

(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2004-531225

(P2004-531225A)

(43) 公表日 平成16年10月14日(2004.10.14)

|                            |                   |                  |
|----------------------------|-------------------|------------------|
| (51) Int. Cl. <sup>7</sup> | F I               | テーマコード (参考)      |
| C 1 2 N 15/09              | C 1 2 N 15/00     | 4 B O 2 4        |
| A 6 1 K 45/00              | A 6 1 K 45/00     | 4 B O 6 4        |
| A 6 1 P 35/00              | A 6 1 P 35/00     | 4 C O 8 4        |
| A 6 1 P 43/00              | A 6 1 P 43/00     | 1 O 1            |
| C 1 2 P 21/02              | C 1 2 P 21/02     | Z C C C          |
|                            | 審査請求 未請求 予備審査請求 有 | (全 130 頁) 最終頁に続く |

(21) 出願番号 特願2002-559580 (P2002-559580)  
 (86) (22) 出願日 平成14年1月25日 (2002.1.25)  
 (85) 翻訳文提出日 平成15年7月24日 (2003.7.24)  
 (86) 国際出願番号 PCT/US2002/002344  
 (87) 国際公開番号 W02002/059293  
 (87) 国際公開日 平成14年8月1日 (2002.8.1)  
 (31) 優先権主張番号 60/264, 147  
 (32) 優先日 平成13年1月25日 (2001.1.25)  
 (33) 優先権主張国 米国 (US)

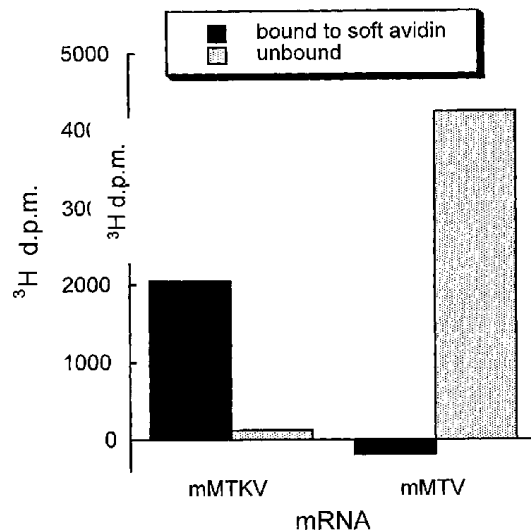
(71) 出願人 503265577  
 アントニー シー. フォースター  
 アメリカ合衆国 02421 マサチュー  
 セッツ、レクシントン、マレット ロード  
 173  
 (71) 出願人 503265599  
 スティーブン シー. ブラックロウ  
 アメリカ合衆国 02138 マサチュー  
 セッツ、ケンブリッジ、アッシュ ストリ  
 ート 16  
 (74) 代理人 100067817  
 弁理士 倉内 基弘  
 (74) 代理人 100085774  
 弁理士 風間 弘志

最終頁に続く

(54) 【発明の名称】 ペプチド、蛋白質及び疑似ペプチド合成のための方法及び組成物

(57) 【要約】

本発明は、翻訳因子 E F - P、W、W2 又はレスキューの添加を必要としない、単純化され、高度に精製された前進型の翻訳系である。新規の翻訳プロセスが現在のところ別経路で合成されている全てのペプチド、蛋白質及び疑似ペプチドに新規で改良された経路を提供する。



## 【特許請求の範囲】

## 【請求項 1】

細胞を含まない再構成翻訳系において、翻訳因子及び tRNA 種であって外因的に添加された mRNA をそれぞれのコドンでの高選択的な取り込みでもって翻訳してペプチド産物又は該系が非天然型アミノ酸若しくはアミノ酸アナログを結合した tRNA の 1 種以上を含むときには疑似ペプチド産物を形成できるものを含む、細胞を含まない再構成翻訳系。

## 【請求項 2】

調製物が翻訳因子 EF-P、W、W2 又はレスキューを実質的に含まない、ペプチド産物生成用の請求項 1 に記載の翻訳系。

## 【請求項 3】

疑似ペプチド産物の主鎖に取り込まれる非天然型アミノ酸又はアミノ酸アナログを結合した tRNA 種の 1 種以上を含む、疑似ペプチド産物生成用の請求項 1 に記載の翻訳系。

## 【請求項 4】

調製物が翻訳因子 EF-P、W、W2 又はレスキューを実質的に含まない、疑似ペプチド産物生成用の請求項 2 に記載の翻訳系。

## 【請求項 5】

アミノ酸アナログが -シアノアラニン、カナバニン、ジエンコル酸、ノルロイシン、3-ホスホセリン、ホモセリン、ジヒドロキシフェニルアラニン、5-ヒドロキシトリプトファン、1-メチルヒスチジン、3-メチルヒスチジン、アシルグリシン（又はそのアルキル誘導体）、O-メチルセリン、ピオチニルリシン、ピオチニルシステイン（又はその他のピオチン標識アミノ酸）、シクロヘキシルアラニン、ホモグルタメート、D-アラニン（又はその他の D-アミノ酸）、N-メチルグリシン（又はその他の N-メチルアミノ酸）、-N-メチルリシン及び 21 種の天然アミノ酸又は非天然アミノ酸の放射性同位元素誘導体よりなる群から選択される請求項 2 に記載の翻訳系。

## 【請求項 6】

1 種以上の外因的に添加された mRNA 種をさらに含む請求項 1 に記載の翻訳系。

## 【請求項 7】

翻訳因子及び tRNA 種であって外因的に添加された mRNA を翻訳して疑似ペプチド産物を形成できるものを含む細胞を含まない翻訳系において、その調製物が

(a) 1 種以上の活性な野生型アミノアシル tRNA と該野生型アミノアシル tRNA を合成する能力とを欠損しており、

(b) 該活性のない tRNA 種の代わりに非天然アミノ酸種又はアミノ酸アナログを結合した少なくとも 1 種の外因性アミノアシル tRNA を含む、細胞を含まない翻訳系。

## 【請求項 8】

複数の異なる mRNA 種を含む請求項 7 に記載の翻訳系。

## 【請求項 9】

(a) 外因的に添加された RNA を翻訳してペプチド又は疑似ペプチドを形成できる細胞を含まない再構成翻訳系を形成させるために混合され得る請求項 1 に記載の翻訳系の成分及び (b) それに関する説明書からなるキット。

## 【請求項 10】

(a) 外因的に添加された RNA を翻訳して疑似ペプチドを形成できる細胞を含まない翻訳系を形成させるために混合され得る請求項 7 に記載の翻訳系の成分及び (b) それに関する説明書からなるキット。

## 【請求項 11】

次の段階：

(a) 請求項 1 に記載の細胞を含まない翻訳系を準備し、

(b) 該翻訳系を 1 種以上の外因性 mRNA 種と接触させ、

(c) 該翻訳系のペプチド若しくは疑似ペプチド産物を単離及び/又は同定することを含むペプチド又は疑似ペプチドの生成方法。

10

20

30

40

50

## 【請求項 12】

翻訳系の tRNA 種の 1 種以上に疑似ペプチド産物の主鎖に取り込まれる非天然型アミノ酸又はアミノ酸アナログを結合させる請求項 11 に記載の方法。

## 【請求項 13】

前記方法を少なくとも 100 種の異なる RNA 種のライブラリーについて実施し、しかもペプチド又は疑似ペプチド産物を翻訳系から触媒活性又は結合活性に基づいて同定又は単離する、請求項 11 に記載の方法。

## 【請求項 14】

mRNA を前記翻訳系で *in vitro* 転写によって生成させる請求項 11 に記載の方法。

## 【請求項 15】

ペプチド又は疑似ペプチド産物を、該産物をコードする外因性 mRNA の共有結合付加物として形成させる、請求項 11 に記載の方法。

## 【請求項 16】

翻訳系を異なる外因性 mRNA 種のライブラリーと接触させて少なくとも  $10^3$  種の異なる配列のペプチド又は疑似ペプチド産物の多様な集団を生じさせる請求項 11 に記載の方法。

## 【請求項 17】

少なくとも  $10^8$  種の異なる配列を生じさせる請求項 16 に記載の方法。

## 【請求項 18】

請求項 11 に記載の方法によって同定若しくは合成されたペプチド又は疑似ペプチド。

## 【請求項 19】

次の段階：

- (a) 請求項 7 に記載の細胞を含まない翻訳系を準備し、
- (b) 該翻訳系を 1 種以上の外因性 mRNA 種と接触させ、
- (c) 該翻訳系の疑似ペプチド産物を単離及び / 又は同定することを含む疑似ペプチドの生成方法。

## 【請求項 20】

次の段階：

- (a) 請求項 11 に記載の方法によって所望の生物学的活性を有する翻訳系のペプチド又は疑似ペプチド産物を同定し、
- (b) 段階 (a) で同定されたペプチド又は疑似ペプチド産物の生物学的活性をさらに試験管内又は生体内アッセイによって確認し、
- (c) 段階 (b) で確認された産物又はさらにそのアナログの動物における有効性及び毒性についての治療プロファイリングを行い、
- (d) 許容できる治療プロファイルを有する段階 (c) で同定された 1 種以上の産物を含む医薬品を処方することを含む薬剤発見ビジネスを行う方法。

## 【請求項 21】

販売用の前記医薬品を流通させるための流通システムを確立する追加の段階を含み、随意として該医薬品を販売するための販売部門を設立することを含む、請求項 20 に記載の方法。

## 【請求項 22】

次の段階：

- (a) 請求項 11 に記載の方法によって所望の生物学的活性を有する翻訳系のペプチド又は疑似ペプチド産物を同定し、
- (b) 段階 (a) で同定されたペプチド又は疑似ペプチド産物の生物学的活性をさらに試験管内又は生体内アッセイによって確認し、
- (c) 随意として段階 (b) で確認された産物又はさらにそのアナログの動物における有効性及び毒性についての治療プロファイリングを行い、
- (d) 確認された産物のさらなる薬剤開発についての権利を第三者にライセンス供与する

10

20

30

40

50

こと

を含む薬剤発見ビジネスを行う方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、翻訳因子EF-P、W、W2又はレスキュー(rescue)の添加を必要としない単純化され、高度に精製された前進的な翻訳系に関する。

【背景技術】

【0002】

リガンドの認識及び結合は、免疫認識、細胞シグナリング、細胞連絡、転写、翻訳、細胞内シグナリング及び触媒反応、即ち酵素反応のようなほとんど全ての生物学的プロセスを調節する。ホルモン、増殖因子及び神経伝達物質のようなアゴニストとして作用し又はリガンド活性を作動若しくは拮抗させることができる天然又は非天然のリガンド分子(このものは、B細胞(抗体仲介)免疫又はT細胞(細胞仲介)免疫を誘導し、化学反応を触媒でき、或いは遺伝子発現を転写又は翻訳レベルで調節できる)を同定及び合成するための技術には、長年にわたって関心が寄せられている。このようなリガンドの大部分は、蛋白質、ペプチド及び疑似ペプチド(peptidomimetic)である。

10

【0003】

リガンド及び薬剤発見への伝統的なアプローチは、運と激務の混ぜ合わせに大きく依存している。動物及び植物組織からの天然産物又は発酵培養基の産物をスクリーニングすること、或いは保管された合成分子を無作為スクリーニングすることが新規のリード化合物を同定するために最も生産的な手段であった。

20

【0004】

しかしながら、新薬の探索における最近の動向は、新薬発見に関して可能性のある新規のリード化合物源としてのコンビナトリアル・ライブラリーを調製することに集中している。「コンビナトリアル・ケミストリー」というこの新しい分野の核心は、非生合成又は生合成のいずれかで調製でき且つ生物活性ごとに様々な形式でスクリーニングできる異なった分子を集めることである。非生合成技術、例えば、コード化、空間的アドレッシング及び/又はデコンボリューションの使用によって、ペプチド、疑似ペプチド及び非ペプチドベースの分子のコンビナトリアル・ライブラリーがバッチプロセスで合成でき、そして重要なことに、このライブラリーの個々の要素の分子同一性は、薬剤スクリーニング形式で確認できる(例えば、「Lam他(1993)Gene 137, 13-16」、Doolley他(1994)Science 266, 2019-2022)。非生合成ライブラリーは、生物学的単量体(例えば天然のアミノ酸及びヌクレオチド)及びそれらの誘導体に制限されないという利点を有するが、このライブラリーは、数週間のうちにスクリーニングできる分子の数が制限されるという不利点を有する(通常せいぜい $10^5 \sim 10^8$ 個であり、これは、対象の標的に対して高い親和性のリガンドを同定するには小数すぎる分子である)(「Roberts(1999)Curr. Op. Chem. Biol. 3, 268-273」、「Wilson他(2001)PNAS 98, 3750-3755」)。しかしながら、生合成ライブラリーは、多くの場合この制限をほとんど受けない。なぜならば、 $10^{15}$ 個の異なるペプチド、RNA又はDNA分子を数週間のうちにスクリーニングできるというこのようなライブラリーの例が存在するからである(Roberts, 上記)。これは、個々の生合成ライブラリー要素を、多くの場合時として定向進化と呼ばれるダーウィンの進化論に類似したプロセスである関連した突然変異誘発段階(例えば、親和性成熟、突然変異誘発PCR又はDNAシャフリング(Roberts, 上記))でもって反復選択し且つ増幅させることによって達成される。

30

40

【0005】

部分的に又は完全に無作為化されたプールからの蛋白質の単離を可能にした従来の方法の多くは、そのようなin vivo段階によるものであった。この種の方法には、モノクローナル抗体技術(「Milstein, Sci. Amer. 243:66(1980)」及

50

び「Schultz他, J. Chem. Engng. News 68:26(1990)」、ファージディスプレイ(「Smith, Science 228:1315(1985)」、「Parmley及びSmith, Gene 73:305(1988)」並びに「McCafferty他, Nature 348:552(1990)」、ペプチド・lacrプレッサー融合(Cull他, PNAS 89:1865(1992))及び古典的な遺伝学的選択が含まれる。これらの方法のそれぞれは、蛋白質と核酸の間のトポロジカルな結合に依拠する。なぜならば、核酸しか複製できないからである。しかして、蛋白質の情報保持され且つ読み込み可能な核酸の形で回収できる。

#### 【0006】

他の蛋白質選択技術は、in vivo 段階なしに実施される。多くの場合リボソームディスプレイと呼ばれる失速翻訳法は、リボソーム及びmRNAとさらに複合体を形成する新生蛋白質鎖のいくらかの特性ごとに選択を行う技術である(「川崎, 米国特許第5658754号」、「Tuerk及びGold, Science 249:505(1990)」、「Irvine他, J. Mol. Biol. 222:379(1991)」、「Korman他, PNAS 79:1844-1848(1982)」、「Mattheakis他, PNAS 91:9022-9026(1994)」、「Mattheakis他, Meth. Enzymol. 267:195(1996)」並びに「Hanes及びPluckthun, PNAS 94:4937(1997)」。mRNA・蛋白質融合法又はmRNAディスプレイ法(「根本他, FEBS Lett. 414, 405-408(1997)」、「柳川他, 米国特許第6228994号」、「Szostak他, 米国特許第6281344号、同6261804号、同6258558号、同6214553号及び同6207446号」並びに「Roberts及びSzostak, PNAS 94, 12297-12302(1997)」)は、mRNAをDNA/プロマイシンリンカーによってその蛋白質産物に直接共有結合させる。終止コドンの存在によっては損なわれない「裸の」mRNA・ペプチド融合体を合成させるための方法は、ペプチドをミセルの状態ですれらのペプチドをリボソームから分離させ、次いでそれらの特異的mRNAに再結合できるような方法でもって合成させることである(例えば、ストレプトアビジン配列を含む蛋白質は、ピオチニル化mRNAに結合するであろう。土井及び柳川(1999) FEBS Lett. 475, 227-230)。

#### 【0007】

しかしながら、この先行技術の「天然」(L-)ペプチドライブラリー技術は、多くの不都合がある。第1に、ほぼ完全にキラルな単量体(アミノ酸)からなるこのライブラリーは、このキラル単量体の鏡像異性体を欠失している。例えば、L-ペプチドライブラリーについて、この20種の天然型アミノ酸は、幅広い立体性、電子性及び官能性の基を与えるが、C-炭素のキラリティーは、先行技術のディスプレイ技術によって達成できる三次元形状の空間を事実上制限する。また、L-ペプチドライブラリーは、標的物との非共有結合又は共有結合複合体を形成させるのに有用であり得る多数の一般的な有機化学官能基も欠失しており(例えば、アルケン、アルキル尿素、ハロゲン化アルキル及びケトン)、しかも「非天然」アミノ酸(予め合成されたものか又は理論上のものかのいずれか)で達成できる非常に大きい付加形状の多様性も欠如している。さらに、治療剤として、天然L-アミノ酸のペプチドは、通常それらの非天然鏡像異性体(D-ペプチド)又はアナログよりも好ましいわけではない。なぜならば、L-ペプチドは、in vivo プロセッシングのため乏しい薬物動態学的プロファイルによって使用が制限され得るからである。例えば、L-ペプチドは、動物への投与後にプロテアーゼによって迅速に分解され得、しかしてより高い有効投与量を必要とする。さらに、薬用ペプチドは、患者の強力な免疫原性応答を誘導でき、さらにその迅速なクリアランスに寄与し、また有害であるかもしれない炎症反応も生じさせる。この薬用ペプチドの分解を回避するための一方法は、レトロインベルソアナログ(Sisto他の米国特許第4522752号を参照)、レトロエナンチオアナログ(Goissis他のJ. Med. Chem. 19:1287-90(1976)を参照)のような非加水分解性のペプチドアナログ、トランス-オレフィン誘導體(Shu

e 他の Tetrahedron Letters 28:3225 (1987) を参照) 及びホスホネート誘導体 (Loots 他「ケミストリーアンドバイオロジー」(エスコムサイエンス出版, ライデン, 1988, p118) の「ペプチド」を参照) を生成させることであった。しかしながら、多くの場合、ペプチドの主鎖は、蛋白質分解に耐性のある疑似ペプチドにするために変質される。このようにすると、得られた疑似ペプチドは、天然ペプチド主鎖と標的受容体との間のある種の結合性接触の欠如並びに二面角などの変化による該ペプチドに対する立体空間の変化によって生物活性が低下する欠点を有し得る。別の問題は、ほとんど全ての L - ペプチドがその親水性のために生体膜を横断しないことである。対照的に、D - ペプチド及び N - メチルアミノ酸のようなその他の非天然アミノ酸を含むペプチド (疑似ペプチド) はプロテアーゼ耐性を増大させ、そして疑似ペプチド薬のシクロスポリン A は、膜を横断でき且つ経口的に利用できる。なぜならば、部分的にはあるが、このものは、天然のペプチド結合よりも疎水性であるいくつかの N - メチルペプチド結合を含有するからである (「Zawadzke 及び Berg (1992) J. Am. Chem. Soc. 114:4002」、Walsh 他 (1992) J. Biol. Chem. 267, 13115 - 13118)。残念ながら、D - ペプチドライブラリーのような化学的に合成された (非生合成の) 疑似ペプチドライブラリー (Lam 他 (上記)、Dooly (上記)) は上で論じたライブラリーサイズの制限を受け、鏡像ファージ又はリボソームディスプレイ (「Schumacher 他 (1996) Science 271, 1854 - 1857」、Eckert 他 (1999) Cell 99, 103 - 115」及び「Forster 他 PCT 公開 WO 97/35194」) のような疑似ペプチドライブラリーのサイズ制限を克服するための方法論的戦略は、鏡像異性体の標的物を化学的に合成するという厄介な必要条件によって制限される。

10

20

**【0008】**

蛋白質、ペプチド及び疑似ペプチドは、一般に、3つの異なる方法で合成され、それぞれ固有の制限を有している。

**【0009】**

1. 合成ペプチド化学は、長さが約30残基までの非常に多様な疑似ペプチドを高収率且つ高純度で合成するために日常的に使用され得る (Eckert 他, 上記)。しかしながら、この方法は、非効率的なカップリング段階、精製の問題及び折りたたみの困難性のためにより長い産物については非効率的又は非実用的である。また、所望の生成物の反応性側鎖の全てについて相溶性のある保護基が必要であるため、合成が制限される。さらに合成疑似ペプチドは、反復選択、増幅及び変異 (進化) のために遺伝学的にコードされ得ず、合成疑似ペプチドライブラリーの複雑性を最適な薬剤発見には少なすぎる約  $10^8$  種の分子に制限させる。

30

**【0010】**

2. 生細胞を使用する *in vitro* 翻訳が、遺伝学的にコードされた天然又は組換え DNA 配列からの短い又は長い蛋白質を効率的に合成し且つ翻訳後に修飾するために広く使用されている。

しかしながら、合成は、その遺伝子産物が有毒である場合には役に立たず、特にその蛋白質が封入体中で発現する場合には精製が困難であり得、しかも復元の問題もあり得る。最も重要なことには、この方法は、複数の非天然アミノ酸を選択的に取り込むことができない、或いは翻訳後修飾プロセス (例えば、プロテアーゼ触媒プロセッシング又は分解) を制御することができない欠点がある。

40

**【0011】**

3. 粗製細胞抽出物による *in vitro* 翻訳は、一般に、毒性の問題を克服し (しかし、翻訳後修飾を制御しない)、より容易な精製及び折りたたみを生じさせることができ、しかも人工のサブストラート RNA を使用して蛋白質当たり1個の非天然アミノ酸の選択的な取り込みを可能にさせる (Noren 他 (1989) Science 244, 182 - 188)。

しかしながら、この方法による非天然アミノ酸の取り込みは、*in vivo* 系よりも収率が非

50

常に低い難点がある。なぜならば、この方法は、終結因子と競合する本質的に非効率なサブレッサー tRNA に依存するからである。100 個以上の異なる非天然アミノ酸が個々に取り込まれたが (例えば、Mendel 他 (1995) *Annu. Rev. Biophys. Biomol. Struct.* 24, 435 - 462)、この戦略は、3 種の終止 (ナンセンス) コドンの一つ (UAG コドン) のみでの蛋白質当たり 1 個だけの非天然アミノ酸の選択的な取り込みに制限された。これは、20 種の異なるアミノアシル tRNA シンテターゼの tRNA 結合活性及びブルーフリーディング活性によって触媒される天然アミノ酸によるアミノ酸 (センス) コドンでの競合のため及びリボソームによるリードスルーのため第 2 の終止コドン (UGA) を使用しようとする試みが失敗したためである (Cloward 他 (1996) *Chem. and Biol.* 3, 1033 - 1038)。

10

## 【0012】

また、非天然アミノ酸をセンスコドンに一般的な態様で選択的に取り込もうとする多くの試みも失敗した。例えば、非天然型アミノ酸を取り込むために最も一般的に使用される方法では、天然アミノ酸の高比活性放射性同位元素誘導体を *in vitro* 翻訳によって取り込んで放射性標識された蛋白質を合成させる場合に、この放射性アミノ酸の比活性が、通常、粗製翻訳系に存在するアミノ酸の非標識型による取り込みのための競合によってこの非標識型を添加された非標識アミノ酸プールから与えないにもかかわらず減少することは周知である。プロメガ社によって、類似するアナログの希釈の結果が別のレポーター群であるピオチン標識リシンの取り込み用の市販キットを使用して得られている (Trasce and (商標) 非放射性翻訳系の付属文献)。さらに、粗製の翻訳抽出液を濾過して天然アミノ酸を除去し、次いでリシンを除く天然アミノ酸の全てを追加し且つアミノ酸アナログが付加されたリシン tRNA を追加すると、リシンアナログ対リシンがわずかに 1 : 3 ~ 1 : 4 の比で取り込まれた (Crowley 他 (1993) *Cell* 73, 1101 - 1115)。アミノ酸アナログの取り込みの低い選択性は、ある種の用途には十分であるが (Rothschild 他, 米国特許第 5643722 号)、遺伝学的にコードされた特異的な疑似ペプチド配列の増幅及び特徴付けを必要とするような多くの用途には明らかに適さない。2 種の異なるフレームシフトサブレッサー tRNA を使用して 2 種の異なる非天然アミノ酸を取り込めることが証明され (芳坂他 (1999) *JACS* 121, 12194 - 12195)、多くの同一の非天然アミノ酸が Phe アミノアシル tRNA シンテターゼに対して特異的な阻害剤を使用して取り込まれた (Baldini 他 (1988) *Biochem.* 27, 7951 - 7959)。しかしながら、これらの方法の両方は、多くの異なる非天然アミノ酸を単一の疑似ペプチドに取り込むのに必要な態様には一般化できない。粗製及び *in vivo* 翻訳に固有のこれらの制限を克服するために、遺伝暗号を直交 tRNA 及び直交非天然核酸塩基対に基づいて拡張させるための綿密な戦略が提案されたが、単一の試験管内改変終止コドン (Bain 他 (1992) *Nature* 356, 537 - 539) を超える開発は、技術的に非常に困難であることが分かった (Service (2000) *Science* 289, 232 - 235)。

20

30

## 【0013】

本発明者は、この問題が純粋な *in vitro* 翻訳系を使用して解決できるかもしれないと考えた。非天然アミノ酸と天然アミノ酸又は終結因子との間の競合は、ある種のアミノ酸、tRNA、アミノアシル tRNA シンテターゼ及び/又は終結因子のようなある種の成分を除外することによって回避できたかもしれない。不運にも、mRNA 依存性ポリペプチド合成のための最低限の必要条件是、含まれる高分子が多いために定義するのが困難であった。精製された成分からの翻訳の再構成が *E. coli* について達成されたが、必要とされる翻訳因子の数は、いまだ議論の余地がある。

40

## 【0014】

ワイスバッハ研究所によって構築された最初の精製翻訳系は、4 種の *E. coli* mRNA を、高度に塩洗浄されたリボソーム、開始因子 (部分的な IF-1 依存性)、伸長因子 Tu (EF-TuH) 及びアミノアシル tRNA シンテターゼの群に強く依存して、met-tRNA<sub>i</sub><sup>fMet</sup> ホルミルトランスフェラーゼ及び伸長因子 G (EF-G) に部分的

50

に依存して、伸長因子Ts (EF-Ts) 又は終結因子に全く依存しないで効率的に翻訳した (Kung 他 (1978) Arch. Biochem. and Biophys. 187, 457-463)。非常に多くの精製成分を維持し且つ微量の汚染物を除去する際の困難性のために、追加の一般的な翻訳因子についてのサーチは、fMet-tRNA<sub>i</sub><sup>f</sup> Me<sup>t</sup> 及び1又は2種の伸長体アミノアシルtRNAによるジペプチド又はトリペプチドの合成へとこの系を単純化させることによって促進され、それによってアミノアシルtRNAシンターゼのための必要条件が回避された (Weissbach 他 (1984) Biotechniques 2, 16-22)。

#### 【0015】

Ganozaが率いる第2グループは、後者の単純化された方法を試験管内で付加された全tRNA及び終結因子を使用してより長いペプチドに拡大させたときに、バクテリオファージMS2及びf1の翻訳がEF-P、W及びレスキュー (rescue) と呼ばれる3種の追加の因子に依存することを発見した (「Green 他 (1985) Biochem. Biophys. Res. Com. 126, 792-798」、 「Ganoza 他 (1985) PNAS 82, 1648-1652」)。これらの因子を存在させないと、連続移動性が非効率的になった。例えば、ヘキサペプチド合成反応ではジ、トリ、テトラ及びペントペプチドの休止産物又は早期終結産物が支配的であった。dead/W2 (Wよりも数kD大きい) と呼ばれる恐らくは関連のある翻訳因子及びEF-Pもクローン化され、これらのものは最大限の成長にとって必要であり、そして真核細胞の開始因子に相通的である (「青木 他 (1991) Nucleic Acids Res. 19, 6215-6220」、 「青木 他 (1997) J. Biol. Chem. 272, 32254-32259」、 「Lu 他 (1999) Int. J. Biochem. Cell Biol. 31, 215-229」)。

#### 【0016】

Ganozaの結果とは明らかに相反して、その他2つのグループが精製成分をEF-P、W、W2又はレスキューを加えることなく使用してアミノアシルtRNA基質から短いペプチドを合成することを報告したが、これら2つのグループは、それらの系の連続移動性又はそれらのリボソームの純度を直接記録していなかった (「Stade 他 (1995) Nucleic Acids Res. 23, 2371-2380」、 「Pavlov 他 (1997) EMBO J. 16, 4134-4141」)。この矛盾が本当ならば、これは単に解釈に関して推測しているにすぎない。例えば、EF-P、W及びレスキューは、リボソーム調製物から精製できるため (Ganoza 他 (1996) Biochemie 78, 51-61)、Ganozaのグループが使用したのとはかなり異なる手順によって調製された後者2つのグループが使用したリボソームは、EF-P、W、W2及び/又はレスキューで汚染されていた可能性がある。これは問題が多い。なぜならば、EF-P、W、W2及び/又はレスキューによる汚染は、アミノアシルtRNAシンターゼ及び終結因子のような、望ましくない反応の原因となり得るより多数の蛋白質による汚染を暗示しかねないからである。一方、EF-P、W、W2及び/又はレスキューは、Ganozaの系では単に効率的な連続移動性のために必要とされ得るにすぎない。

#### 【0017】

可能ならば、ペプチド又は蛋白質を純粋な翻訳系によってEF-P、W (多くの場合W2と呼ばれる) 及びレスキューを添加することなく合成できることが望ましい。なぜならば、これらの蛋白質は、機能に関してあまりよく理解されておらず、結果的にそれらの活性を検査するのが困難となり、しかしてその前に活性な蛋白質を精製するのが困難になるからである。さらに、W (又はW2) の実際寸法及びWとW2が同一蛋白質の誘導体に相当するのかどうかについての議論があり、また、レスキューについての遺伝子をさらにクローン化しなければならない。

#### 【発明の開示】

#### 【課題を解決するための手段】

#### 【0018】

10

20

30

40

50

本発明は、翻訳因子 E F - P、W、W<sub>2</sub> 又はレスキューの添加を必要としない単純化され、高度に精製された前進的な翻訳系である。新規な翻訳プロセスは、これまでのところ別経路によって合成されていた全てのペプチド及び蛋白質に新規でもすれば改善された経路を提供する。このプロセスは、蛋白質、ペプチド及び疑似ペプチド合成のために上記された方法 1、2 及び 3 に固有の制限を克服する。

【0019】

ある好ましい具体例では、この精製された系は、インスリン、成長ホルモン又はエリスロポエチンのようなペプチド又は蛋白質リガンド又は触媒の合成のために使用できる。

【0020】

別の好ましい具体例では、この精製系は、粗製の細胞抽出液に依存する既存のリボソーム及び mRNA ディスプレイ系とは異なり、「純粋なリボソームディスプレイ」及び「純粋な mRNA ディスプレイ」選択実験のために使用できる。ペプチドの翻訳後修飾の欠如が予期されること、しばしば蛋白質分解の問題を生じさせるプロテアーゼがないこと及び選択段階で汚染物による競合がないことのような、純粋系でペプチド及び蛋白質ディスプレイを実施することに関連するいくつかの利点がある。追加の利点として、以下のことが挙げられる。

(i) リボヌクレアーゼが存在しないこと(本発明の純粋な *E. coli* 系で測定された放射性 mRNA の長期間安定性によって立証される(結果は示さない))は、様々な粗製系、特に *E. Coli* (Robertson, 上記) で観察される mRNA 分解に関わる問題を回避させる。mRNA が翻訳及び選択され得る前に分解されないことは明らかに重要である。

(ii) 純粋な再構成系は翻訳終結因子によって汚染されないことが予期される。実際に、本発明の系は、終結因子の添加によって刺激される。最も一般的な粗製ディスプレイ系を使用する研究者が未完成ペプチドの急速な放出を回避するために選択前又は選択前の mRNA 結合体への融合前に mRNA から終止コドンを除くことが必要であることを発見した(融合及び選択の両方は、緩慢なプロセスである)。この終止コドンの除去は、通常、個々の mRNA の場合には特別の突然変異誘発段階を必要とし、しかも特に全ての終止コドンを変異させることが不可能である天然 mRNA ライブラリー又は合成コンビナトリアルライブラリーについては多くの問題がある。このライブラリーにおける終止コドンに関連する問題は、蛋白質の小さなサブドメインのライブラリー(このものは、全長の蛋白質を欠いており、十分に現れないカルボキシ末端ドメインを有し、正確に折りたたむ能力に対して多くの不適切な境界をコードし、しかも非天然のオープンリーディングフレームからの多数の配列を含有する)を無作為に生成させることによって、或いは終止コドン含有する無作為ライブラリーの要素を選択することによって回避されてきた(それによって、合成反応及び多様性の大部分を浪費する。「Choi 他(2000) *J. Mol. Biol.* 297, 309-319」)。精製リボソーム及び精製 mRNA ディスプレイ系では、ペプチジル tRNA が 1 日以上にわたってリボソームと安定に結合したままであることができる(翻訳が完了した後に、温度を低下させ及び/又は塩濃度を増加させる場合に安定していることが特に有利である(Schaffitzel 他(1999) *J. Immunol. Methods* 231, 119-135))：リボソームは、tRNA が与えられないセンスコドンで伸長を停止させるか(それによって終結因子との競合を完全に回避させる)又は終止コドンで停止させるか又は終止コドンを欠失した mRNA の末端で停止させるかのいずれかである。しかして、天然 mRNA からの全長翻訳産物のライブラリーが本発明により調製でき、そして、このような発現ライブラリーは、*in vitro* 選択に直接付され、或いはこれらのものは、特に、ハイブリダイゼーションによってゲノム及びプロテオミクス研究のための DNA マイクロチップに対応できる。

(iii) この純粋な再構成系は、終止コドンを欠失した mRNA から合成されたペプチドを分解させる tmRNA 系によって汚染されないことが予期される。粗製系を使用する研究者がこの tmRNA 系を阻害するように試みることを重要であることを発見した(Hanes 及び Pluckhoun(1997) *PNAS* 94, 4937-4942)。

## 【0021】

他の好ましい具体例では、本発明は、特定の疑似ペプチド（ペプチドアナログ）のmRNA指令合成を一般的な態様でできるようにし、得られる疑似ペプチドの多様性及び長さを大いに増大させることができる。見込まれるものとしては、既存の疑似ペプチドリガンド及び疑似ペプチド薬（シクロスポリンAのようなリボソームによらないで生合成されたりリガンドを含めて）並びにそれらの誘導体が挙げられる。

## 【0022】

他の具体例では、本発明は、in vitro 進化によって（例えば、上記の純粋なりボソームディスプレイ又は純粋なmRNAディスプレイを使用することによって）触媒、リガンド及び薬剤発見のために疑似ペプチド産物を遺伝学的にコードできるようにする。疑似ペプチドの特異的な合成は、既存の粗製リボソーム及びmRNAディスプレイの手法では不可能である。なぜならば、粗製翻訳系では、天然アミノ酸が非天然アミノ酸と取り込みのために競合するからである。

10

## 【0023】

本発明は、所望の特性を有するペプチド及び疑似ペプチドの単離を容易にさせる。一つ的具体例では、この方法は、標的分子のためのリガンドを同定することに関する。例として、標的分子にはペプチド、核酸、炭水化物及び非重合性分子、例えば、ステロイド、イノシトール、脂溶性ビタミン、テルペン、アセトゲニム、神経伝達物質又は遷移状態のアナログが含まれる。好ましい具体例では、標的分子は蛋白質である。蛋白質の標的は、例示すると、受容体、酵素、DNA結合蛋白質又は蛋白質複合体又はスクリーニングできる活性を保持しているそれらの部分若しくはドメインであることができる。

20

## 【0024】

好ましい具体例では、主題の方法を使用して少なくとも $10^3$ 種の異なる配列、好ましくは少なくとも $10^8$ 種の異なる配列、最も好ましくは少なくとも $10^{15}$ 種の異なる配列の試験ペプチド又は疑似ペプチドの多様な集団を生成させる。

## 【0025】

本発明のさらに別の態様は、主題の方法によって同定されたペプチド及び疑似ペプチドのような化合物及びそれらの使用に関する。また、これらのものには、このようなペプチド及び疑似ペプチドの結合体及び誘導体も含まれる（例えば、結合したペプチド又は疑似ペプチド配列を膜を横切って効率的に輸送できるようにする陽イオン性ペプチド配列への結合体）（Moore及びRosbash（2001）Science 294, 1841-1842）。

30

## 【0026】

本発明の別の態様は、ペプチド又は疑似ペプチドの合成及び/又は進化用のキットに関する。

## 【0027】

図面の説明

図1：E. coliからの5種のhisタグE. coli翻訳因子の過剰発現及び精製。15%ゲルでのSDS-PAGE後に、試料をクマシーブルーで染色した。Uは誘導されなかった全細胞であり、IはIPTG誘導された全細胞であり、Pは $Ni^{2+}$ ビーズから溶出した精製蛋白質であり、Mは分子量マーカーの蛋白質（寸法はkDで示される）である。

40

## 【0028】

図2：リボソーム指令ペプチド合成段階の概略図である。矢印によって示された3種の酵素反応は、開始（上）、第1伸長段階（右）、次の転位及び伸長段階（下）である。ペプチド産物は、分析のために塩基触媒加水分解によってペプチジルtRNAから放出され得る。産物のGDP及びPiは示されていない。Eは出口部位であり、Pはペプチジル部位、Aはアミノアシル部位である。

## 【0029】

図3：短鎖mRNA鋳型。最長mRNAを合成するのに使用したDNAプライマーと鋳型

50

の対を上部に示している。また、本発明の精製系からの予測される翻訳産物も示している（3'の終止コドンGAA (Glu)及びUUC (Phe)についてのアミノアシルtRNAは使用しなかった）。bKはピオチン標識されたリシンであり、S-Dはシャイン・ダルガーノリボソーム結合部位である。

#### 【0030】

図4：精製hisタグ翻訳系におけるmRNAからのオリゴペプチド合成速度の特徴付け。fMTVは、IF1H、IF2H、IF3H、EF-TuH、EF-GH及び0.020 A<sub>260</sub>/μlのリボソームを含有する翻訳系でペプチド産物に取り込まれた<sup>3</sup>H-バリンによって測定された。三角形：翻訳は、プレインキュベートされた開始成分をプレインキュベートされた伸長成分と混合することによって開始された。四角形：翻訳は、翻訳混合物を0 から37 に移すことによって開始された。アリコートをして1分から始める所定の時間にNaOHで終了させた。ペプチド産物のd.p.m.は、37 のインキュベーション前に終了されたアリコートで得られたd.p.m.を引き算することによって計算された。代表的な実験からの個々のデータのポイントがプロットされ、推定される偏差は20%未満であった。プレインキュベーション反応の曲線に対して接線を引いて定常状態の速度を推定する。

