1 種の組換え直交アミノアシル tRNA シンテターゼ (O-RS) の生産方法は更に、(d) 少なくとも 1 種の組換え O-RS を単離する段階と、(e) 少なくとも 1 種の組換え O-RS から誘導される第 2 組の (場合により突然変異) O-RS を作製する段階と、(f) O-tRNA を優先的にアミノアシル化することが可能な突然変異 O-RS が得られるまで段階 (b) 及び (c) を反復する段階を含む。場合により、段階 (d) ~ (f) を例えば少なくとも約 2 回反復する。1 側面では、少なくとも 1 種の組換え O-RS から誘導される第 2 組の突然変異 O-RS は突然変異誘発 (例えばランダム突然変異誘発、部位特異的突然変異誘発、組換え又はその任意組合せ) により作製することができる。

【0042】

上記方法における選択/スクリーニング段階 (例えばポジティブ選択/スクリーニング段階 (b)、ネガティブ選択/スクリーニング段階 (c) 又はポジティブ及びネガティブ選択/スクリーニング段階 (b) 及び (c) の両方) は、場合により選択/スクリーニングストリンジェンシーの変動を含む。別の態様では、ポジティブ選択/スクリーニング段階 (b)、ネガティブ選択/スクリーニング段階 (c) 又はポジティブ及びネガティブ選択/スクリーニング段階 (b) 及び (c) の両方はレポーターの使用を含み、レポーターは蛍光活性化細胞ソーティング (FACS) により検出されるか又はレポーターは発光により検出される。場合により、レポーターは細胞表面、ファージディスプレイ等に提示され、非天然アミノ酸又は類似体に対する親和性又は触媒活性に基づいて選択される。1 態様では、突然変異シンテターゼは細胞表面、ファージディスプレイ等に提示される。

【0043】

本明細書に記載する方法は場合により非天然アミノ酸が O-メチル-L-チロシン、L-3-(2-ナフチル)アラニン、3-メチルフェニルアラニン、O-4-アリル-L-チロシン、4-プロピル-L-チロシン、トリ-O-アセチル-GlcNAc β -セリン、L-Dopa、フッ素化フェニルアラニン、イソプロピル-L-フェニルアラニン、p-アジド-L-フェニルアラニン、p-アシル-L-フェニルアラニン、p-ベンゾイル-L-フェニルアラニン、L-ホスホセリン、ホスホノセリン、ホスホノチロシン、p-ヨードフェニルアラニン、p-ブロモフェニルアラニン、p-アミノ-L-フェニルアラニン、及びイソプロピル-L-フェニルアラニンから選択される態様を含む。本明細書に記載する方法により生産された組換え O-RS も本発明に含まれる。

【0044】

組換え直交 tRNA (O-tRNA) の生産方法は、(a) 第 1 の生物に由来する少なくとも 1 種の tRNA (例えばサプレッサー tRNA) から誘導される突然変異体 tRNA のライブラリーを作製する段階と、(b) 第 1 の生物に由来する RS の不在下に第 2 の生物に由来するアミノアシル tRNA シンテターゼ (RS) によりアミノアシル化される (場合により突然変異体) tRNA のライブラリーを選択 (例えばネガティブ選択) 又はスクリーニングすることにより tRNA (場合により突然変異体) のプールを提供する段階と、(c) 導入した直交 RS (O-RS) によりアミノアシル化されるメンバーを tRNA (場合により突然変異体) のプールから選択又はスクリーニングすることにより、少なくとも 1 個のセクターコドンを認識し、第 2 の生物に由来する RS により効率的に認識されず、O-RS により優先的にアミノアシル化される少なくとも 1 種の組換え O-tRNA を提供する段階を含む。所定態様では、少なくとも 1 種の tRNA はサプレッサー tRNA であり、及び/又は天然もしくは非天然塩基のユニーク 3 塩基コドン、ナンセンスコドン、レアコドン、非天然コドン、少なくとも 4 塩基を含むコドン、アンバーコドン、オーカーコドン又はオパール終止コドンを含む。1 態様では、組換え O-tRNA は改善された直交性をもつ。当然のことながら、所定態様では O-tRNA は場合により改変の必要なしに第 1 の生物から第 2 の生物に導入される。各種態様において、第 1 の生物と第 2 の生物は同一又は異なり、場合により原核生物 (例えば *Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*、*Escherichia coli*、*Halobacterium* 等)

10

20

30

40

50

、真核生物、哺乳動物、真菌、酵母、始原菌、真正細菌、植物、昆虫、原生生物等から選択される。更に、組換え tRNA は場合により非天然アミノ酸によりアミノアシル化され、非天然アミノ酸は天然又は遺伝子操作により *in vivo* 生合成される。非天然アミノ酸は場合により少なくとも第 1 又は第 2 の生物の増殖培地に添加する。

【0045】

1 側面では、アミノアシル tRNA シンテターゼによりアミノアシル化される（場合により突然変異体）tRNA のライブラリーを選択（例えばネガティブ選択）又はスクリーニングする段階（段階（b））は、少なくとも 1 個のセクターコドンを含む毒性マーカー遺伝子（又は毒性もしくは静的物質の生産を誘導する遺伝子又は生物に必須の遺伝子であって、少なくとも 1 個のセクターコドンを含む遺伝子）と（場合により突然変異体）tRNA のライブラリーを第 2 の生物に由来する複数の細胞に導入し、少なくとも 1 種の直交 tRNA 又は非機能的 tRNA を含む（場合により突然変異体）tRNA のプールを含む生存細胞を選択することを含む。例えば、生存細胞は比較細胞密度アッセイを使用することにより選択することができる。

10

【0046】

別の側面では、毒性マーカー遺伝子は 2 個以上のセクターコドンを含むことができる。前記方法の別の態様では、毒性マーカー遺伝子はリボスクレアーゼバルナーゼ遺伝子であり、リボスクレアーゼバルナーゼ遺伝子は少なくとも 1 種のアンバーコドンを含む。場合により、リボスクレアーゼバルナーゼ遺伝子は 2 種以上のアンバーコドンを含むことができる。

20

【0047】

1 態様では、導入した直交 RS（O-RS）によりアミノアシル化されるメンバーを（場合により突然変異体）tRNA のプールから選択又はスクリーニングする段階は、薬剤耐性遺伝子（例えば少なくとも 1 個のセクターコドン（例えば少なくとも 1 個のアンバーコドン）を含むβラクターゼ遺伝子）又は生物に必須の遺伝子又は毒性物質の解毒を誘導する遺伝子を含むポジティブ選択又はスクリーニングマーカー遺伝子と、O-RS と、（場合により突然変異体）tRNA のプールを第 2 の生物に由来する複数の細胞に導入し、選択又はスクリーニング物質（例えば抗生物質）の存在下に増殖した生存又はスクリーニング済み細胞を同定することにより、O-RS によりアミノアシル化され、少なくとも 1 個のセクターコドンに応答してポジティブマーカー遺伝子によりコードされる翻訳産物にアミノ酸を挿入する少なくとも 1 種の組換え tRNA をもつ細胞プールを提供することを含む。別の態様では、選択及び／又はスクリーニング物質濃度を変動させる。本発明の方法により生産された組換え O-tRNA も本発明に含まれる。

30

【0048】

特異的 O-tRNA/O-RS 対の生産方法も提供する。本方法は、（a）第 1 の生物に由来する少なくとも 1 種の tRNA から誘導される tRNA のライブラリーを作製する段階と、（b）第 1 の生物に由来する RS の不在下で第 2 の生物に由来するアミノアシル tRNA シンテターゼ（RS）によりアミノアシル化される（場合により突然変異体）tRNA のライブラリーをネガティブ選択又はスクリーニングすることにより、（場合により突然変異体）tRNA のプールを提供する段階と、（c）導入した直交 RS（O-RS）によりアミノアシル化されるメンバーを（場合により突然変異体）tRNA のプールから選択又はスクリーニングすることにより、少なくとも 1 種の組換え O-tRNA を提供する段階を含む。少なくとも 1 種の組換え O-tRNA はセクターコドンを認識し、第 2 の生物に由来する RS により効率的に認識されず、O-RS により優先的にアミノアシル化される。本方法は更に、（d）第 3 の生物に由来する少なくとも 1 種のアミノアシル tRNA シンテターゼ（RS）から誘導される（場合により突然変異体）RS のライブラリーを作製する段階と、（e）非天然アミノ酸と天然アミノ酸の存在下に少なくとも 1 種の組換え O-tRNA を優先的にアミノアシル化するメンバーを突然変異体 RS のライブラリーから選択又はスクリーニングすることにより、活性（場合により突然変異体）RS のプールを提供する段階と、（f）非天然アミノ酸の不在下に少なくとも 1 種の組換え O-

40

50

tRNAを優先的にアミノアシル化する活性（場合により突然変異体）RSのプールをネガティブ選択又はスクリーニングすることにより、非天然アミノ酸に特異的な少なくとも1種の組換えO-RSと少なくとも1種の組換えO-tRNAを含む少なくとも1種の特異的O-tRNA/O-RS対を提供する段階を含む。前記方法により生産された特異的O-tRNA/O-RS対も本発明に含まれる。例えば、特異的O-tRNA/O-RS対としては、例えばmutRNA^{Tyr}-mut^{Tyr}RS対（例えばmutRNA^{Tyr}-SS12^{Tyr}RS対）、mutRNA^{Leu}-mut^{Leu}RS対、mutRNA^{Thr}-mut^{Thr}RS対、mutRNA^{Glu}-mut^{Glu}RS対等を挙げることができる。更に、前記方法は第1の生物と第3の生物が同一である（例えば*Methanococcus jannaschii*）場合を含む。

10

【0049】

第2の生物のin vivo翻訳系で使用するための直交tRNA-tRNAシンテターゼ対の選択方法も本発明に含まれる。本方法はマーカー遺伝子と、tRNAと、第1の生物から単離又は誘導されるアミノアシルtRNAシンテターゼ（RS）を第2の生物に由来する第1組の細胞に導入する段階と、マーカー遺伝子とtRNAを第2の生物に由来する複製組の細胞に導入する段階と、第1組と複製組の細胞を選択又はスクリーニング物質の存在下に増殖させ、第1組では生存するが、複製組では生存することができない細胞を選択するか、又は第1組では特異的スクリーニング応答を示すが、複製組では特異的スクリーニング応答を示すことができない細胞を選択し、第2の生物のin vivo翻訳系で使用するための直交tRNA-tRNAシンテターゼ対を含む生存又はスクリーニング済細胞を提供する段階を含む。1態様では、比較と選択又はスクリーニングはin vivo相補アッセイを含む。選択又はスクリーニング物質濃度を変動させてもよい。

20

【0050】

本発明の生物は多様な生物と多様な組合せを含む。例えば、本発明の方法の第1の生物と第2の生物は同一でも異なってもよい。1態様では第1の生物は場合により原核生物（例えば*Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*、*Halobacterium*、*Escherichia coli*、*A. fulgidus*、*P. furiosus*、*P. horikoshii*、*A. pernix*、*T. thermophilus*等）である。あるいは、第1の生物は場合により例えば植物（例えば単子葉植物又は双子葉植物等の複雑な植物）、藻類、原生生物、真菌類（例えば酵母等）、動物（例えば哺乳動物、昆虫、節足動物等）等の真核生物でもよい。別の態様では、第2の生物は*Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*、*Halobacterium*、*Escherichia coli*、*A. fulgidus*、*Halobacterium*、*P. furiosus*、*P. horikoshii*、*A. pernix*、*T. thermophilus*等の原核生物である。あるいは、第2の生物は例えば酵母、動物細胞、植物細胞、真菌、哺乳動物等の真核生物でもよい。各種態様において、第1の生物と第2の生物は異なる。

30

【0051】

本発明の（上記）各種方法は場合により選択又はスクリーニングが1種以上のポジティブ又はネガティブ選択又はスクリーニング（例えばアミノ酸透過性の変化、翻訳効率の変化及び翻訳忠実度の変化）を含む態様を含む。更に、1種以上の変化は場合によりタンパク質を生産するために直交tRNA-tRNAシンテターゼ対を使用する生物における1個以上の遺伝子の突然変異に基づく。本発明の選択及び／又はスクリーニングは場合により1種以上の選択遺伝子内又は1種以上のスクリーニング遺伝子内で少なくとも2個のセクターコドンを使用する態様を含む。このような多重セクターコドンは場合により同一遺伝子内又は異なるスクリーニング／選択遺伝子内に含まれる。更に、場合により使用する多重セクターコドンは場合により異なるセクターコドンであるか、又は同一型のセクターコドンを含む。

40

【0052】

50

本発明の付加特徴はキットである。例えば、キットは1種以上の上記翻訳系（例えば細胞）と、1種以上の非天然アミノ酸と、例えば適当なパッケージング材料と、キットの成分を保持するための容器と、本発明の方法を実施するための説明書及び／又は他の成分を含むことができる。同様に、例えばキットの成分を保持するための容器と、本発明の方法を実施するための説明書及び／又は他の成分を含むキットの形態で翻訳系の生産物（例えば非天然アミノ酸を含むEPO類似体等のタンパク質）を提供することもできる。

【0053】

図1は、非天然アミノ酸の部位特異的 *in vivo* タンパク質組込みを模式的に示す。直交アミノアシル *tRNA* シンテターゼは直交 *tRNA* を非天然アミノ酸でアミノアシル化する。アシル化された直交 *tRNA* は所期タンパク質をコードする遺伝子に導入されたセクターコドン（例えばユニークコドン）により指定される位置に非天然アミノ酸を挿入する。

10

【0054】

図2A及びBは、非天然アミノ酸でアミノアシル化する活性シンテターゼの選択方法の例を模式的に示す。図2Aは非天然アミノ酸特異性をもつアミノアシル *tRNA* シンテターゼの選択／スクリーニングの概要を示す。ポジティブ選択では、天然又は非天然アミノ酸特異性をもつ活性シンテターゼを同定し、ネガティブ選択では、天然アミノ酸特異性をもつシンテターゼを排除する。直交 *tRNA* に非天然アミノ酸を負荷するシンテターゼのみが選択／スクリーニング後に生存することができる。図2BはO-*tRNA* を非天然アミノ酸で優先的にアミノアシル化するシンテターゼの選択／スクリーニングの1態様を模式的に示す。例えば、直交サプレッサー *tRNA* を含むベクターと突然変異RSのライブラリーのメンバーとセクターコドン（例えばアンバーコドン）をもつポジティブ選択マーカー（例えばβラクタマーゼ）を含む発現ベクターを生物に導入し、選択物質（例えばアンピシリン）の存在下に増殖させる。ポジティブ選択マーカーが発現すると、細胞は選択物質中で生存することができる。生存細胞は天然又は非天然アミノ酸（aa）をO-*tRNA* に負荷することが可能なシンテターゼをコードする。発現ベクターの第2の株と、1個以上のセクターコドン（例えばTAG）をもち、発現すると細胞を死滅させるネガティブ選択マーカー（例えばバルナーゼ等の毒性遺伝子）を含む発現ベクターに活性シンテターゼを形質転換する。細胞を非天然アミノ酸の不在下に増殖させる。所与シンテターゼがO-*tRNA* を天然アミノ酸でアミノアシル化するならば、ネガティブ選択マーカーが発現され、細胞は死滅する。シンテターゼがO-*tRNA* を優先的にアミノアシル化するならば、非天然アミノ酸は存在せず、細胞は生存するのでネガティブ選択マーカーは発現されない。こうして、所望非天然コドンでO-*tRNA* を優先的にアミノアシル化する少なくとも1種の直交シンテターゼが得られる。

20

30

【0055】

図3は、チロシン類似体の特異的ライブラリーを作製するための部位特異的突然変異を示す。

【0056】

図4は、前記類似体の特異的ライブラリーを作製するためのペンタフルオロフェニルアラニン選択用コンセンサス配列を示す。

40

【0057】

図5は、1種の生物（例えば大腸菌）に由来する1個のドメイン（例えばCPIドメイン）を他の生物のシンテターゼ（例えば*Methanococcus jannaschii* TyrRS）に移植する方法を模式的に示す。

【0058】

図6は、キメラ*Methanococcus jannaschii*／大腸菌シンテターゼの構築を模式的に示す。

【0059】

図7は、キメラシンテターゼ（例えば*Methanococcus jannaschii*／大腸菌シンテターゼ）のライブラリーの作製を模式的に示す。

50

【0060】

図8は、内在シンテターゼ（例えば大腸菌シンテターゼ）の弱い基質であり、効率的に所期コグネイトシンテターゼにより負荷されるサプレッサーtRNAの選択の1例を模式的に示す。突然変異tRNAライブラリーのメンバーを含む発現ベクターと、1個以上のセクターコドンをもつネガティブ選択マーカー（例えばバルナーゼ等の毒性遺伝子）を含む別のベクターを生物の細胞に導入する。ネガティブ選択の生存細胞は生物に直交性であるか又は非機能的な突然変異tRNAをコードする。生存細胞からのベクターを単離し、セクターコドンを含むポジティブ選択マーカー（例えばβラクタマーゼ遺伝子）と共に他の細胞に形質転換する。選択物質（例えばアンピシリン）とtRNAと同一起源の生物（例えば*Methanococcus jannaschii*）に由来するRSの存在下に細胞を増殖させる。この選択の生存細胞は細胞のシンテターゼ（例えば大腸菌シンテターゼ）に直交性であり、tRNAと同一起源に由来するRSによりアミノアシル化される突然変異体tRNAをコードする。

【0061】

図9A及びBは突然変異アンチコドンループtRNAライブラリーを模式的に示す。図9Aは突然変異全ループライブラリーを示し、図9Bは*Methanococcus jannaschii* tRNA_{TyrCUA}のライブラリーを示す。ランダム突然変異ヌクレオチド（N）を白抜きで示す。

【0062】

図10は、水素結合以外の力により対合する非天然塩基対（PICS:PICS, 3MN:3MN, 7AI:7AI, Dipic:Py）の構造の例を模式的に示す。

【0063】

図11は、サプレッサーtRNAのネガティブ選択法の結果のグラフであり、所定時間における3種の構築物のうちの1種を含む生存細胞の百分率を示し、ベクター（例えばプラスミドpSCB2）により導入したバルナーゼ遺伝子中の2個のアンバーの抑圧に基づいて決定した。このプラスミドは2個のアンバーコドンを含むバルナーゼ遺伝子をコードする。GMM液体培地で選択を行い、20mMアラビノースを使用してバルナーゼ発現を誘導する。3種の構築物は以下の通りである。（1）○はサプレッサーtRNAを含まない対照プラスミドを示す。（2）△はプラスミドpAC-YYG1上のサプレッサーtRNAを示す。（3）□はプラスミドpAC-JY上のサプレッサーtRNAを示す。

【0064】

図12は、βラクタマーゼ遺伝子中のアンバーコドンの抑圧に基づくポジティブ選択を示す増殖ヒストグラム。サプレッサーtRNAをコードするベクター（例えばpACプラスミド）とシンテターゼをコードするベクター（例えばpBLAM-JYRS）を生物（例えば大腸菌DH10B細胞）に同時形質転換する。各種濃度（0、100及び500μg/ml）のアンピシリンを加えた液体2×YT培地でシンテターゼと各種pACプラスミドを含む細胞を増殖させた結果をAに示し、図中、pACはサプレッサーtRNAを含まない対照プラスミドであり、pAC-YYG1はサプレッサーtRNAを含むプラスミドであり、pAC-JYはサプレッサーtRNAを含むプラスミドである。Bは500μg/mlアンピシリンを加えた2×YT寒天プレートを使用した同一構築物のポジティブ選択を示す。3種の構築物は以下の通りである。（1）○はサプレッサーtRNAを含まない対照プラスミドを示す。（2）△はプラスミドpAC-YYG1上のサプレッサーtRNAを示す。（3）□はプラスミドpAC-JY上のサプレッサーtRNAを示す。

【0065】

図13は、アンチコドンループ及び全ループライブラリーから選択した突然変異体サプレッサーtRNAのDNA配列を示す。JYは野生型*Methanococcus jannaschii* tRNA_{CUA}TyrCUAを表す。

【0066】

図14は、TyrRSの活性部位の立体図を模式的に示す。B. stearoothermophilus TyrRSに由来する残基を示す。*Methanococcus ja*

n n a s c h i i *T y r R S*からの対応する残基は *T y r*³² (*T y r*³⁴)、*G l u*¹⁰⁷ (*A s n*¹²³)、*A s p*¹⁵⁸ (*A s p*¹⁷⁶)、*I l e*¹⁵⁹ (*P h e*¹⁷⁷) 及び *L e u*¹⁶² (*L e u*¹⁸⁰) である (括弧内は *B. s t e a r o t h e r m o p h i l u s* *T y r R S* 残基を表す)。

【0067】

図15は、*T y r R S*の活性部位を模式的に示す。*B. s t e a r o t h e r m o p h i l u s* *T y r R S*に由来する残基も示す。*M e t h a n o c o c c u s* *j a n n a s c h i i* *T y r R S*からの対応する残基は *T y r*³² (*T y r*³⁴)、*A s p*¹⁵⁸ (*A s p*¹⁷⁶)、*I l e*¹⁵⁹ (*P h e*¹⁷⁷)、*L e u*¹⁶² (*L e u*¹⁸⁰) 及び *A l a*¹⁶⁷ (*G l n*¹⁸⁹) である (括弧内は *B. s t e a r o t h e r m o p h i l u s* *T y r R S* 残基を表す)。

【0068】

図16A及びBは、本発明の成分 (例えば直交シンテターゼ) を作製するために使用した *F A C S* 選択及びスクリーニング法の1例を模式的に示す。図16Aは直交シンテターゼライブラリーと *O - t R N A* の発現ベクター (ライブラリープラスミド) と、1個以上のセクターコドン (例えば *T A G*) を含む *T 7 R N A* ポリメラーゼ/*G F P* レポーター系 (レポータープラスミド) のベクター (例えばプラスミド) を模式的に示す。図16Bはポジティブ選択/ネガティブスクリーニングスキームを模式的に示し、非天然アミノ酸の存在下と不在下及び選択物質の存在下と不在下で細胞を増殖させ、スクリーニング過程で蛍光細胞は非蛍光細胞をスクリーニングし、「+」と○が夫々蛍光細胞と非蛍光細胞に対応する。

【0069】

図17A、B、C及びDは、増幅性蛍光レポーター系を示す。図17Aはスクリーニングに使用することができるベクター (例えば *p R E P* 等のプラスミド) を模式的に示し、*T 7 R N A* ポリメラーゼ転写は *a r a* プロモーターにより制御され、タンパク質発現は遺伝子内の各種位置のアンバーコードンの抑圧に依存する。レポーター発現 (例えば *G F P u v* 発現) は *T 7 R N A* ポリメラーゼにより制御される。レポーターベクター (例えばプラスミド *p R E P*) は直交シンテターゼ/*t R N A* 対の発現用ベクター (例えば *C o l E 1* プラスミド) との併用に適合可能である。図17Bは *p R E P* (1~12) 内の *T 7 R N A* ポリメラーゼ遺伝子構築物の組成と蛍光増加を示す。構築物番号を各構築物の左側に示す。蛍光増加は各構築物の右側に示し、*p R E P* (1~12) と *p Q* 又は *p Q D* を含む細胞の蛍光をフルオロメトリーにより測定し、細胞濃度補正比として計算した。遺伝子内のアンバー突然変異の位置を示す。図17Cは *p R E P* (10) と *p Q D* (上段) 又は *p Q* (下段) を含む細胞のサイトメトリー分析を示す。図17Dは *p R E P* (10) を含む、各種大腸菌サプレッサー *t R N A* を発現する細胞のフルオロメトリー分析を示す。「なし」は細胞がサプレッサー *t R N A* を含まないことを示す。

【0070】

図18は、表面エピトープへの非天然アミノ酸の組込みのファージ選択を模式的に示す。例えば、突然変異体シンテターゼライブラリーを導入した大腸菌に、表面タンパク質をコードする遺伝子に終止コドンをもつファージを感染させる。活性シンテターゼを含むファージはファージ表面に非天然アミノ酸を提示し、固定化モノクローナル抗体により選択される。

【0071】

図19は、非天然アミノ酸特異性をもつ提示されたシンテターゼ (例えばファージに提示されたシンテターゼ) をスクリーニングするために使用する分子 (例えばアミノアシルアデニル酸中間体の固定化アミノアルキルアデニル酸類似体) の1例を模式的に示す。

【0072】

図20は、各種生物に由来する各種直交対のアンピシリン耐性を示すグラフ。本図はセクターコドン (例えばアンバーコドン) を含むレポーター遺伝子 (例えば β ラクターゼ遺伝子) と (セクターアンチコドンを含む) サプレッサー *t R N A* を各々含むレポーター

一構築物を使用した直交対の検出例を示し、サプレッサー tRNA は各種生物（例えば *A. fulgidus*、*Halobacterium* NRC-1、*P. furiosus*、*P. horikoshii* 及び *Methanococcus jannaschii*）に由来することができる。レポーター構築物と各種生物（例えば *M. thermoautotrophicum*、*Methanococcus jannaschii*、*P. horikoshii*、*A. pernix*、*A. fulgidus*、*Halobacterium* NRC-1 及び *Escherichia coli*）からクローニングしたシンテターゼを細胞に形質転換する。各種濃度の選択物質（例えばアンピシリン）中で細胞を増殖させる。例えば *in vivo* 相補アッセイを使用して直交 tRNA / RS 対をもつ細胞を選択する。本図から明らかなように、2 種の系が大腸菌シンテターゼで観察されるよりも有意に高い抑圧レベルを示した。それらは *M. thermoautotrophicum* と *Methanococcus jannaschii* である。

10

【0073】

図 21A 及び B は、ロイシル tRNA のアンチコドンループを突然変異（例えばランダム突然変異）させ、例えば選択段階の組合せ（例えば β ラクタマーゼに基づく選択とバルナーゼに基づく選択）を使用してセクターコドン（例えばレポーター遺伝子中のアンバーコドン）のより効率的な抑圧について選択（図 21B）することにより作製した *Halobacterium* NRC-1 由来突然変異アンバーサプレッサー tRNA を示す。図 21B は元のアンバー突然変異体、最適化アンチコドンループ及び最適化アクセプターシステムの 3 種の突然変異体 tRNA 構築物を単独又は RS（例えば Mt LRS）と併用した場合の β ラクタマーゼアンバー抑圧系の IC₅₀ 値（ $\mu\text{g}/\text{ml}$ アンピシリン）を示す。最適化アンチコドンと最適化アクセプターシステムは β ラクタマーゼ選択段階で最高値を示した。

20

【0074】

図 22 は、基本コドンの tRNA サプレッサーを示す。本図に示す tRNA サプレッサーは *Halobacterium* NRC-1 TTG tRNA に由来するライブラリーから単離し、アンチコドンループの 8 ヌクレオチドランダムを突然変異させ、A184 部位に AGGA コドンをもつ β ラクタマーゼ遺伝子を含むレポーター構築物を用いてアンピシリン選択した。

【0075】

図 23A～D は、成功した各進化実験からの優性シンテターゼ変異体の活性を示す。図 23A は pREP / YC-JYC UA と指定シンテターゼ変異体を含む細胞を対応する非天然アミノ酸の存在下（+）又は不在下（-）で増殖させ、長波長紫外線を照射した場合の写真である。図 23B は pREP / YC-JYC UA と指定シンテターゼ変異体を含む細胞を対応する非天然アミノ酸の存在下（左）又は不在下（右）で増殖させた場合のフルオロメトリー分析を示す。図 23C は pREP / YC-JYC UA と指定シンテターゼ変異体を含む細胞を対応する非天然アミノ酸の存在下又は不在下で増殖させた場合の Cm IC₅₀ 分析を示す表である。図 23D は pBAD / JYAMB-4 TAG と指定シンテターゼ変異体を含む細胞を対応する非天然アミノ酸の存在下（+）又は不在下（-）で増殖させた場合のタンパク質発現分析を示す。

30

40

【0076】

図 24 は、ネガティブ FACS スクリーニング（OAY-RS（1, 3, 5））又はネガティブバルナーゼ選択（OAY-RS（B））を使用して得られた OAS-RS 変異体の活性比較を示す。pREP / YC-JYC UA と指定シンテターゼ変異体を含む細胞を対応する非天然アミノ酸の存在下（立体棒、左）又は不在下（立体棒、右）で増殖させ、フルオロメトリー分析した。非天然アミノ酸の不在下に対して存在下で増殖させた細胞の蛍光の細胞濃度補正比として蛍光増加（背後棒）を計算した。

【0077】

図 25A～B は、*M. jannaschii* Tyr RS の進化を誘導するための多目的レポータープラスミド系の成分を示す。図 25A はプラスミド pREP / YC-JYC U

50

Aを示す。プラスミド pREP/YC-JYCUAはプラスミド pBK及び変異体との併用に適合可能である。図25BはM. jannaschii TyrRSの進化のターゲットとして使用した非天然アミノ酸の構造を示す。

【0078】

図26は、プラスミド pREP/YC-JYCUAを使用してアミノアシル tRNA シンテターゼを進化させるためのストラテジーを示す。蛍光及び非蛍光細胞を夫々黒とグレーで示す。

【0079】

図27は、Thermus thermophilusに由来するスレオニル tRNA シンテターゼを示す。

10

【0080】

図28は、T. thermophilus 直交スレオニル tRNA/R Sのための直交 tRNA の作製を示す。

【0081】

図29は、本発明で利用される典型的非天然アミノ酸を示す。

【0082】

図30は、本発明で利用される典型的非天然アミノ酸を示す。

【0083】

図31は、本発明で利用される典型的非天然アミノ酸を示す。

【発明を実施するための最良の形態】

20

【0084】

序文

【0085】

タンパク質は光合成や視覚からシグナル伝達や免疫応答までのほぼ全生物プロセスの岐路に位置する。これらの複雑な機能は規定一次配列に配置された比較的単純な20個の構成単位から構成されるポリアミド系ポリマーから得られる。

【0086】

本発明は非天然アミノ酸をタンパク質に部位特異的に直接 in vivo 組込むのに使用する方法と組成物を含む。重要な点は、非天然アミノ酸を20種の共通アミノ酸の1種に置換するのでなく遺伝子レパートリーに加えるという点である。本発明は非天然アミノ酸をタンパク質に組込むために生合成機構により使用される成分の作製方法、同定方法及びこれらの成分を含む組成物を提供する。本発明は例えば(i)任意タンパク質の任意所望位置に1種以上の非天然アミノ酸を部位選択的に挿入することができ、(ii)原核細胞と真核細胞のいずれにも適用可能であり、(iii)大量の精製突然変異体タンパク質の作製に加えて突然変異体タンパク質の in vivo 試験が可能であり、(iv)多様な非天然アミノ酸の任意のものをタンパク質に in vivo 組込むように応用できる。従って、1つの特定ポリペプチド配列に非天然アミノ酸の多数の異なる部位選択的挿入が可能である。このような挿入は場合により全てが同一型である(例えばポリペプチドの多重点に同一種の複数の非天然アミノ酸を挿入する)か、又は場合により多様な型である(例えばポリペプチドの多重点に異なる非天然アミノ酸種を挿入する)。

30

40

【0087】

定義

【0088】

本発明を詳細に説明する前に、本発明は特定組成物又は生物系に限定されず、当然のことながら種々に適用できることを理解すべきである。また、本明細書で使用する用語は特定態様のみの説明を目的としており、限定的ではないことも理解すべきである。本明細書と特許請求の範囲に使用する単数形は明示しない限り複数形も含む。従って、例えば「分子」と言う場合には場合により2個以上のこのような分子の組合せ等も含む。

【0089】

特に記載しない限り、全科学技術用語は本発明が属する分野で通常使用されていると同一

50

の意味をもつものとする。本発明の趣旨では、以下の用語は次のように定義される。

【0090】

本明細書においてタンパク質及び／又はタンパク質配列は共通の祖先タンパク質又はタンパク質配列から天然又は人工的に誘導される場合に「相同」である。同様に、核酸及び／又は核酸配列は共通の祖先核酸又は核酸配列から天然又は人工的に誘導される場合に相同である。例えば、入手可能な任意突然変異誘発法により任意天然核酸を改変し、1個以上のセクターコドンを加えることができる。この突然変異誘発した核酸は発現されると、1種以上の非天然アミノ酸を含むポリペプチドをコードする。突然変異法は当然のことながら更に1個以上の標準コドンを変異させ、得られる突然変異体タンパク質の1個以上の標準アミノ酸も変異させることができる。相同性は一般に2種以上の核酸又はタンパク質（又はその配列）間の配列類似度から推定される。相同性の判定に有用な配列間類似度の厳密な百分率は該当核酸及びタンパク質により異なるが、通常は25%程度の低い配列類似度を使用して相同性を判定する。もっと高いレベルの配列類似度、例えば30%、40%、50%、60%、70%、80%、90%、95%又は99%以上を使用して相同性を判定することもできる。配列類似度百分率の決定方法（例えばデフォルトパラメーターを使用するBLASTP及びBLASTN）は本明細書に記載し、一般に入手可能である。

10

【0091】

「優先的にアミノアシル化する」なる用語はO-RSが非天然アミノ酸でO-tRNAをアミノアシル化する効率が天然tRNA又はO-tRNAを作製するために使用する出発材料に比較して例えば約70%、約75%、約85%、約90%、約95%、又は約99%以上であることを意味する。その後、非天然アミノ酸は高い忠実度で成長中のポリペプチド鎖に組込まれ、例えば所与セクターコドンで約75%以上、所与セクターコドンで約80%以上、所与セクターコドンで約90%以上、所与セクターコドンで約95%以上、又は所与セクターコドンで99%以上の効率である。

20

【0092】

「セクターコドン」なる用語は翻訳プロセスでO-tRNAにより認識されるが、内在tRNAにより認識されないコドンを意味する。O-tRNAアンチコドンループはmRNA上のセクターコドンを認識し、そのアミノ酸（例えば非天然アミノ酸）をポリペプチド内のこの部位に組込む。セクターコドンとしては例えばナンセンスコドン（例えばアンバー、オーカー及びオパールコドン等の終止コドン）、4塩基以上のコドン、天然又は非天然塩基対から誘導されるコドン等を挙げることができる。所与系ではセクターコドンは更に天然3塩基コドンの1種を含むことができ、内在系は前記天然3塩基コドンを使用せず、このような系としては例えば天然3塩基コドンを認識するtRNAを欠失する系や、天然3塩基コドンがレアコドンである系が挙げられる。

30

【0093】

本明細書において「直交」なる用語は所期系（例えば翻訳系、例えば細胞）により低効率で使用される分子（例えば直交tRNA（O-tRNA）及び／又は直交アミノアシルtRNAシンテターゼ（O-RS））を意味する。直交とは直交tRNA及び／又は直交RSが所期翻訳系で機能できないか又は効率が低く、例えば20%未満、10%未満、5%未満、又は例えば1%未満であることを意味する。例えば、直交tRNAが所期翻訳系で所期翻訳系の任意内在RSをアミノアシル化する効率は内在RSによる内在tRNAのアミノアシル化に比較して低下するか又はゼロである。別の例では、直交RSが所期翻訳系で任意内在tRNAをアミノアシル化する効率は内在RSによる内在tRNAのアミノアシル化に比較して低下するか又はゼロである。「直交性の改善」とは出発材料又は天然tRNAもしくはRSに比較して直交性が増していることを意味する。

40

【0094】

「相補的」なる用語は直交対の成分であるO-tRNAとO-RSが共同して機能できる（例えばO-RSがO-tRNAをアミノアシル化する）ことを意味する。

【0095】

50

「から誘導される」なる用語は生物から単離された成分、単離して改変した成分、又は生物からの成分の情報を使用して作製（例えば化学的に合成）された成分を意味する。

【0096】

「翻訳系」なる用語は成長中のポリペプチド鎖（タンパク質）に天然アミノ酸を組込むために必要な成分を意味する。例えば、成分としてはリボソーム、tRNA、シンテターゼ、mRNA等を挙げることができる。本発明の成分は *in vivo* でも *in vitro* でも翻訳系に加えることができる。

【0097】

「不活性RS」なる用語はそのコグネイトtRNAをアミノ酸でアミノアシル化できないように突然変異させたシンテターゼを意味する。

10

【0098】

「選択物質」なる用語はこのような物質が存在すると、集団から所定成分を選択できる物質（例えば抗生物質、光波長、抗体、栄養素等）を意味する。選択物質は例えば濃度、強度等を変動させることができる。

【0099】

「ポジティブ選択マーカー」なる用語はこのような物質が存在する（例えば発現、活性化等される）と、ポジティブ選択マーカーをもたない生物からポジティブ選択マーカーをもつ生物を識別するマーカーを意味する。

【0100】

「ネガティブ選択マーカー」なる用語はこのような物質が存在する（例えば発現、活性化等される）と、（例えば所望特性をもつ生物から）所望特性をもたない生物を識別できるマーカーを意味する。

20

【0101】

「レポーター」なる用語は本発明に記載する成分を選択するために使用することができる成分を意味する。例えば、レポーターとしてはグリーン蛍光タンパク質、ホタルルシフェラーゼタンパク質、 β -gal/lacZ（ β ガラクトシダーゼ）、Adh（アルコールデヒドロゲナーゼ）等の遺伝子等を挙げることができる。

【0102】

「効率的に認識されない」なる用語はある生物からのRSがO-tRNAをアミノアシル化する効率が例えば約10%未満、約5%未満、又は約1%未満であることを意味する。

30

【0103】

「真核生物」なる用語は系統発生分類の真核生物に属する生物を意味し、例えば動物（例えば哺乳動物、昆虫、爬虫類、鳥類等）、絨毛虫、植物、真菌類（例えば酵母等）、鞭毛虫、微孢子虫、原生生物等である。更に、「原核生物」なる用語は系統発生分類の真正細菌（例えば*Escherichia coli*、*Thermus thermophilus*等）と始原菌（例えば*Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*、*Halobacterium*（例えば*Haloferax volcanii*及び*Halobacterium*種NRC-1）、*A. fulgidus*、*P. furiosus*、*P. horikoshii*、*A. pernix*等）に属する非真核生物を意味する。

40

【0104】

「サプレッサーtRNA」は所与翻訳系でメッセンジャーRNA（mRNA）の読取りを改変するtRNAである。サプレッサーtRNAは例えば終止コドン、4塩基コドン又はレアコドンを介して読み取ることができる。

【0105】

詳細な説明

【0106】

本発明は非天然アミノ酸をタンパク質に *in vivo* 組込むための生合成翻訳機構の新規成分のための方法と組成物に関する。具体的には、直交tRNAと、直交RSと、直交tRNA／直交RS対を含む組成物と作製方法が提供される。これらの成分は、宿主細胞

50

に導入すると、非天然アミノ酸を所期ポリペプチド（タンパク質）に *in vivo* 組込むために細胞の翻訳系で使うことができる。例えば、こうして部位特異的非天然アミノ酸突然変異誘発、又は場合によりランダム非天然アミノ酸突然変異誘発を行うことができる。直交 tRNA はセクターコドンに応答して非天然アミノ酸を送達し、直交シンテターゼは直交 tRNA を非天然アミノ酸で優先的にアミノアシル化する。O-RS は直交 tRNA を 20 種の共通アミノ酸で効率的にアミノアシル化しない。直交対の作製及び同定方法も提供する。

【0107】

非天然アミノ酸の部位特異的 *in vivo* タンパク質組込みを図 1 に示す。セクターコドン（例えばユニークコドン）を所期遺伝子に導入する。遺伝子は mRNA に転写され、リボソームで通常の翻訳が開始する。内在シンテターゼは ATP の存在下に内在 tRNA を天然アミノ酸（aa）でアミノアシル化する。直交 tRNA は ATP の存在下に直交シンテターゼにより非天然アミノ酸で酵素的にアミノアシル化される。リボソームがセクターコドンにぶつかると、セクターアンチコドン（例えばユニークアンチコドン）を含むように改変された直交 tRNA は非天然アミノ酸として突然変異をデコードすることができ、非天然アミノ酸を組込んだ全長産物まで翻訳が進行する。

10

【0108】

直交アミノアシル tRNA シンテターゼ、O-RS

【0109】

非天然アミノ酸を特異的に *in vivo* 組込むためには、20 種の共通アミノ酸ではなく所望非天然アミノ酸のみを tRNA に負荷するようにシンテターゼの基質特異性を改変する。直交シンテターゼが無差別な場合には、ターゲット位置に天然アミノ酸と非天然アミノ酸の混合物を含む突然変異体タンパク質となる。例えば、p-F-Phe を部位特異的に組込む試みの 1 つとして、酵母アンバーサプレッサー tRNA Phe CUA / フェニルアラニル-tRNA シンテターゼ対が p-F-Phe 耐性 Phe 栄養要求性大腸菌株で使用された。例えば R. Furter, Protein Sci., 7: 419 (1998) 参照。酵母 Phe RS は p-F-Phe に対して高い基質特異性をもたないので、増殖培地に過剰の p-F-Phe を添加しても突然変異誘発部位は p-F-Phe が 64 ~ 75 % で残りは Phe と Lys として翻訳された。更に、Phe コドン位置に 7 % の p-F-Phe が検出され、内在大腸菌 Phe RS は Phe 以外に p-F-Phe も組込むと

20

30

【0110】

本発明の組成物は直交アミノアシル tRNA シンテターゼ（O-RS）を含み、O-RS は場合により *in vivo* において直交 tRNA（O-tRNA）を非天然アミノ酸で優先的にアミノアシル化する。1 態様において、O-RS は配列番号 4 ~ 34（表 5 参照）及びその相補的ポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸を含む。別の態様において、O-RS は天然アミノ酸（例えば 20 種の公知アミノ酸の 1 種）よりも非天然アミノ酸に対して改善又は強化された酵素特性をもち、例えば K_m が低い、 k_{cat} が高い、 k_{cat} / K_m が高い等である。典型的 O-tRNA 及び O-RS 分子の配列は実施例 10 に記載する。

40

【0111】

O-RS の生産方法は野生型シンテターゼの骨格から突然変異体シンテターゼのプールを作製した後、20 種の共通アミノ酸に比較した非天然アミノ酸に対する特異性に基づいて突然変異 RS を選択することを基本とする。このようなシンテターゼを単離するために、本発明の選択方法は（i）1 回目の選択からの所期シンテターゼの活性を低くすることが

50

でき、集団を小さくできるので高感度であり、(ii) 各回の選択で選択ストリンジェンシーを変えることが望ましいので「調節可能」であり、(iii) 汎用性であるため、種々の非天然アミノ酸に使用できる。

【0112】

本発明は例えばシンテターゼの活性部位、シンテターゼの編集メカニズム部位、シンテターゼの種々のドメインを組み合わせるにより種々の部位等でシンテターゼを突然変異させ、選択法を適用することにより直交アミノアシル tRNA シンテターゼを作製するための方法を提供する。図 2A はポジティブ選択後にネガティブ選択を行う組合せに基づく *in vivo* 選択／スクリーニングストラテジーを模式的に示す。ポジティブ選択では、ポジティブマーカーの非必須位置に導入したセクターコドンを抑圧し、ポジティブ選択圧下に細胞を生存させる。従って、天然及び非天然アミノ酸両者の存在下で生存する細胞は直交サプレッサー tRNA に天然又は非天然アミノ酸を負荷する活性シンテターゼをコードする。ネガティブ選択では、ネガティブマーカーの非必須位置に導入したセクターコドンを抑圧し、天然アミノ酸特異性をもつシンテターゼを除去する。ネガティブ選択とポジティブ選択後に生存している細胞は直交サプレッサー tRNA を非天然アミノ酸のみでアミノアシル化（負荷）するシンテターゼをコードする。その後、これらのシンテターゼを例えば DNA シャフリング又は他の再帰的突然変異誘発法により更に突然変異誘発することができる。当然のことながら、他の態様では、本発明は場合により例えば O-RS、O-tRNA、対等を同定するために段階の順序を変え、ネガティブ選択／スクリーニング後にポジティブ選択／スクリーニングを行うか、又は逆でも任意組合せが可能である。

【0113】

例えば、図 2B 参照。図 2B ではセクターコドン（例えば TAG）をもつレポーター遺伝子（例えば β ラクターゼ等の抗生物質耐性遺伝子）にセクターコドン（例えばアンバーコドン）を配置する。突然変異 RS のライブラリーのメンバーを含む発現ベクターにこのレポーター遺伝子を配置する。この発現ベクターと直交 tRNA（例えば直交サプレッサー tRNA）を含むベクターを細胞に導入し、選択物質（例えばアンピシリン等の抗生物質培地）の存在下に増殖させる。シンテターゼが所定アミノ酸でサプレッサー tRNA をアミノアシル化（負荷）できる場合のみにセクターコドンはデコードされ、細胞は抗生物質培地で生存する。

【0114】

非天然アミノ酸の存在下にこの選択を適用し、ある程度のアミノアシル化能をもつシンテターゼをコードするシンテターゼ遺伝子を不活性なシンテターゼ遺伝子から選別する。得られるシンテターゼプールは 20 種の天然アミノ酸又は非天然アミノ酸を負荷している可能性がある。非天然アミノ酸のみに負荷するシンテターゼを更に選択するために、第 2 の選択（例えばネガティブ選択）を適用する。この場合には、ネガティブ選択マーカーと O-tRNA を含む発現ベクターと、突然変異 RS ライブラリーのメンバーを含む発現ベクターを使用する。このネガティブ選択マーカーは少なくとも 1 個のセクターコドン（例えば TAG）を含む。これらの発現ベクターを別の細胞に導入し、非天然アミノ酸の不在下で場合により選択物質（例えばテトラサイクリン）の存在下に増殖させる。ネガティブ選択では、天然アミノ酸に特異性をもつシンテターゼが直交 tRNA に負荷するので、ネガティブマーカーのセクターコドンが抑圧され、細胞死に至る。非天然アミノ酸を加えないので、非天然アミノ酸に特異性をもつシンテターゼは生存する。例えば、セクターコドン（例えば終止コドン）をレポーター遺伝子（例えばバルナーゼ等の毒性タンパク質をコードする遺伝子）に導入する。シンテターゼが非天然アミノ酸の不在下でサプレッサー tRNA に負荷することができるならば、細胞は毒性遺伝子産物を翻訳することにより死滅する。両者の選択／スクリーニングを通過する生存細胞は直交 tRNA に非天然アミノ酸を特異的に負荷するシンテターゼをコードする。

【0115】

1 態様では、少なくとも 1 種の組換え直交アミノアシル tRNA シンテターゼ（O-RS

）の生産方法は、（a）第1の生物に由来する少なくとも1種のアミノアシル tRNA シンテターゼ（RS）から誘導される突然変異体 RS のライブラリーを作製する段階と、（b）非天然アミノ酸と天然アミノ酸の存在下に直交 tRNA（O-tRNA）をアミノアシル化するメンバーを突然変異体 RS のライブラリーから選択することにより活性突然変異体 RS のプールを提供する段階と、（c）非天然アミノ酸の不在下に O-tRNA を優先的にアミノアシル化する活性突然変異体 RS のプールをネガティブ選択することにより、O-tRNA を非天然アミノ酸で優先的にアミノアシル化する少なくとも1種の組換え O-RS を提供する段階を含む。場合により、選択されたシンテターゼ遺伝子に突然変異誘発（例えばランダム突然変異誘発、組換え等）により更に突然変異を導入して第2世代シンテターゼライブラリーを作製し、これを使用して所望活性をもつ突然変異体シンテターゼが得られるまで更に選択する。前記方法により生産された組換え O-RS も本発明に含まれる。以下に説明するように、本発明の直交 tRNA / シンテターゼ対は場合により第1の生物に由来する成分を第2の生物に導入することによっても作製される。

【0116】

1 態様では、RS は不活性 RS である。不活性 RS は活性 RS を突然変異させることにより作製することができる。例えば、不活性 RS は少なくとも約5個のアミノ酸を別のアミノ酸（例えばアラニン）に突然変異させることにより作製することができる。

【0117】

突然変異体 RS のライブラリーは当分野で公知の種々の突然変異誘発技術を使用して作製することができる。例えば、突然変異体 RS は部位特異的突然変異、ランダム点突然変異、*in vitro* 相同組換え、キメラ構築物等により作製することができる。1 態様では、シンテターゼの編集部位に突然変異を導入し、編集メカニズムを妨害及び／又は基質特異性を改変させる。例えば図3及び図4参照。図3はチロシン類似体の特異的ライブラリーを作製するための部位特異的突然変異を示す。図4は前記類似体の特異的ライブラリーを作製するためのペンタフルオロフェニルアラニン選択用コンセンサス配列を示す。突然変異体 RS のライブラリーはキメラシンテターゼライブラリー（例えばキメラ *Methanococcus jannaschii* / 大腸菌シンテターゼのライブラリー）も含む。あるシンテターゼのドメインに別のシンテターゼからのドメインを付加又は交換することができる。図5は1種の生物（例えば大腸菌）に由来する1個のドメイン（例えば CPI ドメイン）を他の生物のシンテターゼ（例えば *Methanococcus jannaschii* Tyr RS）に移植する方法を模式的に示す。CPI は大腸菌 Tyr RS からヒト Tyr RS に移植することができる。例えば Wakasugi, K. ら, EMBO J. 17: 297-305 (1998) 参照。図6はキメラ *Methanococcus jannaschii* / 大腸菌シンテターゼの構築を模式的に示し、図7はキメラシンテターゼ（例えば *Methanococcus jannaschii* / 大腸菌シンテターゼ）のライブラリーの作製を模式的に示す。例えば Sieber ら, Nature Biotechnology, 19: 456-460 (2001) 参照。キメラライブラリーを各種特性についてスクリーニングし、例えば発現されたインフレイムのメンバー、所期シンテターゼの活性をもたないメンバー、及び／又は所期シンテターゼの活性を示すメンバーをスクリーニングする。

【0118】

1 態様では、ポジティブ選択段階は、少なくとも1個のセクターコドン（例えばアンバーコドン）を含むポジティブ選択マーカー（例えば抗生物質耐性遺伝子等）と突然変異体 RS のライブラリーを複数の細胞に導入し、選択物質の存在下に複数の細胞を増殖させ、ポジティブ選択マーカー中で少なくとも1個のセクターコドンを抑圧することにより選択物質の存在下に生存する細胞を選択し、活性突然変異体 RS のプールを含むポジティブ選択済み細胞のサブセットを提供することを含む。場合により、選択物質濃度を変動させてもよい。

【0119】

1 態様では、ネガティブ選択段階は、少なくとも1個のセクターコドンを含むネガティ

ブ選択マーカー（例えばクロラムフェニコールアセチルトランスフェラーゼ（CAT）遺伝子等の抗生物質耐性遺伝子）をポジティブ選択からの活性突然変異体RSのプールと共に第2の生物の複数の細胞に導入し、非天然アミノ酸と選択物質を加えた第1の培地で生存するが、非天然アミノ酸と選択物質を加えない第2の培地では生存することができない細胞を選択することにより、少なくとも1種の組換えO-RSをもつ生存細胞を提供することを含む。場合により、選択物質濃度を変動させる。

【0120】

上記第1の培地と第2の培地としては例えば直接レプリカプレート法が挙げられる。例えば、ポジティブ選択を通過した後に、アンピシリン又はクロラムフェニコールの存在下で非天然アミノ酸の不在下に細胞を増殖させる。非天然アミノ酸を加えたレプリカプレートから生存しない細胞を単離する。第2のネガティブ選択株への形質転換は不要であり、表現型が分かる。他の潜在的選択マーカーに比較すると、抗生物質耐性に基づくポジティブ選択は抗生物質濃度を変えることにより選択ストリンジェンシーを調節することができ、細胞が生存可能な最高抗生物質濃度をモニターすることにより抑圧効率を比較することができる。更に、増殖法は集積法である。このため所望表現型を迅速に蓄積することができる。

10

【0121】

別の態様では、活性突然変異体RSのプールをネガティブ選択する段階は、ポジティブ選択段階（b）からの活性突然変異体RSのプールを単離し、少なくとも1個のセクターコドンを含むネガティブ選択マーカー（例えばリボヌクレアーゼバルナーゼ遺伝子等の毒性マーカー遺伝子）と活性突然変異体RSのプールを第2の生物の複数の細胞に導入し、非天然アミノ酸を加えない第1の培地で生存するが、非天然アミノ酸を加えた第2の培地では生存することができない細胞を選択することにより、非天然アミノ酸に特異的な少なくとも1種の組換えO-RSをもつ生存細胞を提供することを含む。場合により、ネガティブ選択マーカーは2個以上のセクターコドンを含む。

20

【0122】

1側面では、ポジティブ選択はポジティブ選択マーカー（例えばCAT遺伝子内にセクターコドン（例えばアンバー終止コドン）を含むクロラムフェニコールアセチルトランスフェラーゼ（CAT）遺伝子）の抑圧に基づくので、クロラムフェニコールをポジティブ選択圧として適用することができる。更に、非天然アミノ酸の存在下と不在下で本明細書に記載するようなポジティブマーカーとネガティブマーカーの両者としてCAT遺伝子を使用することができる。場合により、セクターコドンを含むCAT遺伝子をポジティブ選択に使用しネガティブ選択マーカー（例えば、少なくとも1個以上のセクターコドンを含むバルナーゼ遺伝子等の毒性マーカー）をネガティブ選択に使用する。

30

【0123】

別の側面では、ポジティブ選択は細胞をアンピシリン耐性にするβ-ラクタマーゼ遺伝子の非必須位置におけるセクターコドンを抑圧に基づき、ネガティブマーカーとしてリボヌクレアーゼバルナーゼを使用するネガティブ選択を使用する。ペリプラズムに分泌されるβ-ラクタマーゼに対して、CATは細胞質に局在し、更にアンピシリンは殺菌性であるが、クロラムフェニコールは静菌性である。

40

【0124】

組換えO-RSを更に突然変異させ、選択することができる。1態様では、少なくとも1種の組換え直交アミノアシルtRNAシンテターゼ（O-RS）の生産方法は更に（d）少なくとも1種の組換えO-RSを単離する段階と、（e）少なくとも1種の組換えO-RSから誘導される第2組の突然変異O-RSを作製する段階と、（f）O-tRNAを優先的にアミノアシル化することができる突然変異O-RSが得られるまで段階（b）及び（c）を反復する段階を含むことができる。場合により、段階（d）～（f）を例えば少なくとも約2回反復する。1側面では、第2組の突然変異O-RSは突然変異誘発（例えばランダム突然変異誘発、部位特異的突然変異誘発、組換え又はその組合せ）により作製することができる。

50

【0125】

上記方法における選択段階（例えばポジティブ選択段階（b）、ネガティブ選択段階（c）又はポジティブ選択段階（b）とネガティブ選択段階（c）の両者）のストリンジエンシーは場合により選択ストリンジエンシーの変動を含む。例えば、バルナーゼは極毒性タンパク質であるので、バルナーゼ遺伝子に導入するセクターコドンの数を変えることによりネガティブセクターのストリンジエンシーを制御することができる。本発明の1側面では、初期選択で所望活性を低くできるのでストリンジエンシーを変動できる。即ち、初期選択には低ストリンジエンシー選択基準を適用し、後期選択にはよりストリンジエントな基準を適用する。

【0126】

本発明では例えばO-RS、O-tRNA及びO-tRNA/O-RS対に他の型の選択も使用できる。例えば、ポジティブ選択段階（b）、ネガティブ選択段階（c）又はポジティブ選択段階（b）とネガティブ選択段階（c）の両者はレポーターを使用することができ、レポーターは蛍光活性化細胞ソーティング（FACS）により検出される。例えば、ポジティブ選択マーカー（例えばCAT遺伝子内にセクターコドン（例えばアンバー終止コドン）を含むクロラムフェニコールアセチルトランスフェラーゼ（CAT）遺伝子）を使用してまずポジティブ選択を行った後にネガティブ選択スクリーニングを行い、ネガティブマーカー（例えばT7 RNAポリメラーゼ遺伝子）内の位置でセクターコドン（例えば2個以上）を抑圧できないものを選択することができる。1態様では、ポジティブ選択マーカーとネガティブ選択マーカーを同一ベクター（例えばプラスミド）に配置20
することができる。ネガティブマーカーの発現はレポーター（例えばグリーン蛍光タンパク質（GFP））の発現を誘導する。選択とスクリーニングのストリンジエンシーは変動させることができ、例えばレポーターを蛍光発光させるために必要な光強度を変動させることができる。別の態様では、FACによりスクリーニングされるポジティブ選択マーカーとしてレポーターを使用してポジティブ選択を行った後にネガティブ選択スクリーニングを行い、ネガティブマーカー（例えばバルナーゼ遺伝子）内の位置でセクターコドン（例えば2個以上）を抑圧できないものを選択することができる。

【0127】

場合により、レポーターを細胞表面、ファージディスプレイ等に提示する。細胞表面提示（例えばOmpA細胞表面提示系）は大腸菌細胞表面における特定エピトープ（例えば例えば外膜ポリンOmpAに融合したポリオウイルスCペプチド）の発現に依存する。エピトープは翻訳中にタンパク質メッセージ中のセクターコドンが抑圧されるときにのみ細胞表面に提示される。従って、提示されるペプチドはライブラリー内の突然変異体アミノアシルトRNAシンテターゼの1種により認識されるアミノ酸を含み、特定非天然アミノ酸を含むペプチドに対する抗体を用いて対応するシンテターゼ遺伝子を含む細胞を単離30
することができる。OmpA細胞表面提示系はファージディスプレイの代用としてGeorgiouらにより開発され、改良された。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) 参照。

【0128】

本発明の他の態様は、選択段階の1段階以上をin vitroで実施することを含む。その後、選択した成分（例えばシンテターゼ及び／又はtRNA）を非天然アミノ酸のin vivo組込みに使用する細胞に導入することができる。

【0129】

直交tRNA

【0130】

10

20

30

40

50

直交 tRNA (O-tRNA) の組成物も本発明の特徴であり、例えば O-tRNA はセレクトターコドンを認識し、O-tRNA は直交アミノアシル tRNA シンテターゼにより非天然アミノ酸で優先的にアミノアシル化される。1 態様では、O-tRNA は配列番号 4 ~ 34 (表 5 参照) 及びその相補的ポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸を含む。

【0131】

本発明は組換え直交 tRNA (O-tRNA) の生産方法も提供する。例えば、所望 RS に対するその親和性を維持しながら tRNA の直交性を改善するために、本方法は夫々コグネイトシンテターゼの不在下と存在下で突然変異体サプレッサー tRNA ライブラリーを用いたネガティブ選択とポジティブ選択を組合せる。図 8 参照。ネガティブ選択ではセレクトターコドン(10)をマーカー遺伝子(例えばバルナーゼ遺伝子等の毒性遺伝子)の非必須位置に導入する。例えば *Methanococcus jannaschii* に由来する突然変異 tRNA ライブラリーのメンバーが内在宿主(例えば大腸菌)シンテターゼによりアミノアシル化される(即ち宿主、例えば大腸菌シンテターゼに非直交性である)場合には、セレクトターコドン(例えばアンバーコドン)が抑圧され、毒性遺伝子産物が生産され、細胞死に至る。直交 tRNA 又は非機能的 tRNA をもつ細胞は生存する。次に、セレクトターコドン(例えばアンバーコドン)をポジティブマーカー遺伝子(例えばβ-ラクタマーゼ遺伝子等の薬剤耐性遺伝子)に配置したポジティブ選択で生存細胞を選択する。これらの細胞はコグネイト RS をもつ発現ベクターも含む。これらの細胞を選択物質(例えばアンピシリン)の存在下に増殖させる。次に、同時発現したコグネイトシンテターゼによりアミノアシル化され、このセレクトターコドンに応答してアミノ酸を挿入することができる tRNA を選択する。非機能的 tRNA 又は所期シンテターゼにより認識することができない tRNA を含む細胞は抗生物質に感受性である。従って、(i) 内在宿主(例えば大腸菌)シンテターゼの基質でなく、(ii) 所期シンテターゼによりアミノアシル化(20)することができ、(iii) 翻訳能力のある tRNA はどちらの選択でも生存する。

【0132】

組換え O-tRNA の生産方法は、(a) 第 1 の生物に由来する少なくとも 1 種の tRNA (例えばサプレッサー tRNA) から誘導される突然変異体 tRNA のライブラリーを作製する段階と、(b) 第 1 の生物に由来する RS の不在下に第 2 の生物に由来するアミノアシル tRNA シンテターゼ(RS)によりアミノアシル化される突然変異体 tRNA (30)のライブラリーをネガティブ選択することにより突然変異体 tRNA のプールを提供する段階と、(c) 導入した直交 RS (O-RS) によりアミノアシル化されるメンバーを突然変異体 tRNA のプールから選択することにより、セレクトターコドン(10)を認識し、第 2 の生物に由来する RS により効率的に認識されず、O-RS により優先的にアミノアシル化される少なくとも 1 種の組換え O-tRNA を提供する段階を含む。1 態様では、組換え O-tRNA は改善された直交性をもつ。

【0133】

突然変異 tRNA のライブラリーを構築する。例えば図 9 参照。tRNA の所望ループ(例えばアンチコドンループ(Dアーム、Vループ、TΨCアーム)又はループ組合せ又は全ループ)の特定位置(例えば非保存位置又は保存位置)、ランダム位置又は両者の組合(40)せに突然変異を導入することができる。tRNA のキメラライブラリーも本発明に含まれる。場合により各種生物(例えば真正細菌又は始原菌等の微生物)に由来する tRNA シンテターゼのライブラリー(例えば天然ダイバーシティを含むライブラリー)(例えば Shortらの米国特許第 6, 238, 884 号、Schallenberg らの米国特許第 5, 756, 316 号、Petersen らの米国特許第 5, 783, 431 号、Thompson らの米国特許第 5, 824, 485 号、及び Shortらの米国特許第 5, 958, 672 号参照)を構築して直交対をスクリーニングすることに留意されたい。

【0134】

1 態様では、アミノアシル tRNA シンテターゼによりアミノアシル化される突然変異体 (50)

R N A のライブラリーをネガティブ選択する段階（段階（b））は、少なくとも 1 個のセレクトターコドンを含む毒性マーカー遺伝子と突然変異体 t R N A のライブラリーを第 2 の生物に由来する複数の細胞に導入し、少なくとも 1 種の直交 t R N A 又は非機能的 t R N A を含む突然変異体 t R N A のプールを含む生存細胞を選択することを含む。例えば、毒性マーカー遺伝子は場合によりリボヌクレアーゼバルナーゼ遺伝子であり、リボヌクレアーゼバルナーゼ遺伝子は少なくとも 1 個のアンバーコドンを含む。場合により、リボヌクレアーゼバルナーゼ遺伝子に 2 個以上のアンバーコドンを含むことができる。生存細胞は例えば比較細胞密度アッセイを使用することにより選択することができる。

【0135】

1 態様では、導入した直交 R S（O-R S）によりアミノアシル化されるメンバーを突然変異体 t R N A のプールから選択する段階は、少なくとも 1 個のセレクトターコドンを含む薬剤耐性遺伝子（例えば β ラクタマーゼ遺伝子、例えば少なくとも 1 個のアンバーコドンを含む β ラクタマーゼ遺伝子）を含むポジティブ選択マーカー遺伝子と、O-R S と、突然変異体 t R N A のプールを第 2 の生物に由来する複数の細胞に導入し、選択物質（例えば抗生物質）の存在下に増殖した生存細胞を選択することにより、O-R S によりアミノアシル化され、少なくとも 1 個のセレクトターコドンに応答してポジティブマーカー遺伝子によりコードされる翻訳産物にアミノ酸を挿入する少なくとも 1 種の組換え t R N A をもつ細胞プールを提供することを含む。別の態様では、選択物質濃度を変動させる。本方法により生産された組換え O-t R N A も本発明に含まれる。

【0136】

O-R S の作製について上述したように、選択段階のストリンジェンシーは変動可能である。更に、本明細書に記載する他の選択／スクリーニング法（例えば F A C S、細胞及びファージディスプレイ）も使用できる。

【0137】

セレクトターコドン

【0138】

本発明のセレクトターコドンはタンパク質生合成機構の遺伝コドン枠を拡張する。例えば、セレクトターコドンとしては例えばユニーク 3 塩基コドン、ナンセンスコドン、（例えばアンバーコドン又はオパールコドン等の終止コドン）、非天然コドン、4 塩基（以上）のコドン等が挙げられる。例えば 1 個以上、2 個以上、3 個以上等の多数のセレクトターコドンを所望遺伝子に導入することができる。更に、当然のことながら、こうして複数の異なる（又は類似又は同一の）非天然アミノ酸を（即ち複数のセレクトターコドンの使用により）アミノ酸に正確に組込むことができる。

【0139】

64 種の遺伝コドンは 20 種のアミノ酸と 3 種の終止コドンをコードする。翻訳終結には 1 個の終止コドンしか必要ないので、他の 2 個は主に非タンパク生産アミノ酸をコードするために使用することができる。アンバー終止コドンである U A G は *in vitro* 生合成系とアフリカツメガエル卵母細胞で使用して非天然アミノ酸の組込みに成功している。3 種の終止コドンのうちで U A G は大腸菌での使用頻度が最も低い終止コドンである。大腸菌株には U A G を認識して天然アミノ酸を挿入する天然サプレッサー t R N A を含むものもある。更に、これらのアンバーサプレッサー t R N A は慣用タンパク質突然変異誘発でも使用されている。各種はその天然アミノ酸の各コドンを優先的に使用するので、本発明のセレクトターコドンの設計／選択には場合によりこのような優先性を利用する。

【0140】

本明細書では非天然アミノ酸について記載するが、当然のことながら特定セレクトターコドンに応答して天然アミノ酸を組込むために同様のストラテジーを使用することもできる。即ち、本明細書全般に記載する非天然アミノ酸の負荷と同様にセレクトターコドンを認識する直交 t R N A に天然アミノ酸を負荷するようにシンテターゼを改変することができる。

【0141】

1 態様では、本方法は非天然アミノ酸の *in vivo* 組込みに終止コドンであるセレクト

ターコドンを使用する。例えば、終止コドン（例えばUAG）を認識するO-tRNAを作製し、O-RSにより所望非天然アミノ酸でアミノアシル化する。このO-tRNAは天然アミノアシル-tRNAシントターゼにより認識されない。慣用部位特異的突然変異誘発を使用して終止コドン（例えばTAG）をタンパク質遺伝子内の所期部位に導入することができる。例えばSayers, J. R., Schmidt, W. Eckstein, F. 5', 3' Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis. *Nucleic Acids Res*, 791-802 (1988) 参照。O-RS、O-tRNA及び突然変異体遺伝子をin vivo融合すると、UAGコドンにตอบสนองして非天然アミノ酸が組込まれ、特定位置に非天然アミノ酸を含むタンパク質が得られる。 10

【0142】

非天然アミノ酸のin vivo組込みは宿主（例えば大腸菌）にさほど影響を与えずに行うことができる。例えば、UAGコドンの抑圧効率はO-tRNA（例えばアンバーサプレッサーtRNA）と（UAGコドンに結合してリボソームから増殖ペプチドの放出を開始する）放出因子1（RF1）の競合に依存するので、抑圧効率は例えばO-tRNA（例えばサプレッサーtRNA）の発現レベルを高めるか又はRF1欠損株を使用することにより調節することができる。更に、生物又は生物のゲノムの一部にランダム突然変異誘発を行い、例えば本明細書に記載するレポーター系の1種を使用して適正な選択を実施することにより抑圧効率と非天然アミノ酸取込みを調節することができる。 20

【0143】

非天然アミノ酸はレアコドンでコードすることもできる。例えば、in vitroタンパク質合成反応でアルギニン濃度を下げると、レアアルギニンコドンAGGはアラニンでアシル化された合成tRNAによるAlaの挿入に有効であることが分かっている。例えばC. H. Ma, W. Kudlicki, O. W. Odom, G. Kramer and B. Hardesty, *Biochemistry*, 32: 7939 (1993) 参照。この場合には、合成tRNAは大腸菌に少量種として存在する天然tRNA^{Arg}と競合する。生物によっては全三重項コドンを使用しないものもある。*Micrococcus luteus*で割り当てられないコドンAGAがin vitro転写／翻訳抽出物へのアミノ酸の挿入に使用されている。例えばA. K. KowalとJ. S. Oliver 30, *Nucl. Acid. Res.*, 25: 4685 (1997) 参照。本発明の成分はこれらのレアコドンを用いてin vivo使用するために作製することができる。

【0144】

セクターコドンは更に4塩基以上のコドン（例えば4、5、6、又は7塩基以上のコドン）も含む。4塩基コドンの例としては例えばAGGA、CUAG、UAGA、CCCU等が挙げられる。5塩基コドンの例としては例えばAGGAC、CCCCU、CCUC、CUGA、CUACU、UAGGC等が挙げられる。例えば、突然変異O-tRNA（例えば特殊フレームシフトサプレッサーtRNA）とアンチコドンループ（例えば少なくとも8~10ntアンチコドンループ）の存在下では、4塩基以上のコドンは単一アミノ酸として読取られる。他の態様では、アンチコドンループは例えば少なくとも4塩基コドン、少なくとも5塩基コドン、又は少なくとも6塩基コドン又はそれ以上をデコードすることができる。4塩基コドンは256種が考えられるので、4塩基以上のコドンを使用すると同一細胞で多重非天然アミノ酸をコードすることができる。J. Christopher Andersonら, *Exploring the Limits of Codon and Anticodon Size, Chemistry and Biology*, Vol. 9, 237-244 (2002); Thomas J. Magliery, *Expanding the Genetic Code: Selection of Efficient Suppressors of Four-base Codons and Identification of "Shifty" Four-base Codons with a Library Approach in* 40 50

Escherichia coli, J. Mol. Biol. 307:755-769 (2001) も参照されたい。

【0145】

本発明の方法はフレームシフト抑圧に基づく拡張コドンを使用する。4塩基以上のコドンは例えば1個又は多数の非天然アミノ酸を同一タンパク質に挿入することができる。例えば、*in vitro* 生合成法を使用して非天然アミノ酸をタンパク質に組込むために4塩基コドンが使用されている。例えばC. H. Ma, W. Kudlicki, O. W. Odom, G. Kramer and B. Hardesty, Biochemistry, 1993, 32, 7939 (1993); 及びT. Hohsaka, D. Kajihara, Y. Ashizuka, H. Murakami and M. Sisido, J. Am. Chem. Soc., 121:34 (1999) 参照。2個の化学的にアシル化したフレームシフトサプレッサートRNAを使用して2-ナフチルアラニンとリジンのNB D誘導体を同時にストレプトアビジンに*in vitro* 組込むためにCGGGとAGGUが使用されている。例えばT. Hohsaka, Y. Ashizuka, H. Sasaki, H. Murakami and M. Sisido, J. Am. Chem. Soc., 121:12194 (1999) 参照。*in vivo* 研究では、MooreらがNCUAアンチコドンをもつtRNA^{Leu}誘導体がUAGNコドンを抑圧できるかどうかを試験し(NはU、A、G又はCであり得る)、四重項UAGAはUCUAアンチコドンをもつtRNA^{Leu}により13から26%の効率でデコードすることができるが、0又は-1フレームでは殆どデコードできないことを見出した。B. Moore, B. C. Persson, C. C. Nelson, R. F. Gesteland and J. F. Atkins, J. Mol. Biol., 298:195 (2000) 参照。1態様では、本発明ではレアコドン又はナンセンスコドンに基づく拡張コドンを使用し、他の望ましくない部位でのミスセンス読み飛ばしとフレームシフト抑圧を減らすことができる。

10

20

【0146】

翻訳バイパス系を使用して非天然アミノ酸を所望ポリペプチドに組込むこともできる。1翻訳バイパス系では、大きい配列を遺伝子に挿入するが、タンパク質に翻訳しない。この配列はリボソームに配列を飛び越させて挿入の下流の翻訳を再開するための合図として機能する構造を含む。

【0147】

非天然アミノ酸をポリペプチドに組込むための他の上記方法の代用又は併用してトランス翻訳系を使用することができる。この系は大腸菌に存在するtmRNAと呼ぶ分子を利用する。このRNA分子はアラニルトRNAに構造的に近縁であり、アラニルシンテターゼによりアミノアシル化される。tmRNAとtRNAの相違はアンチコドンループが特殊な大きい配列で置換されていることである。この配列は鋳型としてtmRNA内でコードされるオープンリーディングフレームを使用して停止した配列でリボソームに翻訳を再開させることができる。本発明では、直交シンテターゼで優先的にアミノアシル化して非天然アミノ酸を負荷することができる直交tmRNAを作製することができる。この系を使用する遺伝子を転写することにより、リボソームは特定部位で停止し、非天然アミノ酸はこの部位に導入され、その後、直交tmRNA内にコードされた配列を使用して翻訳が再開する。

30

40

【0148】

セクターコドンは場合により非天然塩基対を含む。これらの非天然塩基対は更に既存遺伝アルファベットを拡張する。塩基対が1個増えると、三重項コドンの数は64から125に増す。3番目の塩基対の性質としては安定的且つ選択的な塩基対合、高い忠実度で効率的なポリメラーゼによるDNAへの酵素組込み、及び新生非天然塩基対の合成後の効率的な連続プライマー伸長が挙げられる。方法と組成物に応用可能な非天然塩基対については、例えばHiraora, An unnatural base pair for incorporating amino acid analogues into protein, Nature Biotechnology, 20:177-182 (2

50

002)に記載されている。他の関連文献は以下に挙げる。

【0149】

in vivo使用では、非天然ヌクレオシドは膜透過性であり、リン酸化して対応する三リン酸塩を形成する。更に、増加した遺伝情報は安定しており、細胞酵素により破壊されない。Bennerらによる従来の研究ではカノニカルワトソンクリック対とは異なる水素結合パターンを利用しており、そのうちで最も注目に価する例はイソC：イソG対である。例えばC. Switzer, S. E. Moroney and S. A. Benner, J. Am. chem. Soc., 111:8322 (1989); J. A. Piccirilli, T. Krauch, S. E. Moroney and S. A. Benner, Nature, 1990, 343:33 (1990); 及びE. T. Kool, Curr. Opin. Chem. Biol., 4:602 (2000) 参照。これらの塩基は一般に天然塩基とある程度ミスペアであり、酵素により複製することができない。Koolらは水素結合の代わりに塩基間の疎水性充填相互作用により塩基対の形成を誘導できることを示した。E. T. Kool, Curr. Opin. Chem. Biol., 4:602 (2000); 及びK. M. GuckianとE. T. Kool, Angew. Chem. Int. Ed. Engl., 36, 2825 (1998) 参照。上記全要件を満足する非天然塩基対を開発する試みとして、Schultz, Romerbergらは一連の非天然疎水性塩基を体系的に合成し、試験している。PICS：PICS自己対(図10参照)が天然塩基対よりも安定であることが判明し、大腸菌DNAポリメラーゼIのKlenowフラグメント(KF)によりDNAに効率的に組込むことができる。例えばD. L. McMin, A. K. Ogawa, Y. Q. Wu, J. Q. Liu, P. G. Schultz and F. E. Romersberg, J. Am. Chem. Soc., 121:11586 (1999); 及びA. K. Ogawa, Y. Q. Wu, K. L. McMin, J. Q. Liu, P. G. Schultz and F. E. Romersberg, J. Am. Chem. Soc., 122:3274 (2000) 参照。生物機能に十分な効率と選択性でKFにより3MN：3MN自己対を合成することができる。例えばA. K. Ogawa, Y. Q. Wu, M. Berger, P. G. Schultz and F. E. Romersberg, J. Am. Chem. Soc., 122:8803 (2000) 参照。しかし、どちらの塩基も後期複製用チェーンターミネーターとして作用するに止まっている。PICS自己対を複製するために使用できる突然変異体DNAポリメラーゼが最近開発された。更に、7AI自己対も複製できる。例えばE. J. L. Tae, Y. Q. Wu, G. Xia, P. G. Schultz and F. E. Romersberg, J. Am. Chem. Soc., 123:7439 (2001) 参照。Cu(II)と結合すると安定な対を形成する新規メタロ塩基対Dipic：Pyも開発されている。E. Meggers, P. L. Holland, W. B. Tolman, F. E. Romersberg and P. G. Schultz, J. Am. Chem. Soc., 122:10714 (2000)。拡張コドンと非天然コドンは天然コドンに本質的に直交性であるので、本発明の方法はその直交tRNAを作製するためにこの性質を利用することができる。

【0150】

直交tRNAと直交アミノアシルtRNAシンテターゼの対

【0151】

直交対はO-tRNA(例えばサプレッサーtRNA、フレームシフトtRNA等)とO-RSから構成される。O-tRNAは上述のように内在シンテターゼによりアシル化されず、セクターコドンをデコードすることができる。O-RSは例えば拡張アンチコドンループでO-tRNAを認識し、O-tRNAを非天然アミノ酸で優先的にアミノアシル化する。直交対と直交対の組成物の作製方法も本発明に含まれる。多重直交tRNA／シンテターゼ対の開発により、各種コドンを使用して多重非天然アミノ酸を同一ポリペプチド／タンパク質に同時に組込むことが可能になった。

【0152】

10

20

30

40

50

本発明において、方法と関連組成物は非天然アミノ酸をタンパク質に *in vivo* 組込むことができる直交対 (O-tRNA/O-RS) の作製に関する。例えば、本発明の O-tRNA の組成物は直交アミノアシル tRNA シンテターゼ (O-RS) を含むことができる。1 態様において、O-tRNA と O-RS は相補的とすることができる (例えば直交 O-tRNA/O-RS 対)。対の例としては mutRNA Tyr-mutTyrRS 対 (例えば mutRNA Tyr-SS12TyrRS 対)、mutRNA Leu-mutLeuRS 対、mutRNA Thr-mutThrRS 対、mutRNA Glu-mutGluRS 対等が挙げられる。別の態様において、本発明の直交対は直交 tRNA-アミノアシル tRNA シンテターゼ対の所望特性を含み、本発明の対の特性をもたない大腸菌由来 mutRNA Glu-mutGluRS、酵母由来 mutRNA Asp-mutAspRS 又は酵母由来 mutRNA PheCUA-mutPheRS 以外のものである。

10

【0153】

O-tRNA と O-RS は起源と宿主名で記載されている多様な生物に由来する天然 tRNA 及び/又は RS を突然変異させることにより誘導することができる。1 態様において、O-tRNA と O-RS は少なくとも 1 種の生物に由来する。別の態様では、O-tRNA は第 1 の生物に由来する天然に存在するか又は突然変異させた天然 tRNA を突然変異させることにより誘導し、O-RS は第 2 の生物に由来する天然に存在するか又は突然変異させた天然 RS を突然変異させることにより誘導する。

【0154】

本発明は特異的 O-tRNA/O-RS 対の作製方法も提供する。これらの方法は直交 tRNA/RS 対を作製するために試みられた他のストラテジーについて後述する問題を解決する。具体的には、本発明の方法は、(a) 第 1 の生物に由来する少なくとも 1 種の tRNA から誘導される tRNA のライブラリーを作製する段階と、(b) 第 1 の生物に由来する RS の不在下で第 2 の生物に由来するアミノアシル tRNA シンテターゼ (RS) によりアミノアシル化される突然変異体 tRNA のライブラリーをネガティブ選択することにより、突然変異体 tRNA のプールを提供する段階と、(c) 導入した直交 RS (O-RS) によりアミノアシル化されるメンバーを突然変異体 tRNA のプールから選択することにより、少なくとも 1 種の組換え O-tRNA を提供する段階を含む。少なくとも 1 種の組換え O-tRNA はセクターコドンを認識し、第 2 の生物に由来する RS により効率的に認識されず、O-RS により優先的にアミノアシル化される。本方法は更に、(d) 第 3 の生物に由来する少なくとも 1 種のアミノアシル tRNA シンテターゼ (RS) から誘導される少なくとも 1 種の組換え O-tRNA RS のライブラリーを作製する段階と、(e) 非天然アミノ酸と天然アミノ酸の存在下に少なくとも 1 種の組換え O-tRNA を優先的にアミノアシル化するメンバーを突然変異体 RS のライブラリーから選択することにより、活性突然変異体 RS のプールを提供する段階と、(f) 非天然アミノ酸の不在下に少なくとも 1 種の組換え O-tRNA を優先的にアミノアシル化する活性突然変異体 RS のプールをネガティブ選択することにより、非天然アミノ酸に特異的な少なくとも 1 種の組換え O-RS と少なくとも 1 種の組換え O-tRNA を含む少なくとも 1 種の特異的 O-tRNA/O-RS 対を提供する段階を含む。本発明の方法により生産された対も本発明に含まれる。

20

30

40

【0155】

既存大腸菌 tRNA/シンテターゼ対からの直交 tRNA/シンテターゼ対の作製は従来試みられている。この方法は他のシンテターゼに対する直交性と翻訳能力を維持しながら tRNA-シンテターゼ界面のヌクレオチドを突然変異させることによりそのコグネイトシンテターゼに対する tRNA の親和性を失わせる。出発鋳型としてコグネイト野生型シンテターゼを使用し、組換え直交 tRNA のみを認識する突然変異体シンテターゼを得る。tRNA Glu2 と複合化した大腸菌グルタミンイル-tRNA シンテターゼ (GluRS) の X 線結晶構造分析によると、GluRS と特異的に接触する 3 個の部位 (「ノブ」) が tRNA Glu2 で同定された。例えば D. R. Liu, T. J. Magliery and P. G. Schultz, Chem. Biol., 4: 685 (1997);

50

及びD. R. Liu, T. J. Magliery, M. Pasternak and P. G. Schultz, Proc. Natl. Acad. Sci. U S A, 94:10092 (1997) 参照。これらの部位をtRNAで突然変異させ、ノブ1、2及び3の可能な全組合せを含む突然変異体サプレッサーtRNAを作製し、GlnRSによるin vitroアミノアシル化とコリスミン酸ムターゼのアンバー突然変異体のin vitro抑圧により個々に試験した。全3個のノブの突然変異をもつ突然変異体tRNA (O-tRNA) は全内在大腸菌シンテターゼと直交性であり、翻訳能力があることが判明した。次に、DNAシャフリングとオリゴヌクレオチド誘導突然変異誘発を複数回実施して突然変異体GlnRSのライブラリーを作製した。大腸菌株BT235を使用してこれらの突然変異体酵素がO-tRNAをin vivoアシル化する能力について選択した。突然変異体GlnRSがO-tRNAにグルタミンを負荷する場合のみにlacZのゲノムアンバーコドンを抑圧することができ、BT235細胞はラクトース最少培地で増殖することができる。各選択後に生存している数個の突然変異体シンテターゼを精製し、in vitroアッセイした。野生型(wt) tRNA Glnアシル化と突然変異体シンテターゼによるO-tRNAアシル化の比は複数回の選択後に有意に低下した。他方、野生型大腸菌tRNA Gln2よりも効率的にO-tRNAに負荷する突然変異体大腸菌GlnRSは得られなかった。7回のDNAシャフリングと選択後に得られた最良の突然変異体は野生型tRNA Glnの9分の1の率でしかO-tRNAをアシル化しない。しかし、野生型tRNAが非天然アミノ酸で誤ってアシル化されると致死性表現型となる可能性があるので、これらの実験は所望特性をもつシンテターゼ候補(例えば野生型tRNAをアシル化しないシンテターゼ)を生産することができなかった。更に、tRNA内の突然変異はアミノアシル化と翻訳の両者に関して複雑で非付加的方法で相互作用する。D. R. Liu, T. J. Magliery and P. G. Schultz, Chem. Biol., 4:685 (1997) 参照。従って、所望特性をもつ機能的対を提供するには一般に代替方法が使用されている。

【0156】

直交tRNA／シンテターゼ対を作製する第2のストラテジーは別の生物に由来するtRNA／シンテターゼ対を大腸菌に導入する。異種シンテターゼ候補の性質としては、例えば大腸菌tRNAに負荷しない等が挙げられ、異種tRNA候補の性質としては、例えば大腸菌シンテターゼによりアシル化されない等が挙げられる。更に、異種tRNAから誘導されるサプレッサーtRNAは全大腸菌シンテターゼに直交性である。Schimmelらは大腸菌GlnRS (EcGlnRS) がSaccharomyces cerevisiae tRNA Glnをアシル化しないことを報告した (EcGlnRSはSaccharomyces cerevisiae GlnRS (ScGlnRS) がもつN末端RNS結合ドメインをもたない)。E. F. WhelihanとP. Schimmel, EMBO J., 16:2968 (1997) 参照。その後、Saccharomyces cerevisiaeアンバーサプレッサーtRNA Gln (Sc tRNA GlnCUA) を分析し、同様にEcGlnRSの基質でないか否かを調べた。in vitroアミノアシル化アッセイによるとその通りであり、in vitro抑圧試験によるとSc tRNA GlnCUAは翻訳能力がある。例えばD. R. LiuとP. G. Schultz, Proc. Natl. Acad. Sci. U S A, 96:4780 (1999) 参照。ScGlnRSがin vitroで大腸菌tRNAをアシル化せず、Sc tRNA GlnCUAのみをアシル化することも判明した。ScGlnRSが大腸菌でSc tRNA GlnCUAをアミノアシル化する程度もin vivo相補アッセイにより評価された。アンバーナンセンス突然変異がβラクタマーゼ遺伝子の許容部位に導入された。アンバーサプレッサーtRNAによる突然変異の抑圧は全長βラクタマーゼを生じ、細胞にアンピシリン耐性を付与するはずである。Sc tRNA GlnCUAのみが発現される場合に細胞は20 μg/mLアンピシリンのIC₅₀を示し、内在大腸菌シンテターゼによるアシル化が殆どないことを示し、Sc tRNA GlnCUAがScGlnRSと同時に発現される場合には細胞は約500 μg/mLアンピシリンのIC₅₀を獲得し、ScG

l n R Sが大腸菌でS c t R N A G l n C U Aを効率的にアシル化することを示す。D. R. L i uとP. G. S c h u l t z, Proc. Natl. Acad. Sci. U S A, 96:4780 (1999) 参照。S a c c h a r o m y c e s c e r e v i s i a e t R N A G l n C U A / G l n R Sは大腸菌に直交性である。