10

#### 【0031】

図5：mRNAのMTVを有する精製翻訳系における翻訳因子の依存性。淡い棒：ペプチド産物に取り込まれた<sup>3</sup>H-バリン (fMTVの測定値)は、IF2H、IF3H、EF-TuH及びEF-GHへの強い依存性を示している (IF1Hは、これらの翻訳系から除外された。材料及び方法を参照されたい)。黒い棒：同一の産物に取り込まれた<sup>14</sup>C-トレオニン (fMTとfMTVと一緒に測定したもの)。0 から37 に移すことによって開始された30分の翻訳でのペプチド合成は、全d.p.m.からmRNAなしの対照反応で得られたd.p.m. (最大のd.p.m.の1.3%)を引き算することによって計算された。得られた合成産物の最大濃度は、T及びV両者について0.12 μMであった。

20

#### 【0032】

図6：MTTVのmRNA鋳型から産生された産物のHPLC分析。二重標識翻訳のペプチド産物は、まず、tRNAから塩基により放出され、次いで標識されていないマーカープепチドと混合された。この混合物を酸性化させ、微量遠心分離して不溶性の物質を除去し、そして分析のために注入前に10kDフィルターによって微量遠心分離した (材料及び方法参照)。マーカープепチドの溶出位置は、クロマトグラムの上に示されている。黒丸は<sup>14</sup>C-トレオニンの全d.p.m.である。白丸は同一画分の<sup>3</sup>H-バリンの全d.p.m.である。産物の合成量は、30 μl当たり2.1 pmol (70 nM)であった。

30

#### 【0033】

図7：配列の包含は、精製翻訳系でのオリゴペプチド産物の合成を向上させる。mRNAであるMVT (丸)、スクランブルされたMTV (四角形)及びを欠失したMVT (三角形)からのfMVT合成の速度の間の比較。IF1H、IF2H、IF3H、EF-TuH及びEF-GHを含有する二重標識翻訳系を図5に記載されるように分析した。30分で合成されたオリゴペプチドの最大濃度は、それぞれのアミノアシルtRNAのうち0.5 μMを使用して0.25 μMであった。

40

#### 【0034】

図8：精製hisタグ翻訳系を使用した非天然アミノ酸を含むペプチドの合成及び選択。ピオチン標識lys-tRNA<sup>lys</sup>、fmet-tRNA<sub>i</sub><sup>fmet</sup>、thr-tRNA<sub>3</sub><sup>thr</sup>及び<sup>3</sup>H-val-tRNA<sub>1</sub><sup>val</sup>基質並びにmRNAであるMTKV又はMTV (翻訳産物については図3を参照されたい)のいずれかを含有する翻訳混合物を37 で30分間インキュベートした。ペプチド及びアミノ酸をtRNA及びリボソームから塩基により放出させ、中和し、そしてこの混合物を軟質のアビジンビーズと共にインキュベートしてピオチン含有分子を結合させた。結合された<sup>3</sup>H (黒い棒：<sup>3</sup>H-バリンに共有結合したピオチ

50

ン標識リシンを含有する産物の測定値)をカウントする前にこのビーズを4回洗浄してピオチニル化していない分子を除去した。プールされた洗浄物をろ過し、酸性化させ、そして陽イオン交換カラムに通して結合していない<sup>3</sup>H(淡い棒:<sup>3</sup>H-バリンを含有し、ピオチン標識リシン又はリシンを含有しないホルミル化ペプチド産物の測定値)をカウントした。結合した及び結合していない<sup>3</sup>Hのd.p.m.は、mRNAなしの対照反応で得られたd.p.m.(最大の結合及び非結合のd.p.m.のそれぞれ23%及び15%)を引き算した後にプロットしている。ピオチン標識lys-tRNA<sup>lys</sup>がmRNAのMTKVの翻訳から除外されたときに、<sup>3</sup>Hのビーズへの結合は観察されない(図示しない)。

【0035】

10

図9: Cys-tRNA<sup>cys</sup>の化学的ピオチニル化。

【0036】

図10: ピオチニル-Cys-tRNA<sup>cys</sup>の精製。ピオチニル-Cys-tRNAは、図16に示すように電気泳動され(以下参照)、その産物は、適切なバンド(即ち、レーン7又は9の大きい方のバンド)を切り出し、4で溶出させることによって精製された。tRNAにCysを付加させたものとさせていないもの、ピオチニル化させたもの(pH6.0又は6.9のいずれかで)とさせていないもの及びその後アミノアシル結合をTris-HCl pH8.8で加水分解させたものとさせていないものを含む対照(レーン1~6、8、10)は、ピオチニル化がCysに特異的であり、しかもCys-tRNAがピオチニル-Cys-tRNAと共移動しなかったことを示した。

20

【0037】

図11: 精製系を使用した精製ピオチニルCysのfM-T-bC-V疑似ペプチドへの取り込み。翻訳は、mMTCV、fMとTとVとをそれぞれ結合したtRNA及び非結合の、Cys結合の又は精製ピオチニル-Cys結合のtRNAを含み、対照はmRNAがないものであった。選択は図8に示すような軟質アビジンによるものであった。<sup>3</sup>Hペプチドのピオチニル-Cys基質との不完全な結合は、軟質アビジン(ほぼ10<sup>-7</sup>Mの非常に高いK<sub>d</sub>を有するように誘導体化されたアビジン)の単一のピオチンを含有するペプチドに対する低い親和性のためであると思われた。

【0038】

図12: 隣接する多数の非天然アミノ酸をfM-T-bC-bC-V疑似ペプチドに取り込むためのアッセイ。この実験は、図11に示すように実施された。軟質アビジンによる結合は、1個以上のピオチンを含有するペプチドに対する非常に低い親和性のために予想されるように、fM-T-bC-bC-Vについては完全であったが、fM-T-bC-Vについては不完全であったことに留意されたい。

30

【0039】

図13: 全tRNAと15種の異なる<sup>14</sup>C標識アミノ酸の混合物(ニュー・イングランド・ニュークレア社)との結合活性(TCA沈殿法によって測定される)についての本発明の純粋な翻訳系のアッセイ(「蛋白質ポリメラーゼ」)。添加されたシンテターゼは、tRNAを含まない粗製のアミノアシルtRNAシンテターゼ細胞抽出物からなる。

【0040】

40

図14: 選択のコドンに対して特異的な非天然アミノ酸を結合したアミノアシルtRNAを合成するために一般化できる方法。合成オリゴデオキシリボヌクレオチドから調製された本発明の組換えDNAクローンから試験管内で合成された伸長体tRNA<sup>Asn-CA</sup>(図16のレーン5)は、矢印で示される天然tRNA<sup>Asn</sup>からの実質的な塩基改変物を含有する。このtRNAのアンチコドンは、大きい文字で示されている。アミノNVOCで保護された天然アミノ酸は、pdcA上で化学的にアミノアシル化され(右上参照)、次いでtRNA<sup>Asn-CA</sup>(FokI切断鑄型のランオフ転写法によって形成された)にT4RNAリガーゼで連結された(図16、レーン6)。この方法は一般化できる。なぜならば、アミノアシルtRNAシンテターゼ又は天然tRNAが全く必要とされなかったからである。

50

## 【0041】

図15：左から右に tRNA<sup>Asn</sup>(T) tRNA<sup>Asn</sup>(S) tRNA<sup>Asn</sup>(V) と命名された tRNA<sup>Asn</sup>(N) の3種のアンチコドン変異体。これら tRNA の新規なアンチコドンは、これらの tRNA が認識するコドンの上に大きい文字で示されている。この遺伝暗号は、コドンが、どちらのアミノ酸(天然又は非天然)がそれぞれの tRNA<sup>Asn-CA</sup> に連結するために選択されるのかを直ちに指定するように再設計された。

## 【0042】

図16：非結合 tRNA (レーン1、3及び5) 及びアミノアシル tRNA 基質 (レーン2、4及び6) の酸/尿素ポリアクリルアミドゲル電気泳動 (Varshney 他 (1991) J. Biol. Chem. 266, 24712-24718)。このゲル系において、遊離の RNA 種 (レーン1、3及び5) の観察された移動度は、アミノアシル化 (レーン2、4) によって、或いはアミノアシル化 CA ジヌクレオチド (レーン6) の連結によって抑制された。

10

## 【0043】

図17：隣接する aG アミノ酸を疑似ペプチドに選択的に取り込むための一般的な方法。この実験は、図8及び11と同様に実施したが、ただし選択段階は実施しなかった。ここで、tRNA<sup>Asn</sup>(N) は Asn をベースとして tRNA<sup>AsnB</sup> と略記されていることに留意されたい。

## 【0044】

図18：純粋リボソームディスプレイ。これは、ペプチドと疑似ペプチドの反復合成及び選択(定向進化)のために使用できる Mattheakis 他(上記)の粗製 in vitro 系(生体物質なしの)の純粋なタイプである。疑似ペプチドは、選択されるためには非放出 de novo 合成ペプチドを取り囲むほぼ 100 の長さのリボソームトンネルを超える程度に長いことが必要である。mRNA (及びリボソーム)からのペプチドの分離は、終止コドン、終結因子又はある種のアミノアシル tRNA を除外し且つリボソームを立ち往生させる抗体を使用することによって防止できる。

20

## 【0045】

図19：本発明のスペーサー mRNA 及びそれらのコードペプチド産物。

## 【0046】

図20：本発明の精製系を使用した図19のスペーサー mRNA の翻訳。ペプチド産物に取り込まれた伸長体バリン(<sup>3</sup>H 標識)対開始体ホルミルメチオニン(<sup>35</sup>S 標識)の比率は、二重標識 d.p.m. 計数プログラムを使用して TCA 沈殿産物を分析することによって測定された。観察された直線のグラフは、高い連続移動性について予期されるものである。

30

## 【0047】

図21：純粋 mRNA ディスプレイ。これは、Mattheakis 他(図18)の粗製 in vitro 系を基にした根本他(上記)並びに Roberts 及び Szostak (上記)の粗製 in vitro 系の純粋なタイプであって、mRNA をプロマイシン(Pm)と結合させてその mRNA をそれらのペプチド又は疑似ペプチド産物に共有結合的に融合させるようにできるという点で異なるものである。しかして、この mRNA と疑似ペプチドの融合体は、選択前にその他の翻訳成分から精製できるため、非常に短い疑似ペプチドをリボソームトンネルによって遮蔽することなく表示できる。この融合反応は緩やかであるため、終止コドンを含む mRNA を使用するとき終結因子を除外して終結因子が触媒するペプチドの放出を防止することが重要である。典型的には、リボソームを mRNA に結合したデオキシヌクレオチド配列によって立ち往生させる。本発明者は、mMTKV 及び mMTCV (図8及び11)を、注文合成された「添え木」DNA (mRNA 及び pdA<sub>27</sub>dCdCPm のいずれかの末端にハイブリダイズさせるように設計された TTTTTTTTTTAAATTC AAC) 及び T4 DNA リガーゼを使用して pdA<sub>27</sub>dCdCPm に効率的に連結させ(トリ連結)、次いで融合及び選択実験に使用するためにこの結合体をゲル精製した (Roberts 及び Szostak (1997) PNAS 94, 12297

40

50

- 1 2 3 0 2 )。ここで、図中の X は、C 又は b C ( 図 9 ) 又は b K ( 図 2 2 ) のような容易に選択できるアミノ酸である。

【 0 0 4 8 】

図 2 2 : プロメガ社製の T r a n s c e n d ビオチニル - リシン t R N A <sup>lys</sup> ( E . c o l i 、材料及び方法を参照)。

【 発明を実施するための最良の形態 】

【 0 0 4 9 】

#### 発明の詳細な説明

##### ( I ) 概説

本発明は、複数の異なる天然及び非天然アミノ酸残基を高度に制御され且つ一般化できる態様で特異的に取り込めるようにするための精製成分から再構成された *in vitro* 翻訳系に関する。ある種の野生型のアミノ酸及び終結因子による望ましくない競合を取り除くため、遺伝学的にコードされた小さい及び長い疑似ペプチドの分子及びライブラリーを効率的且つ特異的に合成することが可能である。 10

【 0 0 5 0 】

本発明は、斯界の多数の研究者による長年にわたる相当な努力にもかかわらず未解決のままであった遺伝学的にコードされた疑似ペプチドを合成することのできる生合成方法に対する長く切望された必要性のために開発された。

【 0 0 5 1 】

また、本発明は、効率的な連続移動性のために E F - P、W、W 2 又はレスキューの添加を必要としない、単純化され、高度に精製された翻訳系に対する長く切望された必要性のためにも開発された。斯界におけるこのような系の最も高度に精製されたタイプでの翻訳の連続移動性の問題及び明らかな解答のないこの分野における 1 0 年間の研究を基にすれば、成功する見込みは小さいように思われる。 20

【 0 0 5 2 】

##### ( II ) 用語の定義

ここで、便宜上、明細書、実施例及び添付された特許請求の範囲で使用したある種の使用語を収録する。

【 0 0 5 3 】

「蛋白質」とは、1 個以上のペプチド結合によって結合された任意の 2 個以上の天然型アミノ酸を意味する。ここでは、「蛋白質」及び「ペプチド」が互換性を持つように使用される。 30

【 0 0 5 4 】

用語「ペプチド」とは、それらの単量体がアミド結合によって互いに結合された天然アミノ酸 ( - アミノ酸 ) であるオリゴマーをいう。ペプチドは、2 個以上のアミノ酸単量体の長さであるが、たいていは 5 ~ 1 0 個のアミノ酸単量体の長さであり、さらにはそれよりも長い、即ち 2 0 個まで又はそれ以上のアミノ酸であることができるが、2 0 個のアミノ酸よりも長いペプチドは、「ポリペプチド」と呼ぶのが適当である。用語「蛋白質」とは、斯界に周知のものであり、いくつかの生物学的機能を有する非常に大きなポリペプチド又は会合した相同若しくは異種ポリペプチドをいう。本発明の目的のために、用語「ペプチド」、「ポリペプチド」及び「蛋白質」は、3 つのタイプ全てが翻訳系によって合成できるときにはたいてい置き換えることができ、そのため一括して蛋白質と呼ばれる。 40

【 0 0 5 5 】

「疑似ペプチド」とは、非天然主鎖を有する産物及びさらにエステル、チオエステル若しくはその他の結合で部分的に又は全体的に置換されたアミド ( ペプチド ) 結合を有するもの ( M e n d e l、前述 ) を含めて、1 種以上の非天然アミノ酸 ( 例えば、非天然側鎖、非天然キラリティー、N - 置換アミノ酸若しくは - アミノ酸 )、非天然トポロジー ( 例えば、環状若しくは分岐 ) 又は非天然化学誘導体 ( 例えば、メチル化若しくは末端がブロックされたもの ) を含むペプチドアナログ、或いは天然アミノ酸を含むペプチド以外の任意の分子であってリボソームによって合成されるものを意味する。 50

## 【0056】

用語「アミノ酸残基」及び「ペプチド残基」とは、アミノ酸又はペプチド分子であってそのカルボキシル基の-OHがないもの（C末端が結合されたもの）又はそのアミノ基のプロトンがないもの（N末端が結合されたもの）を意味する。一般的に、アミノ酸及び保護基を示すためにここで使用される略称は、IUPAC-IUB生化学命名法委員会の提言に基づくものである（Biochemistry（1972）11：1726-1732参照）。ペプチド中のアミノ酸残基は、次のように略記する。アラニンはAla又はA、システインはCys又はC、アスパラギン酸はAsp又はD、グルタミン酸はGlu又はE、フェニルアラニンはPhe又はF、グリシンはGly又はG、ヒスチジンはHis又はH、イソロイシンはIle又はI、リシンはLys又はK、ロイシンはLeu又はL、メチオニンはMet又はM、アスパラギンはAsn又はN、プロリンはPro又はP、グルタミンはGln又はQ、アルギニンはArg又はR、セリンはSer又はS、トレオニンはThr又はT、バリンはVal又はV、トリプトファンはTrp又はW及びチロシンはTyr又はYである。ホルミルメチオニンはfMet又はfMと略記する。用語「残基」とは、相当する - アミノ酸からそのカルボキシル基のOH部分及び - アミノ基のH部分を取り除くことによって誘導される基を意味する。用語「アミノ酸側鎖」とは、「K. D. Kopple, "ペプチド及びアミノ酸", W. A. Benjamin社, ニューヨーク及びアムステルダム, 1966年, 2及び33ページ」によって定義されるように、-CH(NH<sub>2</sub>)COOH部分を除いたアミノ酸の一部分のことである。一般的なアミノ酸のこのような側鎖の例は、-CH<sub>2</sub>CH<sub>2</sub>SCH<sub>3</sub>（メチオニンの側鎖）、-CH<sub>2</sub>(CH<sub>3</sub>)-CH<sub>2</sub>CH<sub>3</sub>（イソロイシンの側鎖）、-CH<sub>2</sub>CH(CH<sub>3</sub>)<sub>2</sub>（ロイシンの側鎖）又は-H（グリシンの側鎖）である。

10

20

## 【0057】

ある具体例では、本発明の用途で使用されるアミノ酸は、蛋白質中に見いだされる天然型のアミノ酸である。特に好適なアミノ酸側鎖には、次の21種の天然アミノ酸：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、グルタミン酸、グルタミン、グリシン、ヒスチジン、イソロイシン、ロイシン、リシン、メチオニン、フェニルアラニン、プロリン、セレノシステイン、セリン、トレオニン、トリプトファン、チロシン及びバリンの側鎖から選択されるものが含まれる。しかしながら、本発明は、ここで言及される任意の特定のアミノ酸のアナログ、誘導體及び同族体を使用することを意図する。例えば、本発明は、放射性アミノ酸アナログ、側鎖を長くし又は短くすると同時に重合のためのカルボキシル、アミノ又はその他の反応性の前駆体官能基をさらに与えるアミノ酸アナログ並びに異なる側鎖（適切な官能基を有する）を有するアミノ酸アナログを使用することを意図する。例えば、主題の疑似ペプチドは、例えば - シアノアラニン、カナバニン、ジエンコル酸、ノルロイシン、3-ホスホセリン、ホモセリン、ジヒドロキシフェニルアラニン、5-ヒドロキシトリプトファン、1-メチルヒスチジン、3-メチルヒスチジン、アリルグリシン（又はそのアルキン誘導體）、O-メチルセリン、ピオチニルリシン、ピオチニルシステイン（又はその他のピオチン標識アミノ酸）、シクロヘキシルアラニン、ホモグルタメート、D-アラニン（又はその他のD-アミノ酸）、N-メチルグリシン（又はその他のN-メチルアミノ酸）、-N-メチルリシンとしてのアミノ酸アナログ及び21種の天然アミノ酸又は非天然アミノ酸の放射性同位元素誘導體を包含できる。ここで好適なその他の天然型アミノ酸又は非天然型アミノ酸は、当業者であれば容易に認識され且つ本発明の範囲に含められる。

30

40

## 【0058】

用語「キラル」とは、鏡像の相手を重ね合わせることができないという特性を有する分子をいい、一方、用語「アキラル」とは、その鏡像の相手に重ね合わせることができる分子をいう。

## 【0059】

用語「立体異性体」とは、同一の化学構成を有する化合物であるが、原子又は基の空間配置に関して異なるものをいう。特に、「鏡像異性体」とは、互いの鏡像を重ね合わせること

50

ができない化合物の2種の立体異性体をいう。一方、「ジアステレオマー」とは、非対称である2つ以上の中心を有する立体異性体であってその分子が互いに鏡像でないものをいう。キラル中心の命名に関して、用語「D」及び「L」立体配置は、IUPACの推奨によって定義されるようなものである。この用語の使用に関して、ジアステレオマー、ラセミ化合物及び鏡像異性体は、ペプチド調製物の立体化学を説明するためにその一般的な状況で使用されるであろう。

【0060】

用語「D-アミノ酸」及び「L-アミノ酸」とは、それぞれ、グリセルアルデヒドの予想される立体異性体に関する規則による絶対配置を意味する。しかして、所定の異性体による偏光面の回転方向にかかわらず、立体化学的にL-グリセルアルデヒドに関わる全ての立体異性体はL-と表記され、D-グリセルアルデヒドに関わるものはD-と表記される。トレオニン及びイソロイシンの場合には、2つの立体化学中心、即ち、C<sub>2</sub>及びC<sub>3</sub>原子がある。ここで使用されるD-トレオニン及びD-イソロイシンは、好ましくは、両方のキラル部位においてこれらのアミノ酸のL-立体異性体の立体配置とは逆である(鏡像異性である)立体配置を有する。例えば、これらのものは完全に鏡像である。グリシンは一般的に生じる唯一のアキラルなアミノ酸である。グリシンのようなアキラルなアミノ酸残基の存在は、そのキラリティーの名称に影響を及ぼさない。

10

【0061】

自然界で *de novo* 合成される蛋白質の全てのキラルなアミノ酸(例えば、「天然型」)は、L-アミノ酸である。

20

【0062】

「D-鏡像異性体」又は「D-ペプチド鏡像異性体」とは、L-アミノ酸ではなく、D-アミノ酸残基からなるペプチドをいう。

【0063】

用語「合成」とは、試験管内での化学的又は酵素的合成による産生をいう。

【0064】

DNAとは、2個以上の共有結合された天然型又は修飾デオキシリボヌクレオチドの配列を意味する。

【0065】

「RNA」とは、2個以上の共有結合された天然型又は修飾リボヌクレオチドの配列を意味する。この用語に包含される修飾RNAの2つの例は、ホスホリボチオエート型RNA及び天然の修飾塩基を含有する「トランスファーRNA」である。

30

【0066】

「トランスファーRNA」又は「tRNA」とは、ペプチド又は疑似ペプチド前駆体をmRNAへの部分的な塩基の対合によって特定される態様でリボソームに運搬できる任意のRNA分子を意味する。

【0067】

「翻訳開始配列」とは、機能性のリボソーム進入部位を与えることができる任意の配列を意味する。細菌の系では、この領域は、時としてリボソーム結合配列又はシャイン・ダルガーノ配列と呼ばれる。

40

【0068】

「メッセンジャーRNA」又は「mRNA」とは、「翻訳開始配列」を含有する任意の核酸を意味する。

【0069】

「再構成翻訳系」とは、*in vitro* 翻訳、例えばmRNA依存性蛋白質合成を行うことのできる反応混合物(a)及び細胞溶解物の翻訳系又は麦芽抽出物の翻訳系に見いだされる汚染性蛋白質の10%未満を有し、さらに好ましくはこのような汚染性蛋白質の5%未満又はさらに1%未満を有することを特徴とする反応混合物(b)をいう。ある好ましい具体例では、主題の再構成翻訳系は、組換えによって生産された及び/又は精製された蛋白質を混合することによって生成される。

50

## 【0070】

「開始コドン」とは、蛋白質コード配列の始まりを表す3種の塩基を意味する。一般的には、これらの塩基は、AUG（又はATG）であるが、この態様で使用できる任意のその他の塩基のトリプレットは置換されていてよい。

## 【0071】

「終止コドン」とは、蛋白質コード配列の終わりを表す3種の塩基を意味する。一般的には、これらの塩基は、UAG、UAA又はUGA（ここで、UはTで置換されていてよい）であるが、この態様で使用できる任意のその他の塩基のトリプレットは置換されていてよい。

## 【0072】

「休止配列」とは、リボソームにその翻訳速度の遅延及び停止を生じさせる核酸配列、例えばDNA配列を意味する。

## 【0073】

「ペプチドアクセプター」とは、リボソームのペプチジルトランスフェラーゼ機能の触媒活性によって、成長しつつある蛋白質鎖のカルボキシル末端に付加できる任意の分子を意味する。典型的には、このような分子は、(i)ヌクレオチド又はヌクレオチド様の部分（例えば、アデノシン又はアデノシナンalog（N-6アミノ部位でのジメチル化が許容できる））、(ii)アミノ酸又はアミノ酸様の部分（例えば、20種のD-又はL-アミノ酸又はそれらの任意のアミノ酸アナログのいずれか（例えば、O-メチルチロシン又は「Eilman他, Meth. Enzymol. 202:301, 1991」に記載されたアナログのうち任意のもの））及び(iii)これら2種の間での結合（例えば、エステル、アミド又は3'の位置若しくはあまり好ましくはないが2'の位置でのケトン結合）（好ましくは、この結合は、天然リボヌクレオチドのコンホメーションによる環の縮みを有意に摂動させない）を含有する。また、ペプチドアクセプターは求核基も有し、このものは、アミノ基、ヒドロキシル基又はスルフヒドリル基であることができるが、これらに制限されるわけではない。さらに、ペプチドアクセプターは、ヌクレオチドの疑似体、アミノ酸の疑似体又はヌクレオチド・アミノ酸の結合構造の疑似体から構成されていてよい。

## 【0074】

「それぞれのコドンでの高選択的な取り込み」とは、コドンに相当するペプチド又は疑似ペプチド中のある位置でのアミノ酸残基の少なくとも80%の選択的な取り込み、より好ましくは少なくとも90、95又はさらには98%の選択的な取り込みを意味する。

## 【0075】

ペプチドアクセプターが蛋白質コード配列の「3'末端」に配置されるとは、ペプチドアクセプター分子がその蛋白質コード配列の最後のコドンの後に配置されることを意味する。この用語には、蛋白質コード配列の3'末端に正確に配置されるペプチドアクセプター分子並びにコード配列又は非コード配列を介在させることによって最後のコドンから隔てられるもの（例えば、休止部位に相当する配列）が含まれるが、これに限定されるものではない。また、この用語には、コード又は非コード配列がペプチドアクセプター分子に追従する（即ち、3'からそれに）構築物も含まれる。さらに、この用語には、蛋白質コード配列に共有結合される（直接的に又は核酸配列を介在させることによって間接的に）ペプチドアクセプター分子並びに蛋白質コード配列にいくつかの非共有結合の手段によって、例えば、蛋白質コード配列の3'末端に又はその近傍に結合する第2の核酸配列を使用するハイブリダイゼーションによって結合し、しかもそれ自体がペプチドアクセプター分子に結合するものが含まれるが、これに限定されるわけではない。

## 【0076】

ペプチドアクセプターに「共有結合する」とは、ペプチドアクセプターが「蛋白質コード配列」に共有結合によって直接的に、或いは別の共有結合した配列（例えば、休止部位に相当するDNA）によって間接的に結合することを意味する。

## 【0077】

10

20

30

40

50

「mRNAディスプレイ」とは、共有又は非共有結合によって、mRNAの翻訳からのペプチド又は疑似ペプチド産物をそのペプチド又は疑似ペプチド産物に関わる同種の（コグネイト）mRNA又はcDNA又はヌクレオチド配列に結合させるための任意の方法を意味する。一般的に使用される結合はプロマイシンを含む。

【0078】

「変化した機能」とは、ある分子の機能における任意の定性的な又は定量的な変化を意味する。

【0079】

ここで使用するとき、「結合の相手」とは、所望のmRNA・ペプチド/疑似ペプチドのリボソーム複合体又は融合体の一部分に対する特異的で共有結合的又は非共有結合的な親和性を有する任意の分子を意味する。結合の相手の例として、抗原/抗体の対、蛋白質/阻害剤の対、受容体/リガンドの対（例えば、ホルモン受容体/ペプチドホルモンの対のような細胞表面受容体/リガンドの対）、酵素/基質の対（例えば、キナーゼ/基質の対）、レクチン/炭水化物の対、オリゴマー又はヘテロオリゴマー蛋白質集合体、DNA結合蛋白質/DNA結合部位の対、RNA/蛋白質の対及び核酸の二本鎖、ヘテロ二本鎖又は連結鎖並びにRNA・蛋白質融合体の任意の部分と1個以上の共有又は非共有結合（例えば、ジスルフィド結合）を形成することができる任意の分子よりなる構成要素が挙げられるが、これらに限定されない。

10

【0080】

用語「リガンド」とは、蛋白質、例えば受容体のような特定の標的によって認識される分子をいう。標的が結合する又はそれと反応する任意の薬剤が「リガンド」と呼ばれるため、この用語には、酵素の基質及び触媒反応の反応体が含まれる。用語「リガンド」とは、任意の特定の分子の大きさ又は問題の物質が標的と結合でき、さもなくばそれと相互作用できること以外のその他の構造的特徴又は組成的特徴を意味するものではない。「リガンド」は、標的が結合する天然のリガンド又はアゴニスト若しくはアンタゴニストとして作用できる機能性アナログとしての役割を果たし得る。

20

【0081】

用語「基質」とは、酵素のリガンドであって触媒として作用し且つ酵素によって生成物に化学的に変換されるものをいう。

【0082】

用語「受容体」とは、所定のリガンドに対して親和性を有する分子をいう。

30

【0083】

用語「固体担体」とは、硬質又は半硬質の表面を有する材料をいう。このような材料は、好ましくは、小型のビーズ、ペレット、ディスク、チップ、皿状物、マルチウェルプレート、ウエハーなどの形をとるであろうが、その他の形を使用してもよい。いくらかの具体例では、この材料の少なくとも一つの表面は、実質的に平坦である。用語「表面」とは、ある固形物上の任意の一般的には二次元構造をいい、表面が無くならなければ段差、隆起、屈曲、テラスなどがあってもよい。

【0084】

ここで使用するとき、「集団」とは、1種以上の分子（例えば、1種以上のRNA、DNA、RNA・蛋白質又はRNA・疑似ペプチド融合分子）を意味する。本発明の方法は、所望ならば多数の候補分子から始める選択を容易にするものであるから、本発明に従う「集団」とは、好ましくは、少なくとも $10^3$ 種の異なる配列、さらに好ましくは少なくとも $10^8$ 種の異なる配列、そして最も好ましくは少なくとも $10^{15}$ 種の異なる配列を意味する。

40

【0085】

用語「無作為ペプチドライブラリー」とは、無作為の又は半無作為のペプチド集団並びにこれらの無作為ペプチドを含有する融合蛋白質の集団（応用できるように）を意味する。疑似ペプチドの1種以上の残基が非天然型アミノ酸様の部分であるということ認識した上で、同様の意味が用語「無作為疑似ペプチドライブラリー」に与えられる。

50

## 【0086】

「選択」とは、ある分子をある集団内のその他の分子から実質的に分離することを意味する。ここで使用されるときに、「選択」段階は、この選択段階の後に所望の分子を集団中の望ましくない分子に対して少なくとも2倍、好ましくは10倍、さらに好ましくは100倍、最も好ましくは1000倍以上富化させる。ここに示されるときに、選択段階を何度でも繰り返してよく、また、異なるタイプの選択段階を所定の方法に組み合わせてもよい。

## 【0087】

試験ペプチド又は疑似ペプチドを標的と結合させることに関して、句「個々に選択的な態様」及び「個々に選択的な結合」とは、標的の分子同一性に対して特異的な結合及びそれに依存した結合をいう。

10

## 【0088】

用語「差異結合手段」並びに「親和性選択」及び「親和性富化」とは、ペプチド又は疑似ペプチドディスプレイライブラリーの構成要素を、このライブラリーのディスプレイパッケージのそれぞれについて試験ペプチド又は疑似ペプチドを標的に結合する異なる能力に基づいて分離させることをいう。ディスプレイ分子による標的の差異結合は、標的と特に相互作用する分子を相互作用しない分子から親和性によって分離するのに使用できる。親和性選択の例として、アフィニティークロマトグラフィー、沈殿法、蛍光表示式細胞分取法及びブランクリフト法が挙げられる。以下に記載するように、アフィニティークロマトグラフィーには、例えば固定化された標的蛋白質を使用するパンニング技術が含まれる。

20

## 【0089】

用語「レポーター群」又は「タグ」とは、他の分子に結合したとき及び化学的な分離プロセスで利用されたときに容易に検出できる原子、化合物又は生体分子若しくは複合体をいう。レポーター群は、例えば、蛍光性若しくは放射性原子又は1個以上のこのような原子を含有する化合物であることができる。

## 【0090】

アミノアシル tRNA シンテターゼの「ブルーフリーディング」活性とは、同種でない（非コグネイト）アミノ酸（天然又は非天然）を認識し、次いでこれをシンテターゼのアミノアシル化された同種の tRNA アイソアクセプターから除去するというこの酵素の加水分解触媒活性を意味する。

30

## 【0091】

## (III) 代表的な具体例

一般に、本発明の方法は、ペプチド又は疑似ペプチドを生成させる *in vitro* 又は *in situ* 転写/翻訳プロトコルからなる。これは、mRNA 分子を天然型及び/又は非天然型アミノ酸又はアミノ酸アナログを結合した1種以上の tRNA 分子で合成し且つ *in vitro* 又は *in situ* 翻訳することによって達成される。本発明者は、細菌の翻訳が EF-P、W、W2 又はレスキューを添加することなく再構成できることを発見した。

## 【0092】

一般に、好ましい最小の翻訳系は、次の高分子成分：リボソーム、mRNA、アミノアシル tRNA 及び翻訳因子である IF2H、IF3H、EF-TuH、EF-GH を含む。IF1H と EF-Ts は促進性であるので、しばしば添加される。更なる別法を以下に詳述する。

40

## 【0093】

標識開始因子の高レベル過剰発現用のクローンが入手できないため、精製翻訳系を再構成するために大量の高度に精製された因子を再現可能に調製することが大きな技術的障害であった。実際に、多くの研究が主要成分の資質に依存していた。この問題を克服するために、本発明者は、(His)<sub>6</sub>タグ(hisタグと呼ばれる)を有する3種の E. coli 開始因子全てをサブクローニングし且つ過剰産生させ、精製翻訳系でそれぞれの因子の活性を検査した。

## 【0094】

50

his タグ E. coli 翻訳因子のサブクローニング、過剰発現及び精製。

公表された E. coli の IF1、IF2 及び IF3 発現用クローンは、多数の初期研究には有用ではあるものの、親和精製できず且つ一つの例外を除いて熱誘導できる非標識因子のためのものである (「Calogero 他 (1987) Mol Gen Genet 208, 63-9」、  
「Laalami 他 (1991) J Mol Biol 220, 335-49」、  
「De Bellis 及び Schwartz (1990) Nucleic Acids Res 18, 1311」、  
「Mortensen 他 (1991) Biochimie 73, 983-9」、  
「Brombach 及び Pon (1987) Mol Gen Genet 208, 94-100」)。IF3 の過剰発現について考慮すべき重要な事項は、希少な AUU 開始コドンが存在することである (Brombach 及び Pon (1987) Mol Gen Genet 208, 94-100)。fmet-tRNA<sub>i</sub><sup>fmet</sup> と対合する開始コドンは、IF3 によって直接校正され (Meinzel 他 (1999) J Mol Biol 290, 825-37)、それによって生体内で自身の遺伝子の翻訳を IF3 仲介フィードバック抑制することができる (Brombach 及び Pon (1987) Mol Gen Genet 208, 94-100)。従って、発現レベルを増大させ且つ3種の開始因子全ての精製を単純化させるために、本発明者は、開始コドンを PCR 法によって AUG (CAC)<sub>6</sub> 配列で置き換え、そのコード配列を pET 由来の発現用プラスミドにサブクローニングした (材料及び方法参照)。得られた his タグクローンは、非毒性であり且つ溶解物の可溶性画分中でそれら因子 (IF1H、IF2H 及び IF3H と呼ぶ) を非常に高いレベルで過剰産生した (図1、レーン1、2、5、6、8、9)。さらに、N末端 his<sub>6</sub> タグを有する E. coli の EF-Tu 及び EF-G (EF-TuH 及び EF-GH と呼ぶ) を市販のクローンを使用して過剰発現させた (図1、レーン12、13、15、16 (「Hwang 他 (1997) Arch Biochem Biophys 348, 157-62」、  
「Semenko 他 (1996) Proc Natl Acad Sci USA 93, 12183-8」))。また、Ni<sup>2+</sup> 精製開始因子及び伸長因子のゲル分析も図1に示す (レーン3、7、10、14 及び 17)。IF2H、IF3H、EF-TuH 及び EF-GH は、真正の E. coli 蛋白質と共に移動し (データは示さない)、IF1H は、分子量マーカとの比較に基づいて予期された電気泳動度を有していた (図1、レーン1~4)。IF2 のより短い形の IF2-2 は、おそらく IF2H と共に過剰発現したが (図1、レーン6)、内部翻訳開始機構によるその合成 (Sacerdot 他 (1992) J Mol Biol 225, 67-80) は his<sub>6</sub> タグを取り込まないであろうから、IF2H と共に精製しなかった (図1、レーン7)。

#### 【0095】

開始アッセイにおける his タグ開始因子の依存性。

3種の開始因子の活性及び純度を、リボソーム : fmet-tRNA<sub>i</sub><sup>fmet</sup> : ApUpG のトリヌクレオチド複合体構成によって測定した (図2、頂部線)。3x 塩洗浄リボソームを有する開始複合体構成中での his タグ開始因子の依存性は、ネイティブ因子について報告された依存性に匹敵した (表1、  
「Kung 他 (1974) Arch Biochem Biophys 162, 578-84」、  
「Dubnof 及び Maitra (1972) J Biol Chem 247, 2876-83」)。IF3 依存の変差は、異なる 70S リボソーム調製物中の異なる量の遊離 30S リボソームサブユニットに起因する可能性がある。なぜならば、IF3 は、このアッセイで 70S リボソームの代わりに遊離 30S サブユニットを使用すると、開始複合体構成物を強くは刺激しないからである (Canonaco 他 (1987) Biochimie 69, 957-63)。開始因子がリボソームと共に精製される傾向があれば (Kung 他 (1974) Arch Biochem Biophys 162, 578-84)、この結果もリボソームの純粋さを裏付けるものである。AUG トリヌクレオチド鋳型を MTTV mRNA 鋳型と非常に低濃度 (0.3 μM) で置き換えると (図3、以下に説明される)、同等の収率の複合体構成が可能になる (データは示さない)。

## 【0096】

トリペプチド合成におけるhisタグ因子の依存性。

本発明者は、次に、開始複合体が図2に示される成分を使用する精製翻訳系で伸長を受けることができたかどうかを試験した。本発明者が選択したmRNAのデザイン(Pavlov他(1997) *Biochimie* 79, 415-22)は、最適なシャイン・ダールガーノ配列及び翻訳エンハンサー(Olins及びRangwala(1989) *J Biol Chem* 264, 16973-6)を含む。本発明のデザイン(図3)は、標準的な18merオリゴデオキシリボヌクレオチドにハイブリダイズされる一本鎖の長い合成DNAの鋳型から直接合成される程度に短いmRNAを、クローン化する必要なくコードするという利点を有する(Milligan及びUhlenbeck(1989) *Methods in Enzymology* 180, 51-62)。この鋳型の翻訳は、リボソームがあるコドンに転位し、この場合に同種の(コグネイト)tRNAが全く供給されないとき完了する(Weissbach他(1984) *BioTechniques* 2, 16-22)。

10

## 【0097】

図4に示すように、3x塩洗浄リボソーム及びMTV mRNA(図3)を使用して図2に示される成分の全てによりオリゴペプチド合成を再構成すると、<sup>3</sup>H標識バリンの単離ペプチド産物への取り込みによって判断されるように、fMTVの合成に帰着する。合成のタイムコースがブレインキュベートされた開始因子の混合物をブレインキュベートされた伸長因子の混合物と結合させることによって開始されるときには(図4の三角形、凡例参照)、最初の1分以内での産物合成の迅速な初期バースト及び定常状態での緩やかな合成速度が存在する。10分間のブレインキュベーションなしでは(図4の四角形、凡例参照)、最初の1分以内では産物がほとんど合成されず、また合成は、時間経過につれてほぼ直線である(図7も参照)。

20

## 【0098】

図5は、<sup>14</sup>C標識トレオニン及び<sup>3</sup>H標識バリンのペプチド産物への取り込みに基づいて、*in vitro* 翻訳に必要なそれぞれの因子(IF1はこの実験では除外された)への依存性を示している。この完全な系は、fMTV合成について予期されるように、1.0のトレオニン/バリン(T/V)比のペプチド産物を産生した。任意の1種の因子の除外は、fMTV合成を劇的に減少させ(図5)、ネイティブ因子について報告された依存性に匹敵する依存性を与える(「Robakis他(1981) *Proc Natl Acad Sci USA* 78, 4261-4」、*Cenatiempo*他(1982) *Arch Biochem Biophys* 218, 572-8)。翻訳因子EF-GHの除外は、転位におけるEF-Gの既知の主要な役割と一致して翻訳をトリペプチド合成からfMTジペプチド合成に転換させた(図5)。ジペプチドの合成はIF1の添加を必要とせず(Robakis他(1981) *Proc Natl Acad Sci USA* 78, 4261-4)、しかも本発明の翻訳系にIF1Hを含めても全体的な収率の劇的な増加には至らないが、IF1Hは、前に報告されたネイティブIF1の研究(Robakis他(1981) *Proc Natl Acad Sci USA* 78, 4261-4)と一致して、ブレインキュベーションなしで実施される翻訳の最初の数分間にfMTV合成の速度を2.5倍促進させる(データは示さない)。従って、この翻訳の結果は、開始アッセイからの知見を確認し且つ拡大するものである(表1)。翻訳系のその他の高分子成分を除外した追加の対照の翻訳は、予期された依存性を示す(材料及び方法参照)。1回だけ洗浄されたリボソームは、図5に示されるタイプの翻訳アッセイでは不十分な因子依存性を示した(データは示さない)。

30

40

## 【0099】

テトラペプチド合成。

次に、本発明者は、開始及び伸長の全ての段階が高度に精製された系で起こり得ることを示すために、テトラペプチド合成についてのこの系の適合性を調査した。トリペプチド合成とは対照的に、テトラペプチド合成は、脱アシル化tRNAのリボソーム出口部位から

50

の解離なしにはあり得ない (Wilson 及び Nollner (1998) Cell 92, 337-49)。テトラペプチド fMTTV をコードする鋳型を使用して (図3)、二重標識産物の合成を逆相 HPLC で評価した (図6)。放射性ペプチドを化学合成された非放射性標準物との共移動に基づいて同定した。ペプチジル分離範囲の主要放射性ピークは、予期される T/V 比に近い比を有する fMTTV テトラペプチドに相当する (この範囲内の  $^{14}\text{C}$  又は  $^3\text{H}$  放射活性の 80~85%)。2つの副ピークは、予期されるように、 $^3\text{H}$  が全く取り込まれていない休止又は早期終結産物である fMT 及び fMTT に相当する。fMT 及び fMTT は、共に、ペプチジル範囲内の  $^{14}\text{C}$  放射活性の 12% を含有し、これはモルベースでは結合産物 fMTTV、fMTT 及び fMT の 20% に相当する。残りの2つの副ピーク (24 及び 50 分での) は、おそらく fMTV の誘導體 (例えば、メチオニン酸化産物又はホルミル化されていないペプチド)、その他のペプチジル産物及び/又はペプチドでなはなない放射性汚染物に相当する。従って、精製された histag テトラペプチド翻訳系は、大多数ではあるが完全に前進的ではなく、全長テトラペプチド産物の収率は、ペプチド産物の 80% に等しい。この詳細な分析は、添加された EF-P、W 及びレスキューがない場合の連続移動性が以前に報告されたもの (Ganoza 他 (1985) Proc Natl Acad Sci USA 82, 1648-52) よりも非常に高いことを示した。

#### 【0100】

本発明の *in vitro* 翻訳系によって効率的に合成されるペプチドのさらなる例を表2に示す。予期される 7-mer の全長 fMTTTTV ペプチド産物の合成は非常に前進的であった。これは、早期終結産物の圧倒性が、測定された 6.4 の値よりもさらに大きい T/V 比を生じさせたためである。

#### 【0101】

オリゴペプチド合成に及ぼす上流での mRNA 変異の影響。

翻訳開始に関して未解決の問題は、精製系における配列の影響である。従って、スクランピング又は欠失させる効果を、mRNA のスクランブル MV T と欠失 MV T (図3) 及び mRNA の MV T に対して飽和する鋳型濃度 ( $1\ \mu\text{M}$ ) を使用して、開始速度を制限する条件下 (図4参照) で決定した。図7は、Uリッチな配列をスクランピングすると fMTV 合成が低下し、この配列を欠失させると産物合成の初速度が5倍低下することを示している。mRNA 欠失 MV T の濃度がそれよりも高い1.4倍までであると、fMV T の収率を実質的に増加させることができなかつた (データは示さない)。

#### 【0102】

非天然アミノ酸の取り込み及び選択。

本発明のテトラペプチド合成方式は、力学的実験又は選択実験用のリボソームによる非天然アミノ酸の取り込みについての試験を含めて、多くのタイプの開始及び伸長アッセイに直接適用できる。例えば、基質である  $^3\text{H}$ -fmet-tRNA<sub>i</sub><sup>fmet</sup>、 $^{14}\text{C}$ -thr-tRNA<sub>3</sub><sup>thr</sup>、ピオチン標識 lys-tRNA<sup>lys</sup> (プロメガ社) 及び  $^3\text{H}$ -val-tRNA<sub>1</sub><sup>val</sup> を使用して mRNA の mMTKV を翻訳すると (図3)、mRNA のリシンコドンとピオチン標識 lys-tRNA<sup>lys</sup> とに依存する態様で産物を軟質アビジンビーズ (プロメガ社) と共に精製することによって評価されるように、ピオチンと  $^3\text{H}$ -バリンの両方を含有するペプチド産物 (fM-T-bK-V) を産生した (図8)。この実験は、この精製翻訳系が容易に選択できる大きな非天然アミノ酸を取り込むことができることを示している。

#### 【0103】

ピオチン・アビジン相互作用の  $K_d$  ( $10^{-15}\text{M}$ ) が標的に非共有結合的に且つ1価として結合される任意のリガンドで知られている最も低いものの一つであれば、本発明者が合成したピオチン含有疑似ペプチドは、そのコグネイト標的 (アビジン) の1種に対して親和性を有し、これはそれと同様の長さのどんな天然ペプチドのそのコグネイト標的に対する親和性よりも大きいことが予期される。非常に高い親和性のリガンドの同定がライブラリースクリーニング実験に最も望ましいならば、非天然アミノ酸を使用できるリガンドスク

リーニング法の天然アミノ酸しか使用できない方法を超える見込まれる利点は、本発明のピオチニルリシン含有ペプチドについての上記の例によって例示される。

【0104】

また、図8の実験に類似する実験もピオチニルCys-tRNA<sup>Cys</sup>で実施したが、ただし、この場合には、使用したピオチニル-アミノアシルtRNA種を他の手段で精製した。純粋なtRNA<sup>Cys</sup>アイソアクセプター(Subriden RNA)にtRNAを含まない粗製のシンテターゼ抽出物を使用してCysを結合させ、得られたCys-tRNA<sup>Cys</sup>を、公開された方法(大塚他(1997)Nucleic Acid Symp. 37, 125-126)を多少改変してピオチンで化学的に標識し(図9)、ピオチニル-Cys-tRNA<sup>Cys</sup>をゲル精製し(図10)、次いでピオチニル-Cys(bCと略記する)を疑似ペプチドに取り込ませ、そして軟質アビジンに結合させた(図11及び12)。図12の結果は、本発明のin vitro系が特に2種の隣接する大きな非天然アミノ酸を取り込むことができる可能性があることを示唆している。

10

【0105】

重要なことには、図11の非結合tRNA<sup>Cys</sup>による翻訳は、<sup>3</sup>Hの取り込みを全く与えなかったため、この精製系は結合活性がなく且つ外因性の基質に特異的であったことを示している。また、これは、アミノアシルtRNAシンテターゼを含まない純粋な系について予期されるように、非結合tRNA<sup>Thr</sup>又は非結合tRNA<sup>Val</sup>によって「種をまかれる」翻訳のためであることも示している。このシンテターゼはE. coliの最も数の多い蛋白質の1種であり、しかもリボソーム調製物はそれらの精製のために使用できる(Ganoza他(1996)Biochemie 78, 51-61)ので、本発明の純粋な翻訳系がシンテターゼ活性を欠損していたことを確かめるためにさらなる実験を行った。図13は、本発明の「蛋白質ポリメラーゼ」が汚染性のアミノアシルtRNAシンテターゼを欠損していたことを裏付けた。

20

【0106】

ペプチド及び疑似ペプチド産物の精製。

斯界において標準的ないくつかの別法(非変性法又は変性法のいずれか)がペプチジルtRNAから遊離のペプチド又は疑似ペプチドを放出させるために使用できる。この方法の例として、塩基による化学的加水分解(例えば、図4~8)、終結因子(表3)又はペプチジルtRNA加水分解酵素(Karimi他(1998)J. Mol. Biol. 281, 241-252のように精製され且つ使用される)による酵素的触媒作用及びプロマイシン抗生物質又はその誘導體(図21)による求核攻撃が挙げられるが、これらに限定されない。ペプチド又は蛋白質精製のための斯界において標準的な多数の方法のうち任意の方法(非変性法又は変性法)がペプチド及び疑似ペプチド産物の精製のために使用できる。これらの方法の例として、固形物担体を使用する親和性精製(例えば、図8、11及び12の軟質アビジンビーズとピオチンの間の相互作用を使用する)、クロマトグラフィー(例えば、図4及び5の陽イオン交換クロマトグラフィー若しくは図6の逆相HPLCを使用する)又は溶液からの沈殿(例えば、図20のTCAを使用する)が挙げられるが、これらに限定されない。

30

【0107】

高いことが示された本発明の翻訳系の純度(表1、図1、5、11及び13)並びにEF-P、W及びレスキューが前進型翻訳のために明らかに重要でないことにもかかわらず、これらの因子の1種以上がリボソームを汚染しているという議論(発明の背景参照)を除外するためにさらなる実験を行った。リボソームからEF-P、W及びレスキューを除去するための公開された方法(Green他(1985)Biochem. Biophys. Res. Comm. 126, 792-798)は、リボソームを5倍にまでそれぞれの遠心分離の間に高濃度塩(0.5~1M)で一晩洗浄することによってペレット化することである(塩洗浄が多ければ多いほどリボソームの純度は高い)。本発明のリボソームは3回の高濃度塩洗浄で4倍にペレット化されたため、本発明のリボソームはGanozaのものに匹敵する純度を有していよう。それにもかかわらず、追加の(4回目の)高濃度

40

50

塩洗浄を取り入れた異なるバッチのリボソームを調製し (Goanzaの最も高度に洗浄されたりボソームと同一の総数)、さらに、最後の150000gの遠心分離物を、ペレット化物を容易に取り出すために2段階に分けた。第1段階では、遠心分離を最大速度で1分間にわたって行い、得られたペレットを処分した(この物質は260nmでの低い吸光度に基づく主にリボソームではないものであった)。第2段階では、上澄み液からリボソームを長期間にわたる遠心分離によってペレット化させた。これら4×洗浄リボソームの純度は、開始及び翻訳のための翻訳因子への大きな依存に基づいて非常に高いものであることが立証された(表1及び図5のようにアッセイされた)。追加の対照として、本発明の翻訳因子の全て(このものはNiアフィニティークロマトグラフィーによって予め精製された)をさらに追加のゲルろ過クロマトグラフィー段階によって別々に精製した。これら4×洗浄リボソームとNi/ゲル精製因子の組合せは、101個のアミノ酸(図20参照)程度に長いペプチドの合成では活性且つ前進的であり、しかもこの追加の精製段階は、3×洗浄リボソーム及びNi精製因子と比較して本発明に有意に影響を及ぼさないことが分かった。これらの研究は、本発明の翻訳系がEF-P、W及びレスキューに依存しないことを確認するものである。

10

## 【0108】

本発明の他の具体例は、EF-P、W及びレスキューを欠失した本発明の単純化された翻訳系のバージョンであって、その成分の1種以上が実質的に濃度調節され又は完全に除外されたものである。例えば、たいていの場合、PEGを除外すると収率がほぼ25%だけ減少し、効率的な翻訳はIF1なしで生じる(表2、図5)。本発明のその他の標準的な翻訳開始因子及び伸長因子を含めることは、一般的に使用される条件下では効率のために重要であるが、産物合成にとって必須ではない(図5)。実際に、モデル系での効率的な翻訳は、細菌の開始因子(即ち、IF1、IF2及びIF3)をより高濃度の陽イオンで置き換える(例えば、 $Mg^{2+}$ 又はポリアミン、Wagner他(1982)Eur. J. Biochem, 122, 193-197)ならば、この開始因子のいずれかなしに可能である。

20

## 【0109】

本発明のさらに別の具体例は、上に詳述した本発明の単純化された精製翻訳系のバージョンであって、翻訳に関与し又は翻訳を刺激することが知られているその他の精製高分子及び小分子のうち1種以上が添加されたものである。これらは、細胞性の全tRNA又はそれらの画分、細胞性の全アミノアシルtRNA又はそれらの画分、合成型の結合又は非結合tRNA、20種の天然アミノ酸のそれぞれに対するアミノアシルtRNAシンターゼのうち1種以上、Met-tRNA<sub>i</sub><sup>Met</sup>ホルミルトランスフェラーゼ(メチオニルトRNAトランスフェラーゼとも呼ばれる)、N<sup>10</sup>-ホルミルトHFシンターゼ及びTHF誘導体、伸長因子Ts(材料及び方法参照)、終結因子(RF1、RF2、RF3及びRRF又はRF4)、転写と翻訳を対で行うためのDNA鋳型とRNAポリメラーゼ、RNAヘリカーゼ、シャペロン(Hardesty他(1999)Curr Opin Struct Biol 9, 111-4)、ショ糖密度勾配遠心分離法のような異なる手順により精製されたりボソーム(サブユニットへの分離を含めて)、ポリアミンを含めて「ポリミックス緩衝液」の成分(Jelenc及びKurland(1979)PNAS 76, 3174-3178)並びにクレアチンキナーゼ、ミオキナーゼ及び/又はピロホスファターゼ(清水他(2001)Nat. Biotech. 19, 751-755)を有する系のような本発明のピルビン酸キナーゼ系とは異なるエネルギー関連系を包含するが、これらに限定されるべきではない。また、非天然アミノ酸の取り込みを改善させる本発明の組換え非標識IF1、IF2及びIF3(材料及び方法参照)又はEF-Tu誘導体のような天然成分の非標識の又は変異されたタイプを添加し、或いはこれらで置き換えることも純粋な翻訳系では可能である。

30

40

## 【0110】

例えば、図6及び11のような本発明の単純化された精製翻訳系は、細菌の終結因子の添加によって刺激される(表3)。本発明の翻訳にシンターゼ(例えば、Thr及びVa

50

1に対するシンテターゼ)をアミノ酸ThrとVal及びATPと共に添加すると、それぞれの非結合tRNAからそれぞれのアミノアシルtRNAを生成及び再生成できる(示さない)。DNA鋳型、NTP及びRNAポリメラーゼを添加すると、転写と翻訳を対で行うことができた(表2)。また、ポリエチレングリコールのようなその他の分子の添加でも刺激されることが分かった(示さない)。また、細胞から酸・フェノール抽出によって単離された全アミノアシルtRNA(「Varshney他(1991)J. Biol. Chem. 266, 24712-24718」、以下の図16、レーン2も参照)も本発明の精製翻訳に有効である。基質調製のための別法は、脱アシル化された全tRNAを翻訳前に試験管内で結合させることである(Green他(1985)Biochem. Biophys. Res. Comm. 126, 792-798)。後者の方法は、適切なtRNA上に結合した放射性標識された又は非天然のアミノ酸の取り込みを可能にさせるように、ある種のアイソアクセプターtRNAを選択的に消耗させ又は不活性化させる(例えば、アイソアクセプターに特異的なDNAオリゴマーとRNAアーゼHを使用して(神田他(1998)FEBS Lett. 440, 273-276))ことにさらに容易に適合できるという利点を有する。

10

20

30

40

50

#### 【0111】

仮特許出願の6ヶ月後に、精製翻訳系におけるEF-P、W及びレスキューのこの可欠性に関してさらなる証拠が公開された(所望ならば清水他(2001)Nat. Biotech. 19, 751-755を参照されたい)。本発明の好ましい系とは対照的に、これらの本発明系の変形系は、アミノアシルtRNAシンテターゼを含有していた。翻訳は、よく特徴付けられた翻訳因子及びシンテターゼの全ての組換えタイプでもってEF-P、W、W2及びレスキューを添加することなく効率的に再構成されていた。しかしながら、リボソームは、Ganoza及び本発明の方法とは別の方法(より少ない塩洗浄を使用する)によって調製され、そして20種の異なるシンテターゼのうち3種への依存性が不十分であった。これは、EF-P、W、W2及びレスキューによる汚染の可能性を高める。それにもかかわらず、これらの因子の大部分への強い依存性が報告され、いくらかの蛋白質を効率的に合成した。

#### 【0112】

シンテターゼの認識を回避するように変異された化学結合サブレッサーtRNAを使用してバリンを効率的に取り込むためにある種の終止コドンが補充されたが、この文献の主張に反して、非天然アミノ酸の取り込みは試験されなかった(清水他(2001)Nat. Biotech. 19, 751-755)。場合によっては、この本発明の系を変形した系でのサブレッサーtRNAによる非天然アミノ酸の取り込みは、ある種の終結因子との競合をこれらのものを単に除外することによって回避できるという利点を有するかもしれない。しかしながら、粗製抽出物を使用する非天然アミノ酸の取り込みのための既存の系(発明の背景参照)と同様に、この戦略は、おそらく3種の終止コドンのうち1種のみ(UAGコドン)での1蛋白質当たり1個だけのタイプの非天然アミノ酸の取り込みに制限される。これは、添加された20種の異なるアミノアシルtRNAシンテターゼのtRNA結合活性とブルーフリーディング活性によって触媒されるセンスコドンでの天然アミノ酸による競合のため及びリボソームによるリードスルーのため第2の終止コドン(UGA)を使用する試みが失敗したためである(Cload他(1996)Chem. and Biol. 3, 1033-1038)。場合によっては、このような系で複数の異なる非天然アミノ酸の取り込みを可能にさせるために、このサブレッサーアンチコドンは、ある種のセンスコドンを認識させるように変異できたかもしれず(塩基対合によって)、これらのコドンでの競合は、同種の(コグネイト)天然アミノ酸及びシンテターゼを除外することによって回避できたかもしれない。しかしながら、20種のシンテターゼの大部分にとっては、そのコグネイトtRNAアイソアクセプター基質のアンチコドンは最も重要な認識要素であるため、サブレッサーtRNAのアンチコドンが変化すると望ましくないシンテターゼの認識に至り(時として交差認識の予測できない特異性によって)、しかしてブルーフリーディング及び/又は天然アミノ酸の結合に至ると思われる(Giege他(

1998) *Nucleic Acid Res.* 26, 5017-5035)。さらに、シンテターゼの認識を防止することを意図した tRNA の変異は、EF-Tu 及び / 又はリボソームの認識効率をも減少させ、それによって運ばれたアミノ酸の取り込み効率を減少させる可能性がある。これらのシンテターゼ認識の困難性及び 20 種類もの数の異なるシンテターゼ蛋白質を過剰発現させることに関わる困難性のために (Swartz (2001) *Nat. Biotech.* 19, 732-733)、本発明者は、非天然アミノ酸による本発明の試験のためにシンテターゼを含まない翻訳系を選んだ。

#### 【0113】

本発明者は、化学的に結合された tRNA を使用して数種の非天然アミノ酸の取り込みを試験した。第 1 段階は、非天然アミノ酸による化学的な誤アシル化を可能にさせるように、末端 CA ジヌクレオチドを欠失した合成伸長体 tRNA を一般化できる態様で構築することであった。現在の技術は、シンテターゼのいずれかによる結合及びブルーフリーディングを避けるように特別に設計された人工サプレッサー tRNA に依存している。本発明の精製系では、本発明者は、tRNA 塩基修飾活性の予期される欠如について見込まれる効果のみを問題にした。なぜならば、このような修飾は機能にとって重要であり得 (Bjork 他 (1999) *FEBS Lett.* 452, 47-51)、粗製翻訳系は合成 tRNA を修飾し得るからである (Claesson 他 (1990) *FEBS Lett.* 273, 173-6)。新規の試験の雛形として、本発明者は、E. Coli の tRNA<sup>Asn</sup> (図 14、大橋他 (1976) *Nucleic Acids Res.* 3, 3369-3376) 及び tRNA<sup>Ala</sup> (以下に議論する。Pickering 他 (1991) *Nucleic Acids Res.* 19, 5749-5754) を選択した。tRNA<sup>Asn</sup> の 5' 末端配列を T7 RNA ポリメラーゼによる最適な転写のために変異させたが、M1 RNA 又は RNアーゼ P を使用して合成非修飾 tRNA 前駆体を処理するような突然変異誘発とは別の戦略もある (Forster 及び Altman (1990) *Science* 249, 783-786)。また、両者の tRNA のアンチコドンも改変されたコドン認識特性を有するいくつかの変異体を生じさせるように変異させた (本発明の tRNA<sup>Asn</sup> 変異体のうち 3 種を図 15 に示す。tRNA によって認識されるアミノ酸コドンはその括弧内に示されている)。非天然アミノ酸であるアリルグリシン (aG、或いは 2P と略記される) を NVOC 基でアミノ保護し、そして T4 RNA リガーゼを使用して標準的且つ一般的な戦略で (材料及び方法参照。「Thorson 他 (1988) *Methods in Molecular Biology* 77, 43-47」、Steward 及び Chamberlin (1998) *Methods in Molecular Biology* 77, 325-354) tRNA<sup>Asn</sup> (N) に連結して (図 14) ゲル上で予期される移動度でもって移動する種を得た (図 16、レーン 6)。

#### 【0114】

NVOC aG - tRNA<sup>Asn</sup> (N) のアミノ基を紫外線光分解法によって脱保護し、その aG - tRNA<sup>Asn</sup> (N) を mMTNV 鑄型を含有する純粋な翻訳反応に添加した。aG は、下流の <sup>3</sup>H - V の取り込みを可能にするように N コドンにうまく取り込まれたが、mMTNV の対照の翻訳よりも低い収率を示した (図 17)。このより低い収率は、主にこの系を飽和させるのには不十分な aG - tRNA<sup>Asn</sup> (N) を使用するためであることが分かった。なぜならば、aG - tRNA<sup>Asn</sup> の RNA 濃度を 1 μM にまで倍加させる (転写及び連結反応からの aG - tRNA<sup>Asn</sup> (N) の収率が 100% であると見なして翻訳系中のそれぞれの天然アミノアシル tRNA 濃度を倍加させる) と、収率が倍増し且つ約 65% の連続移動率でアナログを取り込んだ (アナログの前の <sup>14</sup>C - T 及びその後の <sup>3</sup>H - V の取り込みに基づく) からである。mMTNNV の翻訳は、隣接したアナログの取り込みを示し、この第 2 のアナログの取り込みについての連続移動率 (%) は、これらのアナログ基質制限条件下では第 1 のアナログと同様であった。一般的な偏在性 CCA 3' 配列で終結することを除いて図 14 の tRNA<sup>Asn</sup> (N) と同一構造を有する非結合 tRNA による mMTNV 及び mMTNNV (BstNI で切断された鑄型のランオフ転写によって生じる) の翻訳は、取り込みを全く与えなかったことに留意されたい。これは、この精

製系がこの変異  $tRNA^{Asn}(N)$  についての結合活性を欠損しており、しかも N コドンでの外因性基質の取り込みに対して特異的であったことを示している。追加の対照として、非結合  $tRNA^{Asn}(N)$  は、mMTV の翻訳を抑制しなかった (図 17)。

#### 【0115】

センスコドンが選択的なアミノ酸の取り込みに一般的な態様で奪い取られたという確証を得たならば、次の段階は、aG 及びその他の非天然アミノ酸の取り込み効率を最適化することなので、O-メチルセリン (mS) 及びアラニンの NVOC アミノ保護型を調製した (材料及び方法参照)。aG、mS 又は Ala で化学的にアミノアシル化された高濃度 (7  $\mu$ M まで) の様々な  $tRNA^{Asn}$  (図 14 及び 15) を使用することによって、Thr- $tRNA^{Thr}$  で得られた取り込み効率の 90% 以上の取り込み効率を得られた (示さない)。

10

#### 【0116】

興味深いことに、非修飾 aG- $tRNA^{Ala}(N)$  であって、公表された合成  $tRNA^{Ala}$  (Pickling 他 (1991) *Nucleic Acids Res.* 19, 5749-5754) と比較してそのアンチコドンに GUU 置換を有するものは、翻訳が全体的に不活性であった。しかしながら、粗製の  $tRNA$  を含まない細胞抽出物を使用して Ala を結合させると、効率的に取り込まれる Ala- $tRNA^{Ala}(N)$  を生じた。この活性の獲得は、aG が  $tRNA^{Asn}(N)$  に良好に作用したので恐らくアミノ酸の相違によるものではなく、粗製抽出物の塩基修飾活性によるものと思われる。この結果は、 $tRNA$  の塩基修飾がある種の  $tRNA$  の翻訳活性にとって極めて重要であり、しかも適切な原型  $tRNA$  の初期選択が試行錯誤を必要とする自明でないプロセスであるという本発明の仮説と一致する。別の問題は、ある種の非修飾アミノアシル  $tRNA$  が標準的な精製手順中に変性によって不活性化されることである (Harrington 他 (1993) *Biochemistry* 32, 7617-7622)。それにもかかわらず、いったん本発明の  $tRNA^{Ala}(N)$  のような安定な活性の非修飾  $tRNA$  が同定されたならば、アンチコドン変異体であってその他のコドンを予測できる態様で認識するものを作るのは容易である (図 15、下部参照)。

20

#### 【0117】

表 4 は、5 個の連続的な非天然アミノ酸 (aG) の 7mer 疑似ペプチド産物内への添加された aG- $tRNA^{Asn}(T)$  に完全に依存する態様での典型的な取り込みを示している。これは、長さの点だけでなく、より重要なことには一般化及び有用性の点でも、2 個の連続的な非天然アミノ酸について公表された記録 (芳坂他 (1999) *JACS* 121, 12194-12195) を向上させている。十分に異なるアミノ酸単量体を与えると、無作為の 7mer 疑似ペプチドライブラリーは、実質的な疑似ペプチド多様性をコードできる。4 種の異なる非修飾  $tRNA^{Asn}$  変異体であってその T コドンについて mS 及びその N、S 及び V コドンについて aG で化学的にアミノアシル化されたものを使用することが図 14 及び 15 に示されており、全長産物がカルボキシ末端  $^3H-E$  の取り込みに基づいて合成された (表 5)。これは、単一産物中に複数且つ異なる非天然アミノ酸が取り込まれた証拠を与える。付加された mS への完全な依存性は、mS が実際に取り込まれ、しかもこの T コドンが  $tRNA^{Asn}(T)$  に対して特異的である (即ち、図 14 及び 15 に示される密接に関連したアンチコドン変異体の  $tRNA^{Asn}(N)$ 、 $tRNA^{Asn}(S)$  及び  $tRNA^{Asn}(V)$  によって誤読されない) ということを示した。

30

40

#### 【0118】

非天然アミノ酸によるセンスコドン置換という本発明の戦略は一般化できるものであり、しかして 61 種のセンスコドン全てに拡張すべきである。自然型 (21 種の天然アミノ酸のみ) によるコドン使用の有意な重複を与えるならば、単一の mRNA を実質的に 21 種以上の異なるアミノ酸に翻訳できるアミノアシル  $tRNA$  ライブラリーを創作することが可能であろう。3 種の終止コドンは本発明の系では必ずしも必要ではないが、これら全てはサブレッサーにより翻訳され得るかもしれないので、見込まれるリードスルーの問題 (Cload 他 (1996) *Chem. and Biol.* 3, 1033-1038) は、

50

リードスルーを引き起こす天然のアミノアシル tRNA を除外することによって克服される。AUG 開始コドンと fMet 開始体アミノ酸を効果的に置換すると、伸長から一連の異なる分子間接触が維持されるが、これは既知の許容できる置換に基づいて生体内及び試験管内でできる(「Picking 他(1991) Nucleic Acids Res. 19, 5749-5754」、「Wu 及び Raj Bhandary (1997) J. Biol. Chem. 272, 1891-1895」)。

【0119】

A. in vitro ライブラリー

リボソーム上で発現するペプチドアナログのコンビナトリアル・ライブラリーからのペプチド又は疑似ペプチドの試験管内での定向進化は、任意の選択の標的分子に結合するリガンド又は薬剤候補物の迅速な進化のために効果的な方法である。粗製抽出物で実施されるリボソームディスプレイ(Mattheakis 及び Dower (1995) PCT WO 95/11922)に関連する本発明の「純粋なリボソームディスプレイ」(図18)の例は、次の段階及びそれらの別法からなる。

【0120】

DNA ライブラリーの構築：

無作為化された又は部分的に無作為化されたコード配列を含む合成オリゴデオキシリボヌクレオチドを化学的に合成し、RNA ポリメラーゼプロモーター(例えば、バクテリオファージ RNA ポリメラーゼ)、翻訳開始配列及び開始コドンを含むオリゴデオキシリボヌクレオチドにハイブリダイズさせ、DNA ポリメラーゼで伸長させ、DNA リガーゼによ

【0121】

段階1(図18)：合成 DNA ライブラリーを RNA ポリメラーゼで mRNA ライブラリーに転写する。

【0122】

段階2：mRNA ライブラリーを、本発明の精製されたリボソーム、翻訳因子及びある種の(全てではない)野生型アミノアシル tRNA とアミノ酸アナログを結合した特別に合成された tRNA との混合物を含有する再構成アミノアシル tRNA プールでペプチドアナログに翻訳する。好ましくは、それぞれの mRNA コドンは、該プール中のユニークな結合 tRNA によって解読されるため、この mRNA の配列は、ユニークなペプチドアナログ配列を規定している。例えば、図19の本発明のスペーサー配列であってそのいくつかはリボソームトンネルを横断するほどに長いものは、良好な連続移動率で合成された(図20)。蛋白質合成を、例えば、終止コドンを欠失している鑄型の末端に到達させることによって、終結因子を除外することによって、供給されるアミノアシル tRNA が全くないコドン(図19及び20の最長産物の場合には Asn コドンのような)に到達させることによって又は抗生物質で立ち往生させることによって停止させる。

【0123】

段階3：無作為ペプチドアナログを含有するリボソームを遠心分離によって単離し、次いで固定化された標的物と共にインキュベートし、そして結合していないリボソームを洗い流す。結合したリボソームを解離させ(例えば EDTA で)、放出された選択 mRNA を精製する。

【0124】

段階4：cDNA 合成及び PCR 増幅。Error-prone PCR を使用して突然変異を導入する。

【0125】

段階1~4の反復：これらの選択及び増幅段階を必要に応じて1回以上繰り返す(以下参照)。

【0126】

10

20

30

40

50

段階 5 (示さない) : 増幅された DNA を、選択され且つ増幅された疑似ペプチドの構造を推定するために、DNA シーケンス法による分析用のプラスミドにクローン化させる。選択されたコンセンサ配列の同定は、反復した選択及び増幅 (段階 1 ~ 4) の十分なラウンドが使用され、しかもリガンドがこの実験によって同定されたという証拠となる。さらに、サイクルを使用してより高い親和性を有するリガンドを生じさせることができる。

【0127】

リボソームの非存在下での選択段階を可能にする方法の改変 :

DNA 鋳型を、タンパク合成中にペプチドアナログに導入されるときに、mRNA 又は mRNA に結合する分子 (例えば、ペプチドタグに対して高い親和性を有する抗体に共有結合するハイブリダイズした相補 DNA プライマー) に対して高い親和性を有するペプチドタグをコードするように修飾する。このとき、このペプチド・mRNA を選択段階の前にリボソームから分離してもよい。これらの方法及びその他の可能な別法は今までに記載されている (Mattheakis 及び Dower, 上記、土井及び柳川 (1999) FEBS Lett. 457, 227-230)。別法として、mRNA は、プロマイシンで 3' 末端標識されていてもよく、そのため「純粋な mRNA ディスプレイ」用のペプチドに直接結合できる (図 21、以下参照)。

10

【0128】

Szostak 他の PCT 公開 WO 00/047775 号及び WO 98/31700 号 (所望ならば参照されたい) は、対象のペプチド及び疑似ペプチドをコードする RNA 分子に共有結合される該ペプチド及び疑似ペプチドの形を生じさせるために、本発明に容易に適合できる方法を記載している。即ち、本発明は、疑似ペプチドであってその自身の mRNA の 3' 末端に共有結合したものの、即ち、RNA・疑似ペプチド融合体を生成させるプロトコルを提供する。

20

【0129】

これは、mRNA 分子をその 3' 末端に結合したペプチドアクセプターと共に合成及び *in vitro* 又は *in situ* 翻訳することによって達成される。一つの好ましいペプチドアクセプターは、成長しつつあるペプチド鎖の C 末端に付加し且つ翻訳を終結させるヌクレオシドアナログのプロマイシンである。一つの好ましい設計では、ある DNA 配列が遺伝子情報の末端とリボソームをオープンリーディングフレームの末端で停止させるためのペプチドアクセプターとの間に含まれ、このペプチドアクセプター (例えば、プロマイシン) がペプチジル tRNA 連結体の加水分解前に初期のペプチド鎖を受容する追加時間を与える (図 21)。

30

【0130】

所望ならば、得られた RNA・疑似ペプチド融合体は、選択及び増幅のラウンドを繰り返してもよい。なぜならば、このコード配列の情報は、逆転写及び増幅 (例えば、PCR 増幅並びに 3SR 又は TSA のような RNA ベースの増幅技術を含めてその他の増幅技術によって) によって回復できるからである。次いで、増幅された核酸を転写させ、修飾させ、*in vitro* 又は *in situ* 翻訳させて次の選択ラウンド用の mRNA・疑似ペプチド融合体を生成させる。この選択及び増幅の複数ラウンドを実施する能力は、非常に希少な分子 (例えば、 $10^5$  種の構成分子からの 1 種の所望の分子) を富化及び単離するのを可能にする。このことは、結果として、新規な又は改良された、特に任意の標的を実質的に認識し又は所望の化学反応を触媒するペプチド及び疑似ペプチドの単離を可能にする。

40

【0131】

従って、ある態様では、本発明は、(a) 候補の RNA 分子集団であって、これらのいずれかが翻訳開始配列及びある候補の蛋白質コード配列に作動可能に連結した開始コドンを含み、これらのいずれかが該候補蛋白質コード配列の 3' 末端でペプチドアクセプターに作動可能に連結したものを与え、(b) 天然型及び / 又は非天然型アミノ酸の存在下に該候補蛋白質コード配列を *in vitro* 又は *in situ* 翻訳させて候補の RNA・疑似ペプチド融合体の集団を産生させ、(c) 所望の RNA・疑似ペプチド融合体を選択し、それによって所望の疑似ペプチドを選択する段階を含む、所望の蛋白質又は疑似ペプチドの選択

50

方法に関する。

【0132】

上記方法の好ましい具体例では、候補RNA分子の集団は、少なくとも $10^2$ 種、好ましくは、少なくとも $10^5$ 種、より好ましくは $10^{10}$ 種又は $10^{15}$ 程度の異なるRNA分子を含む。重要なことには、in situ 翻訳反応は、好ましくは、粗製翻訳系ではなく再構成された精製混合物中で実施される。選択段階は、疑似ペプチドを固定化された結合の相手に結合させること又は疑似ペプチドの機能活性についてアッセイすることを含む。

【0133】

別の関連する態様では、本発明は、ここに記載される選択方法のいずれかを実施するためのキットに関する。

【0134】

最後の態様では、本発明は、固定化された本発明の疑似ペプチドの配置を含むマイクロチップに関する。

【0135】

B. 標的分子

標的分子は、実質的には、本発明のペプチド又は疑似ペプチドとの相互作用が有用であり得る任意の分子であることができる。ある種の具体例では、標的分子は、核酸(DNA又はRNA)、蛋白質、脂質、炭水化物などのような生体高分子である。

【0136】

ポリペプチドスクリーニング用標的物を選択する際に考慮され得る因子には、溶解度、ペプチド鎖の長さ、翻訳後修飾の必要条件若しくはコファクターの付加及び/又はこの標的がベースとしている蛋白質のモノマー又はオリゴマーの性質が含まれる。一般に、ペプチドの標的は、可溶性で、部分的に精製され又は純粋であり、しかも斯界に標準的な方法によって固体担体上に固定化されていることが望ましいであろう。

【0137】

従って、本発明は、SH2ドメイン、SH3ドメイン、アンキリン様リピート、WD40モチーフ、Kunitz型阻害剤のドメイン、EFG様ドメインのような増殖因子様ドメイン、クリングルドメイン、フィブロネクチンフィンガー様ドメイン、ヘパリン結合ドメイン、デスドメイン、TRAFドメイン、プレクストリン相同(PH)ドメイン、ITAM、キナーゼドメインのような触媒ドメイン、ホスファターゼドメイン、ホスホリパーゼドメイン、グアニンヌクレオチド交換因子(GEF)ドメイン及びヒドロラーゼドメイン(例えばプロテアーゼドメイン)又はDNA結合ドメイン、例えばロイシンジッパー、ジンクフィンガー及びヘリックス・ループ・ヘリックスモチーフのようなドメイン構造に相当する(例えば、これらを含む)スクリーニング用標的物を意図する。