【0157】

その後、このストラテジーはt R N A ^{A s p} / A s p R S系に応用された。S a c c h a r o m y c e s c e r e v i s i a e t R N A ^{A s p}は大腸菌シンテターゼに直交性であることが知られている。例えばB. P. D o c t o rとJ. A. M u d d, J. Biol. Chem., 238:3677 (1963); 及びY. K w o kとJ. T. W o n g, Can. J. Biochem., 58:213 (1980) 参照。上記i n v i v o β
 ラクタマーゼアッセイによると、これ(S c t R N A A s p C U A)に由来するアンバー
 サプレッサーt R N Aも大腸菌で直交性であることが分かった。他方、t R N A ^{A s p}の
 アンチコドンはA s p R Sの必須認識エレメントであり(例えばR. G i e g e, C. F l o r e n t z, D. K e r n, J. G a n g l o f f, G. E r i a n i and D. M o r a s, Biochimie, 78:605 (1996) 参照)、このアンチコドン
 をC U Aに突然変異させるとA s p R Sに対するサプレッサーの親和性が失われる。大腸
 菌A s p R S E 9 3 K突然変異体は野生型A s p R Sよりもほぼ一桁良好に大腸菌アンバー
 サプレッサーt R N A A s p C U Aを認識することが判明した。例えばF. M a r t i n, 'Thesis', U n i v e r s i t e L o u i s P a s t e u r, S t r a s b o u r g, F r a n c e, 1995参照。S a c c h a r o m y c e s c e r e v
 i s i a e A s p R S (E188K)に関連突然変異を導入するとS c t R N A A s p
 C U Aに対するその親和性は回復すると予想された。i n v i t r oアミノアシル化実
 験によると、S a c c h a r o m y c e s c e r e v i s i a e A s p R S (E18
 8 K)突然変異体は大腸菌t R N Aをアシル化しないが、中程度の効率でS c t R N A A
 s p C U Aに負荷することが判明した。例えばM. P a s t r n a k, T. J. M a g l i e r y and P. G. S c h u l t z, Helv. Chim. Acta, 83:2
 277 (2000) 参照。S c t R N A A s p C U A / S c A s p R S (E188K)は大腸菌で別の直交対として機能できるが、活性は弱い。

【0158】

同様のアプローチとして、直交シンテターゼとして異種シンテターゼを使用するが、直交
 t R N Aとしては同一生物又は近縁生物の突然変異体イニシエーターt R N Aを使用する
 方法がある。R a j B h a n d a r yらはヒトイニシエーターt R N A f M e tのアンバー
 突然変異体がE c G l n R Sを同時発現させる場合のみに大腸菌G l n R Sによりアシル
 化され、酵母細胞でアンバーサプレッサーとして機能することを発見した。A. K. K
 o w a l, C. K o h r e r and U. L. R a j B h a n d a r y, Proc. Natl. Acad. Sci. U S A, 98:2268 (2001) 参照。即ち、この
 対は酵母で使用される直交対である。また、大腸菌シンテターゼに対して不活性な大腸菌
 イニシエーターt R N A f M e tアンバー突然変異体も発見された。この突然変異体t R
 N Aに負荷する突然変異体酵母T y r R Sを選択し、大腸菌で直交対を得た。A. K. K
 o w a lら(2001) 前出参照。

【0159】

原核生物と真核生物のt R N A T y r / T y r R S対には有意な差があり、原核t R N A
 T y rのアイデンティティーエレメントは長い可変アームを含むが、真核t R N A T y r
 のアームは短い。更に、真核t R N A T y rはC1:G72ポジティブ認識エレメントを
 含むが、原核t R N A T y rはこのようなコンセンサス塩基対をもたない。i n v i t
 r o試験によると、S a c c h a r o m y c e s c e r e v i s i a eとヒトのt R N
 A T y rは細菌シンテターゼによりアミノアシル化することができず、そのT y r R Sが
 細菌t R N Aをアミノアシル化しないことも分かった。例えばK. W a k a s u g i, C.
 L. Q u i n n, N. T a o and P. S c h i m m e l, EMBO J., 17
 :297 (1998); 及びT. A. K l e e m a n, D. W e i, K. L. S i m p s

on and E. A. First, *J. Biol. Chem.*, 272:14420 (1997) 参照。しかし、これらの直交性の有望な全特徴にも拘わらず、*in vivo* βラクターゼ相補アッセイによると、*Saccharomyces cerevisiae* とヒトのいずれに由来するアンバーサプレッサー tRNA Tyr CUA も大腸菌で直交性でないことが判明した。例えば L. Wang, T. J. Magliery, D. R. Liu and P. G. Schultz, *J. Am. Chem. Soc.*, 122:5010 (2000) 参照。サプレッサー tRNA を大腸菌シンテターゼによりアシル化できるか否かはアンチコドンの単一ヌクレオチドの変異 (G34→C34) による。

【0160】

本発明の方法を使用して上記所望対及び対の成分を進化させ、所望特性、例えば O-tRNA を非天然アミノ酸で優先的にアミノアシル化できるという特性をもつ直交 tRNA/シンテターゼ対を作製する。 10

【0161】

資源及び宿主生物

【0162】

例えば少なくとも第1の生物又は同一でも異なってもよい少なくとも2種の生物に由来する直交 tRNA-RS 対は各種宿主生物 (例えば第2の生物) で使用することができる。本発明の方法の第1の生物と第2の生物は同一でも異なってもよい。1態様では第1の生物は例えば *Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*、*Halobacterium*、*Escherichia coli*、*A. fulgidus*、*Halobacterium*、*P. furiosus*、*P. horikoshii*、*A. pernix*、*T. thermophilus* 等の原核生物である。あるいは、第1の生物は例えば植物 (例えば単子葉植物又は双子葉植物等の複雑な植物)、藻類、原生動物、真菌類 (例えば酵母等)、動物 (例えば哺乳動物、昆虫、節足動物等) 等の真核生物でもよい。別の態様では、第2の生物は *Methanococcus jannaschii*、*Methanobacterium thermoautotrophicum*、*Halobacterium*、*Escherichia coli*、*A. fulgidus*、*Halobacterium*、*P. furiosus*、*P. horikoshii*、*A. pernix*、*T. thermophilus* 等の原核生物である。あるいは、第2の生物は例えば植物、真菌類、動物等の真核生物でもよい。 20 30

【0163】

上述のように、対の各成分は同一生物に由来するものでも異なる生物に由来するものでもよい。例えば、tRNA は例えば始原菌 (例えば *Methanococcus jannaschii* 及び *Halobacterium* NRC-1) 又は真正細菌 (例えば大腸菌) 等の原核生物に由来し、シンテターゼは同一又は異なる原核生物 (例えば *Methanococcus jannaschii*、*Archaeoglobus fulgidus*、*Methanobacterium thermoautotrophicum*、*P. furiosus*、*P. horikoshii*、*A. pernix*、*T. thermophilus*、*Halobacterium*、*Escherichia coli* 等) に由来するものとする。例えば植物 (例えば単子葉植物又は双子葉植物等の複雑な植物)、藻類、原生動物、真菌類 (例えば酵母等)、動物 (例えば哺乳動物、昆虫、節足動物等) 等の真核源も使用できる。 40

【0164】

第2の生物の *in vivo* 翻訳系で使用するための直交 tRNA-tRNA シンテターゼ対の選択方法も本発明に含まれる。本方法はマーカー遺伝子と、tRNA と、第1の生物から単離又は誘導されるアミノアシル-tRNA シンテターゼ (RS) を第2の生物に由来する第1組の細胞に導入する段階と、マーカー遺伝子と tRNA を第2の生物に由来する複製組の細胞に導入する段階と、第1組と複製組の細胞を選択物質の存在下に増殖させ、第1組では生存するが、複製組では生存することができない細胞を選択し、第2の生 50

物の *in vivo* 翻訳系で使用するための直交 tRNA-tRNA シンテターゼ対を含む生存細胞を提供する段階を含む。1 態様では、比較と選択は *in vivo* 相補アッセイを含む。別の態様では選択物質濃度を変動させる。

【0165】

例えば、シンテターゼに対する tRNA の認識部位であるアイデンティティエレメントが存在する場所に基づいて tRNA / シンテターゼ対を選択することができる。例えば、アイデンティティエレメントがアンチコドンの外側にある場合に tRNA / シンテターゼ対を選択する（例えば始原菌 *Methanococcus jannaschii* に由来する tRNA^{Tyr} / TyrRS 対）。この TyrRS はその tRNA^{Tyr} のアンチコドンループと結合する非保存ドメインの大半を欠失しているが、G1 : C72 をもつ tRNA から C1 : G72 をもつ tRNA を識別することができる。更に、*Methanococcus jannaschii* TyrRS (Mj TyrRS) は *Saccharomyces cerevisiae* をアミノアシル化するが、大腸菌天然 tRNA をアミノアシル化しない。例えば B. A. Steer と P. Schimmel, J. Biol. Chem., 274 : 35601 (1999) 参照。少なくとも 1 個のセクターコドンをもつレポーター遺伝子（例えば β ラクターゼ遺伝子）を含む発現ベクターを使用する *in vivo* 相補アッセイによると、*Methanococcus jannaschii* tRNA^{Tyr} CUA (Mj tRNA^{Tyr} CUA) のみを発現する細胞は $IC_{50} = 55 \mu g / mL$ アンピシリンまで生存し、Mj tRNA^{Tyr} CUA とその TyrRS を同時発現する細胞は $IC_{50} = 1220 \mu g / mL$ アンピシリンまで生存する。Mj tRNA^{Tyr} CUA は大腸菌で Sc tRNA^{Gln} CUA ($IC_{50} = 20 \mu g / mL$) よりも直交性が低い、Mj TyrRS はそのコグネイトアンバーサプレッサー tRNA に対するアミノアシル化活性が高い。例えば L. Wang, T. J. Magliery, D. R. Liu and P. G. Schultz, J. Am. Chem. Soc., 122 : 5010 (2000) 参照。その結果、*Methanococcus jannaschii* / TyrRS は大腸菌で直交対であるとみなされ、*in vivo* 翻訳系で使用するに選択することができる。

【0166】

非天然アミノ酸

【0167】

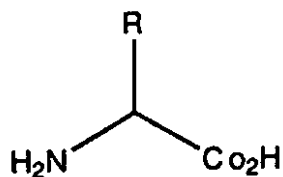
本発明の方法では多様な非天然アミノ酸を使用することができる。非天然アミノ酸は非天然アミノ酸の所望特性（例えば非天然アミノ酸の機能）に基づいて選択することができる、例えば毒性、生体分布又は半減期、構造的性質、分光学的性質、化学及び／又は光化学的性質、触媒性質、他の分子との（例えば共有又は非共有）反応性等のタンパク質の生物学的性質を改変する機能が挙げられる。

【0168】

本明細書において非天然アミノ酸とはセレノシステインと 20 種の遺伝的にコードされる α アミノ酸（即ちアラニン、アルギニン、アスパラギン、アスパラギン酸、システイン、グルタミン、グルタミン酸、グリシン、ヒスチジン、イソロイシン、ロイシン、リジン、メチオニン、フェニルアラニン、プロリン、セリン、スレオニン、トリプトファン、チロシン、バリン）以外の任意アミノ酸、改変アミノ酸又はアミノ酸類似体を意味する。 α アミノ酸の一般構造は式 I :

【化 1】

I



10

により表される。非天然アミノ酸は一般に式 I をもつ任意構造であり、式中、R 基は 20 種の天然アミノ酸で使用されている以外の任意置換基である。20 種の天然アミノ酸の構造については、例えば L. Stryer 著 Biochemistry, 第 3 版, 1988, Freeman and Company, New York を参照されたい。本発明の非天然アミノ酸は上記 20 種の α アミノ酸以外の天然化合物でもよいことに留意されたい。本発明の非天然アミノ酸は一般に側鎖のみが天然アミノ酸と異なるので、本発明の非天然アミノ酸は天然タンパク質で形成されると同様に他のアミノ酸（例えば天然又は非天然アミノ酸）とアミド結合を形成する。他方、非天然アミノ酸は天然アミノ酸と異なる側鎖基をもつ。例えば、式 I 中の R は場合によりアルキル、アリール、アシル、ケト、アジド、ヒドロキシル、ヒドラジン、シアノ、ハロ、ヒドラジド、アルケニル、アルキニル、エーテル、チオール、セレノ、スルホニル、硼酸、ボロン酸、ホスホ、ホスホノ、ホスフィン、複素環、エノン、イミン、アルデヒド、エステル、チオ酸、ヒドロキシルアミン、アミノ基等又はその任意組合せを含む。他の該当非天然アミノ酸としては、光架橋基を含むアミノ酸、スピラベル付きアミノ酸、蛍光アミノ酸、金属結合性アミノ酸、金属含有アミノ酸、放射性アミノ酸、新規官能基をもつアミノ酸、他の分子と共有又は非共有的に相互作用するアミノ酸、フォトケージド及び／又は光異性化可能なアミノ酸、ビオチン又はビオチン類似体含有アミノ酸、グリコシル化アミノ酸（例えば糖鎖置換セリン）、他の糖鎖修飾アミノ酸、ケト含有アミノ酸、ポリエチレングリコール又はポリエーテルを含むアミノ酸、重原子置換アミノ酸、化学分解性又は光分解性アミノ酸、天然アミノ酸に比較して延長側鎖（例えばポリエーテル又は例えば炭素長が約 5 もしくは約 10 を越える長鎖炭化水素）をもつアミノ酸、炭素結合糖含有アミノ酸、レドックス活性アミノ酸、アミノチオ酸含有アミノ酸、及び 1 個以上の毒性部分を含むアミノ酸が挙げられるが、これらに限定されない。

20

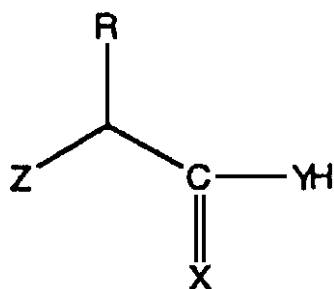
30

【0169】

新規側鎖を含む非天然アミノ酸に加え、非天然アミノ酸は場合により例えば式 I I 及び I I I :

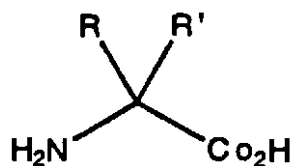
【化 2】

II



10

III



20

の構造により表されるような修飾主鎖構造を含み、式中、Zは一般にOH、NH₂、SH、NH-R'又はS-R'を含み、XとYは同一でも異なってもよく、一般にS又はOであり、RとR'は場合により同一又は異なり、一般に式Iをもつ非天然アミノ酸について上記に記載したR基と同一の基及び水素から選択される。例えば、本発明の非天然アミノ酸は場合により式I及びI'により表されるようにアミノ又はカルボキシル基に置換を含む。この種の非天然アミノ酸としては、例えば20種の共通天然アミノ酸に対応する側鎖又は非天然側鎖をもつα-ヒドロキシ酸、α-チオ酸、α-アミノチオカルボキシル酸が挙げられるが、これらに限定されない。更に、α-炭素の置換は場合によりL、D又はα、α-ジ置換アミノ酸（例えばD-グルタミン酸、D-アラニン、D-メチル-α-チロシン、アミノ酪酸等）を含む。他の構造としては環状アミノ酸（例えばプロリン類似体や、3、4、6、7、8及び9員環プロリン類似体）、β及びγアミノ酸（例えば置換β-アラニン及びγ-アミノ酪酸）が挙げられる。

30

【0170】

例えば、多数の非天然アミノ酸は天然アミノ酸（例えばチロシン、グルタミン、フェニルアラニン等）に由来する。チロシン類似体としてはパラ置換チロシン、オルト置換チロシン、及びメタ置換チロシンが挙げられ、置換チロシンはアセチル基、ベンゾイル基、アミノ基、ヒドラジン、ヒドロキシルアミン、チオール基、カルボキシ基、イソプロピル基、メチル基、C₆-C₂₀直鎖又は分枝鎖炭化水素、飽和又は不飽和炭化水素、O-メチル基、ポリエーテル基、ニトロ基等を含む。更に、多重置換アリール環も考えられる。本発明のグルタミン類似体としてはα-ヒドロキシ誘導体、γ置換誘導体、環状誘導体及びアミド置換グルタミン誘導体が挙げられるが、これらに限定されない。フェニルアラニン類似体の例としてはヒドロキシ基、メトキシ基、メチル基、アリル基、アセチル基等を置換基に含むメタ置換フェニルアラニンが挙げられるが、これらに限定されない。非天然アミノ酸の特定例としてはO-メチル-L-チロシン、L-3-(2-ナフチル)アラニン、3-メチルフェニルアラニン、O-4-アリル-L-チロシン、4-プロピル-L-チロシン、トリ-O-アセチル-GlcNAcβ-セリン、L-Dopa、フッ素化フェニルア

40

50

ラニン、イソプロピル-L-フェニルアラニン、p-アジド-L-フェニルアラニン、p-アシル-L-フェニルアラニン、p-ベンゾイル-L-フェニルアラニン、L-ホスホセリン、ホスホノセリン、ホスホノチロシン、p-ヨードフェニルアラニン、p-ブロモフェニルアラニン、p-アミノ-L-フェニルアラニン、及びイソプロピル-L-フェニルアラニン等が挙げられるが、これらに限定されない。各種非天然アミノ酸の構造を図（例えば図29、30及び31）に示す。

【0171】

一般に、本発明の非天然アミノ酸は20種の天然アミノ酸では得られない付加特性を提供するように選択又は設計される。例えば、非天然アミノ酸は場合により（例えばこれらのアミノ酸を組込む）タンパク質の生物学的性質を改変するように設計又は選択される。例えば、非天然アミノ酸をタンパク質に組込むことにより、場合により毒性、生体分布、溶解度、安定性（例えば熱、加水分解、酸化、酵素耐性、分解等）、精製と処理し易さ、構造的性質、分光学的性質、化学及び／又は光化学的性質、触媒活性、酸化還元電位、半減期、他の分子との（例えば共有又は非共有）反応性等の性質を改変する。

【0172】

非天然アミノ酸の更に詳細な説明は参考資料として本明細書に組込む2002年4月19日付け対応出願（発明の名称" *In vivo Incorporation of Unnatural Amino Acids*"、代理人整理番号54-000120PC/US）に記載されている。

【0173】

突然変異体 tRNA 及び O-RS 及び O-tRNA / O-RS 対の使用

【0174】

本発明の組成物と本発明の方法により生産された組成物は場合により細胞に組込む。その後、本発明の O-tRNA / O-RS 対又は個々の成分を宿主系の翻訳機構で使用し、非天然アミノ酸をタンパク質に組込むことができる。対応特許出願（発明の名称" *In vivo Incorporation of Unnatural Amino Acids*"、代理人整理番号54-000120PC/US、発明者Schultzら）はこの方法について記載しており、参考資料として本明細書に組込む。例えば、O-tRNA / O-RS 対を宿主（例えば大腸菌）に導入すると、対は外部から増殖培地に添加可能な非天然アミノ酸（例えばO-メチル-L-チロシン等の合成アミノ酸）をセクターコードン（例えばアンバーナンセンスコードン）に応答してタンパク質（例えばジヒドロ葉酸レダクターゼやEPO等の治療用タンパク質）に *in vivo* 組込むことができる。場合により、本発明の組成物を *in vitro* 翻訳系又は *in vivo* 系に組込むこともできる。

【0175】

核酸及びポリペプチド配列変異体

【0176】

上記及び下記に記載するように、本発明は核酸ポリヌクレオチド配列及びポリペプチドアミノ酸配列（例えばO-tRNA及びO-RS）と、例えば前記配列を含む組成物と方法を提供する。前記配列（例えばO-tRNA及びO-RS）の例を本明細書に開示する。しかし、当業者に自明の通り、本発明は本明細書（例えば実施例）に開示する配列に限定されない。当業者に自明の通り、本発明は本明細書に記載する機能をもつ（例えば（例えばO-tRNA又はO-RSをコードする））多数の非関連配列も提供する。

【0177】

同様に当業者に自明の通り、開示配列の多数の変異体も本発明に含まれる。例えば、機能的に同一配列である開示配列の保存変異も本発明に含まれる。少なくとも1種の開示配列にハイブリダイスする核酸ポリヌクレオチド配列の変異体も本発明に含むものとする。例えば標準配列比較法により本明細書に開示する配列のユニークサブ配列であるとみなされる配列も本発明に含まれる。

【0178】

10

20

30

40

50

保存変異

【0179】

遺伝コードの縮重により、「サイレント置換」（即ちコードされるポリペプチドに変化を生じない核酸配列の置換）はアミノ酸をコードする全核酸配列の暗黙の特徴である。同様に、「保存アミノ酸置換」はアミノ酸配列中の1又は数個のアミノ酸を高度に類似する特性をもつ別のアミノ酸で置換するものであり、このような置換も開示構築物と高度に類似するとは自明である。各開示配列のこのような保存変異は本発明の1つの特徴である。

【0180】

特定核酸配列の「保存変異」とは同一又は本質的に同一のアミノ酸配列をコードする核酸を意味し、核酸がアミノ酸配列をコードしない場合には本質的に同一の配列を意味する。当業者に自明の通り、コードされる配列中の単一アミノ酸又は低百分率（一般に5%未満、より一般には4%、2%又は1%未満）のアミノ酸を置換、付加又は欠失させる個々の置換、欠失又は付加の結果としてアミノ酸1個を欠失するか、アミノ酸1個が付加するか、又はアミノ酸1個が化学的に類似するアミノ酸1個で置換される場合には、これらの改変は「保存改変変異」である。従って、本発明に示すポリペプチド配列の「保存変異」はポリペプチド配列のアミノ酸の低百分率、一般に5%未満、より一般には2%又は1%未満を同一保存置換基の保存可能に選択したアミノ酸で置換する置換を含む。最後に、非機能的配列の付加のように核酸分子のコードされる活性を変えない配列の付加も基本核酸の保存変異である。

【0181】

表1 — 保存置換基

1	アラニン (A)	セリン (S)	スレオニン (T)	
2	アスパラギン酸 (D)	グルタミン酸 (E)		
3	アスパラギン (N)	グルタミン (Q)		
4	アルギニン (R)	リジン (K)		
5	イソロイシン (I)	ロイシン (L)	メチオニン (M)	バリン (V)
6	フェニルアラニン (F)	チロシン (Y)	トリプトファン (W)	

【0182】

核酸ハイブリダイゼーション

【0183】

比較ハイブリダイゼーションを使用して（本発明の核酸の保存変異を含む）本発明の核酸を同定することができ、この比較ハイブリダイゼーション法は本発明の核酸を識別する好適な方法である。更に、高、超高及び超々高ストリンジェンシー条件下で配列番号1～3又は配列番号4～34（表5参照）に示す核酸とハイブリダイズするターゲット核酸も本発明の1つの特徴である。このような核酸の例としては所定核酸配列と比較して1又は数個のサイレント又は保存核酸置換をもつものが挙げられる。

【0184】

完全にマッチする相補的ターゲットに比較して少なくとも1/2の割合でプローブにハイブリダイズする場合、即ちマッチしないターゲット核酸の任意のものとのハイブリダイゼーションに観測されるシグナル対ノイズ比の少なくとも約5倍～10倍で完全にマッチするプローブが完全にマッチする相補的ターゲットと結合する条件下におけるプローブとターゲットのハイブリダイゼーションに比較してシグナル対ノイズ比が少なくとも1/2である場合に試験核酸はプローブ核酸に特異的にハイブリダイズすると言う。

【0185】

核酸は一般に溶液中で会合するときに「ハイブリダイズ」する。核酸は水素結合、溶媒排除、塩基スタッキング等の種々の十分に特性決定された物理化学的力によりハイブリダイズする。核酸ハイブリダイゼーションの詳しい手引きは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) 及びAusubel, 前出に記載されている。HamesとHiggins (1995) *Gene Probes 1* IRL Press at Oxford University Press, Oxford, 英国 (HamesとHiggins 1) 及びHamesとHiggins (1995) *Gene Probes 2* IRL Press at Oxford University Press, Oxford, 英国 (HamesとHiggins 2) はDNAとRNA（オリゴヌクレオチドを含む）の合成、標識、検出及び定量について詳細に記載している。

【0186】

サザン又はノーザンブロットでフィルター上で100を越える相補的残基をもつ相補的核酸のハイブリダイゼーションを行うためのストリンジェントハイブリダイゼーション条件の1例は、50%ホルマリンにヘパリン1mgを加え、42℃で一晩のハイブリダイゼーションである。ストリンジェント洗浄条件の1例は65℃、0.2×SSCで15分間である（SSC緩衝液の説明についてはSambrook, 前出参照）。多くの場合には高ストリンジェンシー洗浄の前に低ストリンジェンシー洗浄でバックグラウンドプローブシグナルを除去する。低ストリンジェンシー洗浄の1例は40℃、2×SSCで15分間である。一般に、シグナル対ノイズ比が特定ハイブリダイゼーションアッセイで非関連プローブに観測されるよりも5倍（以上）である場合に特異的ハイブリダイゼーションが検出されたとみなす。

【0187】

サザン及びノーザンハイブリダイゼーション等の核酸ハイブリダイゼーション実験に関して「ストリンジェントハイブリダイゼーション洗浄条件」は配列依存的であり、種々の環境パラメーターにより異なる。核酸ハイブリダイゼーションの詳しい手引きはTijssen (1993), 前出やHamesとHiggins 1及び2に記載されている。ストリンジェントハイブリダイゼーション及び洗浄条件は任意試験核酸について経験により

容易に決定することができる。例えば、高ストリンジェントハイブリダイゼーション及び洗浄条件を決定するには、総合選択基準に合致するまで（例えばハイブリダイゼーション又は洗浄における温度上昇、塩濃度低下、界面活性剤濃度増加及び／又はホルマリン等の有機溶媒濃度増加により）ハイブリダイゼーション及び洗浄条件を漸増させる。例えば、マッチしないターゲットとプローブのハイブリダイゼーションに観測されるシグナル対ノイズ比の少なくとも約5倍でプローブが完全にマッチする相補的ターゲットと結合するまでハイブリダイゼーション及び洗浄条件を漸増させる。

【0188】

「超ストリンジェント」条件は特定プローブの熱融点（ T_m ）に等しくなるように選択する。 T_m は（規定イオン強度及びpH下で）試験配列の50%が完全にマッチするプローブとハイブリダイズする温度である。本発明の趣旨では、一般に規定イオン強度及びpHで特定配列の T_m よりも約5℃低くなるように「高ストリンジェント」ハイブリダイゼーション及び洗浄条件を選択する。

10

【0189】

「超高ストリンジェンシー」ハイブリダイゼーション及び洗浄条件は完全にマッチする相補的ターゲット核酸とプローブを結合させる場合のシグナル対ノイズ比がマッチしないターゲット核酸の任意のものとのハイブリダイゼーションに観測されるシグナル対ノイズ比の少なくとも10倍になるまでハイブリダイゼーション及び洗浄条件のストリンジェンシーを増す条件である。完全にマッチする相補的ターゲットのシグナル対ノイズ比の少なくとも1/2で前記条件下でプローブとハイブリダイズする場合にターゲット核酸は超高ストリンジェンシー条件下でプローブと結合すると言う。

20

【0190】

同様に、該当ハイブリダイゼーションアッセイのハイブリダイゼーション及び／又は洗浄条件を漸増させることにより更に高いレベルのストリンジェンシーを決定することもできる。例えば、完全にマッチする相補的ターゲット核酸とプローブを結合させる場合のシグナル対ノイズ比がマッチしないターゲット核酸の任意のものとのハイブリダイゼーションに観測されるシグナル対ノイズ比の少なくとも10倍、20倍、50倍、100倍又は500倍以上になるまでハイブリダイゼーション及び洗浄条件のストリンジェンシーを増す。完全にマッチする相補的ターゲットのシグナル対ノイズ比の少なくとも1/2で前記条件下でプローブとハイブリダイズする場合にターゲット核酸は超々高ストリンジェンシー条件下でプローブと結合すると言う。

30

【0191】

ストリンジェント条件下で相互にハイブリダイズしない核酸でも、これらの核酸によりコードされるポリペプチドが実質的に同一である場合には実質的に同一である。これは、例えば遺伝コードに許容される最大コドン縮重を使用して核酸のコピーを作製する場合に該当する。

【0192】

ユニークサブ配列

【0193】

1側面では、本発明は本明細書に開示するO-tRNA及びO-RSの配列（例えば配列番号1～3又は配列番号4～34（表5参照））から選択される核酸にユニークサブ配列を含む核酸を提供する。ユニークサブ配列は任意公知O-tRNA及びO-RS核酸配列に対応する核酸に比較してユニークである。例えばデフォルトパラメーターに設定したBLASTを使用してアラインメントを実施することができる。任意ユニークサブ配列は例えば本発明の核酸を同定するためのプローブとして有用である。

40

【0194】

同様に、本発明は本明細書に開示するO-RSの配列（例えば配列番号35～66（表5参照））から選択されるポリペプチド中にユニークサブ配列を含むポリペプチドを含む。この場合には、ユニークサブ配列は公知ポリペプチド配列の任意のものに対応するポリペプチドに比較してユニークである。

50

【0195】

本発明はO-RSの配列から選択されるポリペプチド中のユニークサブ配列をコードするユニークコーディングオリゴヌクレオチドとストリンジェント条件下でハイブリダイズするターゲット核酸も提供し、この場合、ユニークサブ配列は対照ポリペプチド（例えば本発明のシンテターゼを得るために例えば突然変異させた親配列）の任意のものに対応するポリペプチドに比較してユニークである。ユニーク配列は上記のように決定する。

【0196】

配列比較、一致度及び相同道

【0197】

2種以上の核酸又はポリペプチド配列に関して「一致」又は「一致度」百分率なる用語は2種以上の配列又はサブ配列を最大限に対応するように対比及び整列させ、以下に記載する配列比較アルゴリズム（又は当業者に入手可能な他のアルゴリズム）の1種を使用するか又は目視により測定した場合に相互に同一であるか又は同一のアミノ酸残基もしくはヌクレオチドの百分率が特定値であることを意味する。 10

【0198】

2種以上の核酸又はポリペプチド（例えばO-tRNAもしくはO-RS又はO-RSのアミノ酸配列をコードするDNA）に関して「実質的に一致」なる用語は2種以上の配列又はサブ配列を最大限に対応するように対比及び整列させ、配列比較アルゴリズムを使用するか又は目視により測定した場合にヌクレオチド又はアミノ酸残基一致度が少なくとも約60%、好ましくは80%、最も好ましくは90~95%であることを意味する。この 20
ような「実質的に一致」する配列は一般に実際の起源が記載されていなくても「相同」であるとみなす。少なくとも約50残基長の配列の領域にわたって「実質的一致」が存在することが好ましく、少なくとも約100残基長の配列の領域がより好ましく、少なくとも約150残基又は比較する2配列の全長にわたって配列が実質的に一致していることが最も好ましい。

【0199】

配列比較及び相同性決定には、一般に一方の配列を参照配列としてこれに試験配列を比較する。配列比較アルゴリズムを使用する場合には、試験配列と参照配列をコンピューターに入力し、必要に応じてサブ配列座標を指定し、配列アルゴリズムプログラムパラメーターを指定する。こうすると、配列比較アルゴリズムは指定プログラムパラメーターに基づ 30
いて参照配列に対して試験配列の配列一致度百分率を計算する。

【0200】

比較のための最適な配列アラインメントは例えばSmith & Waterman, Adv. Appl. Math. 2:482 (1981)の局所相同性アルゴリズム、Needleman & Wunsch, J. Mol. Biol. 48:443 (1970)の相同性アラインメントアルゴリズム、Pearson & Lipman, Proc. Nat'l

l Acad. Sci. USA 85:2444 (1988)の類似性探索法、これらのアルゴリズムのコンピューターソフトウェア(Wisconsin Genetics Software Package, Genetics Computer Group 40
, 575 Science Dr., Madison, WIのGAP、BESTFIT、FASTA及びTFASTA)、又は目視（一般にAusubelら、後出参照）により実施することができる。

【0201】

配列一致度及び配列類似度百分率を決定するのに適したアルゴリズムの1例はAltschulら, J. Mol. Biol. 215:403-410 (1990)に記載されているBLASTアルゴリズムである。BLAST分析を実施するためのソフトウェアはNational Center for Biotechnology Information (www.ncbi.nlm.nih.gov/)から公に入手可能である。このアルゴリズムはデータベース配列中の同一長さの単語と整列した場合に所定の正の閾値ス 50

コア T と一致するか又はこれを満足するクエリー配列中の長さ W の短い単語を識別することによりまず高スコア配列対 (H S P) を識別する。T を隣接単語スコア閾値と言う (A l t s c h u l ら, 前出)。これらの初期隣接単語ヒットをシードとして検索を開始し、これらの単語を含むもっと長い H S P を探索する。次に、累積アラインメントスコアを増加できる限り、単語ヒットを各配列に沿って両方向に延長する。ヌクレオチド配列の場合にはパラメーター M (1 対のマッチ残基のリワードスコア、常に > 0) 及び N (ミスマッチ残基のペナルティスコア、常に < 0) を使用して累積スコアを計算する。アミノ酸配列の場合には、スコアリングマトリクスを使用して累積スコアを計算する。累積アラインメントスコアがその最大到達値から量 X だけ低下するか、累積スコアが 1 カ所以上の負スコア残基アラインメントの累積によりゼロ以下になるか、又はどちらかの配列の末端に達したら各方向の単語ヒットの延長を停止する。B L A S T アルゴリズムパラメーター W、T 及び X はアラインメントの感度と速度を決定する。B L A S T I N プログラム (ヌクレオチド配列用) は語長 (W) 11、期待値 (E) 10、カットオフ 100、M = 5、N = 4、及び両鎖の比較をデフォルトとして使用する。アミノ酸配列用として、B L A S T P プログラムは語長 (W) 3、期待値 (E) 10、及び B L O S U M 62 スコアリングマトリクスをデフォルトとして使用する (H e n i k o f f & H e n i k o f f (1989) Proc. Natl. Acad. Sci. USA 89:10915 参照)。

10

【0202】

配列一致度百分率の計算に加え、B L A S T アルゴリズムは 2 配列間の類似性の統計分析も実施する (例えば K a r l i n & A l t s c h u l, Proc. Natl. Acad. Sci. USA 90:5873-5787 (1993) 参照)。B L A S T アルゴリズムにより提供される類似性の 1 尺度は 2 種のヌクレオチド又はアミノ酸配列間に偶然にマッチが起こる確率を示す最小合計確率 (P (N)) である。例えば、試験核酸を参照核酸に比較した場合の最小合計確率が約 0.1 未満、より好ましくは約 0.01 未満、最も好ましくは約 0.001 未満である場合に核酸は参照核酸に類似しているとみなす。

20

【0203】

免疫反応性によるポリペプチドの特性決定

【0204】

本発明のポリペプチドは (例えば本発明の翻訳系で合成されるタンパク質の場合には非天然アミノ酸を含み、例えば本発明の新規シンテターゼの場合には標準アミノ酸の新規配列を含む) 種々の新規ポリペプチドを提供するので、ポリペプチドは例えばイムノアッセイで認識することができる新規構造特徴も提供する。本発明のポリペプチドに特異的に結合する抗血清の作製とこのような抗血清と結合するポリペプチドも本発明の特徴の 1 つである。

30

【0205】

例えば、本発明は配列番号 35~66 (表 5 参照) の 1 種以上から選択されるアミノ酸配列を含む免疫原に対して作製した抗体又は抗血清に特異的に結合するか又は特異的に免疫反応性であるシンテターゼタンパク質を含む。他の相同体との交差反応性をなくすために、野生型 M e t h a n o c o c c u s j a n n a s c h i i (M. j a n a s c h i i) チロシルシンテターゼ (T y r R S) 等の入手可能なシンテターゼ (例えば「コントロール」ポリペプチド) を用いて抗体又は抗血清をサブトラクションする。野生型 M e t h a n o c o c c u s j a n n a s c h i i (M. j a n a s c h i i) チロシルシンテターゼ (T y r R S) が核酸に対応する場合には、核酸によりコードされるポリペプチドを作製し、抗体/抗血清サブトラクション目的に使用する。

40

【0206】

典型的なフォーマットの 1 例では、イムノアッセイは配列番号 35~66 (表 5 参照) の 1 種以上に対応する配列の 1 種以上又はその実質的サブ配列 (即ち記載する全長配列の少なくとも約 30%) を含む 1 種以上のポリペプチドに対して作製したポリクローナル抗血清を使用する。配列番号 35~66 (表 5 参照) に由来する潜在的ポリペプチド免疫原群

50

を以下の文中では「免疫原性ポリペプチド」と総称する。得られた抗血清を場合により対照シンテターゼ相同体に対する交差反応性が低くなるように選択し、ポリクローナル抗血清をイムノアッセイで使用する前に例えば対照シンテターゼ相同体の1種以上に免疫吸着させることによりこのような交差反応性を除去する。

【0207】

イムノアッセイで使用する抗血清を作製するためには、免疫原性ポリペプチドの1種以上を本明細書に記載するように作製し、精製する。例えば、組換えタンパク質を組換え細胞で生産することができる。標準マウス免疫プロトコールに従って近交系マウス（マウスの仮想遺伝子一致により結果の再現性が高いのでこのアッセイで使用）に標準アジュバント（例えばフロイントアジュバント）と共に免疫原性タンパク質を免疫する（抗体作製、イムノアッセイフォーマット及び特異的免疫反応性を測定するために使用可能な条件の標準解説については、例えばHarlowとLane（1988）Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York参照。本明細書に示した他の抗体関連文献もこの場合の免疫反応性によるポリペプチドの特性決定に適用することができる）。あるいは、本明細書に開示する配列から誘導される1種以上の合成又は組換えポリペプチドをキャリアタンパク質に結合させて免疫原として使用してもよい。

10

【0208】

ポリクローナル血清を収集し、イムノアッセイ（例えば固体支持体に固定化した免疫原性タンパク質の1種以上を使用する固相イムノアッセイ）で免疫原性ポリペプチドに対する力価を測定する。 10^6 以上の力価をもつポリクローナル抗血清を選択し、プールし、対照シンテターゼポリペプチドでサブトラクションし、高力価ポリクローナル抗血清サブトラクションプールを作製する。

20

【0209】

高力価ポリクローナル抗血清サブトラクションプールを比較イムノアッセイで対照相同体に対する交差反応性について試験する。この比較アッセイでは、高力価ポリクローナル抗血清が免疫原性シンテターゼと結合する場合のシグナル対ノイズ比が対照シンテターゼ相同体と結合する場合の少なくとも約5～10倍となるように、サブトラクション高力価ポリクローナル抗血清に差別的な結合条件を決定する。即ち、アルブミン又は脱脂粉乳等の非特異的競合剤を加えるか及び／又は塩条件、温度及び／又はその他を調節することにより結合反応のストリンジェンシーを調整する。これらの結合条件は、試験ポリペプチド（免疫原性ポリペプチド及び／又は対照ポリペプチドに比較するポリペプチド）がサブトラクションポリクローナル抗血清プールに特異的に結合するか否かを調べる後期アッセイで使用される。特に、差別的結合条件下で対照シンテターゼ相同体の少なくとも2～5倍のシグナル対ノイズ比と、免疫原性ポリペプチドの少なくとも約1.5倍のシグナル対ノイズ比を示す試験ポリペプチドは公知シンテターゼに比較して免疫原ポリペプチドと実質的な構造類似性をもつので、本発明のポリペプチドである。

30

【0210】

別の例では、試験ポリペプチドの検出に競合結合フォーマットのイムノアッセイを使用する。例えば、自明の通り、対照ポリペプチドに免疫吸着させることにより抗血清混合物プールから交差反応性抗体を除去する。次に、免疫原性ポリペプチドを固体支持体に固定化し、支持体をサブトラクション抗血清プールに暴露する。試験タンパク質をアッセイに加え、サブトラクション抗血清プールとの結合を競合させる。試験タンパク質が固定化タンパク質に対してサブトラクション抗血清プールとの結合を競合する能力と、アッセイに加えた免疫原性ポリペプチドが結合を競合する能力とを比較する（免疫原性ポリペプチドは抗血清プールとの結合を固定化免疫原性ポリペプチドと有効に競合する）。標準計算を使用して試験タンパク質の交差反応性百分率を計算する。

40

【0211】

平行アッセイで場合により対照タンパク質がサブトラクション抗血清プールとの結合を競合する能力を免疫原性ポリペプチドが抗血清との結合を競合する能力と比較測定する。こ

50

の場合も、標準計算を使用して対照ポリペプチドの交差反応性百分率を計算する。試験ポリペプチドの交差反応性百分率が対照ポリペプチドの少なくとも5～10倍である場合又は試験ポリペプチドの結合が免疫原性ポリペプチドの結合とほぼ同一範囲である場合に、試験ポリペプチドはサブトラクション抗血清プールに特異的に結合すると言う。

【0212】

一般に、本明細書に記載する競合的結合イムノアッセイでは任意試験ポリペプチドを免疫原性及び／又は対照ポリペプチドに比較するために免疫吸着抗血清プールを使用することができる。この比較を行うためには、免疫原性、試験及び対照ポリペプチドを各々広い濃度範囲でアッセイし、サブトラクション抗血清と例えば固定化した対照、試験又は免疫原性タンパク質との結合の50%を阻害するために必要な各ポリペプチドの量を標準技術により決定する。競合アッセイで結合に必要な試験ポリペプチドの量が免疫原性ポリペプチドの必要量の2倍未満であり、対照ポリペプチドの少なくとも約5～10倍である場合には試験ポリペプチドは免疫原性タンパク質に対して作製した抗体に特異的に結合すると言う。

10

【0213】

特異性の付加試験として、得られる免疫原性ポリペプチドサブトラクション抗血清プールと免疫吸着に使用する免疫原性ポリペプチドの結合が殆ど又は全く検出できなくなるまで場合により抗血清プールを（対照ポリペプチドではなく）免疫原性ポリペプチドに完全に免疫吸着させる。この完全に免疫吸着した抗血清を次に試験ポリペプチドとの反応性について試験する。反応性が殆ど又は全く観察されない場合（即ち完全に免疫吸着した抗血清と免疫原性ポリペプチドの結合に観察されるシグナル対ノイズ比の2倍以下）には、試験ポリペプチドは免疫原性タンパク質により誘導される抗血清に特異的に結合する。

20

【0214】

一般技術

【0215】

分子生物学技術について記載している一般教科書としてはBergerとKimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA (Berger); Sambrookら, Molecular Cloning—A Laboratory Manual (第3版), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2000 (「Sambrook」); 及び Current Protocols in Molecular Biology, F. M. Ausubelら, 編, Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (2002年補遺) (「Ausubel」) が挙げられる。これらの教科書は突然変異誘発、ベクターの使用、プロモーター及び他の多くの関連事項について記載しており、例えば直交tRNA、直交シンテターゼ及びその対の作製について記載している。

30

40

【0216】

例えば新規シンテターゼ又はtRNAを生産するために、本発明では各種突然変異誘発を使用する。これらの方法としては部位特異的、ランダム点突然変異誘発、相同組換え（DNAシャフリング）、ウラシル含有鋳型を使用する突然変異誘発、オリゴヌクレオチド誘導突然変異誘発、ホスホロチオエート修飾DNA突然変異誘発、ギャップデュプレクスDNAを使用する突然変異誘発等が挙げられるが、これらに限定されない。他の利用可能な方法としては点ミスマッチ修復、修復欠損宿主株を使用する突然変異誘発、制限一選択及び制限一精製、欠失突然変異誘発、完全遺伝子合成による突然変異誘発、2本鎖切断修復等が挙げられる。例えばキメラ構築物を使用する突然変異誘発も本発明に含まれる。1態様では、天然分子又は改変もしくは突然変異させた天然分子の公知情報（例えば配列、配

50

列比較、物性、結晶構造等)により突然変異誘発を実施することができる。

【0217】

本明細書に記載する上記教科書と実施例はこれらの手順について記載している。以下の刊
行物とその引用文献にも情報が記載されている。Sieberら, Nature Biotechnology, 19:456-460 (2001); Lingら, Approaches to DNA mutagenesis: an overview, Anal Biochem. 254 (2):157-178 (1997); Daleら, Oligonucleotide-directed random mutagenesis using the phosphorothioate method, Methods Mol. Biol. 57:369-374 (1996); I. A. Lorimer 10
, I. Pastan, Nucleic Acids Res. 23, 3067-8 (1995); W. P. C. Stemmer, Nature 370, 389-91 (1994); Arnold, Protein engineering for unusual environments, Current Opinion in Biotechnology 4:450-451 (1993); Bassら, Mutant Trp repressors with new DNA-binding specificities, Science 242:240-245 (1988); Fritzら, Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro, Nucl. Acids Res. 16:6987-6999 (1988); Kramerら, Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations, Nucl. Acids Res 16:7207 (1988); Sakama 20
r & Khorana, Total synthesis and expression of a gene for the α -subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin), Nucl. Acids Res. 14:6361-6372 (1988); Sayersら, Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis, Nucl. Acids Res. 16:791-802 (1988); Sayersら, 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; Carter, Improved oligonucleotide-directed mutagenesis using M13 vectors, Methods in Enzymol. 154:382-403 (1987); Kramer & Fritz 40
Oligonucleotide-directed construction of mutations via gapped duplex DNA, Methods in Enzymol. 154:350-367 (1987); Kunkel, The efficiency of oligonucleotide directed mutagenesis, in Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D. M. J. 編, Springer Verlag, Berlin) (1987); Kunkelら, Rapid and efficient site-specific mutagenesis without phenotypic selection, Methods in Enzymol. 154, 367-382 (1987); Zoller & 50

Smith, Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template, Methods in Enzymol. 154:329-350 (1987); Carter, Site-directed mutagenesis, Biochem. J. 237:1-7 (1986); Eghtedarzadeh & Henikoff, Use of oligonucleotides to generate large deletions, Nucl. Acids Res. 14:5115 (1986); 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); Nakamaye & Eckstein, Inhibition of restriction endonuclease NciI cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis, Nucl. Acids Res. 14:9679-9698 (1986); Wells⁵, Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin, Phil. Trans. R. Soc. Lond. A 317:415-423 (1986); Botstein & Shortle, Strategies and applications of in vitro mutagenesis, Science 229:1193-1201 (1985); Carter⁵, Improved oligonucleotide site-directed mutagenesis using M13 vectors, Nucl. Acids Res. 13:4431-4443 (1985); Grundstrom⁵, Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis, Nucl. Acids Res. 13:3305-3316 (1985); Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection, Proc. Natl. Acad. Sci. USA 82:488-492 (1985); Smith, In vivo mutagenesis, Ann. Rev. Genet. 19:423-462 (1985); Taylor⁵, The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA, Nucl. Acids Res. 13:8749-8764 (1985); Taylor⁵, The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA, Nucl. Acids Res. 13:8765-8787 (1985); Wells⁵, Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites, Gene 34:315-323 (1985); Kramer⁵, The gapped duplex DNA approach to oligonucleotide-directed mutation construction, Nucl. Acids Res. 12:9441-9456 (1984); Kramer⁵, Point Mismatch Repair, Cell 38:879-887 (1984); Nambiar⁵, Total synthesis and cloning of a gene codi

ng for the ribonuclease S protein, Science 223:1299-1301 (1984); Zoller & Smith, Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, Methods in Enzymol. 100:468-500 (1983); 及び 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) 10
)。上記方法の多くに関する更に詳細な説明は Methods in Enzymology Volume 154 に記載されており、この文献には各種突然変異誘発法に伴うトラブルシューティング問題の有用な解決方法も記載されている。

【0218】

例えば本発明の突然変異誘発（例えばシンテターゼのライブラリーの突然変異又は tRNA の改変）に使用するオリゴヌクレオチドは一般に Beaucage と Caruthers, Tetrahedron Letts. 22(20):1859-1862 (1981) により記載されている固相ホスホロアミダイトリエステル法に従い、例えば Nee dham-VanDevanter ら, Nucleic Acids Res., 12: 6159-6168 (1984) に記載されているような自動合成器を使用して化学的に 20
合成される。

【0219】

更に、The Midland Certified Reagent Company (mcrc@oligos.com)、The Great American Gene Company (www.genco.com)、ExpressGen Inc. (www.expressgen.com)、Operon Technologies Inc. (Alameda, CA)、その他多数の各種販売会社からほぼ任意核酸（及び標準又は標準外を問わずほぼ任意標識核酸）をオーダーメイド又は標準注文することができる。

【0220】

本発明は直交 tRNA/R S 対を介する非天然アミノ酸の in vivo 組込みに用いる 30
宿主細胞及び生物にも関する。例えばクローニングベクター又は発現ベクターとすることができる本発明のベクターで宿主細胞を遺伝子組換え（例えば形質転換、形質導入又はトランスフェクション）する。ベクターは例えばプラスミド、細菌、ウイルス、裸のポリヌクレオチド又はポリヌクレオチドコンジュゲートの形態とすることができる。ベクターはエレクトロポレーション（From ら, Proc. Natl. Acad. Sci. USA 82, 5824 (1985)）、ウイルスベクターによる感染、核酸を小ビーズもしくは粒子のマトリックスに埋込むか又は表面に付着させて小粒子形態で高速射入する（Klein ら, Nature 327, 70-73 (1987)）等の標準方法により細胞及び 40
／又は微生物に導入する。Berger、Sambrook 及び Ausubel には各種の適当な形質転換法が記載されている。

【0221】

組換え宿主細胞は例えばスクリーニング段階、プロモーター活性化又は形質転換細胞選択等の操作に合うように適宜改変した慣用栄養培地で培養することができる。これらの細胞を場合によりトランスジェニック生物で培養することができる。

【0222】

（例えば後期核酸分離のための）例えば細胞分離及び培養に関する他の有用な文献としては Freshney (1994) Culture of Animal Cells, a Manual of Basic Technique, 第3版, Wiley-Liss, New York とその引用文献; Payne ら (1992) Plant Cell 50

and Tissue Culture in Liquid Systems John Wiley & Sons, Inc, New York, NY; Gamburg & Phillips (編) (1995) Plant Cell, Tissue and Organ Culture; Fundamental Method Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York; 及び Atlas & Parks (編) The Handbook of Microbiological Media (1993) CRC Press, Boca Raton, FL が挙げられる。

【0223】

ターゲット核酸を宿主細胞に導入する方法としては数種の周知方法が入手可能であり、本発明ではその任意のものを使用することができる。これらの方法としては、DNAを含む細菌プロトプラストとレシピエント細胞の融合、エレクトロポレーション、遺伝子銃及びウイルスベクターによる感染等が挙げられる。本発明のDNA構築物を含むプラスミドの数を増幅するためには細菌細胞を使用することができる。対数期まで細菌細胞を増殖させると、細菌内のプラスミドを当分野で公知の各種方法により分離することができる（例えば Sambrook 参照）。更に、細菌からプラスミドを精製するために多数のキットが市販されている（例えば Easy Prep（登録商標）、Flexi Prep（登録商標）（いずれも Pharmacia Biotech）；Strata Clean（登録商標）（Stratagene）；及び QIAprep（登録商標）（Qiagen））。分離精製したプラスミドを更に操作して他のプラスミドを作製し、細胞をトランスフェクトするために使用するか又は生物に感染させるために関連ベクターに組込む。典型的ベクターは転写及び翻訳ターミネーターと、転写及び翻訳開始配列と、特定ターゲット核酸の発現の調節に有用なプロモーターを含む。ベクターは場合により少なくとも1個の独立ターミネーター配列と、真核生物又は原核生物又は両者（例えばシャトルベクター）でカセットの複製を可能にする配列と、原核系と真核系の両者の選択マーカーを含む包括的発現カセットを含む。ベクターは原核生物、真核生物、又は好ましくは両者での複製と組みに適している。Giliman & Smith, Gene 8:81 (1979)；Robertsら, Nature, 328:731 (1987)；Schneider, B. ら, Protein Expr. Purif. 6435:10 (1995)；Ausubel, Sambrook, Berger (いずれも前出) 参照。クローニングに有用な細菌とバクテリオファージのカタログは例えば ATCC から入手でき、例えば ATCC から刊行された The ATCC Catalogue of Bacteria and Bacteriophage (1992) Gherna ら (編) が挙げられる。その他のシーケンシング、クローニング及び分子生物学の他の側面の基本手順と基礎理論事項も Watson ら (1992) Recombinant DNA 第2版, Scientific American Books, NY に記載されている。

【0224】

以下、実施例により本発明を説明するが、これらの実施例により発明を限定するものではない。

【実施例1】

【0225】

実施例1—*Methanococcus jannaschii* 由来 tRNA の直交性の改善

【0226】

タンパク質翻訳の忠実度を高くするために必要な tRNA-シンテターゼ相互作用の複雑な性質により、直交 tRNA-シンテターゼ対の合理的設計は困難である。本実施例は所定の異種間 tRNA-シンテターゼ対の交差認識が弱いことを利用し、高い直交性をもつ tRNA の後期 *in vivo* 進化と組合せる方法に関する。L. Wang と P. G. Schultze, Chem. Biol., 8:883 (2001) も参照されたい。具体的には、*Methanococcus jannaschii* tRNA^{Tyr} に由来する

アンバーサプレッサー tRNA のライブラリーを作製した。バルナーゼ遺伝子におけるアンバーナンセンス突然変異の抑圧に基づくネガティブ選択により、内在大腸菌アミノアシル tRNA シンテターゼの基質である tRNA TyrCUA をプールから除去した。次に、残りの tRNA TyrCUA のうちでコグネイト *Methanococcus jannaschii* チロシル tRNA シンテターゼ (TyrRS) の存在下に βラクタマーゼ遺伝子のアンバーナンセンスコドンを抑圧できるものを選択した。*Methanococcus jannaschii* tRNA TyrCUA よりも弱い大腸菌シンテターゼの基質であるが、効率的に *Methanococcus jannaschii* TyrRS により負荷することができる 4 個の突然変異体サプレッサー tRNA を選択した。突然変異体サプレッサー tRNA TyrCUA は *Methanococcus jannaschii* TyrRS と共に非天然アミノ酸の *in vivo* タンパク質組込みに有用な直交 tRNA シンテターゼ対を提供する。 10

【0227】

始原菌である *Methanococcus jannaschii* の tRNA Tyr は大腸菌 tRNA Tyr と異なるアイデンティティーエレメントをもつ。特に、大腸菌 tRNA Tyr はアクセプターステムに G1C72 対をもつが、*Methanococcus jannaschii* tRNA Tyr は C1G72 対をもつ。*Methanococcus jannaschii* tRNA Tyr に由来するアンバーサプレッサー tRNA は効率的に大腸菌シンテターゼによりアミノアシル化されないが、大腸菌でタンパク質翻訳に効率的に機能することが判明した。例えば L. Wang, T. J. Magliery, D. R. Liu, P. G. Schultz, A new functional suppressor tRNA/aminoacyl-tRNA synthetase pair for the *in vivo* incorporation of unnatural amino acids into proteins. J. Am. Chem. Soc., 122:5010-5011 (2000) 参照。更に、最低限のアンチコドンループ結合ドメインしかもたない *Methanococcus jannaschii* TyrRS は大腸菌 tRNA をアミノアシル化しないが、それ自体のサプレッサー tRNA TyrCUA を効率的にアミノアシル化する。例えば B. A. Steer, P. Schimmel, Major anticodon-binding region missing from an archaeobacterial tRNA synthetase, J. Biol. Chem. 274 (1999) 35601-35606; 及び Wang ら, (2000), 前出参照。 20 30

【0228】

大腸菌でこのサプレッサー tRNA の直交性を試験するために、βラクタマーゼの許容部位 (Ala184) にアンバーコドンを導入した。例えば D. R. Liu, P. G. Schultz, Progress toward the evolution of an organism with an expanded genetic code. Proc. Natl. Acad. Sci. USA 96:4780-4785 (1999) 参照。大腸菌シンテターゼにより負荷することができる tRNA はアンバーコドンを抑圧し、アンピシリンの存在下で細胞を生存させる。*Methanococcus jannaschii* tRNA TyrCUA は IC₅₀ 値 56 μg/ml アンピシリンで βラクタマーゼ遺伝子内のアンバーコドンを抑圧する。Wang ら, (2000), 前出参照。これに対して、*Saccharomyces cerevisiae* tRNA Gln2 に由来する直交 tRNA GlnCUA を同一アッセイで試験すると IC₅₀ は 21 μg/ml アンピシリンである。Liu & Schultz, (1999), 前出参照。サプレッサー tRNA の不在下の大腸菌の IC₅₀ は 10 μg/ml アンピシリンである。この結果は *Methanococcus jannaschii* tRNA TyrCUA が tRNA GlnCUA よりも良好な大腸菌シンテターゼの基質であることを示している。従って、非天然アミノ酸を大腸菌中のタンパク質に送達するために *Methanococcus jannaschii* tRNA TyrCUA を *in vivo* 使用するならば 40 50

、天然アミノ酸も大腸菌シンテターゼにより誤って負荷される可能性があり、異種アミノ酸組込みを生じる恐れがある。

【0229】

内在大腸菌シンテターゼによる認識を防ぐために「ネガティブ認識決定基」の導入により *Methanococcus jannaschii* tRNA^{Tyr}CUAの直交性の改善を行った。これらの突然変異はtRNAとそのコグネイト*Methanococcus jannaschii* TyrRS又はリボソームとの相互作用を強く妨害するものであってはならない。*Methanococcus jannaschii* TyrRSはアンチコドン結合ドメインの大半を欠失しているので（例えばB. A. Steer, P. Schimmel, Major anticodon-binding region missing from an archaeobacterial tRNA synthetase, J. Biol. Chem. 274:35601-35606 (1999) 参照）、tRNAのアンチコドンループに導入した突然変異はTyrRS認識に最小の影響しかないと予想される。4個のランダム突然変異ヌクレオチドをもつアンチコドンループライブラリーを構築した。図9参照。各種大腸菌tRNAのアイデンティティーエレメントの組合せと位置は多様であるので、付加位置に突然変異を導入すると、所望特性をもつ突然変異体tRNAをみつける確率を増すことができる。そこで、全tRNAループの非保存位置に突然変異を含む第2のライブラリー（全ループライブラリー）も構築した。図9参照。保存ヌクレオチドはランダム突然変異させず、tRNAの「L」字形構造を安定化する三次相互作用を維持した。例えばG. Dirheimer, G. Keith, P. Dumas, E. Westhof, Primary, secondary, and tertiary structures of tRNAs, in: D. Soll, U. L. RajBhandary (編), tRNA Structure, Biosynthesis, and Function, ASM Press, Washington, DC, 1995, pp. 93-126; 及びR. Giege, M. Sissler, C. Florentz, Universal rules and idiosyncratic features in tRNA identity, Nucleic Acids Res. 26:5017-5035 (1998) 参照。ステムヌクレオチドを1個置換するには補償突然変異が必要であるのでステムヌクレオチドも突然変異させなかった。11個のヌクレオチド(C16, C17, U17a, U20, C32, G37, A38, U45, U47, A59及びU60)をランダム突然変異させた。図9参照。このライブラリーの理論寸法は約 4.19×10^6 であり、突然変異体ライブラリーを完全にカバーするように約 1.93×10^8 コロニー形成単位の寸法のライブラリーを構築した。

【0230】

本方法ではGibco/BRLから入手した大腸菌株（例えばDH10B）を使用した。サプレッサーtRNA発現プラスミドはプラスミド（例えばpAC123）から誘導した。例えばD. R. Liu, T. J. Magliery, M. Pasternak, P. G. Schultz, Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo, Proc. Natl. Acad. Sci. U.S.A. 94:10091-10097 (1997) 参照。ネガティブ選択用プラスミドは後述するようなプラスミド（例えばpBATS、pYsupA38B2及びpYsupA38B3）から誘導した。例えばK. Gabriel, W. H. McClain, A set of plasmids constitutively producing different RNA levels in *Escherichia coli*, J. Mol. Biol. 290 (1999) 385-389; 及びLiu & Schultz, (1999), 前出参照。

【0231】

直交性が増した*Methanococcus jannaschii* tRNAライブラリー 50

リーのメンバーを選択するために、コグネイトシンテターゼの不在下と存在下のネガティブ選択とポジティブ選択の組合せを使用した。図8参照。ネガティブ選択では、セレクトターコドン（例えばアンバーナンセンス）をネガティブマーカー遺伝子（例えば毒性遺伝子）の例えば非必須位置に導入する。突然変異（例えばサプレッサー）tRNAライブラリーのメンバーが内在（例えば大腸菌）シンテターゼによりアミノアシル化される（即ち大腸菌シンテターゼに対して直交性でない）ならば、セレクトターコドンは抑圧され、毒性遺伝子産物が生産され、細胞死に至る。直交tRNA又は非機能的tRNAをもつ細胞のみが生存することができる。その後、全生存細胞をポジティブ選択し、セレクトターコドン（例えばアンバーコドン）をポジティブ選択マーカー（例えば薬剤耐性遺伝子）の例えば非必須位置に挿入する。その後、同時発現したコグネイトシンテターゼによりアミノアシル化され、このセレクトターコドンに応答してアミノ酸を挿入することができるtRNAを選択する。非機能的tRNA又は所期シンテターゼにより認識することができないtRNAを含む細胞は抗生物質に感受性である。従って、（1）内在大腸菌シンテターゼの基質でなく、（2）所期シンテターゼによりアミノアシル化することができ、（3）翻訳能力のあるtRNAのみが両者選択後に生存している。

【0232】

その天然阻害剤であるバルスターの不在下に大腸菌で生産した場合のバルナーゼの毒性を利用したネガティブ選択を採用した。例えばR. W. Hartley, Barnase and barstar. Expression of its cloned inhibitor permits expression of a cloned ribonuclease, J. Mol. Biol. 202:913-915 (1988) 参照。バルナーゼの三次元構造の分析に基づいてバルナーゼ遺伝子の非必須位置にアンバーコドンを導入した。例えばLiu & Schultz, (1999), 前出参照。バルナーゼは自家毒性が非常に強いので、バルナーゼを発現するプラスミドに低コピー数pSC101起点を導入した。更に、選択のストリンジェンシーを調節するためにアンバーコドンの数を変えて試験した。プラスミドpSCB2を使用し、Gln2とAsp44に2個のアンバー終止コドンをもつバルナーゼ突然変異体を発現させ、プラスミドpSCB3には更にGly65にもアンバー終止コドンを付加した。

【0233】

ネガティブ選択には、以下のオリゴヌクレオチドLW115: 5' - ATGCA TGC TGC A T T A A T G A A T C G G C C A A C G - 3' 及びLW116: 5' - T C C C C G C G G A G G T G G C A C T T T T C G G G G - 3' を使用してpBATSからβラクタマーゼ遺伝子とpSC101起点を含むPCRフラグメントを作製した。SacIIとSphIで消化することにより夫々pYsupA38B2とpYsupA38B3から2個（残基Gln2及びAsp44）又は3個（残基Gln2、Asp44及びGly65）のアンバーコドンを含むバルナーゼをコードするDNAを得た。上記フラグメントをライゲートし、プラスミドpSCB2及びpSCB3を得た。バルナーゼの発現はアラビノースで誘導した。in vivo発現のために各種サプレッサーtRNAをコードする遺伝子を2個のオーバーラップする合成オリゴヌクレオチド（Operon, Alameda, CA, 米国）からKlenow伸長により構築し、pAC123のEcoRI-PstI部位間に挿入し、lppプロモーターとrrnCターミネーターの制御下に転写するpAC-YYG1とpAC-JYを夫々作製した。pAC-CmはtRNAをもたない対照プラスミドである。ネガティブ選択条件を最適化するために、pSCB2又はpSCB3を含むコンピテントDH10B細胞をエレクトロポレーションによりpAC-Cm、pAC-YYG1及びpAC-JYで別々に形質転換した。単コロニーを釣菌し、クロラムフェノール（Cm, 34 µg/ml）とアンピシリン（Amp, 100 µg/ml）を加えた2×YTで増殖させた。一晚増殖させた細胞培養液を1%グリセロールと0.3 mMロイシンを加えた最少培地（GMM L）で2回洗浄し、CmとAmpを加えたGMM LにOD600=0.1まで再懸濁した。30℃で10分間再生後に第1の培養液（第1組）に20 mMアラビノースを加えてバルナーゼの発現を誘導し、第2の培養液（第2組）

にはアラビノースを加えなかった。各時点で少量の培養液を希釈し、CmとAmpを加えた2×YT寒天に撒き、細胞密度を測定した。サプレッサーtRNAライブラリーのネガティブ選択のために、pSCB2を含むDH10B細胞にライブラリーを含むpACプラスミドを形質転換した。SOC培地を加えて細胞をクエンチし、30℃で1時間再生した後、リン酸緩衝液とGMM Lで洗浄し、GMM L 1リットル中で培養した。30℃で30分間再生後にCm、Amp及び20mMアラビノースを加えた。36時間後に細胞をペレット化し、pACプラスミドを単離し、アガロースゲル電気泳動により精製した。

【0234】

選択条件を最適化するために、大腸菌シンテターゼにより正しく認識されることが分かっている2個のサプレッサーtRNAを使用した。Saccharomyces cerevisiaeに由来する突然変異体サプレッサーtRNA Tyr (pAC-YYG1で発現させたsc-tRNA Tyr CUA)はβラクタマーゼ遺伝子のアンバーコドン (Ala184TAG)を抑圧し、大腸菌細胞のIC₅₀が12μg/mlアンピシリンであり、Methanococcus jannaschiiに由来するサプレッサーtRNA Tyr (pAC-JYで発現させたmj-tRNA Tyr CUA)は大腸菌細胞のIC₅₀が56μg/mlである。例えばWangら, (2000), 前出参照。これに対して、Saccharomyces cerevisiae tRNA Gln2に由来するサプレッサーtRNA Gln CUAを同一アッセイで試験するとIC₅₀=21μg/mlアンピシリンであり、in vitro及びin vivoで大腸菌シンテターゼに直交性であることが判明した。例えばLiu & Schultz, (1999), 前出参照。従って、mj-tRNA Tyr CUAを発現する細胞を排除し且つsc-tRNA Tyr CUAを発現する細胞の増殖を許容するネガティブ選択は非直交サプレッサーtRNAを排除する。1%グリセロールと0.3mMロイシンを加えた液体最少培地 (GMM L) にプラスミドpSCB2とpACプラスミドを維持するのに適した抗生物質を加えて細胞を増殖させた。第1組の細胞 (第1組) にはバルナーゼの発現を誘導するためにアラビノースを加え、第2組にはアラビノースを加えなかった。選択後に生存している細胞画分を第1組と第2組の細胞密度の比により決定した。図11参照。対照プラスミドpAC-Cm (サプレッサーtRNAなし) とプラスミドpAC-YYG1を含む細胞は生存したが、プラスミドpAC-JYを含む細胞は大半が死滅した。プラスミドpSCB3を使用した場合には、プラスミドpAC-JYを含む細胞は24時間以内に増殖し始めた。従って、上記ライブラリー選択条件下で2個のアンバーコドンを含むバルナーゼ遺伝子をコードするpSCB2を使用してネガティブ選択を実施した。

【0235】

ポジティブ選択には、オリゴヌクレオチドLW104: 5'-GGAAATTC CATTA GGACGAATTTGA AATG-3' 及びLW105: 5'-AAACTGCAGT TATAATCTCTTTCTA AATTGGCTC-3'を使用してpBLAM-YQRSのNdeI-PstI部位間にpBSA50からのMethanococcus jannaschii TyrRS遺伝子を挿入することによりプラスミド (例えばpBLAM-JYRS) を構築した。例えばSteerら, (1999), 前出、及びLiu & Schultz, (1999), 前出参照。ポジティブ選択条件を最適化するために、pBLAM-JYRSを含むコンピテントDH10B細胞をpAC-Cm、pAC-YYG1及びpAC-JYで別々に形質転換した。単コロニーを釣菌し、Cmとテトラサイクリン (Tet, 40μg/ml) を加えた2×YTで増殖させた。液相選択では、出発時のOD₆₀₀が0.1となるようにCmとTetを加えた2×YTで一晩細胞培養液を希釈した。各種濃度のAmpを加え、OD₆₀₀により細胞増殖をモニターした。プレート選択では、CmとTetと一方の組には500μg/ml Ampも加えた2×YT寒天プレート2組に細胞約103~105個を撒いた。突然変異体tRNAライブラリーの選択では、pBLAM-JYRSを含むコンピテントDH10B細胞にネガティブ選択からの細胞から単離したpACプラスミドを形質転換した。細胞を37℃で45分間再生させ、Cm、Tet及び500μg/ml Ampを加えた各2×YT寒天プレートに細胞約1

0.5個を撒いた。24時間後にコロニーを釣菌し、Cm、Tet及び200 µg/ml Ampを加えた2×YT 6 ml中で再増殖させた。DNAを単離し、アガロースゲル電気泳動により精製した。

【0236】

ポジティブ選択はTEM-1 βラクタマーゼ遺伝子のAla184位に導入したアンバー終止コドンの抑圧に基づいて行う。プラスミドpBLAM-JYRSはMethanococcus jannaschiiチロシルtRNAシンテターゼの遺伝子とAla184にアンバー突然変異をもつβラクタマーゼをコードする。ネガティブ選択後に生存している細胞から単離したpACプラスミドをpBLAM-JYRSと共に大腸菌DH10B細胞に同時形質転換した。非機能的tRNA又はMethanococcus jannaschiiシンテターゼの弱い基質であるtRNAを含む細胞は死滅し、シンテターゼにより負荷することができるtRNAをもつ細胞は生存する。ポジティブ選択の実施可能性を試験するために、Methanococcus jannaschii TyrRSの存在下に2種のモデルサプレッサーtRNAを試験した。sc-tRNA TyrCUAはG1:C72塩基対をもち、効率的にMethanococcus jannaschii TyrRSにより負荷されない。これをAla184アンバーβラクタマーゼ突然変異体と細胞で同時発現させると、細胞はIC₅₀=18 µg/mlアンピシリンまで生存した。他方、Methanococcus jannaschii tRNA TyrCUAとコグネイトTyrRSを含む細胞はIC₅₀=1220 µg/mlアンピシリンまで生存した。例えばWangら, (2000), 前出参照。モデルポジティブ選択をまず液体2×YT培地で試した。各種濃度のアンピシリンを加えた液体2×YT培地でpBLAM-JYRSと各種pACプラスミドを含む細胞を増殖させた結果を図12Aに示す。mj-tRNA TyrCUAで形質転換した細胞はアンピシリン濃度が高いと速い速度で増殖した。細胞を24時間以上増殖させると、pAC-Cm又はpAC-YYG1で形質転換した細胞も飽和まで増殖した。従って、ポジティブ選択をプレートで10³~10⁵/プレートの初期細胞密度で実施した。図12B参照。生存率(アンピシリンを加えないプレートに対するアンピシリンを加えたプレートのコロニー数)は初期細胞密度を変えても有意に変化せず、増殖時間を通して安定であった。アンピシリンプレートでのポジティブ選択の結果、mj-tRNA TyrCUAを発現させた細胞が優先的に増殖した。従って、ライブラリー選択には液体培地でなくプレートでポジティブ選択を行った。

【0237】

アンチコドンループライブラリーを構築するために使用した2個のオーバーラップするオリゴヌクレオチド(tRNA配列を下線で示す): LW125: 5'-GGAA TTC-3'及びLW126: 5'-AAAACTGCAG-3' (配列中、Nは等モルのA、C、T又はGである)を使用することにより突然変異体tRNAのライブラリーを作製した。全ループライブラリーのオリゴヌクレオチド配列はLW145: 5'-GGAA TTC-3'及びLW146: 5'-AAAACTGCAG-3'である。これらの遺伝子を上記と同様にpAC123に挿入し、tRNAライブラリーを得た。

【0238】

アンチコドンループ及び全ループ両者のライブラリーでネガティブ及びポジティブ選択を上記のようにタンデムで実施した。選択したサプレッサーtRNAを単離し、pBLAMを含む大腸菌DH10Bに再形質転換し、大腸菌シンテターゼに対するtRNAの直交性を試験した。次に、pBLAM-JYRSを含む大腸菌にtRNAを再形質転換し、tRNAがMethanococcus jannaschii TyrRSによりどの程度の効率で負荷されたかを試験した。アンチコドンループライブラリーの1回のネガティブ及びポジティブ選択から得られたクローンを配列決定した処、3個の独立したtRNAが単離されたことが分かった。図13参照。pBLAMと同時形質転換した場合には、IC₅₀値は全てMethanococcus jannaschii tRNA TyrCUAよりも低く、より弱い大腸菌シンテターゼの基質であることが分かった。

【0239】

突然変異体 A A 2 も *Methanococcus jannaschii* TyrRS に対して親和性が非常に高かった。この突然変異体 tRNA は大腸菌で安定に維持することができたが、理由は不明であるが増殖速度が低下した。恐らくこのために、いずれもランダム突然変異領域の外側に突然変異をもつ突然変異体 A A 3 及び A A 4 が生じた。A A 3 又は A A 4 を含む細胞は正常に増殖した。しかし、A A 3 と A A 4 は *Methanococcus jannaschii* TyrRS の比較的弱い基質であった。

【0240】

全ループライブラリーを使用する 2 回のネガティブ及びポジティブ選択から 4 個の独立 tRNA を選択した。図 13 参照。いずれも親 *Methanococcus jannaschii* tRNA TyrCUA よりも弱い大腸菌シンテターゼの基質であったが、*in vivo* β ラクタマーゼアッセイによると依然として *Methanococcus jannaschii* TyrRS により効率的に負荷された。表 2 参照。最良突然変異体 J17 を発現する細胞の IC_{50} 値は $12 \mu g/ml$ アンピシリンであり、*Saccharomyces cerevisiae* に由来する直交 tRNA GlnCUA を発現させた細胞 ($21 \mu g/ml$ アンピシリン) よりも更に低い値であった。J17 を *Methanococcus jannaschii* TyrRS と同時発現させると、細胞は $436 \mu g/ml$ アンピシリンの IC_{50} 値まで生存し、選択窓 (TyrRS の存在下の IC_{50} と TyrRS の不在下の IC_{50} の比) は 35 倍となった。更に、これらの突然変異体 tRNA の発現は大腸菌細胞の増殖に影響を与えなかった。

【0241】

表 2. 選択したサプレッサー tRNA の *in vivo* β ラクタマーゼアッセイ

10

20

サプレッサー-TyrRS	IC ₅₀ (μg/mLアンピシリン)	
	pBLAMと同時 発現	pBLAM-JYRS と同時発現
mj-tRNATyrCUA	56	1220
tRNATyrCUAなし	10	10
アンチコドンループライブラリーから選択した突然変異体 tRNA		
AA2	22	1420
AA3	10	110
AA4	12	135
全ループライブラリーから選択した突然変異体 tRNA		
両者選択後に生存している突然変異体 tRNA		
J15	30	845
J17	12	436
J18	20	632
J22	14	459
ネガティブ選択後のみに生存している突然変異体 tRNA		
N11	11	16
N12	9	18
N13	10	12
N16	9	9

プラスミド pBLAMを使用してAla184にアンバーコドンをもつβラクタマーゼ遺伝子を発現させ、プラスミド pBLAM-JYRSを使用してアンバー突然変異体と Methanococcus jannaschii の TyrRS を発現させた。サプレッサー tRNA は pAC プラスミドでコードさせ、アッセイで pBLAM 又は pBLAM-JYRS と同時形質転換した。

【 0 2 4 2 】

選択したサプレッサー tRNA の特性を確認するために、クロラムフェニコールアセチルトランスフェラーゼ (CAT) 遺伝子のアンバーコドンの抑圧に基づく別の *in vivo* アッセイで試験した。ペリプラズムに分泌される β -ラクタマーゼに対して、CAT は細胞質に局在する。更にアンピシリンは殺菌性であるが、クロラムフェニコールは静菌性である。下表 3 に示すように、選択したサプレッサー tRNA は更に CAT アッセイで直交性であり、CAT 選択に適していることが判明した。

【 0 2 4 3 】

表 3. 選択したサプレッサー tRNA の *in vivo* クロラムフェニコールアセチル
トランスフェラーゼアッセイ

サプレッサー-TyrRS	IC ₅₀ (μg/mLクロラムフェニコール)	
	pYC単独	pYC+pBK-JYRS
mj-tRNATyrCUA	27	308
tRNATyrCUAなし	3	3
J15	11	297
J17	4	240
J18	6	284
J22	5	271

10

pYCプラスミドはAsp112にアンバーコドンをもつクロラムフェニコールアセチルトランスフェラーゼと表の左欄に示した各種サプレッサー-tRNAをコードした。pBK-JYRSを使用して*Methanococcus jannaschii*のTyrRSを発現させた。

【0244】

βラクターゼ遺伝子のアンバーコドンの抑圧に基づく*in vivo*相補アッセイを従来の記載通りに実施した。例えばLiu & Schultz, (1999), 前出; 及びWangら, (2000), 前出参照。クロラムフェニコールアセチルトランスフェラーゼ(CAT)アッセイではpACYC184のCAT遺伝子のAsp112をアンバーコドンで置換し、pACMD112TAGとした。例えばM. Pasternak, T. J. Magliery, P. G. Schultz, A new orthogonal suppressor tRNA/aminoacyl-tRNA synthetase pair for evolving an organism with an expanded genetic code, *Helv. Chim. Acta* 83: 2277-2286 (2000) 参照。lppプロモーターとrrnCターミネーターの制御下にサプレッサー-tRNAをコードする遺伝子をaPCプラスミドからNcoIとAvaIで切出し、予め消化しておいたpACMD112TAGに挿入し、夫々プラスミドpYC-JY、pYC-J15、pHC-J17、pYC-J18及びpYC-J22を得た。pBR322の誘導体であるプラスミドpBK-JYRSを使用し、大腸菌GlnRSプロモーター及びターミネーターの制御下に*Methanococcus jannaschii* TyrRSを発現させた。広い濃度範囲のクロラムフェニコールを増殖培地に加えて滴定し、pYCプラスミド単独又はpYCをpBK-JYRSと同時形質転換した大腸菌DH10B細胞の生存を調べ、曲線からIC₅₀値を補間した。

20

30

【0245】

比較のために、ネガティブ選択のみを通過した4個のコロニーをランダムに釣菌し、*in vivo*相補アッセイを使用してtRNAを試験した。pBLAMで形質転換した場合にはいずれもIC₅₀値が非常に低く、ネガティブ選択が良好に機能していることが判明した。表2参照。pBLAM-JYRSと同時形質転換した場合にもIC₅₀値は非常に低く、ポジティブ選択が*Methanococcus jannaschii* TyrRSにより負荷できないtRNAを排除するように機能していることが判明した。

40

【0246】

選択したtRNAのDNA配列を分析した処、ヌクレオチド置換の特徴的パターンが明らかになった。図13参照。ネガティブ選択とポジティブ選択の両者を通過したtRNAはいずれもC32とT60が不変であり、G37がAに突然変異し、T17がA又はGに突然変異していた。半保存変異として、A38からC又はAへの突然変異、T45からT又はAへの突然変異、T47からG又はTへの突然変異もあった。他の突然変異には明白な共通パターンがなかった。ネガティブ選択のみを通過した20個のtRNAも配列決定した処(そのうちの4個を図13に示す)、いずれも上記共通突然変異の少なくとも1個を

50

欠いていることが判明した。

【0247】

選択した突然変異体サプレッサー tRNA における好適ヌクレオチドは次の役割を果たすことができる。(i) 大腸菌シンテターゼによる認識のネガティブ決定基として機能することができる。(ii) *Methanococcus jannaschii* TyrRS によるアミノアシル化のアイデンティティーエレメントを識別することができる。又は (iii) tRNA の抑圧効率を増すように tRNA と大腸菌翻訳機構の相互作用を最適化することもできる。アンチコドンループ及び全ループの両ライブラリーから選択した tRNA に G37A 突然変異が認められたことに注目すべきである。この突然変異は 37 位の アデニンがアンバー抑圧効率を増すことを示した従来の報告に一致している。例えば M. Yarus, Translational efficiency of transfer RNA's: Use of an expanded anticodon, Science 218:646-652 (1982); D. Bradley, J. V. Park, L. Soll, tRNA^{2Gln} Su+2 mutants that increase amber suppression, J. Bacteriol. 145:704-712 (1982); 及び L. G. Kleina, J. Masson, J. Normanly, J. Aberlson, J. H. Miller, Construction of Escherichia coli amber suppressor tRNA genes. II. Synthesis of additional tRNA genes and improvement of suppressor efficiency, J. Mol. Biol. 213:705-717 (1990) 参照。Fechter らは *Methanococcus jannaschii* tRNA^{Tyr} の完全なアイデンティティーセットが 6 ヌクレオチド (C1G72A73 及びアンチコドン G34U35A36) であることを最近報告している。P. Fechter, J. Rüdinger-Thirion, M. Tukulio, R. Giege, Major tyrosine identity determinants in Methanococcus jannaschii and Saccharomyces cerevisiae tRNA^{Tyr} are conserved but expressed differently, Eur. J. Biochem. 268:761-767 (2001) 参照。従って、選択した全突然変異体サプレッサーにおける C32 と T60 の存在は *Methanococcus jannaschii* TyrRS による認識に必要ない。大腸菌 tRNA は 4 個の tRNA が C である以外は全てが 60 位に T をもつ。M. Sprinzl, C. Horn, M. Brown, A. Loudovitch, S. Steinberg, Compilation of tRNA sequences and sequences of tRNA genes, Nucleic Acids Res. 26:148-153 (1998) 参照。酵母 tRNA^{Phe} の結晶構造によると、ヌクレオチド 60 は他のヌクレオチドと相互作用しない。J. L. Sussman, S. R. Holbrook, R. W. Warrant, G. M. Church, S. H. Kim, Crystal structure of yeast phenylalanine transfer RNA. 1. Crystallographic refinement, J. Mol. Biol. 123:607-630 (1978) 参照。従って、T60 は大腸菌翻訳機構との生産的相互作用のために TC ループの形状を維持すると思われる。TC と保存された C61 の間にヌクレオチドが挿入されるとグリシン tRNA が読み枠をシフトするので、TC ループ構造の変化は翻訳忠実性に影響があると思われる。D. J. O'Mahony, B. H. Hims, S. Thompson, E. J. Murgola, J. F. Atkins, Glycine tRNA mutants with normal anticodon loop size cause 1 frameshifting, Proc. Natl. Acad. Sci. USA 86:7979-7983 (1989) 参照。C32 の役割は明らかでないが、大腸菌 tRNA の 32 位は T、C 及び A を含み、2 個の大腸菌 tRNA^{Tyr} は C32 をもつ。17a 位

については、tRNA^{Thr}のみがこの位置にAをもつ。

【0248】

選択した全サプレッサーtRNAは*Methanococcus jannaschii* tRNA^{Tyr}CUAよりも弱い大腸菌シンテターゼの基質であるので、大腸菌に導入すると負荷の誤りが少なくなる。これらのtRNAは宿主細胞の増殖に悪影響を与えずに大腸菌で安定に維持することもできる。更に、効率的に*Methanococcus jannaschii* TyrRSにより負荷することができる。これらの全特性により、突然変異体サプレッサーtRNAは*Methanococcus jannaschii* TyrRSと共に非天然アミノ酸の選択的in vivoタンパク質組込みに使用する堅牢な直交tRNA-シンテターゼ対を形成することができる。J17突然変異体サプレッサーtRNAと組換え突然変異体TyrRSを使用し、20種の共通アミノ酸に匹敵する忠実度でTAGコドンに応答してO-メチル-L-チロシンが送達された。L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Expanding the genetic code of Escherichia coli*, Science, 292: 498-500 (2001) 参照。

10

【実施例2】

【0249】

実施例2-mutRNA^{Tyr}/CUAに非天然アミノ酸O-メチル-L-チロシンを負荷するようにTyrRSを突然変異

【0250】

遺伝的にコードされるアミノ酸の数を拡大するユニークなトランスファーRNA (tRNA) -アミノアシルtRNAシンテターゼ対が大腸菌で作製されている。この対を大腸菌に導入すると、外部から増殖培地に添加した合成アミノ酸O-メチル-L-チロシンをアンバーナンセンスコドンに応答してタンパク質にin vivo組込むことができる。非天然アミノ酸を含むジヒドロ葉酸レダクターゼの分析によると、翻訳忠実度は99%を上回る。このアプローチは20種の共通アミノ酸に存在しない新規な構造的、化学的及び物理的特性をもつ多様なアミノ酸を導入するための生きた細胞の遺伝子レパートリーを増加する一般方法を提供する。

20

【0251】

異種間アミノアシル化が非効率的であり、場合により、アンチコドンループがシンテターゼ認識の主要決定因子ではない場合には、別の生物に由来する対を導入することにより大腸菌で直交tRNA/シンテターゼ対を作製することができる。このような候補対の1例は*Methanococcus jannaschii*のチロシルtRNA/シンテターゼ対であり、*Methanococcus jannaschii*はtRNA^{Tyr}アイデンティティーエレメントが大腸菌tRNA^{Tyr}と異なり（特に、アクセプターステムの第1番目の塩基対が大腸菌ではGCであるが、*Methanococcus jannaschii*ではCGである）、チロシルシンテターゼ（TyrRS）が最低限のアンチコドンループ結合ドメインしかもたない始原菌である。例えばB. A. Steer, & P. Schimmel, J. Biol. Chem. 274: 35601-6 (1999) 参照。更に、*Methanococcus jannaschii* TyrRSは編集メカニズムをもたない（例えばJakubowski & Goldman, Microb. iol. Rev., 56: 412 (1992) 参照）ので、tRNAにライゲートした非天然アミノ酸を校正しないと思われる。*Methanococcus jannaschii* TyrRSはそのコグネイトtRNA^{Tyr}に由来するアンバーサプレッサーtRNAを効率的にアミノアシル化する（例えばWangら, (2000), J. Am. Chem. Soc., 前出参照）が、大腸菌tRNAをアミノアシル化しない（例えばSteer & Schimmel, (1999), 前出参照）。更に、*Methanococcus jannaschii* tRNA^{CUA}^{Tyr}は大腸菌シンテターゼの弱い基質であるが、大腸菌でタンパク質翻訳に効率的に機能する。例えばWangら, (2000), J. Am. Chem. Soc., 前出参照。

30

40

50

【0252】

大腸菌シンテターゼによる直交 tRNA、Methanococcus jannaschii tRNA_{CUA}^{Tyr} の認識を更に弱めるために、Methanococcus jannaschii TyrRS と直接相互作用しない tRNA の 11 個のヌクレオチド (C16, C17, U17a, U20, C32, G37, A38, U45, U47, A59 及び U60) をランダムに突然変異させ、サプレッサー tRNA ライブラリーを作製した。この tRNA ライブラリーをネガティブ選択 (例えばバルナーゼ遺伝子等の毒性レポーター遺伝子におけるアンバー突然変異の抑圧) に向け、大腸菌シンテターゼによりアミノアシル化される tRNA を除去した後、効率的に Methanococcus jannaschii TyrRS によりアミノアシル化される tRNA をポジティブ選択した (例えば β ラクタマーゼ遺伝子等のレポーター遺伝子におけるアンバー突然変異の抑圧)。

【0253】

得られたサプレッサー tRNA の直交性をベクター上のレポーター (例えばプラスミド pBLAM に担持した TEM-1 β ラクタマーゼ遺伝子) の非必須位置 (例えば Ala184) のアンバー終止コドンの抑圧に基づく in vivo 相補アッセイにより試験した。形質転換したサプレッサー tRNA が内在大腸菌シンテターゼによりアミノアシル化されると、アンピシリンの存在下で細胞増殖する。Methanococcus jannaschii tRNA_{CUA}^{Tyr} とレポーター構築物 pBLAM で形質転換した大腸菌は 55 μg/mL アンピシリンの存在下に生存する。ライブラリーから選択した最良突然変異体サプレッサー tRNA (mtRNA_{CUA}^{Tyr}) を発現させると、僅か 12 μg/mL のアンピシリンでも細胞は生存し、サプレッサー tRNA の不在下でも同様の結果が得られる。突然変異体サプレッサー tRNA はヌクレオチド置換 C17A、U17aG、U20C、G37A 及び U47G を含んでいた。Methanococcus jannaschii TyrRS をこの mtRNA_{CUA}^{Tyr} と同時発現させると、細胞は 440 μg/mL アンピシリンで生存する。従って、mtRNA_{CUA}^{Tyr} は Methanococcus jannaschii tRNA_{CUA}^{Tyr} よりも弱い内在シンテターゼの基質であるが、効率的に Methanococcus jannaschii TyrRS によりアミノアシル化される。

【0254】

mtRNA_{CUA}^{Tyr} に所望非天然アミノ酸を負荷するように直交 TyrRS のアミノ酸特異性を改変するために、TyrRS 突然変異体のライブラリーを作製し、スクリーニングした。Bacillus stearothermophilus に由来する相同 TyrRS の結晶構造 (例えば、P. Brick, T. N. Bhat, D. M. Blow, J. Mol. Biol., 208:83 (1988)) に基づき、結合チロシンのアリアル環のパラ位から 6.5 Å 以内にある Methanococcus jannaschii TyrRS の活性部位の 5 個の残基 (Tyr³²、Glu¹⁰⁷、Asp¹⁵⁸、Ile¹⁵⁹ 及び Leu¹⁶²) を突然変異させた。図 14 参照。これらの残基をまず全てアラニンに突然変異させ、得られた不活性 Ala₅TyrRS を鋳型として使用し、オリゴヌクレオチドプライマーによるポリメラーゼ連鎖反応 (PCR) ランダム突然変異誘発を行った。

【0255】

例えば、カナマイシン耐性をもつ pBR322 由来プラスミドであるプラスミド pBK-JYRS で大腸菌 GlnRS プロモーター及びターミネーターの制御下に TyrRS 遺伝子を発現させた。残基 Tyr³²、Glu¹⁰⁷、Asp¹⁵⁸、Ile¹⁵⁹ 及び Leu¹⁶² を部位特異的突然変異誘発により Ala で置換し、プラスミド pBK-JYA5 を得た。突然変異部位に NNK (N=A+T+G+C、K=G+T、及び M=C+A) をもつ 8 種のオリゴヌクレオチド (例えばオリゴヌクレオチド LW157 5'-GGAAATTC CATATGGACGAATTTGAAATG-3'、LW164 5'-GTATTTT TCCACTTGTTC AAAACCTATMNNAGCAGATTTTTCATCTTTTTCATCTTTT TTTTTCATCTTTT TTTTAA AAC-3')

10

20

30

40

50

、LW159 5' - TAGGTTTGTGAACCAAGTGGTAAAAATAC - 3'
、LW165 5' - CATTCAGTGTAATAATCCTTATCAAGCTGGA
AMNNACTTCCATAAACATATTTTGGCCTTTAAAC - 3'
、LW161 5' - TCCAGCTTGATTAAGGATTAATACACTGAATG
- 3'、LW167 5' - CATCCCTCCAACCTGCAACATCAACGCC
MNNATAATGMNNMNNATTAACCTGCATTATTGGATAGATA
AC - 3'
、LW163 5' - GCGTTGATGTTGTCAGTTGGAGGGATG - 3'、
及びLW105 5' - AAACTGTCAGTTATAAATCTCTTTCTAATTG
GCTC - 3'

10

(Operon, CA)) を使用してAla₅ TyrRS突然変異体 (pBK-JYA5) のPCR増幅を行い、NdeI-PstIで消化したpBK-JYA5に再ライゲートしてTyrRSライブラリーを得た。ライゲートしたベクターを大腸菌DH10Bコンピテント細胞に形質転換し、 1.6×10^9 コロニー形成単位 (cfu) のライブラリーを得た。40個のランダムに釣菌したコロニーからのTyrRS遺伝子を配列決定した処、ランダムに突然変異させたNNK位置に塩基置換はなく、他の予想外の突然変異もないことが確認された。ライブラリーをマキシプレップにより増幅し、スーパーコイルDNAを使用して選択株pYC-J17を形質転換した。

【0256】

次に、クロラムフェニコールアセチルトランスフェラーゼ (CAT) 遺伝子の非必須位置 (例えばAsp112) のアンバーコドンの抑圧に基づき、突然変異直交TyrRSのライブラリーをポジティブ選択にかけた。例えばM. Pasternak, T. J. Maglery, P. G. Schultz, Helv. Chim. Acta, 83:2277 (2000) 参照。突然変異体TyrRSライブラリーとmtRNA_{CUA}^{Tyr} 遺伝子で形質転換した細胞を、非天然アミノ酸を加えた培地で細胞を増殖させ、各種濃度のクロラムフェニコールの存在下で生存するものを選択した。突然変異体TyrRSが直交mtRNA_{CUA}^{Tyr} に天然又は非天然アミノ酸を負荷するならば、細胞はCATを生産し、生存する。次に生存細胞をクロラムフェニコールの存在下で非天然アミノ酸の不在下に増殖させた。非天然アミノ酸を加えたレプリカプレートから生存しなかった細胞 (例えば直交mtRNA_{CUA}^{Tyr} に非天然アミノ酸を負荷する突然変異体TyrRSをコードする細胞) を単離した。突然変異体TyrRS遺伝子をこれらの細胞から単離し、DNAシャフリングによりin vitro組換えし、クロラムフェニコール濃度を増加する後期選択のために大腸菌に再形質転換した。

20

30

【0257】

パラヒドロキシ基をメトキシ基で置換したチロシン類似体を選択に使用した。場合により、他のチロシン類似体も選択に使用することができ、例えばアリール環のパラ位に各種官能基 (アセチル、アミノ、カルボキシル、イソプロピル、メチル、O-メチル及びニトロ等) をもつチロシン類似体が挙げられる。例えば、Asp₁₁₂TAGCAT突然変異体もコードするpACYC184誘導体であるプラスミドpYC-J17でlppプロモーターとrrnCターミネーターの制御下にmtRNA_{CUA}^{Tyr} をコードする遺伝子が大腸菌DH10B細胞で発現させた。pYC-J17を導入した大腸菌DH10Bコンピテント細胞にTyrRSライブラリーをコードするスーパーコイルDNAを形質転換し、元のライブラリーを完全にカバーする 3×10^9 cfuよりも大きい寸法のライブラリーを得た。次に、1%グリセロールと0.3mMロイシンを含む最少培地 (GMM L) プレートに17µg/mLテトラサイクリン (Tet)、25µg/mLカナマイシン (Kan)、50µg/mLクロラムフェニコール (Cm) 及び1mM非天然アミノ酸を加えて細胞を培養した。37℃で44時間培養後にO-メチル-L-チロシンを加えたプレート上のコロニーをプールし、プラスミドを単離し、pYC-J17を導入した大腸菌DH10Bコンピテント細胞に再形質転換し、形質転換細胞を50µg/mL Cm上でポジティブ選択した。コロニー (96) をプレートから個々に釣菌し、液体GMM L培地100mL

40

50

で希釈し、各種濃度のCmを加えたKan/Tet GMM Lプレート2組に画線した。第1組のプレートはO-メチル-L-チロシンを加えず、Cm濃度を10~25 µg/mLとし、第2組のプレートは1mM O-メチル-L-チロシンと50 µg/mL Cmを加えた。第1組のプレートで15 µg/mLのCmの存在下に増殖しなかったコロニーのレプリカを第2組のプレートから釣菌した。TyrRS遺伝子を含むプラスミドを精製し、断片化反応のMg²⁺を10mM Mn²⁺に代えた以外はStemmerのプロトコールを使用してDNAシャフリングにより*in vitro*組換えした。W. P. C. Stemmer, Nature 370, 389-91 (1994); 及びI. A. Lorrimer, I. Pastan, Nucleic Acids Res. 23, 3067-8 (1995) 参照。次に、予め消化しておいたpBK-JYA5ベクターにライブラリーを再ライゲートし、典型寸法 $8 \times 10^8 \sim 3 \times 10^9$ cfuの第2世代TyrRSライブラリーを得た。ライブラリーからランダムに選択した30個を配列決定した。DNAシャフリングにより導入された突然変異誘発率は0.35%であった。このライブラリーを次回選択用選択株に形質転換した後、シャフリングした。第2組のプレートではポジティブ選択のCm濃度を2回目に80 µg/mL、3回目に120 µg/mLに上げ、第1組のプレートではCm濃度を変えなかった。3回のDNAシャフリング後にコロニーは第1組のプレートではCm濃度20~25 µg/mLで増殖し始め、TyrRS突然変異体が天然アミノ酸を基質として受容していることを示した。従って、2回のDNAシャフリング後に選択した最良クローンを詳細に特性決定した。

【0258】

2回の選択とDNAシャフリングを実施し、クロラムフェニコール中の生存が増殖培地への1mM O-メチル-L-チロシン添加に依存するクローンを得た。O-メチル-L-チロシンの不在下では、突然変異体TyrRSを導入した細胞は1%グリセロール、0.3mMロイシンを含む最少培地(GMM L)プレートに15 µg/mLクロラムフェニコールを加えると生存できなかった。細胞は1mM O-メチル-L-チロシンの存在下では125 µg/mLクロラムフェニコールを加えたGMM Lプレートで増殖することができた。液体GMM Lでも同様の結果が得られた。対照として、mtRNA_{CUA}^{Tyr}と不活性Ala₅ TyrRSを導入した細胞は1mM O-メチル-L-チロシンの有無に拘わらず、使用した最低クロラムフェニコール濃度でも生存しなかった。図14参照。1mM O-メチル-L-チロシン自体を加えても大腸菌の増殖率はさほど変わらない。

【0259】

mtRNA_{CUA}^{Tyr}にO-メチル-L-チロシンを負荷する突然変異体TyrRSの配列を分析した処、以下の突然変異が判明した。Tyr³²→Gln³²、Asp¹⁵⁸→Ala¹⁵⁸、Glu¹⁰⁷→Thr¹⁰⁷ 及びLeu¹⁶²→Pro¹⁶²。図14参照。相同B. stearothermophilus TyrRSのX線結晶構造によると、Tyr³²、Asp¹⁵⁸と基質チロシン間の水素結合網が失われるとチロシンと突然変異体TyrRSの結合に不利であると思われる。実際に、B. stearothermophilus TyrRSのAsp¹⁷⁶(Methanococcus jannaschiiのAsp¹⁵⁸に対応)の突然変異は不活性な酵素を生じる。例えばG. D. P. Gray, H. W. Duckworth, A. R. Fersht, FEBS 318:167-71 (1993) 参照。同時に、Asp¹⁵⁸→Ala¹⁵⁸ 及びLeu¹⁶²→Pro¹⁶² 突然変異はO-メチル-L-チロシンのメチル基を更に基質結合キャビティに延ばす疎水性ポケットを形成する。アデニル酸のリボース又はリン酸基に結合する活性部位の他の重要な触媒残基は2回のDNAシャフリング後も変化しなかった。

【0260】

突然変異体TyrRSを触媒とするO-メチル-L-チロシン及びチロシンとアデノシン三リン酸(ATP)のアデニル酸形成の動態を37℃でピロリン酸交換アッセイにより*in vitro*分析した。例えば、そのC末端に6個のヒスチジンを付けた突然変異体TyrRSをプラスミドpQE-60(Qiagen, CA)にクローニングし、プラスミドpQE-mJYRSを作製した。製造業者のプロトコール(Qiagen, CA)に従

って固定化金属アフィニティークロマトグラフィーによりタンパク質を精製した。100 mM TrisHCl (pH 7.5)、10 mM KF、5 mM MgCl₂、2 mM ATP、2 mM NaPPi、0.1 mg/mL ウシ血清アルブミン、約0.01 μ Ci/mL [³²P] NaPPi 及び各種濃度のチロシン又はO-メチル-L-チロシンを含む反応混合物中で37℃でピロリン酸 (PPi) 交換を行った。精製突然変異体 TyrRS を加えて反応を開始し、定期的にアリコートを取分け、0.2 M NaPPi、7% 過塩素酸及び2% 活性炭で反応停止した。活性炭を濾過し、10 mM NaPPi (pH 2) で洗浄した後、活性炭吸着ATPの³²P濃度をシンチレーション計数により測定した。非線形回帰分析を使用してミカエリス・メンテンの式の直接フィットにより k_{cat} 及び K_m 値を計算した。

10

【0261】

チロシンのミカエリス定数 (K_m) ($5833 \pm 902 \mu M$) はO-メチル-L-チロシン ($443 \pm 93 \mu M$) の約1.3倍であり、チロシンの触媒速度定数 (k_{cat}) ($1.8 \pm 0.2 \times 10^{-3} s^{-1}$) はO-メチル-L-チロシン ($14 \pm 1 \times 10^{-3} s^{-1}$) の8分の1である。従って、O-メチル-L-チロシンに対する突然変異体 TyrRS の k_{cat}/K_m 値はチロシンの約100倍である。大腸菌におけるチロシンの生理的濃度は約80 μM であり、突然変異体 TyrRS のチロシンに対する K_m 値 ($5833 \mu M$) よりもはるかに低い。細胞中のO-メチル-L-チロシンの濃度は K_m ($443 \mu M$) と同等以上であると予想される。

【0262】

本実施例は遺伝的にコードされる付加的アミノ酸を受容するように大腸菌のタンパク質合成機構を補強できることを示す。新規アミノ酸を直接生きた細胞でタンパク質に導入できると、タンパク質及び細胞機能の研究の新規ツールとなると共に、天然タンパク質に比較して特性を増進したタンパク質を作製することが可能になる。本明細書に記載する方法は新規分光学的、化学的、構造的等の性質をもつ他のアミノ酸にも適用できる。大腸菌リボソームはin vitroタンパク質合成を使用して多様な側鎖をもつアミノ酸をタンパク質に組み込むことができることが示されている。例えば、C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, P. G. Schultz, Science 244, 182-8 (1989) 参照。付加的直交tRNA/システターゼ対 (例えばD. R. Liu, P. G. Schultz, Proc. Natl. Acad. Sci. USA 96, 4780-5 (1999); 及びA. K. Kowal, C. Kohrer, U. L. RajBhandary, Proc. Natl. Acad. Sci. U. S. A., 98: 2268 (2001) 参照)、4塩基コドン (例えばT. J. Magliery, J. C. Anderson, P. G. Schultz, J. Mol. Biol. 307: 755 (2001); 及びB. Moore, B. C. Persson, C. C. Nelson, R. F. Gesteland, J. F. Atkins, J. Mol. Biol., 298: 195 (2000) 参照) 及び本明細書に記載する他のセクターコドンは組み込み可能なアミノ酸の数と範囲を更に広げることができる。真核細胞の直交対も本明細書に記載する方法により作製することができる。

20

30

【0263】

参考資料として本明細書に組み込む対応特許出願 (発明の名称 "In vivo Incorporation of Unnatural Amino Acids", 代理人整理番号54-000120PC/US) も参照されたい。同出願は上記系を使用してジヒドロ葉酸レダクターゼ (DHFR) のO-メチル-L-チロシン突然変異体を作製する実施例を記載している。

40

【実施例3】

【0264】

実施例3-mutRNA Tyr/CUAに非天然アミノ酸L-3-(2-ナフチル)アラニンを負荷するようにTyrRSを突然変異

【0265】

50

本実施例は第2の非天然アミノ酸L-3-(2-ナフチル)アラニンを生物(例えば大腸菌)でタンパク質に組み込むために使用可能な別の直交対に関する。この非天然アミノ酸をタンパク質に組み込む直交対を作製するために使用する方法を以下に記載する。この非天然アミノ酸のタンパク質組み込みに関するより詳細な記載は、参考資料として本明細書に組み込む対応特許出願(発明の名称"*In vivo incorporation of unnatural amino acids*", 代理人整理番号54-000120PC/US)を参照されたい。

【0266】

非天然アミノ酸をコードするためにアンバー終止コドンとその対応する直交アンバーサプレッサートRNAであるmutRNA_{CUA}^{Tyr}を選択した。上記及びWang & Schultztz, *Chem. Biol.* 8:883-890(2001)に記載されているように、非天然アミノ酸特異性をもつ直交シンテターゼを生産するための出発点として*Methanococcus jannaschii*チロシルtRNAシンテターゼ(TyrRS)を使用した。このTyrRSは内在大腸菌tRNAをアミノアシル化しない(例えばSteer & Schimmel, *Biol. Chem.*, 274:35601-35606(1999)参照)が、mutRNA_{CUA}^{Tyr}をチロシンでアミノアシル化する(例えばWang, Magliery, Liu, Schultztz, *J. Am. Chem. Soc.*, 122:5010-5011(2000)参照)。L-3-(2-ナフチル)アラニンはチロシンと構造が大きく異なり、新規充填特性をもつ可能性があることからこの試験に選択した。mutRNA_{CUA}^{Tyr}に20種の共通アミノ酸は負荷せずにL-3-(2-ナフチル)アラニンを負荷するようにTyrRSのアミノ酸特異性を改変するために、*Methanococcus jannaschii* TyrRS突然変異体のライブラリーを作製し、スクリーニングした。*Bacillus stearothermophilus*からの相同TyrRSの結晶構造の分析(Brick, Bhat, Blow, *J. Mol. Biol.*, 208:83-98(1989)参照)に基づき、チロシンのアリール環のバラ位から7Å以内にある*M. jannaschii* TyrRSの活性部位の5個の残基(Tyr³²、Asp¹⁵⁸、Ile¹⁵⁹、Leu¹⁶²及びAla¹⁶⁷)を突然変異させた。図15参照。L-3-(2-ナフチル)アラニンに特異的なシンテターゼはWang, Brock, Herberich, Schultztz, *Science*, 292:498-500(2001)に報告されている突然変異体TyrRSライブラリーから選択しなかった。後期選択における野生型シンテターゼ汚染を減らすために、(Ala¹⁶⁷を除く)これらの残基をまず全てアラニンに突然変異させた。得られた不活性Ala₅ TyrRS遺伝子を鋳型として使用し、対応する部位にランダム突然変異をもつオリゴヌクレオチドを用いてポリメラーゼ連鎖反応(PCR)によりランダム突然変異誘発を行った。

【0267】

突然変異体TyrRSライブラリーをまずクロラムフェニコールアセチルトランスフェラーゼ(CAT)遺伝子の非必須位置(Asp¹¹²)のアンバー終止コドンの抑圧に基づくポジティブ選択にかけた。1mM L-3-(2-ナフチル)アラニンと80μg/mLクロラムフェニコールを加えた最少培地で突然変異体TyrRSライブラリーとmutRNA_{CUA}^{Tyr}で形質転換した細胞を増殖させた。細胞は突然変異体TyrRSが天然アミノ酸又はL-3-(2-ナフチル)アラニンでmutRNA_{CUA}^{Tyr}をアミノアシル化する場合にしか生存することができない。次に、クロラムフェニコールの存在下で非天然アミノ酸の不在下に生存細胞を増殖させた。生存しなかった細胞はmutRNA_{CUA}^{Tyr}にL-3-(2-ナフチル)アラニンを負荷する突然変異体をコードすると思われるので、非天然アミノ酸を加えたレプリカプレートから釣菌した。3回のポジティブ選択後にネガティブスクリーニングを行った後、CAT遺伝子のAsp¹¹²TAGコドンの抑圧に基づく*in vivo*アッセイを使用して4個の突然変異体TyrRSを特性決定した。

【0268】

表4. 突然変異体TyrRSの*in vivo*クロラムフェニコールアセチルトランス

フェラーゼアッセイ^a

突然変異体Ty r R S

I C₅₀ (μ g / mL クロラムフェニコール)L-3-(2-ナフ
チル)-AlaなしL-3-(2-ナフ
チル)-Ala添加

Ty r R Sなし

4

4

野生型Ty r R S

240

240

選択後

S1-Ty r R S

30

120

S2-Ty r R S

30

120

S3-Ty r R S

25

110

S4-Ty r R S

35

100

DNAシャフリング後

SS12-Ty r R S

9

150

^a p Y C-J 17 プラスミドを使用してmutRNA_{CUA}^{Tyr} 遺伝子とAsp112にアンバー終止コドンをもつクロラムフェニコールアセチルトランスフェラーゼ遺伝子を発現させた。p B K プラスミドを使用してTy r R Sを発現させ、大腸菌DH10Bにp Y C-J 17 と同時形質転換した。各種濃度のクロラムフェニコールで滴定してG M M L プレート上の細胞生存を試験した。

10

20

30

40

50

【0269】

L-3-(2-ナフチル)アラニンの不在下では、選択したTy r R SとmutRNA_{CUA}^{Tyr}を発現する細胞は1%グリセロールと0.3mMロイシンを含む最少培地プレート(G M M L プレート)で25~35μg/mLクロラムフェニコールの存在下に生存し、L-3-(2-ナフチル)アラニンの存在下では、細胞はG M M L プレートで100~120μg/mLクロラムフェニコールの存在下に生存した。Ty r R Sの不在下のI C₅₀値(4μg/mLクロラムフェニコール)に比較してこれらの結果は選択したTy r R SがL-3-(2-ナフチル)アラニンを受容するが、天然アミノ酸もある程度まで負荷することを示している。上記表4参照。

【0270】

天然アミノ酸に対する突然変異体Ty r R Sの活性を更に弱めるために、上記4種の突然変異体遺伝子を鋳型として使用してDNAシャフリングを1回行った。得られた突然変異体Ty r R Sライブラリーを更に2回ポジティブ選択とネガティブスクリーニングにかけた。天然アミノ酸に対する活性が大幅に低下し(I C₅₀=9μg/mLクロラムフェニコール)且つL-3-(2-ナフチル)アラニンに対する活性が増加(I C₅₀=150μg/mLクロラムフェニコール)した1個の突然変異体Ty r R S(SS12-Ty r R S)が得られた。表4参照。

【0271】

得られた進化型SS12-Ty r R Sは以下の突然変異をもつ。Ty r³²→Leu³²、Asp¹⁵⁸→Pro¹⁵⁸、Ile¹⁵⁹→Ala¹⁵⁹、Leu¹⁶²→Gln¹⁶²、及びAla¹⁶⁷→Val¹⁶⁷。図15参照。相同B. stearoothermophilus Ty r R Sの結晶構造によると、Ty r³²→Leu³²とAsp¹⁵⁸→Pro¹⁵⁸の突然変異はTy r³²、Asp¹⁵⁸と天然基質チロシン間の水素結合を失わせる可能性があるのでチロシンとSS12-Ty r R Sの結合に不利であると考えられる。殆どの残基はL-3-(2-ナフチル)アラニンの結合を助長すると予想される疎水性側鎖をもつアミノ酸に突然変異される。野生型Methanococcus jannaschii Ty r R Sと進

化型 S S 1 2 - T y r R S の結晶構造は入手可能な方法により決定することができる。

【0272】

本明細書と対応出願（発明の名称 " *In vivo incorporation of unnatural amino acids* " , 代理人整理番号 54-000120 PC/US）に記載した細胞増殖、タンパク質発現及び質量分析実施例によると、mutRNA_{CUA}^{Tyr} / S S 1 2 - T y r R S 対は天然アミノ酸に匹敵する忠実度でアンバーコードンに応答して L - 3 - (2 - ナフチル) アラニンをタンパク質に選択的に挿入することができた。Wang, Brock, and Schultz, *Adding L - 3 - (2 - Naphthyl) alanine to the genetic code of E. coli*, J. Am. Chem. Soc., (2002) 124 (9) : 1836 - 7 も参照されたい。チロシンと構造的に異なるアミノ酸に関する上記結果は、本明細書に記載する方法が多様な非天然アミノ酸に適用可能であることを裏付けるものである。

【実施例4】

【0273】

実施例4 - mutRNA^{Tyr} / CUA に負荷するように T y r R S を突然変異させ、所望特性をもつ突然変異 T y r R S を他の方法（例えば FACS 及びファージディスプレイ・パンニング）によりスクリーニング

【0274】

上記のようなレポーター遺伝子とタンパク質を *in vivo* FACS スクリーニング、抗体検出、*in vitro* ファージディスプレイ・パンニング等と併用して直交対を選択することもできる。例えば、Wang & Schultz, *Expanding the genetic code*, Chem. Commun., 1 : 1 - 11 (2002) 参照。

【0275】

例えば、シンテターゼをスクリーニングするために、例えばグリーン蛍光タンパク質 (GFP) をレポーターとして汎用蛍光活性化細胞ソーティング (FACS) によるスクリーニングが開発されている。図 16 A 及び B 参照。シンテターゼ活性は GFP の発現を誘導する T7 RNA ポリメラーゼ内のセレクトターコードン（例えばアンバー終止コードン (TAG) ）の抑圧により報告される。例えば、本発明の選択 / スクリーニング方法の別の例を示す図 26 参照。アンバーコードンが抑圧される場合のみに細胞は機能的 T7 RNA ポリメラーゼを生産し、GFP を発現することができるので、細胞は蛍光発光する。ポジティブスクリーニングでは、直交 tRNA に天然又は非天然アミノ酸を負荷する活性シンテターゼをコードする蛍光細胞を回収する。その後、選択した細胞を希釈し、非天然アミノ酸の不在下で増殖させた後、例えば非天然アミノ酸のみに特異性をもつシンテターゼを発現する非蛍光細胞を FACS により選別する。図 17 A、B、C 及び D はグルタミンシンテターゼを使用したセレクトターコードン（例えばアンバーコードン）の抑圧を示す。蛍光強度の測定閾値を設定することにより、ポジティブ及びネガティブ両者のスクリーニングのストリンジエンシーを簡便に制御することができる。

【0276】

特定非天然アミノ酸に特異的な直接ポジティブ選択法として、ファージ表面に提示される非天然アミノ酸に対するモノクローナル抗体の高親和性を利用した方法も開発されている。図 18 参照。M. Pasternak と P. G. Schultz, Bioorg. Med. Chem., 9 : 2373 (2001) 参照。例えば、ファージ産生がアンバー終止コードンの抑圧を必要とするように、アンバー突然変異をもつ C3 ペプチドを VSCM13 ファージコートタンパク質 pIII の N 末端に融合する。直交サプレッサー tRNA とシンテターゼライブラリーを発現するファージミドを含む細胞に C3 TAG ファージを感染させる。活性シンテターゼにより C3 TAG が抑圧され、そのコグネイトアミノ酸がファージ表面に提示される。その後、ファージプールを非天然アミノ酸に対する固定化モノクローナル抗体と共にインキュベートし、非天然アミノ酸に特異的なシンテターゼをもつファージのみを単離する。模擬選択では、A sp を含むエピトープに対する抗体を使用して A

s nを提示するファージのプールからA s pを提示するファージを300倍以上に集積した。

【0277】

数種のin vitroスクリーニング方法も使用することができる。このような方法の1例では、突然変異体シンテターゼのライブラリーをファージに提示し、ファージ粒子をアミノアシルアデニル酸中間体の固定化アミノアルキルアデニル酸類似体に対してパンニングする。図19参照。例えば、Methanococcus jannaschii TyrRSをM13ファージのpIIIコートタンパク質に融合した。このファージはチロシルアデニル酸のアミノアルキルアデニル酸類似体に対してパンニング後に非関連抗体を提示する対照ファージの1000倍に集積した。出発時のTyrRSファージ集団の0.1~1%しかTyrRSタンパク質を提示しないので、実際の集積倍率は 10^5 ~ 10^6 にすることができる。

【実施例5】

【0278】

実施例5—始原菌ロイシルtRNAシンテターゼ対の作製

【0279】

始原菌Methanococcus jannaschii由来ロイシルtRNAシンテターゼとして、始原菌ロイシルtRNAに由来するアンバー及びフレームシフトサプレッサーtRNAをアミノアシル化することができるが、大腸菌に内在するtRNAをアミノアシル化しないロイシルtRNAシンテターゼを同定した。本明細書に記載する選択ストラテジーを使用して、このシンテターゼにより選択的に負荷される高度に活性なtRNA基質を同定した。突然変異体シンテターゼライブラリーを作製し、非天然アミノ酸を選択的に負荷することが可能な突然変異体シンテターゼを選択することができる。

【0280】

アンチコドンでCUAアンチコドンで置換してアンバーサプレッサーtRNAとした5個の異なる始原菌ロイシルtRNAに由来するアンバーコドンとサプレッサーtRNAをシンテターゼもつβラクタマーゼレポーター遺伝子を構築した。7個の異なるロイシルtRNAをクローニングし、レポーター構築物で同時形質転換した。3個のシンテターゼはシンテターゼをもたない対照よりもシンテターゼの存在下でアンピシリン耐性が高く、これらの系を更に試験した。図20参照。

【0281】

次段階では宿主tRNAと相互作用せずにサプレッサーtRNAに負荷するシンテターゼを決定した。Methanobacterium thermoautotrophicumとMethanococcus jannaschiiの2種の系を選択して発現させ、陽性対照としてHalobacteriumに由来する精製tRNAと大腸菌全tRNAでin vitroアミノアシル化を実施した。Methanococcus jannaschiiシンテターゼは大腸菌tRNAに有効に負荷することができたが、Methanobacterium thermoautotrophicumシンテターゼはHalobacterium tRNAに特異的であることが判明した。

【0282】

抑圧系の効率を高めるように更に改良した。アンチコドンループのA³⁷はロイシルtRNAシンテターゼではG³⁷であった。この突然変異は各種真核細胞とHalobacteriumで非コグネイトシンテターゼによるアミノアシル化に対するネガティブ決定基であると共に、(Halobacteriumではそうでないが)酵母でアミノアシル化のポジティブ決定基であることが分かった。A³⁷は効率的抑圧の重要な要件であることも分かった。アンチコドンループをランダムに突然変異誘発し、より効率的な抑圧について選択した。G³⁷をAに突然変異させると、より効率的なサプレッサーとなり、突然変異しないものに比較して20倍の濃度のアンピシリンを抑圧することができた。図21参照。

【0283】

大腸菌で他のシンテターゼにより優先的に負荷されないようにtRNAを改良するために

、tRNAのアクセプターステムをランダム突然変異誘発した。ポジティブ／ネガティブ選択を使用して*Methanobacterium thermoautotrophicum* RSの不在下に負荷されないtRNAを同定した。

【0284】

選択した突然変異tRNAを観察した処、ロイシルアミノアシル化の必須ポジティブ決定基であることが先の全系で判明した識別子塩基A³⁷は観察した全突然変異tRNAで保存されていた。直交性が改善された全突然変異tRNAでC³:C⁷⁰塩基対も保存されていた。観察された最良の突然変異体tRNAはシンテターゼの不在下のアミノアシル化が約3分の1となり、*Methanobacterium thermoautotrophicum* RSの存在下の抑圧が実際に増加した。

10

【0285】

例えば3塩基コドンでなく4塩基コドンを抑圧することが可能な変異体も作製した。4塩基コドンは一度に3塩基でなく一度に4塩基の遺伝コードをデコードすることが可能になり、64種しかアミノ酸をコードできない場合に256種をコードできる。4塩基コドンの使用に伴う問題はtRNAのアンチコドンループの拡張が必要であり、殆どの系が攪乱を受ける可能性があることである。しかし、ロイシル系の初代AGGAサプレッサーが同定された。これは8塩基のアンチコドンループをランダム突然変異誘発し、AGGA-βラクターゼレポーター系で選択を行うことにより作製した。図22参照。

【0286】

シンテターゼの編集機構も突然変異させ、編集機能を除去した。ロイシル系は他の数種のシンテターゼと同様に（少なくとも）2個の活性部位をもつ。一方の部位はATPによるアミノ酸の活性化を行い、シンテターゼと共に酵素結合アミノアシルアデニル酸を形成した後、アミノ酸をtRNAの3'末端に転移する。他方、第2の部位はロイシン以外のアミノ酸をtRNAから加水分解することができない。ロイシン系はメチオニンとイソロイシンにこの転移後編集機能を発揮することが知られており、非天然アミノ酸にも同様に機能する場合がある。

20

【0287】

まず、編集ドメインを除去した。編集ドメインを6個のタンデムランダムアミノ酸のライブラリーで置換した。βラクターゼ中の終止コドンの抑圧に基づくポジティブ選択を使用した。多数の機能的シンテターゼが得られたが、これらのシンテターゼを精製しようと試みたが、いずれの場合も何も検出することができず、これらの全シンテターゼは温度感受性表現型を示し、編集ドメインの欠失の結果として不安定なタンパク質となったことが示唆された。

30

【0288】

次に、編集ドメインに点突然変異を行った。編集ドメインの触媒コアは種間で十分に保存されており、少なくとも分枝鎖アミノ酸の類では異なるアミノ酸でも保存されている。これらの保存部位の数箇所は従来突然変異されており（例えばT→P）、編集機能を欠損することが分かっている。数種の公知突然変異体に類似する*Methanobacterium thermoautotrophicum* RSの突然変異体を構築すると共に、T214に由来する20メンバーNNKライブラリーも作製した。タンパク質を発現させ、ロイシンとメチオニンによるアミノアシル化を*in vitro*試験した。従来同定されている突然変異のうちで本発明の系に転移可能なものは皆無であったが、T214ライブラリーから望ましい突然変異が同定された。ロイシンを負荷することが可能な2個の突然変異体T214S及びT214Qが同定された。これらの突然変異体のうちでT214Qのみはメチオニンを負荷することができた。T214Sは明らかにメチオニンを削除する機能を維持しているが、Gln突然変異体はこの機能を失っていた。

40

【0289】

次に*Thermus thermophilus*ロイシルシンテターゼについて解明されている結晶構造に基づいてライブラリーを設計した。ロイシンアミノアルキルアデニル酸類似体アデノシンインヒビターのロイシン側鎖を活性部位に結合した。ロイシン側鎖結合

50

ポケットの周囲の6部位をランダムアミノ酸で置換し、より大きいライブラリーを作製した。その後、例えばポジティブ／ネガティブ2段階選択を行うことにより、このライブラリーからのシンテターゼをスクリーニングすると、非天然アミノ酸を選択的に負荷することが可能なシンテターゼを同定することができる。

【実施例6】

【0290】

実施例6-4塩基コドンを効率的に抑圧するtRNAの同定

【0291】

4塩基コドンで効率的に抑圧する突然変異tRNAを同定するために併用アプローチを使用した。T. J. Magliery, J. C. Anderson and P. G. Schultz, J. Mol. Biol., 307:755 (2001) 参照。βラクタマーゼ遺伝子のセリンコドンを4個のランダムヌクレオチドで置換したレポーターライブラリーを構築した。次に、アンチコドンループ(7nt)を8又は9個のランダムヌクレオチドで置換した大腸菌の誘導体から構成される突然変異tRNA(例えばサプレッサーtRNA)サプレッサーライブラリーを作製した。これらの2つのライブラリーを交配すると、4塩基配列を単一コドンとしてデコードする適当なフレームシフトサプレッサーtRNAが全長βラクタマーゼを翻訳し、細胞はアンピシリン耐性となった。より高いアンピシリン濃度での生存は、対応するtRNAが4塩基コドンに対してより高い抑圧効率をもつことを意味する。この選択を使用し、天然三重項コドンサプレッサーに近い効率をもつカノニカル4塩基コドンのみをデコードする4種の四重項コドンAGGA、CUAG、UAGA及びCCCUとそのコグネイトサプレッサーtRNAを同定した。この戦略を使用して新規5及び6塩基コドンサプレッサーも選択した。Anderson, Magliery, Schultz, Exploring the Limits of Codon and Anticodon Size, Chemistry & Biology, 9:237-244 (2002) 参照。新規に同定されたものも含めてこれらの拡張コドンは多重非天然アミノ酸のin vitro組込みとin vivoタンパク質突然変異誘発に有用であると思われる。

【実施例7】

【0292】

実施例7-pアミノフェニルアラニンの直交tRNAシンテターゼの作製

【0293】

pアミノフェニルアラニン(pAF)の直交シンテターゼを作製するために、*Methanococcus jannaschii*チロシルtRNAシンテターゼ(TyrRS)と突然変異体チロシンアンバーサプレッサーtRNA(TyrCUAmutRNA)の対を出発点として使用した。Wang, L., Magliery, T. J., Liu, D. R. & Schultz, P. G. A new functional suppressor tRNA/aminoacyl-tRNA synthetase pair for the in vivo incorporation of unnatural amino acids into proteins. J. Am. Chem. Soc. 122:5010-5011 (2000); 及びWang, L. & Schultz, P. G. Chem. and Biol. 8:883 (2001) 参照。pAFを受容し且つ20種の共通アミノ酸のいずれをも受容しないように*Methanococcus jannaschii* TyrRSのアミノ酸特異性を改変することによりpAF特異的シンテターゼ(pAFRS)を作製した。ポジティブ選択とネガティブスクリーニングの組合せを使用し、5カ所(Tyr³²、Glu¹⁰⁷、Asp¹⁵⁸、Ile¹⁵⁹及びLeu¹⁶²)にランダムアミノ酸を含むTyrRS変異体12のライブラリーからpAFRS酵素を同定した。Wang, L., Brock, A., Herberich, B. & Schultz, P. G. Expanding the genetic code of Escherichia coli. Science 292:498-500 (2001) 参照。選択とスクリーニングの両者に単一レポータープラスミドを使用した。

例えば、レポータープラスミドはCAT、T7 RNAポリメラーゼ、GFP及び^{TyrCUA}mutRNAの遺伝子とTet耐性選択マーカを含むpREP(2)/YC-JYCUAである。CAT遺伝子はD112位にTAGコドン置換を含む。T7 RNAポリメラーゼ遺伝子は7アミノ酸N末端リーダーペプチドを含み、M1とQ107にTAG置換を含む。

【0294】

ポジティブ選択はpAF又は内在アミノ酸によるクロラムフェニコールアセチルトランスフェラーゼ(CAT)遺伝子内の許容位置のTAGコドンの抑圧に基づいて行う。例えばWang, L. ら, (2001)前出; 及びPasternak, M., Maglieri, T. J. & Schultz, P. G. A new orthogonal suppressor tRNA/aminoacyl-tRNA synthetase pair for evolving an organism with an expanded genetic code. Helvetica Chimica Acta 83:2277(2000)参照。pAFを含む液体培地でTyrRSライブラリーとレポータープラスミドを含む細胞を増殖させ、クロラムフェニコール(Cm)の存在下で生存する細胞を選択した。例えば、ポジティブ選択では、35 µg/ml Kn、25 µg/ml Tet、75 µg/ml Cm及び1 mM pAF(Sigma)を加えたGMM最少培地で細胞を増殖させた。

【0295】

ネガティブスクリーニングはpAFの不在下でT7 RNAポリメラーゼ遺伝子内の許容位置の2個のTAG終止コドンを抑圧できないことに基づいて行う。全長T7 RNAポリメラーゼの発現はグリーン蛍光タンパク質(GFP)の発現を誘導する。ポジティブ選択からの細胞をpAFとCmの不在下で増殖させた後、蛍光活性化細胞ソーティング(FACS)を使用して蛍光の欠如をスクリーニングする。例えば、ネガティブスクリーニングでは、35 µg/ml Kn、25 µg/ml Tet及び0.002%アラビノースを加えたGMM培地で細胞を増殖させた。Coherent Enterprise I IオンレーザーでBDIS FAVantage TSO細胞ソーターを使用してFACSを実施した。励起波長は351 nmとし、575/25 nm帯域フィルターで発光を検出した。細胞を集め、TetとKnを加えたLB少なくとも10容量で希釈し、飽和まで増殖させた。

【0296】

液体培地でのポジティブ選択2回と、ネガティブスクリーニング1回と、更に1回の液体培地でのポジティブ選択と、プレート上でのポジティブ選択1回の後に所望pAFRSが同定された。pAFRS酵素は野生型TyrRSに対して5個の突然変異を含む(Y32T, E107T, D158P, I159L及びL162A)。pAFの不在下では、選択したpAFRSとレポータープラスミドを発現する細胞のIC₅₀はGMM最少培地プレートで10 µg/ml Cmであった。1 mM pAFの存在下ではIC₅₀は120 µg/ml Cmであった。従って、pAFはUAGコドンを選択的に抑圧している。

【実施例8】

【0297】

実施例8—蛍光活性化細胞ソーティングを使用したアミノアシルtRNAシンテターゼの進化

【0298】

FACに基づくスクリーニングシステムを使用してアミノ、イソプロピル又はアリル含有チロシン類似体を受容する3種の高度選択性シンテターゼ変異体を迅速に進化させた。このシステムでは多目的レポータープラスミドを使用してポジティブ及びネガティブ両者の選択圧をかけ、シンテターゼ活性を容易且つ定量的に評価した。クロラムフェニコールアセチルトランスフェラーゼ(CAT)マーカーによりM. jannaschiiチロシル-tRNAシンテターゼ(TyrRS)の活性のポジティブ選択を行った。T7ポリメラーゼ/GFPレポーター系により、非天然アミノ酸の存在下と不在下で増殖させた細胞内

のシンテターゼ活性を評価した。蛍光活性化細胞ソーティング (FACS) を使用して天然アミノ酸を受容するシンテターゼ変異体をスクリーニングし、目視とフルオロメトリー分析によりシンテターゼ活性を夫々定性及び定量的に評価した。

【0299】

増幅性蛍光レポーター系の設計。生きた細胞でシンテターゼ活性を評価する汎用スクリーニングシステムを開発する試みは、シンテターゼ変異体集団に適用する選択圧、特にネガティブ選択圧をより精密に制御したいという要望に端を発した。多数のシンテターゼ変異体の活性を評価するためにシステムを使用しなければならないので、高スループットスクリーニングが可能なレポーターが追求された。更に、シンテターゼ活性を容易に定性及び定量評価できるようなレポーターが必要であった。これらの要件を満たすために、蛍光スクリーニングが考案された。このシステムは大腸菌での発現に最適化されたグリーン蛍光タンパク質の変異体である GFPuv のシンテターゼ依存性生産に基づいて考案された (Cramer, A., Whitehorn, E. A., Tate, E. & Stemmer, W. P., Nature Biotechnol. 1996, 14, 315-319 参照)。この蛍光体は FACS とフルオロメトリーで使うことができ、更に固体培地と液体培地で目視にも利用できる。このシステムはセクター (例えばアンバーナンセンスコドン) のシンテターゼ依存性抑圧により蛍光シグナルを発生するように設計された。レポーターの感度を最大にするために、他の増幅可能な細胞内レポーター系と同様に GFPuv レポーター遺伝子の発現を誘導するように構成された T7 RNA ポリメラーゼ遺伝子内にアンバーコドン配置することにより増幅可能にした (Lorincz, M., Roederer, M., Diwu, Z., Herzenberg, L. A., Nolan, G. P., Cytometry, 1996, 24, 321-329; 及び Zlokarnik, G., Negulescu, P. A., Knapp, T. E., Mere, L., Burrell, N., Feng, L., Whitney, M., Roemer, K. & Tsien, R. Y., Science, 1998, 279, 84-88 参照)。アンバーコドンを含む T7 RNA ポリメラーゼに合うように RNA 転写産物の生産を容易に最適化できるようにするために、T7 RNA ポリメラーゼ遺伝子をアラビノースプロモーターの制御下においた。

【0300】

T7 RNA ポリメラーゼ / GFPuv レポーター系の最適化。アラビノースプロモーターの制御下にアンバーコドンを含む T7 RNA ポリメラーゼ変異体を発現させると共に T7 プロモーターの制御下に GFPuv 遺伝子を発現させるように中コピーレポータープラスミド pREP を設計した (図 17a)。シンテターゼ依存性蛍光増加を最適化するように設計した 12 種の一連の T7 RNA ポリメラーゼ変異体 (図 17b) を pREP に挿入してプラスミド pREP (1~12) を作製した。全変異体は 7 アミノ酸 (MTMITVH) の N 末端リーダー配列と 1~3 個のアンバー終止コドン (TAG) を含んでいた。変異体 1~3 はリーダー配列のすぐ下流の 1 位の元のメチオニン (M1) を夫々 1、2 及び 3 個のアンバー終止コドンで置換していた。変異体 4~9 は構造のループ領域に位置すると予想される D10、R96、Q107、A159、Q169 又は Q232 が夫々アンバーコドンで置換していた (Jeruzalmi, D. & Steitz, T. A., EMBO J., 1998, 17, 4101-4113)。変異体 10~12 は M1 位と Q107、A159 又は Q232 位が夫々アンバー終止コドンで置換していた。直交グルタミニル tRNA シンテターゼと S. cerevisiae 由来グルタミン tRNA_{CUA} を含む適合可能なプラスミドを使用して生きた細胞のフルオロメトリーとフローサイトメトリーによりプラスミド構築物の蛍光増加を検討した。プラスミド pREP (1~12) は各種レベルのシンテターゼ依存性蛍光増加を生じることが判明し、最良構築物 pREP (10) は不活性突然変異体に比較して野生型シンテターゼを含む細胞でフルオロメトリーにより 220 倍の蛍光 (図 17c) を示し、サイトメトリーにより ~400 倍の蛍光中央値 (図 17d) を示した。pREP (10) 内のアンバーコドンに対応する位置の各種官能基を置換した処、T7 RNA ポリメラーゼ内の 107 位の許容性が高いことが判明

した。

【0301】

多目的レポータープラスミドの構築。M. jannaschii TyrRSの変異体を選択及びスクリーニングするために使用する多目的プラスミドを構築するために、プラスミドpREP(10)をプラスミドpYC-J17(Wang, L., Brock, A., Herberich, B. & Schultz, P. G., Science, 2001, 292, 498-500参照)と融合し、pREP/YC-JUCUA(図25a)を得た。O-メチル-L-チロシン(OMY)の組込みに選択的なM. jannaschii TyrRSの変異体を発現する適合可能なプラスミド(pBK-mJYRS; Wang, L., Brock, A., Herberich, B. & Schultz, P. G., Science, 2001, 292, 498-500)を使用してプラスミドpREP/YC-JYCUAの機能をアッセイした。pREP/YC-JYCUAとpBK-mJYRSを導入した細胞をOMYの存在下に増殖させた処、プラスミドpYC-J17を使用して得られると同一の120 $\mu\text{g}/\text{ml}$ のクロラムフェニコール(Cm)IC₅₀値を示し、フルオロメトリーで測定した場合にOMYの存在下に増殖させた細胞の蛍光強度は不在下の330倍であった。

【0302】

M. jannaschii チロシルtRNAシンテターゼの基質特異性の進化。チロシンのフェノール側鎖以外の化学基を選択的に受容するようにM. jannaschii TyrRSのアミノ酸側鎖結合ポケットを進化させることができることが報告されている(Wang, L., Brock, A., Herberich, B. & Schultz, P. G., Science, 2001, 292, 498-500; Wang, L., Brock, A., & Schultz, P. G., J. Am. Chem. Soc. 2002, 124, 1836-1837参照)。本発明者らはp-イソプロピルフェニルアラニン(pIF)、p-アミノフェニルアラニン(pAF)、p-カルボキシルフェニルアラニン(pCF)又はO-アリルチロシン(OAT)の4種の新規官能基(図25b)を酵素に受容させることによりM. jannaschii TyrRSによる非天然アミノ酸組込みの一般性を更に検討しようと試みた。チロシル環のパラ位に結合ポケットを形成すると考えられる残基であるY32、E107、D158、I159及びL162位にランダム変異をもつM. jannaschii TyrRS変異体のライブラリー(Wang, L., Brock, A., Herberich, B. & Schultz, P. G., Science, 2001, 292, 498-500)を、プラスミドpREP/YC-JYCUAを含む細胞に導入した。ライブラリーダイバーシティ $\sim 10^9$ をもつこれらの細胞を使用してpIF、pAF、pCF又はOATに選択的なシンテターゼ変異体を同定する4種の進化実験を開始した(図25b)。Cmと4種の非天然アミノ酸の1種の存在下に細胞培養液を飽和まで増殖させることによりポジティブ選択を2サイクル行った。2サイクル目のポジティブ選択後に細胞アリコートを取り出し、アミノ酸又はCmを加えない新しい培養液に接種するために使用し、培養液を再び飽和まで増殖させた。この時点で蛍光発光する細胞は20種の天然アミノ酸の1種を受容できるシンテターゼ変異体を含むと思われる。各株から約 10^8 個の細胞をFACSによるネガティブスクリーニングにかけ、天然アミノ酸を受容するシンテターゼ変異体を排除した。蛍光発光しない細胞を集め、飽和まで増殖させることにより増幅した。これらの増幅細胞を使用して新規培養液に接種し、非天然アミノ酸とCmを加えた液体培地で最終サイクルのポジティブ選択にかけた。飽和まで増殖後に0、30、60又は100 $\mu\text{g}/\text{mL}$ Cmと適当な非天然アミノ酸0又は1 mMを加えた培地に各細胞集団を播種した。

【0303】

進化シンテターゼ変異体の同定と特性決定。pIF、pAF及びOATを加えたCmプレートはアミノ酸を加えないプレートの10~100倍の数の蛍光コロニーを生じた。これに対して、pCF集団のプレートはpCF添加の有無に拘わらず同一数の蛍光コロニーを生じた。非天然アミノ酸を加えたプレートからのpIF、pAF及びOAT集団の各々に

ついて大きいほうから10個の蛍光コロニーを釣菌し、非天然アミノ酸の存在下又は不在下に液体培地で飽和まで増殖させた。携帯長波長紫外線ランプを使用して蛍光発生の定性評価を目視により行った(図23a)。

【0304】

蛍光に有意差を生じるクローンに対応するシンテターゼ変異体を配列決定した。pIF及びpAF集団からの全10個のクローンは同一配列であったが、OAT集団からは3種の異なるクローンが同定された。全クローンで5箇所のランダム突然変異部位にアミノ酸置換があり、pIF-tRNAシンテターゼ(pIF-RS)変異体には更に2箇所の置換があった。各クローンの活性を定量評価した。非天然アミノ酸の存在下又は不在下に液体培地で増殖させた細胞の蛍光をフルオロメトリーにより測定した(図23b)。非天然ア

10

【0305】

4位にアンバーコドンを含むミオグロブリン遺伝子を使用して非天然アミノ酸を含むタンパク質の生産を評価した。pIF、pAF又はOATの存在下又は不在下にpIF-RS、pAF-RS又はOMT-RS変異体を夫々使用して遺伝子を細胞で発現させた(図23d)。タンパク質収率は全3種の変異体で同等であり、非天然アミノ酸含有細胞培養液1リットル当たりタンパク質1~2mgであった。他方、非天然アミノ酸の不在下に増殖させた培養液ではタンパク質生産はほとんど検出できなかった。タンパク質をエレクトロスプレー質量分析により分析した処、夫々の質量はpIF含有タンパク質18457.40±0.81(予想値18457.28)、pAF含有タンパク質18430.30±0.27(予想値18430.21)であった。Cm^rIC₅₀、フルオロメトリー及びタンパク質発現分析を使用して得られた活性測定値はよく相関していたが、pIF-RSの活性はフルオロメトリーによると多少低いようである。他のアッセイに比較して不釣り合いに低いpIF-RS変異体のフルオロメトリー測定値は、レポーター内のアンバー位置の見掛けの許容性にも拘わらずT7 RNAポリメラーゼがpIF類似体の組込み後に部分的に不安定になる可能性を示している(図17c参照)。

20

【0306】

多目的レポーター系の有用性。本明細書に記載するレポーター系はポジティブ選択とネガティブスクリーニングの両者に単一の多目的プラスミドを使用できるので、ポジティブ選択とネガティブ選択で交互にプラスミドを交換する必要がなくなる。全体でたった3回のポジティブ選択と1回のネガティブスクリーニングで所望非天然アミノ酸を選択的に受容するシンテターゼ変異体を同定することができた。これらの特徴によりほんの数日間に進化実験を実施することができる。このスクリーニングシステムを使用すると、非天然アミノ酸を加えた寒天プレートを使用して活性シンテターゼ変異体を容易に同定することができ、変異体のアミノ酸特異性を個々にアッセイすることができる。

30

【0307】

上述のように、T7 RNAポリメラーゼ/GFP系はシンテターゼ変異体の活性を定量的に比較するために使用することができる。CATとバルナーゼに基づくポジティブ/ネガティブ選択を使用して本実施例に記載する3種のOAT-RSクローンと同一ライブラリーから独立して得られる別のOAT-RSクローンを入手できるので、各々に由来するシンテターゼ変異体について2種の異なる進化系を比較することが可能になる。この分析によると、ポジティブ選択とネガティブスクリーニング後に得られる3種のクローンはOATの存在下ではやや低い蛍光レベルを示すが、非天然アミノ酸の不在下では~10分の1のバックグラウンドレベルを示す。従って、非天然アミノ酸の不在下に対して存在下で増殖させた細胞の蛍光増加はバルナーゼを使用するポジティブ/ネガティブ選択後に得られるOAT-RSクローンを発現する細胞に比較して選択とスクリーニング後に得られるOAT-RS(1)を発現する細胞が約6倍となる。この例が代表例であるか否かははっきりしないが、これらのデータはT7 RNAポリメラーゼ/GFP系が天然アミノ酸基質に対して無差別なシンテターゼ変異体を選択する際にストリンジェンシーを高くできる

40

50

ことを示唆している。しかし、非天然アミノ酸は進化型シンテターゼ変異体との結合を競合し、天然アミノ酸の結合が減るので、非天然アミノ酸の不在下に比較して存在下で増殖させた細胞の蛍光が増加するのは非天然アミノ酸組込みの忠実度の限界が低くなることを意味すると予想される。更に、忠実度が高いことは明らかに望ましいが、所望用途に応じて忠実度と総シンテターゼ活性のどちらかをとらなければならないと思われる。

【0308】

アミノアシル tRNA シンテターゼ進化の一般性。従来の報告と本実施例に記載する結果から明らかなように、M. jannaschii TyrRS のアミノ酸側鎖結合ポケットは非常に適応性がある。チロシンのフェノール側鎖の代わりに各種官能基を受容するように酵素を進化させることができ、これは高い選択性で可能である。本願では、チロシン環のパラ位にヒドロキシルの代わりにアミン、イソプロピル又はアリルエーテル官能基を受容するように酵素を進化できることが立証された。pCF 非天然アミノ酸を受容することが可能な酵素変異体を同定することはできなかった。pCF アミノ酸を受容するようにシンテターゼを進化させる第2の試みも成功しなかった。LC/MS 分析を使用すると、非天然アミノ酸の存在下に増殖させた大腸菌細胞のトルエン処理後に pCF を検出することができず、pCF は細胞に輸送されないか又は導入後に代謝されることを示唆している。

10

【0309】

本実施例に記載した3種の進化実験の成功例のうちで2個以上の活性クローンを同定できたのは OAT-RS の進化のみであった。OAT-RS 進化は最高活性シンテターゼ変異体を生じた実験でもあった。これらの結果は、所定のアミノ酸特異性を他よりも選択し易いことを示唆している。これは20種の天然アミノ酸のコンテキストで各種非天然アミノ酸を選択的に認識するのは比較的困難であることが一因であると思われる。例えば、pAF は構造的及び電子的にチロシンと似ているために OAT よりも選択的に認識しにくいと思われる。このために、pAF-RS クローンよりも多数の OAT-RS クローンが同定され、pAF-RS クローンが最良の OAT-RS クローンよりも低活性であるのだと考えられる。

20

【0310】

プラスミド構築。プラスミド pREP (図17a) は T7 RNA ポリメラーゼ遺伝子を含む BamHI/ApaI オーバーラップ PCR フラグメントを rrnB 転写終結領域の上流に挿入した後に pBAD/Myc-HisA プラスミドからの araC 遺伝子及び ara プロモーター領域 (Invitrogen; T7 RNA ポリメラーゼ遺伝子の転写制御用) と GFPuv 遺伝子 (Clontech; T7 ターミネーター領域の上流で T7 プロモーターの下流) を含む ApaI/AhdI オーバーラップ PCR フラグメントをプラスミド pACYC177 (New England Biolabs) の AhdI/BamHI 部位間に挿入することにより構築した。プラスミド pREP (1~12) は記載位置にアンバー突然変異を含むオーバーラップ PCR フラグメントで T7 RNA ポリメラーゼの HpaI/ApaI フラグメントを置換することにより構築した。プラスミド pREP/YC-JYCUA は pREP (10) からの AfeI/SacII フラグメントと pYC-J17 (Wang, L., Brock, A., Herberich, B. & Schultz, P. G., Science, 2001, 292, 498-500) からの EarI (平滑化)/SacII フラグメントのライゲーションにより構築した。蛍光スクリーニング用プラスミド pQ を含む細胞に形質転換後に所望構築物を同定した。

30

40

【0311】

プラスミド pQ は lac プロモーター領域の下流で大腸菌 QRS 終結領域の上流に ScQRS を含む AatII/SalI オーバーラップ PCR フラグメントと、lpp プロモーター領域の下流で rrnC 終結領域の上流に S. cerevisiae tRNA (CUA)^{Gln} を含む SalI/AvaiI オーバーラップ PCR フラグメントと、pBR322 (New England Biolabs) の AvaiI/AatII フラグメントの三

50

重ライゲーションにより構築した。プラスミド p Q D は B a m H I と B g l I I の間の p Q フラグメントを S c Q R S (D 2 9 1 A) 突然変異体の B a m H I / B g l I I フラグメントで置換することにより構築した。

【0312】

プラスミド p B A D / J Y A M B - 4 T A G はプラスミド p Y C - J 1 7 (W a n g , L . , B r o c k , A . , H e r b e r i c h , B . & S c h u l t z , P . G . , S c i e n c e , 2 0 0 1 , 2 9 2 , 4 9 8 - 5 0 0) と、M j Y t R N A _{C U A} の遺伝子と p B A D プロモーターと挿入遺伝子のヘテロ発現用クロニング領域を含む p B A D / M y c - H i s A (I n v i t r o g e n) のハイブリッドである p B A D / Y C - J Y C U A プラスミドに C 末端 6 H i s タグを含むミオグロブリンの S 4 アンバー突然変異体の P C R フラグメントを挿入することにより構築した。

10

【0313】

フルオロメトリー及びサイトメトリー分析。所望プラスミドを含む単コロニーを使用し、適当な抗体と 0 . 0 0 2 % アラビノースと必要に応じて適当な非天然アミノ酸を加えた G M M L 培養液 2 m L に接種した。培養液を飽和まで増殖させ、細胞 (2 0 0 μ L) をペレット化してリン酸緩衝食塩水 (P B S) 1 m L に再懸濁した。6 0 0 n m 吸光度により細胞密度を分析し、F l u o r o M a x - 2 フルオロメーターを使用して励起波長 3 9 6 n m 、測定波長 5 0 5 n m で蛍光強度を測定した。P B S に懸濁した細胞をサイトメトリーにより分析した。p R E P (1 0) の T 7 ポリメラーゼ遺伝子内のアンバー位置の許容性を評価するために、レポータープラスミドを一連のサプレッサー株に形質転換した後、フルオロメトリーにより分析した。

20

【0314】

アミノアシル t R N A シンテターゼ変異体の進化。Y 3 2 、 E 1 0 7 、 D 1 5 8 、 I 1 5 9 及び L 1 6 2 位をランダム突然変異させた M . j a n n a s c h i i T y r R S 変異体 (W a n g , L . , B r o c k , A . , H e r b e r i c h , B . & S c h u l t z , P . G . , S c i e n c e , 2 0 0 1 , 2 9 2 , 4 9 8 - 5 0 0) を、p R E P / Y C - J Y C U A を含む D H 1 0 B 大腸菌細胞 (L i f e T e c h n o l o g i e s) に形質転換させ、ダイバーシティ ~ 1 0 ⁹ のライブラリーを作製した。形質転換細胞を S O C 培地で 6 0 分間 3 7 ° C で再生させ、L B 培地で飽和まで増殖させた。1 回目のポジティブ選択を開始するために、ライブラリー培養液 2 m L をペレット化して G M M L 培地に再懸濁し、2 5 μ g / m L テトラサイクリン (T e t) 、3 5 μ g / m L カナマイシン (K n) 及び 1 m M p I F 、p A F 、p C F 又は O A Y を加えた G M M L 5 0 0 m L に接種した。3 7 ° C で 3 時間培養後に終濃度 7 5 μ g / m L となるように C m を加え、細胞を飽和まで (~ 4 8 時間) 増殖させた。2 回目のポジティブ選択では、T e t 、K n 、7 5 μ g / m L C m 及び 1 m M p I F 、p A F 、p C F 又は O A Y を加えた G M M L 1 0 0 m L に 1 回目のポジティブ選択からの細胞 (5 0 0 μ L) を接種し、3 7 ° C で飽和まで (~ 2 4 ~ 3 6 時間) 増殖させた。ネガティブスクリーニングに備え、T e t 、K n 及び 0 . 0 2 % アラビノース (A r a) を加えた G M M L 培地 2 5 m L に 2 回目のポジティブ選択からの細胞 (1 0 0 μ L をペレット化して G M M L に再懸濁) を接種し、3 7 ° C で飽和まで (~ 2 4 時間) 増殖させた。A r a で誘導して非天然アミノ酸の不在下に増殖させた細胞 (1 m L) をペレット化し、リン酸緩衝食塩水 (P B S) 3 m L に再懸濁した。C o h e r e n t E n t e r p r i s e I I イオンレーザーを 3 5 1 n m で励起し、5 7 5 / 2 5 n m 帯域フィルターで発光を検出することにより B D I S F A C V a n t a g e T S O 細胞ソーターを使用して G F P u v を発現しない細胞を選別した。細胞を集め、T e t と K n を加えた L B 少なくとも 1 0 容量で希釈し、飽和まで増殖させた。3 回目のポジティブ選択を開始するために、ネガティブスクリーニングからの細胞 1 0 0 μ L をペレット化し、G M M L に再懸濁し、T e t 、K n 及び 1 m M p I F 、p A F 、p C F 又は O A Y を加えた G M M L 2 5 m L に接種した。3 時間 3 7 ° C で培養後に終濃度 7 5 μ g / m L となるように C m を加え、細胞を飽和まで (~ 2 4 時間) 増殖させた。3 回目のポジティブ選択後に、T e t 、K n 、0 . 0 0 2 % A r a 、0 . 7 5 又は 1 0 0 μ g / m L

30

40

50

Cm、及び0又は1mMpIF、pAF、pCF又はOAYを加えたGMM L／寒天に細胞を撒き、37℃で48時間増殖させた。

【0315】

非天然アミノ酸を含有するタンパク質の発現と特性決定。pBAD／JYAMB-4TAGと適当なpBKプラスミドで同時形質転換したDH10B細胞を使用し、KnとTetを加えたGMM Lスター培養液100mLに接種し、飽和まで増殖させた。Kn、Tet、0.002%Ara、5μM FeCl₃及び所望非天然アミノ酸（又はなし）を加えた培養液500mLにスター培養液50mLを接種し、飽和まで（～18時間）増殖させた。培養液をベレット化し、音波処理し、QiaExpress ionist（QiaGen）Hisタグ精製キットのプロトールに従ってミオグロブリンタンパク質を分離した。タンパク質を12→20%グラジエントSDSポリアクリルアミドゲル電気泳動とエレクトロスプレー質量分析により分析した。 10

【実施例9】

【0316】

実施例9－直交tRNA／スレオニルtRNAシンテターゼ対

【0317】

本実施例は直交tRNA／スレオニルtRNAシンテターゼ対の作製を例証する。図27はThermus thermophilusに由来するスレオニルtRNAシンテターゼを示す。このシンテターゼは2個のN末端編集ドメインと、触媒ドメインと、C末端アンチコドン結合ドメインをもつ（659アミノ酸）。T. thermophilusシンテターゼに由来する直交シンテターゼを作製するために、編集ドメインと、N1又はN1とN2をシンテターゼから除去し、N短縮T. thermophilus ThrRS（475アミノ酸）を作製した。このシンテターゼは同一触媒活性をもつが、校正活性をもたない。N短縮シンテターゼの活性をスクリーニングした。N短縮シンテターゼは大腸菌tRNAをアミノアシル化しなかった。 20

【0318】

T. thermophilus tRNA Thrは大腸菌スレオニルtRNAシンテターゼの基質であることが分かったので、T. thermophilus tRNA Thrを突然変異させて直交対を作製した。図28はtRNAに行った突然変異を示す。具体的には、C2G71をA2U71に突然変異させた。in vitro負荷実験によると、この突然変異体は大腸菌スレオニルtRNAシンテターゼの基質ではないが、T. thermophilusスレオニルtRNAシンテターゼの良好な基質である。アンバーサプレッサーtRNAを作製するために、突然変異C2G71→A2U71及びG34G35U36→C34G35U36を含む別の突然変異体も構築した。C2G71→A2U71に加えてアンチコドンループも変異させた他の突然変異体tRNAも作製し、TGA、ACA、ACAA、AGGA、CCCT、TAGA及びCTAG等の3及び4塩基コドンを抑圧した。これらの全tRNAは（A2U71をもつ）野生型tRNA Thrほど良好な基質ではなかったが、T. thermophilusスレオニルtRNAシンテターゼのアンチコドン結合部位を突然変異させることにより改良することができる。 30

【実施例10】

【0319】

実施例10－典型的O-tRNA及びO-RSの配列

【0320】

典型的O-tRNAは配列番号1～3及び／又はその相補的ポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸を含む。添付書類1、表5参照。同様に、典型的O-RSは配列番号35～66から構成される群から選択されるアミノ酸配列を含むポリペプチドと、配列番号4～34及びその相補的ポリヌクレオチド配列から構成される群から選択されるポリヌクレオチド配列を含む核酸によりコードされるポリペプチドから構成される群から選択されるポリペプチドを含む。 40

【0321】

当然のことながら、本明細書に記載する実施例及び態様は例示の目的に過ぎず、これらの記載に鑑みて種々の変形又は変更が当業者に想到され、このような変形又は変更も本願の精神及び範囲と特許請求の範囲に含む。本明細書に引用した全刊行物、特許及び特許出願は全目的でその開示内容全体を参考資料として本明細書に組込む。

【0322】

配列 番号	添付書類1 表5：配列	注	tRNA 又はRS
1		M. jannaschii mtRNA _{CUA} ^{Tyr}	tRNA
2		HLAD03；最適化アンバーサプレッサーtRNA	tRNA
3		HL325A；最適化AGGAフレームシフトサプレッサーtRNA	tRNA
4		突然変異体 TyrRS (LWJ16)	RS
5		p-iPr-PheRS	RS
6		p-NH ₂ -PheRS(1)	RS
7		p-NH ₂ -PheRS(2)	RS
8		p-NH ₂ -PheRS(3a)	RS
9		p-NH ₂ -PheRS(3b)	RS
10		O-Allyl-TyrRS(1)	RS
11		O-Allyl-TyrRS(3)	RS

配列	添付書類1	注	tRNA
----	-------	---	------

番号	表5：配列		又はRS
12		O-Allyl-TyrRS(4)	RS
13		p-Br-PheRS	RS
14		p-Az-PheRS(1)	RS
15		p-Az-PheRS(3)	RS
16		p-Az-PheRS(5)	RS
17		m-アシルフェニルアラニンをタンパク質 に取込むための突然変異体シンテターゼ (ケトン3-4)	RS
18		m-アシルフェニルアラニンをタンパク質 に取込むための突然変異体シンテターゼ (ケトン3-7)	RS
19		m-アシルフェニルアラニンをタンパク質 に取込むための突然変異体シンテターゼ (ケトン4-1)	RS

10

20

30

配列 番号	添付書類1 表5：配列	注	tRNA 又はRS
20		m-アシルフェニルアラニンをタンパク質	RS

40

		に取込むための突然変異体シンテターゼ (ケトン5-4)	
2 1		m-アシルフェニルアラニンをタンパク質 に取込むための突然変異体シンテターゼ (ケトン6-8)	R S
2 2		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ (OMe 1-6)	R S
2 3		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ (OMe 1-8)	R S
2 4		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ (OMe 2-7)	R S
2 5		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ (OMe 4-1)	R S

10

20

30

配列 番号	添付書類1 表5：配列	注	tRNA 又はRS
2 6		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ	R S

40

		(OMe 4-8)	
27		p-O-アリルチロシンをタンパク質に組むための突然変異体シンテターゼ (アリル)	RS
28		p-ベンゾイル-L-フェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ (p-BpaRS (H6))	RS
29		p-アジドフェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ (p-Az-PheRS (3))	RS
30		p-アジドフェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ (p-Az-PheRS (6))	RS
31		p-アジドフェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ (p-Az-PheRS (20))	RS

10

20

30

				配列	添付書類
1	注	tRNA			
番号	表5：配列			又はRS	
32		p-アジドフェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ		RS	

40

		(p-Az-PheRS (24))	
3 3		Archaeoglobus fulgidus ロイシル t R N A シンテターゼ (AFLRS)	R S
3 4		Methanobacterium thermoautorophicum ロ イシル t R N A シンテターゼ (MtLRS)	R S
3 5		突然変異体 TyrRS (LWJ16)	R S
3 6		TyrRS (SS12)	R S
3 7		p-iPr-PheRS	R S
3 8		p-NH ₂ -PheRS (1)	R S
3 9		p-NH ₂ -PheRS (2)	R S
4 0		p-NH ₂ -PheRS (3a)	R S
4 1		p-NH ₂ -PheRS (3b)	R S

10

20

30

配列 番号	添付書類 1 表 5 : 配列	注	t R N A 又は R S
4 2		0-アリル-TyrRS (1)	R S
4 3		0-アリル-TyrRS (3)	R S

40

4 4		0-アリル-TyrRS (4)	R S
4 5		p-Br-PheRS	R S
4 6		p-Az-PheRS (1)	R S
4 7		p-Az-PheRS (3)	R S
4 8		p-Az-PheRS (5)	R S
4 9		m-アシルフェニルアラニンをタンパク質 に取込むための突然変異体シンテターゼ (ケトン 3-4)	R S
5 0		m-アシルフェニルアラニンをタンパク質 に取込むための突然変異体シンテターゼ (ケトン 3-7)	R S

10

20

配列 番号	添付書類 1 表 5 : 配列	注	tRNA 又はRS
5 1		m-アシルフェニルアラニンをタンパク質 に取込むための突然変異体シンテターゼ (ケトン 4-1)	R S
5 2		m-アシルフェニルアラニンをタンパク質 に取込むための突然変異体シンテターゼ	R S

30

40

		(ケトン5-4)	
53		m-アシルフェニルアラニンをタンパク質 に取込むための突然変異体シンテターゼ (ケトン6-8)	RS
54		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ (OMe1-6)	RS
55		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ (OMe1-8)	RS
56		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ (OMe2-7)	RS

10

20

配列 番号	添付書類1 表5:配列	注	tRNA 又はRS
57		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ (OMe4-1)	RS
58		m-メトキシフェニルアラニンをタンパク 質に取込むための突然変異体シンテターゼ (OMe4-8)	RS

30

40

59		p-O-アシルチロシンをタンパク質に組み込むための突然変異体シンテターゼ (アシル)	RS
60		p-ベンゾイル-L-フェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ (p-BpaRS (H6))	RS
61		p-アジドフェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ (p-Az-PheRS (3))	RS
62		p-アジドフェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ (p-Az-PheRS (6))	RS

配列 番号	添付書類1 表5：配列	注	tRNA 又はRS
63		p-アジドフェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ (p-Az-PheRS (20))	RS
64		p-アジドフェニルアラニンを取込むためのアミノアシル tRNA シンテターゼ (p-Az-PheRS (24))	RS

10

20

30

40

6 5		Archaeoglobus fulgidusロイシル tRNA シンテターゼ(AFLRS)	RS
6 6		Methanobacterium thermoautotrophicumロ イシル tRNAシンテターゼ(MtLRS)	RS

10

【図面の簡単な説明】

【0323】

【図1】非天然アミノ酸の部位特異的 in vivoタンパク質組込みを模式的に示す。

【図2A】非天然アミノ酸でアミノアシル化する活性シンテターゼの選択方法の例を模式的に示す。図2Aは非天然アミノ酸特異性をもつアミノアシル tRNAシンテターゼの選択／スクリーニングの概要を示す。

【図2B】非天然アミノ酸でアミノアシル化する活性シンテターゼの選択方法の例を模式的に示す。図2BはO-tRNAを非天然アミノ酸で優先的にアミノアシル化するシンテターゼの選択／スクリーニングの1態様を模式的に示す。

【図3】チロシン類似体の特異的ライブラリーを作製するための部位特異的突然変異を示す。 20

【図4】前記類似体の特異的ライブラリーを作製するためのペンタフルオロフェニルアラニン選択用コンセンサス配列を示す。

【図5】1種の生物（例えば大腸菌）に由来する1個のドメイン（例えばCPIドメイン）を他の生物のシンテターゼ（例えばMethanococcus jannaschii TyrRS）に移植する方法を模式的に示す。

【図6】キメラMethanococcus jannaschii／大腸菌シンテターゼの構築を模式的に示す。

【図7】キメラシンテターゼ（例えばMethanococcus jannaschii／大腸菌シンテターゼ）のライブラリーの作製を模式的に示す。 30

【図8】内在シンテターゼ（例えば大腸菌シンテターゼ）の弱い基質であり、効率的に所期コグネイトシンテターゼにより負荷されるサプレッサー tRNAの選択の1例を模式的に示す。

【図9】図9A及びBは突然変異アンチコドンループ tRNAライブラリーを模式的に示す。図9Aは突然変異全ループライブラリーを示し、図9BはMethanococcus jannaschii tRNA TyrCUAのライブラリーを示す。

【図10】水素結合以外の力により対合する非天然塩基対（PICS：PICS，3MN：3MN，7AI：7AI，Dipic：Py）の構造の例を模式的に示す。

【図11】サプレッサー tRNAのネガティブ選択法の結果のグラフであり、所定時間における3種の構築物のうちの1種を含む生存細胞の百分率を示す。 40

【図12】βラクタマーゼ遺伝子中のアンバーコードンの抑圧に基づくポジティブ選択を示す増殖ヒストグラム。

【図13】アンチコドンループ及び全ループライブラリーから選択した突然変異体サプレッサー tRNAのDNA配列を示す。

【図14】TyrRSの活性部位の立体図を模式的に示す。

【図15】TyrRSの活性部位を模式的に示す。

【図16】図16A及びBは本発明の成分（例えば直交シンテターゼ）を作製するために使用したFACS選択及びスクリーニング法の1例を模式的に示す。図16Aは直交シンテターゼライブラリーとO-tRNAの発現ベクター（ライブラリープラスミド）と、1個以上のセクターコードン（例えばTAG）を含むT7RNAポリメラーゼ／GFプレポ 50

ーター系（レポータープラスミド）のベクター（例えばプラスミド）を模式的に示す。図 1 6 B はポジティブ選択／ネガティブスクリーニングスキームを模式的に示す。

【図 1 7 A】増幅性蛍光レポーター系を示す。図 1 7 A はスクリーニングに使用することができるベクター（例えば、p R E P 等のプラスミド）を模式的に示す。

【図 1 7 B】増幅性蛍光レポーター系を示す。図 1 7 B は p R E P（1～12）内の T 7 R N A ポリメラーゼ遺伝子構築物の組成と蛍光増加を示す。

【図 1 7 C】増幅性蛍光レポーター系を示す。図 1 7 C は p R E P（10）と p Q D（上段）又は p Q（下段）を含む細胞のサイトメトリー分析を示す。

【図 1 7 D】増幅性蛍光レポーター系を示す。図 1 7 D は p R E P（10）を含み、各種大腸菌サプレッサー t R N A を発現する細胞のフルオロメトリー分析を示す。

10

【図 1 8】表面エピトープへの非天然アミノ酸の組み込みのファージ選択を模式的に示す。

【図 1 9】非天然アミノ酸特異性をもつ提示されたシンテターゼ（例えばファージに提示されたシンテターゼ）をスクリーニングするために使用する分子（例えばアミノアシルアデニル酸中間体の固定化アミノアルキルアデニル酸類似体）の 1 例を模式的に示す。

【図 2 0】各種生物に由来する各種直交対のアンピシリン耐性を示すグラフ。本図はセレクトターコドン（例えばアンバーコドン）を含むレポーター遺伝子（例えばβラクタマーゼ遺伝子）と（セレクトターアンチコドンを含む）サプレッサー t R N A を各々含むレポーター構築物を使用した直交対の検出例を示す。

【図 2 1】図 2 1 A 及び B はロイシル t R N A のアンチコドンループを突然変異（例えばランダム突然変異）させ、例えば選択段階の組合せ（例えばβラクタマーゼに基づく選択とバルナーゼに基づく選択）を使用してセレクトターコドン（例えばレポーター遺伝子中のアンバーコドン）のより効率的な抑圧について選択（図 2 1 B）することにより作製した *Halobacterium* N R C - 1 由来突然変異アンバーサプレッサー t R N A を示す。

20

【図 2 2】基本コドンの t R N A サプレッサーを示す。

【図 2 3】図 2 3 A～D は成功した各進化実験からの優性シンテターゼ変異体の活性を示す。図 2 3 A は p R E P / Y C - J Y C U A と指定シンテターゼ変異体を含む細胞を対応する非天然アミノ酸の存在下（+）又は不在下（-）で増殖させ、長波長紫外線を照射した場合の写真である。図 2 3 B は p R E P / Y C - J Y C U A と指定シンテターゼ変異体を含む細胞を対応する非天然アミノ酸の存在下（左）又は不在下（右）で増殖させた場合のフルオロメトリー分析を示す。図 2 3 C は p R E P / Y C - J Y C U A と指定シンテターゼ変異体を含む細胞を対応する非天然アミノ酸の存在下又は不在下で増殖させた場合の C m I C₅₀ 分析を示す表である。図 2 3 D は p B A D / J Y A M B - 4 T A G と指定シンテターゼ変異体を含む細胞を対応する非天然アミノ酸の存在下（+）又は不在下（-）で増殖させた場合のタンパク質発現分析を示す。

30

【図 2 4】ネガティブ F A C S スクリーニング（O A Y - R S（1, 3, 5））又はネガティブバルナーゼ選択（O A Y - R S（B））を使用して得られた O A S - R S 変異体の活性比較を示す。

【図 2 5】図 2 5 A～B は *M. jannaschii* T y r R S の進化を誘導するための多目的レポータープラスミド系の成分を示す。図 2 5 A はプラスミド p R E P / Y C - J Y C U A を示す。図 2 5 B は *M. jannaschii* T y r R S の進化のターゲットとして使用した非天然アミノ酸の構造を示す。

40

【図 2 6】プラスミド p R E P / Y C - J Y C U A を使用してアミノアシル t R N A シンテターゼを進化させるためのストラテジーを示す。

【図 2 7】*Thermus thermophilus* に由来するスレオニル t R N A シンテターゼを示す。

【図 2 8】*T. thermophilus* 直交スレオニル t R N A / R S のための直交 t R N A の作製を示す。

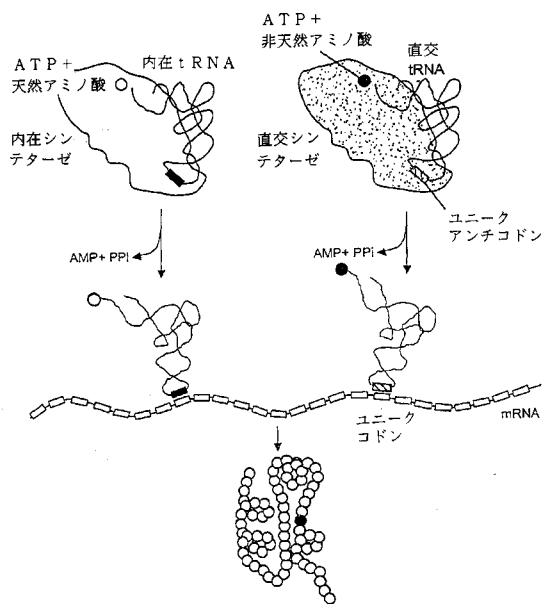
【図 2 9】本発明で利用される典型的非天然アミノ酸を示す。

【図 3 0】本発明で利用される典型的非天然アミノ酸を示す。

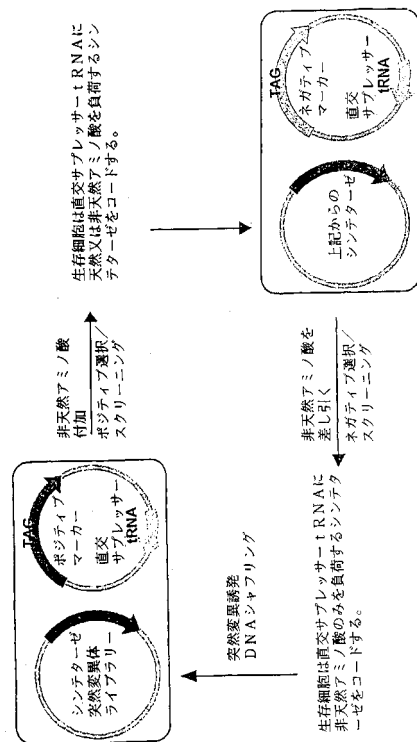
50

【図 3 1】 本発明で利用される典型的非天然アミノ酸を示す。

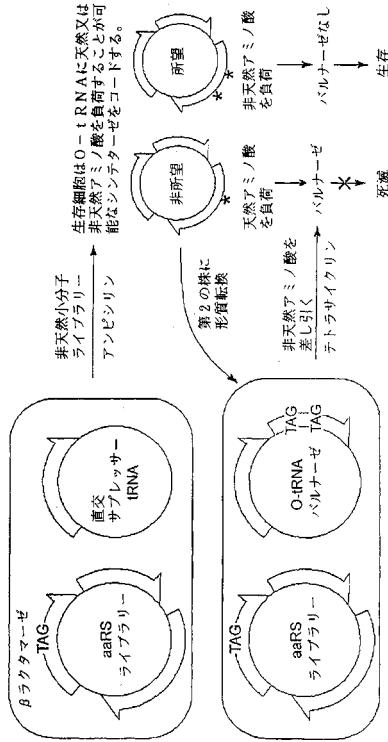
【図 1】



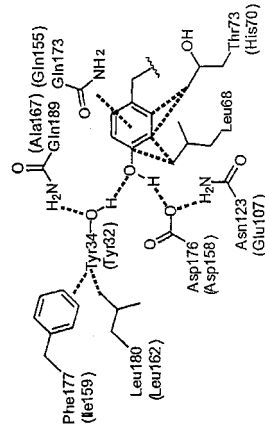
【図 2 A】



【図 2 B】



【図 3】



野生型

32 67 70 107 155 160 167

Tyr -- Ala Asp Leu His -- Glu -- Gln Val Asn Asp Ile His Tyr Leu Gly Val Asp Val Ala