【0138】

対象の蛋白質が膜貫通蛋白質である場合には、スクリーニング用標的物は、可溶性の細胞外又は細胞質ドメインから得ることができる。例示のために、スクリーニング用標的物は、グアニリルシクラーゼ、シトキン受容体、チロシンキナーゼ受容体又はセリン/トレオニンキナーゼ受容体の細胞外ドメインに相当し得る。その他の具体例では、スクリーニング用標的物は、リガンド結合活性を保持しているG蛋白質共役型受容体(GCR)の可溶性部分に相当し得る。例えば、上記のように、GCRの膜貫通部分の間の細胞外ループのうちのある種のものは、溶液状でなくなったときでさえもリガンド結合活性を保持することが示された。さらに別の具体例では、スクリーニング用標的物は、リボソームのような脂質二重層又はその他の小胞中で再構成でき(例えば、Kalva Kolanu他(1990) *Biotechniques* 11:248並びにHuangの米国特許第4957735号及び同第4708933号を参照)、この脂質/蛋白質結合体をスクリーニング用標的物として使用できる。

【0139】

単なる例示の目的のために、以下の蛋白質標的物を対象の方法に使用するために記載する。

10

20

30

40

50

## 【0140】

一具体例では、見込まれる治療用標的物は、neu受容体ファミリーからの受容体である。米国の女性において、乳癌は最も一般的な癌であり且つ癌死亡率が肺ガンに次いで単独で第2位である。主な乳癌の標的は、185kDの膜貫通ホスホ糖タンパク質のチロシンキナーゼであるneu/erbB-2/HER-2である(Shih他、(1981) Nature 290:261)。neu癌遺伝子の増幅又は過剰発現は、乳房及び卵巣の約30%で発生し、これは一次治療のまずい対応に相関することが分かっている(Slamon他(1987) Science 235:177及びHayes他(1993) Annals of Oncology 4:807)。NIH 3T3細胞にneu癌遺伝子をトランスフェクトすると形質転換を生じ(Shih他、上記)、活性化されたneu癌遺伝子をマウスに導入すると全乳房上皮組織の形質転換(Muller他、(1988) Cell 54:105)を生じ又は乳房腫瘍が確率的に出現する(Bouchard他(1989) Cell 57:931)。

## 【0141】

乳癌がneu又はヘレグリンファミリーの一員のようなneuのリガンドに結合する分子によって抑制できるという証拠が蓄積している(Holmes他、1992, Nature 256:1205)。neuの細胞外ドメインに結合するMAb及びそれらの放射性標識結合体は、neu陰性細胞クローンの選択なしに培養液中及びヌードマウス内で乳癌細胞の増殖を遅延させる(DeSantes他、(1992) Cancer Res. 52:1916及びDrebin他(1986) PNAS 83:9129)。高度乳癌治療のための診療所で使用される組換えヒト化MAbであるヘルセプチン(Colomer他(2001) Cancer Investigation 19, 49-56)のような当該MAb及びそれらの誘導体は、neuの情報伝達経路を変更し且ついくつかの異なる方法で腫瘍の増殖に影響を及ぼし得る。これらのものは、(i)neuを過剰に刺激し、それによって分化を生じさせ(Bacus他(1992) Cancer Res. 52:2580)、(ii)neuのホモ又はヘテロ二量体化を妨げ、それによってneuを不活性化させ(Caraway他(1994) Cell 78:5)、(iii)neuの細胞性インターナリゼーション及びダウンレギュレーションを生じさせ(Tagliabue他(1991) Int. J. Cancer 47:93)、(iv)細胞表面又は細胞質に共役細胞傷害性放射性ヌクレオチド又は毒素を送達し、或いは(v)リガンドがneuに結合しないようにする。

## 【0142】

本発明によって誘導され得る疑似ペプチドは、例えばneuのそのリガンド又はその他の蛋白質との結合を競争的に開裂させ又はneuに関わる酵素活性のアロステリックな活性化を阻害することによってneuの生物学的機能を抑制するのに有用であり得る。別法として、これらのものは、neuの過剰刺激又はダウンレギュレーションを生じさせるアゴニストとして有用であり得る。

## 【0143】

さらに別の見込まれる標的は、インターロイキン8(IL-8)である。IL-8は、好中球の化学誘引物質及び活性化物質であり且つ様々な急性及び慢性の炎症性疾患に関与している(Murphy(1994) Annu. Rev. Immunol. 12:593-633)。ヒトIL-8は、腫瘍壊死因子のような因子、インターロイキン1及びリポ多糖による誘導に基づいて単球、繊維芽細胞、角化細胞及び内皮細胞によって産生される72アミノ酸残基のポリペプチドである(Murphy、上記)。IL-8のある種のアナログは、試験管内で好中球の活性化(走化性、エキソサイトーシス及び呼吸バースト)を阻害することによってIL-8アンタゴニストとして作用するため、IL-8抑制剤が炎症性疾患に対する治療上の可能性を有し得ることを示唆する(Moser他(1993) J. Biol. Chem. 268:7125-7128)。

## 【0144】

単量体のIL-8ペプチドは、試験管内で20MのK<sub>d</sub>でもって二量体を形成する(Pa 50

olini他(1994) *J Immunology* 153:2704及びBurrows他(1994) *Biochemistry* 33:12741-12745)ため、この単量体及び/又は二量体は生体内で活性であることが可能である。二量体化できない変異体は、試験管内での機能アッセイでは活性である(Rajaratnam他(1994) *Science* 264:90)。興味深いことに、IL-8二量体の三次元構造のNMR及びX線測定(Cloure他(1990) *Biochemistry* 29:1689-1696及びBaldwin他(1991) *PNAS* 88:502)により、このものがMHCのクラスI及びII蛋白質のペプチド結合の溝に類似していることが明らかになった(Bjorkman他(1987) *Nature* 329:506)ため、IL-8二量体は、MHC分子と同様の態様で *in vitro* 選択ペプチド又は疑似ペプチド配列に結合できる可能性があると考えられる。従って、IL-8単量体に加えて、その二量体は、魅力的な標的である。

10

## 【0145】

IL-8経路の場合には、IL-8とは別の標的は、IL-8受容体の機能性断片である。それぞれ39個及び44個のアミノ酸のヒトIL-8タイプ1受容体及びウサギIL-8タイプ1受容体のさらに小さい断片は、IL-8結合アッセイでは機能的である(Gayle他(1993) *J Biol Chem* 268:7283-7289)。従って、最も大きな受容体のファミリーの一部である7種の膜貫通受容体が見込まれる標的である。

## 【0146】

細胞性の癌原遺伝子であるc-mycは、細胞増殖及び形質転換に関与するか、プログラム細胞死(アポトーシス)の誘導にも関わる。c-Myc蛋白質は、カルボキシル末端塩基性領域/ヘリックス・ループ・ヘリックス(HLH)/ロイシンジッパー(LZ)ドメインを有する転写活性化物質である。このものは、HLH/LZ蛋白質のMaxとヘテロ二量体を形成し、DNAのEボックス要素に結合した後に遺伝子発現をトランスに活性化する。Max蛋白質は、現在までに分析された多くのその生物学的機能にとって必須のc-Mycのパートナーである。例えば、Mycは、DNAに結合し且つその癌遺伝子活性を果たすためにはMaxと共にヘテロ二量体化しなければならない。

20

## 【0147】

本発明によれば、主題の方法は、MycとMaxのようなその他の蛋白質との間の複合体の形成を阻害できる及び/又はMyc複合体が遺伝子内のmyc応答配列に結合するのを阻害できるペプチド及び疑似ペプチドを誘導するために使用できる。Myc・Max及びMax・Max二量体の全合成は、Canne他(1995) *J Am Chem Soc* 117:2998-3007に記載されている。

30

## 【0148】

TGFの全合成は、Woo他(1989) *Protein Eng* 3:29-37に記載されており、これは別の見込まれる標的分子を与えるものである。

## 【0149】

主題の方法に使用するために誘導できるさらに別の標的は、細胞接着、組織器質化及び創傷治癒に関わる糖蛋白質のフィブロネクチンである。フィブロネクチン分子の全合成は、例えば、Williams他(1994) *J Am Chem Soc* 116:10797-10798に記載されている。

40

## 【0150】

これまでに、ヒト免疫不全ウイルス1型(HIV-1)の主要gag蛋白質であるp24の発現は、HIV-1に感染した細胞の表面上で持続することが示された(西野他(1992) *Vaccine* 10:677-683)。p24のC末端の100アミノ酸断片の全合成は、Mascagni他(1990) *Tetrahedron Lett* 31:4637-4640)によって説明され、p24蛋白質のその部分を使用してスクリーニング用標的物を生成することができる。

## 【0151】

50

同様に、HIVプロテアーゼは、すべて化学合成手段によって合成され（kent他PCT公開WO93/20098）且つ触媒活性の阻害剤並びにこのプロテアーゼに関わる蛋白質・蛋白質相互作用の阻害剤を開発するためのユニークな標的となる。

#### 【0152】

上に例示された標的物の多くは化学合成によって調製できるが、これは決して必要条件又はさらに優先的に選択されるものではない。本発明の利点は、生物学的供給源からの精製、生合成又は組換えDNAベースの発現のような任意の他の方法を使用して部分的に精製された又は純粋な標的物を調製できることである。

#### 【0153】

##### 材料及び方法

hisタグ及び非標識E. coli IF1、IF2及びIF3蛋白質の過剰発現用プラスミドの構築。

N末端メチオニンのすぐ後に6個のヒスチジン挿入物をそれぞれ含むE. coliの開始因子コード配列を公表されたプラスミドからPCRによって合成し、pET24a（Novagen社）由来のベクターにサブクローニングした。E. coli優先コドン（infAコドンの代わりに）の人工配列によってコードされるネイティブIF1配列を含むプラスミドpXR201を、R. Spurio及びC. Guallerzi両氏から快く提供してもらい（Calogero他（1987）Mol Gen Genet 208, 63-9）、所定のpAF1Hにサブクローニングした。infBによってコードされるネイティブIF2配列を含むプラスミドpSL4を、S. Laalami及びM. Grunberg-manago両氏から快く提供してもらい（Laalami他（1991）J Mol Biol 220, 335-49）、所定のpAF2Hにサブクローニングした。infCによってコードされるネイティブIF3配列を含むプラスミドpDD1を、N. brot及びI. Schwartz両氏から快く提供してもらい（De Bellis及びSchwartz（1990）Nucleic Acids Res 18, 1311）、所定のpAF3Hにサブクローニングした。この3種のサブクローンの配列は、制限酵素消化と配列分析の組み合わせによって特徴付けられ、第2アミノ酸の前のTATACA/TATG(CAC)<sub>6</sub>から始まる（連結部位に下線を施した）。最後のアミノ酸の後にTAAG/AATTCGAGCTCCGTCGA/42bp欠失/AGATCC配列が続き、この配列の残余は、pET24aからのものである。また、IF1、IF2及びIF3の非標識バージョンをクローン化し且つ過剰発現させるために類似の方法も使用した。

#### 【0154】

E. coli翻訳因子蛋白質の過剰発現及び精製。

TufAの最初の2つのコドン間に挿入されたhis<sub>6</sub>配列を含むhisタグE. coli EF-TuをコードするプラスミドpHTA7（E. coli BL21(DE3)内の）及びTsfの最初の2つのコドン間に挿入されたhis<sub>6</sub>配列を含むプラスミドpHTSは、Y. W. Hwang及びD. Miller両氏が提供してくれた（Hwang他（1997）Arch Biochem Biophys 348, 157-62参照）。his<sub>6</sub>を含めて約30個のアミノ酸のN末端伸長を含む、hisタグE. coli EF-GをコードするプラスミドpRSET/EF-G(His)(pLysSプラスミドと共にE. coli BL21(DE3)細胞内の）は、A. Savelbergh及びW. Wintermeyer両氏が提供してくれた（Semenkov他（1996）Proc Natl Acad Sci USA 93, 12183-8）。本発明の3種のhisタグ開始因子のサブクローン（E. coli BL21(DE3)pLysSの：Novagen社）及び提供された3種のクローンの発現をIPTGによって誘導した。これらの因子全てを主として可溶性細胞画分中で発現させ、標準的なプロトコルを使用してNiNTAアガロースカラム（Qiagen社）から段階溶出によって精製した（ただし、10μMのGDPをEF-TuHのための最後の透析段階まで含めた）。全ての因子を緩衝液A（100mMのTris-HCl pH 7.4, 1mMのMgCl<sub>2</sub>, 1mMの

10

20

30

40

50

DTT)に透析した。沈殿したIF3Hを5Mの尿素に再溶解させることによって回収し(Hershey他(1977)Arch Biochem Biophys 182, 626-38)、希釈し、次いで100mMのNH<sub>4</sub>Clを含有する緩衝液Aに透析した。ネイティブIF2の過剰発現及び精製中に他の者によって観察された広範囲にわたる蛋白質分解性の分解(Mortensen他(1991)Biochimie 73, 983-9)とは異なり、組換えIF2Hは、蛋白質分解に不安定ではなかった。純粋なE. coli終結因子(RF)のRF1、RF2、RF3及びRRFを、記載されているように(Yu他(1998)J. Mol. Biol. 284, 579-590)調製した。全ての因子を-80で保存した。これらのものを、EF-TuHを除いて活性を損失させることなく何度も解凍し、解凍後には4で保存し且つ数週間以内に使用した。

10

## 【0155】

E. coliリボソームの精製(Kung他(1974)Arch Biochem Biophys 162, 578-84)。  
対数増殖中期にまで増殖したSOLR(商標)細胞(Stratagene社)を緩衝液B(60mMのKOAc, 14mMのMg(OAc)<sub>2</sub>, 10mMのTris-HOAc, 1mMのDTT, pH7.9)に再懸濁し、超音波処理し、10000gで遠心分離した。上澄液を30000gで遠心分離し、次いで得られた上澄液を150000gで遠心分離した。このリボソームペレットを1MのNH<sub>4</sub>Clを含有する緩衝液B中で4で一晩攪拌し、次いで150000gで再ペレット化することによって洗浄した。この洗浄を2回以上繰り返して3×洗浄リボソームを与え、このものを緩衝液C(10mMのMg(OAc)<sub>2</sub>, 1mMのTris-HCl pH7.4, 1mMのDTT)に再懸濁し、-80で保存した。

20

## 【0156】

mRNAの合成。

mRNAをオリゴデオキシリボヌクレオチド(Research Genetics社)から転写し(Milligan及びUhlenbeck(1989)Methods in Enzymology 180, 51-62)、ゲル電気泳動により精製した(Foster及びSymons(1987)Cell 49, 211-20)。鑄型が比較的長いために、最適な転写を可能にするにはブロックされたオリゴヌクレオチドの合成の後に長期の脱保護時間が必要であった(12時間)。

30

## 【0157】

アミノアシルtRNAの調製。

純粋なE. coli tRNAアイソアクセプターは、SubridenのRNAからのものであった。それぞれのアイソアクセプターは、E. coliの全tRNA(Plenum社)から3つのカラムクロマトグラフィー段階を使用してメーカーにより製造された。第1の分画はBDセルロースを使用し、第2はpH7でのDEAEセファデックスを使用し、第3はpH5でのDEAEセファデックス又はセファロースを使用した。天然のアミノアシルtRNAをこれらのアイソアクセプターから次のように調製した。高特異的活性<sup>3</sup>H-fmet-tRNA<sub>i</sub><sup>fmet</sup>(表1において24000d.p.m./pmol)及び低特異的活性<sup>3</sup>H-又は<sup>35</sup>S-fmet-tRNA<sub>i</sub><sup>fmet</sup>(数百d.p.m./pmol、全ての他の研究のために使用される)を、MetRS、met-tRNA<sub>i</sub><sup>fmet</sup>ホルミルトランスフェラーゼ及びN<sup>5,10</sup>-メチルTHF(N<sup>10</sup>-ホルミルTHFシンテターゼを使用して合成された)を使用して記載されるように(Robakis他(1981)Proc Natl Acad Sci USA 78, 4261-4)調製した。<sup>14</sup>C-thr-tRNA<sub>3</sub><sup>thr</sup>(510d.p.m./pmol)及び<sup>3</sup>H-val-tRNA<sub>1</sub><sup>val</sup>(図5~7において21000d.p.m./pmol、図4及び8において28000d.p.m./pmol)を、記載されるように(Robakis他(1981)Proc Natl Acad Sci USA 78, 4261-4)、ThrRS、ValRS又は、低特異的活性val-tRNAについては、150000gの上澄液から0.3MのKClでDEAEセファロースから段階溶出することによって部分的に精製され

40

50

た全 *E. coli* シンテターゼ (Kung 他 (1975) *J Biol Chem* 250, 1556-62 参照) の tRNA を含まない調製物でもってアミノアシル化した。

【0158】

ピオチン標識 lys-tRNA<sup>lys</sup> (Transcend (商標) tRNA、図22) をプロメガ社から購入した。この材料は、このメーカーによって、*E. coli* の全 tRNA から全シンテターゼの粗製調製物を使用してリシンを結合させ、lys-tRNA<sup>lys</sup> をイオン交換クロマトグラフィーにより富化させ、そして非電荷の13個の炭素の長さのスペーサーを介してピオチンに化学的に結合させることによって調製された。

【0159】

NVOCアミノ保護アミノアシル tRNA を斯界に標準的な方法によって調製し、保存し、使用した (「Thorson 他 (1988) *Methods in Molecular Biology* 77, 43-73」、*Steward* 及び *Chamberlin* (1998) *Methods in Molecular Biology* 77, 325-354」に詳述されている)。手短に言えば、pdCpA をアデノシンとデオキシチジン誘導体から化学的に合成した。種々の非天然アミノ酸を標準的な業者 (例えば、Sigma 社、Aldrich 社、Fluka 社、Bachem 社及び Novabiochem 社) から保護されていない形で購入した。aG、mS 及び Ala のような本発明者が最も一般的に使用した NVOCアミノ酸は、化学的な保護を必要とする反応性のアミノ酸側鎖を欠失していたため、これを保護されていないアミノ酸から NVOC-Cl によるアミノ保護によって直接合成した。標準法によって合成された追加の側鎖保護を有するものを含めて、適切に保護された全てのアミノ酸をシアノメチルエステル合成によって活性化させ、次いで pdCpA に結合させた。得られた NVOC 保護アミノアシル化 pdCpA 化合物 (及びそれらの RNA との結合体) は、水溶液中に -80 の暗所でほぼ 5 の pH でもって保存したときに安定であった。NVOC 保護アミノアシル化 pdCpA 化合物を、末端 CA ジヌクレオチドを欠失した様々な tRNA 誘導体に T4 RNA リガーゼで連結し、そして結合体を精製し、保存した (上記参照)。得られた NVOC 保護アミノアシル tRNA を翻訳に使用する前に直ちに UV 照射によって脱保護した。

【0160】

his タグ *E. coli* 翻訳因子蛋白質のアッセイ。

開始因子アッセイ (Kung 他 (1974) *Arch Biochem Biophys* 162, 578-84) は、受注合成された ApUpG RNA 鋳型 (TriLink Biotechnologies 社) を使用した。0.95 μM の IF1H、0.15 μM の IF2H、0.78 μM の IF3H、0.029 A<sub>260</sub> / μl の 3× 洗浄リボソーム (33 nM が翻訳において活性であると見積もられる。以下参照)、0.29 μM の <sup>3</sup>H-fmet-tRNA<sub>i</sub><sup>fmet</sup>、150 μM の AUG 及び 0.4 mM の GTP を含んでなる (50 mM の Tris-HCl pH 7.4, 100 mM の NH<sub>4</sub>Cl, 5 mM の Mg(OAc)<sub>2</sub>, 2 mM の DTT) 混合溶液を 37 で 10 分間インキュベートした。希釈後、この混合液を素早くニトロセルロースで濾過して開始複合体結合 fmet-tRNA<sub>i</sub><sup>fmet</sup> を非結合種から分離した。鋳型を欠いた反応を非特異的な結合のための対照として使用した (表1に関して最大の d.p.m. の 29%、<sup>3</sup>H-fmet-tRNA<sub>i</sub><sup>fmet</sup> の 70% を結合したより高 (飽和) 濃度を使用して最大 d.p.m. の 9%)。EF-Ts に対する EF-Tu の高い親和性のため、EF-Ts の活性アッセイ (Weissbach 及び Pestka (1977) *Molecular Mechanisms of Protein Biosynthesis*, Academic Press, ニューヨーク, NY) を EF-TuH と共に実施したところ、約 2% の共精製レベルを示した。EF-TuH 活性を GDP の結合によって測定し (Weissbach 及び Pestka (1977) *Molecular Mechanisms of Protein Biosynthesis*, Academic Press, ニューヨーク, NY)、EF-GH 活性をリボソーム依存型 GTP アーゼアッセイによって測定した (Weissbach 及び Pestka (1977) *Molecular Mechanisms of Prote*

in Biosynthesis, Academic Press, ニューヨーク, NY)。

【0161】

翻訳。

翻訳混合物の成分を公表された実験法(「Robakis他(1981) Proc Natl Acad Sci USA 78, 4261-4」、 「Cenatiempo他(1982) Arch Biochem Biophys 218, 572-8」)から改変した。5xプレミックス緩衝液を、180mMのTris-HOAc(pH7.5)、50mMの3,3-ジメチルグルタル酸ナトリウム(pH6.0)、180mMのNH<sub>4</sub>OAc、10mMのDTT、140mMのホスホエノールピルビン酸カリウム(pH6.6)、195mMのKOAc及び4mMのスペルミジン・3HClを含有する溶液からこれをNaOHでpH6.8に調整することによって調製した。代表例として、MTTV mRNA 鋳型(図6、30µlの総容量)を翻訳する際の混合物は、1xプレミックス緩衝液、9.5mMのMg(OAc)<sub>2</sub>、1mMのGTP、14ng/µlのピルベートキナーゼ、4.3%のPEG8000、0.95µMのIF1H、0.15µMのIF2H、0.78µMのIF3H、3.1µMのEF-TuH、0.88µMのEF-GH、0.029A<sub>260</sub>/µlの3x洗浄リボソーム(33nMが翻訳の際に活性であると見積られる。図4参照)、0.29µMの<sup>3</sup>H-fmet-tRNA<sub>i</sub><sup>fmet</sup>、0.58µMの<sup>14</sup>C-thr-tRNA<sub>3</sub><sup>thr</sup>、0.29µMの<sup>3</sup>H-val-tRNA<sub>1</sub><sup>val</sup>及び1µMのmRNAを包含した。この混合物を37で50分間インキュベートしたが、いくつかの翻訳は1分程であった(図4)。また、いくつかの翻訳のための設定は、開始混合物(GTP、IF1F、IF2H、IF3H、リボソーム、<sup>3</sup>H-fmet-tRNA<sub>i</sub><sup>fmet</sup>及びmRNA)及び伸長因子混合物(GTP、EF-TuH、EF-GH、<sup>14</sup>C-thr-tRNA<sub>3</sub><sup>thr</sup>及び<sup>3</sup>H-val-tRNA<sub>1</sub><sup>val</sup>(図4))の37で10分間のプレインキュベーションを包含した。

【0162】

HPLC分析(図6)のために、ペプチド及びアミノ酸を1MのNaOH(6µl)の添加及び37で20分間のインキュベーションによりtRNAから放出させた。次いで、非標識マーカーのペプチド(Research Genetics社)を添加し(H<sub>2</sub>O中に10µg/µlの合わせたペプチド濃度で18µl)、この溶液を氷酢酸(5µl)で酸性化し、微量遠心分離し、次いで上澄液をMicrocon10限外濾過装置(10kDカットオフ、Amicon社)によって微量遠心分離した。濾過液の一部分(20µl)をC18逆相カラム(Vydac社)上に添加し、0.1%のTFAを含有する0~31.5%のH<sub>2</sub>O/MeCN勾配でもって1ml/分で溶出させ、229nmの吸光度で検出し、二重標識d.p.m.プログラム(Packard社)を使用して42個の滴下画分をシンチレーション計数した。

【0163】

fMTV及びfMVT合成の分析のために、ホルミル化されたペプチド及びホルミルメチオニンを塩基加水分解、酸性化及び小型カラム上での陽イオン交換クロマトグラフィーによってホルミル化されていないアミノ酸から分離した(Peacock他(1984) Proc Natl Acad Sci USA 81, 6009-13)。

【0164】

対照の翻訳。

追加の翻訳を実施して翻訳因子以外の成分への依存性を評価した。<sup>3</sup>H-fmet-tRNA<sub>i</sub><sup>fmet</sup>又は<sup>14</sup>C-thr-tRNA<sub>3</sub><sup>thr</sup>の除外は、<sup>3</sup>H-valの産物への取り込みに基づいてmRNAMTVからのfMTV合成を停止させる。<sup>14</sup>C-thr-tRNA<sub>3</sub><sup>thr</sup>を除外した実験において、非結合tRNA<sub>3</sub><sup>thr</sup>を添加すると精製系で予期されるような測定可能なfMTV合成を再構成しなことから、翻訳条件下ではThrRS活性がないことを示している(さらなる対照として、thr-tRNA<sub>3</sub><sup>thr</sup>の添加はfMTV合成を阻害しなかった)。<sup>3</sup>H-val-tRNA<sub>1</sub><sup>val</sup>の除外は、<sup>14</sup>C-thr-tRNA<sub>3</sub><sup>thr</sup>の産

物への取り込みに基づいて mRNA の MVT からの fMTV 合成 (図 4) を停止させる。mRNA MVT 及び mRNA mMV (図 3) 産物中の T/V 比は、予期されるように、それぞれ約 1.0 及び 0 である。リボソームを除外すると、最も低いバックグラウンド放射能測定値が得られ、mRNA を欠失した翻訳は、ホルミル化ペプチド産物を時間経過と共に蓄積しない。ウサギ筋肉のピルベートキナーゼ (シグマ社製のこの蛋白質 (Jelenic 及び Kurland (1979) Proc Natl Acad Sci USA 76, 3174-8) は、SDS-PAGE で単一の主要バンドとして移動する) を除外すると、産物の収率が 50% だけ減少する。Mg<sup>2+</sup> の標準濃度 (9.5 mM) が最適であるが、翻訳はそれよりも高い及び低い濃度で効率的に起こり得る。

【0165】

本研究の一部は、NIH 承認 K08-C A80833 によって裏付けられた。アメリカ合衆国政府は本発明に関して一定の権利を有する。

【0166】

【表 1】

表 1  
開始複合体形成についての因子依存性

| 除外した開始因子 | 開始複合体 <sup>a</sup>  |                                |   |
|----------|---------------------|--------------------------------|---|
|          | His タグ <sup>b</sup> | ネイティブ<br>(Kung 他) <sup>c</sup> | ネイティブ (Dubnoff<br>及び Maitra) <sup>d</sup> |
| なし       | 100                 | 100                            | 100                                       |
| IF1      | 48                  | 61                             | 36  |
| IF2      | 0                   | 4                              | 7   |
| IF3      | 7                   | 17                             | 48  |

<sup>a</sup> GTP の存在下での <sup>3</sup>H 標識 fmet-tRNA<sub>i</sub><sup>fmet</sup> のリボソームへの最大の結合 (%)。

<sup>b</sup> 精製された his タグ開始因子、ApUpG 鑄型及び 5mM の Mg<sup>2+</sup> (材料及び方法参照) で実施した。特に開始複合体に結合した fmet-tRNA<sub>i</sub><sup>fmet</sup> の最大濃度は 19nM であった。

<sup>c</sup> 精製されたネイティブの開始因子、ポリ(U:G) (3:1) 鑄型及び 10mM の Mg<sup>2+</sup> で実施した。

<sup>d</sup> 精製されたネイティブの開始因子、ApUpG 鑄型及び 5mM の Mg<sup>2+</sup> で実施した。

【0167】

【表 2】

表 2

7mer ペプチド合成の連続移動性の測定。翻訳因子である IF2、IF3、EF-Tu 及び EF-G を含む 5.1 μl の組合わせ転写/翻訳物からのペプチド産物の d. p. m. は pmol に変換している。

| 予期される<br>Thr/Val の<br>mRNA 鑄型によってコー<br>ドされるペプチド | <sup>14</sup> C-Thr<br>(pmol) | <sup>3</sup> H-Val<br>(pmol) | 測定された<br><sup>14</sup> C-Thr/ <sup>3</sup> H-Val |   |
|---|-------------------------------|------------------------------|--|---|
| fMet Thr Thr Thr Thr Thr Val                    | 0.45                          | 0.07                         | 6.4  | 5 |
| fMet Thr Val                                    | 0.06                          | 0.06                         | 1.0  | 1 |
| fMet Val  | 0.008                         | 0.07                         | 0.11   | 0 |

【0168】

【表 3】

表 3

純粋なE. coli終結因子(RF)であるRF1、RF2、RF3及びRRFによる純粋な翻訳系の刺激(Yu他(1998)J. Mol. Biol. 284, 579-590;材料及び方法参照)。mRNAは、終止コドンに欠失しているか(mMTTV)又は第4のアミノ酸コドンの後に直接終止コドンにコードしているか(mMTCV<sub>UAG</sub>)のいずれかであった。産物のdpm:ペプチド産物に取り込まれたトリチウム化バリンの1分当たりの放射性崩壊値。全ての翻訳は、翻訳因子IF1、IF2、IF3、EF-Tu、EF-G及びEF-Tsを包含した。

| mRNA                 | 供給されたアミノアシルtRNA | RF | 産物のdpm | ペプチド産物 |
|----------------------|-----------------|----|--------|--------|
| mMTTV                | fM、T、V          | -  | 4597   | fMTTV  |
| mMTTV                | fM、T、V          | +  | 7187   | fMTTV  |
| mMTCV <sub>UAG</sub> | fM、T、C、V        | -  | 2676   | fMTCV  |
| mMTCV <sub>UAG</sub> | fM、T、C、V        | +  | 3866   | fMTCV  |

10

## 【 0 1 6 9 】

## 【 表 4 】

表4

連続的な5個の非天然アミノ酸の疑似ペプチド産物への取り込み。アリルグリシンはaGと略記する。その構造は図14に示している。aGを結合した人工tRNA<sup>Asn</sup>は、Tコドンを認識するためaG-tRNA(T)と命名する(図15)。全d. p. m.:小型カラムから溶出した全<sup>3</sup>Hd. p. m.(材料及び方法参照)。全ての翻訳は、翻訳因子IF1、IF2、IF3、EF-Tu、EF-G及びEF-Tsを包含した。

20

| mRNA      | 供給されたアミノアシルtRNA | 全dpm | 疑似ペプチド産物            |
|-----------|-----------------|------|---------------------|
| mMTTTTTTV | fM、aG-tRNA(T)、V | 2721 | fM-aG-aG-aG-aG-aG-V |
| mMTTTTTTV | fM、T、V          | 7685 | fMTTTTTTV           |
| mMTTTTTTV | fM、V            | 1116 | なし(バックグラウンドdpm)     |
| -         | fM、V            | 1113 | なし(バックグラウンドdpm)     |

## 【 0 1 7 0 】

## 【 表 5 】

表5

2種の異なるタイプの非天然アミノ酸の疑似ペプチド産物への取り込み。O-メチルセリンはmSと略記する。4種の人工アミノアシルtRNA:mS-tRNA(T)、aG-tRNA(N)、aG-tRNA(S)及びaG-tRNA(V)を括弧内に与えられるそれぞれの人工tRNAによって認識されるそれぞれのmRNAコドンと共に使用した。高度に標識された<sup>3</sup>Hアミノ酸はEであった。全d. p. m.:小型カラムから溶出した全<sup>3</sup>Hd. p. m.。全ての翻訳は、翻訳因子IF1、IF2、IF3、EF-Tu及びEF-Gを包含した。

30

| mRNA    | 供給されたアミノアシルtRNA                                  | 全dpm | 疑似ペプチド産物         |
|---------|--|------|------------------|
| mMTNSVE | fM、mS-tRNA(T)、aG-tRNA(N)、aG-tRNA(S)、aG-tRNA(V)、E | 2353 | fM-mS-aG-aG-aG-E |
| mMTNSVE | fM、aG-tRNA(N)、aG-tRNA(S)、aG-tRNA(V)、E            | 899  | なし(バックグラウンド)     |
| -       | fM、mS-tRNA(T)、aG-tRNA(N)、aG-tRNA(S)、aG-tRNA(V)、E | 1033 | なし(バックグラウンド)     |
| mMTTV   | fM、T、V   | 7100 | fMTTV            |

40

## 【 図面の簡単な説明 】

50

【0171】

【図1】E. coliからの5種のhisタグE. coli翻訳因子の過剰発現及び精製

。【図2】リボソーム指令ペプチド合成段階の概略図。

【図3】短鎖mRNA鑄型。

【図4】精製hisタグ翻訳系におけるmRNAからのオリゴペプチド合成速度の特徴付け。

【図5】mRNAのMTVを有する精製翻訳系における翻訳因子の依存性。

【図6】MTTVのmRNA鑄型から産生された産物のHPLC分析。

【図7】精製翻訳系でのオリゴペプチド産物の合成。

10

【図8】精製hisタグ翻訳系を使用した非天然アミノ酸を含むペプチドの合成及び選択

。【図9】Cys-tRNA<sup>Cys</sup>の化学的ピオチニル化。

【図10】ピオチニル-Cys-tRNA<sup>Cys</sup>の精製。

【図11】精製系を使用した精製ピオチニルCysのfM-T-bC-V疑似ペプチドへの取り込み。

【図12】隣接する多数の非天然アミノ酸をfM-T-bC-bC-V疑似ペプチドに取り込むためのアッセイ。

【図13】全tRNAと15種の異なる<sup>14</sup>C標識アミノ酸の混合物(ニュー・イングランド・ニュークリア社)との結合活性(TCA沈殿法によって測定される)についての本発明の純粋な翻訳系のアッセイ(「蛋白質ポリメラーゼ」)。

20

【図14】選択のコドンに対して特異的な非天然アミノ酸を結合したアミノアシルtRNAを合成するために一般化できる方法。

【図15】左から右にtRNA<sup>Asn</sup>(T)tRNA<sup>Asn</sup>(S)tRNA<sup>Asn</sup>(V)と命名されたtRNA<sup>Asn</sup>(N)の3種のアンチコドン変異体。

【図16】非結合tRNA(レーン1、3及び5)及びアミノアシルtRNA基質(レーン2、4及び6)の酸/尿素ポリアクリルアミドゲル電気泳動。

【図17】隣接するaGアミノ酸を疑似ペプチドに選択的に取り込むための一般的な方法

。【図18】純粋リボソームディスプレイ。

30

【図19】本発明のスペーサーmRNA及びそれらのコードペプチド産物。

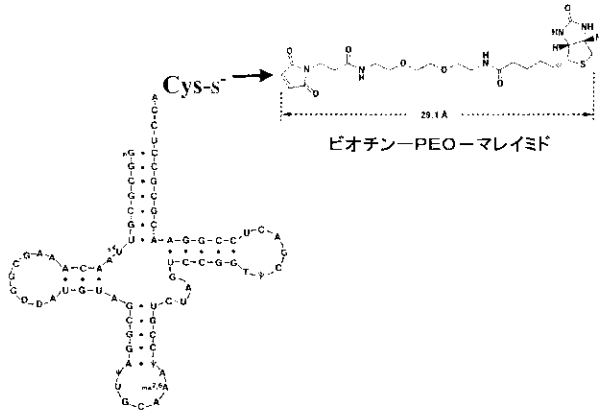
【図20】本発明の精製系を使用した図19のスペーサーmRNAの翻訳。

【図21】純粋mRNAディスプレイ。

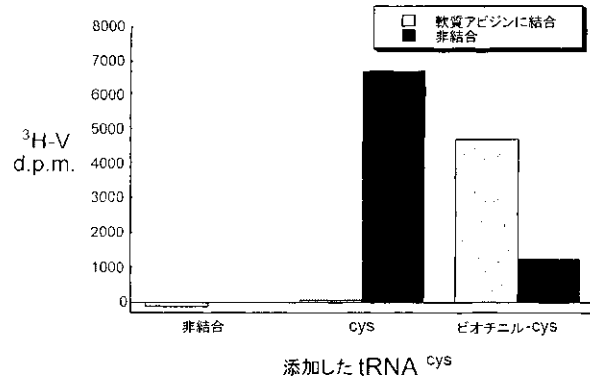
【図22】プロメガ社製のTranscendピオチニル-リシンtRNA<sup>Lys</sup>。



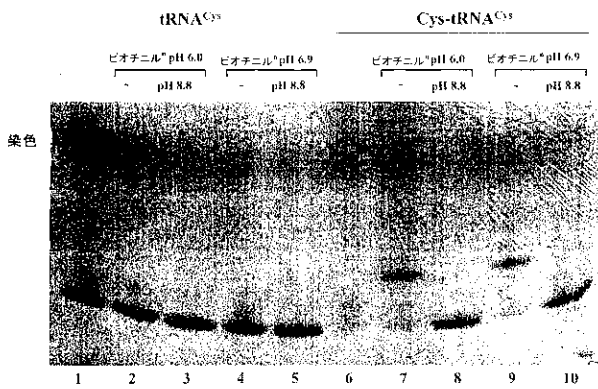
【図 9】



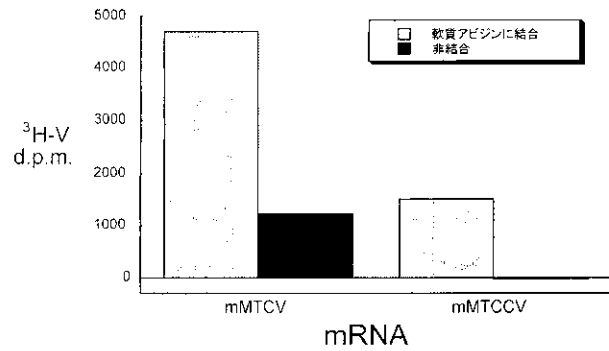
【図 11】



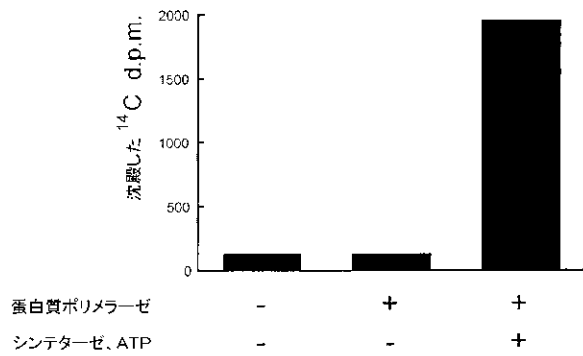
【図 10】



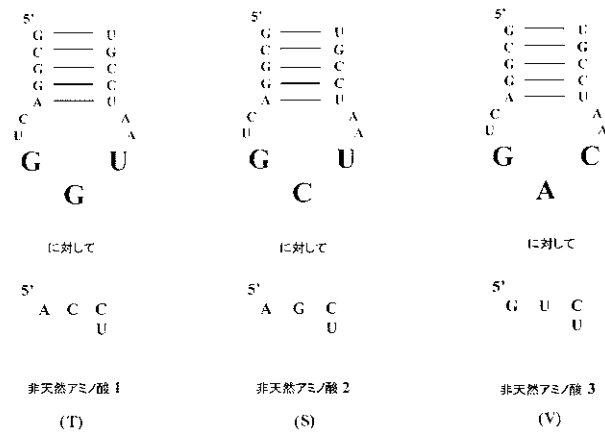
【図 12】



【図 13】



【図 15】





【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  
1 August 2002 (01.08.2002)

PCT

(10) International Publication Number  
WO 02/059293 A2

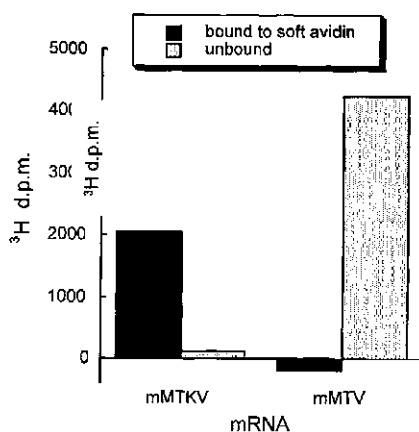
- (51) International Patent Classification: C12N 15/00
- (21) International Application Number: PCT/US02/02344
- (22) International Filing Date: 25 January 2002 (25.01.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Date: 25 January 2001 (25.01.2001) US 60/264,147
- (71) Applicants and (72) Inventors: FORSTER, Anthony, C. [AUS]; 173 Market Road, Lexington, MA 02421 (US); BLACKLOW, Stephen, C. [US]; 116 Tower Road, Lincoln, MA 01773 (US)
- (74) Agents: VINCENT, Matthew, P. et al.; Ropes & Gray, Patent Group, One International Place, Boston, MA 02110 (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, 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, MY, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TL, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (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, UJ, 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: without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: PROCESS AND COMPOSITIONS FOR PEPTIDE, PROTEIN AND PEPTIDOMIMETIC SYNTHESIS

WO 02/059293 A2



(57) Abstract: The present invention is a simplified, highly-purified, processive translation system that does not require the addition of translation factors (E-P, W, W2 or rescue). A new translation process offers new, potentially improved, routes to all peptides, proteins and peptidomimetics currently synthesized by alternative routes.