ライブラリー-1 NNN -- Ala Asp Leu His -- NNN -- Gln Val Asn NNN NNN His Tyr NNN Gly Val Asp Val Ala

ライブラリー-2 NNN -- Ala Asp Leu His -- Glu -- Gln Val Asn NNN NNN His Tyr NNN Gly Val Asp Val Ala

ライブラリー-3 NNN -- NNN Asp Leu NNN -- Glu -- NNN Val Asn NNN Ile His Tyr Leu Gly Val Asp Val NNN

【図 4】

野生型

32 67 70 107 155 160 167

Tyr -- Ala Asp Leu His -- Glu -- Gln Val Asn Asp Ile His Tyr Leu Gly Val Asp Val Ala

ライブラリー-1 NNN -- Ala Asp Leu His -- NNN -- Gln Val Asn NNN NNN His Tyr NNN Gly Val Asp Val Ala

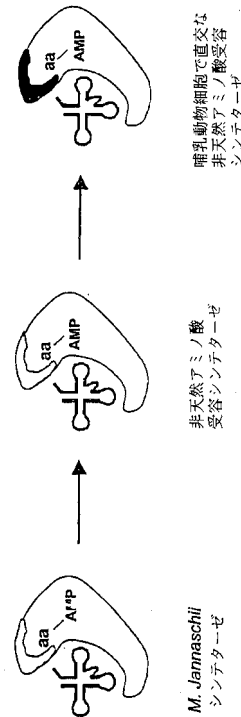
ライブラリー-2 NNN -- Ala Asp Leu His -- Glu -- Gln Val Asn NNN NNN His Tyr NNN Gly Val Asp Val NNN

ライブラリー-3 NNN -- NNN Asp Leu NNN -- Glu -- NNN Val Asn NNN Ile His Tyr Leu Gly Val Asp Val NNN

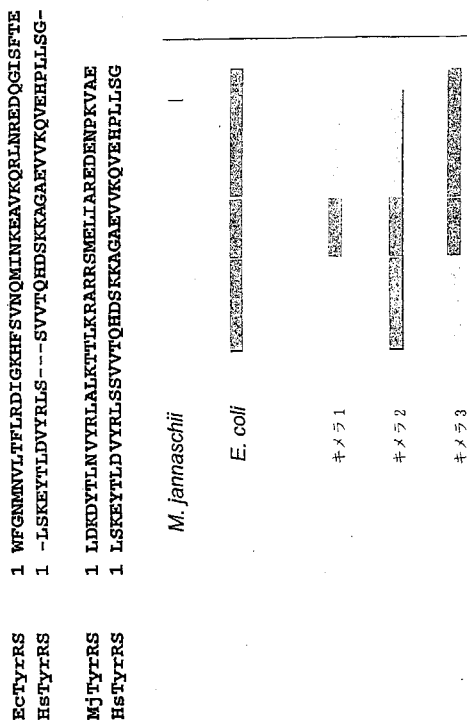
コンセンサス配列

Ile -- Gln Asp Leu Tyr -- Glu -- Ala Val Asn Ala Ile His Tyr Leu Gly Val Asp Val Leu

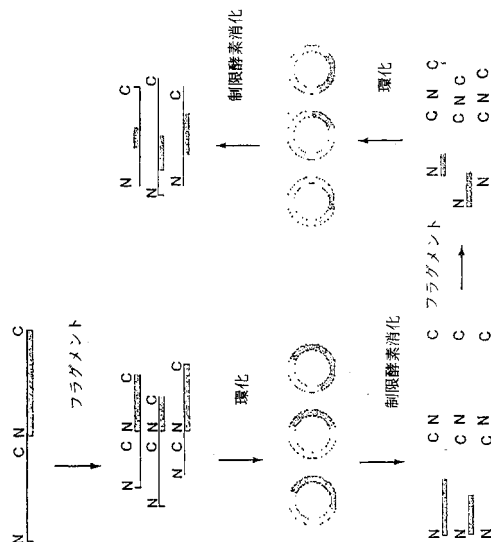
【図 5】



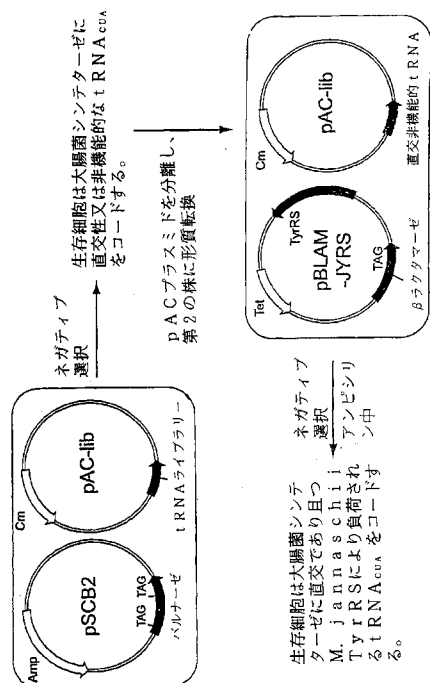
【図 6】



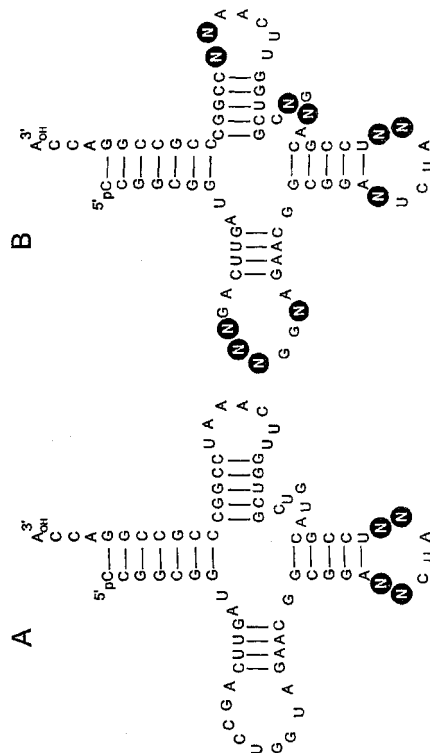
【図 7】



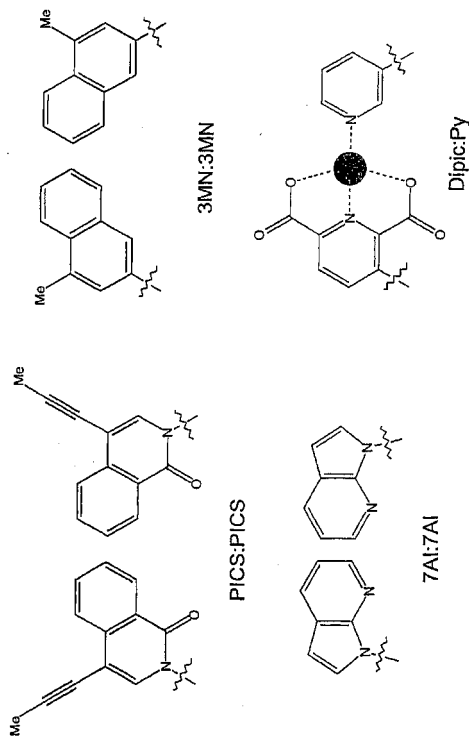
【図 8】



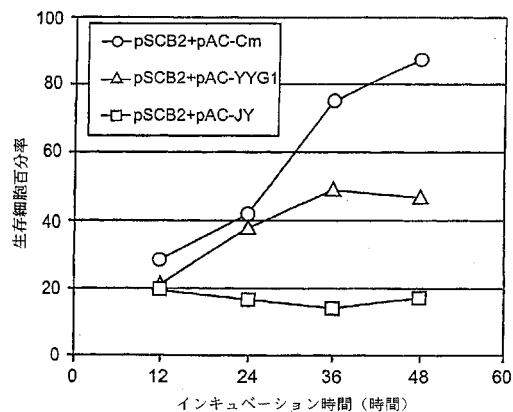
【図 9】



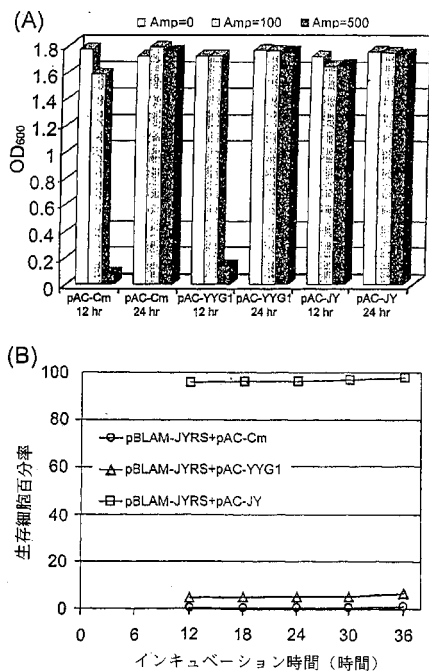
【図 10】



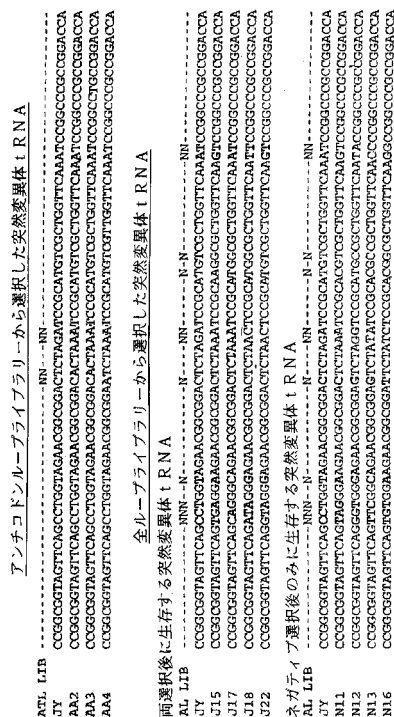
【図 11】



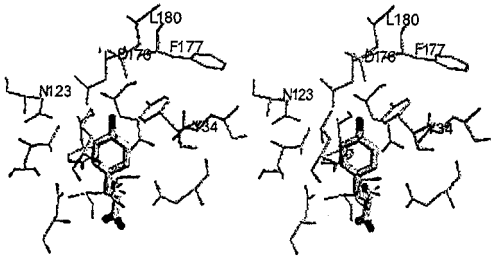
【図 12】



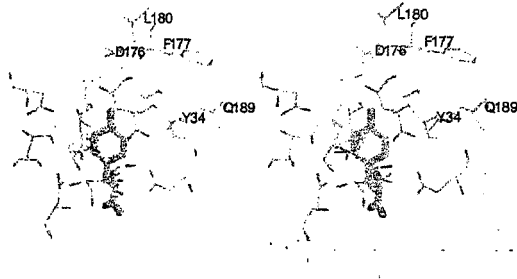
【図 13】



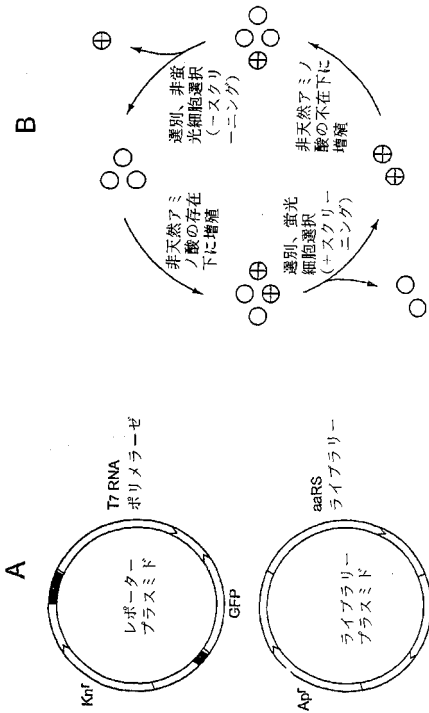
【図 1 4】



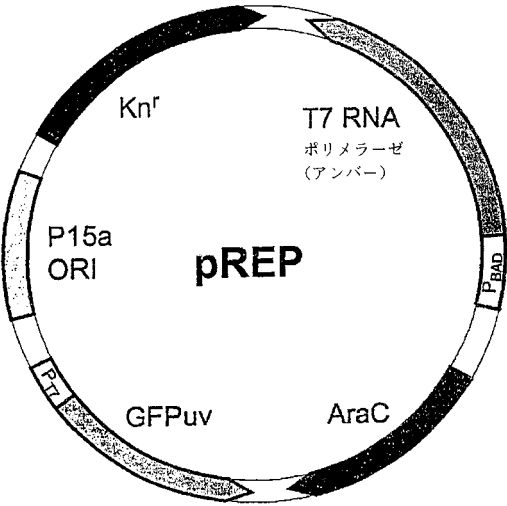
【図 1 5】



【図 1 6】



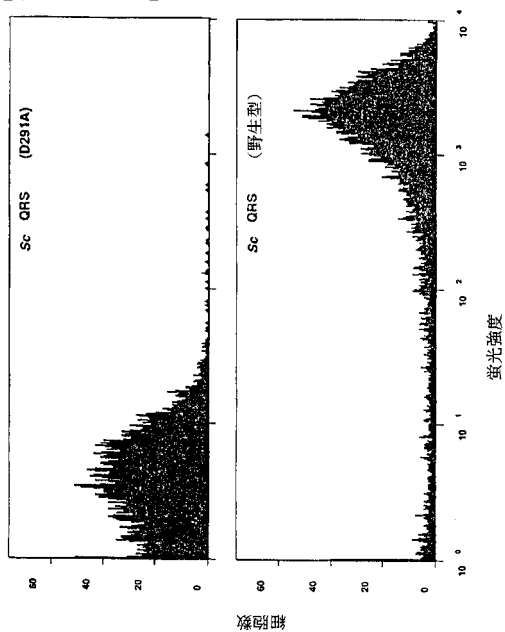
【図 1 7 A】



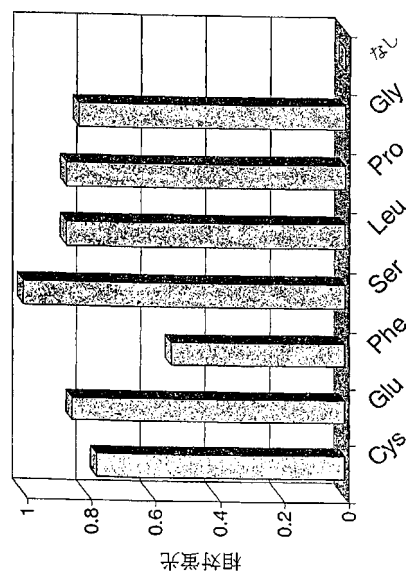
【図 1 7 B】

構築物番号		蛍光増加 (倍)	
1	M1	12	
2	M1	-	
3	M1	-	
4	D10	4	
5	R96	-	
6	Q107	32	
7	A159	6	
8	Q189	-	
9	Q232	20	
10	M1 Q107	220	
11	M1 A159	48	
12	M1 Q232	82	

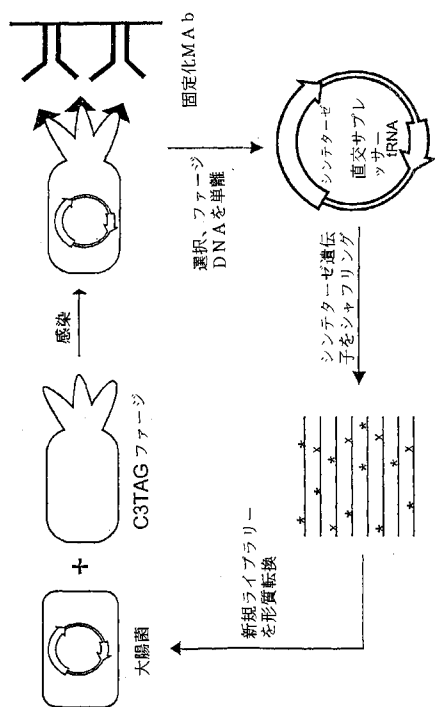
【図 17 C】



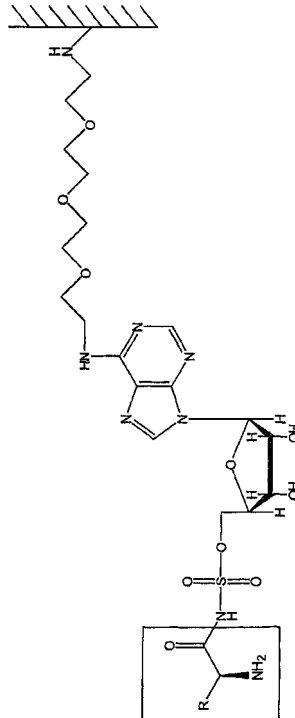
【図 17 D】



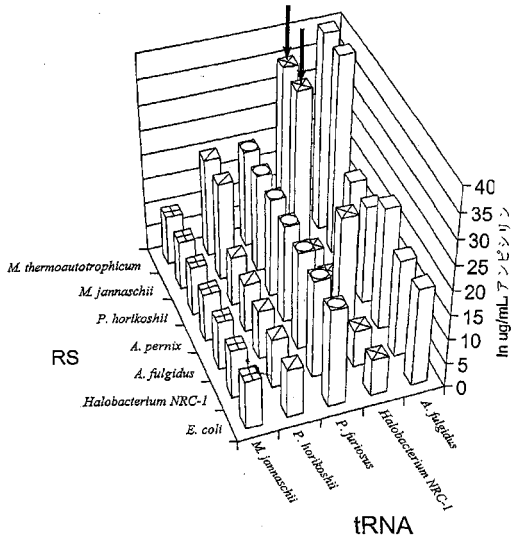
【図 18】



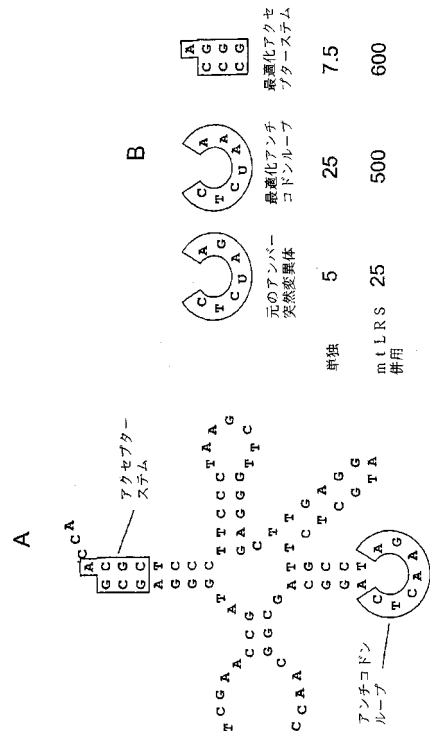
【図 19】



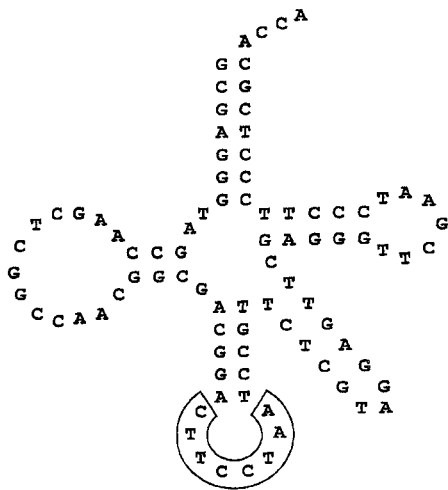
【図 20】



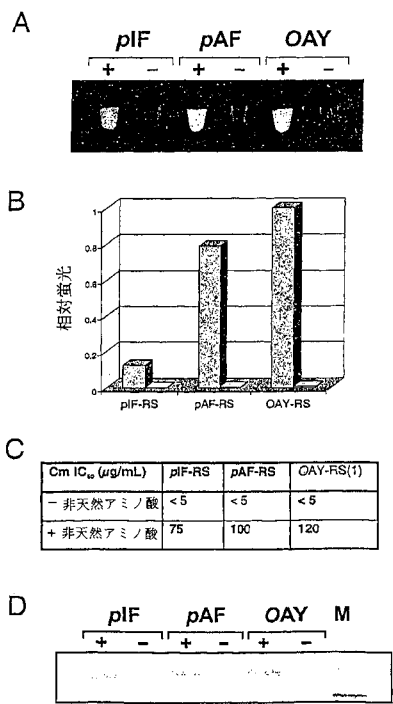
【図 21】



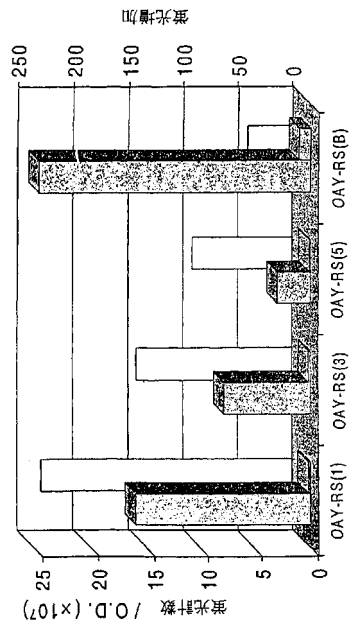
【図 22】



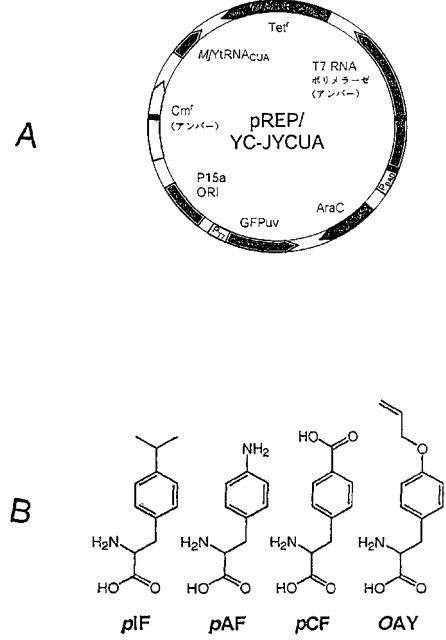
【図 23】



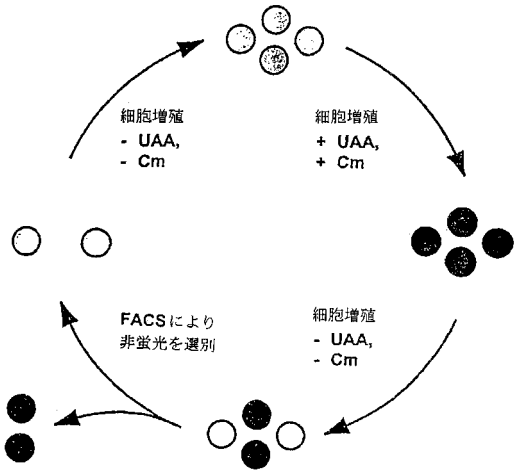
【図 2 4】



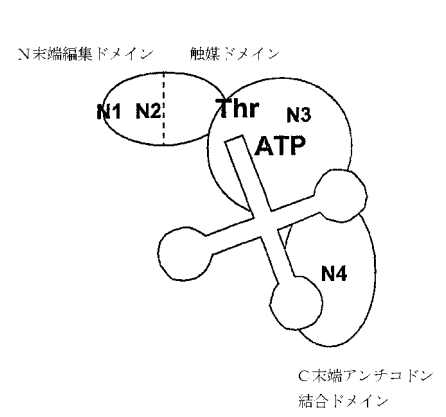
【図 2 5】



【図 2 6】



【図 2 7】



【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number
WO 02/086075 A2

- (51) International Patent Classification: C12N 92130 (US). SANTORO, Stephen, William [US/US]; 12668-63 Carmel Country Rd., San Diego, CA 92130 (US). ZHANG, Zhiwen [CN/US]; 7045 Charmant Drive, #33, San Diego, CA 92122 (US).
- (21) International Application Number: PCT/US02/12635
- (22) International Filing Date: 19 April 2002 (19.04.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/285,030 19 April 2001 (19.04.2001) US
60/355,514 6 February 2002 (06.02.2002) US
- (71) Applicant (for all designated States except US): THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): SCHULTZ, Peter [US/US]; 1650 La Jolla Ranch Road, La Jolla, CA 92037 (US). WANG, Lei [CN/US]; 8730 Costa Verde Blvd., # 2320, San Diego, CA 92122 (US). ANDERSON, John, Christopher [US/US]; 3922 8th Avenue, Apt. 11, San Diego, CA 92103 (US). CHIN, Jason, W. [GB/US]; 4249 Nobel Drive #23, San Diego, CA 92122 (US). LIU, David, R. [US/US]; 1 Fox Run Lane, Lexington, MA 02420 (US). MAGLIERY, Thomas, J. [US/US]; 333 Mill Rd., North Haven, CT 06473 (US). MEGGERS, Eric, L. [DE/US]; 1534 Pine Street, Apt. 2F, Philadelphia, PA 19102 (US). MEHL, Ryan, Aaron [US/US]; 3576 31st Street, San Diego, CA 92104 (US). PASTRNAK, Miro [SK/US]; 11686 Carmel Creek Road, Apt. 107, San Diego, CA 92130 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GH, GM, GR, GU, HT, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KI, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NI, SN, TD, TG).
- Published:
without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/086075 A2

(54) Title: METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGONAL TRNA-AMINOACYL-TRNA SYNTHETASE PAIRS

(57) Abstract: This invention provides compositions and methods for generating components of protein biosynthetic machinery including orthogonal tRNAs, orthogonal aminoacyl-tRNA synthetases, and orthogonal pairs of tRNAs/synthetases. Methods for identifying orthogonal pairs are also provided. These components can be used to incorporate unnatural amino acids into proteins in vivo.

WO 02/086075

PCT/US02/12635

METHODS AND COMPOSITION FOR THE PRODUCTION OF
ORTHOGONAL tRNA-AMINOACYL-tRNA SYNTHETASE PAIRS

CROSS-REFERENCE TO RELATED APPLICATIONS

- 5 [0001] This application claims priority to United States provisional patent application Serial No. 60/285,030, filed April 19, 2001, and United States patent application Serial No. 60/355,514, filed February 6, 2002, the specifications of which are incorporated herein in their entirety.

**STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY
10 SPONSORED RESEARCH AND DEVELOPMENT**

- [0002] The invention was made with United States Government support under Grant No. 6502573 from the Office of Naval Research and Grant No. GM2159 from the National Institutes. The United States Government has certain rights in the invention.

FIELD OF THE INVENTION

- 15 [0003] The invention relates to the field of translation biochemistry. In particular, the invention relates to methods for producing mutated orthogonal tRNAs, mutated orthogonal aminoacyl-tRNA synthetases, and pairs thereof. The invention also provides methods for identifying orthogonal pairs, which are used for the incorporation of unnatural amino acids into proteins in vivo, and related compositions.

BACKGROUND OF THE INVENTION

- 20 [0004] Proteins carry out virtually all of the complex processes of life, from photosynthesis to signal transduction and the immune response. To understand and control these intricate activities, a better understanding of the relationship between the structure and function of proteins is needed.
- 25 [0005] Unlike small organic molecule synthesis wherein almost any structural change can be made to influence functional properties of a compound, the synthesis of

WO 02/086075

PCT/US02/12635

proteins is limited to changes encoded by the twenty natural amino acids. The genetic code of every known organism, from bacteria to human, encodes the same twenty common amino acids. These amino acids can be modified by post-translational modification of proteins, e.g., glycosylation, phosphorylation or oxidation, or in rarer instances, by the enzymatic modification of aminoacylated suppressor tRNAs, e.g., in the case of selenocysteine. Nonetheless, polypeptides, which are synthesized from only these 20 simple building blocks, carry out all of the complex processes of life.

[0006] Both site-directed and random mutagenesis, in which specific amino acids in a protein can be replaced with any of the other nineteen common amino acids, have become important tools for understanding the relationship between the structure and function of proteins. These methodologies have made possible the generation of proteins with enhanced properties, including stability, catalytic activity and binding specificity. Nevertheless, changes in proteins are limited to the 20 common amino acids, most of which have simple functional groups. See Knowles, J. R. *Tinkering with enzymes: what are we learning?* Science, 236:1252-1258 (1987); and, Zoller, M. J., Smith, M. *Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors*, Methods Enzymol, 100:468-500 (1983). By expanding the genetic code to include additional amino acids with novel biological, chemical or physical properties, the properties of proteins, e.g., the size, acidity, nucleophilicity, hydrogen-bonding, hydrophobic properties, etc., can be modified as compared to a protein composed of only amino acids from the 20 common amino acids, e.g., as in a naturally occurring protein.

[0007] Several strategies have been employed to introduce unnatural amino acids into proteins. The first experiments involved the derivatization of amino acids with reactive side-chains such as Lys, Cys and Tyr, for example, the conversion of lysine to N^ε-acetyl-lysine. Chemical synthesis also provides a straightforward method to incorporate unnatural amino acids, but routine solid-phase peptide synthesis is generally limited to small peptides or proteins with less than 100 residues. With the recent development of enzymatic ligation and native chemical ligation of peptide fragments, it is possible to make larger proteins, but such methods are not easily scaled. See, e.g., P. E. Dawson and S. B. H. Kent, Annu. Rev. Biochem., 69:923 (2000). A general in vitro biosynthetic method in which a suppressor tRNA chemically acylated with the desired unnatural amino acid is added to an in vitro extract capable of supporting protein biosynthesis, has been

WO 02/086075

PCT/US02/12635

used to site-specifically incorporate over 100 unnatural amino acids into a variety of proteins of virtually any size. *See, e.g.*, V. W. Cornish, D. Mendel and 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); and, J. D. Bain, C. G. Glabe, T. A. Dix, A. R. Chamberlin, E. S. Dila, *Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide*, J. Am. Chem. Soc. 111 8013-8014 (1989). A broad range of functional groups has been introduced into proteins for studies of protein stability, protein folding, enzyme mechanism, and signal transduction. Although these studies demonstrate that the protein biosynthetic machinery tolerates a wide variety of amino acid side chains, the method is technically demanding, and yields of mutant proteins are low.

[0008] Over 50 years ago, it was found that many analogs of natural amino acids inhibit the growth of bacteria. Analysis of the proteins produced in the presence of these amino acid analogs revealed that they had been substituted for their natural counterparts to various extents. *See, e.g.*, M. H. Richmond, Bacteriol. Rev., 26:398 (1962). This occurs because the aminoacyl-tRNA synthetase, the enzyme responsible for the attachment of the correct amino acid to its cognate tRNA, cannot rigorously distinguish the analog from the corresponding natural amino acid. For instance, norleucine is charged by methionyl-tRNA synthetase, and *p*-fluorophenylalanine is charged by phenylalanine-tRNA synthetase. *See*, D. B. Cowie, G. N. Cohen, E. T. Bolton and H. de Robichon-Szulmajster, Biochim. Biophys. Acta, 1959, 34:39 (1959); and, R. Munier and G. N. Cohen, Biochim. Biophys. Acta, 1959, 31:378 (1959).

[0009] An in vivo method, termed selective pressure incorporation, was later developed to exploit the promiscuity of wild-type synthetases. *See, e.g.*, N. Budisa, C. Minks, S. Alefelder, W. Wenger, F. M. Dong, L. Moroder and R. Huber, FASEB J., 13:41 (1999). An auxotrophic strain, in which the relevant metabolic pathway supplying the cell with a particular natural amino acid is switched off, is grown in minimal media containing limited concentrations of the natural amino acid, while transcription of the target gene is repressed. At the onset of a stationary growth phase, the natural amino acid is depleted and replaced with the unnatural amino acid analog. Induction of expression of the recombinant protein results in the accumulation of a protein containing the unnatural

WO 02/086075

PCT/US02/12635

- analog. For example, using this strategy, *o*, *m* and *p*-fluorophenylalanines have been incorporated into proteins, and exhibit two characteristic shoulders in the UV spectrum which can be easily identified, *see, e.g.*, C. Minks, R. Huber, L. Moroder and N. Budisa, Anal. Biochem., 284:29 (2000); trifluoromethionine has been used to replace methionine in bacteriophage λ lysozyme to study its interaction with chitooligosaccharide ligands by ^{19}F NMR, *see, e.g.*, H. Duewel, E. Daub, V. Robinson and J. F. Honek, Biochemistry, 36:3404 (1997); and trifluoroisoleucine has been inserted in place of leucine, resulting in increased thermal and chemical stability of a leucine-zipper protein. *See, e.g.*, Y. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado and D. A. Tirrell, Angew. Chem. Int. Ed. Engl., 40:1494 (2001). Moreover, selenomethionine and telluromethionine are incorporated into various recombinant proteins to facilitate the solution of phases in X-ray crystallography. *See, e.g.*, W. A. Hendrickson, J. R. Horton and D. M. Lemaster, EMBO J., 9:1665 (1990); J. O. Boles, K. Lewinski, M. Kunkle, J. D. Odom, B. Dunlap, L. Lebiada and M. Hatada, Nat. Struct. Biol., 1:283 (1994); N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann and R. Huber, Eur. J. Biochem., 230:788 (1995); and, N. Budisa, W. Kambrock, S. Steinbacher, A. Humm, L. Prade, T. Neufeld, L. Moroder and R. Huber, J. Mol. Biol., 270:616 (1997). Methionine analogs with alkene or alkyne functionalities have also been inserted efficiently, allowing for additional modification of proteins by chemical means. *See, e.g.*, J. C. M. van Hest and D. A. Tirrell, FEBS Lett., 428:68 (1998); J. C. M. van Hest, K. L. Kiick and D. A. Tirrell, J. Am. Chem. Soc., 122:1282 (2000); and, K. L. Kiick and D. A. Tirrell, Tetrahedron, 56:9487 (2000).
- [0010]** The success of this method depends on the recognition of the unnatural amino acid analogs by aminoacyl-tRNA synthetases, which, in general, requires high selectivity to insure the fidelity of protein translation. Therefore, the range of chemical functionality accessible via this route is limited. For instance, although thiaproline can be incorporated quantitatively into proteins, oxaproline and selenoproline cannot. *See*, N. Budisa, C. Minks, F. J. Medrano, J. Lutz, R. Huber and L. Moroder, Proc. Natl. Acad. Sci. U S A, 95:455 (1998). One way to expand the scope of this method is to relax the substrate specificity of aminoacyl-tRNA synthetases, which has been achieved in a limited number of cases. For example, it was found that replacement of Ala²⁹⁴ by Gly in *Escherichia coli* phenylalanyl-tRNA synthetase (PheRS) increases the size of substrate binding pocket, and results in the acylation of tRNA^{Phe} by *p*-Cl-phenylalanine (*p*-Cl-Phe). *See*, M. Ibba, P. Kast and H. Hennecke, Biochemistry, 33:7107 (1994). An *Escherichia*

WO 02/086075

PCT/US02/12635

coli strain harboring this mutant PheRS allows the incorporation of *p*-Cl-phenylalanine or *p*-Br-phenylalanine in place of phenylalanine. See, e.g., M. Ibba and H. Hennecke, FEBS Lett., 364:272 (1995); and, N. Sharma, R. Furter, P. Kast and D. A. Tirrell, FEBS Lett., 467:37 (2000). Similarly, a point mutation Phe130Ser near the amino acid binding site of

- 5 *Escherichia coli* tyrosyl-tRNA synthetase was shown to allow azatyrine to be incorporated more efficiently than tyrosine. See, F. Hamano-Takaku, T. Iwama, S. Saito-Yano, K. Takaku, Y. Monden, M. Kitabatake, D. Soll and S. Nishimura, J. Biol. Chem., 275:40324 (2000).
- [0011] The fidelity of aminoacylation is maintained both at the level of substrate discrimination and proofreading of non-cognate intermediates and products. Therefore, an alternative strategy to incorporate unnatural amino acids into proteins in vivo is to modify synthetases that have proofreading mechanisms. These synthetases cannot discriminate and therefore activate amino acids that are structurally similar to the cognate natural amino acids. This error is corrected at a separate site, which deacylates the mischarged amino acid from the tRNA to maintain the fidelity of protein translation. If the proofreading activity of the synthetase is disabled, structural analogs that are misactivated may escape the editing function and be incorporated. This approach has been demonstrated recently with the valyl-tRNA synthetase (ValRS). See, V. Doring, H. D. Mootz, L. A. Nangle, T. L. Hendrickson, V. de Crecy-Lagard, P. Schimmel and P. Marliere, Science, 292:501
- 20 (2001). ValRS can misaminoacylate tRNA^{Val} with Cys, Thr, or aminobutyrate (Abu); these noncognate amino acids are subsequently hydrolyzed by the editing domain. After random mutagenesis of the *Escherichia coli* chromosome, a mutant *Escherichia coli* strain was selected that has a mutation in the editing site of ValRS. This edit-defective ValRS incorrectly charges tRNA^{Val} with Cys. Because Abu sterically resembles Cys (–SH group of Cys is replaced with –CH₃ in Abu), the mutant ValRS also incorporates Abu into
- 25 proteins when this mutant *Escherichia coli* strain is grown in the presence of Abu. Mass spectrometric analysis shows that about 24% of valines are replaced by Abu at each valine position in the native protein.

[0012] At least one major limitation of the methods described above is that all sites corresponding to a particular natural amino acid throughout the protein are replaced. The extent of incorporation of the natural and unnatural amino acid may also vary – only in rare cases can quantitative substitution be achieved since it is difficult to completely

30

WO 02/086075

PCT/US02/12635

deplete the cognate natural amino acid inside the cell. Another limitation is that these strategies make it difficult to study the mutant protein in living cells, because the multi-site incorporation of analogs often results in toxicity. Finally, this method is applicable in general only to close structural analogs of the common amino acids, again because
 5 substitutions must be tolerated at all sites in the genome.

[0013] Solid-phase synthesis and semi-synthetic methods have also allowed for the synthesis of a number of small proteins containing novel amino acids. For example, see the following publications and references cited within: Crick, F.J.C., Barrett, L. Brenner, S. Watts-Tobin, R. *General nature of the genetic code for proteins*, Nature, 192:1227-
 10 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., 59:5914-5919 (1966); Kaiser, E. T. *Synthetic approaches to biologically active peptides and proteins including enzymes*, Acc. Chem. Res., 47-54 (1989); Nakatsuka, T., Sasaki, T., Kaiser, E. T. *Peptide segment coupling catalyzed by the*
 15 *semisynthetic enzyme thiosubtilisin*, J. Am. Chem. Soc., 109:3808-3810 (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-
 20 157 (1987); and, 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):243-247(1994).

[0014] Chemical modification has been used to introduce a variety of unnatural side chains, including cofactors, spin labels and oligonucleotides into proteins in vitro.
 25 See, e.g., Corey, D. R., Schultz, P. G. *Generation of a hybrid sequence-specific single-stranded deoxyribonuclease*, Science, 283(4832):1401-1403 (1987); Kaiser, E. T., Lawrence D. S., Rokita, S. E. *The chemical modification of enzymatic specificity*, 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. B., M. L. A new enzyme containing a synthetically formed active site. *Thiol-subtilisin*. J. Am. Chem. Soc., 88:3153-3154 (1966); and, Pollack, S.J., Nakayama, G. Schultz, P. G.

WO 02/086075

PCT/US02/12635

Introduction of nucleophiles and spectroscopic probes into antibody combining sites, Science, 242(4881):1038-1040 (1988).

- [0015] Alternatively, biosynthetic methods that employ chemically modified aminoacyl-tRNAs have been used to incorporate several biophysical probes into proteins synthesized in vitro. See the following publications and references cited within: Brunner, J. *New Photolabeling and crosslinking methods*, Annu. Rev. Biochem., 62:483-514 (1993); and, Krieg, U. C., Walter, P., Hohnson, A. E. *Photocrosslinking of the signal sequence of nascent preprolactin of the 54-kilodalton polypeptide of the signal recognition particle*, Proc. Natl. Acad. Sci. 83(22):8604-8608 (1986).
- 10 [0016] Previously, it has been shown that unnatural amino acids can be site-specifically incorporated into proteins in vitro by the addition of chemically aminoacylated suppressor tRNAs to protein synthesis reactions programmed with a gene containing a desired amber nonsense mutation. Using these approaches, one can substitute a number of the common twenty amino acids with close structural homologues, e.g.,
- 15 fluorophenylalanine for phenylalanine, using strains auxotrophic for a particular amino acid. See, e.g., Noren, C. J., Anthony-Cahill, Griffith, M.C., Schultz, P. G. *A general method for site-specific incorporation of unnatural amino acids 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., Dimala, E. S. *Biosynthetic site-specific Incorporation of a non-natural amino acid into a polypeptide*, J. Am. Chem. Soc., 111:8013-8014
- 20 (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 acids site-specifically into proteins*, Methods in Enz., 301-336 (1992); and, Mendel, D., Cornish, V.W. & Schultz, P. G. *Site-Directed Mutagenesis with an Expanded Genetic Code*, Annu. Rev. Biophys. Biomol. Struct. 24, 435-62 (1995).
- 25 [0017] For example, a suppressor tRNA was prepared that recognized the stop codon UAG and was chemically aminoacylated with an unnatural amino acid. Conventional site-directed mutagenesis was used to introduce the stop codon TAG, at the site of interest in the protein gene. See, e.g., Sayers, J. R., Schmidt, W. Eckstein, F. 5', 3'
- 30 *Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis*, Nucleic Acids Res., 16(3):791-802 (1988). When the acylated suppressor tRNA and the mutant gene were combined in an in vitro transcription/translation system, the unnatural amino

WO 02/086075

PCT/US02/12635

acid was incorporated in response to the UAG codon which gave a protein containing that amino acid at the specified position. Experiments using [^3H]-Phe and experiments with α -hydroxy acids demonstrated that only the desired amino acid is incorporated at the position specified by the UAG codon and that this amino acid is not incorporated at any other site in the protein. See, e.g., Noren, et al, *supra*; and, Ellman, J. A., Mendel, D., Schultz, P. G. *Site-specific incorporation of novel backbone structures into proteins*, Science, 197-200 (1992).

[0018] In general, these in vitro approaches are limited by difficulties in achieving site-specific incorporation of the amino acids, by the requirement that the amino acids be simple derivatives of the common twenty amino acids or problems inherent in the synthesis of large proteins or peptide fragments.

[0019] Microinjection techniques have also been use incorporate unnatural amino acids into proteins. See, e.g., 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 and H. A. Lester, Science, 268:439 (1995); and, D. A. Dougherty, Curr. Opin. Chem. Biol., 4:645 (2000). A *Xenopus* oocyte was coinjected with two RNA species made in vitro: an mRNA encoding the target protein with a UAG stop codon at the amino acid position of interest and an amber suppressor tRNA aminoacylated with the desired unnatural amino acid. The translational machinery of the oocyte then inserted the unnatural amino acid at the position specified by UAG. This method has allowed in vivo structure-function studies of integral membrane proteins, which are generally not amenable to in vitro expression systems. Examples include the incorporation of a fluorescent amino acid into tachykinin neurokinin-2 receptor to measure distances by fluorescence resonance energy transfer, see, e.g., G. Turcatti, K. Nemeth, M. D. Edgerton, U. Meseth, F. Talbot, M. Peitsch, J. Knowles, H. Vogel and A. Chollet, J. Biol. Chem., 271:19991 (1996); the incorporation of biotinylated amino acids to identify surface-exposed residues in ion channels, see, e.g., J. P. Gallivan, H. A. Lester and D. A. Dougherty, Chem. Biol., 4:739 (1997); the use of caged tyrosine analogs to monitor conformational changes in an ion channel in real time, see, e.g., J. C. Miller, S. K. Silverman, P. M. England, D. A. Dougherty and H. A. Lester, Neuron, 20:619 (1998); and, the use of α -hydroxy amino acids to change ion channel backbones for probing their gating mechanisms, see, e.g., P. M. England, Y. Zhang, D. A. Dougherty and H. A. Lester,

WO 02/086075

PCT/US02/12635

Cell, 96:89 (1999); and, T. Lu, A. Y. Ting, J. Mainland, L. Y. Jan, P. G. Schultz and J. Yang, Nat. Neurosci., 4:239 (2001).

[0020] However, there are limitations microinjection method, e.g., the suppressor tRNA has to be chemically aminoacylated with the unnatural amino acid in vitro, and the acylated tRNA is consumed as a stoichiometric reagent during translation and cannot be regenerated. This limitation results in poor suppression efficiency and low protein yields, necessitating highly sensitive techniques to assay the mutant protein, such as electrophysiological measurements. Moreover, this method is only applicable to cells that can be microinjected.

10 [0021] The ability to incorporate unnatural amino acids directly into proteins in vivo offers the advantages of high yields of mutant proteins, technical ease, the potential to study the mutant proteins in cells or possibly in living organisms and the use of these mutant proteins in therapeutic treatments. The ability to include unnatural amino acids with various sizes, acidities, nucleophilicities, hydrophobicities, and other properties into
15 proteins can greatly expand our ability to rationally and systematically manipulate the structures of proteins, both to probe protein function and create new proteins or organisms with novel properties. However, the process is difficult, because the complex nature of tRNA-synthetase interactions that are required to achieve a high degree of fidelity in protein translation. Therefore, improvements to the process are needed to provide more
20 efficient and effective methods to alter the biosynthetic machinery of the cell. The present invention addresses these and other needs, as will be apparent upon review of the following disclosure.

SUMMARY OF THE INVENTION

[0022] The present invention provides compositions of components used in protein
25 biosynthetic machinery, which include orthogonal tRNA-aminoacyl-tRNA synthetase pairs and the individual components of the pairs. Methods for generating and selecting orthogonal tRNAs, orthogonal aminoacyl-tRNA synthetases, and pairs thereof that can use an unnatural amino acid are also provided. Compositions of the invention include novel orthogonal tRNA-aminoacyl-tRNA synthetase pairs, e.g., mutRNATyr-mutTyrRS pairs,
30 mutRNALeu-mutLeuRS pairs, mutRNAThr-mutThrRS pairs, mutRNAGlu-mutGluRS pairs, and the like. The novel orthogonal pairs can be used to incorporate an unnatural

WO 02/086075

PCT/US02/12635

amino acid in a polypeptide in vivo. Other embodiments of the invention include selecting orthogonal pairs.

[0023] Compositions of the present invention include an orthogonal aminoacyl-tRNA synthetase (O-RS), where the O-RS preferentially aminoacylates an orthogonal
5 tRNA (O-tRNA) with an unnatural amino acid, optionally, in vivo. In one embodiment, the O-RS comprises a nucleic acid comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO: 4-34 (*see*, Table 5) and a complementary polynucleotide sequence thereof. In another embodiment, the O-RS has improved or enhanced enzymatic properties, e.g., the K_m is higher or lower, the k_{cat} is higher or lower,
10 the value of k_{cat}/K_m is higher or lower or the like, for the unnatural amino acid compared to a naturally occurring amino acid, e.g., one of the 20 known amino acids.

[0024] The unnatural amino acids of the present invention encompass a variety of substances. For example, they optionally include (but are not limited to) such molecules as: an O-methyl-L-tyrosine, an L-3-(2-naphthyl)alanine, a 3-methyl-phenylalanine, an O-
15 4-allyl-L-tyrosine, a 4-propyl-L-tyrosine, a tri-O-acetyl-GlcNAc β -serine, an L-Dopa, a fluorinated phenylalanine, an isopropyl-L-phenylalanine, a *p*-azido-L-phenylalanine, a *p*-acyl-L-phenylalanine, a *p*-benzoyl-L-phenylalanine, an L-phosphoserine, a phosphoserine, a phosphotyrosine, a *p*-iodo-phenylalanine, a *p*-bromophenylalanine, a *p*-amino-L-phenylalanine, and an isopropyl-L-phenylalanine. Additionally, other
20 examples optionally include (but are not limited to) an unnatural analogue of a tyrosine amino acid; an unnatural analogue of a glutamine amino acid; an unnatural analogue of a phenylalanine amino acid; an unnatural analogue of a serine amino acid; an unnatural analogue of a threonine amino acid; an alkyl, aryl, acyl, azido, cyano, halo, hydrazine, hydrazide, hydroxyl, alkenyl, alkynyl, ether, thiol, sulfonyl, seleno, ester, thioacid, borate,
25 boronate, phospho, phosphono, phosphine, heterocyclic, enone, imine, aldehyde, hydroxylamine, keto, or amino substituted amino acid, or any combination thereof; an amino acid with a photoactivatable cross-linker; a spin-labeled amino acid; a fluorescent amino acid; an amino acid with a novel functional group; an amino acid that covalently or noncovalently interacts with another molecule; a metal binding amino acid; a metal-
30 containing amino acid; a radioactive amino acid; a photocaged amino acid; a photoisomerizable amino acid; a biotin or biotin-analogue containing amino acid; a glycosylated or carbohydrate modified amino acid; a keto containing amino acid; an

WO 02/086075

PCT/US02/12635

amino acid comprising polyethylene glycol; an amino acid comprising polyether; a heavy atom substituted amino acid; a chemically cleavable or photocleavable amino acid; an amino acid with an elongated side chain; an amino acid containing a toxic group; a sugar substituted amino acid, e.g., a sugar substituted serine or the like; a carbon-linked sugar-containing amino acid; a redox-active amino acid; an α -hydroxy containing acid; an amino thio acid containing amino acid; an α,α disubstituted amino acid; a β -amino acid; and a cyclic amino acid other than proline.

[0025] The present invention also includes a polypeptide comprising an amino acid sequence encoded by a coding polynucleotide sequence which is selected from: a coding polynucleotide sequence selected from SEQ ID NO: 4-34 (*see*, Table 5 for sequences); a coding polynucleotide sequence encoding a polypeptide selected from SEQ ID NO: 35-66 a polynucleotide sequence which hybridizes under highly stringent conditions over substantially the entire length of such polynucleotide sequences; and complementary sequences of any of such sequences. Additionally, such polypeptide optionally encodes an orthogonal aminoacyl tRNA synthetase and/or an amino acid sequence selected from SEQ ID NO:35 to SEQ ID NO:66.

[0026] The present invention also includes a nucleic acid comprising a polynucleotide sequence selected from the group consisting of: a polynucleotide sequence selected from SEQ ID NO:1 to SEQ ID NO:3 (or a complementary polynucleotide sequence thereof) and a polynucleotide sequence which hybridizes under highly stringent conditions over substantially the entire length of such polynucleotide sequences. Such nucleic acids also include wherein the polynucleotide sequence comprises an orthogonal tRNA and/or wherein the polynucleotide sequence forms a complementary pair with an orthogonal aminoacyl-tRNA synthetase (which optionally is selected from the those whose sequence is listed in SEQ ID NO:35 to SEQ ID NO:66).

[0027] Compositions of an orthogonal tRNA (O-tRNA) are also included, where the O-tRNA recognizes a selector codon and wherein the O-tRNA is preferentially aminoacylated with an unnatural amino acid by an orthogonal aminoacyl-tRNA synthetase. In one embodiment, the O-tRNA comprises a nucleic acid comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO: 1-3 (*see*, Table 5) and a complementary polynucleotide sequence thereof.

WO 02/086075

PCT/US02/12635

[0028] Selector codons of the present invention expand the genetic codon framework of protein biosynthetic machinery. For example, a selector codon includes, e.g., a unique three base codon (composed of natural or unnatural bases), a nonsense codon (such as a stop codon, e.g., an amber codon, or an opal codon), an unnatural codon, a rare codon, a codon comprising at least four bases, a codon comprising at least five bases, a codon comprising at least six bases, or the like.

[0029] In one embodiment, the O-tRNA (optionally comprising within compositions) can include an orthogonal aminoacyl-tRNA synthetase (O-RS), e.g., where the O-tRNA and the O-RS are complementary, e.g., an O-tRNA/O-RS pair. In one embodiment, a pair comprises e.g., a mutRNATyr-mutTyrRS pair, such as mutRNATyr-SS12TyrRS pair, a mutRNALeu-mutLeuRS pair, a mutRNAThr-mutThrRS pair, a mutRNAGlu-mutGluRS pair, or the like. In another embodiment, the pair is other than a mutRNAGln-mutGlnRS derived from *Escherichia coli*, a mutRNAAsp-mutAspRS derived from yeast or a mutRNAPheCUA-mutphenylalanineRS from yeast, where these pairs do not possess the properties of the pairs of the present invention.

[0030] The O-tRNA and the O-RS can be derived by mutation of a naturally occurring tRNA and RS from a variety of organisms. In one embodiment, the O-tRNA and O-RS are derived from at least one organism, where the organism is a prokaryotic organism, e.g., *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *P. furiosus*, *P. horikoshii*, *A. permix*, *T. thermophilus*, or the like. Optionally, the organism is a eukaryotic organism, e.g., plants (e.g., complex plants such as monocots, or dicots), algae, fungi (e.g., yeast, etc), animals (e.g., mammals, insects, arthropods, etc.), insects, protists, or the like. Optionally, the O-tRNA is derived by mutation of a naturally occurring tRNA from a first organism and the O-RS is derived by mutation of a naturally occurring RS from a second organism. In one embodiment, the O-tRNA and O-RS can be derived from a mutated tRNA and mutated RS.

[0031] The O-tRNA and the O-RS also can optionally be isolated from a variety of organisms. In one embodiment, the O-tRNA and O-RS are isolated from at least one organism, where the organism is a prokaryotic organism, e.g., *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *P. furiosus*, *P. horikoshii*, *A. permix*, *T. thermophilus*, or the like. Optionally, the organism is

WO 02/086075

PCT/US02/12635

a eukaryotic organism, e.g., plants (e.g., complex plants such as monocots, or dicots), algae, fungi (e.g., yeast, etc), animals (e.g., mammals, insects, arthropods, etc.), insects, protists, or the like. Optionally, the O-tRNA is isolated from a naturally occurring tRNA from a first organism and the O-RS is isolated from a naturally occurring RS from a second organism. In one embodiment, the O-tRNA and O-RS can be isolated from one or more library (which optionally comprises one or more O-tRNA and/or O-RS from one or more organism (including those comprising prokaryotes and/or eukaryotes).

[0032] In another aspect, the compositions of the present invention can be in a cell. Optionally, the compositions of the present invention can be in an in vitro translation system.

[0033] Methods for generating components of the protein biosynthetic machinery, such as O-RSs, O-tRNAs, and orthogonal O-tRNA/O-RS pairs that can be used to incorporate an unnatural amino acid are provided in the present invention. Methods for selecting an orthogonal tRNA-tRNA synthetase pair for use in in vivo translation system of an organism are also provided. The unnatural amino acids and selectors codons used in the methods are described above and below.

[0034] Methods for producing at least one recombinant orthogonal aminoacyl-tRNA synthetase (O-RS) comprise: (a) generating a library of (optionally mutant) RSs derived from at least one aminoacyl-tRNA synthetase (RS) from a first organism, e.g., a prokaryotic organism, such as *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *P. furiosus*, *P. horikoshii*, *A. pernix*, *T. thermophilus*, or the like; (b) selecting (and/or screening) the library of RSs (optionally mutant RSs) for members that aminoacylate an orthogonal tRNA (O-tRNA) in the presence of an unnatural amino acid and a natural amino acid, thereby providing a pool of active (optionally mutant) RSs; and/or, (c) selecting (optionally through negative selection) the pool for active RSs (e.g., mutant RSs) that preferentially aminoacylate the O-tRNA in the absence of the unnatural amino acid, thereby providing the at least one recombinant O-RS; wherein the at least one recombinant O-RS preferentially aminoacylates the O-tRNA with the unnatural amino acid. Recombinant O-RSs produced by the methods are also included in the present invention.

[0035] In one embodiment, the RS is an inactive RS. The inactive RS can be generated by mutating an active RS. For example, the inactive RS can be generated by

WO 02/086075

PCT/US02/12635

mutating at least about 1, at least about 2, at least about 3, at least about 4, at least about 5, at least about 6, or at least about 10 or more amino acids to different amino acids, e.g., alanine.

5 [0036] Libraries of mutant RSs can be generated using various mutagenesis techniques known in the art. For example, the mutant RSs can be generated by site-specific mutations, random mutations, diversity generating recombination mutations, chimeric constructs, and by other methods described herein or known in the art.

[0037] In one embodiment, selecting (and/or screening) the library of RSs (optionally mutant RSs) for members that are active, e.g., that aminoacylate an orthogonal
10 tRNA (O-tRNA) in the presence of an unnatural amino acid and a natural amino acid, includes: introducing a positive selection or screening marker, e.g., an antibiotic resistance gene, or the like, and the library of (optionally mutant) RSs into a plurality of cells, wherein the positive selection and/or screening marker comprises at least one selector codon, e.g., an amber, ochre, or opal codon; growing the plurality of cells in the presence
15 of a selection agent; identifying cells that survive (or show a specific response) in the presence of the selection and/or screening agent by suppressing the at least one selector codon in the positive selection or screening marker, thereby providing a subset of positively selected cells that contains the pool of active (optionally mutant) RSs. Optionally, the selection and/or screening agent concentration can be varied.

20 [0038] In one aspect, the positive selection marker is a chloramphenicol acetyltransferase (CAT) gene and the selector codon is an amber stop codon in the CAT gene. Optionally, the positive selection marker is a β -lactamase gene and the selector codon is an amber stop codon in the β -lactamase gene. In another aspect the positive screening marker comprises a fluorescent or luminescent screening marker or an affinity
25 based screening marker (e.g., a cell surface marker).

[0039] In one embodiment, negatively selecting or screening the pool for active RSs (optionally mutants) that preferentially aminoacylate the O-tRNA in the absence of the unnatural amino acid includes: introducing a negative selection or screening marker with the pool of active (optionally mutant) RSs from the positive selection or screening
30 into a plurality of cells of a second organism, wherein the negative selection or screening marker comprises at least one selector codon (e.g., an antibiotic resistance gene, e.g., a chloramphenicol acetyltransferase (CAT) gene); and, identifying cells that survive or show

WO 02/086075

PCT/US02/12635

a specific screening response in a 1st media supplemented with the unnatural amino acid and a screening or selection agent, but fail to survive or to show the specific response in a 2nd media not supplemented with the unnatural amino acid and the selection or screening agent, thereby providing surviving cells or screened cells with the at least one recombinant O-RS. For example, a CAT identification protocol optionally acts as a positive selection and/or a negative screening in determination of appropriate O-RS recombinants. For instance, a pool of clones is optionally replicated on growth plates containing CAT (which comprises at least one selector codon) either with or without one or more unnatural amino acid. Colonies growing exclusively on the plates containing unnatural amino acids are thus regarded as containing recombinant O-RS. In one aspect, the concentration of the selection (and/or screening) agent is varied. In some aspects the first and second organisms are different. Thus, the first and/or second organism optionally comprises: a prokaryote, a eukaryote, a mammal, an *Escherichia coli*, a fungi, a yeast, an archaeobacterium, a eubacterium, a plant, an insect, a protist, etc. In other embodiments, the screening marker comprises a fluorescent or luminescent screening marker or an affinity based screening marker.

[0040] In another embodiment, screening or selecting (e.g., negatively selecting) the pool for active (optionally mutant) RSs includes: isolating the pool of active mutant RSs from the positive selection step (b); introducing a negative selection or screening marker, wherein the negative selection or screening marker comprises at least one selector codon (e.g., a toxic marker gene, e.g., a ribonuclease barnase gene, comprising at least one selector codon), and the pool of active (optionally mutant) RSs into a plurality of cells of a second organism; and identifying cells that survive or show a specific screening response in a 1st media not supplemented with the unnatural amino acid, but fail to survive or show a specific screening response in a 2nd media supplemented with the unnatural amino acid, thereby providing surviving or screened cells with the at least one recombinant O-RS, wherein the at least one recombinant O-RS is specific for the unnatural amino acid. In one aspect, the at least one selector codon comprises about two or more selector codons. Such embodiments optionally can include wherein the at least one selector codon comprises two or more selector codons, and wherein the first and second organism are different (e.g., each organism is optionally, e.g., a prokaryote, a eukaryote, a mammal, an *Escherichia coli*, a fungi, a yeast, an archaeobacteria, a eubacteria, a plant, an insect, a protist, etc.). Also, some aspects include wherein the negative selection marker comprises a ribonuclease

WO 02/086075

PCT/US02/12635

barnase gene (which comprises at least one selector codon). Other aspects include wherein the screening marker optionally comprises a fluorescent or luminescent screening marker or an affinity based screening marker. In the embodiments herein, the screenings and/or selections optionally include variation of the screening and/or selection stringency.

- 5 [0041] In one embodiment, the methods for producing at least one recombinant orthogonal aminoacyl-tRNA synthetase (O-RS) can further comprise: (d) isolating the at least one recombinant O-RS; (e) generating a second set of O-RS (optionally mutated) derived from the at least one recombinant O-RS; and, (f) repeating steps (b) and (c) until a mutated O-RS is obtained that comprises an ability to preferentially aminoacylate the O-tRNA. Optionally, steps (d)-(f) are repeated, e.g., at least about two times. In one aspect, 10 the second set of mutated O-RS derived from at least one recombinant O-RS can be generated by mutagenesis, e.g., random mutagenesis, site-specific mutagenesis, recombination or a combination thereof.

- [0042] The stringency of the selection/screening steps, e.g., the positive 15 selection/screening step (b), the negative selection/screening step (c) or both the positive and negative selection/screening steps (b) and (c), in the above-described methods, optionally includes varying the selection/screening stringency. In another embodiment, the positive selection/screening step (b), the negative selection/screening step (c) or both the positive and negative selection/screening steps (b) and (c) comprise using a reporter, 20 wherein the reporter is detected by fluorescence-activated cell sorting (FACS) or wherein the reporter is detected by luminescence. Optionally, the reporter is displayed on a cell surface, on a phage display or the like and selected based upon affinity or catalytic activity involving the unnatural amino acid or an analogue. In one embodiment, the mutated synthetase is displayed on a cell surface, on a phage display or the like.

- 25 [0043] The methods embodied herein optionally comprise wherein the unnatural amino acid is selected from, e.g.: an O-methyl-L-tyrosine, an L-3-(2-naphthyl)alanine, a 3-methyl-phenylalanine, an O-4-allyl-L-tyrosine, a 4-propyl-L-tyrosine, a tri-O-acetyl-GlcNAc β -serine, an L-Dopa, a fluorinated phenylalanine, an isopropyl-L-phenylalanine, a *p*-azido-L-phenylalanine, a *p*-acyl-L-phenylalanine, a *p*-benzoyl-L-phenylalanine, an L-phosphoserine, a phosphoserine, a phosphonotyrosine, a *p*-iodo-phenylalanine, a *p*-bromophenylalanine, a *p*-amino-L-phenylalanine, and an isopropyl-L-phenylalanine. A 30

WO 02/086075

PCT/US02/12635

recombinant O-RS produced by the methods herein is also included in the current invention.

[0044] Methods for producing a recombinant orthogonal tRNA (O-tRNA) include:

- (a) generating a library of mutant tRNAs derived from at least one tRNA, e.g., a suppressor tRNA, from a first organism; (b) selecting (e.g., negatively selecting) or screening the library for (optionally mutant) tRNAs that are aminoacylated by an aminoacyl-tRNA synthetase (RS) from a second organism in the absence of a RS from the first organism, thereby providing a pool of tRNAs (optionally mutant); and, (c) selecting or screening the pool of tRNAs (optionally mutant) for members that are aminoacylated by an introduced orthogonal RS (O-RS), thereby providing at least one recombinant O-tRNA; wherein the at least one recombinant O-tRNA recognizes a selector codon and is not efficiently recognized by the RS from the second organism and is preferentially aminoacylated by the O-RS. In some embodiments the at least one tRNA is a suppressor tRNA and/or comprises a unique three base codon of natural and/or unnatural bases, or is a nonsense codon, a rare codon, an unnatural codon, a codon comprising at least 4 bases, an amber codon, an ochre codon, or an opal stop codon. In one embodiment, the recombinant O-tRNA possesses an improvement of orthogonality. It will be appreciated that in some embodiments, O-tRNA is optionally imported into a first organism from a second organism without the need for modification. In various embodiments, the first and second organisms are either the same or different and are optionally chosen from, e.g., prokaryotes (e.g., *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Escherichia coli*, *Halobacterium*, etc.), eukaryotes, mammals, fungi, yeasts, archaeobacteria, eubacteria, plants, insects, protists, etc. Additionally, the recombinant tRNA is optionally aminoacylated by an unnatural amino acid, wherein the unnatural amino acid is biosynthesized in vivo either naturally or through genetic manipulation. The unnatural amino acid is optionally added to a growth medium for at least the first or second organism.

- [0045]** In one aspect, selecting (e.g., negatively selecting) or screening the library for (optionally mutant) tRNAs that are aminoacylated by an aminoacyl-tRNA synthetase (step (b)) includes: introducing a toxic marker gene, wherein the toxic marker gene comprises at least one of the selector codons (or a gene that leads to the production of a toxic or static agent or a gene essential to the organism wherein such marker gene

WO 02/086075

PCT/US02/12635

comprises at least one selector codon) and the library of (optionally mutant) tRNAs into a plurality of cells from the second organism; and, selecting surviving cells, wherein the surviving cells contain the pool of (optionally mutant) tRNAs comprising at least one orthogonal tRNA or nonfunctional tRNA. For example, surviving cells can be selected by using a comparison ratio cell density assay.

[0046] In another aspect, the toxic marker gene can include two or more selector codons. In another embodiment of the methods, the toxic marker gene is a ribonuclease barnase gene, where the ribonuclease barnase gene comprises at least one amber codon. Optionally, the ribonuclease barnase gene can include two or more amber codons.

[0047] In one embodiment, selecting or screening the pool of (optionally mutant) tRNAs for members that are aminoacylated by an introduced orthogonal RS (O-RS) can include: introducing a positive selection or screening marker gene, wherein the positive marker gene comprises a drug resistance gene (e.g., β -lactamase gene, comprising at least one of the selector codons, such as at least one amber stop codon) or a gene essential to the organism, or a gene that leads to detoxification of a toxic agent, along with the O-RS, and the pool of (optionally mutant) tRNAs into a plurality of cells from the second organism; and, identifying surviving or screened cells grown in the presence of a selection or screening agent, e.g., an antibiotic, thereby providing a pool of cells possessing the at least one recombinant tRNA, where the at least recombinant tRNA is aminoacylated by the O-RS and inserts an amino acid into a translation product encoded by the positive marker gene, in response to the at least one selector codons. In another embodiment, the concentration of the selection and/or screening agent is varied. Recombinant O-tRNAs produced by the methods of the present invention are also included.

[0048] Methods for generating specific O-tRNA/O-RS pairs are provided.

Methods include: (a) generating a library of mutant tRNAs derived from at least one tRNA from a first organism; (b) negatively selecting or screening the library for (optionally mutant) tRNAs that are aminoacylated by an aminoacyl-tRNA synthetase (RS) from a second organism in the absence of a RS from the first organism, thereby providing a pool of (optionally mutant) tRNAs; (c) selecting or screening the pool of (optionally mutant) tRNAs for members that are aminoacylated by an introduced orthogonal RS (O-RS), thereby providing at least one recombinant O-tRNA. The at least one recombinant O-tRNA recognizes a selector codon and is not efficiently recognized by the RS from the

WO 02/086075

PCT/US02/12635

second organism and is preferentially aminoacylated by the O-RS. The method also includes (d) generating a library of (optionally mutant) RSs derived from at least one aminoacyl-tRNA synthetase (RS) from a third organism; (e) selecting or screening the library of mutant RSs for members that preferentially aminoacylate the at least one recombinant O-tRNA in the presence of an unnatural amino acid and a natural amino acid, thereby providing a pool of active (optionally mutant) RSs; and, (f) negatively selecting or screening the pool for active (optionally mutant) RSs that preferentially aminoacylate the at least one recombinant O-tRNA in the absence of the unnatural amino acid, thereby providing the at least one specific O-tRNA/O-RS pair, wherein the at least one specific O-tRNA/O-RS pair comprises at least one recombinant O-RS that is specific for the unnatural amino acid and the at least one recombinant O-tRNA. Specific O-tRNA/O-RS pairs produced by the methods are included. For example, the specific O-tRNA/O-RS pair can include, e.g., a mutRNATyr-mutTyrRS pair, such as a mutRNATyr-SS12TyrRS pair, a mutRNALeu-mutLeuRS pair, a mutRNAThr-mutThrRS pair, a mutRNAGlu-mutGluRS pair, or the like. Additionally, such methods include wherein the first and third organism are the same (e.g., *Methanococcus jannaschii*).

[0049] Methods for selecting an orthogonal tRNA-tRNA synthetase pair for use in an in vivo translation system of a second organism are also included in the present invention. The methods include: introducing a marker gene, a tRNA and an aminoacyl-tRNA synthetase (RS) isolated or derived from a first organism into a first set of cells from the second organism; introducing the marker gene and the tRNA into a duplicate cell set from a second organism; and, selecting for surviving cells in the first set that fail to survive in the duplicate cell set or screening for cells showing a specific screening response that fail to give such response in the duplicate cell set, wherein the first set and the duplicate cell set are grown in the presence of a selection or screening agent, wherein the surviving or screened cells comprise the orthogonal tRNA-tRNA synthetase pair for use in the in vivo translation system of the second organism. In one embodiment, comparing and selecting or screening includes an in vivo complementation assay. The concentration of the selection or screening agent can be varied.

[0050] The organisms of the present invention comprise a variety of organism and a variety of combinations. For example, the first and the second organisms of the methods of the present invention can be the same or different. In one embodiment, the organisms

WO 02/086075

PCT/US02/12635

are optionally a prokaryotic organism, e.g., *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *P. furiosus*, *P. horikoshii*, *A. pernix*, *T. thermophilus*, or the like. Alternatively, the organisms optionally comprise a eukaryotic organism, e.g., plants (e.g., complex plants such as monocots, or dicots), algae, protists, fungi (e.g., yeast, etc), animals (e.g., mammals, insects, arthropods, etc.), or the like. In another embodiment, the second organism is a prokaryotic organism, e.g., *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *Halobacterium*, *P. furiosus*, *P. horikoshii*, *A. pernix*, *T. thermophilus*, or the like. Alternatively, the second organism can be a eukaryotic organism, e.g., a yeast, a animal cell, a plant cell, a fungus, a mammalian cell, or the like. In various embodiments the first and second organisms are different.

[0051] The various methods of the invention (above) optionally comprise wherein selecting or screening comprises one or more positive or negative selection or screening, e.g., a change in amino acid permeability, a change in translation efficiency, and a change in translational fidelity. Additionally, the one or more change is optionally based upon a mutation in one or more gene in an organism in which an orthogonal tRNA-tRNA synthetase pair are used to produce such protein. Selecting and/or screening herein optionally comprises wherein at least 2 selector codons within one or more selection gene or within one or more screening gene are use. Such multiple selector codons are optionally within the same gene or within different screening/selection genes. Additionally, the optional multiple selector codons are optionally different selector codons or comprise the same type of selector codons.

[0052] Kits are an additional feature of the invention. For example, the kits can include one or more translation system as noted above (e.g., a cell), one or more unnatural amino acid, e.g., with appropriate packaging material, containers for holding the components of the kit, instructional materials for practicing the methods herein and/or the like. Similarly, products of the translation systems (e.g., proteins such as EPO analogues comprising unnatural amino acids) can be provided in kit form, e.g., with containers for holding the components of the kit, instructional materials for practicing the methods herein and/or the like.

WO 02/086075

PCT/US02/12635

BRIEF DESCRIPTION OF THE DRAWINGS

[0053] **Figure 1** schematically illustrates site-specific incorporation of unnatural amino acids into proteins in vivo. An orthogonal aminoacyl-tRNA synthetase aminoacylates an orthogonal tRNA with an unnatural amino acid. The acylated orthogonal tRNA inserts the unnatural amino acid at the position specified by a selector codon, e.g., a unique codon, which is introduced into the gene encoding a protein of interest.

[0054] **Figure 2, Panel A and Panel B**, schematically illustrates examples of selection methods for active synthetases that aminoacylate with unnatural amino acids.

- 10 **Panel A** illustrates the general selection/screen for aminoacyl-tRNA synthetases with unnatural amino acids specificities. In the positive selection, active synthetases with either natural or unnatural amino acid specificities are identified; in the negative selection, synthetases with natural amino acid specificities are eliminated. Only synthetases charging the orthogonal tRNA with the unnatural amino acid can survive both
- 15 selections/screens. **Panel B** schematically illustrates one embodiment of the selection/screen for synthetases preferentially aminoacylating an O-tRNA with an unnatural amino acid. For example, expression vectors containing an orthogonal suppressor tRNA and a member of a library of mutated RS with a positive selection marker, e.g., β -lactamase, with a selector codon, e.g., an amber codon, are introduced into
- 20 an organism and grown in the presence a selector agent, e.g., ampicillin. The expression of the positive selection marker allows the cell to survive in the selection agent. Survivors encode synthetases capable of charging any natural or unnatural amino acid (aa) onto the O-tRNA. The active synthetases are transformed into a second strain in the expression vector, and an expression vector with a negative selection marker, e.g., a toxic gene, such
- 25 as barnase, that when expressed kills the cells, with one or more selector codons, e.g., TAG. The cells are grown without the unnatural amino acid. If the synthetase provided aminoacylates the O-tRNA with a natural amino acid, the negative selection marker is expressed and the cell dies. If the synthetase preferentially aminoacylates the O-tRNA, no negative selection marker is expressed, because there is no unnatural amino acid and the
- 30 cell lives. This provides at least one orthogonal synthetase that preferentially aminoacylates the O-tRNA with the desired unnatural codon.

WO 02/086075

PCT/US02/12635

[0055] **Figure 3** illustrates site-specific mutations to generate directed libraries for tyrosine analogues.

[0056] **Figure 4** illustrates a consensus sequence for pentafluorophenylalanine selection to generate directed libraries for these analogues.

- 5 [0057] **Figure 5** schematically illustrates the transplantation of one domain, e.g., the CPI domain, from one organism, e.g., *Escherichia coli*, to the synthetase of other organism, e.g., *Methanococcus jannaschii* TyrRS.

[0058] **Figure 6** schematically illustrates the construction of chimeric *Methanococcus jannaschii*/*Escherichia coli* synthetases.

- 10 [0059] **Figure 7** schematically illustrates the generation of a library of chimeric synthetases, e.g., *Methanococcus jannaschii*/*Escherichia coli* synthetases.

[0060] **Figure 8** schematically illustrates an example for selection of suppressor tRNAs that are poor substrates for an endogenous synthetases, e.g., an *Escherichia coli* synthetase, and that are charged efficiently by a cognate synthetase of interest. Expression

- 15 vectors that contain a member of a mutated tRNA library and another vector with a negative selection marker, e.g., a toxic gene, such as barnase, with one or more selector codons are introduced into a cell of an organism. Survivors of the negative selection encode mutated tRNAs that are either orthogonal to the organism or non-functional. The vectors from the survivors are isolated and transformed into other cells along with a
20 positive selection marker, e.g., β -lactamase gene, with a selector codon. The cells are grown in the presence of a selection agent, e.g., ampicillin, and an RS from an organism from the same source, e.g., *Methanococcus jannaschii*, as the tRNA. Survivors of this selection encode mutant tRNA that are orthogonal to the cell's synthetases, e.g., *Escherichia coli*'s synthetases, and aminoacylated by RS from the same source as the
25 tRNA.

[0061] **Figure 9, Panel A and B**, schematically illustrates a mutated anticodon-loop tRNA library, **Panel A**, and a mutated all-loop library, **Panel B**, from *Methanococcus jannaschii* tRNA_{TyrCUA}. Randomly mutated nucleotides (N) are shaded in black.

- [0062] **Figure 10** schematically illustrates examples of structures of unnatural base
30 pairs which pair by forces other than hydrogen bonding (PICS:PICS, 3MN:3MN, 7AI:7AI, Dipic:Py).

WO 02/086075

PCT/US02/12635

[0063] **Figure 11** is a graph of results of a negative selection method for suppressor tRNAs, which shows the percentage of surviving cells containing one of three constructs, for a given amount of time based on the suppression of two amber codons in the barnase gene introduced by a vector, e.g., plasmid pSCB2. This plasmid encodes the barnase gene containing two amber codons. Selections are carried out in GMM liquid medium, and 20 mM of arabinose is used to induce barnase expression. Three constructs are indicated by the following: (1) a circle which represents a control plasmid with no suppressor tRNA; (2) a triangle which represents a suppressor tRNA on plasmid, pAC-YYG1; and, (3) a square which represents a suppressor tRNA on plasmid, pAC-JY.

[0064] **Figure 12** displays growth histograms, illustrating positive selection based on the suppression of an amber codon in the β -lactamase gene. A vector encoding a suppressor tRNA, e.g., pAC plasmid, is cotransformed with a vector encoding a synthetase, e.g., pBLAM-JYRS, in an organism, e.g., *Escherichia coli* DH10B cells. The growth of cells harboring synthetase and different pAC plasmids in liquid 2X YT medium with various concentrations of ampicillin, e.g., 0, 100 and 500 μ g/ml, is shown in **Panel A**, where pAC is a control plasmid with no suppressor tRNA, where pAC-YYG1 is a plasmid with a suppressor tRNA, and where pAC-JY is a plasmid with a suppressor tRNA. **Panel B** shows positive selection of the same constructs using 2X YT agar plates with 500 μ g/ml ampicillin. Three constructs are indicated by the following: (1) a circle which represents a control plasmid with no suppressor tRNA; (2) a triangle which represents a suppressor tRNA on plasmid, pAC-YYG1; and, (3) a square which represents a suppressor tRNA on plasmid, pAC-JY.

[0065] **Figure 13** illustrates DNA sequences of mutant suppressor tRNAs selected from anticodon-loop and all-loop library. JY stands for the wild-type *Methanococcus jannaschii* tRNACUATyrCUA.

[0066] **Figure 14** schematically illustrates a stereo view of the active site of TyrRS. Residues from *B. stearothermophilus* TyrRS are illustrated in the figure. Corresponding residues from *Methanococcus jannaschii* TyrRS are Tyr³²(Tyr³⁴), Glu¹⁰⁷(Asn¹²³), Asp¹⁵⁸(Asp¹⁷⁶), Ile¹⁵⁹(Phe¹⁷⁷), and Leu¹⁶²(Leu¹⁸⁰) with residues from *B. stearothermophilus* TyrRS in parentheses.

[0067] **Figure 15** schematically illustrates a view of the active site of TyrRS. Residues from *B. stearothermophilus* TyrRS are illustrated in the figure. Corresponding

WO 02/086075

PCT/US02/12635

residues from *Methanococcus jannaschii* TyrRS are Tyr³²(Tyr³⁴), Asp¹⁵⁸(Asp¹⁷⁶), Ile¹⁵⁹(Phe¹⁷⁷), Leu¹⁶²(Leu¹⁸⁰) and Ala¹⁶⁷(Gln¹⁸⁹) with residues from *B. stearrowthermophilus* TyrRS in parentheses.

[0068] **Figure 16, Panel A and Panel B** schematically illustrate an example of FACS-based selection and screening methods used to generate a component of the present invention, e.g., orthogonal synthetase. **Panel A** schematically illustrates vectors, e.g., plasmids, for expression of orthogonal synthetase library and O-tRNA (library plasmid) and for the T7 RNA polymerase/GFP reporter system (reporter plasmid), with one or more selector codons, e.g., TAG. **Panel B** schematically illustrates positive selection/negative screen scheme, where the cells are grown the presence and absence of the unnatural amino acid, the presence and absence of a selection agent, and screened for fluorescing cells and non-fluorescing cells in the screening process, where the "+" and empty circles correspond to fluorescing and non-fluorescing cells, respectively.

[0069] **Figure 17, Panel A, Panel B, Panel C and Panel D** illustrates an amplifiable fluorescence reporter system. **Panel A** schematically illustrates vectors that can be used in the screen, e.g., plasmids, such as pREP, where T7 RNA polymerase transcription is controlled by the ara promoter; protein expression depends on suppression of amber codons at varying locations in the gene. Reporter expression, e.g., GFPuv expression is controlled by T7 RNA polymerase. The reporter vector, e.g., plasmid pREP, is compatible for use with a vector for expressing an orthogonal synthetase/tRNA pair, e.g., a Cole1 plasmid. **Panel B** illustrates compositions and fluorescence enhancement of T7 RNA polymerase gene constructs within pREP (1-12). The construct number is indicated to the left of each. Fluorescence enhancements, indicated to the right of each construct, are calculated as the cell concentration-corrected ratio of fluorescence, as measured fluorimetrically, of cells containing pREP(1-12) and pQ or pQD. The position of the amber mutations within a gene are indicated. **Panel C** illustrates cytometric analysis of cells containing pREP (10) and either pQD (top) or pQ (bottom). **Panel D** illustrates fluorimetric analyses of cells containing pREP (10) and expressing various *Escherichia coli* suppressor tRNAs. "None" indicates that the cells contain no suppressor tRNA.

[0070] **Figure 18** schematically illustrates phage-based selection for the incorporation of unnatural amino acids into a surface epitope. For example, *Escherichia*

WO 02/086075

PCT/US02/12635

coli carrying the mutant synthetase library are infected by phage with a stop codon in a gene encoding a surface protein. Phage containing an active synthetase display the unnatural amino acid on the phage surface and are selected with immobilized monoclonal antibodies.

- 5 [0071] **Figure 19** schematically illustrates an example of a molecule, e.g., immobilized aminoalkyl adenylate analog of the aminoacyl adenylate intermediate, used to screen displayed synthetases, e.g., phage-displayed synthetases, with unnatural amino acid specificity.

- [0072] **Figure 20** is a graph illustrating ampicillin resistance of various orthogonal
10 pairs from a variety of organisms. The figure illustrates an example of finding an orthogonal pair using a reporter constructs, each containing a reporter gene, e.g., a β -lactamase gene, with a selector codon, e.g., an amber codon, and a suppressor tRNA (with a selector anticodon), where the suppressor tRNA can be from a variety of organisms, e.g., *A. fulgidus*, *Halobacterium NRC-1*, *P. furiosus*, *P. horikoshii*, and *Methanococcus*
15 *jannaschii*. The reporter constructs and cloned synthetases from different organisms, e.g., *M. thermoautotrophicum*, *Methanococcus jannaschii*, *P. horikoshii*, *A. permix*, *A. fulgidus*, *Halobacterium NRC-1*, and *Escherichia coli* are transformed into a cell. Cells are grown in various concentrations of a selector agent, e.g., ampicillin. Cells possessing an orthogonal tRNA/RS pair are selected, e.g., using an in vivo complementation assay. As
20 shown, two systems showed suppression levels significant higher than was observed with *Escherichia coli* synthetase. They are *M. thermoautotrophicum* and *Methanococcus jannaschii*.

- [0073] **Figure 21, Panel A and Panel B**, illustrates mutated amber suppressor
tRNAs from a *Halobacterium NRC-1*, which are generated by mutating, e.g., randomizing,
25 the anticodon loop of the leucyl tRNA and selecting (**Panel B**) for more efficient suppression of a selector codon, e.g., an amber codon in a reporter gene(s), e.g., using a combination of selection steps, such as selection based on β -lactamase and selection based on barnase. Panel B illustrates IC50 values in $\mu\text{g/ml}$ of ampicillin for a β -lactamase amber
suppression system with three mutant tRNA constructs, original amber mutant, optimized
30 anticodon loop, and optimized acceptor stem, alone or with an RS, e.g., M₁LRS. The optimized anticodon and optimized acceptor stem gave the highest values in the β -lactamase selection step.

WO 02/086075

PCT/US02/12635

- [0074] **Figure 22** illustrates a tRNA suppressor for a base codon. The tRNA suppressor illustrated in this figure was isolated from a library derived from the *Halobacterium NRC-1* TTG tRNA, where the anticodon loop was randomized with 8 nucleotides and subjected to ampicillin selection with a reporter construct containing a β -lactamase gene with an AGGA codon at the A184 site.
- [0075] **Figure 23 Panels A-D**, illustrates the activity of the dominant synthetase variant from each successful evolution experiment. **Figure 23A** is a photograph illustrating long-wavelength ultraviolet illumination of cells containing pREP/YC-JYCUA and the indicated synthetase variant, grown in either the presence (+) or absence (-) of the corresponding unnatural amino acid. **Figure 23B** illustrates a fluorimetric analysis of cells containing pREP/YC-JYCUA and the indicated synthetase variant, grown in either the presence (left) or absence (right) of the corresponding unnatural amino acid. **Figure 23C** is a table that illustrates a Cm IC₅₀ analysis of cells containing pREP/YC-JYCUA and the indicated synthetase variant, grown in either the presence or absence of the corresponding unnatural amino acid. **Figure 23D** illustrates a protein expression analysis from cells containing pBAD/JYAMB-4TAG and the indicated synthetase variant, grown in either the presence (+) or absence (-) of the corresponding unnatural amino acid.
- [0076] **Figure 24**, illustrates activity comparisons of OAY-RS variants derived using a negative FACS-based screen (OAY-RS(1,3,5)) or negative barnase-based selection (OAY-RS(B)). Cells containing pREP/YC-JYCUA and the indicated synthetase variant were grown in either the presence (solid block, left) or absence (solid block, right) of the corresponding unnatural amino acid and analyzed fluorimetrically. Fluorescence enhancement (bar, back) is calculated as the cell concentration-corrected ratio of fluorescence of cells grown in the presence versus the absence of unnatural amino acid.
- [0077] **Figure 25, Panels A-B**, illustrate components of the multipurpose reporter plasmid system for directing the evolution of *M. jannaschii* TyrRS. **Figure 25A** illustrates plasmid pREP/YC-JYCUA. Plasmid pREP/YC-JYCUA is compatible for use with plasmid pBK and variants. **Figure 25B** illustrates structures of unnatural amino acids used as targets for the evolution of *M. jannaschii* TyrRS.
- [0078] **Figure 26** illustrates the strategy for the evolution of an aminoacyl-tRNA synthetase using plasmid pREP/YC-JYCUA. Fluorescent and non-fluorescent cells are shown in black and white, respectively.