**WO 02/059293 A2**



---

*For two-letter codes and other abbreviations, refer to the "Guidance Note: on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Process and Compositions for Peptide,  
Protein and Peptidomimetic Synthesis**

**Background of the Invention**

5 The recognition and binding of ligands regulates almost all biological processes, such as immune recognition, cell signaling and communication, transcription and translation, intracellular signaling, and catalysis, *i.e.*, enzyme reactions. There is a long-standing interest in the art to identify and synthesize natural or unnatural ligand molecules which act as agonists or which can agonize or  
10 antagonize the activity of ligands such as hormones, growth factors, and neurotransmitters; which induce B-cell (antibody-mediated) or T-cell (cell-mediated) immunity; which can catalyze chemical reactions; or which can regulate gene expression at the level of transcription or translation. A large proportion of such ligands are proteins, peptides, and peptidomimetics.

15 The traditional approach to ligand and drug discovery relies heavily on a mixture of serendipity and hard work. Screening natural products from animal and plant tissues, or the products of fermentation broths, or the random screening of archived synthetic molecules have been the most productive avenues for the identification of new lead compounds.

20 However, recent trends in the search for novel pharmacological agents have focused on the preparation of combinatorial libraries as potential sources of new leads for drug discovery. At the heart of this new field of "combinatorial chemistry" is a collection of differing molecules which can be prepared either non-biosynthetically or biosynthetically and screened for biological activity in a variety  
25 of formats. Through the use of non-biosynthetic techniques, *e.g.*, encoding, spatially addressing and/or deconvolution, combinatorial libraries of peptides, peptidomimetics and non-peptide-based molecules can be synthesized by batch processes and, importantly, the molecular identity of individual members of the library can be ascertained in a drug screening format (*e.g.* Lam *et al.* (1993) *Gene*  
30 137, 13-16; Dooley *et al.* (1994) *Science* 266, 2019-2022). While non-biosynthetic libraries have the advantage of being unrestricted to biological monomers (such as natural amino acids and nucleotides) and their derivatives, they have the

disadvantage of being limited in the number of molecules that may be screened within several weeks: usually  $10^5$  to  $10^8$  at most, which is too few molecules for favoring the identification of high affinity ligands for a target of interest (Roberts (1999) *Curr. Op. Chem. Biol.* 3, 268-273; Wilson *et al.* (2001) *PNAS* 98, 3750-3755). Biosynthetic libraries, however, often do not suffer from this limitation because there are examples of such libraries that enable  $10^{15}$  different peptide, RNA or DNA molecules to be screened within several weeks (Roberts, *supra*). This is achieved by reiterative selection and amplification of individual biosynthetic library members, often with associated mutagenesis steps (*e.g.* affinity maturation, mutagenic PCR, or DNA shuffling (Roberts, *supra*)) in a process analogous to Darwinian evolution, sometimes termed directed evolution.

Many prior methods that allowed the isolation of proteins from partially or fully randomized pools did so through an *in vivo* step. Methods of this sort include monoclonal antibody technology (Milstein, *Sci. Amer.* 243:66 (1980); and Schultz *et al.*, *J. Chem. Engng. News* 68:26 (1990)), phage display (Smith, *Science* 228:1315 (1985); Parmley and Smith, *Gene* 73:305 (1988); and McCafferty *et al.*, *Nature* 348:552 (1990)), peptide-lac repressor fusions (Cull *et al.*, *PNAS* 89:1865 (1992)), and classical genetic selections. Each of these methods relies on a topological link between the protein and the nucleic acid, since only nucleic acids can be replicated. Thus, the information of the protein is retained and can be recovered in readable, nucleic acid form.

Alternative protein selection technologies are performed without *in vivo* steps. The stalled translation method, often termed ribosome display, is a technique in which selection is for some property of a nascent protein chain that is still complexed with the ribosome and its mRNA (Kawasaki U.S. Patent 5,658,754; Tuerk and Gold, *Science* 249:505 (1990); Irvine *et al.*, *J. Mol. Biol.* 222:739 (1991); Korman *et al.*, *PNAS* 79:1844-1848 (1982); Matheakis *et al.*, *PNAS* 91:9022-9026 (1994); Matheakis *et al.*, *Meth. Enzymol.* 267:195 (1996); and Hanes and Pluckthun, *PNAS* 94:4937 (1997)). The mRNA-protein fusion method or mRNA display (Nemoto *et al.* (1997) *FEBS Lett.* 414, 405-408; Yanagawa *et al.* US Patent 6228994; Szostak *et al.* US Patents 6281344, 6261804, 6258558, 6214553, and 6207446; Roberts and Szostak (1997) *PNAS* 94, 12297-12302) covalently couples

the mRNA directly to its protein product via a DNA/puromycin linker. A method for synthesizing "naked" mRNA-peptide fusions that is not compromised by the presence of stop codons is to synthesize peptides in micelles in such a way that they can dissociate from the ribosomes and then rebind to their specific mRNAs (e.g. proteins containing streptavidin sequences will bind to biotinylated mRNA; Doi and Yanagawa (1999) *FEBS Lett.* 457, 227-230).

The prior art "natural" (L-) peptide library techniques, however, suffer from a number of disadvantages. First, the libraries, which consist almost entirely of chiral monomers (amino acids) lack the enantiomers of the chiral monomers. For example, with L-peptide libraries, while the 20 naturally occurring amino acids provide a wide range of steric, electronic and functional groups, the chirality of the C-alpha carbon effectively limits the three-dimensional shape space which is accessible by the prior art display technology. L-peptide libraries also lack a number of common organic chemistry functional groups which may be helpful for forming non-covalent or covalent complexes with targets (e.g. alkene, alkyl urea, alkyl halide, and ketone), and lack the enormous additional shape diversity achievable with "unnatural" amino acids (either previously synthesized or theoretical). Moreover, as therapeutic agents, peptides with natural L-amino acids are often less preferable than their unnatural enantiomers (D-peptides) or analogs because L-peptides can be limited in use by poor pharmacokinetic profiles due to *in vivo* processing. For example, L-peptides can be rapidly degraded by proteases after administration to an animal, thus requiring a higher effective dose. Furthermore, pharmaceutical peptides can elicit strong immunogenic responses in patients, further contributing to their rapid clearance and also causing inflammatory reactions that may be toxic. One approach to preventing the degradation of the therapeutic peptide has been to generate non-hydrolyzable peptide analogs such as retro-inverso analogs (c.f., Sisto *et al.* U.S. Patent 4,522,752), retro-enantio analogs (c.f., Goissis *et al.* (1976) *J Med Chem* 19:1287-90); trans-olefin derivatives (c.f., Shue *et al.* (1987) *Tetrahedron Letters* 28:3225); and phosphonate derivatives (c.f., Loots *et al.*, in *Peptides: Chemistry and Biology*, (Escam Science Publishers, Leiden, 1988, p. 118). However, in most instances the backbone of the peptide is altered in order to render the peptidomimetic resistant to proteolysis. In doing so, the resulting

peptidomimetic can suffer from decreased bioactivity through loss of certain binding contacts between the natural peptide backbone and target receptor, as well as changes in the steric space relative to the peptide due to alteration in dihedral angles and the like. Another problem is that almost all L-peptides do not cross biological membranes readily because of their hydrophilicity. In contrast, D-peptides and peptides containing other unnatural amino acids (peptidomimetics) such as N-methyl amino acids have increased resistance to proteases, and the peptidomimetic drug Cyclosporin A can cross membranes and is orally available, in part because it contains several N-methyl peptide linkages which are more hydrophobic than natural peptide linkages (Zawadzke and Berg (1992) *J Am Chem Soc* 114:4002; Walsh *et al.* (1992) *J. Biol. Chem.* 267, 13115-13118). Unfortunately, chemically synthesized (non-biosynthetic) peptidomimetic libraries, such as D-peptide libraries (Lam *et al.*, *supra*; Dooley *et al.*, *supra*) suffer the limitation of library size discussed above, and methodological tricks to overcome the size limit of peptidomimetic libraries, such as mirror-image phage or ribosome display (Schumacher *et al.* (1996) *Science* 271, 1854-1857; Eckert *et al.* (1999) *Cell* 99, 103-115; Forster *et al.* PCT publication WO97/35194, are limited by the onerous requirement of chemically synthesizing an enantiomeric target.

Proteins, peptides and peptidomimetics are currently synthesized in three different ways, each with their own inherent limitations:

1. Synthetic peptide chemistry can be used routinely for the synthesis in high yield and purity of very diverse peptidomimetics of up to about 30 residues in length (Ecker *et al.*, *supra*).

However, the method is inefficient or impractical for longer products because of inefficient coupling steps, purification problems, and folding difficulties. There are also synthetic restrictions because of the need for compatible protecting groups for all of the reactive side chains in a desired product. Furthermore, synthetic peptidomimetics cannot be genetically encoded for reiterative selection, amplification, and mutation (evolution), limiting the complexity of synthetic peptidomimetic libraries to about  $10^8$  molecules, too few for optimal drug discovery.

2. *In vivo* translation using living cells is widely used for the efficient synthesis and posttranslational modification of short or long proteins from a genetically encoded natural or recombinant DNA sequence.

5 However, synthesis may be inefficient if the gene product is toxic, and there may be difficult purification and refolding problems, particularly if the protein is expressed in inclusion bodies. Most importantly, the method suffers from the inability to incorporate multiple unnatural amino acids selectively or control the post-translation modification process (e.g. protease-catalysed processing or degradation).

10 3. *In vitro* translation with crude cell extracts generally overcomes the toxicity problem (but does not control post-translational modifications), may result in easier purification and folding, and allows the selective incorporation of a single unnatural amino acid per protein using an artificial suppressor tRNA (Noren *et al.* (1989) *Science* 244,182-188).

15 However, the incorporation of an unnatural amino acid by this approach usually suffers from much lower yields than *in vivo* systems because it relies on inherently inefficient suppressor tRNAs competing with termination factors. Although over one hundred different unnatural amino acids have been incorporated on an individual basis (e.g. Mendel *et al.* (1995) *Annu. Rev. Biophys. Biomol. Struct.* 24, 435-462), this strategy has been restricted to selective incorporation of only a single unnatural amino acid per protein at only one of the three termination (nonsense) codons (the UAG codon) because of competition at amino acid (sense) codons from natural amino acids catalysed by the tRNA charging and proofreading activities of the twenty different aminoacyl tRNA synthetases, and because an attempt to use a second termination codon (UGA) failed due to readthrough by the ribosome (Cloud *et al.* (1996) *Chem. and Biol.* 3, 1033-1038).

20 Many attempts to incorporate unnatural amino acids selectively at sense codons in a generalizable manner have also failed. For example, in the most commonly used method for unnatural amino acid incorporation, where a high-specific-activity, radioactive-isotope derivative of a natural amino acid is incorporated by *in vivo* translation to synthesize a radiolabelled protein, it is well known that the specific activity of the radioactive amino acid is always substantially

reduced by competition for incorporation by the unlabelled version of the amino acid present in the crude translation system, despite withholding the unlabelled version from the added unlabelled amino acid pool. Analogous analog dilution results are obtained by the Promega company using their commercially available kit for incorporation of another reporter group, biotin-labelled lysine (literature 5 accompanying Transcend™ non-radioactive translation systems). Furthermore, filtration of a crude translation extract to remove natural amino acids followed by supplementation with all of the natural amino acids except lysine and supplementation with a lysine tRNA charged with an amino acid analog resulted in 10 incorporation of lysine analog to lysine at a ratio of only 1:3 to 1:4 (Crowley *et al.* (1993) *Cell* 73, 1101-1115). While a low selectivity of amino acid analog incorporation is sufficient for certain applications (Rothschild *et al.*, US patent 5,643,722) it is clearly incompatible with many applications such as that requiring the amplification and characterization of genetically encoded specific 15 peptidomimetic sequences. It has proved possible to incorporate two different unnatural amino acids using two different frameshifting suppressor tRNAs (Hohsaka *et al.* (1999) *JACS* 121, 12194-12195), and many identical unnatural amino acids have been incorporated using an inhibitor specific for Phe aminoacyl-tRNA synthetase (Baldini *et al.* (1988) *Biochem.* 27, 7951-7959). However, both of these 20 methods are not generalizable in the manner necessary for the incorporation of many different unnatural amino acids into a single peptidomimetic. In order to overcome these restrictions inherent in crude and *in vivo* translations, an elaborate strategy for expansion of the genetic code based on orthogonal tRNAs and orthogonal unnatural nucleic acid base pairs has been proposed, but development beyond a single *in vitro*-engineered termination codon (Bain *et al.* (1992) *Nature* 356, 537-539) has proved 25 to be too challenging technically (Service (2000) *Science* 289, 232-235).

We envisioned that this problem potentially may be solved by using a pure *in vitro* translation system. Competition between unnatural amino acids and natural amino acids or termination factors could potentially be avoided by the omission of 30 certain components such as certain amino acids, tRNAs, aminoacyl tRNA synthetases and/or termination factors. Unfortunately, the minimal requirements for mRNA-dependent polypeptide synthesis have been difficult to define because of the

large number of macromolecules involved. Reconstitution of translation from purified components has been achieved for *E. coli*, but the number of translation factors required remains controversial.

The first purified translation system, constructed by the Weissbach laboratory, efficiently translated four *E. coli* mRNAs with strong dependencies on high salt-washed ribosomes, initiation factors (partial IF1 dependency), elongation factor Tu (EF-TuII), and groups of aminoacyl-tRNA synthetases, and partial dependencies on met-tRNA<sup>Met</sup> formyltransferase and elongation factor G (EF-G), with no dependency on elongation factor Ts (EF-Ts) or termination factors (Kung *et al.* (1978) *Arch. Biochem. and Biophys.* 187, 457-463). Because of the difficulties in maintaining so many purified components and in removing trace contaminants, the search for additional general translation factors was facilitated by simplifying the system to di- or tripeptide synthesis from fMet-tRNA<sup>fMet</sup> and one or two elongator aminoacyl-tRNAs, thereby avoiding the requirement for aminoacyl-tRNA-synthesizing enzymes (Weissbach *et al.* (1984) *Biotechniques* 2, 16-22).

When a second group, led by Ganoza, extended the latter simplified approach to longer peptides using *in vitro*-charged total tRNA and release factors, translation of bacteriophages MS2 and  $\phi$ 1 were found to be dependent on three additional factors, termed EF-P, W and rescue (Green *et al.* (1985) *Biochem. Biophys. Res. Com.* 126, 792-798; Ganoza *et al.* (1985) *PNAS* 82, 1648-1652). The absence of these factors resulted in inefficient processivity. For example, there was a predominance of di-, tri-, tetra- and pentapeptide pausing or premature termination products in hexapeptide synthesis reactions. A possibly related translation factor termed dead/W2 (several kD bigger than W) and also EF-P have been cloned, are necessary for maximal growth, and are homologous to eukaryotic initiation factors (Aoki *et al.* (1991) *Nucleic Acids Res.* 19, 6215-6220; Aoki *et al.* (1997) *J. Biol. Chem.* 272, 32254-32259; Lu *et al.* (1999) *Int. J. Biochem. Cell Biol.* 31, 215-229).

In apparent conflict with the results of Ganoza, two other groups have reported synthesis of short peptides from aminoacyl-tRNA substrates using purified components without the addition of EF-P, W, W2 or rescue, although these two groups did not directly document the processivity of their systems or the purity of their ribosomes (Stade *et al.* (1995) *Nucleic Acids Res.* 23, 2371-2380; Pavlov *et al.*

(1997) *EMBO J.* 16, 4134-4141). If the discrepancy is real, one can only speculate as to the explanation. For example, because EF-P, W, and rescue can be purified from ribosome preparations (Ganoza *et al.* (1996) *Biochemie* 78, 51-61), it is possible that the ribosomes used by the latter two groups, prepared by very different procedures from that used by Ganoza's group, were contaminated with EF-P, W, W2 and/or rescue. This is problematic because contamination with EF-P, W, W2 and/or rescue likely implies contamination with more abundant proteins, such as aminoacyl-tRNA synthetases and termination factors, that could cause unwanted reactions. Alternatively, EF-P, W, W2 and/or rescue may only be required for efficient processivity in Ganoza's system.

The ability to synthesize peptides or proteins from a pure translation system without added EF-P, W (sometimes called W2) and rescue is desirable, if possible, because these proteins are not well understood in terms of function, resulting in difficulty in assaying their activities and therefore following the purification of active protein. Furthermore, there is controversy with respect to the actual size of W (or W2) and whether W and W2 represent derivatives of the same proteins, and the gene for rescue is yet to be cloned.

#### **Summary of the Invention**

The present invention is a simplified, highly-purified, processive translation system that does not require the addition of translation factors EF-P, W, W2 or rescue. A new translation process offers new, potentially improved, routes to all peptides and proteins currently synthesized by alternative routes. This process overcomes the limitations inherent in methods 1, 2 and 3 described above for protein, peptide and peptidomimetic synthesis.

In one preferred embodiment, the purified system can be used for the synthesis of peptide or protein ligands or catalysts, such as insulin, growth hormone or erythropoietin.

In another preferred embodiment, the purified system can be used for "pure ribosome display" and "pure mRNA display" selection experiments, in contrast to existing ribosome and mRNA display systems which rely on crude cell extracts. There are several advantages associated with performing peptide and protein display in a pure system, such as an expected lack of post-translational modification of

peptides, a lack of proteases which often cause protein degradation problems, and a lack of competition from contaminants in the selection steps. Additional advantages include:

5 (i) The absence of ribonucleases (demonstrated by measured long-term stability of radioactive mRNA in our pure *E. coli* system (results not shown)) avoids problems associated with mRNA degradation observed in various crude systems, especially *E. coli* (Roberts, *supra*); it is obviously important that the mRNA not be degraded before it can be translated and selected.

10 (ii) It is expected that the pure reconstituted system is not contaminated by translation termination factors. Indeed, our system is stimulated by addition of termination factors. Workers using the most popular crude display systems have found it necessary to remove stop codons from mRNAs to avoid rapid release of nascent peptides before either selection or before fusion to the mRNA conjugate before selection (both fusion and selection are slow processes). The removal of stop  
15 codons usually requires special mutagenesis steps in the case of individual mRNAs, and is more problematic for natural mRNA libraries or synthetic combinatorial libraries where it is impossible to specifically mutate all stop codons. The problem with stop codons in libraries has been circumvented by either randomly generating libraries of small subdomains of proteins (which lack full-length proteins, have  
20 under-represented carboxy-terminal subdomains, encode many inappropriate boundaries with respect to ability to fold correctly, and contain an abundance of sequences from unnatural open reading frames) or by selecting out members of random libraries that contain stop codons (thereby wasting a major portion of the synthetic reaction and diversity; Cho *et al.* (2000) *J. Mol. Biol.* 297, 309-319). In  
25 the purified ribosome or purified mRNA display systems, the peptidyl-tRNA can remain stably associated with the ribosome for more than a day (stability is especially favored if, after translation is complete, the temperature is lowered and/or the salt concentration is increased (Schaffitzel *et al.* (1999) *J. Immunol. Methods* 231, 119-135)); either the ribosome stops elongating at a sense codon for which no  
30 tRNA is provided (thereby avoiding competition with termination factors altogether), or it stops at a stop codon, or it stops at the end of mRNAs lacking a stop codon. Thus, libraries of full-length translation products from natural mRNAs

can be prepared with this invention, and such expression libraries can be either directly subjected to *in vitro* selection or they can be spatially addressed by hybridization to a DNA microchip for genomic and proteomic studies.

(iii) It is expected that the pure reconstituted system is not contaminated by the tmRNA system that degrades peptides synthesized from mRNAs lacking a stop codon. Workers using the crude system have found it important to try to inhibit this tmRNA system (Hanes and Pluckthun (1997) *PNAS* 94, 4937-4942).

In another preferred embodiment, the invention enables the mRNA-directed synthesis of specific peptidomimetics (peptide analogs) in a generalizable manner, greatly increasing the diversity and length of peptidomimetics available. Possibilities include existing peptidomimetic ligands and drugs (including non-ribosomally biosynthesized ligands, such as Cyclosporin A) and derivatives thereof.

In another embodiment, the invention enables the genetic encoding of peptidomimetic products for catalyst, ligand and drug discovery by *in vitro* evolution (e.g. by using pure ribosome display or pure mRNA display described above). Specific synthesis of peptidomimetics is not possible in existing crude ribosome and mRNA display methodologies because natural amino acids compete with unnatural amino acids for incorporation in crude translation systems.

The invention facilitates the isolation of peptides and peptidomimetics with desired properties. In one embodiment, the method is directed to identifying ligands for a target molecule. Exemplary target molecules include peptides, nucleic acids, carbohydrates and non-polymeric molecules, such as steroids, inositols, lipid soluble vitamins, terpenes, acetogenins, neurotransmitters, or a transition state analog. In a preferred embodiment, the target molecule is a protein. The protein target can be, to illustrate, a receptor, an enzyme, a DNA-binding protein or a protein complex, or a portion or domain thereof which retains a screenable activity.

In preferred embodiments, the subject method is used to generate variegated population of test peptides or peptidomimetics of at least  $10^3$  different sequences, though more preferably at least  $10^8$  different sequences, and most preferably at least  $10^{15}$  different sequences.

Yet another aspect of the invention relates to compounds, such as peptides and peptidomimetics, identified by the subject method, and their uses. This also

includes conjugates and derivatives of such peptides and peptidomimetics (e.g. conjugation to cationic peptide sequences that enable efficient transport across membranes of attached peptide or peptidomimetic sequences (Moore and Roshash (2001) *Science* 294, 1841-1842).

5 Another aspect of the invention relates to kits for synthesis and/or evolution of peptides or peptidomimetics.

#### **Description of the Drawings**

Fig. 1: Over-expression and purification of five his-tagged *E. coli* translation factors from *E. coli*. After SDS-PAGE on 15% gels, samples were stained with  
10 Coomassie Blue. U, uninduced total cells; I, IPTG-induced total cells; P, purified protein eluted from Ni<sup>2+</sup> beads; M, molecular weight marker proteins (sizes indicated in kD).

Fig. 2: Schematic illustrating steps in ribosome-directed peptide synthesis. The three enzymatic reactions depicted by arrows are initiation (top), the first  
15 elongation step (right), and subsequent translocation and elongation steps (bottom). Peptide products can be released from the peptidyl-tRNAs by base-catalyzed hydrolysis for analysis. Products GDP and P<sub>i</sub> are not shown. E, exit site; P, peptidyl site; A, aminoacyl site.

Fig. 3: Short mRNA templates. The DNA primer-template pair used to  
20 synthesize the longest mRNA is illustrated at the top. The predicted translation products from our purified system are also shown (aminoacyl tRNAs for the 3' terminal codons GAA (Glu) and UUC (Phe) were not used). bK: biotin-labeled-lysine. S-D: Shine and Dalgarno ribosome binding site.

Fig. 4: Characterization of oligopeptide synthesis rates from mRNA MTV in  
25 a purified his-tagged translation system. fMTV was measured by <sup>3</sup>H-valine incorporated into peptide products in translations containing IF1H, IF2H, IF3H, EF-TuH, EF-GH and 0.020 A<sub>260</sub>/μl ribosomes. *Triangles*: translations were started by mixing preincubated initiation components with preincubated elongation components. *Squares*: translations were started by transferring the translation mix  
30 from 0°C to 37°C. Aliquots were terminated with NaOH at the indicated times beginning at 1 min. Peptide product d.p.m. was calculated by subtracting d.p.m. obtained in aliquots terminated before 37°C incubation. Individual data points from

representative experiments are plotted, with variations estimated to be less than 20%. A tangent line to the preincubation reaction curve is drawn to estimate the steady state rate.

Fig. 5: Translation factor dependencies in the purified translation system with mRNA MTV. *Light bars*:  $^3\text{H}$ -valine incorporated into peptide products (a measure of fMTV) shows strong dependencies on IF2H, IF3H, EF-TuH and EF-GH (IF1H was omitted from these translations; see Materials and Methods). *Dark bars*:  $^{14}\text{C}$ -threonine incorporated into the same products (a combined measure of fMT and fMTV). Peptide synthesis in a 30 minute translation started by transfer from 0°C to 37°C was calculated by subtracting d.p.m. obtained in a control reaction lacking mRNA (1.3% of maximal d.p.m.) from total d.p.m. The maximum concentration of synthetic product obtained was 0.12  $\mu\text{M}$  for both T and V.

Fig. 6: HPLC analysis of products produced from the MTTV mRNA template. The peptide products of a dual-labeled translation were first released from the tRNAs with base and then mixed with unlabeled marker peptides. The mixture was acidified, microcentrifuged to remove insoluble material, and microcentrifuged through a 10 kD filter before injection for analysis (see Materials and Methods). The elution positions of marker peptides are indicated above the chromatogram. *Filled circles*:  $^{14}\text{C}$ -threonine total d.p.m. *Open circles*:  $^3\text{H}$ -valine total d.p.m. in the same fractions. The amount of product synthesized was 2.1 pmol in 30  $\mu\text{l}$  (70 nM).

Fig. 7: Inclusion of the epsilon sequence enhances the synthesis of oligopeptide product in a purified translation system. Comparison between rates of fMVT synthesis from mRNAs MVT (*circles*), scrambled-epsiMVT (*squares*) and  $\Delta$ epsiMVT (*triangles*). Dual-labeled translations containing IF1H, IF2H, IF3H, EF-TuH and EF-GH were analysed as described in Fig. 5. The maximum concentration of oligopeptide synthesized in 30 minutes was 0.25  $\mu\text{M}$  using 0.5  $\mu\text{M}$  of each aminoacyl tRNA.

Fig. 8: Synthesis and selection of peptides containing an unnatural amino acid using the purified his-tagged translation system. Translation mixes containing biotin-labeled-lys-tRNA<sup>lys</sup>, fmet-tRNA<sup>fmet</sup>, thr-tRNA<sup>thr</sup> and  $^3\text{H}$ -val-tRNA<sup>val</sup> substrates and either mRNA MTKV or MTV (see Fig. 3 for translation products) were incubated at 37°C for 30 minutes. The peptides and amino acids were released

from the tRNAs and ribosomes with base, neutralized, and the mixtures incubated with soft avidin beads to bind biotin-containing molecules. The beads were washed four times to remove unbiotinylated molecules before counting bound  $^3\text{H}$  (*dark bars*: a measure of products containing biotin-labeled-lysine covalently linked to  $^3\text{H}$ -valine). The pooled washes were filtered, acidified, and passed through a cation exchange column to count unbound  $^3\text{H}$  (*light bars*: a measure of formylated peptide products containing  $^3\text{H}$ -valine without biotin-labeled-lysine or lysine). Bound and unbound  $^3\text{H}$  d.p.m. are plotted after subtracting d.p.m. obtained in a control reaction lacking mRNA (23% and 15% of maximal bound and unbound d.p.m., respectively).  
5 When biotin-labeled-lys-tRNA<sup>Cys</sup> is omitted from a translation of mRNA MTKV, binding of  $^3\text{H}$  to the beads is not observed (not shown).

Fig. 9. Chemical biotinylation of Cys-tRNA<sup>Cys</sup>

Fig. 10. Purification of biotinyl-Cys-tRNA<sup>Cys</sup>. Biotinyl-Cys-tRNA was electrophoresed as in Fig. 16 (see below), and the product purified by cutting out the appropriated band (*i.e.* the major band in lane 7 or 9) and elution at 4°C. Controls including +/- charging of the tRNA with Cys, +/- biotinylation (at either pH 6.0 or 6.9), and +/- subsequent hydrolysis of the aminocyl linkage with Tris-HCl pH 8.8 (lanes 1-6,8,10) showed that biotinylation was specific to the Cys, and that Cys-tRNA did not comigrate with biotinyl-Cys-tRNA.  
15

Fig. 11. Incorporation of purified biotinyl-Cys into fM-T-bC-V peptidomimetic using the purified system. Translations contained mMTCV, tRNAs charged with fM, T, V, and either uncharged, Cys-charged or purified biotinyl-Cys-charged tRNA, with controls lacking mRNAs. Selection was with Soft Avidin as in Fig. 8. Incomplete binding of  $^3\text{H}$ -peptide with biotinyl-Cys substrate was likely due to the low affinity of Soft Avidin (an avidin derivatised to have a much higher  $K_d$  of approximately  $10^{-7}\text{M}$ ) for peptides containing a single biotin.  
20

Fig. 12. Assay for incorporation of adjacent large unnatural amino acids into fM-T-bC-bC-V peptidomimetic. The experiment was carried out as in Fig. 11. Note binding by Soft Avidin was complete for fM-T-bC-bC-V but incomplete for fM-T-bC-V, as expected for the much higher affinity for peptides containing more than one biotin.  
25

Fig. 13. Assay of our pure translation system ("protein polymerase") for charging activity (measured by TCA precipitation) with total tRNA and a mixture of fifteen different  $^{14}\text{C}$ -labelled amino acids (New England Nuclear). The added synthetases consisted of a tRNA-free crude aminoacyl-tRNA synthetase cell extract.

5 Fig. 14. A generalizable approach for the synthesis of aminoacyl-tRNAs charged with unnatural amino acids specific for the codon(s) of choice. The elongator: tRNA<sup>Asn-CA</sup>, synthesized *in vitro* (Fig. 16, lane 5) from our recombinant DNA clone prepared from synthetic oligodeoxyribonucleotides, contains substantial base alterations from the natural tRNA<sup>Asn</sup> sequence that are indicated by arrows.  
10 The anticodon of the tRNA is indicated with large letters. An amino NVOC-protected unnatural amino acid was chemically aminoacylated on pAcA (see upper right) and then ligated to the tRNA<sup>Asn-CA</sup> (produced by run-off transcription of P<sub>o</sub>k I cut template) with T4 RNA ligase (Fig. 16, lane 6). The approach is generalizable because no aminoacyl-tRNA-synthetase or natural tRNA was required.

15 Fig. 15. Three anticodon mutants of tRNA<sup>Asn(N)</sup> termed, from left to right, tRNA<sup>Asn(T)</sup>, tRNA<sup>Asn(S)</sup>, tRNA<sup>Asn(V)</sup>. The new anticodons of the tRNAs are indicated with large letters above the codons that they recognise. The genetic code has been redesigned so that the codons now specify whichever amino acid (natural or unnatural) is chosen to be ligated onto each tRNA<sup>Asn-CA</sup>.

20 Fig. 16: Acid/urea polyacrylamide gel electrophoresis (Varshney *et al.* (1991) *J. Biol. Chem.* 266, 24712-24718) of uncharged tRNAs (lanes 1, 3 and 5) and aminoacyl tRNA substrates (lanes 2, 4 and 6). In this gel system, the observed mobility of the free RNA species (lanes 1, 3 and 5) is retarded by aminoacylation (lanes 2, 4) or by ligation of an aminoacylated CA dinucleotide (lane 6).

25 Fig. 17. A generalizable approach for the selective incorporation of adjacent  $\alpha\text{G}$  amino acids into peptidomimetics. The experiment was carried out as in Figs. 8 and 11, but without a selection step. Note that tRNA<sup>Asn(N)</sup> is abbreviated here as tRNA<sup>AsnB</sup> for Asn-based.

30 Fig. 18: Pure ribosome display. This is the pure version of the crude *in vitro* system (without living material) of Mattheakis *et al.* (*supra*) that can be used for the reiterative synthesis and selection (directed evolution) of peptides and peptidomimetics. The peptidomimetics need to be long enough to traverse the

approximately 100-Angstrom-long ribosome tunnel that surrounds unreleased *de novo*-synthesized peptides in order to be selected. Dissociation of the peptide from the mRNA (and ribosome) can be prevented by omitting stop codons, release factors or certain aminoacyl-tRNAs, and by using antibiotics that stall the ribosomes.

5 Fig. 19: Our spacer mRNAs and their encoded polypeptide products.

Fig. 20. Translation of spacer mRNAs of Fig. 19 using our purified system. The ratios of elongator valine ( $^3\text{H}$ -labelled) to initiator formylmethionine ( $^{35}\text{S}$ -labelled) incorporated into peptide products was measured by analysis of TCA-precipitated products using a dual-labelled-d.p.m. counting program. The linear plot  
10 observed is that expected for high processivity.

Fig. 21: Pure mRNA display. This is a pure version of the crude *in vitro* system (without living material) of Nemoto *et al.* (*supra*) and Roberts and Szostak (*supra*) based on that of Mattheakis *et al.* (Fig. 18) that differs in that the mRNAs are conjugated with puromycin (Pm) to enable covalent fusion of the mRNAs to  
15 their peptide or peptidomimetic products. Thus, the mRNA-peptidomimetic fusion may be purified from other translation components before selection, enabling very short peptidomimetics to be displayed without masking by the ribosome tunnel. Since the fusion reaction is slow, it is important to omit release factors when using mRNAs containing stop codons to prevent release factor-catalysed peptide release.  
20 Typically, ribosomes are stalled by a deoxynucleotide sequence conjugated to the mRNA. We ligated mMTKV and mMTCV (Figs. 8 and 11) to pdA<sub>17</sub>dCdCPm (Trilink) efficiently using a custom-synthesized "splint" DNA (TTTTTTTTTAAATCAAC, designed to hybridize to the ends of either mRNA and also the pdA<sub>17</sub>dCdCPm), and T4 DNA ligase, then gel-purified the conjugates for  
25 use in fusion and selection experiments (Roberts and Szostak (1997) *PNAS* 94, 12297-12302) where X in the figure is a readily selectable amino acid such as C or bC (Fig. 9) or bK (Fig. 22).

Fig. 22. Promega's Transcend biotinyl-lysine tRNA<sup>Lys</sup> (*E. coli*; see Materials and Methods).

30

**Best Mode for Carrying Out the Invention:****Detailed Description of the Invention****(I) Overview**

The present invention relates to an *in vitro* translation system that has been reconstituted from purified components to enable the specific incorporation of multiple different natural and unnatural amino acid residues in a highly controlled and generalizable manner. Because of the removal of unwanted competition from certain wild-type amino acids and termination factors, the efficient and specific synthesis of genetically encoded small and long peptidomimetic molecules and libraries is possible.

This invention was developed because the long-felt need for a biosynthetic method capable of synthesizing genetically encoded peptidomimetics remained unsolved, despite considerable effort over many years by many workers skilled in the art.

This invention was also developed because of a long-felt need for a simplified highly purified translation system that does not require the addition of EF-P, W, W2 or rescue for efficient processivity. Based on the problems with processivity of translation in the most highly purified versions of such systems in the prior art and on the decades of research in the field without a clear solution, there appeared to be a low expectation of success.

**(II) Definition of Terms**

For convenience, certain terms employed in the specification, examples, and appended claims are collected here.

By a "protein" is meant any two or more naturally occurring amino acids, joined by one or more peptide bonds. "Protein" and "peptide" are used interchangeably herein.

The term "peptide" refers to an oligomer in which the monomers are natural amino acids (alpha-amino acids) joined together through amide bonds. Peptides are two or more amino acid monomers long, but more often are between 5 to 10 amino acid monomers long and can be even longer, *i.e.* up to 20 amino acids or more, although peptides longer than 20 amino acids are more likely to be called "polypeptides." The term "protein" is well known in the art and usually refers to a

very large polypeptide, or set of associated homologous or heterologous polypeptides, that has some biological function. For purposes of the present invention the terms "peptide," "polypeptide," and "protein" are largely interchangeable, as all three types can be synthesized by the translation system, and so are collectively referred to as peptides.

By "peptidomimetic" is meant a peptide analog containing one or more unnatural amino acids (e.g. unnatural side chains, unnatural chiralities, N-substituted amino acids, or beta amino acids), unnatural topologies (e.g. cyclic or branched) or unnatural chemical derivatives (e.g. methylated or terminally blocked), or any molecule, other than a peptide containing natural amino acids, that is synthesized by a ribosome, including those products that have unnatural backbones and even those with partially or totally substituted amide (peptide) bonds with ester, thioester or other linkages (Mendel, *supra*).

By the terms "amino acid residue" and "peptide residue" is meant an amino acid or peptide molecule without the  $-OH$  of its carboxyl group (C-terminally linked) or the proton of its amino group (N-terminally linked). In general the abbreviations used herein for designating the amino acids and the protective groups are based on recommendations of the IUPAC-IUB Commission on Biochemical Nomenclature (see *Biochemistry* (1972) 11:1726-1732). Amino acid residues in peptides are abbreviated as follows: Alanine is Ala or A; Cysteine is Cys or C; Aspartic Acid is Asp or D; Glutamic Acid is Glu or E; Phenylalanine is Phe or F; Glycine is Gly or G; Histidine is His or H; Isoleucine is Ile or I; Lysine is Lys or K; Leucine is Leu or L; Methionine is Met or M; Asparagine is Asn or N; Proline is Pro or P; Glutamine is Gln or Q; Arginine is Arg or R; Serine is Ser or S; Threonine is Thr or T; Valine is Val or V; Tryptophan is Trp or W; and Tyrosine is Tyr or Y. Formylmethionine is abbreviated as fMet or fM. By the term "residue" is meant a radical derived from the corresponding  $\alpha$ -amino acid by eliminating the OH portion of the carboxyl group and the H portion of the  $\alpha$ -amino group. The term "amino acid side chain" is that part of an amino acid exclusive of the  $-CH(NH_2)COOH$  portion, as defined by K. D. Kopple, "Peptides and Amino Acids", W. A. Benjamin Inc., New York and Amsterdam, 1966, pages 2 and 33; examples of such side chains of the common amino acids are  $-CH_2CH_2SCH_3$  (the side chain of

methionine),  $-\text{CH}_2(\text{CH}_2)_3\text{CH}_2\text{CH}_2\text{CH}_3$  (the side chain of isoleucine),  $-\text{CH}_2\text{CH}(\text{CH}_3)_2$  (the side chain of leucine) or  $-\text{H}$  (the side chain of glycine).

In certain embodiments, the amino acids used in the application of this invention are those naturally occurring amino acids found in proteins. Particularly suitable amino acid side chains include side chains selected from those of the following 21 natural amino acids: alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, selenocysteine, serine, threonine, tryptophan, tyrosine, and valine. However, the present invention specifically contemplates the use of analogs, derivatives and congeners of any specific amino acid referred to herein. For example, the present invention contemplates the use of radioactive amino acid analogs, amino acid analogs wherein a side chain is lengthened or shortened while still providing a carboxyl, amino or other reactive precursor functional group for polymerization, as well as amino acid analogs having variant side chains (with appropriate functional groups). For instance, the subject peptidomimetic can include an amino acid analog as for example,  $\beta$ -cyanoalanine, canavanine, djenkolic acid, norleucine, 3-phosphoserine, homoserine, dihydroxyphenylalanine, 5-hydroxytryptophan, 1-methylhistidine, 3-methylhistidine, allyl glycine (or its alkyne derivative), O-methyl-serine, biotinyl-lysine, biotinyl-cysteine (or other biotin-labelled amino acids) cyclohexylalanine, homoglutamate, D-alanine (or other D-amino acids), N-methyl glycine (or other N-methyl amino acids), epsilon-N-methyl-lysine, and radioisotope derivatives of the 21 natural amino acids or unnatural amino acids. Other naturally or non-naturally occurring amino acids which are suitable herein will be recognized by those skilled in the art and are included in the scope of the present invention.

The term "chiral" refers to molecules which have the property of non-superimposability of the mirror image partner, while the term "achiral" refers to molecules which are superimposable on their mirror image partner.

The term "stereoisomers" refers to compounds which have identical chemical constitution, but differ with regard to the arrangement of the atoms or groups in space. In particular, "enantiomers" refer to two stereoisomers of a compound which are non-superimposable mirror images of one another. "Diastereomers", on the

other hand, refers to stereoisomers with two or more centers of dissymmetry and whose molecules are not mirror images of one another. With respect to the nomenclature of a chiral center, terms "D" and "L" configuration are as defined by the IUPAC Recommendations. As to the use of the terms, diastereomer, racemate  
5 and enantiomer will be used in their normal context to describe the stereochemistry of peptide preparations.

The terms "D-amino acid" and "L-amino acid" each denote an absolute configuration by convention relative to the possible stereoisomers of glyceraldehyde. Thus, all stereoisomers that are stereochemically related to L-  
10 glyceraldehyde are designated L-, and those related to D-glyceraldehyde are designated D-, regardless of the direction of the rotation of plane of polarized light by the given isomer. In the case of threonine and isoleucine, there are two stereochemical centers, *i.e.*, the C $\alpha$  and the C $\beta$  atoms. The D-threonine and D-  
isoleucine employed herein preferably have stereochemistries at both chiral sites which are opposite (enantiomeric) to the stereochemistry of the L-enantiomers of  
15 those amino acids, *e.g.*, they are complete mirror images. Glycine is the only commonly occurring achiral amino acid. The presence of achiral amino acid residues such as glycine do not affect the designation of its chirality.

All chiral amino acids in protein synthesized *de novo* in nature, *e.g.*,  
20 "naturally occurring" are L-amino acids.

A "D-enantiomer" or "D-peptide enantiomer" refers to a peptide comprised of D-amino acid residues, as opposed to L-amino acids.

The term "synthetic" refers to production by *in vitro* chemical or enzymatic synthesis.

25 By "DNA" is meant a sequence of two or more covalently bonded, naturally occurring or modified deoxyribonucleotides.

By "RNA" is meant a sequence of two or more covalently bonded, naturally occurring or modified ribonucleotides. Two examples of modified RNAs included within this term are phosphorothioate RNA, and "transfer RNA" containing natural  
30 modified bases.

By "transfer RNA" or "tRNA" is meant any RNA molecule that can deliver peptide or peptidomimetic precursors to the ribosome in a manner specified by partial base-pairing to an mRNA.

5 By a "translation initiation sequence" is meant any sequence which is capable of providing a functional ribosome entry site. In bacterial systems, this region is sometimes referred to as a ribosome-binding or Shine-Dalgarno sequence.

By "messenger RNA" or "mRNA" is meant any nucleic acid containing a "translation initiation sequence".

10 A "reconstituted translation system" refers to a reaction mixture (a) capable of performing *in vitro* translation, e.g., mRNA-dependent protein synthesis, and (b) characterized by having less than 10 percent of the contaminating proteins found in cell lysate translation systems or wheat germ extract translation systems, and more preferably having less than 5 percent or even less than 1 percent of such contaminating proteins. In certain preferred embodiments, the subject reconstituted  
15 translation system is generated by admixing recombinantly produced and/or purified proteins.

By a "start codon" is meant three bases which signal the beginning of a protein coding sequence. Generally, these bases are AUG (or ATG); however, any other base triplet capable of being utilized in this manner may be substituted.

20 By a "stop codon" is meant three bases which signal the end of a protein coding sequence. Generally, these bases are UAG, UAA or UGA (where U may be substituted by T); however, any other base triplet capable of being utilized in this manner may be substituted.

By a "pause sequence" is meant a nucleic acid sequence which causes a  
25 ribosome to slow or stop its rate of translation, such as a DNA sequence.

By a "peptide acceptor" is meant any molecule capable of being added to the carboxyl-terminus of a growing protein chain by the catalytic activity of the ribosomal peptidyl transferase function. Typically, such molecules contain (i) a nucleotide or nucleotide-like moiety (for example, adenosine or an adenosine analog  
30 (dimethylation at the N-6 amino position is acceptable)), (ii) an amino acid or amino acid-like moiety (for example, any of the 20 D- or L-amino acids or any amino acid analog thereof (for example, O-methyl tyrosine or any of the analogs described by

Ellman *et al.*, *Meth. Enzymol.* **202**:301, 1991), and (iii) a linkage between the two (for example, an ester, amide, or ketone linkage at the 3' position or, less preferably, the 2' position); preferably, this linkage does not significantly perturb the pucker of the ring from the natural ribonucleotide conformation. Peptide acceptors may also  
5 possess a nucleophile, which may be, without limitation, an amino group, a hydroxyl group, or a sulfhydryl group. In addition, peptide acceptors may be composed of nucleotide mimetics, amino acid mimetics, or mimetics of the combined nucleotide-amino acid structure.

By "highly selective incorporation at each codon", it is meant at least 80  
10 percent selective incorporation of an amino acid residue at a position in the peptide or peptidomimetic corresponding to the codon, more preferably at least 90, 95 or even 98 percent selective incorporation.

By a peptide acceptor being positioned "at the 3' end" of a protein coding sequence is meant that the peptide acceptor molecule is positioned after the final  
15 codon of that protein coding sequence. This term includes, without limitation, a peptide acceptor molecule that is positioned precisely at the 3' end of the protein coding sequence as well as one which is separated from the final codon by intervening coding or non-coding sequence (for example, a sequence corresponding to a pause site). This term also includes constructs in which coding or non-coding  
20 sequences follow (that is, are 3' to) the peptide acceptor molecule. In addition, this term encompasses, without limitation, a peptide acceptor molecule that is covalently bonded (either directly or indirectly through intervening nucleic acid sequence) to the protein coding sequence, as well as one that is joined to the protein coding sequence by some non-covalent means, for example, through hybridization using a  
25 second nucleic acid sequence that binds at or near the 3' end of the protein coding sequence and that itself is bound to a peptide acceptor molecule.

By "covalently bonded" to a peptide acceptor is meant that the peptide acceptor is joined to a "protein coding sequence" either directly through a covalent bond or indirectly through another covalently bonded sequence (for example, DNA  
30 corresponding to a pause site).

By "mRNA-display" is meant any method for coupling, via covalent or non-covalent linkage(s), the peptide or peptidomimetic product from translation of a

mRNA to its cognate mRNA or cDNA or nucleotide sequence related to the peptide or peptidomimetic product. A commonly used linkage contains puromycin.

By an "altered function" is meant any qualitative or quantitative change in the function of a molecule.

5 By "binding partner," as used herein, is meant any molecule which has a specific, covalent or non-covalent affinity for a portion of a desired mRNA-peptide/peptidomimetic ribosomal complex or fusion. Examples of binding partners include, without limitation, members of antigen/antibody pairs, protein/inhibitor pairs, receptor/ligand pairs (for example cell surface receptor/ligand pairs, such as  
10 hormone receptor/peptide hormone pairs), enzyme/substrate pairs (for example, kinase/substrate pairs), lectin/carbohydrate pairs, oligomeric or heterooligomeric protein aggregates, DNA binding protein/DNA binding site pairs, RNA/protein pairs, and nucleic acid duplexes, heteroduplexes, or ligated strands, as well as any molecule which is capable of forming one or more covalent or non-covalent bonds  
15 (for example, disulfide bonds) with any portion of an RNA-protein fusion.

The term "ligand" refers to a molecule that is recognized by a particular target, such as a protein, *e.g.*, a receptor. Any agent bound by or reacting with a target is called a "ligand," so the term encompasses the substrate of an enzyme and the reactants of a catalyzed reaction. The term "ligand" does not imply any particular  
20 molecular size or other structural or compositional feature other than that the substance in question is capable of binding or otherwise interacting with a target. A "ligand" may serve either as the natural ligand to which the target binds or as a functional analog that may act as an agonist or antagonist.

The term "substrate" refers to a ligand of an enzyme which is catalytically acted on and chemically converted by the enzyme to product(s).  
25

The term "receptor" refers to a molecule that has an affinity for a given ligand.

The term "solid support" refers to a material having a rigid or semi-rigid surface. Such materials will preferably take the form of small beads, pellets, disks, chips, dishes, multi-well plates, wafers or the like, although other forms may be  
30 used. In some embodiments, at least one surface of the material will be substantially flat. The term "surface" refers to any, generally two-dimensional, structure on a solid

material and may have steps, ridges, kinks, terraces, and the like without ceasing to be a surface.

As used herein, by a "population" is meant more than one molecule (for example, more than one RNA, DNA, RNA-protein or RNA-peptidomimetic fusion molecule). Because the methods of the invention facilitate selections which begin, if desired, with large numbers of candidate molecules, a "population" according to the invention preferably means at least  $10^3$  different sequences, though more preferably at least  $10^8$  different sequences, and most preferably at least  $10^{15}$  different sequences.

The term "random peptide library" refers to a population of random or semi-random peptides, as well as a population of fusion proteins containing those random peptides (as applicable). A similar meaning is given to the term "random peptidomimetic library", but with the understanding that one or more residues of the peptidomimetic are non-naturally occurring amino acid-like moieties.

By "selecting" is meant substantially partitioning a molecule from other molecules in a population. As used herein, a "selecting" step provides at least a 2-fold, preferably, a 10-fold, more preferably, a 100-fold, and, most preferably, more than 1000-fold enrichment of a desired molecule relative to undesired molecules in a population following the selection step. As indicated herein, a selection step may be repeated any number of times, and different types of selection steps may be combined in a given approach.

The phrases "individually selective manner" and "individually selective binding", with respect to binding of a test peptide or peptidomimetic with a target, refers to binding specific for, and dependent on, the molecular identity of the target.

The language "differential binding means", as well as "affinity selection" and "affinity enrichment", refer to the separation of members of the peptide or peptidomimetic display library based on the differing abilities of test peptides or peptidomimetics on each of the display packages of the library to bind to the target. The differential binding of a target by molecules of the display can be used in the affinity separation of molecules which specifically interact with the target from those which do not. Examples of affinity selection means include affinity chromatography, precipitation, fluorescence activated cell sorting, and plaque lifts.

As described below, the affinity chromatography includes panning techniques using, e.g., immobilized target proteins.

The term "reporter group" or "tag" refers to an atom, compound, or biological molecule or complex that can be readily detected when attached to other molecules and exploited in chemical separation processes. A reporter group can be, for example, a fluorescent or radioactive atom or a compound containing one or more such atoms.

By "proofreading" activity of an aminoacyl tRNA synthetase is meant the hydrolytic catalytic activity of the enzyme that recognises and then removes non-cognate amino acids (natural or unnatural) from the aminoacylated cognate tRNA isoacceptors of the synthetase.

### (III) Exemplary Embodiments

In general, the inventive method consists of an *in vitro* or *in situ* transcription/translation protocol that generates peptides or peptidomimetics. This is accomplished by synthesis and *in vitro* or *in situ* translation of an mRNA molecule with one or more tRNA molecules that are charged with naturally and/or non-naturally occurring amino acids or amino acid analogs. We have discovered that bacterial translation can be reconstituted without added EF-P, W, W2, or rescue.

In general, a preferred minimal translation system includes the following macromolecular components: ribosomes, mRNA, aminoacyl tRNAs, and translation factors IF2H, IF3H, EF-TuI and EF-GH. IF1H and EF-Ts are stimulatory, and are often added. Additional alterations are detailed below.

The lack of availability of clones designed for high-level over-expression of tagged initiation factors has made the reproducible preparation of large quantities of highly purified factors to reconstruct a purified translation system a major technical challenge. Indeed, many studies have relied on gifts of key components. In order to overcome this problem, we subcloned and overproduced all three *E. coli* initiation factors with (His)<sub>6</sub> tags (termed his-tags) and tested the activity of each factor in a purified translation system.

*Subcloning, over-expression and purification of his-tagged E. coli translation factors.* Published clones for the expression of *E. coli* IF1, IF2 and IF3, though useful in numerous initiation studies, are for untagged factors that cannot be

affinity purified, and, with one exception, are thermally inducible (Calogero *et al.* (1987) *Mol Gen Genet* 208, 63-9; Laalami *et al.* (1991) *J Mol Biol* 220, 335-49; De Bellis and Schwartz (1990) *Nucleic Acids Res* 18, 1311; Mortensen *et al.* (1991) *Biochimie* 73, 983-9; Brombach and Pon (1987) *Mol Gen Genet* 208, 94-100). An important consideration for IF3 over-expression is the presence of the rare AUU initiation codon (Brombach and Pon (1987) *Mol Gen Genet* 208, 94-100). Initiation codon pairing with fmet-tRNA<sub>i</sub><sup>fmet</sup> is directly proofread by IF3 (Meinzel *et al.* (1999) *J Mol Biol* 290, 825-37), thereby enabling IF3-mediated feedback repression of translation of its own gene *in vivo* (Brombach and Pon (1987) *Mol Gen Genet* 208, 94-100). Thus, to increase the expression levels and to simplify purification of all three initiation factors, we replaced the initiation codons with an AUG(CAC)<sub>6</sub> sequence by PCR, and subcloned the coding sequences into a pET-derived expression plasmid (see Materials and Methods). The resulting his-tagged clones were non-toxic and overproduced the factors (termed IF1H, IF2H and IF3H) at very high levels in the soluble fraction of the lysate (Fig. 1, lanes 1,2,5,6,8,9). In addition, *E. coli* EF-Tu and EF-G with N-terminal his<sub>6</sub> tags (termed EF-TuH and EF-GH) were over-expressed using available clones (Fig. 1, lanes 12,13,15,16, (Hwang *et al.* (1997) *Arch Biochem Biophys* 348, 157-62; Semenov *et al.* (1996) *Proc Natl Acad Sci U S A* 93, 12183-8)). Gel analysis of the Ni<sup>2+</sup>-purified initiation and elongation factors is also shown in Fig. 1 (lanes 3, 7, 10, 14 and 17). IF2H, IF3H, EF-TuH and EF-GH comigrated with samples of the authentic *E. coli* proteins (data not shown), and IF1H had the expected electrophoretic mobility based on comparison with molecular weight markers (Fig. 1, lanes 1-4). IF2-2, a shorter form of IF2, was probably over-expressed together with IF2H (Fig. 1, lane 6), but, because its synthesis by internal translation initiation (Sacerdot *et al.* (1992) *J Mol Biol* 225, 67-80) would not incorporate a his<sub>6</sub> tag, it did not copurify with IF2H (Fig. 1, lane 7).

*Dependencies of his-tagged initiation factors in an initiation assay.* The activity and purity of the three initiation factors were measured by ribosome: fmet-tRNA<sub>i</sub><sup>fmet</sup>: ApUpG trinucleotide complex formation (Fig. 2, top line). Dependencies of the his-tagged initiation factors in initiation complex formation with the 3 x salt-washed ribosomes were comparable to those reported for native factors (Table 1;

Kung *et al.* (1974) *Arch Biochem Biophys* 162, 578-84; Dubnoff and Maitra (1972) *J Biol Chem* 247, 2876-83). The variation in IF3 dependence may result from different amounts of free 30S ribosomal subunits in the different 70S ribosome preparations because IF3 does not strongly stimulate initiation complex formation when free 30S subunits are substituted for 70S ribosomes in the assay (Canonaco *et al.* (1987) *Biochimie* 69, 957-63). Given that initiation factors tend to copurify with ribosomes (Kung *et al.* (1974) *Arch Biochem Biophys* 162, 578-84), the results also attest to the purity of the ribosomes. Substitution of the AUG trinucleotide template with the MTV mRNA template (Fig. 3; described below) at a much lower concentration (0.3  $\mu$ M) enables complex formation with equivalent yield (data not shown).

*Dependencies of his-tagged factors in tripeptide synthesis.* We next tested whether initiation complexes were competent to undergo elongation in a purified translation system using the components depicted in Fig. 2. The mRNA design we selected (Paylov *et al.* (1997) *Biochimie* 79, 415-22) contains an optimal Shine-Dalgarno sequence and an epsilon translational enhancer (Olins and Rangwala (1989) *J Biol Chem* 264, 16973-6). Our design (Fig. 3) has the advantage of encoding mRNAs short enough to be synthesized directly from a single long synthetic DNA template hybridized to a standard 18-mer oligodeoxyribonucleotide without the need for cloning (Milligan and Uhlenbeck (1989) *Methods in Enzymology* 180, 51-62). Translations of the templates are completed when the ribosomes translocate to a codon for which there is no supplied cognate aminoacyl tRNA (Weissbach *et al.* (1984) *BioTechniques* 2, 16-22).

As shown in Fig. 4, reconstitution of oligopeptide synthesis with all of the components shown in Fig. 2 using 3 x salt-washed ribosomes and the MTV mRNA (Fig. 3) results in the synthesis of tMTV, as judged by the incorporation of  $^3$ H-labeled valine into the isolated peptide products. When the time course of synthesis is begun by combining a preincubated initiation mix with a preincubated elongation factor mix (triangles in Fig. 4; see legend), there is a rapid initial burst of product synthesis within the first minute and a slower rate of synthesis at steady state. Without the 10 minute preincubations (squares in Fig. 4; see legend), there is little

product synthesis within the first minute, and synthesis is fairly linear with time (see also Fig. 7).

Fig. 5 illustrates the dependence on each factor necessary for *in vitro* translation (IF1 was omitted in this experiment), based on the incorporation of <sup>14</sup>C-labeled threonine and <sup>3</sup>H-labeled valine into peptide products. The complete system yielded peptide products with a threonine/valine (TV) ratio of 1.0, as expected for fMTV synthesis. Omission of any one factor dramatically reduces fMTV synthesis (Fig. 5), giving dependencies comparable to those reported for the native factors (Robakis *et al.* (1981) *Proc Natl Acad Sci U S A* 78, 4261-4; Ceniempo *et al.* (1982) *Arch Biochem Biophys* 218, 572-8). The omission of translation factor EF-GH switched translation from tripeptide synthesis to fMT dipeptide synthesis (Fig. 5), consistent with the known primary role of EF-G in translocation. Although synthesis of dipeptides does not require addition of IF1 (Robakis *et al.* (1981) *Proc Natl Acad Sci U S A* 78, 4261-4), and although the inclusion of IF1H in our translations does not lead to a dramatic increase in overall yield, IF1H does stimulate the rate of fMTV synthesis 2.5 fold during the first few minutes of translations performed without preincubation (data not shown), consistent with previously reported studies with native IF1 (Robakis *et al.* (1981) *Proc Natl Acad Sci U S A* 78, 4261-4). Thus, the translation results confirm and extend the findings from the initiation assay (Table I). Additional control translations omitting other macromolecular components of the translation system give the expected dependencies (see Materials and Methods). Ribosomes washed only once gave poor factor dependencies in translation assays of the type illustrated in Fig. 5 (data not shown).

**Tetrapeptide synthesis.** We next investigated the suitability of this system for the synthesis of tetrapeptides to show that all steps of initiation and elongation could occur in a highly purified system. In contrast to tripeptide synthesis, tetrapeptide synthesis is not possible without dissociation of deacylated tRNA from the exit site of the ribosome (Wilson and Noller (1998) *Cell* 92, 337-49). Using a template encoding the tetrapeptide fMTTV (Fig. 3), synthesis of dual-labeled products was assessed by reversed phase HPLC (Fig. 6). Radioactive peptide products were identified based on their co-migration with chemically synthesized

non-radioactive standards. The predominant radioactive peak in the peptidyl separation range corresponds to the fMTTV tetrapeptide (80-85% of the  $^{14}\text{C}$  or  $^3\text{H}$  radioactivity in this range) with a T/V ratio close to that expected. Two minor peaks correspond to the pausing or premature termination products fMT and fMTT, with no  $^3\text{H}$  incorporated, as expected. fMT and fMTT together contain 12% of the  $^{14}\text{C}$  radioactivity in the peptidyl range, equivalent to 20% of the combined products fMTTV, fMTT and fMT on a molar basis. The remaining two minor peaks (at 24 and 50 min.) presumably correspond to derivatives of fMTTV (*e.g.* methionine oxidation products or unformylated peptide), other peptidyl products and/or non-peptide radioactive contaminants. Thus, the purified his-tagged tetrapeptide translation system is predominantly, but not completely, processive with yields of full-length tetrapeptide products equal to 80% of the peptide products. This detailed analysis demonstrated much higher processivity in absence of added EF-P. W and rescue than that previously reported (Ganoza *et al.* (1985) *Proc Natl Acad Sci U S A* 82, 1648-52).

Further examples of peptides synthesized efficiently by our *in vitro* translation system are shown in the Table 2. The synthesis of the expected 7-mer full-length fMTTTTV peptide product was highly processive because a predominance of premature termination products would have resulted in a much higher T/V ratio than the measured value of 6.4.

*Effects of upstream mRNA mutations on oligopeptide synthesis.* An unresolved question in translation initiation is the influence of the epsilon sequence in a purified system. Therefore, the effect of scrambling or deleting epsilon was determined using the mRNAs scramble-epsiMVT and delta-epsiMVT (Fig. 3) and template concentrations (1  $\mu\text{M}$ ) that are saturating for mRNA MVT (data not shown) under conditions where initiation should be rate-limiting (see Fig. 4). Fig. 7 shows that scrambling of the U-rich epsilon sequence results in decreased fMVT synthesis, and deletion of the sequence results in a five-fold decrease in the initial rate of product synthesis. Concentrations of mRNA delta-epsiMVT up to 14 fold higher failed to substantially increase the yield of fMVT (data not shown).

*Incorporation and selection of an unnatural amino acid.* Our tetrapeptide synthesis format is directly amenable to many types of initiation and elongation

assays, including the testing of unnatural amino acids for incorporation by ribosomes for mechanistic or selection experiments. For example, translation of mRNA MMTKV (Fig. 3) using the substrates  $^3\text{H}$ -fmet-tRNA<sup>fmet</sup>,  $^{14}\text{C}$ -thr-tRNA<sup>thr</sup>, biotin-labeled-lys-tRNA<sup>lys</sup> (Promega) and  $^3\text{H}$ -val-tRNA<sup>val</sup> yielded peptide product containing both biotin and  $^3\text{H}$ -valine (M-T-bK-V), as judged by product copurification with soft avidin beads (Promega) in a manner dependent on a lysine codon in the mRNA and on biotin-labeled-lys-tRNA<sup>lys</sup> (Fig. 8). This experiment demonstrates that the purified translation system is capable of incorporating an easily selectable, large, unnatural amino acid.

Given that the  $K_d$  of the biotin-avidin interaction ( $10^{-15}$  M) is one of the lowest known for any ligand bound non-covalently and monovalently to its target, it is expected that our synthesized biotin-containing peptidomimetic has an affinity for one of its cognate targets (avidin) that is greater than the affinity of any natural peptide of similar length for its cognate target. Given that the identification of very high affinity ligands is most desirable in library screening experiments, the potential advantage of a ligand-screening method that can use unnatural amino acids over a method that can only use natural amino acids is illustrated by the above example of our peptide containing biotinyl-lysine.

An analogous experiment to that of Fig. 8 has also been performed with biotinyl-Cys-tRNA<sup>Cys</sup>, except that, in this case, the biotinyl-aminoacyl-tRNA species used had been purified by alternative procedures. Pure tRNA<sup>Cys</sup> isoacceptor (Subriden RNA) was charged with Cys using a tRNA-free crude synthetase extract, the resulting Cys-tRNA<sup>Cys</sup> was chemically labelled with biotin (Fig. 9) by minor modification of a published method (Ohtsuka et al. (1997) *Nucleic Acids Symp.* 37, 125-126), biotinyl-Cys-tRNA<sup>Cys</sup> was gel-purified (Fig. 10), and then biotinyl-Cys (abbreviated bC) was incorporated into peptidomimetics and bound to Soft Avidin (Figs. 11 and 12). The results in Fig. 12 suggest that our *in vitro* system may be able to incorporate specifically two adjacent large unnatural amino acids.

Importantly, the translation with uncharged tRNA<sup>Cys</sup> in Fig. 11 gave no incorporation of  $^3\text{H}$ , demonstrating that the purified system lacked charging activity and was specific for exogenous substrate. This has also been demonstrated for translations "seeded" with uncharged tRNA<sup>Thr</sup> or uncharged tRNA<sup>Val</sup> (not shown), as

5 expected for an aminoacyl-tRNA-synthetase-free pure system. Given that the synthetases are amongst the most abundant proteins of *E. coli* and that ribosome preparations can be used for their purification (Ganoza *et al.* (1996) *Biochemie* 78, 51-61), additional experiments were performed to verify that our pure translation system lacked synthetase activity. Fig. 13 confirmed that our "protein polymerase" lacked contaminating aminoacyl-tRNA synthetases.

*Purification of peptide and peptidomimetic products.* Several alternative methods (either non-denaturing or denaturing) standard in the art are available for the release of free peptides or peptidomimetics from peptidyl-tRNA, including, but not limited to, chemical hydrolysis with base (*e.g.* Figs. 4-8), enzymatic catalysis by release factors (Table 3) or peptidyl-tRNA hydrolase enzyme (purified and used as in Karimi *et al.* (1998) *J. Mol. Biol.* 281, 241-252) and nucleophilic attack by the puromycin antibiotic or its derivatives (Fig. 21). Any of the numerous methods (either non-denaturing or denaturing) standard in the art for peptide or protein purification can be used for purification of peptide and peptidomimetic products, including, but not limited to, affinity purification using a solid support (*e.g.* using the interaction between soft avidin beads and biotin in Figs. 8, 11 and 12), chromatography (*e.g.* using cation exchange chromatography in Figs. 4 and 5, or reversed phase HPLC in Fig. 6), or precipitation from solution (*e.g.* using TCA in Fig. 20).

20 Despite the high demonstrated purity of our translation system (Table 1, Figs. 1, 5, 11, and 13) and the apparent dispensability of EF-P, W and rescue for processive translation, additional experiments were performed to rule out the argument that one or more of these factors contaminates the ribosomes (see Background of the Invention). The published method for removing EF-P, W and rescue from ribosomes (Green *et al.* (1985) *Biochem. Biophys. Res. Comm.* 126, 792-798) is to pellet the ribosomes up to five times with an overnight wash in high salt (0.5 - 1 M) between each centrifugation (the more salt washes, the higher the purity of the ribosomes). Since our ribosomes had been pelleted four times with three high salt washes, our ribosomes should have a comparable purity to Ganoza's. Nevertheless, a different batch of ribosomes was prepared that incorporated an additional (fourth) high salt wash (the same total number as Ganoza's most highly

washed ribosomes), and, in addition, the final 150 000 g centrifugation was split into two steps to remove easily pelleted material. In the first step, centrifugation was performed at the maximum speed for one minute, and the resulting pellet was discarded (this material was predominantly non-ribosomal based on a low absorbance at 260 nm). In the second step, the ribosomes from the supernatant were pelleted by extended centrifugation. The purity of these 4 x washed ribosomes was demonstrated to be very high based on strong dependencies on translation factors for initiation and translation (assayed as in Table 1 and Fig. 5). As an additional control, all of our translation factors, which had previously been purified by Ni-affinity chromatography, were further purified individually by an additional gel-filtration chromatography step. The combination of these 4 x washed ribosomes and Ni/gel-purified factors was found to be active and processive in synthesis of peptides as long as 101 amino acids (see Fig. 20 below), and the additional purification steps do not significantly affect our translations in comparison with the 3 x washed ribosomes and Ni-purified factors. These studies confirm that our translation system is not dependent on EF-P, W and rescue.

Other embodiments of the invention are versions of our simplified purified translation system lacking EF-P, W and rescue in which one or more of the components have been substantially adjusted in concentration or even omitted entirely. For example, PEG is often omitted, resulting in only an approximately 25% decrease in yield, and efficient translation occurs without IF1 (Table 2; Fig. 5). Inclusion of our other standard translation initiation and elongation factors, although important for efficiency under commonly used conditions, is not essential for product synthesis (Fig. 5). Indeed, efficient translation in model systems is possible without any of the bacterial initiation factors (*i.e.* IF1, IF2 and IF3) if they are substituted by higher concentrations of cations (*e.g.*  $Mg^{2+}$  or polyamines; Wagner *et al.* (1982) *Eur. J. Biochem.* 122, 193-197).

Still further embodiments of the invention are versions of our simplified purified translation system described in detail above in which one or more of other purified macromolecules and small molecules known to be involved in, or to stimulate, translation have been added. These include, but should not be limited to, cellular total tRNA or fractions thereof, cellular total aminoacyl-tRNA or fractions

thereof, synthetic charged or uncharged tRNAs, one or more of the aminoacyl-tRNA synthetases for each of the twenty natural amino acids, Met-tRNA<sub>Met</sub>, formyltransferase (also called methionyl-tRNA transformylase), N<sup>10</sup>-formyl THF synthetase and TIF derivatives, elongation factor Ts (see Materials and methods), release factors (RF1, RF2, RF3, and RRF or RF4), DNA templates and RNA polymerases for coupled transcription and translation, RNA helicases, chaperones (Hardesty *et al.* (1999) *Curr Opin Struct Biol* 9, 111-4), ribosomes purified by different procedures (including separation into subunits) such as sucrose-density-gradient-centrifugation, components of "polymix buffer" including polyamines (Jelenc and Kurland (1979) *PNAS* 76, 3174-3178) and energy-related systems that differ from our pyruvate kinase system such as those with creatine kinase, myokinase and/or pyrophosphatase (Shimizu *et al.* (2001) *Nat. Biotech.* 19, 751-755). Addition of, or substitution with, untagged or mutated versions of natural components, such as our recombinant untagged IF1, IF2 and IF3 (see Materials and Methods) or an EF-Tu derivative with improved incorporation of unnatural amino acids, or altered ribosomes is also possible in a pure translation system.

For example, our simplified purified translation system, such as that of Figs. 6 and 11, is stimulated by addition of bacterial release factors (Table 3). The addition of synthetases to our translations, such as those for Thr and Val, together with amino acids Thr and Val and ATP, enabled generation of, and regeneration of, the respective aminoacyl-tRNAs from the respective uncharged tRNAs (not shown). The addition of DNA template, NTPs and RNA polymerase enabled coupled transcription and translation (Table 2). Addition of other molecules, such as polyethylene glycol, also proved stimulatory (not shown). Total aminoacyl tRNA, isolated from cells by acid phenol extraction (Varshney *et al.* (1991) *J. Biol. Chem.* 266, 24712-24718; see also Fig. 16, lane 2 below), is also active in our purified translations. An alternative approach for substrate preparation is to charge deacylated total tRNA *in vitro* before translation (Green *et al.* (1985) *Biochem. Biophys. Res. Comm.* 126, 792-798). The latter approach has the advantage of being more readily suitable for the selective depletion or inactivation of certain tRNA isoacceptors (*e.g.* using isoacceptor-specific DNA oligos and RNase H (Kanda *et*

*et al.* (1998) *FEBS Lett.* 440, 273-276)) to allow incorporation of radiolabelled or unnatural amino acids charged on appropriate tRNAs.

Six months after our Provisional Patent Application, further evidence was published (Shimizu *et al.* (2001) *Nat. Biotech.* 19, 751-755; incorporated by reference herein) for the dispensability on EF-P, W and rescue in a purified translation system. Their version of our system, in contrast to our preferred system, contained aminoacyl-tRNA synthetases. Translation was reconstituted efficiently with recombinant versions of all of the well-characterized translation factors and synthetases, without added EF-P, W, W2 and rescue. However, the ribosomes were prepared by a different method from Ganoza's and ours (using less salt washing), and the dependencies on three of the twenty different synthetases were incomplete, raising the possibility of contamination by EF-P, W, W2 and rescue. Nevertheless, strong dependencies on most of the factors were reported, as was efficient synthesis of several proteins.

A stop codon was recruited to incorporate valine efficiently using a chemically charged suppressor tRNA mutated to avoid synthetase recognition, but, contrary to claims in the publication, incorporation of an unnatural amino acid was not tested (Shimizu *et al.* (2001) *Nat. Biotech.* 19, 751-755). Potentially, the incorporation of an unnatural amino acid by a suppressor tRNA in this version of our system has the advantage that competition with certain termination factors can be circumvented by simply omitting them. However, like existing systems for unnatural amino acid incorporation that use crude extracts (see Background of the Invention), this strategy is likely to be restricted to incorporation of a single type of unnatural amino acid per protein at only one of the three termination codons (the UAG codon) because of competition from natural amino acids at sense codons catalysed by the tRNA charging and proofreading activities of the twenty different added aminoacyl tRNA synthetases, and because an attempt to use a second termination codon (UGA) failed due to readthrough by the ribosome (Cload *et al.* (1996) *Chem. and Biol.* 3, 1033-1038). Potentially, to allow incorporation of multiple different unnatural amino acids in such a system, the suppressor anticodon could be mutated to recognise certain sense codons (by base pairing), and competition at those codons potentially could be circumvented by omitting the

cognate natural amino acids and synthetases. However, for most of the twenty synthetases, the anticodons of their cognate tRNA isoacceptor substrates are the most important recognition elements, so alterations in the anticodon of the suppressor tRNA are likely to lead to unwanted synthetase recognition (sometimes with an unpredictable specificity of cross-recognition) and therefore proofreading and/or charging with a natural amino acid (Giege *et al.* (1998) *Nucleic Acids Res.* 26, 5017-5035). In addition, tRNA mutations designed to prevent synthetase recognition may also decrease the efficiency of EF-Tu and/or ribosome recognition, thereby decreasing the efficiency of incorporation of the carried amino acid. Because of these difficulties with synthetase recognition and also the difficulties associated with over-expressing as many as twenty different synthetase proteins (Swartz (2001) *Nat. Biotech.* 19, 732-733), we preferred a synthetase-free translation system for our tests with unnatural amino acids.

We have tested incorporation of several unnatural amino acids using chemically charged tRNAs. The first step was to construct a synthetic elongator tRNA lacking the terminal CA dinucleotide to allow chemical misacylation with an unnatural amino acid in a generalisable manner. Current technology relies on artificial suppressor tRNAs that have been specially engineered to prevent charging and proofreading by any of the synthetases. In our pure system, our only concern was the possible effects of an expected lack of tRNA base-modifying activities because such modifications can be important for function (Bjork *et al.* (1999) *FEBS Lett.* 452, 47-51), and crude translation systems can modify synthetic tRNAs (Claesson *et al.* (1990) *FEBS Lett.* 273, 173-6). As new test prototypes, we chose *E. coli* tRNA<sup>Asp</sup> (Fig. 14; Ohashi *et al.* (1976) *Nucleic Acids Res.* 3, 3369-3376) and tRNA<sup>Ala</sup> (discussed below; Picking *et al.* (1991) *Nucleic Acids Res.* 19, 5749-5754). The 5' terminal sequence of tRNA<sup>Asp</sup> was mutated for optimal transcription by T7 RNA polymerase, although alternative strategies to mutagenesis exist, such as the use of M1 RNA or RNase P to process synthetic unmodified tRNA precursors (Forster and Altman (1990) *Science* 249, 783-786). The anticodons of both tRNAs were also mutated to create several variants with altered codon recognition properties (three of our tRNA<sup>Asp</sup> mutants are shown in Fig. 15, with the amino acid codons recognised by the tRNAs indicated in brackets). An unnatural amino acid,

allylglycine (aG, sometimes alternatively abbreviated 2P), was amino-protected with an NVOC group and ligated onto the tRNA<sup>Asp</sup>(N) (Fig. 14) using T4 RNA ligase in a standard and generalizable strategy (see Materials and Methods; Thorson *et al.* (1988) *Methods in Molecular Biology* 77, 43-73; Steward and Chamberlin (1998) *Methods in Molecular Biology* 77, 325-354) to give a species that migrated on a gel with the expected mobility (Fig. 16, lane 6).