WO 02/086075

PCT/US02/12635

[0079] Figure 27 illustrates a threonyl-tRNA synthetase from *Thermus thermophilus*.

[0080] Figure 28 illustrates the generation of an orthogonal tRNA for a *T. thermophilus* orthogonal threonyl-tRNA/RS.

5 [0081] Figure 29 illustrates exemplary unnatural amino acids as utilized in the current invention.

[0082] Figure 30 illustrates exemplary unnatural amino acids as utilized in the current invention.

10 [0083] Figure 31 illustrates exemplary unnatural amino acids as utilized in the current invention.

DETAILED DESCRIPTION

[0084] Introduction

[0085] Proteins are at the crossroads of virtually every biological process, from photosynthesis and vision to signal transduction and the immune response. These
15 complex functions result from a polyamide based polymer consisting of twenty relatively simple building blocks arranged in a defined primary sequence.

[0086] The present invention includes methods and composition for use in the site-specific incorporation of unnatural amino acids directly into proteins in vivo. Importantly, the unnatural amino acid is added to the genetic repertoire, rather than substituting for one
20 of the common 20 amino acids. The present invention provides methods for generating, methods for identifying and compositions comprising the components used by the biosynthetic machinery to incorporate an unnatural amino acid into a protein. The present invention, e.g., (i) allows the site-selective insertion of one or more unnatural amino acids at any desired position of any protein, (ii) is applicable to both prokaryotic and eukaryotic
25 cells, (iii) enables in vivo studies of mutant proteins in addition to the generation of large quantities of purified mutant proteins, and (iv) is adaptable to incorporate any of a large variety of non-natural amino acids, into proteins in vivo. Thus, in a specific polypeptide sequence a number of different site-selective insertions of unnatural amino acids is possible. Such insertions are optionally all of the same type (e.g., multiple examples of
30 one type of unnatural amino acid inserted at multiple points in a polypeptide) or are

WO 02/086075

PCT/US02/12635

optionally of diverse types (e.g., different unnatural amino acid types are inserted at multiple points in a polypeptide).

[0087] DEFINITIONS

5 **[0088]** Before describing the present invention in detail, it is to be understood that this invention is not limited to particular compositions or biological systems, which can, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting. As used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "a molecule" optionally includes a combination of two or more such molecules, and the like.

[0089] Unless defined otherwise, all scientific and technical terms are understood to have the same meaning as commonly used in the art to which they pertain. For the purpose of the present invention, the following terms are defined below.

15 **[0090]** As used herein, proteins and/or protein sequences are "homologous" when they are derived, naturally or artificially, from a common ancestral protein or protein sequence. Similarly, nucleic acids and/or nucleic acid sequences are homologous when they are derived, naturally or artificially, from a common ancestral nucleic acid or nucleic acid sequence. For example, any naturally occurring nucleic acid can be modified by any available mutagenesis method to include one or more selector codon. When expressed, this mutagenized nucleic acid encodes a polypeptide comprising one or more unnatural amino acid. The mutation process can, of course, additionally alter one or more standard codon, thereby changing one or more standard amino acid in the resulting mutant protein as well. Homology is generally inferred from sequence similarity between two or more
20 nucleic acids or proteins (or sequences thereof). The precise percentage of similarity between sequences that is useful in establishing homology varies with the nucleic acid and protein at issue, but as little as 25% sequence similarity is routinely used to establish homology. Higher levels of sequence similarity, e.g., 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or 99% or more can also be used to establish homology. Methods for
25 determining sequence similarity percentages (e.g., BLASTP and BLASTN using default parameters) are described herein and are generally available.

WO 02/086075

PCT/US02/12635

- [0091] The term "preferentially aminoacylates" refers to an efficiency, e.g., about 70% efficient, about 75% efficient, about 85% efficient, about 90%, about 95%, about 99% or more efficient, at which an O-RS aminoacylates an O-tRNA with an unnatural amino acid compared to a naturally occurring tRNA or starting material used to generate the O-tRNA. The unnatural amino acid is then incorporated into a growing polypeptide chain with high fidelity, e.g., at greater than about 75% efficiency for a given selector codon, at greater than about 80% efficiency for a given selector codon, at greater than about 90% efficiency for a given selector codon, greater than about 95 % efficiency for a given selector codon, or greater than about 99% efficiency for a given selector codon.
- [0092] The term "selector codon" refers to codons recognized by the O-tRNA in the translation process and not recognized by an endogenous tRNA. The O-tRNA anticodon loop recognizes the selector codon on the mRNA and incorporates its amino acid, e.g., an unnatural amino acid, at this site in the polypeptide. Selector codons can include, e.g., nonsense codons, such as, stop codons, e.g., amber, ochre, and opal codons; four or more base codons; codons derived from natural or unnatural base pairs and the like. For a given system, a selector codon can also include one of the natural three base codons, wherein the endogenous system does not use said natural three base codon, e.g., a system that is lacking a tRNA that recognizes the natural three base codon or a system wherein the natural three base codon is a rare codon.
- [0093] As used herein, the term "orthogonal" refers to a molecule (e.g., an orthogonal tRNA (O-tRNA) and/or an orthogonal aminoacyl tRNA synthetase (O-RS)) that is used with reduced efficiency by a system of interest (e.g., a translational system, e.g., a cell). Orthogonal refers to the inability or reduced efficiency, e.g., less than 20% efficient, less than 10% efficient, less than 5% efficient, or e.g., less than 1% efficient, of an orthogonal tRNA and/or orthogonal RS to function in the translation system of interest. For example, an orthogonal tRNA in a translation system of interest aminoacylates any endogenous RS of a translation system of interest with reduced or even zero efficiency, when compared to aminoacylation of an endogenous tRNA by the endogenous RS. In another example, an orthogonal RS aminoacylates any endogenous tRNA in the translation system of interest with reduced or even zero efficiency, as compared to aminoacylation of the endogenous tRNA by an endogenous RS. "Improvement in

WO 02/086075

PCT/US02/12635

orthogonality" refers to enhanced orthogonality compared to a starting material or a naturally occurring tRNA or RS.

[0094] The term "complementary" refers to components of an orthogonal pair, O-tRNA and O-RS that can function together, e.g., the O-RS aminoacylates the O-tRNA.

- 5 [0095] The term "derived from" refers to a component that is isolated from an organism or isolated and modified, or generated, e.g., chemically synthesized, using information of the component from the organism.

[0096] The term "translation system" refers to the components necessary to incorporate a naturally occurring amino acid into a growing polypeptide chain (protein).

- 10 For example, components can include ribosomes, tRNAs, synthetases, mRNA and the like. The components of the present invention can be added to a translation system, in vivo or in vitro.

[0097] The term "inactive RS" refers to a synthetase that have been mutated so that it no longer can aminoacylate its cognate tRNA with an amino acid.

- 15 [0098] The term "selection agent" refers to an agent that when present allows for a selection of certain components from a population, e.g., an antibiotic, wavelength of light, an antibody, a nutrient or the like. The selection agent can be varied, e.g., such as concentration, intensity, etc.

- [0099] The term "positive selection marker" refers to a marker than when present, 20 e.g., expressed, activated or the like, results in identification of an organism with the positive selection marker from those without the positive selection marker.

- [0100] The term "negative selection marker" refers to a marker than when present, e.g., expressed, activated or the like, allows identification of an organism that does not possess the desired property (e.g., as compared to an organism which does possess the 25 desired property).

[0101] The term "reporter" refers to a component that can be used to select components described in the present invention. For example, a reporter can include a green fluorescent protein, a firefly luciferase protein, or genes such as β -gal/lacZ (β -galactosidase), Adh (alcohol dehydrogenase) or the like.

WO 02/086075

PCT/US02/12635

[0102] The term "not efficiently recognized" refers to an efficiency, e.g., less than about 10%, less than about 5%, or less than about 1%, at which a RS from one organism aminoacylates O-tRNA.

[0103] The term "eukaryote" refers to organisms belonging to the phylogenetic domain Eucarya such as animals (e.g., mammals, insects, reptiles, birds, etc.), ciliates, plants, fungi (e.g., yeasts, etc.), flagellates, microsporidia, protists, etc. Additionally, the term "prokaryote" refers to non-eukaryotic organisms belonging to the Eubacteria (e.g., *Escherichia coli*, *Thermus thermophilus*, etc.) and Archaea (e.g., *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium* such as *Haloferax volcanii* and *Halobacterium* species *NRC-1*, *A. fulgidus*, *P. furiosus*, *P. horikoshii*, *A. permix*, etc.) phylogenetic domains

[0104] A "suppressor tRNA" is a tRNA that alters the reading of a messenger RNA (mRNA) in a given translation system. A suppressor tRNA can read through, e.g., a stop codon, a four base codon, or a rare codon.

15 [0105] DISCUSSION

[0106] The present invention relates to methods and compositions for new components of biosynthetic translational machinery that allows for the incorporation of unnatural amino acids into proteins in vivo. Specifically, compositions comprising and methods for generating orthogonal tRNAs and orthogonal-RS and orthogonal tRNAs/orthogonal-RS pairs are provided. These components, when introduced into a host cell, can be used in the translation system of the cell to incorporate an unnatural amino acid in vivo into a polypeptide (protein) of interest. For example, this can provide site-specific unnatural amino acid mutagenesis; or, optionally, random unnatural amino acid mutagenesis. The orthogonal tRNA delivers the unnatural amino acid in response to a selector codon and the orthogonal synthetase preferentially aminoacylates an orthogonal tRNA with the unnatural amino acid. The O-RS does not efficiently aminoacylate the orthogonal tRNA with any of the common twenty amino acids. Methods for making and identifying orthogonal pairs are also provided.

[0107] The site-specific incorporation of unnatural amino acids into proteins in vivo is schematically illustrated in **Figure 1**. A selector codon, e.g., a unique codon, is introduced into a gene of interest. The gene is transcribed into mRNA and conventional

WO 02/086075

PCT/US02/12635

translation begins on the ribosome. Endogenous synthetases aminoacylate endogenous tRNAs with natural amino acids (aa) in the presence of ATP. An orthogonal tRNA is enzymatically aminoacylated by an orthogonal synthetase with an unnatural amino acid in the presence of ATP. When the ribosome encounters a selector codon, an orthogonal tRNA, which is modified to contain a selector anticodon, e.g., a unique anticodon, it is able to decode the mutation as an unnatural amino acid, and translation proceeds to the full-length product with the incorporated unnatural amino acid.

[0108] Orthogonal aminoacyl tRNA synthetase, O-RS

[0109] In order to specifically incorporate an unnatural amino acid in vivo, the substrate specificity of the synthetase is altered so that only the desired unnatural amino acid, but not any common 20 amino acids are charged to the tRNA. If the orthogonal synthetase is promiscuous, it will result in mutant proteins with a mixture of natural and unnatural amino acids at the target position. For instance, in an attempt to site-specifically incorporate *p*-F-Phe, a yeast amber suppressor tRNA^{PheCUA}/phenylalanyl-tRNA synthetase pair was used in a *p*-F-Phe resistant, Phe auxotrophic *Escherichia coli* strain. See, e.g., R. Furter, *Protein Sci.*, 7:419 (1998). Because yeast PheRS does not have high substrate specificity for *p*-F-Phe, the mutagenesis site was translated with 64-75% *p*-F-Phe and the remainder as Phe and Lys even in the excess of *p*-F-Phe added to the growth media. In addition, at the Phe codon positions, 7% *p*-F-Phe was found, indicating that the endogenous *Escherichia coli* PheRS incorporates *p*-F-Phe in addition to Phe. Because of its translational infidelity, this approach is not generally applicable to other unnatural amino acids. Modification of the substrate specificity of a synthetase was expected to be difficult due to the high intrinsic fidelity of the natural synthetases and the fact that unnatural amino acids are not required for any cellular function. The present invention solves this problem and provides composition of, and methods for, generating synthetases that have modified substrate specificity, such as an unnatural amino acid. Using the components of the present invention, the efficiency of incorporation of an unnatural amino acid into is, e.g., greater than about 75%, greater than about 85%, greater than about 95%, greater than about 99% or more.

[0110] Compositions of the present invention include an orthogonal aminoacyl-tRNA synthetase (O-RS), where the O-RS preferentially aminoacylates an orthogonal

WO 02/086075

PCT/US02/12635

tRNA (O-tRNA) with an unnatural amino acid, optionally, in vivo. In one embodiment, the O-RS comprises a nucleic acid comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO: 4-34 (see, Table 5) and a complementary polynucleotide sequence thereof. In another embodiment, the O-RS has improved or enhanced enzymatic properties, e.g., the K_m is lower, the k_{cat} is higher, the value of k_{cat}/K_m is higher or the like, for the unnatural amino acid compared to a naturally occurring amino acid, e.g., one of the 20 known amino acids. Sequences of exemplary O-tRNA and O-RS molecules can be found in Example 10.

[0111] Methods for producing an O-RS are based on generating a pool of mutant synthetases from the framework of a wild-type synthetase, and then selecting for mutated RSs based on their specificity for an unnatural amino acid relative to the common twenty. To isolate such a synthetase, the selection methods of the present invention are: (i) sensitive, as the activity of desired synthetases from the initial rounds can be low and the population small; (ii) "tunable", since it is desirable to vary the selection stringency at different selection rounds; and, (iii) general, so that it can be used for different unnatural amino acids.

[0112] The present invention provides methods to generate an orthogonal aminoacyl tRNA synthetase by mutating the synthetase, e.g., at the active site in the synthetase, at the editing mechanism site in the synthetase, at different sites by combining different domains of synthetases, or the like, and applying a selection process. **Figure 2, Panel A** schematically illustrates an in vivo selection/screen strategy, which is based on the combination of a positive selection followed by a negative selection. In the positive selection, suppression of the selector codon introduced at a nonessential position(s) of a positive marker allows cells to survive under positive selection pressure. In the presence of both natural and unnatural amino acids, survivors thus encode active synthetases charging the orthogonal suppressor tRNA with either a natural or unnatural amino acid. In the negative selection, suppression of a selector codon introduced at a nonessential position(s) of a negative marker removes synthetases with natural amino acid specificities. Survivors of the negative and positive selection encode synthetases that aminoacylate (charge) the orthogonal suppressor tRNA with unnatural amino acids only. These synthetases can then be subjected to further mutagenesis, e.g., DNA shuffling or other recursive mutagenesis methods. Of course, in other embodiments, the invention optional

WO 02/086075

PCT/US02/12635

can utilize different orders of steps to identify (e.g., O-RS, O-tRNA, pairs, etc.), e.t., negative selection/screening followed by positive selection/screening or vice versa or any such combinations thereof.

[0113] For example, *see*, Figure 2, Panel B. In Figure 2, Panel B, a selector codon, e.g., an amber codon, is placed in a reporter gene, e.g., an antibiotic resistance gene, such as β -lactamase, with a selector codon, e.g., TAG. This is placed in an expression vector with members of the mutated RS library. This expression vector along with an expression vector with an orthogonal tRNA, e.g., a orthogonal suppressor tRNA, are introduced into a cell, which is grown in the presence of a selection agent, e.g., antibiotic media, such as ampicillin. Only if the synthetase is capable of aminoacylating (charging) the suppressor tRNA with some amino acid does the selector codon get decoded allowing survival of the cell on antibiotic media.

[0114] Applying this selection in the presence of the unnatural amino acid, the synthetase genes that encode synthetases that have some ability to aminoacylate are selected away from those synthetases that have no activity. The resulting pool of synthetases can be charging any of the 20 naturally occurring amino acids or the unnatural amino acid. To further select for those synthetases that exclusively charge the unnatural amino acid, a second selection, e.g., a negative selection, is applied. In this case, an expression vector containing a negative selection marker and an O-tRNA is used, along with an expression vector containing a member of the mutated RS library. This negative selection marker contains at least one selector codon, e.g., TAG. These expression vectors are introduced into another cell and grown without unnatural amino acids and, optionally, a selection agent, e.g., tetracycline. In the negative selection, those synthetases with specificities for natural amino acids charge the orthogonal tRNA, resulting in suppression of a selector codon in the negative marker and cell death. Since no unnatural amino acid is added, synthetases with specificities for the unnatural amino acid survive. For example, a selector codon, e.g., a stop codon, is introduced into the reporter gene, e.g., a gene that encodes a toxic protein, such as barnase. If the synthetase is able to charge the suppressor tRNA in the absence of unnatural amino acid, the cell will be killed by translating the toxic gene product. Survivors passing both selection/screens encode synthetases specifically charging the orthogonal tRNA with an unnatural amino acid.

WO 02/086075

PCT/US02/12635

[0115] In one embodiment, methods for producing at least one recombinant orthogonal aminoacyl-tRNA synthetase (O-RS) include: (a) generating a library of mutant RSs derived from at least one aminoacyl-tRNA synthetase (RS) from a first organism; (b) selecting the library of mutant RSs for members that aminoacylate an orthogonal tRNA (O-tRNA) in the presence of an unnatural amino acid and a natural amino acid, thereby providing a pool of active mutant RSs; and, (c) negatively selecting the pool for active mutant RSs that preferentially aminoacylate the O-tRNA in the absence of the unnatural amino acid, thereby providing the at least one recombinant O-RS; wherein the at least one recombinant O-RS preferentially aminoacylates the O-tRNA with the unnatural amino acid. Optionally, more mutations are introduced by mutagenesis, e.g., random mutagenesis, recombination or the like, into the selected synthetase genes to generate a second-generation synthetase library, which is used for further rounds of selection until a mutant synthetase with desired activity is evolved. Recombinant O-RSs produced by the methods are included in the present invention. As explained below, orthogonal tRNA/synthetase pairs or the invention are also optionally generated by importing such from a first organism into a second organism.

[0116] In one embodiment, the RS is an inactive RS. The inactive RS can be generated by mutating an active RS. For example, the inactive RS can be generated by mutating at least about 5 amino acids to different amino acids, e.g., alanine.

[0117] The library of mutant RSs can be generated using various mutagenesis techniques known in the art. For example, the mutant RSs can be generated by site-specific mutations, random point mutations, in vitro homologous recombination, chimeric constructs or the like. In one embodiment, mutations are introduced into the editing site of the synthetase to hamper the editing mechanism and/or to alter substrate specificity. See, e.g., **Figure 3** and **Figure 4**. **Figure 3** illustrates site-specific mutations to generate directed libraries for tyrosine analogues. **Figure 4** illustrates a consensus sequence for pentafluorophenylalanine selection to generate directed libraries for these analogues. Libraries of mutant RSs also include chimeric synthetase libraries, e.g., libraries of chimeric *Methanococcus jannaschii*/*Escherichia coli* synthetases. The domain of one synthetase can be added or exchanged with a domain from another synthetase. **Figure 5** schematically illustrates the transplantation of one domain, e.g., the CPI domain, from one organism, e.g., *Escherichia coli*, to the synthetase of other organism, e.g., *Methanococcus*

WO 02/086075

PCT/US02/12635

jannaschii TyrRS. CPI can be transplanted from *Escherichia coli* TyrRS to *H. sapiens* TyrRS. See, e.g., Wakasugi, K., et al., *EMBO J.* 17:297-305 (1998). **Figure 6** schematically illustrates the construction of chimeric *Methanococcus*

jannaschii/*Escherichia coli* synthetases and **Figure 7** schematically illustrates the

5 generation of a library of chimeric synthetases, e.g., *Methanococcus*

jannaschii/*Escherichia coli* synthetases. See, e.g., Sieber, et al., *Nature Biotechnology*, 19:456-460 (2001). The chimeric library is screened for a variety of properties, e.g., for members that are expressed and in frame, for members that lack activity with a desired synthetase, and/or for members that show activity with a desired synthetase.

10 **[0118]** In one embodiment, the positive selection step includes: introducing a positive selection marker, e.g., an antibiotic resistance gene, or the like, and the library of mutant RSs into a plurality of cells, wherein the positive selection marker comprises at least one selector codon, e.g., an amber codon; growing the plurality of cells in the presence of a selection agent; selecting cells that survive in the presence of the selection
15 agent by suppressing the at least one selector codon in the positive selection marker, thereby providing a subset of positively selected cells that contains the pool of active mutant RSs. Optionally, the selection agent concentration can be varied.

[0119] In one embodiment, negative selection includes: introducing a negative selection marker with the pool of active mutant RSs from the positive selection into a
20 plurality of cells of a second organism, wherein the negative selection marker is an antibiotic resistance gene, e.g., a chloramphenicol acetyltransferase (CAT) gene, comprising at least one selector codon; and, selecting cells that survive in a 1st media supplemented with the unnatural amino acid and a selection agent, but fail to survive in a 2nd media not supplemented with the unnatural amino acid and the selection agent,
25 thereby providing surviving cells with the at least one recombinant O-RS. Optionally, the concentration of the selection agent is varied.

[0120] The 1st and 2nd media described above can include, e.g., a direct replica plate method. For example, after passing the positive selection, cells are grown in the presence of either ampicillin or chloramphenicol and the absence of the unnatural amino
30 acid. Those cells that do not survive are isolated from a replica plate supplemented with the unnatural amino acid. No transformation into a second negative selection strain is needed, and the phenotype is known. Compared to other potential selection markers, a

WO 02/086075

PCT/US02/12635

positive selection based on antibiotic resistance offers the ability to tune selection stringency by varying the concentration of the antibiotic, and to compare the suppression efficiency by monitoring the highest antibiotic concentration cells can survive. In addition, the growth process is also an enrichment procedure. This can lead to a quick accumulation of the desired phenotype.

[0121] In another embodiment, negatively selecting the pool for active mutant RSs includes: isolating the pool of active mutant RSs from the positive selection step (b); introducing a negative selection marker, wherein the negative selection marker is a toxic marker gene, e.g., a ribonuclease barnase gene, comprising at least one selector codon, and the pool of active mutant RSs into a plurality of cells of a second organism; and selecting cells that survive in a 1st media not supplemented with the unnatural amino acid, but fail to survive in a 2nd media supplemented with the unnatural amino acid, thereby providing surviving cells with the at least one recombinant O-RS, wherein the at least one recombinant O-RS is specific for the unnatural amino acid. Optionally, the negative selection marker comprises two or more selector codons.

[0122] In one aspect, positive selection is based on suppression of a selector codon in a positive selection marker, e.g., a chloramphenicol acetyltransferase (CAT) gene comprising a selector codon, e.g., an amber stop codon, in the CAT gene, so that chloramphenicol can be applied as the positive selection pressure. In addition, the CAT gene can be used as both a positive marker and negative marker as describe herein in the presence and absence of unnatural amino acid. Optionally, the CAT gene comprising a selector codon is used for the positive selection and a negative selection marker, e.g., a toxic marker, such as a barnase gene comprising at least one or more selector codons, is used for the negative selection.

[0123] In another aspect, positive selection is based on suppression of a selector codon at nonessential position in the β -lactamase gene, rendering cells ampicillin resistant; and a negative selection using the ribonuclease barnase as the negative marker is used. In contrast to β -lactamase, which is secreted into the periplasm, CAT localizes in the cytoplasm; moreover, ampicillin is bacteriocidal, while chloramphenicol is bacteriostatic.

[0124] The recombinant O-RS can be further mutated and selected. In one embodiment, the methods for producing at least one recombinant orthogonal aminoacyl-tRNA synthetase (O-RS) can further comprise: (d) isolating the at least one recombinant

WO 02/086075

PCT/US02/12635

O-RS; (e) generating a second set of mutated O-RS derived from the at least one recombinant O-RS; and, (f) repeating steps (b) and (c) until a mutated O-RS is obtained that comprises an ability to preferentially aminoacylate the O-tRNA. Optionally, steps (d)-(f) are repeated, e.g., at least about two times. In one aspect, the second set of mutated
 5 O-RS can be generated by mutagenesis, e.g., random mutagenesis, site-specific mutagenesis, recombination or a combination thereof.

[0125] The stringency of the selection steps, e.g., the positive selection step (b), the negative selection step (c) or both the positive and negative selection steps (b) and (c), in the above described-methods, optionally include varying the selection stringency. For
 10 example, because barnase is an extremely toxic protein, the stringency of the negative selection can be controlled by introducing different numbers of selector codons into the barnase gene. In one aspect of the present invention, the stringency is varied because the desired activity can be low during early rounds. Thus, less stringent selection criteria are applied in early rounds and more stringent criteria are applied in later rounds of selection.

[0126] Other types of selections can be used in the present invention for, e.g., O-RS, O-tRNA, and O-tRNA/O-RS pair. For example, the positive selection step (b), the negative selection step (c) or both the positive and negative selection steps (b) and (c) can include using a reporter, wherein the reporter is detected by fluorescence-activated cell sorting (FACS). For example, a positive selection can be done first with a positive
 20 selection marker, e.g., chloramphenicol acetyltransferase (CAT) gene, where the CAT gene comprises a selector codon, e.g., an amber stop codon, in the CAT gene, which followed by a negative selection screen, that is based on the inability to suppress a selector codon(s), e.g., two or more, at positions within a negative marker, e.g., T7 RNA polymerase gene. In one embodiment, the positive selection marker and the negative
 25 selection marker can be found on the same vector, e.g., plasmid. Expression of the negative marker drives expression of the reporter, e.g., green fluorescent protein (GFP). The stringency of the selection and screen can be varied, e.g., the intensity of the light need to fluorescence the reporter can be varied. In another embodiment, a positive selection can be done with a reporter as a positive selection marker, which is screened by
 30 FACS, followed by a negative selection screen, that is based on the inability to suppress a selector codon(s), e.g., two or more, at positions within a negative marker, e.g., barnase gene.

WO 02/086075

PCT/US02/12635

[0127] Optionally, the reporter is displayed on a cell surface, e.g., on a phage display or the like. Cell-surface display, e.g., the OmpA-based cell-surface display system, relies on the expression of a particular epitope, e.g., a poliovirus C3 peptide fused to an outer membrane porin OmpA, on the surface of the *Escherichia coli* cell. The epitope is displayed on the cell surface only when a selector codon in the protein message is suppressed during translation. The displayed peptide then contains the amino acid recognized by one of the mutant aminoacyl-tRNA synthetases in the library, and the cell containing the corresponding synthetase gene can be isolated with antibodies raised against peptides containing specific unnatural amino acids. The OmpA-based cell-surface display system was developed and optimized by Georgiou et al. as an alternative to phage display. See, 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).

[0128] Other embodiments of the present invention include carrying one or more of the selection steps in vitro. The selected component, e.g., synthetase and/or tRNA, can then be introduced into a cell for use in in vivo incorporation of an unnatural amino acid.

[0129] **Orthogonal tRNA**

[0130] Compositions of an orthogonal tRNA (O-tRNA) are also a feature of the invention, e.g., where the O-tRNA recognizes a selector codon and the O-tRNA is preferentially aminoacylated with an unnatural amino acid by an orthogonal aminoacyl-tRNA synthetase. In one embodiment, the O-tRNA comprises a nucleic acid comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO: 4-34 (see, Table 5) and a complementary polynucleotide sequence thereof.

[0131] Methods for producing a recombinant orthogonal tRNA (O-tRNA) are provided herein. For example, to improve the orthogonality of a tRNA while preserving its affinity toward a desired RS, the methods include a combination of negative and positive selections with a mutant suppressor tRNA library in the absence and presence of the cognate synthetase, respectively. See, Figure 8. In the negative selection, a selector codon(s) is introduced in a marker gene, e.g., a toxic gene, such as barnase, at a nonessential position. When a member of the mutated tRNA library, e.g., derived from *Methanococcus jannaschii*, is aminoacylated by endogenous host, e.g., *Escherichia coli* synthetases (i.e., it is not orthogonal to the host, e.g., *Escherichia coli* synthetases), the

WO 02/086075

PCT/US02/12635

selector codon, e.g., an amber codon, is suppressed and the toxic gene product produced leads to cell death. Cells harboring orthogonal tRNAs or non-functional tRNAs survive. Survivors are then subjected to a positive selection in which a selector codon, e.g., an amber codon, is placed in a positive marker gene, e.g., a drug resistance gene, such as a β -lactamase gene. These cells also contain an expression vector with a cognate RS. These cells are grown in the presence of a selection agent, e.g., ampicillin. tRNAs are then selected for their ability to be aminoacylated by the coexpressed cognate synthetase and to insert an amino acid in response to this selector codon. Cells harboring non-functional tRNAs, or tRNAs that cannot be recognized by the synthetase of interest are sensitive to the antibiotic. Therefore, tRNAs that: (i) are not substrates for endogenous host, e.g., *Escherichia coli*, synthetases; (ii) can be aminoacylated by the synthetase of interest; and (iii) are functional in translation survive both selections.

[0132] Methods of producing a recombinant O-tRNA include: (a) generating a library of mutant tRNAs derived from at least one tRNA, e.g., a suppressor tRNA, from a first organism; (b) negatively selecting the library for mutant tRNAs that are aminoacylated by an aminoacyl-tRNA synthetase (RS) from a second organism in the absence of a RS from the first organism, thereby providing a pool of mutant tRNAs; and, (c) selecting the pool of mutant tRNAs for members that are aminoacylated by an introduced orthogonal RS (O-RS), thereby providing at least one recombinant O-tRNA; wherein the at least one recombinant O-tRNA recognizes a selector codon and is not efficiently recognized by the RS from the second organism and is preferentially aminoacylated by the O-RS. In one embodiment, the recombinant O-tRNA possesses an improvement of orthogonality.

[0133] Libraries of mutated tRNA are constructed. See, for example, Figure 9. Mutations can be introduced at a specific position(s), e.g., at a nonconservative position(s), or at a conservative position, at a randomized position(s), or a combination of both in a desired loop of a tRNA, e.g., an anticodon loop, (D arm, V loop, T ψ C arm) or a combination of loops or all loops. Chimeric libraries of tRNA are also included in the present invention. It should be noted that libraries of tRNA synthetases from various organism (e.g., microorganisms such as eubacteria or archaebacteria) such as libraries comprising natural diversity (such as libraries that comprise natural diversity (see, e.g.,

WO 02/086075

PCT/US02/12635

U.S. Patent No. 6,238,884 to Short et al. and references therein, U.S. Patent No. 5,756,316 to Schallenberger et al; U.S. Patent No. 5,783,431 to Petersen et al; U.S. Patent No. 5,824,485 to Thompson et al; and U.S. Patent No. 5,958,672 to Short et al), are optionally constructed and screened for orthogonal pairs.

- 5 [0134] In one embodiment, negatively selecting the library for mutant tRNAs that are aminoacylated by an aminoacyl-tRNA synthetase (step (b) above) includes: introducing a toxic marker gene, wherein the toxic marker gene comprises at least one of the selector codons and the library of mutant tRNAs into a plurality of cells from the second organism; and, selecting surviving cells, wherein the surviving cells contain the
- 10 pool of mutant tRNAs comprising at least one orthogonal tRNA or nonfunctional tRNA. For example, the toxic marker gene is optionally a ribonuclease barnase gene, wherein the ribonuclease barnase gene comprises at least one amber codon. Optionally, the ribonuclease barnase gene can include two or more amber codons. The surviving cells can be selected, e.g., by using a comparison ratio cell density assay.
- 15 [0135] In one embodiment, selecting the pool of mutant tRNAs for members that are aminoacylated by an introduced orthogonal RS (O-RS) can include: introducing a positive selection marker gene, wherein the positive selection marker gene comprises a drug resistance gene, e.g., a β -lactamase gene, comprising at least one of the selector codons, e.g., a β -lactamase gene comprising at least one amber stop codon, the O-RS, and
- 20 the pool of mutant tRNAs into a plurality of cells from the second organism; and, selecting surviving cells grown in the presence of a selection agent, e.g., an antibiotic, thereby providing a pool of cells possessing the at least one recombinant tRNA, wherein the recombinant tRNA is aminoacylated by the O-RS and inserts an amino acid into a translation product encoded by the positive marker gene, in response to the at least one
- 25 selector codons. In another embodiment, the concentration of the selection agent is varied. Recombinant O-tRNAs produced by the methods are included in the present invention.
- [0136] As described above for generating O-RS, the stringency of the selection steps can be varied. In addition, other selection/screening procedures, which are described herein, such as FACs, cell and phage display can also be used.

- 30 [0137] Selector Codons

WO 02/086075

PCT/US02/12635

- [0138] Selector codons of the present invention expand the genetic codon framework of protein biosynthetic machinery. For example, a selector codon includes, e.g., a unique three base codon, a nonsense codon, such as a stop codon, e.g., an amber codon, or an opal codon, an unnatural codon, a four (or more) base codon or the like. A number of selector codons can be introduced into a desired gene, e.g., one or more, two or more, more than three, etc. Additionally, it will be appreciated that multiple different (or similar or identical) unnatural amino acids can thus be incorporated precisely into amino acids (i.e., through use of the multiple selector codons).
- [0139] The 64 genetic codons code for 20 amino acids and 3 stop codons. Because only one stop codon is needed for translational termination, the other two can in principle be used to encode nonproteinogenic amino acids. The amber stop codon, UAG, has been successfully used in in vitro biosynthetic system and in *Xenopus* oocytes to direct the incorporation of unnatural amino acids. Among the 3 stop codons, UAG is the least used stop codon in *Escherichia coli*. Some *Escherichia coli* strains contain natural suppressor tRNAs, which recognize UAG and insert a natural amino acid in response to UAG. In addition, these amber suppressor tRNAs have been widely used in conventional protein mutagenesis. Different species preferentially use different codons for their natural amino acids, such preferentiality is optionally utilized in designing/choosing the selector codons herein.
- [0140] Although discussed with reference to unnatural amino acids herein, it will be appreciated that a similar strategy can be used incorporate a natural amino acid in response to a particular selector codon. That is, a synthetase can be modified to load a natural amino acid onto an orthogonal tRNA that recognizes a selector codon in a manner similar to the loading of an unnatural amino acid as described throughout.
- [0141] In one embodiment, the methods involve the use of a selector codon that is a stop codon for the incorporation of unnatural amino acids in vivo. For example, an O-tRNA is generated that recognizes the stop codon, e.g., UAG, and is aminoacylated by an O-RS with a desired unnatural amino acid. This O-tRNA is not recognized by the naturally occurring aminoacyl-tRNA synthetases. Conventional site-directed mutagenesis can be used to introduce the stop codon, e.g., TAG, at the site of interest in the protein gene. See, e.g., Sayers, J. R., Schmidt, W. Eckstein, F. 5', 3' *Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis*. *Nucleic Acids Res*, 791-

WO 02/086075

PCT/US02/12635

802 (1988). When the O-RS, O-tRNA and the mutant gene are combined in vivo, the unnatural amino acid is incorporated in response to the UAG codon to give a protein containing the unnatural amino acid at the specified position.

[0142] The incorporation of unnatural amino acids in vivo can be done without significant perturbation of the host, e.g., *Escherichia coli*. For example, because the suppression efficiency for the UAG codon depends upon the competition between the O-tRNA, e.g., the amber suppressor tRNA, and the release factor 1 (RF1) (which binds to the UAG codon and initiates release of the growing peptide from the ribosome), the suppression efficiency can be modulated by, e.g., either increasing the expression level of O-tRNA, e.g., the suppressor tRNA, or using an RF1 deficient strain. Additionally, suppression efficiency and unnatural amino acid uptake by carrying out random mutagenesis on an organism or on a portion of an organism's genome and performing proper selection using, e.g., one of the reporter systems described herein.

[0143] Unnatural amino acids can also be encoded with rare codons. For example, when the arginine concentration in an in vitro protein synthesis reaction is reduced, the rare arginine codon, AGG, has proven to be efficient for insertion of Ala by a synthetic tRNA acylated with alanine. See, e.g., C. H. Ma, W. Kudlicki, O. W. Odom, G. Kramer and B. Hardesty, *Biochemistry*, 32:7939 (1993). In this case, the synthetic tRNA competes with the naturally occurring tRNA^{Arg}, which exists as a minor species in *Escherichia coli*. Some organisms do not use all triplet codons. An unassigned codon AGA in *Micrococcus luteus* has been utilized for insertion of amino acids in an in vitro transcription/translation extract. See, e.g., A. K. Kowal and J. S. Oliver, *Nucl. Acid. Res.*, 25:4685 (1997). Components of the present invention can be generated to use these rare codons in vivo.

[0144] Selector codons also comprise four or more base codons, such as, four, five six or more. Examples of four base codons include, e.g., AGGA, CUAG, UAGA, CCCU and the like. Examples of five base codons include, e.g., AGGAC, CCCCUC, CUAGA, CUACU, UAGGC and the like. For example, in the presence of mutated O-tRNAs, e.g., a special frameshift suppressor tRNAs, with anticodon loops, e.g., with at least 8-10 nt anticodon loops, the four or more base codon is read as single amino acid. In other embodiments, the anticodon loops can decode, e.g., at least a four-base codon, at least a five-base codon, or at least a six-base codon or more. Since there are 256 possible

WO 02/086075

PCT/US02/12635

four-base codons, multiple unnatural amino acids can be encoded in the same cell using the four or more base codon. *See also*, J. Christopher Anderson et al., *Exploring the Limits of Codon and Anticodon Size*, Chemistry and Biology, Vol. 9, 237-244 (2002); Thomas J. Magliery, *Expanding the Genetic Code: Selection of Efficient Suppressors of*
 5 *Four-base Codons and Identification of "Shifty" Four-base Codons with a Library Approach in Escherichia coli*, J. Mol. Biol., 307: 755-769 (2001).

[0145] Methods of the present invention include using extended codons based on frameshift suppression. Four or more base codons can insert, e.g., one or multiple
 10 unnatural amino acids into the same protein. For example, four-base codons have been used to incorporate unnatural amino acids into proteins using in vitro biosynthetic methods. *See, e.g.*, C. H. Ma, W. Kudlicki, O. W. Odom, G. Kramer and B. Hardesty, Biochemistry, 1993, 32, 7939 (1993); and, T. Hoshaka, D. Kajihara, Y. Ashizuka, H. Murakami and M. Sisido, J. Am. Chem. Soc., 121:34 (1999). CGGG and AGGU were
 15 used to simultaneously incorporate 2-naphthylalanine and an NBD derivative of lysine into streptavidin in vitro with two chemically acylated frameshift suppressor tRNAs. *See, e.g.*, T. Hoshaka, Y. Ashizuka, H. Sasaki, H. Murakami and M. Sisido, J. Am. Chem. Soc., 121:12194 (1999). In an in vivo study, Moore et al. examined the ability of tRNA^{Leu} derivatives with NCUA anticodons to suppress UAGN codons (N can be U, A, G, or C), and found that the quadruplet UAGA can be decoded by a tRNA^{Leu} with a UCUA
 20 anticodon with an efficiency of 13 to 26% with little decoding in the 0 or -1 frame. *See*, B. Moore, B. C. Persson, C. C. Nelson, R. F. Gesteland and J. F. Atkins, J. Mol. Biol., 298:195 (2000). In one embodiment, extended codons based on rare codons or nonsense codons can be used in present invention, which can reduce missense readthrough and frameshift suppression at unwanted sites.

25 [0146] A translational bypassing system can also be used to incorporate an unnatural amino acid in a desired polypeptide. In a translational bypassing system, a large sequence is inserted into a gene but is not translated into protein. The sequence contains a structure that serves as a cue to induce the ribosome to hop over the sequence and resume translation downstream of the insertion.

30 [0147] Alternatively, or in combination with others methods described above to incorporate an unnatural amino acid in a polypeptide, a trans-translation system can be used. This system involves a molecule called tmRNA present in *Escherichia coli*. This

WO 02/086075

PCT/US02/12635

RNA molecule is structurally related to an alanyl tRNA and is aminoacylated by the alanyl synthetase. The difference between tmRNA and tRNA is that the anticodon loop is replaced with a special large sequence. This sequence allows the ribosome to resume translation on sequences that have stalled using an open reading frame encoded within the tmRNA as template. In the present invention, an orthogonal tmRNA can be generated that is preferentially aminoacylated with an orthogonal synthetase and loaded with an unnatural amino acid. By transcribing a gene by the system, the ribosome stalls at a specific site; the unnatural amino acid is introduced at that site, and translation resumes using the sequence encoded within the orthogonal tmRNA.

10 [0148] Selector codons also optionally include unnatural base pairs. These unnatural base pairs further expand the existing genetic alphabet. One extra base pair increases the number of triplet codons from 64 to 125. Properties of third base pairs include stable and selective base pairing, efficient enzymatic incorporation into DNA with high fidelity by a polymerase, and the efficient continued primer extension after synthesis of the nascent unnatural base pair. Descriptions of unnatural base pairs which can be adapted for methods and compositions include, e.g., Hirao, et al., *An unnatural base pair for incorporating amino acid analogues into protein*, *Nature Biotechnology*, 20:177-182 (2002). Other publications are listed below.

20 [0149] For in vivo usage, the unnatural nucleoside is membrane permeable and is phosphorylated to form the corresponding triphosphate. In addition, the increased genetic information is stable and not destroyed by cellular enzymes. Previous efforts by Benner and others took advantage of hydrogen bonding patterns that are different from those in canonical Watson-Crick pairs, the most noteworthy example of which is the iso-C:iso-G pair. See, e.g., C. Switzer, S. E. Moroney and S. A. Benner, *J. Am. Chem. Soc.*, 111:8322 (1989); and, J. A. Piccirilli, T. Krauch, S. E. Moroney and S. A. Benner, *Nature*, 1990, 343:33 (1990); and E. T. Kool, *Curr. Opin. Chem. Biol.*, 4:602 (2000). These bases in general mispair to some degree with natural bases and cannot be enzymatically replicated. Kool and co-workers demonstrated that hydrophobic packing interactions between bases can replace hydrogen bonding to drive the formation of base pair. See, E. T. Kool, *Curr. Opin. Chem. Biol.*, 4:602 (2000); and, K. M. Guckian and E. T. Kool, *Angew. Chem. Int. Ed. Engl.*, 36, 2825 (1998). In an effort to develop an unnatural base pair satisfying all the above requirements, Schultz, Romesberg and co-workers have systematically synthesized

WO 02/086075

PCT/US02/12635

- and studied a series of unnatural hydrophobic bases. The PICS:PICS self-pair, which is shown in **Figure 10**, is found to be more stable than natural base pairs, and can be efficiently incorporated into DNA by the Klenow fragment of *Escherichia coli* DNA polymerase I (KF). *See, e.g.*, D. L. McMin, A. K. Ogawa, Y. Q. Wu, J. Q. Liu, P. G. Schultz and F. E. Romesberg, J. Am. Chem. Soc., 121:11586 (1999); and, A. K. Ogawa, Y. Q. Wu, D. L. McMin, J. Q. Liu, P. G. Schultz and F. E. Romesberg, J. Am. Chem. Soc., 122:3274 (2000). A 3MN:3MN self-pair can be synthesized by KF with efficiency and selectivity sufficient for biological function. *See, e.g.*, A. K. Ogawa, Y. Q. Wu, M. Berger, P. G. Schultz and F. E. Romesberg, J. Am. Chem. Soc., 122:8803 (2000).
- 10 However, both bases act as a chain terminator for further replication. A mutant DNA polymerase has been recently evolved that can be used to replicate the PICS self pair. In addition, a 7AI self pair can be replicated using a combination of KF and pol β polymerase. *See, e.g.*, E. J. L. Tae, Y. Q. Wu, G. Xia, P. G. Schultz and F. E. Romesberg, J. Am. Chem. Soc., 123:7439 (2001). A novel metallobase pair, Dipic:Py, has also been
- 15 developed, which forms a stable pair upon binding Cu(II). *See*, E. Meggers, P. L. Holland, W. B. Tolman, F. E. Romesberg and P. G. Schultz, J. Am. Chem. Soc., 122:10714 (2000). Because extended codons and unnatural codons are intrinsically orthogonal to natural codons, the methods of the present invention can take advantage of this property to generate orthogonal tRNAs for them.
- 20 **[0150] Orthogonal tRNA and Orthogonal aminoacyl-tRNA synthetase pairs**
- [0151]** An orthogonal pair is composed of an O-tRNA, e.g., a suppressor tRNA, a frameshift tRNA, or the like, and an O-RS. The O-tRNA is not acylated by endogenous synthetases and is capable of decoding a selector codon, as described above. The O-RS recognizes the O-tRNA, e.g., with an extended anticodon loop, and preferentially
- 25 aminoacylates the O-tRNA with an unnatural amino acid. Methods for generating orthogonal pairs along with compositions of orthogonal pairs are included in the present invention. The development of multiple orthogonal tRNA/synthetase pairs can allow the simultaneous incorporation of multiple unnatural amino acids using different codons into the same polypeptide/protein.
- 30 **[0152]** In the present invention, methods and related compositions relate to the generation of orthogonal pairs (O-tRNA/O-RS) that can incorporate an unnatural amino acid into a protein in vivo. For example, compositions of O-tRNAs of the present

WO 02/086075

PCT/US02/12635

invention can comprise an orthogonal aminoacyl-tRNA synthetase (O-RS). In one embodiment, the O-tRNA and the O-RS can be complementary, e.g., an orthogonal O-tRNA/O-RS pair. Examples of pairs include a mutRNATyr-mutTyrRS pair, such as a mutRNATyr-SS12TyrRS pair, a mutRNALeu-mutLeuRS pair, a mutRNAThr-mutThrRS pair, a mutRNAGlu-mutGluRS pair, or the like. In one embodiment, an orthogonal pair of the present invention comprises the desired properties of the orthogonal tRNA-aminoacyl-tRNA synthetase pair and is other than a mutRNAGln-mutGlnRS derived from *Escherichia coli*, a mutRNAAsp-mutAspRS derived from yeast or a mutRNAPheCUA-mutphenylalanineRS from yeast, where these pairs do not possess the properties of the pairs of the present invention.

[0153] The O-tRNA and the O-RS can be derived by mutation of a naturally occurring tRNA and/or RS from a variety of organisms, which are described under sources and hosts. In one embodiment, the O-tRNA and O-RS are derived from at least one organism. In another embodiment, the O-tRNA is derived by mutation of a naturally occurring or mutated naturally occurring tRNA from a first organism and the O-RS is derived by mutation of a naturally occurring or mutated naturally occurring RS from a second organism.

[0154] Methods for generating specific O-tRNA/O-RS pairs are also provided in the present invention. These methods solve the problems discussed below for the other strategies that were attempted to generate orthogonal tRNA/RS pairs. Specifically, methods of the present invention include: (a) generating a library of mutant tRNAs derived from at least one tRNA from a first organism; (b) negatively selecting the library for mutant tRNAs that are aminoacylated by an aminoacyl-tRNA synthetase (RS) from a second organism in the absence of a RS from the first organism, thereby providing a pool of mutant tRNAs; (c) selecting the pool of mutant tRNAs for members that are aminoacylated by an introduced orthogonal RS (O-RS), thereby providing at least one recombinant O-tRNA. The at least one recombinant O-tRNA recognizes a selector codon and is not efficiently recognized by the RS from the second organism and is preferentially aminoacylated by the O-RS. The method also includes: (d) generating a library of mutant RSs derived from at least one aminoacyl-tRNA synthetase (RS) from a third organism; (e) selecting the library of mutant RSs for members that preferentially aminoacylate the at least one recombinant O-tRNA in the presence of an unnatural amino acid and a natural

WO 02/086075

PCT/US02/12635

amino acid, thereby providing a pool of active mutant RSs; and, (f) negatively selecting the pool for active mutant RSs that preferentially aminoacylate the at least one recombinant O-tRNA in the absence of the unnatural amino acid, thereby providing the at least one specific O-tRNA/O-RS pair, where the at least one specific O-tRNA/O-RS pair comprises at least one recombinant O-RS that is specific for the unnatural amino acid and the at least one recombinant O-tRNA. Pairs produced by the methods of the present invention are also included.

[0155] Previously, generation of an orthogonal tRNA/synthetase pair from an existing *Escherichia coli* tRNA/synthetase pair was attempted. The method involves eliminating the tRNA's affinity toward its cognate synthetase by mutating nucleotides at the tRNA-synthetase interface while preserving its orthogonality to other synthetases and its ability to function in translation. Using the cognate wild-type synthetase as the starting template, a mutant synthetase is then evolved that uniquely recognizes the engineered orthogonal tRNA. Based on an analysis of the X-ray crystal structure of *Escherichia coli* glutamyl-tRNA synthetase (GlnRS) complexed with tRNA^{Gln}2, three sites ("knobs") in tRNA^{Gln}2 were identified which make specific contacts with GlnRS. See, e.g., D. R. Liu, T. J. Magliery and P. G. Schultz, *Chem. Biol.*, 4:685 (1997); and, D. R. Liu, T. J. Magliery, M. Pastrnak and P. G. Schultz, *Proc. Natl. Acad. Sci. U S A*, 94:10092 (1997). These sites were mutated in the tRNA, and mutant suppressor tRNAs containing all possible combinations of knobs 1, 2, and 3 were generated and tested individually by in vitro aminoacylation with GlnRS and in vitro suppression of amber mutants of chorismate mutase. A mutant tRNA (O-tRNA) bearing all three-knob mutations was shown to be orthogonal to all endogenous *Escherichia coli* synthetases and competent in translation. Next, multiple rounds of DNA shuffling together with oligonucleotide-directed mutagenesis were used to generate libraries of mutant GlnRS's. These mutant enzymes were selected for their ability to acylate the O-tRNA in vivo using *Escherichia coli* strain BT235. Only if a mutant GlnRS charges the O-tRNA with glutamine can the genomic amber codon in lacZ be suppressed, enabling BT235 cells to grow on lactose minimal media. Several mutant synthetases surviving each round of selection were purified and assayed in vitro. The ratio of wild-type (wt) tRNA^{Gln} acylation to O-tRNA acylation by mutant synthetase decreased significantly upon multiple rounds of selection. However, no mutant *Escherichia coli* GlnRS's have been evolved that charge the O-tRNA more efficiently than wild-type *Escherichia coli* tRNA^{Gln}2. The best mutant evolved after

WO 02/086075

PCT/US02/12635

seven rounds of DNA shuffling and selection acylates the O-tRNA at only one-ninth the rate of wt tRNA^{Gln}. However, these experiments failed to produce a synthetase candidate with the desired properties, e.g., a synthetase that does not acylate any wt tRNA, since misacylation of a wt tRNA with an unnatural amino acid could result in a lethal

- 5 phenotype. In addition, the mutations within the tRNA interact in complicated, non-additive ways with respect to both aminoacylation and translation. See, D. R. Liu, T. J. Magliery and P. G. Schultz, *Chem. Biol.*, 14:685 (1997). Thus, alternative methods are typically used to provide a functional pair with the desired properties.
- [0156] A second strategy for generating an orthogonal tRNA/synthetase pair
- 10 involves importing a tRNA/synthetase pair from another organism into *Escherichia coli*. The properties of the heterologous synthetase candidate include, e.g., that it does not charge any *Escherichia coli* tRNA, and the properties of the heterologous tRNA candidate include, e.g., that it is not acylated by any *Escherichia coli* synthetase. In addition, the suppressor tRNA derived from the heterologous tRNA is orthogonal to all *Escherichia*
- 15 *coli* synthetases. Schimmel et al. reported that *Escherichia coli* GlnRS (EcGlnRS) does not acylate *Saccharomyces cerevisiae* tRNA^{Gln} (EcGlnRS lacks an N-terminal RNA-binding domain possessed by *Saccharomyces cerevisiae* GlnRS (ScGlnRS)). See, E. F. Whelihan and P. Schimmel, *EMBO J.*, 16:2968 (1997). The *Saccharomyces cerevisiae* amber suppressor tRNA^{Gln} (ScRNA^{Gln}CUA) was then analyzed to determine whether it
- 20 is also not a substrate for EcGlnRS. In vitro aminoacylation assays showed this to be the case; and in vitro suppression studies show that the ScRNA^{Gln}CUA is competent in translation. See, e.g., D. R. Liu and P. G. Schultz, *Proc. Natl. Acad. Sci. U.S.A.*, 96:4780 (1999). It was further shown that ScGlnRS does not acylate any *Escherichia coli* tRNA, only the ScRNA^{Gln}CUA in vitro. The degree to which ScGlnRS is able to aminoacylate
- 25 the ScRNA^{Gln}CUA in *Escherichia coli* was also evaluated using an in vivo complementation assay. An amber nonsense mutation was introduced at a permissive site in the β -lactamase gene. Suppression of the mutation by an amber suppressor tRNA should produce full-length β -lactamase and confer ampicillin resistance to the cell. When only ScRNA^{Gln}CUA is expressed, cells exhibit an IC₅₀ of 20 μ g/mL ampicillin,
- 30 indicating virtually no acylation by endogenous *Escherichia coli* synthetases; when ScRNA^{Gln}CUA is coexpressed with ScGlnRS, cells acquire an IC₅₀ of about 500 μ g/mL ampicillin, demonstrating that ScGlnRS acylates ScRNA^{Gln}CUA efficiently in

WO 02/086075

PCT/US02/12635

Escherichia coli. See, D. R. Liu and P. G. Schultz, Proc. Natl. Acad. Sci. U S A, 96:4780 (1999). The *Saccharomyces cerevisiae* tRNA^{Gln}CUA/GlnRS is orthogonal to *Escherichia coli*.

[0157] This strategy was later applied to a tRNA^{Asp}/AspRS system.

- 5 *Saccharomyces cerevisiae* tRNA^{Asp} is known to be orthogonal to *Escherichia coli* synthetases. See, e.g., B. P. Doctor and J. A. Mudd, J. Biol. Chem., 238:3677 (1963); and, Y. Kwok and J. T. Wong, Can. J. Biochem., 58:213 (1980). It was demonstrated that an amber suppressor tRNA derived from it (SctRNA^{Asp}CUA) is also orthogonal in
- 10 *Escherichia coli* using the in vivo β -lactamase assay described above. However, the anticodon of tRNA^{Asp} is a critical recognition element of AspRS, see, e.g., R. Giege, C. Florentz, D. Kern, J. Gangloff, G. Eriani and D. Moras, Biochimie, 78:605 (1996), and mutation of the anticodon to CUA results in a loss of affinity of the suppressor for AspRS. An *Escherichia coli* AspRS E93K mutant has been shown to recognize *Escherichia coli* amber suppressor tRNA^{Asp}CUA about an order of magnitude better than wt AspRS. See,
- 15 e.g., F. Martin, 'Thesis', Universite Louis Pasteur, Strasbourg, France, 1995. It was speculated that introduction of the related mutation in *Saccharomyces cerevisiae* AspRS (E188K) might restore its affinity for SctRNA^{Asp}CUA. It was determined that the *Saccharomyces cerevisiae* AspRS(E188K) mutant does not acylate *Escherichia coli* tRNAs, but charges SctRNA^{Asp}CUA with moderate efficiency as shown by in vitro aminoacylation experiments. See, e.g., M. Pastnak, T. J. Magliery and P. G. Schultz,
- 20 Helv. Chim. Acta, 83:2277 (2000). Although the SctRNA^{Asp}CUA/ScAspRS(E188K) can serve as another orthogonal pair in *Escherichia coli*, it possesses weak activity.

- [0158] A similar approach involves the use of a heterologous synthetase as the orthogonal synthetase but a mutant initiator tRNA of the same organism or a related
- 25 organism as the orthogonal tRNA. RajBhandary and coworkers found that an amber mutant of human initiator tRNA^{fMet} is acylated by *Escherichia coli* GlnRS and acts as an amber suppressor in yeast cells only when EcGlnRS is coexpressed. See, A. K. Kowal, C. Kohrer and U. L. RajBhandary, Proc. Natl. Acad. Sci. U S A, 98:2268 (2001). This pair thus represents an orthogonal pair for use in yeast. Also, an *Escherichia coli* initiator
- 30 tRNA^{fMet} amber mutant was found that is inactive toward any *Escherichia coli* synthetases. A mutant yeast TyrRS was selected that charges this mutant tRNA, resulting in an orthogonal pair in *Escherichia coli*. See, A. K. Kowal, et al, (2001), *supra*.

WO 02/086075

PCT/US02/12635

[0159] The prokaryotic and eukaryotic tRNA^{Tyr}/TyrRS pairs have significant differences: the identity elements of prokaryotic tRNA^{Tyr} include a long variable arm in contrast to the short arm of eukaryotic tRNA^{Tyr}. In addition, eukaryotic tRNA^{Tyr} contains a C1:G72 positive recognition element whereas prokaryotic tRNA^{Tyr} has no such consensus base pair. In vitro studies have also shown that tRNA^{Tyr} of *Saccharomyces cerevisiae* and *H. sapiens* cannot be aminoacylated by bacterial synthetases, nor do their TyrRS aminoacylate bacterial tRNA. See, e.g., K. Wakasugi, C. L. Quinn, N. Tao and P. Schimmel, *EMBO J.*, 17:297 (1998); and, T. A. Kleeman, D. Wei, K. L. Simpson and E. A. First, *J. Biol. Chem.*, 272:14420 (1997). But, in spite of all these promising features for orthogonality, in vivo β -lactamase complementation assays showed that the amber suppressor tRNA^{Tyr}CUA derived from both *Saccharomyces cerevisiae* and *H. sapiens* are not orthogonal in *Escherichia coli*. See, e.g., L. Wang, T. J. Magliery, D. R. Liu and P. G. Schultz, *J. Am. Chem. Soc.*, 122:5010 (2000). The susceptibility of the suppressor tRNA to acylation by *Escherichia coli* synthetases is due to the change of one single nucleotide in the anticodon (G34 to C34).

[0160] Using the methods of the present invention, the pairs and components of pairs desired above are evolved to generate orthogonal tRNA/synthetase pairs that possess desired characteristic, e.g., that can preferentially aminoacylate an O-tRNA with an unnatural amino acid.

20 [0161] Source and Host Organisms

[0162] The orthogonal tRNA-RS pair, e.g., derived from at least a first organism or at least two organisms, which can be the same or different, can be used in a variety of organisms, e.g., a second organism. The first and the second organisms of the methods of the present invention can be the same or different. In one embodiment, the first organism is a prokaryotic organism, e.g., *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *Halobacterium*, *P. furiosus*, *P. horikoshii*, *A. pernix*, *T. thermophilus*, or the like. Alternatively, the first organism is a eukaryotic organism, e.g., plants (e.g., complex plants such as monocots, or dicots), algae, protists, fungi (e.g., yeast, etc), animals (e.g., mammals, insects, arthropods, etc.), or the like. In another embodiment, the second organism is a prokaryotic organism, *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, *Escherichia coli*, *A. fulgidus*, *Halobacterium*, *P. furiosus*, *P. horikoshii*, *A. pernix*, *T.*

WO 02/086075

PCT/US02/12635

thermophilus, or the like. Alternatively, the second organism can be a eukaryotic organism, e.g., plants, fungi, animals, or the like.

- [0163] As described above, the individual components of a pair can be derived from the same organism or different organisms. For example, tRNA can be derived from a
- 5 prokaryotic organism, e.g., an archaeobacterium, such as *Methanococcus jannaschii* and *Halobacterium NRC-1* or a eubacterium, such as *Escherichia coli*, while the synthetase can be derived from same or another prokaryotic organism, such as, *Methanococcus jannaschii*, *Archaeoglobus fulgidus*, *Methanobacterium thermoautotrophicum*, *P. furiosus*, *P. horikoshii*, *A. pernix*, *T. thermophilus*, *Halobacterium*, *Escherichia coli* or the like.
- 10 Eukaryotic sources can also be used, e.g., plants (e.g., complex plants such as monocots, or dicots), algae, protists, fungi (e.g., yeast, etc.), animals (e.g., mammals, insects, arthropods, etc.), or the like.

- [0164] Methods for selecting an orthogonal tRNA-tRNA synthetase pair for use in an in vivo translation system of a second organism are also included in the present
- 15 invention. The methods include: introducing a marker gene, a tRNA and an aminoacyl-tRNA synthetase (RS) isolated or derived from a first organism into a first set of cells from the second organism; introducing the marker gene and the tRNA into a duplicate cell set from the second organism; and, selecting for surviving cells in the first set that fail to survive in the duplicate cell set, where the first set and the duplicate cell set are grown in
- 20 the presence of a selection agent, and where the surviving cells comprise the orthogonal tRNA-tRNA synthetase pair for use in the in vivo translation system of the second organism. In one embodiment, comparing and selecting includes an in vivo complementation assay. In another embodiment, the concentration of the selection agent is varied.

- 25 [0165] For example, a tRNA/synthetase pair can be chosen based on where the identity elements, which are recognition sites of the tRNA for the synthetase, are found. For example, a tRNA/synthetase pair is chosen when the identity elements are outside of the anticodon, e.g., the tRNA^{Tyr}/TyrRS pair from the archaeobacterial *Methanococcus jannaschii*. This TyrRS is missing most of the non-conserved domain binding for the
- 30 anticodon loop of its tRNA^{Tyr}, but can discriminate tRNA with C1:G72 from that with G1:C72. Furthermore, the *Methanococcus jannaschii* TyrRS (MjTyrRS) aminoacylates *Saccharomyces cerevisiae* but not *Escherichia coli* crude tRNA. See, e.g., B. A. Steer and

WO 02/086075

PCT/US02/12635

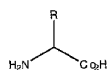
- P. Schimmel, *J. Biol. Chem.*, 274:35601 (1999). Using an in vivo complementation assay with an expression vector containing a reporter gene, e.g., β -lactamase gene, with at least one selector codon, cells expressing the *Methanococcus jannaschii* tRNA^{Tyr}CUA (*Mj* tRNA^{Tyr}CUA) alone survive to an IC₅₀ of 55 μ g/mL ampicillin; cells coexpressing *Mj* tRNA^{Tyr}CUA with its TyrRS survive to an IC₅₀ of 1220 μ g/mL ampicillin. Although *Mj* tRNA^{Tyr}CUA is less orthogonal in *Escherichia coli* than the Sc tRNA^{Gln}CUA (IC₅₀ 20 μ g/mL), the *Mj* TyrRS has higher aminoacylation activity toward its cognate amber suppressor tRNA. See, e.g., L. Wang, T. J. Magliery, D. R. Liu and P. G. Schultz, *J. Am. Chem. Soc.*, 122:5010 (2000). As a result, *Methanococcus jannaschii* /TyrRS is identified as an orthogonal pair in *Escherichia coli* and can be selected for use in an in vivo translation system.

[0166] **Unnatural Amino Acids**

- [0167] A wide variety of unnatural amino acids can be used in the methods of the invention. The unnatural amino acid can be chosen based on desired characteristics of the unnatural amino acid, e.g., function of the unnatural amino acid, such as modifying protein biological properties such as toxicity, biodistribution, or half life, structural properties, spectroscopic properties, chemical and/or photochemical properties, catalytic properties, ability to react with other molecules (either covalently or noncovalently), or the like.

- [0168] As used herein an "unnatural amino acid" refers to any amino acid, modified amino acid, or amino acid analogue other than selenocysteine and the following twenty genetically encoded α -amino acids: alanine, arginine, asparagine, aspartic acid, cysteine, glutamine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, valine. The generic structure of an α -amino acid is illustrated by Formula I:

I



An unnatural amino acid is typically any structure having Formula I wherein the R group is any substituent other than one used in the twenty natural amino acids. See, e.g., any biochemistry text such as *Biochemistry* by L. Stryer, 3rd ed. 1988, Freeman and Company, New York, for structures of the twenty natural amino acids. Note that, the unnatural

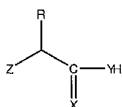
WO 02/086075

PCT/US02/12635

amino acids of the present invention may be naturally occurring compounds other than the twenty alpha-amino acids above. Because the unnatural amino acids of the invention typically differ from the natural amino acids in side chain only, the unnatural amino acids form amide bonds with other amino acids, e.g., natural or unnatural, in the same manner in which they are formed in naturally occurring proteins. However, the unnatural amino acids have side chain groups that distinguish them from the natural amino acids. For example, R in Formula I optionally comprises an alkyl-, aryl-, acyl-, keto-, azido-, hydroxyl-, hydrazine, cyano-, halo-, hydrazide, alkenyl, alkynyl, ether, thiol, seleno-, sulfonyl-, borate, boronate, phospho, phosphono, phosphine, heterocyclic, enone, imine, aldehyde, ester, thioacid, hydroxylamine, amino group, or the like or any combination thereof. Other unnatural amino acids of interest include, but are not limited to, amino acids comprising a photoactivatable cross-linker, spin-labeled amino acids, fluorescent amino acids, metal binding amino acids, metal-containing amino acids, radioactive amino acids, amino acids with novel functional groups, amino acids that covalently or noncovalently interact with other molecules, photocaged and/or photoisomerizable amino acids, amino acids comprising biotin or a biotin analogue, glycosylated amino acids such as a sugar substituted serine, other carbohydrate modified amino acids, keto containing amino acids, amino acids comprising polyethylene glycol or polyether, heavy atom substituted amino acids, chemically cleavable and/or photocleavable amino acids, amino acids with an elongated side chains as compared to natural amino acids, e.g., polyethers or long chain hydrocarbons, e.g., greater than about 5 or greater than about 10 carbons, carbon-linked sugar-containing amino acids, redox-active amino acids, amino thioacid containing amino acids, and amino acids comprising one or more toxic moiety.

[0169] In addition to unnatural amino acids that contain novel side chains, unnatural amino acids also optionally comprise modified backbone structures, e.g., as illustrated by the structures of Formula II and III:

II

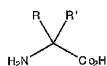


-54-

WO 02/086075

PCT/US02/12635

III



wherein Z typically comprises OH, NH₂, SH, NH-R', or S-R'; X and Y, which may be the same or different, typically comprise S or O, and R and R', which are optionally the same or different, are typically selected from the same list of constituents for the R group described above for the unnatural amino acids having Formula I as well as hydrogen. For example, unnatural amino acids of the invention optionally comprise substitutions in the amino or carboxyl group as illustrated by Formulas II and III. Unnatural amino acids of this type include, but are not limited to, α-hydroxy acids, α-thioacids α-aminothiocarboxylates, e.g., with side chains corresponding to the common twenty natural amino acids or unnatural side chains. In addition, substitutions at the α-carbon optionally include L, D, or α-α-disubstituted amino acids such as D-glutamate, D-alanine, D-methyl-O-tyrosine, aminobutyric acid, and the like. Other structural alternatives include cyclic amino acids, such as proline analogues as well as 3, 4, 6, 7, 8, and 9 membered ring proline analogues, β and γ amino acids such as substituted β-alanine and γ-amino butyric acid.

[0170] For example, many unnatural amino acids are based on natural amino acids, such as tyrosine, glutamine, phenylalanine, and the like. Tyrosine analogs include para-substituted tyrosines, ortho-substituted tyrosines, and meta substituted tyrosines, wherein the substituted tyrosine comprises an acetyl group, a benzoyl group, an amino group, a hydrazine, an hydroxyamine, a thiol group, a carboxy group, an isopropyl group, a methyl group, a C₆ - C₂₀ straight chain or branched hydrocarbon, a saturated or unsaturated hydrocarbon, an O-methyl group, a polyether group, a nitro group, or the like. In addition, multiply substituted aryl rings are also contemplated. Glutamine analogs of the invention include, but are not limited to, α-hydroxy derivatives, γ-substituted derivatives, cyclic derivatives, and amide substituted glutamine derivatives. Example phenylalanine analogs include, but are not limited to, meta-substituted phenylalanines, wherein the substituent comprises a hydroxy group, a methoxy group, a methyl group, an allyl group, an acetyl group, or the like. Specific examples of unnatural amino acids include, but are not limited

WO 02/086075

PCT/US02/12635

to, O-methyl-L-tyrosine, an L-3-(2-naphthyl)alanine, a 3-methyl-phenylalanine, an O-4-allyl-L-tyrosine, a 4-propyl-L-tyrosine, a tri-O-acetyl-GlcNAc β -serine, an L-Dopa, a fluorinated phenylalanine, an isopropyl-L-phenylalanine, a *p*-azido-L-phenylalanine, a *p*-acyl-L-phenylalanine, a *p*-benzoyl-L-phenylalanine, an L-phosphoserine, a

5 phosphoserine, a phosphotyrosine, a *p*-iodo-phenylalanine, a *p*-bromophenylalanine, a *p*-amino-L-phenylalanine, and an isopropyl-L-phenylalanine, and the like. The structures of a variety of non-limiting unnatural amino acids are provided in the figures, e.g., Figures 29, 30, and 31.

[0171] Typically, the unnatural amino acids of the invention are selected or

10 designed to provide additional characteristics unavailable in the twenty natural amino acids. For example, unnatural amino acid are optionally designed or selected to modify the biological properties of a protein, e.g., into which they are incorporated. For example, the following properties are optionally modified by inclusion of an unnatural amino acid into a protein: toxicity, biodistribution, solubility, stability, e.g., thermal, hydrolytic,

15 oxidative, resistance to enzymatic degradation, and the like, facility of purification and processing, structural properties, spectroscopic properties, chemical and/or photochemical properties, catalytic activity, redox potential, half-life, ability to react with other molecules, e.g., covalently or noncovalently, and the like.

[0172] Further details regarding unnatural amino acids are described in

20 corresponding application, "In vivo Incorporation of Unnatural Amino Acids", attorney docket number 54-000120PC/US, filed April 19, 2002, which is incorporated herein by reference.

[0173] Use of mutant tRNA and O-RS and O-tRNA/O-RS pairs

[0174] The compositions of the present invention and compositions made by the

25 methods of the present invention optionally are in a cell. The O-tRNA/O-RS pairs or individual components of the present invention can then be used in a host system's translation machinery, which results in an unnatural amino acid being incorporated into a protein. The corresponding patent application "In vivo Incorporation of Unnatural Amino Acids", attorney docket number 54-000120PC/US by Schultz, et al. describes this process

30 and is incorporated herein by reference. For example, when an O-tRNA/O-RS pair is introduced into a host, e.g., *Escherichia coli*, the pair leads to the in vivo incorporation of an unnatural amino acid, e.g., a synthetic amino acid, such as O-methyl-L-tyrosine, which

WO 02/086075

PCT/US02/12635

can be exogenously added to the growth medium, into a protein, e.g., dihydrofolate reductase or a therapeutic protein such as EPO, in response to a selector codon, e.g., an amber nonsense codon. Optionally, the compositions of the present invention can be in an in vitro translation system, or in an in vivo system(s).

5 [0175] Nucleic acid and polypeptide sequence variants

[0176] As described above and below, the invention provides for nucleic acid polynucleotide sequences and polypeptide amino acid sequences, e.g., O-tRNAs and O-RSs, and, e.g., compositions and methods comprising said sequences. Examples of said sequences, e.g., O-tRNAs and O-RSs are disclosed herein. However, one of skill in the art will appreciate that the invention is not limited to those sequences disclosed herein. One of skill will appreciate that the present invention also provides many related and unrelated sequences with the functions described herein, e.g., encoding an O-tRNA or an O-RS.

[0177] One of skill will also appreciate that many variants of the disclosed sequences are included in the invention. For example, conservative variations of the disclosed sequences that yield a functionally identical sequence are included in the invention. Variants of the nucleic acid polynucleotide sequences, wherein the variants hybridize to at least one disclosed sequence, are considered to be included in the invention. Unique subsequences of the sequences disclosed herein, as determined by, e.g., standard sequence comparison techniques, are also included in the invention.

20 [0178] Conservative variations

[0179] Owing to the degeneracy of the genetic code, "silent substitutions" (*i.e.*, substitutions in a nucleic acid sequence which do not result in an alteration in an encoded polypeptide) are an implied feature of every nucleic acid sequence which encodes an amino acid. Similarly, "conservative amino acid substitutions," in one or a few amino acids in an amino acid sequence are substituted with different amino acids with highly similar properties, are also readily identified as being highly similar to a disclosed construct. Such conservative variations of each disclosed sequence are a feature of the present invention.

[0180] "Conservative variations" of a particular nucleic acid sequence refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or, where the nucleic acid does not encode an amino acid sequence, to essentially identical

WO 02/086075

PCT/US02/12635

sequences, *see*, Table 1 below. One of skill will recognize that individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids (typically less than 5%, more typically less than 4%, 2% or 1%) in an encoded sequence are "conservatively modified variations" where the alterations result in the deletion of an amino acid, addition of an amino acid, or substitution of an amino acid with a chemically similar amino acid. Thus, "conservative variations" of a listed polypeptide sequence of the present invention include substitutions of a small percentage, typically less than 5%, more typically less than 2% or 1%, of the amino acids of the polypeptide sequence, with a conservatively selected amino acid of the same conservative substitution group. Finally, the addition of sequences which do not alter the encoded activity of a nucleic acid molecule, such as the addition of a non-functional sequence, is a conservative variation of the basic nucleic acid.

[0181] Table 1 -- Conservative Substitution Groups

1	Alanine (A)	Serine (S)	Threonine (T)	
2	Aspartic acid (D)	Glutamic acid (E)		
3	Asparagine (N)	Glutamine (Q)		
4	Arginine (R)	Lysine (K)		
5	Isoleucine (I)	Leucine (L)	Methionine (M)	Valine (V)
6	Phenylalanine (F)	Tyrosine (Y)	Tryptophan (W)	

15 [0182] **Nucleic Acid Hybridization**

[0183] Comparative hybridization can be used to identify nucleic acids of the invention, including conservative variations of nucleic acids of the invention, and this comparative hybridization method is a preferred method of distinguishing nucleic acids of the invention. In addition, target nucleic acids which hybridize to the nucleic acids represented by SEQ ID NO:1-3 or SEQ ID NO:4-34 (*see*, Table 5) under high, ultra-high and ultra-ultra high stringency conditions are a feature of the invention. Examples of such nucleic acids include those with one or a few silent or conservative nucleic acid substitutions as compared to a given nucleic acid sequence.

WO 02/086075

PCT/US02/12635

[0184] A test nucleic acid is said to specifically hybridize to a probe nucleic acid when it hybridizes at least $\frac{1}{2}$ as well to the probe as to the perfectly matched complementary target, i.e., with a signal to noise ratio at least $\frac{1}{2}$ as high as hybridization of the probe to the target under conditions in which the perfectly matched probe binds to the perfectly matched complementary target with a signal to noise ratio that is at least about 5x-10x as high as that observed for hybridization to any of the unmatched target nucleic acids.

[0185] Nucleic acids "hybridize" when they associate, typically in solution. Nucleic acids hybridize due to a variety of well characterized physico-chemical forces, such as hydrogen bonding, solvent exclusion, base stacking and the like. An extensive guide to the hybridization of nucleic acids is found in 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), as well as in Ausubel, *infra*. Hames and Higgins (1995) *Gene Probes 1* IRL Press at Oxford University Press, Oxford, England, (Hames and Higgins 1) and Hames and Higgins (1995) *Gene Probes 2* IRL Press at Oxford University Press, Oxford, England (Hames and Higgins 2) provide details on the synthesis, labeling, detection and quantification of DNA and RNA, including oligonucleotides.

[0186] An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formalin with 1 mg of heparin at 42°C, with the hybridization being carried out overnight. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (*see*, Sambrook, *infra* for a description of SSC buffer). Often the high stringency wash is preceded by a low stringency wash to remove background probe signal. An example low stringency wash is 2x SSC at 40°C for 15 minutes. In general, a signal to noise ratio of 5x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.

[0187] "Stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and northern hybridizations are sequence dependent, and are different under different environmental parameters. An extensive

WO 02/086075

PCT/US02/12635

- guide to the hybridization of nucleic acids is found in Tijssen (1993), *supra*, and in Hames and Higgins, 1 and 2. Stringent hybridization and wash conditions can easily be determined empirically for any test nucleic acid. For example, in determining highly stringent hybridization and wash conditions, the hybridization and wash conditions are
- 5 gradually increased (e.g., by increasing temperature, decreasing salt concentration, increasing detergent concentration and/or increasing the concentration of organic solvents such as formalin in the hybridization or wash), until a selected set of criteria are met. For example, the hybridization and wash conditions are gradually increased until a probe binds to a perfectly matched complementary target with a signal to noise ratio that is at least 5x
- 10 as high as that observed for hybridization of the probe to an unmatched target.
- [0188] "Very stringent" conditions are selected to be equal to the thermal melting point (T_m) for a particular probe. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the test sequence hybridizes to a perfectly matched probe. For the purposes of the present invention, generally, "highly stringent" hybridization and wash
- 15 conditions are selected to be about 5° C lower than the T_m for the specific sequence at a defined ionic strength and pH.
- [0189] "Ultra high-stringency" hybridization and wash conditions are those in which the stringency of hybridization and wash conditions are increased until the signal to noise ratio for binding of the probe to the perfectly matched complementary target nucleic
- 20 acid is at least 10x as high as that observed for hybridization to any of the unmatched target nucleic acids. A target nucleic acid which hybridizes to a probe under such conditions, with a signal to noise ratio of at least ½ that of the perfectly matched complementary target nucleic acid is said to bind to the probe under ultra-high stringency conditions.
- 25 [0190] Similarly, even higher levels of stringency can be determined by gradually increasing the hybridization and/or wash conditions of the relevant hybridization assay. For example, those in which the stringency of hybridization and wash conditions are increased until the signal to noise ratio for binding of the probe to the perfectly matched complementary target nucleic acid is at least 10X, 20X, 50X, 100X, or 500X or more as
- 30 high as that observed for hybridization to any of the unmatched target nucleic acids. A target nucleic acid which hybridizes to a probe under such conditions, with a signal to

WO 02/086075

PCT/US02/12635

noise ratio of at least ½ that of the perfectly matched complementary target nucleic acid is said to bind to the probe under ultra-ultra-high stringency conditions.

[0191] Nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are

- 5 substantially identical. This occurs, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

[0192] Unique subsequences

[0193] In one aspect, the invention provides a nucleic acid which comprises a unique subsequence in a nucleic acid selected from the sequences of O-tRNAs and O-RSs

- 10 disclosed herein, *e.g.*, SEQ ID NO:1-3 or SEQ ID NO:4-34 (*see*, Table 5). The unique subsequence is unique as compared to a nucleic acid corresponding to any previously known O-tRNA or O-RS nucleic acid sequence, *e.g.*, as found in Genbank. Alignment can be performed using, *e.g.*, BLAST set to default parameters. Any unique subsequence is useful, *e.g.*, as a probe to identify the nucleic acids of the invention.

- 15 [0194] Similarly, the invention includes a polypeptide which comprises a unique subsequence in a polypeptide selected from the sequences of O-RSs disclosed herein, *e.g.*, SEQ ID NO:35-66 (*see*, Table 5). Here, the unique subsequence is unique as compared to a polypeptide corresponding to any of known polypeptide sequence.

- [0195] The invention also provides for target nucleic acids which hybridizes under
20 stringent conditions to a unique coding oligonucleotide which encodes a unique subsequence in a polypeptide selected from the sequences of O-RSs wherein the unique subsequence is unique as compared to a polypeptide corresponding to any of the control polypeptides. Unique sequences are determined as noted above.

[0196] Sequence comparison, identity, and homology

- 25 [0197] The terms "identical" or percent "identity," in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence, as measured using one of the sequence comparison algorithms described below (or other algorithms
30 available to persons of skill) or by visual inspection.

WO 02/086075

PCT/US02/12635

[0198] The phrase "substantially identical," in the context of two nucleic acids or polypeptides (*e.g.*, DNAs encoding an O-tRNA or O-RS, or the amino acid sequence of an O-RS) refers to two or more sequences or subsequences that have at least about 60%, preferably 80%, most preferably 90-95% nucleotide or amino acid residue identity, when compared and aligned for maximum correspondence, as measured using a sequence comparison algorithm or by visual inspection. Such "substantially identical" sequences are typically considered to be "homologous," without reference to actual ancestry. Preferably, "substantial identity" exists over a region of the sequences that is at least about 50 residues in length, more preferably over a region of at least about 100 residues, and most preferably the sequences are substantially identical over at least about 150 residues, or over the full length of the two sequences to be compared.

[0199] For sequence comparison and homology determination, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

[0200] Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Natl. Acad. Sci. USA 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by visual inspection (*see generally*, Ausubel *et al.*, *infra*).

[0201] One example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul *et al.*, J. Mol. Biol. 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either

WO 02/086075

PCT/US02/12635

match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see, Henikoff & Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915).

[0202] In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, *Proc. Natl. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

[0203] **Defining Polypeptides by Immunoreactivity**

[0204] Because the polypeptides of the invention provide a variety of new polypeptide sequences (e.g., comprising unnatural amino acids in the case of proteins synthesized in the translation systems herein, or, e.g., in the case of the novel synthetases herein, novel sequences of standard amino acids), the polypeptides also provide new

WO 02/086075

PCT/US02/12635

structural features which can be recognized, e.g., in immunological assays. The generation of antisera which specifically bind the polypeptides of the invention, as well as the polypeptides which are bound by such antisera, are a feature of the invention.

[0205] For example, the invention includes synthetase proteins that specifically
 5 bind to or that are specifically immunoreactive with an antibody or antisera generated against an immunogen comprising an amino acid sequence selected from one or more SEQ ID NO:35-66 (see, Table 5). To eliminate cross-reactivity with other homologues, the antibody or antisera is subtracted with available synthetases, such as the wild-type *Methanococcus jannaschii* (*M. jannaschii*) tyrosyl synthetase (TyrRS), e.g., the "control"
 10 polypeptides. Where the the wild-type *Methanococcus jannaschii* (*M. jannaschii*) tyrosyl synthetase (TyrRS) corresponds to a nucleic acid, a polypeptide encoded by the nucleic acid is generated and used for antibody/antisera subtraction purposes.

[0206] In one typical format, the immunoassay uses a polyclonal antiserum which was raised against one or more polypeptide comprising one or more of the sequences
 15 corresponding to one or more of SEQ ID NO:35-66 (see, Table 5) or a substantial subsequence thereof (i.e., at least about 30% of the full length sequence provided). The set of potential polypeptide immunogens derived from SEQ ID NO:35-66 (see, Table 5) are collectively referred to below as "the immunogenic polypeptides." The resulting antisera is optionally selected to have low cross-reactivity against the control synthetase
 20 homologues and any such cross-reactivity is removed, e.g., by immunoabsorption, with one or more of the control synthetase homologues, prior to use of the polyclonal antiserum in the immunoassay.

[0207] In order to produce antisera for use in an immunoassay, one or more of the immunogenic polypeptides is produced and purified as described herein. For example,
 25 recombinant protein can be produced in a recombinant cell. An inbred strain of mice (used in this assay because results are more reproducible due to the virtual genetic identity of the mice) is immunized with the immunogenic protein(s) in combination with a standard adjuvant, such as Freund's adjuvant, and a standard mouse immunization protocol (see, e.g., Harlow and Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor
 30 Publications, New York, for a standard description of antibody generation, immunoassay formats and conditions that can be used to determine specific immunoreactivity. Additional references and discussion of antibodies is also found herein and can be applied

WO 02/086075

PCT/US02/12635

here to defining polypeptides by immunoreactivity). Alternatively, one or more synthetic or recombinant polypeptide derived from the sequences disclosed herein is conjugated to a carrier protein and used as an immunogen.

- [0208] Polyclonal sera are collected and titered against the immunogenic
- 5 polypeptide in an immunoassay, for example, a solid phase immunoassay with one or more of the immunogenic proteins immobilized on a solid support. Polyclonal antisera with a titer of 10^6 or greater are selected, pooled and subtracted with the control synthetase polypeptides to produce subtracted pooled titered polyclonal antisera.
- [0209] The subtracted pooled titered polyclonal antisera are tested for cross
- 10 reactivity against the control homologues in a comparative immunoassay. In this comparative assay, discriminatory binding conditions are determined for the subtracted titered polyclonal antisera which result in at least about a 5-10 fold higher signal to noise ratio for binding of the titered polyclonal antisera to the immunogenic synthetase as compared to binding to the control synthetase homologues. That is, the stringency of the
- 15 binding reaction is adjusted by the addition of non-specific competitors such as albumin or non-fat dry milk, and/or by adjusting salt conditions, temperature, and/or the like. These binding conditions are used in subsequent assays for determining whether a test polypeptide (a polypeptide being compared to the immunogenic polypeptides and/or the control polypeptides) is specifically bound by the pooled subtracted polyclonal antisera.
- 20 In particular, test polypeptides which show at least a 2-5x higher signal to noise ratio than the control synthetase homologues under discriminatory binding conditions, and at least about a $\frac{1}{2}$ signal to noise ratio as compared to the immunogenic polypeptide(s), shares substantial structural similarity with the immunogenic polypeptide as compared to known synthetases, and is, therefore a polypeptide of the invention.
- 25 [0210] In another example, immunoassays in the competitive binding format are used for detection of a test polypeptide. For example, as noted, cross-reacting antibodies are removed from the pooled antisera mixture by immunoabsorption with the control polypeptides. The immunogenic polypeptide(s) are then immobilized to a solid support which is exposed to the subtracted pooled antisera. Test proteins are added to the assay to
- 30 compete for binding to the pooled subtracted antisera. The ability of the test protein(s) to compete for binding to the pooled subtracted antisera as compared to the immobilized protein(s) is compared to the ability of the immunogenic polypeptide(s) added to the assay

WO 02/086075

PCT/US02/12635

to compete for binding (the immunogenic polypeptides compete effectively with the immobilized immunogenic polypeptides for binding to the pooled antisera). The percent cross-reactivity for the test proteins is calculated, using standard calculations.