The amino group of the NVOCaG-tRNA<sup>Asp</sup>(N) was deprotected by ultraviolet photolysis, and the aG-tRNA<sup>Asp</sup>(N) added to a pure translation reaction containing mMTNV template. aG was successfully incorporated at the N codon to allow downstream <sup>3</sup>H-V incorporation, but gave a lower yield than the mMTV control translation (Fig. 17). The lower yield was found to be predominantly due to using insufficient aG-tRNA<sup>Asp</sup>(N) to saturate the system, as doubling the aG-tRNA<sup>Asp</sup>(N) RNA concentration to 1 μM (double the concentration of each natural aminoacyl tRNA in the translation, assuming 100% yield of aG-tRNA<sup>Asp</sup>(N) from the transcription and ligation reactions) doubled the yield and incorporated the analog with about 65% processivity (based on incorporation of <sup>14</sup>C-T before the analog and <sup>3</sup>H-V after; result not shown). Translation of mMTNNV demonstrated incorporation of adjacent analogs, with the % processivity for incorporation of the 2<sup>nd</sup> analog similar to the first under these analog substrate-limited conditions. Note that the translations of mMTNV and mMTNNV with an uncharged tRNA having the same structure as the tRNA<sup>Asp</sup>(N) of Fig. 14, except ending with the normal ubiquitous CCA 3' sequence (produced by run-off transcription of BstNI cut template) gave no incorporation, demonstrating that the purified system lacked charging activity for this mutant tRNA<sup>Asp</sup>(N), and was specific for incorporation of exogenous substrate at the N codon. As an additional control, uncharged tRNA<sup>Asp</sup>(N) did not inhibit translation of mMTV (Fig. 17).

Having established that sense codons could be commandeered in a generalisable manner for selective unnatural amino acid incorporation, the next step was to optimise the efficiency of incorporation of aG and also other unnatural amino acids, so the NVOC amino-protected forms of O-methylserine (mS) and alanine were prepared (see Materials and Methods). Using higher concentrations (up to 7 μM) of various tRNA<sup>Asp</sup>s (Figs. 14 and 15) chemically aminoacylated with aG,

mS, or Ala, incorporation efficiencies of greater than 90% of that obtained with Thr-tRNA<sup>Thr</sup> were obtained (not shown).

Interestingly, the unmodified aG-tRNA<sup>A(N)</sup>, which has a GUC substitution in its anticodon in comparison with the published synthetic tRNA<sup>Ala</sup> (Picking *et al.* (1991) *Nucleic Acids Res.* 19, 5749-5754), was totally inactive in translations. However, charging with Ala using a crude tRNA-free cell extract produced Ala-tRNA<sup>A(N)</sup> that was incorporated efficiently. The gain of activity was probably not due to the difference in amino acid because aG worked well on tRNA<sup>Asn(N)</sup>, but more likely due to the base-modifying activity of the crude extract. The results are consistent with our hypothesis (above) that tRNA base modifications are crucial for translation activity of certain tRNAs and that initial selection of an appropriate prototype tRNA is a non-obvious process requiring trial and error. Another problem is that certain unmodified aminoacyl tRNAs are rendered inactive by denaturation during standard purification procedures (Harrington *et al.* (1993) *Biochemistry* 32, 7617-7622). Nevertheless, once a stable active unmodified tRNA is identified, such as our tRNA<sup>A(N)</sup>, the production of anticodon variants that recognise other codons in a predictable manner is straightforward (Fig. 15; see below).

Table 4 shows a representative incorporation of five successive unnatural amino acids (aG's) within a 7-mer peptidomimetic product in a manner completely dependent on added aG-tRNA<sup>Asn(T)</sup>. This improves upon the published record of two successive unnaturals (Hohsaka *et al.* (1999) *JACS* 121, 12194-12195), not only in length, but more importantly in generalizability and utility. Given enough different amino acid monomers, a library of random 7-mer peptidomimetics can encode substantial peptidomimetic diversity. Using the four different unmodified tRNA<sup>Asn</sup> variants shown in Figs. 14 and 15, chemically aminoacylated with mS for the T codon and aG for the N, S, and V codons, full-length product was synthesized based on incorporation of carboxy-terminal <sup>3</sup>H-E (Table 5). This provides evidence for incorporation of multiple and different unnatural amino acids in a single product. The complete dependency on added mS established that mS was indeed incorporated, and that this T codon is specific for tRNA<sup>Asn(T)</sup> (*i.e.* not misread by the closely related anticodon variants tRNA<sup>Asn(N)</sup>, tRNA<sup>Asn(S)</sup>, and tRNA<sup>Asn(V)</sup> shown in Figs. 14 and 15).

Our strategy of sense codon replacement with unnatural amino acids is a generalisable one and should therefore be extendable to all 61 sense codons. Given the significant redundancy in codon usage by nature (only 21 natural amino acids), it should be possible to create an aminoacyl tRNA library that can translate a single mRNA into substantially more than 21 different amino acids. The three termination codons are dispensable in our system, but potentially all could be translated with suppressors, with potential readthrough problems (Cload *et al.* (1996) *Chem. and Biol.* 3, 1033-1038) overcome by omission of the natural aminoacyl tRNA(s) that are responsible for readthrough. Effective replacement of the AUG initiation codon and fMet initiator amino acid is subject to maintenance of a different set of intermolecular contacts from elongation, but it is also possible based on known allowable substitutions *in vivo* and *in vitro* (Picking *et al.* (1991) *Nucleic Acids Res.* 19, 5749-5754; Wu and RajBhandary (1997) *J. Biol. Chem.* 272, 1891-1895).

*A. In vitro libraries*

The directed evolution *in vitro* of peptides or peptidomimetics from a combinatorial library of peptide analogs expressed on ribosomes is a powerful method for the rapid evolution of ligands or drug candidates that bind to any target molecule of choice. An example of our "pure ribosome display" (Fig. 18), related to ribosome display performed in crude extracts (Matheakis and Dower (1995) PCT WO 95/11922) consists of the following steps and variants thereof:

**Construction of DNA library:** a synthetic oligodeoxyribonucleotide containing a randomized or partially randomized coding sequence is chemically synthesized, hybridized to an oligodeoxyribonucleotide containing an RNA polymerase promoter (*e.g.* a bacteriophage RNA polymerase), translation initiation sequence and start codon, extended with DNA polymerase, ligated by DNA ligase to a plasmid restriction fragment encoding an open reading frame (*e.g.* a repetitive spacer amino acid sequence such as our longer sequences shown in Fig. 19), purified on a gel, and quantitated.

**Step 1** (Fig. 18): The synthetic DNA library is transcribed with an RNA polymerase into an mRNA library.

**Step 2:** The mRNA library is translated into peptide analogs with our purified ribosomes and translation factors and a reconstituted aminoacyl-tRNA pool

containing a mixture of certain (but not all) wild-type aminoacyl-tRNAs and specially synthesized tRNAs charged with amino-acid analogs. Preferably, each mRNA codon is decoded by a unique charged tRNA in the pool, so the sequence of the mRNA defines a unique peptide analog sequence. For example, our spacer sequences of Fig. 19, some of which are long enough to traverse the ribosome tunnel, have been synthesized with good processivity (Fig. 20). Protein synthesis is stopped (e.g. by reaching the end of a template lacking stop codons, or by omitting release factors, or by reaching a codon for which there is no supplied aminoacyl tRNA (such as an Asn codon in the case of the longest products in Figs. 19 and 20), or by stalling with an antibiotic.

**Step 3:** Ribosomes containing the random peptide analogs are isolated by centrifugation and then incubated with an immobilized target, and unbound ribosomes are washed away. Bound ribosomes are dissociated (e.g. with EDTA), and the released selected mRNA purified.

**Step 4:** cDNA synthesis and PCR amplification. Error-prone PCR may be used to introduce mutations.

**Repetition of steps 1-4:** These selection and amplification steps are repeated one or more times, as necessary (see below).

**Step 5 (not shown):** Amplified DNA is cloned into a plasmid for analysis by DNA sequencing to deduce the structures of selected and amplified peptidomimetics. The identification of a selected consensus sequence(s) is evidence that sufficient rounds of reiterative selection and amplification (steps 1-4) have been employed and that ligands have been identified by the experiment. Further cycles can be used to evolve ligands with higher affinities.

**Modification of the method that enables a selection step in the absence of ribosomes:** The DNA template is modified to encode a peptide tag that, when introduced into the peptide analog during protein synthesis, has a high affinity for mRNA or a molecule bound to the mRNA (e.g. a hybridized complementary DNA primer covalently linked to an antibody which has a high affinity for the peptide tag). The peptide-mRNA could then be separated from ribosomes before the selection step. These approaches and other possible modifications have been described (Mattheakis and Dower, *supra*; Doi and Yanagawa (1999) *FEBS Lett.*

457, 227-230). Alternatively, the mRNA could be 3'-end-labelled with puromycin so that it could be directly linked to the peptide for "pure mRNA display" (Fig. 21; see below).

The Szostak *et al.* PCT publications WO00/047775 and WO98/31700 (incorporated by reference herein) describe methods which can be readily adapted in the present invention in order to generate forms of the subject peptides and peptidomimetics which are covalently linked to the RNA molecule by which they are encoded. That is, the present invention provides a protocol that generates peptidomimetic covalently linked to the 3' end of its own mRNA, *i.e.*, an RNA-peptidomimetic fusion.

This is accomplished by synthesis and *in vitro* or *in situ* translation of an mRNA molecule with a peptide acceptor attached to its 3' end. One preferred peptide acceptor is puromycin, a nucleoside analog that adds to the C-terminus of a growing peptide chain and terminates translation. In one preferred design, a DNA sequence is included between the end of the message and the peptide acceptor which is designed to cause the ribosome to pause at the end of the open reading frame, providing additional time for the peptide acceptor (for example, puromycin) to accept the nascent peptide chain before hydrolysis of the peptidyl-tRNA linkage (Fig. 21).

If desired, the resulting RNA-peptidomimetic fusion allows repeated rounds of selection and amplification because the coding sequence information may be recovered by reverse transcription and amplification (for example, by PCR amplification as well as any other amplification technique, including RNA-based amplification techniques such as 3SR or TSA). The amplified nucleic acid may then be transcribed, modified, and *in vitro* or *in situ* translated to generate mRNA-peptidomimetic fusions for the next round of selection. The ability to carry out multiple rounds of selection and amplification enables the enrichment and isolation of very rare molecules, *e.g.*, one desired molecule out of a pool of  $10^{15}$  members. This in turn allows the isolation of new or improved peptides and peptidomimetics which specifically recognize virtually any target or which catalyze desired chemical reactions.

Accordingly, in one aspect, the invention features a method for selection of a desired protein or peptidomimetic, involving the steps of (a) providing a population of candidate RNA molecules, each of which includes a translation initiation sequence and a start codon operably linked to a candidate protein coding sequence and each of which is operably linked to a peptide acceptor at the 3' end of the candidate protein coding sequence; (b) *in vitro* or *in situ* translating the candidate protein coding sequences in the presence of natural and/or non-naturally occurring amino acids to produce a population of candidate RNA-peptidomimetic fusions; and (c) selecting a desired RNA-peptidomimetic fusion, thereby selecting the desired peptidomimetic.

In preferred embodiments of the above methods, the population of candidate RNA molecules includes at least  $10^2$ , preferably, at least  $10^5$ , more preferably,  $10^{10}$ , or as many as  $10^{15}$  different RNA molecules; importantly, the *in vitro* translation reaction is preferably carried out in a reconstituted purified mixture, not a crude translation system; the selection step involves binding of the peptidomimetic to an immobilized binding partner or assaying for a functional activity of the peptidomimetic.

In another related aspect, the invention features kits for carrying out any of the selection methods described herein.

In a final aspect, the invention features a microchip that includes an array of immobilized peptidomimetics of the present invention.

#### *B. Target Molecules*

The target molecule can be virtually any molecule for which interaction with a peptide or peptidomimetic of the present invention may be useful. In certain embodiments, the target molecule is a biopolymer, such as a nucleic acid (DNA or RNA), a protein, a lipid, a carbohydrate or the like.

In choosing a polypeptide screening target, factors which can be considered include solubility, peptide chain length, requirement of post-translational modifications, or addition of co-factors, and/or monomeric or oligomeric nature of protein(s) upon which the target is based. In general, it will be desirable that the polypeptide target be soluble, partially purified or pure, and immobilized on a solid support by methods standard in the art.

Accordingly, the present invention contemplates screening targets which correspond to (e.g. include) such domain structures as: SH2 domains; SH3 domains; ankyrin-like repeats; WD40 motifs; Kunitz-type inhibitor domains; growth factor-like domains such as EGF-like domains; Kringle domains; fibronectin finger-like domains; heparin-binding domains; death domains; TRAF domains; pleckstrin  
5 homology (PH) domains; ITAMs; catalytic domains such as kinase domains; phosphatase domains; phospholipase domains; guanine nucleotide exchange factor (GEF) domains; and hydrolase domains (such as protease domains); or DNA binding domains such as leucine zippers, zinc fingers and helix-loop-helix motifs.

10 Where the protein of interest is a transmembrane protein, the screening target can be derived from a soluble extracellular or cytoplasmic domain. To illustrate, the screening target can correspond to the extracellular domain of a guanylyl cyclase, a cytokine receptor, a tyrosine kinase receptor, or a serine/threonine kinase receptor. In other embodiments, the screening target can correspond to a soluble portion of a  
15 G-protein coupled receptor (GCR) which retains ligand binding activity. For example, as described above, certain of the extracellular loops between the transmembrane portions of the GCRs have been shown to retain ligand binding activity even when provided free in solution. In still other embodiments, the screening target can be reconstituted in a lipid bilayer, such as a liposome or other  
20 vesicle (see, for example, Kalva Kolanu *et al.* (1990) *Biotecniques* 11:248; and The Huang U.S. Patents 4,957,735 and 4,708,933) and the lipid/protein combination used as the screening target.

Merely for purposes of illustration, the following protein targets are described for use in the subject method.

25 In one embodiment, potential therapeutic targets are receptors from the *neu* receptor family. In women in the U.S.A., breast cancer is the most common cancer and is only second to lung cancer in the number of cancer deaths. A prime breast cancer target is *neu/erbB-2/HER-2*, a 185 kD trans-membrane phosphoglycoprotein tyrosine kinase (Shih *et al.* (1981) *Nature* 290:261). Amplification or over  
30 expression of the *neu* oncogene occurs in about 30% of breast and ovarian adenocarcinomas, a finding that correlates with a poor response to primary therapy (Slamon *et al.* (1987) *Science* 235:177; and Hayes *et al.* (1993) *Annals of Oncology*

4:807). Transfection of NIH 3T3 cells with the neu oncogene results in transformation (Shih *et al.*, *supra*), and introduction of an activated neu oncogene into mice results in the transformation of the entire mammary epithelium (Muller *et al.* (1988) *Cell* 54:105) or the stochastic appearance of mammary tumors (Bouchard *et al.* (1989) *Cell* 57:931).

Evidence is accumulating that breast cancer may be inhibited by molecules that bind to neu or a ligand of neu, such as members of the heregulin family (Holmes *et al.* 1992 *Nature* 256:1205). MAbs and their radiolabelled conjugates that bind to the extracellular domain of neu retard the growth of breast cancer cells in culture and in nude mice without the selection of neu-negative cell clones (DeSantes *et al.* (1992) *Cancer Res.* 52:1916; and Drebin *et al.* (1986) *PNAS* 83:9129). Such MAbs and their derivatives, such as the recombinant humanised MAb, Herceptin, that is used in the clinic for advanced breast cancer treatment (Colomer *et al.* (2001) *Cancer Investigation* 19, 49-56), may alter the neu signal transduction pathway and affect tumor growth in several different ways. They may (i) over stimulate neu, thereby causing differentiation (Bacus *et al.* (1992) *Cancer Res.* 52:2580), (ii) prevent homo- or hetero-dimerization of neu, thereby inactivating neu (Caraway *et al.* (1994) *Cell* 78:5), (iii) cause cellular internalization and down regulation of neu (Tagliabue *et al.* (1991) *Int. J. Cancer* 47:93), (iv) deliver conjugated cytotoxic radionuclides or toxins to the cell surface or cytoplasm, or (v) prevent binding of a ligand to neu.

The peptidomimetics which can be derived by the present invention can be useful for inhibiting the biological function of neu by, for example, competitively disrupting the binding of neu with its ligand or other protein, or preventing allosteric activation of an enzymatic activity associated with neu. Alternatively, they may be useful as agonists causing over stimulation or down-regulation of neu.

Yet another potential target is Interleukin-8 (IL-8). IL-8 is a chemoattractant and activator of neutrophils, and has been implicated in a wide range of acute and chronic inflammatory diseases (Murphy (1994) *Annu. Rev. Immunol.* 12:593-633). Human IL-8 is a 72-amino-acid-long polypeptide produced by monocytes, fibroblasts, keratinocytes and endothelial cells upon induction by factors such as tumor necrosis factor, interleukin-1, and lipopolysaccharides (Murphy, *supra*).

Certain analogs of IL-8 act as IL-8 antagonists *in vitro* by inhibiting neutrophil activation (chemotaxis, exocytosis and respiratory burst), suggesting that anti-IL-8 agents may have therapeutic potential for inflammatory diseases (Moser *et al.* (1993) *J. Biol. Chem.* 268: 7125-7128).

5       The monomeric IL-8 peptide forms dimers *in vitro* with a  $K_d$  of 20 $\mu$ M (Paolini *et al.* (1994) *J Immunology* 153: 2704; and Burrows *et al.* (1994) *Biochemistry* 33:12741-12745), so it is possible that the monomer and/or the dimer are active *in vivo*. Mutants that cannot dimerize are active in functional assays *in vitro* (Rajaratnam *et al.* (1994) *Science* 264:90). Interestingly, NMR and X-ray  
10       determination of the three-dimensional structure of the IL-8 dimer (Cloue *et al.* (1990) *Biochemistry* 29:1689-1696; and Baldwin *et al.* (1991) *PNAS* 88:502) revealed that it resembled the peptide-binding groove of MHC class I and II proteins (Bjorkman *et al.* (1987) *Nature* 329:506), so it is conceivable that IL-8 dimers may be able to bind to *in vitro*-selected peptide or peptidomimetic sequences in a manner  
15       similar to the MHC molecules. Therefore, the IL-8 dimer, in addition to the monomer, is an attractive target.

      In the case of the IL-8 pathway, an alternative target from IL-8 is a functional fragment of the IL-8 receptor. Even small fragments of the human and rabbit IL-8 Type 1 receptor of 39 and 44 amino acids, respectively, are functional in  
20       IL-8 binding assays (Gayle *et al.* (1993) *J Biol Chem* 268:7283-7289). Thus, members of the largest receptor family, the seven transmembrane receptors, are potential targets.

      The cellular proto-oncogene c-myc is involved in cell proliferation and transformation but is also implicated in the induction of programmed cell death  
25       (apoptosis). The c-Myc protein is a transcriptional activator with a carboxyl-terminal basic region/helix-loop-helix (HLH)/leucine zipper (LZ) domain. It forms heterodimers with the HLH/LZ protein Max and transactivates gene expression after binding DNA E-box elements. The protein Max is the obligatory partner of c-Myc for many of its biological functions analyzed to date. For instance, Myc must  
30       heterodimerize with Max to bind DNA and perform its oncogenic activity.

      According to the present invention, the subject method can be used to derive peptides and peptidomimetics which can inhibit formation of complexes between

Myc and other proteins such as Max, and/or which can inhibit the binding of a Myc complex to a myc-responsive element in a gene. The total synthesis of Myc-Max and Max-Max dimers are described by Canne *et al.* (1995) *J Am Chem Soc* 117:2998-3007.

5 The total synthesis of TGF $\beta$  has been described by Woo *et al.* (1989) *Protein Eng* 3:29-37, and provides another possible target molecule.

Yet another target which can be derived for use in the subject method is fibronectin, a glycoprotein involved in cell adhesion, tissue organization and wound healing. The total synthesis of fibronectin modules is described by, for example,  
10 Williams *et al.* (1994) *J Am Chem Soc* 116:10797-10798.

It has been previously shown that the expression of human immunodeficiency virus type 1 (HIV-1) major gag protein, p24, is persistent on the surface of HIV-1-infected cells (Nishino *et al.* (1992) *Vaccine* 10:677-683). The total synthesis of a C-terminal 100 amino acid fragment of p24 is described by  
15 Mascagni *et al.* (1990) *Tetrahedron Lett* 31:4637-4640, and that portion of the p24 protein can be used to generate a screening target.

Likewise, the HIV protease has been synthesized by total chemical synthetic means (Kent *et al.* PCT Publication WO93/20098) and provides a unique target for developing inhibitors of the catalytic activity as well as inhibitors of protein-protein  
20 interactions involving the protease.

While many of the targets illustrated above may be prepared by chemical synthesis, this is by no means a requirement or even a preference. An advantage of our invention is that any other method may be used to prepare a partially purified or pure target, such as purification from its biological source, biosynthesis, or  
25 recombinant DNA-based expression.

#### MATERIALS AND METHODS

*Construction of plasmids for the over-expression of his-tagged and untagged E. coli IF1, IF2 and IF3 proteins.* *E. coli* initiation factor coding sequences, each containing an insertion of six histidines immediately after the N-terminal methionine, were synthesized by PCR from published plasmids and sub-cloned into  
30 a vector derived from pET24a (Novagen). Plasmid pXR201 containing the native IF1 sequence encoded by an artificial sequence of *E. coli*-preferred codons (instead

of *infA* codons) was kindly supplied by R. Spurio and C. Gualerzi (Calogero *et al.* (1987) *Mol Gen Genet* 208, 63-9) and sub-cloned to give pAF1H. Plasmid pSL4 containing the native IF2 sequence encoded by *infB* was kindly supplied by S. Laalami and M. Grunberg-Manago (Laalami *et al.* (1991) *J Mol Biol* 220, 335-49) and sub-cloned to give pAF2H. Plasmid pDD1 containing the native IF3 sequence encoded by *infC* was kindly supplied by N. Brot and I. Schwartz (De Bellis and Schwartz (1990) *Nucleic Acids Res* 18, 1311) and sub-cloned to give pAF3H. The sequences of the three subclones, characterized by a combination of restriction digests and sequence analyses, begin (ligation sites underlined) with  
10 TATACA/TATG(CAC)<sub>6</sub> before the second amino acid; the final amino acid is followed by the sequence TAAG/AATTCGAGCTCCGTCGA/42 bp deletion/AGATCC, and the remainder of the sequence is from pET24a. Analogous methods were also used to clone and over-express untagged versions of IF1, IF2 and IF3.

15 *Over-expression and purification of E. coli translation factor proteins.* Plasmid pHTA7 (in *E. coli* BL21(DE3)) encoding his-tagged *E. coli* EF-1 $\alpha$ , containing a his<sub>6</sub> sequence inserted between the first two codons of *TufA*, and plasmid pHIS, containing a his<sub>6</sub> sequence inserted between the first two codons of *Tsf* was kindly supplied by Y.-W. Hwang and D. Miller (see Hwang *et al.* (1997) *Arch Biochem Biophys* 348, 157-62). Plasmid pRSET/EF-G(His) (in *E. coli* BL21(DE3) cells together with the pLysS plasmid) encoding his-tagged *E. coli* EF-G (EF-GH), containing an N-terminal extension of about 30 amino acids including his<sub>6</sub>, was kindly supplied by A. Savelsbergh and W. Wintermeyer (Semenkov *et al.* (1996) *Proc Natl Acad Sci U S A* 93, 12183-8). Expression of our three his-tagged  
25 initiation factor subclones (in *E. coli* BL21(DE3)pLysS; Novagen) and the three supplied clones was induced with IPTG. All the factors were expressed predominantly in the soluble cellular fractions and purified by step elution from Ni-NTA agarose columns using standard protocols (Qiagen), except that 10  $\mu$ M GDP was included up to the last dialysis step for EF-TuH. All factors were dialysed  
30 against buffer A (10 mM Tris-HCl pH 7.4, 1 mM MgCl<sub>2</sub>, 1 mM DTT). Precipitated IF3H was recovered by redissolving in 5M urea (Hershey *et al.* (1977) *Arch Biochem Biophys* 182, 626-38), diluted and then dialysed against buffer A

containing 100 mM NH<sub>4</sub>Cl. In contrast to the extensive proteolytic degradation observed by others during the over-expression and purification of native IF2 (Mortensen *et al.* (1991) *Biochimie* 73, 983-9), recombinant IF2II was not proteolytically labile. Pure *E. coli* release factors (RFs) RF1, RF2, RF3, and RRF, were prepared as described (Yu *et al.* (1998) *J. Mol. Biol.* 284, 579-590). All factors were stored at -80°C. They were thawed many times without loss of activity, except for EF-TuH, which was stored at 4°C after thawing and used within a few weeks.

*Purification of E. coli ribosomes* (Kung *et al.* (1974) *Arch Biochem Biophys* 162, 578-84). SOLR™ cells (Stratagene) grown to mid-log phase were resuspended in buffer B (60 mM KOAc, 14 mM Mg(OAc)<sub>2</sub>, 10 mM Tris-HOAc, 1 mM DTT, pH 7.9), sonicated, and centrifuged at 10 000 g. The supernatant was centrifuged at 30 000 g, and the resulting supernatant was then centrifuged at 150 000 g. The ribosomal pellet was washed by stirring in buffer B containing 1M NH<sub>4</sub>Cl at 4°C overnight and then repelleting at 150 000 g. The washing was repeated twice more to give 3 x washed ribosomes which were resuspended in buffer C (10 mM Mg(OAc)<sub>2</sub>, 1 mM Tris-HCl pH 7.4, 1 mM DTT) and stored at -80°C.

*Synthesis of mRNAs.* mRNAs were transcribed (Milligan and Uhlenbeck (1989) *Methods in Enzymology* 180, 51-62) from synthetic oligodeoxyribonucleotides (Research Genetics) and purified by gel electrophoresis (Forster and Symons (1987) *Cell* 49, 211-20). Because the templates are relatively long, extended deprotection times were necessary (12 h) following the synthesis of the blocked oligonucleotides to enable optimal transcription.

*Preparation of aminoacyl tRNAs.* Pure *E. coli* tRNA isoacceptors were from Subriden RNA. Each isoacceptor was prepared by the manufacturer from *E. coli* total tRNA (Plenum) using three column chromatography steps. The first fractionation used BD cellulose, the second, DEAE Sephadex at pH 7, and the third, DEAE Sephadex at pH 5 or Sepharose. Natural aminoacyl tRNAs were prepared from these isoacceptors as follows. High specific activity <sup>3</sup>H-fmet-tRNA<sub>f</sub><sup>fmet</sup> (24 000 d.p.m./pmol in Table 1) and low specific activity <sup>3</sup>H- or <sup>35</sup>S-fmet-tRNA<sub>f</sub><sup>fmet</sup> (a few hundred d.p.m./pmol; used for all other studies) were prepared as described (Robakis *et al.* (1981) *Proc Natl Acad Sci U S A* 78, 4261-4) using MetRS, met-tRNA<sub>f</sub><sup>fmet</sup> formyltransferase and N<sup>5,10</sup>-methenyl THF (synthesized using N<sup>10</sup>-formyl

THF synthetase.  $^{14}\text{C}$ -thr-tRNA $^{\text{thr}}$  (510 d.p.m./ $\mu\text{mol}$ ) and  $^3\text{H}$ -val-tRNA $^{\text{val}}$  (21 000 d.p.m./ $\mu\text{mol}$  in Figs. 5-7; 28 000 d.p.m./ $\mu\text{mol}$  in Figs. 4 and 8) were aminoacylated as described (Robakis *et al.* (1981) *Proc Natl Acad Sci U S A* 78, 4261-4) with ThrRS, ValRS or, for the lower specific activity val-tRNA, a tRNA-free preparation of total *E. coli* synthetases partially purified from an 150 000 g supernatant by step elution with 0.3M KCl from DEAE Sepharose (see Kung *et al.* (1975) *J Biol Chem* 250, 1556-62).

Biotin-labeled-lys-tRNA $^{\text{lys}}$  (Transcend $^{\text{TM}}$  tRNA; Fig. 22) was purchased from Promega. This material was prepared by the manufacturer from *E. coli* total tRNA by charging with lysine using a crude preparation of total synthetases, enriching for lys-tRNA $^{\text{lys}}$  by ion exchange chromatography, and chemical coupling to biotin via an uncharged, 13-carbon-long spacer.

NVOC-amino protected aminoacyl tRNAs were prepared, stored and used by methods standard in the art (detailed in Thorson *et al.* (1988) *Methods in Molecular Biology* 77, 43-73; Steward and Chamberlin (1998) *Methods in Molecular Biology* 77, 325-354). Briefly, pdCpA was chemically synthesised from adenosine and a deoxycytidine derivative. The various unnatural amino acids were purchased in unprotected form from standard suppliers (*e.g.* Sigma, Aldrich, Fluka, Bachem and Novabiochem). Our most commonly used NVOC-amino acids, such as aC, mS and Ala, lacked reactive amino acid side chains requiring chemical protection, so they were synthesized directly from the unprotected amino acids by amino protection with NVOC-Cl. All suitably protected amino acids, including those with additional side-chain protections synthesized by the standard methods, were activated by cyanomethyl ester synthesis and then coupled to pdCpA. The resulting NVOC-protected aminoacylated pdCpA compounds (and also their conjugates with RNAs) were stable when stored in an aqueous solution with a pH of approximately 5 in the dark at  $-80^{\circ}\text{C}$ . NVOC-protected aminoacylated pdCpA compounds were ligated by T4 RNA ligase to various tRNA derivatives lacking the terminal CA dinucleotide, and conjugates were purified and stored (see above). The resulting NVOC-protected aminoacyl tRNAs were deprotected by UV-irradiation immediately prior to use in translations.

Assay of his-tagged *E. coli* translation factor proteins. Initiation factor assays (Kung *et al.* (1974) *Arch Biochem Biophys* 162, 578-84) used custom-synthesized  $\Delta$ pUpG RNA template (TriLink Biotechnologies). A mixture containing 0.95  $\mu$ M IF1H, 0.15  $\mu$ M IF2H, 0.78  $\mu$ M IF3II, 3 x washed ribosomes at 5 0.029  $A_{260}/\mu$ l (33 nM estimated to be active in translation; see below), 0.29  $\mu$ M  $^3$ H-fmet-tRNA $_{f}^{\text{fmet}}$ , 150  $\mu$ M AUG and 0.4 mM GTP in 50 mM Tris-HCl pH 7.4, 100 mM NH $_4$ Cl, 5 mM Mg(OAc) $_2$  and 2 mM DTT was incubated at 37°C for 10 min. After dilution, the mixture was rapidly filtered through nitrocellulose to separate initiation complex-bound fmet-tRNA $_{f}^{\text{fmet}}$  from unbound species. Reactions that

10 lacked template were used as controls for non-specific binding (29% of maximal d.p.m. for Table 1; 9% of maximal d.p.m. using higher (saturating) concentrations of ribosomes that bound 70% of  $^3$ H-fmet-tRNA $_{f}^{\text{fmet}}$ ). Because of the high affinity of EF-Tu for EF-Ts, EF-Ts activity assays were performed (Weissbach and Pestka (1977) *Molecular Mechanisms of Protein Biosynthesis, Academic Press, New York, NY*) with EF-TuH, indicating a copurification level of about 2%. EF-TuI activity was measured by GDP binding (Weissbach and Pestka (1977) *Molecular Mechanisms of Protein Biosynthesis, Academic Press, New York, NY*), and EF-GH activity was measured by a ribosome-dependent GTPase assay (Weissbach and Pestka (1977) *Molecular Mechanisms of Protein Biosynthesis, Academic Press, New York, NY*).

20 *Translations.* The components of translation mixes were adapted from published work (Robakis *et al.* (1981) *Proc Natl Acad Sci U S A* 78, 4261-4; Cenatiempo *et al.* (1982) *Arch Biochem Biophys* 218, 572-8). 5 x Premix buffer was prepared from a solution containing 180 mM Tris-HOAc (pH 7.5), 50 mM sodium 3,3-dimethyl-glutarate (pH 6.0), 180 mM NH $_4$ OAc, 10 mM DTT, 140 mM potassium phosphoenolpyruvate (pH 6.6), 195 mM KOAc and 4 mM spermidine $\cdot$ 3HCl by adjusting it to pH 6.8 with NaOH. As a representative example, the mixture in the translation of the MTTV mRNA template (Fig. 6, 30  $\mu$ l total volume) included 1 x premix buffer, 9.5 mM Mg(OAc) $_2$ , 1 mM GTP, 14 ng/ $\mu$ l

30 pyruvate kinase, 4.3% PEG 8000, 0.95  $\mu$ M IF1H, 0.15  $\mu$ M IF2H, 0.78  $\mu$ M IF3H, 3.1  $\mu$ M EF-TuH, 0.88  $\mu$ M EF-GH, 3 x washed ribosomes at 0.029  $A_{260}/\mu$ l (33 nM estimated to be active in translation; see Fig. 4), 0.29  $\mu$ M  $^3$ H-fmet-tRNA $_{f}^{\text{fmet}}$ , 0.58

$\mu\text{M}$   $^{14}\text{C}$ -thr-tRNA $_3^{\text{thr}}$ ,  $0.29 \mu\text{M}$   $^3\text{H}$ -val-tRNA $_1^{\text{val}}$ , and  $1 \mu\text{M}$  mRNA. This mixture was incubated at  $37^\circ\text{C}$  for 50 min., although some translations were as short as 1 min. (Fig. 4). The set up for some translations also included preincubations at  $37^\circ\text{C}$  for 10 min. of an initiation mix (GTP, IF1H, IF2H, IF3H, ribosomes,  $^3\text{H}$ -fmet-tRNA $_1^{\text{fmet}}$  and mRNA) and an elongation factor mix (GTP, EF-TuH, EF-GH,  $^{14}\text{C}$ -thr-tRNA $_3^{\text{thr}}$  and  $^3\text{H}$ -val-tRNA $_1^{\text{val}}$  (Fig. 4)).

For HPLC analysis (Fig. 6), peptides and amino acids were released from tRNAs by addition of  $1\text{M}$  NaOH ( $6 \mu\text{l}$ ) and incubation at  $37^\circ\text{C}$  for 20 min. Unlabeled marker peptides (Research Genetics) were then added ( $18 \mu\text{l}$  with a combined peptide concentration of  $10 \mu\text{g}/\mu\text{l}$  in  $\text{H}_2\text{O}$ ), the solution was acidified with glacial acetic acid ( $5 \mu\text{l}$ ), microcentrifuged, and the supernatant was then microcentrifuged through a Microcon 10 ultrafiltration device (10 kD cut-off; Amicon). A portion of the filtrate ( $20 \mu\text{l}$ ) was loaded onto a C18 reversed phase column (Vydac) and eluted with a 0 to 31.5%  $\text{H}_2\text{O}/\text{MeCN}$  gradient containing 0.1% TFA at  $1 \text{ ml}/\text{min}$ . with detection by absorbance at 229 nm and scintillation counting of 42-drop fractions using a dual-labeled d.p.m. program (Packard).

For analysis of the fMTV and fMVT syntheses, formylated peptides and formyl methionine were separated from unformylated amino acids by base hydrolysis, acidification, and cation-exchange chromatography on mini-columns (Peacock *et al.* (1984) *Proc Natl Acad Sci US A* 81, 6009-15).

*Control translations.* Additional translations were performed to assess the dependence on components other than the translation factors. Omission of  $^3\text{H}$ -fmet-tRNA $_1^{\text{fmet}}$  or  $^{14}\text{C}$ -thr-tRNA $_3^{\text{thr}}$  abolishes fMTV synthesis from mRNA MTV, based on incorporation of  $^3\text{H}$ -val into product. In the experiment omitting  $^{14}\text{C}$ -thr-tRNA $_3^{\text{thr}}$ , addition of uncharged tRNA $_3^{\text{thr}}$  does not reconstitute measurable fMTV synthesis, demonstrating a lack of ThrRS activity under translation conditions, as expected for a purified system (as a further control, addition of uncharged thr-tRNA $_3^{\text{thr}}$  was not inhibitory to fMTV synthesis). Omission of  $^3\text{H}$ -val-tRNA $_1^{\text{val}}$  abolishes fMVT synthesis from mRNA MVT (Fig. 4), based on incorporation of  $^{14}\text{C}$ -thr-tRNA $_3^{\text{thr}}$  into product. The T/V ratios in the mRNA MVT and mRNA mMV (Fig. 3) products are about 1.0 and 0, respectively, as expected. Omitting ribosomes gives the lowest background radioactivity measurements, and translations lacking

WO 02/059293

50

PCT/US02/02344

mRNA do not accumulate formylated peptide products with time. Omitting rabbit muscle pyruvate kinase (the protein (Jelenc and Kurland (1979) *Proc Natl Acad Sci U S A* 76, 3174-8) from Sigma migrates as a single major band on SDS-PAGE) decreases the yield of product by 50%. The standard concentration of  $Mg^{2+}$  (9.5 mM) is optimal, but translation can occur efficiently at higher and lower concentrations.

WO 02/059293

51

PCT/US02/02344

Table 1

Factor dependencies for initiation complex formation.

5

| Initiation factor omitted | Initiation complex <sup>a</sup> |   |   |
|---------------------------|---------------------------------|---|---|
|                           | His-tagged <sup>b</sup>         | Native (Kang <i>et al.</i> ) <sup>c</sup> | Native (Dubroff and Mitra) <sup>d</sup> |
| None                      | 100                             | 100                                       | 100                                     |
| IF1                       | 48                              | 61  | 36                                      |
| IF2                       | 0                               | 4   | 7                                       |
| IF3                       | 7                               | 17  | 48                                      |

15

<sup>a</sup> % Maximal binding of <sup>3</sup>H-labeled fmet-tRNA<sub>fmet</sub> to ribosomes in the presence of GTP.

<sup>b</sup> Performed with purified his-tagged initiation factors, ApUpG template and 5 mM Mg<sup>2+</sup> (see Materials and Methods). The maximum concentration of fmet-tRNA<sub>fmet</sub> specifically bound into initiation complexes was 19 nM.

20

<sup>c</sup> Performed with purified native initiation factors, poly(U:G) (3:1) template and 10 mM Mg<sup>2+</sup>.

<sup>d</sup> Performed with purified native initiation factors, ApUpG template and 5 mM Mg<sup>2+</sup>.