[0211] In a parallel assay, the ability of the control proteins to compete for binding to the pooled subtracted antisera is optionally determined as compared to the ability of the immunogenic polypeptide(s) to compete for binding to the antisera. Again, the percent cross-reactivity for the control polypeptides is calculated, using standard calculations. Where the percent cross-reactivity is at least 5-10x as high for the test polypeptides as compared to the control polypeptides and or where the binding of the test polypeptides is approximately in the range of the binding of the immunogenic polypeptides, the test polypeptides are said to specifically bind the pooled subtracted antisera.

[0212] In general, the immunoabsorbed and pooled antisera can be used in a competitive binding immunoassay as described herein to compare any test polypeptide to the immunogenic and/or control polypeptide(s). In order to make this comparison, the immunogenic, test and control polypeptides are each assayed at a wide range of concentrations and the amount of each polypeptide required to inhibit 50% of the binding of the subtracted antisera to, e.g., an immobilized control, test or immunogenic protein is determined using standard techniques. If the amount of the test polypeptide required for binding in the competitive assay is less than twice the amount of the immunogenic polypeptide that is required, then the test polypeptide is said to specifically bind to an antibody generated to the immunogenic protein, provided the amount is at least about 5-10x as high as for the control polypeptide.

[0213] As an additional determination of specificity, the pooled antisera is optionally fully immunosorbed with the immunogenic polypeptide(s) (rather than the control polypeptides) until little or no binding of the resulting immunogenic polypeptide subtracted pooled antisera to the immunogenic polypeptide(s) used in the immunosorption is detectable. This fully immunosorbed antisera is then tested for reactivity with the test polypeptide. If little or no reactivity is observed (i.e., no more than 2x the signal to noise ratio observed for binding of the fully immunosorbed antisera to the immunogenic polypeptide), then the test polypeptide is specifically bound by the antisera elicited by the immunogenic protein.

[0214] General Techniques

WO 02/086075

PCT/US02/12635

[0215] General texts which describe molecular biological techniques, which are applicable to the present invention, such as cloning, mutation, cell culture and the like, include Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook et al., 5 Molecular Cloning - A Laboratory Manual (3rd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2000 ("Sambrook") and 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., (supplemented through 10 2002) ("Ausubel"). These texts describe mutagenesis, the use of vectors, promoters and many other relevant topics related to, e.g., the generation of orthogonal tRNA, orthogonal synthetases, and pairs thereof.

[0216] Various types of mutagenesis are used in the present invention, e.g., to produce novel synthetases or tRNAs. They include but are not limited to site-directed, random point mutagenesis, homologous recombination (DNA shuffling), mutagenesis 15 using uracil containing templates, oligonucleotide-directed mutagenesis, phosphorothioate-modified DNA mutagenesis, mutagenesis using gapped duplex DNA or the like. Additional suitable methods include point mismatch repair, mutagenesis using repair-deficient host strains, restriction-selection and restriction-purification, deletion mutagenesis, mutagenesis by total gene synthesis, double-strand break repair, and the like. 20 Mutagenesis, e.g., involving chimeric constructs, are also included in the present invention. In one embodiment, mutagenesis can be guided by known information of the naturally occurring molecule or altered or mutated naturally occurring molecule, e.g., sequence, sequence comparisons, physical properties, crystal structure or the like.

[0217] The above texts and examples found herein describe these procedures as well as the following publications and references cited within: Sieber, et al., *Nature Biotechnology*, 19:456-460 (2001); 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); I. A. Lorimer, I. Pastan, Nucleic Acids Res. 23, 3067-8 (1995); W. P. C. Stemmer, 30 Nature 370, 389-91 (1994); Arnold, *Protein engineering for unusual environments*, Current Opinion in Biotechnology 4:450-455 (1993); Bass et al., *Mutant Trp repressors with new DNA-binding specificities*, Science 242:240-245 (1988); Fritz et al.,

WO 02/086075

PCT/US02/12635

- 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., Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations, Nucl. Acids Res. 16: 7207 (1988);
- 5 Sakamar and Khorana, Total synthesis and expression of a gene for the α -subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin), Nucl. Acids Res. 14: 6361-6372 (1988); Sayers et al., Y-T 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
- 10 endonucleases in the presence of ethidium bromide, (1988) Nucl. Acids Res. 16: 803-814; Carter, Improved oligonucleotide-directed mutagenesis using M13 vectors, Methods in Enzymol. 154: 382-403 (1987); Kramer & Fritz Oligonucleotide-directed construction of mutations via gapped duplex DNA, Methods in Enzymol. 154:350-367 (1987); Kunkel, The efficiency of oligonucleotide directed mutagenesis, in Nucleic Acids & Molecular
- 15 Biology (Eckstein, F. and Lilley, D.M.J. eds., Springer Verlag, Berlin)) (1987); Kunkel et al., Rapid and efficient site-specific mutagenesis without phenotypic selection, Methods in Enzymol. 154, 367-382 (1987); 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); Carter, Site-directed mutagenesis, Biochem. J.
- 20 237:1-7 (1986); Eghtedarzadeh & Henikoff, Use of oligonucleotides to generate large deletions, Nucl. Acids Res. 14: 5115 (1986); 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); Nakamaye & Eckstein, Inhibition of restriction endonuclease *Nci* I cleavage by phosphorothioate groups and its
- 25 application to oligonucleotide-directed mutagenesis, Nucl. Acids Res. 14: 9679-9698 (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); Botstein & Shortle, Strategies and applications of in vitro mutagenesis, Science 229:1193-1201(1985); Carter et al., Improved oligonucleotide site-directed mutagenesis using M13 vectors, Nucl. Acids
- 30 Res. 13: 4431-4443 (1985); Grundström et al., Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis, Nucl. Acids Res. 13: 3305-3316 (1985); Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection, Proc. Natl. Acad. Sci. USA 82:488-492 (1985); Smith, In vitro mutagenesis, Ann. Rev. Genet.

WO 02/086075

PCT/US02/12635

- 19:423-462(1985); 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: 8765-8787 (1985); Wells et al.,
- 5 *Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites*, Gene 34:315-323 (1985); Kramer et al., *The gapped duplex DNA approach to oligonucleotide-directed mutation construction*, Nucl. Acids Res. 12: 9441-9456 (1984); Kramer et al., *Point Mismatch Repair*, Cell 38:879-887 (1984); Nambiar et al., *Total synthesis and cloning of a gene coding for the ribonuclease S protein*, Science 223: 1299-
- 10 1301 (1984); Zoller & Smith, *Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors*, Methods in Enzymol. 100:468-500 (1983); and 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). Additional details on many of the above methods can be found in
- 15 Methods in Enzymology Volume 154, which also describes useful controls for troubleshooting problems with various mutagenesis methods.

[0218] Oligonucleotides, e.g., for use in mutagenesis of the present invention, e.g., mutating libraries of synthetases, or altering tRNAs, are typically synthesized chemically according to the solid phase phosphoramidite triester method described by Beaucage and

20 Caruthers, Tetrahedron Letts. 22(20):1859-1862, (1981) e.g., using an automated synthesizer, as described in Needham-VanDevanter et al., Nucleic Acids Res., 12:6159-6168 (1984).

[0219] In addition, essentially any nucleic acid can be custom or standard ordered from any of a variety of commercial sources, such as The Midland Certified Reagent

25 Company (mcrc@oligos.com), The Great American Gene Company (www.genco.com), ExpressGen Inc. (www.expressgen.com), Operon Technologies Inc. (Alameda, CA) and many others.

[0220] The present invention also relates to host cells and organisms for the in vivo incorporation of an unnatural amino acid via orthogonal tRNA/RS pairs. Host cells

30 are genetically engineered (e.g., transformed, transduced or transfected) with the vectors of this invention, which can be, for example, a cloning vector or an expression vector. The vector can be, for example, in the form of a plasmid, a bacterium, a virus, a naked

WO 02/086075

PCT/US02/12635

- polynucleotide, or a conjugated polynucleotide. The vectors are introduced into cells and/or microorganisms by standard methods including electroporation (From et al., Proc. Natl. Acad. Sci. USA 82, 5824 (1985), infection by viral vectors, high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface (Klein et al., Nature 327, 70-73 (1987)). Berger, Sambrook, and Ausubel provide a variety of appropriate transformation methods.
- [0221] The engineered host cells can be cultured in conventional nutrient media modified as appropriate for such activities as, for example, screening steps, activating promoters or selecting transformants. These cells can optionally be cultured into transgenic organisms.
- [0222] Other useful references, e.g. for cell isolation and culture (e.g., for subsequent nucleic acid isolation) include Freshney (1994) Culture of Animal Cells, a Manual of Basic Technique, third edition, Wiley- Liss, New York and the references cited therein; Payne et al. (1992) Plant Cell and Tissue Culture in Liquid Systems John Wiley & Sons, Inc. New York, NY; Gamborg and Phillips (eds.) (1995) Plant Cell, Tissue and Organ Culture: Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) and Atlas and Parks (eds.) The Handbook of Microbiological Media (1993) CRC Press, Boca Raton, FL.
- [0223] Several well-known methods of introducing target nucleic acids into bacterial cells are available, any of which can be used in the present invention. These include: fusion of the recipient cells with bacterial protoplasts containing the DNA, electroporation, projectile bombardment, and infection with viral vectors, etc. Bacterial cells can be used to amplify the number of plasmids containing DNA constructs of this invention. The bacteria are grown to log phase and the plasmids within the bacteria can be isolated by a variety of methods known in the art (*see*, for instance, Sambrook). In addition, a plethora of kits are commercially available for the purification of plasmids from bacteria, (*see*, e.g., EasyPrep™, FlexiPrep™, both from Pharmacia Biotech; StrataClean™, from Stratagene; and, QIAprep™ from Qiagen). The isolated and purified plasmids are then further manipulated to produce other plasmids, used to transfect cells or incorporated into related vectors to infect organisms. Typical vectors contain transcription and translation terminators, transcription and translation initiation sequences, and promoters useful for regulation of the expression of the particular target nucleic acid. The

WO 02/086075

PCT/US02/12635

vectors optionally comprise generic expression cassettes containing at least one independent terminator sequence, sequences permitting replication of the cassette in eukaryotes, or prokaryotes, or both, (e.g., shuttle vectors) and selection markers for both prokaryotes and eukaryotic systems. Vectors are suitable for replication and integration in prokaryotes, eukaryotes, or preferably both. See, Gilman & Smith, Gene 8:81 (1979); Roberts, et al., Nature, 328:731 (1987); Schneider, B., et al., Protein Expr. Purif. 6435:10 (1995); Ausubel, Sambrook, Berger (*all supra*). A catalogue of Bacteria and Bacteriophages useful for cloning is provided, e.g., by the ATCC, e.g., The ATCC Catalogue of Bacteria and Bacteriophage (1992) Gherna et al. (eds.) published by the ATCC. Additional basic procedures for sequencing, cloning and other aspects of molecular biology and underlying theoretical considerations are also found in Watson et al. (1992) Recombinant DNA Second Edition Scientific American Books, NY.

EXAMPLES

[0224] The following examples are offered to illustrate, but not to limit the claimed invention.

[0225] Example 1- Improvement of Orthogonality of a tRNA from *Methanococcus jannaschii*

[0226] Because of the complex nature of tRNA-synthetase interactions that are required to achieve a high degree of fidelity in protein translation, the rational design of orthogonal tRNA-synthetase pairs is difficult. This example describes methods that exploit the poor cross recognition of some interspecies tRNA-synthetase pairs, coupled with subsequent in vivo evolution of tRNAs with enhanced orthogonality. See, also, L. Wang and P. G. Schultz, Chem. Biol., 8:883 (2001). Specifically, a library of amber suppressor tRNAs derived from *Methanococcus jannaschii* tRNA^{Tyr} was generated. tRNA^{Tyr}CUAs that are substrates for endogenous *Escherichia coli* aminoacyl-tRNA synthetases were deleted from the pool by negative selection based on suppression of amber nonsense mutations in the barnase gene. The remaining tRNA^{Tyr}CUAs were then selected for their ability to suppress amber nonsense codons in the β -lactamase gene in the presence of the cognate *Methanococcus jannaschii* tyrosyl-tRNA synthetase (TyrRS). Four mutant suppressor tRNAs were selected that are poorer substrates for *Escherichia coli* synthetases than *Methanococcus jannaschii* tRNA^{Tyr}CUA, but still can be charged

WO 02/086075

PCT/US02/12635

efficiently by *Methanococcus jannaschii* TyrRS. The mutant suppressor tRNA^{Tyr}CUA together with the *Methanococcus jannaschii* TyrRS provide a useful orthogonal tRNA-synthetase pair for the in vivo incorporation of unnatural amino acids into proteins.

[0227] The tRNA^{Tyr} of *Methanococcus jannaschii*, an archaeobacterium, has
5 different identity elements from those of *Escherichia coli* tRNA^{Tyr}. In particular, the
Escherichia coli tRNA^{Tyr} has a G1C72 pair in the acceptor stem while the
Methanococcus jannaschii tRNA^{Tyr} has a C1G72 pair. An amber suppressor tRNA
derived from *Methanococcus jannaschii* tRNA^{Tyr} was shown not to be efficiently
aminoacylated by the *Escherichia coli* synthetases, but functions efficiently in protein
10 translation in *Escherichia coli*. See, e.g., L. Wang, T.J. Magliery, D. R. Liu, P. G. Schultz,
A new functional suppressor tRNA/aminoacyl-tRNA synthetase pair for the in vivo
incorporation of unnatural amino acids into proteins, J. Am. Chem. Soc., 122:5010-5011
(2000). In addition, the *Methanococcus jannaschii* TyrRS, which has only a minimalist
anticodon-loop-binding domain, does not aminoacylate *Escherichia coli* tRNAs, but still
15 efficiently aminoacylates its own suppressor tRNA^{Tyr}CUA. See, e.g., B. A. Steer, P.
Schimmel, *Major anticodon-binding region missing from an archaeobacterial tRNA*
synthetase, J. Biol. Chem., 274 (1999) 35601-35606; and, Wang et al., (2000), *supra*.

[0228] To test the orthogonality of this suppressor tRNA in *Escherichia coli*, an
amber codon was introduced at a permissive site (Ala184) in the β -lactamase gene. See,
20 e.g., D. R. Liu, P. G. Schultz, *Progress toward the evolution of an organism with an*
expanded genetic code, Proc. Natl. Acad. Sci. USA 96:4780-4785 (1999). Those tRNAs
that can be charged by *Escherichia coli* synthetases will suppress the amber codon and
allow cells to live in the presence of ampicillin. The *Methanococcus jannaschii*
tRNA^{Tyr}CUA suppresses the amber codon in the β -lactamase gene with an IC₅₀ value of
25 56 μ g/ml ampicillin. See Wang et al., (2000), *supra*. In contrast, the orthogonal
tRNA^{Gln}CUA derived from *Saccharomyces cerevisiae* tRNA^{Gln}2 has an IC₅₀ of 21
 μ g/ml ampicillin when tested in the same assay. See Liu & Schultz, (1999), *supra*. The
IC₅₀ for *Escherichia coli* in the absence of any suppressor tRNA is 10 μ g/ml ampicillin.
This result shows that the *Methanococcus jannaschii* tRNA^{Tyr}CUA is a better substrate
30 for *Escherichia coli* synthetases than the tRNA^{Gln}CUA. Consequently, if the
Methanococcus jannaschii tRNA^{Tyr}CUA is used in vivo to deliver unnatural amino acids

WO 02/086075

PCT/US02/12635

into proteins in *Escherichia coli*, it can also be mischarged with natural amino acids by *Escherichia coli* synthetases, leading to heterogeneous amino acid incorporation.

[0229] The improvement of the orthogonality of the *Methanococcus jannaschii* tRNA^{TyrCUA} was accomplished by the introduction of 'negative recognition

- 5 determinants' to prevent recognition by endogenous *Escherichia coli* synthetases. These mutations should not strongly interfere with the tRNA's interaction with its cognate *Methanococcus jannaschii* TyrRS or the ribosome. Since *Methanococcus jannaschii* TyrRS lacks most of the anticodon-binding domain, see, e.g., B. A. Steer, P. Schimmel, *Major anticodon-binding region missing from an archaeobacterial tRNA synthetase*, 1. Biol. Chem. 274:35601-35606 (1999), mutations introduced at the anticodon loop of the tRNA are expected to have a minimal effect on TyrRS recognition. An anticodon-loop library with four randomized nucleotides was constructed. See Figure 9. Given the various combinations and locations of identity elements for various *Escherichia coli* tRNAs, mutations at additional positions can increase the likelihood of finding a mutant tRNA with the desired properties. Thus, a second library containing mutations at nonconserved positions in all of the tRNA loops (all-loop library) was also constructed. See Figure 9. Conserved nucleotides were not randomized so as to maintain the tertiary interactions that stabilize the 'L'-shaped structure of the tRNA. See, e.g., G. Dirheimer, G. Keith, P. Dumas, E. Westhof, *Primary, secondary, and tertiary structures of tRNAs*, in: D. Söll, U. L. RajBhandary (eds.), *tRNA Structure, Biosynthesis, and Function*, ASM Press, Washington, DC, 1995, pp. 93-126; and, R. Giegé, M. Sissler, C. Florentz, *Universal rules and idiosyncratic features in tRNA identity*, Nucleic Acids Res. 26:5017-5035 (1998). Stem nucleotides were also not mutated since substitution of one such nucleotide requires a compensatory mutation. The 11 nucleotides (C16, C17, U17a, U20, C32, G37, A38, U45, U47, A59, and U60) were randomized. See, Figure 9. The theoretical size of this library is about 4.19×10^6 , and a library with a size of about 1.93×10^8 colony-forming units was constructed to ensure complete coverage of the mutant library.

- [0230] The methods used an *Escherichia coli* strain, e.g., DH10B, which was obtained from Gibco/BRL. Suppressor tRNA expression plasmids were derived from a plasmid, e.g., pAC123. See, e.g., D. R. Liu, T. J. Magliery, M. Pastnak, P. G. Schultz, *Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of*

WO 02/086075

PCT/US02/12635

unnatural amino acids into proteins in vivo, *Proc. Natl. Acad. Sci. USA* 94:10091-10097 (1997). Plasmids for negative selections were derived from plasmids, e.g., pBATS, pYsupA38B2 and pYsupA38B3 as described below. See, e.g., K. Gabriel, W.H. McClain, *A set of plasmids constitutively producing different RNA levels in Escherichia coli*, *J. Mol. Biol.* 290 (1999) 385-389; and, Liu & Shultz, (1999), *supra*.

[0231] To select for a member of the *Methanococcus jannaschii* tRNA library with enhanced orthogonality, a combination of negative and positive selections in the absence and presence of the cognate synthetase was used. See Figure 8. In the negative selection, selector codon(s), e.g., amber nonsense, are introduced in a negative marker gene, e.g., a toxic gene, at e.g., a nonessential position. When a member of the mutated, e.g., suppressor, tRNA library is aminoacylated by endogenous (e.g., *Escherichia coli*) synthetases (i.e. it is not orthogonal to the *Escherichia coli* synthetases), the selector codon is suppressed and the toxic gene product produced leads to cell death. Only cells harboring orthogonal tRNAs or nonfunctional tRNAs can survive. All survivors are then subjected to a positive selection in which a selector codon, e.g., an amber codon, is placed in a positive selection marker, e.g., drug resistance gene at, e.g., a nonessential position. tRNAs are then selected for their ability to be aminoacylated by the coexpressed cognate synthetase and to insert an amino acid in response to this amber codon. Cells harboring nonfunctional tRNAs, or tRNAs that cannot be recognized by the synthetase of interest will be sensitive to antibiotic. Therefore, only tRNAs that (1) are not substrates for endogenous *Escherichia coli* synthetases; (2) can be aminoacylated by the synthetase of interest; (3) are functional in translation will survive both selections.

[0232] A negative selection was chosen that takes advantage of the toxicity of barnase when produced in *Escherichia coli* in the absence of its natural inhibitor barstar. See, e.g., R.W. Hartley, Barnase and barstar. *Expression of its cloned inhibitor permits expression of a cloned ribonuclease*, *J. Mol. Biol.* 202:913-915 (1988). Amber codons were introduced at nonessential positions in the barnase gene based on analysis of the three-dimensional structure of barnase. See, e.g., Liu & Schultz, (1999), *supra*. Because of barnase's extreme autotoxicity, a low copy number pSC101 origin was placed in the plasmid expressing barnase. In addition, different numbers of amber codons were tested to modulate the stringency of the selection. Plasmid pSCB2 was used to express a barnase

WO 02/086075

PCT/US02/12635

mutant with two amber stop codons at Gln2 and Asp44; plasmid pSCB3 contained an additional amber stop codon at Gly65.

[0233] For negative selection, a PCR fragment containing the β -lactamase gene and the pSC101 origin was generated from pBATS using the following oligonucleotides:

- 5 LW115, 5'-ATGCATGCTGCATTAAATGAATCGGCCAACG-3'; LW116, 5'-
TCCCCGCGGAGGTGGCACTTTTCGGGG-3'. DNA encoding barnase containing two
(residues Gln2 and Asp44) or three (residues Gln2, Asp44 and Gly65) amber codons were
obtained from pYsupA38B2 and pYsupA38B3, respectively, by digestion with SacII and
SphI. Ligation of the above fragments afforded plasmids pSCB2 and pSCB3. The
10 expression of barnase was under arabinose induction. Genes encoding different
suppressor tRNAs for in vivo expression were constructed from two overlapping synthetic
oligonucleotides (Operon, Alameda, CA, USA) by Klenow extension and inserted
between the EcoRI and PstI sites of pAC123 to generate pAC-YYG1 and pAC-JY,
respectively, placing transcription under control of the lpp promoter and the rmC
15 terminator. pAC-Cm is the control plasmid without any tRNA. To optimize the negative
selection conditions, competent DH10B cells harboring pSCB2 or pSCB3 were
transformed by electroporation with pAC-Cm, pAC-YYG1, and pAC-JY, separately.
Single colonies were picked and grown in 2xYT with chloramphenicol (Cm, 34 μ g/ml)
and ampicillin (Amp, 100 μ g/ml). Cell cultures grown overnight were washed twice with
20 minimal media containing 1% glycerol and 0.3 mM leucine (GMML), and resuspended in
GMML with Cm and Amp to an OD600 of 0.1. After recovering at 30°C for 10 min, into
one culture (set 1) was added 20 mM of arabinose to induce the expression of barnase; no
arabinose was added to the second culture (set 2). At different time points, a small amount
of cell culture was diluted and plated on 2xYT agar with Cm and Amp to measure cell
25 density. For negative selections of the suppressor tRNA libraries, the pAC plasmids
containing the library were transformed into DH10B cells harboring pSCB2. Cells were
quenched by addition of SOC medium and recovered at 30°C for 1 hour, then were
washed with phosphate buffer and GMML, and cultured in 1 l GMML. After recovering
at 30°C for 30 min, Cm, Amp, and 20 mM arabinose were added. After 36 hours, cells
30 were pelleted and pAC plasmids were isolated and purified by agarose gel electrophoresis.

[0234] To optimize the selection conditions, two suppressor tRNAs were used that are known to be poorly recognized by the *Escherichia coli* synthetases. A mutant

WO 02/086075

PCT/US02/12635

suppressor tRNA^{Tyr} derived from *Saccharomyces cerevisiae* (sc-tRNA^{Tyr}CUA, expressed in pAC-YYG1) suppresses the amber codon (Ala184TAG) in the β -lactamase gene, affording an IC₅₀ value of 12 μ g/ml ampicillin for *Escherichia coli* cells; and the suppressor tRNA^{Tyr} derived from *Methanococcus jannaschii* (mj-tRNA^{Tyr}CUA, expressed in pAC-JY) affords an IC₅₀ of 56 μ g/ml ampicillin for host cells. *See, e.g.,* Wang et al, (2000), *supra*. For comparison, the suppressor tRNA^{Gln}CUA derived from *Saccharomyces cerevisiae* tRNA^{Gln}2 has an IC₅₀ of 21 μ g/ml ampicillin when tested in the same assay, and has been demonstrated to be orthogonal to *Escherichia coli* synthetases in vitro and in vivo. *See, e.g.,* Liu & Schultz, (1999), *supra*. Therefore, a negative selection that eliminates cells expressing mj-tRNA^{Tyr}CUA, but allows the growth of cells expressing sc-tRNA^{Tyr}CUA deletes non-orthogonal suppressor tRNAs. Cells were grown in liquid minimal media containing 1% glycerol and 0.3 mM leucine (GMML) with appropriate antibiotics to maintain plasmid pSCB2 and the pAC plasmid. Arabinose was added to one set of cells (set 1) to induce the expression of the barnase, while in set 2 no arabinose was added. The fraction of cells surviving the selection was determined by the ratio of cell densities in set 1 relative to set 2. *See Figure 11*: cells harboring the control plasmid pAC-Cm (without suppressor tRNA) and plasmid pAC-YYG1 survived, while cells harboring plasmid pAC-JY largely died. When plasmid pSCB3 was used, cells harboring plasmid pAC-JY started to grow in 24 hours. Therefore, the negative selection was carried out using pSCB2, which encodes the barnase gene containing two amber codons under the above conditions for the library selection.

[0235] For positive selection, a plasmid, e.g., pBLAM-JYRS was constructed by inserting the *Methanococcus jannaschii* TyrRS gene from pBSA50 between NdeI and PstI sites of pBLAM-YQRS using oligonucleotides LW104, 5'-GGAATTCATTAGGACGAATTGAAATG-3'; and LW105, 5'-AAACTGCAGTTATAATCTCTTCTAATTGGCTC-3'. *See, e.g.,* Steer, et al., (1999), *supra*; and Liu & Schultz, (1999), *supra*. To optimize the positive selection conditions, competent DH10B cells harboring pBLAM-JYRS were transformed with pAC-Cm, pAC-YYG1, and pAC-JY, separately. Single colonies were picked and grown in 2 $\times$ YT with Cm and tetracycline (Tet, 40 μ g/ml). In liquid selections, overnight cell cultures were diluted into 2 $\times$ YT with Cm and Tet at a starting OD₆₀₀ of 0.1. Various concentrations of Amp were added, and cell growth was monitored by OD₆₀₀. In plate selections, approximately 103 to 105 cells were plated on two sets of 2 $\times$ YT agar plates containing

WO 02/086075

PCT/US02/12635

Cm and Tet, one set of which contained 500 μ g/ml Amp. For selections involving the mutant tRNA library, pAC plasmids isolated from the cells from the negative selection were transformed into competent DH10B cells harboring pBLAM-JYRS. Cells were recovered at 37°C for 45 minutes, and approximately 105 cells were plated onto each 2xYT agar plate containing Cm, Tet and 500 μ g/ml of Amp. After 24 hours, colonies were picked and re-grown in 6 ml 2xYT containing Cm, Tet and 200 μ g/ml of Amp. DNA was isolated and pAC plasmid was purified by agarose gel electrophoresis.

[0236] The positive selection is based on suppression of an amber stop codon introduced at position Ala184 in the TEM-1 β -lactamase gene. Plasmid pBLAM-JYRS encodes the gene for the *Methanococcus jannaschii* tyrosyl-tRNA synthetase and a β -lactamase with an amber mutation at Ala184. pAC plasmids isolated from cells surviving the negative selection were cotransformed with pBLAM-JYRS into *Escherichia coli* DH10B cells. Cells harboring nonfunctional tRNAs or tRNAs that are poor substrates for the *Methanococcus jannaschii* synthetase die; those with tRNAs that can be charged by the synthetase survive. To test the feasibility of the positive selection, two model suppressor tRNAs were tested in the presence of *Methanococcus jannaschii* TyrRS. The sc-tRNA^{Tyr}CUA has a G1:C72 base pair and is not charged efficiently by *Methanococcus jannaschii* TyrRS. When they were coexpressed in cells with the Ala184amber β -lactamase mutant, cells survived to an IC₅₀ of 18 μ g/ml ampicillin. In contrast, cells containing the *Methanococcus jannaschii* tRNA^{Tyr}CUA and the cognate TyrRS survive to an IC₅₀ of 1220 μ g/ml ampicillin. See, e.g., Wang, et al., (2000), *supra*. The model positive selection was first tried in liquid 2xYT medium. The growth of cells harboring pBLAM-JYRS and different pAC plasmids in liquid 2xYT medium with various concentrations of ampicillin are shown in **Figure 12, Panel A**. Cells transformed with the mj-tRNA^{Tyr}CUA grew at a faster rate and at higher concentrations of ampicillin. If cells were grown longer than 24 hours, cells transformed with either pAC-Cm or pAC-YYG1 also grew to saturation. Therefore, the positive selection was carried out on plates with initial cell densities between 103 and 105 per plate. See **Figure 12, Panel B**. The survival ratio (number of colonies on plates with ampicillin relative to plates without ampicillin) did not change significantly with different initial cell densities, and was stable over the growth time. The positive selection on ampicillin plates resulted in preferential growth of

WO 02/086075

PCT/US02/12635

cells with mj-tRNA^{Tyr}CUA expressed. Therefore, for the library selection the positive selection was carried out on plates instead of in liquid medium.

[0237] The library of mutant tRNAs was generated by using the sequences of the two overlapping oligonucleotides used to construct the anticodon-loop library are (the tRNA sequence underlined): LW125, 5'-GGAATTC-3'; LW126, 5'-AAAAGTGCAG-3' (where N is equimolar of A, C, T or G). The sequences of oligonucleotides for the all-loop library are: LW145, 5'-GGAATTC-3' and LW146, 5'-AAAAGTGCAG-3'. These genes were inserted into pAC123 similarly as described above to afford the tRNA libraries.

10 [0238] The negative and positive selections were carried out in tandem as described above on both the anticodon-loop and all-loop libraries. The selected suppressor tRNAs were isolated and retransformed into *Escherichia coli* DH10B harboring pBLAM to test the tRNA's orthogonality to *Escherichia coli* synthetases. The tRNAs were then retransformed into *Escherichia coli* harboring pBLAM-JYRS to test how efficiently the tRNA was charged by *Methanococcus jannaschii* TyrRS. Sequencing of the clones resulting from one round of negative and positive selection of anticodon-loop library revealed that three independent tRNAs were isolated. See Figure 13. When cotransformed with pBLAM, all had lower IC₅₀ values than the parent *Methanococcus jannaschii* tRNA^{Tyr}CUA, indicating they are poorer substrates for *Escherichia coli* synthetases.

[0239] Mutant AA2 also had very high affinity for *Methanococcus jannaschii* TyrRS. Although this mutant tRNA could be stably maintained in *Escherichia coli*, it slowed the growth rate of cells for unknown reasons. This effect likely led to the emergence of mutants AA3 and AA4, which both had a mutation outside of the randomization region. Cells harboring AA3 or AA4 grew normally. Nevertheless, AA3 and AA4 were relatively poor substrates for the *Methanococcus jannaschii* TyrRS.

[0240] Four independent tRNAs were selected from two rounds of negative and positive selections using the all-loop library. See Figure 13. All were poorer substrates for the *Escherichia coli* synthetase than the parent *Methanococcus jannaschii* tRNA^{Tyr}CUA, yet were still efficiently charged by the *Methanococcus jannaschii* TyrRS as shown by the in vivo β -lactamase assay. See Table 2. The IC₅₀ value for cells expressing the best mutant, J17, was 12 μ g/ml ampicillin, which is even lower than that of

WO 02/086075

PCT/US02/12635

cells with the orthogonal tRNA^{Gln}CUA derived from *Saccharomyces cerevisiae* expressed (21 µg/ml ampicillin). When J17 was coexpressed with the *Methanococcus jannaschii* TyrRS, cells survived to an IC₅₀ value of 436 µg/ml ampicillin, providing a selection window (ratio of IC₅₀ value with TyrRS to IC₅₀ value without TyrRS) of 35-fold. In addition, the expression of all these mutant tRNAs did not affect the growth of *Escherichia coli* cells.

[0241] Table 2. In vivo β-lactamase assay of selected suppressor tRNAs

Suppressor tRNA		IC ₅₀ (µg/ml of ampicillin)	
		Coexpressed with pBLAM	Coexpressed with pBLAM-JYRS
10	mj-tRNA ^{Tyr} CUA	56	1220
	No tRNA ^{Tyr} CUA	10	10
Mutant tRNAs selected from anticodon-loop library			
	AA2	22	1420
	AA3	10	110
15	AA4	12	135
Mutant tRNAs selected from all-loop library			
Mutant tRNAs surviving both selections			
	J15	30	845
	J17	12	436
20	J18	20	632
	J22	14	459
Mutant tRNAs surviving negative selection only			
	N11	11	16
	N12	9	18
25	N13	10	12
	N16	9	9

Plasmid pBLAM was used to express the β-lactamase gene with an amber codon at Ala184; plasmid pBLAM-JYRS expressed the amber mutant and the TyrRS of *Methanococcus jannaschii*. Suppressor tRNAs were encoded on pAC plasmid and cotransformed with pBLAM or pBLAM-JYRS in the assay.

[0242] To confirm the properties of the selected suppressor tRNAs, they were tested in another in vivo assay based on the suppression of an amber codon in the chloramphenicol acetyltransferase (CAT) gene. In contrast to β-lactamase which is

WO 02/086075

PCT/US02/12635

secreted into the periplasm, CAT localizes in the cytoplasm. Moreover, ampicillin is bacteriocidal while chloramphenicol is bacteriostatic. As shown in Table 3 below, the selected suppressor tRNAs also were orthogonal in the CAT assay, indicating their suitability for CAT selections.

- 5 [0243] Table 3. In vivo chloramphenicol acetyltransferase assay of selected suppressor tRNAs

Suppressor tRNA	IC ₅₀ (μg/ml of chloramphenicol)	
	pYC only	pYC+pBK-JYRS
mj-tRNA ^{Tyr} CUA	27	308
10 No tRNA ^{Tyr} CUA	3	3
J15	11	297
J17	4	240
J18	6	284
J22	5	271

- 15 pYC plasmids encoded the chloramphenicol acetyltransferase gene with an amber codon at Asp112 and different suppressor tRNAs listed in the left column of the table. pBK-JYRS was used to express the TyrRS of *Methanococcus jannaschii*.

- [0244] The in vivo complementation assay which is based on suppression of an amber codon in the β-lactamase gene was carried out as described. See, e.g., Liu & Schultz, (1999), *supra*; and, Wang, et al., (2000), *supra*. In the chloramphenicol acetyltransferase (CAT) assay, an amber codon was substituted for Asp112 in the CAT gene of pACYC184 to afford pACMD112TAG. See, e.g., M. Pastnak, T.J. Magliery, P. G. Schultz, *A new orthogonal suppressor tRNA/aminoacyl-tRNA synthetase pair for evolving an organism with an expanded genetic code*, *Helv. Chim. Acta* 83:2277-2286 (2000). The genes encoding the suppressor tRNAs under the control of the lpp promoter and rrmC terminator were excised from pAC plasmids with NcoI and Aval, and inserted into the pre-digested pACMD112TAG to afford plasmids pYC-JY, pYC-J15, pYC-J17, pYC-J18, and pYC-J22, respectively. Plasmid pBK-JYRS, a derivative of pBR322, was used to express the *Methanococcus jannaschii* TyrRS under the control of the Escherichia coli GlnRS promoter and terminator. The survival of Escherichia coli DH10B cells transformed with pYC plasmid alone or cotransformed with pYC and pBK-JYRS was
- 20
- 25
- 30

WO 02/086075

PCT/US02/12635

titrated against a wide range of chloramphenicol concentrations added to the growth media, and IC₅₀ values were interpolated from the curves.

[0245] For comparison, four colonies were randomly picked that passed the negative selection only, and tested the tRNAs using the in vivo complementation assay.

- 5 All of them had very low IC₅₀ values when transformed with pBLAM, indicating the negative selection worked well. See Table 2. The IC₅₀ values were also low when cotransformed with pBLAM-JYRS, revealing that the positive selection functions to delete tRNAs that cannot be charged by the *Methanococcus jannaschii* TyrRS.

- [0246] Analysis of the DNA sequences of the selected tRNAs yielded a
10 characteristic pattern of nucleotide substitutions. See Figure 13. tRNAs that passed both negative and positive selections all had C32 and T60 unchanged, while G37 was mutated to A, and T17a was mutated to either A or G. Some semi-conserved changes included mutation of A38 to either C or A; mutation of T45 to either T or A; mutation of T47 to either G or T. Other mutations had no obvious common pattern. Twenty (20) tRNAs that
15 passed the negative selection only were also sequenced, four of which are shown in Figure 13, and found they all lacked at least one of the common mutations listed above.

- [0247] The preferred nucleotides in the selected mutant suppressor tRNAs can play the following roles: (i) they can function as negative determinants for recognition by the *Escherichia coli* synthetases; (ii) they can be identity elements for aminoacylation by
20 *Methanococcus jannaschii* TyrRS; or (iii) they can also optimize the tRNA's interaction with *Escherichia coli*'s translational machinery so as to increase the suppression efficiency of the tRNA. It is noteworthy that the G37A mutation was found in tRNAs selected from both the anticodon-loop and all-loop library. This mutation is consistent with previous studies that showing that adenine at position 37 enhances amber suppression efficiency.
25 See, e.g., M. Yarus, *Translational efficiency of transfer RNA's: Use of an expanded anticodon*, *Science* 218:646-652 (1982); D. Bradley, J.V. Park, L. Soll, *tRNA^{2Gln} Su+2 mutants that increase amber suppression*, *J. Bacteriol.* 145:704-712 (1981); and, L. G. Kleina, J. Masson, J. Normanly, J. Abelson, J.H. Miller, *Construction of Escherichia coli amber suppressor tRNA genes. II. Synthesis of additional tRNA genes and improvement of
30 suppressor efficiency*, *J. Mol. Biol.* 213:705-717 (1990). Fechter et al. recently reported that the complete identity set for *Methanococcus jannaschii* tRNA^{Tyr} is six nucleotides (C1G72, A73, and anticodon G34U35A36). See P. Fechter, J. Rudinger-Thirion, M.

WO 02/086075

PCT/US02/12635

- Tukalo, R. Giegé, *Major tyrosine identity determinants in Methanococcus jannaschii and Saccharomyces cerevisiae tRNA^{Tyr} are conserved but expressed differently*, Eur. J. Biochem. 268:761-767 (2001). The presence of C32 and T60 in all selected mutant suppressors therefore is not required for recognition by *Methanococcus jannaschii* TyrRS.
- 5 All *Escherichia coli* tRNAs have T at position 60 except four tRNAs which have C. See, M. Sprinzl, C. Horn, M. Brown, A. Loudovitch, S. Steinberg, *Compilation of tRNA sequences and sequences of tRNA genes*, Nucleic Acids Res. 26:148-153 (1998). Based on the crystal structure of yeast tRNA^{Phe}, nucleotide 60 does not interact with other nucleotides. See J. L. Sussman, S. R. Holbrook, R. W. Warrant, G. M. Church, S.H. Kim,
- 10 *Crystal structure of yeast phenylalanine transfer RNA. I. Crystallographic refinement*, J. Mol. Biol. 123:607-630 (1978). Thus, T60 may maintain the shape of the TC loop for productive interaction with the *Escherichia coli* translational machinery. The change of the TC loop structure may affect translational fidelity, as the insertion of a nucleotide between T60 and the conserved C61 enables a glycine tRNA to shift reading frame. See,
- 15 D. J. O'Mahony, B.H. Hims, S. Thompson, E.J. Murgola, J.F. Atkins, *Glycine tRNA mutants with normal anticodon loop size cause 1 frameshifting*, Proc. Natl. Acad. Sci. USA 86:7979-7983 (1989). The role of C32 is not obvious - position 32 in *Escherichia coli* tRNAs includes T, C, and A, and two *Escherichia coli* tRNA^{Tyr}s do have C32. As for position 17a, only tRNA^{Thr} has an A at this position.
- 20 [0248] All of the selected suppressor tRNAs are poorer substrates for *Escherichia coli* synthetases relative to the *Methanococcus jannaschii* tRNA^{Tyr}CUA, resulting in less mischarging when introduced into *Escherichia coli*. These tRNAs can also be stably maintained in *Escherichia coli* without adverse effects on the growth of host cells. Moreover, they can still be charged efficiently by *Methanococcus jannaschii* TyrRS. All
- 25 these properties make the mutant suppressor tRNA together with the *Methanococcus jannaschii* TyrRS a robust orthogonal tRNA-synthetase pair for the selective incorporation of unnatural amino acids into proteins in vivo. The J17 mutant suppressor tRNA and an engineered mutant TyrRS has been used to deliver O-methyl-L-tyrosine in response to a TAG codon with a fidelity rivaling that of the common 20 amino acids. See, L. Wang, A.
- 30 Brock, B. Herberich, P. G. Schultz, *Expanding the genetic code of Escherichia coli*, Science, 292:498-500 (2001).

WO 02/086075

PCT/US02/12635

[0249] Example 2- Mutating TyrRS so that it charges the mutRNA Tyr/CUA with an unnatural amino acid, O-methyl-L-tyrosine

[0250] A unique transfer RNA (tRNA)-aminoacyl tRNA synthetase pair has been generated that expands the number of genetically encoded amino acids in *Escherichia coli*.

5 When introduced into *Escherichia coli*, this pair leads to the in vivo incorporation of the synthetic amino acid O-methyl-L-tyrosine, added exogenously to the growth medium, into protein in response to an amber nonsense codon. The fidelity of translation is greater than 99%, as determined by analysis of dihydrofolate reductase containing the unnatural amino acid. This approach provides a general method for increasing the genetic repertoire of
10 living cells to include a variety of amino acids with novel structural, chemical and physical properties not found in the common twenty amino acids.

[0251] An orthogonal tRNA/synthetase pair in *Escherichia coli* can be generated by importing a pair from a different organism, if cross-species aminoacylation is inefficient, and, optionally, the anticodon loop is not a key determinant of synthetase
15 recognition. One such candidate pair is the tyrosyl tRNA/synthetase pair of *Methanococcus jannaschii* (*Methanococcus jannaschii*), an archaeobacterium whose tRNA^{Tyr} identity elements differ from those of *Escherichia coli* tRNA^{Tyr} (in particular, the first base pair of the acceptor stem is GC in *Escherichia coli* and CG in *Methanococcus jannaschii*), and whose tyrosyl synthetase (TyrRS) has only a minimalist
20 anticodon loop binding domain. See, e.g., B. A. Steer, & P. Schimmel, *J. Biol. Chem.* 274:35601-6 (1999). In addition, the *Methanococcus jannaschii* TyrRS does not have an editing mechanism, see, e.g., Jakubowski & Goldman, *Microbiol. Rev.*, 56:412 (1992), and therefore should not proofread an unnatural amino acid ligated to the tRNA. The *Methanococcus jannaschii* TyrRS efficiently aminoacylates an amber suppressor tRNA
25 derived from its cognate tRNA^{Tyr}, see, e.g., Wang, et al., (2000 *J. Am. Chem. Soc.*, *supra*), but does not aminoacylate *Escherichia coli* tRNAs, see, e.g., Steer & Schimmel, (1999), *supra*. Moreover, the *Methanococcus jannaschii* tRNA^{Tyr}_{CUA} is a poor substrate for the *Escherichia coli* synthetases but functions efficiently in protein translation in *Escherichia coli*. See, e.g., Wang, et al., (2000 *J. Am. Chem. Soc.*, *supra*.

30 **[0252]** To further reduce recognition of the orthogonal tRNA, *Methanococcus jannaschii* tRNA^{Tyr}_{CUA}, by *Escherichia coli* synthetases, eleven nucleotides of the tRNA that do not interact directly with the *Methanococcus jannaschii* TyrRS (C16, C17, U17a, U20,

WO 02/086075

PCT/US02/12635

C32, G37, A38, U45, U47, A59 and U60) were randomly mutated to generate a suppressor tRNA library. This tRNA library was passed through a negative selection (e.g., suppression of amber mutations in a toxic reporter gene, e.g., barnase gene), which removes tRNAs that are aminoacylated by *Escherichia coli* synthetases, and then a
 5 positive selection for tRNAs that are efficiently aminoacylated by *Methanococcus jannaschii* TyrRS (e.g., suppression of amber mutations in a reporter gene, e.g., β -lactamase gene).

[0253] The orthogonal nature of the resulting suppressor tRNAs was tested by an *in vivo* complementation assay, which is based on suppression of an amber stop codon at a
 10 nonessential position (e.g., Ala184) of a reporter gene on a vector, e.g., the TEM-1 β -lactamase gene carried on plasmid pBLAM. Aminoacylation of a transformed suppressor tRNA by any endogenous *Escherichia coli* synthetase results in cell growth in the presence of ampicillin. *Escherichia coli* transformed with *Methanococcus jannaschii* tRNA^{Tyr}_{CUA} and the reporter construct, pBLAM, survive at 55 μ g/mL ampicillin. When the
 15 best mutant suppressor tRNA (mtRNA^{Tyr}_{CUA}) selected from the library was expressed, cells survived at only 12 μ g/mL ampicillin; similar values are obtained in the absence of any suppressor tRNA. The mutant suppressor tRNA contained the following nucleotide substitutions: C17A, U17aG, U20C, G37A, and U47G. When the *Methanococcus jannaschii* TyrRS is coexpressed with this mtRNA^{Tyr}_{CUA}, cells survive at 440 μ g/mL
 20 ampicillin. Thus, the mtRNA^{Tyr}_{CUA} is a poorer substrate for the endogenous synthetases than the *Methanococcus jannaschii* tRNA^{Tyr}_{CUA} but is still aminoacylated efficiently by the *Methanococcus jannaschii* TyrRS.

[0254] To alter the amino acid specificity of the orthogonal TyrRS so that it charges the mtRNA^{Tyr}_{CUA} with a desired unnatural amino acid, a library of TyrRS mutants
 25 was generated and screened. Based on the crystal structure of the homologous TyrRS from *Bacillus stearothermophilus*, see, e.g., P. Brick, T. N. Bhat, D. M. Blow, *J. Mol. Biol.*, 208:83 (1988), five residues (Tyr³², Glu¹⁰⁷, Asp¹⁵⁸, Ile¹⁵⁹ and Leu¹⁶²) in the active site of *Methanococcus jannaschii* TyrRS which are within 6.5 Å of the para position of the aryl ring of bound tyrosine were mutated. See, Figure 14. These residues were all
 30 initially mutated to alanine, and the resulting inactive Ala₅ TyrRS was used as a template for polymerase chain reaction (PCR) random mutagenesis with doped oligonucleotides.

WO 02/086075

PCT/US02/12635

[0255] For example, the TyrRS gene was expressed under the control of *Escherichia coli* GlnRS promoter and terminator in plasmid pBK-JYRS, a pBR322 derived plasmid with kanamycin resistance. Residues Tyr³², Glu¹⁰⁷, Asp¹⁵⁸, Ile¹⁵⁹ and Leu¹⁶² were substituted with Ala by site-directed mutagenesis to provide plasmid pBK-JYA5. Eight (8) oligonucleotides with NNK (N=A + T + G + C and K= G + T, and M= C + A), e.g., oligonucleotides LW157 5'-GGAATTCATATGGACGAATTTGAAATG-3', LW164 5'-GTATTTTACCACCTGGTTCAAACCTATMNNAGCAGATTTTTCATCTTTTTCATCTTTTAAAC-3', LW159 5'-TAGGTTTTGAACCAAGTGGTAAATAC-3', LW165 5'-CATTCAGTGTATAATCCTTATCAAGCTGGAAMNACTTCCATAACATATTTTGCCTTTAAAC-3', LW161 5'-TCCAGCTTGATAAGGATTATACA CTGAATG-3', LW167 5'-CATCCCTCCAAGTCAACATCAACGCCMNATA ATGMNMMNATTAACTGCATTATTGGATAGATAAC-3', LW163 5'-GCGT TGATGTTGCAGTTGGAGGGATG-3', and LW105 5'-AAACTGCAGTTATAAT CTCTTCTAATTGGCTC-3' (Operon, CA) at the mutation sites were used for PCR amplification of the Ala₅ TyrRS mutant (pBK-JYA5) and ligated back into the NdeI-PstI-digested pBK-JYA5 to afford the TyrRS library. The ligated vectors were transformed into *Escherichia coli* DH10B competent cells to yield a library of 1.6 X 10⁹ colony forming unit (cfu). The TyrRS genes from 40 randomly picked colonies were sequenced to confirm that there was no base bias at the randomized NNK positions and no other unexpected mutations. The library was amplified by maxiprep, and supercoiled DNA was used to transform the selection strain pYC-J17.

[0256] A positive selection was then applied to the library of mutated orthogonal TyrRS that is based on suppression of an amber stop codon at a nonessential position (e.g., Asp112) in the chloramphenicol acetyltransferase (CAT) gene. See, e.g., M. Pastnak, T.J. Magliery, P. G. Schultz, *Helv. Chim. Acta*, 83:2277 (2000). Cells transformed with the mutant TyrRS library and mtRNA_{CUA}^{Tyr} gene were grown in media containing the unnatural amino acid and selected for their survival in the presence of various concentrations of chloramphenicol. If a mutant TyrRS charges the orthogonal mtRNA_{CUA}^{Tyr} with any amino acid, either natural or unnatural, the cell produces CAT and survives. The surviving cells were then grown in the presence of chloramphenicol and in the absence of the unnatural amino acid. Those cells that did not survive, e.g., which encode mutant TyrRS's that

WO 02/086075

PCT/US02/12635

charge the orthogonal mtRNA^{Tyr}_{CUA} with an unnatural amino acid, were isolated from a replica plate supplemented with the unnatural amino acid. The mutant TyrRS genes were isolated from these cells, recombined in vitro by DNA shuffling, and transformed back into *Escherichia coli* for further rounds of selection with increasing concentrations of chloramphenicol.

[0257] A tyrosine analogue with the para hydroxyl group substituted with the methoxy group was used in the selection. Optionally, other tyrosine analogues can also be used in selection, e.g., tyrosine analogues with different functional groups at the para position of the aryl ring (acetyl, amino, carboxyl, isopropyl, methyl, O-methyl and nitro, etc.). For example, the gene encoding mtRNA^{Tyr}_{CUA} was expressed in *Escherichia coli* DH10B cells under the control of the lpp promoter and rrnC terminator in plasmid pYC-J17, a pACYC184 derivative that also encodes the Asp₁₁₂ TAG CAT mutant. Supercoiled DNA encoding the TyrRS library was transformed into *Escherichia coli* DH10B competent cells containing pYC-J17 to yield a library of size greater than 3×10^9 cfu, ensuring complete coverage of the original library. Cells were then plated on minimal media plates containing 1% glycerol and 0.3 mM leucine (GMML) with 17 µg/mL tetracycline (Tet), 25 µg/mL kanamycin (Kan), 50 µg/mL of chloramphenicol (Cm), and 1 mM unnatural amino acid. After incubation at 37°C for 44 hours, colonies on plates supplied with O-methyl-L-tyrosine were pooled, plasmids were isolated and retransformed into *Escherichia coli* DH10B competent cells containing pYC-J17, and the transformed cells were positively selected on 50 µg/mL of Cm. Colonies (96) were individually picked from the plate, diluted into 100 µL of liquid GMML media, and streaked onto two sets of Kan/Tet GMML plates with various concentration of Cm. No O-methyl-L-tyrosine was added to plate set 1 and the concentration of Cm was varied from 10 – 25 µg/mL; plate set 2 contained 1mM O-methyl-L-tyrosine and 50 µg/mL of Cm. Replicates of colonies that did not grow on 15 µg/mL of Cm in plate set 1 were picked from plate set 2. Plasmids containing the TyrRS gene were purified and recombined in vitro by DNA shuffling using Stemmer's protocol with the exception of 10 mM Mn²⁺ instead of Mg²⁺ in the fragmentation reaction. See, W. P. C. Stemmer, *Nature* 370, 389-91 (1994); and, I. A. Lorimer, I. Pastan, *Nucleic Acids Res.* 23, 3067-8 (1995). The library was then religated into predigested pBK-JYA5 vector to afford a second generation TyrRS library with a typical size of 8×10^8 to 3×10^9 cfu. Thirty randomly selected members from the library

WO 02/086075

PCT/US02/12635

- were sequenced. The mutagenic rate introduced by DNA shuffling was 0.35%. This library was transformed into the selection strain for the next round of selection followed by shuffling. The concentration of Cm in the positive selection and in plate set 2 was raised to 80 µg/mL for the second round and 120 µg/mL for the third round; the concentration of Cm in plate set 1 was unchanged. After three rounds of DNA shuffling, colonies began to grow on 20 – 25 µg/mL Cm in plate set 1, indicating that the TyrRS mutants were accepting natural amino acids as substrates. Therefore, the best clone selected after two rounds of DNA shuffling was characterized in detail.
- [0258]** Two rounds of selection and DNA shuffling were carried out and a clone was evolved whose survival in chloramphenicol was dependent on the addition of 1mM O-methyl-L-tyrosine to the growth media. In the absence of O-methyl-L-tyrosine, cells harboring the mutant TyrRS were not viable on minimal media plates containing 1% glycerol, 0.3 mM leucine (GMML), and 15 µg/mL of chloramphenicol. Cells were able to grow on GMML plates with 125 µg/mL chloramphenicol in the presence of 1mM O-methyl-L-tyrosine. Similar results were obtained in liquid GMML. As a control, cells with the *mtRNA^{Tyr}_{CUA}* and the inactive Ala₅ TyrRS did not survive at the lowest concentration of chloramphenicol used, either in the presence or absence of 1mM O-methyl-L-tyrosine. See **Figure 14**. Addition of 1mM O-methyl-L-tyrosine itself does not significantly affect the growth rate of *Escherichia coli*.
- [0259]** Analysis of the sequence of the mutant TyrRS that charges the *mtRNA^{Tyr}_{CUA}* with O-methyl-L-tyrosine revealed the following mutations: Tyr³²→Gln³², Asp¹⁵⁸→Ala¹⁵⁸, Glu¹⁰⁷→Thr¹⁰⁷, and Leu¹⁶²→Pro¹⁶². See **Figure 14**. Based on the x-ray crystal structure of the homologous *B. stearothermophilus* TyrRS, the loss of the hydrogen-bonding network between Tyr³², Asp¹⁵⁸ and substrate tyrosine can disfavor binding of tyrosine to the mutant TyrRS. Indeed, mutation of Asp¹⁷⁶ (which corresponds to Asp¹⁵⁸ in *Methanococcus jannaschii*) of *B. stearothermophilus* TyrRS yields inactive enzyme. See, e.g., G.D.P. Gray, H.W. Duckworth, A. R. Fernst, FEBS Lett. 318:167 (1993). At the same time, the Asp¹⁵⁸→Ala¹⁵⁸ and Leu¹⁶²→Pro¹⁶² mutations create a hydrophobic pocket that allows the methyl group of O-methyl-L-tyrosine to extend further into the substrate-binding cavity. Other important catalytic residues in the active site, which bind to the ribose or the phosphate group of the adenylate, were unchanged after two rounds of DNA shuffling.

WO 02/086075

PCT/US02/12635

- [0260]** Kinetics of adenylate formation of O-methyl-L-tyrosine and tyrosine with adenosine triphosphate (ATP) catalyzed by the mutant TyrRS were analyzed *in vitro* using a pyrophosphate-exchange assay at 37°C. For example, the mutant TyrRS gene with six histidines at its C-terminus was cloned into plasmid pQE-60 (Qiagen, CA) to generate
- 5 plasmid pQE-mJYRS. Protein was purified by immobilized metal affinity chromatography according to manufacture's protocol (Qiagen, CA). Pyrophosphate (PPi) exchange was carried out at 37°C in a reaction mixture containing 100 mM TrisHCl (pH7.5), 10 mM KF, 5 mM MgCl₂, 2 mM ATP, 2 mM NaPPi, 0.1 mg/mL bovine serum albumin, approximately 0.01 μCi/μL [³²P]NaPPi, and various concentrations of tyrosine or
- 10 O-methyl-L-tyrosine. Reactions were initiated with the addition of the purified mutant TyrRS, and aliquots were periodically taken and quenched with 0.2 M NaPPi, 7% perchloric acid, and 2% activated charcoal. The charcoal was filtered and washed with 10 mM NaPPi (pH2), then measured by scintillation counting to determine the ³²P levels in charcoal-adsorbed ATP. Values of k_{cat} and K_m were calculated by direct fitting of the
- 15 Michaelis-Menten equation using nonlinear regression analysis.
- [0261]** The Michaelis constant (K_m) for tyrosine (5833 ± 902 μM) is approximately 13-fold higher than that for O-methyl-L-tyrosine (443 ± 93 μM), and the catalytic rate constant (k_{cat}) for tyrosine ($1.8 \pm 0.2 \times 10^{-3} \text{ s}^{-1}$) is eightfold less than that for O-methyl-L-tyrosine ($14 \pm 1 \times 10^{-3} \text{ s}^{-1}$). Thus, the value of k_{cat}/K_m of the mutant TyrRS
- 20 for O-methyl-L-tyrosine is about 100-fold higher than that of tyrosine. The physiological concentration of tyrosine in *Escherichia coli* is about 80 μM, which is far below K_m value (5833 μM) of the mutant TyrRS for tyrosine. Presumably, the concentration of O-methyl-L-tyrosine in cells is comparable or greater than the K_m (443 μM).
- [0262]** This example shows that it is possible to augment the protein biosynthetic
- 25 machinery of *Escherichia coli* to accommodate additional genetically encoded amino acids. The ability to introduce novel amino acids into proteins directly in living cells will provide new tools for studies of protein and cellular function and can lead to generation of proteins with enhanced properties compared to a naturally occurring protein. The methods described here can be applied to other amino acids with novel spectroscopic, chemical,
- 30 structural or the like properties. The *Escherichia coli* ribosome has been shown to be able to incorporate amino acids with a wide array of side chains into proteins using *in vitro* protein synthesis. See, e.g., C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, P. G.

WO 02/086075

PCT/US02/12635

Schultz, *Science* 244, 182-8 (1989). Additional orthogonal tRNA/synthetase pairs, *see*, e.g., D. R. Liu, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* 96, 4780-5 (1999); and, A. K. Kowal, C. Kohrer, U.L., RajBhandary, *Proc. Natl. Acad. Sci. U.S.A.*, 98:2268 (2001), as well as four base codons, *see*, e.g., T. J. Magliery, J. C. Anderson, P. G. Schultz, *J. Mol. Biol.*, 307:755 (2001); and, B. Moore, B. C. Persson, C. C. Nelson, R. F. Gesteland, J. F. Atkins, *J. Mol. Biol.*, 298:195 (2000), and other selector codons described herein, can further expand the number and scope of amino acids that can be incorporated. Orthogonal pairs for eukaryotic cells can also be generated by the methods provided herein.

[0263] *See also* corresponding patent application "In vivo Incorporation of Unnatural Amino Acids" attorney docket number 54-000120PC/US which is incorporated herein by reference. This application describes an example of the generation of an O-methyl-L-tyrosine mutant of dihydrofolate reductase (DHFR) using the above-described system.

[0264] **Example 3- Mutating TyrRS so that it charges the mutRNA Tyr/CUA with an unnatural amino acid, L-3-(2-Naphthyl)alanine**

[0265] This example provides another orthogonal pair that can be used to incorporate a second unnatural amino acid, L-3-(2-Naphthyl)alanine into proteins in an organism, e.g., *Escherichia coli*. An example of the methods used to generate the orthogonal pair that incorporates the unnatural amino acid into proteins is described below. More details describing the incorporation of the unnatural amino acid into a protein can be found in corresponding patent application "In vivo incorporation of unnatural amino acid" attorney docket number 54-000120PC/US incorporated herein by reference.

[0266] An amber stop codon and its corresponding orthogonal amber suppressor tRNA, $\text{mu tRNA}_{\text{CUA}}^{\text{Tyr}}$, were selected to encode an unnatural amino acid. As described above, and *see* Wang & Schultz, *Chem. Biol.* 8:883-890 (2001). The *Methanococcus jannaschii* tyrosyl-tRNA synthetase (TyrRS) was used as the starting point for the generation of an orthogonal synthetase with unnatural amino acid specificity. This TyrRS does not aminoacylate any endogenous *Escherichia coli* tRNAs, *see*, e.g., Steer & Schimmel, *J. Biol. Chem.*, 274:35601-35606 (1999), but aminoacylates the $\text{mu tRNA}_{\text{CUA}}^{\text{Tyr}}$ with tyrosine. *See*, e.g., Wang, Magliery, Liu, Schultz, *J. Am. Chem. Soc.*, 122:5010-5011 (2000). L-3-(2-naphthyl)-alanine was chosen for this study since it represents a significant structural

WO 02/086075

PCT/US02/12635

- perturbation from tyrosine and may have novel packing properties. To change the amino acid specificity of the TyrRS so that it charges the mu rRNA_{CUA}^{Tyr} with L-3-(2-naphthyl)-alanine and not any common 20 amino acids, a library of *Methanococcus jannaschii* TyrRS mutants was generated and screened. On the basis of an analysis of the crystal structure of the homologous TyrRS from *Bacillus stearothermophilus*, see, Brick, Bhat, Blow, J. Mol. Biol., 208:83-98 (1989), five residues (Tyr³², Asp¹⁵⁸, Ile¹⁵⁹, Leu¹⁶², and Ala¹⁶⁷) in the active site of *Methanococcus jannaschii* TyrRS that are within 7 Å of the para position of the aryl ring of tyrosine were mutated. See Figure 15. No synthetases specific for L-3-(2-naphthyl)alanine were selected from the mutant TyrRS library reported in Wang, Brock, Herberich, Schultz, Science, 292:498-500 (2001). To reduce the wild-type synthetase contamination in the following selection, these residues (except Ala¹⁶⁷) were first all mutated to alanine. The resulting inactive Ala₅ TyrRS gene was used as a template for polymerase chain reaction (PCR) random mutagenesis with oligonucleotides bearing random mutations at the corresponding sites.
- 15 [0267] The mutant TyrRS library was first passed through a positive selection based on suppression of an amber stop codon at a nonessential position (Asp¹¹²) in the chloramphenicol acetyltransferase (CAT) gene. Cells transformed with the mutant TyrRS library and the mu rRNA_{CUA}^{Tyr} gene were grown in minimal media containing 1 mM L-3-(2-naphthyl)-alanine and 80 µg/mL chloramphenicol. Cells can survive only if a mutant
- 20 TyrRS aminoacylates the mu rRNA_{CUA}^{Tyr} with either natural amino acids or L-3-(2-naphthyl)-alanine. The surviving cells were then grown in the presence of chloramphenicol and the absence of the unnatural amino acid. Those cells that did not survive must encode a mutant TyrRS that charges the mu rRNA_{CUA}^{Tyr} with L-3-(2-naphthyl)-alanine, and were picked from a replica plate supplied with the unnatural amino acid. After three rounds of positive
- 25 selection followed by a negative screen, four TyrRS's were characterized using an in vivo assay based on the suppression of the Asp¹¹²TAG codon in the CAT gene.

[0268] Table 4 In vivo chloramphenicol acetyltransferase assay of mutant TyrRS.^a

Mutant TyrRS	IC ₅₀ (µg/mL of chloramphenicol)	
	No L-3-(2-naphthyl)-Ala	Add L-3-(2-naphthyl)-Ala

30

WO 02/086075		PCT/US02/12635
no TyrRS	4	4
wt TyrRS	240	240
<i>After selection</i>		
S1-TyrRS	30	120
S2-TyrRS	30	120
S3-TyrRS	25	110
S4-TyrRS	35	100
<i>After DNA shuffling</i>		
SS12-TyrRS	9	150

10 ^a A pYC-J17 plasmid was used to express the mu_{RNA_{CCA}} gene and the chloramphenicol acetyltransferase gene with an amber stop codon at Asp112. A pBK plasmid was used to express TyrRS, and was cotransformed with pYC-J17 into *Escherichia coli* DH10B. Cell survival on GMML plates was titrated in the presence of different concentrations of chloramphenicol.

15 [0269] In the absence of L-3-(2-naphthyl)-alanine, cells expressing the selected TyrRS and the mu_{RNA_{CCA}} survived in 25 to 35 µg/mL chloramphenicol on minimal media plates containing 1% glycerol and 0.3 mM leucine (GMML plate); in the presence of L-3-(2-naphthyl)-alanine, cells survived in 100 to 120 µg/mL chloramphenicol on GMML plates. Compared to the IC₅₀ value in the absence of any TyrRS (4 µg/mL

20 chloramphenicol), these results indicate that the selected TyrRS's accept L-3-(2-naphthyl)-alanine, but also still charge natural amino acids to some degree. *See* Table 4 above.

[0270] To further reduce the activity of the mutant TyrRS toward natural amino acids, one round of DNA shuffling was carried out using the above four mutant genes as templates. The resulting mutant TyrRS library was passed through two additional rounds

25 of positive selections and negative screens. One mutant TyrRS (SS12-TyrRS) was evolved, whose activity for natural amino acids was greatly reduced (IC₅₀=9 µg/mL chloramphenicol) while its activity toward L-3-(2-naphthyl)-alanine was enhanced (IC₅₀=150 µg/mL chloramphenicol). *See* Table 4.

WO 02/086075

PCT/US02/12635

- [0271] The evolved SS12-TyrRS has the following mutations: Tyr³²→Leu³², Asp¹⁵⁸→Pro¹⁵⁸, Ile¹⁵⁹→Ala¹⁵⁹, Leu¹⁶²→Gln¹⁶², and Ala¹⁶⁷→Val¹⁶⁷. See Figure 15. Based on the crystal structure of the homologous *B. stearothermophilus* TyrRS, the mutations of Tyr³²→Leu³² and Asp¹⁵⁸→Pro¹⁵⁸ can result in the loss of hydrogen bonds between Tyr³², Asp¹⁵⁸, and the native substrate tyrosine, thus disfavoring the binding of tyrosine to SS12-TyrRS. Most residues are mutated to amino acids with hydrophobic side chains, which are expected to favor binding of L-3-(2-naphthyl)-alanine. The crystal structure of the wild-type *Methanococcus jannaschii* TyrRS and the evolved SS12-TyrRS can be determined by available methods.
- 10 [0272] The mu tRNA^{Tyr}_{CUA}/SS12-TyrRS pair was capable of selectively inserting L-3-(2-naphthyl)-alanine into proteins in response to the amber codon with fidelity rivaling that of the natural amino acids based on cell growth, protein expression and mass spectrometry examples described herein and in corresponding application "In vivo incorporation of unnatural amino acids" attorney docket number 54-000120PC/US. See also, Wang, Brock, and Schultz, *Adding L-3-(2-Naphthyl)alanine to the genetic code of E. coli*, *J. Am. Chem. Soc.*, (2002) 124(9):1836-7. This result, which involves an amino acid that is structurally distinct from tyrosine, confirms that the methods described herein are generalizable to a variety of unnatural amino acids.
- 20 [0273] **Example 4- Mutating TyrRS so that it charges the mutRNA Tyr/CUA and screening for the mutated TyrRS with the desired properties by other methods, e.g., FACs and Phage display and panning**
- [0274] Orthogonal pairs can also be selected by using reporter genes and proteins as described above, along with in vivo FACS screening, antibody detection, in vitro phage display and panning, or the like. See, Wang & Schultz, *Expanding the genetic code*, *Chem. Commun.*, 1:1-11 (2002).
- 25 [0275] For example, a general fluorescence-activated cell sorting (FACS) based screen has been developed with, e.g., green fluorescent protein (GFP) as the reporter, to screen for synthetases. See Figure 16, Panel A, and Panel B Synthetase activity is reported by suppression of the selector codon, e.g., an amber stop codon (TAG) within T7 RNA polymerase, which drives the expression of GFP. See, e.g., Figure 26 for another example of selection/screening methods of the invention. Only when the amber codons are suppressed can cells produce functional T7 RNA polymerase and express GFP,
- 30

WO 02/086075

PCT/US02/12635

rendering cells fluorescent. In the positive screen, fluorescent cells are collected which encode active synthetases charging the orthogonal tRNA with either natural or unnatural amino acids. The selected cells are then diluted and grown in the absence of the unnatural amino acid, and then sorted by FACS for cells without fluorescence, e.g., that express synthetases with specificities for unnatural amino acids only. **Figure 17, Panel A, Panel B Panel C and Panel D** illustrates suppression of a selector codon, e.g., an amber codon, using glutamine synthetase. By setting the collection threshold of the fluorescence intensity, the stringency of both positive and negative screen can be conveniently controlled.

- 10 **[0276]** A direct positive selection specific for a particular unnatural amino acid has also been developed which exploits the high affinity of a monoclonal antibody for an unnatural amino acid displayed on a phage surface. See **Figure 18**. See, M. Pastmak and P. G. Schultz, Bioorg. Med. Chem., 9:2373 (2001). For example, a C3 peptide with an amber mutation is fused to the N-terminus of VSCM13 phage coat protein pIII, such that
- 15 phage production requires suppression of the amber stop codon. Cells harboring a phagemid that expresses an orthogonal suppressor tRNA and a synthetase library are infected with the C3TAG phage. An active synthetase results in suppression of C3TAG and display of its cognate amino acid on the phage surface. The phage pool is then incubated with immobilized monoclonal antibodies directed against the unnatural amino
- 20 acid to isolate only those phage carrying the synthetase specific for the unnatural amino acid. In a simulated selection, phage displaying Asp were enriched over 300-fold from a pool of phage displaying Asn using antibodies raised against the Asp-containing epitope.

- [0277]** Several in vitro screen methods can also be used. In one such method, a library of mutant synthetases is displayed on the phage, and the phage particles are panned
- 25 against immobilized aminoalkyl adenylate analogs of the aminoacyl adenylate intermediate. See **Figure 19**. For example, *Methanococcus jannaschii* TyrRS was fused to the pIII coat protein of M13 phage. This phage was enriched 1000-fold over a control phage displaying an unrelated antibody after panning against the aminoalkyl adenylate analog of tyrosyl adenylate. Given that only 0.1 to 1% of the starting TyrRS phage
- 30 population displays the TyrRS protein, the actual enrichment factor can be as high as 10^5 to 10^6 .

[0278] **Example 5- Generating an archaeal leucyl-tRNA synthetase pair**

WO 02/086075

PCT/US02/12635

- [0279] A leucyl-tRNA synthetase from an archaebacterium, *Methanobacterium thermoautotrophicum*, was identified that can aminoacylate amber and frameshift suppressor tRNAs derived from archaeal leucyl tRNAs, but does not aminoacylate any tRNAs native to *Escherichia coli*. Using a selection strategy described in the present invention, highly active tRNA substrates were identified that are selectively charged by the synthetase. Mutant libraries of synthetases can be generated and selected for that are capable of selectively charging unnatural amino acids.
- [0280] β -lactamase reporter genes were constructed with amber codons and suppressor tRNAs derived from five different archaeal leucyl tRNAs for which the anticodon was replaced with a CUA anticodon to make amber suppressor tRNAs. Seven different leucyl tRNA synthetases were cloned and were cotransformed with reporter constructs. Three synthetases gave higher levels of survival on ampicillin in the presence of the synthetase than controls lacking synthetase, and these systems were examined further. See, Figure 20.
- [0281] The next step involved determination of a synthetase that charges the suppressor tRNA without interacting with host tRNA. The two chosen systems, *Methanobacterium thermoautotrophicum* and *Methanococcus jannaschii* were expressed, and aminoacylation was performed in vitro on purified tRNA from *Halobacterium* as a positive control, and for *Escherichia coli* total tRNA. It was found that the *Methanococcus jannaschii* synthetase was able to effectively charge *Escherichia coli* tRNA, but the *Methanobacterium thermoautotrophicum* synthetase was specific towards the *Halobacterium* tRNA.
- [0282] Further improvements were made to increase the efficiency of the suppression system. The A³⁷ site of the anticodon loop was a G³⁷ in the leucyl tRNA synthetases. This mutation has been shown to be a negative determinant against aminoacylation by non-cognate synthetases in various eukaryotic cells and *Halobacterium*, and also a positive determinate for aminoacylation in yeast, but not in *Halobacterium*. A³⁷ was also shown to be a key requirement for efficient suppression. The anticodon loop was randomly mutagenized and selected for more efficient suppression. Mutating G³⁷ to A, resulted in a more efficient suppressor, which could suppress 20 fold higher concentrations of ampicillin compared to the un-mutated version. See, Figure 21.

WO 02/086075

PCT/US02/12635

[0283] To improve the tRNA so that is not preferentially charged by other synthetases in *Escherichia coli*, the acceptor stem of the tRNA was randomly mutagenized. A positive/negative selection was used to identify tRNAs that would not be charged in the absence of *Methanobacterium thermoautotrophicum* RS.

- 5 [0284] Amongst the selected mutated tRNAs observed, all conserved the discriminator base, A⁷³, which has been shown in all previous systems to be a critical positive determinate for leucyl aminoacylation. Also conserved was a C³:G⁷⁰ base pair amongst all hits that had improved orthogonality. The best mutant tRNA observed gave about a 3-fold decrease in aminoacylation without synthetase and actually an increase in
10 suppression in the presence of *Methanobacterium thermoautotrophicum* RS.

- [0285] Variants were also made that could suppress four-base codons instead of, e.g., three base codons. Four base codons offer the possibility of decoding the genetic code four bases at a time, for which 256 things could be encoded rather than 3 at a time, where only 64 amino acids can be encoded. The difficulty with using four-base codons is
15 that they require expansion of the anticodon loop for the tRNA, a perturbation which most systems are unlikely to accept. However, a first generation AGGA suppressor for the leucyl system was identified. This was generated by randomly mutagenizing the anticodon loop with 8 bases and performing selection with an AGGA- β -lactamase reporter system. See Figure 22.

- 20 [0286] The editing mechanism of the synthetase was also mutated to eliminate the editing function. The leucyl system, like several other synthetases has (at least) two active sites. One site performs activation of the amino acid with ATP to form an enzyme bound aminoacyl adenylate in complex with the synthetase, and then transfer of the amino acid onto the 3' terminus of the tRNA. A second site, however, is able to hydrolyze the amino
25 acid from the tRNA if it is not leucine. The leucine system is known to perform this post-transfer editing function for methionine and isoleucine, and it optionally does this to unnatural amino acids as well.

- [0287] Initially, the editing domain was deleted. The editing domain was replaced with a library of 6 tandem random amino acids. A positive selection was used, which was
30 based on suppression of a stop codon in β -lactamase. Many functional synthetases were obtained, but upon trying to purify the synthetases, no material in any cases could be

WO 02/086075

PCT/US02/12635

detected, and all of these synthetases displayed a temperature sensitive phenotype suggesting that the deletion of the editing domain resulted in a less stable protein.

[0288] Next, point mutations were made in the editing domain. The catalytic core of the editing domain is well conserved across species and even for different amino acids, at least for the family of branched chain amino acids. Several of these conserved sites have previously been mutated, for example a T→P mutation, and found to knock out editing function. Mutants of *Methanobacterium thermoautotrophicum* RS were constructed that were similar to several known mutants, and also a 20 member NNK library derived from T214 was made. Proteins were expressed and examined in vitro for aminoacylation with leucine and methionine. None of the previously identified mutations were transferable to our system, but a desirable mutation was identified from the T214 library. Two mutants were identified that were capable of charging with leucine, T214S and T214Q. Of these mutations, only T214Q was capable of charging methionine. The T214S mutant apparently retains the ability to edit out methionine whereas the Gln mutant has lost this function.

[0289] A library was then designed based on the crystal structure that has been solved for the *Thermus thermophilus* leucyl synthetase. The leucine side chain of the leucine aminoalkyl adenylate analog adenosine inhibitor was bound in the active site. Six sites surrounding the leucine side chain-binding pocket were replaced with randomized amino acids to create a larger library. The synthetases from this library can then be screened, e.g., by performing positive/negative double sieve selections, to identify synthetases capable of charging unnatural amino acids selectively.

[0290] **Example 6- Identification of tRNAs that efficiently suppress four-base codons**

[0291] A combinatorial approach was used to identify mutated tRNAs that efficiently suppress four-base codons. See, T. J. Magliery, J. C. Anderson and P. G. Schultz, *J. Mol. Biol.*, 307:755 (2001). A reporter library was constructed in which a serine codon in the β -lactamase gene was replaced by four random nucleotides. A mutated tRNA, e.g., suppressor tRNA, suppressor library was then generated that consists of derivatives of *Escherichia coli* with the anticodon loop (7 nt) replaced with eight or nine random nucleotides. When these two libraries are crossed, an appropriate frameshift suppressor tRNA that decodes the four-base sequence as a single codon results in

WO 02/086075

PCT/US02/12635

translation of full-length β -lactamase, rendering the cells resistant to ampicillin. Survival at higher concentrations of ampicillin indicates that the corresponding tRNA has higher suppression efficiency for the four-base codon. Using this selection, four quadruplet codons AGGA, CUAG, UAGA, and CCCU and their cognate suppressor tRNAs were identified that decode only the canonical four-base codon with efficiencies close to that of natural triplet codon suppressors. Novel five- and six-base codon suppressors have also been selected using this strategy. See, Anderson, Magliery, Schultz, *Exploring the Limits of Codon and Anticodon Size*, Chemistry & Biology, 9:237-244 (2002). These extended codons, some of which are newly identified, can be useful for the incorporation of multiple unnatural amino acids in vitro and for in vivo protein mutagenesis.

[0292] Example 7- Generation of an orthogonal tRNA-synthetase for *p*-aminophenylalanine

[0293] To generate an orthogonal synthetase pair for *p*-aminophenylalanine (pAF), the *Methanococcus jannaschii* tyrosyl-tRNA synthetase (TyrRS) and mutant tyrosine amber suppressor tRNA (TyrCUA mutRNA) pair were used as a starting point. See, Wang, L., Magliery, T. J., Liu, D. R. & Schultz, P. G. *A new functional suppressor tRNA/aminoacyl-tRNA synthetase pair for the in vivo incorporation of unnatural amino acids into proteins*. J. Am. Chem. Soc. 122:5010-5011 (2000); and, Wang, L. & Schultz, P. G. Chem. and Biol. 8:883 (2001). The pAF specific synthetase (pAFRS) was generated by modifying the amino acid specificity of the *Methanococcus jannaschii* TyrRS to accept pAF and not any of the common twenty amino acids. A combination of positive selections and negative screens was used to identify the pAFRS enzyme from a library of TyrRS variants 12 containing random amino acids at five positions (Tyr³², Glu¹⁰⁷, Asp¹⁵⁸, Ile¹⁵⁹, and Leu¹⁶²). See, Wang, L., Brock, A., Herberich, B. & Schultz, P. G. *Expanding the genetic code of Escherichia coli*. Science 292:498-500 (2001). A single reporter plasmid was used for both selection and screening. For example, the reporter plasmid is pREP(2)YC-JYCUA, which contains the genes for CAT, T7 RNA polymerase, GFP, and TyrCUA mutRNA, and a selectable marker for Tet resistance. The CAT gene contains a TAG codon substitution at position D112. The T7 RNA polymerase gene contains a seven-amino acid N-terminal leader peptide and TAG substitutions at M1 and Q107.

[0294] The positive selection is based on suppression of a TAG codon at a permissive position within the chloramphenicol acetyltransferase (CAT) gene by either

WO 02/086075

PCT/US02/12635

pAF or an endogenous amino acid. See, e.g., Wang et al. (2001), *supra*; and, Pastrnak, M., Magliery, T. J. & Schultz, P. G. *A new orthogonal suppressor tRNA/aminoacyl-tRNA synthetase pair for evolving an organism with an expanded genetic code*. Helvetica Chimica Acta 83:2277 (2000). Cells containing the TyrRS library and reporter plasmid were grown in liquid culture containing pAF and selected for survival in the presence of chloramphenicol (Cm). For example, for the positive selection, cells were grown in GMML minimal media containing 35 $\mu\text{g/ml}$ Kn, 25 $\mu\text{g/ml}$ Tet, 75 $\mu\text{g/ml}$ Cm, and 1mM pAF (Sigma).

[0295] The negative screen is based on the inability to suppress in the absence of pAF two TAG stop codons at permissive positions within the T7 RNA polymerase gene. Expression of full length T7 RNA polymerase drives expression of green fluorescent protein (GFP). Cells from the positive selection were grown in the absence of pAF and Cm, and then screened using fluorescence activated cell sorting (FACS) for a lack of fluorescence. For example, for the negative screen, cells were grown in GMML media containing 35 $\mu\text{g/ml}$ Kn, 25 $\mu\text{g/ml}$ Tet, and 0.002 % arabinose. FACS was carried out using a BDIS FACVantage TSO cell sorter with a Coherent Enterprise II ion laser. The excitation wavelength was 351 nm and emission was detected using a 575/25 nm bandpass filter. Collected cells were diluted into at least 10 volumes of LB, containing Tet and Kn, and grown to saturation.

[0296] The desired pAFRS was identified following two rounds of positive selection in liquid media, one round of negative screening, another round of positive selection in liquid media, and one round of positive selection on plates. The pAFRS enzyme contains five mutations relative to the wild type TyrRS (Y32T, E107T, D158P, I159L, and L162A). In the absence of pAF, the IC_{50} of cells expressing the selected pAFRS and reporter plasmid was 10 $\mu\text{g/ml}$ Cm on GMML minimal media plates. The IC_{50} was 120 $\mu\text{g/ml}$ Cm with 1 mM pAF. Thus, pAF is selectively suppressing the UAG codon.

[0297] **Example 8-Evolution of an Aminoacyl-tRNA Synthetase using fluorescence-activated cell sorting.**

[0298] A FACS based screening system was used to rapidly evolve three highly selective synthetase variants that accept amino-, isopropyl-, or allyl-containing tyrosine analogues. The system included a multipurpose reporter plasmid used for application of

WO 02/086075

PCT/US02/12635

both positive and negative selection pressure and for the facile and quantitative evaluation of synthetase activity. A chloramphenicol acetyl transferase (CAT) marker allowed positive selection for activity of the *M. jannaschii* tyrosyl-tRNA synthetase (TyrRS). A T7 polymerase/GFP reporter system allowed assessment of synthetase activity within cells grown in both the presence and absence of an unnatural amino acid. Fluorescence activated cell sorting (FACS) was used to screen against synthetase variants that accept natural amino acids, while visual and fluorimetric analyses were to assess synthetase activity qualitatively and quantitatively, respectively.

[0299] Design of an amplifiable fluorescence reporter system. Efforts to develop a versatile screening system for the assessment of synthetase activity in living cells initially arose out of a desire for a greater degree of control over the selective pressure applied to populations of synthetase variants, especially negative selective pressure. As the system was to be used to assess the activities of large numbers of synthetase variants, a reporter was sought that would be amenable to high-throughput screening. In addition, a reporter that would allow for facile qualitative and quantitative evaluation of synthetase activity was desired. To meet these requirements, a fluorescence-based screen was designed. The system was based on the synthetase-dependent production of GFPuv, a variant of the green fluorescent protein that has been optimized for expression in *E. coli* (see, Cramer, A., Whitehorn, E. A., Tate, E. & Stemmer, W. P., *Nature Biotechnol.* 1996, 14, 315-319). This fluorophore is amenable to use in FACS and fluorimetry, as well as visual inspection on plates and in liquid culture. The system was designed such that synthetase-dependent suppression of selector, e.g., amber nonsense codons would result in the production of a fluorescence signal. In order to maximize the sensitivity of the reporter, it was made amplifiable by placement of the amber codons within the gene for T7 RNA polymerase, which was designed to drive expression of the GFPuv reporter gene in analogy to other amplifiable intracellular reporter systems (see, Lorincz, M., Roederer, M., Diwu, Z., Herzenberg, L. A., Nolan, G. P. *Cytometry*, 1996, 24, 321-329; and Zlokarnik, G., Negulescu, P. A., Knapp, T. E., Mere, L., Burres, N., Feng, L., Whitney, M., Roemer, K. & Tsien, R. Y., *Science*, 1998, 279, 84-88). The T7 RNA polymerase gene was placed under control of the arabinose promoter in order to allow facile optimization of the production of the RNA transcript for amber codon-containing T7 RNA polymerase.

WO 02/086075

PCT/US02/12635

[0300] Optimization of the T7 RNA polymerase/GFPuv reporter system. A medium-copy reporter plasmid, pREP, was designed to express amber-containing T7 RNA polymerase variants under control of the arabinose promoter and the GFPuv gene under control of the T7 promoter (**Figure 17a**). A series of twelve T7 RNA polymerase variants, designed to optimize synthetase-dependent fluorescence enhancement (**Figure 17b**), were inserted into pREP to create plasmids pREP(1–12). All variants contained an N-terminal leader sequence of seven amino acids (MTMITVH) and 1–3 amber stop codons (TAG). Variants 1–3 contained one, two, and three amber stop codons, respectively, substituted for the original methionine at position one (M1), just downstream of the leader sequence. Variants 4–9 contained an amber codon substituted for D10, R96, Q107, A159, Q169, or Q232, respectively, which were predicted to be located in loop regions of the structure (*see*, Jeruzalmi, D. & Steitz, T. A., EMBO J., 1998, 17, 4101-4113). Variants 10–12 contained amber stop codons substituted at positions M1 and either Q107, A159, or Q232, respectively. Plasmid constructs were evaluated by fluorimetry and flow cytometry of live cells for fluorescence enhancement using a compatible plasmid containing the orthogonal glutaminyl-tRNA synthetase and Glutamine tRNA_{CUA} from *S. cerevisiae*. Plasmids pREP(1–12) were found to provide varying levels of synthetase-dependent fluorescence enhancement, with the best construct, pREP(10) exhibiting 220-fold greater fluorescence by fluorimetry (**Figure 17c**) and ~400-fold greater median fluorescence by cytometry (**Figure 17d**) in cells containing the wild type synthetase versus an inactive mutant. Substitution of a variety of functional groups at positions corresponding to the amber codons within pREP(10) demonstrate that position 107 within T7 RNA polymerase is highly permissive.

[0301] Construction of a multipurpose reporter plasmid. In order to construct a multipurpose plasmid to be used both for selecting and screening variants of a *M. jannaschii* TyrRS, plasmid pREP(10) was combined with plasmid pYC-J17 (*see*, Wang, L., Brock, A., Herberich, B. & Schultz, P. G., Science, 2001, 292, 498-500) to obtain pREP/YC-JYCUA (**Figure 25a**). Plasmid pREP/YC-JYCUA was assayed for function with a compatible plasmid expressing a variant of *M. jannaschii* TyrRS (pBK-mJYRS; Wang, L., Brock, A., Herberich, B. & Schultz, P. G., Science, 2001, 292, 498-500) selective for incorporating *O*-Methyl-Tyrosine (OMY). Cells containing pREP/YC-JYCUA and pBK-mJYRS, grown in the presence of OMY, exhibited a chloramphenicol (Cm) IC₅₀ value of 120 µg/µl, identical to that obtained using plasmid pYC-J17, and a

WO 02/086075

PCT/US02/12635

fluorescence enhancement of 330-fold for cells grown in the presence versus the absence of OMY, as measured by fluorimetry.

[0302] Evolution of the substrate specificity of the *M. jannaschii* tyrosyl-tRNA synthetase.

Results have shown that the amino acid side chain binding pocket of the *M. jannaschii* TyrRS can be evolved to selectively accommodate chemical groups other than the phenol side chain of tyrosine (see, Wang, L., Brock, A., Herberich, B. & Schultz, P. G., Science, 2001, 292, 498-500; Wang, L., Brock, A. & Schultz, P. G. J. Am. Chem. Soc. 2002, 124, 1836-1837). We sought to further explore the generality of unnatural amino acid accommodation by *M. jannaschii* TyrRS by challenging the enzyme to accept four new functionalities: *p*-Isopropyl-Phenylalanine (pIF), *p*-Amino-Phenylalanine (pAF), *p*-Carboxyl-Phenylalanine (pCF), or *O*-Allyl-Tyrosine (OAT) (Figure 25b). A library of *M. jannaschii* TyrRS variants containing randomizations at positions Y32, E107, D158, I159, and L162 (Wang, L., Brock, A., Herberich, B. & Schultz, P. G., Science, 2001, 292, 498-500), residues thought to form the binding pocket for the *para* position of the tyrosyl ring, was introduced into cells containing plasmid PREP/YC-JYCUA. These cells, encompassing a library diversity of $\sim 10^9$, were used to begin four evolution experiments to identify synthetase variants selective for pIF, pAF, pCF, or OAT (Figure 25b). Two cycles of positive selection were carried out by allowing the cell cultures to grow to saturation in the presence of Cm and one of the four unnatural amino acids. Cell aliquots were removed following the second cycle of positive selection and used to inoculate a new culture containing no added amino acid or Cm, and the culture was again allowed to grow to saturation. At this point, cells that fluoresce are likely to contain synthetase variants that can accept one of the 20 natural amino acids. Approximately 10^8 cells from each line were subjected to negative screening using FACS in order to eliminate natural amino acid-accepting synthetase variants. The non-fluorescent cells were collected and amplified through growth to saturation. These amplified cells were used to inoculate a new culture for a final cycle of positive selection in liquid culture containing unnatural amino acid and Cm. Following growth to saturation, each population of cells was plated on media containing 0, 30, 60, or 100 $\mu\text{g/mL}$ Cm and either 0 or 1 mM of the appropriate unnatural amino acid.

[0303] Identification and characterization of evolved synthetase variants. Cm plates supplemented with pIF, pAF, and OAT produced 10-100-fold greater numbers of

WO 02/086075

PCT/US02/12635

fluorescent colonies than plates containing no added amino acid. In contrast, plates for the *pCF* population produced the same number of fluorescent colonies with or without addition of *pCF*. The ten largest fluorescent colonies were picked for each of the *pIF*, *pAF*, and *OAT* populations from unnatural amino acid-containing plates and grown to saturation in liquid media with or without added unnatural amino acid. A qualitative assessment of fluorescence production was made visually with the use of a hand-held long-wavelength ultraviolet lamp (**Figure 23a**).

[0304] Synthetase variants corresponding to clones producing significant differences in fluorescence were sequenced. All ten clones from the *pIF* and *pAF* populations had identical sequences, while three different clones were identified from the *OAT* population. Amino acid changes occurred within the five randomized sites in all clones, with the exception of two additional substitutions within the *pIF*-tRNA synthetase (*pIF*-RS) variant. The activities of the different clones were quantitatively assessed. Fluorescence was measured fluorimetrically for cells grown in liquid culture in the presence or absence of unnatural amino acid (**Figure 23b**). The Cm IC_{50} s were determined by plating the cells on varying concentrations of Cm in the presence or absence of unnatural amino acid (**Figure 23c**).

[0305] A myoglobin gene containing an amber codon in the fourth position was used to assess the production of unnatural amino acid-containing protein. The gene was expressed in cells, using the *pIF*-RS, *pAF*-RS, or *OMY*-RS variant, respectively, in either the presence or absence of *pIF*, *pAF*, or *OAT* (**Figure 23d**). Protein yields were comparable for all three variants, ranging from 1–2 milligrams of protein per liter of unnatural amino acid-containing cell culture. In contrast, protein production was virtually undetectable in cultures grown in the absence of unnatural amino acid. Proteins were analyzed by electrospray mass spectrometry, giving masses of 18457.40 ± 0.81 (18457.28 expected) for the *pIF*-containing protein, and 18430.30 ± 0.27 (18430.21 expected) for the *pAF*-containing protein. Activity measurements obtained using the Cm IC_{50} , fluorimetry, and protein expression analyses correlated well, however the activity of the *pIF*-RS appears to be somewhat underestimated by fluorimetry. As compared to other assays, the disproportionately low fluorimetry measurement for the *pIF*-RS variant, suggests that T7 RNA polymerase may be partially destabilized upon incorporation of the *pIF* analogue,

WO 02/086075

PCT/US02/12635

despite the apparent permissivity of the amber positions within the reporter (*see, Figure 17c*).

[0306] Utility of the multipurpose reporter system. The reporter system described here allows the use of a single multipurpose plasmid for both positive selection and negative screening, obviating the need to shuttle plasmids between alternating rounds of positive and negative selection. A total of only three rounds of positive selection and one round of negative screening were required to enable the identification of synthetase variants that selectively accept desired unnatural amino acids. These features allow evolution experiments to be carried out in a matter of days. The screening system can be used to readily identify active synthetase variants using agar plates containing unnatural amino acid and to individually assay the amino acid specificity of the variants.

[0307] As described above, the T7 RNA polymerase/GFP system can be used to quantitatively compare the activities of synthetase variants. The availability of the three OAT-RS clones described here and a different OAT-RS clone derived independently from the same library using a positive/negative selection based on CAT and barnase allows the possibility of comparing the two different evolution systems in terms of the synthetase variants resulting from each. This analysis reveals that the three clones derived from positive selection and negative screening exhibit slightly lower levels of fluorescence in the presence of OAT, but ~10-fold lower background levels in the absence of the unnatural amino acid. The fluorescence enhancement for cells grown in the presence versus the absence of the unnatural amino acid is thus about 6-fold higher for cells expressing OAT-RS(1) from selection and screening than for cells expressing the OAT-RS clone derived from positive/negative selection using barnase. Although it is not clear whether this example is representative, these data suggest that the T7 RNA polymerase/GFP system may allow more stringency in selecting against synthetase variants that are promiscuous towards natural amino acid substrates. However, the fluorescence enhancement for cells grown in the presence versus the absence of an unnatural amino acid is expected to represent a lower limit for the fidelity of unnatural amino acid incorporation, as competition of unnatural amino acids for being bound by an evolved synthetase variant would reduce binding of natural amino acids. Moreover, although high fidelity is clearly desirable, there is likely to be a trade-off between fidelity and overall synthetase activity, which may depend on the desired application.

WO 02/086075

PCT/US02/12635

[0308] **Generality of aminoacyl tRNA synthetase evolution.** Previous results and those presented here demonstrate that the amino acid side chain binding pocket of the *M. jannaschii* TyrRS is quite malleable. The enzyme can be evolved to accommodate a variety of functionalities in place of the phenol side chain of tyrosine and can do so with high selectivity. In this application it was demonstrated that enzyme can be evolved to accommodate an amine, isopropyl, or allyl ether functionality at the *para* position of the tyrosine ring, instead of hydroxyl. It was not possible to identify an enzyme variant that could accept the *pCF* unnatural amino acid. A second attempt to evolve a synthetase to accept the *pCF* amino acid was also unsuccessful. Using LC/MS analysis, *pCF* could not be detected upon toluenization of *E. coli* cells grown in the presence of the unnatural amino acid, suggesting that *pCF* is not transported into cells or that it is metabolized upon entry.

[0309] Of the three successful evolution experiments described here, only the evolution of the OAT-RS resulted in the identification of more than one active clone. The OAT-RS evolution was also the experiment that produced the most active synthetase variant. These results suggest that some amino acid specificities may be easier to select for than others. This could be due, in part, to the relative difficulty of selectively recognizing different unnatural amino acids in the context of the 20 natural amino acids. It may be, for example, that *pAF*, due to its structural and electronic similarities to tyrosine, is more difficult to selectively recognize than OAT. This would explain why a greater number of OAT-RS clones were identified than *pAF*-RS clones and why the *pAF*-RS clone is less active than the best OAT-RS clone.

[0310] **Plasmid Construction.** Plasmid pREP (Figure 17a) was constructed by insertion of a *Bam*HI/*Apa*LI overlap PCR fragment containing the T7 RNA polymerase gene upstream of an *rrnB* transcription termination region, followed by an *Apa*LI/*Ahd*I overlap PCR fragment containing the *araC* gene and *ara* promoter region from the pBAD/Myc-His A plasmid (Invitrogen; for transcriptional control of the T7 RNA polymerase gene) and the GFPuv gene (Clontech; upstream of the T7 terminator region and downstream of the T7 promoter) between the *Ahd*I/*Bam*HI sites of plasmid pACYC177 (New England Biolabs). Plasmids pREP(1-12) were constructed by replacement of an *Hpa*I/*Apa*LI fragment of T7 RNA polymerase with overlap PCR fragments containing amber mutations at the positions described. Plasmid pREP/YC-

WO 02/086075

PCT/US02/12635

JYCUIA was constructed by ligation of an *AfeI/SacII* fragment from pREP(10) and an *EatI*(blunted)/*SacII* fragment from pYC-J17 (Wang, L., Brock, A., Herberich, B. & Schultz, P. G., *Science*, 2001, 292, 498-500). The desired construct was identified following transformation into cells containing plasmid pQ screening for fluorescence.

- 5 [0311] Plasmid pQ was constructed by triple ligation of a *AatII/SalI* overlap PCR fragment containing the *ScQRS* downstream of the *lac* promoter region and upstream of the *E. coli* QRS termination region, a *SalI/Aval* overlap PCR fragment containing the *S. cerevisiae* tRNA(CUA)^{Gln} downstream of the *lpp* promoter region and upstream of an *rrnC* termination region, and the *Aval/AatII* fragment of pBR322 (New England Biolabs).
- 10 Plasmid pQD was constructed by replacement of pQ fragment between *Bam*III and *Bgl*II with a *Bam*HI/*Bgl*II fragment of the *ScQRS* (D291A) mutant.

- [0312] Plasmid pBAD/JYAMB-4TAG was constructed by insertion of a PCR fragment of the S4Amber mutant of myoglobin, containing a C-terminal 6His-tag, into the pBAD/YC-JYCUIA plasmid, a hybrid of plasmid pYC-J17 (Wang, L., Brock, A., Herberich, B. & Schultz, P. G., *Science*, 2001, 292, 498-500) and pBAD/Myc-His A (Invitrogen) containing the gene for *MjYtRNA*_{CUA}, and the pBAD promoter and cloning regions for heterologous expression of an inserted gene.

- [0313] **Fluorimetric and cytometric analyses.** Single colonies containing desired plasmids were used to inoculate 2-mL GMMML cultures containing the appropriate antibiotics, 0.002% Arabinose, and an appropriate unnatural amino acid, if desired.
- 20 Cultures were grown to saturation and cells (200 μ L) were pelleted and resuspended in 1 mL phosphate-buffered saline (PBS). Cell concentrations were analyzed by absorbance at 600 nm and fluorescence levels were measured at 505 nm with excitation at 396 nm using a FluoroMax-2 fluorimeter. Cells suspended in PBS were analyzed cytometrically. To
- 25 evaluate the permissivity of the amber positions within the T7 polymerase gene of pREP(10), the reporter plasmid was transformed into a panel of suppressor strains, which were subsequently analyzed fluorimetrically.

- [0314] **Evolution of aminoacyl-tRNA synthetase variants.** *M. jannaschii* TyrRS variants randomized at positions Y32, E107, D158, I159, and L162 (Wang, L., Brock, A., Herberich, B. & Schultz, P. G., *Science*, 2001, 292, 498-500) were transformed into
- 30 DH10B *E. coli* cells (Life Technologies) containing pREP/YC-JYCUIA to generate a library with a diversity of $\sim 10^9$. Transformants were allowed to recover in SOC medium

WO 02/086075

PCT/US02/12635

for 60 min at 37°C, and were grown to saturation in LB medium. To begin an initial positive selection, 2 mL of library culture, pelleted and resuspended in GMML medium, was used to inoculate 500 mL of GMML containing 25 µg/mL Tetracycline (Tet), 35 µg/mL Kanamycin (Kn), and 1 mM *pIF*, *pAF*, *pCF*, or *OAY*. After incubation for 3 hours at 37°C, Cm was added to a final concentration of 75 µg/mL and cells were grown to saturation (~48 hours). For the second positive selection, a 100-mL GMML culture containing Tet, Kn, 75 µg/mL Cm, and 1 mM *pIF*, *pAF*, *pCF*, or *OAY* was inoculated with cells from the initial positive selection (500 µL) and grown to saturation at 37°C (~24–36 hours). In preparation for negative screening, a 25-mL GMML culture containing Tet, Kn, and 0.02% arabinose (Ara) was inoculated with cells from the second positive selection (100 µL, pelleted and resuspended in GMML) and grown to saturation at 37°C (~24 hours). Ara-induced cells grown in the absence of unnatural amino acids (1 mL) were pelleted and resuspended in 3 mL of phosphate-buffered saline (PBS). Cells were sorted for lack of expression of GFPuv using a BDIS FACVantage TSO cell sorter with a Coherent Enterprise II ion laser with excitation at 351 nm and emissions detected using a 575/25 nm bandpass filter. Collected cells were diluted in at least 10 volumes of LB, containing Tet and Kn, and grown to saturation. To begin the third round of positive selection, 100 µL of cells from the negative screen were pelleted, resuspended in GMML, and used to inoculate 25 mL of GMML containing Tet, Kn, and 1 mM *pIF*, *pAF*, *pCF*, or *OAY*. After incubation for 3 hours at 37°C, Cm was added to a final concentration of 75 µg/mL and cells were grown to saturation (~24 hours). Following the third positive selection, cells were plated on GMML/agar containing Tet, Kn, 0.002% Ara, 0, 75, or 100 µg/mL Cm, and 0 or 1 mM *pIF*, *pAF*, *pCF*, or *OAY*, and grown for 48 hours at 37°C.

[0315] Expression and characterization of unnatural amino acid-containing proteins. DH10B cells co-transformed with pBAD/JYAMB-4TAG and the appropriate pBK plasmid were used to inoculate a 100-mL GMML starter culture containing Kn and Tet, which was grown to saturation. A 500-mL culture containing Kn, Tet, 0.002% Ara, 5 µM FeCl₃, and the desired unnatural amino acid (or none) was inoculated with 50 mL of the starter culture and grown to saturation (~18 hours). Cultures were pelleted, sonicated, and the myoglobin protein isolated according to the protocol of the QiaExpressionist (Qiagen) His-tag purification kit. Proteins were analyzed electrophoretically on a 12–20% gradient SDS polyacrylamide gel and by electrospray mass spectrometry.

WO 02/086075

PCT/US02/12635

[0316] Example 9- Orthogonal tRNA/Threonyl-tRNA synthetase Pair.

[0317] This example illustrates the generation of an orthogonal tRNA/Threonyl-tRNA synthetase pair. **Figure 27** illustrates a threonyl-tRNA synthetase from *Thermus thermophilus*. This synthetase has two N-terminal editing domains, a catalytic domain and a C-terminal anticodon binding domain (659 amino acids). To generate the orthogonal synthetase based on the *T. thermophilus* synthetase, the editing domain(s), N1 or N1 and N2 was deleted from the synthetase to generate an N-truncated *T. thermophilus* ThrRS (475 amino acids). This synthetase has the same catalytic activity but lacks the proofreading activity. The N-truncated synthetase was screened for activity. The N-truncated synthetase did not aminoacylate *Escherichia coli* tRNA.

[0318] Because, the *T. thermophilus* tRNA^{Thr} was found to be a substrate for *Escherichia coli* Threonyl-tRNA synthetase, the *T. thermophilus* tRNA^{Thr} was mutated in order to generate an orthogonal pair. **Figure 28** illustrates the mutations made in the tRNA. Specifically, C2G71 was mutated to A2U71. In vitro charging experiments demonstrate that this mutant is not a substrate for the *E. coli* Threonyl-tRNA synthetase but is a good substrate for the *T. thermophilus* Threonyl-tRNA synthetase. Another mutant was also constructed, which included the following mutations: C2G71→A2U71 and G34G35U36→C34G35U36 in order to generate an amber suppressor tRNA. Other mutant tRNAs with modified anticodon loops in addition to C2G71→A2U71 were also generated to suppress three and four base codons such as TGA, ACCA, ACAA, AGGA, CCCT, TAGA, and CTAG. All these tRNAs were not as good as substrate as the wild type tRNA^{Thr} (with A2U71) but can be improved by mutating the anticodon binding site of the *T. thermophilus* Threonyl-tRNA synthetase.

[0319] Example 10- Sequences of exemplary O-tRNAs and O-RSs.

[0320] Exemplary O-tRNAs comprise a nucleic acid comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO:1-3 and/or a complementary polynucleotide sequence thereof. See, Table 5, Appendix 1. Similarly, example O-RS include polypeptides selected from the group consisting of: a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 35-66 and a polypeptide encoded by a nucleic acid comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO:4-34 and a complementary polynucleotide sequence thereof.

WO 02/086075

PCT/US02/12635

[0321] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, 5 and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

WO 02/086075

PCT/US02/12635

CLAIMSWhat is claimed is:

1. A composition comprising an orthogonal aminoacyl-tRNA synthetase (O-RS), wherein the O-RS preferentially aminoacylates an orthogonal tRNA
5 (O-tRNA) with an unnatural amino acid.
2. The composition of claim 1, wherein the O-RS comprises a nucleic acid comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO: 4-34 and a complementary polynucleotide sequence thereof.
3. The composition of claim 1, wherein the O-RS aminoacylates the
10 O-tRNA with the unnatural amino acid in vivo.
4. The composition of claim 1, wherein the unnatural amino acid is selected from the group consisting of: an O-methyl-L-tyrosine, an L-3-(2-naphthyl)alanine, a 3-methyl-phenylalanine, an O-4-allyl-L-tyrosine, a 4-propyl-L-tyrosine, a tri-O-acetyl-GlcNAc β -serine, an L-Dopa, a fluorinated phenylalanine, an isopropyl-L-phenylalanine, a
15 *p*-azido-L-phenylalanine, a *p*-acyl-L-phenylalanine, a *p*-benzoyl-L-phenylalanine, an L-phosphoserine, a phosphoserine, a phosphotyrosine, a *p*-iodo-phenylalanine, a *p*-bromophenylalanine, a *p*-amino-L-phenylalanine, and an isopropyl-L-phenylalanine.
5. The composition of claim 1, wherein the unnatural amino acid is selected from the group consisting of: an unnatural analogue of a tyrosine amino acid; an
20 unnatural analogue of a glutamine amino acid; an unnatural analogue of a phenylalanine amino acid; an unnatural analogue of a serine amino acid; an unnatural analogue of a threonine amino acid; an alkyl, aryl, acyl, azido, cyano, halo, hydrazine, hydrazide, hydroxyl, alkenyl, alkynyl, ether, thiol, sulfonyl, seleno, ester, thioacid, borate, boronate, phospho, phosphono, phosphine, heterocyclic, enone, imine, aldehyde, hydroxylamine, keto, or amino substituted amino acid, or any combination thereof; an amino acid with a
25 photoactivatable cross-linker; a spin-labeled amino acid; a fluorescent amino acid; an amino acid with a novel functional group; an amino acid that covalently or noncovalently interacts with another molecule; a metal binding amino acid; a metal-containing amino acid; a radioactive amino acid; a photocaged amino acid, a photoisomerizable amino acid;

WO 02/086075

PCT/US02/12635

a biotin or biotin-analogue containing amino acid; a glycosylated or carbohydrate modified amino acid; a keto containing amino acid; an amino acid comprising polyethylene glycol; an amino acid comprising polyether; a heavy atom substituted amino acid; a chemically cleavable or photocleavable amino acid; an amino acid with an elongated side chain; an amino acid containing a toxic group; a sugar substituted amino acid; a sugar substituted serine; a carbon-linked sugar-containing amino acid; a redox-active amino acid; an α -hydroxy containing acid; an amino thio acid containing amino acid; an α,α disubstituted amino acid; a β -amino acid; and a cyclic amino acid other than proline.

10 6. The composition of claim 1, wherein the O-RS has one or more improved or enhanced enzymatic properties, selected from the groups consisting of: K_m and K_{cat} , for the unnatural amino acid as compared to a natural amino acid.

 7. A polypeptide comprising an amino acid sequence encoded by a coding polynucleotide sequence, the coding polynucleotide sequence selected from the group consisting of:

- 15 a) a coding polynucleotide sequence selected from SEQ ID NO:4-34;
- b) a coding polynucleotide sequence that encodes a polypeptide selected from SEQ ID NO:35-66;
- 20 c) a polynucleotide sequence which hybridizes under highly stringent conditions over substantially an entire length of a polynucleotide sequence of (a) or (b); and,
- d) a complementary sequence of (a), (b), or (c).

 8. The polypeptide of claim 7, wherein the encoded polypeptide encodes an orthogonal aminoacyl tRNA synthetase.

25 9. A polypeptide comprising an amino acid sequence selected from SEQ ID NO:35-66.

 10. A nucleic acid comprising: a polynucleotide sequence selected from the group consisting of:

WO 02/086075

PCT/US02/12635

- a) a polynucleotide sequence selected from SEQ ID NO:1 to SEQ ID NO:3, or a complementary polynucleotide sequence thereof; and,
- b) a polynucleotide sequence which hybridizes under highly stringent conditions over substantially the entire length of polynucleotide sequence (a).
- 5
11. The nucleic acid of claim 10, wherein the polynucleotide sequence comprises an orthogonal tRNA.
12. The nucleic acid of claim 10, wherein the polynucleotide sequence forms a complementary pair with an orthogonal aminoacyl-tRNA synthetase.
- 10
13. The nucleic acid of claim 12, wherein the polypeptide sequence of the orthogonal aminoacyl-tRNA synthetase is selected from the group consisting of SEQ ID:35 to SEQ ID NO:66.
14. A composition comprising an orthogonal tRNA (O-tRNA), wherein the O-tRNA recognizes a selector codon and wherein the O-tRNA is preferentially
- 15 aminoacylated with an unnatural amino acid by an orthogonal aminoacyl-tRNA synthetase.
15. The composition of claim 14, wherein the O-tRNA comprises a nucleic acid comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO: 1-3 and a complementary polynucleotide sequence thereof.
- 20
16. The composition of claim 14, wherein the selector codon comprises a unique three base codon, a nonsense codon, a rare codon, an unnatural codon, or at least a four base codon.
17. The composition of claim 14, further comprising the orthogonal aminoacyl-tRNA synthetase (O-RS).
- 25
18. The composition of claim 17, wherein the O-tRNA and the O-RS are complementary.

WO 02/086075

PCT/US02/12635

19. The composition of claim 17, wherein the composition comprises a mutRNATyr-mutTyrRS pair.
20. The composition of claim 19, wherein the composition comprises a mutRNATyr-SS12TyrRS pair.
- 5 21. The composition of claim 17, wherein the composition comprises a mutRNALeu-mutLeuRS pair.
22. The composition of claim 17, wherein the composition comprises a mutRNAThr-mutThrRS pair.
23. The composition of claim 17, wherein the composition comprises a
10 mutRNAGlu-mutGluRS pair
24. The composition of claim 17, wherein the O-tRNA and the O-RS are derived by mutation of a naturally occurring tRNA and an RS from at least one organism, wherein the at least one organism is a prokaryotic organism.
25. The composition of claim 24, wherein the at least one organism is
15 selected from the group consisting of: *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, and a *Halobacterium*.
26. The composition of claim 17, wherein the O-tRNA and the O-RS are derived by mutation of a naturally occurring tRNA and RS from at least one organism, wherein the at least one organism is a eukaryotic organism.
- 20 27. The composition of claim 26, wherein the at least one organism is selected from the group consisting of: yeasts, mammals, fungi, insects, plants and protists.
28. The composition of claim 17, wherein the O-tRNA is derived by mutation of a naturally occurring tRNA from a first organism and the O-RS is derived by mutation of a naturally occurring RS from a second organism.
- 25 29. The composition of claim 17, wherein the O-tRNA and the O-RS are isolated from at least one organism, wherein the at least one organism is a prokaryotic organism.

WO 02/086075

PCT/US02/12635

30. The composition of claim 29, wherein the at least one organism is selected from the group consisting of: *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, and a *Halobacterium*.
31. The composition of claim 17, wherein the O-tRNA and the O-RS
5 are isolated from at least one organism, wherein the at least one organism is a eukaryotic organism.
32. The composition of claim 31, wherein the at least one organism is selected from the group consisting of: yeasts, mammals, fungi, insects, plants and protists.
33. The composition of claim 17, wherein the O-tRNA is isolated from
10 a first organism and the O-RS is isolated from a second organism.
34. The composition of claim 17, wherein one or more of the O-tRNA and the O-RS is isolated from one or more library, which one or more library comprises an O-tRNA or an O-RS from one or more organism.
35. The composition of claim 34, wherein the one or more organism
15 comprises a prokaryote or a eukaryote.
36. The composition of claim 17, wherein the composition is in a cell.
37. The composition of claim 17, wherein the composition comprises an in vitro translation system.
38. A method for producing at least one recombinant orthogonal
20 aminoacyl-tRNA synthetase (O-RS), the method comprising:
- (a) generating a library of RSs derived from at least one aminoacyl-tRNA synthetase (RS) from a first organism;
- (b) selecting or screening the library of RSs for members that aminoacylate an orthogonal tRNA (O-tRNA) in the presence of an unnatural amino acid and a natural amino acid, thereby providing a pool of active RSs; and,
25
- (c) selecting or screening the pool for active RSs that preferentially aminoacylate the O-tRNA in the absence of the unnatural amino acid, thereby providing

WO 02/086075

PCT/US02/12635

the at least one recombinant O-RS; wherein the at least one recombinant O-RS preferentially aminoacylates the O-tRNA with the unnatural amino acid.

39. The method of claim 38, wherein the first organism is a prokaryotic organism.
- 5 40. The method of claim 38, wherein the first organism is selected from the group consisting of: *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, and a *Halobacterium*.
41. The method of claim 38, wherein the at least one RS is an inactive RS.
- 10 42. The method of claim 41, where the inactive RS is generated by mutating an active RS.
43. The method of claim 42, wherein the inactive RS comprises at least about 1, at least about 2, at least about 3, at least about 4, about 5, at least about 6, or at least about 10 or more amino acids mutated to alanine amino acids.
- 15 44. The method of claim 38, wherein the library of RSs comprises a library of mutant RS.
45. The method of claim 44, wherein step (a) comprises random mutation.
46. The method of claim 38, wherein the step (b) comprises:
- 20 introducing a positive selection or screening marker and the library of RSs into a plurality of cells, wherein the positive selection or screening marker comprises at least one selector codon;
- growing the plurality of cells in the presence of a selection or screening agent;
- 25 identifying cells that survive or show a specific screening response in the presence of the selection or screening agent by suppressing the at least one selector codon

WO 02/086075

PCT/US02/12635

in the positive selection or screening marker, thereby providing a subset of positively selected or screened cells that contains the pool of active RSs.

47. The method of claim 46, wherein the selection or screening agent concentration is varied.

5 48. The method of claim 46, wherein the at least one selector codon comprises an amber codon, an ochre codon, or an opal stop codon.

49. The method of claim 46, wherein the positive selection marker is a chloramphenicol acetyltransferase (CAT) gene and wherein the at least one selector codon is an amber stop codon in the CAT gene.

10 50. The method of claim 46, wherein the positive selection marker is a β -lactamase gene and wherein the at least one selector codon is an amber stop codon in the β -lactamase gene.

51. The method of claim 46, wherein the positive screening marker comprises a fluorescent or luminescent screening marker.

15 52. The method of claim 46, wherein the positive screening marker comprises an affinity based screening marker.

53. The method of claim 38, wherein the step (c) comprises:

introducing a negative selection or screening marker with the pool of active RSs from step (b) into a plurality of cells of a second organism, wherein the negative

20 selection or screening marker gene comprises at least one selector codon;

identifying cells that survive or show a specific screening response in a 1st media supplemented with the unnatural amino acid and a selection or screening agent, but fail to survive or to show the specific screening response in a 2nd media not supplemented with the unnatural amino acid and the selection or screening agent, thereby providing

25 surviving or screened cells with the at least one recombinant O-RS.

54. The method of claim 53, wherein concentration of the selection or screening agent is varied.

WO 02/086075

PCT/US02/12635

55. The method of claim 53, wherein the first and the second organism are different.
56. The method of claim 53, wherein the first and second organism comprise a prokaryote, a eukaryote, a mammal, an *Escherichia coli*, a fungi, a yeast, an archaeobacterium, a eubacterium, a plant, an insect, or a protist.
57. The method of claim 53, wherein the negative selection marker comprises a chloramphenicol acetyltransferase (CAT) gene comprising at least one selector codon.
58. The method of claim 53, wherein the screening marker comprises a fluorescent or luminescent screening marker.
59. The method of claim 53, wherein the screening marker comprises an affinity based screening marker.
60. The method of claim 38, wherein step (c) comprises:
isolating the pool of active RSs from step (b);
introducing a negative selection or screening marker, wherein the negative selection or screening marker gene comprises at least one selector codon, and the pool of active RSs into a plurality of cells of a second organism;
identifying cells that survive or show a specific screening response in a 1st media not supplemented with the unnatural amino acid, but fail to survive or to show the specific screening response in a 2nd media supplemented with the unnatural amino acid, thereby providing surviving or screened cells with the at least one recombinant O-RS, wherein the at least one recombinant O-RS is specific for the unnatural amino acid.
61. The method of claim 60, wherein the at least one selector codon comprises two or more selector codons.
62. The method of claim 60, wherein the first and the second organism are different.

WO 02/086075

PCT/US02/12635

63. The method of claim 60, wherein the first and second organism comprise a prokaryote, a eukaryote, a mammal, an *Escherichia coli*, a fungi, a yeast, an archaeobacteria, a eubacteria, a plant, an insect, or a protist.
64. The method of claim 60, wherein the negative selection marker
5 comprises a ribonuclease barnase gene comprising at least one selector codon.
65. The method of claim 60, wherein the screening marker comprises a fluorescent or luminescent screening marker.
66. The method of claim 60, wherein the screening marker comprises an affinity based screening marker.
- 10 67. The method of claim 38, wherein step (b), (c) or both steps (b) and (c), comprise varying a selection or screening stringency.
68. The method of claim 38, further comprising:
(d) isolating the at least one recombinant O-RS;
(e) generating a second set of O-RS derived from the at least one
15 recombinant O-RS, wherein the second set of O-RS comprises a set of mutated O-RS; and,
(f) repeating steps (b) and (c) until a mutated O-RS is obtained that comprises an ability to preferentially aminoacylate the O-tRNA.
69. The method of claim 68, further comprising repeating steps (d)-(f) at least about two times.
- 20 70. The method of claim 68, wherein step (e) comprises random mutagenesis, site-specific mutagenesis, recombination or any combination thereof.
71. The method of claim 38, wherein step (b), (c) or both steps (b) and (c) comprise using a reporter, wherein the reporter is detected by fluorescence-activated cell sorting (FACS) or wherein the reporter is detected by luminescence.

WO 02/086075

PCT/US02/12635

72. The method of claim 38, wherein the step (b), (c) or both steps (b) and (c) comprise using a reporter, wherein the reporter is displayed on a cell surface or on a phage display.

73. The method of claim 38, wherein the unnatural amino acid is
 5 selected from the group consisting of: an O-methyl-L-tyrosine, an L-3-(2-naphthyl)alanine, a 3-methyl-phenylalanine, an O-4-allyl-L-tyrosine, a 4-propyl-L-tyrosine, a tri-O-acetyl-GlcNAc β -serine, an L-Dopa, a fluorinated phenylalanine, an isopropyl-L-phenylalanine, a *p*-azido-L-phenylalanine, a *p*-acyl-L-phenylalanine, a *p*-benzoyl-L-phenylalanine, an L-phosphoserine, a phosphoserine, a phosphotyrosine, a *p*-iodo-phenylalanine, a *p*-bromophenylalanine, a *p*-amino-L-phenylalanine, and an isopropyl-L-phenylalanine.
 10

74. The at least one recombinant O-RS produced by the method of claim 38.

75. A method for producing a recombinant orthogonal tRNA (O-tRNA), the method comprising:

15 (a) generating a library of tRNAs derived from at least one tRNA from a first organism;

(b) selecting or screening the library for tRNAs that are aminoacylated by an aminoacyl-tRNA synthetase (RS) from a second organism in the absence of a RS from the first organism, thereby providing a pool of tRNAs; and,

20 (c) selecting or screening the pool of tRNAs for members that are aminoacylated by an introduced orthogonal RS (O-RS), thereby providing at least one recombinant O-tRNA; wherein the at least one recombinant O-tRNA recognizes at least one selector codon and is not efficiently recognized by the RS from the second organism and is preferentially aminoacylated by the O-RS.

25 76. The method of claim 75, wherein the at least one tRNA is a suppressor tRNA.

WO 02/086075

PCT/US02/12635

77. The method of claim 75, wherein the selector codon comprises a unique three base codon of natural or unnatural bases, a nonsense codon, a rare codon, an unnatural codon or at least a four base codon.
78. The method of claim 75, wherein the selector codon comprises an
5 amber codon, an ochre codon, or an opal stop codon.
79. The method of claim 75, wherein the at least one recombinant O-tRNA possesses an improvement of orthogonality.
80. The method of claim 75, wherein the first organism is a prokaryotic organism.
- 10 81. The method of claim 80, wherein the prokaryotic organism is selected from the group consisting of: *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, and a *Halobacterium*.
82. The method of claim 75, wherein the first and second organism
15 comprise a prokaryote, a eukaryote, a mammal, an *Escherichia coli*, a fungi, a yeast, an archaeobacteria, a eubacteria, a plant, an insect, or a protist.
83. The method of claim 75, wherein the first and the second organism are different.
84. The method of claim 75, wherein the recombinant tRNA is aminoacylated by an unnatural amino acid, wherein the unnatural amino acid is
20 biosynthesized in vivo either naturally or through genetic manipulation.
85. The method of claim 75, wherein the unnatural amino acid is added to a growth media of one or more of the first or second organism.
86. The method of claim 75, wherein the step (b) comprises:
25 introducing a toxic marker gene or a gene that leads to the production of a toxic or static agent or a gene essential to the organism, wherein the marker gene comprises the at least one selector codon and the library of mutant tRNAs into a plurality of cells from the second organism; and,

WO 02/086075

PCT/US02/12635

selecting surviving cells, wherein the surviving cells contain the pool of tRNAs comprising at least one orthogonal tRNA or nonfunctional tRNA.

87. The method of claim 86, wherein selecting surviving cells comprises a comparison ratio cell density assay.

5 88. The method of claim 86, wherein the selector codon comprises two or more selector codons.

89. The method of claim 86, wherein the marker gene is a ribonuclease barnase gene, wherein the ribonuclease barnase gene comprises at least one amber codon.

10 90. The method of claim 89, wherein the ribonuclease barnase gene comprises two or more amber codons.

91. The method of claim 75, wherein the step (c) comprises:

15 introducing a positive selection or screening marker gene, wherein the positive marker gene comprises a drug resistance gene or a gene essential to the organism, or a gene that leads to detoxification of a toxic agent, comprising at least one of the selector codons, the O-RS, and the pool of tRNAs into a plurality of cells from the second organism; and,

20 identifying surviving or screened cells grown in the presence of a selection or screening agent, thereby providing a pool of cells possessing the at least one recombinant tRNA, wherein the at least recombinant tRNA is aminoacylated by the O-RS and inserts an amino acid into a translation product encoded by the positive marker gene, in response to the at least one selector codons.

92. The method of claim 91, wherein the selection agent is an antibiotic.

93. The method of claim 91, wherein concentration of the selection or screening agent is varied.

25 94. The method of claim 91, wherein the drug resistance gene is a β -lactamase gene.

WO 02/086075

PCT/US02/12635

95. The method of claim 94, wherein the β -lactamase gene comprises at least one amber stop codon.

96. The at least one recombinant O-tRNA produced by the method of claim 75.

5 97. A method for producing at least one specific O-tRNA/O-RS pair, the method comprising:

(a) generating a library of tRNAs derived from at least one tRNA from a first organism;

(b) negatively selecting or screening the library for tRNAs that are
10 aminoacylated by an aminoacyl-tRNA synthetase (RS) from a second organism in the absence of a RS from the first organism, thereby providing a pool of tRNAs;

(c) selecting or screening the pool of tRNAs for members that are aminoacylated by an introduced orthogonal RS (O-RS), thereby providing at least one recombinant O-tRNA; wherein the at least one recombinant O-tRNA recognizes a selector
15 codon and is not efficiently recognized by the RS from the second organism and is preferentially aminoacylated by the O-RS;

(d) generating a library of RSs derived from at least one aminoacyl-tRNA synthetase (RS) from a third organism;

(e) selecting or screening the library of RSs for members that preferentially
20 aminoacylate the at least one recombinant O-tRNA in the presence of an unnatural amino acid and a natural amino acid, thereby providing a pool of active RSs; and,

(f) negatively selecting or screening the pool for active RSs that preferentially aminoacylate the at least one recombinant O-tRNA in the absence of the unnatural amino acid, thereby providing the at least one specific O-tRNA/O-RS pair,
25 wherein the at least one specific O-tRNA/O-RS pair comprises at least one recombinant O-RS that is specific for the unnatural amino acid and the at least one recombinant O-tRNA.

WO 02/086075

PCT/US02/12635

98. The specific O-tRNA/O-RS pair produced by the method of claim 97.
99. The method of claim 97, wherein the at least one specific O-tRNA/O-RS pair is a mutRNATyr-mutTyrRS pair.
- 5 100. The method of claim 97, wherein the first and the third organism are same.
101. The method of claim 97, wherein the first organism and the third organism are *Methanococcus jannaschii*.
102. A method for identifying an orthogonal tRNA-tRNA synthetase pair
10 for use in an in vivo translation system of a second organism, the method comprising:
introducing a marker gene, a tRNA and an aminoacyl-tRNA synthetase (RS) isolated or derived from a first organism into a first set of cells from the second organism;
introducing the marker gene and the tRNA into a duplicate cell set from the
15 second organism; and
selecting or screening for surviving cells or for cells showing a specific screening response in the first set that fail to survive or show said response in the duplicate cell set, wherein the first set and the duplicate cell set are grown in the presence of a selection or screening agent, wherein the surviving or screened cells comprise the
20 orthogonal tRNA-tRNA synthetase pair for use in the in vivo translation system of the second organism.
103. The method of claim 102, wherein the comparing and selecting or screening comprises an in vivo complementation assay.
104. The method of claim 102, wherein concentration of the selection or
25 screening agent is varied.
105. The method of claim 102, wherein the first organism is a prokaryotic organism.

WO 02/086075

PCT/US02/12635

106. The method of claim 102, wherein the second organism is a prokaryotic organism.
107. The method of claim 102, wherein the first and second organism are different.
- 5 108. The method of claim 102, wherein the first organism is selected from the group consisting of: *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, and a *Halobacterium*.
109. The method of claim 102, wherein the second organism is *Escherichia coli*.
- 10 110. The method of claim 38, 75, 97, or 102, wherein the selecting or screening comprises one or more positive or negative selection or screening chosen from the groups consisting of: a change in amino acid permeability, a change in translation efficiency, and a change in translational fidelity, and wherein the one or more change is based upon a mutation in one or more gene in an organism in which an orthogonal tRNA-
15 tRNA synthetase pair are used to produce protein.
111. The method of claim 38, 75, 97, or 102, wherein the selecting or screening comprises selecting or screening at least 2 selector codons within one or more selection gene or within one or more screening gene.
112. The method of claim 111, wherein the at least 2 selector codons are
20 in the same selection gene or the same screening gene.
113. The method of claim 111, wherein the at least 2 selector codons are in different selection or screening genes.
114. The method of claim 111, wherein the at least 2 selector codons comprise different selector codons.
- 25 115. The method of claim 111, wherein the at least 2 selector codons comprise the same selector codons.

WO 02/086075

PCT/US02/12635

1/35

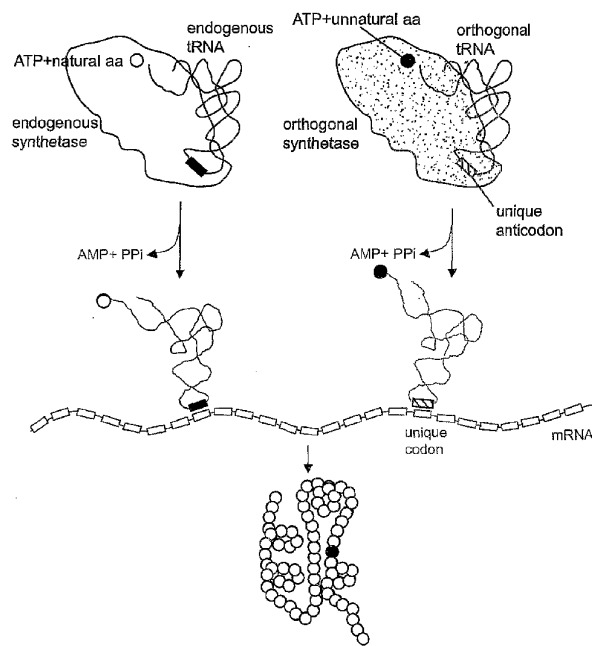


Fig. 1

WO 02/086075

PCT/US02/12635

2/35

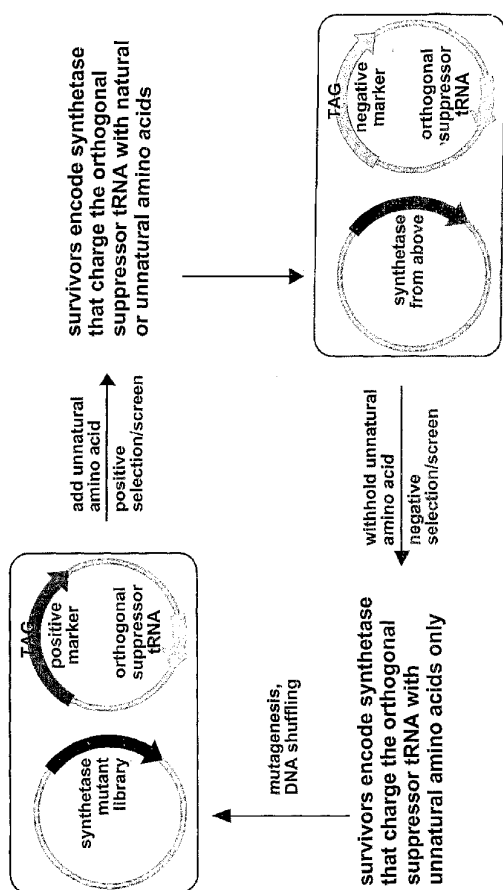


Fig. 2A

3/35

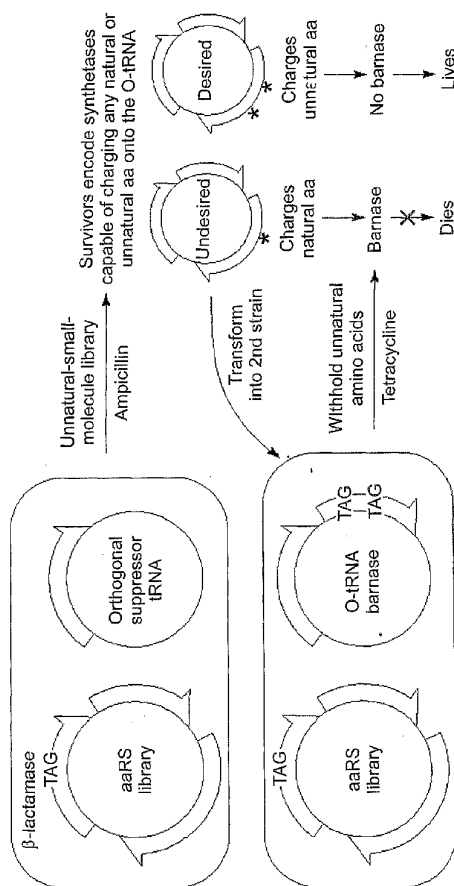
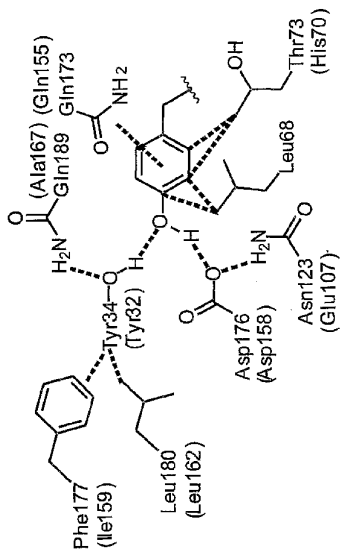


Fig. 2B

WO 02/086075

PCT/US02/12635

4/35



Wild Type	32	67	70	107	155	160	167
Library 1	Tyr --	Ala Asp Leu His	--	Glu --	Gln Val Asn Asp Ile His	Tyr Leu Gly Val Asp Val Ala	
Library 2	NNN --	Ala Asp Leu His	--	NNN --	Gln Val Asn NNN NNN His Tyr NNN Gly Val Asp Val Ala		
Library 3	NNN --	Ala Asp Leu His	--	Glu --	Gln Val Asn NNN NNN His Tyr NNN Gly Val Asp Val NNN		
Library	NNN --	NNN Asp Leu NNN	--	Glu --	NNN Val Asn NNN Ile His Tyr Leu Gly Val Asp Val NNN		

Fig. 3

WO 02/086075

PCT/US02/12635

5/35

		32	67	70	107	155	160	167
Wild Type		Tyr	-- Ala Asp Leu His	-- Glu	-- Gln Val Asn Asp Ile His	Tyr Leu Gly Val Asp Val Ala		
Library 1	NNN	-- Ala Asp Leu His	-- NNN	-- Gln Val Asn NNN NNN His	Tyr NNN Gly Val Asp Val Ala			
Library 2	NNN	-- Ala Asp Leu His	-- Glu	-- Gln Val Asn NNN NNN His	Tyr NNN Gly Val Asp Val NNN			
Library 3	NNN	-- NNN Asp Leu NNN	-- Glu	-- NNN Val Asn NNN Ile His	Tyr Leu Gly Val Asp Val NNN			

Consensus sequence

Ile -- Gln Asp Leu Tyr -- Glu -- Ala Val Asn Ala Ile His Tyr Leu Gly Val Asp Val Leu

Fig. 4

6/35

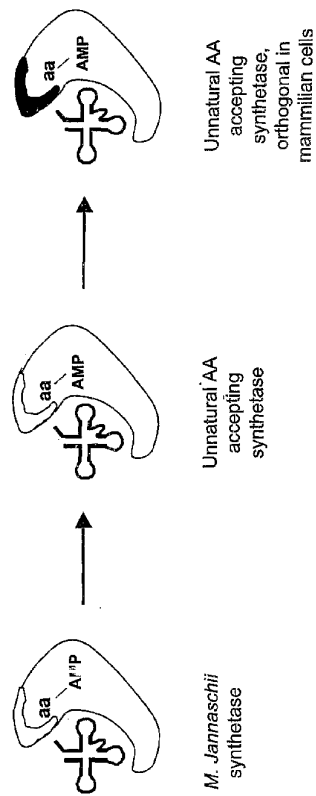


Fig. 5

7/35

EctYtRS 1 WFGNMNVLTFLLRDIGKHFVSNQMINKKAVKQRLNREDQGISTFE
HstYtRS 1 -LSKEYTLDVYRLS---SVVTQHDSKKAGAEVVKQVEHPHLLSG-

MjYtRS 1 LDKDYTLNVYRLALKTTIKRARRSMELIAREDENPKVAE
HstYtRS 1 LSKEYTLDVYRLSSVVTQHDSKKAGAEVVKQVEHPHLLSG

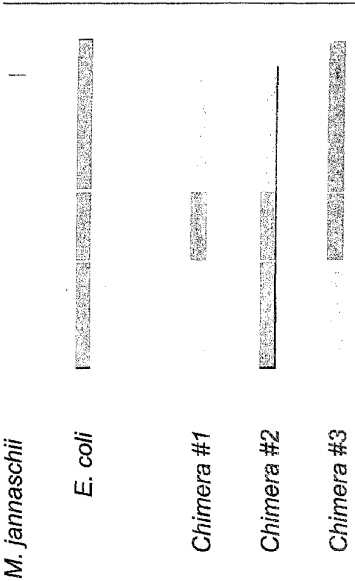


Fig. 6

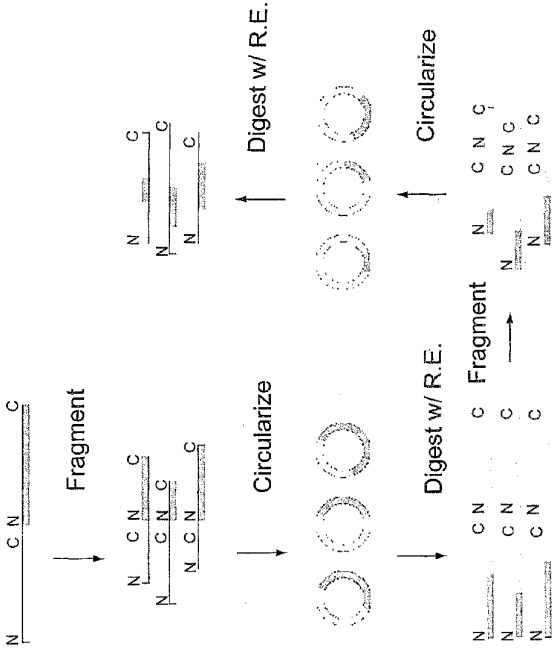


Fig. 7

9/35

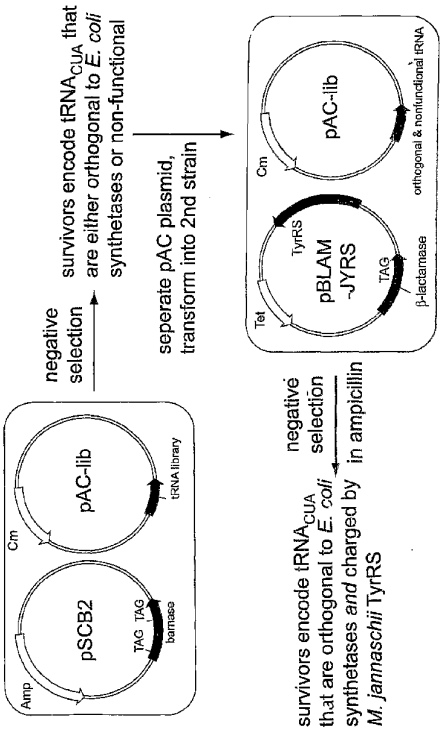
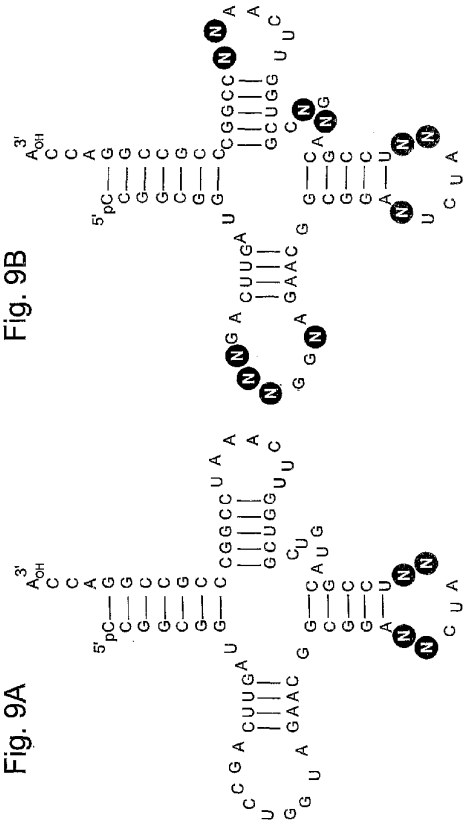


Fig. 8



WO 02/086075

PCT/US02/12635

11/35

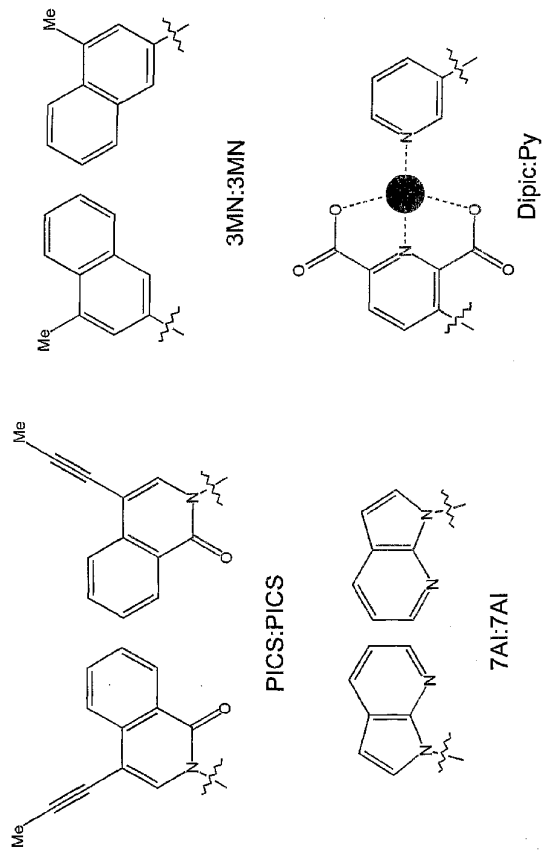


Fig. 10

WO 02/086075

PCT/US02/12635

12/35

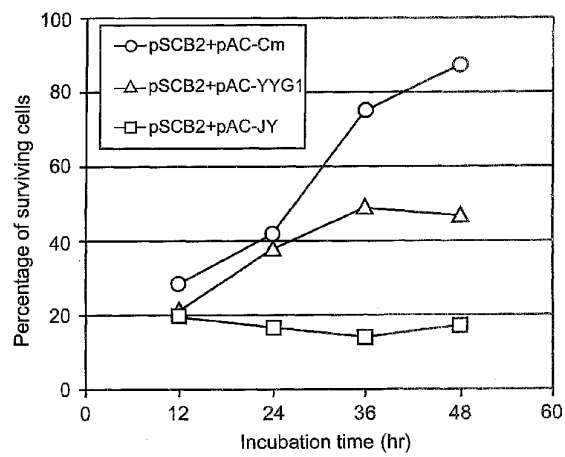


Fig. 11

WO 02/086075

PCT/US02/12635

13/35

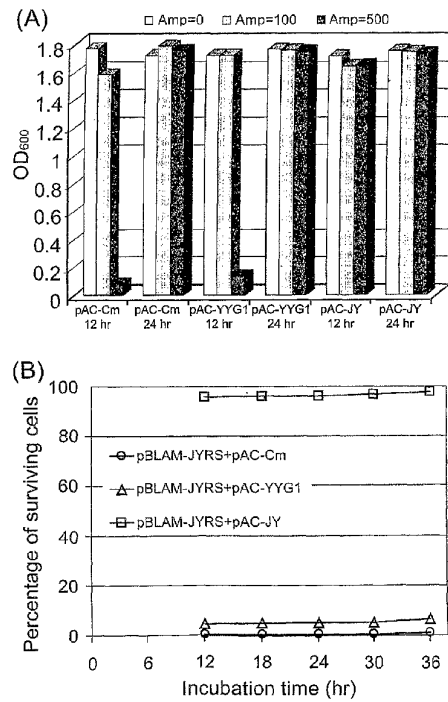


Fig. 12

WO 02/086075

PCT/US02/12635

14/35

Mutant tRNAs selected from anticodon-loop library

ATL L1B -----NN-----
JY CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
AA2 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
AA3 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
AA4 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA

Mutant tRNAs selected from all-loop libraryMutant tRNAs surviving both selections

AL L1B -----NN-----N-----N-----N-----N-----
JY CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
J15 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
J17 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
J18 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
J22 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA

Mutant tRNAs surviving negative selection only

AL L1B -----NN-----N-----N-----N-----N-----
JY CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
N11 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
N12 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
N13 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA
N16 CCGCGGTAGTTCAGCCTGTAGAACCGCGGACTCTAGATCCGATGTCTGGTTCAAATCCGGCCCGCGGACCA

Fig. 13

WO 02/086075

PCT/US02/12635

15/35

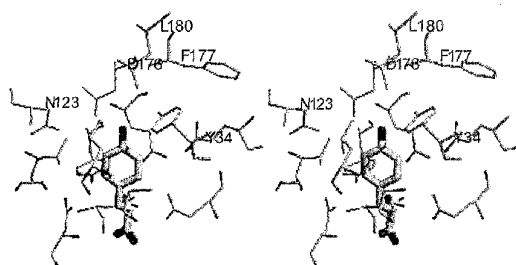


Fig. 14

WO 02/086075

PCT/US02/12635

16/35

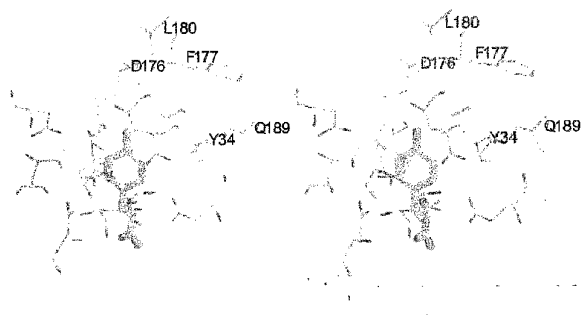


Fig. 15

17/35

Fig. 16B

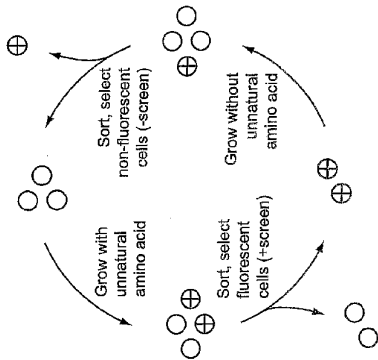
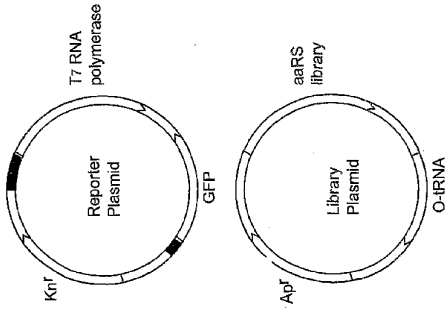


Fig. 16A



WO 02/086075

PCT/US02/12635

18/35

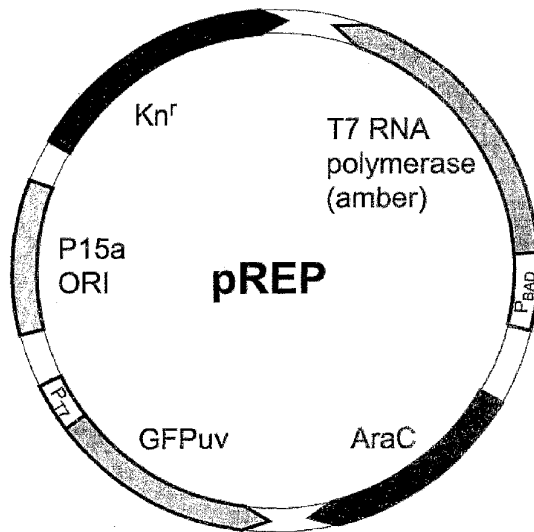


Figure 17A

19/35

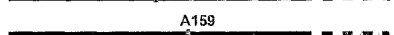
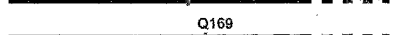
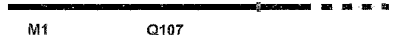
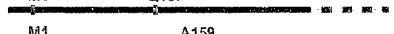
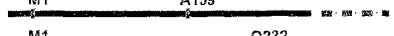
Construct Number		Fluorescence Enhancement (fold)
1		12
2		-
3		-
4		4
5		-
6		32
7		6
8		-
9		20
10		220
11		48
12		82

Figure 17B

WO 02/086075

PCT/US02/12635

20/35

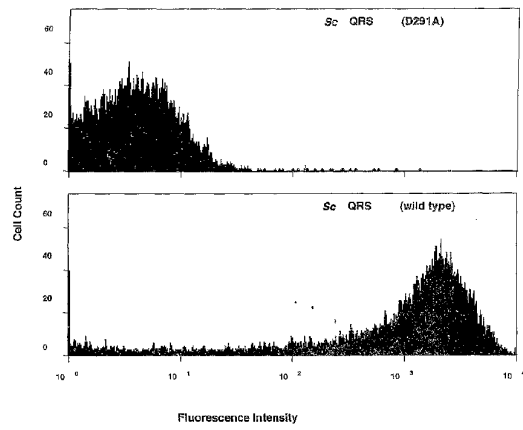


Figure 17C

WO 02/086075

PCT/US02/12635

21/35

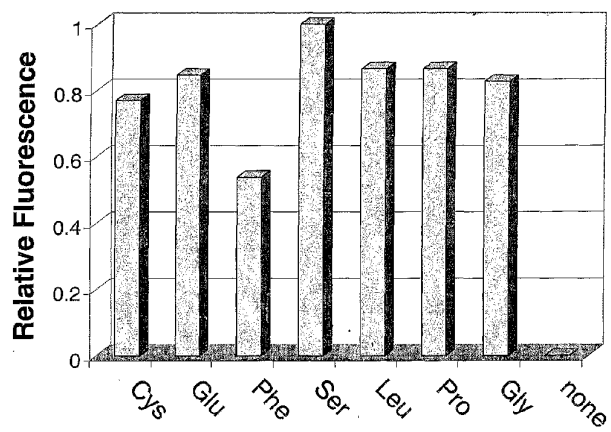


Figure 17D

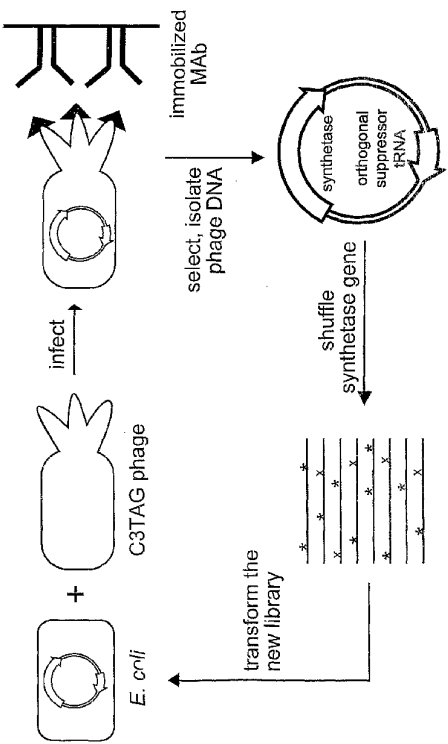


Fig. 18

WO 02/086075

PCT/US02/12635

23/35

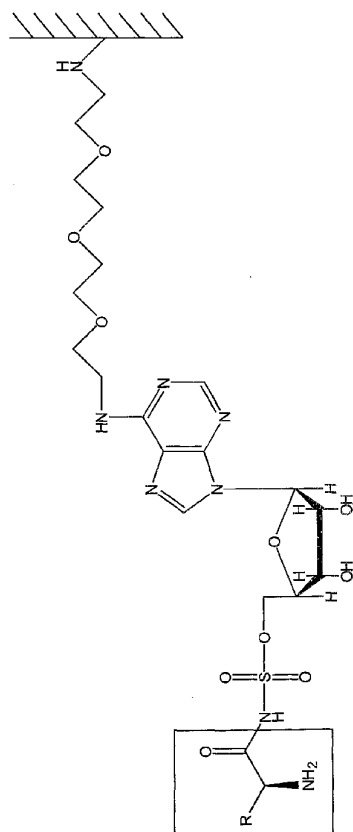


Fig. 19

WO 02/086075

PCT/US02/12635

24/35

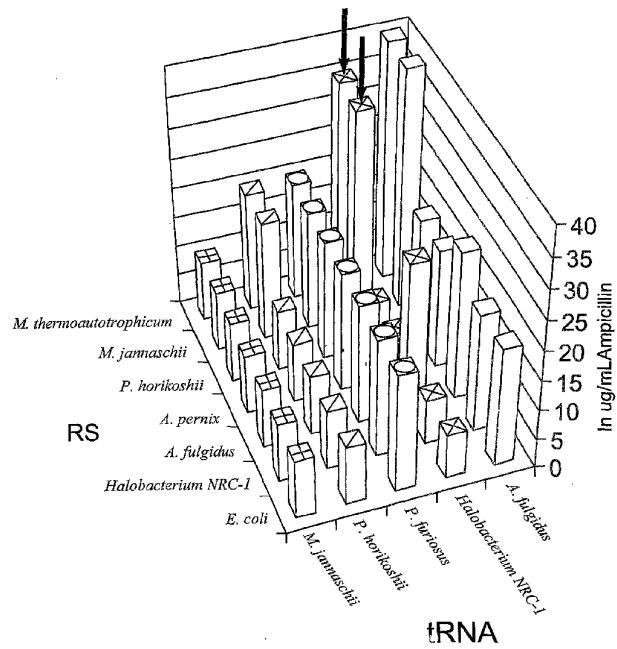


Fig. 20

25/35

Fig. 21A

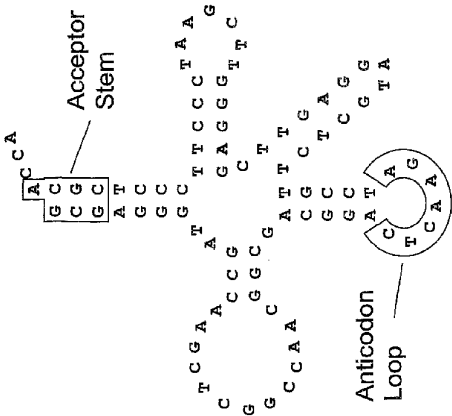


Fig. 21B

	Original Amber Mutant	5	25	Alone
	Optimized Anticodon Loop	25	500	With MLRS
	Optimized Acceptor Stem	7.5	600	

Fig. 23A

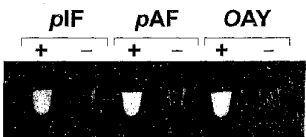


Fig. 23B

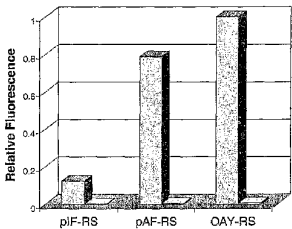
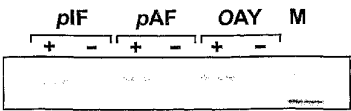


Fig. 23C

Cm IC ₅₀ (μg/mL)	pIF-RS	pAF-RS	OAY-RS(1)
- Unnatural	< 5	< 5	< 5
+ Unnatural	75	100	120