25

WO 02/059293

52

PCT/US02/02344

Table 2

5 Measurement of processivity of synthesis of a 7-mer peptide. Peptide product  
 d.p.m. from 5.1  $\mu$ l coupled transcription/translations containing translation factors  
 IF2, IF3, EF-Tu and EF-G are converted to pmoles.

| 10 | Peptide encoded by<br>Thr/Val | $^{14}\text{C}$ -Thr<br>(pmol) | $^3\text{H}$ -Val<br>(pmol) | $^{14}\text{C}$ -Thr/ $^3\text{H}$ -Val<br>measured |   |
|----|-------------------------------|--------------------------------|-----------------------------|---|---|
| 15 | mRNA template<br>expected     |                                |                             |   |   |
|    | Met-Thr-Thr-Thr-Thr-Thr-Val   | 0.45                           | 0.07                        | 6.4   | 5 |
|    | Met-Thr-Val                   | 0.06                           | 0.06                        | 1.0   | 1 |
|    | Met-Val                       | 0.008                          | 0.07                        | 0.11  | 0 |

WO 02/059293

53

PCT/US02/02344

Table 3

- 5 Stimulation of pure translation system by pure *E. coli* release factors (RFs) RF1, RF2, RF3, and RRF (Yu *et al.* (1998) *J. Mol. Biol.* 284, 579-590; see Materials and Methods). mRNAs either lacked a stop codon (mMTTV) or encoded a stop codon (mMTCV<sub>UAG</sub>) directly following the fourth amino acid codon. Product dpm: radioactive disintegrations per minute of tritiated valine incorporated into peptide products. All translations contained translation factors IF1, IF2, IF3, EF-Tu, EF-G and EF-Ts.
- 10

| mRNA                    | Supplied aminoacyl tRNAs | RFs | Product dpm | Peptide product |
|-------------------------|--------------------------|-----|-------------|-----------------|
| 15 mMTTV                | fM <sub>T</sub> , V      | -   | 4597        | fMTTV           |
| mMTTV                   | fM <sub>T</sub> , V      | +   | 7187        | fMTTV           |
| mMTCV <sub>UAG</sub>    | fM <sub>T</sub> , C, V   | -   | 2676        | fMTCV           |
| 20 mMTCV <sub>UAG</sub> | fM <sub>T</sub> , C, V   | +   | 3866        | fMTCV           |

WO 02/059293

54

PCT/US02/02344

Table 4

- 5 Incorporation of five successive unnatural amino acids into a peptidomimetic product. Allylglycine is abbreviated as aG; its structure is shown in Fig. 14. The artificial tRNA<sup>Asp</sup> charged with aG is termed aG-tRNA(T) because it recognises a T codon (Fig. 15). Total d.p.m.: total <sup>3</sup>H d.p.m. eluted from mini-column (see Materials and Methods). All translations contained translation factors IF1, IF2, IF3, EF-Tu, EF-G and EF-Ts.

|    | mRNA product     | Supplied aminoacyl-tRNAs | Total dpm | Peptidomimetic   |
|----|------------------|--------------------------|-----------|------------------|
| 15 | mMTTTTTV<br>aG-V | fM,aG-tRNA(T),V          | 2721      | fM-aG-aG-aG-aG-  |
|    | mMTTTTTV         | fM, T, V                 | 7685      | fMTTTTTV         |
| 20 | mMTTTTTV<br>dpm) | fM, V                    | 1116      | none (background |
|    | -<br>dpm)        | fM, V                    | 1113      | none (background |

WO 02/059293

55

PCT/US02/02344

Table 5

5 Incorporation of two different types of unnatural amino acids into a peptidomimetic product. O-methyl serine is abbreviated as mS. Four artificial aminoacyl tRNAs were used: mS-tRNA(T), aG-tRNA(N), aG-tRNA(S), and aG-tRNA(V), with the respective mRNA codons recognised by each artificial tRNA given in parentheses. The highly labelled <sup>3</sup>H-amino acid was E. Total d.p.m.: total <sup>3</sup>H d.p.m. eluted from mini-column. All translations contained translation factors IF1, IF2, IF3, EF-Tu, and EF-G.

| mRNA product       | Supplied aminoacyl-tRNAs                                 | Total dpm | Peptidomimetic    |
|--------------------|--|-----------|-------------------|
| 15 mMTNSVE<br>aG-E | fM, mS-tRNA(T), aG-tRNA(N),<br>aG-tRNA(S), aG-tRNA(V), E | 2353      | fM-mS-aG-aG-      |
| 20 mMTNSVE         | fM, aG-tRNA(N),<br>aG-tRNA(S), aG-tRNA(V), E             | 899       | none (background) |
| 25 -               | fM, mS-tRNA(T), aG-tRNA(N),<br>aG-tRNA(S), aG-tRNA(V), E | 1033      | none (background) |
| mMTTV              | fM, T, V   | 7100      | fMTTV             |

**Claims:**

1. A reconstituted cell-free translation system comprising translation factors and tRNA species capable of translating exogenously added mRNA(s) with highly selective incorporation at each codon to form a peptide product, or a peptidomimetic product when the system includes one or more of the tRNA species charged with a non-naturally occurring amino acid or amino acid analog.  
5
- 10 2. The translation system of claim 1 for generating a peptide product, which preparation is substantially free of the translation factors EF-P, W, W2 or rescue.
- 15 3. The translation system of claim 1 for generating a peptidomimetic product, including one or more of the tRNA species charged with a non-naturally occurring amino acid or amino acid analog that is incorporated into the backbone of the peptidomimetic product.
- 20 4. The translation system of claim 2 for generating a peptidomimetic product, which preparation is substantially free of the translation factors EF-P, W, W2 or rescue.
- 25 5. The translation system of claim 2, wherein the amino acid analog is selected from the group consisting of  $\beta$ -cyanoalanine, canavanine, djenkolic acid, norleucine, 3-phosphoserine, homoserine, dihydroxyphenylalanine, 5-hydroxytryptophan, 1-methylhistidine, 3-methylhistidine, allyl glycine (or its alkyne derivative), O-methyl-serine, biotinyl-lysine, biotinyl-cysteine (or other biotin-labelled amino acids) cyclohexylalanine, homoglutamate, D-alanine (or other D-amino acids), N-methyl glycine (or other N-methyl amino acids), epsilon-N-methyl-lysine, and radioisotope derivatives of the 21 natural amino acids or unnatural amino acids.  
30
- 35 6. The translation system of claim 1, further including one or more exogenously added mRNA species.

WO 02/059293

57

PCT/US02/02344

7. A cell-free translation system comprising translation factors and tRNA species capable of translating exogenously added mRNA(s) to form a peptidomimetic product, which preparation
- 5 (a) lacks one or more active wild-type amino acyl tRNAs and ability to synthesize said wild-type amino acyl tRNA,
- (b) includes at least one exogenous amino acyl tRNA charged with a non-natural amino acid species or amino acid analog, replacing said inactive tRNA species.
- 10 8. The translation system of claim 7, including a plurality of different mRNA species.
9. A kit comprising: (a) components of the translation system of claim 1, which can be admixed to form the reconstituted cell-free translation system capable of translating exogenously added RNA to form a peptide or peptidomimetic; (b) instructions associated there with.
- 15 10. A kit comprising: (a) components of the translation system of claim 7, which can be admixed to form the cell-free translation system capable of translating exogenously added RNA to form a peptidomimetic; (b) instructions associated there with.
- 20 11. A method for generating a peptide or peptidomimetic comprising:
- (a) providing a cell-free translation system of claim 1;
- 25 (b) contacting the translation system with one or more exogenous mRNA species;
- (c) isolating and/or identifying peptide or peptidomimetic products of the translation system.
- 30 12. The method of claim 11, wherein one or more of the tRNA species of the translation system is charged with a non-naturally occurring amino acid or amino acid analog that is incorporated into the backbone of a peptidomimetic product.

13. The method of claim 11, wherein the method is carried out on a library of at least 100 different RNA species, and peptide or peptidomimetic products are identified or isolated from the translation system based on catalytic or binding activity.
- 5
14. The method of claim 11, wherein the mRNA is generated by in vitro transcription in the translation system.
15. The method of claim 11, wherein the peptide or peptidomimetic products are formed as a covalent adduct of the exogenous mRNA by which said products are encoded.
- 10
16. The method of claim 11, wherein the translation system is contacted with a library of different exogenous mRNA species to generate variegated population of peptides or peptidomimetics products of at least  $10^3$  different sequences.
- 15
17. The method of claim 16, wherein at least  $10^8$  different sequences are produced.
- 20
18. A peptide or peptidomimetic identified or synthesized by the method of claim 11.
19. A method for generating a peptidomimetic comprising:
- 25
- (a) providing a cell-free translation system of claim 7;
  - (b) contacting the translation system with one or more exogenous mRNA species;
  - (c) isolating and/or identifying peptidomimetic products of the translation system.
- 30
20. A method for conducting a drug discovery business, comprising:
- (a) by the method of claim 11, identifying peptide or peptidomimetic products of the translation system having a desired biological activity;

WO 02/059293

59

PCT/US02/02344

- (b) validating the biological activity of the peptide or peptidomimetic products identified in step (a) by further in vitro or in vivo assay;
- (c) conducting therapeutic profiling of products validated in step (b), or further analogs thereof, for efficacy and toxicity in animals; and
- 5 (d) formulating a pharmaceutical preparation including one or more products identified in step (c) as having an acceptable therapeutic profile.
21. The method of claim 20, including an additional step of establishing a distribution system for distributing the pharmaceutical preparation for sale, and may optionally include establishing a sales group for marketing the pharmaceutical preparation.
- 10
22. A method for conducting a drug discovery business, comprising:
- 15 (a) by the method of claim 11, identifying peptide or peptidomimetic products of the translation system having a desired biological activity;
- (b) validating the biological activity of the peptide or peptidomimetic products identified in step (a) by further in vitro or in vivo assay;
- 20 (c) (optionally) conducting therapeutic profiling of products validated in step (b), or further analogs thereof, for efficacy and toxicity in animals; and
- (d) licensing, to a third party, the rights for further drug development of validated products.
- 25

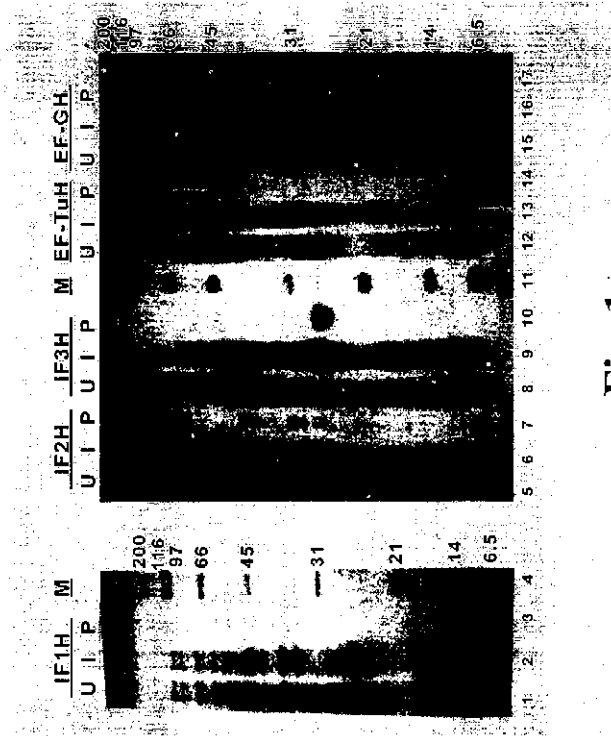


Fig. 1

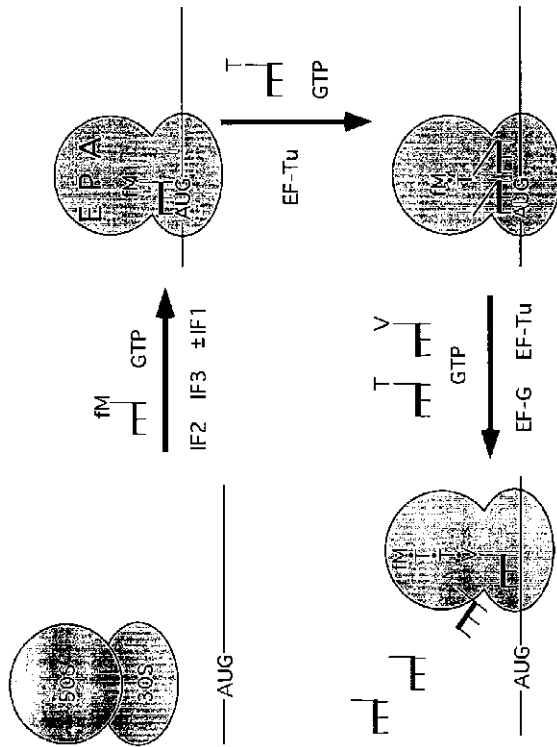


Fig. 2



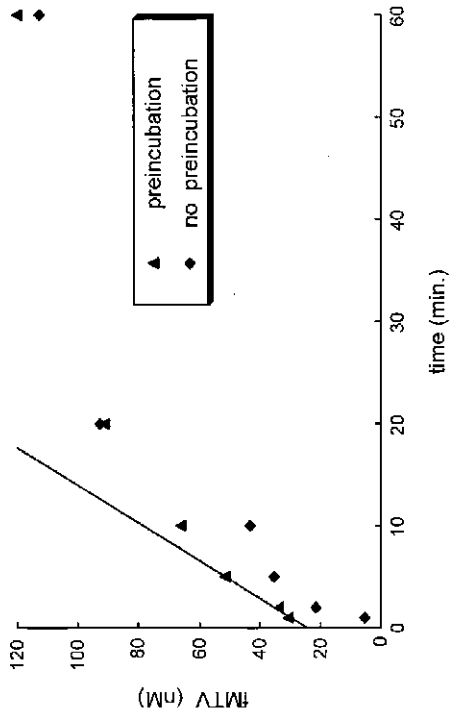


Fig. 4

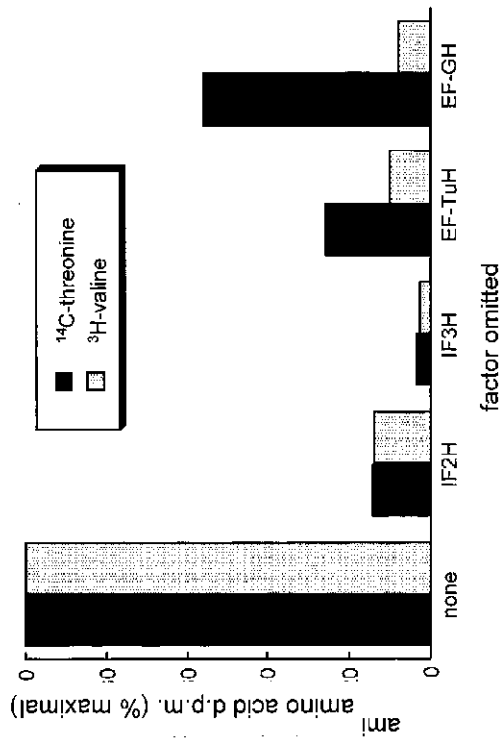
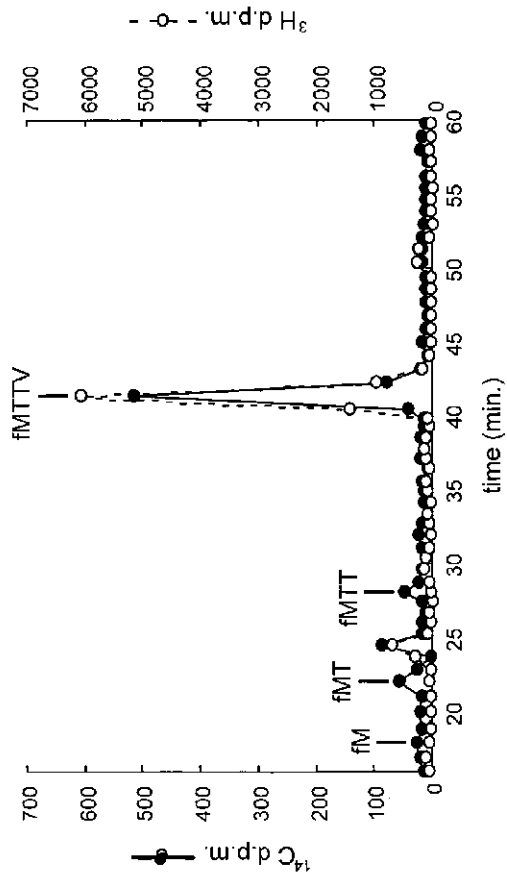


Fig. 5



**Fig. 6**

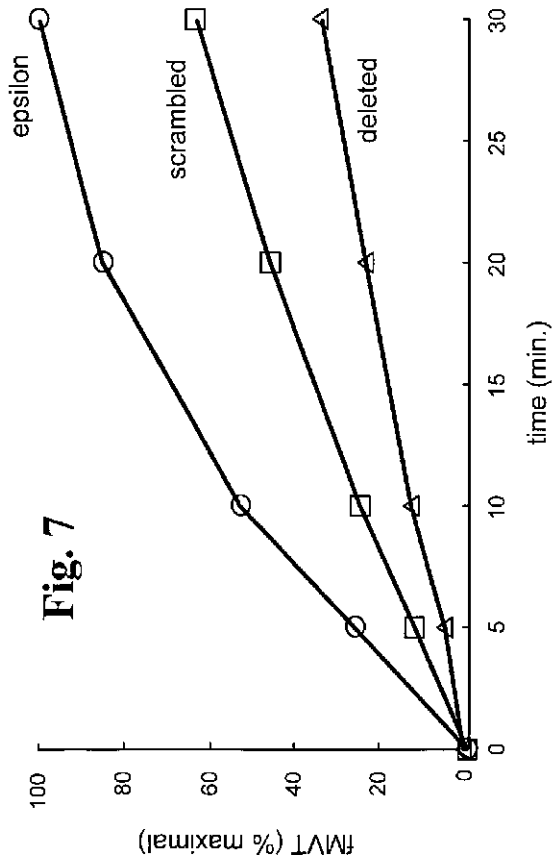
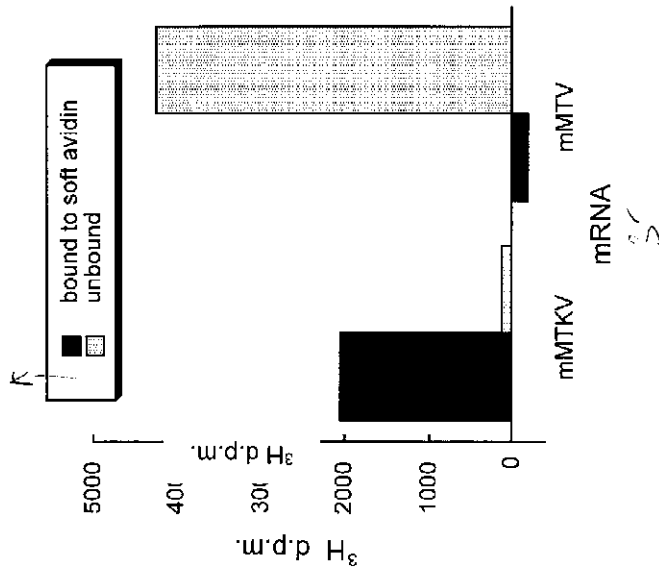


Fig. 8





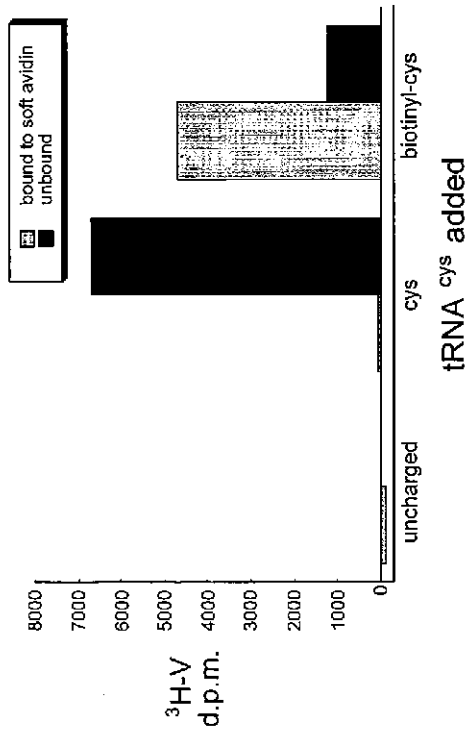


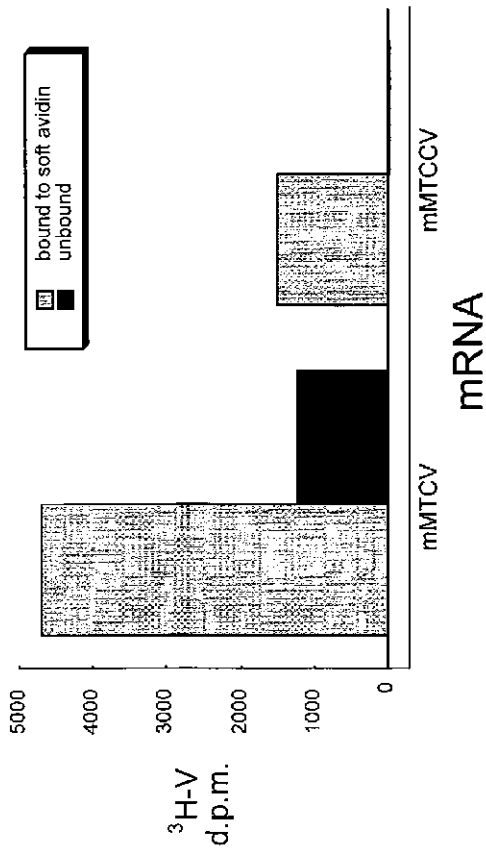
WO 02/059293

11/22

PCT/US02/02344

**Fig. 11**





**Fig. 12**

WO 02/059293

13/22

PCT/US02/02344

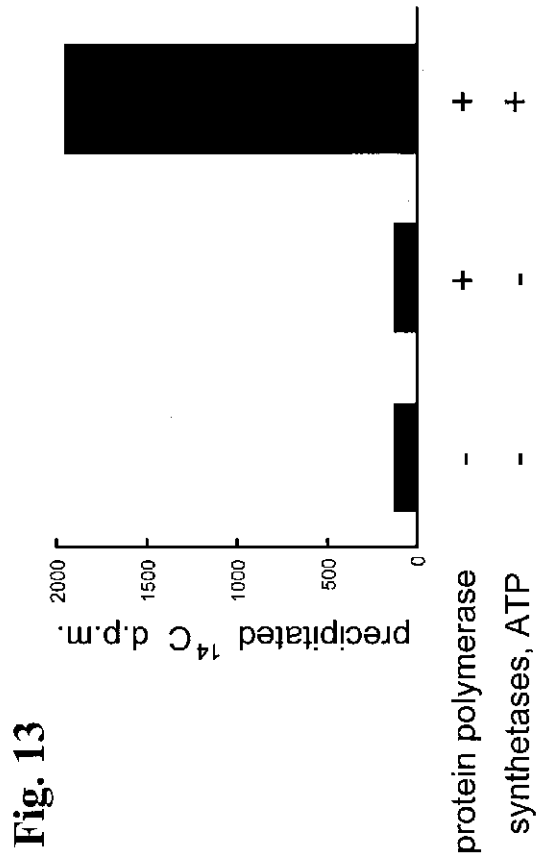
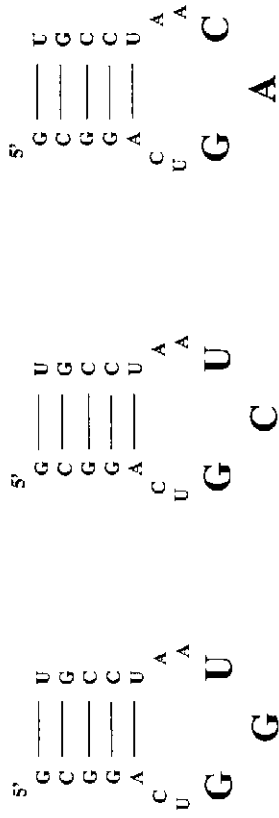
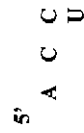


Fig. 13

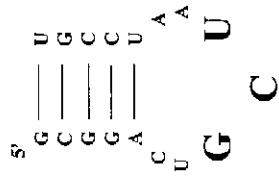




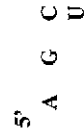
for



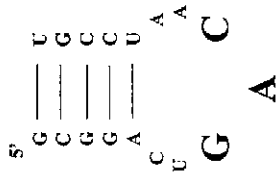
Unnatural amino acid 1  
 (was T)



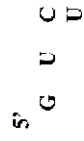
for



Unnatural amino acid 2  
 (was S)



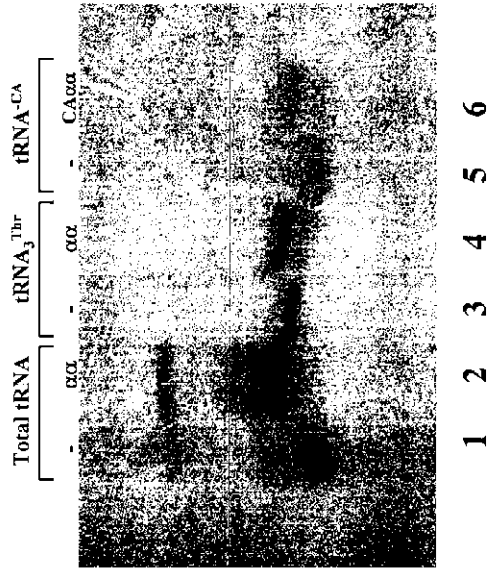
for



Unnatural amino acid 3  
 (was V)

Fig. 15

**Fig. 16**



WO 02/059293

17/22

PCT/US02/02344

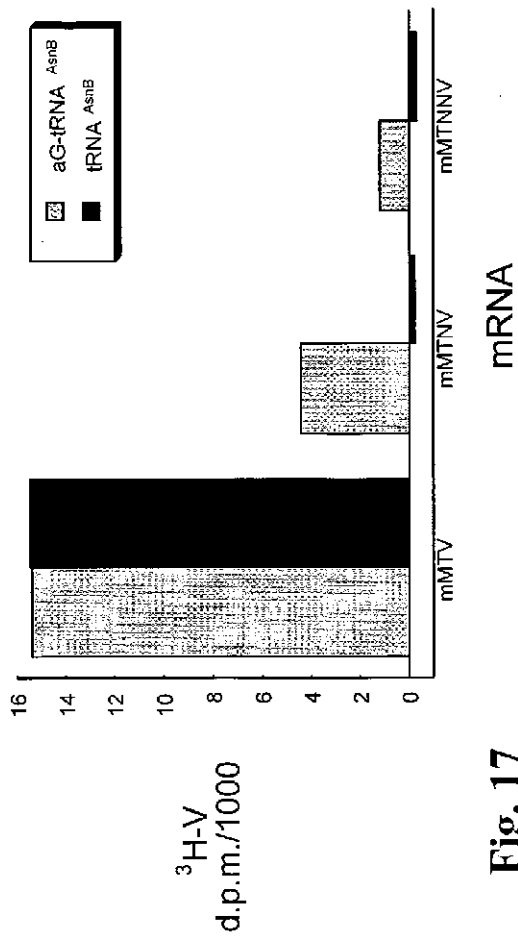
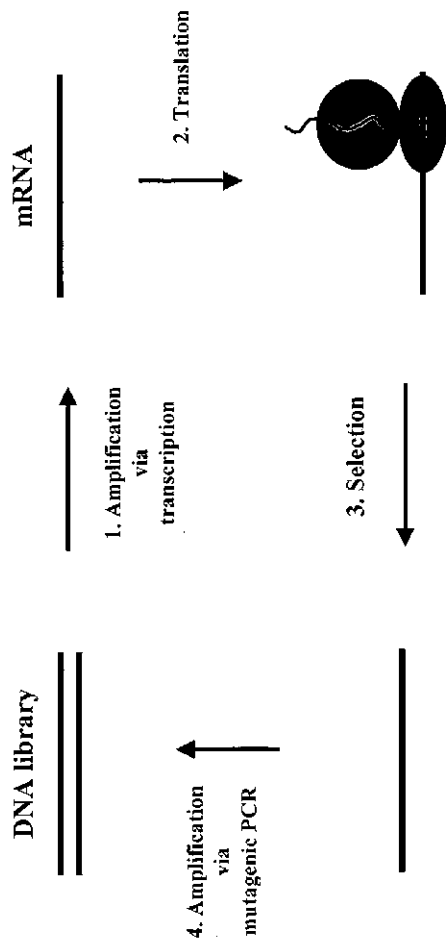


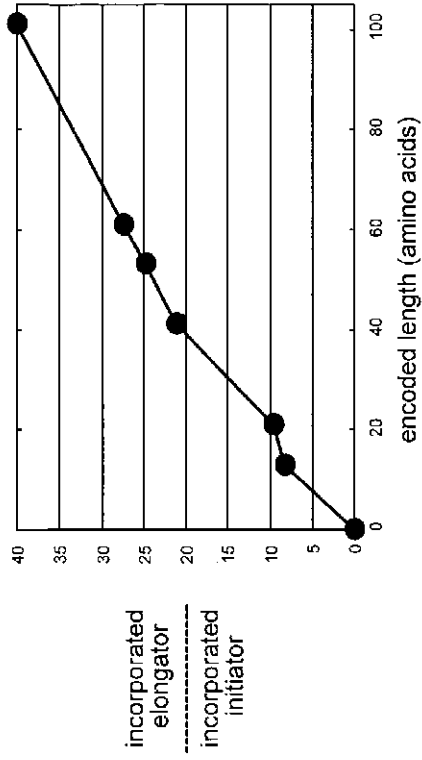
Fig. 17

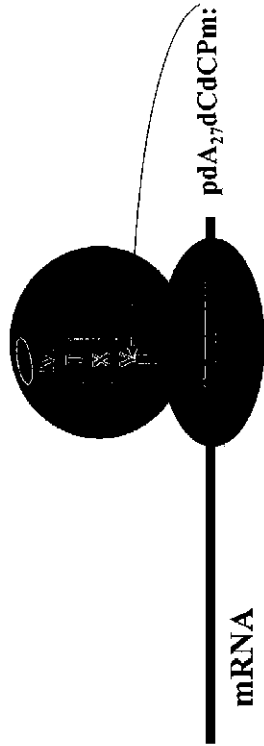


**Fig. 18**



**Fig. 20**





**Fig. 21**

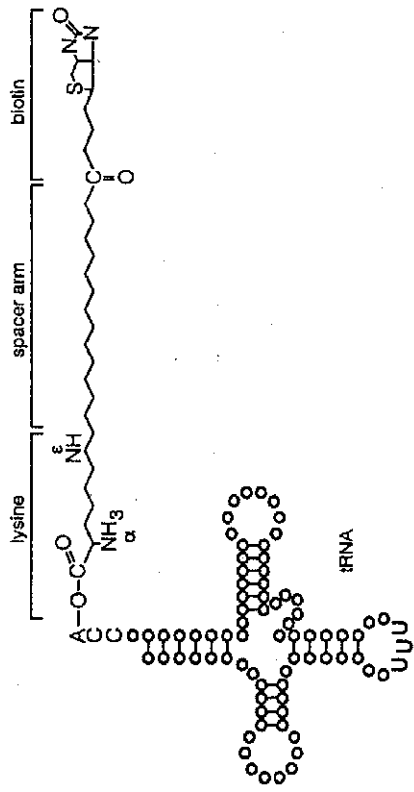


Fig. 22

【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  
1 August 2002 (01.08.2002)

PCT

(10) International Publication Number  
WO 02/059293 A3

(51) International Patent Classification: C12N 15/00, C07K 2400 (74) Agents: VINCENT, Matthew, P. et al.; Ropes & Gray, Patent Group, One International Place, Boston, MA 02110 (US).

(21) International Application Number: PCT/US02/02344

(22) International Filing Date: 25 January 2002 (25.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 60/261,137 25 January 2001 (25.01.2001) US

(71) Applicants and (72) Inventors: FORSTER, Anthony, C. (AU/US); 177 Market Road, Lexington, MA 02421 (US); BLACKLOW, Stephen, C. (US/US); 116 Tower Road, Lincoln, MA 01773 (US).

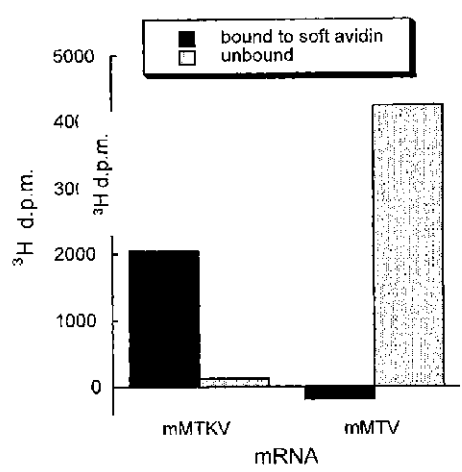
(81) Designated States (national): AL, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GR, GU, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NZ, PL, PT, RO, RU, SD, SE, 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, 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, PL, SE, TR), OAPI patent (BF, BI, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: PROCESS AND COMPOSITIONS FOR PEPTIDE, PROTEIN AND PEPTIDOMIMETIC SYNTHESIS

WO 02/059293 A3



(57) Abstract: The present invention is a simplified, highly purified, processive translation system that does not require the addition of translation factors E1, E2, E3, E4 or release. A new translation process process offers new, potentially improved, routes to all peptides, proteins and peptidomimetics currently synthesized by alternative routes.

WO 02/059293 A3



Published:  
— with international search report

*For issue-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.*

(88) Date of publication of the international search report:  
30 May 2003

## 【 国際調査報告 】

| INTERNATIONAL SEARCH REPORT  |  | International Application No.<br>PCT/US 02/02344                 |
|--|--|--|
| A. CLASSIFICATION OF SUBJECT MATTER<br>IPC 7 C12N15/00 C07K2/00  |  |  |
| According to International Patent Classification (IPC) or to both national classification and IPC  |  |  |
| B. FIELD(S) SEARCHED<br>Minimum documentation searched (classification system followed by classification symbols)<br>IPC 7 C12N C07K   |  |  |
| Documentation searched other than minimum documentation to the extent that such documents are included in the lists searched   |  |  |
| Electronic data base consulted during the international search (name of data base and, where practical, search terms used)<br>EPO-internal, WPI Data, PAJ, BIOSIS, EMBASE  |  |  |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT   |  |  |
| Category*  | Character of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.  |
| X  | PAVLOV ET AL: "Fast recycling of Escherichia coli ribosomes requires both ribosome recycling factor (RRF) and release factor RF3". THE EMBO JOURNAL, OXFORD UNIVERSITY PRESS, VOL. 16, NR. 13, PAGE(S) 4134-4141 XP001145784 abstract<br>page 4135, column 1, paragraph 3<br>page 4140, column 1, paragraph 4 -column 2, paragraph 4<br>---<br>-/- | 1-22   |
| <input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input type="checkbox"/> Patent family members are listed in annex.   |  |  |
| * Special categories of cited documents:<br>*A* document defining the general state of the art which is not considered to be of particular relevance<br>*E* earlier document but published on or after the international filing date<br>*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)<br>*O* document relating to an oral disclosure, use, exhibition or other means<br>*P* document published prior to the international filing date but later than the priority date claimed<br>*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<br>*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<br>*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<br>*Z* document member of the same patent family |  |  |
| Date of the actual completion of the international search<br>21 February 2003  |  | Date of mailing of the international search report<br>04/03/2003 |
| Name and mailing address of the ISA<br>European Patent Office P.O. Box 5516 Patentplan 2<br>NL - 2280 HV Rijswijk<br>Tel: (+31-70) 340-2040, Tx: 31 651 epo nl<br>Fax: (+31-70) 340-5016   |  | Authorized officer<br>Chavanne, F                                |

Form PCT/ISA/210 (second sheet) (July 1994)

## INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/US 02/02344

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT |   |          |
|--|---|----------|
| Category *   | Relevant to claim No.   |          |
| X  | <p>MURPHY E C JR. ET AL: "Cell-Free Synthesis of Rauscher Murine Leukemia Virus Gag and Gag-Pol Precursor Poly Proteins from Virion 35 S RNA in a mRNA-Dependent Translation System Derived from Mouse Tissue Culture Cells", VIROLOGY, VOL. 86, NR. 2, PAGE(S) 329-343 XP009005933<br/>ISSN: 0042-6822<br/>abstract<br/>page 330, column 1, paragraph 2 -column 2, paragraph 1<br/>page 331, column 1, paragraph 3 -column 2, paragraph 1<br/>page 336, column 1, paragraph 2 -column 2, paragraph 1<br/>page 340, column 1, paragraph 2 -column 2, paragraph 1<br/>figures 5,10<br/>---</p> | 1,3,6-22 |
| X  | <p>CLOAD S T ET AL: "Development of improved tRNAs for in vitro biosynthesis of proteins containing unnatural amino acids", CHEMISTRY AND BIOLOGY (LONDON), VOL. 3, NR. 12, PAGE(S) 1033-1038 XP009005959<br/>ISSN: 1074-5521<br/>page 1033, column 1, paragraph 1<br/>page 1034, column 1, paragraph 2 -page 1036, column 1, paragraph 2<br/>page 1037, column 2, paragraph 3<br/>---</p>  | 1,3,6-22 |
| X  | <p>KARGINOV A V ET AL: "Facile characterization of translation initiation via nonsense codon suppression", NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, VOL. 27, NR. 16, PAGE(S) 3283-3290 XP001145472<br/>ISSN: 0305-1048<br/>abstract<br/>page 3283, column 2, paragraph 1 - paragraph 2<br/>page 3284, column 2, paragraph 2 - paragraph 4<br/>page 3288, column 2, paragraph 1 - paragraph 2<br/>-----</p>  | 1,3,6-22 |

## フロントページの続き

|                          |               |            |
|--------------------------|---------------|------------|
| (51)Int.Cl. <sup>7</sup> | F I           | テーマコード(参考) |
| // C 0 7 K 5/10          | C 0 7 K 5/10  |            |
| C 0 7 K 14/47            | C 0 7 K 14/47 |            |

(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,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,RO,R U,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

(72)発明者 アントニー シー・フォスター  
 アメリカ合衆国 0 2 4 2 1 マサチューセッツ、レクシントン、マレット ロード 1 7 3

(72)発明者 スティーブン シー・ブラックロウ  
 アメリカ合衆国 0 1 7 7 3 マサチューセッツ、リンカン、タワー ロード 1 1 6

Fターム(参考) 4B024 AA01 AA20 BA19 BA21 BA36 BA63 CA05 CA11 CA12 DA06  
 FA01 HA03  
 4B064 AG02 AG13 AG20 AG23 AG32 CC01 CC04 CC24 CD13 CD20  
 CD30 CE10  
 4C084 AA17 NA14 ZB262 ZC022  
 4H045 AA20 BA01 BA13 CA05 CA41 DA01 DA02 DA50 DA51 DA56  
 EA22 EA29 FA70 FA74 GA21 GA23 GA25 GA26