Fig. 23D



28/35

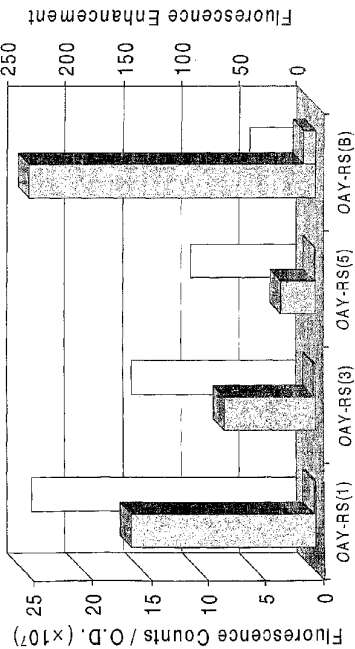


Figure 24

WO 02/086075

PCT/US02/12635

29/35

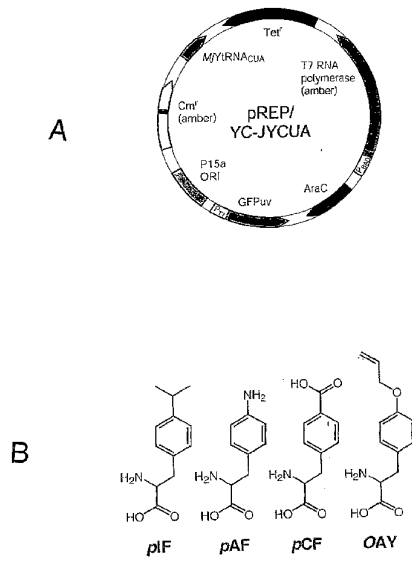


Figure 25

WO 02/086075

PCT/US02/12635

30/35

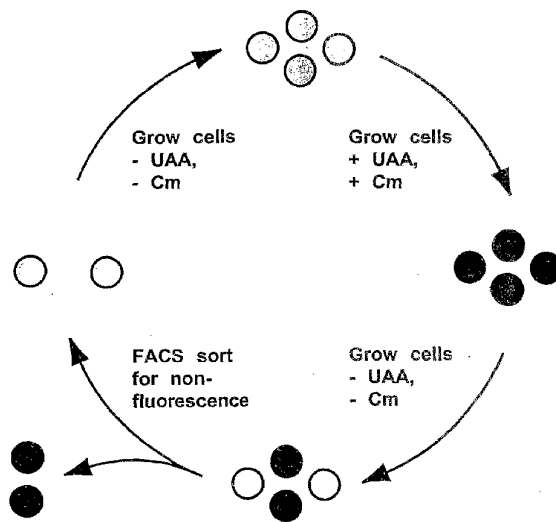


Figure 26

31/35

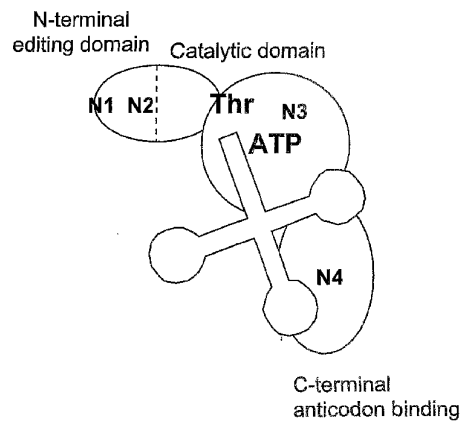


Fig. 27

WO 02/086075

PCT/US02/12635

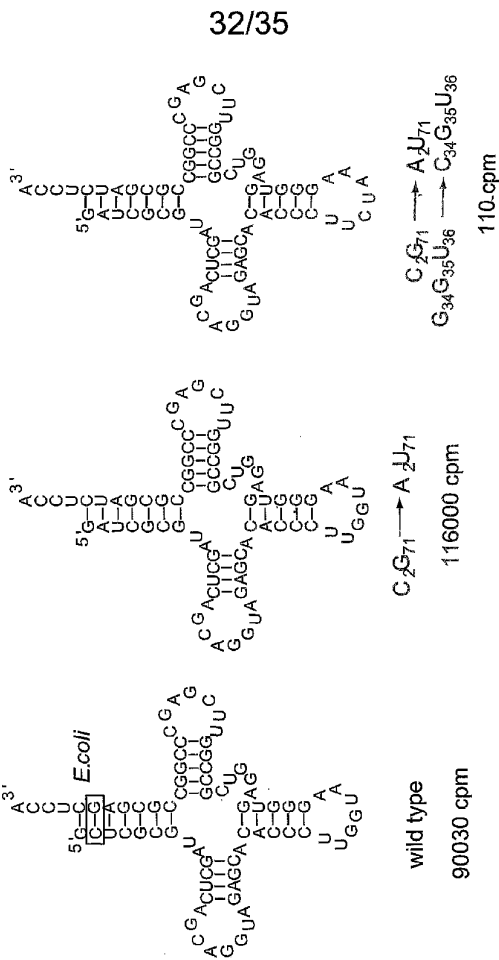


Fig. 28

WO 02/086075

PCT/US02/12635

33/35

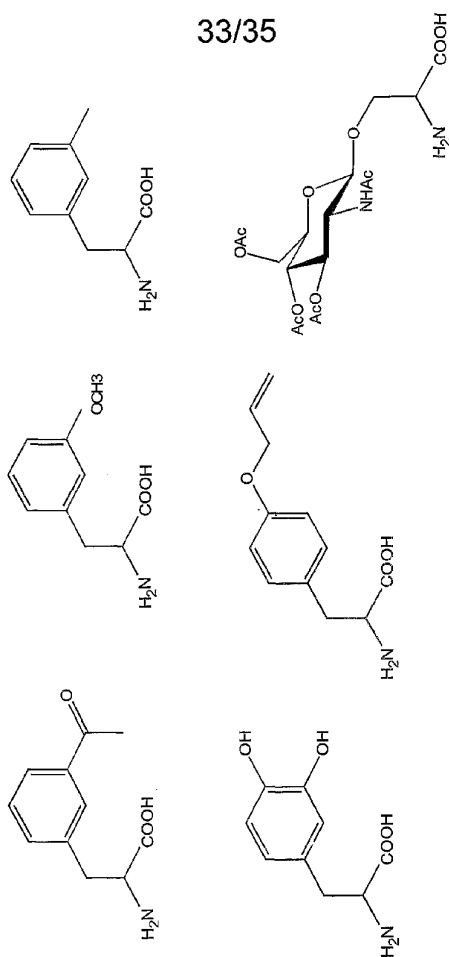


Fig. 29

WO 02/086075

PCT/US02/12635

34/35

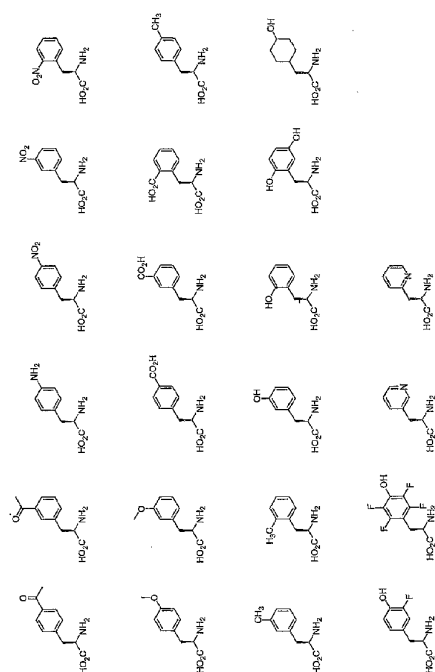


Fig. 30

WO 02/086075

PCT/US02/12635

35/35

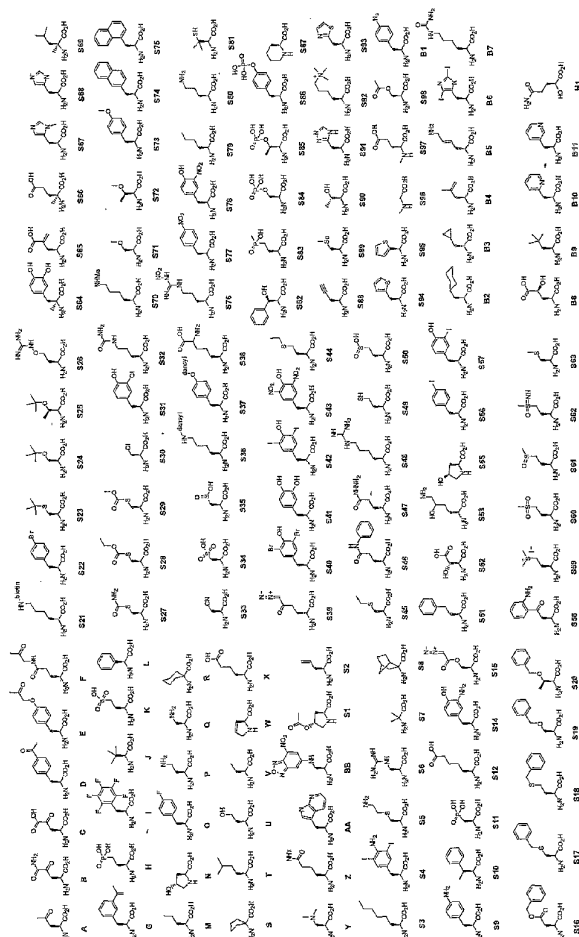


Fig. 31

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
1	CCGCGCGTGTTCAGCAGGCGAGAACGCGGACCTCTAAATCCGATCGCGCTGGTTC AAATCCGCGCCCGCGGACCA	<i>M. jannaschii</i> Ty mtRNA _{CUA}	tRNA
2	CCACAGGATAG CCAAGCTCGG CCAACGGCGA CGGACTCTAA ATCCGTTCTC GTAGGAGTTC GAGGGTTCGA ATCCCTTCCC TGGGACCA	HLAD03; an optimized amber suppressor tRNA	tRNA
3	GCAGAGGTAG CCAAGCTCGG CCAACGGCGA CGGACTCTCT AATCCGTTCT CTAGGAGTTC GAGGGTTCG AATCCCTCCC CTGGACCA	HL325A; an optimized AGGA frameshift suppressor tRNA	tRNA
4	ATGGACGAATTTGAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTGCTCAATAGGTTTGAACCAAGT GGTAAATACATTTAGGCGATTATCTCCAAATAAAGAGATGATTTGATTTACAAAT GCTGGAATTTGAATGATTTATGTTGGCTGATTTACAGGCTATTTAAACAGAGAA GGAGAGTTGGATGAGATTAGAAATAAGGAGATTATACAAATAAGTTTGAAGCA ATGGGGTTAAAGCGCAAAATATGTTTATGGAGTACTTCCAGCTTGTAAGGATTAT ACACTGATGCTCTATGATGCTGCTTTAAAAACTACCTTAAAAAGAGCAGAGAGAGT ATGGAATTTATGCAAGCGAGGTGAAATCTCAAGGTTCTCTAGATTTCTATCCA ATAATGCGAGGTTAATGCAATCTATATCTGCGCTTGTCTTCCAGTTGGAGGAGC GAGCAGAGAGAAATACATGTTTACGAGGAGGCTTTACCAAAAGAGTTGTTTGT ATTCACAAACCTTCTTAAAGGTTTGAAGGAGAGAGAGAGATGATTTCCAAAA GGGAATTTATAGCTCTTATGCTCTCCGAGAGAGATTAGGCTTAGAGTAAGGAA GCTACTGCGCAGCTGGAGTTTGAAGGAAATCCAAATAAGGAGATGCTTAAATAC TTCTCTGAATATCTTTTAACTAAAGAGCAGAGAAATTTGGTGGAGATTGACAC CTGAGTCTCTGGAGGTTTGAAGGTTTATTTAAATAAGAGATTTCTCTCAATG GATTTAAATAAGCTGATGCTGAGGACTTATAAGATTATAGAGCCAATTAGAAAG AGATTTATA	mutant TyrRS (LWJ16)	RS
5	ATGGACGAATTTGAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTGCTCAATAGGTTTGAACCAAGT GGTAAATACATTTAGGCGATTATCTCCAAATAAAGAGATGATTTGATTTACAAAT GCTGGAATTTGAATGATTTATGTTGGCTGATTTACAGGCTATTTAAACAGAGAA GGAGAGTTGGATGAGATTAGAAATAAGGAGATTATACAAATAAGTTTGAAGCA ATGGGGTTAAAGCGCAAAATATGTTTATGGAGTACTTCCAGCTTGTAAGGATTAT ACACTGATGCTCTATGATGCTGCTTTAAAAACTACCTTAAAAAGAGCAGAGAGAGT ATGGAATTTATGCAAGCGAGGTGAAATCTCAAGGTTCTCTAGATTTCTATCCA TAATGCAAGTTAAAGGTTATCAATATCTTGGCTGATTTGCAAGTTGGAGGAGTGG AGCAGAGAAATATGACATGTTAGCAGAGGAGCTTTTACCAAAAGAGTTTGTGTA TTCTCAACCTCTTCTTACAGGTTTGGATGAGAGAGAGAGAGATTCTTCAAAAG GGAAATTTATAGCTCTTATGATGCTCTCCGAGAGAGATTAGGCTTAGAGTAAGGAA GCTACTGCGCAGCTGGAGTTTGAAGGAAATCCAAATAAGGAGATGCTTAAATAC TTCTCTGAATGCTTTTACCAATAAAGGCGAGAAATTTGGTGGAGATTGACAG TTAATAGCTATGAGGAGTTAGAGGTTTATTTAAATAAGGATTTAGAGCCAATTAGAAAG GATTA	p-Pr-PheRS	RS
6	ATGGACGAATTTGAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTGCTCAATAGGTTTGAACCAAGT GGTAAATACATTTAGGCGATTATCTCCAAATAAAGAGATGATTTGATTTACAAAT GCTGGAATTTGAATGATTTATGTTGGCTGATTTACAGGCTATTTAAACAGAGAA GGAGAGTTGGATGAGATTAGAAATAAGGAGATTATACAAATAAGTTTGAAGCA ATGGGGTTAAAGCGCAAAATATGTTTATGGAGTACTTCCAGCTTGTAAGGATTAT ACACTGATGCTCTATGATGCTGCTTTAAAAACTACCTTAAAAAGAGCAGAGAGAGT ATGGAATTTATGCAAGCGAGGTGAAATCTCAAGGTTCTCTAGATTTCTATCCA TAATGCAAGTTAAAGGTTATCAATATCTTGGCTGATTTGCAAGTTGGAGGAGTGG AGCAGAGAAATATGACATGTTAGCAGAGGAGCTTTTACCAAAAGAGTTTGTGTA TTCTCAACCTCTCTTACAGGTTTGGATGAGAGAGAGAGAGATTCTTCAAAAG GGAAATTTATAGCTCTTATGATGCTCTCCGAGAGAGATTAGGCTTAGAGTAAGGAA GCTACTGCGCAGCTGGAGTTTGAAGGAAATCCAAATAAGGAGATGCTTAAATAC TTCTCTGAATGCTTTTACCAATAAAGGCGAGAAATTTGGTGGAGATTGACAG TTAATAGCTATGAGGAGTTAGAGGTTTATTTAAATAAGGATTTAGAGCCAATTAGAAAG GATTA	p-NH ₂ -PheRS(1)	RS
7	ATGGACGAATTTGAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTGCTCAATAGGTTTGAACCAAGT GGTAAATACATTTAGGCGATTATCTCCAAATAAAGAGATGATTTGATTTACAAAT GCTGGAATTTGAATGATTTATGTTGGCTGATTTACAGGCTATTTAAACAGAGAA GGAGAGTTGGATGAGATTAGAAATAAGGAGATTATACAAATAAGTTTGAAGCA ATGGGGTTAAAGCGCAAAATATGTTTATGGAGTACTTCCAGCTTGTAAGGATTAT ACACTGATGCTCTATGATGCTGCTTTAAAAACTACCTTAAAAAGAGCAGAGAGAGT ATGGAATTTATGCAAGCGAGGTGAAATCTCAAGGTTCTCTAGATTTCTATCCA TAATGCAAGTTAAAGGTTATCAATATCTTGGCTGATTTGCAAGTTGGAGGAGTGG AGCAGAGAAATATGACATGTTAGCAGAGGAGCTTTTACCAAAAGAGTTTGTGTA TTCTCAACCTCTCTTACAGGTTTGGATGAGAGAGAGAGAGATTCTTCAAAAG GGAAATTTATAGCTCTTATGATGCTCTCCGAGAGAGATTAGGCTTAGAGTAAGGAA GCTACTGCGCAGCTGGAGTTTGAAGGAAATCCAAATAAGGAGATGCTTAAATAC TTCTCTGAATGCTTTTACCAATAAAGGCGAGAAATTTGGTGGAGATTGACAG TTAATAGCTATGAGGAGTTAGAGGTTTATTTAAATAAGGATTTAGAGCCAATTAGAAAG GATTA	p-NH ₂ -PheRS(2)	RS
8	ATGGACGAATTTGAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTGCTCAATAGGTTTGAACCAAGT GGTAAATACATTTAGGCGATTATCTCCAAATAAAGAGATGATTTGATTTACAAAT GCTGGAATTTGAATGATTTATGTTGGCTGATTTACAGGCTATTTAAACAGAGAA GGAGAGTTGGATGAGATTAGAAATAAGGAGATTATACAAATAAGTTTGAAGCA ATGGGGTTAAAGCGCAAAATATGTTTATGGAGTACTTCCAGCTTGTAAGGATTAT ACACTGATGCTCTATGATGCTGCTTTAAAAACTACCTTAAAAAGAGCAGAGAGAGT ATGGAATTTATGCAAGCGAGGTGAAATCTCAAGGTTCTCTAGATTTCTATCCA TAATGCAAGTTAAAGGTTATCAATATCTTGGCTGATTTGCAAGTTGGAGGAGTGG AGCAGAGAAATATGACATGTTAGCAGAGGAGCTTTTACCAAAAGAGTTTGTGTA TTCTCAACCTCTCTTACAGGTTTGGATGAGAGAGAGAGATTGATTTCTCAAAAG GGAAATTTATAGCTCTTATGATGCTCTCCGAGAGAGATTAGGCTTAGAGTAAGGAA GCTACTGCGCAGCTGGAGTTTGAAGGAAATCCAAATAAGGAGATGCTTAAATAC TTCTCTGAATGCTTTTACCAATAAAGGCGAGAAATTTGGTGGAGATTGACAG TTAATAGCTATGAGGAGTTAGAGGTTTATTTAAATAAGGATTTAGAGCCAATTAGAAAG GATTA	p-NH ₂ -PheRS(3a)	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
	GGAGAGTTGGATGAGATTACAAATAGGAGATTATACAAAAAGTTTGAAGCA ATGGGTTAAAGGCAAAATATTATTAGGAGTGGATTCCAGCTTGATAGGATTAT ACACTGAAATGCTATAGATTGGCTTTAAAAATACCTTAAAGAGCAAGAGAGGT ATGGAACCTATAGAGAGAGAGGATGAAATCCAAAGGTTGCTGAAGTTATCTATCCAA TAAGCGAGTTATCGGCGGATATCCCTGCTTGAGGTTCAGTTGAGAGAGTGG AGCAGAGAAAAATACACATGTAGCAAGGAGCTTTTACCAAAAAGGTTGTTGTA TTACAAACCTGCTTTAACGGGTTGGATGGAGAGAGAAAGATGAGTTCTCAAAG GGATTTTATAGCTGTTGATGATCTTCCAGAGAGATTAGAGCTAGATTAAGAGAG CATACTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATGAGAGATAGCTAAATACT TCCCTGAATATCCCTTAAACCTAAAGAGGCGAGAAAAATTTGGTGGAGATTGACAG TTAATAGCTATGGGAGTTAGAGAGTTTATTTAAAAATAGGAAATTCATCCATGG ATTAAAAATGCTTAGCTGAGAGACTTATAAGATTTCAGAGCCATAGAGAGA GATTA		
9	ATGGACGAATTTGAATGATTAAGAGAAACACATCTGAATATTACGCGAGGAGAG TTAGAGAGGTTTTAAAAAAGAGGAAATCTCTGATATAGGTTTGAAGCAGGT GGTAAATACATTTAGGCGATTATCTCCAAATAAAAAGATGATGATTACAAAT GCTGGATTGATATATATATGTTGGCTGATTTACAGCCTTATTAACAGAGAA GGAGAGTTGGTGGAGTTAGAAATAGGAGATTACAAAGAGTTTGAAGAGCA ATGGGGTTAAAGGCAAAATGTTTATGGAAGTCTTCCAGCTTGAAGAGATTAT ACACTGAATGCTTATAGATGGCTTTAAAAATACCTTAAAGAGCAAGAGAGAGT ATGGAACCTTATAGAGAGAGAGATGAAATCCAAAGGTTGCTGAAGTTATCTCCAA TATTCAGCTTATAGCAGGCTTTGAGAGGCTTTGATTTGAGTTGGAGAGATGG AGCAGAGAAAAATACACATGTAGCAAGGAGCTTTTACCAAAAAGGTTGTTGTA TTACAAACCTGCTTTAACGGGTTGGATGGAGAGAGAAAGATGAGTTCTCAAAG GGATTTTATAGCTGTTGATGATCTTCCAGAGAGATTAGAGCTAGATTAAGAGAG CATACTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATGAGAGATAGCTAAATACT TCCCTGAATATCCCTTAAACCTAAAGAGGCGAGAAAAATTTGGTGGAGATTGACAG TTAATAGCTATGGGAGTTAGAGAGTTTATTTAAAAATAGGAAATTCATCCATGG ATTAAAAATGCTTAGCTGAGAGACTTATAAGATTTCAGAGCCATAGAGAGA GATTA	p-NH ₂ -PheRS(3b)	RS
10	ATGGACGAATTTGAATGATTAAGAGAAACACATCTGAATATTACGCGAGGAGAG TTAGAGAGGTTTTAAAAAAGAGTGAATATCTGCTGATAGGTTTGAAGCAGGT GGTAAATACATTTAGGCGATTATCTCCAAATAAAAAGATGATGATTACAAAT GCTGGATTGATATATATATGTTGGCTGATTTACAGCCTTATTAACAGAGAA GGAGAGTTGGTGGAGTTAGAAATAGGAGATTACAAAGAGTTTGAAGAGCA ATGGGGTTAAAGGCAAAATGTTTATGGAAGTCTTCCAGCTTGAAGAGATTAT ACACTGAATGCTTATAGATGGCTTTAAAAATACCTTAAAGAGCAAGAGAGAGT ATGGAACCTTATAGAGAGAGAGATGAAATCCAAAGGTTGCTGAAGTTATCTCCAA TATTCAGCTTATAGCAGGCTTTGAGAGGCTTTGATTTGAGTTGGAGAGATGG AGCAGAGAAAAATACACATGTAGCAAGGAGCTTTTACCAAAAAGGTTGTTGTA TTACAAACCTGCTTTAACGGGTTGGATGGAGAGAGAAAGATGAGTTCTCAAAG GGATTTTATAGCTGTTGATGATCTTCCAGAGAGATTAGAGCTAGATTAAGAGAG CATACTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATGAGAGATAGCTAAATACT TCCCTGAATATCCCTTAAACCTAAAGAGGCGAGAAAAATTTGGTGGAGATTGACAG TTAATAGCTATGGGAGTTAGAGAGTTTATTTAAAAATAGGAAATTCATCCATGG ATTAAAAATGCTTAGCTGAGAGACTTATAAGATTTCAGAGCCATAGAGAGA GATTA	O-Allyl-TyrRS(1)	RS
11	ATGGACGAATTTGAATGATTAAGAGAAACACATCTGAATATTACGCGAGGAGAG TTAGAGAGGTTTTAAAAAAGAGTGAATATCTGCTGATAGGTTTGAAGCAGGT GGTAAATACATTTAGGCGATTATCTCCAAATAAAAAGATGATGATTACAAAT GCTGGATTGATATATATATGTTGGCTGATTTACAGCCTTATTAACAGAGAA GGAGAGTTGGTGGAGTTAGAAATAGGAGATTACAAAGAGTTTGAAGAGCA ATGGGGTTAAAGGCAAAATGTTTATGGAAGTCTTCCAGCTTGAAGAGATTAT ACACTGAATGCTTATAGATGGCTTTAAAAATACCTTAAAGAGCAAGAGAGAGT ATGGAACCTTATAGAGAGAGAGTGAATCCAAAGGTTGCTGAAGTTATCTCCAA TATTCAGCTTATAGCAGGCTTTGAGAGGCTTTGATTTGAGTTGGAGAGATGG AGCAGAGAAAAATACACATGTAGCAAGGAGCTTTTACCAAAAAGGTTGTTGTA TTACAAACCTGCTTTAACGGGTTGGATGGAGAGAGAAAGATGAGTTCTCAAAG GGATTTTATAGCTGTTGATGATCTTCCAGAGAGATTAGAGCTAGATTAAGAGAG CATACTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATGAGAGATAGCTAAATACT TCCCTGAATATCCCTTAAACCTAAAGAGGCGAGAAAAATTTGGTGGAGATTGACAG TTAATAGCTATGGGAGTTAGAGAGTTTATTTAAAAATAGGAAATTCATCCATGG ATTAAAAATGCTTAGCTGAGAGACTTATAAGATTTCAGAGCCATAGAGAGA GATTA	O-Allyl-TyrRS(3)	RS
12	ATGGACGAATTTGAATGATTAAGAGAAACACATCTGAATATTACGCGAGGAGAG TTAGAGAGGTTTTAAAAAAGAGTGAATATCTGCTGATAGGTTTGAAGCAGGT GGTAAATACATTTAGGCGATTATCTCCAAATAAAAAGATGATGATTACAAAT GCTGGATTGATATATATATGTTGGCTGATTTACAGCCTTATTAACAGAGAA GGAGAGTTGGTGGAGTTAGAAATAGGAGATTACAAAGAGTTTGAAGAGCA ATGGGGTTAAAGGCAAAATGTTTATGGAAGTCTTCCAGCTTGAAGAGATTAT ACACTGAATGCTTATAGATGGCTTTAAAAATACCTTAAAGAGCAAGAGAGAGT ATGGAACCTTATAGAGAGAGATGAATCCAAAGGTTGCTGAAGTTATCTCCAA TATTCAGCTTATAGCAGGCTTTGAGAGGCTTTGATTTGAGTTGGAGAGATGG AGCAGAGAAAAATACACATGTAGCAAGGAGCTTTTACCAAAAAGGTTGTTGTA TTACAAACCTGCTTTAACGGGTTGGATGGAGAGAGAAAGATGAGTTCTCAAAG GGATTTTATAGCTGTTGATGATCTTCCAGAGAGATTAGAGCTAGATTAAGAGAG CATACTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATGAGAGATAGCTAAATACT TCCCTGAATATCCCTTAAACCTAAAGAGGCGAGAAAAATTTGGTGGAGATTGACAG TTAATAGCTATGGGAGTTAGAGAGTTTATTTAAAAATAGGAAATTCATCCATGG ATTAAAAATGCTTAGCTGAGAGACTTATAAGATTTCAGAGCCATAGAGAGA GATTA	O-Allyl-TyrRS(4)	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
13	ATGGAAGCAATTTGAAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAAGAG TTAAGAGAGGTTTAAAGAAAGATGAAATCTCTCATATAGGTTTGAACCAAGT GCTAAATACATTTAGGCGATTTCTCCAAATATAAGAGTATGATTATCAAAAT GCTGATTTGATATATTTATTTGCTGATTTACACCTTTTAAACCGAA GGAGGTTGGATGAGATTAAGAAATAGGAGATTATACAAAGATTTTGAAGCA ATCGGTTTAAAGCAAAATGTTTATGGAAGTAMTTCACCTTGTATAGGATTAT ACCTGATTTCTTGAATGCTTTAAAJAATCTCTTAAAGAGCAAGAGGAT ATCGAATTTATAGAGAGGATGAAATCCAAAGTTGCTGAGTTATCTATCCAA TAATGCAAGTTAATCCGAGCATATCAGGCTGTGATGTTGAGTTGGAGGAGTGG AGCAGGAAATATCAGATTTTACAGAGGAGCTTTTACAAAGATTTGTTTGA TTCCAAACCTGCTTAAAGGTTGGATGAGAGGAAGATGATCTCTTAAAG GGAAATTTADAGCTGTGATGACTCTCCAGAGAGATTAAGGCTATAGATAAGAAAG CATATGCCAGCTGCTGATTTGTGAAGAAATCCAAATAGAGTATGCTAAATCT TCTTGAATATCTTTAAACCAAAAGGCGCAAAATTTGGTGGAGTTTGACAG TTAATAGCTATGAGGATTAAGAGTTTATTTAAAGATTAAGGATTTGATCCAAATGG ATTTAAAGATGCTGATGCTGAGGACTTATAAGATTTAGAGCCATTTAGAAAGA GATTA	p-Br-PheRS	RS
14	ATGGAAGCAATTTGAAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAAGAG TTAAGAGAGGTTTAAAGAAAGATGAAATCTCTGCTATAGGTTTGAACCAAGT GCTAAATACATTTAGGCGATTTCTCCAAATATAAGAGTATGATTATCAAAAT GCTGATTTGATATATTTATTTGCTGATTTACACCTTTTAAACCGAA GGAGGTTGGATGAGATTAAGAAATAGGAGATTATACAAAGATTTTGAAGCA ATCGGTTTAAAGCAAAATGTTTATGGAAGTAMTTCACCTTGTATAGGATTAT ACCTGATTTCTTGAATGCTTTAAAJAATCTCTTAAAGAGCAAGAGGAT ATCGAATTTATAGAGAGGATGAAATCCAAAGTTGCTGAGTTATCTATCCAA TAATGCAAGTTAATCCGAGCATATCAGGCTGTGATGTTGAGTTGGAGGAGTGG AGCAGGAAATATCAGATTTTACAGAGGAGCTTTTACAAAGATTTGTTTGA TTCCAAACCTGCTTAAAGGTTGGATGAGAGGAAGATGATCTCTTAAAG GGAAATTTADAGCTGTGATGACTCTCCAGAGAGATTAAGGCTATAGATAAGAAAG CATATGCCAGCTGCTGATTTGTGAAGAAATCCAAATAGAGTATGCTAAATCT TCTTGAATATCTTTAAACCAAAAGGCGCAAAATTTGGTGGAGTTTGACAG TTAATAGCTATGAGGATTAAGAGTTTATTTAAAGATTAAGGATTTGATCCAAATGG ATTTAAAGATGCTGATGCTGAGGACTTATAAGATTTAGAGCCATTTAGAAAGA GATTA	p-Az-PheRS(1)	RS
15	ATGGAAGCAATTTGAAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAAGAG TTAAGAGAGGTTTAAAGAAAGATGAAATCTCTGCTATAGGTTTGAACCAAGT GCTAAATACATTTAGGCGATTTCTCCAAATATAAGAGTATGATTATCAAAAT GCTGATTTGATATATTTATTTGCTGATTTACACCTTTTAAACCGAA GGAGGTTGGATGAGATTAAGAAATAGGAGATTATACAAAGATTTTGAAGCA ATCGGTTTAAAGCAAAATGTTTATGGAAGTAMTTCACCTTGTATAGGATTAT ACCTGATTTCTTGAATGCTTTAAAJAATCTCTTAAAGAGCAAGAGGAT ATCGAATTTATAGAGAGGATGAAATCCAAAGTTGCTGAGTTATCTATCCAA TAATGCAAGTTAATCCGAGCATATCAGGCTGTGATGTTGAGTTGGAGGAGTGG AGCAGGAAATATCAGATTTTACAGAGGAGCTTTTACAAAGATTTGTTTGA TTCCAAACCTGCTTAAAGGTTGGATGAGAGGAAGATGATCTCTTAAAG GGAAATTTADAGCTGTGATGACTCTCCAGAGAGATTAAGGCTATAGATAAGAAAG CATATGCCAGCTGCTGATTTGTGAAGAAATCCAAATAGAGTATGCTAAATCT TCTTGAATATCTTTAAACCAAAAGGCGCAAAATTTGGTGGAGTTTGACAG TTAATAGCTATGAGGATTAAGAGTTTATTTAAAGATTAAGGATTTGATCCAAATGG ATTTAAAGATGCTGATGCTGAGGACTTATAAGATTTAGAGCCATTTAGAAAGA GATTA	p-Az-PheRS(3)	RS
16	ATGGAAGCAATTTGAAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAAGAG TTAAGAGAGGTTTAAAGAAAGATGAAATCTCTGCTATAGGTTTGAACCAAGT GCTAAATACATTTAGGCGATTTCTCCAAATATAAGAGTATGATTATCAAAAT GCTGATTTGATATATTTATTTGCTGATTTACACCTTTTAAACCGAA GGAGGTTGGATGAGATTAAGAAATAGGAGATTATACAAAGATTTTGAAGCA ATCGGTTTAAAGCAAAATGTTTATGGAAGTAMTTCACCTTGTATAGGATTAT ACCTGATTTCTTGAATGCTTTAAAJAATCTCTTAAAGAGCAAGAGGAT ATCGAATTTATAGAGAGGATGAAATCCAAAGTTGCTGAGTTATCTATCCAA TAATGCAAGTTAATCCGAGCATATCAGGCTGTGATGTTGAGTTGGAGGAGTGG AGCAGGAAATATCAGATTTTACAGAGGAGCTTTTACAAAGATTTGTTTGA TTCCAAACCTGCTTAAAGGTTGGATGAGAGGAAGATGATCTCTTAAAG GGAAATTTADAGCTGTGATGACTCTCCAGAGAGATTAAGGCTATAGATAAGAAAG CATATGCCAGCTGCTGATTTGTGAAGAAATCCAAATAGAGTATGCTAAATCT TCTTGAATATCTTTAAACCAAAAGGCGCAAAATTTGGTGGAGTTTGACAG TTAATAGCTATGAGGATTAAGAGTTTATTTAAAGATTAAGGATTTGATCCAAATGG ATTTAAAGATGCTGATGCTGAGGACTTATAAGATTTAGAGCCATTTAGAAAGA GATTA	p-Az-PheRS(5)	RS
17	ATGGAAGCAATTTGAAATGATAAGAGAAACACATCTGAATTTATCAGCGAGGAAGAG TTAAGAGAGGTTTAAAGAAAGATGAAATCTCTGCTATAGGTTTGAACCAAGT GCTAAATACATTTAGGCGATTTCTCCAAATATAAGAGTATGATTATCAAAAT GCTGATTTGATATATTTATTTGCTGATTTACACCTTTTAAACCGAA GGAGGTTGGATGAGATTAAGAAATAGGAGATTATACAAAGATTTTGAAGCA ATCGGTTTAAAGCAAAATGTTTATGGAAGTAMTTCACCTTGTATAGGATTAT ACCTGATTTCTTGAATGCTTTAAAJAATCTCTTAAAGAGCAAGAGGAT ATCGAATTTATAGAGAGGATGAAATCCAAAGTTGCTGAGTTATCTATCCAA TAATGCAAGTTAATCCGAGCATATCAGGCTGTGATGTTGAGTTGGAGGAGTGG AGCAGGAAATATCAGATTTTACAGAGGAGCTTTTACAAAGATTTGTTTGA TTCCAAACCTGCTTAAAGGTTGGATGAGAGGAAGATGATCTCTTAAAG GGAAATTTADAGCTGTGATGACTCTCCAGAGAGATTAAGGCTATAGATAAGAAAG CATATGCCAGCTGCTGATTTGTGAAGAAATCCAAATAGAGTATGCTAAATCT TCTTGAATATCTTTAAACCAAAAGGCGCAAAATTTGGTGGAGTTTGACAG TTAATAGCTATGAGGATTAAGAGTTTATTTAAAGATTAAGGATTTGATCCAAATGG ATTTAAAGATGCTGATGCTGAGGACTTATAAGATTTAGAGCCATTTAGAAAGA GATTA	Mutmt synthetases to incorporate m-acyl phenylalanine into proteins (Ketone 3-4)	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
18	GTAAATAGCTATAGGAGTATAGAGATTTATTAATAATAGGAATTCATCAATG GATTTAAATAATGCTGTAGCTGAAGAATTTATAAGATTTAGAGCAATTAGAAG AGATTTATAA ATGAGAGGATTTGAATATGAAGAAGAAACATCTGAATTTACAGGAGGAGAG TTAAGAGAGGTTTAAATAAGATTAATAATCTGCTTACAGGCTTTGACCAAGT GTTAAATAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTACAAAT GCTGGATTTGATTAATATATATTTGCTGCTGATTTACAGGCTTTTAAACAGAAA GAGAGGTTGATTTAGATTTGAATAATAGGAGATTTATACAAATAAGTTTGTGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTCTATTCCAGCTTGATAAGGATTA ACACTGAATGCTATAGATTTGGCTTTAAAACTACCTTAAAAAGAGCAAGAGAGG ATGGAGCTTTATACAGAGAGAGATTAATAATCAAGGTTGCTTAAGATTTATCTCCA ATATCTCAGGTTTAAATATATCTATATACAGGAGTGTATGTTCCAGTTGGAGGATG GAGCAGAGAAAATACACATGTTAGCAAGGAGGCTTTTACCAATAAGGTTGTTTGT ATTACACACCCCTGCTTAAAGGCTTTGATGAGAGAGAGAGATGATTTTCAAAA GGCAATTTTATAGCTTTGATGATCTTCCGAAAGGATTTAGGCTTAAAGAAA GCATATCTGCCCAGCTGGAGTTGTTGAAGGAATCCATAATGAGATAGCTAAATAC TTCTTTGAATATCTTTTAACTTAAAGAGCCAGAAAATTTGCTGAGATTTGACA GTTTATAGCTTTAGAGGTTGAGAGTTTATTTAAATAAGGAATTCCTCCAAATG GATTTAAATAATGCTGTAGCTGAAGAATTTATAAGATTTAGAGCAATTAGAAG AGATTTATAA	Mutant synthetase to incorporate m-acyl phenylalanine into proteins (Ketone 3-7)	RS
19	ATGAGCAATTTGAATATGAAGAAGAAACATCTGAATTTATCAGGAGGAGAG TTAAGAGAGGTTTAAATAAGATTAATAATCTGCTTAAATAGGTTTGAACCAAGT GTTAAATAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTACAAAT GCTGGATTTGATTAATATATATTTGCTGCTGATTTACAGGCTTTTAAACAGAAA GAGAGGTTGATTTAGATTTGAATAATAGGAGATTTATACAAATAAGTTTGTGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTA ACACTGAATGCTATAGATTTGGCTTTAAAACTACCTTAAAAAGAGCAAGAGAGG ATGCAACCTTTTATGGAAGAGAGGATTAATAATCCAAAGGTTGCTGAGTTTATCTCA ATAATGAGGTTTAAAGATTTATCTTTTAAAGGCTTGAAGTTTGAAGTTGAGAGGAT GAGCAGAGAAAATACACATGTTAGCAAGGAGGCTTTTACCAATAAGGTTGTTTGT ATTACACACCTTTGCTTAAAGGCTTTGATGAGCTTCCGAAAGGATTTAGGCTTAA GGGAATTTTATAGCTTTGATGAGCTTCCGAAAGGATTTAGGCTTAAAGTAAAGAAA GCATATCTGCCCAGCTGGAGTTGTTGAAGGAATCCATAATAGGAGATAGCTAAATAC TTCTTTGAATATCTTTTAACTTAAAGAGCCAGAAAATTTGCTGAGATTTGACA GTTTATAGCTTTAGAGGTTGAGAGTTTATTTAAATAAGGAATTCCTCCAAATG GATTTAAATAATGCTGTAGCTGAAGAATTTATAAGATTTAGAGCAATTAGAAG AGATTTATAA	Mutant synthetase to incorporate m-acyl phenylalanine into proteins (Ketone 4-1)	RS
20	ATGAGCAATTTGAATATGAAGAAGAAACATCTGAATTTATCAGGAGGAGAG TTAAGAGAGGTTTAAATAAGATTAATAATCTGCTTAAATAGGTTTGAACCAAGT GTTAAATAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTACAAAT GCTGGATTTGATTAATATATATTTGCTGCTGATTTACAGGCTTTTAAACAGAAA GAGAGGTTGATTTAGATTTGAATAATAGGAGATTTATACAAATAAGTTTGTGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTA ACACTGAATGCTATAGATTTGGCTTTAAAACTACCTTAAAAAGAGCAAGAGAGG ATGGAACCTTTATGCAAGAGAGGATTAATAATCCAAAGGTTGCTGAGTTTATCTCA ATAATGAGGTTTAAAGATTTATCTTTTAAAGGCTTGAAGTTTGAAGTTGAGAGGAT GAGCAGAGAAAATACACATGTTAGCAAGGAGGCTTTTACCAATAAGGTTGTTTGT ATTACACACCTTTGCTTAAAGGCTTTGATGAGCTTCCGAAAGGATTTAGGCTTAA GGGAATTTTATAGCTTTGATGAGCTTCCGAAAGGATTTAGGCTTAAAGTAAAGAAA GCATATCTGCCCAGCTGGAGTTGTTGAAGGAATCCATAATAGGAGATAGCTAAATAC TTCTTTGAATATCTTTTAACTTAAAGAGCCAGAAAATTTGCTGAGATTTGACA GTTTATAGCTTTAGAGGTTGAGAGTTTATTTAAATAAGGAATTCCTCCAAATG GATTTAAATAATGCTGTAGCTGAAGAATTTATAAGATTTAGAGCAATTAGAAG AGATTTATAA	Mutant synthetase to incorporate m-acyl phenylalanine into proteins (Ketone 5-4)	RS
21	ATGAGCAATTTGAATATGAAGAAGAAACATCTGAATTTATCAGGAGGAGAG TTAAGAGAGGTTTAAATAAGATTAATAATCTGCTTAAATAGGTTTGAACCAAGT GTTAAATAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTACAAAT GCTGGATTTGATTAATATATATTTGCTGCTGATTTACAGGCTTTTAAACAGAAA GAGAGGTTGATTTAGATTTGAATAATAGGAGATTTATACAAATAAGTTTGTGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTA ACACTGAATGCTATAGATTTGGCTTTAAAACTACCTTAAAAAGAGCAAGAGAGG ATGGAACCTTTATGCAAGAGAGGATTAATAATCCAAAGGTTGCTGAGTTTATCTCA ATAATGAGGTTTAAAGATTTATCTTTTAAAGGCTTGAAGTTTGAAGTTGAGAGGAT GAGCAGAGAAAATACACATGTTAGCAAGGAGGCTTTTACCAATAAGGTTGTTTGT ATTACACACCTTTGCTTAAAGGCTTTGATGAGCTTCCGAAAGGATTTAGGCTTAA GGGAATTTTATAGCTTTGATGAGCTTCCGAAAGGATTTAGGCTTAAAGTAAAGAAA GCATATCTGCCCAGCTGGAGTTGTTGAAGGAATCCATAATAGGAGATAGCTAAATAC TTCTTTGAATATCTTTTAACTTAAAGAGCCAGAAAATTTGCTGAGATTTGACA GTTTATAGCTTTAGAGGTTGAGAGTTTATTTAAATAAGGAATTCCTCCAAATG GATTTAAATAATGCTGTAGCTGAAGAATTTATAAGATTTAGAGCAATTAGAAG AGATTTATAA	Mutant synthetase to incorporate m-acyl phenylalanine into proteins (Ketone 6-8)	RS
22	ATGAGCAATTTGAATATGAAGAAGAAACATCTGAATTTATCAGGAGGAGAG TTAAGAGAGGTTTAAATAAGATTAATAATCTGCTTAAATAGGTTTGAACCAAGT GTTAAATAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTACAAAT GCTGGATTTGATTAATATATATTTGCTGCTGATTTACAGGCTTTTAAACAGAAA GAGAGGTTGATTTAGATTTGAATAATAGGAGATTTATACAAATAAGTTTGTGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTA ACACTGAATGCTATAGATTTGGCTTTAAAACTACCTTAAAAAGAGCAAGAGAGG ATGGAACCTTTATGCAAGAGAGGATTAATAATCCAAAGGTTGCTGAGTTTATCTCA ATAATGAGGTTTAAAGATTTATCTTTTAAAGGCTTGAAGTTTGAAGTTGAGAGGAT GAGCAGAGAAAATACACATGTTAGCAAGGAGGCTTTTACCAATAAGGTTGTTTGT ATTACACACCTTTGCTTAAAGGCTTTGATGAGCTTCCGAAAGGATTTAGGCTTAA GGGAATTTTATAGCTTTGATGAGCTTCCGAAAGGATTTAGGCTTAAAGTAAAGAAA GCATATCTGCCCAGCTGGAGTTGTTGAAGGAATCCATAATAGGAGATAGCTAAATAC TTCTTTGAATATCTTTTAACTTAAAGAGCCAGAAAATTTGCTGAGATTTGACA GTTTATAGCTTTAGAGGTTGAGAGTTTATTTAAATAAGGAATTCCTCCAAATG GATTTAAATAATGCTGTAGCTGAAGAATTTATAAGATTTAGAGCAATTAGAAG AGATTTATAA	Mutant synthetase to incorporate m-methoxy phenylalanine into proteins (OMe 1-6)	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
23	GGGAATTTTATGACCTGTTGATGACTCTCCAGAGAGATTTAGGGCTTANGATAAGAAAG GCAATCTGCCAGCTGGAGTCTGTGAAGGAAATCCAAATATGGAGATAGCTAAATAC TTCTTTGGAATATCTTTAAACCAATAAAGGCCAGAAATTTGCTGGAGATTTGACA GTTAAAGCTATGAGAGCTTTAGAGATTTTATTTAAATAAGGATTTCAATCAATG GATTTAAATAATGCTGAGCTGAGAGCTTATAAGATTTTAGAGCCATTTAGAAAG AGATTTATAA ATGAGACGAAATTTGAATGATAAAGGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTCTCAATAGAGTTTGAACCAAGT GGTAAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTGATTTACAAAT GCTGGATTTGATATAATTTATTTGTTGCTGATTTACAGGCTTATTAACAGGAA GGAGAGTTGGATGAGCTTGAATAATAGAGATTTATTAACAAAGTTTGTAGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTTAT ACATCTGAATCTCTATGATTTGCTTTAAATACTACCTTAAATAAGAGCAGAGAGAGT ATCGAGCTTTATGCAAGAGAGAGGAAATCAAGAGTTCTTGAAGTTTCTATCTCA ATATCCAGGTTTATGATTTATCTTTTAGGCGTTGATGTTGAGCTTGGAGGGGATG GAGCAGAGAAATACACATGTTAGCAGAGGAGGTTTACCAAAAGAGGTTGTTTGT ATTCACACCTCTGCTTAAAGGCTTTGATGAGAGAGAGAGATGATTTCTCAAA GGAGATTTTCTCTCTCTGATTTGCTTCCAGAGAGGTTTACCTTGAAGCAAGAA GCATCTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATATGGAGATAGCTAAATAC TTCTCTGATATCTTTTAAACATAAAGGCCAGAAATTTGGTGGAGATTTGACA GTTAAAGCTTTGAGAGGTTTGAAGGTTTATTTAAATAAGAGATTTCAATCAATG GATTTAAATAATGCTGAGCTGAGAGCTTATAAGATTTTAGAGCCATTTAGAAAG AGATTTATAA	Mutant synthetase to incorporate m-methoxy phenylalanine into proteins (OMe 1-8)	RS
24	ATGAGACGAAATTTGAATGATAAAGGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTCTCAATAGAGTTTGAACCAAGT GGTAAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTGATTTACAAAT GCTGGATTTGATATAATTTATTTGTTGCTGATTTACAGGCTTATTAACAGGAA GGAGAGTTGGATGAGCTTGAATAATAGAGATTTATTAACAAAGTTTGTAGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTTAT ACATCTGAATCTCTATGATTTGCTTTAAATACTACCTTAAATAAGAGCAGAGAGAGT ATCGAGCTTTATGCAAGAGAGAGGAAATCAAGAGTTCTTGAAGTTTCTATCTCA ATATCCAGGTTTATGATTTATCTTTTAGGCGTTGATGTTGAGCTTGGAGGGGATG GAGCAGAGAAATACACATGTTAGCAGAGGAGGTTTACCAAAAGAGGTTGTTTGT ATTCACACCTCTGCTTAAAGGCTTTGATGAGAGAGAGATGATTTCTCAAA GGAGATTTTATGCTCTTATGACTCTCCAGAGAGATTTAGGCTTGAATTAAGAA GCATCTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATATGGAGATAGCTAAATAC TTCTCTGATATCTTTTAAACATAAAGGCCAGAAATTTGGTGGAGATTTGACA GTTAAAGCTTTGAGAGGTTTGAAGGTTTATTTAAATAAGAGATTTCAATCAATG GATTTAAATAATGCTGAGCTGAGAGCTTATAAGATTTTAGAGCCATTTAGAAAG AGATTTATAA	Mutant synthetase to incorporate m-methoxy phenylalanine into proteins (OMe 2-7)	RS
25	ATGAGACGAAATTTGAATGATAAAGGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTCTCAATAGAGTTTGAACCAAGT GGTAAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTGATTTACAAAT GCTGGATTTGATATAATTTATTTGTTGCTGATTTACAGGCTTATTAACAGGAA GGAGAGTTGGATGAGCTTGAATAATAGAGATTTATTAACAAAGTTTGTAGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTTAT ACATCTGAATCTCTATGATTTGCTTTAAATACTACCTTAAATAAGAGCAGAGAGAGT ATCGAGCTTTATGCAAGAGAGAGGAAATCAAGAGTTCTTGAAGTTTCTATCTCA ATATCCAGGTTTATGATTTATCTTTTAGGCGTTGATGTTGAGCTTGGAGGGGATG GAGCAGAGAAATACACATGTTAGCAGAGGAGGTTTACCAAAAGAGGTTGTTTGT ATTCACACCTCTGCTTAAAGGCTTTGATGAGAGAGAGATGATTTCTCAAA GGAGATTTTATGCTCTTATGACTCTCCAGAGAGATTTAGGCTTGAATTAAGAA GCATCTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATATGGAGATAGCTAAATAC TTCTCTGATATCTTTTAAACATAAAGGCCAGAAATTTGGTGGAGATTTGACA GTTAAAGCTTTGAGAGGTTTGAAGGTTTATTTAAATAAGAGATTTCAATCAATG GATTTAAATAATGCTGAGCTGAGAGCTTATAAGATTTTAGAGCCATTTAGAAAG AGATTTATAA	Mutant synthetase to incorporate m-methoxy phenylalanine into proteins (OMe 4-1)	RS
26	ATGAGACGAAATTTGAATGATAAAGGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTCTCAATAGAGTTTGAACCAAGT GGTAAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTGATTTACAAAT GCTGGATTTGATATAATTTATTTGTTGCTGATTTACAGGCTTATTAACAGGAA GGAGAGTTGGATGAGCTTGAATAATAGAGATTTATTAACAAAGTTTGTAGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTTAT ACATCTGAATCTCTATGATTTGCTTTAAATACTACCTTAAATAAGAGCAGAGAGAGT ATCGAGCTTTATGCAAGAGAGAGGAAATCAAGAGTTCTTGAAGTTTCTATCTCA ATATCCAGGTTTATGATTTATCTTTTAGGCGTTGATGTTGAGCTTGGAGGGGATG GAGCAGAGAAATACACATGTTAGCAGAGGAGGTTTACCAAAAGAGGTTGTTTGT ATTCACACCTCTGCTTAAAGGCTTTGATGAGAGAGAGATGATTTCTCAAA GGAGATTTTATGCTCTTATGACTCTCCAGAGAGATTTAGGCTTGAATTAAGAA GCATCTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATATGGAGATAGCTAAATAC TTCTCTGATATCTTTTAAACATAAAGGCCAGAAATTTGGTGGAGATTTGACA GTTAAAGCTTTGAGAGGTTTGAAGGTTTATTTAAATAAGAGATTTCAATCAATG GATTTAAATAATGCTGAGCTGAGAGCTTATAAGATTTTAGAGCCATTTAGAAAG AGATTTATAA	Mutant synthetase to incorporate m-methoxy phenylalanine into proteins (OMe 4-8)	RS
27	ATGAGACGAAATTTGAATGATAAAGGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAATAAGATGAAATCTCTCAATAGAGTTTGAACCAAGT GGTAAATACATTTAGGGCATTTATCTCCAAATAAATAAGATGATTTGATTTACAAAT GCTGGATTTGATATAATTTATTTGTTGCTGATTTACAGGCTTATTAACAGGAA GGAGAGTTGGATGAGCTTGAATAATAGAGATTTATTAACAAAGTTTGTAGAGCA ATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTTAT ACATCTGAATCTCTATGATTTGCTTTAAATACTACCTTAAATAAGAGCAGAGAGAGT ATCGAGCTTTATGCAAGAGAGAGGAAATCAAGAGTTCTTGAAGTTTCTATCTCA ATATCCAGGTTTATGATTTATCTTTTAGGCGTTGATGTTGAGCTTGGAGGGGATG GAGCAGAGAAATACACATGTTAGCAGAGGAGGTTTACCAAAAGAGGTTGTTTGT ATTCACACCTCTGCTTAAAGGCTTTGATGAGAGAGAGATGATTTCTCAAA GGAGATTTTATGCTCTTATGACTCTCCAGAGAGATTTAGGCTTGAATTAAGAA GCATCTGCCAGCTGGAGTTGTTGAAGGAAATCCAAATATGGAGATAGCTAAATAC TTCTCTGATATCTTTTAAACATAAAGGCCAGAAATTTGGTGGAGATTTGACA GTTAAAGCTTTGAGAGGTTTGAAGGTTTATTTAAATAAGAGATTTCAATCAATG GATTTAAATAATGCTGAGCTGAGAGCTTATAAGATTTTAGAGCCATTTAGAAAG AGATTTATAA	Mutant synthetase to incorporate p-O-allyl tyrosine into proteins (Allyl)	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
28	ATTAATGAGGTTTAATTGCGACATTAATTAGGCTTGAATGTCAGATTGGAGGGATG GAGGAGAGGAAATATACACATCTTACCGAGGAGCTTTACCAAAAGGTTGTGT ATTCACACCCCTGCTTAACGGGTTTGGATGGAGGAGAAAGATGATTCCTCAAAA GGGAATTTTATAGCTGTGAGAGCTTCCAGAGAGATTTAGGCTAGATTAAGGAAA GCTATCTCCCGACGTGAGCTTTGGAGAGAAATCCATTAATGGAGATGGTAAATG TTCTTGAATATCTTTAAACCAAAAAGGCCAGAAAATTTGGTGGAGATTTGACA GTTAATAGCTATGAGAGATTTAGAGGTTTATTAATAAATAGGAATTCATCCAAATG GNTTTAAAAATGCTGTAGCTGAGAACTTTAAAGATTTTGAAGCCATTAGAAAG AGATTATA ATGGAGCAATTTGAATGATAAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAAGAAAGATGAATAATCTCCGAGATAGGTTTGAACCAAGT GCTGAATATACATTTAGGCGATTTATCTCCAAATAAAGATGATTTGATTAACAAAT GCTGAGTTGATATATTAATTTGTGTGCTGATTTACACGCTATTTAAACAGAAA GGAGATTTGGATGAGATTAGAAAATAGGAGATTATACAAATAAGTTTGTGAGCA ATGGGTTTAAAGCAAAATATTTATGGAATTCCTTCCAGCTTGTATAGAGATAT ACACTGAATGCTATAGATTGCTTTAAAACTACCTTAAAGAGCAAGAGAGAGT ATGGAATTTATGAGAGAGAGATGAATAATCCAAAGTTGCTGAAGTTATCTATCCAA TATGCACTTAAACAGAGCAATTTATGCGCTTGAAGTTTGCAGTTGGAGAGTGG AGCAGAGAAAATACACATCTTACCAAGGAGCTTTTACCAAAAAGGTTGTGTGTA TTTACAAACCTGCTTAAACGGTTTGGATGGAGAGAGGAGATGAGTTCTTCAAAAG CGAATTTATAGCTGTTGATGAGCTTCCAGAGAGATTTAGGCTTAAAGATAGAGAG CATACTGCCAGCTGGAGTTCTTGAAGGAAATCCAAATAAGGAGATAGTAAATAT TCTTGAATATCTTTAAACGATAAAGAGCCAGAAAATTTGGTGGAGATTTGAGAG TTAATAGCTTGTAGAGATTTAGAGGTTTATTAATAAATAGGAATTCATCCAAATG ATTTAAAAAATGCTGTAGCTGAGAGCTTTAAGAGTTTAAAGCAATTTGAAGAG GATTA	Aminoacyl tRNA synthetase for the incorporation of p-benzoyl-L-phenylalanine (p-BpaRS(H6))	RS
29	ATGGAGCAATTTGAATGATAAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAAGAAAGATGAATAATCTCCGAGATAGGTTTGAACCAAGT GCTGAATATACATTTAGGCGATTTATCTCCAAATAAAGATGATTTGATTAACAAAT GCTGAGTTGATATATTAATTTGTGTGCTGATTTACACGCTATTTAAACAGAAA GGAGATTTGGATGAGATTAGAAAATAGGAGATTATACAAATAAGTTTGTGAGCA ATGGGTTTAAAGCAAAATATTTATGGAATTCCTTCCAGCTTGTATAGAGATAT ACACTGAATGCTTATAGATTGCTTTAAAGATCTCTTAAAGAGCAAGAGAGAGT ATGGAATTTATGAGAGAGAGATGAATAATCCAAAGTTGCTGAAGTTATCTATCCAA TATGCACTTAAACAGAGCAATTTATGCGCTTGAAGTTTGCAGTTGGAGAGTGG AGCAGAGAAAATACACATCTTACCAAGGAGCTTTTACCAAAAAGGTTGTGTGTA TTTACAAACCTGCTTAAACGGTTTGGATGGAGAGAGGAGATGAGTTCTTCAAAAG CGAATTTATAGCTGTTGATGAGCTTCCAGAGAGATTTAGGCTTAAAGATAGAGAG CATACTGCCAGCTGGAGTTCTTGAAGGAAATCCAAATAAGGAGATAGTAAATAT TCTTGAATATCTTTAAACGATAAAGAGCCAGAAAATTTGGTGGAGATTTGAGAG TTAATAGCTTGTAGAGATTTAGAGGTTTATTAATAAATAGGAATTCATCCAAATG ATTTAAAAAATGCTGTAGCTGAGAGCTTTAAGAGTTTAAAGCAATTTGAAGAG GATTA	Aminoacyl tRNA synthetase for the incorporation of p-azido-phenylalanine (p-Az-PheRS(3))	RS
30	ATGGAGCAATTTGAATGATAAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAAGAAAGATGAATAATCTCCGAGATAGGTTTGAACCAAGT GCTGAATATACATTTAGGCGATTTATCTCCAAATAAAGATGATTTGATTTACAAAT GCTGAGTTGATATATTAATTTGTGTGCTGATTTACACGCTATTTAAACAGAAA GGAGATTTGGATGAGATTAGAAAATAGGAGATTATACAAATAAGTTTGTGAGCA ATGGGTTTAAAGCAAAATATTTATGGAATTCCTTCCAGCTTGTATAGAGATAT ACACTGAATGCTTATAGATTGCTTTAAAGATCTCTTAAAGAGCAAGAGAGAGT ATGGAATTTATGAGAGAGAGATGAATAATCCAAAGTTGCTGAAGTTATCTATCCAA TATGCACTTAAACAGAGCAATTTATGCGCTTGAAGTTTGCAGTTGGAGAGTGG AGCAGAGAAAATACACATCTTACCAAGGAGCTTTTACCAAAAAGGTTGTGTGTA TTTACAAACCTGCTTAAACGGTTTGGATGGAGAGAGGAGATGAGTTCTTCAAAAG CGAATTTATAGCTGTTGATGAGCTTCCAGAGAGATTTAGGCTTAAAGATAGAGAG CATACTGCCAGCTGGAGTTCTTGAAGGAAATCCAAATAAGGAGATAGTAAATAT TCTTGAATATCTTTAAACGATAAAGAGCCAGAAAATTTGGTGGAGATTTGAGAG TTAATAGCTTGTAGAGATTTAGAGGTTTATTAATAAATAGGAATTCATCCAAATG ATTTAAAAAATGCTGTAGCTGAGAGCTTTAAGAGTTTAAAGCAATTTGAAGAG GATTA	Aminoacyl tRNA synthetase for the incorporation of p-azido-phenylalanine (p-Az-PheRS(6))	RS
31	ATGGAGCAATTTGAATGATAAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAAGAAAGATGAATAATCTCCGAGATAGGTTTGAACCAAGT GCTGAATATACATTTAGGCGATTTATCTCCAAATAAAGATGATTTGATTTACAAAT GCTGAGTTGATATATTAATTTGTGTGCTGATTTACACGCTATTTAAACAGAAA GGAGATTTGGATGAGATTAGAAAATAGGAGATTATACAAATAAGTTTGTGAGCA ATGGGTTTAAAGCAAAATATTTATGGAATTCCTTCCAGCTTGTATAGAGATAT ACACTGAATGCTTATAGATTGCTTTAAAGATCTCTTAAAGAGCAAGAGAGAGT ATGGAATTTATGAGAGAGAGATGAATAATCCAAAGTTGCTGAAGTTATCTATCCAA TATGCACTTAAACAGAGCAATTTATGCGCTTGAAGTTTGCAGTTGGAGAGTGG AGCAGAGAAAATACACATCTTACCAAGGAGCTTTTACCAAAAAGGTTGTGTGTA TTTACAAACCTGCTTAAACGGTTTGGATGGAGAGAGGAGATGAGTTCTTCAAAAG CGAATTTATAGCTGTTGATGAGCTTCCAGAGAGATTTAGGCTTAAAGATAGAGAG CATACTGCCAGCTGGAGTTCTTGAAGGAAATCCAAATAAGGAGATAGTAAATAT TCTTGAATATCTTTAAACGATAAAGAGCCAGAAAATTTGGTGGAGATTTGAGAG TTAATAGCTATGAGAGTTAGAGGTTTATTAATAAATAGGAATTCATCCAAATG ATTTAAAAAATGCTGTAGCTGAGAGCTTTAAGAGTTTAAAGCAATTTGAAGAG GATTA	Aminoacyl tRNA synthetase for the incorporation of p-azido-phenylalanine (p-Az-PheRS(20))	RS
32	ATGGAGCAATTTGAATGATAAAGAGAAACACATCTGAATTTATCAGCGAGGAGAG TTAAGAGAGGTTTAAAGAAAGATGAATAATCTCCGAGATAGGTTTGAACCAAGT GCTGAATATACATTTAGGCGATTTATCTCCAAATAAAGATGATTTGATTTACAAAT GCTGAGTTGATATATTAATTTGTGTGCTGATTTACACGCTATTTAAACAGAAA GGAGATTTGGATGAGATTAGAAAATAGGAGATTATACAAATAAGTTTGTGAGCA ATGGGTTTAAAGCAAAATATTTATGGAATTCCTTCCAGCTTGTATAGAGATAT ACACTGAATGCTTATAGATTGCTTTAAAGATCTCTTAAAGAGCAAGAGAGAGT ATGGAATTTATGAGAGAGAGATGAATAATCCAAAGTTGCTGAAGTTATCTATCCAA TATGCACTTAAACAGAGCAATTTATGCGCTTGAAGTTTGCAGTTGGAGAGTGG AGCAGAGAAAATACACATCTTACCAAGGAGCTTTTACCAAAAAGGTTGTGTGTA TTTACAAACCTGCTTAAACGGTTTGGATGGAGAGAGGAGATGAGTTCTTCAAAAG CGAATTTATAGCTGTTGATGAGCTTCCAGAGAGATTTAGGCTTAAAGATAGAGAG CATACTGCCAGCTGGAGTTCTTGAAGGAAATCCAAATAAGGAGATAGTAAATAT TCTTGAATATCTTTAAACGATAAAGAGCCAGAAAATTTGGTGGAGATTTGAGAG TTAATAGCTATGAGAGTTAGAGGTTTATTAATAAATAGGAATTCATCCAAATG ATTTAAAAAATGCTGTAGCTGAGAGCTTTAAGAGTTTAAAGCAATTTGAAGAG GATTA	Aminoacyl tRNA synthetase for the incorporation of p-azido-phenylalanine (p-	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
	ATCGGCGTTAAAGCAAAATATGTTTATGGAGTTCTCCAGCTTGAATAGGATTAT ACCTCGAATGTTTATGATTTGCTTTTAAATCTCTTTAAATAGAGCGAGGAGT ATGGACTTATATAGAGAGAGATGAAATCCAAAGGTGCTGAAGTTATCTATCCAA TATGACAGTTTATCCACTGATTTATCAGGCGCTTGATGTTGCACTTGGAGGATG AGCGAGAAATATACATGTTGACAGAGGAGCTTTTACGAAAGAGCTTTGTTA TTCCAAACCCCTCTTAAACGGTTTGGATGGAGAGGAAGATGATCTTCAAAAG GGAAATTTATAGCTGTTGATGACTCTCCAGAGAGATAGGCTAAGATAAAGAAAG CACTGTGCCCACTGGAGCTTTTGGAGAGAACTCAATATAGGATAGCTTAAGTACT TCCTTGAATATCCCTTAAACATAAAGGCCAGAAATCTGGTGGAGTTTGAACG TTATATAGCTATGAGGAGTTAGAGAGTTTATTTAAATAAAGAAATGCAATCCAAATG ACTTAAAAATGCTGTAGCTGAGAACTTATAAGATTTAGAGCAATAGAAAGA GATTA	Az-PheRS(24)	
33	ATGAGCGATT TCAGGATAT TTGGGGAAG TGGCAGAAG CGTGGAGAA GGACAGATT TTGGATCCG ATCTATATGA GAAGGAGAG TTTTCTCTCA CAATCTCTA TCTTTACTT ATGGAAATC TCGACGCGA TGCACAGGA ACCTTCACAA TTGGCGATGC CTGCGCAGA TACATGAGA TGAAGGGCTA CAACGTTCTC TTTCCTCTG GCTTTCATGT TACGGGACCC CCAATCATTC CGCTTTCGCG CCTGATGAC AAGAGGAGG AGAGACCAT AGAGGTTTAC ACCAATATCC ATGACGTTCC GCTGGAGGAC TTGCTTCAGC TCACACTCC AGGAAATATC GTTGAATACT TCTCAAGGA GCGCTGAGG GCTTGAAGA CGATAGCTT CTCCATGAC TGGAGAGGG TTTTCAACAC AACCAATGA GAGTATCAGA GATTATCTGA GTGGCAGTAC TGAAGCTCA AGGACTTGG CGTGAATGTC AAGGGCACCC ACCGCTCAG ATACTGCCCC CAGACACAGA ATCTGTTGTA AGACACGAC CTCTCTCTG GGGAGAGGC AACTATTGTT GAAATTCACG TTATTAAGGT CAGCTTGGA GATGAGACT TCAATTTCT CTGTCAACT CTCCGTCCG AAACCGTGT TTGAGTCAAG AACATCTGG TAAAGCCGAC AACCTACCTA ATTCGCGAG TGAATGGGA AAGGTGTTT CTGAGCAGG AGCTTACGA GATCTCTCC TACACCGGGA AAAAGCTAG CTGTCTGGAG GAGTTGATG CTTCGAGTT CTTCGCGAAG TACGTCAATG TCCGCTGCTT AAACGAAATA GTGCGCATTC TGCTCTCAGA GTTGTGTGAC ACGACACATC CAACAGGATC TTGTGAGAG GTTCCCGGAC AGCTCTCTT TGACTCTGCT GCAATGAGG ACTTGAAGAG AGACGAGGA ACCTGSGGA AGTACCGAAT TGACAAAGC GTTGTAGAGA GCATTAAGCC AATATCTCTG ACTAAGAGCG ACATTAAGAG TTCTCTGCT GAGAGCTTAA TAAAGAGCT TGGATGTAAG AGCCAGAAAG ACAAGGAGT GCTGTAATAG GCAACCAAGA CCCTCTACAA GAAGGAGTAC CACACGGAA TCATGCTGGA CAACACATG AATATCTCTG GATTAAGAT TTCTGAGCG AGGWAAGAG TTCTATGAGA TTTGTATTAAG CTGCTCTGG GGAATTTCT CACAGATTC AGCGAAGCC CCATATCTG CAGTTCGGA AGGAGTGGG TTGTAAAGT TGTTAGGAC CAGTGTCTTC TGAATCTCT CACACAGAG TGGAAAGGA AGTTCCTGAA TCACTCTGGA AGGTGAGA TATCTCTGGA CTACTACAG GAGAGTCTA GGAACAGAT TTGATGGCTC AGGACAGAG CTTGTGCGAG AAGGAAGGG CTTGGACGA GAATTCCTG GATTAAGAG TGGTCACTG AGAGCTTTTC AGATCTACCA ATCTGATGCT CTACTCTCTC CTCTGCGAG TACATCAAG CAGGATTCCT CAAGCGCGAG AACATGACTC CCGAGTCTCT GACTACGTC CTGCTGAGCA AAGTGAAGT TGGGAAGTT GCGGAAGCTT CAATACTGAG CTGTGAATTA ATCCAGCAG TCGAGGAGG CTTCAGTAC AGATATCCG TTGACCTAAG AAGCAGTGGC TACCTCTCC CATTCTCCG CCAGATAAGT GCGCGAGGCG TACCTCTTCC ACCAGTCCG TCAAGCTTGA GGGCAAGAG ATTAGCAGAG AATTCCTTA AAGCGATGAC ATGAGAGAGG CCGTCCAGCA ATATGCTGCG GATGTACGA GCTCTACAT CCTCCAGCT GCAGATACG ACAGGATCC GACTTGAAG AGCAGAGAG TTGAAGGCT TGAACACAC CTCAAGAGT TTTCACACT CTCAAGAGG AATCTCTCA AGAGCTTGG AGACTTACA ACCTCGACC GCTGCTCTT GAGCAGATG CAGAGGCGAA TAAAGGAAT GAGGAGAGCT ATGGAACAC TGCAGCAG TACGCGCTG AATGCGCTC TCTTCAGCT CATGAAGAG CTGATGAGT ATCTGCGAG AGGAGTGGG AACCTCGCTA TAATCTGGA GACTGATC AAGCTCTCTG CCGCTTTGCG TCGACACATT TCGAGGAGG GTTGGCATC GAGGATGAG AGTACGCTCA GCTTCAGAG CTACTAGAA TACGACAAA CCGGTTTGA CAGAGGCGG GAGGAATTT AGGATACCT CCGAAACCT GTTGAAGACA TTTCAAGAAAT CAGAGAGTT GTTACGATG CGAAAGAGT TTACATTGCT CCGCGCGAAG ACTGAGAGT TAAAGGAGA AAGTCTCTG CTGAAGCGG GAGTCTGGG GAGCGATGA AGCAGCTTAT GCGAGACGAG GAGCTTAGGA AGTCTGAAA AGAAGTGTCA AATTCTGCTA AGAAGATTTT CAAGACAGA AAGAAGCTGA TGTGATTTAA GAGTGGGAA GTTCTGAGG AGAACCTGAA ATTTATTGAG ATGAGACCG GACTGAAGCT TATCTCTGAT ACTCAGAGG TTCTCGAGA GAGAGAGAG CAGGAGTTTC GGGCGAGCC GCGGATTTT GTTCTTAA GTGATATTG AAGGAAATG GCGTATAGA TGGAGAGATG CTGGCAATTT TCAAGTCTT CTTATGAGA GAGAAAGAT ATCTCTACA GCGCTTACG CTTACCCAG TGTTCGATG CACATAGGAC ACGGAGGAC CTACACTGTC CTGATGCTT ATGACAGTTC CAGAGAGATG CAGGCTTACA ACGTCTCTTT TCTC TGGCC TGGCATTTCT TACGAGGCC TGTCTTAGG ATAGCGGGA GATTCAGAG GAAGATCCC TGGACCTCA AATCTACAG GAGGTTCAC AGGCTCCCG AGATGAGCT TGAAGCTTC AGTGACCTG AGTACATAT TGAATACTT AGCAGGAT ACCGCTCTG TATGAGAGT ATGGCTACT CMTGCAATG GAGCGTAAA TCCAAACCT CAGATCTAC CTACAGAGG TTCATACAT GGCAGATAAG GAAGCTGAGG GACCTTGCTC TCTAAGGAA GAGGCCCAT CTTGTAAAGT ACTGCGCTGA ATGTGAAGC CTTCTGAGT ACCTTGTCT CTTTGAAGT GAGGCTGCTG CCAATACCA GCTCACCTC CTCAATTTCA AACTTGGAGA CTCACACTG GTGCGACCA CTTTCAAGCC GAGAGAAATC TATGGGCGA CCAACCTCTG GCTGAACCTT GATGAGGAT	Archaeoglobus fulgidus leucyl tRNA-synthetase (AFLRS)	RS
34	GTGATATTG AAGGAAATG GCGTATAGA TGGAGAGATG CTGGCAATTT TCAAGTCTT CTTATGAGA GAGAAAGAT ATCTCTACA GCGCTTACG CTTACCCAG TGTTCGATG CACATAGGAC ACGGAGGAC CTACACTGTC CTGATGCTT ATGACAGTTC CAGAGAGATG CAGGCTTACA ACGTCTCTTT TCTC TGGCC TGGCATTTCT TACGAGGCC TGTCTTAGG ATAGCGGGA GATTCAGAG GAAGATCCC TGGACCTCA AATCTACAG GAGGTTCAC AGGCTCCCG AGATGAGCT TGAAGCTTC AGTGACCTG AGTACATAT TGAATACTT AGCAGGAT ACCGCTCTG TATGAGAGT ATGGCTACT CMTGCAATG GAGCGTAAA TCCAAACCT CAGATCTAC CTACAGAGG TTCATACAT GGCAGATAAG GAAGCTGAGG GACCTTGCTC TCTAAGGAA GAGGCCCAT CTTGTAAAGT ACTGCGCTGA ATGTGAAGC CTTCTGAGT ACCTTGTCT CTTTGAAGT GAGGCTGCTG CCAATACCA GCTCACCTC CTCAATTTCA AACTTGGAGA CTCACACTG GTGCGACCA CTTTCAAGCC GAGAGAAATC TATGGGCGA CCAACCTCTG GCTGAACCTT GATGAGGAT	Methanobacterium thermoautotrophicum leucyl tRNA-synthetase (MLRS)	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
	ATGTGAGGTT TGAACAGGTT GTTGAAGGTT GGATAAAGA CAGGCGTGGC GTTGATATTC TTTTACACCA GAAATCTGAC CTCAGAGGTT CCGGTGAGCT CAACCCCGGG GACCTGATAG GGAATGTGCTT GGAGAAATCT GTACGCGGCT AGGAACACCC CATTACTCCG GCTTCTCTTG TTGACCTGTA ATATGCCACA GGTGTGTGTT TCTCTGTCCT TGACACACCC CTTGACGACT TCAATAGCCCT TGAGGACCTC AGGACAGACC ATGAATCTCT TGAATAGGTA GATCTTGGG ATGTGCTTGC TGAATATTGAG CCGGTGAATG TCATAGAGAT GATGGCTTAC GGTGAATCTC GGGCGGCGCA GGTATATAGG AATATTGTGT TCAGAAACCA GGAAGACCTC GCTCTTGAGC ATGCTATGCG GGAAGCTGAC AGATGCGG ATGCGAGGGG TGTATAGAGC AGCCACATCC CTGTCTATGG AGTATAGAG GTCTTGTAGG CCGGTGAGGT CATGCTGTAT GAATGAGAG ACCAGGCGCT TGCAGTGG AGTATGAT TCGCTGAGG ACCTCTTADA TGGCGGTGGC GTGCGAGGTG GGTGTGAGG GTCATGAGG ACCAGTGTCT CATGAAGTAC TCTGATGAG CCGTGAAGGA CTTGCGCCAC AGTGTGCTCG ATGGCATGAA GATATATCC GAGAGGTC GGGCAGCTT TGAATATGAC ATGCGCTGCG TCAATGACTG GGCATGTTCA AGGAGATAG GCTTGAAC AAGCTGCGC TGGATGAGA GGTGATCAT CGAACCCCTC ACAGACTCAA CAATCTACAT GCTATATGAC ACATGCGAC ACCGCTCAG GGAATGATGAT GCGGAGAGA TGGACGATGA GTTCTTGTAT GGCATATCCC TGAATGATG AGCACTCTT GAGGATCTCA GAGAGGATTT CCGTACTGCT TACCCCTTGT ACTGAGGCT CTCTCAAG GACCTCATAG GCAATCACT GACATTCAT ATATTCACCT ACCAAGCTAT ATCTCTGAG TCAAGGTGCG CCGGAGGCT TGTGCTCTTT GATATGAGCT TTTCTGAGGG CAACAAGATG TCATCTCTCA AGGCAACCT CATATCTCTG AGGATGATCA TGAGAGAGCA CGTCTCAGAC GTGCTGCGG TCTCTCTCAT GTCTCTGACA GAGCAATGCG AGGACTTCTG CTGAGAGGAG AGTGAAGTCA TCGGACCCCG CAGGAGGAT GAATGTTCA GGAATTCGG AGAGAGGCTC TCAAGTATCC TGAATGATAG GCCAGCTCTC AGTGAAGTTA CTGCACTGTA ACTGAAAC TTATGAGGA GTGAGATGAT GGTGAGCTG AACCAAGGA TACGTGAAC CAACAAGGCC CTTGATCAT TCCAGACAAG AAGGCACTT CAGGAGGAC TCTATCTCTT TAAAGAGAT GTTGACCAT ACCTGACCTG TGTGAGGAT AGAGGTGAG ATGAGTTAA ATCTGTCTT GCAAACTTTC TGAAGCTCTG GATTAAGCTC ATGCTCTCAT TCAACCTTA CACTGCTGAG GAGATGTGG AGAGGTATG TGTGAGGCT TTTGTAGCAG AGCTGCAAT GCTGACTCTC TGAGATGATG CAGAGAGGCG GATATGCG GTCTGAGGAG AGATGCTCCA GATACACTCT AGAGACTCTC AGAATCAT GAAGATCTTT GATTCACCC CAGAGAGGCT CCACATATAC ACCTACCAA AATGAGATG GATATGCTA AGGTCTGAG CAGAGGTAG AAACTAGAT AGGCTCTCA TAAKSGAG GCTTCACT GAGGCTCTC AGATATCAT GAGAGGCTT GCTGAATTTG TAAAGAGGAT CATCAGGAG CTGTGAAAT CAGAGGTAG GGTGATGAG GATACAGGCT TACTCATGTA TGCATCTGAT TCAATGAT CAGAGGTGAG AGTCAAGGCT GTGATACCA GCACACAGA CTATGACCTT GAAACAGG CTGTAAGTGC CGTCTCTCTG AAGCAGCCA TATACCTTGA AGA		
35	MDPFTIKRNTSII SEBELREVLRVKKDEKSAQIGFEPGKTHLGHVLTQKMDLQ AGFDITLLADLHAYLNQKGLDETRKIGDYNKKVPEAMGLKAKYVYGSFQLDKDY TLNVTRALAKTTLKRAKRSMEIAREDENPKVAEVIYPTMGVNPVHYAGVDVAVGGM EQKTHMLARELLPKVVCINHPVLTGLDGGKMSKGNFIAVDSPEIRAKIKK AYCPAGVVEQNPIMETAKYFLVPLTLKRPKFGGDLTVNSYELESLEFNKELHFM DEKNVAEELIKLEPFRRL	mutant TyrRS (LWJ16)	RS
36	MDPFTIKRNTSII SEBELREVLRVKKDEKSAQIGFEPGKTHLGHVLTQKMDLQ AGFDITLLADLHAYLNQKGLDETRKIGDYNKKVPEAMGLKAKYVYGSFQLDKDY TLNVTRALAKTTLKRAKRSMEIAREDENPKVAEVIYPTMGVNPVHYAGVDVAVGGM EQKTHMLARELLPKVVCINHPVLTGLDGGKMSKGNFIAVDSPEIRAKIKK AYCPAGVVEQNPIMETAKYFLVPLTLKRPKFGGDLTVNSYELESLEFNKELHFM DEKNVAEELIKLEPFRRL	TyrRS (SS12)	RS
37	MDPFTIKRNTSII SEBELREVLRVKKDEKSAQIGFEPGKTHLGHVLTQKMDLQ AGFDITLLADLHAYLNQKGLDETRKIGDYNKKVPEAMGLKAKYVYGSFQLDKDY TLNVTRALAKTTLKRAKRSMEIAREDENPKVAEVIYPTMGVNPVHYAGVDVAVGGM EQKTHMLARELLPKVVCINHPVLTGLDGGKMSKGNFIAVDSPEIRAKIKK AYCPAGVVEQNPIMETAKYFLVPLTLKRPKFGGDLTVNSYELESLEFNKELHFM DEKNVAEELIKLEPFRRL	p-iPr-PheRS	RS
38	MDPFTIKRNTSII SEBELREVLRVKKDEKSAQIGFEPGKTHLGHVLTQKMDLQ AGFDITLLADLHAYLNQKGLDETRKIGDYNKKVPEAMGLKAKYVYGSFQLDKDY TLNVTRALAKTTLKRAKRSMEIAREDENPKVAEVIYPTMGVNPVHYAGVDVAVGGM EQKTHMLARELLPKVVCINHPVLTGLDGGKMSKGNFIAVDSPEIRAKIKK AYCPAGVVEQNPIMETAKYFLVPLTLKRPKFGGDLTVNSYELESLEFNKELHFM DEKNVAEELIKLEPFRRL	p-NH ₂ -PheRS(1)	RS
39	MDPFTIKRNTSII SEBELREVLRVKKDEKSAQIGFEPGKTHLGHVLTQKMDLQ AGFDITLLADLHAYLNQKGLDETRKIGDYNKKVPEAMGLKAKYVYGSFQLDKDY TLNVTRALAKTTLKRAKRSMEIAREDENPKVAEVIYPTMGVNPVHYAGVDVAVGGM EQKTHMLARELLPKVVCINHPVLTGLDGGKMSKGNFIAVDSPEIRAKIKK AYCPAGVVEQNPIMETAKYFLVPLTLKRPKFGGDLTVNSYELESLEFNKELHFM DEKNVAEELIKLEPFRRL	p-NH ₂ -PheRS(2)	RS
40	MDPFTIKRNTSII SEBELREVLRVKKDEKSAQIGFEPGKTHLGHVLTQKMDLQ AGFDITLLADLHAYLNQKGLDETRKIGDYNKKVPEAMGLKAKYVYGSFQLDKDY TLNVTRALAKTTLKRAKRSMEIAREDENPKVAEVIYPTMGVNPVHYAGVDVAVGGM EQKTHMLARELLPKVVCINHPVLTGLDGGKMSKGNFIAVDSPEIRAKIKK AYCPAGVVEQNPIMETAKYFLVPLTLKRPKFGGDLTVNSYELESLEFNKELHFM DEKNVAEELIKLEPFRRL	p-NH ₂ -PheRS(3a)	RS
41	MDPFTIKRNTSII SEBELREVLRVKKDEKSAQIGFEPGKTHLGHVLTQKMDLQ AGFDITLLADLHAYLNQKGLDETRKIGDYNKKVPEAMGLKAKYVYGSFQLDKDY TLNVTRALAKTTLKRAKRSMEIAREDENPKVAEVIYPTMGVNPVHYAGVDVAVGGM EQKTHMLARELLPKVVCINHPVLTGLDGGKMSKGNFIAVDSPEIRAKIKK AYCPAGVVEQNPIMETAKYFLVPLTLKRPKFGGDLTVNSYELESLEFNKELHFM DEKNVAEELIKLEPFRRL	p-NH ₂ -PheRS(3b)	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
42	AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL# MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	O-Allyl-TyrRS(1)	RS
43	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	O-Allyl-TyrRS(3)	RS
44	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	O-Allyl-TyrRS(4)	RS
45	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	p-Br-PheRS	RS
46	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	p-Az-PheRS(1)	RS
47	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	p-Az-PheRS(3)	RS
48	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	p-Az-PheRS(5)	RS
49	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	Mutant synthetase to incorporate m-acyl phenylalanine into proteins (Ketone 3-4)	RS
50	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	Mutant synthetase to incorporate m-acyl phenylalanine into proteins (Ketone 3-7)	RS
51	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	Mutant synthetase to incorporate m-acyl phenylalanine into proteins (Ketone 4-1)	RS
52	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	Mutant synthetase to incorporate m-acyl phenylalanine into proteins (Ketone 5-4)	RS
53	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	Mutant synthetase to incorporate m-acyl phenylalanine into proteins (Ketone 6-8)	RS
54	MDEPFMIKRNSETISSEELREVLKKDEKSAITGPFPSGKTHLGHYLQIKMIDLQN AGFDIILLADLHAYLNQKGEIDETRIKIDYNNKVFPAAGLAKAYVGSFQLDKDY TLNVTALAKTTLKARRSMELIAREDENPKVAEVIPTMQVNDHYLGVDAVGGM BQRKTHMLARELLPKVVCVTHNPVLTGLDGGKMSSSKGNFIAVDSPEETRAKIKK AYCPAGVVEGNPIMETAKYFLYPLTIKRPKFGGDLTVNSYSELESFKNKELHFM DLKNAVAELIKILPEIRKRL#	Mutant synthetase to incorporate m-methoxy phenylalanine into	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
	AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	proteins (Ome 1-6)	
55	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Mutant synthetase to incorporate m-methoxy phenylalanine into proteins (Ome 1-8)	RS
56	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Mutant synthetase to incorporate m-methoxy phenylalanine into proteins (Ome 2-7)	RS
57	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Mutant synthetase to incorporate m-methoxy phenylalanine into proteins (Ome 4-1)	RS
58	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Mutant synthetase to incorporate m-methoxy phenylalanine into proteins (Ome 4-8)	RS
59	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Mutant synthetase to incorporate p-O-allyl tyrosine into proteins (Allyl)	RS
60	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Aminoacyl tRNA synthetase for the incorporation of p-benzoyl-L-phenylalanine (p-BnaRS(H6))	RS
61	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Aminoacyl tRNA synthetase for the incorporation of p-azido-phenylalanine (p-Az-PheRS(3))	RS
62	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Aminoacyl tRNA synthetase for the incorporation of p-azido-phenylalanine (p-Az-PheRS(6))	RS
63	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Aminoacyl tRNA synthetase for the incorporation of p-azido-phenylalanine (p-Az-PheRS(20))	RS
64	MDSEFMIRKNTSEIISEELREVLKDKDEKSAIGFEPGSKIHGLHYLQIKKIDLDQN AGPDIIILLADLHAYLNKGELDEIRKIGDYNKVFPMGLKAKYVYGSFQLDKDY TLNVYRLAKTTTKRARRSMELIAREDENPKVAEVIYPIQVNDIHYLGVDVAVGGM EQKIKHMLARELLPKVVCINHPVLTGLDGEGBMSSKGNFIAVDDSPETRAKIKK AYCPAGVVEGNPIMETAKYFLEYPLTIRKPEKFGDGLTVNSYELESFKNKELHPM DLKNAVASELKIILEPTIRKRL#	Aminoacyl tRNA synthetase for the incorporation of p-azido-phenylalanine (p-Az-PheRS(24))	RS
65	NSDFRIIEEK WQKWEKORI FESDPNBAEK FFLTIPYPL NGLNHAGHTR FTFTGDFAFR YMRMGYNVL FPLGPHVTGT FIIGLAEILIA KRDEKTEPV TKRHDVPLED LLQUTPEKI VEIIFREALO ALEISIGYSID MRRVETITDE EYRPFLEMY WKRELGLIV KSHPEYCP HDQNPVEZHD ILAGSEACIV EFTVIRFRL DGDITPCAT LRSTVEGVT NIKVKPTIYV IAAVDKWP VSRAXEKL/ YTRKVRLL EVDASQFGK YIVVPLNRK VFILPAEVD TMNAGVMS VPIAPDIA AKEDKDEE TLAKYDIES VSRKTPYL IKFDIGVPA EKLIRELVK SQDKKELDK ATPTTYKCEY HIGIMLDNTM NVAGMKVBA KRVIEDLVK LGLGDFYEF SERFVICROG TKCVKCVVD CNFLATYSSE NKSEVLNLE KGLTTPYTK EEPNKLHGL HKAACARRG	Archaeoglobus fulgidus leucyl tna-synthetase (AFLRS)	RS

WO 02/086075

PCT/US02/12635

SEQ ID #	Appendix 1: Table 5: Sequences	Notes	tRNA or RS
66	LGTRIPMKRS WLRELSNST IYMAFYIAK YIMAGLLAAE NMTVEPLZTV LLKREWEKCV AEASKLVRL IQOICDEFEY WTVDLASSG KOLAVANGHLE YLFHHVATFP PDKMPRAIAV NGVVSLEGKK MSKSGGPLLT MGRAVQYQGA DVTRLIYLLHA AFDSDADWK SRVVEGLAH LKRFYNLVKS NYLKEVGSLT CLDFWVGRM QRAIKETVE MDNLQTEAV NNAFFELAND VRYLARGGE NLAILLDWI KILAPFABHI CEELNHLKHD STVSESYPS YDETRVDSEA ERIEEYLRL VEDIQIKKF VSDAKEVYIA PAEDWKVRAA KVVAESGDVG EAMQVMQDS ELKIKREVS NPVKIYKTR KILMLYKWS VLQNDKATIE NETGKYILD TQRPVEEKKR QAVPGKPAIV VA* VDIERKWRDR WRDAGIFQAD PDREKIFLT VAYFPFSGAM HIGHRTTIV PDYAFPKRM GGHNVLFMA MHVTSAPVIG IARRLQKCP WILKITERGVH RVPDELERF GDFEYVTF SRELSSTED MGSLLMRRS EKTQDFYSR FIOMQIRKLK DLGLVRGAH PVKYQPCEN PVSDRDLGG EGVAIMQULT LKFKLGDSYL VAITPRPET YGATNLWLF DSDVVRVEYG GERWISRAA VDNLSHGLD LKVSQVWNG DLGMCVENP ITQDHPITL ASFQREYAT GVVFSVPAMA PAFIALDEL RTDHELLERY GLEDVVDIE PVNVLAVDGY GEPFAEVIE KFGVRMQEDP RLEDATGSLY KISHARGVMS SHLPVVGEMX VSSAREVTAD ELKQCLADE NYEFAERPII CROGRCVTR VSGDQWPMY SDDAWKDLAH RCLDMKILP SEVRANFEY ITWINDWACS RRIGLATRLP WDERMIEPL TDSITYNAYT TIAHRLSEMO AGEMDDEPFD AIFLDSGTF EDLREPRYH YLIDRLSAK DLIGNHLTFH FHSALTFE GWRHGAVPF GMGLLEGNM GSSKGNVILL RDALEKHGAD VVRLFLMSSA EFWQDFDWR SEVLGTRRI EMPREGERV SGILLDGRPV SEVTPAEPES FIGHMMGQL NQILREKRA LRSFQKRAV QALILLAKD VDYTLARVSG RVDEYASVL ANVLHAWLRD MAPPIYTAE EGMHEVYCG VAAEAQWTF SDALESDVQ VAEEMVQNTV RDIQELMKIL GSTPERVHIY TSKRWKMDVL RVAABVGLD MOSTMGRVSA EGHNMREY AEPVRLIRD LGESEVTVID EYSVMDASD VIRESEGANV VIREKTYDP EGVANLPL KATVLE*	Methanobacterium thermoautotrophicum leucyl tRNA-synthetase (Mtd.RS)	RS

【国際公開パンフレット（コレクトバージョン）】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
31 October 2002 (31.10.2002)

PCT

(10) International Publication Number
WO 2002/086075 A3

(51) International Patent Classification: **C12N 9/00**, 12668-63 Carmel Country Rd., San Diego, CA 92130
C07H 21/04 (US). **ZHANG, Zhiwen** [CN/US]; 12751 Via Terceto, San
Diego, CA 92130 (US).

(21) International Application Number: PCT/US2002/012635 (74) Agents: **QUINE, Jonathan, Alan** et al.; Quine Intellec-
tual Property Law Group, P.C., P.O. Box 458, Alameda, CA
94501 (US).

(22) International Filing Date: 19 April 2002 (19.04.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 19 April 2001 (19.04.2001) US
60/285,030 6 February 2002 (06.02.2002) US
60/355,514

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.

(71) Applicant (*for all designated States except US*): **THE
SCRIPPS RESEARCH INSTITUTE** [US/US]; 10550
North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **SCHULTZ, Peter**
[US/US]; 1650 La Jolla Ranch Road, La Jolla, CA 92037
(US). **WANG, Lei** [CN/US]; 8730 Costa Verde Blvd., #
2320, San Diego, CA 92122 (US). **ANDERSON, John,
Christopher** [US/US]; 3922 8th Avenue, Apt. 11, San
Diego, CA 92103 (US). **CHIN, Jason, W.** [GB/US]; 4249
Nobel Drive #23, San Diego, CA 92122 (US). **LIU, David,
R.** [US/US]; 1 Fox Run Lane, Lexington, MA 02420 (US).
MAGLIERY, Thomas, J. [US/US]; 333 Mill Rd., North
Haven, CT 06473 (US). **MEGGERS, Eric, L.** [DE/US];
1534 Pine Street, Apt. 2F, Philadelphia, PA 19102 (US).
MEHL, Ryan, Aaron [US/US]; 3576 31st Street, San
Diego, CA 92104 (US). **PASTRNAK, Miro** [SK/US];
11686 Carmel Creek Road, Apt. 107, San Diego, CA
92130 (US). **SANTORO, Stephen, William** [US/US];

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW).
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM).
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR). OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:

— with international search report
— before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments

(88) Date of publication of the international search report:
24 June 2004

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGANAL TRNA-AMINOACYLTRNA SYN-
THETASE PAIRS

(57) Abstract: This invention provides compositions and methods for generating components of protein biosynthetic machinery
including orthogonal tRNAs, orthogonal aminoacyl-tRNA synthetases, and orthogonal pairs of tRNAs/synthetases. Methods for
identifying orthogonal pairs are also provided. These components can be used to incorporate unnatural amino acids into proteins in
vivo.



WO 2002/086075 A3

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US02/12635
A. CLASSIFICATION OF SUBJECT MATTER IPC(7) : C12N 9/00; C07H 21/04 US CL : 435/183; 536/24.1 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 435/183; 536/24.1 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) C.Aplus		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database NCBI, GenBank Accession Number B64348, BULT et al. "Complete genome sequence of the methanogenic archaeon, <i>Methanococcus jannaschii</i> ," Gene Sequence, 03 June 1996.	7
X,P --- Y,P	WANG et al. A general approach for the generation of orthogonal tRNAs. Chemistry & Biology, September 2001, Vol. 8, pages 883-890.	1,3-9,17-19,24,25,28-30,34-36,74,98
A	LIU et al. Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo. Proc. Natl. Acad. Sci. September 1997, Vol. 94, pages 10092-10097.	1,3,6,17,18,26,31,36
A	LIU et al. Progress toward the evolution of an organism with an expanded genetic code. Proc. Natl. Acad. Sci. April 1999, Vol. 96, pages 4780-4785.	1,3,6,17,18,26,31,36
A	KOWAL et al. Twenty-first aminoacyl-tRNA synthetase-suppressor tRNA pairs for possible use in site-specific incorporation of amino acid analogues into proteins in eukaryotes and in eubacteria. Proc. Natl. Acad. Sci. 27 February 2001, Vol. 98, No. 5, pages 2268-2273.	1,3,6,17,18,26,31,36
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents:		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"A" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 27 January 2004 (27.01.2004)	Date of mailing of the international search report 05 MAY 2004	
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703) 305-3230	Authorized officer Kathleen M Kerr <i>J. Roberts for</i> Telephone No. (571) 272-1600	

INTERNATIONAL SEARCH REPORT		PCT/US02/12635
Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)		
This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1.	☐	Claim Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
2.	☐	Claim Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3.	☐	Claim Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)		
This International Searching Authority found multiple inventions in this international application, as follows: Please See Continuation Sheet		
1.	☐	As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.	☐	As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.	☐	As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.	☒	No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-9, 17-37, 74, 98 (SEQ ID NOs: 4 and 35 only)
Remark on Protest		
	☐	The additional search fees were accompanied by the applicant's protest.
	☐	No protest accompanied the payment of additional search fees.

PCT/US02/12635

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claim(s) 1-9, 17-37, 74, and 98, drawn to orthogonal aminoacyl-tRNA synthetases (O-RSs) and compositions thereof.

Group II, claim(s) 10-37, 96, and 98, drawn to orthogonal tRNAs (O-tRNA) and compositions thereof.

Group III, claim(s) 38-73, 97, 99-101, and 110-115, drawn to methods of producing O-RSs using two successive screening/selecting procedures.

Group IV, claim(s) 75-95, 97, 99-101, and 110-115, drawn to methods of producing O-tRNAs using two successive screening/selecting procedures.

Group V, claim(s) 102-115, drawn to methods for identifying an O-tRNA-RS synthetase pair for use in an in vivo translation system of a second organism.

The inventions listed as Groups I-V do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons. The technical feature of Group I is the combination of the functional and structural requirements of the claimed polypeptides, orthogonal tRNA synthetases. This technical feature is not a special technical feature because the structural requirements broadly claimed in Claim 7 without any functional limitations do not contribute novelty as a whole over the prior art. Particularly, SEQ ID NOs: 35-64 are mutants of *M. jannaschii* tyrosyl-tRNA synthetase (GenBank Accession Number Q57834) and the DNA encoding Q57834 will hybridize under highly stringent conditions to DNA encoding SEQ ID NOs:35-64, as required to meet the limitations in Claim 7, item c.

The products of Group II do not share or have a corresponding technical feature to that of Groups I or IV because the tRNA molecules of Group II, which are made in the methods of Group IV, have different structures and functions to that of tRNA synthetases of Group I.

The methods of Group III do not share or have a corresponding special technical feature to that of Group I because, as noted above, the technical feature holding the products of Group I together is not a special technical feature that, considered as a whole, teaches above the art. When a special technical feature is lacking in Group I, different categories of inventions (i.e., product, process of making that product, process of using that product) are not grouped together since technical features do not define unity of invention. Thus, methods of making the products, Group III, are not grouped together with Group I.

The methods of Group V do not share or have a corresponding special technical feature to that of Group I because the products of Group I are neither made nor used in the methods since the phrase, "preferentially aminoacylates", which is a requirement in the technical feature of Group I, is not a requirement in the methods of Group V.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack unity of invention because they are not so linked as to form a single general inventive concept under PCT Rule 13.1.

In order for more than one species to be examined, the appropriate additional examination fees must be paid. The species are as follows:

- In Group I, species are SEQ ID NOs: 35-66 as encoded by SEQ ID NOs: 4-34.
- In Group II, species are SEQ ID NOs: 1-3.

The claims are deemed to correspond to the Group I species listed above in the following manner:

Claims 2 and 7-9 are drawn to each of the thirty-two species listed above for Group I.

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

PCT/US02/12635

The following claim(s) are generic: 1, 3-6, 17-37, 74, and 98.

The claims are deemed to correspond to the Group II species listed above in the following manner:

Claims 10-13 are drawn to each of the three species listed above for Group II.

The following claim(s) are generic: 14-37, 74, and 98.

The species listed above do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons.

The technical feature of the species of Group I is the structural requirement of the claimed polypeptides, orthogonal tRNA synthetases. This technical feature is not a special technical feature because, as broadly claimed in Claim 7, the structures as a whole do not contribute novelty over the prior art. SEQ ID NOs: 35-64 are mutants of *M. jannaschii* tyrosyl-tRNA synthetase (GenBank Accession Number Q57834) and the DNA encoding Q57834 will hybridize under highly stringent conditions to DNA encoding SEQ ID NOs: 35-64, as required to meet the limitations in Claim 7, item c.

The technical feature of the species of Group II is the structural requirement of the claimed orthogonal tRNAs. This technical feature is not a special technical feature because, as broadly claimed in Claim 10, the structures as a whole do not contribute novelty over the prior art. SEQ ID NOs: 1-3 are mutants of *M. jannaschii* tRNA (GenBank Accession Number J11U_B) and the RNA of J11U_B will hybridize under highly stringent conditions to RNA that is SEQ ID NOs: 1-3, as required to meet the limitations in Claim 10, item b.

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN, TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE, GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,R O,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

(72)発明者 レイ・ワン

アメリカ合衆国 カリフォルニア州 9 2 1 2 2 サン ディエゴ ナンバー 2 3 2 0 コスタ
ヴァード ブールヴァード 8 7 3 0

(72)発明者 ジョン・クリストファー・アンダーソン

アメリカ合衆国 カリフォルニア州 9 2 1 0 3 サン ディエゴ アpartment 1 1 エイ
ス アヴェニュー 3 9 2 2

(72)発明者 ジェイソン・ダブリュー・・チン

アメリカ合衆国 カリフォルニア州 9 2 1 2 2 サン ディエゴ ノーベル ドライヴ ナンバ
ー 2 3 4 2 4 9

(72)発明者 デイヴッド・アール・・リウ

アメリカ合衆国 マサチューセッツ州 0 2 4 2 0 レキシントン フォックス ラン レーン
1

(72)発明者 トーマス・ジェー・・マグライリー

アメリカ合衆国 コネティカット州 0 6 4 7 3 ノース ヘヴン ミル ロード 3 3 3

(72)発明者 エリック・エル・・メッゲルス

アメリカ合衆国 ペンシルベニア州 1 9 1 0 2 フィラデルフィア アpartment 2 エフ
パイン ストリート 1 5 3 4

(72)発明者 ライアン・アaron・メール

アメリカ合衆国 カリフォルニア州 9 2 1 0 4 サン ディエゴ サーティファースト ストリ
ート 3 5 7 6

(72)発明者 ミロ・パストラナック

アメリカ合衆国 カリフォルニア州 9 2 1 3 0 サン ディエゴ アpartment 1 0 7 カ
ーメル クリーク ロード 1 1 6 8 6

(72)発明者 ステイーヴン・ウィリアム・サントロ

アメリカ合衆国 カリフォルニア州 9 2 1 3 0 サン ディエゴ カーメル カントリー ロー
ド 1 2 6 6 8-6 3

(72)発明者 ツィーウェン・チャン

アメリカ合衆国 カリフォルニア州 9 2 1 2 2 サン ディエゴ ナンバー 3 3 チャーマント
ドライヴ 7 0 4 5

Fターム(参考) 4B024 AA20 BA07 CA02 CA11 DA06 FA10 HA01

4B050 CC04 DD02 KK12 KK20

4B063 QA06 QQ05 QQ21 QQ50 QQ80 QR01 QR35 QR49 QR74